nature neuroscience

Cardinal rules: visual orientation perception reflects knowledge of environmental statistics

Ahna R Girshick[1,2], Michael S Landy[1,2] & Eero P Simoncelli[1–4]

Humans are good at performing visual tasks, but experimental measurements have revealed substantial biases in the perception of basic visual attributes. An appealing hypothesis is that these biases arise through a process of statistical inference, in which information from noisy measurements is fused with a probabilistic model of the environment. However, such inference is optimal only if the observer's internal model matches the environment. We found this to be the case. We measured performance in an orientation-estimation task and found that orientation judgments were more accurate at cardinal (horizontal and vertical) orientations. Judgments made under conditions of uncertainty were strongly biased toward cardinal orientations. We estimated observers' internal models for orientation and found that they matched the local orientation distribution measured in photographs. In addition, we determined how a neural population could embed probabilistic information responsible for such biases.

© 2011 Nature America, Inc. All rights reserved.

Perception has been described as a process of unconscious inference, generating a best guess about the structure of the scene given the available sensory measurements and an internal model of the world[1,2]. When sensory information is degraded, reliance on an internal model becomes advantageous. When reaching for the light switch in a dim room, for example, the knowledge that it is likely to be near the door is usually beneficial.

The Bayesian framework provides a quantitative formulation of the inference problem and has been used to explain aspects of human perception[3–5], cognition[6] and visuo-motor control[7,8]. A Bayesian observer's internal model of an environmental attribute is represented using a prior probability distribution. To achieve optimal inference, the observer's prior should be matched to the actual distribution of the attribute in the environment, which we refer to as the environmental distribution. Here we provide direct evidence that this requirement is satisfied for the perceptual task of estimating the orientation of local image structure.

To test the correspondence between the environmental distribution and observer's prior, one must know or estimate these two distributions. Many previous studies have exposed observers to stimuli drawn from a known distribution[7,8] and compared their behavior to that of an optimal observer with full knowledge of that distribution. Such an approach is limited by the ability of humans to internalize the distribution in the time frame of the experiment and leaves open the question of what prior knowledge humans use under natural conditions. Alternatively, one can attempt to derive an environmental distribution from known properties of the environment and the image-formation process. For example, sunlight comes from above, and humans appear to make use of this information[9]. However, a precise description of the illumination in an arbitrary scene is extraordinarily complex, and development of a direct expression for the illumination distribution is therefore likely to be intractable. As another alternative, one can attempt to estimate the quantity of interest directly from a large collection of photographic images, generating a distribution by binning values into histograms[4,10] or fitting a parametric form[9,11,12]. This process can also be difficult; measurement of the distribution of retinal image motion, for example, requires a head-mounted video camera and eye-tracker and an accurate algorithm for estimating motion from video footage. Local image orientation, on the other hand, is relatively easy to estimate, and we used a histogram of such estimates to approximate the environmental distribution.

Determination of an observer's internal prior distribution can also be difficult. One can assume a particular parametric form and use it to fit perceptual bias data[9,11], but this is only useful if the observer's prior can be well approximated by the chosen form. We used a recently developed methodology for estimating a nonparametric prior from measurements of perceptual bias and variability[13] and found that the recovered observer's prior and measured environmental distribution are well matched, thus providing direct evidence that humans behave according to the rules of Bayesian inference in estimating the orientation of local image structure. In particular, we measured the distribution of local orientations in a collection of photographic images and found it to be strongly non-uniform, exhibiting a preponderance of cardinal orientations. We also measured human observers' bias and variability when comparing oriented stimuli under different uncertainty conditions and found strong biases toward the cardinal axes when stimuli were more uncertain. The prior that best explains the observed pattern of bias and variability is well matched to the environmental distribution. Finally, the most distinctive and well understood property of neurons in primary visual cortex is their selectivity for local orientation, and we demonstrated how a population of such neurons with inhomogeneities matching those previously reported can give rise to both the observed perceptual biases and discriminability.

[1]Department of Psychology, New York University, New York, New York, USA. [2]Center for Neural Science, New York University, New York, New York, USA. [3]Howard Hughes Medical Institute, New York University, New York, New York, USA. [4]Courant Institute for Mathematical Sciences, New York University, New York, New York, USA. Correspondence should be addressed to A.R.G. (ahna.girshick@gmail.com).

Received 28 December 2010; accepted 5 April 2011; published online 5 June 2011; doi:10.1038/nn.2831

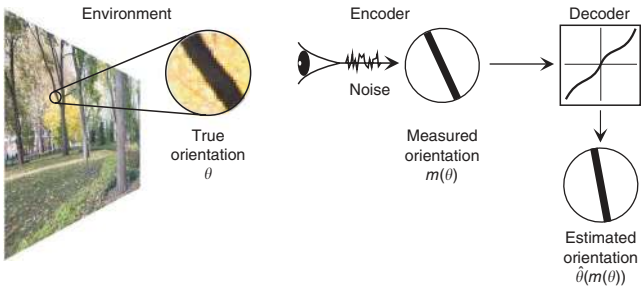

Figure 1 Observer model for local two-dimensional orientation estimation. Each local edge has a true image orientation, θ. In the encoding stage, the observer obtains a visual measurement, $m(\theta)$, corrupted by sensory noise. In the decoding stage, a function (the estimator, black curve) is applied to the measurement to produce the estimated orientation $\hat{\theta}(m(\theta))$. Because of the sensory noise, the estimated orientation will exhibit variability across repeated presentation of the same stimulus, and may also exhibit a systematic bias relative to the true orientation.

RESULTS

We assumed that the observer's sensory measurements, internal prior model and estimates are related to each other through an encoder-decoder observer model[13] (**Fig. 1**). Given a contour or edge with true orientation θ in the retinal image, the observer makes an internal measurement, $m(\theta)$, which is corrupted by sensory noise. Repeated presentations of the same stimulus result in slightly different measurements, and this collection of measurements can be described with a conditional probability function, $p(m|\theta)$. An estimator (decoder) is a function, typically nonlinear, that maps the noisy measurement to an estimate of the true orientation, $\hat{\theta}(m(\theta))$. A Bayesian observer uses a particular estimator, optimized to minimize expected error for the given measurement noise and prior model (**Fig. 2** and Online Methods).

The Bayesian encoder-decoder model provides a framework for understanding two fundamental types of perceptual error: bias and variability[13]. Perceptual variability arises from the sensory (encoder) noise, which is propagated through the estimator, and limits the precision with which the observer can discriminate stimuli. Perceptual bias is the average mismatch between the perceived and true orientations and primarily arises because a Bayesian estimator prefers interpretations that have higher prior probability from amongst those consistent with the measurement. The origin of these two types of error makes it clear that one cannot predict one from the other. In particular, the direction and magnitude of bias are determined by the slope of the prior, which need not have any relationship to the variability.

Psychophysical measurements of bias and variability

We conducted an experiment to estimate observers' prior knowledge about orientation. We asked observers to compare the average orientations of two arrays of oriented Gabor patches (Gaussian-windowed sinusoidal gratings). We assumed that observers make local orientation estimates for each patch and then averaged these to obtain an overall estimate of orientation for each array[14]. The low-noise stimuli (L; **Fig. 3a**) had identical orientations, whereas the high-noise stimuli (H; **Fig. 3a**) were variable in orientation (individual Gabor orientations were drawn from a distribution with an s.d. of about 22 deg; see Online Methods). Observers viewed the two stimuli simultaneously, symmetrically displaced to the right and left of fixation, and were asked to indicate whether the mean orientation of the right stimulus was clockwise or counter-clockwise relative to that of the left stimulus. Comparisons were made between three stimulus combinations: low noise versus low noise, high noise versus high noise, and high noise versus low noise.

The discrimination thresholds are of interest because they can be related to the s.d. of the internal measurement distributions (see Online Methods). The two same-noise conditions (low noise versus low noise and high noise versus high noise; **Fig. 3b** and **Supplementary Fig. 1a**) provide an estimate of measurement noise

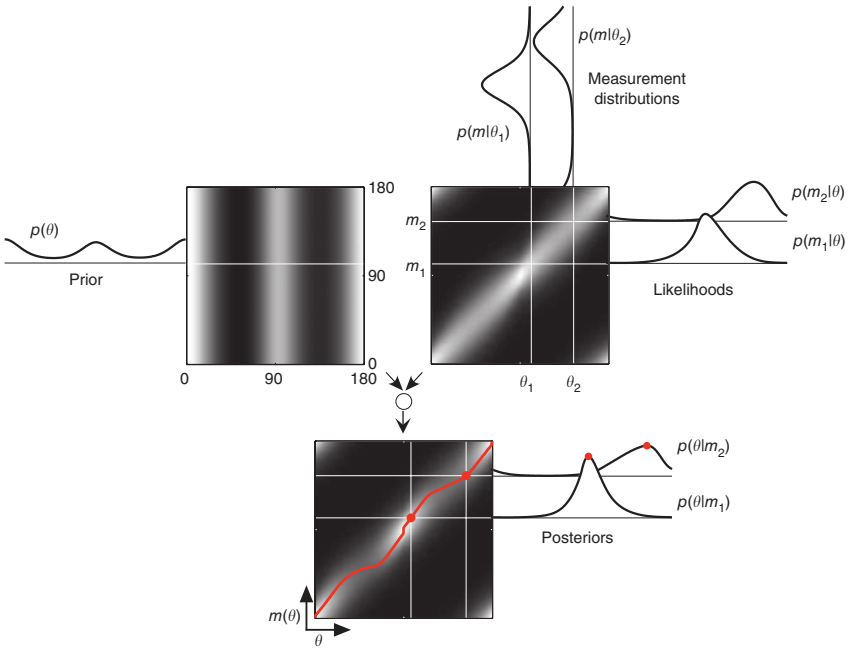

Figure 2 Derivation of the estimator $\hat{\theta}(m(\theta))$. In all three grayscale panels, the horizontal axis is stimulus orientation (θ), the vertical axis is the measured orientation ($m(\theta)$) and the intensity corresponds to probability. Upper left, the mean observer's prior, raised to the power of 2.25 and re-normalized for visibility, is independent of the measurements (that is, all horizontal slices are identical). Upper right, the conditional distribution, $p(m|\theta)$. Vertical slices indicate measurement distributions, $p(m|\theta_1)$ and $p(m|\theta_2)$, for two particular stimuli θ_1 and θ_2. The widths of the measurement distributions are the average of those for the low- and high-noise conditions for the mean observer (multiplied by a factor of 10 for visibility). Horizontal slices, $p(m_1|\theta)$ and $p(m_2|\theta)$, describe the likelihood of the stimulus orientation, θ, for the particular measurements, m_1 and m_2. Note that the likelihoods are not symmetric, as the measurement distribution width depends on the stimulus orientation. Bottom, the posterior distribution is computed using Bayes' rule, as the normalized product of the prior and likelihood (top two panels). Horizontal slices correspond to posterior distributions $p(\theta|m_1)$ and $p(\theta|m_2)$, which describe the probability of a stimulus orientation given two particular measurements. Red dots indicate maximum a posteriori estimates (the modes of the posterior) for these two likelihoods, $\hat{\theta}(m_1)$ and $\hat{\theta}(m_2)$. Circular mean estimates yield similar results (see **Supplementary Fig. 2**). The red curve shows the estimator $\hat{\theta}(m)$ computed for all measurements. An unbiased estimator would correspond to a straight line along the diagonal.

© 2011 Nature America, Inc. All rights reserved.

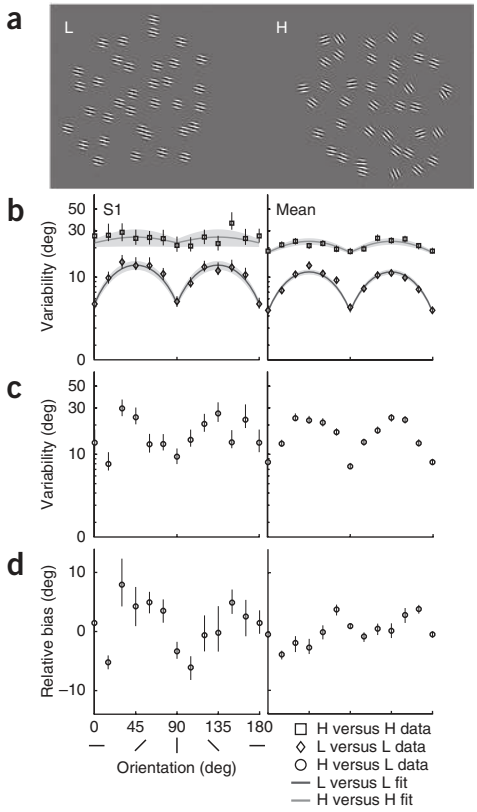

Figure 3 Stimuli and experimental results. (**a**) Stimuli are arrays of oriented Gabor functions (contrast increased for illustrative purposes). Left, a low-noise stimulus (L). Right, a high-noise stimulus (H) with mean orientation slightly more clockwise. Observers indicated whether the right stimulus was oriented counter-clockwise or clockwise relative to the left stimulus. (**b**) Variability for the same-noise conditions for representative subject S1 (left) and the mean subject (right), expressed as the orientation discrimination threshold (that is, JND). Mean subject values are computed by pooling raw choice data from all five subjects. Error bars indicate 95% confidence intervals. Dark gray and light gray curves are fitted rectified sinusoids, used to estimate the widths of the underlying measurement distributions. Pale gray regions indicate ± 1 s.d. of 1,000 bootstrapped fits. (**c**) Cross-noise (high noise versus low noise) variability data (circles). The horizontal axis is the orientation of the high-noise stimulus. (**d**) Relative bias, expressed as the angle by which the high-noise stimulus must be rotated counter-clockwise so as to be perceived as having the same mean orientation as the low-noise stimulus.

for each class of stimuli. For the low-noise stimuli, all subjects exhibited better discrimination at the cardinals, a well-studied behavior known as the oblique effect[15]. As there was no noise in the stimuli, these inhomogeneities must arise from non-uniformity in the amplitude of the internal noise at different orientations. This effect was diminished with the high-noise stimuli, for which the inhomogeneous internal noise is presumably dominated by external stimulus noise. As expected, discrimination thresholds were significantly higher for the high-noise stimuli than the low-noise stimuli for all subjects (98% of all just noticeable differences (JNDs) across orientations and subjects, sign test $P \approx 0$). The cross-noise variability data (**Fig. 3c** and **Supplementary Fig. 1b**) show a moderate oblique effect whose strength lies between that of the low noise versus low noise and high noise versus high noise conditions (98% of high noise versus low noise JNDs are larger than low noise versus low noise JNDs, sign test $P \approx 0$; 73% of high noise versus low noise JNDs are smaller than high noise versus high noise JNDs, sign test $P < 0.0005$).

A non-uniform prior will cause a bias in estimation. Biases are not observable when comparing same-noise stimuli, as both stimuli presumably have the same bias. Cross-noise comparisons can be used to estimate relative bias[13] (that is, the difference between the low- and high-noise biases) by computing the difference between the mean orientation of the two stimuli when they are perceived to be equal. This represents the counter-clockwise rotation that must be applied to the high-noise stimulus to perceptually match the orientation of the low-noise stimulus (**Fig. 3d** and **Supplementary Fig. 1c**). All subjects showed a systematic bimodal relative bias, indicating that a high-noise stimulus was perceived to be oriented closer to the nearest cardinal orientation (that is, vertical or horizontal) than the low-noise stimulus of the same orientation. The relative bias was 0 at the cardinal and oblique orientations, and as large as 12 deg in between. These relative biases suggest that perceived orientations are attracted toward the cardinal directions and repelled from the obliques, and that these effects are stronger for the high-noise stimuli (**Fig. 4**).

Estimation of observers' likelihood and prior

If our human observers are performing Bayesian inference, what is the form of the prior probability distribution that they are using? We assume that our observers select the most probable stimulus according to the posterior density $p(\theta|m)$ (known as the maximum a posteriori estimate). We noted that the circular mean of the posterior produced similar estimates, as the posterior distributions are only slightly asymmetric (**Supplementary Fig. 2**). According to Bayes' rule, the posterior is the product of the prior $p(\theta)$ and the likelihood function $p(m|\theta)$, normalized so that it integrates to 1. We assume that the decoder is based on the correct likelihood function, which is simply the measurement noise distribution, interpreted not as a probability distribution over measurements but as a function of the stimulus for a particular measurement. That is, we assume the observer knows and takes into account the uncertainty of each type of stimulus[16] (see Online Methods).

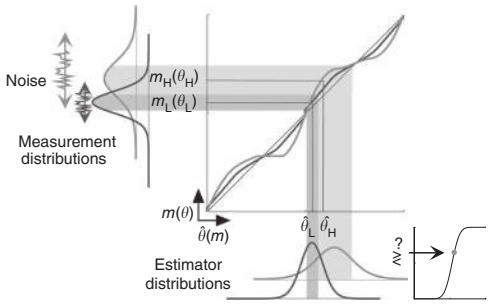

Figure 4 Example cross-noise comparison. The vertical axis is the measured orientation, $m(\theta)$, and the horizontal axis is estimated stimulus orientation, $\hat{\theta}(m(\theta))$. Measurements corresponding to low-noise stimuli, $m_L(\theta_L)$ (dark gray), or high-noise stimuli, $m_H(\theta_H)$ (light gray), enter on the left. Each measurement is transformed by the appropriate nonlinear estimator (solid curves) into an estimate (bottom). The estimators correspond to those of the mean observer exaggerated for illustration as in **Figure 2**. The high-noise estimator exhibits larger biases than the low-noise estimator. The sensory noise of the measurements propagates through the estimator, resulting in estimator distributions (note these should not be confused with the posteriors). Comparison of these distributions produces a single point on the psychometric function.

© 2011 Nature America, Inc. All rights reserved.

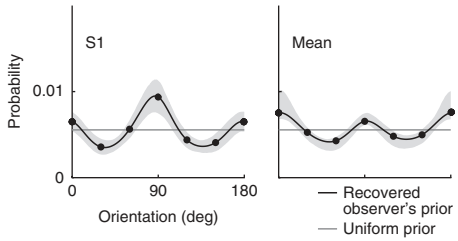

Figure 5 Recovered priors for subject S1 and mean subject. The control points of the piecewise cubic spline (see Online Methods) are indicated by black dots. The gray error region shows ± 1 s.d. of 1,000 bootstrapped estimated priors.

The observer model (**Fig. 1**) provides a link between the likelihood and prior and the two experimentally accessible aspects of perceptual behavior: bias and variability. Perceptual variability is caused by variability in the estimates, $\hat{\theta}(m(\theta))$, which arises from variability in the measurements, $m(\theta)$. Relative bias corresponds to the difference in orientation between two stimuli of different uncertainty, $\theta_L - \theta_H$, whose estimates are (on average) equal, $\hat{\theta}_H(m_H(\theta_H)) = \hat{\theta}_L(m_L(\theta_L))$. Note that the two estimator functions, $\hat{\theta}_H$ and $\hat{\theta}_L$, are dependent on noise level. These relationships allow us to estimate the likelihood width and prior (as functions of orientation) from the experimentally measured bias and variability[13] (see Online Methods). Specifically, we obtained the likelihood functions directly from the same-noise variability data (**Fig. 3b** and **Supplementary Fig. 1a**). We represented the prior as a smooth curve and determined its shape for each observer by maximizing the likelihood of the raw cross-noise data. The recovered priors of all observers were bimodal, with peaks at the two cardinal orientations (**Fig. 5** and **Supplementary Fig. 3**).

Environmental orientation distribution

It has been suggested that the prevalence of vertical and horizontal orientations in the environment is the underlying cause of the anisotropy of orientation discriminability (that is, the oblique effect)[17]. Orientation content in images is often studied by averaging the Fourier amplitude spectrum over all spatial scales[18,19]. For the purposes of our study, we defined the environmental distribution as the probability distribution over local orientation in an ensemble of visual images[17], measured at a spatial scale roughly matched to peak human sensitivity (approximately the same as the scale of our experimental stimuli).

We obtained our measurements from a large database of photographs of scenes of natural content. We estimated the local image gradients by convolution with a pair of rotation-invariant filters[20], identified strongly oriented regions, computed their dominant orientations (**Fig. 6a**) and formed histograms of these values. The resulting estimated environmental distribution indicates a predominance of cardinal orientations (**Fig. 6b**). This is consistent with the orientation priors that we recovered from our human subjects (**Fig. 5** and **Supplementary Fig. 3**), and therefore explains the cardinal biases in their perception. We chose the spatial scale that corresponds most closely to our 4 cycles per deg experimental stimuli and human peak spatial frequency sensitivity of 2–5 cycles per deg[21]. We found that this choice did not have a strong effect on the results: the dominance of cardinal orientations was similar across spatial scales (**Supplementary Fig. 4**).

Observers' priors versus the environmental distribution

We compared the estimated human observers' priors and environmental distribution, both directly (as probability distributions) and in terms of their predicted perceptual effects (bias and variability in cross-noise comparisons). The observers' prior probability distributions and the environmental distribution all had local maxima at the cardinals and minima at the obliques, and the heights of the peaks and troughs were quite similar (**Fig. 7a**). We computed perceptual predictions of the trial-by-trial behavior of the Bayesian encoder-decoder model by comparing simulated responses to each pair of stimuli shown to our observers. We found that the relative variability (**Fig. 7b** and **Supplementary Fig. 1b**) and bias (**Fig. 7c** and **Supplementary Fig. 1c**) are similar for a model that uses either the environmental distribution or the human observer's prior, and both closely resemble the human behavior.

To assess the strength of this result, we also considered the null hypothesis that observers use a uniform prior (equivalent to assuming that observers perform maximum-likelihood estimation). A Bayesian-observer model with a uniform prior does not produce the distinct bimodal relative bias (**Fig. 7c** and **Supplementary Fig. 1b**). Instead, this model either produces no bias (for example, mean subject and subjects S1, S3, S4 and S5) or a small relative bias away from the cardinal orientations (for example, subject S2). This repulsive relative bias is a result of the asymmetrical shape of the likelihoods near the cardinals, which pushes the low-noise estimates toward the cardinals more than the high-noise estimates. Furthermore, the uniform-prior observer predicts little or no oblique effect for the cross-noise condition, unlike the human observers (**Fig. 7b** and **Supplementary Fig. 1c**). This indicates that the human observers' biases cannot arise purely from inhomogeneities in sensory noise but require a non-uniform prior.

We also compared the ability of Bayesian encoder-decoder models with different priors to explain the raw experimental data. We computed the log likelihoods of the two non-uniform prior models and linearly rescaled them so that a value of 0 corresponds to the uniform-prior model (degrees of freedom = 0) and a value of 1 corresponds to the raw psychometric fits (degrees of freedom = 24; **Fig. 7d**). In general, a Bayesian observer with the recovered observer's prior (degrees of freedom = 6) performed quite well, often on a par with the raw psychometric fits to the data. For the mean observer, a Bayesian observer using the environmental distribution (degrees of freedom = 0) as a prior predicted the data even better than using the observer's recovered prior and better than the psychometric fits. It is important to note that these models are not nested; the recovered observer's prior is constrained to a family of smooth shapes (see Online Methods) and cannot fully capture the peakedness of the environmental distribution. These results provide strong support of the hypothesis that human observers use prior knowledge of the non-uniform orientation statistics of the environment.

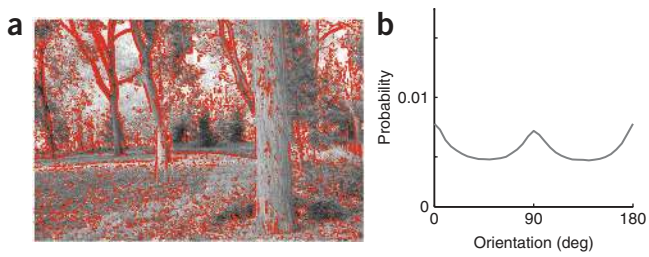

Figure 6 Natural image statistics. (**a**) Example natural scene from **Figure 1**, with strongly oriented locations marked in red. (**b**) Orientation distribution for natural images (gray curve).

© 2011 Nature America, Inc. All rights reserved.

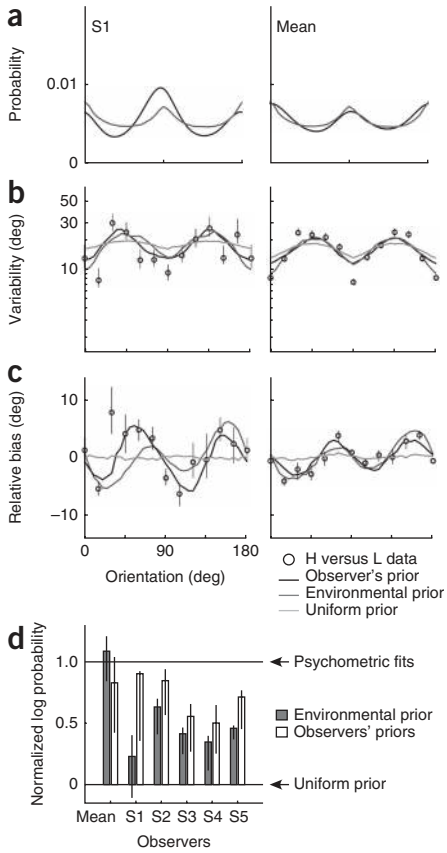

Figure 7 Comparison of human observers' priors and environmental distribution for subject S1 (left) and the mean subject (right). (**a**) Human observers' priors (black curves, from **Fig. 5**) and environmental distribution from natural images (medium gray curve, from **Fig. 6b**). (**b**) Cross-noise variability data (circles, from **Fig. 3c**) with predictions of the two Bayesian-observer models using each of the three priors shown in **a** and the uniform prior. The uniform prior predicts little or no effect of stimulus orientation on discrimination (light gray curves). In contrast, both the environmental prior (medium gray curves) and the recovered human observers' priors (black curves) predict better discrimination at the cardinals, as seen in the human observers. (**c**) Relative bias data (circles, from **Fig. 3d**) with the predictions of Bayesian-observer models using three priors shown in **a**. The uniform prior predicts no bias or a small bias in the opposite direction (for example, Subject S2 in **Supplementary Fig. 1c**). In contrast, both non-uniform priors predict the bimodal bias exhibited by human observers. (**d**) Normalized log likelihood of the data for Bayesian-observer models using two different priors: environmental distribution and the recovered observer's prior. Error bars denote the 5th and 95th percentiles from 1,000 bootstrap estimates. Values greater than 1 indicate performance better than that of the raw psychometric fits, whereas values less than 0 indicate performance worse than that obtained with a uniform prior.

© 2011 Nature America, Inc. All rights reserved.

DISCUSSION

We found that humans exploit inhomogeneities in the orientation statistics of the environment and use them when making judgments about visual orientation. Although previous work has largely focused on variability in orientation estimation (that is, the oblique effect), our results emphasize the importance of bias, which is essential for studying the prior used by the observer[13]. Our findings directly demonstrate a systematic perceptual bias toward the cardinals, which had been hypothesized in previous work[22]. We used the measured bias and variability to estimate observers' internal prior distributions and found that these were well matched to estimates of the environmental distribution, exhibiting significant peaks at the cardinal orientations. In addition, when used as prior probabilities in a Bayesian-observer model, both distributions accurately predicted our human observers' perceptual biases.

Intuitively, one might have expected that a large bias would be a direct consequence of large variability, but our results indicate that this is not the case. Our observers' orientation discriminability was worst at oblique angles and best at the cardinals (**Fig. 3b**), but their biases were approximately 0 at both the obliques and cardinals and largest in between (**Fig. 3d**). These results are consistent with a Bayesian observer that is aware of the environmental prior: the bias is approximately equal to the product of the variability and the slope of the (log) prior[13]. It is small at the cardinals because the variability is small (and the prior slope reverses sign), and it is small at the oblique orientations because the prior slope is small (and again, reverses sign). The well-known 'tilt aftereffect', in which adaptation to an oriented stimulus causes a subsequent reduction in variability and increase in bias, provides another example in which bias and variability do not covary[23].

Our observer model (**Fig. 1**) can be extended in a number of ways. For example, the decoding stage is deterministic, but could be made stochastic (that is, incorporating additional noise). In the context of simulating our experimental task, this additional variability would act as a sort of 'decision noise', as has been used in may previous models of perceptual decision making[24]. This would, of course, entail an additional parameter that would need to be constrained by data. Our model (as well as our experiment) is also currently limited to estimation of retinal orientation. However, perception is presumably intended to facilitate one's interactions with a three-dimensional world of objects and surfaces. A full Bayesian model of orientation perception should incorporate priors on the three-dimensional orientation of contours, observer orientation and perspective image formation, and would perhaps aim to recover aspects of the three-dimensional scene[25]. From this perspective, one can think of our model as effectively capturing a two-dimensional marginal of a full three-dimensional orientation prior.

A critical advantage of the Bayesian modeling framework is that the fundamental ingredients (likelihood, prior) have distinct meanings that extend beyond fitting the data in our experiments. As such, the model makes testable predictions. Our perceptual measurements were obtained under a specific set of viewing conditions, including stimulus size, eccentricity, duration, contrast and spatial frequency, which jointly determine the measurement noise. If we were to repeat these measurements having altered one of the viewing conditions that does not covary with orientation in the environment (for example, contrast), we would expect subject responses to be consistent with the same orientation prior[26]. For changes in viewing conditions that do covary with orientation in the environment (for example, spatial frequency; **Supplementary Fig. 4**), we expect that a multi-dimensional prior (for example, ref. 10) would be required to explain perceptual biases.

The timescale over which priors are developed is an important open question in Bayesian modeling of perception. In some cases, priors appear to adapt over relatively short timescales[7,8]. In contrast, our observers seemed to use priors that are matched to the statistics of natural scenes, as opposed to human-made or blended scenes (**Supplementary Fig. 4**), perhaps suggesting that they are adapted over very long timescales. Nevertheless, the variation in the recovered priors of our observers might reflect differences in their previous perceptual experience.

The question of whether orientation priors are innate or learned may also be related to the development of the oblique effect.

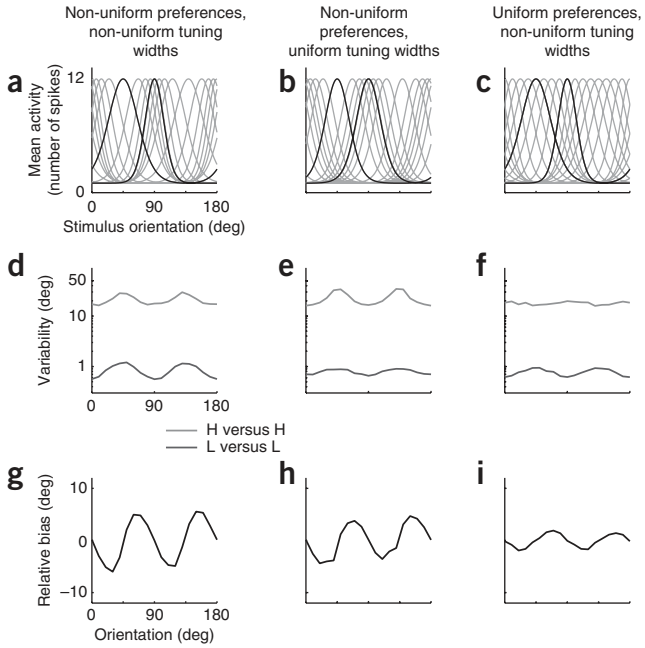

Figure 8 Simulations of neural models with non-uniform encoder and population vector decoder. (**a**) Tuning curves of an encoder population with non-uniform orientation preferences and non-uniform tuning widths based on neurophysiology (only a subset of neurons shown). Neurons preferring 45 deg and 90 deg stimuli are highlighted in black. (**b**) Tuning curves of a population with non-uniform preferences and uniform widths. (**c**) Tuning curves of a population with uniform preferences and non-uniform widths. (**d**–**f**) Variability for the same-noise conditions for the populations in **a**–**c**: low noise versus low noise (dark gray) and high noise versus high noise (light gray). (**g**–**i**) Relative bias for the cross-noise condition (high noise versus low noise) for the populations in **a**–**c**. The fully non-uniform population (**a**) produces variability and bias curves similar to those exhibited by humans (**Fig. 3b,d** and **Supplementary Fig. 1a,c**).

In support of the innate hypothesis is evidence of the oblique effect in 12-month-old human infants[27], and a large variation in the strength of the oblique effect amongst observers raised in similar human-made environments[28]. In support of the hypothesis that the prior is learned in our lifetimes is evidence that the strength of the oblique effect continues to grow from childhood through adulthood[29], and that early visual deprivation affects neural orientation sensitivity in kittens[30] and humans[31]. Both of these hypotheses could explain the slight individual differences that we observed in the strength of the oblique effect and shape of the prior. Lastly, the oblique effect has been shown to be consistent with retinal, rather than gravitational, coordinates[32], whereas our environmental prior is most naturally associated with gravitational coordinates (assuming all photographs in the database were taken with the camera held approximately level). The simplest interpretation is that the prior emerges from retinal stimulation, but because humans tend to keep their heads vertical, retinal coordinates are generally matched to gravitational coordinates.

Finally, we consider the physiological instantiation of our encoder-decoder observer model (**Fig. 1**). The encoding stage of the model is most naturally associated with orientation-selective neurons in primary visual cortex (area V1). Non-uniformities in orientation discriminability (the oblique effect) have been posited to arise because of non-uniformities in the representation of orientation in the V1 population. Specifically, a variety of measurements (single-unit recording[33,34], optical imaging[35] and functional magnetic resonance imaging[36]) have shown that cardinal orientations are represented by a disproportionately large fraction of V1 neurons, and that those neurons also tend to have narrower tuning curves[33]. Additional non-uniformities have been reported in V1 and elsewhere (for example, variations in gain, baseline firing rate or correlations in responses), and these may also contribute to non-uniformities in perceptual discriminability.

Now consider the Bayesian decoding stage. For a uniform population of encoder neurons with Poisson spiking statistics, the log-likelihood may be expressed as the sum of the log tuning curves, weighted by the spike counts[37,38] (for a non-uniform population, this needs to be corrected by subtracting the sum of the tuning curves). A decoder population could compute the log-posterior at each orientation by computing this weighted sum and adding the log prior. The desired estimate would be the orientation associated with the decoder neuron having maximum response (that is, winner take all). Although this is an explicit computation of the Bayesian estimator, and would therefore generate the same perceptual biases seen in our subjects, several aspects of this implementation seem implausible[39]. In particular, the decoder utilizes an entire population of cells to 'recode' the information in the encoder population; the decoder must have complete knowledge of the encoder tuning curves as well as the prior, and the estimate is obtained with a winner-take-all mechanism that is highly sensitive to noise.

As an alternative, we wondered whether the optimal mapping from an encoder population to an estimate might be approximated by a simpler, more plausible computation (also see ref. 40). We simulated an encoder population whose spiking responses were generated as samples from independent Poisson distributions, with mean rates (in response to an oriented stimulus) determined by a set of orientation tuning curves with non-uniformities approximating those found in V1 (refs. 33–35,41) (**Fig. 8**). We used a standard population vector decoder[42], which computes the sum of directional vectors associated with the preferred orientation of each encoder cell, weighted by the response of that cell. This decoder is robust to noise (compared with winner take all), does not require any knowledge of the prior and requires only the preferred orientation of each cell in the encoding population (as opposed to the entire set of tuning curves). We used the full neural encoder-decoder model to simulate per-trial decisions on the same stimuli shown to our human subjects. From these, we computed discriminability for the same-noise conditions and bias for the cross-noise conditions. Notably, we found that this generic decoder, when combined with the non-uniform encoding population (**Fig. 8a**), produced bias and discriminability curves (**Fig. 8d**,**g**) that were quite similar to those of our human subjects (**Fig. 7b**,**d**). The narrower tuning near the cardinals leads to a substantial decrease in discrimination thresholds; eliminating the non-uniform tuning widths (**Fig. 8b**) greatly reduces the oblique effect for the low-noise stimuli (**Fig. 8e**). The proportionally larger number of cells around the cardinals leads to a disproportionate influence on the weighted sum decoder, inducing biases toward the cardinals; eliminating the non-uniform orientation preferences (**Fig. 8c**) substantially reduces the amplitude of the bias (**Fig. 8i**). Finally, eliminating both non-uniformities (that is, an equal-spaced population of neurons with identical tuning, as is commonly assumed in the population coding literature[37,43,44]) produces constant bias and discriminability curves, as expected (data not shown).

The fact that this observer model is able to approximate the bias behavior of the optimal Bayesian decoder implies that the non-uniformities of the encoder implicitly capture the properties of the prior in such a way that they can be properly used by a population-vector decoder[40]. Recent theoretical work posits that the non-uniformities

© 2011 Nature America, Inc. All rights reserved.

of neural populations reflect a strategy by which the brain allocates its resources (neurons and spikes) so as to optimally encode stimuli drawn from an environmental distribution[45]. Our findings suggest that encoder non-uniformities may also serve the role of enabling a decoder to perform Bayesian inference without explicit knowledge of the prior[39]. If so, this solution could provide a universal mechanism by which sensory systems adapt themselves to environmental statistics, allowing for optimal representation and extraction of sensory attributes under limitations of neural resources.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/natureneuroscience/.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS
We wish to thank D. Ganguli for helpful discussions on image statistics and population coding, T. Saarela for the monitor calibration and N. Solomon for assisting with the experiments. This work was funded by US National Institutes of Health grants EY019451 (A.R.G.) and EY16165 (M.S.L.), and the Howard Hughes Medical Institute (E.P.S.).

AUTHOR CONTRIBUTIONS
A.R.G., M.S.L. and E.P.S. contributed to the design of the experiments, design of the analyses and the writing of the manuscript. A.R.G. conducted the experiments and performed the analyses.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Published online at http://www.nature.com/natureneuroscience/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Helmholtz, H. *Treatise on Physiological Optics* (Thoemmes Press, Bristol, UK, 2000).
2. El-Bizri, N. A philosophical perspective on Alhazen's optics. *Arab. Sci. Philos.* **15**, 189–218 (2005).
3. Knill, D.C. & Richards, W. *Perception as Bayesian Inference* (Cambridge University Press, Cambridge, UK, 1996).
4. Geisler, W.S., Perry, J.S., Super, B.J. & Gallogly, D.P. Edge co-occurrence in natural images predicts contour grouping performance. *Vision Res.* **41**, 711–724 (2001).
5. Simoncelli, E.P. *Distributed Analysis and Representation of Visual Motion*. PhD thesis, Massachusetts Institute of Technology (1993).
6. Griffiths, T.L., Kemp, C. & Tenenbaum, J.B. Bayesian models of cognition. in *Cambridge Handbook of Computational Psychology* (ed. Sun, R.) 59–100 (Cambridge University Press, 2008).
7. Körding, K.P. & Wolpert, D.M. Bayesian integration in sensorimotor learning. *Nature* **427**, 244–247 (2004).
8. Tassinari, H., Hudson, T.E. & Landy, M.S. Combining priors and noisy visual cues in a rapid pointing task. *J. Neurosci.* **26**, 10154–10163 (2006).
9. Mamassian, P. & Goutcher, R. Prior knowledge on the illumination position. *Cognition* **81**, 1–9 (2001).
10. Burge, J., Fowlkes, C.C. & Banks, M.S. Natural-scene statistics predict how the figure-ground cue of convexity affects human depth perception. *J. Neurosci.* **30**, 7269–7280 (2010).
11. Weiss, Y., Simoncelli, E.P. & Adelson, E.H. Motion illusions as optimal percepts. *Nat. Neurosci.* **5**, 598–604 (2002).
12. Brainard, D.H. *et al.* Bayesian model of human color constancy. *J. Vis.* **6**, 1267–1281 (2006).
13. Stocker, A.A. & Simoncelli, E.P. Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* **9**, 578–585 (2006).
14. Dakin, S.C. Information limit on the spatial integration of local orientation signals. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **18**, 1016–1026 (2001).
15. Appelle, S. Perception and discrimination as a function of stimulus orientation: the "oblique effect" in man and animals. *Psychol. Bull.* **78**, 266–278 (1972).
16. Landy, M.S., Goutcher, R., Trommershäuser, J. & Mamassian, P. Visual estimation under risk. *J. Vis.* **7**, 1–15 (2007).
17. Coppola, D.M., Purves, H.R., McCoy, A.N. & Purves, D. The distribution of oriented contours in the real world. *Proc. Natl. Acad. Sci. USA* **95**, 4002–4006 (1998).
18. van der Schaaf, A. & van Hateren, J.H. Modelling the power spectra of natural images: statistics and information. *Vision Res.* **36**, 2759–2770 (1996).
19. Switkes, E., Mayer, M.J. & Sloan, J.A. Spatial frequency analysis of the visual environment: anisotropy and the carpentered environment hypothesis. *Vision Res.* **18**, 1393–1399 (1978).
20. Farid, H. & Simoncelli, E.P. Differentiation of discrete multidimensional signals. *IEEE Trans. Image Process.* **13**, 496–508 (2004).
21. Campbell, F.W. & Robson, J.G. Application of Fourier analysis to the visibility of gratings. *J. Physiol. (Lond.)* **197**, 551–566 (1968).
22. Tomassini, A., Morgan, M.J. & Solomon, J.A. Orientation uncertainty reduces perceived obliquity. *Vision Res.* **50**, 541–547 (2010).
23. Schwartz, O., Hsu, A. & Dayan, P. Space and time in visual context. *Nat. Rev. Neurosci.* **8**, 522–535 (2007).
24. Shadlen, M.N., Britten, K.H., Newsome, W.T. & Movshon, J.A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**, 1486–1510 (1996).
25. Cohen, E.H. & Zaidi, Q. Fundamental failures of shape constancy resulting from cortical anisotropy. *J. Neurosci.* **27**, 12540–12545 (2007).
26. Maloney, L.T. & Mamassian, P. Bayesian decision theory as a model of human visual perception: testing Bayesian transfer. *Vis. Neurosci.* **26**, 147–155 (2009).
27. Birch, E.E., Gwiazda, J., Bauer, J.A. Jr. Naegele, J. & Held, R. Visual acuity and its meridional variations in children aged 7–60 months. *Vision Res.* **23**, 1019–1024 (1983).
28. Timney, B.N. & Muir, D. Orientation anisotropy: incidence and magnitude in Caucasian and Chinese subjects. *Science* **193**, 699–701 (1976).
29. Mayer, M.J. Development of anisotropy in late childhood. *Vision Res.* **17**, 703–710 (1977).
30. Blakemore, C. & Cooper, G.F. Development of the brain depends on the visual environment. *Nature* **228**, 477–478 (1970).
31. Mitchell, D.E., Freeman, R.D., Millodot, M. & Haegerstrom, G. Meridional amblyopia: evidence for modification of the human visual system by early visual experience. *Vision Res.* **13**, 535–558 (1973).
32. Orban, G.A., Vandenbussche, E. & Vogels, R. Human orientation discrimination tested with long stimuli. *Vision Res.* **24**, 121–128 (1984).
33. Li, B., Peterson, M.R. & Freeman, R.D. The oblique effect: a neural basis in the visual cortex. *J. Neurophysiol.* **90**, 204–217 (2003).
34. De Valois, R.L., Yund, E.W. & Hepler, N. The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res.* **22**, 531–544 (1982).
35. Chapman, B. & Bonhoeffer, T. Overrepresentation of horizontal and vertical orientation preferences in developing ferret area 17. *Proc. Natl. Acad. Sci. USA* **95**, 2609–2614 (1998).
36. Furmanski, C.S. & Engel, S.A. An oblique effect in human primary visual cortex. *Nat. Neurosci.* **3**, 535–536 (2000).
37. Jazayeri, M. & Movshon, J.A. Optimal representation of sensory information by neural populations. *Nat. Neurosci.* **9**, 690–696 (2006).
38. Zhang, K., Ginzburg, I., McNaughton, B.L. & Sejnowski, T.J. Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J. Neurophysiol.* **79**, 1017–1044 (1998).
39. Simoncelli, E.P. Optimal estimation in sensory systems. in *The Cognitive Neurosciences IV* (ed. Gazzaniga, M.) 525–535 (MIT Press, 2009).
40. Fischer, B.J. Bayesian estimates from heterogeneous population codes. *Int. Jt. Conf. Neural Netw.* 1–7 (2010).
41. Ringach, D.L., Shapley, R.M. & Hawken, M.J. Orientation selectivity in macaque V1: diversity and laminar dependence. *J. Neurosci.* **22**, 5639–5651 (2002).
42. Georgopoulos, A.P., Schwartz, A.B. & Kettner, R.E. Neuronal population coding of movement direction. *Science* **233**, 1416–1419 (1986).
43. Ma, W.J., Beck, J.M., Latham, P.E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**, 1432–1438 (2006).
44. Seung, H.S. & Sompolinsky, H. Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. USA* **90**, 10749–10753 (1993).
45. Ganguli, D. & Simoncelli, E.P. Implicit encoding of prior probabilities in optimal neural populations. in *Advances in Neural Information Processing Systems 23 NIPS*. (eds. Williams, C., Lafferty, J., Zemel, R., Shawe-Taylor, J. & Culotta, A.) (MIT Press, Cambridge, Massachusetts, 2010).

© 2011 Nature America, Inc. All rights reserved.

ONLINE METHODS

Psychophysical experiments. Five subjects participated in this study. Subjects S2 and S3 were A.R.G. and M.S.L., respectively. Subjects S1, S4 and S5 were naive to the aims of the experiment and were compensated for their time at $10 per h. All subjects had normal or corrected vision. Subjects stabilized their heads on a chin/forehead rest and viewed the stimuli binocularly at a distance of 57 cm in a dark room.

Each stimulus consisted of an array of Gabor patches positioned on a randomly jittered grid. Each Gabor was composed of a high-contrast 4 cycles per deg sinusoidal grating windowed by a circular Gaussian whose full width at half height was 0.41 deg. The Gabor patches were on a gray background and the monitor was gamma linearized. The entire stimulus array subtended 10 deg and contained roughly 37 Gabor patches. Two new random stimuli were generated for each trial and displayed simultaneously 2.5 deg to the left and right of the fixation point. Stimuli were visible for 0.75 s (0.5 s for subject S2). For each stimulus, Gabor orientations were either identical (low noise) or drawn from a Gaussian distribution (high noise) with an s.d. individually predetermined for each subject (see below). The three comparison conditions were randomly interleaved: low noise versus low noise, high noise versus high noise, and high noise versus low noise.

A central fixation point was presented between trials and we did not control for eye movements. Randomly interleaved staircases (two up/one down and two down/one up) adjusted the mean orientation of the comparison stimulus, whereas the mean orientation of the standard stimulus remained fixed. On each trial, the standard was randomly positioned on the left or right, with the comparison on the other side. Subjects completed at least 7,200 trials divided into 36 conditions (12 orientations of the standard for the three stimulus combinations) over at least six sessions. Subject S3 participated in an earlier version that was blocked by orientation. Before each session, subjects completed a few practice trials in the high noise versus high noise condition with auditory feedback. These trials were not included in analysis, and there was no feedback during the main experiment.

For each condition, we fit a cumulative Gaussian psychometric function to the data using a maximum-likelihood criterion, resulting in an estimate of the mean (relative bias) and s.d., σ_P (JND). Because we used a two-interval task[46], the s.d. of the underlying likelihood is $\sigma = \sigma_P / \sqrt{2}$.

In an initial experiment, we determined a level of orientation variability for the high noise condition such that each subject's high noise versus high noise JNDs were roughly threefold larger than their low noise versus low noise JNDs. Subjects first ran the low noise versus low noise condition with the standard oriented at 105 deg. We then fit a cumulative Gaussian and calculated its s.d., σ_L. Next, we fixed the two stimuli at 105 deg and 105 + $3\sigma_L$ deg and used two randomly interleaved two down/one up staircases to adjust the amount of orientation noise (applied to both stimuli). We fit a reflected cumulative Gaussian to these data (which ranged from 100% to 50% correct) and estimated the noise level that yielded 76% correct performance. This noise level was used for the high-noise condition in the main experiment (24.3, 21.8, 26.7, 19.2 and 20.0 deg for the five subjects, respectively).

We also ran a control experiment to ensure that the relative bias was not induced by the rectangular frame of the monitor. Subject S2 viewed monocularly through a 16.8-degree circular aperture, 20 cm from the eyes, with a black hood that obscured everything except the stimuli. The stimuli were presented centrally in two temporal intervals of 0.4 s each, with a 0.25-s mask after each, consisting of non-oriented 4 cycles per deg noise. The relative bias was identical to the main experiment, ruling out the concern of an artifact.

Estimation of the likelihood functions. Figure 2 shows the formulation of the likelihood functions. The measurement distribution $p(m|\theta_1)$ describes the probability of a sensory measurement given a particular stimulus value and is shown as a two-dimensional grayscale image. We assume that sensory noise is drawn from a von Mises distribution (described below) whose variance can be related to the same-noise variability data. The measurement distributions are vertical slices through this two-dimensional function, whereas the likelihoods are horizontal slices. Note that the likelihood function, $p(m_1|\theta)$, is typically not normalized as a probability distribution. Although the measurement distributions are assumed to be symmetric, their width depends on orientation, and the resulting likelihood functions are generally asymmetric (see **Fig. 2**)[47]. **Supplementary Figure 5** shows the corresponding likelihoods for the mean subject.

To estimate the measurement noise, we fit functions to the same-noise JND data (**Fig. 3b** and **Supplementary Fig. 1a**). We used rectified two-cycle sinusoids peaking at the obliques for the low noise versus low noise JNDs ($J_L(\theta) = \alpha_L|\sin(2\theta)| + \beta_L$) and high noise versus high noise JNDs ($J_H(\theta) = \alpha_H|\sin(2\theta)| + \beta_H$). We used least-squares minimization to estimate the parameters α_L, β_L, α_H and β_H for each subject, and computed confidence intervals for these estimates using the s.d. of estimates obtained on 1,000 bootstrap samples of the data. The prior will have a similar effect on the two stimuli in the same-noise conditions, leading to a negligible same-noise bias, as seen in our data. Our model of the measurement distribution is a von Mises function that peaks every 180 deg, $P(m|\theta) \propto e^{\kappa\cos(2(\theta-m))}$, where κ is a concentration parameter, whose value was chosen so as to produce JND values matching the fitted $j(\theta)$.

Converting sensory measurements to Bayesian estimates. Figure 2 shows how the Bayesian framework can be used to describe the decoding process. The estimator's shape depends on both the measurement noise and prior; the bias is approximately the product of the likelihood width and the slope of the log of the prior[13]. The Bayesian decision-making process that produces simulated psychometric data is depicted in **Figure 4**. The observer makes measurements $m_L(\theta_L)$ and $m_H(\theta_H)$, which are transformed by noise-appropriate estimators into values $\hat{\theta}_L$ and $\hat{\theta}_H$. Repeated measurements fluctuate due to sensory noise, resulting in a measurement distribution. This distribution is propagated through the estimator, resulting in a distribution of estimates. These distributions are then compared using signal detection theory[46], resulting in a single point on a psychometric function.

Estimation of the observers' priors. The human observers' priors (**Fig. 5** and **Supplementary Fig. 3**) were estimated by fitting the behavior of a Bayesian observer to the human data in the cross-noise condition. To avoid any preconceptions of what the prior ought to look like, we used a globally nonparametric model. Our only constraints were that it integrate to 1, that it be periodic with period 180 deg, and that the log prior be smooth and continuous. A previous study[13] used a globally nonparametric model, in which the log prior was approximated as linear over the central span of the likelihood. This assumption could be wrong in our case, because the likelihoods are quite broad in a few cases (for example, up to 30 deg s.d.). Instead, we used cubic splines to smoothly interpolate between values of $\log(P(\theta_i))$ at neighboring control points θ_i and θ_{i+1}

$$\log(P(\theta)) = a_i(\theta-\theta_i)^3 + b_i(\theta-\theta_i)^2 + c_i(\theta-\theta_i) + d_i$$

with the constraint that the neighboring splines' first and second derivatives matched at the common boundary θ_i. The parameters of the model were the values of $\log(P(\theta_i))$ at six control points θ_i = {30, 60, 90, 120, 150, 180 deg}. Changing the number of control points resulted in similar recovered priors.

Given a candidate prior $P(\theta)$ for each stimulus pair used in the experiment, we simulated 1,000 trials by drawing measurement samples from the appropriate measurement distributions. Each measurement leads to a likelihood function (the corresponding row of the precomputed matrix; **Fig. 2** and **Supplementary Fig. 5**), which is multiplied by the prior and maximized to obtain the estimate. The 1,000 pairs of estimates were then compared to obtain the binary model response (counter-clockwise or clockwise), and the average of this set of binary decisions yielded a single point on the Bayesian observer model–generated psychometric function (**Fig. 4**).

We estimated the best-fitting parameters of the log prior $\{\log(P(\theta_i))\}$ for each human subject by maximizing the probability of their observed data, given the above model-generated psychometric functions. We performed bootstrapping by randomly sampling, with replacement, the raw cross-noise response data 1,000 times (this is separate from the 1,000 measurement samples described above). If the subject performed n trials for a particular pair of orientations, then n responses were sampled from the corresponding data. The variation of the estimated priors across bootstrap replications is shown as the gray regions in **Figure 5**.

Estimating the environmental distribution. The distribution of orientations was computed from a publicly available image database (Olmos, A. & Kingdom, F.A.A., McGill Calibrated Color Image Database, http://tabby.vision.mcgill.ca/, 2004) containing 653 2,560 × 1,920 TIF photographs of natural scenes, with a pixel size of 0.0028 deg. We gamma-linearized the images using camera parameters provided

© 2011 Nature America, Inc. All rights reserved.

with the database, and converted to XYZ color space. Each image was normalized by its mean luminance and only luminance information (Y channel) was used. We decomposed each image into six spatial resolutions using a Gaussian pyramid[48]. At each scale, we used five-tap, rotation-invariant derivative filters[20] to compute gradients (x and y derivative pairs) centered on each pixel. We combined these into an orientation tensor[49], the local covariance matrix of the gradient vectors computed by averaging their outer products over square regions of 5×5 pixels. We calculated the tensor's eigenvector decomposition (that is, principal component analysis), and from that computed three quantities for each pixel: the energy (the sum of eigenvalues), the orientedness (the eigenvalue difference divided by the sum) and the dominant orientation (the angle of the leading eigenvector). Tensors centered on the two pixels closest to the borders were discarded. We created a histogram of dominant orientations for tensors that exceeded two thresholds; their orientedness was greater than 0.8 and their energy exceeded the 68th percentile of all the energy of the corresponding scale. We verified that our histograms were only weakly dependent on these two threshold values. Subthreshold tensors were typically located in patches of sky, or in internal regions of nontextured objects. The histogram for each scale was converted to a probability density by dividing by the total number of supra-threshold tensors, thus providing an estimate of the probability distribution over orientation. We selected the scale whose derivative filters best matched the peak spatial frequency sensitivity of humans[21] (2–5 cycles per deg) and the spatial frequency of the stimulus (4 cycles per deg). We also estimated the orientation distributions of scenes of primarily human-made content and blended scenes (for example, an image of a tree and a house). These distributions also exhibit a preponderance of cardinal orientations, but the heights of the peaks differ (**Supplementary Fig. 4**).

Simulated neural model. We created simulated populations of 60 orientation-tuned V1 neurons for each Gabor. We randomly sampled 8 of the 37 Gabors from the stimulus, consistent with human perceptual pooling[14]. The neural responses were characterized by independent Poisson noise, with expected number of spikes bounded between 1 and 12. The populations had either uniform or non-uniform tuning-curve widths and either uniform or non-uniform preferred orientations, resulting in four different models. We used von Mises tuning curves, which have been shown to provide good fits to empirical tuning curves[50]. For the models with uniform tuning-curve widths, the s.d. of the tuning curves was 17 deg, which is typical of V1 orientation–tuned neurons[34,41]. Non-uniform tuning widths were chosen according to a von Mises function (plus a constant) centered at the obliques with a concentration parameter of 42 deg. This produced s.d. consistent with neurophysiology in both average tuning width (17 deg) and ratio of oblique to cardinal tuning widths (approximately 3:2)[34,41]. Non-uniform preferred orientations were drawn from a von Mises distribution modified to peak at 0 and 90 deg with a s.d. of 35 deg, producing a ratio of the density of oblique to cardinal neurons consistent with neurophysiology (approximately 9:5)[33,34]. Simulations were run for the same conditions as shown to the human subjects: 12 orientations for each of three stimulus combinations. On every trial, we computed population vector estimates[42] for the orientations of both stimuli, and a decision was made as to whether the comparison was counter-clockwise or clockwise of the standard.

46. Green, D.M. & Swets, J.A. *Signal Detection Theory and Psychophysics* (Wiley, Oxford, UK, 1966).
47. Stocker, A.A. & Simoncelli, E.P. Visual motion aftereffects arise from a cascade of two isomorphic adaptation mechanisms. *J. Vis.* **9**, 1–14 (2009).
48. Burt, P. & Adelson, E. The Laplacian pyramid as a compact image code. *IEEE Trans. Commun.* **31**, 532–540 (1983).
49. Granlund, G.H. & Knutsson, H. *Signal Processing for Computer Vision* (Kluwer Academic Publishers, Norwell, Massachusetts, 1995).
50. Swindale, N.V. Orientation tuning curves: empirical description and estimation of parameters. *Biol. Cybern.* **78**, 45–56 (1998).

© 2011 Nature America, Inc. All rights reserved.

doi:10.1038/nn.2831