Cue combination on the circle and the sphere

Richard F. Murray
Department of Psychology and Centre for Vision Research, York University, Toronto, Ontario, Canada

Yaniv Morgenstern
Department of Psychology and Centre for Vision Research, York University, Toronto, Ontario, Canada

Bayesian cue combination models have been used to examine how human observers combine information from several cues to form estimates of linear quantities like depth. Here we develop an analogous theory for circular quantities like planar direction. The circular theory is broadly similar to the linear theory but differs in significant ways. First, in the circular theory the combined estimate is a nonlinear function of the individual cue estimates. Second, in the circular theory the mean of the combined estimate is affected not only by the means of individual cues and the weights assigned to individual cues but also by the variability of individual cues. Third, in the circular theory the combined estimate can be less certain than the individual estimates, if the individual estimates disagree with one another. Fourth, the circular theory does not have some of the closed-form expressions available in the linear theory, so data analysis requires numerical methods. We describe a vector sum model that gives a heuristic approximation to the circular theory's behavior. We also show how the theory can be extended to deal with spherical quantities like direction in three-dimensional space.

Keywords: computational modeling, space and scene perception, depth

Citation: Murray, R. F., & Morgenstern, Y. (2010). Cue combination on the circle and the sphere. *Journal of Vision, 10*(11):15, 1–11, http://www.journalofvision.org/content/10/11/15, doi:10.1167/10.11.15.

Introduction

Images often contain several cues to object properties like depth and shape, and many studies have investigated how the human visual system combines multiple cues to arrive at a single estimate of such properties (Landy, Maloney, Johnston, & Young, 1995; MacKenzie, Wilcox, & Murray, 2008; Yuille & Bülthoff, 1996). Previous studies have addressed properties that are linear, in the sense that they take a value on the real number line. Here we develop a theory of cue combination for circular quantities like planar direction and spherical quantities like direction in three-dimensional space. We have found this theory useful for examining how the light-from-above prior is combined with contextual cues to lighting direction (Morgenstern & Murray, 2009; Morgenstern, Murray, & Harris, in preparation), and it may also be useful for investigating other circular and spherical properties like hue and direction of motion. We begin by outlining the modified weak fusion theory of linear cue combination (Landy et al., 1995), and then we develop an analogous theory of cue combination on the circle. The circular theory is mostly a straightforward extension of the linear theory, but there are some important differences, which we discuss. We present a vector sum model that gives a heuristic description of the circular theory's behavior. Finally, we show how the theory can be extended to deal with cue combination on the sphere.

Cue combination on the line

We take the theory of modified weak fusion to be the canonical model of linear cue combination, as it is the most widely used and the most thoroughly tested (Landy et al., 1995; Maloney & Landy, 1989). In this section, we review the statistical basis of modified weak fusion, and in the next section, we show how the theory can be modified to handle circular quantities.

Suppose that an image contains multiple cues to the depth of a point. Modified weak fusion assumes that the observer has a normal random variable for each cue, say M for motion and D for disparity. On every trial, the observer has a sample m_i from M and d_i from D, and these samples are metric depth estimates. The observer knows the standard deviations of M and D, but not the means. Bayes' theorem gives the posterior probability distribution on possible values of depth z, conditioned on the samples m_i and d_i:

$$P(z|m_i, d_i) = \frac{P(m_i, d_i|z)P(z)}{P(m_i, d_i)}. \quad (1)$$

Here $P(z)$ is the prior probability of a depth value z, and $P(m_i, d_i)$ is the marginal probability of samples m_i and d_i,

doi: 10.1167/10.11.15 Received June 10, 2010; published September 16, 2010 ISSN 1534-7362

i.e., $P(m_i, d_i) = \int P(m_i, d_i \mid z)P(z)dz$. If M and D are conditionally independent, Equation 1 becomes

$$= \frac{P(m_i|z)P(d_i|z)P(z)}{P(m_i)P(d_i)}. \quad (2)$$

If the motion and disparity cues are normal random variables with means equal to the true depth value (i.e., the cues are unbiased), then we can use the normal probability density function $g(x; \mu, \sigma)$ and the standard deviations $\sigma_M = std[M]$ and $\sigma_D = std[D]$ to write Equation 2 as

$$= \frac{g(m_i; z, \sigma_M)g(d_i; z, \sigma_D)P(z)}{P(m_i)P(d_i)}, \quad (3)$$

$$= \frac{g(z; m_i, \sigma_M)g(z; d_i, \sigma_D)P(z)}{P(m_i)P(d_i)}. \quad (4)$$

Using Equation A1 (see Appendix A) for the pointwise product of two normal density functions, this becomes

$$= \frac{g(m_i - d_i; 0, (\sigma_M^2 + \sigma_D^2)^{1/2})g(z; s_i, \sigma_S)P(z)}{P(m_i)P(d_i)}, \quad (5)$$

where

$$s_i = \frac{(1/\sigma_M^2)m_i + (1/\sigma_D^2)d_i}{(1/\sigma_M^2) + (1/\sigma_D^2)}, \quad (6)$$

and

$$\frac{1}{\sigma_S^2} = \frac{1}{\sigma_M^2} + \frac{1}{\sigma_D^2}. \quad (7)$$

The maximum-likelihood depth estimate is s_i in Equation 6, as this is the value of z that maximizes the likelihood term $g(z; s_i, \sigma_S)$ in Equation 5. Thus the maximum-likelihood decision variable is a random variable S that is a weighted average of cues M and D, with weights determined by the variance of the cues. Modified weak fusion asserts that the observer uses this maximum-likelihood estimate as a decision variable when making depth judgements. In a weaker form, the theory says that the observer's decision variable is a weighted average of M and D, but not necessarily the optimal weighted average as in Equation 6. When deriving the normative cue combination rule in Equation 6, we assumed that the cues M and D are unbiased, but the theory leaves open the possibility that in human vision the cues are in fact biased. Furthermore, if the prior $P(z)$ is not uniform over the range of interest, the theory allows that observer may use the maximum a posteriori estimate, which is the value of z that maximizes $g(z; s_i, \sigma_S)P(z)$ in Equation 5 and can differ from the maximum-likelihood estimate. Finally, when depth estimates from different cues disagree strongly, one or more of the estimates may be corrupt, so the theory includes a robustness mechanism that reduces the weight assigned to cues that have very different values than more reliable cues.

Cue combination on the circle

Sometimes circular quantities can be treated adequately with linear models, but in general they require methods of their own (Fisher, 1993). For example, the arithmetic mean of angles is not a good measure of central tendency (consider the mean of 1° and 359°), and linear probability densities are not intrinsically periodic like circular densities. Here we develop a Bayesian approach to circular cue combination by following the outline of cue combination in the previous section, substituting circular probability distributions for linear distributions.

Consider a scene where lighting directions are restricted to the frontoparallel plane, like the hands of a clock. Suppose that images of this scene have two cues to lighting direction, such as shading and cast shadows. As in the linear model, our goal is to find the most probable value of the quantity of interest, here the lighting direction, after having observed the available cues. We model each cue as an independent von Mises random variable (Fisher, 1993), say H for shading and C for cast shadows. The von Mises distribution (Figure 1) is a circular analogue of the normal distribution, with a location parameter μ indicating the mean and the angle of maximum probability, and a concentration parameter κ indicating the narrowness of the peak at μ. The concentration parameter κ is the circular analogue of the "reliability" σ^{-2} of a normal distribution, and when a von Mises distribution is narrow ($\kappa > 2$) it is well approximated by a normal distribution with variance $\sigma^2 = 1/\kappa$. The von Mises probability density function is

$$f_{VM}(\theta; \mu, \kappa) = \frac{\exp(\kappa \cos(\theta - \mu))}{2\pi I_0(\kappa)}. \quad (8)$$

Here I_0 is the modified Bessel function of the first kind and order zero. Following the linear theory, we assume that the observer knows the concentration parameters κ_H and κ_C of cues H and C, but not the means μ_H and μ_C.

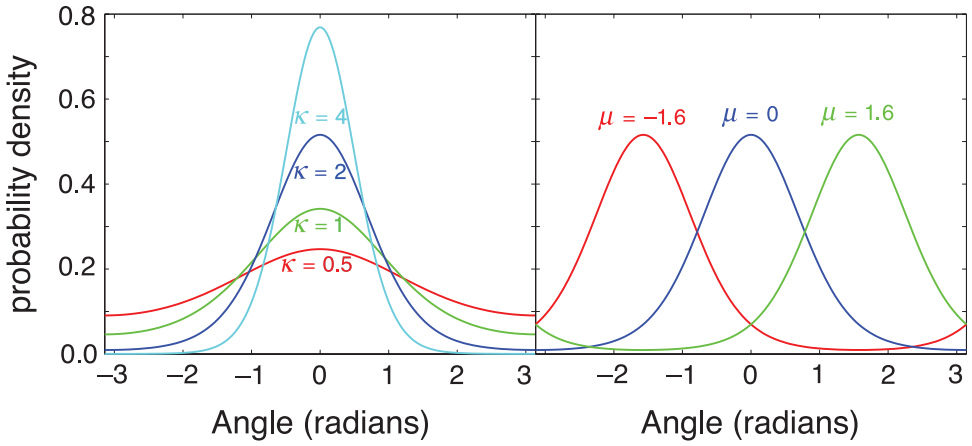

Figure 1. Examples of the von Mises probability density function. In the left panel, $\mu = 0$, and in the right panel, $\kappa = 2$.

On each trial, the observer has a sample h_i from H and c_i from C. Bayes' theorem gives the posterior distribution on lighting directions θ:

$$P(\theta|h_i, c_i) = \frac{P(h_i, c_i|\theta)P(\theta)}{P(h_i, c_i)}, \tag{9}$$

$$= \frac{P(h_i|\theta)P(c_i|\theta)P(\theta)}{P(h_i)P(c_i)}, \tag{10}$$

$$= \frac{f_{VM}(h_i; \theta, \kappa_H) f_{VM}(c_i; \theta, \kappa_C) P(\theta)}{P(h_i)P(c_i)}, \tag{11}$$

$$= \frac{f_{VM}(\theta; h_i, \kappa_H) f_{VM}(\theta; c_i, \kappa_C) P(\theta)}{P(h_i)P(c_i)}. \tag{12}$$

With Equation A4 (see Appendix A) for the pointwise product of two von Mises densities, this becomes

$$= \frac{f_{VM}(\theta; l_i, \kappa_L) P(\theta) I_0(\kappa_L)}{2\pi P(h_i)P(c_i)I_0(\kappa_H)I_0(\kappa_C)}, \tag{13}$$

where

$$l_i = h_i + \arctan(\sin(c_i - h_i), (\kappa_H/\kappa_C) + \cos(c_i - h_i)), \tag{14}$$

and

$$\kappa_L = \sqrt{\kappa_H^2 + \kappa_C^2 + 2\kappa_H\kappa_C\cos(c_i - h_i)}. \tag{15}$$

Here $\arctan(y, x)$ is the four-quadrant inverse tangent. The location parameter l_i is the maximum-likelihood estimate of lighting direction, and the concentration κ_L represents the certainty of this estimate. Thus the maximum-likelihood decision variable is a random variable L that is a nonlinear function of cues H and C, as in Equation 14. Following the linear model, we can extend the combination rule by allowing the concentration parameters κ_H and κ_C in Equation 14 (i.e., the analogues of $1/\sigma_M^2$ and $1/\sigma_D^2$ in Equation 6) to be replaced by arbitrary positive weights, thus allowing a broader and generally nonoptimal range of cue combination strategies, and by allowing for an effect of the prior $P(\theta)$. In most of what follows, we will allow the decision variable L to be calculated with concentrations replaced by arbitrary positive weights w_H and w_C:

$$L = H + \arctan(\sin(C - H), (w_H/w_C) + \cos(C - H)). \tag{16}$$

We will mostly discuss the behavior predicted by this decision variable, as the prior does not introduce complications specific to the circular model.

Here we find the first important difference from the linear model: the combined decision variable L does not belong to the same family of probability distributions (i.e., von Mises) as the individual cues H and C. In fact, it does not belong to any standard family we know of, and we have no closed-form expression for its distribution. Numerical methods are necessary to find the distribution of the combined decision variable L from the means, concentrations, and weight parameters of individual cues.

There is, however, a useful closed-form approximation. In Appendix A, we show that the decision variable L can be approximated as a von Mises random variable L^* with parameters

$$\mu_{L^*} = \mu_H + \arctan(\sin(\mu_C - \mu_H), (w_H/w_C) + \cos(\mu_C - \mu_H)), \tag{17}$$

$$\kappa_{L^*} = \frac{w_H + w_C}{w_H^2/\kappa_H + w_C^2/\kappa_C} \cdot \sqrt{w_H^2 + w_C^2 + 2w_H w_C \cos(\mu_C - \mu_H)}. \quad (18)$$

When the weights are optimal ($w_H = \kappa_H$, $w_C = \kappa_C$), this approximation simplifies to

$$\mu_{L^*} = \mu_H + \arctan(\sin(\mu_C - \mu_H), (\kappa_H/\kappa_C) + \cos(\mu_C - \mu_H)), \quad (19)$$

$$\kappa_{L^*} = \sqrt{\kappa_H^2 + \kappa_C^2 + 2\kappa_H \kappa_C \cos(\mu_C - \mu_H)}. \quad (20)$$

Figure 2 shows the probability densities of the decision variable L (calculated using Monte Carlo methods) and the approximation L^*, for several means, concentrations, and weights for the individual cues H and C. Figure 3 shows the circular means and circular standard deviations (Fisher, 1993) of L and L^* for a wider range of parameters. The approximation is reasonably good, at least as an aid to understanding the model's qualitative behavior. It is least accurate when approximately equally weighted cues indicate opposite directions; in this case, L^* has a uniform distribution, whereas L is bimodal with peaks halfway between the two opposed directions.

Key properties of the circular model

This circular cue combination model has some similarities to and some important differences from the linear model. As in the linear model, the combined estimate L has a value partway between the cues H and C. Unlike in the linear model, the combined estimate is not simply a weighted average of the cues. The left-hand column of Figure 3 shows that the mean of the combined estimate is approximately a weighted average of the cue means when the cue means are similar, but when the discrepancy between cues is large, the combined estimate is closer to the more heavily weighted cue. For example, when the shading cue is weighted more strongly than the cast shadow cue ($w_H > w_C$), the cast shadow cue pulls the combined estimate away from the shading cue over a limited range, but when the discrepancy is large the combined estimate smoothly reverts to the shading cue. The robustness mechanism in the linear model causes similar behavior. Robust behavior without a dedicated robustness mechanism for rejecting discrepant cues is not unique to circular models: non-Gaussian models of cue combination on the line, particularly those based on probability distributions with heavy tails, can also behave robustly without dedicated robustness mechanisms (Girshick & Banks, 2009; Knill, 2007).

In the linear model, the combined estimate is always more certain than either of the cues, and the certainty of the combined estimate does not depend on the discrepancy between the cues (see Equation 7). Neither of these properties holds in the circular model, where the width of the combined distribution increases with the discrepancy between the cues (Figure 3, right-hand column). Equation 18 shows that in the von Mises approximation, the combined concentration decreases as a function of $|\mu_C - \mu_H|$. Equation 20 shows that in the von Mises approximation with optimal weights, when the cues agree ($\mu_H = \mu_C$) the combined concentration κ_L equals the sum of the cue concentrations, $\kappa_H + \kappa_C$ (just as reliabilities sum in the linear model, i.e., $\sigma_S^{-2} = \sigma_M^{-2} + \sigma_D^{-2}$; see Equation A10), and when cues indicate opposite directions κ_L equals the difference between the cue concentrations, $|\kappa_H - \kappa_C|$.

Using the von Mises approximation

The linear model has the convenient property that the mean of the combined decision variable depends only on the cue means and weights, not on the cue standard deviations (except in the optimal version of the linear theory, where the weights are determined by the cue standard deviations). For instance, in a cue conflict experiment on perceived depth, if a motion cue has twice the weight of a disparity cue, $w_M = 2w_D$, then motion perturbations will have twice the effect on perceived depth as disparity perturbations, regardless of the standard deviations of the motion and disparity cues (Young, Landy, & Maloney, 1993). Equation 17 shows that in the von Mises approximation to the circular model, the decision variable mean is independent of the cue concentrations, as in the linear model. In the exact circular model, though, the cue concentrations do affect the mean of the decision variable. Figure 4 shows the mean of the decision variable L for several cue means, concentrations, and weights, along with the von Mises approximation. First, we see that relative to the von Mises approximation, the decision variable mean is biased toward the mean of the cue with the higher concentration (Figure 4a). Second, again relative to the von Mises approximation, the decision variable mean is biased toward the mean of the cue with the greater weight (Figure 4b). (These two biases may be in opposite directions.) Third, these two biases are strongest when the difference between the cue means is large (Figures 4c and 4d) and when the concentrations are low (Figure 4, all panels).

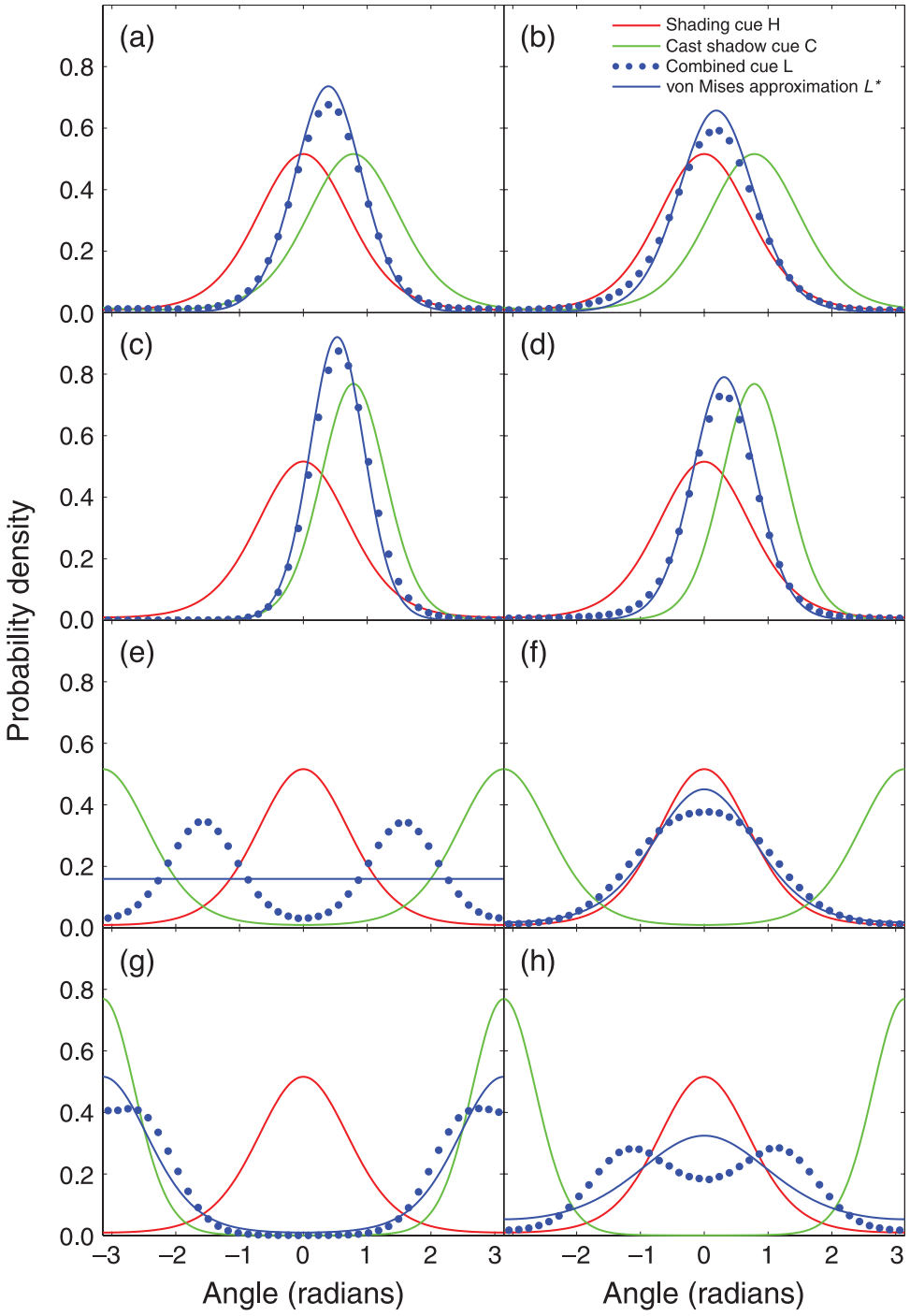

Figure 2. Examples of probability density functions for individual and combined cues, and the von Mises approximation. The shading cue (red line) has concentration $\kappa_H = 2$ in all panels. The cast shadow cue (green line) has concentration $\kappa_C = 2$ in the first and third rows, and $\kappa_C = 4$ in the second and fourth rows. In the left-hand column, the shading and cast shadow cues are combined optimally (weights equal concentrations), and in the right-hand column, the shading cue is given three times the optimal weight.

These biases are important to understand when using the von Mises approximation. Consider a cue conflict experiment where a fit of the von Mises approximation suggests that shading cues have twice the weight of cast shadow cues. Figure 4a shows that the decision variable mean is biased toward the cue with the higher concentration, and the von Mises approximation does not take this into account, so if the shading cue is more concentrated then its apparent weight will be inflated. Similarly, Figure 4b shows that at low concentrations the decision variable

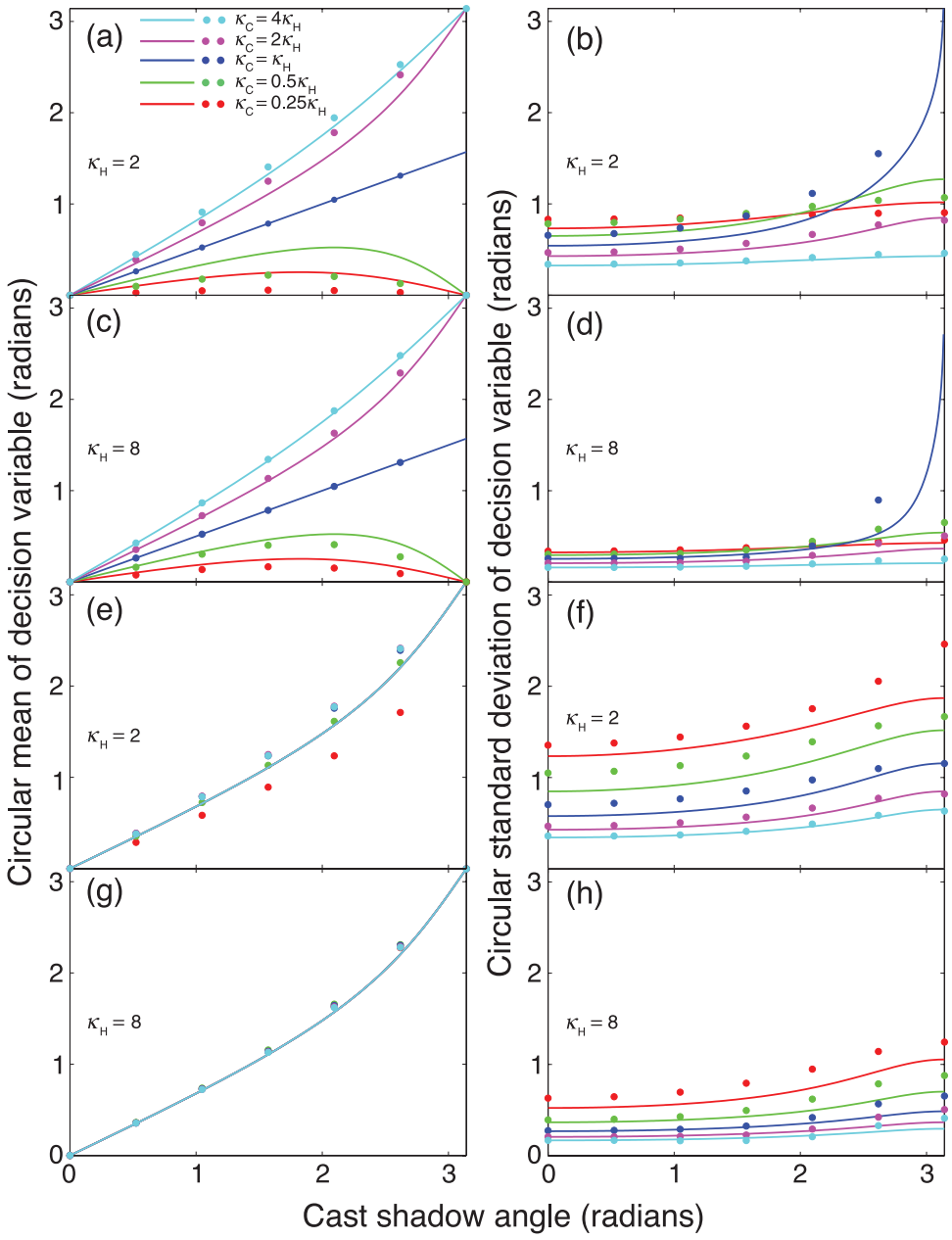

Figure 3. Circular means and circular standard deviations of the combined lighting direction estimate L and the von Mises approximation L^*. The data points are means and standard deviations of the decision variable L, and the solid lines are von Mises approximations. The shading cue mean is 0 radians, and the cast shadow cue mean varies along the x-axis. In each panel, the shading cue concentration κ_H has the value indicated, and each color corresponds to a cast shadow cue concentration κ_C at a multiple of κ_H, as indicated in the legend. In the top four panels, the cues are combined optimally. In the bottom four panels, the shading cue weight is $w_H = 1$ and the cast shadow cue weight is $w_C = 2$. In the bottom two left-hand panels, all five von Mises approximation curves are superimposed, and most of the mean and standard deviation points from different concentrations are superimposed as well, indicating that the decision variable mean depends mostly on the cue weights and means, with only a small effect of cue concentrations.

mean is biased toward the cue with the greater weight, which will make a heavily weighted shading cue appear to have an even greater weight. Fortunately, Figure 4 shows that these biases are weak as long as the cue concentrations are reasonably large (e.g., $\kappa > 2$). Cue concentrations can be estimated from the slopes of psychometric functions, e.g., the slope of a psychometric function for discriminating between lighting directions defined by shading gives an upper limit on the noisiness of the shading cue, since discrimination performance is limited

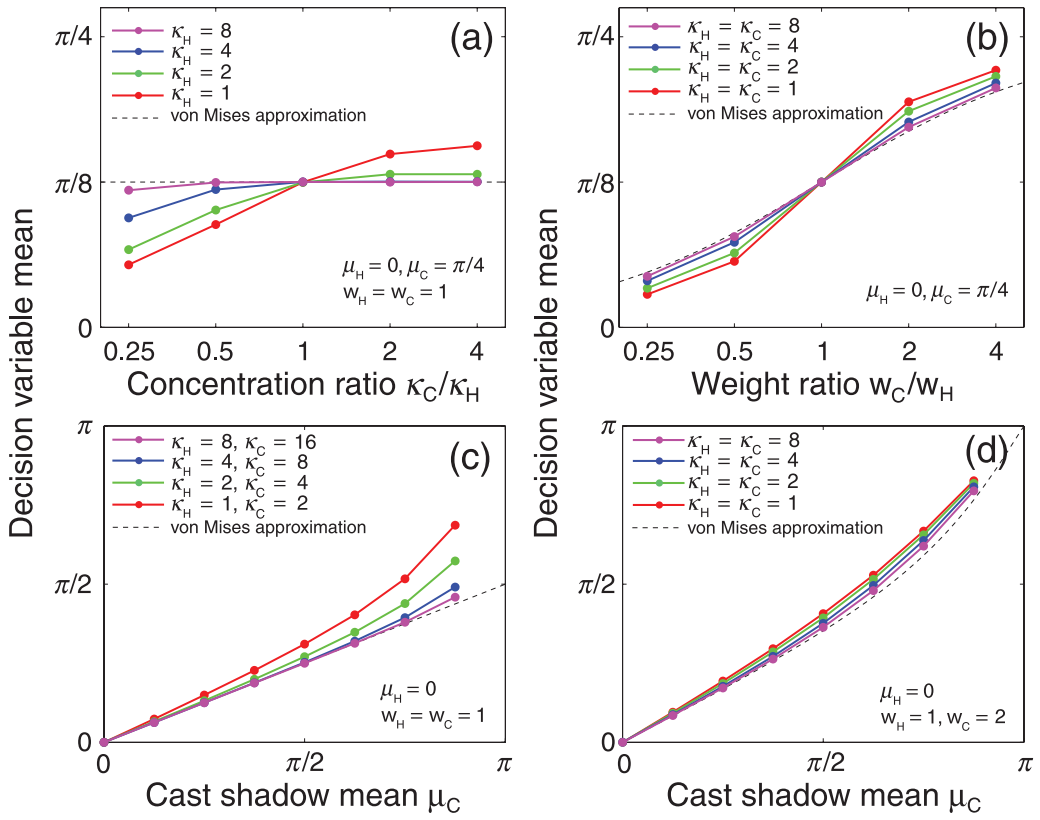

Figure 4. Biases in the decision variable mean at low cue concentrations. (a) The mean is biased toward the more concentrated cue. (b) The mean is biased toward the more heavily weighted cue. (c) The bias toward the more concentrated cue increases with the difference between the cue means. (d) The bias toward the more heavily weighted cue increases with the difference between the cue means. In all cases, the biases are small when the cue concentrations are large (e.g., $\kappa > 2$).

by noise in the shading cue plus any other internal decision noise. Independent estimates like this can show whether cues in a given experiment are concentrated enough to make the biases illustrated in Figure 4 negligible.

A vector sum heuristic

Equations 19 and 20 describe vector addition in polar coordinates, where μ is the vector's direction and κ is its length. That is, the sum of two polar-coordinate vectors (κ_H, μ_H) and (κ_C, μ_C) is the polar-coordinate vector (κ_{L*}, μ_{L*}). This means we can visualize the von Mises approximation to optimally weighted circular cue combination as a vector sum, with a vector for each cue pointing in the direction μ of the cue, and the length of the vector equal to the cue's concentration parameter κ. The vector representing the von Mises approximation to the combined distribution is the sum of the vectors representing the individual cues. This heuristic is only as good as the von Mises approximation, so it fails when we combine equally strong cues indicating opposite directions. In this case, the cue vectors sum to the zero vector, i.e., $\kappa_{L*} = 0$, representing the uniform distribution, whereas the combined decision variable L is bimodal, as mentioned earlier. The vector sum heuristic also fails when observers use nonoptimal cue weights, as the von Mises approximation for nonoptimal weights does not describe vector addition in polar coordinates; specifically, Equation 18 does not give the correct length of the summed vector.

Interestingly, Mittelstaedt (1983, 1986) arrived empirically at a vector sum model of the effects of visual, body orientation, and gravitational cues on the direction of subjective vertical, and Dyde, Jenkin, and Harris (2006) and Jenkin, Dyde, Jenkin, Howard, and Harris (2003; Jenkin et al., 2005; Jenkin, Jenkin, Dyde, & Harris, 2004) developed a similar model of the direction of perceptual upright. In these models, each cue is represented by a vector indicating the cue's direction, with more certain cues represented by longer vectors, and the direction of subjective vertical or perceptual upright is found by summing the cue vectors. In the experiments that test these models, cues vary around a circle of possible directions. The vector sum models and the experimental findings that support them reproduce key properties of

our Bayesian circular model. They show robustness-like effects, where a weak cue's effect declines as it becomes discrepant with stronger cues (e.g., compare Mittelstaedt's (1983) Figure 1 with our Figure 3, left-hand column). They also show that combined directional estimates are more certain when individual cues are consistent than when they indicate different directions (e.g., Jenkin et al., 2004, Table 1). Furthermore, Dyde et al. show that cue weights are sometimes assigned optimally, i.e., the length of each cue vector is inversely proportional to the variance of the corresponding cue. Our Bayesian model suggests that these vector sum models are not arbitrary constructions, but rather they succeed because observers make maximum-likelihood directional estimates based on noisy directional cues, using computations that take account of the circular nature of the cues.

The fact that taking the pointwise product of two von Mises densities is equivalent to taking a vector sum also suggests an alternate representation of optimal circular cue combination. Equation 14 shows that the *angle* of a combined direction estimate is a nonlinear function of the *angles* of the individual cues. However, optimal circular cue combination can be represented as a simple weighted sum, if we represent individual cues as *vectors* instead of angles. As in Equation 14, on each trial we have a shading cue angle h_i and its associated concentration κ_H, and a cast shadow cue angle c_i and its concentration κ_C. We can find the maximum-likelihood combined angle estimate using Equation 14. Alternatively and equivalently, we can represent the two cues as polar-coordinate vectors (κ_H, h_i) and (κ_C, c_i), and then the combined estimate is represented by the polar vector (κ_L, l_i) that is the vector sum of the cue vectors. Of course, we must still use Equations 14 and 15, which are equations for vector summation in polar coordinates, to find the summed vector (κ_L, l_i). Nevertheless, this alternate view highlights a similarity to optimal linear cue combination, where the maximum-likelihood estimate is also a weighted sum of individual cues.

Scope of the circular model

Circular data can sometimes be handled adequately with linear models (e.g., Hillis, Ernst, Banks, & Landy, 2002; Knill, 1998), so when are circular models necessary? When circular data occupy only a small segment of a circle, that segment can be approximated as a line, and the fact that the data space is periodic may be unimportant. When data occupy most of the 360° range of the circle, circular models are necessary. In the context of cue combination, this means that a circular model is appropriate when the probability distributions that characterize individual cues have very different means, or low concentrations. Cue combination experiments on slant estimation have typically examined how observers resolve small discrepancies between individual slant cues and have used stimuli that allow slant to be estimated precisely (e.g., Hillis et al., 2002; Knill, 1998). In such cases, linear models are adequate. In contrast, experiments on the light-from-above prior have used lighting cues that cover a broad range of directions, sometimes with large discrepancies between the prior and the cues (Adams, 2007; Morgenstern & Murray, 2009), and furthermore, it is not known whether the light-from-above prior is best characterized by a high- or low-concentration probability distribution. Similarly, experiments on subjective vertical and perceptual upright (Jenkin et al., 2003, 2005, 2004; Mittelstaedt, 1983, 1986) have examined conflicts between cues that indicate very different directions. Under these conditions, a circular model is appropriate.

Other circular distributions

Many properties of this Bayesian circular cue combination model are not immediate consequences of working on the circle instead of the line but depend on our choice of the von Mises distribution instead of some other circular distribution. For instance, Figure 2e shows that when two equally weighted cues indicate opposite directions, the model's decision variable is bimodal, with modes halfway between the two opposed directions. We have found that a circular model based on the wrapped Cauchy distribution (Fisher, 1993) responds very differently, producing a bimodal decision variable with modes in the two opposed cue directions. The von Mises distribution is a reasonable starting point for circular models, as it has some of the properties that make the normal distribution fundamental on the line, e.g., the von Mises is the circular distribution that has the greatest entropy (and so in a sense is the most random) for a given mean and variance (Mardia, 1972). Nevertheless, circular cue combination experiments may find that other distributions describe behavioral data more accurately. For example, as just discussed, the von Mises and Cauchy models predict very different trial-by-trial response distributions for equally weighted, directly opposed cues, and these predictions can be tested by examining the relevant response distributions from human observers.

Cue combination on the sphere

This model has an obvious generalization to higher dimensions. The von Mises distribution on the circle is a special case of the von Mises–Fisher distribution on the

n-dimensional sphere (Fisher, Lewis, & Embleton, 1987), which in three dimensions has probability density function

$$f_F(\mathbf{x};\boldsymbol{\mu},\kappa) = \frac{\kappa}{4\,\pi \sinh \kappa} \exp(\kappa \mathbf{x} \bullet \boldsymbol{\mu}). \quad (21)$$

Here $\mathbf{x}$ is a unit vector, $\boldsymbol{\mu}$ is a unit-vector location parameter indicating the peak direction, κ is a scalar concentration parameter, and $\bullet$ is the vector dot product. Suppose the shading cue $\mathbf{H}$ and the cast shadow cue $\mathbf{C}$ are now unit vectors ranging over all three-dimensional directions (and so we write them in bold). Repeating the derivation in Equations 9 to 15 with a von Mises–Fisher distribution shows that the maximum-likelihood combined lighting direction estimate is

$$\mathbf{L} = \frac{\kappa_H \mathbf{H} + \kappa_C \mathbf{C}}{\|\kappa_H \mathbf{H} + \kappa_C \mathbf{C}\|}. \quad (22)$$

Here $\| \|$ is Euclidean vector length. As in the linear and circular models, we can generalize this combination rule by replacing the concentrations with arbitrary positive weights.

This spherical model has many properties of the circular model. We state a few without proof. The combined decision variable is not a von Mises–Fisher random variable but can be approximated as one, and the parameters of the approximation can be found using the approach followed in Appendix A for the circular model, with spherical distributions replacing circular distributions. The spherical model shows robustness-like effects, where the less reliable of two highly discrepant cues tends to be ignored. The certainty of the combined spherical estimate decreases as the discrepancy between the individual cues grows. Finally, the vector sum heuristic approximates optimal cue combination on the sphere.

Conclusion

Linear cue combination models have extended our understanding of vision from conditions where only one task-relevant cue is available to more complex conditions where several cues provide useful information. Here we have shown that these models can be adapted to circular and spherical domains, but that some care is necessary as several familiar properties of linear models do not extend to circular and spherical models. Results from some common tools, such as cue conflict studies, must be interpreted differently when working with circular and spherical cues. The circular and spherical models we have described should provide a useful framework for understanding how multiple information sources contribute to observers' estimates of properties like direction in a plane or in three-dimensional space, as well as other properties like hue that are organized periodically.

Appendix A

Combination formulas

In the main text and in this appendix, we use the following expressions for the pointwise product of normal and von Mises densities:

$$g(x;\mu_1,\sigma_1)g(x;\mu_2\sigma_2) = g(\mu_1 - \mu_2; 0, (\sigma_1^2 + \sigma_2^2)^{1/2}) \cdot g(x;\mu_3,\sigma_3), \quad (A1)$$

where

$$\mu_3 = \frac{(1/\sigma_1^2)\mu_1 + (1/\sigma_2^2)\mu_2}{(1/\sigma_1^2) + (1/\sigma_2^2)}, \quad (A2)$$

and

$$\frac{1}{\sigma_3^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}, \quad (A3)$$

$$f_{VM}(\theta;\mu_1,\kappa_1)f_{VM}(\theta;\mu_2,\kappa_2) = \frac{I_0(\kappa_3)}{2\pi I_0(\kappa_1)I_0(\kappa_2)} \cdot f_{VM}(\theta;\mu_3,\kappa_3), \quad (A4)$$

where

$$\mu_3 = \mu_1 + \arctan(\sin(\mu_2 - \mu_1), (\kappa_1/\kappa_2) + \cos(\mu_2 - \mu_1)), \quad (A5)$$

and

$$\kappa_3 = \sqrt{\kappa_1^2 + \kappa_2^2 + 2\kappa_1\kappa_2\cos(\mu_2 - \mu_1)}. \quad (A6)$$

Optimal weights

In the linear model described in Equations 1 to 7, the motion and disparity cues have densities $g(x;\, \mu_M,\, \sigma_M)$ and

$g(x;\ \mu_D,\ \sigma_D)$. Equation A1 shows that the pointwise product of these densities is proportional to $g(x;\ \mu_T,\ \sigma_T)$, where

$$\mu_T = \frac{(1/\sigma_M^2)\mu_M + (1/\sigma_D^2)\mu_D}{(1/\sigma_M^2) + (1/\sigma_D^2)}, \tag{A7}$$

$$\sigma_T^2 = \frac{1}{(1/\sigma_M^2) + (1/\sigma_D^2)}. \tag{A8}$$

We can find the mean and variance of the linear maximum-likelihood decision variable S from Equation 6:

$$E[S] = \frac{(1/\sigma_M^2)\mu_M + (1/\sigma_D^2)\mu_D}{(1/\sigma_M^2) + (1/\sigma_D^2)}, \tag{A9}$$

$$VAR[S] = \frac{1}{(1/\sigma_M^2) + (1/\sigma_D^2)}. \tag{A10}$$

Thus, in the linear model, the pointwise product of the densities of the individual cues (Equations A7 and A8) is equal to the density of the maximum-likelihood decision variable S (Equations A9 and A10), up to a normalization constant.

As discussed in the main text, the circular maximum-likelihood decision variable S is not von Mises. The von Mises approximation to the optimally weighted decision variable (Equations 19 and 20) is the pointwise product of the von Mises densities of the individual cues, as can be seen by comparing the approximation to Equations 12 to 15. As we have just shown, this approximation is exact in the linear model, so it should work well in the circular model under conditions where the circular model approximates the linear model: large cue concentrations and small differences between cue means. Figure 3 supports this approximation and shows that even at low concentrations and large differences between the means, the approximation still captures the models' behavior qualitatively, e.g., the width of the combined decision variable grows with the difference between the cue means.

Arbitrary weights

In the linear model, when we replace the optimal weights in Equation 6 with arbitrary positive weights, the decision variable is

$$S = \frac{w_M m_i + w_D d_i}{w_M + w_D}, \tag{A11}$$

which has mean and variance

$$E[S] = \frac{w_M\mu_M + w_D\mu_D}{w_M + w_D}, \tag{A12}$$

$$VAR[S] = \frac{w_M^2\sigma_M^2 + w_D^2\sigma_D^2}{(w_M + w_D)^2}. \tag{A13}$$

The decision variable would have the same distribution if, instead of suboptimally combining M and D, the observer optimally combined two random variables X and Y with means $\mu_X = \mu_M$ and $\mu_Y = \mu_D$, and variances

$$\sigma_X^2 = \frac{w_M^2\sigma_M^2 + w_D^2\sigma_D^2}{w_M(w_M + w_D)}, \tag{A14}$$

$$\sigma_Y^2 = \frac{w_M^2\sigma_M^2 + w_D^2\sigma_D^2}{w_D(w_M + w_D)}. \tag{A15}$$

That is, when an observer suboptimally combines cues M and D using weights w_M and w_D, the decision variable's density is the pointwise product of $g(x;\ \mu_X,\ \sigma_x)$ and $g(x;\ \mu_Y,\ \sigma_Y)$, up to a normalization factor.

We will use the preceding fact about the linear model to approximate the decision variable in the circular model when the observer uses arbitrary weights. When the observer combines cues H and C suboptimally as in Equation 16, we approximate the decision variable by taking the pointwise product of von Mises distributions U and V that have means $\mu_U = \mu_H$ and $\mu_V = \mu_C$, and concentrations chosen by analogy with Equations A14 and A15 and using the approximation $\kappa = \sigma^{-2}$, which is valid at large concentrations (Fisher, 1993):

$$\kappa_U = \frac{w_H(w_H + w_C)}{w_H^2/\kappa_H + w_C^2/\kappa_C}, \tag{A16}$$

$$\kappa_V = \frac{w_C(w_H + w_C)}{w_H^2/\kappa_H + w_C^2/\kappa_C}. \tag{A17}$$

Equation A4 shows that the pointwise product of these densities is proportional to a von Mises density function with parameters

$$\mu_W = \mu_H + \arctan(\sin(\mu_C - \mu_H), (w_H/w_C) + \cos(\mu_C - \mu_H)), \tag{A18}$$

$$\kappa_W = \frac{w_H + w_C}{w_H^2/\kappa_H + w_C^2/\kappa_C} \cdot \sqrt{w_H^2 + w_C^2 + 2w_H w_C\cos(\mu_C - \mu_H)}. \tag{A19}$$

This is the approximation in Equations 17 and 18.

Acknowledgments

We thank Minjung Kim and two anonymous reviewers for comments on the manuscript. We also thank Laurence Harris for comments, particularly the suggestion that we investigate the relationship between the Bayesian circular model and the vector sum model.

Commercial relationships: none.
Corresponding author: Richard F. Murray.
Email: rfm@yorku.ca.
Address: Centre for Vision Research, York University, 4700 Keele Street, CSEB 0009, Toronto, Ontario M3J 1P3, Canada.

References

Adams, W. J. (2007). A common light-prior for visual search, shape, and reflectance judgments. *Journal of Vision, 7*(11):11, 1–7, http://www.journalofvision.org/content/7/11/11, doi:10.1167/7.11.11. [PubMed] [Article]

Dyde, R. T., Jenkin, M. R., & Harris, L. R. (2006). The subjective visual vertical and the perceptual upright. *Experimental Brain Research, 173,* 612–622.

Fisher, N. I. (1993). *Statistical analysis of circular data*. New York: Cambridge University Press.

Fisher, N. I., Lewis, T., & Embleton, B. J. J. (1987). *Statistical analysis of spherical data*. New York: Cambridge University Press.

Girshick, A. R., & Banks, M. S. (2009). Probabilistic combination of slant information: Weighted averaging and robustness as optimal percepts. *Journal of Vision, 9*(9):8, 1–20, http://www.journalofvision.org/content/9/9/8, doi:10.1167/9.9.8. [PubMed] [Article]

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: Mandatory fusion within, but not between, senses. *Science, 298,* 1627–1630.

Jenkin, H. L., Dyde, R. T., Jenkin, M. R., Howard, I. P., & Harris, L. R. (2003). Relative role of visual and non-visual cues in determining the direction of "up": Experiments in the York tilted room facility. *Journal of Vestibular Research, 13,* 287–293.

Jenkin, H. L., Dyde, R. T., Zacher, J. E., Zikovitz, D. C., Jenkin, M. R., Allison, R. S., et al. (2005). The relative role of visual and non-visual cues in determining the perceived direction of "up": Experiments in parabolic flight. *Acta Astronautica, 56,* 1024–1032.

Jenkin, H. L., Jenkin, M. R., Dyde, R. T., & Harris, L. R. (2004). Shape-from-shading depends on visual, gravitational, and body-orientation cues. *Perception, 33,* 1453–1461.

Knill, D. C. (1998). Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture. *Vision Research, 38,* 2635–2656.

Knill, D. C. (2007). Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *Journal of Vision, 7*(7):5, 1–24, http://www.journalofvision.org/content/7/7/5, doi:10.1167/7.7.5. [PubMed] [Article]

Landy, M. S., Maloney, L. T., Johnston, E. J., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35,* 389–412.

MacKenzie, K. J., Murray, R. F., & Wilcox, L. M. (2008). The intrinsic constraint approach to cue combination: An empirical and theoretical evaluation. *Journal of Vision, 8*(8):5, 1–10, http://www.journalofvision.org/content/8/8/5, doi:10.1167/8.8.5. [PubMed] [Article]

Maloney, L. T., & Landy, M. S. (1989). A statistical framework for robust fusion of depth information. In W. A. Pearlman (Ed.), *Visual communications and image processing IV. Proceedings of the SPIE* (vol. 1199, pp. 1154–1163). Bellingham, WA: SPIE.

Mardia, K. V. (1972). *Statistics of directional data*. London: Academic Press.

Mittelstaedt, H. (1983). A new solution to the problem of the subjective vertical. *Naturwissenschaften, 70,* 272–281.

Mittelstaedt, H. (1986). The subjective vertical as a function of visual and extraretinal cues. *Acta Psychologica, 63,* 63–85.

Morgenstern, Y., & Murray, R. F. (2009). Contextual lighting cues can override the light-from-above prior [Abstract]. *Journal of Vision, 9*(8):65, 65a, http://www.journalofvision.org/content/9/8/65, doi:10.1167/9.8.65.

Morgenstern, Y., Murray, R. F., & Harris, L. (in preparation). The light-from-above prior is weak.

Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research, 33,* 2685–2696.

Yuille, A. L., & Bülthoff, H. H. (1996). Bayesian decision theory and psychophysics. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian inference*. New York: Cambridge University Press.