Published in final edited form as:

Eur J Neurosci. 2010 May ; 31(10): 1721–1729. doi:10.1111/j.1460-9568.2010.07207.x.

Visual-vestibular cue integration for heading perception: Applications of optimal cue integration theory

Christopher R. Fetsch¹, Gregory C. DeAngelis^{2,*}, and Dora E. Angelaki^{1,*}

¹ Dept. of Anatomy & Neurobiology, Washington University School of Medicine, St. Louis, MO 63110

² Dept. of Brain and Cognitive Sciences, Center for Visual Science, University of Rochester, Rochester, NY 14627

Abstract

The perception of self-motion is crucial for navigation, spatial orientation, and motor control. In particular, estimation of one's direction of translation, or heading, relies heavily on multisensory integration in most natural situations. Visual and non-visual (e.g., vestibular) information can be used to judge heading, but each modality alone is often insufficient for accurate performance. It is not surprising, then, that visual and vestibular signals converge frequently in the nervous system, and that these signals interact in powerful ways at the level of behavior and perception. Early behavioral studies of visual-vestibular interactions consisted mainly of descriptive accounts of perceptual illusions and qualitative estimation tasks, often with conflicting results. In contrast, cue integration research in other modalities has benefited from the application of rigorous psychophysical techniques, guided by normative models that rest on the foundation of ideal-observer analysis and Bayesian decision theory. Here we review recent experiments that have attempted to harness these so-called optimal cue integration models for the study of self-motion perception. Some of these studies used nonhuman primate subjects, enabling direct comparisons between behavioral performance and simultaneously recorded neuronal activity. The results indicate that humans and monkeys can integrate visual and vestibular heading cues in a manner consistent with optimal integration theory, and that single neurons in area MSTd show striking correlates of the behavioral effects. This line of research and other applications of normative cue combination models should continue to shed light on mechanisms of self-motion perception and the neuronal basis of multisensory integration.

Keywords

visual; vestibular; human; macaque; psychophysics

Introduction

To effectively navigate and interact with objects in the environment, we must accurately estimate parameters of our own motion. Self-motion perception is computationally challenging for the nervous system, largely because it involves the integration of a complex array of visual and non-visual sensory cues. Since the pioneering work of James Gibson (1950; 1954; 1958), it has been widely recognized that visual cues provide a rich source of information about our movements through the environment. Specifically, as an observer

Corresponding author: Dora Angelaki, Dept. of Anatomy and Neurobiology, Box 8108, Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110, Tel: (314) 747-5529; Fax: (314) 747-4370, angelaki@pcg.wustl.edu.

*Equal author contributions

moves relative to a world-fixed scene, the resulting pattern of full-field retinal motion, known as optic flow, can be informative about the instantaneous direction of self-motion (heading). In the simplest example, translation with some forward component generates an expanding flow field, and the position of the focus of expansion corresponds to the observer's heading (Gibson, 1958). However, the real world is not so simple: translating observers regularly undergo movements of their eyes and head (Royden *et al.*, 1992; Royden *et al.*, 1994; Banks *et al.*, 1996; Crowell *et al.*, 1998), and may also experience motion of other objects in the scene (Gibson, 1954; Warren & Saunders, 1995; Royden & Hildreth, 1996; Warren & Rushton, 2009). Both of these factors greatly complicate the analysis of retinal motion as it pertains to self-motion.

A large body of work has been devoted to understanding the mechanisms by which the brain might compensate for these disruptions of the optic flow field and recover the true heading (for reviews see Lappe *et al.*, 1999; Warren, 2003; Britten, 2008). However, another (non-mutually exclusive) strategy would be to combine visual information with inertial signals that specify the motion of the head in space, such as from the vestibular otolith organs (Guedry, 1974; Fernandez & Goldberg, 1976b; a; Benson *et al.*, 1986). As the translational counterpart to the rotation-sensing semicircular canals, the otolith organs (utricle and saccule) transduce inertial acceleration by way of hair cells oriented in a variety of directions in the horizontal and vertical planes, respectively. Thus, otolith afferents provide directional linear acceleration signals that could be used to estimate heading, and unlike vision, these signals are unaffected by eye movements or object motion. The usefulness of otolith signals for heading perception might seem to be limited by their inability to distinguish translational motion from changes in orientation relative to gravity (the tilt-translation ambiguity; Angelaki *et al.*, 1999; Merfeld *et al.*, 1999; Angelaki *et al.*, 2004), but in fact this ambiguity is resolved in the brain by incorporating angular velocity signals from the semicircular canals (reviewed by Angelaki & Cullen, 2008; Angelaki & Yakusheva, 2009). However, other limitations of the vestibular system, such as its failure to encode sustained constant-velocity motion (Fernandez & Goldberg, 1976a; Büttner & Waespe, 1981; Si *et al.*, 1997), suggest that the nervous system cannot rely solely on vestibular signals for heading perception, just as it cannot rely solely on vision. Thus, a sensible strategy would be to combine visual and vestibular information when necessary to overcome the shortcomings of each modality on its own.

We should emphasize that, in addition to the vestibular system, other non-visual modalities such as somatosensation, proprioception, and audition may play important roles in heading perception. Somatosensory and proprioceptive cues, in addition to efference copy signals, are especially critical for the guidance of active movements (Mittelstaedt & Mittelstaedt, 2001; Sun *et al.*, 2004; Angelaki & Cullen, 2008) which comprise most natural self-motion. In this review we will focus on vestibular cues during passive movement, but we acknowledge that even controlled laboratory experiments cannot always rule out contributions from these other modalities.

Optimal cue integration theory

In general, the primary benefit of cue integration stems from the fact that all sensory information carries some degree of uncertainty, placing limits on perceptual performance. When multiple independent sources of information (or 'cues') are available, combining them can reduce this uncertainty and improve performance. This basic insight has been successfully applied in a number of theoretical and psychophysical studies of cue integration (e.g., Landy *et al.*, 1995; Jacobs, 1999; van Beers *et al.*, 1999; Landy & Kojima, 2001; Ernst & Banks, 2002; van Beers *et al.*, 2002; Knill & Saunders, 2003; Oruc *et al.*, 2003; Alais & Burr, 2004; Hillis *et al.*, 2004; Jürgens & Becker, 2006; MacNeilage *et al.*, 2007; Burr *et al.*, 2009; Ma *et al.*, 2009), each grounded in mathematical principles borrowed from ideal

observer analysis, signal detection theory, and/or statistical (Bayesian) decision theory. One of the simplest of such models (Cochran, 1937; Clark & Yuille, 1990; Landy *et al.*, 1995) posits that the lowest-variance unbiased estimate (denoted $\hat{S}$), given a set of n sensory cues, is one that takes a weighted linear sum of each single-cue estimate ($\hat{S}_i$):

$$\hat{S} = \sum_{i=1}^n w_i \hat{S}_i \quad (1)$$

The weights w_i are proportional to each cue's *reliability* (defined as its inverse variance, $1/\sigma^2$). Thus, for the case of two cues A and B, the optimal estimate is:

$$\hat{S}_{AB} = w_A \hat{S}_A + w_B \hat{S}_B \quad (2)$$

with weights given by:

$$w_A = \frac{1/\sigma_A^2}{1/\sigma_A^2 + 1/\sigma_B^2} \\ w_B = \frac{1/\sigma_B^2}{1/\sigma_A^2 + 1/\sigma_B^2} \quad (3)$$

The reliability of the combined estimate is simply the sum of the individual cue reliabilities:

$$\frac{1}{\sigma_{AB}^2} = \frac{1}{\sigma_A^2} + \frac{1}{\sigma_B^2}, \text{ or} \\ \sigma_{AB}^2 = \frac{\sigma_A^2 \sigma_B^2}{\sigma_A^2 + \sigma_B^2} \quad (4)$$

A distinct but related approach uses Bayes' rule to define optimal cue integration (for reviews see Yuille & Bülthoff, 1996; Mamassian *et al.*, 2002; Knill & Pouget, 2004). Bayesian models typically state that the goal of an optimal observer is to determine the value of some variable of interest, X , that maximizes the posterior distribution $P(X|A,B)$. This distribution describes the probability of each possible value of X , given the information provided by cues A and B. According to Bayes' rule, the posterior is proportional to the product of the likelihood function, $P(A,B|X)$, and the prior, $P(X)$. The former quantifies the relative likelihood of obtaining the sensory data for different values of X , and the latter is the probability of observing each value of X before receiving the sensory data at hand. If the cues have independent sources of noise, the joint likelihood can be parsed into separate likelihoods for A and B, and Bayes' rule becomes:

$$P(X|A, B) \propto P(A|X) * P(B|X) * P(X) \quad (5)$$

If one further assumes that the distributions are Gaussian, the implication of Eq. 5 is that the maximum a posteriori (MAP) estimate is a weighted sum of the single-cue estimates (the peaks of their respective likelihood functions) and the peak of the prior. Thus, the prior acts as an additional 'cue', contributing to the estimate in proportion to its inverse variance.

Priors allow Bayesian models to better account for behavioral performance by incorporating top-down expectations or biases about the stimuli, whether they are learned during the course of a task (Körding & Wolpert, 2004) or otherwise instantiated during development or through life experience (e.g., Weiss *et al.*, 2002; Stocker & Simoncelli, 2006). However, in

many cases the prior is assumed to be approximately uniform (e.g., Ernst & Banks, 2002; Alais & Burr, 2004; Hillis *et al.*, 2004). With this assumption, the predictions from the Bayesian framework become identical to those described by Eqs. 1–4: cues are weighted in proportion to their reliability, and the combined estimate is more precise than either single-cue estimate.

Regardless of the specifics, a key advantage of these models is that they often make clear predictions about behavior that are testable with standard psychophysical methods, provided the experimenter can manipulate each cue appropriately. Later in the review we describe some of these experiments in more detail, but first we will briefly summarize some of the earlier work on the interplay between visual and vestibular signals in perception.

Behavioral evidence for visual-vestibular interactions

The earliest demonstration of a perceptual link between the visual and vestibular modalities comes from the study ofvection, first reported by Ernst Mach in 1875. Vection refers to an illusory percept of self-motion induced by visual motion (Mach, 1875; Brandt *et al.*, 1972; Berthoz *et al.*, 1975; Dichgans & Brandt, 1978; Howard, 1982). The illusion is usually observed after several seconds of viewing a constant-velocity, large-field stimulus, similar to what one might observe while physically translating or rotating within a world-fixed environment. Subjects in these early experiments were often asked to make a verbal or joystick report of their perceived motion, i.e., either a yes-no detection task or an analog speed estimation task.

Perhaps owing to the variety of methodologies and lack of a unifying theoretical framework, there remains substantial controversy in the literature as to which stimulus parameters are necessary and/or sufficient to produce vection (e.g., Zacharias & Young, 1981; Probst *et al.*, 1985; Howard & Heckmann, 1989; Telford & Frost, 1993; Tarita-Nistor *et al.*, 2006). We speculate that some of these discrepancies could be resolved within the context of an optimal cue integration framework, but no published studies to our knowledge have attempted such an approach. In the model of Zacharias and Young (1981) for example, visual and vestibular cues were proposed to combine in a weighted fashion to estimate self-rotation velocity. However, the weights were determined not by cue reliability *per se*, but by factors such as the frequency content of the stimuli and the magnitude of the conflict between the two cues (Zacharias & Young, 1981; see also Wright *et al.*, 1995). Although the notion of frequency-based weighting has been disputed (Probst *et al.*, 1985), future studies could attempt to reconcile these findings by explicitly acknowledging the probabilistic nature of the task, and by seeking to measure and manipulate quantities such as cue reliability or the prior probability of real self-motion.

To our knowledge, the specific task of estimating heading from visual and vestibular cues was not systematically examined until a set of experiments by Telford, Howard, and Ohmi (Telford *et al.*, 1995; Ohmi, 1996). In the latter study, subjects' heading estimates showed greater precision for visual cues (average standard deviation $\approx 5.8^\circ$) compared with vestibular cues (s.d. $\approx 8.9^\circ$). Performance improved slightly in the combined visual-vestibular condition (s.d. $\approx 5.0^\circ$), but this difference did not reach significance. When cues were placed in conflict, subjects' perceived heading was dominated by either the visual or the vestibular cue, depending on the magnitude of the conflict. With a directional conflict of 30° , heading estimates were aligned with the visual cue, whereas with a conflict of 180° their estimates followed the vestibular cue (Ohmi, 1996). Another study (Harris *et al.*, 2000) used a match-to-sample task, asking subjects to estimate their distance traveled. Results showed that physical motion (e.g., vestibular) cues dominated the estimation of self-motion, contrary to the findings of Ohmi (1996). Taken together, these studies hint at a functional

interaction between visual and non-visual cues for heading perception. However, interpretation of these results is hampered by the lack of a rigorous psychophysical and modeling approach, and by the use of pointing and matching tasks that can introduce motor noise or systematic biases.

More recently, several groups have begun to use Bayesian or ideal-observer models, in conjunction with quantitative psychophysics, to study visual-vestibular cue integration. Jürgens and Becker (2006) studied the estimation of angular displacement, observing that subjects showed reduced variance when multiple cues (combinations of visual, vestibular, and kinesthetic) were available. The results were also consistent with the presence of a prior for some 'default' angular velocity. Consistent with Bayesian schemes, the influence of the prior diminished as the subjects' uncertainty was reduced by the addition of multiple cues. Notably, these authors suggested that a Bayesian prior may in fact explain a more general psychophysical phenomenon known as the 'range effect' (Poulton, 1975), in which subjects are biased to report estimates near the center of the range of presented stimuli.

MacNeilage and colleagues (2007) also applied a Bayesian approach to the problem of spatial orientation, detailing the specific contributions of optic flow and other visual signals for interpreting the gravito-inertial force detected by the otolith organs (i.e., resolving the tilt-translation ambiguity). They found that (a) estimates of tilt and acceleration were more precise when multiple cues were available as compared with any single-cue condition, and (b) more reliable visual cues were weighted more highly. This study further illustrated the utility of Bayesian theory for incorporating prior information into perceptual estimates, showing that known spatial orientation illusions are consistent with priors favoring zero linear acceleration and angular velocity (MacNeilage *et al.*, 2007; see also Laurens & Droulez, 2007). A recent series of experiments employing the 'subjective visual vertical' (SVV) task (De Vrijer *et al.*, 2008; 2009; Vingerhoets *et al.*, 2009) extended these findings and supported the existence of an upright prior. These and other modern computational methods for modeling spatial orientation are discussed further in a comprehensive review by MacNeilage *et al.* (2008).

In the domain of heading perception, our laboratory has recently introduced the principles of optimal cue integration to combined behavioral and neurophysiological experiments in rhesus monkeys (Gu *et al.*, 2008; Fetsch *et al.*, 2009). Using a motion platform and projector apparatus (Fig. 1A), Gu *et al.* (2008) trained monkeys to report their perceived heading as defined by inertial motion ('vestibular' condition), optic flow ('visual' condition; Fig. 1B), or a combination of both cues ('combined' condition). Note that previous labyrinthectomy experiments (Gu *et al.*, 2007) demonstrated severe and long-lasting performance deficits in the 'vestibular' condition, suggesting that the cues used by the animals in this condition were primarily of vestibular origin (rather than somatosensory or proprioceptive).

In each trial, the animal was presented with a 2-second heading trajectory with a Gaussian velocity profile. The direction of translation was varied in logarithmic steps around straight ahead (0°), and the subjects' task was to indicate their perceived heading via a leftward or rightward saccade (Fig. 1C). Psychometric functions were plotted as the proportion of rightward choices as a function of heading angle (leftward = negative, rightward = positive). The data were fit with a cumulative Gaussian function, and the psychophysical threshold was taken as the standard deviation (σ) of this best-fitting function. Figure 2A shows data from an example session. Note the steeper slope of the psychometric function in the combined condition (blue curve), which corresponds to a lower threshold, i.e., greater precision of the animal's heading estimates. In 6 out of 7 monkeys tested to date, average thresholds were significantly lower in the combined condition than either unimodal

condition, and the majority were not significantly different from the optimal prediction defined by Eq. 4 (Fig. 2B; data from Gu *et al.*, 2008 and Fetsch *et al.*, 2009).

Although this study demonstrated a perceptual improvement (reduction in variance) consistent with optimal integration of visual and vestibular heading cues, it could not address the prediction of cue reweighting in proportion to reliability (Eq. 3). To test this prediction, we modified the heading discrimination task in two ways (Fetsch *et al.*, 2009): (i) placing the cues in conflict on a proportion of trials (Fig. 3A), and (ii) varying the relative reliability of the cues by changing the motion coherence of the optic flow stimulus. This type of experimental design, sometimes known as a perturbation analysis, has been used previously in several human psychophysical studies of cue integration in other sensory modalities (e.g., Young *et al.*, 1993; Landy *et al.*, 1995; Ernst & Banks, 2002; Alais & Burr, 2004). The basic strategy is to measure psychophysical thresholds (σ) in the single-cue conditions, including multiple levels of cue reliability. These single-cue thresholds, taken as a proxy for cue reliability, are used to compute *predicted* weights according to Eq. 3. The predictions are then compared to *actual* (empirical) weights computed from the shifts of the psychometric functions during cue-conflict.

Average psychometric functions for one monkey are plotted for the single-cue conditions in Fig. 3B. Note that the two levels of motion coherence shown here were chosen such that the visual thresholds would be larger (24% coherence) or smaller (96% coherence) than the vestibular threshold (vestibular cue reliability was held constant). In this case, the vestibular threshold was 2.0° , and the visual thresholds were 2.8° and 1.1° for 24% and 96% coherence, respectively. From these threshold values, the model predicts that performance in the cue-conflict conditions should exhibit a higher vestibular weight (0.66) at 24% coherence and a lower vestibular weight (0.22) at 96% coherence. Plotted separately are the measured psychometric functions from the combined, cue-conflict conditions (Fig. 3C–D, symbols and solid curves), where each curve represents a different value of the conflict angle Δ . Choice data from positive Δ trials (visual to the right, vestibular to the left) are plotted in purple, while data from negative Δ trials (visual left, vestibular right) are in green. For each coherence level, the horizontal shift of the psychometric function, measured as the change in the point of subjective equality (PSE), is normalized by the magnitude of the perturbation ($\Delta/2$) to compute the actual weights that subjects assigned to each cue (Landy *et al.*, 1995; Ernst & Banks, 2002; Alais & Burr, 2004). For example, when Δ is positive, a rightward shift of the PSE corresponds to a high vestibular weight, while a leftward shift corresponds to a low vestibular weight (and vice versa when Δ is negative). The actual vestibular weights for this subject were 0.78 (24% coherence) and 0.36 (96% coherence), which were in the expected direction but slightly greater than the predictions of 0.66 and 0.22. For visual comparison, the predicted weights were converted to psychometric functions, shown as dashed curves in Fig. 3C–D. Note that the solid curves in Fig. 3C show a larger shift compared to the dashed curves, reflecting the fact that the actual vestibular weight (0.78) was greater than the predicted weight (0.66). In Fig. 3D, the shift of the solid curves was smaller than predicted, which also indicates a larger actual (0.36) vs. predicted (0.22) vestibular weight.

Figure 4 summarizes the predicted and actual weights for 5 human (A) and 5 monkey (B) subjects (Fetsch *et al.* 2009). All subjects demonstrated robust changes in actual weights as a function of motion coherence, and some subjects' weights were very close to the optimal predictions. However, the pattern illustrated in Fig. 3 was characteristic of the data as a whole: on average, both humans and monkeys appeared to over-weight the vestibular cue (or equivalently, under-weight the visual cue). Vestibular over-weighting was also observed in a similar study in humans carried out by Butler and colleagues (J. S. Butler, J. L. Campos, H. H. Bühlhoff, and S. T. Smith, unpublished observations), suggesting that this result is

robust and not a function of the specific monkey training regimen or instructions given to human subjects in our experiments. This modest but systematic deviation from optimality remains to be explained by future studies. One speculation is that the reliability of the optic flow cue cannot be estimated perfectly from data collected in the single-cue visual condition. Strictly speaking, this condition does not isolate the visual cue because the vestibular system remains active and indicates that the subject is stationary (i.e., it is a type of cue-conflict). Since the predicted weights depend on estimates of single-cue reliability, such a misestimation could give rise to apparent deviations from optimality. Nevertheless, the most critical observation is that all subjects show clear changes in cue weights as a function of cue reliability, indicating that both monkeys and humans dynamically re-weight visual and vestibular cues on a trial-by-trial basis during self-motion perception (Fetsch *et al.* 2009). These findings, together with those of Gu *et al.* (2008), establish for the first time an animal model that exhibits the principal features of cue integration represented by Eqs. 3 and 4.

Neural correlates of visual-vestibular cue integration

Having developed a paradigm for near-optimal visual-vestibular integration in monkeys, where in the brain should we expect to find neuronal populations that could mediate these behavioral effects? Unlike in other primary sensory systems, vestibular signals converge with visual, proprioceptive, and motor-related signals as early as one or two synapses from the periphery, in the brainstem vestibular nuclei (Henn *et al.*, 1974; Robinson, 1977; Waespe & Henn, 1977; Daunton & Thomsen, 1979) and vestibulo-cerebellum (Waespe *et al.*, 1981; Waespe & Henn, 1981; Markert *et al.*, 1988). However, putative responses to visual (optokinetic) stimuli within these subcortical structures may be more likely to be involved in reflexive eye movements or gaze stabilization (optokinetic nystagmus (OKN), vestibulo-ocular reflex (VOR), and/or smooth pursuit) rather than self-motion perception per se. In fact, a recent study (Bryan & Angelaki, 2009) showed relatively weak convergence of visual and vestibular inputs in the vestibular and fastigial nuclei when animals were required to fixate a head-fixed target (suppressing OKN). This lack of visual-vestibular convergence was also observed in vestibular-responsive areas of the thalamus (Meng & Angelaki, 2009). Similarly, areas traditionally recognized as ‘vestibular cortex’ (e.g., 2v and PIVC; Schwarz & Fredrickson, 1971; Grüsser *et al.*, 1990; Fukushima, 1997) are largely unresponsive to the types of optic flow stimuli used in the behavioral experiments described above (A. Chen & DEA, unpublished observations), and thus seem unlikely to play a substantial role in visual-vestibular integration for heading perception.

Perhaps surprisingly, two strong candidates have emerged in studies of parietal and extrastriate visual cortex, specifically the ventral intraparietal (VIP) and dorsal medial superior temporal (MSTd) areas. MSTd and VIP neurons have large receptive fields and directional selectivity for optic flow patterns that simulate self-motion (Tanaka *et al.*, 1986; Tanaka & Saito, 1989; Duffy & Wurtz, 1991; 1995; Schaafsma & Duysens, 1996; Bremmer *et al.*, 2002). In addition, they show some degree of compensation for distortions of the flow field due to real or simulated eye movements (Bradley *et al.*, 1996; Page & Duffy, 1999; Zhang *et al.*, 2004; Bremmer *et al.*, 2009), and they have been causally linked to heading perception from optic flow in microstimulation experiments (Britten & van Wezel, 1998; T. Zhang & K. H. Britten, unpublished observations). Of particular relevance here, at least 50–60% of MSTd neurons are also selective for physical translation in darkness (Duffy, 1998; Bremmer *et al.*, 1999; Gu *et al.*, 2006). This result suggested the presence of vestibular signals that may be useful for heading perception, and thus the potential for integration with visual (optic flow) signals. The vestibular origin of these responses was later confirmed by experiments which showed that MSTd cells were no longer tuned during inertial motion following bilateral labyrinthectomy (Gu *et al.*, 2007; Takahashi *et al.*, 2007). Unlike otolith

afferents, but similar to recently reported neurons in the cerebellar nodulus and uvula (Yakusheva *et al.*, 2007), MSTd cells respond selectively to translation and not tilt relative to gravity (Liu & Angelaki, 2009). Furthermore, the temporal dynamics of vestibular responses in MSTd are generally consistent with the encoding of stimulus velocity (Gu *et al.*, 2006; Liu & Angelaki, 2009), consistent with the dynamics of visual motion responses.

Taken together, these findings suggest a role for MSTd in visual-vestibular cue integration for heading perception. Ultimately, however, understanding the neural basis of multisensory integration requires measuring neuronal responses while cue integration is known to be taking place behaviorally. This was done recently by Gu *et al.* (2008), who identified a sub-population of MSTd neurons (those with congruent visual and vestibular preferences) that appear well suited to account for the behavioral result shown in Fig. 2. Heading tuning curves from one such ‘congruent’ cell are shown in Fig. 5A, where firing rate is plotted as a function of heading angle for the vestibular (black), visual (red), and combined (blue) conditions. The tuning became steeper in the combined condition, suggesting that the neuron increased its sensitivity to small changes in heading angle around straight forward. To quantify neuronal sensitivity in a manner comparable to behavior, we used signal detection theory (ROC analysis; Green & Swets, 1966; Bradley *et al.*, 1987; Britten *et al.*, 1992) to generate neurometric functions for each cell in the three stimulus conditions. This procedure quantifies the ability of an ideal observer to perform the same task as the monkey, based on the activity of a given cell. As with the psychometric data, we fitted these neurometric data with cumulative Gaussian functions (Fig. 5B) and defined the neuronal threshold as the standard deviation (σ) of the Gaussian. For the example neuron, the neuronal threshold was smallest in the combined condition (blue curve), confirming that the cell indeed increased its ability to discriminate small variations in heading angle when both cues were provided. Across the population of congruent cells, this decrease in neuronal thresholds was remarkably similar to the prediction from optimal cue integration models (Fig. 5C, compare blue and green bars). This suggests that congruent cells in MSTd satisfy one major prediction of optimal cue integration theory (Eq. 4).

Cue reweighting and the effects of reliability on MSTd responses

It remains to be demonstrated whether and how individual neurons implement the other major prediction of optimal cue integration schemes, the dynamic reweighting of cues based on reliability (Eq. 3). Our ongoing experiments examine the responses of MSTd neurons while monkeys perform the cue-conflict version of the heading discrimination task (Fig. 3A). One way to examine neuronal reweighting of cues is to carry out the same analysis depicted in Fig. 3, except using neurometric functions (computed via ROC analysis, as in Fig. 5B) in place of psychometric functions. Thus, like the behavioral analysis (Fetsch *et al.*, 2009), one can compute ‘predicted’ weights for a single neuron (derived from single-cue neuronal thresholds) and compare them with ‘actual’ (optimal) weights computed from the shifts of the neurometric functions based on cue-conflict trials. Preliminary results using this analysis strategy suggest that most congruent neurons in MSTd exhibit reweighting effects that are consistent with the optimal prediction defined by Eq. 3 (CRF, GCD & DEA, unpublished observations), although it is premature to draw firm conclusions regarding these findings.

If MSTd neurons show robust correlates of the dynamic cue reweighting effects observed in behavior (Figs. 3–4), this raises interesting questions regarding the underlying neuronal computations required to achieve this property. Recent theoretical work has suggested that a population of multisensory neurons with Poisson-like spiking statistics can account for optimal cue reweighting simply by summing linearly their inputs (Ma *et al.*, 2006; Ma *et al.*, 2008). In this scenario, the weights that each neuron applies to its unisensory inputs need not

change with cue reliability. Instead, summing the inputs with fixed weights automatically leads to optimal reweighting at the level of the population code, provided that response gain varies with stimulus strength (i.e., stimulus strength has a multiplicative effect on the entire tuning curve). However, in contrast with one prediction of Ma *et al.* (2006), a recent study in passively fixating animals (Morgan *et al.*, 2008) suggests that single neurons may alter the weights that they apply to their unisensory inputs as a function of cue reliability. Morgan *et al.* (2008) presented many combinations of visual and vestibular heading stimuli to MSTd neurons, including cue-conflict conditions, and attempted to predict the responses to combined (bimodal) stimuli from the single-cue responses. They showed that combined responses were well fit by a weighted linear sum of single-cue responses, using the equation $r_{\text{comb}} = w_{\text{ves}} * r_{\text{ves}} + w_{\text{vis}} * r_{\text{vis}} + c$. Here, w_{ves} and w_{vis} are the vestibular and visual weights (coefficients) that the neuron applies to its unimodal inputs, each r is the firing rate response in a given stimulus condition, and c is a constant. Interestingly, however, the weights of the best-fitting linear model varied with the relative reliability of the cues: as motion coherence of the optic flow was increased, visual weights also increased (Fig. 6A) and vestibular weights decreased (Fig. 6B).

Thus, there exists a mechanism, either within single neurons or perhaps more likely within the network, that appears to alter the weights that MSTd neurons apply to their visual and vestibular inputs. Preliminary results suggest that changes in neuronal weights also occur on a trial-by-trial basis during the cue-conflict heading discrimination task (CRF, GCD, & DEA, unpublished observations). However, it remains unclear how these changes in weights (Morgan *et al.*, 2008) might be related to our preliminary findings of a single-cell correlate of behavioral cue reweighting (measured with ROC-based neurometric functions, as discussed above). Further analyses and modeling efforts are needed to better understand the relationship between these distinct measures of neuronal cue reweighting, and how each might contribute to the behavior of the animal.

Ideally, future studies will be able to explore more directly how optimal cue integration is achieved at the population level. The overarching hypothesis exemplified by the work of Ma *et al.* (2006, 2008) is that cue reliability – perhaps in the form of a probability distribution over stimuli – is encoded implicitly by the activity of sensory neuronal populations (Foldiak, 1993; Sanger, 1996; Deneve *et al.*, 2001; Knill & Pouget, 2004; Ma *et al.*, 2006; Ma *et al.*, 2008). Additional single- and multi-electrode recordings, coupled with data-driven population modeling, are needed to conclusively test this hypothesis. However, even if one takes for granted that an implicit representation of cue reliability (uncertainty) exists in the responses of sensory neurons, an important follow-up question is whether and how this representation is decoded into a more explicit form (e.g., Jazayeri & Movshon, 2006; Jazayeri, 2008) that can be used to guide decisions and predict their outcomes.

Conclusions

In this review we have attempted to summarize traditional and modern approaches to the problem of cue integration for self-motion perception, as well as recent work investigating its possible neuronal basis. The application of ideal-observer models and Bayesian decision theory in this field should continue to yield new insights into how the brain solves the inherently multisensory problems of spatial orientation and navigation through complex environments. Our goal here was not to provide a detailed account of these models, but rather to highlight one example of their potential for expanding our knowledge of cross-modal interactions in the nervous system. Specifically, we suggest that a more complete understanding of visual-vestibular interactions – traditionally confined to the study of illusions, postural control, and eye movements – could be obtained through the application

of rigorous psychophysical and computational techniques more commonly used in other modalities.

We also suggest that heading perception is a good model for investigating neuronal mechanisms of multisensory integration in general, for two main reasons. First, there are well-defined brain structures that receive both visual and vestibular signals related to self-motion (e.g., macaque areas MSTd and VIP). Second, the task is well suited for using tools developed in classic studies of visual motion discrimination (Britten *et al.*, 1992; Shadlen *et al.*, 1996), including a principled method for manipulating cue reliability (motion coherence), and the use of signal-detection theory to compare neuronal and psychophysical performance. Vestibular research has lagged behind other sensory systems, particularly vision and somatosensation (Parker & Newsome, 1998; Luna *et al.*, 2005), in terms of understanding the information contained in neuronal responses, the uncertainty associated with that information, and how signals might be read out to inform perceptual decisions. Future studies in this vein could be useful for revealing the neural underpinnings of not only heading perception, but also classical phenomena such as vection, the somatogravic and oculogravic illusions (Clark & Graybiel, 1963; Wolfe & Cramer, 1970; Previc *et al.*, 1992), and biases in subjective verticality (Howard, 1982; De Vrijer *et al.*, 2009).

Acknowledgments

The work was supported by NIH-EY019087 and NIH-DC007620 (to DEA), NIH-EY016178 (to GCD), and NIH Institutional National Research Service Award 5-T32-EY13360-07. We thank Yong Gu and Paul R. MacNeilage for helpful suggestions on the manuscript, and Steven M. Nelson for creating Fig. 1B.

References

- Alais D, Burr D. The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol.* 2004; 14:257–262. [PubMed: 14761661]
- Angelaki DE, Cullen KE. Vestibular system: the many facets of a multimodal sense. *Annu Rev Neurosci.* 2008; 31:125–150. [PubMed: 18338968]
- Angelaki DE, McHenry MQ, Dickman JD, Newlands SD, Hess BJ. Computation of inertial motion: neural strategies to resolve ambiguous otolith information. *J Neurosci.* 1999; 19:316–327. [PubMed: 9870961]
- Angelaki DE, Shaikh AG, Green AM, Dickman JD. Neurons compute internal models of the physical laws of motion. *Nature.* 2004; 430:560–564. [PubMed: 15282606]
- Angelaki DE, Yakusheva TA. How vestibular neurons solve the tilt/translation ambiguity. Comparison of brainstem, cerebellum, and thalamus. *Ann N Y Acad Sci.* 2009; 1164:19–28. [PubMed: 19645876]
- Banks MS, Ehrlich SM, Backus BT, Crowell JA. Estimating heading during real and simulated eye movements. *Vision Res.* 1996; 36:431–443. [PubMed: 8746233]
- Benson AJ, Spencer MB, Stott JR. Thresholds for the detection of the direction of whole-body, linear movement in the horizontal plane. *Aviat Space Environ Med.* 1986; 57:1088–1096. [PubMed: 3790028]
- Berthoz A, Pavard B, Young LR. Perception of linear horizontal self-motion induced by peripheral vision (linear vection) basic characteristics and visual-vestibular interactions. *Exp Brain Res.* 1975; 23:471–489. [PubMed: 1081949]
- Bradley A, Skottun BC, Ohzawa I, Sclar G, Freeman RD. Visual orientation and spatial frequency discrimination: a comparison of single neurons and behavior. *J Neurophysiol.* 1987; 57:755–772. [PubMed: 3559700]
- Bradley DC, Maxwell M, Andersen RA, Banks MS, Shenoy KV. Mechanisms of heading perception in primate visual cortex. *Science.* 1996; 273:1544–1547. [PubMed: 8703215]
- Brandt T, Dichgans J, Koenig E. Perception of self-rotation (circular vection) induced by optokinetic stimuli. *Pflugers Arch.* 1972; 332(Suppl 332):R398.

- Bremmer F, Duhamel JR, Ben Hamed S, Graf W. Heading encoding in the macaque ventral intraparietal area (VIP). *Eur J Neurosci.* 2002; 16:1554–1568. [PubMed: 12405970]
- Bremmer F, Kubischik M, Pekel M, Hoffmann KP, Lappe M. Visual selectivity for heading in monkey area MST. *Exp Brain Res.* 2009
- Bremmer F, Kubischik M, Pekel M, Lappe M, Hoffmann KP. Linear vestibular self-motion signals in monkey medial superior temporal area. *Ann N Y Acad Sci.* 1999; 871:272–281. [PubMed: 10372078]
- Britten KH. Mechanisms of self-motion perception. *Annu Rev Neurosci.* 2008; 31:389–410. [PubMed: 18558861]
- Britten KH, Shadlen MN, Newsome WT, Movshon JA. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci.* 1992; 12:4745–4765. [PubMed: 1464765]
- Britten KH, van Wezel RJ. Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nat Neurosci.* 1998; 1:59–63. [PubMed: 10195110]
- Bryan AS, Angelaki DE. Optokinetic and vestibular responsiveness in the macaque rostral vestibular and fastigial nuclei. *J Neurophysiol.* 2009; 101:714–720. [PubMed: 19073813]
- Burr D, Banks MS, Morrone MC. Auditory dominance over vision in the perception of interval duration. *Exp Brain Res.* 2009
- Büttner U, Waespe W. Vestibular nerve activity in the alert monkey during vestibular and optokinetic nystagmus. *Exp Brain Res.* 1981; 41:310–315. [PubMed: 7215492]
- Clark B, Graybiel A. Contributing factors in the perception of the oculogravic illusion. *Am J Psychol.* 1963; 76:18–27. [PubMed: 14021494]
- Clark, JJ.; Yuille, AL. Data fusion for sensory information processing systems. Kluwer; Boston, MA: 1990.
- Cochran WG. Problems arising in the analysis of a series of similar experiments. *Journal of the Royal Statistical Society.* 1937; 4 (Suppl):102–118.
- Crowell JA, Banks MS, Shenoy KV, Andersen RA. Visual self-motion perception during head turns. *Nat Neurosci.* 1998; 1:732–737. [PubMed: 10196591]
- Daunton N, Thomsen D. Visual modulation of otolith-dependent units in cat vestibular nuclei. *Exp Brain Res.* 1979; 37:173–176. [PubMed: 488213]
- Deneve S, Latham PE, Pouget A. Efficient computation and cue integration with noisy population codes. *Nat Neurosci.* 2001; 4:826–831. [PubMed: 11477429]
- De Vrijer M, Medendorp WP, Van Gisbergen JA. Shared computational mechanism for tilt compensation accounts for biased verticality percepts in motion and pattern vision. *J Neurophysiol.* 2008; 99:915–930. [PubMed: 18094098]
- De Vrijer M, Medendorp WP, Van Gisbergen JA. Accuracy-precision trade-off in visual orientation constancy. *J Vis.* 2009; 9(9):1–15. [PubMed: 19271919]
- Dichgans, J.; Brandt, T. Visual-vestibular interaction: Effects on self-motion perception and postural control. In: Held, R.; Leibowitz, HW.; Teuber, HL., editors. *Handbook of Sensory Physiology.* Springer-Verlag; Berlin: 1978. p. 756-804.
- Duffy CJ. MST neurons respond to optic flow and translational movement. *J Neurophysiol.* 1998; 80:1816–1827. [PubMed: 9772241]
- Duffy CJ, Wurtz RH. Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. *J Neurophysiol.* 1991; 65:1346–1359. [PubMed: 1875244]
- Duffy CJ, Wurtz RH. Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *J Neurosci.* 1995; 15:5192–5208. [PubMed: 7623145]
- Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature.* 2002; 415:429–433. [PubMed: 11807554]
- Fernandez C, Goldberg JM. Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. I. Response to static tilts and to long-duration centrifugal force. *J Neurophysiol.* 1976a; 39:970–984. [PubMed: 824412]

- Fernandez C, Goldberg JM. Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. II. Directional selectivity and force-response relations. *J Neurophysiol.* 1976b; 39:985–995. [PubMed: 824413]
- Fetsch CR, Turner AH, DeAngelis GC, Angelaki DE. Dynamic reweighting of visual and vestibular cues during self-motion perception. *J Neurosci.* 2009; 29:15601–15612. [PubMed: 20007484]
- Foldiak, P. The ‘ideal homunculus’: statistical inference from neural population responses. In: Eeckman, FH.; Bower, JM., editors. *Computation and Neural Systems*. Kluwer Academic Publishers; Norwell, MA: 1993. p. 55-60.
- Fukushima K. Corticovestibular interactions: anatomy, electrophysiology, and functional considerations. *Exp Brain Res.* 1997; 117:1–16. [PubMed: 9386000]
- Gibson, JJ. *The Perception of the Visual World*. Houghton-Mifflin; Boston, MA: 1950.
- Gibson JJ. The visual perception of objective motion and subjective movement. *Psychol Rev.* 1954; 61:304–314. [PubMed: 13204493]
- Gibson JJ. Visually controlled locomotion and visual orientation in animals. *Br J Psychol.* 1958; 49:182–194. [PubMed: 13572790]
- Green, DM.; Swets, JA. *Signal Detection Theory and Psychophysics*. Wiley; New York: 1966.
- Grüsser OJ, Pause M, Schreier U. Vestibular neurones in the parieto-insular cortex of monkeys (*Macaca fascicularis*): visual and neck receptor responses. *J Physiol.* 1990; 430:559–583. [PubMed: 2086774]
- Gu Y, Angelaki DE, DeAngelis GC. Neural correlates of multisensory cue integration in macaque MSTd. *Nat Neurosci.* 2008; 11:1201–1210. [PubMed: 18776893]
- Gu Y, DeAngelis GC, Angelaki DE. A functional link between area MSTd and heading perception based on vestibular signals. *Nat Neurosci.* 2007; 10:1038–1047. [PubMed: 17618278]
- Gu Y, Watkins PV, Angelaki DE, DeAngelis GC. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J Neurosci.* 2006; 26:73–85. [PubMed: 16399674]
- Guedry, FE. Psychophysics of vestibular sensation. In: Kornhuber, HH., editor. *Handbook of Sensory Physiology, The Vestibular System*. Springer-Verlag; New York: 1974.
- Harris LR, Jenkin M, Zikovitz DC. Visual and non-visual cues in the perception of linear self-motion. *Exp Brain Res.* 2000; 135:12–21. [PubMed: 11104123]
- Henn V, Young LR, Finley C. Vestibular nucleus units in alert monkeys are also influenced by moving visual fields. *Brain Res.* 1974; 71:144–149. [PubMed: 4206917]
- Hillis JM, Watt SJ, Landy MS, Banks MS. Slant from texture and disparity cues: optimal cue combination. *J Vis.* 2004; 4:967–992. [PubMed: 15669906]
- Howard, IP. *Human Visual Orientation*. John Wiley; New York: 1982.
- Howard IP, Heckmann T. Circular vection as a function of the relative sizes, distances, and positions of two competing visual displays. *Perception.* 1989; 18:657–665. [PubMed: 2602091]
- Jacobs RA. Optimal integration of texture and motion cues to depth. *Vision Res.* 1999; 39:3621–3629. [PubMed: 10746132]
- Jazayeri M, Movshon JA. Optimal representation of sensory information by neural populations. *Nat Neurosci.* 2006; 9:690–696. [PubMed: 16617339]
- Jazayeri M. Probabilistic sensory recoding. *Curr Opin Neurobiol.* 2008; 18:431–437. [PubMed: 18809493]
- Jürgens R, Becker W. Perception of angular displacement without landmarks: evidence for Bayesian fusion of vestibular, optokinetic, podokinesthetic, and cognitive information. *Exp Brain Res.* 2006; 174:528–543. [PubMed: 16832684]
- Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 2004; 27:712–719. [PubMed: 15541511]
- Knill DC, Saunders JA. Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Res.* 2003; 43:2539–2558. [PubMed: 13129541]
- Körding KP, Wolpert DM. Bayesian integration in sensorimotor learning. *Nature.* 2004; 427:244–247. [PubMed: 14724638]

- Landy MS, Kojima H. Ideal cue combination for localizing texture-defined edges. *J Opt Soc Am A Opt Image Sci Vis.* 2001; 18:2307–2320. [PubMed: 11551065]
- Landy MS, Maloney LT, Johnston EB, Young M. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Res.* 1995; 35:389–412. [PubMed: 7892735]
- Lappe M, Bremmer F, van den Berg AV. Perception of self-motion from visual flow. *Trends Cogn Sci.* 1999; 3:329–336. [PubMed: 10461195]
- Laurens J, Droulez J. Bayesian processing of vestibular information. *Biol Cybern.* 2007; 96:389–404. [PubMed: 17146661]
- Liu S, Angelaki DE. Vestibular signals in macaque extrastriate visual cortex are functionally appropriate for heading perception. *J Neurosci.* 2009; 29:8936–8945. [PubMed: 19605631]
- Luna R, Hernandez A, Brody CD, Romo R. Neural codes for perceptual discrimination in primary somatosensory cortex. *Nat Neurosci.* 2005; 8:1210–1219. [PubMed: 16056223]
- Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. *Nat Neurosci.* 2006; 9:1432–1438. [PubMed: 17057707]
- Ma WJ, Beck JM, Pouget A. Spiking networks for Bayesian inference and choice. *Curr Opin Neurobiol.* 2008; 18:217–222. [PubMed: 18678253]
- Ma WJ, Zhou X, Ross LA, Foxe JJ, Parra LC. Lip-reading aids word recognition most in moderate noise: a Bayesian explanation using high-dimensional feature space. *PLoS One.* 2009; 4:e4638. [PubMed: 19259259]
- Mach, E. *Grundlinien der Lehre von den Bewegungsempfindungen.* W. Engelmann; Leipzig: 1875.
- MacNeilage PR, Banks MS, Berger DR, Bulthoff HH. A Bayesian model of the disambiguation of gravito-inertial force by visual cues. *Exp Brain Res.* 2007; 179:263–290. [PubMed: 17136526]
- MacNeilage PR, Ganesan N, Angelaki DE. Computational approaches to spatial orientation: from transfer functions to dynamic Bayesian inference. *J Neurophysiol.* 2008; 100:2981–2996. [PubMed: 18842952]
- Mamassian, P.; Landy, MS.; Maloney, LT. Bayesian modeling of visual perception. In: Rao, RPN.; Olshausen, BA.; Lewicki, MS., editors. *Probabilistic models of the brain: perception and neural function.* MIT Press; Cambridge, MA: 2002. p. 13–36.
- Markert G, Büttner U, Straube A, Boyle R. Neuronal activity in the flocculus of the alert monkey during sinusoidal optokinetic stimulation. *Exp Brain Res.* 1988; 70:134–144. [PubMed: 3261254]
- Meng H, Angelaki DE. Responses of ventral posterior thalamus neurons to three-dimensional vestibular and optic flow stimulation. *J Neurophysiol.* 2009 Epub ahead of print.
- Merfeld DM, Zupan L, Peterka RJ. Humans use internal models to estimate gravity and linear acceleration. *Nature.* 1999; 398:615–618. [PubMed: 10217143]
- Mittelstaedt ML, Mittelstaedt H. Idiothetic navigation in humans: estimation of path length. *Exp Brain Res.* 2001; 139:318–332. [PubMed: 11545471]
- Morgan ML, DeAngelis GC, Angelaki DE. Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron.* 2008; 59:662–673. [PubMed: 18760701]
- Ohmi M. Egocentric perception through interaction among many sensory systems. *Brain Res Cogn Brain Res.* 1996; 5:87–96. [PubMed: 9049074]
- Oruc I, Maloney LT, Landy MS. Weighted linear cue combination with possibly correlated error. *Vision Res.* 2003; 43:2451–2468. [PubMed: 12972395]
- Page WK, Duffy CJ. MST neuronal responses to heading direction during pursuit eye movements. *J Neurophysiol.* 1999; 81:596–610. [PubMed: 10036263]
- Parker AJ, Newsome WT. Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci.* 1998; 21:227–277. [PubMed: 9530497]
- Poulton EC. Observer bias. *Appl Ergon.* 1975; 6:3–8. [PubMed: 15677161]
- Previc FH, Varner DC, Gillingham KK. Visual scene effects on the somatogravic illusion. *Aviat Space Environ Med.* 1992; 63:1060–1064. [PubMed: 1456916]
- Probst T, Straube A, Bles W. Differential effects of ambivalent visual-vestibular-somatosensory stimulation on the perception of self-motion. *Behav Brain Res.* 1985; 16:71–79. [PubMed: 3875354]

- Robinson DA. Linear addition of optokinetic and vestibular signals in the vestibular nucleus. *Exp Brain Res.* 1977; 30:447–450. [PubMed: 413730]
- Royden CS, Banks MS, Crowell JA. The perception of heading during eye movements. *Nature.* 1992; 360:583–585. [PubMed: 1461280]
- Royden CS, Crowell JA, Banks MS. Estimating heading during eye movements. *Vision Res.* 1994; 34:3197–3214. [PubMed: 7975351]
- Royden CS, Hildreth EC. Human heading judgments in the presence of moving objects. *Percept Psychophys.* 1996; 58:836–856. [PubMed: 8768180]
- Sanger TD. Probability density estimation for the interpretation of neural population codes. *J Neurophysiol.* 1996; 76:2790–2793. [PubMed: 8899646]
- Schaafsma SJ, Duysens J. Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. *J Neurophysiol.* 1996; 76:4056–4068. [PubMed: 8985900]
- Schwarz DW, Fredrickson JM. Tactile direction sensitivity of area 2 oral neurons in the rhesus monkey cortex. *Brain Res.* 1971; 27:397–401. [PubMed: 4994681]
- Seung HS, Sompolinsky H. Simple models for reading neuronal population codes. *Proc Natl Acad Sci U S A.* 1993; 90:10749–10753. [PubMed: 8248166]
- Shadlen MN, Britten KH, Newsome WT, Movshon JA. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci.* 1996; 16:1486–1510. [PubMed: 8778300]
- Si X, Angelaki DE, Dickman JD. Response properties of pigeon otolith afferents to linear acceleration. *Exp Brain Res.* 1997; 117:242–250. [PubMed: 9419070]
- Stocker AA, Simoncelli EP. Noise characteristics and prior expectations in human visual speed perception. *Nat Neurosci.* 2006; 9:578–585. [PubMed: 16547513]
- Sun HJ, Campos JL, Chan GS. Multisensory integration in the estimation of relative path length. *Exp Brain Res.* 2004; 154:246–254. [PubMed: 14685814]
- Takahashi K, Gu Y, May PJ, Newlands SD, DeAngelis GC, Angelaki DE. Multimodal coding of three-dimensional rotation and translation in area MSTd: comparison of visual and vestibular selectivity. *J Neurosci.* 2007; 27:9742–9756. [PubMed: 17804635]
- Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E. Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J Neurosci.* 1986; 6:134–144. [PubMed: 3944614]
- Tanaka K, Saito H. Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol.* 1989; 62:626–641. [PubMed: 2769351]
- Tarita-Nistor L, Gonzalez EG, Spigelman AJ, Steinbach MJ. Linear vection as a function of stimulus eccentricity, visual angle, and fixation. *J Vestib Res.* 2006; 16:265–272. [PubMed: 17726279]
- Telford L, Frost BJ. Factors affecting the onset and magnitude of linear vection. *Percept Psychophys.* 1993; 53:682–692. [PubMed: 8332434]
- Telford L, Howard IP, Ohmi M. Heading judgments during active and passive self-motion. *Exp Brain Res.* 1995; 104:502–510. [PubMed: 7589301]
- van Beers RJ, Sittig AC, Gon JJ. Integration of proprioceptive and visual position-information: An experimentally supported model. *J Neurophysiol.* 1999; 81:1355–1364. [PubMed: 10085361]
- van Beers RJ, Wolpert DM, Haggard P. When feeling is more important than seeing in sensorimotor adaptation. *Curr Biol.* 2002; 12:834–837. [PubMed: 12015120]
- Vingerhoets RA, De Vrijer M, Van Gisbergen JA, Medendorp WP. Fusion of visual and vestibular tilt cues in the perception of visual vertical. *J Neurophysiol.* 2009; 101:1321–1333. [PubMed: 19118112]
- Waespe W, Büttner U, Henn V. Visual-vestibular interaction in the flocculus of the alert monkey. I. Input activity. *Exp Brain Res.* 1981; 43:337–348. [PubMed: 6266856]
- Waespe W, Henn V. Neuronal activity in the vestibular nuclei of the alert monkey during vestibular and optokinetic stimulation. *Exp Brain Res.* 1977; 27:523–538. [PubMed: 404173]

- Waespe W, Henn V. Visual-vestibular interaction in the flocculus of the alert monkey. II. Purkinje cell activity. *Exp Brain Res.* 1981; 43:349–360. [PubMed: 6266857]
- Warren PA, Rushton SK. Optic flow processing for the assessment of object movement during ego movement. *Curr Biol.* 2009; 19:1555–1560. [PubMed: 19699091]
- Warren, WH. Optic Flow. In: Chalupa, LM.; Werner, JS., editors. *The Visual Neurosciences*. MIT Press; Cambridge, MA: 2003. p. 1247-1259.
- Warren WH, Saunders JA. Perceiving heading in the presence of moving objects. *Perception.* 1995; 24:315–331. [PubMed: 7617432]
- Weiss Y, Simoncelli EP, Adelson EH. Motion illusions as optimal percepts. *Nat Neurosci.* 2002; 5:598–604. [PubMed: 12021763]
- Wolfe JW, Cramer RL. Illusions of pitch induced by centripetal acceleration. *Aerosp Med.* 1970; 41:1136–1139. [PubMed: 5458192]
- Wright WG, DiZio P, Lackner JR. Vertical linear self-motion perception during visual and inertial motion: more than weighted summation of sensory inputs. *J Vestib Res.* 2005; 15:185–195. [PubMed: 16286700]
- Yakusheva TA, Shaikh AG, Green AM, Blazquez PM, Dickman JD, Angelaki DE. Purkinje cells in posterior cerebellar vermis encode motion in an inertial reference frame. *Neuron.* 2007; 54:973–985. [PubMed: 17582336]
- Young MJ, Landy MS, Maloney LT. A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Res.* 1993; 33:2685–2696. [PubMed: 8296465]
- Yuille, AL.; Bülthoff, HH. Bayesian decision theory and psychophysics. In: Knill, DC.; Richards, W., editors. *Perception as Bayesian Inference*. Cambridge University Press; New York, NY: 1996. p. 123-162.
- Zacharias GL, Young LR. Influence of combined visual and vestibular cues on human perception and control of horizontal rotation. *Exp Brain Res.* 1981; 41:159–171. [PubMed: 6970678]
- Zhang T, Heuer HW, Britten KH. Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. *Neuron.* 2004; 42:993–1001. [PubMed: 15207243]

Figure 1.

Apparatus and stimuli used to study visual-vestibular cue integration. (A) ‘Virtual reality’ motion platform and projector setup providing independent control of visual (optic flow) and vestibular (inertial motion) cues to heading. (B) Depiction of the optic flow stimulus used by Gu *et al.* (2008) and Fetsch *et al.* (2009). The image is a composite of 5 sequential screen captures from the flow field (generated with OpenGL using in-house software) during simulated forward motion. (C) In the heading discrimination task, the monkey fixates a central target (panel 1) and is presented with a 2-s heading trajectory (inertial motion and/or optic flow) having a small leftward or rightward component (panel 2). The fixation point then disappears (panel 3), and the animal is required to report his perceived heading (‘left’ or ‘right’) by making a saccadic eye movement to one of two choice targets (panel 4).

Figure 2.

Monkey behavioral performance in the multimodal heading discrimination task (Gu *et al.*, 2008; Fetsch *et al.*, 2009). (A) Example psychometric functions for one behavioral session (vestibular condition: black circles and curve; visual condition: red circles and curve; combined condition: blue triangles and curve). Psychophysical threshold was defined as the standard deviation (σ) of the fitted cumulative Gaussian function. (B) Comparison of the average psychophysical thresholds obtained in the combined condition with that predicted based on statistically optimal cue integration (Eq. 4). Data are shown for 7 animals. Error bars indicate 95% confidence intervals.

Figure 3.

Cue-conflict version of the heading discrimination task (Fetsch *et al.*, 2009). (A) On two-thirds of trials, a conflict was introduced between visual and vestibular cues by moving the motion platform along a slightly different trajectory than the one being simulated with optic flow. In the schematic (left panel), the platform moves to the left while the optic flow indicates movement to the right. The conflict angle Δ could be positive (visual to the right, vestibular to the right) or negative (visual to the right, vestibular to the left). Heading angle on the cue-conflict trials (right panel, angle between dotted black lines) was defined as the midpoint between the vestibular and visual heading. (B) Psychometric functions for one monkey subject tested in the vestibular condition (black curve and circles) and in the visual condition at two coherence levels (24%: dotted red curve and triangles; 96%: solid red curve and squares). (C) Combined condition data at 24% motion coherence, plotted separately for the three conflict angles: $\Delta = -4^\circ$, 0° , $+4^\circ$ (green, orange, and purple curves, respectively). The leftward shift of the green curve and rightward shift of the purple curve indicate vestibular dominance (high vestibular weight, low visual weight). (D) Same as C, but for 96% motion coherence. Here the PSEs shift in the opposite direction, indicating visual dominance. The direction of these shifts was predicted by the relative single-cue thresholds (Eq. 3; optimal predictions shown by the dashed curves), but their magnitude suggests over-weighting of the vestibular cue.

Figure 4.

Summary of predicted and actual cue weights in the cue-conflict heading discrimination task (Fetsch *et al.*, 2009). Average vestibular weight ($1 - \text{visual weight}$) is plotted as a function of visual motion coherence for 5 human subjects (A) and 5 monkey subjects (B). Error bars represent $\pm$ SEM across subjects. Predicted weights (open symbols) were computed from Eq. 3 using single-cue thresholds, and actual weights (filled symbols) were computed from the shift of the PSE relative to the magnitude of cue-conflict (see Fig. 3C–D and associated text). The over-weighting of vestibular cues (filled symbols above open symbols) was statistically significant ($p < 0.01$) for both species.

Figure 5.

Heading sensitivity of MSTd neurons (Gu *et al.*, 2008). (A) Tuning curves (firing rate as a function of heading angle) for an example MSTd neuron with congruent visual and vestibular preferences (i.e., a stronger response for leftward headings vs. rightward headings across all conditions). (B) Neurometric functions computed by ROC analysis from the firing rate data plotted in A. Smooth curves show best-fitting cumulative Gaussian functions, and the neuronal thresholds (standard deviation, σ , of the fitted Gaussian) are indicated for each condition. (C) Average neuronal thresholds for the sub-population of congruent cells in MSTd, showing that thresholds in the combined condition (blue) were lower (i.e., neuronal

sensitivity was greater) than either single-cue condition (black and red), and not significantly different from the statistically optimal prediction (green).

Figure 6.

Effects of motion coherence (cue reliability) on the linear combination rule in MSTd (Morgan *et al.*, 2008). Visual (A) and vestibular (B) weights (coefficients) for each MSTd neuron were derived by fitting a simple linear weighted sum model to the bimodal responses, which included all possible combinations of visual and vestibular heading trajectories in the horizontal plane in 45-degree increments ($8 \times 8 = 64$ combinations; see Morgan *et al.*, 2008 for details). Each line connecting a pair of data points indicates a single neuron tested at multiple coherence levels, as indicated by the abscissa. Overall, visual weights significantly increased and vestibular weights significantly decreased as a function of increasing coherence.