Multimodal integration: Resolving sensory ambiguities to build novel representations

Andrea M. Green[1] and **Dora E. Angelaki**[2]

[1]Dépt. de Physiologie, Université de Montréal, Montréal, Québec, Canada

[2] Dept. of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri, USA

Abstract

Summary—Multisensory integration plays several important roles in the nervous system. One is to combine information from multiple complementary cues to improve stimulus detection and discrimination. Another is to resolve peripheral sensory ambiguities and create novel internal representations that do not exist at the level of individual sensors. Here we focus on how ambiguities inherent in vestibular, proprioceptive and visual signals are resolved to create behaviorally useful internal estimates of our self-motion. We review recent studies that have shed new light on the nature of these estimates and how multiple, but individually ambiguous, sensory signals are processed and combined to compute them. We emphasize the need to combine experiments with theoretical insights to understand the transformations that are being performed.

Keywords

vestibular; visual; computation; reference frame; eye movement; sensorimotor; reafference; spatial motion; multisensory; self-motion; spatial orientation

Introduction

One of the most widely recognized benefits of multisensory integration is the improvement in accuracy, precision or reaction times brought about by the simultaneous presentation of two or more sensory cues during sensory discrimination and/or detection tasks (e.g., see [1-4] for reviews). These interactions are often formulated using a probabilistic framework, whereby sensory cues are weighted in a Bayesian optimal fashion by taking into account both their reliability and our prior experiences [4-7,8*,9**]. However, there is another important benefit of multisensory integration: the information provided by an individual sensor is often ambiguous and can be resolved only by combining cues from multiple sensory sources. Unlike the first, more widely appreciated multisensory integration benefit, where each cue provides complementary information about the stimulus, here the brain needs to create unique internal representations that otherwise do not exist at the level of individual sensors. Given the complexity of such internal representations, computational hypotheses are essential to help

Correspondence should be addressed to: Dr. Andrea Green, Dépt. de Physiologie, Université de Montréal, 2960 Chemin de la Tour, Rm. 4141, Montréal, Québec, Canada H3T 1J4, andrea.green@umontreal.ca, tel: 514-343-6111 ext. 3301, fax: 514-343-2111.

understand both *what* information a neural population encodes and *how* the new representation is created.

As a proof of principle, here we review recent studies that examine how multimodal visual, vestibular and proprioceptive cues are integrated to create distinct representations of self-motion (i.e., how we move relative to the outside world). First we describe a unique sensory ambiguity faced by the vestibular system. We then discuss additional ambiguities, including the distinction of exafference versus reafference and the relationship between reference frames and the construction of novel internal representations that do not exist at the sensory periphery.

Multimodal integration for the estimation of inertial motion and spatial orientation

An example of multimodal integration necessary to resolve a peripheral sensory ambiguity and create novel sensory representations is found in the vestibular system. The ambiguity arises because the otolith organs, our sensors that detect linear acceleration, transduce both inertial (translational, $\mathbf{t}$) and gravitational ($\mathbf{g}$) accelerations, such that they sense net gravito-inertial acceleration ($\mathbf{a} = \mathbf{t} - \mathbf{g}$) [10-12]. Thus, as illustrated in Fig. 1a, the firing rates of otolith afferents are ambiguous in terms of the type of motion they encode: they could reflect either translation (e.g., lateral motion) or a head reorientation relative to gravity (e.g., roll tilt). Clearly, if the brain relied only on otolith information, one's actual motion could not be correctly identified.

Behavioral studies in humans and monkeys have shown that the brain resolves this sensory ambiguity by combining otolith signals with extra-otolith rotational signals that arise from either semicircular canal [13-16] or visual cues [17,18]. Specifically, rotational signals can provide an independent estimate of head reorientation relative to gravity ($\mathbf{g}$, tilt). The latter can then be subtracted from the otolith-driven, net acceleration signal ($\mathbf{a}$) to extract an estimate of translational motion ($\mathbf{t}$) (Fig. 2). Neural correlates of the solution to this sensory ambiguity have been described in Purkinje cell activity in the caudal vermis of monkeys [19**] as illustrated in Fig. 1b: modulation was strong during translation but weak during tilt. Individual neurons in the vestibular and deep cerebellar nuclei typically reflected only a partial solution (Fig. 1c) [10,20,21].

Guided by theoretical insights [22,23], these experiments also showed *how* multimodal sensory signals from the otoliths and canals were processed. To solve the ambiguity problem, a simple linear summation of sensory signals is inappropriate. Instead, the way otolith and canal signals must combine is head-orientation-dependent: head-centered canal signals must be transformed into a spatially-referenced estimate of the earth-horizontal rotation component ($\boldsymbol{\omega}_{EH}$) [19**, 20,22,23]. In addition, because for small tilt angles otolith activity is proportional to tilt *position*, whereas canals encode angular *velocity*, canal-derived estimates of $\boldsymbol{\omega}_{EH}$ should be temporally integrated ($\mathbf{g} \approx \int \boldsymbol{\omega}_{EH}$; Fig. 2) to match up with otolith signals [14,22]. Indeed, translation-encoding Purkinje cells in the caudal vermis show evidence for both of these transformations [19**].

Collectively, these studies have emphasized the importance of combining experiments with theoretical approaches to understand how individual sensory cues must be transformed to create new but valuable central representations. Next we consider how similar issues arise in considering the problems of distinguishing active from passive movement.

Exafference versus reafference: Distinguishing visual and vestibular motion signals that are self-generated from those that are externally applied

An important challenge for sensory systems is to differentiate between sensory inputs that arise from changes in the world and those that result from our own voluntary actions. As pointed out by von Helmholtz [24], this dilemma is notably experienced during eye movements: although targets rapidly jump across the retina as we move our eyes to make saccades, we never see the world move over our retina. Yet, tapping on the canthus of the eye to displace the retinal image results in an illusionary shift of the visual world. To address this problem von Holst and Mittelstaedt [25] proposed the "principle of reafference", where a copy of the expected sensory activation from a motor command is subtracted from the actual sensory signal, thereby eliminating the portion of the sensory signal resulting from the motor command (termed "reafference") and isolating the portion of the sensory signal arising from external stimuli (termed "exafference"). The latter can then be used to generate a perception of events in the outside world. In the case of estimating self-motion, the brain must eliminate visually- or vestibularly-driven motion information that results from our own head and eye movements. This problem can be considered as a sensory ambiguity that needs to be resolved, as no sensor can distinguish between exafference and reafference on its own; only by combining sensory signals with information about intended motor action can this ambiguity be resolved.

Recent neural recording studies in monkeys [26-30] have provided solid evidence that, unlike vestibular afferents, second-order vestibular neurons in the brainstem have attenuated responses (or none at all) to vestibular stimuli generated by active head movements. Importantly, they also delineated how such a selective elimination of vestibular signals is achieved. Neither neck motor efference copy nor proprioception cues alone were sufficient to selectively suppress reafference [29,30]. Instead, in a clever experiment Roy and Cullen [30] eliminated the vestibular signal during an active head movement by simultaneously passively rotating the head in the opposite direction, thereby ensuring that even though the head moved relative to the body it remained stationary in space. This approach unmasked the presence of an underlying reafference "cancellation" signal that was generated only when the activation of neck proprioceptors matched the motor-generated expectation (Fig. 3).

The problem of distinguishing exafference versus reafference is equally important for interpreting visual signals. Visual information, such as optic flow, can be used to judge self-motion under certain conditions [31,32]. However, visual signals alone are typically insufficient because retinal patterns of optic flow are confounded or contextually modified by changes in eye or head position. Psychophysical studies indicate that despite this confound human observers can accurately judge heading from optic flow, even while making smooth pursuit eye movements that distort the flow field [33-35]. This ability implies that the neural processing of optic flow signals must somehow take into account and cancel the visual motion caused by self-generated eye and head movements.

A number of recent neurophysiological studies in monkeys have investigated whether such compensation occurs in extrastriate visual areas such as the medial superior temporal (MST) and ventral intraparietal (VIP) areas. Neurons in these areas have large receptive fields and are selective for the speed and direction of visual motion stimuli. Consequently, areas such as MST and VIP appear to be optimal for the analysis of whole-field optic flow stimuli such as those experienced during self-motion (see review by [36]). Inaba et al. [37**] and Chukoskie and Movshon [38**] compared the speed tuning in medial temporal (MT) and MST cells when optic flow stimuli were presented during fixation versus pursuit. They found that whereas most MT cell responses were more consistent with encoding the total optic flow signal (i.e., exafference and reafference) the tuning curves of many MST cells showed at least partial compensation for the optic flow component caused by the eye movement. However, such

compensation was incomplete for most MST neurons, suggesting that further computations take place in downstream areas [38**]. Other studies have measured how MST and VIP estimates of simulated heading direction (i.e., encoding of the focus of expansion of an optic flow field) are affected by pursuit eye movements [39-44]. These studies have also reported partial, but incomplete, compensation for shifts in the focus of expansion on the retina caused by pursuit eye movements.

The exact nature of the reafference "cancellation" signal used in these cortical visual areas, and the extent to which it resembles that uncovered in the vestibular brainstem, nonetheless remains poorly characterized. Future studies should thus more quantitatively examine *what* visual motion parameters MST/VIP populations (and potentially those in other visual motion-sensitive areas) encode and *how* efference copy and/or eye proprioceptive information must be processed to create these representations.

Sensory integration for novel internal representations and the role of reference frames

As outlined above, peripheral sensory representations are often limited by the particular properties of the sensors. For the example topic of self-motion we highlight in this review, note that peripheral vestibular signals encode the motion of the *head*. However, daily tasks such as navigating during locomotion, executing appropriate postural responses and planning voluntary limb movements often require knowledge of the orientation and motion of the *body*. Vestibular signals provide ambiguous information in this regard, as the vestibular sensors are stimulated in a similar fashion regardless of whether the head moves alone or the head and body move in tandem. Similarly, neck proprioceptors are stimulated similarly if the body moves under the stationary head or the head moves with respect to the body. Signals from both sensory sources must therefore be combined to distinguish motion of the body from motion of the head (Fig. 4a). A complicating factor in combining signals from the two sources is that they are attached to different body parts such that the motion signals they encode depend on how the head is statically oriented with respect to the body. For example, because the vestibular sensors are fixed in the head, the way in which individual sensors are stimulated for a given direction of body motion is head-orientation-dependent. This suggests that, to correctly interpret the relationship between the pattern of sensory vestibular activation and body motion, vestibular signals are likely to undergo a reference frame transformation from a head-centered towards a body-centered reference frame. Such a computation requires a nonlinear interaction between dynamic vestibular estimates of head motion and static neck proprioceptive estimates of head orientation with respect to the body (Fig. 4a). Importantly, the combination of dynamic neck proprioceptive signals that indicate a given head/body motion might also vary as a function of current head orientation and thus might need to undergo similar types of transformations to match up spatially with vestibular signals (e.g., see [45**])

In keeping with the need for such reference frame transformations, a mixture of head and body reference frames for encoding vestibular signals has been reported in the rostral medial region of the deep cerebellar nuclei, known as the rostral fastigial nuclei [46,47]. These experiments measured neural responses to vestibular stimulation in monkeys when the head and body were moved in tandem as the static orientation of the head relative to the body was systematically varied. They showed that neural responses varied with head position in a manner consistent with a partial transformation of vestibular signals from head to body-centered coordinates. For example, as shown by Kleine et al. [46] (Fig. 4b), neural responses to rotation about a head-centered (pitch/roll) axis typically varied as a function of body re head orientation (Fig 4b; compare responses across each row) but were often similar when compared for rotations about a common body-centered axis (Fig. 4b; compare responses in blue and green boxes) illustrating that many cells encoded vestibular signals in a frame closer to body- than head-centered.

However, while these experiments examined in what reference frame vestibular signals were expressed, they did not address *what* these neurons ultimately encode. In fact, neurons could be coding either head or body motion in this altered (body-centered) reference frame.

Brooks and Cullen [45**] recently explicitly addressed the question of *what* motion representation is encoded by rostral fastigial neurons by testing their responses to different combinations of vestibular and proprioceptive stimulation during passive rotation about a single axis. They showed that approximately half of the neurons encoded head and the other half encoded body motion; that is, they responded both to vestibular stimulation when the head and body were passively moved in tandem and to neck proprioceptive stimulation when the body was passively moved beneath the head (Fig. 4c; left, middle columns). Most importantly, when the head was passively moved relative to the stationary body, they showed that proprioceptive and vestibular signals combined precisely to cancel one another out (Fig. 4c; right column).

In considering these experimental results it is important to emphasize that what neurons encode (i.e., head versus body motion) and the reference frames in which these signals are represented may be distinct and somewhat independent properties. For example, as noted above, neurons that show a transformation towards a body-centered reference frame could either encode head motion (of purely vestibular origin) or body motion (a new signal created by combining vestibular and proprioceptive information). Conversely, a computed estimate of body motion need not be encoded in a body-centered reference frame but could instead be represented in a frame closer to head-centered or neck muscle-centered. At present, the relationship between the signal coded by rostral fastigial neurons and the reference frame in which it is expressed remains unknown.

It is also important to make conceptual distinctions between (1) what is being encoded (e.g., do neurons carry information about passive head motion or passive body motion?); (2) whether there is a distinction between exafference and reafference (e.g., do neurons respond similarly to the same sensory stimulus – be it head or body motion - when it is actively versus passively generated?) and (3) the reference frames in which the coded signal is represented. Addressing question 1 requires testing neural responses for multiple combinations of sensory stimuli (all of which are either active or passive), whereas question 2 requires comparing those responses during passive versus active movement. Finally, identifying the reference frames in which central signals are encoded (question 3) requires testing for invariance of spatial tuning of a given sensory signal or internal motion estimate as static eye, head or body position is varied.

Despite clear conceptual distinctions, these three issues have often been confounded in the interpretation of experimental results. For example, referring back to optic flow responses in areas MST and VIP, conclusions regarding what signal is carried by individual neurons (e.g., exafference versus reafference) have often been confounded with the issue of reference frames. On one hand, the tuning of some MSTd/VIP neurons was shown to reflect a partial compensation for smooth eye and/or head movements [37**,38**,39-42,48], a finding that shows that these neurons distinguish at least partially between exafference and reafference. Yet, several studies have used such observations to infer that rather than encoding optic flow in eye-centered coordinates, neurons are computing a head, body or world-centered optic flow representation [39,44,48]. Notably, despite an at least partial compensation for pursuit eye movements (i.e., an exafference/reafference distinction), when Fetsch et al. [49] explicitly quantified the spatial tuning curves of MST neurons at different initial eye positions (the experimental manipulation appropriate for characterizing reference frames) they demonstrated that optic flow was actually encoded in an eye-centered reference frame. Thus, although in general, reference frame transformations, the creation of novel sensory representations, and distinguishing exafference from reafference can all occur within the same populations of cells,

these conceptual distinctions are important before we can thoroughly understand what neurons encode and how these representations are generated by multisensory integration.

Conclusions

These examples illustrate an important benefit of multisensory integration: the resolution of peripheral sensory ambiguities and the creation of novel internal representations. Importantly, we have emphasized that appropriate interpretation of experimental results needs to be accompanied by the formulation of clear hypotheses encompassing both what neurons should encode and how multimodal signals should be combined. The resolution of sensory ambiguities and construction of novel sensory representations may provide an explanation for why multimodal integration is often observed even at the earliest sensory processing stages in brain areas typically considered to be associated with a particular sensory modality. These novel representations might appear at the level of individual neurons or be encoded only at the level of a neural population. Synthesizing the multitude of diverse observations is particularly difficult as the functional role of many regions in complex tasks still remains unclear, but represents an exciting challenge for future work.

Acknowledgments

References

1. Stein BE, Stanford TR. Multisensory integration: current issues from the perspective of the single neuron. Nat Rev Neurosci 2008;9:255–266. [PubMed: 18354398]

2. Driver J, Noesselt T. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. Neuron 2008;57:11–23. [PubMed: 18184561]

3. Bulkin DA, Groh JM. Seeing sounds: visual and auditory interactions in the brain. Curr Opin Neurobiol 2006;16:415–419. [PubMed: 16837186]

4. Angelaki DE, Gu Y, Deangelis GC. Multisensory integration: psychophysics, neurophysiology, and computation. Curr Opin Neurobiol. 2009

5. Yuille, AL.; Bülthoff, HH., editors. Bayesian decision theory and psychophysics. New York: Cambridge University Press; 1996.

6. Mamassian, P.; Landy, MS.; Maloney, LT., editors. Bayesian modeling of visual perception. Cambridge, MA: MIT Press; 2002.

7. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 2004;27:712–719. [PubMed: 15541511]

8. 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] * Human psychophysical studies have provided support for a Bayesian probabilistic framework for cue integration by showing that like an optimal bayesian observer human subjects often combine sensory cues by weighting them in proportion to their reliability. This study is the first to show that monkeys also perform such dynamic cue weighting. Monkeys trained to perform a heading discrimination task were shown to rapidly reweight visual and vestibular cues according to their reliability, albeit not always in a Bayesian optimal fashion.

9. Gu Y, Angelaki DE, Deangelis GC. Neural correlates of multisensory cue integration in macaque MSTd. Nat Neurosci 2008;11:1201–1210. [PubMed: 18776893] ** This is the first study to examine multisensory integration in neurons while trained animals performed a behavioral task that requires perceptual cue integration. This study showed that nonhuman primates, like humans, can combine cues near-optimally, and it identified a population of neurons in the areas MSTd that may form the neural substrate for perceptual integration of visual and vestibular cues to heading.

10. 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]

11. 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 1976;39:970–984. [PubMed: 824412]

12. Fernandez C, Goldberg JM. Physiology of peripheral neurons innervating otolith organs of the squirrel monkey III Response dynamics. J Neurophysiol 1976;39:996–1008. [PubMed: 824414]

13. 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]

14. Green AM, Angelaki DE. Resolution of sensory ambiguities for gaze stabilization requires a second neural integrator. J Neurosci 2003;23:9265–9275. [PubMed: 14561853]

15. Merfeld DM, Park S, Gianna-Poulin C, Black FO, Wood S. Vestibular perception and action employ qualitatively different mechanisms II VOR and perceptual responses during combined Tilt&Translation. J Neurophysiol 2005;94:199–205. [PubMed: 15730979]

16. Merfeld DM, Zupan L, Peterka RJ. Humans use internal models to estimate gravity and linear acceleration. Nature 1999;398:615–618. [PubMed: 10217143]

17. MacNeilage PR, Banks MS, Berger DR, Bulthoff HH. A Bayesian model of the disambiguation of gravitoinertial force by visual cues. Exp Brain Res 2007;179:263–290. [PubMed: 17136526]

18. Zupan LH, Merfeld DM. Neural processing of gravito-inertial cues in humans IV Influence of visual rotational cues during roll optokinetic stimuli. J Neurophysiol 2003;89:390–400. [PubMed: 12522188]

19. 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] ** This study provides the strongest evidence to date that brainstem-cerebellar circuits resolve the tilt/translation ambiguity by performing a set of theoretically predicted computations (Fig. 2). In contrast to earlier studies which revealed a distributed representation of translation in the vestibular and fastigial nuclei [10,20,21], this new study identified a neural correlate for translation at the level of individual neurons in the posterior vermis. Most importantly, it provided evidence (for the range of motions considered) that the canal signals which contribute to this translation estimate have been transformed, as theoretically required, from the head-centered reference frame of the vestibular sensors into a spatially-referenced estimate of the earth-horizontal component of rotation.

20. Green AM, Shaikh AG, Angelaki DE. Sensory vestibular contributions to constructing internal models of self-motion. J Neural Eng 2005;2:S164–179. [PubMed: 16135882]

21. Shaikh AG, Green AM, Ghasia FF, Newlands SD, Dickman JD, Angelaki DE. Sensory convergence solves a motion ambiguity problem. Curr Biol 2005;15:1657–1662. [PubMed: 16169488]

22. Green AM, Angelaki DE. An integrative neural network for detecting inertial motion and head orientation. J Neurophysiol 2004;92:905–925. [PubMed: 15056677]

23. Green AM, Angelaki DE. Coordinate transformations and sensory integration in the detection of spatial orientation and self-motion: from models to experiments. Prog Brain Res 2007;165:155–180. [PubMed: 17925245]

24. von Helmholtz, H. Handbuch der Physiologischen Optik [Treatise on Physiological Optics]. Rochester: JPC Southall; 1925. Opt Soc Am

25. von Holst E, Mittelstaedt H. Das reafferenzprinzip. Naturwissenschaften 1950;37:464–476.

26. Cullen KE, Minor LB. Semicircular canal afferents similarly encode active and passive head-on-body rotations: implications for the role of vestibular efference. J Neurosci 2002;22:RC226. [PubMed: 12040085]

27. Jamali M, Sadeghi SG, Cullen KE. Response of vestibular nerve afferents innervating utricle and saccule during passive and active translations. J Neurophysiol 2009;101:141–149. [PubMed: 18971293]

28. McCrea RA, Gdowski GT, Boyle R, Belton T. Firing behavior of vestibular neurons during active and passive head movements: vestibulo-spinal and other non-eye-movement related neurons. J Neurophysiol 1999;82:416–428. [PubMed: 10400968]

29. Roy JE, Cullen KE. Selective processing of vestibular reafference during self-generated head motion. J Neurosci 2001;21:2131–2142. [PubMed: 11245697]

30. Roy JE, Cullen KE. Dissociating self-generated from passively applied head motion: neural mechanisms in the vestibular nuclei. J Neurosci 2004;24:2102–2111. [PubMed: 14999061]

31. Gibson, JJ. The perception of the visual world. Boston: Houghton Mifflin; 1950.

32. Warren, WH., Jr, editor. Optic Flow. Cambridge, MA: MIT Press; 2003.

33. Banks MS, Ehrlich SM, Backus BT, Crowell JA. Estimating heading during real and simulated eye movements. Vision Res 1996;36:431–443. [PubMed: 8746233]

34. Royden CS, Banks MS, Crowell JA. The perception of heading during eye movements. Nature 1992;360:583–585. [PubMed: 1461280]

35. Warren WH Jr, Hannon DJ. Eye movements and optical flow. J Opt Soc Am A 1990;7:160–169. [PubMed: 2299447]

36. Britten KH. Mechanisms of self-motion perception. Annu Rev Neurosci 2008;31:389–410. [PubMed: 18558861]

37. Inaba N, Shinomoto S, Yamane S, Takemura A, Kawano K. MST neurons code for visual motion in space independent of pursuit eye movements. J Neurophysiol 2007;97:3473–3483. [PubMed: 17329625] ** This study examined the responses of MT and MST neurons to a moving visual stimulus (optic flow) while monkeys performed a smooth pursuit task and directly compared them to those observed when the same visual stimulus was presented during fixation. The study showed that whereas MT neuron responses were best correlated with visual motion on the retina, MST neurons were better correlated with motion in the world independent of eye movement, providing evidence that MST cells selectively suppress responses to self-induced retinal slip (reafference).

38. Chukoskie L, Movshon JA. Modulation of visual signals in macaque MT and MST neurons during pursuit eye movement. J Neurophysiol 2009;102:3225–3233. [PubMed: 19776359] ** This study investigated the extent to which MT and MST neurons distinguish optic flow patterns due to passive stimulation (exafference) from those due to active eye movement (reafference) by comparing the speed tuning of these neural populations to optic flow stimulation during fixation and pursuit. The study showed that many MST neurons show changes in their sensitivities to optic flow during pursuit that are consistent with an at least partial compensation for the portion of the optic flow stimulus presumed to be due to the eye movement itself.

39. 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]

40. Page WK, Duffy CJ. MST neuronal responses to heading direction during pursuit eye movements. J Neurophysiol 1999;81:596–610. [PubMed: 10036263]

41. Shenoy KV, Bradley DC, Andersen RA. Influence of gaze rotation on the visual response of primate MSTd neurons. J Neurophysiol 1999;81:2764–2786. [PubMed: 10368396]

42. Shenoy KV, Crowell JA, Andersen RA. Pursuit speed compensation in cortical area MSTd. J Neurophysiol 2002;88:2630–2647. [PubMed: 12424299]

43. Lee B, Pesaran B, Andersen RA. Translation speed compensation in the dorsal aspect of the medial superior temporal area. J Neurosci 2007;27:2582–2591. [PubMed: 17344395]

44. 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]

45. Brooks JX, Cullen KE. Multimodal integration in rostral fastigial nucleus provides an estimate of body movement. J Neurosci 2009;29:10499–10511. [PubMed: 19710303] ** This is the first study to identify a neural correlate for trunk motion in the rostral fastigial nucleus of the macaque cerebellum. It showed that about half of the neurons in this area combine vestibular and neck proprioceptive signals to construct an estimate of body motion.

46. Kleine JF, Guan Y, Kipiani E, Glonti L, Hoshi M, Buttner U. Trunk position influences vestibular responses of fastigial nucleus neurons in the alert monkey. J Neurophysiol 2004;91:2090–2100. [PubMed: 15069099]

47. Shaikh AG, Meng H, Angelaki DE. Multiple reference frames for motion in the primate cerebellum. J Neurosci 2004;24:4491–4497. [PubMed: 15140919]

48. Ilg UJ, Schumann S, Thier P. Posterior parietal cortex neurons encode target motion in world-centered coordinates. Neuron 2004;43:145–151. [PubMed: 15233924]

49. Fetsch CR, Wang S, Gu Y, Deangelis GC, Angelaki DE. Spatial reference frames of visual, vestibular, and multimodal heading signals in the dorsal subdivision of the medial superior temporal area. J Neurosci 2007;27:700–712. [PubMed: 17234602]

50. Angelaki DE, Yakusheva TA, Green AM, Dickman JD, Blazquez PM. Computation of Egomotion in the Macaque Cerebellar Vermis. Cerebellum. 2009
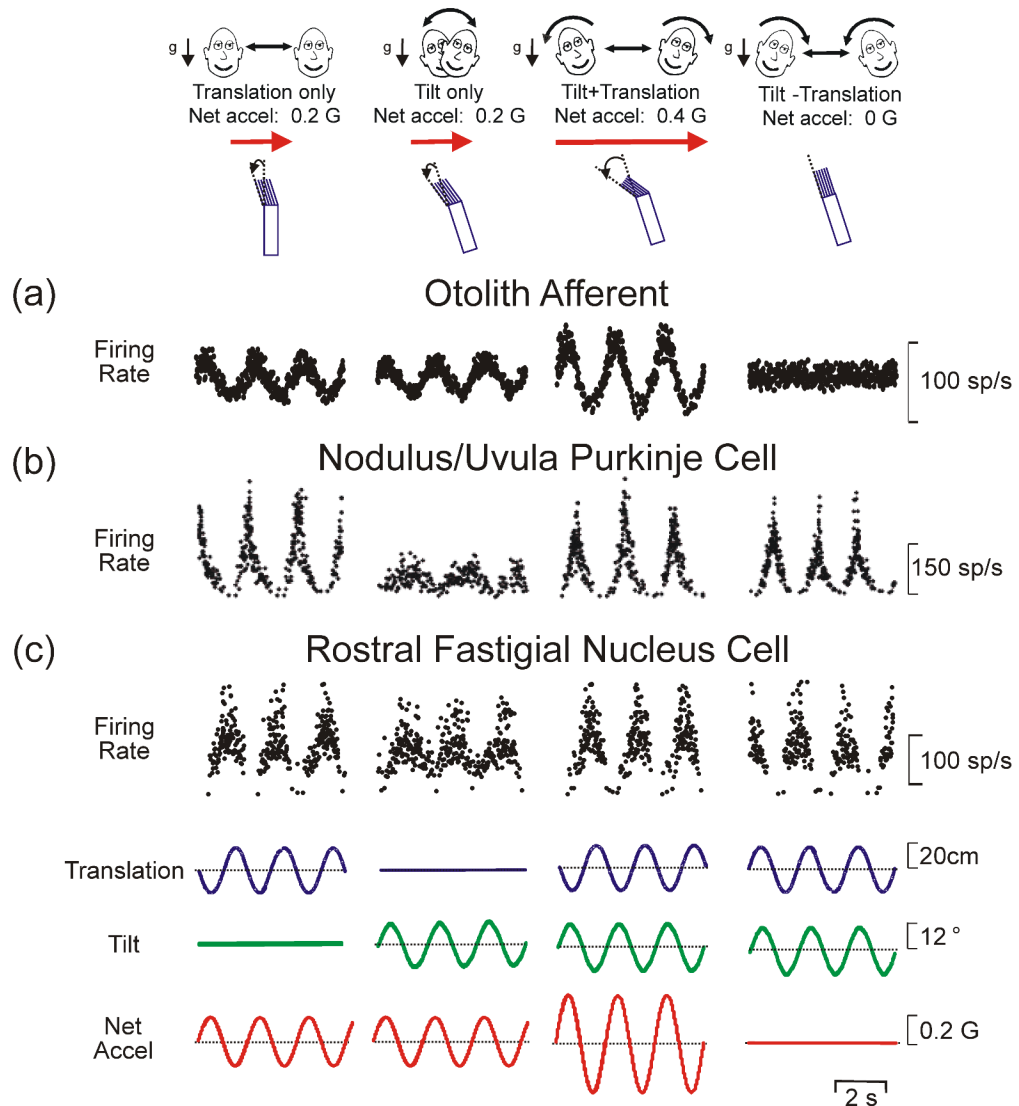
Fig. 1.
Evidence for a neural resolution to the tilt/translation ambiguity. Responses of (**a**) an otolith afferent, (**b**) a Purkinje cell in the nodulus/uvula region of the caudal vermis and (**c**) a neuron in the rostral fastigial nucleus during translation, tilt and combinations of these stimuli presented either in phase to double the net acceleration ("Tilt+Translation") or out-of-phase to cancel it out ("Tilt–Translation"). Unlike otolith afferents (**a**), which provide ambiguous motion information because their responses always reflect net acceleration, nodulus/uvula Purkinje cells (**b**) selectivily encode translation [19**]. Deep cerebellar and vestibular nuclei cells (**c**) show intermediate responses, thus reflecting a partial solution to the ambiguity [10, 20,21]. Neural data are replotted with permission from Angelaki et al. [10,50], Yakusheva et al. [19**] and Green et al. [20].
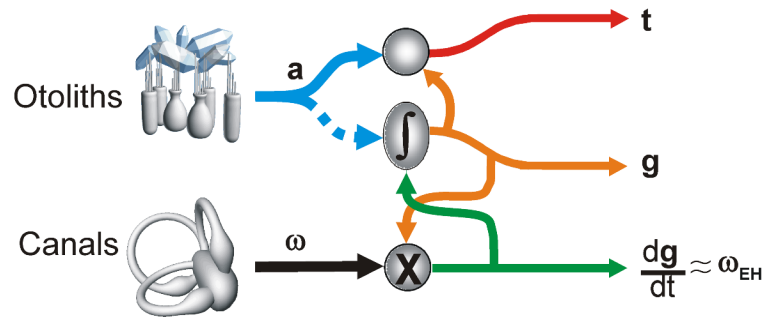
Fig. 2.
Schematic representation of the theoretical computations to solve the tilt/translation ambiguity. Head-centered angular velocity, $\boldsymbol{\omega}$ (e.g., from the canals) combines nonlinearly (multiplicatively) with a current estimate of gravitational acceleration (i.e., tilt) to compute the rate of change of the gravity vector relative to the head ($d\mathbf{g}/dt = -\boldsymbol{\omega}\times\mathbf{g}$ where "×" represents a vector cross product). For small amplitude rotations, $d\mathbf{g}/dt$ represents the earth-horizontal component of rotation, ω_{EH} (green). Integrating ($\int$) $d\mathbf{g}/dt$ and taking into account initial head orientation (e.g., from static otolith signals; dotted blue) yields an updated estimate of gravitational acceleration, $\mathbf{g}$ (orange; $\mathbf{g} = -\int\boldsymbol{\omega}\times\mathbf{g}\ dt$). This $\mathbf{g}$ estimate can be combined with the net acceleration signal, $\mathbf{a}$ (blue; from the otoliths) to calculate translational acceleration, $\mathbf{t}$ (red). This schematic is based on solving the equation $\mathbf{t}=\mathbf{a}-\int\boldsymbol{\omega}\times\mathbf{g}\ dt$ (e.g., see [22,23]).
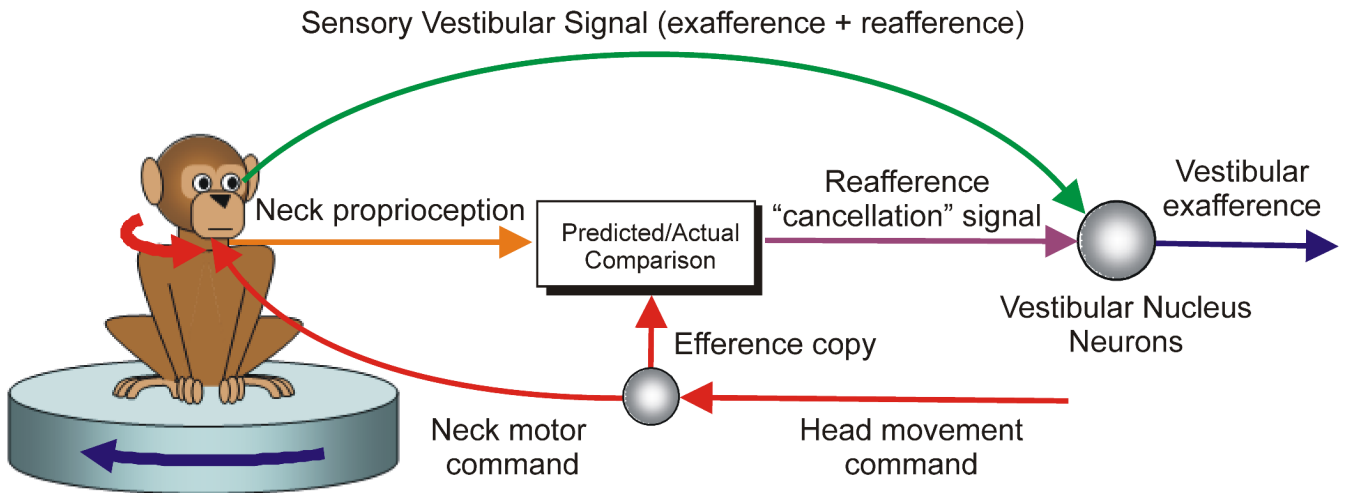
Fig. 3.
Schematic illustration of the proposal of Roy and Cullen [29] for how active and passive head movements could be distinguished in vestibular neurons. During an active head movement, an efference copy of the neck motor command (red) is used to compute the expected sensory consequences of that command. This predicted signal is compared with the actual sensory feedback from neck proprioceptors (orange). When the actual sensory signal matches the prediction it is interpreted as being due to an active head movement and is used to generate a reafference "cancellation" signal (purple; output of "Actual/Predicted Comparison" box) which selectively suppresses vestibular signals that arise from self-generated movements. By matching active head velocity with a simultaneous passive head rotation in the opposite direction, Roy and Cullen [29] eliminated most or all of the sensory vestibular contribution (green) to central activities during active head movement, unmasking for the first time the presence of this "cancellation" signal.
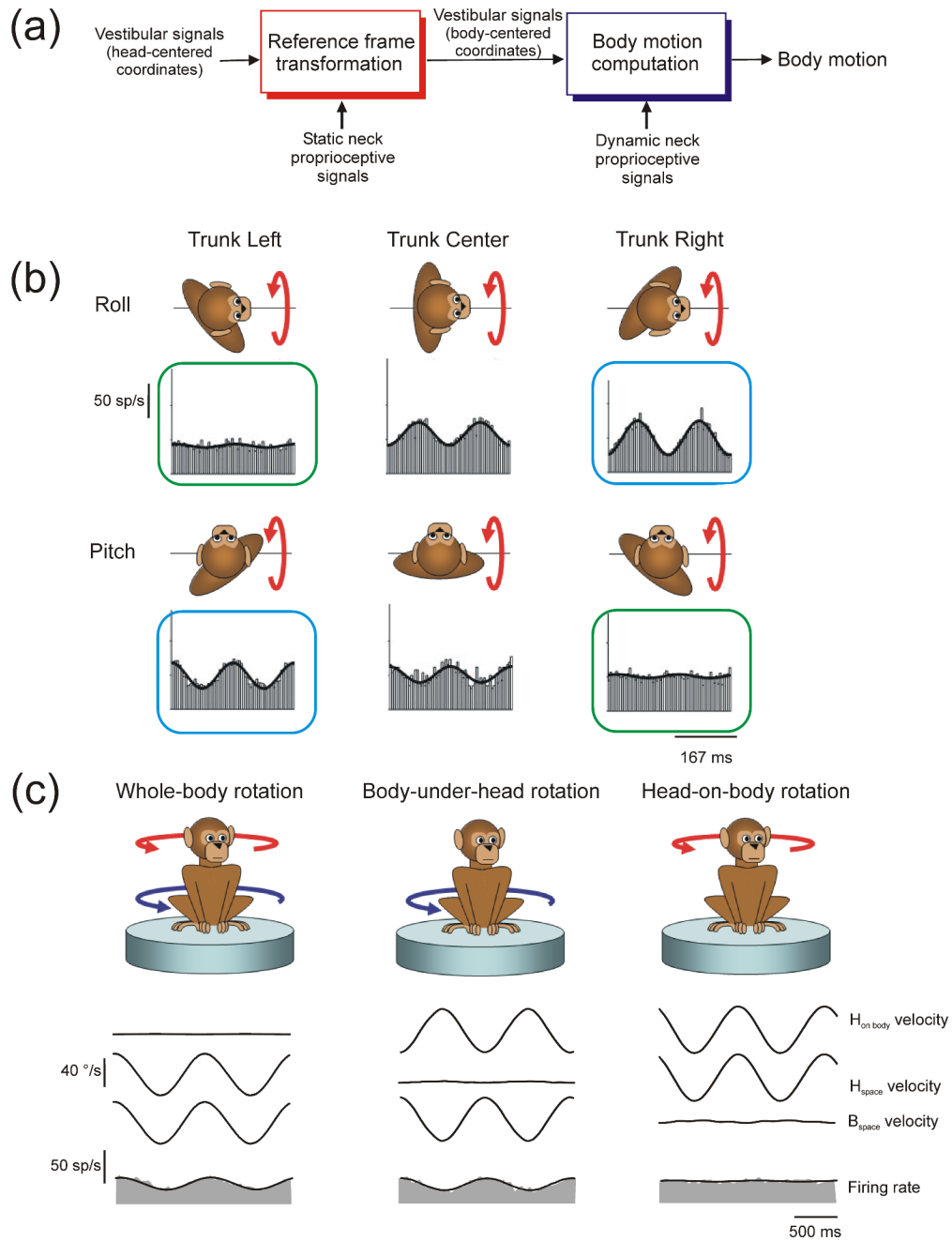
Fig. 4.
Computations to estimate body motion by combining vestibular and neck proprioceptive signals. (**a**) Two required computational steps: 1) "Reference frame transformation" (left) transforms head-centered vestibular estimates of motion into a body-centered reference frame; vestibular signals must be combined non-linearly (multiplicatively) with *static* proprioceptive estimates of head-on-body position. 2) "Body motion computation" (right) involves combining vestibular estimates of motion with *dynamic* proprioceptive signals to distinguish body motion from head motion with respect to the body. For descriptive purposes the two sets of computations are illustrated serially as distinct processing stages. However, both computations could occur in tandem within the same populations of neurons. (**b**) Reference frame experiment

in which head versus body-centered reference frames for encoding vestibular signals were dissociated by examining rostral fastigial neuron responses to pitch and roll rotation for different trunk re head orientations (i.e., trunk left, center and right). Blue and green boxes indicate rotations about common body-centered axes. Data replotted with permission from Kleine et al. [46] (**c**) Evidence for coding of body motion in the rostral fastigial nuclei [45**]. The example cell exhibited a robust response to body motion both during passive whole-body rotation that stimulated the semicircular canals (left) and during passive body-under-head rotation that stimulated neck proprioceptors, but did not respond to head-on-body rotation (right), illustrating that vestibular and priopriceptive signals combined appropriately to distinguish body motion. Data replotted with permission from Brooks and Cullen [45**].