Behavioral/Systems/Cognitive

Multiple Reference Frames for Motion in the Primate Cerebellum

Aasef G. Shaikh, Hui Meng, and Dora E. Angelaki
Department of Neurobiology, Washington University School of Medicine, St. Louis, Missouri 63110

Knowledge of body motion through space is necessary for spatial orientation, self-motion perception, and postural control. Yet, sensory afferent signals may not directly provide such information to the brain. Because motion detected by the vestibular end organs is encoded in a head-fixed frame of reference, a coordinate transformation is thus required to encode body motion. In this study, we investigated whether cerebellar motion-sensitive neurons encode the translation of the body through space. We systematically changed both the direction of motion relative to the body and the static orientation of the head relative to the trunk. The activities of motion-sensitive neurons in the most medial of the deep cerebellar nuclei, the rostral fastigial nucleus, were compared with those in the brainstem vestibular nuclei. We found a distributed representation of reference frames for motion in the rostral fastigial nucleus, in contrast to cells in the vestibular nuclei, which primarily encoded motion in a head-fixed reference frame. This differential representation of motion-related information implies potential differences in the functional roles of these areas.

Key words: self-motion; cerebellum; frame of reference; sensorimotor; multisensory; gain field; coordinate transformation

Introduction
The vestibular system forms an essential component in the sensing and generation of movement. Our inertial motion sensors, the otolith organs, located in the inner ear cavity, are highly conserved throughout evolution and function to provide an accurate moment-to-moment estimate of our motion through space. Being fixed in the head, the corresponding signals carried to the brain by primary vestibular afferents detect self-motion in a head-fixed coordinate system. However, a reference frame transformation might take place centrally because day-to-day functions often require knowledge of body position, orientation, and movement. For example, systematic alterations in vestibulo-spinal reflex properties after altered static orientations of head-on-trunk have been reported (Nashner and Wolfson, 1974; Kennedy and Inglis, 2002). In addition, information on body orientation and movement is also important for the perception of self-motion and location of objects in extra-personal space (Lackner and Graybiel, 1978; Mergner et al., 1991, 1992).

How does the brain acquire access to an internal estimate of body motion if this information is not provided by primary vestibular afferents? By combining otolith signals, which encode translation in head coordinates, with neck proprioceptive information that signals the position of the head relative to the body, a coordinate transformation could take place, whereby the brain constructs an internal estimate of body motion through space.

Extensive convergence of vestibular and somatosensory signals has been reported in motion-sensitive areas of the brainstem vestibular nuclei (VN), cerebellar cortex (e.g., anterior and posterior cerebellar vermis), and deep cerebellar nuclei (Boyle and Pompeiano, 1981; Anastasopoulos and Mergner, 1982; Wilson et al., 1990; Manzoni et al., 1998, 1999; Gdowski and McCrea, 1999, 2000; McCrea et al., 1999). At present, it is not known whether this convergence reflects an underlying reference frame transformation to compute body motion in space. Here, we specifically address one of the possible functions of these multisensory interactions. Results provide novel evidence that a population of neurons in the most medial of the deep cerebellar nuclei, the rostral fastigial nucleus (FN), encode the translation of the body through space. In contrast, motion-sensitive cells in the VN seem to maintain a head-fixed reference frame for encoding self-motion.

Materials and Methods
Single-unit activities in the rostral FN and rostral medial and lateral VN were recorded in two juvenile monkeys (Macaca Fascicularis), which were chronically implanted with an eye coil and a circular delrin ring for head stabilization. A guide tube delrin platform with a staggered array of holes (Angelaki and Dickman, 2000; Dickman and Angelaki, 2002) was stereotaxically secured to the skull inside the head ring. To provide better access to the medial cerebellar nuclei, the platform was slanted in the horizontal plane 10° from anterior to posterior and 10° from left to right. All surgical procedures were performed under sterile conditions in accordance with institutional and National Institutes of Health guidelines.

During experiments, the monkeys were seated in a primate chair with their bodies secured with multiple shoulder and lap belts while their extremities were tied to the chair. Their head ring was fixed to the primate chair through a 2.5 inch center bore-bearing piece that could be locked in place at different head-on-trunk orientations. The primate chair was then secured inside the inner frame of a vestibular turntable consisting of a three-dimensional rotator on top of a 2 m linear sled (Acutronics, Pittsburgh, PA). The linear acceleration was measured with a three-axis...
accelerometer mounted on the primate chair. Thus, as their trunk and extremities were softly restrained to the chair, the motion measured by the accelerometer reflected the motion of the body through space. We point out that in this presentation, body refers to the single block of torso that always faced forward relative to the primate chair.

For each recording session, the eye coil signals and the linear accelerometer output were low-pass filtered (200 Hz; 6-pole bessel), digitized at a rate of 833.33 Hz (model 1401, 16-bit resolution; Cambridge Electronics Design, Cambridge, UK), and stored for off-line analysis. Extracellular recordings from single FN and VN neurons were obtained with epoxy-coated tungsten microelectrodes inserted into 26-gauge guide tubes. Neural activity was amplified, filtered (300 Hz to 6 kHz), and passed through a BAK Instruments (Germantown, MD) dual-time-amplitude window discriminator. Single-unit spikes triggered acceptance pulses (BAK window discriminator) that were stored on a computer using the event channel of a Cambridge Electronics Design (model 1401) data acquisition system. In addition, eye movement signals, stimuli, and neural activity were stored on a digital audio tape recorder for off-line spike discrimination. Stimulus protocols and data acquisition were computer controlled using scripts written for the Spike2 (Cambridge Electronics Design) software environment.

During initial experiments in each animal, the abducens nuclei were first identified bilaterally. The recording sites in the FN and VN were then identified on the basis of their stereotaxic location relative to the abducens nuclei and the fourth ventricle. All neural responses were obtained from vestibular-only (VO) neurons. These neurons were recorded from rostral portions of the medial and lateral VN within 1 mm of areas where eye movement-sensitive cells were recorded (Angelaki et al., 2001; Dickman and Angelaki, 2002). In the FN, neurons were recorded in the rostral portions of the nucleus, anterior to all pursuit-related and saccade-related cells that we consistently encountered in the caudal FN (Gardner and Fuchs, 1975; Buttnet et al., 1991). To characterize cells as VO or eye movement-sensitive, a standard protocol was used, including rotations, fixations, and smooth pursuit eye movements (Angelaki et al., 2001). VN and FN cells that were sensitive to linear motion also were tested using different directions of translation (0, 30, 60, 90, 120, 150, and 180°; defined relative to the body) (see Fig. 1), with the head secured at three different positions relative to the trunk (see Fig. 1, monkey schematic drawing). Initially, the animal’s head was straight-ahead relative to the body (n = 0°) (see Fig. 1, blue drawing). Subsequently, using the specially constructed head holder, the head was secured 30° to the left relative to the animal’s body (h = 30°) (see Fig. 1, green drawing), and the six-direction translation protocol was repeated. Finally, the motion protocol was also delivered with the animal’s head secured 30° to the right (h = −30°) (see Fig. 1, red drawing). The order in which these three head-on-trunk positions were tested was pseudo-randomized for different neurons. Such reorientations of the head in relation to the body were intended to dissociate between body and head coordinates. Data were collected during sinusoidal translational motion stimulation either at 0.5 or 2 Hz (±0.2 g, with g = 9.81 m/sec²), whichever frequency provided the largest cell modulation. Results were similar at the two frequencies and thus are presented together.

Data analyses were performed off-line using custom-written scripts in Matlab (MathWorks, Natick, MA). For each action potential, an occurrence time was logged. Instantaneous frequency (IFR) was then calculated as 1/interspike interval and assigned to the middle of the interval. For each stimulus cycle (e.g., n cycle), an integral (n = 1) times the period was subtracted from the timing for all of the instantaneous frequency values for that specific cycle. The result was to “fold” all instantaneous frequency values into a single stimulus cycle. This procedure provides no averaging, because all spike occurrences are represented in time. Neural response amplitude and phase were then determined by fitting a sine function to both response and stimulus using a nonlinear least squares (Levenberg–Marquardt) minimization algorithm. Portions of the response cycle in which no spikes occur (silencing or rectification) does not pose a problem to the fitting procedure, because only positive IFR values exist in the response cycle and are weighted in the least-squares estimations. Examples of such fits have been illustrated for two cells in Figure 1. Neural response gains were expressed as spikes/second/gravity. Phase was the difference (in degrees) between peak neural activity and peak linear acceleration.

Neural response gains and phases were subsequently plotted as a function of the direction of translation and fitted (Levenberg–Marquardt optimization method) with a two-dimensional spatiotemporal model that represents a more generalized model than cosine tuning (Angelaki and Dickman, 2000). This fitting procedure allowed the estimation of four parameters: direction, gain, and phase of the maximum response as well as the minimum response gain (Angelaki, 1991; Angelaki and Dickman, 2000). The goodness of fit was based on the variance-accounted-for (VAF), computed as VAF = 1 − [var (model − data)/var (data)]. The fitting procedure was first applied to each head-on-trunk position separately (analysis step I). This allowed estimation of all four response parameters, separately for each head-on-tree head-on-trunk position. This first step in analysis was necessary to investigate which, if any, spatial tuning properties changed as a function of head-on-body position. Once it was established that the only systematic difference was a spatial shift of the tuning curves, data in all three head-on-trunk positions were also fitted simultaneously by assuming that the maximum response orientation for each cell was shifted through an angle, ±30°, for the rotated head-on-trunk positions (analysis steps II and III). Thus, in a second stage of analysis (step II), all data for each cell were fitted simultaneously with each of two four-parameter models corresponding to the body-fixed reference frame model (λ = 0; i.e., no spatial shift of maximum response direction for different head-on-trunk positions) or the head-fixed reference frame model (λ = 1; i.e., a ±30° systematic shift in the maximum response direction for different head-on-trunk positions). Finally, because it is possible that cells exhibit an in-between behavior, a third fitting procedure was used, where λ was allowed to vary and was one of the parameters fitted (analysis step III). This latter λ-variable model had a total of five parameters. A sequential F test (Draper and Smith, 1998) was used to ascertain the statistical significance of increasing the model parameters by comparing the errors associated with the λ variable and each of the two λ-fixed models while allowing for differences in their degrees of freedom. The sequential F test was the statistical test used to quantify, on a cell-by-cell basis, whether the tuning of each neuron was best described by the head-fixed/body-fixed or an intermediate reference frame model. For both the λ-fixed and λ-variable models (analysis steps II and III), all other parameters of cell tuning (e.g., maximum and minimum response gain and phase) were independent of head-on-trunk position.

Results

The head and body coordinate systems were dissociated by systematically varying both the direction of motion and the static orientation of the head relative to the body (Fig. 1). Neural activities were recorded as animals were translated along different directions (0, 30, 60, 90, 120, 150, and 180°; defined relative to their body) in the horizontal plane with the head fixed to the trunk in one of three different positions: straight-ahead (Fig. 1, blue traces) (h = 0°) as well as rotated 30° to the left (h = 30°) or right (h = −30°) (Fig. 1, green and red traces, respectively). Accordingly, if a neuron encodes the motion of the body, its spatial tuning should be independent of the change in head position. In contrast, if a cell exclusively detects the motion of the body independently of body motion, its preferred movement direction should systematically shift to the left or right to reflect the shifted direction of motion in head coordinates (Fig. 1, monkey head orientation drawing). Neurons with properties reflecting either the body or the head coordinate systems were encountered in the rostral fastigial nucleus, as illustrated for the two cells in Figure 1A (body coordinates) and Figure 1B (head coordinates). For the neuron illustrated in Figure 1A, the directions of maximum and minimum neural response modulation (0 and 90° motion directions, respectively) were the same for all three head-on-body positions. In contrast, for the neuron in Figure 1B, the directions of maximum and minimum response modulation shifted for the
three different head-on-body positions, such that they would remain fixed relative to the head. This shift is better seen for the minimum response direction, which shifts from 90° (head-on-body position) to 120° (h = 30°) or 60° (h = −30°) head-on-trunk positions. Superimposed solid lines represent best-fit sine functions. 

To describe these differences, neural response gains and phases were plotted as a function of motion direction separately for each of the three head-on-body positions (Fig. 2). The spatial tuning curves for the three head-on-body positions superimpose for the neuron shown in Figure 2A, suggesting that this cell encodes the translational motion of the body (i.e., encoding linear acceleration in body coordinates). In contrast, the spatial tuning for the neuron in Figure 2B shifts for the three head-on-body positions, indicating that this cell encodes the translational motion of the head. To quantify these observations, a spatiotemporal tuning model was used to fit neural gain and phase as a function of motion direction (Angelaki, 1991; Angelaki and Dickman, 2000). Three different analysis steps were performed. First, the spatiotemporal model was fitted separately to data for each head-on-trunk position (analysis step I). This allowed an independent estimation of the gain and phase of neural firing in response to motion along the maximum response direction of the neuron for each head-on-trunk position. Figure 3A illustrates these values for each of the two rotated head-on-body positions (h = ±30°) plotted versus the corresponding values when the head was straight-ahead relative to the body (h = 0°). There was no systematic dependence of any of these parameters on static neck inputs, as illustrated by linear regression (Fig. 3A, solid lines; dashed lines illustrate 95% confidence intervals). The main effect of changing the head-on-body position was a systematic shift in the spatial tuning curves, including the maximum and minimum response directions. Figure 3B illustrates the mean spatial tuning curve shift, plotted separately for the cerebellar and brainstem populations (gray squares vs open circles, respectively). Only the VN data were consistent with the spatial shift expected from a head-fixed reference frame coding of motion. The difference in the spatial shift of the tuning curves for FN and VN neurons was statistically significant (F(1,31) = 11.4; p < 0.01).

Because other parameters of cell firing rate properties remained independent of changes in the head-on-trunk position (including the maximum gain of the cells), the spatiotemporal model was next used to fit the neural gain and phase dependence on motion direction for all three head-on-trunk positions simultaneously. According to the population behaviors in Figure 3B, the maximum sensitivity direction of each cell was assumed to shift through an angle ±30° for the rotated head-on-trunk positions. Thus, in a second analysis procedure (step II; see Materials and Methods), the value of λ was fixed to either λ = 0 (body-fixed reference frame model; e.g., no changes in the direction of maximum response gain as a function of head-on-trunk orientation) or λ = 1 (head-fixed reference frame model; i.e., allowing for a ±30° shift in the maximum response direction as a function of head-on-trunk position). The VAF values for each of the head-fixed and body-fixed reference frame models were compared for each cell in Figure 4A. In the figure, each symbol corresponds to one cell in either the VN (open circles) or the FN (gray squares), and the dotted line illustrates the unity–slope line. For all but one of the VN neurons, the head-fixed reference frame model gave as good or higher VAF values than the corresponding body-fixed reference frame model (Fig. 4A; all but one of the open circles fall in the top left half of the diagram). In contrast, only approximately one-third of the FN neurons was better fitted with the head-fixed reference frame model. The majority of the FN cells better complied with the body-fixed reference frame model. However, many neurons were equally well fitted by either model, illustrating the need for the third analysis procedure.

Thus, to further investigate the reference frame used by FN and VN neurons to encode motion, an additional analysis was performed where the value of λ was allowed to vary and was actually one of the parameters fitted for data of each cell (analysis step III; see Materials and Methods). This resulted in a model with one additional parameter. To quantify whether this higher parameter model fitted the data better compared with the lower parameter head-fixed or body-fixed model (step II), a sequential
The distributions for the spatial shift parameter, suggesting an intermediate reference frame. The difference in the head-fixed nor the body-fixed coordinate system predictions, and five VN cells) did not favor either one of the head-fixed reference frame. The remaining neurons (eight FN neurons and a subpopulation of FN cells encode motion in a body reference frame. The rostral fastigial nucleus represents the main output of the medial zone of the anterior vermis (Voogd, 1989), the role of which in vestibular–somatosensory interactions has received strong experimental support (Manzoni et al., 1998, 1999). In fact, Manzoni et al. (1999) have previously reported that static neck input modulates the responses of anterior vermis Purkinje cells during complex vestibular stimulation in decerebrate cats. The present results show that these vestibular–somatosensory interactions implement a coordinate transformation to estimate motion of the body through space. The outcome of the reference frame change is observed in the firing rates of a subpopulation of FN neurons, an area that represents a main cerebellar outflow to the spinal cord, premotor brainstem centers, and the thalamocortical system.

Vestibular–somatosensory convergence

The concept of vestibular–proprioceptive interactions was originally introduced by von Holst and Mittelstaedt (1950). Several lines of evidence have supported this idea. First and foremost, control of muscular coordination in relationship to movement is crucial for maintaining posture. Vestibular–proprioceptive interactions have been reported for limb control (Manzoni et al., 1983; Ezure and Wilson, 1984; Wilson et al., 1986) and neck reflexes (Peterson et al., 1985). In addition, sensations of tilting and falling, as well as altered perception of spatial orientation can be induced by electrical stimulation or vibration of the neck (Karnath et al., 1994; Ivanenko et al., 1999). Additional psychophysical experiments using trunk rotation relative to the head have demonstrated that normal human subjects derive trunk motion perception from a combination of vestibular and neck cues (Mergner et al., 1983, 1991). In addition, vibration of muscles or muscle tendons causes proprioceptive “misinformation” producing an illusory sensation of movement (Goodwin et al., 1972). In normal subjects, vibration applied to the neck muscles causes a displacement of body orientation, a shift of subjective midline, and a consistent error in pointing and body sway (Biguer et al., 1988; Pyykkö et al., 1989; Karnath et al., 1994).

The interaction between vestibular and joint afferents already occurs at the level of the vestibular nuclei, the first central synapse of primary vestibular afferents (Boyle and Pompeiano, 1981; Anastasopoulos and Mergner, 1982; Wilson et al., 1990). Many VN neurons respond to both vestibular and proprioceptive stimulation (Brink et al., 1980; Kasper et al., 1988). Given such extensive vestibular–proprioceptive convergence in the VN and the fact that the paramedian zone of the anterior vermis projects directly to the lateral vestibular nucleus (Voogd, 1989), the observation that the VN cell population described here used a motion reference frame that was not body-fixed might appear surprising. Several studies have demonstrated recently that there are differences in the VN neuron responses during rotations in which...
the head was fixed relative to the body versus rotations where the head was allowed to rotate freely (Gdowski and McCrea, 1999; McCrea et al., 1999; Roy and Cullen, 2001). Yet, the majority of VN neurons were closer to head than body coordinates. It is possible that if a clear body reference frame exists in the VN for rotational or translational movements, the location of such neurons might be in more posterior areas than those explored here.

Role of fastigial neurons in motion detection

The present experiments, reporting a specific and systematic spatial shift in the tuning curves (without gain or phase changes), represent the first evidence regarding the existence of a body reference frame for translational motion in a subpopulation of FN neurons. Extensive vestibular-proprioceptive interactions that might underlie reference frame transformations also have been reported in rostral fastigial neurons during rotational motion (Buttner et al., 2003). The rostral fastigial nucleus, among all deep cerebellar nuclei, is the one most important in the central processing of motion-related vestibular information. Anatomical and physiological studies have separated the fastigial nucleus into rostral and caudal subdivisions (Noda et al., 1990; Buttner et al., 1991). Neurons of the caudal fastigial nucleus are modulated during saccadic or smooth pursuit eye movements (Gardner and Fuchs, 1975; Buttner et al., 1991; Fuchs et al., 1993). In contrast, the rostral division is populated by neurons that are modulated by vestibular stimulation but do not show sensitivity to eye movements (Gardner and Fuchs, 1975; Buttner et al., 1991; Siebold et al., 1997). Previous studies established that rostral FN neurons respond to sinusoidal rotation in horizontal and vertical planes (Gardner and Fuchs, 1975; Siebold et al., 1997, 1999) as well as to translational movements (Zhou et al., 2001). Given the afferent and efferent projections of rostral FN neurons, it is believed that they are involved in vestibulospinal control, including regulation of gait and postural mechanisms (Gardner and Fuchs, 1975; Buttner et al., 1991; Siebold et al., 1997).

Knowledge of the orientation and motion of the body in space is important for the control of body, limb, and neck posture as well as motion perception and spatial localization of objects in extra-personal space (Mergner et al., 1991, 1992; Horak et al., 1994). All of these functions could be served by projections of FN neurons. For example, the main fastigial efferent projections are
to spinal-related brainstem structures as well as direct projections to the spinal cord, where they directly synapse on spinal motoneurons (Batton et al., 1977; Fukushima et al., 1977; Wilson et al., 1977). The activities of limb musculature for postural responses during movement have been shown to correlate with body, but not head, displacement as required for postural stability, suggesting that vestibulospinal reflexes might be organized in body coordinates (Manzoni et al., 1998). The same study also showed that intact activity in the anterior vermis is necessary for this function. The spatially shifted reference frames characterizing the tuning of FN neurons might represent a neural substrate for this processing. In addition to spinal projections, the FN nucleus, similar to all other cerebellar nuclei, projects to the thalamus (Batton et al., 1977; Nakano et al., 1980; Asanuma et al., 1983). Although the cortical targets of the FN remain unknown, some studies have suggested FN projections to parietal cortex (Sasaki et al., 1976; Amino et al., 2001) as well as to frontal cortex regions and multimodal visual areas (Kiyohou and Kawaguchi, 1987). Fastigial-thalamocortical projections could thus mediate spatial perception functions in which knowledge of body motion in space is essential.

References


Fuchs AF, Robinson FR, Straube A (1993) Role of the caudal fastigial nu-