scientific reports

OPEN

Natural statistics of human head orientation constrain models of vestibular processing

Christian B. Sinnott[1✉], Peter A. Hausamann[2] & Paul R. MacNeilage[1]

Head orientation relative to gravity determines how gravity-dependent environmental structure is sampled by the visual system, as well as how gravity itself is sampled by the vestibular system. Therefore, both visual and vestibular sensory processing should be shaped by the statistics of head orientation relative to gravity. Here we report the statistics of human head orientation during unconstrained natural activities in humans for the first time, and we explore implications for models of vestibular processing. We find that the distribution of head pitch is more variable than head roll and that the head pitch distribution is asymmetrical with an over-representation of downward head pitch, consistent with ground-looking behavior. We further suggest that pitch and roll distributions can be used as empirical priors in a Bayesian framework to explain previously measured biases in perception of both roll and pitch. Gravitational and inertial acceleration stimulate the otoliths in an equivalent manner, so we also analyze the dynamics of human head orientation to better understand how knowledge of these dynamics can constrain solutions to the problem of gravitoinertial ambiguity. Gravitational acceleration dominates at low frequencies and inertial acceleration dominates at higher frequencies. The change in relative power of gravitational and inertial components as a function of frequency places empirical constraints on dynamic models of vestibular processing, including both frequency segregation and probabilistic internal model accounts. We conclude with a discussion of methodological considerations and scientific and applied domains that will benefit from continued measurement and analysis of natural head movements moving forward.

Head orientation relative to gravity shapes statistical regularity of sensory stimulation across modalities[1]. More specifically, there is gravity-dependent structure in the sensory environment, and the orientation of the head relative to gravity determines how this structure is sampled by head-fixed sensory organs like the eyes, ears, and vestibular organs. Statistical regularity in sensory stimulation is therefore jointly mediated by regularity in head orientation and regularity in gravity-dependent environmental structure, and the combination of these two factors shapes sensory processing. At the encoding stage, more resources are devoted to encoding most common stimuli, and at the decoding stage, prior distributions across encountered stimuli shape how behaviorally and perceptually relevant states are estimated[2,3]. Thus, it is necessary to characterize statistical regularity in head orientation in order to better understand how this regularity shapes sensory encoding and decoding in the nervous system.

Prior research that comprehensively characterizes how the head is oriented relative to gravity during everyday activities is currently lacking in the literature. Among existing studies of natural head movement, the statistics of head orientation relative to gravity seems to be neglected[1]. Instead, previous work has characterized vestibular stimulation, that is the linear acceleration and angular velocity experienced by the head during natural tasks in humans[4,5] as well as in non-human primates and rodents[6]. One very recent study does report statistics of head roll relative to gravity, but only during prescribed activities[7]. Similarly, one study reports statistics of head pitch during three tasks performed for 5 min each[8]. To our knowledge, no studies report statistics of head orientation relative to gravity in freely-acting participants.

One reason that head orientation relative to gravity is seldom reported is that it is difficult to measure during unconstrained natural behavior. Previous methods to track angular head position are either confined to the lab, or they are unable to provide robust positional measurement due to gravitoinertial ambiguity[1,9,10]. A recent solution which allows unconstrained, positional tracking is simultaneous localization and mapping (SLAM). SLAM and variants like visual-inertial simultaneous localization and mapping (VI-SLAM) fuses information

[1]Department of Psychology, University of Nevada, Reno 89557, USA. [2]Department of Electrical and Computer Engineering, Technical University of Munich, 80333 Munich, Germany. ✉email: csinnott@nevada.unr.edu

from multiple sensors such as cameras and inertial measurement units (IMUs), allowing for robust positional tracking. SLAM and VI-SLAM have recently been used to track head movement during natural behaviors[10–12]. Here we use a commercial VI-SLAM solution, the Intel RealSense T265 (T265), which has been validated against an optical tracking volume to track human head movement[10].

We use the T265 to measure statistics of natural head orientation during unconstrained natural activities over 5 h of continuous recording. We then use these measures to evaluate a Bayesian model of vestibular sensory processing and perception[7,13–15]. In particular, we seek to explain observed, predominantly attractive biases in perception of head and body orientation[14,16]. Previous efforts to model these biases have used Bayesian frameworks where statistical regularities can influence perception via Bayesian priors. These priors are typically modeled as Gaussian, with the variability of the prior left as a free parameter[7,14]. In the current study we explore using empirical measures of the statistics of head pitch and roll to determine priors in the Bayesian model.

In addition, we address a longstanding question in the vestibular literature regarding tilt-translation processing. The otolith organs respond in equivalent fashion to gravitational and inertial acceleration[17,18]. Resolution of this gravito-inertial ambiguity underlies not only spatial orientation perception, but also low-level behaviors such as reflexive eye movements[19,20] and postural control[21,22], which require distinct responses depending on whether stimulation is induced by tilt of the head on the body or translation of the head through the world.

One simple method that has been proposed to distinguish tilt from translation is frequency segregation[23,24], whereby lower frequency stimuli are interpreted as tilt and higher frequency stimuli are interpreted as translation. Consistent with the notion of frequency-dependent processing, recent neurophysiological work has shown that regular and irregular otolith afferents transmit more information about low and high-frequency naturalistic stimuli, respectively[22]. While this does not mean that regular and irregular afferents exclusively process tilt and translation, respectively, it does demonstrate that neural coding is frequency dependent and it underscores the importance of characterizing natural variations in stimulation frequency. Even internal model accounts of tilt-translation processing, which are typically proposed as an alternative to simple frequency segregation, should be shaped by the frequency content of naturalistic inertial stimuli in a statistically optimal fashion. However, to date, empirical data about the relative power of tilt and translation as a function of frequency during natural behavior has been lacking. For the first time, we evaluate relative power (i.e. power spectra) of tilt and translation components during natural everyday behavior and identify a crossing-point (a central parameter of several models) below and above which tilt and translation components have most power, respectively, thereby placing empirical constraints on models of vestibular processing. Overall, this work demonstrates the value of measuring statistics of behavior and stimulation to constrain models of sensory processing.

Methods

Ten participants (6 male, 3 female, 1 nonbinary identifying) were recruited to participate in this study. Data collection and all other procedures were approved by the Institutional Review Board at the University of Nevada, Reno and carried out in accordance with relevant guidelines and regulations. Researchers obtained separate, written informed consent from participants for use of participant likenesses or identifying images in figures in the current study (e.g. Fig. 1). This separate informed consent process was also approved by the Institutional Review Board at the University of Nevada, Reno. We asked participants to wear the T265 over a continuous 5-h recording period and to engage in natural behaviors that they were comfortable to have recorded. We also instructed participants to avoid activities that would compromise or be impeded by the recording equipment (e.g. high-impact sports), and to periodically re-calibrate the T265 every 30 min by slowly nodding and shaking their head five times (see "Calibration" section). All participants gave verbal informed consent in-line with data collection procedures approved by the Institutional Review Board at the University of Nevada, Reno.

Figure 1. Recording equipment. The T265 measures 6 degree of freedom (6DOF) position of the camera relative to the world, and is worn on the head by a participant over 5 h. The camera plugs into a lightweight laptop worn in a backpack.

Hardware. Recording equipment consisted of the T265 tracking camera which was worn by participants on their head using a commercially available elastic strap harness (Fig. 1). This off-the-shelf tracking camera uses VI-SLAM to estimate position of the camera relative to the world. It contains multiple sensors including two global shutter fisheye cameras, an accelerometer, and a gyroscope. The cameras sample video at 30 Hz, with an individual resolution of 848 × 800 pixels and a combined diagonal field of view of 173°, while the gyroscope records at 200 Hz and the accelerometer at 62.5 Hz. The visual data from the cameras and the inertial data collected from the gyroscope and accelerometer are processed on an internal board with a proprietary VI-SLAM algorithm, yielding 6DOF odometry at 200 Hz. The current study uses odometry data down-sampled to 62.5 Hz, matching the sampling rate of the accelerometer to allow disambiguation of acceleration driven by gravity and self-motion. We also sub-sampled video at 1/60 Hz (1 frame per minute), in order to annotate activities for a separate experiment. This tracking camera has been validated for use in tracking human head movement[10].

The T265 was connected via USB cable to a lightweight, portable laptop (Dell Latitude 5300). The laptop had an 8th generation Intel i7 CPU and 16 GB of DDR4 RAM. The laptop was carried by the participant in a backpack. Data acquisition software was programmed in Python 3.

Calibration. To transform data from camera to head coordinates we measured the offset between the two reference frames at the beginning and end of recording. We first recorded a 15 s segment in which participants held their head in a static position such that Reid's baseline, an anatomical reference line based on the external meatus of the ear and canthus of the eye, was perpendicular to the direction of gravity. Then, participants slowly nodded their head up-and-down five times (pitch movement) and shook their head left-to-right five times (yaw movement). In addition, participants performed periodic re-calibrations roughly every 30 min consisting of only the nodding and shaking segments. This was to measure and correct for any slippage of the device during the recording period.

The calibration movements are used to calculate the offset of camera and head reference frames. First, we solve for the rotation matrix that aligns pitch and yaw angular velocities measured during the calibration period with the horizontal and vertical axes, respectively. Next, we calculate the rotation about the pitch axis that best aligns the horizontal plane with a plane parallel to Reid's baseline, where said plane is defined as the plane perpendicular to the average direction of gravity during the Reid's baseline segment of the calibration. Data is ultimately expressed in this kinematically and anatomically-defined head-centered reference frame that is consistent across participants.

Pre-processing. We performed manual and automated pre-processing on our data prior to further analysis. This was done to remove artifacts driven by epochs where tracking failure occurred, to remove segments used for calibration, and to parse data into separate "low" and "high" velocity categories. Time points of the calibration segments at the beginning and end of the recording period were manually annotated. After the experiment, we visually inspected the angular and linear velocity time-series data to identify re-calibration segments that occurred approximately every 30 min as well as epochs with clear tracking failure.

Additionally, exclusion was conducted based on the confidence metric generated by the T265 during motion tracking. The metric ranges from 0 (poor) to 3 (good) and all positional data with a confidence of less than 3 was excluded from further analysis. Data excluded in this way often consisted of momentary tracking failures. Finally, we noticed loss of tracking can lead to artifacts with momentarily high linear velocity estimates. Therefore, all data frames with estimated linear velocity above a threshold of 3.05 meters per second were also excluded. This threshold was selected based on reported average jogging speed among a sample of young adults[25].

Finally, data were parsed into separate low and high linear velocity categories. This was done in order to observe potential differences in head orientation during more stationary (e.g. during static standing, sitting, or lying positions) versus more mobile epochs (e.g. during locomotion). We chose a cutoff value of 0.75 m/s by calculating the linear velocity norm observed during a subset of selected stationary and walking epochs, and selecting the value that best split the difference between the typical value observed during stationary activities (approximately 0 m/s) and the typical value observed during walking (approximately 1-1.5 m/s). All data with an instantaneous velocity norm value below this cutoff were parsed as low velocity, while all data with an instantaneous velocity norm value equal to or above this threshold were parsed as high velocity.

We do not perform any further pre-processing of time-series data such as smoothing or filtering. We assume that the T265 performs some sort of smoothing and/or filtering on the positional data it provides, but we are unable to verify this as the T265 VI-SLAM algorithm is proprietary. Furthermore, previous work verifying use of the T265 for estimation of human head motion did not filter or smooth data given by the T265 prior to comparison against gold-standard optical tracking positional estimates[10]. After all pre-processing, 41.95 h of data were used for further processing. Of these 41.95 h of data, 38.7 h of data (92.25%) were designated as low velocity while 3.19 (7.75%) were designated as high velocity.

Modeling perception of head orientation. We used a simple, static Bayesian modeling framework to investigate how empirically observed distributions of head roll and pitch might shape biases in perception of pitch and roll[7,13–15]. An introduction to this approach is presented in the Supplemental Material section titled "Bayesian model of perception of head orientation." Previous research has observed "attractive" bias for head roll and pitch perception at most eccentricities[14,16]. This bias is termed "attractive" in a Bayesian framework, because perceptual estimates are attracted towards the mean of the prior distribution which is assumed to have maximal probability at upright. Functionally, this attractive bias towards the prior results in perceptual underestimation relative to the true stimulus value.

To generate model predictions of head roll and pitch perception bias, we generate kernel density estimates (KDEs) from distributions of head roll and pitch orientations across all participants, with kernel bandwidth determined by Scott's Rule (Supplementary Eq. S7). These KDEs are used as prior probability distributions in our Bayesian model, and multiplied with likelihood distributions at every whole degree value in a range of $+/-120°$ for head roll and $+/-90°$ for head pitch. These ranges are selected because they correspond to ranges of pitch and roll perception measured in previous psychophysical research[14,16], and we compare our model prediction with these results. Noise is applied to the likelihood functions as eccentricity increases either linearly (linear model, see Supplementary Eq. S4) or non-linearly (shear model, see Supplementary Eq. S5). The parameter that determines this multiplicative noise on the likelihood is the only free parameter of the model. The mean of the resultant posterior distribution is used as a point estimate for perception of a given roll or pitch angle. The value for the single free parameter that yielded the best-fitting model in each case was found by minimizing the residual standard error (RSE):

$$RSE = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{y}_1)^2} \tag{1}$$

In order to gain a better understanding of how the shape of the prior impacts patterns of bias in head orientation, the modeling described above was then repeated for a variety of both empirical and parametric prior distributions. For empirical priors, progressively more smoothed versions were generated by increasing the bandwidth of the kernel used to generate the KDEs (Fig. S1). Results of this modeling are presented in the Supplemental Material. For parametric priors, we explored the effects of using symmetrical Gaussian priors for roll (Fig. 6a), allowing us to compare our results to previous similar studies[7,14]. We also explored the effects of using a skew normal distribution for pitch (Fig. 7a) in order to examine how well this shape of prior can explain asymmetrical perceptual biases observed for pitch. Results of modeling with parametric priors are presented below in the main text. Parameters for all priors as well as values for model free parameters and goodness of fit are summarized in Supplementary Tables S2 and S3.

Results

Descriptive statistics of head orientation distributions.

While roll and pitch are circular variables, neither roll nor pitch distributions were significantly wrapped (i.e. upside-down orientations were not observed, Fig. 2a), so we opted to use linear rather than circular descriptive statistics. Head roll (Fig. 2b) was centered close to zero and had lower variability (M = 0.5806°, SD = 6.2108°). Head pitch (Fig. 2c) was biased downward and had relatively high variability (M = -1.7701°, SD = 16.8167°). To more completely characterize these distributions, we also calculated higher moments, including skewness and excess kurtosis, using the stats.describe function in the Python 3 library scipy (formulae in Supplementary Table S1). Both roll ($\mu_3 = 0.1211$) and pitch ($\mu_3 = -0.009$) distributions had little skewness. Finally, the roll distribution had much greater excess kurtosis ($\mu_4 = 7.2592$) than the pitch distribution ($\mu_4 = 1.7467$).

We also calculate moments of distributions (Supplementary Table S1) and generate KDEs for head roll and pitch during low-velocity (<.75 m/s, Supplementary Fig. S7) and high-velocity ($\geq$.75 m/s, Supplementary Fig. S8) epochs. Generally, high-velocity head roll and pitch distributions had decreased excess kurtosis relative to that observed from data across all velocities. The high-velocity head pitch distribution also showed increased skewness and bias relative to the distribution generated from head pitch data across all velocities. Low-velocity head pitch and roll distributions differed very little from their analogues generated from data of all velocities, as low-velocity data made up the majority of the total dataset (92.25% of all data). Proportions of low and high-velocity data per participant are summarized in Supplementary Tables S5 and S6.
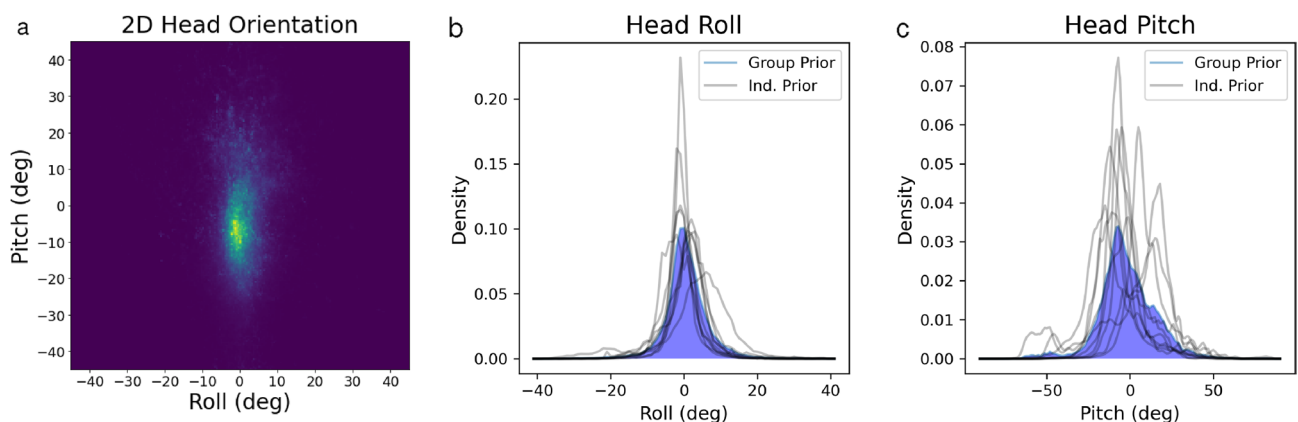

Figure 2. 2D head orientation measured across all participants (**a**). Marginal KDEs for roll (**b**) and pitch (**c**) are also plotted. The KDEs plotted in blue represent the distributions across all participants, while black traces represent KDEs for individual participants.

Power spectra for tilt and translation.   Power spectra of measured linear acceleration were calculated in order to compare relative power of the gravitational (tilt) and inertial (translation) components as a function of frequency. Power was calculated using the scipy.stats library in Python 3, with a sampling frequency (fs) of 62.5 and a fast Fourier transform length (nfft) of 512 samples. Along each axis, a crossing point was observed (Fig. 3). Below this critical frequency, gravitational acceleration exhibited the majority of power while above this frequency inertial acceleration exhibited greater power. These crossing points were observed at 1.148, 0.626, and 0.633 Hz for linear acceleration along the X- (nasal-occipital) (Fig. 3a), Y- (interaural) (Fig. 3b), and Z- (dorsal-ventral) (Fig. 3c) axes, respectively. We conducted similar analyses separately for low- (Supplementary Fig. S5) and high-velocity (Supplementary Fig. S6) epochs. Generally, crossing points were at higher frequencies across all axes for low-velocity epochs relative to all velocity epochs, while they were at slightly lower frequencies during high-velocity epochs. Furthermore, we observe a dominant peak at 2 Hz during high-velocity epochs along the Z-axis, likely driven by preferred stepping frequency during locomotion.

Modeling perception of head orientation.   Bayesian models are fit to psychophysical data by adjusting the value of one free parameter (σ), which represents the multiplicative noise on the likelihood. This multiplicative noise varies as either a linear (linear noise model, Supplementary Eq. S4) or sinusoidal (shear noise model, Supplementary Eq. S5) function of the presented tilt angle. We determined the value that provided the
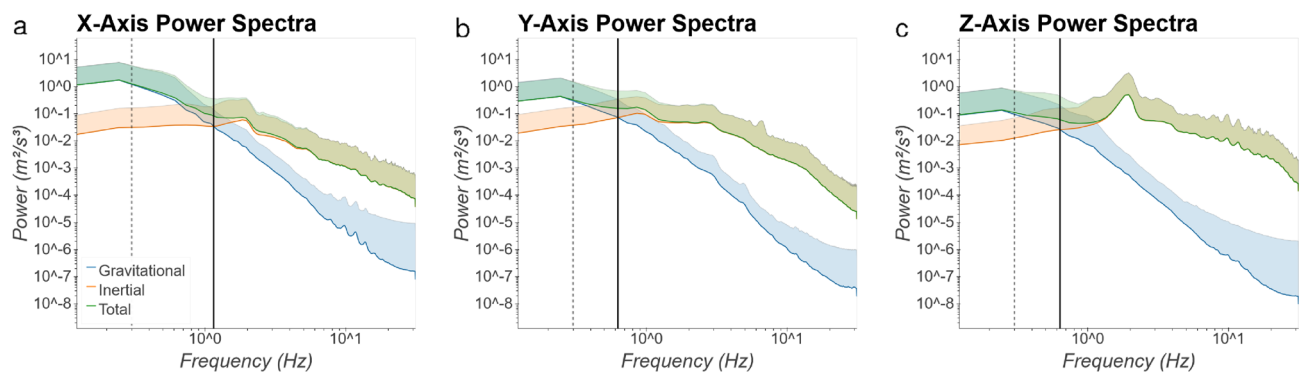

Figure 3. Power spectra for gravitational (blue), inertial (orange), and sum total (green) linear acceleration along X- (nasal-occipital), Y- (interaural), and Z- (dorsal-ventral) axes. Figure axes are log-scaled. Crossing points in power between gravitational and inertial acceleration are observed at 1.148, 0.626, and 0.633 Hz for X-, Y-, and Z-axes, respectively (black). Previously observed crossing points at 0.3 Hz from[24] are plotted for comparison (dashed grey).
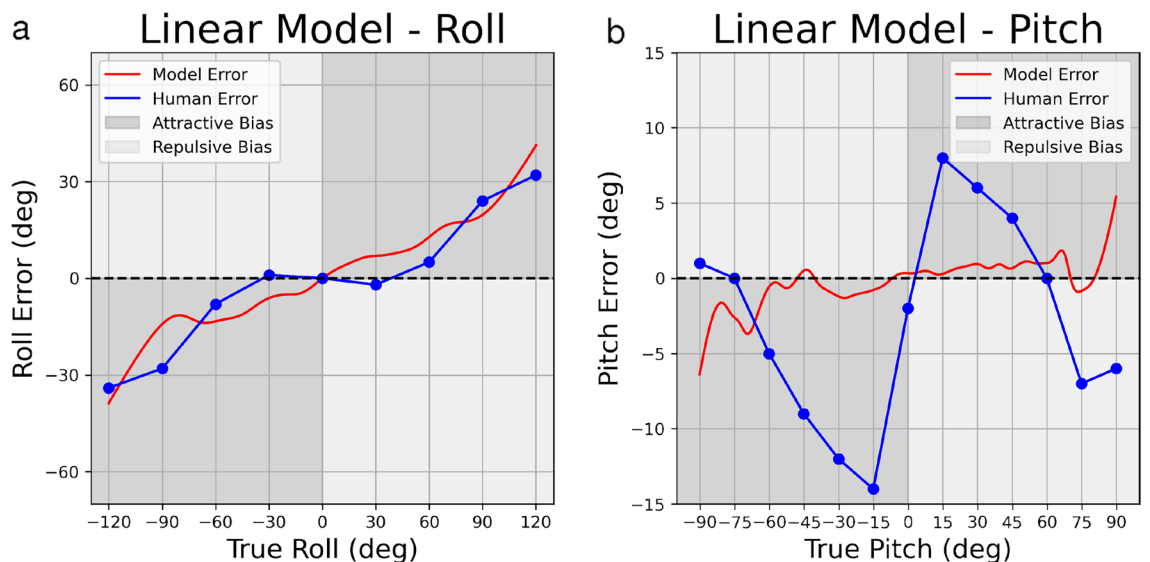
Figure 4. Psychophysical data (blue) and predictions of the linear noise models (red) for roll (**a**) and pitch (**b**). The presented stimulus orientation is plotted on the X-axis and estimation error (bias) is plotted on the Y-axis. Ticks and gridlines on the X-axis represent angles for which psychophysical data (blue) were previously reported: roll data is replotted from[14], pitch data is replotted from[16]. Data between gridlines are linearly interpolated.
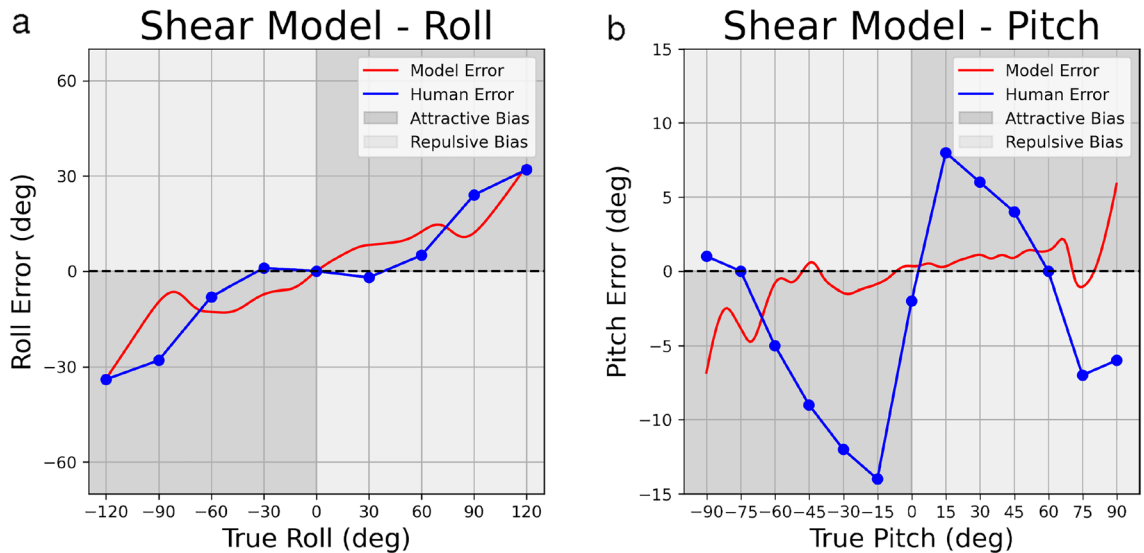
Figure 5. Utricular shear models for roll (5a) and pitch (5b). X- and Y- axes represent true orientation eccentricity and orientation eccentricity error, respectively. X-axis ticks and gridlines represent eccentricities where psychophysical data were reported in previous data, while psychophysical data between gridlines are linearly interpolated.
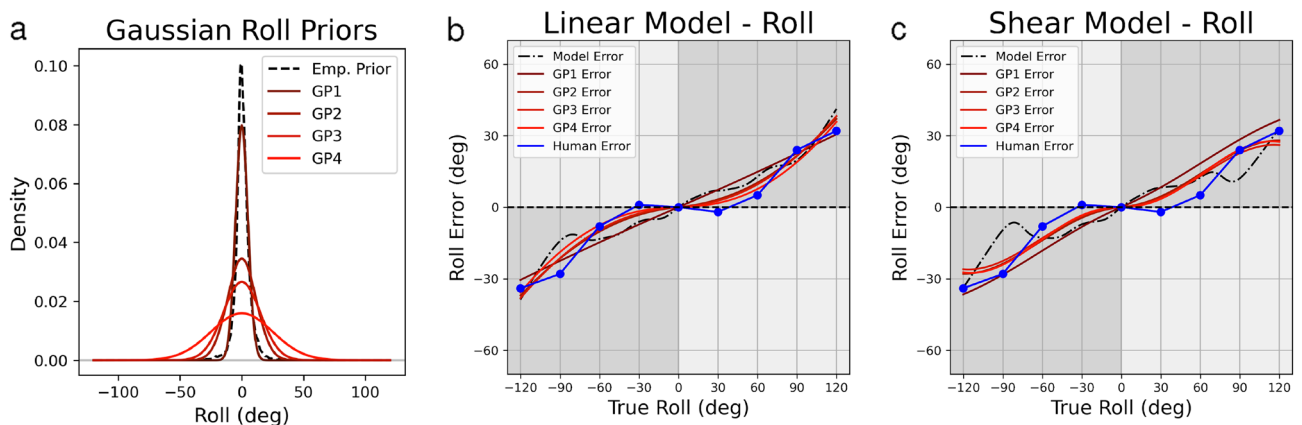
Figure 6. Head roll perception error modeling using Gaussian priors. Gaussian priors are plotted in (**a**) in red, while the empirical roll prior is plotted in dashed black. Improved performance is seen when using these as priors in the linear model (**b**) and the shear model (**c**). Model predictions using a given prior are plotted in dashed black and red, while psychophysical roll perception data from previous work are plotted in dark blue.

best model fit by selecting the sigma value that minimizes distance (RSE) between the predicted and observed perceptual error.

For roll perception, we compare model predictions to data from[14] and we achieve a fit with 6.435 RSE (with σ = 0.113) for the linear noise model (Fig. 4a), and 8.346 RSE (with σ = 0.013) for the shear noise model (Fig. 5a). While the fits are not perfect, the pattern of model bias roughly parallels the pattern of human perceptual bias, with attractive bias that tends to increase with increasing roll angle.

For pitch perception, we compare model predictions to data from[16] and we achieve a fit with 6.779 RSE (with σ = 0.001) for the linear noise model (Fig. 4b), and 6.795 RSE (with σ = 0.01) for the shear noise model (Fig. 5b). In contrast to results for roll perception, the best-fitting model for pitch perception cannot capture the observed pattern of human bias, particularly the repulsive biases observed at extreme pitch angles. The best-fitting model predicts minimal biases, achieved by setting low noise values for the likelihood.

In addition to fitting the model using the empirical priors shown in Fig. 2b,c, we generated several additional empirical and parametric priors to further investigate how the shape of the prior impacts patterns of perceptual bias. First we examined the effect of smoothing the empirical priors by increasing the bandwidth of the kernel used to generate the KDEs (Supplementary Fig. S1). This resulted in some slight smoothing of perceptual biases, but did not generally improve model fits (Supplementary Fig. S2, Table S2). These results are discussed in greater detail in Supplemental Material.

Next, we investigated modeling of roll perception using parametric Gaussian priors in comparison to empirical priors. All Gaussian priors were centered at zero, with each prior differing from others only by its variability
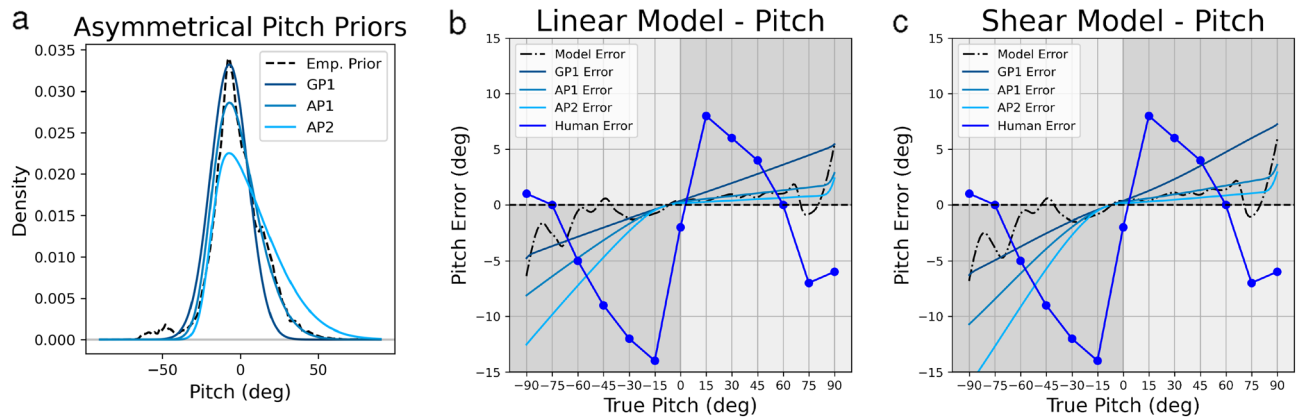
Figure 7. Head pitch perception error modeling using Gaussian and skew normal priors. Gaussian and skew normal priors are plotted in (**a**) in light blue, while the empirical roll prior is plotted in dashed black. Increasingly asymmetric predictions are seen when using these as priors in the linear model (**b**) and the shear model (**c**). Model predictions using a given prior are plotted in dashed black and light blue, while psychophysical roll perception data from previous work are plotted in dark blue.

(Fig. 6a, Supplementary Table S3). In both the linear and shear models, use of a Gaussian prior predicts observed roll perception bias better than our empirical prior (Fig. 6), consistent with results of Willemsen et al.[7]. Use of a Gaussian prior with variance equal to 15 in our linear model results in a fit with RSE of 3.016 (Fig. 6b), compared to our empirical prior with RSE of 6.438. We observed a similar trend in the shear model, as using a Gaussian prior with variance equal to 25 resulted in a fit with RSE = 4.157 (Fig. 6c), roughly twice as good as the fit achieved with our empirical prior (RSE = 8.346). Other Gaussian priors performed better than the empirical prior, though not as well as priors listed explicitly here (Supplementary Table S3).

Unlike head roll, head pitch perception is asymmetrical, with observed biases that are roughly twice as large for backward compared to forward pitch (Fig. 7, Human Error). We suspected that asymmetrical biases might be explained by asymmetry in the pitch prior, which is visible in the pitch KDE (Fig. 2c). However, an asymmetric pattern of bias was not predicted by the best-fitting model (Figs. 4b, 5b) due to the above-mentioned repulsive biases. Therefore, in order to demonstrate that an asymmetric prior predicts asymmetric biases, we modeled pitch perception using two different skew normal distributions, as well as a symmetric Gaussian distribution (Fig. 7a). All three distributions differed with respect to mean, standard deviation and skewness parameters (Supplementary Table S3). Functionally, these differences resulted in distributions roughly centered about a common mode observed in our empirical pitch prior.

As expected, the skew normal distributions predict asymmetric patterns of bias in the expected direction. Greater variability on forward compared to backward pitch, as observed in the empirical pitch prior, predicts biases that are greater for backward pitch than for forward pitch, consistent with human perceptual performance. This confirms that asymmetrical priors can predict asymmetric patterns of bias. While the direction of asymmetric bias is consistent, the pattern of predicted bias does not correspond well with observed pitch perception bias from previous research. This is due to the inability of our Bayesian model to predict the observed repulsive biases.

Discussion

In the present study, we measured head orientation relative to gravity during natural, everyday behaviors of ten participants over approximately 5 h each. The aim was to characterize distributions of head orientation and relate these via Bayesian modeling to known biases in perception of upright. Below we first discuss the nature of these distributions, then we discuss the outcome of modeling efforts, and finally we discuss the dynamics of head orientation and its relation to overall linear acceleration measured at the head. We finish by reviewing certain methodological considerations as well as future research directions.

Characteristics of head orientation distributions. Cumulative across-subject pitch and roll distributions were generally centered near upright. For individual subjects, there was considerable variability in the location of the peak for pitch (grey lines, Fig. 2c). This could reflect calibration variability, differences in activities that were sampled for each participant, or individual differences in biomechanics. We suspect individual biomechanical differences are mostly responsible because most datasets included a significant amount of stationary activity and we expect calibration noise was minimal (see Methodological Considerations). Peaks for individual roll distributions, on the other hand, exhibited much less variability across all subjects.

In terms of variability, the cumulative pitch distribution was much more variable than roll. This reflects biomechanical differences in these dimensions, namely greater range of head-on-body flexion and extension as well as full-body flexion and extension for pitch relative to roll. Additionally there is a greater prevalence of natural variations in pitch during natural locomotion and orienting behaviors. In particular, previous work has robustly demonstrated the need for humans to gaze downwards when walking in order to find stable future footholds and prevent falling or injury, particularly in challenging environments that place high demands on postural control

during locomotion[26–28]. To maintain downward gaze on footholds in the lower visual field, it follows that a persistent downward head pitch is needed. The high-velocity head pitch KDE, which captures primarily locomotion and other movement-based behaviors, supports this notion and shows increased downward bias and asymmetry relative to the low-velocity KDE generated from predominantly stationary behaviors.

This persistent downward head pitch is captured in the shape of the cumulative pitch distribution, which appears more asymmetrical than our observed roll distribution, and also more asymmetrical than distributions measured in previous work investigating naturalistic head pitch[8]. This asymmetry is not captured by the standard skewness metric, though patterns of bias predicted using our empirical pitch prior appeared qualitatively asymmetric.

The pitch and roll distributions also differed in their kurtosis. The roll distribution exhibited greater excess kurtosis than the pitch distribution, meaning that the roll distribution has relatively long tails; extreme roll values were observed, but these observations were relatively infrequent. This has implications for the modeling results discussed below.

Modeling perception of head orientation.
Perception of upright has previously been modeled in a Bayesian framework with a prior distribution for head-in-space orientation that is centered at a value of zero tilt relative to gravity[7,13–15,29]. In this previous work, prior and likelihood are typically modeled as Gaussian distributions with parameters determining the variability on these distributions left free. Here, we instead use empirically-measured distributions as priors, an approach that enables more rigorous testing of the Bayesian framework (e.g.[30]); it allows for non-Gaussian priors and eliminates prior variability as a free parameter. Indeed, empirically-measured distributions appear qualitatively different from Gaussian distributions, with pitch appearing asymmetric and roll exhibiting high excess kurtosis: characteristics that Gaussian distributions cannot capture. The central focus of the modeling work presented here is to explore how these differences in the shape of the prior impact predicted perceptual biases.

We fit two versions of a Bayesian model, one in which variability increases linearly with increasing tilt angle and one in which it increases with the sine of the angle, i.e. proportional to the shear force at the utricle[31,32]. The multiplicative factor for increasing variability on the likelihood is the only free parameter of these models. The additive factor is taken from previous work modeling roll perception[14]. The same additive factor was used for head pitch modeling because perception of head pitch has not been modeled previously in this manner. Modeling results did not reveal substantial differences in goodness of fit between the linear and shear models.

For roll perception, modeled biases increase with increasing roll angle, but exhibit a different trend than the measured bias. The measured bias remains near zero out to roughly 30°, then increases rapidly out to 90°; a trend known as the E-effect[33]. In contrast, the modeled bias (with both linear and non-linear noise) increases in a manner that is roughly constant with increasing tilt angle. Other models with Gaussian priors have included additional free parameters, such as degree of uncompensated ocular torsion, which allow these models to better capture the detailed shape of the perceived bias curve (e.g.[7,15]). Nevertheless, a simple Bayesian model with an empirical prior for upright provides a reasonable qualitative fit with only a single free parameter.

In a recent study similar to ours, Willemsen et al. also measure the empirical roll prior using a smaller dataset (6 subjects, 12 min of prescribed activities) and explore a greater range of model variations with several additional free parameters[7]. They obtain the best fit with a Gaussian rather than an empirical prior. Here, we replicate their finding, with better goodness of fit observed when using Gaussian (Fig. 6) rather than the empirical prior (Figs. 4a, 5a). Willemsen et al. note that the empirical prior leads to posterior probability distributions and thus perceptual estimates with variability that is actually increased relative to the likelihood, and that this is due to the excess kurtosis in the empirical prior compared to the Gaussian. The implication is that the nervous system may use a Gaussian approximation of an empirical prior to allow less variable roll estimates at more extreme angles. This intriguing finding does not detract from the importance of measuring empirical priors, but it does raise followup questions about the nature of this hypothesized approximation process. Specifically, Willemsen et al. suggest that noise on vestibular estimates of head orientation could lead to a neural representation of the roll prior that is more Gaussian in shape due to the central limit theorem, even if the empirically measured prior exhibits excess kurtosis[7]. Use of a Gaussian prior in our roll models (Fig. 6) confirms Willemsen and colleagues' finding that a Gaussian prior better approximates observed roll perception bias than use of the high kurtosis, empirical roll prior.

For pitch perception, both linear and non-linear noise models struggle to capture the detailed shape of the perceptual bias curve. This is because these models tend to predict attractive biases (i.e. underestimation) that increase with increasing tilt angle, as observed for roll. Pitch biases, in contrast, increase sharply out to $+/-15°$, then decrease again, becoming repulsive (i.e. overestimation) at angles greater than 60° (Figs. 4b, 5b). Consequently, a least-square fit using the linear and non-linear noise models settles on minimal multiplicative noise factors that predict little bias.

Despite the poor fit of the model to pitch perception data, we wanted to investigate whether the asymmetry observed in the empirical pitch prior, namely greater variability for forward compared to backward pitch (Fig. 2c), could explain the asymmetry in pitch perception, that is larger biases for backward compared to forward pitch. To this end, we generated predictions of the Bayesian model using a skew normal distribution for the prior (Fig. 7). Results confirmed that an asymmetrical prior predicts asymmetrical biases in pitch perception that qualitative match observed patterns, but once again, the Bayesian model fails to predict the detailed pattern of observed biases.

Generally, we are left to consider why the Bayesian framework can capture patterns of bias observed for roll but not pitch perception. An alternative modeling approach that can better explain repulsive biases may be needed. Specifically, it has been shown that coupling the Bayesian decoding modeled here with an efficient

encoding stage can predict combinations of both attractive and repulsive biases[2,3], as observed for pitch perception. Another possible explanation for the poor match between predicted and observed biases in pitch perception is that the patterns of observed biases are inaccurate or otherwise not representative. We fit to data from[16] which measured bias by asking participants to adjust their own body pitch to given target angles at 15° increments between 0° and 90°. Different patterns of bias may be observed using different methods, e.g. when perceived pitch is indicated using a visual probe[34]. Generally, in comparison to roll perception, there is a lack of studies that have attempted to measure pitch perception. Additional perceptual studies may reveal alternative patterns of bias that are more amenable to modeling using the standard Bayesian framework.

Finally, we suspect that differences in statistics across individuals (for example the substantial differences in pitch distributions shown in Fig. 2c) might result in differences in individual biases as well. In future work, it would be possible to test this hypothesis for pitch perception specifically, for example by measuring asymmetry of an individual's natural pitch distribution and testing how well this asymmetry predicts asymmetry in perceptual biases for that individual. Such a relationship would be particularly important to investigate in a clinical setting where biomechanical constraints associated with certain disorders could shape stimulus statistics and thus manifest as differences in perception of upright and perhaps even balance performance and fall risk.

Power spectra for tilt and translation: implications for models of vestibular processing.

While the dynamics of total linear acceleration have been reported previously, the separate dynamics of gravitational (tilt) and inertial (translation) components during natural everyday activities over long time periods in humans have not been reported previously. As expected, we see acceleration due to gravity contribute more to the total power at lower frequencies and power due to inertial acceleration dominate at higher frequencies. Additionally, we see a peak in power at approximately 2 Hz along the Z-axis, corresponding well with previous work that demonstrates a strong peak of linear acceleration of the head along the vertical axis with 2 Hz stepping frequency[35].

Power spectra of gravitational and inertial components of otolith stimulation are relevant for processing of ambiguous otolith stimulation, known as the tilt-translation ambiguity. An early suggestion for solving this ambiguity is frequency segregation, whereby low and high frequency otolith stimulation are interpreted as due to tilt and translation, respectively[36]. This simple heuristic approach has been successful in modeling otolith-ocular responses in both squirrel monkeys[23] and humans[24], with estimated cutoff frequencies of 0.5 and 0.3 Hz, respectively. Later studies confirmed that frequency segregation could explain ocular responses in humans, but with a significantly lower cutoff frequency of 0.07 Hz[19,20].

Observed crossing points in the power spectra (Fig. 3) where power transitions from predominantly gravitational to predominantly inertial acceleration provide empirical data to constrain selection of cutoff frequencies for frequency segregation models. We observe crossing points at frequencies that are higher than the values suggested by previous studies. Across the data set as a whole, these values are 1.148, 0.626, and 0.633 Hz for X-, Y-, and Z-axes, respectively. However, if we consider only high-velocity epochs these values fall to 1.101, 0.566, and 0.525 Hz for X, Y, and Z-axes, respectively (see Supplemental Material), which is more in line with previous research.

For resolving tilt-translation ambiguity, a more contemporary alternative to frequency segregation is multi-modal (i.e. canal-otolith) integration[17]. While previous research has reported that reflexive eye movements can be explained by frequency segregation, perception was best explained based on internal models of canal-otolith interactions[19,20]. Recent models of canal-otolith interaction for perception have found that sensory integration is most crucial in the range of 0.2–0.5 Hz[37]. We note that the differing natural dynamics of gravitational and inertial components also have a role in shaping the performance of such statistically optimal multi-modal models.

On the one hand, noise characteristics of afferent sensory signals have recently been shown to be related to natural statistics of angular velocity stimulation[38], and should thus influence measurement noise in the Kalman filter framework. The same should hold true for afferent linear acceleration signals and their relation to natural statistics of tilt and translation. Along these lines, recent analysis of otolith afferent responses to naturalistic movements reveal that regular and irregular afferents convey more information about naturalistic tilt and translation movements, respectively, precisely because of the different dynamics of these natural movements (lower vs higher frequency modulation) and the different response dynamics of these populations[22]. Neural populations that respond selectively to tilt and translation[34,39] may also exhibit distinct frequency-dependent responses in line with natural statistics.

On the other hand, the distributions of input motion should also shape the modeled process noise of the Kalman filter. Ultimately, the Kalman gain, and thus the performance of such statistically optimal models as a whole, depends on the ratio of process to measurement noise which in turn depends on natural stimulus dynamics[40]. Further research should consider how best to constrain such models with measurements of natural head movement statistics.

Methodological considerations.

Several recent studies have reported statistics of natural head movements, and it is therefore important to note differences across studies in terms of recording methods, coordinate frame conventions, and sampling procedures. In the current study, we use a commercially available VI-SLAM system which we have previously validated against an optical tracking system[10]. This method overcomes gravito-inertial ambiguity inherent in data reported by several previous studies[4,5]. The ambiguity inherent in IMU data may be overcome using an IMU with magnetometer and filtering methods[7], but the accuracy of tilt and translation estimates derived from filtering should be validated in some way[9].

In the present study, we report data in an anatomical reference frame, i.e. relative to Reid's baseline, which we define as the line running from the canthus of the eye through the middle of the meatus of the ear. This is similar to, but not necessarily identical to, the line defined by the Frankfurt (or Frankfort) plane used as a reference in previous studies[4,5], which runs through the bottom of the orbit and through the upper point of auditory canal.

The alignment of data to this reference plane or line is subject to error because it is based on approximate visual alignment by the experimenter of either the sensor on the head or the head relative to gravity (current study). An alternative method would be to adopt a purely kinematic reference frame defined only by the two planes traced when the subject naturally shakes and nods the head, the third plane being identified as the one perpendicular to the intersection of the first two. This coordinate frame would not rely on identifying Reid's baseline and would therefore have the advantage of avoiding errors introduced by approximate visual alignment and may be a more natural choice when studying head movement.

Finally, it is worth noting differences across studies in terms of what types of activities are sampled. The current study measured unprescribed activity over a longer time period; to our knowledge, this is the only study to report statistics of head orientation during unprescribed activities of daily life. Previous studies measuring unprescribed activities have reported head movements, as captured by an IMU, but not orientation[9,35]. Others report head movement and/or orientation during prescribed activities such as walking, climbing/descending stairs, running, and hopping[5,7]. Dynamic activities are likely over-represented in such studies and this likely impacts the reported statistics. It remains an open question as to what type of sampling is best-suited to which type of scientific inquiry.

One advantage of long-term, unprescribed sampling is that it is more likely to be representative of natural behavior, and data can always be partitioned post-hoc based on approximate human activity recognition (see Supplemental Material). This reveals that difference across behavioral modes can be significant and raises the possibility of mode- or activity dependent (e.g. locomotion-dependent) neural processing[41]. In the current study we make a first attempt at post-hoc activity recognition with a simple heuristic based on each sample's instantaneous velocity norm, splitting data into low- and high-velocity categories. This is a coarse method which doesn't take into account transient decreases or increases in linear velocity norm that might happen during activities such as walking, but nonetheless is useful to see how statistics of head orientation may change with activity.

Conclusion.   Here we report for the first time the statistics of head orientation relative to gravity as well as power spectra of naturally-experienced gravitational and inertial acceleration during everyday activity in humans. We show that these measures have implications for perception of head orientation, as well as processing of dynamic vestibular stimulation. Measures of head orientation are more broadly relevant because they not only constrain models of spatial orientation and vestibular processing, but also determine how the nervous system samples gravity-dependent visual and auditory structure in the environment. Future work is needed to investigate how these measures vary across different groups, such as children and clinical populations, across activities, as well as across species[1]. It is also interesting to consider how statistical measures such as these can inform more technological endeavors, such as predictive methods for head or gaze tracking that are relevant for emerging virtual and augmented reality technologies[42]. We expect that increased availability of head tracking data in the future will contribute to advances across these scientific and application domains.

Data availability

The dataset used and/or analysed during the current study is available from the corresponding author on reasonable request.

References

1. MacNeilage, P. R. Characterization of Natural Head Movements in Animals and Humans. In *The Senses: A Comprehensive Reference* (eds Fritzsch, B. & Straka, H.) 6987 (Elsevier, Academic Press, 2020).
2. Wei, X.-X. & Stocker, A. A. A bayesian observer model constrained by efficient coding can explain "anti-bayesian" percepts. *Nat. Neurosci.* **18**, 1509–1517 (2015).
3. Wei, X.-X. & Stocker, A. A. Lawful relation between perceptual bias and discriminability. *Proc. Natl. Acad. Sci.* **114**, 10244–10249 (2017).
4. Carriot, J., Jamali, M., Chacron, M. J. & Cullen, K. E. Statistics of the vestibular input experienced during natural self-motion: Implications for vestibular processing. *J. Neurosci.* **34**, 8347–8357 (2014).
5. Carriot, J., Jamali, M., Cullen, K. E. & Chacron, M. J. Envelope statistics of self-motion signals experienced by human subjects during everyday activities: Implications for vestibular processing. *PLoS ONE* **12**, 1–24 (2017).
6. Carriot, J., Jamali, M., Chacron, M. J. & Cullen, K. E. The statistics of the vestibular input experienced during natural self-motion differ between rodents and primates. *J. Physiol.* **595**, 2751–2766 (2017).
7. Willemsen, S. C. M. J., Wijdenes, L. O., van Beers, R. J., Koppen, M. & Medendorp, W. P. Natural statistics of head roll: Implications for bayesian inference in spatial orientation. *J. Neurophysiol.* **128**, 1409 (2022).
8. Schwabe, L. & Blanke, O. The vestibular component in out-of-body experiences: A computational approach. *Front. Hum. Neurosci.* **2**, 1–10 (2008).
9. Hausamann, P., Daumer, M., MacNeilage, P. R. & Glasauer, S. Ecological momentary assessment of head motion: Toward normative data of head stabilization. *Front. Neurosci.* **13**, 179 (2019).
10. Hausamann, P., Sinnott, C. B., Daumer, M. & MacNeilage, P. R. Evaluation of the intel realsense t265 for tracking natural human head motion. *Sci. Rep.* **11**, 1–13 (2021).
11. Sinnott, C. B., Dang, T., Papachristos, C., Alexis, K. & MacNeilage, P. R. Statistical characterization of heading stimuli in natural environments using slam. *J. Vis.* **18**, 41 (2018).
12. Kumar, A., Pundlik, S., Peli, E. & Luo, G. Comparison of visual slam and imu in tracking head movement outdoors. *Behav. Res. Methods* https://doi.org/10.3758/s13428-022-01941-1 *(2022).*
13. MacNeilage, P. R., Banks, M. S., Berger, D. R. & Bülthoff, H. H. A bayesian model of the disambiguation of gravitoinertial force by visual cues. *Exp. Brain Res.* **179**, 263–290 (2006).
14. de Vrijer, M., Medendorp, P. & Van Gisbergen, J. A. M. Accuracy-precision trade-off in visual orientation constancy. *J. Vis.* **9**, 1–15 (2009).

15. Clemens, I. A. H., De Vrijer, M., Selen, L. P. J., Van Gisbergen, J. A. M. & Medendorp, W. P. Multisensory processing in spatial orientation: An inverse probabilistic approach. *J. Neurosci.* **31**, 5365–5377 (2011).
16. Cohen, M. M. & Larson, C. A. Human spatial orientation in the pitch dimension. *Percep. Psychophys.* **16**, 508 (1974).
17. Glasauer, S. Interaction of semicircular canals and otoliths in the processing structure of the subjective zenith. *Ann. N Y Acad. Sci.* **656**, 847 (1992).
18. Angelaki, D. E., Wei, M. & Merfeld, D. M. Vestibular discrimination of gravity and translational acceleration. *Ann. N Y Acad. Sci.* **942**, 114–127 (2001).
19. Merfeld, D. M., Park, S., Gianna-Poulin, C., Black, F. O. & Wood, S. Vestibular perception and action employ qualitatively different mechanisms: i: Frequency response of vor and perceptual responses during translation and tilt. *J. Neurophysiol.* **94**, 186–198 (2005).
20. Merfeld, D. M., Park, S., Gianna-Poulin, C., Black, F. O. & Wood, S. Vestibular perception and action employ qualitatively different mechanisms: ii: vor and perceptual responses during combined tilt and translation. *J. Neurophysiol.* **94**, 199–205 (2005).
21. Fitzpatrick, R. C. & Day, B. L. Probing the human vestibular system with galvanic stimulation. *J. Appl. Physiol.* **96**, 2301–2316 (2004).
22. Jamali, M., Carriot, J., Chacron, M. K. & Cullen, K. E. Coding strategies in the otolith system differ for translational head motion vs. static orientation relative to gravity. *eLife* **8**, 1–24 (2019).
23. Paige, G. D. & Tomko, D. L. Eye movement responses to linear head motion in the squirrel monkey: iL Basic characteristics. *J. Neurophysiol.* **65**, 1170–1182 (1991).
24. Wood, S. J. Human otolith-ocular reflexes during off-vertical axis rotation: Effect of frequency on tilt-translation ambiguity and motion sickness. *Neurosci. Lett.* **323**, 41–44 (2002).
25. Barreira, T. V., Rowe, D. A. & Kang, M. Parameters of walking and jogging in healthy young adults. *Int. J. Exerc. Sci.* **3**, 4–13 (2010).
26. Matthis, J. S., Yates, J. L. & Hayhoe, M. M. Gaze and the control of foot placement when walking in natural terrain. *Curr. Biol.* **28**, 1224-1233.e5 (2018).
27. Pelz, J. B. & Rothkopf, C. Chapter 31 Oculomotor Behavior in Natural and Man-made Environments. In *Eye Movements* (eds Van Gompel, R. P. *et al.*) 661–676 (Elsevier, Oxford, 2007).
28. Hollands, M. A., Marple-Horvat, D. E., Henkes, S. & Rowan, A. K. Human eye movements during visually guided stepping. *J. Mot. Behav.* **27**, 155–163 (1995).
29. Eggert, T. *Der Einfluss orientierter Texturen auf die subjektive visuelle Vertikale und seine systemtheoretische Analyse.* Ph.D. thesis, Technische Universität München (1998).
30. Girshick, A. R., Landy, M. S. & Simoncelli, E. P. Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nat. Neurosci.* **14**, 926–932 (2011).
31. Schöne, H. On the role of gravity in human spatial orientation. *Aerosp. Med.* **35**, 764–772 (1964).
32. Clark, T. K., Newman, M. C., Karmali, F., Oman, C. M. & Merfeld, D. M. Chapter 5 - mathematical models for dynamic, multisensory spatial orientation perception. In Ramat, S. & Shaikh, A. G. (eds.) *Mathematical Modelling in Motor Neuroscience: State of the Art and Translation to the Clinic. Ocular Motor Plant and Gaze Stabilization Mechanisms*, vol. 248 of *Progress in Brain Research*, 65–90, https://doi.org/10.1016/bs.pbr.2019.04.014 (Elsevier, 2019).
33. Mueller, G. E. Über das aubertsche phaenomenon. *Z. Psychol. Physiol. Sinnesorg.* **49**, 109–246 (1916).
34. Laurens, J., Meng, H. & Angelaki, D. E. Neural representation of orientation relative to gravity in the macaque cerebellum. *Neuron* **80**, 1508–1518 (2013).
35. MacDougall, H. G. & Moore, S. T. Marching to the beat of the same drummer: The spontaneous tempo of human locomotion. *J. Appl. Physiol.* **99**, 1164–1173 (2005).
36. Mayne, R. A Systems Concept of the Vestibular Organs. In *Vestibular System Part 2: Psychophysics, Applied Aspects and General Interpretations* (ed. Kornhuber, H. H.) 493–580 (Springer, Berlin, Heidelberg, 1974).
37. Lim, K., Karmali, F., Nicoucar, K. & Merfeld, D. M. Perceptual percision of passive body tilt is consistent with statistically optimal cue integration. *J. Neurophysiol.* **117**, 2037–2052 (2017).
38. Carriot, J., Cullen, K. E. & Chacron, M. J. The neural basis for violations of weber's law in self-motion perception. *Proc. Natl. Acad. Sci.* **118**, e2025061118 (2021).
39. Yakusheva, T. A. *et al.* Purkinje cells in posterior cerebellar vermis encode motion in an inertial reference frame. *Neuron* **54**, 973–985 (2007).
40. Karmali, F., Whitman, G. T. & Lewis, R. F. Bayesian optimal adaptation explains age-related huamn sensorimotor changes. *J. Neurophysiol.* **119**, 509–520 (2018).
41. Niell, C. M. & Stryker, M. P. Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* **65**, 472–479 (2010).
42. Liu, W. *et al.* Tlio: Tight learned inertial odometry. *IEEE Robot. Autom. Lett.* **5**, 5653–5660 (2020).

Acknowledgements

Author contributions

P.M. and C.S. conceived the experiments, C.S. conducted the experiment, C.S., P.H. and P.M. analyzed results, C.S. prepared all figures. C.S. and P.M. wrote the main manuscript text. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-32794-z.

Correspondence and requests for materials should be addressed to C.B.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.