

University of Nevada, Reno

**A quantitative investigation of natural head movement and its contribution to spatial  
orientation perception**

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of  
Philosophy in Psychology

by

Christian B. Sinnott

Dr. Paul R. MacNeilage/Thesis Advisor

May, 2023

Copyright by Christian B. Sinnott 2023

All Rights Reserved



THE GRADUATE SCHOOL

We recommend that the dissertation  
prepared under our supervision by

**CHRISTIAN B. SINNOTT**

entitled

**A quantitative investigation of natural head movement and its  
contribution to spatial orientation perception**

be accepted in partial fulfillment of the  
requirements for the degree of

**Doctor of Philosophy**

Paul R. MacNeilage, Ph.D.  
*Advisor*

Mark Lescroart, Ph.D.  
*Committee Member*

Michael A. Webster, Ph.D.  
*Committee Member*

Fang Jiang, Ph.D.  
*Committee Member*

Eelke Folmer, Ph.D.  
*Graduate School Representative*

Markus Kimmelmeier, Ph.D., Dean  
*Graduate School*

May, 2023

## ABSTRACT

Movement is ubiquitous in everyday life. As we exist in a physical world, we constantly account for our position in it relative to other physical features: both at a conscious, volitional level and an unconscious one. Our experience estimating our own position accumulates over the lifespan, and it is thought that this experience (often referred to as a prior) informs current perception of spatial orientation. Broadly, this perception of spatial orientation is rapidly performed by the nervous system by monitoring, interpreting, and integrated sensory information from multiple sense organs. To do this efficiently, the nervous system likely represents this sensory information in a statistically optimal manner.

Some of the most important information for spatial orientation perception comes from visual and vestibular sensation, which rely on sensory organs located in the head. While statistical information about natural visual and vestibular stimuli have been characterized, natural head movement and position, which likely drives correlated dynamics across head-located senses, has not. Furthermore, sensory cues essential to spatial orientation perception are directly affected by head movement specifically. It is likely that measurement of these sensory cues taken during natural behaviors sample a significant portion of the total behaviors that comprise ones' prior.

In this dissertation, I present work quantifying characteristics of head orientation and heading, two dimensions of spatial orientation, over long-duration recordings of natural behavior in humans. Then, I use these to generate priors for Bayesian modeling frameworks which successfully predict observed patterns of orientation and heading

perception bias. Given the ability to predict some patterns of bias (head roll and heading azimuth) particularly well, it is likely our data are representative of real behaviors that comprise previous experience the nervous system may have.

Natural head orientation and heading distributions reveal several interesting trends that open future lines of research. First, head pitch demonstrates large amounts of inter-subject variability; likely this is due to biomechanical differences, but as these remain relatively stable over the lifespan these should bias head movements. Second, heading azimuth appears to vary significantly as a function of task. Heading azimuth distributions during low velocities (which predominantly consist of stationary activities like standing or sitting) are strongly multimodal across all subjects, while azimuth distributions during high velocities (predominantly consisting of locomotion) are unimodal with relatively low variance. Future work investigating these trends, as well as implications these trends and data have for sensory processing and other applications is discussed.

## Table of Contents

ABSTRACT.....	i
Chapter 1: General Introduction.....	1
Motivation .....	1
Technical Background .....	5
Coordinate Systems .....	5
Rigid Body and Reference Frames .....	8
Degrees of freedom, position, velocity, acceleration .....	9
Head-Tracking .....	14
Historical Methods.....	14
Inertial Measurement Units.....	15
Outside-in Optical Tracking Systems .....	17
Inside-out Positional Estimation .....	18
Vestibular Anatomy .....	19
Semicircular Canals .....	20
Otolith Organs (Utricle and Sacculle).....	22
Head-Centered Sensory Dimensions and Physiological Representations .....	24
Head Orientation .....	24
Heading .....	27
Head Orientation and Heading Perception: Previous Research and Observed Bias .....	30
Previous Head Orientation Research .....	30
Head Orientation Bias.....	32
Previous Heading Research.....	34
Vestibular Heading Bias .....	34
Sensory Processing Models.....	35
Internal Observer Models .....	36
Static Bayesian Models .....	38
Dynamic/Recursive Bayesian Estimators .....	44
Scope.....	50
References .....	52
Chapter 2: Natural statistics of human head orientation constrain models of vestibular processing .....	57

Methods.....	60
Results.....	64
Discussion .....	69
Data availability.....	76
References .....	76
Acknowledgements.....	77
Author contributions .....	77
Additional information.....	77
Chapter 3: Natural statistics of human heading are non-Gaussian and reveal task-dependent differences .....	79
Abstract.....	79
Introduction .....	81
Methods.....	86
Hardware .....	86
Calibration.....	87
Pre-processing.....	88
Descriptive statistics and prior generation .....	89
Modeling heading perception.....	91
Results.....	95
Descriptive Statistics .....	95
Modeling .....	98
Discussion .....	101
Characteristics of Heading Distributions.....	102
Modeling Human Heading Perception.....	104
Limitations .....	105
Future Directions .....	108
Conclusion.....	109
References .....	110
CHAPTER 4: General Discussion .....	113
Summary .....	113
Future Work .....	114
Head Pitch Perception .....	114

Heading Perception.....	116
Dynamic Models and Priors .....	119
Human Activity Recognition and Clinical Diagnostics.....	121
Conclusions .....	123
References .....	124
Appendix A:.....	126
References .....	136

## List of Tables

Table 1: Central Moments of Low- and High-Velocity Head Orientation .....	133
Table 2: Smoothed Prior Model Parameters .....	134
Table 3: Simulated Prior Model Parameters.....	135
Table 4: Subject-wise Central Moments of Head Orientation.....	135
Table 5: Subject-wise Central Moments of Low- and High-Velocity Head Roll .....	136
Table 6: Subject-wise Central Moments of Low- and High-Velocity Head Pitch .....	136

## List of Chapter 1 Figures

Figure 1: 1D Cartesian Coordinate System .....	5
Figure 2: 2D Cartesian Coordinate System .....	6
Figure 3: 3D Cartesian Coordinate System .....	7
Figure 4: 3D Spherical Coordinate System .....	8
Figure 5: Rigid Body Transformation .....	9
Figure 6: Vestibular Gross Anatomy .....	22
Figure 7: Bayesian Model with Symmetric Likelihood .....	40
Figure 8: Bayesian Model with Asymmetric Likelihood .....	42

## List of Chapter 2 Figures

Figure 1: Recording Equipment .....	61
Figure 2: 2D Head Orientation .....	65
Figure 3: Linear Acceleration Power Spectra.....	66
Figure 4: Linear Noise Head Orientation Model Results.....	66
Figure 5: Utricular Shear Noise Head Orientation Model Results .....	67
Figure 6: Gaussian Prior Head Roll Model Results.....	67
Figure 7: Skew-Normal Prior Head Pitch Model Results.....	68
Figure 8: Bayesian Model with Asymmetric Likelihood .....	42

## List of Chapter 3 Figures

Figure 1: 2D Heading.....	96
Figure 2: 2D Low Velocity Heading .....	97
Figure 3: 2D High Velocity Heading.....	98
Figure 4: Linear Noise Heading Model Results .....	99
Figure 5: Utricular Shear Heading Model Results .....	100
Figure 6: Efficient Bayesian Heading Model Results.....	101

## List of Chapter 4 Figures

Figure 1: Efficient Bayesian Low- and High-Velocity Models.....	117
---	-----

## List of Appendix A Figures

Figure 1: Smoothed Roll and Pitch Prior Distributions .....	128
Figure 2: Smoothed Roll Model Results.....	128
Figure 3: Effect of Prior Smoothing on Posterior Calculation.....	129
Figure 4: Smoothed Pitch Model Results.....	130
Figure 5: Low-Velocity Power Spectra .....	130
Figure 6: High-Velocity Power Spectra .....	131
Figure 7: Low-Velocity 2D Head Orientation .....	131
Figure 8: High-Velocity 2D Head Orientation .....	132

## Chapter 1: General Introduction

### Motivation

Head movement is ubiquitous in everyday life. While performing purposeful movements like locomotion or playing a sport, the head moves in volitional ways. Even when at rest, or standing quietly, the head remains moving as posture is maintained and perturbed by disturbances as small as one's own heartbeat and respiration. A large motivation for moving one's head is to move the sensory organs contained in the head into positions that better sample stimuli of interest in the environment, or to maintain those organs' current orientation towards some object of interest. These sampling and orientation movements comprise a major part of how vertebrates orient themselves spatially in the environment.

The head is home to the sensory organs for three sensory systems critical to spatial orientation perception: the visual system (eyes), the auditory system (ears), and the vestibular system (otolith organs and semicircular canals). As the head moves, the information that these sensory organs are exposed to will systematically change.

Furthermore, movement of the organs may facilitate higher-resolution sampling of relevant stimuli: for example, moving one's eyes to place a target within the foveated, high-resolution area of one's field of view. While human movement has been broadly studied for thousands of years, systematic characterization of how the head orients and moves, particularly in relation to head-centered sensory organs, is a more recent line of research.

Extensive research has been conducted investigating how the head moves in stereotyped tasks, particularly in the context of how the head and eyes move together. However, how

the head moves during unconstrained, natural behaviors of humans has not been well-studied to date. Understanding and characterizing natural human head movement is important across multiple disciplines and applications. With respect to sensory neuroscience, quantifying natural head behavior is essential to solving long-studied questions of how the human nervous system encodes and disambiguates noisy, ambiguous sensory information. Two important dimensions of spatial orientation perception are the instantaneous direction of gravity as well as the direction of the organism's self-motion. As these are largely informed by the visual and vestibular organs, how the head is oriented relative to gravity and the direction of self-motion plays a major role in the physical characteristics of the stimulation that comprises these cues.

The perception of ones' own head orientation relative to gravity and ones' heading (direction of linear self-motion) is systematically biased, however. Orientation relative to gravity is systematically underestimated as one moves further from completely upright. In contrast, heading exhibits systematic alternation between underestimation and overestimation across different directions. This bias may be explained in part by the previous experience an organism has in moving its head relative to gravity and the direction of its own self-motion, but this experience is likely comprised of activities that, until recently, were difficult or impossible to rigorously quantify in controlled laboratory settings.

Although important for questions of sensation and perception, measurement of natural head movement is broadly applicable outside of this domain. Understanding natural head movement can give insight on the range of movement the human body must work against

in order to stabilize sensory organs during activities like locomotion or athletics. Stereotyped behaviors like gait are reflected in head movement and can give insight on clinical conditions or disease states in people with sensory deficits, neurodegenerative diseases, or acute trauma. Characterization of head movement can also constrain solutions to perform predictive tracking in applications like virtual or augmented reality, as well as provide new data for human activity recognition methods as wearable technologies like smartglasses or other head-mounted interfaces become more ubiquitous in everyday life.

The following sections of this introduction will discuss technical definitions, methods, and background necessary to understand previous work focusing on two specific, head-centered dimensions or quantities: head orientation relative to gravity and heading. This introduction will also introduce the reader to methods used to model patterns of perceptual bias observed in perception of head orientation and heading direction, which are or can be constrained by natural head orientation and heading statistics to varying degrees. This will set the stage for the reader to fully understand work presented in subsequent chapters of this dissertation.

Chapter 2 focuses on published work characterizing head orientation relative to gravity during natural, everyday behaviors. We observe and measure head roll and head pitch in ten human subjects over five-hour recording sessions. Due to our use of inside-out positional tracking methods to perform this observation, we are not only able to characterize head orientation distributions but also able to parse linear acceleration transduced by the otoliths into separate gravitational and inertial components. Both

measurements have important implications for sensory processing models. We use head orientation distributions to constrain a set of static Bayesian models which predict classic patterns of bias in spatial orientation perception and achieve reasonable explainability with models constrained by natural statistics.

Chapter 3 focuses on related work using the same dataset to characterize heading azimuth and elevation. We measure azimuth and elevation in the same group of subjects and find that distributions are significantly non-Gaussian. As in Chapter 3, we use these data to constrain a contemporary efficient Bayesian model to predict patterns of observed heading azimuth and elevation perception bias. Furthermore, selective parsing of data based on linear velocity magnitude suggests that characteristics of these distributions may change as a function of task or mode: epochs with high magnitude that are largely comprised of locomotion have strongly different characteristics than distributions derived from all epochs.

The conclusion of this dissertation focuses on discussion of these findings and what they reveal about avenues of future research. We identify two novel avenues of research which relate individual perceptual performance to natural statistics: measurement of pitch perception bias, both at a group and individual level, as well as measuring the effect of speed on heading perception. Further discussion will focus on possible applications of data collected in the preceding two chapters outside the domain of theoretical research. These data have potential value with respect to head-based human activity recognition, which is expected to increase in prominence as wearable devices worn on the head (e.g. smartglasses or augmented reality displays) enter the market.

## Technical Background

The following section will establish language and definitions used throughout this dissertation concerning observation of position and motion, as well as an overview of methods used to estimate human head motion in research over the last century.

### Coordinate Systems

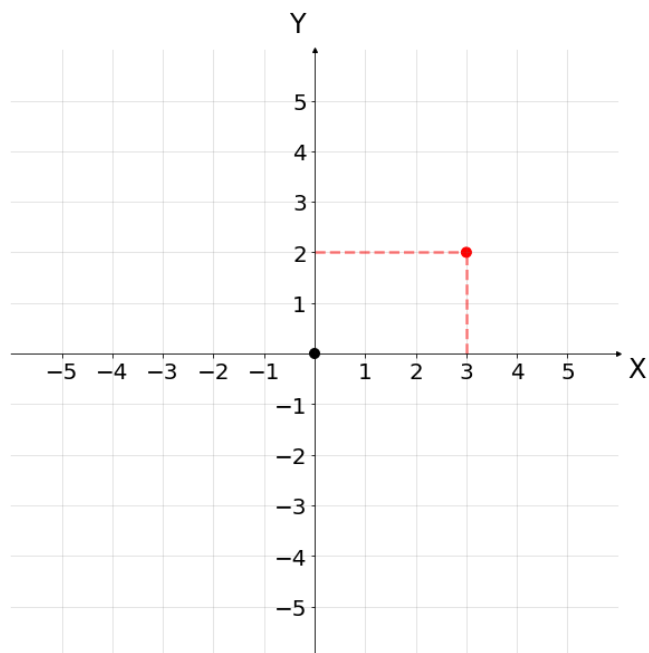
Motion and position are always described relative to some observer or other point.

Understanding coordinate systems comprises a necessary first step towards understanding observation of motion in a three-dimensional space. Generally, a coordinate system is a framework that uses a set of numbers (coordinates) to unambiguously determine position of some point of interest relative to another fixed point. Typically, coordinate systems have some arbitrary start point where all coordinates are zero, referred to as the origin ( $O$ ). A common coordinate system is a simple number line:



**Figure 1:** An example of a 1-dimensional Cartesian coordinate system (number line) with origin in black (at 0) and a point in red (at 3).

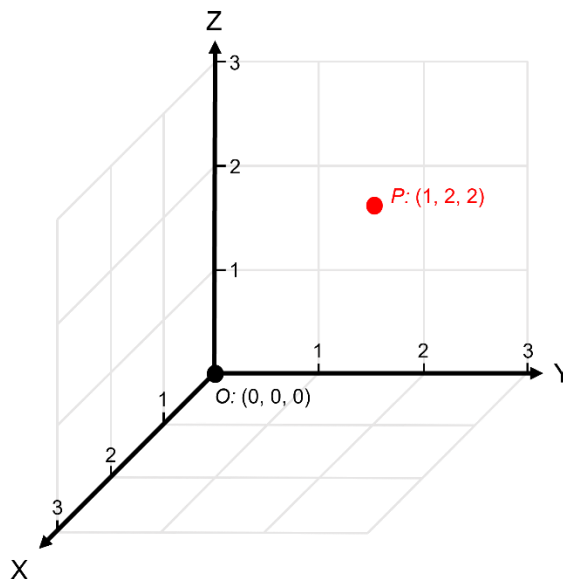
Here, a position is defined by a single coordinate: the signed distance of a given point from the origin along the number line. A number line can be divided into two halves, a positive and negative half, describing which side of the origin a given point ( $P$ ) falls on. Likely, the two-dimensional version of this coordinate system is also familiar to readers:



**Figure 2:** An example of a 2-dimensional Cartesian coordinate system (Cartesian plane). The origin is plotted in black at  $(0, 0)$  while a point in the plane is plotted in red at  $(3, 2)$ . Dashed lines indicate distance from the origin along each axis.

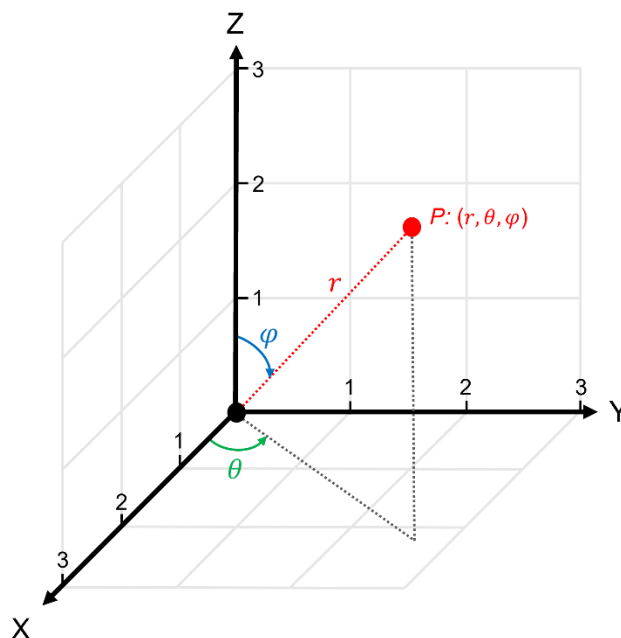
A simple description of a two-dimensional Cartesian coordinate system is that it is two number lines that are perpendicular to each other. The point where these lines or axes meet is the origin and takes an ordered pair or tuple with values  $(0, 0)$  as its coordinates. Values in a set of coordinates are real numbers that correspond with the distance of a given point on one of these axes relative to the origin: the first value corresponds with the distance of the point on the first axis (the X-axis, or less frequently, the abscissa), while the second value describes the distance of the point on the second axis (Y-axis or ordinate). Like the number line, the Cartesian plane can be divided into four quadrants which can describe each combination of sign present in the coordinate values.

The natural extension of this coordinate system is into a three-dimensional space:



**Figure 3:** An example of a 3-dimensional Cartesian coordinate system (Cartesian volume). The origin is plotted in black at  $(0, 0, 0)$  while a point in the volume is plotted in red at  $(1, 2, 2)$ .

Like before, this can be defined in terms of lower-dimensional Cartesian coordinate systems as three orthogonal axes, which create three orthogonal planes that intersect at a single origin. This third axis is the Z-axis (or applicate) and adds a third value to our set of coordinates, which are now an ordered triplet. A three-dimensional Cartesian volume can be divided into eight octants corresponding with all possible permutations of sign along all three axes. An alternative method of representing the location of a point in 3-dimensional space is with a spherical coordinate system:



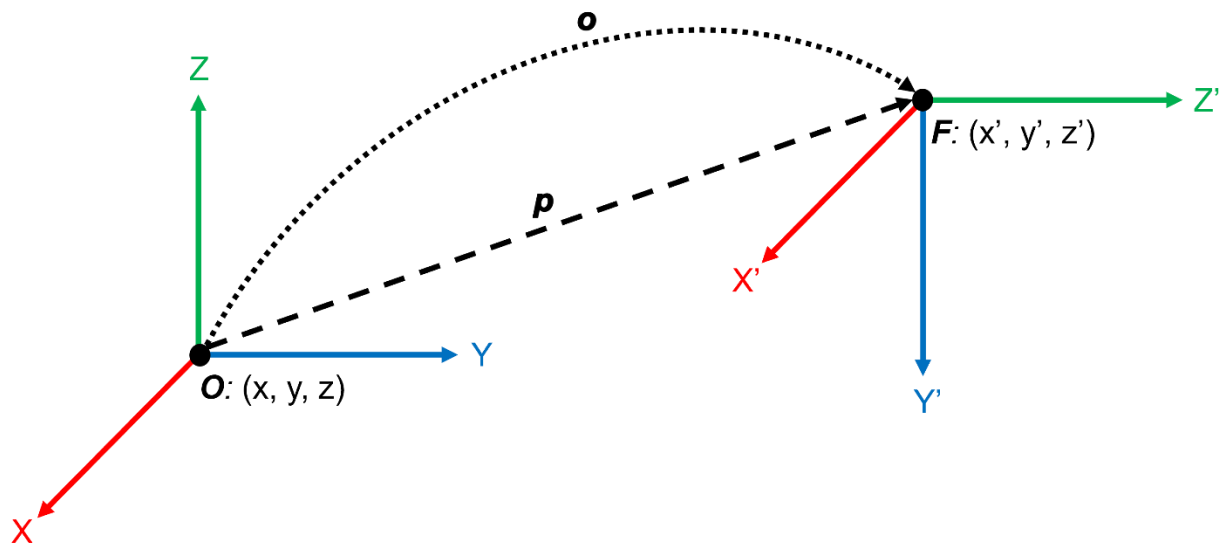
**Figure 4:** An example of a 3-dimensional Cartesian coordinate system with points represented using spherical coordinates  $(r, \theta, \varphi)$ .

Rather than the location of a given point being described as the distance from the origin along each axis, in a spherical coordinate system a point's location is described by three other values. The first is the radial distance of the point from the origin (denoted with  $r$  or sometimes  $\rho$ ), the second is the angle of the point relative to a fixed direction (azimuth,  $\theta$ ), and the last is the angle of projection orthogonal to the zenith direction and passing through the origin (elevation/inclination,  $\varphi$ ). While this is a different representation of a point's location in 3-D space, it is useful when describing the angle of a point's movement to its location relative to the origin point.

### Rigid Body and Reference Frames

Given some moving object, an object's position can be described relative to the position of a reference frame attached to that object. A reference frame is a coordinate system that

is attached to the object with a six degrees of freedom position that is invariant as the object moves. This reference frame also must be considered a rigid body: an object with functionally zero deformation. This must be satisfied as any deformation of the reference frame will result in change in its position with an ambiguous origin (i.e. positional change driven by deformation or by object positional change). An intuitive way to think about this is that rather than describing a point's distance relative to the origin in a coordinate system, reference frame transformation describes the distance of one origin corresponding to one coordinate system from another point on another coordinate system.



**Figure 5:** Representation of the relationship between two reference frames,  $O$  and  $F$ .  $F$  is translated some distance in 3-dimensional space relative to  $O$  using the position vector  $p$ .  $F$  is also rotated 90 degrees about its x-axis ( $X'$ , in red) relative to  $O$ . This rotation is described using the orientation vector  $o$ .

#### Degrees of freedom, position, velocity, acceleration

Movement of a rigid body has six degrees-of-freedom (DOF) in a three-dimensional space. Three of these DOF describe linear position (hereafter referred to as position)

which changes via translation along each of the three axes through space. Movement along the fore-aft axis is surge, along the up-down axis is heave, and about the left-right axis is sway. The other three DOF describe angular position (orientation), which changes via rotation about each axis. Rotation about the fore-aft axis is roll, about the up-down axis is yaw, and about the left-right axis is pitch.

### **Linear Position**

Position of a rigid body can be represented as a 3D vector from the origin of an object to the origin of an associated reference frame:

$$p = \overrightarrow{OF} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix}$$

### **Linear Position Derivatives and Other Measures**

Linear velocity ( $v$ ), the first derivative of linear position, describes change in linear position over time. This can be expressed with respect to velocity along a single axis of the reference frame with origin  $p$ :

$$v_x(t) = \frac{\partial p_x(t)}{\partial t}$$

$$v_y(t) = \frac{\partial p_y(t)}{\partial t}$$

$$v_z(t) = \frac{\partial p_z(t)}{\partial t}$$

Magnitude of velocity, speed, can be calculated using the Euclidean vector norm:

$$\|v(t)\| = \sqrt{v_x(t)^2 + v_y(t)^2 + v_z(t)^2}$$

Using each axis' velocity as well as the norm, we can convert direction of velocity out of a Cartesian representation into a spherical representation, corresponding with relevant heading angles, azimuth ( $\theta$ ) and elevation ( $\phi$ ):

$$\theta = -\arctan2(v_y(t), v_x(t))$$

$$\phi = \arcsin\left(\frac{v_z(t)}{\|v(t)\|}\right)$$

Linear acceleration ( $a$ ) is the second derivative of linear position and first derivative of linear velocity. It describes change in linear velocity over time:

$$a_x(t) = \frac{\partial v_x(t)}{\partial t}$$

$$a_y(t) = \frac{\partial v_y(t)}{\partial t}$$

$$a_z(t) = \frac{\partial v_z(t)}{\partial t}$$

### **Angular Position (Orientation)**

Angular position (hereafter referred to as orientation) of a rigid body can be represented in many ways and is defined as the rotation required to position the rigid body's orientation in-line with some moving object of interest's orientation. A common representation is through use of Euler angles, as well as common related variants including Tait-Bryan rotations and Davenport chained rotations, which describe three consecutive rotations about rigid-body fixed axes. Under certain conditions, these ordered

rotations can result in gimbal lock and loss of a degree of freedom when rotation leads to one axis becoming parallel with another. A solution to this is use of an alternative convention, the quaternion:

$$o = o_w + o_x i + o_y j + o_z k = \begin{pmatrix} o_w \\ o_x \\ o_y \\ o_z \end{pmatrix}$$

Quaternions provide a representation of orientation and rotation using complex numbers which provide an unambiguous account of an object's orientation ( $o$ ). A large advantage of quaternion representation over other representations is that they are not susceptible to gimbal lock. They consist of both real ( $o_w$ ) and imaginary ( $o_x i + o_y j + o_z k$ ) components, and when normalized into a unit quaternion or versor ( $q$ ), such that:

$$\|q\| = 1$$

unit quaternions can be differentiated to estimate angular velocity and acceleration.

### Angular Position Derivatives

Angular velocity is the first derivative of angular position and describes the change in orientation over time. This can be calculated about each Cartesian axis directly from the unit quaternion representation:

$$\omega_x(t) = \frac{2}{\Delta t} \left( q_w(t)q_x(t + \Delta t) - q_x(t)q_w(t + \Delta t) - q_y(t)q_z(t + \Delta t) + q_z(t)q_y(t + \Delta t) \right)$$

$$\omega_y(t) = \frac{2}{\Delta t} \left( q_w(t)q_y(t + \Delta t) + q_x(t)q_z(t + \Delta t) - q_y(t)q_w(t + \Delta t) - q_z(t)q_x(t + \Delta t) \right)$$

$$\omega_z(t) = \frac{2}{\Delta t} \left( q_w(t)q_z(t + \Delta t) - q_x(t)q_y(t + \Delta t) + q_y(t)q_x(t + \Delta t) - q_z(t)q_w(t + \Delta t) \right)$$

Angular velocity can also be estimated by first converting from the unit quaternion representation into a 3-2-1 (yaw, pitch, roll) Euler angle representation:

$$\phi = \arctan2\left(2(q_w q_x + q_y q_z), 1 - 2(q_x^2 + q_y^2)\right)$$

$$\theta = \frac{-\pi}{2} + 2\arctan2\left(\sqrt{1 + 2(q_w q_y - q_x q_z)}, \sqrt{1 - 2(q_w q_y - q_x q_z)}\right)$$

$$\psi = \arctan2\left(2(q_w q_z + q_x q_y), 1 - 2(q_y^2 + q_z^2)\right)$$

Then by taking the first derivative of the angle of interest, angular velocity along each axis can be estimated:

$$\omega_x(t) = \frac{\partial\psi(t)}{\partial t}$$

$$\omega_y(t) = \frac{\partial\theta(t)}{\partial t}$$

$$\omega_z(t) = \frac{\partial\phi(t)}{\partial t}$$

It should be noted that this will be susceptible to gimbal lock. More information including a general method for estimating angular velocity from the unit quaternion orientation representation can be found at (Bernardes and Viollet, 2022).

### **Angular Acceleration**

Angular acceleration ( $\alpha$ ) is the second derivative of orientation and first derivative of angular velocity, and describes change in angular velocity over time:

$$\alpha_x(t) = \frac{\partial\omega_x(t)}{\partial t}$$

$$\alpha_y(t) = \frac{\partial \omega_y(t)}{\partial t}$$

$$\alpha_z(t) = \frac{\partial \omega_z(t)}{\partial t}$$

## Head-Tracking

### Historical Methods

Historically, head movement in mammals including humans has been measured in restrictive laboratory settings (for review, see MacNeilage 2020). Early measurement methods were mechanical, using cables that were connected to the head of a subject on one end, while connected to a potentiometer on the other (Bartz, 1966; Waters, Morris, and Perry, 1973). Other early optical methods used photography or video recordings of the head in front of some known, scaled geometry such as a checkerboard background (Dunlap and Mowrer, 1930). Finally, methods using magnetic search coils to estimate six degree-of-freedom head position with high spatial resolution have been used successfully since the 1970s (Collewijn, 1977).

While many of these estimation methods provide sound, relatively low-error estimates of head position, they are constrained to laboratory settings which may be space-limited. Additionally, recording methods may not be capable of capturing all six DOF of head position; for example, a camera tracking movement along the sagittal plane of the head (comprised of the fore-aft and up-down axes) will not be able to reliably estimate movement along the left-right axis: in this way, they are task-limited. Methods robust

enough to collect data during less constrained activities necessitated development of less constraining solutions.

### Inertial Measurement Units

More contemporary, portable methods of motion estimation have been marked by development and widespread adoption of micro-electromechanical system (MEMS)-based inertial measurement units (IMUs). In contrast to large IMUs mounted on spacecraft and aircraft that, while high-performing, are not human-portable; MEMS IMUs provide acceptable performance with a very portable form factor. All IMUs typically contain a set of accelerometers associated with each axis that can estimate linear acceleration along that axis, as well as a set of gyroscopes associated with each axis that estimate angular velocity about each axis. While there are several methods to construct accelerometers and gyroscopes, typical MEMS-IMUs estimate linear acceleration and angular velocity by estimating the extent of compression or extension of springs attached to the proof mass and enclosure experienced as a result of force applied to the proof mass.

With respect to an accelerometer, this force is the acceleration applied to the proof mass as the IMU moves, as well as the constant acceleration imparted to the proof mass as a result of gravity. With a gyroscope, the force is the Coriolis force acting perpendicular to the direction of motion and axis of rotation a gyroscope is experiencing. An additional sensor present in many MEMS-IMUs is a magnetometer which estimates the strength and direction of a local magnetic field, which can be used as a proxy to estimate an earth-fixed reference frame.

As accelerometers and gyroscopes measure linear acceleration and angular velocity, respectively, further processing is necessary to generate robust estimates of position and orientation. An intuitive first step may be integration of these signals to recover position and orientation. Unfortunately, estimates rendered by the sensors in an IMU are susceptible to noise from various sources; integration of this noisy signal will result in accumulation of error over time and is often referred to as drift. Furthermore, accelerometers are incapable of distinguishing between gravity-driven acceleration and acceleration driven by self-motion due to Einstein's equivalence principle.

Integration of the linear acceleration signal alone is insufficient to recover linear position. More sophisticated methods to recover position and orientation from IMU estimates involve fusing information from multiple sources (aptly named sensor fusion), reducing overall noise or uncertainty in the signal and mitigating drift. A ubiquitous method for this is via Kalman filtering, and the use of these methods to estimate position and orientation is broadly known as attitude estimation.

A Kalman filter is a type of dynamic or recursive Bayesian estimator which utilizes prior information about a system's state as well as measurements from sensors associated with that system to constrain future system state estimates (prediction step) and correct for error between predicted and currently observed measurements (update step). Kalman filters also have high utility as a method for estimating perceptual bias, and detailed explanation of the Kalman filter will occur later in this section of the dissertation. More relevant to this section is the ability for Kalman filters to estimate orientation of an IMU

from linear acceleration and angular velocity estimated by the triaxial accelerometer and gyroscope.

Kalman filter-based inertial tracking methods have been used extensively in human motion estimation. Single and multiple IMUs have been used successfully to estimate head-based movement dynamics (MacDougall and Moore, 2005, Carriot et al., 2014; Carriot et al., 2017; Hausamann et al., 2019), as well as larger, whole-body kinematics (Bonnen et al., 2021; Matthis et al., 2022). Utility of this approach has increased with variants like the unscented or extended Kalman filter (see LaViola, 2003), which are able to incorporate non-linearities into prediction and update steps. More recently, machine learning-augmented approaches that incorporate large training sets of human locomotion data can also boost performance of the Kalman filter.

#### Outside-in Optical Tracking Systems

While not portable, another contemporary method of positional tracking uses multiple cameras facing inward on a single region of three-dimensional space, referred to as a tracking volume. While this is an extension of classic photographic tracking methods, use of multiple cameras, improved camera sampling rate, and computer vision-based techniques to identify tracking markers allow for high spatial and temporal resolution recordings of 6 DOF position of a rigid body. Rigid bodies are typically constructed and tracked through use of passive reflectors that reflect infrared light emitted from recording cameras, though other solutions exist in which active emitters produce infrared light that cameras record. With respect to human positional tracking, markers are often placed on key anatomical landmarks (head, joints, etc.): allowing for estimation of a human

skeleton and inference of anthropometric data such as gait. As a result, this method has been used to estimate head movement during various activities like locomotion over the ground (Pozzo et al., 1989) or on a treadmill (Bloomberg et al., 1992).

While confined to a calibrated tracking volume, modern marker-based optical tracking systems are often used in research and commercial applications due to their high accuracy and precision. Relative to previous methods, another benefit is that markers do not hinder human movement to the extent that mechanical methods attached to the subject may. Nonetheless, further development of these methods to develop markerless tracking methods is underway. Robust computer vision-based methods like OpenPose (Cao et al., 2021) and DeepLabCut (Mathis et al., 2018) are capable of positional estimation of organisms including humans, mice, and drosophila, though comparison of these methods with more established methods of robust positional tracking has not been conducted yet.

#### Inside-out Positional Estimation

Another contemporary method of positional tracking is through a method of sensor fusion called simultaneous localization and mapping (SLAM). SLAM-based methods are primarily used in autonomous robotics applications, and as its name implies, SLAM algorithms work to construct some representation of an agent's surroundings (mapping) as well as estimate the 6 DOF position of the agent within that representation (localization). SLAM algorithms are used to fuse information from multiple redundant and/or distinct sensors including ones discussed here (IMUs, video cameras) as well as other sensors like sonar, LIDAR, GPS, laser, sensors related to the agent's kinematics

(e.g. wheel sensor, limb sensors) and even relative strength of nearby Wi-Fi access points.

Visual-inertial SLAM (VI-SLAM) is an emerging variant of SLAM-based methods and fuses sensor data from one or more IMUs and cameras in a manner akin to Kalman filter-based estimation from one or more IMUs. After parameters representing the initial state of the system and sensors are set, predicted and updated system states are determined through positional estimates rendered from not only one or more IMUs as described previously, but also from visual odometry generated from one or more cameras facing the environment. The VI-SLAM pipeline will demarcate visual features perceived to be static as “landmarks”, and as the agent moves through the environment, movement of these landmarks relative to the agent is estimated. A survey and review of different VI-SLAM methods can be found at (Fuentes-Pacheco, Ruiz-Ascencio, and Rendón-Mancha, 2015).

VI-SLAM and SLAM methods in human movement recording presents exciting opportunities to answer questions related to head position during natural activities.

Positional estimates allow for disambiguation of the gravito-inertial acceleration experienced by IMUs and the vestibular organs. Subsequent work has begun utilizing SLAM or VI-SLAM to estimate head position or related estimates like velocity during natural activities (Sinnott et al., 2018; Hausamann et al., 2021; Kumar et al., 2022).

## Vestibular Anatomy

The following section discusses anatomy of the vestibular sense organs: the semicircular canals, the utricle, and saccule. The vestibular organs can be considered as biomechanical IMUs in a sense. The utricle and saccule are analogous to a triaxial linear accelerometer,

while the semicircular canals are similar to a triaxial gyroscope but respond to angular acceleration rather than velocity.

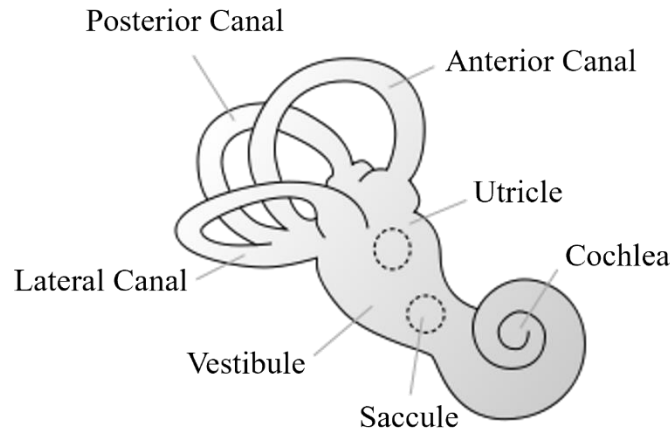
### Semicircular Canals

The semicircular canals are responsible for sensing and transducing angular acceleration of the head. Each set of vestibular organs (one near each ear) has three semicircular canals which are arranged nearly orthogonally with respect to one another. The lateral or horizontal canal is, as the name suggests, roughly in-line with the horizontal plane while the head is held in an upright position; it is chiefly responsible for transducing yaw rotations of the head. The anterior and posterior canals, in contrast to the horizontal canal, are offset approximately 45 degrees relative to the sagittal and coronal planes and are responsible for joint transduction of pitch and roll rotations. Each of these canals converges in an ampulla, a spherical structure roughly three times the diameter of a single semicircular canal which contains a crista ampullaris – a mound of hair cells with projections into the cupula. The cupula is a gelatinous glycoprotein matrix which has otoconia - small calcium carbonate crystals - embedded within it, as the gelatinous structure alone is too light to reliably deflect the hair cells of the crista ampullaris.

The hair cells deflected by the cupula consist of two main types differing in morphology, physiology, and innervation (for review, see McInturff, Burns, and Kelley, 2018; Burns and Stone, 2017). Type I hair cells are cup- or flask-shaped and have a dense envelope of calyceal nerve endings at one end, while Type II hair cells are cylindrical in shape and have relatively low-density synapses that interface with them. Functionally, Type I hair cells exhibit faster firing dynamics with greater sensitivity than Type II hair cells.

Interestingly, Type I hair cells develop in the embryo and are present at birth, while Type II hair cells develop post-natally. Each hair cell contains a single large kinocilium surrounded by a bundle of 30-100 smaller stereocilia. The arrangement of stereocilia to kinocilium on a single hair cell is not symmetric; typical arrangement consists of a kinocilium at one end, with a trail of smaller stereocilia behind it which appears as a taper wedge when viewed along the sagittal plane. Within the semicircular canals, all hair cells are oriented in the same direction.

As the head rotates, after a brief (3-5 ms) delay the endolymph within one or more semicircular canals moves and depresses their respective cupulae in one direction or another. This depression of the cupula deflects hair cells within the crista ampullaris and depending on the direction of cupular depression relative to the orientation of hair cells: either depolarizes the cells, increasing their firing rate or hyperpolarizes them, thereby decreasing their firing rate. Due to their physical arrangement, the semicircular canals function in a “push-pull” fashion: head rotation that evokes a depolarization response in one set of semicircular canals in one ear will evoke a hyperpolarization response in the other. Due to this, the nervous system can discern direction of angular acceleration between both sets of canals in all three dimensions.



**Figure 6:** Gross structure of the bony labyrinth, which contains the semicircular canals, the utricle, the saccule, as well as the cochlea: a key structure in auditory perception. The vestibule describes the region of the labyrinth shared by the utricle and saccule. Image is common domain from Wikipedia (2023).

#### Otolith Organs (Utricle and Saccule)

The otolith organs, comprised of the utricle and saccule, transduce linear acceleration rather than angular and are co-located near the semicircular canals in the vestibule, a portion of the bony labyrinth of the inner ear which is around 4 millimeters. The utricle is slightly larger than the saccule and primarily responds to head movements along the horizontal plane; the saccule, meanwhile, primarily responds to head movements along the sagittal plane. As with the semicircular canals, the otolith organs have a region wherein their sensory epithelium is contained called the macula. In contrast with the crista ampullaris, the hair cells present in the macula are oriented in multiple directions dependent on their place within the macula itself.

The macula has two main regions and consists of the central striola and the peripheral extrastriola. Like the crista ampullaris, the macula also has Type I and Type II hair cells. These have similar morphological, physiological, and functional differences as their

counterparts in the crista ampullaris and are present in both the striola and extrastriola of the macula. The striola is demarcated by a line where the orientation of hair cells suddenly changes; this line of polarity reversal allows for changes in the direction of linear acceleration to be detected. Different areas of the macula also project to different areas of the brain: afferents from the lateral part of the utricle project chiefly to the cerebellum, while those from the medial aspect of the utricle project to the vestibular nuclei in the brainstem. Like those in the crista ampullaris, the hair cells in the macula project into a gelatinous body with otoliths called the otolithic membrane, which moves as endolymph presses against it.

As the head experiences linear acceleration, the otolithic membranes above each maculae shifts concordantly after a short delay and deflects hair cells present in the maculae. If the head is tilted, gravity will cause this shift of the otolithic membrane; head movement driven by self-motion will also shift this membrane in a similar manner. Due to Einstein's equivalence principle, the movement of the otolithic membrane is due to the sum of gravitational and inertial acceleration; a major challenge the nervous system faces is segregation of gravito-inertial acceleration into separate components. There are several proposed mechanisms by which the nervous system performs this segregation; generally, these involve integration of sensory information from multiple sources and/or modalities, or assumptions about characteristics of the physical stimuli (e.g. frequency) the vestibular organs transduce. More detailed discussion of proposed modeling frameworks which work with varying utility are in (MacNeilage et al., 2008; Clark et al., 2019), as well as later in this introduction.

## Head-Centered Sensory Dimensions and Physiological Representations

This dissertation focuses primarily on characterization of two dimensions of head-located sensory information used for spatial orientation, locomotion, postural control, and other behaviors. The first is head orientation relative to gravity, while the second is vestibular heading. Both are modulated by head movement, drive correlated dynamics of extravestibular sensory information (e.g. vision, proprioception, audition), and are relatively uncharacterized during natural behaviors.

### Head Orientation

The first sensory cue focused on in this dissertation is head orientation relative to gravity (head orientation for short). Head orientation is simply the 3 DOF angular position of the head expressed in a gravity-based reference frame. For most people in their daily lives, this is changed mainly through head-on-body movement initiated by the cephalomotor plant and carried out by the 26 muscles in the neck and head. Head orientation relative to gravity can be changed in other ways, as well: active whole-body changes of position (e.g. lying down), as well as passive whole-body changes (e.g. maneuvering an aircraft) can change orientation of the head relative to gravity. This will, in turn, affect sensory information across multiple modalities used to estimate head orientation and spatial orientation perception broadly.

An important cue to estimating head orientation relative to gravity is the vestibular system. Head orientation directly affects the direction of gravity in a head-based reference frame, changing the pattern of stimulation the otolith organs receive.

Momentary changes in head orientation also excite the semicircular canals responsible for sensing angular acceleration of the head. It is postulated that the nervous system compares and integrates information from both otolith organs and semicircular canals to determine the extent to which linear acceleration experienced by the vestibular system is driven by gravity and by self-motion.

Work in mammals has observed two main regions in the cerebellum that are important for head orientation control and perception. Namely, the uvula and nodulus in multiple mammals shows sensitivity to angular acceleration of the head (Laurens, Meng, and Angelaki, 2013; Yakusheva et al., 2007; Precht et al., 1976). Damage to this area can result in hindered head orientation-dependent ocular behavior, like the vestibulo-ocular reflex (Angelaki and Hess, 1995). Of note is work from Yakusheva and colleagues, which found that Purkinje cells present in these regions selectively modulate activity driven by the semicircular canals only when there are changes in head orientation relative to gravity in *upright* orientations.

Perception of head orientation relative to gravity is not informed solely by vestibular information. Arguably the most important extravestibular information used to estimate head orientation is visual. The physical visual environment tends to demonstrate systematic structure both in natural and man-made environments, as structures (e.g. buildings, trees, rock formations) must account for gravity in some way in order to stay upright. This is generally referred to as visual polarization, and organisms can use this information to infer the direction of gravity in a purely visual manner. Any given object's visual pole can be thought of as a structure's identifiable axis that is parallel to gravity.

As the head rotates, the visual scene available for the organism to sample will change, and information in this scene including visual polarity will also change. Comparison of visual information to vestibular is likely performed by the nervous system to improve reliability of the gravity estimate by disambiguating vestibular signals (MacNeilage et al., 2007; Zupan and Merfeld, 2003).

Vision is not the only source of extr vestibular information used to estimate head orientation, however. Proprioception is likely a source of information used to estimate head orientation relative to gravity, though the importance of this sensory modality is relatively understudied. Previous work has documented compelling orientation illusions by tilting a furnished room around human subjects (Howard and Hu, 2001). They find that illusions are strongest when the visual axis of objects in the scene as well as the participant's "body axis" (Mittelstaedt, 1983) are aligned.

Efference copies, copies of motor commands generated by the nervous system and sent to other, non-motor areas (von Holst and Mittelstaedt, 1950), likely play a role in vestibular disambiguation and are thought to be used to correct and update motor error which could be driven by gravity. Efference copies of cephalomotor signals used to move the head on the body are sent to multiple cortical and subcortical regions important for vestibular and visual processing including the ventral and medial intraparietal areas (VIP, MIP; Klam and Graf, 2006) as well as first vestibular projection neurons (Cullen and Roy, 2004). Efference copies may help improve perceptual stability by allowing the nervous system an accurate account of error that is driven by self-generated behaviors (e.g. movement).

## Heading

The second dimension of head-based information focused on in this dissertation is heading. Heading refers to the instantaneous direction of self-motion in a head-based reference frame. This can be estimated by measuring the direction of the linear velocity norm across all three axes of the head-based reference frame. Several areas of the brain have been observed responding to vestibular heading stimuli.

One of the main areas is the dorsal region of the medial superior temporal area (MSTd) in primates. Individual neurons in this region demonstrate sensitivity to vestibular heading stimulation (linear self-motion). Furthermore, some of these are selective, preferring vestibular stimulation, while others are sensitive to vestibular and visual heading information (Gu et al., 2006; Gu et al., 2010). Other areas in primates include the parieto-insular vestibular cortex (PIVC; Chen et al, 2021), and earlier work temporarily deactivating this region using muscimol, a selective GABA-alpha agonist, found that heading perception was greatly diminished in macaque monkeys. Ventral interparietal area (VIP) is also implicated in vestibular heading perception broadly (Chen et al., 2011), with later work determining it plays a role in vestibular heading discrimination (Chen et al., 2013). Research in humans has attempted to identify analogous regions in humans: these include the precuneus (medial Brodmann area 7), retrosplenial complex (RSC; Brodmann areas 29 and 30), and medial parietal cortex (Brodmann area 31; for review, see Baumann and Mattingley, 2010).

As with head orientation, extr vestibular information is also used to estimate heading. Vision plays a critical role in heading perception through optic flow: the systematic

deformation of the visual environment as an organism moves through it (Lappe et al., 1999; Gibson, 1950). Optic flow can be driven by head-in-world movement, as well as eye-in-head movement. Environmental position and movement modulates optic flow patterns as well, because the depth structure present in the visual scene will cause nearer objects to move faster and further objects to move slower (Lappe et al., 1999; Longuet-Higgins and Pradzny, 1980).

Characteristics of self-motion driven by head movement will determine distinct patterns of optic flow. Assuming the eyes are fixed in the head, or that the eyes are fixated on a head-based object: simple angular head movements (e.g. a change in head yaw or pitch) will cause uniform, whole-field changes in optic flow with the appearance of flow in the opposite direction of head movement. Simple linear head movements (often driven by linear whole-body movement like locomotion) will cause distinct laminar patterns of optic flow, with minimal optic flow at the heading direction (focus of expansion), and increasingly larger patterns of optic flow with increasing eccentricity from the focus of expansion. Natural movement is often complex and combines both linear and angular components simultaneously: these can result in complex patterns of optic flow that appear to spiral around the focus of expansion (Matthis, 2022; Longuet-Higgins and Pradzny, 1980; Gibson, 1950).

As optic flow is a visual cue, movement and positioning of the eye in the head (and by proxy, the location of gaze in the world) will distinctly affect patterns of optic flow. The vestibulo-ocular reflex (VOR) serves to compensate for linear and angular head movement by moving the eyes opposite the direction of head movement, in order to

maintain fixation of gaze on an object in the visual environment. In the case of angular head movement (rotational VOR), the eyes move opposite the head on the same axis of rotation (e.g. a rightward yaw head movement of 20 degrees elicits a leftward conjugate eye movement of 20 degrees). While yaw and pitch are well-compensated for, rotational VOR cannot compensate well for head roll due to limitations of the ability of the extraocular muscles to perform eye torsion. Compensation for linear head movement (translational VOR) is performed in a similar manner, with eye movement opposite head movement, but the extent of eye rotation to compensate for linear head movement is dependent on the depth of the object a person is fixated upon. As mobile recording technology and eye-tracking methods have improved, more work has begun to focus on joint head-eye characteristics that drive optic flow at the retinal level (Longuet-Higgins and Pradzny, 1980; Matthis et al., 2022)

Other sources of extr vestibular information have not been definitively identified through physiological or behavioral methods, but it is likely that somatosensation and proprioception play some role, even if relatively small. Somatosensation and proprioception play an important role in estimation of path length, which is modulated by heading (Sun et al., 2004; Mittelstaedt and Mittelstaedt, 2001). The nervous system may use similar information or pathways to aid heading estimation; potentially, somatosensation and proprioception augment optic flow parsing by separating flow into components driven by self-motion and object motion.

Heading perception overall appears highly multimodal. Many areas of the brain which show activity elicited by vestibular heading stimuli show similar activity evoked by optic

flow patterns. VIP (Zhang and Britten, 2011; Chen et al., 2011) and PIVC (Chen et al., 2021) both show activity elicited by optic flow fields. Interestingly, primary motor cortex does as well (Merchant et al., 2001), lending credence to the idea that somatosensation and/or proprioception may aid in optic flow parsing. Further support is rendered by the observation that areas like PIVC and VIP demonstrate connectivity to proprioceptive and somatosensory regions in the brain (Chen et al., 2021; Zhang and Britten, 2011).

## Head Orientation and Heading Perception: Previous Research and Observed Bias

### Previous Head Orientation Research

While research documenting characteristics of natural stimuli is ubiquitous in perception research, focus on natural characteristics of head movement and position is somewhat lacking. While there is a body of research investigating dynamics of linear acceleration of the head (see Carriot et al., 2017), little research investigates characteristics of head movements made during natural activities. Characteristics of head orientation relative to gravity is not well-documented in the current research literature (MacNeilage, 2020), as IMU-based methods face difficulties when disambiguating acceleration driven by gravity and translation. Most related work focuses on vestibular stimulation experienced by humans completing natural activities (Carriot et al., 2014, Carriot et al., 2017). Notable exceptions include studies performed by Schwabe and Blanke in 2008, and Willemsen and colleagues in 2022 measuring aspects of head orientation using relatively unobtrusive equipment.

In both cases, researchers assign a group of natural activities for participants to complete for a short period of time while tracking head orientation. Willemsen and colleagues use a series of IMUs placed on the pelvis, chest, arms, and head to estimate head orientation as they perform five natural activities for two minutes each: standing, sitting, walking, climbing and descending stairs, and running. They show that head roll during each of these activities is distributed in a non-Gaussian manner, with high amounts of kurtosis (Willemsen et al., 2022).

In contrast, Schwabe and Blanke utilize outside-in optical tracking to estimate orientation of the head along the pitch axis during three natural behaviors for 4-5 minutes each. Participants completed a “standing” activity where they visually inspected the experimental room without locomoting or moving their lower body, a “waiting” activity where they paced around the motion capture volume within the experimental room, and an “action” activity where they pantomimed tennis-like movements within the motion capture volume. Akin to Willemsen and colleagues, data from Schwabe and Blanke suggest that head pitch is not normally distributed either with some amount of asymmetry present across all subjects and tasks. Furthermore, their data appear to show task-dependent differences in statistics, with the head pitch distribution derived from “standing” activity having higher variance than either other activity.

Taken together, these studies suggest that head roll and head pitch do not follow Gaussian distributions. This has important implications for probabilistic modeling approaches which assume various model parameters, including those representing head roll or pitch directly, to be Gaussian distributed. A potential caveat to these studies,

however, is that they rely on data from people performing prescribed activity; it may not be representative of true natural behaviors. At time of writing, no research exists characterizing head roll and head pitch during unprescribed natural behaviors.

### Head Orientation Bias

Head orientation perception bias is robust across individuals and was first observed in the late 1800s with seminal work conducted by Aubert (1861). In this study, Aubert noted that increasing roll of the entire body (including head), then asking subjects to report their own orientation by aligning a visual line to vertical absent any other visual cues resulted in increasing underestimation of one's own orientation along the roll axis. In the 160 years since this work, this phenomenon has become known as the Aubert effect (or A-effect) and titrated at increasing ranges and increments of roll tilt. This body of work has shown that, for tilt ranges greater than 60 degrees, this A-effect is robustly observed; roll tilts of less than 60 degrees can result in little error or sometimes overestimation of one's own orientation (Müller, 1916). This phenomenon of overestimation is referred to as the E-effect, and in the context of roll perception has small magnitude relative to the A-effect (de Vrijer, Medendorp, and van Gisbergen, 2009).

Although generally understudied, pitch perception bias has also been previously observed and exhibits patterns akin to roll perception in some ways (DeLage, 1886; Cohen and Larson, 1974). Pitch perception bias differs from roll perception bias in several important ways, however. First, pitch perception bias is generally smaller in terms of magnitude than that observed for roll. Pitch perception bias has relatively small magnitude, ranging from  $\pm 5^\circ$  between fully supine (face up), upright, and fully prone (face down) positions

(Cohen and Larson, 1974). Magnitude for roll bias over a similar range of eccentricities (left-ear-down to upright to right-ear-down) is on the order of  $\pm 20^\circ$  degrees (de Vrijer et al., 2009).

Second, patterns of underestimation (A-effect) and overestimation (E-effect) present in roll perception bias are not present in pitch perception bias. Where roll is overestimated at small angles close to upright and underestimated at angles far from upright, pitch is underestimated most at near upright angles, with decreasing estimation error until nearly prone or supine. At these near prone or supine pitch orientations, overestimation is observed (Cohen and Larson, 1974). Third, magnitude of error is greatest at near upright pitch angles ( $\pm 15^\circ$ ), while magnitude of error is greatest at high roll angles ( $> 60^\circ$ ) that are far from upright.

Finally, patterns of pitch perception bias appear more asymmetric than patterns of roll perception bias. In terms of magnitude, bias for supine pitch angles appears roughly twice as large as bias for prone pitch angles (Cohen and Larson, 1974). In terms of alternation between underestimation and overestimation, overestimation appears to occur at roughly 75 degrees and greater for supine angles, while for prone angles it occurs at roughly 60 degrees. Explanations for why pitch perception bias differs dramatically from roll perception bias have not been proposed, and indeed pitch perception broadly has been understudied since the 1980s. Head yaw perception is generally not studied, as the major vestibular cue to head orientation, change in the direction of gravity, will remain constant with pure yaw movements.

### Previous Heading Research

Along the lines of research investigating characteristics of human head orientation during everyday activities, natural characteristics of human vestibular heading is also understudied. Most research investigating natural statistics of human heading focuses on the visual component of heading perception: optic flow. This follows an extensive line of work attributed to a few key researchers in the 1940s and 1950s (Gibson, 1950).

Relatively contemporary work focuses on how optic flow is modulated by joint head and eye movement, as well as environmental characteristics (Longuet-Higgins and Pradzny, 1980; Matthis et al., 2022). Relatively recent research has sought to characterize human vestibular heading during natural behavior, but this was in a few subjects during a stereotyped walking task (Sinnott et al., 2018). To date, no other research has published characteristics of human vestibular heading stimuli during natural activity.

### Vestibular Heading Bias

Heading perception bias has been studied since the 1990s, though early work is characterized by only measuring across a limited range of forward heading azimuth and elevation directions relative to straight ahead in the visual modality ( $\pm 40^\circ$  in Warren and Kurtz, 1992;  $\pm 45^\circ$  in D'Avossa and Kersten, 1996). More contemporary work seeks to measure heading perception bias across the entire range of possible directions and through use of both visual stimulation (optic flow) and vestibular stimulation (translation of subject). Work by Cuturi and MacNeilage in 2013 focused on measuring heading perception bias across the entire horizontal (azimuthal) plane, while work by Crane a year later focused on measuring heading perception bias across what he described as the

“vertical” planes: the sagittal (corresponding with elevation) and the coronal (Crane 2014), which is analogous to all elevation angles at purely leftward and rightward azimuth angles.

With respect to heading azimuth, both vestibular and visual stimulation evoke a pattern of oscillation between initial underestimation at eccentricities between straight-forward and purely rightward or leftward, then change to overestimation at backward heading angles (Cuturi and MacNeilage, 2013). This pattern is stronger for visual heading stimulation, with increased bias magnitude ( $\pm 15^\circ$ ) relative to vestibular heading stimulation ( $\pm 7^\circ$ ). Similar trends are observed for vestibular heading elevation perception (Crane 2014).

### Sensory Processing Models

A major approach used to understand sensory perception across multiple senses is to model observed patterns of perceptual bias using probabilistic modeling. Many approaches have been used to understand human head orientation and heading perception, as reliable patterns of bias have been repeatedly observed, and these approaches fall under a few broad categories. Internal and observer models propose that the nervous system represents physical laws present in our world in some manner and compares observed stimulus values with stimulus values expected as a result of this representation. Bayesian models utilize Bayes' law to parameterize an organism's previous experience with stimuli of interest, as well as uncertainty both at the level of the organism and the stimulus itself to investigate how uncertainty in sensory estimates change as a result of statistical structure of stimuli and experience.

An extension of these models, termed recursive or dynamic Bayesian estimators, can be viewed as a synthesis of the previous two approaches. These estimators have an internal framework which represents physical laws or environmental characteristics as well as relationships between multiple sensory modalities which is set to some initial state. This framework is constrained by parameters represented by probability distributions which can modulate levels of uncertainty in the estimate and represent a base level of uncertainty analogous to an organisms' prior experience. Then, as sensory information enters this framework, the system updates in a recursive manner by comparing current sensory information with its previous state, generating a prediction used in the next iteration.

This section serves as an introduction to each of these three classes of modeling framework. More extensive reviews of these models have been conducted as they are broadly useful across multiple sensory modalities and perceptual phenomena. Reviews of these models in the context of self-motion perception and spatial orientation perception can be found at (MacNeilage, 2020; Clark et al., 2019; MacNeilage et al., 2008). This section will review these modeling frameworks in the context of vestibular processing.

### Internal Observer Models

Some of the earliest accounts of how the nervous system parses sensory information is through an internal model, and vestibular information sensed by the otoliths is no different. Put generally, this approach posits that the nervous system has some internal representation of physical principles or the physical world, and compares estimates generated from sensory information to estimates that this physical model predicts. If there

is any discrepancy between the two, then the nervous system may update itself to better align estimates the vestibular system generates to ones generated by this physical representation.

Internal observer models have been used to estimate both motor and sensory dynamics (MacNeilage et al 2008; Tin and Poon, 2005): both in forward (where sensory information informs subsequent motor action) and backward directions (where motor output informs expected stimulation of sensory organs). Of greatest relevance to this dissertation is a subset of this approach, which proposes that the nervous system represents physical laws of the environment. The otolith organs responsible for sensing linear acceleration can be thought of as biological accelerometers; due to Einstein's equivalence principle, however, all accelerometers are only capable of sensing the sum force of linear acceleration ( $f$ ) driven by gravity ( $g$ ) and by inertia ( $a$ ):

$$f = g - a$$

This is broadly referred to as gravitoinertial (GI) ambiguity, and computational methods by which humans and other organisms parse inertial and gravitational components from summed GI force is a topic of interest within the field. In a normally functioning organism, it is rarely the case that inertial force is zero. To try and recover orientation relative to gravity ( $\dot{g}$ , the derivative of  $g$  with respect to time) from summed GI, it follows that the nervous system may leverage how semicircular canal estimates of angular velocity ( $\omega$ ) change with head rotation:

$$\dot{g} = -\omega \times g$$

While a simple example, there is no prescribed constraint with respect to how many inputs, outputs, or processing steps an internal observer model must have. Furthermore, modeling of vestibular sensory dynamics, given that the vestibular organs are only capable of estimating linear and angular acceleration, is often nonlinear. Subsequently, most internal model-based approaches of vestibular processing are relatively complex to account for these nonlinearities. Extensive work from the Merfeld research group has focused on development of internal models in non-human primates (Merfeld, Young, Oman, Shelhamer 1993; Merfeld, 1995) as well as humans (Merfeld, Zupan, and Peterka, 1999; Zupan and Merfeld, 2003; Zupan and Merfeld, 2005). This work is preceded by work conducted in the 1980s by Mittelstaedt detailing mechanisms by which the nervous system might represent a body-centric axis denoted as “the idiotropic vector” (1983), and extensions of this work detail how this body axis may interact with head-centered axes (1986, 1989) and visual information.

While these models can predict perceptual phenomena well, they have several shortcomings. First, they often have multiple free parameters that are optimized to bolster model performance. Second, they do not provide a method of accounting for noise or uncertainty at the level of input or output. Overall, this results in reduced relevance for explaining biological processes; both a stimulus and its percept are assumed to have some level of noise and tuning of free parameters seldom occurs with any sort of biological constraint.

### Static Bayesian Models

Another approach to model human perception is by leveraging Bayes' Law to predict perception based off an organisms' previous experience as well as a current stimulus that they are attempting to perceive. These are broadly referred to as Bayesian models.

Bayes's law states that a posterior probability distribution ( $P(A|B)$ ) is equal to the product of the likelihood distribution ( $P(B|A)$ ) and the prior distribution ( $P(A)$ ) divided by the marginal distribution ( $P(B)$ ).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

As the marginal distribution only serves to scale the numerator, this can be simplified further to:

$$P(A|B) \propto P(B|A)P(A)$$

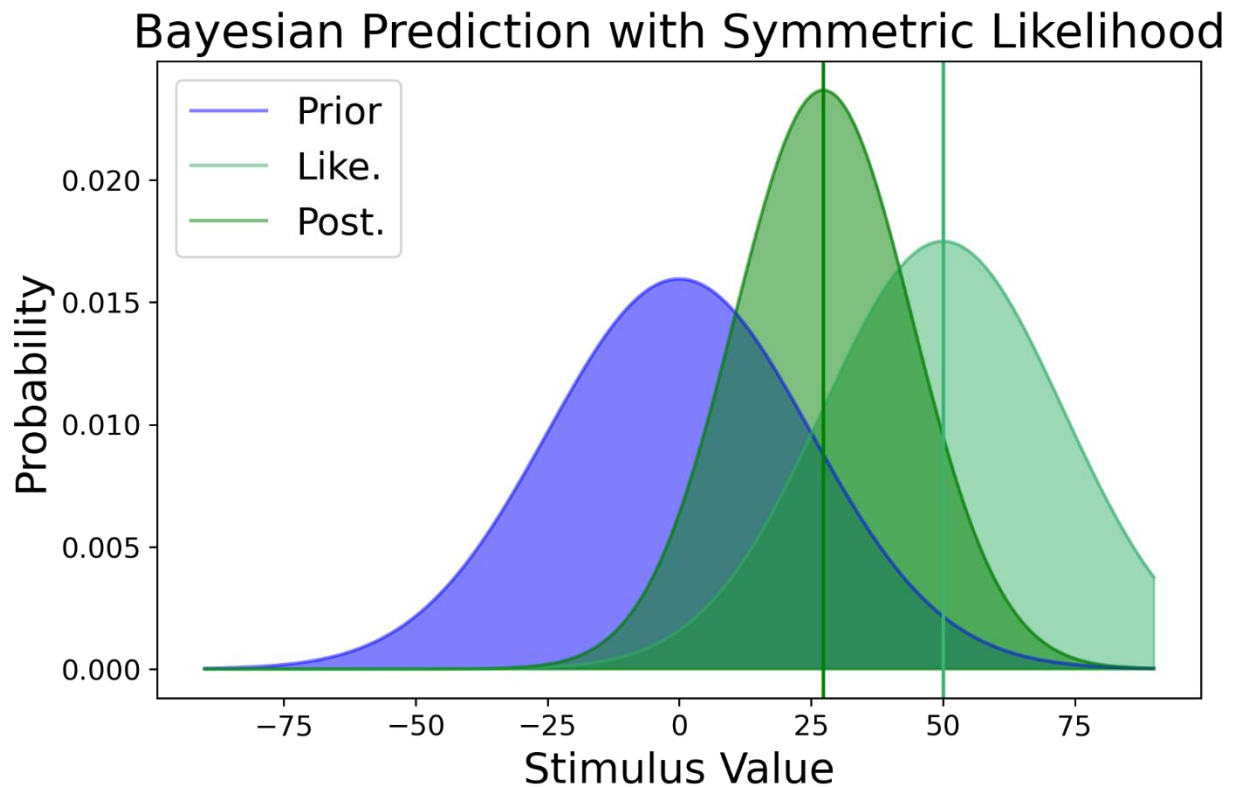
Here, some posterior distribution is proportional to the product of likelihood and prior distributions. In the context of sensory perception this is often re-written using theta ( $\theta$ ) and theta-hat ( $\hat{\theta}$ ) as terms representing sensory information and sensory estimation, respectively:

$$P(\theta|\hat{\theta}) \propto P(\hat{\theta}|\theta)P(\theta)$$

In this case, the likelihood distribution can be considered the sensory evidence available to an organism; some true stimulus value with an amount of noise. The prior, meanwhile, represents all previous experience an organism has with this type of sensory evidence.

The resultant posterior distribution represents the perceived stimulus value. Most approaches using a static Bayesian model assume these distributions to follow a Gaussian

distribution, with variance of that distribution representing noise present in sensory information or estimation of that information. Typically, a central tendency (most often mode or mean) of the posterior distribution is used as a point estimate for a given stimulus value.



**Figure 7:** Static Bayesian model with Gaussian prior, likelihood, and posterior distributions. The prior is centered on 0, while the likelihood has a mode of 50 (representing a stimulus value of 50 with some amount of noise or uncertainty, light green line). The mean of the resultant posterior, used as a point estimate for perception of the stimulus value denoted by the likelihood, is approximately 28. As this value is somewhere between the modes of the prior and likelihood, this is termed “attractive” bias (relative to the prior), representing an underestimation of the true stimulus value represented by the likelihood.

Use of static Bayesian models are ubiquitous in sensory perception research across multiple individual and integrated sensory modalities including vision (MacNeilage et al.,

2007), the vestibular system (de Vrijer et al., 2009), audition (Battaglia, Jacobs, and Aslin, 2003), and somatosensation (Ernst and Banks, 2002). Focusing on spatial orientation perception, extensive work by Medendorp and colleagues has focused on the problem of static head roll perception (Vingerhoets et al., 2008; de Vrijer et al., 2008; de Vrijer et al., 2009; Willemsen et al., 2022).

Recent approaches have extended this model by adding an efficient encoding stage which constrains noise on the likelihood (Wei and Stocker, 2015). This stage is intended to represent how the nervous system encodes information in a non-uniform, but statistically optimal manner. Noise on the likelihood is assumed to be constrained by the shape of the prior distribution driven by natural stimulus statistics, and maximizes mutual information shared between some sensory measurement and a true stimulus variable:

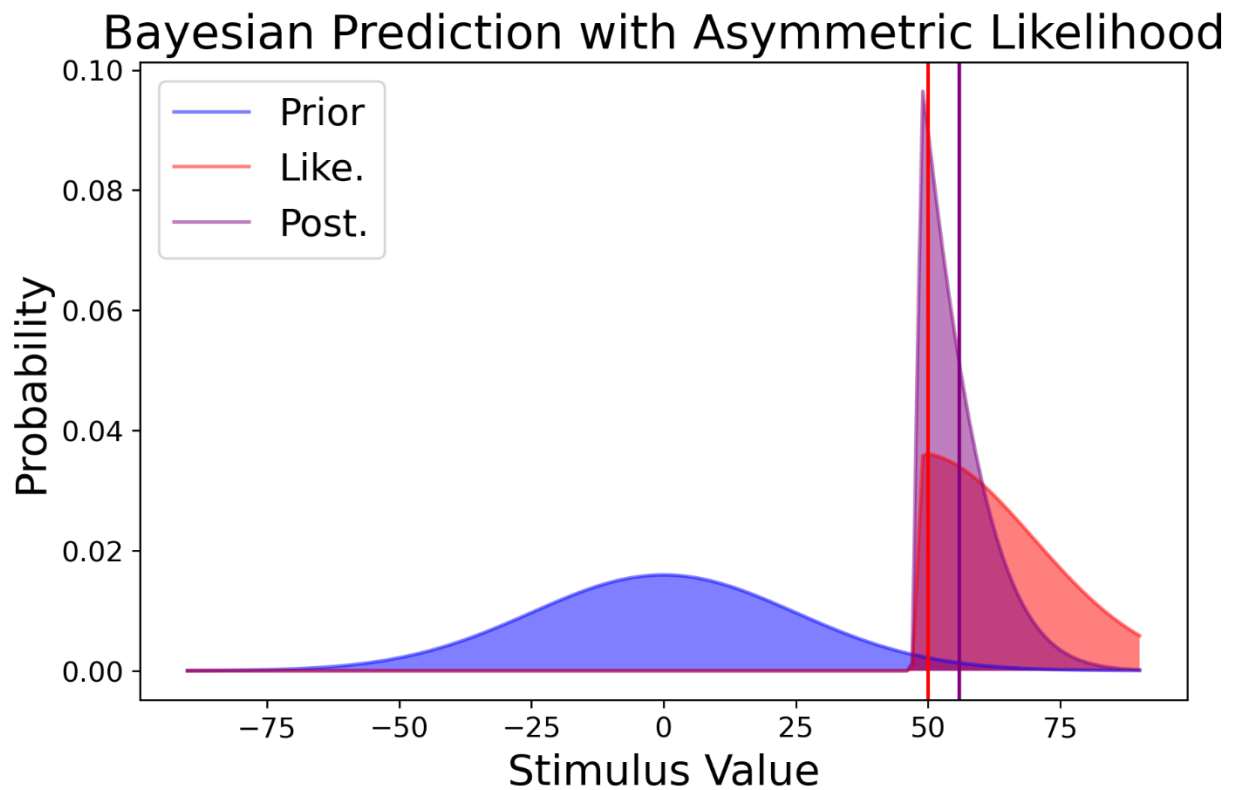
$$P(\theta) \propto \sqrt{J(\theta)}$$

Mutual information is represented through Fisher information, which reflects encoding accuracy and coding resources allocated towards representation of some stimulus value ( $\sqrt{J(\theta)}$ ). This accuracy and allotment of resources is proportional to the prior probability distribution which describes all possible stimulus values ( $P(\theta)$ ). As a result, the more frequent the stimulus value is in the prior, the more resources are allocated to it, and the greater the precision of its subsequent encoding.

Additional constraint is necessary to determine noise on the likelihood function, however. Wei and Stocker add a second assumption made in this model which is based on a mapping function that transforms a stimulus value out of a stimulus space and into a

sensory space based on the cumulative probability distribution derived from the prior. This sensory space is assumed to have uniform Fisher information across all stimulus values, and subsequently noise around the expected likelihood function in this space is homogeneous and symmetric about the stimulus value.

Greater detail on construction and derivation of this model can be found in (Wei and Stocker 2015), with extensions of this model to predict a variety of sensory phenomena at (Wei and Stocker 2017; Ni and Stocker, 2023). The upshot of this model is that likelihoods can have noise modulated in one of two ways, and subsequently likelihoods for a stimulus value that is relatively infrequent in the prior can be asymmetric. This can result in conditions or patterns of predicted perception bias that change between over- and underestimation.



**Figure 6:** Efficient Bayesian model with Gaussian prior (blue), asymmetric likelihood (red), and subsequently asymmetric posterior (purple) distributions. The prior has a max of 0, while the likelihood's mode represents a stimulus value of 50 (red line). While modes of likelihood and posterior are closely aligned, the combination of a skewed posterior and use of its mean (approximately 56, purple line) as a point estimate results in overestimation of the stimulus value. This is often termed “repulsive” bias, as the mean of the posterior appears to be “repelled” away from the likelihood and mode.

There are several criticisms towards the use of static Bayesian models. First, they rely upon parameterization of terms representing an organism's previous experience with a given stimulus type (the prior) and the probability of an organism estimating a stimulus value with some amount of noise given the probability of that stimulus value occurring at all (the likelihood). Empirically observed values of these parameters are not always available, and a common compromise is to use an unbiased Gaussian distribution as a prior. Variability of the prior and likelihood, representing range of the organisms' previous experiences are left free, allowing fit to be optimized.

Use of Gaussian distributions is mathematically convenient, and a good intuition as many phenomena follow Gaussian distributions: but sensory phenomena may not always be Gaussian distributed, as indicated by the limited work measuring natural head orientation (Willemsen 2022, Schwabe and Blanke, 2008). The use of Gaussian distributions also has a knockdown effect on subsequent posteriors generated. Posteriors are always Gaussian, and generally lie somewhere in the stimulus space between the prior and the likelihood: always resulting in attractive bias. Furthermore, static Bayesian models are not capable of predicting dynamic signals or phenomena. While quite capable of modeling perception of stimuli in static situations (e.g. perception of static tilt, constant luminance, etc.), static Bayesian models are not capable of accounting for changes in stimulus magnitude or uncertainty over time.

### Dynamic/Recursive Bayesian Estimators

To begin addressing shortcomings of static Bayesian models, an extension of Bayesian methods can be broadly described as dynamic or recursive Bayesian estimators. One of the most prominent approaches in this category of model is the Kalman filter. Linear Kalman filters (the focus of this section) use a series of measurements from one or more sensors to estimate a system's otherwise unobservable state over time with some amount of uncertainty. It is capable of fusing information from multiple sources to improve state estimation, making it broadly useful to understand unimodal and multimodal sensory systems which may integrate information from multiple sources.

The following example works through a linear Kalman filter with equations derived and adapted from (Faragher, 2012). In a linear Kalman filter, two separate components represent the overall state of some system, and the measurement of that systems' state. The system's state is assumed to evolve from its prior state and is represented with the following equation:

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t + \mathbf{w}_t$$

In this equation, the unobservable state of a system ( $\mathbf{x}_t$ ) at time ( $t$ ) is estimated by adding the product of the state transition matrix ( $\mathbf{F}_t$ ) to the system's previous state ( $\mathbf{x}_{t-1}$ ), the product of the control input matrix ( $\mathbf{B}_t$ ) and control input parameters ( $\mathbf{u}_t$ ), and a vector representing process noise for each parameter in the system's state ( $\mathbf{w}_t$ ). Noise is assumed to be unbiased and drawn from a multivariate normal distribution with covariance denoted by a covariance matrix ( $\mathbf{Q}_t$ )

Control inputs can be broadly defined as any sort of input, whether that be one directly entered through a control interface, or measurement from a sensor that is otherwise not part of the state vector. Absent either type of these inputs (i.e. a system comprised of sensors used to estimate the system's state), this term can be eliminated entirely. Active monitoring or measurement of system parameters is also a key component, and can be completed using the following equation:

$$\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t$$

Here, measurements of the system ( $\mathbf{z}_t$ ) are completed by transforming parameters representing the system's state into a measurement domain ( $\mathbf{H}_t$ ) with some amount of noise ( $\mathbf{v}_t$ ). Noise is assumed to be unbiased and Gaussian, with covariance determined by a separate covariance term ( $\mathbf{R}_t$ ). Kalman filters rely on two steps to estimate state and parameters of the model.

The first consists of a **prediction** step, where:

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \hat{\mathbf{x}}_{t-1|t-1} + \mathbf{B}_t \mathbf{u}_t$$

Here, the system's current state given the previous state ( $\mathbf{x}_{t|t-1}$ ) is determined through combining the previous state's estimate with the previous system state's parameters applied to said estimate ( $\mathbf{F}_t \hat{\mathbf{x}}_{t-1|t-1}$ ) and any control inputs with associated parameters. As the system state is not directly observable, it is estimated with some amount of uncertainty. This uncertainty can be represented with a Gaussian probability distribution function, with variance associated with the state's estimate given by an associated covariance matrix ( $\mathbf{P}_{t|t-1}$ ):

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^T + \mathbf{Q}_t$$

The second consists of an **update** step, where:

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t (\mathbf{z}_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1})$$

The updated estimate of the current state ( $\hat{\mathbf{x}}_{t|t}$ ) is determined by comparing the last predicted state ( $\hat{\mathbf{x}}_{t|t-1}$ ) and the measurement residual ( $\mathbf{z}_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1}$ ). The updated state estimate has variance determined by:

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t \mathbf{H}_t \mathbf{P}_{t|t-1}$$

Both portions of the update step involve  $\mathbf{K}$ , representing Kalman gain. This value can be dynamically changed based on:

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^T + \mathbf{R}_t)^{-1}$$

Kalman filters have several criticisms, however. As the name suggests, linear Kalman filters assume the process or system being modeled is linear. Previous work suggests that certain aspects of gravito-inertial disambiguation may be highly nonlinear (MacNeilage et al., 2008), leading to poor performance or poor explanatory power when using this approach for modeling sensory processing. This is alleviated by various extensions of the linear Kalman filter to work for non-linear systems, the two most prominent being the extended and unscented Kalman filters.

Another criticism of the Kalman filter for modeling sensory processing is the use of one or more Kalman gains which affect update steps of the model. While these are often calculated using an optimized method which reduces uncertainty of the state

representation, they are ultimately arbitrary and any value a Kalman gain takes is not constrained by any known biological process. Other parameters governing noise at various stages of the filter ( $\mathbf{Q}_t, \mathbf{R}_t$ ) and used to inform calculation of Kalman gain are often statically set by the experimenter as well.

Finally, the Kalman filter does not consider all previous measurement and state estimates; it only considers the current and most recent estimates. It is likely that biological systems consider all previous experience of the organism in some way, which is not well-represented in current Kalman filtering methods. Nonetheless, Kalman filter-based methods perform well when predicting perception both within the visual and vestibular systems alone, as well as in larger models which incorporate information from multiple sensory modalities (for review, see Clark 2019).

Another related approach separate from the Kalman filter is the particle filter. A particle filter, also known as a sequential Monte Carlo method, operates in a manner like a Kalman filter. Both Kalman and particle filters attempt to estimate an otherwise unobservable state of a system in a recursive manner. The following example works through a non-linear estimation problem without control inputs using particle filtering, with equations derived from (Elfring, Torta, and van de Molengraft, 2021).

Like the Kalman filter, there are two major components. There is a process model which associates a system's current state ( $\mathbf{x}_t$ ) with its previous state ( $\mathbf{x}_{t-1}$ ) using a nonlinear function ( $\mathbf{f}_t$ ):

$$\mathbf{x}_t = \mathbf{f}_t(\mathbf{x}_{t-1}, \mathbf{v}_{t-1})$$

As well as a measurement model, which associates the current state with an expected or predicted measurement ( $\mathbf{z}_t$ ) using another nonlinear function ( $\mathbf{h}_t$ ):

$$\mathbf{z}_t = \mathbf{h}_t(\mathbf{x}_t, \mathbf{n}_t)$$

Parameters associated with process noise ( $\mathbf{v}_t$ ), and measurement noise ( $\mathbf{n}_t$ ) are all present. To better model uncertainty in state estimation, various parameters can be represented with probability distribution functions. Bayes' rule can then be used to recursively update an estimated state given previous measurements, predictions, and state estimates. The state estimate, referred to as a posterior distribution in Bayesian terms, is estimated with the following probability distribution function:

$$p(\mathbf{x}_t | \mathbf{z}_{1:t})$$

Use of the colon operator here indicates that the probability distribution representing the system's current state is conditional upon the sequence of all previous measurement noise states. The sequence of all previous states and measurements is indicated by:

$$p(\mathbf{x}_t | \mathbf{x}_{1:t-1}, \mathbf{z}_{1:t-1}) = p(\mathbf{x}_t | \mathbf{x}_{t-1})$$

$$p(\mathbf{z}_t | \mathbf{x}_{1:t}) = p(\mathbf{z}_t | \mathbf{x}_t)$$

These are Markov processes which assume that the complete sequence of all previous states or measurements is summarized sufficiently with only the previous step's estimated state or measurement. The model then updates in a two-stage sequence. The first stage involves computation of the prior and, analogous to the Kalman filter, is the **prediction** step:

$$p(\mathbf{x}_t | \mathbf{z}_{1:t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{z}_{1:t-1}) d\mathbf{x}_{t-1}$$

As the equation and name implies, this is the current predicted system state given all previous measurements. The likelihood, representing the probability of a current measurement given the current predicted state, is used with the prior and marginal to calculate the posterior representing the updated state estimate, analogous to the Kalman filter's **update** step:

$$p(\mathbf{x}_t | \mathbf{z}_{1:t}) = \frac{p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{1:t-1})}{p(\mathbf{z}_t | \mathbf{z}_{1:t-1})}$$

The marginal distribution ( $p(\mathbf{z}_t | \mathbf{z}_{1:t-1})$ ) is effectively the probability of the current measurement and can be computed via:

$$p(\mathbf{z}_t | \mathbf{z}_{1:t-1}) = \int p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{1:t-1}) d\mathbf{x}_t$$

Extensive review of particle filtering broadly, as well as derivation of these equations can be found in (Elfring, Torta, and van de Molengraft, 2021). Particle filters also have several criticisms, which revolve primarily around their computational complexity: particularly for higher dimensional system states. Particle filters are also particularly vulnerable to poor optimization if the initial particles do not describe the system state well.

Relative to linear Kalman filters, particle filters have several advantages in the context of perceptual modeling. First, the assumption of Gaussian distribution of noise is relaxed. Second, process and measurement models are non-linear rather than linear, which more

faithfully recreates vestibular processing dynamics. Despite these advantages, the particle filter is relatively underused in the context of vestibular processing (Laurens and Droulez, 2007; Karmali and Merfeld, 2012).

## Scope

One of the main aims of this dissertation is to observe and quantitatively characterize head movement during natural, unconstrained activities. To date, little research has been conducted aimed at measuring head movement during natural activities. Of the extant research, most, if not all consists of participants completing activities prescribed by experimenters. No study, to our knowledge, observes head movement in freely moving humans completing unprescribed activities. Two projects discussing collection of this data over long periods of recording are presented and discussed in the following chapters.

Another aim of this dissertation is to use natural head movement collected in this manner to constrain models of sensory processing. As previously mentioned, perception of one's head orientation relative to gravity, and one's heading direction are systematically biased in different ways. Various modeling approaches have been used to explain or understand these patterns of bias, and many rely directly on some theoretical representation of an organism's total previous experience. Using natural head movements that human subjects make as an empirical representation of their previous experience, we hope to constrain Bayesian models of sensory processing that leverage prior experience in this way in a biologically valid manner, improving their explanatory power.

The second chapter of the dissertation describes observation of natural head orientation relative to gravity in freely moving humans, and use of that data to constrain multiple

models of vestibular sensory processing. We set out to measure head orientation in a group of participants wearing head-tracking equipment for five-hour recording sessions. Distributions of head roll and head pitch relative to gravity are then transformed to probability distribution functions using kernel density estimation. Subsequently, they are used as priors in two variants of a standard Bayesian observer model. The predictions these models generate are then compared to pre-existing psychophysical work measuring head orientation perception bias.

From the same dataset we can compute a person's heading from moment to moment. The third chapter focuses on measurement and use of this data in modeling heading perception bias. We iterate on the previous project by using both a standard Bayesian model as well as a more contemporary efficient Bayesian model. Due to heading perception exhibiting patterns of attractive and repulsive bias, we anticipate that performance of the efficient Bayesian model will be better than the standard Bayesian model.

The fourth chapter of this dissertation will focus on discussing applications of these data to identify and answer open research questions in sensory processing. We identify potential lines of research for both head pitch and vestibular heading perception based off of two observations: head pitch exhibits marked asymmetry with high levels of individual variability, and statistics of heading azimuth appear to change dramatically with speed. Experiments are proposed which probe the effect of potential individual differences in experience (through pitch perception) as well as task- or mode-dependent sensory processing (through vestibular heading perception). Applications of these data outside

psychological research are also discussed: namely, utility of these data for applications in virtual and/or augmented reality, as well as clinical diagnostics and human activity recognition.

## References

- Angelaki, D. E. and Hess, B. J. M. (1995). Inertial representation of angular motion in the vestibular system of rhesus monkeys. II. Otolith-controlled transformation that depends on an intact cerebellar nodulus. *Journal of Neurophysiology*, 73 (5), 1729-1751.
- Aubert, H. (1861). Eine scheinbare bedeutende Drehung von Objecten bei Neigung des Kopfes nach rechts oder links. *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin*, 20, 381-393.
- Bartz, A. E. (1966). Eye and head movements in peripheral vision: Nature of compensatory eye movements. *Science*, 152 (3729), 1644-1645
- Battaglia, P. W., Jacobs, R. A. and Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America, A: Optics, Image Science & Vision*, 20 (7), 1391-1397.
- Baumann, O. and Mattingley, J. B. (2010). Medial parietal cortex encodes perceived heading direction in humans. *Journal of Neuroscience*, 30 (39), 12897-12901.
- Bernardes, E. and Viollet, S. (2022). Quaternion to Euler angles conversion: A direct, general, and computationally efficient method. *PLoS ONE*, 17 (11), 1-13.
- Bloomberg, J. J., Reschke, M. F., Hubner, W. P. and Peters, B. T. (1992). The effects of target distance on eye and head movement during locomotion. *Annals of the New York Academy of Sciences*, 656 (1), 699-707.
- Bonnen, K., Matthis, J. S., Gibaldi, A., Banks, M. S., Levi, D. M. & Hayhoe, M. (2021). Binocular vision and the control of foot placement during walking in natural terrain. *Scientific Reports*, 11 (20881), 1-12.
- Burns, J. C. and Stone, J. S. (2017). Development and regeneration of vestibular hair cells in mammals. *Seminars in Cell & Developmental Biology*, 65, 96-105.
- Cao, Z., Hidalgo, G., Simon, T., Wei, S. and Sheikh, Y. (2021). OpenPose: Realtime multi-person 2d pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43 (1), 172-186.
- Carriot, J., Jamali, M., Chacron, M. J. and Cullen, K. E. (2014). Statistics of the vestibular input experienced during natural self-motion: Implications for neural processing. *The Journal of Neuroscience*, 34 (24), 8347-8357.
- Carriot, J., Jamali, M., Cullen, K. E. and Chacron, M. J. (2017). Envelope statistics of self-motion signals experienced by human subjects during everyday activities: Implications for vestibular processing. *PLoS One*, 1-24.
- Chen, A., DeAngelis, G. C. and Angelaki, D. E. (2011). Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. *The Journal of Neuroscience*, 31 (33), 12036-12052.

- Chen, A., DeAngelis, G. C. and Angelaki, D. E. (2013). Functional specializations of the ventral intraparietal area for multisensory heading discrimination. *The Journal of Neuroscience*, 33 (8), 3567-3581.
- Chen, A., Zeng, F., DeAngelis, G. C. and Angelaki, D. E. (2021). Dynamics of heading and choice-related signals in the parieto-insular vestibular cortex of macaque monkeys. *The Journal of Neuroscience*, 41 (14), 3254-3265.
- Clark, T. K., Newman, M. C., Karmali, F., Oman, C. M., and Merfeld, D. M. (2019). Chapter 5 – Mathematical models for dynamic, multisensory spatial orientation perception. In Ramat, S. and Shaik, A. G. (Eds.), *Mathematical Modeling in Motor Neuroscience: State of the Art and Translation to the Clinic. Ocular Motor Plant and Gaze Stabilization Mechanisms*, volume 248 of *Progress in Brain Research* (pp. 65-90). Elsevier.
- Cohen, M. M. and Larson, C. A. (1974). Human spatial orientation in the pitch dimension. *Perception and Psychophysics*, 16 (3), 508-512.
- Collewijn, H. (1977). Eye- and head movements in freely moving rabbits. *Journal of Physiology*, 266, 471-498.
- Crane, B. T. (2014). Human visual and vestibular heading perception in the vertical planes. *Journal of the Association for Research in Otolaryngology*, 15, 87-102.
- Cullen, K. E. and Roy, J. E. (2004). Signal processing in the vestibular system during active versus passive head movements. *Journal of Neurophysiology*, 91, 1919-1933.
- Cuturi, L. F. and MacNeilage, P. R. (2013). Systematic biases in human heading estimation. *PLoS One*, 8 (2), 1-11.
- D'Avossa, G. and Kersten, D. (1996). Evidence in human subjects for independent coding of azimuth and elevation for direction of heading from optic flow. *Vision Research*, 36 (18), 2915-2924.
- de Vrijer, M., Medendorp, W. P. and van Gisbergen, J. A. M. (2008). Shared computational mechanism for tilt compensation accounts for biased verticality percepts in motion and pattern vision. *Journal of Neurophysiology*, 99 (2), 915-930.
- de Vrijer, M., Medendorp, W. P. and van Gisbergen, J. A. M. (2009). Accuracy-precision trade-off in visual orientation constancy. *Journal of Vision*, 9 (2), 1-15.
- DeLage, Y. (1886). Études expérimentales sur les illusions statiques et dynamiques de direction pour servir à déterminer les fonctions des canaux demicirculaires de l'oreille interne. *Archives de Zoologie Experimentale et Generale*, 4, 535-624.
- Dunlap, K. and Mowrer, O. H. (1930). Head movements and eye functions of birds. *Journal of Comparative Psychology*, 11 (1), 99-113.
- Elfring, J., Torta, E. and van de Molengraft, R. (2021). Particle filters: A hands-on tutorial. *Sensors*, 21 (438), 1-28.
- Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415, 429-433.
- Faragher, R. (2010). Understanding the basis of the Kalman filter via a simple and intuitive derivation [lecture notes]. *IEEE Signal Processing Magazine*, 29 (5), 128-132.
- Fuentes-Pacheco, J., Ruiz-Ascencio, J. and Rendón-Mancha, J. M. (2015). Visual simultaneous localization and mapping: A survey. *Artificial Intelligence Reviews*, 43, 55-81.
- Gibson, J. J. (1950). *The perception of the visual world*. Greenwood Press.

- Gu., Y., Watkins, P. V., Angelaki, D. E. and DeAngelis, G. C. (2006). Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *The Journal of Neuroscience*, 26 (1), 73-85.
- Gu., Y., Fetsch, C. R., Adeyemo, B., DeAngelis, G. C. and Angelaki, D. E. (2010). Decoding of MSTd population activity accounts for variations in the precision of heading perception. *Neuron*, 66 (4), 596-609.
- Hausamann, P., Daumer, M., MacNeilage, Paul R. and Glasauer, S. (2019). Ecological momentary assessment of head motion: Toward normative data of head stabilization. *Frontiers of Human Neuroscience*, 13 (179), 1-13.
- Hausamann, P., Sinnott, C. B., Daumer, M. and MacNeilage, P. R. (2021). Evaluation of the Intel RealSense T265 for tracking natural human head motion. *Scientific Reports*, 11 (12496), 1-12.
- Howard, I. P. and Hu, G. (2001). Visually induced reorientation illusions. *Perception*, 30, 583-600.
- Karmali, F. and Merfeld, D. M. (2012). A distributed, dynamic, parallel computational model: The role of noise in velocity storage. *Journal of Neurophysiology*, 108 (2), 390-405.
- Klam, F. and Graf, W. (2006). Discrimination between active and passive head movements by macaque ventral and medial intraparietal cortex neurons. *Journal of Physiology*, 574 (2), 367-386.
- Kumar, A., Pundlik, S., Peli, E. and Luo, G. (2022). Comparison of visual SLAM and IMU in tracking head movement outdoors. *Behavior Research Methods*, 1-13.
- Lappe, M., Bremmer, F. and van den Berg, A. V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Science*, 3 (9), 329-336.
- LaViola, J. J. (2003). A comparison of unscented and extended Kalman filtering for estimating quaternion motion. *Proceedings of the 2003 American Control Conference*, 3, 2435-2440.
- Laurens, J. and Droulez, J. (2007). Bayesian processing of vestibular information. *Biological Cybernetics*, 96 (4), 389-404.
- Laurens, J., Meng, H. and Angelaki, D. E. (2013). Neural representation of orientation relative to gravity in the macaque cerebellum. *Neuron*, 80, 1508-1518.
- Longuet-Higgins, H. C. and Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 208 (1173), 385-397.
- MacDougall, H. G. and Moore, S. T. (2005). Marching to the beat of the same drummer: The spontaneous tempo of human locomotion. *Journal of Applied Physiology*, 99, 1164-1173.
- MacNeilage, P. R., Banks, M. S., Berger, D. R. and Bühlhoff, H. H. (2007). A Bayesian model of the disambiguation of gravito-inertial force by visual cues. *Experimental Brain Research*, 179, 263-290.
- MacNeilage, P. R., Ganesan, N. and Angelaki, D. E. (2008). Computational approaches to spatial orientation: From transfer functions to dynamic Bayesian inference. *Journal of Neurophysiology*, 100, 2981-2996.
- MacNeilage, P. R. (2020). Characterization of natural head movements in animals and humans. In Fritzsche, B. and Straka, H. (Eds.), *The Senses: A Comprehensive Reference* (pp. 69-87). Elsevier.
- Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W. and Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21, 1281-1289.

- Matthis, J. S., Muller, K. S., Bonnen, K. L. and Hayhoe, M. M. (2022). Retinal optic flow during natural locomotion. *PLoS Computational Biology*, 1-37.
- McInturff, S., Burns, J. C. and Kelley, M. W. (2018). Characterization of spatial and temporal development of Type I and Type II hair cells in the mouse utricle using new cell-type-specific markers. *Biology Open*, 7, 1-13.
- Merchant, H., Battaglia-Mayer, A. and Georgopoulos, A. P. (2001). Effects of optic flow in motor cortex and area 7a. *Journal of Neurophysiology*, 86, 1937-1954
- Merfeld, D. M., Young, L. R., Oman, C. M. and Shelhamer, M. J. (1993) A multidimensional model of the effect of gravity on the spatial orientation of the monkey. *Journal of Vestibular Research*, 3 (2), 141-161.
- Merfeld, D. M. (1995). Modeling the vestibulo-ocular reflex of the squirrel monkey during eccentric rotation and roll tilt. *Experimental Brain Research*, 106, 123-134.
- Merfeld, D. M., Zupan, L. and Peterka, R. J. (1999). Humans use internal models to estimate gravity and linear acceleration. *Nature*, 398, 615-618.
- Mittelstaedt, H. (1983). A new solution to the problem of the subjective vertical. *Naturwissenschaften*, 70, 272-281.
- Mittelstaedt, H. (1986). The subjective vertical as a function of visual and extraretinal cues. *Acta Psychologica*, 63 (1-3), 63-85.
- Mittelstaedt, H. (1989). The role of the pitched-up orientation of the otoliths in two recent models of the subjective vertical. *Biological Cybernetics*, 61 (6), 405-416.
- Mittelstaedt, M. and Mittelstaedt, H. (2001). Idiothetic navigation in humans: Estimation of path length. *Experimental Brain Research*, 139, 318-332.
- Müller, G. E. (1916). Über das Aubertsche Phänomen. *Zeitschrift für Sinnes – Physiologie*, 49, 109-246.
- Ni, L. and Stocker, A. A. (2023). Efficient sensory encoding predicts robust averaging. *Cognition*, 232,
- Pozzo, T., Berthoz, A. and Lefort, L. (1989). Head kinematic during various motor tasks in humans. *Progress in Brain Research*, 80, 377-383.
- Precht, W., Simpson, J. I. and Llinás, R. (1976). Responses of Purkinje cells in rabbit nodulus and uvula to natural vestibular and visual stimuli. *Pflügers Archiv*, 367, 1-6
- Schwabe, L. and Blanke, O. (2008). The vestibular component in out-of-body experiences: A computational approach. *Frontiers in Human Neuroscience*, 2 (17), 1-10.
- Sinnott, C., Dang, T., Papachristos, C., Alexis, K. and MacNeilage, P. (2018). Statistical characterization of heading stimuli in natural environments using SLAM. *Journal of Vision*, 18 (10), 41.
- Sun, H., Campos, J. L., Young, M. and Chan, G. S. W. (2004). The contributions of static visual cues, nonvisual cues, and optic flow in distance estimation. *Perception*, 33, 49-65.
- Tin, C. and Poon, C. (2005). Internal models in sensorimotor integration: Perspectives from adaptive control theory. *Journal of Neural Engineering*, 2, S147-S163.
- Vingerhoets, R. A. A., Medendorp, W. P. and van Gisbergen, J. A. M. (2008). Body-tilt and visual verticality perception during multiple cycles of roll rotation. *Journal of Neurophysiology*, 99 (5), 2264-2280.
- von Holst, E. and Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften*, 37, 464-476.

- Warren, W. H. and Kurtz, K. J. (1992). The role of central and peripheral vision in perceiving the direction of self-motion. *Perception and Psychophysics*, 51 (5), 443-454.
- Waters, R. L., Morris, J., and Perry, J. (1973). Translational motion of the head and trunk during normal walking. *Journal of Biomechanics*, 6 (2), 167-170
- Wei, X. X. and Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain ‘anti-Bayesian’ percepts. *Nature Neuroscience*, 18 (10), 1509-1517.
- Wei, X. X. and Stocker, A. A. (2017). Lawful relation between perceptual bias and discriminability. *Proceedings of the National Academy of Science*, 114 (38), 10244-10249.
- Willemsen, S. C. M. J., Wijdenes, L. O., van Beers, R. J., Koppen, M., and Medendorp, W. P. (2022). Natural statistics of head roll: Implications for Bayesian inference in spatial orientation. *Journal of Neurophysiology*, 128, 1409-1420.
- Yakusheva, T. A., Shaikh, A. G., Green, A. M., Blazquez, P. M., Dickman, J. D. and Angelaki, D. E. (2007). Purkinje cells in posterior cerebellar vermis encode motion in an inertial reference frame. *Neuron*, 54, 973-985.
- Zhang, T. and Britten, K. H. (2011). Parietal area VIP causally influences heading perception during pursuit eye movements. *The Journal of Neuroscience*, 31 (7), 2569-2575.
- Zupan, L. H. and Merfeld, D. M. (2003). Neural processing of gravito-inertial cues in humans. IV. Influence of visual rotational cues during roll optokinetic stimuli. *Journal of Neurophysiology*, 89 (1), 390-400.
- Zupan, L. H. and Merfeld, D. M. (2005). An internal model of head kinematics predicts the influence of head orientation on reflexive eye movements. *Journal of Neural Engineering*, 2, S180-S197.

## Chapter 2: Natural statistics of human head orientation constrain models of vestibular processing



# OPEN Natural statistics of human head orientation constrain models of vestibular processing

Christian B. Sinnott <sup>1</sup>✉, Peter A. Hausmann <sup>2</sup> & Paul R. MacNeilage <sup>1</sup>

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 gravito-inertial 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<sup>1</sup>. 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

---

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

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<sup>2,3</sup>. 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<sup>1</sup>. Instead, previous work has characterized vestibular stimulation, that is the linear acceleration and angular velocity experienced by the head during natural tasks in humans<sup>4,5</sup> as well as in non-human primates and rodents<sup>6</sup>. One very recent study does report statistics of head roll relative to gravity, but only during prescribed activities<sup>7</sup>. Similarly, one study reports statistics of head pitch during three tasks performed for 5 min each<sup>8</sup>. 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 gravito-inertial ambiguity<sup>1,9,10</sup>. 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 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<sup>10-12</sup>. 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<sup>10</sup>.

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<sup>7,13-15</sup>. In particular, we seek to explain observed, predominantly attractive biases in perception of head and body orientation<sup>14,16</sup>. 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<sup>7,14</sup>. 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<sup>17,18</sup>. Resolution of this gravito-inertial ambiguity underlies not only spatial orientation perception, but also low-level behaviors such as reflexive eye movements<sup>19,20</sup> and postural control<sup>21,22</sup>, 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<sup>23,24</sup>, 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<sup>22</sup>. 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<sup>10</sup>. 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<sup>25</sup>.

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<sup>10</sup>. 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<sup>7,13-15</sup>. 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<sup>14,16</sup>. 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  $\pm 120^\circ$  for head roll and  $\pm 90^\circ$  for head pitch. These ranges are selected because they correspond to ranges of pitch and roll perception measured in previous psychophysical research<sup>14,16</sup>, 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):

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

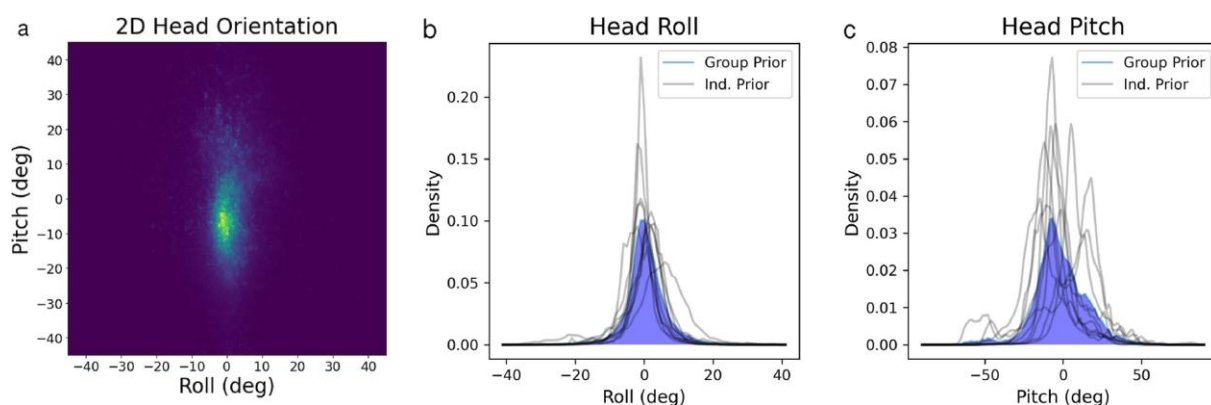
In order to gain a better understanding of how the shape of the prior impacts patterns of bias in head orientation, the modeling described above was then repeated for a variety of both empirical and parametric prior distributions. For empirical priors, progressively more smoothed versions were generated 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<sup>7,14</sup>. 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 ( $M = 0.5806^\circ$ ,  $SD = 6.2108^\circ$ ). Head pitch (Fig. 2c) was biased downward and had relatively high variability ( $M = -1.7701^\circ$ ,  $SD = 16.8167^\circ$ ). 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.

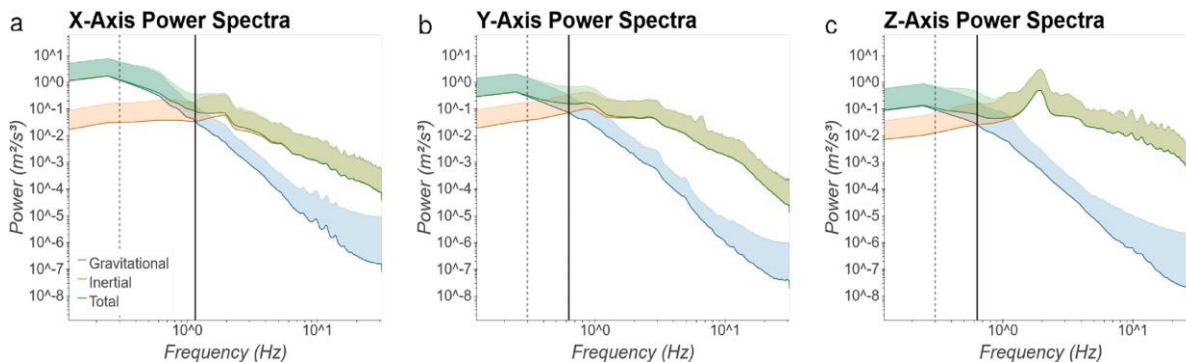


**Figure 2.** 2D head orientation measured across all participants (a). Marginal KDEs for roll (b) and pitch (c) are also plotted. The KDEs plotted in blue represent the distributions across all participants, while black traces represent KDEs for individual participants.

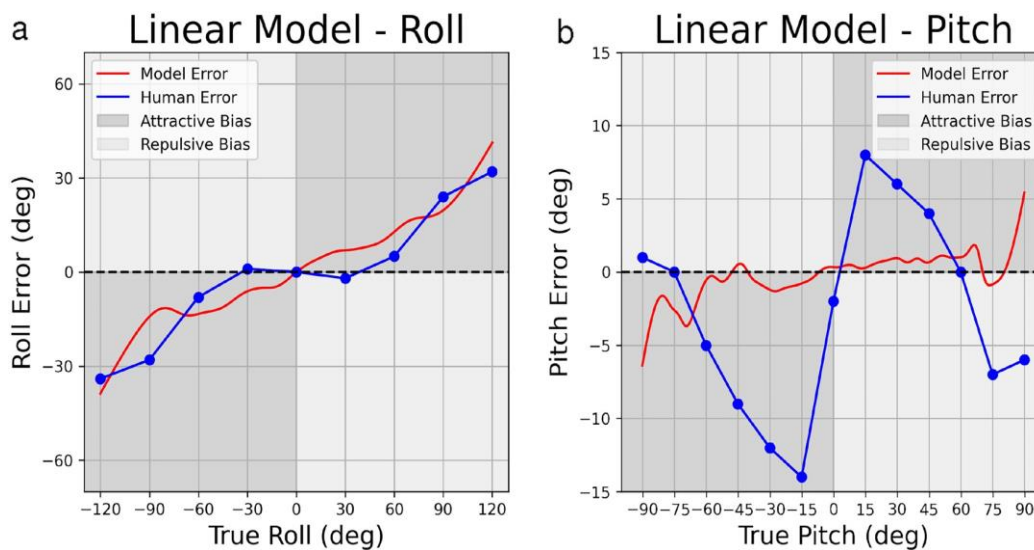
**Power spectra for tilt and translation.** Power spectra of measured linear acceleration were calculated in order to compare relative power of the gravitational (tilt) and inertial (translation) components as a function of frequency. Power was calculated using the `scipy.stats` library in Python 3, 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- (dorsalventral) (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 ( $\sigma$ ), 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

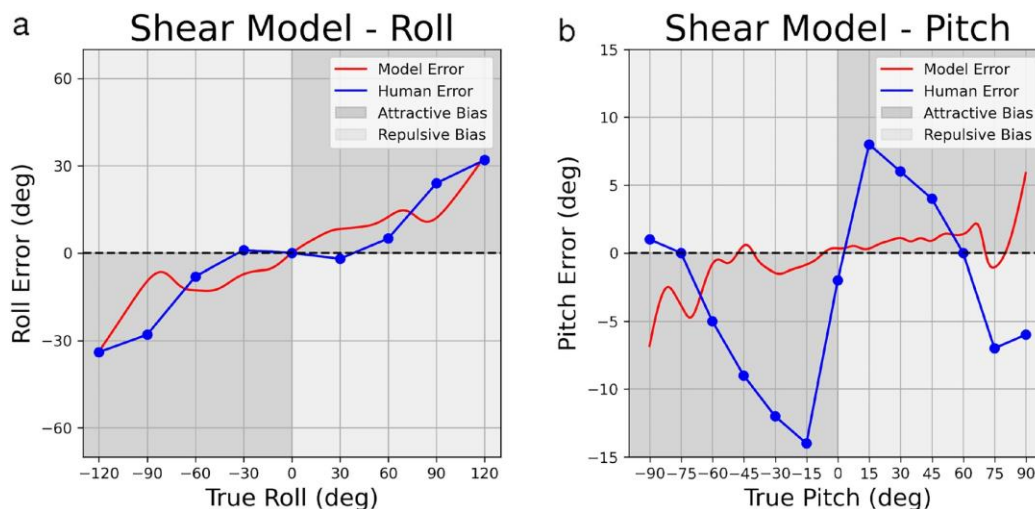
value that provided the best model fit by selecting the sigma value that minimizes distance (RSE) between the predicted and observed perceptual error.



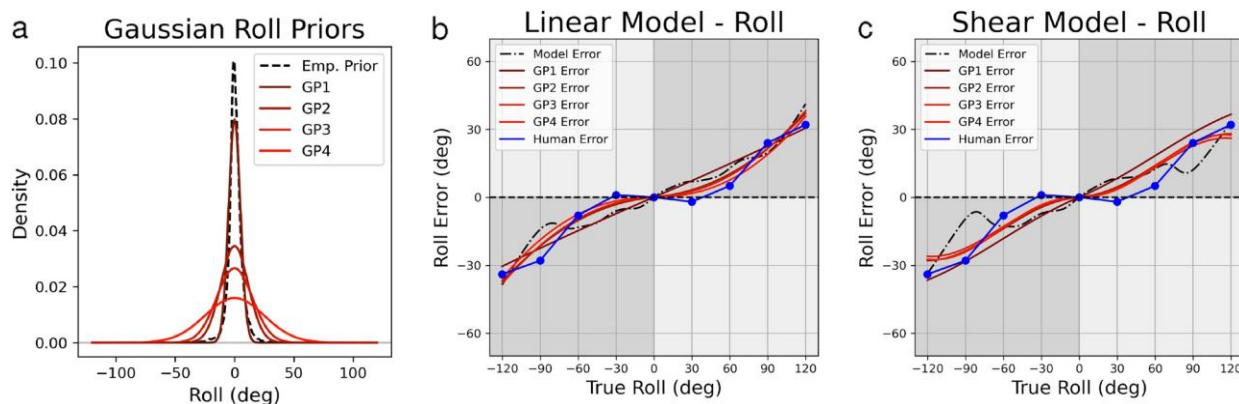
**Figure 3.** Power spectra for gravitational (blue), inertial (orange), and sum total (green) linear acceleration along X- (nasal-occipital), Y- (interaural), and Z- (dorsal-ventral) axes. Figure axes are log-scaled. Crossing points in power between gravitational and inertial acceleration are observed 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<sup>24</sup> are plotted for comparison (dashed grey).



**Figure 4.** Psychophysical data (blue) and predictions of the linear noise models (red) for roll (4a) and pitch (4b). 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<sup>14</sup>, pitch data is replotted from<sup>16</sup>. 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 (6a) 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 (5b) and the shear model (5c). 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.

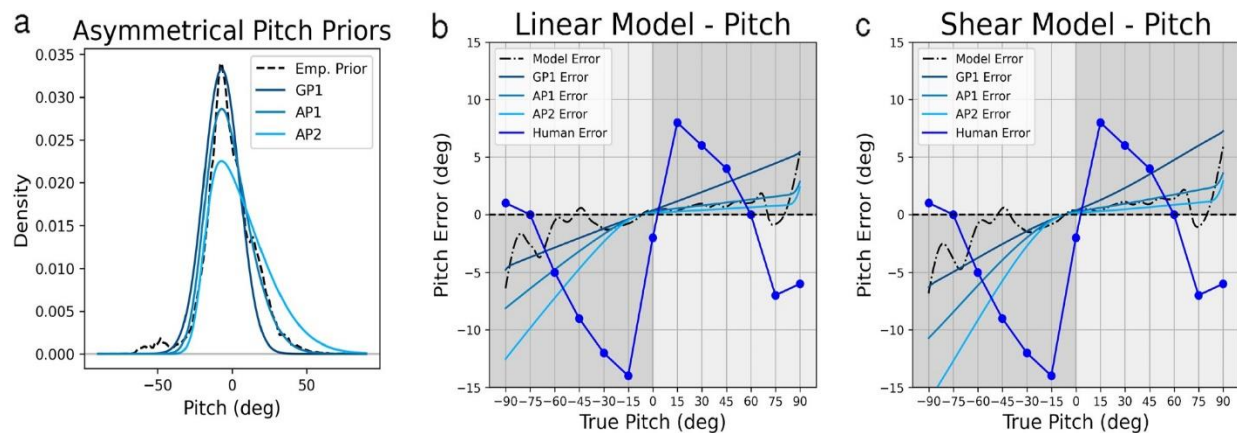
For roll perception, we compare model predictions to data from<sup>14</sup> and we achieve a fit with 6.435 RSE (with  $\sigma = 0.113$ ) for the linear noise model (Fig. 4a), and 8.346 RSE (with  $\sigma = 0.013$ ) for the shear noise model (Fig. 5a). While the fits are not perfect, the pattern of model 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<sup>16</sup> and we achieve a fit with 6.779 RSE (with  $\sigma = 0.001$ ) for the linear noise model (Fig. 4b), and 6.795 RSE (with  $\sigma = 0.01$ ) 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 (Fig. 6a, Supplementary Table S3).



**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.

---

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.<sup>7</sup>. 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<sup>26-28</sup>. 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<sup>8</sup>. 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<sup>7,13-15,29</sup>. 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.<sup>30</sup>); 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<sup>31,32</sup>. 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<sup>14</sup>. 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<sup>33</sup>. 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.<sup>7,15</sup>). 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<sup>7</sup>. 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 follow-up 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<sup>7</sup>. 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  $\pm 15^\circ$ , then decrease again, becoming repulsive (i.e. overestimation) at angles greater than  $60^\circ$  (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<sup>2,3</sup>, 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<sup>16</sup> which measured bias by asking participants to adjust their own body pitch to given target angles at  $15^\circ$  increments between  $0^\circ$  and  $90^\circ$ . Different patterns of bias may be observed using different methods, e.g. when perceived pitch is indicated using a visual probe<sup>34</sup>. 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<sup>35</sup>.

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<sup>36</sup>. This simple heuristic approach has been successful in modeling otolith-ocular responses in both squirrel monkeys<sup>23</sup> and humans<sup>24</sup>, 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<sup>19,20</sup>.

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 multimodal (i.e. canal-otolith) integration<sup>17</sup>. 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<sup>19,20</sup>. 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<sup>37</sup>. 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<sup>38</sup>, 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<sup>22</sup>. Neural populations that respond selectively to tilt and translation<sup>34,39</sup> 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<sup>40</sup>. 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<sup>10</sup>. This method overcomes gravito-inertial ambiguity inherent in data reported by several previous studies<sup>4,5</sup>. The ambiguity inherent in IMU data may be overcome using an IMU with magnetometer and filtering methods<sup>7</sup>, but the accuracy of tilt and translation estimates derived from filtering should be validated in some way<sup>9</sup>.

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<sup>4,5</sup>, 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<sup>9,35</sup>. Others report head movement and/or orientation during prescribed activities such as walking, climbing/descending stairs, running, and hopping<sup>5,7</sup>. 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<sup>41</sup>. 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<sup>1</sup>. 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<sup>42</sup>. 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 analyzed during the current study is available from the corresponding author on reasonable request.

Received: 24 December 2022; Accepted: 2 April 2023

Published online: 11 April 2023

## References

- MacNeilage, P. R. Characterization of Natural Head Movements in Animals and Humans. In *The Senses: A Comprehensive Reference* (eds Fritzsche, B. & Straka, H.) 6987 (Elsevier, Academic Press, 2020).
- 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).
- Wei, X.-X. & Stocker, A. A. Lawful relation between perceptual bias and discriminability. *Proc. Natl. Acad. Sci.* **114**, 10244–10249 (2017).
- 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).
- 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).
- 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).
- 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).
- Schwabe, L. & Blanke, O. The vestibular component in out-of-body experiences: A computational approach. *Front. Hum. Neurosci.* **2**, 1–10 (2008).
- 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).
- 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).
- 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).
- 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).
- MacNeilage, P. R., Banks, M. S., Berger, D. R. & Bühlhoff, H. H. A bayesian model of the disambiguation of gravitoinertial force by visual cues. *Exp. Brain Res.* **179**, 263–290 (2006).
- de Vrijer, M., Medendorp, P. & Van Gisbergen, J. A. M. Accuracy-precision trade-off in visual orientation constancy. *J. Vis.* **9**, 1–15 (2009).
- 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).
- Cohen, M. M. & Larson, C. A. Human spatial orientation in the pitch dimension. *Percept. Psychophys.* **16**, 508 (1974).
- Glasauer, S. Interaction of semicircular canals and otoliths in the processing structure of the subjective zenith. *Ann. N Y Acad. Sci.* **656**, 847 (1992).
- Angelaki, D. E., Wei, M. & Merfeld, D. M. Vestibular discrimination of gravity and translational acceleration. *Ann. N Y Acad. Sci.* **942**, 114–127 (2001).
- 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).
- 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).
- Fitzpatrick, R. C. & Day, B. L. Probing the human vestibular system with galvanic stimulation. *J. Appl. Physiol.* **96**, 2301–2316 (2004).
- 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).
- 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).
- 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).
- 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).
- 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).
- 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, multi-sensory 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

This research was supported by NSF under Grant Number OIA-1920896 and NIH under Grant Number P20 GM103650.

## 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](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.  
© The Author(s) 2023

## Chapter 3: Natural statistics of human heading are non-Gaussian and reveal task-dependent differences

### Abstract

Heading, the direction of linear velocity in a head-based reference frame, affects visual and vestibular cues used by organisms for locomotion and navigation. Moment-to-moment linear acceleration of the head transduced by the otolith organs, as well as optic flow experienced from the combination of head, eye, and environmental movement are both used by the nervous system to estimate heading but are themselves modulated by head movement. The nervous system must encode visual, vestibular, and potentially other signals in an efficient manner, then decode them to provide a percept that the organism experiences. Steps in this encoding/decoding process are believed to rely upon assumptions made about the statistical structure of information from the stimuli, and as a result of these assumptions; human heading perception is biased, with previous work observing overestimation of heading in certain directions away from straight-forward.

Previous Bayesian modeling approaches have been used to predict these patterns of bias, with greater success using Bayesian models that incorporate efficient encoding analogous to the nervous system. All Bayesian models rely upon a representation of an organism's previous experience (a prior); as quantification of human heading during natural activity is limited models often rely upon theoretical, parametric representations of an organism's previous experience. Here we present quantification of human vestibular heading during unprescribed, natural activities for the first time. Ten participants perform unprescribed, everyday tasks during five-hour recording sessions.

Gaussian prior distributions centered at straight-forward are often used to represent previous experience. In contrast to this assumption, our empirical measures show that both heading azimuth and heading elevation are non-Gaussian. Heading azimuth and elevation both show dominant peaks near straight-forward, but both exhibit heavier tails than a Gaussian. Additionally, heading azimuth appears multimodal, with peaks at leftward and rightward heading directions. These data constitute an empirically-derived representation of previous heading experience, i.e. a Bayesian prior for heading perception. These distributions are used in a classic Bayesian model as well as a contemporary efficient Bayesian model to predict previously observed patterns of human heading bias.

## Introduction

A major cue for spatial orientation perception is heading: the direction of linear velocity in a head-based reference frame. Accurate heading estimation is key to spatial orientation perception broadly, and is critical for several important behaviors in humans, non-human primates, and other organisms. These behaviors include volitional, high-level behaviors like locomotion and navigation (Cuturi and MacNeilage, 2013) and low-level reflexive behaviors like postural control. Perception of one's heading in combination with estimating one or more other organisms' heading directions also effects behaviors such as collision avoidance (Sweeny, Haroz, and Whitney, 2012)

Sensory information for heading is multimodal, with the most reliable heading information given through vestibular and visual systems. Heading systematically modulates the moment-to-moment sensation of linear acceleration transduced by the otolith organs via changing the direction of the linear acceleration vector relative to the head. Areas sensitive to vestibular heading, or implicated in vestibular heading processing, can be found in multiple cortical areas in mammals: parieto-insular ventral cortex (PIVC, Chen et al., 2016; Chen et al., 2021), ventral intraparietal area (VIP, Britten, 2008; Chen, DeAngelis and Angelaki, 2011; Chen, DeAngelis and Angelaki, 2013), dorsal area of the medial superior temporal region (MSTd; Britten, 2008; Gu et al., 2010), and retrosplenial cortex (RSC; Cho and Sharp, 2001).

Alongside the vestibular system, vision is critical to heading perception. Optic flow, the systematic visual displacement of world-fixed environmental features induced by motion, is an important sensory cue used to estimate heading direction (Lappe,

Bremmer, and van den Berg, 1999). Optic flow is often elicited through self-motion of the body, head, and eyes, but also affected by environmental depth structure and object movement (Longuet-Higgins and Pradzny, 1980). Simple linear or angular head movement, as well as complex head movement combining linear and angular components result in characteristic patterns of optic flow that can be used to estimate one's heading as they move through an environment.

Visual heading responses are often found in the same areas that exhibit vestibular heading responses. MSTd, VIP, and PIVC all show activity elicited by visual stimuli resembling optic flow fields (Chen et al., 2011; Zhang and Britten, 2011; Chen et al., 2021). Primary motor cortex, another area where optic flow elicits neuronal activity, is seemingly not excited by vestibular heading stimuli (Merchant, Battaglia-Mayer, and Georgopoulos, 2001). Other sensory modalities which contribute to heading perception specifically have not yet been definitively identified. However, other work has shown the importance of somatosensory and proprioceptive systems when estimating path length (Mittelstaedt and Mittelstaedt, 2001; Sun, Campos, and Chan, 2004) which may be related to heading.

A benefit of estimating heading using multiple sources of information across senses is a reduction of uncertainty, as a single source of information may be ambiguous on its own. The otolith organs transduce the sum gravito-inertial acceleration rather than the inertial component corresponding with heading, and object motion in the environment can result in less coherent patterns of optic flow: reducing the reliability of a heading estimate given through vision alone. The primary consequence of this reduction in

uncertainty is increased perceptual sensitivity, which increases heading estimation accuracy and precision. Previous research has identified MSTd as a region where visual-vestibular sensory integration occurs in a statistically optimal manner for heading perception (Angelaki, Gu and DeAngelis, 2011; Fetsch et al., 2011; Gu, Angelaki, and DeAngelis, 2009).

Likely, this observed increase in performance due to sensory integration is driven in part by disambiguating sensory stimulation driven by self-motion from stimulation caused by environmental or object motion. Information from multiple sensory modalities may be used by the nervous system to parse optic flow into components driven by self-motion and object motion (Peltier, Angelaki and DeAngelis, 2020). Information from vision and the semicircular canals also aids the otoliths in disambiguating linear acceleration driven by tilt relative to gravity and acceleration from translation (Yakusheva, Blazquez and Angelaki, 2008). Recent work posits that disambiguation of sensory information occurs concurrently with multisensory integration in MSTd and VIP (Zhang et al., 2019). Evidence suggests that while one neuronal subpopulation (congruent neurons) in these areas performs integration, another (opposite neurons) segregates multisensory information (Zhang et al., 2019). While direct evidence of this function in these areas have not been observed, recent work investigating middle temporal (MT) area with similar neuronal subpopulations finds strong evidence that congruent and opposite neurons function in this way when parsing self-motion-driven optic flow and object-motion optic flow (Kim, Angelaki and DeAngelis, 2022).

Despite the availability of information provided by multiple sensory systems, multisensory integration reducing uncertainty in that information, and the ability of the nervous system to parse that information back into separate components, previous research demonstrates heading to be systematically biased across sensory modalities (Cuturi and MacNeilage, 2013; Crane, 2014). Work investigating visual and vestibular heading azimuth and elevation perception has shown patterns of bias that alternate between attraction towards the assumed most-frequently encountered heading value of straight-forward (functional underestimation of the true stimulus value), and repulsion from straight-ahead (functional overestimation). This pattern persists across sensory modalities, with vestibular and visual bias exhibiting the same alternation. Visual bias relative to vestibular bias appears to exhibit greater magnitude with greater underestimation and overestimation present in periods where attractive and repulsive bias are observed (Cuturi and MacNeilage, 2013; Crane, 2014).

To understand how biases in heading perception emerge, many researchers utilize Bayesian methods to estimate or predict performance of a human observer experiencing motion in a given heading direction (for review, see Fetsch et al. 2013), as the nervous system seems to represent information probabilistically (Knill and Pouget, 2004). These methods range from static Bayesian models that rely directly on Bayes' law to estimate some posterior distribution given a likelihood and prior distribution (MacNeilage et al., 2007), to recursive Bayesian estimators like Kalman filters (MacNeilage, 2008; Clark, 2019) and particle filters (Laurens and Droulez, 2007). While these methods can perform well when modeling human behavior, they rely on multiple assumptions about the statistical structure of heading stimuli experienced by the organism over the lifespan.

Namely, prior distributions of heading are assumed to be unbiased (i.e. peaked at straight-forward) and Gaussian. While this is a reasonable assumption, it is untested as little previous work has estimated vestibular heading during selected natural activities which would inform such a prior (Sinnott et al., 2018). A major criticism of these modeling approaches is that they rely on several free parameters that are manipulated by the modeler to achieve the best fit. In other words, given enough free parameters, any pattern of bias can be fit. To address some of these criticisms, recent work has described an “efficient” Bayesian model which couples an efficient encoding stage with a Bayesian decoding stage (Wei and Stocker, 2015). Later work evaluating this model across multiple sensory phenomena finds that it can predict observed repulsive patterns of bias in heading perception using a parametric Von Mises distribution as a prior (Wei and Stocker, 2017).

In the current study, we set out to characterize empirically observed distributions of heading azimuth and elevation that could be used to constrain generation of prior distributions used in any of these modeling approaches. In contrast with other work measuring natural statistics of vestibular stimuli (e.g. Carriot et al., 2007), rather than record relatively short (e.g. 5-10 minute) bursts of specific activities we record heading during unconstrained, unprescribed, natural activities over long (5-hour) recording sessions. Then, we use distributions of heading azimuth and elevation to generate prior probability distributions that are used in two modeling approaches to predict human heading perception bias: a standard Bayesian model used to predict bias in other sensory cues related to spatial orientation perception (de Vrijer et al., 2009; Willemsen et al.,

2022), and an efficient Bayesian model which has been used to predict perceptual bias across multiple sensory modalities (Wei and Stocker, 2015; Wei and Stocker, 2017).

## Methods

We recruited ten participants (6 male, 3 female, 1 nonbinary identifying) to participate in the study. We asked participants to wear the Intel RealSense T265 (T265) on their heads over a continuous recording period lasting five hours, and complete natural behaviors that they were comfortable with being recorded. We instructed participants to avoid activities that would compromise the equipment or be impeded by the recording equipment itself. Participants were also instructed to re-calibrate the T265 every thirty minutes during the recording session. All participants provided informed consent in-line with procedures approved by the Institutional Review Board at the University of Nevada, Reno.

Once calibration was complete, we instructed participants to wear the T265 while completing typical, everyday tasks. Participants were instructed to avoid activities that would compromise equipment integrity (e.g. swimming) or privacy concerns. If participants needed to perform activities that might compromise privacy, they were able to use a privacy shutter that functionally disabled video recording. We also instructed participants to repeat part of the calibration procedure every 30 minutes during the recording. After five hours, participants returned to the lab, repeated the full calibration procedure, and removed the equipment.

## Hardware

We tracked heading using the Intel RealSense T265 tracking camera. The T265 is an off-the-shelf sensor which performs positional tracking via visual-inertial simultaneous localization and mapping (VI-SLAM). To facilitate this, the T265 is equipped with multiple sensors: two stereoscopic, global shutter, fisheye cameras (848x800 pixel resolution sampled at 30 Hz, 173° field of view), a gyroscope (200 Hz), and an accelerometer (62.5 Hz). All data from these sensors are fused and processed internally on the device using a proprietary algorithm, producing a six degree-of-freedom (DOF) positional estimate at 200 Hz. We down-sample this estimate to 62.5 Hz, matching the accelerometer for analyses presented in the previous chapter. Previous research has validated the T265 for tracking human head movement (Hausamann et al., 2021). Alongside these data, video is recorded from the T265's stereoscopic cameras at 1/30 Hz (1 frame per minute) to augment activity classification for a separate experiment.

The T265 is connected via USB to a laptop (Dell Latitude 5300). The laptop has 16GB of DDR4 RAM and an 8th-generation Intel i7 CPU. This laptop was carried by participants using a backpack. Participants wore the T265 using an elastic strap harness and custom, 3D-printed mounting bracket. Available on the mounting bracket is a privacy shutter which allows participants to perform behaviors they may not be comfortable completing while video-recorded (e.g., viewing sensitive information).

### Calibration

In the current study we use a kinematic transformation procedure, detailed in the previous chapter, to transform data from a camera-based reference frame into a head-based one. First, we record a 15-second segment where participants hold their head still

and in a position where Reid's baseline was perpendicular to gravity. Next, participants are instructed to slowly pitch their head upwards and downwards five times (e.g., nodding their head "yes") followed by yawing their head leftwards and rightwards five times (e.g., "shaking" their head "no"). Participants repeated the yaw and pitch movements approximately every thirty minutes during the recording, allowing for correction of slippage the T265 may experience during the recording.

### Pre-processing

In order to remove artifacts from our data we complete a series of automated and manual exclusion steps. First, we exclude data where tracking confidence is low (and presumably, data is noisy and/or erroneous). The T265 provides a proprietary rating of instantaneous tracking confidence, presumably based on factors like velocity magnitude given from the T265's inertial measurement unit (IMU), stability of visual landmarks used by the (VI-SLAM) algorithm, or other factors. The exact details how this confidence estimate are generated are unknown; regardless, the T265's confidence estimate ranges from 0 (poor) to 3 (good). All data with a confidence rating below 3 are excluded from further analysis.

Next, we exclude data with unreasonably high linear velocity magnitudes. We noticed when visually inspecting timeseries plots of linear velocity that occasional, transient tracking failures manifested as high confidence estimates of extremely high linear velocity magnitude (e.g. 100000 meters per second). As these would not be removed through confidence-based exclusion, we choose a linear velocity threshold of 3.05 m/s based from previous work observing average jogging speed in young adults

(Barreira, Rowe, and Kang, 2010). All data below this threshold are used for further analysis, while all above it are removed.

Finally, we selectively parse data into high- and low-velocity categories using instantaneous velocity magnitude. This is done as a first, coarse means of separating vestibular heading data into activity-based categories. We use a threshold of 0.75 m/s, with any data with magnitude greater than this value categorized as high-velocity activity. Data with magnitude lesser than this value are categorized as low-velocity activity. While this is an arbitrary threshold, it was selected as an intermediate velocity between locomotion (which is generally on the order of 0.9-1.2 m/s) and stationary activities like standing (which generally are below 0.5 m/s). As a result, most data in the high-velocity category consists of walking, while low-velocity data comprises standing, sitting, lying down, or other activities where linear velocity of the head is relatively low.

We do not smooth, filter, or otherwise process data estimated by the T265. Presumably, the T265 smooths and/or filters positional data that it provides to the end user, but as the device and data estimation pipeline are proprietary, we cannot confirm this. Previous work has evaluated the T265 for estimating human head motion (Hausamann et al., 2021), finding it performed well during human locomotion when measured against an optical tracking space. Authors of that study did not smooth or filter estimates given by the T265, so we choose not to in the current study. After pre-processing, we had 41.95 hours of data remaining for further analysis.

Descriptive statistics and prior generation

Previous work often uses linear descriptive statistics such as arithmetic mean and standard deviation to quantitatively describe aspects of a given sampling distribution. More recent work has used the first four moments of a distribution, including arithmetic mean, variance, skewness, and kurtosis to quantitatively describe a sampling distribution's shape (Willemsen et al., 2022). Our data are circular, however; unlike head orientation, which does not appear to wrap and can be well described by linear methods (e.g. arithmetic mean, variance, central moments), circular data necessitate use of methods appropriate for calculating parameters of circular sampling distributions. We use the circular mean and circular standard deviation to quantitatively describe heading azimuth and elevation distributions. Circular mean can be calculated using the following formula:

$$\bar{\alpha} = \arctan2\left(\sum_{i=1}^n \sin(\alpha_i), \sum_{i=1}^n \cos(\alpha_i)\right)$$

In this case, the mean azimuth or elevation angle ( $\bar{\alpha}$ ) is the two-argument arctangent of the sum of sines of all azimuth or elevation angles and the sum of cosines of all azimuth or elevation angles. Likewise, the circular standard deviation of a distribution can be calculated in the following way:

$$\bar{S} = \sqrt{-2\ln(\bar{R})}$$

$$\bar{R} = \sqrt{\left(\frac{1}{n}\sum_{i=1}^n \sin(\alpha_i)\right)^2 + \left(\frac{1}{n}\sum_{i=1}^n \cos(\alpha_i)\right)^2}$$

Circular standard deviation is the square root of -2 times the natural log of the resultant ( $\bar{R}$ ). The resultant is the square root of the sum of squares of the sine of all heading angles, plus the sum of squares of the cosine of all heading angles. Use of these methods is necessary as heading azimuth appears to wrap around all 360 degrees of the unit circle.

To use our empirically observed heading distributions, another processing step is necessary to generate a continuous probability function that can be used as a prior in a Bayesian model. We use kernel density estimation to smooth heading sampling distributions using a Gaussian kernel with a bandwidth ( $h$ ) = 0.04. This value is determined by Scott's rule (Scott, 1992), a heuristic where the data's length ( $n$ ) and dimensionality ( $d$ ) is used to calculate an appropriate kernel bandwidth:

$$h = n^{-\frac{1}{d+4}}$$

### Modeling heading perception

We use two approaches that utilize empirical observations of heading azimuth and elevation to model heading perception bias observed in previous research (Cuturi and MacNeilage, 2013; Crane, 2014). First, we use a standard Bayesian modeling framework which has been used in successfully model bias in head orientation perception in both the previous chapter as well as other, published work (Willemsen et al., 2022). This model predicts patterns of bias that are predominantly "attractive" towards a prior: this is functionally equivalent to underestimation of a true stimulus value.

Given some prior at  $0^\circ$ , posterior distributions generated as a product of the prior and a symmetric likelihood will have a mean that is closer to  $0^\circ$  than the mode of the likelihood distribution. We apply sensory noise in the model by increasing variability of the likelihood with heading angle. As absolute heading angle moves away from straight-forward, likelihood variability increases linearly (linear model) or non-linearly (utricle shear model) with the sine of the heading angle. This can result in a pattern of attractive bias that increases in magnitude with heading angle. A full description of this model, with equations, can be found in Appendix A.

However, previously observed heading bias is not entirely attractive. Perception of heading azimuth and elevation exhibit "repulsive" biases for certain heading directions (Cuturi and MacNeilage, 2013; Crane, 2014). Rather than underestimating the true stimulus value, participants often overestimate the true stimulus during psychophysical heading tasks. These are often termed "anti-Bayesian" biases or behaviors as they largely conflict with the expected set of predictions made by a standard Bayesian model. Due to the pattern of attractive and repulsive biases, we anticipate that a standard Bayesian model would perform poorly given an unbiased, symmetric prior and a symmetric likelihood. As a result, we also use a more recent modeling approach which combines Bayesian decoding (as seen in our standard Bayesian model) with efficient encoding (Wei and Stocker, 2015).

Wei and Stocker argue that while standard Bayesian models predict perceptual bias well, they are ultimately incomplete as they do not account for how sensory information is represented in the nervous system (2015). As such, explanatory power of

this modeling approach is reduced as model parameter selection or other model specifications are somewhat arbitrary, a criticism levied at Bayesian modeling approaches previously (Bowers and Davis, 2012). Wei and Stocker address this criticism by incorporating an efficient encoding step which generates a representation of sensory information, that is then interpreted with a Bayesian decoder.

While this model uses a Bayesian decoding stage analogous to our standard Bayesian model, parameters of the likelihood distribution used in this stage are determined through two constraints that function as an efficient encoding stage analogous to efficient encoding present in neuronal populations. A more thorough description of the model can be found in the general introduction of the dissertation. The efficient encoding stage assumes that any sensory representation of a stimulus and the stimulus value itself share a maximal amount of mutual information. Mutual information is represented as Fisher information, which can be more generally defined as the amount of information derived about a parameter using a given random variable. Authors of this model derive that the square root of the Fisher information for a stimulus value is proportional to the prior of stimulus values:

$$P(\theta) \propto \sqrt{J(\theta)}$$

This serves to quantify the encoding performance for some representation of stimulus value. As a given stimulus value has high probability within the natural environment, more resources are allocated to it which increases the accuracy of the internal representation of the stimulus value.

Further constraint is needed to fully define the uncertainty and shape present in the likelihood. This is done through a step that represents the transfer and remapping of stimulus values out of a natural stimulus space and into an internal sensory representation space using a mapping function derived from the cumulative probability distribution of the prior ( $P(\theta)$ ). This space is assumed to be uniform with respect to Fisher information across the full range of possible stimulus values, and a subsequent assumption given this homogeneous Fisher information is that noise of a likelihood function represented in this space is symmetric.

After the likelihood distribution is generated in this homogeneous sensory space, it is remapped into the stimulus space using the inverse of the previous mapping function. The result of remapping into this space is that when a likelihood is generated at a value that is infrequent in the prior distribution, noise applied to the likelihood is asymmetric with a tail facing away from the prior distribution. Later work would validate use of this model in predicting perception of multiple stimuli across multiple sensory modalities (Wei and Stocker, 2017) as well as for decision-making in the context of credit assignment (Ni and Stocker, 2023). In contrast to the standard Bayesian model, we predict that this model will result in a better qualitative and quantitative fit to existing psychophysical data measuring bias in heading perception. Greater detail on the construction and derivation of model equations can be found in (Wei and Stocker, 2015).

A major innovation we make with both the standard and efficient Bayesian models is the use of empirically derived heading azimuth and elevation distributions as priors, rather than theoretical priors which maximize model fit. While the efficient

Bayesian model is designed with use of measured prior statistics in mind, natural statistics of heading that comprise such a prior have not been well-characterized. All model variants are fit to previous psychophysical work measuring perceptual biases in vestibular heading azimuth (Cuturi and MacNeilage, 2013) and heading elevation (Crane, 2014) using a least-squares approach that minimizes distance of the models' predictions from observed human data. As a metric for distance, we use residual standard error (RSE):

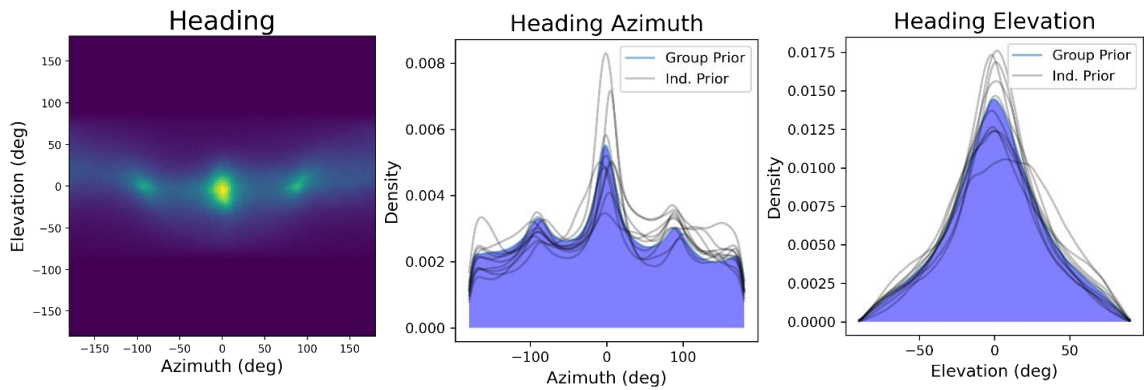
$$RSE = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (y_i - \bar{y}_i)^2}$$

For the standard Bayesian model, we search the parameter space for a single parameter value that provides a fit with minimal distance between model prediction and human data. In the case of both standard Bayesian models, this is a single free parameter representing signal dependent noise ( $\sigma_{signal}$ ). In contrast with the standard Bayesian model, we search the parameter space manually for the efficient Bayesian model for two noise parameter values that provide the minimal distance between model prediction and human data while also providing good qualitative fit. This can result in sub-optimal performance with respect to our quantitative metric RSE. These two free noise parameters represent stimulus noise ( $\sigma_{stim}$ ) and sensory noise ( $\sigma_{sens}$ ).

## Results

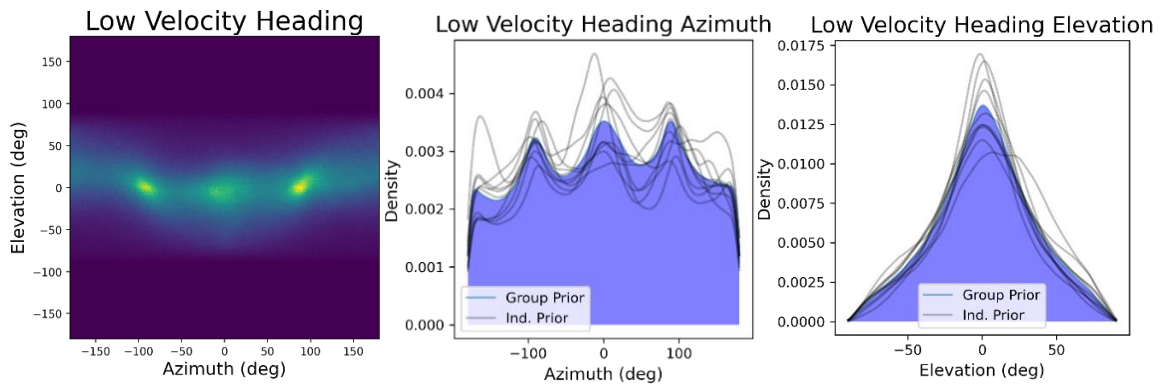
### Descriptive Statistics

Both heading azimuth and heading elevation distributions have means near  $0^\circ$  or straight-forward, though both mean azimuth ( $\bar{\alpha} = 11.75^\circ$ ) and mean elevation ( $\bar{\alpha} = 1.272^\circ$ ) appear to be slightly biased (Figure 1). Heading elevation is quite variable ( $\bar{S} = 18.68$ ), but variability for heading azimuth is substantially larger ( $\bar{S} = 229.66$ ). Qualitatively, heading azimuth is a multimodal distribution; its largest peak is near  $0^\circ$  (straight-forward), while additional peaks appear at  $\pm 90^\circ$  (leftward and rightward). Heading elevation has a more conventional, unimodal distribution with its peak near  $0^\circ$ .



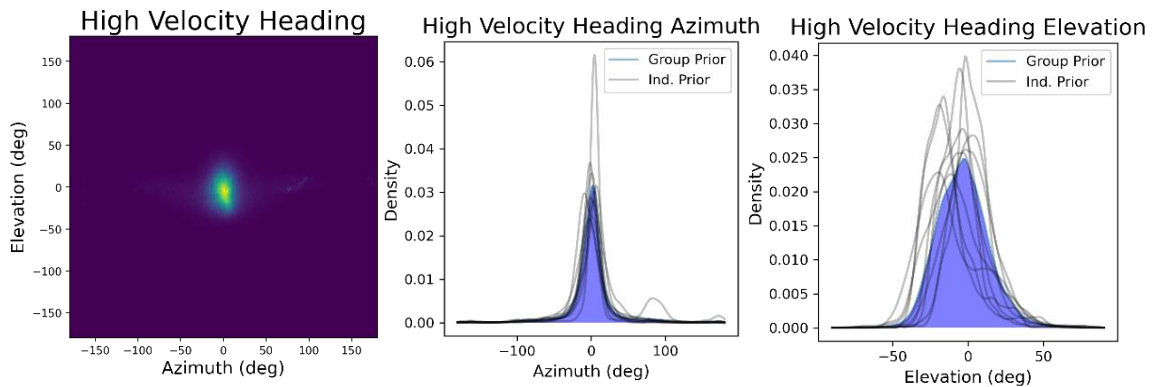
**Figure 1:** 2D heading measured across all participants and velocity ranges (left pane). Marginal kernel density estimates (KDEs) for heading azimuth (middle) and elevation (right) are also plotted. The KDEs plotted in blue represent estimates generated from distributions across all participants, while black traces represent individual KDEs.

Next, we visualize and describe circular mean and variance for heading azimuth and elevation distributions generated from low-velocity epochs (Figure 2). Circular means for these epochs were similar to those for all velocity epochs (azimuth,  $\bar{\alpha} = 22.41^\circ$ ; elevation,  $\bar{\alpha} = 1.67^\circ$ ). In contrast, circular standard deviation was increased (azimuth,  $\bar{S} = 130.42$ ; elevation,  $\bar{S} = 22.41$ ). The low velocity heading azimuth distribution's central peak at  $0^\circ$  was smaller than the azimuth distribution across all epochs. Rather than a strong central peak at  $0^\circ$  with secondary peaks at  $\pm 90^\circ$ , peaks present in low-velocity heading azimuth were fairly similar in size.



**Figure 2:** 2D heading measured across all participants during low-velocity epochs (left). Marginal KDEs during low-velocity epochs for heading azimuth (middle) and elevation (right) are also plotted. The KDEs plotted in blue represent estimates generated from distributions across all participants, while black traces represent individual KDEs.

Finally, high-velocity heading distributions generated from high-velocity epochs was different from distributions generated from all data (Figure 3). Circular means were near zero as observed for overall distributions (azimuth  $\bar{\alpha} = 0.88^\circ$ ; elevation  $\bar{\alpha} = -3.38^\circ$ ). However, standard deviation decreased sharply (azimuth  $\bar{S} = 27.53$ ; elevation  $\bar{S} = 17.31$ ). The shape of heading azimuth and elevation distributions also changed dramatically. The distribution for heading elevation became more platykurtic, with reduced weight in its tails; heading azimuth became unimodal and approached a Gaussian distribution.

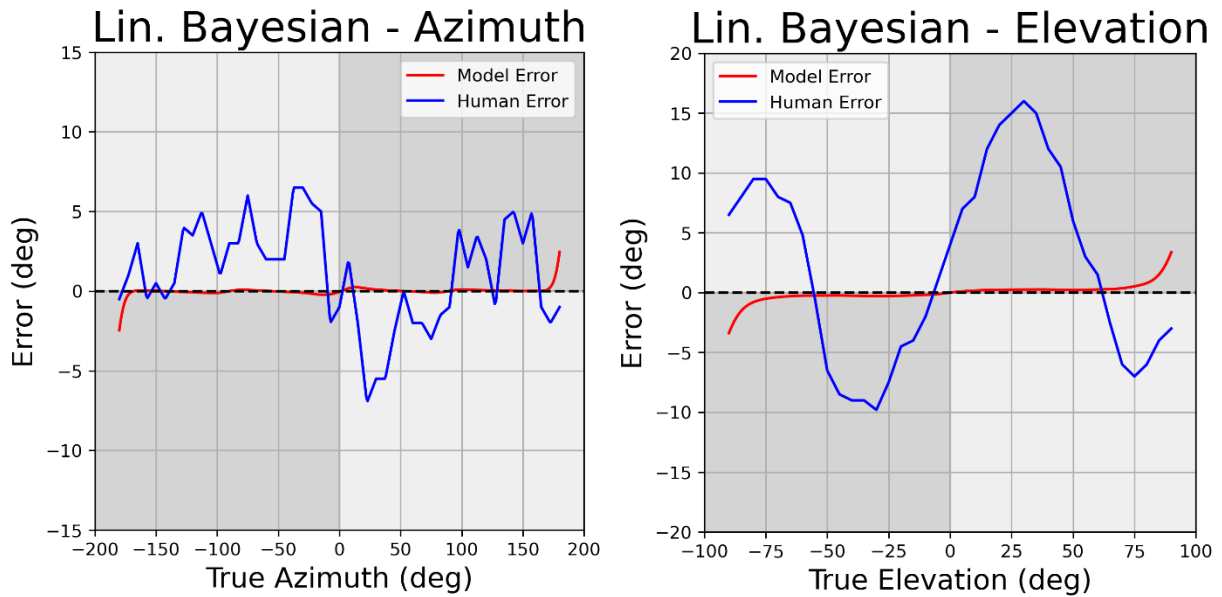


**Figure 3:** 2D heading measured across all participants during high-velocity epochs (left). Marginal KDEs during low-velocity epochs for heading azimuth (middle) and elevation (right) are also plotted. The KDEs plotted in blue represent estimates generated from distributions across all participants, while black traces represent individual KDEs.

### Modeling

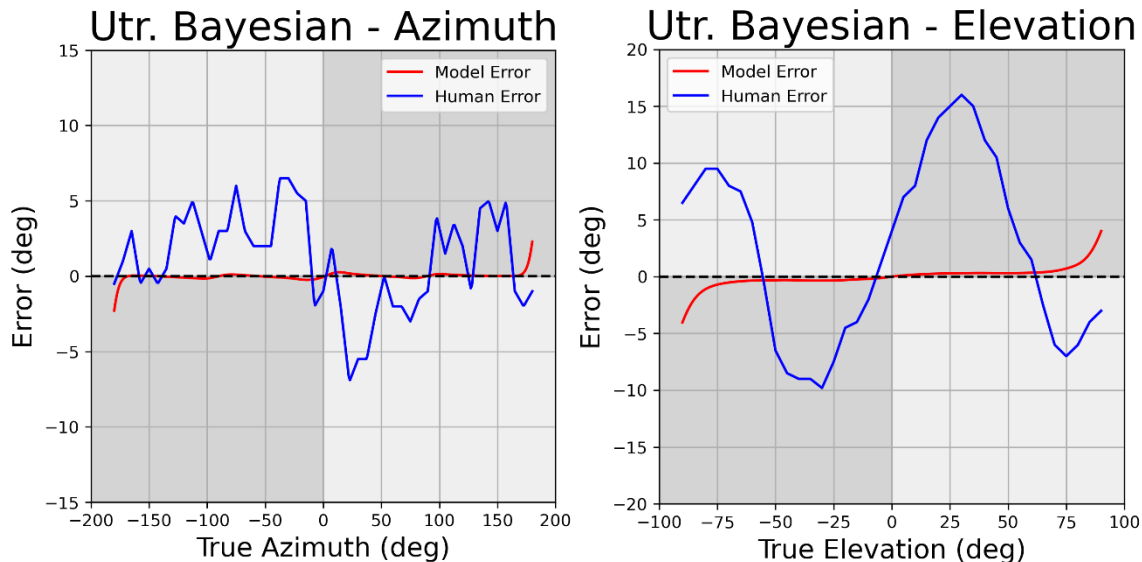
Fit of the standard Bayesian models to perceptual bias in heading relies upon a single parameter describing the level of signal-dependent noise. Optimized model fits resulted in a largely flat pattern of bias across most azimuth and elevation angles, due to this model's inability to predict repulsive bias (Figures 4 and 5). When fitting heading azimuth perception, linear and utricular noise variants yielded  $RSE = 3.288$  and  $RSE =$

3.289, respectively. The qualitative fit with both variants of the standard Bayesian model was poor, converging to minimal signal-dependent noise value.



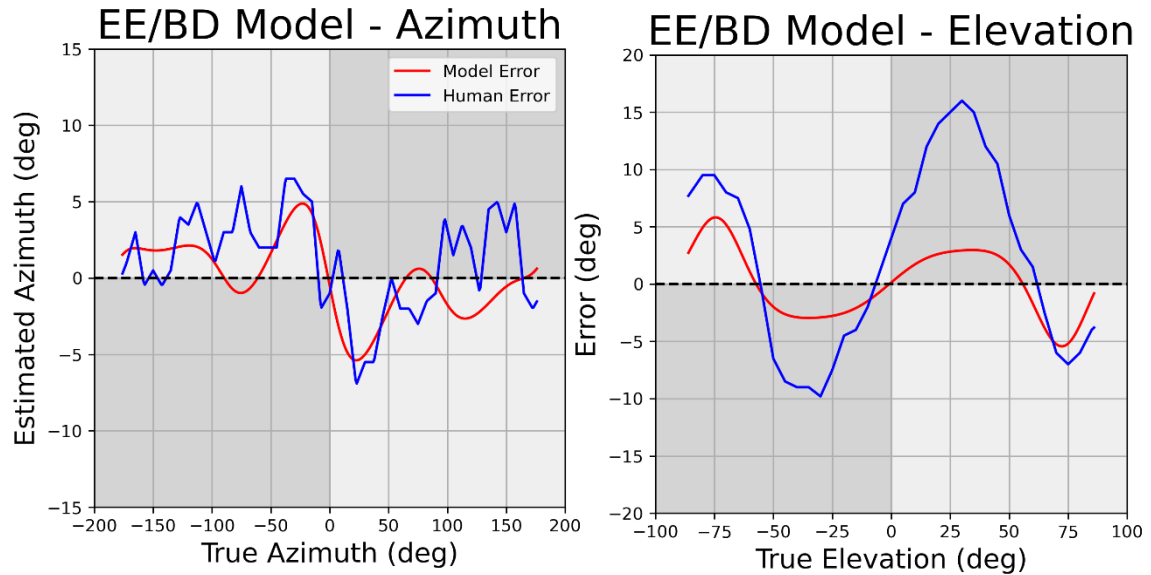
**Figure 4:** Linear noise models for heading azimuth (left) and heading elevation (right). X- and Y- axes represent true heading angle and heading angle estimation error, respectively. Dark grey regions of the plot represent repulsive bias (overestimation error), while light grey regions represent attractive bias (underestimation error).

When fitting heading elevation perception, similar patterns of minimal bias were fit for linear ( $RSE = 8.417$ ) and utricular ( $RSE = 8.476$ ) noise variants of the standard Bayesian model. RSE values were larger compared to azimuth due to relatively larger bias magnitudes.



**Figure 5:** Utricular shear noise models for heading azimuth (left pane) and heading elevation (right pane). X- and Y- axes represent true heading angle and heading angle estimation error, respectively. Dark grey regions of the plot represent repulsive bias (overestimation error), while light grey regions represent attractive bias (underestimation error). Model fit is evaluated by calculating RSE for model predictions at each stimulus value (red) and observed perceptual bias (in blue).

Use of the efficient Bayesian model requires setting values of two free parameters representing the strength of noise present at the encoding and decoding stages of the model. In contrast with either of the standard Bayesian models, the efficient Bayesian model is able to achieve a qualitatively reasonable fit for heading azimuth perception, with a quantitatively improved fit ( $RSE = 2.857$ ). A similar trend was observed for performance of the heading elevation model: a qualitative fit that began to approximate observed patterns of bias, with modest improvements in quantitative fit relative to either standard Bayesian model ( $RSE = 6.056$ ).



**Figure 6:** Efficient encoding/Bayesian decoding (EE/BD) models for heading azimuth (left pane) and heading elevation (right pane). X- and Y- axes represent true heading angle and heading angle estimation error, respectively. Dark grey regions of the plot represent repulsive bias (overestimation error), while light grey regions represent attractive bias (underestimation error). Model predictions (red) are plotted against previously observed patterns of heading perception bias (blue).

## Discussion

In this study we measured heading in ten participants completing unprescribed, naturalistic tasks during long-term recording sessions. Distributions of natural human heading azimuth and elevation have not been well-characterized in the literature, and a primary aim of this study was measuring these natural distributions. A secondary aim was to use these distributions in two Bayesian modeling frameworks to relate the characteristics of natural heading distributions to previously observed patterns of perceptual bias. We conclude with consideration of the current study's limitations as well as future directions of research.

### Characteristics of Heading Distributions

Heading azimuth and heading elevation deviate from parametric unimodal (e.g. Gaussian) distributions in manners that are consistent across all subjects. First, heading azimuth and elevation both exhibit high levels of variability, with azimuth and elevation values present across the full range of possible values. Second, heading azimuth exhibits multiple modes, with peaks in straight-forward, purely leftward, and purely rightward directions. Previous work has assumed priors for heading to have a dominant mode at straight-forward, but subsequent peaks in other cardinal directions were unexpected. Likely, this is driven by the inclusion of multiple different activities with both higher velocities consistent with conventional heading (e.g. during volitional self-motion) as well as lower velocities (e.g. during stationary periods where the head moves slightly).

An interesting trend that emerges from heading azimuth is that there appears to be an overall bias towards rightward trajectories. Furthermore, there appears to be greater inter-subject variability for rightward trajectories than leftward. We do not have a clear explanation for this trend but provide a few possible explanations. First, it is possible that this is driven by calibration error in some way but given the alignment of the head-based reference frame to gravity via Reid's plane, any sort of error or misalignment of the reference frame to gravity would likely manifest as misalignment in heading elevation rather than azimuth. As we do not see a pattern of bias driven through misalignment in heading elevation, we do not believe calibration error is driving any potential rightward bias in heading azimuth.

An alternative explanation is that bias may be driven by activities people are comfortable performing while recording. There are several plausible natural activities such as skateboarding, sitting in an office chair and moving laterally, or walking forward with head rotation (e.g. to speak to someone walking with you) that can result in rightward heading direction. Visual inspection of sparsely-sampled (1/60 Hz) video from the T265 revealed that participants did complete activities such as these, but no single, obvious task emerged that could explain this bias. Further work is necessary to determine the extent to which these types of activities are present in the data we collected.

A more compelling trend emerges when parsing heading azimuth data by velocity magnitude. Observed heading azimuth distributions derived from low-velocity epochs show reduced power in each mode relative to the rest of the data; in other words, they more closely approximate a uniform distribution. The heading azimuth distribution derived from only high-velocity epochs changes dramatically, becoming unimodal and showing substantially reduced variability. Likely, modes in rightward and leftward directions are being driven by stereotyped "stationary" activity such as postural sway, while the dominant peak towards zero is driven by activity like locomotion which is predominantly forward moving.

Ostensibly, use of alternative low- and high-velocity heading distributions as empirical priors would result in different patterns of predicted bias using the efficient Bayesian model. In the case of the low-velocity azimuth distribution, there are multiple peaks but also support throughout the entire range of azimuth values. In the context of the efficient Bayesian model, this should result in relatively low-noise likelihood

distributions with little, if any asymmetry away from straight-forward. This does beg the question: how can repulsion be present if there is no singular mode to be repelled away from? How does repulsion or attraction work in the context of Bayesian models with a multimodal prior? Use of the high-velocity azimuth distribution as a prior, in contrast, should result in strongly repulsive biases as likelihoods at directions away from straight-ahead are generated.

### Modeling Human Heading Perception

We first fit a standard Bayesian model using kernel density estimates generated from our empirically observed heading azimuth and heading elevation data as priors. As we anticipated, this model was unable to predict previously observed patterns of perceptual bias for heading azimuth and elevation. This is due to this model's inability to generate repulsive (overestimation) patterns of bias with large magnitude given a symmetric likelihood and relatively symmetric priors. As a result, both heading azimuth and elevation models converge to a set of bias predictions that are flat with relatively small bias magnitude. While this fit is quantitatively reasonable, as indicated by low RSE, the fit is qualitatively poor with an inability to match patterns of previously observed repulsive bias.

Next, we used an efficient Bayesian model that combines Bayesian probabilistic decoding with an efficient encoding stage. In contrast to our standard Bayesian model which is only capable of generating symmetric likelihoods with noise that increases with eccentricity, this model variant can generate asymmetric likelihoods. Under certain conditions, this can result in patterns of bias that are repulsive or vary between repulsive

and attractive as previous research observes in heading perception bias. Accordingly, we can achieve reasonable model fits for both heading azimuth and elevation perceptual bias.

Previous work demonstrated that patterns of repulsive biases are present in heading azimuth and elevation perception (Cuturi and MacNeilage, 2013; Crane, 2014), and this pattern of bias is difficult to model using standard Bayesian methods. Efficient Bayesian modeling methods are better able to capture repulsive bias generally, as well as the alternation between attractive and repulsive bias present in heading perception (Wei and Stocker, 2017). This holds true when using a non-parametric, empirically derived prior probability distribution as we demonstrate in the current study. While imperfect, empirically derived heading priors can be used in efficient Bayesian models to predict patterns of heading perception bias that fit observed patterns reasonably well.

### Limitations

One possible limitation is our selection of calibration method. As heading is the direction of linear velocity in a head-centered reference frame, it is necessary to register the linear velocity estimate given by the T265 to a given participant's head position. We do this through a kinematic calibration method (Sinnott, Hausamann, and MacNeilage, 2023) that relies on systematic head rotation and visual alignment of anatomical landmarks (Reid's Plane) to construct a kinematic reference frame that data is transformed into. It is likely that this calibration method introduces some amount of noise as alignment of Reid's plane may slightly differ from participant to participant. Future kinematic calibration methods may be made purely kinematic, such that they do not rely on visual alignment, thereby reducing calibration noise.

Another limitation may be due to our data collection protocol. In contrast to previous work, we purposely decided not to prescribe any specific behaviors for participants to complete over a recording session. This was to avoid possible sampling biases from pre-selecting activities experimenters deem important or "natural" but has an unintended consequence of moving sampling bias from the experimenter to the participant. It is difficult to discern whether this sampling bias is better than that induced directly by the experimenter's choice of activities, but it is hard to argue that sampling bias is not present at all. Further sampling bias may be driven by presence of the recording equipment itself.

In order to record head movements, participants had to wear portable lab equipment while they completed various activities. Certain activities were excluded by nature of them being impossible to complete safely while wearing equipment (e.g., swimming) or without compromising participant privacy (e.g., purchasing item with credit card). Although participants had the ability to protect privacy through use of the camera shutter, it is likely that participants, whether consciously or not, only performed a subset of natural activities that they felt comfortable doing while recorded. Future studies may alleviate this concern by collecting large amounts of data, combining prescribed and unprescribed activities, and/or recording head position in a more inconspicuous manner.

A methodological limit concerns our method of fitting models to psychophysical data. In the current study, we fit model predictions to the overall pattern of bias observed in previous research (Cuturi and MacNeilage, 2013; Crane, 2014). While this is not an incorrect method per se, it is likely that qualitatively and/or quantitatively closer fits

could be achieved by using models to simulate individual subject performance. This requires the use of individual subject response data; this data from previous studies we used to fit our models was not available to us.

Another major limitation is related to our method of model fitting through titration of noise parameter values. With respect to the standard Bayesian model, we automatically search through a range of parameter values and find the one that minimizes distance (via RSE) between the models' predicted patterns of bias and patterns previously observed in prior work. This results in a relatively flat pattern of predictions with overall low error because heading azimuth and elevation exhibit alternating attractive and repulsive patterns of bias. A different metric of distance or fit that is appropriate to use with non-linear functions could potentially perform better. It should be noted however that RSE, unlike other common measures (such as  $R^2$ ), is appropriate for evaluating fit of all non-linear functions.

It is possible that use of a similar method to search through both noise parameter values in the efficient Bayesian model converges on a similar solution, as it is possible to generate predicted fits that emulate those of a standard Bayesian model by effectively eliminating noise applied to the likelihood at the encoding stage. Unfortunately, for reasons unclear to us, not all sets of parameter values result in plausible patterns of predicted bias from the efficient encoding model. In some cases, certain sets of parameter values result in extreme bias values that are not physiologically possible. Further work is needed to understand how implementation of this model with our empirical prior and parameter selection can result in these erroneous predictions.

A final consideration is how we define heading. While it is typically agreed upon that heading refers to the direction of linear velocity in a head-based reference frame, it is less clear whether all periods in which the head moves should be associated with heading. Classically, heading is often referred to during times when an organism is moving with an appreciable amount of speed along some path (e.g., during locomotion). However, even during periods of time or activities that may be considered "stationary", the head and otolith organs experience linear movement. As one stands still, respiration and postural sway will cause changes in the head's linear position in the world. Likewise, a momentary, volitional rotation of the head will cause some linear displacement of the otolith organs due to the center of rotation of the head being at the vertebral column. While direction of linear velocity during these time points may not align with the conventional definition of heading used in the context of locomotion or navigation, estimating direction of linear velocity during these time points is nevertheless biologically relevant.

### Future Directions

A first avenue for future research would be modeling individual subject performance, or bias, in a manner constrained by that same individual's empirically measured prior. In the current study we observe some individual differences in heading azimuth and heading elevation distributions. In particular, there appears to be high inter-subject variability for rightward heading azimuth and for high-velocity heading elevation. It may be possible to predict individual differences in heading perception bias from these differences in heading azimuth and elevation prior.

Likely, some of the inter-subject variability we observe is driven by differences in tasks completed by each subject, as well as by anthropometric differences among subjects in our study. While we did not have participants record their own activities during the experiment, it may be possible to recover, reconstruct, or otherwise estimate activities participants completed from moment-to-moment. We attempt this heuristically by parsing data into low- and high-velocity categories based on instantaneous linear velocity magnitude, which serves reasonably well to separate locomotion from other activities. As we see a large difference in the statistical structure of high-velocity heading relative to heading collected at epochs of all velocities, it is likely that some other activities have relatively different structure as well.

A logical extension of this line of work is to investigate task-dependent perceptual bias by probing parameters of heading that differ in one task versus another. One way to do this is through investigating heading bias as a function of magnitude (e.g. speed). This is relatively simple to implement for measuring visual heading (via optic flow of varying magnitudes), but difficult with respect to vestibular heading as most means of applying linear translation to a subject have a limited range of motion. A Bayesian modeling approach could predict qualitatively different patterns of bias with our multimodal heading azimuth distribution observed at all velocity epochs versus our unimodal, nearly Gaussian heading azimuth distribution observed during high-velocity epochs.

### Conclusion

In the current study we present natural statistics of heading observed during natural, unconstrained tasks. These data can constrain models of sensory processing with

respect to vestibular and visual heading perception, with varying degrees of success.

Collection of this data has highlighted multiple avenues of future research, as statistics of heading appear to differ both by subject and potentially by task. In particular, heading statistics observed during periods of activity that are predominantly locomotion vary greatly with respect to periods of stationary activity (e.g. sitting, standing quietly).

## References

- Angelaki, D. E. and Cullen, K. E. (2008). Vestibular system: The many facets of a multimodal sense. *Annual Review in Neuroscience*, 31, 125-150.
- Sweeny, T. D., Haroz, S. and Whitney, D. (2012). Reference repulsion in the categorical perception of biological motion. *Vision Research*, 64, 26-34
- Barreira, T. V., Rowe, D. A. and Kang, M. (2010). Parameters of walking and jogging in healthy young adults. *International Journal of Exercise Science*, 3 (1), 4-13.
- Bowers, J. S., & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, 138(3), 389–414.
- Britten, K. H. (2008). Mechanisms of self-motion perception. *Annual Review of Neuroscience*, 31, 389-410.
- Chen, A., DeAngelis, G. C. and Angelaki, D. E. (2011). Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. *The Journal of Neuroscience*, 31 (33), 12036-12052.
- Chen, A., DeAngelis, G. C. and Angelaki, D. E. (2013). Functional specializations of the ventral intraparietal area for multisensory heading discrimination. *The Journal of Neuroscience*, 33 (8), 3567-3581.
- Chen, A., Gu, Y., Liu, S., DeAngelis, G. C. and Angelaki, D. E. (2016). Evidence for a casual contribution of macaque vestibular, but not intraparietal, cortex to heading perception. *The Journal of Neuroscience*, 36 (13), 3789-3798.
- Chen, A., Zeng, F., DeAngelis, G. C. and Angelaki, D. E. (2021). Dynamics of heading and choice-related signals in the parieto-insular vestibular cortex of macaque monkeys. *The Journal of Neuroscience*, 41 (14), 3254-3265.
- Cho, J. and Sharp, P. E. (2001). Head direction, place, and movement correlates for cells in the rat retrosplenial cortex. *Behavioral Neuroscience*, 115(1), 3–25.
- Clark, T. K., Newman, M. C., Karmali, F., Oman, C. M., and Merfeld, D. M. (2019). Chapter 5 – Mathematical models for dynamic, multisensory spatial orientation perception. In Ramat, S. and Shaik, A. G. (Eds.), *Mathematical Modeling in Motor Neuroscience: State of the Art and Translation to the Clinic. Ocular Motor Plant and Gaze Stabilization Mechanisms*, volume 248 of *Progress in Brain Research* (pp. 65-90). Elsevier.
- Crane, B. T. (2014). Human visual and vestibular heading perception in the vertical planes. *Journal of the Association for Research in Otolaryngology*, 15, 87-102.

- Cuturi, L. F. and MacNeilage, P. R. (2013). Systematic biases in human heading estimation. *PLoS One*, 8 (2), 1-11.
- de Vrijer, M., Medendorp, W. P. and van Gisbergen, J. A. M. (2009). Accuracy-precision trade-off in visual orientation constancy. *Journal of Vision*, 9 (2), 1-15.
- Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415, 429-433.
- Fetsch, C. R., DeAngelis, G. C. and Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience*, 14 (6), 1-27.
- Gu., Y., Watkins, P. V., Angelaki, D. E. and DeAngelis, G. C. (2006). Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *The Journal of Neuroscience*, 26 (1), 73-85.
- Gu., Y., Angelaki, D. E. and DeAngelis, G. C. (2009). Neural correlates of multi-sensory cue integration in macaque area MSTd. *Nature Neuroscience*, 11 (10), 1201-1210.
- Gu., Y., Fetsch, C. R., Adeyemo, B., DeAngelis, G. C. and Angelaki, D. E. (2010). Decoding of MSTd population activity accounts for variations in the precision of heading perception. *Neuron*, 66 (4), 596-609.
- Hausamann, P., Sinnott, C. B., Daumer, M. and MacNeilage, P. R. (2021). Evaluation of the Intel RealSense T265 for tracking natural human head motion. *Scientific Reports*, 11 (12496), 1-12.
- Knill, D. C. and Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27 (12), 712-719.
- Lappe, M., Bremmer, F. and van den Berg, A. V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Science*, 3 (9), 329-336.
- Laurens, J. and Droulez, J. (2007). Bayesian processing of vestibular information. *Biological Cybernetics*, 96 (4), 389-404.
- Longuet-Higgins, H. C. and Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 208 (1173), 385-397.
- MacNeilage, P. R., Banks, M. S., Berger, D. R. and Bühlhoff, H. H. (2007). A Bayesian model of the disambiguation of gravito-inertial force by visual cues. *Experimental Brain Research*, 179, 263-290.
- MacNeilage, P. R., Ganesan, N. and Angelaki, D. E. (2008). Computational approaches to spatial orientation: From transfer functions to dynamic Bayesian inference. *Journal of Neurophysiology*, 100, 2981-2996.
- Merchant, H., Battaglia-Mayer, A. and Georgopoulos, A. P. (2001). Effects of optic flow in motor cortex and area 7a. *Journal of Neurophysiology*, 86, 1937-1954.
- Mittelstaedt, M. and Mittelstaedt, H. (2001). Idiopathic navigation in humans: Estimation of path length. *Experimental Brain Research*, 139, 318-332.
- Peltier, N. E., Angelaki, D. E. and DeAngelis, G. C. (2020). Optic flow parsing in the macaque monkey. *Journal of Vision*, 20 (10),
- Sinnott, C., Dang, T., Papachristos, C., Alexis, K. and MacNeilage, P. (2018). Statistical characterization of heading stimuli in natural environments using SLAM. *Journal of Vision*, 18 (10), 41.

- Sinnott, C. B., Hausamann, P. A. and MacNeilage, P. R. (2023). Natural statistics of human head orientation constrain models of vestibular processing. *Scientific Reports*, 13 (5882), 1-12.
- Scott, D. W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons.
- Sun, H., Campos, J. L., Young, M. and Chan, G. S. W. (2004). The contributions of static visual cues, nonvisual cues, and optic flow in distance estimation. *Perception*, 33, 49-65.
- Wei, X. X. and Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain ‘anti-Bayesian’ percepts. *Nature Neuroscience*, 18 (10), 1509-1517.
- Wei, X. X. and Stocker, A. A. (2017). Lawful relation between perceptual bias and discriminability. *Proceedings of the National Academy of Science*, 114 (38), 10244-10249.
- Willemsen, S. C. M. J., Wijdenes, L. O., van Beers, R. J., Koppen, M., and Medendorp, W. P. (2022). Natural statistics of head roll: Implications for Bayesian inference in spatial orientation. *Journal of Neurophysiology*, 128, 1409-1420.
- Yakusheva, T., Blazquez, P. M. and Angelaki, D. E. (2008). Frequency-selective coding of translation and tilt in macaque cerebellar nodulus and uvula. *Journal of Neuroscience*, 28 (40), 9997-10009.
- Zhang, T. and Britten, K. H. (2011). Parietal area VIP causally influences heading perception during pursuit eye movements. *The Journal of Neuroscience*, 31 (7), 2569-2575.
- Zhang, W., Wang, H., Chen, A., Gu, Y., Lee, T. S., Wong, K.Y.M. and Wu, S. (2019). Complementary congruent and opposite neurons achieve concurrent multisensory integration and segregation. *eLife*, 8 (e43753), 1-33.

## CHAPTER 4: General Discussion

### Summary

Understanding the distribution of stimulus values that can comprise an organism's previous experience with a given stimulus is critical for understanding perceptual processing with the framework of a quantitative model. We show evidence that head orientation and heading, two head-based sources of information used for spatial orientation perception, are not normally distributed; all distributions differ dramatically from a Gaussian distribution, violating previous assumptions of a prior distribution's shape that are ubiquitous in the modeling literature. These have major implications for patterns of predicted perceptual performance, particularly in the context of Bayesian probabilistic modeling frameworks.

We use these natural distributions of heading and head orientation to generate prior probability distributions that empirically constrain Bayesian models of processing, and show that models using these distributions as a prior predict observed patterns of perceptual bias well in some cases. Head roll and heading azimuth are well accounted for when using an empirically derived prior with standard and efficient Bayesian models, respectively. A poorer fit is achieved when using these empirical priors to model heading elevation; qualitative fit of the model is reasonable, but magnitude of predicted bias appears smaller than observed bias by some scale factor. Fit for pitch perception is poorest, with the standard Bayesian model used to model performance unable to account for the specific pattern of bias magnitude and repulsion present in previous work measuring pitch perception.

Trends in observed natural head movement reveal future lines of research. Head pitch appears highly asymmetric and exhibits high inter-subject variability, while heading appears to change drastically as a function of speed (and by proxy, task). Outside of perception research, measurements of natural activities that comprise such an organism's experience also have high utility in applications such as human activity recognition, augmented reality, and clinical healthcare.

## Future Work

### Head Pitch Perception

Data presented in Chapter 2 reveals an interesting trend that has not been previously observed within perception literature. Head pitch relative to gravity appears asymmetrical across all participants and exhibits high levels of inter-participant variability. Some participants have a pitch distribution which is biased towards downward pitch, while others exhibit upward pitch bias. Likely, these indicate anthropometric differences between individuals, which constrain the range of pitch angles the head makes during stereotyped natural behaviors such as locomotion or standing upright. As a result, a person's prior experience will likely be shaped by this over the course of the lifespan.

Little previous research investigates perception of head or body pitch relative to gravity, and it remains a broadly understudied line of research in contrast to head roll relative to gravity. One study conducted by Cohen and Larson in 1974 found asymmetric patterns of pitch perception, while another more contemporary study conducted in 2023

found no significant difference between forwards and backwards pitch perception in complete darkness (Serpellet et al., 2023), though this was over a limited range of pitch values ( $\pm 15^\circ$ ) and only over 12 trials.

To some extent the lack of pitch research is driven by difficulty in applying vestibular or whole-body pitch stimuli to human subjects. Cohen and Larson used hardware resembling a large gimbal that functioned as a motorized tilt table to apply pitch positions to the head and body, while Serpellet and colleagues used a full helicopter flight simulator. It is likely that any task probing vestibular pitch perception will rely on use of a mechanical method like those listed above that may be rather expensive to deploy and maintain.

Visual pitch stimulation is also hard to apply; unlike roll perception, which has classic visual tasks (subjective visual vertical, rod-and-frame task), methods which use visual stimuli to probe pitch perception have not been established. A potential means of probing pitch perception visually may be through use of tasks resembling those used to probe orientation illusions along the roll axis (Howard and Hu, 2001; Allison et al., 1999). In these studies, researchers used a purpose-built facility that allowed them to rotate a room around a subject standing, sitting, or lying supine. Similarly, future researchers could use a similar task in virtual reality where participants must adjust the visual scene such that it is perceived as level (or perpendicular to gravity). This visual scene could have high fidelity, in line with original research discussed earlier with many visual features, or could be relatively impoverished with few, abstract features like shape primitives.

To some extent, individual differences in one's head pitch prior can be modeled. Work presented in the supplemental materials of Chapter 2 (Appendix A) show results of modulating asymmetry in a prior with subsequent modeling of pitch perception bias. In a standard Bayesian model, this results in decreasing bias magnitude for pitch values which fall on the long tail of the asymmetric, skewed distribution. Large bias values are predicted on the other side of the distribution that has little support for highly eccentric pitch values. Due to the pattern of pitch bias observed in previous work (Cohen and Larson, 1974), poor fit is achieved with a standard Bayesian model regardless of how the prior is shaped. It is likely that use of the efficient Bayesian model could fit observed patterns of pitch better due to its ability to predict repulsive patterns of bias.

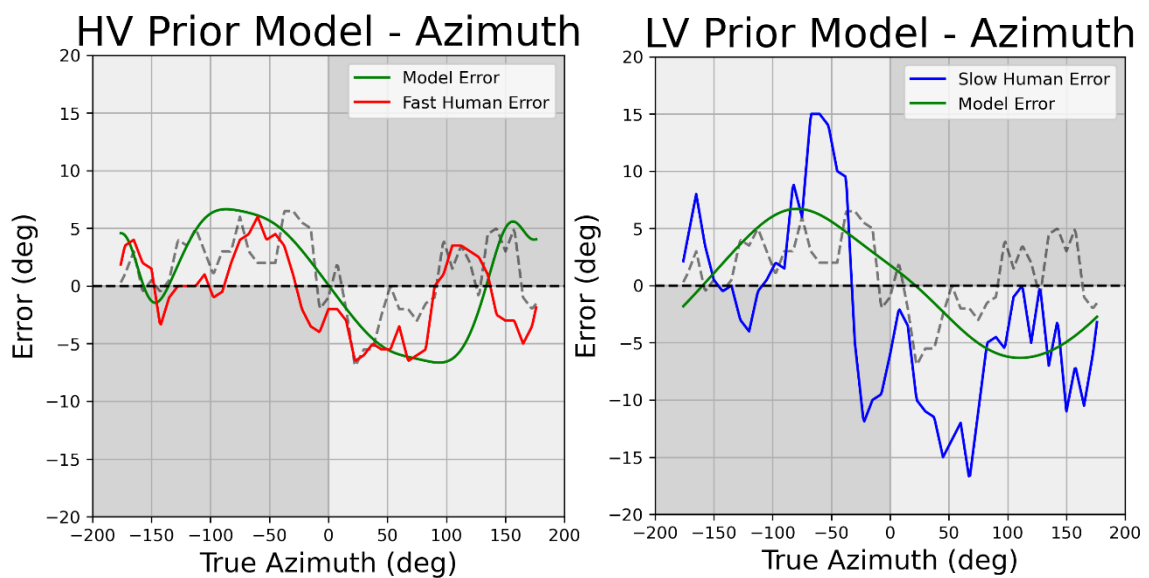
Regardless of specific modeling approach, a large component is the psychophysical values the model is fit to. More work is necessary to understand characteristics of pitch perception, including patterns of bias, before modeling work can be thoroughly validated. Serpollet and colleagues perform recent work measuring pitch perception error, but over a limited range of pitch angles (2023). This is a necessary first step; future work with greater participant and trial count, as well as a greater range of pitch values should be performed to begin fully characterizing human pitch perception.

### Heading Perception

Chapter 3 showed distributions of heading change dramatically as a function of speed. During low-velocity ( $< .75$  m/s) epochs, heading azimuth exhibits a multimodal distribution with peaks at straight-forward ( $0^\circ$ ) as well as purely leftward and rightward ( $\pm 90^\circ$ ). During high-velocity epochs, heading azimuth exhibits a unimodal, nearly

Gaussian distribution with its sole peak at straight-forward and variability greatly reduced relative to the multimodal distribution. Use of either of these priors in a standard or efficient Bayesian model should result in major differences in predicted bias; an open question is whether these changes in modeled performance would reflect true differences in human performance when judging heading direction while modulating speed.

Previous work has investigated whether patterns of vestibular heading perception change as a result of changing speed of heading stimuli. Unpublished work from Cuturi (2013) indicates that speed can affect heading perception. Performance seems to be worse when judging vestibular heading during slow movement, while performance increases when movement is faster. Bias magnitude increases when judging slow movement (e.g. greater over- or underestimation), and variability in the judgment increases. In the context of a standard Bayesian model, can these differences be explained by the empirically observed prior alone?



**Figure 1:** Efficient Bayesian models using high-velocity (left) and low-velocity (right) heading azimuth priors to predict high (red) and low-speed (blue) patterns of bias from (Cuturi 2013). X- and Y- axes represent true heading angle and heading angle estimation error, respectively. Dark grey regions of the plot represent attractive bias (underestimation error) while light grey regions represent repulsive bias (overestimation error). Heading azimuth from published work (Cuturi and MacNeilage, 2013) are plotted in the dashed black line.

Seemingly, the answer is no. With the low- and high-velocity heading distributions observed in chapter 3, we can generate priors and use them in an efficient Bayesian model. Magnitude of heading azimuth bias observed using the low-velocity prior is relatively small across all heading directions: likely due to the representation of all heading directions in the prior. Bias modeled using the high-velocity prior, in contrast, has reasonable magnitude at near straight-forward directions before becoming very noisy at high eccentricities. While models using different priors predict different patterns of bias, more work is needed to disentangle the effects of prior shape and likelihood variability when these distributions are non-Gaussian in an efficient Bayesian model.

Published work notes that the biases observed previously (Cuturi and MacNeilage, 2013) show patterns consistent with a Bayesian model “with priors peaked at the lateral directions” (pp. 61, Cuturi, 2013). Physiological work suggests that cells across MSTd have a non-uniform distribution of preferred heading direction, with greatest representation in the leftward and rightward directions corresponding to peaks in our heading prior (Gu et al., 2010). Supposing a non-uniform prior with peaks at leftward and rightward heading directions, and a high representation of neuron populations in MSTd with greatest selectivity for leftward and rightward heading directions; it stands to reason that observed and predicted biases would be smallest near these directions.

Indeed, vestibular heading bias is near zero for the lateral directions (Cuturi and MacNeilage, 2013), and modeling work presented in Chapter 3 using the efficient Bayesian model reflects this fact. Bias magnitude is near zero across the entire range of vestibular heading directions, however, making this finding far from conclusive. With respect to the shape of the prior, it is not fully clear how a multi-modal heading azimuth prior with peaks at lateral and forward directions determines the ultimate shape of predicted bias.

Repulsive biases are often described in heading perception and assume an unbiased prior with a mode at straight-forward. In this context, overestimation of a stimulus value truly does represent repulsion away from the prior and can be explained by efficient encoding. In the case of a multimodal prior, particularly like one we observe for low-velocity heading azimuth, it could be the case that bias is not repulsive but actually attractive towards modes with the greatest strength in the lateral directions. The extent of repulsive bias that is driven by one of these versus the other is hard to disambiguate, though. Modeling using separate high- and low-velocity priors does seem to reveal that a strong central mode is needed for appropriate bias magnitude, while support in lateral directions of the heading prior is also necessary to ensure bias magnitude is low throughout the range of possible values.

### Dynamic Models and Priors

A major application of data measured in chapters 2 and 3 is to constrain dynamic models of vestibular or multisensory processing. In approaches like the Kalman filter and particle filter, the system is initialized with some specific set of parameters which are

updated as the filter operates. A straightforward application of these data is to use them to initialize either of these filters. This may result in greater estimation performance early in the filter's runtime and avoid issues particular to particle filters wherein they converge on a poor initial set of particles and use them for subsequent updates.

Due to how these estimators function, the system's initial state and parameters are quickly replaced by updated values as the system progresses through the update and comparison steps present in each of these filters. It is generally assumed that the human nervous system has some persistent representation of past experience with a given stimulus that shapes current perception. Current implementations of particle filters account for this by assuming all previous states are accounted well by a Markov process, and as such only the immediately previous state implicitly contains information from all previous states. The Kalman filter, in contrast, doesn't account for previous information or experience before the initial system state.

Investigation of the best way to incorporate some overarching "prior" that shapes dynamic estimates given by the Kalman and particle filters may provide insight into how the nervous system switches in a mode- or task-dependent manner. It could be the case that the nervous system has some "general" representation of a prior based on all experience, which is rapidly updated for a specific task or mode that an organism is operating in, in a manner analogous to sensory adaptation. After the task or mode is complete, the state decays to a baseline state estimate that is heavily influenced by the prior. Future work to implement a prior derived from natural stimulus statistics in a

dynamic Bayesian manner will likely be a large inflection point for future work in perception.

### Human Activity Recognition and Clinical Diagnostics

Another promising application for data and findings of this dissertation is in human activity recognition (HAR). HAR broadly refers to the automated estimation of human behaviors based off sensor data such as video or inertial data from an IMU. A large use-case of HAR methods and hardware is in both acute and preventative healthcare, as sensors used to recognize falls in older adults are a long established, highly-specialized form of HAR. As smartphones and wearables have become ubiquitous, HAR has also become ubiquitous as millions of users wear or carry body-fixed sensors such as smartwatches that produce data usable for HAR. Developing methods of HAR that are not body-fixed (such as a smartphone in a pocket) is an active line of research (for review, see Straczekiewicz, James, and Onnela, 2021; Guidoux et al., 2014), as it is highly likely people hold or wear devices in different positions on the body or move them over the course of their operation.

A burgeoning subset of wearables consist of smartglasses or other head-mounted devices, particularly as companies work to develop augmented reality devices for commercial and recreational sectors. As these devices are used by more and more people, a future source of data used to categorize human activities will likely be head-centered inertial and/or positional data like those presented in this study. While it is likely that these data will be readily usable for categorizing acute events involving large-magnitude or velocity head movement (such as a fall), fine-grained analysis of activities like

walking or exercise may require knowledge of how head movement differs from limb or other bodily movement. Collecting data like those in this study is the first step towards understanding these differences in natural activities, but immediate work to categorize these data based on tasks that participants performed is a second, critical step.

Categorized or labelled head movement data will allow adaptation of extant machine learning based HAR methods to head movement data. While heuristic or threshold-based methods have been used for HAR, most contemporary methods are machine learning-based and may require a labelled training set (for review, see Preece et al., 2009). It is likely that extant datasets that provide labeled human activities alongside visual or inertial data aren't suitable for training new, head-based machine learning models for HAR as they don't have sensors mounted or worn on the head.

From a design or usability perspective, accurate parsing of human activities using head-based data can constrain content displayed to users of head-based wearables such as smartglasses. For instance, one can imagine the situation where a person begins exercising with such a wearable, only for the wearable to correctly recognize this person's activity and dynamically display exercise relevant information such as heart rate, a run timer, and local jogging routes or waypoints. A natural extension of this methodology would be the use of head-based HAR for clinical monitoring or measurement.

Clinical applications and head-based HAR methods are heavily constrained by the sensor suite available on such a head-based wearable. A head-based wearable could detect a wearer's fall through monitoring IMU data, for instance, and if connected to a

smartphone or similar device could automatically call a service provider. More uncommon sensors, such as a thermometer, could advise the user if a sustained elevated temperature indicative of a fever is detected.

## Conclusions

In total, work presented in this dissertation quantifies characteristics of head orientation relative to gravity and heading. Head roll is unimodally distributed with high excess kurtosis, while head pitch is unimodal, but also asymmetric and highly variable. Heading azimuth exhibits multiple modes in its distribution, with peaks at straight forward, leftward, and rightward while heading elevation is unimodal with high variability. At time of writing, no previous work has characterized head orientation and heading in this way from long-term, unconstrained, natural activities in humans.

In this dissertation we also present work using these data to constrain models of vestibular sensory processing. Using a standard Bayesian model, we are able to predict observed patterns of head roll bias well, by using our observed distribution of natural head roll as a prior. Using an efficient Bayesian model, we are also able to predict observed patterns of heading azimuth bias well using our heading azimuth distribution as a prior as well. Performance of the standard Bayesian model is poor when predicting head pitch perception, but this performance is confounded by relatively little research investigating head pitch perception as well as the specific model we use in Chapter 2. Modeling heading elevation perception results in acceptable qualitative fit, with a moderate recreation of observed patterns of heading elevation perception bias: but scale

of the model's predictions is poor as observed bias has greater magnitude than that predicted by our model.

Unexpectedly, we observe several interesting trends in our natural head movement data that may lead to future lines of research. First, distributions of head pitch relative to gravity appear to differ significantly on a participant-wise basis. Some participants pitch distributions exhibit modes at downward pitch angles, while others have modes at upward pitch angles. Likely these correspond to anatomical or biomechanical differences between participants and are relatively stable across the lifespan, and it is possible that these factors constrain or bias one's pitch prior. Future work investigating pitch perception broadly (as it is understudied) as well as linking one's pattern of pitch bias with their relatively individual prior is promising.

Another interesting trend is that we observe dramatic differences in vestibular heading distributions when looking at heading distributions generated from high ( $>0.75$  m/s) velocity and low ( $<0.75$ ) velocity data. Across all velocities, heading azimuth exhibits a multimodal distribution with the most dominant mode at straight-forward, and secondary modes at leftward and rightward heading directions. Heading azimuth at low velocities exhibits a similar multimodal distribution, but modes at all three directions described above are nearly equal in terms of frequency. In contrast, heading azimuth at high velocities is a unimodal distribution that is nearly Gaussian.

## References

Cuturi, L. F. and MacNeilage, P. R. (2013). Systematic biases in human heading estimation. *PLoS One*, 8 (2), 1-11.

- Cuturi, L. F. (2013). *Intrinsic and Induced Biases in Self-Motion Perception*. [Doctoral dissertation]. Ludwig-Maximilians-Universität München.
- Gu., Y., Fetsch, C. R., Adeyemo, B., DeAngelis, G. C. and Angelaki, D. E. (2010). Decoding of MSTd population activity accounts for variations in the precision of heading perception. *Neuron*, *66* (4), 596-609.
- Guidoux, R., Duclos, M., Fleury, G., Lacomme, P., Lamaudière, N., Manenq, P., Paris, L., Ren, L. and Rousset, S. (2014). A smartphone-driven methodology for estimating physical activities and energy expenditure in free living conditions. *Journal of Biomedical Informatics*, *52*, 271-278.
- Preece, S. J., Goulermas, J. Y., Kenney, L. P. J., Howard, D., Meijer, K. and Crompton, R. (2009). Activity identification using body-mounted sensors—a review of classification techniques. *Physiological Measurement*, *30* (4), R1.
- Strackiewicz, M., James, P. and Onnela, J. (2021). A systematic review of smartphone-based human activity recognition methods for health research. *npj Digital Medicine*, *4* (148), 1-15.
- Serpollet, D. K., Hartnagel, D., James, Y., Buffat, S., Vayatis, N., Bargiotas, I. and Vidal, P. (2023). Tilt perception is different in the pitch and roll planes in human. *Physiological Reports*, *11* (3), 1-11.

Appendix A:  
**Natural statistics of human head orientation  
 constrain models of vestibular processing**  
*Supplementary material*

Christian B. Sinnott<sup>1,\*</sup>, Peter A. Hausamann<sup>2</sup>, and Paul R. MacNeilage<sup>1</sup>

<sup>1</sup>University of Nevada, Department of Psychology, Reno, 89557, United States of America

<sup>2</sup>Technical University of Munich, Department of Electrical and Computer Engineering, Munich, 80333, Germany

\*csinnott@nevada.unr.edu

**Bayesian model of perception of head orientation**

Bayes' theorem states that the conditional probability of an event  $x$  given event  $y$  (posterior distribution) is equal to the product of the conditional probability of  $y$  given  $x$  (likelihood) and the probability of event  $x$  (prior), divided by the probability of event  $y$  (marginal).

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \quad (1)$$

To model orientation perception, we substitute  $\theta$  for  $x$ :

$$P(\theta|y) = \frac{P(y|\theta)P(\theta)}{P(y)} \quad (2)$$

$P(\theta)$  represents the prior probability across head orientation values based on the organism's prior experience.  $P(y|\theta)$  is the likelihood, that is the probability of the current sensory information  $y$  given true head orientation  $\theta$ . The posterior distribution  $P(\theta|y)$  governs the perceptual estimate resulting from the model. It represents the conditional probability of true head orientation  $\theta$  given current sensory information  $y$ . The marginal distribution in the denominator,  $P(y)$ , only serves to scale the product of the prior and likelihood distributions<sup>1</sup>, so the equation can be simplified to:

$$P(\theta|y) \propto P(y|\theta)P(\theta) \quad (3)$$

In Eq. 3, the posterior distribution is proportional to the product of the likelihood and prior distributions. Our priors for head pitch and head roll models are kernel density estimates (KDEs) generated using our observed head pitch and head roll data. KDEs were generated using the `gaussian_kde` function in the `scipy.stats` library. Kernel bandwidth was determined using Scott's rule<sup>2</sup> (see Eq. 7).

In addition to these empirical priors, we also model the likelihood. Previous work shows increased variability of perceptual estimates as head orientation eccentricity increases, presumably due to vestibular sensory noise that increases with tilt angle. We model this noise on the likelihood in two ways.

$$f(\alpha, \sigma, \theta) = \alpha + \sigma \times |\theta| \quad (4)$$

The first is through a linear increase in noise as orientation eccentricity increases, shown in Eq. (4). Alpha is a constant representing a baseline level of noise determined from previous research<sup>3,4</sup>, while sigma is a Weber fraction denoting the proportion of signal dependent noise. Alpha is added to the product of sigma and the absolute value of theta; the orientation angle. Second, we increase noise with eccentricity non-linearly by using a sinusoidal function and incorporating two constants, G and K, to better simulate the effect of utricular shear experienced during linear acceleration of the head<sup>5</sup> (Eq. (5)).

$$f(\alpha, \sigma, \sin(|\theta|), K, G) = \alpha + \sigma \times K \times G \times \sin(|\theta|) \quad (5)$$

Alpha, sigma, and theta are the same terms here as they are in the linear noise model. G represents the constant linear acceleration imparted by gravity, while K is a scaling constant determined in previous research<sup>5</sup>. We use both these equations to apply noise to the likelihood distribution used in each of our models.

To generate model predictions for each value of sigma and each tilt angle, we take the mean of the posterior distribution where the posterior is calculated as the product of the noisy likelihood function (with noise determined by  $\sigma$  and Eq. (4) and (5) and our empirical prior, according to Bayes' theorem (Eq. (3)). Both of our models only contain one free parameter which is the level of signal dependent noise denoted by  $\sigma$ . Both linear and utricular shear models are first fit to extant psychophysical data observing bias in roll perception<sup>4</sup> by selecting the sigma value that minimizes distance between the observed error from psychophysical data and predicted error from the model. We repeat the same process using pitch perception data<sup>6</sup>. Our metric for distance is residual square error (RSE), calculated as:

$$RSE = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (6)$$

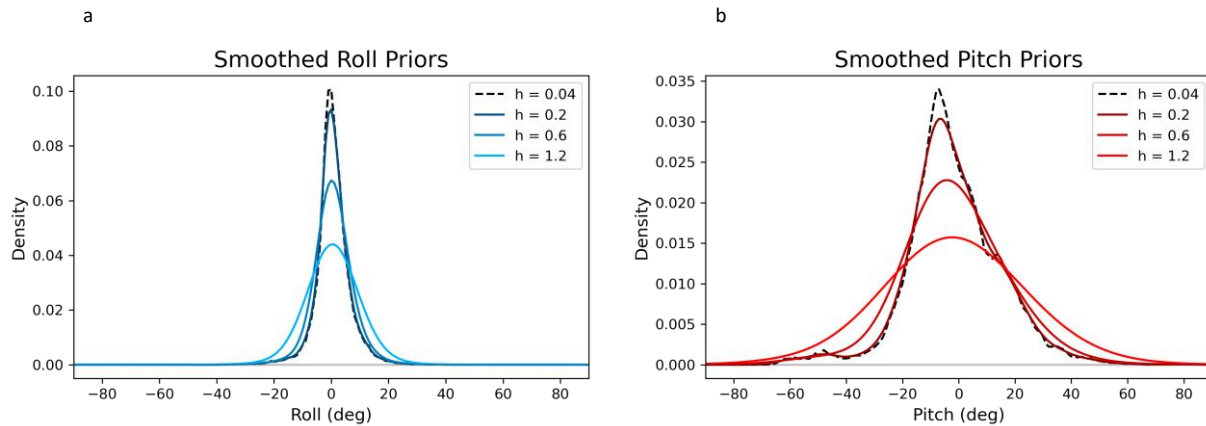
### Modeling with smoothed empirical priors

To understand the effect of differing prior statistics on our model's predictions, we perform a subset of modeling with empirical KDEs that are generated using a Gaussian kernel with increasingly larger bandwidth  $h$ . In the current study, empirical priors used for modeling in the main text use a value of  $h$  that is calculated using Scott's Rule:

$$h = n^{-\frac{1}{d+4}}$$

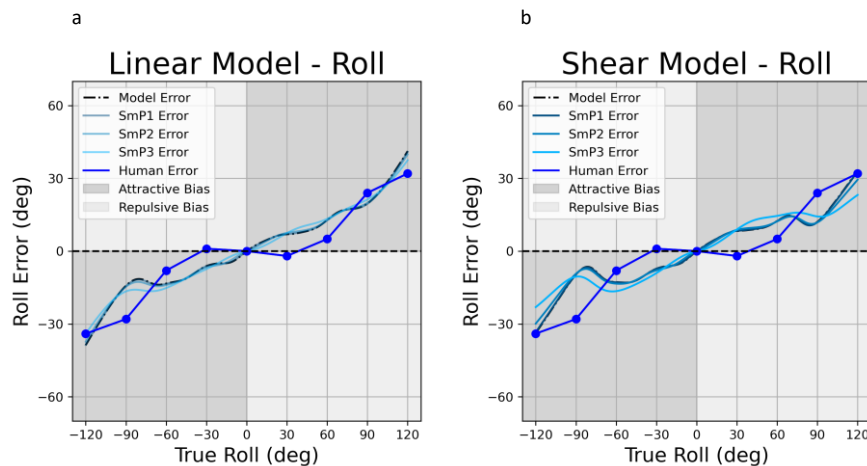
(7)

In this equation,  $n$  is equal to the length of the data to be smoothed, while  $d$  is equal to the number of dimensions in the data. Using this heuristic,  $h$  is approximately 0.04 for both empirical roll and pitch priors. By increasing the bandwidth used in KDEs of our roll and pitch priors, we can create versions of these distributions that have decreasing amounts of excess kurtosis and asymmetry relative to distributions generated with an  $h$  value determined by Scott's Rule (Fig. S1).



Supplemental Figure S1. Empirical roll (S1a) and pitch (S1b) priors smoothed with increasingly larger Gaussian kernel bandwidths ( $h$ ). The dashed black line represents the KDE with bandwidth size calculated using Scott's Rule that was used in main analyses, while colored lines represent KDEs with increasingly larger bandwidths.

By using these smoothed distributions as priors in either of our models, we can probe model performance as a function of increased Gaussian kernel smoothing, and by proxy: reduced asymmetry and excess kurtosis. Like before, we find the value of  $\sigma$  that minimizes RSE between the model's prediction error using a given prior and observed psychophysical error from pre-existing work. Reducing excess kurtosis in this way for our roll distribution does not appear to drastically effect model fit, as every linear model using more smoothed versions of the empirical roll prior converged to use the same factor for multiplicative signal dependent noise ( $\sigma = 0.113$ ) and had similar amounts of RSE (Fig. S2). Similar fits were achieved with all variants of prior using the utricular shear model ( $\sigma = 0.013$ - $0.014$ ).

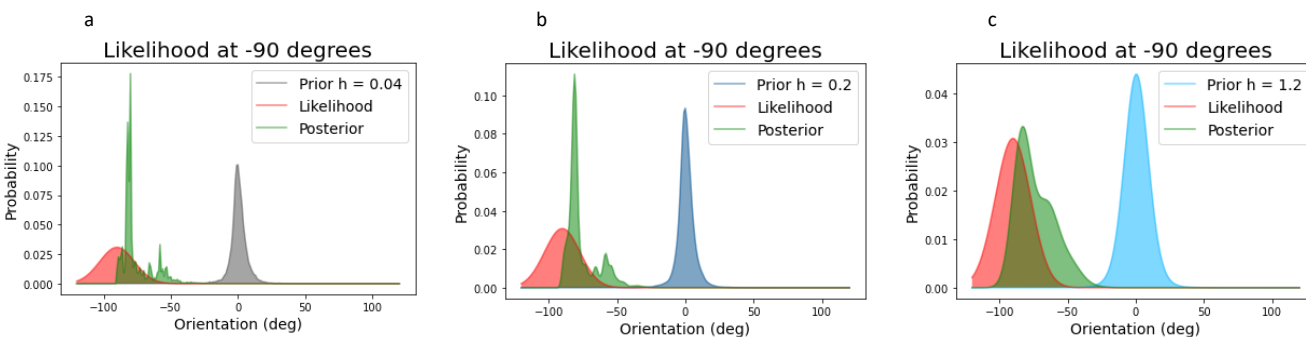


Supplemental Figure S2. Bayesian modeling results using roll priors of varying bandwidth for both linear (S2a) and shear (S2b) models. Original model results using the bandwidth determined by Scott's rule are plotted in dashed

black; results using increasingly smoothed priors are plotted in light blue. Psychophysical data are plotted in dark blue.

An interesting, unanticipated consequence of this process of smoothing is that local extrema observed in posteriors using the original, least smoothed prior are also smoothed, resulting in a slight change to the underlying pattern of bias that our models predict. Given how likelihood distributions are generated, we conclude that this is driven by local extrema in the empirical prior. For roll in particular, a bimodal posterior is calculated at certain eccentricities which may explain the particular pattern of roll bias observed in our models (Fig. S3).

Performing the same smoothing and fitting process for pitch, we see a reduction in the qualitative asymmetry of the models' predictions. However, both variants of our static Bayesian observer model are not able to predict a pattern of bias coinciding with the observed pattern of pitch perception bias in previous work<sup>6</sup>. Regardless of the level of smoothing used on the empirical pitch prior, our model minimizes to a set of predictions that have low error across nearly all pitch eccentricities. As with roll, every linear model using more smoothed versions of the empirical pitch prior converged to use the same factor for multiplicative signal dependent noise ( $\sigma = 0.001$ ) and had similar amounts of RSE. Every shear model using various smoothed pitch priors demonstrated a similar pattern, both with respect to reduced asymmetry in the model prediction as well as a convergence to minimal signal dependent noise ( $\sigma = 0.001$ ). All prior parameters, model parameters, and model performance metrics can be found in Table S2.

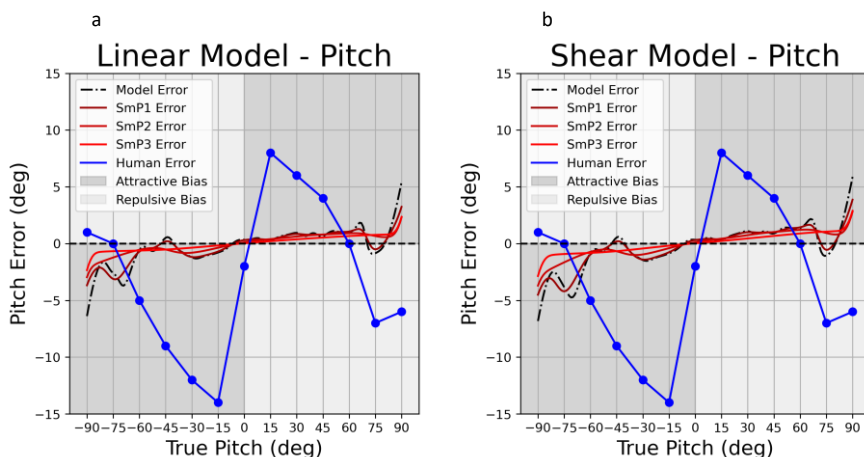


Supplemental Figure S3. Exemplar posteriors generated with likelihood at -90 degrees (left ear down) roll using priors generated using Gaussian kernels of varying bandwidth in the linear model. Signal dependent noise used in likelihood generation is the same across all three priors ( $\sigma = 0.001$ ). Bandwidth increases from 0.04 at far left (S3a, same as prior used in main text), 0.2 in center (S3b), and 1.2 at far right (S3c).

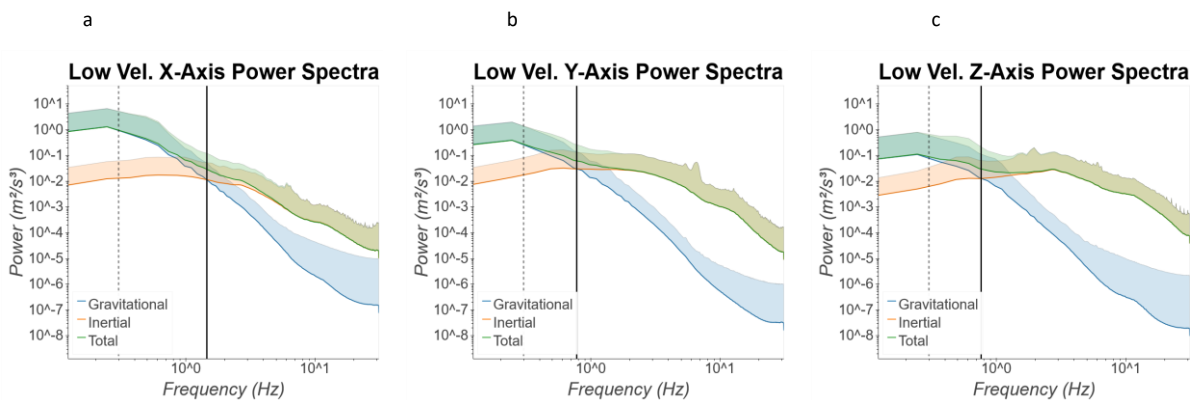
## Gravitational and inertial acceleration power spectra during low versus high velocity epochs

After data were separated into separate high- and low-velocity epochs, we conducted power spectra analyses on both sets of epochs. Like before, crossing points were observed along all three axes, though the exact point of these crossing points differed from crossing points observed across all data. During low-velocity epochs, we observed crossing points at approximately 1.451 Hz (Fig. S5a), 0.775 Hz (Fig. S5b), and 0.761 Hz (Fig. S5c) along X-, Y-, and Z-axes, respectively. During high-velocity epochs, we observed crossing points at approximately 1.108 Hz (Fig. S6a), 0.553 Hz (Fig. S6b), and 0.53 Hz (Fig. S6c) along X-, Y-, and Z-axes. Additionally, power spectra estimated during high-velocity epochs showed a number of transient peaks through the frequency space along all axes, with a notable peak at approximately 2 Hz

along X- and Z-axes corresponding with preferred stepping frequency<sup>7</sup>. No similar peaks were observed in the power spectra estimated during low-velocity epochs.



Supplemental Figure S4. Bayesian modeling results using pitch priors of varying bandwidth for both linear (S4a) and shear (S4b) models. Original model results using the bandwidth determined by Scott's rule are plotted in dashed black; results using increasingly smoothed priors are plotted in red. Psychophysical data are plotted in dark blue.



Supplemental Figure S5. Power spectra calculated from gravitational (blue), inertial (orange), and sum total linear acceleration (green) during low velocity ( $<0.75$  m/s) epochs along X- (S5a), Y- (S5b), and Z-axes (S5c). Figure axes are log-scaled. Approximate crossing points in relative power between gravitational and inertial acceleration are observed at 1.451 Hz, 0.775 Hz, and 0.761 Hz for X-, Y-, and Z-axes, respectively.

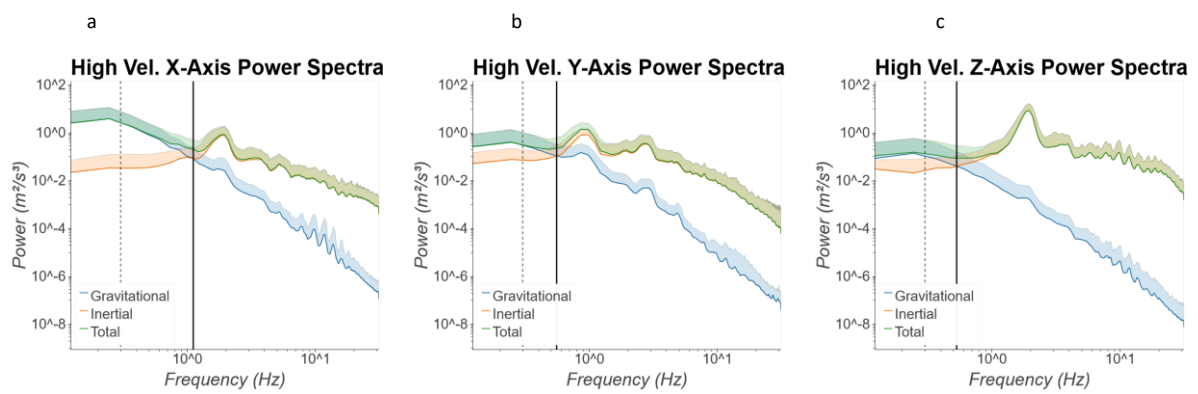
### Head orientation during low versus high velocity epochs

In addition to measuring head roll and head pitch across all participants and velocities, we measured head roll and head pitch across all participants during low- and high-velocity epochs. Given the proportion of total data that are comprised of low-velocity epochs (92.25% of all data), group-level distributions of head orientation as well as group-level KDEs of head roll and pitch look similar to their analogues across all data (Figure 2, main text). Conversely, the high-velocity head orientation distribution as well as high-velocity KDEs of head roll and pitch qualitatively differ from those measured across all velocity epochs (Fig. S8a). With respect to head roll, there appeared to be greater weight in the tails of the KDE relative to the KDE

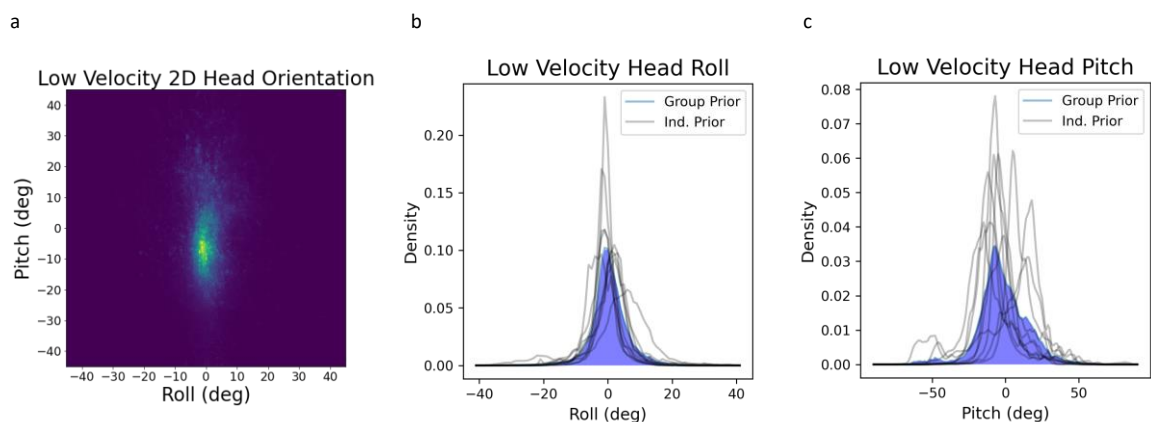
generated for roll across all velocities (Fig. S8b). Head pitch during high-velocity epochs appeared to be shifted further downward relative to head pitch across all velocities, both at a group and individual level (Fig. S8c). As high-velocity epochs primarily locomotion, this downwards shift is likely driven by concordant gaze-down behavior which functions to visually identify stable footholds<sup>8,9</sup>.

We also calculated moments of pitch and roll distributions during low and high-velocity epochs across all participants.

Descriptions of these moments as well as formulae used to calculate them can be found in Table S1. In general, these low-velocity histograms resembled histograms generated from both low- and high-velocity epochs. During low-velocity epochs, pitch was slightly biased downward ( $\mu_1 = -1.574^\circ$ ) while average roll was closer to zero ( $\mu_1 = 0.475^\circ$ ). Pitch had much higher variance ( $\mu_2 = 288.39^\circ$ ) than roll ( $\mu_2 = 39.673^\circ$ ). Both pitch ( $\mu_3 = -0.036$ ) and roll ( $\mu_3 = 0.144$ ) had little skewness during low-velocity epochs, while roll ( $\mu_4 = 7.297$ ) had greater excess kurtosis than pitch ( $\mu_4 = 1.808$ ).



Supplemental Figure S6. Power spectra calculated from gravitational (blue), inertial (orange), and sum total linear acceleration (green) during high velocity ( $\geq 0.75$  m/s) epochs along X- (S6a), Y- (S6b), and Z-axes (S6c). Figure axes are log-scaled. Approximate crossing points in relative power between gravitational and inertial acceleration are observed at 1.108 Hz, 0.553 Hz, and 0.53 Hz for X-, Y-, and Z-axes, respectively.

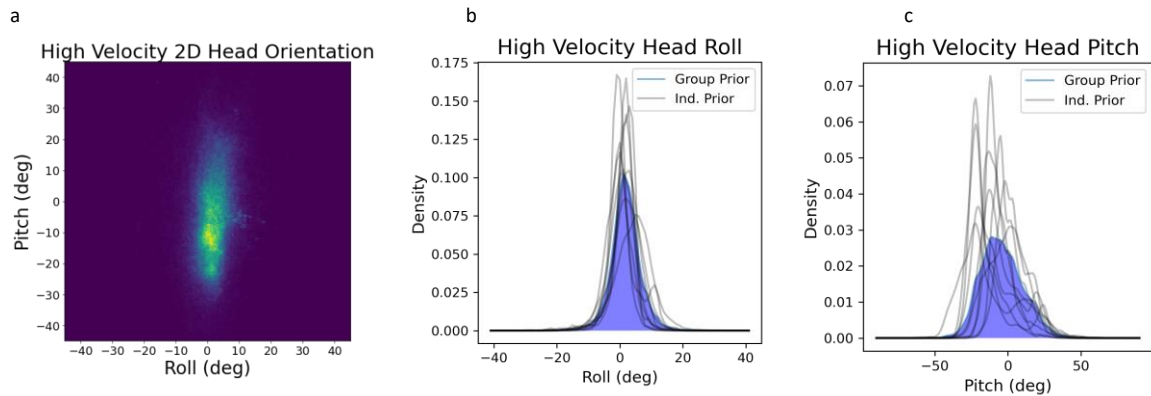


Supplemental Figure S7. 2D head orientation measured across all participants during low-velocity epochs. Marginal kernel density estimates (KDEs) for roll and pitch during low-velocity epochs are also plotted. The KDEs plotted in blue represent the distributions across all participants, while black traces represent KDEs for individual participants.

Histograms generated from high-velocity data looked qualitatively different from histograms generated from both low- and high-velocity data, and this is reflected in moments calculated from high-

velocity roll and pitch distributions. In particular, pitch appeared to show greater downward bias relative to the pitch KDE generated from all data. For high-velocity epochs, both roll ( $\mu_1 = 1.832^\circ$ ) and pitch ( $\mu_1 = -4.418^\circ$ ) were more biased away from upright. Pitch ( $\mu_2 = 208.98^\circ$ ) continued to show higher variance than roll ( $\mu_2 = 23.85^\circ$ ), and both pitch ( $\mu_3 = 0.331$ ) and roll ( $\mu_3 = 0.02$ ) showed low levels of skewness. Roll ( $\mu_4 = 4.028$ ) measured during these epochs continued to have higher excess kurtosis than pitch ( $\mu_4 = 0.143$ ).

We calculate moments of roll and pitch distributions for each subject as well (Tables S5, S6). Since we did not perform modeling using low- or high-velocity KDEs, we did not calculate their moments.



Supplemental Figure S8. 2D head orientation measured across all participants during high-velocity epochs. Marginal kernel density estimates (KDEs) for roll and pitch during high-velocity epochs are also plotted. The KDEs plotted in blue represent the distributions across all participants, while black traces represent KDEs for individual participants.

Moment	Description	Formula	Low Vel. Roll	Low Vel. Pitch	High Vel. Roll	High Vel. Pitch
Mean	The sum of a set of numbers divided by the count of that set.	$\mu_1 = \frac{1}{n} \sum_{i=1}^n$	0.475	-1.574	1.832	-4.148
Variance	The extent to which a set of numbers is spread out from the mean.	$\mu_2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$	39.673	288.39	23.85	208.98
Skewness	The extent to which a given distribution is asymmetrical or otherwise lopsided.	$\mu_3 = \frac{m_3}{m_3^{3/2}}$ <p>where:</p> $m_i = \frac{1}{n} \sum_{i=1}^n (x[n] - \bar{x})^i$	0.144	-0.036	0.02	0.331
Kurtosis	The weight of a given distribution's tails.	$\mu_3 = \frac{m_4}{m_2^{4/2}} - 3$ <p>where:</p> $m_i = \frac{1}{n} \sum_{i=1}^n (x[n] - \bar{x})^i$	7.297	1.808	4.028	0.143

**Supplemental Table S1.** Descriptions, formulae, and values for first four central moments used to describe low-velocity and high-velocity pitch and roll histograms. Moments are calculated from group-level data plotted in 2d histograms in [S7a](#) and [S8a](#)

Prior Type	Bandwidth ( $h$ )	Model	Multiplicative Noise ( $\sigma$ )	RSE
Emp. Roll 1	0.04	Linear	0.113	6.438
Emp. Roll 2	0.2	Linear	0.113	6.418
Emp. Roll 3	0.6	Linear	0.113	6.38
Emp. Roll 4	1.2	Linear	0.113	6.261
Emp. Roll 1	0.04	Shear	0.013	8.346
Emp. Roll 2	0.2	Shear	0.013	8.392
Emp. Roll 3	0.6	Shear	0.013	8.74
Emp. Roll 4	1.2	Shear	0.014	9.612
Emp. Pitch 1	0.04	Linear	0.001	6.774
Emp. Pitch 2	0.2	Linear	0.001	6.687
Emp. Pitch 3	0.6	Linear	0.001	6.713
Emp. Pitch 4	1.2	Linear	0.001	6.835
Emp. Pitch 1	0.04	Shear	0.001	6.795
Emp. Pitch 2	0.2	Shear	0.001	6.712
Emp. Pitch 3	0.6	Shear	0.001	6.713
Emp. Pitch 4	1.2	Shear	0.001	6.845

**Supplemental Table S2.** Kernel density estimates parameters, model parameters, and goodness of fit metrics for model predictions depicted in [S2a](#), [S2b](#), [S4a](#), and [S4b](#).)

Prior Type	Location	Scale	Asymmetry	Model	Multiplicative Noise ( $\sigma$ )	RSE
Gaussian Roll 1	0	5	N/A	Linear	0.001	5.736
Gaussian Roll 2	0	11.56	N/A	Linear	0.041	3.212
Gaussian Roll 3	0	15	N/A	Linear	0.062	3.016
Gaussian Roll 4	0	25	N/A	Linear	0.122	4.157
Gaussian Pitch 1	-7	12	N/A	Linear	0.001	7.105
Asymmetric Pitch 1	-18	22	2.5	Linear	0.001	6.782
Asymmetric Pitch 2	-19	32	5	Linear	0.001	7.199
Gaussian Roll 1	0	5	N/A	Shear	0.001	7.439
Gaussian Roll 2	0	11.56	N/A	Shear	0.007	4.585
Gaussian Roll 3	0	15	N/A	Shear	0.01	4.437
Gaussian Roll 4	0	25	N/A	Shear	0.021	4.157
Gaussian Pitch 1	-7	12	N/A	Shear	0.001	7.498
Asymmetric Pitch 1	-18	22	2.5	Shear	0.001	7.118
Asymmetric Pitch 2	-19	32	5	Shear	0.001	7.872

**Supplemental Table S3.** Distribution parameters, model parameters and goodness of fit metrics for simulation models using Gaussian and skew normal priors in main text. Location and scale parameters correspond with mean and variance of Gaussian distributions and other parameters were used to generate Gaussian and skew-normal distributions in Python 3 using Scipy library.

Subject	$\mu_{1Roll}$	$\mu_{2Roll}$	$\mu_{3Roll}$	$\mu_{4Roll}$	$\mu_{1Pitch}$	$\mu_{2Pitch}$	$\mu_{3Pitch}$	$\mu_{4Pitch}$
001	2.53	41.13	0.96	6.04	-5.8	200.74	0.7	1.05
002	0.43	34.46	-0.41	1.5	3.5	409.84	-0.85	0.19
003	0.73	30.57	-0.23	12.68	-11.11	462.81	-0.63	0.5
004	-0.48	11.8	-0.68	48.64	1.21	132.37	0.89	1.96
005	-1.81	27.41	-0.65	4.43	-2.05	158.67	1.41	4.01
006	5.45	59.04	0.08	4.9	10.77	177.56	1.54	4.79
007	-1.87	23.09	4.81	4.12	13.06	179.23	-0.05	1.1
008	-0.34	11.72	0.72	8.9	-5.69	75.3	1.38	4.84
009	-0.34	32.98	0.74	10.56	-8.21	168.42	1.7	4.71
010	-0.72	73.82	-1.65	3.46	-12.95	217.49	0.38	0.94

**Supplemental Table S4.** Moments of individual roll and pitch distributions across all epochs.

Subject	Low Vel. Data %	High Vel. Data %	$\mu_{1LVR}$	$\mu_{2LVR}$	$\mu_{3LVR}$	$\mu_{4LVR}$	$\mu_{1HVR}$	$\mu_{2HVR}$	$\mu_{3HVR}$	$\mu_{4HVR}$
001	80.7%	19.3%	2.56	45.85	0.98	5.49	2.41	21.25	0.32	6.93
002	98.6%	1.4%	0.41	34.67	-0.4	1.48	1.76	17.71	-0.83	3.36
003	94.8%	5.2%	0.7	31.55	-0.22	12.44	1.41	13.25	-0.69	9.33
004	92.9%	7.1%	-0.65	11.68	-0.66	52.89	1.76	7.98	-1.43	16.6
005	94.9%	5.1%	-1.88	27.89	-0.65	4.42	-0.55	17.02	-0.26	2.68
006	90.4%	9.6%	5.58	61.08	0.08	4.94	4.18	38	-0.2	1.83
007	92.2%	7.8%	-2.17	21.68	-0.13	4.71	1.34	26.82	0.35	1.56
008	97.8%	2.2%	-0.34	11.79	0.74	8.94	-0.55	8.67	-0.49	3.94
009	88.8%	11.2%	-0.43	34.37	0.84	10.78	0.36	21.01	-0.59	4
010	91.9%	8.1%	-0.96	77.19	-1.61	3.19	2.1	27.47	-0.77	2.29

**Supplemental Table S5.** Proportion of low- and high-velocity data for each subject, as well as moments of individual low( $\mu_{nLVR}$ ) and high-velocity ( $\mu_{nHVR}$ ) roll distributions.

Subject	Low Vel. Data Prop.	High Vel. Data Prop.	$\mu_{1LVP}$	$\mu_{2LVP}$	$\mu_{3LVP}$	$\mu_{4LVP}$	$\mu_{1HVP}$	$\mu_{2HVP}$	$\mu_{3HVP}$	$\mu_{4HVP}$
001	80.7%	19.3%	-6.98	202.95	0.89	1.61	-0.87	161.29	-0.02	-0.21
002	98.6%	1.4%	3.75	407.97	-0.88	0.27	-14.73	204.22	1.67	1.92
003	94.8%	5.2%	-10.82	477.27	-0.67	0.46	-17.09	127.05	1.63	4.21
004	92.9%	7.1%	1.81	129.99	0.91	2.1	-6.67	96.7	1.23	2.04
005	94.9%	5.1%	-1.76	151.86	1.55	4.58	-7.4	254.61	0.72	-0.51
006	90.4%	9.6%	11.44	177.55	1.62	5.02	4.43	133.16	0.88	1.21
007	92.2%	7.8%	14.04	170.27	-0.11	1.53	0.35	121.32	1.38	2.85
008	97.8%	2.2%	-5.66	75.45	1.38	4.89	-7.18	66.44	1.08	1.86
009	88.8%	11.2%	-8.6	162.13	1.85	5.64	-5.13	207.83	14.42	0.81
010	91.9%	8.1%	-12.94	203.23	-1.61	3.19	-12.97	381.13	0.59	-0.65

**Supplemental Table S6.** Proportion of low- and high-velocity data for each subject, as well as moments of individual low( $\mu_{nLVP}$ ) and high-velocity ( $\mu_{nHVP}$ ) pitch distributions.

## References

1. MacNeilage, P. R., Ganesan, N. & Angelaki, D. E. Computational approaches to spatial orientation: From transfer functions to dynamic bayesian inference. *J. Neurophysiol.* 100, 2891–2996 (2008).
2. Scott, D. W. *Multivariate Density Estimation: Theory, Practice and Visualization* (John Wiley & Sons, Inc., New York, New York, United States of America, 1992).
3. Vingerhoets, R. A. A., Medendorp, W. P. & van Gisbergen, J. A. M. Body-tilt and visual verticality perception during multiple cycles of roll rotation. *J. Neurophysiol.* 99, 2265–2280 (2008).
4. de Vrijer, M., Medendorp, P. & Van Gisbergen, J. A. M. Accuracy-precision trade-off in visual orientation constancy. *J. Vis.* 9, 1–15 (2009).
5. Schone, H. On the role of gravity in human spatial orientation. *Aerosp. Medicine* 35, 764–772 (1964).

6. Cohen, M. M. & Larson, C. A. Human spatial orientation in the pitch dimension. *Percept. Psychophys.* 16 (1974).
7. 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).
8. 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).
9. Pelz, J. B. & Rothkopf, C. Chapter 31 - oculomotor behavior in natural and man-made environments. In Van Gompel, R. P., Fischer, M. H., Murray, W. S. & Hill, R. L. (eds.) *Eye Movements*, 661–676 (Elsevier, Oxford, 2007).