

# A Bayesian approach to modelling spatial orientation during roll rotation

*Maarten van der Heijden*

Supervisors: Rens Vingerhoets<sup>1,2</sup>, Jan van Gisbergen<sup>1</sup> and Pieter Medendorp<sup>2,3</sup>

<sup>1</sup>Department of Biophysics, Radboud University Nijmegen

<sup>2</sup>Nijmegen Institute for Cognition and Information, Radboud University Nijmegen

<sup>3</sup>FC Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen

## Abstract

Spatial orientation in darkness relies mostly on information from the semicircular canals and the otoliths. The canals sense angular orientation and the otoliths can detect both tilt and linear acceleration caused by translation. One hypothesis on how the inherently ambiguous otolith signal can be interpreted is the canal-otolith interaction model. This model has recently been reformulated in a Bayesian framework by Laurens and Droulez (2007). We have explored whether an extended version of this model can account for body tilt and for visual-verticality perception during roll rotation. In separate experiments, we tested five subjects on their ability to estimate subjective body tilt (SBT) and the subjective visual vertical (SVV) during constant velocity roll rotation. In addition we measured the SVV under static conditions. Subjects performed relatively well in the SBT task, showing random errors but no systematic bias. In the SVV tasks all subjects made large systematic errors that could be classified in two response modes consistent with A-effects and E-effects reported in the literature. Model simulations showed that the SBT data could be reproduced quite well with the original Bayesian model. After extension with a tilt prior, the model could account for the static SVV data as well. When the same model was applied to the dynamic SVV data, it could partially fit the A-effects but not the E-effects. In conclusion, the model can explain certain aspects of spatial orientation. However, with regard to the dynamic SVV, refinement of the model is necessary, whether this can be achieved by adjusting the set of model parameters remains a subject for further investigation.

*Keywords:* Spatial-orientation; Bayesian estimation; Vestibular system

---

*Correspondence to:* M. van der Heijden *E-mail:* M.vanderHeijden@gmail.com

## Introduction

In normal conditions, when visual, somatosensory and vestibular cues are available, our brain is quite good at deciding how we are oriented. Although panoramic visual information plays an important role, we still have a vivid awareness of the direction of gravity when we are tilted in complete darkness. Under these conditions, we depend on the vestibular system to provide sensory signals indicating the direction of gravity and self-motion. But these vestibular cues are not without limitations, as the next section will show.

## Neurobiology of the vestibular system

The balance system, located in the inner ear on both sides of the head, consists of two small organs: the semicircular canals and the otoliths. In simple terms, the canals sense rotation whereas the otoliths detect both tilt and translation.

The semicircular canals consist of fluid-filled rings that respond to angular acceleration when the rotation axis has a component perpendicular to the plane of the canal. Figure 1a illustrates their arrangement in three approximately orthogonal planes which allows rotation detection about any axis in space. The canals contain hair cells embedded in a gelatinous structure called the cupula. As figure 1b shows, the cupula acts as a hinged gate that rotates when the fluid pushes against it. This change in cupula position causes bending of the cilia on top of the hair cells whose signal is then sent along the vestibular nerve to the brain. During rotation at a constant speed the pressure gradually decreases and restoring elastic forces pull the cupula back to its original position. Since the canal signal is proportional to cupula deflection, it shows a corresponding decay during constant velocity rotation. Although rotational acceleration is the driving force, biophysical models have shown that the canals function as mechanical integrators so that their signal effectively codes head velocity (Young, 1984).

The otoliths consist of two organs, the utricle and the saccule, which are arranged in two perpendicular planes. The former lies approximately in a horizontal plane relative to the head, whereas the latter is oriented vertically. Both are made of a gelatinous mass containing crystals on top of an

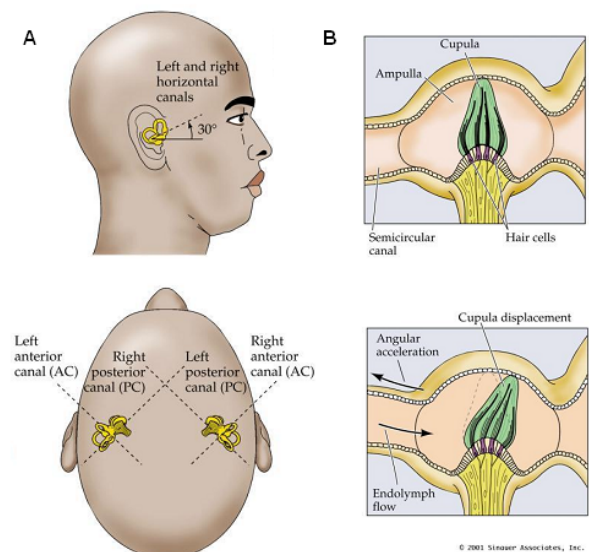


Figure 1: A: Position and orientation of the semicircular canals in the head. B: Cupula deflection is caused by the change in endolymph fluid pressure as a result of angular acceleration. The mechanical integrator characteristics of the cupula cause a signal in the vestibular nerve coding angular velocity.

array of hair cells. When subjected to linear acceleration, inertia keeps the crystals in place causing the cilia of the hair cells to bend, generating a signal in the vestibular nerve. Since gravity works on the crystals when you are tilted, the otoliths also react to tilt and therefore detect the combined effect of tilt and acceleration forces, known as the gravito-inertial force (GIF). As a result their signal is ambiguous (Young, 1984), which complicates spatial orientation considerably, because a particular otolith signal can signify either tilt or linear acceleration, or a combination. See figure 2 for an example. How the brain solves the otolith ambiguity problem is still a major question.

## Deterministic spatial-orientation models

One hypothesis states that the otolith ambiguity is solved by filtering the otolith signal. The high frequency components are interpreted as caused by translation while low frequencies are considered to be the result of tilt (Paige & Seidman, 1999; Telford

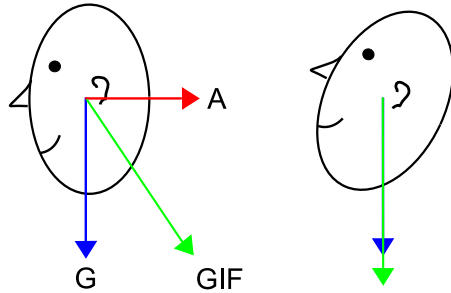


Figure 2: Otolith ambiguity is an effect of the sensitivity of the otoliths to both tilt and translation (Gravito-inertial force ( $GIF$ ) is the vector sum of gravity ( $G$ ) and acceleration ( $A$ )). For example, when a pilot is taking off from a aircraft-carrier but is still level on the runway, otolith ambiguity may cause a false percept of tilt due to the large acceleration. Trying to compensate for this illusory effect usually does not bode well for the pilot.

et al., 1997; Paige & Tomko, 1991). An alternative hypothesis is the canal-otolith interaction model (Angelaki et al., 1999; Merfeld & Zupan, 2002). This hypothesis states that information from the canals is used to interpret the otolith signal. The basic structure of this model is depicted in figure 3. How can the canal signal, which codes head rotation, be used to solve the otolith ambiguity problem? Rotation about the pitch or roll axis (see figure 4 for axis definitions) changes the orientation of the otoliths with respect to gravity. The otolith signal corresponding to the rotation detected by the canals is interpreted as caused by tilt, the unexplained residual is interpreted as translation.

Recently, the canal-otolith interaction model has been tested by Vingerhoets et al. (2006; 2007). In complete darkness, subjects were rotated in yaw at a constant velocity about an axis tilted relative to the direction of gravity. These conditions are called off-vertical axis rotation (OVAR). During the rotation subjects had to indicate at regular intervals whether a moving dot moved slower or faster than their perceived self-motion. All subjects initially perceived rotation, but gradually developed an illusory percept of being translated along a cone. This percept corresponds nicely to the predictions of the canal-otolith interaction model, because the decaying canal signal causes the otolith signal to be inter-

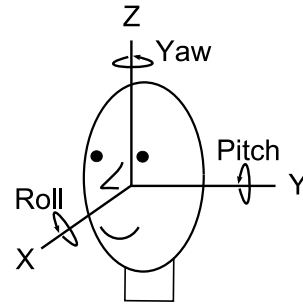


Figure 4: The axes of rotation are defined relative to the head. Yaw rotation is taken as rotation about the vertical z-axis, pitch is defined as rotation about the y-axis and roll denotes sideways rotation about the x-axis.

preted as translation. Since the otolith signal is a combination of tilt and translation, overestimation of translation should lead to an underestimation of tilt. The model thus predicts that body tilt is underestimated during illusory translation, which is indeed what Vingerhoets et al. found. Qualitatively the model and the measurement results were in agreement. However the size of the translation percept was smaller than predicted and tilt was underestimated more than would be expected from the model predictions. Vingerhoets and colleagues thus extended the model with a leaky integrator in the translation pathway and an egocentric bias mechanism in the tilt pathway. The extended model is shown in figure 5.

Vingerhoets et al. (2006, 2007) wondered whether these additional processing stages might be reinterpreted in a Bayesian framework, that formulates the underlying vestibular signal processing in probabilistic terms. This could possibly lead to a more elegant model, without changing the central principle of canal-otolith interaction.

## Bayesian models

When signal processing is described in terms of probabilities, Bayes' rule provides an optimal way to combine prior assumptions with incoming information. Because signals are often noisy, this addition of prior assumptions increases performance. For example, speed perception is influenced by the

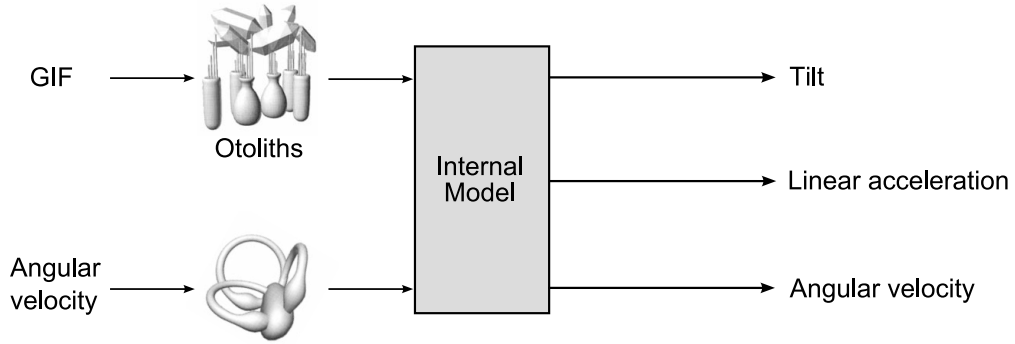


Figure 3: Schematic representation of the canal-otolith interaction hypothesis. An internal model is used to combine the canal and otolith signals in order to disambiguate the otolith signal.

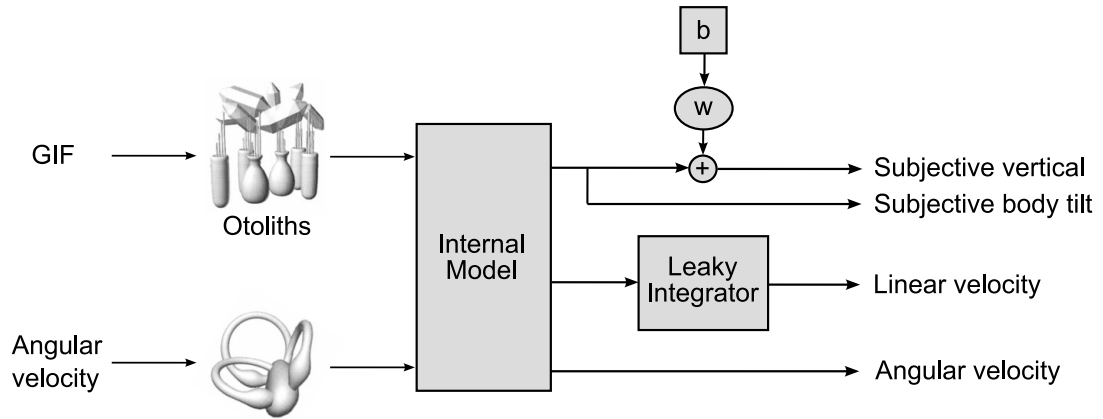


Figure 5: Extended canal-otolith interaction model. The original model has been extended with a leaky integrator in the translation pathway to account for the magnitude of the translation percept. A bias mechanism that pulls the percept of verticality towards the long body axis is added to the tilt pathway to account for tilt underestimation in verticality estimation. Body tilt is estimated in a separate pathway not influenced by the bias mechanism.

contrast of moving stimuli and low contrast stimuli are often seen to be moving slower than their actual speed. This can be explained within a Bayesian framework, as is illustrated in figure 6 (from Carandini, (2006)). The brain has an a priori assumption (the prior) that it is most likely that nothing is moving, represented by a probability distribution centred around zero. Furthermore, using knowledge of the sensory signals and their noise characteristics, the probability that a sensory signal corresponds to a given stimulus speed (the likelihood) can be measured (figure 6a). The combination of the prior and the likelihood leads to the posterior distribution (figure 6b) from which the actual speed percept is derived (figure 6c) (Carandini, 2006). One way to calculate the single point estimate is using the weighted sum over the whole distribution, which is also called the expectancy or expected value. Panel 6d shows that when contrast is low, the uncertainty over the speed is higher, which leads to a broader distribution. Because of the uncertainty the prior has more influence, causing the posterior distribution to shift towards zero (figure 6e), which explains why lower contrast stimuli are perceived to be moving slower (figure 6f). A Bayesian approach thus provides an elegant way to explain signal processing.

Interestingly, Laurens and Droulez (2007) recently reformulated the canal-otolith interaction model in terms of Bayesian probability theory, taking into account the uncertainty inherent in neural signals and providing a role for prior information on the probability of self-motion. The model estimates the self-motion percept, based on two inputs representing the canal and otolith signals. A schematic representation of the model is shown in figure 7. Here follows a brief outline of this model, a more comprehensive description will be provided in the next section. The model uses an internal model of the canal dynamics to estimate rotational velocity from the canal signal. By integrating the estimated rotational velocity, a new head orientation is calculated. The head orientation signal, in combination with the otolith signal, is then used by an internal model of the otoliths to calculate linear acceleration. These estimates are further influenced by priors on angular rotation and linear acceleration, which were chosen to be centred around zero, in line with the idea that we are mostly stationary.

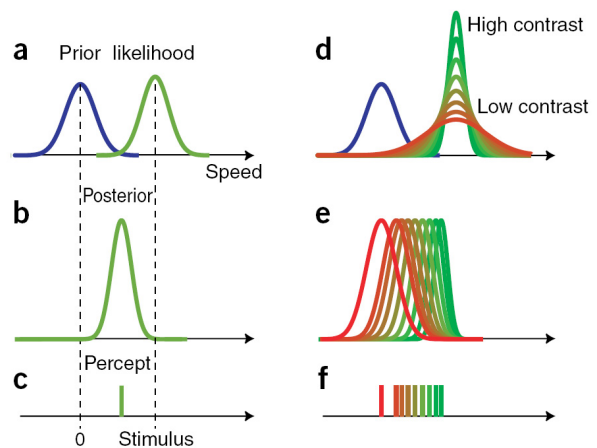


Figure 6: Illustration of a Bayesian model of speed perception. A-C: Basic principle of a Bayesian model, illustrating that the posterior distribution is the combination of the prior and the likelihood. The speed percept is the maximum a posteriori value. D-F: High contrast stimuli are easier to perceive than low contrast stimuli, thus the likelihood of the former is more peaked than the likelihood of the latter. The prior gains more influence at lower contrast, which is why low contrast stimuli are perceived to move slower. From Carandini, 2006.

## Goal of this study

Our main objective is to test whether the tilt bias mechanism proposed by Vingerhoets et al. (2007), can be reinterpreted as a prior in a Bayesian model. Accordingly, we explored whether a prior on tilt can account for the bias in visual-verticality perception. In order to test verticality perception we used a subjective visual vertical (SVV) paradigm, both under dynamic conditions of constant velocity roll rotation and under static tilt conditions. Consistent with the literature, we observed that at modest tilt angles the SVV deviates in the direction of body tilt (A-effect) (Schöne, 1964; Haes, 1970; Van Beuzekom & Van Gisbergen, 2000), while at near inverted positions the SVV is biased away from tilt direction (E-effect) (Kaptein and Van Gisbergen 2004, 2005).

Previous studies using static SVV conditions (Kaptein & Van Gisbergen, 2004; Van Beuzekom & Van Gisbergen, 2000; Mast & Jarchow, 1996) have suggested that the tilt bias mechanism is not in-

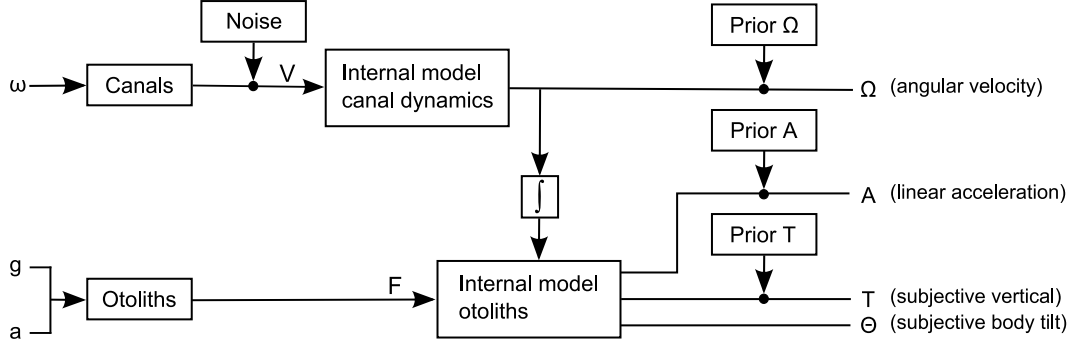


Figure 7: Schematic representation of the Bayesian canal-otolith interaction model. The noisy canal signal ( $V$ ) is processed in an internal model of canal dynamics, which leads to an estimate of angular velocity. The integrated angular velocity estimate is combined with the otolith signal ( $F$ ) in order to estimate body orientation and linear acceleration. Priors on angular velocity ( $\Omega$ ) and linear acceleration ( $A$ ) implement the initial assumption of stationarity. The prior on tilt ( $T$ ) is needed to account for the bias found in verticality perception.

involved in body tilt perception. To test whether this assumptions also holds under dynamic conditions we collected subjective body tilt estimates (SBT) during roll rotation, in separate experiments.

Here we present an extended version of the Bayesian model by Laurens and Droulez (2007), that accounts for SBT and static SVV results quite well, but has been less successful in reproducing the dynamic SVV results.

## Model description

### Model equations

Our Bayesian model, an extended version of the model by Laurens and Droulez (2007), is shown schematically in figure 7. The model has two inputs, a rotational canal input ( $\vec{V}$ ) and the gravito-inertial force signal ( $\vec{F}$ ), provided by the otoliths. Based on these inputs the model estimates the self-motion percept, which consists of rotational velocity ( $\vec{\Omega}$ ), linear acceleration ( $\vec{A}$ ) and head orientation ( $\Theta$ ).

Rotational velocity ( $\vec{\Omega}$ ) can be described as a three dimensional yaw, pitch and roll vector, where each component indicates the velocity around the

z-,y- and x-axis respectively (Figure 4). Similarly, linear acceleration ( $\vec{A}$ ) is a vector describing the acceleration in the x-,y- and z-direction. Head orientation ( $\Theta$ ) is represented by a matrix with each column a unit vector pointing along the respective axis of the coordinate system when the head is upright. This matrix also serves to transform vectors from egocentric to geocentric coordinate frames. The head orientation at a given point in time depends on the previous head orientation, updated by the current rotation vector  $\vec{\Omega}$ , entered into a standard rotation matrix  $R$ :

$$\Theta_{t+\delta t} = R(\vec{\Omega}/\delta t)\Theta_t \quad (1)$$

The canal signal is calculated using the canal dynamics found in Laurens and Droulez (2007, see also Raphan and Cohen, 2002):

$$\frac{d\vec{C}}{dt} = \frac{1}{T_c}\vec{C} - T_{can}\frac{d\vec{\Omega}}{dt} \quad (2)$$

In this equation,  $\vec{C}$  denotes cupula deflection,  $T_c$  is the cupula time constant, taken as 4s following Raphan and Cohen (2002), and  $T_{can}$  represents a matrix correcting for the non-orthogonality of the canals.

The canal signal is used to estimate rotational velocity, but because there are stochastic fluctuations in neural processes, the canal system is modelled by a noisy signal ( $\vec{V}$ ):

$$\vec{V} = \vec{C} + \sigma_v \cdot \vec{\eta} \quad (3)$$

with  $\vec{\eta}$  a random vector drawn from a Gaussian distribution with zero mean and unit variance. The estimated rotational velocity is integrated to obtain the current head orientation which is then used to estimate the direction of gravity relative to the head. As the otoliths react to gravito-inertial force (GIF), the vector sum of gravity and acceleration, linear acceleration can be calculated by subtracting the recovered gravity signal:

$$\vec{F} = \Theta^{-1} \vec{C} - \vec{A} \quad (4)$$

Following Laurens and Droulez (2007) we have chosen not to include noise on the otoliths to minimise the number of free parameters in the model. Estimating self-motion from the sensory information requires inverting equation (1) to (4).

## Bayesian inference

If all information was unambiguous so that exact values for the variables could be calculated, it would be possible for the brain to derive a veridical spatial orientation percept. However, because of the random noise interference, the exact values of the underlying variables are unknown, allowing only an approximation by statistical estimates, represented by probability distributions. Bayesian theory specifies the necessary inferences to generate these probability distributions over the model variables.

Bayesian theory is a branch of probability theory stating that a probability can be used as a degree of belief in a proposition. The general rule used in this theory is called Bayes' rule, which reads:  $P(X, Y) = P(X | Y)P(Y) = P(Y | X)P(X)$ . In the present context the probabilities of interest would look like this:

$$\overbrace{P(\text{state} | \text{input})}^{\text{posterior}} = \frac{\overbrace{P(\text{input} | \text{state})}^{\text{likelihood}} \overbrace{P(\text{state})}^{\text{prior}}}{P(\text{input})}$$

In this equation, the posterior probability is the degree of belief in a self-motion percept given some

input from the vestibular sensors. This equals the likelihood of the input given the current state times the prior probability of being in that state. Dividing by the probability of the input normalises the posterior probability.

In the model we use Bayesian inference to calculate the probability of the state variables  $\Omega$ ,  $A$  and  $\Theta$  given the sensory inputs  $V$  and  $F$ . With Bayes' rule this can be written as:

$$P(\Omega_t, A_t, \Theta_t | V_{0:t}, F_{0:t}) = \frac{1}{P(V_t, F_t)} P(V_t, F_t | \Omega_t, A_t, \Theta_t, V_{0:t-\delta t}, F_{0:t-\delta t}) \times P(\Omega_t, A_t, \Theta_t | V_{0:t-\delta t}, F_{0:t-\delta t}) \quad (5)$$

The probability of the state depends on the sensory input which in turn depends on the change between the previous and the current state. Since the current state follows from the previous state we can rewrite the probability as follows:

$$P(\Omega_t, A_t, \Theta_t | V_{0:t}, F_{0:t}) = \frac{1}{P(V_t, F_t)} P(V_t, F_t | \Omega_t, A_t, \Theta_t) \times P(\Omega_t, A_t, \Theta_t | \Omega_{t-\delta t}, A_{t-\delta t}, \Theta_{t-\delta t}) \times P(\Omega_{t-\delta t}, A_{t-\delta t}, \Theta_{t-\delta t} | V_{0:t-\delta t}, F_{0:t-\delta t}) \quad (6)$$

where the second term recursively depends on the previous state. As comparing the last term of (6) with the first will show, it is identical except for the time point.

The probability of the states that are computed depends on the prior and the likelihood of the sensory signal. Since the state estimates have to satisfy the canal dynamics (2) and the internal model of the otoliths (4), only estimates with a likelihood of 1 are computed and as a result the probability depends only on the priors. The original model has priors centred on zero rotation and acceleration. As the Results section will show, this model is too limited to describe our SVV data. We therefore added an additional prior on tilt as the Bayesian analogy of the tilt bias mechanism introduced in the canal-otolith interaction model by Vingerhoets et al. (2007). All priors are modelled by Gaussian distributions centred at zero with standard deviations  $\sigma_\Omega = 30^\circ/\text{s}$  for rotation and  $\sigma_A = 5\text{m/s}^2$  for acceleration (Laurens & Droulez, 2007). The standard deviation of the tilt prior is a free parameter

to account for the intersubject variability in our experimental data.

The noise on the canal signal has a Gaussian distribution with zero mean and standard deviation  $\sigma_v = 10^\circ/\text{s}$ . For each state estimate a random noise vector is generated at every time step to represent the uncertainty in neural processing.

## Implementation

The model progresses from the current state to the next by calculating the inverse of equations (1) through (4). The probability of the new state is derived from the recursive probability of previous states and cannot be calculated analytically due to the nonlinearity of the dynamics (the rotation matrix in equation (1)). To deal with this problem, Laurens and Droulez (2007) applied an estimation technique known as ‘particle filtering’, which is essentially a Sequential Monte Carlo method. At each discrete time point (0.1s) a number of estimates (or ‘particles’) is used to represent possible model states. These particles, which are samples from the distribution of interest, serve as an approximative substitute for the analytical solution. The reliability of the probability distribution estimate depends on the number of samples. Our simulations were run with 3000 samples, which should give a reasonable approximation. Each estimate is assigned a weight equal to the product of the probabilities of the state variables given the priors. This weight indicates how likely the estimate is under the model assumptions. In practise this weight is the product of the evaluation of Gaussian distributions at  $|W|$  and  $|A|$ . In simulations where we used a tilt prior (see Methods) an additional Gaussian for tilt angle was used. Each time step independent noise samples are generated and new weights are computed, indicating the probability of the state variable estimate. Subsequently, a resampling method draws a number of particles from the previous set with a sampling probability equal to the weight of the estimate. As a result, more likely estimates will be sampled more often than less likely ones, which prevents the set of estimates from drifting toward unlikely states. After resampling the probabilities are renormalized to yield a sum of one. The new set of particles is then used to calculate the estimates for the next time step, and so on.

All variables are initialised to zero at the start of the simulation, which conforms to the real situation that subjects are aware that they are not moving and upright at the start of the trial.

## Methods

### Subjects

Five subjects, four male and one female aged between 22 and 64 years gave written informed consent to participate in the experiments. Except for subject JG, all participants were naive with respect to experimental goals. All subjects were free of known vestibular disorders.

### Setup

A computer controlled vestibular chair was configured to rotate subjects in roll. In the chair, subjects were secured with safety belts and the head was fixated with a padded adjustable helmet. The rotation axis of the chair was aligned with the cyclopean eye. In the experiments where subjects estimated their body tilt (SBT paradigm), a LED was flashed in front of the subject to prompt a verbal body tilt report. In the verticality perception experiments (SVV paradigm), a luminous line replaced the LED. The line, which rotated in the fronto-parallel plane, served to determine the subject’s subjective visual vertical. It was computer controlled with an angular resolution of  $0.5^\circ$  and was located approximately 90cm in front of the subject.

## Experiments

We tested two aspects of spatial orientation. First, the ability to judge visual line orientations during both static and dynamic roll tilt (SVV task). Second, the ability to estimate body tilt during dynamic roll rotation (SBT task). All experiments took place in complete darkness.

### SBT paradigm

In the dynamic subjective body tilt (SBT) task subjects were rotated about their roll axis at  $30^\circ/\text{s}$  for three cycles plus  $60^\circ$  (38s for a total of  $1140^\circ$  rotation), after which they were brought back to the



upright position. Between runs subjects remained upright for 30s with the lights switched on.

During the run subjects reported their perceived body tilt on a clock scale, as if their body was the minute hand, when the LED was flashed. Trials were spaced at varying random intervals 3 to 5 seconds apart. Runs alternated between clockwise and anticlockwise. Each subject performed 40 runs in each direction, yielding a total of 80 runs. Verbal responses were written down and recorded to allow checking afterwards.

### SVV paradigms

Experiments were performed under two conditions, which we will refer to as the static and dynamic SVV paradigm. During the dynamic SVV experiment subjects were rotated 1140° clockwise or anticlockwise at 30°/s, as in the SBT experiment. At regular intervals a line was flashed for 2ms and subjects reported the orientation of the line in world centric coordinates on a clock scale. Intervals of 15° were tested, but during a single run trials were spaced at 60°. All intervals were covered by varying the first test angle through 15° multiples between 0° and 45°.

In the static SVV experiment subjects were rotated at 30°/s to a final tilt angle between 0° and 360° chosen randomly at 15° intervals. Upon reaching the tilt angle there was a 30s wait period to allow rotational signals to decay. Subsequently, subjects had to estimate the orientation of ten consecutively flashed lines of randomly varied orientations.

### Data analysis

In both paradigms subjects reported orientations on a clock scale with minute precision, which were converted to angles in degrees. For further data analysis we then defined SBT errors as the difference between the actual chair rotation angle and the tilt angle reported by the subject, as shown in figure 8a. Errors in the clockwise direction, seen from behind the subject, were taken positive. As a result, positive SBT errors reflect tilt underestimation when rotating clockwise.

Figure 8b shows that SVV errors were calculated as the difference between the orientation of the line with respect to the world and the orientation estimated by the subject. Errors in the clockwise direc-

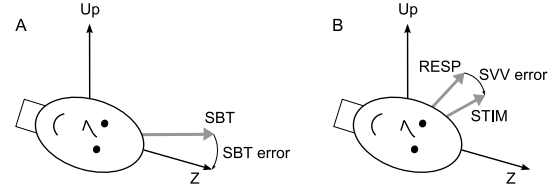


Figure 8: A: SBT errors are defined as the angle between the actual rotation angle and the body tilt estimate (SBT). B: SVV errors are defined as the angle between the real line orientation (STIM) and the reported line orientation (RESP).

tion were again taken positive. Hence, an A-effect occurring after less than half a cycle of clockwise rotation results in a positive SVV, while an E-effect would be negative.

As the results section will show, the SVV data consist of two distinct response modes. Therefore the data were split into two clusters using the k-means clustering algorithm (Matlab 7, The MathWorks). The first cluster represents responses biased toward the long body axis, as if tilt is underestimated (A-cluster). The second cluster, which denotes errors in the opposite direction (E-cluster), is found at near-inverted orientations under static conditions and in a wider range of orientations under dynamic conditions.

### Model simulations and data fitting

We used Matlab 7 (The MathWorks) to simulate the model described in Model description. Model predictions and collected verticality and tilt perception data were compared as follows. The simulations resulted in probability distributions over the state variables  $\Omega$ ,  $A$  and  $\Theta$ . From the head orientation matrix  $\Theta$  we computed a probability distribution over roll angles. This distribution was used to calculate the expectancy of all estimates. The expectancy is defined as the average over all estimates weighted with their respective probabilities.

For the comparison of model predictions and collected data we used the root mean squared error (RMSE) as a goodness of fit measure. The RMSE was calculated as the square root of the mean squared distance between the data and the corre-

sponding model prediction, hence a lower RMSE indicates a better fit.

For each subject we searched for the best fitting tilt prior standard deviation. We kept the value of the noise standard deviation at  $10^\circ/\text{s}$  and the prior width for rotation and acceleration fixed at  $30^\circ/\text{s}$  and  $5\text{m/s}^2$  respectively, as proposed by Laurens and Droulez (2007). For the additional prior on tilt we tested values of  $15^\circ$ ,  $20^\circ$ ,  $25^\circ$ ,  $30^\circ$ ,  $35^\circ$  and  $40^\circ$ . Separate model runs were performed with these six parameter values, which were then fit to the individual subject data. Furthermore we assumed that the prior on tilt was only used during the SVV but not during the SBT task. This assumption is based on data from Kaptein and Van Gisbergen (2004), Van Beuzekom and Van Gisbergen (2000) and Mast and Jarchow (1996), who found systematic errors in the SVV task but not in the SBT task. Figure 7 illustrates this hypothesis by showing two separate pathways for SBT and SVV. The tilt prior only influences the SVV pathway.

For the static SVV model fits we performed separate model runs for each final tilt angle. Due to computational limitations we restricted ourselves to final angles at  $30^\circ$  intervals. Also, separate model runs were performed to obtain the model estimates for the E-cluster data. For these E-cluster fits a prior centred on  $180^\circ$  was used, with standard deviations in the same  $15^\circ$  to  $40^\circ$  range. However, centring the Gaussian prior distribution around  $180^\circ$  is not straightforward due to rotational circularity. For example  $190^\circ$  roll rotation in one direction equals  $170^\circ$  rotation in the opposite direction, therefore  $170^\circ$  and  $190^\circ$  should be equiprobable. This was solved by using a Von Mises or circular normal distribution (Evans et al., 2000), which is a special case of a Gaussian distribution wrapped around a circle and normalised to sum to one over a  $2\pi$  domain.

## Results

In our experiments we measured the dynamic and static subjective visual vertical (SVV) and the dynamic subjective body tilt (SBT) during roll rotation. These measurements will be compared with predictions of the Bayesian model using the best-fit parameters.

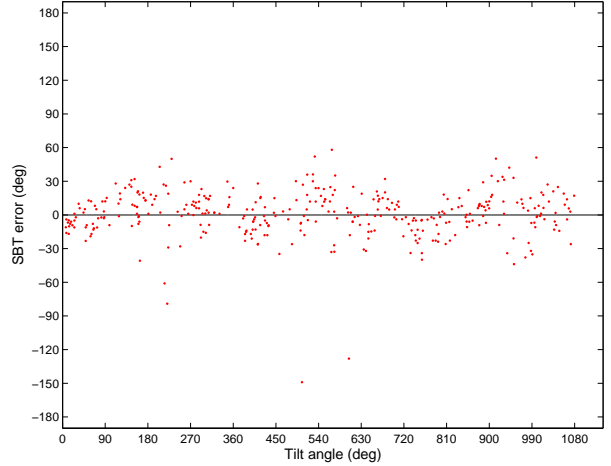


Figure 9: SBT data for a typical subject (SB). Red dots indicate the SBT error at a particular tilt angle.

### Dynamic SBT

In the dynamic SBT experiments, subjects had to verbally estimate their body tilt angle during three cycles of roll rotation. Earlier studies (Kaptein & Van Gisbergen, 2004; Mast & Jarchow, 1996; Mittelstaedt, 1983) showed subjects to be quite good at estimating body tilt under static conditions. As can be seen in figure 9, which plots the difference between the actual tilt angle and the estimated tilt angle in subject SB, we obtained a similar result under dynamic conditions. Errors in this typical subject are quite small, with no indication of systematic deviations. The figure also shows that error magnitudes are relatively constant over the three rotation cycles, indicating that there is no cumulative effect of rotating multiple cycles. We also found no systematic errors when data were pooled over cycles. These results generalise to all subjects, as shown in figure 10. The mean absolute error over all subjects is  $13.7^\circ$  with a standard deviation of  $12.2^\circ$ , corresponding to a modest  $2 \pm 2$  minutes on a clock scale. These results are consistent with earlier findings indicating that subjects are not biased when estimating their body tilt.

If we compare this to the model results using the parameters published by Laurens and Droulez (2007), we find similar results. For all subjects predicted errors are in the same range as real data. Contrary to what is found in subjects, the model predicts a gradually developing small phase delay.

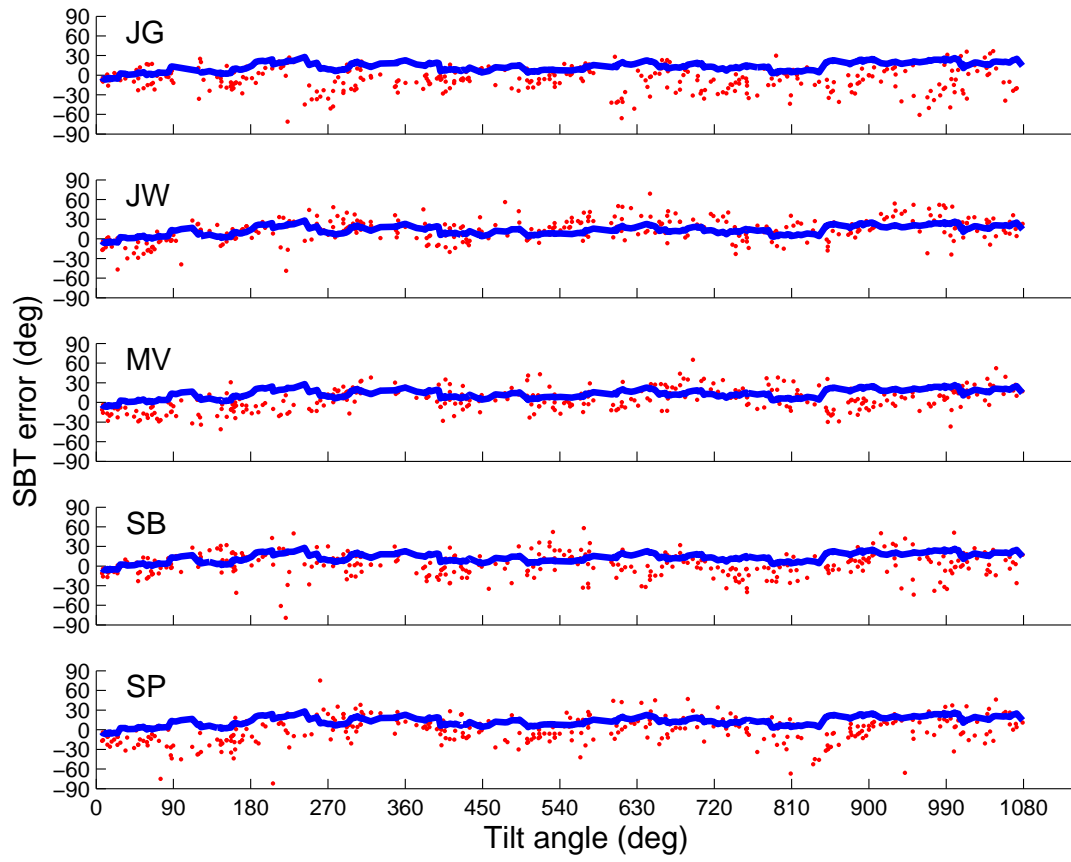


Figure 10: SBT data for all subjects (red) and model fits (blue). Subjects perform relatively well on this task. The model predicts a small phase delay which is not present in the experimental data.

This effect can be seen in figure 10, where especially for subject JG most data points (red) lie below the model curve (blue).

Because the SBT task was modelled without free parameters, the same model prediction is used for all subjects. The Root Mean Squared Error (RMSE) values from each subject can be found in table 1. All subjects have similar RMSE values, which are relatively large due to the accumulation of small random variations in the data that are not predicted by the model.

## Static SVV

Compared to the small random SBT errors the results of the static SVV experiment show a different picture with a pronounced pattern of systematic errors, as shown in figure 11. We observed an A-effect increasing with tilt angle up to  $\pm 135^\circ$  (shown in red), which is in line with the description found in literature (Kaptein and Van Gisbergen, 2004, 2005). At still larger tilt angles we found a sharp transition to the E-effect (shown in black), with errors having the opposite sign. In the  $150^\circ$  range subjects tend to overestimate tilt angle while around  $210^\circ$  they tend to underestimate tilt angle, indicating a clear response bias towards an attractor point at  $180^\circ$ . In other words, these E-cluster errors show a bias in orientation percept in the direction of the feet, whereas the A-cluster errors are biased in the direction of the head.

The two distinct clusters that emerged were separated by cluster analysis, as is described in the Methods section. The exact switch point from A to E-cluster is subject dependent. For example in figure 12, which shows all subject data as well as the model fits, it can be seen that the onset of the E-cluster for subject JG occurs at greater tilt angles than for the other subjects. The size of the systematic error is subject dependent with subject JG showing the most extreme deviations. Other subjects show considerably less pronounced systematic errors.

The blue lines in figure 12 demonstrate that the model can be tuned to fit the static SVV A-effect, while the green lines show the E-effect fit. To achieve this we used individually fitted tilt priors for both the A and E-cluster. For most subjects the A-effects can be fit with a tilt prior centred on zero and a standard deviation of  $40^\circ$ . For the subject

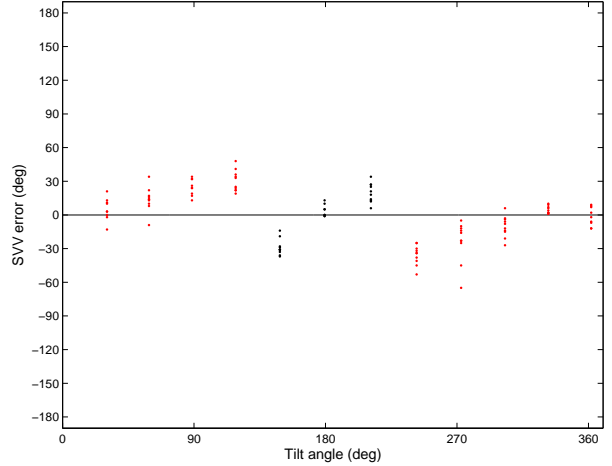


Figure 11: The static SVV data shows two distinct clusters of responses, corresponding to A-effects (red) and E-effects (black).

with the largest A-effect (JG), a prior with a standard deviation of  $35^\circ$ , indicating more bias towards zero, gives the best fit.

The errors in the E-cluster were fitted with a prior centred on  $180^\circ$  and a subject-dependent width. The best fit E-cluster prior width was  $15^\circ$  for all subjects except for MV. That fitting the E-cluster data required a narrower prior may indicate that subjects relied less on their tilt sensors than in the tilt range of the A-cluster. The prior width and RMSE fit results can be found in table 1. RMSE values for the static SVV data are smaller than those for the SBT data, as the systematic errors are fit quite accurately.

## Dynamic SVV

We also measured the SVV during three cycles of roll rotation. Results from a typical subject are shown in figure 13. Qualitatively, the dynamic results show a similar error pattern as under static conditions and again the responses can be subdivided into two clusters. There are however also clear differences. First, the errors under dynamic conditions are substantially larger. Second, the onset of the E-effect in the dynamic experiments occurs at a smaller tilt angle than what is found in the static situation. In addition responses tend to be bistable, meaning that A and E-clusters occupy

Subject	SBT	SVV static				SVV dynamic	
	RMSE	$\sigma_A$	RMSE	$\sigma_E$	RMSE	$\sigma_A$	RMSE
JG	32.9	35	14.6	15	12.8	15	20.1
JW	27.6	40	13.0	15	10.5	40	32.7
MV	27.1	40	16.8	40	16.9	15	36.9
SB	32.7	40	26.8	15	19.4	15	36.4
SP	32.3	40	10.0	15	10.3	40	38.2

Table 1: Fit parameters and goodness of fit measure (RMSE (degree)) for all subjects.  $\sigma_A$  (degree) denotes the prior width used for fitting the A-cluster,  $\sigma_E$  (degree) is the standard deviation used for fitting the E-cluster.

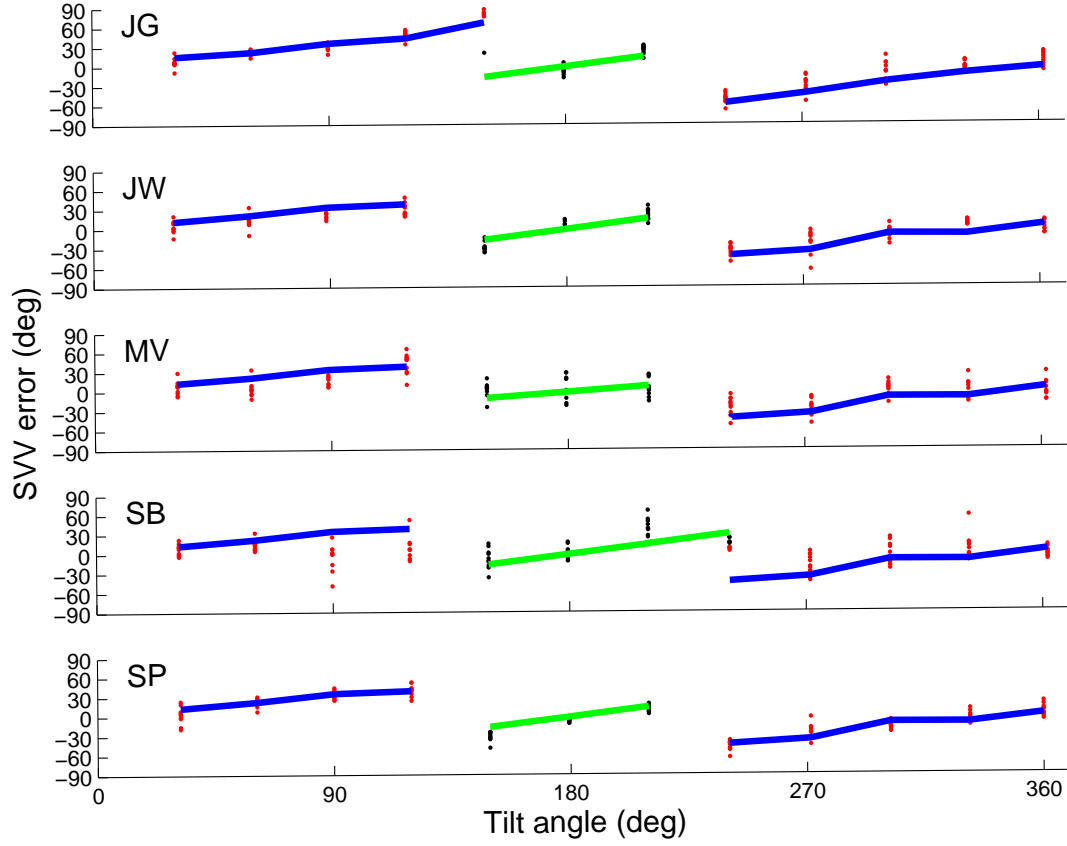


Figure 12: Static SVV data for all subjects with A-cluster data in red, E-cluster data in black, A-cluster model fits in blue and E-cluster model fits in green. Clearly the model is capable of reproducing the experimental data quite accurately.

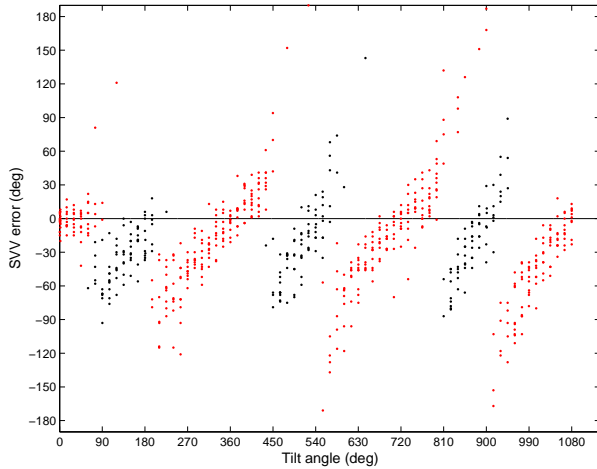


Figure 13: The dynamic SVV data is similar to the static SVV data in that it can be divided into two clusters (A-cluster and E-cluster shown in red and black respectively), but by contrast shows larger errors, an earlier onset of the E-cluster and a bistable response distribution.

overlapping tilt ranges. In separate runs subjects report tilt angles corresponding to A and E-effects when in the overlapping tilt range. The A-effect is seen in roughly the first  $90^\circ$ , at greater tilt angles both A and E-effects occur. The errors in the E-cluster again indicate an orientation percept biased in the direction of the feet. Beyond inverted positions most answers are once more in the A-cluster and errors decrease gradually as the body rotates towards the upright position. Figure 14, showing data from all subjects, demonstrates that intersubject variability is large, with subjects JG and SB making extreme errors up to  $180^\circ$ , but other subjects performing better. Additionally, the tilt range of the A-clusters varies greatly. The tilt range of the first A-cluster for example, varies between  $90^\circ$  and  $180^\circ$  for subjects MV and JG respectively.

Fitting these error patterns turned out to be more difficult than those of the static paradigm, even when we focussed on only simulating the A-cluster. As shown in figure 14 we were able to match model and data for the first  $180^\circ$  of rotation. However, at this point, the model proved incapable of reproducing the switch from positive to negative SVV errors. For static SVV this was unproblematic because all angles were tested independently, but

under dynamic conditions the transitions between clusters is an additional complication. Modelling this switch between clusters is especially problematic for subjects JW and SP who make the least extreme errors. The large errors found in subject JG, SB and to a lesser extent those in MV are fitted quite well. These subjects made large errors up to  $180^\circ$ , indicating that the orientation percept is upright when the actual position is upside down. For these subjects a narrow prior gave good fitting results (see table 1). Subjects JW and SP were fitted with a prior width of  $40^\circ$  which resulted in similar RMSE values as the other subjects, visual inspection shows however that the error patterns are not reproduced accurately. The errors in the E-cluster could not be reproduced accurately and are therefore not shown.

## Discussion

We have tested spatial orientation by means of tilt and verticality perception measurements during static and dynamic roll rotation and modelled the results in a Bayesian framework. We will summarise the experimental and model results followed by an evaluation of the model parameters and assumptions.

## Overview of experimental results

The dynamic subjective body tilt (SBT) paradigm has shown that subjects are able to estimate their body tilt during roll rotation quite accurately. These results are similar to the static SBT measurements by Kaptein and Van Gisbergen (2004). Also the static subjective visual vertical (SVV) results are in accordance with their data, showing a transition from A-effect to E-effect around  $135^\circ$  roll rotation in both directions. In the dynamic SVV paradigm, this shift in response mode starts at smaller tilt angles but is less sharply defined so that both response modes overlap and a pattern of bistable verticality percepts emerges. That bistability is found more under dynamic conditions than under static conditions may be a consequence of the fact that in the static paradigm ten consecutive lines were tested in a single run, while in the dynamic paradigm these measurements were divided over multiple runs. The relatively large intersub-

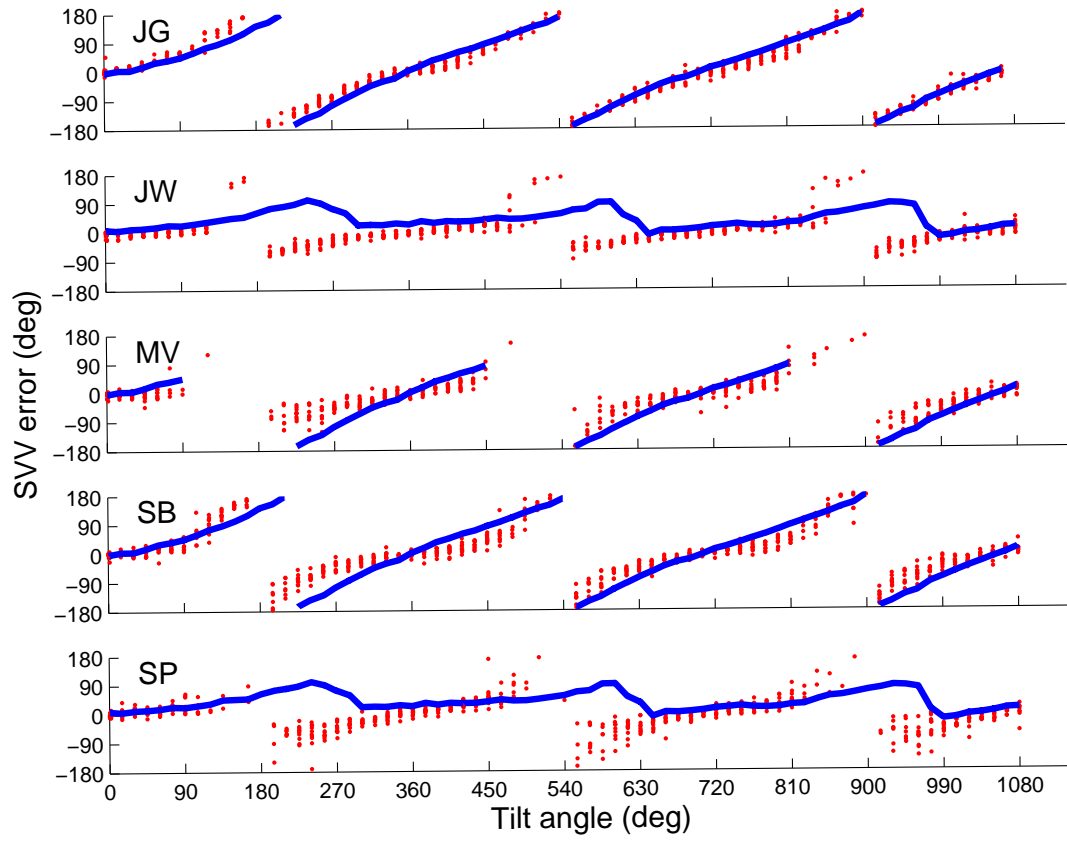


Figure 14: The dynamic SVV data (red) for all subjects shows that intersubjects variability is large. Also the model fits (blue) are less accurate than those of the static SVV data, especially in subjects JW and SP.

ject variability can perhaps be explained by differences in noise characteristics. Subjects that make larger SBT errors should then also have larger SVV errors due to a higher noise value. However, a test of the correlation between noise in the SBT task and the errors in the SVV paradigm did not reach significance (SBT-static SVV,  $r^2 = -.18, P > .7$ ; SBT-dynamic SVV,  $r^2 = .5, P > .3$ ). This may be due to the fact that only five subjects were measured and with more data a correlation can possibly be found between the noise in the body tilt estimation and the size of the systematic SVV errors.

## Modelling results

The model simulations yielded reasonable approximations of the measured percepts, except for the dynamic SVV data. As expected from reports in the literature (Kaptein & Van Gisbergen, 2004) the dynamic SBT percept can be explained without using the tilt bias. The model predictions reasonably mimic the experimental data, except for a small phase delay that was not found experimentally.

For the SVV paradigm, the tilt prior that we introduced enabled the model to predict A-effects in the static SVV task but only to a limited extent in the dynamic SVV task. Intersubject variability can be explained by adjusting the standard deviation of the tilt prior. Further differences are attributable to noise.

The E-effect in the static SVV data could be fit with a separate tilt prior, thus adding another free parameter. With our current model it appeared impossible to explain the A and E-effect with a single prior distribution. With the additional prior, the E-effects observed in the static SVV task could be explained quite well. We have performed model simulations with a bi-model prior, but this did not result in the required error patterns. This is probably due to a problem similar to what also occurred when trying to fit the dynamic SVV data, which will be described below.

Fitting the dynamic SVV data turned out to be problematic. Especially for subjects that perform relatively well on this task, the data are not matched by the model simulations. Although the errors in these subjects are small, the cluster switch results in quite large changes in orientation perception (see figure 14). Subject SP, for example, switches from approximately  $40^\circ$  errors to  $-70^\circ$  er-

rors, resulting in a  $110^\circ$  jump. The model uses the dynamics (equations (1) to (4)) to calculate the estimates and will only allow large changes in orientation when a corresponding canal signal is present. During constant velocity rotation the canal signal decays, which means that changes in orientation percept are mainly caused by noise. The large change in orientation necessary to account for the cluster switch can therefore not be reproduced. As noted in the results section, data from subjects with large dynamic SVV errors can be reproduced much better by the model. This can be explained by considering that the required change in the orientation percept when switching modes becomes smaller when errors approach  $180^\circ$ . Because of rotational circularity,  $180^\circ$  errors are equal to  $-180^\circ$  errors, hence a cluster switch resulting in errors going from  $180^\circ$  to  $-180^\circ$  does not require a jump in orientation percept.

Although RMSE values are roughly similar for all subjects in the dynamic SVV paradigm, the quality of the fit differs considerably. For further exploration of model parameter values, a different goodness of fit measure might be considered to overcome the inconclusiveness of the RMSE results.

## Model parameters

There are two classes of model parameters. First the noise characteristics and second the prior information characterisation. Guedry (1974) performed measurements of the response threshold for stimulus detection. The magnitude of the stimulus is inversely related to the time required for detection. The detection time is called response latency. These measurements form the basis for the noise parameter adopted in our model. Laurens and Droulez (2007) used a Gaussian noise distribution with a standard deviation of  $10^\circ/\text{s}$  on the canal signal to simulate response latencies and compared these to the values found by Guedry. Based on this comparison they concluded that  $10^\circ/\text{s}$  is a reasonable value for the canal noise.

We have performed some preliminary simulations with lower noise values, which are also in the range of response latencies reported by Guedry (1974) but this did not influence the results enough to warrant another free parameter. If however further investigations would demonstrate a correlation between the SBT error and the size of the SVV error, it



may be worthwhile to fit individual noise values to subjects, since it seems reasonable to assume that different subjects have different noise characteristics given the variability between subjects in our experiments.

Remarkably, in the original model by Laurens and Droulez (2007) the noise on the canals is generated using independent random samples. From a biological point of view it is rather unlikely that there is no time correlation in the noise characteristics, because changes between time points would then be completely arbitrary. Although we currently have not changed this model property, it should be relatively straightforward to implement time correlated noise generation. It remains a subject for further research whether this will make a difference for the model performance.

The other model parameters are the priors, of which the original model used two: on  $\Omega$  and  $A$  respectively. These priors correspond to the assumption of stationarity (no rotation and no linear acceleration), which enables estimating body tilt but cannot account for errors in verticality perception. By adding a prior on tilt, reflecting a prior belief that being upright is more likely than being tilted, we can simulate the errors in tilt perception. We have kept the width of the Gaussian priors on  $\Omega$  and  $A$  the same as reported by Laurens and Droulez (2007) at  $30^\circ/\text{s}$  for  $\Omega$  and  $5\text{m/s}^2$  for  $A$ . The width of the prior on tilt is the free parameter used for fitting, in the  $15\text{--}40^\circ$  range. This range is based on a common-sense estimate of what range of tilt angles is likely to be encountered, given that for a Gaussian distribution 95% of the values lie within  $2\sigma$  from the mean. Due to the computational load of performing the best fit parameter search, we limited ourselves to the six reported values for the width of the tilt prior:  $15^\circ$ ,  $20^\circ$ ,  $25^\circ$ ,  $30^\circ$ ,  $35^\circ$  and  $40^\circ$  and found that with these values we could reproduce the static SVV data and parts of the dynamic SVV data. However, exploring more values might lead to better results, especially because the best fits are often found with standard deviations at the extremes of the tested range. This remains a topic for further research.

## Bayesian model evaluation

Modelling biological systems in a Bayesian framework has been successful (e.g. Stocker and Simon-

celli, 2006) but also has some problems. Especially determining which model parameters are plausible and formulating assumptions in terms of priors is hard to justify since experimental verification is problematic. Another complication with Bayesian models is that the computation and the interpretation of probability distributions still lacks a firm neurophysiological basis.

The probability distributions emerging at the output side of the current model, for example, cannot be compared directly to experimental data. A further difficulty is that various methods can be used to obtain a single estimate from the probability distribution as a representation of the resulting percept. The value of this single estimate is relatively arbitrary because no point in the distribution has any special status in a Bayesian framework. Although the single most likely point can be calculated (e.g. maximum a posteriori), this disregards the information present in the complete distribution. As an alternative the expectancy of the distribution can be used as the single point estimate. This way all information in the distribution is taken into account proportionally to the probability. We have used the Bayesian approach of calculating the expectancy.

A more serious problem is that often the shape of a prior is unknown. As a result, computationally favourable solutions like Gaussian distributions are used, because these can be evaluated easily. Whether Gaussian distributions are a reasonable assumption from a neurophysiological viewpoint however, is questionable. Probably, better results can be obtained by explicitly estimating the model parameters. If model parameters can be found experimentally, or at least be confined to some range, the model's performance can be improved without increasing complexity. Some work has been done in this direction (Ernst & Banks, 2002; Stocker & Simoncelli, 2006), but not in the vestibular domain.

MacNeilage et al. (2007) discusses some of the difficulties in the vestibular domain and suggests the following. In order to estimate the model parameters the likelihood of the sensory inputs should be measured. He proposes a 2-alternative forced choice paradigm with stimuli that isolate either canal or otolith activity. The resulting psychometric data should be used to estimate the just noticeable difference as a measure for the sensory signal likelihood. In addition he proposes to use the

method of Stocker and Simoncelli (2006) to disentangle the influence of the likelihood from the prior. Under the assumption of higher variability in the prior than in the sensory likelihood, you can take advantage of the fact that the just noticeable difference will be affected more by the likelihood, while the point on the psychometric curve that signifies perceptual equivalence between the reference and test stimulus will be influenced more by the prior. See MacNeilage et al. (2007) for a more in depth discussion on this topic.

## Conclusion

We explored a Bayesian implementation of the canal-otolith interaction model, as extended by Vingerhoets et al. (2007). Spatial orientation performance of human subjects was tested with body tilt and verticality perception tasks. The model developed by Laurens and Droulez (2007), with our extension in the form of a prior on tilt, shows promise as a framework that can provide a new perspective on spatial orientation percepts. More work is needed to see if it is possible to find a more optimal set of model parameters that accounts better for some aspects of dynamic SVV responses.

## References

- Angelaki, D. E., McHenry, M. Q., Dickman, J. D., Newlands, S. D., & Hess, B. J. (1999). Computation of inertial motion: neural strategies to resolve ambiguous otolith information. *J Neurosci*, 19(1), 316-327.
- Carandini, M. (2006). Measuring the brain's assumptions. *Nat Neurosci*, 9(4), 468-470.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429-433.
- Evans, M., Hastings, N., & Peacock, B. (2000). von mises distribution. In *Statistical distributions* (3 ed., p. 189-191). New York: Wiley.
- Guedry, F. (1974). Psychophysics of vestibular sensation. In H. Kornhuber (Ed.), *Handbook of sensory physiology* (p. 2-154). New York: Springer-Verlag.
- Haes, H. Udo de. (1970). Stability of apparent vertical and ocular countertorsion as a function of lateral tilt. *Percept Psychophys*, 8, 137-142.
- Kaptein, R. G., & Van Gisbergen, J. A. (2004). Interpretation of a discontinuity in the sense of verticality at large body tilt. *J Neurophysiol*, 91(5), 2205-2214.
- Kaptein, R. G., & Van Gisbergen, J. A. (2005). Nature of the transition between two modes of external space perception in tilted subjects. *J Neurophysiol*, 93(6), 3356-3369.
- Laurens, J., & Droulez, J. (2007). Bayesian processing of vestibular information. *Biol Cybern*, 96(4), 389-404.
- MacNeilage, P. R., Banks, M. S., Berger, D. R., & Bulthoff, H. H. (2007). A bayesian model of the disambiguation of gravito-inertial force by visual cues. *Exp Brain Res*, 179(2), 263-290.
- Mast, F., & Jarchow, T. (1996). Perceived body position and the visual horizontal. *Brain Res Bull*, 40(5-6), 393-398.
- Merfeld, D. M., & Zupan, L. H. (2002). Neural processing of gravito-inertial cues in humans. iii. modeling tilt and translation responses. *J Neurophysiol*, 87(2), 819-833.
- Mittelstaedt, H. (1983). A new solution to the problem of the subjective vertical. *Naturwissenschaften*, 70(6), 272-281.
- Paige, G. D., & Seidman, S. H. (1999). Characteristics of the vor in response to linear acceleration. *Ann N Y Acad Sci*, 871, 123-135.
- Paige, G. D., & Tomko, D. L. (1991). Eye movement responses to linear head motion in the squirrel monkey. *J Neurophysiol*, 65(5), 1170-1196.
- Raphan, T., & Cohen, B. (2002). The vestibulo-ocular reflex in three dimensions. *Exp Brain Res*, 145(1), 1-27.
- Schöne, H. (1964). On the role of gravity in human spatial orientation. *Aerosp med*, 35, 764-772.
- Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nat Neurosci*, 9(4), 578-585.
- Telford, L., Seidman, S. H., & Paige, G. D. (1997). Dynamics of squirrel monkey linear vestibulo-ocular reflex and interactions with fixation distance. *J Neurophysiol*, 78(4), 1775-1790.
- Van Beuzekom, A. D., & Van Gisbergen, J. A. (2000). Properties of the internal representation of gravity inferred from spatial-direction

- and body-tilt estimates. *J Neurophysiol*, 84(1), 11-27.
- Vingerhoets, R. A., Medendorp, W. P., & Van Gisbergen, J. A. (2006). Time course and magnitude of illusory translation perception during off-vertical axis rotation. *J Neurophysiol*, 95(3), 1571-1587.
- Vingerhoets, R. A., Van Gisbergen, J. A., & Medendorp, W. P. (2007). Verticality perception during off-vertical axis rotation. *J Neurophysiol*, 97(5), 3256-3268.
- Young, L. (1984). Perception of the body in space: mechanisms. In B. J.M. & M. V.B. (Eds.), *Handbook of physiology - the nervous system iii* (p. 1023-1066). Bethesda, Maryland: American Physiological Society.