Optimal Estimation of Supporting-Ground Orientation for Multi-Segment Body Based on Otolith-Canal Fusion

Karim A. Tahboub

Abstract—This article discusses the problem of estimating the orientation of inclined ground on which a human subject stands based on information provided by the vestibular system consisting of the otolith and semicircular canals. It is assumed that body segments are not necessarily aligned and thus forming an open kinematic chain. The semicircular canals are analogous to technical gyrometers providing a measure of the angular velocity whereas the otolith analogues to a technical accelerometer provide a measure of the translational acceleration. Two solutions are proposed and discussed. The first is based on a stand-alone Kalman filter that optimally fuses the two measurements based on their dynamic characteristics and their noise properties. In this case, no body dynamic model is needed. In the second solution, a central extended disturbance observer that incorporates a body dynamic model (internal model) is employed. The merits of both solutions are discussed and demonstrated by experimental and simulation results.

Keywords—Kalman filter, orientation estimation, otolith-canal fusion, vestibular system.

I. INTRODUCTION

When we walk, jump or run, our vestibular system detects motion of the head in space and in turn generates reflexes that are crucial for our daily activities, such as stabilizing the visual axis (gaze) and maintaining head and body posture. In addition, the vestibular system provides us with our subjective sense of movement and orientation in space. The vestibular sensory organs are located in the petrous part of the temporal bone in close proximity to the cochlea, the auditory sensory organ. The vestibular system is comprised of two types of sensors: the two otolith organs (the saccule and utricle), which sense linear movement (translation), and the three semicircular canals, which sense rotation in three planes. The receptor cells of the otolith and semicircular canals send signals through the vestibular nerve fibers to the neural structures that control eye movements, posture, and balance.

The utricle and the saccule of the otolith organ are sensitive to linear acceleration. The sensory epithelium of these organs consists of hair cells that release transmitter and thus produce activity in vestibular nerve fibers. The cilia which emerge from the hair cells are embedded in a gelatinous matrix containing solid CaCO3 crystals (the otoconia) which overlay the cells. In response to linear acceleration, the crystals are left behind due to their inertia. The resultant bending of the cilia causes either excitation or inhibition of hair cells. Being sensitive to acceleration, the otolith organs detect the direction and magnitude of the resultant of gravity acceleration and transient linear accelerations due to head or body movement. As expected from Einstein’s equivalence principle, it is not possible to distinguish the two parts of this resultant acceleration based on the otolith organ alone.

The three semicircular canals, arranged in three orthogonal planes, are sensitive to angular (rotational) acceleration. Each canal is comprised of a circular tube containing fluid continuity, interrupted at the ampulla (that contains the sensory epithelium) by a water tight, elastic membrane called the cupula. Similar to the otoliths, the sensory cells exhibit release of neurotransmitters that is modified by the direction of cupula deflection. Although the stimulus on the semicircular canals is angular acceleration, the neural output from the sensory cells represents the velocity of rotation. This mathematical integration of the input signal is due to the mechanics of the canals; the viscous properties of the fluid are in part due to the small size of the canal (diameter of ~0.3 mm). By combining the input of the three canals, the brain can create a 3-dimensional representation of the vector of the instantaneous speed of head rotation relative to space.

The information encoded by the vestibular system becomes strongly multisensory and multimodal at the central stages of processing. The vestibular nuclei receive inputs from a wide range of spinal, cortical, cerebellar, and other brainstem structures in addition to direct inputs from the vestibular afferents. Recent studies have emphasized the importance of extra-vestibular signals in shaping the ‘simple’ sensory-motor transformations that mediate vestibulo-ocular and vestibulospinal reflexes. The multisensory and multimodal interactions that occur in vestibular processing also play an essential role in higher-order vestibular functions, like self-motion perception and spatial orientation [1].

In more technical words, the general problem to be solved can be summarized as follows. Given the information provided by the vestibular system (three dimensional linear acceleration produced by the otolith organs and three dimensional angular velocity produced by the semicircular canals including its integration mechanism) and any available other sensory relevant signals (visual, proprioceptive, and
velocity drifts represent low frequency signals. To a drift in the generated angular velocity output. This canals at the acceleration level which leads through integration like acceleration phenomenon. So, noise hits the semicircular endolymph molecules in the semicircular canals cause a noise-signals in terms of their frequency characteristics and noise quantities of interest, calls for a deep understanding of these signals in estimating the body orientation and hence other orientation calls for estimating the head absolute orientation. Thus estimating the ground orientation, see Fig. 1. The joint positions are accessible through proprioception. Thus estimating the ground orientation calls for estimating the head absolute orientation.

In the sequel, it is assumed that a human subject, modeled in the sagittal plane as a two-segment body (free ankle and hip joints), stands (moves) on an inclined surface of unknown orientation, see Fig. 1. The joint positions are accessible through proprioception. Thus estimating the ground orientation calls for estimating the head absolute orientation.

A. Dynamic Characteristics of the Vestibular System

The effective employment of canal and otolith generated signals in estimating the body orientation and hence other quantities of interest, calls for a deep understanding of these signals in terms of their frequency characteristics and noise distributions.

On the physiological level, thermal movements of the endolymph molecules in the semicircular canals cause a noise-like acceleration phenomenon. So, noise hits the semicircular canals at the acceleration level which leads through integration to a drift in the generated angular velocity output. This velocity drifts represent low frequency signals.

Fig. 1 Schematic diagram for a two-segment body on an inclined surface

On the other hand, the otolith system reacts to translational acceleration (gravitational-inertial acceleration GIA) and produces a signal proportional to translational acceleration. Gravity can be identified as the constant part of the signal whereas the inertial acceleration as the transient part. It is worth noting here that humans’ accelerations encountered normally tend to be periodic and for short time. Accordingly, low frequency component of otolith acceleration is interpreted as tilt of the head (or body) relative to gravity. Furthermore, high frequency tilts create high frequency variations in the GIA and at the same time they activate the canals. Neural activity along the otolith pathways is also subject to stochastic fluctuations, the amount of noise added to the otolith signal is small [6].

II. ESTIMATION PARADIGMS

A. Estimation based on a stand-alone Kalman filter

The method presented here is based on a method developed for and tested on a single-segment humanoid robot [9]. It first assumes that the vestibular human system can serve as a stand-alone inclinometer [10]. The need of sensory fusion is due to sensor dynamics and noise characteristics of the two components of the vestibular systems, the otolith and the semicircular canals. So the goal of the fusion is to optimally filter out noise (including drift) by utilizing the frequency complementarity of the two sensors in obtaining a broadband estimate. The method presented in [9] and followed here is based on a stand-alone Kalman filter that explicitly models the semicircular canal drift as one of the filter states. It is worth mentioning here that the estimation of spatial orientation by humans has been the subject of intensive research where different models and methodologies have been proposed. The review of MacNeilage et al. [11] summarizes the field milestones.

The stand-alone Kalman filter approach pays particular attention to noise and drift contaminating the canal and otolith measurements. The angular velocity obtained by the
semicircular canals is the combination of:

$$\omega_s = \omega + \omega_{\text{drift}} + \nu_c$$  \hspace{1cm} (1)

where $\omega$ is the actual angular velocity ($\omega = \dot{\phi} + \dot{\theta}_1 + \dot{\theta}_2$), $\nu_c$ is a measurement Gaussian white noise of variance $\delta_{\nu_c}$, and $\omega_{\text{drift}}$ is the velocity drift that is assumed to be the integration of another Gaussian white noise $\nu_{\text{drift}}$ of variance $\delta_{\nu_{\text{drift}}}$ as:

$$\dot{\omega}_{\text{drift}} = \nu_{\text{drift}}$$  \hspace{1cm} (2)

On the other hand, there are two otolith components that can be employed, these are $a_x$ and $a_y$. However, if a linear model is used, then the normal component $a_y$ does not provide useful information. Thus, only the tangential component $a_x$ is considered. The tangential acceleration measurement as measured by the otolith is:

$$a_x = a_x + \nu_a$$  \hspace{1cm} (3)

where $a_x$ is the actual translational tangential acceleration, $\nu_a$ is a measurement Gaussian white noise of variance $\delta_{\nu_a}$. After linearization, the tangential acceleration has three components: the first is due to platform translation, the second due to body angular acceleration while the third is due to the tilt of the gravity vector:

$$a_x = (L_2 + L_5) \dot{\theta}_1 + (L_2 + L_5) \dot{\theta}_2 + L_5 \dot{\theta}_1 + g(\theta_1 + \theta_2 + \theta_3)$$  \hspace{1cm} (4)

$$\approx \delta_{\omega} + g \varphi$$

where $L_2$ is the distance between the ankle and hip joints, $L_5$ is the distance between the hip joint and the vestibular organ, $h_5$ is the summation of the two distances, $\theta_i$ indicates the ground orientation, $\theta_1$ the ankle joint rotation, $\theta_2$ the hip joint rotation, $\varphi$ is the orientation of the head in space. The approximation is done by assuming that the whole body rotates about the ankle joint. To avoid the use of any kinematic parameter and since the angular acceleration cannot continue in one direction for longer periods of time (due to physical limitations), only the gravitational part is considered and the rest is handled as a high-frequency noise. Thus Eq. 3 becomes:

$$a_x = g \varphi + \nu_a$$  \hspace{1cm} (5)

Equations 1 and 5 represent the sensor dynamics of the vestibular system (in one dimension) whereas the process dynamics is given by:

$$\dot{\omega} = \nu_{\omega}$$

$$\dot{\omega}_{\text{drift}} = \nu_{\text{drift}}$$  \hspace{1cm} (6)

in which $\omega$ is considered to be solely generated by a process Gaussian white noise $\nu_{\omega}$ of variance $\delta_{\omega}$. The above equations can be collected together in state-space form as:

$$\begin{bmatrix}
\dot{\varphi} \\
\dot{\omega} \\
\dot{\omega}_{\text{drift}} \\
\dot{\nu}_c \\
\dot{\nu}_a
\end{bmatrix} =
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\varphi \\
\omega \\
\omega_{\text{drift}} \\
\nu_c \\
\nu_a
\end{bmatrix} +
\begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\begin{bmatrix}
\nu_{\omega} \\
\nu_{\text{drift}} \\
\nu_c \\
\nu_a
\end{bmatrix}$$  \hspace{1cm} (7)

The state vector $x_s = [\varphi \ \omega \ \omega_{\text{drift}}]$ can be optimally estimated by a Kalman filter based on the two measurements $y_v = [\omega_x \ a_x]$ and given the process ($w$) and measurements ($v$) noise variances. In other words, the two measurements can be fused optimally to find the best possible estimates in the presence of process and measurement noise. The Kalman filter is realized by:

$$\begin{bmatrix}
\dot{\varphi} \\
\dot{\omega} \\
\dot{\omega}_{\text{drift}} \\
\dot{\nu}_c \\
\dot{\nu}_a
\end{bmatrix} =
\begin{bmatrix}
0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\varphi \\
\omega \\
\omega_{\text{drift}} \\
\nu_c \\
\nu_a
\end{bmatrix} +
\begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
\begin{bmatrix}
\nu_{\omega} \\
\nu_{\text{drift}} \\
\nu_c \\
\nu_a
\end{bmatrix}$$  \hspace{1cm} (8)

where $\hat{x}_s$ is the optimal estimate of $x_s$, and $L_{KF}$ is the static Kalman gain matrix that is obtained by solving the algebraic Riccati equation:

$$F_s \cdot P + P \cdot F_s^T + Q - P \cdot H_v \cdot R^{-1} \cdot H_v^T \cdot P = 0$$  \hspace{1cm} (10)

and

$$L_{KF} = P \cdot H_v \cdot R^{-1}$$  \hspace{1cm} (11)

with $R$ being the covariance matrix of measurement noise and $Q$ being the covariance matrix of process noise. Both matrices are assumed to be diagonal with the variances of different noise signals placed on the diagonal:

$$Q =
\begin{bmatrix}
0 & 0 & 0 \\
0 & \delta_w & 0 \\
0 & 0 & \delta_{\text{drift}}
\end{bmatrix}$$  \hspace{1cm} (12)
Based on the estimates $\hat{\phi}$ and $\hat{\theta}$, the platform tilt variables are estimated by employing the proprioceptive measurement of the joint angles and velocities as

$$\hat{\theta}_1 = \hat{\phi} - \dot{\theta}_2 - \dot{\theta}_1$$

and

$$\hat{\theta}_2 = \hat{\phi} - \dot{\theta}_2 - \dot{\theta}_2$$

### B. Estimation based on an extended disturbance observer

The estimation method just presented, which is motivated by the assumption that the vestibular system fusion is a stand-alone process, does not make use of the kinetic model to its full extent. It would be feasible to exploit the internal kinematic and kinetic models through an extended disturbance observer that incorporates all available measurements. This idea is already adopted by previous studies. For example, Oman [13] and Glasauer and Merfeld [14] employed the formalism of observer theory where they assumed that the brain uses an internal representation of the motion variables and of the geometrical and physical relationships that link them. The novel idea in this article is the employment of the extended disturbance observer to estimate the external disturbance (ground inclination).

The linearized body dynamics model is abstracted as [9]:

$$\dot{x} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{M^{-1}K}{a} & 0 & 0 \\ \frac{M^{-1}}{b} & \frac{M^{-1}B}{c} & \frac{F_d}{c} \end{bmatrix} x + \begin{bmatrix} 0 \\ \frac{1}{a} \\ \frac{1}{b} \end{bmatrix} u$$

where $F_d = \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix}$ is the external disturbance vector arising due to ground motion, $B_\theta$ and $N$ are the corresponding disturbance effect and state disturbance matrices. The extended disturbance observer is based on the assumption that the external unknown disturbance $\dot{\theta}$ can be considered as a state by assuming that it can be approximated by a piece-wise constant function [12]. In other words:

$$\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$$

leading to the extended state-space representation:

$$\dot{x} = \begin{bmatrix} A & N \\ 0 & A_\theta \end{bmatrix} x + \begin{bmatrix} B \\ 0 \end{bmatrix} u$$

with $y_c = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} \omega \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}$ as measurements. The extended-state vector $x_c$ can be estimated given the measurements, if and only if the pair $(A_\theta, C)$ is observable which is straightforward to prove in this case. The assumption that the external disturbance is piece-wise constant calls for a fast observer. The extended disturbance observer is implemented by:

$$\dot{x}_c = A_\theta x_c + L_c(y_c - C_\theta x_c)$$

where the observer gain matrix $L_c$ can be found by pole placement for example.

### III. EXPERIMENTAL AND SIMULATION RESULTS

To test the validity of the proposed methods, a special-purpose postural humanoid robot “PostuRob” which was built at the Department of Neurology – University of Freiburg is used for the purposes of this study, see Fig. 2. PostuRob is composed of two main segments: the body including the trunk and legs resembling a single inverted pendulum and the feet that rest freely on an in-house-built motion platform. It is capable of combining translational and rotational motions. PostuRob itself can produce voluntary lean motions in the presence or absence of platform motions. It is equipped with a gyrometer and an accelerometer (Types ADXRS401 and ADXL203, respectively, Analog Devices, Norwood, USA) placed at a height of about 0.325m from the ankle joint (center of rotation).

#### A. Stand-alone Kalman filter

**A.1 Experimental results of PostuRob**

To demonstrate the validity and performance of the proposed method, a set of experiments has been performed. These include a pure platform rotation, pure robot rotation, and a combination of both.

Figures 3 and 4 show the accelerometer and gyrometer noisy signals when the robot and platform were completely at rest. Although one can identify some sensor dynamics in both signals, it is assumed that the shown signals are pure Gaussian noises. Observing the measured signals when the robot and platform were at rest and offline experimenting with the Kalman filter yielded the following variances:

$$Q = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 10^{-3} & 0 \\ 0 & 0 & 10^{-6} \end{bmatrix}$$
Fig. 2 Photograph of humanoid robot ‘PostuRob’ standing on motion platform.

and

\[
R = \begin{bmatrix} 10^{-4} & 0 \\ 0 & 10 \end{bmatrix}
\]  

(21)

which lead to the filter gains:

\[
L_{KF} = \begin{bmatrix} 9.9 \times 10^{-1} & 8.0 \times 10^{-3} \\ 9.9 \times 10^{-1} & 3.2 \times 10^{-4} \\ 1.0 \times 10^{-3} & -3.1 \times 10^{-4} \end{bmatrix}
\]  

(22)

These gains indicate that the Kalman filter relies more on the gyrometer measurement in comparison to the accelerometer due to the explicit approximation of Eq. 24 that considers the accelerometer as a noisy inclinometer.

First, the robot is commanded to follow a step function of 5 degrees on a stationary platform. Figure 5 shows the actual and estimated angular displacement of the robot. The estimation based on the presented fusion algorithm demonstrates a fast response to the sudden changes as well as a good steady-state characteristics.

Second, the robot is commanded to follow a sinusoidal trajectory of 5 degrees at a frequency of 0.2 Hz. Figure 6 shows the actual and estimated angular displacement corresponding to that motion. Once again, the estimation exhibits accuracy and good dynamic characteristics. Here, the actual angular displacement is obtained by an optical encoder attached to the rotation joint.

Fig. 3 Accelerometer measurement noise as measured when the robot was at rest.

Fig. 4 Gyrometer measurement noise as measured when the robot was at rest.

The frequency response of estimation is tested. A standard pseudorandom ternary sequence (PRTS) of numbers [13] is used for commanding the robot to move on a stationary platform. Briefly, the stimulus was created from a 242-length PRTS sequence by assigning a rotational velocity waveform a fixed value of +v, 0, or –v [°/s] according to the PRTS sequence for a duration of \( \Delta t = 0.25 \) s. The duration of each stimulus cycle was 60.5 s. The mathematical integration of
this PRTS velocity waveform gave a position waveform with which the platform was commanded to lean forward or backward. The PRTS stimulus has a spectral bandwidth (0.02 – 2 Hz) with the velocity waveform having spectral and statistical properties approximating a white noise stimulus. The transfer function between actual and estimated orientation together with the coherence function are calculated by correlating the two. Figures 7 and 8 show the experimentally identified magnitude and phase frequency response curves of the estimation transfer function.

**Fig. 5** Measured and estimated body angular displacement.

**Fig. 6** Measured and estimated body angular displacement.

**Fig. 7** Magnitude frequency response of Kalman filter body-orientation estimation.

**Fig. 8** Phase frequency response of Kalman filter body-orientation estimation.

**A. 2 Simulation results for multi-segment humanoid**

The Kalman filter detailed above is employed to estimate the platform orientation for a multi-segment humanoid, PostuRob II which is built with an extra hip joint to form a three-segment humanoid. Figures 9 and 10 show the estimate of the platform tilt variables. Platform tilt and its angular velocity are obtained directly from the Kalman filter based on the gyrometer and accelerometer measurements after subtracting ankle and hip angles and velocities according to Eqs. 14-15.
B. Extended disturbance observer

The other option of estimating the disturbance through an extended disturbance observer is, as well, experimented. Here, a linearized model of process kinematics and dynamics as well as that of sensory measurements is imbedded as an internal model in the extended disturbance observer. Based on Eqs. 17-18 and given

\[
\begin{bmatrix}
\theta_2 & \theta_3 & \dot{\theta}_2 & \dot{\theta}_3 & \omega_x & \omega_y & \omega_z
\end{bmatrix}^T
\]

as measurements and by assigning the extended disturbance observer poles, the observer gain matrix is found to be:

\[
L_e = \begin{bmatrix}
1.0 \times 10^4 & -2.0 \times 10^1 & -9.7 \times 10^9 & ... \\
-2.0 \times 10^1 & 1.2 \times 10^1 & 5.1 \times 10^0 & ... \\
-2.0 \times 10^0 & -8.0 \times 10^0 & 1.5 \times 10^2 & ... \\
-6.8 \times 10^1 & 2.8 \times 10^1 & -1.3 \times 10^9 & ... \\
-8.2 \times 10^0 & -2.5 \times 10^0 & 4.3 \times 10^0 & ... \\
9.7 \times 10^0 & -1.4 \times 10^0 & -1.7 \times 10^2 & ... \\
8.5 \times 10^1 & -1.19 \times 10^1 & -1.3 \times 10^1 & ... \\
-9.4 \times 10^0 & 8.7 \times 10^0 & 1.8 \times 10^5 & ... \\
2.1 \times 10^0 & -1.2 \times 10^0 & -6.4 \times 10^7 & ... \\
6.6 \times 10^0 & -7.3 \times 10^0 & 5.5 \times 10^4 & ... \\
1.8 \times 10^2 & 7.3 \times 10^1 & -5.8 \times 10^4 & ... \\
5.4 \times 10^1 & -1.0 \times 10^2 & 1.0 \times 10^4 & ... \\
-1.6 \times 10^2 & 1.6 \times 10^2 & 2.7 \times 10^4 & ... \\
-1.2 \times 10^3 & 1.2 \times 10^3 & 2.4 \times 10^3 & ...
\end{bmatrix}
\]  
\tag{24}

The relative high gains corresponding to three angular velocity measurements (columns 3, 4, and 5) indicate that this extended disturbance observer relies primarily on these measurements in estimating the external disturbances. Figures 11-13 show the estimates of the ground inclination variables (inclination, angular velocity, and angular acceleration) as directly obtained by the extended disturbance observer. Here, all estimates are smooth and accurate to a satisfactory limit.

IV. CONCLUSION

To stabilize posture against non-stationary supporting grounds, humans need to have an estimate of the inclination and possibly the motion of the ground. For this, proprioception, vision, and vestibular systems contribute in an optimal estimation setup that take into consideration the
In this article, optimal fusion paradigms of the otolith and semicircular signals are presented. The first paradigm is based on the assumption that the vestibular system can be considered as a stand-alone inclinometer that does necessarily require further information. Here, the two signal based on their noise characteristics (low-frequency semicircular angular velocity noise versus high frequency otolith acceleration noise) are optimally fused through a Kalman filter. Experimental and simulation demonstrate the validity of the proposed algorithm. On the other hand, the second paradigm is based on the understanding that the estimation of the absolute head orientation in space together with external disturbances can be achieved by centrally incorporating all available information (internal model). In this context, it is possible to distinguish between self-produced motion and motions arising due to external disturbances (including ground motion). Thus, it is anticipated that this estimation option provides better results especially when self motion and external disturbances are superimposed. Simulation results of a multi-segment body demonstrate this anticipated result.

The two fusion algorithms, which are derived from by inspecting the biological system, are employed for a humanoid equipped by an accelerometer and a gyrometer. Simulation and experimental results show the suitability of the two biologically-inspired algorithms for technical systems.

REFERENCES