BAYESIAN RECURSIVE ESTIMATION ON THE ROTATION GROUP

Sofia Suvorova[1], *Stephen Howard*[2], *Bill Moran*[1]

[1] Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia
[2] Defence Science and Technology Organization, Edinburgh, South Australia 5111
Email:{ssuv, wmoran}@unimelb.edu.au, Stephen.Howard@dsto.defence.gov.au

ABSTRACT

Tracking of the orientation of a rigid body based on directional measurements is a key issue in many applications. Configurations in this sense are precisely representable as elements of the rotation group $\mathcal{SO}(3)$, and the issue devolves to one of tracking on this group, for which and algorithm is described here. Its novelty derives from the use of maximum entropy distributions on these groups as models for the priors, and from the approximation algorithms that permit numerical implementation of such a model. These solutions can be written in a recursive form. While the general ideas apply in all dimensions, the focus of this paper is on the important 3-dimensional case. It is impossible to compute the exact solution; instead, obtained here is a highly effective approximation. It is shown that, in contrast with other approaches, the algorithm described here produces outputs which are both very accurate and statistically meaningful.

Keywords: Recursive Estimation, Orientation, Rotation Group, Von Mises Fisher Distribution

1. INTRODUCTION

A large number of applications require tracking of orientation of rigid bodies or linked systems of rigid bodies. These arise, in particular, in tracking of satellites and space junk, estimation of robot configurations, autonomous UAVs, and in tracking of human and animal body motions. Models of this kind arise wherever at least part of the dynamics of the system under observation can be characterized as preserving distance from a fixed centre. A mathematically correct model for such problems represents dynamical changes in the system in terms of elements of the group $\mathcal{SO}(3)$ of rotations on $\mathbb{R}^3$; that is, matrices R satisfying $RR^T = I$. Uncertainty about these dynamical changes is expressed in terms of probability distributions (or more specifically, densities) on that group. The current state of the system is described by such a density. Tracking in this context requires a method for updating this representation, taking account of ongoing measurements. Choice of a family of distributions to represent the uncertainty is always problematic once one departs from the linear Gaussian model on $\mathbb{R}^3$, where the Kalman filter, using Gaussian noise models, is demonstrably optimal for natural measures of performance.

Despite its obvious importance, the literature on this topic, *per se*, is sparse. Most of the known approaches to the problem rely on linear approximations to the state space $\mathcal{SO}(3)$ and the dynamical model. These methods are discussed in [7,13]. Other approaches use linear estimation methods to track rotation given by rotation matrices or quaternions [11, 12]. For this formulation, the additive Gaussian noise assumption of the Kalman filter is not appropriate simply because the addition of Gaussian noise to the rotation matrix will not in general result in another rotation matrix. These methods need to use some ways of "projecting" the results back into the correct space; namely, the group $\mathcal{SO}(3)$. In practice, as a result of the incorrect choice of the noise model, these algorithms require very frequent re-initialization of the tracker.

Perhaps the closest approach to ours is *the complementary tracker* as described in [8]. This overcomes the problems enunciated above to some degree, but does not provide a mechanism for computing the error in the estimator. In effect, it is a constant gain filter. As a result, only *ad hoc* internal methods for correction of the tracker are available.

Maximal entropy distributions on the rotation group have densities with respect to Haar measure (the unique probability measure invariant under left and right translation in the group) of the form

$$\Pr(R; A) = \alpha(A)e^{\mathrm{Tr}(AR)}, \quad R \in \mathcal{SO}(3) \tag{1}$$

where the matrix $A \in \mathbb{R}^{3\times3}$ of rank ≥ 2 encompasses both a *rotational mean* and a measure of spread, known as *concentration*. That these are indeed the maximal entropy distributions subject to a constraint of the form

$$Z = \int_{\mathcal{SO}(3)} R \Pr(R; A)\,[dR], \tag{2}$$

for a specified matrix Z, follows by a straightforward Lagrangian extremal argument.

The *polar decomposition* of the matrix A in (1) as $\hat{R}^T A_0$ provides a version of the *mean* of the distribution $\Pr(R; A)$, $\hat{R} \in \mathcal{SO}(3)$, and an elliptical component, A_0 of the polar decomposition, regarded as the *concentration matrix* [2]. The matrix polar decomposition always exists and is unique. The *partition function* $\alpha(A)$ is a normalizing factor, so that

$$\int_{\mathcal{SO}(3)} \Pr(R; A)\,[dR] = 1, \tag{3}$$

where $[dR]$ denotes integration with respect to the Haar measure on the rotation group. The term $\alpha(A)$ can be shown to be the reciprocal of a *hypergeometric function* formed using the matrix of eigenvalues of the concentration matrix A, denoted by Σ. The hypergeometric function in question is, in fact, ${}_0F_1(\frac{3}{2}; \frac{1}{4}\Sigma^2)$, and this will be important for later calculations. The reader is referred to [4] for a deep discussion of some of these concepts. Indeed, as *von Mises-Fisher* distributions, maximal entropy distributions (1) have been widely studied [4–6, 9].

The collection of von Mises-Fisher matrix distributions has the pleasing property of being closed under point-wise products, as can be easily verified. However, dynamical updates of a given von Mises-Fisher distribution, based on, say, a (noisy) constant angular velocity model, do not retain the von Mises-Fisher character. This property, characteristic of the linear Gaussian model, fails for such distributions, as it does for many linear tracking schemes that do not strictly accord with the conditions for the Kalman filter. As in the non-Kalman linear case, when this happens, an approximation is needed to project back into this class, so as to produce a recursive filter. We describe such a technique in this paper. One desirable characteristic is that the spread of the approximate posterior should be worse (that is, is more uncertain) than is the case for the distribution without approximation. This is an important characteristic, since, especially in contrast with the extended Kalman filtering, this filter effectively assumes no more than is given by the underlying model and the observations. We have not been able to establish this for the three dimensional case treated here, but in another publication will demonstrate it for the two dimensional case. Nonetheless, Monte Carlo simulations suggest that the result is close to true in three dimensions.

2. TRACKING ON $\mathcal{SO}(3)$

In three dimensions, the orientation of an object is represented by an element of the special orthogonal group $\mathcal{SO}(3)$ corresponding to the rotation of the object from a specified but arbitrary initial position. The matrix von Mises-Fisher Distributions, defined in (1), are the maximal entropy probability distributions $\Pr(R; A)$ over rotation matrices. We note that the natural volume of the manifold underlying $\mathcal{SO}(3)$ is $16\pi^2$, and write the Haar measure of $\mathcal{SO}(3)$ in terms of exponentials of elements of the tangent space. This uses Rodrigues' rotation formula [14], which relates the rotation matrix and an angular velocity vector $\mathbf{w} = (w_1, w_2, w_3)^T$, that can be regarded as a member of the tangent space via

$$\mathbf{w} \longleftrightarrow \begin{pmatrix} 0 & w_1 & -w_2 \\ -w_1 & 0 & w_3 \\ w_2 & -w_3 & 0 \end{pmatrix} = [\mathbf{w}]. \tag{4}$$

Rodrigues' formula gives for $R = e^{[\mathbf{w}]}$

$$R = I + \sin(\theta)[\epsilon] + (1 - \cos(\theta))[\epsilon]^2 \tag{5}$$

where $\theta = \sqrt{\mathbf{w}^T\mathbf{w}}$ is an angle of rotation around an axis $\epsilon = \frac{\mathbf{w}}{\theta}$. One can verify that

$$[dR] = \frac{1 - \cos\theta}{8\pi^2\theta^2} \prod_{i=1}^{3} dw_i, \qquad \text{for } |\theta| \leq \pi, \tag{6}$$

2.1. Mean and Concentration Matrices on $\mathcal{SO}(3)$

To derive the tracking algorithm it will be necessary to compute integrals of the form (2), where $\Pr(R; A)$ is as in Eq. (1). An important property of Z is given by the following simple Lemma 1.

Lemma 1 *Z is diagonal if the matrix A is diagonal.*

The proof is a straightforward consequence of the fact that a matrix is diagonal if and only if it commutes with all diagonal matrices, and is omitted.

If the matrix A is non-diagonal, it is decomposable as a product $A = \hat{R}^T V \Sigma V^T$, where $\hat{R}$ and $V\Sigma V^T$ are the polar and elliptic components, respectively, of A. Here Σ is the diagonal matrix of eigenvalues [6] $\kappa_1, \kappa_2, \kappa_3$ of A. This decomposition keeps the $\hat{R}$ as a rotation and so it is possible that Σ is not positive semi-definite to compensate. If A is singular this decomposition may not be unique. The first moment is computed as

$$Z = V f(\Sigma) V^T \hat{R}, \tag{7}$$

where $f(\Sigma)$ is a diagonal matrix with the elements on the diagonal given by [6],

$$f_i = \frac{\partial}{\partial \kappa_i} \log {}_0F_1(\frac{3}{2}; \frac{1}{4}\Sigma^2) \text{ for } i = 1, 2, 3. \tag{8}$$

These elements (f_i) are computed in the tangent space of $\mathcal{SO}(3)$ at the identity; that is, in the Lie algebra [10] of $\mathcal{SO}(3)$, which consists of the skew-symmetric matrices. By Equation (6),

$${}_0F_1(\frac{3}{2}; \frac{1}{4}\Sigma^2) = \frac{1}{8\pi^2} \int_{|\theta| \leq \pi} \frac{1 - \cos\theta}{\theta^2} e^{\mathrm{Tr}(\Sigma \mathbf{R})} d\mathbf{w}. \tag{9}$$

Using Rodrigues' formula and observing that $\mathrm{Tr}(\Sigma[\mathbf{w}]) = 0$, we obtain

$$\begin{aligned} {}_0F_1(\frac{3}{2}; \frac{1}{4}\Sigma^2) &= \frac{1}{8\pi^2} e^{\mathrm{Tr}(\Sigma)} \\ &\int_{|\theta| \leq \pi} \frac{1 - \cos\theta}{\theta^2} e^{-\frac{1}{2}\mathbf{w}^T \Sigma' \mathbf{w} \frac{1-\cos\theta}{\theta^2}} d\mathbf{w}. \end{aligned} \tag{10}$$

where Σ' as a diagonal matrix

$$\Sigma' = \begin{pmatrix} \kappa_2 + \kappa_3 & 0 & 0 \\ 0 & \kappa_1 + \kappa_3 & 0 \\ 0 & 0 & \kappa_1 + \kappa_2 \end{pmatrix}. \tag{11}$$

A very simple Taylor approximation gives $\frac{1-\cos\theta}{\theta^2} \approx \frac{1}{2}$; also we assume that for large values on the diagonal of the matrix Σ' the spread of the Gaussian function $e^{-\frac{1}{2}\mathbf{w}^T\Sigma'\mathbf{w}}$ is essentially contained in the ball of radius π. These, taken together, allow us to compute the following (again the details are omitted):

$$\begin{aligned} f(\Sigma) &\approx I - \\ &\frac{1}{2}\begin{pmatrix} \frac{1}{\kappa_1+\kappa_3} + \frac{1}{\kappa_1+\kappa_2} & 0 & 0 \\ 0 & \frac{1}{\kappa_2+\kappa_3} + \frac{1}{\kappa_1+\kappa_2} & 0 \\ 0 & 0 & \frac{1}{\kappa_1+\kappa_3} + \frac{1}{\kappa_2+\kappa_3} \end{pmatrix}. \end{aligned} \tag{12}$$

As stated earlier, it would be worthwhile to show that the approximate distribution has a wider "spread" than the actual distribution. Unfortunately, it is not apparent from the formulae whether the actual $f(\Sigma)$ is greater or smaller than its approximation in Eq. (12). On the other hand, we have evaluated the integral in Eq. (2) using a Monte Carlo method and orthogonal matrices uniformly distributed on $SO(3)$, using the technique described in [3], for a large number of concentration matrices and compared the result with the approximate solution. It appears from these numerical experiments that the approximate matrix is not always larger then the actual matrix, but the relative error does not exceed 1% if any one of the eigenvalues $\kappa_i \geq 10$. This phenomenon deserves more attention and we intend to continue to investigate it.

2.2. Bayesian Prediction

In our approach, the posterior distribution to time $k-1$ of the rotation matrix R_{k-1} is taken to be a matrix von Mises-Fisher distribution with mean $\hat{R}_{k-1}$ and concentration matrix A_{k-1}, i.e

$$\Pr(R_{k-1}|Z_{k-1}, \ldots, Z_1) = \alpha(A_{k-1}) e^{\mathrm{Tr}(\hat{R}_{k-1}^T A_{k-1} R_{k-1})}, \tag{13}$$

where $Z_{k-1}, \ldots, Z_1$ are the measurements collected at times between 1 and $k-1$. Assume that at time k the orientation changes by the application of another random rotation matrix P_k distributed according to another matrix von Mises-Fisher distribution. Thus the predicted (prior) orientation is just a product of two random rotations $R_k = P_k R_{k-1}$. The state dynamics transition probability is written in terms of distribution of matrix P_k

$$\Pr(R_k|R_{k-1}) = \Pr(P_k; B_k) = \alpha(B_k) e^{\mathrm{Tr}(\hat{P_k}^T B_k R_k R_{k-1}^T)}, \quad (14)$$

where $\hat{P_k}^T$ is a mean and $B_k = U_B \Sigma_B U_B^T$ is a concentration matrices of the dynamics. The prior distribution of R_k is then

$$\Pr(R_k|Z_{k-1}, \ldots, Z_1) = \alpha(B_k)\alpha(A_{k-1}) \int_{\mathcal{SO}(3)} e^{\mathrm{Tr}(\hat{R}_{k-1}^T A_{k-1} R)} e^{\mathrm{Tr}(\hat{P_k}^T B_k R_k R^T)} [dR]. \quad (15)$$

All we need now is to approximate this distribution with a matrix von Mises-Fisher distribution, so that a recursive process can be used to estimate and track the orientation over time. By the method described in the previous section, the first moment of R_k is

$$Z = U_B f(\Sigma_B) U_B^T \hat{P}_k U_A f(\Sigma_A) U_A^T \hat{R}_{k-1}, \quad (16)$$

where $f(\Sigma_B)$ and $f(\Sigma_A)$ are defined in Eq. (8). The matrix Z in Eq. (16) can be polar decomposed and written as

$$Z = U_C f(\Sigma_C) U_C^T \hat{R}_k. \quad (17)$$

Hence, the distribution of R_k is approximated by the following von Mises-Fisher distribution

$$\Pr(R_k|Z_{k-1}, \ldots, Z_1) = \alpha(C_k) e^{\mathrm{Tr}(\hat{Q}_k C_k R_k)}, \quad (18)$$

where $\hat{Q}_k$ is a mean, $C_k = U_C \Sigma_C U_C^T$ is concentration matrix and Σ_C is computed as an inverse of $f(\Sigma_C)$. It is interesting to notice, that since matrices in Eq. (16) do not in general commute, the mean of the product of two random (with von Mises-Fisher distribution) rotations is not a product of the means of the two rotations.

2.3. Bayesian update

The tracking is performed in some inertial frame identified by several inertial objects. The measurements are given by the directional measurements to these objects. Clearly, at least two separate objects are needed. For example, in an inertial measurement unit these are given by the earth gravity and earth magnetic center. Other possibilities could for example be given by directional measurements to relatively distant objects, such as buildings, geographic objects (mountains), stars, etc. The measurements are unit vectors stacked as columns of matrix Z_k. The likelihood function of these measurements Z_k is

$$\mathrm{L}(Z_k|R_k) = \alpha(XMZ_k^T) e^{\mathrm{Tr}(XMZ_k^T R_k)}, \quad (19)$$

where columns of the matrix X are *reference* vectors in a global coordinate system, and M is a diagonal matrix with measurement concentration parameters on the diagonal (see [1]). Bayes' rule enables computation of the posterior distribution from the measurement likelihood and the prior distribution $\Pr(R_k|Z_{k-1}, \ldots, Z_1)$ (Eq. 18); that is

$$\Pr(R_k|Z_k, Z_{k-1}, \ldots, Z_1) = c\,\mathrm{L}(Z_k|R_k)\Pr(R_k|Z_{k-1}, \ldots, Z_1) = c e^{\mathrm{Tr}(XMZ_k^T R_k + \hat{Q}_k^T C R_k)} = \alpha(A_k) e^{\mathrm{Tr}(\hat{R}_k^T A_k R_k)}, \quad (20)$$

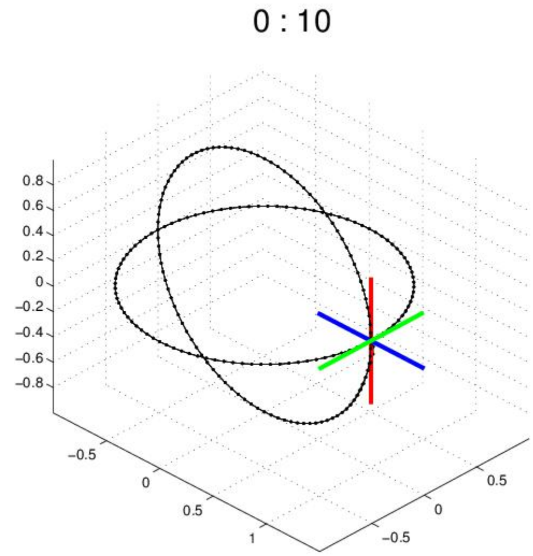

Fig. 1. *The trajectory (normalized) of the end point of x-axis of an IMU*

where the new parameters $\hat{R}_k$ and A_k are computed as the polar decomposition of matrix $XMZ_k^T + \hat{Q}_k^T C_k$.

3. EXPERIMENTAL RESULTS

An experiment was conducted with measurements from a three-axis inertial measurement unit (IMU) with a 100 Hz sampling rate; it was hand held and performed three full rotations around the z-axis, y-axis, and x-axis of the IMU in turn. The trajectory of the endpoint of the unit vector in the x-direction is shown in Figure 1. The measurements were processed, for comparison, with an extended Kalman filter (EKF), implemented to track the Euler angles of the orientation, and by the $\mathcal{SO}(3)$ tracker described here. The position error for both algorithms is shown in Figure 2. Over 10 seconds of the experiment the position errors are almost the same for both algorithms. What is interesting, though, is that the error covariance matrix in our $\mathcal{SO}(3)$ tracker remains essentially constant (See Figure 3), but the EKF's error covariance matrix shrinks in size rapidly, as it shown in Figure 4. This behavior is caused by singularities in the Jacobian for some configurations of Euler angles. Once the error covariance becomes sufficiently small, the EKF had to be re-initialized. Over longer time periods the position error for the EKF tends to *drift*, whereas the position error in the case of the $\mathcal{SO}(3)$ filter remains consistent.

In a second experiment the IMU was placed on a rotating turntable, used for playing vinyl discs, at 33.3 rotations per minute over about 16 minutes with a 100 Hz sampling rate. The position errors are shown in Figure 5. In particular, it shows the drift of the EKF.

4. CONCLUSIONS

A tracking algorithm, specifically designed to track orientation of three dimensional objects, with directional measurements is described here. It uses a Bayesian formalism and the matrix von Mises-Fisher distributions on the rotation group $\mathcal{SO}(3)$. A solution for this

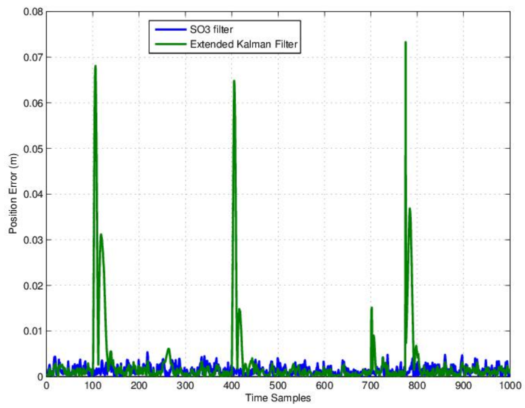

Fig. 2. *Position error (m) for both $\mathcal{SO}(3)$ and EKF, over time period of* 10 *seconds the errors are mostly similar in magnitude*

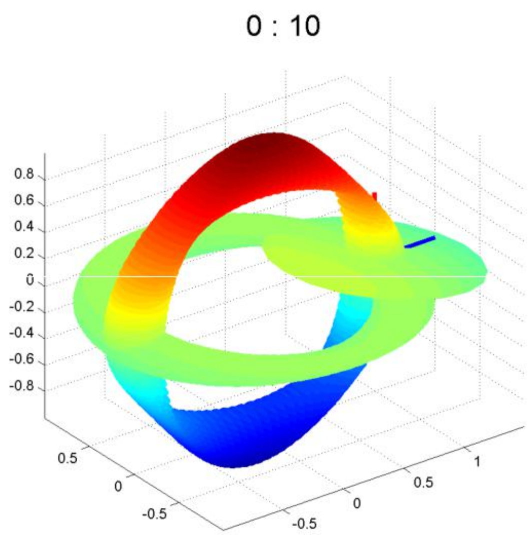

Fig. 3. *Error covariance with $\mathcal{SO}(3)$ filter, represented at each point on trajectory by an ellipsoid.*

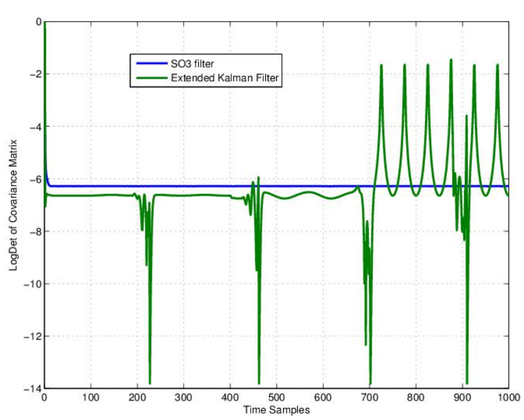

Fig. 4. *Log of determinant of error covariance matrix for both $\mathcal{SO}(3)$ and EKF. EKF has been re-initialized four times in* 10 *second period*

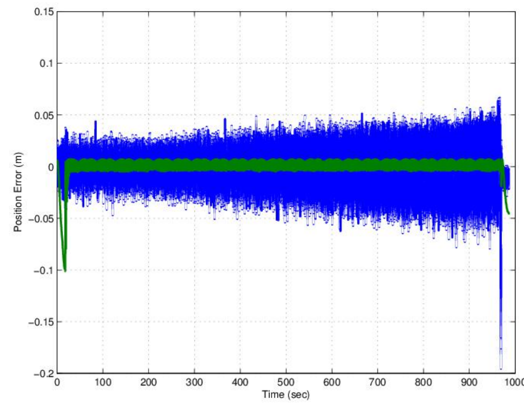

Fig. 5. *Position error for the turntable experiment, shows the drift in EKF*

problem is derived using appropriate approximations which, at least numerically, exhibit essentially no unwarranted assumptions about the measurements. The resulting tracking algorithm is elegant, fast, accurate, does not require re-initialization, and can track for long periods of time. It gives realistic estimates of the spread of the posterior distribution, and hence a correct level of confidence in the estimates of the orientation.

Future work will focus on the uninformative nature of the approximation, and on application to a number of problems including human motion and orientation of small UAVs.

5. REFERENCES

[1] I.V.L. Clarkson, S.D. Howard, W. Moran, D. Cochran, and M.L. Dawson. Maximum-likelihood and best invariant orientation estimation. In *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*, pages 1996–2000. IEEE, 2010.

[2] T.D. Downs. Orientation statistics. *Biometrika*, 59(3):665–676, 1972.

[3] AF Fossum and RM Brannon. Sandia report.

[4] A.T. James. Distributions of matrix variates and latent roots derived from normal samples. *The Annals of Mathematical Statistics*, pages 475–501, 1964.

[5] P.E. Jupp and K.V. Mardia. Maximum likelihood estimators for the matrix von Mises-Fisher and Bingham distributions. *The Annals of Statistics*, 7(3):599–606, 1979.

[6] C.G. Khatri and K.V. Mardia. The von Mises-Fisher matrix distribution in orientation statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 95–106, 1977.

[7] EJ Lefferts, F.L. Markley, and MD Shuster. Kalman filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 5(5):417–429, 1982.

[8] R. Mahony, T. Hamel, and J.M. Pflimlin. Nonlinear complementary filters on the special orthogonal group. *Automatic Control, IEEE Transactions on*, 53(5):1203–1218, 2008.

[9] M. Moakher. Means and averaging in the group of rotations. *SIAM Journal on Matrix Analysis and Applications*, 24(1):1–16, 2002.

[10] A.F. Nikiforov, V.B. Uvarov, and R.P. Boas. *Special functions of Mathematical Physics*. Birkhäuser Basel Boston, 1988.

[11] N. Phuong, H.J. Kang, Y.S. Suh, and Y.S. Ro. A DCM based orientation estimation algorithm with an inertial measurement unit and a magnetic compass. *Journal of Universal Computer Science*, 15(4):859–876, 2009.

[12] A.M. Sabatini. Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing. *Biomedical Engineering, IEEE Transactions on*, 53(7):1346–1356, 2006.

[13] S. Stančin and S. Tomažič. Angle estimation of simultaneous orthogonal rotations from 3D gyroscope measurements. *Sensors*, 11(9):8536–8549, 2011.

[14] E.W. Weisstein et al. Rodrigues rotation formula. *From MathWorld–A Wolfram Web Resource. http://mathworld. wolfram. com/RodriguesRotationFormula. html*, 2006.