Three-dimensional Simultaneous Shape and Pose Estimation for Extended Objects Using Spherical Harmonics

Gerhard Kurz, Florian Faion, Florian Pfaff, Antonio Zea, and Uwe D. Hanebeck

Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics,
Karlsruhe Institute of Technology (KIT), Germany

e-mail: kurz.gerhard@gmail.com, florian.faion@gmail.com, florian.pfaff@kit.edu, antonio.zea@kit.edu, uwe.hanebeck@ieee.org

Abstract

We propose a new recursive method for simultaneous estimation of both the pose and the shape of a three-dimensional extended object. The key idea of the presented method is to represent the shape of the object using spherical harmonics, similar to the way Fourier series can be used in the two-dimensional case. This allows us to derive a measurement equation that can be used within the framework of nonlinear filters such as the UKF. We provide both simulative and experimental evaluations of the novel techniques.

1. Introduction

We consider the problem of simultaneously estimating the position and shape of an extended object. This problem is very widespread in robotics and autonomous navigation whenever a robot encounters unknown objects whose shape is not known a priori.

The problem of extended object tracking has been considered in literature by a number of authors [1, 2, 3, 4, 5, 6, 7]. In some cases, the shape of the object is assumed to be known a priori, in other cases certain parameters (e.g., the radius of a circle, the length of the major and minor axis of an ellipse) are assumed to be unknown, and sometimes the entire shape of the object has to be estimated. One of the key challenges in extended object tracking is the question of how to associate measurements stemming from the object to a particular measurement source on the object. Common solutions are the greedy association model (GAM) [8] that associates each measurement to the most likely source and the spatial distribution model (SDM) [9] that assumes a probability distribution of how likely each possible source is.

It is not only desirable to track the movement of the object but to estimate its shape as well, thus creating a model of the observed object. The techniques presented in this paper can even be applied to nonrigid objects whose shape is subject to changes over time. Another application of the proposed methods consists in the classification of objects, where objects in the environment can be identified based on their estimated shape by comparing with a database of known objects. Moreover, the shape of an object can be used for grasp planning in robotic manipulation.

The main deficiency of most approaches in literature is that they are limited to two-dimensional problems. To address this limitation, we propose a three-dimensional method in this paper.

When estimating the shape of an extended object, a key question is the choice of the representation of the shape. In literature, Fourier series [3], different types of splines [10, 11], as well as Gaussian Processes [12, 5, 7, 13] have been proposed for the 2D case. In the following, we will rely on a three-dimensional generalization of the Fourier series approach. This approach—in 2D as well as in
3D—represents star-convex\textsuperscript{1} shapes by defining a function that maps every direction to the distance from the so-called star point to the boundary of the object along that direction. An illustration of this concept can be seen in Fig. 1. Another 3D approach in literature is the extrusion random hypersurface model proposed in [14]. However, this method is limited to estimation of objects that can be well represented using (scaled) extrusions, such as rotationally symmetric objects. Furthermore, there is an approach based on extended Gaussian images by Sun et al. [15] that can in principle be generalized to the three-dimensional case.

The key idea in this paper consists in representing the shape of an extended object using so-called spherical harmonics. Spherical harmonics can be seen as a counterpart of Fourier series defined on the unit sphere [16, 17, 18, 19], [20, Sec. 4]. Thus, they allow approximation of arbitrary functions defined on the sphere and can be used to elegantly represent the shape of the object under consideration. As a result, estimating the shape is equivalent to estimating the parameters of the weights of the different basis functions. Spherical harmonics have previously been used as a representation of spherical probability distributions [21], as rotation invariant shape descriptors [22], and for view-based robot localization [23].

2. Spherical Harmonics

Spherical harmonics are special functions on the surface of a sphere and resemble Fourier series on the circle. In this section, we will give a concise introduction into the topic of spherical harmonics.

2.1. Basis Functions

First, we introduce the basis functions that are defined on the sphere. In this paper, we use spherical coordinates according to the following convention

\begin{align*}
  x &= r \sin(\theta) \cos(\phi), \\
  y &= r \sin(\theta) \sin(\phi), \\
  z &= r \cos(\theta),
\end{align*}

where \( \phi \in [0, 2\pi), \theta \in [0, \pi], r \geq 0 \).

There are various different conventions for the definition of spherical harmonics, which mostly differ by constant factors and sign. We use the definition by Blanco et al. [24], which in turn is based on Chisholm’s work [20].

---

\textsuperscript{1}A set is star-convex if there exists a star point inside the set such that for every point in the set the line connecting this point to the star point is entirely contained within the set.
There are complex and real versions of spherical harmonics. The basis functions for complex spherical harmonics are defined as \[ Y^m_l(\theta,\phi) = (-1)^m N^m_l P^m_l(\cos(\theta)) \exp(i m \phi) \]
with constants $N^m_l$ and associated Legendre polynomials $P^m_l(\cdot)$, where $l \in \mathbb{N}_0, m \in \mathbb{Z}, -l \leq m \leq l$. This yields a triangular structure as illustrated in Fig. 2. Thus, the number of basis functions up to order $L$ is $(L+1)^2$. The constant factors $N^m_l$ are given according to \[ N^m_l = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}}. \]
Furthermore, the associated Legendre polynomials $P^m_l$ [24, Ch. 8] are defined according to \[ P^m_l(x) = \frac{1}{2^l l!} (1-x^2)^{l/2} \left( \frac{\partial^l}{\partial x^l} \right)^m (x^2 - 1)^l. \]
Following [24, eq. (6)], we now define the real spherical harmonics according to
\[
S^m_l(\theta, \phi) = \begin{cases} 
(-1)^m \sqrt{2} (Y^m_l(\theta, \phi) + Y^{-m}_l(\theta, \phi)) & m > 0 \\
Y^m_l(\theta, \phi) & m = 0 \\
(-1)^m \sqrt{2} (Y^{-m}_l(\theta, \phi) - Y^m_l(\theta, \phi)) & m < 0
\end{cases}
\]
where $\overline{a}$ refers to the complex conjugate of a complex number $a \in \mathbb{C}$. This expression can be simplified according to
\[
S^m_l(\theta, \phi) = \begin{cases} 
N^m_l P^m_l(\cos(\theta)) \sqrt{2} \cos(m \phi) & m > 0 \\
N^m_l P^0_l(\cos(\theta)) & m = 0 \\
N^{-m}_l P^{-m}_l(\cos(\theta)) \sqrt{2} \sin(-m \phi) & m < 0
\end{cases}
\]
These basis functions are illustrated in Fig. 2 for $l = 0, \ldots, 4$. 

Figure 2: Basis functions for spherical harmonics. Positive values are depicted in red, negative values in blue.
2.2. Series Representation

Using the basis functions defined above, we can use a series representation to approximate functions on the unit sphere. This is similar to the way a Fourier series approximates periodic functions using cosine and sine basis functions of different frequencies.

The real-valued series representation using spherical harmonics basis functions is given by

\[ f(\theta, \phi) = \sum_{l=0}^{L} \sum_{m=-l}^{l} w_l^m S_l^m(\theta, \phi), \]

i.e., each basis function is weighted with a coefficient \( w_l^m \). Thus, a star-convex shape in \( \mathbb{R}^3 \) with star point \([0, 0, 0]^T\) can be represented in spherical coordinates with \( f(\theta, \phi) \).

In order to achieve an efficient implementation, it is possible to drop all constants from \( S_l^m(\theta, \phi) \), i.e., all terms that only depend on \( l \) and \( m \) and include them in the weight \( w_l^m \). In that case, the uncertainty of the weighting parameter needs to be scaled accordingly, i.e., its covariance matrix must be divided by the square of the omitted constants.

3. Bayesian Extended Object Tracking

The system state \( x_k \) at time step \( k \) contains the parameters determining the shape of the extended object under consideration as well as further information, such as its position. The shape is represented using the weighting coefficients of the spherical harmonics series. Thus, we can define the state of an object at position \([p_1, p_2, p_3]^T\) as

\[ x = [p_1, p_2, p_3, w_0^0, w_1^0, w_1^1, \ldots]^T, \]

where we omit the time index \( k \). Depending on the application, it is possible to augment the state with other quantities such as orientation, velocity, or angular velocity.

3.1. System Model

The state vector evolves according to a system model

\[ x_{k+1} = a_k(x_k, u_k) + w_k, \]

where \( w_k \) is zero-mean Gaussian noise. The most basic model is a random walk model where \( a_k(x_k, u_k) = x_k \). If the object is rotating at a known angular velocity, this can be modeled as an input \( u_k \), which leads to a nonlinear system model. Here, the rotation function \( a_k(\cdot, \cdot) \) performs the rotation by computing new spherical harmonics coefficients as described in [24]. This method is similar to shifting a Fourier series by changing its phase.

If the angular velocity is not known, it is also possible to include it into the state vector and to estimate it along with the other parameters. Note, however, that this entails a certain ambiguity because a change in orientation can be misconstrued as a change in shape and vice versa.

3.2. Measurement Model

The measurement equation can be derived using the spherical harmonics series representation. First, we shift the measurement \( y_k \) by \([p_1, p_2, p_3]^T\) to obtain \( \tilde{y}_k = y_k - [p_1, p_2, p_3] \). Then, we convert the coordinates of \( \tilde{y}_k \) into polar coordinates. Finally, we obtain the equation

\[ \tilde{z} = \tilde{z}/\|\tilde{z}\| \sum_{l=0}^{L} \sum_{m=-l}^{l} w_l^m S_l^m(\theta_z, \phi_z) + \nu, \]

where we omit the time indices for improved readability. Here, \( \theta_z \) and \( \phi_z \) are polar coordinates of \( \tilde{z} \) and \( \nu \) is zero-mean Gaussian noise. The labeled term corresponds to the measurement source,
i.e., the point on the objects surface that the measurement is assumed to originate from based on the greedy association model. Observe that $\hat{z}$ occurs on both sides of the measurement equation. This is common in extended object tracking because the measurement source depends on the actual measurement. To implement this equation, we first use the known measurement to determine the measurement source and assume it to be a constant. Then, we perform the update as usual, since the right side does not depend on $\hat{z}$ anymore.

This equation can be used as the measurement equation in a suitable nonlinear filter. For example, one can use a nonlinear version of the Kalman filter such as the UKF [26] or the S$^2$KF [27].

In the proposed measurement equation, the association problem is solved using a greedy association model (GAM), where each measurement is projected onto the surface of the estimated object along the line from the measurement to the object’s star point. This approach works well in most cases. However, it has been shown that greedy association models can cause biased estimates of the object’s size if the noise is very large compared to the size of the object [8]. To address this issue, the proposed approach could be extended using the partial information model (PIM) as described in [1] to eliminate the bias.

4. Simulations

In this section, we illustrate the proposed method in multiple simulations.

4.1. Cuboid

First, we apply our approach to estimation of a cuboid of size $3 \times 1 \times 1$ units. For this purpose, we generate measurements at each time step by sampling a predefined number of random points from the cuboid’s surface and adding isotropic Gaussian measurement noise with a variance of $10^{-2}$. In the following simulations, we always use 60 measurements per time step. Then, we perform recursive estimation using an unscented Kalman filter (UKF). The same filter was used in all simulations and experiments in this paper. To quantify the accuracy of the result, we consider the Intersection over Union (IoU), also known as the Jaccard index [28], i.e., the volume of the intersection of the cuboid...
and the estimated shape divided by the volume of their union. The IoU is always in [0,1], where 1 corresponds to an exact match of the two objects.

In Fig. 3, we depict how the estimated shape converges toward the true shape of the cube over time. The measurements at each time step are shown as black dots. It can be seen that the estimate of the shape is quite poor at first and gets better with every time step. After a few time steps, the shape estimate converges to a fairly accurate representation of the true shape of the cuboid.

It is obvious that the accuracy of the shape estimate depends on the order $L$ of the spherical harmonics, and thus, the number of coefficients used to represent the shape. This effect is illustrated in Fig. 4, where we show the estimated shape after convergence for different values of $L$. As is obvious, a higher order allows a more accurate representation of the shape, especially around the corners of the cube.

In order to illustrate the effect of the system model on the estimation results, we perform an experiment with a rotating cuboid. For this experiment, the cube rotates around a fixed axis with an angular rate of 10 degrees per time step. We compare the results from a random walk system model and a system model that considers the angular velocity as a known input and performs an appropriate rotation of the spherical harmonics coefficients according to [24]. The results of this experiment are shown in Fig. 5. As expected, the shape is estimated much more accurately with the proper system model that considers the rotation.
4.2. Complex 3D Model

Now, we apply the algorithm used in the previous section to a more complicated 3D model. For this purpose, we use a model of the famous Utah teapot. Our version of the model has a size of about $3.5 \times 2.2 \times 1.8$ units. For our simulation, we sample 200 random points from the surface of the 3D model at each time step and add isotropic Gaussian measurement noise with a variance of $10^{-2}$ to each point.

The estimated shape after convergence is shown for different numbers of coefficients in Fig. 6. Once again, it is clearly visible how a larger number of coefficients (i.e., a higher order) allows representation of more complex shapes.

5. Experiment

In order to evaluate the proposed approach in a real scenario, we consider simultaneous tracking and shape estimation of a moving watering can. The watering can is a well-suited object because it has a very characteristic shape and can be approximated quite well using a star-convex model despite possessing non-star convex parts such as the handle.

To obtain measurements, we recorded the setting simultaneously with two Kinect 2.0 sensors located on opposite sides of the object, such that measurements from both directions can be obtained (see Fig. 7). These sensors are commonly used in various robotics applications. The extrinsic parameters of the Kinect cameras are calibrated using a checkerboard method. Both depth and color images were recorded during the experiment. An overview of our setup is shown in Fig. 8.

We perform segmentation using a bounding box on the 3D points and color thresholding in HSV space. Afterward, we remove outliers that have too few other points close by. This way, we obtain 3D points that belong to the tracked object. Each Kinect records at a frame rate of 30 Hz, so we obtain a total of 60 point clouds per second. Then, these 3D points are used as measurements for the spherical harmonics algorithm. We assume that the measurements are affected by isotropic Gaussian noise. As an initialization for the location, we use the center of mass of the point cloud at the first time step. The shape is initialized as a sphere, i.e., all $w_m^n$ except $w_0^0$ were zero. For comparison, we implemented the random matrix approach, which is limited to estimation of ellipsoidal shapes. Furthermore, we implemented the alpha shape approach, which can be seen as a generalization of the convex hull. However, it does not consider noise and is not able to fuse information recursively.

---

2Nik Clark, Utah teapot (solid), https://www.thingiverse.com/thing:852078
The results of our experiment are shown Fig. 9. It can be seen that a higher number of coefficients once again leads to a more detailed reconstruction. As far as tracking of the location is concerned, all approaches seem to be able to follow the movement of the watering can quite well. The shape estimate of the random matrices approach is very poor as that method is limited to ellipsoidal shapes. While the shape estimate obtained by the alpha volume approach is much more detailed, it heavily suffers from noise and outliers. Also, the estimate can change significantly from one time step to the next.

6. Conclusion

We have presented a novel method for extended object tracking with simultaneous shape estimation in three dimensions. The key idea of this method is to represent star-convex objects using spherical harmonics and to estimate their coefficients using standard nonlinear estimation techniques.

Through an evaluation of the proposed method in several simulations as well as a real experiment, we have shown its applicability to a number of different scenarios. The experiments show that a larger number of coefficients allows for a more accurate shape representation. We also illustrate how a known angular velocity can be incorporated. Furthermore, the advantages of the proposed approach compared with random matrices, a method limited to ellipsoidal shapes, as well as alpha shapes has been shown.

Future work may include the use of Gaussian processes [12] instead of spherical harmonics, development of methods that can handle shapes that are not star-convex, and research on methods that can take advantage of symmetries. An extension of the approach based on level sets proposed in [33] to three dimensions might also be interesting to consider. In order to estimate volumetric 3D objects, i.e., objects where measurements are obtained from anywhere within the object rather than just its surface, it would also be possible to combine the random hypersurface model (RHM) idea from [3] with the spherical harmonics approach.

Acknowledgment

This work was partially supported by the German Research Foundation (DFG) under grant HA 3789/13-1.
Figure 9: Experimental results for order 6 (top row), 10 (second row), and 15 (third row), as well as random matrices (fourth row), and alpha shapes (fifth row) at different time steps. All units are in meters. The measurements of the current time step are shown as black dots.
References

[1] F. Faion, A. Zea, M. Baum, and U. D. Hanebeck, “Partial Likelihood for Unbiased Extended Object Tracking,” in Proceedings of the 18th International Conference on Information Fusion (Fusion 2015), Washington D. C., USA, Jul. 2015.

[2] ——, “Symmetries in Bayesian Extended Object Tracking,” Journal of Advances in Information Fusion, vol. 10, no. 1, pp. 13–30, Jun. 2015.

[3] M. Baum and U. D. Hanebeck, “Shape Tracking of Extended Objects and Group Targets with Star-Convex RHMs,” in Proceedings of the 14th International Conference on Information Fusion (Fusion 2011), Chicago, Illinois, USA, Jul. 2011.

[4] ——, “Fitting Conics to Noisy Data Using Stochastic Linearization,” in Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2011), San Francisco, California, USA, Sep. 2011.

[5] N. Wahlström and E. Ozkan, “Extended Target Tracking Using Gaussian Processes,” IEEE Transactions on Signal Processing, vol. 63, no. 16, pp. 4165–4178, 2015.

[6] M. Feldmann, D. Franken, and W. Koch, “Tracking of Extended Objects and Group Targets Using Random Matrices,” IEEE Transactions on Signal Processing, vol. 59, no. 4, pp. 1409–1420, 2011.

[7] T. Hirscher, A. Scheel, S. Reuter, and K. Dietmayer, “Multiple Extended Object Tracking Using Gaussian Processes,” in Proceedings of the 19th International Conference on Information Fusion (Fusion 2016), Heidelberg, Germany, Jul. 2016.

[8] F. Faion, A. Zea, and U. D. Hanebeck, “Reducing Bias in Bayesian Shape Estimation,” in Proceedings of the 17th International Conference on Information Fusion (Fusion 2014), Salamanca, Spain, Jul. 2014.

[9] K. Gilholm and D. Salmond, “Spatial Distribution Model for Tracking Extended Objects,” IEE Proceedings – Radar, Sonar and Navigation, vol. 152, no. 5, pp. 364–371, Oct. 2005.

[10] G. Kurz and U. D. Hanebeck, “Dynamic Surface Reconstruction by Recursive Fusion of Depth and Position Measurements,” Journal of Advances in Information Fusion, vol. 9, no. 1, pp. 13–26, Jun. 2014.

[11] A. Zea, F. Faion, and U. D. Hanebeck, “Tracking Elongated Extended Objects Using Splines,” in Proceedings of the 19th International Conference on Information Fusion (Fusion 2016), Heidelberg, Germany, Jul. 2016.

[12] S. Dragiev, M. Toussaint, and M. Gienger, “Gaussian Process Implicit Surfaces for Shape Estimation and Grasping,” in 2011 IEEE International Conference on Robotics and Automation (ICRA), May 2011, pp. 2845–2850.

[13] E. Ozkan, N. Wahlström, and S. J. Godsill, “Rao-Blackwellised Particle Filter for Star-Convex Target Tracking Models,” in Proceedings of the 19th International Conference on Information Fusion (Fusion 2016), Heidelberg, Germany, Jul. 2016.

[14] A. Zea, F. Faion, and U. D. Hanebeck, “Tracking Extended Objects using Extrusion Random Hypersurface Models,” in Proceedings of the IEEE ISIF Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF 2014), Bonn, Germany, Oct. 2014.

[15] L. Sun, X. R. Li, and J. Lan, “Modeling of Extended Objects Based on Support Functions and Extended Gaussian Images for Target Tracking,” IEEE Transactions on Aerospace and Electronic Systems, vol. 50, no. 4, pp. 3021–3035, 2014.

[16] K. Atkinson and W. Han, Spherical Harmonics and Approximations on the Unit Sphere: An Introduction, ser. Lecture Notes in Mathematics. Springer Science & Business Media, 2012, vol. 2044.

[17] C. Müller, Spherical Harmonics, ser. Lecture Notes in Mathematics. Heidelberg, Germany: Springer-Verlag, 1966, vol. 17.

[18] H. Groemer, Geometric Applications of Fourier Series and Spherical Harmonics, ser. Encyclopedia of Mathematics and its Applications. Cambridge: Cambridge University Press, 1996, vol. 61.

[19] T. M. MacRobert, Spherical Harmonics: An Elementary Treatise on Harmonic Functions with Applications, 2nd ed. Dover Publications, 1948.

[20] C. D. H. Chisholm, Group Theoretical Techniques in Quantum Chemistry. London: Academic Press, 1976.

[21] F. Pfaff, G. Kurz, and U. D. Hanebeck, “Filtering on the Unit Sphere Using Spherical Harmonics,” in Proceedings of the 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2017), Daegu, Republic of Korea, Nov. 2017.

[22] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, “Rotation Invariant Spherical Harmonic Representation of 3D Shape Descriptors,” in Eurographics Symposium on Geometry Processing, vol. 6, 2003, pp. 156–164.

[23] H. Friedrich, D. Dederscheck, K. Krajsk, and R. Mester, “View-based Robot Localization Using Spherical Harmonics: Concept and First Experimental Results,” Pattern Recognition, pp. 21–31, 2007.

[24] M. A. Blanco, M. Flörez, and M. Bermejo, “Evaluation of the Rotation Matrices in the Basis of Real Spherical Harmonics,” Journal of Molecular Structure (Theochem), vol. 419, pp. 19–27, 1997.

[25] M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 10th ed. New York: Dover, 1972.

[26] S. J. Julier and J. K. Uhlmann, “Unscented Filtering and Nonlinear Estimation,” Proceedings of the IEEE, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[27] J. Steinbring, M. Pander, and U. D. Hanebeck, “The Smart Sampling Kalman Filter with Symmetric Samples,” Journal of Advances in Information Fusion, vol. 11, no. 1, pp. 71–90, Jun. 2016.

[28] M. Levandowsky and D. Winter, “Distance Between Sets,” Nature, vol. 234, no. 5323, pp. 34–35, 1971.
[29] C. S. Bamji, P. O’Connor, T. Elkhatib, S. Mehta, B. Thompson, L. A. Prather, D. Snow, O. C. Akkaya, A. Daniel, A. D. Payne, T. Perry, M. Fenton, and V. H. Chan, “A 0.13 µm CMOS System-on-Chip for a 512 × 424 Time-of-Flight Image Sensor With Multi-Frequency Photo-Demodulation up to 130 MHz and 2 GS/s ADC,” *IEEE Journal of Solid-State Circuits*, vol. 50, no. 1, pp. 303–319, Jan. 2015.

[30] L. Yang, L. Zhang, H. Dong, A. Arelawi, and A. E. Saddik, “Evaluating and Improving the Depth Accuracy of Kinect for Windows v2,” *IEEE Sensors Journal*, vol. 15, no. 8, pp. 4275–4285, Aug. 2015.

[31] P. Fankhauser, M. Bloesch, D. Rodriguez, R. Kaestner, M. Hutter, and R. Siegwart, “Kinect v2 for mobile robot navigation: Evaluation and modeling,” in *Proc. Int Advanced Robotics (ICAR) Conf.*, Jul. 2015, pp. 388-394.

[32] N. Akkiraju, H. Edelsbrunner, M. Facello, P. Fu, E. Mucke, and C. Varela, “Alpha Shapes: Definition and Software,” in *Proceedings of the 1st International Computational Geometry Software Workshop*, vol. 63, 1995, p. 66.

[33] A. Zea, F. Faion, M. Baum, and U. D. Hanebeck, “Level-Set Random Hyper Surface Models for Tracking Non-Convex Extended Objects,” in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.