# Tracking Extended Objects using Extrusion Random Hypersurface Models

**Antonio Zea**, **Florian Faion**, and **Uwe D. Hanebeck**

Intelligent Sensor-Actuator-Systems Laboratory (ISAS)

Institute for Anthropomatics and Robotics

Karlsruhe Institute of Technology (KIT), Germany

antonio.zea@kit.edu, florian.faion@kit.edu, uwe.hanebeck@ieee.org

*Abstract*—As sensor resolution increases, the accuracy and robustness of tracking algorithms can be improved by incorporating more information about the shape of the target object. This raises the need for simple and robust shape models capable of describing detailed objects. In this paper we propose an approach based on *Random Hypersurface Models* that interprets target shapes as scaled extrusions. This is achieved by combining projection-based models with probabilistic approaches, integrating the strengths of both mechanisms. As extruded shapes such as bottles, boxes, or containers can be extensively found in everyday situations, this approach can be applied for tracking in a large variety of environments.

*Index Terms*—Extended object tracking, extrusions, shape models, cylinder, solid of revolution.

## I. INTRODUCTION

Object tracking classically works with the assumption that the target shape consists of a single point. However, modern sensors are typically capable of resolving multiple measurement sources, allowing tracking approaches that exploit information about the target shape to yield more robust and accurate results. In many cases the shape parameters are not available beforehand, requiring both pose and shape to be estimated simultaneously. This raises the need for simple yet powerful shape models that can take advantage of the information available through the measurements.

For simple shapes, the mechanisms used to describe them can be divided into two classes. In the first class are models based on *fitting*, i.e., minimizing a distance function to a projection, used in particular for conics [1]–[3]. In the second class are approaches that develop probabilistic models explicitly considering *all* sources [4], e.g., in estimators for line segments [5].

For more complex shapes, however, there are two interconnected issues. On the one hand, detailed models allow for more flexibility in describing a variety of shapes. For example, active contours are often used in computer vision to describe complex objects [6]. On the other hand, simpler models have the benefit of being robust and yielding high performance. For instance, rigid object tracking [7] and articulations using basic shapes [8] are popular in literature. As explicit probabilistic models tend to be intractable for detailed shapes, fitting techniques are generally used instead.

However, an alternative approach consists of *constructing* detailed objects using simple shapes, allowing the combination



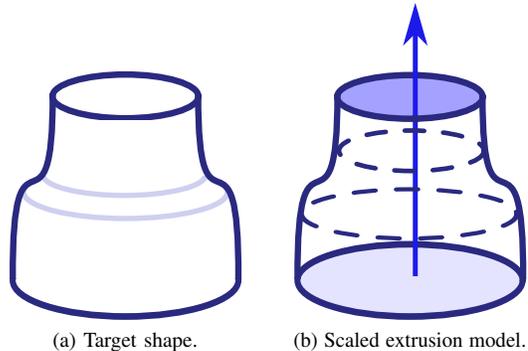(a) Target shape.      (b) Scaled extrusion model.

Figure 1: Modeling an extended target as a scaled extrusion.

of probabilistic and projection-based techniques. One of these is called *Random Hypersurface Models* (RHMs) [9], which describe detailed shapes as the transformations of a base shape. Using RHMs, shapes such as filled disks [10], arbitrary filled star-convex [11] and non-convex shapes [12] have been modeled. In addition, cylinders were used to track objects in 3D point clouds [13].

In this paper, we propose a straightforward mechanism to extend RHMs in order to model scaled extrusions such as in Fig. 1. As extruded objects are easy to manufacture, they are widely present in a large variety of environments. These include familiar objects such as bottles, containers, metal bars, tubes, and many others. In Fig. 1, for example, a target object (Fig. 1a) is seen as the scaled extrusion of a base shape (Fig. 1b).

This paper is structured as follows. First, Sec. II details the problem formulation, and then Sec. III describes how to model extended objects. Implementation details are explained in Sec. V, and an evaluation is presented in Sec. VI. Finally, Sec. VII concludes this paper.

## II. PROBLEM FORMULATION

This paper deals with estimating the parameters of a target object based on received noisy point measurements.

The target parameters for the time step $k$ are contained in the state $\underline{x}_k$. Thus, the probability density $p(\underline{x}_k)$ represents the uncertain knowledge of the parameters. The received point measurements $\mathcal{Y}_k = \{\underline{y}_{k,0}, \cdots, \underline{y}_{k,l}\}$ are assumed to stem from the boundary of the target shape. Each measurement

(a) Spatial distribution model.　　(b) Greedy association model.

Figure 2: Commonly used shape models.



(a) Correct shape.　　(b) Incorrect size.　　(c) Incorrect pose.
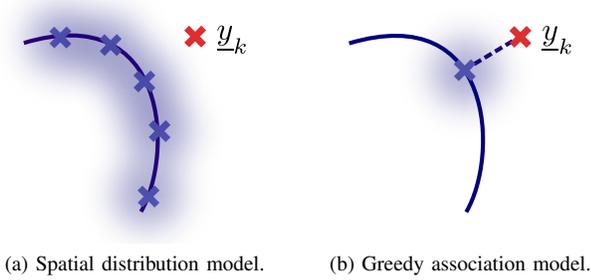
Figure 3: Issues using GAMs.

$\underline{y}_{k,i} \in \mathbb{R}^d$, where $d$ is generally two or three, is described in Cartesian coordinates. It is not assumed that the number of measurements contains any information about the target. Note that the measurement index will be dropped unless needed.

Each measurement $\underline{y}_k \in \mathcal{Y}_k$ is related to the state $\underline{x}_k$ through a *measurement model*, which can be divided into a *sensor model* and a *shape model*. On the one hand, the sensor model assumes that $\underline{y}_k$ originates from a measurement source $\underline{z}_k$ drawn from the boundary of the target shape, described by the set $\mathcal{Z}(\underline{x}_k)$, and then is disturbed during observation by an additive noise term $\underline{v}_k$. It is assumed that $\underline{v}_k$ is drawn from a Gaussian distribution with zero mean and covariance matrix $\mathbf{C}_k^v$, so that

$$p(\underline{y}_k \mid \underline{z}_k) = \mathcal{N}(\underline{y}_k - \underline{z}_k; \underline{0}, \mathbf{C}_k^v) \tag{1}$$

holds. On the other hand, the shape model describes how a measurement source $\underline{z}_k$ is generated in function of $\underline{x}_k$, and depends on the characteristics of the target shape $\mathcal{Z}(\underline{x}_k)$. This can be described using the probability density $p(\underline{z}_k \mid \underline{x}_k)$.

The relation between a given measurement $\underline{y}_k$ and the state $\underline{x}_k$ can be described probabilistically using a likelihood $p(\underline{y}_k \mid \underline{x}_k)$. It is assumed that each measurement noise term is drawn independently from other noise terms and from the state.

## III. MODELING EXTENDED OBJECTS

A probabilistic measurement model that relates $\underline{y}_k$ to the state parameters $\underline{x}_k$ can be described by combining the sensor and shape models from Sec. II, in the form of

$$p(\underline{y}_k \mid \underline{x}_k) = \int_{\mathcal{Z}(\underline{x}_k)} p(\underline{y}_k \mid \underline{z}_k) \cdot p(\underline{z}_k \mid \underline{x}_k) \, \mathrm{d}\underline{z}_k \ , \tag{2}$$

which considers all possible sources $\underline{z}_k$ from the target shape $\mathcal{Z}(\underline{x}_k)$. This model, which explicitly assigns each source a probability of generating a measurement, is called a *spatial distribution model* [4] (SDM), shown in Fig. 2a. The advantage of this approach is that it is able to incorporate information such as object-to-sensor geometry to improve estimation, by adjusting $p(\underline{z}_k \mid \underline{x}_k)$ to specify which sources are more likely to be measured. However, this also raises the requirement that $p(\underline{z}_k \mid \underline{x}_k)$ needs to be explicitly known, which in cases of artifacts and unexpected occlusions might be extremely
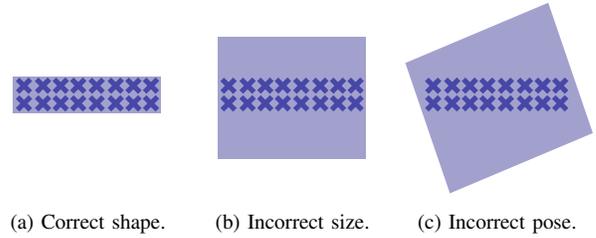
difficult. In addition, evaluating (2) requires considering all possible sources, which may be intractable or inefficient.

In contrast, an approach commonly used in fitting [15] consists of simplifying $p(\underline{z}_k \mid \underline{x}_k)$ by considering only a single source, denoted as the *projection* $\pi(\underline{y}_k, \mathcal{Z}(\underline{x}_k))$. This projection is selected as the closest source in $\mathcal{Z}(\underline{x}_k)$ to the measurement $\underline{y}_k$, based on a given distance function. The task for the estimator is, then, to minimize the distances between measurements and their projections. The projection $\pi(\cdot, \cdot)$ can be determined, for example, by selecting the source with the smallest Euclidian or Mahalanobis distance to the measurement. In this work, we denote these shape models as *greedy association models* (GAMs) [14], shown in Fig. 2b.

GAMs can be seen as a special case of SDMs where all the probability mass is assumed to be concentrated on the projection, assumed to be the real source. Thus, as this approach does not require the probability density to be explicitly defined for all sources, it can deal better with occlusions and artifacts. As an example, let the distance function be the difference between the measurement $\underline{y}_k$ and its projection $\pi(\underline{y}_k, \mathcal{Z}(\underline{x}_k))$. Then, the probability density $p(\underline{z}_k \mid \underline{x}_k)$ can be simplified to

$$p(\underline{z}_k \mid \underline{x}_k) = \delta(\underline{z}_k - \pi(\underline{y}_k, \mathcal{Z}(\underline{x}_k))) \ , \tag{3}$$

where $\delta(\cdot)$ is the Dirac delta function. After plugging this into (2) using (1), we obtain the simple likelihood

$$p(\underline{y}_k \mid \underline{x}_k) = \mathcal{N}(\underline{y}_k - \pi(\underline{y}_k, \mathcal{Z}(\underline{x}_k)); \underline{0}, \mathbf{C}_k^v) \ .$$

However, this crude approximation introduces two issues in GAMs. On the one hand, we have the problem of *estimation bias* [14], which appears as a consequence of the fact that, in many cases, $\pi(\underline{y}_k, \mathcal{Z}(\underline{x}_k)) \neq \underline{z}_k$ due to the effect of noise and shape curvature. On the other hand, it cannot exploit additional information about the distribution of sources on the shape, which leads to issues with filled shapes [5], [14] as seen in Fig. 3. In the three cases, the distance between the sources (dark blue) and the estimated shapes (light blue) is always 0. This means that the estimator has no way of finding out which estimated shape is the correct. As explicit SDMs take into account the distribution of sources in the shape, this issue is not present in those models.

As both the explicit and greedy treatments of $p(\underline{z}_k \mid \underline{x}_k)$ bring their own advantages and disadvantages, it is worth examining how both approaches can be combined.

## IV. EXTRUSION RANDOM HYPERSURFACE MODELS

In this paper, we explore the approach of modeling extended objects as *scaled extrusions*. Informally, constructing an extrusion can be seen as taking a given *base shape* (Fig. 4a) and translating it along a given axis. The result, denoted as the *extruded shape*, is in turn the union of all these translated shapes (Fig. 4b). For example, a cylinder can be seen as an extrusion using a circular base shape. For simplicity, it is assumed that the axis is normal to the base shape plane. In addition, the base shape is assumed to be flat, i.e., there exists a plane that contains all of its points.

This concept can be generalized by translating and *scaling* the base shape, where the scaling parameter of the translated shape varies in function of its position (Fig. 4c). We denote this process as *scaled extrusion*. In this way, solids of revolution can be seen as scaled extruded shapes using a circular base shape.

The key idea is, then, to take the simple model of the base shape and use it to describe more complex, scale extruded shapes. This section is concerned with developing a probabilistic shape model for these scaled extrusions.

### A. Random Hypersurface Models

In order to model scaled extrusions, we introduce the concept of *Random Hypersurface Models* (RHMs), which can be interpreted as a combination of spatial distributions and greedy association models. The key idea is to develop a generative model for a target shape $\mathcal{Z}(\underline{x}_k)$ by *transforming* a simpler base shape, denoted as $\mathcal{Z}_b(\underline{x}_k)$. The transformation, determined by the parameter $\tau \in \mathcal{T}$, can be arbitrary and yields the transformed shape $\mathcal{Z}_b^\tau(\underline{x}_k)$. In addition, each transformed shape is assigned an explicit probability $p(\tau)$ of generating a measurement. In general, the base shape is modeled as a GAM.

Thus, the process of generating a measurement source can be described in the following two steps. In the first step, a transformation parameter $\tau$ is randomly selected considering $p(\tau)$. Then, in the second step, a measurement source is selected from $\mathcal{Z}_b^\tau(\underline{x}_k)$ according to its shape model.

Using RHMs, the term $p(\underline{z}_k \mid \underline{x}_k)$ from (2) can be written as

$$p(\underline{z}_k \mid \underline{x}_k) = \int_{\tau \in \mathcal{T}} p(\underline{z}_k \mid \underline{x}_k, \tau) \cdot p(\tau) \, \mathrm{d}\tau \ , \qquad (4)$$

where $p(\underline{z}_k \mid \underline{x}_k, \tau)$ is determined by the shape model of the corresponding transformed shape.

### B. Extrusions as RHMs

The idea of RHMs can be easily extended to describe scaled extrusions by interpreting the transformation of the base shape $\mathcal{Z}_b(\underline{x}_k)$ as a combination of translations and scalings.

The transformed shapes $\mathcal{Z}_b^\tau(\underline{x}_k)$ will be denoted as *slices* (Fig. 5a). For each $\tau \in [0, 1]$ the corresponding slice can be described as

$$\mathcal{Z}_b^\tau(\underline{x}_k) := \{s(\tau) \cdot \underline{z}_k + \tau \cdot \underline{l}_k \mid \underline{z}_k \in \mathcal{Z}_b(\underline{x}_k)\},$$



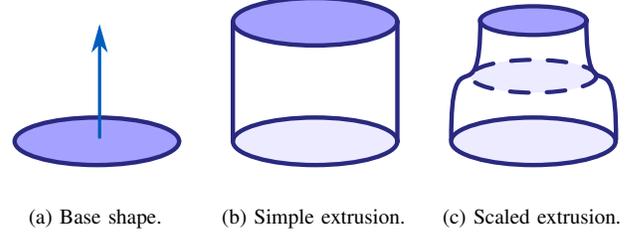(a) Base shape.　　(b) Simple extrusion.　　(c) Scaled extrusion.

Figure 4: Extrusion models.

where $\underline{l}_k$ denotes the *extrusion axis*, and $s(\tau)$ is the *lateral function* that determines the scaling coefficient. Finally, a given $p(\tau)$, denoted as the slice probability density, determines how probable it is that the corresponding slice generates a measurement.

Thus, the parameters to be estimated consist of

$$\underline{x}_k = \left[\underline{x}_k^b, \underline{x}_k^s, x_k^h\right] \ , \qquad (5)$$

where $\underline{x}_k^b$ denotes the parameters of the base shape, $\underline{x}_k^s$ the parameters of the lateral function, and $x_k^h$ determines the height, i.e., the length of the extrusion axis. This approach encompasses both straightforward extrusions when $s(\tau)$ is constant, as handled in [13], and solids of revolution when $\mathcal{Z}_b(\underline{x}_k)$ is a circle. Of course, (5) can be extended to describe additional parameters required by other models.

As this approach is a combination of GAMs and SDMs, it also combines their advantages. On the one hand, as the slices are modeled as GAMs using projections, it retains the robustness in cases of occlusions. On the other hand, as the extrusion axis is modeled as an explicit SDM, we are able to correctly estimate the height and avoid the issues from Fig. 3.

## V. IMPLEMENTATION

In the implementation, we will focus on circular scaled extrusions, i.e., scaled extruded shapes with circles as base shapes.

### A. Extrusion Slices

The base shape is a circle embedded in three dimensions, described with the parameters

$$\underline{x}_k^b = \left[\underline{x}_k^p, \underline{x}_k^o\right] \ ,$$

where $\underline{x}_k^p$ denotes the position of the circle center, and $\underline{x}_k^o$ the orientation. The radius of the base circle is assumed to be 1, as the scaling will be determined by $s(\tau)$ instead.

The shape model for the slices is a GAM, i.e., for a given measurement $\underline{y}_k$ only a single source is considered. This source is selected as the Mahalanobis projection of $\underline{y}_k$ onto the slice, determined from

$$\pi(\underline{y}_k, \mathcal{Z}_b^\tau(\underline{x}_k)) = \underset{\underline{z}_k \in \mathcal{Z}_b^\tau(\underline{x}_k)}{\arg\min} \ (\underline{z}_k - \underline{y}_k)^T (\mathbf{C}_k^v)^{-1} (\underline{z}_k - \underline{y}_k) \ .$$

Note that for isotropic noise this is reduced to a simple Euclidian projection. For non-isotropic noise, a polygonal

approximation of the slice can be used which yields an efficient, closed-form projection.

In order to deal with issues of *bias*, the selected source can be corrected using the approach from [14]. Note that for non-isotropic noise, the estimated parameters may still be slightly biased.

The measurement equation consists simply of comparing $\underline{y}_k$ to its projection onto $\mathcal{Z}_b^\tau(\underline{x}_k)$, in a similar fashion to (3), yielding

$$p(\underline{z}_k \,|\, \underline{x}_k, \tau) = \delta(\underline{z}_k - \pi(\underline{y}_k, \mathcal{Z}_b^\tau(\underline{x}_k))) \ , \qquad (6)$$

where $\delta(\cdot)$ is the Dirac delta function.



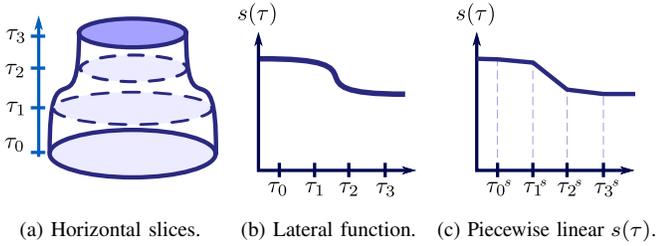(a) Horizontal slices.    (b) Lateral function.    (c) Piecewise linear $s(\tau)$.

Figure 5: Modeling the lateral surface.

## B. The Lateral Function

The lateral function $s(\tau)$ determines the coefficient for the scaling of the slice $\tau$ (Fig. 5b). As the base shape is circular, the extruded shape can be seen as a solid of revolution whose lateral surface is determined by $s(\tau)$. Options to parameterize the lateral function can be Bezier curves, Bernstein polynomials [15] or Fourier series [11]. For simplicity, in this paper we propose to use piecewise linear functions. This allows the easy description of lateral surfaces that are vertically straight, such as cylinders, truncated cones and their combinations, present in many real-life extrusions. In addition, it is straightforward to increase or decrease the level of vertical detail simply by adding or removing a section.

An example implementation is by homogeneously splitting the interval $[0, 1]$ into multiple sections with coordinates $\{\tau_0^s, ..., \tau_m^s\}$, where $\tau_i^s = \frac{i}{m}$. Then, let the shape parameters have the form

$$\underline{x}_k^s = \left[ x_{k,0}^s, ..., x_{k,m}^s \right] \ ,$$

where each $x_{k,i}^s$ determines the value of the lateral function at the coordinate $\tau_i^s$, i.e., $s(\tau_i^s) = x_{k,i}^s$ (Fig. 5c). For the other values of $\tau$, the scaling is obtained from linear interpolations of the nearest sections.

## C. Modeling the Extrusion

Next, we develop a probabilistic model for the scaled extrusion process. For convenience, we define the slice projection

$$\pi_\tau(\underline{y}_k, \underline{x}_k) := \pi(\underline{y}_k, \mathcal{Z}_b^\tau(\underline{x}_k)) \ .$$

Then, by adjusting (6), we obtain

$$p(\underline{z}_k \,|\, \underline{x}_k, \tau) = \delta(\underline{z}_k - \pi_\tau(\underline{y}_k, \underline{x}_k)) \ ,$$

which, when plugged into (4), results in

$$p(\underline{z}_k \,|\, \underline{x}_k) = \int_0^1 \delta(\underline{z}_k - \pi_\tau(\underline{y}_k, \underline{x}_k)) \cdot p(\tau) \ \mathrm{d}\tau \ ,$$

which, in turn, can be inserted into (2), yielding

$$p(\underline{y}_k \,|\, \underline{x}_k) = \int_0^1 \mathcal{N}(\underline{y}_k - \pi_\tau(\underline{y}_k, \underline{x}_k); \underline{0}, \mathbf{C}_k^v) \cdot p(\tau) \ \mathrm{d}\tau \ , \quad (7)$$

so that the likelihood can be described as a weighted Gaussian convolution around the slice projections.

While some forms of $s(\tau)$ and $p(\tau)$ allow for a closed-form evaluation of (7), in general this is untractable. However, given the small range of $\tau$, an implementation can use simple numerical approximations and still obtain good results.

## D. Slice Probabilities

As a reminder, each slice was modeled as a GAM in order to allow the estimator to remain robust even if parts of the slice are occluded. This is particularly useful given that a sensor usually can only observe one side of the object. However, the extrusion axis was modeled using the explicit probability density $p(\tau)$, in order to estimate the height correctly. Because of this, we still require information about the probability of each slice to generate a measurement.

Assuming a low vertical curvature in the lateral surface, a probability density in the form $p(\tau) = \mathcal{U}(\tau; 0, 1)$, i.e., with a uniform distribution in $[0, 1]$, might suffice. In other cases, we need to take into account how the object is observed by the sensors, such as an RGBD camera using a pinhole model or a laser scanner. Thus, a simple approximation is by using the intuitive notion that the probability for each slice to generate measurements depends on its size. From this, we obtain the probability density

$$p(\tau) \approx c \cdot s(\tau) \ ,$$

where $c$ is a normalization constant.

## E. Extrusion Caps

While the focus of this paper is on the lateral surface, it may happen that the top and bottom caps are also visible. In this case, measurements generated from these parts will have an influence on $p(\tau)$.

An approach to solve this issue is by using circular RHMs [10], which work similarly to the proposed approach, by *shrinking* a circular boundary to describe a filled disk. The likelihood of the whole shape is then the weighted sum of the cap disk likelihood and the lateral surface likelihood (7). The required weights are proportional to how much of the surface area of each component is visible to the sensor.

*F. Estimator*

A recursive Bayesian estimator can be developed in a straightforward way. Such an estimator generally has two steps. First, an *update* step is realized using Bayes' rule, in the form of

$$p(\underline{x}_k \mid \underline{y}_k) \propto p(\underline{y}_k \mid \underline{x}_k) \cdot p(\underline{x}_k) \ ,$$

where $p(\underline{x}_k)$ is the prior knowledge of the state and $p(\underline{y}_k \mid \underline{x}_k)$ is the likelihood from (7). Second, a *prediction* step is executed using a given system model.

The update step can be implemented with a variety of filters, such as the PGF [16] or the GPF [17], which deal with explicit likelihoods. An implementation is also possible using non-linear Kalman Filters such as the UKF [18] or the $S^2KF$ [19].

## VI. EVALUATION

The evaluation consists of tracking the object described in Fig. 6a using synthetic data. The target has a height of $1\,m$, and the scaling is $0.5\,m$ at the bottom and $0.1\,m$ at the top. The base shape is a circle, and the extrusion axis is assumed to lie in the $xy$-plane. The object is observed from above, i.e., the camera points in the direction of the negative $z$-axis. Thus, only the lateral surface was visible.
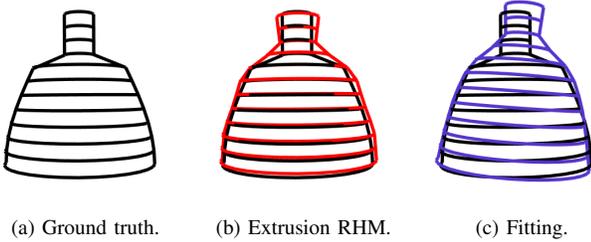


| (a) Ground truth. | (b) Extrusion RHM. | (c) Fitting. |

Figure 6: Ground truth and estimated shape.

The state has the form

$$\underline{x}_k = \left[\underline{x}_k^p, x_k^\alpha, \underline{x}_k^v, x_k^\omega, x_k^h, x_{k,0}^s, ..., x_{k,5}^s\right] \ ,$$

where $\underline{x}_k^p$ represents the position and $\underline{x}_k^v$ the velocity. The object can be rotated around the $z$-axis, with $x_k^\alpha$ determining the angle and $x_k^\omega$ the angular velocity. For both the position and the angle, a constant velocity model is assumed. The height is described with $x_k^h$. The lateral function is modeled as shown in Sec. V-B, using a piecewise linear function which split the lateral surface into six sections. Thus, each $x_{k,i}^s$ determines the scaling for each section.

For estimation, a PGF [16] with 120 samples is used, i.e., 8 samples per state dimension. The state is assumed to be Gaussian, i.e., the estimator treats the uncertain state as $p(\underline{x}_k) = \mathcal{N}(\underline{x}; \hat{\underline{x}}_k, \mathbf{C}_k^x)$. For the initial state at $k = 0$, the position $\underline{x}_k^p$ is $1\,m$ away from the ground truth, with rotation $x_k^\alpha = 0$, and with $x_{k,i}^s = 0.3\,m$ for all scalings. The initial velocities are zero. The initial uncertainty is $\mathbf{C}_0^x = 10^{-2} \cdot \mathbf{I}$.

At each time step, 20 measurement sources are generated from the boundary of the ground truth target. Each measurement

source is then corrupted using zero mean additive Gaussian noise. Of interest are situations with relatively high, non-isotropic noise, similar to that encountered using RGBD cameras. Because of this, the noise covariance matrix had the form $\mathbf{C}_k^v = \eta^2 \cdot \mathrm{diag}(1, 1, 5)$, rotated randomly.

This section is divided in two parts. First, the static scenario compares the new approach with direct *fitting*, and serves to illustrate the height problem. Then, the dynamic scenario evaluates the proposed approach using a moving target.

*A. Static Scenario*

In this scenario, we compared the introduced mechanism with the classical approach of fitting [15], i.e., minimizing the distance of the measurements to the shape. Both the fitting and the extrusion estimators used the same parameters. However, the fitting update assumed a shape model based on GAMs, i.e., only the measurement projection on the shape was considered as a source. The noise parameter was $\eta^2 = 2.5 \cdot 10^{-4}$, i.e., the largest error component had a standard deviation of about $3.5\,cm$. At each time step, a Gaussian zero mean process noise with covariance matrix $10^{-4} \cdot \mathbf{I}$ was assumed.

A snapshot for the extrusion estimator can be seen in Fig. 6b, and for fitting in Fig. 6c. A major problem with fitting is that, given that only the sides are visible, the height could not be estimated at all, which had a catastrophic effect on every other parameter. A regularization mechanism can be found using ideas of *Active Contours* [6], by shrinking it a small amount at each time step. However, finding suitable parameters is extremely difficult, given that they depend on factors such as the number of measurements, measurement quality, process noise, etc, all of which can change at any moment. A representative case which illustrates this issue can be found in Fig. 7a. Even if the shape parameters and the orientation were incorrect, the fitting approach was still able to correctly find the target position, as seen in Fig. 7b.
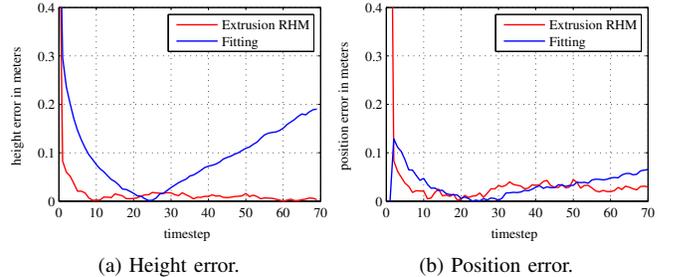


| (a) Height error. | (b) Position error. |

Figure 7: Results of the static scenario.

*B. Dynamic Scenario*

In the dynamic scenario, the target was moving in a circular arc, as shown in Fig. 8a. The position of the ground truth shape moved along a circle of radius $4\,m$, at a rate of $1$ degree each time step. Simultaneously, the true shape was rotating so that the extrusion axis always pointed to the center of the rotation circle. Note that, in consequence, the constant velocity model

(a) Position error.      (b) Height error.



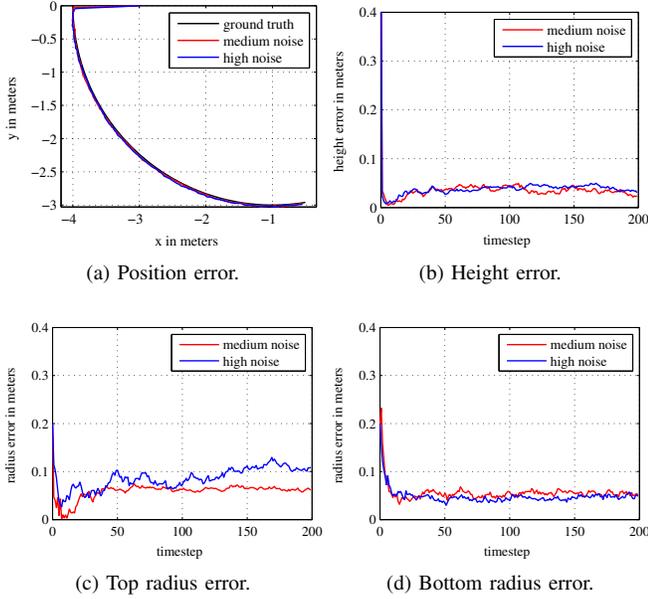(c) Top radius error.      (d) Bottom radius error.

Figure 8: Results of the dynamic scenario.

could approximate but not fully capture the circular motion. The noise parameters were $\eta^2 = 2.5 \cdot 10^{-4}$ for the middle noise runs, and $\eta^2 = 2.5 \cdot 10^{-3}$ for the high noise runs. In other words, the largest error components had a standard deviation of about $3.5$ cm and $11$ cm respectively. At each time step, a Gaussian zero mean process noise with covariance matrix $10^{-4} \cdot \mathbf{I}$ was assumed.

The mean results of 30 runs are shown in Fig. 8. The position was never further than $4$ cm away from the ground truth, as seen in Fig. 8a. The height error (Fig. 8b) remained about $3$ cm for both noise levels, indicating an issue of bias most likely caused by the incomplete motion model. Considering the ground truth height of $1$ m, those are still very good results. Given that the noise was relatively large in relation to the top radius of $0.1$ m, it is not very surprising that the results were not as stable there (Fig. 8c). A similar artifact can be seen in the top radius of Fig. 6b. For the bottom radius of $0.5$ m, a constant error of $5$ cm could be seen (Fig. 8d), likely due to a bias issue in a similar way to the height.

## VII. CONCLUSION

In this paper, we presented an approach to track extended objects by modeling them as scaled extrusions of a base shape. This allows the description of commonly used objects such as bottles, cans, containers, etc. To achieve this, we combined both probabilistic and projection-based techniques in order to exploit the advantages of both approaches, including robustness against occlusions and the ability to incorporate information about the distribution of measurement sources. An implementation using circles as base shapes was presented and evaluated using synthetic data. The results showed that the proposed approach

could easily track the target object even with high noise while also estimating its shape.

## REFERENCES

[1] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, no. 5, pp. 476–480, 1999.

[2] W. Gander, G. H. Golub, and R. Strebel, "Least-squares fitting of circles and ellipses," *Bulletin of the Belgian Mathematical Society Simon Stevin*, vol. 3, no. 5, pp. 63–84, 1996.

[3] M. Baum and U. D. Hanebeck, "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.

[4] K. Gilholm and D. Salmond, "Spatial distribution model for tracking extended objects," in *Radar, Sonar and Navigation, IEE Proceedings-*, vol. 152, no. 5. IET, 2005, pp. 364–371.

[5] M. Werman and D. Keren, "A Bayesian method for fitting parametric and nonparametric models to noisy data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 5, pp. 528–534, 2001.

[6] A. Blake and M. Isard, *Active Contours: The Application of Techniques from Graphics,Vision,Control Theory and Statistics to Visual Tracking of Shapes in Motion*, 1st ed. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1998.

[7] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Robotics-DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.

[8] S. Knoop, S. Vacek, and R. Dillmann, "Sensor fusion for 3d human body tracking with an articulated 3d body model," in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, May 2006, pp. 1686–1691.

[9] M. Baum and U. D. Hanebeck, "Random Hypersurface Models for Extended Object Tracking," in *Proceedings of the 9th IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2009)*, Ajman, United Arab Emirates, Dec. 2009.

[10] M. Baum, B. Noack, and U. D. Hanebeck, "Extended Object and Group Tracking with Elliptic Random Hypersurface Models," in *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*, Edinburgh, United Kingdom, Jul. 2010.

[11] 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, 2011.

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

[13] F. Faion, M. Baum, and U. D. Hanebeck, "Tracking 3D Shapes in Noisy Point Clouds with Random Hypersurface Models," in *Proceedings of the 15th International Conference on Information Fusion (Fusion 2012)*, Singapore, Jul. 2012.

[14] 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.

[15] P. Lancaster and K. Salkauskas, "Curve and surface fitting. an introduction," *London: Academic Press, 1986*, vol. 1, 1986.

[16] J. Steinbring and U. D. Hanebeck, "Progressive Gaussian Filtering Using Explicit Likelihoods," in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, Jul. 2014.

[17] M. F. Huber and U. D. Hanebeck, "Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation," in *Proceedings of the 17th IFAC World Congress (IFAC 2008)*, vol. 17, no. 2, Seoul, Republic of Korea, Jul. 2008.

[18] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," in *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*. IEEE, 2000, pp. 153–158.

[19] J. Steinbring and U. D. Hanebeck, "S2KF: The Smart Sampling Kalman Filter," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.