Tracking Simplified Shapes Using a Stochastic Boundary

Antonio Zea*, **Florian Faion***, **Marcus Baum**†, and **Uwe D. Hanebeck***
*Intelligent Sensor-Actuator-Systems Laboratory (ISAS)
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT), Germany
†University of Connecticut, USA
antonio.zea@kit.edu, florian.faion@kit.edu, baum@engineer.uconn.edu, uwe.hanebeck@ieee.org

***Abstract*—When tracking extended objects, it is often the case that the shape of the target cannot be fully observed due to issues of visibility, artifacts, or high noise, which can change with time. In these situations, it is a common approach to model targets as simpler shapes instead, such as ellipsoids or cylinders. However, these simplifications cause information loss from the original shape, which could be used to improve the estimation results. In this paper, we propose a way to recover information from these lost details in the form of a *stochastic boundary*, whose parameters can be dynamically estimated from received measurements. The benefits of this approach are evaluated by tracking an object using noisy, real-life RGBD data.**

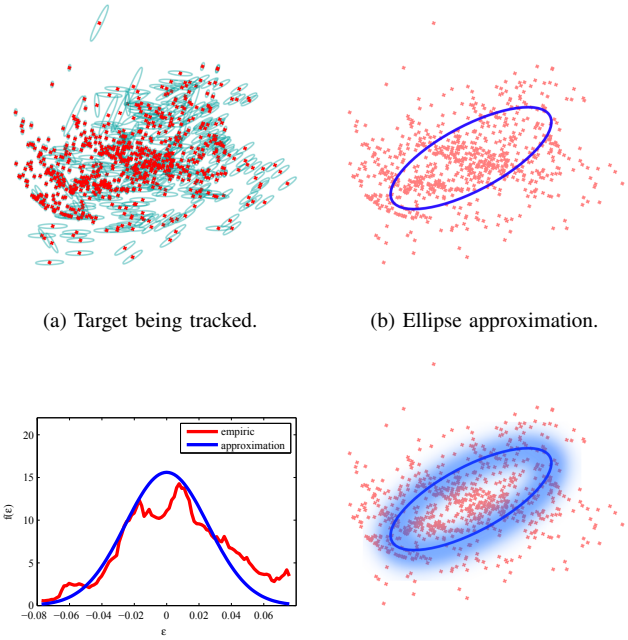
I. Introduction

When tracking an extended target based on real sensor data, the following two issues are encountered. First, measurements are distorted by *noise*. Depending on the sensor model, each measurement may have distinct noise characteristics, and assumptions of isotropy may not be justified, e.g., in RGBD cameras [1]. Second, issues such as high noise, artifacts, occlusions, and incomplete sensor models can make it impossible to find a model to appropriately describe the target (Fig. 1a), In addition, the characteristics of these issues may change with time. In these cases, instead of estimating a complex shape, a more robust approach is to approximate it as a simpler shape instead (Fig. 1b). However, this model simplification ignores information from the original shape.

In this paper we propose a Bayesian approach for tracking simplified shape objects that incorporates information from these lost details obtained from the estimation errors (Fig. 1c). For this purpose, we exploit the simplification uncertainty by using a *stochastic boundary* together with *Random Hypersurface Models* [2] (Fig. 1d). The parameters of this boundary can either be calculated a priori, allowing their easy incorporation into a tracker as additive noise, or estimated as multiplicative noise from received measurements.

Related work can be found, for instance, in curve fitting [3]. On the one hand, there are approaches using direct least squares distance minimization, particularly with ellipses [4], [5], which generally carry assumptions of isotropic noise. On the other hand, stochastic approaches such as [6]–[8] employ recursive Kalman filtering. The concept of a stochastic boundary is related to estimation bias, explored for conics in [7], [8].

This paper is structured as follows. A detailed problem formulation is given in Sec. II. Then, we explain the proposed approach in Sec. III. This is followed by the evaluation in Sec. IV. Finally, Sec. V concludes the paper.



(a) Target being tracked. (b) Ellipse approximation.

(c) Distance error histogram. (d) Stochastic boundary.

Figure 1: Target whose true shape cannot be modeled. Measurements in red, noise in cyan as 90%-confidence ellipses.

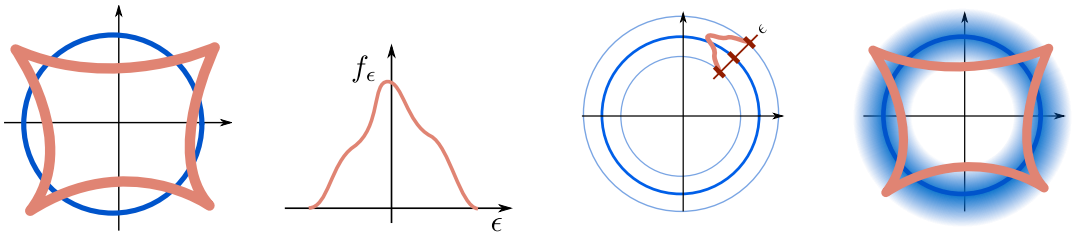
II. Problem Formulation

First, we need to define the state, the measurement model, and the shape model.

The state at the time step k is denoted as $\underline{x}_k$, and contains the shape parameters being estimated. However, the state may also contain further parameters, such as velocity or acceleration, depending on other models. In general, Bayesian estimators treat the state as the random vector $\underline{\boldsymbol{x}}_k$.

A. Shape Model

A shape $\mathcal{S}$ is an arbitrary compact set of points in $\mathbb{R}^n$. We distinguish between the true target shape $\mathcal{T}_k$, from which measurement sources are generated, and the estimated shape



(a) Shape simplification. (b) Distribution of $\underline{\boldsymbol{\epsilon}}_k$. (c) Level-sets from $\mathcal{L}(\underline{\boldsymbol{\epsilon}}_k)$. (d) Stochastic boundary.

Figure 2: Stochastic boundary for simplified shape (blue) derived from true target shape (red).

$\mathcal{S}(\underline{x}_k)$. Measurement sources $\underline{\tilde{z}}_k$ are assumed to be generated from $\mathcal{T}_k$ independently from each other.

B. Measurement Model

At time step k, the list of point measurements $\mathcal{Y}_k = \{\underline{y}_{k,0}, \ldots, \underline{y}_{k,m}\}$ from the target becomes available. Each measurement $\underline{y}_{k,i}$ is modeled as having been generated by the measurement source $\underline{\tilde{z}}_{k,i}$, both related through

$$\underline{y}_{k,i} = \underline{\tilde{z}}_{k,i} + \underline{v}_{k,i} \ ,$$

where $\underline{v}_{k,i}$ represents a Gaussian noise term introduced during observation, and is modeled as a realization of $\underline{\boldsymbol{v}}_{k,i} \sim \mathcal{N}(\underline{0}, \mathbf{C}^v_{k,i})$. For simplicity, the index i will be dropped unless needed.

III. Stochastic Boundary

The underlying problem considered in this paper is known as *curve fitting*. More specifically, tracking $\mathcal{T}_k$ consists of finding the parameters $\underline{x}_k$ that best fit the measurement equation

$$h_\tau(\underline{x}_k, \underline{y}_k, \underline{v}_k) = \underline{0} \ ,$$

which relates $\underline{x}_k$ to a received measurement $\underline{y}_k$. The measurement function h_τ is usually implemented in the form of a distance function, such as the Euclidian signed distance.

However, estimating $\underline{x}_k$ is made difficult by the fact that measurement quality can vary dramatically both in space and time, due to conditions of visibility, occlusions, artifacts, and noise. Fig. 1a shows an example of a such a target being tracked. In these cases, the true target shape $\mathcal{T}_k$ cannot be appropriately modeled.

Because of these issues, the usual approach is to track an approximation of $\mathcal{T}_k$ instead, using a less complex but more robust model. For example, in Fig. 1b, the target is tracked with a much simpler ellipse. The *simplified* model follows the new measurement equation

$$h_s(\underline{x}_k, \underline{y}_k, \underline{v}_k) = \underline{0} \ . \tag{1}$$

Because of the simplification, $\mathcal{S}(\underline{x}_k)$ will get close, but not necessarily converge to $\mathcal{T}_k$.

Note that this paper is not concerned with finding the best simplified model. Instead, given a simplified model, we propose an approach to improve estimation by recovering parts of the lost information, as described in the following sections.

A. Simplification Error

In this section, we consider the effect that the simplification has on measurement sources. Taking into account parameters such as sensor-to-object geometry, we model the distribution of possible measurement sources in $\mathcal{T}_k$ as the random variable $\underline{\tilde{\boldsymbol{z}}}_k$. In general, because of the aforementioned issues, the distribution of $\underline{\tilde{\boldsymbol{z}}}_k$ is not directly known, and changes with time.

In the absence of noise, i.e., assuming $\underline{v}_k = \underline{0}$, we can derive a shape function for $\mathcal{S}(\underline{x}_k)$ in the form of

$$g_s(\underline{x}_k, \underline{z}) := h_s(\underline{x}_k, \underline{z}, \underline{0}) \ .$$

In this way, it can be seen that for every $\underline{z} \in \mathbb{R}^n$ it holds that

$$\underline{z} \in \mathcal{S}(\underline{x}_k) \Leftrightarrow g_s(\underline{x}_k, \underline{z}) = \underline{0} \ .$$

Thus, the measurement equation from (1) is incomplete, as it holds for points in $\mathcal{S}(\underline{x}_k)$, but not in general for measurement sources in $\mathcal{T}_k$. Then, for a given $\underline{x}_k$, we define

$$\underline{\boldsymbol{\epsilon}}_k := g_s(\underline{x}_k, \underline{\tilde{\boldsymbol{z}}}_k) \ , \tag{2}$$

as the *simplification error*, representing the uncertainty introduced by the simplification. The concept of simplification error is shown in Fig. 2a, where a shape (in red) is being simplified using a circle (in blue). Fig. 2b shows the distribution of $\underline{\boldsymbol{\epsilon}}_k$ using the signed Euclidian distance, under the assumption that $\underline{\tilde{\boldsymbol{z}}}_k$ draws sources uniformly from $\mathcal{T}_k$.

A useful concept to visualize $\underline{\boldsymbol{\epsilon}}_k$ is the idea of *level-sets*. For this paper, we define the level-set $\mathcal{L}(\underline{c})$ of $\mathcal{S}(\underline{x}_k)$ as the region where g_s takes the value $\underline{c}$, i.e.,

$$\mathcal{L}(\underline{c}) := \{ \ \underline{z} \in \mathbb{R}^n \mid g_s(\underline{x}_k, \underline{z}) = \underline{c} \ \} \ .$$

It can be seen that $\mathcal{L}(\underline{0})$ corresponds to $\mathcal{S}(\underline{x}_k)$. For example, the level-sets of the simplified shape in Fig. 2a are also circles, as illustrated in Fig. 2c.

Combining level-sets with the simplification error leads to the random set $\mathcal{L}(\underline{\boldsymbol{\epsilon}}_k)$, which we denote as the *stochastic boundary*. Fig. 2d shows an example of a stochastic boundary, where the color intensity denotes the probability of the corresponding $\underline{\epsilon}_k$ being drawn.

Finally, we extend (1) using (2) to incorporate the stochastic boundary, yielding the new measurement equation

$$h_\epsilon(\underline{x}_k, \underline{y}_k, \underline{v}_k, \underline{\epsilon}_k) := h_s(\underline{x}_k, \underline{y}_k, \underline{v}_k) - \underline{\epsilon}_k = \underline{0} \ . \tag{3}$$

The simplification error is related to other model uncertainties in literature. However, in general, these deal with

discrepancies in the state $\underline{\boldsymbol{x}}_k$, such as bias-aware filters [9], or are concerned with shape uncertainty in function of the shape parameters [6]. Instead, the proposed concept deals with estimating the uncertainty in the results of the measurement function h_s, which can change in time.

B. Generative Model

The generative model for measurement sources from the true shape can be considered as a *spatial distribution* according to $\underline{\tilde{\boldsymbol{z}}}_k$, which is generally unknown. For a stochastic boundary, a generative model can be developed using *Random Hypersurface Models* (RHMs) [2], which describe how single measurement sources are generated from a target shape.

The key idea of an RHM is to interpret measurement sources as being generated by *transformations* of the boundary $\mathcal{S}(\underline{x}_k)$. This transformation is parameterized by the random vector $\underline{\boldsymbol{t}}_k$. In this way, a transformation factor $\underline{t}_k$ drawn from $\underline{\boldsymbol{t}}_k$ yields the transformed shape $\mathcal{S}_t(\underline{x}_k)$, and the probability of $\underline{t}_k$ determines how probable it is that a measurement source is in $\mathcal{S}_t(\underline{x}_k)$. How measurement sources are drawn from within $\mathcal{S}_t(\underline{x}_k)$ is arbitrary.

The stochastic boundary $\mathcal{L}(\underline{\boldsymbol{\epsilon}}_k)$ can be visualized as a generative model in the form of an RHM, where the transformation factor is drawn from $\underline{\boldsymbol{\epsilon}}_k$, and the transformed shape for a given $\underline{\epsilon}_k$ is $\mathcal{L}(\underline{\epsilon}_k)$. It can be seen that for any given $\underline{\tilde{z}}_k$ drawn from $\underline{\tilde{\boldsymbol{z}}}_k$, using (2) we can find a corresponding $\underline{\epsilon}_k$ drawn from $\underline{\boldsymbol{\epsilon}}_k$. Also, by definition, any point from $\mathcal{L}(\underline{\epsilon}_k)$ will return in (2) a value of $\underline{\epsilon}_k$. In fact, we can see that

$$g_s(\underline{x}_k, \underline{\tilde{\boldsymbol{z}}}_k) = g_s(\underline{x}_k, \mathcal{L}(\underline{\boldsymbol{\epsilon}}_k))$$

holds. Therefore, for a filter using (3), drawing sources from $\underline{\tilde{\boldsymbol{z}}}_k$ is equivalent to drawing from the stochastic boundary $\mathcal{L}(\underline{\boldsymbol{\epsilon}}_k)$. This can be visualized in Fig. 2d. In this way, propagating on h_s the sources drawn from the true target shape (red) yields the same values as the sources drawn from the level-sets of the simplified shape (faded blue circles). This shows that the proposed model can describe the target shape much more closely.

C. Estimating the Simplification Error

In some cases, the distribution parameters of $\underline{\boldsymbol{\epsilon}}_k$ can be calculated analytically, especially if $\underline{\tilde{\boldsymbol{z}}}_k$ can be appropriately approximated. However, $\underline{\tilde{\boldsymbol{z}}}_k$ can substantially change with time depending on characteristics such as sensor positioning and occlusion. In consequence, $\underline{\boldsymbol{\epsilon}}_k$ will change too, so that a-priori estimations of $\underline{\boldsymbol{\epsilon}}_k$ may be insufficient. In this case, a better approach is to estimate it from received measurements instead.

A way to achieve this is to approximate $\underline{\boldsymbol{\epsilon}}_k$ as Gaussian distributed. A simple parameterization has the form

$$\underline{\boldsymbol{\epsilon}}_k \approx \sigma_k \cdot \underline{\boldsymbol{\nu}} + \underline{\mu_k} \ , \tag{4}$$

where $\underline{\boldsymbol{\nu}} \sim \mathcal{N}(\underline{0}, \mathbf{I})$. In this way, the parameter $\underline{\mu_k}$ represents the mean, and σ_k the standard deviation. Finally, we incorporate σ_k and $\underline{\mu_k}$ as additional parameters into the state $\underline{x}_k$.

A problem with estimating $\underline{\mu_k}$ is that it conflicts with the extent parameters of $\mathcal{S}(\underline{x}_k)$, especially under inappropriate initialization. This can lead to pathological cases where the filter prioritizes fitting $\underline{\boldsymbol{\epsilon}}_k$ instead of correcting $\underline{\boldsymbol{x}}_k$. Thus, when estimating the target extent, $\underline{\mu_k}$ can be approximated as $\underline{0}$.

D. Incorporating the Stochastic Boundary

The stochastic boundary approach can easily be plugged into an estimator by modeling $\underline{\boldsymbol{\epsilon}}_k$ as an additional noise parameter. If the statistics of the stochastic boundary are known or have been estimated a priori, and are assumed not to change substantially, (3) shows that $\underline{\boldsymbol{\epsilon}}_k$ can be incorporated as an additive noise term. In particular, for estimators using measurement functions of the type $\underline{y}_k = f(\underline{x}_k) + \underline{v}_k$, (3) can be rewritten as

$$h_\epsilon(\underline{x}_k, \underline{y}_k, \underline{v}_k, \underline{\epsilon}_k) = \underline{y}_k - f(\underline{x}_k) - \underline{v}_k - \underline{\epsilon}_k = \underline{0} \ ,$$

and thus, incorporating a stochastic boundary is equivalent to simply using $\underline{\boldsymbol{v}}_k + \underline{\boldsymbol{\epsilon}}_k$ as measurement noise. In many cases, similar additive terms for $\underline{\boldsymbol{v}}_k$ can also be derived from $\underline{\boldsymbol{\epsilon}}_k$ for implicit measurement functions, depending on h_s.

Otherwise, if $\underline{\boldsymbol{\epsilon}}_k$ is unknown or changes with time, the approximation from (4) can be used. In this way, σ_k can be incorporated into the state $\underline{\boldsymbol{x}}_k$ and be estimated using multiplicative noise approaches such as [10].

IV. Evaluation

This section presents the evaluation of the approach using real-life RGBD data from a sensor network of eight Kinect devices. The idea was to evaluate the following three approaches:

1) without a stochastic boundary,
2) with a fixed, previously estimated $\underline{\boldsymbol{\epsilon}}_k$, and
3) where the parameters of $\underline{\boldsymbol{\epsilon}}_k$ are dynamically estimated.

The sensor model for Kinects proposed in [1] was used. The target was tracked using a recursive Bayesian estimator, which processed measurements sequentially. Specifically, the *S2KF* [11] was used, with 4 samples per dimension, in particular to deal with the multiplicative noise in Approach 3.

The following applied to all approaches. The state $\underline{\boldsymbol{x}}_k$ was approximated as Gaussian distributed, with initialization mean $\underline{x}_0 = \underline{0}$ and covariance matrix $\mathbf{C}_0^x = 10^{-4} \cdot \mathbf{I}$. Each time step lasted about 0.033 s and consisted of about 150 measurements, and a Gaussian process noise of $\underline{\boldsymbol{w}} \sim \mathcal{N}(\underline{0}, 10^{-6} \cdot \mathbf{I})$ was assumed at each time step.

A. Shape Model

The target shape was a person, and their sensor measurements were projected onto the XY-plane, i.e., the floor. The simplified shape $\mathcal{S}(\underline{x}_k)$ was an ellipse with parameters

$$\underline{\boldsymbol{x}}_k = \left[\underline{\boldsymbol{c}}_k, \underline{\boldsymbol{r}}_k, \boldsymbol{\theta}_k, \underline{\boldsymbol{v}}_k, \boldsymbol{\omega}_k, \boldsymbol{\sigma}_k\right]^T \ ,$$

where $\underline{\boldsymbol{r}}_k$ contains the axis lengths. The pose was determined by the translation $\underline{\boldsymbol{c}}_k$, rotation $\boldsymbol{\theta}_k$, with corresponding velocity parameters $\underline{\boldsymbol{v}}_k$ and $\boldsymbol{\omega}_k$. The simplification error was approximated as a Gaussian distribution as in (4), with $\boldsymbol{\sigma}_k$ as the standard deviation, and assuming $\mu_k = 0$. The shape function was implemented using the signed Euclidian distance $\mathrm{dist}_P(\underline{x}_k, \underline{z})$. The idea was to approximate the ellipse as a polygon, and then calculate the distance between a point $\underline{z}$ and the nearest point in this polygon [12]. The sign of dist_P is negative if $\underline{z}$ is inside the shape. This leads to the measurement equation

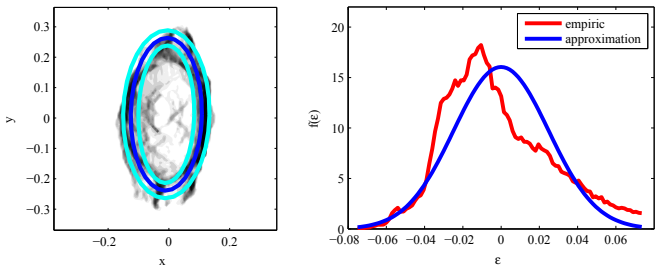
$$h_e(\underline{x}_k, \underline{y}_k, \underline{v}_k, \nu) = \mathrm{dist}_P(\underline{x}_k, \underline{y}_k - \underline{v}_k) - \sigma_k \cdot \nu \ ,$$

where $\boldsymbol{\nu} \sim \mathcal{N}(0, 1)$ is an additional noise term.

B. Results

The evaluation was divided into two scenarios. Scenario A was used to estimate basic parameters needed from the target, and Scenario B evaluated the three approaches.

1) Scenario A: In Scenario A, we estimated the shape parameters of four targets in a low-noise scenario, using two sensors at a distance of about 1.5 m from the target. In average, the estimated ellipse had the dimensions $\hat{\underline{r}}^m = [0.117\text{ m}, 0.251\text{ m}]$, and the simplification error had a standard deviation $\hat{\sigma}^m = 0.025$ m. Fig. 3a shows an example estimation, with the estimated ellipse in blue, and $\mathcal{L}(-\sigma_k)$ and $\mathcal{L}(\sigma_k)$ in cyan. Fig. 3b presents a histogram of $\underline{\epsilon}_k$. Units are in meters.

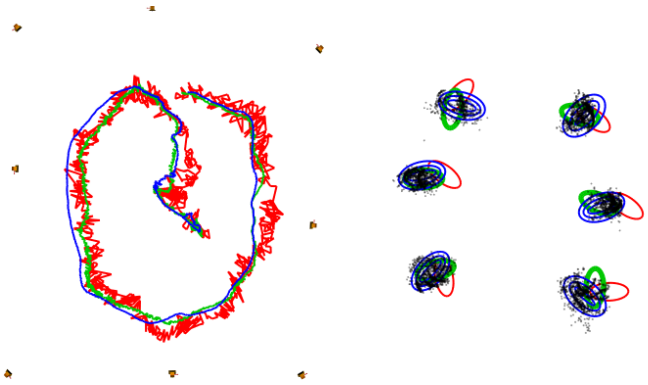


(a) Histogram and estimate. (b) Simplification error.

Figure 3: Example estimation for Scenario A.

2) Scenario B: Scenario B consisted of tracking a person moving in a room of size 4.5 m × 5.5 m, with eight cameras (yellow) positioned as shown in Fig. 4a. Measurements were highly noisy and contained clutter, given that Kinects only support distances up to about 2 m. For visualization, contrast Fig. 3a at 1.5 m to the same target in Fig. 1a at about 3 m. Due to the narrow field of view, in many places the target was only seen by the camera in the opposite side of the room.

Because of these issues, all three approaches used a fixed size of $\underline{r}_k = \hat{\underline{r}}^m$. For Approach 2, the stochastic error standard deviation was fixed to $\sigma_k = \hat{\sigma}^m$. Both $\hat{\underline{r}}^m$ and $\hat{\sigma}^m$ were taken from Scenario A. The scenario spanned a time of about 20 s.



(a) Estimated centers. (b) Example snapshots.

Figure 4: Experiment setup and results for Scenario B.

Fig. 4a shows the results of the estimated centers, with Approach 1 in red, Approach 2 in green, and Approach 3 in blue. In Fig. 4b, selected snapshots of the estimators can be seen, with measurements in black. For Approach 2 and Approach 3, $\mathcal{L}(-\sigma_k)$ and $\mathcal{L}(\sigma_k)$ are also pictured. It can be seen that, on the one hand, Approach 1 mostly failed to follow the center and rotation. On the other hand, Approach 2 and Approach 3 were both much more robust. However, given that $\underline{\epsilon}_k$ changed substantially with time, Approach 2 had issues following the rotation, especially in parts (like Fig. 4b left) where there were issues of visibility and occlusion, causing the estimated shape to spuriously rotate. Approach 3 also had issues due to the low measurement quality, at times reaching $\hat{\sigma}_k \approx 0.06$ m. However, it still showed much better results at the cost of estimating one single additional parameter.

V. Conclusion

In this paper, we proposed an approach to improve estimation results when using a simplified approximation of a shape. The key idea was to exploit the simplification error, a measure of the uncertainty introduced by the approximation. Using this approach, we developed a generative model for the simplified shape in the form of a stochastic boundary. We showed that the simplification error could be easily incorporated into an estimator as a noise parameter, and estimated dynamically if necessary. Our evaluation also demonstrated how this approach leads to more robust results.

References

[1] F. Faion, S. Friedberger, A. Zea, and U. D. Hanebeck, "Intelligent Sensor-Scheduling for Multi-Kinect-Tracking," in *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2012)*, Vilamoura, Algarve, Portugal, Oct. 2012, pp. 3993–3999.

[2] 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.

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

[4] 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.

[5] 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.

[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] Z. Zhang, "Parameter estimation techniques: a tutorial with application to conic fitting," *Image and Vision Computing*, vol. 15, no. 1, pp. 59 – 76, 1997.

[8] 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.

[9] J.-P. Drecourt, H. Madsen, and D. Rosbjerg, "Bias aware kalman filters: Comparison and improvements," *Advances in Water Resources*, vol. 29, no. 5, pp. 707 – 718, 2006.

[10] M. Baum, F. Faion, and U. D. Hanebeck, "Modeling the Target Extent with Multiplicative Noise," in *Proceedings of the 15th International Conference on Information Fusion (Fusion 2012)*, Singapore, Jul. 2012.

[11] 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, 2013.

[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.