Bayesian Estimation of Line Segments

Florian Faion\(^*\), Antonio Zea\(^*\), Marcus Baum\(^\dagger\), and Uwe D. Hanebeck\(^*\)

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

\(^\dagger\)University of Connecticut, USA

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

Abstract—A popular approach when tracking extended objects with elongated shapes, such as ships or airplanes, is to approximate them as a line segment. Despite its simple shape, the distribution of measurement sources on a line segment can be characterized in many radically different ways. The spectrum ranges from Spatial Distribution Models that assume a distinct probability for each individual source, to Greedy Association Models as used in curve fitting, which do not assume any distribution at all. In between these border cases, Random Hypersurface Models assume a distribution over subsets of all sources. In this paper, we compare Bayesian estimators based on these different models. We point out their advantages and disadvantages and evaluate their performance by means of illustrative examples with synthetic and real data using a Linear Regression Kalman Filter.

I. INTRODUCTION

In this work, we consider the task of tracking extended objects with elongated shapes by approximating them as line segments. In contrast to a simple point model, a line segment model allows for additionally estimating parameters such as orientation and length naturally. Using this rough approximation is a reasonable approach when dealing with few measurements and high sensor noise. Furthermore, line segments can also be embedded into extrusion models for more complex 3D shapes [1], when measurement quality and quantity increases. Hence, application areas range from aerial and naval surveillance, over robotics and telepresence, up to consumer entertainment systems. In this work, our intention is to compare three different strategies of modeling line segments [3–7] and to investigate their specific properties. In doing so, we show that these strategies essentially differ in their assumptions on the distribution of the underlying measurement sources.

Incorporating knowledge about the source distribution may have a strong impact on the estimation result, in both directions. As an example, let us consider the task of tracking a pen based on a Microsoft Kinect (see Fig. 1), where 3D point measurements are segmented according to their brightness in the IR-image. Usually, these measurements originate from sources that are uniformly distributed along the pen and incorporating this knowledge improves the estimation result. However, due to segmentation errors, parts of the pen are occasionally invisible to the sensor, which makes the distribution of measurement sources hard to predict. In these cases, it may be better to omit the concept of source distribution at all, as shown in [8]. Note that special considerations on clutter [4], [5] and kinematics [9] are out of the scope of this work. Other approaches to extended object tracking, dealing with different shapes and their source distributions, are proposed in [10–13].

This paper is structured as follows. After a brief problem statement in Sec. II, we introduce the different modeling strategies in Sec. III and their implementation in Sec. IV. In Sec. V we compare the different models and evaluate their usability in a synthetic estimation task and in the tracking scenario from Fig. 1. We conclude in Sec. VI with a summary.

II. PROBLEM STATEMENT

Given: At each time step \(k\), a sensor measures the elongated object as a list of 2D or 3D noisy points \(Y_k = \{y_{i,k}\}_{i = 1,\ldots,n_k}\), where the potentially time-varying number \(n_k\) of measurements is considered to not contain any information about the object extent. Rather, measurements are assumed to originate, statistically independently from another, from unknown sources \(\tilde{z}_{i,k} \in \tilde{Z}(x_k)\) on a line segment, whose pose and extent is specified by the state vector \(x_k\). We assume additive sensor noise according to

\[
y_{i,k} = \tilde{z}_{i,k} + w_{i,k},
\]

(a) Experimental setup
(b) Pen in detail view
(c) Snippet of IR-image

Figure 1: Tracking a pen as a line segment with a Kinect.
where the noise variable $w_{i,k}$ is drawn from the Gaussian distribution $\mathcal{N}(0, C_{w_{i,k}})$, and assumed to be independent from state and measurement sources.

**Desired:** Then, the task is to continuously determine the parameters $\xi_k$, which correspond to the line segment that most likely has generated the measurements $\mathcal{Y}_k$. In particular, we are interested in the position vector $x_k$ and considering each measurement individually. For clarity, time indices $k$ and measurement indices $i$ will be omitted whenever possible.

**Separation of Shape and Pose:** As indicated in Fig. 2, we distinguish between two coordinate frames: the object frame (black) that represents the object without any pose information, i.e., unrotated and centered on the origin, and the world frame (gray) that represents how the object is seen from the outside. Unless otherwise stated, all geometric considerations are assumed to refer to the object frame.

### III. BAYESIAN LINE SEGMENT ESTIMATION

In this section, we derive a Bayesian estimator for the line segment in order to demonstrate the different modeling strategies. The knowledge of the state is modeled by a prior probability distribution $p(x)$. Then, tracking consists of two alternating steps. First, the prediction step lets the distribution $p(x)$ evolve over time according to a system model. Second, the measurement update step incorporates new measurement points $\mathcal{Y}$ according to Bayes’ rule

$$p(x|\mathcal{Y}) \propto p(\mathcal{Y}|x) \cdot p(x) ,$$

where $p(\mathcal{Y}|x)$ is the likelihood of measuring $\mathcal{Y}$. The statistical independence between the measurements allows for separating the likelihood

$$p(\mathcal{Y}|x) = \prod_i p(y_i|x) ,$$

and considering each measurement individually.

If the originating source $\tilde{z}(x)$ for a measurement $y$ and a state $x$ was known, the likelihood would immediately follow from the additive noise model (1) according to

$$p(y|x) = \mathcal{N}(y; \tilde{z}(x), C_w) .$$

However, due to the noise, the true sources are generally unknown and cannot be recovered exactly from the measurements, which is known as the association problem.

We distinguish between three strategies to deal with this problem by deriving approximate likelihoods. For this purpose, we introduce a continuous index parameter $s \in [-1, 1]$ that iterates through all possible sources $\tilde{z}(x, s) \in \tilde{Z}(x)$ according to

$$\tilde{z}(x, s) = [s \cdot r, 0]^T \quad (2)$$

for the 2D case, and $\tilde{z}(x, s) = [s \cdot r, 0, 0]^T$ for the 3D case, respectively. As a reminder, $r$ was defined to be half the length of the line segment, so that $s = -1$ and $s = 1$ refer to the left and right edges, respectively. Note that $s$ defines an individual likelihood for each source by $p(y|\tilde{z}(x, s)) = \mathcal{N}(y; \tilde{z}(x, s), C_w)$ as illustrated in Fig. 3.

**A. Spatial Distribution Model**

The first strategy to deal with the association problem is to define a probability distribution $p(s)$ that specifies the probability of each source $\tilde{z}(x, s)$ to be measured. Then, the likelihood is obtained by marginalizing over all source hypotheses

$$p(y|\tilde{z}) = \int_{-1}^1 \mathcal{N}(y; \tilde{z}(x, s), C_w) \cdot p(s) \, ds . \quad (3)$$

This model [3–6] is known as Spatial Distribution Model (SDM). As an example, Fig. 5(a,c,e) illustrate different distributions $p(s)$ for the sources from Fig. 4a that can be modeled in an SDM. As can be seen, each source hypothesis $\tilde{z}(x, s)$ is considered to be the true measurement source and weighted according to its probability. This individual treatment is indicated by using different colors for each source in Fig. 4a. However, especially in real world applications, it may happen that knowledge about the source distribution is not available or non-trivial to calculate [14]. This raises the need for approaches that depend less on $p(s)$.

**B. Greedy Association Model**

The second strategy is to drop any prior assumption upon the distribution $p(s)$ and greedily associate the measurement to the “best” of all sources on the line segment. In [8], it was shown that “best” in this context generally means that the source yields the highest individual likelihood

$$p(y|\tilde{z}) = \max_{s \in [-1, 1]} \mathcal{N}(y; \tilde{z}(x, s), C_w) . \quad (4)$$

Figure 2: Given: Measurements $\mathcal{Y}_k$ of unknown sources $\tilde{Z}(x_k)$ on the line segment at time $k$. Desired: state parameters $\xi_k$ of the line segment.

Figure 3: Individual likelihoods, dependent on $s$. As an example, Fig. 5(a,c,e) illustrate different distributions $p(s)$ for the sources from Fig. 4a that can be modeled in an SDM.
This *Greedy Association Model* (GAM) is popular in curve-fitting [15] and can be derived from an SDM by imposing special assumptions, as shown in [8]. In contrast to the SDM, the GAM imposes that a specific source must have generated the measurement exclusively and drops all other hypotheses. All sources are connected and drawn in the same color in Fig. 4c, in order to indicate that they are not treated individually. Note that the GAM from (4) does not offer any way to incorporate knowledge about the distribution of measurement sources. However, this ignorance turns the task of estimating the length of a line segment into an ill-posed problem [3] as will be discussed in Sec. V.

C. Random Hypersurface Model

The third strategy combines ideas from the SDM and the GAM in order to arrive at a model, known as *Random Hypersurface Model* (RHM) [7], [16] that overcomes their inherent drawbacks. This combination consist of assuming a probability distribution over partitions of the shape to describe how likely it is that a measurement originates from a source within the partition. Then, within these partitions, a greedy association is performed to select the source which yields the highest individual likelihood.

In the case of a line segment, we proposed [7] to partition the shape into pairs of symmetric sources \( \tilde{z}(x, -s^*) \) and \( \tilde{z}(x, s^*) \), with \( s^* \in [0, 1] \) being the index parameter that iterates through all partitions. Fig. 4b schematically shows the partitions by coloring and connecting lines. Over all of the partitions a probability distribution \( p(s^*) \) is defined, as qualitatively shown in Fig. 5(b,d,f). Within the partitions, the left or right source is greedily associated to the measurement, depending on which one yields the higher likelihood. Formally, the RHM likelihood then can be calculated according to

\[
p(y|x) = \int_{s^*}^1 \max_{s \in \{-s^*, s^*\}} (N(y; \tilde{z}(x, s), C_w)) \cdot p(s^*) \, ds^*
\]

In contrast to the SDM, the integration range of the RHM is reduced by a factor two. This becomes an advantage in sample-based estimators, as it doubles the effective spatial sample resolution [7]. However, this increased resolution is achieved at the cost of coarsening the source distribution, as we will illustrate in Sec. V.

IV. IMPLEMENTATION

In this section, we show how to implement Bayesian estimators for line segments based on the SDM (3), GAM (4), and RHM (5) from the previous section. In doing so, we assume the state parameters \( \tilde{z} \) encode the “radius” \( r \), orientation matrix \( R \), and position \( \ell \) of the line segment.

In order to derive an approximate measurement update, we use *statistical linearization*, which was already successfully applied to the estimation of line segments based on an SDM [6] and RHM [7]. For this purpose, we need to specify a generative measurement function in the form of \( y = h(x, w) \) for each of the models.

A. Measurement Equation (SDM)

The SDM from (3) directly translates to the measurement equation

\[
y = h(x, w, s) = R \cdot \tilde{z}(x, s) + \ell + w.
\]

As measurements \( y \) are generally given in world coordinates, the sources (being defined in object coordinates (2)) have to be converted according to \( R \cdot \tilde{z}(x, s) + \ell \). In (6), \( w \) is the additive measurement noise and \( s \) is a multiplicative noise variable whose characteristics describes the occurrence of measurement sources along the line segment [3–6].

B. Measurement Equation (GAM)

For the GAM, deriving an explicit generative measurement model would yield

\[
y = R \cdot \tilde{z}(x, \hat{s}) + \ell + w,
\]

which requires finding the particular index \( \hat{s} \) that refers to the source \( \tilde{z}(x, \hat{s}) \) with the maximum individual likelihood (4). In [7], it was pointed out that \( \tilde{z}(x, \hat{s}) \) is often approximated by the point on the line segment that is closest to the measurement \( y \) in terms of an appropriate distance measure. However, this calculation depends on the current measurement \( y \) such that the right hand side of (7) is not longer independent of the measurement.

In order to resolve this dependence, let us first rearrange the explicit generative measurement model (7) to the implicit expression

\[
\bar{y} = R \cdot \tilde{z}(x, \hat{s}) + \ell + w - y
\]

Note that \( \bar{z} := R^T \cdot (y - w - \ell) \) can be understood as the measurement minus the noise, converted to object coordinates. Based on (8), we can define an implicit measurement equation in the form of \( \bar{y} = h(x, y, w) \), where \( \bar{y} \) is interpreted as a constant pseudo-measurement, and the original measurement acts as a function parameter. Finally, approximating \( \tilde{z}(x, \hat{s}) \) in (8) with the point on the line segment that is closest to the measurement yields for a line segment in 2D

\[
\bar{y} = h(x, y, w) = \begin{cases} [0, z_2]^T & \text{if } |z_1| < r \\ [||z_1|-r, z_2]^T & \text{otherwise} \end{cases}
\]
The 3D version simply requires adding $z_3$ as another dimension. From a geometric view, the generative model (9) produces $0$-valued differences between $(\frac{y - w}{v - w})$ and their closest point on the line segment, which is given by $\bar{z}$. In [7], we showed that these vector valued-differences are often approximated by by scalar-valued distances by taking the Euclidean norm. There are two important aspects to note. First, the measurement equation in (9) is independent of $s$, such that there is no way to incorporate the source distribution in the generative model. And second, $w$ now is a non-additive noise variable.

C. Measurement Equation (RHM)

Similar to the GAM, we can define an implicit generative model for the RHM [7], e.g., in 2D, according to

$$\Omega = h(\bar{z}, \bar{y}, w, s^*) = |z_1|, |z_2|^T - \bar{z}(\bar{w}, s^*) \quad (10)$$

Again, the 3D version simply requires adding $z_3$ as another dimension. In contrast to (9), which produces $0$-valued differences for $\bar{z}$ lying on the entire line segment, (10) produces $0$ only for $\bar{z}$ being equal to $\bar{z}(\bar{w}, s^*)$ or $\bar{z}(\bar{w}, -s^*)$ for a given instance of the index $s^* \in [0, 1]$. Note that in (10), both $w$ and $s$ are non-additive noise variables.

D. Measurement Update Using Statistical Linearization

Based on (6), (9), and (10) approximate measurement updates can be derived as follows. Preliminary, we require knowledge of the probability distributions of all involved random variables: That is for all models, state $p(\bar{z})$ and measurement noise $p(w)$, and, $p(s)$ for the SDM and $p(s^*)$ for the RHM. Then, the mean $\mu_x$ and covariance matrix $C_x$ of the measurement equation $h$, as well as the cross-covariance $C_{xh}$ between $\bar{z}$ and $h$ need to be calculated, e.g., based on deterministic sampling [17], [18]. Finally, the updated mean $\mu_x^e$ and covariance matrix $C_x^e$ can be calculated using the well-known Kalman formulas. For the SDM, we obtain

$$\mu_x^e = \mu_x + K(p - \mu_h)$$
$$C_x^e = C_x - KC_hK^T$$

with $K = C_{xh}C_h^{-1}$ being the Kalman gain. Due to the pseudo measurement $\Omega$, the calculation of the mean slightly differs for the GAM and the RHM

$$\mu_x^e = \mu_x + K(0 - \mu_h) \quad .$$

V. Comparison

In this section, we point out specific properties of the presented modeling strategies and evaluate their usability in a real-life tracking problem. We consider $p(\bar{z})$ to be Gaussian in all experiments, and employ the sample-based filter from [18] with five samples per dimension. In order to set up an SDM-estimator, two requirements have to be fulfilled.

Requirement 1 (Quadratic Extension)

In [6], it was shown that linear estimators are not suitable for estimating the length of the line segment based on the SDM from (6), as the statistical linearization decorrelates the measurement from the length. This can be intuitively seen by deriving the entries of the cross-covariance matrix between the expected measurement $h$ and radius parameter $r$, which all are $0$. For resolving this issue, a quadratic extension that augments the measurement $[y_1; y_2; y_3]^T$ was proposed in [6].

Requirement 2 (Sigma Points)

Further, it was found in [7] that even with quadratic extension, an SDM-estimator requires a specific type of samples. While the Unscented Kalman Filter [17] does not have an effect on the length, the Smart Sampling Kalman Filter [18] allows for length estimation.

When using a GAM, another specific extension is required.

Requirement 3 (Regularization)

As already mentioned, length estimation based on the GAM of a line segment is ill-posed. This issue was discussed in [3] and is due to the fact that a GAM-estimator does not penalize overestimated lengths, as only the distances from the measured points to their closest sources on the line segment are minimized. For illustration, consider a scenario with, e.g., two measured points. From the viewpoint of a GAM-estimator it does not matter; how long the estimated segment is, as long it intersects with both points. Inspired by the ideas of active contours [3], [19], regularization by minimizing the internal energy can be incorporated into the estimator, in order to resolve this issue. Regularization can be implemented by a process model that contracts the length of the segment in each prediction step. In contrast to the GAM-estimator, SDM- and RHM-estimators both penalize overestimated lengths as they assume measurement sources originate from the line segment according to a specific distribution.

A. Distribution of Sources

As the three approaches essentially differ in their assumptions on the source distributions, we evaluate the effect of varying the distribution. Specifically, we compare the performance
of an SDM-, RHM-, and GAM-estimator when measurement sources \(r \cdot s\) with \(s \in [−1, 1]\) originate along the line segment according to

- a uniform distribution \(p(s) = \frac{1}{2}\),
- a truncated Gaussian distribution \(p(s) \propto \mathcal{N}(s; 0, \frac{1}{2})\), and
- a triangle distribution \(p(s) = \frac{1}{2} \cdot (s + 1)\).

It is interesting to note that the uniform and the Gaussian distribution are symmetric with respect to the line segment, while the triangle distribution is not.

By design, the SDM-estimator directly allows for incorporating these distributions, as qualitatively depicted in Fig. 5(a,c,e). In contrast, the RHM aggregates the probabilities of symmetric sources, such that the uniform and triangle distribution yield identical distributions \(p(s^*)\), as qualitatively shown in Fig. 5(b,f). In the GAM, the underlying distributions are completely ignored.

In order to evaluate the estimation performance when using different models, we simulated a static line segment in 2D with length 1, centered in \((1, 2)\), and rotated by the angle \(\pi/4\). Then, 500 measurements were drawn from each source distribution and distorted with additive Gaussian noise according to (1), with \(C_w = 10^{-2} \cdot I\). Based on these measurements, 2D position, rotation angle, and the length of the line segment were estimated. We performed 100 Monte-Carlo runs of the experiment, initialized with covariance matrix \(C_{x_0} = I\), and mean \(x_0\) as the ground truth with an additive distortion drawn from \(\mathcal{N}(0, C_{s_0})\).

Result: Fig. 6 shows the root mean squared error RMSE of the estimated length when using the different models. For the uniform and Gaussian distribution, the RHM- and the SDM-estimator both converge to the correct value, where the RHM is slightly faster (see Fig. 6a and Fig. 6b). In contrast, the GAM-estimator converges slowly to a biased length, which depends on the regularization parameter and the source distribution, making the task of adjusting parameters non-trivial. When looking at the triangle distribution in Fig. 6c, the SDM-estimator shows a fast convergence. As before, and for the same reason, the GAM converges to a biased length. Interestingly, the RHM also converges to a biased value, as it cannot distinguish between the uniform and the triangle distribution.

As the position is highly correlated with the length, the graphs for their RMSE qualitatively coincide with the length. In case of length convergence, the value for the position RMSE is about 0.02 for the uniform and the Gaussian distribution and about 0.08 for the triangle distribution. In contrast, the angle parameter seems to be at most slightly affected by the length, as all estimators produce comparable RMSEs of about 1.25° for all distributions. Tab. I gives a qualitative summary of the presented results. Ticks mean that the model works for the specific source distribution, or is robust against inaccuracies in the source distribution, as will be investigated in Sec. V-B. For the GAM-estimator, we additionally put the ticks into brackets to indicate that finding parameters for the regularization is non-trivial.

<table>
<thead>
<tr>
<th></th>
<th>requirements</th>
<th>uniform</th>
<th>Gaussian</th>
<th>triangle</th>
<th>robust</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDM</td>
<td>1.2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>RHM</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GAM</td>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table I: Summary of the considered modeling approaches.

B. Tracking a Pen

Next, we consider the experiment from Fig. 1a, where the pose (position and orientation) of a pen should be tracked with a Microsoft Kinect, while drawing a path on a WACOM tablet. Specifically, a stream of registered IR and depth images with 640 × 480 pixels was captured by the Kinect, and then, 3D point clouds were extracted. For segmentation, the pen body was covered by a reflective sheet (see Fig. 1b), which could be easily extracted from the IR image (Fig. 1c). At each time step, 40 point measurements from the segmented pen were randomly selected and assigned a covariance matrix according to the sensor model from [20]. The pen is known to have a total length of 14 cm and the distance from Kinect to tablet plane is about 110 cm, where the tablet has dimensions of 30.5 cm × 23.5 cm. For the estimation task, we considered a uniform source distribution along the pen and the length to be known. Thus, we estimated the 3D position, as well as the pitch and yaw angle (roll angle of a line segment is not observable). In addition, we used a constant velocity model for position and angles, requiring them to be estimated as well.

Result: Occasionally, reflection artifacts and overexposure in the IR image caused incomplete point clouds, which violated the modeled source distribution. Fig. 7, where the intersection of the estimated line segment with the tablet plane is drawn against the ground truth path (colored in black). The RHM shows the smallest estimation error, especially in the horizontal parts. However, the distance from the pen tip to the intersection with the tablet plane (which should be 0 cm), the GAM performs best with an avg. error of 0.3 cm in contrast to \(~0.7\) cm of the SDM and RHM. That is, the GAM is robust against the varying source distribution, which is caused by the miss-segmentation. It happens that half of the pen is missing due to overexposure, causing the SDM- and RHM-estimators to adjust the position along the pen line.

VI. Conclusion

In this work, we compared three strategies to derive a source model, which can be used in a Bayesian estimator for the pose and length of an elongated object by approximating it as a line segment. We found that accurately estimating the length requires assumptions on the source distribution. Then, given a distribution that is symmetric with respect to the line segment, the RHM converges slightly faster than the SDM. A second advantage is that the RHM does not have any requirements to the filter being used. However, in cases of more complex distributions, as for example an asymmetric triangular distribution, only the SDM-estimator converges to the true length. Using a GAM-estimator might be useful in situations where the length is known, but the source
distribution is unknown, e.g., due to miss-segmentation in the pen tracking experiment.

REFERENCES


