On Sensor Model Design Choices for Humanoid Robot Localization

Stefan Tasse, Matthias Hofmann, and Oliver Urbann

Robotics Research Institute
Section Information Technology
TU Dortmund University
44221 Dortmund, Germany

Abstract. The development of estimation systems based on Kalman filters requires several design choices. Among others, these are the methods used for linearization, coordinate systems for measurement representations, and approximations such as how to handle multiple simultaneous observations per time step. This paper evaluates these different choices with respect to their influence on the system's estimation quality and points out simple yet effective solutions. Camera-based localization for a humanoid robot is chosen as an example application and the localization benefit of different approaches is evaluated using real and simulated feature perceptions.

1 Introduction

Localization is essential for mobile robots. When facing the task of designing a localization algorithm for an autonomous robot, one may pick from a vast number of different approaches. While there exist many different strategies such as multi-angulation methods [1] or constraint based localization [2], most algorithms follow the concepts of recursive Bayesian filtering [3]. The two main representatives of this category are Kalman and particle filters. Gutmann and Fox state the common impression that "Markov localization is more robust than Kalman filtering while the latter can be more accurate than the former" in [4]. An additional argument for particle filtering is the easy representation of multimodal belief states. However, Gaussian mixtures allow the same for Kalman filters, and recently such multiple model Kalman filters have been applied with great success even on robot platforms with very limited resources [5,6], allowing superior localization quality and robustness to comparable particle filters, but at lower computational costs in case of [6].

Kalman filters have been applied in many tasks and are covered extensively in literature. Designing a Kalman filter for localization is therefore not a significant challenge. However, this filter will often not perform to its full potential. The implementation process presents several design choices, some of which are discussed frequently, while others are generally neglected or only mentioned briefly. Specialized references such as [7] as well as the most standard books [3,8] leave the impression that the most important decision is whether to address the system's

non-linearity by Taylor series expansion such as in the Extended Kalman filter (EKF) or by use of the Unscented transform as in the Unscented Kalman filter (UKF). This will be addressed only briefly in Section 2. Other influences will be discussed in the course of this paper in Section 3 and 4. Those may seem trivial at first and their importance not obvious, but their significant influence on the outcome will be shown for the example of localization for a humanoid robot. As such, it is this paper's main contribution to point out simple design choices which will lead to significant localization quality improvement with minimal effort.

2 Addressing Non-Linearity by Taylor Series Expansion or Unscented Transform

The Kalman filter in its original form is optimal for systems which fulfill a number of assumptions, such as only involvement of zero mean Gaussian noise and a known and linear system to model. This is rarely given for practical applications, since most systems of interest are non-linear in one aspect or another. The Kalman concept is popular and successful nonetheless, which is due to the possibility to linearize the non-linear models around the current estimate. This provides a decent enough approximation to allow the tracking.

In general, two different concepts are commonly used: the Extended and the Unscented Kalman filter. The Extended Kalman filter employs a Taylor series for linearization, which in effects means to simply substitute Jacobi matrices for the linear transformations in the original Kalman filter equations. This method is and has been widely used for the last four decades. See [3] for further details. Of course the linearization may result in different approximation qualities of the uncertainty propagation, depending on each individual use case. Further limitations of this approach arise in cases of discontinuous systems or such with singularities. Additionally, it is often perceived by developers that "calculating Jacobian matrices can be a very difficult and error-prone process" [7] due to the manual derivation of the Jacobians and possible translation errors in their subsequent conversion to code.

The Unscented Kalman filter offers a different approach to estimate the different expectations necessary to apply the standard Kalman equations, namely the predictions of state and observation and the cross-covariance between the two. This is done by deterministically sampling the state space around the current mean and covariance, applying the non-linear transformation to those sigma points, and then recovering the transformed mean and covariance. Significant improvements of applying the unscented transformation compared to analytical linearization have been shown in [7].

Those findings have led to the impression that choosing an Unscented Kalman filter instead of an Extended Kalman filter will be a major source of improvement in most systems, and that this will be the main design choice in developing an estimation system for a given application. In the course of this paper we will show that much easier alterations may have much bigger effects.

3 Measurement Coordinate System Choices

To illustrate the effect of measurement coordinate system choices, we assume as an example application the problem of estimating the localization of a humanoid robot, so the state to be estimated is the robot's pose $x = (p_x, p_y, p_\theta)^T$. The robot perceives point features on the ground around it, e.g. by means of processing images recorded by one or several cameras mounted in its head. Each point feature corresponds to a landmark with known global position $l = (l_x l_y)^T$. Those expected and actual perceptions, $\overline{z}$ and z respectively, can by expressed in different coordinate systems, each of which may be used to formulate the sensor model of the Kalman filter.

In the following, let

$$\Omega(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \tag{1}$$

be the rotation around α and (l_x, l_y) the global coordinates of a known landmark which is part of the robot's map of the environment. The time index t is omitted in all following equations for the sake of simplicity.

3.1 Measurements in Cartesian Coordinates

As the localization problem is expressed as an orientation and a position in global Cartesian coordinates, a first intuitive choice is to express a measurement on the ground around the robot in robot-centric Cartesian coordinates as shown in Figure 1. The sensor model to calculate the expected measurement $\overline{z} = (z_x, z_y)^T$

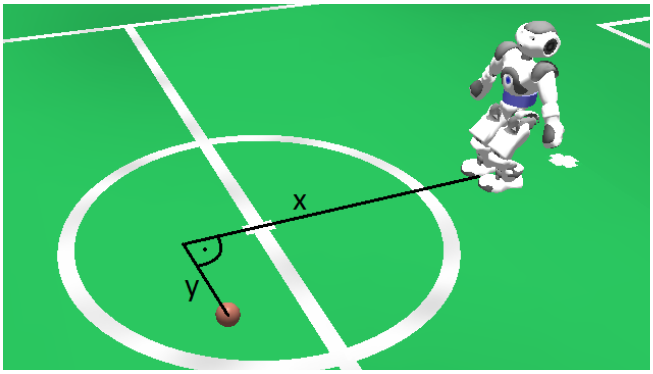

Fig. 1: Observation given in euclidean coordinates.

for the current robot pose $x = (p_x, p_y, p_\theta)^T$ and a correspondence to the land-

mark with known global position $l = (l_x l_y)^T$ is then given in Equation 2 and the corresponding Jacobi matrix in Equation 3.

$$\overline{z} = \begin{pmatrix} z_x \\ z_y \end{pmatrix} = h(x,l) = \Omega(-p_\theta) \cdot \left[\begin{pmatrix} l_x \\ l_y \end{pmatrix} - \begin{pmatrix} p_x \\ p_y \end{pmatrix} \right] \tag{2}$$

$$\begin{aligned} H = \frac{\partial h(x,l)}{\partial x} &= \begin{pmatrix} \frac{\partial z_x}{\partial p_x} & \frac{\partial z_x}{\partial p_y} & \frac{\partial z_x}{\partial p_\theta} \\ \frac{\partial z_y}{\partial p_x} & \frac{\partial z_y}{\partial p_y} & \frac{\partial z_y}{\partial p_\theta} \end{pmatrix} \\ &= \begin{pmatrix} -\cos p_\theta & -\sin p_\theta & -(l_x - p_x)\sin p_\theta + (l_y - p_y)\cos p_\theta \\ sinp_\theta & -\cos p_\theta & -(l_x - p_x)\cos p_\theta - (l_y - p_y)\sin p_\theta \end{pmatrix} \end{aligned} \tag{3}$$

3.2 Measurements in Cylindrical Coordinates

Measurements can also be expressed in cylindrical coordinates, i.e. range and bearing, to indicate the distance and direction of the observed feature (cf. Figure 2). This is often the first choice of those familiar with laser scanners or

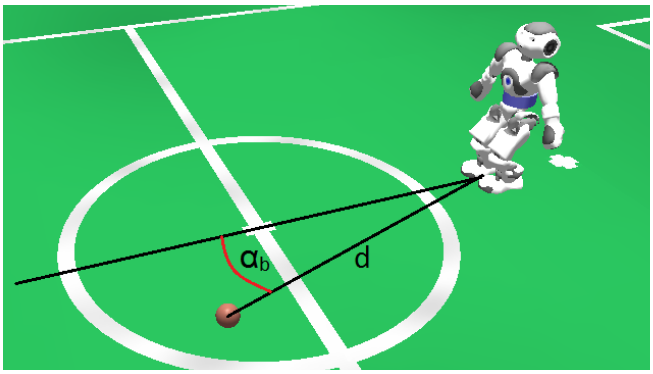

Fig. 2: Observation given as range and bearing.

developers of robot-centric path planning algorithms. In this case, the Sensor model function and Jacobi matrix are given by Equation 4 and 5, respectively, with the abbreviation $d^2 = (l_x - p_x)^2 + (l_y - p_y)^2$.

$$\overline{z} = \begin{pmatrix} z_r \\ z_b \end{pmatrix} = h(x,l) = \begin{pmatrix} \sqrt{(l_x - p_x)^2 + (l_y - p_y)^2} \\ \text{atan2}(l_y - p_y, l_x - p_x) - p_\theta \end{pmatrix} \tag{4}$$

$$\begin{aligned} H = \frac{\partial h(x,l)}{\partial x} &= \begin{pmatrix} \frac{\partial z_r}{\partial p_x} & \frac{\partial z_r}{\partial p_y} & \frac{\partial z_r}{\partial p_\theta} \\ \frac{\partial z_b}{\partial p_x} & \frac{\partial z_b}{\partial p_y} & \frac{\partial z_b}{\partial p_\theta} \end{pmatrix} \\ &= \begin{pmatrix} (-l_x + p_x)d^{-1} & (-l_y + p_y)d^{-1} & 0 \\ (l_y - p_y)d^{-2} & (-l_x + p_x)d^{-2} & -1 \end{pmatrix} \end{aligned} \tag{5}$$

3.3 Measurements in Spherical Coordinates

A third coordinate system choice is given by using the vertical and horizontal angles α_1 and α_2 as indicated in Figure 3. While the meaning of the vertical

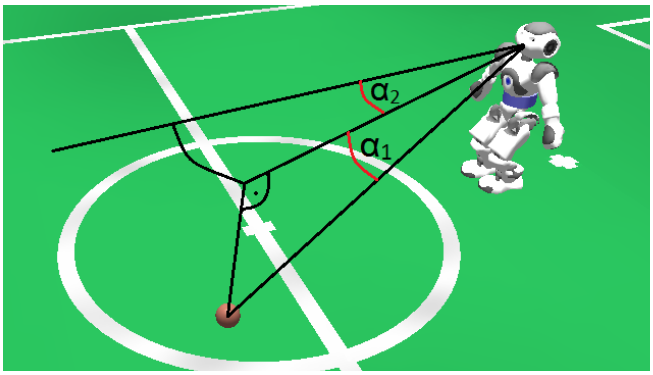

Fig. 3: Observation given in angular coordinates.

angle may not be intuitive for any direct further use, this is the coordinate system which is closest to the actual perception process in this example. With the same abbreviation of $d^2 = (l_x - p_x)^2 + (l_y - p_y)^2$ as used above and the height of the camera h_{camera}, the sensor model function and Jacobi matrix are given in Equation 6 and 7.

$$\begin{aligned} \overline{z} &= \begin{pmatrix} z_{\alpha_1} \\ z_{\alpha_2} \end{pmatrix} = h(x, l, h_{camera}) \\ &= \begin{pmatrix} \text{atan2}(h_{camera}, \sqrt{(l_x - p_x)^2 + (l_y - p_y)^2}) \\ \text{atan2}(l_y - p_y, l_x - p_x) - p_\theta \end{pmatrix} \end{aligned} \tag{6}$$

$$\begin{aligned} H = \frac{\partial h(x,l)}{\partial x} &= \begin{pmatrix} \frac{\partial z_{\alpha_1}}{\partial p_x} & \frac{\partial z_{\alpha_1}}{\partial p_y} & \frac{\partial z_{\alpha_1}}{\partial p_\theta} \\ \frac{\partial z_{\alpha_2}}{\partial p_x} & \frac{\partial z_{\alpha_2}}{\partial p_y} & \frac{\partial z_{\alpha_2}}{\partial p_\theta} \end{pmatrix} \\ &= \begin{pmatrix} \frac{h_{camera}(l_x - p_x)}{d(h_{camera}^2 + d^2)} & \frac{h_{camera}(l_y - p_y)}{d(h_{camera}^2 + d^2)} & 0 \\ (l_y - p_y)d^{-2} & (-l_x + p_x)d^{-2} & -1 \end{pmatrix} \end{aligned} \tag{7}$$

3.4 Experimental Comparison

Two experiments are set up to compare the effects of the sensor model design choices described so far.

Simulated Perception A simulation is set up to test the correctness of the implementation and the conformity with related work's results. Localization algorithms are used with the above mentioned different linearization and coordinate system choices and parametrized using fixed measurement covariances, which were chosen to be optimal for each approach separately. This simulation assumes a humanoid robot with noisy odometry and a perception process which measures randomly distributed landmarks with unique correspondences and contains errors mainly from the cameras unknown orientation, i.e. the errors originate from normally distributed noise in the spherical coordinate system. As expected, Figure 4 shows the localization to be best using this spherical representation. Furthermore, the classical example of transforming between spherical/cylindrical and Cartesian coordinates is handled much better by the UKF than by the EKF as predicted for example by [7].

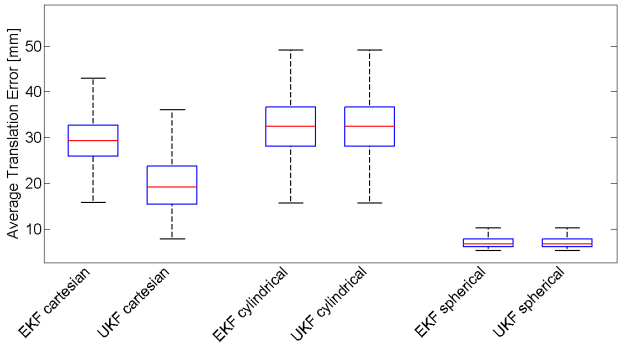

Fig. 4: Comparison between localization quality using different linearization approaches and sensor model coordinate systems with simulated perceptions.

Real Perception Process on the Nao To evaluate the impact on a real system, observations are recorded using a Nao, a humanoid robot which is 58 cm tall and equipped with two cameras in its head with non-overlapping fields of view. The environment is a robot soccer field as used in the Standard Platform League. Any ambiguous observations are associated with maximum likelihood correspondences based on the true robot position. The perception process also produces sporadic false positives. Those sets of observations and correspondences together with artificially generated odometry errors serve as input for all different configurations, whereas each generated set is processed by all approaches so that the random component in the input presents no source of bias. Note that these localization results will not diverge due to the usual problem of wrong correspondence choices once the position estimate contains a certain error, as this

experiment is set up to test the sensor models, not the correctness of correspondence choices. The odometry errors contain white noise and a drift component, as this is the usual behavior of real Nao robots which are worn out or even heated up slightly asymmetrically.

An important factor for each algorithm is its parametrization. All different approaches in this experiment use the same motion update and the same process noise, which is chosen to be a certain amount above the artificially generated white noise component to compensate the drift. In common applications the measurement noise magnitudes are normally constants which are part of the parametrization and subject to a tuning process by the developer. Here, they are optimized separately for the approaches using a randomly picked measurement subset which is not used for the following evaluation afterwards. Thus each approach is performing with the parametrization which empirically provides the least squared localization errors.

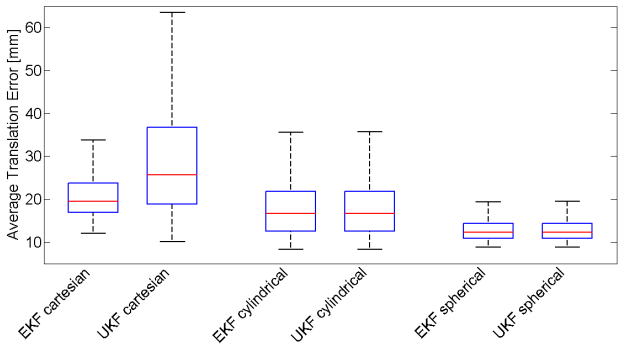

Fig. 5: Comparison between localization quality using different linearization approaches and sensor model coordinate systems with real observations recorded on a Nao.

Figure 5 shows the distribution of the sums of localization errors for 1000 sets of measurements and odometry errors. It can be seen that the effect of the coordinate system choice is in general more significant than the distinction between Extended or Unscented Kalman filter. These real world results mostly verify the tendency of the assumptions in the simulated experiments, but also show discrepancies for example in the results using Cartesian coordinates. This implies that the underlying process is not fully described by assuming only normally distributed angular errors in the camera orientation. Expressing the measurement in spherical coordinates, which is intuitively the closest to the underlying process of perception, still clearly outperforms the other coordinate systems' sensor models. To use these results as a basis for development recommendations, the

EKF/UKF choice is clearly second to the angular coordinate representation of the robot's measurements.

3.5 Hybrid Modifications

The empirical results above raise the question if already implemented systems which did not use the spherical coordinate system for the sensor model design can still make use of this information. One possibility is to adapt the measurement noise covariance matrix to better reflect the properties of the perception process, e.g. to scale the uncertainty depending on the distance of the observed feature. This has not led to any significant improvements in case of the Cartesian representation, for which more complex modification would be necessary to adapt it to reflect the spherical coordinate system's properties. The cylindrical representation however offers an easy improvement.

Both the cylindrical and the spherical coordinate system already share the horizontal angle; they only differ in distance against vertical angle. Applying the knowledge that errors in the distance mainly result from variations in said vertical angle, it is possible to derive an appropriate scaling factor β for the distance measurement's uncertainty.

$$z_r = \frac{h_{camera}}{\sin z_{\alpha_1}} \tag{8}$$

$$\frac{\partial z_r}{\partial z_{\alpha_1}} = -\frac{h_{camera}}{\sin^2 z_{\alpha_1}} \cdot \cos z_{\alpha_1} \tag{9}$$

$$\beta \propto \frac{h_{camera}}{\sin^2 \operatorname{atan2}(h_{camera}, z_r)} \cdot \cos \operatorname{atan2}(h_{camera}, z_r) \tag{10}$$

Equation 8 gives the relation between range observation z_r and vertical angle z_{α_1} and Equation 9 denotes their partial derivative. Therefore, using the first rows of Equation 6 and 4, the scaling factor β in Equation 10 can be derived. Scaling the (newly tuned) expected range error with β or the corresponding entry in the measurement covariance matrix with β^2 results in the hybrid localization approach with cylindrical coordinates and distance scaled measurement covariance in Figure 6 and 7. While this one presents a significant improvement over the cylindrical coordinates with constant measurement covariance and comes close to the approach in spherical coordinates, the latter one is still superior.

4 Multiple Simultaneous Measurements

Practical Kalman implementations rarely go by the theory of one motion update and one sensor update per time step. Instead there are usually many time steps in which no observation is made, so the sensor update is omitted. In other time steps, several observations are made at once, i.e. several different features are detected in the same time step. The common implementation is usually to

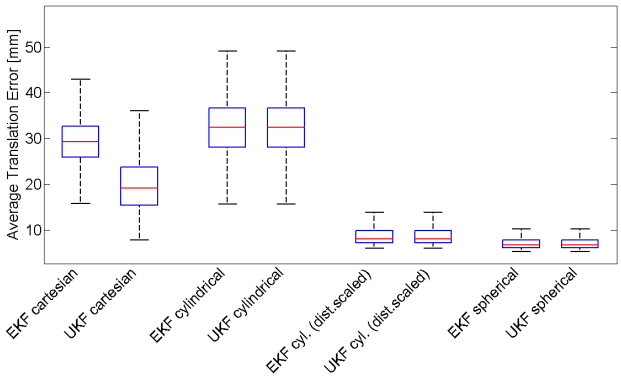

Fig. 6: Comparison between localization quality using different linearization approaches and sensor model coordinate systems with simulated perceptions.

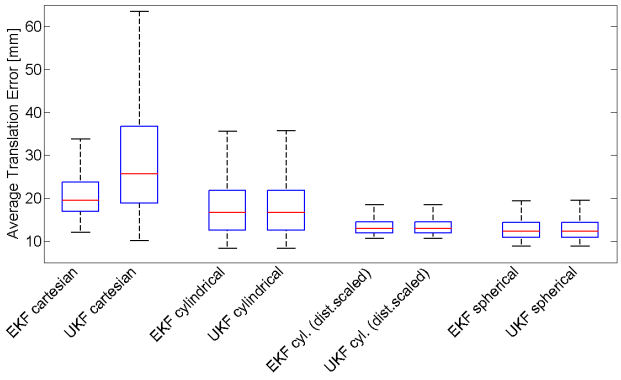

Fig. 7: Comparison between localization quality using different linearization approaches and sensor model coordinate systems with real observations recorded on a Nao.

execute several consecutive sensor updates. The quality of this approximation however depends on the perception process by which those features have been observed.

Now consider 2-dimensional feature observations as described above, each with a separate measurement covariance as in Equation 11.

$$C = \begin{pmatrix} s_1^2 & 0 \\ 0 & s_2^2 \end{pmatrix} \tag{11}$$

The stochastically correct sensor update for n detected features would be to execute a single $2n$-dimensional measurement update instead of n separate 2-dimensional updates. In case the different measurements are stochastically independent of each other, i.e. all off-diagonal inter-feature entries of the $2n \times 2n$ measurement covariance are zero, then a single $2n$-dimensional measurement update is approximated well by n 2-dimensional updates.

If multiple measurements originate from the same perception source and are correlated, then this simple approximation neglects potentially useful information and consequently looses in approximation quality. Taking a humanoid robot with camera based perception again as in Section 3.4, multiple observations originate from the processing of a single camera image and it stands to reason that the main source of measurement error is the inaccurately estimated camera orientation due to the walking motion. Such simultaneous measurements would therefore contain nearly the same angular errors. Assuming a spherical coordinate representation for the measurements as described in Section 3.3, the resulting covariance for 2 simultaneous observations is given in Equation 12 with γ close to 1, while $\gamma = 0$ would neglect any dependence between both observations.

$$C' = \begin{pmatrix} C & \gamma C \\ \gamma C & C \end{pmatrix} \tag{12}$$

Figure 8 shows evaluations with simulated test runs consisting exclusively of multiple observations per time step, and illustrates the differences in localization quality for iterative execution of 2-dimensional sensor updates, for $2n$-dimensional updates which neglect the covariance (i.e. with $\gamma = 0$), and for $2n$-dimensional updates with full covariances as in Equation 12. All sensor updates in this example utilize spherical coordinate representations for the observations. As expected, the multiple 2-dimensional updates are an appropriate approximation as long as the separate measurements are independent. When observations are correlated, then significant benefits can be drawn from the information encoded in the full covariance matrix. Note that some Unscented Kalman filter implementations may become unstable for a γ too close to 1, as C' will still be a valid covariance matrix and therefore positive semi-definite, but very close to not being positive definite any more, which will cause the frequently used Cholesky decomposition to become numerically instable.

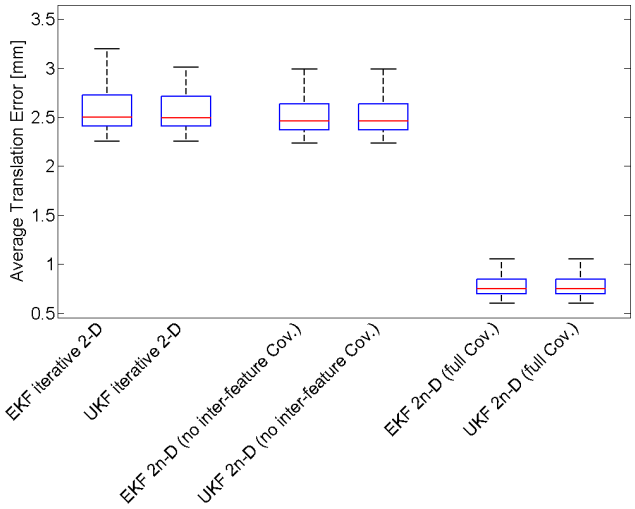

Fig. 8: Comparison between different methods to handle multiple measurements at one time step.

5 Conclusion

This paper gives an overview about common design choices which researchers face when developing localization algorithms based on Kalman filters. The most prominent choice between the Extended and Unscented Kalman filter is widely discussed in common literature, but this is by far overrated, which has been illustrated in the previous sections using the example application of camera-based humanoid robot localization. The choice of the measurement's coordinate system representation is mostly disregarded in most publications as well as in common educational books, but provides a simple way to improve localization quality. The same holds for the correct handling of simultaneous observations originating from processing the same camera image.

As such, this paper provides the means for a better understanding of different approximations' impacts when applying Kalman filters, and presents simple guidelines to develop optimal solutions for localization and tracking systems.

References

1. Betke, M., Gurvits, L.: Mobile robot localization using landmarks. Robotics and Automation, IEEE Transactions on **13**(2) (April 1997) 251 –263
2. Ghring, D., Mellmann, H., Burkhard, H.D.: Constraint based world modeling in mobile robotics. In: Proc. IEEE International Conference on Robotics and Automation ICRA 2009. (2009) 2538–2543
3. Thrun, S., Burgard, W., Fox, D.: Probabilistic Robotics (Intelligent Robotics and Autonomous Agents). The MIT Press (2005)

4. Gutmann, J.S., Fox, D.: An Experimental Comparison of Localization Methods continued. In: Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on. Volume 1. (2002) 454 – 459 vol.1
5. Quinlan, M., Middleton, R.: Multiple Model Kalman Filters: A Localization Technique for RoboCup Soccer. In Baltes, J., Lagoudakis, M., Naruse, T., Ghidary, S., eds.: RoboCup 2009: Robot Soccer World Cup XIII. Volume 5949 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg 276–287
6. Jochmann, G., Kerner, S., Tasse, S., Urbann, O.: Efficient multi-hypotheses unscented kalman filtering for robust localization. In Röfer, T., Mayer, N.M., Savage, J., Saranli, U., eds.: RoboCup 2011: Robot Soccer World Cup XV. Lecture Notes in Computer Science. Springer Berlin / Heidelberg (2012) to appear
7. Julier, S., Uhlmann, J.: Unscented filtering and nonlinear estimation. Proceedings of the IEEE **92**(3) (March 2004) 401 – 422
8. Siegwart, R., Nourbakhsh, I.: Introduction to autonomous mobile robots. Intelligent Robotics and Autonomous Agents. The MIT Press (2004)