Robust Localization Using Context-Aware Filtering

Radoslav Ivanov*, Nikolay Atanasov†, Miroslav Pajic†, Insup Lee* and George Pappas†
*Department of Computer and Information Science
†Department of Electrical and Systems Engineering
University of Pennsylvania
{rivanov, atanasov, pajic, lee, pappasg}@seas.upenn.edu

***Abstract*—In this paper we develop a robot localization technique that incorporates discrete *context measurements*, in addition to standard continuous state measurements. Context measurements provide binary information about detected events in the robot's environment, e.g., a building is recognized using image processing or a known radio station is received. Such measurements can only be detected from certain positions and can, therefore, be correlated with the robot's state. We investigate two specific examples where context measurements are especially useful – an urban localization scenario where GPS measurements are not reliable as well as the capture of the RQ-170 Sentinel drone in Iran, where GPS measurements were spoofed. By focusing on a specific class of probability of context detection functions, we derive a closed-form Gaussian mixture filter that is precise, captures context, and has the theoretical properties of the Kalman filter. Finally, we provide simulations of the urban localization scenario with an unmanned ground vehicle and show that the proposed context-aware filter is more robust and more accurate than the conventional extended Kalman filter, which only uses continuous measurements.**

I. Introduction

Precise localization, i.e., the problem of estimating a robot's position and orientation relative to a given map, is an essential task for any autonomous robot. It is a necessary condition for the accomplishment of many goals, including navigation, mapping and human-robot interaction. Consequently, as part of the Simultaneous Localization and Mapping (SLAM) problem, it has received a lot of attention in the academic community.

Localization is made challenging by several factors. In particular, available sensors are often insufficiently precise for achieving the required level of localization accuracy [13]. In addition, accurate motion models are difficult to derive and even when they exist, they may be highly nonlinear and discontinuous, thus making it only possible to develop approximate and suboptimal algorithms [9]. Finally, data association, i.e., determining the correspondence between objects on the map and received sensor measurements, introduces additional uncertainty when moving through space [15].

In this paper, we address the first and (part of) third problems by incorporating *context measurements*. We define context in a robot's environment as additional information that is not encoded in the robot's state but may be correlated with it. Examples of context measurements include recognition of nearby known buildings (e.g., using image processing), detection of radio signals, detection of sensor faults, etc. What these measurements have in common is that they are all binary, i.e., either 0 or 1, and they have a known probability (or one that can be learned) of occurring given the system state.

More specifically, we focus on two applications that are used as motivating and running examples. Localization and lane following in urban environments with high-rise buildings is known to be a difficult problem [13] due to the imprecision of GPS [12]. Thus, we propose to include context measurements of recognized nearby buildings (with known location on the map) as a way of improving current localization techniques.

The second example is the RQ-170 Sentinel drone that was captured in Iran [20, 23]. While the details of the capture are not publicly available, it is widely believed that, through jamming the GPS signal, hijackers were able to mislead the drone into believing it is landing on friendly territory. At the same time, however, the drone was in a position to receive Iranian frequency modulation (FM) radio signals; incorporating that context measurement in its localization algorithm may have avoided the drastic mislocalization.

While it is common to use contextual cues for object localization [4], such information is used less frequently to improve robot localization. Previous works in context-aware filtering in robotics use scene categorization [6, 18, 21] and appearance [5, 27] as well as object class information [1, 2, 3, 7, 22] to improve localization and mapping. One of the first systems that uses both spatial and context representations was proposed by Galindo et al. [6]. A spatial hierarchy contains camera images, local metric maps, and the environment topology, while a semantic hierarchy represents concepts and relations which allow room categories to be inferred based on object detections. Paletta et al. [19] utilize visual context in images to improve localization performance in the presence of occlusions. Yi et al. [28, 29] use semantic descriptions of distance and bearing in a contextual map for active localization.

Most existing algorithms cannot handle both continuous and discrete measurements as they result in nonlinearities and

This material is based on research sponsored by DARPA under agreement number FA8750-12-2-0247. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA or the U.S. Government.

This research was supported in part by Global Research Laboratory Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning (2013K1A1A2A02078326).

This work was supported by TerraSwarm, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA.

discontinuities that make precise analysis challenging. One way of addressing this problem is by incorporating discrete (object class) observations in the metric localization via a set-based particle filter that handles missed detections, false alarms, and data association [2]. Its drawback compared to our approach is that it uses particles to represent the uncertainty in the robot's location and may have particle depletion problems in high-dimensional spaces. In contrast, Gaussian distributions are better at representing high dimensional states, yet existing techniques for Guassian filtering and smoothing cannot handle discrete observations. For example, SLAM++ [22] takes advantage of the Guassian distribution of spatial variables (e.g., object poses) but does not include object detections (e.g., discrete class information) in the optimization. Alternatively, Bao et al. [3] develop a detailed motion model incorporating camera parameters, object geometry, and object classes, which however results in large and expensive Monte Carlo optimization. Finally, context-aware filtering is similar to target tracking in that measurements arrive at irregular intervals and from different objects [14]; while general models exist in this setting, they lead to very computationally expensive calculations and distributions that cannot be computed in closed form.

Instead, our filter uses Gaussian Mixtures to represent high-dimensional states and is derived in closed form (without any approximation for linear models), hence it is precise and fast. More specifically, we focus on filtering with a class of probability of detection functions that capture the underlying physical properties of context measurements in the applications described above. These functions are defined as scaled Gaussian probability density functions, which means that the probability of getting a detection (e.g., of a radio signal) is high when the robot is close to the object (e.g., beacon) and decreases quickly when the robot moves away.

Using these probability functions, we utilize a Bayesian filter and show that the resulting probability distribution of the robot's state is a Gaussian Mixture, i.e., thus resembling a conventional Kalman filter. Returning to the motivating application, we show that our filter performs well at the problem of localization in an urban environment. Through simulation of a ground vehicle equipped with a GPS and a camera, we show that the context-aware filter performs much better than the extended Kalman filter (EKF) in the presence of big variance and bias in GPS measurements.

The contributions of this work are: (1) formalizing context measurements and their incorporation in robot observation models; (2) developing a closed-form context-aware filter with Gaussian Mixtures that is fast and precise and has the theoretical properties of the Kalman filter; (3) illustrating the advantages of the context-aware filter over the EKF through simulation of the urban environment localization scenario.

This paper is organized as follows. Section II formulates the problem considered in this paper, and Section III presents our solution. In Section IV we provide the case study, while Section V contains a discussion on generalizing our approach to other domains. Finally, Section VI concludes the work.

II. Problem Formulation

Consider a robot with known nonlinear discrete-time dynamics of the form

$$x_{k+1} = f(x_k, u_k) + w_k, \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state (e.g., position, orientation, velocity and acceleration), $u_k \in \mathbb{R}^p$ is the applied input, and $w_k \sim \mathcal{N}(0, Q)$ is Gaussian process noise. The model can also be specified in terms of the probability density function (pdf) of the state x_{k+1}, conditioned on the previous state and input:

$$p_f(x_{k+1} \mid x_k, u_k). \tag{2}$$

The system has two kinds of sensors available to it: state and context. State sensors measure (subsets of) the state directly (e.g., an odometer measures speed). We assume that there is a known observation model for the state sensors of the form

$$y_k = h(x_k) + v_k, \tag{3}$$

where we denote state sensors' measurements by $y_k \in \mathbb{R}^m$ and $v_k \sim \mathcal{N}(0, R)$ is Gaussian measurement noise.

Context sensors, on the other hand, do not measure the robot's state but rather provide binary information about its context. Context can be represented as a finite set $\mathcal{C} = \{c_1, \ldots, c_N\}$, where each c_i is a context element that can be detected by a context sensor from certain system states; example context elements include a building that can be detected through image processing or an FM radio signal that can only be detected within a certain radius of a beacon. For each i, a measurement b_k^i is received that is equal to 1 if c_i is detected and 0 otherwise.[1] We do not assume knowledge about a context element, other than the probability of detection given a state, denoted by $p_d(c_i \mid x_k)$, i.e.,

$$b_k^i = \begin{cases} 1 & w.p. \quad p_d(c_i \mid x_k) \\ 0 & w.p. \quad 1 - p_d(c_i \mid x_k), \end{cases} \tag{4}$$

where p_d is a function of the robot state. Thus, the set of context measurements is $b_k \in \{0,1\}^N$. We assume that, conditioned on the state, context measurements are mutually independent and are independent from state measurements.

Problem. *Given the robot model defined in* (1)-(4) *and a prior pdf* $p_{k|k}(x) := p(x_k \mid u_{0:k-1}, y_{0:k}, b_{0:k})$*, the goal is to compute the posterior pdf*

$$p_{k+1|k+1}(x) := p(x_{k+1} \mid u_{0:k}, y_{0:k+1}, b_{0:k+1})$$

describing the robot's state given all available measurements and inputs.

[1] As part of future work, we will also consider the case when b_{k+1} provides a class measurement, i.e., it can take on values in a finite set of possible classes, e.g., building types.

III. Approach

The robot model described in Section II naturally lends itself to a Bayesian approach, i.e., one that can be divided into a predict and update stage of the form

$$\textbf{Predict: } p_{k+1|k}(x) = \int p_f(x \mid z, u_k) p_{k|k}(z) dz,$$
$$\textbf{Update: } p_{k+1|k+1}(x) = \eta_{k+1} p_o(y_{k+1}, b_{k+1} \mid x) p_{k+1|k}(x),$$

where $p_o(y_{k+1}, b_{k+1} \mid x_{k+1})$ is the joint pdf of all available measurements (state and context) given the state and η_{k+1} is a normalization constant [26].

Computing the densities in the Bayesian filter exactly, however, is not possible in general, with the notable exception being the linear Gaussian case, which leads to the Kalman filter. This difficulty is usually overcome with an EKF or other linear regression Kalman filters [9] by computing linear Taylor series approximations, and using the approximation in the ordinary Kalman filter. It is important to highlight that an EKF cannot be directly applied to the problem considered in this work because the context measurement observation model is discontinuous and cannot be differentiated.

In this paper, rather than using a method to approximate any observation model, we focus on a particular class of probability of detection functions (i.e., $p_d(c_i \mid x_k)$ in (4)).

Assumption. *We consider inverse-exponential functions that are defined as scaled Gaussian pdf's of the form*

$$p_d(c_i \mid x_k) = e^{-\frac{1}{2}(G_i x_k - \theta_i)^T V_i^{-1} (G_i x_k - \theta_i)}, \tag{5}$$

which are parameterized by $\theta_i \in \mathbb{R}^q$ and $V_i \in \mathbb{R}^{q \times q}$, the analogues of a mean and a covariance in a multivariate Gaussian pdf, and $G_i \in \mathbb{R}^{q \times n}$, which can be thought of as a selection matrix when $q < n$. This probability is 1 when $G_i x_k = \theta_i$ and approaches 0 when $G_i x_k - \theta_i$ gets large.[2]

We argue that this class of detection probabilities captures a lot of different scenarios observed in reality. To return to our running examples, the probability of detecting a building decreases quickly as the distance from the camera to the building is increased [2]. In addition, it may be easier to detect a building from some angles than it is from others, and this may also be captured by the above formulation. Similarly, the probability of detecting a radio FM signal also decreases as the receiver gets farther away from the transmitter [8].

Having fixed (5) as the form of the probability of context detection, with θ_i and V_i known (or potentially learned from data) for each c_i, we now show how to derive a closed-form localization filter. Note that the result below holds for linear systems; for nonlinear systems such as the one considered in this paper, linearization techniques can be applied such as a linear regression Kalman filter or an EKF as follows:

$$F_k = \left.\frac{\partial f(x_k, u_k)}{\partial x_k}\right|_{\hat{x}_{k|k}, u_k} \tag{6}$$

$$B_k = \left.\frac{\partial f(x_k, u_k)}{\partial u_k}\right|_{\hat{x}_{k|k}, u_k} \tag{7}$$

$$H_k = \left.\frac{\partial h(x_k)}{\partial x_k}\right|_{\hat{x}_{k|k-1}}. \tag{8}$$

After incorporating the binary measurements, we can show that the posterior distribution is a Gaussian Mixture (GM), assuming the prior is also a GM. A GM is a distribution whose pdf is a weighted sum of Gaussian pdf's of the form

$$g(x) = \sum_{i=1}^{M} w_i \phi(x; \mu_i, \Sigma_i), \tag{9}$$

where $\phi(x; \mu_i, \Sigma_i)$ is the pdf of a Gaussian distribution with mean μ_i and covariance matrix Σ_i, and w_i are weights such that $\sum_{i=1}^{M} w_i = 1$. GM's have been shown to have interesting properties: due to their linear relation to Gaussian pdf's, they can be propagated with a bank of Kalman filters but they can be also used to approximate any pdf with finitely many discontinuities [9].

Proposition 1. *Consider a system with a linear motion model*

$$x_{k+1} = F_k x_k + B_k u_k + w_k,$$

linear state observation model

$$y_k = H_k x_k + v_k,$$

and context observation $p_d(c_i \mid x_k)$ expressed as in (5). Assuming that the state prior $p_{k|k}$ is a Gaussian Mixture, then the predicted and updated pdf's, $p_{k+1|k}$ and $p_{k+1|k+1}$ respectively, are also Gaussian Mixtures ***without any approximation****.*[3]

Proof: Note that, unlike the conventional EKF that has a predict and an update stage, the proposed filter has three steps: prediction, continuous update and discrete update. There is also an optional mixture reduction step discussed at the end of the section.

A. Predict

For the predict stage, we note that

$$\begin{aligned} p_{k+1|k}(x) &= \sum_{i=1}^{M} w_i \int \phi(x; F_k z + B_k u_k, Q) \phi(z; \mu_i, \Sigma_i) dz \\ &= \sum_{i=1}^{M} w_i \phi(x; F_k \mu_i + B_k u_k, F_k \Sigma_i F_k^T + Q) \\ &= \sum_{i=1}^{M} w_i \phi(x; \mu_i^p, \Sigma_i^p), \end{aligned}$$

which is the usual form of the Kalman filter predict equations (e.g., see [11]). The resulting distribution is again a GM.

[2]Note that the function in the exponent in (5) does not need to be linear in x_k. If it is not linear, i.e., it has the form $p_d(c_i \mid x_k) = \exp\{-\frac{1}{2}(g_i(x_k))^T V_i^{-1} (g_i(x_k))\}$, it can be linearized by computing the Jacobian matrix of g_i and expressed in the same form as (5).

[3]Note that if the original system is nonlinear and a linear approximation is used such as in (6)-(8), then the filter derived in this section is an approximation as well.

B. Continuous Update

As described above, we perform the update separately for state (continuous) and context (discrete) sensors. Upon receiving a measurement y_{k+1}, the continuous update is:

$$
\begin{aligned}
p^c_{k+1|k+1}(x) &= \frac{p(y_{k+1} \mid x)p_{k+1|k}(x)}{\int p(y_{k+1} \mid z)p_{k+1|k}(z)dz} \\
&= \frac{\phi(y_{k+1}; H_k x, R)\sum_{i=1}^M w_i \phi(x; \mu^p_i, \Sigma^p_i)}{\int \phi(y_{k+1}; H_k z, R)\sum_{j=1}^M w_j \phi(z; \mu^p_j, \Sigma^p_j)dz} \\
&= \sum_{i=1}^M \left(\frac{w_i \gamma^c_i}{\alpha^c}\right) \frac{\phi(y_{k+1}; H_k x, R)\phi(x; \mu^p_i, \Sigma^p_i)}{\int \phi(y_{k+1}; H_k z, R)\phi(z; \mu^p_i, \Sigma^p_i)dz} \\
&= \sum_{i=1}^M w^c_i \phi(x; \mu^c_i, \Sigma^c_i),
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha^c &:= \sum_{i=1}^M w_i \gamma^c_i \\
\gamma^c_j &:= \int \phi(y_{k+1}; H_k z, R)\phi(z; \mu^p_j, \Sigma^p_j)dz \\
&= \phi(y_{k+1}; H_k \mu^p_j, H_k \Sigma^p_j H_k^T + R) \\
\mu^c_j &:= \mu^p_j + K^c_j(y_{k+1} - H_k \mu^p_j) \\
\Sigma^c_j &:= (I - K^c_j H_k)\Sigma^p_j \\
K^c_j &:= \Sigma^p_j H_k^T (H_k \Sigma^p_j H_k^T + R)^{-1}.
\end{aligned}
$$

Note that the posterior distribution is also a GM with the same number of elements but with possibly rescaled weights.

C. Discrete Update

For the discrete update, to simplify notation we assume that there is only one context element, with corresponding parameters θ, V and G, such that $b_{k+1} \in \{0, 1\}$, though the approach can be straightforwardly modified to include multiple binary measurements through repeated updates.

First note that the posterior distribution depends on whether b_{k+1} is 0 or 1 as the probabilities of getting either one are different. Consider first the case when $b_{k+1} = 1$:

$$
\begin{aligned}
p_{k+1|k+1}(x) &= \frac{p(b_{k+1} = 1 \mid x)p^c_{k+1|k+1}(x)}{\int p(b_{k+1} = 1 \mid z)p^c_{k+1|k+1}(z)dz} \\
&= \frac{\phi(\theta; Gx, V)\sum_{i=1}^M w^c_i \phi(x; \mu^c_i, \Sigma^c_i)}{\int \phi(\theta; Gz, V)\sum_{j=1}^M w^c_j \phi(z; \mu^c_j, \Sigma^c_j)dz} \\
&= \sum_{i=1}^M \left(\frac{w^c_i \gamma^d_i}{\alpha^d}\right) \frac{\phi(\theta; Gx, V)\phi(x; \mu^c_i, \Sigma^c_i)}{\int \phi(\theta; Gz, V)\phi(z; \mu^c_i, \Sigma^c_i)dz} \\
&= \sum_{i=1}^M \left(\frac{w^c_i \gamma^d_i}{\alpha^d}\right) \phi(x; \mu^d_i, \Sigma^d_i),
\end{aligned}
$$

where α^d, γ^d_j, μ^d_j, Σ^d_j and K^d_j are defined similar to their continuous analogues.

Finally, when $b_{k+1} = 0$, the update becomes

$$
\begin{aligned}
p_{k+1|k+1}(x) &= \frac{(1 - p(b_{k+1} = 1 \mid x))\, p^c_{k+1|k+1}(x)}{\int (1 - p(b_{k+1} = 1 \mid z))\, p^c_{k+1|k+1}(z)dz} \\
&= \sum_{i=1}^M \frac{w^c_i (1 - p(b_{k+1} = 1 \mid x))\, \phi(x; \mu^c_i, \Sigma^c_i)}{1 - \sum_{j=1}^M w^c_j \int p(b_{k+1} = 1 \mid z)\phi(z; \mu^c_i, \Sigma^c_i)dz} \\
&= \sum_{i=1}^M \frac{w^c_i}{1 - \sum_{j=1}^M w^c_j \beta_j} \phi(x; \mu^c_i, \Sigma^c_i) + \\
&\quad + \frac{-w^c_i \beta_i}{1 - \sum_{j=1}^M w^c_j \beta_j} \phi(x; \mu^d_i, \Sigma^d_i)
\end{aligned}
$$

where $\beta_j := \gamma^d_j \sqrt{(2\pi)^q \det(V)}$. ■

Thus, we have inductively shown that for the probability of detection function considered in this paper, the localization filter can be computed in closed form and results in a GM distribution of the posterior. Note that the number of elements in the GM doubles every time $b_k = 0$, thus an additional step may be necessary in order to bound the number of elements.

D. Mixture Reduction

In the proof of Proposition 1 we derived an exact form for the posterior distribution of the state but it results in a potentially exponential number of elements in the GM. There are many ways of reducing these elements available in the literature (e.g., see [24]) ranging from simply choosing the elements with highest weights to merging or discarding elements based on distances between them. These techniques cannot be directly applied to the GM in this paper because they assume all weights are positive, whereas the GM we developed can have negative weights as well.

In such cases, another pruning method that has been developed in the literature is the Gibbs Sampler [25]. The Gibbs Sampler draws random samples from the distribution and approximates it with a GM with a desired number of elements. In order to sample from a distribution with negative weights such as the one developed in this paper, accept-reject sampling may be utilized [16].

IV. Simulations

This section presents simulations of the urban localization scenario using the context-aware filter developed in Section III. In particular, the experimental platform is the LandShark robot; it is a differential drive vehicle that is equipped with a GPS and a camera. The LandShark is usually used in critical missions on enemy territory, such as carrying injured people, hence precise localization is crucial during its operation.

The simulated scenario is as follows. The LandShark is moving through an urban environment and trying to reach different waypoints as part of its mission. The city's map is available, in addition to the coordinates of various known buildings. The localization problem is made difficult because GPS measurements have a large variance and a large bias to the North. Therefore, the LandShark is also using its camera and trying to recognize nearby buildings through image processing.

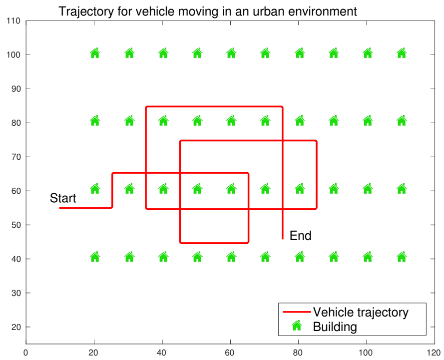

Fig. 1: Entire LandShark trajectory.

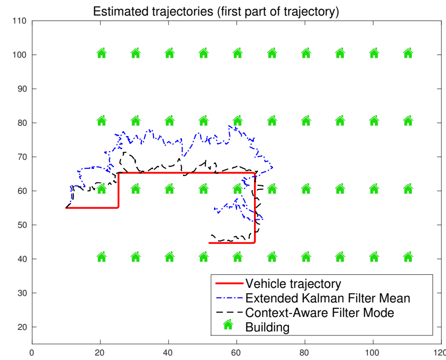

Fig. 2: Estimated trajectories.

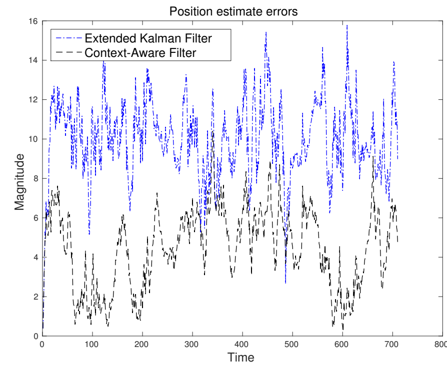

Fig. 3: Position errors by each filter.

The actual LandShark trajectory, including the map, is shown in Figure 1. The robot makes several tours around the city streets, in between the buildings, while visiting its waypoints. Note that the differential drive means that any turn causes the motion model to be nonlinear.

To perform localization, we use the proposed context-aware filter with both GPS and context measurements (i.e., recognized buildings). Since the posterior distribution of the state is a Gaussian Mixture, the mean of the distribution may not be the best estimate of the state as it may not even correspond to a peak in the pdf. Thus, we use the mode, i.e., the highest peak, as the estimate of the state. The utilized pruning technique is to keep the elements with highest weights since that was sufficient in this application. We will investigate other techniques (e.g., Gibbs Sampler) in future work.

To evaluate the performance of the filter, we compare it with an EKF using only GPS for localization. The resulting estimates for each filter are shown on Figure 2. Only the estimates for the first part of the path are shown to avoid clutter on the graph. As can be seen from the Figure, the EKF's estimates are greatly affected by the GPS bias to the North and also tend to vary sharply due to the large variance of GPS. The context-aware filter, on the other hand, is generally much closer to the actual path and is much less sensitive to the bias and variance of GPS.

For further comparison, Figure 3 presents the position errors incurred by each filter for the whole trajectory. This Figure also supports the fact that the context-aware filter is much closer to the ground truth and does not vary greatly from one round to the next. Therefore, we conclude that context-aware localization is robust to severe sensor errors and is a promising direction for future research and incorporation into other robotics applications where current sensing infrastructures are insufficient.

V. Discussion

Note that the approach developed in this paper can be extended to many other applications and domains. In general, it applies to any state estimation problem in which discrete observations exist that are correlated with the system's state. A notable example is the medical domain because there are multiple alarms (which can be considered as detection events) occurring constantly in modern operating rooms. These alarms can vary from thersholding alarms to more sophisticated techniques that target specific events [10].

These two types of alarms can be modeled with different classes of probability of detection functions. Specifically, the former can be modeled with thresholding functions (e.g., what is the probability that a patient's blood oxygen saturation is below 90% given the current state) whereas the latter are harder to model directly, hence a higher-level class of functions may prove better suited (e.g., what is the probability a glucose-based meal detector will raise an alarm given the current patient state).

The probability function considered in this paper is of the second type – it does not specify a probability model of how the used image processing algorithm would detect a certain object; similarly, it does not model how FM signals are detected by a radio receiver. Instead, it stipulates that both of these detections are more likely from certain positions and not so likely from others. This class of functions is in some sense more general because it does not require expert knowledge of the underlying process but only of how it manifests during the robot's operation.

As future work, we will also incorporate the first class of functions in our filtering model. These probability functions model the fact that a certain signal (e.g., blood oxygen saturation, speed) exceeds a specified threshold. Such functions may work better in certain scenarios because they model the underlying physical process directly, whereas the one developed in this paper focuses on input-output relationships between a state and probability of detection. Possible forms of these functions include sigmoids and more general logit/probit functions [17] that have known integration formulas and can be included in the filtering algorithm.

VI. Conclusion

In this paper, we considered the problem of localization in the presence of large sensor imprecisions. We focused on two specific applications in which localization suffers greatly from wrong GPS measurements. We noted that, in addition to continuous sensors, modern robots also have access to discrete observations of their surrounding context that can improve the localization algorithms. By fixing a specific class of probability of detection of these context events, we derived a closed-form context-aware filter. We showed that this filter performs much better than current EKF techniques for the applications considered in this work. As part of future work, we intend to consider a larger class of probability of detection functions and extend the filter to other domains.

References

[1] R. Anati, D. Scaramuzza, K. Derpanis, and K. Daniilidis. Robot localization using soft object detection. In *IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 4992–4999, 2012.

[2] N. Atanasov, M. Zhu, K. Daniilidis, and G. Pappas. Semantic localization via the matrix permanent. In *Robotics: Science and Systems*, 2014.

[3] S. Bao and S. Savarese. Semantic structure from motion. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 2025–2032, 2011.

[4] M. Blaschko and C. Lampert. Object localization with global and local context kernels. In *British Machine Vision Conference (BMVC)*, pages 63.1–63.11, 2009.

[5] M. Cummins and P. Newman. Fab-map: Probabilistic localization and mapping in the space of appearance. *International Journal of Robotics Research (IJRR)*, 27 (6):647–665, 2008.

[6] C. Galindo, A. Saffiotti, S. Coradeschi, P. Buschka, J. Fernandez-Madrigal, and J. Gonzalez. Multi-hierarchical semantic maps for mobile robotics. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 2278–2283, 2005.

[7] D. Gálvez-López, M. Salas, J. Tardós, and J. Montiel. Real-time monocular object slam. *arXiv:1504.02398*, 2015.

[8] A. Goldsmith. *Wireless Communications*. Cambridge university press, 2005.

[9] M. Huber. *Nonlinear Gaussian Filtering: Theory, Algorithms, and Applications*, volume 19. KIT Scientific Publishing, 2015.

[10] R. Ivanov, J. Weimer, A. Simpao, M. Rehman, and I. Lee. Early detection of critical pulmonary shunts in infants. In *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*, ICCPS '15, pages 110–119. ACM, 2015.

[11] R. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82 (1):35–45, 1960.

[12] T. Kos, I. Markezic, and J. Pokrajcic. Effects of multipath reception on GPS positioning performance. In *ELMAR, 2010 PROCEEDINGS*, pages 399–402. IEEE, 2010.

[13] J. Levinson, M. Montemerlo, and S. Thrun. Map-based precision vehicle localization in urban environments. In *Robotics: Science and Systems*, volume 4, page 1. Citeseer, 2007.

[14] R. Mahler. *Statistical Multisource-Multitarget Information Fusion*. Artech House, Inc., 2007.

[15] M. Montemerlo and S. Thrun. Simultaneous localization and mapping with unknown data association using fastslam. In *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*, volume 2, pages 1985–1991, Sept 2003.

[16] R. Neal. Slice sampling. *Annals of statistics*, pages 705–741, 2003.

[17] H. Nickisch and C. Rasmussen. Approximations for binary gaussian process classification. *Journal of Machine Learning Research (JMLR)*, 9(Oct):2035–2078, 2008.

[18] A. Nüchter and J. Hertzberg. Towards semantic maps for mobile robots. *Robotics and Autonomous Systems*, 56(11):915–926, 2008.

[19] L. Paletta, S. Frintrop, and J. Hertzberg. Robust localization using context in omnidirectional imaging. In *IEEE Int. Conf. on Robotics and Automation (ICRA)*, volume 2, pages 2072–2077, 2001.

[20] S. Peterson and P. Faramarzi. Iran hijacked US drone, says Iranian engineer. *Christian Science Monitor, December*, 15, 2011.

[21] A. Pronobis. *Semantic Mapping with Mobile Robots*. dissertation, KTH Royal Institute of Technology, 2011.

[22] R. Salas-Moreno, R. Newcombe, H. Strasdat, P. Kelly, and A. Davison. Slam++: Simultaneous localisation and mapping at the level of objects. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 1352–1359, 2013.

[23] D. Shepard, J. Bhatti, and T. Humphreys. Drone hack. *GPS World*, 23(8):30–33, 2012.

[24] H. Sorenson. *Kalman Filtering: Theory and Application*. IEEE Press, 1985.

[25] E. Sudderth, A. Ihler, W. Freeman, and A. Willsky. Nonparametric belief propagation. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 605–612, 2003.

[26] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT press, 2005.

[27] J. Wang, H. Zha, and R. Cipolla. Coarse-to-fine vision-based localization by indexing scale-invariant features. *IEEE Trans. on Systems, Man, and Cybernetics*, 36(2): 413–422, 2006.

[28] C. Yi, I. Suh, G. Lim, and B. Choi. Active-semantic localization with a single consumer-grade camera. In *IEEE Int. Conf. on Systems, Man and Cybernetics*, pages 2161–2166, 2009.

[29] C. Yi, I. Suh, G. Lim, S. Jeong, and B. Choi. Cognitive representation and bayeisan model of spatial object contexts for robot localization. In *Advances in Neuro-Information Processing*, volume 5506 of *Lecture Notes in Computer Science*, pages 747–754. Springer, 2009.