2016 IEEE International Conference on Robotics and Automation (ICRA)
Stockholm, Sweden, May 16-21, 2016

Localization accuracy estimation with application to perception design

Jan Rohde[1], Jan Erik Stellet[1], Holger Mielenz[1], J. Marius Zöllner[2]

***Abstract*— Landmark-based localization in dynamic environments poses high demands on the perception system of a mobile robot. The pose estimate generally has to fulfill specific accuracy requirements which might be necessitated by dependent systems, such as behavior planning. Thus, in this contribution we focus on the model-based derivation of perception requirements, i.e. detectable landmark types and minimum detection rates, to enable global localization with a specified upper bound on uncertainty. To this end, we utilize stochastic geometry to accurately capture and explicitly consider characteristics of the dynamic environment (e.g. occlusions), and the perception system (e.g. missed detections). From this point our contributions are twofold: i) We propose an analytical model of upper bounds on localization uncertainty. For continuous pose tracking, the Kalman filter equations for intermittent observations are considered and ii) perception requirements, i.e. minimum detection rates, based on specified upper bounds on pose estimation uncertainty are derived. Monte Carlo simulations are used to demonstrate the performance of the proposed methods.**

I. Introduction

The operation of automated mobile robots in dynamic outdoor environments places outstanding requirements on the perception system. We understand the perception system as the unity of exteroceptive sensors, their configuration and the set of algorithms for feature extraction from raw sensor measurements. Furthermore, we define features as landmarks, if they are reproducibly detectable. Hence, landmark types range from features, e.g. SIFT and its extension GLOH [1] or ORB [2], to real objects, e.g. traffic signs or active light sources [3], and can contain mixtures of different types. Due to the type specific spatial distribution and density of the landmarks in the operation environment, the selection of perception algorithms directly influences the number of detectable landmarks. Hence, the robustness of the localization subsystem against occlusions and missed detections strongly depends on the choice of the perception algorithms.

Frameworks for landmark-based global localization rely on the availability of a sufficient rate of landmarks that are detected by the perception system and matched to a global map. Hence, occlusions caused by other dynamic objects in the scene, e.g. vehicles and pedestrians, can dramatically reduce the performance of pose estimation accuracy and dependability. The performance of the perception system can be characterized by its ability to extract a specific number of features (e.g. by considering different types of landmarks) from the environment with sensor system specific measurement accuracy. We argue, that in mobile robotics for dynamic outdoor environments, e.g. urban areas, the distribution of a specific type of landmark is generally fixed and the usage of artificial landmarks is likely to be infeasible. Additionally, the configuration and performance of the sensory system might be predefined as well, especially when commercial sensors are used.

Under the aforementioned constraints, this work considers requirements on a perception system on the algorithm layer, i.e. minimum true positive rates of detection algorithms for a carefully chosen set of landmark types. We address the important and yet mainly unresolved question of how mandatory upper bounds on localization uncertainty, e.g. imposed by a behavior planning system, influence these requirements. To this point, the design of a perception system for automated robot operation in cluttered, e.g. urban, environments should

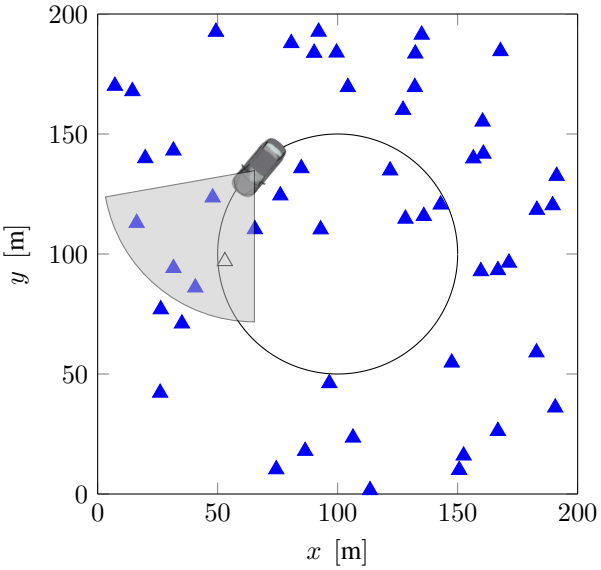

Fig. 1: An example environment with a realization of landmark locations (blue triangles) according to a Poisson point process is depicted. The automated vehicle has to perform localization with given accuracy along a planned trajectory (black line) by means of a fixed sensor configuration (FoV: grey area). Our framework enables the derivation of maximum missed detection (white triangle) rates for the perception system under consideration of a stochastic landmark distribution, uncertainty in landmark position measurements and the map accuracy.

[1]Jan Rohde, Jan Erik Stellet and Holger Mielenz are with Robert Bosch GmbH, Corporate Research, Vehicle Safety and Assistance Systems, 71272 Renningen, Germany

[2]Johann Marius Zöllner is with Research Center for Information Technology (FZI), 76131 Karlsruhe, Germany

take landmark distribution characteristics, occlusions and missed detections into account.

In a first **contribution**, we present a model that describes the influence that environment characteristics have on the localization uncertainty. Furthermore, we propose an environment model which is based on stochastic point processes (SPP). Since in general the exact positions of the landmarks cannot be known beforehand, they are assumed to be stochastic. In this work we will consider the special case of Poisson point processes (PPP, Sec. III-A). Our stochastic environment model allows us to incorporate different feature densities and occlusion rates into the localization uncertainty estimation. Afterward, we outline the connection between this SPP-based model and the measurement uncertainty. Lastly, this framework is combined with the theory of Kalman filtering with intermittent observations [4]. We then evaluate the interconnection between the SPP parameters and the pose estimation system's compliance to user defined estimation uncertainty bounds. By inverting the model, it becomes possible to estimate constraints on a perception system for a specific environment and robotic system.

The **outline** of this paper is the following. A selection of related work is presented in Sec. II. It is followed by a technical overview of the considered problem in Sec. III. Therein, we focus on the SPP-based environment model and its implications concerning Kalman filtering with intermittent observations. In Sec. IV and V we present a localization performance model which forms the basis for the design of perception parameters, i.e. landmark detection rates. This model is evaluated and applied to perception design in Sec. VI. The conclusion in Sec. VII finishes this contribution.

II. Related Work

Estimating the localization performance of a robot system in a specific environment has found wide interest in the field of robotics. Censi proposed a method for estimating the range-finder based localization [5] and pose tracking [6] covariance. Contributions like these are essential for several applications, such as the one presented by Beinhofer *et al.* [7]. They propose an algorithm for the near-optimal placement of artificial landmarks along a given trajectory. We argue, that in many outdoor applications it might be infeasible to use such artificial landmarks. Hence, the landmark distribution takes a more stochastic character. We explicitly consider the stochastic environmental characteristics and integrate them into the Kalman filtering framework.

Within the last decades, stochastic geometry and its special case of SPPs have been instrumented for the design of wireless communication networks [8] and distributed sensing. Modeling the environment, or specific aspects of it, as random finite sets has become popular within the field of multitarget tracking in the last decade as well. Starting with Mahler's outstanding work which laid the foundation of the so-called finite set statistics (FISST) [9] and led to the successful implementation of, e.g., PHD-Filters for feature-based SLAM in outdoor environments [10]. In our contribution we model the arrangements of features and corresponding occlusion rates and missed detections in dynamic environments as superposed SPPs.

Due to occlusions, missed detections and sparseness of features, a pose fixation from map matching cannot be guaranteed in every time step. In control theoretic works this is known as an intermittent observations problem [4]. One of the first works on Kalman filtering with intermittent observations was published by Sinopoli *et al.* in [4] and is still undergoing further extensions and generalization (e.g. [11], [12]). Sinopoli *et al.* proved, that a bound on the observation probability exists, under which the estimation covariance is unbounded. In this work, the observation probability is given by the limited field of view (FoV) of the sensor, the perception algorithms and the characteristics of the landmark configuration. In [13], [14] an upper performance bound on multi-robot multitarget tracking accuracy is proposed. We extend and generalize their results to the case of intermittent observation. In combination with a new closed-form estimate of the pose fixation covariance, this allows us to efficiently estimate upper bounds on the pose tracking performance of a given robotic system in dynamic environments.

Although appropriate perception is of fundamental importance for the design of mobile robotic systems, to the authors' knowledge a small amount of related contributions are available. Tarabanis *et al.* [15] provide a thorough study on sensor planning in computer vision for industrial robotics. Their work focuses on the automatic planning of sensor configuration, i.e. orientation and mounting position, and the control of the camera optics (e.g. focus). In addition to the dissenting application domain, the presented approaches do not consider localization performance constraints for the derivation of perception requirements. Bansal *et al.* [16] provide an empirical investigation of the effects that extrinsic sensor parameters have on vehicle localization accuracy. As a result, they state advantageous parameter assignments. In contrast to [16], our approach is independent of a specific sensor technology or setup and presented in closed form. As opposed to [15], [16], we approach the perception design on the algorithmic as opposed to the hardware level. In the authors' previous work [17], a framework for requirement estimation for perception in an urban setting has been proposed. Pose estimates have been considered in single time steps, that is without filtering over time. The work at hand now extends the previous study to filtered estimates for entire trajectories.

III. Background and Problem Formulation

The goal in localization is to estimate the global pose $\mathbf{x}_k := \begin{bmatrix} x_k & y_k & \theta_k \end{bmatrix}^\top$ of a dynamic system:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{u}_k + \mathbf{w}_k \,, \qquad \mathbf{w}_k \sim \mathcal{N}\big(\mathbf{0}, \mathbf{Q}\big) \tag{1}$$

with covariance $\mathbf{Q}$ of odometry measurements $\mathbf{u}_k$. Global pose measurements $\mathbf{z}_k$ with a covariance $\mathbf{\Sigma}_{\mathbf{z},k}$ are obtained from map matching. A linear relation to the state is assumed:

$$\mathbf{z}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \,, \qquad \mathbf{v}_k \sim \mathcal{N}\big(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{z},k}\big) \,. \tag{2}$$

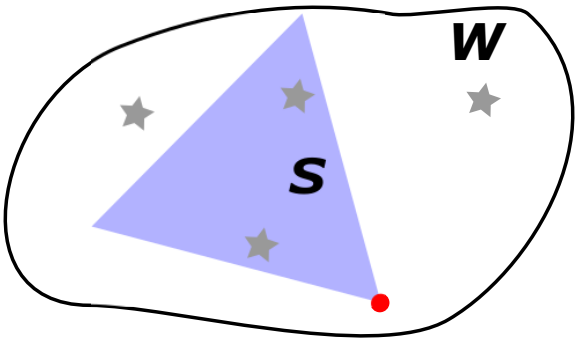

Fig. 2: A stochastic point process is used to model the global landmark distribution which is defined on the world space $\boldsymbol{\mathcal{W}}$. Samples of the global SPP represent feasible map realizations. The global SPP is observed on $\boldsymbol{\mathcal{S}} \subset \boldsymbol{\mathcal{W}}$, i.e. the field of view of the mobile robot's exteroceptive sensors.

Due to the environment's stochastic characteristics (Sec. III-A), e.g. occlusions, sparse features and missed detection, the probability λ of obtaining a pose fixation in time step k is smaller than one. In this case, a state estimate $\hat{\mathbf{x}}$ can be obtained with a Kalman filter for intermittent observations which will be introduced in Sec. III-B.

A. Stochastic geometry based measurement model

To capture the stochastic nature of a dynamic environment in a closed-form model, we utilize stochastic point processes (SPP), which are a part of stochastic geometry. In this contribution, we assume that landmark positions are realizations of a Poisson point process (PPP), a special case of SPPs. This assumption has several implications which lead to convenient analytic expressions. The realizations occur in the continuous world state space $\boldsymbol{\mathcal{W}} = \mathbb{R}^d$, where d might take the values 2 (flat world assumption), as assumed throughout this work, or 3 in the general 3-dimensional case. A realization of a SPP is obtained by drawing a number of N landmarks from a probability distribution (Poisson for PPPs) and then independently drawing N landmark locations from an arbitrary second distribution. A precise localization map is now defined as a set $\mathcal{M} = (M, \{\mathbf{m}_1, \dots \mathbf{m}_M\})$, $\mathcal{M} \in \mathcal{E}(\boldsymbol{\mathcal{W}})$ of landmark locations $\mathbf{m}_i$. The intensity of a specific type of landmarks can be estimated from real measurements, e.g. by expectation maximization. Since we utilize static environment features, these intensity values can be reused for different depth sensor technologies, their configurations and driving maneuvers.

Landmark occurrences can be observed in the arbitrarily shaped FoV $\boldsymbol{\mathcal{S}} \subset \boldsymbol{\mathcal{W}}$ of the exteroceptive sensor setup, as shown in Fig. 2. For homogeneous PPPs the observed properties are independent of the sensor position and hence efficient model formulations can be obtained. This approach can be easily extended to a multi-sensor setup with possibly overlapping FoVs. $\boldsymbol{\mathcal{S}}$ comprises $N \geq 0$ true landmark positions $\mathbf{x}_1, \dots, \mathbf{x}_N$. Measurements $\mathbf{s}_1, \dots, \mathbf{s}_N$ of the positions are disturbed by zero-mean Gaussian measurement noise with covariance $\boldsymbol{\Sigma}_{\mathbf{s}_i}$. The event space $\mathcal{E}(\boldsymbol{\mathcal{S}})$ contains the realization of the PPP which are denoted:

$$\xi = (N,\ \{\mathbf{x}_1, \dots, \mathbf{x}_N\})\ , \qquad \xi \in \mathcal{E}(\boldsymbol{\mathcal{S}})\ . \tag{3}$$

For each of $l = 1, \dots, L$ landmark types, a stochastic point process $\boldsymbol{\Psi}_l$ with corresponding intensity function $\psi_l(\mathbf{x})$ exists, that describes the landmark intensity for a given position $\mathbf{x}$ in space. The overall landmark distribution is then given by the superposition:

$$\psi(\mathbf{x}) = \sum_{l=1}^{L} \psi_l(\mathbf{x})\ . \tag{4}$$

For $\psi_l(\mathbf{x}) = \psi_l$ a PPP is homogeneous and samples of landmark locations are drawn from an uniform distribution. In the remainder of this work, homogeneous PPPs are considered for the sake of readability.

In dynamic environments, occlusions and missed detections can lead to a significant decrease in landmark-based localization performance. Therefore, we explicitly model occlusions by independent Bernoulli thinning of the underlying point process. For every landmark position $\mathbf{x} \in \boldsymbol{\mathcal{W}}$ let $\rho(\mathbf{x})$, $0 \leq \rho(\mathbf{x}) \leq 1$ be the probability that a landmark located at $\mathbf{x}$ is occluded. The resulting point process has the intensity function:

$$\upsilon(\mathbf{x}) = \rho(\mathbf{x})\,\psi(\mathbf{x}) = \rho(\mathbf{x}) \sum_{l=1}^{L} \psi_l(\mathbf{x})\ . \tag{5}$$

This forms the overall SPP $\boldsymbol{\Upsilon}$ with intensity $\upsilon(\mathbf{x})$ in the proposed environment model. Hence, occlusions are modeled as part of the environment and can also occur outside any sensor FoV. The probability of the occlusion of a landmark can vary, e.g. can be high for narrow roads with parked cars and low for main streets with small traffic.

Since the uncertainty in the map matching result depends on the realizations of the SPP, it can be defined as a function on the event space:

$$\boldsymbol{\Sigma}_{\mathbf{z},k} : \mathcal{E}(\boldsymbol{\mathcal{S}}) \to \mathbb{R}^{d \times d},\ \xi \to \hat{\mathbf{x}}\ . \tag{6}$$

As a result, the probability $\mathrm{P}(N = 0)$ is usually not zero and depends on $\boldsymbol{\Upsilon}$. Hence, one has to deal with intermittent observations, when not enough landmarks are visible to obtain a pose fixation. Secondly, the measurement uncertainty strongly depends on the number and positions of landmark within $\boldsymbol{\mathcal{S}}$, i.e. on $\boldsymbol{\Upsilon}$. Both aspects will be considered when deriving perception requirements in Sec. VI.

B. Kalman filtering with intermittent observations

The modified equations for Kalman filtering with intermittent observations [4] allow for explicit consideration of situations where not enough landmarks are visible to perform map matching. The probability density function of the measurement noise is given as follows [4]:

$$p(\mathbf{v}_k|\gamma_k) = \begin{cases} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{z},k})\ , & \gamma_k = 1 \\ \mathcal{N}(\mathbf{0}, \sigma \mathbf{I}_{3\times 3})\ , & \gamma_k = 0 \end{cases}\ . \tag{7}$$

The probability for obtaining a measurement is defined as $\mathrm{P}(\gamma_k = 1) = \lambda$. It is directly coupled to the characteristics

of the underlying SPP, the sensor setup, as well as the utilized map matching algorithm. For $\sigma \to \infty$, the modified Kalman equations are then given as [4]:

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{A}\hat{\mathbf{x}}_{k|k} + \mathbf{u}_k \tag{8a}$$
$$\mathbf{P}_{k+1|k} = \mathbf{A}\mathbf{P}_{k|k}\mathbf{A}^\top + \mathbf{Q}_k \tag{8b}$$
$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \gamma_{k+1}\mathbf{K}_{k+1}\left(\mathbf{z}_{k+1} - \mathbf{C}\hat{\mathbf{x}}_{k+1|k}\right) \tag{8c}$$
$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \gamma_{k+1}\mathbf{K}_{k+1}\mathbf{C}\mathbf{P}_{k+1|k} \tag{8d}$$
$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k}\mathbf{C}^\top\left(\mathbf{C}\mathbf{P}_{k+1|k}\mathbf{C}^\top + \mathbf{\Sigma}_{\mathbf{z},k}\right)^{-1} . \tag{8e}$$

In the remainder we will focus on the uncertainty in the filtered translation estimate denoted $\mathbf{P}_{\hat{\mathbf{t}},k}$.The time evolution of the covariance is expressed by the modified algebraic Riccati equation (MARE) which results from the stated KF equations [4]:

$$\begin{aligned}\mathbf{P}_{k+1|k+1} =& \mathbf{A}\mathbf{P}_{k|k}\mathbf{A}^\top + \mathbf{Q}_k - \lambda\mathbf{A}\mathbf{P}_{k|k}\mathbf{C}^\top \\ &\cdot\left(\mathbf{C}\mathbf{P}_{k|k}\mathbf{C}^\top + \mathbf{\Sigma}_{\mathbf{z},k}\right)^{-1}\mathbf{C}\mathbf{P}_{k|k}\mathbf{A}^\top .\end{aligned} \tag{9}$$

The first part of the problem can be described as finding the steady state solution of the MARE, given

- a stochastic point process which describes the landmark distribution,
- an error model for landmark location measurements and
- a model for localization uncertainty after map matching in a time step.

The inverse model of the steady state bound is then used to calculate a minimum detection rate for a given stochastic point process, sensor model and localization requirements. The result will then be evaluated in comprehensive Monte Carlo simulation experiments.

IV. Derivation of the pose fixation statistic

In this section, measurement uncertainty propagation is performed to the level of a pose fixation, i.e. the outcome of map matching. One main contribution is the derivation of an analytic expression for $\mathbf{\Sigma}_{\mathbf{z},k}$ in dependence of a given landmark configuration. Subsequently, in Sec. V, this stochastic model is utilized to calculate steady state upper bounds on the covariance of a filtered pose estimate.

For every landmark location $\mathbf{x}_i \in \mathcal{E}(\boldsymbol{\mathcal{S}})$, the position measurement $\mathbf{s}_i$ in the sensor's reference frame can be obtained with a sensor specific covariance $\mathbf{\Sigma}_{\mathbf{s}_i}$. A least-squares-optimal pose fixation can be obtained by the point-based matching of a landmark measurement with a global localization map. Information about the Gaussian, possibly anisotropic, measurement noise of the sensor can be included by a weighting matrix $\mathbf{W}_i$. This leads to the general formulation of the matrix-weighted Procrustes problem (MWPP) [19] whose solution provides an estimate of the rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$ of a robot relative to the map. Under the assumption of a given localization map and known point correspondences for the measured landmark positions the MWPP is given by [19]:

$$J = \sum_{i=1}^{N}\left(\mathbf{R}\mathbf{s}_i + \mathbf{t} - \mathbf{m}_i\right)^\top \mathbf{W}_i\left(\mathbf{R}\mathbf{s}_i + \mathbf{t} - \mathbf{m}_i\right) , \tag{10a}$$
$$\mathbf{R} = \begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix} . \tag{10b}$$

If the weighting matrices $\mathbf{W}_i$ do not depend on the unknown parameters $\mathbf{R}, \mathbf{t}$, then their estimates $\hat{\mathbf{R}}, \hat{\mathbf{t}}$ which minimize the cost function satisfy the following condition [19]:

$$\sum_{i=1}^{N}\mathbf{W}_i\mathbf{m}_i = \sum_{i=1}^{N}\mathbf{W}_i\left(\hat{\mathbf{R}}\mathbf{s}_i + \hat{\mathbf{t}}\right) . \tag{11}$$

This is known as the matrix-weighted centroid-coincidence (MWCC) theorem. Thus, an explicit expression for the optimal translation estimate $\hat{\mathbf{t}}$ is obtained, given $\hat{\mathbf{R}}$:

$$\hat{\mathbf{t}} = \left(\sum_{i=1}^{N}\mathbf{W}_i\right)^{-1}\cdot\left(\sum_{i=1}^{N}\mathbf{W}_i\mathbf{m}_i - \mathbf{W}_i\hat{\mathbf{R}}\mathbf{s}_i\right) . \tag{12}$$

Noting that the first factor serves as a normalization constant, the abbreviation $\mathbf{W}_\mathrm{s} := \sum_{i=1}^{N}\mathbf{W}_i$ is used hereafter to refer to this sum of the weighting matrices. With (12) inserted into (10), the cost function reads:

$$J = \sum_{i=1}^{N}\mathbf{a}_i^\top\mathbf{W}_i\mathbf{a}_i , \quad \text{where} \tag{13}$$
$$\mathbf{a}_i = \underbrace{\left[\mathbf{m}_i - \mathbf{W}_\mathrm{s}^{-1}\sum_{j=1}^{N}\mathbf{W}_j\mathbf{m}_j\right]}_{=:\mathbf{c}_i\left(\{\mathbf{m}_j|j=1,\ldots,N\}\right)} - \underbrace{\left[\mathbf{R}\mathbf{s}_i - \mathbf{W}_\mathrm{s}^{-1}\sum_{j=1}^{N}\mathbf{W}_j\mathbf{R}\mathbf{x}_j\right]}_{=:\mathbf{c}_i\left(\{\mathbf{R}\mathbf{x}_j|j=1,\ldots,N\}\right)} .$$

For notational convenience, a centering operator $\mathbf{c}_i(\cdot)$ has been introduced.

Since a closed-form solution that minimizes (13) does not exist in the general case, an approach based on the implicit function theorem (e.g., [20]) is employed to derive the covariance of the estimate:

$$\mathbf{\Sigma}_{\mathbf{z},k} = \mathrm{cov}\left[\hat{\mathbf{t}}^\top, \hat{\theta}\right]^\top \simeq \mathrm{blkdiag}\left[\mathbf{\Sigma}_\mathbf{t}, \sigma_\theta^2\right] \text{ with} \tag{14a}$$
$$\sigma_\theta^2 = \sum_{l=1}^{N}\left(\nabla_{\mathbf{s}_l}\hat{\theta}\right)\mathbf{\Sigma}_{\mathbf{s},l}\left(\nabla_{\mathbf{s}_l}\hat{\theta}\right)^\top + \left(\nabla_{\mathbf{m}_l}\hat{\theta}\right)\mathbf{\Sigma}_{\mathbf{m},l}\left(\nabla_{\mathbf{m}_l}\hat{\theta}\right)^\top ,$$
$$\begin{aligned}\mathbf{\Sigma}_\mathbf{t} =& \sum_{l=1}^{N}\left\{\left(\nabla_{\mathbf{s}_l}\hat{\mathbf{t}}\right)\mathbf{\Sigma}_{\mathbf{s},l}\left(\nabla_{\mathbf{s}_l}\hat{\mathbf{t}}\right)^\top + \left(\nabla_{\mathbf{m}_l}\hat{\mathbf{t}}\right)\mathbf{\Sigma}_{\mathbf{m},l}\left(\nabla_{\mathbf{m}_l}\hat{\mathbf{t}}\right)^\top\right\} \\ &+ \left(\frac{\partial}{\partial\theta}\hat{\mathbf{t}}\right)\sigma_\theta^2\left(\frac{\partial}{\partial\theta}\hat{\mathbf{t}}\right)^\top .\end{aligned} \tag{14b}$$

$\mathbf{\Sigma}_{\mathbf{s},l}$ and $\mathbf{\Sigma}_{\mathbf{m},l}$ denote the covariances of the landmark location measurements and the corresponding locations in the map compared to the landmark positions in the real environment, respectively. For precise 2-dimensional maps, these expressions can be simplified, since $\mathbf{\Sigma}_{\mathbf{m}_i} \simeq \mathbf{0}_{2\times 2}$. The necessary derivatives in (14) are obtained as follows:

$$\nabla_{\mathbf{s}_l}\hat{\theta} = \left(\frac{\partial^2}{\partial\theta^2}J\right)^{-1}\cdot\left(\frac{\partial}{\partial\theta}\nabla_{\mathbf{s}_l}J\right) , \tag{15a}$$
$$\nabla_{\mathbf{m}_l}\hat{\theta} = \left(\frac{\partial^2}{\partial\theta^2}J\right)^{-1}\cdot\left(\frac{\partial}{\partial\theta}\nabla_{\mathbf{m}_l}J\right) \tag{15b}$$

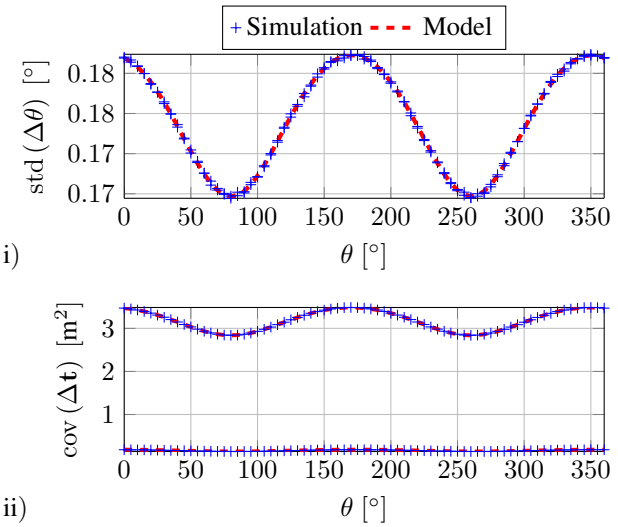

Fig. 3: i) Variances of $\Delta\mathbf{t}$ and ii) $\hat{\mathbf{t}}$ in x and y direction relative to the map reference frame. The model gives an exact estimate of the data obtained by Monte Carlo simulation. The evaluation was performed with three landmarks within the sensor field of view, an inaccurate map with $\mathbf{\Sigma}_{\mathbf{m},k} = 0.1 \cdot \mathbf{I}_{2\times 2}$ and a measurement noise according to the model $\mathbf{\Sigma}_{\mathbf{s}_i}$.

with

$$\frac{\partial^2}{\partial\theta^2} J = 2\sum_{i=1}^{N} \mathbf{c}_i^\top \left(\{\mathbf{R}'\mathbf{s}_j | j = 1,\ldots,N\}\right) \cdot \mathbf{W}_i \cdot \mathbf{c}_i \left(\{\mathbf{R}'\mathbf{s}_j | j = 1,\ldots,N\}\right) \tag{16}$$

and with $\mathbf{R}'$ as the derivative of $\mathbf{R}$ with respect to θ. The Jacobians with respect to $\mathbf{s}$ and θ are:

$$\frac{\partial}{\partial\theta}\nabla_{\mathbf{s}_l} J = 2\mathbf{c}_l^\top \left(\{\mathbf{R}'\mathbf{s}_j | j = 1,\ldots,N\}\right)\mathbf{W}_l\mathbf{R} \tag{17a}$$

$$\frac{\partial}{\partial\theta}\nabla_{\mathbf{m}_l} J = -2\mathbf{c}_l^\top \left(\{\mathbf{R}'\mathbf{s}_j | j = 1,\ldots,N\}\right)\mathbf{W}_l\,. \tag{17b}$$

And finally the Jacobians of $\hat{\mathbf{t}}$ read

$$\frac{\partial}{\partial\theta}\hat{\mathbf{t}} = -\mathbf{W}_\mathrm{s}^{-1}\sum_{i=1}^{N}\mathbf{W}_i\mathbf{R}'\mathbf{s}_i\,,$$
$$\nabla_{\mathbf{s}_l}\hat{\mathbf{t}} = -\mathbf{W}_\mathrm{s}^{-1}\mathbf{W}_l\mathbf{R}\,, \qquad \nabla_{\mathbf{m}_l}\hat{\mathbf{t}} = \mathbf{W}_\mathrm{s}^{-1}\mathbf{W}_l\,. \tag{18}$$

This model generalizes a previous result [17] for the scalar-weighted procrustes problem, i.e. $\mathbf{W}_i = w_i\mathbf{I}_{2\times 2}\ \forall i$.

The model was evaluated in Monte Carlo simulation with 10^6 iterations and a stereo-vision measurement model [18]:

$$\mathbf{\Sigma}_{\mathbf{s}_i} = \frac{1}{(c_\mathrm{k}\cdot b_\mathrm{w})^2} s_{i,x}^2 \begin{bmatrix} s_{i,x}^2\sigma_d^2 & s_{i,x}s_{i,y}\sigma_d^2 \\ s_{i,x}s_{i,y}\sigma_d^2 & s_{i,y}^2\sigma_d^2 + b_\mathrm{w}^2\sigma_u^2 \end{bmatrix}\,. \tag{19}$$

Here, the camera model is defined by the base-width b_w, focal length in pixels c_k and measurement noise variances σ_d^2, σ_u^2 in disparity and image column, respectively. The covariance is strongly distance-dependent. The results for uncertainties in $\hat{\theta}$ and $\hat{\mathbf{t}}$ under varying map-to-sensor rotation angles θ are shown in Fig. 3.

The proposed model will be utilized for efficient numerical integration. The measurement uncertainty $\mathbf{\Sigma}_{\mathbf{z},k}$ is a function of the realization of the PPP and hence stochastic as well. The function of translational map matching accuracy over homogeneous intensities υ are shown in Fig. 4. Hence, the probability density function $p(\mathbf{\Sigma}_z|\Upsilon)$ can be derived. We approximate this density function by α-quantiles obtained by Monte Carlo methods. For the calculation of the steady state upper bounds on localization uncertainty in Sec. V, we are interested in α-quantile values of $\mathbf{\Sigma}_\mathbf{z}$:

$$\mathbf{QU}(\alpha, \mathbf{\Upsilon}) = \mathbf{\Sigma}_{\mathbf{z},\alpha} :\ \mathrm{P}\left(\mathbf{\Sigma}_{\mathbf{z},k} \leq \mathbf{\Sigma}_{\mathbf{z},\alpha}\right) \geq \alpha\ \forall\ k\,. \tag{20}$$

As a metric to compare two covariance matrices, we considered their respective maximal of eigenvalue. For the evaluation of $\mathbf{QU}(\alpha, \mathbf{\Upsilon})$ we will resort to numerical integration.

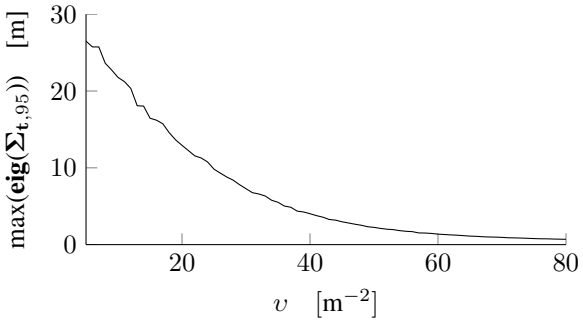

Fig. 4: The accuracy of the pose measurement after map matching strongly depends on the intensity υ of a homogeneous PPP.

V. Steady State Upper Bounds on Pose Estimates

As opposed to [4] we are interested in finding an upper bound on the state uncertainty. By calculating the steady state covariance $\mathbf{P}_{ss}$ for the MARE and a worst-case measurement covariance $\mathbf{\Sigma}_{\mathbf{z},\alpha}$ one obtains a probabilistic upper bound on the state covariance. Mirzaei *et al.* [13] proved that, for $\mathbf{\Sigma}_{\mathbf{z},u} \geq \mathbf{\Sigma}_{\mathbf{z},k}$ and $\mathbf{Q}_u \geq \mathbf{Q}_k$, for all $k \geq 0$, the solution to the MARE (9) with initial condition $\mathbf{P}_0^u$, satisfies $\mathbf{P}_k^u \geq \mathbf{P}_k$, $\forall k \geq 0$. We extend and generalize this lemma to the case of intermittent observations. For the discrete-time odometry motion model [21] with odometry measurement $[\delta_{x,k}\ \delta_{y,k}\ \delta_{\theta_k}]$

$$x_{k+1} = x_k + \delta_{x,k}\cos(\delta_{\theta_k}) \tag{21a}$$

$$y_{k+1} = y_k + \delta_{y,k}\sin(\delta_{\theta_k}) \tag{21b}$$

we can now state upper performance bounds. Since $\mathbf{A} = \mathbf{I}$, this leads to a simplification of the MARE. Hence, the steady-state uncertainty in pose-estimation can be stated in closed form:

$$\mathbf{P}_{ss}^u = \mathbf{Q}_u^{\frac{1}{2}}\mathbf{U}_u\mathbf{diag}\left(\frac{1}{2\lambda} + \sqrt{\frac{1}{4\lambda^2} + \frac{1}{\lambda\lambda_i}}\right)\mathbf{U}_u^\top\mathbf{Q}_u^{\frac{1}{2}} \tag{22}$$

$$\mathbf{Q}_u^{\frac{1}{2}}\mathbf{C}^\top\mathbf{\Sigma}_{\mathbf{z},u}^{-1}\mathbf{C}\mathbf{Q}_u^{\frac{1}{2}} = \mathbf{U}_u\mathbf{diag}(\lambda_i)\,\mathbf{U}_u^\top\,. \tag{23}$$

We define the worst-case measurement uncertainty by $\mathbf{\Sigma}_{\mathbf{z},u} = \mathbf{\Sigma}_{\mathbf{z},\alpha}$. This is a generalization of the results presented in [14] for the case of $\lambda \neq 1$, i.e. Kalman filtering with intermittent observations. Due to spatial limitations the proof is not stated but is easily found. [4] states a similar result for the case of scalar systems and static Kalman gain. For any given PPP λ is calculated as follows.

$$\lambda = \mathrm{P}\big(N \geq N_{\min}\big) = 1 - \sum_{N=0}^{N_{\min}-1} \mathrm{P}\big(N\big) \tag{24}$$

$$= 1 - \sum_{N=0}^{N_{\min}-1} \frac{\left(\int_{\boldsymbol{\mathcal{S}}} \upsilon\,(\mathbf{s})\,d\mathbf{s}\right)^{N}}{N!} \exp\left(-\int_{\boldsymbol{\mathcal{S}}} \upsilon\,(\mathbf{s})\,d\mathbf{s}\right) \tag{25}$$

The minimal number of required landmark measurements $N_{\min}$ to execute map matching depends on the localization framework. In the considered case one has $N_{\min} = 3$ for two-dimensional pose estimation. Thus, for a homogeneous PPP λ is given as:

$$\lambda = 1 - \left(\frac{1}{0!} + \frac{\upsilon A(\mathcal{S})}{1!} + \frac{(\upsilon A(\mathcal{S}))^2}{2!}\right) \exp\left(-\upsilon A(\mathcal{S})\right) \tag{26}$$

with the area $A(\mathcal{S})$ covered by the FoV.

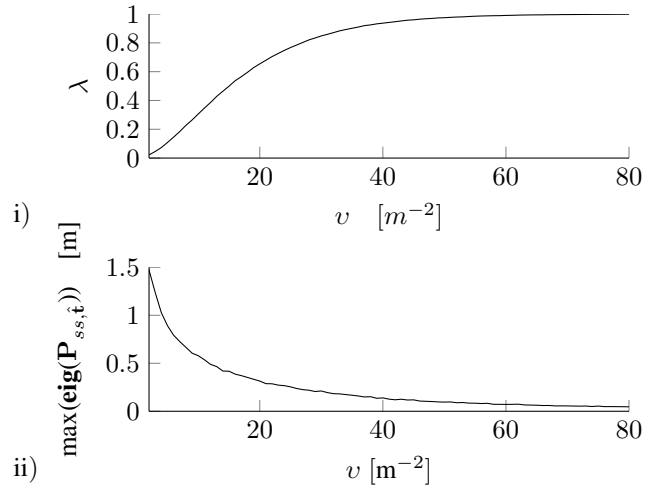

Fig. 5: i) Influence of the landmark intensity υ on the probability λ of obtaining a pose observation. ii) The upper bound in translation estimation uncertainty $\max(\mathbf{eig}(\mathbf{P}_{ss,\hat{\mathbf{t}}}))$ shows a strong dependence on the value of υ.

VI. Experimental evaluation and application

The stochastic geometry framework from Chapter III-A in combination with (14) and (22) enable a detailed understanding of the interplay between characteristics of the environment and pose tracking performance. This makes it possible to study a huge bandwidth of perception aspects.

A. Evaluation of localization performance bound

In this section, the localization performance bound is verified. To this end, Monte Carlo simulation is utilized. Simulation is conducted for 10^5 realizations of the stochastic

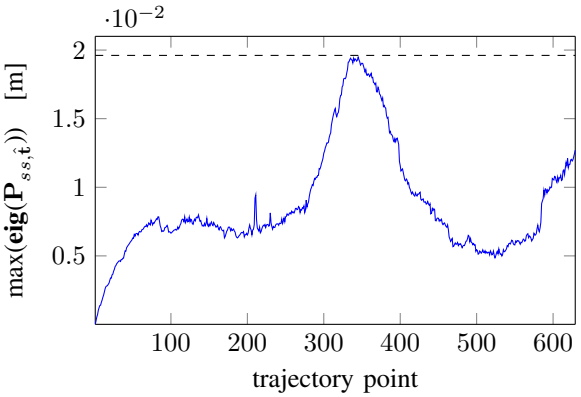

Fig. 6: The estimate on the upper bound (dashed line) of localization uncertainty is validated by the conducted simulation. For simulation, we used the trajectory from Fig. 1 with $\upsilon = 50$ per unit area of $100\,\mathrm{m}^2$ and assumed the localization map to be accurate.

landmark distribution and noise in odometry measurements. For pose tracking the modified Kalman filter equations from (8) are used in combination with the odometry motion model presented in (21a) and (21b). For model evaluation we consider a stationary, homogeneous Poisson point process with the constant intensity $\upsilon = 50$ conditioned on an area of $100\,\mathrm{m}^2$. The localization map is assumed to be precise, i.e. $\mathbf{\Sigma}_{\mathbf{m}} = \mathbf{0}$. For this experiment we assume $\mathbf{\Sigma}_{\mathbf{z},u} = \mathbf{\Sigma}_{\mathbf{z},99}$ and $\mathbf{Q}_u = \mathbf{diag}((0.04\,\mathrm{m})^2,\ (0.04\,\mathrm{m})^2,\ (0.04\,)^2)$ which equals the 2σ values of the odometry noise used in simulation. For the given landmark distribution, the steady state upper bound on localization uncertainty is calculated according to the proposed model. Fig. 1 shows the circular trajectory with a radius of $50\,\mathrm{m}$ in an area of $200\,\mathrm{m}$ side length. The results from Fig. 6 show that the calculated model bound gives a good estimate of the systems performance in the given environment. The uncertainty in vehicle pose estimation, measured as the maximal eigenvalue of $\mathbf{P}_{ss,\hat{\mathbf{t}}}$, remains below the estimated bound.

B. Application to the derivation of minimum detection rates

The proposed model is now utilized to determine perception algorithm requirements. Three exemplary landmark types with different homogeneous intensities $\psi_1 = 30$, $\psi_2 = 15$ and $\psi_3 = 15$ are considered. The upper bound on position estimation uncertainty is evaluated depending on the detection probability $1 - \rho$ and position-independent missed detection rate ρ, respectively. The results from Fig. 7 reveal that the capability of detecting landmarks of type 1 does not result in a sufficient localization performance. In combination with landmark type 2, the localization covariance is sufficient for high detection rates close to 1. Detecting all three landmark types results in an significantly increased robustness of the localization system against occlusions ($\rho_{max} = 0.25$).

VII. Conclusion

In this contribution we presented a new approach for performance estimation for localization in dynamic environ-

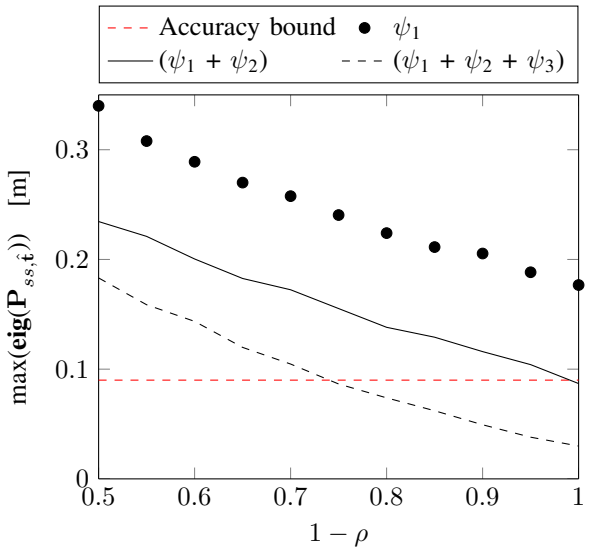

Fig. 7: The robustness of the localization system against occlusions is highly dependent on the amount of extractable environment information. The detection of landmark type 1 (solid line) never yields a sufficient localization performance. Detection of all landmark types (black dashed line) results in an increased robustness against occlusions ($\rho_{\max} = 0.25$) compared to the combination of type 1 and 2.

ments. Landmark locations are modeled as realizations of a stochastic point process. This enabled a detailed characterization of the interdependence of environment attributes and perception requirements. Based on the implicit function theorem we proposed a model of uncertainty in global pose estimation. This model in combination with the characteristics of the environment were then incorporated into the framework of Kalman filtering with intermittent observations. Solving the resulting MARE enabled the calculation of steady-state upper bounds on global localization uncertainty. Utilization of the inverse localization accuracy model enabled the derivation of sensing performance bounds for given environment characteristics and minimum detection rates. Additionally we underlined the importance of a correct choice of landmark detection algorithms by showing their direct influence on the robustness against missed detections and occlusions.

REFERENCES

[1] K. Mikolajczyk, C. Schmid, "A performance evaluation of local descriptors," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 27, no. 10, pp. 1615–1630, 2005.

[2] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, "ORB: An efficient alternative to SIFT or SURF," Computer Vision (ICCV), IEEE International Conference on, pp. 2564–2571, 2011.

[3] P. Nelson, W. Churchill, I. Posner, P. Newman, "From dusk till dawn: Localisation at night using artificial light sources," Robotics and Automation (ICRA), IEEE International Conference on, pp. 5245–5252, 2015.

[4] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan, S.S. Sastry, "Kalman filtering with intermittent observations," Automatic Control, IEEE Transactions on, vol. 49, no. 9, pp. 1453–1464, 2004.

[5] A. Censi, "On achievable accuracy for range-finder localization," Robotics and Automation (ICRA), IEEE International Conference on, pp. 4170–4175, 2007.

[6] A. Censi, "On achievable accuracy for pose tracking," Robotics and Automation (ICRA), IEEE International Conference on, pp. 1–7, 2009.

[7] M. Beinhofer, J. Müller, A. Krause, W. Burgard, "Robust landmark selection for mobile robot navigation," Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on, pp. 3637–2643, 2013.

[8] M. Haenggi, J.G. Andrews, F. Baccelli, O. Dousse, M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," Selected Areas in Communications, IEEE Journal on, vol.27, no.7, pp. 1029–1046, 2009.

[9] R.P.S. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," Aerospace and Electronic Systems, IEEE Transactions on, vol. 39, no. 4, pp. 1152–1178, 2003.

[10] J. Mullane, Ba-Ngu Vo, M.D. Adams, Ba-Tuong Vo, "A Random-Finite-Set Approach to Bayesian SLAM," Robotics, IEEE Transactions on, vol. 27, no. 2, pp. 268–282, 2011.

[11] Ling Shi, M. Epstein, R.M. Murray, "Kalman Filtering Over a Packet-Dropping Network: A Probabilistic Perspective," Automatic Control, IEEE Transactions on, vol. 55, no. 3, pp. 594–604, 2010.

[12] M. Hernandez, B. Ristic, A. Farina, L. Timmoneri, "A comparison of two Cramér-Rao bounds for nonlinear filtering with Pd<1," Signal Processing, IEEE Transactions on, vol. 52, no. 9, pp. 2361–2370, 2004.

[13] F.M. Mirzaei, A.I. Mourikis, S.I. Roumeliotis, "On the Performance of Multi-robot Target Tracking," Robotics and Automation (ICRA), IEEE International Conference on, pp. 3482–3489, 2007.

[14] F. M. Mirzaei. A. I. Mourikis, and S. I. Roumeliotis, "Analysis of positioning uncertainty in cooperative localization and target tracking (CLATT)." Dept. of Comp. Sci., Univ. of Minnesota, Tech. Rep., 2005.

[15] K.A. Tarabanis, P.K. Allen, R.Y. Tsai, "A survey of sensor planning in computer vision," Robotics and Automation, IEEE Transactions on, vol. 11, no. 1, pp. 86–104, 1995.

[16] A. Bansal, H. Badino, D. Huber, "Understanding how camera configuration and environmental conditions affect appearance-based localization," Intelligent Vehicles Symposium Proceedings (IV), IEEE, pp. 800–807, 2014.

[17] J. Rohde, J. E. Stellet, H. Mielenz, J. M. Zöllner, "Model-based derivation of perception accuracy requirements for vehicle localization in urban environments," Intelligent Transportation Systems (ITSC), IEEE International Conference on, 2015.

[18] H. Badino, "Binocular Ego-Motion Estimation for Automotive Applications," Ph.D. dissertation, Goethe University, 2008.

[19] Juyang Weng, P. Cohen, N. Rebibo, "Motion and structure estimation from stereo image sequences," Robotics and Automation, IEEE Transactions on, vol. 8, no. 3, pp. 362–382, 1992.

[20] A. Censi, "An accurate closed-form estimate of ICP's covariance," Robotics and Automation (ICRA), 2007 IEEE International Conference on, pp. 3167–3172, 2007.

[21] S. Thrun, W. Burgard, D. Fox, "Probabilistic Robotics," MIT Press, 2006.