Gaussian Process Mapping of Uncertain Building Models with GMM as Prior

Qianqian Zou, Claus Brenner and Monika Sester

Abstract—**Mapping with uncertainty representation is required in many research domains, especially for localization. Although there are many investigations regarding the uncertainty of the pose estimation of an ego-robot with map information, the quality of the reference maps is often neglected. To avoid potential problems caused by the errors of maps and a lack of uncertainty quantification, an adequate uncertainty measure for the maps is required. In this letter, uncertain building models with abstract map surfaces using Gaussian Processes (GPs) are proposed to describe the map uncertainty in a probabilistic way. To reduce the redundant computation for simple planar objects, extracted facets from a Gaussian Mixture Model (GMM) are combined with an implicit GP map, also employing local GP-block techniques. The proposed method is evaluated on LiDAR point clouds of city buildings collected by a mobile mapping system. Compared to the performance of other methods such as OctoMap, GP Occupancy Map (GPOM), Bayesian Generalized Kernel OctoMap (BGKOctoMap), Local automatic relevance determination Hilbert map (LARD-HM) and Gaussian Implicit Surface map (GPIS), our method achieves a higher Precision-Recall AUC for the evaluated buildings.**

Index Terms—**Mapping, Probability and statistical methods, Laser-based, Uncertainty representation.**

I. INTRODUCTION

MAPS play a crucial role in assisting autonomous systems, such as robots or automated vehicles, to comprehend unknown and complex scenes, enabling them to locate and navigate themselves accurately and safely. However, the uncertainties associated with maps significantly impact the performance of localization, navigation security, and collision avoidance of the autonomous system. With laser scanners, the environment can be represented in a straightforward way with point clouds, and further modelled using structured maps or abstract surface models [1]. The common maps are, e.g., high-definition (HD) vector maps [2], acquired with high geometric and semantic accuracy, 2D or 3D-grids with occupancy probabilities, such as OctoMaps [3], or 3D-City models with various levels of detail. Uncertainty sources in the environment map are inevitable due to, e.g., ambiguous environment information, sparsity, occlusions and noise in the measurements [4]. Hence, when the map is treated as an

The authors are with the Institut of Cartography and Geo-informatics in Leibniz University Hannover, Appelstraße 9a 30167 Hannover, Germany (email: qianqian.zou@ikg.uni-hannover.de; claus.brenner@ikg.uni-hannover.de; monika.sester@ikg.uni-hannover.de).

ideal perfect reference in localization and path-planning tasks without considering its quality, this can result in severe errors, e.g., the wrongly modelled occluded buildings might cause a bias in the estimated pose. Additionally, one may use multiple sensors to capture the environment or may have the demand of updating the existing maps with new data. In those scenarios, a good uncertainty measure of the maps created by different sensors and agents is essential for data fusion and motion planning purposes [5]. Although many existing studies about uncertainty models are found in automated driving related fields, they are mostly for pose estimation. There is a lack of sufficient research describing the accuracy and precision of maps.

In the field of simultaneous localization and mapping (SLAM), uncertainty-aware tasks require a reasonable expression of map uncertainty. There are some existing methods estimating the uncertainty of maps to improve localization accuracy or make the results more robust [1], [6], [7]. For example, Biber and Straßer [6] proposed to represent the environment using normal distribution transform (NDT) maps, where the map uncertainty is represented by a grid of distributions. However, it assumes independent discontinuous distributions of the neighboring cells and it has been found to result in a higher uncertainty at cell boundaries [8]. Javanmardi et al. [1] introduced the idea to build abstract vectors and planar surface maps of buildings and ground, where the uncertainty is given by the normal distribution generated from vectors and planes. Building maps with fixed interval uncertainties have been used in hybrid interval-probabilistic localization [7]. Occupancy maps with uncertain occupied cells are also widely used in localization, whose uncertainties are represented by the occupancy probabilities.

All these existing studies prove the importance of quantifying the uncertainty of maps. However, it remains challenging to map the continuous space from noisy sparse point clouds and assign reasonable uncertainties to it. There is still a research gap in measuring the spatial uncertainty of maps in urban areas. There are many man-made structures with relatively regular geometrical shapes, where buildings are often the most important ones for localization, e.g. as used in [1], [7]. Thus, a detailed spatial uncertainty representation of building models in urban scenes is of great interest to explore.

The choice of an adequate uncertainty measure in environment mapping is critical. To tackle this problem, the uncertainty and completeness of a map can be described in a probabilistic fashion. GPs with their data-driven and probabilistic nature have proven to be a powerful tool for the quantification of uncertainty in various research fields. Recently, GPIS maps

[9], [10] have been proposed for continuous environmental mapping. The posterior mean and covariance kernel of a GP serve as the best estimation of the surface distance field and the uncertainty measure, respectively. GP Occupancy Mapping (GPOM) [12] demonstrates how to predict a continuous occupancy map from sparse noisy point clouds with GP inference. The occupancy probability of unexplored space with no rays passing through or hitting is regressed by spatial neighbors in a GP, including well-calibrated uncertainties. It provides the possibility to capture the underlying correlation between spatially neighboring cells. The continuous inference gives the generalizability to building a map of desired resolution for diverse applications. However, the drawback of GP lies in the high computational cost due to matrix inversion operations, and this drawback limits the applicability of GPs for large datasets. In order to achieve an acceptable computational efficiency, many approximation approaches such as the partitioning of the spatial world into subsets, the fusion of local GPs with Bayesian Committee Machines (BCMs) or a sparse approximation of GPs have been proposed [11]–[15]. To avoid the unreliable prediction from successive BCM updates, the Bayesian generalized kernel OctoMap (BGKOctoMap) [16] leverages Bayesian kernel inference and sparse kernels to perform a stable inference-based occupancy mapping. Jadidi et al. [17] expanded the occupancy mapping to probabilistic geometric frontiers computed efficiently using the gradient of the GP occupancy map.

Hilbert Maps, as another continous mapping technique, have been introduced by Ramos et al. [18], [19] to efficiently map the complex real world by operating on a high-dimensional feature vector. With efficient stochastic gradient descent optimization, it can achieve comparable performance to GP mapping with less time. Inspired by GP, Duong et al. [20] introduced a probabilistic formulation that utilizes a sparse set of relevance vectors to model obstacle boundaries.

However, these methods lack proper uncertainty quantification for their mapping. The uncertainty inherent in a GP map has not been sufficiently investigated, which can be exploited to distinguish unknown areas without enough information. This highlights the need to evaluate the reliability of the inferred occupancy probabilities. This issue has been addressed in recent work by Pearson [21]. Additionally, occasional discontinuities in the environment can pose challenges for GP-based mapping, as GPs are often used for continuous targets. This effect is not well-studied in the current literature on GP mapping either.

Another group of research is learning structural surface models from measurements to probabilistically represent the perceived space. Thrun et al. [22] fitted a set of rectangular flat surfaces to compose three-dimensional (3D) maps, with a group of parameters optimized by Expectation Maximization (EM) and the number of planes estimated by a Bayesian prior. With similar spirit but higher fidelity in approximating diverse arbitrary environments, Gaussian Mixture Models (GMMs) are employed as a semi-parametric tool with large numbers of components to extract the planar models [23] or obtain a high-fidelity representation of sensor observations [8], [24], [25], estimating Gaussian mixtures instead of 3D planes.

Nevertheless, the number of components has a great impact on the performance of GMM-based approaches and it is a non-trivial task to choose a proper number.

In this work, GP is used to obtain an inference-based surface model of structured buildings with GMM priors, leveraging the flexibility of non-parametric methods in arbitrary structures as well as the scalability of parametric approaches in large environments. When exploring large urban scenes, buildings are one of the most important structures for localization, which have relatively simple geometry and are easy to extract. GMMs are first applied to extract the main planar surfaces with a few components, while GPs are used subsequently to carve the irregular parts. In GP inference, the sets of local GPs corresponding to certain planar objects are applied to adapt the discontinuities and reduce redundant computation, while the sparse kernel and data block partitioning techniques similar to the prior work [13], [14] are utilized to further improve the computational efficiency.

The contributions of this work are the following:

- We map uncertain building models in a probabilistic manner using laser scanning point clouds, which can be applied in various downstream tasks and can be updated probabilistically. This method addresses the issue of a lack of uncertainty quantification for reference maps. The merit of this uncertainty representation is revealed in the experiments where better results are achieved when using accurate inference.
- GMM planes and local GP blocks are combined to yield a map with efficient surface prediction and to maintain a faithful uncertainty description. Both GMM planes and GP functions represent the statistical uncertainty of the mapping process.
- We evaluate the mapping performance and the faithfulness of the uncertainty description with real-world LiDAR point clouds, collected by a mobile mapping system (MMS). The Precision-Recall (PR) curve and the area under the curve (AUC) are used as metrics to evaluate the results of OctoMap, fast GPOM, BGKOctoMap, LARD-HM, GPIS, and our method. A comparison to the receiver operating characteristic (ROC) AUC is also discussed.

The rest of the letter is structured as follows. Section II defines the problem we aim to solve as well as the representation of building maps with uncertainty measures. In Section III, a brief introduction to GMMs and GPs is presented. We explain how GPs are used to model non-planar surfaces with a GMM prior, representing the uncertain building façade surface. The strategies for parameterizing the discontinuities, reducing computational efforts, and outlier filtering are explained in detail. In Section IV, experiments using real LiDAR data collected by MMS are presented, and qualitative results as well as a quantitative evaluation with state-of-the-art benchmarks are shown. Finally, we provide our conclusions and an outlook for future work in Section V. Figure 1 illustrates the overall workflow of the mapping process and evaluation.

II. PROBLEM STATEMENT

The objective is to produce maps for automotive robots in outdoor scenarios, which consist of building models including

Figure 1. Overall workflow.

quantitative uncertainty measures for each surface point, as illustrated in Figure 2. More specifically, the goal of the mapping problem is to estimate continuous façade surface: determining the relative surface depth and its uncertainty at any given query point.

Data sources are building points obtained from LiDAR range measurements. (Classification of the raw measurements to extract the building points is performed during a preprocessing step.) Due to the measurement process with a MMS, there are invisible areas occluded by other objects (cars, trees) or also self-occlusion. The resulting map should take these factors into account by properly describing the uncertainty measures. Such a representation also allows for integrating data from different sensors of different accuracy.

Although buildings are 3D objects, they can be modelled as being composed of a set of facades, each of which consists of planar protrusions, extrusions, slopes (non-vertical planes) and non-planar shapes, leading to a 2.5D representation.

Each individual façade of a building is then modelled with two parts: (1) regular planar parts are segmented by a GMM; (2) deviations from planar segments are modelled by GPs to capture arbitrary shapes. The planar parts can be modelled using plane parameters and boundary points indicating the location and the shape of each local planar facet. These parameters are optimized in GMM to provide a prior surface estimation. In this way, the main façade plane, as well as protrusions and extrusions can be specified with the corresponding plane layers resulting from a GMM. The remaining non-planar parts are represented as relative surfaces via non-parametric approaches, namely GP in our case, which describes the surface as a function with depth values of each point. Given a point $\boldsymbol{x} \in \mathbb{R}^2$ on the local projected surface plane, the distance s to the prior mean surface is:

$$s = f(\boldsymbol{x}) \begin{cases} = 0 & : \quad \text{on the mean surface,} \\ < 0 & : \quad \text{on an extrusion,} \\ > 0 & : \quad \text{on a protrusion.} \end{cases} \quad (1)$$

This GP surface representation provides an estimate of depth regarding the local planar segment at any location on the surface. Combining both components, the surface depth $d(\boldsymbol{x})$ is given by:

$$d(\boldsymbol{x}) = g(\boldsymbol{x}) + f(\boldsymbol{x}), \quad (2)$$

where $g(\boldsymbol{x})$ is the expected plane depth from the GMM and $f(\boldsymbol{x})$ denotes the local surface distance from Equation (1).

III. UNCERTAIN BUILDING SURFACE MODEL

In addition to the depth value, the uncertainty is represented probabilistically using GMM and GP components. Figure 3 visualizes the effect of each stage in the processing chain. To generate uncertain building models, the building points are first extracted by a semantic classification from calibrated MMS point clouds. For each individual building, each façade plane is extracted by using a Random Sample Consensus (RANSAC) approach [26], where points are allocated to the corresponding façades. Then, an uncertain model of façades can be explored with GMMs and local GPs, illustrated in Figure 3a, 3b, 3c, 3d and 3e. The ground truth of this exemplary building is provided for comparison, as shown in Figure 3f.

A. Gaussian Mixture Models for planar surfaces

A façade is mainly composed of local planar objects, which can be modelled by the plane normal, location, and the boundary of the facets. The plane normal, i.e. the orientation of the building façade, is determined by a principal component analysis (PCA), using the eigenvector belonging to the smallest component. It serves as the depth direction, which decides the local frame of the façade used in the following modelling. Many map representations such as 3D CityLoD2 models only map the building façade with one plane and ignore the extrusions and protrusions. Often, the uncertainty is not specified at all, or if it is, is given in terms of a single, global uncertainty measure only, e.g. as used in [7]. This is not sufficient for many tasks requiring high accuracy. A GMM is applied to cluster the points with different depth values to depth layers. Each depth layer is described by one Gaussian component in the GMM:

Figure 2. Depiction of a building model where for each location an uncertainty can be queried (indicated by color).

$$p(d) = \sum_{k=1}^{K} \pi_k \mathcal{N}(d|\mu_k, \sigma_k^2), \quad (3)$$

(a) Surface depth layers clustered by GMMs (5 components for each façade, i.e., 5 layers)

(b) Training points (red) in GPs

(c) Partitioning into local GPs

(d) Uncertainties of the estimated surface

(e) Estimated surface: sparse or occluded areas in (a) are inferred with large uncertainty.

(f) Ground truth

Figure 3. Visualization of different stages of the mapping process.

where d is the surface depth of a point, K being the number of Gaussian components, which is determined by the number of peaks detected in the histogram of all depth values; (π_k, μ_k, σ_k) are the GMM parameters for the k-th depth layer, π_k is the prior probability or weight, representing the importance of the component, and μ_k and σ_k are the mean and standard deviation of the Gaussian distribution.

The parameters $\{(\pi_k, \mu_k, \sigma_k)\}_{k=1}^K$ are optimized by maximizing the log-likelihood using the EM algorithm. Figure 4 shows an example of the GMM results. The Gaussian component with the highest prior probability is the main façade plane.

Nevertheless, the limits of GMM are: 1) Non-planar surfaces cannot be modelled. 2) If the number of points on different layers varies greatly, meaning the training for different components is not balanced, the method might fail to model the small protrusions and extrusions. 3) Slopes might be simply modelled as a vertical plane. The non-parametric approach GP, on the other hand, is capable to model them in a data-driven fashion. It is, therefore, introduced here to solve these problems.

Figure 4. GMM for plane modelling: the left picture shows a façade profile, i.e., all scanned points of a façade in a view from the side; The Gaussian components (right) with different colors represent each depth layer. The one with highest peak is the main façade plane.

B. Gaussian Processes

A GP $f(x) \sim GP(\mu(x), k(x, x'))$ is a popular Bayesian nonlinear regression method in machine learning. It can be seen as a multivariate normal distribution with infinite variables in a continuous domain, e.g. a continuous space. The distribution of any finite subset of those variables is still a joint Gaussian distribution. It is characterized by a mean function $\mu(x)$ and a covariance function $k(x, x')$, which is also called the kernel function. The prior GP mean function is often set to zero. There are many possible choices for the kernel function and in this letter, automatic relevance determination (ARD) is used as the prior GP kernel function:

$$k_0(\boldsymbol{x}, \boldsymbol{x}') = \sigma_p^2 \exp\left(-\frac{1}{2}\sum_{m=1}^2 b_m(x_m - x'_m)^2\right), \quad (4)$$

where $\boldsymbol{\theta} = (\sigma_p, b_1, b_2)$ are the hyper-parameters for the kernel function, which are optimized by maximum likelihood, given a measured dataset.

Provided a measurement dataset $\boldsymbol{D} = (\mathbf{X}, \mathbf{y})$, where $\mathbf{X} = \{\boldsymbol{x}_i\}_{i=1}^N$ are N input vectors and $\boldsymbol{y} = \{y_i\}_{i=1}^N$ are the corresponding target values measured with additive noise $\eta_i \sim \mathcal{N}(0, \sigma_\eta^2)$, we have the following likelihood:

$$p(\boldsymbol{y}|\mathbf{X}, \boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{y}|\mathbf{K}_N + \sigma_\eta^2 \mathbf{I}), \quad (5)$$

where $\mathbf{K}_N = k_0(\mathbf{X}, \mathbf{X})$ is the covariance matrix of the input points with $N \times N$ entries calculated by the prior kernel function.

The rest of the space $\mathbf{X}_*$ can be predicted with a multivariate Gaussian distribution conditioned on the observations $\boldsymbol{D}$. The posterior GP mean and kernel are calculated by [27]:

$$\mu(\boldsymbol{x}_*) = \mu_0(\boldsymbol{x}_*) + \mathbf{k}_*^T(\mathbf{K}_N + \sigma_\eta^2\mathbf{I})^{-1}(\boldsymbol{y} - \mu_0(\mathbf{X})), \quad (6)$$

$$k(\boldsymbol{x}_*, \boldsymbol{x}'_*) = k_0(\boldsymbol{x}_*, \boldsymbol{x}'_*) - \mathbf{k}_*^T(\mathbf{K}_N + \sigma_\eta^2\mathbf{I})^{-1}\mathbf{k}'_* + \sigma_\eta^2, \quad (7)$$

where the prior mean $\mu_0(\boldsymbol{x})$ is set to zero, $[\mathbf{k}_*]_N = k_0(\mathbf{X}, \boldsymbol{x}_*)$ and $[\mathbf{k}'_*]_N = k_0(\mathbf{X}, \boldsymbol{x}'_*)$ are the prior covariance between the N input training points and the predicting points.

To utilize a GP for generating an uncertain building model from dense point clouds, there are mainly two problems. First, mapping methods using GPs have to consider the discontinuity of the environment while the GP often encodes the continuity and smoothness assumptions on targets. It is hard to model functions, which exhibit arbitrarily large derivatives, with a GP. Unfortunately, large derivatives often appear in environment mapping. This issue is also addressed in [12], [14], where they tried special kernels to adapt the discontinuities, e.g. Matérn covariance functions. However, the improvement is limited. Second, with large training datasets, the GP regression faces a computational problem due to the inversion matrix calculation having time complexity $O(N^3)$. To capture the discontinuity and reduce the computational cost of GPs in the case of a building model, the two strategies applied here are: 1) using the GMM results as prior, and 2) partitioning data into blocks. These strategies are explained in detail in the following sections.

C. Using the GMM as prior

Using the GMM as prior offers several advantages. The GMM is used to "explain" the planar building parts. Thus, only the remaining arbitrary surfaces have to be modelled by GP, namely non-planar points, points of non-horizontal normals and outlier points too far away from the mean. In this way, the GMM information reduces the number of training points for GPs, as only the irregular points are used for GP training.

Another advantage of using GMM as prior is to provide parameters to the prior distributions of GPs. The estimated μ-values from the GMM serve as prior means in GP modelling, denoting the prior surface depth expectation of each planar layer, while the variances parameterise the prior covariance functions and depict the local data variability. The non-planar points which fall into a local facet on a certain layer are modelled by a local GP with this prior mean. The $g(\boldsymbol{x})$ in Equation (2) is substituted by the mean μ_k of a GMM component. Each local GP model has its own hyper-parameters and estimates the predictions only for the corresponding area. All local GPs compose a GP list for inference.

Different mean and covariance functions adapt GPs to the local facets, capturing and parameterising the discontinuities on the facade surface, caused by the sharp change between two depth layers. With this process, the entire continuous space is also divided into multiple local areas of GMM facets, modelled by GPs individually, which reduces the size of the kernel matrix as well.

There might be cases, where there are regions without any training data and not covered by any local GPs. Then, a global GP prior is included in our GP list. The mean value of this GP is determined by the mean of the Gaussian component with the largest weight in the GMM. Its prior variance can be set as the variance of the entire façade plane, representing the overall variability across all layers of the facade.

Initially, we considered weighted averaging of the inferred results from adjacent local GP segments in empty areas, using either BCM [28] or Product of Expert (PoE) [29]. However, in our case, the fusion did not significantly improve the results, and the resulting decrease in uncertainty was not desirable. Furthermore, the use of mixtures of GPs would have impacted efficiency. Therefore, in this letter, mixtures of local GPs are not utilized.

We classify the points that cannot be effectively modelled by the GMM into three categories: (1) non-planar points, (2) non-vertical slopes whose normal vectors deviates considerably from the normal of the main plane, and (3) points that are significantly outside their Gaussian distributions, i.e. not within the 95% confidence interval ($1.96\sigma_k$) of a certain component. Figure 3b illustrates an example, where the red points are selected as training points and model the GP implicit surface as a function of the continuous surface depth estimate. These points are selected according to the rules regarding the depth value d and the normal vector $\mathbf{n} = \{n_x, n_y, n_z\}$:

$$|d - \mu_k| > 1.96\sigma_k, \tag{8}$$

$$n_z(\boldsymbol{x}) < \cos(\alpha_z), \tag{9}$$

where μ_k and σ_k are the parameters of the k-th Gaussian layer, and $n_z(\boldsymbol{x})$ is the component of the normal $\mathbf{n}$, which is perpendicular to the main façade plane. α_z is the angle threshold for the normal $\mathbf{n}$ deviating from the normal of the assigned GMM plane. In this letter, $\alpha_z = 25°$ is applied.

The estimate for σ_p in the kernel of a local GP is set as the standard deviation of the corresponding planar patch σ_k. Optimization of other hyper-parameters is done on a small subset of the training data, involving a reasonable choice of the initial estimates:

- The initial $\{b_1, b_2\}$ are set to the same value, i.e. isotropic. We set $\frac{1}{b_1} = \frac{1}{b_2} = 0.0025\,m^2$, observed from the investigation of the impact of the length-scale $l^2 = \frac{1}{b_1} = \frac{1}{b_2}$ on kernels and the empirical value from the existing study [12];
- σ_η reflects the precision of the sensors, point cloud alignment noise and random environmental noise. In this case, $\sigma_\eta = 0.02\,m$.

D. Applying GP-block techniques

Although we reduced the training points in the last section, the computational efficiency can be further improved by the data partitioning and GP-block techniques, which are inspired by [13]. In practice, the spatial correlation between two points decreases with their distance. This property is naturally captured by the length-scale hyper-parameter $\boldsymbol{b} = (b_1, b_2)$ in the covariance functions of the GPs. As the points far from each other have negligible influence that can be ignored in the covariance matrix, only close points have significant covariance and need to be considered. An idea to exploit this property is to apply a sparse covariance function [15], by setting the covariance to zero when the distance is larger than a threshold. It also indicates the possible data partition of the training data. The entire training samples can be divided into local blocks within a certain range, as proposed in [13], [14].

In those methods, the block size is set as the predefined scale value of the sparse kernel while in our case, the block size is determined by the distance corresponding to the threshold c_{min} of covariance values, calculated by Equation (4), when σ_p is set to 1 and $b_1 = b_2 = b$:

$$\exp\left(-\frac{1}{2}\sum_{m=1}^{2} b(x_m - x'_m)^2\right) \geq c_{min}, \qquad (10)$$

where $c_{min} = 0.0001$ in this case. The block size is equal to $|\boldsymbol{x} - \boldsymbol{x}'|$ when it makes both sides equal. The extended block of a given block is defined as the central block with its neighboring blocks. In GP regression, the points from the extended block are the training data for the central block being under consideration.

E. Outlier filtering

To generate a GP inference robust to outliers, a model is required for outlier processing. There might be outliers in the training samples due to false classifications (e.g. points of road signs in the vicinity of a building) or other unexpected reasons. A statistical test on a GP with Gaussian likelihood is applied to check the probabilistic significance and identify the outliers, e.g. a Chi-squared test. In this work, the Chi-squared test with one degree of freedom is applied to remove outliers.

The procedure of outlier filtering is as follows.

1) The GP is trained normally with all training data and the mean and variance of each training point are inferred,
2) The Chi-squared test-statistic is calculated using $v_i^2 = (d_i - \mu_i)^2/\sigma_i^2$,
3) The test-statistics are compared to the χ^2 value given a certain α-quantile level and the points with test-statistic values larger than χ^2 value will be removed as outliers,
4) The GP is trained again with the filtered training data and the test points are predicted with mean and variance.

The above process may as well be repeated to iteratively trim the outliers, which, however, will lead to an increased computation time. Therefore, in this case, we only perform the test once in our experiment. The α-percentile level in the test serves as a threshold to filter outliers, indicating the strictness of outlier detection. High values of α will see many points as outliers, while a small α gives more tolerance to the points deviating from the mean and might miss the outliers.

IV. EXPERIMENTS AND EVALUATION

A. Dataset

The GP-based representation of a building surface is evaluated using real-world LiDAR point clouds collected by a Riegl VMX-250 mobile mapping system for the urban environment, in Hanover, Germany. The scanner measurements are specified as $5\,mm$ precision and $10\,mm$ accuracy. However, the accuracy of the GPS sensor in the system is much worse than the laser scanners. Point cloud alignment is thus used to correct the GPS trajectories in preprocessing and the data noise is given by a $2\,cm$ standard deviation, resulting from the alignment [30]. In the experiments, only building points were used, which were extracted using the deep learning based method [31].

In total, we used five measurement campaigns for the same urban areas collected on different dates to obtain very dense point clouds. 2,239,670 points from one measurement campaign were used for modelling the uncertain map and the remaining data (11,694,768 points) was used as ground truth for testing and evaluation. In one single measurement campaign, the number of points for one building façade is often larger than 10,000 and can reach up to 450,000 in our dataset. The investigated urban area is around $55 \times 93 \times 22\,m^3$.

B. Experiments

The building façades obtained from the RANSAC, are firstly transformed individually to the local frame of their main planes, where a map point is a 2.5-dimensional point $\mathcal{P} = \{\boldsymbol{x}, d\}$, with the location $\boldsymbol{x} \in \mathbb{R}^2$ on the plane and the depth value d in the direction of the plane normal (derived by the PCA using all façade points).

In the local façade frame, the planar points are used as i.i.d. samples to optimize the parameters of the GMM. Training points selected according to Equation (8) and (9) are used to derive the posterior GP-based models. For different building façades, the amount of training points varies. Depending on the building structures, 20 to 50% of the points are selected as training points in the experiment. The initial hyper-parameters are given based on the rules in Section III.C. The initial σ_η is chosen as $2\,cm$ for this dataset.

Figure 5 shows the qualitative result of modelled buildings with their associated uncertainties. In the original measurements, certain areas are occluded, where the inferred surfaces exhibit high uncertainties. Among the occluded regions, those located on simple façades with smaller prior variances, have lower uncertainties, since the variance of the entire façade is indicative of the potential surface variability, and hence, the possible errors. In general, the buildings that have been observed with denser measurements are modelled with lower uncertainties. In practical localization applications, users are able to select only the accurate part of maps or use the entire reference map with weights indicating the uncertainty. The benefits of only using maps with low uncertainties in occupancy maps are investigated in the following evaluation section.

(a) Surface inference   (b) Accurate inference

Figure 5. Qualitative results of the building models: uncertainties are denoted by colors. Redder colors indicate higher uncertainties while blue colors denote low uncertainties.

C. Evaluation

To evaluate the proposed method, existing state-of-the-art approaches are used as benchmarks, namely OctoMap [3], GPOM [14], BGKOctoMap [16], LARD-HM [18] and GPIS map [9]. To make the comparison possible, we first transform our abstract surface map to a 3D occupancy map, where probabilistic least-squares classification is applied and the occupancy probabilities are obtained through a cumulative Gaussian density function, inspired by previous research [12], [13]. The details can be found in the paper [13].

The metrics used for evaluation are 1) Precision-Recall (PR) curves and 2) Area under the curve (AUC). Note that some previous work used the ROC curve and the corresponding AUC as the metrics. However, it is not proper to use ROC in highly-skewed data, where a large change in the number of false positives can only lead to a small change in the false positive rate used in ROC [32]. This occurs in our case, as in outdoor scenarios, the number of free samples is much larger than the number of occupied ones. This issue has also been addressed in previous research [33]. Alternatively, the PR-curve is more suitable for highly imbalanced data [32], especially when the positive (occupied) case is of more concern. Precision compares the false positives with true positives, and therefore, captures the detection errors of a large number of false positives. Unlike the case with ROC curves, the goal of PR curves is to reach the upper right corner, instead of the upper left corner.

Figure 6 compares the AUC of different approaches in two resolutions ($0.1\,m$ and $0.05\,m$ grids). In general, GP-based maps have a good performance, among which our proposed method (green and orange curves) achieves the highest AUC values and is robust to small resolutions. The green curves denote the case when we only consider the accurate inference, i.e. areas with small uncertainty. The variance threshold is selected as 0.01 in this case. Areas with inference variance larger than 0.01 are considered being unknown. The figure shows that improvement is obtained by using only the accurate parts and the corresponding maps (green curves) give the best performance. This shows the value of the uncertainty representation for maps and indicates the potential application of this uncertainty model in downstream uncertainty-aware tasks.

The uncertainty evaluation is conducted by plotting AUC changes across varying uncertainty thresholds and distances of query points from the nearest observations. The proposed work is compared with the benchmark - GPIS map, shown in Figure 7. The results reveal a clear monotonic trend in the AUC values as the uncertainty threshold varies. The proposed method shows robustness compared to GPIS for the areas far from the observations, where the uncertainty is high. The AUC values remain consistently high, indicating strong performance. In contrast, GPIS exhibits a higher AUC only when a small number of predicted points, which are close to the original training points, are retained, as shown in Figure 7b. However, our method maintains a high AUC even with a larger number of predicted points, suggesting its ability to deliver reliable predictions for unobserved areas around observation points.

(a) Resolution: 0.1m   (b) Resolution: 0.05m

Figure 6. PR curves and AUC comparison: the AUC value of each curve is given in the legend.

(a)   (b)

Figure 7. Uncertainty evaluation: change of AUC with varying uncertainty thresholds and distances between the predicted points and original training points.

D. Comprehensive comparison

The GPOM and BGK models use fixed values for prior variance, noise, or length scale, which may not adequately support accurate spatial correlation and data variability. Additionally, the fixed prior variance and kernels struggle to model sharp changes in voxel occupancy, potentially impacting the final posterior occupancy probability. Similarly, the GPIS benchmark sets pre-fixed values for prior variance and mean, introducing inaccuracies in surface estimation, particularly for regions far from the observations. The fixed variance also leads to potential overestimation or underestimation of uncertainties. Thus, the results are sensitive to these fixed priors.

By contrast, our proposed method addresses these limitations by leveraging the GMM for obtaining a more tailored prior variance and mean, and employing local GPs for capturing the spatial correlations and uncertainties in a more adaptive manner. This allows for a more accurate estimation of façade surfaces.

E. Computational time comparison

Regarding computational time, BGK and GPOM reach the fastest speed for mapping, as shown in Figure 8. Our proposed method demonstrates comparable training times to benchmarks, thanks to the assistance of the GMM prior. During testing, it may require more time than BGK, GPOM, and LARD-HM but remains faster than GPIS. While we acknowledge the importance of efficient online mapping, our goal is to deliver a mapping framework that prioritizes accuracy and uncertainty awareness, enhancing map-based localization tasks.

(a) Training time   (b) Testing time

Figure 8. Computational time.

V. CONCLUSIONS

In summary, this letter proposed a method to generate building models as reference maps with associated uncertainties. The GMM and GP approaches are combined to map the planar parts and non-planar surfaces together. Local GP and data partition techniques are introduced in GP to reduce the computational cost and parameterise the discontinuity of the building surface. In the experiments, the abstract surface map is transformed into an occupancy map representation and then is compared with state-of-the-art benchmarks. Our evaluation shows that the local GP with GMM prior yields the highest AUC values, indicating its potential for utilizing uncertainty characterizations in future applications.

The generated map will be applied for localization tasks in the future and the positioning results will be compared with other methods, such as NDT-based localization. The approach also has a large potential to be used for updating and refining existing models from new data or other sensors; this will also be studied in future work.

REFERENCES

[1] E. Javanmardi, Y. Gu, M. Javanmardi and S. Kamijo, "Autonomous vehicle self-localization based on abstract map and multi-channel LiDAR in urban area", *IATSS research*, 43(1), pp.1-13, 2019.
[2] C. Zhang, L. Liu, Z. Xue, K. Guo, K. Yang, R. Cai and Z. Li, "Robust LiDAR Localization on an HD Vector Map without a Separate Localization Layer," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 5536-5543.
[3] A. Hornung, K.M. Wurm, M. Bennewitz, C. Stachniss and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous robots*, 34(3), pp.189-206, 2013.
[4] R. Wang, J. Peethambaran and D. Chen, "LiDAR Point Clouds to 3-D Urban Models: A Review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(2), pp. 606-627, 2018.
[5] Y. Huang and K. Gupta, "RRT-SLAM for motion planning with motion and map uncertainty for robot exploration," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008, pp. 1077-1082.
[6] P. Biber and W. Straßer, "The normal distributions transform: A new approach to laser scan matching," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2003, pp. 2743-2748.
[7] A. Ehambram, L. Jaulin and B. Wagner, "Hybrid Interval-Probabilistic Localization in Building Maps," *IEEE Robotics and Automation Letters*, 7(3), pp. 7059-7066, 2022.
[8] S. Srivastava and N. Michael, "Efficient, multifidelity perceptual representations via hierarchical gaussian mixture models," *IEEE Transactions on Robotics*, 35(1), pp.248-260, 2018

[9] B. Lee, C. Zhang, Z. Huang and D. D. Lee, "Online Continuous Mapping using Gaussian Process Implicit Surfaces," in *IEEE International Conference on Robotics and Automation*, 2019, pp. 6884-6890.
[10] J. -P. A. Ivan, T. Stoyanov and J. A. Stork, "Online Distance Field Priors for Gaussian Process Implicit Surfaces," in *IEEE Robotics and Automation Letters*, 7(4), pp. 8996-9003, 2022.
[11] K. Mouhyemen, A. Patel, and A. Chatterjee. "Multi-sparse gaussian process: Learning based semi-parametric control," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020, pp. 5327-5334.
[12] S.T. O'Callaghan and F.T. Ramos. "Gaussian process occupancy maps," *The International Journal of Robotics Research*, 31(1), 42-62, 2012.
[13] S. Kim and J. Kim, "GPmap: A unified framework for robotic mapping based on sparse Gaussian processes," *Field and service robotics*, pp. 319-332, 2015.
[14] J. Wang and B. Englot, "Fast, accurate gaussian process occupancy maps via test-data octrees and nested bayesian fusion," in *IEEE International Conference on Robotics and Automation*, 2016, pp. 1003-1010.
[15] A. Melkumyan and F.T. Ramos, "A sparse covariance function for exact Gaussian process inference in large datasets," in *international joint conference on artificial intelligence*, 2009.
[16] K. Doherty, J. Wang and B. Englot, "Bayesian generalized kernel inference for occupancy map prediction," in *IEEE International Conference on Robotics and Automation*, 2017, pp. 3118-3124.
[17] M. G. Jadidi, J. V. Miró, R. Valencia and J. Andrade-Cetto, "Exploration on continuous Gaussian process frontier maps," in *2014 IEEE International Conference on Robotics and Automation*, 2014, pp. 6077-6082.
[18] R. Senanayake and F. Ramos, "Bayesian hilbert maps for dynamic continuous occupancy mapping," in *Conference on Robot Learning (CoRL)*, 2017, pp. 458–471.
[19] V. Guizilini and F. Ramos, "Large-scale 3D scene reconstruction with hilbert maps," in *IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 3247–3254.
[20] T. Duong, M. Yip and N. Atanasov, "Autonomous Navigation in Unknown Environments With Sparse Bayesian Kernel-Based Occupancy Mapping," *IEEE Transactions on Robotics*, 38(6), pp. 3694-3712, 2022.
[21] E. Pearson, K. Doherty and B. Englot, "Improving obstacle boundary representations in predictive occupancy mapping," *Robotics and Autonomous Systems*, 153, p.104077, 2022.
[22] S. Thrun, C. Martin, Y. Liu, D. Hahnel, R. Emery-Montemerlo, D. Chakrabarti and W. Burgard, "A real-time expectation-maximization algorithm for acquiring multiplanar maps of indoor environments with mobile robots," *IEEE Transactions on Robotics and Automation*, 20(3), pp.433-443, 2004.
[23] R.T. Marriott, A. Pashevich and R. Horaud, "Plane-extraction from depth-data using a Gaussian mixture regression model," *Pattern Recognition Letters*, 110, pp.44-50, 2018.
[24] C. O'Meadhra, W. Tabib and N. Michael, "Variable resolution occupancy mapping using gaussian mixture models," *IEEE Robotics and Automation Letters*, 4(2), pp.2015-2022, 2018.
[25] W. Tabib, K. Goel, J. Yao, M. Dabhi, C. Boirum and N. Michael, "Real-Time Information-Theoretic Exploration with Gaussian Mixture Model Maps," in *Robotics: Science and Systems*, pp. 1-9, 2019.
[26] R. Schnabel, R. Wahl and R. Klein, "Efficient RANSAC for point-cloud shape detection," in *Computer graphics forum*, 26(2), pp. 214-226, 2007.
[27] C. E. Rasmussen, "Gaussian processes in machine learning," in *Summer School on Machine Learning*, Berlin, Germany, 2003, pp. 63–71.
[28] H. Liu, J. Cai, Y. Wang, YS. Ong, "Generalized robust Bayesian committee machine for large-scale Gaussian process regression," in *International Conference on Machine Learning*, 2018, pp. 3131-3140.
[29] H., Geoffrey E. "Training products of experts by minimizing contrastive divergence," *Neural Computation*, 14 (8):1771–1800, 2002.
[30] C. Brenner, "Scalable estimation of precision maps in a mapreduce framework," in *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM Press, New York, USA, 2016, pp. 1-10.
[31] T. Peters and C. Brenner, "Automatic Generation of Large Point Cloud Training Datasets Using Label Transfer." in *39th Annual Scientific and Technical Conference of the DGPF eV*, in Vienna, Austria., 2019.
[32] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proceedings of the 23rd international conference on Machine learning*, 2006, pp. 233-240.
[33] Y. Miao, A. Hunter and I. Georgilas, "Parameter reduction and optimisation for point cloud and occupancy mapping algorithms," *Sensors*, 21(21), p.7004, 2021.