A Diversified Generative Latent
Variable Model for WiFi-SLAM

Hao Xiong, Dacheng Tao

Centre for Artificial Intelligence, University of Technology Sydney, Australia
hao.xiong@student.uts.edu.au
dacheng.tao@uts.edu.au

Abstract

WiFi-SLAM aims to map WiFi signals within an unknown environment while simultaneously determining the location of a mobile device. This localization method has been extensively used in indoor, space, undersea, and underground environments. For the sake of accuracy, most methods label the signal readings against ground truth locations. However, this is impractical in large environments, where it is hard to collect and maintain the data. Some methods use latent variable models to generate latent-space locations of signal strength data, an advantage being that no prior labeling of signal strength readings and their physical locations is required. However, the generated latent variables cannot cover all wireless signal locations and WiFi-SLAM performance is significantly degraded. Here we propose the diversified generative latent variable model (DGLVM) to overcome these limitations. By building a positive-definite kernel function, a diversity-encouraging prior is introduced to render the generated latent variables non-overlapping, thus capturing more wireless signal measurements characteristics. The defined objective function is then solved by variational inference. Our experiments illustrate that the method performs WiFi localization more accurately than other label-free methods.

1 Introduction

Robot localization plays a crucial role in applications such as activity recognition, surveillance, and content-aware computing. Owing to the portability, low price, and high accuracy of WiFi access points, WiFi-based simultaneous localization and mapping (SLAM) has been extremely useful for mobile device localization. Exploiting wireless signal strength for robot localization is a relatively novel but increasingly useful research field. Here we propose a new localization approach that utilizes WiFi access points and aims to improve robot localization.

There is no definite link between the WiFi signal strength and location of a mobile device; therefore, it is difficult to generate a good likelihood model of signal strength measurements that can be used to accurately predict the physical location of a mobile device. Most methods seek to develop robust likelihood models from calibration data collected in a given environment, where calibration data refers

to a set of wireless signal strength measurements associated with their ground truth location from a mobile device. Such approaches have proven to be extraordinarily effective for robot localization.

Nevertheless, these methods are highly dependent on labeled ground truth location data, restricting their flexibility and applicability. Put another way, mobile devices cannot be localized based on raw, unlabeled signal strength data. Although SLAM methods have been developed that simultaneously estimate the location and map an environment, the nature of wireless signal strength prohibits the use of standard SLAM in this scenario.

Ferris *et al.* (Ferris, Fox, and Lawrence 2007) presented Gaussian process latent variable model (GPLVM)-based WiFi-SLAM to perform localization without the need for ground truth-labeled physical locations. They projected high-dimensional signal strength data into two-dimensional latent variables, where the two-dimensional latent variables were considered the x-y-coordinates of the robot. GPLVM-based WiFi-SLAM also incorporated three types of constraints: locations→signal strengths, signal strengths→locations and locations→locations. As a consequence, this technique produced a more accurate signal strength map that could be further used to estimate the physical locations of mobile devices.

WiFi-SLAM employs GPLVMs (Lawrence 2004) to project signal strength measurements into a simpler latent structure. Then, the inferred two-dimensional latent structure is used to find the x-y coordinates of the WiFi access points. In practice, the WiFi access points constantly send signals from widely distributed locations. Therefore, GPLVM is needed to generate the latent variables of the x-y coordinates that span the WiFi signal space as diversely as possible. However, the GPLVMs in WiFi-SLAM create highly redundant latent structures without taking diversity into account. This is because GPLVM places large probability masses on the places where most WiFi access points are located and, conversely, small probability masses on the places where the wireless signal are weak.

Rather than creating latent variables with preferences and redundancy, a unique group of latent variables would ideally be created on the full characteristics of the signal strength measurements. In this paper, we propose DGLVM to place a diversity-encouraging prior on the independent latent dis-

tributions. To achieve this, we construct a probability product kernel to measure the similarities between the distributions of latent variables. A probability diversity prior is then obtained by adding the determinant on the constructed product kernel, where the determinant of this Gram matrix ensures greater diversity in the created latent variables to span all possible WiFi signals areas. To solve the objective function, variational inference is first exploited to obtain a lower bound and, after maximizing the lower bound, the solution is derived. We experiment on a set of wireless signal measurements with ground truth locations recorded. Our results demonstrate the robustness and effectiveness of this approach for localization.

We first review existing works on SLAM techniques in Section 2. A brief introduction to GPLVM-based WiFi-SLAM is provided in Section 3, before demonstrating our proposed model in Section 4. Finally, we perform experiments and evaluate our method in Section 5. We conclude in Section 6.

2 Related Work

SLAM techniques can be broadly classified into several categories: signal-based SLAM, laser SLAM, and visual SLAM.

In signal-based SLAM, the ground truth locations of the training signal strength measurements are probably unknown. Therefore, WiFi-SLAM (Ferris, Fox, and Lawrence 2007) was proposed to project the signal strength measurements into low-dimensional latent variables, where the latent variables are regarded as the x-y locations of the signal strength. For dimension reduction, approaches such as GPLVM (Lawrence 2004) and action respecting embeddings (ARE) (M. Bowling and Milstein 2005) can be employed for non-linear dimension reduction. Furthermore, since successive signal strength readings are based on a time series, the time-series constraint is introduced into Gaussian process dynamical models (Wang, Fleet, and Hertzmann 2006). However, this localization method is inaccurate when only signal measurement information is available. Other methods seek to label ground truth locations on the signal strength measurements for training. For instance, (Bahl and Padmanabhan 2000; Haeberlen et al. 2004; Letchner, Fox, and LaMarca 2005) directly model signal propagation through space. Conversely, rather than modeling signal propagation, (Haeberlen et al. 2004; Letchner, Fox, and LaMarca 2005) model the signal strength measurements at locations of interest. In reality it transpires that obstacles like walls and furnitures exist in the space that compromise modeling accuracy.

In laser-based SLAM, some methods apply Kalman filter-based techniques to address SLAM issues. For instance, extended Kalman filters effectively correlate motion and sensor models to estimate the landmark maps and robot poses. However, one drawback of such techniques is that the sensor model is non-linear, so the estimated pose is less accurate. Therefore, unscented Kalman filter SLAM (Lee, Lee, and Kim 2006; Kim, Sakthivel, and Chung 2008; Kim, Kim, and Chung 2011) approaches have been proposed to improve pose and map estimation. Other works have focused on particle-filter based SLAM such as Rao-Blackwellized particle filters (RBPF), in which each particle in RBPF represents a possible robot trajectory and a map. Some RBPF extensions (Grisetti, Stachniss, and Burgard 2007; 2005) aim to describe more efficient map representations and reduce the number of particles to improve the motion model.

Visual SLAM overcomes the limitations of SLAM with the aid of RGB-D, stereo, or monocular cameras. By jointly combining visual feature points and depth measurements, most RGB-D SLAM methods are likely to estimate the transformations between successive frames with iterative cloud points (Besl and McKay 1992; Rusinkiewicz and Levoy 2001; Fitzgibbon 2003) and PnP estimation (Lowe 2004; Bay et al. 2008; Rublee et al. 2011). They also employ relocalization techniques in case the robot revisits the same places. These methods inevitably employ the g^2o framework (Kuemmerle et al. 2011) for global optimization and to obtain the robot poses. The early monocular SLAM employed filtering techniques (Davison et al. 2007; Civera, Davison, and Montiel 2008; Chiuso et al. 2002) to simultaneously map and localize a camera. However, these methods turned out to be time consuming and error-prone. Therefore, other works (Mouragnon et al. 2006; Klein and Murray 2007; Strasdat, Montiel, and Davison 2012; Engel, Schops, and Cremers 2014) perform SLAM using a small number of key frames and, in doing so, simultaneously accomplish camera tracking and mapping in real time. To further enhance accuracy, the relocalization technique (Klein and Murray 2008) has been employed using the low-resolution thumbnails of the key frames.

3 GPLVM for WiFi-SLAM

In this section, we describe the basic concepts of GPLVM (Lawrence 2004). In doing so, we illustrate how GPLVM can be used to estimate a likelihood model of wireless signal strength measurements for WiFi-SLAM.

In the case of WiFi-SLAM techniques, let $Y \in R^{n \times p}$ (with columns $\{y_{:,j}\}_{j=1}^{p}$) be the wireless signal strength measurements, where n, p are the number and dimensionality of each wireless signal strength measurement, respectively. Wireless signal strength measurements are supposed to have their own locations. Since the ground truth locations are unknown, the low-dimensional latent variables $X \in R^{n \times q}$ in GPLVM can be used to denote these locations in the 2D plane. Then, we have:

$$p(Y \mid X) = \prod_{j=1}^{p} p(y_{:,j} \mid X), \qquad (1)$$

where $y_{:,j}$ represents the jth column of Y and

$$p(y_{:,j} \mid X) = \mathcal{N}(y_{:,j} \mid 0, K_{ff} + \sigma^{-1}I_n). \qquad (2)$$

Here, the covariance matrix K_{ff} refers to a kernel function. Since the observed data are noisy, we take the Gaussian observation noise variance σ^{-1} into consideration. The covariance between two latent variables is defined using the expo-

nentiated quadratic (RBF)-based kernel:

$$k(x_{i,:}, x_{k,:}) = \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{j=1}^{q}\alpha_j(x_{i,j} - x_{k,j})^2\right), \quad (3)$$

where σ_f^2 is the signal variance, $x_{i,:}$ is the ith row of X, and $K_{ff} = k(x_{i,:}, x_{k,:})$. Moreover, α_j is a length scale that determines the strength of the correlation between locations. Both σ_f^2 and α_j are parameters that control the smoothness of the kernel functions.

An example GP signal strength model for one access point is shown in Fig. 1. Based on the GPLVM framework, Ferris *et al.* (2007) proposed the following constraints to render WiFi-SLAM more robust.

Location → signal strength In WiFi-SLAM, similar locations are supposed to have similar signal strengths. This property can essentially be controlled by the kernel function in Eq. (3) because it measures the similarity between latent variables.

Signal Strength → location Likewise, similar signal strengths should correspond to similar locations. GPLVM-based WiFi-SLAM applied back constraints (Lawrence and Quinonerocandela 2006) to ensure that the close points in Y are also close in X.

Location → location Due to the fact that the signal measurements are collected in sequence, the dynamic prior is incorporated into the model. Consequently, locations that follow each other during collection should be near each other in the generated latent space.

Then, the likelihood function that models the observed wireless signal strength measurements is:

$$p(Y) = \int p(Y \mid X)p(X)\mathrm{d}X. \quad (4)$$

The scaled conjugate gradient (SCG) is exploited to solve the objective function. The interested readers should refer to (Lawrence 2004) for more details.

4 The Diversified Generative Latent Variable Model

In WiFi-SLAM, Ferris *et al.* (2007) incorporated GPLVM with three constraints to project the signal measurements into low-dimensional latent variables. These latent variables are regarded as physical locations of the mobile devices. However, the latent variables generated in WiFi-SLAM are highly redundant and overlapping, reducing the capacity of the model to capture all possible wireless signal areas.

We propose the diversified generative latent variable model (DGLVM). We construct a diversity prior on a set of component-specific distributions $q(x_i|\phi_i)_{i=1}^{I}$, where x_i refers to a latent variable. The functionality of distributions $q(x_i|\phi_i)_{i=1}^{I}$ are introduced below. Here, a positive-definite kernel K_ϕ is constructed between the distributions $q(x_i|\phi_i)_{i=1}^{I}$, such that these distributions are non-overlapping and are even inclined to be more diverse. Furthermore, variational inference is exploited to solve the objective function. The created latent variables x_i should explain substantially different regions and lead to enhanced,
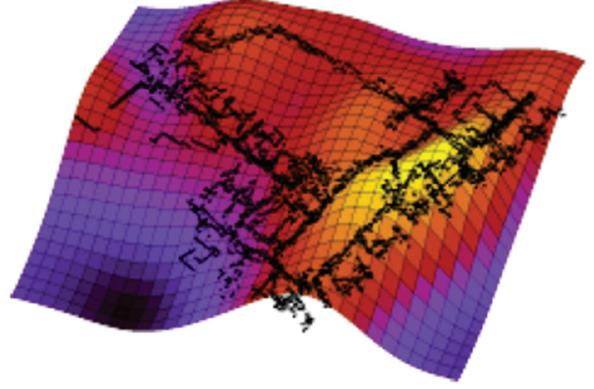

Figure 1: The signal strength propagation of one access point modeled by GPLVM (Ferris, Fox, and Lawrence 2007). The black trace is the ground truth path.

non-redundant feature extraction of wireless signal measurements (shown in Fig. 2).

The positive-definite kernel $K : \Phi \times \Phi \to R$ is defined on a space Φ. The whole process aims to select a random subset of possible ϕ from Φ. Meanwhile, the probability density associated with a particular finite $\phi \in \Phi$ is defined as:

$$p(\phi \in \Phi) = |K_\phi|, \quad (5)$$

where $|K_\phi|$ is a positive-definite Gram matrix and ϕ includes all the parameters of the distributions $q(x_i|\phi_i)_{i=1}^{I}$.

Therefore, the objective function is:

$$F(\theta) = \log P(Y) + \lambda \log |K_\phi|, \quad (6)$$

where $\theta = \{\phi_i, \sigma_f^2, \sigma, \alpha_1, ..., \alpha_j\}$ are the hyperparameters in our proposed model and some of them use the same symbols as in GPLVM. Furthermore, λ is used to balance the weights between measurements of likelihood and the diversity-encouraging prior.

The Kernel-based Diversity Prior

In our model, the kernel K_ϕ consists of a positive-definite correlation function $R(\phi_i, \phi_j)$ and the "prior kernel" $\sqrt{\pi(\phi_i)\pi(\phi_j)}$. Then, the kernel can be expressed as:

$$K(\phi_i, \phi_j) = R(\phi_i, \phi_j)\sqrt{\pi(\phi_i)\pi(\phi_j)}, \quad (7)$$

where $R(\phi_i, \phi_i) = 1$.

Here, the kernel-based diversity prior allows for repulsion, where we use the probability product kernel to construct each element in the kernel matrix to define the repulsion. Therefore, every kernel element is expressed as the inner product of probability distributions:

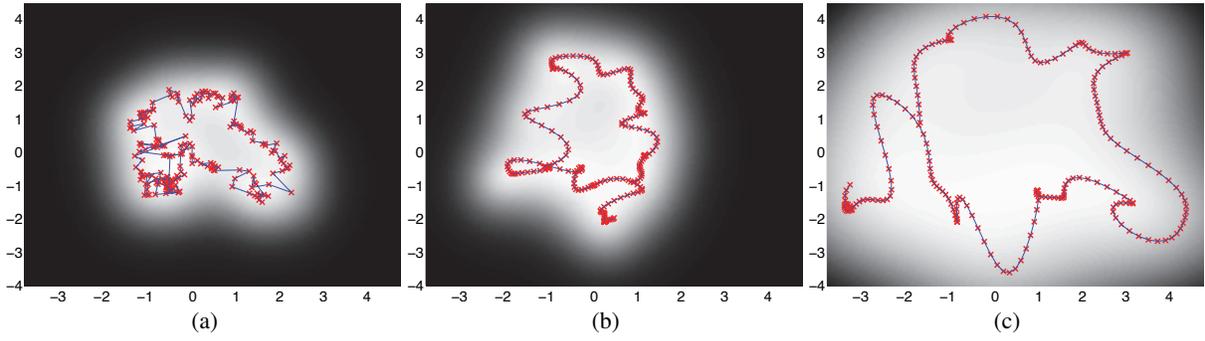$$K(\phi_i, \phi_j; \rho) = \int_\chi f(x|\phi_i)^\rho f(x|\phi_j)^\rho dx, \quad (8)$$

Figure 2: Latent variables generated by our model with different prior types on a set of wireless signal measurements. (a) Without dynamic and diversity-encouraging priors. (b) With dynamic prior only. (c) With both dynamic and diversity-encouraging priors.

where $\rho > 0$.

In the general case of the Gaussian distributions $f_1 = N(\mu_1, \Sigma_1)$ and $f_2 = N(\mu_2, \Sigma_2)$, the product kernel is:

$$K(f_1, f_2) =$$
$$\int \frac{1}{\sqrt{(2\pi)^{\rho D}|\Sigma_1|^{\rho}}} \exp\left(-\frac{\rho}{2}(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1)\right)$$
$$\frac{1}{\sqrt{(2\pi)^{\rho D}|\Sigma_2|^{\rho}}} \exp\left(-\frac{\rho}{2}(x-\mu_2)^T\Sigma_2^{-1}(x-\mu_2)\right) dx,$$
$$(9)$$

where D refers to the dimensionality of x. For simplicity, let $\rho = 1$, the probability product kernel becomes:

$$K(f_1, f_2) = (2\pi)^{-\frac{D}{2}}|\hat{\Sigma}|^{\frac{1}{2}}(|\Sigma_1||\Sigma_2|)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\Big(\right.$$
$$\left. \mu_1^T\Sigma_1^{-1}\mu_1 + \mu_2^T\Sigma_2^{-1}\mu_2 - \hat{\mu}^T\hat{\Sigma}\hat{\mu}\Big)\right), \quad (10)$$

where $\hat{\Sigma} = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}$ and $\hat{\mu} = \Sigma_1^{-1}\mu_1 + \Sigma_2^{-1}\mu_2$.

In our approach, the probability product kernel is employed to measure the variational distributions $f_1 = q(x_i|\mu_i, S_i)$ and $f_2 = q(x_j|\mu_j, S_j)$. Note that the covariance matrices S_i and S_j are diagonal matrices. Then, we have:

$$K(f_1, f_2) = (2\pi)^{-\frac{D}{2}} \left(\prod_{d=1}^{D} \frac{1}{S_{id} + S_{jd}}\right)^{\frac{1}{2}} \exp\Big($$
$$-\frac{1}{2}\sum_{d=1}^{D} \frac{(\mu_{id} - \mu_{jd})^2}{S_{id} + S_{jd}}\Big). \quad (11)$$

Here, the normalized variant $R(\phi_i, \phi_j)$ can be derived as follows:

$$R(f_1, f_2) = K(f_1, f_2)/\sqrt{K(f_1, f_1)K(f_2, f_2)}. \quad (12)$$

Then, the correlation kernel becomes:

$$R(f_1, f_2) = \left(\prod_{d=1}^{D} \frac{2\sqrt{S_{id}S_{jd}}}{S_{id} + S_{jd}}\right)^{\frac{1}{2}} \exp\Big($$
$$-\frac{1}{2}\sum_{d=1}^{D} \frac{(\mu_{id} - \mu_{jd})^2}{S_{id} + S_{jd}}\Big). \quad (13)$$

Now, the product kernel for the diversity prior is constructed. We move on to introduce how to solve the proposed objective function.

Variational Inference

In our proposed model, we wish to maximize the $F(\theta)$

$$F(\theta) = \log \int p(Y, F, U, X) \mathrm{d}X\mathrm{d}F\mathrm{d}U + \log|K_\phi|^\lambda. \quad (14)$$

Here, the inducing points U are to render the objective function solvable. Then, the log likelihood function can be further factorized as:

$$F(\theta) = \log \int \prod_{j=1}^{p} p(y_{:,j}|f_{:,j})\Big(p(f_{:,j}|u_{:,j}, X)p(X)\mathrm{d}X\Big)$$
$$p(u_{:,j})\mathrm{d}U\mathrm{d}F + \log|K_\phi|^\lambda. \quad (15)$$

Note that integration over X is unfeasible since X is an input, in a rather complex non-linear manner, of $p(f_{:,j}|u_{:,j}, X)$, which contains the kernel matrix K_{ff}.

Thus, the variational distribution $q(F, U, X)$ is applied to approximate the true posterior $P(F, U, X|Y)$ with the form:

$$q(F, U, X) = \left(\prod_{j=1}^{p} p(f_{:,j}|u_{:,j}, X)q(u_{:,j})\right)q(X). \quad (16)$$

Here, $q(X)$ is a variational distribution that follows:

$$q(X) = \prod_{i=1}^{n} q(x_{i,:}|\phi_i) = \prod_{i=1}^{n} N(x_{i,:}|\mu_{i,:}, S_i), \quad (17)$$

where the variational distribution q(X) is exploited to construct the diversity prior in the last section and each covariance matrix S_i is diagonal.

In terms of Jensen's inequality, the lower bound $F\big(q(X), q(U)\big)$ of the objective function can be derived by:

$$F\big(q(X), q(U)\big) =$$
$$\int q(F, U, X) \log \frac{p(Y)}{q(F, U, X)} \mathrm{d}X\mathrm{d}F\mathrm{d}U + \log|K_\phi|^\lambda. \quad (18)$$

After inserting Eq. (16) into Eq. (18), we have:

$$F\big(q(X), q(U)\big) = \int \prod_{j=1}^{p} p(f_{:,j}|u_{:,j}, X)q(u_{:,j})q(X) \Bigg($$

$$\log \frac{\prod_{j=1}^{p} p(y_{:,j}|f_{:,j})p(u_{:,j})p(X)}{\prod_{j=1}^{p} q(u_{:,j})q(X)}\Bigg)\mathrm{d}X\mathrm{d}F\mathrm{d}U + \log |K_\phi|^\lambda.$$

$$(19)$$

For simplicity, let $\langle \cdot \rangle_p$ be shorthand for the expectation with respect to the distribution p. Then:

$$F(q(X), q(U)) =$$

$$\sum_{j=1}^{p} \Bigg(\int q(u_{:,j})q(X) \langle \log p(y_{:,j}|f_{:,j}) \rangle_{p(f_{:,j}|u_{:,j},X)} \mathrm{d}X \mathrm{d}u_{:,j}$$

$$+ \langle \log \frac{p(u_{:,j})}{q(u_{:,j})} \rangle_{q(u_{:,j})} \Bigg) - KL\big(q(X)||p(X)\big) + \lambda \log |K_\phi|.$$

$$(20)$$

To simplify the sum term in Eq. (20), let the formulation within the sum notation be denoted by:

$$\hat{F}_j\big(q(X), q(U)\big) = \langle \log \frac{p(u_{:,j})}{q(u_{:,j})} \rangle_{q(u_{:,j})}$$

$$+ \int q(u_{:,j})q(X) \langle \log p(y_{:,i}|f_{:,j}) \rangle_{p(f_{:,j}|u_{:,j},X)} \mathrm{d}X \mathrm{d}u_{:,j}.$$

$$(21)$$

Now, the lower bound $F\big(q(X), q(U)\big)$ consists of three parts: $\lambda \log |K_\phi|$, $\hat{F}_j\big(q(X), q(U)\big)$, and KL divergence $KL\big(q(X)||p(X)\big)$.

We then move on to compute $KL\big(q(X)||p(X)\big)$. Since both $q(X)$ and $p(X)$ are Gaussian distributions, the KL term can be easily calculated:

$$KL\big(q(X)||p(X)\big) = \frac{1}{2}\sum_{i=1}^{n} tr\big(\mu_{i,:}\mu_{i,:}^T + S_i - \log S_i\big) - \frac{nq}{2}.$$

$$(22)$$

To derive $\hat{F}_j\big(q(X), q(U)\big)$, it is essential to compute $\langle \log p(y_{:,i}|f_{:,j}) \rangle_{p(f_{:,j}|u_{:,j},X)}$.

$$\langle \log p(y_{:,i}|f_{:,j}) \rangle_{p(f_{:,j}|u_{:,j},X)} = \log N(y_{:,j}|K_{fu}K_{uu}^{-1}u_{:,j}, \sigma^2 I_n)$$

$$- \frac{1}{2\sigma^2} tr(K_{ff} - K_{fu}K_{uu}^{-1}K_{uf}).$$

$$(23)$$

After inserting Eq. (23) into $\hat{F}_j\big(q(X), q(U)\big)$, $\hat{F}_j\big(q(X), q(U)\big)$ can be easily derived and expressed as:

$$\hat{F}_j(q(X), q(U)) =$$

$$\frac{1}{2\sigma^2} tr\big(\langle K_{ff} \rangle_{q(X)}\big) - \frac{1}{2\sigma^2} tr\big(K_{uu}^{-1}\langle K_{uf}K_{fu} \rangle_{q(X)}\big)$$

$$+ \int q(u_{:,j}) \log \frac{e^{\langle \log N(y_{:,j}|a_j, \sigma^2 I_n) \rangle_{q(X)}} p(u_{:,j})}{q(u_{:,j})} \mathrm{d}u_{:,j}.$$

$$(24)$$

$\hat{F}_j(q(X), q(U))$ can be upper bounded by $\hat{F}_j(q(X))$ after applying the reversing Jensen's inequality to the KL-like quantity containing $q(u_{:,j})$:

$$\hat{F}_j(q(X)) =$$

$$\frac{1}{2\sigma^2} tr\big(\langle K_{ff} \rangle_{q(X)}\big) - \frac{1}{2\sigma^2} tr\big(K_{uu}^{-1}\langle K_{uf}K_{fu} \rangle_{q(X)}\big)$$

$$+ \log \int e^{<\log N(y_{:,j}|a_j, \sigma^2 I_p)>_{q(X)}} p(u_{:,j})\mathrm{d}u_{:,j}. \quad (25)$$

Now $q(U)$ is optimally eliminated, $\hat{F}_j(q(X))$ can be calculated as follows:

$$\hat{F}_j(q(X)) = \Bigg[\log \frac{\sigma^{-n}|K_{uu}|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}|\sigma^{-2}\psi_2 + K_{uu}|^{\frac{1}{2}}} e^{-\frac{1}{2}y_{:,j}^T W y_{:,j}} \Bigg]$$

$$- \frac{\psi_0}{2\sigma^2} + \frac{1}{2\sigma^2} tr(K_{uu}^{-1}\psi_2), \quad (26)$$

where $\psi_0 = tr\big(\langle K_{ff} \rangle_{q(X)}\big)$, $\psi_1 = \langle K_{fu} \rangle_{q(X)}$, $\psi_2 = \langle K_{uf}K_{fu} \rangle_{q(X)}$, and $W = \sigma^{-2}I_n - \sigma^{-4}\psi_1(\sigma^{-2}\psi_2 + K_{uu}^{-1})\psi_1^T$.

With $\hat{F}_j\big(q(X), q(U)\big)$ and $KL\big(q(X)||p(X)\big)$ in hand, we can optimize the parameters θ in our model using a gradient-based algorithm.

5 Results

In this section, we compare our proposed method with Isomap- and GPLVM- based WiFi-SLAM for robot localization. GPLVM-based WiFi-SLAM was first proposed to estimate the location of a robot based on WiFi signals without labeled calibration data. Unlike calibrated data-based methods, these methods directly project the signal measurements into latent variables, which are considered to be the physical locations of the robot. Here, Isomap (Tenenbaum, de Silva, and Langford 2000) is normally used to initialize the latent variables of GPLVM. Localization methods with labeled calibration data are not considered further since they are known to have significantly higher accuracy.

The dataset used here is presented in (Ferris, Fox, and Lawrence 2007). This dataset contains two traces with a range of around 250 meters that were collected over one floor of a university building. To collect these data, several WiFi access points were installed to send signals, and a robot walked around inside the building and received signals and recorded the corresponding locations. The ground truth path is shown in Fig. 3 (d). Isomap is first illustrated to estimate the robot's locations. It can clearly be seen that Isomap does not capture the overall topological structure of the path and fails to demonstrate clear definitions of intersections and the correct path alignment.

As shown in Fig. 3, GPLVM-based WiFi-SLAM provides a finer path resolution with greater topological similarity to the ground truth trace. However, the ground truth trace includes four right-angled intersections. GPLVM-based WiFi-SLAM does not capture this characteristic and simplifies them as non-right-angled intersections. The root cause of this phenomenon is that the latent variables generated by GPLVM are not sufficiently diverse and tend to concentrate

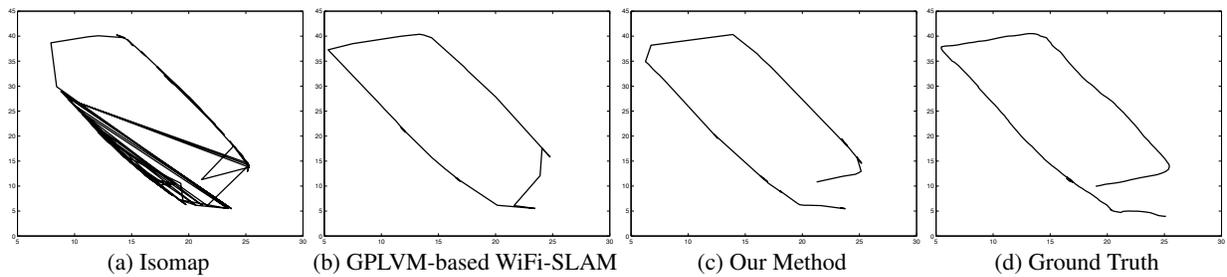| (a) Isomap | (b) GPLVM-based WiFi-SLAM | (c) Our Method | (d) Ground Truth |

Figure 3: The results of three localization methods on one of our traces. The x,y axes are measured in meters. In comparison with other works, our result has more similar topological structure with the ground truth path.

on the most common trace characteristics. In this case, the trace consists of several straight lines and the common characteristics are the straight line paths. As a result, the less frequent feature, i.e., the right-angled intersection, is not considered in the latent variables. The latent variables generated by GPLVM do not capture all aspects of the underlying locations of the signal strength measurements.

In light of this, our method exploits a diversity-encouraging prior to train a more robust model for accurate localization. Here, we choose the dimensionality of latent variable q and parameter λ to be 10 and 0.01, respectively. In comparison with the other techniques, the trace of our method has a normal topological structure. More importantly, our technique can generate other structures present in the ground truth trace. The diversity prior appears to be crucial for capturing more comprehensive features and characteristics of the signal measurements.

However, the mapping between the latent coordinate and the ground truth coordinate is not clearly defined. Therefore, the technique described in (Ferris, Fox, and Lawrence 2007) is exploited to evaluate the localization accuracy. Specifically, given a new signal measurement, its latent variable is used to calculate the distance with training latent variables. Then, the closest point in the latent space to the test latent variable is selected. The ground truth location of the closest point is then used as the test location.

To evaluate localization accuracy, we cross-validate each DGLVM result by performing localization with the remaining test trace. The mean localization error over six localization runs is reported in Fig. 4. Although the localization accuracy is not high enough, this method represents an improvement on WiFi-SLAM techniques that do not use labeled calibration data.

6   Conclusion

Here we propose DGLVM to estimate the locations of WiFi signal strength readings in the case where ground truth locations are unknown during training. By exploiting a diversity-engouraging prior, the location estimation accuracy is significantly improved. Such a prior ensures that the latent variables, namely the physical locations of the robot, are selected without preference or redundancy. Finally, DGLVM can construct a sensor model that can be used for WiFi localization.
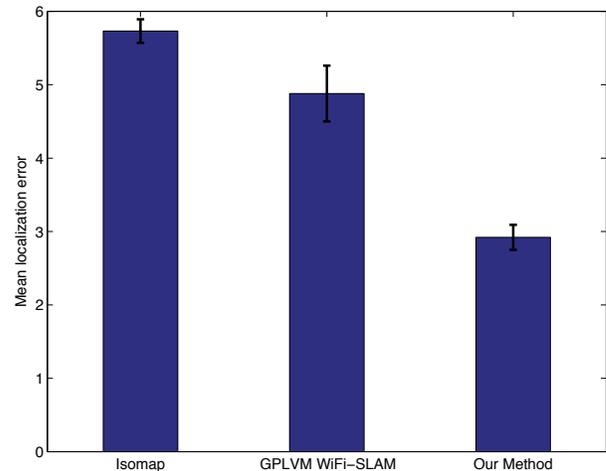

Figure 4: For each method, the error bar illustrates the mean localization error in meters over six localization runs.

Although promising, our model is still not applicable for localization and mapping in complicated and large-scale environments. This is because wireless signals are not constantly stable either indoors or outdoors. To tackle these issues, we will go on to incorporate a variety of sensors (e.g., Kinect, IMU, accelerometers) into our model. Our ultimate aim is to equip any type of mobile device with our WiFi-SLAM technique to undertake various tasks in the future.

7   Acknowledgments

References

Bahl, P., and Padmanabhan, V. 2000. Radar: An in-building rf-based user location and tracking system. In *IEEE Infocom*. MIT Press.

Bay, H.; Ess, A.; Tuytelaars, T.; and Gool, L. V. 2008. Speeded-up robust features (SURF). *Computer Vision and Image Understanding* 110:346–359.

Besl, P. J., and McKay, H. D. 1992. A method for registra-

tion of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2):239–256.

Chiuso, A.; Favaro, P.; Jin, H.; and Soatto, S. 2002. Structure from motion causally integrated over time. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(4):523–535.

Civera, J.; Davison, A. J.; and Montiel, J. M. M. 2008. Inverse depth parametrization for monocular SLAM. *IEEE Transactions on Robotics* 24(5):932–945.

Davison, A. J.; Reid, I. D.; Molton, N. D.; and Stasse, O. 2007. MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(6):1052–1067.

Engel, J.; Schops, T.; and Cremers, D. 2014. LSD-SLAM: Large-scale direct monocular SLAM. In *European Conference on Computer Vision*, 834–849.

Ferris, B.; Fox, D.; and Lawrence, N. 2007. WiFi-SLAM using Gaussian process latent variable model. In *Proc. of the National Conference on Artificial Intelligence (AAAI)*, 2480–2485.

Fitzgibbon, A. W. 2003. Robust registration of 2D and 3D point sets. *Image Vision Computing* 21(13-14):1145–1153.

Grisetti, G.; Stachniss, C.; and Burgard, W. 2005. Improving grid-based SLAM with rao-blackwellized particle filters by adaptive proposals and selective resampling. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*.

Grisetti, G.; Stachniss, C.; and Burgard, W. 2007. Improved techniques for grid mapping with rao-blackwellized particle filters. *IEEE Transactions on Robotics* 23:34–46.

Haeberlen, A.; Flannery, E.; A.M. Ladd, A. Rudys, D. W.; and Kavraki, L. 2004. Practical robust localization over large-scale 802.11 wireless networks. In *Proc. of the Tenth ACM International Conference on Mobile Computing and Networking (MOBICOM)*. MIT Press.

Kim, C.; Kim, H.; and Chung, W. K. 2011. Exactly rao-blackwellized unscented particle filters for SLAM. In *Proc. of the International Conference on Robotics and Automation (ICRA)*, 3589–3594. MIT Press.

Kim, C.; Sakthivel, R.; and Chung, W. K. 2008. Unscented FastSLAM: a robust and efficient solution to the SLAM problem. *IEEE Transactions on Robotics* 24(4):808–820.

Klein, G., and Murray, D. 2007. Parallel tracking and mapping for small AR workspaces. In *IEEE and ACM International Symposium on Mixed and Augmented Reality*, 225–234.

Klein, G., and Murray, D. 2008. Improving the agility of keyframe-based SLAM. In *European Conference on Computer Vision*, 802–815.

Kuemmerle, R.; Grisetti, G.; Strasdat, H.; Konolige, K.; and Burgard, W. 2011. g2o: A general framework for graph optimization. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 3607–3613.

Lawrence, N. D., and Quinonerocandela, J. 2006. Local distance preservation in the GP-LVM through back constraints. In *International Conference on Machine Learning*.

Lawrence, N. D. 2004. Gaussian process latent variable models for visualisation of high dimensional data. In *Neural Information Processing Systems*, 329–336.

Lee, S.; Lee, S.; and Kim, D. 2006. Recursive unscented Kalman filtering based SLAM using a large number of noisy observations. *International Journal of Control, Automation, and Systems* 4(6):736–747.

Letchner, J.; Fox, D.; and LaMarca, A. 2005. Large-Scale localization from wireless signal strength. In *Proc. of the National Conference on Artificial Intelligence (AAAI)*. MIT Press.

Lowe, D. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2):91–110.

M. Bowling, D. Wilkinson, A. G., and Milstein, A. 2005. Subjective localization with action respecting embedding. In *International Symposium of Robotics Research*, 1441–1448. MIT Press.

Mouragnon, E.; Lhuillier, M.; Dhome, M.; Dekeyser, F.; and Sayd, P. 2006. Real time localization and 3D reconstruction. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, 363–370.

Rublee, E.; Rabaud, V.; Konolige, K.; and Bradski, G. 2011. ORB: an efficient alternative to SIFT or SURF. In *IEEE International Conference on Computer Vision*.

Rusinkiewicz, S., and Levoy, M. 2001. Efficient variants of the ICP algorithm. In *International Conference on 3-D Digital Imaging and Modeling*.

Strasdat, H.; Montiel, J. M. M.; and Davison, A. J. 2012. Visual SLAM: Why filter? *Image and Vision Computing* 30(2):65–67.

Tenenbaum, J. B.; de Silva, V.; and Langford, J. C. 2000. A global geometric framework for nonlinear dimensionality reduction. *Science* 290(5500):2319–2323.

Wang, J. M.; Fleet, D. J.; and Hertzmann, A. 2006. Gaussian process dynamical models. In *NIPS*, 1441–1448. MIT Press.