공학석사학위논문

비지도식 베이지안 온라인 학습을 이용한 미지 환경에서의 다중 로봇 탐사 기법

Unsupervised Bayesian Online Learning for Multi-Agent Exploration in an Unknown Environment

2017 년 2 월

서울대학교 대학원

기계항공공학부

임 진 홍

비지도식 베이지안 온라인 학습을 이용한 미지 환경에서의 다중 로봇 탐사 기법

Unsupervised Bayesian Online Learning for Multi-Agent Exploration in an Unknown Environment

지도교수 김 현 진

이 논문을 공학석사 학위논문으로 제출함

2017 년 2 월

서울대학교 대학원

기계항공공학부

임 진 홍

임진홍의 공학석사 학위논문을 인준함

2017 년 2 월

위 원 장	김 유 단
부위원장	김 현 진
위 원	박 찬 국

Unsupervised Bayesian Online Learning for Multi-Agent Exploration in an Unknown Environment

A Thesis

by

JINHONG LIM

Presented to the Faculty of the Graduate School of
Seoul National University
in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE

School of Mechanical & Aerospace Engineering

Seoul National University

Supervisor : Professor H. Jin Kim

February 2017

to my

MOTHER, FATHER, and BROTHER

with love

Abstract

Unsupervised Bayesian Online Learning for Multi-Agent Exploration in an Unknown Environment

Jinhong, Lim
Department of Mechanical & Aerospace Engineering
The Graduate School
Seoul National University

Exploring an unknown environment with multiple robots is an enabling technology for many useful applications. This paper investigates decentralized motion planning for multi-agent exploration in a field with unknown distributions such as received signal strength (RSS) and terrain elevation. We present both supervised with RSS distribution and unsupervised methods with terrain data. The environment is modelled with a Gaussian process using Bayesian online learning by sharing the information obtained from the measurement history of each robot. Then we use the mean function of the Gaussian process to infer the multiple source locations or peaks of the distribution. The inferred locations of sources or peaks are modelled as the probability distribution using Gaussian mixture-probability hypothesis density (GM-PHD) filter. This modelling enables nonparametric approximation of mutual information between peak locations and future robot positions. We combine the variance function of the Gaussian process and the mutual information to design an informative and noise-robust planning algorithm for multiple robots. At the end, the proposed algorithm is extended by applying an unsupervised method with Dirichlet process mixture of Gaussian processes. The experimental performance of supervised method and unsupervised method are analysed by comparing with the variance-based planning algorithm. The experimental results show that the proposed algorithm learns the unknown environmental distribution more accurately and faster.

Keyword : Bayesian nonparametric methods, unsupervised learning, Dirichlet process, online Gaussian process, mutual information, GM-PHD filter, active sensing, decentralized multi-agent

Student Number : 2015-20788

Table of Contents

List of Figures

1

Introduction

This paper proposes a multi-agent exploration algorithm in the following three steps: learning the RSS distribution model, modelling the source locations as probability distribution using PHD, and planning the path by computing mutual information and combining with variance. The corresponding details of the three steps are as follows.

First, using the RSS measurement history of all robots, the RSS distribution model of the entire space is obtained, including the unexplored area, by learning the Gaussian process model. The model is described in terms of means and variances over the entire space [1], and lowering the variances is used as one of the objectives considered in this paper.

Then, we model the signal source location as a probability distribution using PHD by predicting the source location from mean function of the Gaussian process model [2]. Among various kinds of implementations of propagating the PHD, which is called PHD filter [3], the Gaussian mixture-probability hypothesis density (GM-PHD) filter is one of the most suitable forms for real-time implementation, because the GM-PHD filter assumes the prior distribution of the target sample to be Gaussian [4], [5]. Thus, among various extensions [6],[7], we engage the idea of approximating the prior to Gaussian distribution, and applying measurement-based spatial prior on the birth process to enhance the multi-target localization performance [8].

Finally, the local optimal movements of the robots are derived from the policy using combined

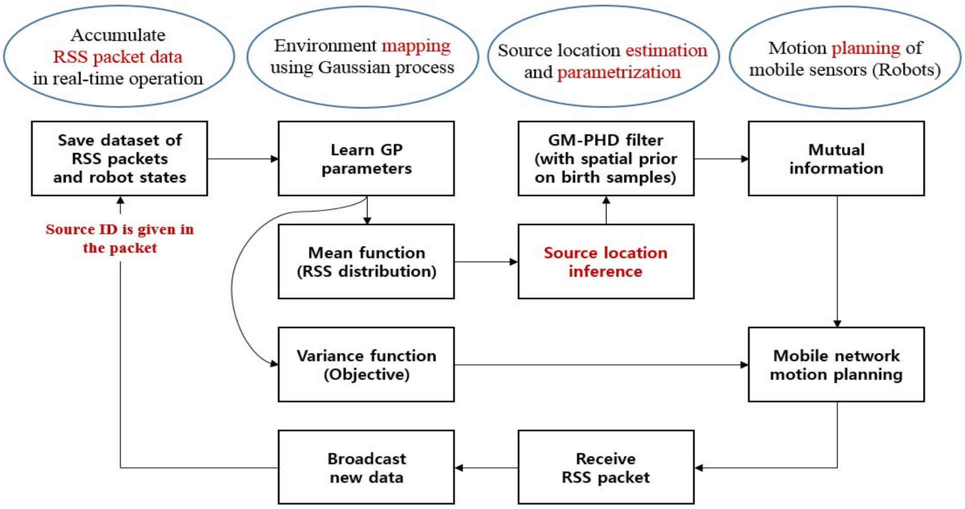


Figure 1.1: Flow of the proposed algorithm (Supervised method).

two objectives. Using the estimated PHD of source location and the future poses of robots, mutual information between them is calculated in nonparametric way [9], [10]. By maximizing the sum of the two normalized objectives, i.e. variance and mutual information, the robots select the most informative actions from a randomly sampled action set. At the end, the proposed algorithm is extended by applying an unsupervised method with Dirichlet process mixture of Gaussian processes.

Using single quantity among the two, i.e. entropy and mutual information, has both pros and cons. Entropy, which is equal to the variance function of Gaussian process model, is an useful objective to find the unexplored region, so that the variance is reduced the most by measuring data at the region. However, the variance-based policy shows poor performance when the amount of data is not enough, especially for noisy signals. Besides, the importance of the data at each position is not reflected in the robots' motion. On the other hand, mutual information between the RSS source location and future movements of the robots focuses on detecting the sources, where the measurements around the sources are the most important data in regression settings. Also, the mutual information tends to lead mobile robots to move toward the sources with even less amount of data. The mutual information, however, is not desirable under poor conditions for computing the mutual information, such as absence of samples or large computational cost. We

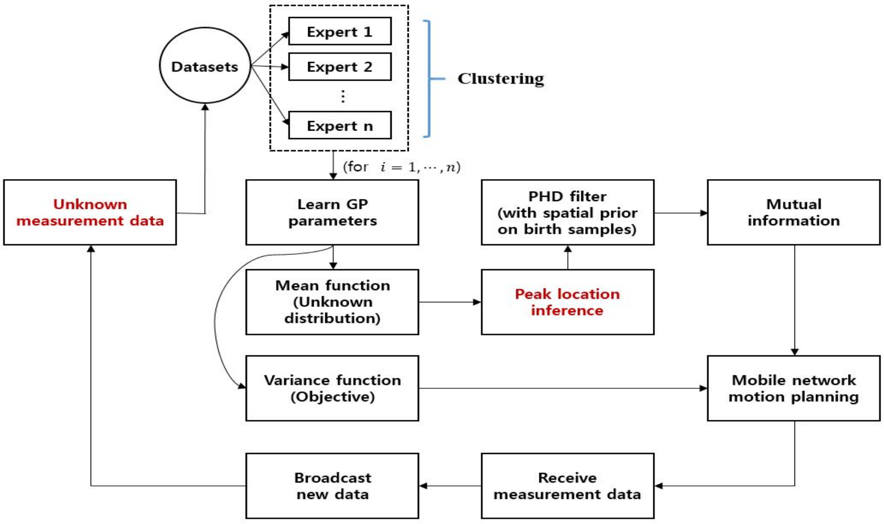


Figure 1.2: Flow of the proposed algorithm (Unsupervised method).

try to compensate the problems of each objective by the advantages of the other by fusing them.

1.1 Literature review

Planning informative actions is an essential problem for active sensing tasks such as autonomous surveillance, mapping and target localization. Especially, engaging multiple robots to cooperatively perform these active sensing tasks has been well studied. Most of the proposed methods are based on the information theory such as entropy and mutual information, because these quantities well represent the informativeness using the measurement history and future poses of sensors.

Many approaches are investigated to implement the information-theory-based methods. Among them, some recent studies employ nonparametric approaches to derive entropy [11] and mutual information [12]. In Gaussian process-based terrain learning [11] algorithm, the motion planning of multiple robots is performed using the variance which is derived from online learning of Gaussian process, where the variance is considered as entropy. Besides, in mutual information-based target localization [12] algorithm, mutual information is calculated using the detection model of laser scanner and approximating the integral using probability hypothesis density (PHD).

The information-theory-based method is highly applicable to various kinds of sensors. [13] proposed the visual sensor-based 3-D mapping using a single quadrotor. Besides, information-based localization using single or multiple robots is performed using downward facing camera [14], [15], laser scanner [12], and many kinds of sensors such as bearing-only or range-only sensors [9], [16]. Furthermore, recognition of environmental signal distribution based on measurement history has also been performed [11], [17], [18].

Among the various kinds of sensors, we consider a received signal strength (RSS) sensor for learning the unknown RSS distribution in this paper. The distribution of RSS field is essential for RSS-based indoor localization [19], [2] in case of mobile-phone-based commercial advertisement and RSS-based search-and-rescue in GPS-restricted space. For such applications, the accurate RSS distribution map is important to utilize as fingerprint or propagation-model-based likelihood. Thus, we present a decentralized RSS distribution learning framework for a team of autonomous mobile robots.

1.2 Thesis contribution

The main contribution of this paper can be summarized as follows. (1) We fuse variance and mutual information to design noise-robust and efficient policy for multi-agent exploration. Unlike most previous research that use a single quantity among entropy and mutual information, this paper combines them for the settings where Gaussian process is used to model the environmental distribution. (2) Moreover, we engage GM-PHD filter which uses the diffuse spatial prior of the birth process, so the samples of the filter are generated over the entire space. Although [12] also engaged GM-PHD filter to compute mutual information, the used traditional GM-PHD filter is not suitable when inference of the source locations is uncertain or fluctuates. (3) At the end, we extend the proposed algorithm to unsupervised method by applying Dirichlet process mixture of Gaussian processes. An actual hardware experiment is performed to find out how the fused two objectives compensate the problems of each other.

1.3 Thesis outline

The rest of the paper is composed as follows. Chapter 2 shows the estimation process of the environmental model and Chapter 3 presents localization and parametrization of the source locations. Then, we describe the decentralized multi-agent control laws in Chapter 4, and extend the proposed algorithm to unsupervised method in Chapter 5. Finally, the experimental results of supervised method and unsupervised method are presented in Chapter 6, which is followed by the conclusions.

2

Gaussian process model

In this Chapter, we present the estimation process of the environmental RSS distribution. To obtain the unknown distribution without obtaining an extremely large amount of data over the entire space, we used Gaussian process to predict the distribution with relatively small amount of data even over unexplored region. Section 2.1 shows the online learning of RSS distribution using Gaussian process, and Section 2.2 presents the hyperparameter optimization by maximizing the log-likelihood.

2.1 Gaussian process

Using the data sampled by robots, we build Gaussian process model [1] that represents the environmental distribution. In our case, let $\mathbf{x}_r^k \in X^k$ be the 2-D position vector of the r-th robot at iteration k, $\mathbf{p}_r^k \in P^k$ be the packet which contains RSS data of m access points in the order of their ID received by r-th robot at iteration k, and $y_r^k \in \mathbf{y^k}$ be the total sum of RSS data measured by r-th robot at iteration k, where $r = 1, \cdots, R$. Then, by continuously sharing and saving data, the training dataset at iteration k can be made as $D = \{(X^1, \mathbf{y}^1), (X^2, \mathbf{y}^2), \cdots, (X^k, \mathbf{y}^k)\}$, which contains entire RSS data history measured by all robots.

From the dataset, Gaussian process is used to obtain the predictive distribution over the arbitrary positions $\mathbf{x}_{i*} \in X_*$, where $i = 1, \cdots, n$. We set the arbitrary positions to be the fixed

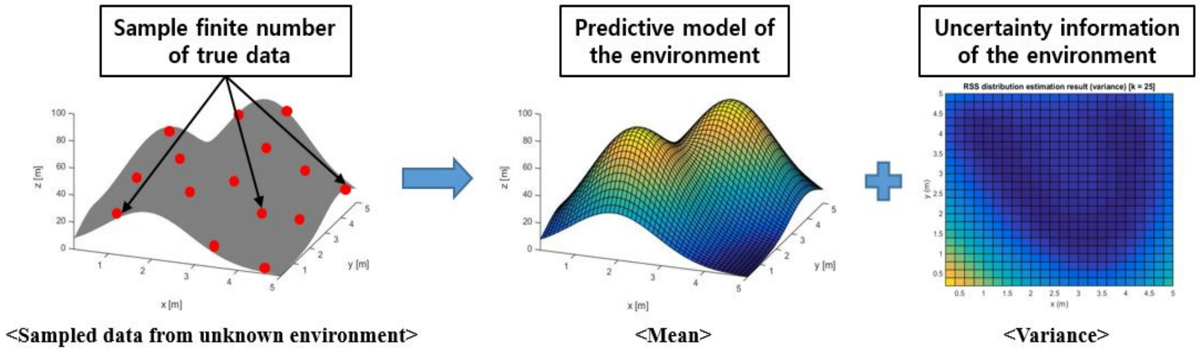


Figure 2.1: Concept of environment learning using Gaussian process regression.

entire workspace of robots, where the number of points is n, so that the source location can be estimated by maximum likelihood principle. Although the large size of the arbitrary position set causes more computational cost for calculating kernel matrix, the matrix can be computed in advance because it is fixed for the whole procedure. In this work, Gaussian kernel function in equation(2.1) is used, with Kronecker's delta $\delta_{\mathbf{x}\mathbf{x}'} = 1\ (if\ \mathbf{x} = \mathbf{x}')$.

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{|\mathbf{x} - \mathbf{x}'|^2}{2l^2}\right) + \sigma_n^2 \delta_{\mathbf{x}\mathbf{x}'}. \tag{2.1}$$

Assuming zero mean and variance of σ_n^2 for noise of target function, let the target function be $s = f(\mathbf{x}) + \epsilon$, and a prediction $s_* = f(\mathbf{x}_*)$. Then the joint distribution is,

$$\begin{bmatrix} s \\ s_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(\bar{X}, \bar{X}) & K(\bar{X}, X_*) \\ K(X_*, \bar{X}) & K(X_*, X_*) \end{bmatrix}\right), \tag{2.2}$$

where $\bar{X} = \left[X^1, X^2, \cdots, X^k\right]$, $[K(X, X')]_{uv} = k(x_u, x_v^{'})$ and $x_u \in X$, $x_v \in X^{'}$, i.e. $K(\bar{X}, \bar{X})$ is the $kR \times kR$ covariance matrix computed for all pairs between training positions, $K(\bar{X}, X_*)$ is the $kR \times n$ covariance matrix computed for all pairs of training positions and arbitrary positions, and $K(X_*, X_*)$ is the $n \times n$ covariance matrix computed for all pairs between arbitrary positions. The predictive distribution of each access point is then given as equations (2.3) and (2.4) with $m \times kR$ matrix $\bar{P} = [P^1, P^2, \cdots, P^k]$ which contains all the shared RSS packet history.

$$\mu_* = K(\bar{X}, X_*)^T K(\bar{X}, \bar{X})^{-1} \bar{P}^T \tag{2.3}$$

$$\Sigma_* = K(X_*, X_*) - K(\bar{X}, X_*)^T K(\bar{X}, \bar{X})^{-1} K(\bar{X}, X_*) \tag{2.4}$$

2.2 Hyperparameter optimization

Let $\theta = [\sigma_f^2, l^2, \sigma_n^2]^T$ be the hyperparameter, which is the only tunable set in the Gaussian process. The optimal hyperparameter can be obtained by maximizing the log-marginal likelihood,

$$\begin{aligned} &\log p(\mathbf{y}|X, \theta) = \\ &\quad -\tfrac{1}{2}\mathbf{y}^T K(\bar{X}, \bar{X})^{-1}\mathbf{y} - \tfrac{1}{2}\log K(\bar{X}, \bar{X}) - \tfrac{n}{2}\log 2\pi \end{aligned} . \tag{2.5}$$

The log-marginal likelihood can be maximized using gradient method, and by optimizing the hyperparameter, the best model that fits observed data the most is obtained as equation (2.3). The mean function (2.3) is also used as the likelihood distribution of the RSS source location which is discussed in Section 2.2. The variance function (2.4) is used as a part of the control law of robots which is discussed in Section 3.2.

3

Parametrization of signal source location

In this chapter, we model the RSS source location as the probability distribution using GM-PHD filter. To compute the mutual information, two random variables are needed. For our settings, the event that the RSS source locations exist at certain positions is one of the two events. Thus, in Section 3.1, the deterministic source locations which are inferred from the mean function of the Gaussian process. Then, the parametrization of the deterministic location to a random variable with GM-PHD filter is described in Section 3.2.

3.1 Conventional GM-PHD filter

The process of GM-PHD filter consists of prediction and measurement update like other typical filters such as Kalman filter or particle filter. At the measurement update step, we obtained measurements from the mean RSS distribution function (2.3) of Gaussian process model. Let $\mathbf{z}_j^k \in Z^k$ be the inferred location of j-th RSS source at iteration k. Since the model of RSS versus distance can be described as the inverse proportion curve in log-scale, the location of sources can be inferred by finding the location that has the maximum predicted mean RSS among n arbitrary positions.

$$\mathbf{z}_j^k = \arg\max_{\mathbf{x}_*} \left[\mu_*(\mathbf{x}_*)\right]_j \quad for\ \mathbf{x}_* \in X_* \tag{3.1}$$

At the prediction step, the traditional GM-PHD filter propagates several predicted intensities,

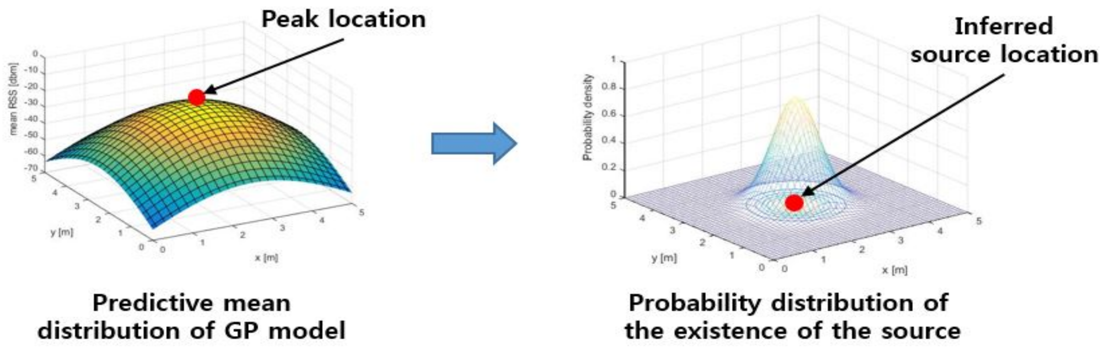


Figure 3.1: Parametrization of source location estimation result.

which are called birth, spawn, and survive samples respectively. Let $\hat{\mathbf{x}}_j^k \in \hat{X}^k$ be the estimated position vector of j-th source at iteration k, and the system update and measurement update functions be,

$$g(\hat{\mathbf{x}}_j^k) = N\left(\hat{\mathbf{x}}_j^k; G\hat{\mathbf{x}}_j^{k-1}, Q^k\right) \tag{3.2}$$

$$h(\mathbf{z}_j^k) = N\left(\mathbf{z}_j^k; H\hat{\mathbf{x}}_j^k, R^k\right), \tag{3.3}$$

where G and H are 2×2 identity transition and observation matrices respectively, and Q^k and R^k are the 2×2 covariance matrices. Then the intensities of each sample are computed as,

$$v_{\gamma,k|k-1} = \sum_{(j)} w_\gamma^{(j)} N(m_\gamma^{(j)}, \mathbf{P}_\gamma) \tag{3.4}$$

$$v_{\beta,k|k-1} = \sum_{(j)} \sum_{(l)} \{w_{k-1}^{(j)} w_\beta^{(j)} N(m_{\beta,k|k-1}^{(j)(l)}, \mathbf{P}_\beta)\} \tag{3.5}$$

$$v_{s,k|k-1} = p_s \sum_{(j)} w_{k-1}^{(j)} N(m_{s,k|k-1}^{(j)}, \mathbf{P}_s), \tag{3.6}$$

where γ, β, s stand for birth, spawn, survive samples respectively, $m_{s,k|k-1}^{(j)} = G\hat{\mathbf{x}}_{(j)}^{k-1}$, $m_{\beta,k|k-1}^{(j)}$ is the position vector obtained by adding Gaussian random noise to $m_{s,k|k-1}^{(j)}$, $\mathbf{P}$ is the covariance matrix of the distribution of each sample, m_γ is the matrix which contains position vectors extracted from the preset spatial prior, and w is the weight of corresponding samples. Then, the final estimated target posterior intensity is updated as,

$$v_{k|k-1} = v_{\gamma,k|k-1} + v_{\beta,k|k-1} + v_{s,k|k-1} \tag{3.7}$$

$$v_k = (1 - p_{D,k})v_{k|k-1} + \sum v_{D,k}, \tag{3.8}$$

where $v_{D,k}$ is the term reflecting that the measurements are correctly detected. Detailed explanation on propagation of PHD can be found in Section III-B of [4] including mixture and pruning steps.

3.2 Spatial prior on the birth process

The traditional method that presets the spatial prior on birth samples is not suitable for our case, because the locations of the RSS sources are unknown in advance, and the estimates of the locations are highly fluctuating. Thus, we apply observation-based spatial prior to the birth process. Then the intensities (3.7) and (3.8) are changed to,

$$v_{k|k-1} = v_{\beta,k|k-1} + v_{s,k|k-1} \tag{3.9}$$

$$v_k = (1 - p_{D,k})v_{k|k-1} + \sum v_{D,k} + \sum_{(j)} w_\gamma^{(j)} N(\mathbf{z}_{(j)}^k, \mathbf{P}_\gamma), \tag{3.10}$$

by introducing a normalizing factor w_0. Details about usage of the normalizing factor can be found in Section 4 of [8]. Using the posterior intensities (3.10), which are composed of weights and mean positions of the sources, estimated locations can be obtained. Also, the final estimates are used for nonparametric approximation of mutual information, which is discussed in Section 3.1.

4

Information-based multi-agent control

In this chapter, we present the policy for multiple robots to find the optimal movements by using the variance which is discussed in chapter 2 and the mutual information which is covered in section 4.1. In section 4.1, the sample-based method of computing the mutual information is described. For real-time application, we used sample-based computation of the mutual information by generating samples with GM-PHD filter. Finally, the proposed policy is shown in section 4.2 in detail.

4.1 Nonparametric computation of mutual information

One of the objectives considered for cooperative multi-agent exploration in this paper is mutual information between estimated RSS source locations and future measurement of robots. The mutual information is defined as,

$$\begin{aligned} I\left[\hat{\mathbf{X}}^k;\mathbf{B}\right] &= \int\int p(\hat{X}^k, B)\log\frac{p(\hat{X}^k, B)}{p(\hat{X}^k)p(B)}\delta\hat{X}^k\delta B \\ &= H\left[\mathbf{B}\right] - H\left[\mathbf{B}|\hat{\mathbf{X}}^k\right], \end{aligned} \tag{4.1}$$

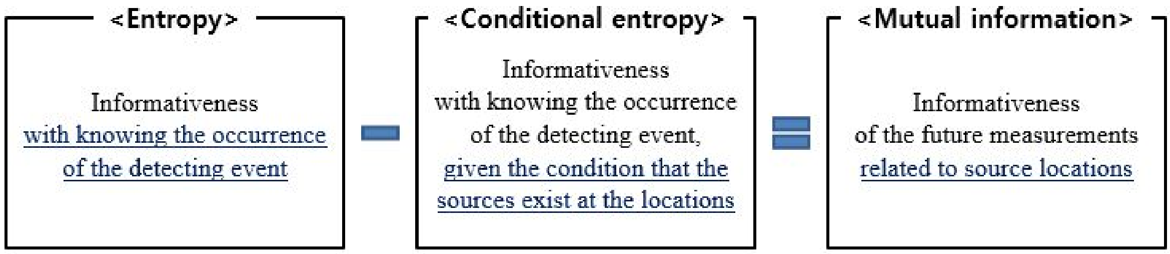


Figure 4.1: Conceptual definition of the mutual information in this paper.

where, $H[\mathbf{B}]$ is the entropy, $H[\mathbf{B}|\hat{\mathbf{X}}^k]$ is the conditional entropy of the future measurements, and $b \in B$ is the binary event,

$$b = \begin{cases} 0 & (Z^k = \phi) \\ 1 & (else) \end{cases}, \tag{4.2}$$

where the binary event denotes whether each robot detects at least one RSS source or not. As shown as equation (4.2), $b = 0$ denotes that a robot does not detect any of the RSS source, and $b = 1$ represents the other case. By introducing such event, the future positions of each robot can be modelled as random samples with probability distribution, so the mutual information becomes the measure of the difference between two probability distributions, future robot positions and estimated RSS source locations (3.10). In this paper, we set probability of detection to be,

$$p_d(\hat{\mathbf{x}}^k) = \min\left(1, \frac{r_t + 1}{r + 1}\right) \times \mathbf{1}\left(r < r_{\max}\right), \tag{4.3}$$

where r is the distance between a robot and an estimated RSS source location, r_t is the preset threshold range, and r_{max} is the preset maximum range. Such setting has physical meaning for RSS propagation model, since the RSS tends to steeply increase near the source. Furthermore, the setting is also expected to be applied to other sensors [12], [11] such as laser scanner, and magnetic sensors, because it is important to sample the information near the peaks in estimating an unknown model using regression. Then, the entropy and conditional entropy can be derived as,

$$H\left[\mathbf{B}\right] = -\left\langle p(b^{1:R}), \ln p(b^{1:R})\right\rangle \tag{4.4}$$

$$H\left[\mathbf{B}|\hat{\mathbf{X}}^k\right] = -\int p(\hat{X}^k) \sum_{b\in\{0,1\}} p(b^{1:R}|\hat{X}^k) \ln p(b^{1:R}|\hat{X}^k)\delta\hat{X}^k, \tag{4.5}$$

where,

$$p(b=0|\hat{X}^k) = \prod_{\hat{X}^k}\left(1 - p_d(\hat{\mathbf{x}}_j^k)\right)$$

$$p(b^1, b^2, \cdots, b^R) = \int \prod_{r\in C_0} p(b^r = 0|\hat{X}^k) \times \prod_{r\in C_1}\left(1 - p(b^r = 0|\hat{X}^k)\right) \times p(\hat{X}^k)\delta\hat{X}^k.$$

The computations for integral and the operator $\langle \bullet, \bullet \rangle$ are done by the means of nonparametric approximations, i.e. total sum of the weighted sample sets, superscript of b is the indicator of each robot, C_0 represents the estimated samples that are not detected, C_1 denotes the estimated samples that are detected by at least one robot, and $b^{1:R}$ represents all possible combinations of binary event b.

4.2 Concatenated objective-based control policy

In this section, we present the concatenated control policy for multiple robots considering both variance and mutual information, which are discussed in Section 2.1 and Section 3.1 respectively. These objectives are widely used in many related studies such as [11], [12], [20], but most of the studies utilized only one objective of the two. Although [20] estimated the RSS source location, only one RSS source is considered in the algorithm and a single optimal policy is randomly selected between variance-based one and source location estimation-based one. In real-world applications, the number of RSS source is usually unknown and the received packet length is not fixed because of the varying accessibility to RSS sources at the specific position. We consider such issues using the random sample-based nonparametric methods.

Let the future position of a robot be $\mathbf{q}_r \in Q_r$ and d be the preset radius. Then the future positions are sampled from the set

$$Q_r = \left\{ \mathbf{q}_r \in Q_r \;\middle|\; \begin{array}{c} \left\|\mathbf{x}_r^k - \mathbf{q}_r\right\|_2 = d, \\ \mathbf{q}_r \in \{collision\ free\} \end{array} \right\}. \tag{4.6}$$

We sample six candidates of local goal positions from the set (4.6). To define the control policy to select the optimal one from the candidates, let the normalized form of variance and mutual information be $\sigma(\mathbf{q}) \in \bar{\Sigma}$ and $I(\mathbf{q}) \in \bar{\mathbf{I}}$ respectively. Then the optimal future position is,

$$\mathbf{q}_r^* = \arg\max_{\mathbf{q}_r} \ \{\sigma(\mathbf{q}_r) + I(\mathbf{q}_r)\}. \tag{4.7}$$

The variance term leads to the largest reduction of the variance over the entire field, and the mutual information term contributes to selecting the most informative position that increases the probability of detecting RSS sources. The algorithm is set to be terminated if the mean normalized variance is converged. The criterion, however, can be interrupted by occasional bad results of hyperparameter optimization. In such cases, the maximum number of iterations is set according to the approximate size of the exploring space.

Algorithm 1 Multi-Agent Exploration

$\bar{X} \leftarrow NULL$, $\bar{Y} \leftarrow NULL$, $\bar{P} \leftarrow NULL$, $\hat{X} \leftarrow NULL$

compute $K(X_*, X_*)$

for r = 1,2,$\cdots$,R (*parallel implementation*) **do**

  while StopCriteria **do**

    $X_{team}, \mathbf{y}_{team}, P_{team} \leftarrow GetInformation$

    $\mathbf{p}_r^k \leftarrow ReceivePacket$

    $y_r^k \leftarrow sum(\mathbf{p}_r^k)$

    $\bar{X} \leftarrow [\bar{X}, \mathbf{x}_r^k, \bar{X}_{team}]$

    $\bar{Y} \leftarrow [\bar{Y}, y_r^k, \mathbf{y_{team}}]$

    $\bar{P} \leftarrow [\bar{P}, \mathbf{p}_r^k, P_{team}]$

    $\theta^* \leftarrow OptimizeHyperparameter(\bar{X}, \bar{Y})$

    $\mu_*, \Sigma_* \leftarrow PredictiveDistributioin(X_*, \bar{P}, \theta^*)$

    $\mathbf{z}_j^k \leftarrow \arg\max_{\mathbf{x}_*} [\mu_*(\mathbf{x}_*)]_j \quad (forj = 1, \cdots, m)$

    $\hat{X}^k = GM - PHDfilter(\hat{X}^{k-1}, Z^k)$

    $Q_r \leftarrow SampleNextPosition(\mathbf{x_r^k})$

    $\mathbf{q}_r^* = \arg\max_{\mathbf{q}_r} \{\sigma(\mathbf{q}_r) + I(\mathbf{q}_r)\}$

    $BroadcastInformation(\mathbf{x}_r^k, \mathbf{q}_r^*, y_r^k, P_r^k)$

    $\mathbf{x}_r^{k+1} \leftarrow \mathbf{q}_r^*$

5
Unsupervised implementation

In this chapter, we extend the proposed algorithm by applying the unsupervised method by with Dirichlet process mixture of Gaussian processes. In Chapter 2, the RSS packet which involves IDs of each signal source is considered as the training data. However, the IDs become not available, if the other kinds of data are considered such as acoustic distribution, magnetic field, terrain information, and pollution level. In such cases, the IDs can be inferred using Dirichlet process mixture of Gaussian processes by the means of maximum a posteriori. It is assumed that the unknown environment is the mixture of several models, which can be modelled using several Gaussian processes.

5.1 Dirichlet process mixture of Gaussian processes

Dirichlet process mixture is one of the most popular stochastic process models, which provide prior probabilities for clustering problems [21], [18]. Furthermore, Dirichlet process is typically used with Gaussian process in pair, because both of them are nonparametic and scalable in continuous space. In this paper, Dirichlet process mixture of Gaussian processes is used in two reasons: (1) Manual tuning of any parameter is not required, if the concentration parameter α, which is the only tunable parameter, is optimized by a data-based sampling technique. (2) The number of Gaussian process models is automatically increased or decreased as the complexity of

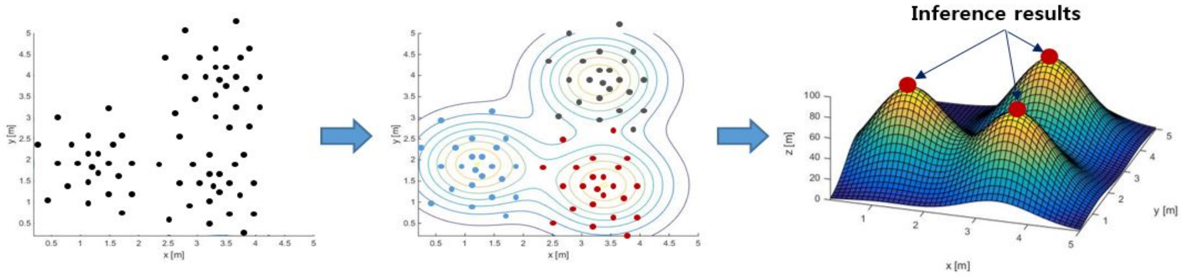


Figure 5.1: Unsupervised method for peak location inference.

the data grows. We call each Gaussian process model among the mixture, an expert.

In this section, let the dataset be $D = \{(X^1, \mathbf{y}^1), (X^2, \mathbf{y}^2), \cdots, (X^k, \mathbf{y}^k)\}$, which is also discussed in Section 2.1. Among the dataset, let the input data which is obtained by r-th robot at k'-th iteration be the $\mathbf{x}_{rk'}$ when the current iteration number is k. Also, let $c_{\mathbf{x}_{rk'}}$ be the label of the data $\mathbf{x}_{rk'}$, which indicates that the data $\mathbf{x}_{rk'}$ is involved in the $c_{\mathbf{x}_{rk'}}$-th expert for the whole dataset D. We use expression j for the indicators, because in unsupervised settings, the indicators play the same role as the source IDs which are discussed in Section 2.1. Then, the prior probabilities for the cases that a data $\mathbf{x}_{rk'}$ is involved in one of the existing experts, and that the data $\mathbf{x}_{rk'}$ is involved in a new expert, is given as,

$$p\left(c_{\mathbf{x}_{rk'}} = j \,|\, c_{D\backslash \mathbf{x}_{rk'}}, \theta_j\right) = \begin{cases} \frac{n_{\mathbf{x}_{rk'}j}}{|D|-1+\alpha} & (if\ j \leq J) \\ \frac{\alpha}{|D|-1+\alpha} & (if\ j = J+1) \end{cases} \tag{5.1}$$

where,

$$n_{\mathbf{x}_{rk'}j} = (|D|-1) \frac{\sum\limits_{\mathbf{x}_{near} \in D_N} \left\{k\left(\mathbf{x}_{rk'}, \mathbf{x}_{near}\right) \delta_{j\, c_{\mathbf{x}_{near}}}\right\}}{\sum\limits_{\mathbf{x}_{near} \in D_N} k\left(\mathbf{x}_{rk'}, \mathbf{x}_{near}\right)} \tag{5.2}$$

$$D_N = \{\mathbf{x}_{near} \in D\backslash \mathbf{x}_{rk'} \,|\, \|\mathbf{x}_{rk'} - \mathbf{x}_{near}\|_2 \leq \eta\} \tag{5.3}$$

i.e. D_N is the set of data which are located less than η distant from $\mathbf{x}_{rk'}$ for some distance $\eta > 0$. Besides, $k(\bullet, \bullet)$ is the predefined kernel function in Section 2.1, α is the concentration parameter, and J is the current maximum number of experts. The equation 5.1 means that for $j = 1, 2, \cdots, J$, a data $\mathbf{x}_{rk'}$ is involved in the j-th expert proportional to the number of neighboring data with the same indicator as $c_{\mathbf{x}_{rk'}}$ and within the range of η. Also, for $j = J+1$, the data $\mathbf{x}_{rk'}$ is involved in the j-th Gaussian process model proportional to the concentration parameter α. Thus, the concentration parameter α is the major factor that balances the total number of experts.

In the proposed algorithm, the indicators of the initial data measured by each robot are set to be identical. Also, we assume that each robot is labelled with the same indicator which the last data was assigned to, and that the newly obtained data is initially labelled with the robot's indicator. Then, the indicators are iteratively re-inferred by Gibbs sampling. In Gibbs sampling process, a data is assigned to an indicator by the means of maximum a posteriori. The posteriors are obtained by updating the Dirichlet process priors of equation 5.1 using Bayes' rule as,

$$p\left(c_{\mathbf{x}_{rk'}} = j | c_{D\backslash\mathbf{x}_{rk'}}, \theta_j, D\right) \propto \begin{cases} p\left(y_{rk'} \left| c_{rk'} = j, D\backslash\mathbf{x}_{rk'}, \theta_j \right.\right) p\left(c_{\mathbf{x}_{rk'}} = j | c_{D\backslash\mathbf{x}_{rk'}}, \theta_j\right) & (if\ j \leq J) \\ p\left(y_{rk'} \left| c_{rk'} = j, \theta_j \right.\right) p\left(c_{\mathbf{x}_{rk'}} = j | c_{D\backslash\mathbf{x}_{rk'}}, \theta_j\right) & (if\ j = J+1) \end{cases} \tag{5.4}$$

where $p\left(y_{rk'} \left| c_{rk'} = j, D\backslash\mathbf{x}_{rk'}, \theta_j \right.\right)$ is the probability of $y_{rk'}$ given the Gaussian process model which is learned with the set $D_j\backslash\mathbf{x}_{rk'}$, and $p\left(y_{rk'} \left| c_{rk'} = j, \theta_j \right.\right)$ is the probability of $y_{rk'}$ where a new model is given by the parameters which are randomly sampled from preset priors. The dataset is divided into several experts as a result of repeating the Gibbs sampling, by iteratively selecting maximum posterior given other conditions fixed such as indicators or hyparparameters of other data. Then, the same process starting from the Section 2.1 is performed until the termination criteria is satisfied.

5.2 Parameter optimization with adaptive rejection sampling

The only tunable parameter of Dirichlet process, the concentration parameter α, can be optimized by drawing from the posterior distribution. With a vague prior of inverse Gamma shape, the prior probability of concentration parameter α is given as,

$$p(\alpha) \propto \alpha^{-\frac{3}{2}} \exp\left(-\frac{1}{2\alpha}\right) \tag{5.5}$$

Also, the likelihood for the concentration parameter α can be derived as,

$$p\left(|D_1|, |D_2|, \cdots, |D_m| \;\; |\alpha\right) = \frac{\alpha^m \Gamma(\alpha)}{\Gamma\left(|D| + \alpha\right)} \tag{5.6}$$

where, $|D_j|$ is the number of data which are assigned to j-th expert. Thus, the posterior probability on the concentration parameter α is given as,

$$p\left(\alpha \,|m, |D|\right) \propto \frac{\alpha^{m-\frac{3}{2}} \exp\left(-\frac{1}{2\alpha}\right) \Gamma(\alpha)}{\Gamma\left(|D| + \alpha\right)} \tag{5.7}$$

As shown in equation 5.7, the posterior is dependent only on the number of data and experts. Noticing that the distribution $p\left(\log(\alpha)\; |m, |D|\right)$ is log-concave, the concentration parameter can be sample from the distribution using adaptive rejection sampling.

Algorithm 2 Unsupervised Multi-Agent Exploration

$\bar{X} \leftarrow NULL$, $\bar{Y} \leftarrow NULL$, $\hat{X} \leftarrow NULL$

compute $K(X_*, X_*)$

for r = 1,2,$\cdots$,R (*parallel implementation*) **do**

  while StopCriteria **do**

    $X_{team}, \mathbf{y}_{team}, \mathbf{c}_{team} \leftarrow GetInformation$

    $y_r^k \leftarrow ReceiveMeasurement$

    $\bar{X} \leftarrow [\bar{X}, \bar{X}_{team}, \mathbf{x}_r^k]$

    $\bar{Y} \leftarrow [\bar{Y}, \mathbf{y_{team}}, \mathbf{y_r^k}]$

    $\mathbf{c}^k \leftarrow [\mathbf{c}^{k-1}; \mathbf{c}_{team}; c_r^{k-1}]$

    $\mathbf{c}^k \leftarrow GibbsSampling(\bar{X}, \bar{Y}, \mathbf{c}^k)$

    $c_r^k \leftarrow \mathbf{c}^k(end)$

    $\theta^* \leftarrow OptimizeHyperparameter(\bar{X}, \bar{Y})$

    $\mu_*, \Sigma_* \leftarrow PredictiveDistributioin(X_*, \bar{Y}, \mathbf{c}^k, \theta^*)$

    $\mathbf{z}_j^k \leftarrow \arg\max_{\mathbf{x}_*} \left[\mu_*(\mathbf{x}_*)\right]_j \quad (forj = 1, \cdots, m)$

    $\hat{X}^k = GM - PHDfilter(\hat{X}^{k-1}, Z^k)$

    $Q_r \leftarrow SampleNextPosition(\mathbf{x_r^k})$

    $\mathbf{q}_r^* = \arg\max_{\mathbf{q}_r} \{\sigma(\mathbf{q}_r) + I(\mathbf{q}_r)\}$

    $BroadcastInformation(\mathbf{x}_r^k, \mathbf{q}_r^*, y_r^k, c_r^k)$

    $\mathbf{x}_r^{k+1} \leftarrow \mathbf{q}_r^*$

6

Simulation and experiment

In this chapter, we validate the performance of the proposed algorithm in several ways. First, we analyze the experimental results of supervised method with RSS distribution data in section 6.1. Then, the simulation using the unsupervised method is performed with terrain data, and the result is discussed in section 6.2. By applying the method to the terrain data in the simulation, we show the flexibility of the proposed algorithm. Finally the experimental result of unsupervised method is presented in section 6.3 with RSS distribution data.

6.1 Experimental settings and results for supervised method

6.1.1 Experimental settings

We apply the proposed algorithm to the experiment for learning the environmental RSS distribution and finds the locations of RSS sources. The experimental setup is composed of two Pioneer P3-DX's which are equipped with a laptop; UBee430 motes including six source nodes and two base nodes; and a Vicon system which provides the poses of each robot. The robots share the information through TCP/IP communication using Wi-Fi, and receive Vicon data from the host PC. Each robot is controlled using ROS program with the main algorithm running on Matlab program. The UBee430 motes are programmed using TinyOS software.

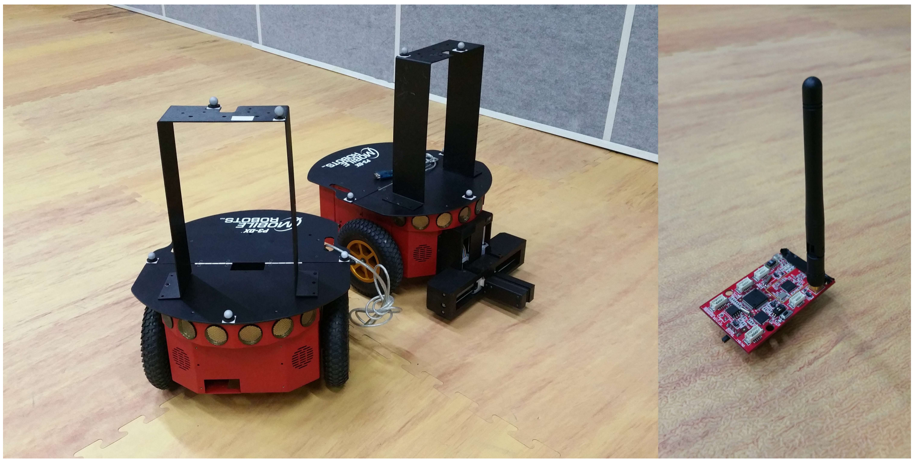


Figure 6.1: Experimental components: Pioneer P3-DX (left) and UBee430 mote (right).

We set the validation area to be a 5m $\times$ 5m square region, and step size of the robot movement to be 20cm. Also, we obtain the ground truth RSS distributions with the resolution of 20cm for each node, which are extremely noisy as shown on Figure 6.2. The ground truth data is obtained by averaging 10 measurements at each positions. We provide the known ground truth for the robots as the measurements in the experiment. The covariance matrices of GM-PHD filter including the system model and observation model are set to be $[0.2^2, 0;\ 0, 0.2^2]$, and the probabilities are set to be $p_s = 0.8$, $p_D = 0.9$, $w_\gamma = 0.6$, $w_\beta = 0.4$, and $w_\gamma^0 = 0.7$ which is the normalizing factor for observation-based birth prior. Also, the thresholds used in computing mutual information are set to be $r_t = 0.2$ and $r_{max} = 6$.

6.1.2 RSS distribution learning experiment result

By comparing with the entropy-based method, the performance of the proposed algorithm is validated in two aspects: the accuracy of RSS source localization and RSS distribution learning; and convergence speed of hyperparameters and RMS error. In Figure 6.4, the hyperparameters converge faster when both objectives are used. The algorithm based on concatenated policy converges after visiting about 10 sampling points, while the entropy-based algorithm converges after visiting about 17 waypoints. When only the variance is used, sampling measurements in the most informative position at the beginning becomes difficult due to the absence of prior knowledge about the environment, which leads to poor performance in finding the optimal position. Otherwise, the concatenated policy finds the informative position by maximizing mutual information. This result suggests that the mutual information term leads to growth in robustness against the noisy measurement.

In Figures. 6.5 and 6.6, we perform 10 validations using each policy, and show mean RMS error histories of the RSS source localization and RSS distribution learning. In average, the RMS error reduces faster and converges to a lower value in the concatenated policy case. Although the converged values seem to be similar in Figure 6.6, the source localization performance indicates that the distribution estimation result of the concatenated policy is more desirable. The larger source localization error means that the distribution estimation for the peaks is not desirable, where the peak area is the most informative and dominant part for RSS model. The major portion of the converged RMS error in the concatenated policy comes from the extremely noisy ground truth. During the experiment, some undesirable situations are observed, for example, all the estimated samples of GM-PHD filter are eliminated at the pruning step because of very unsteady prediction results. The robots, however, still find the suboptimal action by selecting the position with maximum entropy. Overall, the proposed algorithm well performs the task, as the two objectives compensate the issues of each other.

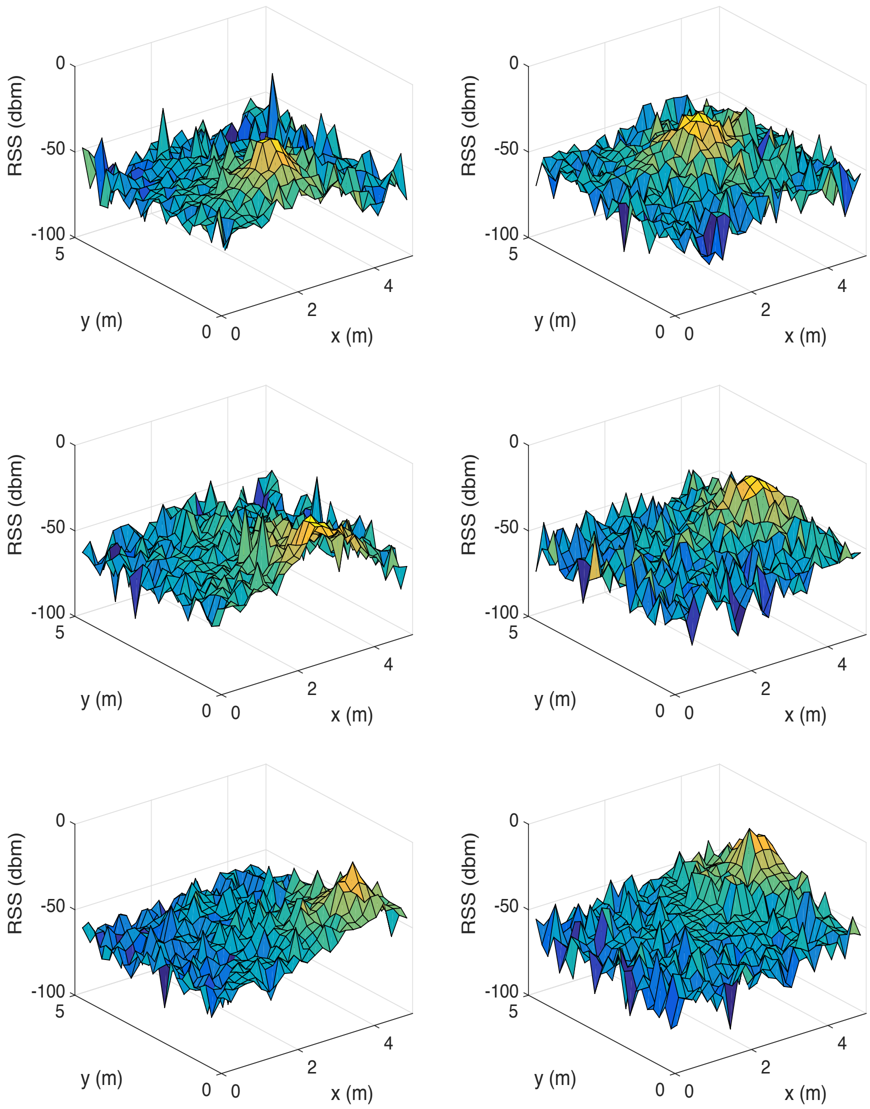


Figure 6.2: Ground truth RSS distributions of six RSS nodes.

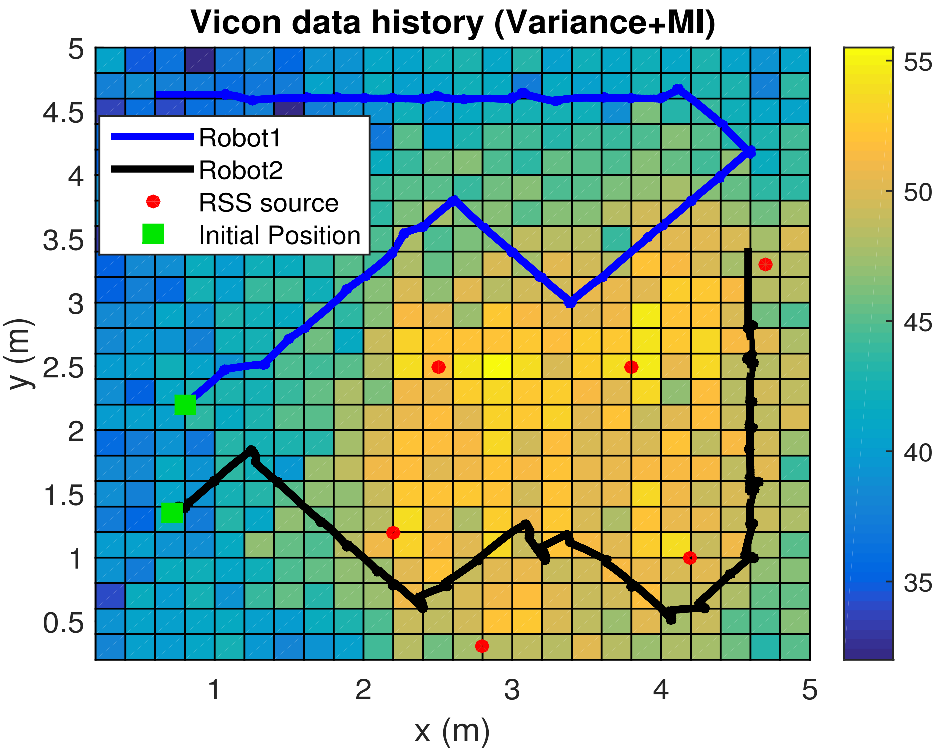


Figure 6.3: Trajectories of each robot in the experiment.

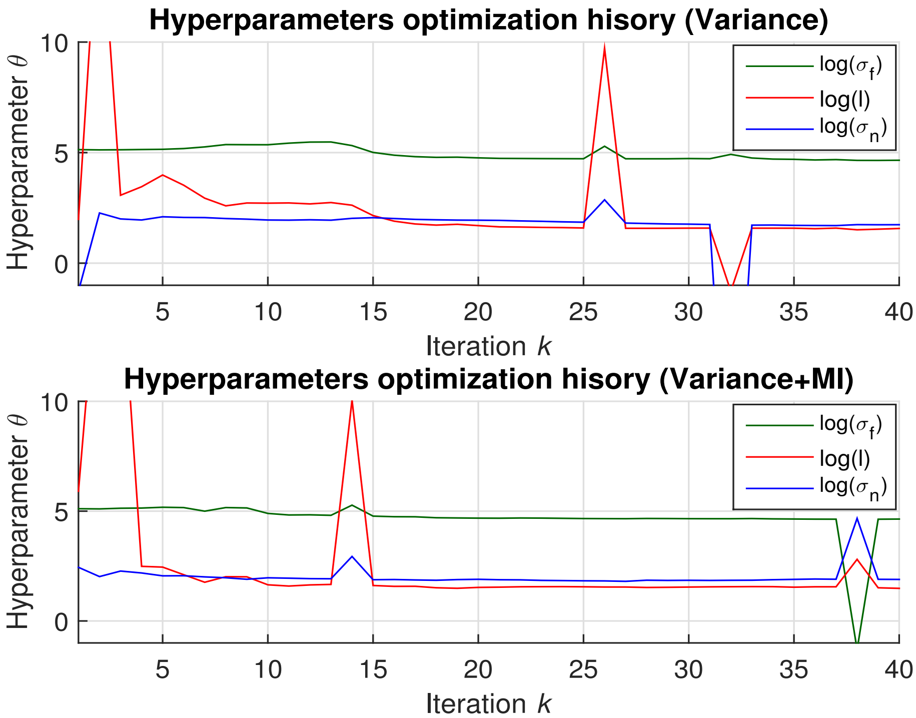


Figure 6.4: Hyperparameters optimization histories of variance-based algorithm (upper) and the proposed method (lower).

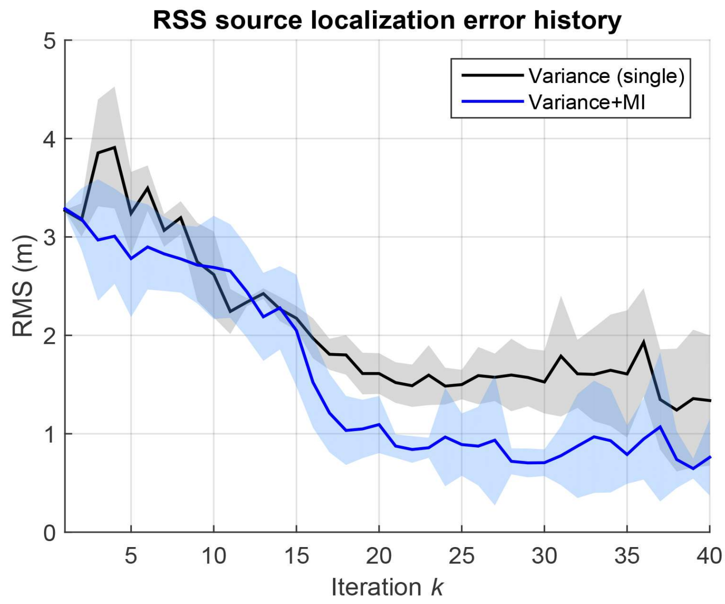


Figure 6.5: Comparison of RSS source localization error histories with 10 trials.

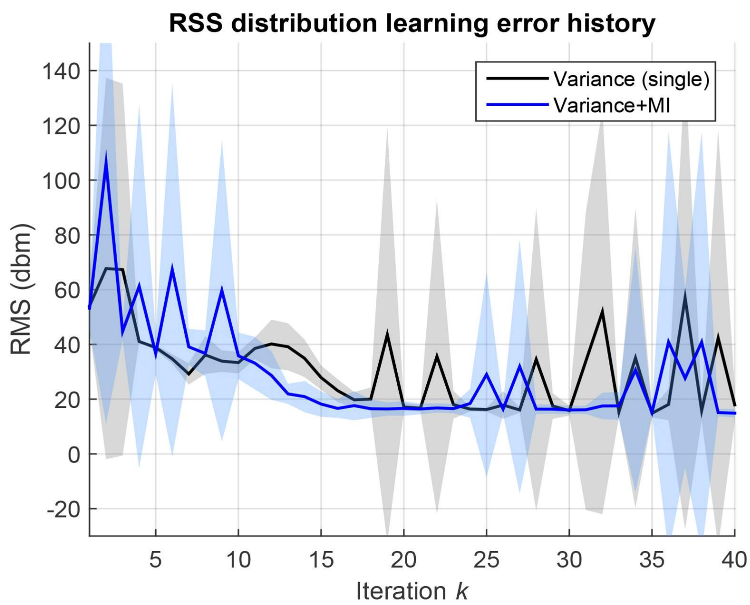


Figure 6.6: Comparison of RSS distribution model error histories against the ground truth with 10 trials.

6.2 Terrain mapping simulation settings and results for unsupervised method

6.2.1 Simulation settings

For performance validation of unsupervised method, we performed simulation with an open dataset which is provided by the United States geographical survey (USGS). The dataset, which is used in the simulation, is the digital elevation model (DEM) named 'n39w113'. The original data is shown in Figure 6.7. We extract a part of the data from the center as shown in Figure 6.8. Then, we apply Gaussian filter to the original data for smoothing, because the original elevation data are too coarsely digitized. The blurred result is also shown in Figure 6.8, and we use the result as the ground truth data in the simulation. The simulation settings are composed of four quadrotors and an unknown terrain environment. The quadrotors are equipped with a sensor, which can measure the height of the terrain at a specific location, such as a sonar sensor or a laser range sensor. We assume that the quadrotors are able to localize themselves, and the task is to learn the unknown model of the terrain. All parameter settings are identical to those of experimental settings, which is discussed in section 6.1.1. Besides, all of the quadrotors are assigned with an identical indicator initially, which means that only one Gaussian process model exists at the beginning.

6.2.2 Terrain mapping simulation result

The performance of the proposed unsupervised algorithm is compared with that of the variance-based algorithm. In Figures 6.9 and 6.10, the final estimated terrain model is presented by iterating 40 sampling steps, i.e. total 160 data is obtained by quadrotors. As shown in the figures, the terrain model learned using the proposed algorithms is more similar to the ground truth which is shown in Figure 6.8. Also, one of the quadrotors with proposed algorithm continuously search the area near the peak locations, while variance-based algorithm is not able to lead quadrotors to the area which is already searched.

Although the RMS error histories in Figure 6.11 too fluctuates because of several undesirable hyperparameter optimization, the informativeness of the data can be more clearly compared by reconstructing the terrain models with the hyperparameters being fixed. After the exploration is finished, we reproduced the estimated terrain model using the measurement histories at each iteration, by fixing the hyperparameters to those with the lowest RMS error. Then, the RMS error histories are shown in Figure 6.12. Although variance-based algorithm shows lower RMS error until 20 iterations, the proposed algorithm converged to lower RMS error. However, it cannot be said that the performance of the first 20 iterations means variance-based algorithm is more informative, because the algorithm does not consider the shape of the terrain. That is, variance-based algorithm would have searched the same area as this, even when the peak locations were changed to the bottom part, and would have shown higher RMS error at the beginning. Thus, the RMS error histories indicates that the proposed algorithm obtained more informative data, because the RMS error converged to lower value.

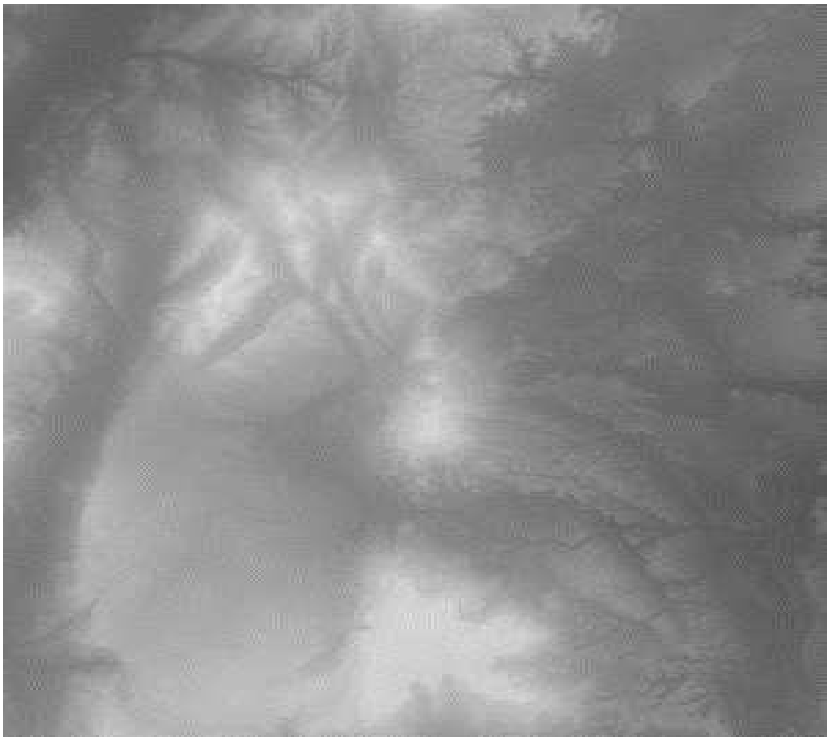
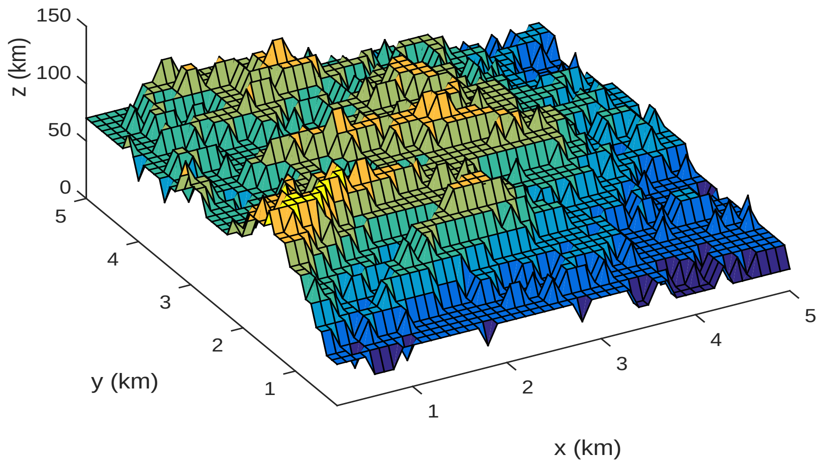


Figure 6.7: A raw terrain dataset from U.S Geographical Survey (USGS). (Website : 'http://eros.usgs.gov/find-data')



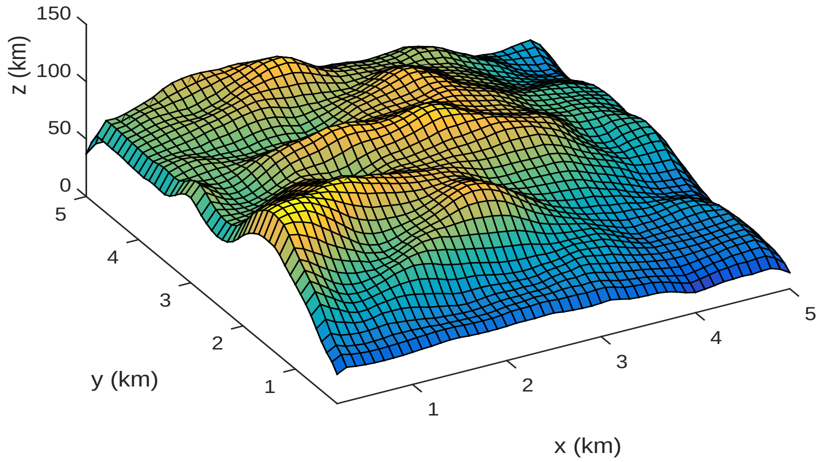


Figure 6.8: The selected area which is extracted from the center (above) and the filtered terrain dataset (bottom). The Gaussian filter is applied for smoothing digitized raw terrain data.

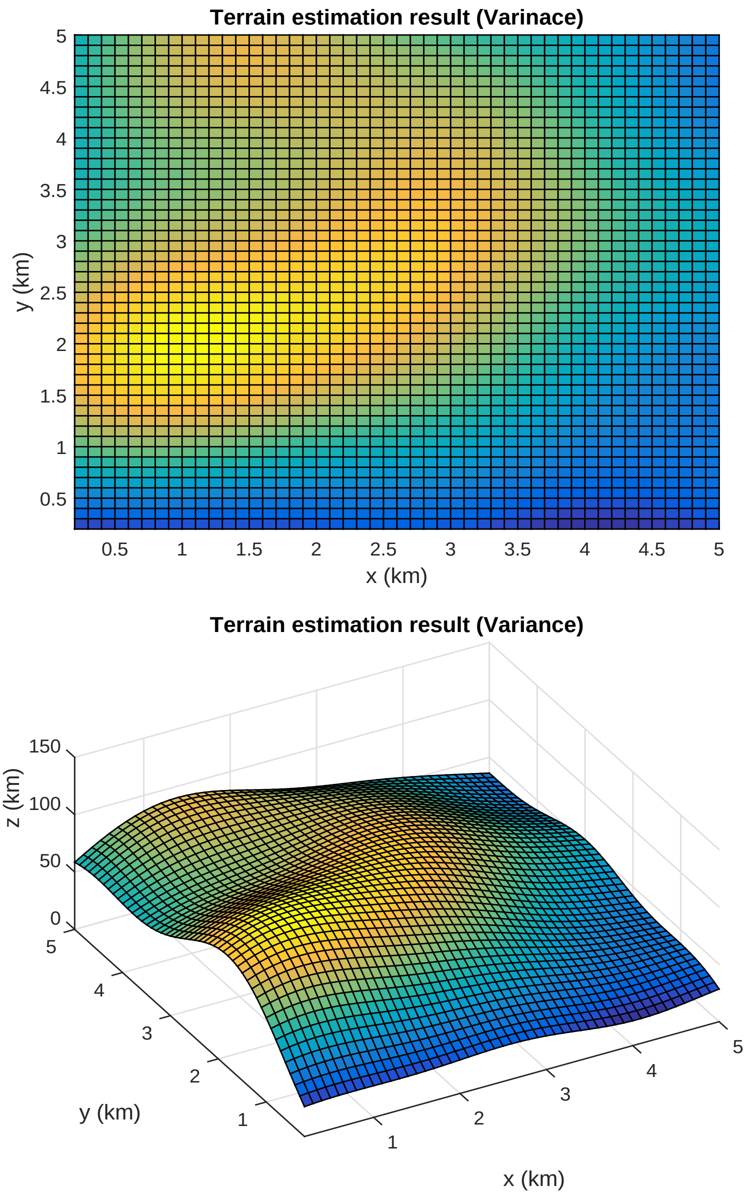


Figure 6.9: Terrain estimation result. (Variance-based with lowest RMS error)

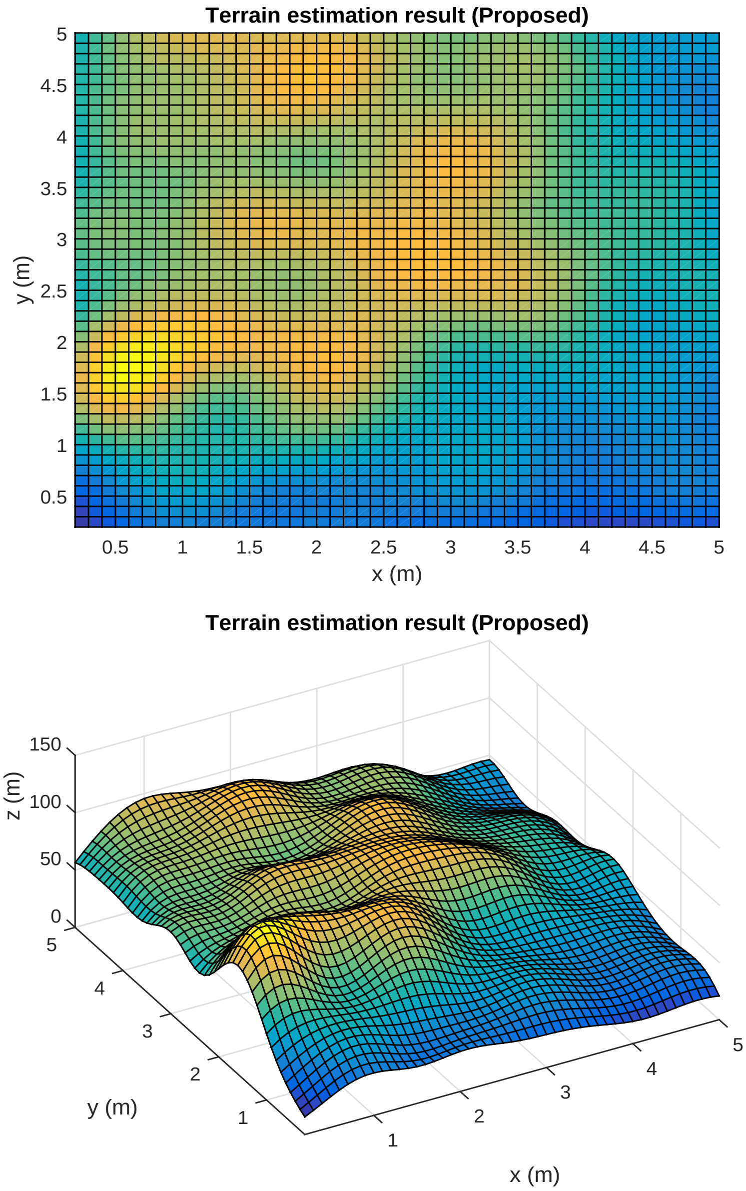


Figure 6.10: Terrain estimation result. (Proposed algorithm)

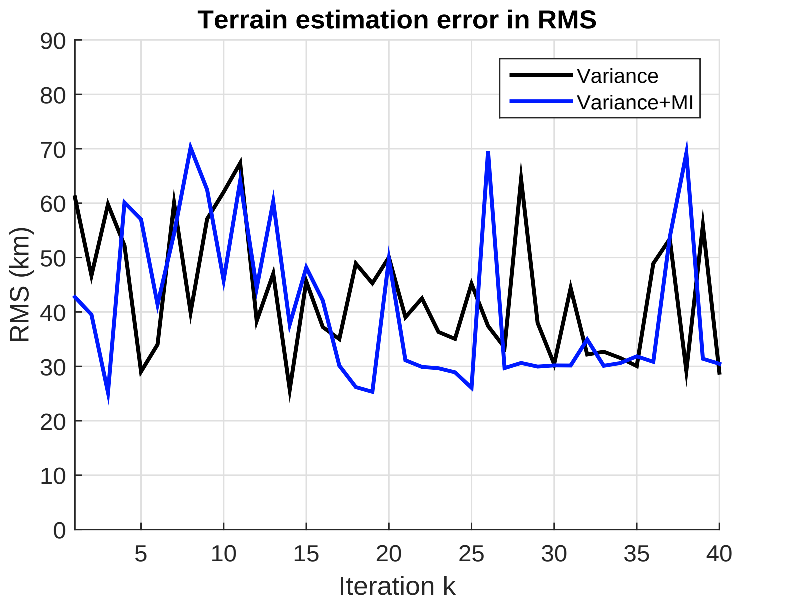


Figure 6.11: Comparison of terrain elevation estimation error in RMS.

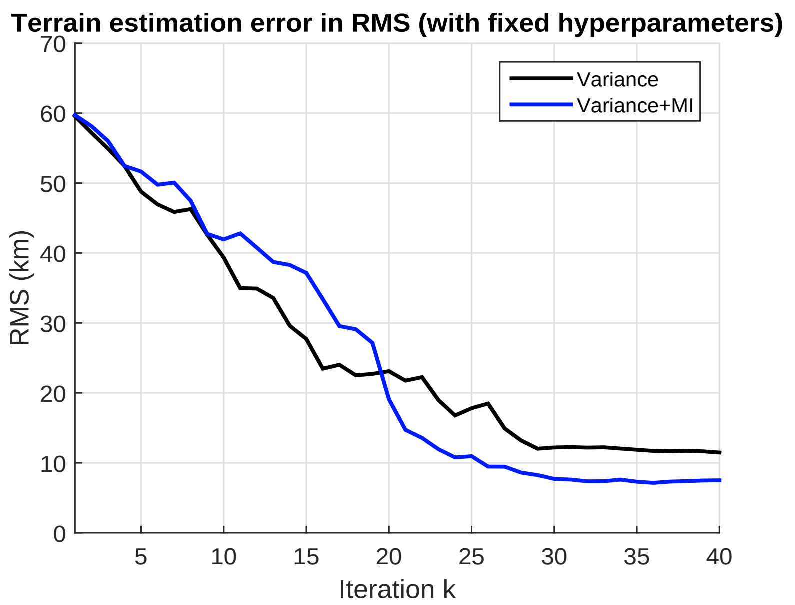


Figure 6.12: Comparison of terrain elevation estimation error in RMS. (with hyperparameter being fixed)

6.3 RSS distribution mapping experimental settings and results for unsupervised method

6.3.1 Experimental settings

To confirm real-world applicability of the unsupervised method, we validated the performance of the proposed algorithm with RSS distribution mapping scenario. Unlike supervised method of RSS distribution mapping which is discussed in section 6.1.1, we assumed the signal source ID is not given in the dataset. This kind of conditions can be found when base node continuously receives bad packets which contain inaccurate ID information. In such cases, the unsupervised method can be applied by considering the total sum of the RSS packet received at each position. Then, the RSS distribution mapping task can be performed by learning the distribution of the total sum of RSS measurements sent from each signal sources. The ground truth RSS distribution is shown in Figures 6.13 and 6.14, and all of the other parameter and hardware settings are equivalent to those of section 6.1.1.

6.3.2 RSS distribution mapping experimental result

The performance of the proposed algorithm is compared with the variance-based algorithm in terms of RMS error of the RSS distribution estimation result in this section. In Figure 6.15, the trajectories of two mobile robots in the experiment are presented, where the position data is obtained by Vicon. Some outliers are observed near the boundaries of the map, but the portion of the phenomena is extremely low so that the condition did not affect the movement of robots at all.

In Figures 6.16 and 6.17, the environmental learning task results are presented. The figures indicate that the performances of the two algorithms are similar, because the shape of the RSS distribution is very simple. In Figure 6.18, the graphs tend to fluctuate because of the fluctuating hyperparameters at each iteration, which was also discussed in Section 6.2.2. Thus, we also compared the performance by fixing the hyperparameters in Figure 6.19. It can be inferred that the proposed algorithm shows better performance, since our algorithm converges to lower RMS error. Also, the experimental result in Figure 6.15 indicates that the two robots tend to collect more data around the peak locations while also reducing the total uncertainty of the environment. Thus we confirmed the applicability of the proposed unsupervised method to real-world applications.

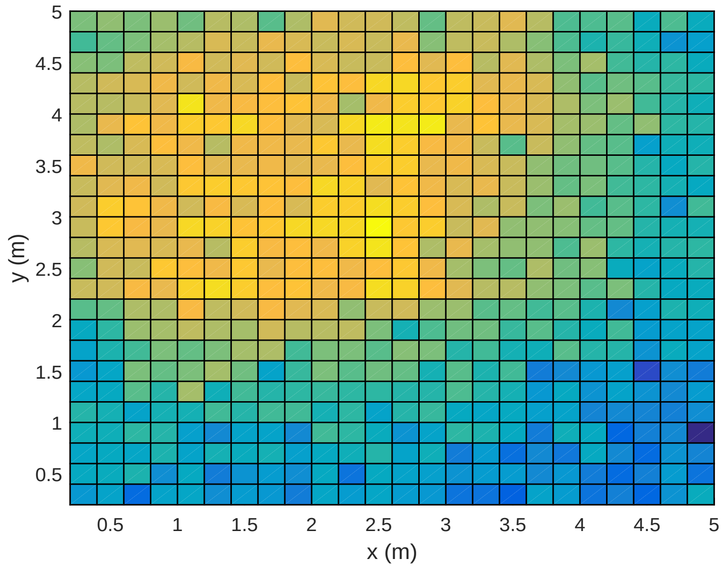


Figure 6.13: Ground truth RSS distribution (Top view).

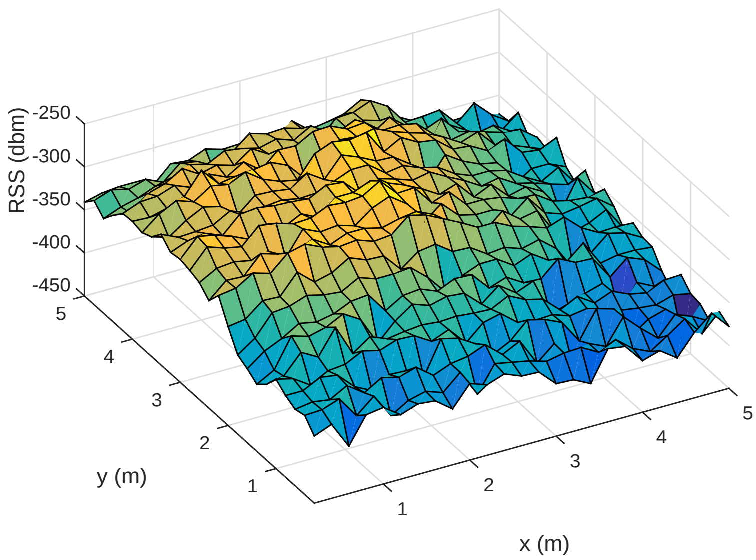


Figure 6.14: Ground truth RSS distribution (Side view).

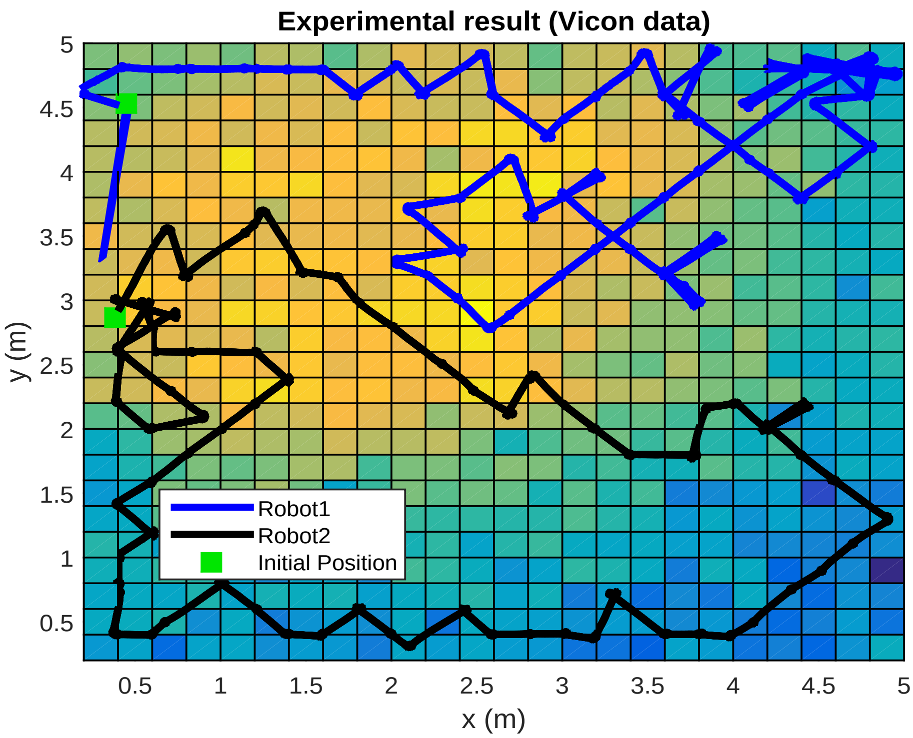


Figure 6.15: Trajectories of each robot in the experiment.

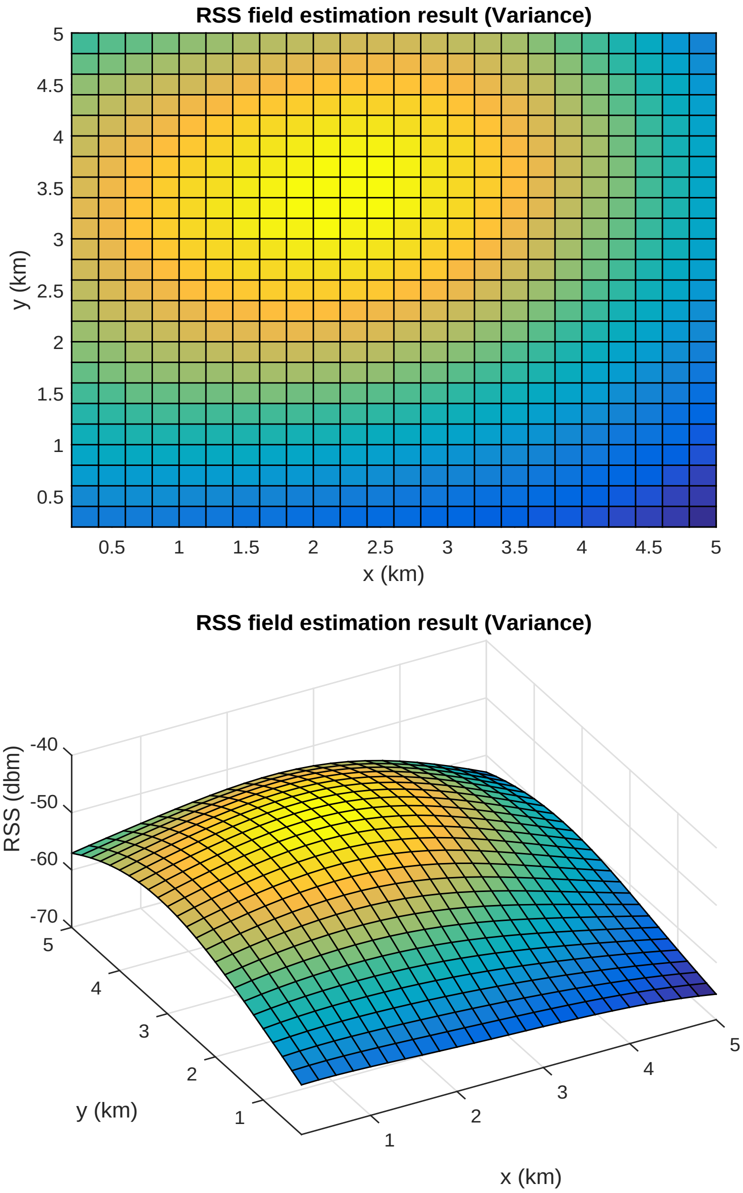


Figure 6.16: RSS distribution estimation result. (Variance-based algorithm)

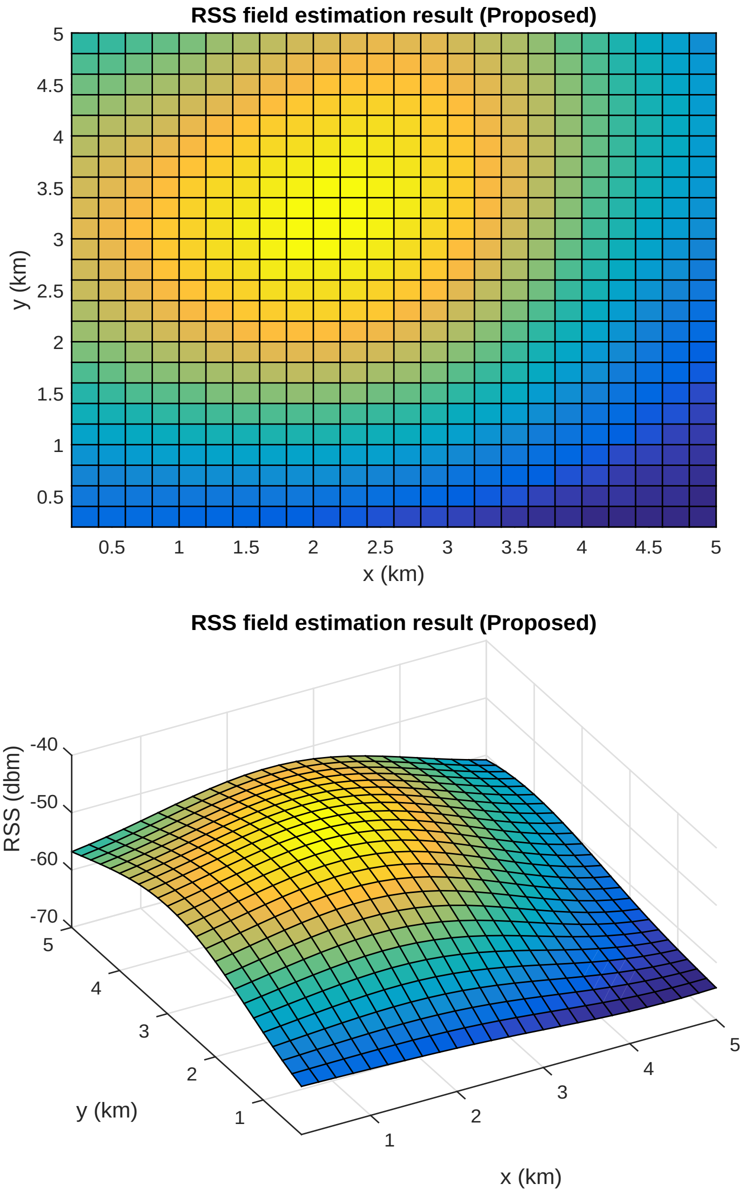


Figure 6.17: RSS distribution estimation result. (Proposed algorithm)

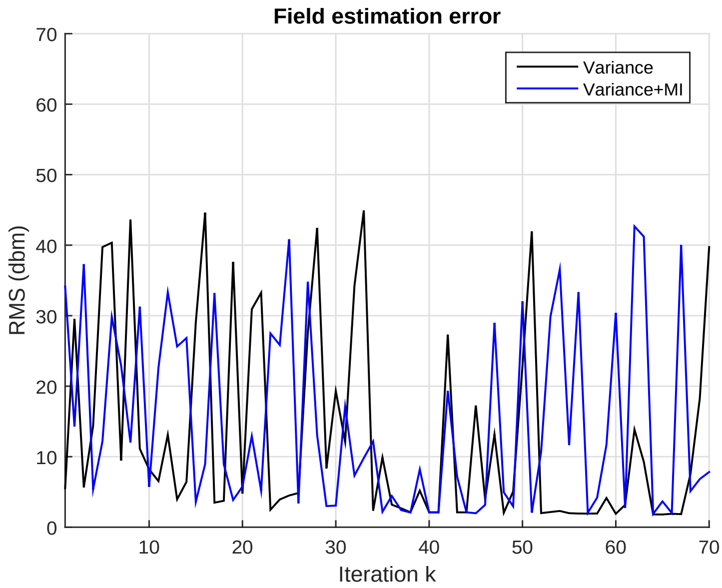


Figure 6.18: Comparison of RSS distribution estimation error in RMS.

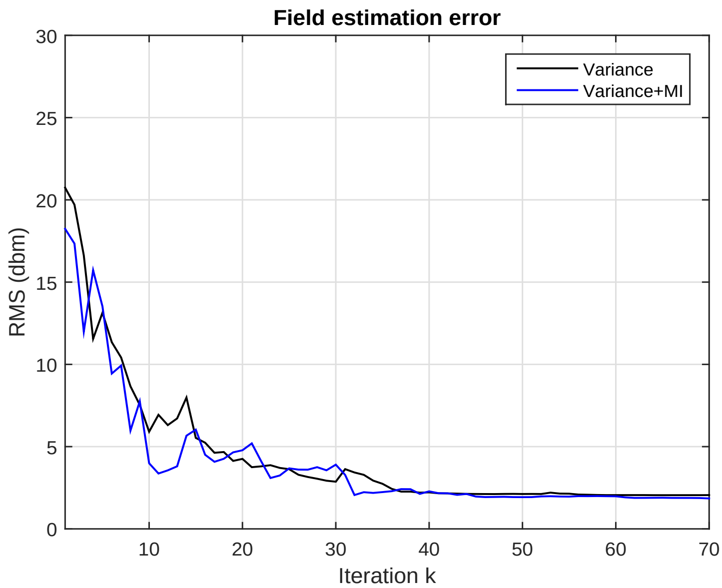


Figure 6.19: Comparison of RSS distribution estimation error in RMS. (with hyperparameter being fixed)

7

Conclusion

In this paper, a noise-robust planning policy for multi-agent exploration in an unknown environment is proposed, using a concatenated policy by fusing two objectives, i.e. variance and mutual information. Also, nonparametric methods are properly applied by using measurement-based spatial prior on the birth proces of GM-PHD filter. The algorithm is applied to decentralized mobile robots to find the most informative positions, while learning the unknown RSS distribution and localizing the source locations. Then, the proposed algorithm is also applied to the unknown terrain data, by applying unsupervised method with Dirichlet process mixture of Gaussian processes. We compared the obtained policy with variance-based one for performance validation. The experimental result of supervised method shows that the proposed algorithm provides more accurate estimations, and that the estimations converge faster. Moreover, the experiment suggests that the concatenated policy shows more robustness to the noisy signal with two objectives compensating problems of each other. Also, the simulation and experiment results of unsupervised method present the flexibility of the proposed algorithm. For the future work, adding the self-localization step may make the algorithm more flexible.

References

[1] C. E. Rasmussen, "Gaussian processes for machine learning," 2006.

[2] B. Ferris, D. Haehnel, and D. Fox, "Gaussian processes for signal strength-based location estimation," in *In proc. of robotics science and systems.* Citeseer, 2006.

[3] B.-N. Vo, S. Singh, and A. Doucet, "Sequential monte carlo methods for multitarget filtering with random finite sets," *IEEE Transactions on Aerospace and electronic systems*, vol. 41, no. 4, pp. 1224–1245, 2005.

[4] B.-N. Vo and W.-K. Ma, "The gaussian mixture probability hypothesis density filter," *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4091–4104, 2006.

[5] D. E. Clark, K. Panta, and B.-N. Vo, "The gm-phd filter multiple target tracker," in *2006 9th International Conference on Information Fusion.* IEEE, 2006, pp. 1–8.

[6] B.-T. Vo, B.-N. Vo, and A. Cantoni, "The cardinalized probability hypothesis density filter for linear gaussian multi-target models," in *Information sciences and systems, 2006 40th annual conference on.* IEEE, 2006, pp. 681–686.

[7] G. Battistelli, L. Chisci, C. Fantacci, A. Farina, and A. Graziano, "Consensus cphd filter for distributed multitarget tracking," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 3, pp. 508–520, 2013.

[8] J. Houssineau and D. Laneuville, "Phd filter with diffuse spatial prior on the birth process with applications to gm-phd filter," in *Information Fusion (FUSION), 2010 13th Conference on.* IEEE, 2010, pp. 1–8.

[9] B. Charrow, N. Michael, and V. Kumar, "Cooperative multi-robot estimation and control for radio source localization," *International Journal of Robotics Research*, vol. 33, no. 4, pp. 569–580, 2014.

[10] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, 2012.

[11] A. Viseras, T. Wiedemann, C. Manss, L. Magel, J. Mueller, D. Shutin, and L. Merino, "Decentralized multi-agent exploration with online-learning of gaussian processes," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 4222–4229.

[12] P. Dames and V. Kumar, "Autonomous localization of an unknown number of targets without data association using teams of mobile sensors," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 3, pp. 850–864, 2015.

[13] B. Charrow, G. Kahn, S. Patil, S. Liu, K. Goldberg, P. Abbeel, N. Michael, and V. Kumar, "Information-theoretic planning with trajectory optimization for dense 3d mapping," in *Proceedings of Robotics: Science and Systems*, 2015.

[14] P. Dames, P. Tokekar, and V. Kumar, "Detecting, localizing, and tracking an unknown number of moving targets using a team of mobile robots," *Intl. Sym. Robot. Research*, 2015.

[15] W. Kim, H. Lee, and H. J. Kim, "Predictive modeling of time-varying environmental information for path planning," in *2013 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 2013, pp. 3639–3644.

[16] G. M. Hoffmann and C. J. Tomlin, "Mobile sensor network control using mutual information methods and particle filters," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 32–47, 2010.

[17] S. Bai, J. Wang, F. Chen, and B. Englot, "Information-theoretic exploration with bayesian optimization," 2016.

[18] R. Ouyang, K. H. Low, J. Chen, and P. Jaillet, "Multi-robot active sensing of non-stationary gaussian process-based environmental phenomena," in *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2014, pp. 573–580.

[19] S. He and S.-H. G. Chan, "Wi-fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.

[20] J. Fink and V. Kumar, "Online methods for radio signal mapping with mobile robots," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on.* IEEE, 2010, pp. 1940–1945.

[21] Y. W. Teh, "Dirichlet process," in *Encyclopedia of machine learning.* Springer, 2011, pp. 280–287.

국 문 초 록

다수의 로봇을 이용한 미지 환경에서의 탐사 문제는, 다른 다양한 기술에 접목할 수 있는 정보를 수집하는데 있어서 매우 중요하다. 이에 본 논문에서는 수신신호세기 분포나 지형 분포와 같은 미지 환경에서 다수 로봇의 분산 경로 계획 기법을 다룬다. 이와 같은 설정을 위해서 수신신호세기 분포에 대한 지도식 기법과 지형 분포와 같은 일반적인 분포에 대한 비지도식 기법을 제시하였다. 미지의 환경은 각 로봇들이 습득한 데이터들을 공유하고 축적하여, 가우시안 프로세스를 이용한 온라인 학습을 통해 모델링하였다. 그리고 가우시안 프로세스의 평균 함수를 이용하여 신호 발생기의 위치를 추론하고, 추론 결과를 Gaussian mixture probability hypothesis density (GM-PHD) 필터를 이용하여 확률분포화 하였다. 또한, 가우시안 프로세스의 분산 함수와 상호정보량을 결합함으로써 잡음이 심한 데이터에 대해서도 강인하고 더 많은 정보를 내포한 데이터를 획득할 수 있는 탐사 기법을 제안하였다. 끝으로 디리클레 프로세스를 통한 가우시안 프로세스 혼합을 이용한 비지도식 기법을 적용함으로써 더 유연한 알고리즘으로 확장하였다. 제안된 알고리즘의 성능 검증을 위하여 지도식 기법과 비지도식 기법에 대한 시뮬레이션 및 실험 결과를 분석함으로써, 환경 분포 모델을 더 정확하고 빠르게 학습하는 것을 확인하였다.

주요어 : 베이지안 비모수 기법, 비지도식 학습, 디리클레 프로세스, 온라인 가우시안 프로세스, 상호정보량, GM-PHD필터, 능동 검출, 분산적 다중 로봇 제어

학번 : 2015-20788