International Journal of Control, Automation, and Systems (2012) 10(4):1-9
DOI

ISSN:1598-6446 eISSN:2005-4092
http://www.springer.com/12555

Stochastic Adaptive Sampling for Mobile Sensor Networks using Kernel Regression

Yunfei Xu and Jongeun Choi*

Abstract: In this paper, we provide a stochastic adaptive sampling strategy for mobile sensor networks to estimate scalar fields over surveillance regions using kernel regression, which does not require *a priori* statistical knowledge of the field. Our approach builds on a Markov Chain Monte Carlo (MCMC) algorithm, *viz.*, the fastest mixing Markov chain under a quantized finite state space, for generating the optimal sampling probability distribution asymptotically. The proposed adaptive sampling algorithm for multiple mobile sensors is numerically evaluated under scalar fields. The comparison simulation study with a random walk benchmark strategy demonstrates the excellent performance of the proposed scheme.

Keywords: Fastest mixing reversible markov chain, mobile sensor networks, stochastic adaptive sampling.

1. INTRODUCTION

Due to recent global climate changes, there is a growing number of environmental scientists who are interested in monitoring changes in detrimental environmental variables such as harmful algal blooms and cyanobacteria in oceans, in lakes, and public water systems [1]. Monitoring hazardous materials could be needed in non-convex surveillance regions such as airports and public transportation [2]. Mobile sensor networks and robotic technology provide a viable solution to such environmental monitoring [3-7]. Adaptive sampling can be utilized by a mobile sensor network such that later sampling is optimally scheduled considering the earlier sampling to improve the quality of the prediction [4,5]. In practice, mobile sensor networks using parametric regression could be limited in the sense that they require *a priori* knowledge about the model structure of the scalar fields [3,8]. Mobile sensing agents using nonparametric Gaussian processes also require a priori knowledge on the covariance functions [6].

To deal with the aforementioned problems, it is important to design theoretically-sound, adaptive sampling algorithms for mobile sensor networks using nonparametric regression such that they operate with no prior information about the scalar field. Due to the nature of the environmental monitoring, the difficulty of covering a large non-convex surveillance region by a limited number of mobile sensors with mobility constraints (such as range of movements and obstacles) has to be addressed as well.

Kernel regression and local linear regression techniques [9,10] provide an effective way for mobile sensors to perform adaptive sampling without prior information about the field of interest. For instance, an adaptive sampling strategy based on the local linear regression for a robotic boat to reconstruct a field was developed in [7]. However, due to the nonparametric approach without a key statistical structure (such as a covariance function in a Gaussian process), kernel regression instead requires a large number of samples for a good level of prediction. The asymptotic properties of local linear regression for univariate and multivariate cases were investigated in [11-13]. The estimation results using local linear regression highly depend on the choice of bandwidths. A large bandwidth may cause a large bias whereas a small bandwidth may result in a large variance [14]. An effective bandwidth selector for local least squares regression was proposed in [10]. Plug-in bandwidths for multivariate kernel density estimation was introduced in [15]. The optimal bandwidths for multivariate local linear regression were established in [14]. Two classes of variable bandwidth selectors (balloon and sample-point selectors) for kernel density estimation which are more flexible as compared to the fixed bandwidths were introduced in [16]. In kernel regression, its sampling (continuous) probability distribution can be considered as a weighting function to be optimized for minimizing the estimation error. For a fixed number of sampling positions, it can be viewed as an optimal resource allocation problem. In this problem,

Manuscript received May 18, 2011; revised November 10, 2011; accepted May 19, 2012. Recommended by Editorial Board member Pinhas Ben-Tzvi under the direction of Editor Young Il Lee.

This work has been supported by the National Science Foundation through CAREER Award CMMI-0846547. This support is gratefully acknowledged.

Yunfei Xu is with the Department of Mechanical Engineering, Michigan State University, MI 48824, USA (e-mail: xuyunfei@egr.msu.edu).

Jongeun Choi is with the Department of Mechanical Engineering and the Department of Electrical and Computer Engineering, Michigan State University, MI 48824, USA (e-mail: jchoi@egr.msu.edu).

* Corresponding author.

Springer

sampling positions have to be optimally allocated for the scalar field of interest such that the resulting kernel regression with the corresponding optimal variable bandwidths produces the minimal estimation error in average.

A Markov Chain Monte Carlo (MCMC) algorithm such as Metropolis-Hastings and maximum-degree Markov chains [17] can be used to coordinate mobile agents to generate the optimal sampling (discrete) probability distribution in a mobility-constraints encoded graph. MCMC based stochastic rules to generate way-points for the agents throughout their surveillance missions have been developed in [18]. In [19], a decentralized strategy was proposed to reallocate a swarm of robots to multiple tasks following a target distribution which is optimized for the fast convergence subject to transition constraints. The convergence rate of the Markov chain to the prescribed equilibrium probability distribution is related to the second largest eigenvalue modulus (SLEM) of the Laplacian matrix associated with the transition probability [18,20]. Finding the fastest mixing reversible Markov chain (FMRMC) on a given graph was successfully formulated as a convex optimization problem in [20].

The contribution of the paper is as follows. First, we formulate and solve an optimal sampling problem for a set of target points to estimate an unknown scalar field over a (possibly non-convex) surveillance region using kernel regression. Next, we propose a stochastic adaptive sampling strategy for mobile sensor networks taking into account mobility constraints, and we show that it achieves the optimal sampling probability distribution asymptotically. Our approach builds on a MCMC algorithm, *viz.,* the FMRMC, under a quantized finite state space for generating the optimal (discrete) probability distribution asymptotically. In particular, a partition of the surveillance region is generated to obtain a finite state space for a Markov chain. To design a Markov chain, an associated graph shall be constructed by taking into account mobility constraints of sensing robots. For a single mobile robot, in each iteration, the robot randomly moves to a neighboring partitioned cell by the FMRMC, and its sampling position is randomly selected within the selected cell according to the optimal (continuous) probability distribution. An adaptive sampling algorithm for multiple mobile sensors to cover a large surveillance region is designed and numerically simulated under scalar fields. The excellent performance of the proposed algorithm as compared to a random walk benchmark is demonstrated in the simulation study.

This paper is organized as follows. In Section 2, local linear regression and its mean square error are reviewed. In Section 3, a concept of target positions and its associated optimal sampling distribution are introduced. A stochastic strategy for mobile sensors for the optimal sampling and its asymptotic convergence properties are presented in Section 4. Stochastic adaptive sampling for multiple robotic sensors over a large surveillance region is provided in Section 5. Section 6.1 provides numerical experiments to demonstrate the effectiveness of the proposed algorithm.

Standard notation is used throughout the paper. Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$, $\mathbb{Z}$, $\mathbb{Z}_{\geq 0}$, $\mathbb{Z}_{>0}$ denote, respectively, the sets of real, non-negative real, positive real, integer, non-negative integer, and positive integer numbers. $\mathscr{A} \cup \mathscr{B}$ and $\mathscr{A} \cap \mathscr{B}$ denote, respectively, the union and intersection of the sets $\mathscr{A}$ and $\mathscr{B}$. Let $\mathbb{E}(x)$ and $\mathrm{Var}(x)$ denote, respectively, the expectation and the variance of x. Other notation will be explained in due course.

2. LOCAL LINEAR REGRESSION

In this section, we briefly review local linear regression and its bandwidth [13].

Let $\{X^{(i)}, Y^{(i)}\}_{i=1}^n$ be the collection of measurements from robotic sensors. $X^{(i)} = (X_1^{(i)}, \cdots, X_d^{(i)})^T \in \mathbb{R}^d$ is a sampling location, which is a random vector generated by a density function $p(x)$, and $Y^{(i)} \in \mathbb{R}$ is a scalar environmental variable of interest. Consider the following multivariate regression model for the environmental scalar field

$$Y^{(i)} = f(X^{(i)}) + \varepsilon,$$

where ε is independent of $X^{(i)}$ satisfying $\mathbb{E}(\epsilon) = 0$ and $\mathrm{Var}(\varepsilon) = \sigma^2$. Note that $f : \mathbb{R}^d \mapsto \mathbb{R}$ and $\sigma^2 > 0$ is the noise level. The local linear estimator of f at a given point $x = (x_1, \cdots, x_d)^T \in \mathbb{R}^d$ is obtained by applying a first-order Taylor expansion of the function f at point x for all the data points in $\{X^{(i)}\}_{i=1}^n$ and solving a least squares problem locally weighted by kernels. The estimator $\hat{f}(x)$ is the first element of $\beta = (\beta_0, \cdots, \beta_d)^T \in \mathbb{R}^{d+1}$ that minimizes

$$\sum_{i=1}^{n} \left\{ Y^{(i)} - \beta_0 - \sum_{j=1}^{d} \beta_j (X_j^{(i)} - x_j) \right\}^2 K_H(X^{(i)} - x),$$

where $K_H(u) = |H|^{-1/2} K(H^{-1/2} u)$ with a bandwidth matrix $H^{1/2} \in \mathbb{R}^{d \times d}$. K is a symmetric, compactly supported, univariate probability kernel that satisfies $\int K(u) du = 1$. Hence, assuming that $X_x^T W_x X_x$ is nonsingular, the local linear estimator is given by

$$\hat{f}(x; H) = e_1^T \left(X_x^T W_x X_x \right)^{-1} X_x^T W_x Y, \tag{1}$$

where

$$X_x = \begin{bmatrix} 1 & (X^{(1)} - x)^T \\ \vdots & \vdots \\ 1 & (X^{(n)} - x)^T \end{bmatrix}, \quad Y = \begin{bmatrix} Y^{(1)} & \cdots & Y^{(n)} \end{bmatrix}^T,$$

$W_x = \mathrm{diag}\left\{ K_H(X^{(1)} - x), \cdots, K_H(X^{(n)} - x) \right\}$, and

$e_1 = (1, 0, \cdots, 0)^T \in \mathbb{R}^{d+1}$.

Since $p(x)$ is a density that generates $\{X^{(i)}\}$, we have $\int p(x) dx = 1$, $\mathscr{H}_f(x)$ be the Hessian matrix of function

$f(x)$. The asymptotic conditional mean squared error (MSE) at a point x in the domain of interest is given by [13]

$$\begin{aligned}\mathrm{MSE}(x) &= \mathbb{E}\{[\hat{f}(x)-f(x)]^2 \mid X^{(1)},\cdots,X^{(n)}\}\\ &= \frac{R(K)\sigma^2}{n\,|H|^{1/2}\;p(x)} + \frac{1}{4}\mu_2^2(K)\mathrm{tr}^2\{H\mathscr{H}_f(x)\} \quad (2)\\ &\quad + o_p\{n^{-1}|H|^{-1/2} + \mathrm{tr}^2(H)\},\end{aligned}$$

where $R(K)=\int K(u)^2du$, and $\mu_2(K)I=\int uu^TK(u)du.$

The estimation quality highly depends on the selection of the bandwidth matrix. The bandwidth matrix in (2) is fixed for simplicity. In this paper, we use the so called balloon bandwidth selector in which the bandwidth is a function of x. The bandwidth matrix has been parameterized by $H^{1/2}(x)=h(x)I$. The optimal bandwidth that minimizes the MSE can be obtained by

$$h(x)=\left(\frac{dR(K)\sigma^2}{np(x)\mu_2^2(K)\mathrm{tr}^2\{\mathscr{H}_f(x)\}}\right)^{1/(d+4)}. \qquad (3)$$

In the next section, we introduce the concept of target positions and its associated optimal sampling distribution.

3. OPTIMAL SAMPLING DISTRIBUTION

In this paper, positions of interest will be referred to as target positions. The introduction of target positions is motivated by the fact that the potential environmental concerns should be monitored with a higher resolution. For instance, the target positions can be assigned at the interface of a factory and a lake, sewage systems, or polluted beaches. Thus, the introduction of target positions, which can be arbitrarily specified by a user, provides a flexible way to define a geometrical shape of a subregion of interest in a surveillance region. Without *a priori* knowledge on such concerns, we will take uniformly distributed grid points as target positions.

According to the selected target positions $\mathscr{T}=\{U^{(i)}\}_{i=1}^n$, we now quantize the region of interest $D\in\mathbb{R}^d$ into n finite sets so that the stochastic sampling strategy can be derived in a finite state space. For a given compact and closed region of interest D, we design a partition of D, denoted by $\Lambda(D):=\{D_i\}_{i=1}^n$, which satisfies

$$D=\bigcup_{i=1}^n D_i, \mathrm{Int}D_i\cap\mathrm{Int}D_j=\varnothing \text{ for } i\neq j, U^{(i)}\in\mathrm{Int}D_i,$$

where Int A denotes the interior of A. Hence $\{D_i\}_{i=1}^n$ is a collection of nonempty subsets of D whose interior is disjoint and whose union is D.

In our approach, the performance cost function is chosen to be the Average Mean Squared Error (AMSE) on n target positions given by

$$\mathrm{AMSE}=\frac{1}{n}\sum_{i=1}^n \mathrm{MSE}(U^{(i)}).$$

By plugging in the optimal bandwidths that minimize MSE, the AMSE $J(p)$ is obtained by

$$J(p):=c\sum_{i=1}^n\left(\frac{\mathrm{tr}^d\{\mathscr{H}_f(U^{(i)})\}}{n^2p^2(U^{(i)})}\right)^{2/(d+4)}, \qquad (4)$$

where c is a constant.

Problem 1: The goal of optimal sampling is to minimize the AMSE by selecting an optimal continuous probability distribution from a family of continuously differentiable probability density functions $p(x)$ with

$$\int_{x\in D}p(x)dx=1. \qquad (5)$$

Our proposed approach to solve Problem 1 is given as follows. From (4), we parameterize the probability distribution by $p_i:=p(U^{(i)})$. We also pre-select the *weighted area* of D_i denoted by ω_i, which provides the following constraint.

$$\omega_i:=\frac{\int_{x\in D_i}p(x)dx}{p_i}. \qquad (6)$$

Hence the probability of x being in D_i is simply $\omega_i p_i$ and $\sum_{i=1}^n\omega_ip_i=1$ is always satisfied due to (5).

The optimal parameter set $\{p_i^\star\}_{i=1}^n$ is derived in the following proposition.

Proposition 2: Consider the following optimization

$$\begin{aligned}&\text{minimize} \quad J(p)/c\\ &\text{subject to} \quad \sum_{i=1}^n\omega_ip_i=1.\end{aligned}$$

The optimal solution is given by

$$p_i^\star=\frac{\frac{1}{\omega_i}\mathrm{tr}^{2d/(d+8)}\{\mathscr{H}_f(U^{(i)})\}}{\sum_{j=1}^n(\frac{\omega_j}{\omega_i})^{4/(d+8)}\mathrm{tr}^{2d/(d+8)}\{\mathscr{H}_f(U^{(j)})\}}. \qquad (7)$$

Notice that $p_i^\star$ is proportional to $\mathrm{tr}^{2d/(d+8)}\{\mathscr{H}_f(U^{(i)})\}$ when $\omega_i=\omega_j,\forall i,j.$

Proof: Note that $J(p)\geq 0$. By introducing Lagrange multipliers λ, we define the Lagrange function as

$$L(p,\lambda)=\sum_{i=1}^n\left(\frac{\mathrm{tr}^d\{\mathscr{H}_f(U^{(i)})\}}{n^2p_i^2}\right)^{2/(d+4)}+\lambda\left(\sum_{i=1}^n\omega_ip_i-1\right).$$

Equation (7) is obtained by solving $\nabla_{p,\lambda}L(p,\lambda)=0$. □

We introduce the following definition.

Definition 3: Consider a pair of a collection of target positions and its associated partition, i.e., $(\mathscr{T},\Lambda(D))$. A continuously differentiable probability distribution that satisfies (6) and (7) will be referred to as an *optimal* sampling continuous probability distribution $p^\star(x)$ with respect to $\mathscr{T}$.

4. STOCHASTIC STRATEGY FOR OPTIMAL SAMPLING

In this section, we propose a stochastic strategy under which mobile sensors with mobility constraints can be coordinated to generate the optimal sampling probability distribution $p^\star(x)$ in Definition 3 asymptotically. We first present a strategy and then prove its asymptotic convergence properties.

Consider the quantized optimal discrete probability distribution $\pi = (\pi_1, \cdots, \pi_n)^T$:

$$\pi_i := \omega_i p_i^\star = \int_{x \in D_i} p^\star(x) dx, \tag{8}$$

where ω_i is the weighted area of D_i. Our stochastic rule consists of two steps. In the first step, while satisfying mobility constraints, a robotic sensor randomly moves to a cell, e.g., D_i in a quantized state space to generate the discrete probability distribution π asymptotically. In the second step, when the robotic sensor is assigned to D_i by the first step, its sampling position is randomly generated proportionally to the optimal sampling continuous probability distribution $p^\star(x)$ over D_i.

To develop such a stochastic rule for the first step, we use a Markov Chain Monte Carlo (MCMC) algorithm [17] to randomly assign a mobile sensor over a partitioned region by synthesizing a Markov chain that has the target distribution π as its equilibrium distribution.

Consider an undirected graph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ with a vertex set $\mathcal{V} = \{1, \cdots, n\}$ and an edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Each vertex is associated to a cell of the surveillance region D. Hence, D_i is indexed by a vertex index $i \in \mathcal{V}$. The edge set $\mathcal{E}$ shall be defined to reflect the mobility constraints of a mobile sensor and obstacles in the surveillance region. For instance, we could have $(i, j) \in \mathcal{E}$, if $D_i \cap D_j \neq \emptyset$, which means that if the two partitions are neighbors, there is an (undirected) edge between two corresponding vertices (see Fig. 1). We also assume that each vertex has a self-loop, i.e., $(i,i) \in \mathcal{E}$ for $\forall i \in \mathcal{V}$, which allows a mobile sensor to stay in the same cell with some probability. The state at time $t \in \mathbb{Z}_{\geq 0}$ is denoted by $q(t) \in \mathcal{V}$. We define a Markov chain by a transition probability $P \in \mathbb{R}_{\geq 0}^{n \times n}$ as follows.

$$P_{ij} := \begin{cases} \Pr(q(t+1) = j \mid q(t) = i)) > 0, & \forall (i, j) \in \mathcal{E}; \\ \Pr(q(t+1) = j \mid q(t) = i)) = 0, & \forall (i, j) \notin \mathcal{E}. \end{cases}$$

A Markov chain on a graph $\mathcal{G}$ with the transition probability P is called *reversible* with respect to an equilibrium probability distribution π on $\mathcal{G}$ if

$$\pi_i P_{ij} = \pi_j P_{ji}, \quad \forall i, j \in \mathcal{V}.$$

To maximize the rate of convergence to the target distribution π, we solve the fastest mixing reversible Markov chain (FMRMC) problem [20] in which the transition probability matrix P is designed to maximize the mixing rate of the reversible Markov chain. Hence, by the FMRMC, the associated discrete probability distribution on the vertices approaches to the equilibrium (optimal) probability distribution π as rapidly as possible.

We now design the FMRMC that minimizes the second largest eigenvalue modulus (SLEM), which determines the mixing rate, by solving the following convex optimization problem [20].

$$\begin{aligned} &\text{minimize} \quad \left\| \Pi^{1/2} P \Pi^{-1/2} - r r^T \right\|_2 \\ &\text{subject to} \quad P \geq 0, \quad P1 = 1, \quad \Pi P = P^T \Pi \\ &\qquad\qquad\quad P_{ij} = 0, \quad \forall (i, j) \notin \mathcal{E}, \end{aligned} \tag{9}$$

where $\Pi = \mathrm{diag}(\pi)$, $r = (\sqrt{\pi_1}, \cdots, \sqrt{\pi_n})^T$, the inequality $P \geq 0$ means element-wise (i.e., $P_{ij} \in \mathbb{R}_{\geq 0}, \forall i, j$) and $1 \in \mathbb{R}^n$ is a vector of all ones. This optimization problem can be solved globally and efficiently using standard SDP solvers [20].

We summarize our stochastic strategy for optimal sampling by the following theorem.

Theorem 4: Consider our stochastic strategy that consists of two steps:

Step 1: *Mobility among cells* is driven by the FMRMC with the transition probability P obtained from (9) over the specified graph $\mathcal{G}$ and the partition $\Lambda(D)$.

Step 2: *Mobility within a cell* is driven by random sampling proportionally to $p^\star(x)$ in the assigned cell from Step 1, i.e., if $q(t) = i$, then the sampling position $x \in D_i$ is randomly distributed by $\tilde{p}_i^\star(x)$:

$$\tilde{p}_i^\star(x) := \frac{p^\star(x)}{\int_{\mu \in D_i} p^\star(\mu) d\mu}, \tag{10}$$

such that $\int_{x \in D_i} \tilde{p}_i^\star(x) dx = 1$.

A mobile sensor under this stochastic rule will then satisfy the mobility constraints and generate a sampling distribution that converges to the optimal continuous probability distribution $p^\star(x)$ asymptotically.

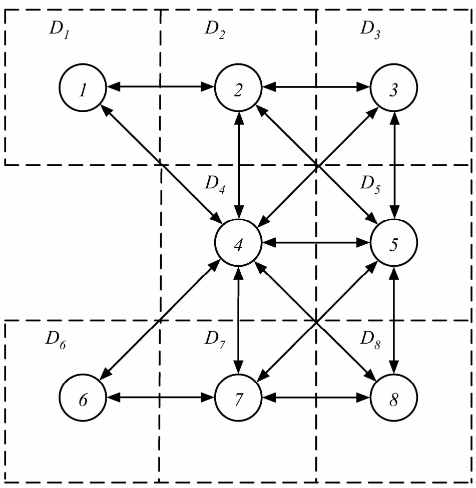


Fig. 1. A graph over a partition $\{D_i\}$ of D.

Proof: Consider a mobile sensor under this stochastic rule. The mobility constraints are satisfied since the edge set $\mathscr{E}$ of the associated graph $\mathscr{G}$ is constructed to reflect the mobility constraints. The discrete probability distribution generated by the FMRMC converges to π, which satisfies (8), i.e., $\pi_i = \omega_i p_i^\star = \int_{\mu \in D_i} p^\star(\mu) d\mu$, $\forall i \in \mathscr{V}$ asymptotically. For D_i, $\forall i \in \mathscr{V}$, the sampling probability distribution over D_i for given $q = i$, converges to $\tilde{p}_i^\star(x)$ asymptotically. Therefore, the sampling distribution $p(x)$ converges to the optimal continuous probability distribution $\pi_i \tilde{p}_i^\star(x) = p^\star(x)$ for $x \in D_i$ asymptotically. □

In the case when multiple mobile sensors are coordinated by the same Markov chain independently, the probabilistic collision avoidance capability of sampling positions of multiple mobile sensors is provided by the following proposition.

Proposition 5: Suppose m sensors are assigned to D_i, i.e., $q(t) = i$ for all m sensors at time $t \in \mathbb{Z}_{>0}$. The probability of the collision of any two sampling points of mobile sensors is zero.

Proof: Let $B_r[s]$ be a closed ball of radius $r > 0$ centered at a point s defined by $B_r[s] := \{x \in D_i \mid \| x - s \| \le r\}$. The sampling position X of sensor k can occur any point $s \in D_i$. The probability that the sampling position Y of sensor ℓ belongs to a closed ball $B_r[s] \subseteq D_i$ is $\Pr(Y \in B_r[s]) = \int_{Y \in B_r[s]} \tilde{p}_i^\star(Y) dY$. Two random variables X, Y are independent, therefore, the probability that two sampling positions of sensors collide is

$$\Pr(X = Y) = \lim_{r \to 0} \int_{s \in D_i} \int_{Y \in B_r[s]} \tilde{p}_i^\star(X = s) \tilde{p}_i^\star(Y) dY dX$$
$$= \lim_{r \to 0} \int_{s \in D_i} \tilde{p}_i^\star(X = s) dX \int_{Y \in B_r[s]} \tilde{p}_i^\star(Y) dY = 0.$$

This argument can be extended for m sensors. □

In practice, robots with non-negligible volume will be coordinated to visit the assigned sampling points in sequence to avoid collision.

This stochastic strategy for optimal sampling requires the knowledge of the Hessian matrix used in (7) for calculating $\{p_i^\star\}_{i=1}^n$ to produce the optimal distribution π. One way to deal with this issue is to use the estimated Hessian matrix from a set of initially observed samples.

5. STOCHASTIC ADAPTIVE SAMPLING

In this section, we provide an adaptive sampling algorithm for a mobile sensor network to cope with a large surveillance region D with respect to the mobility range of sensors. In this case, it is necessary to use multiple mobile sensors that are distributed efficiently over D. Often an attempt to cover a relatively large area by a single mobile sensor yields an infeasibility in the optimization in (9). To address this problem, robotic sensor j will take charge of subregion j, denoted by S_j, which is defined by a union of its cells $S_j := \bigcup_{i \in \mathscr{I}_j} D_i$, where $\mathscr{I}_j$ is the vertex set of the connected subgraph $\mathscr{G}_j = (\mathscr{I}_j, \mathscr{E}_j)$ generated by S_j considering mobility constraints of robot j. We also constrain this collection of subregions to be a partition of D, i.e.,

$$D = \bigcup_{i=1}^{m} S_i, \mathrm{Int}S_j \cap \mathrm{Int}S_\ell = \varnothing \text{ for } j \neq \ell.$$

The key constraint of designing subregions is that the desired sampling probability over each subregion has to be equal to $1/m$ with some small quantization error. In this way, the equilibrium distributions of m FMRMC's will produce the near-optimal discrete sampling distribution. A simplest method of partitioning the region D into $m = 2^k$, $k \in \mathbb{Z}_{>0}$ subregions $\{S_j\}_{j=1}^m$ is shown in Table 1. However, this approach can be used only when m is the power of 2 and the quantization error is not guaranteed to be minimal. On the other hand, any appropriate equitable partitioning algorithm such as a quantized gossip consensus algorithm [21] can be applied for equitable partitioning of the region D into m subregions $\{S_j\}_{j=1}^m$.

We now provide a stochastic adaptive sampling algorithm for multiple robotic sensors as described in Table 2.

Table 1. The simplest equitable partitioning algorithm.

Input:	(1) The partition $\Lambda(D)$ (2) The number of agents (subregions) $m = 2^k$ (3) The optimal discrete probability distribution π
Output:	The partition $\{S_j\}_{j=1}^m$
1: **for** $i = 1{:}k$ **do**	
2:	Divide each region (the entire region when $i = 1$) into subregions over which has the similar sampling probability.
3: **end for**	

6. EXAMPLES

6.1. Simulated examples

We consider a scenario in which $m = 4$ mobile sensors perform the estimation task. Two static scalar fields to be estimated are shown in Fig. 2. The analytical expression of the scalar fields in Fig. 2(a) and (b) are

$$f_1(x) = \exp\left\{ -\frac{(x_1 - 4.5)^2 + (x_2 - 5.5)^2}{10} \right\},$$

and

$$f_2(x) = \frac{1}{2} \sin(0.1\pi x_1) \sin(0.2\pi x_2),$$

respectively. The surveillance region D is given by $D = (0,10)^2$. The measurement noise level σ is chosen

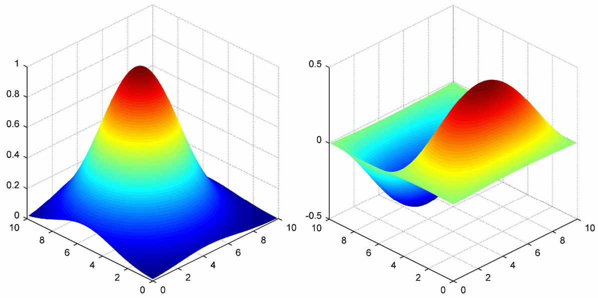
Table 2. Stochastic adaptive sampling for multi-robots.

Input:
(1) Target positions $\mathscr{T}$
(2) The partition $\Lambda(D)$ and the weighted areas $\{\omega_i\}_{i=1}^{n}$
(3) The number of agents m
(4) The number of initial samples N_0
(5) The number of transitions N in each time period
(6) Stop time $t_{\max}$

Output: The estimated field $\hat{f}(x)$ at each time step

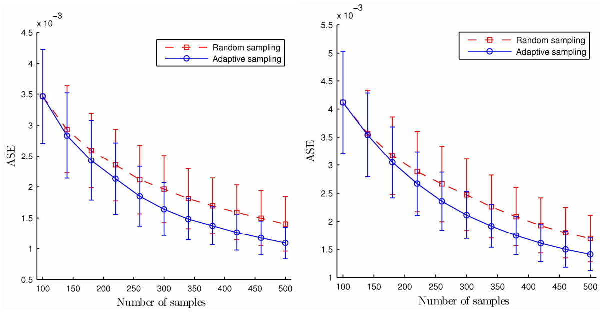
Mobile agents take N_0 initial samples at $t = 0$.
Estimate the Hessian matrix $\mathscr{H}_f(x)$ at $\mathscr{T}$.
Calculate the optimal probability distribution $p_i^\star$ using (6) and (7).
the region D into m subregions $\{S_j\}_{j=1}^{m}$ for m mobile sensors such that desired sampling probability over each subregion is equal to $1/m$ with some minimal quantization error.
for each subregion j **do**
    Calculate the equilibrium probability distribution $\pi^{(j)}$ using (8).
    Find the transition probability matrix $P^{(j)}$ using the FMRMC optimization in (9).
end for
repeat
    Estimate the density function $p(x)$ using a fixed bandwidth h_p.
    Calculate the optimal bandwidth h_{opt} using (3).
    Estimate the field $\hat{f}(x)$ at target positions $\{U^{(i)}\}_{i=1}^{n}$ using local linear regression in (1).
    for $(k = 1; k <= N; k++)$ **do**
        Update sampling positions of sensors according to the transition probability matrix $P^{(j)}$ and $\{\tilde{p}_i^\star(x) \mid i \in I_j\}$ in (10).
        Take new samples at new locations.
    end for
    Set $t = t + 1$.
until $t = t_{\max}$

to be 0.1 (about 10 % of the maximal value changes in the fields). We consider 10×10 grid points on D as the target positions. The region D is partitioned into n =100 cells D_i, each of which has a pre-selected value ω_i =1. At time $t = 0$, $N_0 = 100$ evenly distributed initial samples are collected by these m mobile sensors, which are used for estimating the Hessian matrix $\mathscr{H}_f(x)$. Then each robot collects N =10 samples during each time period based on the proposed algorithm in Table 2. The maximal range of the movement is set to be 3. The field D is partitioned into m rectangular subregions $\{S_j\}_{j=1}^{m}$ such that the desired sampling probability over each subregion is nearly equal to 1/4. A benchmark is created by a stochastic sampling scheme in which each of m robotic sensors uses a random walk to choose a cell to travel in the finite state space and the sampling position is uniformly distributed inside the cell. Each of m mobile sensors is launched at the one of the four vertex cells in D for both proposed and benchmark schemes. To have a fair comparision, the same parameters for the mobility constraints and the initial sampling positions are used for



(a) (b)

Fig. 2. The scalar fields (a) $f_1(x)$ and (b) $f_2(x)$ used in the simulation study.



(a) (b)

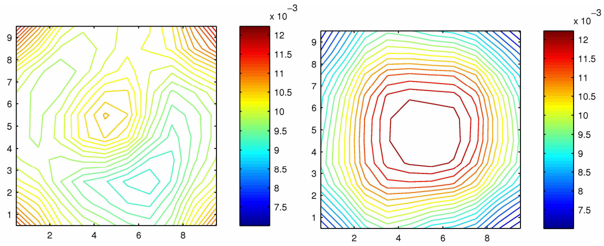
Fig. 3. Simulation results for the adaptive sampling scheme as compared to the random walk scheme for (a) $f_1(x)$ and (b) $f_2(x)$.

our scheme as well as the benchmark. To evaluate our scheme, the simulations of two schemes are repeated 100 times for both fields and the results of our scheme are compared against those of the benchmark in terms of the Average Squared Error (ASE):

$$ASE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{f}(U^{(i)}) - f(U^{(i)})\right)^2,$$

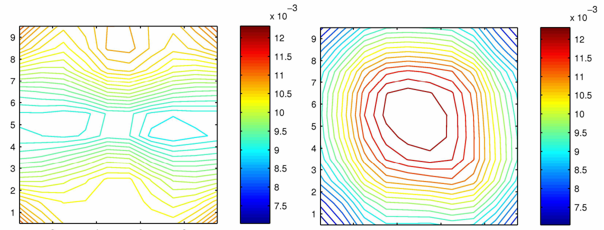
The sample means of the ASE (AMSE) for our adaptive sampling (blue circles) and the benchmark (red squares) along with the corresponding one standard deviation bars for $f_1(x)$ and $f_2(x)$ are shown in Figs. 3(a) and (b), respectively. By the nature of the kernel regression, the sample mean of ASE decreases as the number of samples increases for both schemes. It is important to notice that the proposed scheme decreases the AMSE with a faster rate than the random walk scheme at the beginning, which can be useful for the under-sampled data. The averages of the estimated densities of 500 samples for $f_1(x)$ and $f_2(x)$ over 100 trials are shown in Figs. 4 and 5, respectively. It can be clearly seen that the density is high at the part of the field where there are large changes in scalar values. The percentage improvements of the AMSE using the adaptive sampling scheme with respect to the random walk benchmark scheme are about 24 % and 20 % with 500 samples for $f_1(x)$ and $f_2(x)$, respectively.

Recall that our objective was to reduce the ASE at target positions $\{U^{(i)}\}_{i=1}^{n}$. However, the kernel regres-



(a) (b)

Fig. 4. The estimated density functions of $f_1(x)$ for (a) the random walk scheme and (b) the adaptive sampling scheme.

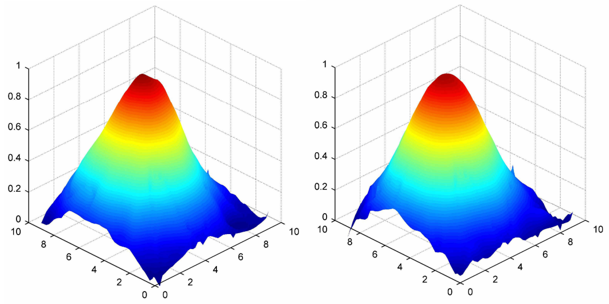


(a) (b)

Fig. 5. The estimated density functions of $f_2(x)$ for (a) the random walk scheme and (b) the adaptive sampling scheme.

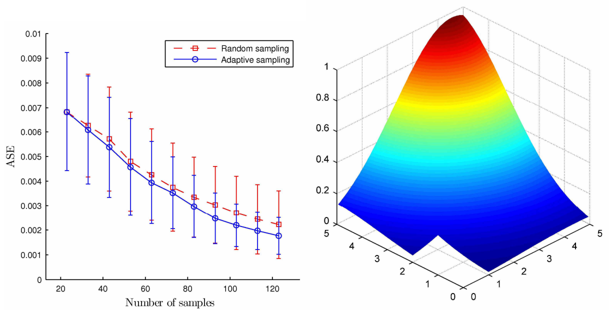
sion approach in (1) with the optimal kernel bandwidth in (3) can estimate the scalar field at any point in D, which is different from the pre-selected target positions. Hence, we estimate the field $f_1(x)$ at a set of 50×50 grid points over $(0.5, 9.5)^2$. The estimated fields $\hat{f}_1(x)$ with 500 samples using the random walk scheme and the adaptive sampling scheme are shown in Figs. 6(a) and (b), respectively. Although the objective was to minimize the estimation error at finite target positions, i.e., 10×10 grid points, the estimated scalar values at the new 50×50 grid points using adaptive sampling strategy is very similar to the true field as shown in Fig. 2(a). Notice also that the kernel regression technique does not produce good estimates on the points near the boundary in general.

To demonstrate that our scheme can be used in a non-convex region, we consider a scenario in which only a single mobile sensor (i.e., m =1) performs the estimation of a scalar field in a non-convex surveillance region as shown in Fig. 7(a). The analytical expression of the scalar field is the same as $f_1(x)$. The non-convex surveillance region D is given by $D = (0,5)^2 \setminus (0,1)\times(0,2)$, where \ denotes the set subtraction. The region D is partitioned into n = 23 cells D_i, each of which has a pre-selected value ω_i = 1. At time t = 0, N_0 = 23 evenly distributed initial samples are collected by the mobile sensors, which are used for estimating the Hessian matrix $\mathscr{H}_f(x)$. Then the robot collects N =10 samples during each time period based on the proposed algorithm. The maximal range of the movement is set to be 3. A benchmark is created by a stochastic sampling scheme in which a robotic sensor uses a random walk to choose a



(a) (b)

Fig. 6. The estimated field $\hat{f}_1(x)$ with 500 samples using (a) the random walk scheme and (b) the adaptive sampling scheme.



(a) (b)

Fig. 7. (a) The scalar field over a non-convex surveillance region used in the simulation study. (b) Simulation results for the adaptive sampling scheme as compared to the random walk scheme on a non-convex field.

cell to travel in the finite state space and finally selects a uniformly distributed sampling position over the cell. The sensor is launched randomly in D for both proposed and benchmark schemes.

The sample means of the ASE (AMSE) for our adaptive sampling (blue circles) and the benchmark (red squares) along with the corresponding one standard deviation bars are shown in Fig. 7(b). The percentage improvement of the AMSE using the adaptive sampling scheme with respect to the random walk benchmark scheme is about 18 % with 123 samples, which demonstrates that the proposed algorithm consistently performs better as compared to the benchmark for an unknown field in a non-convex surveillance region. Notice that in this case, the improvement by using our approach compared to the random strategy is smaller due to the slower convergence rate caused by using only one mobile sensor.

6.2. Real-life example

Consider a case in which a single mobile sensing robot, e.g., a robotic helicopter [22], measures the depth value of a rectangular terrain. The true depth values of the terrain on grid points are obtained by a Kinect sensor [23] as shown in Fig. 9(a). The depth sensor inside the Kinect consists of an infrared laser projector combined with a monochrome CMOS sensor, which captures video

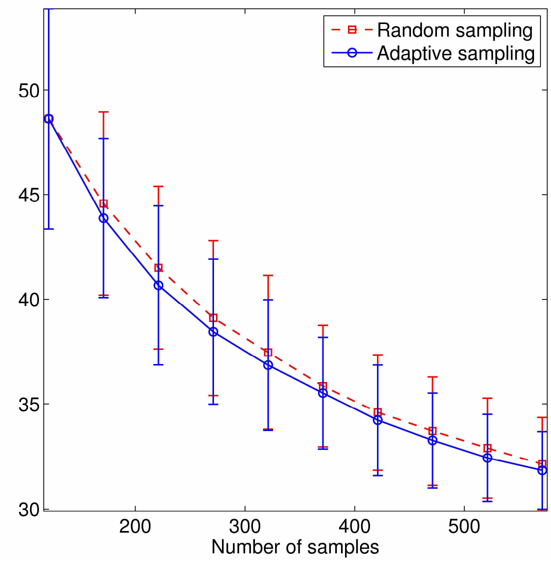
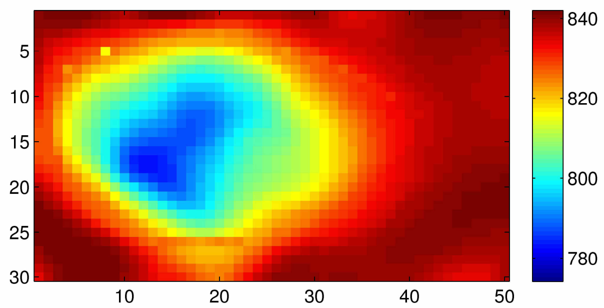
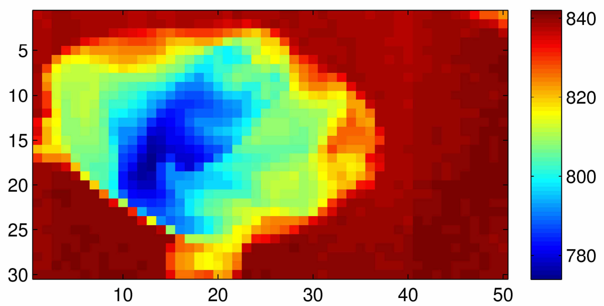


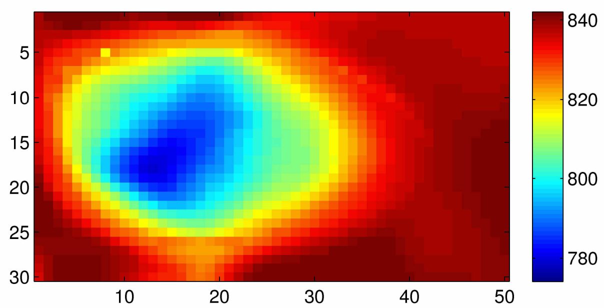
Fig. 8. Simulation results for the adaptive sampling scheme as compared to the random walk scheme on a true depth map.



(a)



(b)



(c)

Fig. 9. (a) The depth map of the terrain. The estimated field using (b) the random scheme, and (c) the stochastic adaptive sampling scheme.

data in 3D under any ambient light conditions. The data are obtained via the USB connection through an open source driver.

The surveillance region D is given by D = (1,50) ×(1,30). The measurement noise level is set to be 10. The region D is partitioned into n =1500 cells, each of which has a preselected value ω_i = 1. At time t = 0, N_0 = 121 evenly distributed initial samples are collected by the mobile sensor, which are used for estimating the Hessian matrix. Then the robot collects N = 50 samples during each time period based on the proposed algorithm and predicts the field at the next time instance. Similarly as in the previous examples, a random sampling scheme is used as a benchmark.

The experiment is repeated 100 times and the averaged mean square error for both scheme are shown in Fig. 8. It is clear that the purposed scheme yields a faster decreasing rate of the ASE. The estimated fields using the random scheme and the adaptive sampling scheme are shown in Figs. 9(b) and (c), respectively.

7. CONCLUSION

We have proposed a stochastic adaptive sampling scheme in which mobile sensors use kernel regression to estimate an unknown scalar field in a possibly large, non-convex surveillance region. The fastest mixing reversible Markov chain (FMRMC) was used to guarantee the fastest convergence of the sampling probability distribution to the optimal sampling probability distribution for the mobile sensor network with mobility constraints. In contrast to the schemes proposed in [3,6,8], the proposed sampling approach using the nonparametric kernel regression does not need to know *a priori* statistical knowledge of the unknown scalar field of interest. Simulation results showed the excellent performance of the proposed scheme as compared to a random walk benchmark on different scenarios. The future work shall focus on developing scalable adaptive sampling schemes for a class of resource-constrained mobile sensor networks to deal with a large number of measurements.

REFERENCES

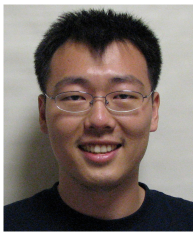
[1] I. Chorus and J. Bartram, *Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring, and Management*, Sponpress, 1999.

[2] S. Binder, "Deaths, injuries, and evacuations from acute hazardous materials releases," *Am. J. Public Health*, vol. 79, no. 8, pp. 1042-1044, August 1989.

[3] J. Choi, S. Oh, and R. Horowitz, "Distributed learning and cooperative control for multi-agent systems," *Automatica*, vol. 45, no. 12, pp. 2802-2814, 2009.

[4] E. Fiorelli, N. E. Leonard, P. Bhatta, D. A. Paley, R. Bachmayer, and D. M. Fratantoni, "Multi-AUV control and adaptive sampling in Monterey Bay," *IEEE Journal of Oceanic Engineering*, vol. 31, no. 4, pp. 935-948, 2006.

[5] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. E. Davis, "Collective motion, sensor networks, and ocean sampling," *Proc. of the IEEE*, vol. 95, no. 1, pp. 48-74, 2007.
[6] Y. Xu and J. Choi, "Adaptive sampling for learning Gaussian processes using mobile sensor networks," *Sensors*, vol. 11, no. 3, pp. 3051-3066, 2011.
[7] B. Zhang and G. S. Sukhatme, "Adaptive sampling for estimating a scalar field using a robotic boat and a sensor network," *Proc. of the IEEE International Conference on Robotics and Automation*, pp. 3673-3680, 2007.
[8] J. Choi, J. Lee, and S. Oh, "Biologically-inspired navigation strategies for swarm intelligence using spatial Gaussian processes," *Proc. of the International Federation of Automatic Control World Congress, Presented in the Invited Session: Sensing, Estimation, Consensus and Control Over Networks*, 2008.
[9] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829-836, 1979.
[10] D. Ruppert, S. J. Sheather, and M. P. Wand, "An effective bandwidth selector for local least squares regression," *Journal of the American Statistical Association*, vol. 90, pp. 1257-1270, 1995.
[11] J. Fan, "Design-adaptive nonparametric regression," *Journal of the American Statistical Association*, vol. 87, pp. 998-1004, 1992.
[12] J. Fan, "Local linear regression smoothers and their minimax efficiencies," *The Annals of Statistics*, vol. 21, pp. 196-216, 1993.
[13] D. Ruppert and M. P. Wand, "Multivariate locally weighted least squares regression," *The Annals of Statistics*, vol. 22, pp. 1346-1370, 1994.
[14] L. Yang and R. Tschernig, "Multivariate bandwidth selection for local linear regression," *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, vol. 61, no. 4, pp. 793-815, 1999.
[15] M. P. Wand and M. C. Jones, "Multivariate plugin bandwidth selection," *Computational Statistics*, vol. 9, no. 2, pp. 97-116, 1994.
[16] S. R. Sain and D. W. Scott, "On locally adaptive density estimation," *Journal of the American Statistical Association*, vol. 91, no, 436, pp. 1525-1534, 1996.
[17] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine Learning*, vol. 50, no. 1-2, pp. 5-43, 2003.
[18] J. Grace and J. Baillieul, "Stochastic strategies for autonomous robotic surveillance," *Proc. of the 44th IEEE Conference on Decision and Control, and European Control Conference*, pp. 2200-2205, 2005.
[19] S. Berman, A. Halasz, M. A. Hsieh, and V. Kumar, "Optimized stochastic policies for task allocation in swarms of robots," *IEEE Trans. on Robotics*, vol. 25, no. 4, pp. 927-937, 2009.
[20] S. Boyd, P. Diaconis, and L. Xiao, "Fastest mixing Markov chain on a graph," *SIAM Review*, vol. 46, no. 4, pp. 667-690, 2004.
[21] A. Kashyap, T. Başsar, and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192-1203, 2007.
[22] R. Miller, O. Amidi, and M. Delouis, "Arctic test flights of the CMU autonomous helicopter," *Proc. of the Association for Unmanned Vehicle Systems International*, 1999.
[23] Microsoft, Official website of Kinect for Xbox 360. http://www.xbox.com/en-US/kinect.



Yunfei Xu received his Ph.D. degree in Mechanical Engineering from Michigan State University in 2011. He also received his B.S. and M.S. degrees in Automotive Engineering from Tsinghua University, Beijing, China, in 2004 and 2007, respectively. Currently, he is a postdoctoral researcher in the Department of Mechanical Engineering at Michigan State University. His current research interests include environmental adaptive sampling algorithms, Gaussian processes and statistical learning algorithms with applications to robotics and mobile sensor networks.

Jongeun Choi received his Ph.D. and M.S. degrees in Mechanical Engineering from the University of California at Berkeley in 2006 and 2002, respectively. He also received a B.S. degree in Mechanical Design and Production Engineering from Yonsei University at Seoul, Republic of Korea in 1998. He is currently an Assistant Professor with the Departments of Mechanical Engineering and Electrical and Computer Engineering at the Michigan State University. His research interests include adaptive, distributed and robust control and statistical learning algorithms, with applications to self-organizing systems, mobile robotic sensors, environmental adaptive sampling, and biomedical problems. He was a recipient of an NSF CAREER Award in 2009. Dr. Choi is a member of IEEE and ASME.