Biologically-inspired Navigation Strategies for Swarm Intelligence using Spatial Gaussian Processes ⋆

Jongeun Choi * **Joonho Lee** ** **Songhwai Oh** ***

* *Department of Mechanical Engineering, Michigan State University, East Lansing, MI 48824-1226 USA (Tel: 517-432-3164; e-mail: jchoi@egr.msu.edu).*
** *Department of Mechanical Engineering, Michigan State University, East Lansing, MI 48824-1226 USA, (e-mail: leejoon8@msu.edu).*
*** *Electrical Engineering and Computer Science, School of Engineering, University of California, Merced, CA 95344-2039 USA (e-mail: songhwai.oh@ucmerced.edu)*

Abstract: This paper presents a novel class of self-organizing sensing agents that form a swarm and learn the static spatial process of interest through noisy measurements from neighbors for various global goals. The spatial phenomenon of interest is modeled by a Gaussian process. Each sensing agent maintains its own prediction of the Gaussian process based on measurements from neighbors. A set of biologically inspired navigation strategies are derived by exploiting the predictive posterior statistics. A unified way to prescribe a global goal for the group of agents so that a high-level behavior builds on a set of low-level simple behavior modules. As a result, collective mobility of agents emerges from a specified global goal. The proposed cooperatively learning control consists of motion coordination based on the recursive estimation of an unknown field of interest with measurement noise. The convergence properties of the proposed coordination algorithm for different situations and global goals are investigated by a simulation study.

Keywords: Multi-agent systems; Estimation and filtering; Sensor networks; Gaussian processes.

1. INTRODUCTION

In recent years, significant enhancements have been made in the areas of sensor networks and mobile sensing agents. Emerging technologies have been reported on coordination of mobile sensing agents (Cortes et al. (2004, 2005); Jadbabie et al. (2003); Tanner et al. (2003); Olfati-Saber (2006); Ren and Beard (2005, 2004); Choi et al. (2007); Martinez (2007); Cortes (2007); Graham and Cortes (2007)). Mobile sensing agents form an ad-hoc wireless communication network in which each agent usually operates under a short communication range, with limited memory and computational power. Mobile sensing agents are often spatially distributed in an uncertain surveillance environment. Among challenging problems of distributed coordination of mobile sensing agents, gradient climbing over an unknown field of interest has attracted much attention of control engineers. This has numerous applications including homeland security, toxic-chemical plume tracing and environmental monitoring. For instance, the most common approach to toxic-chemical plume tracing has been biologically inspired *chemotaxis* (Adler (1966); Dhariwal et al. (2004)), in which a mobile sensing agent is driven according to a local gradient of the chemical plume concentration. The cooperative network of agents that performs adaptive gradient climbing in a distributed environment was presented in Őgren et al. (2004); Leonardo and Robinson (2003). The centralized network can adapt its configuration in response to the sensed environment in order to optimize its gradient climb.

Many of the mobility of the mobile agents can be designed for a certain field of interest. Recently distributed interpolation schemes for field estimation by mobile sensor networks are developed by Martinez (2007). Gradient climbing swarming sensing agents for tracing the maximum of a noisy field via radial basis function learning are proposed and their convergence properties are analyzed by Choi et al. (2007). Our motivation is to design the mobility of sensing agents for various tasks by intelligently dealing with uncertainty in the prediction of a spatial phenomenon based on online measurements, and exploiting the predictive posterior statistics.

In our approach, the physical phenomenon in the surveillance region will be specified by a spatial Gaussian process. A Gaussian process (or Kriging in geostatistics) has been widely used as a nonlinear regression technique to estimate and predict geostatistical data (Cressie (1986, 1991); Gibbs and MacKay; MacKay (1998); Rasmussen and Williams (2006)). It is often used to predict spatial processes in meteorology, ecological systems, and environmental transport phenomena. A Gaussian process with an infinite number of random variables in a compact region $\mathcal{R}$ naturally generalizes a Gaussian distribution with finite number of random variables. A Gaussian process $z(s) \sim \mathcal{GP}(\mu(s), \mathcal{K}(s, s')),\ s, s' \in \mathcal{R}$ is specified by its mean function $\mu(\cdot)$ and a symmetric positive definite covariance function $\mathcal{K}(\cdot,\cdot)$. For Gaussian processes, the joint distribution of random variables in the subset $\mathcal{A} \subset \mathcal{R}$ is Gaussian. This property enables us to predict physical values, such as temperature and plume concentration, at any spatial point with predicted uncertainty level. Recently near-optimal static sensor placements with a mutual information criterion in Gaussian processes were proposed by Guestrin et al. (2005). Distributed Kriged Kalman filter for spatial estimation based on mobile sensor networks are developed by Cortes (2007). Asymptotic optimality of multicenter Voronoi configurations for random field estimation is reported by Graham and Cortes (2007).

⋆ This work was partially supported by Intramural Research Grants Program (IRGP) from Michigan State University.

The contribution of this paper is to develop a novel class of self-organizing multi-agent systems that sample measurements in spatially distributed manner to perform a given task by exploiting predictive posterior statistics from the recursive estimation of a physical process. The collective mobility of sensing agents are designed for various performance criterions depending on the tasks for the sensing agents. Inspired by biological behaviors such as "tracing food", "predator avoidance", and "environment exploration", a set of navigation strategies for swarming agents is precisely prescribed based on the recursively estimated Gaussian process. After identifying a set of strategies, we present a unified way to describe a global goal for agents. Hence a high-level behavior builds on a set of low-level simple behavior modules. The convergence properties of the proposed coordination algorithm for various situations and global goals are investigated by a simulation study. In this way, the distributed and scalable control law can be derived without knowledge of the field of interest in the environment. This is different from other coordination algorithms. The proposed cooperatively learning control consists of motion coordination based on the recursive estimation of parameters in a Gaussian process. Our strategy of cooperative learning control can be applied to a large class of coordination algorithms for mobile agents to deal with the field of interest that requires to be recursively estimated.

This paper is organized as follows. In Section 2, we briefly review the mobile sensing network model, notations related to a graph, and artificial potentials to form a swarming behavior. A recursive learning algorithm for estimating parameters and predicting a Gaussian process is presented in Section 3. Section 4 explains the biologically inspired navigation and a unified way to prescribe the global goal for agents. In Section 5, the resulting cooperatively learning control is described. In Section 6, we numerically test agents with different global performance criterions with respect to several configurations and different Gaussian processes.

2. PRELIMINARIES

In this section, we explain notations and concepts that will arise throughout the paper.

2.1 Mobile Sensing Agent Network

First, we explain the mobile sensing network and sensor models used in this paper. Let N_s be the number of sensing agents distributed over the surveillance region $\mathcal{R} \subset \mathbb{R}^2$. Assume that $\mathcal{R}$ is a compact set. The identity of each agent is indexed by $\mathcal{I} := \{1, 2, \cdots, N_s\}$. Let $q_i(t) \in \mathcal{R}$ be the location of the i-th sensing agent at time $t \in \mathbb{Z}_+$ and let $q := \mathrm{col}(q_1, q_2, \cdots, q_{N_s}) \in \mathbb{R}^{2N_s}$ be the configuration of the swarm system. The discrete time, high-level dynamics of agent i is modeled by

$$\begin{cases} q_i(t+1) = q_i(t) + \epsilon p_i(t), \\ p_i(t+1) = p_i(t) + \epsilon u_i(t) \end{cases}, \qquad (1)$$

where $q_i, p_i, u_i \in \mathbb{R}^2$ are, respectively, the position, the velocity, and the input of the mobile agent and ϵ is the iteration step size (or sampling time). We assume that the measurement $y(q_i(t), t)$ of sensor i includes the scalar value of the Gaussian process $z(q_i(t), t)$ and sensor noise $w(t)$, at its position $q_i(t)$ and some measurement time t,

$$y(q_i(t), t) := z(q_i(t), t) + w(t). \qquad (2)$$

2.2 A Graph

The group behavior of mobile sensing agents and their complicated interactions with neighbors are best treated by a graph with edges. Let $G(q) := (\mathcal{I}, \mathcal{E}(q))$ be a communication graph such that an edge $(i, j) \in \mathcal{E}(q)$ if and only if agent i can communicate with agent $j \neq i$. We assume that each agent can communicate with its neighboring agents within a limited transmission range given by a radius of r. Therefore, $(i, j) \in \mathcal{E}(q)$ if and only if $\|q_i(t) - q_j(t)\| \leq r$. We define the neighborhood of agent i with a configuration of q by $N(i, q) := \{j : (i, j) \in \mathcal{E}(q),\ i \in \mathcal{I}\}$. The adjacency matrix $A := [a_{ij}]$ of an undirected graph G is a symmetric matrix such that $a_{ij} = k_3 > 0$ if vertex i and vertex j are neighbors and $a_{ij} = 0$ otherwise, where k_3 is a positive scalar. The scalar graph Laplacian $L = [l_{ij}] \in \mathbb{R}^{N_s \times N_s}$ is a matrix defined as $L := D(A) - A$, where $D(A)$ is a diagonal matrix whose diagonal entries are row sums of A, i.e., $D(A) := \mathrm{diag}(\sum_{j=1}^{N_s} a_{ij})$. The 2-dimensional graph Laplacian is defined as $\hat{L} := L \otimes I_2$, where $\otimes$ is the Kronecker product. A quadratic disagreement function (Olfati-Saber (2006)) can be obtained via the Laplacian $\hat{L}$:

$$\Psi_G(p) := p^T \hat{L} p = \frac{1}{2} \sum_{(i,j) \in \mathcal{E}(q)} a_{ij} \|p_j - p_i\|^2, \qquad (3)$$

where $p := \mathrm{col}(p_1, p_2, \cdots, p_{N_s}) \in \mathbb{R}^{2N_s}$.

2.3 A Swarming Behavior

In order for agents to sample measurements of a scalar field at spatially distributed locations simultaneously, a group of mobile agents will be coordinated by a flocking algorithm (Olfati-Saber (2006); Tanner et al. (2003); Choi et al. (2007)). We use attractive and repulsive smooth potentials similar to those used in Tanner et al. (2003); Olfati-Saber (2006); Choi et al. (2007) to generate a swarming behavior. To enforce a group of agents to satisfy a set of algebraic constraints $\|q_i - q_j\| = d$ for all $j \in \mathcal{N}(i, q)$, we introduce a collective potential

$$U_1(q) := \sum_i \sum_{j \neq i} U_{ij}(\|q_i - q_j\|^2) = \sum_i \sum_{j \neq i} U_{ij}(r_{ij}), \qquad (4)$$

where $r_{ij} := \|q_i - q_j\|^2$. U_{ij} in (4) is defined by

$$U_{ij}(r_{ij}) := \frac{1}{2} \left(\log(\alpha + r_{ij}) + \frac{\alpha + d^2}{\alpha + r_{ij}} \right), \text{ if } r_{ij} < d_0^2, \qquad (5)$$

otherwise (i.e., $r_{ij} \geq d_0^2$), it is defined according to the gradient of the potential, which will be described shortly. Here $\alpha, d \in \mathbb{R}_+$ and $d < d_0$. The gradient of the potential w.r.t. q_i for agent i is given by

$$\begin{aligned} \nabla U_1(q_i) &:= \frac{\partial U_1(q)}{\partial \tilde{q}_i}\Big|_{\tilde{q}_i = q_i} = \sum_{j \neq i} \frac{\partial U_{ij}(r)}{\partial r}\Big|_{r = r_{ij}} (q_i - q_j) \\ &= \begin{cases} \sum_{j \neq i} \frac{(r_{ij} - d^2)(q_i - q_j)}{(\alpha + r_{ij})^2} & \text{if } r_{ij} < d_0^2 \\ \sum_{j \neq i} \rho \left(\frac{\sqrt{r_{ij}} - d_0}{|d_1 - d_0|} \right) \frac{\|d_0^2 - d^2\|}{(\alpha + d_0^2)^2} (q_i - q_j) & \text{otherwise,} \end{cases} \end{aligned} \qquad (6)$$

where $\rho : \mathbb{R}_+ \to [0, 1]$ is the bump function

$$\rho(z) := \begin{cases} 1, & z \in [0, h); \\ \frac{1}{2}\left[1 + \cos\left(\pi \frac{(z-h)}{(1-h)}\right)\right], & z \in [h, 1]; \\ 0, & \text{otherwise,} \end{cases}$$

that smoothly varies from 1 to 0 as the scalar input increases. In equations (4), (5), and (6), α was introduced to prevent the reaction force from diverging at $r_{ij} = \|q_i - q_j\|^2 = 0$. This potential yields a reaction force that is attracting when the agents are too far and repelling when a pair of two agents are too close. It has an equilibrium point at a distance of d. We also introduce a potential U_2 to model the environment. U_2 enforces each agent to stay inside the closed and connected surveillance region $\mathcal{R}$ and prevents collisions with obstacles in $\mathcal{R}$. We construct U_2 such that it is radially unbounded in q, i.e.,

$$U_2(q) \to \infty \text{ as } \|q\| \to \infty. \tag{7}$$

Define the total artificial potential by

$$U(q) := k_1 U_1(q) + k_2 U_2(q), \tag{8}$$

where $k_1 > 0$ and $k_2 > 0$ are weighting factors.

3. LEARNING AGENTS FOR GAUSSIAN PROCESSES

In this section, we introduce a recursive learning algorithm for each mobile sensing agent to estimate the Gaussian process with a nonzero mean $z(s,t) := \mu(s) + \varsigma(s,t) \sim \mathcal{GP}(\mu(s), \mathcal{K}(s,t;s',t'))$ where $s(t), s'(t') \in \mathcal{R}$ and $t, t' \in \mathbb{Z}_+$. (see Rasmussen and Williams (2006)). Sensors sample noisy measurements of the process $z(s,t), s \in \mathcal{R}$ (see Cressie and Wikle (2002); Wikle and Cressie (1999)),

$$y(s,t) = z(s,t) + w(t), \tag{9}$$

where the sensor noise is Gaussian white noise $w \sim \mathcal{N}(0, \sigma_w^2)$. The mean field $\mu(\cdot)$ is a linear function of an unknown parameter vector θ:

$$\mu(\nu) := \sum_{j=1}^{m} \phi_j(\nu)\theta_j = \phi^T(\nu)\theta, \tag{10}$$

where $\phi^T(\nu)$ and θ are respectively by

$$\phi^T(\nu) := [\, \phi_1(\nu) \ \phi_2(\nu) \ \cdots \ \phi_m(\nu) \,],$$
$$\theta := [\, \theta_1 \ \theta_2 \ \cdots \ \theta_m \,]^T.$$

$\phi_j(\nu)$ are smooth Gaussian kernels given by

$$\phi_j(s) := \frac{1}{Z}\exp\left(\frac{-\|s - \nu_j^c\|^2}{\sigma_\mu^2}\right), \tag{11}$$

where σ_μ is the width of the Gaussian basis and Z is a normalizing constant. ν_j^c for $j \in \{1, \cdots, m\}$ are uniformly distributed in the surveillance region $\mathcal{R}$. The zero mean Gaussian process $\varsigma(s,t)$ is given by $\varsigma(s,t) \sim \mathcal{GP}(0, \mathcal{K}(s,t;s',t'))$ with a covariance matrix $\mathcal{K}(s_i,t_i;s_j,t_j) := \kappa(s_i,s_j)\delta_{(t_i,t_j)}$, where $\delta_{(\cdot,\cdot)}$ is the Kronecker delta and

$$\kappa(s_i,s_j) := \frac{1}{Z_\kappa}\exp\left(\frac{-\|s_i - s_j\|^2}{\sigma_\kappa^2}\right). \tag{12}$$

We also specify *a prior* over θ by: $\theta \sim \mathcal{N}(\theta_0, \Sigma_\theta(0))$. Suppose that at iteration time $t \in \mathcal{T} := \{0,1,2,\cdots\}$, agent i can collect observations $Y(t) := [y(s_1(t),t), \cdots, y(s_n(t),t)]^T$ taken at the n sites $\{s_1(t), \cdots, s_n(t)\}$ by itself and its $n-1$ number of neighbors, then we have:

$$Y(t) = \Phi(t)\theta + v(t) \in \mathbb{R}^n, \tag{13}$$

where $\Phi(t) := [\phi(s_1(t)), \cdots, \phi(s_n(t))]^T$, $\text{rank}\Phi(t) = n$ [1] and $v(t) \sim \mathcal{N}(0, \Sigma_v(t))$. We assume that agent i can compute the covariance function by

$$\Sigma_v(t) := [\kappa(s_i,s_j)] + \text{diag}(\sigma_{w1}^2, \cdots, \sigma_{wn}^2) \in \mathbb{R}^{n\times n}, \tag{14}$$

where σ_{wj}^2 are bounded and are due to the sum of sensor noise and communication noise between agent i and neighboring agents.

The standard unbiased minimum mean square error (MMSE) estimation (Kay (1993)) can be derived in a recursive fashion. Starting from $\hat{\theta}(0) = \theta_0$ and $\Sigma_{\tilde{\theta}}(0) := \Sigma_\theta(0)$, we have:

$$\begin{aligned} K_f(t) &= \Sigma_{\tilde{\theta}}(t-1)\Phi^T(t)\left[\Phi(t)\Sigma_{\tilde{\theta}}(t-1)\Phi^T(t) + \Sigma_v(t)\right]^{-1}, \\ \hat{\theta}(t) &= \hat{\theta}(t-1) + K_f(t)\left[Y(t) - \Phi(t)\hat{\theta}(t-1)\right], \\ \Sigma_{\tilde{\theta}}(t) &= [I_m - K_f(t)\Phi(t)]\,\Sigma_{\tilde{\theta}}(t-1), \end{aligned} \tag{15}$$

[1] There exists a measure zero $s \in \mathcal{R}^n$ that makes $\text{rank}\Phi(t) < n$.

where $\tilde{\theta}(t) := \hat{\theta}(t) - \theta$ and $\Sigma_{\tilde{\theta}}(t) := \mathbb{E}\left[\tilde{\theta}(t)\tilde{\theta}^T(t)\right]$ denote the estimation error vector and the error covariance matrix respectively. $I_m, 0_m \in \mathbb{R}^{n\times n}$ are the identity and zero matrices respectively. For a fixed θ, we have the following :

$$\begin{aligned} \Sigma_z(t) &:= \text{Cov}(z(s,t), z(s,t)|\theta) = \kappa(s,s), \\ \Sigma_Y(t) &:= \text{Cov}(Y(t), Y(t)|\theta) = \Sigma_v(t), \\ \Sigma_{Yz}(t) &= \Sigma_{zY}^T(t) := \text{Cov}(Y(t), z(s,t)|\theta) = \psi(s), \end{aligned} \tag{16}$$

where $\text{Cov}(x,y) := \mathbb{E}(x - \mathbb{E}x)(y - \mathbb{E}y)^T$ and $\psi(s) := [\kappa(s_i,s)] \in \mathbb{R}^n$. We also have $\theta|Y_{\le t} \sim \mathcal{N}(\hat{\theta}(t), \Sigma_{\tilde{\theta}}(t))$, where $Y_{\le t} := \{Y(t), \cdots, Y(1)\}$. The posterior predictive distribution of $z(s,t)$ conditioned on $Y_{\le t}$ can be obtained by marginalizing $p(z(s,t)|\theta, Y_{\le t})$ over $p(\theta|Y_{\le t})$:

$$z(s,t|t) := z(s,t)\,|\,Y_{\le t} \sim \mathcal{N}\left(\hat{z}(s,t|t), \sigma^2(s,t|t)\right), \tag{17}$$

where $\hat{z}(s,t|t) := \mathbb{E}\{z(s,t|t)\}$ is:

$$\begin{aligned} \hat{z}(s,t|t) &:= \phi^T(s)\hat{\theta}(t) + \Sigma_{zY}(t)\Sigma_Y^{-1}(t)\left(Y(t) - \Phi(t)\hat{\theta}(t)\right), \\ &= \phi^T(s)\hat{\theta}(t) + \psi^T(s)\Sigma_v^{-1}(t)\left(Y(t) - \Phi(t)\hat{\theta}(t)\right), \end{aligned}$$

and $\sigma^2(s,t|t)$ is given by

$$\begin{aligned} &\sigma^2(s,t|t) := \Sigma_z(t) - \Sigma_{zY}(t)\Sigma_Y^{-1}(t)\Sigma_{zY}^T(t) \\ &+ (\phi^T(s) - \Sigma_{zY}\Sigma_Y^{-1}\Phi(t))\Sigma_{\tilde{\theta}}(t)(\phi^T(s) - \Sigma_{zY}\Sigma_Y^{-1}\Phi(t))^T \\ &= \kappa(s,s) - \psi^T(s)\Sigma_v^{-1}(t)\psi(s) \\ &+ [\phi^T(s) - \psi^T(s)\Sigma_v^{-1}\Phi(t)]\Sigma_{\tilde{\theta}}(t)[\phi^T(s) - \psi^T(s)\Sigma_v^{-1}\Phi(t)]^T. \end{aligned}$$

The last term is due to using the MMSE estimate $\hat{\theta}(t)$ as compared to applying a simple kriging or a prediction of the Gaussian process for a known θ. This formulation is a popular way to embed a finite number of deterministic kernels to represent a mean trend (See Wikle and Cressie (1999); Blight and Ott (1975); Cressie and Wikle (2002); Rasmussen and Williams (2006); Cortes (2007)). This algorithm combines parametric and nonparametric estimations, which is robust w.r.t. possible mismatches in the selected radial basis functions that parameterize the mean trend. In the following section, navigation strategies based on the spatial prediction and the estimated uncertainty in (17) are presented.

4. NAVIGATION STRATEGIES

The collective mobility of the group of sensing agents shall be designed according to an appropriate performance criterion based on current estimation of the Gaussian process in (17). To rapidly trace a plume source, swarming sensing agents can climb the gradient of the estimated plume field (see Choi et al. (2007)). For estimation of Gaussian processes, sensors can be placed at the highest entropy location of the Gaussian process, which will give us the maximal information gain after sensing at the location (see Cressie (1991)). For instance, swarming mobile agents can sample more measurements at places where the estimation error variance is large. Guestrin et al. (2005) proposed to place static sensors sequentially according to the mutual information criteria for Gaussian processes.

4.1 Biologically Inspired Navigation

Depending on the tasks for the sensing agents, the collective mobility of sensing agents is designed to maximize the specified performance criterion. We propose a set of useful, biologically inspired navigation modes (tracing, avoidance, and exploration): (i) for tracing, the agents can climb the gradient of the estimated field:

Table 1. A list of common goals and their performance related smooth objective functions.

Goals	Smooth objective functions
Avoidance (Negative prediction)	$\beta_0 = -\hat{z}(s,t\|t)$
Tracing (Prediction)	$\beta_1 = \hat{z}(s,t\|t)$
Exploration (Variance)	$\beta_2 = \sigma^2(s,t\|t)$
Exploration (Entropy)	$\beta_3 = H(z(s,t\|t))$

$$\begin{aligned}\frac{d}{dt}q_i(t) &= \nabla_s(\hat{z}(s,t|t))\big|_{s=q_i(t)} \\ &= {\phi'}^T(q_i(t))\hat{\theta}(t) \\ &+ {\psi'}^T(q_i(t))\Sigma_v^{-1}(t)(Y(t)-\Phi(t)\hat{\theta}(t)) : \text{ Tracing;}\end{aligned} \tag{18}$$

where ${\phi'}^T(q_i) := \nabla_s\phi(s)|_{s=q_i}, {\psi'}^T(q_i) := \nabla_s\psi(s)|_{s=q_i} \in \mathbb{R}^{2\times m}$, (ii) for avoidance, we also have:

$$\frac{d}{dt}q_i(t) = -\nabla_s(\hat{z}(s,t|t))\big|_{s=q_i(t)} : \text{ Avoidance;} \tag{19}$$

(iii) for exploration, the agents can climb the gradient of the estimated variance

$$\begin{aligned}&\frac{d}{dt}q_i(t) = \nabla_s(\sigma^2(s,t|t))\big|_{s=q_i(t)} \\ &= -2{\psi'}^T(q_i)\Sigma_v^{-1}(t)\psi(q_i) + 2[{\phi'}^T(q_i) - {\psi'}^T(q_i)\Sigma_v^{-1}\Phi(t)] \\ &\Sigma_{\tilde{\theta}}(t)[\phi^T(q_i) - \psi^T(q_i)\Sigma_v^{-1}\Phi(t)]^T : \\ &\text{Exploration (variance);}\end{aligned} \tag{20}$$

or differential entropy of the Gaussian process:

$$\begin{aligned}\frac{d}{dt}q_i(t) &= \nabla_s H(z(s,t|t))\big|_{s=q_i(t)} \\ &= \nabla_s\left(\frac{1}{2}\ln(2\pi e\sigma^2(s,t|t))\right)\Bigg|_{s=q_i(t)} \\ &= \frac{\nabla_s(\sigma^2(s,t|t))\big|_{s=q_i(t)}}{2\sigma^2(q_i(t),t|t)} : \text{ Exploration (entropy).}\end{aligned} \tag{21}$$

By using (20) and (21), we expect the variance and the entropy of the Gaussian process in the surveillance region $\mathcal{R}$ to decrease.

Notice that prediction in (17) and gradients in (18), (19), (20) and (21) are smooth functions of a location s, which ensures the existence of extreme values over a compact set. Complicated protocols can be designed, for instance, when an agent obtains an estimate within a specified error tolerance, it can make a decision to locate the maximum of the estimated field of interest (Choi et al. (2008)). In Table 1, a list of common goals and their performance related smooth objective functions are summarized.

4.2 The Parameterization of a Global Goal

The optimal balance between exploitation and exploration is commonly observed in biological searchers (Grünbaum (1998); Vergassola et al. (2007)) as well as in learning theory. The balance can be achieved as the convex combination of different objective functions. For instance, a global objective function can be parameterized in the following way:

$$J_i(\Lambda(t);s,t) := \frac{\sum_{k=1}^3 \lambda_{ik}(t)\beta_{ik}(s,t)}{\sum_{k=1}^3 \lambda_{ik}(t)}, \text{ for all } i\in\mathcal{I}, \tag{22}$$

where $\beta_{i1}(s,t) := \hat{z}_i(s,t|t)$, $\beta_{i2}(s,t) := \sigma_i^2(s,t|t)$ and $\beta_{i3}(s,t) := 0.1H(z_i(s,t|t))$ are specifically chosen for the later simulation study. The global objective function is a function of each agent's navigation strategy $\Lambda(t) := [\lambda_{ik}(t)] \in \mathbb{R}_+^{|\mathcal{I}|\times 3}$ that sets the individual weights on all possible objective functions (typical ones are shown in Table 1).

5. COOPERATIVELY LEARNING CONTROL

Each of the mobile agents receives measurements from neighbors, then updates its estimation of the Gaussian process in $\mathcal{R}$ via the recursive algorithm presented in (15) and an update in (17). Subsequently, based on the update of a gradient of an objective function in (22), the control for its coordination will be decided. The update of Gaussian process of agent i at its position $q_i(t)$ is given by:

$$\begin{aligned}K_{fi}(t) &= \Sigma_{\tilde{\theta}i}(t-1)\Phi_i^T(t)\left[\Phi_i(t)\Sigma_{\tilde{\theta}i}(t-1)\Phi_i^T(t)+\Sigma_{vi}(t)\right]^{-1}, \\ \hat{\theta}_i(t) &= \hat{\theta}_i(t-1) + K_{fi}(t)\left[Y_i(t) - \Phi_i(t)\hat{\theta}_i(t-1)\right], \\ \Sigma_{\tilde{\theta}i}(t) &= [I_m - K_{fi}(t)\Phi_i(t)]\Sigma_{\tilde{\theta}i}(t-1),\end{aligned} \tag{23}$$

where $Y_i(t)$ is the collection of collaboratively measured data at iteration time t. Based on the gradient of the performance function $\nabla J_i(\cdot;\cdot,\cdot)$ in (22) updated by (23), a distributed control for agent i is decided by

$$\begin{aligned}u_i(t) := \Big\{ &-\nabla U(q_i(t)) - k_{di}p_i(t) \\ &+ \sum_{j\in N(i,q(t))} a_{ij}(q(t))(p_j(t)-p_i(t)) + k_4\nabla J_i(\Lambda(t);q_i(t),t)\Big\},\end{aligned} \tag{24}$$

where $k_4 \in \mathbb{R}_+$ is a gain for the gradient of a global objective function and $k_{di} \in \mathbb{R}_+$ is a gain for the velocity feedback. The first term in (24) is the gradient of (8) which attracts agents while avoiding collisions among them. Also it restricts the movements of agents inside $\mathcal{R}$. Appropriate artificial potentials can be added to $U(q_i)$ for agents to avoid obstacles in $\mathcal{R}$. The third term in (24) is an effort for agent i to match its velocity with those of neighbors. This term is also called a "velocity consensus" and serves as a damping force among agents. Incorporating the closed-loop discrete time model in (1) along with the proposed control in (24) gives

$$\begin{aligned}q_i(t+1) &= q_i(t) + \epsilon p_i(t) \\ p_i(t+1) &= p_i(t) + \epsilon\Big\{ -\nabla U(q_i(t)) - k_{di}p_i(t) \\ &\quad - \nabla\Psi_G(p_i(t)) + k_4\nabla J_i(\Lambda(t);q_i(t),t)\Big\},\end{aligned} \tag{25}$$

where the iteration step or the sampling time ϵ is sufficiently small so that the trajectories of states may be approximated by the associated ODE (see more details in Kushner and Yin (1997)). If we consider p_i in (25) as an input to the single integrator dynamics, we can use the projection algorithm to ensure that the states remain inside a compact set.

6. SIMULATION RESULTS

To demonstrate our learning agents, we applied the control (24) to a stationary Gaussian process under various global goals generated by (22). The estimate of the unknown density was updated once per iteration. Nine agents were launched with the equilibrium distance $d = 0.8$. For the simulation study, we used $\sigma_\kappa = 0.8$ and $Z_\kappa = 2$ for the width and the normalization constant of the covariance function in (14). The recursive estimation in (23) starts with $\theta_0 = 0$. Hereafter plots contain updated parameters of agent 1 only along with trajectories of all agents.

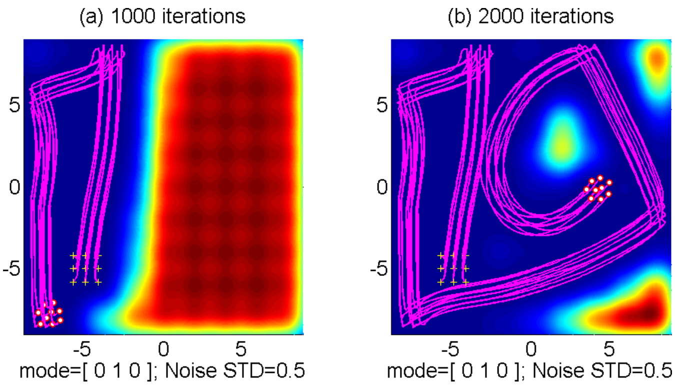

Fig. 1. The plot of $\sigma^2(s,t|t)$ updated by agent 1 (blue-lowest, red-highest). Variance driven exploration with the sensor noise $\sigma_w = 0.5$.

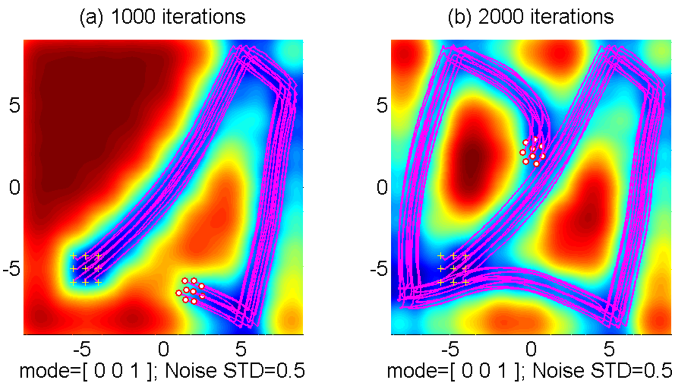

Fig. 2. The plot of $H(z(s,t|t))$. Entropy driven exploration with the sensor noise $\sigma_w = 0.5$.

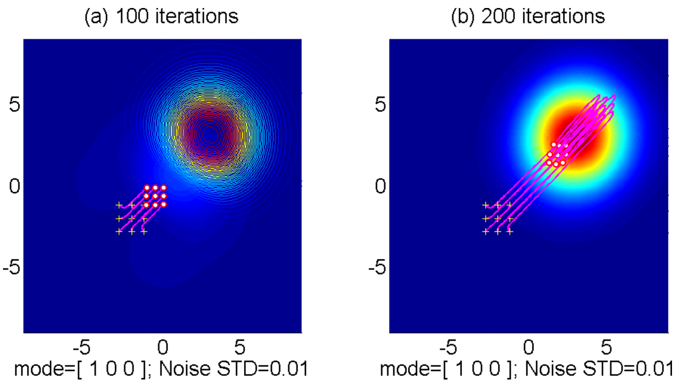

Fig. 3. The plot of $\hat{z}(s,t|t)$. Exploitation (tracing) under a small sensor noise level $\sigma_w = 0.01$. The error field between $\hat{z}(s,t|t)$ and $z(s,t)$ is plotted by colored contours.

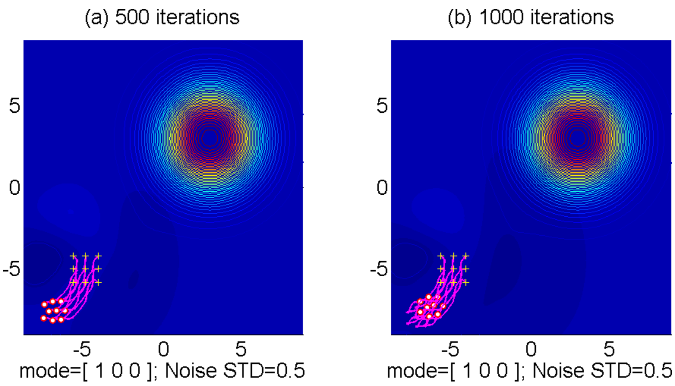

Fig. 4. The plot of $\hat{z}(s,t|t)$. Failed exploitation (tracing) under a larger sensor noise level $\sigma_w = 0.5$. The error field is plotted by colored contours.

6.1 Exploration

Group exploration or surveillance can be performed by having large weights on either $\beta_{i2}(\cdot,\cdot)$ or $\beta_{i3}(\cdot,\cdot)$ in (22). We have used the gradient that achieves the maximum norm of the gradient for objective functions evaluated around the location of each agent. Fig. 1 shows $\sigma^2(s,t|t)$ of agent 1 with the sensor noise $\sigma_w = 0.5$ after some iterations. In this case, the group navigation strategy is chosen as the variance driven exploration described by $\Lambda(t) = [0\,1\,0] \otimes 1_{|\mathcal{I}|} \in \mathbb{R}^{|\mathcal{I}|\times 3}$, where $1_n := [1 \cdots 1]^T \in \mathbb{R}^{n\times 1}$. In the same way, the entropy driven exploration strategy ($\Lambda(t) = [0\,0\,1] \otimes 1_{|\mathcal{I}|}$) is shown with $H(z(s,t|t))$ in Fig. 2. As shown in Figs. 1 and 2, agents with the exploratory group behavior tend to visit unexplored regions. Variants of such exploratory strategies can be straightforwardly developed in a distributed and scalable fashion since each agent can easily find the location of maximal entropy or variance in the surveillance region $\mathcal{R}$ (see Choi et al. (2008)). Different path planning strategies can be used for reducing the uncertainty in the prediction of the Gaussian process.

6.2 Tradeoff between exploitation and exploration

To see the existence of optimal balance between exploitation and exploration, we consider a problem of searching for the maximum of the spatial Gaussian process in (9). We study the following three situations: (A) If the sensor noise level is small and the initial positions of agents are close enough to estimate the gradients of fields, agents utilizing only the tracing strategy ($\Lambda(t) = [1\,0\,0] \otimes 1_{|\mathcal{I}|}$) quickly find the maximum of the field successfully as shown in Fig. 3. The colored, dotted-lines in Fig. 3-(a) represent the error between the estimated field and that of the true field. (B) If the sensor noise level is large and the initial positions are relatively far away from the maximum point, agents with the same tracing strategy only ($\Lambda(t) = [1\,0\,0] \otimes 1_{|\mathcal{I}|}$) can not find the maximum even after 1000 iterations as shown in Fig. 4. This is because that values of gradients at locations far from the maximum are very small and the estimates of such gradients are noisy. (C) Under the same harsh configuration as above, agents utilize a strategy that combines tracing and exploration modes ($\Lambda(t) = [1\,0\,2] \otimes 1_{|\mathcal{I}|}$). These agents start exploring the surveillance region and find the maximum after about 700 iterations as depicted in Fig. 5. In this case, the RMS error values over the finite number of spatially and uniformly sampled estimates w.r.t iterations is plotted in Fig. 6. The observation we made clearly indicates that there exists optimal balance among biologically inspired navigation modes. Therefore, learning capability of such balance $\Lambda(t)$ for a given situation is important, which can be dealt by another adaptation mechanism.

7. CONCLUSIONS

This paper presented a novel class of self-organizing autonomous sensing agents that form a swarm and learn through noisy cooperative measurements from neighbors for various global goals. The learning mechanism is based on a spatial Gaussian process. A set of biologically inspired navigation strategies are devised by exploiting the predictive posterior statistics. A unified way to prescribe a global goal for the group of agents so that a high-level behavior builds on a set of low-level simple behavior modules. As a result, collective mobility of agents emerges from the specified global goal. The proposed

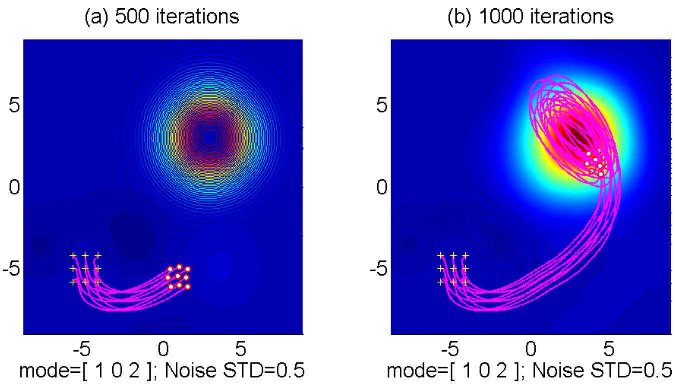

Fig. 5. The plot of $\hat{z}(s,t|t)$. Mixed exploitation (tracing) and entropy driven exploration under a larger sensor noise level $\sigma_w = 0.5$ The error field is plotted by colored contours.

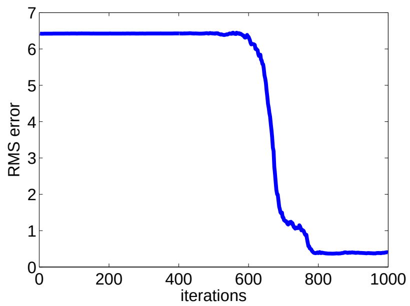

Fig. 6. The plot of RMS error v.s. iterations.

cooperatively learning control consists of motion coordination based on the recursive estimation of an unknown field of interest with measurement noise. Our strategy of the cooperative learning control can be applied to a large class of coordination algorithms for mobile agents in a situation where the field of interest is not known a priori and is to be estimated.

REFERENCES

Julius Adler. Chemotaxis in bacteria. *Journal of Supramolecular Structure*, 4(3):305–317, 1966.

B. J. N. Blight and L. Ott. A bayesian approach to model inadequacy for polynomial regression. *Biometrika*, 19(3): 79–88, 1975.

Jongeun Choi, Joonho Lee, and Songhwai Oh. Swarm Intelligence for Achieving the Global Maximum using Spatio-Temporal Gaussian Processes. *to appear in the 2008 American Control Conference (ACC)*, 2008.

Jongeun Choi, Songhwai Oh, and Roberto Horowitz. Cooperatively learning mobile agents for gradient climbing. In *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007.

J. Cortes, S. Martinez, and F. Bullo. Spatially-distributed coverage optimization and control with limited-range interactions. *ESAIM. Control, Optimisation and Calculus of Variations*, 11:691–719, 2005.

J. Cortes, S. Martinez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*, 20(2):243–255, 2004.

Jorge Cortes. Distributed kriged kalman filter for epatial estimation. *submitted to IEEE Transactions on Automatic Control*, May 2007.

Noel Cressie. Kriging nonstationary data. *Journal of the American Statistical Association*, 81(395):625–634, September 1986.

Noel Cressie and Christopher K. Wikle. *Encyclopedia of Environments*, volume 4, chapter Space-time Kalman filter, pages 2045–2049. John Wiley and Sons, Ltd, Chichester, 2002.

Noel A. C. Cressie. *Statistics for Spatial Data*. A Wiley-Interscience Publication, John Wiley and Sons, Inc., 1991.

Amit Dhariwal, Gaurav S. Sukhatme, and Aristides A. G. Requicha. Bacterium-inspired robots for environmental monitoring. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2004.

M. N. Gibbs and D. J. C. MacKay. Efficient implementation of Gaussian processes for interpolation.

R. Graham and Jorge Cortes. Asymptotic optimality of multicenter voronoi configurations for random field estimation. *submitted to IEEE Transactions on Automatic Control*, August 2007.

Daniel Grünbaum. Schooling as a strategy for taxis in a noisy environment. *Evolutionary Ecology*, 12:503–522, 1998.

Carlos Guestrin, Andreas Krause, and Ajit Paul Singh. Near-optimal sensor placement in gaussian processes. In *Proceeding of the 22nd International Conference on Machine Learning*, 2005.

Petter Őgren, Edward Fiorelli, and Naomi Ehrich Leonard. Cooperative control of mobile sensor networks: Adaptive gradient climing in a distributed envrironment. *IEEE Transaction on Automatic Control*, 49(8):1292, August 2004.

A. Jadbabie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomuous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, June 2003.

Steven M. Kay. *Fundamentals of Statistical Signal Processing : Estimation Theory*, volume 1 of *Signal Processing Series*. Prentice Hall, Inc., Upper Saddle River, New Jersey, 1993.

Harold J. Kushner and G. George Yin. *Stochastic Approximation Algorithms and Applications*. Springer, 1997.

N. Leonardo and A. Robinson. Adaptive sampling and forecasting plan AOSN-II MB'03 project. Technical report, Available: http://www.princeton.edu/ dcsl/aosn/documents/ASFP.pdf, 2003.

D. J. C. MacKay. Introduction to Gaussian processes. In C. M. Bishop, editor, *Neural Networks and Machine Learning*, NATO ASI Series, pages 133–166. Kluwer, 1998.

Sonia Martinez. Distributed interpolation schemes for field estimation by mobile sensor networks. *submitted to IEEE Transactions on Control Systems Technology*, July 2007.

Olfati-Saber. Flocking for multi-agent dynamic systems: Algorithm and theory. *IEEE Transactions on Automatic Control*, 51(3):401–420, March 2006.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, Massachusetts, London, England, 2006.

Wei Ren and Randal W. Beard. Trajectory tracking for unmanned air vehicles with velocity and heading rate constraints. *IEEE Transactions on Control Systems Technology*, 12(5):706–716, September 2004.

Wei Ren and Randal W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, 50(5):655–661, May 2005.

Herbert G. Tanner, Ali Jadbabaie, and George J. Pappas. Stability of flocking motion. Technical report, University of Pennsylvania, 2003.

Massimo Vergassola, Emmanuel Villermaux, and Boris I. Shraiman. 'infotaxis' as a strategy for searching without gradients. *Naure*, 445(25):406–409, 2007.

Christopher K. Wikle and Noel Cressie. A dimension-reduced approach to space-time kalman filtering. *Biometrika*, 86(4): 815–829, 1999.