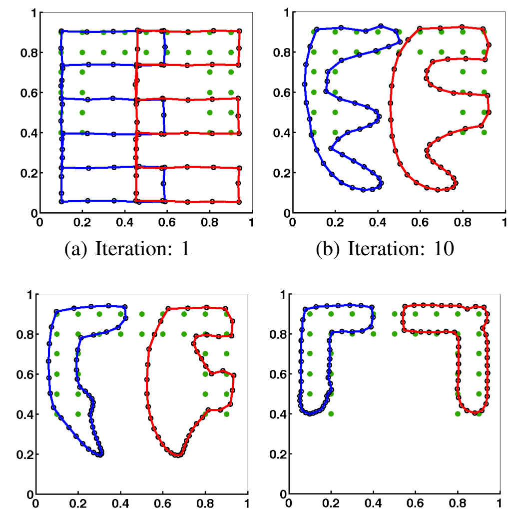
Life-long Informative Paths for Sensing Unknown Environments

Daniel E. Soltero Mac Schwager Daniela Rus

***Abstract*—In this paper, we have a team of robots in a dynamic unknown environment and we would like them to have accurate information about the environment for all time. That is, the error between the robots model of the environment and the actual environment must be bounded for all time, despite dynamic changes in the environment. We present an adaptive control law for the robots to shape their paths to locally optimal configurations for persistently sensing an unknown dynamical environment. Persistent sensing tasks are concerned with controlling the trajectories of mobile robots to act in a growing field in the environment in a way that guarantees that the field remains bounded for all time. With the adaptive controller, as the robots travels through their paths, they identify the areas where the environment is dynamic and shape their paths to sense these areas. A Lyapunov-like stability proofs is used to show that, under the proposed controller, the paths converge to locally optimal configurations according to a Voronoi-based coverage task, referred to as informative paths. Simulated and experimental results support the proposed approach.**



(a) Iteration: 1

(b) Iteration: 10

(c) Iteration: 30

(d) Iteration: 1130

Fig. 1: Path shaping phase of the multi-robot system hardware implementation. The paths, shown as the blue and red lines connect all the waypoints, shown as black circles. The dynamic regions (points of interest) are represented as green dots. As time passes by, the robots shape their paths so they can cover the dynamic regions of the environment.

I. Introduction

Robots operating in dynamic and unknown environments are often faced with the problem of deciding where to go to get the most relevant information for their current task. Given a dynamic environment that is unknown, and a group of robots, each with a sensor to measure the environment, we want to find a set of paths for the robots that will maximize information gathering in long-term operation. To achieve this, the robots need to do three things: 1) learn the structure of the environment by identifying the areas within the environment that are dynamic and the rate of change for these areas; 2) compute paths which allow them to sense the parts of the environment that have high rate of change; and 3) control their speeds along these paths for all time so that they always have time-accurate information of the environment. These paths are called *informative paths* since they focus on driving the robots through locations in the environment where the sensory information is important.

In this paper[1] we introduce the informative path planning problem and present a new control algorithm for generating closed informative paths in unknown environments. The key insight is inspired by [2]. The robots move along their paths, marking the areas they observe as dynamic or static. As the robots discover the static/dynamic structure of the environment, they reshape their paths to avoid visiting static areas and focus on sensing dynamic areas. An example of this reshaping process for two robots can be seen in Figure 1.

The adaptive controller has two key features. The first is an adaptation law that the robots use to learn how sensory information is distributed in the environment, through parameter estimation. The second feature is a Voronoi-based coverage controller (building upon [2]) that performs the reshaping of the paths based on the robots' estimates of the space.

The informative path problem has many applications. For example, a surveillance task of a city where the robots could learn the regions where crime is frequently committed and generate paths so that they can sense these crime regions more frequently and sense the safer regions less frequently. In this paper we are interested in using it to achieve and improve *persistent sensing tasks* [3] in unknown environments, where the robots are assumed to have sensors with finite footprints.

More specifically, in a persistent sensing task we wish for the robots with finite sensor footprints to conduct their information gathering while guaranteeing a bound on the difference between the robots' current model of the environment and the actual state of the environment for all time and all locations. Since their sensors have finite footprints, the robots cannot collect the data about all of the environment at once. As data about a dynamic region becomes outdated, the robots must return to that region repeatedly to collect new data. In previous work [3], a persistent sensing controller calculates the speeds

This material is based upon work supported in part by ONR-MURI Award N00014-09-1-1051, the NSF Graduate Research Fellowship Award 0645960 & The Boeing Company.

D. E. Soltero and D. Rus are with the Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA (soltero@mit.edu, rus@csail.mit.edu). M. Schwager is with the Mechanical Engineering Department, Boston University, Boston, MA (schwager@bu.edu).

[1]This paper's content is extracted from the first author's Master's thesis [1].

of the robots at each point along given paths (referred to as the *speed profiles*) in order to perform a persistent sensing task, i.e. to prevent the robots' model of the environment from becoming too outdated. The intuition behind this speed controller is to visit faster changing areas more frequently than slower changing areas.

The persistent sensing problem is defined in [3] as an optimization problem whose goal is to keep a time changing field as low as possible everywhere for all time. We refer to this field as the *accumulation function*. The accumulation function grows where it is not covered by the robots' sensors, indicating a growing need to collect data at that location, and shrinks where it is covered by the robots' sensors, indicating a decreasing need for data collection. A stabilizing speed profile is one which maintains the height of the accumulation function bounded for all time. In this paper we show that our computed informative paths can be used in conjunction with the stabilizing speed profiles from the persistent sensing controller to locally optimize a persistent sensing task.

The contributions of this paper are:

- a provably stable adaptive controller for learning the location of dynamic events in an environment and simultaneously computing informative paths for these events in order to locally optimize a persistent sensing task. We call this the *persistent informative controller*;
- simulation and hardware implementation.

A. Relation to Previous Work

Most of the previous work in path planning/generation focuses on reaching a goal while avoiding collision with obstacles, e.g. [4], or on computing an optimal path according to some metric, e.g. [5]. Other works have focused on probabilistic approaches to path planning, e.g. [6]. Prior work in adaptive path planning, e.g. [7], considers adapting the robot's path to changes in the robot's knowledge of the environment. In this paper, we focus on generating paths that allow the robots to perpetually travel through regions of interest in an unknown environment. We use adaptive control tools to create a novel algorithm for computing informative paths. We use an approach based on Voronoi partitions, similar to [8]. However, contrary to [8], the environment, although unknown, is not random, and we are not concerned with optimizing the trajectory of the agents to minimize the predictive variance, but rather optimizing the location of agents for a coverage task in an unknown environment.

The adaptive controller relevant to this thesis was introduced in [2], where a group of agents were coordinated to place themselves in locally optimal locations to sense an unknown environment. This paper builds upon this previous work, but with some significant changes and additions. The Voronoi partitions are used by the robots to change the location of the agents, which are now waypoints defining the robots' paths. The paths must be closed paths and must provide good sensing locations for the robots. Also, we wish the robots to perform persistent sensing along these paths.

The persistent sensing concept motivating this paper was introduced in [3], where a linear program was designed to calculate the robots' speed profiles in order for them to perform a persistent sensing task, that is, maintain the height of the growing accumulation function bounded. The robots were assumed to have full knowledge of the environment and were given a pre-designed path. In this paper we alleviate these two assumptions by having the robots learn the environment through parameter estimation, and use this information to shape their paths into useful paths. These two alleviations provide a significant step towards persistent sensing in dynamic environments.

In Section II we set up the problem, present locational optimization tools using a Voronoi-based approach, and present a basis function approximation of the environment. In Section III, we introduce the persistent informative controller and prove stability of the system under this controller. Finally, in Section IV we provide simulated and hardware results.

II. Problem Formulation

We assume we are given multiple robots whose tasks are to sense an unknown dynamical environment while traveling along their individual closed paths, each consisting of a finite number of waypoints. The goal is for the robots to identify the areas within the environment that are dynamic and compute paths that allow them to jointly sense the dynamic areas. A formal mathematical description of the problem follows.

Let there be N robots, identified by $r \in \{1, \ldots, N\}$, in a convex, bounded area $Q \subset \mathbb{R}^2$. An arbitrary point in Q is denoted $\mathbf{q}$. Robot r is located at position $p_r \in Q$ and travels along its closed path consisting of a finite number $n(r)$ of waypoints. The position of the i^{th} waypoint in robot r's path is denoted p_i^r, $i \in \{1, \ldots, n(r)\}$. Let $\{p_i^r, \ldots, p_{n(r)}^r\}$ be the configuration of robot r's path and let V_i^r be a Voronoi partitions of Q, with the i^{th} waypoint position in robot r's path as the generator point, defined as

$$V_i^r = \{\mathbf{q} \in Q : \|\mathbf{q} - p_i^r\| \leq \|\mathbf{q} - p_{i'}^{r'}\|, \ \forall (r', i') \neq (r, i)\},$$
$$r, r' \in \{1, \ldots, N\}, \ i \in \{1, \ldots, n(r)\}, \ i' \in \{1, \ldots, n(r')\}, \quad (1)$$

where, $\|\cdot\|$ denotes the l^2-norm. We assume that the robot is able to compute the Voronoi partitions based on the waypoint locations, as is common in the literature [2].

Since each path is closed, each waypoint i along robot r's path has a previous waypoint $i-1$ and next waypoint $i+1$ related to it, which are referred to as the neighbor waypoints of i. Note that for each r, $i+1 = 1$ for $i = n(r)$, and $i-1 = n(r)$ for $i = 1$. Once a robot reaches a waypoint, it continues to move to the next waypoint along its path, in a straight line interpolation.[2]

We assume that the network of robot's in the system is a connected network, i.e. the graph where each robot is a node and an edge represents communication between two nodes is a connected graph. This connected network corresponds to the robots' abilities to communicate with each other.

[2]We assume that the waypoints do not outrun the robots under the action of the informative path controller.

A. Environment Structure

A sensory function, defined as a map $\phi : Q \mapsto \mathbb{R}_{\geq 0}$ (where $\mathbb{R}_{\geq 0}$ refers to non-negative scalars) determines the constant rate of change of the environment at point $\mathbf{q} \in Q$. The function $\phi(\mathbf{q})$ is not known by the robots, but each robot is equipped with a sensor to make a point measurement of $\phi(p_r)$ at the robot's position p_r.

The interpretation of the sensory function $\phi(\mathbf{q})$ may be adjusted for a broad range of applications. It can be any kind of weighting of importance for points $\mathbf{q} \in Q$. In this paper we treat it as a rate of change in a dynamic environment.

B. Locational Optimization

Building upon [2], we can formulate the cost incurred by the multi-robot system over the region Q as

$$\begin{aligned}\mathcal{H} \;=\; & \sum_{r=1}^{N}\sum_{i=1}^{n(r)} \int_{V_i^r} \frac{W_s}{2}\|\mathbf{q} - p_i^r\|^2 \phi(\mathbf{q}) d\mathbf{q} \\ & + \sum_{r=1}^{N}\sum_{i=1}^{n(r)} \frac{W_n}{2}\|p_i^r - p_{i+1}^r\|^2, \end{aligned} \tag{2}$$

where $\|\mathbf{q} - p_i^r\|$ can be interpreted as the unreliability of the sensory function value $\phi(\mathbf{q})$ when robot r is at p_i^r, and $\|p_i^r - p_{i+1}^r\|$ can be interpreted as the cost of a waypoint being too far away from a neighboring waypoint for robot r's path. W_s and W_n are constant positive scalar weights assigned to the sensing task and neighbor distance, respectively. Note that unreliable sensing and distance between neighboring waypoints are expensive. The first term couples all robots and their respective waypoints, making them work together to cover the environment. The second term couples the waypoints, making a path for each robot. A formal definition of informative paths for multiple robots follows.

Definition II.1 (Informative Paths for Multiple Robots)**.** A collection of informative paths for a multi-robot system corresponds to a set of waypoint locations for each robot that locally minimize (2).

Next we define three properties analogous to mass-moments of rigid bodies. The mass, first mass-moment, and centroid of V_i^r are defined as $M_i^r = \int_{V_i^r} W_s \phi(\mathbf{q}) d\mathbf{q}$, $L_i^r = \int_{V_i^r} W_s \mathbf{q} \phi(\mathbf{q}) d\mathbf{q}$, $C_i^r = \frac{L_i^r}{M_i^r}$, respectively.

Also, let $e_i^r = C_i^r - p_i^r$. From [1], we have $\frac{\partial \mathcal{H}}{\partial p_i^r} = -M_i^r e_i^r - W_n(p_{i+1}^r + p_{i-1}^r - 2p_i^r)$, and an equilibrium is reached when $\frac{\partial \mathcal{H}}{\partial p_i^r} = 0$. Assigning to each waypoint dynamics of the form

$$\dot{p}_i^r = u_i^r, \tag{3}$$

where u_i^r is the control input, we propose the following control law for the waypoints to converge to an equilibrium configuration: $u_i^r = \left(K_i^r(M_i^r e_i^r + \alpha_i^r)\right)/\beta_i^r$, where $\alpha_i^r = W_n(p_{i+1}^r + p_{i-1}^r - 2p_i^r)$, $\beta_i^r = M_i^r + 2W_n$ and K_i^r is a uniformly positive definite matrix and potentially time-varying to improve performance. Note that $\beta_i^r > 0$ and has the nice effect of normalizing the weight distribution between sensing and staying close to neighboring waypoints.

C. Sensory Function Approximation

We assume that the sensory function $\phi(\mathbf{q})$ can be parameterized as an unknown linear combination of a set of known basis functions. That is,

Assumption II.2 (Matching condition). $\exists a \in \mathbb{R}^m_{\geq 0}$ and $\mathcal{K} : Q \mapsto \mathbb{R}^m_{\geq 0}$, where $\mathbb{R}^m_{\geq 0}$ is a vector of non-negative entries, such that

$$\phi(\mathbf{q}) = \mathcal{K}(\mathbf{q})^T a, \tag{4}$$

where the vector of basis functions $\mathcal{K}(\mathbf{q})$ is known by the robot, but the parameter vector a is unknown.

Let $\hat{a}_r(t)$ be robot r's approximation of the parameter vector a. Then, $\hat{\phi}_r(\mathbf{q}) = \mathcal{K}(\mathbf{q})^T \hat{a}_r$ is robot r's approximation of $\phi(\mathbf{q})$. Then, we define the mass moment approximations as

$$\hat{M}_i^r = \int_{V_i^r} W_s \hat{\phi}_r(\mathbf{q}) d\mathbf{q}, \qquad \hat{L}_i^r = \int_{V_i^r} W_s \mathbf{q} \hat{\phi}_r(\mathbf{q}) d\mathbf{q},$$

$$\hat{C}_i^r = \frac{\hat{L}_i^r}{\hat{M}_i^r}.$$

Additionally, we can define $\tilde{a}_r = \hat{a}_r - a$, and the sensory function error, and mass moment errors as

$$\tilde{\phi}_r(\mathbf{q}) = \hat{\phi}_r(\mathbf{q}) - \phi(\mathbf{q}) = \mathcal{K}(\mathbf{q})^T \tilde{a}_r, \tag{5}$$

$$\tilde{M}_i^r = \hat{M}_i^r - M_i^r = \int_{V_i^r} W_s \mathcal{K}(\mathbf{q})^T d\mathbf{q}\ \tilde{a}_r, \tag{6}$$

$$\tilde{L}_i^r = \hat{L}_i^r - L_i^r = \int_{V_i^r} W_s \mathbf{q} \mathcal{K}(\mathbf{q})^T d\mathbf{q}\ \tilde{a}_r, \tag{7}$$

$$\tilde{C}_i^r = \frac{\tilde{L}_i^r}{\tilde{M}_{ir}}. \tag{8}$$

Finally, in order to compress the notation, let $\mathcal{K}_{p_r}(t)$ and $\phi_{p_r}(t)$ be the value of the basis function vector and the value of ϕ at robot r's position $p_r(t)$, respectively.

III. Informative Paths for Persistent Sensing

A. Relation to Persistent Sensing Tasks

Previous work [3] developed a way to calculate the speed of robots traveling along known static paths in order to sense the environment and maintain the height of the accumulation function bounded for all time. We assume each robot is equipped with a sensor with a finite footprint $F_r(p_r) = \{\mathbf{q} \in Q : \|\mathbf{q} - p_r\| \leq \rho\}$[3] when the robot is at location p_r, where ρ is a constant positive scalar. The accumulation function grows where it is not covered by any robot sensor, and shrinks otherwise. Mathematically, the accumulation function, referred to as $Z(\mathbf{q}, t)$ at time t for point $\mathbf{q}$, evolves according to

$$\dot{Z}(\mathbf{q}, t) = \begin{cases} \phi(\mathbf{q}) - \displaystyle\sum_{r \in \mathcal{N}_{\mathbf{q}}(t)} c_r(\mathbf{q}), & \text{if } Z(\mathbf{q}, t) > 0, \\ \Big(\phi(\mathbf{q}) - \displaystyle\sum_{r \in \mathcal{N}_{\mathbf{q}}(t)} c_r(\mathbf{q})\Big)^+, & \text{if } Z(\mathbf{q}, t) = 0, \end{cases} \tag{9}$$

[3] Any footprint shape can be used, and the footprint size does not have to be the same for all robots. For simplicity, we use a circular footprint with same size for all robots.

where $\mathcal{N}_{\mathbf{q}}(t)$ is the set of robots whose sensor footprints are over the point $\mathbf{q}$ at time t, i.e. $\mathcal{N}_{\mathbf{q}}(t) = \{r \mid \mathbf{q} \in F_r(p_r(t))\}$.

We would like to use the stability criterion for a persistent sensing task, defined in [3], and plug it into the adaptive controller, such that the control action increases the stability margin of the persistent sensing task through time. However, since the robots do not know the environment, but have estimates of it, each robot r uses the estimated version of this stability criterion, given by

$$\hat{\phi}_r(\mathbf{q},t) - \sum_{r'=1}^{N} \frac{\tau_c^{r'}(\mathbf{q},t)}{T_{r'}(t)} c_{r'}(\mathbf{q}) = \hat{s}_r(\mathbf{q},t) < 0, \quad \forall \mathbf{q} \mid \hat{\phi}_r(\mathbf{q},t) > 0. \tag{10}$$

where $\hat{\phi}_r(\mathbf{q},t)$ (the estimated sensory function) is the estimated rate at time t at which the accumulation function grows at point $\mathbf{q}$, the constant scalar $c_r(\mathbf{q})$ is the rate at which the accumulation function shrinks when robot r's sensor is covering point $\mathbf{q}$ (and is known by the robots), $T_r(t)$ is the time it takes robot r to complete the path at time t, and $\tau_c^r(\mathbf{q},t)$ is the time robot r's sensor covers point $\mathbf{q}$ along the path at time t. These two last quantities are calculated by the robots with the speed profiles. Each robot also knows its estimated stability margin at time t, defined as $\hat{S}_r(t) = -(\max_{\mathbf{q}} \hat{s}_r(\mathbf{q},t))$. A stable persistent task at time t is one in which $S(t)$ (the true version of $\hat{S}_r(t)$) is positive, which means the robot is able to maintain the height of the accumulation function bounded at all points $\mathbf{q}$. Note that only points $\mathbf{q}$ that satisfy $\hat{\phi}_r(\mathbf{q},t) > 0$ are considered in a persistent sensing task since it is not necessary to persistently sense a point that has no sensory interest. Points that satisfy this condition are referred to as *points of interest*.[4]

In [3], a linear program was given which can calculate the speed profiles for the paths at time t that maximize $\hat{S}_r(t)$ (or $S_r(t)$ for ground-truth), for all r. Hence, from this point on, we assume that the robots know these maximizing speed profiles and use them to obtain $\hat{s}_r(\mathbf{q},t),\ \forall \mathbf{q},\ \forall t$.

B. Persistent Informative Controller

We design an adaptive control law and prove that it causes the paths to converge to a locally optimal configuration according to (2), while improving the stability margin of the persistent sensing task and causing all of the robots' estimates of the environment to converge to the real environment description. All of the robots' estimates of the environment converge to a same estimate due to a consensus term [2] used in their adaptive laws.

Since the robots do not know $\phi(\mathbf{q})$, but have estimates $\hat{\phi}_r(\mathbf{q})$, the control law becomes

$$u_i^r = \frac{K_i^r(\hat{M}_i^r \hat{e}_i^r + \alpha_i^r)}{\hat{\beta}_i^r}, \tag{11}$$

where

$$\begin{aligned} \alpha_i^r &= W_n(p_{i+1}^r + p_{i-1}^r - 2p_i^r), & (12)\\ \hat{\beta}_i^r &= \hat{M}_i^r + 2W_n, & (13)\\ \hat{e}_i^r &= \hat{C}_i^r - p_i^r. & (14) \end{aligned}$$

Additionally, to incorporate the stability criterion for persistent sensing tasks from (10), let the waypoints have new dynamics of the form

$$\dot{p}_i^r = I_i^r u_i^r, \tag{15}$$

where u_i^r is defined in (11), and

$$I_i^r = \begin{cases} 0, & \text{if } \frac{\partial \hat{S}_r}{\partial p_i^r}^T u_i^r < 0 \text{ and } t - t_u^{r,i} > \tau_{\text{dwell}}, \\ 1, & \text{otherwise,} \end{cases} \tag{16}$$

τ_{dwell} is a design parameter, and $t_u^{r,i}$ is the most recent time at which I_i^r, switched from zero to one (switched "up"). Equations (15) and (16) ensure that the estimated stability margin for the persistent sensing task does not decrease.

Remark III.1. For positive $\epsilon \to 0$, $\frac{\partial \hat{S}_r}{\partial p_i^r}(t)$ is not always defined when $\arg\max_{\mathbf{q}} \hat{s}_r(\mathbf{q}, t-\epsilon) \neq \arg\max_{\mathbf{q}} \hat{s}_r(\mathbf{q}, t+\epsilon)$. In such cases $\frac{\partial \hat{S}_r}{\partial p_i^r}(t)$ refers to $\frac{\partial \hat{S}_r}{\partial p_i^r}(t+\epsilon)$.

The parameter $\hat{a}_r$ is adjusted according to

$$\begin{aligned} \Lambda_r &= \int_0^t w_r(\tau)\mathcal{K}_{p_r}(\tau)\mathcal{K}_{p_r}(\tau)^T d\tau, & (17)\\ \lambda_r &= \int_0^t w_r(\tau)\mathcal{K}_{p_r}(\tau)\phi_{p_r}(\tau) d\tau, & (18) \end{aligned}$$

where $w_r(t)$ is a positive constant scalar if $t < \tau_{w_r}$ and zero otherwise, and τ_{w_r} is some positive time at which part of the the adaptation for robot r shuts down to maintain Λ_r and λ_r bounded. Let

$$b_r = \sum_{i=1}^{n(r)} \int_{V_i^r} W_s \mathcal{K}(\mathbf{q})(\mathbf{q} - p_i)^T d\mathbf{q}\ \dot{p}_i^r, \tag{19}$$

$$\dot{\hat{a}}_{\text{pre}_r} = -b_r - \gamma(\Lambda_r \hat{a}_r - \lambda_r) - \zeta \sum_{r'=1}^{N} l_{r,r'}(\hat{a}_r - \hat{a}_{r'}), \tag{20}$$

where $\zeta > 0$ is a consensus scalar gain, and $l_{r,r'}$ can be interpreted as the strength of the communication between robots r and r' and is defined as

$$l_{r,r'} = \begin{cases} D_{\max} - \|p_r - p_{r'}\|, & \text{if } \|p_r - p_{r'}\| \leq D_{\max} \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

Since $a(j) \geq 0,\ \forall j$, we enforce $\hat{a}_r(j) \geq 0,\ \forall r, \forall j$, by the projection law [2],

$$\dot{\hat{a}}_r = \Gamma(\dot{\hat{a}}_{\text{pre}_r} - I_{\text{proj}_r}\dot{\hat{a}}_{\text{pre}_r}), \tag{22}$$

where $\Gamma \in \mathbb{R}^{m\times m}$ is a diagonal positive definite adaptation gain matrix, and the diagonal matrix I_{proj_r} is defined element-wise as

$$I_{\text{proj}_r}(j) = \begin{cases} 0, & \text{if } \hat{a}_r(j) > 0, \\ 0, & \text{if } \hat{a}_r(j) = 0 \text{ and } \dot{\hat{a}}_{\text{pre}_r}(j) \geq 0, \\ 1, & \text{otherwise,} \end{cases} \tag{23}$$

[4]We assume the environment contains a finite number of points of interests. These finite points could be the discretization of a continuous environment.

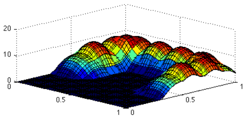


Fig. 2: Environment for the multi-robot hardware implementation.

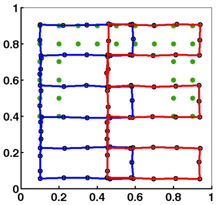


Fig. 3: Initial paths for the multi-robot hardware implementation

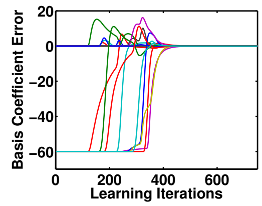


Fig. 4: Basis coefficient error

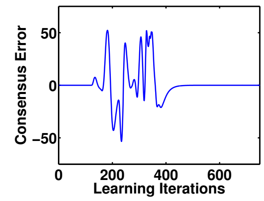


Fig. 5: Consensus error

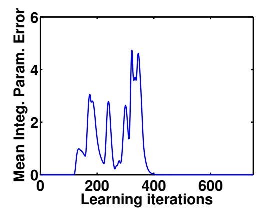


Fig. 6: Integral parameter error

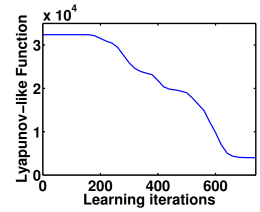


Fig. 7: Lyapunov-like function in learning phase

where (j) denotes the j^{th} element for a vector and the j^{th} diagonal element for a matrix.

Theorem III.2 (Convergence Theorem for Persistent Sensing by Multiple Robots)**.** *Under Assumption II.2, with waypoint dynamics specified by (15), control law specified by (11), and adaptive law specified by (22), we have*

(i) $\lim_{t\to\infty} I_i^r(t)\|\hat{M}_i^r(t)\hat{e}_i^r(t)+\alpha_i^r(t)\|=0,$
$\forall r\in\{1,\ldots,N\},\ \forall i\in\{1,\ldots,n(r)\},$

(ii) $\lim_{t\to\infty}\|\tilde{\phi}_{p_r}(\tau)\|=0,$
$\forall r\in\{1,\ldots,N\},\ \forall\tau\mid w_r(\tau)>0,$

(iii) $\lim_{t\to\infty}(\hat{a}_r-\hat{a}_{r'})=0, \qquad \forall r,r'\in\{1,\ldots,N\}.$

We refer the reader to [1] to view the details of the proof. Properties (ii) and (iii) from Theorem III.2 together imply that the robots will learn the true sensory function for the environment if the union of their trajectories while their weights are positive is rich enough. Therefore, we can design the initial waypoint locations such that, between all robots, most of the dynamic unknown environment is explored (see Figure 3).

The stability margin can theoretically worsen while I_i^r, for some r,i, cannot switch from one to zero because it is waiting for $t-t_u^{r,i}>\tau_{\text{dwell}}$. However, τ_{dwell} can be selected arbitrarily small and, in practice, any computer will enforce a τ_{dwell} due to discrete time steps. Therefore, it is not a practical restriction. As a result, intuitively, (i) from Theorem III.2 means that $\lim_{t\to\infty}\|\hat{M}_i^r(t)\hat{e}_i^r(t)+\alpha_i^r(t)\|=0$ only if this helps the persistent sensing task. Otherwise $\lim_{t\to\infty} I_i^r(t)=0$, meaning that the persistent sensing task will not benefit if the i^{th} waypoint in robot r's path moves.

IV. Implementation and Results

We simulated the system many times and executed an implementation with two quadrotors more than 10 times. Here we present a case for $N=2$ robots, $n(r)=44$ waypoints, $\forall r$. A fixed-time step numerical solver is used with a time step of 0.01 seconds and $\tau_{\text{dwell}}=0.009$. The environment parameters are $\sigma=0.4$ and $\rho_{\text{trunc}}=0.2$, $a(j)=60$, for $j\in\{3,4,5,10,15,20,23,24,25\}$, and $a(j)=0$ otherwise. The environment created with these parameters can be seen in Figure 2. The parameters $\hat{a}_r$, Λ_r and λ_r, for all r are initialized to zero. The parameters for the controller are $K_i^r=90,\ \forall i,r$, $\Gamma=\text{identity}$, $\gamma=3000$, $W_n=6$, $W_s=150$, $w_r=3$, $\forall r$ and $\rho=0.12$. In addition, $D_{\max}$ is assumed to be very large, so that $l_{r,r'}(t)\zeta=20,\ \forall r,r',\ \forall t$. The environment is discretized into a 10×10 grid and only points in this grid with $\hat{\phi}_r(\mathbf{q})>0$ are used as points of interest in (10). By only using this discretized version of the environment, the running time for experiments is greatly reduced. For more sensitive systems, this grid can be refined.

The environment and, therefore, sensor measurements are simulated. The growing behavior of the accumulation function over the environment follows the description from (9). As an implementation detail, although $\rho=0.12$ for purposes of the persistent informative controller, a value of $\rho=0.126$ (5% increase) was used on the physical robot to consume the accumulation function in the environment, which allows to overcome the effects of small tracking errors from the quadrotors and the effects of the discretization of the path for the persistent sensing task.

The initial paths can be seen in Figure 3, where each robot has a "zig-zagging" path across a portion of the environment, and between both robots, most of the environment is initially traversed. We first allowed the robots to go through their initial paths without reshaping them so that they can sample most of the space and learn the distribution of sensory information in the environment. Therefore, we present results in two separate phases: 1) learning phase, and 2) path shaping phase. The learning phase corresponds to the robots traveling through their whole paths once, without reshaping them, in order to learn the environment. The path shaping phase corresponds to when (15) is used to reshape the paths into informative paths, and starts after the learning phase is done by both robots. In the path shaping phase, $w_r=0,\ \forall r$.

A. Learning Phase

In the learning phase, we can see from Figure 4 that $\tilde{\phi}_r(\mathbf{q})\to0$, $\forall\mathbf{q}\in Q$ for one of the robots. Since the consensus error converges to zero in accordance with (iii) from Theorem III.2 and shown in Figure 5, then we can conclude that the adaptation laws cause $\tilde{\phi}_r(\mathbf{q})\to0$, $\forall\mathbf{q}\in Q$, $\forall r$. This means

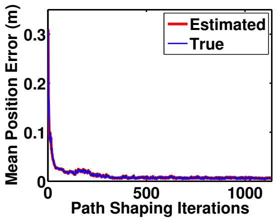


Fig. 8: Mean waypoint position error

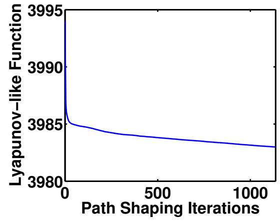


Fig. 9: Lyapunov-like function in path shaping phase

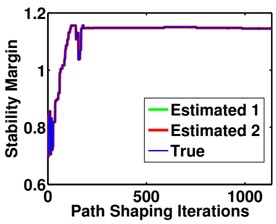


Fig. 10: Persistent sensing task's stability margin

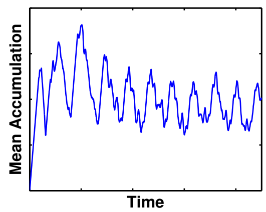


Fig. 11: Mean accumulation function value

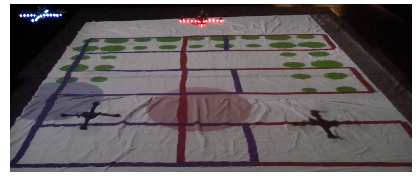
that the union of both robots' trajectories was rich enough to generate accurate estimates for all of the environment. Figure 6 shows that the mean over both robots of $\int_0^t w_r(\tau)(\tilde{\phi}_{p_r}(\tau))^2 d\tau$ converges to zero, in accordance with (ii) from Theorem III.2. Finally, for this learning phase, we see in Figure 7 that the Lyapunov-like function $\mathcal{V}_4$ is monotonically non-increasing.

B. Path Shaping Phase

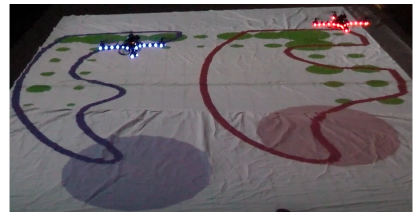
Figure 12 shows snapshots of the multi-robot implementation during the path shaping phase. We can see how the path evolves under this controller in the path shaping phase in Figures 1a to 1d. Figure 8 shows quantity $I_i^r(t)\|\hat{M}_i^r(t)\hat{e}_i^r(t)+\alpha_i^r(t)\|$ converging to zero, $\forall i, r$ in accordance to (i) from Theorem III.2. Figure 9 shows the Lyapunov-like function $\mathcal{V}_4$ monotonically non-increasing and reaching a limit. Figure 10 shows the persistent sensing task's stability margin increasing through time, as expected. The chattering in the stability margin is due to the discretization of the system, and can be reduced by shortening the length of time steps. Finally, Figure 11 shows the mean over all points of interest of the value of the accumulation function over time. As shown, this value initially increases on average due to the initialization of the system. Later, it starts to decrease and reaches an approximate steady-state behavior that corresponds to the locally optimal final configuration of the system. The hardware implementation was run for more than 10 times, generating informative paths that are practically identical to the simulated cases.

V. Conclusion

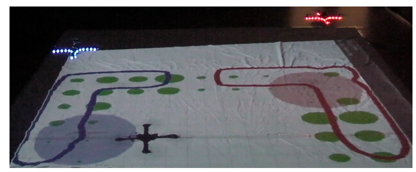
This paper uses a Voronoi-based coverage approach, building upon previous work in [2], to generate an adaptive controller for robots to learn the environment sensory function through parameter estimation and shape their paths into informative paths that can be used for persistent sensing, i.e. locally optimal paths for sensing dynamic regions in the environment and performing persistent sensing tasks. The persistent informative controller was implemented with two quadrotor robots, generating results that support the theory.



(a) Iteration 0



(b) Iteration 20



(c) Iteration 1130

Fig. 12: Three snapshots of the hardware implementation of the multi-robot system during the path shaping phase at different iteration values. The paths, shown as the blue and red lines, connects all the waypoints corresponding to each robot. The points of interest in the environment are shown as green dots, and the size of a green dot represents the value of the accumulation function at that point. The robots are the blue-lit and red-lit quadrotors, and their sensor footprints are represented by the colored circles under them.

References

[1] D. E. Soltero, "Generating informative paths for persistent sensing in unknown environments," Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, June 2012. [Online]. Available: http://groups.csail.mit.edu/drl/wiki/images/1/1e/SolteroThesisSM.pdf

[2] M. Schwager, D. Rus, and J.-J. Slotine, "Decentralized, adaptive coverage control for networked robots," *The International Journal of Robotics Research*, vol. 28, no. 3, pp. 357–375, 2009. [Online]. Available: http://ijr.sagepub.com/content/28/3/357.abstract

[3] S. L. Smith, M. Schwager, and D. Rus, "Persistent robotic tasks: Monitoring and sweeping in changing environments," *Robotics, IEEE Transactions on*, vol. PP, no. 99, pp. 1 –17, 2011.

[4] A. Piazzi, C. Guarino Lo Bianco, and M. Romano, "η^3-splines for the smooth path generation of wheeled mobile robots," *Robotics, IEEE Transactions on*, vol. 23, no. 5, pp. 1089 –1095, oct. 2007.

[5] K. Kyriakopoulos and G. Saridis, "Minimum jerk path generation," in *Robotics and Automation, 1988. Proceedings., 1988 IEEE International Conference on*, apr 1988, pp. 364 –369 vol.1.

[6] L. Kavraki, P. Svestka, J.-C. Latombe, and M. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *Robotics and Automation, IEEE Transactions on*, vol. 12, no. 4, pp. 566 –580, aug 1996.

[7] C. Cunningham and R. Roberts, "An adaptive path planning algorithm for cooperating unmanned air vehicles," in *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, vol. 4, 2001, pp. 3981 – 3986 vol.4.

[8] R. Graham and J. Cortes, "Adaptive information collection by robotic sensor networks for spatial estimation," *Automatic Control, IEEE Transactions on*, vol. PP, no. 99, p. 1, 2011.