Technical Report

Department of Computer Science
and Engineering
University of Minnesota
4-192 Keller Hall
200 Union Street SE
Minneapolis, MN 55455-0159 USA

TR 13-010

Sensor Planning for a Symbiotic UAV and UGV system for Precision Agriculture

Pratap Tokekar, Joshua Vander Hook, David Mulla, Volkan Isler

March 25, 2013

Revised: August 9, 2013

Sensor Planning for a Symbiotic UAV and UGV system for Precision Agriculture

Pratap Tokekar, Joshua Vander Hook, David Mulla and Volkan Isler

***Abstract*— We study the problem of coordinating an Unmanned Aerial Vehicle (UAV) and Unmanned Ground Vehicle (UGV) to collect data for a precision agriculture application. The ground and aerial measurements collected by the system are used for estimating Nitrogen (N) levels across a farm field. These estimates in turn guide fertilizer application. The capability to apply the right amount of fertilizer at the right time can drastically reduce fertilizer usage which is desirable from an environmental and economic standpoint.**

We propose to use a symbiotic UAV and UGV system in which the UGV is capable of muling the UAV to various deployment locations. This would allow the system to overcome the short battery life of a typical UAV. Our goal is to estimate N levels over the field and assign each point in the field into classes indicating N-deficiency levels. Towards building such a system, the paper makes the following contributions: First, we present a method to identify points whose probability of being misclassified is above a threshold, termed as Potentially Mislabeled (PML). Second, we study the problem of planning the UAV path to visit the maximum number of PML points subject to its energy budget. The novelty of our formulation is the capability of the UGV to mule the UAV to deployment points. Third, we introduce a new path planning problem in which the UGV must take a measurement near each PML point visited by the UAV. The goal is to minimize the total time spent in traveling and taking measurements. For both problems, we present constant-factor approximation algorithms. Finally, we demonstrate the utility of the system and our algorithms with simulations which use manually collected data from the field as well as realistic energy models for the UAV and the UGV.

I. Introduction

Precision agriculture offers to improve crop productivity and farm profitability through improved management of farm inputs, leading to better environmental quality [1]. For example, by measuring Nitrogen levels in the soil across a farm and applying the right level of nitrogen at the right time and place, it is possible to reduce fertilizer usage by 25 percent without affecting corn yield [2].

One of the key components of precision agriculture is data collection. At present, there are two primary approaches to data collection for precision agriculture: remote sensing and manual data collection. Satellite and aerial remote sensing are severely limited by cloud cover [3]. Satellite images may not be available at desired times (the update frequency of satellite images can be anywhere from 3 to 26 days). Remote sensing from a manned aerial device (either flown by a pilot or radio-controlled by a professional operator) is costly and difficult to plan against weather conditions. Further, soil moisture, crop height, and pest infestations cannot be measured remotely in a vegetated crop. Manual methods involve the collection of data by humans who either gather the data by walking through the field, or guide a vehicle equipped with a sensor [4]. Manual data collection process can be tedious and time-consuming and we are working on building a robotic system which uses low-cost Unmanned Aerial Vehicles and Unmanned Ground Vehicles (UGVs) working in conjunction to collect such data.

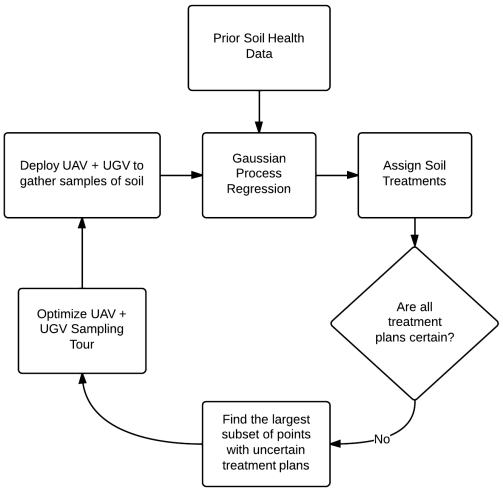


Fig. 1. The problem formulation: Precision Agriculture using the UAV+UGV system. Prior soil health data is available, but a soil treatment plan needs to be developed. If practitioners are uncertain of the treatment plan associated with some part of the field, mobile robots can provide immediate information about the soil health in that region. The concept of assigning treatments based on real-time data is called Precision Agriculture.

Our system will provide multiple on-demand sensing capabilities, and combine the strengths of ground and aerial robots: Ground robots are capable of traveling long distances, carrying large loads and measuring soil data. On the other hand, small aerial vehicles can take images from a low altitude but have limited battery life and the images may need to be complemented with ground measurements. Our proposed solution is to use a UGV to deploy a small, inexpensive UAV at carefully selected measurement locations. As the UAV is taking aerial pictures over a small region, the UGV will take soil measurements. The UAV will then land on the UGV which will take the UAV to the next deployment location.

In this paper, we develop path planning algorithms for the UAV+UGV system for the application of estimating the

P. Tokekar, J. Vander Hook, and V. Isler are with the Department of Computer Science & Engineering, University of Minnesota, U.S.A. {tokekar,jvander,isler} at cs.umn.edu

D. Mulla is with the Deparment of Soil, Water and Climate, University of Minnesota, U.S.A. mulla003 at umn.edu

Nitrogen (N) levels in an agriculture plot. We start with a prior N level map of the field (e.g., obtained from satellite imagery). Our goal is to identify regions with N stress (deficiency) by assigning a label based on the estimated N level to each point. For this purpose, we first show how points with high probability of mislabeling error can be identified (see Section IV). The system is then charged with taking both aerial and soil measurements near these points. The main bottleneck in the system's lifetime is the battery life of the UAV. Therefore we seek to maximize the number of points that will be visited by the UAV+UGV subject to this battery constraint. Note that even though the UGV can mule the UAV between the measurement locations, landing on to the UGV and ascending to a predetermined height consumes energy. Our main contribution in Section V is to show how to compute a series of UAV deployments which visit at least a constant factor of the points visited by an optimal algorithm.

Taking a soil measurement with the UGV is likely to be time-consuming. However, we can reduce the data collection time through careful planning for the measurement locations by combining measurements of nearby points. Our second contribution is an approximation algorithm for solving a novel variant of the Traveling Salesperson Problem with Neighborhoods (TSPN), which we call the Sampling TSPN Problem, that considers both travel and measurement cost. Finally, armed with these two algorithms, we demonstrate the benefit of using the UAV+UGV system in a precision agriculture application through simulations using real data collected from a corn field (Section VI). Figure 1 shows an overview of our approach.

The rest of the paper is organized as follows. We begin by presenting the problem formulation in Section II. The related work in this area is presented in Section III. We present the method used to estimate the N-map and finding the potentially mislabeled points in Section IV. The sensor planning algorithms and their analysis are presented in Section V. Simulation results based on field data are presented in Section VI. We finally conclude with a discussion of future work in Section VII.

II. Problem Formulation

Our operating environment is a farm plot, which we discretize into a set of points $\mathcal{X} = \{x_1, x_2, \cdots, x_n\}$. As is common, we will use a Gaussian Process to estimate the N levels from prior measurements [5]. For each point we associate a most likely Nitrogen estimate as $N(x)$, with variance of the estimate as $\sigma(x)$. Our task is to find regions in the plot with similar N levels. For example, the task can be to classify each point in the plot into three labels: low N, medium N, and high N. In general, we are given a set of labels, and each label i is specified by a minimum and maximum N level, l_i^-, l_i^+ respectively. Since we do not have access to the true N levels and instead have a distribution $N(x)$, we associate with each label a probability of being correct. We define $\mathrm{P}_{lj}(x_i)$ as the probability that the label j for point x_i is correct $\mathrm{P}_{lj}(x_i) = \mathrm{P}(l_j^- \leq N(x_i) < l_j^+)$. The details for computing this probability are given in Section IV. Labels can then be assigned to points based on which is most likely to be correct, given the estimates of N levels at each point. We use the shorter notation $\mathrm{P}_l(x_i)$ to denote the probability of the most likely label.

We define Potentially Mislabeled (PML) points as all points in $\mathcal{X}$ for which the probability of the most-likely label being correct is below a user-desired value $\mathrm{P}_d \in (0, 1)$.

$$\mathcal{X}_{pml} = \{x_i \in \mathcal{X} : \mathrm{P}_l(x_i) \leq \mathrm{P}_d\}. \quad (1)$$

We can increase the probability of the label being correct by taking soil and aerial measurements near the PML points, as we will discuss in Section IV.

One simple strategy would be to take measurements at every PML point until we are sufficiently certain of that point's true label. However, we operate under two practical constraints: The UAV has a limited battery budget (denoted by B) and the UGV requires some non-zero time (denoted by C_g) to take a soil measurement. Our objective is to maximize the number of PML points sampled with the UAV and UGV subject to the UAV's battery constraint.

The UAV spends some part of its energy budget for each take-off and landing. We denote the average of these energy costs by C_a, so that a combined take-off and landing takes $2C_a$. We assume that the UAV and UGV travel at unit speed and the energy required to travel is proportional to the travel time. Hence, we use distance, time and energy interchangeably. Non-unit speeds can be easily accommodated in our analysis. If τ_a is a set of N deployments for the UAV, then the total cost of the UAV tour is given as $\mathrm{len}(\tau_a) + 2N \cdot C_a$, where $\mathrm{len}(\tau_a)$ is the sum of Euclidean lengths of paths in all deployments. Our problem can be concisely stated as,

Problem 1 (UAV coverage): Find a UAV tour τ_a consisting of one or more paths (each of which is associated with a take-off and landing location), to sample the maximum number of PML points, such that the cost of the tour is not greater than the UAV budget.

Given the PML points visited by the UAV, our next objective is to design a UGV path that obtains soil measurements in the least time. The spatial correlation of soil properties means that nearby points are likely to have the same N level. Hence, as described in Section IV, the location of the measurement relative to any PML point should be within the radius defined by the spatial correlation of the measured property. The UGV can thus combine measurement locations for multiple points if their radii overlap. The cost to take measurements is included as an additive time cost C_g for each measurement. The UGV is assumed to have sufficient fuel for the entire trip, but the time cost must be minimized. If τ_g is a UGV tour with N measurement locations, then the cost of the tour is given as $\mathrm{len}(\tau_g) + N \cdot C_g$. Our second optimization problem is then,

Problem 2 (UGV cost): Given the set of points visited by the UAV and a radius associated with each point, find a UGV tour of minimum cost that obtains at least one measurement within the radius of each point.

We now review the related work in this area.

III. Related Work

The problem of designing sensor trajectories (or the related problem of selecting sensor locations) has recently received much attention. Low et al. [6] presented a control law to minimize the probability of misclassification in a Gaussian Process. The authors enforce measurements to be taken continuously, and sensors to only move along a 4-connected grid. Zhang and Sukhatme [7] presented an adaptive search algorithm for finding the optimal sensor path to estimate a scalar field. Song et al. [8] presented an algorithm to localize multiple radio sources using a mobile robot. They presented upper bounds on the time required to localize the sources up to a desired probability. In all these works, the sensing model is assumed to be continuous (i.e. no time cost), unlike our work where we penalize measurements explicitly.

Instead of labeling certainty, Krause et al. proposed Mutual Information as a measure of uncertainty [9]. An algorithm to place sensing locations was given which can closely approximate the optimal increase in Mutual Information. The work was extended to mobile sensor routing in [10], and multiple robots in [11]. Since we are designing algorithms for a heterogeneous sensor network, and use different objective function, these results are not directly applicable.

The problem of finding a tour to maximize the number of points visited (rewards) subject to a budget is known as the *orienteering problem*. Given a weighted graph $G = \{E, V\}$, the orienteering problem is to find a path such that the sum of the weights of the edges traversed is less than the budget, and the sum of the weights of the visited vertices is maximized. Blum et al. [12] presented a 4-approximation to the orienteering problem for complete graphs with metric edges. We show in Section V how to model the problem of selecting most PML points as an orienteering problem on a complete graph with metric edges.

Once the subset of PML points are found, the goal is to find a UGV tour to visit them with the least cost. The classical problem of visiting a set of sites on the plane with the shortest length path is known as the Traveling Salesperson Problem (TSP). Relevant to our application is the variant of TSP, known as TSP with Neighborhoods (TSPN), where the robot need not visit each site exactly but may instead visit any point in each site's neighborhood (subset of the plane). Dumitrescu and Mitchell [13] presented an 11.15-approximation algorithm when the neighborhoods are possibly-overlapping unit disks centered at each site. One of the main differences in our problem and the standard formulation of TSPN is that our cost is not just the traveling time of the tour, but also the total time taken for obtaining soil measurements. As we will discuss in Section V, finding a minimum length/time path does not necessarily ensure that the robot takes fewer soil measurements, and the cost for Problem 2 is not necessarily minimized.

Bhadauria et al. [14] studied the problem of computing a minimum time data collection tour for k robots tasked with wirelessly collecting data from deployed sensors by visiting a point in the sensor's communication range. In their model, robots spends time for both traveling and downloading data from robots. In this problem, the robot has to separately query each sensor whereas in our model, the robot can combine soil measurements for multiple points by sampling the intersection of their neighborhoods.

In [15], Alt et al. studied the problem of covering a given set of points with k radio antenna with circular ranges, where the algorithm has to choose the center and radius r_i for each circle. They consider a cost function which is a weighted sum of the length of the tour and the sum of r_i^α for each disk (α models the transmission power for the antennas). The main difference between this problem formulation and ours is that we do not require the number of samples (i.e. k) to be fixed, and instead penalize higher k in the cost function.

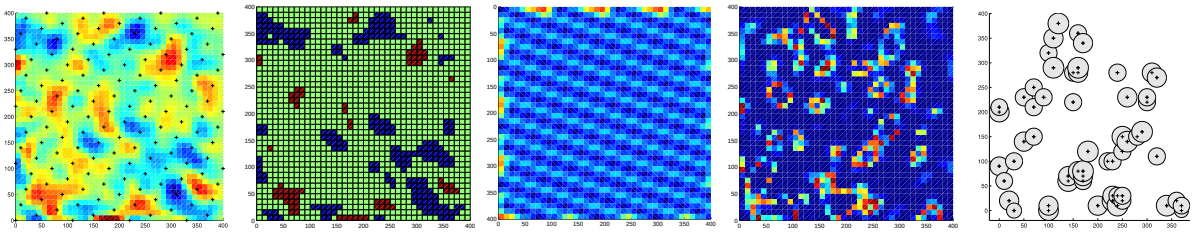
Recently, there has been a significant interest in developing cooperative aerial and ground/surface/underwater robot systems. Grocholsky et al. [16] described a system with coordinating aerial and ground vehicles for the application of detecting and locating targets. Sujit and Saripalli [17] studied the problem of exploring an area to detect targets using an UAV and inspecting the targets with Autonomous Underwater Vehicles (AUV). The authors compared in simulations three strategies to address the trade-off between quickly exploring the environment for all targets, and minimizing the latency between detection with UAVs and inspection with AUVs. Tanner [18] presented control laws for the UGVs to form a grid of sensors and UAVs to fly in a formation over the grid, such that a target moving on the ground can be detected if it moves from one grid cell to the other.

The main difference between existing literature and our work is that we explicitly consider that the UAV can be carried between takeoff locations by the UGV in the sensor planning phase. The resulting plan found by our algorithm may consist of multiple deployments for the UAV, which increases its coverage with limited battery.

IV. Finding Potentially Mislabeled Points

In this section, we first show how to find points with unacceptably high mislabel probability using the N level map. Then, we find the maximum variance any point can have and still be labeled with desired certainty. These results directly yield a set of PML points and a radius within which it is sufficient to obtain a measurement to improve its mislabel probability below the desired threshold.

We use a Gaussian Process (GP) to predict the mean and covariance of the N levels at any point. A Gaussian Process assumes all points are jointly-Gaussian, with a function $K(a, b, \theta)$ to determine the covariance between points a and b, with hyperparameters θ. A prior estimate is required, which in our application is commonly available from a satellite image. Let the points at which prior measurements are obtained be $\mathcal{X}_p$, with estimated N levels $N(\mathcal{X}_p)$. Let μ_p and σ_p be the mean and variance of the set of prior measurements. We have the following equations for mean



(a) (b) (c) (d) (e)

Fig. 2. A generated random field using the GP parameters learned from the soil dataset. (a) The ground-truth data. (b) A GP hypothesis was constructed from a sparse sampling of the data in (a), and the data was partitioned into three labels: low, med, high. (c) The variance of the sampling. (d) The mislabel probability. Note that it is high in many places, even though the variance is roughly uniform and low since the mislabel probability also depends on the value $N(x)$. (e) The points at which the labeling certainty is below P_d, and the corresponding rings described in Lemma 2.

and variance at any specific point set as [5],

$$N(x_i) = \mu_p + K_\theta(x_i, \mathcal{X}_p) K_\theta(\mathcal{X}_p, \mathcal{X}_p)^{-1} \left(N(\mathcal{X}_p) - \mu_p\right)$$
$$\sigma(x_i)^2 = \sigma_p^2 - K_\theta(x_i, \mathcal{X}_p) K_\theta(\mathcal{X}_p, \mathcal{X}_p)^{-1} K_\theta(x_i, \mathcal{X}_p)^T \quad (2)$$

where $K_\theta(a, b)$ is the covariance function used. In agriculture literature, these equations appear in slightly different form, known as geostatistical equations [19]. We will henceforth consider the squared exponential covariance function, but our method generalizes to other functions as well. Then,

$$K_\theta(a, b) = \sigma_f^2 \cdot e^{\left(\frac{1}{2l^2}||a-b||_2^2\right)} + \delta_{a,b}\sigma_n^2 \quad (3)$$

where $\delta_{a,b} = 1$ if $a = b$, and zero otherwise. We will assume the Gaussian Process has been trained, and the hyperparameters $\theta = \{l, \sigma_n, \sigma_f\}$ are known.

Let l_j be the current label at point x_i. Let $\Phi(a)$ be the standard Gaussian Cumulative Distribution Function evaluated at a. Also, we will define the term $l_j^\star$ to be the "closest" labeling threshold, for each label, to the current nitrogen estimate at point x_i as follows,

$$l_j^\star(x_i) = \min\left(|N(x_i) - l_j^-|, |N(x_i) - l_j^+|\right) \quad (4)$$

Due to the Gaussian noise assumption, we can find an upper bound on the probability the point is mislabeled as follows.

Lemma 1 (Mislabel Probability): A point x_i, with label l_j has a likelihood of being mislabeled P_{ml}, which satisfies,

$$\mathrm{P}_{ml}(x_i) \leq 2 \cdot \Phi\left(\frac{-l_j^\star}{\sigma(x)}\right) \quad (5)$$

where $l_j^\star$ is defined as the closest labeling boundary to $N(x)$.

Proof: We are trying to bound the "two-tail" probability mass of the Gaussian random variable x_i. We will use the following relationships. For any $0 < a \leq b$, and $y \sim \mathcal{N}(\mu, \sigma)$,

$$\mathrm{P}(y > b) + \mathrm{P}(y < -b) \leq \mathrm{P}(y > a) + \mathrm{P}(y < -a)$$
$$\mathrm{P}(y > b) + \mathrm{P}(y < -b) = 2 \cdot \Phi(\frac{\mu - b}{\sigma})$$

Then it follows that,

$$\mathrm{P}_{ml}(x_i) = \mathrm{P}(N(x_i) < l_j^-) + \mathrm{P}(N(x_i) > l_j^+)$$
$$\mathrm{P}_{ml}(x_i) \leq 2 \cdot \max[\mathrm{P}(N(x_i) < l_j^-), \mathrm{P}(N(x_i) > l_j^+)]$$
$$\mathrm{P}_{ml}(x_i) \leq 2 \cdot \mathrm{P}(N(x_i) < -|l^\star|) = \Phi\left(\frac{-l_j^\star}{\sigma(x)}\right)$$

■

From this lemma we know the mislabel probability is a function of both the current estimate $N(x)$ and the variance $\sigma(x_i)$. Of these quantities, we have control over $\sigma(x_i)$, since it depends on the distance of measurement locations from x_i. We now show how close to x_i a measurement should be to have enough impact on the mislabel probability. First, we show how low the variance must be for the current value of $N(x)$ to be labeled with high probability.

Corollary 1 (Desired Variance): For any point, $x_i \in \mathcal{X}$, let $l_j^\star(x_i)$ be the closest label boundary to $N(x)$, as defined in Equation 4. If the variance in the estimate is less than

$$\sigma(x_i) \leq \frac{l_j^\star}{\Phi^{-1}(1 - \frac{1}{2}P_d)} \quad (6)$$

then the labeling certainty is greater than P_d.

Due to the spatial correlation from Equation 2, we can find the maximum distance from x_i for a measurement to have the desired uncertainty reduction to satisfy Equation 1.

Lemma 2 (Measurement Radius): Let x_i be a point with high label uncertainty. For a sensor S with measurement error σ_s, we can take a measurement from any point z_S satisfying

$$||z_S - x_i||_2^2 \leq -l^2 \log_e[(\sigma^2(x) - \sigma_s^2)(\sigma_f^2 + \sigma_s^2)\sigma_f^{-4})] \quad (7)$$

to reduce $\mathrm{P}_{ml}(x_i)$ to below P_d.

Proof: The proof follows by manipulating the covariance function to find a desired reduction in variance. From the definition of the Gaussian Process covariance (Equation 2), we have that,

$$\sigma_d^2 = \sigma(x)^2 - K(x, z)[K(z, z) + \sigma_n^2]^{-1} K(x, z)^T$$

Where σ_d^2 is the variance from Corollary 1 and $\sigma(x)^2$ is the current variance at the target PML point. The function $K(\cdot, \cdot)$ is the covariance function given by Equation 3. Because we use a single measurement point for the vector z, and only a single point for x, the covariance of the measurement point is simply $\sigma_f^2 + \sigma_n^2$. Substituting and re-arranging terms we have the following.

$$\sigma(x)^2 - \sigma_d^2 = K(x_i, z_S)[K(z_S, z_S) + \sigma_n^2]^{-1} K(x_i, z_S)^T$$
$$\sigma(x)^2 - \sigma_d^2 = \frac{\sigma_f^4}{\sigma_f^2 + \sigma_n^2} e^{-\frac{1}{l^2}||x - z_S||_2^2}$$

The desired result follows by taking the natural log and solving for the distance. ■

This lemma can be easily extended to handle measurements from two types of sensors (aerial and ground).

Thus, for every PML point (which does not satisfy Equation 6), we can find a maximum distance from the point, as a function of the sensor noise and current uncertainty, from which a sample must be obtained to obtain sufficiently small variance on the point x_i. An example of a field, the field labels, and the points with high mislabel probability are shown in Figure 2. For each point, a measurement from inside the specified radius will satisfy Equation 1, by the previous lemma. Next we present algorithms to find a tour of these measurement locations.

V. Path Optimization

In this section, we first describe our algorithm for finding the UAV+UGV tours to visit the most number of PML points subject to the UAV battery budget. Our algorithm operates in two phases: we first find the subset of PML points to be visited by the UAV (refer to Problem 1). Then we find the tour (i.e. order and sampling locations) for the UGV (refer to Problem 2). The tour for the UGV also includes take-off and landing locations for the UAV. We present the description and analysis for each of these two phases next.

A. Computing the UAV Tour

We reduce the problem of finding the maximum subset of PML points to be visited by the UAV to the orienteering problem. We begin with a graph $G(V, E, \pi, w)$ with weights ($w(u, v)$) on edges, and rewards $\pi(v)$ on the vertices. The objective in the orienteering problem is to find a tour of a subset of vertices collecting maximum reward, with the constraint that the sum of weights of edges on the tour is less than a given budget. In the following, we show how to create such a graph for our problem.

First consider the simpler case of finding the maximum subset of points in a UAV-only system. For simplicity, let the camera footprint be a single point for now. The UAV tour will consist of a single path with one take-off and landing location. We build a graph with PML points as the vertices. We add an edge to G between every pair of points with weight equal to the Euclidean distance between the points. Each vertex is given a unit reward. The budget for the UAV equals the battery lifetime minus $2C_a$ to account for the single takeoff and landing. The solution for the orienteering problem for this instance will be a path traversing a set of PML points. The length of the path will be less than or equal to the available travel budget for the UAV. The reward equals the total number of PML points visited by the path.

Since the edge weights are Euclidean distances, this graph is a complete metric graph. Blum et al. [12] presented a 4-approximation for orienteering problems on undirected metric graphs. Applying this algorithm to the graph we constructed above will yield a UAV tour visiting at least $1/4^{\text{th}}$ of the PML points visited by the optimal algorithm.

Now consider the case of a UAV+UGV system. The UGV can transport the UAV between two PML locations, without affecting the UAV's battery life. We will modify the edge weights accordingly. Furthermore, since the UAV carries a camera with a footprint of diameter C, it can sample a point without flying directly over it. Hence, we will also modify the set of vertices. The detailed construction of the input graph for the orienteering problem is as follows.

(1) Create a square grid of resolution $C/\sqrt{2}$ over the plane. Each point in $\mathcal{X}_{pml}$ is associated with its nearest grid location (Figure 3(a)). Store the number (denoted by $\pi(v)$) of PML points associated with a grid location.
(2) Let V be the set of grid vertices with at least one PML point associated. For each $v \in V$, let $\pi(v)$ be the number of associated PML points (Figure 3(a)).
(3) Build a complete undirected graph $G = \{V, E, \pi, w\}$. For each edge between $(u, v) \in V$, add a weight $w(u, v) = \min\{d(u, v), 2C_a\}$. This implies there are two types of edges between grid points: The UAV can either use the UGV to travel paying only for the ascent/descent ($2C_A$) or travel directly between points paying the distance cost ($d(u, v)$).

Before proceeding, we verify G is still a metric graph. Consider a triple of vertices u, v, w. We know $w(u, v), w(v, w), w(w, u) \leq 2C_a$. It is easy to see the triangle inequality holds when two or three edges have weights equal to $2C_a$. Consider the case when only one edge has weight equal to $2C_a$, say $w(u, v) = 2C_a$. Now, $w(v, w) + w(w, u) = d(v, w) + d(w, u) \geq d(u, v)$. Since $w(u, v) = \min\{2C_a, d(u, v)\} = 2C_a$, we have $d(u, v) \geq 2C_a$. Hence, $w(v, w) + w(w, u) \geq w(u, v)$. And since $w(u, v) = 2C_a$ and $w(v, w), w(w, u) < 2C_a$, $w(u, v) + w(w, u) \geq w(v, w)$ and $w(u, v) + w(v, w) \geq w(w, u)$. For the case when all three edges have weights less than $2C_a$, the weights are equal to Euclidean distances. Hence, weights satisfy triangle inequality in addition to symmetry, identity and non-negativity. Hence, the graph constructed above is a complete metric graph. We can then apply the algorithm in [12] to obtain a 4-approximation to Problem 1.

B. Sampling TSPN

The algorithm in the previous subsection gives us the subset of PML points to be sampled. We must now construct a minimum cost tour for the UGV to visit each point in the subset. Note, for each point we have a radius given by Lemma 2 within which we must obtain a sample. The radius is not necessarily the same for various points, however, in practice, we have found the radii are comparable. Hence, we use the minimum radius (r) to simplify the algorithm.

We must now find a tour for the UGV which visits all disks and takes a measurement within it. In the standard formulation for the Traveling Salesperson with Neighborhoods (for uniform disks) [13] only the travel time is considered. However, recall from Problem 2, the cost of the UGV tour equals the sum of time spent for traveling and the time spent for obtaining all measurements. Thus the UGV must choose

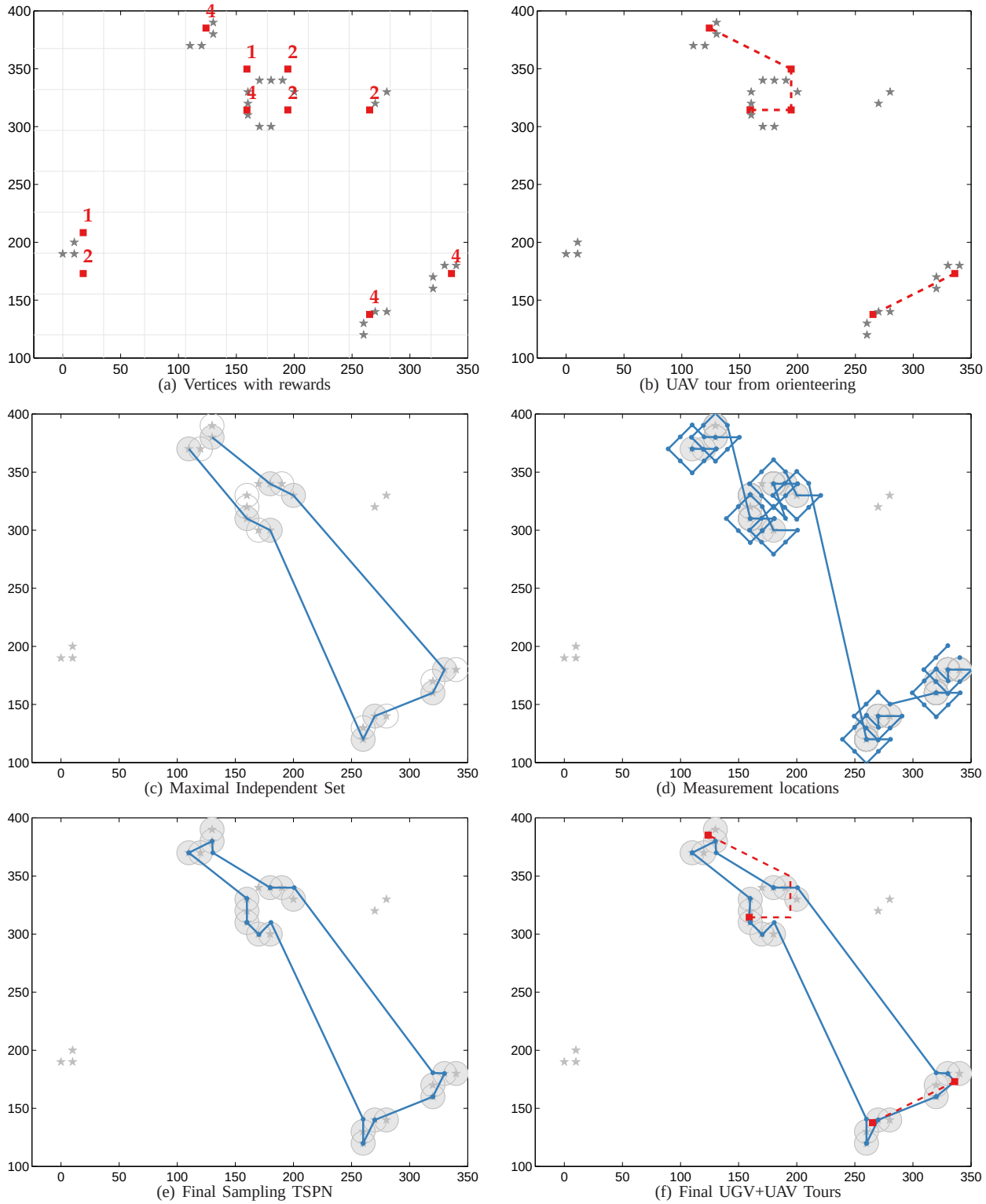


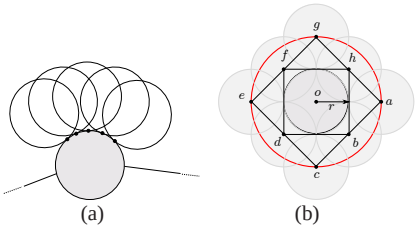
Fig. 3. Path Planning Algorithm. (a) Square grid of resolution $C/\sqrt{2}$. The reward for visiting each grid point (red square) is the number of PML points (gray star) falling within the grid. (b) UAV tour found using orienteering on the graph of grid points. For this instance, UAV budget was 500 secs out of which 200 secs are spent traveling and 240 secs are spent for the 2 ascents/descents. (c), (d), (e) Steps in constructing a Sampling TSPN tour (Algorithm 1) for the UGV. (f) Final UGV tour including UAV take-off locations (red squares).

a measurement location within each disk, while the length of the tour and the number of distinct measurement locations are minimized. Specifically, we are studying the following new variant of the Traveling Salesperson Problem.

Problem 3 (Sampling TSPN): Given a set of disks with uniform radius r and centers at points X, find a tour τ of N

distinct sample locations to minimize the cost $\text{len}(\tau) + C \cdot N$ such that each disk contains a sample location.

We will present a brief description of the algorithm in [13] for the standard TSPN problem and show how to modify it to accommodate the additional cost of measurements. Dumitrescu and Mitchell in [13] first find a maximal independent set (MIS) of non-intersecting disks, from the given set of disks. Then they find a TSP tour visiting the centers of the disks in MIS (Figure 3(c)). This tour intersects the circumference of each disk in the MIS twice. The final tour that visits each disk is constructed as follows. Start from an arbitrary disk in the MIS. Follow the tour in the clockwise order up to the intersection point of the next disk. Follow the circumference of this disk in the clockwise order up to the second intersection point of the same disk. Continue until the starting point is encountered again. Repeat this process in the anti-clockwise direction.



(a) (b)

Fig. 4. (a) The standard TSPN algorithm [13] that visits just the circumference of the disk in the MIS (shown shaded) can be arbitrarily bad when applied to the Sampling TSPN problem. This tour will be forced to take one measurement each for each outer disk and thus have $O(n)$ measurement locations. The optimal algorithm is not forced to move along the circumference. Thus, it can visit a small number of locations where the disks overlap. (b) We modify the TSPN heuristic to visit a fixed number of sites (shown as dots) around each disk in the MIS yielding a constant factor approximation for the Sampling TSPN problem.

This algorithm ensures that the tour intersects all disks by moving along the circumference of each disk in the MIS. However, when the UGV has to stop and obtain a measurement in each disk, restricting the motion to the circumference can be potentially costly in terms of the number of measurements. For example, consider the situation shown in Figure 4(a) where the outer disks intersect with each other only outside the circumference of the disk in MIS. Since the tour in the standard algorithm only moves along the circumference, it will be forced to take one measurement for every outer disk, which can be as large as $O(n)$, where as the optimal algorithm can visit a small number of intersection points, not necessarily on the circumference.

We modify the local strategy (going around the circumference) in order to simultaneously find sampling locations for the tour. We will bound the additional length due to this new local strategy, and bound the number of samples obtained with respect to the optimal. Our algorithm is as follows:

(1) Find an MIS of non-overlapping disks from input disks.
(2) Starting from an arbitrary point, follow the TSP tour of the centers of disks in the MIS (Figure 3(c)).
(3) When a new center (say x) is reached on the TSP tour, before visiting the next center on the tour, first visit the eight sites as shown in Figure 4(b). o is the center of the disk in the MIS. Sites b, d, h, f lie on a square of side $2r$ centered at o. Sites a, c, e, g lie on a square of side $2\sqrt{2}r$ rotated by $\frac{\pi}{4}$. After visiting the last of these sites, we continue towards the center of the next disk in MIS (Figure 3(d)).
(4) Restrict the candidate sampling locations to the set of centers of disks in MIS and the set of eight sites as described above. Denote this set of candidate sampling locations by S.

Algorithm 1: Sampling TSPN Algorithm.

First we will show that the set S defined above intersects all disks in X. Then we will bound the size of S with respect to the size of the optimal number of samples N^*. Finally, we will bound the length of the tour and thus the total cost.

Lemma 3: Let S be the set of all centers of the disks in a MIS of non-overlapping disks X of radius r. Let S also contain the eight sites as described in Algorithm 1. Then, for each disk in X, there exists a point in S lying in its interior.

Proof: Consider Figure 4(b). o is the center of the disk in the MIS. Then all disks that intersect the disk at o, lie within a disk of radius $2r$ from o (shown in red in Figure 4(b)). Denote this outer disk by D_2. Let $S_o = \{o, a, \ldots, h\}$. We can observe that the set of disks of radius r centered at all sites in S_o form a cover of D_2. Hence, any point in D_2 is at a distance of at most r from a site in S_o. Now consider any disk centered at o' that intersects the disk centered at o. $o' \in D_2$, hence there exists a site in S_o at a distance r from o', i.e., lying within the disk at o'. This proves the lemma. ■

Next, we bound the size of S with respect to the number of samples in an optimal Sampling TSPN algorithm. We also show that the total length of our tour is no more than a constant times the length of the tour of an optimal algorithm for the sampling TSPN algorithm.

Lemma 4: If N^* is the number of samples by an optimal algorithm for the Sampling TSPN problem, then $|S| \leq 9N^*$.

Proof: S contains 9 sites per disk in an MIS of non-intersecting disks: one for the center, and eight lying on the two squares as shown in Figure 4(b). Hence, $|S| = 9|\text{MIS}|$. No two disks in the MIS overlap, hence they cannot share any sampling location between them. Hence, $N^* > |MIS|$ and thus $|S| \leq 9N^*$. ■

Lemma 5: Let T_{ALG} be the tour constructed by the algorithm above, and T^* be the tour for the optimal Sampling TSPN algorithm. Then $\text{len}(T_{ALG}) \leq 44T^*$.

Proof: For ease of notation, in this proof we refer both a tour and its length by T, and T^* refers to an optimal tour. The analysis of this proof proceeds similar to that in [13] with modifications where the tour differs. Using Proposition 1 in [13] for non-unit disks, we have $n \leq \left(4 + \frac{4T^*}{\pi r}\right)$.

Denote by T_I and T_C the TSPN tour of the MIS and TSP tour of the center of the MIS respectively. Let n be the total

number of input disks. Now

$$\begin{aligned} T_C^* &\leq T_I^* + 2nr \\ &\leq T^* + 8r + \frac{8T^*}{\pi} \\ &= T^* \left(1 + \frac{8}{\pi}\right) + 8r. \end{aligned}$$

The first inequality follows from the fact that a tour of the centers can be constructed by taking a detour of at most $2r$ for each disk from the tour of the disks. T_{ALG} consists of a TSP tour of the centers of the disks in MIS and a tour of the regular hexagon surrounding each disk. Using the $(1+\epsilon)$-approximation for the TSP tour [20], we get

$$\begin{aligned} T_{ALG} &\leq (1+\epsilon)T_C^* + \sqrt{2}nr + 7\sqrt{2}nr, \\ &\leq (1+\epsilon)\left(\left(1+\frac{8}{\pi}\right)T^* + 8r\right) + 32\sqrt{2}r + \frac{32\sqrt{2}T^*}{\pi}. \end{aligned}$$

The $\sqrt{2}nr$ term in the first inequality comes from the part of moving from the center of the disk in MIS to the closest of the eight sites (from o to h in Figure 4(b)). The $7\sqrt{2}nr$ comes from visiting each of the eight sites (starting from h through a in Figure 4(b)).

Using Lemma 1 from [21], we know that the length of any path that visits three non-overlapping disks of radius r is at least $0.4786r$. Thus when the MIS contains more than 3 disks, we get $T^* \geq 0.4786r$. Therefore,

$$\begin{aligned} T_{ALG} &\leq T^* \left[(1+\epsilon)\left(1+\frac{8}{\pi}\right) + \frac{32\sqrt{2}}{\pi} + 0.47(8+32\sqrt{2}+8\epsilon)\right] \\ &\leq 44T^* \end{aligned}$$

for small ϵ. ■

Based on the above lemmas, we can easily bound the total cost of our algorithm and show that it is at most a constant times that of the optimal cost.

Theorem 1: Algorithm 1 is a valid Sampling TSPN tour with cost at most 44 times that of the optimal algorithm.

Proof: Let C^* be the cost of the optimal algorithm for the Sampling TSPN problem. Therefore, $C^* \geq \text{len}(\tau^*) + N^* \cdot C_g$, where τ^* is the optimal TSPN tour visiting all disks, and N^* is the minimum number of sample locations such that each disk has at least one sample location.

Consider the cost of our algorithm,

$$\begin{aligned} C_{ALG} &= \text{len}(\tau) + N \cdot C_g, \\ &\leq 44\text{len}(\tau^*) + 9N^* \cdot C_g, \\ &\leq 44C^*. \end{aligned}$$

where the first inequality comes from Lemmas 5 and 4. ■

In practice, we do not have to sample at all the nine sites. We can discard sample locations that do not intersect any disk or find a smaller subset (e.g. greedily) that samples all the disks. Then, we can find a TSP tour of just these sampling locations, as shown in Figure 3(e). We have not yet considered the UAV landing and take-off locations in the UGV tour. We add all the take-off locations to S, and find a TSP tour of the combined set of points (Figure 3(f)). Then, after visiting a take-off location, the UGV continues along its tour. Whenever, its time to reach the landing location from the tour becomes equal to the landing time of the UAV, the UGV can deviate from its pre-planned route, visit the landing location, and return to its pre-planned route. The total distance overhead for the UGV is at most the budget of the UAV and in practice, not very significant.

VI. Simulations

In the previous sections, we showed theoretical bounds on the number of PML points selected and the distance traveled by our algorithm with respect to optimal. We expect the UAV+UGV system to sample more PML points as compared to a UAV only system with the same battery limitations. We explore this through simulations using actual system parameters and real data collected from an agricultural plot.

A. System Description

We present the details of the robotic system we are developing to motivate the choice of our simulation parameters. Our UGV is a Husky A200 by Clearpath Robotics [22]. The UGV has a typical battery life of two hours on a single charge. The operating lifetime can be extended to over six hours easily with additional batteries. The UGV will measure soil organic matter as a proxy for soil N supply to the crop using a Minolta SPAD-502 Chlorophyll meter [23].

Our UAV is a Hexa XL by MikroKopter [24]. This UAV can operate for 25 mins. Deploying the UAV to approximately 100 meters height gives the camera a 50 meter diameter coverage with a single image. The UAV takes about 2 minutes to ascend/descend this height. The images include multi-spectral information, such as near-infrared reflectance, which is used to estimate the crop N status [25].

B. Modeling

To generate realistic data, we need a generative model of Nitrogen levels and realistic values for the sampling noise for both systems. We will briefly discuss how we obtained these from a nitrogen remote sensing and soil sampling dataset [25]. The data consists of 1375 soil measurements taken manually in a 50m by 250m corn field, along with corresponding 1m spatial resolution remote sensing images in the green (G), red (R) and near infrared (NIR) portions of the spectrum. The samples were taken along a dense uniform coverage (See Figure 5) and provided the levels of soil organic matter (OM). R and NIR are known to be inversely related to crop N status [25].

We used OM as a proxy for the initial quantity of soil N supplied to the crop. We modeled the UGV as taking direct measurements of OM, corrupted by some sensor noise σ_g, and the UAV as measuring the Normalized Difference Vegetation Index (NDVI), which is a combination of NIR and R levels [26]. We assume the NDVI levels are corrupted by sensor noise σ_a. To model the spatial patterns of the OM levels, we used GP regression over the set of sample points and OM measurement values. This densely-sampled GP defined the hyperparameters which were used to generate

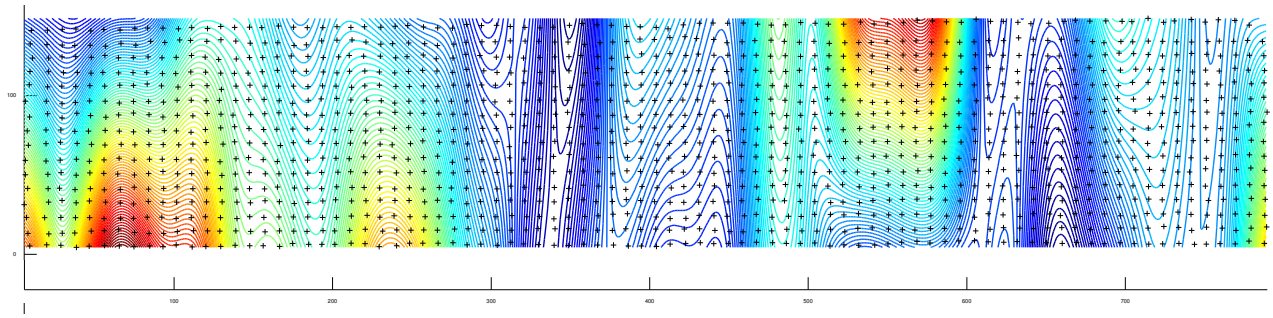


Fig. 5. Soil organic matter data set from [25]. Dense sampling was collected by hand (black crosses) and used to train a Gaussian Process. The resulting estimate of nitrogen levels is shown as the contour map. From this data set we learn the sensor noise values σ_a and σ_g, as well as model the underlying soil organic matter for larger simulations (Figure best viewed in color).

new ground-truth N maps in our simulations. We used the GPML Toolbox [27] for performing the GP regression.

As part of the ground-truth GP regression, we can estimate the sample noise at each point from the data directly (σ_n in Equation 3). We used this value directly as σ_g, since we assumed the robot would have the same sensing capability as the human operators. To estimate σ_a, we built the covariance matrix, $\mathbb{E}[(\text{NDVI} - \mathbb{E}(\text{NDVI}))(\text{OM} - \mathbb{E}(\text{OM}))^T]$. This gives us a 2×2 matrix, $\begin{bmatrix} \sigma^2_{\text{OM}} & \sigma_{\text{OM,NDVI}} \\ \sigma_{\text{NDVI,OM}} & \sigma^2_{\text{NDVI}} \end{bmatrix}$.

From this, we can find the variance in OM given a measurement of NDVI as, $\sigma^2_{\text{OM|NDVI}} = \sigma^2_{\text{NDVI}} - \frac{\sigma^2_{\text{OM,NDVI}}}{\sigma^2_{\text{OM}}}$. We found that $\sigma_{\text{OM|NDVI}} > \sigma_g$. We can now model NDVI measurements directly from measurements of OM when sampling the Gaussian Process model. These UAV and UGV measurement noise variance were found to be $\sigma_a = 0.31$ and $\sigma_g = .05$ respectively, for the dataset.

For the simulations, we formed a prior estimate of OM levels by down sampling each randomly-generated ground-truth N-map by a factor of 20 and fitting a new GP. From this prior GP, we found the PML point set as described in Section IV. We randomly generated 100 N-maps for a 600×400 m field. We then found the PML points using a desired labeling probability of 0.6. The number of PML points in any instance depends on the randomly generated map.

C. Results

We first compare the number of PML points covered by the UAV+UGV system versus an UAV-only system. We use the procedure described in Section V for finding the subset of PML points visited by the UAV only and the UAV+UGV system, subject to the battery constraint of 25 mins. We used the implementation from the SFO Toolbox [28] for finding an orienteering tour, and the Concorde TSP solver [29] as a subroutine in the Sampling TSPN algorithm implementation.

Figure 6 shows a sample run from the simulations. We observe that the UAV-only tour is constrained to only one part of the field, whereas the UAV+UGV system can obtain measurements from farther away locations. This input instance consisted of 75 potentially mislabeled points, the UAV only tour covers 38 points whereas the UAV+UGV tour covers 50 points. Figure 7 shows a histogram of the ratio of the points covered by the UAV+UGV and the UAV only tours for 100 random instances. As expected, the ratio is always greater than 1 as the UAV+UGV system is at least as good as a UAV only system in terms of the number of points visited. Table I shows the effect of varying the budget on the percentage of input PML points visited.

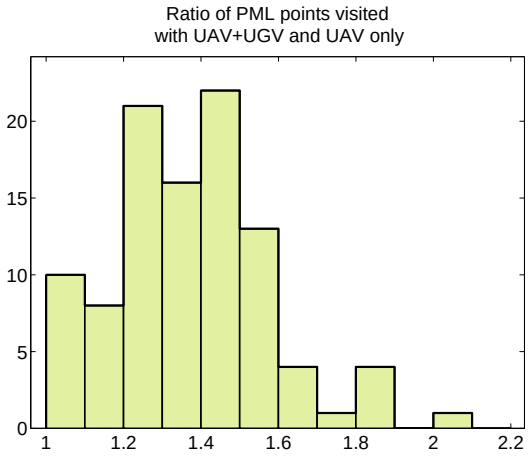


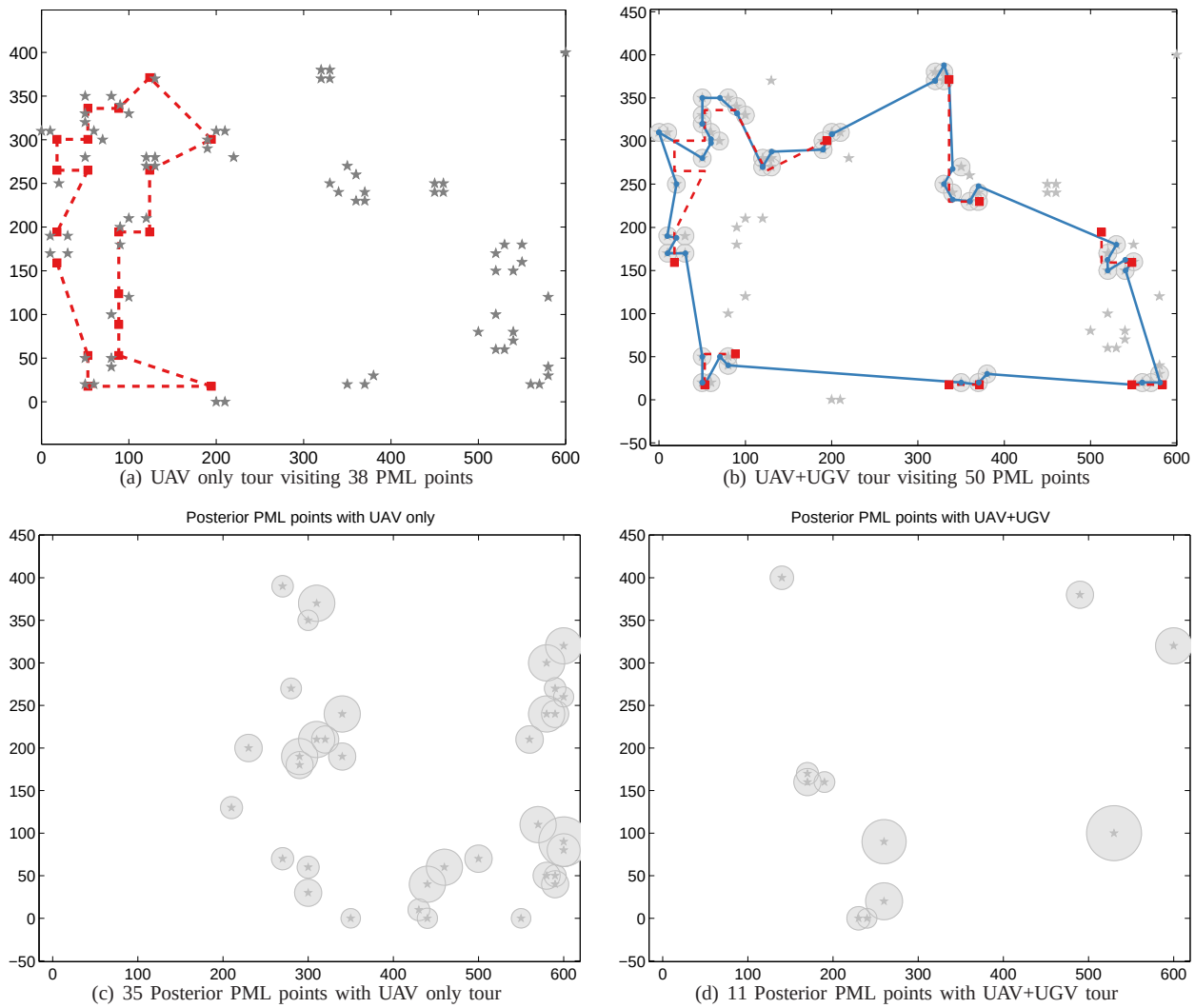
Fig. 7. Histograms of the ratio of number of PML points visited for a UAV+UGV system and a UAV only system for 100 random instances. Both systems are given an equal budget of 25 minutes.

TABLE I

PERCENTAGE OF INPUT PML POINTS VISITED (AVG. OF 30 INSTANCES).

Budget (sec)	UAV only	UAV+UGV
500	19	25
1000	36	49
1500	55	72

The UAV+UGV system can cover points that are spread across the field. Intuitively, if the measurements are distributed across the field, we expect the resulting map (after incorporating the measurements) to have fewer mislabeled points than if all measurements are nearby. After calculating the desired UAV/UGV tours, random measurements for the sensors were sampled directly from OM values given the dense (ground truth) GP. We added noise to the measurements using estimated variances $\sigma_a = 0.31$ and $\sigma_g = 0.5$ as described in Section VI-B. These values were then used to update the prior GP, which was then used to find the posterior PML points. We observe the posterior PML points



(a) UAV only tour visiting 38 PML points

(b) UAV+UGV tour visiting 50 PML points

(c) 35 Posterior PML points with UAV only tour

(d) 11 Posterior PML points with UAV+UGV tour

Fig. 6. Sample simulation instance. (a) & (b) shows the tours found using a UAV only and UAV+UGV system. The input consists of 75 PML points. The UAV+UGV tour consists of 6 subtours. (c) & (d) shows the PML points found in the updated N level map after incorporating aerial and ground measurements. The UGV allows the UAV to transport to farther locations in the plot which is reflected in fewer posterior PML points.

in Figures 6(c) & 6(d). For a fair comparison, we add UGV measurements for each PML point visited by a UAV only tour, in obtaining the updated N level map.

Figure 8 shows a histogram of the ratio of the posterior PML points with a UAV+UGV system and a UAV only system. Since the number of PML points depend on both the variance, and the estimated $N(x)$ values, occasionally there are instances when the number of posterior PML points with UAV only system are lesser than that of UAV+UGV system. However, as we can observe in Figure 8 the UAV+UGV system often outperforms the UAV only system in terms of number of posterior PML points.

VII. Conclusion

In this paper, we studied the problem of designing sensing strategies for obtaining aerial images and soil samples with a UAV+UGV system to estimate the Nitrogen level in a plot. Effective fertilizer treatment plans can be developed by better estimates of Nitrogen levels. Since the battery life of the UAV is limited, the UAV and UGV can only sample a limited number of points. We studied the problem of maximizing the number of points visited by the UAV and UGV. Unlike traditional approaches, our algorithm takes into consideration the situation where the UAV can land on the UGV and thus be carried between points without expending energy. We also studied the problem of minimizing the time for sampling these points with a UGV. We presented a constant-factor approximation algorithm which finds a set of sampling locations and a tour of these locations, such that each point has a sampling location within its disk neighborhood.

An interesting direction for future work is bounding the number of tours required to correctly label the whole field. What separates this from coverage problems is the fact that new PML points are likely to appear during sampling tours, since N estimates change after every sample. This presents an interesting trade-off between using the slower UGV measurements to handle the new PML points, or returning and re-planning with a full UAV budget.

We have also started building the complete system using a

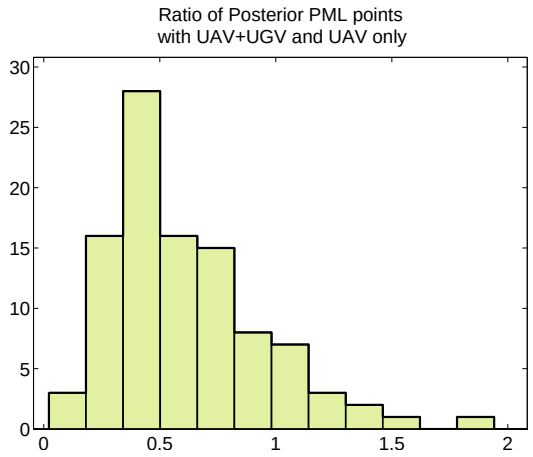


Fig. 8. Histograms of the ratio of number of posterior PML points generated after updating N map with simulated measurements for a UAV+UGV system and a UAV only system for 100 random instances. Both systems are given an equal budget of 25 minutes.

Clearpath Husky A200 ground robot and a hexacopter from MikroKopter. In order to execute the algorithms presented in this paper, additional capabilities such as autonomous landing and soil sampling are necessary. We are committed to developing these capabilities and enabling the use of robots in precision agriculture.

ACKNOWLEDGMENT

This work is supported by NSF Awards #1111638, #0916209, #0917676, #0936710 and a fellowship from the Institute on the Environment at the University of Minnesota.

REFERENCES

[1] D. Mulla, "Mapping and managing spatial patterns in soil fertility and crop yield," in *Soil Specific Crop Management*. American Society of Agronomy, 1993, pp. 15–26.

[2] G. W. Randall and D. J. Mulla, "Nitrate nitrogen in surface waters as influenced by climatic conditions and agricultural practices," *Journal of Environmental Quality*, vol. 30, no. 2, pp. 337–344, 2001.

[3] M. S. Moran, Y. Inoue, and E. Barnes, "Opportunities and limitations for image-based remote sensing in precision crop management," *Remote Sensing of Environment*, vol. 61, no. 3, pp. 319–346, 1997.

[4] V. I. Adamchuk, J. Hummel, M. Morgan, and S. Upadhyaya, "On-the-go soil sensors for precision agriculture," *Computers and Electronics in Agriculture*, vol. 44, no. 1, pp. 71–91, 2004.

[5] C. E. Rasmussen and C. K. Williams, *Gaussian processes for machine learning*. MIT press Cambridge, MA, 2006, vol. 1.

[6] K. H. Low, J. Chen, J. M. Dolan, S. Chien, and D. R. Thompson, "Decentralized active robotic exploration and mapping for probabilistic field classification in environmental sensing," in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, 2012, pp. 105–112.

[7] B. Zhang and G. S. Sukhatme, "Adaptive Sampling for Estimating a Scalar Field using a Robotic Boat and a Sensor Network," *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pp. 3673–3680, Apr. 2007.

[8] D. Song, C.-Y. Kim, and J. Yi, "Simultaneous localization of multiple unknown and transient radio sources using a mobile robot," *Robotics, IEEE Transactions on*, vol. 28, no. 3, pp. 668–680, 2012.

[9] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *The Journal of Machine Learning Research*, vol. 9, pp. 235–284, 2008.

[10] A. Meliou, A. Krause, C. Guestrin, and J. M. Hellerstein, "Nonmyopic informative path planning in spatio-temporal models," in *Proceedings of the 22nd national conference on Artificial intelligence-Volume 1*. AAAI Press, 2007, pp. 602–607.

[11] A. Singh, A. Krause, and W. J. Kaiser, "Nonmyopic Adaptive Informative Path Planning for Multiple Robots," *International Joint Conference on Artificial Intelligence*, pp. 1843–1850, 2008.

[12] A. Blum, S. Chawla, D. R. Karger, T. Lane, A. Meyerson, and M. Minkoff, "Approximation algorithms for orienteering and discounted-reward tsp," *SIAM Journal on Computing*, vol. 37, no. 2, pp. 653–670, 2007.

[13] A. Dumitrescu and J. S. Mitchell, "Approximation algorithms for tsp with neighborhoods in the plane," *Journal of Algorithms*, vol. 48, no. 1, pp. 135–159, 2003.

[14] D. Bhadauria, O. Tekdas, and V. Isler, "Robotic data mules for collecting data over sparse sensor fields," *Journal of Field Robotics*, vol. 28, no. 3, pp. 388–404, 2011.

[15] H. Alt, E. M. Arkin, H. Brönnimann, J. Erickson, S. P. Fekete, C. Knauer, J. Lenchner, J. S. Mitchell, and K. Whittlesey, "Minimum-cost coverage of point sets by disks," in *Proceedings of the twenty-second annual symposium on Computational geometry*. ACM, 2006, pp. 449–458.

[16] B. Grocholsky, J. Keller, V. Kumar, and G. Pappas, "Cooperative air and ground surveillance," *Robotics & Automation Magazine, IEEE*, vol. 13, no. 3, pp. 16–25, 2006.

[17] P. Sujit and S. Saripalli, "An empirical evaluation of co-ordination strategies for an auv and uav," *Journal of Intelligent & Robotic Systems*, vol. 70, pp. 373–384, 2013.

[18] H. G. Tanner, "Switched uav-ugv cooperation scheme for target detection," in *Robotics and Automation, 2007 IEEE International Conference on*. IEEE, 2007, pp. 3457–3462.

[19] D. E. Myers, "Estimation of linear combinations and co-kriging," *Mathematical Geology*, vol. 15, no. 5, pp. 633–637, 1983.

[20] J. S. Mitchell, "Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric tsp, k-mst, and related problems," *SIAM Journal on Computing*, vol. 28, no. 4, pp. 1298–1309, 1999.

[21] O. Tekdas, D. Bhadauria, and V. Isler, "Efficient data collection from wireless nodes under the two-ring communication model," *The International Journal of Robotics Research*, vol. 31, no. 6, pp. 774–784, 2012.

[22] "Clearpath robotics," 2013 (accessed Jan 2013). [Online]. Available: http://clearpathrobotics.com

[23] "Spectrum technologies inc." 2013 (accessed Mar 2013). [Online]. Available: http://www.specmeters.com/

[24] "Mikrokopter," 2013 (accessed Jan 2013). [Online]. Available: http://mikrokopter.de

[25] D. Mulla, M. Beatty, and A. Sekely, "Evaluation of remote sensing and targeted soil sampling for variable rate application of nitrogen," in *Proceedings of 5th International Conference on Precision Agriculture*. American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, 2001.

[26] D. J. Mulla, "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps," *Biosystems Engineering*, vol. 114, no. 4, pp. 358 – 371, 2013, special Issue: Sensing Technologies for Sustainable Agriculture.

[27] C. E. Rasmussen and H. Nickisch, "Gaussian processes for machine learning (gpml) toolbox," *The Journal of Machine Learning Research*, vol. 11, pp. 3011–3015, 2010.

[28] A. Krause, "Sfo: A toolbox for submodular function optimization," *The Journal of Machine Learning Research*, vol. 11, pp. 1141–1144, 2010.

[29] D. Applegate, R. Bixby, V. Chvatal, and W. Cook, "Concorde tsp solver," *URL http://www. tsp. gatech. edu/concorde*, 2006.