Adaptive sampling for environmental field estimation using robotic sensors

Mohammad Rahimi\(^1\), Mark Hansen\(^1,3\), William J. Kaiser\(^1,5\), Gaurav S. Sukhatme\(^1,2\), Deborah Estrin\(^1,4\)

\(^1\)Center for Embedded Networked Sensing, UCLA, Los Angeles, CA 90095
\(^2\)Department of Computer Science, USC, Los Angeles, CA 90089
\(^3\)Department of Statistics, UCLA, Los Angeles, CA 90095
\(^4\)Department of Computer Science, UCLA, Los Angeles, CA 90095
\(^5\)Department of Electrical Engineering, UCLA, Los Angeles, CA 90095

\{mhr@cens.ucla.edu, cocteau@stat.ucla.edu, kaiser@ee.ucla.edu, gaurav@robotics.usc.edu, destrin@cs.ucla.edu\}

Abstract—Monitoring environmental phenomena by distributed sensor sampling confronts the challenge of unpredictable variability in the spatial distribution of phenomena often coupled with demands for a high spatial sampling rate. The introduction of actuation-enabled robotics sensors permits a system to optimize the sampling distribution through runtime adaptation. However, such systems must efficiently dispense sampling points or otherwise suffer from poor temporal response. In this paper we propose and characterize an active modeling system. In our approach, as the robotic sensor acquires measurement samples of the environment, it builds a model of the phenomenon. Our algorithm is based on an incremental optimization process where the robot supports a continuous, iterative process of 1) collecting samples with maximal coverage in the design space, 2) building the environmental model 3) predicting sampling point locations that contribute the greatest certainty regarding the phenomenon 4) and sampling the environment based on a combined measure of information gain and navigation and sampling cost. This can provide significant reductions in the magnitude of field estimation error with a modest navigational trajectory time. We evaluate our algorithm through a simulation, using a combination of static and mobile sensors sampling light illumination field.

Index Terms—Adaptive Sampling, Modeling, Field Estimation

I. INTRODUCTION

Over the last few years, we have witnessed the emergence and rapid maturation of a number of key embedded systems technologies, including reliable wireless communications, compact low-power micro-processor sensors and actuation enabled sensing systems. An example of such system is the Networked Info-mechanical System (NIMS) \cite{1} \cite{2} \cite{3}. NIMS is a mobile robotic sensing platform that has been developed to complement “traditional” fixed sensor deployments. It enables active physical reconfiguration of a diverse spatiotemporal distribution of sensor nodes in three dimensional environments (Fig-1). Its infrastructure provides high precision and reliable mobility to locate sensor nodes as required to configure sensor distributions to the points of minimum sensing uncertainty. However, this new capability is accompanied by many challenges, including the problem of optimizing spatiotemporal sensor sampling to reduce resource costs while recovering the most accurate reconstruction of environmental phenomena.

Consider the case that a robot carries a sensor for analyzing a phenomenon. Examples of such a phenomenon is solar radiation flux that affects natural ecosystem \cite{4}. Solar radiation is spatially filtered by complex ecosystem structure and ultimately affects photosynthesis and plant growth. Characterization of the solar radiation is then of primary interest to understanding the growth and evolution of plants in ecosystem. Other examples of such sensors are temperature, humidity and CO\(_2\). While the robot that carries those sensors can periodically sample the environment systematically or randomly, it can not adapt itself to the irregularities of features of the space such as inhomogeneity or anisotropy. This usu-
ally leads to underutilization of available sampling resources to maximize our knowledge about the phenomena under study. In principle, we would like to model the process by mathematical representation, incorporate prior knowledge in our model and select the best subset of points to represent the environmental field. Here, we assume that prior knowledge is incorporated through the previously sampled values by robot as it navigates across the field or collected by static sensors. One may also use some extra information about the model through previously known studies.

A major challenge in runtime adaptation is that of avoiding the trapping of a system in a local optimum. In the early stages of an adaptive exploration scheme, the selection of initial points has a strong effect on the field estimation performance; so that the addition of new sampling points may lead to biased estimation of the structures in the field. This may lead to increasing the mean squared error (MSE) and misleading any adaptive exploration scheme. To avoid this problem, we begin with a regular exploration scheme that is gradually relaxed as the robot gains more information about the field. Over time, a more adaptive scheme is followed that traces important features of the field. We will later address this problem.

The remainder of this paper develops our approach in more details and shows the result of our experimentation on some light scenes.

II. BACKGROUND

To date, there have been several innovative approaches to estimating a field with a fixed array of sensors. Many of these techniques employ some kind of adaptive querying and hierarchical processing [13], [14]. The background field model for these studies is often taken to be a piecewise-smooth surface, where the pieces are separated by smooth curves. In the statistics literature, we find a recent approach to this problem that utilizes the equivalent of a mobile sensing platform. In [10], the authors consider a smooth “fault” separating two smooth surfaces. A mobile node tracks the fault sequentially. Given an estimate of the separating curve, the node will take measurements in an arc extending out from its current position. These values are then examined for a breakpoint, indicating where the fault lies ahead of the node. In this way, a series of simple hypothesis tests are used to guide the node along the fault. While largely theoretical, this paper does illustrate one benefit of a mobile node: the mean squared error associated with estimating the important features of the field decreases (as we collect more samples) at a rate that is faster than we would expect for a smooth function of a single variable.

The motion dictated in [10] is tied to the idea of tracking faults. The general problem of design for mobile sensors was studied by control theorists nearly 20 years ago. In [11], the classical optimal design framework was extended to mobile nodes that collect data continuously. Versions of the various “alphabet optimality” criteria were developed to describe the best patterns of motion for a node given a specific parametric form for the underlying field. For example, if the field is the density corresponding to two independent normal random variables with a common variance that is increasing in time, the optimal design path could be a ray extending from the origin (the velocity of the node depends on time). In general, these results require knowing the precise parametric form of the field, limiting their applicability. In some cases, however, these results can be used to design sensor paths when prior information exists about the general shape of the field.

The so-called active learning schemes of [8], [9] also build on optimal design principles, but utilize non-parametric field estimators like neural networks or local polynomial regression. This approach involves constructing a second-order approximation to the likelihood and deriving analogs of the classical optimality criteria. As with the continuous sensor work mentioned above, these results implicitly assume zero bias in the estimation procedure (in the case of [11], this is essentially the assumption that the parametric form of the field is known precisely). In our experience, bias is perhaps the dominant component of mean squared error for field estimation, and these results do not seem immediately applicable.

Finally, one last line of research does seem to have considerable impact on the design of a mobile sensing node comes from biology. In [12], [15], simple sensing organisms are endowed with different forms of locomotion and assigned simple rules to respond to their measured “data.” Navigation strategies are then compared based on how well the organism achieves a goal like tracking gradients. Again, the paths discussed in this literature and the general research strategy can prove useful for our work with the NIMS node.

Prior exploration of NIMS adaptive sampling [2] [3] has demonstrated its feasibility for mapping of static phenomena. However, further requirements for characterizing dynamic phenomena, requires a dramatically new approach. As will be described, the method reported here introduces significant advantages in performance arising from a feature-driven design and active modeling of a phenomenon considering the cost of navigation and collecting measurements.

III. MOTIVATION

In this section, we develop a sequential sampling strategy that starts with classical experimental design principles. This initial approach does not incorporate the fact that the node needs to travel to take samples. We next incorporate navigation costs and explore some basic properties of the resulting algorithm. Finally, we add an “interest” measure that allows us to adapt our sampling to the characteristics of the field being sensed.

Let $\Omega$ represent the transect spanned by a NIMS node. In a typical deployment, we will equip the robot with a number of sensors, each able to record possibly different aspects of phenomena occurring within $\Omega$. For the moment, however, will only consider data taken from a single sensor. We will further assume that the transect itself has been instrumented by a set of static sensors of the same variety. Static sensors may be deployed at perimeter of the transect to reduce
boundary effect in the field estimation. In addition, since they provide data almost immediately they contribute to higher rate of reduction of mean square error. Finally, let $T_M$ denote the amount of time the node must be stationary so that the sensors can make a reliable measurement. In practice, this time is typically determined by underlying sampling physics. When collecting observations on CO$_2$ concentration, $T_M$ ranges from three to five minutes, while for light intensity, a PAR (photo-synthetically active radiation) sensor may take data almost continuously.

In this section, we will develop a framework for combining both fixed and NIMS observations to construct an adaptive navigation scheme in which the robotic node explores the field and gradually is directed to regions of the transect exhibiting interesting features. By lingering in these areas and making more measurements, we are better able to resolve strong features and have a more complete view of the field. We demonstrate that this approach provides us with more informative samples in the sense that we can achieve a greater reduction (per unit time) in estimation error than a simple raster scan or other space-filling designs. Implicitly, our setup assumes that the underlying phenomena affecting our measurements change slowly enough so that we can obtain a reasonable estimate of the field. If the scene we are observing changes rapidly in time, we suggest a different strategy entirely. We will return to the incorporation of temporal effects at the end of the paper.

A. Space Filling Designs

Ignoring for the moment the fact that the robotic node actually has to move to a particular location to make a measurement, the task of determining where to take data is essentially a problem of experimental design. Without prior knowledge about the structure of the field we are interested in, it is sensible to try to spread design points throughout the region, leaving as few holes as possible. While there are many approaches to constructing so-called space filling designs, we have chosen to build on a proposal by [16]. Let $d(x, y)$ be the simple Euclidean distance between $x$ and $y$, both in $\Omega$. Let $S$ denote a collection of $n + m$ points, where $n$ is the number of locations where we will use the NIMS node to collect data, and $m$ represents the number of fixed sensors located in $\Omega$. Then, $S$ is a maximin Euclidean distance design if and only if

$$\min_{x,y\in S} d(x, y) \geq \min_{x,y\in S^*} d(x, y)$$

for any other design set $S^*$ that includes the $m$ static sensor locations.

Given a fixed set of static sensors, there are computational methods to determine the optimal placement of the $n$ NIMS observations [17]. In practice, we would like to have a continuously refinable design, meaning that our $n$ can grow in time if the speed of the phenomena warrants it. A fast (approximate) sequential solution are the so-called “coffee-house” designs [18]. Given a set of points $S$, for each $x \in \Omega$ define

$$\text{Condition}(x|S) = \min_{y \in S} d(x, y)$$

To $S$, we add point $x$ for which Condition$(x|S)$ is a maximum. In essence, we are selecting $x$ from among the vertices of the Voronoi tessellation of $S$. This simple procedure has been shown to create designs with some of the same desirable properties as the maximin designs [18]. In terms of our NIMS application, this approach makes intuitive sense providing the amount of time required to measure a phenomenon is large relative to the speed of the node.

We formalize this notion by introducing navigation constraints on this sequential design scheme. Assume that $S$ consists of some number of previously visited points in the transect together with the $m$ static sensor locations. Assume that the node is currently in position $x_0 \in S$; that is, we have just taken a measurement at $x_0$ and need to plan our next move. We will select the point $x$ in $\Omega$ that maximizes the criterion

$$\frac{T_M + d(x_0, x)}{v}$$

where $v$ is the speed with which the node travels and the term in the denominator represents the total time required to travel to a new location and make a measurement. The above formula is a measure of benefit of each potential candidate sample in filling up the holes in space and the cost of vising that point and collecting a measurement. Here, we consider time as our cost metric. In practice one may pick other metrics such as energy or travelling distance.

We see that if $T_M$ is large, the effect of the travel time is small. In such cases robot may pick any point based on its benefit (since cost is universally constant) and we are left with the coffee-house design. As $T_M$ starts to rival travel time, we have to tradeoff gaps or irregularities in the design against their distance from the robot. Notice that for this sampling scheme, we can again restrict our attention to the Voronoi tessellation of the current design set. That is, given a set $S$ of design points, the maximum of (1) for all $x \in \Omega$ is occurs on the edges of the Voronoi tessellation of $S$.

In Fig-2, we illustrate a kind of hierarchy of designs in terms of the size of $T_M$. In the extreme case, $T_M >> 0$, a design is just a series of disconnected, well-separated points. At the other extreme, $T_M \approx 0$, the robot is sampling from the environment in (near) real-time, and our designs become

![Fig. 2. Sampling viewed through three different regimes based on the size of $T_M$, the time required to make a measurement.](image-url)
set by the node greedily applying the criterion (1). Here we
return to this material briefly at the end of the paper. In Fig-3, we show two different sample paths taken
by the node greedily applying the criterion (1). Here we
set $v = 1 m/s$ and $T_M = 10$ with a transect measuring
60m x 150m. In the upper panel, we have 14 equally spaced,
fixed sensors around the perimeter of the transect. In the
lower panel, we have specified a particular region of interest,
forcing the node to explore a circular area in the middle of
the transect. Notice in this case that the first move the node
makes is to sweep out a hexagon and then dive into the center.

B. Feature-driven design

When the phenomena under study evolve slowly relative
to the speed of the collection capabilities of the node, we can
view the problem as an example of function estimation. The
unknown field is a surface over $\Omega$, and given a set of points
at which we have made observations, we can construct an es-
timate using a non-parametric smoother. In [3], we employed
local polynomial regression with a fixed number of nearest
neighbors. This procedure has the property that as more
points are added to a region, the bandwidth of the estimator
automatically decreases, providing greater ability to resolve
features in that region. In that paper, we also noted that in
many sensing situations, the mean squared error in estimating
a field is dominated by bias. By contrast, many active learning
procedures adopt an approximate optimal design framework
that focuses attention on the variance component of mean
squared error [8], [9].

Continuing with a bias-dominated view of the underly-
ing field estimation problem, we have previously proposed
adaptive sampling criterion that places points in regions with
significant misfit. An estimate of the bias is used to guide our
procedure, and points are introduced in batches [3]. This batch
scheme did not, however, directly address node navigation,
and is perhaps most appropriate in situations where $T_M \gg
0$. In addition, each batch of new points were not gracefully
added to the region of high bias; instead a semi-regular design
was simply overlayed on the previous sampling points. This
had the tendency to create clumps and odd structures in the
resulting design.

To address these problems, we now fold our feature-
oriented sampling scheme into the navigation criterion discussed
above. Let $S$ be a set of design points at which we have
taken a series of measurements. We then define $\text{Feature}(x|S)$
to be an interest score for the neighborhood of $x$. In our
applications, we will take $\text{Feature}(x|S)$ to be an estimate
of the bias error in our field estimate using samples observed at
$S$. We then select our next design point $x$ so as to maximize
the combined utility:

$$\frac{\text{Condition}(x|S) + \lambda \text{Feature}(x|S)}{T_M + d(x_0, x)/v},$$

where $\lambda$ is a balancing coefficient. This criterion explicitly
balances our interest in tracking the features with our desire
for a regular design; the smaller $\lambda$, the greater our emphasis
on regularity. In our experiments, we use a local polynomial
fit to estimate the field. We compute residual errors at the
sampled locations and estimate the field error based on the
observed residuals. The result is an estimated error map that
is taken to be $\text{Feature}(x|S)$.

There are many different methods that could be used in
this capacity; for example, [19] considers using the second
derivative of an over-smoothed estimate of the field to de-
rive an estimate of the bias. Naturally, the interest measure
$\text{Feature}(x|S)$ can encode other aspect of the field that we
hope to capture through our sampling.

C. Transitional Phase

As mentioned earlier, we need to strike a reasonable
balance between regularity of the design and adaptation to
strong features in the field. This balance will shift as we
collect more data, with regularity being emphasized early in
the process and adaptation coming later as the robot collects
information. To achieve this, we use an exponentially growing
process that sets $\lambda$ based on the number of samples that has
been collected:

$$\lambda(i) = \lambda_\infty(1 - e^{-i/\tau})$$

where $i$ is the number of collected measurements and
$\tau$ is the growth rate constant. The parameter $\tau$ should be
picked based the expected rate of growing attention to the
features as samples are being collected. This specification
initially assigns $\lambda$ to be zero meaning an emphasis on regular
design and later moves gradually toward a mixed design
with with limiting value of $\lambda_\infty$. In practice picking proper
values for $\lambda$ and $\tau$ is challenging. Typical values that we set
throughout our experimentation are $\lambda_\infty = 1$ and $\tau = 100$

Fig. 3. Two sample paths, each of 70 points, from the sequential design
scheme. In the lower plot, we have specified a circular region of interest.
A. Tools and setup

NIMS technology has been recently deployed in multiple of sites [7] and characterization of forest ecosystem phenomena is underway. However, any performance analysis of high level algorithms is not possible without having a complete understanding of underlying field variable data. We have used the indoor NIMS system, NIMS-LS [3], to create a representation of an actual environment in a controlled condition. The NIMS-LS system allowed us to generate different light patterns by having different arrangements of illumination sources and obstacles. We then use the NIMS mobile sensor system to measure the resulting light field. This combination enabled us to perform dense sampling of the environmental light variable under controlled conditions. This data is then applied as ground truth for verification of our algorithm versus a uniform sampling design (raster scan). Our algorithm is implemented in the R statistical computing environment [20]. This is then available both as a real-time service on the NIMS system, with access to sensors and actuators. It is also available for emulation of NIMS operation and as well as available for post-processing of archived data.

To evaluate our algorithm we subjected it to environmental fields having two extremes in their curvature characteristics: 1) For one limit, the environmental variable field was created by placing a high density of obstacles in the illumination field to emulate the characteristically most complex patterns observed in the natural environment. 2) We then reduced the number of obstacles to represent environments with sparse light segments. Figure-5 shows these phenomena. We configured the transect size to be 8m in length and 2.5m in height and densely sampled the environment at 5cm intervals to generate the data for the performance analysis. The light intensity of the scene varied by a factor of 5.7 from darkest to the lightest regions of the transect.

B. Experiments

The ultimate goal of our sampling method is to best reconstruct the underlying phenomenon in space. To test this we operated the adaptive sampling system algorithms with data input directly from the field variable maps captured by NIMS-LS (Fig-5). As the sampling commenced, We then measured the reconstruction performance of our algorithm as it varied with time. Performance is measured as the Mean Squared Error computed across the entire variable field area. The adaptive and uniform sampling approaches were compared for each example. In all the experiments we set the speed of the robot to be 0.5m/s and the sampling time to be 0.5sec. These values correspond to those of the NIMS mobile platform and its sensors characteristics. In all cases we used a local linear smoothing function provided by the R local regression and likelihood package [22] to reconstruct the variable field as the mobile sensor collected new samples (Fig-7(ai)). We used the estimation package to predict error across the field based on residual error at the sampled points (Fig-7(c)).

Fig-6 shows the distribution of sampling points in the environment. As can be seen, the sampling points are distributed in space near locations where the field variable value spatial derivative is largest. Fig-4 shows the rate of reduction of Mean Squared Error for three different cases of uniform design with 100, 400 and 600 measurements and for the case of adaptive design. It shows that the adaptive case outperforms all the three cases of uniform sampling design. In our algorithm, the initial desire to fill the space guarantees a high rate of improvement in the Mean Squared Error. This warrants a fast transient response that entirely outperforms the two dense uniform design cases and it is comparable with a sparse uniform sampling (100 measurements). Later in time evolution, as the robot relaxes to higher fidelity measurement in the interesting regions, it outperforms the cases of sparse uniform sampling and is only comparable to a very high sampling rate of uniform design with 600 sample measurement points. This suggests that our algorithm achieves a fast response time as well as displaying a low steady state error.

An important characteristic of our design is its smooth reduction in Mean Squared Error(Fig-4) in time. In a uniform design, the biased nature of collecting measurement points (from one side of the phenomenon) may create a high degree of variation in the Mean Squared Error. In our case, however the reduction in Mean Squared Error occurs smoothly. In practice, our algorithm will permit a user to specify a desired fidelity level (MSE value). Then, the algorithm will enable the mobile sensing system to smoothly traverse until the desired degree of fidelity (Mean Squared Error) is reached.

An animation of these experimental results can be found at: http://cens.ucla.edu/~mhr/nims/iros2005/iros.wmv
Fig. 5. The environmental field maps measured using NIMS indoor system. These phenomena have been generated by applying different combination of illumination sources and obstacles to emulate different natural light patterns. NIMS robot then samples the environment. This includes a very high spatial density sampling to generate ground truth dataset for post-processing analysis.

Fig. 6. Spatial distribution of sampling points (a) after 15min for the complex scene and (b) after 10min for the sparse scene. According to the adaptive sampling method, the robot exploration enables it to most frequently visit the regions of highest error. These regions generally occur near rapidly changing “edges” of the variable field.

Fig. 7. (a) Instantaneous estimation and reconstruction of the field by the robot after 10min. The robot continuously reconstructs the field to discover the locations of maximum error. (b) Robot utility after 1min. The robot is located at (6.1m, 1.7m). In the initial phases the utility is dominated by the space filling design (c) The error map generated by applying a local smoothing estimation on residual error. As expected, it shows maxima at the edges observed in the variable field.
V. Future work

Environmental robotics technology such as NIMS are now being applied in critical science and engineering applications in complex environments. However, such environmental field characterizations confronts the challenge of spatiotemporal evolution of the sensing environment. This introduces an unknown level of measurement distortion. This paper describes a new architecture that enhance utilization of the underlying technology by intelligent use of such sensing resources. This architecture incorporates a system that exploit a combination of: 1) regular sampling design in a way that maximize sampling coverage in space and 2) adaptive mobility to actively explore environments and determine sampling points distribution based on the observed existence of the features in the environment.

While the experiments described in the previous sections are not exhaustive, they clearly show that our algorithm can improve the sampling performance of a robotic sensing system. We have shown that our scheme outperforms a uniform sampling design in terms of rate of MSE improvement and its steady state response. We have also demonstrated that our method achieves a balance between the sampling and sensing delay tradeoff incorporating knowledge of mobile sensor traveling time.

Future research is directed to incorporation of these methods into a regular sampling strategy. As phenomena change in time, our new methods will gradually deemphasize the influence of prior measurements. To achieve the proper sample lifetime, our systems will consider both phenomena rate of change along with robot actuation speed.

Early experiments suggest that we can achieve a steady-state design process that fills vacancies in the design as they emerge and still maintain adaptation to field features. We also have experimented with the addition of a probabilistic framework for navigation; rather than pursuing an optimum according to a greedy algorithm, this new approach allows for considering the value and selecting one of several promising directions.

Throughout this paper we have assumed that the phenomena under study are moving slowly enough to permit us to create an accurate estimate of the variable field. In some situations, however, this is not the case. For example, consider solar radiation light measurements of a transect under a forest canopy. The pattern of light might be very stable (where created by the shadow of foliage that are static in nature) or it might be quite dynamic (the pattern of light and shadow formed by foliage moving in a high velocity wind). In the former case, it is sensible to think about a function estimation formed by foliage moving in a high velocity wind). In such a setting, our strategy for sampling is to consider some very regular design. For light, we are currently experimenting with roulettes (spirograph patterns) that provide an unbiased estimate of the size of the light and dark fields.

Finally multi-robot extensions of our algorithm may be useful in applications that require more sampling resources for better characterization of the phenomenon. In principle the approach described here should be extensible to multi-robot examples by incorporating a proper balance between regularity and feature driven design.

VI. Acknowledgement

The authors wish to acknowledge the expert contributions by Steven Liu, Richard Pon, and Lisa Shirachi.

REFERENCES