Sampling Optimization for Networked Underwater Gliders

Shijie Liu*†, Jie Sun*†, Jiancheng Yu*, Aiqun Zhang* and Fumin Zhang‡
*Shenyang Institute of Automation, Chinese Academy of Sciences
Shenyang,110016, P.R.China.
Email: {liushijie,sunjie,yjc,zaq}@sia.cn
†University of Chinese Academy of Sciences
Beijing, 100049, P.R.China.
‡School of Electrical and Computer Engineering, Georgia Institute of Technology
Atlanta, Georgia 30332, USA
Email: fumin@gatech.edu

***Abstract*—This paper focuses on coverage sampling over a given region by a fleet of underwater gliders. The coverage performance is evaluated by the observing uncertainty performed by Objective Analysis (OA). Deployed gliders are commonly restricted to move through transects along which the data collected are convenient for data assimilation. We optimize the initial location and the initial heading of each glider that are deployed into an interested region by using Generic Algorithm (GA). Two cases of environmental sampling tasks are used for verifying the proposed approach.**

I. Introduction

The marine environment is a very complex system in which strong interactions between physical, chemical and biological processes take place. These processes occur on varying scales in time and space, which makes the investigation of them particularly challenging. Emerging Autonomous Underwater Vehicles (AUVs) enable more efficient data acquisition approaches. Underwater gliders, a new type of AUVs, have been deployed to provide collect data in many regions around the world. Underwater glider fleets are often used cooperatively for large-scale ocean observation. In an interesting region, one would like to cover the region with a glider fleet while the goal is to decrease observation uncertainty or estimate error.

Ocean processes change over time and space, from seconds to decades and from millimetres to kilometers. The observation of these dynamic processes requires both temporal and spatial resolutions. In an early work, Willcox et al. [1] used spatial and temporal aliasing arguments to define trade-offs between spatial resolution and survey time for a grid survey with a single AUV. Leonard et al. extended the problem of synoptic sampling with multiple vehicles in [2], [3]. They developed a controller that optimized spacing between multiple vehicles moving along elliptical racetracks. In work [4], Zhang et al. coordinated a fleet of gliders to generate patterns on closed smooth curves. The phase and the distance between gliders are measured for better spatial and temporal area coverage.

To quantify how well an array of gliders sample a given survey region, a performance metric should be designed as the optimization criterion. In [2], this metric is called coverage metric since it specifies the statistic uncertainty as a function of where and when the data is taken. The metric is derived from a data assimilation scheme, Objective Analysis (OA) [5]. Furthermore, the trajectories of gliders are coordinated as collective patterns on certain shaped curves. These curves are parameterized by a limited number of parameters avoiding searching in a large functional space when optimizing individual trajectories.

The intelligent algorithms, such as genetic algorithm, perform well are able to a large state-space, multi-modal state-space, or n-dimensional surface. There are also some works contribute to improving the performance of underwater vehicles. In [6], Alvarez et al. proposes a GA-based approach for path planning of an autonomous underwater vehicle in an ocean environment characterized by strong currents and the space-time variability is enhanced. In a recent research [7], Frolov et al. extended survey from grid-based into full directional by using GA to optimize the heading angle of a vehicle. Instead of covering the entire survey area, the method focuses on revisiting areas of high uncertainty.

This work develops a GA based solution to achieve high value data collection by a fleet of underwater gliders. We discuss two cases of environmental backgrounds: 1) there is no prior knowledge available and 2) we know the true value or predictions of the environment. The value of collected data is demonstrated to decrease the coverage uncertainty and the estimate error in two cases. Different from Leonard's work [2], we use Genetic algorithm for efficient searching. Rather than moving in free angles in Frolov's work, the gliders move along transects after the initial locations and heading have been optimized. That is to response to the requirements of better assimilating the collected data.

The remainder of this paper is organized as follows. To quantify the quality of collected data, we introduce a performance metric and then derive the optimization function in our case in Section II. The proposed approach is based on GA, detailed in Section III, we first simply describe the principle of classical Genetic algorithm and then give the genetic operators for solving our optimization problem. The simulation results are given in Section IV and conclusions are given in Section

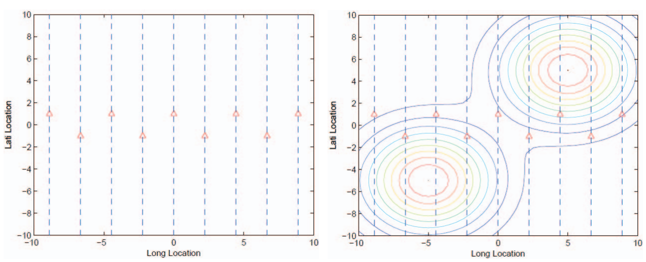
This work is supported by National Natural Science Foundation of China, Grant No. 61233013 and the State Key Laboratory of Robotics at Shenyang Institute of Automation (Grant No. 2014-Z02).

978-1-4673-9724-7/16/$31.00 ©2016 IEEE

V.

II. Coverage Problem Statement and Metric Definition

For large-scale or high-resolution ocean observation, single glider is unable to cover the entire domain. It is preferred to use a fleet of gliders cooperating with each other for spatial and temporal coverage [4]. This often requires that the trajectories of the gliders to be coordinated to avoid under-sampling or over-sampling of certain area. Assume that a fleet of gliders are available for the survey and each glider moves along one transect. Fig.1 illustrates 9 gliders patrolling through a rectangle region, with the blue dashed lines are the transects and the red triangles are gliders. The goal of our research is to find out initial location and heading for each glider while the integral of coverage uncertainty or estimate error is minimum. The minimizing term is accorded with two environmental backgrounds under which the distribution of gliders is performed. In one environment the value of interest is unknown (Fig.1(a)) and in the other is known (Fig.1(b)) with the closed curves are contour of a generated scalar field. Next, derived from [2], we describe the performance metrics for these two environments.



(a) (b)

Fig. 1. 9 gliders patrolling with in a rectangle region while each of them move along one transect. The blue dashed lines are the transects and the red triangles are gliders. (a) is the first environment where no prior knowledge is obtained before. In (b), the closed lines are the contour of a generated scalar field.

Assume that the scalar field (e.g., temperature, salinity) the glider observed at each point $\mathbf{r}$ and at each time t is viewed as a random variable $T(\mathbf{r}, t)$. Let P be the number of observations at one time, based on Objective Analysis (OA) framework, the estimated value of T over the entire region can be expressed as

$$\hat{T}(\mathbf{r}, t) = \bar{T}(\mathbf{r}, t) + \sum_{k=1}^{P} \zeta_k(\mathbf{r}, t)[T_k - T(\mathbf{r}_k, t_k)]$$

with the optimal coefficients

$$\zeta_{kl}(\mathbf{r}, t) = \sum_{k=1}^{P} B(\mathbf{r}, t, \mathbf{r}_l, t_l)$$

Where $B(\mathbf{r}, t, \mathbf{r}_l, t_l)$ is the covariance of the mean of $T(\mathbf{r}, t)$, namely $\bar{T}(\mathbf{r}, t)$, and C^{-1} is the inverse of the $P \times P$ covariance matrix of the data T_k. Then performance matric is the

$$\begin{aligned} \phi \ &= \int dr \int dt (B(\mathbf{r}, t, \mathbf{r}_l, t_l) \\ &- \sum_{k,l=1}^{P} B(\mathbf{r}, t, \mathbf{r}_k, t_k)(C^{-1})_{kl} B(\mathbf{r}_l, t_l, \mathbf{r}, t)) \end{aligned}$$

The metric above depends only on the covariance function $B(\mathbf{r}, t, \mathbf{r}, t)$ and the sampling location $\mathbf{r}_k$ and time t_k. An experiential background covariance is selected as $B(\mathbf{r}, t, \mathbf{r}', t') = \sigma_0 e^{-\frac{(\mathbf{r}-\mathbf{r}')^2}{\sigma} - \frac{(t-t')^2}{\tau}}$. All the gliders travel back and forth along different transects after selecting initial location $\mathbf{r}_0$ and heading direction θ_0. In our work, one glider patrols on one transect and its state can be denoted by $(\mathbf{r}, \theta, t)$. A cycle of one glider is defined starting from $(\mathbf{r}_s, \theta_s, t_s)$ and end when it passes through point $\mathbf{r}_s$ with the same heading direction θ_s at time t_e. The time span during a cycle is $t_c = t_s - t_e$.

In the first environment, there is no prior knowledge of the value of T. The solution of coverage optimization is dedicated to decreasing the covering uncertainty accumulated within a cycle over the entire region. Assume that the gliders surface to communicate their data to processing center every time t_p. Then one glider would have $N = [\frac{t_c}{t_p}]$ surfacing within one cycle. The optimizing function can be expressed as

$$\phi_{np} = \sum_{n=1}^{N+1} \sum^{\mathbf{r}} \phi(n) \tag{1}$$

where $\phi(n)$ denotes the uncertainty of every gridpoint in the survey region at the n_{th} surfacing.

In the second environment, we can get or forecast the value of T. The coverage here is to decrease the estimating error $\Delta T = T - \hat{T}$. The accumulate error in a cycle over the entire region can be expressed as

$$\phi_p = \sum_{n=1}^{N+1} \sum^{\mathbf{r}} \Delta T(n) \tag{2}$$

where $\Delta T(n)$ denotes the estimating error of every gridpoint in the survey region at n_{th} surfacing.

Equation (1) and (2) are functions we are intend to optimize and are selected to be the fitness function in genetic algorithm.

III. Genetic Operations

The genetic algorithm (GA) is a search heuristic that mimics the process of natural selection by a population of candidate solutions (also called individuals, creatures, or phenotypes) [8]. Each candidate solution can be encoded in a DNA-like string and has a set of properties (its chromosomes or genotype) which can be mutated and altered. A typical genetic algorithm requires: 1) a genetic representation of the solution domain and 2) a fitness function to evaluate the solution domain. The evolution is conducted with iterative process starting from a population of randomly generated individuals, with the population in each iteration called a generation. In each generation, the fitness, i.e. the value of the objective function, of every individual in the population is evaluated. The more fit individuals are more likely selected from the current population. Each selected individual's genome is modified (recombined and possibly randomly mutated) to form a new generation which is then used in the next iteration of the algorithm.

Three genetic operators dominate the process of evolution

- Selection - executes after the fitness of a generation has been evaluated. The fitter solutions are typically more likely to be selected to serve as parents to the next generation.

- Crossover - occurs between a pair of “parent” solutions to generate “child” solutions that typically share many of the characteristics of their “parents”.
- Mutation - alters one or more gene values in a chromosome from its initial state, expecting to come to better solution.

This generational process is repeated until a termination condition has been reached. Common terminating conditions include: 1) a solution is found that satisfies minimum criteria; 2) fixed number of generations is reached; 3) no longer produce better results in sequentially a certain steps.

We implemented GA adapted from Genetic Algorithm Toolbox [9] developed by University of Sheffield, UK. Based on the statement in section II, the population indicates the initial position r_x and r_y of gliders and the initial heading direction θ. Since the gliders only move back and forth along their own transects, the directions are quantified by 1 for one direction and -1 for the other, for computational simplicity. A population example of 6 gliders is shown as follows:

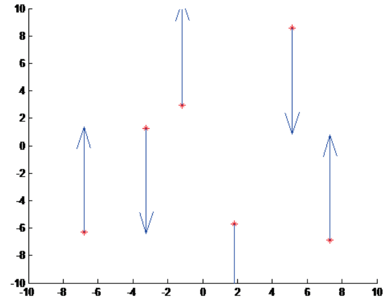
$$\begin{array}{lllllll} r_x[& -9.81 & -0.46 & -4.99 & -3.84 & 9.33 & -5.82 \\ r_y & 0.40 & -5.48 & 1.34 & 9.96 & -7.36 & 9.09 \\ \theta & -1 & -1 & 1 & -1 & -1 & 1 \quad] \end{array}$$

The fitness functions in two cases are equation (1) and (2). Starting from each initial state, gliders move from their own initial location with initial heading direction along transects every time interval with constant speed. The value of the fitness function would be calculated after all the glider complete a cycle. In selection process, 10% less fit solutions will be left out and replaced by new random solutions in order to maintain diversity of population. Two typically crossover techniques exist, One-point crossover and Two-point crossover. The latter, our choice, calls for everything between two points on the parent organism strings is swapped.

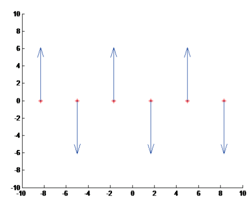
A main drawback of GA is that it may have a tendency to converge towards local optima or even arbitrary points rather than the global optimum. This problem may be alleviated by increasing the rate of mutation, using selection techniques that maintain a diverse population of solutions, or particular solution for specific problem.

IV. Simulation Results

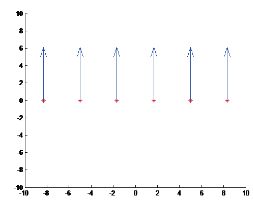
We set the spacial factor and temporal factor in the covariance metric B as $\sigma = 20, \tau = 50$. The performance of GA-based optimization method is first performed in the environment without prior knowledge. Six gliders are considered in surveying the region of interest. Equation (1) is set to be the optimization metric, or the fitness function in GA. Solution achieved by the proposed approach is shown in Fig.2(a). If without optimization, a general intuitive distribution of gliders might be as Fig.2(b) or 2(c). We also calculate the overall uncertainty in these two cases for comparison. The accumulated uncertainty in Fig.2(a) is 2.0716×10^4, in Fig.2(b) is 2.7258×10^4 and in Fig.2(c) is 3.6669×10^4. The result in Fig.2(a) seems not significantly improved comparing with that in Fig.2(b). The reason is that for a homogeneous environment



(a) Optimized solution



(b) As a comparison



(c) As a comparison

Fig. 2. In the first environment, the initial locations and heading directions of gliders in (a) are optimized by GA method, while as comparison, (b) and (c) are following intuitive distribution.

background, the patterns in Fig.2(b) is already a very good solution. GA probably terminates into a local minimum that the result is probably not the optimal one.

If we explore a region that we have already got some knowledge or an unknown region has been observed before, the optimization would focus on minimizing the estimate error. A scalar feature

$$T(x, y) = h \cdot (e^{-\frac{(x-5)^2+(y-5)^2}{20} - \frac{t^2}{50}} + e^{\frac{(x+5)^2+(y+5)^2}{20} - \frac{t^2}{50}})$$

is generated in the region, in Fig.3, and serves as the target of observation.

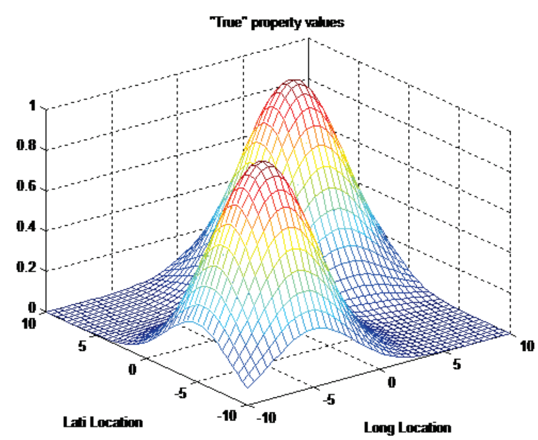
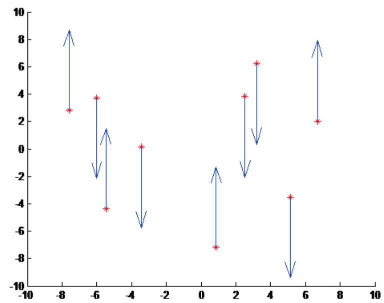


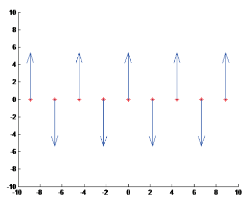
Fig. 3. A generated scalar field $T(x, y)$.

Equation (2) is set to be the optimization metric in this case. The optimization solution is giving in Fig.4(a). As comparisons, intuitive distribution is also shown in Fig.4(b) and 4(c). The accumulated estimate error in Fig.4(a) is 414.0401, while

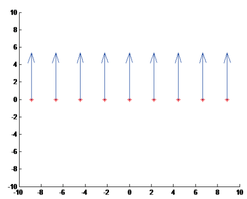
1.9562×10^3 and 6.5120×10^3 in Fig.4(b) and 4(c) respectively. The results reflect significant improvement achieved by the proposed method.



(a) Optimized solution



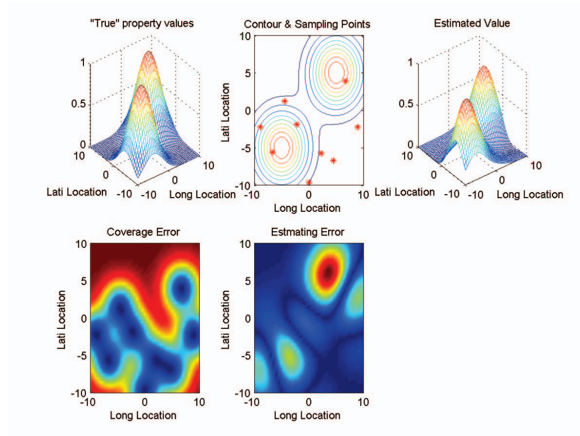
(b) As a comparison



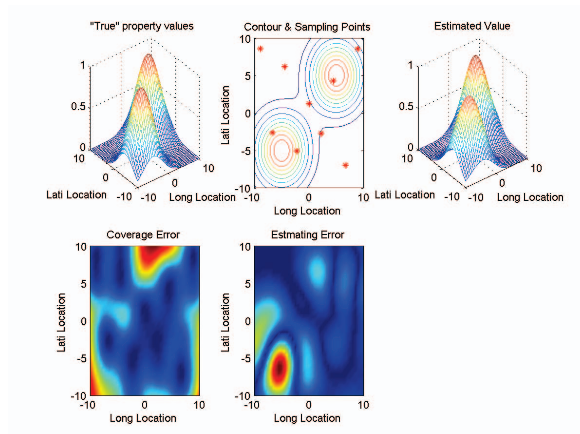
(c) As a comparison

Fig. 4. In the second environment, the initial locations and heading directions of gliders in (a) are optimized by GA method, while as comparison, (b) and (c) are following intuitive distribution.

Given an initial state, let gliders move along transects with velocity of 1. We sequentially plot out the value of generated scalar field, locations of gliders, estimated value, coverage error (uncertainty) and estimating error evolving with time. We give snapshots at two points of time in Fig.5. The virtual surfacing locations have added small noises for more realistic.



(a)



(b)

Fig. 5. Two snapshots of observation state of gliders moving along transects. From left to right, up to bottom sequentially are the value of generated scalar field, locations of gliders, estimated value, coverage error (uncertainty) and estimating error.

V. Conclusions

This paper presents an approach to optimize strategies of ocean covering observation over space and time by a swarm of underwater gliders. The movement pattern of gliders described here is simple but more related to real experiment. The intelligent algorithm based approach provides possibility for efficiently search in more complex situation.

References

[1] J. S. Willcox, J. G. Bellingham, Y. Zhang, and A. B. Baggeroer, "Performance metrics for oceanographic surveys with autonomous underwater vehicles," *IEEE Journal of Oceanic Engineering*, vol. 26, no. 4, pp. 711–725, 2001.

[2] N. E. Leonard, D. a. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. E. Davis, "Collective Motion, Sensor Networks, and Ocean Sampling," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 48–74, jan 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4118466

[3] N. E. Leonard, D. A. Paley, R. E. Davis, D. M. Fratantoni, F. Lekien, and F. Zhang, "Coordinated control of an underwater glider fleet in an adaptive ocean sampling field experiment in Monterey Bay," *Journal of Field Robotics*, vol. 27, no. 6, pp. 718–740, nov 2010. [Online]. Available: http://doi.wiley.com/10.1002/rob.20366

[4] F. Zhang, D. M. Fratantoni, D. A. Paley, J. M. Lund, and N. E. Leonard, "Control of coordinated patterns for ocean sampling," *International Journal of Control*, vol. 80, no. 7, pp. 1186–1199, 2007.

[5] F. Bretherton, R. Davis, and C. Fandry, "A technique for objective analysis and design of oceanographic experiments applied to MODE-73," *Deep Sea Research and Oceanographic Abstracts*, 1976. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0011747176900012

[6] a. Alvarez, a. Caiti, and R. Onken, "Evolutionary path planning for autonomous underwater vehicles in a variable ocean," *IEEE Journal of Oceanic Engineering*, vol. 29, no. 2, pp. 418–429, 2004.

[7] S. Frolov, B. Garau, and J. G. Bellingham, "Can we do better than the grid survey: Optimal synoptic surveys in presence of variable uncertainty and decorrelaton scales," *Journal of Geophysical Research : Oceans*, vol. 119, no. 8, pp. 1–20, 2014.

[8] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, Reading, MA, 1989, no. 2.

[9] A. Chipperfield and P. Fleming, "The matlab genetic algorithm toolbox," in *Applied control techniques using MATLAB, IEE Colloquium on*. IET, 1995, pp. 10–1.