Near-optimal Collecting Data Strategy Based on Ordinary Kiriging Variance*

Xinke Zhu[1,2], Jiancheng Yu[1], Shenzhen Ren[1,2], Xiaohui Wang[1]

[1]Shenyang Institute of Automation, Chinese Academy of Sciences
& State Key Laboratory of Robotics, Shenyang, China

[2]Graduate School of the Chinese Academy of Sciences, Beijing, China

Abstract- **When monitoring spatial phenomena, we are not just interested in measurements at sensed locations but also at locations where no sensors were placed. To estimate the scalar field where no sensors are deployed, we need to interpolate the data. We are interested in how the best sampling design is to be found and best used to draw conclusions about the field as whole. First of all, a performance metric is defined to quantify how well the sampling network collecting data in a given region. Secondly, near-optimal collecting data strategy proposed minimizes the integral of the Kriging variance over the area of interest. Thirdly, several approaches proposed make the optimization more computationally efficient. Finally, the proposed methods are verified respectively by simulation.**

I. INTRODUCTION

In marine environments, accurately measuring quantities such as temperature, salinity, chlorophyll, and concentration of various nutrients is useful to scientists understanding of marine ecosystems. Generally speaking, these quantities of interest are scalar fields and each of them is characterized by a single scalar quantity which varies spatiotemporally. Accordingly, understanding the physical oceanography and how it couples with the biological dynamics is necessary for tackling a number of important open problems.

Mobile sensor networks composed of underwater gliders provide new tools for observing and monitoring the environment [1]. When monitoring spatial ocean process, we are not just interested in observations at sensed locations but also at locations where no sensors were placed. To estimate the scalar field where no sensors are deployed, we need to interpolate the data. Intuitively, the more the measurements near the location estimated is desired, the less the reconstructed errors. In other words, the spatial distribution of the observations affects the estimation error. If the budget for the survey is not limited, then the optimal sampling strategy is the one that just achieves the desired precision. However, the budget is always limited, therefore, how to place the sensors or deploy the mobile sensor networks to collect the data that best reveals the ocean processes of the interest given available resources is a key question.

Fiorelli et al. focused on sampling at smaller spatial scales and addressed the problems of gradient climbing and front tracking in [2]. Leonard et al. designed near-optimal sampling path using objective analysis as performance metric in [1]. Zhang and Sukhatme proposed an adaptive sampling algorithm for estimating a scale field using a robotic boat and a sensor network in [3]. All of them assumed that the communication between the networks is frequent. However, this condition isn't always satisfied under large scales ocean sampling. Krause et al. proposed a near-optimal algorithm to solve a problem similar et al. to ours in [4]. They assumed that the spatial process is Gaussian. However, it is not always universal.

Both the theoretical analysis and the empirical study show that the Kriging technique outperforms other commonly used techniques, which account for spatial correlation are more efficient than those that do not. One of factors affecting Kriging estimation and the associated accuracy is statistical inference model. All kind of methods were tried to refine statistical model such as covariance function or variogram, for example, in [5], [6], and [7]. In this paper we don't discuss about how to calculate covariance function or variogram. Another factor is the spatial configuration of the available observations. Therefore, the Kriging variance can be used to construct an objective function for designing optimal spatial sampling strategies and that problem has attracted the interest of a number of authors.

Van Groenigen discussed the influence of variogram parameters on optimal sampling schemes for mapping by Kriging in [8]. McBratney et al. used variogram form previous data to calculate optimal sampling density, given certain accuracy in [9]. Van Groenigen el al. proposed a spatial simulated annealing method to optimize spatial sampling schemes for obtaining the minimal Kriging variance in [10]. On the one hand, all of approaches above are lacking for taking Kriging computing time into account. Obviously, acquiring every Kriging estimation and variance need to calculate a large linear system and an inverse matrix. Therefore, selecting the optimal sampling locations from a set of candidate locations maybe result in excessive computation time. On the other, none of the approaches above guarantees a constant-factor approximation of the best set of sensor locations in polynomial time.

In this paper, we improve the searching algorithm in two ways, which one is to reduce computation time using Kriging variance updated equation [11], the other is to use heuristic

* This work is supported by the State Key Laboratory of Robotics #PLZ200810 and the Knowledge Innovation Program of the Chinese Academy of Sciences #KZCX2-YW-JS205.

978-1-4244-5222-4/10/$26.00 ©2010 IEEE

algorithm. For the latter, we give a proof which guarantees a constant-factor approximation of the best set of observation locations in polynomial time.

The paper is organized as follows. In section 2, ordinary Kriging technique is represented. In section 3, the sampling performance metric used to design the near-optimal collecting data strategy is defined and we would prove the solution by the near-optimal heuristic algorithm to have a guaranteed minimum performance level of 1−1/e when compared to the optimal solution. Several approaches make the optimization more computationally efficient in section 4. The proposed methods are verified respectively by simulation in section 5. Concluding remarks are in section 6.

II. SCALE FIELD ESTIMATION

This section reviews the geostatistical Kriging procedure for the estimation of spatial processes, see e.g. [12]. It is assumed that the covariance function has calculated from previous data. The scalar field (e.g., temperature, salinity) observed at each point x is viewed as a random variable $z(x)$. According to the difference of the mean function $z(x)$, Kriging can be classified simple Kriging (SK), ordinary Kriging (OK) and universal Kriging (UK). We only discuss OK here.

Suppose the random function $z(x)$ satisfy with second order stationary, and then the covariance function $c[z(x_1), z(x_2)]$ only depends on difference $x_1 - x_2$. Kriging procedure is concerned with the estimation of $z(x)$, an unknown value at known locations x. Now data $z(x_i)\ \ i = 1, \cdots n$ are observed at known locations. For simplicity, we assume that the observation error is zero. The OK estimation and variance at locations x are represented by $\hat{z}[x \mid z(X)]$ and $\sigma^2[\hat{z} \mid X]$, respectively, where $z(X)$ indicates the set of available sampling values at known locations $X = [x_1, \cdots, x_n]^T$. $\hat{z}[x \mid z(X)]$ is best linear unbiased estimation, which is combination of $z(X)$, can be written as:

$$\hat{z}[x \mid z(X)] = \sum_{i=1}^{n} \lambda_i z(x_i) \tag{1}$$

Weights λ_i satisfies with: $$\begin{cases} \sigma^2[\hat{z}(x \mid z(X)] = Min \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases} \tag{2}$$

Using method of Lagrange multiplier, equation (2) can be transformed into:

$$\begin{cases} \sum_{i=1}^{n} \lambda_i c(x_i, x_j) - \mu = c(x, x_j), \ \ j = 1, \cdots n \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases} \tag{3}$$

Where, μ is a Lagrange multiplier, $c(x_i, x_j)$ is the covariance between observation points x_i and x_j, $c(x, x_j)$ is the covariance between location x to be predicted and the observation point x_j. Substituting (3) into (2), we can derive the Kriging variance:

$$\sigma^2[\hat{z}(x \mid X)] = c(x, x) - \sum_{i=1}^{n} \lambda_i c(x, x_i) + \mu \tag{4}$$

Equation (4) can be written in matrix form as:

$$\sigma^2[\hat{z}(x \mid X)] = c(x, x) - V_n^T H_n^{-1} V_n \tag{5}$$

Where, $V_n = [c(x, x_1), c(x, x_2), \cdots c(x, x_n), 1]^T$

$$H_n = \begin{bmatrix} c(x_1, x_1) & \cdots & c(x_1, x_n) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ c(x_n, x_1) & \cdots & c(x_n, x_n) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}$$

III. OPTIMIZING SAMPLING

We are interested in how the best sampling design is to be found and best used to draw conclusions about the field as whole. However, the budget is always limited, therefore, how to select the best locations out of the whole sampling space to collect the data is a key issue.

A. Sampling Performance Metric

In this section we define a performance metric to quantify how well the sampling network collecting data in a given region. Recall that an objective is to estimate an unknown value at known locations according to the measurements at known locations. Therefore, the metric should reflect how a particular collected data set reduces the estimation error.

Kriging variance can be used as a quantitative measure of the impact of the sequence of measurements on the error in the sampling scheme. Form (3) and (4), it shows that the Kriging variance depends only on the distances between x and the n measurement locations x_i, the geometry of the n measurements and the covariance function (variogram). The values of the sample observations have no influence. In other words, if the covariance function is known, the Kriging variance can be predicted for any proposed sampling strategy prior to the actual survey. Kriging is, therefore, an ideal tool for designing optimal spatial sampling strategies. Therefore, the integrated Kriging variance (IKV) is used as the performance metric:

$$g(X) = \int_{x \in d, X \subseteq d} \sigma^2[\hat{z}(x \mid X)] dx \tag{6}$$

Where, d is the continuous sampling space.

In most studies, the Kriging predictions are calculated on a grid where one value represents character of the whole. IKV can be calculated at the grid in discrete formula:

$$g(X) = \sum_{i=1}^{t} \sigma^2[\hat{z}(x_i \mid X)], \ \ \mathrm{x} \in \mathrm{D}, \mathrm{X} \subseteq \mathrm{D} \tag{7}$$

Where, x_i denotes the i^{th} discrete grid, t is the total number of grids, D is discrete sampling space.

B. Searching Algorithm

Let $X^0 = [x_1^0, \cdots, x_k^0]^T$ be previous locations in which the data have collected. Let $X_n = [x_1, \cdots, x_n]^T$ be adding optimal locations need to measure. Given the previous measurements, our goal of adding new measurements is to minimize the integral of the Kriging variance over the sampling space. Therefore, the objective function is defined by:

$$obj(X_n^*) = \underset{X_n \subset D}{Arg\min}[g(X_{k+n})] \tag{8}$$

Here, $X_{k+n} = [X^0; X_n]$

The sampling problem is formulated as selecting n locations out of total number of t candidate locations, which satisfies with (8). There are approximate C_t^n different ways to select n locations form t locations. When using exhaustive searching algorithm, the computing time is unacceptable to a large scales ocean sampling. For example, if a region with size 100×100km is consisted of each grid with size 1×1km, the computing time of finding 100 best grids in the region is enormous.

An often used heuristic searching algorithm is to start from an empty set of locations $X_0^* = \varnothing$, and greedily add locations one at a time until $X_n^* = [x_1^*, \cdots, x_n^*]^T$. Usually, the greedy rule makes $g(X_{k+i})$ decrease the most at iteration. Therefore, the greedy rule is defined as:

$$Obj(x_i*) = Arg\max[g(X_{k+i-1}) - g(X_{k+i})], \quad i \geq 1 \tag{9}$$

Greedy searching algorithm does not guarantee an optimal solution which is only near-optimal. Compared to other searching algorithms, however, greedy algorithm only spends $n \cdot t$ steps searching for the near-optimal solution. In section below, we will analysis the performance level of the solution by greedy algorithm in theory.

C. The Approximation Bound

For the near-optimal solution by the greedy algorithm, we will use submodularity to have a guaranteed minimum performance level when compared to the optimal solution.

Definition 1 [13] Given a finite set D, a real-valued function f on the set of subsets of D is called submodular if

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B), \ \forall A, B \subseteq D$$

Equivalently, a proposition deduced by Nemhauser[13] shows that a set function is submodular if for all $A \subseteq B \subseteq D$, $x \in D - B$, $f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B)$.

However, the function $g(X)$ does not satisfy with the submodularity. We construct a function $f(X_n)$:

$$f(X_n) = g(X^0) - g(X_{k+n}) \tag{10}$$

One can infer:

$$\begin{aligned} Obj(x_i^*) &= Arg\max[f(X_{k+i}) - f(X_{k+i-1})] \\ &= Arg\max[g(X_{k+i-1}) - g(X_{k+i})] \end{aligned} \tag{11}$$

Proposition 1 The set function $f(X_n)$ is monotonic submodular.

Proof. Set $X_{n1} \subseteq X_{n2} \subseteq D$, $x \in D - X_{n2}$. Assuming that the reduced value of adding element x to the set X_{n1} is $\delta_x(X_{k+n1})$, which is called marginal benefit

$$\begin{aligned} \delta_x(X_{k+n1}) &= f(X_{k+n1} \cup \{x\}) - f(X_{k+n1}) \\ &= g(X_{k+n1}) - g(X_{k+n1} \cup \{x\}) \end{aligned} \tag{12}$$

Form ref. [14], we can obtain two equations below:

$$\sigma^2[\hat{z}(x \mid X_n)] = \sigma^2[\hat{z}(x \mid X_{n+1})] + \lambda^2_{n+1|X_{n+1}} \sigma^2[\hat{z}(x_{n+1} \mid X_n)] \tag{13}$$

$$\lambda_{i|X_n} = \lambda_{i|X_{n+1}} + \lambda_{n+1|X_{n+1}} \lambda^*_{i|X_n}, \ i = 1, \cdots, n \tag{14}$$

Where, $\lambda_{i|X_n}$ and $\lambda^*_{i|X_n}$ denotes the weight assigned to the data when predicting $z(x)$ and $z(x_{n+1})$ by the data measured at locations X_n respectively.

$$\begin{aligned} \delta_x(X_{k+n1}) - \delta_x(X_{k+n2}) = g(X_{k+n1}) - g(X_{k+n2} \cup \{x\}) \\ -[g(X_{k+n2}) - g(X_{k+n2} \cup \{x\})] \end{aligned} \tag{15}$$

Substituting (13), (14) into (15), one can get following result:

$$\begin{aligned} \delta_x(X_{k+n1}) - \delta_x(X_{k+n2}) &= \sigma^2[\hat{z}(x \mid X_{k+n1})] \sum_{i=1}^{t} \lambda^2_{i|X_{k+n1} \cup \{x\}} \\ &\quad - \sigma^2[\hat{z}(x \mid X_{k+n2})] \sum_{i=1}^{t} \lambda^2_{i|X_{k+n2} \cup \{x\}} \\ &\geq \sigma^2[\hat{z}(x \mid X_{k+n1})] \sum_{i=1}^{t} [\lambda^2_{i|X_{k+n1} \cup \{x\}} - \lambda^2_{i|X_{k+n2} \cup \{x\}}] \\ &\geq 0 \end{aligned} \tag{16}$$

Lemma 1[13] Let f be a monotone sbumodular set function over a finite set D with $f(\varnothing) = 0$. Let X_G be set of the first n elements chosen by greedy algorithm and $OP = \underset{X_n \subset D}{Max} f(X_n)$, then

$$f(X_G) \geq (1 - \frac{n-1}{n})^n OP$$

According (11), it clearly finds that the near-optimal solution by lemma 1 is equivalent to the solution derived form greed rule (9). Therefore, the value of near-optimal solution by greedy algorithm is at least $1 - [(n-1)/n]^n$ times the optimal value. This bound can be achieved for each n and has a limiting value of $(e-1)/e$.

IV. Accelerating The Searching Algorithm

Although greed algorithm could greatly reduce the number of searching steps, the computing time is very large because choosing every near-optimal location need to calculate t in-

verse matrixes and large linear systems and search for all the candidate locations in the sampling space at iteration. We improve the computational efficiency in two ways which one changes the style of the calculating Kriging variance, another change the process of the searching for the near-optimal locations. Both ways can be combined for even more efficient computation.

A. Updated Kriging Variance Algorithm (UKVA)

Suppose that the iteration has been executed i ($i \geq 1$) steps, the goal of the $(i+1)^{th}$ iteration is to find x_{i+1}^* based on greedy rule (9). In order to find x_{i+1}^*, greedy algorithm need to search $t-i$ ($t \gg i$) steps and every step searching requires recalculating the inverse matrix H_{i+1}. For IKV minimization, where we consider calculating the inverse matrix alone, the complexity of this operation is $O(n^3)$, which consumes much computation time.

According to (5), applying updated Kriging variance equation proposed by Barnes [11], one can get following formula:

$$\sigma^2[\hat{z}(x \mid X_{i+1})] = \sigma^2[\hat{z}(x \mid X_i)] - (r_{i+1} - \hat{r}_{i+1})^2 / D \qquad (17)$$

Where, $r_{i+1} = c(x_{i+1}, x)$ is the covariance between the new location of observation and x; $\hat{r} = \Lambda_i H_i^{-1} V_i$ is the covariance between the estimation error associated with $\hat{z}(x \mid X_i)$ and x; $D = c(x_{i+1}, x_{i+1}) - \Lambda_i H_i^{-1} \Lambda_i^T$ is covariance between the errors associated $\hat{z}(x \mid X_i)$; and $\Lambda_i = [c(x_{i+1}, x_1), \cdots c(x_{i+1}, x_i), 1]$.

Substituting (17) into (9), one can obtain the formula below:

$$Obj(x_{i+1}^*) = Arg\max\{\sum_{i=1}^{t} (r_{i+1} - \hat{r}_{i+1})^2 / D\} \qquad (18)$$

At starting of iteration, calculating inverse matrix H_i only once instead of $t-i$ different inverse matrices H_{i+1} could find x_{i+1}^* based on (18). In other word, applying UKVA could reduce by $t-i-1$ calculating inverse matrix H_{i+1} at each iteration. Therefore, computation time can be reduced greatly while the results form two different methods are fully equivalent.

B. Accelerating Searching Algorithm Using Voronoi Partition (ASAVP)

Whether it is for the objective function (9) or (18), evaluation criteria are consistent, that is, the adding observations could reduce the Kriging variance the most. In other words, the goal is to find sensor placement with least uncertainty after observations. For the large sampling space, the consuming time of choosing the near-optimal sampling locations is unacceptable even using UKVA.

The property of the submodularity, for all $A \subseteq B \subseteq D$, $x \in D - B$, $f(A \cup \{x\}) - f(A) \geq f(B \cup \{x\}) - f(B)$, implies that adding a new sampling x when we only have a small set of sampling A gives us more advantage than adding to a larger set of sampling B. Using this characteristic, we can improve speed of the greedy searching algorithm by choosing the optimal sampling locations in local sampling space instead of the whole sampling space.

The question, however, is how to find such regions. We propose a method of Voronoi partition to obtain the local sampling space in which the optimal sampling lies. The Voronoi partion $V(X) = [v_1, \cdots, v_n]$ of Ω generated by the point $X = [x_1, \cdots x_n]$ is defined by:

$$v_i(x) = \{x \in \Omega \mid \| x - x_i \| \leq \| x - x_j \|, \forall j \neq i\} \qquad (19)$$

Where, let sampling space Ω be a convex polygon. Each $v_i(x)$ is called Voronoi cell. Using previous sampling points plus near-optimal locations obtained by greedy searching algorithm generate Voronoi diagram which is used to partition the sampling space. According to (19), one can infer that there is only one observation in each Voronoi cell. According to submodularity, adding a new sampling into a larger Voronoi cell gives us more advantage than adding a new sampling into a smaller one. Therefore, the optimal sampling location must lie in the largest Voronoi cell. Choosing the optimal sampling locations in the largest Voronoi cell instead of the whole sampling space can save a great deal of computing time. The searching algorithm is detailed in table Ⅰ.

Applying the ASAVP, assume that there are m previous observations and n new adding sampling points needed to be find out of t candidates, the i^{th} iteration only requires approximate $t/(m+i)$ steps searching. It clearly finds that the searching speed of average one of the near-optimal observations gets faster and faster with the number of sampling increasing because the local searching space becomes smaller and smaller.

V. Simulation Results And Disscussion

We performed simulations with the following parameters: sampling space size 100×100km, grid size 2×2km, 10 existing observations at locations (10.3,20.6), (20.2,80.3), (40.5,30.2), (40.7,90.5), (10.4,50.6), (70.3,90.5), (60.7,60.4), (80.7,40.2), (95.6,10.5), (95.3,90.6).We used the covariance function derived during the AOSN 2003 experiment in [1]:

$$c(x_1, x_2) = c_0 \exp(-d^2 / d_0^2) \qquad (20)$$

TABLE Ⅰ
ACCELERATING SEARCHING ALGOTITHM

Name: accelerating searching algorithm Input: 1. searching space Ω (convex polygon) 2. the set of earlier measurements S, $S \subset \Omega$ 3. the number of the new adding measurements, n Output: the near-optimal set of measurements, X

1: $X = \varnothing$
2: for i=1 to n
{
3: use $X \cup S$ generating Voronoi diagram in the whole sampling space Ω
4: calculate all Voronoi cells' area
5: using greedy algorithm, searching for x_i* in the largest Voronoi cell
6: $X \leftarrow X \cup x_i$*
}

Where, $d = |x_1 - x_2|$ is Euclidean distance, $d_0 = 25$ km is spatial decorrelation scale of temperature during Monterey Bay during AOSN 2003. Scaling factor c_0 has no effect on the sampling scheme and let $c_0 = 2$. All the simulations are implemented on a personal computer with hardware: Intel CPU 3.2GHz and 1GB ram and software: Matlab 2008a on Microsoft Windows XP Professional.

A. *Comparing the Results between the Improved and Conventional Algorithm*

1) The Results between UKVA and Basic Kriging: As can be seen from Fig.1, the distribution of variance computed by basic Kriging and UKVA, respectively, is exactly the same based on 10 existing observations. However, the computing time is very different, which is detailed in section B.

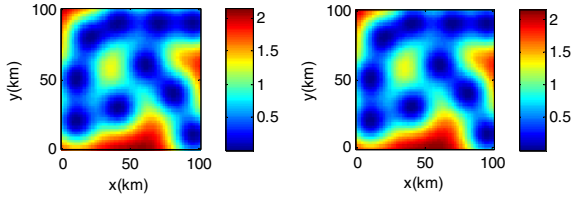

Fig.1 The distribution of the variance; left: variance computed by basic Kriging; right: variance computed by UKVA

2) *The Results of ASAVP and Basic Algorithm*: Fig.2 shows that the sampling point searched by ASAVP at iteration lies in the Voronoi cell whose area is the largest. It obviously finds that the new sampling points searched by ASAVP are nearly consistent with the sampling point searched by basic algorithm in the whole sampling space except for at 4^{th} iteration. The boundary of the largest Voronoi cell traversing the grid in which the near-optimal sampling point lie results in the inconsistency between sampling point searched by ASAVP and basic algorithm. Two ways is advised to avoid such case. The one is to reduce the grid size; the other is to increase a step searching if the optimal sampling point is on boundary of the Voronoi cell. Fig.3 shows the distribution of the variance in the whole sampling space computed based on the optimal sampling points plus 10 earlier observations.

B. *Comparing the Computing time*

We divide the optimizing process into two phases: computing Kriging variance and searching for the optimal sampling location. In the first phase, we compute the Kriging variance using UKVA and basic algorithm respectively. And we search for the optimal sampling location by ASAVP and basic searching algorithm separately in the second phase. The computing time using different methods searching for one and five new adding sampling points is showed in table Ⅱ. When using the classical method (Basic + Basic), it obviously find that the total computing time is proportional to number of sampling points. When using improved method (UKVA + ASAVP), however, average computing time of searching for one sampling point decreases as the number of sampling points increasing. The more sampling points are added, the searching space is smaller, and therefore, the searching speed of average one of the optimal observations is faster and faster.

VI. Conclusions And Future Work

In this paper, it is has been shown how sampling schemes can be optimized by minimizing the integrated Kriging variance. The main contribution of this paper lies in two aspects. One is that we use submodularity to get an approximately bound of near-optimal solution by greedy algorithm. On the other, we propose two approaches to reduce the computation time of searching for the near-optimal sampling location based on known covariance (variogram). The main advantage of improved algorithms is that the searching speed of average one of the near-optimal sampling gets faster and faster with the number of sampling increasing. Therefore, this algorithm is very suitable for a large scales ocean sampling.

While the greedy searching algorithm certainly does not guarantee an optimal solution, but it dose provide a good first guess for other algorithm, such as sequential exchange algorithm, which is our next work. In addition, how to deploy mobile sensor platforms, such as underwater glider, following the optimal path connected the best sampling locations searched by our proposed algorithm to measure the ocean character is our next future work.

TABLE Ⅱ
COMPARING THE COMPUTING TIME

Type of methods	Consuming time (second)	
	1 sampling point	5 sampling points
Basic + Basic	148.9688	746.9625
UKVA + Basic	66.7344	334.8653
Basic + ASAVP	25.8906	98.5469
UKVA + ASAVP	13.8125	41.4219

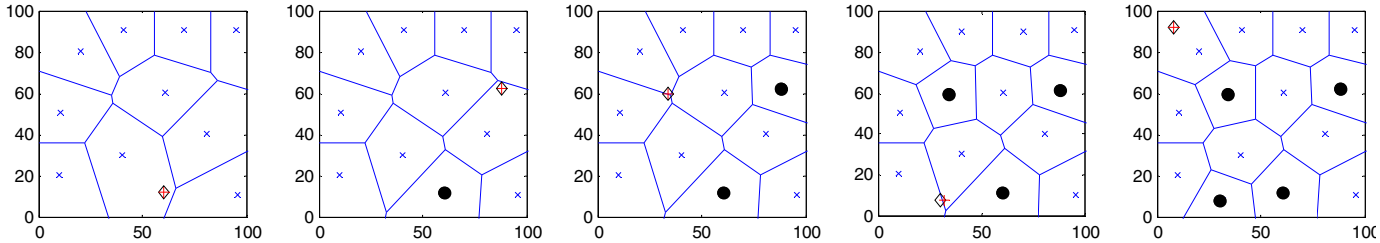

Fig.2 The process of searching for 5 near-optimal sampling locations; × and ● indicate the existing observation locations and the near-optimal sampling locations in previous stage respectively; ◇ and + indicate near-optimal sampling location at i^{th} iteration by ASAVP and basic algorithm respectively.

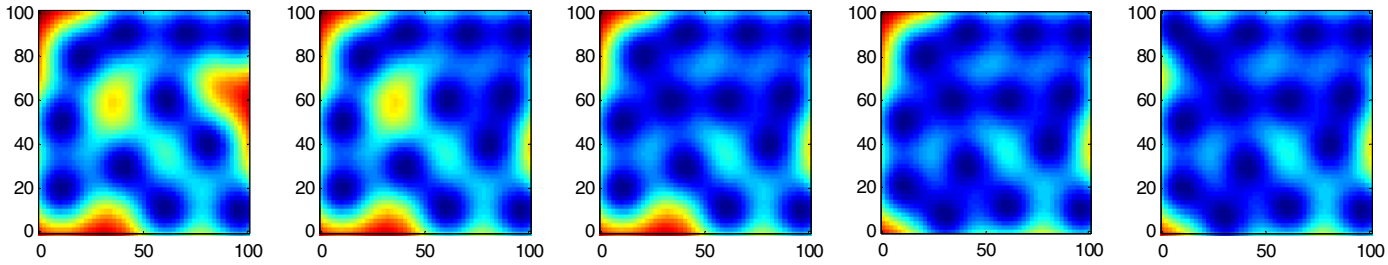

Fig.3 The distribution of the variance in the whole sampling space at iteration

REFERENCES

[1] N. E. Leonard et al., "Collective motion, sensor networks and ocean sampling," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 48-74, 2007.

[2] E. Fiorelli et al., "Multi-AUV control and adaptive samplingin Monterey Bay," *in Proceedings IEEE AUV 2004: Workshop on Multiple AUV Operations*, 2004.

[3] B. Zhang and G. S. Sukhatme, "Adaptive Sampling for Estimating a ScalField using a Robotic Boat and a Sensor Network," *Robotics and Automation, 2007 IEEE International Conference,* April 2007, pp. 3673-3680.

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

[5] R. A. Olea, "A six-step practical approach to semivariogram modelling," *Stoch Environ Res Risk Assess*, vol. 20, pp. 307-318, 2006.

[6] N. Cressie, "Fitting variogram models by weighted least squares," *Mathematical Geology*, vol. 17, no. 5, pp. 563-586, 1985.

[7] P. M. Atkinson and C.D. Lloyd, "Non-stationary variogram models for geostatistical sampling optimisation: An empirical investigation using elevation data," *Computers & Geosciences*, vol. 33, pp. 1285–1300, 2007.

[8] J. W. van Groenigen, "The influence of variogram parameters on optimal sampling schemes for mapping by Kriging," *Geoderma,* vol. 97, pp. 223-236, 2000.

[9] A. B. McBratney, R. Webster, T.M. Burgess, "The design of optimal sampling schemes for local estimation and mapping of regionalized variables: I. theory and method," *Comput.Geosci*, July 1981, pp. 335-365.

[10] J. W. van Groenigen, W. Siderius and A. Stein, "Constrained optimisation of soil sampling for minimisation of the Kriging variance," *Geoderma,* vol. 87, pp. 239-259, 1999.

[11] R.J. Barnes and A.G. Watson, "Efficient updating of Kriging estimations and variances", *Mathematical Geology,* vol. 24, no.1, pp. 129-133, 1992

[12] H. Gao, J. Wang and P. Zhao, "The updated Kriging variance and optimal sample design," *Mathematical Geology,* vol. 28, no. 3, pp. 295-313, 1996.

[13] G.L. Nemhauser and L.A. Wolsey, "An analysis of approximations for maximizing submodular set functions-I," *Mathematical Programming*, 14, pp. 265-294, 1978

[14] X. Emery, "The kriging update equations and their application to the selection of neighboring data," *Comput Geosci*, **no.** 13, pp. 269–280, 2009