



# Planning Placement of Distributed Sensor Nodes to Achieve Efficient Measurement

Yuichi Nakamura<sup>(✉)</sup>, Masaki Ito, and Kaoru Sezaki

Institute of Industrial Science, The University of Tokyo,  
Komaba 4-6-1, Meguro, Tokyo, Japan  
y-nakamura@mcl.iis.u-tokyo.ac.jp

**Abstract.** This paper proposes a method to plan a placement of multiple sensors distributed in a certain area to enable an efficient measurement in terms of the confidence of the interpolation of measured data using kriging. We considered a system where we have some sensors that can move and are distributed in a certain area and a static scaler filed of interest such as a map of temperature in a certain city. We propose a method to plan the placement for the next time step using the value measured until that time by calculating the gradient of kriging variance. For the sake of evaluation of this method, we conducted a simulation of two-step measurement where a scaler filed is created and some sensor nodes are virtually placed. Here, the interpolation with the data from sensor node placement with the proposed method showed better accuracy than that from randomly placed sensors.

**Keywords:** Active sensing · Mobile sensing · Kriging

## 1 Introduction

Smartphones have become very common these days and the percentage of the smart phone holders in Japan has risen from 14.6% to 56.8% in 5 years [1]. Due to the advancement in semiconductor and other related technologies, smartphones today are equipped with more various types of sensors and it can be interpreted that there are an increasing number of sensors connected to the Internet in cities. Consequently, mobile sensing, which is a technique to make use of them, has drawn much attention.

Mobile sensing is useful to conduct a measurement of a data in a wide area in a low cost. For example, when one tries to create a heat map of the temperature in one city, in case of mobile sensing, the cost required is less than conventional fixed sensors because s/he doesn't have to implement all the sensors but has to collect sensor data from smart-phone holders.

In this paper, the focus has been set to a sensing system, or broader class of measurement using mobile sensors, which of course includes smartphones. For example, if garbage collecting trucks are equipped with sensors, that would provide with virtually exhaustive coverage of the city [2]. Automobiles can move faster than a human being and so it is expected that they can ensure broader coverage per sensor.

Although the data acquired from a sensing system can be useful as mentioned above in an example of mobile sensing, such data have to be processed in order for the user to utilize

because the raw data acquired is merely a collect of measured values in discrete points. For example, if a temperature measurement is conducted, it is not such discrete data, but a heat map of temperature that is of use. In other words, a sensing system requires spatial interpolation, or spatially completing the estimated values at unvisited points.

In spatial interpolation, the accuracy has dependency on the placement of sensors. As is intuitively understood, measuring values at two points close would give lower accuracy of spatial interpolation than at two points apart. Therefore, we propose a placement optimization strategy of sensors in a sensor system. In our proposal, we use a technique called kriging for spatial interpolation and kriging variance for the placement strategy.

The rest of the paper is organized as follows. Section 2 describes the background to this paper; Sect. 3 introduces the problem settings of this study; Sect. 4 gives details of the proposed method; Sect. 5 shows the simulation that we did; Sect. 6 shows the prototyping and the direction of future research; and Sect. 7 is the conclusions.

## 2 Related Works

Sensing system or especially mobile sensing has attracted a lot of attention and it is an active area of study and development. As mentioned above, previous research has demonstrated some actual implementation of such system and the method of spatial interpolation, but optimal sensors placement strategies are limited.

The sensor system is implemented for a broad area. In [2], garbage collecting trucks in Fujisawa City, Japan are equipped with sensors such as a GPS sensor, thermometer, barometer and so on. Unlike private vehicles, those trucks go around the city in such a route that covers the entire road. Through actual implementation, they checked that most area of the city can be sensed at least once every garbage-collecting day. In [3], they used pictures taken by a camera installed on automobiles to detect damage on the road. In [4], the acoustic noise level of entire Setagaya City, which is about 60 km<sup>2</sup>, has been sensed in 6 h by 40 people. In [5], the participants walked around city with a mobile radiation detector.

For this kind of measurement, interpolation is important. The authors of [6] used a technique named consensus filter. In consensus filter, for each point that sensor nodes did not visit, the value is estimated by waited sum of the values at neighboring visited points. This method is useful for relatively densely sensed situation but not for sparsely measured case. There are also studies that introduces methods using other background data. In [7], they spatially interpolated the temperature data taking into account other geographical knowledge such as the altitude. In [8], they only used 12 sensors to measure AQI, or air quality index, in the entire LA. Combined with contextual data of geography of each points from OpenStreetMap and used the measured data for learning, their estimation was better than conventional IDW, or inversed distance weight, in cross validation.

There are several works of active sensing with sensing robots such as [9]. They proposed a route optimization for a sensing robot. Here, mutual information of the measurement and interpolation is used to determine the trajectory of a sensing robot. In this work, the focus is set on the case of only one sensing node.

As we have listed above, there are a wide variety of application of sensing system and also a number of interpolation methods but few works have covered the strategy of placing sensors, especially multiple sensors. In this paper, we propose a method to interpolate measured data from multiple sensor nodes and plan the placement of them in the next step to achieve the most optimal placement in terms of the accuracy of the interpolation.

### 3 Problem Settings

In this paper, we propose a method to plan an optimal sensor placement method to a situation as follows:

One would like to map a value of interest in a certain area, such as temperature in Tokyo. For that purpose, multiple movable sensors are deployed in the area. Each sensor conducts the measurement at the same time and send the data to a server. Then each of them moves to the next sensing position, instructed by the server, and measure the value there again.

Here we assumed the scaler field of the true value of the vale of interest do not change between the time of the first measurement and the second. This is because even though the scaler field does change in the time length, in one step of the move the time is short and the change is negligible in the real scenario. In the real use, the measurement will be conducted sequentially beyond the second time but in this paper, two steps of measurement are discussed for simplicity.

## 4 Method

### 4.1 Kriging

Kriging is a geostatistical method to interpolate data using spatial autocorrelation. For a system where sensor nodes  $i$  ( $i = 1, \dots, N$ ) are deployed in  $\mathbf{r}_i$  the field and measurement  $z(\mathbf{r}_i)$  is conducted, semivariance, or the spatial autocorrelation of  $z$  is defined as

$$\gamma(\mathbf{h}) = \frac{1}{2n} \sum_{i=1}^n [z(\mathbf{r}_i + \mathbf{h}) - z(\mathbf{r}_i)]. \quad (1)$$

For actual data, since the measured value at an arbitrary point is not always available, semivariance cannot be calculated for all  $\mathbf{h}$ . It is why variogram, or a plot of  $\gamma(\mathbf{h})$  versus  $h$ , only for known points. By fitting variogram with certain function, we can get the estimated semivariance for arbitrary  $\mathbf{h}$ . There are several common functions used to fit variogram, such as spherical model, exponential model and so on. In this paper, we used the exponential model in this case, which is represented as follows:

$$\hat{\gamma}(h) = p \left[ 1 - \left( 1 - \frac{3h}{r} \right) e^{-\frac{3h}{r}} \right] + n \quad (2)$$

where  $r$  is called range and  $p$  is called sill and  $n$  is called nugget; those three parameters are determined by the fitting.

After getting the estimation of semivariance, we now use this to conduct the interpolation. In kriging, the value is interpolated through weighted sum

$$\hat{z}(r) = \sum_{i=1}^N \lambda_i z(r_i) \quad (3)$$

where the weights  $\lambda_i$  are bound to the normalization

$$\sum_{i=1}^N \lambda_i = 1. \quad (4)$$

In order to minimize the estimation error, which is a sum of the gap squared between the estimated value and the weights at each visited point, are determined using semi-variance, in such a way as minimizing

$$\sum_{i=1}^N \lambda_i \hat{\gamma}(r_i - r_j) + \phi = \hat{\gamma}(r_j - r), \quad \forall j. \quad (5)$$

## 4.2 Kriging Variance Gradient

Then the kriging variance, or the variance of the estimated value, is expressed as:

$$\hat{\sigma}_e^2(\mathbf{r}_0) := \sum_{i=1}^N \hat{\lambda}_i \gamma(\mathbf{r}_i - \mathbf{r}_0) + \phi. \quad (6)$$

From the method above we can obtain the kriging variance at all points including unvisited ones as well as estimated value there. In order for each sensor nodes to move towards the direction so that the next measurement would be the most optimal with regard to minimizing the error for the whole area, we also calculated the gradient  $\mathbf{G}(\mathbf{r})$  of the kriging variance:

$$\mathbf{G}(\mathbf{r}) = \nabla [\hat{\sigma}_e^2(\mathbf{r})]. \quad (7)$$

## 4.3 Placement Determination

It is most optimal to move the sensor nodes towards the direction where the kriging variance is the highest. We used the gradient to find such a direction. Therefore, we determined the placement of the sensor  $i$  for the next time step  $t + \Delta t$  using the gradient vector  $\mathbf{G}(\mathbf{r})$  at time  $t$  as follows:

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta r \frac{\mathbf{G}(\mathbf{r})}{|\mathbf{G}(\mathbf{r})|} \quad (8)$$

where in each step each sensor node moves for the distance  $\Delta r$ .

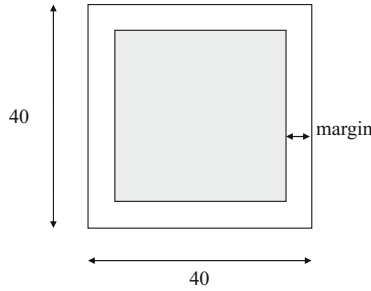
## 5 Simulation

In order to test the validation of the above method, a simulation was conducted and the mean squared error of estimated values at all points in the area was compared with the random determination.

### 5.1 Setting of the Simulation

In the simulation, we made a certain scalar field, which is the true value in this simulation, in the area of interest and interpolated the measured values at five sensor nodes. The simulation was conducted in the following settings.

The area of interest has been formed as  $40 \times 40$  square, filled with  $0.5 \times 0.5$  meshes. The five sensors were deployed in the sensing area, or the grey area in Fig. 1, so that it is not located on the edge of the  $40 \times 40$  square, where gradient of kriging variance is not well-defined.



**Fig. 1.** The field of simulation is  $40 \times 40$  including the margin.

The scalar field we prepared was a sum of three different Gaussian with different means  $\mu$ 's and covariance matrices  $\Sigma$ 's as:

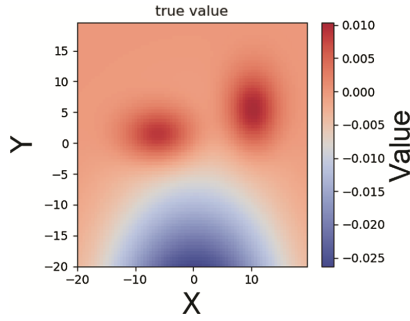
$$z(\mathbf{r}) = \sum_{i=1}^3 \frac{1}{2\pi|\Sigma_i|^{\frac{1}{2}}} \exp \left[ -\frac{1}{2} (\mathbf{r} - \boldsymbol{\mu}_i)^\top \Sigma_i^{-1} (\mathbf{r} - \boldsymbol{\mu}_i) \right]. \quad (9)$$

which is shown in Fig. 2.

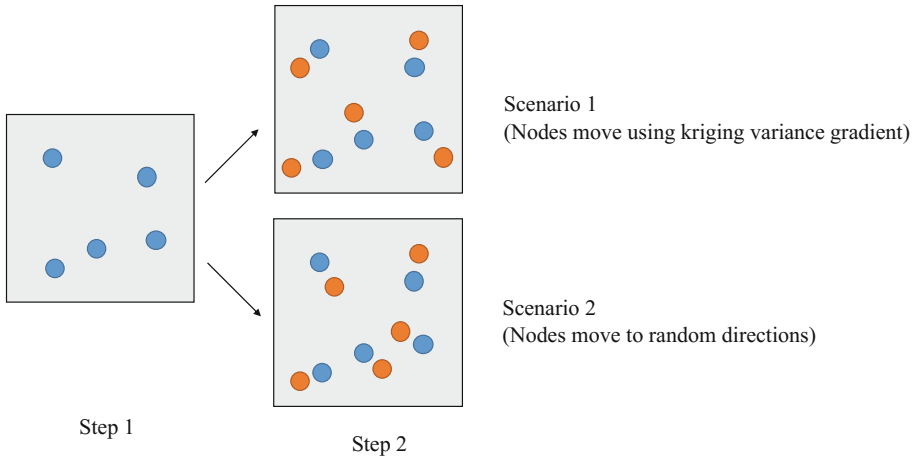
In the simulation, first we deployed five sensors at random in the sensing area and made a measurement (step 1). We then kriged the measured values to obtain the map of the value in the entire field and also the kriging variance map (Fig. 3).

### 5.2 Method

Now we have the kriging variance field and so we used that to determine the placement of the sensor nodes for Step 2. We set two different scenarios.



**Fig. 2.** We made 2-D scalar field by summing three different Gaussians.



**Fig. 3.** In Step 1, we deployed five sensors randomly. In Step 2, we simulated two different scenarios: in Scenario 1, the sensors were moved using kriging variance gradient; in Scenario 2, they were moved to random directions. In Step 2, both data from 5 initial points and 5 new points were used in the kriging.

*Scenario 1.* It is most optimal to move the sensor nodes towards the direction where the kriging variance gradient is the highest in all the directions. Therefore, we determined the placement of the sensor  $i$  for the next time step as follows:

$$r_i(t + \Delta t) = r_i(t) + \Delta r \frac{\nabla V(\mathbf{r})}{|\nabla V(\mathbf{r})|} \tag{10}$$

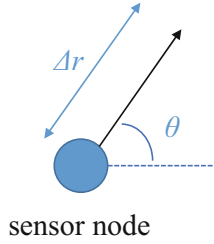
where  $\Delta r$  is the step of the movement.

Practically, in this simulation the kriging was done for discrete grid points. Therefore variance calculated only for discretely distributed points and the gradient is not actually calculated through differentiation but computed using second-order accurate central differences with certain sampling distances. Since the resulting gradient has a dependency on sampling distance  $d_{sample}$ , we compared the results for different  $d_{sample}$ .

*Scenario 2.* For comparison, we also simulated the case where the nodes are moved to a random direction. For this the position of the same sensor node at the next time step is expressed as:

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta r \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad (11)$$

Here,  $\theta$  is a random value between 0 and  $2\pi$  (Fig. 4).



**Fig. 4.** Each sensor node moves by  $\Delta r$  in the direction  $\theta$ .

### 5.3 Evaluation

For the evaluation of these methods, we calculated the sum of error squared at each grid point of the entire area of interest, namely,

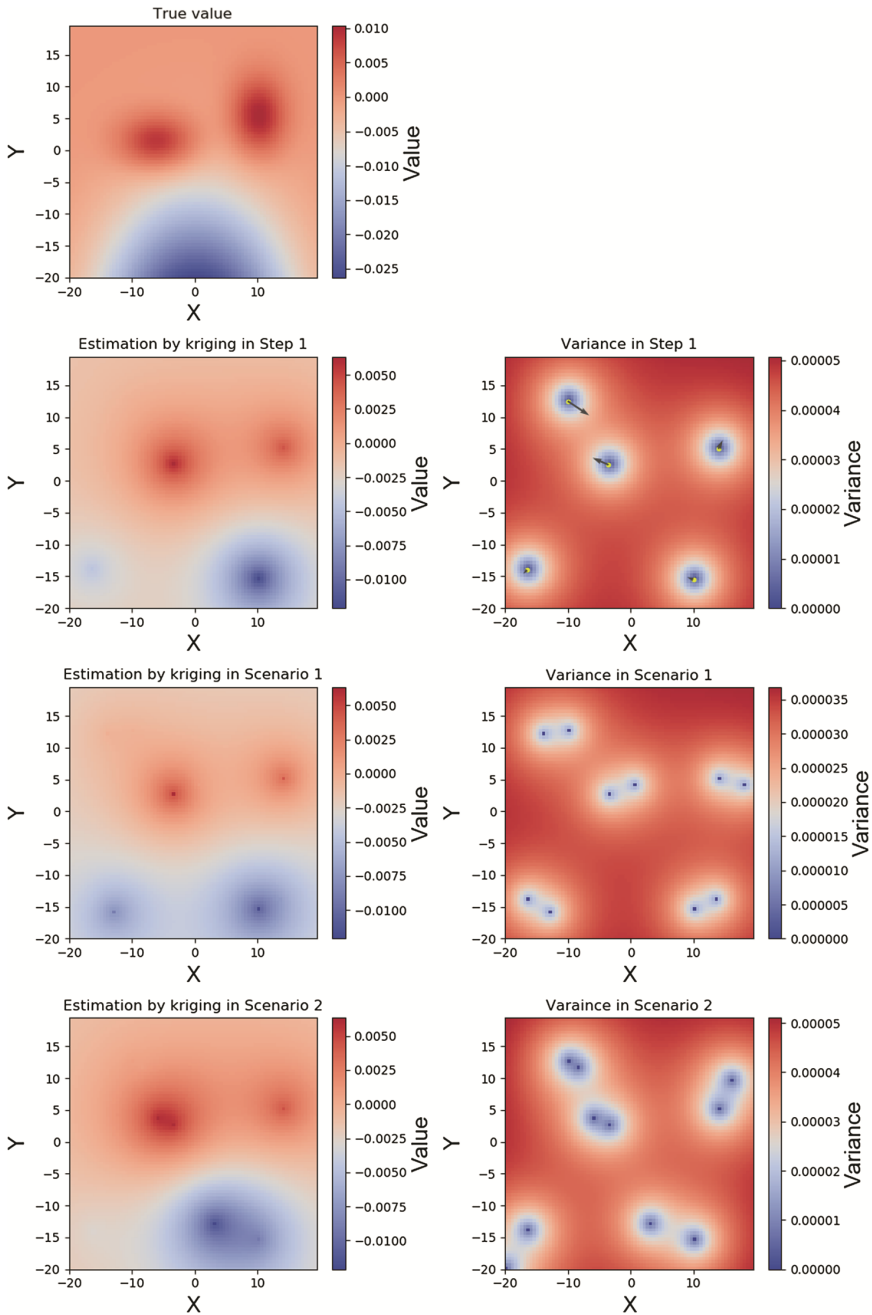
$$\text{Error} = \sum_j [z(\mathbf{r}_j) - \hat{z}(\mathbf{r}_j)]^2 \quad (12)$$

where  $z(\mathbf{r}_j)$  and  $\hat{z}(\mathbf{r}_j)$  are the true value and the interpolated value at the grid point  $j$ , respectively and the summation is done through all the grid points.

In order to examine the effect of  $\Delta r$ , we executed the simulation for different  $\Delta r$ . Also, as mentioned above, scenario 2 was performed for different  $d_{\text{sample}}$ .

In other words, for each initial random sensor deployment (step 2), (1) scenario 1 was performed for different  $\Delta r$ 's and (2) scenario 2 was performed for different  $\Delta r$ 's and different  $d_{\text{sample}}$ 's.

We calculated the error for each scenario with different parameters for 1000 different initial placement of sensors in Step 1. For each error in Scenario 1 was compared with the error in the corresponding simulation in Scenario 2 with regard to  $\Delta r$  and we counted the number of cases where error in scenario 1 is greater than that in scenario 2 out of 1000 simulations for each  $\Delta r$  and  $d_{\text{sample}}$ .



**Fig. 5.** An example of the simulation results. The left figure of the top row is the true value. The second row show is the kriged result (left) and kriging variance with the positions of sensor nodes (yellow dots) and the gradient (arrows). The third row shows the results of Step 2 in Scenario 1 and the forth that in Scenario 2. (Color figure online)



### 5.4 Results

Figure 5 shows the true value and the result of kriging at each step. Figures on the left column are the true value or interpolation result of different steps, the true value on the top figure, Step 1 on the second, Step 2 in Scenario 1 on the third, Step 2 in Scenario 2 at the bottom. On the right column are the kriging variance.

Figure 6 shows the result of the comparison of the error in Scenarios 1 and 2. The x axis indicates  $\Delta r$  and y axis  $d_{\text{sample}}$ . The color bar shows the ratio of the cases where Scenario 1 gives less error than 2, over 1000 simulations. The resulting figure indicates that for certain  $\Delta r$  and  $d_{\text{sample}}$ , Scenario 1 gives more accurate kriging results than Scenario 2 in more than 50% of the cases. Given 100 trials have been done, if the ratio from the result is greater than 51.6%, the actual ratio will be more than half, with 95% confidence. Therefore, the fact that there are cases that the ratio is over 51.6% shows that this proposed method gave significantly better results than random decision.

The ratio is higher especially when  $\Delta r$  is smaller and  $d_{\text{sample}}$  is greater.

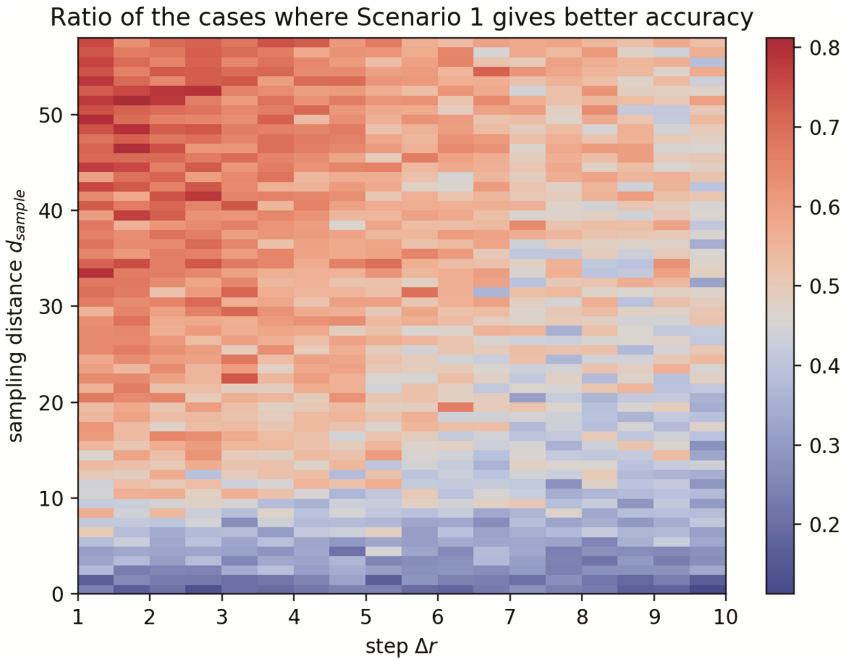


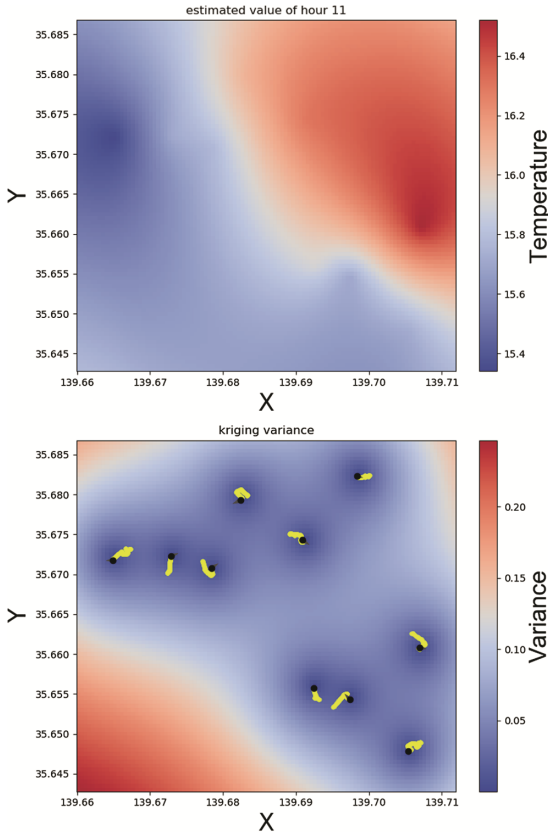
Fig. 6. The ratio of the cases where Scenario 1 gives better accuracy than Scenario 2.

## 6 Prototype

We also utilized the proposed method to actual measured data from the previous work [10]. Here, some portable devices were prepared and used to make measurement around Shibuya City, Tokyo. They are equipped with multiple sensors including thermometer,

hygrometer and air quality sensors such as  $O_3$ , CO,  $NO_2$ , PM2.5. The devices were made compactly so that human beings can hold while walking. Participants of the experiments were asked to walk around the city carrying one of those. 12 people took part in the experiment and those series of data were collected.

Among those data, the atmospheric temperature was picked and kriged. The results of the kriging and the variance are shown in Fig. 6. Here, with the proposed method, as shown by arrows, the next suggested movement of sensors was successfully calculated (Fig. 7).



**Fig. 7.** The result of kriging of the atmospheric temperature data in Shibuya City, Tokyo in 11:00–11:10am on Nov. 14, 2016. The figure on the top is the interpolated map of the temperature and the one below is the kriging variance with gradient shown by arrows. The yellow line in the figure below is the trajectory of each sensor node. (Color figure online)

## 7 Conclusion

In this paper, we proposed a method of sensor placement in a multiple distributed sensor system in order to achieve better accuracy of interpolation using kriging. Then we

demonstrated the method in a simulation and compared with moves with randomly determined direction for validation. As a result, the proposed method was proven to give better accuracy of interpolation in a certain situation, that is, small move step and relatively long sampling distance. We also used actual data from a prototype of a multiple distributed sensor system to execute the proposed method to plan the next step of movement.

As future work, we have three tasks. First one is to extend this method to multiple steps. In this paper we merely focused on two-step measurement. For real usage, it is vital to plan the best trajectory beyond the next move, where it is important not to move different sensors to the same place. This could be realized by introducing a two-body repulsive power among sensors. Second, we have to think about the time dependency of the scalar field of interest. In the simulation above we set a fixed scalar field with regard to time but in the actual world, most values that are measured is not time-invariant. For the last but not least, the method should be carried out for actual sensor placement and movement. By using the prototype mentioned above, it can be achieved.

## References

1. 総務省: 情報通信白書. 日経印刷, Tokyo (2017)
2. Chen, Y., Nakazawa, J., Yonezawa, T., Kawasaki, T., Tokuda, H.: An empirical study on coverage-ensured automotive sensing using door-to-door garbage collecting trucks. In: Proceedings of the 2nd International Workshop Smart - SmartCities 2016, pp. 1–6 (2016)
3. Kawasaki, T., et al.: A method for detecting damage of traffic marks by half celestial camera attached to cars. In: Proceedings of the 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (2015)
4. Sezaki, K.: 都市をセンシングする. 生産研究, **65**(3), 691–697 (2013)
5. Sullivan, C.J.: Radioactive source localization in urban environments with sensor networks and the Internet of Things. In: 2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), pp. 384–388 (2016)
6. Xiao, L., Boyd, S., Lall, S.: A scheme for robust distributed sensor fusion based on average consensus. In: Fourth International Symposium on Information Processing in Sensor Networks, IPSN 2005, pp. 63–70 (2005)
7. Holden, Z.A., et al.: Development of high-resolution (250 m) historical daily gridded air temperature data using reanalysis and distributed sensor networks for the US Northern Rocky Mountains. *Int. J. Climatol.* **36**(10), 3620–3632 (2016)
8. Lin, Y., Pan, F., Habre, R.: Mining public datasets for modeling intra-city PM 2.5 concentrations at a fine spatial resolution. In: SIGSPATIAL 2017 (2017)
9. Le Ny, J., Pappas, G.J.: On trajectory optimization for active sensing in Gaussian process models. In: Proceedings of the 48th IEEE Conference Decision Control Held Jointly with 2009 28th Chinese Control Conference, pp. 6286–6292 (2009)
10. Suzuki, T., Ito, M., Sezaki, K.: Evaluating estimation methods of reconstruction accuracy of negative surveys using field survey data of air pollution. IEICE Technical report, vol. 116, no. 488, pp. 85–87 (2017)