

An Experimental Environment for Optimal Spatial Sampling in a Multi-Robot System

ANSSI KEMPPAINEN^a, TONI MÄKELÄ^a, JANNE HAVERINEN^a and
JUHA RÖNING^a

^a*Department of Electrical and Information Engineering, University of Oulu, Finland*

Abstract.

In our research, we are concerned with sensing the environment using mobile robots. This enables selection of optimal sampling locations in order to produce maximum information about the environment. Selection of sampling locations plays a key role in hospital environments, for example, where humidity and temperature levels or carbon dioxide concentration may require regular monitoring. On the other hand, the accuracy of a regression model depends on the sampling locations, which is significant, for example, in planetary exploration.

In geostatistics optimal spatial sampling strategies search for sensor locations that produce minimal variance in estimates with a restricted number of sensors. Minimal variance is achieved, for example, by minimizing the conditional entropy of unobserved locations, where the environment is modelled using Gaussian processes.

In this paper, we propose an experimental environment for optimal spatial sampling using mobile sensors. Mobility reduces the reliability of sampling locations due to odometer failures, which again reduce the likelihood of the model. We have resolved this problem by building an experimental environment where a ceiling camera vision system provides multi-robot localization in an area measuring approximately 240 x 160 cm. Preliminary experiments compare optimal spatial sampling for both stationary and nonstationary models using scalar measurements of ambient light and magnetic flux density.

Keywords. Non-stationary Gaussian Processes, Distributed Sensing, Environment Modelling

Introduction

Gaussian processes have a long history in geostatistics in the context of interpolation, where the method is known as kriging [10]. Later, the machine learning community adopted Gaussian processes [16] for regression problems, where they enabled reasonable interpretations for predictions of models, yet provided a simple and sufficient choice of models. In the field of robotics, Gaussian processes have recently attracted attention in terrain modelling [9], for example, since they allow prior knowledge of the environment to be taken into account in the model.

Gaussian processes are also useful in distributed sensing scenarios, like in hospital environments, where humidity and temperature levels or carbon dioxide concentration

may require regular monitoring with a restricted number of sensors. On the other hand, in spatial sampling, for example in planetary exploration, the accuracy of the regression model depends on the sampling locations. In our research, we are concerned with sensing the environment using mobile robots. This enables selection of optimal sampling locations in order to produce maximum information about the environment.

In geostatistics, optimal spatial sampling [3] searches for sensor locations that produce minimal variance in estimates with a restricted number of sensors. Correlations between the samples are modelled using a Gaussian process [14], which is a generalization from probability distributions to functions: a mean function describes the expected value and a covariance function, the variance in an estimate at a given location. Minimal variance is obtained by minimizing the conditional entropy of unobserved locations [15], which is analogous to maximizing the entropy of sensor locations. An alternative approach is to maximize mutual information [8] between samples, which leads to samples focused closer to the centre of the sampling area.

In this paper we propose an experimental environment for a study on optimal spatial sampling strategies using mobile robots. The experimental environment is a restricted area where mobile robots can measure ambient light and magnetic flux density and place the measurements onto maps. Mapping and optimal sampling use a centralized control system aided by roof camera-based localization. Control and sensor information between the robots and the control system is supplied through radio communication. Section 1 briefly reviews the theory of Gaussian process regression and Section 2 introduces information about theoretical criteria in optimal spatial sampling. Section 3 presents the experimental environment with some preliminary results of optimal spatial sampling with two different data sets: one from ambient light and the other from magnetic flux density measurements.

1. Gaussian processes

A Gaussian process [14] can be seen as a collection of candidate functions that model the environment. The range of functions is restricted by a covariance function and the most probable function is described by a mean function. Formally, a Gaussian process is defined by a mean function (1)

$$\mu(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})] \quad (1)$$

and a covariance function (2)

$$c(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - \mu(\mathbf{x}))(f(\mathbf{x}') - \mu(\mathbf{x}'))] \quad (2)$$

where $f(\mathbf{x})$ is a random function obeying the Gaussian process (3).¹

$$f(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), c(\mathbf{x}, \mathbf{x}')) \quad (3)$$

We can imagine a function $f(\mathbf{x})$ as a finite dimensional vector \mathbf{y} , where each random variable y^i in some orthonormal basis \mathbf{e}_i corresponds to a function value $f(\mathbf{x}_i)$. Thus,

¹Bold lowercase characters denote vectors and uppercase matrices. Not bolded italic characters stand for scalars and constants.

a Gaussian process defines a joint distribution of random variables $p(y^1, y^2, \dots, y^n)$ with a mean value vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_n)$ and a covariance matrix \mathbf{C}_{nm} with elements $c_{ij} = cov(y^i, y^j)$. In order to define a continuous covariance function, we can use integral operators to define the correlation between bases. Thus, the covariance among function values is given by

$$\begin{aligned} cov(y^i, y^j) &= \mathbb{E}[(f(\mathbf{x}_i) - \mu(\mathbf{x}_i))(f(\mathbf{x}_j) - \mu(\mathbf{x}_j))] \\ &= \mathbb{E}[(f(\mathbf{x}_i) - \mu(\mathbf{x}_i))k(\mathbf{x}_i, \mathbf{x}_j)(f(\mathbf{x}_j) - \mu(\mathbf{x}_j))] \\ &= var(y^i)k(\mathbf{x}_i, \mathbf{x}_j) \end{aligned}$$

where the kernel function $k(\mathbf{x}_i, \mathbf{x}_j)$ must be symmetric, positive and definite.

1.1. Gaussian process regression

A Gaussian process defines a prior probability distribution over functions before any observations are made. This leads to a Bayesian inference, which in this context is known as a Gaussian process regression. In the Gaussian process regression we try to estimate the most likely posterior distribution over functions $f(\mathbf{x})$, conditioned on measurement function $z(\mathbf{x})$. For a multivariate Gaussian distribution we have that the conditional distribution is given by a conditional mean (4)

$$\hat{\mathbf{y}}_n = \boldsymbol{\mu}_n + \mathbf{C}_{nm}(\mathbf{C}_{mm} + \sigma_z^2 \mathbf{I}_{mm})^{-1}(\mathbf{z}_m - \boldsymbol{\mu}_m) \quad (4)$$

and a conditional variance (5)

$$var(\hat{\mathbf{y}}_n - \mathbf{y}_n) = \mathbf{C}_{nn} - \mathbf{C}_{nm}(\mathbf{C}_{mm} + \sigma_z^2 \mathbf{I}_{mm})^{-1} \mathbf{C}_{mn} \quad (5)$$

where σ_z^2 is the measurement variance (independent measurements) and m is the number of measurements.

1.2. Stationarity

A Gaussian process is said to be stationary if its mean and covariance functions are invariant to translations, i.e., $\mu(\mathbf{x}) = A$ and $c(\mathbf{x}, \mathbf{x}') = c(\mathbf{x} - \mathbf{x}')$. The stationarity assumption is somewhat artificial, since nature provides spatial information that is inherently irregular but yet continuous. In order to obtain a more accurate model of the environment, one could try to learn the parameters of the model based on experience. This leads to non-stationary Gaussian processes, where the mean function is defined by parameters and the covariance function, by hyperparameters. A covariance function restricts models to those that do not violate the definition of symmetric positive definiteness.

The literature [14] provides many non-stationary versions of covariance functions. In this study we adopted the *HSK* covariance function (6) [6], [13]

$$c(\mathbf{x}_i, \mathbf{x}_j) = \int_{\mathbb{R}^2} k(\mathbf{x}_i, \mathbf{u})k(\mathbf{x}_j, \mathbf{u})d\mathbf{u} \quad (6)$$

which uses local kernel functions $k(\mathbf{v}, \mathbf{u})$ centred on location \mathbf{v} . If each of these are given through Gaussian kernels, i.e,

$$k(\mathbf{v}, \mathbf{u}) = (2\pi)^{-1/2} |\boldsymbol{\Sigma}_v|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{v} - \mathbf{u})^T \boldsymbol{\Sigma}_v^{-1} (\mathbf{v} - \mathbf{u})\right)$$

the covariance function can be given as (7)

$$c(\mathbf{x}_i, \mathbf{x}_j) = S(\mathbf{x}_i, \mathbf{x}_j) \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^T (\boldsymbol{\Sigma}_i + \boldsymbol{\Sigma}_j)^{-1} (\mathbf{x}_i - \mathbf{x}_j)\right) \quad (7)$$

where $S(\mathbf{x}_i, \mathbf{x}_j) = (2\pi)^{-1/2} |\boldsymbol{\Sigma}_i + \boldsymbol{\Sigma}_j|^{-1/2}$.

2. Optimal spatial sampling

Optimal experimental design for spatial sampling [3] applies the theory of optimum experimental designs [1] to stochastic processes, typically to Gaussian processes. The problem is to select the sampling locations $\mathcal{X}^* = \{\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_m^*\}$ that provide maximum information about the environment. In stationary Gaussian processes, optimal sampling is suggested to correspond to regular sampling with evenly distributed sampling locations [12]. In non-stationary processes one could intuitively argue that in optimal sampling the sampling locations should be distributed so that we get more information from areas where the data are less spatially correlated, and vice versa.

2.1. Maximum Entropy Sampling

In maximum entropy sampling [15] the objective is to minimize uncertainty in unobserved locations. Now we state that the variables \mathbf{u} and \mathbf{v} are members of observed \mathcal{U} and unobserved \mathcal{V} sets, respectively, and for the sets we have that $\mathcal{V} = \mathcal{X} \setminus \mathcal{U}$ where $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots\}$. Minimal uncertainty is obtained by selecting m sampling locations \mathcal{X}^* so that conditional entropy is minimized. This is equivalent to maximizing the entropy of a random function $f(\mathbf{x})$ in sampling locations (8)

$$\begin{aligned} \mathcal{X}^* &= \operatorname{argmin}_{\mathcal{U}} H(f(\mathbf{v})|f(\mathbf{u})) \\ &= \operatorname{argmin}_{\mathcal{U}} (H(f(\mathbf{x})) - H(f(\mathbf{u}))) \\ &= \operatorname{argmax}_{\mathcal{U}} H(f(\mathbf{u})) \end{aligned} \quad (8)$$

where $\mathcal{U} \subset \mathcal{X} : |\mathcal{U}| = m$.

For a Gaussian process $f(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), c(\mathbf{x}, \mathbf{x}'))$, the entropy at locations \mathcal{U} is given by (9)

$$\begin{aligned} H(\mathbf{y}_m) &= - \int p(\mathbf{y}_m) \log p(\mathbf{y}_m) d\mathbf{y}_m \\ &= \frac{1}{2} \log((2\pi e)^m |\mathbf{C}_{mm}|) \end{aligned} \quad (9)$$

where $\mathbf{y}_m = \{y^1, y^2, \dots, y^m\}$ and $y^i = f(\mathbf{u}_i)$. Thus, maximal entropy in the sampling locations is obtained by maximizing the determinant of covariance matrix \mathbf{C}_{mm} , which is one form of D -optimal design [1], [2]. However, this is a NP -hard problem, for which reason the following greedy heuristic [8] is typically used:

Start with an empty optimal set $\mathcal{X}^* = \emptyset$ and a full retrieval set $\mathcal{U} = \mathcal{X}$
Add on each iteration i the location \mathbf{x}_i^* that has the highest conditional entropy $\mathbf{x}_i^* = \operatorname{argmax}_{\mathbf{u}} H(f(\mathbf{u})|f(\mathcal{X}_{i-1}^*))$
Remove on each iteration the optimal location from the retrieval set $\mathcal{U} = \mathcal{U} \setminus \mathbf{x}_i^*$
Finish with an optimal set of locations $\mathcal{X}^* : |\mathcal{X}^*| = m$

which follows from the chain rule of entropy (conditional entropy), i.e.,

$$H(f(\mathcal{X}_i^*)) = H(f(\mathbf{x}_i^*)|H(f(\mathcal{X}_{i-1}^*))) + H(f(\mathcal{X}_{i-1}^*)).$$

In the previous section we found that the conditional distribution of a Gaussian process is given by a multivariate normal distribution with a mean (4) and variance (5), which for a single random variable $y^i = f(\mathbf{u}_i)$ is given by a 1-dimensional normal distribution. As given by (9), we are now searching for the location \mathbf{u}_i that maximizes the conditional variance $\operatorname{var}(y^i|f(\mathcal{X}_{i-1}^*)) = \operatorname{var}(y^i) - \operatorname{cov}(y^i, f(\mathcal{X}_{i-1}^*))\operatorname{var}(f(\mathcal{X}_{i-1}^*))\operatorname{cov}(f(\mathcal{X}_{i-1}^*), y^i)$ on every iteration.

3. Experiments

In our experiments, a distributed multi-robot sensing system was used to collect scalar maps of ambient light and magnetic flux density. These data were further used in simulated optimal spatial sampling for both stationary and non-stationary models. The accuracy of the models was measured by the integrated mean squared error (IMSE) between the models and the maps. We assumed that optimal spatial sampling using non-stationary Gaussian processes would provide more accurate models than stationary processes. The following subsections present the experimental environment, data acquisition and simulation results, respectively.

3.1. Experimental Environment

In spatial modelling, mobile sensors provide an opportunity to select sampling locations that maximize the accuracy of the model. However, mobility reduces the reliability of the sampling locations due to odometer failures, which again reduce the likelihood of the model. Therefore, we built an experimental environment that provides an opportunity to study spatial modelling with known sensor locations. The experimental environment was realized using roof camera-based multi-robot localization, remote-operated robots and a central computer that performed data acquisition, robot control and environment modelling.

The multi-robot localization [11] was based on colour segmentation, which proved to be an efficient and reliable method. In order to localize and identify each robot from a roof camera image, localization discs (Figure 1) were mounted on top of the homogenous robots [4]. The segmentation algorithm exploits an HSV colour space with the V channel discarded for better tolerance against changes in lighting. The hue and saturation of each pixel was compared against preset minimum and maximum thresholds. The robot candidates with a large enough area were found using connected component labelling. The smaller rings under the robot area were sought in a similar manner in order to identify the

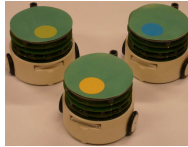


Figure 1. Three different localization discs attached to homogenous robots.

robots. The orientation of the robots was calculated as the angle between the centroids of the rings.

The segmentation algorithm provided multi-robot localization on image coordinates as presented in Figure 2. In order to find a transformation between image and metric coordinates, the camera was calibrated using Heikkilä's Camera calibration toolbox for Matlab [5]. Based on the calibration, one pixel in the image was judged to correspond to ca. 2.3 mm in the metric coordinate system. The remaining distortion, occurring especially near the image corners, was further reduced by nonlinear transformation resulting from curve fitting.

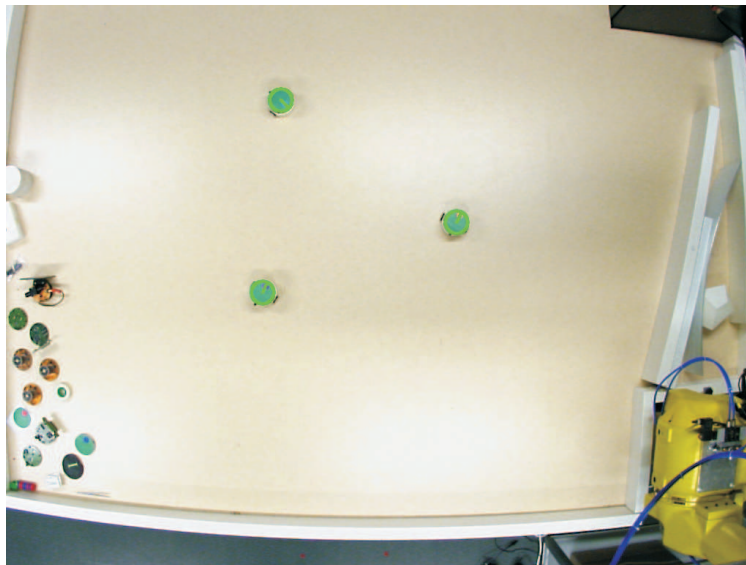


Figure 2. A Roof camera localization in a three-robot system

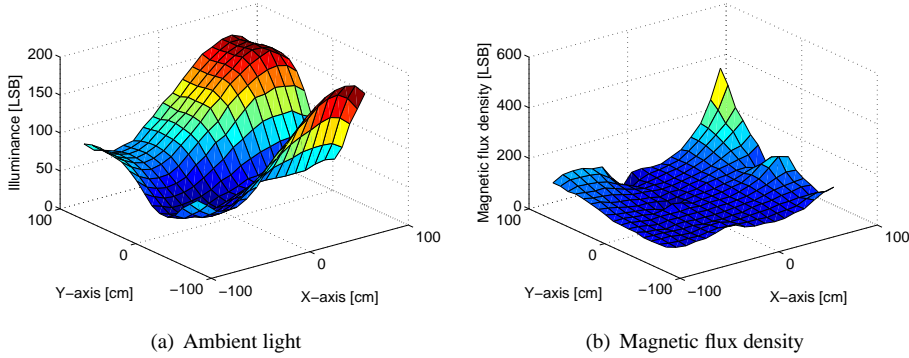
The accuracy of the localization system was evaluated for both position and orientation. Accuracy was measured on a plane in ten positions chosen on the basis of systematic unaligned sampling [7]. Table 1 presents the results of localization accuracy. The localization method reached a frame rate of 18 FPS with a resolution of 1024 x 768 pixels on a 3 GHz P4 desktop computer. The presented experimental environment can provide simultaneous localization for up to six robots with virtually unaffected performance.

Table 1. Observed localization errors

	Position, x axis	Position, y axis	Orientation
Maximum absolute error	3.8cm	4.4cm	13.7 deg.
Mean absolute error	1.6cm	1.7cm	4.7 deg.
Localization standard deviation	0.1cm	0.1cm	2.9 deg.

3.2. Data acquisition

The environmental data were acquired using a group of three robots. The robots moved arbitrarily in an area measuring approximately 180 x 160 cm while measuring ambient light and magnetic flux density. Furthermore, this data was stored and averaged on scalar maps containing 288 samples with a 10 cm grid. These maps are presented in Figure 3(a) and Figure 3(b). The units of ambient light and magnetic flux density were scaled and do not correspond to true luminance [cd/cm²] or magnetic flux density [T].

**Figure 3.** Scalar maps

3.3. Simulation results

Optimal spatial sampling simulations were performed for environmental data collected in advance using robots. The stationary models used a Gaussian kernel (10)

$$k(\mathbf{x}_i, \mathbf{x}_j) = (2\pi)^{-1/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^T \Sigma^{-1} (\mathbf{x}_i - \mathbf{x}_j)\right) \quad (10)$$

where the variance hyperparameter was given a constant value

$$\Sigma = \begin{bmatrix} 4000 & 0 \\ 0 & 4000 \end{bmatrix}.$$

This kernel was also an initial estimate for each local kernel in the non-stationary Gaussian processes. For both ambient light and magnetic flux density simulations, the mean value functions in the Gaussian processes were set to a constant $\mu(\mathbf{x}) = 100$ LSB.

For local kernel adaptation in the non-stationary Gaussian processes we used bounded linear adaptation with evaluation structure tensors [EST], as presented in [9]. However, in our simulation local gradients had to be calculated from interpolated values, in contrast to [9], who were able to use all the collected samples in the adaptation. Especially at the beginning of the sampling, this caused overestimated spatial smoothness in some interpolation areas. Thus, for maximum entropy sampling these areas seemed to provide a small amount of information and they were undersampled. To overcome these difficulties, the learning rate and the hyperparameter boundaries were given values $\eta = 0.2$, $\sigma_{min} = 3000$ and $\sigma_{max} = 5000$.

Figures 4(a) and 4(b) present 50 optimal samples in ambient light and magnetic flux density interpolation, respectively. These samples were selected by maximum entropy sampling for both stationary and non-stationary models. In the stationary models the optimal samples are distributed uniformly over the sampling area, whereas in the non-stationary models the sampling locations are focused more on the areas with the greatest gradients, i.e., on the areas that provide the most information about quantity.

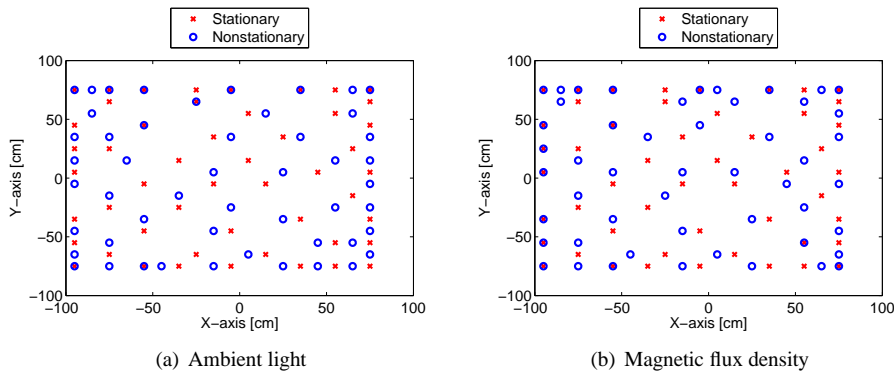


Figure 4. Optimal sampling locations

Figures 5(a) and 5(b) present integrated mean squared errors as a function of the optimal samples up to 50 samples. The integrated mean squared errors were obtained by comparing the interpolation results with the maps. The IMSE of ambient light interpolation was nearly equal after 50 samples for both models which results from the smooth variation in spatial data. The IMSE of magnetic flux density interpolation was slightly smaller for the non-stationary model. However, the difference was not statistically significant. Figures 6(a) and 6(b) present stationary interpolation, and Figures 7(a) and 7(b) nonstationary interpolation in a 50-sample optimal spatial sampling. We believe larger data sets with more variation would bring out the advantages of the non-stationary models.

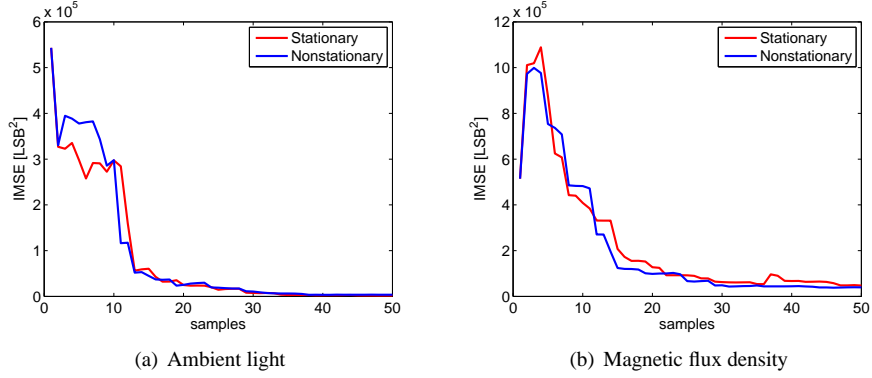


Figure 5. Integrated mean squared errors

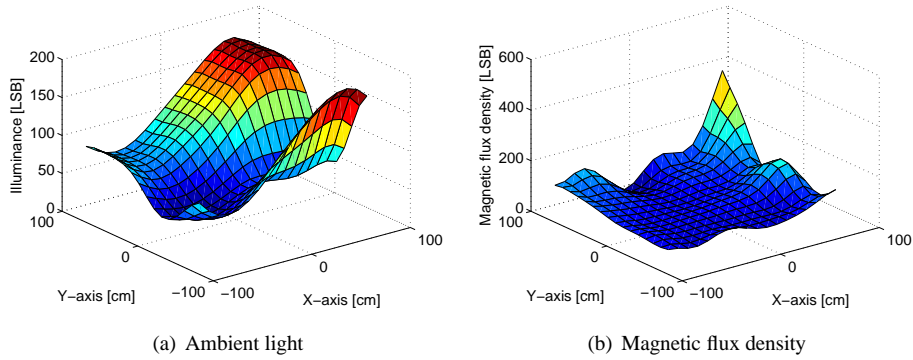


Figure 6. Stationary interpolation

4. Conclusions

In this paper we presented an experimental environment for optimal spatial sampling in a multi-robot system. The environment enabled research on optimal spatial sampling using distributed and mobile sensors. The experimental environment was realized using roof camera-based multi-robot localization, remote-operated robots and a central computer for data acquisition, robot control and environment modelling.

In the simulations, the accuracy of optimal spatial sampling was measured for both the stationary and non-stationary models. Although the results did not show a clear advantage of the non-stationary models, authors believe that given larger and more variable data sets the difference would have been more significant. The simulation results encourage the authors to develop actual optimal spatial sampling methods for multi-robot systems. Such a research area contains not only optimization of information, but also optimization of measurement time, travelled distances, etc. Outdoors, aided with GPS localization, these methods could measure pH levels or the oxygen content of water, for example.

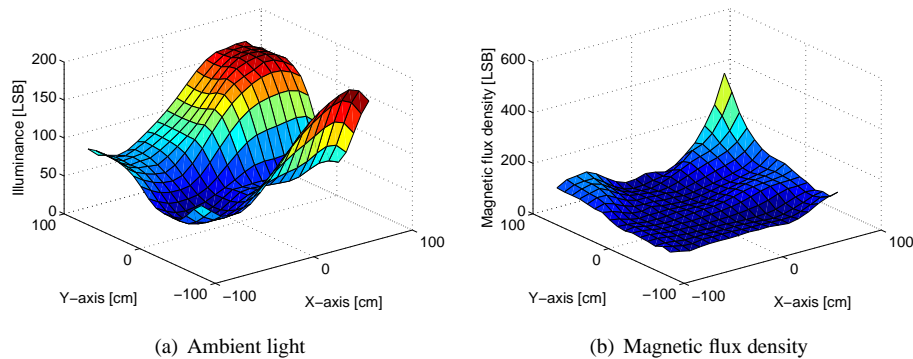


Figure 7. Nonstationary interpolation

Acknowledgements

This work was supported by the Academy of Finland and Infotech Oulu.

References

- [1] A.C. Atkinson and A.N. Donev. *Optimum Experimental Designs*. Oxford University Press, New York, 1992.
- [2] D.D.Cox, L.H. Cox, and K.B. Ensor. Spatial sampling and the environment: some issues and directions. *Environmental and Ecological Statistics*, 4:219–233, 1997.
- [3] V.V. Fedorov. Optimal spatial design: Spatial sampling. *Calcutta Statistical Association Bulletin*, 44, 1994.
- [4] J. Haverinen, M. Parpala, and J. Röning. A miniature mobile robot with a color stereo camera system for swarm robotics research. In *IEEE International Conference on Robotics and Automation (ICRA2005)*, pages 2494–2497, Barcelona, Spain, Apr 18 - 22 2005.
- [5] J. Heikkilä. Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1066–1077, 2000.
- [6] D. Higdon, J. Swall, and J. Kern. Non-stationary spatial modeling. *Bayesian Statistics 6*, eds. J.M. Bernardo et al., Oxford University Press, pages 761–768, 1999.
- [7] L. J. King. *Statistical Analysis in Geography*. Englewood Cliffs (N.J.) : Prentice-Hall, 1969.
- [8] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. Technical report, Carnegie Mellon, Machine Learning Department, 2007.
- [9] T. Lang, C. Plagemann, and W. Burgard. Adaptive non-stationary kernel regression for terrain modeling. In *Proceedings of Robotics: Science and Systems*, Atlanta, GA, USA, June 2007.
- [10] G. Matheron. Kriging, or polynomial interpolation procedures? *CIM transactions*, LXX:240–244, 1967.
- [11] T. Mäkelä, A. Tikanmäki, and J. Röning. A general-purpose vision system for diverse robots. In *Robotics and Applications and Telematics*, Würzburg, Germany, Aug 29 - 8 2007.
- [12] B. Mukherjee. A note on sampling designs for random processes with no quadratic mean derivative. *Australian & New Zealand Journal of Statistics*, 48(3):305–319, 2006.
- [13] C. Paciorek. *Nonstationary Gaussian Processes for Regression and Spatial Modelling*. PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania, 2003.
- [14] C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.
- [15] M.C. Shewry and H.P. Wynn. Maximum entropy sampling. *Applied Statistics*, 46:165–170, 1987.
- [16] C.K.I. Williams and C.E. Rasmussen. Gaussian processes for regression. In David S. Touretzky, Michael C. Mozer, and Michael E. Hasselmo, editors, *Proc. Conf. Advances in Neural Information Processing Systems, NIPS*, volume 8. MIT Press, 1996.