Robot Path Planning Using SIFT and Sonar Sensor Fusion

ALFREDO CHÁVEZ
Aalborg University
Department of Control
Fredrik Bajers Vej 7C, 9220 Aalborg
DENMARK

HECTOR RAPOSO
Aalborg University
Department of Control
Fredrik Bajers Vej 7C, 9220 Aalborg
DENMARK

Abstract: This paper presents a novel map building approach for path planning purposes, which takes into account the uncertainty inherent in sensor measurements. To this end, Bayesian estimation and Dempster-Shafer evidential theory are used to fuse the sensory information and to update the occupancy and evidential grid maps, respectively. The approach is illustrated using actual measurements from a laboratory robot. The sensory information is obtained from a sonar array and the Scale Invariant Feature Transform (SIFT) algorithm. Finally, the resulting two evidential maps based on Bayes and Dempster theories are used for path planning using the potential field method. Both yield satisfying results.

Key–Words: Sensor fusion, mobile robots, stereo vision, sonar, occupancy grids, SIFT, Dempster-Shafer, potential field.

1 Introduction

In the field of mobile robots one of the main requirements is to have the capacity to operate independently in uncertain and unknown environments; fusion of sensory information, map building and path planning are some of the key capabilities the mobile robot has to possess in order to achieve autonomy. Map building is one solution that addresses the problem of aquiring data from sensors; the data must be interpreted and fused by means of sensor models. The fusion process is carried out using standard data fusion methods [1]. The result of the fusion of the sensor information can be used to construct a map of the robot's environment, and the robot can then plan its own path and avoid obstacles. In this paper, the scope is limited to building a map for a laboratory robot by fusing range readings from a sonar array with landmarks extracted from stereo vision images using the SIFT algorithm. We compare two sensor fusion techniques, Bayesian Inference and Dempster-Shafer Evidential theory. These techniques yield so-called Occupancy and Dempster-Shafer grids, respectively, which are internal representations that can be used for robot navigation. Occupancy grids were introduced by Elfes in [2] and [3]. Dempster-Shafer grids were proposed in [4], as an alternative to occupancy grids. Localization can also be implemented, but it is not considered in this paper. The aim of this paper is to show that it is feasible to perform path planning based on the potential fields derived from maps that have been generated using fused range readings from the sonar and the vision system. In the following, in Sections 2 and 3, we will first present an overview of the sensor fusion along with the main contribution of this paper: a novel sensor fusion of Scale Invariant Feature Transform, a recently developed computer vision method, and sonar range readings. In Section 4, we outline how the sensor fusion can be employed to generate potential fields for indoor robot path planning, whereupon we show some experimental results in Section 5. Finally, in Section 6 we sum up the conclusions of this work.

2 Sensor Models

A common sensor used to measure distance is the ultrasonic range finder or sonar. The sonar can measure the distance from the transducer to an object quite accurately. However, it cannot estimate at what angle within the sonar cone the pulse was reflected. Hence, there will be some uncertainty about the angle at which the obstacle was measured.

A wide range of sonar models have been developed in the past years by various researchers, [2], [3], [12] and [14].

Taking the starting point in these methods, we shall define a grid G of cells $C_{x,y}, 1 \leq x \leq x_{max}, 1 \leq y \leq y_{max}$, in front of the sensor. Consider the representation of the sonar beam cone shown in figure 1. Let Ψ denote the top angle of the cone in the horizontal plane and let ϕ denote the (unknown) angle from the centerline of the beam to the grid cell $c_{i,j}$. Let r denote a sonar range measurement and ϵ the mean sonar deviation error. Then $\Phi_e^s(\delta, \phi, r) = F_s(\delta, r)A_n(\phi)$ represents the probability of the cell $c_{i,j}$ (translated from polar coordinates (r, ϕ)) being empty, and $\Phi_o^s(\delta, \phi, r) = \mathcal{O}_s(\delta, r)A_n(\phi)$ represents the probability of the cell $c_{i,j}$ being occupied. The factors F_s and $\mathcal{O}_s$ are given by

$$F_s(\delta, r) = \begin{cases} 1 - \left(\frac{\delta}{r}\right)^2, & \text{if } \delta \in [0, \mu] \\ e^r, & \text{if } \delta \in [\mu, r - \epsilon] \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$\mathcal{O}_s(\delta, r) =$$

$$\begin{cases} \left(\frac{1}{r}\right)\left(1 - \left(\frac{\delta - r}{\epsilon}\right)^2\right), & \text{if } \delta \in [r - \epsilon, r + \epsilon] \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

and

$$A_n(\phi) = \begin{cases} 1 - \left(\frac{2\phi}{\Psi}\right)^2, & \text{if } \phi \in \left[-\frac{\Psi}{2}, \frac{\Psi}{2}\right] \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The value μ in the sonar model represents the minimal measurement and δ is the distance from the sonar to the cell. The former can be depicted in fig 1.

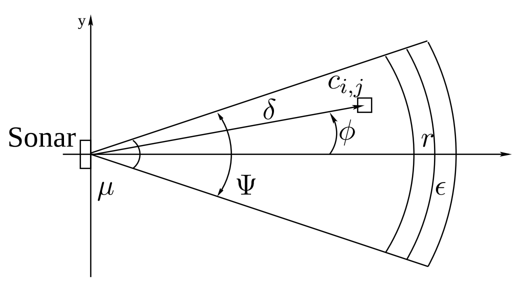


Figure 1: Sonar Model

The other sensor, we shall use for sensor fusion in this study is a stereo vision system. In particular, The Scale Invariant Feature Transform (SIFT) is a method for extracting distinctive invariant features from digital images. The features are invariant to scale and rotation. They also provide a robust matching across a substantial range of affine distortion, change in 3D view point, addition of noise, and change in illumination. Furthermore, the features are distinctive, i.e. they can be matched with a high probability with other features in a large database with many images, [11]. This particular property makes the SIFT method suitable for robot navigation, localization, and mapping. The SIFT algorithm consists of the following major steps, [11].

- Scale-space peak detection: The aim of this step is to find locations in the image that are invariant to scale change in the same image.
- Accurate key-point localization: In this step the position of each point candidate is determined; points with low contrast and poor localization along the edge are removed. This yields a so-called descriptor at each point.
- Majority orientation assignment: This step makes the rotation descriptor invariant. This is done by assigning a consistent orientation to each key-point.
- Computation of the local image descriptor: This step associates each feature point with a 128-element feature vector or interest point descriptor that uniquely identifies that point.

Once the descriptors are found in each image, eg. left and right images, a matching algorithm is applied in both images, figure 2 presents the matching features descriptors which have been identified from a stereo pair of images. Stereo triangulation is implemented in order to get the depth from the pair of cameras to the features [8] and [9]. Due to the factors of quantification and calibration errors, a certain degree of uncertainty must be expected in the triangulation [13].

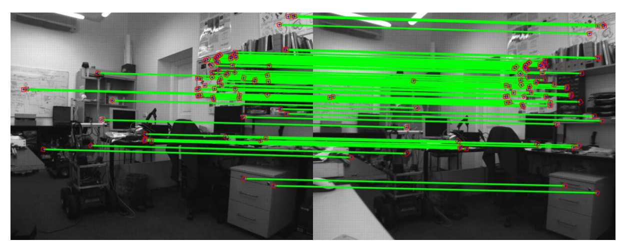


Figure 2: Descriptor matches between left and right images.

3 Sensor Fusion

In the following, we wish to apply two different methods of sensor fusion: Bayes' rule and Dempster-Shafer theory of evidence. In each case, we wish to construct occupancy grids for each sensor type, which will then be used to build up the resulting maps.

In the first case, two rules are applied to obtain the resulting probability, [14]:

- If at least one of the source cells has higher probability that it represents occupied space than a predefined threshold T_o, the probability of the resulting cell is set to 1.

- Otherwise, Bayes' rule is applied to determine whether it is occupied or empty.

More precisely, the resulting grid is computed in two steps. Values in the source grids are modified first using the following formula:

$$P^o(c) = \begin{cases} 1 & \text{for } P^o(c) > T_o, \\ \frac{P^o(c)+T_o-1}{2 \cdot T_o - 1} & \text{for } P^o(c) \in [\frac{1}{2}, T_o] \\ P^o(c) & \text{otherwise} \end{cases} \quad (4)$$

where $\mathcal{P}^o(c)$ is the probability of the cell $(c_{i,j})$ being occupied (o).

The computed values are then applied using Bayes' rule to obtain the occupied fused probability $(P^o_f(c))$ of the cell $(c_{i,j})$ in the resulting grid.

$$P^o_f(c) = \frac{P^o_1(c)P^o_2(c)}{P^o_1(c)P^o_2(c) + (1 - P^o_1(c)(1 - P^o_2(c))} \quad (5)$$

where $P^o_1(c)$ is the modified probability of occupancy from the first sensor and $P^o_2(c)$ is the modified probability of occupancy from the second sensor.

The second case concerns Dempster-Shafer theory of evidence. This theory was proposed by Glenn Shafer [4] as an extension of the work presented in [5] and [6]. Dempster-Shafer theory is mainly characterized by a frame of discernment Θ, a basic probability assignment function, a belief function β and a plausibility function π. These are tied together via the so-called Dempster's rule of combination [7]. Each proposition in Θ is called a singleton. 2^Θ is called the power set of Θ. Any subset of Θ is called a hypothesis.

Applying the notion of frames of discernment to an occupancy grid yields a set of frames $\Theta_{i,j} = \{o, e\}$; where i, j represents an individual cell in the grid. Let A denote the subsets of the power set of $2^{\Theta_{i,j}} = 2^{\{o,e\}} = \big\{\{\varnothing\}, \{o\}, \{e\}, \{o, e\}\big\}$; where $\{\varnothing\}$ and $\{o, e\}$ are the empty and the disjunction or "don't know" subsets, respectively. $\{o\}$ and $\{e\}$ denote the probabilities of the cell being occupied or empty, respectively. The *quantum of belief* is distributed as $\beta(A) = m(\varnothing) + m(o) + m(e) + m(o, e) = 1$ [4]. Finally, the function $m : 2^\Theta \rightarrow [0, 1]$ is called the basic propability assignment, and must satisfy the following criteria.

$$\sum_{A \subset 2^\Theta} m(A) = 1 \quad (6)$$

$$m(\varnothing) = 0 \quad (7)$$

Equation (7) reflects the fact that no belief is assign to $\varnothing$. In order to obtain the total evidence assigned to A, one must add to $m(A)$ the quantities $m(B)$ for all proper subsets B of A.

$$\beta(A) = \sum_{\forall B : B \subseteq A} m(B) \quad (8)$$

In [4], the notion of *plausibility* or upper probability of A is defined as $1 - \beta(\neg A)$; where $(\neg A)$ is used to denote the set theoretic complement of A. $\beta(\neg A)$ is the disbelief of the hypothesis of A. Consequently, $\pi(A)$ can be thought of as the amount of evidence that does not support its negation. All in all, this sums up to

$$\pi(A) = 1 - \beta(\neg A) = 1 - \sum_{\forall B : B \not\subseteq A} m(B) \quad (9)$$

Notice that $\beta(A) \leq \pi(A)$ for any given A.

In general notation, Dempster's rule of combination is:

$$m(C_k) = \frac{\sum_{\forall A_i, B_j \in \Lambda : A_i \cap B_i = C_k ; C_k \neq \emptyset} m(A_i)m(B_j)}{1 - \sum_{\forall A_i, B_j \in \Lambda : A_i \cap B_j = \emptyset} m(A_i)m(B_j)} \quad (10)$$

where k is an appropriate index. When using Dempster's rule of combination to update a grid map for each cell $c_{i,j}$ lying in the main lobe of the sonar model and for each sensor reading r, equation (10) becomes

$$m^G_o = \frac{m^G_o m^S_o + m^G_o m^S_{o,e} + m^G_{o,e} m^S_o}{1 - m^G_e m^S_o - m^G_e m^S_e} \quad (11)$$

$$m^G_e = \frac{m^G_e m^S_e + m^G_e m^S_{o,e} + m^G_{o,e} m^S_e}{1 - m^G_e m^S_o - m^G_e m^S_e} \quad (12)$$

The quantities m^S_o, m^S_e and $m^S_{o,e}$ are obtained from sensor models, while m^G_o, m^G_e and $m^G_{o,e}$ are obtained from the existing grid map. Note that $m^G_{o,e} = 1 - m^G_o - m^G_e$, and $m^S_{o,e} = 1 - m^S_o - m^S_e$. The Dempster-Shafer grid map is initialized as follows:
$m^G_o = 0$, $m^G_e = 0$ and $m^G_{o,e} = 1 \; \forall P_{i,j} \in G$.

4 Path Planning Using Potential Fields

The main idea of potential fields is to discretize the configuration space of the robot $\mathcal{A}$ into a regular grid and search for an appropriate path within that grid. In this approach, the robot is considered as a particle in the configuration space moving under the influence of an artificial potential field U. The potential field consists of the sum of an attractive potential field generated by the goal and a repulsive potential generated by the obstacles [10] and [16].

$$U(\vec{q}) = U_{att}(\vec{q}) + U_{rep}(\vec{q}) \tag{13}$$

where $\vec{q}$ is the current state of the robot (aka. *configuration*). The attractive potential can for instance be represented as a quadratic function with its minimum at the goal configuration:

$$U_{att}(\vec{q}) = \frac{1}{2}\xi\rho^2_{goal}(\vec{q}) \tag{14}$$

in which ξ is a positive scaling factor, ρ_{goal} denotes the Euclidean distance between the current configuration and the goal $\|\vec{q}-\vec{q}_{goal}\|$ and U_{att} is a non-negative field with minimum in $U_{att}(\vec{q}_{goal}) = 0$.

An example of a repulsive potential field function is

$$U_{rep}(\vec{q}) = \begin{cases} \frac{1}{2}\eta\left(\frac{1}{\rho(\vec{q})} - \frac{1}{\rho_0}\right)^2 & \text{if } \rho(\vec{q}) \leq \rho_0, \\ 0 & \text{if } \rho(\vec{q}) > \rho_0, \end{cases} \tag{15}$$

Where η is a positive scaling factor; $\rho_{(\vec{q})}$ are the distances from the current configuration of the robot $\vec{q}$ to the obstacle region $\mathcal{CB}$, e.g. it is the Euclidian distance. ρ_0 is the maximum distance of influence, e.g. it is the distance from the center of the obstacle to the boundary of the obstacle region.

The force to attract and repulse the robot can be obtained from the negated gradient of the potential, as it is showed in (16).

$$\vec{F} = -\nabla U(\vec{q}) = -\begin{bmatrix} \nabla U_{att}(\vec{q}) \\ \nabla U_{rep}(\vec{q}) \end{bmatrix} = -\begin{bmatrix} \vec{F}_{att}(\vec{q}) \\ \vec{F}_{rep}(\vec{q}) \end{bmatrix} \tag{16}$$

As it is stated above, the potential field can be obtained mathematically when the position of the obstacles is precisely identified. The obstacles generate a repulsive potential field which it makes the robot navigate far away from the obstacles. The other option considered in this article consists of moving the robot through the obstacles generated by applying the sensor fusion techniques (Bayes and Dempster-Shafer theorems) to the sensor readings. The attractive potential field is added to the potential field generated from the environment using sensor readings, as it can be depicted in figure 3.

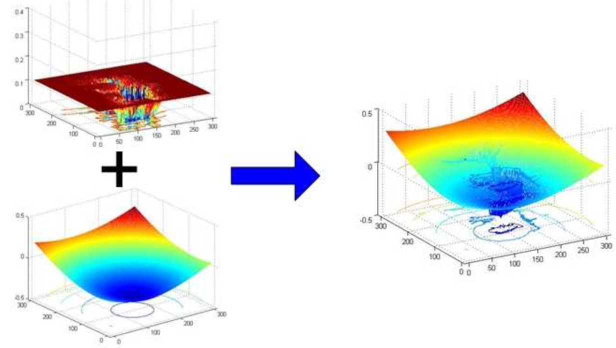


Figure 3: Addition of two potentials (attractive and repulsive) into the potential field.

The algorithm implemented in this section is called ”Depth-first planning” [16]; it mainly consists of constructing single segments starting at the initial configuration of the robot $\vec{q}_{init}$. The direction of each segment is obtained by solving (16); the ”Depth-first planning” algorithm can be depicted as in figure 4. This technique simply follows the steepest descent of the potential function until the goal configuration $\vec{q}_{goal}$ is reached. A drawback of this method is when the mobile robot gets trapped into a local minimum, which is not the case in the present simulation. There are solutions to the local minima problem [17].

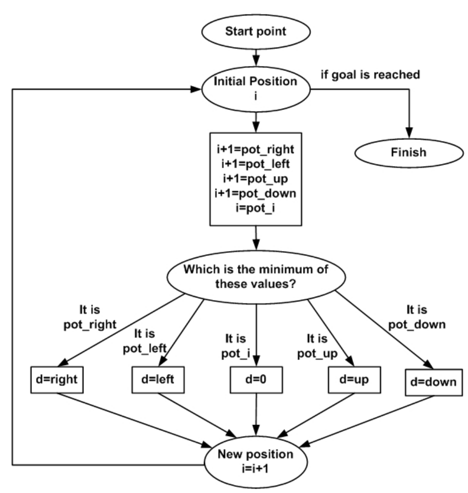
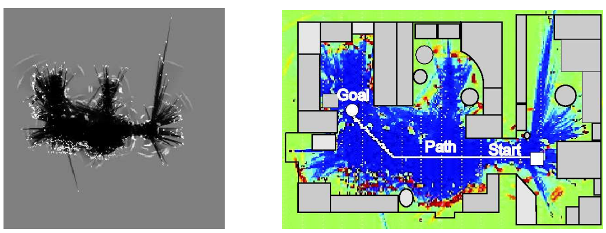


Figure 4: The Algorithm considers the implementation of the path based on the attractive potential of the goal location.

5 Experimental Results

The experiment depicted in figs. 5 to 9 was performed in an office/laboratory environment. Figure 5(1) represents a map which has been constructed by fusing sonar readings with SIFT-features using Bayes' rule; the resulting grid is created by fusing two grids, one grid created from sonar measurements and the other one from SIFT-features. In the resulting map the black color represents the empty area; the white color represents the occupied area, and the gray color represents the probability of $P^o_{i,j} = P^e_{i,j} = 0.5$. Figure 5(2) represents the shape of the laboratory/office and the map from fig 5(1) which is embedded in this map; the dark color represents the empty area. A path is planned in this figure; it connects the start point configuration with the goal point configuration.



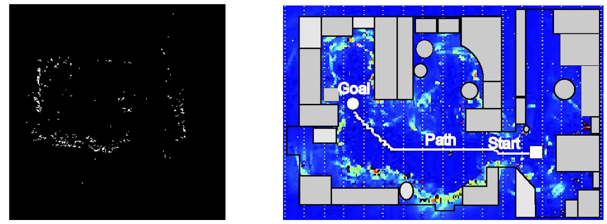
(1) **(2)**

Figure 5: (1) Grid created by fusing two grids, using Bayes' rule, one grid created from sonar measurements and the second one from SIFT-features. (2) Shows the path constructed on the map which has been built up by sensor readings.

Fig. 6(1) presents a grid created from $m^G_{i,j}$(o). The white spots in this map represent the occupied region, the black area outside the white spots represents the lack of evidence, e.g. $\beta(o) = \beta(e) = 0$ and $\beta(o, e) = 1$, and the black area inside the white spots represent the empty region where the robot can plan its path. Fig. 6(2) shows the path.

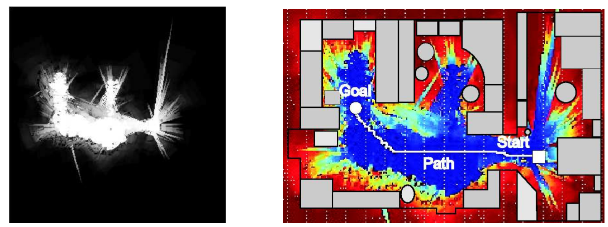
Figure 7(1) shows a map which has been constructed by fusing sonar readings with SIFT-features using Dempster-Shafer evidential theory; the resulting grid shows the $2D$ plot of the $m^G_{i,j}(e)$ region (white color on the map). The black color represents the lack of evidence e.g. $\beta(o) = \beta(e) = 0$ and $\beta(o, e) = 1$. Figure 7(2) shows the path connecting the start point configuration with the goal point configuration.

Fig. 8(1) depicts a grid created from $m^G_{i,j}(o, e)$. The black area represents the lack of evidence, e.g. $\beta(o, e) = 0$, in this area the robot can move around and plan its path. 8(2) shows the path that has been planned in the empty area of the map.



(1) **(2)**

Figure 6: Grid created only from the fusion between SIFT-features and sonar. (1) represents the $2D$ plot which shows the $m^G_{i,j}$(o), (2) Shows the path.



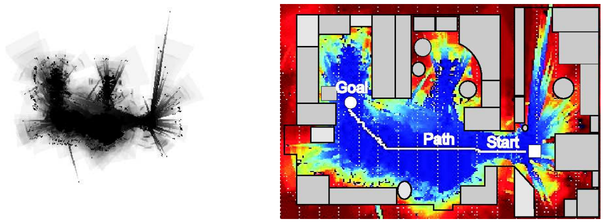
(1) **(2)**

Figure 7: Grid created only from the fusion between SIFT-features and sonar. (1) represents the $2D$ plot which shows the $m^G_{i,j}(e)$, (2) is the $3D$ representation.

Figure 9(1) shows a resulting map which has been constructed by transforming the $m^G_{i,j}(e)$, $m^G_{i,j}(o)$ and $m^G_{i,j}(o, e)$ maps into one map; this map is very similar to the one from the Bayes theory 5(1). Figure 9(2) shows the path connecting the start point configuration with the goal point configuration.

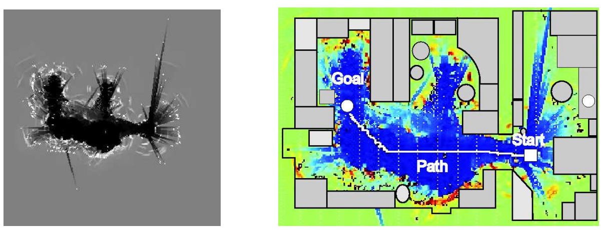
6 Conclusion and Future Work

The work made in this article presents a novel application of sensor data fusion for robot map making and path planning. The work considers the use of Bayes and Demster-Shafer rules of combination to integrate sensor readings from a stero vision system using the SIFT algorithm and a ring of sonars. The experiments were verified with real data in a real indoor environment. The experiment shows that the use of the SIFT algorithm can improve the sonar map and it can be effectively used for robot path planning. Future research work is to apply control strategies to the path planned by the algorithm.



(1) **(2)**

Figure 8: Grid created only from the fusion between SIFT-features and sonar. (1) represents the $2D$ plot which shows the $m^G_{i,j}(o, e)$, (2) depicts the path planned by the algorithm.



(1) **(2)**

Figure 9: Grid created from the transformation of the $(o, e\ and\ (o, e))$ maps from the fusion of the vision and sonar systems into one map. (2) it is the $3D$ representation.

References:

[1] Ren C. Luo and Chih-Chen Yih, "Multisensor Fusion and Integration: Approaches, Applications, and Future Research Directions", *IEEE Sensor Journal*, Vol. 2, No. 2, 1985.

[2] P. Moravec and A. Elfes, "High Resolution Maps from Wide Angle Sonar", *IEEE Journal of Robotics and Automation*, pp. 116–121, 1985.

[3] A. Elfes, "Using Occupancy Grids for Mobile Robot Perception and Navigation", *IEEE Journal of Robotics and Automation*, 1989.

[4] G. Shafer, "A Mathematical Theory of Evidence", *Princeton, NJ:Princeton University. Press*, 1976.

[5] Dempster A. P., "A generalization of Bayesian Inference", *Jornal of the Royal Statistical Society*, Vol. 30, pp. 205–247, 1968.

[6] Dempster A. P., "Upper and Lower probabilities Included by a Multivalued Mappin", *Annals of Mathematical statistics*, Vol. 38, pp. 325–339, 1967.

[7] Daniel Pagnac Eduardo M. Nebot and Hugh Durrant-Whyte, "An Evidential Approach to Map Bulding for Autonomous Vehicles", *IEEE Transactions of Robotics and Automation*, Vol. 14, pp. 623–629, 1998.

[8] Alfredo Chávez and Jan Dimon Bendtsen, "Sensor Fusion -Sonar and Stereo Vision, Using Occupancy Grids and SIFT", *Proceedings of the IEEE International Conference on Dynamics, Instrumentation and Control*, 2006.

[9] Alfredo Chávez and Petr Štěpán, "Sensor Data Fusion", *Proceedings of the IEEE Systems, Man and Cybernetics Society, Conference on Advances in Cybernetics Systems*, pp. 20–25, 2006.

[10] Gregorio Dudek and Michael Jenkin, "Computational Principles of Mobile Robotics", *Cambridge University Press*.

[11] David G. Lowe, "Distinctive Image Features from Scale-Invariant Key points", *Computer Science Department, University of British Colombia, Vancouver, B.C., Canada* Vol 2, pp. 91–110, 2004.

[12] Kurt Konolige, "Improved Occupancy Grids for Map Building", *Autonomous Robots* Vol 4, pp. 351–367, 1997.

[13] Larry Matthies and Steven A. Shafer, "Error Modeling in Stereo Navigation", *IEEE Journal of Robotics and Automation*, Vol 3, pp. 239–248, 1987.

[14] Štěpán,P., Miroslav Kulich and Libor Přeučil, "Robust Data Fusion with Occupancy Grids", *IEEE Transactions on Systems, Man, and Cybernetics* , Vol. 35, pp. 106-115, 2005.

[15] Andrew Howard, Maja J Matarić and Gaurav S Sukhatme "Mobile Sensor Network Deployment using Potential Fields: A Distributed, Scalable Solution to the Area Coverage Problem", *Proceedings of the 6th International Symposium on DARS02, Fukuoka, Japan*, 2002.

[16] Jean Claude Latombe "Robot Motion Planning", *Stanford University*, 1991.

[17] Christopher I. Connolly and J. B. BurnsS Sukhatme "Path Planning Using Laplace's Equation", *In Proceedings of the IEEE International Conference on Robotics and Automation*,2102–2106, 1990.