Incremental Robot Mapping with Fingerprints of Places

Adriana Tapus and Roland Siegwart
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Autonomous Systems Lab
1015 Lausanne, Switzerland
{Adriana.Tapus, Roland.Siegwart}@epfl.ch

***Abstract* – Even today, robot mapping is one of the biggest challenges in mobile robotics. Geometric or topological maps can be used by a robot to navigate in the environment. Automatic creation of such maps is still problematic if the robot tries to map large environments. This paper presents a new method for incremental mapping using fingerprints of places. This type of representation permits a reliable, compact, and distinctive environment-modeling and makes navigation and localization easier for the robot. Experimental results for incremental mapping using a mobile robot equipped with a multi-sensor system composed of two 180° laser range finders and an omni-directional camera are also reported.**

Keywords— mapping, fingerprint of places, cognitive representation

I. Introduction

Mobile robots typically use metric [1, 3, 4] or topological maps [12, 14, 18] of their physical environment to navigate reliably. Approaches using metric maps are suited when it is necessary for the robot to know its location accurately, in terms of metric coordinates. However, the state of the robot can also be represented in a more qualitative manner, similar to the way humans store spatial information, through the use of cognitive maps [20]. These permit an encoding of the spatial relations between relevant locations in the environment through a topological representation. The topological map can be viewed as a graph of places, where at each node the information concerning the visible landmarks and the way to reach other places, connected to it, is stored. The topological representation is compact and allows high-level symbolic reasoning for map building and navigation.

Space representation, perception, localization and mapping are all needed in order to obtain a robust and reliable framework for navigation (i.e., in order to move within an environment, manipulate objects in it, avoid undesirable collisions, etc.). In this context, we introduced in [9], the fingerprint approach (a fingerprint is a circular list of features around the robot) and further showed in [15] that a distinctive space representation combined with the uncertainty of the features can result in good performance in localization.

In this paper, we present a new technique for incremental and automatic topological mapping using fingerprints of places. One of the main difficulties in topological mapping is the automatic detection of new nodes that should be added to the map. Our approach relies on a heuristic that detects whether the current location of the robot is similar to a mapped one or not. The main contribution of this paper is the construction of an automatic topological mapping system relying on fingerprints of places. The closing the loop issue is not discussed in this work. The proposed method permits a reliable and distinctive environment model that can be globally handled in an efficient way.

The reminder of the paper is organized as follows. In Section II, the fingerprint concept, the way it is encoded, generated and combined with the uncertainty of features is described. Section III is dedicated to the new topological mapping approach with fingerprints of places. Experimental results are presented in Section IV. The system uses both, a laser and an omnidirectional camera for feature extraction. Section V exposes a short review of related research on metric, topological and hybrid mapping techniques. Finally, Section VI draws conclusions and discusses further work.

II. The Fingerprint Concept In A Topological Framework

In this work, fingerprints of places characterize the environment. The complete process is described below. This methodology is especially interesting when used within a topological framework and in a multi-modality context.

A. Fingerprint encoding

A fingerprint is a circular list of features, where the ordering of the set matches the relative ordering of the features around the robot. We denote the fingerprint sequence using a list of characters, where each character represents an instance of a specific feature type. In our case we choose to extract color patches and vertical edges from visual information and corners from laser scanner. The letter 'v' is used to characterize an edge, the letters 'A','B','C',...,'P' to represent hue bins and the letter 'c' to characterize a corner feature. Details about the extraction of visual features can be found in [10] and that of features extracted using laser scanner, in [1].

B. Fingerprint generation

Fingerprint generation is performed in three steps, as shown in Figure 1. The extraction of the different features (e.g., vertical edges, corners, color patches) from the sensors is the first step of the fingerprint generation process. The extracted features are ordered in a sequence depending on their angular position (0…360°). In the second step, a new type of feature, the virtual feature 'f' is introduced. This reflects the correspondence between a corner (detected with the laser scanner) and an edge (detected in the unwrapped omnidirectional image). In order to represent large (> 20 degrees, in our case) angular distances between successive fingerprint elements, the notion of an ′empty space′ feature is added. This is denoted in the fingerprint sequence by the character ′n′. In this way, the ordering of the features in a fingerprint sequence becomes highly informative, thereby increasing distinctiveness of fingerprints. This insertion is the last step of the fingerprint generation process. More details can be found in [10].

C. Uncertainty Modeling in the Fingerprint

Sensors are imperfect devices and thus, all obtained measurements are erred. These errors can be accounted for, by associating an uncertainty value to each recorded measurement. This uncertainty value represents the belief in the existence of the measured value/ feature when the robot (its sensor) actually perceives it. In our fingerprint approach, this idea is incorporated by associating every observed feature (for each of the different types of features mentioned above) with an uncertainty measure. More details, with regards to this, may be found in [15]. This uncertainty measure is however, only as accurate as the model that describes it.

III. TOPOLOGICAL MAP BUILDING

While navigating in the environment, the robot first creates and then updates the global topological map. One of the main problems in topological map building is to detect when a new node should be added in the map. Most of the existing approaches to topological mapping place nodes periodically in either space (displacement, Δd) or time (Δt) or alternatively attempt to detect important changes in environment structure. Any of these methods cannot result in an optimal topology. In contrast, our approach is based directly on the differences in the perceived features.

In the following sub-sections, the fingerprint-based approach for incremental and automatic topological mapping is described. In addition, they clearly illustrate how a reliable and distinctive representation of the environment is obtained.

A. New Node Detection

Our method introduces a new node into the map whenever an important change in the environment occurs. This is an improvement over the existing approaches, using some fixed rules based on distance measurements or topology structures, which are limited to specific cases of indoor or outdoor environments. This is possible using the fingerprints of places. A heuristic is applied to compare whether a new location is similar to the last one that has been mapped.

The process of introducing a new node in the topological map is divided into the following sequence of steps:

1) Start with an initial node (i.e. fingerprint f_0)

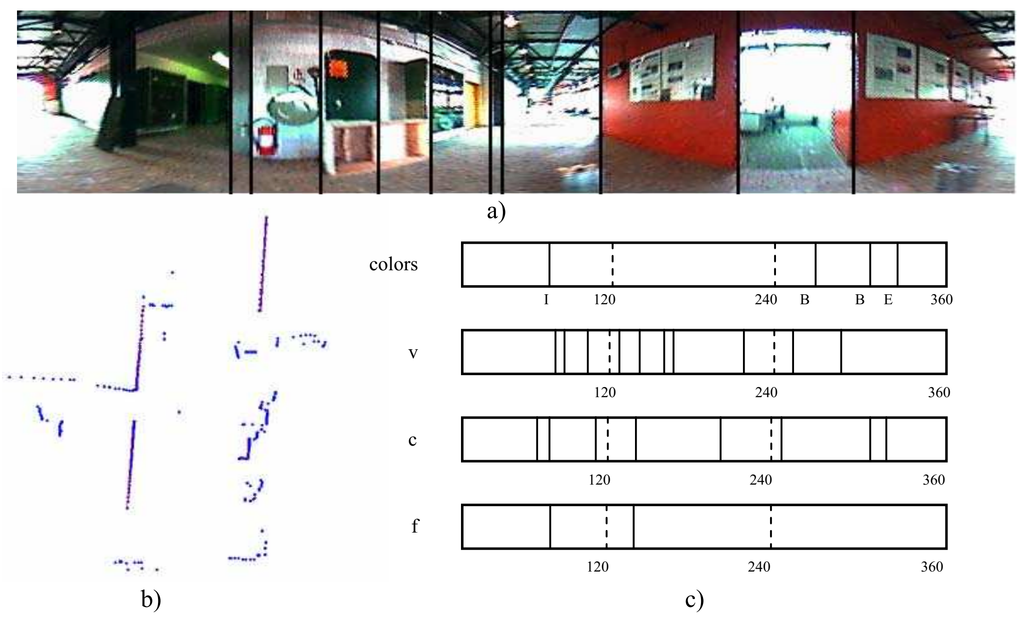

Figure 1: Fingerprint generation. (a) panoramic image with the vertical edges and color patches detection ′v′ and color; (b) laser scan with extracted corners ′c′; (c) the first three images depict the position (0 to 360°) of the colors (I-light blue, B- orange and E-light green), vertical edges and corners, respectively. The forth image describes the correspondence between the vertical edges features and the corner features. By regrouping all this results together and by adding the empty space features, the final fingerprint is: cIfvnvcvfnvvncvnncvBnvBccE

2) Move and at each Δt (time) or Δd (distance), take a new scan with the laser scanner and a new image with the omnidirectional camera and generate the new fingerprint f_i

3) Calculate the probability of matching, *prob_matching,* between the fingerprints f_{i-1} and f_i respectively. Compute the dissimilarity factor, *dissimilarity*.

$$prob\_matching = \mathrm{P}\,(f_i \mid f_{i-1}) \qquad (1)$$

$$dissimilarity(f_i\,, f_{i-1}) = 1 - prob\_matching \qquad (2)$$

4) If $dissimilarity(f_i\,, f_{i-1}) < \theta$ then
   a. Add fingerprint f_i to the current node n_k
   b. Calculate the new mean fingerprint of the node n_k

   Else
   a. A new node n_{k+1} is inserted (added) in the map
   b. Add fingerprint f_i to the node n_{k+1}

5) Repeat from step 2)

In step 4), we defined a threshold θ as the maximum allowable dissimilarity (i.e., 1-*prob_matching*) between the fingerprints. The value of *prob_matching* is calculated with the "global alignment with uncertainty" algorithm [15]. This method is an adaptation of the global alignment algorithm usually used for comparing D.N.A. sequences, introduced by Needleman and Wunsch [13]. The value of the threshold is determined experimentally. The incremental nature of the approach permits incremental additions to the map and yields the most up-to-date map at any time.

As mentioned previously, a step in the construction of the map is the generation of a mean fingerprint for each node. The following sub-section explains this process. It uses the "global alignment with uncertainty" algorithm [15] for fingerprint matching.

B. *Mean Fingerprint Generation*

As stated earlier, a fingerprint is extracted periodically in space (every Δd) or time (every Δt). A node is composed of several similar fingerprints that will be subsequently regrouped into a mean fingerprint. By choosing a suitable threshold θ, the mean fingerprint enables clustering of places into nodes.

As soon as a new fingerprint is added to the current node n_k, the mean fingerprint is updated by constructing the new mean fingerprint between the previous mean fingerprint and the newly introduced fingerprint (see Figure 2). The generation of the mean fingerprint between two fingerprints is performed in several steps, described briefly below. The first step in the mean fingerprint generation process consists of matching the two fingerprints involved. As the orientation of the robot is not known a priori or fixed beforehand, the robot estimates it by considering all the possible permutations of one fingerprint sequence with respect to another. The fingerprint matching algorithm, illustrated in [15], yields their best match. It can be seen in Figure 2 (Step 1), that once the two fingerprint sequences are aligned, they have the same length. In the second step, the mean fingerprint between two consecutively obtained fingerprints of places is computed. The mean fingerprint contains the features that matched during the fingerprint matching process and those with a high probability of existence. For calculating the mean fingerprint for a specific node n_k, these two steps are repeated until all the fingerprints of that node are included in it. Figure 2 describes this process through a simple example. Further enhancements to this method will be incorporated in the near future.

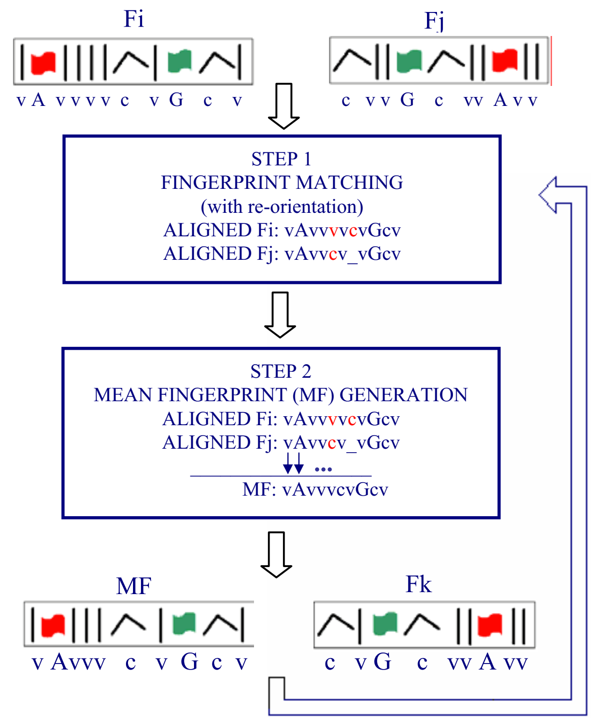

Figure 2: Illustration of the Mean Fingerprint Generation process. Fi, Fj and Fk are three consecutively obtained fingerprints.

The mean-fingerprint associates each node with a unique identifier (i.e., the mean-fingerprint). This enables the construction of a very distinctive and compact representation of the environment.

C. *Map Update*

By using a POMDP (Partially Observable Markov Decision Process), a discrete approximation of a probability distribution over all possible poses in the environment is computed.

The entropy of a probability distribution p is:

$$H(p) = -\sum_{s \in S} ps \log ps, \qquad (3)$$

where S is the finite set of environment states of the POMDP, and $p_s \log p_s = 0$ when $p_s = 0$. The lower the entropy value, the more certain the distribution. When the robot is "confused", the entropy is high.

Therefore, the strategy for the update of the map will be the following:

- When the entropy of the belief state is low enough, the map will be updated and so the fingerprint and the uncertainty of the features will also be updated.
- If the entropy is above a threshold α, then the updating will not be allowed, and we will try to reduce the entropy by continuing the navigation with localization.

If the robot feels confident of its state, it can decide if an extracted feature is new by comparing the observed fingerprint to the fingerprint from the map, corresponding to the most likely state (MLS), computed with the help of the POMDP. This can happen either in an unexplored portion of the environment, or in a known portion, where new features appear due to the environmental dynamics. As the features from the fingerprint come with their uncertainty*,* when a feature is re-observed, the uncertainty of the feature from the map fingerprint is weight averaged with the uncertainty of the extracted one. If the robot does not see an expected feature, the uncertainty is decreased. When the uncertainty of a feature from a map fingerprint is below a minimum threshold, then the feature is deleted, thereby allowing the incorporation of the dynamics in the map thus formed.

IV. Experimental Results

The approach has been evaluated in a portion of our institute building shown in Figure 4. For the experiments, the BIBA robot (see Figure 3), a fully autonomous mobile robot, has been used. Its controller consists of a VME standard backplane with a PowerPC 750 clocked at 400 MHz running XO/2, a hard real-time operating system and a Pentium III running at 700 MHz, with 128 MB RAM, using the Windows 2000 operating system for all interaction tasks. Both computers can communicate with each other over a 3 Mbps local Ethernet and with a central computer over wireless interfaces to allow for monitoring of the state of the robot for supervision.

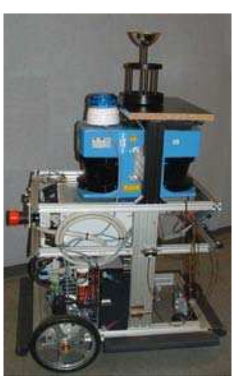

Figure 3: The fully autonomous robot BIBA.

Among its peripheral devices, the most important are the wheel encoders, two 180° laser range finders, five infrared sensors, four ultrasound sensors and an omni-directional camera. The omni-directional camera system uses a mirror-camera system to image 360° in azimuth and up to 110° in elevation.

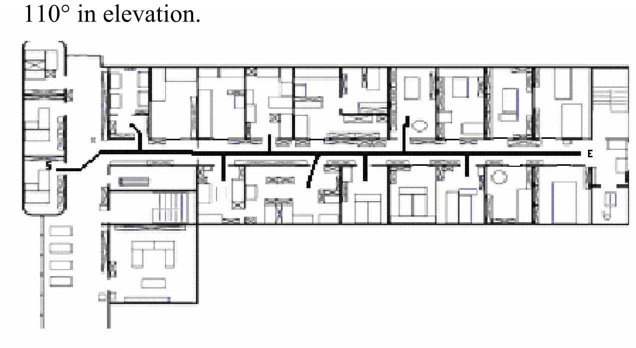

Figure 4: Floor plan of the first environment where the experiments have been conducted. The line shows the path followed by the robot in the environment. The robot starts at the point S and ends at the point E. The trajectory length is 75 m. During this step, the robot collected 500 data sets (i.e. images and scans) from the environment.

The test setup was the following: the robot started at the point S and ended at the point E as illustrated in the Figure 4, the distance traveled being of 75m. While the robot explored the environment, it recorded, at every Δd (distance) (e.g., in our case d =15cm), data readings from sensors (i.e., an image from the omni-directional camera and a scan from the laser scanner) in order to extract the fingerprints. The robot has a ′mid-line following′ behavior in the hallways and ′center of the free space′ behavior in the open spaces. We assume that the position in the room with the maximum free space around it, is the one with the highest probability of extracting numerous and characteristic features. This ensures high distinctiveness of the observation. The map building process was performed off-line.

Figure 5 shows the topological map obtained by the system in our laboratory, superimposed on a coarse map of the environment.

The fingerprints used for this representation contain just the vertical edges and the corners as features. The color patches were not included because they were very sensitive to changes in illumination present in our case. The resulting map is composed of 20 nodes as shown in the Figure 5. Each node is represented by a mean fingerprint which is an aggregation of all the fingerprints composing the respective node. Typically, the nodes are positioned in the rooms and in the hallway. Four cases merit some additional discussion. The first special node is the one in-between Room 2 and Room 3. This node is justified because a door that connects the two hallways is present. A new node is introduced in the hallway between Room 4 and Room 5. The robot detected important changes in the environment due to the vertical pillar present in the hallway node between the Room 7 and Room 8. The door of Room 8 is opened (hallway opening), obstructing the view of the robot and making the environment very different in front

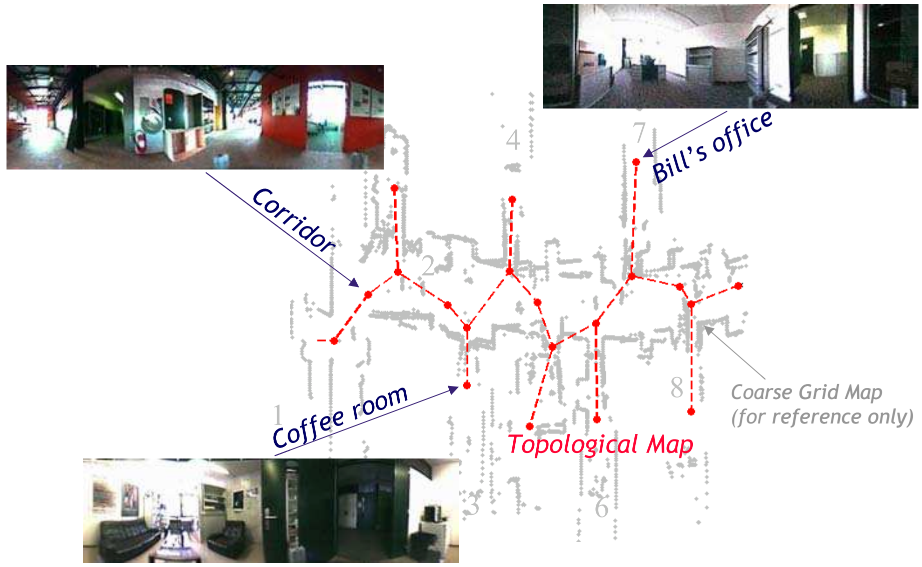

Figure 5: The map of the test environment with the graph representing the topological map. The topological representation is superimposed on a coarse map of the environment.

and behind corridor. Another node that deserves attention is the the door. A new node is therefore automatically introduced by the mapping system. The distance in the corridor between Room 8 and the end point E is quite significant. Since the robot detects distinguishing features due to the changes in this portion of the environment, a new node specifying this is required.

The doors of some rooms remained closed at the time of experimentation; this explains why no node is present in front of and in the respective rooms (sees Figure 4).

The representation thus obtained (see Figure 5) reproduces correctly the structure of the physical space, in a manner that is compatible with the topology of the environment. It also verifies the consistency of the map and permits a distinctive modeling of it. It is important to mention that the map is obtained by using local features only and not by using the topological structure of the environment.

V. Related Work

Various methods have been proposed to represent environments in the framework of autonomous navigation, from precise geometric maps based on raw data or lines to purely topological maps using symbolic descriptions. Each of these methods is optimal with respect to some characteristics but can be very disappointing with respect to other requirements.

Research in map-building can be classified into two major approaches: metric and topological.

Metric maps are spatial representations that have been extensively studied in the robotics community. The stochastic map technique to SLAM [3, 4, 11] and the occupancy grids [18] are typical examples belonging to this kind of space representation. Because these methods are used with high precision sensors, mapping yields a precise localization is accurate. However, metric SLAM (Simultaneous Localization and Mapping) can representation of the environment and consequently become computationally very expensive for large environments. Thrun in [19] proposes probabilistic methods that make the metric mapping process faster and more robust. However, metric approaches also suffer from other shortcomings. One of the main drawbacks of the metric approaches is the cumulative error in the position estimate which renders it unreliable and inaccurate in large environments. Another negative aspect of metric maps is that they are not easily extensible so as to be useable for higher level, symbolic reasoning.

Topological approaches attempt to overcome the drawbacks of geometric methods by modeling space using graphs. Significant progress has been made since the seminal paper by Kuipers [8], where, an approach based on concepts derived from a theory on human cognitive mapping is described as the body of knowledge representing large scale space. Kortenkamp and Weymouth in [7] have also used cognitive maps for topological navigation. They defined the concept of *gateways* which have been used to mark the transition between two adjacent places in the environment. They have used the data from sonars combined with vision information in order to achieve a rich sensory place-characterization. Their work has been an amelioration of Mataric's approach [12], contributing towards the reduction of the perceptual aliasing problem. The improvement is obtained by introducing more sensory information for place representation. A model by Franz, Schölkopf and Mallot [5] was designed to explore open environments within a maze-like structure and to build graph-like representations. Their method has been tested on a real robot equipped with an omni-directional camera. In [6] and [14], the authors have used a model based on a self-organizing map which creates a topological representation of the environment while the robot explores it. Most recently, Beeson et al have used Extended Voronoi Graphs to demonstrate place detection in the context of topological maps [2].

Topological maps are less complex and permit more efficient planning than metric maps. Moreover, they are easier to generate. Maintaining global consistency is also easier in topological maps compared to metric maps. However, the main problems to deal with, when working with topological maps are the perceptual aliasing (i.e., observations at multiple locations are similar) and the automatic establishment of a minimal topology (nodes).

Recently, researchers have integrated both the metric and topological paradigms, thereby obtaining a hybrid system. Thrun, in [19], uses occupancy-grid based maps in order to build the metric map. The topological map is extracted from the grid-based map. Learning a topological representation depends on learning a geometric map, which relies on the odometric capability of the robot. However, in large environments, it is difficult to maintain the consistency of the metric map, due to the drift in the odometry. In [21], Tomatis et al. have conceived a hybrid representation, comprising of a global topological map with local metric maps associated to each node for precise navigation. The authors of [17] have illustrated another hybrid representation, similar to the previously mentioned work. The space is represented as a set of local geometric maps interconnected via a global topological map. The nodes define the visible region of an artificial landmark (i.e., bar-codes placed at a fixed height).

Our method uses fingerprints of places to create a topological model of the environment. The fingerprint approach, by combining the information from all sensors available to the robot, eliminates perceptual aliasing and improves the distinctiveness of places. The topological mapping system, described in this work, relies on fingerprints of places to yield a consistent and distinctive representation of the environment. This fingerprint-based approach is extensible in that it permits spatial cognition beyond just pure navigation.

VI. Conclusions And Future Works

This paper has presented a new technique for automatic and incremental topological mapping with fingerprints of places. The fingerprint provides a compact and distinctive methodology for space representation and place recognition – it permits encoding of a huge amount of place-related information in a single circular sequence of features. This representation is suitable for both indoor and outdoor environments. The experiments verify the efficacy and reliability of our approach. Some future works will focus on the extension of the whole approach towards topological SLAM (Simultaneous Localization and Mapping). In addition, efforts will also be directed towards testing the sensitivity of the currently chosen heuristic, with respect to indoor and outdoor conditions.

Acknowledgments

This work was supported by the European project BIBA IST-2001-32115 EU project.

References

[1] Arras, K.O. and Siegwart, R, Feature Extraction and Scene Interpretation for Map-Based Navigation and Map Building, In Proceedings of the Symposium on Intelligent Systems and Advanced Manufacturing, Pittsburgh, USA, October 13-17,1997.

[2] Beeson, P., Nicholas K. Jong, and Benjamin Kuipers. (2005). Towards Autonomous Topological Place Detection Using the Extended Voronoi Graph. IEEE International Conference on Robotics and Automaton, April 2005.

[3] Castellanos J.A., Tardos J.D. (1999), Mobile Robot Localization and Map Building: Multisensor Fusion Approach, Kluwer.

[4] Dissanayake, Newman, Clark, Durrant-Whyte and Csorba (2001), A Solution to the Simultaneous Localization and Map Building (SLAM) problem, IEEE Trans. On Robotics and Automation, Vol 17, No.3, June.

[5] Franz, M.O., Schölkopf, B., Mallot, H.A. and Bülthoff.(1998). Learning view graphs for robot navigation, Autonomous Robots 5 111-125.

[6] Hafner V.V. (2000), Learning Places in Newly Explored Environments, in Meyer, Berthoz, Floreano, Roitblat and Wilson (Eds.), SAB2000 Proceedings Supplement Book, Publication of the International Society for Adaptive Behavior, Honolulu.

[7] Kortenkamp, D. and Weymouth, T. (1994), Topological mapping for mobile robots using a combination of sonar and vision sensing, In Proceedings of AAAI-94, Seattle, WA.

[8] Kuipers, B. J. (1978), Modeling Spatial Knowledge, Cognitive Science, 2: 129-153, 1978.

[9] Lamon, P., I. Nourbakhsh, B. Jensen and R. Siegwart (2001), Deriving and Matching Image Fingerprint Sequences for Mobile Robot Localization, IEEE International Conference on Robotics and Automation, Seoul, Korea.

[10] Lamon, P., Tapus A., Glauser E., Tomatis N., Siegwart R. (2003), Environmental Modeling with Fingerprint Sequences for Topological Global Localization, In Proceedings of the IEEE International Conference on Intelligent Robot and Systems, Las Vegas, USA, October 27-30.

[11] Leonard J.J, H.F. Durrant-Whyte (1992), Directed Sonar Sensing for Mobile Robot Navigation, Kluwer Academic Publishers, Dordrecht.

[12] Mataric, M. J. (1991), Navigating with a rat brain: A neurobiologically-inspired model for robot spatial representation. In: J.A.Meyer, S.W.Wilson (Eds), From Animals to Animats, MIT Press, Cambridge, MA.

[13] Needleman, S. and Wunsch, C. (1970), A general method applicable to the search for similarities in the amino acid sequence of two proteins, J. Molecular Biology, 48:443-453.

[14] Owen, C. and Nehmzow, U.(1998), Landmark-based navigation for a mobile robot, in : Meyer, Berthoz, Floreano, Roitblat and Wilson (Eds.), From Animals to Animate 5, Proceedings of SAB'98, MIT Press, Cambridge, MA, pp. 240-245.

[15] Tapus A., Tomatis N. and Siegwart R. (2004), Topological Global Localization and Mapping with Fingerprints and Uncertainty, In Proceedings of the International Symposium on Experimental Robotics, Singapore, June 18-21.

[16] Tapus A., Ramel G., Dobler L. and Siegwart R. (2004), Topology Learning and Recognition using Bayesian Programming for Mobile Robot Navigation In Proceedings of the IEEE International Conference on Intelligent Robot and Systems, Sendai, Japan.

[17] Taylor C. J. and Kriegman D.J. (1998), Vision-Based Motion Planning and Exploration Algorithms for Mobile Robots, IEEE Transactions on Robotics and Automation, Vol. 14, No 3, June 1998, pp 417-426.

[18] Thrun, S. (1998), Learning metric-topological maps for indoor mobile robot navigation. In Artificial Intelligence 99(1):21-71.

[19] Thrun, S. (2000), Probabilistic algorithms in robotics. In Artificial Intelligence Magazine 21(4):93-109.

[20] Tolman, E. C. (1948), Cognitive maps in rats and men, Psychological Review, 55:189-208.

[21] Tomatis, N., I. Nourbakhsh, and R. Siegwart (2003). Hybrid simultaneous localization and map building: a natural integration of topological and metric. Robotics and Autonomous Systems 44:3-14.