

Multi-resolution SLAM for Real World Navigation

Agostino Martinelli, Adriana Tapus, Kai Arras, Roland Y. Siegwart

▶ **To cite this version:**

Agostino Martinelli, Adriana Tapus, Kai Arras, Roland Y. Siegwart. Multi-resolution SLAM for Real World Navigation. Paolo Dario ; Raja Chatila. Robotics Research. The Eleventh International Symposium, Springer, pp.442-452, 2005, Springer Tracts in Advanced Robotics, 978-3-540-23214-8. 10.1007/11008941_47 . hal-01015921

HAL Id: hal-01015921
https://hal.science/hal-01015921

Submitted on 27 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Proceedings of the International Symposium of Research Robotics
Siena, Italy, October 2003

Multi-resolution SLAM for Real World Navigation

Agostino Martinelli, Adriana Tapus, Kai Olivier Arras, and Roland Siegwart

Autonomous Systems Lab
Swiss Federal Institute of Technology Lausanne (EPFL)
CH-1015 Lausanne

Multi-resolution SLAM for Real World Navigation

Agostino Martinelli, Adriana Tapus, Kai Olivier Arras, and Roland Siegwart

Autonomous Systems Lab, Swiss Federal Institute of Technology Lausanne (EPFL),
CH-1015 Lausanne

Abstract

In this paper a hierarchical multi-resolution approach allowing for high precision and distinctiveness is presented. The method combines topological and metric paradigm. The metric approach, based on the Kalman Filter, uses a new concept to avoid the problem of the drift in odometry. For the topological framework the fingerprint sequence approach is used. During the construction of the topological map, a communication between the two paradigms is established. The fingerprint used for topological navigation enables also the re-initialization of the metric localization. The experimentation section will validate the multi-resolution-representation maps approach and presents different steps of the method.

1 Introduction

Many methods have been proposed to represent an environment in the framework of autonomous navigation, from precise geometric maps based on raw data or lines up to purely topological maps using symbolic descriptions. Each of these methods is optimal concerning some characteristics but can be very disappointing with respect to other requirements. In particular, none of them is able to cope with the large variety of environments that humans encounter in their daily live.

Hierarchical Multi-Resolution maps

Most current approaches make a trade-off between precision and global distinctiveness. Precision and distinctiveness have a strong link with the level of abstraction of the features used for navigation (fig. 1). Raw data represent the lowest level in the hierarchy of abstraction. Localization and mapping with raw data can result in very high precision of the represented environment, but the required data volume scales very badly with the size of the environment and the distinctiveness of the individual data points is very low. An example of such a approach is Markov localization [9, 15]. The second level of abstraction corresponds to the geometric features (lines, edges). The stochastic map technique to SLAM [6, 8, 14] and the multi-hypothesis localization [3] are typical examples belonging to this level. These approaches still feature high precision with reduced memory requirements, but have shortcomings concerning global distinctiveness and unmodeled events. Partially geometric features correspond to the third level of the hierarchy. Representations using partially geometric features are for examples fingerprints (a sequence of low level features) described in [12,13], but also more bio-inspired approaches [1, 5, 10] using neural networks, e.g. the work by Hafner [10] uses a neural network to create a topological map of the environment based

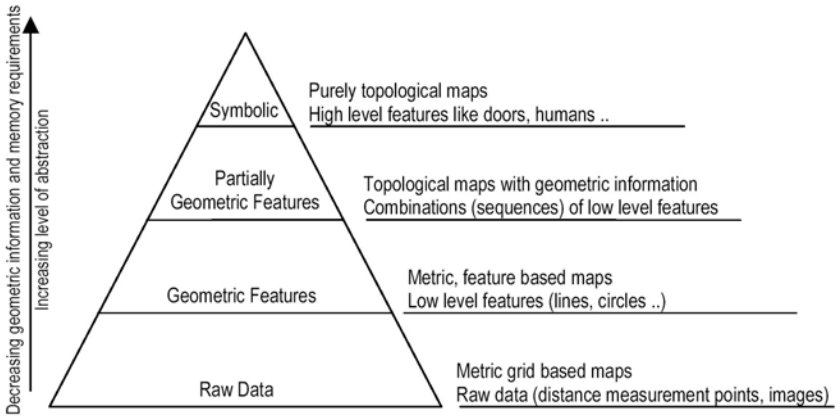


Fig. 1. This figure depicts the hierarchy of abstraction levels. More we go up in the hierarchy, more we reduce the geometric information and more we increase the distinctiveness. For global localization and mapping, high distinctiveness is of importance, whereas for local action, precise geometric relations with the environment come forward.

on images of an omni-directional camera. On the highest abstraction level the environment is represented by a symbolic description. This can be very compact and distinctive, but reliable tools for extraction of high level features are still missing.

The aim of the present work is to suggest and discuss a multilevel representation approach for robust and compact environment representation. This is realized by integrating different level of abstraction and representations in a common navigation framework. For example, to get from the airport to a given room in a hotel, maps of the city streets and of the hotel are necessary. Clearly, these maps are on different scales and different representations (e.g. metric/topological). A crucial point is then the linkage between two maps of different resolution. In the multi-level representation we propose, the mobile robot uses the most suitable level of abstraction, depending on the task it has to accomplish and the certainty about its pose. Moreover, depending on the actual situation, it has to be able to switch from one level to another one (see figure 1).

In this paper we will discuss the possibility to integrate most recent results in mobile robot navigation in a multi-level hierarchical representation and we will present our first attempts towards multi-resolution, multi-representation maps.

2 First attempt for a multi-resolution Map

2.1 Metric Mapping based on a Relative Map Filter

In this section we introduce a new approach to solve the SLAM problem in the framework of the stochastic map based on the concept of the relative map. The aim of this method is to minimize the loop consistency problem (i.e. the global convergence and consistency of the built map in large environment - large meaning when the size of the environment is much larger than the range of the adopted external sensor). Clearly, even with this approach the problem of the consistency of the built map cannot be completely solved and therefore, for very

large environment, it is necessary to combine the metric approach with a topological approach (see the next sections) and several local maps give the global metric description. However, the relative map introduced here increases the convergence properties of the map with respect to the standard approach since the convergence of the map built following the standard approach [8] requires infeasible hypothesis (odometry perfectly modelled and linearity of the observation in the estimated state (see [8] for details)).

The basic idea consists of introducing a map state, which only contains quantities invariant under translations and rotations. This is the only way in order to have an actual decoupling between the robot and the landmark estimation and therefore to do not rely the landmark estimation on the unmodeled error sources in the robot motion. Once a relative map is estimated and the absolute location of a set of landmarks is known (e.g. by using the first observation) it is possible to build the absolute map. Therefore, the entire method contains two algorithms. The former estimates the relative map, the latter builds the absolute map. Only the case of point landmark is here considered, although the same idea could be applied to other kind of landmark. In the sections 2.1.1 and 2.1.2 respectively we describe the two algorithms.

2.1.1 The Relative Map Filter

The state estimated through this filter only contains the distances between the point landmarks. Clearly, the distance is a quantity invariant under translation and rotation, i.e. it is independent of the robot configuration.

Let denote with d the state and with P its covariance matrix. In Fig 2a the vector d contains the marked distances between the *6* landmarks. Clearly, not all of the distances between the *6* landmarks are stored in d because not all the landmarks were observed together at the same time. At a given time step, the observation consists of a set of distances between the landmarks observed by the robot through its external sensor (Fig 2b). Clearly, these distances may be already observed (i.e. can be in the vector d) or may not. Let introduce the following notation:

$$d_{old}=[u, w_{old}]^T \qquad d_{obs}=[w_{obs}, v]^T \tag{1}$$

where d_{old} is the state estimated at a given time step and d_{obs} is the observation at the same time step, containing a set of distances between the landmarks observed by the robot. u contains the distances which are not re-observed (i.e. which do not appear in the vector d_{obs}) and w_{old} contains the distances re-observed (denoted by w_{obs} in the vector d_{obs}). Finally, v, contains the distances observed for the first time at the considered time step.
The covariance matrix of the previous vectors are:

$$P_{old}=\begin{bmatrix} P_{uu} & P_{uw} \\ P_{uw}^T & P_{ww} \end{bmatrix} \qquad P_{obs}=\begin{bmatrix} R_{ww} & R_{wv} \\ R_{wv}^T & R_{vv} \end{bmatrix} \tag{2}$$

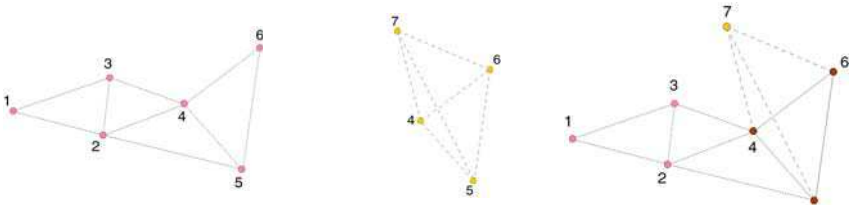


Fig. 2. Relative Map before the observation (a), the observation (b), and the relative map obtained by fusing the information coming from the old map and the observation (c). In all the three figures the map state only contains the indicated distances between the landmarks

We adopt the following notation to denote the estimated quantities, obtained by fusing the old state with the observed one (the new estimated distances are depicted in Fig2c).

$$d_{new}=[u_{new}, w_{new,} v_{new}]^T \tag{3}$$

$$P_{new} = \begin{bmatrix} Pn_{uu} & Pn_{uw} & Pn_{uv} \\ Pn_{uw}^T & Pn_{ww} & Pn_{wv} \\ Pn_{uv}^T & Pn_{wv}^T & Pn_{vv} \end{bmatrix} \tag{4}$$

We obtain the new estimation for the state and its covariance matrix by applying the equations of the Kalman filter.

$$u_{new} = u + P_{uw}\,(P_{ww} + R_{ww})^{-1}(w_{obs} - w_{old}) \tag{5}$$
$$w_{new} = w_{old} + P_{ww}\,(P_{ww} + R_{ww})^{-1}(w_{obs} - w_{old}) \tag{6}$$
$$v_{new} = v + R_{vw}\,(P_{ww} + R_{ww})^{-1}(w_{old} - w_{obs}) \tag{7}$$
$$Pn_{uu} = P_{uu} - P_{uw}\,(P_{ww} + R_{ww})^{-1}\,P_{wu} \tag{8}$$
$$Pn_{uw} = P_{uw} - P_{uw}\,(P_{ww} + R_{ww})^{-1}\,P_{ww} \tag{9}$$
$$Pn_{uv} = 0 \tag{10}$$
$$Pn_{ww} = P_{ww} - P_{ww}\,(P_{ww} + R_{ww})^{-1}\,P_{ww} \tag{11}$$
$$Pn_{wv} = R_{wv} - R_{ww}\,(P_{ww} + R_{ww})^{-1}\,R_{wv} \tag{12}$$
$$Pn_{vv} = R_{vv} - R_{vw}\,(P_{ww} + R_{ww})^{-1}\,R_{wv} \tag{13}$$

2.1.2 Recovering the Absolute Landmark Location

We adopt a simple linear method to recover the absolute landmark locations starting from the absolute location of three or more landmarks and the state estimated by the previous filter, which contains the distances between the landmarks. At a given time step the absolute locations of a set of landmarks are available (we assumed that the absolute coordinates of at least three landmarks are known at the beginning; these coordinates could be provided by the first observation). The aim is to estimate the location of a new landmark denoted by *j*. We extract from the previous set a subset containing the landmarks whose distance from the landmark *j* is provided by the relative filter. Let denote the

locations of these landmarks by (x_i, y_i) and the distance between the landmark j and the landmark i of this subset by d_i. We assumed that the number of the elements of this subset is equal to n. If $n<3$ the absolute location of the landmark j cannot be provided. Clearly, we have for the landmark i

$$d_i^2 = (x_j - x_i)^2 + (y_j - y_i)^2 \tag{14}$$

We therefore obtain a linear system in the unknowns x_j and y_j by considering all the differences $d_i^2 - d_1^2$. By applying recursively this method for all the landmarks (all j) it is possible to get their absolute location

2.2 Topological Map

The topological approach tries to give a compact representation since only distinctive places within the environment are encoded. One of the main problems of this method is the perceptual aliasing (i.e. distinct locations within the environment appearing identical to the robot's sensors). In order to be capable to eliminate the perceptual aliasing problem, the distinctiveness of the space must be improved. To maximize the reliability in navigation, the information from all the sensors that are available to the robot must be used. To solve this problem the notion of fingerprint as described in [12, 13] is used. This characterization of the environment is especially interesting when used within a topological localization and multiple modalities framework. 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 the instance of a specific feature type. In our case we choose to extract color patches and vertical edges from visual information and corners and beacons from laser scanner. We decided to use the letter 'v' to characterize an edge, the letters A,B,C,...,P to represent hue bins, the letter 'c' to characterize a corner feature and the letter 'b' to characterize a beacon feature. Details about the visual features extraction can be found in [13] and laser scanner features extraction can be found in [4].

2.2.1 Fingerprints matching for localization

The string-matching problem is not an easy one. Usually strings don't match exactly because the robot may not be exactly located on a map point and/or some changes in the environment or perception errors occurred. The standard algorithms are quite sensitive to insertion and deletion errors, which cause the string lengths to vary significantly.

Minimum energy algorithm

The approach adopted for sequence matching is inspired by the minimum energy algorithm used in stereo-vision for finding pixels in two images that correspond to the same point of a scene [11]. As in the minimum energy case, the problem can be seen as an optimization problem, where the goal is to find the path that spends the minimum energy to go from the beginning to the end of the first sequence

considering the values of the second one. The similarity between two sequences is given by the resulting minimum energy of traversal.

2.2.2 Fingerprint generation

The fingerprint generation is done in three steps (see Fig. 3). The first step consists in extracting the different features from the panoramic image and the laser scan: vertical edges, patches of colors, corners and beacons. The features are arranged in an array along with their corresponding position (from 0 to 359 degrees). At this stage a new type of feature that reflects a correspondence between a corner and an edge: the feature 'f', is introduced, more details can be found in [13]. The ordering of the features in a fingerprint sequence is highly informative and for that reason the notion of angular distance between two consecutive features will be added. This adds geometric information and increases once again the distinctiveness between the fingerprints. Therefore, we decided to introduce a new type of feature, the empty space feature 'n', for reflecting angular distance. Each 'n' covers the same angle of the scene (20 degrees). So, as many empty spaces as needed must be inserted in order to fill the gap between two consecutive features. This insertion is the last step of the fingerprint generation.

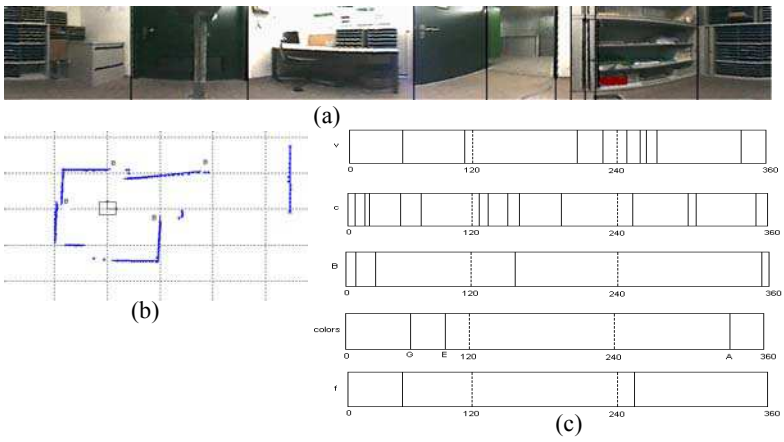


Fig. 3. Fingerprint generation. (a) panoramic image with the vertical edges and color patches detection 'v' and color (b) laser scan with extracted corners 'c' and beacons 'b' (c) the first three images depict the position (0 to 360°) of the vertical edges, the corners, the beacons and the colors (G-green, E-light green, and A-red) respectively. The fifth 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: cbccbnfGcnEnvccncbcvncnnfvvvnccAcb

2.3 Multi-resolution Map

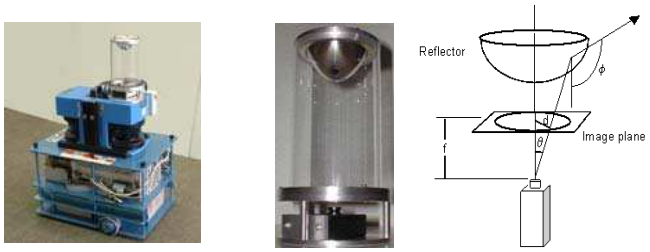
Our approach combines low-level feature for local metric and fingerprints for global topological localization and mapping. Indeed, the combination of these two levels of abstraction is necessary, because the consistency in large environments with stochastic maps is impossible. The relative map approach described in Section 2.1 increases the convergence property. The estimation does not rely on the odometry and therefore on the unmodeled events in the robot motion. However, this method fails when the unmodeled events (e.g. strong collision) are strong enough to cause error in data association. Clearly, the stochastic maps completely fail when there is a kidnapping during the motion.

The main problem that arises when there are several levels of abstraction is the switching between the levels. In our case, this appears when we switch from topological to metric. In order to deal with the switching problem, we introduce a new element in the fingerprint approach. The new fingerprint contains the fingerprint as illustrated in Section 2.2, and the local coordinates of the correspondent beacons as appearing in the local metric map. To enable the generation of this new fingerprint, a communication between the two levels of abstraction is required.

When a kidnapping is performed, in order to find the new location of the robot, the fingerprint approach will be used. Once the location is found, the association between the local coordinates of the beacons (contained in the matched fingerprint) and the observed fingerprint is done. In this way the re-initialization of the metric navigation is accomplished. This multi-representation approach enables SLAM on two levels of abstraction, thus featuring precision and global consistency. A similar concept can be found in [16].

3 Experimental Results

For the experiments, Donald Duck (see Fig. 4a), a fully autonomous mobile robot, has been used.



a) b)

Fig. 4. System used for experimentation: a) The fully autonomous robot Donald Duck. b) The panoramic vision system. The camera has a 640x480 pixels resolution and an equiangular mirror is used so that each pixel in the image covers the same view angle

The robot is equipped with wheel encoders, a 360° laser range finder and an omnidirectional camera. The panoramic vision system depicted in figure 4b uses a mirror-camera system to image 360° in azimuth and up to 110° in elevation. The use of an omnidirectional camera combines the advantages of the SICK laser range finder (e.g. an angle of view of 360°) and the capability of detecting verticals. Twenty-two beacons were placed in the environment. They could be detected by the laser sensor with an accuracy of around 5 cm. There were adopted to create the point landmarks mentioned in section 2. Clearly, instead of them a function able to extract corners from laser scanner could be used.

Metric Map

Figure 5 shows the metric map obtained through the relative map filter described in the section 2.1.1. The robot visited three rooms and estimated at each time step its configuration and the positions of the *22* beacons in the environment. Since the odometry was completely decoupled by the estimation process (as explained in the section 2.1) there is not any drift in the built map and in particular it was possible to adopt a unique global metric map. During the experiment the odometry data were only used to solve the data association problem and not in the estimation process. The initial robot configuration coincides with the origin of the global reference whose axis were chosen coincident with the axis of the robot at the initial time. The grey line shows the robot trajectory computed by using the same algorithm described in the section 2.1.1 and the symbol 'O' is adopted to indicate the absolute beacon position.

Topological map

In order to validate the fingerprint approach, for each of the four rooms (see Figure 5), a fingerprint has been extracted. This experiment has been repeated six times for each room, and the fingerprint has been extracted in different locations of the room. Three fingerprints per room have been included in a database as reference (map). The others 12 fingerprints (3 per room) have been matched to the database for testing the localization. For a given observation (fingerprint), a match is successful if the best match with the database (highest probability) corresponds to the correct room. For our experiments the match was successful for all the test fingerprints and the correct room has always been found with the highest probability, a probability going from 0.67 to 0.92.

Table 1. Probabilities for matches between the database fingerprints and the observed fingerprints

Fingerprint	Room1	Room2	Room4
Room1_1	0.72	0.37	0.50
Room1_2	0.57	0.64	0.39
Room1_3	0.61	0.66	0.44
Room2_1	0.51	0.69	0.38
Room2_2	0.49	0.92	0.48
Room2_3	0.39	0.82	0.49
Room4_1	0.42	0.48	0.65
Room4_2	0.56	0.32	0.79
Room4_3	0.49	0.58	0.55

Kidnapping and metric re-initialization

Three kidnapping were performed, respectively in Room1, Room2 and Room4 (see figure 5). In the Table 1 the results of the matching algorithm presented in Section 2.2.1, are normalized to have a probability (between 0 and 1). As it can be seen the correspondence between the real location and the observed one is always found with a high probability.

Concerning the metric re-initialization, the results are showed in the figure 5, and the estimated robot positions are indicated with symbol '*'.

The error between the metric position that have been found and the real position of the robot is very small, the error going from 1 to 4 cm.

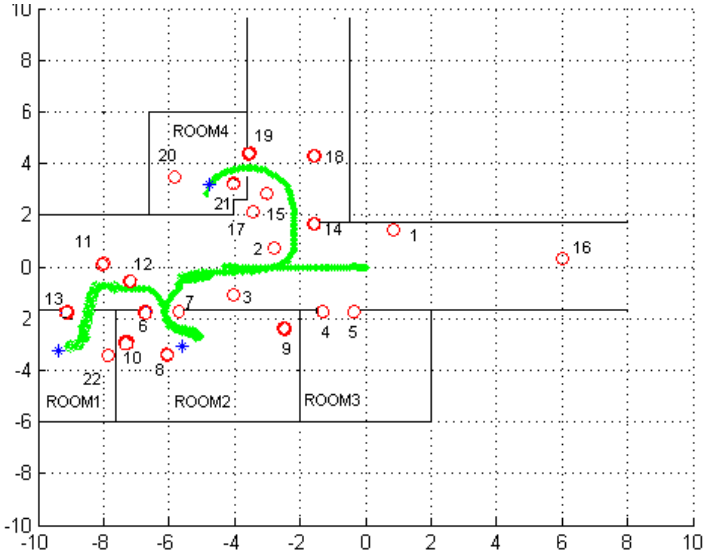


Fig. 5. Local map of the environment with the way traversed by the robot, with the numeroted seen beacons and with the estimated metric position '*'

4 Conclusion and Outlook

This paper presented a multi-level hierarchical representation. The new approach depicted combines metric and topological concept. The two paradigms are on two different levels of abstraction. This kind of representation allows a very compact and computationally efficient representation of the environment for mobile robot navigation. The results of our first attempts towards multi-resolution maps showed the robustness of the approach. Future works will focus on experiments in large environments, on the introduction of an average fingerprint instead of several fingerprints for each node, on a probabilistic fingerprint-matching concept, on an automatic generation of nodes and on the extension of the relative map filter to line landmark.

References

1. Arleo A. and Gerstner W. (2000), Spatial Cognition and Neuro-Mimetic Navigation: a model of Hippocampal Place Cell Activity, Biological Cybernetics, 83, 287-299
2. Alfred V. Aho (1990), Algorithms for finding patterns in strings, In J. van Leeuwen, editor, Handbook of Theoretical Computer Science, chapter 5, pages 254-300. Elsevier Science Publishers B. V.
3. K.O.Arras, Castellanos, J.A., Schilt, M. and Siegwart, R. (2003), Feature-based multi-hypothesis localization and tracking using geometric constraints, Robotics and Autonomous Systems, 1056 (2003) 1-13.
4. K.O.Arras, and Siegwart, R. (1997), Feature Extraction and Scene Interpretation for Map-Based Navigation and Map Building, Proc. of SPIE, Mobile Robotics XII, Vol. 3210, p-42-53.
5. Berthoz A. (1997), Le sens du mouvement, Editions Odile Jacob, Paris.
6. Castellanos J.A., Tardos J.D. (1999), Mobile Robot Localization and Map Building: Multisensor Fusion Approach, Kluwer.
7. J. Crowley (1989), World modeling and position estimation for a mobile robot using ultrasonic ranging, In Proceedings of IEEE International Conference on Robotics and Automation, pp. 674--680, Scottsdale, AZ, May 14-19.
8. 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.
9. Fox, D. (1998), Markov Localization: A Probabilistic Framework for Mobile Robot Localization and Navigation, Ph.D. Thesis, Institute of Computer Science III. Bonn, Germany, University of Bonn.
10. 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.
11. Kanade, T., Ohta, Y. (1985), Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming, IEEE Transactions on pattern analysis and machine intelligence, Vol PALMZ No 3.
12. Lamon, P., I. Nourbakhsh, Jensen B., Siegwart R. (2001), Deriving and Matching Image Fingerprint Sequences for Mobile Robot Localization, In Proceeding of the IEEE International Conference on Robotics and Automation (ICRA), Seoul, Korea.
13. 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.
14. J.J. Leonard, H.F. Durrant-Whyte (1992), Directed Sonar Sensing for Mobile Robot Navigation, Kluwer Academic Publishers, Dordrecht.
15. S. Thrun, D. Fox, W. Burgard, and F. Dellaert (2001), Robust Monte Carlo Localization for Mobile Robots, In Artificial Intelligence (AI).
16. Tomatis N., Nourbakhsh I. and Siegwart R. (2002), Hybrid Simultaneous Localization and Map Building: Closing the Loop with Multi-Hypotheses Tracking, In Proceedings of the IEEE International Conference on Robotics and Automation, Washington DC, USA, May 11 - 15.