

Topological navigation and qualitative localization for indoor environment using multi-sensory perception

P. Ranganathan^a, J.B. Hayet^b, M. Devy^{b,*}, S. Hutchinson^a, F. Lerasle^b

^a Beckman Institute, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA

^b LAAS-CNRS, Toulouse, France

Abstract

This article describes a navigation system for a mobile robot which must execute motions in a building; the robot is equipped with a belt of ultrasonic sensors and with a camera. The environment is represented by a topological model based on a Generalized Voronoi Graph (GVG) and by a set of visual landmarks. Typically, the topological graph describes the free space in which the robot must navigate; a node is associated to an intersection between corridors, or to a crossing towards another topological area (an open space: rooms, hallways, . . .); an edge corresponds to a corridor or to a path in an open space. Landmarks correspond to static, rectangular and planar objects (e.g. doors, windows, posters, . . .) located on the walls. The landmarks are only located with respect to the topological graph: some of them are associated to nodes, other to edges. The paper is focused on the preliminary exploration task, i.e. the incremental construction of the topological model. The navigation task is based on this model: the robot self-localization is only expressed with respect to the graph.

© 2002 Published by Elsevier Science B.V.

Keywords: Topological navigation; Generalized Voronoi Graph; Visual landmarks

1. Introduction

Navigation is a critical task for a mobile robot to allow it to move and act autonomously in its environment. Because internal sensors on the robot are not accurate enough or may give false measurements, a navigation system must be based on exteroceptive sensors like cameras, sonars or laser range finders.

As opposed to the classical methods based on explicit localization of the robot with respect to the environment, other methods [6,9] make the robot localization relative only to discriminant features learned and successively perceived by the robot or

relative to an area (environment modeling in topologically independent areas: corridors, room, . . .) [1]; the continuity of a path is guaranteed by a graph which expresses some relationships between landmarks; for example, landmark A is connected to landmark B only if B is visible from A, or if a sensor-based motion (wall following, visual servoing, for example) can be executed to go from A to B. This kind of approach could be more generally embedded in the family of *qualitative* or *topological* navigation methods. Note that these methods alone will never be sufficient to provide a truly reliable navigation system in a general indoor environment but have to be integrated in a more general, adaptive system as the ones described in [5].

This paper proposes such a topological navigation ability. A service robot must execute motions in an office environment, so the Generalized Voronoi

* Corresponding author. Tel.: +33-5-61336331;

fax: +33-5-61336455.

E-mail address: michel@laas.fr (M. Devy).

Graph (GVG) representation proposed by Choset and coworkers [2,7] could be well adapted to solve the navigation problem; in such a graph, nodes are associated to transitions between areas (corridor crossings, area entrances, doors, . . .); an edge typically corresponds to a path in a corridor or in an open space. In a corridor, the robot motion can be controlled using sonars to maintain the robot on the GVG; the robot localization is expressed according to the GVG (the robot is on this node or is moving on this edge).

Nevertheless, a self-localization problem may occur because this kind of environment is very ambiguous using only sonars whenever human presence or topological modification (an open door) may occur. If the robot is equipped with several sensors—in our experiment, monocular vision and sonars—it can take advantage of different topological representations (visual landmarks and GVG) to validate an hypothesis about a node recognition. Vision gives stable, reliable information from a large part of the environment, which may be helpful in comparison with ultrasonic sensors.

Kortenkamp and Weymouth [4] have already presented results combining these two sensors. Pre-determined forms of *gateways* are searched with sonar and these distinct places are associated with some simple visual landmarks. The learning phase was not done autonomously, and the processing steps of sonar data made the algorithm usable in only orthogonal corridors. Our contribution is two-fold:

- we make the robot *learn autonomously* the environment model, without preliminary guided route traversals and with extended environment structural configurations, although considering only corridors-based environment;
- we use a visual landmark intrinsic representation *independent from the viewpoint* and as stable as possible with respect to illumination, scale changes and small occlusions.

Section 2 proposes an overview on the environment representation and the navigation system. Sections 3 and 4 present our strategy to build a hybrid topological map—landmark-based and GVG; in Section 5, experimental results are commented. Finally in Section 6, discussions about this work and some future researches are considered.

2. Overview of our approach

A topological map represents the robot environment by a graph. Paths are defined as sets of two distinct points, or “places” which must be detected and recognized by the robot using sensors data. These points provide the nodes of the map. Only a few relevant information about the places are required to locate and identify them. The edges between two nodes correspond to navigation operations such as wall following, visual servoing [8], . . . These navigation operations take the robot from one node to another. Such a representation has a lot of advantages:

- the node information may be meet points, landmarks detected by a vision system or any other distinguishing features of the environment that can be reliably recognized by a robot;
- the path between two nodes does not have to be traced exactly; it is sufficient if the robot can traverse a general path (not exactly defined) between two nodes;
- small storage capacities are required, since only information about the nodes are stored;
- there is no need to maintain a global coordinate frame, so this method is suitable for exploring large-scale environments;
- path planning from a topological map can be very fast and without complex computations.

However, two important properties have to be guaranteed: first, the nodes have to be detected and identified with certainty and accuracy, and secondly, the navigation operations must lead the robot from one node to another. In this paper, we focus on the first problem. Our aim is to link some other information in the graph nodes so that failures in the node recognition could not occur.

The GVG (see Fig. 1) has been popularized by Choset et al. [2] with sonar sensors; such a graph can also be built using a laser range finder [10]. The nodes are the so-called meet points and the arcs correspond to the Generalized Voronoi Diagram, i.e. the locus of all points that are equidistant to two object boundaries. At the intersection of two GVG edges, the meet point is defined to be equidistant to at least three points. A key feature that makes the GVG so useful for mobile robot navigation is that it can be constructed incrementally by using only sensor data and line of sight information.

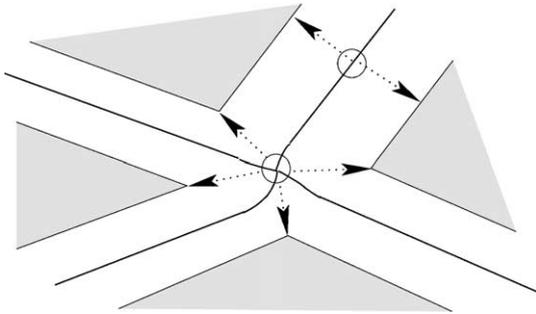


Fig. 1. The GVG: edges and meeting points.

As we have seen it, sonar information may not be sufficient to recognize all nodes. As our robot is navigating in an indoor environment, we will find mainly vertical planes and a lot of vertical structures: doors, windows, posters on walls, That is what justifies our choice of planar, quadrangular landmarks. Furthermore, this kind of 2D primitive is stable and easily recognizable under very different viewpoints and under some basic assumptions: the camera roll angle will remain close to zero and the tilt value remains constant. It means that we are able to find the class of projected verticals directions in images, we will call them “pseudo-vertical” directions.

3. Learning the topological graph

The exploration task consists in going over every path in the environment, memorizing the path connections in a GVG and learning some visual landmarks at the proximity to every meet point. From such a point at least two paths begin. Hereafter the different steps of the exploration task are listed.

Meet point detection. When it goes down an unexplored corridor, the robot is controlled to be on the GVG (between the two closer obstacles, typically the walls of a corridor); a meet point is detected if at least three points appear to be closer from the robot than the other ones in the same point vector acquired by the ultrasonic belt. Using the method proposed in [2], the meet point (point equidistant to the obstacles) is computed and an incremental honing strategy drives the robot to the precise meet point.

Post honing process. Once the meet point has been reached, the robot makes a 360° pan rotation to see if any planar landmark can be detected. If so, the robot identifies it (or not), as described in Section 4, two cases have to be distinguished.

New nodes. If no current landmarks were matched against the previously seen ones, a new node n is created and the “best” landmark (highest saliency/stability) \hat{L}^n is selected to serve as a local

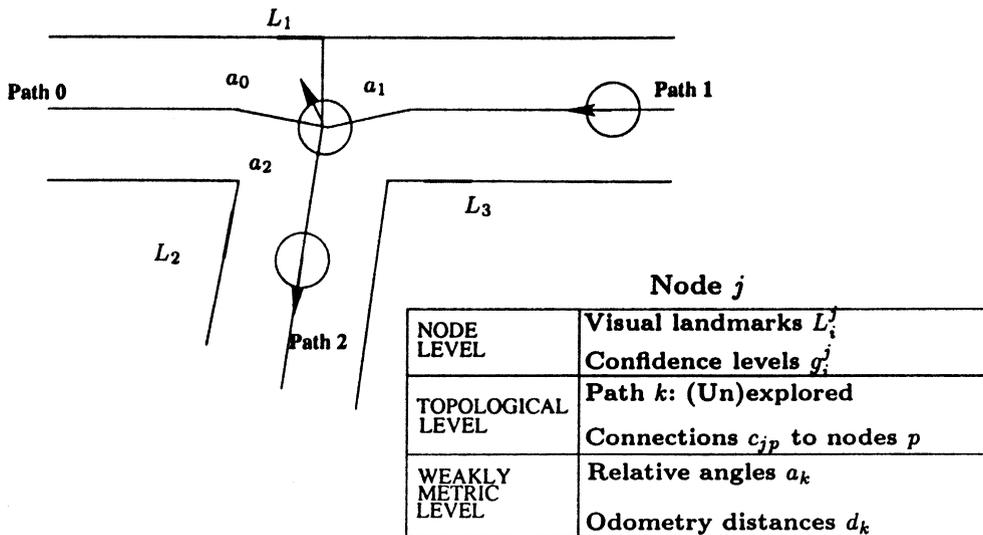


Fig. 2. Meet point and visual landmark detection.

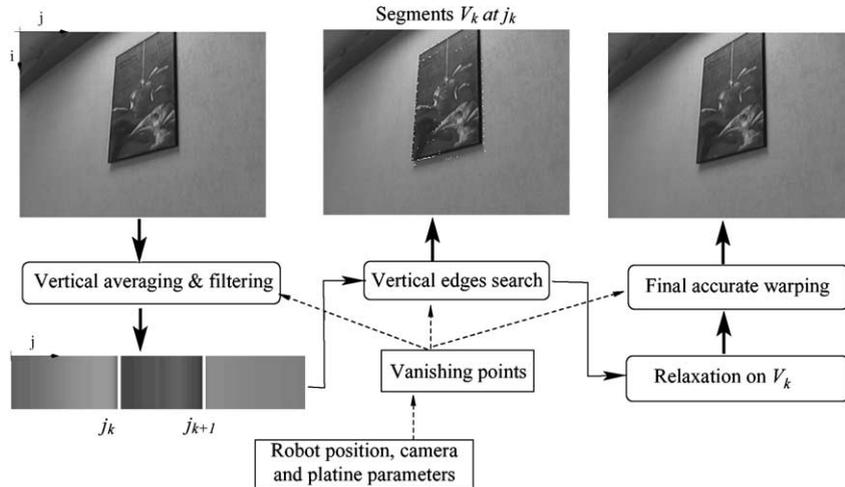


Fig. 3. Landmark detection.

reference orientation in relation to which the robot localizes itself: all departing paths from this node are defined with respect to it (see Fig. 2). Landmarks are noted L_i^j , where i refers to a view i (subscript) at node j (superscript). Node information (landmarks), pure connectivity information (paths set to “explored”) and weakly metric information (rough angles, odometry between nodes) are updated, as described in Fig. 2.

Known node. If a landmark is found to be matched with a previously visited node, the robot localizes itself on it, and the departing paths are identified. The robot departs in one of the unexplored paths. If all paths have been previously explored, the robot looks for the node at the shortest distance from it with unexplored paths and moves in that direction.

4. Learning landmarks

To learn a model of a landmark on an autonomous way, we need on the one hand to set criteria and methods to detect this landmark and on the other hand to build a model that will be reusable for recognition.

4.1. Landmarks detection

The principle of landmark detection is illustrated in Fig. 3, to be read from left to right. The idea is to search the quadrangles pseudo-vertical edges projected on a 1D image resulting from an averaging operator, with

the help of one or more vanishing points. This 1D image is correlated with a step-like reference signal to isolate discontinuities. Each of them allows to get a full pseudo-vertical segment in the 2D image. Indeed, the lines corresponding to the selected j_k j -coordinates are regularly sampled and step-like transitions around these sampled points are searched. After segmentation, we get the approximate segments V_k , as illustrated in Fig. 3.

Among all the selected vertically-oriented segments, indexed by j_k as seen in Fig. 3, we form couples (j_k, j_l) corresponding to potential landmarks thanks to a discrete relaxation scheme. We define *a priori probabilities* for pairs (V_k, V_l) depending on geometric, photogrammetric, projective properties of segments V_k and V_l and *compatibilities* between pairs of segments. As an example, two potential landmarks may be related only by a relationship of full inclusion or of no intersection, as seen in Fig. 4. Discrete

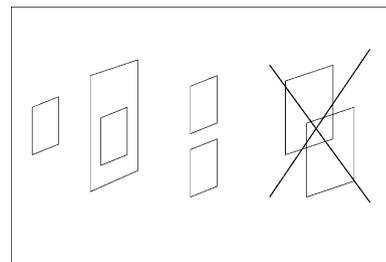


Fig. 4. Accepted configurations.

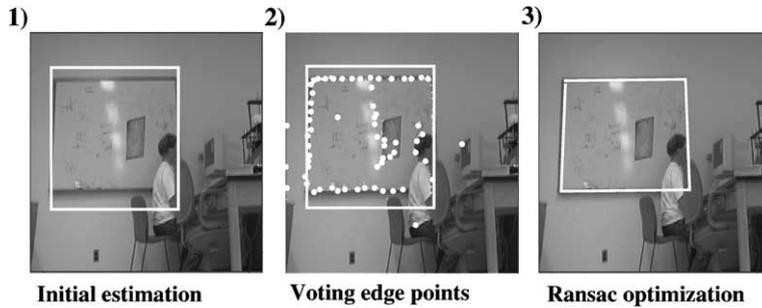


Fig. 5. Landmarks accurate extraction.

relaxation finally gives a list of the most plausible potential landmarks.

4.2. Landmarks accurate extraction

Fig. 5 describes the accurate extraction: the (V_k, V_l) segments vertices define four estimated edges for the so-called landmark. Each of these is regularly sampled and the procedure looks for edge points along the normal direction to the segment by correlation with a step-like signal.

As this 1D signal may be very noisy (slight occlusions for instance), a RANSAC procedure is used as a voting scheme among all the valid correlation maxima we found before so that we do not take into account the too noisy points for the computation of the straight line parameters. In our implementation, 51% of the valid votes are necessary for the selection of given line parameters. Fig. 5 shows an example of extraction where occlusion occurs and is overcome.

4.3. Landmark representation

At each newly discovered node, and then at each recognized one, landmarks are learned or recognized/updated according to the following principles.

In order to handle the perspective distortion problem, the detected set of segments, a quadrangle, is first rectified: we compute an homography H between this quadrangle and a square with a given size (75×75 for instance), as illustrated in Figs. 6 and 7, the “icon” I_p of landmark L_p from view p . H is a 3×3 matrix with eight degrees of freedom. The computation of H is straightforward from the correspondences between the icon vertices and the quadrangle ones, that can be written in a linear system.

We apply two kinds of saliency tests on L_p :

$$(\sigma(I_n) > \sigma_s) \wedge (nh(I_p) > n_s).$$

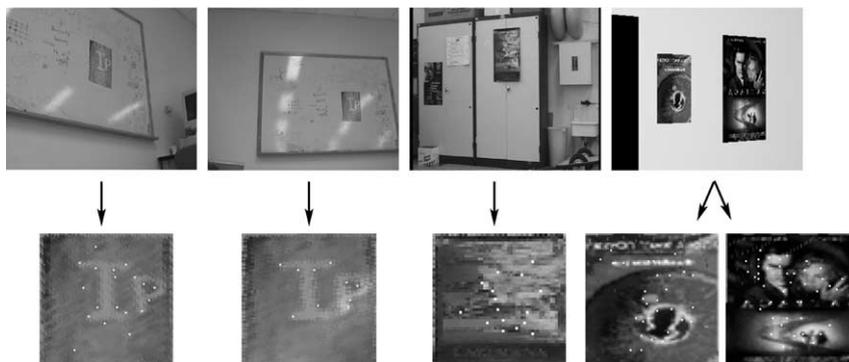


Fig. 6. Icon representations for landmarks.

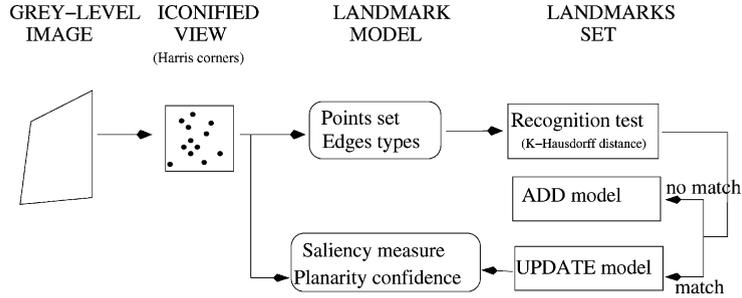


Fig. 7. Recognition and modeling.

σ is a global covariance on the icon I_p ; $nh(I_p)$ of representative Harris corners we can extract in I_p . σ_s and n_s are two thresholds. The set of detected corners will constitute the landmark model. We will denote it by P_p^k for landmark L_p .

Let us consider another landmark L_r to be identified. The similarity between the sets of m_p and m_r corners P_p^k and P_r^l , $1 \leq k \leq m_p$, $1 \leq l \leq m_r$, resulting from L_p and L_r , is evaluated by a Hausdorff partial distance d defined according to:

$$d_1(p, r) = K^{\text{th}}(\min_l \|P_p^k, P_r^l\|) \quad \text{and} \\ d(p, r) = \max(d_1(p, r), d_1(r, p)),$$

where K is a given fraction of $\min(m_p, m_r)$. $d(p, r)$ inferior to a given threshold implies that L_p is identified to L_r . In such a case, the L_r model is updated.

d cannot be used as a planarity indicator: for a given landmark L_r , low values on d may be reached with small viewpoint differences. Our approach to validate planarity consists in testing at the end of the learning phase whether a given candidate landmark has been seen from sufficiently different viewpoints. To quantify it we define the *planarity confidence measure* over all the views p_1, p_2, \dots, p_m corresponding to the landmark L_r by $\text{pcm}(r)$:

$$\text{vcm}(p_1, p_2) = \|\hat{H}_{p_1 p_2} - I\| \quad \text{and} \\ \text{pcm}(r) = \max_{k,l} \text{vcm}(k, l),$$

where $\text{vcm}(p_1, p_2)$ is a *viewpoint change measure*, $\hat{H}^{p_1 p_2}$ is the homography between the two quadrangles in views p_1 and p_2 such that $\hat{H}_{33}^{p_1 p_2} = 1$.

A threshold on the pcm value rejects landmarks L_r on which planarity and stability information are not sufficient.

5. Experimental results

For the preliminary experiment in the Beckman Institute, as illustrated in Fig. 8, only sonars were used. Without vision, we have adopted a global localization technique based on odometer readings; we needed to minimize the number of wheel spins, and even with such a limitation, the node recognition procedure failed very often.

In Fig. 9, experimental results for a complex building environment are presented. The environment is not a regular corridor network; only a partial exploration has been done (see the map in Fig. 9(a): a long corridor with two posters, three doors, a corner and two entrances in a hallway). The robot is a Nomadic XR4000, equipped with a SICK laser range finder, two belts of ultrasonic sensors and a stereo rig mounted on a pan and tilt platform; this robot can be considered as an holonomic robot. Images are acquired only from one static camera.

Although only sonars are used for the GVG incremental construction, we display in Fig. 9(b) both sonar data (points) and laser segments with the robot trajectory. The robot finds two nodes in front of the two large, complex entrances; two posters are found and associated to these nodes. For the moment, the vision module is not activated along the GVG arc, so that the doors are not discovered. Fig. 9(b) shows clearly how noisy are the odometer measurements on the XR4000 robot, so our method relies on them only locally.

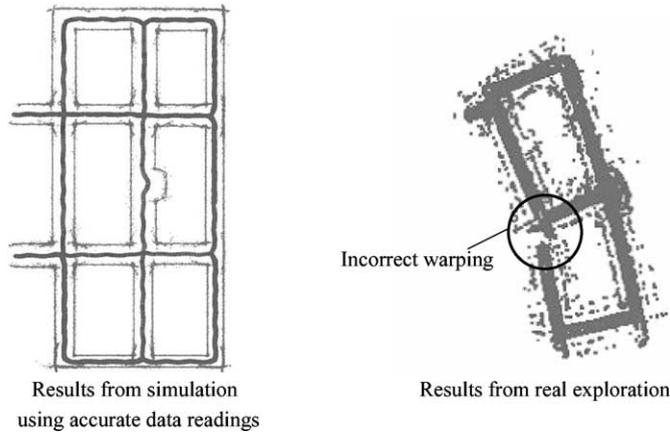


Fig. 8. Experiments in Beckman Institute.

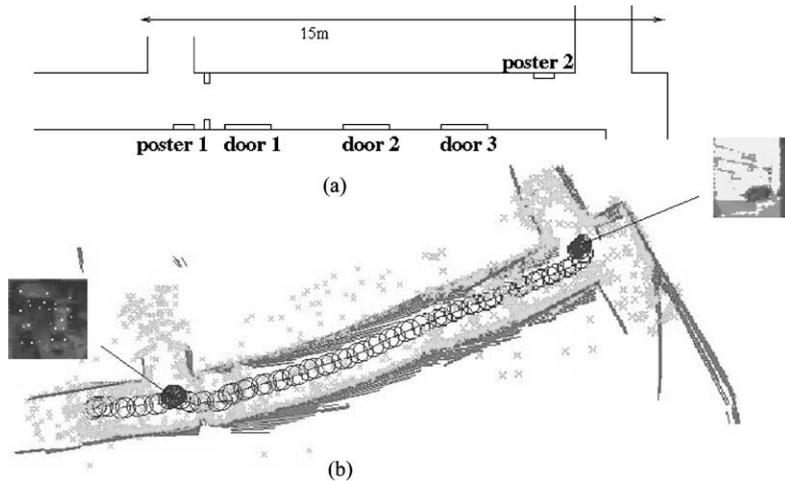


Fig. 9. Experiments: (a) the corridor map; (b) trajectory and sensory data.

6. Discussion and future work

This paper has presented the integration of several topological based representations required for the navigation of a mobile robot in an office environment. It takes advantage both from the Generalized Voronoi Graph model, suitable to represent a network of corridors, and from a landmark-based topological map which has been proposed to get rid of the classical problems which occur with an explicit self-localization with respect to an absolute reference frame. We avoid the use of traditional artificial visual landmarks by using, when available, some salient quadrangles in the scene.

In the current experiment, the environment is more complex than simple sets of orthogonal corridors, so that vision is mandatory to guarantee a good recognition of the nodes. We are currently trying to get some more significant experimental results for this complex environment; however, when dealing with a mixture of corridors, rooms and open spaces, the geometrical characterization of the meet points using sonars is difficult and visual landmarks must be considered to support local reference frames.

In our future work, we intend to improve the exploration task by the use of a Laser Range Finder instead of the ultrasonic sensors, and to add an uncertainty

representation to our model using Hidden Markov Models [3].

References

- [1] H. Bulata, M. Devy, Incremental construction of a landmark-based and topological model of indoor environments by a mobile robot, in: Proceedings of the 1996 IEEE International Conference on Robotics and Automation (ICRA'96), Minneapolis, MN, USA, April 1996.
- [2] H. Choset, K. Nagatani, A. Rizzi, Sensor based planning: using a honing strategy and local map method to implement the Generalized Voronoi Graph, in: Proceedings of the SPIE Technical Conference on Mobile Robots, Pittsburgh, PA, 1997.
- [3] S. Koenig, R. Simmons, Unsupervised learning of probabilistic models for robot navigation, in: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'96), Minneapolis, MN, USA, 1996, pp. 2301–2308.
- [4] D. Kortenkamp, T. Weymouth, Topological mapping for mobile robots using a combination of sonar and vision sensing, in: Proceedings of the AAAI, Seattle, WA, 1994, pp. 979–984.
- [5] B. Kuipers, The spatial semantic hierarchy, *Artificial Intelligence* 119 (2000) 191–233.
- [6] M. Devy, C. Parra, Topological modeling for outdoor mobile robots, in: Proceedings of the Seventh International Symposium on Intelligent Robotic Systems (SIRS'99), Coimbra, Portugal, July 1999.
- [7] K. Nagatani, H. Choset, S. Thrun, Towards exact localization without explicit localization with the Generalized Voronoi Graph, in: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'98), Louvain, Belgium, 1998.
- [8] J. Santos-Victor, R. Vassallo, H.J. Schneebeli, Topological maps for visual navigation, in: Proceedings of the First International Conference on Computer Vision Systems, Las Palmas, Spain, January 1999.
- [9] C. Taylor, D. Kriegman, Vision-based motion planning and exploration algorithms for mobile robots, in: K.Y. Goldberg, D. Halperin, J.-C. Latombe, R.H. Wilson (Eds.), *Algorithmic Foundations of Robotics*, A.K. Peters, Wellesley, MA, 1995.
- [10] D. Van Zwynsvoorde, T. Simeon, R. Alami, Incremental topological modeling using local Voronoi-like graphs, in: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems IROS'2000, Takamatsu, Japan, 2000.

Prashanth Ranganathan was born in 1978 in India. He completed his high school at Sydney Technical High School, and then his Bachelors in Computer Engineering at University of Illinois at Urbana-Champaign. He did his undergraduate thesis in mobile robotics and navigation using sensory devices. He is currently with Microsoft TV as a software design engineer.



Jean-Bernard Hayet got his degree in Computer Science Engineering in 1999 from ENSTA (Paris, France). He also got a DEA degree in Artificial Intelligence and Pattern Recognition from Paris VI University and since 1999 he has been working at the LAAS-CNRS Robotics and Artificial Intelligence Group on a Ph.D. Thesis about the use of visual landmarks for mobile robots in indoor environments.



Michel Devy got his degree in Computer Science Engineering in 1976 from IMAG, in Grenoble (France). He received his Ph.D. in 1980 from LAAS-CNRS in Toulouse (France). Since 1980, he has participated to the Robotics and Artificial Intelligence Group of LAAS-CNRS; his research is devoted to the application of computer vision in Automation and Robotics. He has been involved in numerous national and international projects, about Manufacturing Applications, Mobile Robots for space exploration or for civil safety, 3D Vision for Intelligent Vehicles or Medical Applications. He is now Research Director at CNRS, head for the Perception Area in the Robotics Group of LAAS-CNRS and his main scientific topics concern Perception for Mobile Robots in natural or indoor environments.



Seth Hutchinson received his Ph.D. from Purdue University in 1988. He is an Associate Professor in the Department of Electrical and Computer Engineering at University of Illinois at Urbana-Champaign and a full-time Beckman Institute faculty member in the Artificial Intelligence Group. His fields of professional interest are robotics, computer vision, and artificial intelligence.



Frédéric Lerasle was born in France, in 1969. He received the Diplôme d'Ingénieur from the Centre Universitaires des Sciences et Techniques in 1992 and the Diplôme de Docteur-Ingénieur in 1997 from Blaise Pascal University, Clermont-Ferrand. He is currently a Maître de Conférences in Image Processing at the Paul Sabatier University in Toulouse. In 1997, he joined the Robotics and AI Group of the LAAS. His main fields of research include indoor robot navigation using perception and human-robot interaction.