Mobile Robot Simultaneous Localization and Mapping Using Low Cost Vision Sensors

Vivek A. Sujan[1], Marco A. Meggiolaro[2], and Felipe A.W. Belo[2]

[1]Cummins Engine Company, Columbus, IN 47201, USA
[2]Pontifical Catholic University of Rio de Janeiro, 22453-900, Brazil

Summary. In this work, an information-based iterative algorithm is proposed to plan a mobile robot's visual exploration strategy, enabling it to most efficiently build a graph model of its environment. The algorithm is based on determining the information present in sub-regions of a 2-D panoramic image of the environment from the robot's current location using a single camera fixed on the mobile robot. Using a metric based on Shannon's information theory, the algorithm determines potential locations of nodes from which to further image the environment. Using a feature tracking process, the algorithm helps navigate the robot to each new node, where the imaging process is repeated. A Mellin transform and tracking process is used to guide the robot back to a previous node. The set of nodes and the images taken at each node are combined into a graph to model the environment. By tracing its path from node to node, a service robot can navigate around its environment. Experimental results show the effectiveness of this algorithm.

1 Introduction

In the past decade, mobile service robots have been introduced into various non-industrial application areas such as entertainment, building services, and hospitals. The market for medical robots, underwater robots, surveillance robots, demolition robots, cleaning robots and many other types of robots for carrying out a multitude of services has grown significantly (Thrun 2003). The algorithm complexity of personal and service robots has grown as a result from increased computational performance (Khatib 1999). This growth in algorithm complexity has been in conjunction with growth in hardware costs, a discouraging factor when aiming for large markets. Although hardware costs have declined with respect to their sophistication, this economic trend will still require the replacement of complex hardware architectures by more intelligent and cost-effective systems. Of particular interest here are the environment sensing abilities of the robot, thus algorithms must be developed to facilitate this behavior.

Mobile robot environment mapping falls into the category of Simultaneous Localization and Mapping (SLAM). In SLAM, a robot localizes itself as it maps the environment. To achieve the localization function, landmarks and their relative motions are monitored with respect to the vision systems. Although novel natural landmark selection methods have been proposed (Simhon et al. 1998), most SLAM architectures rely on identifying distinct, recognizable landmarks such as corners or edges in the environment (Taylor et al. 1998). This often limits the algorithms to well-structured environments, with poor performance in highly textured environments.

These algorithms have been implemented for several different sensing methods, such as stereo camera vision systems (Se et al. 2002), laser range sensors (Tomatis et al. 2001), and ultrasonic sensors (Anousaki et al. 1999). Sensing uncertainties have been investigated for single or multi-robot systems (Roumeliotis et al. 2004). However, such relatively expensive sensor suite is not suitable for many commercial applications for service robots in large scale, which would require simplifications in hardware resources, lower setup costs, maintenance costs, computational resources and interface resources. The basic sensor suite of a typical large-scale service robot should include only a single monocular camera fixed to the base, contact (bump) switches around its periphery, and an odometer sensor such as wheel encoders, subject to dead reckoning errors.

In this work, an unknown environment exploration and modeling algorithm is developed to accurately map and navigate a robot agent in a flat-floored environment, subject to the above hardware limitations - using only information from a single monocular camera, wheel encoders and contact switches.

2 Technical Approach

The environment exploration and modeling algorithm proposed here consists of 3 primary components. The overall process is shown in Fig. 1. The mobile robotic agent models the environment as a collection of nodes on a graph. The first component of the algorithm is to identify potential child nodes from a given location, see Fig. 2. At each node the robot conducts a panoramic scan of the environment. This scan is done as a series of 2-D image snapshots using in-place rotations of the base by known angles. Next, an information theoretic metric is used to divide the panoramic image into regions of interest. Using a metric based on Shannon's information theory (Reza 1994), the algorithm determines potential locations of nodes from which to further image the environment. If any region of interest contains sufficient information, then it is identified as a potential child node, which would then need to be explored.

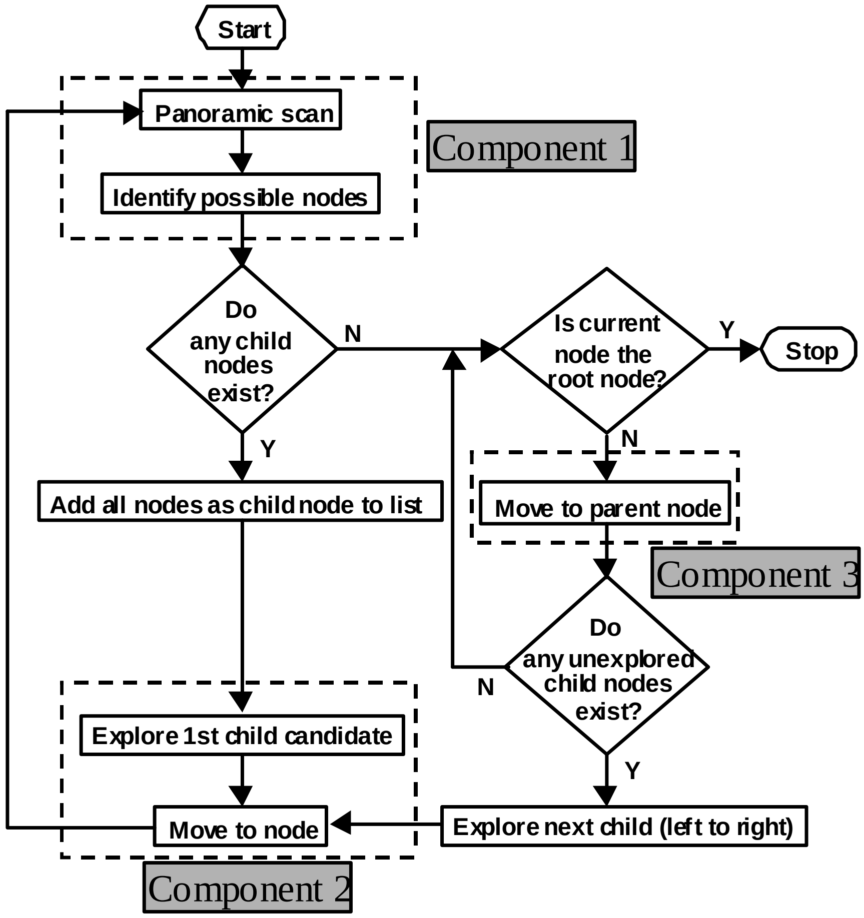

Fig. 1. Algorithm overview

After the list of child nodes is collated, each one is explored sequentially, traversing the graph in a depth-first manner. The second component of the algorithm is to traverse to a child node from the current node. This is achieved by tracking the target node using a simple region growing tracker and a visual servo controller. Heading accuracy is improved using an Extended Kalman Filter. If the node cannot be reached by a straight line due to an obstruction, then the point of obstruction is defined as the new child node. The process of panoramic image development and child node identification and exploration continues.

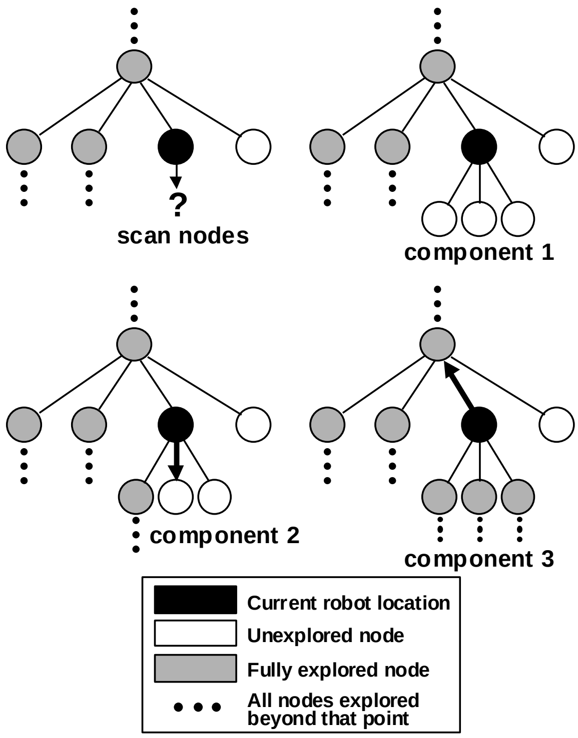

Fig. 2. Key components of the environment exploration and modeling algorithm

The third component of the algorithm is to traverse to a parent node from the current one. To return to a parent node, the robot must first identify the direction of such node. This is done using a Mellin transform (Ruanaidh et al. 1997) to determine if the image that the robot currently sees is what it would expect to see if it were pointed in the correct direction toward the parent node. The expected image is derived from the panoramic scan previously taken at the parent node. Once this direction is established, the robot moves toward the parent node. Visual servo control, based on the correlation between the current image and the one it would see if it were at the parent node pointed in the same direction, governs if the robot has reached the parent node.

An alternative approach to find the parent node direction is done using Scale-Invariant Feature Transforms, SIFT (Se et al. 2002). The features are invariant to image translation, scaling, rotation, and partially invariant to illumination changes and affine or 3D projection. SIFT features from the current and expected views are selected to verify whether the mobile agent has reached the previously explored parent node.

3 Experimental Results

The proposed algorithm is applied to the exploration of an apartment by a single mobile robot agent, adapted from an ER-1 commercial system (Evolution Robotics 2006), see Fig. 3. The system consists of a two-wheeled differential system driven by step motors, equipped with wheel encoders (for odometry), three bump switches for measuring contact, and a single monocular camera mounted to its base. Three infrared sensors are used just as on-off bump switches for measuring contact, without providing any range data. The robot is controlled by a 1.5GHz Pentium IV notebook mounted on its structure.

Fig. 3. Mobile robot experimental system and resulting environment map with node locations

Figure 4 shows two steps in guiding the mobile robot to an unexplored child node, validating the second component of the algorithm. The node is tracked using a simple region growing method. In addition to the primary target (+), two redundant fiducials are selected and tracked (O). Since a single camera cannot provide 3-D data, these two fiducials are automatically selected based on their apparent closeness to the primary target in a 2-D sense. This permits fewer re-identifications of fiducials.

Fig. 4. Robot view tracking to child node (+) with the aid of fiducials (o)

Figure 5 shows an example of a raw image taken by the robot to search for child node candidates, the first component of the algorithm. This image is trimmed and simplified using the proposed information-based quadtree decomposition process. Several key areas in the image have been determined to have high quantities of information present. These areas are further processed to determine the coordinates of the child nodes, see Fig. 5. The identification process has selected nodes that are both useful - such as the ones near the doorway - but has also picked up nodes that may not be very useful - such as the one on the carpet. These latter nodes are often eliminated with low pass filtering in image pre-processing steps.

The process to return to a previously explored node can be seen in Fig. 6. The panoramic image previously taken at the target node is cropped to show the expected view when the robot reaches the target. SIFT features are then automatically chosen and used to correlate the current robot view to the expected one, successfully guiding the robot without the need of stereo vision. A different method, based on invariant Mellin transforms, is also investigated experimentally.

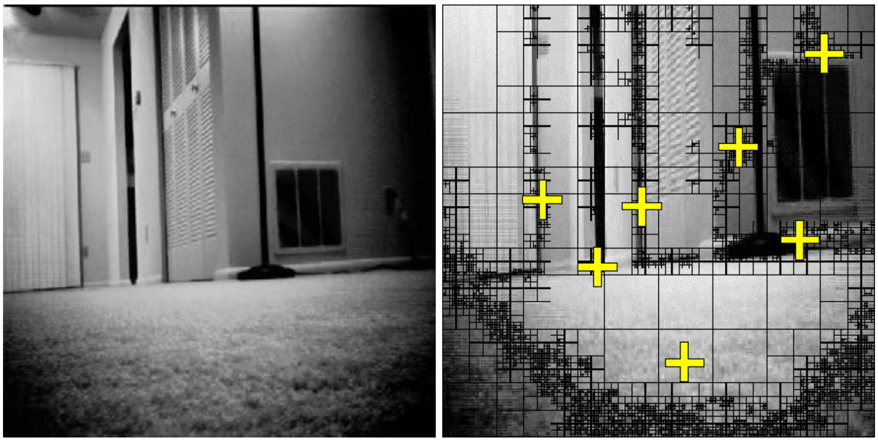

Fig. 5. Raw and processed images taken by the onboard camera, showing information-based quadtree decomposition and identified child node candidates (+)

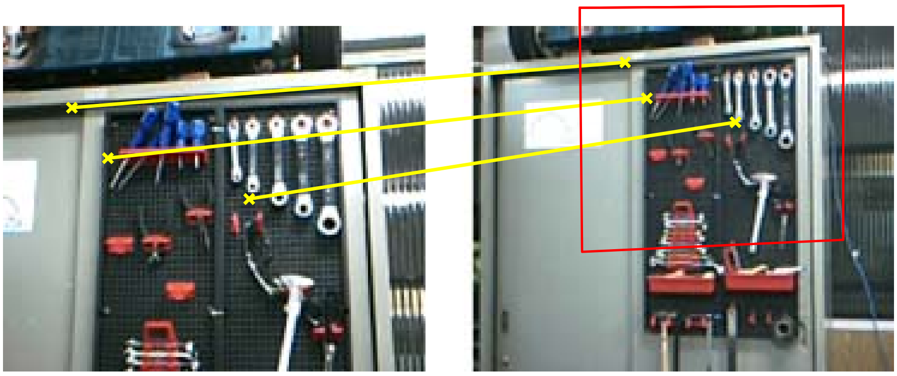

Fig. 6. SIFT features used to reach a parent node, correlating the current view (right) from a child node to the expected view when reaching the parent node (left)

Figure 7 shows the resulting graph obtained from the experimental exploration of the flat-floored apartment. Each node is marked and linked to its parent/child nodes. This graph model is essentially the causal map described by Kuipers (2000), where the panoramic images correspond to views, navigation methods correspond to actions, and nodes correspond to distinctive states. Finally, by tracing its path from node to node, a service robot can then continue to navigate from one node to another through the environment. This map may then be used for navigation by the robot within its environment. Note that the walls and furniture were later added to the figure, since the absence of range sensors or stereo vision prevents the robot from identifying their exact location. However, the robot is able to recognize its environment, including walls and furniture, from the set of panoramic images taken at each landmark.

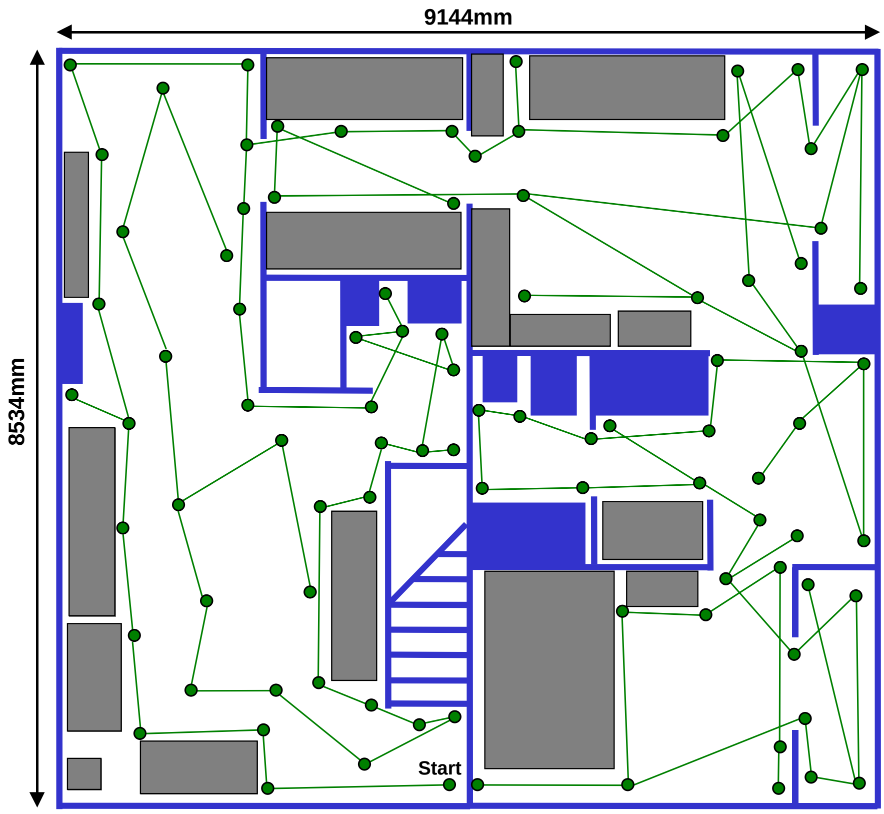

Fig. 7. Mobile robot experimental system and resulting environment map with node locations.

From the performed experiments it is found that the robot is able to successfully map the environment as well as localize itself, arriving at any specified node. Note that the error in reaching a desired node is not a function of how far down the graph the node is, because the Mellin transform or the SIFT features only need to consider the image that the target node has in memory. It has been found that the positioning error of the robot is directly proportional to its average distance to the walls and obstacles in its field of view, and inversely proportional to the camera resolution. For several test runs on our experimental system in indoor environments such as the one in Fig. 7, with distance from walls ranging in average between one and ten meters, and with a (limited) camera resolution of 176 × 144 pixels, an average positioning error of 60mm (RMS) has been obtained. This positioning accuracy can certainly be improved with a better camera resolution.

4 Conclusions

An information-based iterative algorithm has been presented to plan the visual exploration strategy of an autonomous mobile robot using a single base-mounted camera. Experimental studies have been conducted to demonstrate the effectiveness of the entire algorithm. It was found that this algorithm allows a mobile robot to efficiently localize itself using a limited sensor suite, consisting of a single monocular camera, contact sensors, and an odometer, reduced memory requirements - only enough to store one 2-D panoramic image at each node of a graph - as well as modest processing capabilities. Therefore, the presented approach has a potential benefit to significantly reduce the cost of autonomous mobile systems such as indoor personal and service robots.

References

Anousaki GC, Kyriakopoulos KJ (1999) Simultaneous localization and map building for mobile robot navigation. IEEE Robotics & Automation Magazine 6(3):42-53

Evolution Robotics, 2006. Webpage www.evolution.com

Gelb A (1974) Applied optimal estimation. MIT press, Cambridge, MA

Khatib O (1999) Mobile Manipulation: The Robotic Assistant. J. of Robotics and Autonomous Systems 26:175-183

Kuipers B (2000) The Spatial Semantic Hierarchy. Artificial Intelligence 119:191-233

Reza FM (1994) An introduction to information theory. Dover, New York

Roumeliotis SI, Rekleitis IM (2004) Propagation Uncertainty in Cooperative Multirobot Localization: Analysis and Experimental Results. Autonomous Robots 17:41-54

Ruanaidh JO, Pun T (1997) Rotation, scale and translation invariant digital image watermarking. IEEE International Conference on Image Processing (ICIP 97), Santa Barbara, CA, pp.536-539

Se S, Lowe DG, Little J (2002) Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. International Journal of Robotics Research 21(8):735-758

Simhon S, Dudek G (1998) Selecting targets for local reference frames. Proceedings of the IEEE International Conference on Robotics and Automation, pp.2840-2845

Taylor CJ, Kriegman D (1998) Vision-based motion planning and exploration algorithms for mobile robots. IEEE Trans. on Robotics and Automation 14(3):417-426

Thrun S (2003) Robotic mapping: a survey. In: Exploring artificial intelligence in the new millennium, Morgan Kaufmann Publishers Inc., San Francisco, CA, pp.1-35

Tomatis N, Nourbakhsh I, Siegwar R (2001) Simultaneous localization and map building: a global topological model with local metric maps. Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.421-426