

Simultaneous Localization and Mapping for Mobile Robot Teams with Visual Sensors

Vivek Anand Sujan[#], Marco Antonio Meggiolaro^{*}

[#]Advanced Control Division, Cummins Engine Company, Columbus, IN 47201

^{*}Department of Mechanical Engineering, Pontifical Catholic University of Rio de Janeiro Rio de Janeiro, RJ - Brazil 22451-900

Abstract: In field or indoor environments it is often not possible to provide robots or robotic teams with detailed a priori environment and task models. In such environments, robots will need to create a dimensionally accurate geometric model by moving around and scanning the surroundings with their sensors. In the case of robotic teams, there is a further need of cooperatively sharing the acquired data. However, uncertainties in robot locations and sensing limitations/occlusions make this difficult. A novel information-based methodology based on iterative sensor planning and sensor redundancy is presented to build a geometrically consistent dimensional map of the environment and task. The proposed algorithm efficiently repositions the systems' sensing agents using an information theoretic approach and fuses sensory information using physical models to yield a geometrically consistent environment map. This is achieved by utilizing a metric derived from Shannon's information theory to plan the robots' visual exploration strategy, determining optimal sensing poses for the agent(s) mapping a highly unstructured environment. This map is then distributed among the agents (if robotic teams are considered) using an information-based relevant data reduction scheme. This methodology is unique in the application of information theory to enhance the performance of cooperative sensing robot teams. It may be used by multiple distributed and decentralized sensing agents for efficient and accurate environment modeling. The algorithm makes no assumptions of the environment structure. Hence it is robust to robot failure since the environment model being built is not dependent on any single agent frame. It accounts for sensing uncertainty, robot motion uncertainty, environment model uncertainty and other critical parameters, allowing for regions of higher interest getting more attention by the agent(s). The methodology works with mobile robots (or vehicles) with eye-in-hand vision sensors to provide 3-D or 2.5-D information of the environment. The presented methodologies are particularly well suited to unstructured environments, where sensor uncertainty is significant. Simulation and experimental results show the effectiveness of this approach. A cooperative multi-agent sensing architecture is presented and applied to the mapping of a cliff surface using the JPL Sample Return Rover (SRR). The information-based methods are shown to significantly improve mapping efficiency over conventional ones, with the potential benefit to reduce the cost of autonomous mobile systems.

Keywords: Visual mapping, Cooperative Robots, Information Theory, Unstructured Environments, Data Fusion.

1. INTRODUCTION

¹An important goal of robotics research is to develop mobile robot teams that can work cooperatively in unstructured field environments [Baumgartner *et al.* 1998, Huntsberger *et al.* 2000]. Potential tasks include explosive ordinance removal, de-mining and hazardous waste handling, exploration/development of

space, environment restoration, and construction [Baumgartner *et al.* 1998, Huntsberger *et al.* 2001]. For instance, space and planetary robotic missions will require robot scouts to lead the way, by exploring, mapping, seeking or extracting soil and rock samples and eventually constructing facilities in complex terrains. Multiple cooperating robots will be required to set up surface facilities in challenging terrain for in-situ measurements, communications, and to pave the way for human exploration of planetary surfaces. This will require the handling of relatively large objects, such as deploying of solar panels and

* Address correspondence to this author at the Department of Mechanical Engineering, Pontifical Catholic University of Rio de Janeiro Rio de Janeiro, RJ - Brazil 22451-900.
E-mail: meggi@mec.puc-rio.br

sensor arrays, anchoring of deployed structures, movement of rocks, and clearing of terrain. Figure 1 shows an example of a representative system for such task.

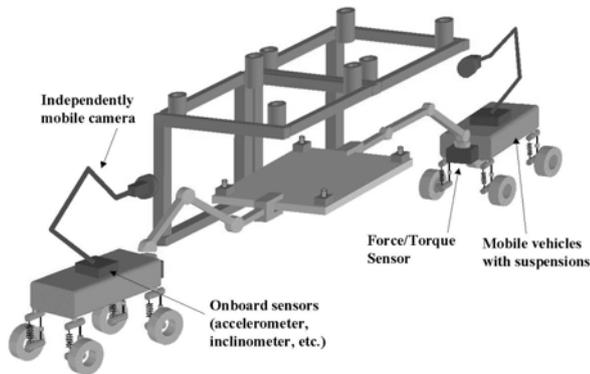


Figure 1: Representative physical system.

The use of robot teams working cooperatively to acquire and share data has been proposed to succeed in those missions [Huntsberger *et al.* 2001, Pirjanian *et al.* 2001, Schenker *et al.* 2001, Sujan *et al.* 2002, Trebi-Ollennu *et al.* 2002]. The control of such systems typically requires models of the environment and task. In unstructured field environments it is often not possible to have such a-priori models. In such cases, the robot needs to construct these from sensory information, usually from vision systems. A number of problems can make this non-trivial. These include the uncertainty of the task in the environment, location and orientation uncertainty in the individual robots, and occlusions, e.g. due to obstacles, work piece, or other robots. If the systems are equipped with vision sensors mounted at a manipulator end-effector (eye-in-hand systems), intelligent planning of the sensor motion can alleviate problems of the occlusions, providing an accurate geometrical model of the task and environment. If the system consists of more than one robot, planning the behavior of these multi-information sharing systems can further improve the system performance. In addition, to date planetary robots missions have been limited to moving over rather benign terrain [Schenker 1998]. These systems are not capable of exploring highly irregular terrain such as cliff surfaces that are potentially geologically rich and hence very interesting for planetary science

[Baumgartner 1998, Huntsberger 2000]. To succeed, robot teams working cooperatively to acquire and share data have been proposed [Pirjanian 2001, Schenker 2001, Sujan 2002, Trebi-Ollennu 2002, Huntsberger 2001].

Another important application of autonomous mobile systems is personal service robots. As the autonomy of personal service robotic systems increases so has their need to interact with their environment. For many applications it is not usually practical to provide robots in advance with valid geometric models of their environment. The robot will need to create these models by moving around and sensing the environment, while minimizing the complexity of the required sensing hardware.

Significant advances have been made in the area of personal mobile robots from the mid 90's to today. More notable but not limited to are [Austin *et al.*, 2002; Bischoff, 2000; Borenstein *et al.*, 1990; Burgard *et al.*, 1998; Colgate *et al.*, 1996; Dario *et al.*, 1997; Dubowsky *et al.*, 2000; Engelberger, 1989; Evans, 1994; Glüer *et al.*, 2000; Graf *et al.*, 2000; Haegele *et al.*, 2001; Han *et al.*, 2002; Jung *et al.*, 1997; Kawamura *et al.*, 1996; Khatib, 1999; Lawitzky, 2000; Marrone *et al.*, 2001; Pettinaro *et al.*, 2002; Schaeffer *et al.*, 1999; Thrun, 1998; Thrun *et al.*, 1999; Ulrich *et al.*, 1997; Wellman *et al.*, 1994; Wong *et al.*, 2000]. In recent years, mobile service robots have been introduced into various non-industrial application areas such as entertainment, building services, and hospitals. They are relieving humans of tedious work with the prospect of 24-hour availability, fast task execution, and cost-effectiveness. 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. In 2001 a total of at least 750,000 robotic units were in use worldwide, of which 389,000 robot units were in Japan, 198,000 in the European Union, 90,000 in North America and 4,500 in South America. In Europe, Germany led with 91,000 robot units, followed by Italy with 39,000, France with 21,000 and the United Kingdom with 12,000. In South America, Brazil led with approximately 90% of the installed units. It is projected that in 2004 some 975,000 systems worldwide will be in use, of which 447,000 will

be in Japan, 306,000 in the European Union and 116,000 in North America [World Robotics, 2001]. Although most of these numbers relate to industrial manipulators, the sales of mobile robots are projected to exceed the sales of factory floor robots by a factor of four, exceeding US\$2 billion [Lavery, 1996]. And unlike the factory floor robot market, the sources for the vast majority of these machines could be U.S. companies.

Service robots for personal and private use are mainly found in the areas of domestic (household) robots, which include vacuum cleaning and lawn-mowing robots, and entertainment robots, including toy and hobby robots. If the technology for personal service robots provides what it has promised, at a competitive price, and if there is a sufficient degree of consumer acceptance, then this can be a very large market.

Due to increased computational performance, algorithm complexity has grown thus providing increased system capability [Austin *et al.*, 2002; Bischoff, 2000; Borenstein *et al.*, 1990; Glüer *et al.*, 2000; Khatib, 1999; Lawitzky, 2000; Wellman *et al.*, 1994; Wong *et al.*, 2000]. This growth in algorithm complexity has been in conjunction with growth in hardware complexity [Bischoff, 2000; Dario *et al.*, 1997; Dubowsky *et al.*, 2000; Kawamura *et al.*, 1996; Schaeffer *et al.*, 1999; Ulrich *et al.*, 1997]. However, the high costs associated with hardware complexity are a discouraging factor. This economic drive has been seen in the last decade, where the performance of industrial and personal robots has radically increased while prices have fallen. A robot sold in 2000 would have cost less than a fifth of what a robot with the same performance would have cost in 1990 [World Robotics, 2001]. 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. As the autonomy of these service robotic systems increases, so has their need to interact with their environment. The most basic interaction a robotic agent may have with its

environment is to sense and navigate through it. Thus algorithms must be developed to facilitate this behavior.

Environment mapping by mobile robots falls into the category of Simultaneous Localization and Mapping (SLAM). In SLAM a robot is localizing itself as it maps the environment. Researchers have addressed this problem for well-structured (indoor) environments and have obtained important results [Anousaki *et al.* 1999, Burschka *et al.* 1997, Castellanos *et al.* 1998, Choset *et al.* 2001, Kruse *et al.* 1996, Thrun *et al.* 2000, Tomatis *et al.* 2001, Victorino *et al.* 2000, Yamauchi *et al.* 1998]. These algorithms have been implemented for several different sensing methods, such as camera vision systems [Castellanos *et al.* 1998, Hager *et al.* 1997, Park *et al.* 1999], laser range sensors [Tomatis *et al.* 2001, Yamauchi *et al.* 1998], and ultrasonic sensors [Anousaki *et al.* 1999, Choset *et al.* 2001]. Sensor movement/placement is usually done sequentially (raster scan type approach), by following topological graphs or using a variety of greedy algorithms that explore regions only on the extreme edges of the known environment [Anousaki *et al.* 1999, Choset *et al.* 2001, Rekleitis *et al.* 2000, Victorino *et al.* 2000, Yamauchi *et al.* 1998]. Geometric descriptions of the environment are modeled in several ways, including generalized cones, graph models and Voronoi diagrams, occupancy grid models, segment models, vertex models, convex polygon models [Choset *et al.* 2001]. The focus of these works is accurate mapping. They do not address mapping efficiency. Researchers have addressed mapping efficiency to a limited amount [Kruse *et al.* 1996]. However, sensing and motion uncertainties are not accounted for. They also generally assume that the environment is effectively flat (e.g. the floor of an office or a corridor) and readily traversable (i.e. obstacles always have a route around them) [Anousaki *et al.* 1999, Choset *et al.* 2001, Thrun *et al.* 2000, Yamauchi *et al.* 1998] and have not been applied to robot teams working in rough planetary environments. Also, prior work has not addressed optimizing the communication between agents for both multi-agent planning and cooperative map-building.

To achieve the localization function, landmarks and their relative motions are monitored with respect to the vision systems. Several localization schemes have been implemented, including topological methods such as generalized Voronoi graphs and global topological maps [Choset *et al.* 2001, Tomatis *et al.* 2001, Victorino *et al.* 2000], extended Kalman filters [Anousaki *et al.* 1999, Park *et al.* 1999], and robust averages [Park *et al.* 1999]. Although novel natural landmark selection methods have been proposed [Hager *et al.* 1997, Simhon *et al.* 1998, Yeh *et al.* 1995], most SLAM architectures rely on identifying landmarks as corners or edges in the environment [Anousaki *et al.* 1999, Castellanos *et al.* 1998, Choset *et al.* 2001, Victorino *et al.* 2000]. This often limits the algorithms to structured indoor-type environments. Others have used human intervention to identify landmarks [Thrun *et al.* 2000].

Some studies have considered cooperative robot mapping of the environment [Jennings *et al.* 1999, Rekleitis *et al.* 2000, Thrun *et al.* 2000]. Novel methods of establishing/identifying landmarks and dealing with cyclic environments have been introduced for indoor environments [Jennings *et al.* 1999, Thrun *et al.* 2000]. In some cases, observing robot team members as references to develop accurate maps is required [Rekleitis *et al.* 2000]. While the work done in this field has had significant impact on robot control architectures, these results largely do not address the problem of cooperative sensing in the context of mobile robots in unknown, unstructured environments. The methods developed to date generally rely on assumptions that include: simple well-known terrains; accurate knowledge of the environment; little or no task uncertainty; sufficient sensing capability and sufficient communication capabilities. For real field environments, these assumptions are often not valid. In general, current research has not solved the problem of controlling multiple mobile robots performing cooperative tasks in unknown environments, where uncertainty, limited sensing capabilities, and incomplete physical models of the system(s)/environment dominate the problem.

This work presents a cooperative multi-agent algorithm for the visual exploration of an unknown environment. The basic approach to the algorithm is given in Figure 2 [Sujan *et al.* 2002, Sujan *et al.* 2003]. This algorithm fuses sensory information from one or multiple agents using physical sensor models, robot models, and environment maps to yield geometrically consistent surrogate information in lieu of missing data due to the environment, task, robot and sensor uncertainties. The multiple agents can have different shapes and kinematic constraints, as long as their models are known. The algorithm falls into the general category of SLAM. The mapping and localization process is as follows. First, each agent efficiently repositions its sensors using an information theoretic approach, in order to optimally fill in uncertain/unknown regions of the environment map, based on maximizing the expected new information obtained to yield a geometrically consistent environment map while minimizing the motions of the robots over the hazardous surfaces. Next, each agent fuses the data to its known environment model by localizing itself with respect to a global fixed reference frame. Finally, each agent shares its known environment map with its team, which is then integrated by the other agents into their own environment maps. The information is then used by the control and planning architecture to plan further movements of the sensors for each agent. A common environment map is built by fusing the data available from the individual robots, providing improved accuracy and knowledge of regions not visible by all robots. Thus, the experiences (measurements) of each robot can become a part of the collective experience of the multi-agent team. The algorithm is unique in using the quantity of information of the environment that it currently has, predicting high information-yielding viewpoints from which to continue exploring the environment. This results in a significantly more efficient exploration process. This algorithm is directly applicable to a team of robots as it is to a single explorer. This generality also extends to data representation formats (e.g. 2-D, 2.5-D and 3-D representations). This is achieved through a unique information representation, sharing and fusion architecture.

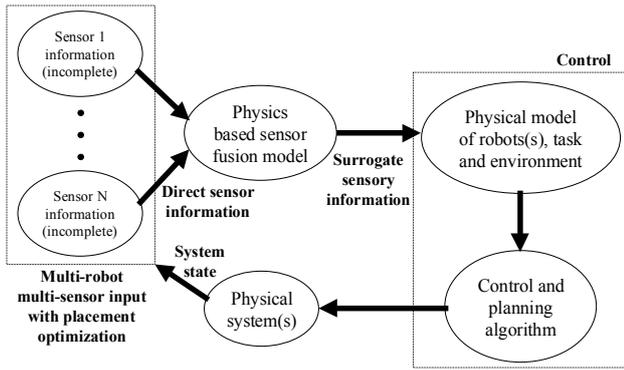


Figure 2: An architecture for multi-robot cooperative sensing.

The algorithm is applied in this study to a team of four robots to cooperatively explore a cliff surface. Figure 3 shows schematically four cooperative robots working in an unstructured field environment to lower one robot down a cliff face that is not accessible by a single robot alone. One robot (Cliff-bot) is lowered down a cliff face on tethers. Two robots (Anchorbots) act as anchor points for the tethers. A fourth robot, RECON-bot (REmote Cliff Observer and Navigator) provides mobile sensing. All the robots are equipped with a limited sensor suite, computational power and communication bandwidths. The Cliff-bot, usually the lightest system, may be equipped with primarily a science sensor suite, and limited sensors for navigation. The RECON-bot serves to observe the environment to be traversed by the Cliff-bot and communicates the data relevant for navigation to the Cliff-bot. The RECON-bot has an independently mobile camera and other onboard sensors to map and observe the environment. Rocks, outcroppings and other robots limit sensing and sensor placement, resulting in uncertainties and occlusions (see Figure 4). There is significant uncertainty in the robots' locations and poses with respect to the environment. Due to these limitations and uncertainties it is difficult or impossible for all robots to independently measure the environment to control the system, therefore the need of an algorithm that includes sensor fusion. In this application, the RECON-bot agent is the JPL Sample Return Rover (SRR), which optimally surveys the cliff surface and transmits the information to other agents. Experimental results compare the proposed architecture to

conventional raster (or sequential) sensing schemes.

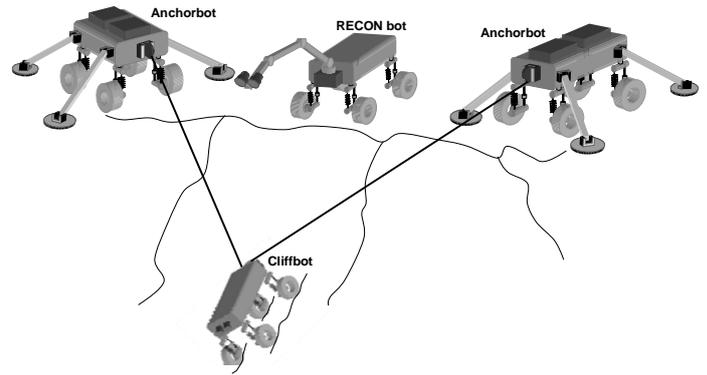


Figure 3: Schematic for a cooperative robot cliff descent.

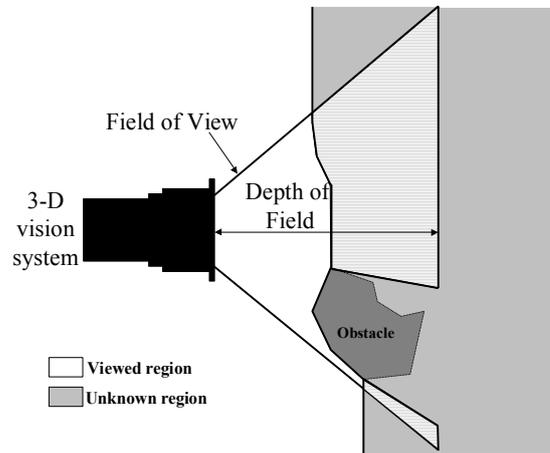


Figure 4: Sensing limitations due to occlusions.

The algorithm used in the cliff exploration task assumes 3-D environment sensing by stereo camera pairs [Sujan, 2002; Sujan *et al.*, 2002]. The onboard sensors include stereo vision systems mounted on independently mobile arms, active mobility of the base, inclinometers, and directional sensors. However, this method proved to be computationally expensive for most applications and onboard resources. Thus the algorithm is modified by reducing the sensed data set from 3-D to 2.5-D, i.e., from (x, y, z) data points to (x, y, h) where h is the elevation. Thus, instead of a true 3-D environment map, the algorithm develops an elevation map [Hunstberger *et al.*, 2003; Schenker *et al.*, 2003; Sujan *et al.*, 2004]. Applications for both mapping methods include rough terrain modeling for

planetary exploration, cooperative task execution, and others. The algorithms are described next.

2. DESCRIPTION OF THE 3-D AND 2.5-D MAPPING ALGORITHMS

The SLAM algorithm here proposed may be broken down as follows. First, the vision sensors cooperatively scan the region around a target, generating a local 3D geometric model. This allows the robots to locate themselves and the obstacles in the target reference frame. Next, these models are used to find optimum environment viewing poses for the multiple vision systems [Sujan *et al.* 2002]. The process is initialized by visually finding the target and robots in a common reference frame. Then, a new pose is found for each of the sensors by defining and optimizing a rating function (RF) over the possible sensor positions, subject to kinematic constraints of the sensor placement mechanisms for the individual robots. This rating function aims to acquire as much new information about the environment as possible with every sensing cycle, while maintaining or improving the map accuracy, and minimizing the exploration time. The process is constrained by selecting goal points that are not occluded and that can be reached by a collision-free traversable path.

The sensors then move to their new poses and acquire 3D data. Based on the sensor mount kinematics, the motion of the sensor is known. However, small motions of the robot base (due to suspension compliance) and errors in sensor mounts lead to additional uncertainties. These are accounted for by measuring common features during the vision sensor motion. Finally, the new data and its associated uncertainty are fused with the current environment map, resulting in an updated probabilistic environment map, which may then be shared with the other sensing agents. In general, for each sensing agent the algorithm consists of four steps, described as follows (see Figure 5).

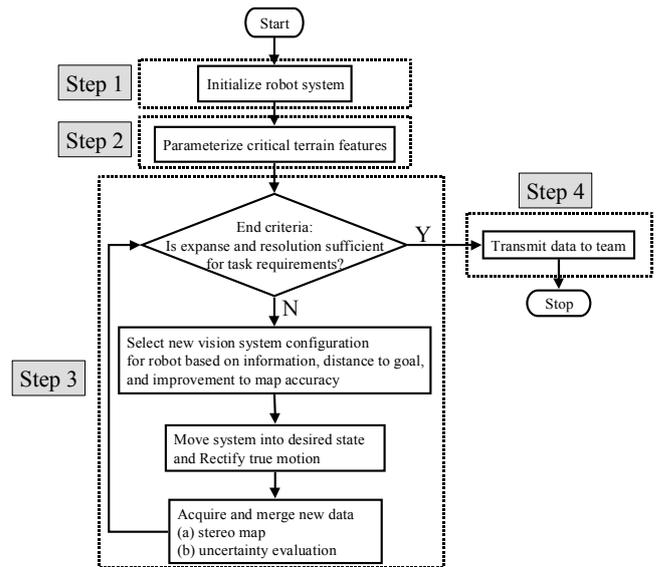


Figure 5: Outline of model building and placement algorithm.

2.1. Step 1: System Initialization

Here the environment map is initialized, the robots are localized, and a first map is generated. The environment may be mapped using several different representations. In this version of the algorithm, both 2.5-D elevation maps and 3-D occupancy grid maps are addressed. In 2.5-D elevation maps, the map is a plane of grid cells where each grid cell value represents the average elevation of the environment at that cell location. Uncertainty in elevation is also maintained with each grid cell. In 3-D occupancy grid maps, the map is modeled as a probabilistic discretized occupancy 3-D grid. Every voxel in the 3-D grid has a value for probability-of-occupancy that ranges from 0 (empty) to 1 (occupied). A value of 0.5 indicates maximum uncertainty in occupancy of the voxel. Each representation has its advantages, as follows. The 2.5-D elevation maps are particularly useful for terrain profiling where there are little or no terrain overhangs. On the other hand, the 3-D occupancy grid maps are more useful in areas that do have a large number of significant overhangs, such as caves, rooms, etc, since 2.5-D maps only provide one data point, the elevation, for each planar grid cell.

The map is built in a fixed reference frame defined by a well-known landmark measurable by all the sensing agents. All robots contributing to or requiring use of the map are localized with respect to the initial map. For the cliff exploration team, the RECON-bot contributes to and uses the

environment map, while the Cliff-bot only uses the environment map. Localization may be achieved by either:

- (a) absolute localization, achieved by mapping a common environment landmark that is visible by all robots; or
- (b) relative localization, done by mapping fiducials on all robots by other robot team members where one robot is selected as the origin.

Relative localization is used in the Cliff-bot application, with the RECON-bot localizing the Cliff-bot with respect to itself (the origin, see Figure 6). Then, each agent initially senses the environment. Absolute localization is also studied in this work, see Section 4.2. Other methods can be used for either absolute or relative localization [Yamauchi *et al.* 1998]. Searching for the target (which forms the absolute origin) by moving the robot sensors can be done in many ways, such as exhaustive raster scanning, random walking, tracking “space filling curves”, and model-based image understanding methods [Tarabanis *et al.* 1995]. In this study, sensor positioning for target searching is done in the same way as sensor positioning for environment model building (described in Step 3, Section 2.3). The absolute origin target is located by matching the known target element geometric CAD model with visual data [Lara *et al.* 1998]. At this stage, the environment model is considered empty, i.e. no points are known. The first stereo range map (including the common target and all objects within the field of view) is taken by each agent. It is assumed that only the geometry of the task elements - such as parts of a solar panel that is to be assembled [Huntsberger *et al.* 2001] - are well known. Obstacles and robot positions are unknown.

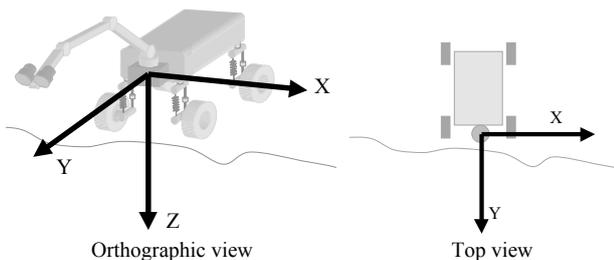
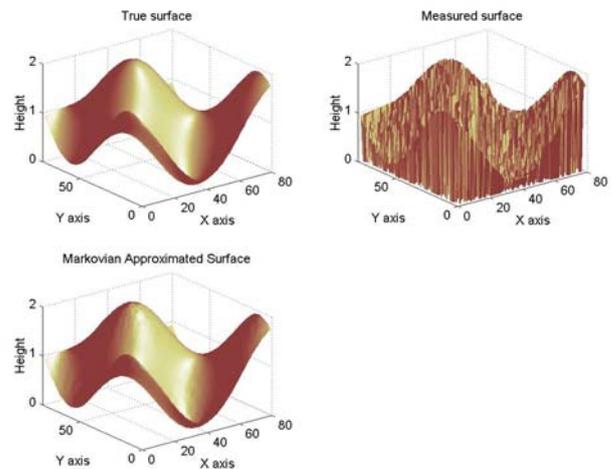


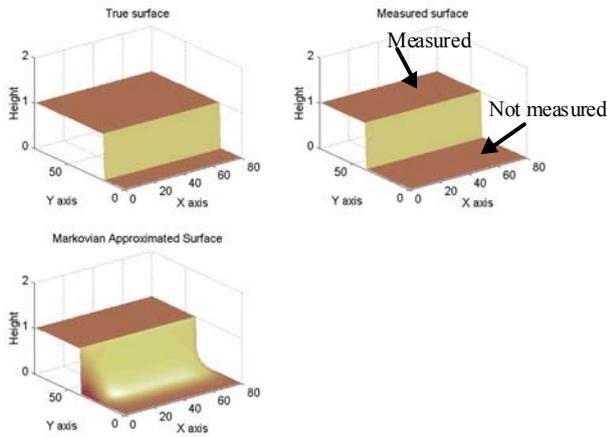
Figure 6: Initial environment map coordinate frame.

2.2. Step 2: Critical Terrain Feature Identification

In some applications, certain regions of the terrain may be critical, requiring early identification, mapping and monitoring as the environment model grows [Huntsberger *et al.* 2003, Schenker 2001, Sujan *et al.* 2003]. An example is determining regions of safe travel for the sensing agents. For this process, the incomplete environment model is temporarily completed by a Markovian approximation for unknown environment grid point values. In the cliff exploration application, identification of the cliff edge by the RECON-bot is critical. The edge is parameterized by the edge of a best-fit non-convex polygon of the local terrain. This permits the RECON-bot to move along the cliff edge without falling over it. In cliff edge parameterization, the surface currently in contact with the RECON-bot is identified in the environment model. This surface is then approximated by a best-fit polygon. The tolerance of the fit is limited by the known rover wheel diameter i.e. fit tolerance equal to the wheel characteristic length divided by the length per pixel. For this process the incomplete environment model is temporarily completed by a Markovian approximation for unknown grid cells. For all unknown points a worst case initial guess is assumed. This value is the lowest elevation value currently in the known model. A nearest measured neighbor average is performed and iterated till convergence. An example of this is shown in Figure 7.



Sample 1: Sinusoidal surface with random unknown cells



Sample 2: Step surface with half unknown cells

Figure 7. Markovian interpolation of unknown regions.

Using the Markovian approximation of the environment, the current rover contact surface (called the plateau) is first identified. This is achieved by setting a height threshold bound to the environment model and projecting the resulting data set onto the XY plane. This is followed by a region growing operation around the current known rover coordinates. Next, the binary image is smoothed by an image closing operation (dilation + erosion). Plateau edge pixels are easily identified at this stage. However, to remove small holes in the plateau, an edge following operation is performed. At this stage there is a single closed loop of boundary pixels. Finally, this set of points is parameterized by a closed polygon. This is initiated by fitting the full set of boundary pixels to a straight line. For any given sub set of boundary pixels that is currently fit to a line, if the error bound on this fit exceeds the prescribed tolerance, then the pixel set is divided into two, and the process is repeated. However, before error bound evaluation, line segments fit to each sub set of boundary pixels are joined to form a closed polygon. The cliff edge parameterization algorithm is outlined in Figure 8.

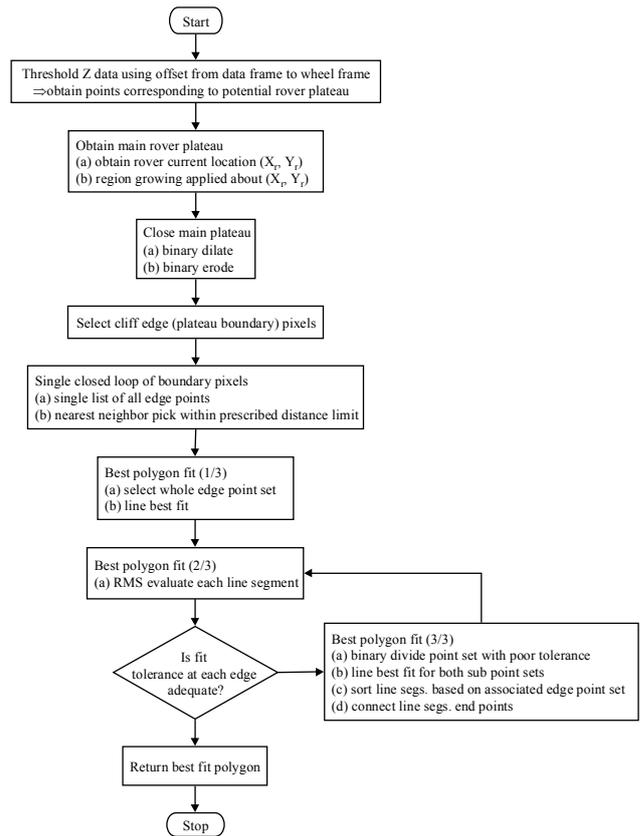
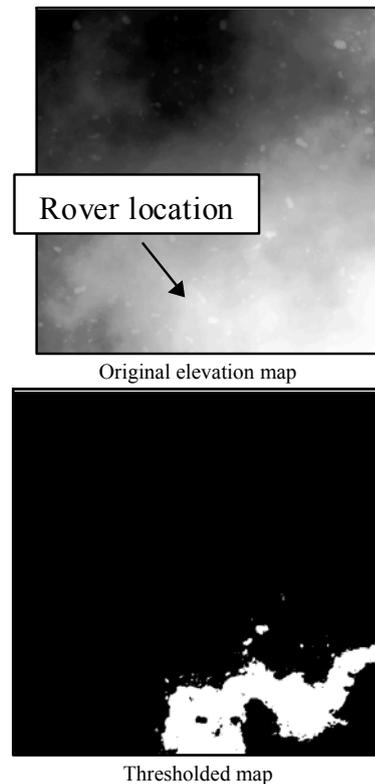


Figure 8. Cliff edge parameterization algorithm flow diagram.

An example of the process is shown in Figure 9 on a simulated Mars-type environment based on Viking I/II Mars lander rock distribution statistics.



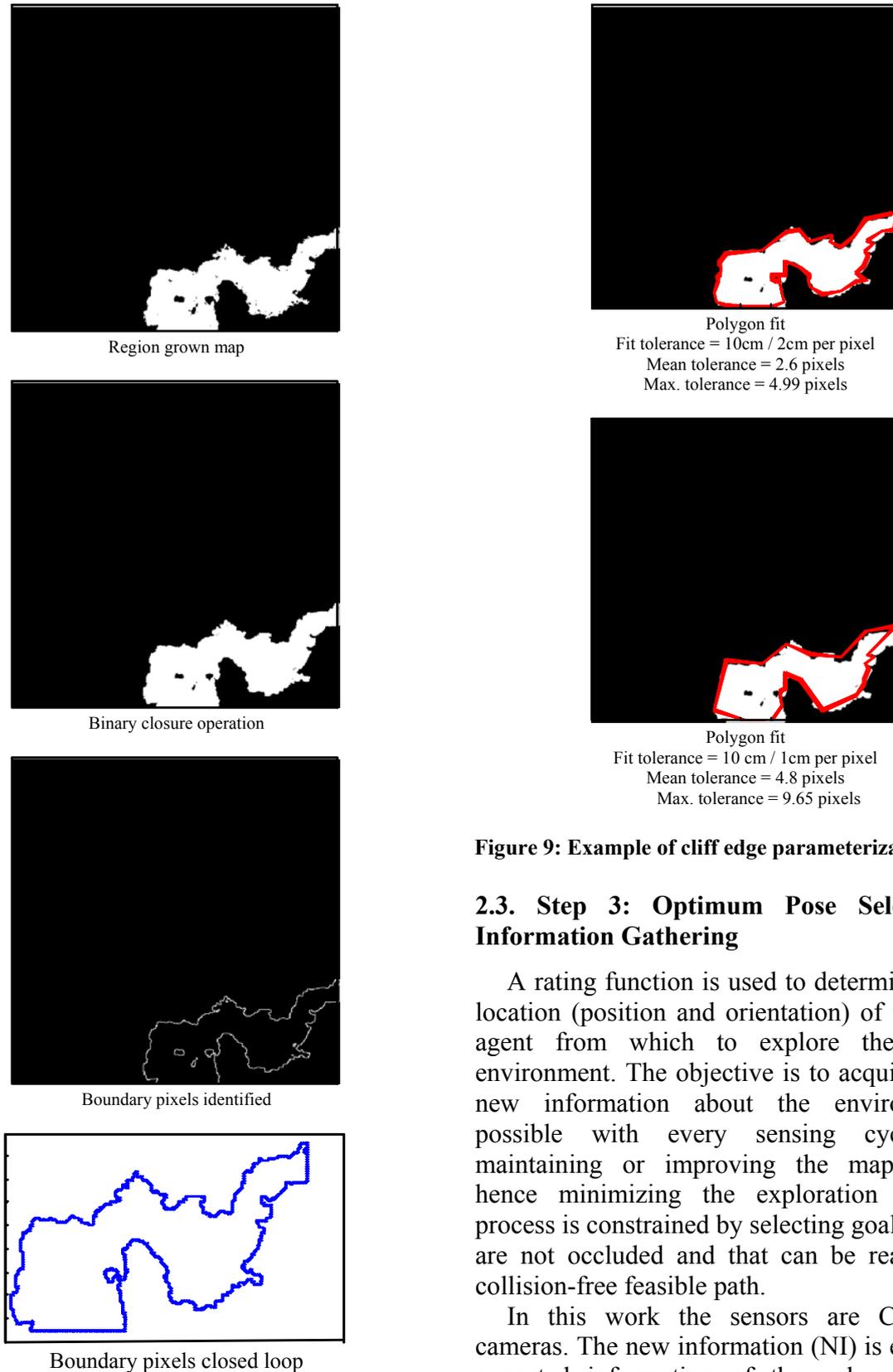


Figure 9: Example of cliff edge parameterization.

2.3. Step 3: Optimum Pose Selection for Information Gathering

A rating function is used to determine the next location (position and orientation) of the sensing agent from which to explore the unknown environment. The objective is to acquire as much new information about the environment as possible with every sensing cycle, while maintaining or improving the map accuracy, hence minimizing the exploration time. The process is constrained by selecting goal points that are not occluded and that can be reached by a collision-free feasible path.

In this work the sensors are CCD stereo cameras. The new information (NI) is equal to the expected information of the unknown/partially known region viewed from the sensor pose under consideration. This is based on the known obstacles from the current environment map, the field of view of the sensor and a framework for quantifying information. Shannon (1948) showed

that the information gained by observing a specific event among an ensemble of possible events may be described by the following function:

$$H(q_1, q_2, \dots, q_n) = -\sum_{k=1}^n q_k \log_2 q_k \quad (1)$$

Where q_k represents the probability of occurrence for the k^{th} event. This definition of information may also be interpreted as the minimum number of states (bits) needed to fully describe a piece of data. Shannon's emphasis was in describing the information content of 1-D signals. In 2-D the gray level histogram of an ergodic image can be used to define a probability distribution:

$$q_i = n_i / N \text{ for } i = 1, \dots, n \quad (2)$$

Where n_i is the number of pixels in the image with gray level i , N is the total number of pixels in the image, and n is the number of possible gray levels. With this definition, the information of an image for which all the q_i are the same - corresponding to a uniform gray level distribution or maximum contrast - is a maximum. The less uniform the histogram, the lower the information. Although this is generally true, it is critical to note that images with ordered patterns may result in the same information content as one with no order. For example, a uniform histogram may be mapped to two very different images, such as a random arrangement of intensity values and a (uniform) smooth color gradient. Intuitively, the former would be expected to contain more information, but using Equations (1) and (2), they result in the same value. This anomaly exists for both the 2.5-D and 3-D models. However, this is readily rectified using conventional lossless image compression algorithms, defined below. Thus, before the information content of a data set can be evaluated, it must be processed by a compression algorithm.

A *compression program* is used to convert data from an easy-to-use format to one optimized for compactness. Likewise, an *uncompression program* returns the information to its original form. Only compression is addressed here since just a measure on the information present after compression is

required, with no needs for decompression of the data. Decompression techniques can be inferred from the compression methods. Refer to [Smith 1999] for complete descriptions.

There are many different forms of compression, classified in various ways. One way to classify the compression techniques is lossless vs. lossy. A lossless technique means that the restored data file is *identical* to the original. This is absolutely necessary for many types of data, for example: executable code, word-processing files, tabulated numbers, etc. In comparison, data files that represent images and other required signals do not have to be kept in perfect condition for storage or transmission. All real world measurements inherently contain a certain amount of noise. If the changes made to these signals resemble a small amount of additional noise, no harm is done. Compression techniques that allow this type of degradation are called lossy. This distinction is important since lossy techniques are much more effective at compression than lossless methods. The higher the compression ratio, the more noise added to the data. A few common methods of lossless compression are Simple Run-length compression, Lossless JPEG, Huffman coding, and Lempel-Ziv-Welch (LZW) compression. An ideal compression algorithm would remove all traces of any pattern in the data. Such an algorithm currently does not exist, however the LZW compression algorithm is well recognized to approach this limit. A thorough review is beyond the scope of this work, but it can be found in [Smith 1999]. Limited studies on several of the above methods have been carried out and results are presented in Section 4.1.

This concept of information content may now be extended to both 2.5-D environment elevation maps as well as 3-D environment occupancy grid maps. The process is constrained by selecting goal points that are not occluded and that can be reached by a collision free traversable path.

2.3.1. Extension to 2.5-D Environment Elevation Map

In this sub-section, the above idea is extended to a 2.5-D signal such as an environment elevation map. In 2.5-D elevation maps, the map is a plane of grid cells where each grid cell value represents the average elevation of the environment at that

cell location. Uncertainty in elevation is also maintained with each grid cell. The new information (NI) content for a given sensor (camera) view pose is given by

$$H(\text{cam}_{x,y,z,\theta_p,\theta_y}) = \sum_i \frac{n_{\text{grid}}^{\text{max}} - n_{\text{grid}}^i}{n_{\text{grid}}^{\text{max}}} * \left\{ \left(\frac{P_V^i}{2} \log_2 \frac{P_V^i}{2} \right) + \left(1 - \frac{P_V^i}{2} \log_2 \left(1 - \frac{P_V^i}{2} \right) \right) \right\} \quad (3)$$

Where H is an information measure summed over all grid cells, i , visible from camera pose $\text{cam}_{x,y,z,\theta_p,\theta_y}$; n_{grid}^i is the number of environment points measured and mapped to cell i ; $n_{\text{grid}}^{\text{max}}$ is the maximum allowable mappings to cell i ; and P_V^i is the probability of visibility of cell i from the camera test pose. Note in Equation (3) that when the number of environment points measured and mapped to cell i reaches the maximum allowable mappings, its contribution to the new information content is zero, as expected. Locations where the elevation is greater than the local average elevation ($+2\sigma$) are considered as unoccupied and form candidate test poses for the vision system, since these locations are least likely to be obstructed.

A single range observation of a point (\bar{x}) is modeled as a 3-D Gaussian probability distribution centered at \bar{x} , based on two important observations. First, the use of the mean and covariance of a probability distribution function is a reasonable form to model sensor data and it is a second order linear approximation [Smith *et al.* 1986]. This linear approximation corresponds to the use of a Gaussian (having all higher moments of zero). Second, from the central limit theorem, the sum of a number of independent variables has a Gaussian distribution regardless of their individual distributions. The standard deviations along the three axes of the distribution correspond to estimates of the uncertainty in the range observation along these axes. These standard deviations are a function of intrinsic sensor parameters (such as camera lens shape accuracy) as well as extrinsic sensor parameters (such as the distance to the observed point or

feature). For most range sensing systems, this model can be approximated as [Sujan *et al.* 2002]

$$\sigma_{x,y,z} \approx S \cdot T_{x,y,z} \cdot L^n \quad (4)$$

Where S is an intrinsic parameter uncertainty constant, $T_{x,y,z}$ is an extrinsic parameter uncertainty constant, L is the distance to the feature/environment point, and n is a constant exponent (typically 2).

The probability of visibility P_V^i of a target cell ($\Delta x, \Delta y$) from the camera location is evaluated by computing the likelihood of occlusion of a ray $\text{ray}_{x,y,z}$ using all the elevations of the obstructions, $\text{Ob}_{x,y,z}$, and the associated uncertainties, $\sigma_{x,y,z}$, at cells lying along the ray path shot through each target in the environment grid to the camera center. From Figure 10, if grid cell i falls within the camera field of view, then its average elevation, $P_{t,x,y,z}$ (obtained either as an average of all measured points mapped to cell i , or as the Markovian approximation of its neighborhood if no points have currently been mapped to cell i) traces a ray to the camera center, $\text{Cam}_{x,y,z}$. Assuming z as the vertical direction, P_V^i is given by the product of the probability of visibility over all cells lying under the light ray from the target cell ($\Delta x, \Delta y$) to the camera location

$$P_V^i = \prod_{\text{ray}} \left\{ \text{sgn}(\text{ray}_z - \text{Ob}_z) * \int_0^{(\text{ray}_z - \text{Ob}_z)} \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma_z^2}\right) dz + 0.5 \right\} \quad (5)$$

Note that an exact solution for P_V^i should be given by the volumetric integral about dx , dy and dz . However, for computational simplicity only an approximation has been applied by considering the single dimension, dz . This definition for NI has an intuitively correct form. Regions with higher visibility and associated higher level of unknowns yield a higher expected NI value. Higher occlusions or better known regions result in lower expected NI values.

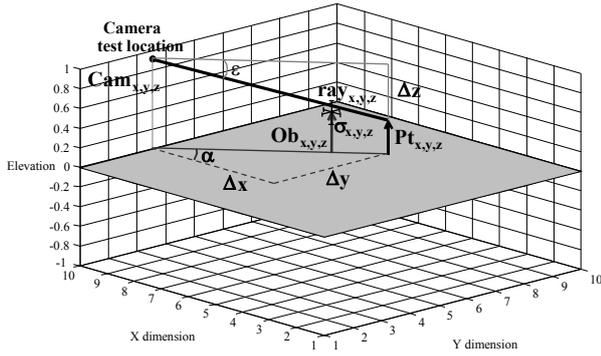


Figure 10: Ray tracing to determine probability of visibility of a grid cell from a given camera configuration.

2.3.2. Extension to 3-D Environment Occupancy Grid

In 3-D occupancy grid maps, the map is modeled as a probabilistic discretized occupancy 3-D grid. Every voxel in the 3-D grid has a value for probability-of-occupancy that ranges from 0 (empty) to 1 (occupied). A value of 0.5 indicates maximum uncertainty in occupancy of the voxel. From the probabilistic geometric environment model, (x,y,z) locations with probability of occupancy smaller than 0.05 (2σ) are considered as unoccupied. Such points form candidate configuration space camera pose coordinates. In such an instance the scene probability distribution for information analysis is still given by Equation (2). However, in this case N is the maximum number of voxels visible by the vision system (limited by the depth of field and the field of view), and n_i is the number of voxels in the scene with gray level i ($i = 1, \dots, n$). The equation is evaluated separately for mapped (known) versus unmapped (unknown) regions.

$$H(\text{cam}_{x,y,z,\theta_p,\theta_y}) = - \left(\sum_{k=1}^n q_k \log_2 q_k \right)_{\text{known}} + \left(\sum_{k=1}^n q_k \log_2 q_k \right)_{\text{unknown}} \quad (6)$$

The possible gray level values are defined as follows. For all unknown/unsampled voxels, an occupancy value $p(\bar{x})_{\text{unknown}}$ may be defined in the

form of a Markovian chain, i.e., $p(\bar{x})$ of a particular voxel is the average value of $p(\bar{x})$ of the neighboring voxels. Intuitively, this results in unknown regions that are mapped as averages of closest known regions. Thus, for all spatial voxels, a gray (probabilistic) occupancy value between 0 and 1 is found. Next, the values for $p(\bar{x})$ are modified as follows:

$$\text{stretching: } p'(\bar{x}) = \begin{cases} \frac{1}{1-p(\bar{x})} \cdot \frac{1}{d_{\text{voxel}}} & \forall p(\bar{x}) < 0.5 \\ \frac{1}{p(\bar{x})} \cdot \frac{1}{d_{\text{voxel}}} & \forall p(\bar{x}) \geq 0.5 \end{cases} + \text{scaling: } p''(\bar{x}) \quad (7)$$

$$= \begin{cases} \frac{p'(\bar{x})-1}{2} & \forall p(\bar{x}) < 0.5 \\ 1 - \frac{p'(\bar{x})-1}{2} & \forall p(\bar{x}) \geq 0.5 \end{cases}$$

Where d_{voxel} is the Euclidean distance of the voxel from the camera coordinate frame. This process causes regions with probability densities closer to 0 or 1 (regions of most certainty) or regions too distant from the camera (high d_{voxel}) to have a reduced effect on the new information expected. Regions that have a probability density closer to 0.5 (regions of least certainty of occupancy) are “stretched out” in the scene probability distribution, thus increasing the new expected information associated with those regions. A uniform discretization of this range of the range of scaled $p''(\bar{x})$ values may be performed. In this case, q_k is defined as the probability distribution of the histogram of the $p''(\bar{x})$ values. With these definitions, q_k is evaluated and the results applied to Equation (6), resulting in a metric for new information (NI). Alternatively, a possibly better choice is a uniform discretization of $p(\bar{x})$, namely p_k with $k = 1, \dots, n$, defining q_k as the ratio between the number of voxels in the scene with $p(\bar{x})$ in the p_k interval and the maximum number of visible voxels. To increase the contribution of regions with higher occupancy uncertainty to the information metric, the term $q_k \log_2 q_k$ of Equation (6) is premultiplied by $-(p_k \log_2 p_k + (1-p_k) \log_2 (1-p_k))$, reflecting the greater expected information available in such regions, resulting in

$$H(q) = - \left(\begin{array}{l} \left(\sum_{k=1}^n \{ -[p_k \log_2 p_k + (1-p_k) \log_2 (1-p_k)] \cdot q_k \log_2 q_k \} \right)_{known} \\ + \left(\sum_{k=1}^n \{ -[p_k \log_2 p_k + (1-p_k) \log_2 (1-p_k)] \cdot q_k \log_2 q_k \} \right)_{unknown} \end{array} \right) \quad (8)$$

Therefore, voxels with $p(\bar{x})$ close to zero or one will result in less expected information in the modified expression above, due to their greater level of certainty. The definitions for NI shown in Equations (3), (6) and (8), do behave in an intuitively correct form. For example, for a given sensor pose, if the field of view is occluded then NI decreases. If every point in the field of view is known to be empty, then NI is equal to zero. NI increases as the number of unknowns in the field of view increases. Further, the new expected information also increases in the presence of regions that are known with median probabilistic values, i.e., values that indicate with least amount of certainty whether a voxel is occupied or not. On the other hand, regions with high probabilistic values for occupancy result in reduced associated information.

2.3.3. Compensating for Poor Data Quality

During the mapping process, some regions that are expected to be visible may not be, because of sensor characteristics (e.g. lack of stereo correspondence due to poor textures or lighting conditions), and inaccuracies in the data model (e.g. expected neighboring cell elevations and uncertainties/occlusions), resulting in an unsuccessful measurement for that specific cell. However, after repeated unsuccessful measurements of cells expected to be visible, it becomes more likely that sensor characteristics are the limitation. In this case, it may be impossible to improve the measurements only by increasing their number. Therefore, instead of dedicating too much time trying to unsuccessfully measure those cells, the robot should lose interest in them and try to explore other areas. This is represented as a data quality function that reduces as the number of unsuccessful measurements of the visible grid cell increases. The information metric associated with such regions is pre-multiplied by an

“interest function”, IF, for the grid cell i given at the k^{th} unsuccessful measurement by

$$\begin{aligned} IF_i^0 &= 1 \\ IF_i^k &= \frac{1}{e^{\beta P_V^i}} \cdot IF_i^{k-1} \end{aligned} \quad (9)$$

where β is a scaling constant determined empirically, with larger values resulting in faster decrease of IF. Note that occluded regions do not translate to low data quality regions, because cells with low P_V^i resulting in an unsuccessful measurement are not as severely penalized as cells with high P_V^i . This permits future “interest” in such regions that may be explored later from another view location.

2.3.4. Data Fusion

The next step in environment map building is to fuse the newly acquired data by each agent with the environment model currently available to that agent. At any time, the sensors on each mobile robot can only observe a small part of their environment. However, measurements obtained from multiple view-points can provide reduced uncertainty, improved accuracy, and increased tolerance in estimating the location of observed objects [Smith *et al.* 1986]. To fuse multiple range measurements of a feature by sensors, a statistical model of sensor uncertainty is employed (see Figure 11). Current and previous range sensor measurements and their uncertainty models can be integrated to give an updated probabilistic geometric model of the environment.

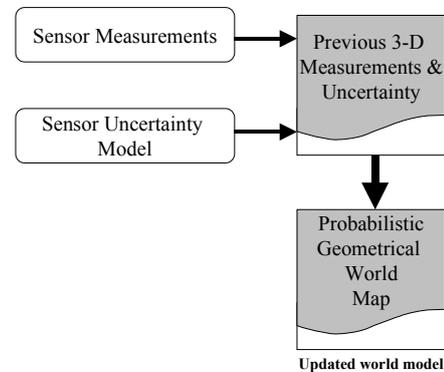


Figure 11: Data fusion with sensor uncertainty.

Each agent only fuses its own newly acquired data to the environment map stored in its memory.

Thus, as the environment map develops on an individual agent level, it needs to be shared and integrated among the team to keep each agent updated. Optimal map sharing protocols for multi agent systems is currently work in progress, i.e. decentralized protocols instructing the team members when and how to share their individual environment maps. However, once an agent shares its map, the other agents fuse this shared map into their own environment maps using the same method to fuse directly measured data, as described below.

Since the environment model has been developed in a fixed frame (see Step 1), all agents contributing to the environment map require identification of their vision system motion with respect to the fixed coordinate frame, i.e., the agents require global localization. This compensation process during the coordinate transformation reduces robot positioning errors, such as sensor motion errors, and vehicle suspension motions, and allows for accurate data fusion from multiple sources. The process for data fusion is as follows. A single spatial point in the fixed reference frame, $\bar{\mathbf{r}}_i$, is related to the image point (u_i, v_i) in the sensor frame by the 4x4 transformation matrix \mathbf{g}_{01} , see Figure 12. Spatial points are selected and tracked based on a Forstner interest operator and a homography transform [Huntsberger 2001].

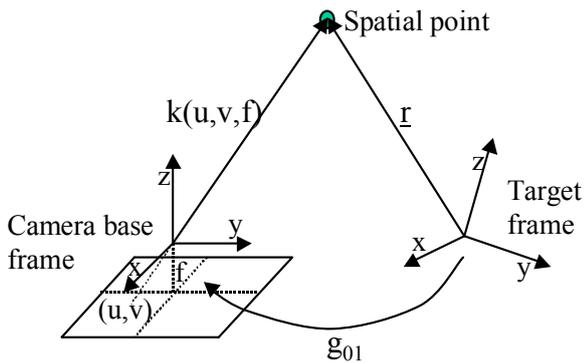


Figure 12: Relationship between sensor and target frames.

For motion calibration of a sensor, \mathbf{g}_{01} needs to be identified:

$$\begin{bmatrix} k_i u_i \\ k_i v_i \\ k_i f \\ 1 \end{bmatrix} = \mathbf{g}_{01} \cdot \bar{\mathbf{r}}_i = \begin{bmatrix} [\mathbf{R}_{01}]_{3 \times 3} & \bar{\mathbf{X}}_{3 \times 1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_i^x \\ r_i^y \\ r_i^z \\ 1 \end{bmatrix} \quad (10)$$

Where \mathbf{R}_{01} is the rotational matrix, $\bar{\mathbf{X}}$ is the translation vector, f is the sensor (camera system) focal length, and k_i is a scaling constant. For computational reasons, it is more convenient to treat the nine rotational components of \mathbf{R}_{01} as independent, rather than a transcendental relation of three independent parameters. Each spatial point gives three algebraic equations, but also introduces a new variable, k_i . This variable is a multiplicative constant to extend the i^{th} image point vector $(u, v, f)_i$ to the i^{th} spatial point in the sensor coordinate frame. These k_i may be found from the disparity pair of the stereo images. For n points it is found that

$$\mathbf{u} = \mathbf{g}_{01} \cdot \mathbf{r} \Rightarrow \begin{bmatrix} k_1 u_1 & k_2 u_2 & & k_n u_n \\ k_1 v_1 & k_2 v_2 & & k_n v_n \\ k_1 f & k_2 f & \dots & k_n f \\ 1 & 1 & & 1 \end{bmatrix} = \mathbf{g}_{01} \cdot \begin{bmatrix} r_1^x & r_2^x & & r_n^x \\ r_1^y & r_2^y & & r_n^y \\ r_1^z & r_2^z & \dots & r_n^z \\ 1 & 1 & & 1 \end{bmatrix} \quad (11)$$

This set of linear equations can be readily solved using conventional techniques. A least mean square error solution is given by

$$\mathbf{g}_{01} = \mathbf{u} \cdot (\mathbf{r}^T \mathbf{r})^{-1} \mathbf{r}^T \quad (12)$$

The rotation matrix, \mathbf{R}_{01} , and the translation vector, $\bar{\mathbf{X}}$, of the sensor frame with respect to the base frame are extracted directly from this solution of \mathbf{g}_{01} . However, for real measured data and associated uncertainty, a larger number of spatial points is required to more correctly identify the geometric transformation matrix, \mathbf{g}_{01} . Given the $(i+1)^{\text{st}}$ spatial and image point, from the above equations, the new estimates of the rotational matrix and translational vector, \mathbf{R}_{i+1} and $\bar{\mathbf{X}}_{i+1}$, can be obtained. A recursive method is used

to determine the mean $\hat{\bar{\mathbf{x}}}$ and covariance C of $\bar{\mathbf{x}}$ and R_{01} based on the previous i measurements as follows.

$$\begin{aligned}\hat{\bar{\mathbf{x}}}_{i+1} &= \frac{(i\hat{\bar{\mathbf{x}}}_i + \bar{\mathbf{x}}_{i+1})}{i+1} \\ C_{\bar{\mathbf{x}}}_{i+1} &= \frac{iC_{\bar{\mathbf{x}}}_i + [\bar{\mathbf{x}}_{i+1} - \hat{\bar{\mathbf{x}}}_{i+1}][\bar{\mathbf{x}}_{i+1} - \hat{\bar{\mathbf{x}}}_{i+1}]^T}{i+1} \\ \hat{\mathbf{R}}_{i+1}^{(l,m)} &= \frac{(i\hat{\mathbf{R}}_i^{(l,m)} + \mathbf{R}_{i+1}^{(l,m)})}{i+1} \\ C_{\mathbf{R}^{(l,m)}}_{i+1} &= \frac{iC_{\mathbf{R}^{(l,m)}}_i + [\mathbf{R}_{i+1}^{(l,m)} - \hat{\mathbf{R}}_{i+1}^{(l,m)}][\mathbf{R}_{i+1}^{(l,m)} - \hat{\mathbf{R}}_{i+1}^{(l,m)}]^T}{i+1}\end{aligned}\quad (13)$$

This method essentially maintains a measure on how certain the sensor motion is with respect to its original configuration (assuming the original configuration is known very precisely with respect to the common reference frame). This sensor pose uncertainty must be accounted for to obtain an estimate on the position uncertainty of a measured point in the environment. Let the measurement $\bar{\mathbf{z}}$ be related to the state vector (actual point position) $\bar{\mathbf{x}}$ by a non-linear function, $\mathbf{h}(\bar{\mathbf{x}})$. The measurement vector is corrupted by a sensor noise vector $\bar{\mathbf{v}}$ of known covariance matrix \mathbf{R} , giving

$$\bar{\mathbf{z}} = \mathbf{h}(\bar{\mathbf{x}}) + \bar{\mathbf{v}} \quad (14)$$

Assume that the measurement of the state vector $\bar{\mathbf{x}}$ is done multiple times. In terms of the current measurement, a Jacobian matrix of the measurement relationship evaluated at the current state estimate is defined as

$$\mathbf{H}_k = \left. \frac{\partial \mathbf{h}(\bar{\mathbf{x}})}{\partial \bar{\mathbf{x}}} \right|_{\bar{\mathbf{x}}=\bar{\mathbf{x}}_k} \quad (15)$$

The state (or position) may then be estimated as follows:

$$\begin{aligned}\mathbf{K}_k &= \mathbf{P}_k \mathbf{H}_k^T [\mathbf{H}_k \mathbf{P}_k \mathbf{H}_k^T + \mathbf{R}_k]^{-1} \\ \hat{\bar{\mathbf{x}}}_{k+1} &= \hat{\bar{\mathbf{x}}}_k + \mathbf{K}_k [\bar{\mathbf{z}}_k - \mathbf{h}(\bar{\mathbf{x}}_k)] \\ \mathbf{P}_{k+1} &= [1 - \mathbf{K}_k \mathbf{H}_k] \mathbf{P}_k\end{aligned}\quad (16)$$

Where \mathbf{P}_k and \mathbf{R}_k are the covariance of the state error and measurement noise after k measurements.

This estimate is known as the Extended Kalman Filter [Gelb 1974]. Using this updated

value for both the measured point $\bar{\mathbf{x}}$ and the absolute uncertainty \mathbf{P} , the measured point may then be merged with the current environment model.

A method to obtain appropriate spatial points is now addressed. Spatial points are a visible set of fiducials that are tracked during sensor motion. As the sensor moves, the fiducials move relative to the sensor, eventually moving out of the sensor view. This requires methods to identify and track new fiducials. Fiducials are selected from the probabilistic environment model based on three criteria: the degree of certainty with which a sampled point is known, the visual contrast of the sampled point with its surroundings, and depth contrast of the sampled point with its surroundings. These are combined into a single fiducial evaluation function (FEF):

$$\text{FEF} = f_1(P(\mathbf{x})) + f_2(C(u, v)) + f_3(H(\mathbf{x})) \quad (17)$$

where:

- $f_1(P(\mathbf{x})) \sim P(\mathbf{x})/r$ is the fiducial certainty, where r is the radius of a sphere centered at the potential fiducial within which neighboring voxels have descending certainty levels (outside this sphere, voxel certainty levels increase, and lower values for r suggest that the region surrounding a potential fiducial is well known, a desirable property);
- $f_2(C(u, v)) \sim \text{contrast } (C) \times \text{window size } (w)$ is the fiducial visual contrast, with contrast defined as:

$$C(u, v) = \frac{I(\mathbf{x}) - \bar{I}_w}{\bar{I}_w} \quad (18)$$

where $I(\mathbf{x})$ is the 2D image intensity value of the potential fiducial at \mathbf{x} , \bar{I}_w is the average intensity of a window centered at the potential fiducial in the 2D image, and w is the maximum window size after which the contrast starts to decrease; and

- $f_3(H(\mathbf{x})) \sim H(\mathbf{x}) \times \text{window size } (w)$ is the fiducial depth contrast, where $H(\mathbf{x})$ here is the maximum spatial frequency (from a 3D Fourier transform) at the potential fiducial within a window, and w is the maximum window size after which the power spectrum (of the 3D Fourier transform) starts shifting to higher frequencies (to simplify computation, this may be approximated with some heuristics).

Additionally, a penalty is added if a potential fiducial is too close to other identified fiducials. Using the identified fiducials, sensor motion can be obtained. Fiducials can be tracked with simple methods such as region growing or image disparity correspondence. An example of this process is shown in Figure 13. The flow diagram of visual system motion identification is shown in Figure 14.

Once the sensor motion has been identified using spatial mapped and tracked fiducials, the newly acquired data may be merged with the existing model as follows. Provided two observations are drawn from a normal distribution, the observations can be merged into an improved estimate by multiplying the distributions. Since the result of multiplying two Gaussian distributions is another Gaussian distribution, the operation is symmetric, associative, and can be used to combine any number of distributions in any order. The canonical form of the Gaussian distribution in n dimensions depends on the standard deviations, $\sigma_{x,y,z}$, a covariance matrix (C) and the mean (\bar{x}) [Stroupe 2000, Smith 1986]:

$$p(\bar{x}' | \bar{y}) = \frac{1}{(2\pi)^{n/2} \sqrt{|C|}} \exp\left(-\frac{1}{2}(\bar{y} - \bar{x}')^T C^{-1} (\bar{y} - \bar{x}')\right),$$

$$\text{where } C = \begin{bmatrix} \sigma_x^2 & \rho_{xy}\sigma_{xy}\sigma_{xy} & \rho_{zx}\sigma_{zx}\sigma_{zx} \\ \rho_{xy}\sigma_{xy}\sigma_{xy} & \sigma_y^2 & \rho_{yz}\sigma_{yz}\sigma_{yz} \\ \rho_{zx}\sigma_{zx}\sigma_{zx} & \rho_{yz}\sigma_{yz}\sigma_{yz} & \sigma_z^2 \end{bmatrix} \quad (19)$$

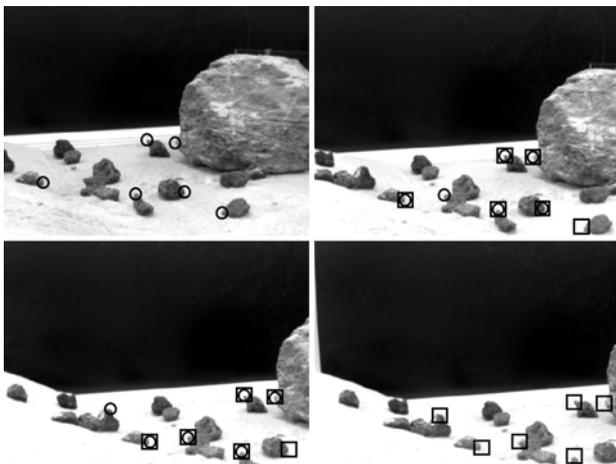


Figure 13: Identification and tracking of 6 fiducials (□= tracked with previous image, ○= tracked with next image).

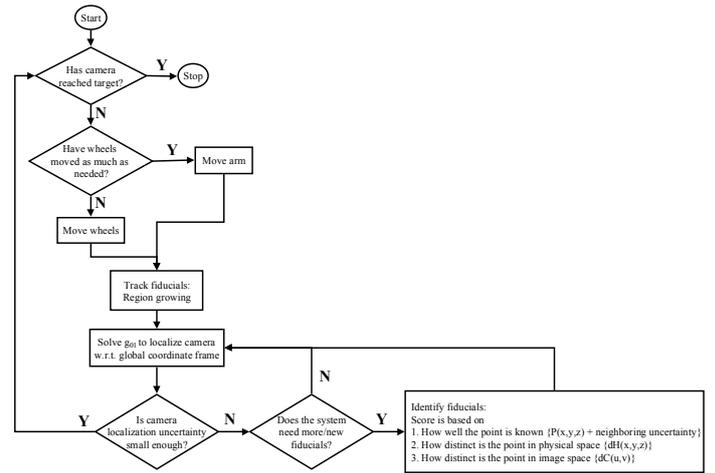


Figure 14: Flowchart for vision system motion identification using scene fiducials.

Where the exponent is called the Mahalanobis distance. For uncorrelated measured data, it is found that $\rho = 0$. The formulation in Equation (19) is in the spatial coordinate frame. However, all measurements are made in the sensor coordinate frame. This problem is addressed through a transformation of parameters from the observation frame to the spatial reference frame as follows:

$$C_{\text{transformed}} = R(-\bar{\theta})^T \cdot C \cdot R(-\bar{\theta}) \quad (20)$$

where $R(\theta)$ is the rotation matrix between the two coordinate frames. The angle of the resulting principal axis can be obtained from the merged covariance matrix [Smith 1986, Stroupe 2000]:

$$C_{\text{merged}} = C_1 (\mathbf{I} - C_1 (C_1 + C_2)^{-1}) \quad (21)$$

Where C_i is the covariance matrix associated with the i^{th} measurement. Additionally, a translation operation is applied to the result from Equation (19), to bring the result into the spatial reference frame. To contribute to the probabilistic occupancy environment model, all measured points corresponding to obstacles are merged. That is, all measured points falling in a particular grid cell contribute to the error analysis associated with that voxel.

Note that adding noisy measurements leads to a noisier result. For example, the sensor pose uncertainty increases as the number of sensor steps increase. With every new step, the current uncertainty is merged with the previous uncertainty to get an absolute uncertainty in sensor pose. However, by merging redundant

measurements leads to a result with less noise.

2.3.5. Camera Pose Rating Function

In addition to maximizing information acquisition, as done in the previous sub-sections, it is also desirable to minimize travel distance and maintain/improve the map accuracy, while being constrained to move along feasible paths. An Euclidean metric in configuration space, with individual weights α_i on each degree of freedom of the sensor pose \bar{c} , is used to define the distance moved by the sensor:

$$d = \left(\sum_{i=1}^n \alpha_i (c_i - c'_i)^2 \right)^{1/2} \quad (22)$$

where c_i and c'_i are the components of the vectors \bar{c} and \bar{c}' of the new and current sensor poses respectively. Here, α_i is set to unity. In general, this parameter reflects the ease/difficulty in moving the vision system in the respective axis. Map accuracy is based on the accuracy of localization of each sensing agent. This may be obtained by adding the localization error (LE) of the agent along the path to the target. Paths containing more promising fiducials for localization result in higher utility in determining both the goal location and the path to the goal. The new information, the travel distance and the net improvement of map accuracy are combined into a single utility rating function (RF) that may be optimized to select the next view pose:

$$RF = (w_{NI} \times NI - w_d \times d - w_{LE} \times LE) \times (1 - P_{x,y,z}) \quad (23)$$

where w_{NI} , w_d and w_{LE} are scaling constants. This rating function can be evaluated and optimized to find the next sensor configuration from which to make future measurements of the environment. Although this choice of rating function is somewhat arbitrary, good results have been obtained. Additional constraints can also be accommodated. The vision system pose selection algorithm is outlined in Figure 15. Note that the movement of the vision system

may require motions by the mobile robot (in addition to manipulator motions). The flowchart in Figure 15 includes a simple path planning approach based on the principle of convex hulls [Sujan *et al.* 2002].

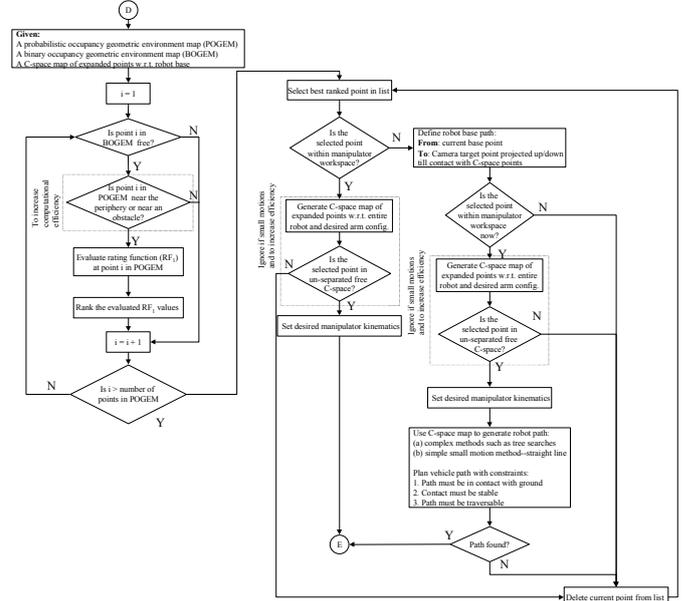


Figure 15: Flowchart for vision system pose selection of environment mapping algorithm.

2.4. Step 4: Map Distribution

After each agent maps and fuses an environment section to the environment map, it needs to distribute this updated map among the other agents. This is required so that each agent may optimally plan its next move and add information to the map. Once completed, the environment map needs to be distributed to the team. For instance, in the cliff exploration task, the Cliff-bot, which has limited sensors for navigation, depends on the RECON-bot to receive information on the acquired map [Huntsberger *et al.* 2001, Huntsberger *et al.* 2003, Sujan *et al.* 2003].

Due to communication bandwidth limitations, such as in the case of NASA/JPL present and near-term rovers, an appropriate data transfer algorithm needs to be developed. For example, during the 1997 Mars Sojourner mission, both the lander and rover carried 9600 baud radio modems, with an effective data rate of 2400bps [NASA 1997]. For the 2003 Mars Exploration Rover (MER) mission the data transfer rates of MER-to-Earth is expected to vary from 3Kbps to 12Kbps and MER-to-orbiter is expected to stay constant at

128Kbps [NASA 2003]. These communication limitations may be further exacerbated with multiple cooperating agents. Thus successful communication requires the reduction of the data set into relevant data i.e. only communicate data that is necessary for task execution.

The data reduction algorithm used here breaks down the environment map into a *quadtree of interest regions*. This is achieved in the 2.5-D model by first reducing the entire elevation map with adaptive decimation. This removes highly insignificant objects, such as small pebbles. The resulting data set is divided into four quadrants. The information content of each quadrant is evaluated using Equations (1) and (2). This information content reflects the amount of variation in the terrain quadrant (where higher information content signifies higher variation in the terrain). Quadrants with high information content are further divided into sub-quadrants and the evaluation process is continued. Once it is determined that a quadrant does not require further subdivision, an average elevation value of the particular quadrant is used for transmission (rather than the elevation of all grid cells within that quadrant). This cutoff threshold of information is based on a critical robot physical parameter (e.g. the wheel diameter). This results in a significantly reduced data set known as the *quadtree of interest regions*. Conventional lossless compression schemes may then be applied to the reduced data set to further reduce the number of transmission bits. The flow diagram of this process is given in Figure 16.

These 4 steps conclude the 3-D and 2.5-D versions of the algorithm. It has been shown above that the concept of information content can be applied to both 2.5-D environment elevation maps as well as 3-D environment occupancy grid maps [Sujan, 2002; Sujan *et al.*, 2002; Sujan *et al.*, 2003]. For 2.5-D and 3-D imaging, the new information (NI) is equal to the expected information of the unknown/partially known region viewed from the sensor pose under consideration. This is based on the known obstacles from the current environment map, the field of view of the sensor, and a framework for quantifying information. The goal is to find sensor poses that

yield high amounts of new information. However, these 2.5-D and 3-D versions of the algorithm require 3-D imaging systems such as stereo camera pairs. This is an undesirable hardware complexity for low-cost commercial applications for service robots in large scale. To address this issue, a 2-D version of the algorithm, which only requires a limited sensor suite, is presented next.

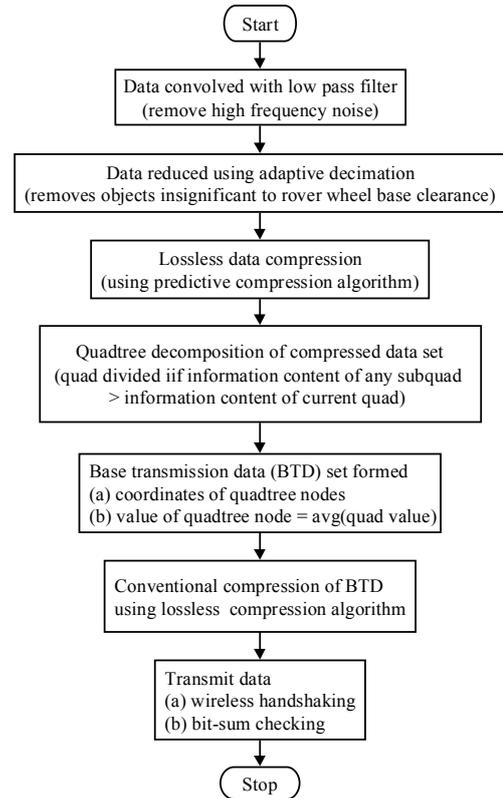


Figure 16: Inter-robot communication flow diagram.

In the next section, experimental results are presented to validate the approaches proposed in this work.

3. RESULTS

In this section, simulations and experimental results are presented to evaluate the proposed methodologies. Section 3.1 evaluates the 3-D mapping procedures described in Section 2.3, while the 2.5-D version of the algorithm is experimentally studied in Section 3.2.

3.1. Simulations on the 3-D Mapping Algorithm

The 3-D version of the proposed algorithm is evaluated from simulations aimed to develop a 3-D probabilistic occupancy model using the rating function for vision system pose selection

described in Section 2.3. Without loss of generality, a planar environment is considered in the 3-D model simulations. Two simulation results are presented: single sensor/robot modeling of an unstructured environment, and two cooperative sensors/robots modeling of an unstructured environment. Five sensor pose selection methods are compared:

- (i) *random pose selection* - the next sensor pose is selected randomly within the known environment;
- (ii) *sequential/raster pose selection* - the next sensor pose is selected as the next free location in the known environment from which measurements have not yet been made;
- (iii) *pose with maximum expected unmapped (new) region* - the next sensor pose is selected as the location with the largest expected new region while accounting for known occlusions;
- (iv) *pose with minimum mapped (old) region (also known as the Frontier strategy)* - the next sensor pose is selected as the location that will map the smallest previously mapped region; and
- (v) *pose with maximum expected information*.

The first two methods reflect commonly used environment mapping schemes [Burschka *et al.* 1997, Castellanos *et al.* 1998, Choset *et al.* 2001, Rekleitis *et al.* 2000, Victorino *et al.* 2000]. The latter three reflect with increased complexity the algorithm developed here. The rating function (RF) cannot be optimized analytically. Furthermore, exhaustive searching though the entire n -point configuration space (a process that takes $O(n)$ time) is computationally very expensive. Here, to reduce the search time, a finite set of goal configurations is employed. This set of goal configurations may be selected in several ways (random, closest to current pose, etc.). For m possible configurations, this process takes $O(m)$ time, where m often is a constant. Thus, while the best goal configuration would be the one maximizing RF, any configuration with a high value for RF should suffice. Such a configuration can be found with reasonable effort.

Figure 17 shows an unknown environment (100m × 100m) with occlusions (black) to be

mapped/modeled. It is assumed that all mobile mapping sensing agents start at the center of this environment. Figures 18-20 show the results of exploring this environment with a single robot (with a 90° field of view, 15m depth of field). Figure 18 shows examples of the area mapped using mapping/modeling methods (i) and (v). Figure 19 shows the average accumulated motion error of the vision system as it explores the environment as a function of traveled distance. Comparing Figures 19(a) and 19(b), it is seen that this error decreases substantially with redundancy of fiducials as well as with proximity of the fiducials. Figures 20(a) and 20(b) show respectively the fraction of the environment mapped and the net distance moved by the vision system for the five mapping methods, using a single mobile vision system. The energy consumption by the system is proportional to the net distance moved by the vision system. Hence it is desirable to have large fraction of the environment mapped with small net displacements.

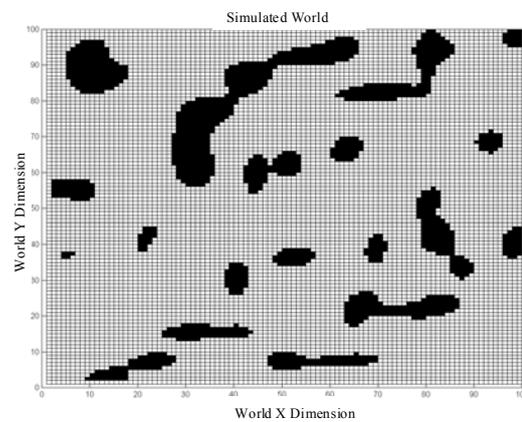


Figure 17: Unknown planar environment.

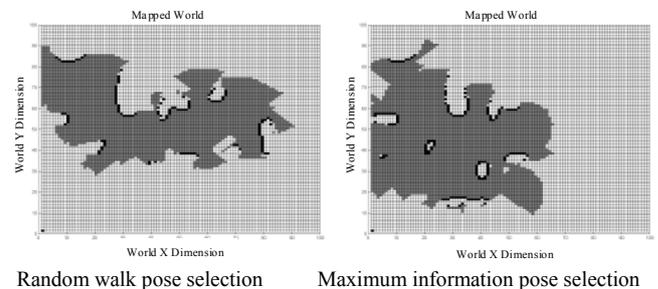


Figure 18: Mapped area by a single vision system (gray = empty space, black = obstacle, white = unknown).

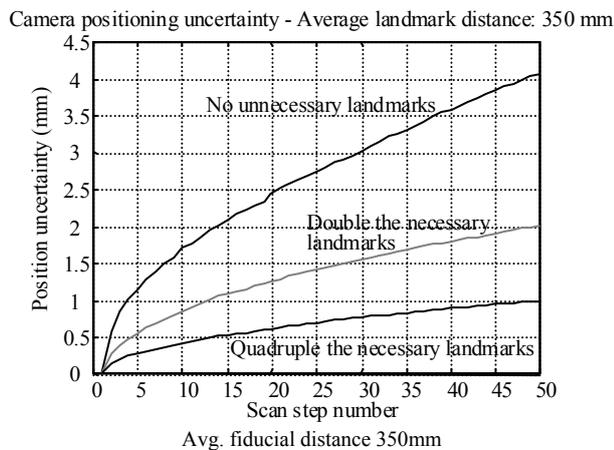
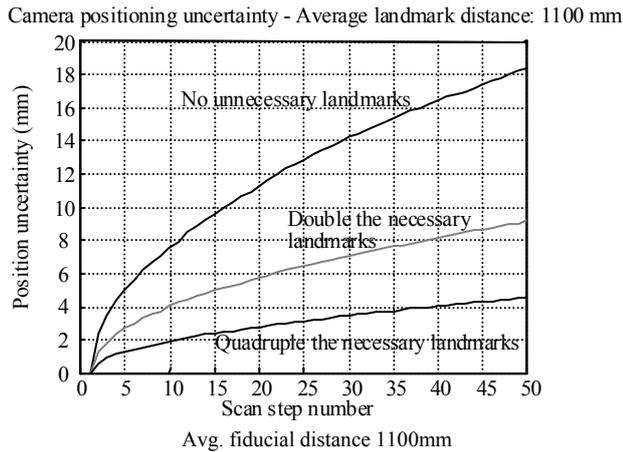


Figure 19: Accumulated r.m.s. translation error of vision system.

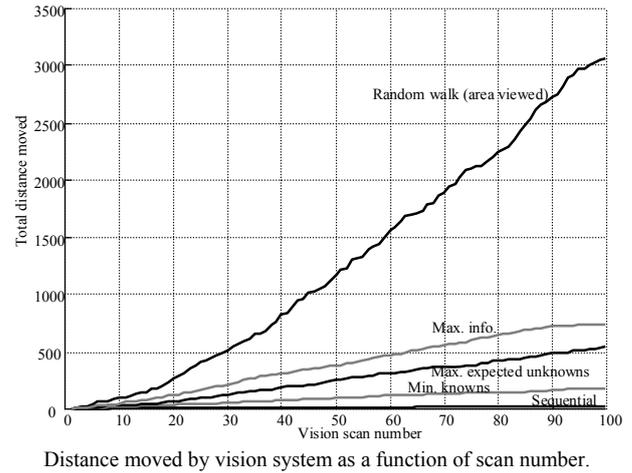
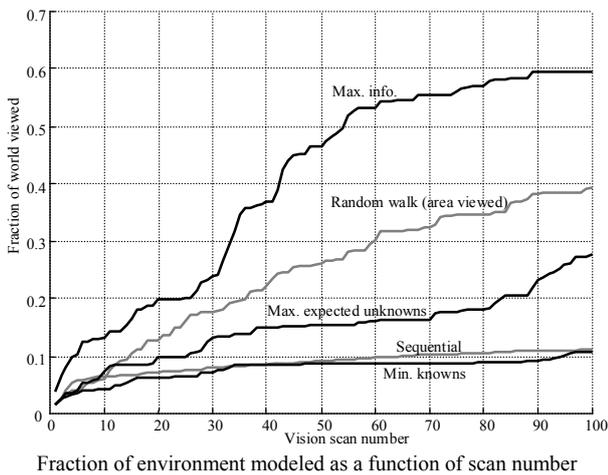


Figure 20: Results of single vision system modeling an unknown environment.

Note that with random walking you have the likelihood of ending up with a solution that is not as good as the information theory approach, but it is better than conventional methods such as (iii) or (iv). The problem with random walk is the huge amount of energy spent in moving around and the computational load in maintaining position information. The maximum expected unmapped method and minimum mapped region methods do not account for extrapolations on known obstacles. Hence although they do consider the obstacle that has been mapped, they don't try to figure out where the obstacle may exist. Additionally, when deciding on amount of world mapped, the information theory approach accounts for the certainty to which it is known. Hence, if certain areas have lower certainty (due to distance) these are often remapped with the random walk and the maximum information methods, but not so with methods (iii) or (iv). The max information approach also results in a slightly higher distance traveled compared to, e.g., the max expected unknowns, however the fraction of world viewed is so much higher that this turns out to be an advantage. If comparing the ratio between fraction of world viewed and distance traveled from these figures, the proposed approach turns out to be the best and random walk the worst based on the obtained results.

Figures 21-22 show the results of exploring the same environment cooperatively using two robots (each with 75° field of view, 10m depth of field). Figure 21 shows examples of the area mapped using mapping/modeling methods (i) and (v).

Figure 22 shows the fraction of the environment mapped in Figure 17 and the net distance moved by the vision system for the five mapping methods, using the two cooperating mobile vision systems. These results show the effectiveness of the information theoretic approach to vision system pose selection in environment modeling.

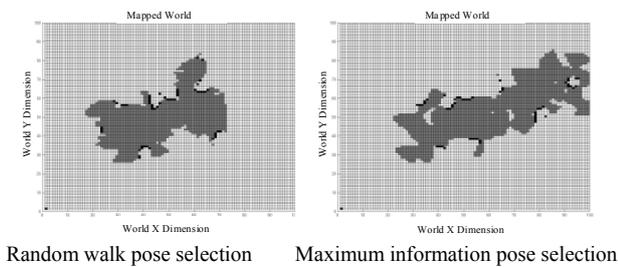
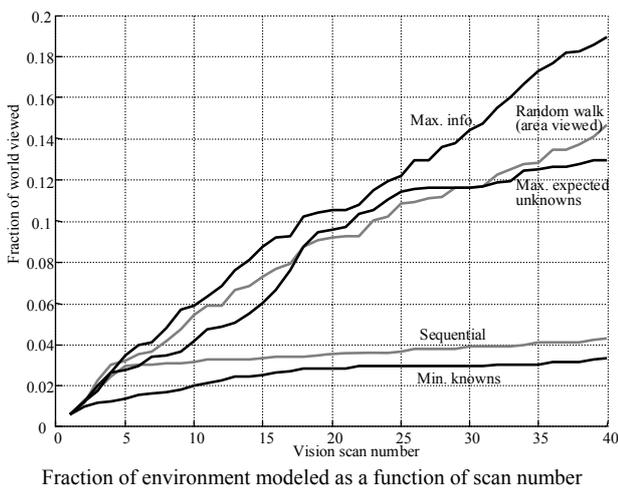
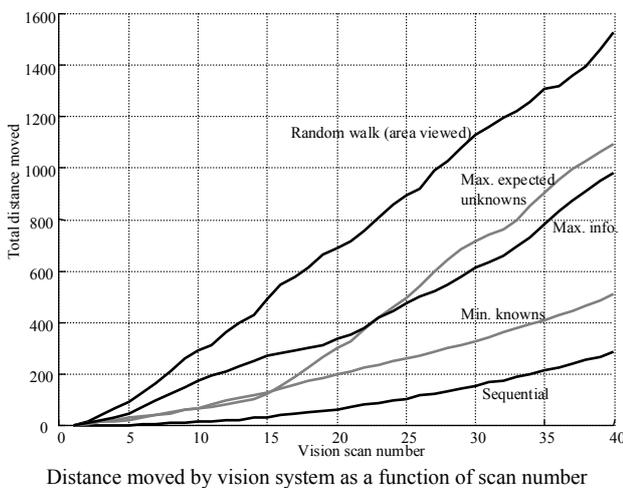


Figure 21. Mapped area by two cooperating vision systems (gray = empty space, black = obstacle, white = unknown).



Fraction of environment modeled as a function of scan number



Distance moved by vision system as a function of scan number

Figure 22: Results for two vision systems modeling an unknown environment.

3.2. Experiments on the 2.5-D Mapping Algorithm

In this section, the 2.5-D version of the presented algorithm is applied to the cooperative exploration of cliff surfaces by a team of four robots. The JPL Sample Return Rover (SRR) serves as the RECON-bot for this application. The SRR is a four-wheeled mobile robot with independently steered wheels and independently controlled shoulder joints. It carries a stereo pair of cameras mounted on a three DOF articulated manipulator. The SRR is equipped with a 266 MHz PC-104 computer platform, operating with VX-Works. Five mapping techniques, including the one developed above, are implemented:

- Method 1: Raster scanning without yaw;
- Method 2: Raster scanning with yaw;
- Method 3: Information-based environment mapping with cliff edge assumed to be a straight line segment;
- Method 4: Information-based environment mapping with cliff edge approximated as a non-convex polygon;
- Method 5: Information-based environment mapping with interest function and cliff edge approximated as a non-convex polygon.

The first two methods reflect commonly used environment mapping schemes [Asada 1990, Burschka 1997, Castellanos 1998, Choset 2001, Kuipers 1991, Rekleitis 2000, Victorino 2000]. The latter three reflect with increased complexity the algorithm developed here.

The experimental setup for the first study in the Planetary Robotics Lab (PRL) at JPL is shown in Figure 23. A recessed sandpit containing several rock piles is mapped. The edge of the sandpit, a vertical drop, acts as the cliff edge. This limits the motion of the RECON-bot to lie in the flat plane behind the cliff edge (see Figure 23). Figure 24 shows the number of environment grid cells explored as a function of the number of stereo imaging steps. From this experimental study, the improved efficiency of the method presented in this paper over conventional raster scanning methods can be seen, with an order of magnitude more points being mapped by Method 5 over those returned from Method 1 for the same number of stereo imaging steps. A significant improvement in efficiency can be seen while

progressing from Method 3 to Method 5. In Method 4, by parameterizing the cliff edge, the rover is able to follow the edge more aggressively, thus covering a larger variety of view points.

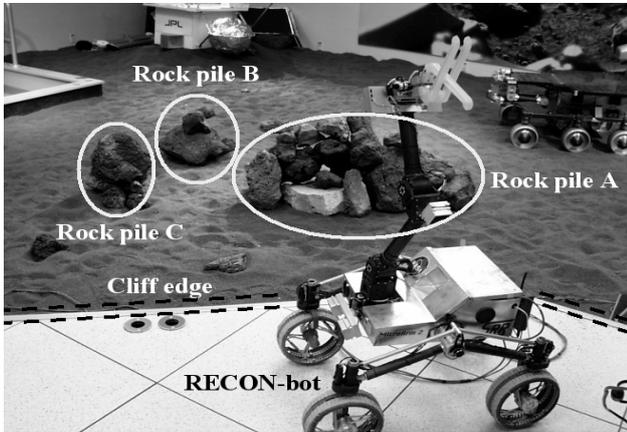


Figure 23: Experimental laboratory setup.

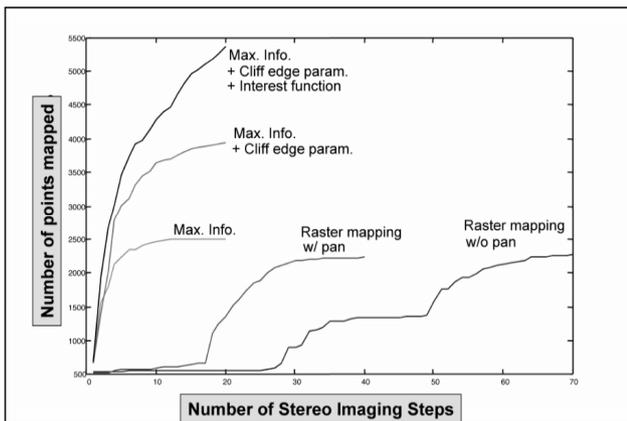
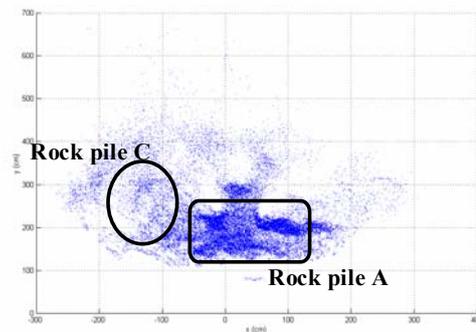


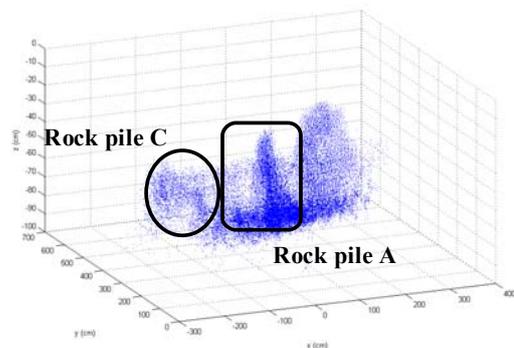
Figure 24: Amount of environment explored.

Figure 25 shows a top view of the environment points mapped using Methods 3 and 5. It is seen that Method 5 takes approximately half the number of steps to map a qualitatively similar region. Further, it is observed that the left region of the sandpit in Figure 23 yields poor data (due to lack of stereo correspondence). Since this region is expected to have high information content (due to lack of occlusions), the algorithm in Method 3 tends to converge to view points looking in that direction. However, in Method 5, the algorithm concludes that the data quality is poor and eventually loses interest in this region. This is seen in Figure 26, which shows the number of expected environment grid cell measurements as opposed to the number obtained. In Method 5,

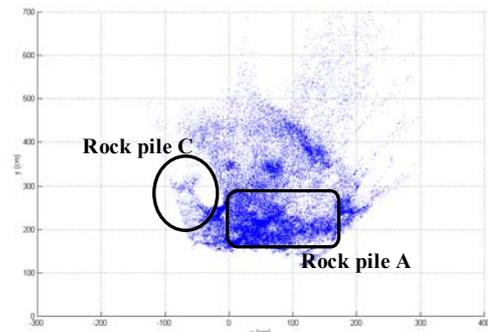
there is reasonable agreement. However, in Method 3, while the expected number of measurements is significant, the obtained number of grid cell measurements drops off to zero. Figure 26(c) shows an example of the projected mapped area as opposed to the true mapped area. Differences exist primarily due to poor imaging (stereo correspondence). However, occlusions and inaccuracies in projected area from local slope variations also contribute to this difference. Figure 27 shows the interest function value obtained from Equation (9) in Method 5 for each environment grid cell. It is seen that regions to the left rapidly lose their interest values with time since they yield low quality data.



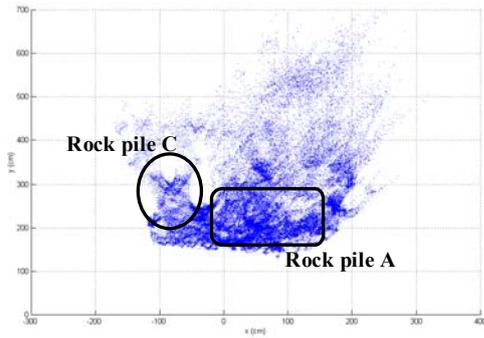
Method 1 – after 70 steps



Method 1 – after 70 steps



Method 3 – after 20 steps



Method 5 – after 10 steps

Figure 25: Top view of mapped points.

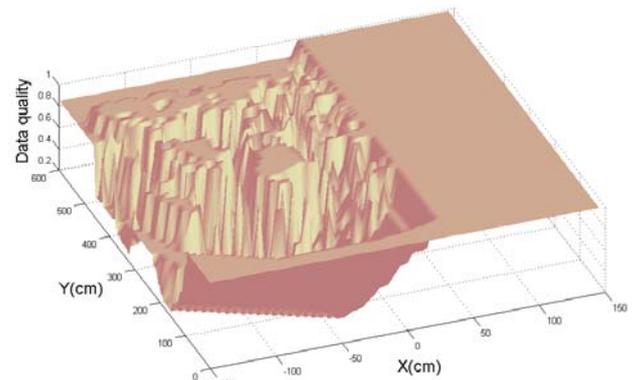
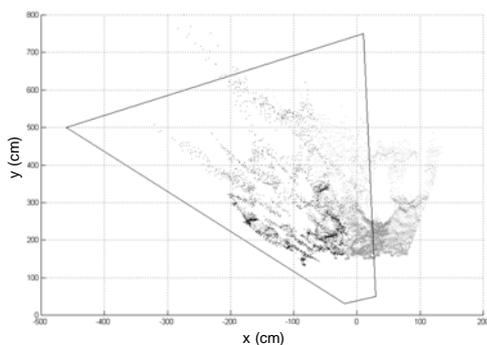
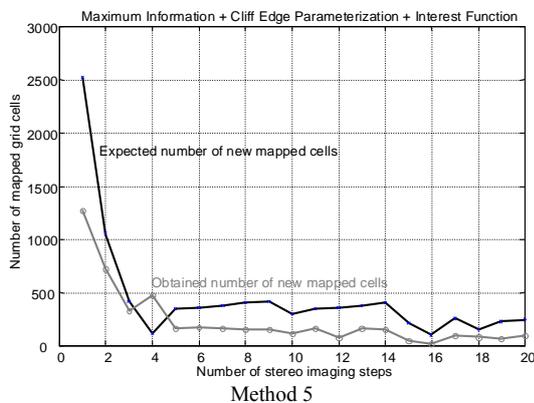
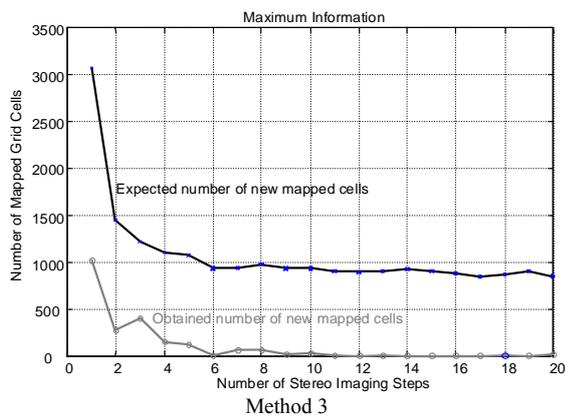


Figure 27: Interest function value after 10 steps using Method 5.



Selection of best view pose – projection of expected mapped region (black line) and acquired region (black points) using known mapped region (gray points)

Figure 26: Comparison of the number of expected new mapped cells versus the number obtained.

Field tests are conducted near the Tujunga Dam in Tujunga, CA on a natural cliff face with a vertical slope of approximately 75°. This setup is seen in Figure 28. This is the physical realization of the conceptual description in Figure 3 of a team of four cooperating robots exploring a cliff surface. Due to time constraints, experimental tests could only be run for Method 3 using the maximum information content and Method 5 using the maximum information content with interest function. The results of the study for 10 imaging steps are shown in Figure 29. Figure 30 shows part of the cliff surface and its corresponding map. Of particular interest to the Cliff-bot is the rock jumble in Figure 30(a), allowing it to choose to avoid it during traversal.

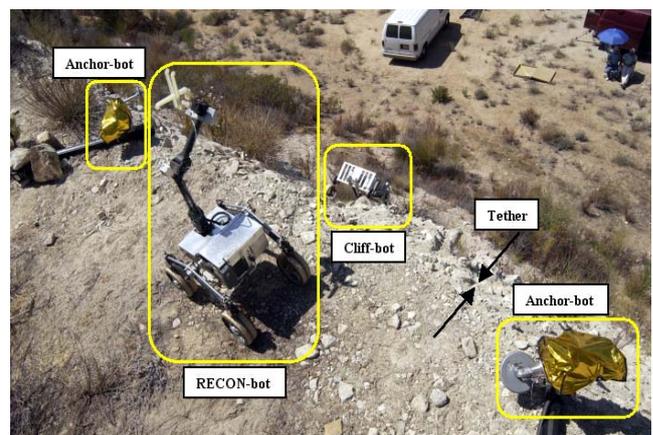


Figure 28: Experimental field system setup.

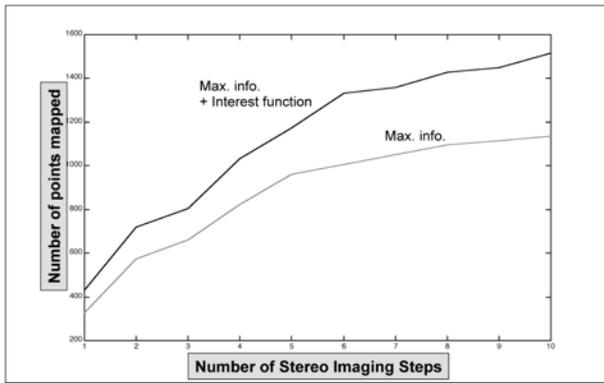
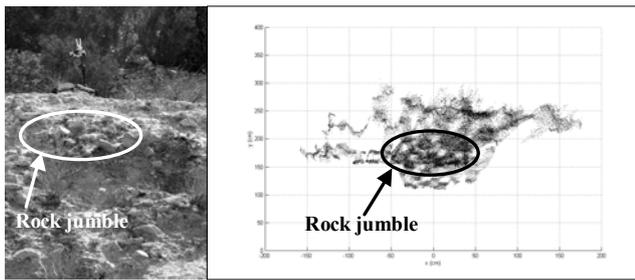


Figure 29. Amount of environment explored.



(a) View of RECON-bot mapping the cliff, (b) Overhead view of surface mapped

Figure 30: Tujung dam cliff site

Figure 31 compares the number of expected environment grid cell measurements and the number obtained for the two methods. Method 5 shows reasonable agreement, while Method 3 results in a large discrepancy. Once again, differences exist primarily due to poor imaging (stereo correspondence). However, occlusions and inaccuracies in projected area from local slope variations also contribute to this difference. Finally, Figure 32 shows the interest function value obtained in Method 5 for each environment grid cell using Equation (9).

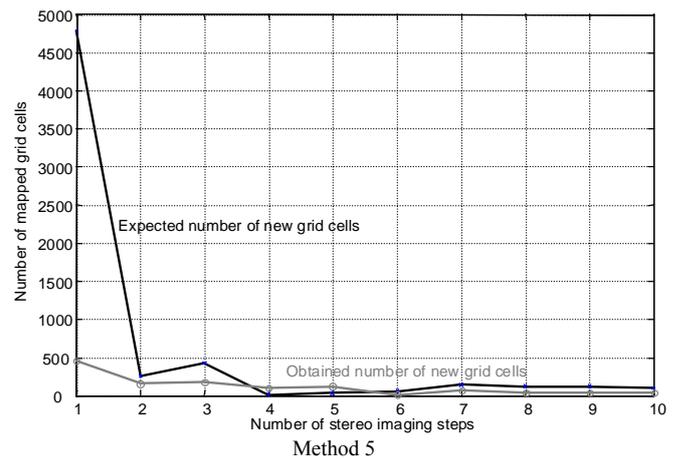
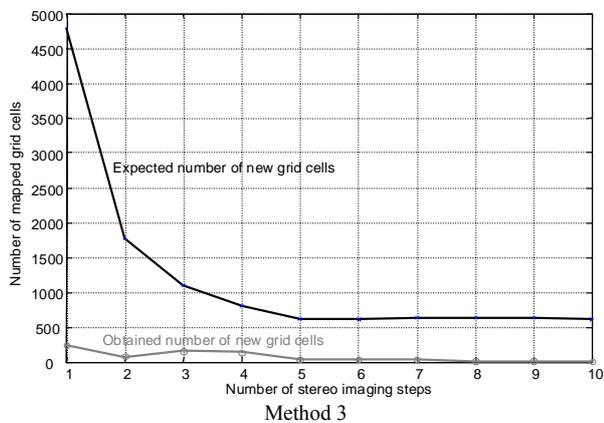


Figure 31: Comparison of the number of expected new mapped cells verses the number obtained.

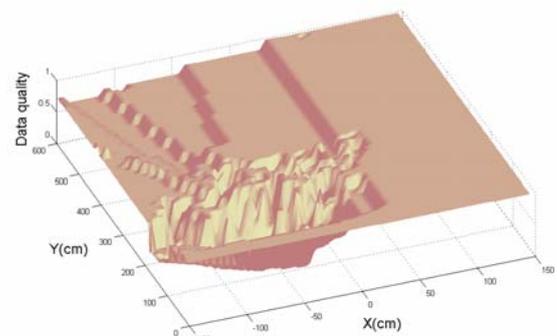


Figure 32. Interest function value after 10 steps using Method 5.

4. CONCLUSIONS

In this work, a new algorithm based on iterative sensor planning and sensor redundancy was proposed to build a geometrically consistent dimensional map of the environment for mobile robot teams with eye-in-hand systems. The aim was to acquire new information that could lead to more detailed and complete knowledge of the environment. Controlling robots to maximize knowledge was performed using Shannon's information theory-based evaluation functions. The work applied information theory to enhance the performance of cooperative sensing robot teams compared with traditional pose selection methods. It may be used by multiple distributed and decentralized sensing agents for efficient and accurate environment modeling. The algorithm makes no assumptions of the environment structure. Hence it is robust to robot failure since the environment model being built is not

dependent on any single agent frame. It accounts for sensing uncertainty, robot motion uncertainty, environment model uncertainty and other critical parameters. It allows for regions of higher interest getting more attention by the agents. Versions of the algorithm based on 3-D and 2.5-D representations of the environment were developed and experimentally validated through two cooperative robot exploration tasks: (i) mapping of an unknown planar environment and (ii) exploration of cliff surfaces. The presented approach has a potential benefit to significantly improve robot self-localization and mapping efficiency, while reducing the cost of autonomous mobile systems.

ACKNOWLEDGMENTS

The support from NASA - JPL is gratefully acknowledged, in special for the experimental data on the cliff exploration task and for the Mars elevation maps from the Viking I/II lander. In particular the contributions of Terry Huntsberger and Paul Schenker are acknowledged.

REFERENCES

- Anousaki, G.C. and Kyriakopoulos, K.J.: Simultaneous localization and map building for mobile robot navigation, *IEEE Robotics & Automation Magazine* 6(3), 1999, pp.42-53.
- Asada, M. *Map building for a mobile robot from sensory data*. *IEEE Transactions on Systems, Man, and Cybernetics*. Volume 37 no. 6. Nov/Dec 1990.
- Austin, D., and Kouzoubov, K. *Robust, Long Term Navigation of a Mobile Robot*. Proceedings of IARP/IEE-RAS Joint Workshop on Technical Challenges for Dependable Robots in Human Environments, October, 2002.
- Baumgartner, E.T., Schenker, P.S., Leger, C., and Huntsberger, T.L.: Sensor-fused navigation and manipulation from a planetary rover, Proceedings of SPIE Symposium on Sensor Fusion and Decentralized Control in Robotic Systems Vol. 3523, 1998.
- Bischoff, R. *Recent advances in the development of the humanoid service robot HERMES*. In 3rd EUREL Workshop and Masterclass - European Advanced Robotics Systems Development, volume I, pages 125-134, Salford, U.K., April 2000.
- Borenstein, J., and Koren, Y. *Real time obstacle avoidance for fast mobile robots in cluttered environments*. Proceedings of the IEEE International Conference on Robotics and Automation, pages 572-577, 1990.
- Burgard, W., Cremers, A., Fox, D., Hahnel, D., Lakemeyer, G., Schulz, D., Steiner, W., and Thrun, S. *The interactive museum tourguide robot*. Proceedings of the Fifteenth National Conference on Artificial Intelligence, Madison, WI, 1998.
- Burschka, D., Eberst, C., and Robl, C.: Vision based model generation for indoor environments, Proceedings of the 1997 IEEE International Conference on Robotics and Automation, Volume 3, Albuquerque, New Mexico, 1997, pp.1940-1945.
- Castellanos, J.A., Martinez, J.M., Neira, J. and Tardos, J.D.: Simultaneous map building and localization for mobile robots: a multisensor fusion approach, Proceedings of 1998 IEEE International Conference on Robotics and Automation, Volume 2, Leuven, Belgium, 1998, pp.1244-1249.
- Choset, H. and Nagatani, K.: Topological simultaneous localization and mapping (SLAM): toward exact localization without explicit localization, *IEEE Transactions on Robotics and Automation* 17(2), 2001, pp.125-137.
- Colgate, J., Wannasuphprasit, W., Peshkin, M. *Cobots: Robots for Collaboration with Human Operators*. Proceedings of the International Mechanical Engineering Congress and Exhibition, Atlanta, GA, DSC-Vol. 58, pp. 433-39, 1996.
- Dario, P., Guglielmelli, E., Laschi, E., Teti, G. *MOVAID: A Mobile Robotic System for Residential Care to Disabled and Elderly People*. Proceedings of the First MobiNet Symposium, Athens, Greece, May 15-16, 1997.

- Dubowsky, S., Genot, F., Godding, S., Kozono, H., Skwersky, A., Yu, H., and Yu, L.S. *PAMM - A Robotic Aid to the Elderly for Mobility Assistance and Monitoring: A "Helping-Hand" for the Elderly*. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 00), San Francisco, CA, vol. 1, pp. 570-576, April 2000.
- Engelberger, J. *Robotics in Service*, MIT Press, 1989
- Evans, J. *HelpMate: An autonomous mobile robot courier for hospitals*. In Proceedings of the International Conference on Intelligent Robots and Systems (IROS 94), pages 1695-1700, Munich, Germany, 1994.
- Gelb, A.: *Applied optimal estimation*, MIT press, Cambridge, Massachusetts, 1974.
- Glüer, D., Schmidt, G. *A New Approach for Context Based Exception Handling in Autonomous Mobile Service Robots*. Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 00), San Francisco, CA, pp 3272-3277. April 2000.
- Graf, B., Schraft, R.D., Neugebauer, J. *A Mobile Robot Platform for Assistance and Entertainment*. Proceedings of ISR-2000, Montreal, Canada.
- Haegele, M., Neugebauer, J. and Schraft, R. *From Robots to Robot Assistants*. Proceedings of the 32nd International Symposium on Robotics, ISR, April 19-21, 2001.
- Hager, G.D., Kriegman, D., Teh, E. and Rasmussen, C.: *Image-based prediction of landmark features for mobile robot navigation*, Proceedings of the 1997 IEEE International Conference on Robotics and Automation, Volume 2, Albuquerque, New Mexico, 1997, pp.1040-1046.
- Han, K., Kim, Y., Kim, J., and Hsia, S. *Internet Control of Personal Robot between KAIST and UC Davis*. Proceedings of the IEEE International Conference on Robotics and Automation, May 2002.
- Huntsberger, T.L., Rodriguez, G., and Schenker, P.S.: *Robotics: challenges for robotic and human Mars exploration*, Proceedings of ROBOTICS2000, Albuquerque, New Mexico, 2000, pp.299-305.
- Huntsberger, T.L., Pirjanian, P. and Schenker, P.S.: *Robotic outposts as precursors to a manned Mars habitat*, Proceedings of the 2001 Space Technology and Applications International Forum (STAIF-2001), Albuquerque, New Mexico, 2001, pp.46-51.
- Huntsberger, T.L., Sujan, V. A., Dubowsky, S., Schenker, P. *Integrated System for Sensing and Traverse of Cliff Faces*. Proceedings of the 2003 SPIE's 17th Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls: Symposium on Unmanned Ground Vehicle Technology V. April 21-24, Orlando, Florida USA.
- Jennings, C., Murray, D., and Little, J.: *Cooperative robot localization with vision-based mapping*, Proceedings of the 1999 IEEE International Conference on Robotics and Automation, Detroit, Michigan, 1999.
- Jung, D., Cheng, G., and Zelinsky, A. *Robot Cleaning: An Application of Distributed Planning and Real-time Vision*. International conference on Field and Service Robotics (FSR97), Canberra, Australia, 1997.
- Kawamura, K., Wilkes, D., Pack, T. *Humanoids: Future Robots for Home and Factory*. Proceedings of the First International Symposium on Humanoid Robots, Waseda University, Tokyo, Japan, October 30 -31, 1996, pp. 53-62.
- Khatib, O. *Mobile Manipulation: The Robotic Assistant*. Journal of Robotics and Autonomous Systems, vol. 26, 1999, pp.175-183.
- Kruse, E., Gutsche, R., and Wahl, F.M.: *Efficient, iterative, sensor based 3-D map building using rating functions in configuration space*, Proceedings of the 1996 IEEE International Conference on Robotics and Automation, Volume 2, Minneapolis, Minnesota, 1996, pp.1067-1072.
- Kuipers, B. and Byun, Y. *A robot exploration and mapping strategy based on semantic hierarchy of spatial representations*. Journal of Robotics and Autonomous Systems. Volume 8, Page(s) 47-63. 1991.

- Lara, B., Althoefer, K., and Seneviratne, L.D.: Automated robot-based screw insertion system, Proceedings Conference of the IEEE of the 24th Annual Industrial Electronics Society, IECON '98, Volume 4, 1998, pp.2440-2445.
- Lavery, D., The Future of Telerobotics. Robotics World, Summer 1996.
- Lawitzky, G. *A navigation system for cleaning robots*. Autonomous Robots, vol. 9, pp. 255--260, 2000
- Marrone, F., Strobel, M. *CleaningAssistant - A Service Robot Designed for Cleaning Tasks*. Proceedings of the 2001 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM '01). 8-11 July 2001, Como, Italy
- NASA:
http://mars.jpl.nasa.gov/MPF/rover/faqs_sojourner.html, 1997.
- NASA:
http://mars.jpl.nasa.gov/mer/mission/comm_data.html, 2003.
- Park, J., Jiang, B. and Neumann, U.: Vision-based pose computation: robust and accurate augmented reality tracking, Proceedings of the 2nd IEEE/ACM International Workshop on Augmented Reality, San Francisco, California, 1999, pp.3-12.
- Pettinaro G.C., Kwee I., Gambardella L.M., Mondada F., Floreano D., Nolfi S., Deneubourg J.-L. and Dorigo M. *SWARM Robotics: A Different Approach to Service Robotics*. Proceedings of the 33rd International Symposium on Robotics, Stockholm, Sweden: International Federation of Robotics, 2002.
- Pirjanian, P., Huntsberger, T.L., and Schenker, P.S.: Development of CAMPOUT and its further applications to planetary rover operations: A multirobot control architecture, Proceedings of SPIE Sensor Fusion and Decentralized Control in Robotic Systems IV, Volume 4571, Newton, Massachusetts, 2001, pp.108-119.
- Rekleitis, I., Dudek, G. and Milios, E.: Multi-robot collaboration for robust exploration, Proceedings of the 2000 IEEE International Conference on Robotics and Automation, Volume 4, San Francisco, California, 2000, pp.3164-3169.
- Reza, F.M., An introduction to information theory, Dover, New York, 1994.
- Schaeffer, C., May, T. *Care-O-bot: A System for Assisting Elderly or Disabled Persons in Home Environments*. Proceedings of Assistive Technology on the Threshold of the New Millennium Amsterdam u.a.: IOS Press, 1999, pp. 340-345 (Assistive Technologie Research Series 6).
- Schenker, P.S., Baumgartner, E.T., Lindemann, R.A., Aghazarian, H., Ganino, A.J., Hickey, G.S., Zhu, D.Q., Matthies, L.H., Hoffman, B.H. and Huntsberger, T.L.: New planetary rovers for long-range Mars science and sample return, Proceedings of SPIE Symposium on Intelligent Robots and Computer Vision XVII: Algorithms, Techniques, and Active Vision, Volume 3522, Boston, Massachusetts, 1998.
- Schenker, P.S., Pirjanian, P., Huntsberger, T.L., Trebi-Ollenu, A., Aghazarian, H., Leger, C., Dubowsky, S., McKee, G.T.: Robotic intelligence for space: Planetary surface exploration, task-driven robotic adaptation, and multirobot cooperation, Proceedings of SPIE Symposium on Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision, Volume 4572, Newton, Massachusetts, 2001.
- Schenker, P., Huntsberger, T., Pirjanian, P., Dubowsky, S., Iagnemma, K., and Sujan, V. *Rover Control for Intelligent and Agile Traverse of Challenging Terrain*. Invited session paper in the Proceedings of the IEEE 11th International Conference on Advanced Robotics (ICAR), June 30 - July 3, 2003, University of Coimbra, Portugal.
- Shannon, C. E. 1948. *A mathematical theory of communication*. The Bell System Technical Journal. Volume 27, Page(s) 379-423. July, 1948.
- Simhon, S. and Dudek, G.: Selecting targets for local reference frames, Proceedings of the 1998 IEEE International Conference on Robotics and Automation, Volume 4, Leuven, Belgium, 1998, pp.2840-2845.

- Smith, R.C. and Cheeseman, P.: On the representation and estimation of spatial uncertainty, *International Journal of Robotics Research* 5(4), 1986, pp.56-68.
- Smith, S.W.: *The Scientist and Engineer's Guide to Digital Signal Processing*, 2nd Edition, California Technical Publishing, San Diego, CA, 1999.
- Stroupe, A.W., Martin, M. and Balch, T.: Merging Gaussian distributions for object localization in multi-robot systems, *Proceedings of the Seventh International Symposium on Experimental Robotics, ISER '00*, Hawaii, 2000.
- Sujan, V.A. *Optimum Camera Placement by Robot Teams in Unstructured Field Environments*. *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, September 22-25, 2002, Rochester, New York.
- Sujan, V.A. and Dubowsky, S.: Visually Built Task Models for Robot Teams in Unstructured Environments, *Proceedings of the 2002 IEEE International Conference on Robotics and Automation, Volume 2*, Washington, D.C., 2002, pp.1782-1787.
- Sujan, V.A., Dubowsky, S., Huntsberger, T., Aghazarian, H., Cheng, Y. and Schenker, P.: Multi Agent Distributed Sensing Architecture with Application to Cliff Surface Mapping, *Proceedings of the 11th International Symposium of Robotics Research (ISRR)*, Siena, Italy, 2003.
- Sujan, V.A., Meggiolaro, M.A. *Intelligent and Efficient Strategy for Unstructured Environment Sensing using Mobile Robot Agents*. Submitted to the *Journal of Intelligent and Robotic Systems*, 2004.
- Tarabanis, K.A., Allen, P.K. and Tsai, R.Y.: A survey of sensor planning in computer vision, *IEEE Transactions on Robotics and Automation* 11(1), 1995, pp.86-104.
- Thrun, S. *When robots meet people: Research directions in mobile robots*. *IEEE Intelligent Systems* 1998.
- Thrun, S., Bennewitz, M., Burgard, W., Cremers, A., Dellaerr, F., Fox, D., Haehnel, D., Rosenberg, C., Roy, N., Schulte, L., and Schulz, D. *Minerva: A second generation mobile tour-guide robot*. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'99)*, May 1999, Detroit, MI, USA
- Thrun, S., Burgard, W. and Fox, D.: A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping, *Proceedings of the 2000 IEEE International Conference on Robotics and Automation, Volume 1*, San Francisco, California, 2000, pp.321-328.
- Tomatis, N., Nourbakhsh, I. and Siegwar, R.: Simultaneous localization and map building: a global topological model with local metric maps, *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, Volume 1*, Maui, Hawaii, 2001, pp.421-426.
- Trebi-Ollennu, A., Das, H., Aghazarian, H., Ganino, A., Pirjanian, P., Huntsberger, T. and Schenker, P.: Mars Rover Pair Cooperatively Transporting a Long Payload, *Proceedings of 2002 IEEE International Conference on Robotics and Automation, Volume 4*, Washington, D.C., 2002, pp.3136-3141.
- Ulrich, I., Mondada, F., and Nicoud, J. *Autonomous Vacuum Cleaner*. *Robotics and Autonomous Systems* 19, 1997, pp. 233-245.
- Victorino, A.C., Rives, P. and Borrelly, J.-J.: Localization and map building using a sensor-based control strategy, *Proceedings of 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems, Volume 2*, Takamatsu, Japan, 2000, pp.937-942.
- Wellman, P., Krovi V., Kumar, V. *An Adaptive Mobility System for the Disabled*. *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, San Diego, May 8-13, 1994
- Wong, S., Coghill, G., MacDonald, B. *Natural landmark recognition using neural networks for autonomous vacuuming robots*. *Proceedings of the 6th International Conference on Control, Automation, Robotics and Vision, ICARCV'00*, Singapore, December 2000.
- World Robotics 2001 - Statistics, Market Analysis, Forecasts, Case Studies and Profitability of*

Robot Investment. Produced by the United Nations Economic Commission for Europe (UNECE) in cooperation with the International Federation of Robotics (IFR). No. GV.E.01.0.16 or ISBN No. 92-1-101043-8

Yamauchi, B., Schultz, A. and Adams, W.: Mobile robot exploration and map-building with continuous localization, Proceedings of 1998 IEEE International Conference on Robotics and Automation, Volume 4, Leuven, Belgium, 1998, pp.3715-3720.

Yeh, E., and Kriegman, D.J.: Toward selecting and recognizing natural landmarks, Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems, Volume 1, Pittsburgh, Pennsylvania, 1995, pp.47-53.