Chapter 9

3D Probabilistic Occupancy Grid to Robotic Mapping with Stereo Vision

Anderson A. S. Souza, Rosiery Maia and Luiz M. G. Gonçalves

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/49050

1. Introduction

Environment mapping is considered an essential skill for a mobile robot in order to actually reach autonomy [1]. The robotic mapping can be defined as the process of acquiring a spatial model of the environment through sensory information. The environment map allows mobile robots to interact coherently with objects and people in this environment. The robot can safely navigate, identify surrounding objects and have flexibility to dealing with unexpected situations. Without a map some important operations could be complex as the determination of objects position in the surroundings of the robot and the definition of the path to be followed. These issues involve the importance of the mapping task be performed correctly, since the acquisition of inaccurate maps can lead to errors in the inference of correct positioning of the robot, resulting in an imperfect implementation of these operations. Therefore there is a mutual dependence between inferring the exact localization of the robot and building an accurate map.

There are several researches done in robotics mapping proposing ways to represent a mapped environment, all of them concerned in dealing with high dimensional mapped environment. The work of Agelopoulu et al. [3] provides a good overview on articles published at the very early period of research about mapping and it reports the several different ways of representing the environment, focusing on the main features of each approach. Thrun [1] proposes a classification of the ways of representing an environment mapped into two main categories: metric and topological representations. The metric representations store geometric properties of the environment, whereas the topological representations describe connectivity between different places. Within the metric representations, the occupancy grid stands out by providing a relatively accurate reproduction of the mapped environment.

Robots can be used to map internal or external environments depending on the task type supported by them. Lee et. al. [4], for example, uses a mobile robot equipped with sonar to build a features-based map. The robot finds points, lines and circles in the environment through processing of the information provided by sonar. Dealing with external

environments, Guivant et. al. [5] has implemented a mobile vehicle to map an environment typical of a farm. The map built by the vehicle was composed of typical objects of the environment, as trees and stakes. However, the majority of the works dealing with the robotics mapping theme do not discuss a generalist alternative to provide the mapping of both internal and external environments. The difficulty is in the detection of characteristics inherent in all kind of environments. It is well known that internal environments are more structured, so that the vast majority of them have common cues, for example, lines, nooks and corners. External terrain mapping depends on the objects that can be highlighted in them (transit cards, buildings, trees, and rocks, among others).

Several types of sensors can be used for carrying out the mapping. The most common are sonar, lasers and cameras. The sonar is attractive because of its low cost. Besides being relatively inexpensive they can be easily found in a commercial market. However, sonar presents significant inaccuracies in the measurements acquired. Lasers are highly accurate and provide the acquisition of detailed maps but they are not attractive because of its high cost. Cameras are sensors that, at each day, are getting cheaper and they make possible the acquisition of a large amount of information surrounding the robot. For these reasons, cameras are playing an important role among the sensors used for robotic mapping.

This chapter proposes the mapping of internal and/or external environments through a system of stereo cameras of low-cost with metric representation of the environmental in a probabilistic occupancy grid. With the stereo vision system the robot can collect information from different places with different types of obstacle, and it does not dependent on the type of environment in which it is located or the type of features inside this place. The mapping algorithm considers a probabilistic modeling for the vision system used by the robot, as well as to its performed movements. With this, the 3D Probabilistic Occupancy Grid to Robotic Mapping with Stereo Vision generates results (maps) consistent with the information obtained by the robot. To attest to the feasibility proposed by that research will be presented preliminary experiments performed.

The article is structured as follows: section 2 presents the problem of environment mapping with robots. The section introduce the occupancy grid mapping. Section 3 contextualizes the same problem with the use of visual sensors (cameras). In section 4 will be presented the proposed mapping with stereo vision and occupation grid representation. Section 5 expose the preliminary results achieved with the mapping algorithm. Finally, the section 6 brings the concluding remarks, directing to the next steps that will be given toward the full implementation of the system.

2. Robotic mapping

To better formalize the problem of robotic mapping, we need to establish some basic assumptions (or restrictions). The first hypothesis is that the robot has proprioceptive and exteroceptive sensors that allow collect data about yourself (position and orientation) and about the environment. The second is that perception system must always provide the exact knowledge of the position and orientation of the robot relative to any coordinate system or frame of global fixed reference adopted.

With these assumptions, we can define the robotic mapping as the problem of building a space model of the environment by the robotic computational system, based on data provided by

the perception system. The Figure 1 illustrates the evolution of the process of building a map. In this illustration, x represents the robot pose (position and orientation), z represents the sensory measurements collected over time, and m represents the map being built iteratively at each time interval.

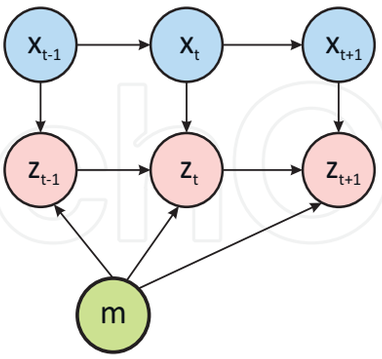

Figure 1. Mapping process.

The hypothesis that the perception system has the exact knowledge of the robot position and robot orientation relative to some global reference adopted is not always true. And in most cases to get the exact position and orientation of the robot inside of your environment is necessary to have a map. Here a conflict arises: to get a consistent map is required exact knowledge of the robot position and orientation, and for this, it is necessary an environment map. Note that there is a dependency between location (inferring the position and orientation of the robot in the environment) and mapping. To resolve this conflict some researchers extend the mapping problem to a simultaneous localization and mapping problem or simply SLAM [2].

From this point, we assume as true the hypothesis that the robot position and orientation are known somehow. Then specifically we approach the mapping problem with known position and known orientation, as proposed by Thrun [6]. It should be noted that even assuming this assumption, the following challenges need to be overcome [1]: sensory inaccuracies, dimensionality of environment, data matching, dynamism of the environment and operating strategy. Then each of these challenges will be detailed.

2.1. Sensory uncertainties

The sensors used for robotic mapping are influenced by several sources of noise that cause errors in his measurements. An important issue in the mapping environments is how to deal with the uncertainties and inaccuracies found in the data provided by these sensors. Inaccuracies can be caused due to many factors intrinsic to the type of device used. In addition, there are also the errors inherent in the robot movements within its environment, which can be caused by systemic factors, such as the uncertainties present in parameters that are part kinematic modeling of the robot (for example, differences in the size of the wheels, erroneous measurement axes) and/or by non systematic factors as uncertainties from unexpected situations (for example, collision with unexpected objects, wheel slip) [7].

The challenge occurs depending on the type of the noise in measurements, in the other words, to model the robotic mapping problem would be relatively easy if the noise in different measurements was statistically independent. However, there is a statistical dependency that

occurs because the errors inherent in the robot movements accumulate over time, and affect how the sensory measurements are interpreted. To overcome this challenge, generally the error sources are modeled or approximated by stochastic functions, allowing the sensory data are handled properly during the mapping process. When these factors are disregarded, we are likely to witness the construction of inconsistent and inaccurate maps.

2.2. Dimensionality

Another challenge that arises in the mapping problem refers to the dimensionality of the environment to be mapped. Imagine an environment typically simple, such as a house with rooms, corridors and kitchen. If it is considered just a topological description of compartments will require a few variables to describe this environment. However, when the goal is to get rich maps in detail with two (2D) or three (3D) dimensions, the complexity to estimate this map is larger and, in addition, it is necessary to increase the space for storing its representation in the computer memory.

Furthermore, cannot forget the larger and more complex the environment, higher is the probability of some kind of error in the robot movements, which can prejudice the quality of built map.

2.3. Matching

During the mapping, normally a same object or obstacle in the environment is perceived several times by the robot, in different moments. Therefore is desirable that this object should be identified as mapped and must be distinguished from those not yet mapped. This problem is known as matching or data association.

The lack of data association in the mapping can generate inconsistencies and differences in maps. An effective scheme of matching must make distinctions between spurious measurements, new measurements and lost measurements together with a basic function to associate the available map with the new measurements. A study of the techniques more used to solve the matching problem can be found in the work of Wijesoma et al. [8], where are presented the general idea of solutions and their respective original references.

A method widely used in matching is the Nearest Neighbor - NN [9]. This method associates a map point to the nearest observation inside a region of validation, based on some distance, which is usually given by the Mahalonobis distance. This method is attractive because the implementation is easy, however it gives poor results. Other robust methods of solving the matching problem can be checked in the literature, for example, the Joint Probabilistic Data Association (JPDA) [10], Joint Compatibility Branch and Bound (JCBB) [11], Multi Hypotheses Tracking (MHT) [12] and "lazy" data association method (a variant of MHT) [13].

2.4. Dynamism of the environment

Another challenge arises when the interest is in mapping dynamic environments or not stationary, such as offices where people travel constantly and manufactures where objects are transported to different places. In these cases, the robotic system can consider such changes as inconsistent measurements taken at a given time. Think about the case where, in a given moment, a robot maps a table in your environment. Then, for some reason, this table is moved

to another location at a later time, and this change of location, which should be changed also on a robot map, does not occur. With this, when the robot go through the place where the table was previously may seem that the robot is in a new location not yet mapped, because the object that should be detected, it is not.

The vast majority of mapping algorithms considers that the process is running in static environments, and this makes the mapping problem of dynamic environments be largely unexplored. However, in recent years, several proposals have emerged to solve this problem. For example, Biswas et al. [14] proposes a mapping algorithm in the occupation grid called Robot Object Mapping Algorithm - ROMA, able to model not stationary environments. The approach assumes that objects move slowly in the environment that can be considered static in relation to time it takes to build a map in occupation grid. The proposed algorithm is able to identify such moving objects, learn the model of them and determine their location at any time. It also estimates the total number of distinct objects in the environment, making the approach applicable to situations where not all objects not stationary are visible at all times. The changes in the environment are detected by a simple technique of differentiation of maps.

Wolf and colleague [15] propose an approach based on differentiation of maps, however the differentiation is made between a map of occupancy grid with static parts, and a map of occupancy grid with dynamic parts. The algorithm proposed is able to detect dynamic objects in the environment and represent them on a map, even when they are outside the robot perceptual field. In a recent work, Baig and collaborators [16] propose the detection of dynamic objects in the mapping process of external environments through the method Detection And Tracking of Moving Objects (DATMO), from a vehicle equipped with laser and odometry.

2.5. Exploration strategy

During the mapping, the robotic system should choose certain paths to be followed. Given that the robot has some knowledge about the world, the central question is: where he should move to get new information? This is done by operating strategy that determines the displacement that the robot should perform to meet all environment. At the process end we have a visited environment map produced from the information provided by the robot sensors.

The choice and implementation of good operating strategy emerges as a fifth challenge for the mapping problem. The exploration assumes that the robot position and robot orientation are precisely known and focuses on bring the robot with efficiency throughout the environment to build the full map [17]. For this, the exploration strategy should consider a model of partially built environment, and from this to plan the movement action to be performed. Stachniss [17] describes the exploration strategy as a combination between the mapping and the planning.

It is important that it be considered the exploration strategy efficiency to avoid unnecessary spending of time and energy. In addition, it is necessary that the robot is able to deal with unexpected situations that may arise during the operation and building of the map. A classic technique of exploration is proposed by Yamauchi [18], which is based on the concept of frontiers (frontier-based exploration), that are regions forming limits between free spaces and unexplored spaces. When the robot moves toward a new frontier, it can understand

unexplored spaces and add new information to your map. As result the mapped territory expands, retreating the limits between known regions and unknown regions. Leaving for successive frontiers, the robot can constantly increase your knowledge of the world. Silva Júnior [19], presents the idea of exploitation based on boundary value problems. Stachniss [17] presents a summary of several exploration techniques developed, comparing your proposal based on coverage maps. In this strategy, the robot exploration actions are selected to provide a reduction of uncertainties that affect the mapping.

2.6. Representation types

There are two main approaches to represent an environment using a mobile robot: the representation based in topological maps and the representation using metric maps [1]. However, some authors prefer change this classification increasing another paradigm called features maps representation [20] [21]. This category, by storing metrics information of the notable objects (features) of the robot environment, is treated here as a metric maps subset. The following are main peculiarities of these paradigms.

2.6.1. Topological maps

Topological maps are computationally represented by the graphs, which describes an arrangement of nodes (or vertices) connected by edges (links or arcs). Typically the graphs describe the free spaces for performing tasks. The nodes correspond to regions that have homogeneous sensory information and the arcs reflect the connection between these regions. Intuitively the use of a graph to describe an environment is a great idea because graph is a compact structure for storage in memory of the robot computational system and can be used to model structures or enumerate processes, such as representation of cities connected by roads, connections of the printed circuit board, and others. The main problem of this representation is the lack of a standard that define which structures are associated with the vertices and which relationships are described as links.

Still, the robot location using topological maps is abstract, this means that, there is no way to define explicitly the robot position and robot orientation. However, it is possible to affirm in which graph node or in which environment regions it is.

The lack of standardization of which elements are considered nodes and edges of graphs is easily seen. For example, Kuipers and colleague [22] uses map nodes to represent places, characterized by sensory data, and the edges represent paths between places, characterized by control strategies. Thrun [23] uses a topological map obtained from a probabilistic occupancy grid partitioned into regions (nodes) separated by close passages (edges). In a recent work [24], the definitions of nodes and edges are made from a topology extraction method based on Generalized Voronoi Diagram, that make the skeletonisation of the images provided by a laser, to produce appropriate topological information.

2.6.2. Metric maps

The metric representation or metric maps reproduce, with a certain accuracy degree, the geometry of the environment in which the robot is inserted. Objects such as walls, obstacles and passages are easily identified in this approach, since the map maintains a good

topographic relationship with the real world. The metric maps are represented by occupancy grids and features maps.

- *Feature-based maps*

The features maps store information of important elements founded in environment specific locations (features). A feature can be understood as "something" easily notable inside the environment such as corners and edges. In images, special properties of the some parts (such as a circle, a line, or a region with particular texture) are usually identified. Lee and colleagues [4] use a robot equipped with a sonar belt to build a features map with lines, points and circular objects. These features are differentiated according with the sonar readings processing.

This does not impede others objects, such as doors, lights, trees, buildings, towers etc, are also used as notable features to be stored in a map. In the work of Guivant [5], for example, a mobile vehicle gives the map of the typical and external environment of a farm, identifying trees as features to be stored in a map.

This map type usually stores geometric information of the chosen features and Detected, as for example, cartesian coordinates or polar coordinates of features relatives to some fixed reference. Santana and Medeiros [7] use floor lines of internal environments as interesting features to be detected. The map was built with the polar coordinates of the straight obtained from imaging by Hough transform. A great advantage of this map type is have a compact representation if compared to occupancy grid representations. However, it also has disadvantages, and the main one is the dependence of a pre-defined procedure to detect and extract environment features [17].

- *Occupancy Grid*

The representation using occupancy grid was initially proposed in 1987 and formalized in 1989 by Alberto Elfes [25][26]. With this representation, the continuous spaces of the environment are discretized, so that the environment is being represented in the configuration of a grid or multi-dimensional matrix (2D or 3D). Each element of the matrix, also called cell, represents a environment location that can be occupied, empty or unexplored.

The occupancy grid representation initially was proposed to map environments in a 2D grid, however, after Elfes [25], Moravec [27] expanded this representation to a 3D discrete configuration innovating also the sensor type used for the construction of the map. Unlike Elfes who employed sonar in his experiments, Moravec used a stereo vision system in the construction of a 3D map. In addition, Moravec also proposed a new approach to indicate the possibility of cell occupancy based on evidence theory of Dempster-Shafer, also differing from the Bayesian probabilistic approach proposed by Elfes. The 3D grid presented was called evidence grid and each cell stores a value that indicates the evidence of occupation.

Another approach of grid representation was presented in the work of Oriolo et al. [28]. In this work, the authors show that it is possible formulate and solve perception problems and planning problems dealing with uncertainties using the set theory (fuzzy logic). The built map by this technique is defined as a fuzzy set, in which each point is associated to a real number that quantifies the possibility of the map pertain to an obstacle. The main advantage is the possibility of using several types of fuzzy operators to model uncertainty and aggregate information from multiple sources. Ribo and Pinz [29] present a comparative

study between the three approaches mentioned, showing their models to the range sensors and checking the treatment of the uncertainties presents in the measurements. However, the cited work is restricted to experiments with robots equipped with sonar inserted in typical office environments.

Other variations of these approaches can be seen in the literature. For example, Konolige [30] proposes a method that is a mathematical refinement of the mapping method presented by Elfes [25], named MUltiple Representation Independent Evidence Log - MURIEL, aiming to control the intrinsic problems of the sonar, such as reflection multiples and redundant readings. Borenstein and Koren [31] presents a method based on histograms. In this method, the main goal is to reduce the computational cost intrinsic to representation based occupancy grids. Another variation can be checked on work Yguel et al. [32]. This work reports the issues of representation and data storage by large maps and, for this, proposes a form of occupancy grid representation based on wavelets: Wavelet Occupancy Grids.

Some changes of the default algorithm presented by Elfes [25] are proposals aiming to improve the quality of built maps, considering computational cost of processing and storage [13], uncertainty treatment [33] and sensors modeling [34]. There are several recent examples in the literature that attest the utility of the representation by occupancy grid in SLAM [36] [38], with applications like surveillance [30] [37], exploration [39], rescue [40], and others tasks.

3. Occupancy grid mapping

The models based on occupancy grid proposed by Elfes [26] is one of the most used metric approaches. The environment is represented as a regular grid of cells, where the value of each cell encodes its state that can be free, occupied, or not mapped (undefined). The occupancy value of a cell is determined using a probabilistic approach that has as input estimated distances to objects calculated from data given by the sensors. Through a bayesian approach, it is possible to update the cell values at every time that a measure is performed. Note that a subset of the whole grid is updated each time. The resulting model can be used directly as a map of the environment in navigation tasks as path planning, obstacle avoidance, and pose estimation. Figure 2 ilustrates the representation of a depth sensor measure in a 2D occupancy grid. Grey cells have unknown occupancy values, white cells are free and black cells are occupied. The main advantages of this method are: it is easy to construct and it can be as accurate as necessary.

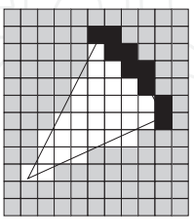

Figure 2. Occupancy grid representation in 2D. All cells of the grid are initialized with gray color meaning that the occupancy value is not known at start. From the readings inside the field of view of the sonar, white cells (free) and black cells (occupied) can be determined.

The construction of a map based on occupancy grid involves the determination of the occupancy probability of each cell. Let us assume that the occupancy probabilities of neighboring cells are decoupled, that is, the probability that a cell is occupied does not affect

the estimation of probability occupancy values for its neighboring cells. Besides this is not an ideal assumption specifically in the case of cells representing the same object it turns easier the implementation of the mapping algorithm without a substantial increasing in the measured error. With this assumption, the probability of occupation of a map M can be factored into the product of the occupancy probability of each cell m_n individually according to Equation 1.

$$P(M|x_{1:t}, z_{1:t}) = \prod_n p(m_n|x_{1:t}, z_{1:t}) \tag{1}$$

The construction of the map M depends on the history of robot localizations $x_{1:t}$ and on the sensors readings $z_{1:t}$ performed in each localization. Updating the map is a process of repeating these readings, which can be performed from other locations with other orientations. In practice, only cells currently inside the field of view of the sensors are to be updated. The occupancy value of a cell m_n is calculated by Equation 2.

$$p(m_n|x_{1:t}, z_{1:t}) = 1 - \frac{1}{1 - e^{l_{t,n}}} \tag{2}$$

with

$$l_{t,n} = l_{t-1,n} + log\frac{p(m_n|x_t, z_t)}{1 - p(m_n)} - log\frac{1 - p(m_n|x_t, z_t)}{p(m_n)} \tag{3}$$

$p(m_n)$ is the occupancy value of the cell m_n previously to any measurement that can be attributed according to the obstacles density in the environment. The probability $p(m_n|x_t, z_t)$ specifies the occupancy probability of cell m_n conditioned to the sensor reading z_t in the time t and depends on the sensor used, and is named inverse sensor model. More details of the standard algorithm for occupancy grid can be found in the work of Thrun [34].

4. Mapping with stereo vision

4.1. Stereo vision

The term stereo vision is generally used in the literature to designate problems where it is necessary to recover real measures of points in a 3D scene from measures performed in their corresponding pixels in two or more images taken from different viewpoints. The determination of the 3D structure of a scene using stereo vision is a well known problem [41]. Formally, a stereo algorithm should solve two problems: the matching (pixel correspondence) and the consequent 3D geometry reconstruction.

The matching problem involves the determination of pairs of pixels, one pixel from each image in each stereo pair, corresponding to the points in the scene that have been projected to the pixels. That is, given a pixel x in the first image, the matching problem is solved by determining its corresponding pixel x' in the second image, for all pixels in the first one. In principle, a search strategy has to be adopted to find the correspondences. Several solutions are proposed generally using restrictions to facilitate the matching. The epipolar geometry is one of these, that makes narrower the search space [41].

By using this restriction, given a pixel in the one image, the search for the corresponding pixel in the other image can be performed in a line thus substantially decreasing the search space (from 2D to 1D). In general, the result of the matching phase is a disparity map that has, for

all corresponding pixels determined, the displacement in images coordinates of each matched pixel in the second image.

The reconstruction of 3D geometry relates to the determination of the scene 3D structures. The 3D position of a point P in space can be calculated by knowing the disparity between positions (in image coordinate system) of the corresponding pair of pixels x and x' given by the disparity map and by knowing the geometry of the stereo vision system. The last is specified in two matrices: M (named external) and M' (internal), which are previously determined. The positions of pixels are used in the matching phase to calculated the disparity map, which is then used directly in the process.

In order to better understand the stereo reconstruction problem, let us assume the system geometry shown in Figure 3. The system has two cameras with projection centers O and O', respectively. The optical axes are perfectly parallel to each other and the two camera images are in the same plane (coplanar). The focal distance f is the distance from each projection center to each image plane, assumed to be the same for both cameras. The baseline b is the distance between the projection centers. The values of b and f are known a priori by using some camera calibration procedure [41].

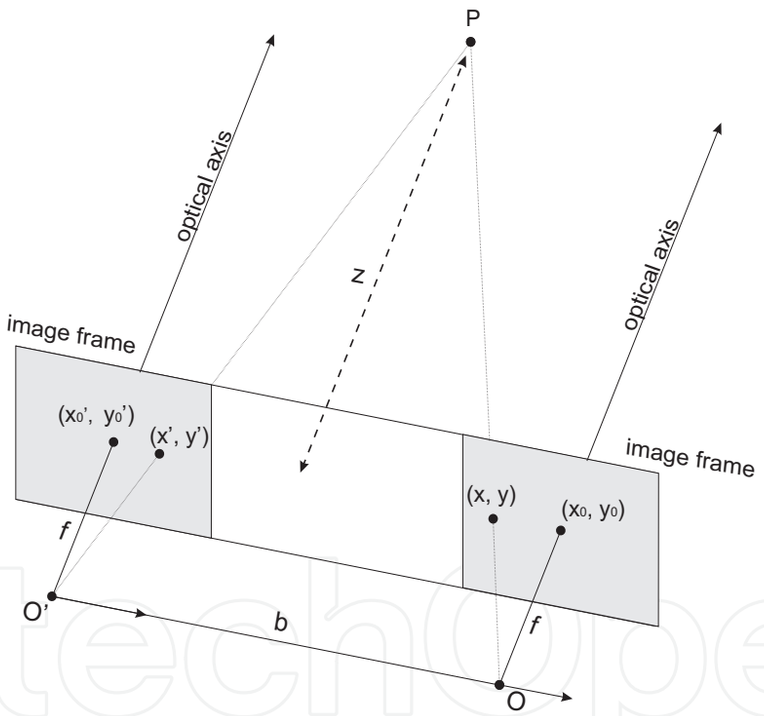

Figure 3. Stereo geometry of a coplanar camera system.

In Figure 3, (x'_o, y'_o) and (x_o, y_o) are the coordinates of the pixels in each image center, that is, in the intersection of the optical axes and the image planes. A point P in the scene is projected in the pixels (x', y') (left) and (x, y) (right) in the images. The distance z_c of the camera system to the point P can be calculated by Equation 4, in camera frame reference:

$$z_c = f.\frac{b}{d} \tag{4}$$

where b is the baseline as said above and d is the disparity between the corresponding pixels, given by $d = x' - x$. Note that there is no disparity in y coordinate because the axes are

parallel. From Equation 3, we can observe that z_c is inversely proportional to d. In practical situations, d is limited considering a maximum and minimum distance of the system to the scene. These distances are empirically defined considering the scene. The other coordinates of point $P_c = (x_c, y_c, z_c)$ in relation to the stereo system can be calculated by using Equations 5 and 6.

$$x_c = z_c.\frac{(x - x0)}{f} \tag{5}$$

$$y_c = z_c.\frac{(y - y0)}{f} \tag{6}$$

The coordinates of a point in the camera coordinate system P_c can be transformed to world frame. To do that, one just has to know the rotation matrix and translation vector that map the camera to world frame, and use Equation 7.

$$P_w = R^T.P_c + T \tag{7}$$

Parameters b, f , (x'_o, y'_o) e (x_o, y_o) are determined through a previous calibration of the stereo system. we remark that even if the system does not follow the restrictions depicted in Figure 3, it is possible to derive a general mathematical model for the problem. In this case, it would be more complex to recover the 3D geometry. Besides, in the calibration process it is also possible to diminish or eliminate errors caused by lens distortions and illumination.

Stereo matching The stereo matching should determine, for all pixels in one image, the pixels that are their homologous in the other image. That is, to determine all pairs of pixels, one pixel from each image, that correspond to the projections of points in the scene. Once determined the matching, disparity d can be calculated as the difference between the coordinates of the corresponding pixels in each image and the depth for all points in the scene can be calculated by using triangle similarity. So determination of matching for all pixels is a fundamental step. In fact, due to occlusions and image errors, we get not completely dense matchings.

Between the several methods used for matching and disparity map calculation, in this work we performed experiments with the two suggested in the work of Scharstein and Szeliski [42]: the block-matching and the graph-cuts. The block-matching determines the correspondence for pairs of pixels with characteristics that are well discernible. Basically, this method gives as result a disparity map in four main steps. The first is a previous filtering of the images to normalize brightness and texture enhancement. The second step is the search for correspondences using the epipolar constraint. This step performs the summation of absolute differences between windows of the same size of both images to find the matching. As the third step, a posterior filtering is performed in order to eliminate false matchings. In the fourth step, the disparity map is calculated for the trustable pixels. This method is considered fast, in fact, its computational complexity depends only on the image size.

The graph-cuts algorithm treats the best matching problem as a problem of energy minimization, including two energy terms. The smoothness energy (SE) measures the smoothness of disparity between neighboring pixels, which should be as smooth as possible. The second term is data energy (ED) that measures how divergent are corresponding pixels with basis on assumed disparity. A weighted graph is constructed with vertices representing the image pixels. The labels (or terminals) are all the possible disparities (or all discrete values in the interval of disparity variation). The edge weights correspond to the energy terms above defined. The graph cuts technique is used to approximate the optimal solution, that computes

corresponding disparity values (each edge of the graph) to each pixel (vertex of the graph) [43].

4.2. Modeling probabilistic mapping with stereo disparity

The acquisition and manipulation of images produces a uncertainty Δd for the error in the coordinates of a pixel. This causes an error factor Δz_c in the estimation of the coordinate z_c calculated from the disparity map, named depth resolution, which is given by Equation 8.

$$\Delta z_c = \frac{z_c^2}{b.f}.\Delta d \tag{8}$$

In order to construct the environment model using occupancy grid, it is necessary to model the used sensor generally named the inverse measurement model. We adopt a modeling that is similar to the one depicted by Andert [44], however we propose to incorporate the errors inherent to robot motion in the calculation of the probability of occupation of a cell. With this modification, we get a map that is more coherent with the sensory data provided by the robot. We can guarantee a limit for the maximum error found in the map. This is extremely important in the treatment of uncertainties encountered in the scene.

The inverse model transform the data provided by the sensors in the information contained in the map. In our case, a distance measured from a specific point in the environment indicates the probability that part of the grid is occupied or free. Distance is calculated directly from disparity. In this way, it is possible to map each point of the space seen by the robot vision system with a valid disparity value into the possible values of elements of the occupancy grid.

In the measurement model, each point of the disparity map can act as a distance sensor measured along the ray defined by the world coordinates of the pixel (in the image plane) and the projection center of the camera. The point in the image plane is given in camera coordinates by $Pc = (x_c, y_c, z_c)$ and the distance of the camera to P_c is calculated by Equation 9 with a sensorial uncertainty given by Equation 10.

$$l_p = \sqrt{x_c^2 + y_c^2 + z_c^2} \tag{9}$$

$$\Delta l_p = \Delta z_c \frac{l_p}{z_c} \tag{10}$$

Each cell pertaining to this ray defined by the projection center and point P_w also in world coordinates must have a probability of occupation updated according to the function of density of probability given by Equation 11:

$$P(m_n|x_{1:t}, z_{1:t}) = P_{occ}(l) + \left(\frac{k}{\Delta l_p \sqrt{2\pi}} + 0.5 - P_{occ}(l) \right) e^{-\frac{1}{2}(\frac{l-l_p}{\Delta l_p})^2} \tag{11}$$

where

$$P_{occ}(l) = \begin{cases} p_{min}, & \text{if } 0 < l \leq l_p \\ 0.5, & \text{if } l > l_p \end{cases} \tag{12}$$

Our proposal is just to incorporate the uncertainty that we have with respect to the robot motion in the calculation of the occupancy probability. With this, the occupancy grid map

gets more coherency with the sensory data acquired by the robot perceptual system. In this way, Equation 8 can be modified resulting in Equation 13:

$$\Delta z_c = \frac{z_c^2}{b.f}.\Delta d + \varepsilon \tag{13}$$

where ε is a function that describe the errors of the linear movements of the robot (systematic errors) that are modeled from repetitive experiments performed previously. This function ε represents the degradation that the motion errors of the robot produce in the certainty that a cell is occupied or not, like in a previous work [45]. So we consider the existence of real uncertainties in the robotic systems.

5. Preliminary experiments

We have performed experiments in order to evaluate and decide the better stereo matching algorithm to be used for the disparity map generation. In these experiments, we use the Minoru 3D stereo vision system mounted on the top of a Pioneer 3-AT robot, as seen in Figure 4. The implementation is done using the Computer Vision OpenCV library.

Figure 4. Pioneer 3-AT robot with stereo cameras Minoru.

The next set of experiments is done in order to evaluate the use of stereo disparity as input to the probabilistic approach in the occupancy grid construction (our main proposal). The stereo setup used is the same mounted on the Pioneer 3-AT equipped with the Minoru 3D. The OpenCV library is used for software developments. As said, the disparity map is estimated using the *graph-cuts* algorithm. The experiments were conducted in a building of Federal University of Rio Grande do Norte, which mixes scenes of outdoor and indoor environments. Figure 5 shows one of the images captured from an scene of a typical corridor and the Figure 6 illustrates a scene with the appearance of the external environment.

Figure 7 shows the result of this experiment. It is shown only one of the slices of the cells of the occupancy grid, the one that has the plane in the center of projection of the camera system parallel to the ground. We remark that the robot could perform stops to better determining some regions with more details.

From these preliminary results of the 2D mapping, the first steps were made to perform 3D mapping. In this experiment, the figure was swept in both coordinate axes so that the 3D information could be inferred. The figures below illustrate a basic experiment of this construction.

In this figure, the red polygon represents the camera and the blue cubes represent the obstacle verified by stereo image processing.

Figure 5. Scene of a corridor.

Figure 6. Scene of outdoor space.

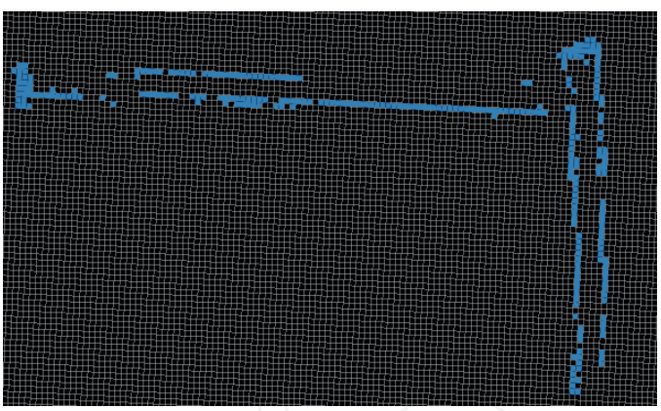

Figure 7. Resulting map.

Figure 8. Captured scene to 3D mapping.

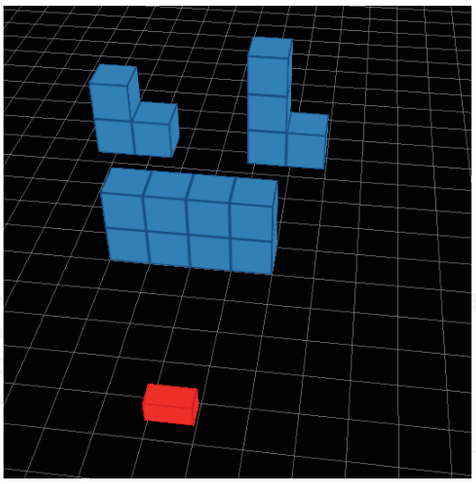

Figure 9. 3D map from stereo vision.

6. Conclusions and future works

We propose a new approach to 3D mapping of environments by a mobile robots with representation based on a probabilistic occupancy grid. Our approach considers using the errors inherent to the robot. Visual information provided by a stereo vision system is included in the modeling. This has resulted in a more robust technique where the error can be well defined and limited, what is very relevant in robotics applications. With our proposal, we have the mapping of environments actually coherent with sensory data provided by the robotic perceptual system. This is one of the main contributions of our work, besides the manner of using stereo vision for 3D mapping using occupancy grid. A 3D representation is much more reliable than a 2D one provided by using only sonar, like in our previous work [45].

As future work, we will go perform experimentation in more complex environments in order to test the coherent construction of 3D occupancy grids. A strategy based on stop points and feature detection (mainly using gradient of disparity) is also being developed in order to determine regions in which a more detailed map is necessary. Further, a more robust modeling of the errors present in the robot movements and on the image processing will be performed in order for this proposal to upgrade to a SLAM application (Simultaneous Localization and Mapping).

Acknowledgements

We thank the University of the State of the Rio Grande do Norte to help in the developing work.

Author details

Anderson A. S. Souza and Rosiery S. Maia
Department of Informatics, University of the State of the Rio Grande do Norte, Natal, Brazil

Luiz M.G. Gonçalves
Department of Computer Engineering, Federal University of the Rio Grande do Norte, Natal, Brazil

7. References

[1] S. Thrun (2002). Robotic mapping: A survey. In Lakemeyer, G. and Nebel, B., editors, Exploring Artificial Intelligence in the New Millenium. Morgan Kaufmann.

[2] G. Grisetti, C. Stachniss, W. Burgard (2007) Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filter, IEEE Trans. on Robotics, Vol. 23.

[3] Angelopoulou E., Hong T., Wu A. (1992) World Model Representation for Mobile Robots, In Proc. of Intelligent Vehicles Š92 Symposium, 293-297.

[4] S. Lee, B. Park, J. Lim, D. Cho (2010) Feature map management for mobile robots in dynamic environments,Robotica, Vol.28, Issue 1, 97-106, Cambridge University Press, 2010.

[5] J. Guivant, F. Masson, E. Nebot (2002) Simultaneous localization and map building using natural features and absolute information, Journal of Robotics and Autonomous Systems, Vol. 40, 79-90.

[6] S. Thrun, D. Fox, W. Burgard (2005) Probabilistic Robotics, MIT Press, Cambridge, MA, 2005.

[7] A. M. Santana, A. A. Medeiros (2009) Simultaneous Localization and Mapping (SLAM) of a Mobile Robot Based on Fusion of Odometry and Visual Data Using Extended Kalman Filter, chapter CONTEMPORARYROBOTICS - Challenges and Solutions, pages 129Ű 146. Intech, Austria.

[8] W. Wijesoma, L. Perera, M. Adams (2006) Toward Multidimensional Assignment Data Association in Robot Localization and Mapping, IEEE Transactions on Robotics, Vol. 22, Issue 2, 350-365.

[9] M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, M. Csorba (2001) A solution to the simultaneous localization and map building (SLAM) problem, IEEE Transaction on Robotics and Automation, Vol. 17, 229-241.

[10] J. Dezert, Y. Bar-Shalom (1993) Joint probabilistic data association for autonomous navigation, IEEE Transaction on Aerospace and Electronic Systems, Vol. 29, Issue 4, 1275-1286.

[11] J. Neira, J. Tardos (2001) Data association in stochastic mapping using the joint compatibility test, IEEE Transaction on Robotics Automation, Vol. 17, 890-897.

[12] D. Maksarov, H. Durrant-Whyte (1995) Mobile vehicle navigation in unknown environments: A multiple hypothesis approach, In Proc. of the IEEE Control Theory and Applications, Vol. 142, Issue 4, 385-400.

[13] D. Hähnel, W. Burgard, S. Thrun (2003) Learning compact 3D models of indoor and outdoor environments with a mobile robot, Robotics and Autonomous Systems, Vol. 44, 15-17.

[14] R. Biswas, B. Limketkai, S. Sanner, S. Thrun (2002) Towards object mapping in non-stationary environments with mobile robots, In. Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Vol.1, 1014-1019.

[15] D. Wolf, G. Sukhatme (2004) Online Simultaneous Localization and Mapping in Dynamic Environments, In Proc. of the IEEE Int. Conf. on Robotics and Automation, Vol.2, 1301-1307.

[16] Q. Baig, T.-D. Vu, O. Aycard (2009) Online localization and mapping with moving objects detection in dynamic outdoor environments, In Proc. of the IEEE 5th Int. Conf. on Intelligent Computer Communication and Processing, 401-408.

[17] C. Stachniss (2009) Robotic Mapping and Exploration, Springer Tracts in Advanced Robotics, Vol. 55.

[18] B. Yamauchi (1997) A frontier-based approach for autonomous exploration, In Proc. of IEEE Int. Symposium on Computational Intelligence in Robotics and Automation, 146-151.
[19] E. Silva Júnior (2003) Navegação exploratória baseada em Problemas de Valores de Contorno, PhD Thesis, Universidade Federal do Rio Grande do Sul - UFRGS.
[20] H. Choset, D. Fox (2004) The World of Mapping, In Proc. of the Workshop on Review of United States Research in Robotics, National Science Foundation (NSF).
[21] R. Rocha (2006) Building Volumetric Maps whit Cooperative Mobile Robots and Useful Information Sharing: A Distributed Control Approach based on Entropy, PhD Thesis, Faculdade de Engenharia da Universidade do Porto, Portugal.
[22] B. Kuipers, Y.-T. Byun (1991) A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations, Robotics and Autonomous Systems, Vol. 8, 47-63.
[23] S. Thrun (1998) Learning metric-topological maps for indoor mobile robot navigation, Artificial Inteligence, vol. 99, Issue 1, 21-71.
[24] H. Cheong, S. Park, S.-K. Park (2008) Topological Map Building and Exploration Based on Concave Nodes, In Proc. of the Int. Conf. on Control, Automation and Systems, 1115-1120.
[25] A. Elfes (1987) Sonar-based real-world mapping and navigation, IEEE Journal of Robotics and Automation, Vol. 3, Issue 3, 249-265.
[26] A. Elfes (1989) Occupancy Grid: A Probabilistic Framework for Robot Perception and Navigation, PhD Thesis, Carnegie Mellon University, Pensylvania, USA.
[27] H. Moravec (1996) Robot Spatial Perception by Stereoscopic Vision and 3D Evidence Grids, CMU Robotics Institute Technical Report, Daimler Benz Research.
[28] G. Oriolo, G. Ulivi, M. Vendittelli (1997) Fuzzy maps: a new tool for mobile robot perception and planning, Robotics Systems, Vol. 14, Issue 3, 179-197.
[29] M. Ribo, A. Pinz (2001) A comparison of three uncertainty calculi for building sonar-based occupancy grids, Robotics and Autonomous Systems, Vol.31.
[30] K. Konolige (1997) Improved occupancy grids for map building, Autonomous Robots Vol.4, 351-367, 1997.
[31] J. Borenstein, Y. Koren (1997) The vector Field histogram - fast obstacle avoidance for mobile robots, IEEE Journal of Robotics and Automation, Vol.7.
[32] M. Yguel, O. Aycard, C. Laugier (2006) Wavelet Occupancy Grids: a Method for Compact Map Building, P. Corke Fields and S. Sukkarieh (Eds.): Field and Services Robotics, 219-230, Springer-Verlag Berlin Heidelberg.
[33] E. Ivanjko, I. Petrovic (2005) Experimental evaluation of occupancy grids maps improvement by sonar data correction, In Proc. of the 13th Mediterranean Conference on Control and Automation, Limassol, Cyprus.
[34] S. Thrun (2003) Learning occupancy grid maps with forward sensor models, Autonomous Robots, Vol.15, 111-127.
[35] D. Hähnel, W.Burgard, S. Thrun (2003) Learning compact 3D models of indoor and outdoor environments with a mobile robot, Robotics and Autonomous Systems, Vol. 44, 15-17.
[36] O. Özisik, S. Yavuz (2008) An Occupancy Grid Based SLAM Method, In Proc. of the IEEE Int. Conf. on Computational Intelligence for Measurement Systems And Applications, 117-119.

[37] A. Kolling, S. Carpin (2008) Extracting surveillance graphs from robot maps, In Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, 2323-2328.
[38] T. Marks, A. Howard, M. Bajracharya, G. Cotrell, L. Mathies (2009) Gamma-SLAM: Visual SLAM in Unstructured Environments Using Variance Grid Maps, Journal of Field Robotics, Vol.26, Issue 1, Wiley, 26-51.
[39] R. Sim, J. Little (2006) Autonomous vision-based exploration and mapping using hybrid maps and Rao-Blackwellised particle filters, In Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, 2082 -2089.
[40] A. Birk, S. Carpin (2006) Rescue robotics - a crucial milestone on the road to autonomous systems, Advanced Robotics, Vol.20, Issue 5, 595-605, 2006.
[41] E. Trucco, A. Verri (1998) Introductory Techniques for 3D Computer Vision. Prentice Hall.
[42] D. Scharstein, R. Szeliski (2002) A taxonomy an evaluation of dense two-frame stereo correspondence algorithms, International Journal of Computer Vision, vol. 47, 7Ű42.
[43] L. Hong, G. Chen (2004) Segment-based stereo matching using graph cuts, In Conference on Computer Vision and Pattern Recognition, vol. 1, 74Ű81.
[44] F. Andert (2009) Drawing stereo disparity images into occupancy grids: Measurement model and fast implementation.Ť in IEEE/RSJ International Conference on Intelligent Robots and Systems, 5191Ű519.
[45] A. Souza, A. Santana, A. Medeiros, L. Gonçalves (2010) Probabilistic Mapping by Fusion of Range-Finders Sensors and Odometry, In Sensor Fusion and Its Applications, Ed.Ciza Thomas, Sciyo.