Self-organizing Mapping of Robotic Environments Based on Neural Networks

Mônica Figueiredo, Silvia Botelho, Paulo Drews and Celina Haffele

Centro de Ciências Computacionais - C3
Universidade Federal do Rio Grande - FURG
Rio Grande - RS - Brazil
monicafigueiredo@furg.br, silviacb@furg.br, paulodrews@furg.br, celinahaffele@yahoo.com.br

Abstract—An important aspect in robotics is mapping environments, it means know free space configuration of the robot and establish landmarks. It enables the robot to calculate its position in the environment. In this context, this paper proposes a method for mapping generic environments (structured or not) based on topological maps (SOM and Growing Cell Structures), which uses self-organizing networks. The results obtained on different dynamic and ambiguous environment demonstrate both generalization and compactness.

Keywords-Hybrid Maps; Topological Mapping; Neural Networks

I. INTRODUCTION

Nowadays, one of the most challenging task for mobile robots is navigate in a complex environment. When performing missions, these vehicles should be able to deal with a wide variety of dynamic and uncertain informations. They also interact with humans beings and, possibly, other robotic systems, and ensure their own safety as well as the people around them [1].

For this purpose, estimating their location, mapping the navigable environment and planning its trajectory are basic preconditions for the conception of autonomous robotic systems. The mapping provides a representation of the environment based on data acquired by the sensors. The process of estimating the pose of a robot is called localization. Finally, the path planning is responsible for sending commands to the robot motors providing a safe navigation, avoiding obstacles, to a determined place [2].

The mapping has been an area widely exploited in robotics and artificial intelligence for at least two decades [3]. The quality of the representative model obtained depends on the characteristics of the environment and sensors. There are several issues that are involved, making it a complex problem, including [3]:

- The high dimensionality of the entities, \( n \), to be mapped: The complexity of the map to be generated depends on the detail level chosen to represent the environment, perhaps having to deal with millions of descriptive elements;
- The correspondence problem: Also known as the problem of data association, which seeks to obtain different representation models to match similar perceptions found in different spatial locations.
- Dynamics of the environment: In dynamic environments, the robots need to deal with changes, where it should be able to detect and represent them.
- Uncertainty of the sensors: The sensors are not completely accurate, the perceived values can be discrepant in relation to the true state of the environment.

In the literature, different mapping approaches are proposed to deal with the mentioned issues. Probabilistic techniques associate the problem of mapping to the localization in order to minimize the cumulative effects of localization imprecision when building the map, and vice versa. During mapping stage the robot must be able to synthesize and explore the map. The map synthesis, stage at which this study focuses, is characterized by building a model that represents the environment, using sensory information and their locations. On the exploration, the robot navigates in the environment to obtain sufficient information to enable its modeling [5],[6],[7],[8].

Bio-inspired methods seek to develop representation models that provide subsidies for exploration and record of the navigable environment without the need for accurate descriptions of their components [9],[10],[2],[11] and able to represent different abstraction levels necessary to missions of different natures in time. Topological maps based on Self-Organizing Networks [12] (Self-organized Maps - SOM) can be used for this goal. These networks allow to reduce the dimensionality of the data from dimension \( m \) to a smaller one, \( n \). Two-dimensional grids (\( n = 2 \)) topologically gather perceptions acquired by the robot, providing representative maps of the navigable environment.

The typical capacities of SOMs, as topological representation, adaptation, generalization and noise tolerance, are relevant features for its use. The complexity of values, due of scalability and diversity of multi-sensors, and the difficulty in relating the obtained perceptual maps with spatial information are aspects that may be suitable for treatment by a SOM. Recent studies using SOMs explore different semantic levels of representation [10], sharing information to be mapped at different levels of abstraction, ranging from low-level spacial geometric information to high-level context dependent cognitive information [9],[10].

The main objective of this paper is to propose an original method for robotic mapping based on different levels of Self-
Organizing maps. In order to obtain representative models of navigable environments by a mobile robots, SOMs are used in different semantic levels for the treatment of sensory information provided by different sensors. This method should be able to handle, at least partially, with problems as scalability, correspondence, presence of noise, dynamicity of the environment and exploration strategies.

II. METHODOLOGY

Requirements: Given the complexity of the mapping problem, the proposal seeks to deal with the following issues:

- Measurement uncertainties: Probabilistic dependence of the measurements leads to incremental errors relating to the spatiality of sensory impressions;
- Scalability: Allows its tractability in terms of processing time and storage space;
- Correspondence: Dealing with similar aspects of perception over the environment, detecting correspondences and optimizing them;
- Dynamicity: The architecture must provide plasticity throughout the building process of the maps, being tolerant to possible modifications on the environment.

A Conceptual Proposal: The mapping problem can be summarized as getting a representation to describe the list of obtained perceptions and its location in space. Assuming the robot has a certain degree of autonomy, these perceptions will be useful to missions that involve low-level tasks, such as control its actuators in order to perform a given path, and even cognitive high-level tasks, e.g. choose the way between two different colors corridors.

This proposal aims to present an architecture able to work with the different cognitive levels associated with environments mapping. The environment is described by a set of maps. These maps individually represent the different elements and cognitive-degree perceived, associated with the context where the robot navigates. The different relationships between these maps are represented by dynamic edges.

The Developed Architecture: Due to the nature of information and current sensor technologies, there are two dimensions of perceptual information processing: i. association with a set of impressions, ii. its pose in space navigated by the robot. The different impressions are handled by individual representations, which shall be composed by a specific type of representation associated with the spatiality of navigated environment (metric information). Due to the adaptive features of dimensional reduction, the SOMs [12] are chosen as the framework for building the maps, and their purpose is simple: integrate similar information. Thus, is proposed a system capable of providing different descriptive maps of the environment, from different sensory sources using SOMs. In order to integrate metrics information with topological abstractions obtained from different perceptions of the environment, Figure 1 presents an overview of the proposed system. It divides the mapping problem in two categories:

Perceptual Maps (Mp): This structure is responsible by the registration, in each abstraction level, of the perceptive information using self-organizing maps. Establishes that each sensory source, $Si$, will have a correspond to a perceptual map $Mpi$. Each node on the map represents a possible perception, in a way that the resulting map $Mpi$ register the universe of possible sensory impressions from that sensory data. It allows the establishment of perceptual groups representing the various elements that compose the environment. Different Perceptual Maps may exist depending on the distinct perceptive natures (size and significance), allowing a multi-sensory representation. Perceptual Maps may also represent different layers of information leading to cognitive representations of higher level abstraction. For example, a first level would be represent descriptors for objects belonging to the scene (corners, spots), and a second level would represent only the set of objects in the environment. Thus, each perceptive map $Mpi$ is associated with a own semantic meaning (sensor type, level of information processing, etc.). The main parameters that needed to be defined in the perceptual maps are: its size ($n \times n$ neurons), the learning rate, which influences the neuron weights adaptation, and the neighbourhood width, which can change the number of neighbour neurons that is adapted.

Spatial Map (Me): This structure is responsible by the registration of spatial information associated with the configuration space navigable by the robot, storing its location. Each node in the map represents the spatial surrounding associated to sensory impressions represented in perceptual maps. It is based in Growing Self-Organizing approach [13], since it is not possible to determine a priori the number of nodes in the map. These representations are associated with the location in the space, where the impressions are located. In addition to addressing the issue of perceptions spatiality, the Space Map serves as an integrator between different Perceptual Maps in the system. The parameters defined to

![Figure 1. The different perceptual information recorded in each perceptual map $Mpi$ are associated with different spatial regions in the spatial map $Me$ by edges.](image-url)
build this map are the learning rate, which influences the neuron weights adaptation, and the insertion threshold of a new node, which verifies the maximum error of the winner neuron in order to insert a new node.

The map association is provided between the spatial information and the different perceptions obtained at the same instant of time, i.e., it represents the perception of a given location. The map is associated in a way that perceptual ambiguities are connected to different nodes of the spatial map, while localization difficulties can be solved using the perceptual information.

The presented system is based on commonly used sensors in terrestrial robotics as an omnidirectional video camera, besides position estimation based on GPS data. However, the proposal is generic enough in a context that allows the use of the architecture for mapping with other sensory devices. The following describes in details the implementation of spatial and perceptual maps.

A. These maps are composed by two steps described below, as feature extraction and map build.

Feature Extraction: Different sensors provide sensory information that are used for mapping by autonomous robots. Generally, dealing with such information, of high dimension \( m \), is a complex problem to be solved. The increased scalability of the system, arising from every small change in robot path, requires the use of selection techniques of relevant information in front of the whole perceptions set captured by the robot. Initially, for the selection of relevant information, a Feature Extraction module is used for dimensional reduction, showed in Figure 2. Thus, each perceptual map \( M_{pi} \) is associated to \( n \) dimensions for representation of descriptors of sensor information.

![Figure 2. Dimensional Reduction of Feature Extraction. A different implementation is used for each sensor.](image)

In this work, only one perceptual map is used and it represents the data from a omnidirectional camera. Figure 3 shows a example of the feature extractor, as described in [9]. First, omnidirectional images of the environment explored by the robot are obtained, after these images are rectified by turning them into panoramic [14], as shown in the step A in Figure 3. In the process, shown at B in the figure, each captured image is comprised of \( k \) pixels grouped into a grid of \( t \) rows and \( q \) columns. For each \( q \) column of an image, we compute the average of its \( t \) pixels. Then, at C, is applied a low pass filter for smoothing through a window of \( x \) elements. Finally, as shown in the step D, there is a re-sampling to reduce the size of the descriptor for a number of samples \( y \), which serve as input to the perceptual map.

![Figure 3. Feature Extraction Process [9].](image)

Other features extractors of images could be used as the SURF [15], SIFT [16] or saliency model [17].

In order to address the dynamics and redundancy associated to sensory informations, is proposed the implementation of the perceptual maps \( M_{pi} \), which receive entries \( S_i \) from its \( i \) sensor. Each node of the map have the same dimension of the input descriptors, i.e. size \( y \).

The dynamic adaptation of the map, facing each new perceptual entry is established through the updating of the weights of each node and its neighbourhood, see [12].

B. Implementing Spatial Maps

Information associated with the location of the robot in the configuration space are handled by the Space Map. As the Perceptual Maps case, the Spatial Map is implemented through SOMs, but using the approach of Growing Cell Structures (GCS) [13]. This implementation was chosen because we do not know, a priori, the size of the spatial map. The whole learning process of the GCS, and the removal process, could be found in [13].

The GCS nodes represent different poses of dimension \( p \) in the configuration space of the robot. Thus, in each round, information \( G_p \) is used as an input to the spatial map in order to adapt the nodes to better represent the spatial information presented. The acquired data representing the spatial information is compared to all nodes existing in space map. Over these nodes is elected a winner (BMU), whose Euclidean distance is shorter when compared with the input, i.e., the node that has the closest similarity to the information presented. The adaptation occurs through two operators:

- **create _ node** After being elected a winner, the Euclidean distance is compared with a threshold. If it is greater than this threshold, a new node is created to represent the new information, since the existing ones are not able to represent it. This new node will have connections to the two nodes closest to him, since the triangular topology of the network should be maintained.
- **adapt _ node** If the Euclidean distance between the BMU and input is less than the threshold set, that means there is a node similar enough to represent that input.
The chosen node is then adapted to fit the information presented.

In this work, the operations with cognitive connections remains opened. It will be focused on future works.

III. Tests and Results

The proposed work was validated using a mobile robots, equipped with sensors, during its navigation in different environments. For this paper, the tests were conducted using the physical simulator Webots (http://www.cyberbotics.com/), which allows physically realistic simulations in 3D configurable environments. Among the models of robots, Webots simulator offers the Pioneer robot, which was used in the tests. A keyboard control system of navigation, that allows remote control of the vehicle, was developed.

During the tests, two sensors are embedded in the robot: a camera and a GPS, to acquire the information from the environment. The pose information obtained from the GPS is the input of the spatial map, while the processed visual information supplies a perceptual map. The omnidirectional camera captures images with resolution $128 \times 128$ pixels, see Figure 4(a). The images are rectified, resulting in panoramic images, according to section II-A. The rectified images are turned into entries of descriptors extraction module see Figure 4(b). In the examples, each image is subdivided into 64 rows and 400 columns, which are represented by a descriptor vector with 16 values. Each array element describes the average value of pixels of the column [10], after passed through a smoothing of a low pass filter with a window of seven elements. The information acquired by GPS is used as entry in the spatial map, since it provides the position $(x, y)$ of the robot.

1) Calibration of Parameters - Baseline Scenario: A basic initial environment was developed for calibration of the system parameters, as shown in Figure 5. Quantization Error and Topographic Error $^1$ were used to evaluate the performance of the generated maps. Based on these two metrics, the parameters obtained are: a $50 \times 50$ neurons SOM representing the perceptual map (corresponding to camera data), with learning rate $n = 0.5$ and width of the neighborhood $= 8$. The parameters in spatial map are: the learning rate $n = 0.5$ and insertion threshold $= 0.005$.

Based on these parameters, were found the spatial and perceptual representations shown in Figure 6(a). The dark map, in the top, represents the perceptual information using the unified distance matrix [18]. The map shown in the bottom of this Figure 6(a) represents the GCS. In this scenario, seven perceptual groups were formed. Some location, which have similar perceptive information, are shown in the upper right corner of the spatial map. Due to its similarity, the nodes are connected to the same perceptual group. Figure 6(b) corresponds to the same map of Figure 6(a), but here the connections were removed and the association between the maps was done through colors. Each spatial location can be connected to more than one perceptual group, due the rotational variance. The color of this node is determined by its main perceptual group (the one with small quantization error). It can be seen by comparing the path taken in Figure 5 that perceptions significantly different are equally represented. The beginning (magenta) and end of the map (green and yellow) are represented by the system as well as similar perceptions (green), at its bifurcation. The system performance is shown in Table I.

2) Matching Scenario: The scenario of Figure 7(a) was developed in order to verify the system performance against the problem of the correspondence. By presenting a uniform arrangement of similar obstacles, this scenario is favorable

\[ ^1 \text{Quantization Error (QE)}- \text{Provides a measure of error associated with how an input pattern is well represented by its winner neuron (BMU) - and the Topographic Error (TE)} - \text{Continuity measure of mapping, i.e., if the two closest neurons of the entry are adjacent, the mapping is locally continuous, if not adjacent, there is a local topographic error. The TE for all mapping is obtained summing the local TE for all entries [4].} \]
(a) Connections between perceptive groups and the spatial map

(b) Space nodes can be connected to more than one perceptual group. The node color is determined by the group in which the perceptual information is closest to the node original input.

Figure 6. Results of Calibration Scenario.

Table I

<table>
<thead>
<tr>
<th>Maps</th>
<th>QE</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial</td>
<td>0.004784</td>
<td>0.029737</td>
</tr>
<tr>
<td>Perceptual</td>
<td>0.385198</td>
<td>0.397015</td>
</tr>
</tbody>
</table>

Table II

<table>
<thead>
<tr>
<th>Map</th>
<th>QE</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial</td>
<td>0.013316</td>
<td>0.029752</td>
</tr>
<tr>
<td>Perceptual</td>
<td>0.416541</td>
<td>0.345435</td>
</tr>
</tbody>
</table>

3) Dynamic environments Scenario: This environment is identical to the baseline scenario, Figure 7(a). A total of 605 sensory samples were taken, which were represented in 480 spatial nodes. As expected, the system compresses the similar amount of insights into a small number of groups, 4 principal groups, in perceptive map, as shown by the colors of spatial nodes in Figure 7(b). The system performance is shown in Table II.

Figure 7. Results of Matching Scenario.

Table III

<table>
<thead>
<tr>
<th>Map</th>
<th>QE</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial</td>
<td>0.015236</td>
<td>0.126482</td>
</tr>
<tr>
<td>Perceptual</td>
<td>0.517431</td>
<td>0.464427</td>
</tr>
</tbody>
</table>
and the use of more sensor data may also be conducted. Another future direction is improve the spatial maps using bio-inspired location system using grid and place cells.

REFERENCES