Using Global Appearance Descriptors to Solve Topological Visual SLAM

Lorenzo Fernández Rojo
Miguel Hernandez University, Spain

Luis Paya
Miguel Hernández University, Spain

Francisco Amoros
Miguel Hernandez University, Spain

Oscar Reinoso
Miguel Hernandez University, Spain

INTRODUCTION

Nowadays, the use of mobile robots has extended to many different environments, where they have to move autonomously to fulfill an assigned task. With this aim, it is necessary that the robot builds a model of the environment and estimates its position using this model. These two problems are often faced simultaneously. This process is known as SLAM (Simultaneous Localization and Mapping) and is very common since when a robot begins moving in a previously unknown environment it must start generating a model from the scratch while it estimates its position simultaneously.

To carry out the SLAM process, the robot has to make use of the different sensors it may be equipped with (e.g. odometry, touch, laser, cameras, etc.). During the last years, the use of cameras has extended thanks to the amount of information they can capture from the environment and their relatively low cost. This work is focused on the use of computer vision.

When a vision sensor is used to solve the mapping and localization tasks, we must take into account that scenes contain a great quantity of information. This way, it is necessary to extract the most relevant information from the scenes as it will allow us to work with a lower number of components. This problem can be approach from two points of view: local and global appearance methods. First, the methods based on local appearance consist in extracting and describing prominent landmarks or regions from the scenes. These methods typically need more computational time to build the map and estimate the position of the robot since it is necessary to extract the distinctive features from each image, describe them and make a complete comparison with the data stored in the map. Second, global-appearance methods describe each scene with a unique descriptor that contains information of the whole appearance. These methods tend to be computationally more efficient.

The SLAM problem can be approached from three different points of view: metric, topological and hybrid metric-topological SLAM. First, the metric approach consists in representing the position of some landmarks in the environment with geometric accuracy, with respect to a reference system. Using this approach, the position of the robot can be estimated with respect to this system. However, these methods often present a high computational cost. Second, the topological approach represents typically the environment as a graph, where the nodes represent prominent localizations (e.g. rooms) and links are the connectivity

DOI: 10.4018/978-1-5225-2255-3.ch597

relations between areas. These approaches offer a more compact representation that supposes a lower computational cost, but the localization process is inherently less accurate. At last, some researchers have started working on hybrid approaches that try to combine the advantages of the metric and the topological approaches.

The main objective of this work is to develop and test an algorithm to solve the SLAM problem using two sources of information: (a) the global appearance of omnidirectional images captured by a camera mounted on the mobile robot and (b) the robot internal odometry. We have decided to use a hybrid metric-topological approach to solve the SLAM problem.

BACKGROUND

The SLAM problem has been extensively studied. Moravec and Elfes (1985) developed one of the first works in this area. They build a metric map by means of wide-angle sonar range measurements and a probabilistic approach. Subsequently laser sensors were introduced to improve the accuracy and computational efficiency of the algorithms. For example, Thrun (2001) develops a SLAM algorithm in which a team of robots builds a map online using laser sensors and a Monte Carlo approach. Lately, the use of cameras in the field of mobile robotics has become widespread due to the numerous advantages they offer (passive sensors, low cost, large amount of information, low power consumption, etc.). Many authors have studied the SLAM problem both using local features (Gil et al., 2010; Valiente et al., 2015) or global appearance (Paya et al., 2014; Berenguer et al. 2015).

The main problem that local features present is the high computational cost of the necessary processes. As a feasible alternative, the use of topological approaches is a field of great interest in the construction of maps by means of the global appearance of visual information, due to the numerous advantages it presents in terms of simplicity and computational cost. For example, Menegatti et al. (2004) carried out a study on robot navigation using the omnidirectional visual information captured from the environment, by using a global appearance descriptor.

Cameras have been used both to solve the SLAM problem metrically, topologically and in a hybrid way. For example, Guerra el al. (2014) and Gil et al. (2010) present two metric visual SLAM approaches using respectively a monocular and a stereo-camera and SURF features. On the other hand, Lui et al. (2012) develop a purely vision-based topological approach that builds an incremental model based on Haar Wavelet and Werner et al. (2009) carried out a task of topological SLAM using vision-based techniques and global appearance. They also make use of omnidirectional images, and furthermore they propose a Bayesian approach that combines the odometry information of the robot with the visual information to improve the accuracy. At last, Tully et al, (2012) develop a unified filtering framework for hybrid SLAM.

In this work we present a comparison of three global appearance descriptors in a process of robotic mapping, using a hybrid metric-topological SLAM.

APPEARANCE-BASED SLAM

In this work, we develop a new approach to hybrid SLAM, trying to obtain the advantages of each method. On the one hand, the topological representation permits building a global compact representation of the environment that permits a quick localization. On the other hand, the metrical representation makes use of the information provided by the topological method to detect loop closures, so that it is possible to correct the possible errors in the position of the robot. In both cases, the global appearance of the omnidirectional scenes captured by the sensor mounted on the robot is used to create the model and localize