Testing Image Segmentation for Topological SLAM with Omnidirectional Images

Anna Romero and Miguel Cazorla
Instituto Universitario Investigación en Informática, Universidad de Alicante
P.O. Box. 99 03080 Alicante
aromero@dccia.ua.es, miguel.cazorla@ua.es

Abstract. Image feature extraction and matching is useful in many areas of robotics such as object and scene recognition, autonomous navigation, SLAM and so on. This paper describes a new approach to the problem of matching features and its application to scene recognition and topological SLAM. For that purpose we propose a prior image segmentation into regions in order to group the extracted features in a graph so that each graph defines a single region of the image. We compare two basic methods for image segmentation, in order to know the effect of segmentation in the result. We have also extend the initial segmentation algorithm in order to take into account the circular characteristics of the omnidirectional image. The matching process will take into account the features and the structure (graph) using the GTM algorithm, modified to take into account the cylindrical structure of omnidirectional images. Then, using this method of comparing images, we propose an algorithm for constructing topological maps. During the experimentation phase we will test the robustness of the method and its ability to construct topological maps. We have also introduced a new hysteresis behavior in order to solve some problems found in the graph construction.

Keywords: Topological Mapping, Graph matching, Visual features.

1 Introduction

The extraction and matching of features and regions is an important area in robotics since it allows, among other things, object and scene recognition and its application to object localization, autonomous navigation, obstacle avoidance, topological SLAM.

The SLAM (Simultaneous Localization And Mapping) problem consists of estimating the position of the robot while building the environment map. The problem is not trivial, since errors in position estimation affect the map and vice versa. In the literature, depending on the form to represent the robot environment, we can talk of two types of SLAM: the Metric SLAM and the Topological SLAM. In the first, the position is determined by a continuous space, i.e. we know exactly what position the robot has on the map (with an assumed error). It is easy to find solutions that include odometry, sonars and lasers ([21][23]). There are less solutions using vision since calculating the exact position is more complicated. In the second type, the different points where you can find the robot are represented by a list of positions, i.e. the map is a discrete set of...
locations which defines a small region on the environment. In this case there are plenty of solutions that use images for the calculations. In [25] they use the images captured by the AIBO robot to learn the topological map. We also find solutions using omnidirectional images such as [26] and [24], [27] where a topological map is constructed using an incremental algorithm.

For both object and scene recognition we need methods of extracting features and/or regions from images. Several solutions in the literature use different methods for extracting the features. In [5] they use an over-segmentation algorithm for splitting the image into small regions. In [6] they combine the Harris corner detector with SIFT descriptor. Many solutions in the literature are based on the combination of a segmentation algorithm with a feature extractor ([5], [12], [10]).

Object recognition requires a manually selected database to describe the objects that the robot must recognize. In the case of scene recognition we could require a scene database as in [11] where it is introduced the concept of “Visual Place Categorization“ (VPC) which consists of identifying the semantic category of one place/room using visual information. However, there are situations requiring no pre-existing database as it is constructed as the robot navigates through the environment ([12], [13]) such as in the SLAM problem.

Affine invariant feature detectors have been shown to be very useful in several computer vision applications, like object recognition and categorization, wide baseline stereo and robot localization. These detection algorithms extract visual features from images that are invariant to image transformations such as illumination change, rotation, scale and slight viewpoint change. High level vision tasks that rely on these visual features are more robust to these transformations and also to the presence of clutter and occlusions. A more detailed survey of the state of the art of visual feature detectors can be found in [7]. In this work, the authors assess the performance of different algorithms for the matching problem, with the Maximally Stable Extremal Regions algorithm (MSER) [8], the Harris affine and the Hessian affine [9] being the best suited for that task.

Several methods are based on a combination of feature detectors (regions, contours and/or invariant points) to improve the matching and taking advantage of the extraction methods used, as well as eliminating some of the problems of the individual methods. However, it has not proposed the creation of structures from the extracted features to check the overall consistency of the matchings but the features are matched one by one without taking into account any possible neighborhood relationships. Some of those methods apply a matching consistency, eliminating cross-matches, those matches that intersects with others. In the case of omnidirectional images that can not be done, due to the circular nature of the images.

In this paper we propose a method for matching features and an algorithm to construct topological maps using this comparison method. For the image comparison method we propose a image pre-processing in two steps: segmentation into regions and invariant feature extraction (using MSER with SIFT descriptors). For image segmentation, we compare the results for two algorithms: JSEG, which provides a good segmentation but it takes so much time to compute, and EGBIS, which is faster than the previous one. We have also extend the initial segmentation algorithm in order to take into account the circular characteristics of the omnidirectional image.