

A Probabilistic Approach to Appearance-Based Localization and Mapping

F. M. Campos¹, L. Correia² and J. M. F. Calado³

Abstract. Recent approaches to appearance-based robot localization and mapping have relied on a topological representation of the environment and loop closing procedures to build correct maps. Because these solutions do not separate the problem of localization in mapped areas from the exploration of new locations, they require that loop-closing detection is performed continuously, inducing an extra computational burden on the robot. In this paper, we introduce a modified version of the Markov localization algorithm which avoids this condition by employing procedures that detect when novel areas are being explored or already mapped areas are revisited.

1 INTRODUCTION

Robots to be deployed in indoor environments should be able to autonomously build a map to support navigation and evolve this representation as they accumulate experience in the same environment. Instead of using it as static structure, it is desirable that robots have mechanisms for expanding the map as new areas of the environment are explored or making it more accurate as it revisits already mapped areas. The problem of SLAM (self-localization and mapping) addresses these issues, as it merges the problem of localization and mapping into a single estimation problem. Recently, a significant number of studies have been devoted to the use of vision as the sensory source to be used on localization and mapping, taking advantage of the high discriminative power of the visual signal, when compared to sonar or laser scanner readings [1][2][3][4][5]. The approach taken in this work differs from previous studies by explicitly separating inference rules for exploration and localization in mapped areas, and by using separate representations for mapped and unmapped positions.

The paper is organized as follows: section 2 presents an overview of the proposed method and the subsequent section details the beliefs representation and the inference mechanisms used. Section 4 presents experimental results, conclusions and directions for further work.

2 METHOD OVERVIEW

The method proposed in this paper extends the Markov localization algorithm by explicitly considering the hypothesis of the robot being in an unmapped state, which is represented metrically. The map representation that is used consists of position samples

extracted from past trajectories, estimated from odometry. Associated with each mapped position is a visual signature taken from that location. In our method, Markov localization is performed over the mapped positions, both in exploration and revisiting situations. In the latter case, Markov localization guarantees correct position estimation and the unmapped hypothesis is not associated with a particular position. However, when the exploration condition is detected, the unmapped hypothesis is represented by an estimate of the current position, which is tracked based on odometry readings, until a loop closing is detected. Additionally, the uncertainty about this position is estimated, allowing for a reasonable distribution of beliefs over the mapped positions, when the robot returns to already mapped areas. At the core of the method are two detection procedures, which trigger the transition from localization on mapped areas and localization during exploration.

3 INFERENCE MECHANISM

Let the domain of beliefs be defined by the set $\{x^0, x^1, \dots, x^n\}$, where x^1 to x^n are the mapped positions and x^0 represents the unmapped position state. In the Markov localization framework an action model is used with the purpose of updating the beliefs based on the known action taken, usually measured by odometry. Denoting the action at time t by u_t , the prediction update is computed by

$$\overline{bel}(x_t) = \sum_{x_{t-1}} p(x_t | x_{t-1}, u_t) bel(x_{t-1}) \quad (1)$$

where $bel(x_{t-1})$ is the belief distribution originating from the previous inference step and $\overline{bel}(x_t)$ is the distribution originating from prediction. The form of the action model, $p(x_t | x_{t-1}, u_t)$, strongly depends on the kind of discretization used for the domain of beliefs. Due to the sparse representation of positions used in our map model, some approximations have to be introduced in this function. Defining x'_t as the coordinates of the position predicted from odometry considering the previous state x_{t-1} , $p(x_t | x_{t-1}, u_t)$ is computed as the value on x_t of a Gaussian centered in x'_t . Rather than representing the uncertainty in the motion measurements, the variance of this Gaussian is intended to reflect the fact that x'_t will not exactly fall on any of the position samples that comprise the map, especially if a coarse sampling is used. Due to the discrete representation of the environment, and the fact that not all areas of the state space are covered by the map, the application of this action model may result in a decrease of the total probability over the map, after the prediction step. In fact, the variation of the probability over the map is not undesired, as it can be used to determine if the robot is moving into mapped or unmapped

¹ Mechanical Engineering Department, Instituto Superior de Engenharia de Lisboa/LabMAG, Portugal, email: fcampos@dem.isel.pt

² LabMAG, Computer Science Department, University of Lisbon, Portugal, email: Luis.Correia@di.fc.ul.pt

³ Mechanical Engineering Department, Instituto Superior de Engenharia de Lisboa/IDMEC, Portugal, email: jcalado@dem.isel.pt

areas. In our modification of the Markov localization algorithm, we consider two inference modes corresponding to the two different situations, which we will refer to as *localization* and *exploration* modes, respectively. While in localization mode, the belief of the unmapped state should be zero and the prediction step is performed by

$$\begin{cases} \overline{bel}(x_t) = \sum_{x_{t-1}=x^1 \dots x^n} p(x_t | x_{t-1}, u_t) bel(x_{t-1}) + \varepsilon, & x_t = x^1, \dots, x^n \\ \overline{bel}(x_t) = 0, & x_t = x^0 \end{cases} \quad (2)$$

Here, ε denotes the amount of probability lost over the map, divided by the number of mapped positions. This means that, in localization mode, the mass of probability that is lost due to approximations is distributed back uniformly over all the positions in the map, in order to guarantee that the final beliefs sum to 1.

When the robot moves into an unmapped area, the probability over the map decreases abnormally, and the algorithm is switched into exploration mode. When entering this mode, x_0 is initialized to the mapped position with the highest probability in the previous step and $bel(x_0)$ is set to ε . In subsequent steps of the exploration mode, the x_0 state is estimated incrementally according to the odometry readings. In this mode, the prediction update is calculated from

$$\begin{cases} \overline{bel}(x_t) = \eta \cdot \sum_{x_{t-1}=x^0 \dots x^n} p(x_t | x_{t-1}, u_t) bel(x_{t-1}), & x_t = x^1, \dots, x^n \\ \overline{bel}(x_t) = \eta \cdot (bel(x_t) + \delta), & x_t = x^0 \end{cases} \quad (3)$$

where δ is the probability loss over the map and η is a normalization constant. Notice that in this update rule the belief of the mapped positions receive a contribution not only from the hypotheses from the map, but also from the unmapped position hypothesis. This requires that an action model of the form $p(x_t | x_{t-1}=x_0, u_t)$ is defined. This function was also defined by a Gaussian, but in this case representing the accumulated uncertainty due to the estimation of the unmapped position. This representation of the uncertainty allows for a sensible distribution of probability over the map positions when the robot is returning to a mapped area, facilitating the correct and prompt detection of loop closing.

In this work, image appearance is characterized by features extracted by the Local Binary Pattern (LBP) operator [6], gathered in the form of a histogram of occurrences. Visual dissimilarity is computed by a histogram distance measure, in this case the χ^2 distance. Following [3], the likelihood of an observation at iteration t , z_t , has been defined as a Gaussian function of the visual distance, defined by

$$p(z_t, x_t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{d(z_t, x_t)}{2\sigma^2}} \quad (4)$$

Here, $d()$ denotes the visual distance function, z_t is the visual signature associated with x_t and σ was set empirically. For the likelihood of the unmapped hypothesis a probability of 1 was used.

4 EXPERIMENTAL RESULTS AND CONCLUSIONS

Two experiments were conducted on the public dataset KTH-IDOL2 [7], in order to validate our method. In the first experiment the robot is given an incomplete map of the environment and must switch between localization and exploration modes due to visiting two unmapped rooms. Results from this experiment showed that the algorithm is able to detect early when the robot is moving beyond the map and also correctly locate itself in the map after returning. However, the detection of loop closing only occurred after the robot entered 1m into the map after visiting the first room and 0.75m after visiting the second. Globally, localization was performed accurately, with a mean error of 0.35m on a map whose positions are sampled with an interval of 0.5m. In a second experiment, the algorithm was tested against a long loop closing covering a distance of approximately 20m, occurring while the environment is being mapped for the first time. In this case, localization at loop closing was again accurately estimated but the distance after entering the map increased to 2.5m. This is due to the fact that after a long loop, the uncertainty about the unmapped position is high, implying that the problem resembles a global localization problem, on which localization algorithms take a longer time to converge to a solution.

The experiments substantiate the validity of the approach both in the situation when the robot travels through a partially mapped environment and in the case the robot must localize and build a map from scratch. Future work should focus on a more extensive validation of the method and on the issue of employing different sensor models in localization and exploration. In fact, at exploration stage, there is more pressure on the sensor model for being sufficiently discriminative, so that incorrect loop closings do not occur. By distinguishing the two modes, there is the opportunity to use more powerful visual features for loop closing detection. These would be suppressed during localization, in order to reduce the processing burden on the robot during this more frequent stage.

REFERENCES

- [1] P. Rybski, S. Roumeliotis, M. Gini, and N. Papanikopoulos, 'Appearance-based mapping using minimalistic sensor models', *Autonomous Robots*, **24**, n. 3. Springer, (2008).
- [2] F. Fraundorfer, C. Engels, and D. Nister, 'Topological mapping, localization and navigation using image collections' *Int. Conf. on Intelligent Robots and Systems*, San Diego, 3872-3877, (2007).
- [3] T. Goedemé, M. Nuttin, T. Tuytelaars, and L. Van Gool, 'Markerless Computer Vision Based Localization using Automatically Generated Topological Maps', *European Navigation Conference GNSS*, (2004).
- [4] O. Booij, B. Terwijn, Z. Zivkovic, and B. Krose, 'Navigation using an appearance based topological map', *IEEE Int. Conf. on Robotics and Automation (ICRA)*, Roma, Italy, 3927-3932, (2007).
- [5] H. Andreasson, T. Duckett, and A. Lilienthal, 'A minimalistic approach to appearance-based visual SLAM', *IEEE Tran. on Robotics*, **24**, 991-1001, (2008).
- [6] T. Ojala M. Pietikainen and D. Harwood, 'A comparative study of texture measures with classification based on featured distribution', *Pattern Recognition*, **29**, 51-59, (1996).
- [7] J. Luo, A. Pronobis, B. Caputo, and P. Jensfelt, 'The KTH-IDOL2 Database', Technical Report CVAP304, Kungliga Tekniska Hogskolan, CVAP/CAS, (2006).