Robotics and Autonomous Systems 72 (2015) 93–104

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Information-based view initialization in visual SLAM with a single omnidirectional camera

David Valiente[a,*], Maani Ghaffari Jadidi[b], Jaime Valls Miró[b], Arturo Gil[a], Oscar Reinoso[a]

[a] Miguel Hernández University, System Engineering and Automation Department, 03202, Elche, Spain
[b] Faculty of Engineering and IT, University of Technology Sydney, NSW 2007, Australia

HIGHLIGHTS

- Novel mechanism for the view initialization process in an EKF-based visual SLAM approach.
- An efficient strategy which accounts for information gain and losses.
- Probabilistic representation of features and correlation learning by Gaussian regression.
- Bounding the uncertainty mitigates the non-linear effects which compromise the solution.
- Accuracy and robustness comparison versus a traditional EKF-based SLAM approach.

ARTICLE INFO

Article history:
Received 19 December 2014
Received in revised form
19 April 2015
Accepted 4 May 2015
Available online 13 May 2015

Keywords:
Visual SLAM
EKF
Omnidirectional images
Gaussian process
Information gain

ABSTRACT

This paper presents a novel mechanism to initiate new views within the map building process for an EKF-based visual SLAM (Simultaneous Localization and Mapping) approach using omnidirectional images. In presence of non-linearities, the EKF is very likely to compromise the final estimation. Particularly, the omnidirectional observation model induces non-linear errors, thus it becomes a potential source of uncertainty. To deal with this issue we propose a novel mechanism for view initialization which accounts for information gain and losses more efficiently. The main outcome of this contribution is the reduction of the map uncertainty and thus the higher consistency of the final estimation. Its basis relies on a Gaussian Process to infer an information distribution model from sensor data. This model represents feature points existence probabilities and their information content analysis leads to the proposed view initialization scheme. To demonstrate the suitability and effectiveness of the approach we present a series of real data experiments conducted with a robot equipped with a camera sensor and map model solely based on omnidirectional views. The results reveal a beneficial reduction on the uncertainty but also on the error in the pose and the map estimate.

1. Introduction

The problem of SLAM poses a challenge in the framework of mobile robot applications. It involves a laborious process that deals simultaneously with the mapping and robot's localization. This fact brings a challenge with regard to complexity, as the procedure is expected to work incrementally and to return a coherent representation of the environment. Besides, the existence of noise sources become accountable for undesired effects which aggravate and jeopardize the final estimation.

Lately, visual sensors have reached a great emergence as the main tool for collecting information in the field of SLAM. They represent a promising option compared to classic sensors such as laser or sonar. They allow us to take the best advantage of cameras due to their low cost, light weight and low consumption principally. Nonetheless, their major benefit turns to be their capability to collect a large amount of visual information. Such a quality is especially remarkable in the case of omnidirectional cameras, whose field of view is maximum. Many approaches have exploited this aspect of single cameras by means of visual descriptors to encode 3D visual landmarks [1–3]. Omnidirectional cameras have also been used within different contexts successfully [4–6].

* Corresponding author. Tel.: +34 96 665 9005; fax: +34 96 665 8979.
E-mail addresses: dvaliente@umh.es (D. Valiente), Maani.Ghaffarijadidi@uts.edu.au (M. Ghaffari Jadidi), Jaime.VallsMiro@uts.edu.au (J. Valls Miró), arturo.gil@umh.es (A. Gil), o.reinoso@umh.es (O. Reinoso).

http://dx.doi.org/10.1016/j.robot.2015.05.005
0921-8890/

Over the last years, great efforts have been made on the study and research of the EKF-based SLAM methods sustained by visual sensors [7,1,8,2,9,10]. The main efforts have been concentrated on the position estimation of a 3D visual landmarks set in a common reference system. These approaches are liable to encounter difficulties in assuring the convergence of the solution, particularly in the presence of non-linear errors. Such errors are usually provoked by sensory input. Omnidirectional sensors are significantly susceptible to cause this issue [11], due to its highly non-linear nature. Their correspondent effects tend to affect severely the data association problem in SLAM [12]. Other offline algorithms [13–15] may be seen as an alternative technique to keep stability under non-linear circumstances for SLAM problems [16]. Within this last group, there are some other authors who take advantage of other iterative optimization techniques embedded in the core of the EKF filter [17,18].

In this approach we rely on an improved version of the EKF which demonstrates its ability to face these common shortcomings commented above. In particular, we employ EKF in a non-iterative way. The most relevant characteristic of this approach is the definition of the map by omnidirectional images (denoted as views), which are captured along the path of the robot and stored with their poses and visual descriptors. This idea is closely related to the concept of submap. Here, a reduced set of views constitutes a compact and simpler representation of the environment compared to traditional 3D landmark map models. The main novelty is a new mechanism for the initialization of views within the map building process, aimed at uncertainty reduction. We make use of the visual information provided by the visual sensor data in order to construct an information distribution model which accounts for information gain and losses. This task is carried out by means of a Gaussian Process (*GP*), which is included within the field of non-parametric Bayesian learning techniques. Application of non-parametric methods, such as *GPs* has recently proven great enhancements on the mapping tasks within the context of autonomous navigation. Continuous frontier maps are obtained by optimizing the process parameters, which reveal important uncertainty reduction [19,20]. Therefore, we propose the training of a *GP* as a tool to establish a bounded uncertainty scheme for our approach. By adopting such a technique, we pursue a positive impact on the uncertainty, which we intend to minimize. As a result, harmful effects that are likely to appear under high uncertainty conditions, such as errors induced by non-linearities and consequently instabilities and convergence difficulties are mitigated. As a consequence, a more robust and consistent map and trajectory estimate is obtained for the visual SLAM problem.

Summarizing, the fundamental aspects and contributions of this approach may be listed as follows:

- A new view initialization mechanism is presented for the map building process within the problem of EKF-based visual SLAM.
- This strategy accounts for information gain and losses more efficiently.
- Probabilistic representation of features and learning their correlations through Gaussian processes regression.
- Bounding the uncertainty leads to the mitigation of harmful effects induced by non-linearities in the framework of EKF-based visual SLAM.

This section has introduced the scope and it has also given a brief outline of the related work. Next, the rest of this paper has been structured in the following manner: Section 2 briefly presents the basic theory of an EKF filter within this framework. Section 3 provides a general explanation to our EKF-based visual SLAM approach. Next, Section 4 exposes the key points of this contribution, which is supported by Gaussian Processes and Information theory. Finally, Section 5 shows the results extracted from real data experiments. They are aimed at testing the validity and reliability of this approach in terms of accuracy and robustness, but they are especially seeking the uncertainty reduction, which is obviously translated into an improvement on the solution convergence. Comparison between this proposal and a former SLAM approach has also been included to support these results. Further discussion and conclusions are addressed in Section 6.

2. EKF

The principle of the EKF [21] is based on the iterative update of an augmented state vector which represents the real time estimate to the problem. Considering our specific visual SLAM case, constituted by a view-based map, the estimate returns the pose of the views in the map and the pose of the robot. Then, the state vector can be defined as:

$$\bar{x}(t) = [x_v, x_{l_1}, x_{l_2}, \ldots, x_{l_N}]^T \tag{1}$$

where x_v represents the current pose of the robot and x_{l_N} the pose of the Nth view in the map. Two linear relations are defined by $F(t)$ and $H_i(t)$ so as to encode the dependency between $\bar{x}(t)$ and the observation measurement $z_i(t)$ respectively. In addition, it is essential to bear in mind the information provided by the odometry of the robot $u(t+1)$, the uncorrelated Gaussian noise introduced into the system $v(t+1)$, and the noise generated by the sensors, $w_i(t)$, being also Gaussian and with covariance $R(t)$.

Three fundamental stages are well differentiated by the EKF to operate. Firstly, a prediction for $\hat{x}(t)$ and $\hat{z}_i(t)$ is proposed. Then the second stage makes use of this prediction to determine the deviation between the prior $\hat{z}_i(t)$ with respect to the real observation $z_i(t)$. This concept is commonly known as the innovation, and its meaning is of paramount significance in the computation of the final solution provided by the filter. Finally, the third stage takes into account the second stage's output to produce the refinement of the estimation obtained during the first stage, seen as an updating step. These three stages may be described by their analytic expressions in the following terms:

- Prediction

$$\hat{x}(t+1|t) = F(t)\hat{x}(t|t) + u(t) \tag{2}$$

$$\hat{z}_i(t+1|t) = H_i(t)\hat{x}(t+1|t) \tag{3}$$

$$P(t+1|t) = F(t)P(t|t)F^T(t) + Q(t) \tag{4}$$

being $P(t|t)$ and $P(t+1|t)$ the covariance matrices which correspond to the uncertainty of the estimation at instants t and $t+1$ respectively. $Q(t)$ is constituted by the noise parameters which characterize the odometry of the wheels of the vehicle.

- Innovation

$$v_i(t+1) = z_i(t+1) - \hat{z}_i(t+1|t) \tag{5}$$

$$S_i(t+1) = H_i(t)P(t+1|t)H_i^T(t) + R_i(t+1) \tag{6}$$

where $S_i(t+1)$ represents the innovation's covariance.

- Update

$$\hat{x}(t+1|t+1) = \hat{x}(t+1|t) + K_i(t+1)v_i(t+1) \tag{7}$$

$$P(t+1|t+1) = P(t+1|t) - K_i(t+1)S_i(t+1)K_i^T(t+1) \tag{8}$$

being $K_i(t+1)$ the gain matrix of the filter which plays the role of weighting. It is computed in the following manner:

$$K_i(t+1) = P(t+1|t)H_i^T(t)S_i^{-1}(t+1). \tag{9}$$

It is worth noting that $Q(t)$ and $R(t)$ have to be initialized. The noise parameters which characterize the odometry are introduced into $Q(t)$ and the experimental accuracy parameters associated with the visual sensor into $R(t)$. In addition, the odometry $u(t)$ is