ARTICLE IN PRESS

Information Sciences xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Information Sciences

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

A modified stochastic gradient descent algorithm for view-based SLAM using omnidirectional images

David Valiente *, Arturo Gil, Lorenzo Fernández, Óscar Reinoso

Miguel Hernández University, System Engineering Department, 03202 Elche, Spain

A R T I C L E I N F O

Article history:
Received 25 March 2013
Received in revised form 25 March 2014
Accepted 29 March 2014
Available online xxxx

Keywords:
Mobile robotics
Visual SLAM
Omnidirectional images
SGD

A B S T R A C T

This paper describes an approach to the problem of Simultaneous Localization and Mapping (SLAM) based on Stochastic Gradient Descent (SGD) and using omnidirectional images. In the field of mobile robot applications, SGD techniques have never been evaluated with information gathered by visual sensors. This work proposes a SGD algorithm within a SLAM system which makes use of the beneficial characteristics of a single omnidirectional camera. The nature of the sensor has led to a modified version of the standard SGD to adapt it to omnidirectional geometry. Besides, the angular unscaled observation measurement needs to be considered. This upgraded SGD approach minimizes the non-linear effects which impair and compromise the convergence of traditional estimators. Moreover, we suggest a strategy to improve the convergence speed of the SLAM solution, which inputs several constraints in the SGD algorithm simultaneously, in contrast to former SGD approaches, which process only constraint independently. In particular, we focus on an efficient map model, established by a reduced set of image views. We present a series of experiments obtained with both simulated and real data. We validate the new SGD approach, compare the efficiency versus a standard SGD and demonstrate the suitability and the reliability of the approach to support real applications.

1. Introduction

In the field of mobile robot applications, the problem of SLAM is a crucial factor, due to the need for a complete representation of the environment, especially for navigation purposes. The objective of building a map entails considerable complexity, since the map has to be built incrementally, while, the localization of the robot inside it needs to be calculated simultaneously. Generating a reliable and coherent map is even more challenging and laborious when sensor data is affected by noise, and this directly impairs the simultaneous estimation of the map and the path followed by the robot.

To date, SLAM approaches have been differentiated according to several factors, such as the way to estimate the representation of the map, the main algorithm for computing a solution and the kind of sensor to extract information from the environment. For instance, several map representations were obtained thanks to the extensive use of laser data range and sonar [8]. In this area, maps were principally generated following two representation models [16,11], corresponding, respectively, to 2D occupancy grid maps based on raw laser and 2D landmark-based maps focused on the extraction of features, described from laser data measurements.

* Corresponding author. Tel.: +34 96 665 9005; fax: +34 96 665 8979.
E-mail addresses: dvaliente@umh.es (D. Valiente), arturo.gil@umh.es (A. Gil), l.fernandez@umh.es (L. Fernández), o.reinoso@umh.es (Ó. Reinoso).

http://dx.doi.org/10.1016/j.ins.2014.03.122
0020-0255/

Please cite this article in press as: D. Valiente et al., A modified stochastic gradient descent algorithm for view-based SLAM using omnidirectional images, Inform. Sci. (2014), http://dx.doi.org/10.1016/j.ins.2014.03.122

More recently, the tendency has turned to the use of visual information by means of digital cameras. Many applications benefit from the use of these sensors, whose characteristics outperform previous sensors such as lasers in terms of the amount of usable information from the environment for building the map. For instance, the approaches that use two calibrated cameras, known as stereo-pairs, in order to extract a set of 3D visual landmarks determined by a visual description [5]. Other approaches simply exploit a single camera to estimate 3D visual landmarks [2,10]. They initialize the coordinates of each 3D landmark relying on the inverse depth parametrization, since there exists a scale uncertainty about the distance to each landmark which cannot be directly retrieved with a single camera. Omnidirectional cameras have also been used alone [15], and some others have even arranged two omnidirectional images, in order to take the best advantage of the wider field of view provided by these cameras.

As important as the kind of sensor and the map representation is the estimation algorithm for a SLAM scheme. It defines the core of the system, as it is responsible for the ultimate solution. Most extensively used are online methods such as EKF [4], Rao-Blackwellized particle filters [11] and offline algorithms, such as, Stochastic Gradient Descent [7].

The combination of data sensors, map representation and the core of the algorithm therefore determines the final effectiveness of a SLAM which seeks reliability and suitability for realistic applications. Great efforts have been made in this field. For example, certain approaches [4,5,3,2,14] have concentrated on estimating of the position of a set of 3D visual landmarks in a main reference system, while, simultaneously, building the map. The main idea lies in the capability of an EKF filter to converge the estimation to an appropriate solution for the SLAM problem. In this same line of EKF usage, [18] has recently proposed a distinctive map representation consisting of a reduced set of image views, determined by their position and orientation in the environment. Such a technique establishes an estimation of a state vector which includes the map and the current localization of the robot at each timestep k. The estimation of the transition between states at k and $k+1$ considers the wheel's odometry as initial estimate, but also the observation measurements gathered by sensors.

Generally, EKF methods are troublesome in the presence of non-linear errors as they have difficulties in maintaining the convergence of the estimation. This situation normally appears in presence of Gaussian errors introduced by the observation measurement, which usually causes data association problems [12]. A visual observation model, such as the omnidirectional, is susceptible to introduce non-linearities and is thus responsible for this kind of errors. By contrast, an offline algorithm such as SGD [1] may deal with this issue caused by non-linearity effects. Similarly, in [20,19,17], parallel approaches are presented to maintain stability in non-linear contexts.

Regarding the basic goals of this study, we present a new visual SLAM approach based on omnidirectional images and sustained by a SGD solver algorithm which helps overcome the harmful effects caused by errors. To achieve this, and depending on the nature of the problem, different aspects have to be taken into consideration so that the research is conducted towards the achievement of new contributions and advantages compared to former applications based on the standard SGD algorithm [13,7,6,1]. Firstly, a map model has to be adapted to the omnidirectional observation. Along the same line, the standard SGD has to be redesigned to be able to work with the omnidirectional geometry of the images, but also considering the nature of the measurement, which lacks scale. This implies that the solution to the problem is not a trivial one. So, the difference between our approach, which uses a different geometrical environment, and all the previous SGD applications, which consider data range observations in a Cartesian measurement system, should be noted. Next, to improve the efficiency of the standard method, in terms of the convergence speed, we propose a modification in the estimation procedure. The traditional models mentioned above, usually process every odometry and observation measurement (denoted as constraints) independently at each iteration step. By contrast, with the aim of finding a valid solution quickly, we propose a strategy based on the simultaneous use of a certain set of information provided by our visual observation measurements. This proposal might appear to be liable to cause an increase in the required computational resources. Nevertheless, we have concentrated on preventing this by updating several stages of the SGD's iterative optimization so as to avoid possible harmful bottleneck handicaps. Therefore, the main expected contributions and advantages of this SLAM approach compared to traditional approaches might be synthesized as it follows:

- An efficient map model established by a reduced set of omnidirectional images.
- A modified SGD solver algorithm adapted to the omnidirectional geometry which is the basis of the proposed SLAM's observation model. Development of the new differential equations related to the observation measurements.
- Improved efficiency of the estimation thanks to the use of simultaneous constraint processing.

The structure of the paper has been divided as it follows: Section 2 depicts the SLAM problem within this framework. Then, Section 3 describes the general specifications of a SGD algorithm, concentrating on the standard SGD. Section 4 details the proposed modification of the standard SGD and the main contributions mentioned above. Next, Section 5 provides both simulated and real data experimental results to validate the model and test its reliability and expected benefits versus traditional methods. Finally, Section 6 analyzes the results to draw a general conclusion.

2. SLAM

A visual SLAM technique is expected to retrieve a feasible estimation of the position of the robot within a certain environment, which also has to be precisely determined by the estimation. In our approach, the map is composed of a set of