ICINCO 2013

10th INTERNATIONAL CONFERENCE ON INFORMATICS
IN CONTROL, AUTOMATION AND ROBOTICS

Proceedings

Volume 2

REYKJAVÍK, ICELAND

29 - 31 JULY, 2013

SPONSORED BY: INSTICC

CO-ORGANIZED BY: HÁSKÓLINN Í REYKJAVÍK REYKJAVIK UNIVERSITY

TECHNICALLY CO-SPONSORED BY: IEEE, IEEE Robotics & Automation Society

CO-SPONSORED BY: IFAC

IN COOPERATION WITH: acm In-Cooperation, SIGART

ICINCO 2013

Proceedings of the
10th International Conference on
Informatics in Control, Automation and Robotics

Volume 2

Reykjavík, Iceland

29 - 31 July, 2013

Sponsored by
INSTICC – Institute for Systems and Technologies of Information, Control and Communication

Co-organized by
Reykjavik University

Technically Co-sponsored by
IEEE-RAS – IEEE Robotics and Automation Society

Co-Sponsored by
IFAC – International Federation of Automatic Control

In Cooperation with
ACM SIGART – Association for Computing Machinery / Special Interest Group on Artificial Intelligence
AAAI – Association for the Advancement of Artificial Intelligence
EUROMICRO

Edited by Jean-Louis Ferrier, Oleg Gusikhin, Kurosh Madani and Jurek Sasiadek

Printed in Portugal

ISBN: 978-989-8565-71-6

Depósito Legal: 361500/13

http://www.icinco.org

icinco.secretariat@insticc.org

CONTENTS

SHORT PAPERS

SLAM of View-based Maps using SGD

David Valiente, Arturo Gil, Francisco Amorós and Óscar Reinoso
System Engineering Department, Miguel Hernández University, 03202, Elche, Spain
{*dvaliente, arturo.gil, famoros, o.reinoso*}*@umh.es*

Keywords: Visual SLAM, Omnidirectional Images, SGD.

Abstract: This work presents a solution for the problem of Simultaneous Localization and Mapping (SLAM) based on a Stochastic Gradient Descent (SGD) technique and using omnidirectional images. In the field of applications of mobile robotics, SGD has never been tested with visual information obtained from the environment. This paper suggests the introduction of a SGD algorithm into a SLAM scheme which exploits the benefits of omnidirectional images provided by a single camera. Several improvements have been introduced to the vanilla SGD in order to adapt it to the case of omnidirectional observations. This new SGD approach reduces the undesired harmful effects provoked by non-linearities which compromise the convergence of the traditional filter estimates. Particularly, we rely on an efficient map representation, conformed by a reduced set of image views. The first contribution is the adaption of the basic SGD algorithm to work with omnidirectional observations, whose nature is angular, and thus it lacks of scale. Former SGD approaches only process one constraint independently at each iteration step. Instead, we think of a strategy which employes several constraints simultaneously as system inputs, with the purpose of improving the convergence speed when estimating a SLAM solution. In this context, we present different sets of experiments which have been carried out seeking for validation of our new approach based on SGD with omnidirectional observations. In addition, we compare our approach with a basic SGD in order to prove the expected benefits in terms of efficiency.

1 INTRODUCTION

The problem of SLAM is an essential aspect in the field of mobile robots, since a representation of the environment is necessary for navigation purposes. The aim of building a map entails a complex process, since the robot needs to build the map in an incremental manner, while, simultaneously, calculating its localization inside it. Achieving a coherent map is problematic due to the noisy sensor data, which affects the simultaneous estimation of the map and the path followed by the robot.

Traditionally, SLAM approaches have been differentiated according to the representation of the map, the basic algorithm to compute a solution and the kind of sensor to gather information of the environment. Some examples reveal the extensive use of laser range sensors to construct maps. In this context, two models were mostly used to estimate map representations: 2D occupancy grid maps based on raw laser (Stachniss et al., 2004), and 2D landmark-based maps focused on the extraction of features, described from laser measurements (Montemerlo et al., 2002).

The use of visual information has lately set a new tendency towards the utilization of cameras. These sensors outperform former sensors such as laser in terms of the amount of valid information to build the map of the environment. Despite the fact that vision sensors are not as precise as laser sensors, they provide a huge variety of information of the scene, as well as they are less expensive, lighter and more efficient in terms of consumption. In the latter group there exists several alternatives according to the number of cameras introduced and their configuration. For instance, stereo-based approaches (Gil et al., 2010), in which two calibrated cameras extract relative measurements of a set of 3D visual landmarks, determined by a visual description. Other approaches present their estimation of 3D visual landmarks by using a single camera. In (Civera et al., 2008; Joly and Rives, 2010) an inverse depth parametrization is carried out to initialize the coordinates of each 3D landmark since the distance to the visual landmark cannot be directly extracted with a single camera. Some other approaches (Jae-Hean and Myung Jin, 2003) have also combined two omnidirectional cameras to exploit the benefits of a wider field, as in the case of a stereo-pair sensor.

Besides the sensor and the representation of the map, the kind of algorithm is the main basis to ob-

tain a solution for the problem of SLAM. In this sense, different SLAM algorithms have been extensively used, differentiating between online methods such as, EKF (Davison and Murray, 2002), Rao-Blackwellized particle filters (Montemerlo et al., 2002) and offline algorithms, such as, for example, Stochastic Gradient Descent (Grisetti et al., 2007).
Thus, the success of an efficient SLAM solution is closely dependent on the suitability of the combination between data sensor information, map representation model and the SLAM algorithm. A large amount of research has been conducted to obtain the map representation, dealing simultaneously with the estimation of the position of a set of 3D visual landmarks expressed in a general reference system (Davison and Murray, 2002; Gil et al., 2010; Andrew J. Davison et al., 2004; Civera et al., 2008). They rely on the capability of an EKF filter to converge to a proper solution for the problem of SLAM. More recently (Valiente et al., 2012), also based on EKF, proposed a different representation of the environment, where the map is formed by the position and orientation of a reduced set of image views in the environment. The most distinctive novelty of this representation is the capability to compute a relative movement between views, which allows to retrieve the localization of the robot and simultaneously it provides a compact representation of the environment by means of a reduced set of views. Such 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 $k+1$ takes into account the wheel's odometry as initial estimate, and the observation measurements gathered by sensors. The major problematic experienced by these methods based on an EKF is the difficulty to keep the convergence of the estimation when gaussian errors appear in the observation, since they usually cause important data association problems (Neira and Tardós, 2001). A visual observation as omnidirectional is susceptible to introduce non-linearties and therefore responsible for those errors.

Some approaches have demonstrated considerable improvements thanks to the assumption of a novel representation of the map (Valiente et al., 2012), composed by a subset of omnidirectional images, denoted as views. Each view is representative of a large area of the map since it may encapsulate a huge amount of visual information in its neighborhood. Considering a reduced set of images, it leads to a simplification of the map representation. The scheme allows for a simple-computation of an observation measurement: given two images acquired at two poses of the robot, a specific subset of significant points between images can be processed in order to get a motion transformation which establishes the movement between the two poses of the robot. This approach mitigates the main EKF's efficiency issue related to the high number of variables needed to estimate a solution at each iteration. Here, the number of landmarks, seen as views, is drastically reduced and thus the time consumption for the state's re-estimation too. On the other hand, the problematic of convergence associated to EKF's algorithms still exists. In this work we rely on the SGD algorithm, which it is known to outperform EKF's SLAM results. Our main goal is to benefit from our visual observation model adapted to this kind of SLAM solution algorithm based on SGD. Note that even with SGD, solution's convergence is not trivial when the observation does not provide range distance values. We also propose a variation in the procedure to estimate the solution. Traditionally, either odometry or observation are independently used in order to iterate up to reach a final solution. By contrast, we suggest the use of the information provided by several observations at the same time to quickly find the SLAM solution. This proposal gives the first impression of being liable to cause an increase of the required computational resources. Nevertheless, we have concentrated on the updating of several stages of the iterative optimization process aiming at prevention from possible harmful bottleneck effects.

The paper is divided as follows: Section 2 explains the SLAM problem in the mentioned context. Then, in Section 3 we detail the specifications of a Maximum Likelihood algorithm such as SGD. Section 4 describes the main novelties provided by this approach. Next, Section 5 shows the experimental results to test the suitability of the model and its benefits. Finally, Section 6 is referred to the result's analysis in order to extract a general conclusion.

2 SLAM

A visual SLAM technique is expected to retrieve a feasible estimation of the position of the robot inside a certain environment, which it has to be well-determined by the estimation as well. In our approach, the map is composed by a set of omnidirectional images obtained from different poses in the environment, denoted as views. These views do not represent any physical landmarks, since they will be constituted by an omnidirectional image captured at the pose $x_l = (\mathrm{x}_l, \mathrm{y}_l, \theta_l)$ and a set of interest points extracted from that image. Such arrangement, allows us to exploit the capability of an omnidirectional image

to gather a large amount of information in a simple image, due to its large field of view. Thus, an important reduction is achieved in terms of number of variables to estimate the solution.

The pose of the mobile robot at time t will be denoted as $x_v = (\mathrm{x}_v, \mathrm{y}_v, \theta_v)^T$. Each view $i \in [1, \ldots, N]$ is constituted by its pose $x_{l_i} = (\mathrm{x}_l, \mathrm{y}_l, \theta_l)_i^T$, its uncertainty P_{l_i} and a set of M interest points p_j expressed in image coordinates. Each point is associated with a visual descriptor d_j, $j = 1, \ldots, M$.

Thus, the augmented state vector is defined as:

$$\bar{x} = \begin{bmatrix} x_v & x_{l_1} & x_{l_2} & \cdots & x_{l_N} \end{bmatrix}^T \tag{1}$$

where $x_v = (\mathrm{x}_v, \mathrm{y}_v, \theta_v)^T$ is the pose of the moving vehicle and $x_{l_N} = (\mathrm{x}_{l_N}, \mathrm{y}_{l_N}, \theta_{l_N})$ is the pose of the N-view that exist in the map.

2.1 Observation Model

According to the view-based representation of the map, it is also necessary to describe an observation model. The versatility of using omnidireccional images makes it possible to extract an observation measurement which defines the motion transformation between two poses (Valiente et al., 2012), as seen in Figure 1. Actually, these poses represent the positions where the robot acquired two specific images. To that effect, only two images with a set of corresponding points between them are required to obtain the transformation. Following, the observation measurement is described:

$$z_t = \begin{pmatrix} \phi \\ \beta \end{pmatrix} = \begin{pmatrix} \arctan\left(\frac{\mathrm{y}_{l_N} - \mathrm{y}_v}{\mathrm{x}_{l_N} - \mathrm{x}_v}\right) - \theta_v \\ \theta_{l_N} - \theta_v \end{pmatrix} \tag{2}$$

where the angle ϕ is the bearing at which the view N is observed and β is the relative orientation between the images. The view N is represented by $x_{l_N} = (\mathrm{x}_{l_N}, \mathrm{y}_{l_N}, \theta_{l_N})$, whereas the pose of the robot is described as $x_v = (\mathrm{x}_v, \mathrm{y}_v, \theta_v)$. Both measurements (ϕ, β) are represented in Figure 1.

3 MAXIMUM LIKELIHOOD MAPPING

3.1 Problem Specifications

A graph-oriented map is composed by a set of nodes defining the poses traversed by the robot and the landmarks initialized into the map. The state vector s_t encodes this representation through a set of variables which are expressed in the following manner:

$$s_t = \begin{bmatrix} (x_0, y_0, \theta_0), (x_1, y_1, \theta_1) \ldots & (x_n, y_n, \theta_n) \end{bmatrix} \tag{3}$$

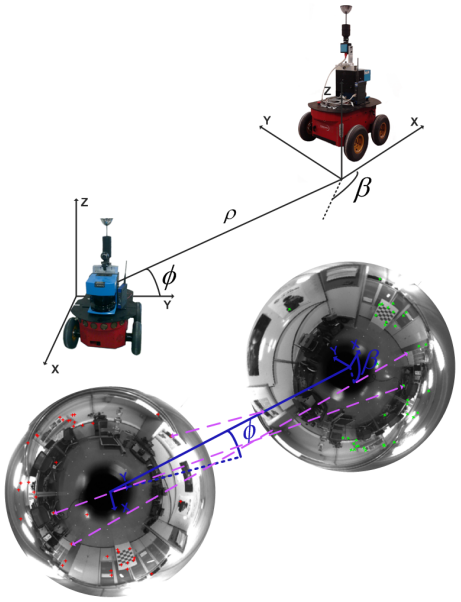


Figure 1: Motion relation between two poses of the robot with two omnidirectional images acquired at that poses. The same motion relation is shown in the images with specific observation variables ϕ and β indicated.

being (x_n, y_n, θ_n) the 3D coordinates in a general reference system. A complementary subset of edges represents the relationships between nodes, by means of either distance measurements generated by the odometry or observations measurements provided by the on board sensors. Both measurements are commonly known as constraints and denoted as δ_{ji}, where j indicates the observed node, seen from node i. The general objective stated by these kind of methods (Olson et al., 2006; Grisetti et al., 2007) is to minimize the error likelihood expressed as:

$$P_{ji}(s) \propto \eta \exp\left(-\frac{1}{2}(f_{ji}(s) - \delta_{ji})^T \Omega_{ji} (f_{ji}(s) - \delta_{ji})\right) \tag{4}$$

being $f_{ji}(s)$ a function dependent on the state s_t and both nodes j and i. The difference between $f_{ji}(s)$ and δ_{ji} expresses the error deviation between nodes. Such error term is weighted by the information matrix:

$$\Omega_{ji} = \Sigma_{ji}^{-1} \tag{5}$$

where Σ_{ji}^{-1} is the associated covariance, aimed at introducing the uncertainty of the measurements. The assumption of logarithmic notation leads to:

$$F_{ji}(s) \propto (f_{ji}(s) - \delta_{ji})^T \Omega_{ji} (f_{ji}(s) - \delta_{ji}) \tag{6}$$

$$= e_{ji}(s)^T \Omega_{ji} e_{ji}(s) = r_{ji}(s)^T \Omega_{ji} r_{ji}(s) \tag{7}$$

being $e_{ji}(s)$ the error determined by $f_{ji}(s)$-$\delta_{ji}(s)$, also denoted as $r_{ji}(s)$ to emphasize its condition of residue. Finally, the global problem seeks the minimization of the objective function which represents the accumulated error:

$$F(s) = \sum_{\langle j,i\rangle \in G} F_{ji}(s) = \sum_{\langle j,i\rangle \in G} r_{ji}(s)^T \Omega_{ji} r_{ji}(s) \quad (8)$$

where $G = \{\langle j_1, i_1\rangle, \langle j_2, i_2\rangle \ldots\}$ defines the subset of particular constraint conforming the map, either pertaining to odometry or observation measurements.

3.2 Problem Solution. SGD

Once the formulation of the problem has been presented, it has to be detailed the Stochastich Gradient Descent algorithm. The basic goal is to compute in an iterative manner a closer estimation to reach a valid solution for the SLAM problem. The basis of a SGD method lays on the minimization of (8) through derivative optimization techniques. The estimated state vector is obtained as:

$$s_{t+1} = s_t + \Delta s \quad (9)$$

where Δs expresses a certain update with respect to s_t, term which is sequentially generated by means of the constraint optimization procedure. It is worth noting that in a general case, this update is calculated independently at each step by using only a simple constraint, that is to say $\Delta s_n = f(\delta_{ji})$. The general expression for the transition between s_t and s_{t+1} has the following form:

$$s_{t+1} = s_t + \lambda \cdot H^{-1} J_{ji}^T \Omega_{ji} r_{ji} \quad (10)$$

- λ is a learning factor to re-scale the term $H^{-1} J_{ji}^T \Omega_{ji} r_{ji}$. Normally, λ takes decreasing values following the criteria $\lambda = 1/n$, where n is the iteration step. This strategy pretends to quickly reach the final solution for higher values of λ. In case it is close to the optimum, lower values of λ are expected to prevent from oscillations around the final solution.
- H is the Hessian matrix, calculated as $J^T \Omega J$, and it represents the shape of the error function through a preconditioning matrix to scale the variations of J_{ji}. According to (Grisetti et al., 2009), H can be computed:

$$H \approx \sum_{\langle i,j\rangle} J_{ji} \Omega_{ji} J_{ji}^T \quad (11)$$

approximating H^{-1} to allow a simpler and faster computation:

$$H^{-1} \approx [diag(H)]^{-1} \quad (12)$$

- J_{ji} is the Jacobian of f_{ji} with respect to s_t, $J_{ji} = \frac{\partial f_{ji}}{\partial s}$. It converts the error deviation into a spacial variation.
- Ω_{ji} is the information matrix associated to a constraint. $\Omega_{ji} = \Sigma_{ji}^{-1}$, being Σ_{ji} the covariance matrix corresponding with the constraint observations δ_{ji}

This scheme updates the estimation by computing the rectification introduced by each constraint at each iteration step respectively. Despite the learning factor to reduce the weight by which each constraint updates the estimation, the procedure can lead to an inefficient method to reach a stable solution, since undesired oscillations may occur due to the stochastic nature of the constraints' selection. For this reason, we propose an optimization process which takes into account several constraints at the same iteration. It might be thought that same drawbacks of a general case could arise, in addition of some other inconveniences such as undesired overloads of time, as a consequence of the simultaneous processing of several constraints at the same iteration. However, we have modified some calculations at specific stages of the algorithm in order to maintain the time requirements and even to reduce them. As a results, we achieved improved convergence ratios in terms of speed. Further details will be provided in the next sections.

4 MODIFIED SGD

The first assumption to consider is the redefined state vector s_t, which will be treated as a set of incremental variables. Pose incremental is defined as:

$$s_t^{inc} = \begin{bmatrix} x_0 \\ y_0 \\ \theta_0 \\ dx_1 \\ dy_1 \\ d\theta_1 \\ \cdots \\ dx_n \\ dy_n \\ d\theta_n \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ \theta_0 \\ x_1 - x_0 \\ y_1 - y_0 \\ \theta_1 - \theta_0 \\ x_2 - x_1 \\ y_2 - y_1 \\ \theta_2 - \theta_1 \\ \cdots \\ x_n - x_{n-1} \\ y_n - y_{n-1} \\ \theta_n - \theta_{n-1} \end{bmatrix} \quad (13)$$

where $(dx_i, dy_i, d\theta_i)$ encode the variation between consecutive poses in coordinates of the global reference system. A global encoding has the main drawback of not being capable to update more than one node and its adjacents per constraint. Regarding a relative codification of the sate, it arises the problem of

non-linearities in J_{ji}. By contrast, an incremental state vector allows a single constraint to generate a variation on every pose. In this framework, Δs (9) affects all poses because the state vector is differentially encoded.

It should be noticed that in this approach we are dealing with a visual observation given by an omnidirectional camera. This fact makes us to adapt the equations defined in the previous section, since the nature of the constraints are not only metrical like odometry's constraints. According to (2), given two nodes, the observation measurement allows us to determine a specific motion transformation between them up to a scale factor. Therefore, the omnidirectional measurements and the incremental representation require the reformulation of several terms involved in the estimation. Following, we detail all the terms which needed to be modified and recalculated.

- The first adaption made was in f_{ji}, differentiating between odometry:

$$f_{j,i}^{odo}(s) = \begin{pmatrix} dx_j \\ dy_j \\ d\theta_j \end{pmatrix} + \begin{pmatrix} dx_{j-1} \\ dy_{j-1} \\ d\theta_{j-1} \end{pmatrix} + \cdots + \begin{pmatrix} dx_i \\ dy_i \\ d\theta_i \end{pmatrix} \quad (14)$$

  where $(dx_j, dy_j, d\theta_j)$ are defined in (13), and visual observation constraints:

$$f_{j,i}^{visual}(s) = \begin{pmatrix} \phi \\ \beta \end{pmatrix} = \begin{pmatrix} \arctan\left(\frac{\mathrm{dy}_j - \mathrm{dy}_i}{\mathrm{dx}_j - \mathrm{dx}_i}\right) - d\theta_i \\ d\theta_j - d\theta_i \end{pmatrix} \quad (15)$$

  where ϕ and β are directly computed as the observation measurement (Valiente et al., 2012) which expresses the relation between the omnidirectional images and the pose codification (13). Inspection of Figure 1 leads to define (2) and similarly (15).

- Second step, according to this reformulation, is to recalculate $J_{ji} = \frac{\partial f_{ji}}{\partial s}$. Please note the importance of considering the indexes of the corresponding nodes, either $j > i$ or $j < i$ since the derivatives vary its form considerably. Furthermore, as seen above, the dimensions of $f_{j,i}$ are different, fact which has to be taken into account to resize the rest of the terms involved in the SGD algorithm.

$$J_{j,i} = \frac{\partial f_{j,i}}{\partial s} = \frac{\partial f_{j,i}(s)}{\partial s} = \left[\frac{\partial f_{j,i}(\phi)}{\partial s}, \frac{\partial f_{j,i}(\beta)}{\partial s}\right] \quad (16)$$

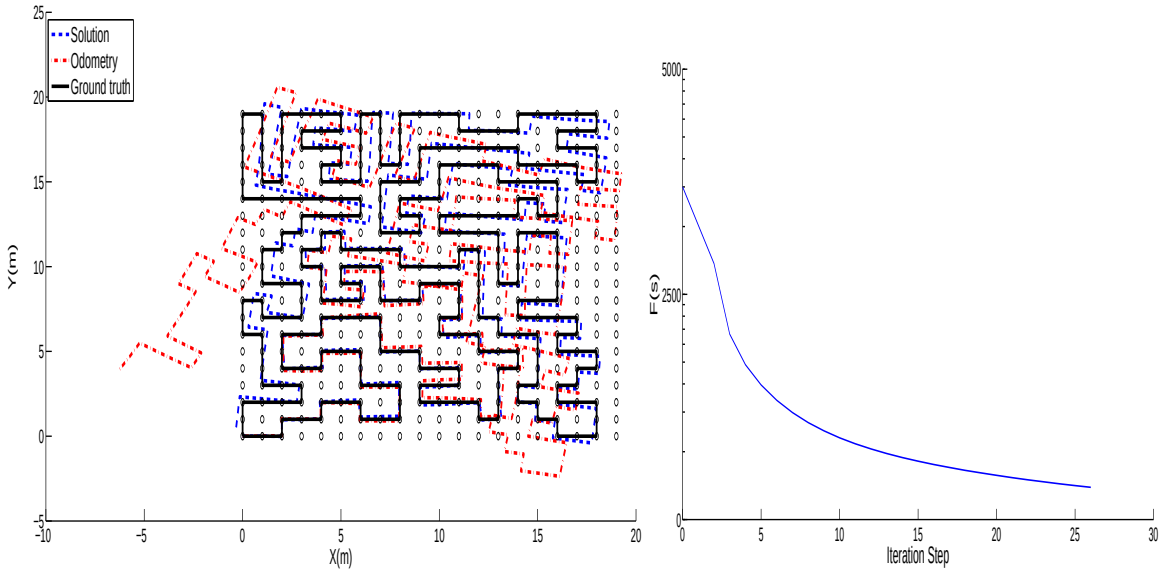
- Thirdly, in this work we suggest the estimation of the new state s_{t+1} by considering several constraints at the same time. With this assumption, we seek for more relevance of constraints' weight when searching for the optimal minimum estimation. It is clear that computing more than one constraint at each algorithm step leads to a certain overload. By contrast, with this approach, we reduce the expensive estimation of H. In a general case H is computed whenever a single constraint is introduced, that is to say, H is computed as many times as constraints exist. In our case we only compute H once for each subset of constraints introduced simultaneously into the system. Consequently we obtain H in a more efficient manner, and therefore we compensate possible time overloads associated.

5 RESULTS

We have carried out three different experimental sets. First, in Section 5.1 we show SLAM results obtained in a simulated scenario to confirm the validity of the new SLAM approach supported by SGD. Then, Section 5.2 presents a comparison of SLAM results using our approach and a basic SGD algorithm, respectively. Finally, in Section 5.3 we present SLAM results in an office-like environment.

5.1 Experimental Set 1

Assuring the convergence of an SLAM algorithm is of paramount importance when a new solver technique is introduced, such SGD. Moreover, the performance of the new method has to consider certain modifications to treat with a visual observation model, which usually adds non-linearities. Figure 2(a) presents a random simulation environment of 20x20m, where the robot traverses 300m approximately. The real path followed by the robot is shown with continuous line, the odometry is represented with dash-dot line, whereas the estimated solution is shown with dash line. A set of views have been placed randomly along the trajectory. The arrangement of these views is controlled by an appearance ratio between images, so as to assure a realistic placement of each view. A grid of circles represents the possible poses where the robot might move to and gather a new view. The number of iterations of the SGD algorithm is 25. As it can be observed in Figure 2(a), starting from a noisy first odometry estimate, the final estimation has been rectified following the tendency of the real path. Figure 2(b) shows the decreasing evolution of the accumulated error probability $P_{ji}(s)$ in (4), expressed in logarithmic terms (8), versus the number of iterations. It can be confirmed the reliability of this new approach to work



(a) (b)

Figure 2: Figure 2(a) presents a map obtained with the proposed approach in an environment of 20x20m. Continuous line shows the ground truth, dash-dot line the odometry and dash line the estimated solution. Figure 2(b) shows the accumulated error probability $F(s)$ along the number of iterations.

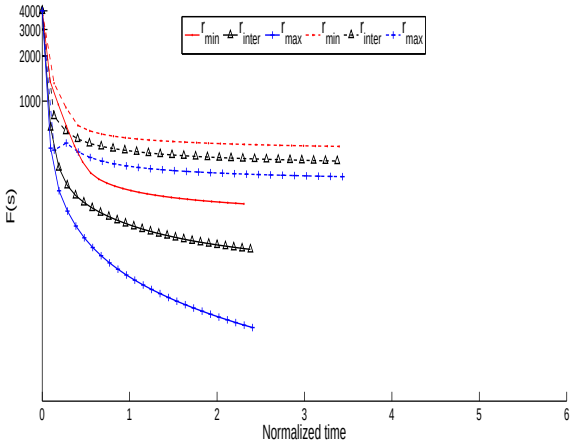
with omnidirecctional observations since it provides a correct solution.

5.2 Comparing Experiments

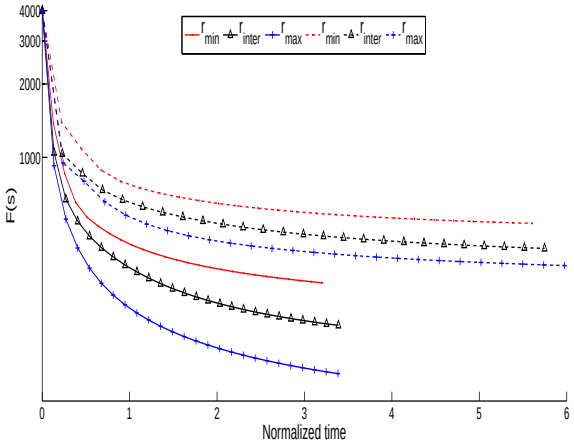
The following experiments has been conducted to compare our approach with a basic SGD in terms of efficiency. We suggest a strategy to introduce simultaneously several constraints into the SGD algorithm. The main goal is to improve the speed by which the method iteratively optimizes up to reach a final solution. In this sense, we have performed a SLAM experiment, where the robot traverses 50m through a given environment. Again, the number of views in the map has been randomly placed by following the same policy explained above. The same experiment has been repeated 200 times by using the same series of odometry inputs, so as to provide mean values which express consistent results. Both approaches, ours and the basic SGD algorithm's have been compared. We have modified the number of views N that the robot is able to observe from each pose. The observation range r of the robot has also been varied. Figure 3 presents results for the accumulated error probability $F(s)$, being the objective function which the SGD algorithm seeks to minimize. We compare the solution obtained with our approach, drawn with continuous line, versus the solution obtained with a basic SGD algorithm, drawn with dashed line. Figures 3(a), 3(b) and 3(c) represent $F(s)$ when the robot observes N=2, N=4 and N=8 views, respectively. Since we look for a fair comparison, the x-axis, originally representing iteration steps, has been transformed to a normalized time variable to generate a trustworthy comparison between both schemes. In terms of efficiency, it may be proved that the solution provided by our method outperforms the solution given by a basic SGD, since the decreasing slope of $F(s)$ is clearly steeper, therefore a quicker convergence speed is demonstrated. This is the main consequence of combining several constraints at each iteration step, instead of using only one like a basic SGD used to. It is also remarkable the relevance of the observation range of the vehicle r. As seen in Figures 3(a), 3(b) and 3(c), long values of r provide a better convergence to the detriment of shorter r. Since the omnidirectional observation is angular, and it lacks of scale, views seen by the robot at long distances in the map permit to compute a more feasible localization.

5.3 Experimental Set 2

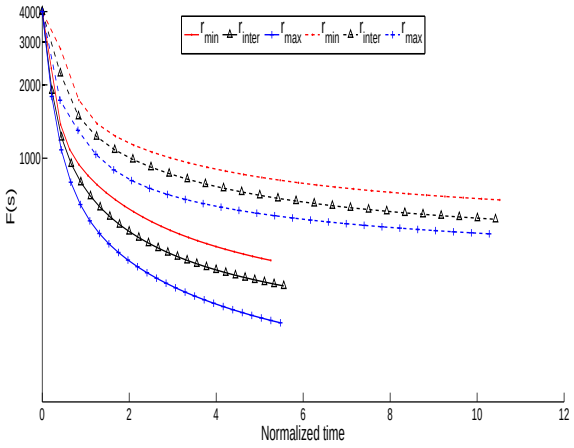
The purpose of this experiment is to deal with a more realistic situation, defined through an office-like environment. Typical obstructor elements like walls or doors have been introduced. Figure 4 describes an environment of 20x40m which the robot moves through. The continuous line represents the real path followed by the robot, the dash-dot line shows the odometry, whereas the estimated solution with our approach is



(a)



(b)



(c)

Figure 3: Accumulated error probability $F(s)$ versus time in a SLAM experiment. Continuous lines show results provided by the proposed solution whereas the dash lines show results provided by a basic SGD solution. Figures 3(a), 3(b) and 3(c) show the results when the number of views observed by the robot is $N = 2$, $N = 4$ and $N = 8$ respectively. Different length observation ranges r_{min}, r_{iter} and r_{max} are defined.

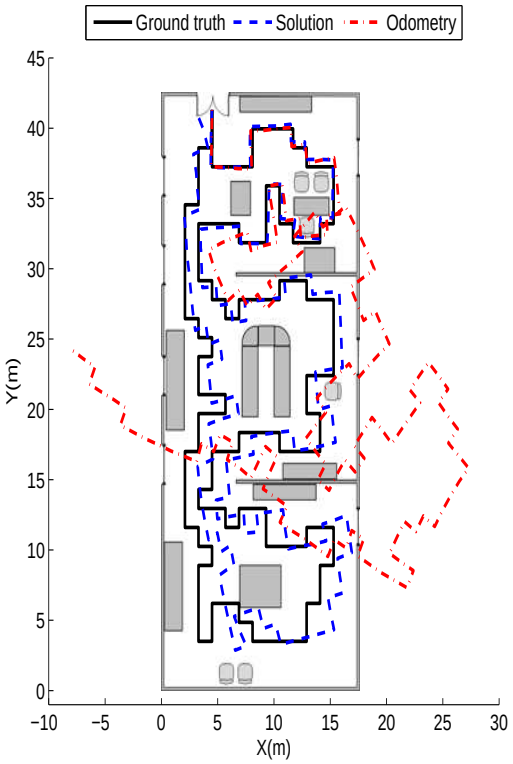


Figure 4: SLAM results in an office-like environment of 20x40m. Continuous line shows the ground truth, dash-dot line the odometry and dash line the estimated solution.

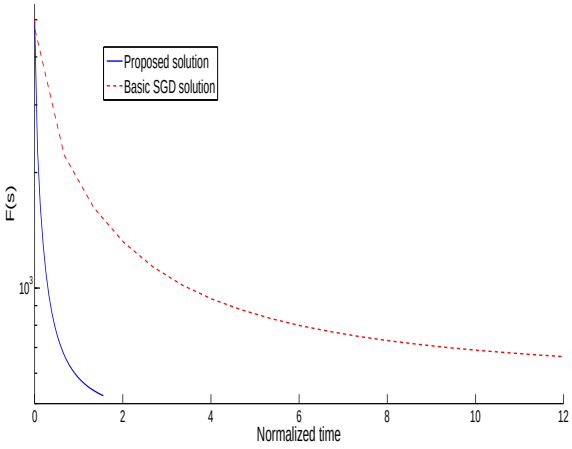


Figure 5: Comparison of the accumulated error probability $F(s)$. Continuous line and dash line show F(s) provided by the proposed approach and a basic SGD algorithm, respectively.

described by a dash line. It may be noticed that in only 15 iterations of the algorithm the robot is able to estimate a quiet reliable solution, whose topology follows the real path's. On the other hand, the error of the odometry grows out of bounds. Figure 5 shows a comparison of the evolution of the accumulated error

$F(s)$ versus the time. Once again, it is revealed the improved capability of our approach to quickly reach a solution, thus involving a better efficiency. In this particular case, it can be observed that our approach requires approximately less than 6 times the computational needs of a basic SGD to reach an optimum value.

6 CONCLUSIONS

This work has presented an approach to the SLAM problem by introducing a SGD algorithm adapted to work with visual observations. The adoption of SGD has been intended to overcome instabilities and harmful effects which compromise the convergence of traditional SLAM algorithms such as filters like EKF. These problems are mainly caused by a visual observation, whose nature is non-linear, especially distinctive in omnidirectional observations. We rely on a visual SLAM approach which builds a map by using a reduced set of omnidirectional image views. The retrieval of the pose of the robot is provided by a single-computation of two views, which directly provides a motion transformation to relate two poses of the robot. A basic SGD algorithm has been modified to work with such unscaled observation model. In addition, a new strategy has been defined to utilize several constraints simultaneously into the same SGD iteration, instead of utilizing only one per iteration as former SGD does. We have shown SLAM results that prove the validity of this new approach with omnidirectional observations. Moreover, we have established a comparison between the results obtained with our approach and those obtained with a basic SGD algorithm. Consequently, the suitability of the approach and its performance has been demonstrated, but also its benefits in the sense of efficiency.

REFERENCES

Andrew J. Davison, A. J., Gonzalez Cid, Y., and Kita, N. (2004). Improving data association in vision-based SLAM. In *Proc. of IFAC/EURON*, Lisboa, Portugal.

Civera, J., Davison, A. J., and Martínez Montiel, J. M. (2008). Inverse depth parametrization for monocular slam. *IEEE Trans. on Robotics*.

Davison, A. J. and Murray, D. W. (2002). Simultaneous localisation and map-building using active vision. *IEEE Trans. on PAMI*.

Gil, A., Reinoso, O., Ballesta, M., Juliá, M., and Payá, L. (2010). Estimation of visual maps with a robot network equipped with vision sensors. *Sensors*.

Grisetti, G., Stachniss, C., and Burgard, W. (2009). Nonlinear constraint network optimization for efficient map learning. *IEEE Trans. on Intelligent Transportation Systems*.

Grisetti, G., Stachniss, C., Grzonka, S., and Burgard, W. (2007). A tree parameterization for efficiently computing maximum likelihood maps using gradient descent. In *Proc. of RSS*, Atlanta, USA.

Jae-Hean, K. and Myung Jin, C. (2003). Slam with omnidirectional stereo vision sensor. In *Proc. of the IROS*, Las Vegas, USA.

Joly, C. and Rives, P. (2010). Bearing-only SAM using a minimal inverse depth parametrization. In *Proc. of ICINCO*, Funchal, Madeira, Portugal.

Montemerlo, M., Thrun, S., Koller, D., and Wegbreit, B. (2002). FastSLAM: a factored solution to the simultaneous localization and mapping problem. In *Proc. of the 18th national conference on Artificial Intelligence*, Edmonton, Canada.

Neira, J. and Tardós, J. D. (2001). Data association in stochastic mapping using the joint compatibility test. *IEEE Trans. on Robotics and Automation*.

Olson, D., Leonard, J., and Teller, S. (2006). Fast iterative optimization of pose graphs with poor initial estimates. In *Proc of ICRA*, Orlando, USA.

Stachniss, C., Grisetti, G., Haehnel, D., and Burgard, W. (2004). Improved Rao-Blackwellized mapping by adaptive sampling and active loop-closure. In *Proc. of the SOAVE*, Ilmenau, Germany.

Valiente, D., Gil, A., Fernández, L., and Reinoso, O. (2012). View-based maps using omnidirectional images. In *Proc. of ICINCO*, Rome, Italy.