Cooperative Robot Localization and Target Tracking based on Least Squares Minimization

Aamir Ahmad[1], Gian Diego Tipaldi[2], Pedro Lima[1] and Wolfram Burgard[2]

***Abstract*—In this paper we address the problem of cooperative localization and target tracking with a team of moving robots. We model the problem as a least squares minimization problem and show that this problem can be efficiently solved using sparse optimization methods. To achieve this, we represent the problem as a graph, where the nodes are robot and target poses at individual time-steps and the edges are their relative measurements. Static landmarks at known position are used to define a common reference frame for the robots and the targets. In this way, we mitigate the risk of using measurements and state estimates more than once, since all the relative measurements are i.i.d. and no marginalization is performed. Experiments performed using a set of real robots show higher accuracy compared to a Kalman filter.**

I. Introduction

Robot localization and target tracking have been areas of extensive research in mobile robotics. When using mobile platforms, these two problems are entangled together since the position of the robots in the environment influences the estimated position of the targets due to the absence of absolute measurements.

Several techniques based on probabilistic filtering have been proposed to address these two problems simultaneously. Single robot approaches have been extensively used, especially for tracking people [1], [2]. Multi-robot teams have also received a lot of attention in addressing this problem in a cooperative fashion, due to the better performance in terms of accuracy and efficiency [3], [4], [5]. The major idea behind cooperative localization and tracking is to explicitly share information among the robots. The idea is to combine localization estimates and target observations to enrich the individual robot beliefs. When sharing robot beliefs, care has to be taken to measure the reliability of the sensed information [6], [7] and to avoid reusing information [8], [4], as this might lead to overconfidence and filter divergence.

All the aforementioned techniques rely on the use of a probabilistic filter. In this paper, we show that least squares minimization approaches can provide more accurate results than filtering methods without any increase in computation time. We formalize the multi-robot cooperative localization and tracking problem within a graph optimization framework, where the tracked objects are treated as a moving landmarks and their states (positions and velocities) are included in the parameters to be estimated. The basic idea of having moving landmarks is not new and has occasionally appeared in the literature [9], [10]. However, the concept has been mainly used to separate static from dynamic landmarks (e.g., a piece of furniture that could be moved around in the office). The novelty of the approach described here lies in the use of dynamic objects as means to improve the localization of the robot team and the targets. Multiple observations of the targets correlate with the estimate of the poses of the robots, thus improving the accuracy of each individual robot estimate. More specifically our approach of multi-robot moving landmark graph optimization (MMG-O) consists of the following steps:

- We create a graph representing poses of the robots and positions and velocities of the targets (nodes) and the observations made by the robots (edges).
- We stack all the observations together to create a single non-linear least squares error function.
- We use a state-of-the-art non-linear least squares solver [11] to calculate the minimum of this error function.

The rest of this paper is organized as follows. After discussion related work in the following section, Section III provides a brief overview of graph based optimization and an in-detail description of our MMG-O approach. In Section IV we then present our experimental results obtained with real robots.

II. Related Work

In the previous decade, several techniques for data fusion and cooperation for multi-robot systems have been investigated. Ong *et al.* [3] introduce a decentralized particle filter (DPF) where the particles are transformed into Gaussian mixture models for communication and fusion. It ensures reduced bandwidth usage and at the same time provides to the whole team a summary of the individual beliefs of the robots. The approach has been later extended [12] to account for correlations between measurements. Those correlations typically arise from the common past information being communicated among the robots in a team. Gohring and Burkhard [5] present a novel approach for cooperative target tracking and localization using visual relations between static and dynamic objects. These relations are independent from the position of the robots in the environment and thus resilient to localization errors. The method, however, is limited to small or limited environments or open space, since the robots need to perceive the static objects continuously to extract the relations. In our previous

[1]A. Ahmad and P. Lima are with the Institute for Systems and Robotics, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal {aahmad,pal}@isr.ist.utl.pt

[2]G. D. Tipaldi and W. Burgard are with the Autonomous Intelligent System Laboratory, Department of Computer Science, Albert-Ludwigs-Universität Freiburg, Germany {tipaldi,burgard}@informatik.uni-freiburg.de

work [7] we modified the particle filter to handle complete and partial occlusions as well as unreliable measurements.

Mourikis and Roumeliotis [13] presented a performance analysis for different existing cooperative localization methods. They derive an upper bound for the position uncertainties as a function of the sensors characteristics and the structure of the graph of relative measurements. They finally demonstrate that an EKF on the joint space of target and robot positions leads to improved accuracy. Whereas our approach is somewhat similar to theirs, the major difference lies in the full least squares formulation instead of a filtering approach. Howard *et al.* [14] introduced a maximum likelihood estimation approach for multi-robot cooperative localization. They combine all the relative measurements between robots in a least squares formulation and use numerical optimization to find their position. We extend their work by including moving targets in the least squares problem. The idea of using teammates as landmarks was further developed by Martinelli *et al.* [15], who used an extended Kalman filter where relative robot observations are treated as additional measurements in the filter. Bahr *et al.* [8] propose a solution for inconsistent estimates that occur due to re-sharing of old data among robots in a team while performing cooperative localization. Jung *et al.* [16] explore an active cooperative target tracking approach, where the objective becomes to control the robot team to move to the next best sensing locations. The approach has been extended by Zhou and Roumeliotis [4] by reducing the complexity from being exponential to linear w.r.t. the number of robots, while keeping a good target accuracy.

Similar ideas for simultaneous localization and tracking have also been extensively studied in a single robot context. A conditional particle filter for localizing a mobile robot in a known environment while at the same time tracking multiple people was presented by Montemerlo *et al.* [2]. This work has later been extended by Schulz *et al.* [1], where a sample-based joint probabilistic data association filters have been presented to deal with uncertainty in data association.

III. Multi-Robot Moving Landmark Graph Optimization

In this paper, we follow the notation of Kümmerle *et al.* [11] to provide a mathematical description of the MMG-O problem in terms of least squares minimization of a sparse system. Let N be the number of robots tracking O objects, treated as a moving landmarks, for T time-steps in an environment containing L static landmarks. Note that, without loss of generality, the formulation can easily be extended for unknown landmarks, given known data associations.

Let $\mathbf{x}_i^n$ be the pose of the n^{th} robot in the team at the i^{th} time-step, $\mathbf{o}_i^o$ the position and velocity of the o^{th} object at time-step i and $\mathbf{l}^l$ the position of the l^{th} landmark. Let $\mathbf{z}_{i,i+1,n}$ be the mean and $\mathbf{\Omega}_{i,i+1,n}$ be the information matrix of a virtual measurement between the pose at time i and $i+1$ of robot n. This can be either coming from odometry or local matching algorithms. Let $\hat{\mathbf{z}}^r(\mathbf{x}_i^n, \mathbf{x}_{i+1}^n)$ be the prediction of the measurement between them. The error term $\mathbf{e}_{i,i+1,n}^r$ described

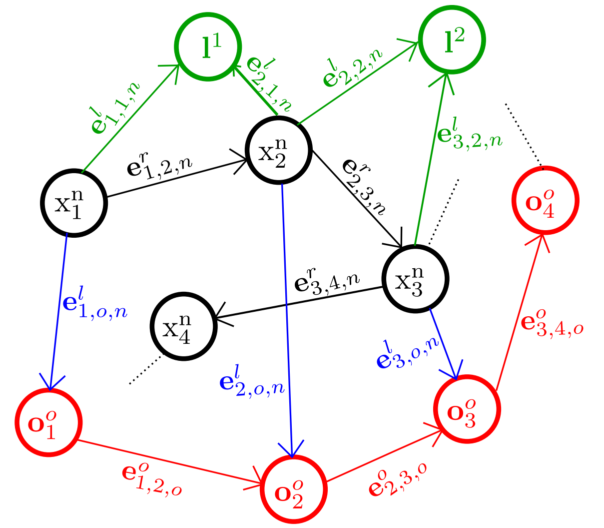


Fig. 1. An example of the pose-graph representation of the MMG-O depicting the n^{th} robot in the team, the o^{th} object (moving landmark), two static landmarks numbered 1 and 2 and the edges connecting all these nodes.

by the edge is

$$\mathbf{e}_{i,i+1,n}^r = \mathbf{z}_{i,i+1,n} - \hat{\mathbf{z}}^r(\mathbf{x}_i^n, \mathbf{x}_{i+1}^n). \tag{1}$$

Let $\mathbf{z}_{i,l,n}$ be the mean and $\mathbf{\Omega}_{i,l,n}$ be the information matrix of a virtual measurement of the landmark l from robot n at time i. Let $\hat{\mathbf{z}}^l(\mathbf{x}_i^n, \mathbf{l}^l)$ be the prediction of the measurement. The error term $\mathbf{e}_{i,l,n}^l$ of the edge is

$$\mathbf{e}_{i,l,n}^l = \mathbf{z}_{i,l,n} - \hat{\mathbf{z}}^l(\mathbf{x}_i^n, \mathbf{l}^l). \tag{2}$$

Similarly, the error $\mathbf{e}_{i,o,n}^l$ between the object o and the robot n at time i is

$$\mathbf{e}_{i,o,n}^l = \mathbf{z}_{i,o,n} - \hat{\mathbf{z}}^l(\mathbf{x}_i^n, \mathbf{o}_i^o). \tag{3}$$

The motion of the objects is modeled using a constant velocity motion model with random acceleration, which leads to the following term $\mathbf{e}_{i,i+1,o}^o$ for the edge between consecutive object positions:

$$\mathbf{e}_{i,i+1,o}^o = \mathbf{o}_{i+1}^o - \mathbf{A}\mathbf{o}_i^o - \nu, \tag{4}$$

where $\mathbf{A}$ is the matrix modeling discrete-time constant velocity and ν is a zero mean error with information matrix $\mathbf{\Omega}_{i,i+1,o}$.

Now let $\mathbf{x}$ be the vector obtained by stacking all the variables. The solution of the MMG-O is the value $\mathbf{x}^*$ that minimizes the following function

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \mathbf{F}(\mathbf{x}), \tag{5}$$

where

$$
\begin{aligned}
\mathbf{F}(\mathbf{x}) \quad = \quad & \sum_{n=1}^{N}\sum_{i=1}^{T-1} {\mathbf{e}^{r}_{i,i+1,n}}^{\top} \boldsymbol{\Omega}_{i,i+1,n} \mathbf{e}^{r}_{i,i+1,n} \\
& + \sum_{n=1}^{N}\sum_{l,i\in\mathcal{C}_l} {\mathbf{e}^{l}_{i,l,n}}^{\top} \boldsymbol{\Omega}^{l}_{i,l,n} \mathbf{e}^{l}_{i,l,n} \\
& + \sum_{n=1}^{N}\sum_{o,i\in\mathcal{C}_o} {\mathbf{e}^{l}_{i,o,n}}^{\top} \boldsymbol{\Omega}^{o}_{i,m,n} \mathbf{e}^{l}_{i,o,n} \\
& + \sum_{o=1}^{O}\sum_{i=1}^{T-1} {\mathbf{e}^{o}_{i,i+1,o}}^{\top} \boldsymbol{\Omega}^{o}_{i,i+1,o} \mathbf{e}^{o}_{i,i+1,o}. \qquad (6)
\end{aligned}
$$

Here, $\mathcal{C}_l$ and $\mathcal{C}_o$ are respectively the set of all the observations between any robot and any static or moving landmark. Fig. 1 provides a visual explanation on how the graph is built. While green nodes represent landmark locations, black nodes indicate the robot poses and red nodes stand for target positions. The red arrows indicate the motion of the individual targets. Robot observations of static and moving objects are shown respectively with green and blue arrows.

The formulation of the final objective function (6) is similar to the one obtained by Grisetti *et al.* [17] and Kümmerle *et al.* [11]. Hence, following their procedure, we can approximate it using a Taylor expansion around an initial guess for $\mathbf{x}$ and applying error minimization via iterative local linearizations or least squares minimization on a manifold.

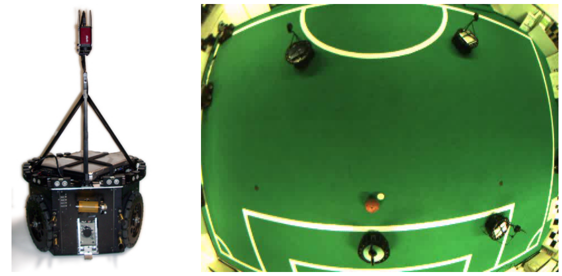
IV. Experiments and Results

We implemented the MMG-O approach using the g^{2}o framework and applied it within a RoboCup Soccer Middle Sized League scenario, where the primary requisites for a team of robots successfully playing soccer are : i) accurate localization for the individual robots and ii) continuous tracking of the soccer ball. The localization and/or ball tracking failures in such a scenario owe much to the field size and symmetry, limited sensor range, occlusions as well as dynamic and fast movements of the robots and/or the ball. This makes the robot soccer scenario an interesting and suitable test bed for evaluating the performance of our MMG-O approach.

A. Test Bed and Experimental Scenario

We used a team of four robots (OMNI1 to OMNI4) equipped with a three-wheeled omni directional drive and a dioptric vision system consisting of a fish-eye lens facing downwards (Fig. 2(a)). Each camera has of resolution of 640x480 pixels. The team of robots was placed in one half of the field (Fig. 2(b)) and was controlled remotely by human users. An orange colored soccer ball was manually moved around in the field. The odometry data of the robots and the images acquired from their cameras were logged on a laptop mounted on each individual robot and used for the experiments.

The experiment lasted for approximately 6 minutes. The odometry data was saved at 40 Hz and the images were stored at approximately 25 Hz. A pair of overhead stereo cameras were used to obtain the ground truth positions of the robots and the ball at 25 Hz. Due to the system used to measure the ground truth, only the position of the ball and the robots could be recorded and neither their orientations nor their velocities. All the observations and the ground-truth data were time-synchronized. A pre-processing step was performed on the images saved by the robots to compute the observations (range and bearing) of six known static landmarks and one orange ball. The graph was initialized using the odometry measurements and the ball observations.



(b)

(b)

Fig. 2. An OMNI robot (a) and a typical scenario on the soccer field (b).

We evaluated the result provided by MMG-O against the ground-truth and compared it with an EKF for cooperative localization and object tracking, similar to the one presented by Mourikis and Roumeliotis [13]. In the EKF, the state vector estimated at any time-step i consists of all the poses of the robots $(\mathbf{x}^1_i, \ldots, \mathbf{x}^{\mathrm{N}}_i)^\top \in \mathbb{R}^{3\times \mathrm{N}}$ and the 2D position and velocity $\mathbf{o}^{\mathrm{ball}}_i \in \mathbb{R}^4$ of the ball. In the prediction step we used the odometry information for the robot poses and for the ball positions we implemented a constant velocity and zero mean white Gaussian acceleration noise model [18]. In the update step, observations of static landmarks and the ball,coming from each robot, are synchronized to obtain a joint measurement vector.

In subsection IV-B we present the results of the MMG-O compared with the EKF on the data from the two robots OMNI1 and OMNI2 for localization and ball tracking. The aim of this experiment is to present a proof of concept of the MMG-O and the higher accuracy achieved by it compared with the EKF based approach.

In subsection IV-C we extended the comparison to all the four robots OMNI1 to OMNI4. The aim here is to show the scalability of our approach while still achieving a higher degree of accuracy over the EKF=based approach.

In the video accompanying this paper we present the results of the MMG-O experiment overlaid on the stream of images from the stereo camera used to compute the ground-truth. Each robot has a uniquely colored plate placed on top of it (OMNI1:Magenta, OMNI2:Brown, OMNI3:Red and OMNI4:Blue). The stereo camera uses these plates for detecting the positions of the robots. The ground-truth is marked in the video frames with an overlaid black circle around the plates on top of the robots and the orange ball except for the

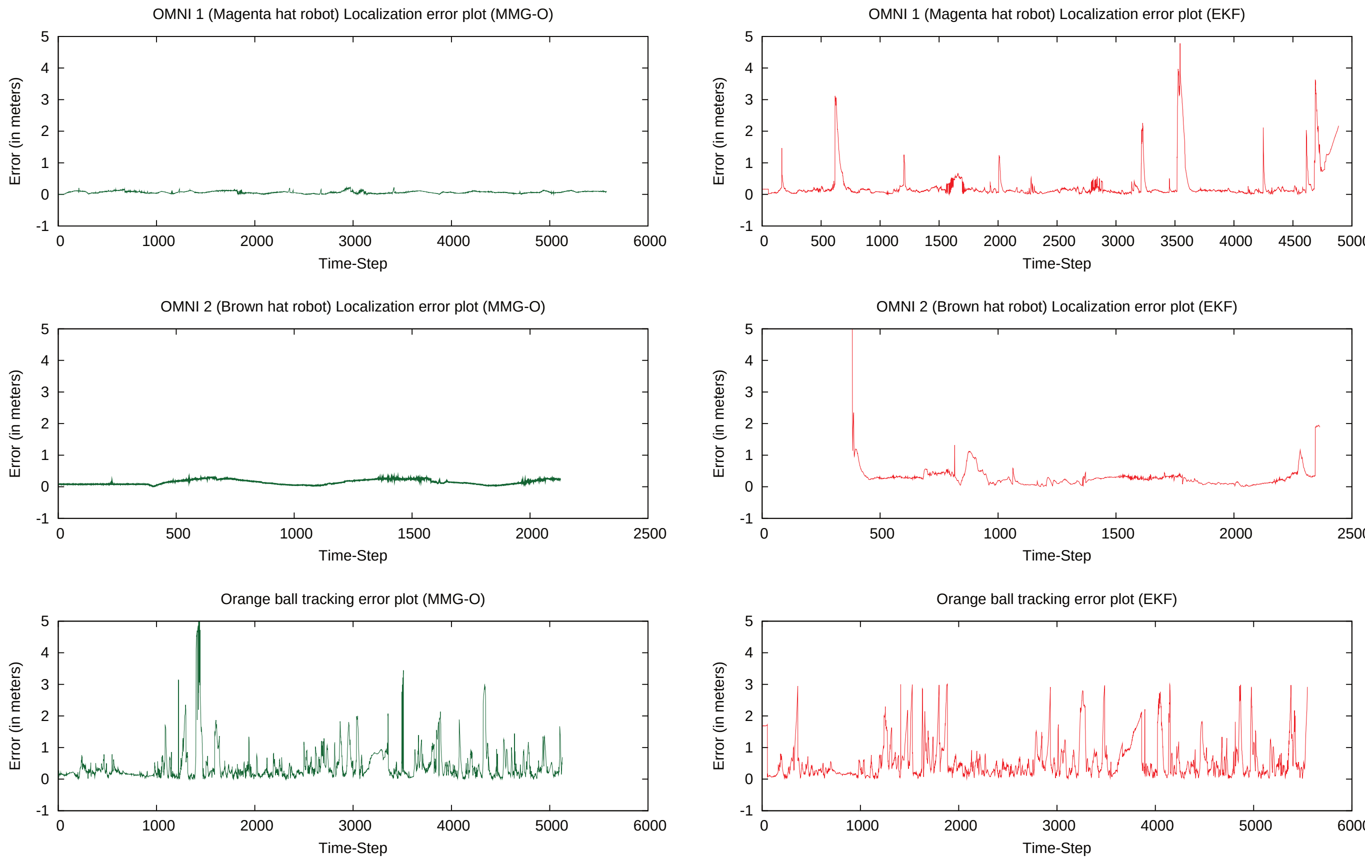


Fig. 3. Experiment with two robots: The X-axis in these plots represent the number of time-steps between two consecutive frames. The Y-axis represents the root mean square error between the ground-truth and the MMG-O (left) or EKF (right) estimate of the robots and the ball. Note that only the time-steps when the ground-truth was available are shown in the plots.

TABLE I
STATISTICAL ESTIMATES OF THE 2 ROBOTS EXPERIMENT RESULTS.

	MMG-O			Cooperative EKF		
	Mean (m)	Median (m)	Variance (m^2)	Mean (m)	Median (m)	Variance (m^2)
OMNI1	0.069	0.066	0.001	0.274	0.123	0.272
OMNI2	0.147	0.127	0.006	0.294	0.261	0.077
Ball	0.426	0.240	0.314	0.583	0.356	0.357

instances during which the markers of the robots or the ball could not be detected due to occlusions. The estimates of the robot positions are marked in the video frames by overlaying a circle of the same color as the plate on to of the robot. These circles are centered at the position estimates for the particular robot and are placed at the same height as of the plate on top of the robot from the ground level to facilitate easy visual comparison. The estimate of the orientations of the robots are indicated by a radial line from the center of the circle. The position estimates of the orange ball are marked by an orange circle overlaid to the video frames.

B. Experiment with Two Robots

Here, we present the results obtained by applying MMG-O and the EKF to the data of OMNI1 and OMNI2. The green-colored plots in the left column of Fig. 3 display the error of the positions of the robots and the ball as estimated by MMG-O with respect to the ground-truth, while the red-colored plots in the right column of Fig. 3 display the respective errors of the EKF based approach. Table I shows the mean and variance of the errors presented in Fig. 3.

From Fig. 3 and Table I we can infer that MMG-O is able to reduce the mean error in position estimates for OMNI1 by a factor of 4, for OMNI2 by a factor of 2, and for the orange ball by a factor of 1.3 compared to the EKF-based approach. Please note that the difference in accuracy is mainly due to the different path of the two robots, showing the benefit of the iterative relinearization of MMG-O.

C. Experiment with Four Robots

The extension of the MMG-O to the four robots case shows more interesting results. Figure 4 presents the error plots of all the four robots and the orange ball for both the MMG-O (green-colored plots on the left column) and the EKF-based approach (red-colored plots on the right column). In the EKF-based approach, the robots (most notably OMNI4) often tend to loose their position due to noisy observations and odometry measurements. However, as soon as better measurements

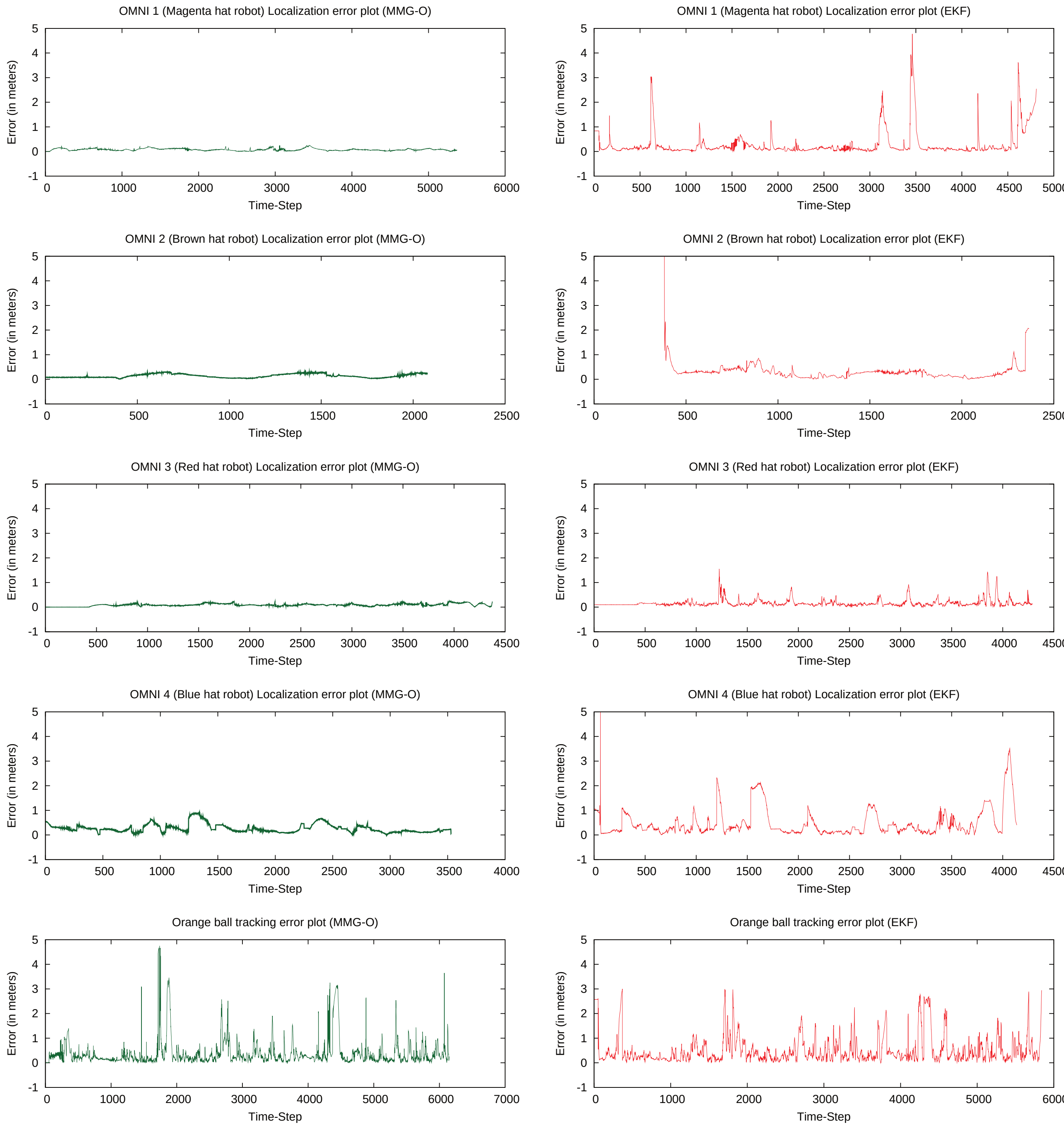


Fig. 4. Experiment with four robots: The X-axis in these plots represent the number of time-steps between two consecutive frames. The Y-axis represents the root mean square error between the ground-truth and the MMG-O (left) or EKF (right) estimate of the robots and the ball. Note that only the time-steps when the ground-truth was available are shown in the plots.

arrive, the filter is updated correctly and the error is reduced. The effect of noise is mitigated in MMG-O thanks to the iterative re-linearizations and the batch processing. This can be observed in the error plots for the EKF-based approach in Fig. 4, when compared to MMG-O. The latter is able to maintain an acceptable estimate for all these robots even in the case of highly noisy measurements and observations. From Fig. 4 and Table II we can infer that MMG-O outperforms

TABLE II
STATISTICAL ESTIMATES FOR THE EXPERIMENTS WITH FOUR ROBOTS.

	MMG-O			Cooperative EKF		
	Mean (m)	Median (m)	Variance (m^2)	Mean (m)	Median (m)	Variance (m^2)
OMNI1	0.077	0.070	0.002	0.298	0.118	0.295
OMNI2	0.141	0.120	0.006	0.296	0.265	0.075
OMNI3	0.090	0.086	0.003	0.166	0.124	0.024
OMNI4	0.267	0.221	0.030	0.493	0.276	0.327
Ball	0.399	0.226	0.327	0.528	0.315	0.334

TABLE III
COMPARISON OF THE TOTAL COMPUTATION TIME TAKEN BY MMG-O AND THE COOPERATIVE EKF APPROACH ON THE FULL DATASET.

	MMG-O	EKF
2 Robot Exp.	13.75s	13.38s
4 Robot Exp.	19.38s	137.95s

the EKF-based approach by a factor of 3.8 for OMNI1, 2.1 for OMNI2, 1.8 for OMNI3, 1.8 for OMNI4 and 1.3 for the orange ball.

D. Computation Time Comparison

For a fair comparison of the computation time, we ran both, the MMG-O and the EKF-based approach, on the same machine (Quad-Core Intel(R) Core(TM) i5 CPU 750 with 2.67GHz and 8GB RAM). Table III presents the total time taken by both implementations on the same dataset. For MMG-O, this reflects the total time taken by the g^2o for the whole graph of the dataset, while it shows the total time taken to iterate once over the whole dataset for the EKF-based approach. The graph for MMG-O consists of 50.000 nodes in the case of four robots and 30.000 in the case of two robots, requiring respectively 87 and 60 iterations to converge. By exploiting the sparsity of the MMG-O graph, our approach is able to scale linearly with respect to the number of robots and targets, while in the EKF formulation the required computation grows quadratically. This makes the use of MMG-O an appealing alternative to EKF for cooperative localization and tracking, being MMG-O more accurate and more scalable with increasing numbers of robots.

V. CONCLUSION

In this paper we presented a novel approach for cooperative localization of a team of robots while jointly tracking moving targets. We model the problem as graph-based optimization, where the poses of the robots, of the moving targets and of the static landmarks are jointly estimated in a least squares minimization framework. We presented a mathematical formulation of this problem described an implementation using g^2o. We tested our approach in a robot soccer scenario and compared it to an EKF-based approach. The results show that our approach leads to increased accuracy in the estimation and to an improved scalability in scenarios in which a higher number of robots is required. In the future, we plan to extend our approach so that it can also deal with unknown data associations and heterogeneous robots with different sensors.

ACKNOWLEDGEMENT

The work described in this paper was partially funded by the Portuguese Government through FCT Project PTDC/EEA-CRO/100692/2008, by the European project RoCKIn: FP7-EU-601012 and by Fundação para a Ciência e a Tecnologia (ISR/IST pluriannual funding) through the PIDDAC Program funds. It has also been supported by the European Commission under the ERC Advanced Grant ERC-267686-LifeNav.

REFERENCES

[1] D. Schulz, W. Burgard, D. Fox, and A. Cremers, "People tracking with mobile robots using sample-based joint probabilistic data association filters," *Int. Journal of Robotics Research*, vol. 22, no. 2, pp. 99–116, 2003.

[2] M. Montemerlo, S. Thrun, and W. Whittaker, "Conditional particle filters for simultaneous mobile robot localization and people-tracking," in *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, vol. 1. IEEE, 2002, pp. 695–701.

[3] L.-L. Ong, B. Upcroft, T. Bailey, M. Ridley, S. Sukkarieh, and H. Durrant-Whyte, "A decentralised particle filtering algorithm for multi-target tracking across multiple flight vehicles," in *IEEE/RSJ Int. Conf. on Intel. Rob. and Sys. (IROS)*, 2006.

[4] K. Zhou and S. Roumeliotis, "Multirobot active target tracking with combinations of relative observations," *IEEE Transactions on Robotics*, vol. 27, no. 4, pp. 678 –695, 2011.

[5] D. Gohring and H.-D. Burkhard, "Multi robot object tracking and self localization using visual percept relations," in *IEEE/RSJ Int. Conf. on Intel. Rob. and Sys. (IROS)*, 2006.

[6] L. Marchetti, D. Nobili, and L. Iocchi, "Improving tracking by integrating reliability of multiple sources," in *Int. Conf. on Information Fusion*, 2008, pp. 1–8.

[7] A. Ahmad and P. U. Lima, "Multi-robot cooperative object tracking based on particle filters," in *European Conference on Mobile Robots (ECMR)*, 2011.

[8] A. Bahr, M. Walter, and J. Leonard, "Consistent cooperative localization," in *IEEE Int. Conf. on Rob. & Aut. (ICRA)*, 2009.

[9] J. Rogers, A. Trevor, C. Nieto-Granda, and H. Christensen, "Slam with expectation maximization for moveable object tracking," in *IEEE/RSJ Int. Conf. on Intel. Rob. and Sys. (IROS)*, 2010.

[10] C.-C. Wang, C. Thorpe, S. Thrun, M. Hebert, and H. Durrant-Whyte, "Simultaneous localization, mapping and moving object tracking," *Int. Journal of Robotics Research*, vol. 26, no. 9, pp. 889–916, 2007.

[11] R. Kümmerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g2o: A general framework for graph optimization," in *IEEE Int. Conf. on Rob. & Aut. (ICRA)*, 2011.

[12] L.-L. Ong, T. Bailey, H. Durrant-Whyte, and B. Upcroft, "Decentralised particle filtering for multiple target tracking in wireless sensor networks," in *Int. Conf. on Information Fusion*, 2008.

[13] A. Mourikis and S. Roumeliotis, "Performance analysis of multirobot cooperative localization," *IEEE Transactions on Robotics*, vol. 22, no. 4, pp. 666–681, 2006.

[14] A. Howard, M. Matark, and G. Sukhatme, "Localization for mobile robot teams using maximum likelihood estimation," in *IEEE/RSJ Int. Conf. on Intel. Rob. and Sys. (IROS)*, 2002.

[15] A. Martinelli, F. Pont, and R. Siegwart, "Multi-robot localization using relative observations," in *IEEE Int. Conf. on Rob. & Aut. (ICRA)*, 2005.

[16] B. Jung and G. Sukhatme, "Cooperative multi-robot target tracking," *Distributed Autonomous Robotic Systems 7*, pp. 81–90, 2006.

[17] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, "A tutorial on graph-based SLAM," *IEEE Transactions on Intelligent Transportation Systems Magazine*, vol. 2, pp. 31–43, 2010.

[18] W. Blair, "Design of nearly constant velocity track filters for tracking maneuvering targets," in *Int. Conf. on Information Fusion*, 2008.