Collaborative Localization of Robotic Wheeled Walkers using Interlaced Extended Kalman Filters

Daniele Fontanelli, David Macii
DII - University of Trento
Via Sommarive 9, 38123, Trento, Italy
Email: {david.macii,daniele.fontanelli}@unitn.it

Payam Nazemzadeh, Luigi Palopoli
DISI - University of Trento
Via Sommarive 5, 38123, Trento, Italy
Email: {payam.nazemzadeh,luigi.palopoli}@unitn.it

***Abstract*—This paper deals with the problem of localizing a group of robots. Each robot is equipped with an autonomous self-localization system based on both high-rate odometers and a front camera detecting sporadically simple visual landmarks placed on the floor at known positions and with a known attitude within a given 2D reference frame. In addition, the robots can share their own states to improve localization accuracy. In this paper, after reviewing the general problem of collaborative localization, two potential strategies based on a distributed Interlaced Extended Kalman Filter (IEKF) are compared. In the first one, localization is refined through high-rate mutual measurements of Euclidean distance between pairs of robots. The distance values are obtained using an omni-directional wireless ranging system of limited accuracy. In the second one, an on-board Kinect-like front RGB-D vision system is assumed to measure both the agent's distance and its relative orientation with respect to other robots in view at a lower rate, but with higher accuracy. To the best of authors' knowledge, this performance comparison is new, since most of previous studies on collaborative localization are overoptimistic as implicitly assume that mutual agent measurements are not intermittent and are available at all sampling times.**

***Keywords*—Localization, position measurements, sensor fusion, estimation, distributed systems, Kalman filters.**

I. Introduction

Indoor localization of mobile platforms is a well-known measurement problem that has been tackled using different technologies, depending on both the nature of the targets to be tracked and the level of accuracy required. Since absolute localization based on Global Navigation Satellite Systems (GNSS) can be hardly used indoors, several alternative techniques have been proposed over the last few years. They range from low-cost (but quite inaccurate) solutions based on Received Signal Strength (RSS) measurements and fingerprinting of standard wireless signals (e.g. WiFi) [1], to expensive systems able to measure the distance from a target with uncertainty of a few cm using the time-of-flight (ToF) values of Ultra-Wide Band (UWB) pulses [2], laser beams [3] or ultrasonic stimuli [4], [5]. In the case of wheeled devices, such as robotic vehicles, localization can also benefit from the use of dead reckoning techniques (e.g. based on odometry or inertial measurement units), which can be combined with absolute positioning data (e.g. collected by vision systems [6]) through proper fusion algorithms [7]–[9].

The localization of multiple robotic platforms (henceforth referred to as *agents*) is typically a difficult problem if it is addressed in a centralized way, especially in non-line-of-sight (NLOS) conditions and when multiple targets are present in the same environment. However, if the problem is tackled in a distributed way, with each agent being endowed of an autonomous localization system, the presence of multiple *agents* can possibly turn into an advantage, provided that different robots are able to cooperate. This general idea, often called *synergic* or *collaborative* localization, has proved to be successful in different contexts involving groups of robots [10], [11].

An early study on synergic localization is presented in [12]. In [13] the authors envision a fully wireless synergic localization system based on the potential ability of clusters of 4G mobile devices to measure their reciprocal distances through a hybrid time of arrival/angle of arrival (TOA/AOA) technique. The case of collaborative localization of wireless mobile platforms has been also addressed in [14], where the so-called *parallel projection method* is used to improve localization accuracy in NLOS conditions. Taniuchi et al. suggest using a spring model to reduce the pose estimation uncertainty associated with distance measures obtained using WiFi and Bluetooth RSS data [15].

In the field of robotics, Fox et al. propose a Markov-based probabilistic method in which each robot's belief about its own position is refined as soon as other robots are detected [16]. A different Markovian approach is adopted in [17]. In this case, first the egocentric measurement data are fused locally to create a Markov chain of robot pose estimates. Then, both inter-robot measurement data and state estimates are transferred to a central server, where localization is refined by minimizing the mean square error of agents' positions.

An alternative statistical method for collaborative localization is instead described in [18]. This relies on a decentralized, real-time particle filter coupled with a reciprocal sampling algorithm to reduce the overall computational burden. In spite of the accuracy improvements achieved by applying computationally demanding optimization strategies to the collaborative localization problem [19], the simplest general technique for fully-distributed, multi-robot localization is still the extended Kalman filter (EKF) [12]. In [20] the update step of a distributed EKF is modified by an algorithm preventing data reuse in order to avoid inconsistent (i.e. overconfident) covariance estimates.

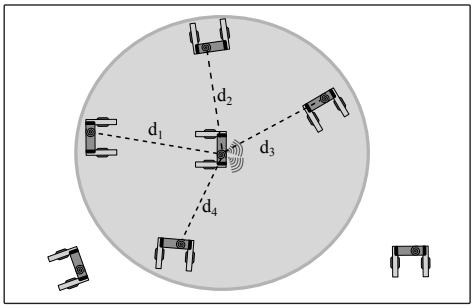
Panzieri et al. address the collaborative localization problem by means of an interlaced extended Kalman filter (IEKF) [21], [22]. The IEKF is inherently distributed, computationally acceptable and easy to implement. Therefore, this

was also used as a starting point for the analysis reported in this paper. However, the implementation of the original IEKF disregards the possibility of both having different sampling rates and intermittent observations. The goal of this paper instead is to investigate to what extent the trade-off between accuracy, and rate of inter-robot relative position measurements can affect localization accuracy. In the following, at first, in Section II, the collaborative localization problem is formalized. Afterwards, in Section III, the IEKF algorithm is recalled and it is generalized in the case of intermittent observations. Section IV reports several simulation results comparing two different techniques to measure the relative position of pairs of agents, i.e. using a low-rate RGB-D camera or, alternatively, a less accurate wireless ranging system collecting data at a higher rate. Finally, Section V concludes the paper and outlines future work.

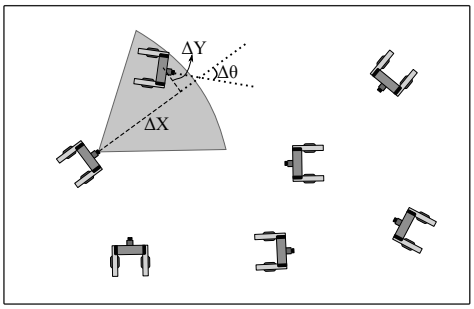
II. Problem formulation and system description

As shortly explained in the Introduction, for *collaborative* or *synergic* localization of a team of agents we mean the ability to refine the position and the heading estimated by each agent in a common reference frame by using both local positioning data and relative distance and/or attitude measures between pairs of devices. The main underlying assumptions are summarized below.

1) N agents can move freely in a large room. The dynamic of each agent does not depend on any other agent, since each user may act independently. The only constraint to robot motion is collision avoidance.
2) The state of each agent i (with $i = 1, \ldots, N$) at time kT_s (T_s being the sampling period used to discretize the system) is represented by vector $\mathbf{p}_k^{(i)} = [x_k^{(i)}, y_k^{(i)}, \theta_k^{(i)}]^T$, where $x_k^{(i)}$ and $y_k^{(i)}$ are the agent's planar coordinates in the chosen reference frame, while $\theta_k^{(i)}$ is the angle between the longitudinal axis of the robot and the $x-$axis of the reference frame.
3) Each agent is able to estimate its own state autonomously (namely without the help of other agents) by fusing odometry data with absolute position and heading measures obtained from special visual landmarks (e.g. QR codes) placed regularly on the floor at the same distance from one another and in such a way that at most one of the them can be detected by a plain monocular camera directed towards ground [23].
4) Besides the sensors used by each robot for its own local state estimation, every agent is supposed to be equipped with two alternative types of exteroceptive sensors for collaborative localization, i.e. either an omni-directional wireless ranging system (*case A*) or a front RGB-D camera (e.g. a Kinect [24]) (*case B*). Both cases are qualitatively shown in Fig. 1(a) and (b), respectively. In *case A* the ranging system is used to measure just the distance between a robot and any other agent located within an (approximately) circular range. On the contrary, in *case B* the stereo vision system is employed to recognize and to measure the relative position between the robot's camera and any other agents located within its detection range.



(a)



(b)

Figure 1. Qualitative overview of the measurement systems mounted on each agent for collaborative localization: an omni-directional wireless ranging systems for distance measurement only (a), and a front RGB-D camera measuring the relative position of two agents (b).

5) All agents are equipped with a radio transceiver ensuring complete connectivity between pairs of robots as well as high-rate and low-latency communication. Long Term Evolution (LTE) wireless modules can indeed meet such requirements [25].

If each robot is modeled as a unicycle-like vehicle, the overall state transition of all agents in the chosen reference frame is given by the following non-linear discrete-time system [21], i.e.

$$\mathbf{p}_{k+1} = \begin{bmatrix} \mathbf{p}_{k+1}^{(1)} \\ \vdots \\ \mathbf{p}_{k+1}^{(N)} \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mathbf{p}_k^{(1)}, \mathbf{u}_k^{(1)}, \boldsymbol{\varepsilon}_k^{(1)}) & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \mathbf{f}(\mathbf{p}_k^{(N)}, \mathbf{u}_k^{(N)}, \boldsymbol{\varepsilon}_k^{(N)}) \end{bmatrix} \tag{1}$$

where

$$\mathbf{p}_{k+1}^{(i)} = \mathbf{f}(\mathbf{p}_k^{(i)}, \mathbf{u}_k^{(i)}, \boldsymbol{\varepsilon}_k^{(i)}) = \mathbf{p}_k^{(i)} + G_k^{(i)}(\mathbf{u}_k^{(i)} + \boldsymbol{\varepsilon}_k^{(i)}), \quad i = 1, \ldots, N, \tag{2}$$

describes the evolution of the $i-$th robot, $\mathbf{u}_k^{(i)} = [\delta s_k^{(i)}, \delta\theta_k^{(i)}]^T$ is the input vector including the linear and angular displacements of the $i-$th agent at time kT_s, $\boldsymbol{\varepsilon}_k^{(i)}$ is the vector including the respective zero-mean noise terms and, finally,

$$G_k^{(i)} = \begin{bmatrix} \cos\theta_k^{(i)} & 0 \\ \sin\theta_k^{(i)} & 0 \\ 0 & 1 \end{bmatrix}. \tag{3}$$

The observation model associated with system (1) includes two types of output functions, i.e.

- the geometrical relationship between the position and heading of each agent and those of any detected visual landmark in the same reference frame;

- the geometrical relationship between the pose of each robot and those of the other $N-1$ agents in the room.

As a result, the overall observation equation at time kT_s becomes

$$\mathbf{z}_k = \begin{bmatrix} \mathbf{z}_k^{(1)} \\ \vdots \\ \mathbf{z}_k^{(N)} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{h}}^{(1)}(\mathbf{p}_k) \\ \vdots \\ \tilde{\mathbf{h}}^{(N)}(\mathbf{p}_k) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\eta}_k^{(1)} \\ \vdots \\ \boldsymbol{\eta}_k^{(N)} \end{bmatrix}, \tag{4}$$

where vector $\mathbf{z}_k^{(i)} = \left[\mathbf{z}_k^{(i,1)}, \ldots, \mathbf{z}_k^{(i,N)}\right]^T$ includes all possible observations from agent i, $\boldsymbol{\eta}_k^{(i)} = \left[\boldsymbol{\eta}_k^{(i,1)}, \ldots, \boldsymbol{\eta}_k^{(i,N)}\right]^T$ is the vector comprising the respective measurement uncertainty contributions, and, finally,

$$\tilde{\mathbf{h}}^{(i)}(\mathbf{p}_k) = \begin{bmatrix} \mathbf{h}(\mathbf{p}_k^{(i)}, \mathbf{p}_k^{(1)}) \\ \vdots \\ \mathbf{h}(\mathbf{p}_k^{(i)}, \mathbf{p}_k^{(i-1)}) \\ \boldsymbol{o}(\mathbf{p}_k^{(i)}) \\ \mathbf{h}(\mathbf{p}_k^{(i)}, \mathbf{p}_k^{(i+1)}) \\ \vdots \\ \mathbf{h}(\mathbf{p}_k^{(i)}, \mathbf{p}_k^{(N)}) \end{bmatrix}, \quad i = 1, \ldots, N, \tag{5}$$

is the vector including all observations from agent i. Note that the $i-$th function of the vector, referred to as $\boldsymbol{o}(\cdot)$, is different from the other elements, as it depends on the geometrical relationship between the position and the heading of each agent and those of a detected visual landmark. On the contrary, each function $\mathbf{h}(\cdot,\cdot)$ consists of M equations and depends on how the state variables of agents $j = 1, \ldots, N$ for $j \neq i$ are actually observed by the $i-$th robot. In the rest of this paper, we will assume that while $\boldsymbol{o}(\cdot)$ is the same in all conditions (see assumption 3), the equations of $\mathbf{h}(\cdot,\cdot)$ differ in *case A* and *case B*, respectively, in accordance with assumption 4.

III. Position estimation algorithm

The nonlinear model based on expressions (1)-(4) can be used to implement an estimation algorithm similar to the IEKF presented in [21]. Since the state evolution (namely the pose) of each robot does not depend on the state of the other robots, the *prediction step* equations of the IEKF are straightforward, as they basically coincide with those of N independent standard EKFs, i.e. [26]

$$\begin{aligned} \hat{\mathbf{p}}_{k+1|k}^{(i)} &= \hat{\mathbf{p}}_{k|k}^{(i)} + G_k^{(i)} \mathbf{u}_k^i \\ \hat{P}_{k+1|k}^{(i)} &= F_k^{(i)} \hat{P}_{k|k}^{(i)} {F_k^{(i)}}^T + G_k^{(i)} Q_k^{(i)} {G_k^{(i)}}^T \end{aligned} \quad i = 1, \ldots, N \tag{6}$$

where $\mathbf{p}_{k+1|k}^{(i)}$ and $\hat{\mathbf{p}}_{k|k}^{(i)}$ denote the predicted and estimated states, respectively, at time kT_s, $P_{k+1|k}^{(i)}$ and $P_{k|k}^{(i)}$ are the corresponding covariance matrices, $F_k^{(i)}$ is the Jacobian of (2) with respect to $\mathbf{p}^{(i)}$ computed at $\hat{\mathbf{p}}_{k|k}^{(i)}$, $G_k^{(i)}$ is defined as in (3), $\mathbf{u}_k^i$ is the vector of the input values at time kT_s and, finally, $Q_k^{(i)}$ is the covariance matrix of $\boldsymbol{\varepsilon}_k^{(i)}$ (i.e. due to odometers). Observe that (6) depends just on local quantities. Therefore, the prediction step equations can be computed locally, i.e. on board of each robot, thus ensuring a fully distributed implementation.

As far as the *update step* is concerned, the corresponding equations are different from those of a standard EKF for two reasons. First of all, due to the definition of the measurement model (4), the updated state estimate of the $i-$th agent and its covariance matrix depend not only on the respective predicted values and on the measurement data, but also on the predicted state and on the predicted covariance of the other agents. Secondly, generally a robot is not able to observe all the other agents simultaneously and can occasionally miss landmarks as well. This means that all observations are inherently intermittent, as they depend on the reading range of the chosen measurement systems and on the distance between each robot and both the other agents and one of the landmarks. In particular, if the distance between pairs of adjacent landmarks is larger than the reading range of the camera, then at most one landmark can be used in the update step. As a result of all the issues above, the *update step* of the proposed IEKF is inherently stochastic in the case considered.

In the classic IEKF formulated in [21], the computation of the innovation term associated with a generic pair of agents (i, j) relies on the predicted state of j regarded as an additional measure. Therefore, both $\hat{\mathbf{p}}_{k+1|k}^{(j)}$ and its covariance matrix $P_{k+1|k}^{(j)}$ have to be transmitted to agent i, thus "interlacing" the two subsystems. In particular, $P_{k+1|k}^{(j)}$ has to be included in the Kalman gain, as it will be shown in the following, to keep into account the fact that $\hat{\mathbf{p}}_{k+1|k}^{(j)}$ is affected by some uncertainty.

The problem of intermittent observations and the related stability issues are deeply analyzed in [27] for simple KFs, but, to the best of authors' knowledge, have never been considered in the case of IEKFs. By extending a similar approach to the system at hand for all pairs of agents i and j, we can define a binary random variable $\gamma_k^{(i,j)}$, to be set to 1 if agent i is able to observe agent j at time kT_s, or 0 otherwise. Similarly, $\gamma_k^{(i,i)}$ is set to 1 if robot i is able to detect a landmark at time kT_s, or 0 otherwise. Starting from the basic *update step* equations of an EKF and assuming to replace the variance of real measurements with a large dummy value anytime $\gamma_k^{(i,j)} = 0$, after some algebraic steps it can be shown that, if the dummy variance tends to infinity [27], then the update equations of the IEKF running on agent i become

$$\begin{aligned} \hat{\mathbf{p}}_{k+1|k+1}^{(i)} &= \hat{\mathbf{p}}_{k+1|k}^{(i)} + K_{k+1}^{(i)} \Gamma_{k+1}^{(i)} [\mathbf{z}_{k+1}^{(i)} - \tilde{\mathbf{h}}^{(i)}(\hat{\mathbf{p}}_{k+1|k})] \\ \hat{P}_{k+1|k+1}^{(i)} &= \hat{P}_{k+1|k}^{(i)} - K_{k+1}^{(i)} \Gamma_{k+1}^{(i)} \tilde{H}_{k+1}^{(i,i)} \hat{P}_{k+1|k}^{(i)} \end{aligned} \tag{7}$$

where, if I_M denotes the unit matrix of size M, then

$$\Gamma_{k+1}^{(i)} = \begin{bmatrix} \gamma_{k+1}^{(i,1)} I_M & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \gamma_{k+1}^{(i,i)} I_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & \gamma_{k+1}^{(i,N)} I_M \end{bmatrix} \tag{8}$$

is an $M \cdot (N-1)+3 \times M \cdot (N-1)+3$ diagonal matrix made of binary random variables (as all observations can be reasonably assumed to be independent); $\tilde{H}_{k+1}^{(i,i)}$ is the Jacobian of $\tilde{\mathbf{h}}^{(i)}(\cdot)$

with respect to $\mathbf{p}^{(i)}$ computed at $\hat{\mathbf{p}}_{k+1|k}$ and

$$K_{k+1}^{(i)}=\hat{P}_{k+1|k}^{(i)}\tilde{H}_{k+1}^{(i,i)^T}\left[\tilde{H}_{k+1}^{(i,i)}\hat{P}_{k+1|k}^{(i)}\tilde{H}_{k+1}^{(i,i)^T}+\tilde{S}_{k+1}^{(i)}+\tilde{R}_{k+1}^{(i)}\right]^{-1} \quad (9)$$

is the Kalman gain of the IEKF running on the $i-$th agent. Observe that (9) comprises two measurement covariance matrices instead of just one, i.e. the block diagonal covariance matrix

$$\tilde{R}_{k+1}^{(i)}=\begin{bmatrix} R_{k+1}^{(i,1)} & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \\ \cdots & \cdots & R_{k+1}^{(i,i)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & R_{k+1}^{(i,N)} \end{bmatrix} \quad (10)$$

and

$$\tilde{S}_{k+1}^{(i)}=\sum_{j=1\wedge j\neq i}^{N}\tilde{H}_{k+1}^{(i,j)}\hat{P}_{k+1|k}^{(j)}\tilde{H}_{k+1}^{(i,j)^T}. \quad (11)$$

Matrix (10) includes both the covariance matrix $R_{k+1}^{(i,j)}$ associated with the relative pose measurements between each pair of agents (i,j) and the covariance matrix $R_{k+1}^{(i,i)}$ due to the absolute position and heading measures injected into the IEKF anytime a visual marker is detected. Matrix (11) instead takes into account the covariances $P_{k+1|k}^{(j)}$ of the states predicted by the agents different from i, with $\tilde{H}_{k+1}^{(i,j)}$ being the Jacobian of (5) with respect to $\mathbf{p}^{(j)}$ for $i\neq j$ and computed at $\hat{\mathbf{p}}_{k+1|k}$.

It is worth emphasizing that expressions (6) and (7) are absolutely general, but their actual implementation depends on the observation model used. With reference to the two models introduced in Section II, in the rest of this paper we will denote with subscripts A and B all the quantities which refer to *case A* and *case B*, respectively. In particular, in *case A* $\tilde{\mathbf{h}}^{(i)}(\mathbf{p}_k)=\tilde{\mathbf{h}}_A^{(i)}(\mathbf{p}_k)=[\mathbf{h}_A(\mathbf{p}_k^{(i)},\mathbf{p}_k^{(1)}),\ldots,\boldsymbol{o}(\mathbf{p}_k^{(i)}),\ldots,\mathbf{h}_A(\mathbf{p}_k^{(i)},\mathbf{p}_k^{(N)})]^T$ where

$$\mathbf{h}_A(\mathbf{p}_k^{(i)},\mathbf{p}_k^{(j)})=\sqrt{(x_k^{(j)}-x_k^{(i)})^2+(y_k^{(j)}-y_k^{(i)})^2} \quad (12)$$

and

$$\boldsymbol{o}(\mathbf{p}_k^{(i)})=\begin{bmatrix} (x_q-x_k^{(i)})\cos\theta_k^{(i)}+(y_q-y_k^{(i)})\sin\theta_k^{(i)} \\ -(x_q-x_k^{(i)})\sin\theta_k^{(i)}+(y_q-y_k^{(i)})\cos\theta_k^{(i)} \\ \theta_q-\theta_k^{(i)} \end{bmatrix}, \quad (13)$$

with (x_q,y_q,θ_q) being the planar coordinates and orientation of the $q-$th visual landmark on the floor in the chosen absolute reference frame. Dually, in *case B* $\tilde{\mathbf{h}}^{(i)}(\mathbf{p}_k)=\tilde{\mathbf{h}}_B^{(i)}(\mathbf{p}_k)=[\mathbf{h}_B(\mathbf{p}_k^{(i)},\mathbf{p}_k^{(1)}),\ldots,\boldsymbol{o}(\mathbf{p}_k^{(i)}),\ldots,\mathbf{h}_B(\mathbf{p}_k^{(i)},\mathbf{p}_k^{(N)})]^T$ with

$$\mathbf{h}_B(\mathbf{p}_k^{(i)},\mathbf{p}_k^{(j)})=\begin{bmatrix} (x_k^{(j)}-x_k^{(i)})\cos\theta_k^{(i)}+(y_k^{(j)}-y_k^{(i)})\sin\theta_k^{(i)} \\ -(x_k^{(j)}-x_k^{(i)})\sin\theta_k^{(i)}+(y_k^{(j)}-y_k^{(i)})\cos\theta_k^{(i)} \end{bmatrix} \quad (14)$$

and $\boldsymbol{o}(\mathbf{p}_k^{(i)})$ is the same as (13). Of course, a similar notation can be extended to all variables used in expressions (6)-(11). For instance, $R_{A,k+1}^{(i,j)}$ and $R_{B,k+1}^{(i,j)}$ denote the covariance matrices associated with the relative measurements between agents i and j in *case A* and *case B*, respectively, at time $(k+1)T_s$. Since such matrices can be assumed to be stationary, the time index can be omitted in the following. Notice that both (12) and (14) are nonlinear, but the number of observations M when another agent is detected is different in either case. In *case A* $M=1$ since low-cost wireless ranging systems can just measure the distance between transmitter and receiver, whereas in *case B* $M=2$ because the relative offsets between agents i and j along axes x and y can be easily extracted from the image collected by the RGB-D camera. On the contrary, measuring the relative orientation between i and j requires more sophisticated image processing algorithms and it is a measurement problem on its own. This is why this kind of measurements has not been included in the present analysis.

A final remark is on communication latency, which may influence measurement uncertainty significantly, due to the difference between the time when the predicted state of agent j is sent to agent i and the moment when the relative pose of j is actually measured by i. If assumption 5 defined in Section II holds, then the impact of communication latencies is negligible, provided that robots move so slowly that their linear and angular displacements during the time interval spent for communication is reasonably small. On the contrary, if the communication latency becomes significant, then the uncertainty contributions affecting all measurements of position, distance and heading should be properly estimated and used to boost the elements of $R_A^{(i,j)}$ and $R_B^{(i,j)}$.

IV. Simulation results

In order to evaluate the positioning accuracy with and without collaborative localization in *case A* and *B*, respectively, the results of some Monte Carlo simulations in different conditions are reported in the following. The main simulation parameters are listed below:

- Duration of each simulated path: about 120 s;
- Number N of agents in the room: between 2 and 10;
- Room size: 100 m^2;
- Number of random paths of each robot: 24;
- Robots linear velocity range: $[0,2]$ m/s (at such speeds assumption 5 holds, so the effect of communication latencies can be assumed to be negligible);
- Robots angular velocity range: $[-\frac{\pi}{2},\frac{\pi}{2}]$ rad/s;
- Odometers sampling period: $T_s=4$ ms;
- Covariance matrix of odometry noise (based on experiments on the field): $Q_k=\text{diag}(2\cdot10^{-8}\text{m}^2,2.2\cdot10^{-7}\text{rad}^2)$. In order to make simulations more realistic a relative (uncompensated) offset within $\pm3\%$ of wheels' angular displacements have been included in the simulated odometry.
- Reading range of the ground-facing monocular camera for landmark detection: 1 m with an aperture angle of about 40°;
- Distance between landmarks on the floor (assuming a regular grid): 2 m;

Table I. RELATIVE FREQUENCY OF DETECTION (EXPRESSED IN %) OF SOME OTHER AGENTS IN *Case A* AND IN *Case B*, RESPECTIVELY.

No. Agents	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Case A	9.8	9.9	10	10	10	10	10	10	10
Case B	0.2	0.4	0.5	0.7	0.9	1.0	1.2	1.3	1.4

- Covariance matrix of the measurement uncertainty associated with landmark detection (based on experiments on the field): $R_A^{(i,i)} = R_B^{(i,i)} = \text{diag}(1.6 \cdot 10^{-3}\text{m}^2,\ 5 \cdot 10^{-5}\text{m}^2,\ 1 \cdot 10^{-3}\text{rad}^2)$.
- Wireless system detection range: about 15 m;
- Rate of wireless range measurements: about 25 Hz;
- Variance of indoor wireless distance measurement data anytime agent j is detected: $R_A^{(i,j)} \approx 0.45$ m^2, in accordance with the experimental results published in the scientific literature [28], [29];
- RGB-D camera reading range (according to Kinect-like specifications): from 0.8 m to 3.5 m with a horizontal aperture angle of about 57°;
- Covariance matrix of stereo camera measurements anytime agent j is detected (based on Kinect's v.2 average accuracy reported in [30]): $R_B^{(i,j)} = \text{diag}(6.1 \cdot 10^{-7} d^{(i,j)^2}\text{m}^2,\ 6.25 \cdot 10^{-6}\text{m}^2)$, with $d^{(i,j)}$ being the distance between agents i and j along the focal axis of the camera.
- Camera and Kinect image acquisition rate: about 10 Hz.

Tab. I shows the simulated probabilities (expressed in %) that agent i detects some other agent in *Case A* and *B*, respectively. The reported values keep into consideration the actual data rates of the sensors employed for collaborative localization. Given that the RGB-D image acquisition and processing rate is notoriously quite low and its reading range is also much more limited than the range of a wireless system, the probability of agent detection in *Case B* is about 10 times smaller than in *Case A*. Observe that while in *Case B* the detection probability slightly grows with the number of agents, in *Case A* it saturates to about 10%. This is reasonable, because, even though the size of the room is smaller than the nominal wireless reading range, the rate of wireless distance measurements is 10 times smaller than the system sampling frequency. Thus, only once out of 10 times at least one of binary variables $\gamma^{(i,j)}$ is equal to 1 for $i \neq j$.

Fig. 2 shows the average root mean square estimation errors (RMSE) of state variables x (a), y (b) and θ (c) of one of the agents as a function of the number of robots N moving freely in the room. First, the RMSE values for each path are computed. Then, they are averaged together. Different bars refer to *Case A*, *Case B* and without collaborative localization, respectively. Notice that the results without collaborative localization do not depend on N, as expected, because the same 24 paths are used in all tests. On the other hand, both collaborative localization strategies greatly enhance the accuracy in estimating the state variables. Of course, such an improvement is more evident when the number of agents grows, due to the availability of a larger amount of relative

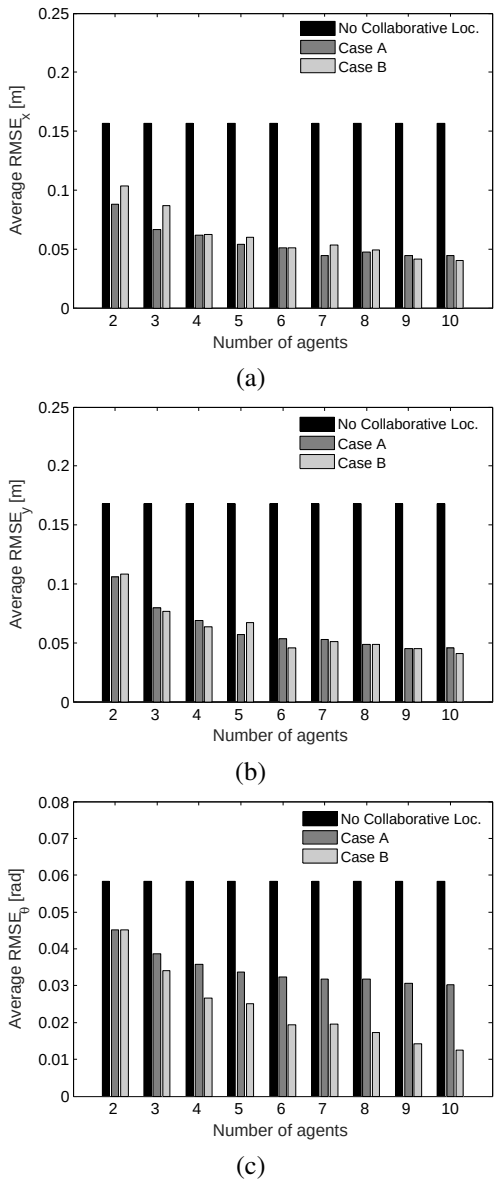


Figure 2. Average RMS estimation errors of state variables x (a), y (b) and θ (c) for one of the agents as a function of the number of robots in the room and in *case A*, *case B* and without collaborative localization, respectively.

measurement data from nearby robots. Quite interestingly, in spite of some fluctuations due to the limited set of simulated paths, the estimation accuracy of x and y in *Cases A*, *B* is comparable for a given N. The greater amount of available information due to both larger wireless coverage and higher data rate indeed compensate the poorer accuracy of wireless ranging. As far as state variable θ is concerned, estimation accuracy is generally clearly better in *Case B*. Also, the accuracy gap compared to *Case A* tends to grow with N, because wireless ranging cannot measure the relative position of two agents in a 2D reference frame.

V. CONCLUSIONS

Collaborative or synergic localization refers to the ability of a group of robots to refine their own estimated positions by using both neighbors' states and relative measurements of

distance and/or position. In this paper, the performances of two alternative collaborative localization strategies, both based on a common underlying Interlaced Extended Kalman Filter (IEKF), are compared through simulations. The reported results confirm that the effectiveness of collaborative localization becomes more evident when the probability of detecting other agents in the environment grows. In the two cases considered, the use of a RGB-D camera seems to be globally preferable to wireless ranging, although the difference is not so impressive as it was expected. Future work will be focused on a more in-depth analysis of the trade-off between these two scenarios, possibly finding an analytical relationship between detection probability, measurement accuracy and target performance.

ACKNOWLEDGMENT

The activities described in this paper have received funding from the European Union's Horizon 2020 Research and Innovation Programme - Societal Challenge 1 (DG CONNECT/H) under grant agreement n° 643644 "ACANTO - A CyberphysicAl social NeTwOrk using robot friends".

REFERENCES

[1] P. Chen, Y. B. Xu, L. Chen, and Z. A. Deng, "Survey of WLAN fingerprinting positioning system," *Applied Mechanics and Materials*, vol. 380, pp. 2499–2505, Aug. 2013.

[2] A. Cazzorla, G. De Angelis, A. Moschitta, M. Dionigi, F. Alimenti, and P. Carbone, "A 5.6-GHz UWB position measurement system," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 3, pp. 675–683, Mar. 2013.

[3] K. Lingemann, A. Nüchter, J. Hertzberg, and H. Surmann, "High-speed laser localization for mobile robots," *Robotics and Autonomous Systems*, vol. 51, no. 4, pp. 275–296, Jun. 2005.

[4] F. Ijaz, H. K. Yang, A. Ahmad, and C. Lee, "Indoor positioning: A review of indoor ultrasonic positioning systems," in *15th International Conference on Advanced Communication Technology (ICACT)*, Jan. 2013, pp. 1146–1150.

[5] A. Stancovici, M. Micea, V. Cretu, and V. Groza, "Relative positioning system using inter-robot ultrasonic localization turret," in *Proc. IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Montevideo, Uruguay, May 2014, pp. 1427–1430.

[6] R. Lins, S. Givigi, and P. Kurka, "Vision-based measurement for localization of objects in 3-D for robotic applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 11, pp. 2950 – 2958, Nov. 2015.

[7] H. Chung, L. Ojeda, and J. Borenstein, "Accurate mobile robot dead-reckoning with a precision-calibrated fiber-optic gyroscope," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 1, pp. 80–84, Feb. 2001.

[8] G. Panahandeh, M. Jansson, and S. Hutchinson, "IMU-Camera data fusion: Horizontal plane observation with explicit outlier rejection," in *Proc. International conference on Indoor Positioning and Indoor Navigation (IPIN)*, Montbeliard, France, Oct. 2013, pp. 406–414.

[9] K. Wang, Y. hui Liu, and L. Li, "A Simple and Parallel Algorithm for Real-Time Robot Localization by Fusing Monocular Vision and Odometry/AHRS Sensors," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 4, pp. 1447–1457, Aug. 2014.

[10] A. Stancovici, S. Indreica, M. Micea, V. Cretu, and V. Groza, "Relative localization methodology for autonomous robots in collaborative environments," in *Proc. IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Minneapolis, MN, USA, May 2013, pp. 1730–1733.

[11] A. Kealy, G. Retscher, N. Alam, A. Hasnur-Rabiain, C. Toth, D. Grejner-Brzezinska, T. Moore, C. Hill, V. Gikas, C. Hide, C. Danezis, L. Bonenberg, and G. Roberts, "Collaborative navigation with ground vehicles and personal navigators," in *Proc. International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Sydney, Australia, Nov. 2012, pp. 1–8.

[12] S. Roumeliotis and G. A. Bekey, "Synergetic localization for groups of mobile robots," in *Proc. of IEEE Conference on Decision and Control*, vol. 4, Sydney, Australia, Dec. 2000, pp. 3477–3482.

[13] S. Frattasi, M. Monti, and R. Prasad, "The synergic localization system (SLS)," in *Proc. Vehicular Technology Conference (VTC)*, vol. 4, Dallas, TX, USA, Sep. 2005, pp. 2753–2755.

[14] R. Buehrer, T. Jia, and B. Thompson, "Cooperative indoor position location using the parallel projection method," in *Proc. International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Zurich, Switzerland, Sep. 2010, pp. 1–10.

[15] D. Taniuchi, X. Liu, D. Nakai, and T. Maekawa, "Spring model based collaborative indoor position estimation with neighbor mobile devices," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 2, pp. 268–277, Mar. 2015.

[16] D. Fox, W. Burgard, H. Kruppa, and S. Thrun, "A probabilistic approach to collaborative multi-robot localization," *Autonomous Robots*, vol. 8, no. 3, pp. 325–344, 2000.

[17] T. Bailey, M. Bryson, H. Mu, J. Vial, L. McCalman, and H. Durrant-Whyte, "Decentralised cooperative localisation for heterogeneous teams of mobile robots," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, May 2011, pp. 2859–2865.

[18] A. Prorok and A. Martinoli, "A reciprocal sampling algorithm for lightweight distributed multi-robot localization," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, San Francisco, CA, USA, Sep. 2011, pp. 3241–3247.

[19] J. Knuth and P. Barooah, "Collaborative localization with heterogeneous inter-robot measurements by Riemannian optimization," in *Proc. International Conference on Robotics and Automation (ICRA)*, Karlsruhe, Germany, May 2013, pp. 1534–1539.

[20] A. Bahr, M. Walter, and J. Leonard, "Consistent cooperative localization," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Kobe, Japan, May 2009, pp. 3415–3422.

[21] S. Panzieri, F. Pascucci, and R. Setola, "Multirobot localisation using interlaced extended Kalman filter," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China, Oct. 2006, pp. 2816–2821.

[22] S. Panzieri, F. Pascucci, L. Sciavicco, and R. Setola, "Distributed cooperative localization," *Journal of Information Technology Research*, vol. 6, no. 3, pp. 49–67, Jul. 2013.

[23] P. Nazemzadeh, F. Moro, D. Fontanelli, D. Macii, and L. Palopoli, "Indoor Positioning of a Robotic Walking Assistant for Large Public Environments," *IEEE Trans. on Instrumentation and Measurement*, vol. 64, no. 11, pp. 2965–2976, Nov 2015.

[24] P. Nazemzadeh, D. Fontanelli, D. Macii, T. Rizano, and L. Palopoli, "Design and performance analysis of an indoor position tracking technique for smart rollators," in *Proc. International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Montbeliard-Belfort, France, Oct. 2013, pp. 1–10.

[25] A. Elnashar and M. El-Saidny, "Looking at LTE in practice: A performance analysis of the LTE system based on field test results," *IEEE Vehicular Technology Magazine*, vol. 8, no. 3, pp. 81–92, Sep. 2013.

[26] T. K. Y. Bar-Shalom, X. Rong Li, *Estimation with Application to Tracking and Navigation – Theory, Algorithm and Software*. John Wiley and Sons, 2001.

[27] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[28] P. Pivato, S. Dalpez, and D. Macii, "Performance evaluation of chirp spread spectrum ranging for indoor embedded navigation systems," in *Proc. of IEEE International Symposium on Industrial Embedded Systems (SIES)*, Karlsruhe, Germany, June 2012, pp. 307–310.

[29] D. Macii, A. Colombo, P. Pivato, and D. Fontanelli, "A data fusion technique for wireless ranging performance improvement," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 1, pp. 27–37, Jan. 2013.

[30] A. Corti, S. Giancola, G. Mainetti, and R. Sala, "A metrological characterization of the kinect V2 time-of-flight camera," *Robotics and Autonomous Systems*, vol. 75, Part B, no. 1, pp. 584 – 594, Jan. 2016.