2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Localization and mapping for aerial manipulation based on range-only measurements and visual markers

Felipe R. Fabresse[1], Fernando Caballero[1], Ivan Maza[1] and Anibal Ollero[1]

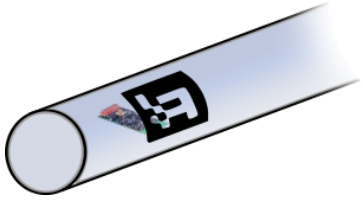
***Abstract*— This paper presents a new approach for aerial robots simultaneous localization and mapping (SLAM) oriented to aerial manipulation applications. The approach is based on the integration of range-only measurements and visual markers detected with the on-board camera. A multiple hypotheses framework is applied for range-only together with visual markers SLAM. This approach allows integrating two different types of sensors that are complementary for localization, mixing the stable and continuous estimation provided by range sensors with the precise but infrequent measurements from the visual markers. Real experiments involving an aerial robot and radio/visual markers have been used to validate the approach.**

I. Introduction

Range sensors have been subject of research in the last decade not only in the domain of localization applications but also for Simultaneous Localization And Mapping (SLAM) systems. These devices make possible the localization of robots or other kind of objects in GPS denied environments such as indoors. In the advantages, they offer a low cost localization solution which does not require a direct line of sight (LOS) between each pair of sensors when employing radio systems like Wifi or Ultra Wide Band (UWB) beacons. In addition, the data association problem is often solved by attaching the sensor identification with the range measurement information. However, the observability of range sensors state depends on the trajectory tracked by the vehicle, making visual information a perfect complement for these measurements in order to solve the ambiguities produced in those situations. In manipulation applications the use of range sensor might be used to give a rough estimation of the elements to be manipulated when their location is completely unknown and there is NLOS between the camera of the vehicle and these elements, whereas visual information provides a fine estimation for manipulation tasks.

Range sensors have been applied in different areas, such as localization of Autonomous Guided Vehicles (AGV) in manufacturing environments [1], outdoor mapping of RFID-based range sensors employing an aerial vehicle [2] and simultaneous localization and mapping of Autonomous Underwater Vehicles (AUV) and ultrasonic transponders [3]. The advantage of not requiring a LOS between each pair of sensors makes them very useful for the localization and identification of structural elements in indoor/outdoor aerial or ground manipulation.

(a) Aerial manipulation



(b) Embedded range sensors and visual markers

Fig. 1. Aerial manipulation of structural elements based on embedded sensors and visual markers.

The information provided by these range sensors is a simple measurement of the distance between two range sensor, which makes their integration in localization and/or mapping applications specially difficult due to the lack of bearing information which leads to multiple location hypotheses. This problem have been researched, specially in the case of 2D RO-SLAM, by researchers who have proposed different methods to solve the multi-hypotheses problem. Thus, to solve the mapping problem [4] proposes a grid-based solution to map the position of different beacons (range sensors with unknown locations). The solution is improved by the employment of a robust outlier rejection method based on a spectral graph partitioning. On the other hand, [2] proposes a probabilistic approach based on a particle filter to get an initial 2D estimation of each beacon position while the aerial vehicle moves around these beacons. Once the particle filter of each beacon has converged into a Gaussian solution, the estimation is switched to an Extended Kalman Filter (EKF). Although the results of this method are accurate, the main drawback is the delayed initialization of the EKF which loses all past range information employed by the particle filter. Later, the same authors proposed a multiple hypotheses solution [5] where the initial 2D localization uncertainty of the beacons is modeled with a Gaussian Mixture Model integrated in an EKF since the first range measurement re-

This work is partially supported by the ARCAS project (FP7-ICT-2011-7-287617) funded by the European Commission of the Seventh Framework Programme and the national projects RANCOM (P11-TIC-7066) and CLEAR (DPI2011-28937-C02-01).

[1]F.R. Fabresse, F. Caballero, I. Maza and A. Ollero are with Grupo de Robotica, Visión y Control, Universidad de Sevilla, Spain. E-mail: fabresse@us.es, fcaballero@us.es, imaza@us.es, aollero@cartuja.us.es

ceived. As new range measurements are received, the system reduces the number of hypotheses in order to decrease the computational load associated to multi-hypotheses solutions while the EKF estimate the real location of the beacon. To model the bearing information, the method is based on the polar parametrization proposed in [6]. The results of this method are accurate and, in contrast to previous work, it allows to initialize the EKF with a single range measurement.

In the domain of Range-only SLAM (RO-SLAM), two novel solutions are presented in [7], [8]. In [7] the authors employ a spectral algorithm based on a SVD decomposition of the observation matrix while in [8] two different approaches are used to solve the RO-SLAM problem: one based on semidefinite programming techniques and other based on convex optimization techniques. Other conventional SLAM frameworks are compared in [9], [10] where the Fast-SLAM approach is considered the most accurate and efficient solution. This Fast-SLAM framework is used in [11], [12], where a particle filter for both localization and mapping problem is used. Two different optimization techniques for the particle filter used to locate the different beacons are applied. Another Fast-SLAM solution is proposed in [13], [14] where the particle filter of the mapping problem is switched into an EKF once each particle filter converges into a Gaussian distribution reducing the computational burden of previous approaches. For 3D RO-SLAM, [15] proposes an undelayed solution which models the multi-modality of the range measurement model with a Gaussian Mixture Model so that each beacon is integrated in an independent EKF. This solution presents accurate results but does not allow the integration of inter-beacon measurements in such a way that cross correlations between beacons can be taken into account. Thus, the authors of [16] propose a 3D RO-SLAM method based on a centralized EKF-SLAM framework which allows the integration of inter-sensor measurements considering the cross correlation between them. In aerial manipulation, keeping the correlation between beacons is crucial when multiple beacons are embedded in the same structural element (see Fig. 1). The method also improves the scalability of the system with a reduced spherical parametrization and an efficient EKF update scheme.

In general, the main drawback of previous works is the observability problem associated to range measurements which depends on the trajectory tracked by the autonomous vehicle. In [17] this issue is solved using visual information to refine the rough estimation performed by a 2D range-only mapping method which is based on a delayed EKF-SLAM. The delayed initialization algorithm of this method is based on a particle filter and the solution proposed uses a Cartesian parametrization of the beacon hypotheses which is less robust than a polar parametrization [6].

The main contributions of this paper are the use of an optimal 3D RO-SLAM solution based on a reduced spherical parametrization integrating not only range measurements but also visual information. On the other hand, this paper will propose two different strategies to reduce the computational load required by multi-hypotheses frameworks by means of integration of visual information, these strategies will be compared an validated with experimental results.

The paper is organized as follows. The first section introduces the aerial manipulation problem focusing in the localization and mapping of the aerial vehicle and the structural elements. A summary of the method employed is given in Sect. II. Section III details the proposed method, explaining the different stages of the Bayesian filter applied. The method is then validated in real experiments in Sect. IV-A. Finally the conclusions in Sect. V close the paper.

II. System overview

The method presented in this paper extends the RO-SLAM method presented by the authors in [16] making it more robust against range measurement outliers and improving the prune strategy. On the other hand, the mapping solution proposed in this paper describes a method to include visual measurements which are based on the detection and localization of visual markers placed on the same structural elements where the range sensors are embedded (see Fig. 1). The method proposed is suitable for aerial manipulation tasks where an aerial robot endowed with a manipulator have to interact with several structural elements. In these cases, the system should localize the aerial robot and, at the same time, it should also map the structural elements to be manipulated. The use of visual markers allow the correction of the coarse estimation about the position of structural elements using the 3D range-only mapping algorithm detailed in the following.

Thus, this paper uses a centralized EKF to estimate not only the 3D position of the aerial robot, but also the 3D position of each beacon embedded in structural elements. This scheme has the advantage of keeping the correlation existing between those beacons embedded in the same structural element, which makes the position estimation more robust. On the other hand, the use of an EKF allows to deal with the non-linearities which appear in this kind of observation models in an efficient way.

III. SLAM with range sensors and visual markers

The solution presented in this paper uses a centralized EKF-SLAM whose state vector $\mathbf{x}$ is composed by the aerial vehicle position $\mathbf{x}_r = [x_r, y_r, z_r]^T$ and the position of m beacons (features of the map $\mathbf{f}_i$):

$$\mathbf{x} = [\mathbf{x}_r^t, \mathbf{f}_1^t, \mathbf{f}_2^t, \ldots, \mathbf{f}_m^t]^T \ . \tag{1}$$

The prediction and correction stages of this filter are described in the following subsections.

A. Aerial robot localization

For the localization of the aerial vehicle, any kinematic and dynamic model of the vehicle can be used. Then to correct the position predicted, this paper proposes to use some range sensors whose positions are known (anchors) as a mean to trilaterate the 3D position of the aerial robot. The positions of the anchors must be placed in a way that the vehicle can trilaterate its position as it moves around the entire scenario.

The equation applied to update the vehicle position in the EKF with a single range measurement is

$$r_i = \sqrt{(x_{a_i} - x_r)^2 + (y_{a_i} - y_r)^2 + (z_{a_i} - z_r)^2}\ , \qquad (2)$$

where $[x_{a_i}, y_{a_i}, z_{a_i}]$ is the known 3D position of the i-th anchor that provided the range measurement r_i.

B. Range-only mapping

As it was mentioned above, the mapping solution proposed in this paper uses a multiple-hypotheses strategy similar to that one described in [16]. Hence the position of the radio beacons is estimated by using range-only information and, in order to tackle with the multi-modality of this observation model, two Gaussian Mixture Models (GMM) are used to integrate all the hypotheses in the EKF since the very first measurement. Then, as new beacons are discovered, their hypotheses are added dynamically into the EKF with only the first range measurement received from the beacons. The state vector of each beacon $\mathbf{f}_i$ is represented using a reduced spherical parametrization (see Fig. 2) which is composed by the position of the aerial vehicle $\mathbf{x}_i = [x_i, y_i, z_i]^T$ from which the first range measurement was received, the range measurement received $\rho_i = r_i$ as the radius of the sphere, n_θ samples of possible azimuth angles θ_i and n_ϕ samples of possible elevation angles ϕ_i. Hence, the state vector of a beacon during the initialization stage is

$$\mathbf{f}_i = [\mathbf{x}_i^t, \rho_i, \theta_{i1}, \theta_{i2}, \ldots, \theta_{in_\theta}, \phi_{i1}, \phi_{i2}, \ldots, \phi_{in_\phi}]^T\ . \qquad (3)$$

It should be noticed how this spherical parametrization of the $n_\theta n_\phi$ hypotheses uses only $4+n_\theta+n_\phi$ parameters instead of $4+n_\theta n_\phi$. As in [16], each sample θ_{ij} is the mean value of one Gaussian mode of the GMM used to model the possible azimuth angles θ_i. The weights of each mode are initialized with the same probability and depends on the number of samples employed. In the same manner, the standard deviation $\sigma_{\theta_{ij}}$ of each sample is initialized uniformly and depends on the number of hypotheses $h = n_\theta n_\phi$ used to model the real uniform spherical distribution. The same applies to the GMM used to model the possible elevation angles ϕ_i. The number of hypotheses h are initialized automatically according to the first measurement received and the optimal density of hypotheses as it is described in [16]. Finally, the covariance of the center $\mathbf{x}_i$ is initialized with the same covariance matrix of the aerial vehicle at the moment of the initialization, and finally, the deviation σ_{ρ_i} represents the thickness of the initial uniform sphere distribution represented in Fig.2

Then, once a beacon has been initialized, the following range measurements are used to update each sample. This range measurements are filtered with a median filter in order to reject sensor outliers and hence making the method more robust. The update scheme used has also been optimized to only use $n_\theta + n_\phi$ equations, instead of the $n_\theta n_\phi$ equations used in other works. This update scheme is composed by $n_\theta + n_\phi$ equations like

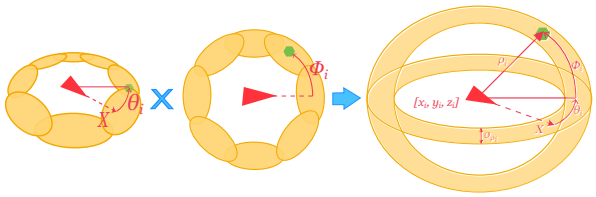


Fig. 2. Spherical parametrization of beacons. The yellow area represents the probability distribution where the feature can be located once a first range measurement is received. The annulus at left represents the GMM used to model the theta samples (azimuth angles), the annulus at center represent the GMM used to model the phi samples (elevation angles) and finally the sphere at right represents the real uniform and spherical distribution modeled with the combination of both GMM. The green object represents the real position of the range-sensor (beacon). The center of the sphere is composed by the position where the robot was located when the first range measurement was received.

$$r_i = \sqrt{(x_{f_i} - x_r)^2 + (y_{f_i} - y_r)^2 + (z_{f_i} - z_r)^2}\ , \qquad (4)$$

where x_{f_i}, y_{f_i} and z_{f_i} stand for

$$\begin{aligned} x_{f_i} &= x_i + \rho_i cos(\theta_i) cos(\phi_i) \\ y_{f_i} &= y_i + \rho_i sin(\theta_i) cos(\phi_i) \\ z_{f_i} &= z_i + \rho_i sin(\phi_i)\ . \end{aligned} \qquad (5)$$

Thus, to update one θ_{ij} sample the method proposed is to substitute the value ϕ_i of (4) by the expected elevation angle $\overline{\phi_i}$ and propagating its equation properly through the Jacobian $\mathbf{H}$. Analogously, $\overline{\theta_i}$ is used to update each sample ϕ_{ij}. More details are given in [16].

On the other hand, to update the weights of both GMMs the equations used followed by normalization are

$$\omega_{\theta_{ij_\theta}} = \omega_{\theta_{ij_\theta}} max(p(r_i|\mathbf{x}_r^t, \mathbf{x}_i^t, \rho_i, \theta_{ij_\theta}, \phi_{ij_\phi})|j_\phi = 1..n_\phi) \qquad (6)$$

and

$$\omega_{\phi_{ij_\phi}} = \omega_{\phi_{ij_\phi}} max(p(r_i|\mathbf{x}_r^t, \mathbf{x}_i^t, \rho_i, \theta_{ij_\theta}, \phi_{ij_\phi})|j_\theta = 1..n_\theta)\ . \qquad (7)$$

In addition to this optimized update strategy, a prune scheme is also used to reduce the computational burden and resources of the method as the filter converges into a single hypotheses. In this paper samples are removed if their weights are lower than a threshold $th = 10^{-10}/h$ where h is the current number of hypotheses. But, in contrast to [16], this paper proposes a merging strategy which join those samples which are similar with respect to their angular distance. Thus, for example, when the angular distance between two samples θ_{ij1} and θ_{ij2} is lower than a certain threshold, then these samples are joined as

$$\theta'_{ij1} = \frac{\omega_{\theta_{ij1}}\theta_{ij1} + \omega_{\theta_{ij2}}\theta_{ij2}}{2} \qquad (8)$$

and the new EKF covariance is updated with the Jacobian $\mathbf{J}$ of the equation (8) as $\mathbf{P}^t = \mathbf{J}\mathbf{P}^t\mathbf{J}$.

The mapping strategy presented is suitable for aerial manipulation to offer a coarse estimation about the location of structural elements to be manipulated, so that the vehicle can approach to these structural elements and refine the position estimation employing an on-board camera and visual markers placed over the same position where radio beacons are embedded (see Fig. 1(b)). The way in which this visual information is integrated into the filter is described in the next subsection.

C. Integration of visual markers

For the detection and localization of visual markers with a monocular camera several solutions have been analyzed and developed, such as the ArUco library [18]. The detection of visual markers is often solved using a hamming codification printed in visual markers which allows its identification after the marker is extracted from the image as follows:

1) Use of a border detection algorithm in the input image.
2) Remove undesired borders: borders without rectangular shapes and borders with few of pixels.
3) Sort corners of detected rectangles in anti-clockwise direction.
4) Remove elements which are very close to each other.
5) Use of homography methods to remove the perspective of markers detected.
6) Validation of the hamming code of markers detected (6x6 grids).

Finally, for those markers detected as valid, the algorithm applies the intrinsic model of the camera to get the 3D position of the marker with respect to the camera reference frame. This relative position is integrated in the EKF. The integration of this relative positions is computed if the filter previously contained an initialized beacon with the same identification code of the marker. If it is the case, the relative position of the beacon is then transformed in order to get the position with respect to the vehicle reference frame V. This frame is aligned with the global reference frame in our approach (i.e. the vehicle is only translated with respect to the global frame but not rotated). Finally the estimated position of the beacon with respect to the vehicle is computed for each hypotheses as

$$\begin{aligned} x_{f_i}^V &= x_i + \rho_i cos(\theta_i) cos(\phi_i) - x_r \\ y_{f_i}^V &= y_i + \rho_i sin(\theta_i) cos(\phi_i) - y_r \\ z_{f_i}^V &= z_i + \rho_i sin(\phi_i) - z_r \ , \end{aligned} \tag{9}$$

where $\mathbf{x}_{f_i}^V$ is the estimated relative position of the beacon with respect to the vehicle which can be compared by the EKF with the measured relative position extracted from visual markers. In that point, this paper propose two different algorithms:

- The first algorithm prune all hypotheses of the beacon detected except the one with higher probability according to the visual measurement received. The visual information is assumed to be very precise with respect to the hypotheses estimation of the filter which have not been pruned yet by the prune strategy explained in Sect. III-B. The following visual measurements are used to update the most likely estimated position of the beacon according to (9).
- The second algorithm updates all existing hypotheses of the beacon detected and their weights in a similar way than the method described in Sect. III-B, but now applying (9) and taking the visual measure received, i.e. each θ_i and ϕ_i sample is updated applying the expected elevation $\overline{\phi_i}$ and azimuth $\overline{\theta_i}$ angles respectively and the weights are updated in this case evaluating the probability $p(v_i|\mathbf{x}_r^t, \mathbf{x}_i^t, \rho_i, \theta_{ij_\theta}, \phi_{ij_\phi})$ where v_i is the visual measurement received. In this method, hypotheses are used as a mean to make the system robust against camera outliers instead of selecting the most likely which may lead to bad hypotheses due to a visual outlier for instance.

In practice, as it will be explained in the next section, both methods are very similar when the aerial robot follows an optimal trajectory. In those cases, when the first visual measurement is received and all the hypotheses have already converged into a single hypotheses the application of both methods is equivalent. The first method reduces considerably the computational burden of the multi-hypotheses scheme used by RO-Mapping algorithm with a single visual measure at the cost of a worst estimation due to visual outliers, whereas the second method is more robust against visual outliers.

IV. Experimental results

Experiments involving an aerial robot, range sensors and visual markers have been conceived in order to validate the approach presented in this paper. The main objective of these experiments is to demonstrate the suitability of the approach when real sensors with noisy measurements are used. The experimental results will be presented at two different levels: pure mapping experiments and full SLAM.

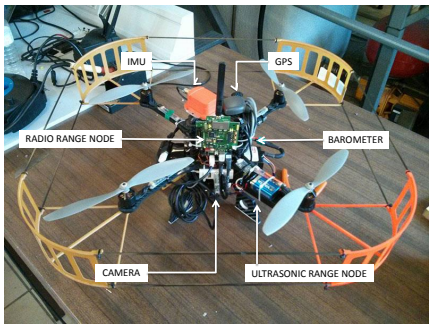
The mapping experiments are oriented to show how the approach behaves when the localization of the aerial robot is solved, and the problem is reduced to map the position of the range nodes in the space. This is a typical situation in indoors experiments when very precise positioning systems such as VICON or OptiTrack are used, or when RTK positioning systems are used outdoors.

SLAM experiments will consider the localization of both the aerial vehicle and the range nodes without the integration of precise localization systems. The paper will show results based only on local sensors (barometer and inertial measurement unit), range sensors and visual markers, and results also integrating standard GPS in single configuration.

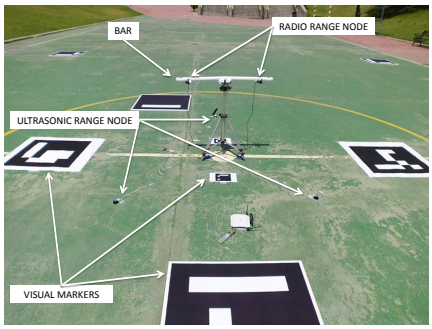
Prior to the discussion of the experimental results, next section describes the experimental setup describing all the elements involved in the experiment: aerial robot, sensors, ground-truth, etc.

A. Experimental setup

The complete experimental setup is composed by the following elements:



(a)



(b)

Fig. 3. Experimental setup. (a) Sensors on-board the aerial robot. The system is equipped with GPS in single configuration, a range sensor based on radio, a range sensor based on ultrasounds, a visual camera looking downwards, an inertial measurement unit and a barometer. (b) Experiments area with several elements deployed: ultrasound and radio range sensors, visual markers at known positions and a bar whose position will be estimated based on range sensors and visual markers.

- An aerial robot with the following on-board sensors (see Fig. 3(a)): GPS in single configuration, an inertial measurement unit providing roll, pitch and yaw, a radio based range sensor, an ultrasound based range sensor, a barometer and a visual camera pointing downwards.
- A bar with two range sensors based on radio and two small visual markers (see Fig. 3(b)).
- Five range sensors based on radio. Three of them are installed in the scenario at known positions, whereas the other two are mounted in the bar (see Fig. 3(b)).
- Three range sensors based on ultrasound placed at known positions in the scenario (see Fig. 3(b)).
- Several visual markers with different size deployed all over the scenario at known positions (see Fig. 3(b)).

Due to payload constraints, the aerial robot was not able to carry a RTK GPS, so the only positioning information was the single GPS on-board. In order to provide a good ground-truth for the position of the aerial robot, the real position of all the visual markers placed on the floor were registered and the images of the camera computed as a RANSAC PnP problem based on Levenberg-Marquardt optimization in order to compute the position of the vehicle with respect to them. The absence of vehicle ground truth in result figures is due to a lack of markers visibility from the vehicle camera. The ArUco library [18] was used to build and detect the markers not only for the ground truth computation but for the detection of bar markers too. This setup allows errors in the order of 10 cm when the markers occupy the 70% of the image or when several markers are detected in the same image.

During the experiments, the aerial robot was flying around and close to the bar in order to detect the visual markers on top of it. Next sections will explain the mapping and SLAM results obtained with Matlab. A video summarizing the experiments and the results obtained with an implementation on ROS framework can be downloaded from http://grvc.us.es/staff/felramfab/icra2014/video.mp4.

B. Mapping results

The goal of this experiment was to estimate the range sensors attached to the bar assuming that the position of the vehicle is known. For this purpose, the robot localization based on visual markers is used as robot true position, and the range nodes and visual markers on the bar are used to estimate their position. This experiment allows showing the convergence and proper behaviour of the approach when the robot position is known, which is very usual in indoor setups or outdoors with accurate position systems such as the RTK position system.

Figure 4 presents the evolution of the hypotheses for both range sensors. All the hypotheses are initialized around the robot position with the first range measurement. After some trilateration (see Fig. 4(b)), most position hypotheses have been deleted for beacon 1, while for beacon 2 have already converged to a single solution. Figure 4(c) shows the results right after the visual marker information attached to each node has been integrated using the prune hypotheses method. It can be seen how both positions converge to a single solution close to the actual position. Finally, after integrating more range and visual information, Figure 4(d) shows the estimated position of both range sensors. It can be seen in the legend of the figure how the errors are 15 cm and 25 cm respectively.

It is important to mention that the minimum distance from the aerial robot to the marker was 1.5 m during the mapping experiments. Visual markers detection on the marker's bar could be improved if the aerial robot were closer, and this improvement can be directly translated to the mapping process.

Both methods for visual marker detection have been implemented (prune hypotheses and weighting), but they yield to the same results with slightly differences, so the weighting results are not shown in the paper. In general, results will be very similar because the hypotheses quickly converge to a single one thanks to the good precision of the visual markers measurements. If the visual marker measurement is subject to outliers, then the weighting method could be more robust. Otherwise the prune hypotheses method is the best option from the computational point of view.

C. SLAM results

This section presents the results when the SLAM filter is used. The method integrate range only measurement from

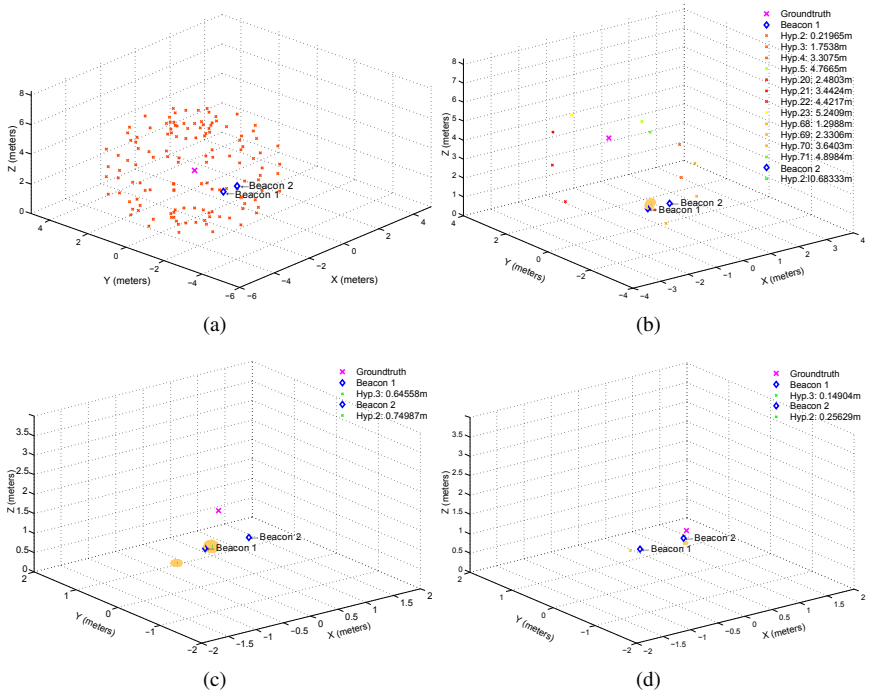


Fig. 4. Mapping results. Evolution of the multiple hypotheses for the position of two nodes attached to the bar. (a) Initialization of the poses into the filter. (b) Most of the hypotheses are pruned after some trilateration of the aerial robot. (c) Result right after visual marker measurements are included. It can be seen how both positions converge to a single hypothesis. (d) Final mapping results.

two different type of sensors at known positions, three radio based sensors with standard deviation of 1.5 m approximately and three ultrasonic devices with standard deviation of 5 mm. The SLAM filter also integrated information from the IMU in the prediction stage together with a model of the aerial robot, and barometric information in order to estimate the altitude. The positions of two radio nodes attached to the bar are also mapped and the position of the visual markers placed on the bar when they are detected are also integrated.

Figure 5 shows the localization results of the SLAM filter. The global RMS error is about 1.7 m, but the error is below this threshold most of the time as it can be seen in the figure. The ground-truth showed in the figure is the estimation provided by the visual marker detector. It can be seen how the estimation follows the real position of the vehicle most of the time with small errors. It is remarkable that no GPS have been used on the robot pose estimation. The localization of the nodes on the bar converges to single hypotheses thanks to the visual marker integration and they are localized with errors of 0.7 and 0.6 m respectively.

Figure 5 shows the results of the same experiment, but integrating a GPS in single configuration with the usual 2.5 m CEP. It can be seen how the RMS error is almost reduced to the half, 1.06 m. The mapping of nodes on the bar is slightly improved in one of them with localization errors of 0.7 and 0.4 m respectively. In general, the SLAM filter behaves better because the prediction model does not constraint enough the space of possible solutions, because we are using a pure predictive model. If other predictions are used, as visual odometry, the results will not differ significantly with respect

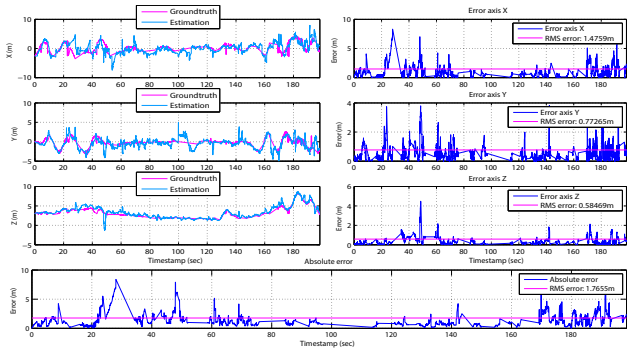


Fig. 5. SLAM results without GPS. Estimated position of the aerial robot with the SLAM filter. The position integrates range measurements from three radio and three ultrasonic devices, an IMU and a barometer for altitude estimation. No GPS is used for position estimation. The ground-truth is computed based on the detection of visual markers placed in the floor at known locations. The graphs show the estimated X, Y and Z values together with the ground truth. The RMS error per axis and the global error are also shown.

to the SLAM without GPS. It is important to remark that from seconds 100 to 120 there was no ground-truth because there were no markers in the field of view of the camera on-board the aerial robot. The mean value has been plotted in order to have an estimation.

V. Conclusions

The paper has presented a simultaneous localization and mapping approach employing range measurements to get a coarse estimation of beacons position and a camera to detect and localize the position of the beacons by using visual

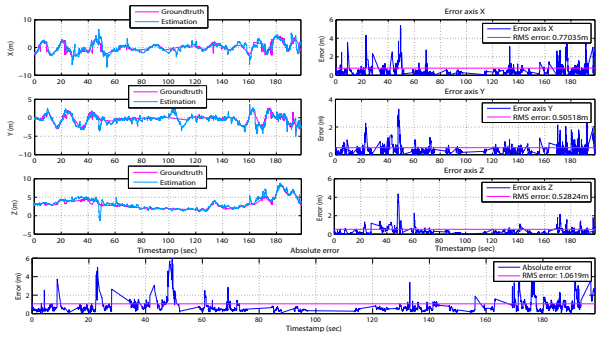


Fig. 6. SLAM results with GPS. Estimated position of the aerial robot with the SLAM filter. The position integrates range measurements from three radio and three ultrasonic devices, an IMU, a barometer for altitude estimation and a GPS in single configuration. The ground-truth is computed based on the detection of visual markers placed in the floor at known locations. The graphs show the estimated X, Y and Z values together with the ground truth. The RMS error per axis and the global error are also shown.

markers attached to the same beacons. This strategy can be applied for aerial manipulation tasks where the aerial robot uses the coarse estimation of the beacons embedded in structural elements for approaching, and then visual information is used to refine the position estimation of structural elements to be manipulated.

The localization approach uses range sensors with prior knowledge of their real position (anchors) to correct the prediction of the aerial vehicle position. On the other hand, the mapping solution relies on an optimal range-only mapping approach based on a reduced spherical parametrization and the use of marker detection and localization. Additionally, two different strategies have been proposed to reduce the computational load of the multi-hypotheses method used to initialize beacons position by integrating visual information and explaining the advantages and disadvantages of both methods. For the validation of the methods, the paper has presented some experimental results based on a real dataset collected with an aerial robot flying around a bar. A coarse estimation is obtained from the range measurements provided by two radio beacons embedded in the bar. The robot finally refines the estimation by approaching to the localized bar. The localization results using only a basic prediction model and five anchors are good. The mapping results have demonstrated how the precise visual information complements very well the stable and continuous range measurements provided by the radio beacons.

Future work will be focused on the use of other distributed solutions which can keep the cross correlation between those elements integrated in the same structural elements while at the same time it improves the scalability of the system with the number of elements to be manipulated by the aerial vehicle. On the other hand, more experiments with a higher number of structural elements and different vehicle maneuvers will be performed.

ACKNOWLEDGEMENT

Fernando Caballero is partially supported by the FROG project (FP7-ICT-2011.2.1) funded by the European Comission.

REFERENCES

[1] M. Gholami, N. Cai, and R. Brennan, "An artificial neural network approach to the problem of wireless sensors network localization," *Robotics and Computer-Integrated Manufacturing*, vol. 29, no. 1, pp. 96–109, Feb. 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0736584512000907

[2] F. Caballero, L. Merino, I. Maza, and A. Ollero, "A particle filtering method for wireless sensor network localization with an aerial robot beacon," in *IEEE International Conference on Robotics and Automation, 2008. ICRA 2008*, May 2008, pp. 596 –601.

[3] P. Newman and J. Leonard, "Pure range-only sub-sea SLAM," in *IEEE International Conference on Robotics and Automation, 2003. Proceedings. ICRA '03*, vol. 2, Sept. 2003, pp. 1921 – 1926 vol.2.

[4] E. Olson, J. Leonard, and S. Teller, "Robust range-only beacon localization," in *In Proceedings of Autonomous Underwater Vehicles*, 2004, p. 6675.

[5] F. Caballero, L. Merino, and A. Ollero, "A general gaussian-mixture approach for range-only mapping using multiple hypotheses," in *IEEE International Conference on Robotics and Automation (ICRA), 2010*, 2010, pp. 4404–4409.

[6] J. Djugash and S. Singh, "A robust method of localization and mapping using only range," in *Experimental Robotics*, ser. Springer Tracts in Advanced Robotics, O. Khatib, V. Kumar, and G. J. Pappas, Eds. Springer Berlin Heidelberg, Jan. 2009, no. 54, pp. 341–351.

[7] B. Boots and G. J. Gordon, "A spectral learning approach to range-only SLAM," *arXiv:1207.2491*, July 2012. [Online]. Available: http://arxiv.org/abs/1207.2491

[8] J. R. Spletzer, *A New Approach to Range-only SLAM for Wireless Sensor Networks*, 2003.

[9] Z. Kurt-Yavuz and S. Yavuz, "A comparison of EKF, UKF, FastSLAM2.0, and UKF-based FastSLAM algorithms," in *2012 IEEE 16th International Conference on Intelligent Engineering Systems (INES)*, June 2012, pp. 37 –43.

[10] G. Tuna, K. Gulez, V. Gungor, and T. Veli Mumcu, "Evaluations of different simultaneous localization and mapping (SLAM) algorithms," in *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*, 2012, pp. 2693–2698.

[11] P. Yang, "Efficient particle filter algorithm for ultrasonic sensor-based 2D range-only simultaneous localisation and mapping application," *IET Wireless Sensor Systems*, vol. 2, no. 4, pp. 394 –401, Dec. 2012.

[12] Z. M. Wang, D. H. Miao, and Z. J. Du, "Simultaneous localization and mapping for mobile robot based on an improved particle filter algorithm," in *International Conference on Mechatronics and Automation, 2009. ICMA 2009*, Aug. 2009, pp. 1106 –1110.

[13] D. Hai, Y. Li, H. Zhang, and X. Li, "Simultaneous localization and mapping of robot in wireless sensor network," in *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS)*, vol. 3, Oct. 2010, pp. 173 –178.

[14] J.-l. Blanco, J. Gonzlez, and J.-a. Fernndez-madrigal, "A pure probabilistic approach to range-only SLAM," Pasedena Conference Center, California, USA, May 2008.

[15] J.-L. Blanco, J.-A. Fernandez-Madrigal, and J. Gonzalez, "Efficient probabilistic range-only SLAM," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008. IROS 2008*, Sept. 2008, pp. 1017 –1022.

[16] F. R. Fabresse, F. Caballero, I. Maza, and A. Ollero, "Undelayed 3D RO-SLAM based on gaussian-mixture and reduced spherical parametrization," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Tokyo Big Sight, Tokyo, Japan, Nov. 2013, pp. 1555–1561.

[17] E. Menegatti, M. Danieletto, M. Mina, A. Pretto, A. Bardella, S. Zanconato, P. Zanuttigh, and A. Zanella, "Autonomous discovery, localization and recognition of smart objects through WSN and image features," in *2010 IEEE GLOBECOM Workshops (GC Wkshps)*, Dec. 2010, pp. 1653 –1657.

[18] "ArUco: a minimal library for augmented reality applications based on OpenCv," Jan. 2013. [Online]. Available: http://www.uco.es/investiga/grupos/ava/node/26