A Decoupled Approach for Simultaneous Stochastic Mapping and Mobile Robot Localization

G. A. Borges M.-J. Aldon

Robotics Department - LIRMM, UMR CNRS/Université Montpellier II, n°. C55060
161, rue ADA. 34392 - Montpellier - Cedex 5 - France. e-mail: {borges,aldon}@lirmm.fr

Abstract

This paper introduces a decoupled approach of concurrent mapping and localization for mobile robots. Its theoretical aspects rely on recent techniques for correct uncertainty handling using stochastic models: covariance intersection and unscented transform. Further, stochastic constraints are considered as a way to minimize map incoherence with respect to the real environment. Experimental results obtained from multisensory data acquired in a large real environment illustrate the performance of the proposed method.

1 Introduction

Concurrent mapping and localization (*CML*) is a very difficult ill-conditioned problem, on which a global environment map is updated using local exteroceptive sensor measurements acquired by mobile robots. Actually, there exist a number of approaches for solving *CML*, aimed at different map representations, as [13] dealing with occupancy grid maps. This paper concerns the stochastic map paradigm, which represents structured environments by geometrical primitives in a global map $\mathcal{M}^G$. In this context, the most used approach for *CML* relies on the Extended Kalman Filter (*EKF*), which keeps a stochastic state vector containing the robot pose and map features estimates, as well as its full-covariance matrix. Predictions of the state vector and its covariance matrix are obtained by dead-reckoning and updated by the *EKF* when exteroceptive measurements represented in a local map $\mathcal{M}^R$ are available [5]. However, this solution to *CML* has some limitations. The first one comes from the maintenance of a full-covariance matrix, which results in most time consuming map update, as well as in large memory requirements. These effects are reduced by the decoupled stochastic mapping approach [12], which divides the environment into multiple overlapped local maps. The second limitation comes from the use of the *EKF* for full-covariance map building. This filter may not perform well in presence of non-linearities, and presents long term divergence [11]. A common solution is to add artificial noise in the linearized models in order to cope with approximation uncertainties. This solution does not solve completely the problem since it can lead the *EKF* to converge to a biased state [8].

An interesting solution has been investigated by Julier and Uhlmann in [14], by using the extended covariance intersection (*ECI*) filter to separate pose estimation and individual map structures update. The *ECI* is a linearized version of covariance intersection filter (*CI*). *CI* guarantees conservative estimation given conservative prediction and measurements, even in the presence of unknown covariances. However, model linearization in the *ECI* can lead to incorrect non-conservative uncertainty handling. Motivated by recent advances on stochastic estimation, we present a new approach for *CML*, which separates pose estimation and map updating.

This paper is organized as follows. Section 2 introduces the environment map structures. Map updating requires an explicit local map representation, constructed as explained in section 3. Section 4 presents the theoretical aspects of this approach, summarized in section 5. Finally, section 6 presents experimental results carried out in a large real environment.

2 Map structures

In this work, map structures are composed of 2-D primitives: infinite lines, represented in polar parametric form $\mathbf{l} = (\rho, \alpha)^T$, and points, given by their Cartesian coordinates $\mathbf{p} = (x, y)$. These primitives are supposed to be perturbed by zero mean Gaussian noise, with covariance matrices $\Lambda_\mathbf{l}$ and $\Lambda_\mathbf{p}$ associated to $\mathbf{l}$ and $\mathbf{p}$, respectively. We further give a high-level identity to the map structures (Fig. 1):

- *Semiplanes* correspond to walls and other planar obstacles. They are represented by (i) an

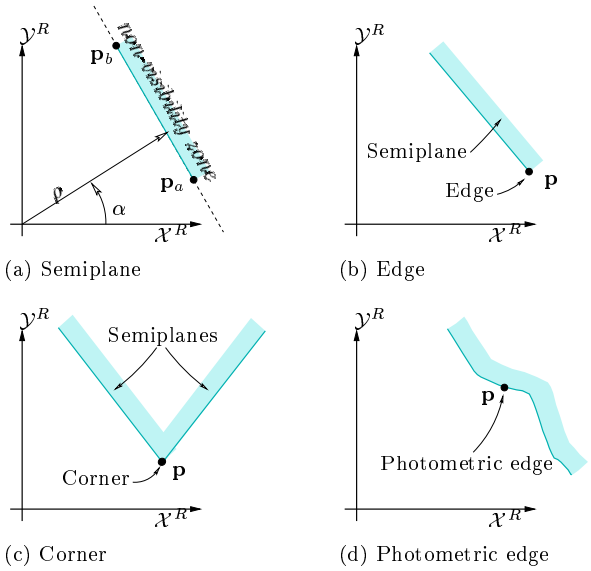


Figure 1: *Map structures.*

infinite line $\mathbf{l}$, (ii) two end-points and (iii) one flag indicating the visibility side of the obstacle;

- *Edges* correspond to the extremities of walls. They are represented by a point $\mathbf{p}$;
- *Corners* correspond to the intersection of two walls or two consecutive planar faces, and are represented by a point $\mathbf{p}$;
- *Photometric edges* correspond to artifacts observed as vertical lines in video images, and that correspond neither to *edges* nor to *corners*. They are represented by a point $\mathbf{p}$.

The above structures allow a rich representation for indoor environments, and the following relations exist among them : (i) every *edge* is associated to a *semiplane*, *i.e.*, the point primitive of the *edge* is over the line primitive of the *semiplane*; (ii) every *corner* is associated to two adjacent *semiplanes*, *i.e.*, the point primitive of the *corner* is on the intersection of the lines of the associated *semiplanes*. Such relations are used as prior information to constrain map updating. Furthermore, given the map structures uncertainty, constraints should be satisfied in the stochastic sense: the uncertainty of the constrained features should also be taken into account. For instance, an *edge* might not be exactly over the line of the associated *semiplane*. It may be close to the *semiplane*, but at a distance which should compatible with the uncertainties of both structures. We believe that respecting such constraints can minimize map divergence.

3 Local map building

The local map $\mathcal{M}^R$ is obtained from a pair of synchronized exteroceptive sensor images $(\mathcal{L}, \mathcal{I})$. $\mathcal{L}$ is the range image provided by a 2-D laser rangefinder. $\mathcal{I}$ is the video image provided by a monochrome camera. Multisensor data fusion consists in registering vertical edges extracted from the video image with the range image map. To do that, we estimate the observation angle ϕ from the rangefinder of any vertical edge observed at column u of the video image. An off-line calibration procedure has been applied to estimate the following relation:

$$\hat{\phi} = \Phi(u). \tag{1}$$

3.1 Image segmentation

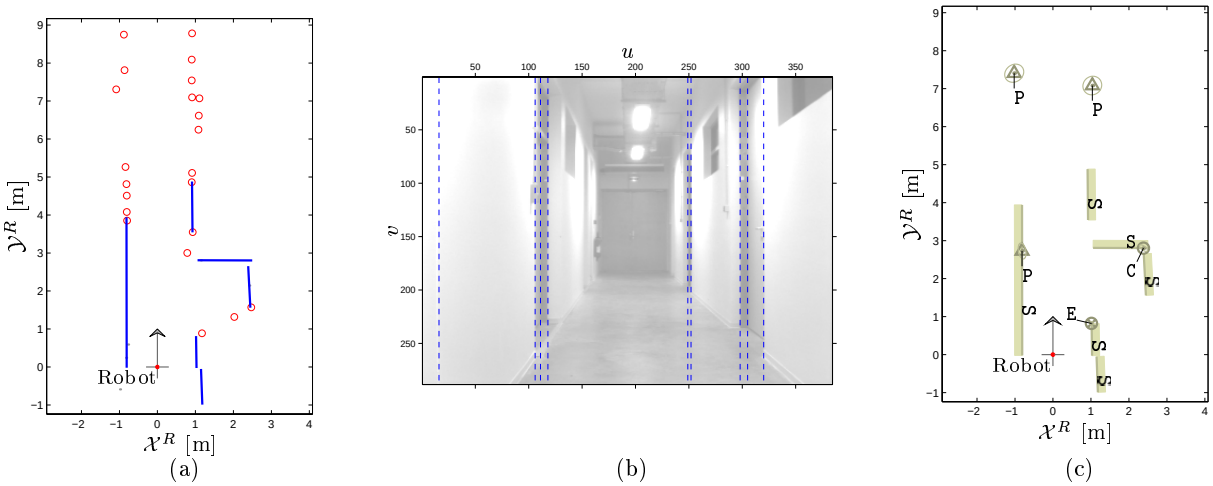
From $\mathcal{L}$, we extract a set Ω^S of line segments by using the fast split-and-merge fuzzy algorithm [2]. Further, laser scan breakpoints are detected using an extended Kalman filter-based approach [4], and stored in Ω^B. Breakpoints correspond to important discontinuities in the laser scan sequence, and may indicate the extremities of local environment surfaces. From image $\mathcal{I}$, similarly to [1], a set Ω^V of vertical line segments is extracted, which correspond to strong vertical contours in the image. .

3.2 Sensor data fusion

Given the sets Ω^S and Ω^B obtained from the range image, an initial local map is built. Such a map is composed of semiplanes, edges and corners only. Semiplanes are obtained from all line segments of Ω^S. Edges are computed from scan points which are (i) support points of line segments, and (ii) breakpoints. Edges are the projections on the support line segment of all scan points which satisfy (i) and (ii). Corners are the intersection of all line segments supported by consecutive scan points which are not breakpoints.

The second phase of local map uses the vertical lines extracted from the video image to update the initial local map structures. This procedure concerns only edges, which are also observed by the camera. Thus, the classical *EKF* is used to update the point coordinates of all edges whose observation angle $\phi_i = \arctan(y_i/x_i)$ is in correspondence with the measurement $\hat{\phi}_j = \Phi(u_j)$ (eq. (1)) Correspondences are verified using the classical χ^2-hypothesis test.

For last, photometric edges are estimated from all vertical lines which did not have correspondences in the initial local map. Since these structures are composed of a point, and the camera can only capture their observation angle $\hat{\phi}$, they are estimated using 100 bootstrap samples [6], obtained from the distribution of all scan points which lie, according to the



(a) (b) (c)

Figure 2: *Local map building example. (a) Range image and features (line segments and breakpoints). (b) Video image and features (vertical lines). (c) Local map (*`S`*: Semiplanes,* `E`*: Edges,* `C`*: Corners, and* `P`*: Photometric edges*

Mahalanobis distance test, in the vertical line angle of view. Photometric edges presenting large covariance matrices are discarded. Such a phenomena may arrive if the local region is very cluttered. In this way, we obtain reliable photometric edges. Figure 2 shows a local map $\mathcal{M}^R$ obtained from laser and video images.

4 Support for global map consistency

4.1 Theoretical tools

In order to achieve consistent map building, the proposed method relies on the following techniques :

1. The use of the Unscented Transform [9] to achieve consistent propagation of mean and covariance of map structures in all coordinate transformations involved in matching or fusing these features. Consistency of the stochastic variables is a necessary condition for consistent state estimation with linear filters;

2. The use of the linear covariance intersection [10] (*CI*) filter for all map structures fusion and constraint satisfaction. Since these procedures are governed by stochastic linear models, this filter guarantees conservative estimates of map features if they are at least consistent. The same can be done with Kalman filtering, but only if the prediction and measurement errors are non-correlated, or if their correlation is known. Since the proposed decoupled approach does not maintain map cross-covariances, even if they exist, Kalman filter cannot guarantee features consistency for the fusion and constraint satisfaction procedures. On the other hand, *CI* provides conservative estimates even in the presence of unknown correlations [10].

4.2 Map constraints satisfaction strategy

As discussed in section 2, there exist constraints which apply to map structures. Such constraints of the type "*a point* $\mathbf{p}$ *is constrained by lines* $\mathbf{l}_1, \mathbf{l}_2, \ldots$", should be satisfied in the stochastic sense since map structures are stochastic variables. Thus, these constraints can written as a linear stochastic model:

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{p} + \mathbf{w}, \tag{2}$$

with $\mathbf{w} \sim \mathcal{N}(0, \mathbf{\Lambda_w})$ representing constraint uncertainty, and

$$\mathbf{y} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} \cos(\alpha_1) & \sin(\alpha_1) \\ \cos(\alpha_2) & \sin(\alpha_2) \\ \vdots & \vdots \end{pmatrix}. \tag{3}$$

In a recent study [7], Kalman filter formalism has been used for constraint satisfaction in stochastic models. However, as stated before, unknown correlations exist between $\mathbf{p}$ and $\mathbf{l}_1, \mathbf{l}_2, \ldots$. Thus, the covariance intersection formalism is used to provide conservative update of the point primitive after constraint application. $\mathbf{\Lambda_w}$ is computed considering the uncertainties propagated from $\mathbf{p}$ and $\mathbf{l}_1, \mathbf{l}_2, \ldots$.

5 Decoupled pose estimation and global map updating

5.1 Map matching

Fast map matching is performed by the well known local matching procedure based on Mahalanobis dis-

tance. Prior pose information is provided by dead-reckoning for pose estimation. In the case of map updating, the prior is the corrected pose computed by the pose estimator.

For each pair of global map and local map semiplanes S^G and S^R, respectively, we compute the global representation $\hat{S}^G$ of S^R. The correspondence is confirmed if the following conditions are satisfied : (i) the Mahalanobis distance between S^G and $\hat{S}^G$ is in the 95% confidence interval of the χ^2_2 distribution; (ii) they are both visible from the prior robot pose; and (iii) they have at least a minimum common length.

For edges, corners and photometric edges, 95% confidence interval of the Mahalanobis distance between their point primitives is the necessary condition to establish a correspondence. Thus, an edge and a corner may be in correspondence, even if the associated semiplanes are not in correspondence.

5.2 Pose estimation

In the evaluations presented in this paper, pose estimation is accomplished by one of the following estimators: an extended Kalman filter (*EKF*)-based approach [1], or an optimal weighted least-squares (*WLS*) estimator [3]. In both approaches, robot pose is represented by $\mathbf{z} = (x, y, \theta)^T$, where x and y are the robot coordinates, and θ its heading. For large environments, less biased pose prediction $\hat{\mathbf{z}}_{k|k-1}$ is accomplished by gyrodometry, a dead-reckoning technique which fuses high precision laser gyrometer readings and odometry. Periodically, a correction cycle is performed after multisensory data acquisition, segmentation and local map building (as previously described). Local map structures are matched against global map ones using the map matching procedure of the previous section. Map correspondences are used to correct the pose prediction using the *EKF* or the iterative re-weighted *WLS* estimator ([3], eq. (39)). Nevertheless, some improvements were done on the *WLS* estimator: (i) all matrix inversions are performed using singular value decomposition, and (ii) the estimates are fused with the predicted ones by means of a linear Kalman filter. These procedures avoid some limitations regarding to the completeness of solutions, where only a solution subspace can be estimated. In the case of *EKF*, conservative estimation was possible after adjusting the filter parameters. In the theoretical point of view, the proposed map building approach requires conservativeness on pose estimation.

5.3 Map update

Map update is based on correspondences between the local and global maps. Using the current robot pose estimate, we apply the procedure of section 5.1 to establish such correspondences. Global map structures having at least a local map correspondence are denoted *observed*, whilst the local ones are *supported* features. Then, the following steps are performed.

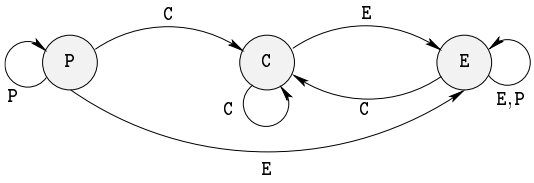


Figure 3: *Transition diagram of map structures.*

New map structures inclusion. All local non-supported structures are systematically included in the global map. Their global parameters are computed from the local ones by using the unscented transform. Local feature relationships are also propagated. For instance, as a local edge is always associated to a local semiplane, in the global map such an edge is associated (i) to the global representation of the local semiplane if it is *not supported*, or (ii) to the global correspondence of the local semiplane if it is *supported*.

Observed structures update. In this phase, *observed* structures updating is accomplished by fusing them with the global representation of the local *supported* structures. Semiplanes are fused first, followed by edges, corners and photometric edges. It should be pointed out that, according to the robot point of view, an environment corner can be captured as an edge or as a photometric edge in the local map. This is why map matching takes into account only low-level geometric information for edges, corners and photometric edges. Thus, the high-level interpretation of these structures may change depending of the local observation, according to the transition diagram of Figure 3. In this diagram, `P`, `C` and `E` stand for photometric edge, corner and edge, respectively. In the circles we have the type of the global map structure, and the labeled arrows represent the structure changing given the local observation type.

When all fusions have finished, map constraints are applied for all global map *edges* and *corners* which were updated, or which had at least one associated semiplane updated.

Free space preservation. The free space associated to a local semiplane is a triangle defined by the semiplane extremities and the robot position. Thus, any global map edge, corner or photometric edge which belong to the free space of a local semiplane

is deleted from the map. If it is a semiplane, only its part included in the free space is deleted. Some care should be taken when deleting map semiplanes. Indeed, when deleting a semiplane, all associated edges should also be deleted, and the associated corners become edges.

Map auto-fusion. In order to minimize the number of multiple map structures representing the same environment artifact, a procedure called map auto-fusion is performed. To do this, we identify multiple structures by applying the map matching procedure within the global map. The difference here is that high level features of different type cannot be fused. Two correspondences X_i^G and X_j^G are fused only if X_i^G is the correspondence of X_j^G with the smallest Mahalanobis distance between them, and *vice-versa*. This procedure finishes by applying map constraints to all modified structures.

6 Experimental results

Figure 4(a) presents a global map built from real data acquired during the exploration of a large environment. In this experiment, 162 images pairs were acquired, at a rate of 0.5 Hz. The total traveled distance was 133 meters, with maximum translational and rotational speeds of 58 m/s and 25 $^\circ/s$, respectively. The maximum time required for the map updating cycle was 135 ms.

One may argue that, given the precision of the laser gyrometer, gyrodometry could be used as pose estimator, without being necessary to use a map-based method. This is particularly true for short traveled distance map building experiments, on which gyrodometry provides accurate pose estimates. However, considering the traveled distance of this experiment, gyrodometry presents divergence after some traveled distance. We propose a comparison of the same decoupled framework in two different configurations: *UTC* (*U*nscented *T*ransform-based covariance propagation and *C*onstraints propagation), and *LNC* (*L*inearization-based covariance propagation and *N*o *C*onstraints propagation). Further, these map building configurations are combined with the pose estimators of section 5.2, resulting in *UTC-EKF*, *UTC-WLS*, *LNC-EKF* and *LNC-WLS*. Indeed, the map of Figure 4(a) is obtained when using the configuration *UTC-WLS*, which presented the best results in this experiment, and in all others we have done. When using *EKF* for pose estimation, we have noticed map incoherence with respect to the environment after some experimentation time. This can be seen in the zone of interest of Figure 4(a), which is magnified in Figure 4(b) for both *UTC* and *LNC* configurations. Nevertheless, we remark that *UTC* presented less biased map building than *LNC*. Figure 4(c) shows the same zone of interest when using *WLS* for pose estimation. The map incoherences of the previous results are not verified, in most part due to a more accurate pose estimation provided by the *WLS* estimator. In this case, the benefits of the *UTC* configuration are verified when using the constructed map for absolute navigation experiments, where the number of failures (*i.e.*, the robot gets lost) is far lower that when using the map built with the *LNC* configuration.

7 Conclusion

In this paper, we described a decoupled approach of simultaneous mapping and robot localization. It has been applied for consistent construction of a high level environment map by a mobile robot equipped with a laser rangefinder and a monochrome camera. The results obtained do not imply that the proposed approach guarantees a long term consistent map building. However, we claim that its strong theoretical support reduces the risk of map divergence. This is reinforced by a comparison, where map divergence has been noticed by the same decoupled framework using classical uncertainty handling methods and different pose estimation approaches.

Acknowledgments

G. A. Borges was supported by Capes (Brasília - Brazil), under grant BEX2280/97-3.

References

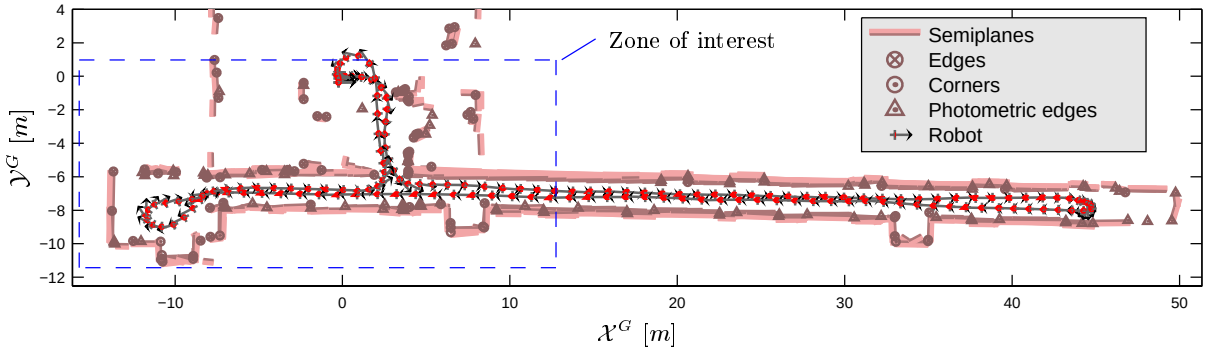
[1] K. O. Arras, N. Tomatis, B. T. Jensen, and R. Siegwart. Multisensor on-the-fly localization: Precision and reliability for applications. *Robotics and Autonomous Systems*, 34:131–143, 2001.

[2] G. A. Borges and M.-J. Aldon. A split-and-merge segmentation algorithm for line extraction in 2-D range images. In 15^{th} *International Conference on Pattern Recognition*, September 2000.

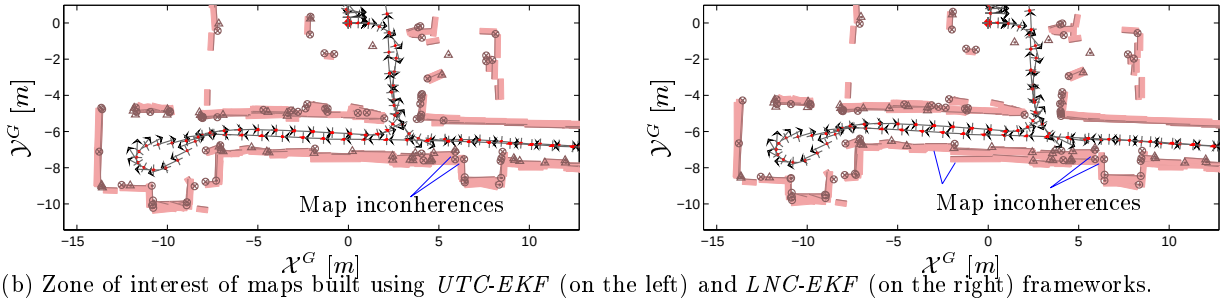
[3] G. A. Borges and M.-J. Aldon. Optimal mobile robot pose estimation using geometrical maps. *IEEE Transactions on Robotics and Automation*, 18(1):87–94, February 2002.

[4] J.A. Castellanos and J.D. Tardós. Laser-based segmentation and localization for a mobile robot. In F. Pin M. Jamshidi and P. Dauchez, editors, *Robotics and Manufacturing: Recent Trends in Research and Applications*, volume 6, pages 101–109. ASME Press, 1996.

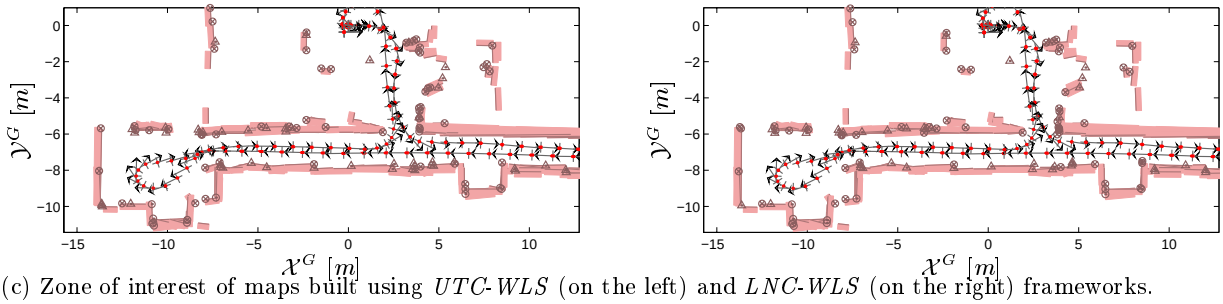
[5] M. W. M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Transactions on Robotics and Automation*, 17(3):229–241, June 2001.



(a) Large real environment global map built in the *UTC-WLS* framework.



(b) Zone of interest of maps built using *UTC-EKF* (on the left) and *LNC-EKF* (on the right) frameworks.



(c) Zone of interest of maps built using *UTC-WLS* (on the left) and *LNC-WLS* (on the right) frameworks.

Figure 4: *Experimental results.*

[6] B. Efron. Bootstrap methods: another look at the jackknife. *Journal of Ann. Stat.*, 7:1–26, 1979.

[7] Y. Hel-Or and Werman M. Constraint-fusion for interpretation of articulated objects. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 39–45, 1994.

[8] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[9] S. J. Julier, J. Uhlmann, and H. F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Robotics and Automation*, 45(3):477–482, 2000.

[10] S. J. Julier and J. K. Uhlmann. A non-divergent estimation algorithm in the presence of unknown correlations. In *American Control Conference*, 1997.

[11] S. J. Julier and J. K. Uhlmann. A counter example to the theory of simultaneous localization and map building. In *IEEE International Conference on Robotics and Automation*, 2001.

[12] J. J. Leonard and H. J. S. Feder. Decoupled stochastic mapping. *IEEE Journal of Oceanic Engineering*, 26(4):561–571, October 2001.

[13] S. Thrun, W. Burgard, and D. Fox. A probabilistic approach to concurrent mapping and localization for mobile robots. *Machine Learning*, pages 29–53, 1998.

[14] J. K. Uhlmann, S. J. Julier, and M. Csorba. Nondivergent simultaneous map-building and localization using covariance intersection. In Scott A. Speigle, editor, *Navigation and Control Technologies for Unmanned Systems II*, pages 2–11. SPIE, 1997.