2009 IEEE International Conference on Robotics and Automation
Kobe International Conference Center
Kobe, Japan, May 12-17, 2009

# Spectral Clustering for Feature-based Metric Maps Partitioning in a Hybrid Mapping Framework

Ricardo Vázquez-Martín, Pedro Núñez, Antonio Bandera and Francisco Sandoval

*Abstract*— **Hybrid maps combine metric and topological information for efficiently managing large-scale environments. In a feature-based mapping framework, this paper describes the application of a spectral clustering approach for automatically detecting the transitions between subsequently traversed local maps. Contrary to recently proposed approaches, this algorithm considers each individual map feature as a node of a graph whose edges link two nodes if they are simultaneously observed. Thus, given a sequence of observations, an auxiliary graph is incrementally built whose edges carry non-negative weights according to the locality of the features. Given a feature, its locality defines the set of features that has been observed simultaneously with it at least once. At each execution of the mapping approach, the feature-based graph is split into two subgraphs using a normalized spectral clustering algorithm. If the graph partition is validated, the algorithm determines that the robot is moving into a new area and a new local map is generated. We have tested the proposed approach in real environments where features are obtained using 2D laser sensors or vision. Experimental results demonstrate the performance of the proposal.**

## I. Introduction

Autonomous navigation is a fundamental ability for mobile robots which requires the integration of different modules. Among them, self-localization and environment mapping are two essential ones, as they are needed at different levels, from low-level control to higher-level strategic decision making or navigation supervision. It is well known that to guarantee bounded errors on its pose estimates, the robot must rely on sensors which can perceive stable environment features. Thus, if the robot manages a spatially consistent map of the environment, it could apply a map-based localization approach to obtain a correct estimation of its pose [1]. On the other hand, if the robot pose is exactly known, it could build a consistent environment map with the perceived data. The mapping and localization tasks are then *intimately tied together* [2], and they must be concurrently solved. The problem of the simultaneous localization and mapping (SLAM) has been extensively addressed by the robotic community in the last years.

Conventional approaches to SLAM rely on a metric, probabilistic representation of the robot pose and map. These metric approaches attempt to represent the spatial distribution of the perceived environment, commonly in the form of a feature map or an occupancy grid. Although they have been successfully employed to map relatively large-sized environments, the main limitation of these techniques is related to the excessive computational complexity associated to these mapping processes. The classical alternative to metric maps is to model the environment using a topological map. Topological maps attempt to capture the spatial connectivity of the environment by representing it as a graph with edges connecting the nodes that designate distinctive places in the environment [3]. These maps typically require reduced storage requirements, but they also usually lack the necessary information to localize arbitrarily (can only localize to nodes in the topological graph). Besides, while probabilistic methods have been extensively investigated for performing inference over the space of metric maps, it is not the same case for topological maps. As a significant exception, the probabilistic topological maps (PTMs) [4] is a sample-based representation that approximates the posterior distribution over topologies given available sensor measurements.

In order to deal with large, complex environments, the internal representation acquired by the robot can be organized as a hierarchy of maps which represent the whole environment at different levels of abstraction. Typically, these hierarchical representations consists of two layers: a metric map and a higher-level topological map. The hybrid approach usually attaches a local metric map to the nodes of a graph-based environment representation, where edges represent coordinates transformation between nodes. Thus, the complexity can be bounded within each local map. The problem is then to define what part of the mapped environment is associated to each topological node. Simhon and Dudek [5] proposed a strategy to create new maps in the presence of feature-rich regions or islands of reliability. Geometrical methods such as generalized Voronoi graphs [6] have been also used to segment the metric space representation. Recently, Zivkovic et al [7] and Blanco et al [8] have proposed to represent the base-level map as a graph, and to decompose it into nodes using efficient approximate solutions to the normalized graph cut criterion. This graph-partitioning method can be generalized for dividing a graph into a variable number of subgraphs [8]. The main disadvantage of these approaches is that the whole base-level map must be built in advance. An alternative solution is to apply this partitioning algorithm periodically to the acquired local map [9]. In other hierarchical approaches, this partitioning process works on-line. Thus, in the atlas framework, Bosse et al [10] proposed to create a new map when the uncertainty of

This work has been partially granted by the Spanish Ministerio de Ciencia y Tecnología (MCYT) and FEDER funds, and by Junta de Andalucía, under projects no. TIN2005-01359 and P07-TIC-03106, respectively

R. Vázquez-Martín, A. Bandera and F. Sandoval are with Department of Electronic Technology, University of Málaga, Campus de Teatinos 29071-Málaga, Spain rvmartin@uma.es, ajbandera@uma.es, fsandoval@uma.es

P. Núñez is with the Department of Signal Theory and Communications, University of Extremadura, Cáceres, Spain pnuntru@unex.es

the robot location grows above some limit. The hierarchical SLAM [11] integrates new features into local maps until a given number is reached. Both partitioning approaches can generate different submaps of a given environment area depending on the robot trajectory across it. In a recent work, Brunskill et al [12] have proposed a new algorithm to automatically decompose a map into submap segments using a spectral clustering approach.

In the framework of a hybrid metric/topological approach to the SLAM problem, this paper describes a partition algorithm which can be used to cluster detected features into groups that give rise to precise local maps. These maps can be built using any kind of sensor, as the only information that the algorithm needs is the set of observed features and the association between them. Assuming that the robot is always located at one environment area, the proposed approach has been integrated into a SLAM algorithm which estimates the robot pose and the local map structure at metric level by means of a feature-based stochastic mapping using an extended Kalman filter (EKF). Map management and data association are then addressed by the SLAM algorithm. This paper is closely related to the previous approaches of Blanco et al [8], Zivkovic et al [7] and Brunskill et al [12]. Like these proposals, our approach decompose the acquired local map into submaps using a spectral clustering algorithm. However, meanwhile all these approaches consider robot poses as the nodes of the local graph to divide, in our proposal the nodes of the local graph are environment features. Besides, all approaches employ different methods to compute the similarity matrix.

Finally, it must be noted that the proposed partitioning approach must be integrated into a hybrid mapping framework. When the robot is turning back or when traversing a loop for the second time, it can reaches a previously visited local map. In these situations, a relocation approach must be used to confirm the revisiting of the local map. In this way, the unbounded growth of the topological representation when the robot moves through loops repeatedly will be avoided [14].

The rest of the paper is organized as follows: after briefly discussing the main aspects of the normalized spectral clustering theory in Section II, Section III presents the proposed algorithm for feature-based map partitioning. Experimental results in Section IV demonstrate the efficiency and precision of the proposed method. Finally, in Section V, we draw the main conclusions of this study and outline future research directions.

## II. NORMALIZED SPECTRAL CLUSTERING

### A. Graph Notation

Let $G = (N, E)$ be an undirected, weighted graph with node set $N = \{n_1, ...n_n\}$ and where each edge between two nodes $n_i$ and $n_j$ has associated a non-negative weight $w_{ij} \geq 0$. The weighted *adjacency matrix* or *similarity matrix* of the graph $G$ is the matrix $W = (w_{ij})_{i,j=1,...n}$. As the graph is undirected, this matrix is symmetric ($w_{ij} = w_{ji}$).

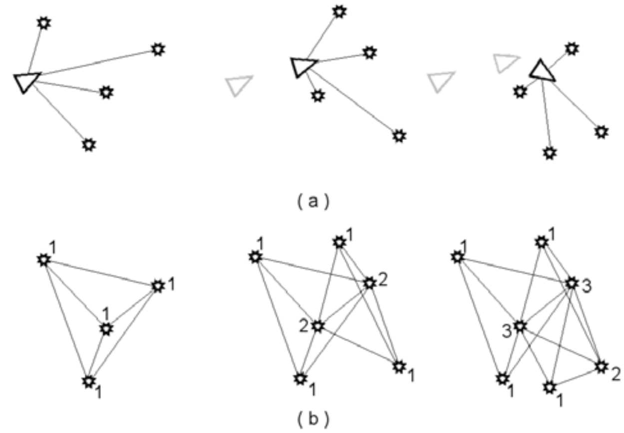


Fig. 1. Covisibility Graph building process. Each node is annotated with the number of observations of the corresponding feature. Edges represent locality of features and their values correspond to the times they have been observed simultaneously (these values have not been drawn).

The degree of a node $n_i \in N$ is defined as

$$d_i = \sum_{j=1}^{n} w_{ij} \tag{1}$$

Then, the *degree matrix* $D$ is defined as the diagonal matrix with the degrees $\{d_i\}_{i=1}^n$ on the diagonal. Given a subset $A \subset N$, $|A|$ denotes the number of nodes in $A$ and vol$(A)$ is a measure of the size of $A$ defined by the weights of its edges ($vol(A) = \sum_{i \in A} d_i$).

A subset $A \subset N$ is connected if any two nodes in $A$ can be joined by a path such that all intermediate nodes also lie in $A$. The subsets $A_1, ...A_k$ are a partition of the graph $G$ if $(A_i \cap A_j)_{i \neq j} = \emptyset$ and $A_1 \cup ... \cup A_k = N$.

### B. Normalized Spectral Clustering Algorithm

Given the graph $G = (N, E)$, according to the proposal of Shi and Malik [13], the normalized spectral clustering algorithm for graph partitioning consists of the following steps:

1) Solve for the two eigenvectors with the smallest eigenvalues of the generalized eigenproblem

$$Lv = \lambda Dv \tag{2}$$

where $L$ is the unnormalized graph Laplacian matrix defined by $L = D - W$.

2) Let $V \in \mathcal{R}^{n\text{x}2}$ be the matrix containing the vectors $v_1$ and $v_2$ as columns.
3) Let $y_i \in \mathcal{R}^2$ be the vector corresponding to the $i$-th row of $V$. Cluster the points $\{y_i\}_{i=1}^n$ into clusters $C_1$ and $C_2$ with the $k$-means algorithm.

Thus, the second smallest eigenvector of the generalized eigenproblem is the real valued solution to the normalized cut problem. The only reason that it is not necessary the solution to the original problem is that the second constraint on $y$ that $y_i$ takes on two discrete values is not automatically satisfied. Relaxing this constraint is what makes this optimization problem tractable in the first place.

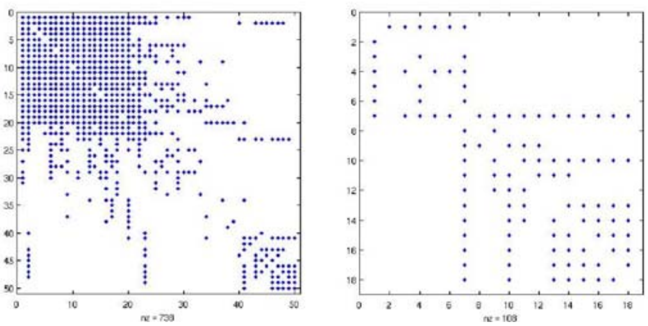

Fig. 2. Similarity Matrices associated to two Covisibility Graphs.

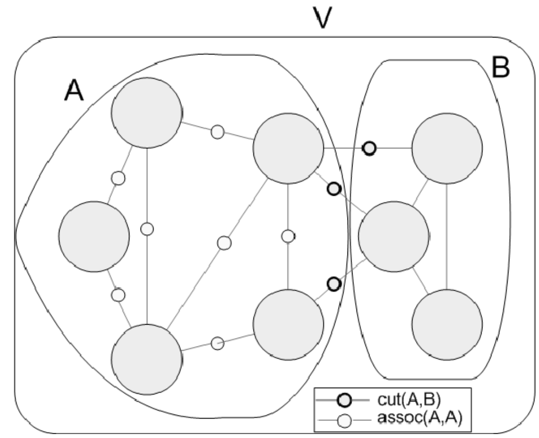


Fig. 3. This example illustrate two subsets $A, B \subset V$ to show the concepts of $cut$ and $assoc$. It can be noted how association involves not only the intergroup cohesion between $A$ and $B$ but the intragroup cohesion of the subset $A$ (see text).

## III. Graph-based Environment Representation

### A. The Covisibility Graph

The proposed approach works in the EKF-SLAM framework, being the environment description based on feature maps. Detected and matched landmarks are used to build an auxiliary graph of robot observations, called the *Covisibility Graph* (CG). In this graph, observations are represented as nodes and those features that has been observed from the same pose of the robot are connected by edges. Given a sequence of robot observations, the CG is incrementally built in each step (see Fig. 1). Here, the critical problem is to match different observations $z_t$ and $z'_t$, taken at time steps $t$ and $t'$. The EKF-SLAM algorithm solve this issue in the Data Association stage, where the observations gathered at each time step are compared to those features stored in the map. Using the set of landmarks updated in the SLAM process, after the Data Association stage, the correspondence problem is solved. Thus, the CG building process is able to take advantage of using a batch data association method, and a reliable multiple data tracking in cluttered environments is achieved [15]. Therefore, this graph is obtained from the set of features given in the update stage of the EKF-SLAM algorithm, and contrary to previous approaches [8] [12] [7] these nodes are related to landmarks and no additional information about the robot pose is directly added.

The CG represents the observed landmarks, taken at different time steps and robot poses, and their relationship. As it has been aforementioned, this relation between observations is the locality of a feature, defined as the set of features that have been seen simultaneously with it at least once [16]. This locality is represented in the CG as edges connecting those features. Once this auxiliary graph is built, the edge weights must be computed in order to define the similarity matrix. For each node in the graph, $n_i$ is the number of observations of the feature $F_i^t$ up to time $t$. If two features $F_i^t$ and $F_j^t$ have been seen in the same observation at time $t$, an edge is created between them or if it exists, i.e. both landmarks have been previously observed at the same time, the edge value $a_{ij}$ is increased. In order to compute an average value related to the locality of features, the *covisibility rate* is defined as:

$$CR_{ij} = \frac{a_{ij}}{\min(n_i, n_j)} \tag{3}$$

here, the edge value is divided by the minimum number of observations of the corresponding features. As it can be seen, the minimum possible value of the *covisibility rate* is zero when two features have not been seen simultaneously, and the maximum value is one, when two features always have been observed simultaneously, at least one of them. So the *covisibility rate* is within the range [0, 1]. This is our similarity function that allows to compute the adjacency or similarity matrix. This matrix is symmetric, non-negative and band diagonal. This last property can be seen in Figure 2. Next subsections explain how this matrix is used to find the map partition.

### B. Graph Partitioning

Once the CG is built and the similarity matrix is computed, the aim is to split the graph in order to minimize the lost of information in the map partition. This problem can be stated as to find a partition of the graph where the edges between different groups have a very low weight and the edges within the same cluster have high values. The degree of dissimilarity between these two clusters can be computed as the total weight of the edges between them, that is, the $cut$ for two disjoint subsets $A, B \subset V$ is:

$$cut(A, B) = \sum_{i \in A, j \in B} w_{ij} \tag{4}$$

The optimal partitioning of a graph is achieved when the cut value is minimized. This leads to cuts of small sets of isolated nodes in the graph. Instead of using this criteria, Shi and Malik [13] proposed the minimization of the normalized cut ($Ncut$). The $Ncut$ value is defined as:

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \tag{5}$$

where:

$$assoc(A, V) = \sum_{u \in A v \in V} w_{uv} \tag{6}$$

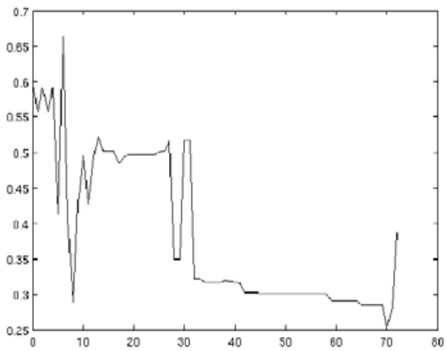


Fig. 4. $Ncut$ values during the mapping process. If it is thresholded, a non-optimal solution can be taken due to the presence of local minimum.

is the total connection from nodes in A to all nodes in the graph. This definition of the disassociation between clusters fulfil both the intergroup (the strength of the edges between clusters) and the intragroup cohesion (the strength of the edges in the same cluster). Note that this definition satisfy:

$$assoc(A,V) = cut(A,B) + assoc(A,A) \tag{7}$$

as it is illustrated in Figure 3.

So, minimizing $Ncut$ values allows to avoid isolated partitions. Although this problem is NP-hard, the statistics literature contains many variants of algorithms to approximate the normalized cut of a graph. Spectral clustering is one of the most efficient and robust. In our case, we have employed the algorithm presented in Section II.

### C. Validation of the Graph Partition

The proposed approach build an auxiliary graph (CG) incrementally from the observations gathered by the robot during the localization and mapping process. During the CG building process, while the robot moves through the environment, the spectral clustering is run in order to find a suitable partition of the map and the transition to a new area. As it has been aforementioned, the objective function is the minimum normalized cut of the graph [8]. Thus, the best partition is given when the minimum $Ncut$ value is found. However, the aim of our approach is to provide a map partition while the robot is moving. If the $Ncut$ value is thresholded, it is possible to fall in a local minima and a wrong partition would be taken (see Figure 4). Another possibility is to study the set of eigenvalues of the similarity matrix. These eigenvalues generally correspond to clusters in the graph [12]. In our case, the spectral clustering algorithm is applied to the map that is being built to find a partition of the graph into only two clusters in order to detect a new area. Therefore, we should evaluate the magnitude of the second eigenvalue ($eigv_2$). In order to define a more robust criterion, these two values are combined. The map partition is valid if the following condition holds

$$Ncut \leq Ncut_{min} \ \ \& \ \ eigv_2 \leq eigv_{2min} \tag{8}$$

where $Ncut_{min}$ and $eigv_{2min}$ are constants. The thresholds for submap generation for the $Ncut$ value and the second eigenvalue have been set experimentally to 0.3 and 0.25, respectively. As it is shown in Figure 5, the same local maps are created even when the robot follows different trajectories through the same environment. In this Figure the information included in the CG is based on feature maps, being these features obtained using a curvature-based environment description from laser scan data [17]. However, the next section shows how this algorithm is able to be used for any sort of features and sensor.

## IV. Experimental Results

The proposed feature-based map partitioning algorithm has been extensively tested in real indoor scenarios for laser rangefinder and stereo cameras. In this section, just one experiment for each type of sensor is shown. The first one is depicted in Figure 6. It has been carried out in a Pioneer 2AT robot platform from ActivMedia equipped with a SICK LMS200. The field of view is $180^o$ in front of the robot and up to 8 m distance. The range samples are spaced every half a degree, all within the same plane. The set of natural features used in the EKF-SLAM algorithm, and subsequently in the submap generation process, has been extracted using a curvature-based algorithm for laser scan data segmentation [17]. The whole segmentation process consists of two stages. The first stage divides the laser scan into clusters of consecutive range readings according to a distance criterion. Then, the second stage calculates the curvature function associated to each cluster and uses it to split each cluster into a set of straight-line and curve segments.

This test is composed of 1500 scans of an office-like environment where the robot was driven through a room and a corridor (see Figure 6), whose trajectory return to a previous area through a different way. The ground-truth is shown in Figure 6a. It has been obtained using the Mapper3 software from ActivMedia Robotics, whose process is based on offline scanmatching techniques applied over the complete set of scans. The local maps generated by the proposed approach are shown in Figure 6b. Here, features belonging to different submaps are depicted in different colors. Some map partition events during the trajectory are shown in Figures 6c-e.

Finally, similar results can be achieved using this submap generation process for vision. In Figure 7 some frames for different trajectories in an indoor environment are shown. These experiments has been conducted using the aforementioned platform which is also equipped with a stereoscopic camera. This stereo system is a STH-MDCS from Videre Design: a compact, low-power colour digital stereo head with an IEEE 1394 digital interface. It consists of two 1.3 megapixel, progressive scan CMOS images mounted in a rigid body, and a 1394 peripheral interface module, joined in an integral unit. Visual landmarks used in the map partitioning process are associated to distinguished regions

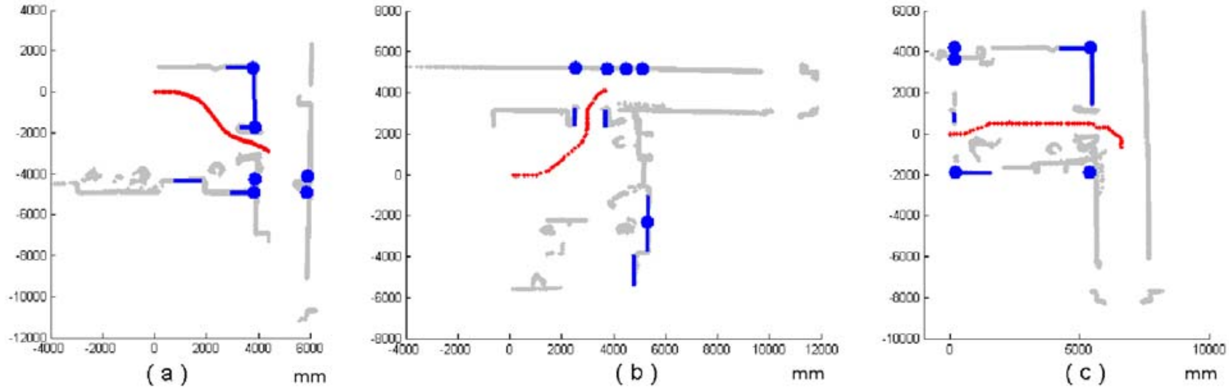


Fig. 5. Several tests in the same environment following different trajectories. The spectral clustering algorithm generates the same submaps in each experiment. When the robot leaves the room and enter in the corridor the thresholds values for $Ncut$ and the second eigenvalue are satisfied.

extracted from a stereo vision system using a perception-based grouping mechanism [18]. In these tests, a new submap is generated when the robot leaves the room (see Figure 7). In order to evaluate the robustness of the proposal, several trials have been conducted on this environment. It can be noted that the algorithm always splits the map at the same location, as can be seen in Figures 7a-c. These partitions correspond to the obtained in the laser experiment in the same environment, as is shown in Figures 6c-e (point B).

## V. Conclusions and Future Work

In this paper a spectral clustering technique for efficient graph partitioning has been used for submaps generation in a hybrid framework. Unlike other approaches, this proposal is based on feature maps without constraints about the sort of features and the sensor used. Experiments with both laser range finder and vision have been successfully conducted. Our proposal builds a covisibility graph, where nodes are related to environment features and not to robot poses, as it has been employed by previous proposals [7] [8] [12]. A near-optimal solution to partition this graph is achieved using spectral clustering. Experimental results show that the definition of the *covisibility rate* is a suitable measure which allows to compute the similarity matrix. This measure is based on the concept of feature locality, those features that are visible from the same robot pose.

The set of submaps generated provides a topological representation of the environment being an important issue to arise a hybrid approach for localization and mapping. Regarding the graph partition, combined thresholds have been used but an adaptive criterion according to the environment will be considered.

## References

[1] G.A. Borges and M.J. Aldon, "Optimal Mobile Robot Pose Estimation using Geometrical Maps", *IEEE Trans. Robotics and Automation*, vol. 18, no. 1, pp. 87–94, 2002.

[2] T. Lemaire, C. Berger, I. Jung and S. Lacroix, "Vision-based SLAM: Stereo and Monocular Approaches", *Int. Journal of Computer Vision*, vol. 74, no. 3, pp. 343-364, 2007.

[3] J. Košecká, F. Li and X. Yang, "Global Localization and Relative Positioning based on Scale-invariant Keypoints", *Robotics and Autonomous Systems*, vol. 52, pp. 27–38, 2007.

[4] A. Ranganathan, E. Menegatti and F. Dellaert, "Bayesian Inference in the Space of Topological Maps", *IEEE Trans. Robotics*, pp. 92–107, 2006.

[5] S. Simhon and G. Dudek, "A Global Topological Map formed by Local Metric Maps", *Proc. Int. Conf. Intell. Robots Syst.*, vol. 3, pp. 1708-1714, 1998.

[6] H. Choset and K. Nagatani, "Topological Simultaneous Localisation and Mapping: Towards Exact Localisation without Explicit Localisation", *IEEE Trans. Robotics and Automation*, vol. 17, no. 2, pp. 125–137, 2001.

[7] Z. Zivkovic, B. Bakker and B. Kröse, "Hierarchical Map Building and Planning based on Graph Partitioning", *Proc. Int. Conf. Robotics and Automation*, pp. 803–809, 2006.

[8] J.L. Blanco, J. González and J.A. Fernández-Madrigal, "Consistent Observation Grouping for Generating Metric-Topological Maps that Improves Robot Localization", *Proc. Int. Conf. Robotics Automation*, pp. 818–823, 2006.

[9] J.L. Blanco, J.A. Fernández-Madrigal and J. González, "A New Approach for Large-Scale Localization and Mapping: Hybrid Metric-Topological SLAM", *Proc. Int. Conf. Robotics Automation*, pp. 2061–2067, 2007.

[10] M. Bosse, P. Newman, J. Leonard, M. Soika, W. Feiten and S. Teller, "An Atlas Framework for Scalable Mapping", *Proc. Int. Conf. Robotics Automation*, pp. 1899–1906, 2003.

[11] C. Estrada, J. Neira and J.D. Tardós, "Hierarchical SLAM: Real-Time Accurate Mapping of Large Environments", *IEEE Trans. Robotics*, vol. 21, no. 4, pp. 588–596, 2005.

[12] E. Brunskill, T. Kollar and N. Roy, "Topological Mapping Using Spectral Clustering and Classification", *Proc. Int. Conf. Intell. Robots and Syst.*, pp. 3491–3496, 2007.

[13] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 8, pp. 888–905, 2000.

[14] C. Valgren, T. Ducket and A. Lilienthal. "Incremental Spectral Clustering and Its Application to Topological Mapping", *Proc. Int. Conf Robotics and Automation*, pp. 4283–4288, 2007.

[15] R. Vázquez-Martín, P. Núñez, J. C. del Toro, A. Bandera and F. Sandoval. "Simultaneous Mobile Robot Localization and Mapping using an Adaptive Curvature-based Environment Description", *Proc. 14th IEEE Mediterranean Electrotechnical Conference*, pp. 343–349, 2008.

[16] J. Neira, J. D. Tardós and J. A. Castellanos, "Linear Time Vehicle Relocation in SLAM", *Proc. Int. Conf. Robotics and Automation*, pp. 427–433, 2003.

[17] P. Núñez, R. Vázquez-Martín, J.C. del Toro, A. Bandera and F. Sandoval. "Natural landmark extraction for mobile robot navigation based on an adaptive curvature estimation", *Robotics and Autonomous Systems* vol. 56, pp. 247–264, 2008.

[18] R. Marfil,A. Bandera, and F. Sandoval. "Perception-based Image Segmentation Using the Bounded Irregular Pyramid", *In: Proc. of the 29th German Soc. Pattern Recognition Symposium (DAGM'07)*, pp. 244–254, 2007.

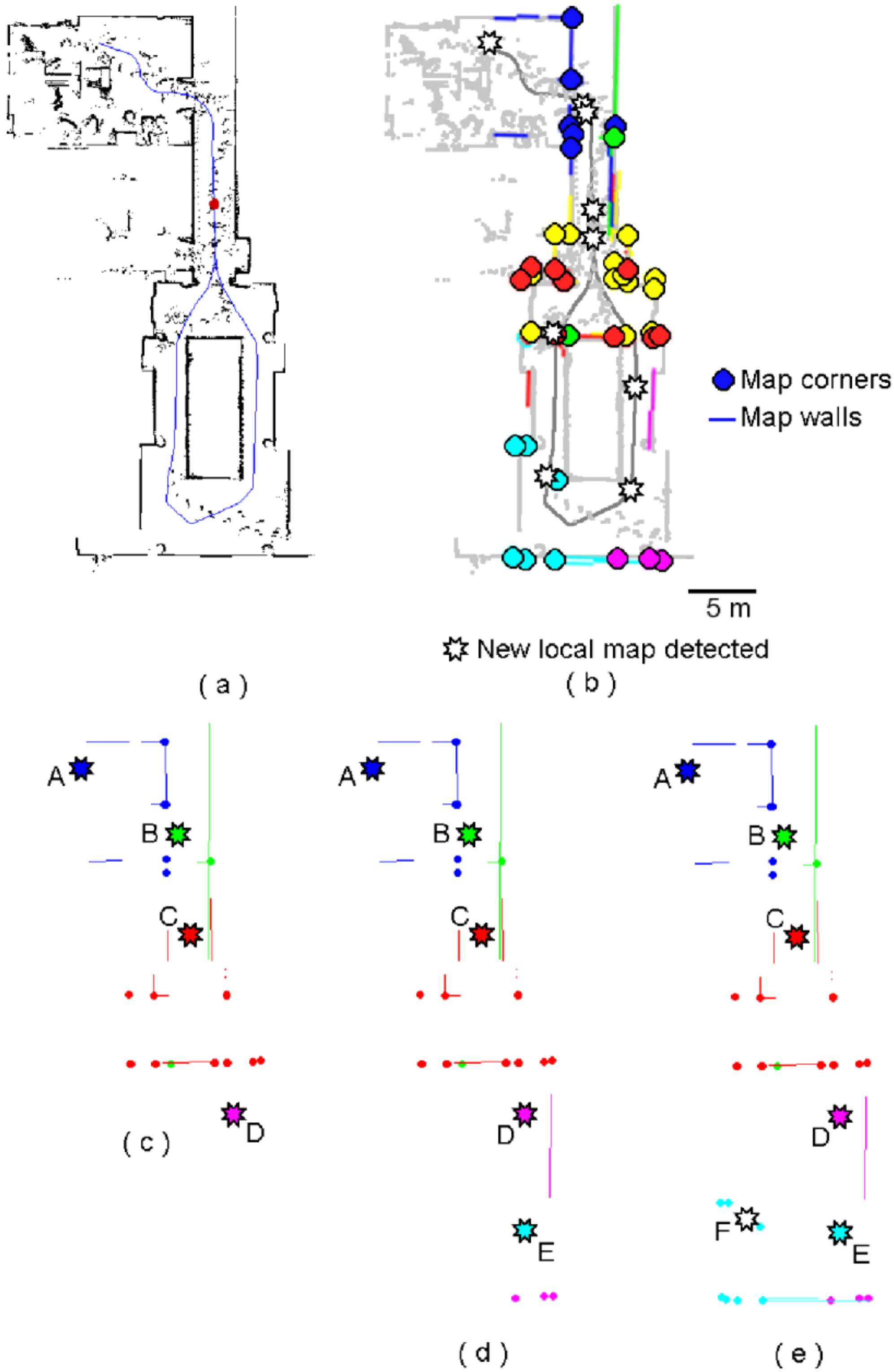


Fig. 6. Experiment in a real indoor environment with a laser range finder. a) Environment layout; b) detected submaps; c) submaps generated during the trajectory up to the third cut, d) and e) following partitions in the trajectory. Features have been obtained using a curvature-based laser scan data segmentation algorithm [17].


Fig. 7. Some experiments using vision. Three experiments in the same scenario in different viewing conditions. Visual landmarks detected are represented as ellipses, and the correspondence between landmarks in different frames with the corresponding index. A new area is detected at the same location as for the laser experiment. a) new area detected in frame #70, b) in frame #104 and c) in frame #71.