ROBOT 2011

ROBÓTICA EXPERIMENTAL

28-29 Noviembre 2011. Escuela Superior de Ingenieros de la Universidad de Sevilla (España)

Libro de Actas

Editores:

A. Ollero
J.R. Martinez de Dios
J.A. Cobano
J.Ferruz
F. Caballero
L. Merino
A. Rodriguez-Castaño

Con la colaboración de los organizadores de sesiones: J. Agirre , J. Amat, M. Armada , R. Aracil, C. Balaguer, L. Basañez, A. Casals, J.A.Castellanos, A. García-Cerezo, A. González, V. Matellán, L. Montano, J. M. M. Montiel, V. Muñoz, J.L. Pons , O. Reinoso, P. Ridao, A. Sanfeliú, P. Sanz y F. Torres.

Comité de Programa

Anibal Ollero (Presidente), *Univ. de Sevilla, FADA-CATEC*
Jon Agirre, *Tecnalia*
Josep Amat, *Univ. Politècnica de Catalunya*
Rafael Aracil, *CAR - Univ. Politécnica de Madrid*
Manuel A. Armada, *CAR - CSIC*
Carlos Balaguer, *Univ. Carlos III de Madrid*
Luís Basañez, *Univ. Politècnica de Catalunya*

Fernando Caballero, *Univ. de Sevilla*
Alicia Casals, IBEC - *Univ. Politècnica de Catalunya*
Jose Ángel Castellanos, *Univ. de Zaragoza*

Josep Lluís de la Rosa, *Univ. de Girona*
Vicente Feliú, *Univ. Castilla la Mancha*
Alfonso García-Cerezo, *Univ. de Málaga*
Fernando Gómez, *Univ. de Huelva*
Antonio González, *Univ. de Granada*
José Ramiro Martínez, *Univ. de Sevilla*

Vicente Matellán, *Univ. de León*
Luis Merino, *Univ. Pablo Olavide*
Luis Montano, *Univ. de Zaragoza*
Victor Muñoz, *Univ. de Málaga*
José Luís Pons, *CAR - CSIC*

Óscar Reinoso, *Univ. Miguel Hernández*
Pere Ridao, *Univ. de Girona*
Miguel Ángel Salichs, *Univ. Carlos III de Madrid*
Alberto Sanfeliú, *Univ. Politècnica de Catalunya*
Pedro Sanz, *Univ. Jaume I*
Rafael Sanz, *Univ. de Vigo*
Fernando Torres, *Univ. de Alicante*
Eduardo Zalama, *Univ. de Valladolid*

Comité Organizador

Publicaciones:	José Antonio Cobano, *Univ. de Sevilla*
Finanzas:	Joaquín Ferruz, *Univ. de Sevilla*
Web/Difusión:	Begoña C. Arrúe, *Univ. de Sevilla*
Gestión local:	Guillermo Heredia, *Univ. de Sevilla*
Registro:	Ángel Rodríguez, *Univ. de Sevilla*

Patrocinadores

Vicerrectorado de Investigación, Universidad de Sevilla.

Escuela Superior de Ingenieros, Universidad de Sevilla

Grupo Temático de Robótica, Comité Español de Automática

Plataforma Tecnológica Española de Robótica

CONET Network of Excellence, FP7

Red Temática en Visión por Computador, Comité Español de Automática

Organizadores

Grupo de Robótica, Visión y Control

Centro Avanzado de Tecnologías Aeroespaciales

Índice general

Actas ROBOT 2011. 28-29 de Noviembre de 2011. Sevilla (España)

Visual Odometry with an Appearance-based Method

Lorenzo Fernández, Luis Payá, Miguel Juliá, Francisco Amorós and Oscar Reinoso

Abstract— **In this paper we present an efficient method to solve the problem of map creation and localization of a mobile robot using omnidirectional images. We propose a real-time algorithm for topological mapping, using as input data only a set of images captured by a single omnidirectional camera mounted at a fixed position on the mobile robot. We use techniques based in the global appearance of the images to compute the topological relationships between locations in the map. When using these methods, it is important to remove redundant information to get an acceptable computational cost when comparing locations. With this aim, we describe each omnidirectional image by a single Fourier descriptor that represents the appearance as well as the relative orientation between images. This algorithm permits computing the relative topological position of a location with respect to the previous one, acting as a visual compass. Experiments were conducted in a variety of environments to study the validity of the proposed visual odometry and topological mapping and to perform an objective comparison between the results obtained using the robot odometry, our visual odometry and the ground truth. We have also checked the time consumption to carry out the process and the geometrical accuracy obtained comparing to the ground truth.**

I. INTRODUCTION

In typical applications where a mobile robot must carry out a task autonomously, it is important that the mobile robot has a map or an internal representation of the environment continuously updated, so that the robot can make decisions about its localization and about the path to follow to move from its current position to the target points. Omnidirectional vision systems are commonly used (sometimes in conjunction with encoder sensors) at this kind of applications due to their relative low cost and the richness of the information they provide. Different representations of the visual information can be used when working with these catadioptric systems, such as the omnidirectional, panoramic and bird eye view images [1]. In this paper, we use the panoramic representation of the scenes as it can offer invariance to ground-plane rotations, and it allows us to use only the information provided by the vision sensor to perform the process.

Severals authors have studied how to use the omnidirectional images both to solve the mapping and the localization problems. We can categorize these solutions into two main groups: feature-based and appearance-based solutions. In the first group, a number of significant points or landmarks from each image are extracted and each point is described using an invariant descriptor. For example, [2] and [3] use SURF features [4] extracted from a set of omnidirectional images to find the localization of the robot in a given map, and [5] uses SIFT features [6] in localization and mapping tasks. In the second group, the whole appearance of the omnidirectional image is represented by a single descriptor, with no local feature extraction. For example [7] presents a method to build a topological map using a Fourier descriptor of omnidirectional images and [8] performs a probabilistic localization in some environments. [9] uses PCA (Principal Components Analysis) features of panoramic images for environment modeling and localization.

When working in unstructured environments, where the creation of appropriate models of recognition can be an arduous chore, it is useful using appearance-based techniques. These techniques offer an intuitive way to construct the map and to get the position. However, as no relevant information is extracted from the images, it is necessary to apply a compression method to reduce the computational cost of the mapping and localization processes. PCA (Principal Components Analysis) is a widely extended method used to extract the most relevant information from a set of images [10]. However, this family of methods is not inherently invariant to the ground-plane orientation of the robot nor an incremental method. Some algorithms can be found in the literature to solve these problems [11], but they usually present a quite high computational cost that makes online topological mapping unfeasible. Other researchers have developed DFT (Discrete Fourier Transform) methods to get the most relevant information from the images [12]. These descriptors present rotational ground-plane invariance and concentrate the most relevant information in the low frequency components of the transformed images. Also, each image descriptor can be computed independently of the rest of images.

Taking these facts into account and based on some prior works [13], [8], [14], we have decided to describe each omnidirectional image by means of a Fourier descriptor. We use the Fourier Signature [12] to compress each image captured. The processing time needed to compute the Fourier Signature is noticeably lower than in other common feature extraction algorithms, and it permits a fast comparison between the images in the map by means of a vector distance measurement. Also, when using the Fourier Signature, we exploit better the invariance to ground-plane rotations in panoramic images, property that will be of utmost importance when deploying our visual compass.

In this paper we present a methodology to build a topological map using the global appearance of the panoramic images to model the topological relationships between successive nodes in the map. We use the Fourier Signature to get a

Departamento de Ingenieria de Sistemas y Automatica, Miguel Hernandez University, Avda. de la Universidad s/n, Elche, Spain
l.fernandez@umh.es; l.paya@umh.es

robust descriptor that allow us to work in real-time. However, the methods described here are independent of the descriptor used to represent the images, and other appearance-based descriptors may also be applied. To represent the distance between two consecutive poses we have used the normalized Euclidean distance between their Fourier Signatures and to get the relative angle between them we have implemented a Fourier-based visual compass. Our main objective consists in evaluating the feasibility of using purely global-appearance methods in these tasks and how the main features of the descriptor influence the final result. The reliability of this method in the navigation tasks will make it useful in more complicated tasks that require an internal representations of the environment, such as the use of an autonomous robot as a guide in a museum or hospital, or the use of a mobile vehicle for the exploration of unknown environments in military search and rescue tasks.

The paper is organized as follows. Section 2 presents the fundamentals of topological mapping approaches. In section 3, we describe the Fourier descriptor and how to use it with omnidirectional images to implement the visual compass. Section 4 deals with the problem of localization and map creation using visual odometry. Next, Section 5 presents the experimental setting and the results obtained. Finally, we present the conclusions and future work in Section 6.

II. TOPOLOGICAL MAP BUILDING. STATE OF THE ART

With respect to the mapping problem we can establish two approaches: *metric* and *topological*. The first one consists in modeling the environment using a metric map obtained with geometrical accuracy when representing the position of the robot in it. For example, [15] describes a sonar-based mapping system developed for mobile robots navigation, [16] presents an approach to carry out the mapping process with a team of mobile robots and visual information and [17] demonstrate how the usage of an odometer with LBA-based (Local Bundle Adjustment) monocular visual SLAM permits improving the pose accuracy in a robot SLAM process. On the other side, topological mapping consists in the creation of maps that represent graphical models of the environment that capture places and their connectivity in a compact form. An example of this approach is presented in [8] where a topological representation of the environment is obtained by applying a method based on the physics of harmonic oscillators. [18] shows how a topological map of the environment can be obtained using fast vision techniques from a sequence of color histograms. Also, [19] describes a probabilistic method for topological SLAM, solving the topological graph loop-closing problem. At last, [20] studies how to build a topological representation of large indoor and outdoor environments using local features extracted from omnidirectional images and the epipolar constraint, and a clustering method to perform localization more efficiently and [14] describes a Monte-Carlo Localization approach using the robot odometry and the appearance of omnidirectional images to localize in a topological map.

Fig. 1. A robot rotation in the ground plane produces a shift in the columns of the panoramic images captured.

Recovering relative robot poses from a set of camera images has been a largely studied problem in recent years. For example [21] presents a system that estimates the motion of a stereo head using a feature tracker or [22] describes a real-time algorithm for computing the ego-motion of a vehicle using as input only omnidirectional images and a homography-based tracker that detects and matches robust scale-invariant features. In these cases, to get a robust and accurate map it is interesting to perform a loop closing process, to correct the cumulative errors along the process. [23] studies how to close the loop by using the omnidirectional visual odometry and a vocabulary tree. This work shows how it is possible to carry out a process of robot localization and mapping simultaneously using as input data only omnidirectional images and a loop-closing process.

In this paper, we face the mapping problem as a *relative camera pose recovering* problem, using the overall appearance of the panoramic images, without any feature extraction process. We describe a real-time algorithm for computing an appearance-based topological map through visual odometry. The main contributions of the work are the development of a visual compass that permits computing the position and orientation of each new location in the map, with a low computational cost; the development of a method to estimate the accuracy of the map built comparing to the real trajectory of the robot and the study of the parameters of the descriptor that influence the process and the final results. Solving the appearance-based visual SLAM (Simultaneous Localization and Mapping) problem is out of the scope of this work. In this proposal, we show how it is possible to obtain the topological relative position of the vehicle at each moment, and in a next phase of the work we will include loop closing to carry out this process of topological visual SLAM.

III. FOURIER SIGNATURE WITH OMNIDIRECTIONAL IMAGES

In this section we present the method we have used to build the appearance descriptor of each omnidirectional image. This descriptor must present a low computational cost, to allow working in real time and it has to be built in an incremental way. We have built our descriptor using the Fourier Signature [12].

A. Fourier Signature

The Fourier Signature presents several advantages among other Fourier-based methods. It is simple to compute, it presents a low computational cost in terms both of computation time and memory required, and it exploits well the invariance against ground-plane rotations using panoramic images. From a panoramic image I^j that has N_x rows and N_y columns, it is possible to obtain the most relevant information from the image by means of the Discrete Fourier Transform [13]. The Fourier Signature consists in expanding each row of the panoramic image $\{a_k\} = \{a_0, a_1, \ldots, a_{N_y-1}\}$ using the Discrete Fourier Transform into the sequence of complex numbers $\{A_k\} = \{A_0, A_1, \ldots, A_{N_y-1}\}$.

The Fourier Signature presents the same properties as the 2D Discrete Fourier Transform. The most important information is concentrated in the low frequency components of each row, so we can work only with the information from the k_1 first columns in the signature ($k_1 < N_y$), and it presents rotational invariance when working with panoramic images. It is possible to prove that if each row of the original image is represented by the sequence $\{a_n\}$ and each row of the rotated image by $\{a_{n-q}\}$ (being q the amount of shift) (Fig. 1), when the Fourier Transform of the shifted sequence is computed, we obtain the same amplitudes A_k than in the non-shifted sequence, and there is only a phase change, proportional to the amount of shift q (1).

$$F[\{a_{n-q}\}] = A_k \exp\left(-j\frac{2\pi q l}{N_y}\right) \qquad l = 0, \ldots, N_y - 1 \tag{1}$$

Thanks to this shift theorem we can separate the computation of the robot position and the orientation. With this aim, we decompose the Fourier Signature in two matrices, one containing the modules $d^j \in \Re^{N_x,k_1}$ and the other the phases $p^j \in \Re^{N_x,k_2}$ of this signature.

Finally, it is interesting to highlight also that the Fourier Signature is an inherently incremental method (what differs from other appearance-based descriptor, such as PCA (Principal Components Analysis) [9]).

B. Visual Compass

Once we have studied the kind of information to store in the database, we have to establish some relationships between the stored poses to carry out the relative camera pose recovering.

When we have the Fourier Signature of two panoramic images that have been captured in two points that are geometrically close (depending on the environment, under $0.8\ m$ in our experiments), it is possible to compute their relative orientation using the shift theorem (1).

Since the Fourier Signature is invariant to ground-plane rotations and there is a relationship between phases of the Fourier Signature of a panoramic image taken at one position and the phases of the Fourier Signature of another panoramic image taken at the same point but with different orientation (eq. 1), we can expand this property and calculate the approximate rotation $\phi^{t+1,t}$ between two panoramic images taken on two consecutive poses. In fig. 2, v^t is the velocity of the robot at time t, v^{t+1} at time $t+1$ and the relative orientation $\phi^{t+1,t}$. The angle obtained corresponds to the rotation the robot has performed in the ground-plane when going from the first to the second point as shown in fig. 2.

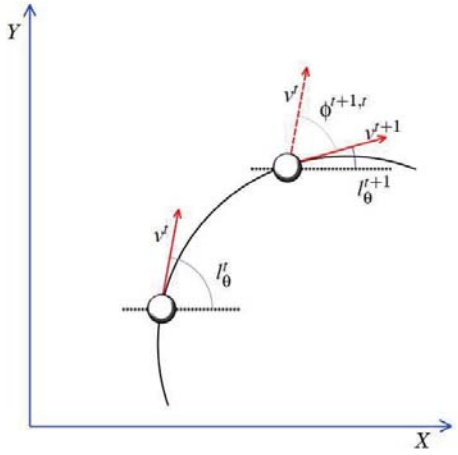


Fig. 2. Relative orientation between two consecutive nodes in the map.

To obtain $\phi^{t+1,t}$ we have implemented a convolution operation between the phases of the Fourier Signatures of the panoramic images of the two poses by applying (1).

IV. TOPOLOGICAL MAP CREATION

A. Map representation

The objective of this section is to show how the map is built as the robot goes through the environment to map. We build a graph-based map where when a new image is captured, a new node is added to the map, and the topological relationships with the previous node are computed using the global appearance information of the scenes. With our procedure, this computation is made online, as the robot is going through the environment, in a simple and robust way.

In this case we consider that our map is composed of a set of nodes $L = \{l^1, l^2, \ldots, l^N\}$. Each node l^j is represented by an omnidirectional image $I^j \in \Re^{N_x,N_y}$ associated and a Fourier descriptor that describes the global appearance of the omnidirectional image, composed of a modules matrix $d^j \in \Re^{N_x,k_1}$ and a phases matrix $p^j \in \Re^{N_x,k_2}$. Also, with the algorithm implemented, we can compute the position $(l^j_{\mathrm{x}}, l^j_{\mathrm{y}})$ and the orientation l^j_θ of each node in the map (with respect to the previous one) thus $l^j = \{(l^j_{\mathrm{x}}, l^j_{\mathrm{y}}, l^j_\theta), d^j, p^j, I^j\}$.

We consider that the robot captures a new image at time $t+1$ and then, the Fourier descriptors d^{t+1} and p^{t+1} are computed. Comparing it with the descriptors of the previously captured image d^t and p^t we can find the topological relationships between these two nodes. We can separate the computation of the robot position and the orientation at time $t+1$ $(l^{t+1}_{\mathrm{x}}, l^{t+1}_{\mathrm{y}}, l^{t+1}_\theta)$ thanks to the shift theorem (1).

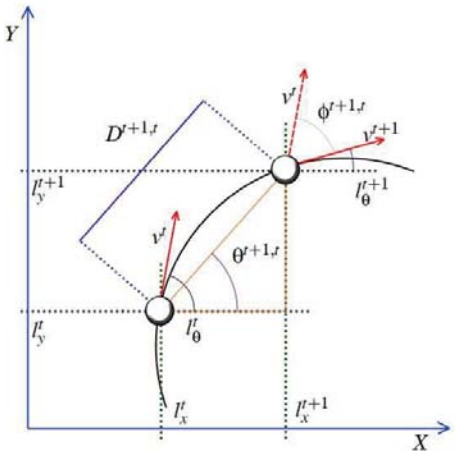


Fig. 3. Position and orientation of the new node in the map computed incrementally from the previous node.

In the surroundings of the point where one image is taken, the distance between Fourier Signatures is approximately proportional to the actual geometrical distance [14]. To compute the difference between the appearance of two scenes, we use the Euclidean distance between the modules of the Fourier signature. If d^i is the Fourier signature of the image I^i and d^j is the Fourier signature of the image I^j, then the distance between scenes i and j is:

$$D^{i,j} = \sqrt{\sum_{u=0}^{N_x} \sum_{v=0}^{k_1} (d^i(u,v) - d^j(u,v))^2} \quad (2)$$

On the other hand, thanks to the visual compass implemented, we can estimate the relative orientation between images. After this process, the position of the current node is computed from the previous node as:

$$l_x^{t+1} = l_x^t + D^{t+1,t} \cdot cos(\theta^{t+1,t}) \quad (3)$$

$$l_y^{t+1} = l_y^t + D^{t+1,t} \cdot sin(\theta^{t+1,t}) \quad (4)$$

$$l_\theta^{t+1} = l_\theta^t + (\theta^{t+1,t}) \quad (5)$$

These relationships are shown graphically in fig. 3.

B. Shape Difference between the resulting map and the actual map

Once the topological map is built from the Visual Odometry data, we need a mechanism to test the performance of our approach. We have decided to evaluate how similar is the layout of the resulting map comparing to the actual map (*Ground Truth* or *Real Map*) of the images captured. As the only information we have used to build the map is the distance $D^{i,j}$ between Fourier Signatures and the relative position angle $\phi^{t+1,t}$ between them, the resulting map is expected to have a similar shape comparing to the original map but with a scale factor, a rotation and a possible reflection. We can remove these effects in the resulting topological map to get an acceptable measure of the shape difference between the original and the resulting map layout [8]. With this aim, we use a method based on the Procrustes analysis. This method is a statistical shape analysis that evaluates shape correspondence [24].

First, we arrange the coordinates of the points that form the real map (where the images were captured) in a matrix $A = [(x_1, y_1)^T, (x_2, y_2)^T, \ldots, (x_n, y_n)^T]^T$ and we arrange the coordinates of the obtained topological map in a matrix $C = [(l_x^1, l_y^1)^T, (l_x^2, l_y^2)^T, \ldots, (l_x^n, l_y^n)^T]^T$. With the Procrustes analysis we can compare the shape of these two sets of points by determining a linear transformation (a translation c, a reflection, an orthogonal rotation T and a scaling b) of the points in the matrix C so that the points in $b \cdot C \cdot T + c$ to best conform to the points in the matrix A.

When the translation, rotation and scaling effects have been removed, the sum of squared errors is a good fit criterion. Thanks to this analysis we can measure how accurate is the layout of the topological map, comparing to the layout where the images were taken.

This analysis has a closed form, as detailed in a [24]. As a result of this process, a parameter $\mu \in [0, 1]$ can be obtained. μ is a measure of the shape correspondence between the sets of points A and C. The lower is μ, the more similar are A and C. We name this parameter *shape difference* along the paper. We use this difference with the only purpose to know the feasibility of our appearance-based Visual Odometry, and its use is possible due to the fact that we know the coordinates of the points in the original map (ground truth).

V. EXPERIMENTS

To carry out the experiments, we have used two different sets of omnidirectional images. The first set has been captured in an office environment, when the robot performs the trajectory shown in fig. 4 (ground truth), which includes a loop closing. Ground truth graphics are obtained with laser and robot odometry fusion. The first set is composed by 200 images with a distance of $10cm$ between images. The second set (fig. 5) has been captured in a laboratory environment. It is composed by 150 images and the image acquisition has been automated so that a new image is captured when the difference with the previous one goes over a threshold.

We have designed a complete set of experiments in order to test the validity of the global appearance-based approach in topological map building. With this aim, we study some features that define the feasibility of the procedures, such as the accuracy of the map built (how similar it is comparing to the actual map) and the computational cost of the method. We also study how these features are influenced by some typical parameters, such as the number of images, the distance between them, the degree of compression during the Fourier Transformation and the number of components used to compute the difference in the visual compass.

Fig. 6 shows an example of the map computed with the first set of images and the actual map (ground truth). This map has been built using all the images of the set

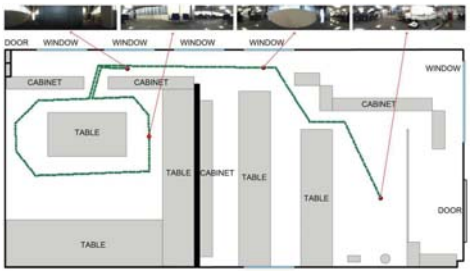


Fig. 4. Trajectory followed by the robot when capturing the first set of images.

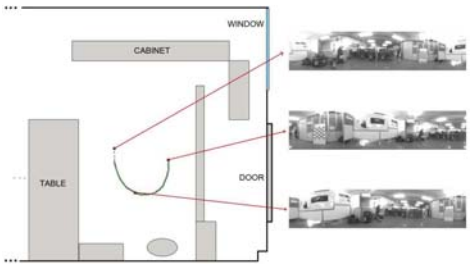


Fig. 5. Trajectory followed by the robot when capturing the second set of images.

(geometrical distance between images equals 10 cm.), $k_1 = 64$ module components to compute distance between Fourier descriptors and $k_2 = 64$ phase components used to compute relative orientations in the visual compass. Fig. 7 shows an example with the second set of images. The visual odometry map has been built with all the images in the set, and $k_1 = k_2 = 64$. In this case, we compare it with the actual map and with the map computed using the odometry of the robot. The map obtained with our visual odometry algorithm clearly outperforms the map obtained with the odometry of the robot.

In the set of experiments we have implemented to study the performance of our algorithm, we have tested the influence of the number of module components to compute the distance $D^{t+1,t}$ (k_1) and the number of components to compute the phase difference (k_2). We also test the influence of the geometrical distance between the points where the images were captured. To study it, we have built the maps using all the images in the sets (step = 1), every two images (step = 2) and one every three images (step = 3).

Fig. 8 shows the shape difference of the resulting map comparing to the actual map when using the first set of images. The step we take to build the map is (a) step = 1, (b) step = 2, (c) step = 3 and (d) step = 4. Fig. 9 shows the same results when building the map with the second set of images. When we use all the images of set 1, the minimum shape distance is around 0.04 when $k_1 = 26$ and $k_2 = 20$. In set 2, this factor is 0.015 when $k_1 = 4$ and $k_2 = 8$. The shape distance tends to increase when k_2 does, due to the fact

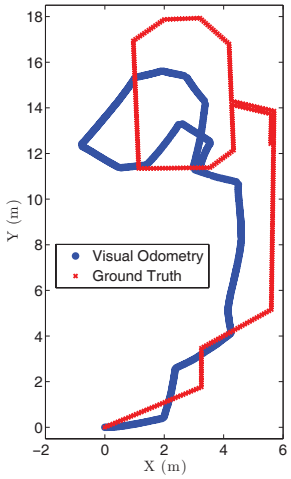


Fig. 6. Example of map built with the visual odometry and the first set of images and ground truth.

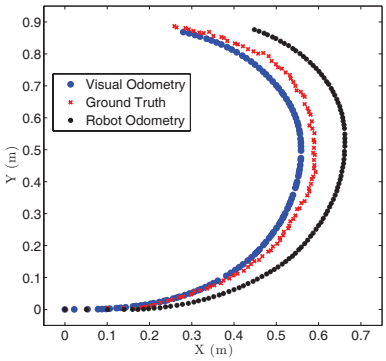


Fig. 7. Example of map built with the visual odometry and the second set of images, ground truth and map built using the robot odometry.

that the first components contain the main information so, the high frequency components may be adding noise to the computation. As far as k_1 is concerned, the tendency is not clear but, in general, the shape distance is quite insensitive to this parameter.

Fig. 10 shows the necessary time to build the complete map using the first set of images. The step we take to build the map is (a) step = 1 and (b) step = 2. Fig. 11 shows the same results when building the map with the second set of images. This time is much more dependent on the number of phase components (k_2) that on the number of module components (k_1). This is due to the fact that the computation of the relative angle is a computationally much more expensive process than computing the Euclidean

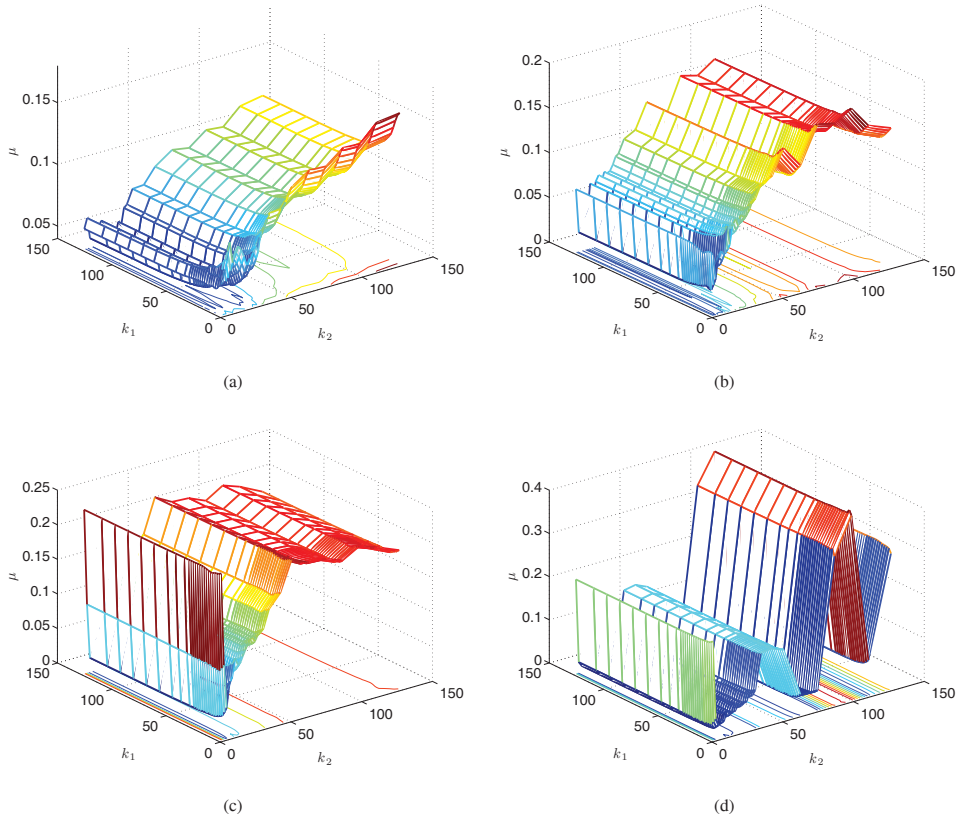


Fig. 8. Shape difference versus number of module components (k_1) and phase components (k_2) for the first set of images taking (a) all the images (b) one every two images, (c) one every three images and (d) one every four images.

distance. This way, if possible, it is more interesting to keep k_2 low so that the process can be carried out online, while the robot is going through the environment to map.

VI. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

In this paper we have studied the applicability of the approaches based on the global appearance of omnidirectional images in topological mapping, using a set of images a robot has captured when running a trajectory in an environment. The main contributions of the paper include the development of a visual compass that allows building a map of the environment online, while the robot is going through the environment, the development of a method to compare the accuracy of the layout of the map computed and the study of the influence of the parameters of the process both in the layout of the resulting map and in the processing time.

To carry out the experiments, we have used two sets of omnidirectional images captured by a catadioptric vision system mounted on the mobile robot. Once the image has been captured, each scene is described through a Fourier-based signature that presents a good performance in terms of amount of memory and processing time, and it is also invariant to ground-plane rotations and an inherently incremental method.

We present a methodology to build graph-based maps of the environment. As we use a topological approach, these maps represent the real world except for a scale factor and a rotation. To make an homogeneous comparison between the map computed and the real map, we have developed a method based in the Procrustes analysis. As shown in the results, when the parameters of the system are correctly tuned, accurate results can be obtained, maintaining a reasonable computational cost. Thus, one can conclude that our algorithm will be used to perform navigation tasks in unknown environments (disaster areas or war) or in familiar surroundings in which it is desirable to perform a repetitive task such as the use of a museum guide robot.

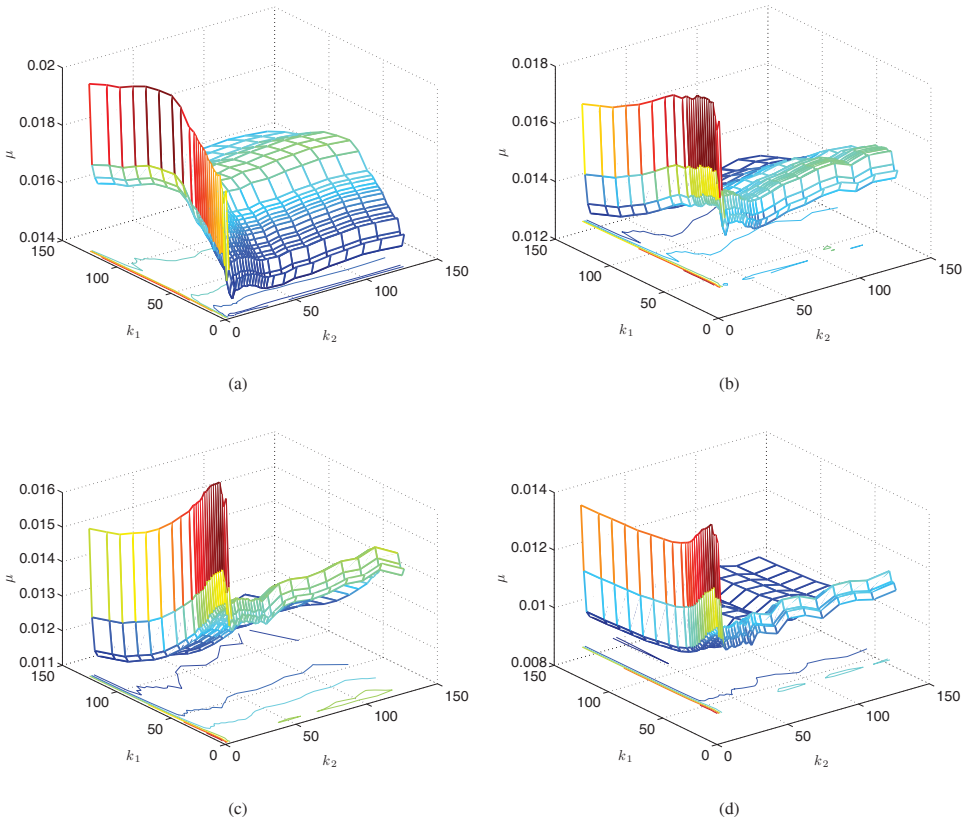


Fig. 9. Shape difference versus number of module components (k_1) and phase components (k_2) for the second set of images taking (a) all the images (b) one every two images, (c) one every three images and (d) one every four images.

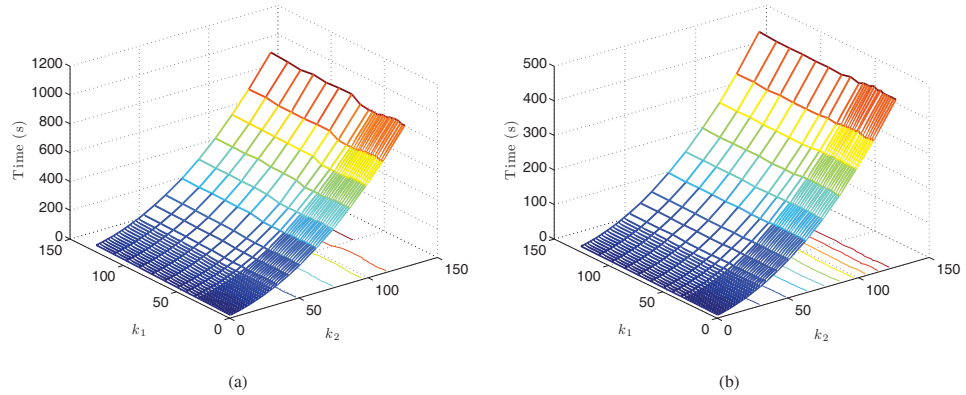
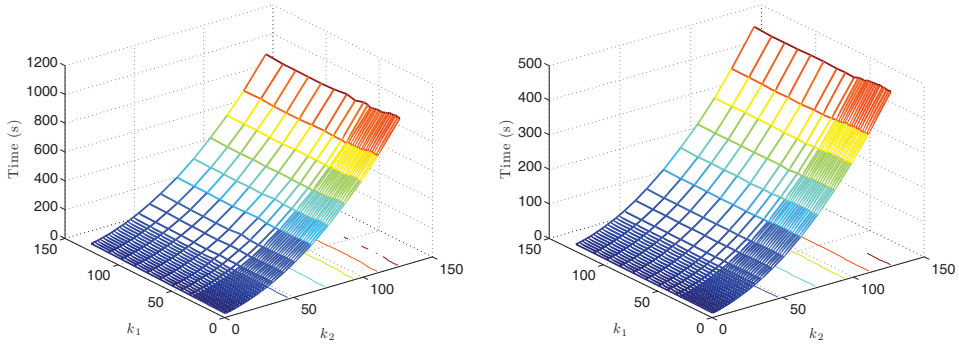


Fig. 10. Necessary time to build the complete map with the first set of images when using (a) all the images (b) one of every two images.



(a) (b)

Fig. 11. Necessary time to build the complete map with the second set of images when using (a) all the images (b) one of every two images.

B. Future Works

We are now working in this approach to build a topological SLAM algorithm (Simultaneous Localization and Map Building) using just the global appearance of omnidirectional images and different map topologies. Some features to deepen in are the initialization of the algorithm and the scale factor to get a representation with geometrical accuracy.

VII. ACKNOWLEDGMENTS

This work has been supported by the Spanish government through the project DPI2010-15308. 'Exploración integrada de entornos mediante robots cooperativos para la creación de mapas 3D visuales y topológicos que puedan ser usados en navegación con 6 grados de libertad'.

REFERENCES

[1] N. Winters, J. Gaspar, G. Lacey and L. Santos, "Omni-directional Vision for Robot Navigation", *in IEEE Workshop on Omnidirectional Vision (OMNIVIS'00)*, 2000.

[2] H. Murillo, J. J. Guerrero and C. Sagues, "Surf features for efficient robot localization with omnidirectional images", *in Robotics and Automation, 2007 IEEE International Conference (ICRA)*, 2007.

[3] C. Valgren and A. Lilienthal, "Sift, surf and seasons: Appearance-based long-term localization in outdoor environments", *in Robotics and Autonomous Systems*, 2010.

[4] H. Bay, A. Ess, T. Tuytelaars and L. V. Gool, "Speeded-up robust features (surf)", *in Comput. Vis. Image Underst. Elsevier Science Inc*, 2008.

[5] S. Se, D. Lowe and J. Little, "Vision-Based Global Localization and Mapping for Mobile Robots", *in IEEE Transactions on Robotics*, 2005, 21(3), pp. 364-375.

[6] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *in International Journal of Computer Vision*, 2004, 2(60), pp. 91-110.

[7] E. Menegatti, M. Zoccarato, M. Pagello and H. Ishiguro, "Image-based monte-carlo localisation with omnidirectional images", *in Robotics and Autonomous Systems*, 2004.

[8] L. Payá, L. Fernández, A. Gil and O. Reinoso, "Map building and monte carlo localization using global appearance of omnidirectional images", *Sensors*, 2010, 10(12), pp. 11468-11497.

[9] B. Kröse, R. Bunschoten, S. T. Hagen, B. Terwijn and N. Vlassis, "Environment Modeling and Localization from an Omnidirectional Vision System", *in IEEE Robotics and Automation Magazine*, 2004, 11(4), pp. 45-52.

[10] M. Kirby, "Geometric data analysis", *Wiley Interscience*, 2001.

[11] M. Jogan and A. Leonardis, "Robust Localization Using Eigenspace of Spinning-Images", *in Proc. of the IEEE Workshop on Omnidirectional Vision*, 2000, pp. 37-44.

[12] E. Menegatti, T. Maeda and H. Ishiguro, "Image-based memory for robot navigation using properties of omnidirectional images", *in Robotics and Autonomous Systems*, 2004, 47: pp. 251-276.

[13] L. Payá, L. Fernández, O. Reinoso, A. Gil and D. Úbeda, "Appearance-based dense maps creation - comparison of compression techniques with panoramic images", *in 6th International Conference on Informatics in Control, Automation and Robotics ICINCO 2009*, 2009, pp. 250-255.

[14] L. Fernández, A. Gil, L. Payá and O. Reinoso, A. Gil and D. Úbeda, "An evaluation of weighting methods for appearance-based monte carlo localization using omnidirectional images", *in IEEE International Conference on Robotics and Automation (ICRA, 2010 Workshop on Omnidirectional Robot Vision)*, 2010.

[15] H. Moravec and A. Elfes, "High resolution maps from wide angle sonar", *in IEEE International Conference on Robotics and Automation*, 1985.

[16] A. Gil, O. Reinoso, M. Ballesta, M. Juliá and L. Payá, "Estimation of visual maps with a robot network equipped with vision sensors", *Sensors*, 2010, 10: pp. 5209-5232.

[17] A. Eudes, M. Lhuillier, S. Naudet-Collette and M. Dhome, "Fast odometry integration in local bundle adjustment-based visual slam", *in International Conference on Pattern Recognition*, 2010, 0: pp. 290-293.

[18] F. Werner, F. Maire and J. Sitte, "Topological slam using fast vision techniques", *in Proceedings of the FIRA RoboWorld Congress 2009 on Advances in Robotics*, 2009, pp. 187-196.

[19] S. Tully, G. Kantor, H. Choset and F. Werner, "A multi-hypothesis topological slam approach for loop closing on edge-ordered graphs", *in Proceedings of the 2009 IEEE/RSJ international conference on Intelligent robots and systems*, 2009, pp. 4943-4948.

[20] C. Valgren and A. Lilienthal, "Incremental spectral clustering and seasons: Appearance-based localization in outdoor environments", *in ICRA 2008. IEEE International Conference on Robotics and Automation*, 2008, pp. 1856 -1861.

[21] D. Nistér, O. Naroditsky and J. Bergen, "Visual odometry for ground vehicle applications", *Journal of Field Robotics*, 2006, 23.

[22] D. Scaramuzza and R. Siegwart, "Appearance guided monocular omnidirectional visual odometry for outdoor ground vehicles", *in IEEE Transactions on Robotics*, 2008, 24(5): pp. 1015-1026.

[23] D. Scaramuzza F. Fraundorfer and M. Pollefeys, "Closing the loop in appearance-guided omnidirectional visual odometry by using vocabulary trees", *in Robot. Auton. Syst*, 2010, 58: pp. 820-827.

[24] G. Seber, "Multivariate observations", *Wiley Interscience*, 1984.