Management of Landmarks in a GIS for an Enhanced Localisation in Urban Areas

Maged Jabbour, Véronique Cherfaoui and Philippe Bonnifait

Heudiasyc UMR CNRS 6599, Université de Technologie de Compiègne, France
BP 20529 - 60205 COMPIEGNE cedex FRANCE
Tel. +33 (0)3 44 23 44 23 - Fax +33 (0)3 44 23 44 77
{Maged.Jabbour, Veronique.Cherfaoui, Philippe.Bonnifait}@hds.utc.fr

Abstract

Abstract – This paper addresses the problem of localizing a vehicle in urban environment by using natural information provided by exteroceptive sensors. For this purpose, sensors need to detect landmarks which have been characterized in a previous passage. As the amount of data can be significantly large, we propose a strategy to manage this information in a GIS (Geographical Information System). We illustrate our developments using visual landmarks made of key images and 3D points that are regrouped in local maps that correspond to the roads of a GIS layer thanks to the use of GPS data and proprioceptive sensors. Real experiments are reported to illustrate the performance of this approach which is robust to GPS outages due to poor satellite visibility in urban areas.

Key Words –Precise localization, natural landmarks, GPS, GIS.

1. Introduction

Dynamic localization with respect to a digital map is an essential task for advanced assistance systems and for intelligent vehicles. As a matter of fact, Geographical Information Systems (GIS) can provide precious data like signposts, speed limits, or hazardous zones like schools, pedestrian crossings, etc.

Positioning systems often rely on the use of a GNSS (GPS today, Glonass or Galileo in near future) that corrects the drift of a high bandwidth dead-reckoning system using odometry and inertial sensors. In urban areas and tunnels, a satellite based localization system suffers from satellite outages and multi-tracks which can decrease significantly the precision and the availability of the estimated positions. Therefore, for continuous, precise and reliable localization, the use of complementary localization systems that use landmarks is a key issue.

There are different kinds of landmarks that depend on the used exteroceptive sensors. They can be classified in the following categories:

- *Active landmarks*. Actives landmarks are beacons that contain active components in order to transmit a signal. They mainly rely on the use of radio-frequency signals (GPS pseudolites, transponders stored in the pavement, Wifi antennas, etc.). Such landmarks are usually distinguishable from each other. In this case, sensors are receivers equipped with an antenna.
- *Passive landmarks*. They are artificial landmarks pertinently located in the environment for localization purposes like magnets in the pavement or reflectors installed on post in curves for instance.
- *Natural landmarks*. Natural landmarks are features of the environment detected by on-board sensors like cameras or ladars [4],[9]. They can correspond to characteristic points (egde of windows for instance), road markings, roof of buildings, posts, curbs, etc.

The management of such landmarks in order to help or replace a GNSS receiver is the issue studied in this paper.

It is known that landmarks have to be associated in *local maps*, a local map being a set of landmarks put together because of

- memory constraints arising from the use of embedded systems,
- the need to download or update a limited amount of data from a distant server,
- the connections that exist between the landmarks, essential to compute a location.

Naturally, the use of a GIS has to be studied in order to manage a large amount of local maps geo-referenced in a global frame.
For intelligent vehicles, a GIS database is usually a set of digitalized roads provided by cartographers like NavTeQ or TeleAtlas. This topological information is very useful for the navigation tasks like path planning and route guidance. The use of this topology is also interesting for the landmark management [5] and consequently the local maps. Indeed, once the location of the vehicle and a road (described by a poly-line in the map) unambiguously map-match, the pertinent landmarks are those associated with this road. Secondly, if the vehicle is autonomous (ie driven by a regulator), it has to follow a predefined trajectory described by a set of connected roads. In this case, the pertinent landmarks are those attached with this path. Therefore, we propose in this work to map-match the landmarks with the road stored in the GIS database. Then, each road will logically define a local map.
The paper is organized as follows. Next section describes the used natural landmarks; section 3 presents the localization and map-matching method. In section 4, the geo-referencing of the landmarks is done. In section 5, the management of landmarks during navigation stage is performed. Last section is dedicated to real experiments with our experimental vehicle.

2. Natural Landmarks

In order to illustrate the concept of natural landmark, let consider visual landmarks with characteristic 3D points as used by Royer et al. [7, 8] for the navigation of a Cycab using a mono-camera at video rate. These landmarks are characteristic points in images called Harris points [3] (cf Fig. 1). Thanks to these points, one can rebuild the 3D pose of the vehicle using a sequence of images in which each landmark has been detected at least in two images. Other methods can use complex landmarks [1] or only 2D points [6].
Let us study how the key points are reference between each other in a sequence of images.

Figure 1: Example of detected visual features

First, the method consists in calculating the characteristic points in each image. Then, by a tracking method, the most relevant points are selected. The following stage consists in calculating the homogeneous coordinates *(x, y, z, t)* for each point in a common reference frame. This reference frame is related to the first image. Then, a projection matrix M_p (3x4) is calculated for each image, which makes it possible to transform the coordinates *(x, y, z, t)* of the characteristic points into coordinates *(u, v, t)* in a chosen image. Therefore, for an urban area, one obtains 3D Harris points, referenced in a local frame,(in other words the frame of the local map), and a list of key images in which these points were detected.
The online localization process during the navigation phase consists in seeking points of interest in the current image. Then, an algorithm tries to match these points with the ones of the nearest key image. Finally, the estimate of the pose of the camera is computed using the matched points between the current view and the key image.
We can remark therefore that a visual landmark is, in this case, a key image with corresponding characteristic points and associated matrices of projection. Moreover, as these key images are associated with 3D poses expressed in a local frame, it is thus not an isolated image which is useful during navigation but the whole set of key images, in which one finds common characteristic points.
We propose to gather this whole of images of references and the reference frame associated in a *local map* containing the following structure as indicated by Table 1.

Local map with characteristic 3D points
Set of *I* landmarks made up of - Key Images - Projection Matrices associated with each key image - Characteristic 3D Points localized in the frame of the map with arcs specifying the images where these points are found.

Table 1: Landmarks for precise localization.

3. Map-Matched Continuous Localization

In order to collect and store natural landmarks, it is necessary to have estimates of the pose of the vehicle, even if these estimates are really imprecise.

In urban areas, GPS suffers from several drawbacks such as multi-tracks and masking: GPS signal is often blocked or reflected by high-rise buildings. Because GPS alone cannot localize continuously the vehicle, it is localized by fusing an odometer, a gyrometer, GPS data and map information. If the GPS satellites signal is blocked by buildings, for example, the evolution model provides a dead-reckoned (DR) estimate the drift of which can be corrected using the map information.

In order to simplify the matching process, we assume that the vehicle follows the pre-computed itinerary. Figure **4** gives an example of a pre-computed itinerary obtained thanks to a usual route calculation functionality of a GIS.

The map-matched localization is done by a pose tracking method based on Kalman filtering. The DR sensor are used in a prediction stage which is corrected by the GPS, if it is available and coherent (no multi-tracks) and the map.

Let us consider a car with front-wheel drive. The mobile frame is chosen with its origin *M* attached to the centre of the rear axle. The *x*-axis is aligned with the longitudinal axle of the car. The vehicle position is represented by (x_k, y_k), the Cartesian coordinates of *M* in a global frame (a projection of geographic data). The bearing angle is denoted θ_k.

If the road is perfectly planar and horizontal, and if the motion is locally circular, the evolution model can be expressed by:

$$\begin{cases} x_{v,k+1} = x_{v,k} + \Delta_k.\cos\left(\theta_{v,k} + \frac{w_k}{2}\right) \\ y_{v,k+1} = y_{v,k} + \Delta_k.\sin\left(\theta_{v,k} + \frac{w_k}{2}\right) \\ \theta_{v,k+1} = \theta_{v,k} + w_k \end{cases} \quad (1)$$

Where $X_{v,k}$ is the vehicle's state vector at the instant k, composed of (x_k, y_k, θ_k), $U_{v,k}$ the vector of the measured inputs consisting of (Δ_k, w_k), Δ_k and w_k being respectively the elementary distance covered by the rear wheels and the elementary rotation of the mobile frame. $\alpha_{v,k}$ is the process noise and γ_k represents the measurement error of the inputs. $\alpha_{v,k}$ and γ_k are assumed to be uncorrelated and zero mean noise.

The values of Δ_k and w_k are computed using the ABS measurements of the rear wheels and a fibre optic gyrometer.

The fusion of GPS, Map information and odometry is done by Kalman Filtering (EKF here but it could be UKF) which uses the prediction/update mechanism. In the prediction step, the car evolves using (2) and the covariance of the error is estimated.

When a GPS position is available, a correction of the predicted pose is performed. In urban areas, GPS suffers from multi-tracks and bad satellite constellation (urban canyoning). So, when a GPS position is available, it is necessary to verify its coherence. For that, the Normalized Innovation Squared (NIS) which has a chi-square distribution is used: a distance d_m is computed between the GPS observation and the state vector.

Let Y_v be the observation vector, μ_v the innovation vector.

$$Y_v = \begin{bmatrix} x_{GPS} \\ y_{GPS} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} . X_v + \beta_{GPS} = C_v . X_v + \beta_{GPS} \quad (2)$$

$$\mu_v = \begin{bmatrix} x_{GPS} - x \\ y_{GPS} - y \end{bmatrix} \quad (3)$$

$$d_m = \mu_v . (P_\mu)^{-1} . \mu_v^T \quad (4)$$

Where P_μ is the covariance of the innovation.

If the computed distance d_m is smaller than a threshold, for instance $\chi^2(0.05,3)$, then the GPS measurement is assumed to be reliable and a

correction of the predicted pose is performed. Otherwise, the dead-reckoned pose provided by the evolution model is kept. Please, note that the GPS noise is not stationary. The GPS measurement error can be estimated in real time by using the *NMEA* sentence *GST*. This information is provided by the TRIMBLE AgGPS132 GPS receiver that has been used in the experiments.
Before launching the acquisition system, an itinerary has been computed and we assume that the vehicle follows exactly this path. Each segment in the GIS map has an Identifier (ID). This ID will be used in the following stages to perform the geo-referencing of the images.
Let us consider the segment selection problem which consists in extracting from the GIS map the most likelihood segment using the predicted state vector. Let (x_{pred}, y_{pred}, θ_{pred}) be the predicted pose. The distance between (x_{pred}, y_{pred}) and some segments of the itinerary that are near to the last fused pose are computed. The segment that has the smallest distance to the predicted pose and whose driving direction corresponds to the heading of the vehicle is considered as the good one. As we will see in section 6, this simple matching strategy gives good results. The matched point is obtained by projecting the estimated position onto the selected segment.
The matched point can be used as a map observation (denoted Y_{vm}) in order to correct the drift the DR estimate if the GPS is unavailable:

$$Y_{vm} = \begin{bmatrix} x_{MAP} \\ y_{MAP} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} . X_v + \beta_{MAP} = \mathrm{C_v} . X_v + \beta_{MAP} \quad (5)$$

Another NIS coherence test is used to verify the map observation coherence before fusing it thanks to a Kalman filter correction stage.

4. Geo-Referencing Natural Landmarks

Let consider the use of natural landmarks detected by a camera as explained in section 2.
After a learning stage, the visual features are detected and stored in an additional layer of the GIS. Then, these features are extracted each time the vehicle navigates in the same area.
In the learning stage, dated images are acquired by cameras connected to an acquisition system, along with other sensors. This includes dead-reckoning sensors (odometers attached on the rear wheel and a KVH fiber optic gyrometer) and GPS data.
Our method map-matches all the images and, afterwards, selects the landmarks (key images).

4.1. Map-Matching the Images

In this stage, map-matched coordinates of the fused poses of the localization algorithm shown in section 3 are associated with the acquired images using timestamps in the same reference time coordinates.
While the timestamp of the image to match falls between the timestamps of two positions belonging to the same road (they have the same ID), the geo-referencing of the image is done by a linear interpolation.
If the two map-matched positions used to the geo-referencing have different ID, the nearest is kept and the image position is projected onto its segment.
This simple strategy is well adapted in practice since the sampling rate of the matched position is significantly high (100 Hz used in the experiments).
The question is now to estimate the covariance matrix of the image. It is well known that imprecision is less dynamic than the state to observe. Therefore, the covariance error matrix associated with each image is considered to be equal to the one of the nearest map-matched pose.

4.2. Regrouping the images in local maps

We proposed to regroup the images in local maps to facilitate their management. For this purpose, the use of ID of the roads is well adapted; two different roads have necessarily different IDs. Indeed, a road is a set of same ID segments with eventually a junction at its beginning or/and at its end.
Moreover, each road can be one-way or double direction (this information is contained in the road attributes).
For one-way roads, a unique local map is built. It contains all the images matched to it. For double direction roads, two maps are built, one for each direction: E2W (East To West) or W2E (West To East). Figure 2 shows double direction roads with images associated to each direction.

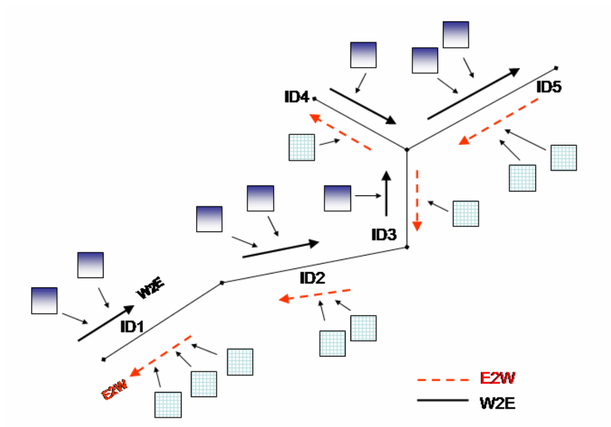

Figure 2 : Associating Images to Road Segments

4.3. Landmarks Characterization

Storing all the acquired images in the GIS database requires a huge amount of disk-space, which is unrealistic for real-time applications. So, chosen landmarks have to be stored. These landmarks are key-images associated with 3D points and connectivity links, as seen in section 2.

The characterization of the visual landmarks is done by post-processing the stored sequence. The main objective is to find enough landmarks for the precise localization. It depends mainly on the curvature of the trajectory. In a straight line, a key image can be characterized every 5 meters while it can be necessary to store them every meter, if the vehicle does a rotation.

In order to manage the landmarks, several attributes need to be added to the ones of Table 1. They are the GIS ID, the direction and, for each landmark, its map-matched position, its covariance and its date of acquisition (see Table 2).

Local map of index "j"
★ GIS ID
★ Direction (E2W or W2E)
★ Set of *I* landmarks made up of
- Key Images
- Projection Matrices associated with each key image
- Characteristic 3D Points localized in the frame of the map with arcs specifying the images where these points are found
- Matched position $Xmap_{j,I}$
- Covariance of this position $Pmap_{j,I}$
- Date of the acquisition

Table 2: Landmarks with extended attributes for their management in a GIS.

5. Landmark Extraction During Navigation

We consider now that the vehicle navigates in the environment learned in a previous stage. As did in the learning stage, a destination point is chosen and an itinerary is computed using the GIS. The vehicle follows this pre-computed itinerary.

In the navigation phase, the system needs high precision for localization. A centimetric precision is obtained when the vehicle can use the pertinent learned landmarks regarding the observed landmarks during navigation.

The topic of this paragraph concerns extraction and selection in the additional layer of GIS of pertinent landmarks at each time *t* of the navigation stage. The methods allowing the control the vehicle are described in [1],[6],[8]. The landmarks management for navigation consists in two parallel tasks:

- Local map extraction
- Visual landmarks selection

In what follows these different tasks are developed.

5.1. Find the right local map: map-matching fused poses

The goal of this task is to obtain a vehicle pose with a metric precision and then to find the appropriate local landmarks map stored in the GIS built landmark layer.

The localization algorithm is the same one used in the learning stage: it fuses GPS, road map information with dead reckoning sensors. A first pose is thus computed and a map-matching is done using the fused pose, then the road ID is retrieved.

The selected road ID and the heading permit to select the adequate local map. If the selected local map is a new one, the visual landmarks are then loaded in the dynamic system memory.

This supervisory task is repeated during the navigation process. It guaranties correct transitions between two local maps.

5.2. Visual landmarks extraction

The previous task provides a local map map_j composed by a set of geo-referenced visual landmarks as previously described. To obtain a precise localization in order, for instance, to control the vehicle, the navigation process needs to use the landmark I having the nearest matched position $X_{mapj,I}$. Two approaches are then possible.

The first method is classically based on the matching of features points (i.e. Harris coin detector) between the current image and all key-images of the local map. The matching of feature points consists in finding corresponding pairs. One example of matching algorithm is standard RANSAC-based method [7]. The selection of the pertinent landmarks depends on matching results. This method is well adapted for tracking process when previous correct localization has been done.

We propose to develop a hybrid method which consists in using when it is possible the current pose of the vehicle in order to select landmarks from the local map. The matched pose of the vehicle (on the GIS and with the driving direction) is compared with the pose of visual landmarks in the GIS. This comparison can be implemented using two methods, by using the landmarks I of the local map "map$_j$" which has been selected previously.

- Compute Euclidian distance between positions

$$D = \left\| \overrightarrow{X_{v,k} - X_{map_j, I}} \right\|$$

- Compute Mahalanobis distances tacking into account the inaccuracy of positions.

$$D_{Mh} = (X_{v,k} - X_{map_j,I})^T.(P_k + P_{map_j,I})^{-1}.(X_{v,k} - X_{map_j,I})$$

Where,

-$X_{v,k} = [x_{v,k}, y_{v,k}]^T$ is the vehicle position at the time k.

-$X_{mapj,I} = [x_{mapj,I}, y_{mapj,I}]^T$ is the absolute position of the landmark I of map$_j$.

-P_k is the covariance matrix of the vehicle position.

-$P_{mapj,I}$ is the covariance matrix of the landmark I.

The selection of pertinent landmarks can be obtained by applying a threshold. If several landmarks are selected, the RANSAC method has to be used to solve the ambiguity.

This second method combines matching poses and matching visual features. It is necessary to initialize the precise localization process during the navigation stage.

6. Experimental Results

Real experiments have been carried out with our experimental car (see Fig. 3) in the downtown area of Compiègne (see Fig. 4) using a KVH fibre optic gyro sampled at 100Hz, an odometer input and a Trimble AgGPS 132 running with a geostationary differential correction (Omnistar).

To test different technologies of cameras and various configurations, four cameras were connected to the acquisition system. Timestamped images were logged at the rate of 15 images per second. The used cameras were a CMOS fisheye one, a CMOS stereo pair (AVT Marlin F-131) and a CCD Sony Firewire (DFW-VL500).

Figure 3 : The Real time acquisition system

The GPS has been tuned in a 3-D only-mode to deliver reliable positions, by setting the threshold of DOP to a low value and the threshold of the SNR to a high value. Such tuning induces a reliable but intermittent positioning in urban areas.

Figure 5 shows the vehicle localization result (in bold) on the map (thin) in the urban environment. This localization results had been used to map-match the acquired key-images. The zoomed view shows key-images positions associated to road segments.

Two weeks later, we carried out navigation experiments. In these experimentations, a driver controlled the vehicle along a defined itinerary.

Figure 4: Overview of Road Map around the experimental field, with the itinerary plotted in bold

Figure 6 shows in bold green the localization result and in thin red the segments where visual local maps has been previously acquired.

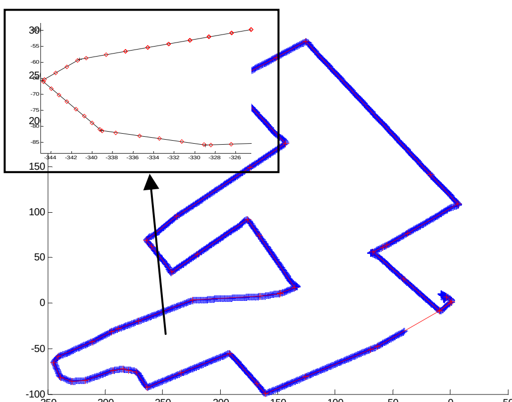

Figure 5: Map Matched results in learning stage

To illustrate the visual landmarks management, we have extracted at each time (in fact at each new image) the best landmarks for localization. For these experimentations, we have computed the Euclidian distance between the vehicle position and the landmarks position. Then, the *best* landmarks mean here the nearest geo-referenced key image.

A graphical interface allows to follows the way of the vehicle during the navigation process. The interface displays the map and adds two positions icons: the current matched position of vehicle is drawn with a blue/yellow arrow and the reference of the key-image selected by the algorithm is drawn with a red point.

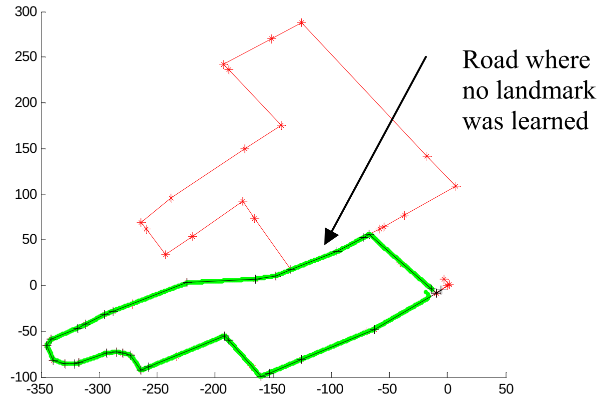

Figure 6: Map-matched vehicle position for the navigation stage

Two additional windows display the current image and the selected key-image. At each position, the interface refreshes the current position, the current image and the key-image, if it changes. Figure 7 displays this interface.

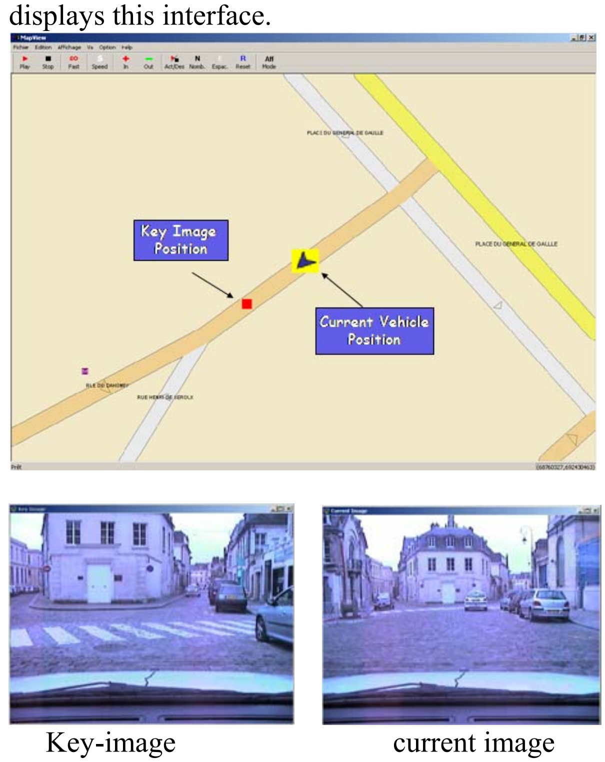

Figure 7: Real time graphical interface.

The time stamped data was saved in order to analyze the results in post-processing and verify the extracted landmarks regarding to the current image. The route was 1.8 km long during 6 mn. At each key image selection, the overlapping between key image and current image was verified. We have considered the selection correct if the overlapping covers the half area of images. The results give 98% overlapping when local map is available. The main errors come from the vehicle rotations at junctions. As said in paragraph 4.3, the key-image learning strategy can be refined at these points to assume visual landmark continuity. Another solution consists in using a special optic like fisheye camera.

7. Conclusion

This article has proposed a method to manage a huge amount of landmarks data in GIS for precise localization in urban areas. The example of visual landmarks with 3D points and connectivity links has been considered and the method has been illustrated with real experiments carried out in the downtown area of the city of Compiegne.

Our proposition is to regroup the landmarks in local maps defined by a map-matching process: a local map is a set of segments which has the same ID in the GIS database. In order to simplify the map-matching process, we use the route calculation of a GIS software to characterize an itinerary which has to be strictly followed by the vehicle during the learning and navigation stages.

Experimental results indicate that this approach is a good candidate to the management of landmarks in urban area since it allows extracting correctly landmarks stored in a previous passage. The weakness of this strategy appears when the curvature of the itinerary is significant, especially when two orthogonal roads make part of the route. In such a case, selecting only the most probable landmark can conduct to a wrong match. For that, the management of the imprecision of the landmarks is a crucial point. It is the main perspective of this research.

Acknowledgment

This research has been carried out in the framework of the project BODEGA supported by the French CNRS program ROBEA from Sept. 2003 to Sept 2005.

References

[1] S. Benhimane and E. Malis, “Real-time image-based tracking of planes using efficient second-order minimization”. IEEE/RSJ Int. Conf. on Intelligent Robots Systems, Sendai, Japan, Oct. 2004.

[2] A. Georgiev and P. K. Allen, “Localization Methods for a Mobile Robot in Urban Environments”, IEEE Trans. on Robotics and Automation, V. 20, N. 5, pp. 851-864, Oct. 2004.

[3] Ch. Harris and M. Stephens, “A Combined Corner and Edge Detector” Proceedings of The Fourth Alvey Vision Conference, , pp 147-151. Manchester, UK. 1988

[4] A. Howard, D. Wolf and G. Sukhatme, “Towards 3D Mapping in Large Urban Environments”, 2004 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems Sept. 28 – Oct. 2, 2004, Sendai, Japan

[5] Y. Kim, M. Pyeon, Y. Eo, “Development of hypermap database for ITS and GIS”, Computers, Environment and Urban Systems, Vol. 24, pp. 45-60, 2000.

[6] A. Remazeilles, F. Chaumette, P. Gros, “Robot motion control from a visual memory”. In IEEE Int. Conf. on Robotics and Automation, ICRA'04, Volume 4, Pages 4695-4700, New Orleans, Louisiane, April 2004.

[7] E. Royer, M. Lhuillier, M. Dhome and T. Chateau, “Towards an alternative GPS sensor in dense urban environment from visual memory”, 15th British Machine Vision Conference, Sept. 2004.

[8] E. Royer, J. Bom, M. Dhome, B. Thuillot, M. Lhuillier, F. Marmoiton, “Outdoor autonomous navigation using monocular vision”. IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 3395-3400, Edmonton, Canada, Aug. 2005.

[9] Th. Weiss, N. Kaempchen and K. Dietmayer, “Precise Ego-Localization in Urban Areas using Laserscanner and High Accuracy Feature Maps”, IEEE Intelligent Vehicle Symposium, Las Vegas, USA, June 6-8, 2005.