VISION-BASED MOBILE ROBOT LOCALIZATION WITH SIMPLE ARTIFICIAL LANDMARKS

Robert Bączyk, Andrzej Kasiński, Piotr Skrzypczyński

Poznań University of Technology,
Institute of Control and Information Engineering,
ul. Piotrowo 3A, PL-60-965 Poznań, Poland

Abstract: This article presents a vision-based method for mobile robot localization in partially known environments. Simple, unobtrusive artificial landmarks are used as external localization aids. The landmarks are labelled with unique code patterns, what enables unambiguous global localization. The analysis is presented which enables prediction of the uncertainty of the position and orientation of the robot relative to the landmark. Experimental evidence is provided, demonstrating that the proposed method can be reliably used over a wide range of relative positions between the camera and the landmark.

Keywords: Autonomous mobile robots, Positioning systems, Computer vision.

1. INTRODUCTION

To navigate successfully in a large-scale environment, mobile robot should know where it is within this environment, i.e. it must localize itself. The pose of a wheeled robot (position and orientation in the global frame $\mathbf{X}_R = [x_r\ y_r\ \theta_r]^T$) is usually estimated by means of odometry. This method alone is insufficient, and the pose has to be corrected using measurements from external sensors. A number of approaches to the sensor-based localization are known from the literature. Many successful robotic systems utilize landmarks, which are distinct features that a robot can recognize from sensory data (Feng *et al.*, 1996). Landmarks can be either natural or artificial. Natural landmarks are the features extracted from the sensory data without any changes being made in the environment. Artificial landmarks are the objects purposefully placed in the environment, such as visual patterns or reflecting tapes. In practice, recognition of natural landmarks by means of range sensors is limited to structured, office-like environments, which contain sufficient number of naturally occurring features (Skrzypczyński, 1998). The ability to detect natural landmarks by computer vision methods is limited because of problems related to changing illumination, occlusions, shadows, etc. However, the use of artificial landmarks as visual cues can improve operational characteristics of vision-based methods compared to the recognition of natural environment features (Yoon and Kweon, 2001). The disadvantage is, that the environment has to be engineered, what in turn limits the flexibility and adaptability to different operational sites. This problem can be avoided by using simple, cheap and unobtrusive landmarks, which can be easily attached to walls and furniture in virtually any environment.

This article presents a mobile robot localization system which uses passive visual landmarks to enhance the recognition capabilities of the onboard camera. The focus is on the evaluation of the spatial localization uncertainty. Both theoretical analysis and presentation of experimental results are given.

2. LANDMARK RECOGNITION

2.1 Landmark design

Landmark-based navigation relies on fast finding of landmarks, which are currently visible in the

field of view of the camera. This is done by segmentation of the available image. The segmentation criteria must be sufficiently reliable and independent on the geometrical deformation of the landmark appearance, due to the random position of the camera w.r.t. robot surrounding. On the other hand, precise determination of the landmark reference points on the image is critical for obtaining accurate enough coordinates, determining the pose of the camera fixed to the mobile robot.

As there is a number of perceptually similar landmarks placed in the workspace of the robot, in order to determine robot position in unambiguous way, each landmark must carry its proper code. This is to enable the identification of particular landmarks, whose location in world-coordinates is a priori known. Mobile robot moves on the floor, so the following assumptions are justified: camera optical axis is parallel to the floor, center of the landmark is at its height (on contrary to (Kasiński and Bączyk, 2001)). As the camera moves in horizontal plane it is worth building the landmark patterns of vertical and horizontal features.

To meet the above requirements landmarks have been designed as paper sheets (210 × 297 mm) printed in appropriate way. Within strong black orthogonal frame, filled with fluorescent green, there is a chessboard placed, made of 9 fields. Particular fields of chessboard are black or green according to the unique code, proper to the given landmark. Central field is always black. In this way one can place up to 256 unique landmarks in the workspace of the robot. Primary detection of landmarks on the current image is based on the segmentation in color-space. Landmark frame gives a strong luminance gradient with respect to its green filling. After thresholding one gets a quadrangle pattern on the output image, whose properties are a secondary criterion for the detection of the landmark. Internal corners of the black frame are treated as reference points and are used for pose evaluation. The particular layout of landmark-code area does not influence the positioning accuracy because this area is separated from reference points.

2.2 Image processing

Robot vision system works on true-color images of 640x480 pixels. Landmark recognition process consists of two steps: (i) detection of ROIs (*Regions Of Interest*) on the available image, (ii) recognition of the landmark in particular ROI and determination of image-coordinates of their reference points.

After Gaussian smoothing of the RGB image, it is transformed into the HSI-color space, in order to enable the following segmentation procedure.

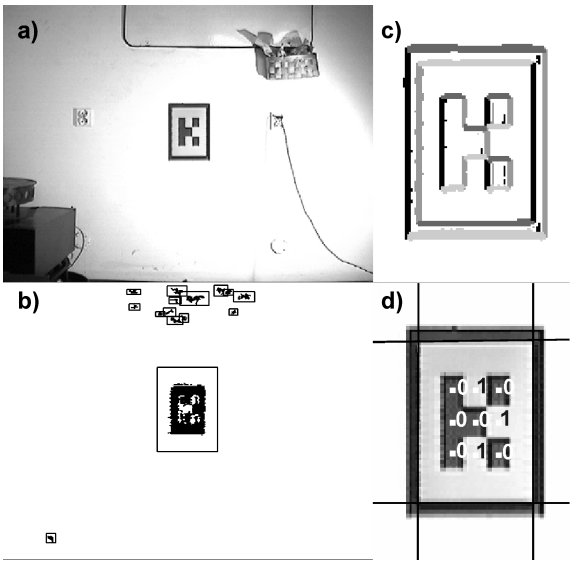

Fig. 1. Acquired image a), ROIs b), domains of edgels c), frame-corners determined by straight lines and landmark-code extracted d)

We are looking for pixels satisfying the predicate $I(x,y) > \vartheta_I \ \& \ H(x,y) > \vartheta_{Hmin} \ \& \ H(x,y) < \vartheta_{Hmax}$. The resulting binary image is subject to connectivity analysis. Each N_4-connected set of pixels is potentially an image of the landmark green-filling. Two additional conditions are checked: compactness of such a set and its cardinality. As a measure of compactness normalized, central, second-order moments are used. This test enables elimination of green artefacts. Moreover, the cardinality of the set representing green landmark must reach a given value, as maximal distance of the camera to the landmark is limited. In this way, problems with the texture having strong green component are eliminated. For each segment, satisfying compactness and cardinality criteria, ROIs are determined, having a form of rectangular windows (Fig. 1). Further processing steps are with respect to the G-component of the original RGB-image and their scope is restricted only to ROIs.

The internal edges of the landmark frame are searched. To that goal, gradient images are calculated, getting potential edgels. The information on the polarity of particular edgels is kept. After thresholding of the gradient image, one obtains domains of edgels. Some of them potentially support the hypothesis on the existence of the line-segments, making pattern of the landmark. Line-segments extraction is based on the following selection criteria: (i) the cardinality of each domain of edgels (small clusters are ignored); (ii) appropriate polarity of edgels (edgels making left or upper edges of the frame should have a positive polarity, negative - in the case of lower or right frame edge); (iii) sufficient slimness (measured with normalized, central, second-order moments of

binary domains); (iv) relative size (line-segments too short w.r.t. the ROI dimensions are ignored). The extracted segments are linked into quadrangle. The distances between ends of the segments should be smaller than given value. For all edgels supporting thus assembled landmark frame a precise location of gradient maxima is performed. Straight lines are fitted to those maxima. Their crossections determine with subpixel accuracy the picture-location of potential frame-corners. Having determined the image coordinates of the landmark frame corners, one can establish centers of chessboard fields. It is checked whether its center is black, if so (hypothesis verification) the landmark code determination follows (values of pixels in the centers of particular chessboard fields are read). All values of thresholds in the described procedure have been experimentally established.

2.3 Robot pose computation

Pin-hole camera model is used. Camera is internally calibrated (its focal length is exactly known) (Heikkilä, 2000). The image-coordinates frame is fixed in center of the image. Image coordinates are expressed in [mm] by taking into account the physical dimensions and vertical/horizontal resolution of CCD-matrix of the camera. X-coordinates (x_1[mm] and x_2[mm]) of the centers of left and right frame edge are determined. Half the lengths (y_1[mm] and y_2[mm]) of the vertical left and right landmark frame-edge are calculated. Those data, obtained from the image, are used to evaluate vector $\mathbf{L} = [l_2\ \varphi_1\ \varphi_2]^T$ determining robot (camera) pose on the plane (Fig. 2).

$$l_2 = vl \frac{\sqrt{\lambda^2 + \left(\frac{x_1 y_2 + x_2 y_1}{y_1 + y_2}\right)^2}}{\frac{2 y_1 y_2}{y_1 + y_2}} \tag{1}$$

where vl is the half-width of the landmark frame,

$$\varphi_1 = -\arctan\left(\frac{\frac{x_1 y_2 + x_2 y_1}{y_1 + y_2}}{\lambda}\right) \tag{2}$$

$$\varphi_2 = \arctan\left(\lambda \frac{y_1 - y_2}{-x_1 y_2 + x_2 y_1}\right) \tag{3}$$

3. UNCERTAINTY EVALUATION

3.1 Problem statement

This section describes how inaccuracies in the recognition of a landmark pattern in the camera image propagate to the uncertainty of the robot pose. The uncertainty analysis uses first order covariance propagation (Smith and Cheeseman, 1987). The robot pose is a random variable, being a subject to an additive, zero-mean Gaussian noise. The pose uncertainty is described by a covariance matrix:

$$\mathbf{C}_R = \begin{bmatrix} \sigma_x & \sigma_{yx} & \sigma_{\theta x} \\ \sigma_{xy} & \sigma_y & \sigma_{\theta y} \\ \sigma_{x\theta} & \sigma_{y\theta} & \sigma_\theta \end{bmatrix}. \tag{4}$$

There are many sources of the uncertainty in the vision-based recognition of objects and patterns (Jain *et al.*, 1995). In this study, we focus on the evaluation of the sensitivity of the obtained pose estimate to the relative position between the robot and the observed landmark. The optical center of the camera is assumed to be in the center of the robot. Moreover, it is assumed that the landmark coordinates in the global frame are certain (this assumption is justified for artificial landmarks). Errors due to signal interference and other electronic noises in the image are not taken into account in the uncertainty prediction procedure, as they do not depend on the spatial configuration of the camera w.r.t. the landmark. All other sources of noise are assumed to be Gaussian. The input data for the computation of the predicted uncertainty are: l_2, φ_1, and φ_2 (Fig. 2). These three values are predicted on the basis of the current robot pose estimate and the knowledge of the position and orientation of the chosen landmark. The same triple $\mathbf{L} = [l_2\ \varphi_1\ \varphi_2]^T$ is then computed from the actual image of the landmark pattern, after taking an image. The prediction procedure enables: (i) computation of the covariance matrix $\mathbf{C}_R$, which is necessary if the robot computes the pose estimate by means of the multi-sensor fusion (Crowley, 1996); (ii) rough estimation of the robot pose uncertainty *before* an image has been taken and processed, what in turn enables the robot to decide which landmark should be used for the localization, and how to plan a path/movement to minimize pose uncertainty (Madsen *et al.*, 1997); (iii) theoretical evaluation of the uncertainty as a function of the robot pose w.r.t. the landmark, in order to find the best and worst configurations, to be preferred or avoided by the robot, respectively.

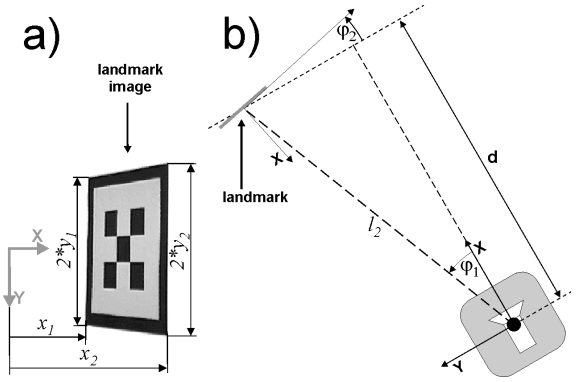

Fig. 2. Image coordinates a), geometric relations between the landmark and the robot b).

3.2 Location uncertainty of the landmark points

The distances and angles between the camera and the landmark are computed from the relations between the known dimensions of the landmark pattern, and the dimensions of the image of this pattern appearing on the CCD matrix. The camera is a tool to measure dimensions of the image. The resolution of measurement is bounded by the CCD matrix pixel size. Although positions of the landmark reference points are computed with subpixel resolution, the standard deviation of this measurement equals the size of a pixel. This is the primary uncertainty, introduced by the limited resolution of the camera, and it is then propagated to the uncertainty of the **L** parameters, and then to the uncertainty of the robot pose $\mathbf{X}_{R_{land}}$. It has been assumed, that the primary uncertainty depends on the errors in computation of the coordinates $[x_1\ x_2]$ and lengths $[y_1\ y_2]$ in the image:

$$\left[\,\Delta x_1\ \Delta y_1\ \Delta x_2\ \Delta y_2\,\right]. \tag{5}$$

Thus, the primary uncertainty is represented by a 4×4 matrix :

$$\mathbf{C}_P = \begin{bmatrix} \sigma_{x1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{y1}^2 & 0 & 0 \\ 0 & 0 & \sigma_{x2}^2 & 0 \\ 0 & 0 & 0 & \sigma_{y2}^2 \end{bmatrix}. \tag{6}$$

The standard deviation σ_y in the y coordinate is defined as the vertical discretization error:

$$\sigma_y = \frac{h_m}{R_h}, \tag{7}$$

where R_h is the vertical resolution of the CCD matrix, and h_m is it's height (here R_h=480, h_m=3.6mm). Similarly, the standard deviation σ_x in the x coordinate is defined as:

$$\sigma_x = \frac{w_m}{R_w}, \tag{8}$$

where R_w is the horizontal resolution of the CCD matrix, and w_m is it's width (here R_w=640, w_m=4.8mm). For the case under study, σ_x=σ_y, and thus all the primary uncertainties are equivalent.

3.3 Uncertainty of the distances and angles to the landmark

The uncertainty of **L** is described by the covariance matrix $\mathbf{C}_L$:

$$\mathbf{C}_L = \begin{bmatrix} \sigma_{l_2} & \sigma_{\varphi_1 l_2} & \sigma_{\varphi_2 l_2} \\ \sigma_{l_2 \varphi_1} & \sigma_{\varphi_1} & \sigma_{\varphi_2 \varphi_1} \\ \sigma_{l_2 \varphi_2} & \sigma_{\varphi_1 \varphi_2} & \sigma_{\varphi_2} \end{bmatrix}. \tag{9}$$

This matrix is computed from the primary uncertainty matrix $\mathbf{C}_P$ of the vector $\mathbf{P} = [x_1\ y_1\ x_2\ y_2]^T$. The parameters of this vector constitute the input data for the procedures calculating **L**. As a transformation from **P** to **L** the equations (1), (2), (3) are used. They contain only the parameters of **P** and constants being the parameters of the landmark and the camera. Because the transformation between **P** and **L** is nonlinear, the covariance matrix $\mathbf{C}_L$ is computed as a first order approximation:

$$\mathbf{C}_L = \mathbf{J}_P \mathbf{C}_P \mathbf{J}_P^T, \tag{10}$$

where $\mathbf{J}_P$ is the Jacobian of this transformation w.r.t. **P**:

$$\mathbf{J}_P = \begin{bmatrix} \frac{\partial l_2}{\partial x_1} & \frac{\partial l_2}{\partial y_1} & \frac{\partial l_2}{\partial x_2} & \frac{\partial l_2}{\partial y_2} \\ \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial y_1} & \frac{\partial \varphi_1}{\partial x_2} & \frac{\partial \varphi_1}{\partial y_2} \\ \frac{\partial \varphi_2}{\partial x_1} & \frac{\partial \varphi_2}{\partial y_1} & \frac{\partial \varphi_2}{\partial x_2} & \frac{\partial \varphi_2}{\partial y_2} \end{bmatrix}. \tag{11}$$

3.4 Robot pose uncertainty in the global frame

The last step in the pose uncertainty evaluation is to show how the uncertain distance and angles to the landmark influence the position and orientation of the robot in the global frame. For notational simplicity, it is assumed here that the observed landmark is located in the origin of the global coordinate system. Position of the camera (and thus the robot) in the global frame is described by the vector $\mathbf{X}_R = [x_r\ y_r\ \theta_r]^T$, and computed from the formula:

$$x_r = l_2 * \cos(\varphi_2 - \varphi_1), \tag{12}$$

$$y_r = l_2 * \sin(\varphi_1 - \varphi_2), \tag{13}$$

$$\theta_r = -\varphi_2. \tag{14}$$

The uncertainty of $\mathbf{X}_R$ is described by the covariance matrix $\mathbf{C}_{R_{land}}$ (4), which is a result of the uncertainty propagation from the vector **L** described by $\mathbf{C}_L$. Because the relation between **L** and $\mathbf{X}_R$ described by (12), (13), (14) is nonlinear, the covariance matrix $\mathbf{C}_{R_{land}}$ is computed from a first order approximation:

$$\mathbf{C}_{R_{land}} = \mathbf{J}_L \mathbf{C}_L \mathbf{J}_L^T, \tag{15}$$

where $\mathbf{J}_L$ is the Jacobian of (12), (13), (14) w.r.t. **L** :

$$\mathbf{J}_L = \begin{bmatrix} \frac{\partial x_r}{\partial l_2} & \frac{\partial x_r}{\partial \varphi_1} & \frac{\partial x_r}{\partial \varphi_2} \\ \frac{\partial y_r}{\partial l_2} & \frac{\partial y_r}{\partial \varphi_1} & \frac{\partial y_r}{\partial \varphi_2} \\ \frac{\partial \theta_r}{\partial l_2} & \frac{\partial \theta_r}{\partial \varphi_1} & \frac{\partial \theta_r}{\partial \varphi_2} \end{bmatrix}. \tag{16}$$

Beside the monocular vision system, the robot uses odometry, and updates it's pose from time

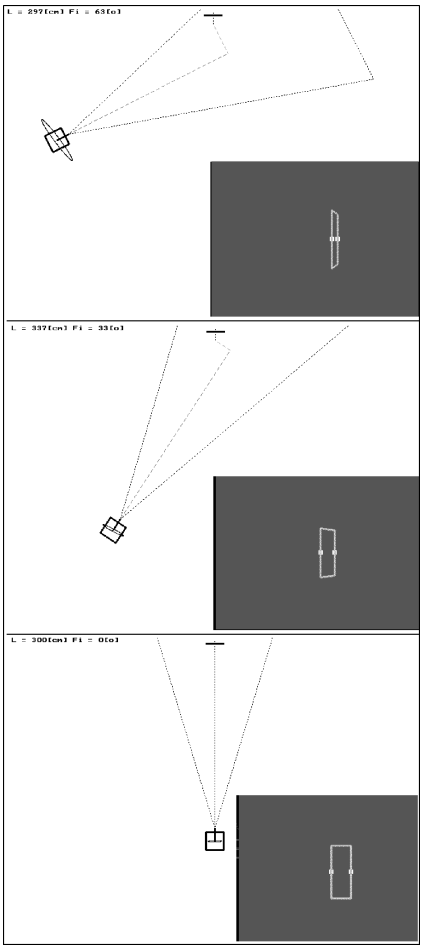

Fig. 3. Influence of the robot position w.r.t. the landmark on the position uncertainty.

to time if a landmark is available. The robot uses a Kalman filter to fuse both pose estimates. In this fusion procedure, pose vectors are weighted by their uncertainties, $\mathbf{C}_{R_{land}}$ and $\mathbf{C}_{R_{odo}}$, respectively, which have been obtained from the above described procedure and from the model of the robot odometry (Crowley, 1996).

The graphical representation of the positional uncertainty of the robot is an equiprobability ellipse (Smith and Cheeseman, 1987). In Fig. 3 the predicted uncertainty ellipses in three different relative positions of the robot are shown, for the 95% probability threshold. The landmark is shown as a black mark in the top of the images. The grey areas on the images contain the predicted shape of the landmark frame, as it is seen by the virtual robot camera. The small white areas on those figures indicate primary uncertainty.

4. EXPERIMENTS AND RESULTS

A series of experiments has been undertaken to evaluate the accuracy of the presented localization method in practice. The robot was equipped with a *Panasonic WV-CP230/GE* colour camera,

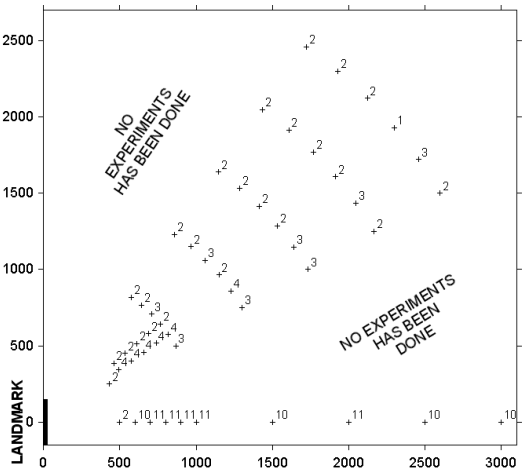

Fig. 4. Camera configurations examined in the experiment.

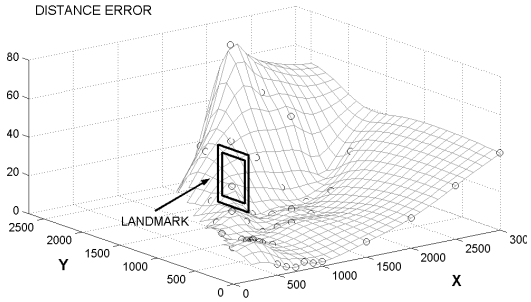

Fig. 5. Mean errors in l_2 (in mm).

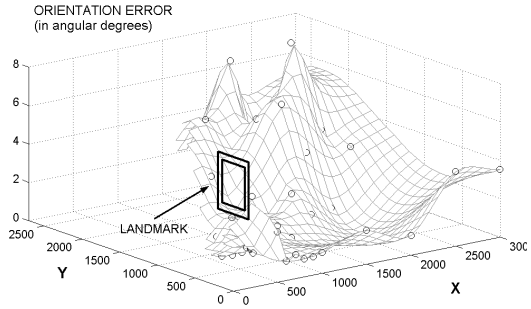

Fig. 6. Mean errors in φ_2 (in deg.).

connected to a low-cost *AverMedia* BT8x8 chip-based frame grabber. The lenses used had the focal length λ=6mm.

The aim of the experiments was to obtain a map of the positioning accuracy as a function of the relative spatial configuration between the robot and the landmark. Figure 4 illustrates the spatial layout of the points, in which the positioning accuracy has been surveyed, and the number of experiments performed for the particular point. Initial experiments have shown that the error map is symmetrical, thus only a half of the area has been surveyed thoroughly. The experiments have been performed in varying lighting conditions.

The results from the localization program have been compared against the ground truth obtained

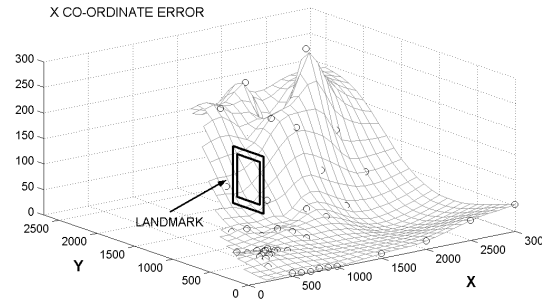

Fig. 7. Mean pose errors in x (in mm).

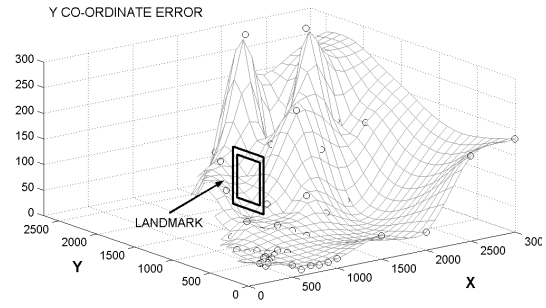

Fig. 8. Mean pose errors in y (in mm).

by measurements performed with an ordinary meter and a coarse grid (cell size 0.5m) painted on the floor. Figures 5 and 6 show the maps of average values of errors for the parameters computed by the localization procedure, i.e. for l_2 and φ_2. Figures 7 and 8 show the errors in the robot position, along the x and y coordinate, respectively.

The experiments have demonstrated, that the positioning errors depend on the distance between the camera and the landmark, and on the angle between the camera optical axis and the landmark plane. According to the experimental results, the robot can made the best use of a landmark which is in front of it, and within a distance of about 2 meters. In such a case the positioning accuracy of 5 cm is achievable in both x and y coordinates, with an orientation error smaller than 2^o. The obtained error maps show some outliers, what can be attributed to errors in the surveyed ground truth and/or to the large image distortions, which have been occasionally observed in the pictures taken with the *AverMedia* frame-grabber. A large number of experiments has to be undertaken in order to explain and eliminate such outliers by statistical methods.

5. CONCLUSIONS

This article presents a novel method for recognition of passive visual landmarks. Owing to this method, a robot can use it's on-board camera for global localization, achieving both high accuracy and robustness against problems typical to the passive vision-based methods. Although the presented landmark recognition method is quite complicated, the experiments have shown that it's localization accuracy outperforms methods based on simple processing of grey-level images (Hallmann and Siemiątkowska, 2001).

Moreover, a theoretical analysis of the image processing error propagation to robot pose has been given. The role of this theoretical model is twofold: it enables computation of the covariance matrix from the actual measurements, and permits prediction of robot pose uncertainty before image processing takes place. The localization accuracy has been evaluated experimentally, confirming the presented error propagation model.

REFERENCES

Crowley J. L. (1996). Mathematical Foundations of Navigation and Perception for an Autonomous Mobile Robot, In: *Reasoning with Uncertainty in Robotics* (L. Dorst, ed.), Springer-Verlag.

Feng, L., Borenstein, J. and Everett, H. (1996). *"Where am I ?" Sensors and Methods for Autonomous Mobile Robot Positioning*, Technical Report, University of Michigan.

Hallmann I. and Siemiątkowska B. (2001). Artificial Landmark Navigation System, *Proc. Int. Symp. Intelligent Robotic Systems*, Toulouse, 219–228.

Heikkilä J. (2000). Geometric Camera Calibration Using Circular Control Points, *IEEE Trans. on Pattern Analysis and Machine Intell.*, **22(10)**.

Jain, R. Kasturi, R. and Schunck B. (1995). *Machine Vision*. McGraw-Hill, New York.

Kasiński A. and Bączyk R. (2001). Robust Landmark Recognition with Application to Navigation, *Proc. Conf. Computer Recognition Systems (KOSYR)*, Wroclaw, 401–407.

Madsen, C., Andersen, C. and Sørensen, J. (1997). A Robustness Analysis of Triangulation-Based Robot Self- Positioning, *Proc. Int. Symp. Intelligent Robotic Systems*, Stockholm, 195–204.

Skrzypczyński P. (1998). Localization of a Mobile Robot Based on Natural Landmarks, *Proc. IFAC Symp. on Intelligent Autonomous Vehicles*, Madrid, 615–620.

Smith R. and Cheeseman P. (1987). On the Estimation and Representation of Spatial Uncertainty, *Int. Journal of Robotics Research*, **5(4)**, 56–68.

Yoon K-J. and Kweon I-S. (2001). Artificial Landmark Tracking Based on the Color Histogram, *Proc. IEEE/RSJ Conf. on Intelligent Robots and Systems*, Maui, 1918–1923.