Recognition of landmarks distorted by fish-eye lens

Barbara Zitová* Homero V. Rios† Juan Manuel Gutierrez† Antonio Marín‡

Abstract

The paper deals with recognition of artificial landmarks using fish-eye lens. The proposed recognition model is based on the algebraic affine moment invariants (AMIs). Landmarks are used for the navigation of autonomous robot, equipped with a fish-eye lens camera. The recognition ability of the AMIs regarding the deformation introduced by the fish-eye lens image acquisition is investigated. The results of the experiments in real situations, which proved the discriminability and the stability of the recognition model, are shown.

Keywords: Landmark recognition; Affine moment invariants; Fish-eye lens; Robot navigation

1. Introduction

The use of robots has an increasing tendency nowadays. Some tasks require mobile robots to act autonomously, without human help. They must have correct notion about their position and what should be their next action. Moravec's Cart (Moravec, 1981) represents one of the first solutions of the robot navigation.

The information about the robot trajectory and the previous representation of the robot's world can be used for the estimation of the current robot position. The robot can construct its own world map based on the data acquired by its sensors (Leonard, Durrant-Whyte and Cox, 1990; Brooks, 1986; Shah and Aggrawal, 1995a). After an incorporation of the robot motion trajectory and the sensor data into the evolving world representation, the robot is able to make a decision about its current position and the next action to take. This approach is efficient in situations when the robot working space is complicated or is not static.

The map of the robot world can be assumed as known a priori in some tasks (Cox and Wilfong, 1990). Representation of the robot's working space is stored in its memory. Hereupon, the robot makes the decision about its current position and next action by matching the new

*Department of Image Processing, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Praha 8, Czech Republic. E-mail: zitova@utia.cas.cz

†Laboratorio Nacional de Informática Avanzada, A. C., Rébsamen 80, C.P. 91090, Xalapa, Veracruz, México. E-mail: hrios@xalapa.lania.mx

‡Facultad de Física e Inteligencia Artificial, Universidad Veracruzana, Xalapa, Veracruz, México. E-mail: amarin@mia.uv.mx

collected sensor data with the stored reference world representation (Kosaka and Kak, 1992; Dulimarta and Jain, 1997).

Landmarks represent one possible solution of the robot position estimation. Objects in the scene, which are found distinctive by the robot are called landmarks. In this approach no representation of the surrounding world is necessary, so complex and often memory demanding descriptions can be avoided. Landmarks can be used to impart certain type of information to the robot (Lewitt, Lawton, Cheldberg and Nelson, 1987). For example, recovery from failure during robot navigation can be based on finding landmarks (Kosaka and Kak, 1992). The landmarks can be either part of the world (Yeh and Kriegman, 1995) or artificial signs placed into the robot's environment (Kortenkamp et al., 1993). They are detected by visual sensors and then compared against the database of landmarks.

A robust recognition system for landmark classification is essential. The appropriate recognition model is chosen according to the type of degradation, which is going to be present in acquired images. For example, the robot equipped with the common type of camera introduces to the landmark acquisition process the projective deformation and additive noise. The recognition model based on affine moment invariants AMIs (Zitova and Flusser, 1999) was proven to be robust under such type of deformation.

In this paper, we present a system for autonomous robot navigation. The robot is equipped with fish-eye lens camera, which provides more information than the conventional type of lenses. Especially, the robot navigation in indoor environments can be improved using the fish-eye lenses because of their wide field of view (approximately 180^o in the diagonal direction) and the ability to obtain information even from very close robot surrounding (Shah and Aggrawal, 1995b). We propose to apply a recognition model based on affine moment invariants (AMIs) (Flusser and Suk, 1993; Flusser and Suk 1994). To prove the applicability of this idea we investigated the stability and robustness of AMIs on the images acquired by a fish-eye lens camera.

Several aspects of the autonomous mobile robot navigation, use of navigation marks and fish-eye lens camera use are mentioned in Section 2. The recognition model based on affine moment invariants (AMIs) is introduced in Section 3. Section 4 deals with experiments we made to test the recognition ability and robustness of the proposed model. Section 5 concludes the described propositions and experiments.

2. Aspects of the mobile robot navigation

As mentioned above, one approach to navigate a mobile robot is to use landmarks. They can be part of the environment, for example sets of vertical edges of objects located in the robot's surrounding. However, it is often much easier to use artificial marks placed in the environment. They are considered as a priori knowledge of the world.

Information which should be imparted to the robot in a particular situation (about robot's position, crossing, obstacle, special task), is bound with certain shape of the mark. Distinctively shaped marks then should be situated in the area of the possible occurrence of corresponding situations. If artificial marks are used, there are no more limitations, produced by the natural layout of objects. Use of natural marks can be inconvenient due to the possible complexity of mark description and their detection in some tasks. Artificial marks can be used in many situations without problems and we can even assume that they are not occluded with an obstacle (we can choose their positions). We confined ourselves to the model of robot navigation problem solution with artificial non-occluded marks in this article.

To obtain good results using the landmark approach, marks should fulfill several conditions. They should be distinctive, meaning that they should be discriminable, and finding them should not be too difficult.

The way of using imparted information is an important aspect of artificial landmarks. Marks can be used directly for robot self-localization in the world map, where the position of each landmark is recorded. On the other hand, the information can be encoded directly into the shape, for example the numbers of vertical and horizontal lines on the mark can determine the x and y coordinates of the robot position, respectively. It is necessary to design mark shapes correctly for a certain task, when using this approach. Their shapes have to offer enough free parameters for encoding desired data (if x and y coordinates of a position has to be encoded, the mark shapes should be able to vary in two parameters).

It is often more suitable to use information, encoded in mark shapes, as a key for the look-up tables. Then we can change data in the look-up tables in case robot's task is changed. Our

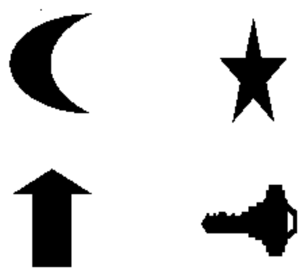

Figure 1: The database of used landmark shapes [MOON, STAR, ARROW, KEY].

mark set consists of four shapes which have different properties of curvature, number of sides, holes, etc., which makes them easy to differentiate visually (see Fig. 1). We could have more shapes in the mark set to provide information to the robot, or instead, we could change their aspect ratio (Zitova and Flusser, 1999), or encode data in the shape itself through the use of bar codes (Becker et al., 1995; Lopez et al., 1999).

3. Recognition model - Affine moment invariants (AMIs)

The recognition model based on invariants was chosen for marks recognition. Invariants describe features of an object, which stay unchanged even in situations, when an object is distorted. The invariants should be chosen according to the type of deformation which is introduced during the image acquisition process. Several kinds of features have been used for object description in recent works, such as shape vectors (Peli, 1981), shape matrices (Goshtasby, 1985), Fourier descriptors (Zahn and Roskies, 1972), differential invariants (Weiss, 1988) and moment invariants

(Hu, 1962; Belkasim, Shridhar and Ahmadi, 1991; Prokop and Reeves, 1992). Most of them can describe binary objects only and, moreover, are invariant only under translation, rotation and scaling of the object. Recently, Flusser and Suk (Flusser and Suk, 1993) have derived *affine moment invariants* (AMIs), which are invariant under *general affine transformations*

$$\begin{aligned} u &= a_0 + a_1 x + a_2 y, \\ v &= b_0 + b_1 x + b_2 y, \end{aligned} \tag{1}$$

((x, y) and (u, v) are the coordinates in the image plane before and after the transformation, respectively). Invariance of the AMIs has been proven theoretically and experimentally in recent papers (see (Flusser and Suk, 1993; Flusser and Suk, 1994)). They are robust enough in the case of additive zero-mean random noise. The AMIs are able to correctly recognize even objects, underlying the *perspective projection*

$$\begin{aligned} u &= (a_0 + a_1 x + a_2 y) \,/\, (1 + c_1 x + c_2 y)\,, \\ v &= (b_0 + b_1 x + b_2 y) \,/\, (1 + c_1 x + c_2 y)\,. \end{aligned} \tag{2}$$

A perspective transformation exactly describes the projection of a 3-D object of a general position into the 2-D image plane when captured by a pinhole camera. When the distance between the camera and the object is significantly larger than the size of the object, the projective transform can be well approximated by the affine transform and the AMIs can provide a correct classification of the objects. Descriptions of experiments proving this can be found in (Zitova and Fluser, 1999).

In our case, the deformation is not just perspective one but includes another type of deformation, introduced by the fish-eye lens (see Fig. 3, 4). Under specific conditions (the size of an object, the distance between the object and the camera) the AMIs are robust under the perspective transform. We assume that under similar conditions the AMIs recognition model can be stable under the fish-eye lens deformation too. This robustness of AMIs is tested experimentally in Section 4, where the experiments we made with the landmark database and the proposed recognition model are described.

The first six AMIs follow in explicit form:

$$\begin{aligned}
I_1 &= \frac{1}{\mu_{00}^4}(\mu_{20}\mu_{02} - \mu_{11}^2), \\
I_2 &= \frac{1}{\mu_{00}^{10}}(\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^2\mu_{12}^2), \\
I_3 &= \frac{1}{\mu_{00}^7}(\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)), \\
I_4 &= \frac{1}{\mu_{00}^{11}}(\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 + 12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} + \\
&\quad + 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} - 8\mu_{11}^3\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + \\
&\quad + 9\mu_{20}\mu_{02}^2\mu_{21}^2 + 12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2), \\
I_5 &= \frac{1}{\mu_{00}^6}(\mu_{40}\mu_{04} - 4\mu_{31}\mu_{13} + 3\mu_{22}^2),
\end{aligned}$$

$$I_6 = \frac{1}{\mu_{00}^9}(\mu_{40}\mu_{04}\mu_{22} + 2\mu_{31}\mu_{22}\mu_{13} - \mu_{40}\mu_{13}^2 - \mu_{04}\mu_{31}^2 - \mu_{22}^3),$$

where μ_{pq} is the central moment of order $(p+q)$. For a 2-D object A it is defined as

$$\mu_{pq} = \int\int_A (x - x_t)^p (y - y_t)^q f(x,y)dxdy, \tag{3}$$

where (x_t, y_t) are the coordinates of the center of gravity of object A and $f(x,y)$ describes an intensity distribution within A. Full derivation of the AMIs can be found in (Flusser and Suk, 1993).

Some general remarks about AMIs should be made here. It can be said that moments of higher orders describe more subtle variations in shape and are more sensitive to noise corruption. The second important thing about them is that AMIs I_2, I_3 and I_4 (mentioned above) have theoretically zero-values in the case of radially symmetric objects. This is caused by the fact that the value of μ_{pq} is zero, when p or q are odd. During experiments, described in Section 4, the AMIs I_2, I_3 and I_4 appeared to be less robust to the fish-eye lens deformation, present in our experiments. They were omitted from the recognition model for this reason.

Generally, the moments and the AMIs are defined for gray-level images with a finite support without any other restrictions. In our case, however, the images of the landmarks are binarized before calculating the invariants in order to eliminate different lighting conditions. Thus, $f(x,y)$ in definition (3) represents the characteristic function of the landmark picture.

We propose a recognition model, which for a given image of a landmark, acquired by the robot sensors, computes the AMIs I_1, I_5 and I_6, and by the minimum distance algorithm finds the closest mark from the database. The ability of this recognition model to make correct decisions for the fish-eye lens acquired images of landmarks is investigated experimentally in Section 4.

4. Experiments

The discriminability of the AMIs was tested on images acquired by the fish-eye lens camera with lens focal length 3.8mm and specification F 1.8, W/IRIS. We use four different shapes of landmarks (see Fig. 1). Images were taken from a 90 cm distance and landmarks were situated as it is shown on Fig. 2. For every shape we acquired 10 images of the landmark at specified positions (Fig. 3 and 4). On the images the deformation caused by the fish-eye lens is visible. For better appreciation we took an image of the floor with regular grid pattern. The image was taken form the same camera distance. The introduced deformation is very well visible on the image (Fig. 5).

The landmarks were segmented from acquired images using a boundary following technique (detected boundaries are shown in Fig. 3 and 4) and a thresholding segmentation. Both approaches are implemented in a semi-automatic way. We are going to make the segmentation fully automatic for the mobile robot navigation project for the future. We just work with the binary images of detected landmarks after the landmark segmentation was done.

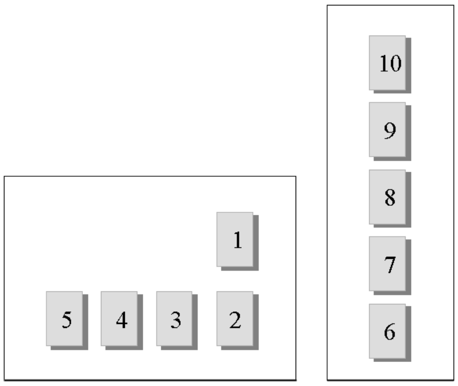

Figure 2: Landmark positions for experiments and position indexing.

AMIs were computed from the segmented landmarks for all 10 positions for each type of landmark shape and for all types of segmentation techniques. The AMIs I_2, I_3 and I_4 (mentioned above) appeared to be very sensitive to the present type of deformation, but the AMIs I_1, I_5 and I_6 confirmed our expectation that AMIs are robust even under the projective and fish-eye lens transformation when the landmarks are not too far away from the camera (in our case of the mobile robot navigation, just the near surrounding information will be used for the planning of the robot movement). The values of AMIs I_1, I_5 and I_6 for all images are shown in Tab. 1. The AMIs I_1, I_5 and I_6 were chosen to form the feature vector for the landmark classification.

	MOON			STAR			ARROW			KEY		
	I_1	I_5	I_6	I_1	I_5	I_6	I_1	I_5	I_6	I_1	I_5	I_6
1	1430	2333	1170	1173	1174	765	776	437	145	1283	1356	854
2	1499	2594	1341	1165	1152	751	785	452	155	1276	1347	830
3	1467	2479	1259	1165	1161	734	770	424	139	1282	1335	844
4	1445	2408	1208	1165	1161	734	781	439	147	1285	1376	846
5	1508	2635	1376	1162	1152	724	780	437	145	1292	1528	1024
6	1430	2243	1130	1170	1174	760	778	438	143	1290	1344	892
7	1412	2162	1052	1165	1153	723	778	442	141	1290	1344	892
8	1426	2230	1087	1169	1160	731	783	452	141	1280	1342	870
9	1394	2072	984	1214	1210	784	784	447	136	1285	1365	932
10	1432	2199	1077	1204	1169	761	779	430	133	1292	1528	1024

Table 1: Values of AMIs I_1, I_5 and I_6, computed from segmented images (10 different images taken at specified positions for every type of landmark shape [MOON, STAR, ARROW, KEY]).

The stability of the AMIs I_1, I_5 and I_6 can be seen in Fig. 6, where the relative variations RV of AMIs are plotted $RV(image) = abs(AMI(image) - AMI(reference))/(AMI(reference)/100))$. The reference image is the position No. 2 from every set [MOON, STAR, ARROW, KEY]. The error is very small according to the level of degradation (see Fig. 3, 4).

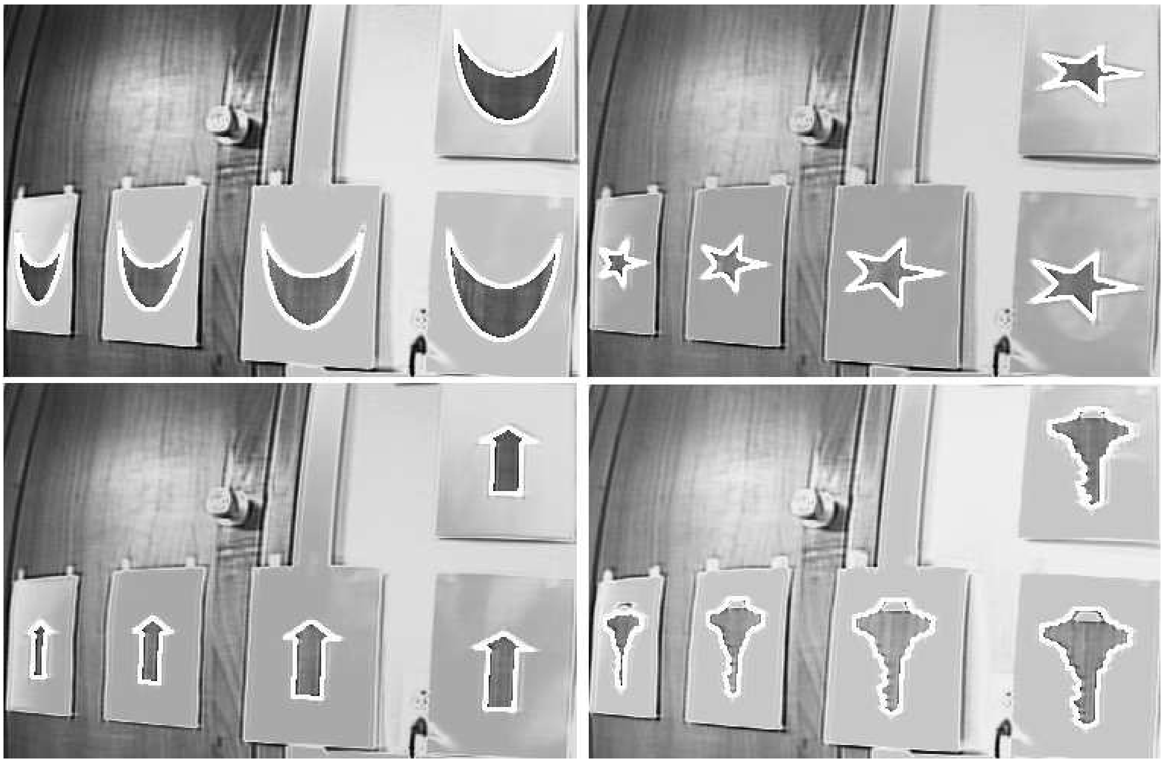

Figure 3: The first five landmark images from every set [MOON, STAR, ARROW, KEY]; the detected landmark boundaries are shown.

The influence of the chosen technique for segmentation was investigated. The maximum relative variation RV of AMIs can vary with the type of the segmentation. It is important to pay attention to the segmentation of landmarks. In case of following chosen segmentation techniques the classification rate was good, as can be seen in the next part of this section.

In Fig. 7 the feature space (I_1, I_5, I_6) is plotted together with the feature vectors, corresponding to the segmented landmarks. The MOON and ARROW clusters are well separated in the feature space, on the other hand the KEY and STAR clusters are close to each other in the feature space. A misclassification can appear in the case of wrong segmentation. The Fig. 8 shows distances of all acquired landmarks from the representatives [MOON, STAR, ARROW, KEY] in the feature space (I_1, I_5, I_6). The distance to the corresponding representatives is much smaller than to other representatives in most cases.

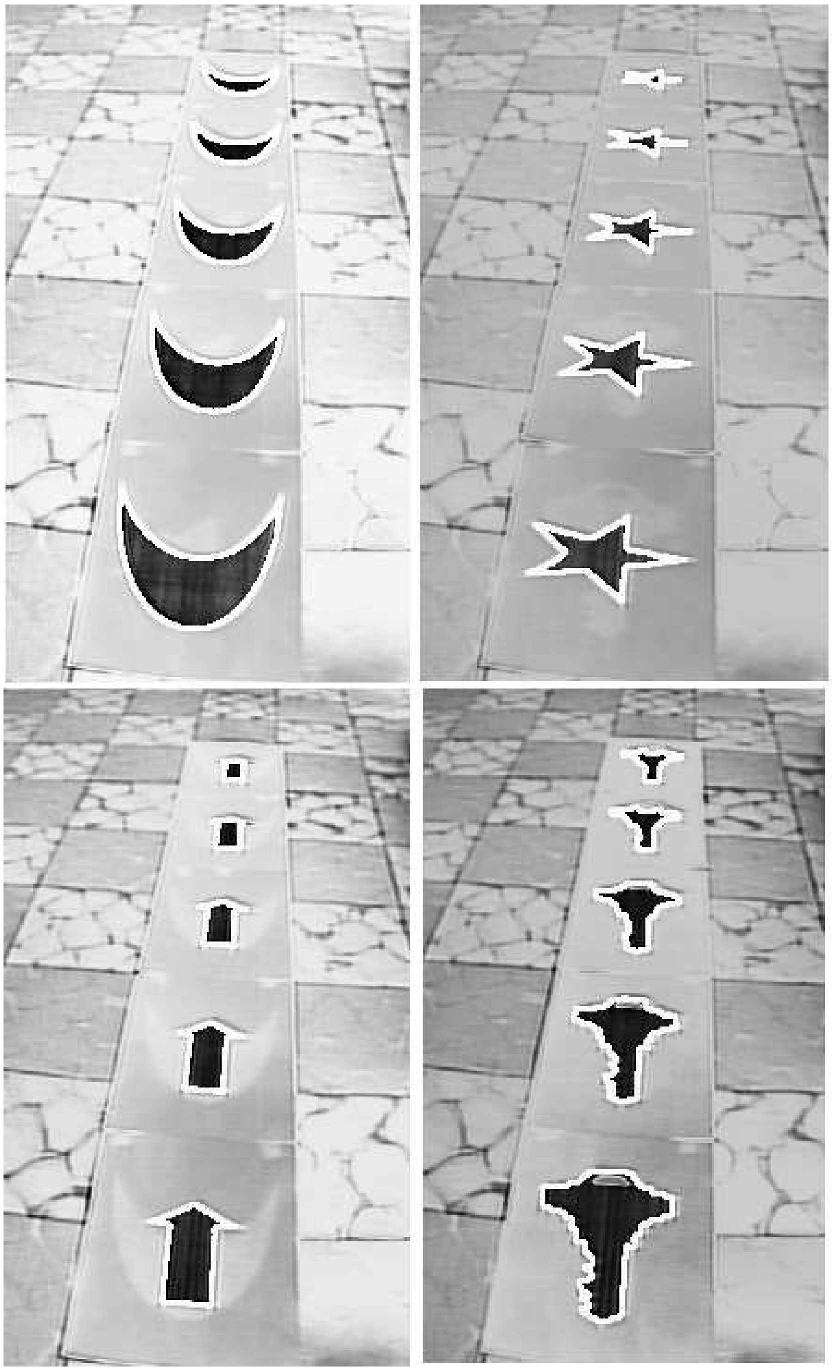

Figure 4: The last five landmark images from every set [MOON, STAR, ARROW, KEY]; the detected landmark boundaries are shown.

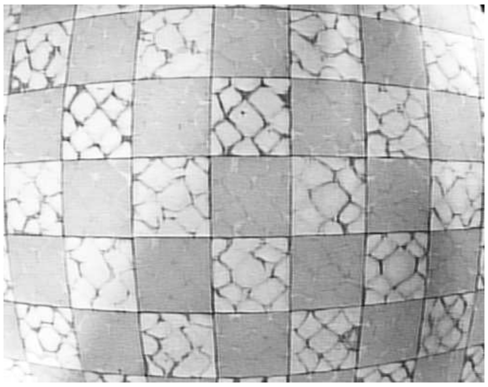

Figure 5: The deformation of a regular grid pattern, introduced by the fish-eye lens.

Image	1	2	3	4	5	6	7	8	9	10
MOON	1	1	1	1	1	1	1	1	1	1
STAR	2	2	2	2	2	2	2	4	4	2
ARROW	3	3	3	3	3	3	3	3	3	3
KEY	4	4	4	4	4	4	4	4	4	4
Image	1	2	3	4	5	6	7	8	9	10
MOON	1	1	1	1	1	1	1	1	1	1
STAR	2	2	2	2	2	2	2	2	2	4
ARROW	3	3	3	3	3	3	3	3	3	3
KEY	4	4	4	4	2	4	2	4	1	1
Image	1	2	3	4	5	6	7	8	9	10
MOON	1	1	1	1	1	1	1	1	1	1
STAR	2	2	2	2	2	2	2	2	2	2
ARROW	3	3	3	3	3	3	3	3	3	3
KEY	4	4	4	4	4	4	4	4	4	4

Table 2: The results of image classification. Every table row shows results of the classification of ten images at the specified positions. The value 1 stands for MOON, 2 for STAR, 3 for ARROW and 4 for KEY shapes, respectively. Three feature vector sets were used, computed from differently segmented images. The first group was segmented using a simple thresholding technique, in the second case a boundary following technique was applied and finally a manual thresholding was realized. The three table sections correspond to the three used thresholding techniques.

The ability of AMIs to make a correct classification of landmark images was tested. The minimum distance between the acquired landmark and the corresponding reference one from

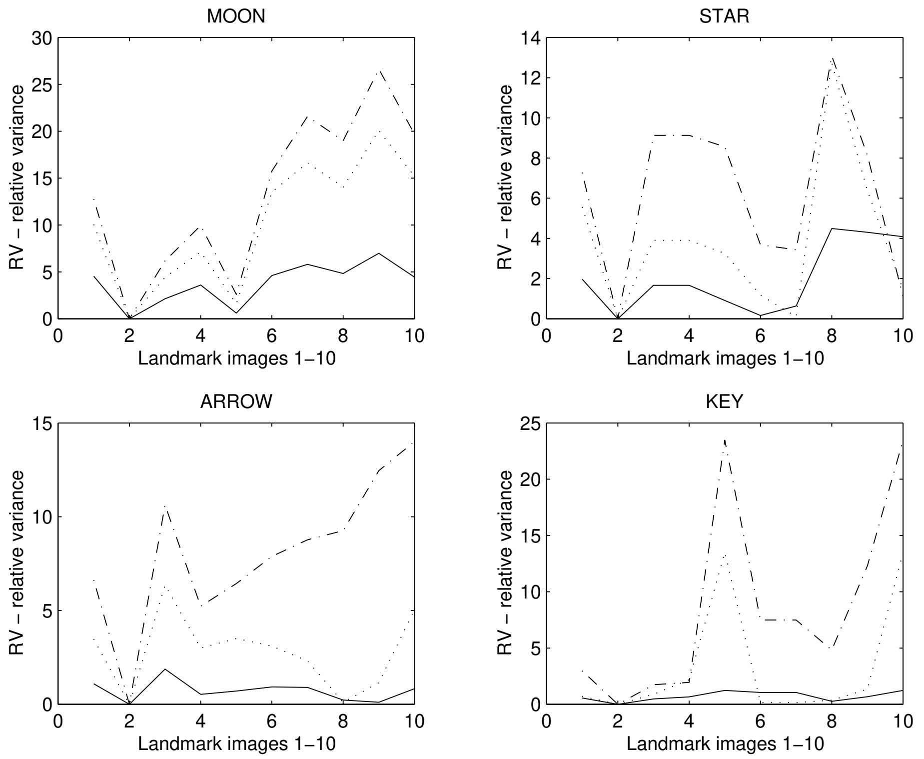

Figure 6: The stability of AMIs I_1, I_5 and I_6; for every type of landmark shape [MOON, STAR, ARROW, KEY] the relative variance RV (defined above) is plotted for I_1, I_5 and I_6 AMIs, the X-axis represents ten specified positions of the landmark, the Y-axis corresponds to the RV value. The I_1, I_5 and I_6 are plotted by solid, dotted and dashdotted lines, respectively.

the database (MOON, STAR, KEY, ARROW) was computed in the feature space (I_1, I_5, I_6). Results of this experiment are shown in Tab. 2. The images were classified using data obtained from images segmented by simple thresholding, boundary following and manual thresholding. In the first case, 38 images out of 40 (95%) were correctly classified, in the second 35 out of 40 (87.5%) and in the last case all images (100%) were correctly classified. It can be seen that the landmarks are classified correctly even when the introduced degradation is perceptible. The proposed recognition model (AMIs) showed sufficient stability and robustness under the perspective transform together with the fish-eye lens deformation.

5. Conclusions

In this paper, a method for the recognition of landmarks acquired by a fish-eye lens camera was introduced. This model is based on affine moment invariants.

The performance of the proposed method was demonstrated by experiments with acquired images of landmarks. Although the AMIs are invariant theoretically just under the affine trans-

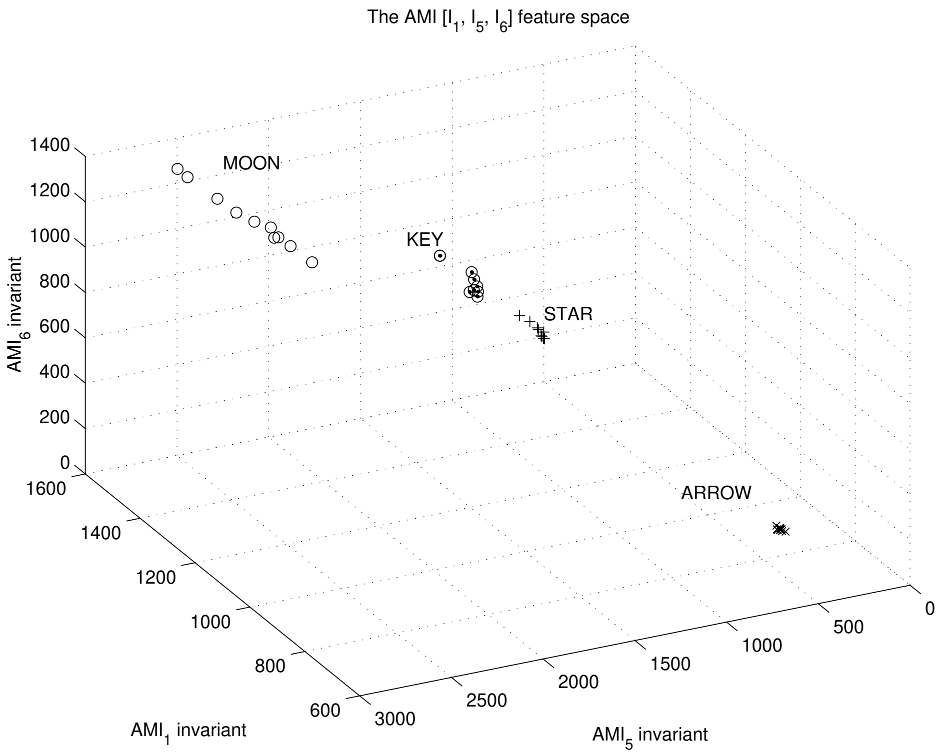

Figure 7: The feature space (I_1, I_5, I_6) with plotted feature vectors, corresponding to the segmented landmarks. Every cluster is labeled with the reference shape type.

form and we worked with images transformed by the projective transform and degraded by the fish-eye lens camera acquisition, AMIs recognition ability is high enough even in that case.

The presented experiments proved the possibility of using the AMIs for the recognition of landmarks for the navigation of the mobile robot, equipped with a fish-eye lens camera.

Acknowledgement

The first author acknowledges the support of the Institute of Information Theory and Automation, Czech Academy of Sciences, and of LANIA and Universidad Veracruzana during her visit to Mexico.

This work was supported by the Mexican National Council of Science and Technology (CONACYT) as project ref. C098-A.

References

Becker, C., Salas, J., Tokusei, K. and Latombe, J.C. (1995). Reliable navigation using landmarks. *IEEE International Conference on Robotics and Automation*, 401-406.

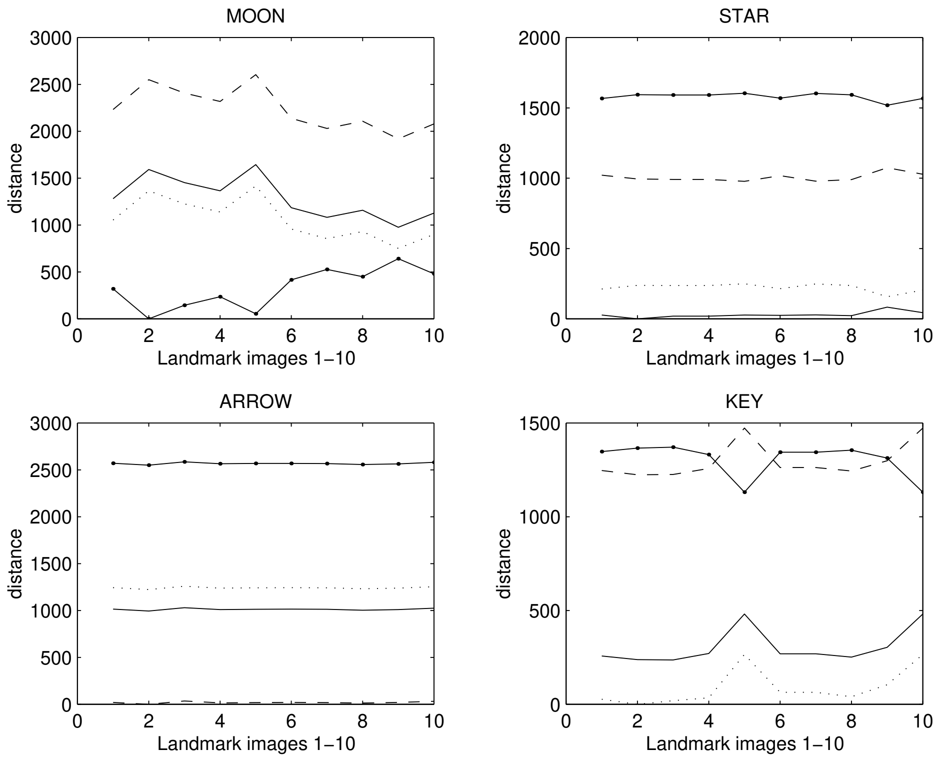

Figure 8: The distance in the feature space (I_1, I_5, I_6) between the reference vector and segmented landmark vectors for every type of landmark shape [MOON, STAR, ARROW, KEY]. The distance from the image to the representatives of MOON, STAR, ARROW and KEY are plotted with points, solid, dashed and dotted lines, respectively.

Belkasim, S.O., Shridhar, M. and Ahmadi, M. (1991). Pattern recognition with moment invariants: A comparative study and results. *Pattern Recognition* 24, 1117-1138.

Brooks, R.A. (1986). A robust layered control system for a mobile robot. *IEEE Trans. On Robotics and Automation* 2(1), 14-23.

Cox, I.J. and Wilfong, G.T. (1990). Autonomous Robot Vehicles, *Springer-Verlag.*

Dulimarta, H.S. and Jain, A.K. (1997). Mobile robot localization in indoor environment. *Pattern Recognition* 30(1), 99-111.

Flusser, J. and Suk, T. (1994). A moment based approach to registration of images with affine geometric distortion, *IEEE Transaction on geoscience and remote sensing* 32(2), 382 - 387.

Flusser, J. and Suk, T. (1993). Pattern recognition by affine moment invariants. *Pattern recognition* 26, 167-174.

Goshtasby, A. (1985). Description and discrimination of planar shapes using shape matrices. *IEEE Trans. Pattern Anal. Mach. Intell.* 7, 738-743.

Hu, M.K. (1962). Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory* 8, 179-187.

Kortenkamp et al. (1993). Integrated Mobile-Robot Design *IEEE Expert* August 1993, 61-73.

Kosaka, A. and Kak, A.C. (1993). Fast vision-guided mobile robot navigation using model-based reasoning and prediction of uncertainties. *CVGIP: Image understanding* 56(3), 271-329.

Leonard, J.J., Durrant-Whyte, H.F. and Cox, I.J. (1990). Dynamic Map Building for an Autonomous Mobile Robot. *IEEE International Workshop on Intelligent Robots and Systems*, 89-95.

Lewitt, T., Lawton, D., Cheldberg, D. and Nelson, P. (1987). Qualitative navigation. *Proc. Image Understanding Workshop*, 447-465.

López, I., Salas, J. and Gordillo, J.L. (1999). Navegación de robots móviles utilizando marcas artificiales. *Soluciones Avanzadas* 58, 54-61.

Moravec, H.P. (1981). Robot Rover Visual Navigation. *Ann Arbor, MI: UMI Research Press*, Michigan.

Peli, T. (1981). An algorithm for recognition and localization of rotated and scaled objects. *Proc. IEEE* 69, 483-485.

Prokop, R.J. and Reeves, A.P. (1992). A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processing* 54, 438-460.

Shah, S. and Aggrawal, J.K. (1995a). Modeling structured environments using robot vision. *Proceedings of 1995 Asian Conference on Computer Vision*, Singapore, 297 - 304.

Shah, S. and Aggrawal, J.K. (1995b). Autonomous Mobile Robot Navigation Using Fish-Eye Lenss. *Proceedings of 1995 Third International Computer Science Conference*, Honkong, December 1995.

Weiss, I. (1988). Projective invariants of shapes. *Proc. DARPA Image Understanding Workshop*, Cambridge, MA, 1125-1134.

Yeh, E. and Kriegman, D.J. (1995). Toward Selecting and Recognizing Natural Landmarks. *Tech.Report 9503*, Yale University.

Zahn, C.T. and Roskies, R.Z. (1972). Fourier descriptors for plane closed curves. *IEEE Trans. Comput.* 21, 269-281.

Zitova, B. and Flusser, J. (1999): Landmark Recognition using Invariant Features. *accepted for publications in Pattern Recognition Letters.*