Improving Self-localization Efficiency In a Small Mobile Robot by Using a Hybrid Field of View Vision System

Submitted: 27th August 2015; accepted 18th September 2015

Marta Rostkowska, Piotr Skrzypczyński

DOI: 10.14313/JAMRIS_4-2015/30

Abstract:
In this article a self-localization system for small mobile robots based on inexpensive cameras and unobtrusive, passive landmarks is presented and evaluated. The main contribution is the experimental evaluation of the hybrid field of view vision system for self-localization with artificial landmarks. The hybrid vision system consists of an omnidirectional, upward-looking camera with a mirror, and a typical, front-view camera. This configuration is inspired by the peripheral and foveal vision co-operation in animals. We demonstrate that the omnidirectional camera enables the robot to detect quickly landmark candidates and to track the already known landmarks in the environment. The front-view camera guided by the omnidirectional information enables precise measurements of the landmark position over extended distances. The passive landmarks are based on QR codes, which makes possible to easily include in the landmark pattern additional information relevant for navigation. We present evaluation of the positioning accuracy of the system mounted on a SanBot Mk II mobile robot. The experimental results demonstrate that the hybrid field of view vision system and the QR code landmarks enable the small mobile robot to navigate safely along extended paths in a typical home environment.

Keywords: *self-localization, artificial landmark, omnidirectional camera*

1. Introduction

An important requirement for any mobile robot is to figure out where it is within its environment. The pose of a wheeled robot (position and orientation $x_R = [x_R\ y_R\ \theta_R]^T$) can be estimated by means of odometry, but this method alone is insufficient [27], and the pose has to be corrected using measurements from external sensors. Although there are many approaches to self-localization known from the literature, nowadays the Simultaneous Localization and Mapping (SLAM) is considered the state-of-the-art approach to obtain information about the robot pose [7].

The SLAM algorithms estimate from the sensory measurements both the robot pose and the environment map, thus they do not need a predefined map of the workspace. This is an important advantage, because obtaining a map of the environment that is suitable for self-localization is often a tedious and time-consuming task. However, the known SLAM algorithms require data from highly precise sensors, such as laser scanners [28], or have high computing power demands, if less precise data (e.g. from passive cameras) are used [8]. Thus, the SLAM approach is rather unsuitable for small mobile robots, such like our SanBot [19], which have quite limited resources with respect to on-board sensing, computing power, and communication bandwidth. Thus, for such a robot an approach to self-localization that does not need to construct a map of the environment, or uses a simple and easy to survey representation of the known area is required. Moreover, the self-localization system should use data from compact and low-cost sensors.

In the context of navigation CCD/CMOS cameras are the most compact and low-cost sensors for mobile robots [6]. However, most of the passive vision-based localization methods fail under natural environmental conditions, due to occlusions, shadows, changing illumination, etc. Therefore, in practical applications of mobile robots artificial landmarks are commonly employed. They are objects purposefully placed in the environment, such as visual patterns or reflecting tapes. Landmarks enhance the efficiency and robustness of vision-based self-localization [29]. It was also demonstrated that simple artificial landmarks are a valuable extension to visual SLAM [3]. An obvious disadvantage is that the environment has to be engineered. This problem can be alleviated by using simple, cheap, expendable and unobtrusive markers, which can be easily attached to walls and various objects. In this research we employ simple landmarks printed in black and white that are based on the matrix QR (Quick Response) codes commonly used to recognize packages and other goods.

In our recent work [21] we evaluated the QR code landmarks as self-localization aids in two very different configurations of the camera-based perception system: an overhead camera that observed a landmark attached on top of a mobile robot, and a front-view camera attached to a robot, which observed landmarks freely placed in the environment. Both solutions enable to localize the robot in real-time with a sufficient accuracy, but both have important practical drawbacks. The overhead camera provides inexpensive means to localize a group of few small mobile robots in a desktop application, but cannot be easily scaled up for larger mobile robots operating in a real environment. The front-view camera with on-board image processing is a self-contained solution for self-localization, which enables the robot to work autonomously, making it independent from possible com-

munication problems. However, the landmarks are detectable and decodable only over a limited range of viewing configurations. Thus, the robot has to turn the front-mounted camera towards the area of landmark location before it starts to acquire an image. In a complicated environment, with possible occlusions this approach may lead to a lot of unnecessary motion. Eventually, the robot can get lost if it cannot find a landmark before the odometry drifts too much.

In this paper we propose an approach that combines to some extent the advantages of the overhead camera and the front-view camera for self-localization with passive landmarks, avoiding the aforementioned problems. We designed an affordable hybrid field of view vision system, which takes inspiration from nature, and resembles the peripheral and foveal vision in animals. The system consists of a low-cost omnidirectional camera and a typical, front-view camera. The omnidirectional component, employing an upward-looking camera and a profiled mirror provides to the robot an analogy of the peripheral vision in animals. It gives the robot the ability to quickly detect interesting objects over a large field of view. In contrast, the front-view camera provides an analog of foveal vision. The robot can focus on details of already detected objects in a much narrower field of view. The cooperation of these two subsystems enables to track in real-time many landmarks located in the environment, without the need to move the robot platform, whereas it is still possible to precisely measure the distances and viewing angle to the already found landmarks.

The reminder of this paper is organized as follows: In the next Section we analyze the most relevant related work. Section 3 introduces the concept and design of the hybrid vision system, whereas the landmarks based on QR codes and the image processing algorithms used in self-localization are described in Section 4. The experimental results are presented in Section 5. Section 6 concludes the paper and presents an outlook of further research.

2. Related Work

The advantages of biologically-inspired vision for robot self-localization have been demonstrated in few papers – for instance Siagnian and Itti [25] have shown that extracting the "gist" of a scene to produce a coarse localization hypothesis, and then refining this hypothesis by locating salient landmark points enables the Monte-Carlo localization algorithm to work robustly in various indoor/outdoor scenarios. However, in this work both the global and the local characteristics of the scene were extracted from typical perspective-view images. One example of a system that is more similar to our approach and mimics the cooperation between the peripheral vision and the foveal vision in humans is given by Menegatti and Pagello [16]. They investigate cooperation between an omnidirectional camera and a perspective-view camera in the framework of a distributed vision system, with the RoboCup Soccer as the target applications. Only simple geometric and color features of the scene are considered in this system. An integrated, self-contained hybrid field of view vision system called HOPS (Hybrid Omnidirectional Pin-hole Sensor), which is quite similar in concept to our design is presented in [5], where the calibration procedure is described that enables to use this sensor for 3D measurements of the scene. Unfortunately, [5] gives no real application examples. Also Adorni *et al.* [1] describe the use of a combined peripheral/foveal vision system including an omnidirectional camera in the context of mobile robot navigation. Their system uses both cameras in a stereo vision setup and implements obstacle detection and avoidance, but not self-localization.

Although the bioinspired vision solutions in mobile robot navigation mostly extract natural salient features, in many practical applications artificial landmarks are employed in order to simplify and speed-up the image processing and to make the detection and recognition of features more reliable [15]. Visual self-localization algorithms are susceptible to errors due to unpredictable changes in the environment [11], and require much computing power to process natural features, e.g. by employing local visual descriptors [24]. The need to circumvent these problems in a small mobile robot that is used for education and requires reliable self-localization, offering only limited computing resources motivated us to enhance the scene by artificial landmarks. Although active beacons can be employed, such like infra-red LEDs [27], most of the artificial visual landmarks are passive. This greatly simplifies deployment of the markers and makes them independent of any power source. Depending on the robot application and the characteristics of the operational environment very different designs of passive landmarks have been proposed [9, 22]. In general, simple geometric shapes can be quickly extracted from the images, particularly if they are enhanced by color [3]. A disadvantage of such simple landmarks is that only very limited information (usually only the landmark ID) can be embedded in the pattern. In contrast, employing in landmark design the idea of barcode, either one-dimensional [4] or two-dimensional [12] makes it possible to easily encode additional information. In particular, matrix codes, that proliferated recently due to their use in smartphone-based applications enable to fabricate much more information-rich landmarks. Moreover, landmarks based on matrix codes are robust to partial occlusion or damage of the content. Landmarks based on matrix codes are unobtrusive – their size can be adapted to the requirements of particular application and environment. As they are monochromatic, they can be produced in a color matching the surroundings, partially blending into the environment. The robotics and computer vision literature provides examples of successful applications of QR codes for mobile robot self-localization. Introducing QR codes into the environment has improved the robustness and accuracy of the 3D-vision-based Monte Carlo self-localization algorithm in a dynamic environment as demonstrated in [14]. The information-carrying capability of matrix codes can be efficiently used for self-localization and communication in a system of many mobile robots [18] and in an intelligent home space for service robot

[13]. An applicability of QR codes for navigation and object labelling has been also demonstrated in [10] on the NAO humanoid robot.

3. Hybrid Field of View Vision System

3.1. Concept and Components

Most of the mobile robots that employ vision for navigation use typical perspective cameras. A perspective camera can observe landmarks located at relatively large distances and positioned arbitrary in the environment within the camera's horizontal field of view. The distance to the robot and orientation of the landmark can be calculated from a single image taken by the perspective camera. Due to practical considerations, working indoors we assume that the landmarks are attached to vertical surfaces, such as walls that dominate man-made environments. Thus, we consider only the angle α between the camera's optical axis and the normal to the landmark's plane in 2D (Fig. 1) In the same camera coordinates the position of the landmark is defined by the distance z_y measured along the camera's optical axis, which is assumed to be coincident with the robot's y_R axis, and the distance d in the robot's x_R axis, computed as the offset between the center of the image (i.e. the optical axis) and the center of the landmark. The distance at which the landmark can be detected and recognized depends on the camera resolution and the physical size of the landmark [21]. The information about the actual landmark size, as well as the position and orientation in the global reference frame $x_L = [x_L\ y_L\ \theta_L]^T$ is encoded in the QR code of the landmark itself, so the robot doesn't need to keep a map of known landmarks in the memory. Therefore, if at least one landmark can be recognized and decoded, the position of the robot and its orientation can be computed. However, in order to find landmarks in the surroundings, the robot has to constantly change its heading, which is inconvenient.

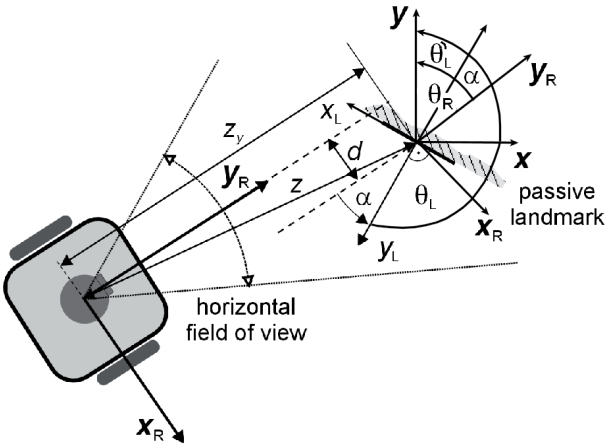

Fig. 1. Geometry of landmark measurements by using the perspective camera

The omnidirectional subsystem combines a standard upward-looking camera with an axially symmetric mirror located above this camera and provides 360° field of view in the horizontal plane. This type of omnidirectional sensor is called catadioptric [23] and can be implemented using mirrors of different vertical profiles: parabolic, hyperbolic, or elliptical. The omnidirectional sensor used in this research has been designed and built within a project run by students, which imposed limitations as to the costs and the used technology. The mirror has been fabricated in a workshop from a single piece of aluminium using a simple milling machine, which limited the achievable curvature of the profile. Thus, a mirror of conical shape with a rounded, parabolic tip was designed (Fig. 2). This profile could be fabricated at acceptable cost using typical workshop equipment. The mirror is hold by a highly transparent acrylic tube over the lens of an upward-looking webcam.

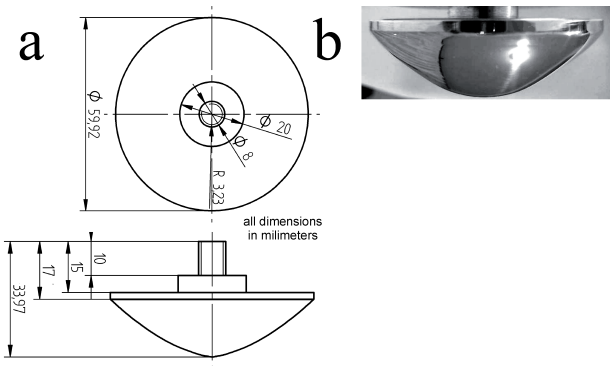

Fig. 2. Conical mirror: a – design of the conical mirror with rounded tip for the omnidirectional vision sensor, b – the fabricated mirror

Omnidirectional camera images represent geometrically distorted environment such as: straight lines are arcs, squares are rectangles. For this reason it is difficult to find characteristic elements which are needed in the localization process. It is therefore necessary to transform images using the single effective viewpoint [26]. Unfortunately, the chosen shape of the mirror makes it hard to achieve the single effective viewpoint property in the sensor. While for hyperbolic or elliptical mirrors this is simply achieved by placing the camera lens at a proper distance from the mirror (at one of the foci of the hyperbola/ellipse), for a parabolic mirror, an orthographic lens must be interposed between the mirror and the camera [2]. This was impossible in the simple design sensor which uses a fixed-lens webcam as the camera. Therefore, it is impossible to rectify the images captured by our omnidirectional camera to geometrically correct planar perspective images [26]. While the captured pictures may be mapped to flat panoramic images covering the 360° field of view, these images are still distorted along their vertical axis, i.e. they do not map correctly all the distances between the objects and the sensor into the vertical pixel locations. However, there are no distortions along the horizontal axis, which allows to recover the angular location of the observed objects with respect to the sensor. In the context of landmark-based positioning it means that while the landmarks can be detected in the omnidirectional images, only their angular locations, but not the distances with respect to the robot can be determined precisely, particularly for more distant landmarks. Moreover, the internal content of

the landmark (QR code) cannot be decoded reliably from the distorted images. Eventually, while the omnidirectional camera is capable of observing the whole proximity of the robot without unnecessary motion, it requires high computing power to rectify the whole images, still giving no guarantee that the geometric measurements of landmark positions are precise enough for self-localization.

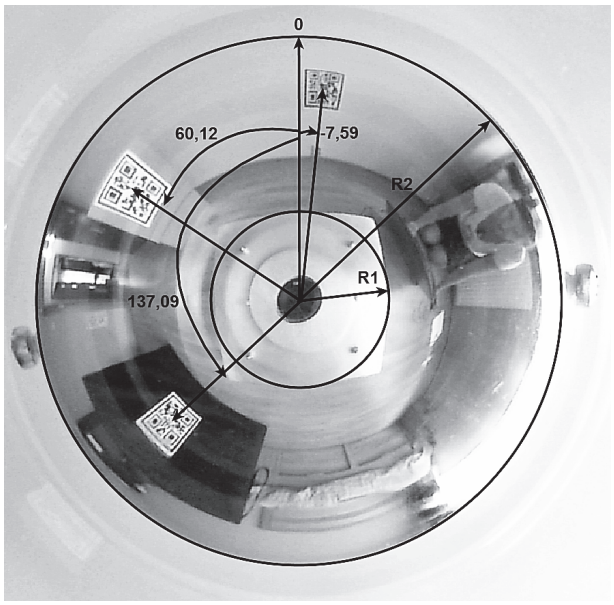

Fig. 3. Exemplary view from the omnidirectional vision component with artificial landmarks in the field of view

The aforementioned properties and limitations of the two camera subsystems resemble the characteristics of the foveal and peripheral vision in animals. This provides a strong argumentation to combine both systems. If the perspective view camera and the omnidirectional camera subsystems are coupled for landmark perception their drawbacks can be mutually compensated to a great extent. The omnidirectional camera can provide 360° view with detection of landmarks, and then guide the perspective camera to the angular coordinates of the found landmarks. The perspective camera can be pointed directly to the landmark at the known angular coordinates, and then can precisely measure its location and read the QR code. It should be noted, that in this cooperation scheme neither full rectification of the omnidirectional images or the perspective correction in the front-view camera images are needed, which significantly decreases the required computing power.

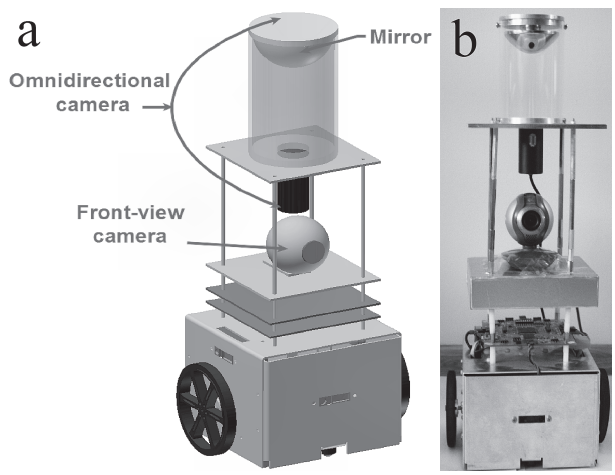

Fig. 4. SanBot Mk II with the hybrid field of view vision system: CAD drawing (a), and a photo of assembled robot (b)+

3.2. Experimental System on the Mobile Robot

The experimental mobile robot with the hybrid field of view vision system is shown in Fig. 4. It is based on the small, differential-drive mobile platform SanBot Mk II [19]. The robot is equipped with the front-view camera and the omnidirectional camera.

The front-view camera is mounted directly to the upper plate of the robot's chassis. It is a Logitech 500 webcam, providing images at the resolution of 1280x1024. The Microsoft LifeCam webcam is used in the omnidirectional sensor. This particular camera has been chosen due to its compact size, high resolution (1280x720), and an easy to use API. The mirror has the diameter of 6 cm, and is located 11 cm above the camera lens. The omnidirectional camera is positioned precisely above the front-view camera. Both cameras stream images at 15 FPS through the USB interface.

In the current experimental setup image processing takes place on a notebook PC. The simple controller board of the SanBot robot receives only the calculated positions of the landmarks that are necessary to compute the motion commands. These data are transferred via a serial (COM) port. The robot schedules the sequence of motions to execute in order to follow the planned path. The robot stops for a moment when taking images, and then obtains the outcome of calculations related to landmark-based self-localization.

4. Landmarks and Self-localization

4.1. Passive Landmarks with Matrix Codes

There are many possibilities to design a passive landmark, but if a CCD/CMOS camera has to be used as the sensor, the landmark should have the following basic properties:

- should be recognizable and decodable over a wide range of viewing ranges and angles;
- should be easily recognizable under changing environment conditions (e.g. variable lighting, partial occlusions);
- its geometry should allow easy and quick extraction from an image;
- it should be easy to prepare, preferably printable in one color;
- it should be unique within the robot's working area, e.g. by containing an encoded ID.

For the small mobile robot positioning we formulate a further requirement related to the limited computing power of the system: the landmarks should be able to carry additional information related to self-localization and navigation, such like the position of the landmark in the global frame, object labels or guidance hints. Such information easily and robustly decodable from the landmark's image helps the robot to navigate without building a map of the environment in memory.

All of the above-listed requirements are met by matrix codes. In our previous work [20] we have experimentally evaluated four types of commercially used matrix codes as candidates for landmarks. The results revealed that among these code types, the most suitable for navigation are the QR codes. The QR codes contain three marker positions (upper left, upper right and lower left), which are additionally separated from the data with white frame. This pattern allows easily recovering the code orientation. Comparing to other considered variants the QR code is also characterized by large size of a single module (i.e. white/black cell). This is an important advantage, which ensures proper measurements, even for long distances. In addition, QR codes are capable of partial error correction if they are damaged or occluded.

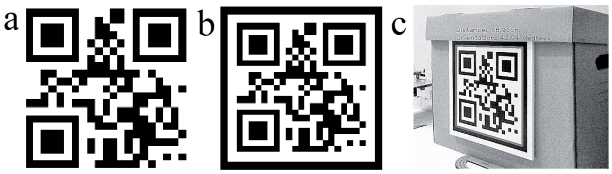

Fig. 5. Exemplary QR code based landmark: QR code encoding value '1' (a), 16.5x16.5 cm landmark with frame (b), recognized landmark in the environment (c)

As a result, our landmark is designed around a standard QR code, which is placed in the center. The code is encompassed with a black frame, which is used for initial sorting of landmark candidates on images and for reducing the number of potential objects, which can be subject to decoding and further processing. Landmarks are monochromatic (usually black-and-white), because they should be extremely low-cost and printable on any material, not only paper. An example of a QR code, in which the ID '1' has been encoded, is shown in Fig. 5a. A complete landmark with the added black frame is depicted in Fig. 5b. The same landmark, recognized and decoded in the environment is shown in Fig. 5c.

The program processing the data from both cameras has been created using C# programming language and Microsoft Visual Studio 2010. The basic image processing part, leading to conversion of the omnidirectional images, extraction of landmark candidates from these images, and computation of the geometric measurements is implemented with the EmguCV library [30], which is a C# port of the well-known OpenCV. The decoding of extracted QR codes is accomplished using the specialized MessagingToolkit library [32].

4.2. Localization of Landmarks on Perspective Camera Images

We assume that landmarks are mounted on rigid surfaces, so that they will not bend, deforming the square frames. Thus, the images from the perspective view camera are assumed to be not distorted. The algorithm of recognition and localization of landmarks observed by the perspective view camera is shown in Fig. 6. The image processing begins with acquiring images from the front-view camera. The images are filtrated, and then the thick frames around the QR code are searched for by extracting candidate rectangles.

If the surface of a landmark is roughly parallel to the camera sensor's surface, there is no perspective deformation of the QR code, and it can be directly processed by the appropriate MessagingToolkit routine. If such a landmark is found, the distance to the robot's camera is calculated. Whenever the camera's optical axis intersects the center of the landmark (i.e. it is located horizontally in the center of the image) the distance is calculated in a simple way:

$$z = \frac{f * h_L}{h_I}, \tag{1}$$

where z is the distance between the landmark and the camera, f is the camera's focal length, h_L is the known vertical dimension (height) of the landmark, and h_I is the observed object's vertical dimension on the image. The viewing angle can be computed from the formula:

$$\propto = \arccos\left(\frac{w_I}{w_L}\right) * \frac{180°}{\pi}, \tag{2}$$

where w_L is the known horizontal dimension (width) of the landmark, and w_I is the observed object's horizontal dimension on the image.

However, if the landmark isn't located in the center of the image (cf. Fig. 1), the distance between the camera and the landmark is calculated from the right-angle triangle made by the distance z_y measured along the camera's optical axis (which is assumed to be coincident with the robot's y_R axis), and the distance d in the robot's x_R axis, computed as the offset between the center of the image and the center of the landmark:

$$d = \frac{d_p * h_L}{h_I}, \tag{3}$$

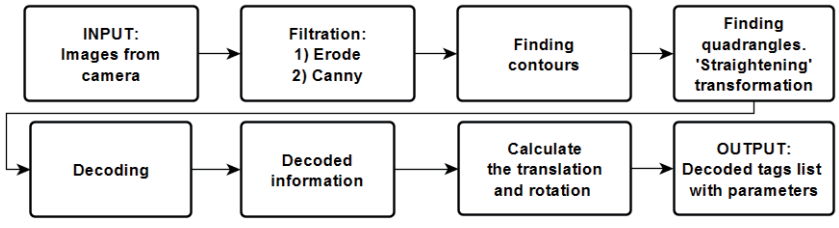

Fig. 6. Landmarks detection and decoding algorithm for the perspective camera

where d_p is the distance in pixels between the center of the image and the center of the landmark's bounding frame.

The viewing angle between the camera's optical axis and the vector normal to the landmark surface is calculated as:

$$\propto = \arctan\left(\frac{d}{z_y}\right) * \frac{180°}{\pi}, \tag{4}$$

where z_y is the perpendicular distance from camera to the landmark, and d is the distance calculated from (3).

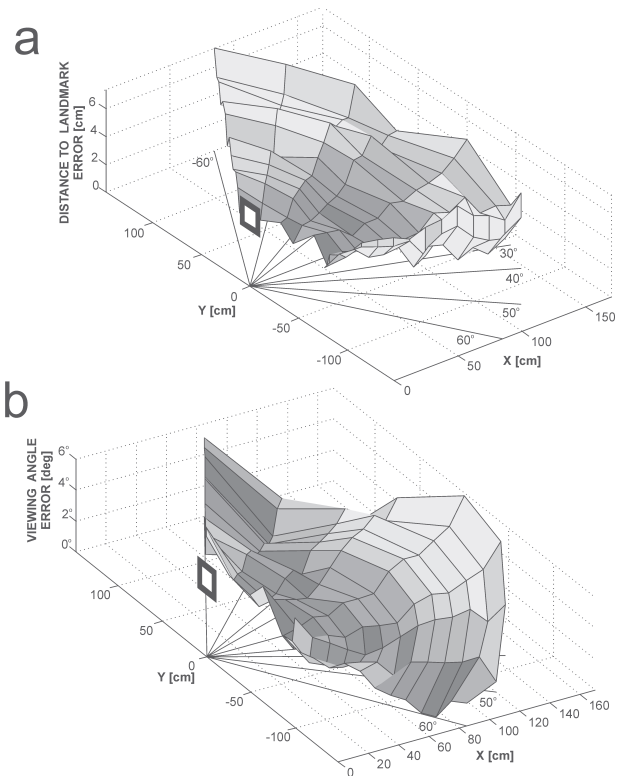

Fig. 7. Spatial distribution of errors in QR code-based landmark localization by the perspective view camera: distance to landmark z errors (a), and viewing angle α errors (b)

However, if in the given viewing configuration of the perspective camera the surface of a landmark is not parallel to the camera's sensor's surface the perspective deformation of the landmark's image has to be corrected before decoding the QR code and calculating the distances and angle from (1)–(4). In such a case the relation between locations of the characteristic points (corners) in 3D and the image plane has to be found in order to properly calculate the landmark's position and rotation. Computations of this relation are described in more details in [21]. We omit these calculations here, because such situations should not occur when self-localization takes place by using both the perspective and omnidirectional cameras, as the perspective camera is set to a proper angular position before taking an image of the landmark. Thus, the viewing angle of the landmark never exceeds 15°. Results of our earlier experiments [21] provide evidence that for such small viewing angles the correction of perspective brings no improvement in the landmark localization, while this procedure is computation intensive.

Quantitative results for the measurements of an exemplary passive landmark (size 20x20 cm) are shown in Fig. 7. The landmark was observed by the perspective camera from distances up to 2 meters and for the viewing angles up to 60°. In this experiment the camera was positioned in such a way that the optical axis always intersected the center of the landmark, thus the d offset was zero. As it could be expected, the distance measurement error grows for larger distances, but it grows also slightly for large viewing angles (Fig. 7a), which could be attributed to the not corrected perspective deformation. As can be seen from the plot in Fig. 7b the measured viewing angle is less precise for large and very small distances. This is probably caused by the procedure searching for the thick black frame, which for very large images of a landmark (small distances) occasionally finds the inner border of the frame instead of the outer one. Average precision of the measurements (over all distances and viewing angles) turned out to be 1.3 cm for the distance and 2° for the viewing angle.

4.3. Recognition of Landmarks on the Omnidirectional Images

As described in Section 3 the low-cost omnidirectional camera geometry and optics do not permit full rectification of the distorted images. Therefore, we use images from the omnidirectional camera only to find potential landmark candidates in the robot's vicinity, and to track the known landmarks.

At the beginning, in order to reduce the amount of processing information, the color image from the camera is transmuted into black-and-white image. The data processing starts by cropping and unwinding the omnidirectional image. Cropping the image relies on selecting the part of the picture which is necessary for recognition of the landmarks. The unwinding procedure is a simple cylinder unrolling. At the beginning the algorithm sets the height and width of the unrolled picture:

$$H = R_2 - R_1$$

$$W = 2\pi\frac{H}{2}, \tag{5}$$

where R_2 is the radius of outer circle, and R_1 is the radius of an inner circle marked in Fig 3. Next, the algorithm starts computing a new position of each pixel in the unrolled image. This procedure is shown in pseudo-code:

Listing 1. Pixel position calculation procedure

```
y = H - 1;
for (x = 0; x < W; x++)
{
 for (y = 0; y < H; y++)
 {
       r = (y/H) * (R2 - R1) + R1;
       theta =(x/W) * 2 * PI;
```

```
        xs = Cx + r * Sin(theta);
        ys = Cy + r * Cos(theta);
        mapx.Data[y, x] = xs;
        mapy.Data[y, x] = ys;
        y--;
    }
    y = H - 1;
}
```

The two operations described above provide the same result for each processed image, so they are executed only once at the beginning of the program. Afterwards, the program starts the procedure of unwinding the picture using the EmguCV *cvRemap* function. This function transforms the source image using the specified map (in our case map_x and map_y from the algorithm in Listing 1).

After the procedure of unwrapping the image, the extreme parts of the image are duplicated in the opposing ends in order to obtain a continuous image in areas where landmarks can appear. The unrolled image is shown in Fig. 8. This image undergoes morphological erosion in order to remove noise. Next, the Canny operator is used to find edges on the image. Among the found edges those that are connected into rectangular shapes are selected. Then, the algorithm eliminates all nested rectangles – those which are located inside other rectangles. The found landmark candidates are marked on the image.

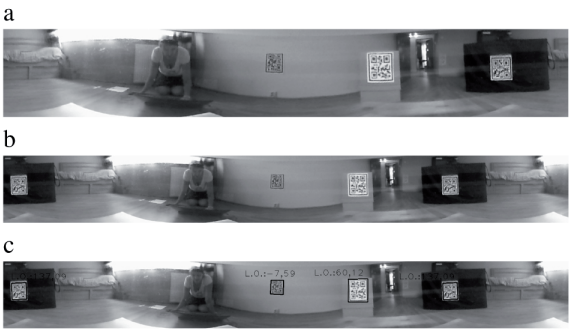

Fig. 8. Results from the omnidirectional camera a – cropped and unwrapped image, b – extended image, c – unwrapped and extended image with marked candidates

The viewing angle of the landmark with respect to the robot's heading is calculated as:

$$\alpha_{absolute} = x_s \frac{360}{W},$$

$$\alpha_{relative} = \alpha_{absolute} - \frac{(W + 2W_d) * 360}{2W}, \quad (6)$$

$$\alpha_{relative} = \begin{cases} \alpha_{relative} - 360 \text{ if } \alpha_{relative} \geq 180 \\ \alpha_{relative} + 360 \text{ if } \alpha_{relative} < 180 \end{cases}$$

where x_s and y_s define center of the landmark, W is width of the unwrapped image, and W_d is width of duplicated part of the picture.

Afterwards, the program makes a list of potential landmarks and relative angles. The algorithm of finding and localization of landmarks in the omnidirectional images is shown in Fig. 9.

4.4. Self-localization with the Hybrid System

The self-localization algorithm based on data from both the omnidirectional and perspective camera is shown in Fig: 10. At the beginning, the program processes only an image from the omnidirectional camera. If the algorithm described in Subsection 4.3 finds a landmark candidate at the viewing angle smaller than ±15°, the program starts processing the image from the front-view camera. This way the robot does not need to aim the perspective camera directly at the landmark, which speeds up the self-localization process.

When the landmark is seen in the angular sector of ±15°, the image from the perspective camera is processed. The program searches for the landmark, decodes it and calculates robot's position and orientation in the external reference frame (cf. Fig. 1). The orientation of the robot θ_R is concatenation of the landmark's orientation in the global coordinates θ_L, and robot's orientation with regard to the landmark α. The orientation is calculated as:

$$\theta_R = \alpha + \theta_L', \quad (7)$$

where θ_L' is θ_L – 180° and α is the angle calculated from (4). The robot's position in the global reference frame is calculated as:

$$x_R = x_L \pm d,$$

$$y_R = y_L \pm z_y, \quad (8)$$

where x_L and y_L define the landmark's position, z_y is the perpendicular distance between the camera and the landmark, and d is the distance calculated from (3). In (8) the plus sign is used to compute the

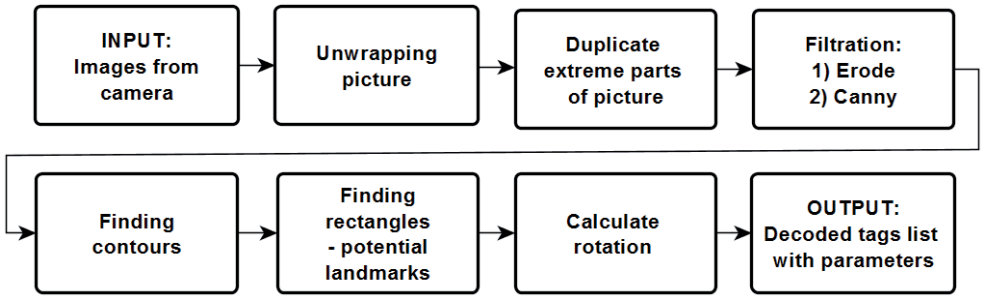

Fig. 9. Landmarks detection for omnidirectional camera

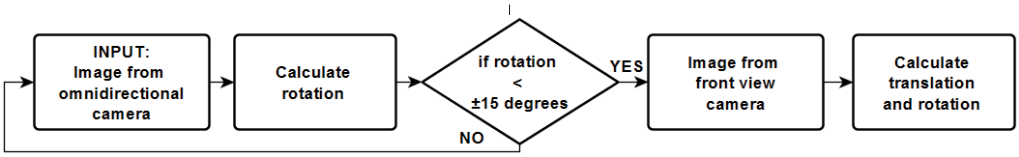

Fig. 10. Landmarks' detection and decoding algorithm for the hybrid system

position in *x*-axis when the landmark is located to the right side of the robot, and minus when it is on the left. At the beginning of calculations the algorithm assumes that the landmark is located in front of the robot, and uses plus sign in (8) to compute the position in *y*-axis. But if the robot has to rotate 180° to decode the landmark, the algorithm uses minus sign.

If the candidate landmark is found at a viewing angle larger than ±15°, the robot turns around until the angle becomes smaller than ±15°. The most common situation is when the algorithm finds more than one potential landmark. In such case the robot turns towards the nearest landmark. Although the front-view camera is capable of recognizing landmarks that are visible at the angles up to ±60°, to ensure robustness, the QR codes are decoded when they are visible at an angle at most ±15°. Images from the front-view camera are processed only if the omnidirectional camera finds a potential landmark. If we cannot find any landmark candidates in the unwrapped image, the robot does not localize itself, and tries to continue using the odometry.

5. Experiments and Results

In order to verify the accuracy of the landmark-based self-localization and the usability of the hybrid vision system in practical scenarios, we have performed several experiments in a typical home environment.

In these experiments we used one SanBot Mk II equipped with the hybrid field of view vision system. The ground truth data about the covered path was collected by manually measuring the robot's

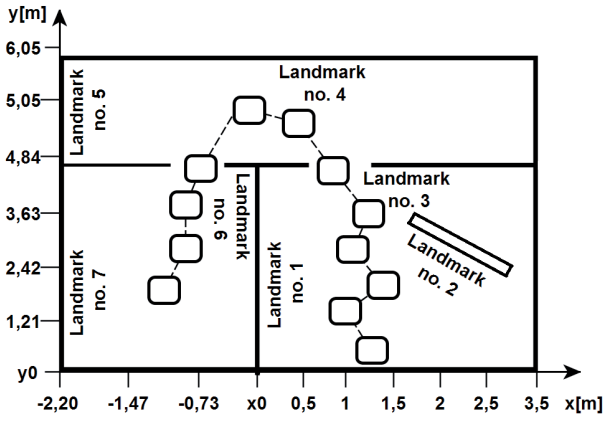

Fig. 11. Robot's path during the experiment. Small squares represent points, at which the robot stops and takes images

Tab. 1. Viewing angle determination results for the omnidirectional camera

Robot stop no.	α[°]	α^g[°]	Δα[°]
1	71.98	72.00	0.02
2	-9.19	-10.00	0.81
3	10.52	9.00	1.52
4	21.33	20.00	1.33
5	4.39	5.00	0.61
6	0.35	0.00	0.35
7	-37.69	-38.00	0.31
8	-3.56	-3.00	0.56
9	-49.88	-50.00	0.12
10	-3.40	-4.00	0.60
11	47.84	50.00	2.16
12	-2.91	-3.00	0.09

2D position with respect to the planned path that was marked on the floor with scotch tape. Here we present the quantitative results for the longest path, spanning three rooms (Fig. 11). During this experiment, the robot covered the planned path ten times, which enabled us to asses the repeatability of the measurements carried out by our vision system.

The test environment has seven landmarks, which contain their localizations and orientations with regard to external system of coordinates. Using only one landmark, its data and trigonometric relations, the robot can calculate its position. For this reason, landmarks in the environment are arranged so that the robot can always see at least one landmark. At the beginning program searches potential landmarks in images from the omnidirectional camera. If the algorithm finds a potential landmark and its absolute viewing angle is less than ±15°, the program starts processing data from the front-view camera. The robot stops near a detected landmark. If the algorithm finds a landmark candidate, but the angle is bigger than ±15°, the robot turns towards the landmark until the angle is

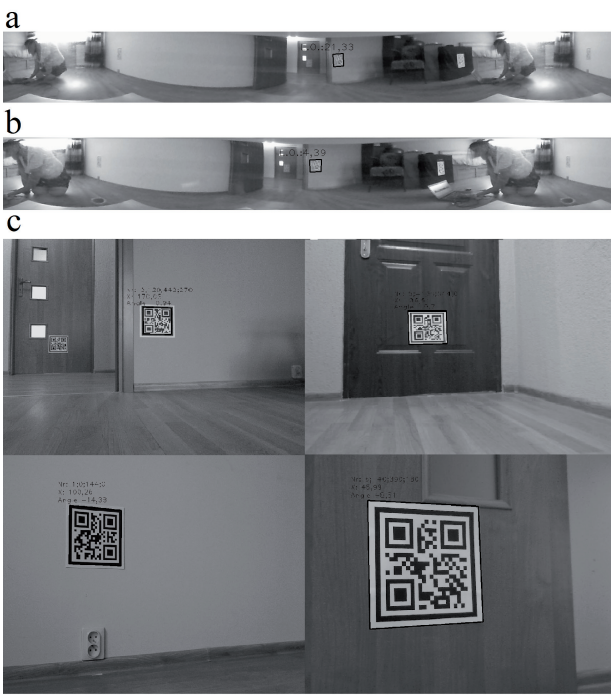

Fig. 12 Exemplary images of a measurement taken at a single robot stop: a – cropped and unwrapped image where the algorithm finds a potential landmark but the angle is larger than 15°, b – cropped and unwrapped image where algorithm finds a potential landmark and the angle is less than 15°, c – images from the perspective camera with marked landmark

smaller than ±15°. Then, the robot updates its pose from the computed localization data and continues to the next via-point on the planned path.

In this experiment the average error of determining the position of the robot was 3 cm in *x*-axis, 5 cm in *y* -axis, and the orientation error was 4°, when using the front-view camera. For the omnidirectional camera the orientation error was only 1°. This enables to compensate the degraded orientation accuracy in the robot pose by using data from the omnidirectional system. Sample images from the measurements are presented in Fig. 12. Results for the omnidirectional camera are shown in Tab. 1, where α denotes the measured angle and α^g the known ground-truth angle. Final results for the perspective camera-based self-localization guided by the omnidirectional camera data are shown in Tab. 2, where x_L, y_L, α_L describe the landmark position in the global frame, x_R, y_R, α_R denote the computed robot's pose in the same global frame, x^g_R, y^g_R, α^g_R are the ground truth coordinates of the robot, $\Delta x_R, \Delta y_R, \Delta\alpha_R$ define the absolute localization errors, and $\sigma x_R, \sigma y_R, \sigma\alpha_R$ the standard deviations of the localization measurements. Both tables contain average results from 10 runs along the same path. These results demonstrate that the system based on a combination of the omnidirectional camera and the perspective camera provides localization accuracy that is satisfactory for home environment navigation, and allows improving the results in comparison to a system using only the front-view camera.

6. Conclusions

This paper presents a new approach to mobile robot self-localization with passive visual landmarks. Owing to the hybrid field of view vision system even a small and simple robot can use passive vision for global self-localization, achieving both high accuracy and robustness against problems that are common in vision-based navigation: occlusions, limited field of view of the camera, and limited range of landmark recognition. The proposed approach enables to use low-cost hardware components and allows simplifying the image processing by avoiding full rectification and geometric correction of the images. The experiments conducted using a mobile robot demonstrated that the omnidirectional component can in most cases determine the viewing angle of a landmark with the accuracy better than 1°, using a partially rectified image. The positional accuracy of robot localization using the hybrid field of view system was in most cases better than 5 cm, which is satisfactory for home or

Tab. 2. Robot self-localization results for the hybrid system

L. no.	x_L [cm]	y_L [cm]	α_L [°]	x_R [cm]	y_R [cm]	α_R [°]	x^g_R [cm]	y^g_R [cm]	α^g_R [°]	Δx_R [cm]	Δy_R [cm]	$\Delta\alpha_R$ [°]	σx_R [cm]	σy_R [cm]	$\sigma\alpha_R$ [°]
1	0.00	144.00	90.00	93.90	105.86	-75.62	97.00	108.00	-65.00	3.10	2.14	10.62	2.00	3.63	9.34
2	235.00	250.00	210.00	123.86	167.89	34.28	120.00	160.00	30.00	3.86	7.89	4.28	2.55	6.97	5.06
3	120.00	442.00	180.00	112.45	282.54	0.94	110.00	288.00	5.00	2.45	5.46	4.06	1.74	4.42	3.09
4	45.00	605.00	180.00	56.75	438.32	4.77	54.00	442.00	3.00	2.75	3.68	1.77	1.35	4.08	2.98
5	-135.00	544.00	90.00	15.85	529.85	-81.30	18.00	522.00	-85.00	2.15	7.85	3.70	1.2	6.90	1.05
6	-40.00	390.00	270.00	-82.49	429.76	98.51	-78.00	434.00	98.00	4.49	4.24	0.51	3.10	5.70	1.51
7	-220.00	222.00	90.00	-125.83	245.82	-81.79	-128.00	240.00	-85.00	2.17	5.82	3.21	1.40	4.92	2.10

office navigation.

However, an omnidirectional camera that provides the single effective viewpoint geometry should allow us to extend the applications of the hybrid system beyond the artificial landmarks. This is a matter of ongoing development. Another direction of further research is the model of measurements uncertainty for the omnidirectional camera. Such a model should enable optimal fusion of the localization data from both cameras (e.g. by means of Kalman filtering), and more efficient planning of the positioning actions [27].

ACKNOWLEDGEMENTS

This work was supported by the Poznań University of Technology Faculty of Electrical Engineering grant DS-MK-141 in the year 2015.

AUTHORS

Marta Rostkowska* – Poznań University of Technology, Institute of Control and Information Engineering, ul. Piotrowo 3A, 60-965 Poznań, Poland,.
E-mail: marta.a.rostkowska@doctorate.put.poznan.pl
Piotr Skrzypczyński – Poznań University of Technology, Institute of Control and Information Engineering, ul. Piotrowo 3A, 60-965 Poznań, Poland.
E-mail: piotr.skrzypczyński @put.poznan.pl

*Corresponding author

REFERENCES

[1] Adorni G., Bolognini L., Cagnoni S., Mordonini M., *A Non-traditional Omnidirectional Vision System with Stereo Capabilities for Autonomous Robots*, LNCS 2175, Springer, Berlin, 2001, 344–355. DOI: 10.1007/3-540-45411-X_36.

[2] Bazin J., *Catadioptric Vision for Robotic Applications*, PhD Dissertation, Korea Advanced Institute of Science and Technology, Daejeon, 2010.

[3] Baczyk R., Kasinski A., "Visual simultaneous localisation and map–building supported by structured landmarks", *Int. Journal of Applied Mathematics and Computer* Science, vol. 20 no. 2, 2010, 281–293. DOI: 10.2478/amcs-2014-0043.

[4] Briggs A., Scharstein D., Braziunas D., Dima C., Wall P. , "Mobile Robot Navigation Using Self-Similar Landmarks". In: *Proc. IEEE Int. Conf. on Robotics and Automation*, San Francisco, 2000, 1428-1434, DOI: 10.1109/ROBOT.2000.844798.

[5] Cagnoni S., Mordonini M., Mussi L., "Hybrid Stereo Sensor with Omnidirectional Vision Capabilities: Overview and Calibration Procedures". In: *Proc. Int. Conf. on Image Analysis and Processing*, Modena, 2007, 99–104, DOI: 10.1109/ICIAP.2007.4362764.

[6] DeSouza G., A. C. Kak, "Vision for Mobile Robot Navigation: A Survey", *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. 24, no. 2, 2002, 237–267. DOI: 10.1109/34.982903.

[7] Durrant-Whyte H. F., Bailey T., "Simultaneous localization and mapping (Part I)", *IEEE Robotics & Automation Magazine*, vol. 13, no. 2, 2006, 99–108. DOI: 10.1109/MRA.2006.1638022.

[8] Davison A., Reid I., Molton N., Stasse O., "MonoSLAM: Real-time single camera SLAM", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, 2007, 1052–1067. DOI: 10.1109/TPAMI.2007.1049.

[9] Fiala M., "Designing highly reliable fiducial markers", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 7, 2010, 1317–1324. DOI: 10.1109/TPAMI.2009.146.

[10] Figat J., Kasprzak W., "NAO-mark vs. QR-code Recognition by NAO Robot Vision". In: *Progress in Automation, Robotics and Measuring Techniques*, vol. 2 Robotics, (R. Szewczyk *et al.*, eds.), AISC 351, Springer, Heidelberg, 2015, 55–64. DOI 10.1007/978-3-319-15847-1_6.

[11] Lemaire T., Berger C., Jung I.-K., Lacroix S., "Vision-based SLAM: Stereo and monocular approaches", *Int. Journal of Computer Vision*, vol. 74, no. 3, 2007, 343–364. DOI: 10.1007/s11263-007-0042-3.

[12] Lin G., Chen X., "A robot indoor position and orientation method based on 2D barcode landmark", *Journal of Computers*, vol. 6, no. 6, 2011, 1191–1197. DOI:10.4304/jcp.6.6.1191-1197.

[13] Lu F., Tian G., Zhou F., Xue Y., Song B., "Building an Intelligent Home Space for Service Robot Based on Multi-Pattern Information Model and Wireless Sensor Networks", *Intelligent Control and Automation*, vol. 3, no. 1, 2012, 90–97. DOI: 10.4236/ica.2012.31011.

[14] McCann E., Medvedev M., Brooks D., Saenko K., "Off the Grid: Self-Contained Landmarks for Improved Indoor Probabilistic Localization". In: *Proc. IEEE Int. Conf. on Technologies for Practical Robot Applications*, Woburn, 2013, 1–6. DOI: 0.1109/TePRA.2013.6556349.

[15] Martínez-Gomez J., Fernańdez-Caballero A., Garciá-Varea I. , Rodriǵuez L., Romero-Gonzalez C., "A Taxonomy of Vision Systems for Ground Mobile Robots", *Int. Journal of Advanced Robotic Systems*, vol. 11, 2014. DOI: 10.5772/58900.

[16] Menegatti E., Pagello E., "Cooperation between Omnidirectional Vision Agents and Perspective Vision Agents for Mobile Robots", *Intelligent Autonomous Systems 7* (M. Gini *et al.*, eds.), IOS Press, Amsterdam, 2002, 231–135, 2002.

[17] Potúcek I., *Omni-directional image processing for human detection and tracking*, PhD Dissertation, Brno University of Technology, Brno, 2006.

[18] Rahim N., Ayob M., Ismail A., Jamil S., "A comprehensive study of using 2D barcode for multi robot labelling and communication", *Int. Journal on Advanced Science Engineering Information Technology*, vol. 2, no. 1, 80–84, 1998.

[19] Rostkowska M., Topolski M., Skrzypczynski P., „A Modular Mobile Robot for Multi-Robot Applications", *Pomiary Automatyka Robotyka*, vol. 17, no. 2, 2013, 288–293.

[20] Rostkowska M., Topolski M., „Usability of matrix barcodes for mobile robots positioning", *Postępy*

Robotyki, Prace Naukowe Politechniki Warszawskiej, Elektronika (K. Tchon, C. Zielinski, eds.), vol. 194, no. 2, 2014, 711–720. (in Polish)

[21] Rostkowska M., Topolski M., "On the Application of QR Codes for Robust Self-Localization of Mobile Robots in Various Application Scenarios". In: *Progress in Automation, Robotics and Measuring Techniques*, (R. Szewczyk *et al.*, eds.), AISC, Springer, Zürich, 2013, 243-252. DOI 10.1007/978-3-319-15847-1_24.

[22] Rusdinar A., Kim J., Lee J., Kim S., "Implementation of real-time positioning system using extended Kalman filter and artificial landmarks on ceiling, Journal of Mechanical Science and Technology", vol. 26, no. 3, 2012, 949–958, DOI: 10.1007/s12206-011-1251-9.

[23] Scaramuzza D., *Omnidirectional vision: from calibration to robot motion estimation*, PhD Dissertation, ETH Zürich, 2008.

[24] Schmidt A., Kraft M., Fularz M., Domagala Z., "The comparison of point feature detectors and descriptors in the context of robot navigation", *Journal of Automation, Mobile Robotics & Intelligent Systems*, vol. 7, no. 1, 2013, 11–20.

[25] Siagian C., Itti L., "Biologically Inspired Mobile Robot Vision Localization", *IEEE Trans. on Robotics*, vol. 25, no. 4, 2009, 1552–3098, DOI: 10.1109/TRO.2009.2022424.

[26] Scharfenberger Ch. N., *Panoramic Vision for Automotive Applications: From Image Rectification to Ambiance Monitoring and Driver Body Height Estimation*, PhD Dissertation, Institute for Real-Time Computer Systems at the Munich University of Technology, Munich, 2010.

[27] Skrzypczynski P., "Uncertainty Models of the Vision Sensors in Mobile Robot Positioning", *Int. Journal of Applied Mathematics and Computer Science*, vol. 15, no. 1, 2005, 73–88.

[28] Skrzypczynski P., "Simultaneous Localization and Mapping: A Feature-Based Probabilistic Approach", *Int. Journal of Applied Mathematics and Computer Scienc*e, vol. 19, no. 4, 2009, 575–588, DOI: 10.2478/v10006-009-0045-z.

[29] Yoon K-J., Kweon I-S., "Artificial Landmark Tracking Based on the Color Histogram". In: *Proc. IEEE/RSJ Conf. on Intelligent Robots and Systems*, Maui, 2001, 1918-19203. DOI: 10.1109/IROS.2001.976354.

[30] *EmguCV*, http://www.emgu.com/wiki/index.php/Main

[31] *OpenCV Documentation*, http://docs.opencv.org

[32] *MessagingToolkit*, http://platform.twit88.com