Dragan Mićić and Danilo Tonić

Indoor Localization Using Fiducial Markers

In this paper, the possibilities of indoor based object localization using AruCo markers were examined. AruCo markers are synthetic square markers with high contrast edges. Together with the black frame, they represent a black and white 8×8 dimension matrix, where black or white fields determine the binary matrix. AruCo's black frame allows for quick detection, while its binary codification allows for its easy identification. The system consists of 8 AruCo markers placed on the walls surrounding 1×1 m space and a camera (Figure 2). The camera was located on an object the location of which in the set space was determined.

Localization is done in two basic steps (Figure 1). The first step involves the detection, segmentation and recognition of markers in camera images. After distortion removing, the image is binarized by the local adaptive threshold method which assigns a binary value to each pixel, depending on whether its value is greater than or less than the mean pixel value in the given environment. After binarization, the contour of objects in the binary image is separated using the Suzuki-Abe algorithm (Suzuki and Abe 1987). The next algorithm conducts the approximation of singular contours composed of pixels to polygons, where only quadruple contours remain candidates for the contour of the markers (Douglas and Peucker 1973). A customized 8×8 mesh is placed on the parts of the images within the selected quadruple contours, the fields of which, depending on whether the majority of the pixels within them have a value of one or zero, determine the binary code. The obtained binary codes are compared with the codes of the set markers. If they do not differ by more than a certain value from their most similar markers, the points of the detected quadruples will be paired with the coordinates of the most similar marker in the coordinate system set in the space.

After identifying the marker, the coordinates of the marker vertices on the image and the coordinates of them in the space are passed to the solvePnP algorithm. The aforementioned algorithm, using the known dimensions of the markers in the space and the angles under which the camera's optical center looks at the marker's vertices, determines the distance of the individual vertices from the camera by geometric equations. The position of the camera is obtained by triangulation of the mentioned distances. The solvePnP algorithm results in a translational vector and a rotation vector that transforms the camera's coordinate system into a coordinate system set up in the space.

A series of images were obtained on which the method was tested. The database consists of 288 images made from 36 points correctly arranged in space. At each point, 8 images are created for 8 different angles. The position and orientation of the camera is varied by a rotating base placed on a surface on which the coordinate system is drawn (Figure 1). Of the 288 im-

ages created, the markers are spotted on 224 of them. A total of 349 markers were observable, of which the marker detection algorithm successfully detected 277.

The position and rotation error is defined as the euclidean distance of the position obtained and the rotation of the camera from the reference values of the same. In the pictures where one marker is spotted, the mean position error is 8 cm and the mean rotation error is 8°. In the pictures where two markers are spotted, the mean position error is 4 cm and the mean rotation error is 4°, while in the images with three markers, the mean position error is 4 cm and the mean rotation error is 5°.

It is noted that localization is much more precise when two or more markers appear in the image. This is due to the very nature of the work of the solvePnP algorithm, which is due to the large number of data that is brought to its input.

The obtained results verify the described method and confirm the hypothesis that a larger number of markers in the image leads to more precise localization.