Recognition of lane markings in factories and self-position estimation method using AR markers

Kento Hisanaga, Shiyuan Yang, Siichi Serikawa

Department of Electrical and Electronic Engineering, Kyushu Institute of Technology, Kitakyushu, Japan

*Corresponding Author: hisanaga.kento426@mail.kyutech.jp

Abstract

In recent years, many unmanned transfer robots have been introduced in factories and warehouses. Functions such as self-position estimation are indispensable for freely operating automated guided vehicles. In this research, we estimated the self-position in the factory, proposed a robot control method using lane marking, and verified the measurement accuracy of the system in real time.

In this study, the camera image is used to read the AR marker and lane markings to calculate the distance between the camera and lane markings and estimate the self-position.

In this study, lane markings and AR markers are photographed horizontally with a camera. The distance is calculated from the position and tilt of the lane markings on the image. When the AR marker is detected, the camera is calibrated to calculate the distance and angle, and the self-position is estimated by comparing it with the actual coordinates.

As an experiment to measure the distance to the lane marking, the distance was calculated by gradually bringing the camera closer to the stationary camera with a thick paper with a thickness of 30 mm, which is likened to the lane marking. In the distance calculation, two experiments were conducted with the camera oriented horizontally and diagonally with respect to the lane marking.

As an experiment of self-position estimation using AR markers, we created a model like a passage in a factory, placed cameras at multiple points, and measured the error from theoretical values. As a method of expressing the self-position, I assigned the x-axis and y-axis to the model in the actual coordinate system and expressed it in two dimensions. In both experiments, in order to verify the accuracy, 100 continuous data were acquired at each point and the variability of the data was investigated.

Keywords: keyword1, keyword2, keyword3.

1. Introduction

In recent years, the delivery record of automated guided vehicle systems has been on the rise, which indicates that the demand for automated guided vehicle systems is increasing. [1]In the future, considering the decrease in workers and the problem of running costs, it is expected that attention will be paid to unmanned factories. [2]However, the current situation is that the spread of automated guided vehicle systems in Japan is lagging behind that in other countries. The reasons for this are that the initial cost for introducing the system is high, and that it is difficult to introduce the system because the passages in Japanese factories and warehouses are narrower than those in overseas factories and warehouses.

Most of the automated guided vehicle systems currently in use are low-priced but monotonous routes, or various routes can be set by installing various sensors, but the price is high. There are many things that end up. Ease and accuracy of route change are indispensable for automatic guided vehicles, and the decrease in efficiency caused by losing sight of the route is the most important problem to avoid. However, it is inevitable that the implementation will be costly due to the complexity of the system for the purpose of improving accuracy. Therefore, in this research, we are developing a method using a camera and an AR marker as a system that can be implemented at low cost and estimates the self-position of the transfer robot in the factory.[4] This time, we focused on real-time self-position estimation of this system and reduction of error.

2. calculate distance and detect self-position

Figure 1: An outline of system configuration in this study

Figure 2: Image captured from the camera

2.1 Summary of detection method

In this study, we use a camera to capture the lane markings and AR markers for measurement.

Figure 1 gives an overview of the system used in this study. Place the camera at the center of the lens on the floor or at a position of 180 [mm], aim the camera so that it is parallel to the floor, and capture the image. Figure 2 shows an image captured from the camera. As you can see from this figure, the lane markings and AR markers in the projected area are detected.

2.2 Preprocessing in image processing

Grayscale transformation, histogram flattening, masking, edge detection, and Hough transform were performed as preprocessing for lane marking detection. Grayscale conversion is performed to speed up image processing. Histogram flattening is done to make it easier to analyze the image. In this experiment, only the lower half of the image captured by the camera has the necessary

information, so the upper half is masked. Edge detection is performed to emphasize the lane markings for the Hough transform. The Hough transform is performed to extract only the dividing line from the image[3].

Camera calibration is performed as a preprocessing for estimating the self-position from the AR marker. Distortion of the camera image can be corrected by performing camera calibration.

2.3 Principle of calculating the distance from the lane marking after the lane marking is detected

This section describes the calculation principle for finding the distance from the lane marking from the image after lane marking detection.

By applying the image input from the camera to the processing flow shown in 2.2, an image as shown in Fig. 4 can be obtained if there is a lane marking in the camera field of view.

Figure 3: Example of image detection image

The obtained image has the x-axis in the horizontal direction and the y-axis in the vertical direction with the upper left as the origin, and is converted to the actual length based on the number of pixels(H_{p1} , H_{p2} , W_{p1} , W_{p2}) of the camera. Figure 4 below shows the relationship between the shooting range of the camera.

Figure 4: About the shooting range of the camera

If the distance z_0 from the lens where the angle of view is cut off when the height of the camera is H_c , it can be calculated from the image sensor size (H_i in the vertical direction, W_i in the horizontal direction) and can be obtained by the following equations (1). The maximum shooting range (H_{max} in the vertical direction and W_{max} in the horizontal direction) at a distance z away from the lens can be obtained by the following equations (2) and (3), respectively.

$$Z_0 = \frac{f \times H_c \times 2}{H_i} \quad (1)$$

$$H_{max} = \frac{z \times H_i}{f} \quad (2)$$

$$W_{max} = \frac{z \times W_i}{f} \quad (3)$$

Figure 5 below shows the positional relationship between the camera of this system and the ground.

Figure 5: Coordinate relationship between camera and ground

The camera is placed on the y-axis with the origin as the reference, and the camera is pointed in the z-axis direction to take a picture. When there is a lane marking in the shooting range, the starting point of the lane marking at the bottom on the screen is defined as $p_1(x_1, y_1)$, and p_2 as $p_2(x_2, y_1 + 20)$.

2.3.1 Calculation of the distance from the lane marking in the side direction of the camera

In this section, the distance from the lane marking in the side direction of the camera is calculated. If H_{z1} and H_{z2} are set as shown in Fig. 5, H_{z1} and H_{z2} can be obtained by equations (4) and (5).

$$H_{z1} = H_{max} \frac{PICTURE_HEIGHT - y_1}{PICTURE_HEIGHT} \quad (4)$$

$$H_{z2} = H_{max} \frac{PICTURE_HEIGHT - (y_1 + 20)}{PICTURE_HEIGHT} \quad (5)$$

The distance $Z1$ between the camera and the point (PICTURE_WIDTH / 2, y_1) and the distance between the point (PICTURE_WIDTH / 2, y_2) are calculated by Eqs. (6) and (7), respectively.

$$z_1 = -H_c \frac{z_0}{H_{z1} - H_c} \quad (6)$$

$$z_2 = -H_c \frac{z_0}{H_{z2} - H_c} \quad (7)$$

Next, let the number of pixels of the distance between the z-axis and the points p_1 and p_2 on the image be W_{p1} and W_{p2} , respectively. Furthermore, by combining Eqs. (8) and (9) with the obtained z_1 and z_2 , the maximum horizontal shooting range W_{max1} , and W_{max2} at each point can be obtained.

$$W_{max1} = \frac{z_1 \times W_i}{f} \quad (8)$$

$$W_{max2} = \frac{z_2 \times W_i}{f} \quad (9)$$

By substituting the values calculated in Eqs. (10) and (11) described below, the distances W_1 and W_2 between the z-axis and the points p_1 and p_2 in the real coordinate system can be obtained, respectively.

$$W_1 = W_{max1} \frac{W_{p1}}{PICTURE_WIDTH} \quad (10)$$

$$W_2 = W_{max2} \frac{W_{p2}}{PICTURE_WIDTH} \quad (11)$$

Next, Fig. 6 shows the coordinate relationship of each point when the camera is viewed from directly above.

Figure 6: Positional relationship when viewed from directly above the camera (1)

From this, the distance W from the lane marking in the side direction of the camera can be calculated by finding the intersection of the line connecting the points and the x -axis.

2.3.2 Calculation of the shortest distance between the camera and the marking line

This section shows the principle of calculating the shortest distance between the camera and the marking line. Figure 7 shows the coordinate relationship of each point when the camera is viewed from directly above.

Figure 7: Positional relationship when viewed from directly above the camera (2)

If the intersection of the straight line connecting points P1 and P2 and the z -axis is P3 (0, z_3), the distance z_3 can be obtained in the same way as z_1 and z_2 . Based on the obtained values, the shortest distance l between the camera and the marking line can be calculated by Eq. (12).

$$l = \frac{z_3 W}{\sqrt{z_3^2 + W^2}} \quad (12)$$

2.4 Principle of camera position estimation after AR marker detection

This section describes the calculation principle for finding the distance between the camera and the AR marker from the image after the AR marker is detected.

2.4.1 Camera calibration

In this study, camera calibration was performed and the parameters of the camera image sensor were estimated. The position of the camera is estimated by using the estimated parameters. Figure 8 shows the checkerboard used for calibration. In this study, in order to improve the accuracy of the parameters, the checkerboard was photographed while moving the camera, 40 sample images were acquired, and the corners of the checkerboard were detected from the images. The camera parameters and distortion parameters are calculated by comparing the coordinates of the corners read from the image with the known local coordinates. The camera parameters and distortion parameters are the values used when performing distortion correction.

When taken with a camera, the correction of radial distortion, which is actually a straight line but is displayed as a curved line, is as shown in Eqs. (13) and (14).

$$x_{corrected} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \quad (13)$$

$$y_{corrected} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \quad (14)$$

Similarly, the tangential distortion, which is the distortion caused by the lens and the image plane not being perfectly parallel, is corrected by Eqs. (15) and (16) below.

$$x_{corrected} = x + [2p_1 xy + p_2(r^2 + 2x^2)] \quad (15)$$

$$y_{corrected} = y + [p_1(r^2 + 2x^2)p_1 xy + 2p_2 xy] \quad (16)$$

(k_1, k_2, p_1, p_2, k_3) in this equation are distortion parameters, respectively.

Next, the internal and external parameters of the camera will be described. The internal parameters, which are camera-specific values, are expressed by the following equation (17).

$$\text{camera matrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3.17)$$

External parameters are rotation and translation parameters for converting the coordinates of a 3D point in one coordinate system to the coordinates in another coordinate system.

Figure 8: Checkerboard used

2.4.2 Camera position estimation from AR markers

In this study, we first estimated the distance of the camera with the center of the AR marker as the center of the 3D coordinate system. Figure 9 shows an image in which the read data is drawn with the 3D coordinate system for the image of the read AR marker.

Figure 9: AR marker that draws a 3D coordinate system

As shown in Fig. 10, in this study, the size and number of bits of the AR marker are set in advance, and the straight-ahead vector and rotation vector for the AR marker of the camera are set using the camera parameters and distortion parameters explained in the previous section. The position $c(x_c, y_c, z_c)$ of the camera in the three-dimensional coordinate system was estimated from the straight-ahead vector.

Figure 10: Position of AR marker and camera in 3D coordinate system

As a premise of this study, the distance l_c between the AR marker and the camera on the xy plane should be obtained, so the height h_c of the camera was set as a fixed value and calculated by Eq. (18).

$$l_c = \sqrt{(x_c^2 + y_c^2 + z_c^2) - h_c^2} \quad (18)$$

Next, find the positional relationship between the AR marker and the camera. The vector $l(x_c, y_c)$ was obtained from the obtained values of x_c and y_c , and the angle θ_c formed by the unit vector $n(1,0)$ in the x-axis direction was obtained from Eq. (19).

$$\cos\theta_c = \frac{x_c}{\sqrt{x_c^2 + y_c^2}} \quad (19)$$

The position of the camera was estimated from the AR marker by adjusting the distance to the camera.

2.5 About real-time system

In this experiment, the processing unit was a 64-bit operating system, the processor was (Intel (R) Core (TM) i5-9400F CPU @ 2.90GHz 2.90GHz), and the webcam BB-SW174WA (for video acquisition). Real-time processing was realized using the manufacturer: Panasonic). Regarding webcams, Fig. 11 below shows the cameras used for real-time operation of this system, and Table 1 shows each parameter. In this system, spider was used as the development environment and python was used as the language to process the video captured by the webcam.

Figure 11: Webcam BB-SW174WA

Table 1: Each parameter of the webcam used

Resolution	640×480[pixel]
Scanned area	3.52 mm (H) ×2.64 mm (V)
Focal length	1.95 mm

3. Experiment

3.1 Overview of lane marking accuracy verification

In this experiment, we created a model to make it easier to move the lane markings and verified the detection accuracy. Even if it is in a stationary state, the effect of light sources such as fluorescent lamps around it on each pixel value changes over time. Therefore, in this section, we verified the data scatter when a certain number of data were continuously acquired in real time when the camera and the lane marking were stationary, and the measurement accuracy of the detection system itself.

Figure 12: Schematic diagram of lane line detection accuracy verification

3.1.1 Calculation of distance from lane marking

Figure 3 shows a schematic diagram of this experiment.

In this experiment, the lane markings are brought closer to the stationary camera by 50 [mm] from a point 400 [mm] ahead when the camera is facing parallel to the lane markings, and the camera becomes the lane markings. On the other hand, when facing diagonally, approach from the point of 700 [mm] by 50 [mm] and record 100 consecutive data at each point. Scattering of recorded data and error from the true value. To make a comparison. Use thick paper with a thickness of 3 [cm] for the lane markings.

3.1.2 Experimental results

The table below summarizes the measurement errors and data scatter from the results of verifying the measurement accuracy of the camera in the stationary state.

Table 2: Measurement error and data scatter when cameras are parallel

theoretical value[mm]	Measured value from the side[mm]		Measured value of the shortest	
	Maximum error	max-min	Maximum error	max-min
400	4	0	4	0
350	18	27	18	26
300	4	0	4	0
250	8	0	8	0
200	11	0	11	0
150	12	0	12	0
100	8	0	8	0

Table 3: Measurement error and data scatter when the camera is tilted

theoretical value[mm]	Measured value from the side[mm]		Measured value of the shortest distance[mm]	
	Maximum error	max-min	Maximum error	max-min
700	8	0	5	0
650	16	0	7	0
600	2	0	11	0
550	10	0	4	0
500	6	0	8	0
450	0	0	3	0
400	13	0	9	0
350	8	0	3	0
300	15	0	12	0
250	11	0	27	35
200	16	0	2	0

From the above results, when the camera is horizontal to the lane marking, the distance can be measured with an error of at most 18 [mm] compared to the theoretical value, so practical lane marking detection is possible. It is considered possible. In addition, when the camera is horizontal to the lane marking, the distance can be measured with an error of at most 27 [mm] compared to the theoretical value, so practical lane marking detection is possible. it is conceivable that. There was almost no scattering of data when the minimum value was subtracted from the maximum value from the measurement results of each distance, and there was almost no scattering when shooting horizontally or at an angle, but to some extent at some distances. Sometimes scattered. This is considered to be an error caused by using thick paper with a thickness of 30 [mm] as the dividing line in this experiment

3.2 AR marker detection accuracy verification

In this experiment, self-position estimation was performed using a model. Even in the stationary state, the influence of

the surrounding light sources such as fluorescent lamps on each pixel value changes with time. Therefore, in this section, we verified the scattering of data when a certain number of data were continuously acquired in real time when the camera and AR marker were stationary, and the measurement accuracy of the self-position estimation system itself.

3.2.1 Experimental method

Figure 3 shows a schematic diagram of the model used in this experiment. In this experiment, the accuracy of the stationary model was verified by moving the camera to five arbitrary points. The orientation of the camera was verified in the direction that may be suitable for transportation in the actual factory. In this experiment, the lower left of Fig. 3 was set as (0,0) on the coordinates, the coordinates corresponding to the center of each marker were assigned, and the positional relationship with the camera was calculated from there to estimate the self-position. In verifying the accuracy of self-position estimation using AR markers, 100 consecutive data were recorded at each point. We will compare the scatter of the recorded data and the error from the true value. Use 50 [mm] x 50 [mm] thick paper for the AR marker.

Figure 13: Schematic diagram of a model using AR markers

3.2.2 Experimental result

The results obtained in this experiment are shown below.

Figure 14: Measurement results and camera position and orientation (500,0)

Figure 15: Measurement results and camera position and orientation (500,500)

Figure 16: Measurement results and camera position and orientation (500,1000)

Figure17: Measurement results and camera position and orientation (1000,1000)

Figure 18: Measurement results and camera position and orientation (2000,1000)

From the above results, regarding the self-position estimation by the AR marker, the maximum error is 16 [mm] in the x direction and 21 [mm] in the y direction, except for the points in the real coordinates (500,1000). Since the self-position can be estimated by, it is considered that practical obstacle detection is possible. For the points in real coordinates (500,1000), the maximum error was 1011 [mm] for the x-axis and 460 [mm] for the y-axis. It is considered that this is because the z axis flipping phenomenon that the z axis is inverted when the AR marker is read has occurred. This phenomenon occurs because it is not possible to judge from the image whether the AR marker is hanging from the ceiling or attached to the ground. This phenomenon is unlikely to occur when the distance to the AR marker is short, so we would like to devise ways to reduce false recognition as much as possible in the future.

4. Conclusions

In this research, in order to allow the robot to estimate its own position in the factory and travel freely to the target point, the approach is to use AR markers and lane markings, which are different from conventional automatic transfer robots, from the lane markings. The distance was measured and the self-position was estimated using the AR marker. The points that can be improved by this system compared to the conventional method are described below.

Since it is possible to recognize the lane markings originally drawn in the factory or warehouse and measure the distance, there is no need to draw a new magnetic tape on the path as in the line tracing method. Since this system uses image processing using a camera, the camera information can be used for other purposes such as obstacle detection. Unlike the conventional self-position estimation method using SLAM, it can be operated immediately by

simply installing an AR marker on the floor. Compared to installing a reflector as a landmark, AR markers can be mounted flexibly because there are few restrictions on the installation location. Unlike the method of installing the QR code on the floor, the self-position can be estimated even if it is far from the QR code, so even if there is an obstacle on the route, the self-position can be estimated after avoiding it. ..

References

- [1]. FNM Surname (2018). Article Title. Journal Title, 10(3), 1–10.
- [2]. F.N.M Surname, Article Title, <https://www.acm.org/publications/proceedings-template>.
- [3]. Shotatsu Keiji(2008), “[Details] Image processing programming” , SB creative, pp.57-98
- [4]. Ronald T. Azuma: “A survey of Augmented Reality”, Presence, Vol.6 No.4, pp.355-385, Aug.1997