IEEJ Journal of Industry Applications
Vol.8 No.6 pp.934–940 DOI: 10.1541/ieejjia.8.934

Paper

Feasibility Study of Vision-based Localization Method for Stopping Control Using Real Environment Data

Atsushi Oda[*a)] Member, Yuhi Tsutsumi[*] Non-member

(Manuscript received Feb. 4, 2019, revised May 14, 2019)

Stopping control is very important in public transportation systems. It is used to precisely match where a vehicle is boarded at a station and the positions of the doors of the vehicle. Although it is necessary to estimate self-localization with high accuracy, the cost of implementing it using balises, which is currently popular, is very high. To accelerate the introduction of stopping control to public transportation systems, a low-cost self-localization estimation technology is required. We propose a vision-based self-localization method for underground railway, which does not use expensive balises as that in conventional method. The method uses image recognition to calculate the relative distance to a ground target, whose position is known, and estimates the position of a vehicle from the absolute position of the target and the relative distance. In this paper, we evaluate the accuracy of the self-localization estimation using the proposed method under outdoor conditions that are more severe compared with underground conditions. From the results of six trials, we confirmed that the accuracy of position estimation of all trials satisfied the required accuracy of 0.8 m, when the distance between the vehicle and ground target was 20 m.

Keywords: stopping control, image recognition, localization

1. Introduction

In recent years, rapid progress has been made on methods for the automated driving of automobiles, and there is a movement to deploy automated driving technology in public transportation systems. In particular, automated driving systems can be easily installed in underground railways because the tracks are exclusive and there is less risk of obstacles. Since public transportation systems are easy for elderly people and users of wheelchairs and strollers to use, the importance of stopping control has been pointed out [(1)]. "Stopping control" controls a vehicle accurately to match the position where the vehicle is entered in the public transportation system and where a passenger waits. In public railways, platform screen doors are being installed to prevent people from falling from the platform and making contact with a train. Even when these doors are installed, stopping control is required to accurately align the opening of the platform screen doors and the doors of the vehicle. In stopping control for public transportation systems, an accuracy of within 1 m is generally required for the stopping position, and highly accurate methods for self-localization estimation are required to achieve this.

One such method for current public transportation systems is to use dedicated balises installed on routes. However, this leads to high costs due to civil engineering work as the balises need to be installed on the routes, and there are also maintenance costs. There is also technology for self-localization estimation that uses the global navigation satellite system (GNSS), but it cannot be introduced to public transportation systems running underground.

Therefore, the authors propose a self-localization estimation method that uses image recognition as a method that can be applied to underground routes with high accuracy and low cost.

In this paper, we verify the accuracy of the self-localization estimation of the proposed method in an actual railway environment and determine whether it is satisfactory.

2. Related Work

2.1 Stopping Control

"Stopping control" controls a vehicle along a target speed profile and stops the vehicle at a predetermined position. A system configuration for achieving this is shown in Fig. 1. The control is composed of three function blocks: estimating vehicle position, calculating target speed, and calculating brake command. The processes for each are as follows.

1) Estimating vehicle-position function: Estimate the vehicle position by integrating the vehicle speed. Since errors due to integral errors or slips are included, periodically correct the vehicle position on the basis of the position information from balises on the ground.
2) Calculating target-speed function: Calculate the target speed profile from the current vehicle position to the target stop point.
3) Calculating brake-command function: Calculate the brake command using proportional control or PID control so that the vehicle speed follows the target speed.

A conceptual diagram of stopping control is shown in Fig. 2. In conventional stopping control, it is common to correct the vehicle position with a balise. In this method, when

a) Correspondence to: Atsushi Oda. E-mail: atsushi.oda.gm@hitachi.com
* Hitachi, Ltd. Research & Development Group
832-2, Horiguchi, Hitachinaka, Ibaraki 312-0034, Japan

© 2019 The Institute of Electrical Engineers of Japan.

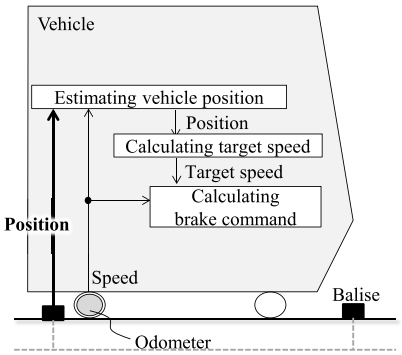


Fig. 1. System configuration of conventional stopping-control system

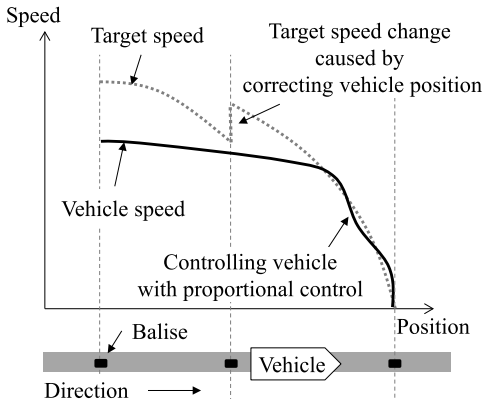


Fig. 2. Conceptual diagram of stopping control

a receiver installed in a vehicle passes over a balise placed on the ground, the stopping control function receives position information from the balise on the ground and updates the vehicle position. Balises are intermittently installed on the route, and the vehicle position between two balises is estimated by integrating the vehicle speed.

The accuracy of stop positions with stopping control is affected by the error in estimating the position of a vehicle and the dispersion of stop positions due to proportional control. As a result of simulating the accuracy of PID control under various vehicle conditions, it is reported that the dispersion of stop positions by PID control is a maximum of 0.2 m [2]. Even in an actual vehicle, the dispersion is within 0.2 m [3][4]. Therefore, if dispersion due to proportional control is defined as 0.2 m, to achieve a stopping accuracy of 1 m, the error in self-localization estimation needs to be 0.8 m or less. Therefore, in this paper, we define the required self-localization estimation accuracy as 0.8 m.

2.2 Self-localization Technologies A previous study on self-location estimation technology is shown in Table 1. Using balises (Table 1, No. 1) is mainstream in the railway field, and it is highly accurate in self-localization estimation. It has high environmental resistance in that the position accuracy does not decrease even in bad weather, such as fog and snow, or at nighttime. The balise method estimates a vehicle position from the integration of speed, and it corrects the position periodically with information from balises. However, as mentioned above, civil engineering work to install balises on the ground is necessary, and there is a problem in that the introduction cost is expensive. Using GNSS

Table 1. Comparison of localization technologies

No.	Method (Sensor type)	Localization accuracy [m]	Features
1	Balise (marker)	<0.1	High accuracy High performance under severe conditions
2	GNSS + gyro	<10	Low cost Abundant knowledge in automotive field Unavailable in tunnels
3	LIDAR	0.01–1.7	Accuracy is affected by environment Abundant knowledge in automotive field
4	Millimeter wave radar	0.5–1.0	Low cost Abundant knowledge in automotive field
5	Speed integration +gyro	<10	Low cost High performance under severe conditions
6	Camera (speed integration)	37 per km	Low cost

is a typical self-localization estimation method in addition to the balise method. GNSS is used to estimate the absolute position in two dimensions, and there is much research on self-localization estimation for public transportation systems with GNSS [5]–[9]. Although such estimation using GNSS is inexpensive, the positional accuracy can degrade due to multi-path, and it cannot be used underground. There are studies that remove the influence of multi-path [10] and fusion gyroscope sensor information on GNSS information when GNSS does not work well [11] (Table 1, No. 2), but the self-localization estimation accuracy is low.

Self-localization estimation technology using light detection and ranging (LIDAR) (Table 1, No. 3) is often used for autonomous cars and robots, and numerous research results have been reported [12]–[15]. There is also an example of estimating the self-localization of public transportation systems [16][17]. At present, it is a problem that the cost of LIDAR is high.

There is a method [18] (Table 1, No. 4) in which the speed is estimated with millimeter waves and self-localization is estimated from the integration of the speed. Although estimating speed with millimeter waves has an advantage in that it is less influenced by wheel slips and locks and the wheel diameter error unlike the general method of estimating a vehicle position by integrating vehicle speed, low estimation accuracy is a problem.

A method that uses the rotation speed of a wheel and an inertial sensor has been proposed (Table 1, No. 5). It [19] estimates a vehicle position mainly by using the rotation speed of a wheel and the speed estimated from the acceleration and angular velocity of the inertial sensor only when wheel slips and locks occur. However, the estimation accuracy is 10 m, which is too low to use in stopping control.

There is also research on estimating speed by using image recognition [20][21] (Table 1, No. 6). One method estimates the

travel distance between equivalent points of two images captured in a given period of time by image recognition, and it calculates the vehicle speed on the basis of information on the travel distance and the travel time. Studies [21] applied to public transportation systems have been reported, but the position estimation accuracy is 37 m per km, and the error is too large for application to stopping control.

2.3 Problems The target of this study is an underground railway, but GNSS is not available underground. Generally, the self-localization accuracy of LIDAR is very high, but the accuracy of LIDAR tends to be low under smooth wall conditions like tunnels [16].

3. Proposed Method

3.1 Concept of Proposed Method In this section, we describe self-localization estimation technology and the formulation of the relative distance calculation in the proposed method. We propose a method that uses relative distance information to ground targets estimated by image recognition for self-location estimation with high accuracy and low cost. This method estimates the self-localization by using both the absolute position of the stopping position target and the relative distance between the stopping position target and a vehicle. The targets are located at designed positions and that's positions are fixed, and the positions are known. Distance recognition technologies by image recognitions have been adopted for the relative distance calculation because there are abundant research cases in the field of automobiles [22]–[27]. There are also informative researches about image recognition technologies in railway system, e.g. [28].

Underground railways are favorable for image recognition because the illuminance of the environment is stabilized by lighting, and they are not affected by weather conditions such as rain.

Figure 3 shows a system configuration of stopping control using self-localization by image recognition. In the proposed method, the position information obtained by image recognition processing is used to correct the vehicle position by using the information from balises. No change occurs in the function units for estimating vehicle position, calculating target speed, and calculating brake command, except that the method of calculating the position information for correction is different. Therefore, the proposed method has an advantage in that it can be applied without greatly changing the software of the conventional stopping control. We can use existing stop target indicators for drivers to estimate the vehicle position, so there is no need to set new targets to the wayside.

Self-localization is estimated by integrating the speed until the self-localization estimation by image recognition is executed (until the stop target is recognized) as in the conventional method.

3.2 Relative Distance Calculation by Image Processing The method of measuring distance by image recognition will be described. In the proposed method, the relative distance between a stop target and a vehicle is estimated by using the area of the stop target detected by an image sensor. That is, the smaller the target area detected by the image sensor is, the longer the relative distance is. Figure 4 shows the relationship between the area of the target detected by an

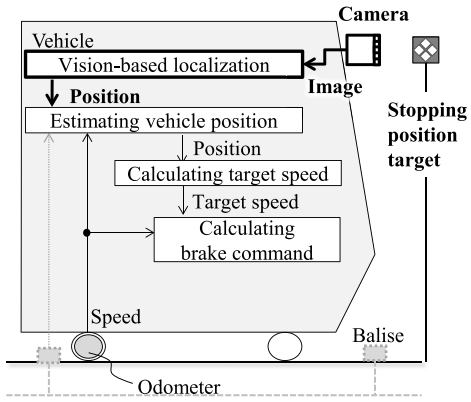


Fig. 3. System configuration of proposed stopping control system

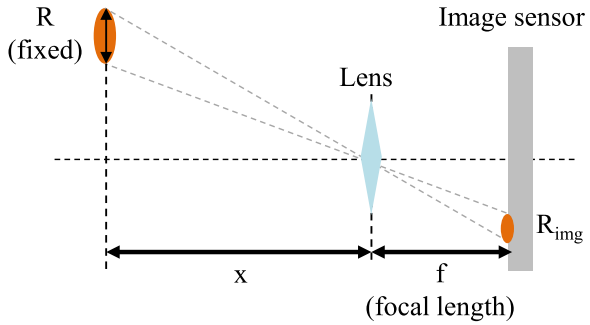


Fig. 4. Relation of relative distance x and radius of target area detected by image sensor

image sensor and the relative distance. The following relationship is derived from Fig. 4.

$$R_{img} = \frac{R \times f}{x} \quad \cdots\cdots (1)$$

Here, R is the length of one side of a stop target, f is the focal length, x is the relative distance, and R_{img} is the length of one side of the stop target detected by an image sensor. Since R and f in (1) are fixed, R_{img} at the relative distance x is proportional to $1/x$. Therefore, the relationship between R_{img} and relative distance x is defined by the following formula.

$$R_{img} = C_1\left(\frac{1}{x}\right) + C_2 \quad \cdots\cdots (2)$$

Here, C_1 and C_2 are coefficients. If the shape of a stop target is square, the area S_{img} of the stop target detected by the image sensor is calculated by (3).

$$S_{img} = {R_{img}}^2 \quad \cdots\cdots (3)$$

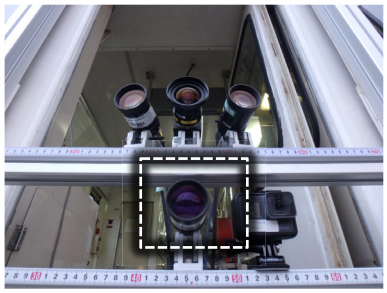
Equation (2) is substituted into (3) to obtain (4).

$$S_{img} = C_1^2\left(\frac{1}{x}\right)^2 + 2C_1C_2\left(\frac{1}{x}\right) + C_2^2 \quad \cdots\cdots (4)$$

The relative distance satisfying the area of the stop target detected by the image sensor is obtained by solving (5).

$$C_1^2\left(\frac{1}{x}\right)^2 + 2C_1C_2\left(\frac{1}{x}\right) + C_2^2 - S_{img} = 0 \quad \cdots\cdots (5)$$

(a) Test car



(b) Installation of camera

Fig. 5. Data acquisition was conducted using real-scale test car and camera was installed in front of test car

4. Data Acquisition for Evaluation

An image acquisition test was conducted to evaluate the self-localization estimation accuracy of the proposed method, and it was carried out under outdoor conditions that are more severe compared with underground conditions. In this chapter, we describe the test conditions.

4.1 Test Condition Image data acquisition tests done outdoors were carried out on railway lines in Hokkaido, Japan. The weather on the test day was sunny, and it was cloudy from time to time, and the illuminance changed periodically. The test was conducted seven times in total over 1 hour from 10 o'clock to 11 o'clock.

Figure 5 shows the test vehicle and the camera, which was installed outside the vehicle. The image size was 1280×960 pixels, the sensor size was 1/3-inch CCD, the frame rate was 15 fps, and the focal length was 30 mm.

4.2 Test Procedure The test was carried out with the following procedure.

1) The train was stopped 10, 20, 30, 40, and 50 m before the stop target, and an image of the stop target was acquired. We also measured the distance to the stop target (hereinafter referred to as the ground truth of relative distance) using the measurement at each stop.
2) Step 1 was carried out seven times in total. Image data acquisition tests done outdoors were carried out on railway lines in Hokkaido, Japan. The weather on the test day was sunny, and it was cloudy from time to time.

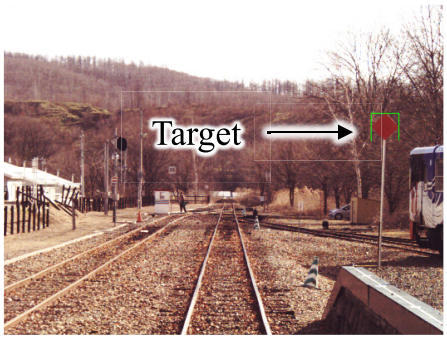
5. Evaluation

5.1 Target Extraction and Target Area Calculation The method of target extraction is as follows.

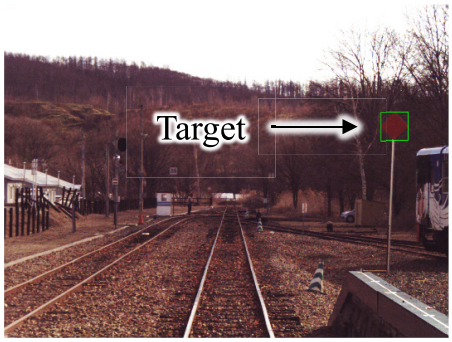
1) Converting image data from RGB to HSV.
2) Setting HSV value range.
3) Extracting pixels having same HSV value as set HSV value (H:332-360, S:31-100, V:0-100).
4) Creating mask image.
5) Converting mask image to grayscale image.
6) Converting grayscale image to binary image.
7) Using closing process.
8) Extracting contours of every white pixel cluster.
9) Calculating area of every cluster using its contour.
10) Extracting cluster of maximum area as target.

The calculating of the stop target area from the image data acquired in the test will be described. First, image data, which is in the RGB format, is converted into data in the HSV format, and a binary image, in which pixels having the same HSV value as the set HSV value are white and the others are black, is created. H:332-360 and S:31-100, V:0-100 are used in this paper. We use red as a color of the target because it is rare color against background of this field. For commercial use, we determine RGB value for target color to be easily separated from the background based on image data of a line where our technology would be installed. Since black noise was observed in the stop target, the closing process, in which black sesame noise is removed, was conducted on the binary image. In the closing process, two processes are conducted the same number of times: a process for replacing a pixel of interest with white if there is a white pixel around the pixel of interest (Dilation) and a process for replacing a pixel of interest with black if there is a black pixel around the pixel of interest (Erosion). In this paper, Dilation was carried out twice; then, Erosion was carried out twice. After the closing process, the largest white area in the binary image was extracted as the target detected by the image sensor, and the number of pixels of the target was measured and set as the area of the target detected by the image sensor. A recognized stop target is shown in Fig. 6. Image processing for seven trials was performed with the same parameters (HSV value, number of times of conducting closing process). Even when backlighting and illuminance changed as shown in Fig. 6, the stop target could be recognized with the same parameters of image recognition processing.

5.2 Derivation of Relative Distance Calculation Formula A quadratic approximation function of the test data was derived by using a Python function, the numpy.polyfit function, to identify the coefficient of (4), which is the relational expression of the target area detected by an image sensor and the relative distance, from the area calculated by using the method that is mentioned in section 5.1 and the ground truth of the relative distance when the image data was recorded. Only the data of the first out of seven tests was used to derive the approximate function. Figure 7 shows the relationship between the area of the stop target in the images (vertical axis) and the reciprocal of the ground truth of the relative distance (horizontal axis) when the images were recorded and the quadratic approximation function. According to equation (4), the relationship between the area of the stop target and the reciprocal of the relative distance is approximately a quadratic relation, and it was confirmed that the test data was expressed by the approximate expression. From the above results, (6) was obtained as an approximate relational expression of the area and the relative distance of the stop target in the images under the current test conditions.



(a) Clear sky condition



(b) Cloudy sky condition

Fig. 6. Result of target detection by image recognition under various conditions

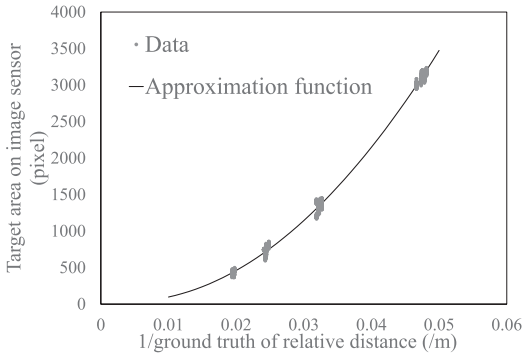


Fig. 7. Approximation function between target area and reciprocal of ground truth of relative distance

$$S_{img} = 1604558\left(\frac{1}{x}\right)^2 - 11834\left(\frac{1}{x}\right) + 56 \cdots\cdots\cdots (6)$$

The solution obtained by substituting the area S_{img} obtained by image recognition into (6) is the estimated relative distance, and (7), which is an estimated relative distance calculation formula, is obtained from the solution formula.

$$x = \frac{2 \times 1604558}{11834 + \sqrt{11834^2 - 6418232\left(56 - S_{img}\right)}} \cdots\cdots\cdots\cdots (7)$$

5.3 Evaluation of Self-localization Accuracy The accuracy of self-localization estimation in the proposed method was evaluated. Under the test conditions, the stop target was out of the frame of the image 10 m before the stop target, so the data 10 m before the stop target was excluded from the evaluation. Also, we evaluated data during an arbitrary 10 seconds at each stop position among the data for the second to the seventh trial except for the first test used for the calculation of (6). In this paper, the accuracy of self-localization estimation is evaluated as a deviation between the estimated relative distance up to the stop target and ground truth of relative distance. Therefore, the estimation accuracy ε is expressed by (8).

$$\varepsilon = |x - X_{real}| \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (8)$$

Here, x is the estimated relative distance, and X_{real} is the ground truth when the image data is acquired. The results are shown in Tables 2 and 3. The maximum absolute value of deviation was about 4 m at 50 m before the target position. It was confirmed that the deviation became smaller as the distance to the stop target shortened, and the maximum absolute value of deviation was about 0.4 m at 20 m before the target position. At 20 m before the target position, the estimation accuracies of all test data satisfied the required accuracy of self-localization estimation, 0.8 m. The root mean squared error (RMSE) also decreased as the vehicle approached the stopping target position, and the maximum RMSE was 0.3 m at the stop position of 20 m. It was confirmed that the proposed method has high accuracy in self-localization estimation for stopping control.

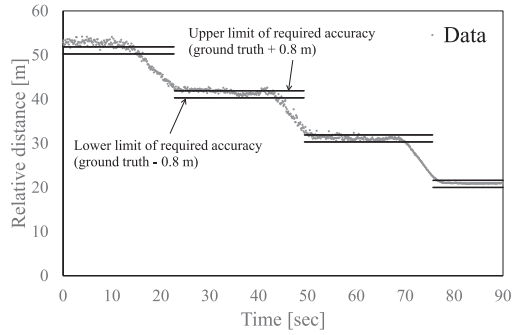
Table 2. Maximum absolute value of deviation of proposed method [m]

No.	20 m	30 m	40 m	50 m
2	0.23	1.26	1.58	3.07
3	0.16	1.21	1.48	2.60
4	0.33	0.87	2.30	2.41
5	0.33	0.82	1.45	2.22
6	0.28	1.33	2.42	0.99
7	0.41	1.70	3.58	4.47

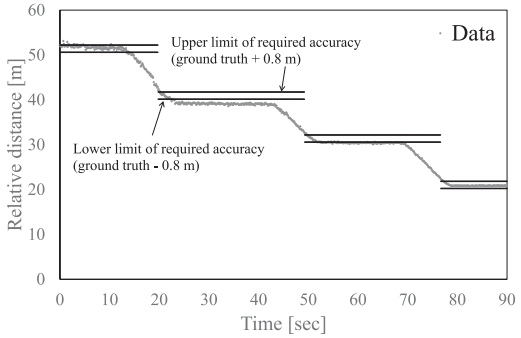
Table 3. RMSE of proposed method [m]

No.	20 m	30 m	40 m	50 m
2	0.10	0.41	0.89	1.89
3	0.07	0.28	0.53	0.98
4	0.14	0.38	1.42	1.00
5	0.13	0.58	1.09	1.39
6	0.19	0.99	1.88	0.44
7	0.25	1.30	2.50	3.03

As a comparison with the proposed method, a method of estimating the self-localization of automobiles by using road direction signs whose positions are known is used [29]. In this method, the relative positions between road signs and a car are derived by matching the feature of the camera image of the road sign and the feature of the corresponding database image using the Oriented FAST and Rotated BRIEF method. The position of an automobile is estimated from the Universal Transverse Mercator (UTM) coordinates of a road sign and the relative position between the road sign and the car. When the relative distance to the sign was 17 m, the RMSE was 0.36 m, and when the distance was 33 m, the RMSE was 2.12 m. Since the relative distances were not exactly the



(a) Trial #2



(b) Trial #6

Fig. 8. Estimation result example: comparison between estimated relative distance and required accuracy. Estimated relative distances meet required accuracy when the relative distance to the sign was 20 m at each trial

same, we cannot compare them directly, but the RMSE of our proposed method was 0.25 m maximum at a relative distance from the target of 20 m and 1.3 m when it was 30 m. It seems that our proposed method performs well compared with the method proposed in paper [29].

Figure 8 shows a comparison between the estimated relative distance and required accuracy. The deviation in the estimated relative distance was relatively large far from the target position, so it does not meet the required accuracy. The deviation converged and satisfied the requirement as the vehicle approached the stopping target position. The estimated relative distance was calculated smoothly even when the vehicle was moving to the next position.

5.4 Considerations From the test results, the accuracy of the self-localization tended to decline (the dispersion became larger) as the distance from the stop target became longer. This is considered to be caused by the fact that the sensitivity of the relative distance to the area in an image is large at a distance, so the relative distance greatly changes even when there are slight area fluctuations due to noise. Under the test conditions in this case, it was found that it was necessary to approach the stop target to about 20 m to stably estimate the vehicle position. When the self-localization is corrected 20 m before the stop target position, there is a possibility that the proportional control of the stopping control may not control the vehicle to the target position accurately. Therefore, when the proposed method is applied to the stopping control of public transportation, it is desirable to install multiple targets on the ground far from the target stopping position so that the stopping control can stop the vehicle accurately.

Alternatively, it is also conceivable to adjust to the focal length where the stable self-localization is obtained so that stopping control can stop the vehicle accurately.

In this study, a simple color sign was used as a stop target, but it is also possible to use a two-dimensional code as a stop target. If the information on the absolute positions of stop targets were recorded in the two-dimensional code, a database of this information would be unnecessary, which would make it easy to introduce stopping control to underground railways.

6. Conclusion

We proposed vision-based localization using image recognition as a method for self-localization with high accuracy and low cost. The accuracy of the self-localization estimation of the proposed method was evaluated by using image data acquired with a test line under outdoor conditions. When the distance from the stop target was longer than 20 m, there was a trial that did not satisfy the required self-localization estimation accuracy for the stopping control of public transportation systems, that is, 0.8 m. We confirmed that, at 20 m before the target position, the estimation accuracies of all test data achieved the required accuracy. In the future, we will carry out a test of stopping control using an actual vehicle that incorporates the proposed method. These results stated above are only for the daytime condition. Therefore we need more verifications about our technology on various severe conditions like night, rain, fog, etc. to apply it to commercial uses.

References

(1) M. Kawamoto: “Next generation urban transport system”, Journal of Society of Automotive Engineers of Japan, Vol.69, No.12, pp.58–61 (2015)
(2) D. Chen, R. Chen, Y. Li, and T. Tang: “Online learning algorithms for train automatic stop control using precise location data of balises”, IEEE Trans. Intelligent Transportation Systems, Vol.14, No.3, pp.1526–1535 (2013)
(3) H. Ando, M. Tanaka, and T. Narita: “TASC device for 1000 series Aonami Line”, [Translated from Japanese], in Proc. J. Cybernet, No.519 (2005)
(4) T. Negishi, M. Takagi, J. Ohya, S. Iba, and H. Nakazawa: “Development of ATO aimed at energy saving and improvement of stopping accuracy”, [Translated from Japanese], in Proc. J. Cybernet, No.513 (2007)
(5) H. Yamamoto: “Application of GPS to railway system [Translated from Japanese]”, Railway Research Review, Vol.67, No.2, pp.30–31 (2010)
(6) L. Jiang, C. Bai-gen, and W. Jian: “Track-constrained GNSS/odometer-based train localization using a particle filter”, in Proc. IEEE Intelligent Vehicles Symposium, pp.877–882 (2016)
(7) T. Mizuma, J. Yoshinaga, and N. Kudo: “Development of new traffic control and signaling system with using GPS”, Report of N.T.S.E.L., No.11 (2007)
(8) A. Albanese and L. Marradi: “The integrity performances of GNSS-based railway user navigation equipment”, in Proc. the 2005 ASME/IEEE Joint Rail Conference, pp.211–218 (2005)
(9) O. Gebauer, W. Pree, and B. Stadlmann: “Autonomously driving trains on open tracks - concepts, system architecture and implementation aspects”, unpublished.
(10) N. Sünderhauf, M. Obst, G. Wanielik, and P. Protzel: “Multipath mitigation in GNSS-based localization using robust optimization”, in Proc. Intelligent Vehicles Symposium, pp.784–789 (2012)
(11) Y. Sakaguchi: “Cooperation technology of a person and a car in the automated driving” [Translated from Japanese], presented at JSAE Symposium, Tokyo, Japan (2016)
(12) E. Muramatsu: “LiDAR technology realizes advanced automated driving”, Journal of Society of Automotive Engineers of Japan, Vol.71, No.3, pp.37–42 (2017)
(13) K. Lingemann, H. Surmann, A. Nuchter, and J. Hertzberg: “Indoor and outdoor localization for fast mobile robots”, in Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.2185–2190 (2004)
(14) R. Zlot and M. Bosse: “Place recognition using keypoint similarities in 2D lidar maps”, Springer Tracts in Advanced Robotics, Vol.54, pp.363–372 (2009)

(15) J. Zhang and S. Singh: "Lidar odometry and mapping in real-time", in Proc. Robotics: Science and Systems Conference (2014)
(16) T. Daoust, F. Pomerleay, and T.D. Barfoot: "Light at the end of the tunnel: High-speed lidar-based train localization in challenging underground environments", in Proc. Conference on Computer and Robot Vision, pp.93–100 (2016)
(17) D. Stein, M. Spindler, J. Kuper, and M. Lauer: "Rail detection using lidar sensors", International Journal of Sustainable Development and Planning, Vol.11, No.1, pp.65–78 (2016)
(18) M. Fukuda, H. Sugawara, A. Gion, Y. Ono, T. Kitano, and T. Shirai: "Development of train location detection methods for signalling", Railway Technical Research Institute Report, Vol.27, No.9, pp.17–22 (2013)
(19) K. Iwata, T. Yamamichi, Y. Sugiyama, T. Kitano, A. Asano, and Y. Hirao: "A distance calculation method by combining inertial sensors and tachometer generators for communications-based train control systems", IEEJ Transactions on Industry Applications, Vol.136, No.12, pp.915–924 (2016)
(20) H. Lategahn and C. Stiller: "Vision-only localization", IEEE Transactions on Intelligent Transportation Systems, Vol.15, No.3, pp.1246–1257 (2014)
(21) B. Bah, E. Jungabel, M. Kowalska, C. Leithäuser, A. Pandey, and C. Vogel: "Odometry for train location", [Online]. Available: https://www.win.tue.nl/casa/meetings/special/ecmi08/odometry.pdf (2009)
(22) T. Shima, H. Matono, S. Kakegawa, and T. Monji: "Image recognition for on-board cameras and distance measurement using stereo camera", Kogaku, Vol.41, No.5, pp.287–291 (2012)
(23) S. Takimoto and T. Ito: "Development of range-finding technology using on-board mono-camera", [Translated from Japanese], SEI Technical Review, Vol.169, pp.82–87 (2006)
(24) L. Bo, X. Zhang, and M. Sato: "Pitch angle estimation using a vehicle mounted camera for vehicle target range measurement", The Journal of the Institute of Image Information and Television Engineers, Vol.69, No.4, pp.169–176 (2015)
(25) G. P. Stein, O. Mano, and A. Shashua: "Vision-based ACC with a single camera: bounds on range and range rate accuracy", in Proc. Intelligent Vehicles Symposium, pp.120–125 (2003)
(26) Z. Sadreddini, T. Çavdar, and H.B. Jond: "A distance measurement method using single camera for indoor environments", in Proc. 39th International Conference on Telecommunications and Signal Processing, pp.462–465 (2016)
(27) M. Milkhalsky and J. Sitte: "Active motion vision for distance estimation in autonomous mobile robots", in Proc. Conference on Robotics, Automation and Mechatronics, pp.1060–1065 (2004)
(28) N. Nagamine and M. Ukai: "A Visibility Check Method for Obstruction Warning Signals by Cab-mounted Camera", Railway Technical Research Institute REPORT, Vol.26, No.7, pp.23–28 (2012)
(29) C. Jang, J.E. Park, and Y. Kim: "Compensating drift of mono-visual odometry using road direction sign database", in Proc. IEEE 20th International Conference on Intelligent Transportation Systems (2017)

Atsushi Oda (Member) received M.S. degrees in engineering from Osaka University, Osaka, Japan, in 2008, and received Ph.D. degree in engineering at the University of Tsukuba, Ibaraki, Japan, in 2019. His research interests include railway systems, especially automatic train operation (ATO) systems and driver assistance systems (DASs) using sensor technologies, mathematical optimization of speed profiles for eco-friendly ATO systems, and auto parameter tuning ATO systems.

Yuhi Tsutsumi (Non-member) received the M.S. degrees in engineering from the department of electrical engineering and bioscience from Waseda University, Japan, in 2015. His research interests include railway systems, especially controlling trains for punctuality and energy-saving, stopping accuracy, and operating trains with sensor systems.