Proceedings of the Seventh International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 15, Paper 24.3
Civil-Comp Press, Edinburgh, United Kingdom, 2026
ISSN: 2753-3239, doi: 10.4203/ccc.15.24.3

Experimental Validation of Onboard Train Localization Using Marker Detection and Accelerometer

K. Nagai[1], W. Ohnishi[1], T. Koseki[1], Y. Setoguchi[2], T. Yamaguchi[2], S. Morita[2] and K. Tanaka[2]

[1] **Graduate School of Engineering, The University of Tokyo, Japan**
[2] **NIPPON SIGNAL CO., LTD., Tokyo, Japan**

Abstract

Train position and velocity are essential for preventing collisions and maintaining appropriate train intervals. This study proposes an onboard train position and velocity estimation method combining a camera and an accelerometer. In the proposed method, trackside gradient markers are detected through image recognition and used to correct the estimated position and compensate for gradient effects in the measured acceleration. This improves the accuracy of both position and velocity estimation. In addition, divergence in the estimated states is reduced by correcting the position, velocity, and acceleration during train stops. The proposed method was validated using data collected in field tests on an actual train. The results show that the root mean square error of the velocity estimation was less than 2 km/h, while that of the position estimation was approximately 10 m. These findings indicate that the proposed method can estimate train position and velocity without additional trackside equipment and can be applied effectively using data from actual train operations.

Keywords: train localization, velocity estimation, image recognition, inertial measurement unit, gravity compensation, marker detection

1 Introduction

Train position and velocity are essential for preventing train collisions and maintaining appropriate train intervals. Currently, train velocity is primarily measured with a tachometer generator (TG). These sensors are mounted on axles and estimate velocity and distance traveled by detecting wheel rotation; however, their accuracy can be affected by wheel slip and slide [1]. For position detection, commonly used methods combine TG with absolute positioning sensors, such as balises or track circuits [2]. However, these trackside facilities require substantial maintenance labor and cost. Therefore, methods for estimating train position and velocity, mainly using onboard sensors, are needed.

In previous studies, absolute positioning has often been performed using Global Navigation Satellite System (GNSS) or Light Detection And Ranging (LiDAR), with an Inertial Measurement Unit (IMU) used to supplement these measurements [3]. However, GNSS is unavailable in tunnel sections, and its accuracy decreases in urban canyons and forested areas [4]. LiDAR also has limitations in environments with few distinctive surrounding structures [5] and under adverse weather conditions [6]. In such conditions, velocity and position are mainly estimated by integrating acceleration measurements. However, integration accumulates errors in the estimated velocity and position, leading to divergence during long-term operation [7]. Therefore, effective use of IMUs is essential for bridging the gaps between absolute position updates.

When using an IMU, random noise in the acceleration sensor can be partially smoothed by integration. In contrast, the effect of gravitational acceleration is critical, making accurate attitude estimation essential. Many studies have investigated attitude estimation from inertial sensor measurements using methods such as complementary filters [8] and Kalman filters (KFs) [9]. However, the fundamental instability of attitude estimation based solely on an IMU remains unresolved. In other words, the accuracy of gravity compensation decreases over time as long as we use only IMUs for attitude estimation. Therefore, it is unavoidable to add some positioning sensors or known attitude information.

In this paper, we validate a method that uses image recognition to simultaneously address two issues: the influence of gravity on acceleration measurements and the divergence of position and velocity estimates. While our previous study [10] focused mainly on velocity estimation, the present paper also examines position estimation. In addition, we refine the estimation algorithm and validate the method using a larger set of experimental data.

The outline is as follows: In Section 2, the proposed sensor fusion method is described. In Section 3, the experimental results are prsented. In Section 4, conclusions and future directions are discussed.

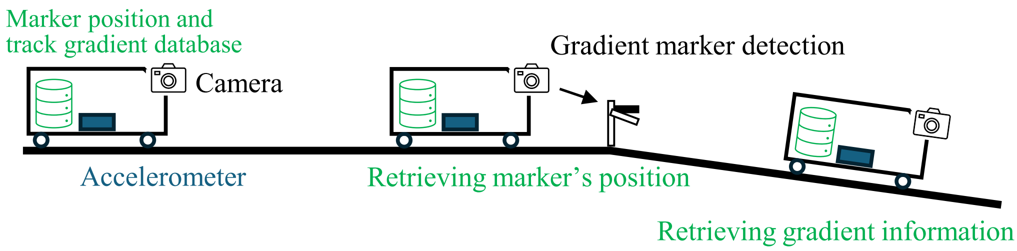

Figure 1: Conceptual overview of the proposed position and velocity estimation method.

2 Methodology

2.1 Basic concept

This section describes the proposed method for estimating position and velocity by combining gradient marker detection based on image recognition with acceleration measurements. As shown in Fig. 1, the train is equipped with a camera for image recognition and an accelerometer. Velocity is obtained by integrating the measured acceleration, and position is then obtained by integrating the velocity. Because position estimation errors accumulate during these integration processes, position correction is performed by combining gradient marker detection with a database of gradient marker positions.

In addition, the acceleration measured in the train's traveling direction includes a gravitational component that depends on the gradient of the track. Therefore, the true acceleration of the train is calculated by retrieving gradient information from a database that stores the gradient at each position along the route and compensating for the effect of gravity. This process improves the accuracy of velocity and position estimation. The overall processing flow is shown in the flowchart in Fig. 2.

2.2 Velocity and position estimation by integrating acceleration

The measured acceleration includes a component related to gravitational acceleration g, which depends on the track's gradient. On Earth, a condition in which the measured acceleration is zero along all three axes corresponds to free fall. In other words, when an accelerometer is stationary on the ground, it measures an acceleration with the same magnitude as gravitational acceleration but in the opposite direction to gravity. To clarify how this affects acceleration measurements in a train, consider a train that is stopped by braking on a track with an inclination angle of θ. The forces acting on the train in this situation are shown in Fig. 3. For better understanding, we add direction information x and z to a_{train} and a_{IMU} only in this subsection. The equation of motion along the slope is given by

$$m \cdot 0 = f_b - mg\sin\theta, \tag{1}$$

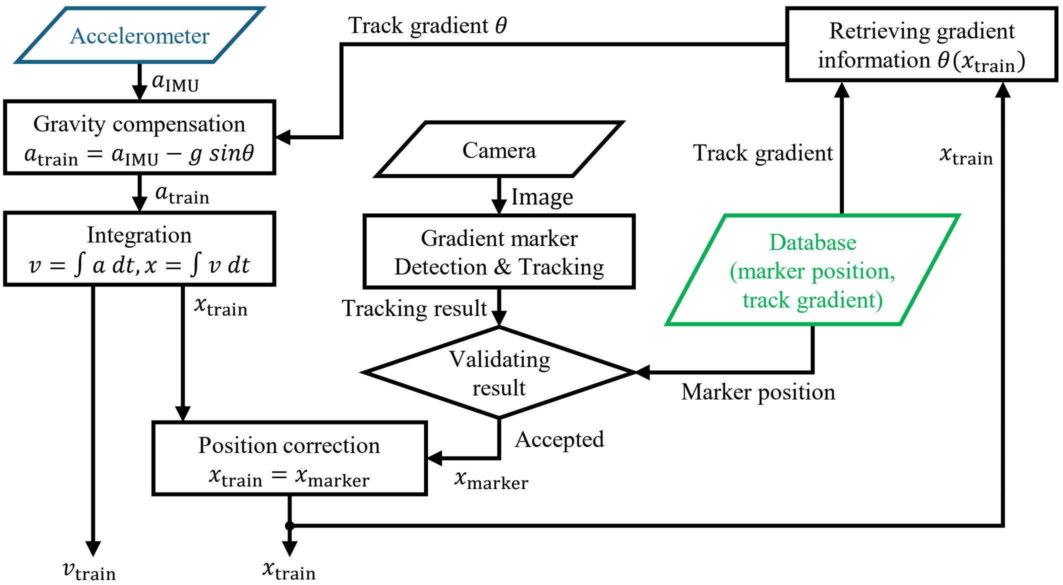

Figure 2: The flowchart of the proposed method.

and thus the true acceleration of the train, a^x_{train}, is zero. However, from the perspective of the accelerometer mounted on the train body, the train is stationary on the inclined track, resisting free fall. Consequently, the accelerometer outputs the measured acceleration in the longitudinal direction of the train as

$$a^x_{\text{IMU}} = \frac{f_b}{m}. \tag{2}$$

From the above equation of motion,

$$\frac{f_b}{m} = g \sin\theta. \tag{3}$$

Therefore, the true longitudinal acceleration of the train, a^x_{train}, can be obtained by subtracting $g\sin\theta$ from the acceleration measured by the accelerometer, a^x_{IMU}, as follows:

$$a^x_{\text{train}} = a^x_{\text{IMU}} - g\sin\theta. \tag{4}$$

After this, a^x_{train} and a^x_{IMU} are denoted as a_{train} and a_{IMU} for simplicity.

If a database contains the track gradient angle θ at each position along the route, θ can be retrieved based on the current estimated position of the train x_{train}. This allows a_{train} to be calculated. By integrating this corrected acceleration, the train velocity v_{train} and position x_{train} can be estimated. Let T_s denote the sampling period of the sensor. Then, at time step k, $v_{\text{train}}(k)$ and $x_{\text{train}}(k)$ are calculated as

$$v_{\text{train}}(k) = v_{\text{train}}(k-1) + a_{\text{train}}(k-1)T_s, \tag{5}$$

$$x_{\text{train}}(k) = x_{\text{train}}(k-1) + v_{\text{train}}(k-1)T_s + \frac{1}{2}a_{\text{train}}(k-1)T_s^2. \tag{6}$$

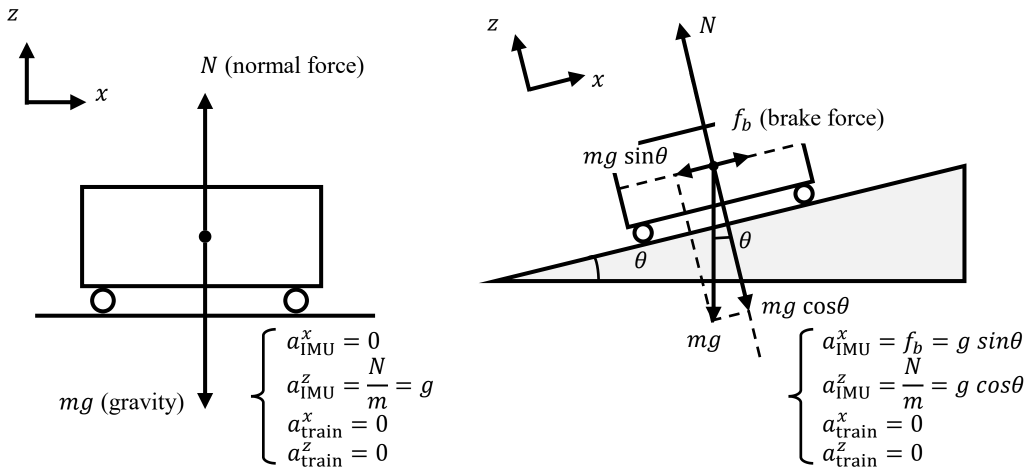

Figure 3: The effect of gravity on an accelerometer. The left figure is on the level section, whereas the right figure shows the situation in the slope area. Both trains are stopping.

It should be noted that detecting whether the train is stopped is relatively straightforward. For example, stop conditions can be detected using information such as notch and brake command signals, door opening and door closing status, or optical flow calculated from camera images. Once the train is detected as stopped, divergence of the estimated states can be prevented by setting $v_{\mathrm{train}} = 0$. This approach is commonly referred to as a zero velocity update (ZUPT) and is widely used in motion estimation research [11]. In addition, because trains repeatedly run on the same track, position correction can also be performed easily while the train is stopped at a station. By using highly accurate position information x_{truth} obtained from other sensors or systems and setting $x_{\mathrm{train}} = x_{\mathrm{truth}}$, divergence of the position estimation can be prevented. This study also adopts these approaches.

2.3 Gradient marker detection by image recognition

The loop in which velocity and position are calculated by integrating acceleration, and in which the gradient angle θ is retrieved from the database based on the estimated position and used for acceleration compensation, is inherently unstable. As positioning errors accumulate during long distance operation, the accuracy of the estimated position decreases. Consequently, the retrieved gradient value may become inaccurate, degrading gravity compensation for acceleration and causing errors in velocity and position estimation to increase rapidly.

To prevent divergence in the estimation process, it is important to periodically obtain highly accurate position information. In particular, accurately detecting locations where the track gradient changes improves gravity compensation. Therefore, in the proposed method, position correction is performed by detecting gradient markers in-

Figure 4: Typical shape and placement of gradient marker in Japan.

stalled at gradient change points using image recognition. In Japanese railways, placing markers where the track gradient changes is required by ministerial ordinance [12]. As shown in Fig. 4, this gradient marker has a uniform shape. This means the proposed method requires no additional trackside equipment compared with current railway operations. This basic concept has already been presented [10]. In this paper, to further improve the practical value of the proposed method, tracking across multiple frames is adopted rather than single frame image recognition. In addition, sensor position offsets are considered to improve estimation accuracy.

In image recognition, tracking refers to the process of following the same object across multiple frames. Because a unique ID is assigned to each tracked object, the object can be considered the same structure as long as its ID remains unchanged. In the proposed method, when a gradient sign with an assigned tracking ID disappears from the image after satisfying specific conditions, the train is judged to have passed the position of the gradient sign. These conditions include the detection duration, the confidence score, the position of the detected object in the image at the time of disappearance, and the consistency between the current estimated position and the gradient marker position. The following paragraphs describe the criteria for each condition and its purpose.

First, the detection duration and the average confidence score are checked to see whether they exceed the thresholds T_{th} and P_{th}, respectively. Detection results with a short duration or a low confidence score are likely to be false detections and should therefore not be used. T_{th} and P_{th} are experimentally decided. In general, these thresholds depend on the image recognition model's performance and the visibility around the markers.

Next, the position in the image at the time of disappearance is considered. Because gradient signs are installed beside the track, the gradient sign should be located near either the left or right edge of the image in the last frame in which it is detected. Therefore, the detection result is accepted when the left or right edge of the bounding box is located within the left side or right side region of the image width, as defined by the threshold A_{th}. The detection is accepted when $w_{\mathrm{left}} < A_{\mathrm{th}}W$ or $w_{\mathrm{right}} > (1 - A_{\mathrm{th}})W$, where W is the image width, w_{left} is the left coordinate of the bounding box, and w_{right} is that of right side. The concept of this process is illustrated in Fig. 5.

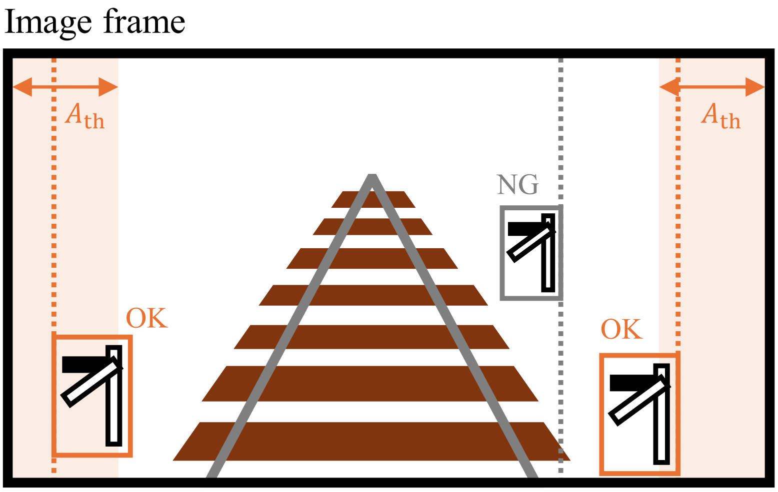

Figure 5: Checking the marker positions in the image at the last frame of marker detection.

Apparently, A_{th} should not contain track and sleeper regions. A_{th} also depends on the field of view of a camera.

Finally, the consistency between the current estimated position and the gradient marker position is verified. The current train position x_{train} at the time when the gradient sign is detected is compared with the nearest gradient marker position x_{marker} retrieved from the database. If the difference between these two positions is within the threshold X_{th}, the estimated position is corrected by overwriting it as

$$x_{\text{train}} = x_{\text{marker}} - o \times d_{\text{cam}}. \tag{7}$$

Note that o is the orientation of the train run, and d_{cam} is the longitudinal offset between the camera and the accelerometer. When the train is running in the outbound direction, $o = 1$. If the train is running in the inbound direction, $o = -1$. X_{th} is related to the accumulated error of integration and the distance between a gradient marker and another one. Through the above sequence of processes, image recognition results can be used to estimate position while reducing the influence of false detections.

3 Experimental validation

3.1 Experimental setup

This section validates the proposed method using data measured on an actual railway vehicle equipped with cameras and an accelerometer. The vehicle consists of three cars, and the sensor arrangement is shown in Fig. 6. A camera is installed on each leading car, while the accelerometer is installed on only one. The accelerometer's

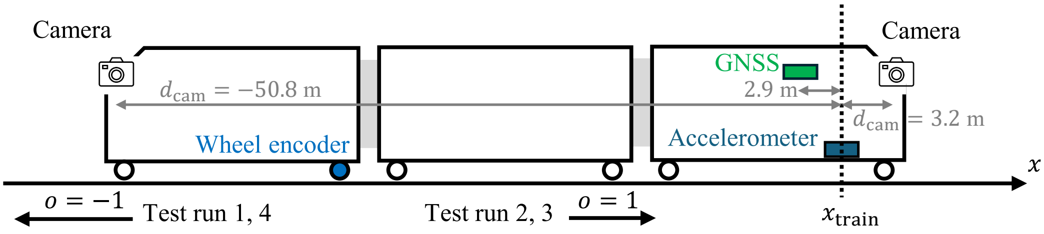

Figure 6: The sensor installations on the railway vehicles in the experiment.

position is defined as the train's reference position. GNSS and wheel encoder data are also acquired as ground truth data for position and velocity. Using this setup, the experiment was conducted on a local railway line in Japan. The total length of the line is about 9.6 km, and we conducted two test runs in each direction. This means four measurements are performed to evaluate the accuracy of the position and velocity estimates. Hereafter, the four runs are referred to as test runs 1, 2, 3, and 4 in the order in which they were conducted.

There are 54 gradient markers along the route for each one-way run. On average, this corresponds to approximately one marker every 180 m. The maximum interval between adjacent gradient markers is approximately 560 m. Therefore, if image recognition is performed correctly, the train position can be corrected at intervals of at most approximately 560 m. The thresholds used to evaluate the image recognition results are set as follows:

$$T_{\mathrm{th}} = 0.5\ \mathrm{s}, \quad P_{\mathrm{th}} = 0.7, \quad A_{\mathrm{th}} = 0.3, \quad X_{\mathrm{th}} = 50\ \mathrm{m}. \tag{8}$$

3.2 Gradient marker detection using image recognition

The image recognition model was developed using video recorded along the same experimental route during runs conducted separately from the data acquisition runs used for validation. The dataset created from this video contained approximately 7000 images of gradient markers, and the image recognition model was trained using YOLOv8 [13].

The results of gradient marker detection using this model are shown in Table 1. Here, TP denotes a gradient marker that was correctly detected and used for position

Table 1: The result of gradient marker detection.

Test run	TP	FP	FN	Precision	Recall
1	43	0	11	1.00	0.796
2	49	1	5	0.980	0.907
3	49	0	5	1.00	0.907
4	35	0	19	1.00	0.648

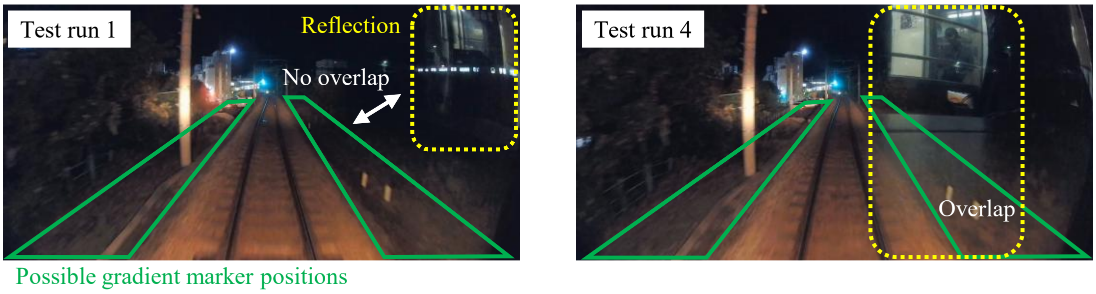

Figure 7: The reflection in the camera image during test run 4.

correction, FP denotes a detection event that did not correspond to a distinct gradient marker, and FN denotes a gradient marker that was not used for position correction. Except for test run 4, approximately 80% to 90% of the gradient markers were successfully used to correct positions. In contrast, the recall decreased substantially only in test run 4; therefore, this result was examined in detail. The detection results from test run 4 showed that multiple tracking IDs were frequently assigned to the same gradient marker, thereby shortening the detection duration. In addition, as shown in Fig. 7, light reflection was observed in the image region where gradient markers appeared, and the video data were less clear than those obtained in the other test runs. Although the decreased recall in test run 4 did not cause a severe degradation of position estimation, this result indicates that the robustness of the image recognition model and sensor setups under adverse lighting conditions remains an important issue for future work. Rather, it could be addressed by improving the sensor installation method or by further training the image recognition model. Furthermore, even when some gradient markers are missed, the gradient angle θ is continuously retrieved from the database based on the current estimated position. Therefore, estimation accuracy can be maintained over relatively long distances.

There is one false positive case in test run 2. This is because one gradient marker has two tracking IDs. This is not a dangerous case because there are only a few meters of positional error in a few seconds.

3.3 Position and velocity estimation by the proposed method

First, the velocity estimation results obtained using the proposed method are discussed. For quantitative evaluation, the RMSE of the velocity for each test run is shown in Table 2. Since ZUPT is adopted, the evaluation is conducted only during periods when the train is running. In each test run, the velocity estimation error is approximately $1 \sim 2$ km/h. This indicates the proposed method provides practically useful velocity estimates.

To examine the estimation behavior during actual operation in detail, the velocity estimation result and the corresponding velocity error for test run 2 are shown in the top and middle figures of Fig. 8. The light blue vertical dotted lines indicate the times

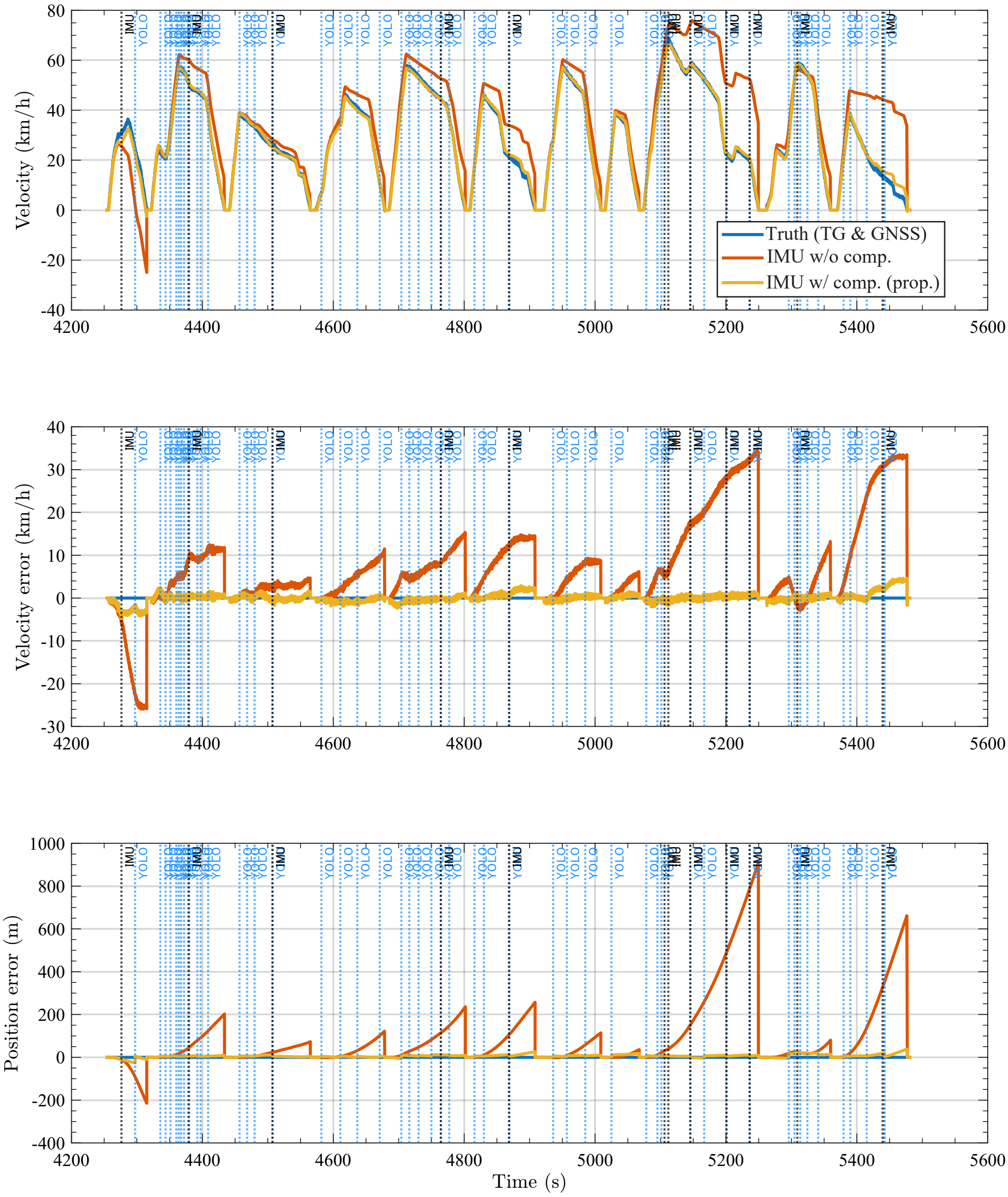

Figure 8: The velocity and position estimation results during test run 2. The top figure shows estimated velocity, the middle one shows estimated velocity error, and the bottom one shows estimated position error.

at which gradient markers were detected by image recognition, whereas the black vertical dotted lines indicate the times at which the gradient information was updated because the estimated current position, x_{train}, passed the installation position of a gradient marker. The results show that the proposed method, indicated by the yellow line, achieves higher estimation accuracy than the case without image recognition

Table 2: RMSEs of velocity estimation.

Test run	IMU w/o comp.	IMU w/ comp. (prop.)
1	11.7 km/h	1.49 km/h
2	13.2 km/h	1.18 km/h
3	12.7 km/h	1.57 km/h
4	10.3 km/h	1.28 km/h

Table 3: RMSEs of position estimation.

Test run	IMU w/o comp.	IMU w/ comp. (prop.)
1	151 m	6.94 m
2	202 m	9.63 m
3	190 m	10.3 m
4	125 m	12.0 m

based position and gradient corrections. In addition, setting the velocity to zero while the train is stopped effectively suppresses divergence of the velocity estimate.

Next, the position estimation results obtained using the proposed method are discussed. As in the velocity evaluation, the RMSE of the position for each test run is shown in Table 3. The results show that the position error in each test run is significantly reduced by the proposed method.

To examine the estimation behavior in detail, the position estimation error for test run 2 is shown in the bottom figure of Fig. 8. As in the velocity plots, the light blue vertical dotted lines indicate the times at which gradient markers were detected by image recognition, whereas the black vertical dotted lines indicate the times at which the gradient information was updated because the estimated current position, x_{train}, passed the installation position of a gradient marker. The results show that the position is corrected when a gradient marker is detected by image recognition, thereby preventing divergence of the estimate. In addition, when gradient compensation is not applied, the estimation error diverges more rapidly than with the proposed method. These results demonstrate that both gradient compensation and position correction in the proposed method are effective for position estimation.

4 Conclusions

In this study, we refined an approach that uses an accelerometer and a camera to estimate train position and velocity, and validated its accuracy using data obtained from actual train tests. The proposed method achieved a velocity estimation accuracy of approximately $1 \sim 2$ km/h. For position estimation, which requires double integration of the measured acceleration, the proposed method achieved an average accuracy of approximately 10 m. These results demonstrate the feasibility of the proposed method under actual railway conditions.

However, estimation accuracy is partially maintained by correcting velocity and position while the train is stopped. Therefore, estimation accuracy may decrease when the train runs for a long period without stopping. Possible solutions for correcting position and velocity during operation include using GNSS only in open sky areas and installing balises. These issues remain topics for future work.

Acknowledgements

We are grateful to the Izu Hakone Railway Company for their support in preparing the experimental vehicle and conducting test runs in the railway environment. This work was supported by JSPS KAKENHI Grant Numbers JP24K07521 and JP25KJ1125. Generative AIs were used to translate and edit sentences.

References

[1] K. Kim, S. -H. Kong, S. -Y. Jeon, “Slip and Slide Detection and Adaptive Information Sharing Algorithms for High-Speed Train Navigation Systems”, IEEE Transactions on Intelligent Transportation Systems, 16, 6, 3193-3203, 2015.

[2] G. Theeg, S. Vlasenko, Railway Signalling & Interlocking, 1st ed., EUrail Press, 2009.

[3] Y. Wang, Y. Lou, W. Song, Z. Tu, Y. Wang, S. Zhang, “Simultaneous Localization of Rail Vehicles and Mapping of Surroundings With LiDAR-Inertial-GNSS Integration”, IEEE Sensors Journal, 22, 14, 14501-14512, 2022.

[4] Y. Chen et al., “An LSTM-Assisted GNSS/INS Integration System Using IMU Recomputed Error Information for Train Localization”, IEEE Transactions on Aerospace and Electronic Systems, 60, 3, 2658-2671, 2024.

[5] E. M. Lee, K. C. Marsim, H. Myung, “LODESTAR: Degeneracy-Aware LiDAR-Inertial Odometry With Adaptive Schmidt-Kalman Filter and Data Exploitation”, IEEE Robotics and Automation Letters, 11, 1, 922-929, 2026.

[6] J. Chang, R. Hu, F. Huang, D. Xu, L. -T. Hsu, “LiDAR-Based NDT Matching Performance Evaluation for Positioning in Adverse Weather Conditions”, IEEE Sensors Journal, 23, 20, 25346-25355, 2023.

[7] M. Ma, Q. Song, Y. Li, Z. Zhou, “A zero velocity intervals detection algorithm based on sensor fusion for indoor pedestrian navigation” 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chengdu, China, 2017.

[8] D. Q. Duong, J. Sun, T. P. Nguyen, L. Luo, “Attitude estimation by using MEMS IMU with Fuzzy Tuned Complementary Filter” 2016 IEEE International Conference on Electronic Information and Communication Technology (ICEICT), Harbin, China, 372-378, 2016.

[9] B. Candan, H. E. Soken, “Robust Attitude Estimation Using IMU-Only Measurements”, IEEE Transactions on Instrumentation and Measurement, 70, 9512309, 1-9, 2021.
[10] K. Nagai, W. Ohnishi, T. Koseki, Y. Setoguchi, T. Yamaguchi, S. Morita, K. Tanaka, “Onboard Train Position and Velocity Estimation using Sign Detection by Image Recognition and Accelerometer”, The papers of Technical Meeting on “Transportation and Electric Railway”, TER-26-064, 2026.
[11] Z. Wang, H. Zhao, S. Qiu, Q. Gao, “Stance-Phase Detection for ZUPT-Aided Foot-Mounted Pedestrian Navigation System”, IEEE/ASME Transactions on Mechatronics, 20, 6, 3170-3181, 2015.
[12] Ministerial Ordinance to Provide Technical Standards for Railways, The Ministry of Land, Infrastructure, Transport and Tourism, Japan, 2025.
[13] G. Jocher, A. Chaurasia, J. Qiu, Ultralytics YOLOv8 (8.0.0) [Computer software], https://github.com/ultralytics/ultralytics, 2023.