2021 IEEE Intelligent Vehicles Symposium (IV)
July 11-17, 2021. Nagoya, Japan

Vehicle Distance Measurement based on Visible Light Communication Using Stereo Cameras

Ruiyi Huang[1], Takaya Yamazato[1], Masayuki Kinoshita[2], Hiraku Okada[1], Koji Kamakura[2], Shintaro Arai[3], Tomohiro Yendo[4] and Toshiaki Fujii[1]

***Abstract*— Visible light communication based intelligent transportation systems (ITS-VLC) show great potential for future urban mobility. This study presents a performance evaluation of range estimation between vehicles and infrastructures in an ITS-VLC system. In the proposed ITS-VLC system, it is easy to simultaneously conduct communication and ranging using stereo cameras. However, the stereo camera calibration becomes a problem during simultaneous communication and ranging due to vehicle vibration. Using the data from LED transmitters and stereo cameras, it can obtain multiple measurements of distance. The monocular-stereo fusion algorithm is applied to visible light ranging in the proposed scheme using particle swarm optimization. We employed real data from the field trial experiment and achieved a ranging accuracy of 60 $\pm$ 1.0 m.**

I. INTRODUCTION

Recently, the rapid development of intelligent transportation system (ITS) has revolutionized the way we think about vehicles. In autonomous driving, the core technology of the future traffic system, visible light communication (VLC) uses the ubiquitous LED lights (e.g., traffic lights, car headlights and taillights) to communicate with vehicles, providing drive assistance information, which will greatly ease the pressure of communication spectrum on the road [1]–[3]. The combination of VLC and driverless cars has shown great potential in the transportation system market, such as the application of optical communication image sensor [4] .

Apart from vehicle communication, vehicle ranging is also an application of VLC [5]. We can use the unique LED-ID to locate vehicles with every LED light on the road. The huge LED network that exists on roadways throughout the world can provide high precision positioning system for vehicles [6], [7].

Compared with 4G/5G and other vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) technologies, VLC does not require additional devices for communication. For realizing vehicle locating, we only need to send unique LED IDs to the vehicles, which is convenient and easy to implement.

In this study, we employ stereo cameras to provide vehicles with simultaneous ranging and communication functions. VLC signals will be transmitted among different vehicles and lighting devices as shown in Fig.1. The stereo cameras are mounted on the cars to receive VLC signals and obtain the depth map from the front view. By receiving the range information, it makes the ITS-VLC system more stable and precise. Moreover, because we use high-speed stereo cameras, vehicle ranging and communication can be completed in a very short time (exceeding 1000 Hz).

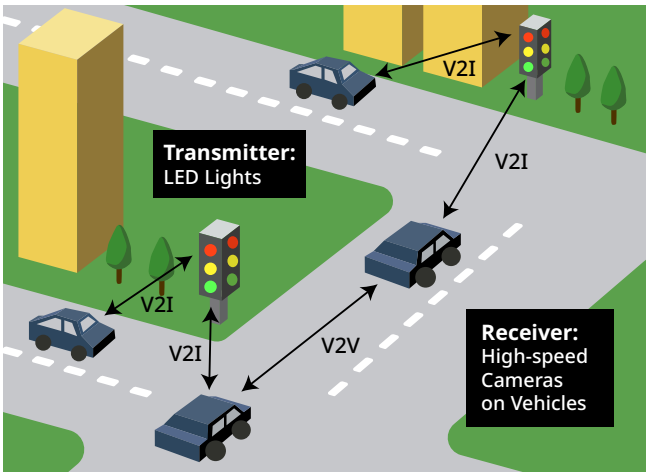

Fig. 1. Diagram of the proposed ITS-VLC system. The LED lights transmit visible light signals between vehicle to infrastructure (V2I) and vehicle to vehicle (V2V). High-speed stereo cameras are mounted on the cars as receivers.

The remainder of this paper is structured as follows. Section II presents a brief overview of the proposed simultaneous VLC and ranging system. Section III illustrates the basic mechanism of monocular ranging and stereo ranging using an LED transmitter. Section IV describes the proposed vehicle ranging approach. The experiment setup is described in Section V and an experimental evaluation is presented in Section VI. Section VII presents the conclusion.

II. SYSTEM OVERVIEW

This section presents an overview of the proposed integrated system and simultaneous VLC and ranging can be achieved. The VLC and ranging data will be output simultaneously and integrated into a single device. In this paper, only the ranging function is discussed. The range estimation is achieved based on visual odometry.

A. System Concept

In our previous study [8], we used a high-speed stereo camera to perform communication and compared the algorithm of maximal ratio combining and selective diversity,

[1]Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8603, JAPAN
[2]Chiba Institute of Technology, 2-17-1, Tsudanuma, Narashino, 275-0016, JAPAN
[3]Okayama University of Science, 1-1, Ridai-cho, Kita-ku, Okayama, 700-0005, JAPAN
[4]Nagaoka Univeresity of Technology, 1603-1, Kamitomioka, Nagaoka, Niigata, 940-2188, JAPAN

2021 IEEE Intelligent Vehicles Symposium (IV) | 978-1-7281-5394-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/IV48863.2021.9575534

exhibiting a symbol error rate (SER) below 10^{-1}. In [9], we showed the performance evaluation of simultaneous ranging and communication; the communication performance shows an SER below 10^{-1}. The largest communication distance was about 60 m [8], [9]. Thus, we have succeeded in VLC using stereo cameras, and now we focus on simultaneous communication and ranging. This study focuses only on the ranging.

In [10], we used a high-speed stereo camera to perform the range estimation with a single LED and the on-off keying (OOK) modulation method to achieve the accuracy of 60 ± 0.1 m. In [9], we investigated the influence of the pulse width modulation (PWM) method, and proposed a brightness selection method to achieve the accuracy of 60 ± 0.15 m; the accuracy without brightness selection was 60 ± 0.5 m. These methods can achieve simultaneous ranging and communication. However, these experiments were performed using a stationary vehicle. This paper marks the first time that we have used a moving ego-vehicle.

This paper is an extension of our previous study. Although high-speed stereo cameras provide extremely rapid refresh rates on simultaneous communication and ranging, calibration between two cameras is relatively not easy due to vehicle vibration. To improve the ranging performance, we introduce a ranging pattern using the LED transmitter with only four image frames. Thus, it is possible to obtain multiple measurements of distance to optimize the ranging results without affecting the communication performance. The combination with ranging patterns is necessary because although stereo cameras can realize simultaneous communication and ranging, their ranging accuracy is lower than that of monocular cameras using the ranging patterns; in contrast, monocular cameras can perform accurate ranging but cannot perform simultaneous VLC.

In this study, we propose a novel algorithm combining single and stereo cameras. By processing the introduced ranging frames, we obtain different measured distances between two cameras. Then, the relative positions of the two cameras are re-calibrated using the particle swarm optimization (PSO) algorithm. The optimized parameters are obtained to achieve better ranging performance. Besides, we accomplished field tests to show that the performance of the proposed method is better than the traditional methods.

B. System Model

Fig. 2 shows a block diagram of the proposed system model.

At the transmitter end, each data packet comprises ranging and data patterns. The inverted pattern was also transmitted for LED detection.

At the receiver end, LED detection is accomplished at the beginning using threshold detection. Then, communication function is achieved using signal demodulation, and the ranging function is conducted using the proposed monocular-stereo fusion algorithm. The data transmission and range estimation are achieved at the same time.

III. Basic Concept of Range Estimation

A. Camera Model

Fig. 3 shows the stereo camera model. In the experiment, the two cameras are placed strictly parallel to the LED at the beginning; however, slight differences in the camera postures may still exist. As shown in Fig. 3, the original point of the world coordinate is assumed to be at the center of the left camera O_L. T denotes the translation vector from left to right camera and R denotes the rotation matrix from left to right camera. T and R are also known as the camera extrinsic matrix.

The relationship between the projected object points of left and right cameras can be represented as

$$P_R = RP_L + T \tag{1}$$

If we denote the coordinates of P_L and P_R as (x_R, y_R, z_R) and (x_L, y_L, z_L), respectively, (1) can be expressed by

$$\begin{bmatrix} x_R \\ y_R \\ z_R \end{bmatrix} = \begin{bmatrix} x_L \\ y_L \\ z_L \end{bmatrix} R + T \tag{2}$$

Through these equations, the stereo images can be rectified using the known T and R; it also means the stereo cameras will be calibrated to the strictly parallel position. T and R are computed using a designed calibration object at the beginning of the experiment using the algorithm in [11].

The next range estimation methods are based on the calibrated stereo cameras. We calibrate the stereo cameras once before starting the experiment. We will rectify the stereo images with an updated extrinsic matrix during the driving process using the proposed method which will be clarified in Section IV. The intrinsic parameters are assumed to be the same in the entire experiment.

B. Monocular Ranging Considering Vehicle Vibration

The algorithm can estimate the range using only a single camera. When importing ranging frames, we use two LEDs placed on the edge of the LED array and only one camera receiver for distance estimation [12].

Fig. 4 shows the LED ranging pattern. It comprises two frames to deal with the noise that vehicle vibration brings to the system when it is undergoing ego-motion. The displacement of 1 and 2 on the image plane (d_1) is used to calculate the vehicle vibration. Combining the displacement of 1 and 3 (d_2), the distance between LED array and vehicle is estimated using the projection model of a single camera and the known LED size.

The estimated range Z_m is given as follows:

$$Z_m = f\frac{b_m}{d_2 - d_1}, \tag{3}$$

where b_m is the actual range between 2 and 3, and f is the focal length of the camera.

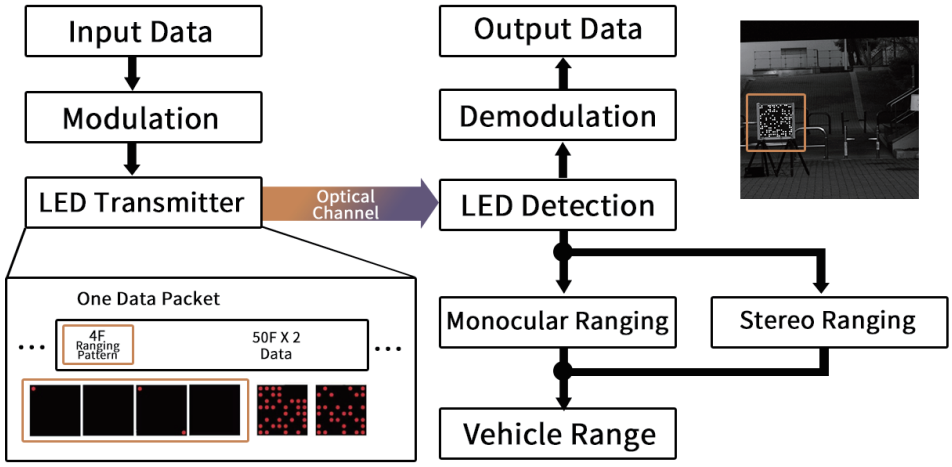

Fig. 2. Block diagram of the system model. The LED transmitter is marked with an orange rectangle in the image shown on the upper right. The communication and ranging functions are conducted separately.

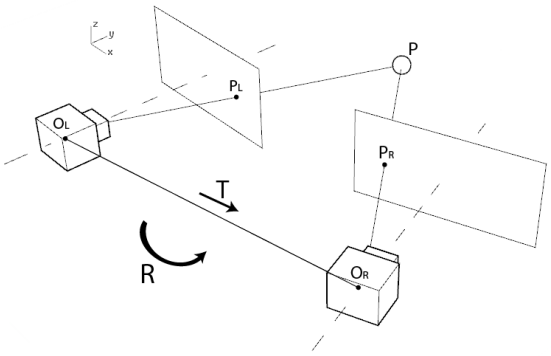

Fig. 3. Stereo Camera Model. O_L and O_R denote the center of the cameras respectively. The object is on point P, while P_L and P_R represent the projected points on the image planes.

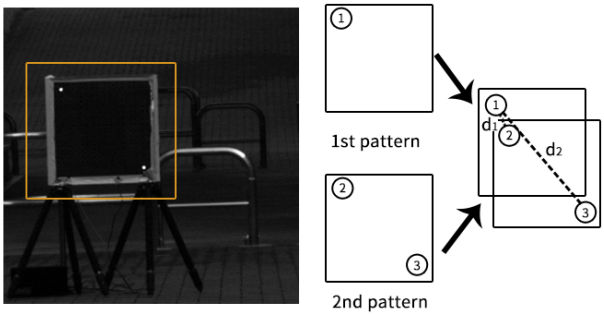

Fig. 4. Ranging pattern using a single camera. d_1 is the displacement of points 1 and 2, and d_2 is the displacement of points 1 and 3.

C. Stereo Ranging with Rectified Stereo Images

When we use two cameras to estimate the ranging, the distance between LEDs and cameras can be obtained through disparity estimation from multiple captures of the same background from different viewpoints, i.e., left and right cameras [13].

If we denote the actual size per pixel as ρ and disparity as d, then, d can be expressed as $d = \rho(x_l - x_r)$, where x_l and x_r represent the LED positions, i.e., P_L and P_R, in the x- axis on the image planes of left and right cameras, respectively.

We can easily calculate the range using the following equation:

$$Z_s = f\frac{b}{\rho(x_l - x_r)}, \tag{4}$$

where Z_s is the estimated range between the object and cameras; f is the focal length of cameras; b is the distance between the two cameras.

Thus, if we can find the corresponding feature point on the image plane of the left and right views, the 3D coordinates of the project can be determined accurately.

In the present study, the image displacement is calculated using phase-only correlation (POC) [14] algorithm.

IV. Monocular-Stereo Fusion Algorithm

A. Overview of the Fusion Algorithm

Fig. 5 shows the block diagram of the proposed monocular-stereo fusion algorithm.

First, we use the stereo image pair to estimate the range. If the image pair contains a ranging pattern, the range will be estimated using monocular images and stereo image

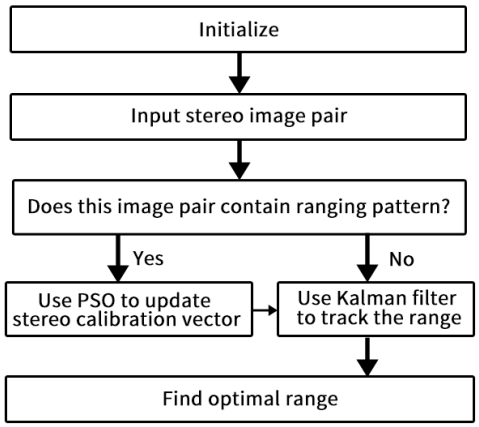

Fig. 5. Block diagram of the proposed algorithm

pair, respectively, and three measurements of range can be obtained. Then, we optimize the stereo calibration vector using PSO. This optimization is conducted every 400 frames. Finally, after the optimization, Kalman filter is applied to track the transmitter and the optimal range is determined.

B. Vehicle Tracking: Kalman Filter

A Kalman filter is used to track the LED transmitter. Equations (5) and (6) represent the Kalman filter function [15]. V_k denotes the full state of the tracked object, and M_k denotes the predicted measurement taken in each stereo image pair.

$$V_{k+1} = AV_k + \sigma_k \tag{5}$$

$$M_k = CV_k + \varepsilon_k \tag{6}$$

Here, σ_k and ε_k denote the plant and observation noise, respectively. The full state of the tracked object is given by

$$V_k = [distance \;\; velocity] \tag{7}$$

The state transition matrix A and the observation matrix C are given as follows:

$$A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{8}$$

The Kalman filter tracks the transmitter in 3D world coordinates.

C. Camera Pose Update: Particle Swarm Optimization

The stereo rectification will be optimized using PSO. The two cameras will re-calibrate every 400 frames with the image frames of ranging pattern as explained in Section IV-A. In this section, we illustrate a brief overview of PSO and the proposed fitness function within PSO.

PSO is an evolutionary optimization algorithm using the swarm intelligence technique proposed by Kennedy and Eberhart [16]. It uses the fitness function and a certain number of candidate particle swarms to determine the optimal solution for the N-dimensional search space. Each particle in the PSO swarm represents a candidate solution in the N-dimensional search space and has a corresponding N-dimensional position and velocity vector. We update the best position for each particle and the overall swarm in every iteration to find the best solution for the corresponding fitness function.

During the camera posture updating phase, we employ the optimized PSO parameters to constrain the stereo disparity estimation. It is achieved by using the monocular range images as prior disparity range images. The optimal calibration vectors rectify the two cameras to the proper relative positions based on epipolar constraint.

Each particle is defined based on the rotation matrix R and translation matrix T between two cameras: the extrinsic matrix. The particles are defined as follows:

$$X = [\alpha, \beta, \gamma, t_x, t_y, t_z] \tag{9}$$

where α,β and γ represent the rotation angles in the x, y and z directions, respectively; t_x,t_y,t_z represent the translation in the x, y and z directions, respectively. The dimension of particle X is six; $X^1 = \alpha, X^2 = \beta, X^3 = \gamma, X^4 = t_x, X^5 = t_y, X^6 = t_z$.

We can calculate the range with each particle using the stereo ranging methods discussed in Section III.

The proposed PSO algorithms are conducted in the following steps:

Step 1: Create a swarm of particles randomly distributed over M.
Step 2: Evaluate each particle's position using fitness function.
Step 3: Update the particle best position for each particle.
Step 4: Update the global best position.
Step 5: Update velocities of each particle.
Step 6: Move particles to their new positions.
Step 7: Repeat Step 2-4 until the number of iterations reaches the designed condition.

The fitness function is defined as

$$f = \sum_{j=1}^{N} (S - I_j)^2 \tag{10}$$

where I_j is the input estimated range, i.e., the ranging results of monocular and stereo vision before conducting PSO ($N = 3$); S is the ranging result calculated for each particle.

We update the velocity V_i^q and position X_i^q using the following equations:

$$\begin{aligned} V_{i+1}^q = \omega V_i^q + c_1 rand(0,1)(pbest - X_i^q) \\ + c_2 rand(0,1)(gbest - X_i^q) \end{aligned} \tag{11}$$

$$X_{i+1}^q = X_i^q + V_i^q; \tag{12}$$

where ω is the inertia weight, and c_1 and c_2 are learning factors. ω was set as 0.4, and c_1 and c_2 were set as 2.

In our experiment, the number of particles M is 10 and the iteration is 15 times. Fig. 6 shows the change of the fitness f with the iteration times at a distance of around 30, 40, 50 and 60 m. We can see from the graph that the value of fitness is slightly different at each trial; for all the testing range, the the best value for the number of iterations is around 15.

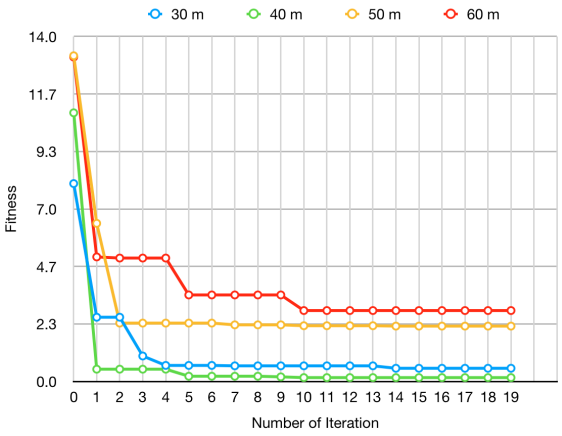

Fig. 6. Fitness as a function of the number of iterations.

TABLE I
EXPERIMENTAL PARAMETERS

Transmitter	
Transmitter	LED Array (SunLED XZM2ACR55W-3)
Size of the LED Array	45 cm × 45 cm
Frequency	500 Hz
Modulation Method	On-off keying
Vehicle Speed	20 km/h
Receiver	
Receiver	Photoron IDP-Express
Shutter Speed	1000 fps
Resolution	512 × 512 pixels
Focal Length	35 mm
Distance of two cameras	0.25 m
Pixel Size	10 μm
Communication Distance	20-70 m
Number of frames used for estimation	
8000	

V. EXPERIMENT SETUP

The performance evaluation in this paper is based on real data from a field trial experiment. In the experiment, we drove the car leaving the LED transmitter from 20 to 70 m, captured the LED transmitter images, and estimated the range between them offline.

The experiment was conducted outdoors in the daytime in Nagoya University. Table 1 presents the experimental parameters. Two cameras were mounted inside the vehicle, as shown in Fig. 7.

The actual distance between the vehicle and LED array is determined by observing the captured images. We recorded the actual distances corresponding to the piles on either side of the road in advance and then deduced the actual distances from the piles' positions on the image planes.

VI. RESULT AND DISCUSSIONS

We compared the proposed approach with two conventional methods and the range estimation results of these three methods are shown in Fig. 8.

The red circle represents the proposed algorithm. The green circle and yellow crosses represent the method of using

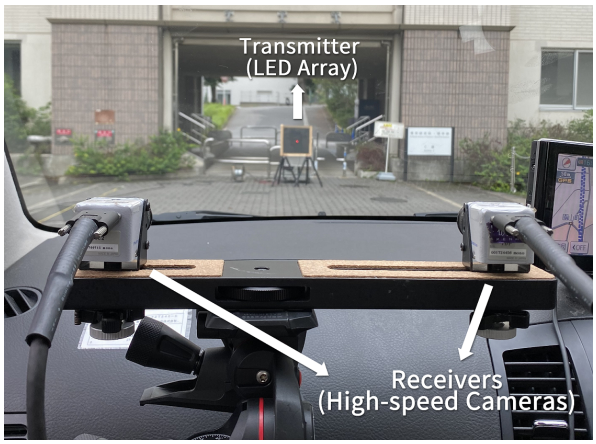

Fig. 7. Experiment setup of the proposed system. A 16 × 16 LED Array is used as the transmitter and two high-speed cameras are used as receivers. The vehicle drives leaving the transmitter.

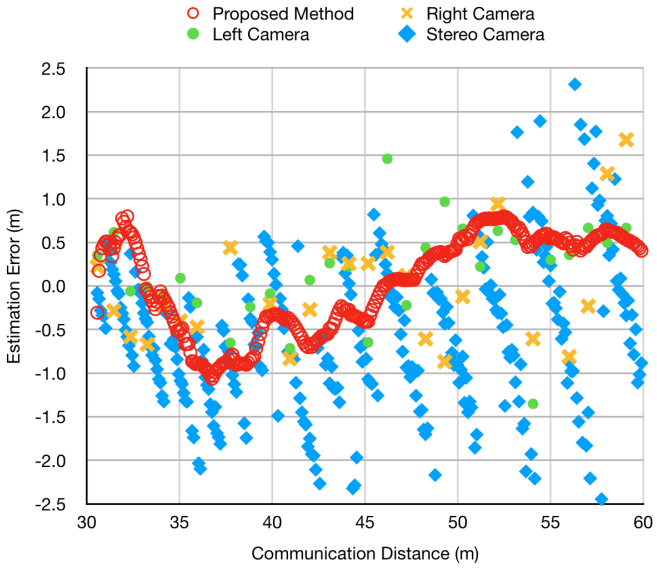

Fig. 8. Estimation error as a function of communication distance. The red circles show the performance of this algorithm; the yellow crosses and green circles show performance for conventional monocular ranging using right and left cameras, respectively; the blue diamonds show performance for the stereo system.

conventional monocular ranging method with left and right camera image, respectively. Finally, blue diamonds show the estimation error for the conventional stereo ranging method.

There is a clear tendency for errors from each method. In the case of the conventional monocular ranging method, the error increases when the range increases due to decreasing LED sizes. For the conventional stereo ranging method, it fluctuates regularly over a certain interval of the communication distance. It is because the LED size was too small to recognize in sub-pixel levels. The estimation error tends to have the best value when the LED size is in integer, meaning that the LED is captured in pixel level.

Fig. 9 shows the estimation error variance for the proposed method. It cannot tell how distance affects accuracy. For the

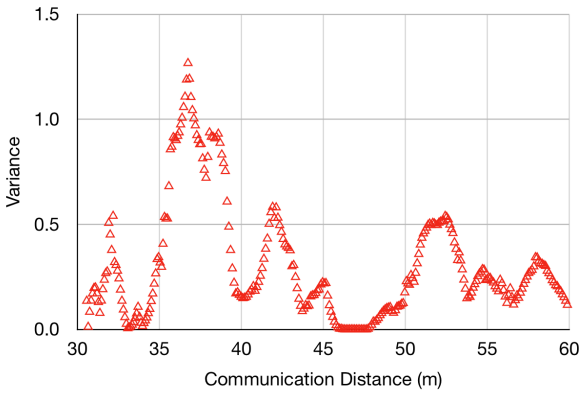

Fig. 9. Variance for the estimation error of proposed method.

proposed method, the error decreases from 35 to 45 m. Other factors might also affect the accuracy, such as defocusing of the cameras.

For the conventional methods, the estimation method of using a single camera and LED ranging pattern is more precise than using stereo cameras; however, the ranging rate is much lower than stereo cameras. It is worth noting that during the ranging process using stereo cameras, it is expected to simultaneously perform the communication; however, only the ranging function is shown here.

By combining the single camera method and stereo camera method, the range estimation becomes more accurate and the refresh rate is fast as 1000 fps in this study.

VII. CONCLUSIONS

We proposed a novel algorithm combining monocular and stereo ranging schemes using an LED transmitter. The range estimation and communication between vehicles and LEDs can be achieved using the proposed system.

We achieved simultaneous communication and ranging using the stereo cameras. However, we obtained poor ranging performance due to inaccurate camera calibration. A single camera with LED ranging pattern can perform precise ranging as it has a calibration object using the LED ranging pattern; however, it cannot conduct simultaneous VLC and the ranging rate is low. Thus, we combined the monocular ranging method with stereo-vision to improve the ranging performance for both ranging methods and enable simultaneous VLC and ranging.

The experimental results of the proposed approach showed that the estimation error is 1.0 m below the communication distance of 60 m, which is better than the conventional methods. Note that our previous works [9], [10] were conducted under the stationary cases, whereas the experiment in the present study was conducted using a moving ego-vehicle. The accuracy for the conventional stereo-vision ranging method in [9] and [10] was 2.5 m in the present study. Besides, the communication function is possible using the proposed scheme by demodulating the VLC data.

In summary, the proposed algorithm can obtain high ranging accuracy using stereo cameras. The ITS-VLC system with stereo cameras can simultaneously transmit VLC data and estimate vehicle distance, thus significantly improving the performance of the ITS-VLC system.

ACKNOWLEDGMENT

The authors would like to thank Prof. Masaaki KATAYAMA of Nagoya University and Assistant Prof. Chedlia BEN NAILA of Nagoya University for their valuable suggestions.

REFERENCES

[1] H. M. La, R. Lim, J. Du, S. Zhang, G. Yan, and W. Sheng, "Development of a small-scale research platform for intelligent transportation systems," *IEEE Trans. Intell. Transport. Syst.*, vol. 13, no. 4, pp. 1753–1762, Dec 2012.

[2] J. S. Kwak and J. H. Lee, "Infrared transmission for intervehicle ranging and vehicle-to-roadside communication systems using spread-spectrum technique," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 1, pp. 12–19, March 2004.

[3] A. Ohmura, T. Yamazato, H. Okada, T. Fujii, S. Arai, and Yendo, "Accuracy improvement by phase only correlation for distance estimation scheme for visible light communications using an LED array and a high-speed camera," *Proc. of 20th World Congress on Intelligent Transport Systems*, Oct. 2013.

[4] T. Yamazato, I. Takai, H. Okada, T. Fujii, T. Yendo, S. Arai, M. Andoh, T. Harada, K. Yasutomi, K. Kagawa, and S. Kawahito, "Image-sensor-based visible light communication for automotive applications," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 88–97, July 2014.

[5] T. Yamazato, A. Ohmura, H. Okada, T. Fujii, T. Yendo, S. Arai, and K. Kamakura, "Range estimation scheme for integrated I2V-VLC using a high-speed image sensor," in *2016 IEEE International Conference on Communications Workshops (ICC)*, May 2016, pp. 326–330.

[6] A. Jovicic, J. Li, and T. Richardson, "Visible light communication: opportunities, challenges and the path to market," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 26–32, December 2013.

[7] R. Roberts, P. Gopalakrishnan, and S. Rathi, "Visible light positioning: Automotive use case," in *IEEE Vehicular Networking Conference (VNC 2010)*, Dec 2010, pp. 309–314.

[8] M. Kinoshita, T. Yamazato, H. Okada, T. Fujii, S. Arai, T. Yendo, and K. Kamakura, "A comparison of reception methods for visible light communication using high-speed stereo cameras," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–5.

[9] R. Huang, M. Kinoshita, T. Yamazato, H. Okada, K. Kamakura, S. Arai, T. Yendo, and T. Fujii, "Performance evaluation of range estimation for image sensor communication using phase-only correlation," *2020 IEEE Global Communications Conference (GLOBECOM) Workshops*, 2020.

[10] M. Kinoshita, K. Kamakura, T. Yamazato, H. Okada, T. Fujii, S. Arai, and T. Yendo, "Stereo ranging method using led transmitter for visible light communication," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[11] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 911–916, 2000.

[12] T. Yamazato, "V2X communications with an image sensor," *Journal of Communications and Information Networks*, vol. to be published., 2017.

[13] D. Nister, O. Naroditsky, and J. Bergen, "Visual odometry," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 1, 2004, pp. I–I.

[14] H. Foroosh, J. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 188–200, Mar 2002.

[15] S. Sivaraman and M. M. Trivedi, "Combining monocular and stereo-vision for real-time vehicle ranging and tracking on multilane highways," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2011, pp. 1249–1254.

[16] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, 1995, pp. 1942–1948 vol.4.