A Localization Method Using Reflected Luminance Distribution

Yoshihiro Yamashita[1(✉)], Shota Shimada[1], Hiromichi Hashizume[2], Hiroki Watanabe[1], and Masanori Sugimoto[1]

[1] Hokkaido University, Sapporo, Japan
{yamash.lab,hiroki.watanabe,sugi}@ist.hokudai.ac.jp,
shimadas@frontier.hokudai.ac.jp
[2] National Institute of Informatics, Tokyo, Japan
has@nii.ac.jp

Abstract. Visible light positioning is expected to become an effective means of indoor localization, but most existing methods require the capture of direct light, which is a significant limitation. In this paper, we propose a novel localization method based on received signal strength, which does not need to use direct light signals but instead uses reflected light from the floor. Our method is based on two observations. First, assuming a flat floor, the reflected light from the light source decays according to a gradient model whose peak is just below the light source. Second, the decay can be estimated effectively even for a floor surface several meters away from the light source. Inspired by these observations, we propose a method for estimating the coordinates of the floor surface and the two-dimensional coordinates of the light receiver (a camera). The method uses pattern matching within the distribution of signal decay measurements obtained via photographs of an arbitrary floor surface. Although the proposed method is vulnerable to shadow effects such as those caused by the camera tripod used in our experiments, we achieved a 90th percentile of less than 32 cm in our offline experiments. After removing the tripod shadows from the captured video manually, the same technique achieved a 90th percentile of 22 cm. To investigate the efficiency of the pattern matching, we also conducted experiments on the relationship between pixel utilization and localization results. In this paper, we also discuss camera posture estimation and power consumption issues.

Keywords: Visible light positioning · Received signal strength (RSS) · Indoor positioning · Non-line of sight (NLoS)

1 Introduction

Indoor localization technology is exerting a major impact on human activities, in the same way that the Global Positioning System [8] revolutionized outdoor

T. Hara and H. Yamaguchi (Eds.): MobiQuitous 2021, LNICST 419, pp. 390–405, 2022.
https://doi.org/10.1007/978-3-030-94822-1_21

navigation. In 2020, indoor localization and indoor navigation technology generated $12 million dollars in revenue in the United States (US) and is predicted to reach US$35 million by 2026 [6]. However, even after decades of research, a simple and robust indoor localization solution is yet to be found [9,12,14,22]. Visible light positioning (VLP) [10,11] has recently been attracting attention because of the widespread use and low price of devices based on light-emitting diodes (LEDs). VLP systems have shown some promise for indoor positioning [7] but the unwieldy nature of cameras causes problems in realistic environments. There are three main challenges.

Limited Coverage of the System. Most methods require line-of-sight (LoS) signals from several LEDs simultaneously, requiring either dense or extremely dense arrangements of LEDs or coverage only in some places.

Difficulty of Capturing Multiple LED Signals. Even inside the coverage area of the system, the user will be forced to move the camera in unnatural ways, particularly if the ceiling is low.

Necessity of High-resolution Imaging. Most existing methods use a high-resolution image to detect light sources directly, which is both power hungry and computationally demanding. Note that this problem would remain an issue in implementing real-time positioning systems in the future.

In this paper, we propose an RSS-based method that does not require direct light, and instead captures reflected light from the floor surface to enable localization. The method identifies each of multiple LEDs installed on the ceiling by performing a discrete Fourier transform (DFT) on the reflected signal and using the modulation frequency as the LED's identity (ID). Normally, positioning using only reflected light would suffer from reduced accuracy because of the decrease in signal-to-noise ratio. Therefore, we aim to maintain high positioning accuracy by adopting pattern matching between the measured decay distribution of luminance values and an ideal distribution.

This paper also addresses the following two issues. First, we found experimentally that the positioning accuracy of our method is sensitive to shadows and that it is difficult to capture signals from the floor without shadows. As the percentage of shadowing in the field of view increases, the positioning accuracy decreases, with positioning becoming almost impossible if the shadows occupy more than half of the field of view. Moreover, if the photographer or camera moves, the shadows also move, making it difficult to remove the shadow from the field of view. Unfortunately, we could not eliminate the shadow effects in real time; instead, we removed the shadow manually offline and compared the accuracy with that before removal. Second, our method uses the entire screen and takes a long time to process, so positioning could not be done in real time. To address this, we investigated the relationship between positioning accuracy and the number of screen pixels used.

We implemented our proposed system and evaluated it in a 3.5-m square room with four LEDs on the ceiling. We estimated the position of each of 64 points (8×8 points) on the floor 100 times. The resulting accuracy was 0.32 m

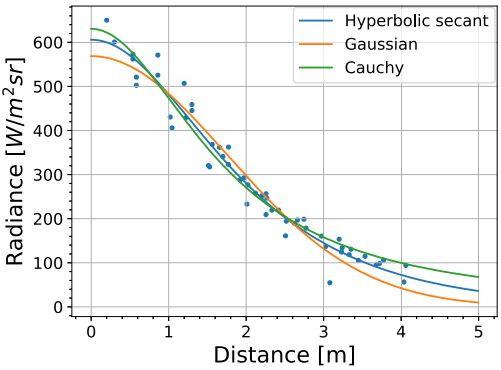


Fig. 1. Experimentally obtained values and distance decay models (Gaussian, hyperbolic secant, Cauchy).

at the 90th percentile, without compensating for shadow effects from the camera tripod. We then applied the offline manual shadow removal method to the same data, achieving an accuracy of 0.22 m at the 90th percentile. Next, we tested the effect of the method on positioning accuracy by reducing the number of pixels used for positioning at two locations with relatively low shadow effects and high positioning accuracy. At one location, we achieved 90th percentile positioning accuracies of 2.2 cm, 4.9 cm, and 48 cm using 1,228,800 pixels, 6,144 pixels, and 1,024 pixels, respectively. From the results of these experiments, we can assess the potential and limiting factors of the proposed method.

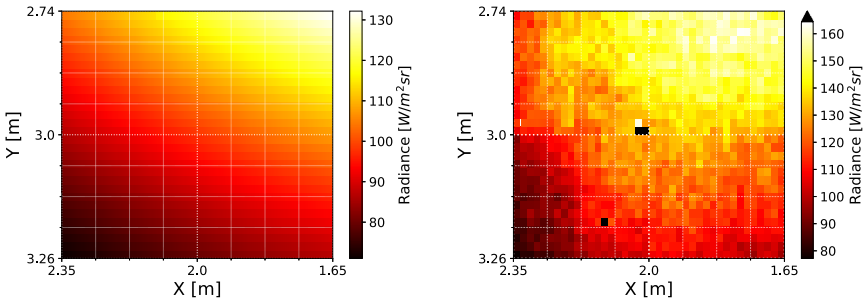
The contributions of this paper are summarized below.

- We have proposed a new positioning method using pattern matching and reflected light from the floor.
- We implemented a prototype of the proposed method, which achieved submeter positioning accuracy.
- We investigated the effects of shadows on the positioning accuracy and the trade-off between the number of pixels used and positioning accuracy.

2 Luminance Distribution Across the Floor

In this section, we describe the preliminary research we conducted prior to developing the proposed method. Our method is based on two observations.

The first is that, assuming a flat floor, the reflected light decays according to a gradient model whose peak is just below the LED. Shimada et al. [18] found that the decay of the reflected light on a Lambertian floor surface can be approximated by the hyperbolic secant distribution of the two-dimensional (2D) distance d_k between the floor surface and the kth LED. The RSS received by the camera from the kth LED is



(a) Distribution of luminance from LED1 obtained by simulation.

(b) Distribution of luminance from LED1 obtained experimentally. Black blocks represent outliers.

Fig. 2. Comparison of experimental data and data obtained by simulation.

$$RSS_k = \frac{C_k}{e^{\frac{\pi}{2}\sigma_k d_k} + e^{-\frac{\pi}{2}\sigma_k d_k}}, \tag{1}$$

where C_k and σ_k are parameters. Figure 1 shows the decay for the hyperbolic secant distribution, Gaussian distribution and Cauchy distribution obtained from the actual data using regression analysis. Note that the amplitude of the signal spectrum decreases as the 2D distance from the LED increases.

Second, the decay of the luminance of the reflected light on the floor can be observed even at a distance of several meters from the point directly under the LED. Figure 2a shows the heat map of the luminance values from LED1 at point $(1.0, 0.5)$, obtained via simulation using video data at 1280×980-pixel resolution with a lens height of 90 cm above the point $(2.0, 3.0)$. The data obtained in our laboratory with the same LED and camera configuration is shown in Fig. 2b. The actual data shows a luminance distribution close to the simulation, indicating that it is possible to estimate the camera coordinates using pattern matching.

Based on these two observations, we propose the new positioning method using floor reflection and pattern matching described in Sect. 3.3.

3 Proposed Method

In our method, signals are received from multiple LEDs installed on the ceiling as shown in Fig. 3. Normally, when multiple LEDs emit light at the same time, the signals overlap on the floor, making RSS detection impossible. To address this, each LED is modulated with a unique carrier frequency, and the RSS is detected by analysis in the frequency domain, using the carrier frequency as the location identity. The LEDs need to be modulated at a high enough frequency that the human eye does not experience flickering. We 100 Hz or higher (denoted F_{100}) as the carrier frequency, because ordinary fluorescent lamps oscillate at a frequency 100 Hz or higher [19]. Section 3.1 explains frequency aliasing and

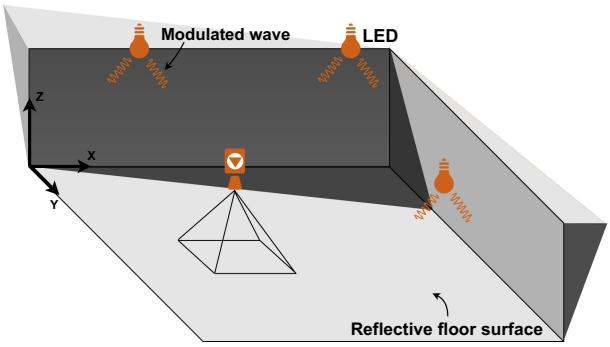


Fig. 3. Overview of the proposed method.

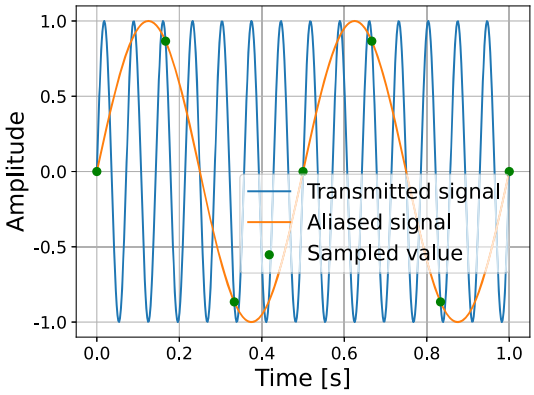


Fig. 4. The 2-Hz aliased signal obtained by sampling a 14-Hz signal 6 Hz.

gives the equation used to find the aliased frequency. In Sect. 3.2, we investigate the relationship between the RSS obtained by analyzing the carrier frequency and the distance from the LED. Section 3.3 estimates the 2D coordinates of an arbitrary point on the floor using pattern matching and the relational equation obtained in Sect. 3.2.

3.1 Frequency Aliasing

The VLP system uses signals modulated at F_{100} to enable identification of each transmitting LED. However, we know from the sampling theorem [16] that we cannot decode the signal accurately because we use a camera with a frequency 50 Hz in our experiment. If the modulated signal is sampled at a frequency below the Nyquist frequency, the aliased frequency obtained via DFT is given by

$$f_a = Min(f_o \ \mathbf{mod}\ f_s, f_s - f_o \ \mathbf{mod}\ f_s), \tag{2}$$

where f_a is the aliased frequency, f_o is the frequency of the transmitted signal, and f_s is the sampling frequency. For example, as shown in Fig. 4, if 14 Hz signal is sampled at a sampling frequency 6 Hz, the aliased frequency f_a is

$$f_a = Min(14 \textbf{ mod } 6, 6 - 14 \textbf{ mod } 6) = Min(2, 4) = 2. \tag{3}$$

This shows that it is possible to identify $f_s/2 - 1$ LEDs, assuming that the aliased frequency of the carrier frequency assigned to each LED is known.

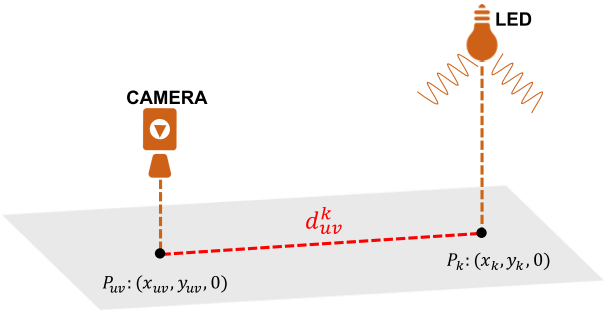


Fig. 5. Two-dimensional distance between camera and LED

3.2 Relationship Between Distance and RSS

Consider the 2D coordinates of the image sensor. Let (u, v) be a randomly selected image sensor, and let $P_{u,v}$ be the point on the floor corresponding to image sensor (u, v). Let P_k be the intersection point between the vertical line down from the kth LED and the floor surface, and let $d^k_{u,v}$ be the 2D distance between $P_{u,v}$ and P_k, as shown in Fig. 5. The $RSS_{u,v}$ value, which is obtained from the image sensor (u, v) using aliasing, decays according to the distance $d^k_{u,v}$. In this section, we develop the relationship between $RSS_{u,v}$ and $d^k_{u,v}$.

The signal emitted by the kth LED at time t can be expressed as follows.

$$b_k(t) = \sin(2\pi t f_k) + \alpha, \tag{4}$$

where α is the DC component keeping $b_k(t)$ always positive and f_k is the carrier frequency assigned to the kth LED. Assuming that the frame rate of the camera is f_c and the exposure time ratio is η, the shutter cycle is $T_c = 1/f_c$, the exposure time is ηT_c, and the time to start capturing the nth image is nT_c. The luminance value $I^n_{u,v}$ received by the image sensor (u, v) when capturing the nth image is

$$I^n_{u,v} = \sum_k \frac{2\pi X(d^k_{u,v})}{T_c} \int_0^{\eta T_c} b_k(t + \delta T_c + nT_c)dt, \tag{5}$$

where δ denotes the delay in shutter timing relative to the signal and $X(d^k_{u,v})$ is the decay function of the luminance value. Note that $X(d^k_{u,v})$ is affected both by the distance $d^k_{u,v}$ and the transfer efficiency between the kth LED and the camera.

The camera captures N ($= f_c$) images per second. The DFT of the video stream $\mathbf{I_{u,v}} = (I^0_{u,v}, I^1_{u,v}, ..., I^{N-1}_{u,v})$ at pixel (u, v) is

$$B^l_{u,v} = \frac{1}{N} \sum_{n=0}^{N-1} I^n_{u,v} e^{\frac{-j2\pi nl}{N}}. \tag{6}$$

Assuming that $\beta_{l,k}$ is obtained by the DFT of $b_k(t)$, $B^l_{u,v}$ is given by

$$B^l_{u,v} = \eta X(d^k_{u,v}) e^{j\pi f_k(2\delta+\eta)} \mathrm{sinc}\,(\pi f_k \eta)\, \beta_{l,k}. \tag{7}$$

By fixing the exposure time ratio to the intrinsic value $\eta = \eta_0$ and using the aliased frequency determined by

$$m_k = Min(f_k \ \mathbf{mod}\ f_c, f_c - f_k \ \mathbf{mod}\ f_c).\,, \tag{8}$$

we can focus on the decay of the amplitude at $l = m_k$. Under these conditions, the amplitude spectrum $|B^{m_k}_{u,v}|$ is affected only by the decay function $X(d^k_{u,v})$. This amplitude spectrum (denoted by $RSS_{u,v,k}$) is then given by

$$\begin{aligned} RSS_{u,v,k} &= |B^{m_k}_{u,v}| \\ &= \eta_0 X(d^k_{u,v}) \mathrm{sinc}\,(\pi f_k \eta_0)\, |\beta_{m_k,k}|. \end{aligned} \tag{9}$$

It is difficult to specify an exact model of the luminance decay $X(d^k_{u,v})$ because the floor is not an ideal Lambert surface [23]. Shimada et al. [18] have shown that $X(d^k_{u,v})$ can be approximated using a hyperbolic secant distribution, as follows.

$$X(d^k_{u,v}) = \frac{C_k}{e^{\frac{\pi}{2}\sigma_k d^k_{u,v}} + e^{-\frac{\pi}{2}\sigma_k d^k_{u,v}}}, \tag{10}$$

where C_k and σ_k are parameters. Therefore, $RSS_{u,v,k}$ can be approximated as

$$RSS_{u,v,k} = \frac{C'_k}{e^{\frac{\pi}{2}\sigma_k d^k_{u,v}} + e^{-\frac{\pi}{2}\sigma_k d^k_{u,v}}}, \tag{11}$$

where $C'_k = \eta_0 C_k \mathrm{sinc}\,(\pi f_k \eta_0)\, |\beta_{m_k,k}|$.

3.3 Pattern Matching

In Sect. 3.2, it was found that the RSS corresponding to an arbitrary image sensor can be calculated as in (11). In this section, we describe a pattern matching process that uses the difference between the calculated RSS and the ideal distribution of luminance. Assuming that the image sensor located at the center of the screen $(0, 0)$ captures the corresponding floor surface (x, y), the floor-surface

position corresponding to the image sensor (u, v) is $(x + W_x u, y + W_y v)$. W_x and W_y are the pixel pitches (i.e., distance between two points on the floor corresponding to a pair of adjacent pixels) in the horizontal and vertical directions, respectively. All pixel pitches can be approximated to the same value, with $W_x = W_y$ if the camera is assumed to be pointing downwards. If the coordinates of the kth LED are (x_k, y_k, z_k), the amplitude spectrum of the luminance of the floor corresponding to (u, v), calculated theoretically, is

$$RSS^{theory}_{u,.v,k} = \frac{C'}{e^{\frac{\pi}{2}\sigma d^k_{u,v}} + e^{-\frac{\pi}{2}\sigma d^k_{u,v}}}, \tag{12}$$

provided that

$$d^k_{u,v} = \sqrt{(x_k - x - W_x u)^2 + (y_k - y - W_y v)^2}. \tag{13}$$

Therefore, we can formulate the error function

$$e_{u,v,k} = RSS^{observed}_{u,v,k} - \frac{C'}{e^{\frac{\pi}{2}\sigma d^k_{u,v}} + e^{-\frac{\pi}{2}\sigma d^k_{u,v}}}, \tag{14}$$

where $RSS^{observed}_{u,v,k}$ is the RSS of the kth LED observed by the image sensor (u, v). Furthermore, we can formulate an equation J for the sum of squares of the error functions.

$$J = \sum_k \sum_{u,v} e^2_{u,v,k}. \tag{15}$$

If we consider J as an evaluation function, we can estimate the (x, y) position where J is minimized as the center of the screen. This problem can be thought of as an unconstrained nonlinear optimization problem, to which the Gauss–Newton method [13] can be applied.

4 Implementation Details

4.1 LED Transmitter

An overview of our experimental implementation is shown in Fig. 6. BXRE-50C4001-B-74-type LEDs from Bridgelux [1] were mounted on a heat sink and used as LED transmitters. Four LEDs were placed on the ceiling at a height of 2.6 m, and the coordinates of LED1, LED2, LED3, and LED4 were (1.0 m, 0.5 m, 2.6 m), (1.0 m, 3.0 m, 2.6 m), (3.2 m, 0.5 m, 2.6 m), and (3.2 m, 3.0 m, 2.6 m), respectively. This was the same arrangement as that of the fluorescent lamps originally installed in the building. Denoting f_k as the frequency of the signal emitted by LEDk, the respective frequencies for LED1, LED2, LED3, and LED4 101 Hz, 106 Hz, 113 Hz, 120 Hz, respectively. Because the frame rate f_c of the camera 50 Hz, the aliased frequencies were m_k = (1 Hz, 6 Hz, 13 Hz, 20 Hz),

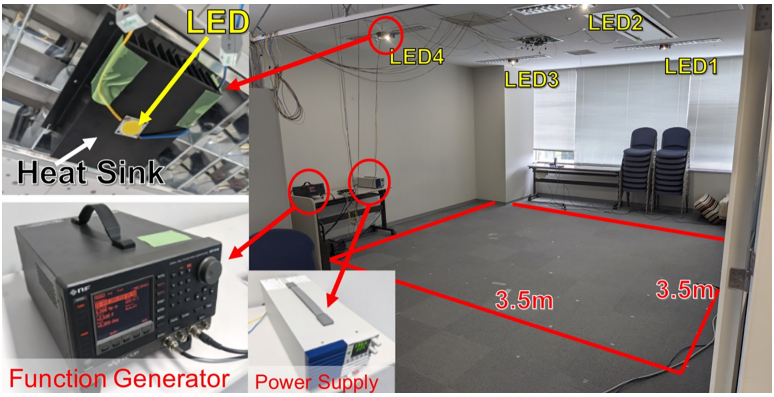


Fig. 6. Overview of the implementation, including LEDs, function generator, and power supply

Fig. 7. Floor material in our conference room

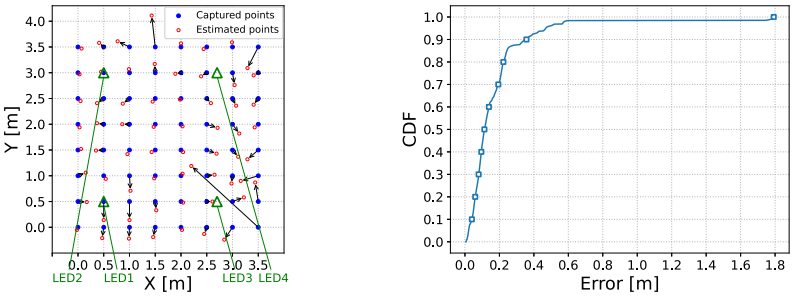
respectively. Each LED emitted a sine wave generated by delta–sigma modulation of a pulse-width-modulated signal. A 5-V signal from a function generator (NF Corporation WF-1948) was amplified to 39.60 V using a power driver and power supply.

4.2 Camera Receiver

We used a Point Grey Flea3 FL3-U3-13S2C camera [4] as the receiver and used the software Flycapture SDK2 [2] to set the camera parameters. The shutter speed and ISO rating were set to avoid luminance saturation, and the frame rate f_c was set 50 Hz. Since the modulation frequency of fluorescent lamps in eastern Japan 100 Hz, this could be treated as a DC component.

4.3 Floor Material

We used the unmodified floor of our conference room, which was covered with 50-cm square nonfluorescent mats without gaps, as shown in Fig. 7. The mat



(a) Positioning-error map and LED placement.

(b) Cumulative distribution function of positioning error.

Fig. 8. Map and cumulative distribution function of positioning error. (Color figure online)

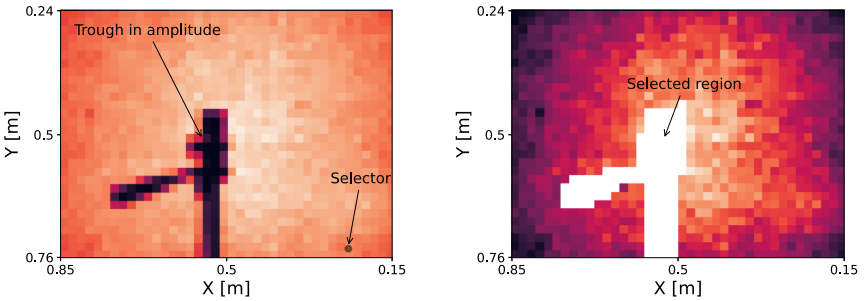
surface texture had a linear grain, whose direction alternated between vertical and horizontal for adjacent mats. Because Shimada et al. [18] had shown that the color of the floor surface can affect the accuracy of positioning, we used gray mats, because gray is the color that was found to be the least affected by noise.

5 Experimental Evaluation

Our experiments to investigate the effectiveness of the proposed method were conducted in a 3.5×3.5-m room. For all experiments, the same 1280×960-pixel camera was used and the luminance values used were calculated from the average of 32×32 pixels. Because the parameters C_k and σ_k in (11) were different for each LED, calibration was performed for (x,y) = (1.0, 0.5), (1.0, 2.0), (2.0, 2.0), (3.0, 1.0), and (3.0, 3.0). During the experiments, it was found that shadows were having a significant impact on accuracy. We therefore repeated the calculations after removing shadows manually from the images, to confirm that the accuracy was improved. Finally, we examined the relationship between the number of pixels used and the positioning accuracy.

5.1 Estimation Using Captured Floor Signals

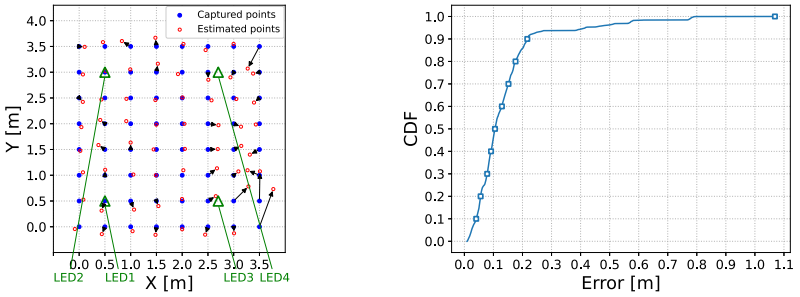
The images from the 1280×960-pixel image sensor were divided into 40×30 (=1200) blocks of 32×32 pixels, with a single luminance value being estimated as the average of the luminance values in each block. Because the distribution of luminance values can vary for each of the four LEDs, the Gauss–Newton method was applied to the total of $40 \times 30 \times 4 = 4,800$ blocks. The resulting solution was updated sequentially, until the solution converged to a point where the distance from the solution in the previous step was less than 0.001 (the "trough"). This was then considered the optimal solution. This process was performed at 64 points (8 points in the x-axis direction and 8 points in the y-axis direction),



(a) Trough in amplitude and shadow selector.

(b) Signal from removed part to be ignored in estimation.

Fig. 9. An example of selecting trough in amplitude



(a) Positioning error map after shadow removal and LED placement.

(b) Cumulative distribution function of positioning error after shadow removal.

Fig. 10. Map and cumulative distribution function of positioning error after shadow removal.

with the error between the obtained optimal solution and the actual coordinates being mapped in Fig. 8a. In this figure, the green triangles represent LEDs, the blue circles represent the coordinates of the actual position, and the red circles represent the average value of 100 positioning attempts.

In Fig. 8a, note that the central area, which is surrounded by LEDs, has higher positioning accuracy than the outer areas. One reason for this is that the positioning results are particularly susceptible to noise for the more remote LEDs. We also found that shadows and other objects have a significant impact on positioning accuracy. For example, the image at (1.5, 3.5) is about 30% covered by shadows, and the image at (3.5, 0.0) is about 50% covered by a pillar, with these two points having the worst overall positioning accuracy. The cumulative distribution function (CDF) of the positioning error is shown in Fig. 8b. The overall result of this experiment was to achieve a positioning error of 0.32 m at the 90th percentile.

Table 1. Trade-off between the number of pixels used and the 90th percentile of positioning error.

	1200 blocks = 1,228,800	300 blocks = 307,200	48 blocks = 49,152	12 blocks = 12,288	6 blocks = 6,144	2 blocks = 2,048	1 blocks = 1,024
$(\mathbf{x}, \mathbf{y}) = (1.5, 2.5)$	0.026 [m]	0.028	0.027	0.04	0.049	0.084	0.655
$(\mathbf{x}, \mathbf{y}) = (1.0, 2.0)$	0.022 [m]	0.020	0.024	0.028	0.049	0.048	0.48

5.2 Estimation Using Captured Floor Signals After Shadow Removal

As mentioned above, shadows and objects greatly reduce the accuracy of positioning. Here, we describe reapplying the proposed method after manually removing the shadows and showing that the accuracy improved as a result.

When shadowing is present, some of the light from the LEDs is blocked and attenuated significantly, resulting in a trough in the amplitude of the luminance. Figure 9a shows the observed amplitude trough and the region selector used to exclude areas with troughs. The luminance values of the removed regions are masked as shown in Fig. 9b and are excluded when applying the Gauss–Newton method. The difference between the optimal solution and the actual coordinates obtained in this way is mapped in Fig. 10a. In this figure, note that the positioning results at the two points (1.5, 3.5) and (3.5, 0.0) were greatly improved. In addition, positioning in the neighborhood of an LED tended to improve the accuracy. This may be because positioning near the LED is more likely to be affected by the shadow of the camera tripod. The CDF of the positioning error is shown in Fig. 10b. As a result, a 90th percentile of 0.22 m was achieved, with the positioning accuracy being improved by nearly 0.1 m after shadow removal. This result confirms that the proposed method is sensitive to shadows and objects, but it also shows that the positioning accuracy can be improved substantially by shadow removal.

5.3 Effect of the Number of Blocks Used

Finally, we investigated the relationship between the number of pixels used and the positioning accuracy. In the previous experiments, the received image was divided into $30 \times 40 = 1200$ blocks. In this experiment, the number of blocks was reduced while maintaining the size of each block (32×32 pixels). The two locations, (1.5, 2.5) and (1.0, 2.0) were unaffected by shadows or objects and were selected for the experiment. Table 1 shows the 90th percentile of the positioning error when the number of pixels is reduced from 1,228,800 to 1,024. Although the positioning accuracy tends to decrease as the number of pixels used decreases, the 90th percentile of positioning error using two blocks remains within 10 cm. This shows that it is possible to implement a positioning system that can operate in real time by reducing the number of pixels used and the computational cost, if necessary.

6 Discussion

6.1 Estimation of Camera Posture

In our proposed method, we obtained the 2D coordinates of the camera without consideration of the camera posture. However, a theoretical analysis is possible for cases where the pitch angle of the camera is considered. In such cases, the camera's pitch angle θ can be varied freely while maintaining the level. The constraint in (13) is then

$$D^k_{u,v,\theta} = \sqrt{(x_k - X^\theta_{u,v})^2 + (y_k - Y^\theta_{u,v})^2}, \tag{16}$$

where:

$$X^\theta_{u,v} = x + W_x u \cos\theta - W_y v \sin\theta \tag{17}$$

$$Y^\theta_{u,v} = y + W_x u \sin\theta + W_y v \cos\theta. \tag{18}$$

It should also be possible to include the roll and yaw angles in the calculation, but we would then not be able to use the approximation that all pitches have the same value and $W_x = W_y$. The specific mathematical formulas are omitted here because of the complexity of the calculations but, theoretically, it should be possible to estimate the camera's posture.

6.2 Power Consumption

The power consumption of the camera we used in our experiments was 3 W. Although the full image-sensing capacity of the camera was used in the proposed method, sufficient positioning accuracy should be achieved even with a much smaller number of pixels, as described in Sect. 5.3. If only the necessary minimum of the image sensor capacity is activated and those signals are processed, the required power consumption can be reduced.

Table 2. Comparisons with proposed VLP method using LEDs

	RefRec	Luxapose	PIXEL	Rajagopal	STARLIT	Proposed
Accuracy	∼0.4 m	∼0.1 m	∼0.3 m	N/A	∼0.75 m	∼0.2 m
Reference	[17]	[10]	[21]	[15]	[20]	
Range [m]	4.0×4.0	1.0×1.0	2.4×1.8	3.9×8.0	7.0×10.0	3.5×3.5
Rolling shutter	No	Yes	No	Yes	Yes	No
Real time	Yes	No	Yes	No	Yes	No
LEDs	4	5	8	4	1	4
Resolution	32×32	7712×5360	120×160	1280×720	1080×1920	1280×960
LoS/NLoS	NLoS	LoS	LoS	NLoS	NLoS	NLoS

7 Related Work

Table 2 compares existing VLP methods with our proposed method. In this table, the term "Accuracy" refers to the 90th percentile of the positioning error. "Rolling shutter" refers to the rolling shutter effect in smartphones. (Customers often do not like the distortion of the image caused by the rolling shutter, which may disappear in the future [5].) "LoS/NLoS" (line of sight/non-line of sight) indicates whether the LED signals are assumed to be acquired directly (LoS) or indirectly (NLoS).

RefRec [17] also utilized floor-reflected light and uses only 32×32 pixels at the center of the screen. It is more sensitive to shadows and objects than our method, and less likely to be applied to camera posture estimation. RefRec achieved a 90th percentile error of 0.42 m in a 4.0×4.0-m conference room. We achieved better positioning accuracy since our method is more generalized. However, RefRec is able to perform real-time positioning.

Luxapose [10] can compute the position and posture of a smartphone by capturing ceiling light signals directly using a camera. The error is less than 10 cm and less than 3°. It uses 7712×5360 pixels in a WindowsPhone 8 smartphone camera as the receiver. The calculation requires a cloud server for high-quality image processing.

PIXEL [21] is a polarization-based localization method. Only 120×160 pixels are required. Measurements require several seconds, and the accuracy is 0.4 m. However, a polarizing filter must be attached to the camera, with the accompanying risk of impairing the original image.

Rajagopal's approach [15] uses light reflected from the floor. It is similar to our idea, but they focus on the rolling shutter distortion to receive the LED's ID from the reflected light. Carrier signals up to 8 kHz can be received with a channel separation 200 Hz. Tag information is transmitted by assigning ON and OFF bits to the different frequencies. The data rate is 10 bps, and up to 29 light sources can be uniquely separated. Positioning accuracy was not discussed in the paper because the research aimed at semantic positioning using differences in packet reception rates. Furthermore, because MATLAB [3] was used for calculation, processing was not in real time.

STARLIT [20] is a VLP system using floor reflections and utilizing RSS. To capture the ID tag from a single LED, the authors leverage the rolling shutter mechanism in a smartphone camera. STARLIT achieved an 80th percentile error of 55 cm in a 72-m^2 room and a response time of 0.87 s on average. They also noted that vertically spreading shadows can cause failures in positioning.

8 Conclusion and Future Work

VLP not only promises to play an active role in the future but is already expanding its use in a variety of fields. In this paper, we propose a new reflected-light-based positioning method using pattern matching across the entire image-capture screen. Using a prototype system, we achieved a 90th percentile error

of 0.22 m after removing shadow effects, demonstrating the effectiveness of the method. We also investigated the trade-off between positioning accuracy and the number of pixels used. These investigations show that our proposed method can achieve high positioning accuracy in NLoS environments. Furthermore, we have shown that it has the potential to achieve positioning with less computational effort and power consumption. In future work, we aim to extend the proposed method to posture estimation and real-time positioning.

Acknowledgement. This research was supported by JSPS Kakenhi Grant Numbers 19H04222 and 20K21781.

References

1. Bridgelux led lighting v series v18b-gen8/v18c-gen 8. https://www.bridgelux.com/products/v-series#specifications. Accessed 19 July 2021
2. Flycapture sdk homepage. https://www.flir.com/products/flycapture-sdk/. Accessed 19 July 2021
3. Matlab homepage. https://www.mathworks.com. Accessed 19 July 2021
4. Point grey research flea3 fl3-u3-13s2c specifications. http://sine.ni.com/apps/utf8/nipc.specs?action=view_specs&asid=1102&pid=11401&tier=3. Accessed 19 July 2021
5. Sony Develops the Industry's First*1 3-Layer Stacked CMOS Image Sensor with DRAM for Smartphones. https://www.sony.net/SonyInfo/News/Press/201702/17-013E/. Accessed 19 July 2021
6. Indoor Location Based Services Market Research Report by Technology, by Industry, by Application - Global Forecast to 2026 - Cumulative Impact of COVID-19. https://www.researchandmarkets.com/reports/4896758/indoor-location-based-services-market-research (2021). Accessed 19 July 2021
7. De Lausnay, S., De Strycker, L., Goemaere, J.P., Nauwelaers, B., Stevens, N.: A survey on multiple access visible light positioning. In: 2016 IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech), pp. 38–42 (2016). https://doi.org/10.1109/EmergiTech.2016.7737307
8. Groves, P.D.: Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems, 2nd edn. [book review]. IEEE Aerosp. Electron. Syst. Maga. **30**(2), 26–27 (2015). https://doi.org/10.1109/MAES.2014.14110
9. Hossain, A.M., Soh, W.S.: A survey of calibration-free indoor positioning systems. Comput. Commun. **66**, 1–13 (2015)
10. Kuo, Y.S., Pannuto, P., Hsiao, K.J., Dutta, P.: Luxapose: indoor positioning with mobile phones and visible light. In: Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, pp. 447–458. Association for Computing Machinery, Maui (2014). https://doi.org/10.1145/2639108.2639109
11. Li, L., Hu, P., Peng, C., Shen, G., Zhao, F.: Epsilon: a visible light based positioning system. In: Proceedings of the 11th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2014, pp. 331–343 (2014)
12. Lymberopoulos, D., Liu, J., Yang, X., Choudhury, R.R., Handziski, V., Sen, S.: A realistic evaluation and comparison of indoor location technologies: experiences and lessons learned. In: Proceedings of the 14th International Conference on Information Processing in Sensor Networks, pp. 178–189. Association for Computing Machinery, Seattle (2015). https://doi.org/10.1145/2737095.2737726

13. Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. **11**(2), 431–441 (1963)
14. Pavel, D., Robert, P.: A survey of selected indoor positioning methods for smartphones. IEEE Commun. Surv. Tutorials **19**(2), 1347–1370 (2016)
15. Rajagopal, N., Lazik, P., Rowe, A.: Visual light landmarks for mobile devices. In: IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks, pp. 249–260 (2014). https://doi.org/10.1109/IPSN.2014.6846757
16. Shannon, C.E.: Communication in the presence of noise. Proc. IRE **37**(1), 10–21 (1949)
17. Shimada, S., Hashizume, H., Sugimoto, M.: RefRec: indoor positioning using a camera recording floor reflections of lights (2020)
18. Shimada, S., Hashizume, H., Sugimoto, M.: Indoor positioning using reflected light and a video camera. In: Proceedings 9th International Conference on Indoor Positioning and Indoor Navigation, Nantes, France, pp. 1–8 (2018)
19. Wilkins, A., Veitch, J., Lehman, B.: Led lighting flicker and potential health concerns: IEEE standard par1789 update. In: 2010 IEEE Energy Conversion Congress and Exposition, pp. 171–178 (2010). https://doi.org/10.1109/ECCE.2010.5618050
20. Yang, F., Li, S., Zhang, H., Niu, Y., Qian, C., Yang, Z.: Visible light positioning via floor reflections. IEEE Access **7**, 97390–97400 (2019). https://doi.org/10.1109/ACCESS.2019.2929160
21. Yang, Z., Wang, Z., Zhang, J., Huang, C., Zhang, Q.: Polarization-based visible light positioning. IEEE Trans. Mob. Comput. **18**(3), 715–727 (2019). https://doi.org/10.1109/TMC.2018.2838150
22. Zafari, F., Gkelias, A., Leung, K.K.: A survey of indoor localization systems and technologies. IEEE Commun. Surv. Tutorials **21**(3), 2568–2599 (2019). https://doi.org/10.1109/COMST.2019.2911558
23. Zhang, H., Yang, F.: Push the limit of light-to-camera communication. IEEE Access **8**, 55969–55979 (2020)