https://doi.org/10.7840/kics.2023.48.11.1378

A Vision-Based Ranging Algorithm Combining Monocular Camera and Laser for Indoor Positioning

Nan Yin◆, Yuxiang Sun*, Zhengyang Zou**, Jae-soo Kim°

ABSTRACT

The core issue for indoor positioning is reducing the ranging error. The traditional Time-of-Flight (ToF) method has errors caused by time error calculation, and the Received Signal Strength Indicator (RSSI) method requires huge calculation and data collection. Moreover, the vision-based binocular ranging method has poor performance in dark environments. Therefore, considering the above weaknesses, we designed a hybrid device based on a Monocular Camera and four Lasers (MC4L), which uses four lasers to mark the target object and record the irradiation spot with a high-definition monocular camera. Due to the angle between the laser emitter and the base, the center point and laser irradiation area will change with distance. We can obtain the corresponding relationship between the irradiated area and real distance by a logarithmic regression algorithm. In addition, we propose a new indoor positioning algorithm based on the MC4L device to control the error within 2.4 cm.

Key Words : Monocular ranging, Binocular ranging, Logarithmic regression, Time-of-Flight

Ⅰ. Introduction

Indoor positioning technology is playing an increasingly important role in daily life and business. However, in contrast to outdoor environments, indoor positioning faces unique challenges, such as the instability of Global Positioning System (GPS) signals, multipath effects, and signal interference. To address these issues, researchers have explored various innovative localization methods and algorithms to achieve accurate real-time location recognition in indoor environments. As a cutting-edge technology in the field of indoor positioning, the monocular ranging algorithm can infer the distance between a target object and camera by analyzing image information from a single camera. Compared with traditional multi-sensor methods, monocular ranging algorithms have obvious advantages in terms of hardware requirements, cost, and deployment. In addition, with the rapid development of fields such as computer vision and deep learning, the accuracy and robustness of monocular ranging algorithms have provided new possibilities for indoor positioning.

In this study, we adopted a combination of a monocular camera and four lasers to measure distance in a dark environment. Distance measurements were conducted by tracking the area change of the four laser irradiation zones in the image. Finally, the linear relationship between the real distance and laser irradiation area was explored using the designed ranging algorithm. The experimental results show that the proposed method has a minimum error of 1.6 cm and average error of 2.4 cm within three meters. The contributions of this study can be summarized as

◆ First Author : Kyungpook National University, Computer Science and Engineering, yinnan1990@gmail.com, 정회원
° Corresponding Author : Kyungpook National University, Computer Science and Engineering, kjs@knu.ac.kr, 종신회원
* Kyungpook National University, Software Technology Research Center, yxsun@knu.ac.kr, 정회원
** Kyungpook National University, Department of French Language and Literature, zzy@knu.ac.kr, 학생회원
논문번호 : 202306-123-B-RN, Received June 13, 2023; Revised August 14, 2023; Accepted August 21, 2023

follows:

1. We designed a measurement device called MC4L, which combines a monocular camera and four laser emitters and costs less than $ 9.
2. We designed a ranging algorithm based on MC4L that extracts the irradiation area. Finally, the logarithmic relationship is established between the real distance and irradiation area. We designed an indoor positioning algorithm based on MC4L to reduce the minimum error to 1.6 cm.
3. Our model only needs to be trained once and does not require any parameters during the calculation. This approach is more robust than other approaches.

The remainder of this paper is organized as follows. In Section II, we analyze existing positioning technologies, indoor positioning approaches, and ranging approaches. In Section III, we introduce the architecture of the MC4L device. In addition, MC4L-based ranging and indoor positioning algorithms are designed. Section IV presents the performance evaluation, including error evaluation and selection of exponent term numbers. Finally, conclusions and future research directions are presented in Section V.

Ⅱ. Related works

2.1 Positioning technologies

Indoor positioning algorithms are designed to determine the location of objects within an indoor environment, such as a building or room. These algorithms use various sensors, including Wi-Fi, Bluetooth, and Radio-Frequency Identification (RFID). Some indoor positioning algorithms also use visual cues, such as cameras or markers, to determine object positions. The Wi-Fi Positioning System (WPS) is a popular indoor positioning algorithm. WPS uses the strength of Wi-Fi signals to determine the position of an object. The system can calculate the position of the object with reasonable accuracy by comparing the strengths of Wi-Fi signals at different locations. Another popular indoor positioning algorithm is Bluetooth Low Energy (BLE) positioning. BLE positioning uses Bluetooth signals instead of Wi-Fi signals to determine the position of an object.

By contrast, outdoor positioning algorithms are designed to determine the location of an object within an outdoor environment, such as a city or rural area. The most well-known outdoor positioning technology is the Global Positioning System (GPS), which uses a satellite network to determine the location of an object. GPS is widely used for outdoor navigation and location-based services. However, GPS has some limitations, particularly in urban environments, where tall buildings can block satellite signals. To overcome these limitations, researchers have proposed alternative outdoor positioning technologies, such as Assisted GPS (A-GPS), which uses network-based positioning to supplement GPS signals.

Hybrid positioning algorithms combining indoor and outdoor positioning technologies have also been developed. These algorithms use a combination of Wi-Fi, GPS, and other sensors to determine the position of an object. Hybrid positioning algorithms can provide more accurate positioning information by combining the advantages of both indoor and outdoor positioning technologies.

2.2 Indoor positioning approaches

A 3D local reference image database is generated by utilizing a dynamic depth acquisition backpack for the 3D modeling of indoor environments[1]. The entire process is mainly divided into Image Retrieval and Pose Estimation (IRPE). A limitation of IRPE is that it requires many images to improve accuracy. To reduce the number of computations, the Simultaneous Localization and Mapping (SLAM)[2] approach was proposed, which can run on a device with a 7-processor CPU and 64 kb RAM. Spread Spectrum Ultrasonic Waves (SSUw)[3] has also been adopted for indoor positioning. The disadvantage of SSUw positioning technology is that it requires a fixed positioning network and three positioning anchor points. Ultra-Wide Band (UWB) is based on three-point positioning; thus, if the number of

positioning base stations is reduced, the positioning accuracy will also be affected. In addition, it can only exhibit excellent performance over short distances (~10 m) and is significantly affected by temperature changes. A novel fingerprint algorithm based on Exponential Effective Signal Mapping (EESM) improves positioning performance by involving channel estimation to create more robust fingerprints[4]. The disadvantage of a fingerprint-based algorithm is that it requires a large amount of data collection in the early stages, and it is difficult to adapt in scenarios with large environmental changes. Hence, to increase computing efficiency, cloud computing is adopted to store and process data. Peng et al.[5] proposed an algorithm based on Iterated Reduced Sigma Point Kalman Filter (IRSPKF) to complete object positioning at a lower computational cost in complex non-linear systems. Subhan et. al.[6] found a relationship between the received power level and distance using a standard radio propagation model. The above-mentioned EESM, IRSPKR and Bluetooth Fingerprinting (GF)[7] are RSSI-based positioning approaches, while SSUw and SLAM are Time-of-Flight (ToF)-based positioning technologies. The IRPE and MC4L methods used in this study can be classified as vision-based positioning approaches. The aforementioned RSSI-based and ToF-based methods can ignore the interference of indoor obstacles. However, the MC4L-based indoor positioning algorithm has difficulty distinguishing whether the front is a wall or an obstacle.

2.3 Ranging approaches

Existing mainstream ranging approaches are mainly divided into ToF- and parallax-based. The principle of ToF is to utilize the round-trip time difference between two transceivers to calculate the object distance, as shown in Equation (1), where c represents the signal speed and t represents the time difference. The ToF method has the advantages of high accuracy (millimeter level), fast measurement speed, and low power consumption. However, it is extremely susceptible to disturbances, such as surface reflections and light angles, in outdoor environments. Moreover, ToF sensors can be complex and expensive, requiring advanced electronics and signal-processing algorithms. Therefore, vision-based ranging algorithms have attracted increasing attention in recent years.

$$distance = \frac{c \times t}{2} \tag{1}$$

Vision-based ranging approaches include monocular ranging[8] and binocular ranging[9-11]. In principle, binocular ranging is similar to how the human eye works. It utilizes the visual difference between two collected images to calculate the distance between itself and obstacles. Fig. 1 shows a schematic of binocular ranging. The distance between the projection centers of the two cameras is the baseline distance and is expressed as B. When the target point T passes through the binocular ranging system composed of two left and right cameras with parallel optical axes, it is imaged at point T^1 of the left camera and point T^2 of the right camera, and its imaging positions are T^d_{left} and T^d_{right}, respectively. Assuming that the focal lengths of the two cameras are both f, the measured distance l can be deduced according to the triangle similarity principle as follows:

$$l = \frac{Bf}{T^d_{left} + T^d_{right}} \tag{2}$$

According to the vision-based ranging principle mentioned above, we can determine whether the target point T can be clearly imaged by the left and right cameras for accurate measurement. In other words,

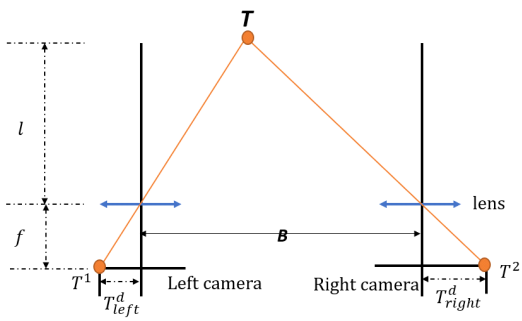

Fig. 1. Schematic of binocular ranging.

calculating parallax in a dark environment is difficult. The monocular ranging method performs better than the stereovision-based ranging method because it does not need to match images in the data preprocessing stage[12]. Compared to binocular ranging, monocular distance measurement has lower environmental requirements and computational consumption. It typically uses tracking changes in image features to calculate the distance between the object and itself.

Ⅲ. Ranging and Indoor Positioning Algorithms

3.1 Ranging Devices

Fig 2 shows the architecture of the designed MC4L device. The device consists of four laser transmitters and a high-definition monocular camera. The four laser emitters are evenly distributed on a circular track, the camera is placed at the center of the circular track, and the track radius is 7.75 cm. The laser transmitters are KY-008, and the camera is a Centechia Webcam. The target object is placed at 120 cm and measured every 5 cm; the entire test distance is 3 m. Capture images are 640 × 480 pixels.

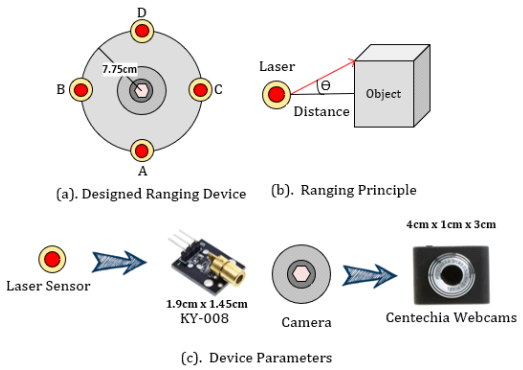

Fig. 2. Architecture of ranging devices.

3.2 Ranging Algorithm

3.2.1 Binarization process

The entire ranging algorithm is divided into four steps. The first step is the binarization process shown in Equation (3). The purpose of the binarization process is to better extract target objects from the original image. x and y in Equation (3) represent the coordinates of the pixel values, and T is a threshold. IFAD the pixel value exceeds the threshold T, it will be set to white; otherwise, it will be black. g represents the input image, and G represents the output image.

$$G(x,y) = \begin{cases} 255, & if \;\; g(x,y) \geq T \\ 0, & otherwise \end{cases} \tag{3}$$

3.2.2 Erosion operation

An erosion operation is used to remove noise from the binary image. The purpose of erosion is to reduce the size of the foreground object and isolate the individual elements of objects. Erosion splits the joined objects apart,, augments the distance between them, or removes some redundant pixels. Equation (4) describes the erosion operation, where (x, y) represents the pixel values of an image. SE is a structuring element with a 3×3 rectangular matrix, and *Eroded_pixel* is the output image after erosion. The erosion operation is performed by sliding a Structuring Element (*SE*) over the image, and the minimum value of the corresponding pixels in the neighborhood is calculated. The minimum value is assigned to the center of the structuring element.

$$Eroded_{pixel(x,y)} = min\{original\_pixel(x+i, y+i) | (i,j) \in SE\} \tag{4}$$

3.2.3 Dilation operation

Dilation operations are used to increase the size of objects in a binary image. This was performed using a 3×3 *SE*, which is a small matrix that defines the neighborhood of each pixel in the eroded image. The dilation formula is as follows:

$$Dilated_{pixel(x,y)} = min\{eroded\_pixel(x+i, y+i) | (i,j) \in SE\} \tag{5}$$

3.2.4 Laser points coordinate extraction and area calculation

Equations (4) and (5) represent image opening processing, with the purpose of removing small and isolated pixels (noise) from the original images and preparing them for extracting laser point coordinates.

Next, we iterate the list of contours and extract the coordinates of the points in each contour. Sub-figure *d* in Fig 3 shows an image acquiring the laser point coordinates and marking them in the original image. The *a*, *b*, *c* and *d* represent the four quadrilateral sides. According to Heron's formula shown in equation (6) and (7), we can obtain the quadrilateral area *S*.

$$z = \frac{a+b+c+d}{2} \tag{6}$$

$$S = \sqrt{(z-a)\times(z-b)\times(z-c)\times(z-d)} \tag{7}$$

3.2.5 logarithmic regression between area and real distance

Logarithmic regression is nonlinear, and the relationship can be modeled using logarithmic transformation. As shown in equation (8), we utilize a cubic logarithmic regression algorithm to find the weight values between area *S* and the real distance *y*, where *b* is the bias. The logarithmic regression model can be written as

$$y = \beta_3 \ln(s)^3 + \beta_2 \ln(s)^2 + \beta_1 \ln(s) + \varepsilon \tag{8}$$

where *y* is real distance, *s* is quadrilateral area, β_3, β_2, and β_1 are the coefficients of the cubic regression equation, and ε is the bias. The coefficients β_3, β_2, and β_1 are estimated using the least squares method.

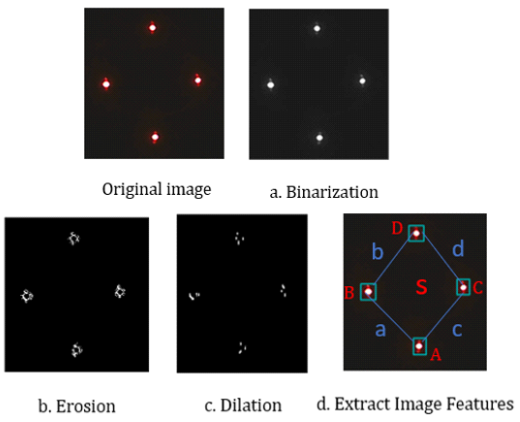

Fig. 3. Flowchart of proposed ranging algorithm.

3.2.6 Indoor Positioning Algorithm

In this section, the experimental environment and parameters are combined to describe the indoor positioning algorithm in detail. Fig 4 shows a 2D diagram of the experimental environment. We randomly placed a target object cube *ABCD* with a length, width, and height of 15 cm in an experimental environment of 3 × 3 m. The perimeter of the cube *ABCD* is placed on four designed ranging devices: *a*, *b*, *c*, and *d*. We can obtain four vertex coordinates (A_x, A_y), (B_x, B_y), (C_x, C_y), and (D_x, D_y) by utilizing the proposed ranging algorithm.

Fig 5 shows the proposed indoor positioning algorithm based on MC4L (Algorithm 1). β and θ are the width and height of the target object *ABCD*, respectively. First, we obtain the distances of $DA(D_Y - A_y)$ and $CB(C_Y - B_y)$ and compare the error between them and object width $\beta(15cm)$ shown in line 1. If the error of *DA* is greater than that of *CB*, then the center point *c* of *CB* will be the ordinate of the target object *ABCD* shown in line 2. Otherwise, the center point *a* of *DA* will be used as the ordinate of the target object *ABCD* shown in line 4. The abscissa of *ABCD* can be obtained by repeating the same operations from lines 5 to 8. Finally, the center point coordinates (p_x, p_y) of *ABCD* can be obtained. Limited by the ranging mode of the MC4L device, the indoor positioning algorithm cannot distinguish whether the measured object is a wall or obstacle. Hence, it can only be positioned in an open environment without any obstacles.

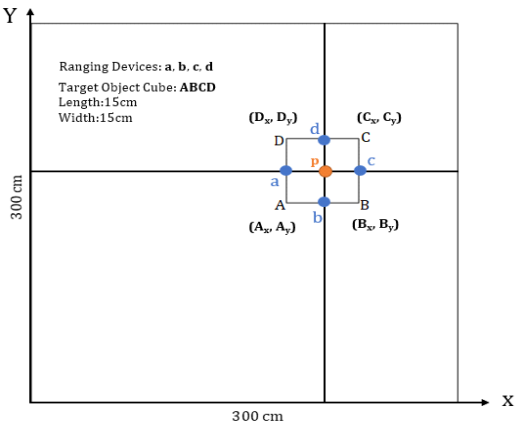

Fig. 4. 2D diagram of experimental environment.

Algorithm 1: Indoor Positioning Algorithm

Input: Vertex coordinates (A_x, A_Y), (B_x, B_Y), (C_x, C_Y) and (D_x, D_Y)

Output: Target object coordinate (p_x, p_y)

if $|(D_y - A_y) - \beta| > |(C_y - B_y) - \beta|$, **then**
  $p_y \leftarrow (C_y - B_y)/2 + B_y$
else
  $p_y \leftarrow (D_y - A_y)/2 + A_y$
if $|(D_x - A_x) - \theta| > |(C_x - B_x) - \theta|$, **then**
  $p_x \leftarrow (C_x - B_x)/2 + B_x$
else
  $p_x \leftarrow (D_x - A_x)/2 + A_x$
Return target object coordinate (p_x, p_y)

Fig. 5. Indoor positioning algorithm based on MC4L.

Ⅳ. Performance Evaluations

4.1 Error evaluation

Fig. 6 shows the errors of indoor positioning based on different devices and positioning algorithms. The experimental results show that the proposed MC4L-based positioning algorithm has the smallest error. Compared with ToF-based SLAM and RSSI-based EESM, its accuracy increased by 84.3% and 94.4%, respectively. Overall, TOF-based methods are better than RSSI-based positioning algorithms. Compared with the IRPE method, which is also based on vision, our proposed MC4L-based ranging algorithm is still 97.1% better. The indoor positioning error can be controlled between 1.6 cm and 2.4 cm based on our proposed MC4L device and indoor positioning algorithm.

Fig. 7 shows the error optimization based on the logarithmic regression model. The X- and Y-axes represent the laser irradiation area and real distance, respectively. The red line indicates the model fitting

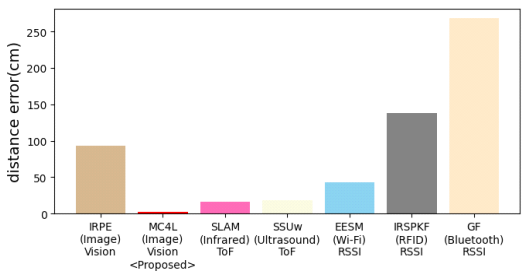

Fig .6. Error comparison of ranging methods based on different devices.

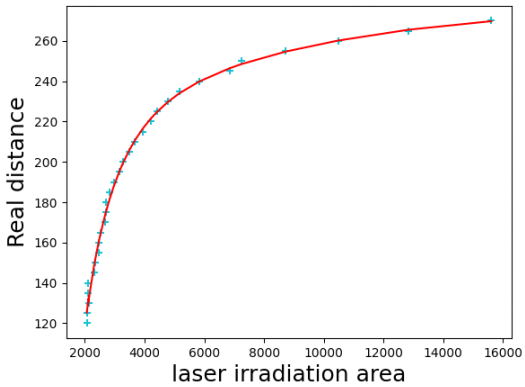

Fig. 7. Error optimization based on logarithmic regression.

effect. The MC4L equipment has a sufficient measurement accuracy within three meters of the fitting effect.

Fig 8 shows the error for different distances. First, the error distribution is controlled within a certain interval and does not increase with distance. It was verified that the proposed algorithm has a certain robustness. Second, the distribution of errors is relatively average and maintained within 3 cm, and the controllability is relatively strong.

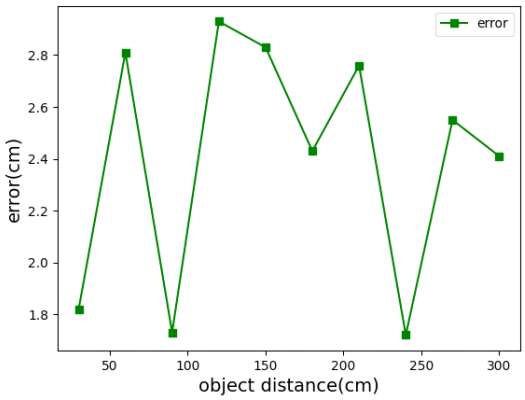

Fig. 8. Error at different distances from the target object.

4.2 Selection of exponent term numbers for logarithmic regression model

The choice of exponent term number is crucial for the logarithmic regression algorithm. To select a suitable model more efficiently, the determination coefficient (R^2) was used as an evaluation benchmark, as shown in Equation (9). R^2 is also called the

goodness of fit, which reflects the degree to which the independent variable x explains changes in the dependent variable y. The closer it is to 1, the better the fitting effect of the logarithmic regression algorithm.

$$R^2(y,\hat{y}) = 1 - \frac{RSS}{TSS} \tag{9}$$

$$TSS = \sum_{i=0}^{n}(y_i - \bar{y}_i)^2 \tag{10}$$

$$RSS = \sum_{i=0}^{n}(y_i - \hat{y}_i)^2 \tag{11}$$

where Total Sum of Squares (*TSS*) indicates the degree of change in y, which is proportional to the variance. The Residual Sum of Squares (*RSS*) represents the residual between the proposed model and real values. Fig. 9 shows the selection of exponent term number, where the X-axis represents the number of exponent terms, and the Y-axis is the determination coefficient R^2. When the number of exponent terms is 3, the determination coefficient is closest to 1.

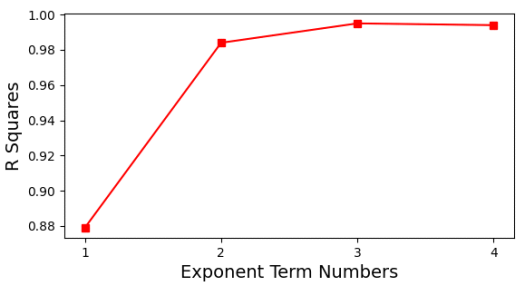

Fig. 9. Selection of exponent term number.

V. Conclusions

In this study, we designed an MC4L device and proposed a new ranging and indoor positioning algorithm based on this device. The error of indoor positioning was reduced to 2.4 cm by combining it with the logarithmic regression model. We not only solved the measurement error caused by the time calculation for ToF-based methods but also reduced the calculation cost compared to RSSI-based methods. In a recent study, Fan et. al[13] proposed an automatic auto-correction method combined with a known point location. In future research, we plan to adapt and optimize this algorithm with MC4L.

References

[1] J. Z. Liang, N. Corso, E. Turner, and A. Zakhor, "Image based localization in indoor environments," in *Proc. Computing for Geospatial Res. and Appl.*, pp. 70-75, San Jose, CA, USA, Jul. 2013. (https://doi.org/10.1109/COMGEO.2013.11)

[2] J. S. Gutmann, P. Fong, L. Chiu, and M. E. Munich, "Challenges of designing a low-cost indoor localization system using active beacons," *Int. Conf. Technol. for Practical Robot Applications*, pp. 22-23, Woburn, MA, USA, Apr. 2013. (https://doi.org/10.1109/TePRA.2013.6556348)

[3] Y. Itagaki, A. Suzuki, and T. Iyota, "Indoor positioning for moving objects using a hardware device with spread spectrum ultrasonic waves," *Int. Conf. Indoor Positioning and Indoor Navig.*, pp. 13-15, Sydney, Australia, Nov. 2012. (https://doi.org/10.1109/IPIN.2012.6418850)

[4] F. Wang, Z. Huang, H. Yu, X. Tian, X. Wang, and J. Huang, "EESM-Based fingerprint algorithm for wi-fi indoor positioning system," *Int. Conf. Commun. in China*, pp. 674-679, Xi'an, China, Aug. 2013. (https://doi.org/10.1109/ICCChina.2013.6671197)

[5] J. Peng, M. Zhu, and K. Zhang, "New algorithms based on sigma point kalman filter technique for multi-sensor integrated rfid indoor/outdoor positioning," *Int. Conf. Indoor Positioning and Indoor Navig.*, pp. 21-23. Guimarães, Portugal, Sep. 2011.

[6] F. Subhan, H. Hasbullah, A. Rozyyev, and S. T. Bakhsh, "Indoor positioning in bluetooth networks using fingerprinting and lateration approach," *Int. Conf. Inf. Sci. and Appl.*, pp. 29-35, Jeju, Korea, Apr. 2011.

(https://doi.org/10.1109/ICISA.2011.5772436)

[7] L. Zhang, X. Liu, J. Song, C. Gurrin, and Z. Zhu, “A comprehensive study of bluetooth fingerprinting-based algorithms for localization,” *Int. Conf. Advanced Inf. Netw. and Appl. Wkshps.*, pp. 300-305, Barcelona, Spain, Mar. 2013. (https://doi.org/10.1109/WAINA.2013.205)

[8] P. Ferrara, et al., “Wide-angle and long-range real time pose estimation: A comparison between monocular and stereo vision systems,” *J. Vis. Commun. Image Represent.*, vol. 48, pp. 159-168, 2017. (https://doi.org/10.1016/j.jvcir.2017.06.008).

[9] L. Xue, M. Li, L. Fan, A. Sun, and T. Gao, “Monocular vision ranging and camera focal length calibration,” *J. Scientific Programming*, vol. 2021, no. 9979111, pp. 1-15, 2021. (https://doi.org/ 10.1155/2021/9979111).

[10] Q. Mu, J. Wei, Z. Yuan, and Y. Yin, “Research on target ranging method based on binocular stereo vision,” in *Proc. Int. Conf. Intell. Computing, Automat. and Appl.*, pp. 81-85, 2021. (https://doi.org/10.1109/ICAA53760.2021.00023).

[11] D. Haifeng and Y. Jun, “Research on robot binocular ranging based on SIFT feature extraction algorithm,” *J. Physics: Conf. Series*, vol. 1607, no. 1, pp. 1-7, 2020. (https://doi.org/10.1088/1742-6596/1607/1/012015).

[12] C. Lu and X. Tang, “Binocular vision-based recognition method for table tennis motion trajectory,” *J. Mobile Inf. Syst.*, vol. 2022, pp. 1-8, 2022. (https://doi.org/10.1155/2022/2093631).

[13] Z. Fan and J. Kim, “A weighted quadrangle positioning algorithm using automatic distance correction for optimizing indoor positioning accuracy,” *J. KICS*, vol. 46, no. 03, pp. 518-525, 2020. (https://doi.org/10.7840/kics.2021.46.3.518)

인 난 (Yan Yin)

2013 : Bachelor’s, Process Equipment and Control Engineering, Beijing University of Chemical Technology

2015 : M.Eng Mechanical Engineering, University of Ottawa

2022~Current : Ph. D. School of Computer of Kyungpook National University

<Research Interests> Machine Learning, Artificial Intelligence, Internet of Things

[ORCID:0009-0007-3506-9592]

순 위 샹 (Yuxiang Sun)

2016 : Bachelor’s, Computer Engineering, Andong National University

2019 : M.Sc, Computer Science and Engineering, Kyungpook National University

2022 : Ph. D. Computer Science and Engineering, Kyungpook National University

2023~Current : Researcher, Software Technology Research Center of Kyungpook National University

<Research Interests> Semantic Web, Entity Alignment, Machine Learning

[ORCID:0000-0003-0165-7664]

쩌우쩡양 (Zhengyang Zou)

2022 : Bachelor’s, Department of French Language and Literature, Kyungpook National University

<Research Interests> Signal Processing, Machine Learning, Internet of Things

[ORCID:0009-0009-6048-9614]

김 재 수 (Jaesoo Kim)

1985 : Bachelor's, Electronic Engineering, Kyungpook National University

1987 : M.Sc Computer Science, Joong-Ang University

1999 : Ph. D. Computer Engineering, Kyungnam University

1987~1996 : Senior Researcher, Korea Electrical Research Institute

2003~2004 : Visiting Professor, The University of Cincinnati, OH, USA

1996~Current : Professor, School of Computer Science and Engineering, Kyungpook National University

<Research Interests> Mobile Computing, Sensor Network, Internet of Things, UAV Network

[ORCID:0000-0003-2541-1669]