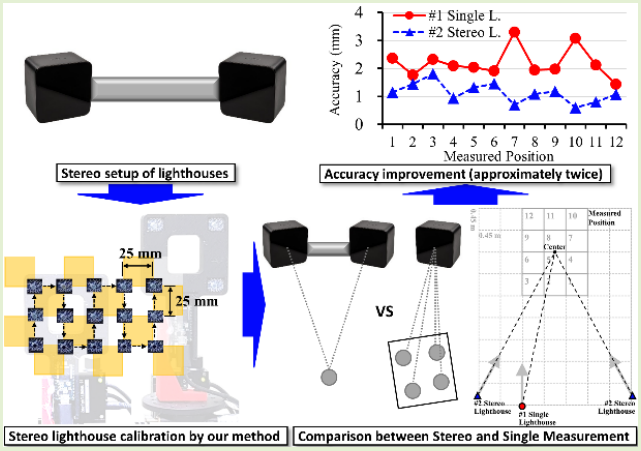
A Comparative Evaluation of a Single and Stereo Lighthouse Systems for 3-D Estimation

Sungtaek Cho, Dongyeon Kim, and Sungon Lee, *Member, IEEE*

***Abstract*—The lighthouse localization system has recently been developed and used for localization in virtual reality (VR). Not only for VR but also for a general indoor positioning systems (IPSs) it has several advantages over existing methods, including low cost, wide detection area, and easy setup. Here, we adopt the stereo configuration of a lighthouse for improved sensing performance and propose a novel calibration method for stereo configuration. For the stereo calibration, the exact positions of sufficient corresponding points in two sensor coordinates need to be determined. A printed checkerboard is widely used for stereo camera systems because it is easy to construct and its accuracy is guaranteed owing to its printing accuracy. However, in the case of the lighthouse system, it is very difficult or impossible to construct a highly accurate calibration board similar to the checkerboard mainly because of manufacturing errors. In this study, we use a receiver sensor and a two-axis linear stage equipped with micrometers. By moving predetermined distances along the x and y directions on the linear stage, we can obtain multiple-point information with high accuracy, which can then be used for the stereo calibration of two lighthouses. In this paper, the calibration and pose estimation procedures are described in detail, and the pose estimation result of the perspective-n-points method is compared with that of the triangulation method. Finally, the pose estimation accuracy of the proposed system is compared with that of a commercial system that is widely used for highly accurate medical applications.**



***Index Terms*—Base station, calibration, lighthouse, pose estimation, virtual reality.**

I. Introduction

IN ORDER to give a user of virtual reality (VR) system highly immersion, to freely interact between a user and the virtual environment, accurate and low latency tracking is one of the most important requirements for the VR system [1]. This accurate and fast tracking is expected to be used in variety of applications not only for entertainment [2] but also for medicine [3], engineering and design [4], military [5], education [6], virtual prototyping [7], and architecture and cultural heritage [8]. To meet the high requirements of tracking, lighthouse localization system of HTC Vive use inertial measurements and light data from the inertial measurement unit (IMU) and the photodiodes in the tracked object like HMD, controller, and tracker [9]. Before emerging of Vive tracking system, there already exist similar trackers that only use light data like Minnesota scanner [10] and indoor global positioning system (iGPS) [11]. Actually, the architecture of iGPS is similar to the second generation lighthouse localization system [11], [12]. The lighthouse localization system has been developed as a suitable device for VR incorporating IMU into existing rotating laser localization technology. Considering only light data, the lighthouse localization system can be largely divided into two parts: lighthouses and a tracked object. A lighthouse periodically emits two types of light source: a synchronization flash, an infrared (IR) sweeping laser plane. And a tracked object calculates its pose using the time interval of the lights from the lighthouse [13].

Several researches that assessed 3-D pose estimation of the lighthouse localization system have been

Manuscript received August 17, 2021; accepted August 29, 2021. Date of publication September 14, 2021; date of current version October 29, 2021. This work was supported in part by the Ministry of Trade, Industry and Energy (MOTIE), Republic of Korea, under Grant 20001856. The associate editor coordinating the review of this article and approving it for publication was Dr. Marco J. Da Silva. *(Corresponding author: Sungon Lee.)*

Sungtaek Cho was with the School of Electrical Engineering, Hanyang University, Ansan 15588, Republic of Korea. He is now with NETXQURE Inc., Seongnam 13486, Republic of Korea (e-mail: savagegard_n@nate.com).

Dongyeon Kim was with the School of Electrical Engineering, Hanyang University, Ansan 15588, Republic of Korea. He is now with SK Hynix, Cheongju-si 28429, Republic of Korea (e-mail: gomdongyeon@gmail.com).

Sungon Lee is with the School of Electrical Engineering, Hanyang University, Ansan 15588, Republic of Korea (e-mail: sungon@hanyang.ac.kr).

Digital Object Identifier 10.1109/JSEN.2021.3112584

conducted [9], [13]–[25]. Niehorster *et al.* [14] firstly conducted experiments with the lighthouse system and announced the results about its accuracy, jitter, and latency. They reported that a sufficient care must be taken to use the system for general research purpose because the lighthouse system has a systemic error of a slanted ground problem. Sitole *et al.* [25] solved the systemic error problem raised by Niehorster *et al.* using an additional tracked object as a global tracking frame. Groves *et al.* [20] conducted experiments using the second generation lighthouse system, commercial optical tracking system, and a hybrid system that combined former two systems for a medical application. They concluded that the lighthouse system could be used for the medical field and added that updated software and upgraded hardware may make a different result from Niehorster's. Borrego *et al.* [18] compared the lighthouse localization system with Oculus Rift (Oculus VR, Irvine, CA) about their working area, accuracy, and jitter.

On the other hand, Azad *et al.* [26] compared stereo and mono configuration of cameras theoretically and experimentally to estimate six degree of freedom (DOF) pose of objects and concluded that pose estimation using stereo method were more robust and accurate. The accuracy of the camera calibration, which obtains intrinsic parameters related to camera itself and extrinsic parameters about geometric relation between two cameras is an important factor that determines the performance of the stereo vision [27]. In 3-D pose estimation with lighthouse system using SteamVR plugin or openVR software development kit (SDK) that is offered by manufacturer's side, the lighthouses are being used in a typical configuration where two base stations are being diagonally mounted above head height at maximum 5 m distance with a tilting angle of 30 to 45 degrees [14], [18]–[25]. It seems that the lighthouses do not use stereo matching method for 3D estimation but use 3D-2D correspondence [15], [19] but multiple base stations were used for the expansion of detecting range. In this configuration of lighthouses, when line of sight from a lighthouse is blocked, the whole lighthouse system still estimates the pose of the object using another lighthouse. This setup and configuration is appropriate for VR because the size of the detecting range is more important than the tracking accuracy. In tracking applications such as medical purpose, the stereo method would be more appropriate than the typical mono method because the accuracy is much more important than the size of the detecting area. In this study, we investigate lighthouses with a stereo method. To the best of our knowledge, there has been no study comparing 3-D pose estimation results of the stereo and mono configuration of the lighthouse.

The aim of this study is two-fold. First, we proposed a novel stereo lighthouse calibration method by adapting an algorithm proposed by Zhang [28] which is widely used for camera calibration. For the calibration, the accurate positions of corresponding points in two lighthouse coordinate frames need to be determined. We use a photodiode and a two-axis linear stage. By moving intended distances along the two different directions on the linear stage, we can get multiple-point information with high accuracy, which can then be used for the stereo lighthouse calibration.

Second, we comparatively evaluated the accuracy and jitter using the single and stereo configuration of lighthouse localization system in static state. For a position estimation, PnP method is used for a single lighthouse and triangulation method is used for a stereo lighthouse. The two methods are described in detail for potential users of this technology. The result of the stereo lighthouse calibration which represents geometric relationship between two lighthouses is neceassary for the triangulation method. For medical or research purpose, we set the tracking area rather smaller like recumbent human body size [29] than the original lighthouse tracking area [14]. Furthermore, we compared the position estimation according to distance and vergence between two lighthouses. Finally, we compared our results with those of other research groups.

This paper is organized as follows. In Section II, the hardware and software used in the experiments are described. In Section III, the pose estimation algorithm using the lighthouse system is explained in detail. We explain the basic operation of the lighthouse, as well as its calibration and pose estimation, using both PnP and triangulation methods. In Section IV, the experimental environment as well as the calibration and position estimation results are described. In Section V, the conclusion of this research is presented.

II. Materials

A. Hardware

Our lighthouse localization measurement system comprises two Vive lighthouses (HTC Corporation, Taoyuan City, Taiwan), a tracked object, and an operating system (Fig. 1). The tracked object was attached to the end-effector of a six-degree-of-freedom device comprising three-axis rotation and translation stages. The tracked object consists of four photodiodes and a micro-controller unit (MCU). The photodiode is TS3633-CM1 (Triad Semiconductor, Inc., Winston-Salem, North Carolina, US), which converts infrared light pulses to electrical pulses. The MCU is Teensy 3.2 (PJRC.COM, LLC., Sherwood, Oregon, US) whose clock frequency is 48 MHz. The three-axis rotation stage is TTR001 (Thorlab, Inc., Newton, New Jersey, US), and it provides $\pm 5^\circ$ adjustment and 0.036° resolution in the pitch and roll directions, and $\pm 10^\circ$ adjustment and 0.03° resolution in the yaw direction. The three-axis automatic translation stage comprises three T-LSM Series (Zaber Technologies Inc., Vancouver, British Columbia, Canada). The moving range of the linear stage in the x and z directions is 100 mm and that in the y direction is 50 mm. The accuracy of the linear stage is 8 μm and its resolution is 50 nm according to the online manual of the manufacturer. For the 3-D reference sensor, Polaris Spectra infrared camera (Northern Digital Inc., Waterloo, Ontario, Canada) was used. Based on its specification, the sensor can measure the 3-D position of infrared reflective spheres with a less than 0.30-mm root-mean-square (RMS) error.

B. Software

MATLAB (Mathworks Inc.) was primarily used for the experiments. Open-source computer vision library (OpenCV) library was used only for the lighthouse calibration [30], and Arduino 1.8.12/Teensyduino 1.51 was used to operate

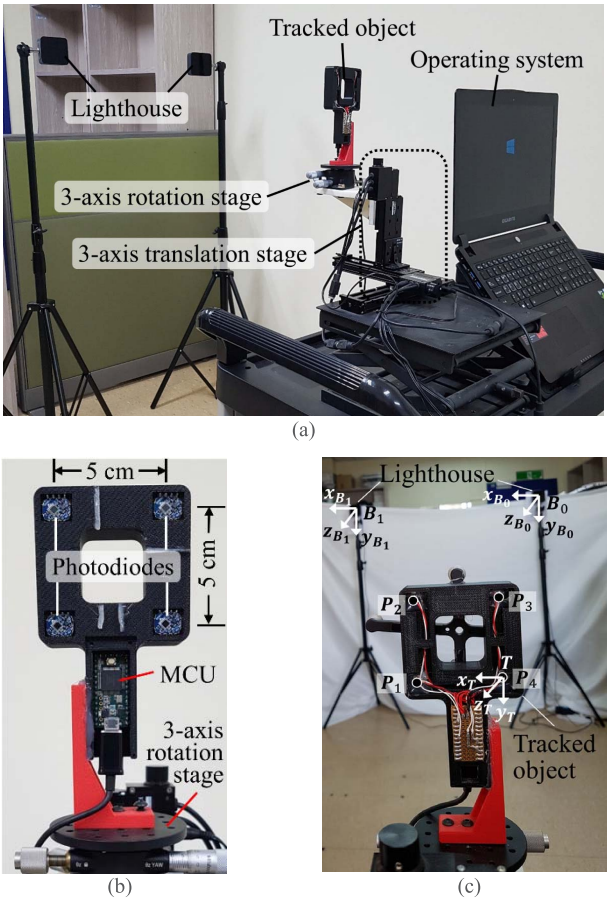


Fig. 1. Components of the localization system and the coordinate frames. (a) 3-D measurement system comprising the lighthouses, tracked object, operating system, and rotation and translation stages. A tracked object is attached on the end-effector of the rotation and translation stages. (b) Components of our tracked object and its geometric information. (c) Coordinate frame of the two lighthouses and the tracked object.

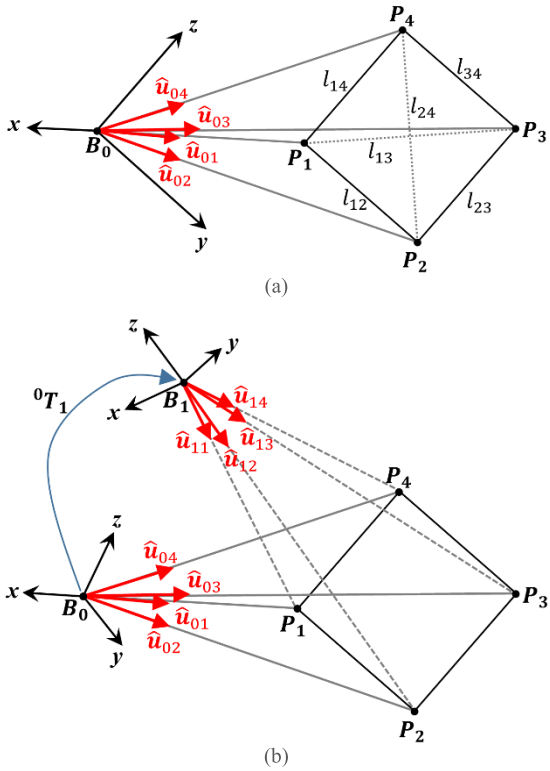


Fig. 2. Two schematic diagrams for pose estimation. (a) PnP method using a single lighthouse and the geometric information between the points. (b) Triangulation method using a stereo lighthouse and the transform between two lighthouses.

the MCU of the tracked object. Furthermore, Zaber Console software was used for the automatic xyz stage.

III. Methods

In this section, we introduce the process for obtaining the position and orientation of the tracked object in detail. PnP method with a single lighthouse and triangulation method with a stereo lighthouse were used to estimate the pose of the tracked object as shown in Fig. 2. The basic operation of the lighthouse, as well as its calibration and pose estimation, is described in the following subsections. Calibration is necessary only for triangulation-based pose estimation.

A. Basic Operation of Lighthouse

In this subsection, we explain the basic operation of a lighthouse system and the mathematical procedure for obtaining a unit vector from the raw data of the lighthouse system. A lighthouse comprises two main parts: an IR array and two rotors [12]. The IR array periodically emits light to synchronize the starting time; one rotor emits vertical laser sweep and the other rotor emits horizontal laser sweep. The four photodiodes of the tracked object detect the synchronized starting time (sync) and beam light, and the MCU saves the interval time between the sync and the beam light of each photodiode [13]. The lighthouse blinks a sync flash and the MCU starts clock counting as soon as the sync light reaches the photodiode. After a short time that varies according to the angular distance between the lighthouse and each photodiode, the laser plane sweeps the photodiode. After this second signal from the photodiode, the MCU ends clock counting (tick_v) and saves it. The counting of horizontal laser plane tick_h is obtained in a similar manner.

The procedure for obtaining the unit vector is as follows [31]. The rotors rotate at 60 revolutions per second [12]. In other words, the rotors complete one revolution in 16,666 μs and half revolution in 8,333 μs. The clock frequency of the MCU is 48 MHz. tick_v and tick_h are converted to vertical and horizontal angle (θ_v, θ_h) between the lighthouse and the photodiode respectively using:

$$\theta_\text{v} = \left(\frac{\text{tick}_\text{v}}{48} - 4000\right) \times \frac{180}{8333} \tag{1}$$

$$\theta_\text{h} = \left(\frac{\text{tick}_\text{h}}{48} - 4000\right) \times \frac{180}{8333} \tag{2}$$

where, 48 relates the clock frequency of the MCU, 4000 is central time offset in μs, 8333 is half revolution time of the rotor in μs, and 180 is used to express θ_v and θ_h in degree.

The next step is to obtain normal vectors of planes from θ_v and θ_h. For example, a lighthouse B_0, its coordinate frame and the photodiode of a tracked object P_1 are shown in Fig. 3a, in which two laser planes reach P_1. In this situation, normal

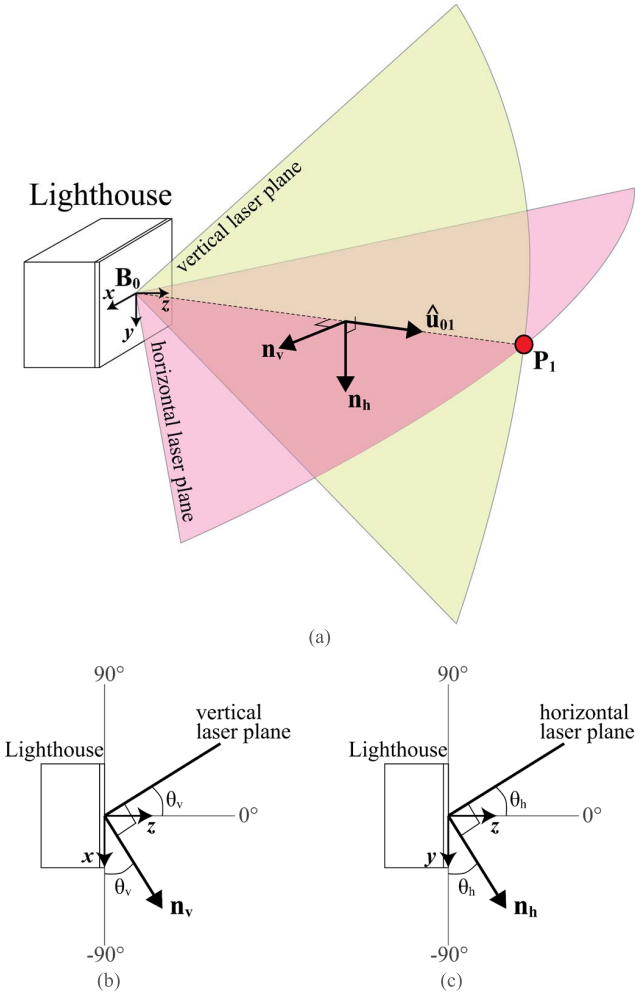


Fig. 3. (a) Unit vector $\hat{u}_{01}$ from origin B_0 of the master lighthouse to photodiode P_1 is described in 3-D view. (b) Normal vector n_v about the vertical plane is described in 2-D view. (c) Normal vector n_h about the horizontal plane is described in 2-D view. The vertical laser plane rotates about the –y axis and the horizontal plane rotates about the x axis.

vectors about vertical and horizontal planes are $\mathbf{n}_v$ and $\mathbf{n}_h$, respectively. From Fig. 3b, 3c, (3) and (4) can be obtained. The unit vector $\hat{u}$ from the lighthouse to the photodiode is obtained using the cross-product of the two normal vectors as given by (5).

$$\mathbf{n}_v = [\cos\theta_v, 0, \sin\theta_v]^T \tag{3}$$

$$\mathbf{n}_h = [0, \cos\theta_h, \sin\theta_h]^T \tag{4}$$

$$\hat{\mathbf{u}} = (\mathbf{n}_v \times \mathbf{n}_h)/\mathbf{n}_v \times \mathbf{n}_h \tag{5}$$

B. Calibration

The purpose of calibration is to identify the geometric relationship between two lighthouses as described in Fig. 2b. The relationship ${}^0\mathbf{T_1}$ comprises a 3×3 rotation matrix $\mathbf{R}$ and a 3×1 translation vector $\mathbf{T}$, and is given by:

$${}^0\mathbf{T_1} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix} \tag{6}$$

The concept of stereo camera calibration was adopted in our lighthouse calibration. Using several pictures of a checkerboard (see Fig. 4a) in various positions and orientations,

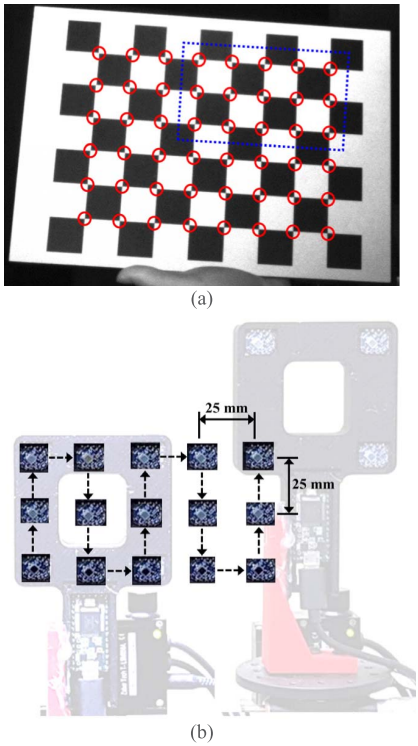


Fig. 4. (a) Checkerboard used for stereo camera calibration. (b) Proposed calibration procedure of the checkerboard. The locus of the photodiode is similar to the 15 points inside the blue-dotted rectangle of the checkerboard. The locus is precisely determined by the automatic xyz stage and one photodiode with an intended distance.

intrinsic and extrinsic parameters of the stereo camera were calculated. The camera calibration accuracy is guaranteed owing to its printing accuracy. For the lighthouse calibration, the xyz stage and one photodiode was substituted for the checkerboard as shown in Fig. 1. Using this apparatus, the unit vectors about 15 points from the two lighthouses were measured (Fig. 4b). Owing to the accurate positioning of the xyz stage, we can obtain the exact data for calibration.

Furthermore, we used an OpenCV function whose core algorithm was based on Zhang's method [28]. Since Zhang's method was initially developed for stereo camera calibration, we performed three steps for the lighthouse calibration. First, we determined the number of point sets. The number of cross points of the checkerboard is 48 as shown in Fig. 4a. We took ten different angle and distance photos of the board to obtain a stable calibration result. Consequently, we used 480 points for the camera calibration. For the lighthouse, we obtained data 36 times in the target area as shown in Fig. 7a. We obtained 540 points for the stereo lighthouse calibration. For reference, 1280-point data were used by Zhang [32] to test his algorithm.

Second, we found the contact points between the unit vectors derived in the previous subsection and the virtual planes as shown in Fig. 5. Assume that an arbitrary plane exists. This plane is apart from the origin with a length f along the z-axis and normal to the z-axis using the relation:

$$Ax + By + Cz + D = 0 \tag{7}$$

where A and B are zero, because the plane is normal to z-axis, C and D are scalar variables determined by f in Fig. 5. Assume

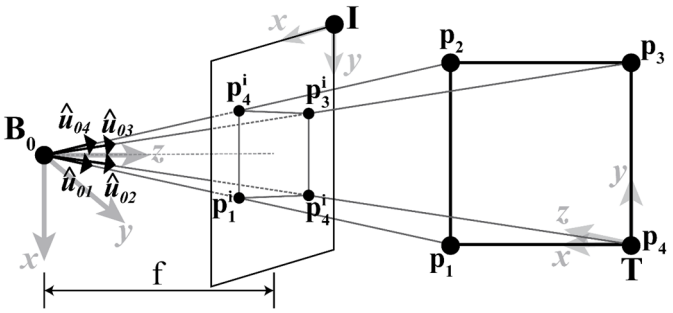


Fig. 5. Intersections (p^i_1, p^i_2, p^i_3, p^i_4) between a virtual 2-D plane I and unit vectors. The virtual plane is perpendicular to the z-axis of the coordinate frame B_0 and placed at a distance f from B_0. The virtual plane acts as an image plane in the camera and the distance f acts as the focal length in the camera. The intersections are used as input of the lighthouse calibration process.

an arbitrary line $\mathbf{t}$ that passes two points $\mathbf{t}_1$ and $\mathbf{t}_2$:

$$\mathbf{t} = \mathbf{t}_1 + \mathrm{u}(\mathbf{t}_2 - \mathbf{t}_1) \tag{8}$$

In this case, the unit vectors from the lighthouses to photodiodes are obtained in the previous subsection. Therefore, $\mathbf{t}_1$ and $\mathbf{t}_2$ are known variables but u is an unknown scalar variable. For better understanding, (8) is rewritten as:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \mathrm{u} \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{bmatrix} = \begin{bmatrix} \mathrm{u}(x_2 - x_1) + x_1 \\ \mathrm{u}(y_2 - y_1) + y_1 \\ \mathrm{u}(z_2 - z_1) + z_1 \end{bmatrix} \tag{9}$$

u can be obtained from (7) and (9). The contact points are then obtained from (9) by substituting the unknown u with the calculated value. Consequently, the contact point between arbitrary planes and the unit vectors is given as:

$$\mathbf{p}^{\mathbf{i}}_{\mathbf{j}} = \tilde{t} + \frac{\mathrm{f}}{2} = \left[x + \frac{\mathrm{f}}{2}, y + \frac{\mathrm{f}}{2} \right]^T \tag{10}$$

where $\tilde{t}$ is a 2×1 vector from (9) without the z-axis value, $\mathbf{p}^{\mathbf{i}}_{\mathbf{j}}$ is a coordinate on an arbitrary plane whose origin is on the upper left corner, $\mathbf{j}$ of $\mathbf{p}^{\mathbf{i}}_{\mathbf{j}}$ indicates a photodiode number and $\mathbf{i}$ is used to distinguish from 3-D coordinate $\mathbf{p_j}$ as shown in Fig. 3.

Finally, we tuned the function for calibration to estimate only the extrinsic parameters for the lighthouses. If intrinsic parameters are not estimated with high accuracy, the function may diverge [30]. Thus, we can obtain a better calibration result when we estimate only the rotation matrix ($\mathbf{R}$) and the translation vector ($\mathbf{T}$) between the two lighthouses. Hence, we obtained ${}^{\mathbf{0}}\mathbf{T_1}$.

C. Pose Estimation

The pose estimation is divided into two steps. In the first step, the positions of four photodiodes ($\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$, $\mathbf{P}_4$) about the world coordinate frame $\mathbf{B}_0$ (Fig. 2) were determined using PnP and triangulation methods. In the second step, the orientation of the tracked object was calculated, given the positions of the photodiodes. Unit vectors $\hat{u}_{01}$, $\hat{u}_{02}$, $\hat{u}_{03}$, and $\hat{u}_{04}$ are from $\mathbf{B}_0$; $\hat{u}_{11}$, $\hat{u}_{12}$, $\hat{u}_{13}$, and $\hat{u}_{14}$ are from $\mathbf{B}_1$.

A major difference between the PnP and triangulation methods is additional information, except the number of required lighthouses. As shown in Fig. 2a, the geometric information of the tracked object ($l_{12}, l_{13}, \ldots, l_{34}$) is necessary for the PnP method, whereas a relationship (${}^{\mathbf{0}}\mathbf{T_1}$) between two lighthouses is required for the triangulation method (Fig. 2b).

1) PnP Method With a Single Lighthouse: According to Leptetit [33], the purpose of PnP method is to find the position and orientation of a camera with its intrinsic parameters and a set of n correspondences between 3-D points and their 2-D projections. PnP problems are solved using iterative and non-iterative methods.

We calculated the positions of the photodiodes using a geometry-based PnP method that is similar to that of Islam *et al.* [13], which is an iterative method. From Fig. 2a, the variables are defined as follows: the distances between the lighthouse and the photodiodes are unknown variables: $\mathrm{L}_1 = \overline{\mathbf{B_0P_1}}$, $\mathrm{L}_2 = \overline{\mathbf{B_0P_2}}$, $\mathrm{L}_3 = \overline{\mathbf{B_0P_3}}$, $\mathrm{L}_4 = \overline{\mathbf{B_0P_4}}$. Furthermore, the distances between the photodiodes are known variables: $\mathrm{l}_{12} = \overline{\mathbf{P_1P_2}}$, $\mathrm{l}_{13} = \overline{\mathbf{P_1P_3}}$, $\mathrm{l}_{14} = \overline{\mathbf{P_1P_4}}$, $\mathrm{l}_{23} = \overline{\mathbf{P_2P_3}}$, $\mathrm{l}_{24} = \overline{\mathbf{P_2P_4}}$, $\mathrm{l}_{34} = \overline{\mathbf{P_3P_4}}$. For the angles, $\theta_{12} = \angle\mathbf{P_1B_0P_2}$, $\theta_{13} = \angle\mathbf{P_1B_0P_3}$, $\theta_{14} = \angle\mathbf{P_1B_0P_4}$, $\theta_{23} = \angle\mathbf{P_2B_0P_3}$, $\theta_{24} = \angle\mathbf{P_2B_0P_4}$, $\theta_{34} = \angle\mathbf{P_3B_0P_4}$. From the six triangles $\mathbf{B_0P_1P_2}$, $\mathbf{B_0P_1P_3}$, $\mathbf{B_0P_1P_4}$, $\mathbf{B_0P_2P_3}$, $\mathbf{B_0P_2P_4}$, and $\mathbf{B_0P_3P_4}$, the following six equations can be derived using cosine law:

$$\begin{aligned}
&\mathrm{L}_1^2 + \mathrm{L}_2^2 - 2\mathrm{L}_1\mathrm{L}_2\cos\theta_{12} - \mathrm{l}_{12}^2 = 0,\\
&\mathrm{L}_1^2 + \mathrm{L}_3^2 - 2\mathrm{L}_1\mathrm{L}_3\cos\theta_{13} - \mathrm{l}_{13}^2 = 0,\\
&\mathrm{L}_1^2 + \mathrm{L}_4^2 - 2\mathrm{L}_1\mathrm{L}_4\cos\theta_{14} - \mathrm{l}_{14}^2 = 0,\\
&\mathrm{L}_2^2 + \mathrm{L}_3^2 - 2\mathrm{L}_2\mathrm{L}_3\cos\theta_{23} - \mathrm{l}_{23}^2 = 0,\\
&\mathrm{L}_2^2 + \mathrm{L}_4^2 - 2\mathrm{L}_2\mathrm{L}_4\cos\theta_{24} - \mathrm{l}_{24}^2 = 0,\\
&\mathrm{L}_3^2 + \mathrm{L}_4^2 - 2\mathrm{L}_3\mathrm{L}_4\cos\theta_{34} - \mathrm{l}_{34}^2 = 0,
\end{aligned} \tag{11}$$

where $\cos\theta_{12,13,\ldots,34}$ are calculated by the dot product of the unit vectors. The dot product is given by (12). Equation (13) can be simplified from (12) because $\hat{\mathbf{u}}_{01}$ and $\hat{\mathbf{u}}_{02}$ are unit vectors.

$$\hat{\mathbf{u}}_{01} \cdot \hat{\mathbf{u}}_{02} = |\hat{\mathbf{u}}_{01}||\hat{\mathbf{u}}_{02}| cos\theta_{12} \tag{12}$$

$$\cos\theta_{12} = \hat{\mathbf{u}}_{01} \cdot \hat{\mathbf{u}}_{02} \tag{13}$$

Equation (11) is a non-linear problem with four unknowns and six equations solved using the Levenberg–Marquardt (LM) algorithm [34], [35]. After determining L_1, $\mathbf{P}_1$ whose origin is $\mathbf{B}_0$ was obtained using (14), and $\mathbf{P}_2$, $\mathbf{P}_3$, and $\mathbf{P}_4$ were calculated in a similar manner.

$$\mathbf{P}_1 = \mathrm{L}_1\hat{\mathbf{u}}_{01} = \left[\mathrm{L}_1\hat{\mathrm{u}}_{01\mathrm{x}}, \mathrm{L}_1\hat{\mathrm{u}}_{01\mathrm{y}}, \mathrm{L}_1\hat{\mathrm{u}}_{01\mathrm{z}}\right]^T \tag{14}$$

2) Triangulation Method With a Stereo Lighthouse: Triangulation method finds the closest points of two unit vectors from two lighthouses about the same photodiode [36]. In Fig. 2b for example, $\hat{\mathbf{u}}_{01}$ from $\mathbf{B}_0$ and $\hat{\mathbf{u}}_{11}$ from $\mathbf{B}_1$ are the two unit vectors, and $\mathbf{P}_1$ is a photodiode. For a stereo lighthouse, $\mathbf{B}_0$ denotes the master lighthouse and world coordinate frame and $\mathbf{B}_1$ is a slave lighthouse. $\mathbf{R}$ and $\mathbf{T}$, which are the results of the calibration process are used here. Let us estimate $\mathbf{P}_1$ using triangulation method. $\mathbf{B}_1$ is set apart from $\mathbf{B}_0$ with $\mathbf{T}$ like

$$\mathbf{B_0} = [0, 0, 0]^{\mathrm{T}}, \quad \mathbf{B_1} = \mathbf{T} \tag{15}$$

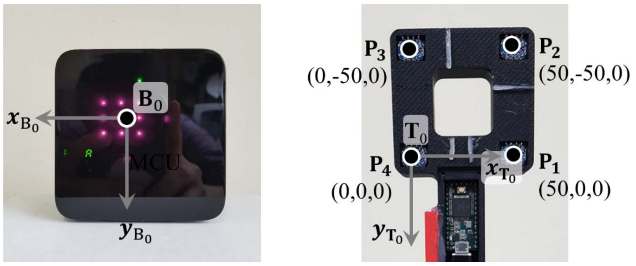


Fig. 6. A coordinate frame of master lighthouse ($\mathbf{B_0}$) is described in (a). The origin of the coordinate frame is in the intersection of the rotating axes of the two rotors. A coordinate frame of the tracked object ($\mathbf{T_0}$) and geometric information of the photodiodes are described in (b).

$\mathbf{R}$ converts $\hat{\mathbf{u}}_{11}$ (a unit vector about $\mathbf{B}_1$) to $\hat{\mathbf{u}}_{11\mathrm{r}}$ (a unit vector about $\mathbf{B}_0$) using:

$$\hat{\mathbf{u}}_{11\mathrm{r}} = \mathbf{R}\hat{\mathbf{u}}_{11} \tag{16}$$

$\mathbf{P}_{01}$ represents position of $\mathbf{P}_1$ about $\mathbf{B}_0$ and $\mathbf{P}_{11}$ represents the position of $\mathbf{P}_1$ about $\mathbf{B}_1$ as follows:

$$\mathbf{w}_0 = \mathbf{B_0} - \mathbf{B_1} \tag{17}$$

$$\begin{aligned} \mathrm{a} &= \hat{\mathbf{u}}_{01}\cdot\hat{\mathbf{u}}_{01}, \\ \mathrm{b} &= \hat{\mathbf{u}}_{01}\cdot\hat{\mathbf{u}}_{11\mathrm{r}}, \\ \mathrm{c} &= \hat{\mathbf{u}}_{11\mathrm{r}}\cdot\hat{\mathbf{u}}_{11\mathrm{r}}, \\ \mathrm{d} &= \hat{\mathbf{u}}_{01}\cdot\mathbf{w}_0, \\ \mathrm{e} &= \hat{\mathbf{u}}_{11\mathrm{r}}\cdot\mathbf{w}_0 \end{aligned} \tag{18}$$

$$\mathbf{P}_{01} = \frac{\mathrm{be}-\mathrm{cd}}{\mathrm{ac}-\mathrm{bb}}\hat{\mathbf{u}}_{01} + \mathbf{B_0}, \qquad \mathbf{P}_{11} = \frac{\mathrm{ac}-\mathrm{bd}}{\mathrm{ac}-\mathrm{bb}}\hat{\mathbf{u}}_{11\mathrm{r}} + \mathbf{B_1} \tag{19}$$

where $\mathbf{w}_0$ and a–e are used to simplify (19). The estimated position of $\mathbf{P}_1$ is obtained as the midpoint of $\mathbf{P}_{01}$ and $\mathbf{P}_{11}$ as given by (20). $\mathbf{P}_2$, $\mathbf{P}_3$, and $\mathbf{P}_4$ can be obtained in a similar manner.

$$\mathbf{P}_1 = (\mathbf{P}_{01} + \mathbf{P}_{11})/2 \tag{20}$$

3) Orientation Estimation: After obtaining the position of $\mathbf{P}_1 \sim \mathbf{P}_4$ about coordinate frame $\mathbf{B}_0$, we can calculate the orientation of the tracked object using Horn's method [37]. This method is used to find the coordinate transform between two coordinate frames using corresponding coordinates of same points in the two different frames. In our case, the coordinates of four photodiodes about coordinate frame $\mathbf{T}_0$ were fixed as described in Fig. 6b. The other coordinates about coordinate frame $\mathbf{B}_0$ were calculated in the previous subsection. The result of Horn's method is a position and orientation between $\mathbf{B}_0$ and $\mathbf{T}_0$. The orientation is represented by a unit quaternion.

IV. Result

A. Experimental Setup

The experimental setup (Fig. 7) shows the hardware positioning. Two red circles represent two lighthouses in the first experiment. The lines of sight of the two lighthouses are parallel. Two blue circles represent two lighthouses in the second experiment. For the second experiment, we enlarged

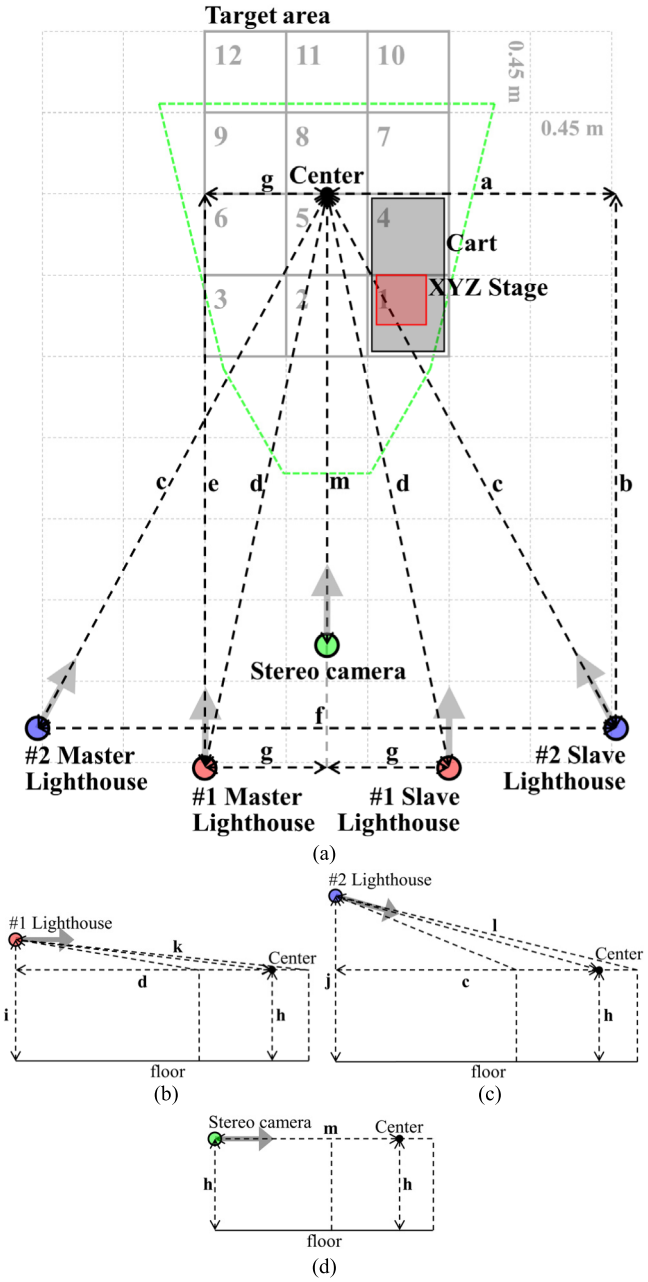


Fig. 7. (a) shows experimental setup in aerial view. Red circles represent the position of the lighthouse at first experiment, blue circles represent the lighthouse at second experiment, and a green circle represents stereo camera. Five gray arrows attaching to circles represent the direction of each tracking system. A cart represented by gray rectangle containing an xyz stage represented by red rectangle with the tracked object moves 12 sections in the target area for the calibration and pose estimation. A green dotted polygon represents detecting area of the stereo camera. (b) shows experimental setup of the first experiment using lighthouse in lateral view. (c) shows experimental setup of the second experiment using lighthouse in lateral view. (d) shows experimental setup using stereo camera in lateral view. All symbols (a–m) representing distance between objects are given in TABLE I.

the distance and the vergence angle between the two lighthouses. The green circle indicates a stereo camera. The master lighthouse of each case was used for pose estimation using PnP method. The distances in Fig. 7 are given in Table I. Some distances (c, d, k, l, m) have minimum and maximum values according to the location of a cart. The cart carries the xyz stage and the tracked object moves from 1 to 12 in the target area of Fig. 7a.

TABLE I
SETTING POSITION OF THE EXPERIMENTAL SETUP

Distance	Setting in m	Distance	Setting in m
a	1.60	i	1.62
b	3.11	j	2.20
c	3.50	k	3.43
c_{min}	2.40	k_{min}	2.46
c_{max}	4.01	k_{max}	3.91
d	3.40	l	3.64
d_{min}	2.43	l_{min}	2.60
d_{max}	3.89	l_{max}	4.13
e	3.33	m	2.45
f	3.20	m_{min}	1.55
g	0.68	m_{max}	2.90
h	1.21		

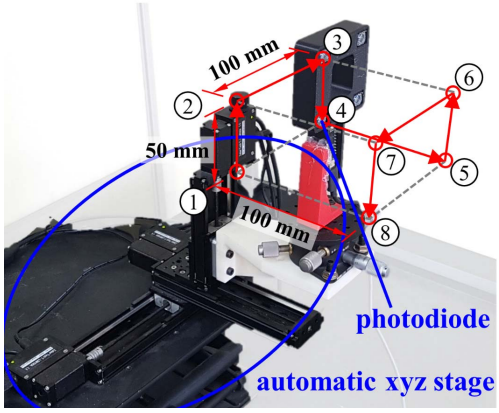


Fig. 8. We use an automatic xyz stage and a photodiode of the tracked object to assess the accuracy and jitter of our system. Using the automatic xyz stage, the photodiode moves precisely and sequently from ① to ⑧. The positions of a photodiode are saved 250 times and then the data is averaged excepting abnormal value.

We set the target area because our goal was to evaluate the accuracy of the stereo lighthouse by comparing it with a commercial stereo camera. We also set the target area to fit the measurement area of the camera system. The tracking area of the stereo camera represented with the green dotted line in Fig. 7a is smaller than the tracking area of the original lighthouse setup. We performed an experiment in the maximum overlapped area between our lighthouse system and the stereo camera.

We now describe how to obtain the position of one photodiode. To avoid confusion, we refer to the position of the cart as target and the position of the tracked object on the xyz stage as position. Thus, there were 12 targets (Fig. 7a) and eight positions (Fig. 8) per target. We used a cart to move from one target to another and the xyz stage to move from one position to another. To obtain the position of one photodiode, we averaged 250 samples at one position. Likewise, we obtained 96 position data in the 12 targets.

B. Result

1) Calibration Result: Two calibration results about two sets of the lighthouse positioning were obtained. Equation (21) represents the relationship of the first experiment represented by red circles in Fig. 7a. The coordinate frames of the lighthouses are shown in Fig. 1c. The orientation part of (21) is compatible with the setup of the two lighthouses for the first experiment as observed in Fig. 1c. The translation part of (21) also accurately represents the distance between the two lighthouses. In the second experiment, the faces of the two lighthouses were tilted to focus on the center of the target area (gray arrow in Fig. 7). Although the calibration result (22) cannot be confirmed intuitively, the norm (3.11 m) of the translation part of (22) is similar to the distance **f** (3.20 m) between the blue circles in Fig. 7a.

$$\begin{bmatrix} 0.9994 & -0.0053 & -0.0351 & 1307.89 \\ 0.0042 & 0.9996 & -0.0295 & 17.59 \\ 0.0352 & 0.0293 & 0.9989 & 135.15 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{21}$$

$$\begin{bmatrix} 0.9343 & 0.3251 & -0.1459 & 382.96 \\ -0.2858 & 0.4392 & -0.8517 & 2680.66 \\ -0.2128 & 0.8375 & 0.5033 & 1537.67 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{22}$$

2) Position Estimation Result: We first define the term accuracy and jitter used in this paper. We obtained the data of eight positions moved by the automatic xyz stage per target (see Fig. 8). With these eight points, we made 28 lines such as $\overline{①②}$, $\overline{①③}$, …, $\overline{⑦⑧}$. We then averaged the difference between the desired distance and the measured distance to obtain accuracy using:

$$E = \frac{1}{N}\sum_{i=1}^{N} |D_i - M_i| \tag{23}$$

where N is 28, the number of lines, D_i indicates the desired distance, M_i represents the measured distance, and $|\cdot|$ is an operator to calculate absolute values. Jitter is a means of measuring the precision of a localization system [20], [38]. We calculated jitter using:

$$J = \frac{1}{L}\sum_{j=1}^{L}\sqrt{\frac{1}{M}\sum_{i=1}^{M}\left\|\mathbf{P}_{ij} - \bar{P}_j\right\|} \tag{24}$$

where L is 8, number of position per target, M is 250, number of sample per position, $\mathbf{P}_{ij}$ is the measured position of ith sample of jth position, $\bar{P}_j$ represents the mean position of 250 samples of jth position, and $\|\cdot\|$ indicates Euclidean distance between two points in 3-D space.

Fig. 9 - Fig. 11 and Table II are results of the position estimation about one photodiode of the tracked object in our

TABLE II
AVERAGE ACCURACIES AND JITTERS ABOUT FIVE EXPERIMENTS

Description	Accuracy [mm]	Mean of Jitter [mm]
#1 Stereo L.[a]	2.781	0.244
#1 Single L.[a]	2.199	3.238
#2 Stereo L.[a]	**1.131**	**0.141**
#2 Single L.[a]	2.960	3.672
Stereo Cam.[b]	**0.488**	**0.036**

[a]Lighthouse, [b]Camera

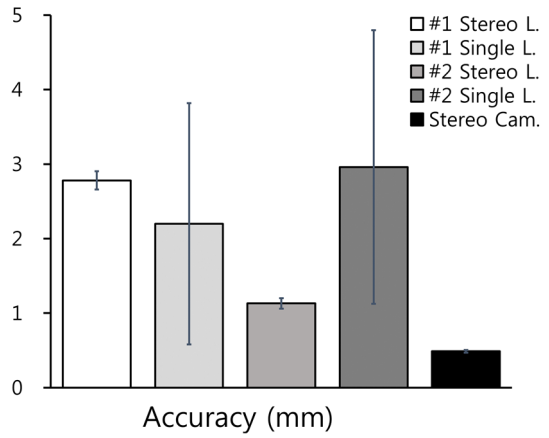


Fig. 9. Average accuracies with jitters about five experiments. We used 96 points in 12 targets for the calculation of the pose estimation at the each experiment. #1 means first experiment, L. means lighthouse, and Cam. means camera. For single lighthouse experiments, we used master lighthouse in each experiment (refer to Fig. 7a). The values of the accuracies and jitters are given in TABLE II.

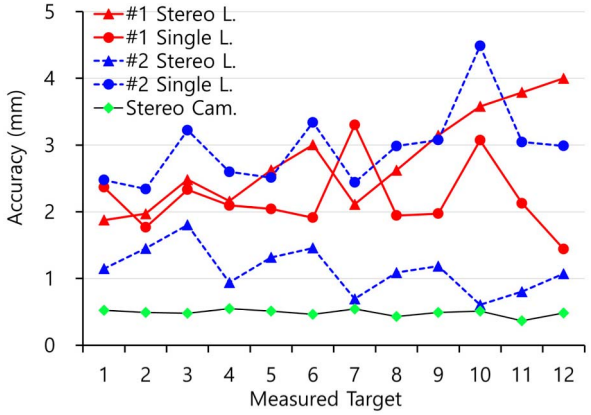


Fig. 10. Accuracies of five experiments in twelve targets. In each target, 8 positions of the photodiode were measured and then 28 lines made from the 8 points were compared to the desired length. The average of the difference between the measured and the desired length is defined as accuracy.

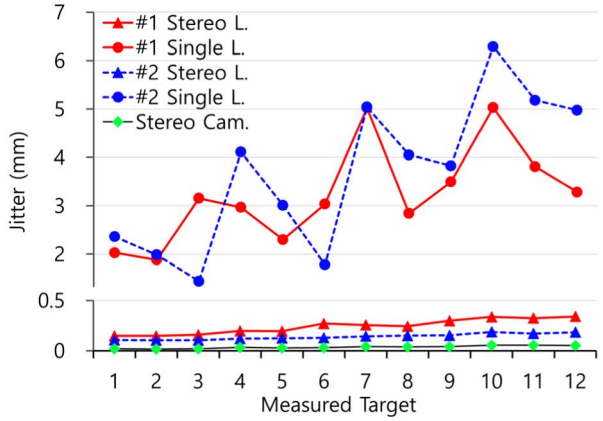


Fig. 11. Jitters of five experiments in twelve targets. 8 positions of the photodiode were measured in each target and 250 times were measured in each position. Jitter is defined as the average of the difference between the measured and averaged position about the 2000 points.

target area (Fig. 7). All the figures and the table show the results of five different experiments for accuracy assessment. Fig. 9 and Table II show the accuracy and the jitter, in which a well-tuned stereo lighthouse (#2 Stereo L.) has a better result than a single lighthouse. The accuracy and jitter of the second experiment for 12 targets using a stereo lighthouse are 1.131 mm and 0.141 mm, respectively, whereas the corresponding best results for a single lighthouse are 2.199 mm and 3.238 mm, respectively, as shown in Table II. The accuracy result about 12 targets is shown in Fig. 10, and that of jitter is shown in Fig. 11.

3) Result Analysis: From the results, we can guess some characteristics of the lighthouse localization system. First, the distance and vergence angle between the lighthouses affect the accuracy of pose estimation. In Fig. 10, the single lighthouse shows a better result than the stereo lighthouse in the first experiment (red solid lines in Fig. 10). However, in the second experiment (blue dotted lines in Fig. 10), the stereo lighthouse has a better result than the single lighthouse. The difference between the first and second experiments is the distance and vergence angle between the lighthouses. Kyt̀o *et al.* [39] showed the relationship between the distance between the cameras and the depth resolution. Meanwhile, Sahabi and Basu [40] analyzed the correlation between the vergence angle and the depth resolution with stereo images. Our results are compatible with those of the two studies.

Second, a stereo lighthouse is more precise than a single lighthouse. As jitter quantitatively represents precision [38], the jitters of the stereo lighthouse show much better results than those of the single lighthouse (Fig. 11). Furthermore, the jitters tend to increase when the distance between the localization system and the tracked object increases.

We initially supposed that we use lighthouse tracking technology as a tracking method for medical or research purpose demanding rather small area but high accuracy. Conclusively, for stereo lighthouse, longer distance and larger vergence angle between lighthouses would make better accuracy at pose estimation. However, optimal vergence angle and distance for best tracking accuracy is not obtained yet. It may require several and intensive experiments with various conditions. For single lighthouse, frontal facing the lighthouse and the tracked object is more critical than other factors. Short distance between the lighthouse and the tracked object is good for tracking considering jitter for both lighthouse configurations.

4) Comparison With Other Research: Before comparing our position estimation result with those of previous studies, we briefly mention something related to the comparison. We selected three papers among several papers dealing with the accuracy of the lighthouse system for quantitative comparison. Because of the difference in the test area, the number of tests, test apparatus, etc., the comparative evaluation has some limits.

Through the comparison, we made several observations. First, the results of our experiments are reliable compared with those of the previous studies. The results of the same stereo camera used by Groves [20] and ours are very similar. The accuracy is 0.49 mm and 0.21 mm, respectively, and the jitter is 0.04 mm and 0.11 mm, respectively (Table III). One reason

TABLE III
ACCURACY AND JITTER COMPARISON AMONG SEVERAL RESEARCH

Paper of method		Accuracy [mm]	Jitter [mm]	# of points	Tracked object
Ours	Single L.[a]	2.20	3.24	96	Our own tracked object
	Stereo L.[a]	1.13	0.14		
	Stereo C.[b]	0.49	0.04		-
Groves [20]	Lighthouse	2.63	0.20	48	VIVE controller
	Stereo C.[b]	0.21	0.11		-
Yang [15]		1.47 (0.79/2.82)[c]	15.21 (4.53/36.58)[c]	60	Their own tracked object
Ameler [23]		0.56	0.21	7	VIVE tracker

[a]Lighthouse, [b]Camera
[c]Results of Yang's study [15] show a large difference according to the distance between the lighthouse and the tracked object. Hence, we divided it into two cases in parentheses.

for the difference in accuracy could be from the different positioning devices that carry the tool detected by the stereo camera. We used the xyz stage, whereas Groves [20] used a computer numerical control (CNC) machine. The difference in moving distance between our experiment (50–150 mm) and Groves' (50 mm) may also have affected the results. The reason for the difference in the jitter results may be due to the difference in the number of samples: 250 samples in ours and 25 samples in Groves' [20].

Second, our proposed stereo method shows compatible accuracy and precision comparing those of other studies. Although Ameler's work [23] shows the best result, the number of test points is relatively small compared to other results (Table III).

Third, the precision of the customized tracked object is lower than that of commercial tracked objects. Our single lighthouse and Yang's [15] result has a higher jitter than Groves' [20] and Ameler's [23]. The difference between the two groups is the tracked object. Both Yang's and our method used fabricated tracked objects, whereas Groves [20] and Ameler [23] used commercial objects. In our opinion, the number of photodiodes and manufacturing quality may explain the difference in jitter. The precision is enhanced when the stereo configuration is applied to the same device and circumstance.

V. CONCLUSION AND FUTURE WORK

To enhance the accuracy of the lighthouse system, we proposed a stereo lighthouse configuration. For this configuration, we proposed a calibration method for a stereo lighthouse using two linear micro-stages and one photodiode. We also described the procedures for obtaining the position of one photodiode using PnP and triangulation methods. Specifically, we compared the accuracy of two lighthouse setups for pose estimation: a single lighthouse with PnP and a stereo lighthouse with triangulation through experiments. The stereo lighthouse with vergence gave the best result among all the tested configurations and showed better accuracy than those of previous researches.

The lighthouse localization system could be used for research purposes under certain circumstances. Niehorster *et al.* reported that the lighthouse system has a slanted reference plane that causes error and the calibration to fix the slanted plane is difficult when tracking is lost. In a small area where the risk of tracking loss is low, the lighthouse system could be used for experiments, albeit with care [14]. Groves *et al.* [20] also insisted that the lighthouse system can be used for research and medical purposes. We agree with the opinion of the two previous research groups. If the conditions of a line of sight and a small detecting area are satisfied, experiments with the lighthouse system are possible. For more accurate experiments, we recommend the stereo configuration of the lighthouse system. However, currently, the accuracy and precision of the lighthouse system are lower than those of the stereo camera.

Our final goal is to improve the accuracy of lighthouse pose estimation up to the level of that of a commercial stereo camera system, which is widely used in medical and research applications. We suggested stereo configuration of lighthouse for better accuracy with a calibration method and also showed that vergence angle and distance between lighthouses influence the estimation accuracy. With improved accuracy, the lighthouse system can be used for various applications, for example, neuro-navigation [41]. The accuracy can be improved by increasing the tracking resolution. To achieve this, the rotation speed of the rotors of the lighthouse needs to be changed or the MCU of the tracked object needs to have a different clock frequency. These will be addressed in our future work.

REFERENCES

[1] G. Burdea and P. Coiffet, *Virtual Reality Technology*. New York, NY, USA: Wiley, 2004, pp. 19–21.
[2] B. H. Thomas, "A survey of visual, mixed, and augmented reality gaming," *Comput. Entertainment*, vol. 10, no. 1, pp. 1–33, Oct. 2012.
[3] S. R. Barber, S. Jain, Y.-J. Son, and E. H. Chang, "Virtual functional endoscopic sinus surgery simulation with 3D-printed models for mixed-reality nasal endoscopy," *Otolaryngol.-Head Neck Surg.*, vol. 159, no. 5, pp. 933–937, Nov. 2018.
[4] L. P. Berg and J. M. Vance, "Industry use of virtual reality in product design and manufacturing: A survey," *Virtual Reality*, vol. 21, no. 1, pp. 1–17, 2017.
[5] K. K. Bhagat, W.-K. Liou, and C.-Y. Chang, "A cost-effective interactive 3D virtual reality system applied to military live firing training," *Virtual Reality*, vol. 20, pp. 127–140, Apr. 2016.
[6] D. M. Markowitz, R. Laha, B. P. Perone, R. D. Pea, and J. N. Bailenson, "Immersive virtual reality field trips facilitate learning about climate change," *Frontiers Psychol.*, vol. 9, p. 2364, Nov. 2018.
[7] H. Li *et al.*, "Integrating design and construction through virtual prototyping," *Autom. Construct.*, vol. 17, no. 8, pp. 915–922, Nov. 2008.
[8] T. P. Kersten *et al.*, "The Selimiye mosque of Edirne, Turkey—An immersive and interactive virtual reality experience using HTC Vive," in *Proc. Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci. (ISPRS)*, vol. XLII-5/W1, May 2017, pp. 403–409.
[9] M. Borges, A. Symington, B. Coltin, T. Smith, and R. Ventura, "HTC Vive: Analysis and accuracy improvement," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 2610–2615. [Online]. Available: https://doi.org/10.1109/IROS.2018.8593707
[10] B. R. Sorensen, M. Donath, G.-B. Yang, and R. C. Starr, "The Minnesota scanner: A prototype sensor for three-dimensional tracking of moving body segments," *IEEE Trans. Robot. Autom.*, vol. 5, no. 4, pp. 499–509, Aug. 1989.
[11] T. M. Hedges, H. Takagi, T. Pratt, and M. J. Sobel, "Position measurement system and method using cone math calibration," U.S. Patent 6 535 282 B2, Mar. 18, 2003. [Online]. Available: https://patents.google.com/patent/US6535282B2/en
[12] A. Yates and J. Selan, "Positional tracking systems and methods," U.S. Patent 10 338 186 B2, Jul. 2, 2019. [Online]. Available: https://patents.google.com/patent/US10338186B2/en

[13] S. Islam, B. Ionescu, C. Gadea, and D. Ionescu, "Indoor positional tracking using dual-axis rotating laser sweeps," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, May 2016, pp. 1–6. [Online]. Available: https://doi.org/10.1109/I2MTC.2016.7520559

[14] D. C. Niehorster, L. Li, and M. Lappe, "The accuracy and precision of position and orientation tracking in the HTC Vive virtual reality system for scientific research," *i-Perception*, vol. 8, no. 3, pp. 1–23, 2017. [Online]. Available: https://doi.org/10.1177/2041669517708205

[15] Y. Yang, D. Weng, D. Li, and H. Xun, "An improved method of pose estimation for lighthouse base station extension," *Sensors*, vol. 17, no. 10, p. 2411, 2017. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/29065509

[16] S. P. Kleinschmidt, C. S. Wieghardt, and B. Wagner, "Tracking solutions for mobile robots: Evaluating positional tracking using dual-axis rotating laser sweeps," in *Proc. ICINCO*, vol. 1. Hannover, Germany, 2017, pp. 155–164.

[17] D. Laurijssen, S. Truijen, W. Saeys, W. Daems, and J. Steckel, "Six-DoF pose estimation using dual-axis rotating laser sweeps using a probabilistic framework," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, 2017, pp. 1–8. [Online]. Available: https://doi.org/10.1109/IPIN.2017.8115913

[18] A. Borrego, J. Latorre, M. Alcañiz, and R. Llorens, "Comparison of oculus rift and HTC Vive: Feasibility for virtual reality-based exploration, navigation, exergaming, and rehabilitation," *Games Health J.*, vol. 7, pp. 151–156. Jan. 2018.

[19] M. J. Pomianek, M. Piszczek, and M. Maciejewski, "Testing the SteamVR trackers operation correctness with the OptiTrack system," in *Proc. 13th Conf. Integr. Opt., Sens., Sens. Struct., Methods*, Aug. 2018, Art. no. 108300D.

[20] L. A. Groves, P. Carnahan, D. R. Allen, R. Adam, T. M. Peters, and E. C. S. Chen, "Accuracy assessment for the co-registration between optical and VIVE head-mounted display tracking," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 7, pp. 1207–1215, Jul. 2019. [Online]. Available: https://doi.org/10.1007/s11548-019-01992-4

[21] E. Luckett, T. Key, N. Newsome, and J. A. Jones, "Metrics for the evaluation of tracking systems for virtual environments," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Mar. 2019, pp. 1711–1716.

[22] G. Verdelet *et al.*, "Assessing spatial and temporal reliability of the VIVE system as a tool for naturalistic behavioural research," in *Proc. Int. Conf. 3D Immersion (IC3D)*, Brussels, Belgium, Dec. 2019, pp. 1–8.

[23] T. Ameler *et al.*, "A comparative evaluation of SteamVR tracking and the OptiTrack system for medical device tracking," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 1465–1470. [Online]. Available: https://doi.org/10.1109/EMBC.2019.8856992

[24] T. A. Jost, G. Drewelow, S. Koziol, and J. Rylander, "A quantitative method for evaluation of 6 degree of freedom virtual reality systems," *J. Biomech.*, vol. 97, Dec. 2019, Art. no. 109379.

[25] S. P. Sitole, A. K. LaPre, and F. C. Sup, "Application and evaluation of lighthouse technology for precision motion capture," *IEEE Sensors J.*, vol. 20, no. 15, pp. 8576–8585, Aug. 2020.

[26] P. Azad, T. Asfour, and R. Dillmann, "Stereo-based vs. monocular 6-DoF pose estimation using point features: A quantitative comparison," in *Autonome Mobile Systeme* (Informatik aktuell). Berlin, Germany: Springer, 2009, pp. 41–48.

[27] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 1997, pp. 1106–1112.

[28] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000. [Online]. Available: https://doi.org/10.1109/34.888718

[29] NDI Medical, Waterloo, ON, Canada. (2013). *Hybrid Polaris Spectra User Guide*. [Online]. Available: https://usermanual.wiki/Pdf/HybridPolarisSpectraUserGuide.95551077.pdf

[30] G. Bradski, "The OpenCV library," *Dr. Dobb's J. Softw. Tools*, vol. 25, no. 11, pp. 120–123, 2000.

[31] A. Shtuchkin. *DIY Position Tracking Using HTC Vive's Lighthouse*. Accessed: Sep. 2021. [Online]. Available: https://github.com/ashtuchkin/vive-diy-position-sensor

[32] Z. Zhang, "A flexible new technique for camera calibration," Microsoft Corp., Redmond, WA, USA, Tech. Rep. MSR-TR-98-71, 2008. [Online]. Available: http://opilab.utb.edu.co/topics-computer-vision/pdfs/PAMI_2000_Zhang.pdf

[33] V. Lepetit, F. Moreno-Noguer, and P. Fua, "EPnP: An accurate O(n) solution to the PnP problem," *Int. J. Comput. Vis.*, vol. 81, no. 2, p. 155, 2009. [Online]. Available: https://doi.org/10.1007/s11263-008-0152-6

[34] R. Fletcher. (1971). *A Modified Marquardt Subroutine for Non-Linear Least Squares*. [Online]. Available: https://www.osti.gov/biblio/4667484

[35] M. Balda, "An algorithm for nonlinear least squares," in *Proc. Tech. Comp. Conf.*, Praha, Czech Republic, Nov. 2007. [Online]. Available: http://dsp.vscht.cz/konference_matlab/MATLAB07/prispevky/balda_m/balda_m.pdf

[36] D. H. Eberly, *3D Game Engine Design: A Practical Approach to Real-Time Computer Graphics*. San Diego, CA, USA: Academic, 2006, pp. 38–77.

[37] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 4, no. 4, pp. 629–642, 1987. [Online]. Available: https://doi.org/10.1364/JOSAA.4.000629

[38] R. Khadem *et al.*, "Comparative tracking error analysis of five different optical tracking systems," *Comput. Aided Surg.*, vol. 5, no. 2, pp. 98–107, Jan. 2000. [Online]. Available: https://doi.org/10.3109/10929080009148876

[39] M. Kytö, M. Nuutinen, and P. Oittinen, "Method for measuring stereo camera depth accuracy based on stereoscopic vision," in *Proc. Three-Dimensional Imag., Interact., Meas. Int. Soc. Opt. Photon.*, vol. 7864, Jan. 2011, Art. no. 78640I. [Online]. Available: https://doi.org/10.1117/12.872015

[40] H. Sahabi and A. Basu, "Analysis of error in depth perception with vergence and spatially varying sensing," *Comput. Vis. Image Understand*, vol. 63, no. 3, pp. 447–461, May 1996. [Online]. Available: https://doi.org/10.1006/cviu.1996.0034

[41] H. Shin, W. Ryu, S. Cho, W. Yang, and S. Lee, "Development of a spherical positioning robot and neuro-navigation system for precise and repetitive non-invasive brain stimulation," *Appl. Sci.*, vol. 9, no. 21, p. 4561, Oct. 2019. [Online]. Available: https://doi.org/10.3390/app9214561

Sungtaek Cho received the B.S. degree in civil engineering from Kyung Hee University, Gyeonggi-do, South Korea, in 2006, and the M.S. degree in mechatronics engineering from Chungnam National University, Deajeon, South Korea, in 2014. He is currently pursuing the Ph.D. degree in electronic engineering with Hanyang University, Ansan, South Korea. From 2014 to 2015, he worked at Korea Institute of Science and Technology (KIST) as an Intern Researcher. From 2015 to 2018, he worked at Industry-University Research Cooperation Foundation, Hanyang University, as a Researcher. His research interests include medical navigation systems, 3D pose estimation systems, and robotics.

Dongyeon Kim received the B.S. and M.S. degrees in electronic engineering from Hanyang University, Gyeonggi-do, South Korea, in 2019. He works as a Test Engineer at SK Hynix. His research in Hanyang University was about pose estimation with photo sensors and verification of sensor.

Sungon Lee (Member, IEEE) received the B.S. degree in mechanical design and product engineering from Seoul National University, Seoul, South Korea, in 1997, the M.S. degree in mechanical engineering from POSTECH, Pohang, South Korea, in 1999, and the Ph.D. degree in mechano-informatics from The University of Tokyo, Tokyo, Japan, in 2008.

From 2010 to 2012, he was a Postdoctoral Research Fellow with the Center for Systems Biology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. From 1999 to 2015, he was a Senior Research Scientist with Korea Institute of Science and Technology, Seoul. Since 2015, he has been with the School of Electrical Engineering, Hanyang University, Ansan, South Korea, where he is currently an Associate Professor. His research interests include biomedical robotics, surgical robots, medical image processing, mobile robots, and motion compensation systems.