

Article

A Mobile Robot Localization via Indoor Fixed Remote Surveillance Cameras [†]

Jae Hong Shim ¹ and Young Im Cho ^{2,*}

¹ Department of Mechatronics Engineering, Korea Polytechnic University, Si-Heung, Gyunggi-do 429-793, Korea; jhshim@kpu.ac.kr

² Department of Computer Engineering, Gachon University, Sung-Nam, Gyunggi-do 461-701, Korea

* Correspondence: yicho@gachon.ac.kr; Tel.: +82-31-750-5800; Fax: +82-31-750-5768

[†] This paper is an extended version of our paper published in *Procedia Computer Science*. Shim, J.H.; Cho, Y.I. Mobile robot localization using external surveillance cameras at indoor. In the Proceedings of the CHARMS 2015 Workshop, Belfort, France, 17–20 August 2015; Volume 56; pp. 502–507.

Academic Editors: Eric T. Matson, Byung-Cheol Min and Donghan Kim

Received: 31 October 2015; Accepted: 2 February 2016; Published: 4 February 2016

Abstract: Localization, which is a technique required by service robots to operate indoors, has been studied in various ways. Most localization techniques have the robot measure environmental information to obtain location information; however, this is a high-cost option because it uses extensive equipment and complicates robot development. If an external device is used to determine a robot's location and transmit this information to the robot, the cost of internal equipment required for location recognition can be reduced. This will simplify robot development. Thus, this study presents an effective method to control robots by obtaining their location information using a map constructed by visual information from surveillance cameras installed indoors. With only a single image of an object, it is difficult to gauge its size due to occlusion. Therefore, we propose a localization method using several neighboring surveillance cameras. A two-dimensional map containing robot and object position information is constructed using images of the cameras. The concept of this technique is based on modeling the four edges of the projected image of the field of coverage of the camera and an image processing algorithm of the finding object's center for enhancing the location estimation of objects of interest. We experimentally demonstrate the effectiveness of the proposed method by analyzing the resulting movement of a robot in response to the location information obtained from the two-dimensional map. The accuracy of the multi-camera setup was measured in advance.

Keywords: localization; mobile robot; surveillance camera; indoor; homography

1. Introduction

Recently, various localization technologies for mobile robots have been studied with respect to acquiring accurate environmental information. Typically, mobile robots have self-organized sensors to obtain environmental information. However, such robots are very expensive to manufacture and have complicated body structures. The structural layout of indoor environments is typically a known state; thus, many localization studies have obtained the required information from external sensors installed on a robot. The infrared light, ultrasonic, laser range finder, RFID (Radio Frequency Identification), and RADAR (Radio Detecting and Ranging) are the popularly used sensors for localization. Hopper *et al.* [1] presented an active sensing system, which uses infrared emitters and detectors to achieve 5–10 m accuracy. However, this sensor system is not suitable for high-speed application as the localization cycle requires about 15 s and always requires line of sight. Ultrasonic sensor uses the time of flight measurement technique to provide location information [2]. However, the ultrasonic sensor requires a great deal of infrastructure for its high effectiveness and accuracy.

Laser distance measurement is executed by measuring the time that it takes for a laser light to be reflected off a target and returned back to the sensor. Since the laser finder is a very accurate and quick measurement device, this device is widely used in many applications. Subramanian *et al.* [3] and Barawid *et al.* [4] proposed a localization method based on a laser finder. The laser finder was used to acquire environment distance information that can be used to identify and avoid obstacles during navigation. However, their high performance relies on high hardware costs. Miller *et al.* [5] presented an indoor localization system based on RFID. RFID based localization used RF tags and a reader with an antenna to locate objects, but detection of each tag can only work over about 4–6 m. Bahl *et al.* [6] and Lin *et al.* [7] introduced the RADAR system, which is a radio-frequency (RF) based system for locating and tracking users inside buildings. The concept of RADAR is to measure signal strength information at multiple stations positioned to provide overlapping field of coverage. It aggregates real measurements with signal propagation modeling to determine object location, thereby enabling location-aware applications. The accuracy of the RADAR system was reported by 2–3 m.

Recently, visual image location systems have been preferred because they are not easily disturbed by other sensors [2,8,9]. Sungho [10] used workspace landmark features as external reference sources. However, this method has insufficient accuracy and is difficult to install and maintain due to the required additional equipment. Kim *et al.* [11] proposed the augmented reality techniques to achieve an average location recognition success rate of 89%, though the extra cost must be considered. Cheoket *et al.* [12] developed a method of localization and navigation in wide indoor areas by using a vision sensor. Though the set-up cost is lower, this method is not easy to implement if users do not have knowledge about the basic concept of electronic circuit analysis. Recently, two camera localization systems have been proposed [13,14]. The concept of this system is that the object distances can be calculated by a triangular relationship from two different images of the cameras. However, to ensure the measuring reliability, the relative coordinates between two cameras must be maintained at the same position. In addition, the set-up cost of the experimental environment is quite expensive due to the use of two cameras.

Nowadays, surveillance systems exist in most modern buildings, and cheap cameras are usually installed around these buildings. Indoor surveillance cameras are typically installed without blind areas, and visual data are transferred to a central data server for processing and analysis. If a mobile robot can determine its position using indoor cameras, it would not require an additional sensor for localization and could be applied to multi-agent mobile robot systems [15,16]. Kuscue *et al.* [17] and Li *et al.* [18] proposed a vision-based localization method using a single ceiling mounted surveillance camera. However, there are several problems that must be addressed prior to the realization of this concept. First, lens distortion arises from the poor-quality lenses in surveillance cameras, and shadow effects are produced by indoor light sources [19,20]. Second, information about occluding objects cannot be obtained using a single camera. Third, calibrations of camera and a map for localization are carried out independently, and it is very time-consuming work.

Herein, we propose a localization method for a mobile robot to overcome the abovementioned problems associated with indoor environments. A two-dimensional map containing robot and object position information is constructed using several neighboring surveillance cameras [21]. The concept of this technique is based on modeling the four edges of the projected image of the field of coverage of the camera and an image processing algorithm of the finding object's center for enhancing the location estimation of objects of interest. This approach relies on coordinate mapping techniques to identify the robot in the environment using multiple ceiling-mounted cameras. It can be applied for localization in complex indoor environments like T- and L-shaped environments. In addition, simultaneous calibration of cameras and a two-dimensional map can be carried out. Via the above modeling process, a 2D map is built in the form of air-view and quite accurate location can be dynamically acquired from a scaled grid of the map. Significant advantages of the proposed localization are its minimal cost, simple calibration and little occlusion, where it needs multiple ceiling-mounted inexpensive cameras that are installed in opposition to each other and wirelessly communicate with the mobile robot and

update its current estimated position. Moreover, we experimentally demonstrate the effectiveness of the proposed method by analyzing the resulting robot movements in response to the location information acquired from the generated map.

2. Two-Dimensional Visual Map by Using the Homograph

2.1. Projected Image Plane for Two Dimensional Map

Indoor surveillance cameras are typically installed to view the same object from opposite directions. Such images contain ground-area information that may be occluded by objects, as shown in Figure 1. Therefore, two object images viewed from opposite directions must be combined into a single image. We have attempted to accomplish this using homography.

Homography is a projection wherein a plane is transformed into another plane in space. A surveillance camera is mounted on a slant to obtain an image, as shown in Figure 2a. To observe the position and size of an object viewed from the camera, the image in Figure 2a is transformed into the air-view image of Figure 2b using homography.

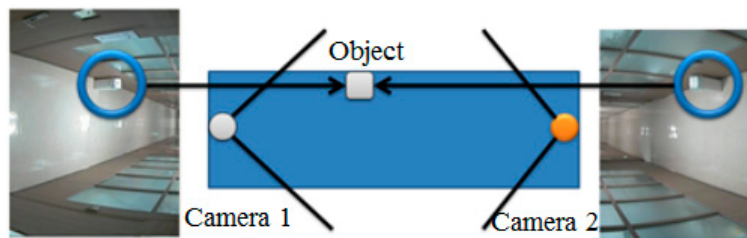


Figure 1. Images viewed from two neighboring indoor surveillance cameras.

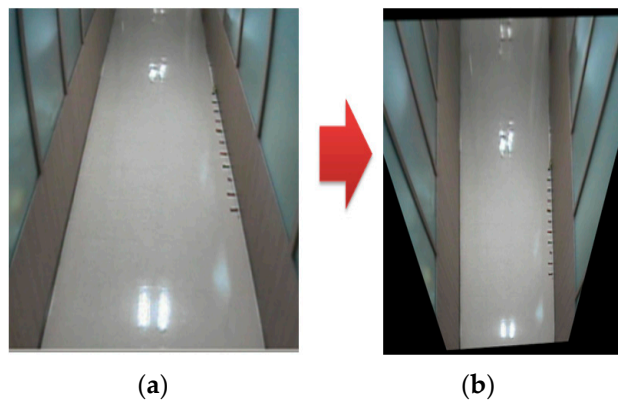


Figure 2. (a) Original image; (b) Air view image of (a).

To transform an original image from a surveillance camera into an air-view image, a feature point Q of the original image is matched with the corresponding point q of the air-view image, as shown in Figure 3. We used a large placard of a chess board to match feature points between the original and air-view images.

Using Equation (1), a homography matrix H is obtained using four points from both plane Q and plane q

$$q = HQ \quad (1)$$

The resulting homographically transformed positions of the same feature points of these two images from two surveillance cameras are combined in a new projected plane, as shown in Figure 4①. This results in a single united plane, as shown in Figure 4②. We can then construct a two-dimensional map, as shown in Figure 4③, by extracting the region of interest (ROI) of the actual floor area from

the single plane. The homography transformation process makes the floor width of the projected image to be spread evenly like the air-view image. At that time, the distortion of the camera could be compensated together.

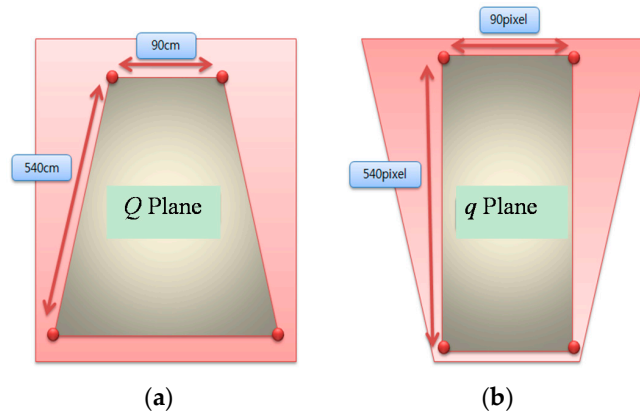


Figure 3. (a) Q plane of the original image; (b) q plane of the air view image of (a).

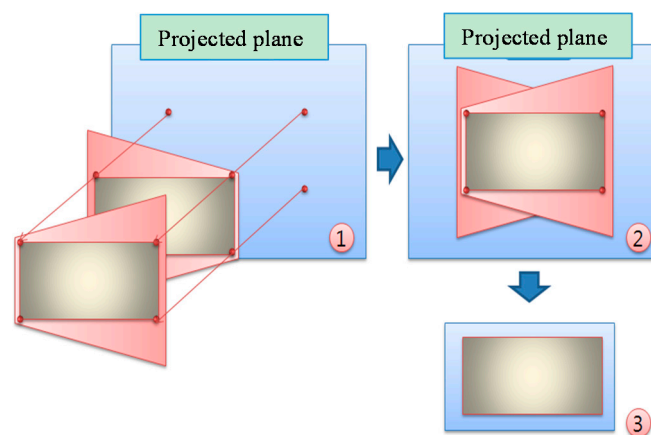


Figure 4. Extracting the two-dimensional map from two projected images by homography. ① Two images on the projected plane; ② A single united plane; ③ An extracted two dimensional map.

If multiple cameras are used for localization in more complex indoor environments, rotational relationships between image planes of the cameras are considered as shown in Equation (3).

$$q = \begin{bmatrix} n_x & o_x & a_x \\ n_y & o_y & a_y \\ n_z & o_z & a_z \end{bmatrix} HQ \quad (2)$$

where $\bar{n} = [n_x \ n_y \ n_z]$, $\bar{o} = [o_x \ o_y \ o_z]$, $\bar{a} = [a_x \ a_y \ a_z]$ are normal, orientation and approach unit vectors, respectively.

2.2. Object Modeling on the Two-Dimensional Map

Here, we present a method to acquire the position and size of an object image on the two-dimensional map. Since two neighboring cameras view an object from opposite directions, their images of the same object differ. However, the floor area occupied by the object is the same in the two images. Therefore, if the rest of the image (except for the floor area) is deleted, we can obtain the actual floor area of the object on the two-dimensional map. Even if an object is looked at from opposite

directions by two cameras, image correspondence on the two-dimensional map can be obtained by adopting area features of the object, *i.e.*, center position and size of the area.

To obtain the object floor area on the two-dimensional map, two projected images are transformed from their original images, as shown in Figure 5. Figure 5a shows the original images from the two cameras, Figure 5b shows binary images of Figure 5a with shadow effects removed, and Figure 5c shows the projected images of Figure 5b obtained by homography. If the coordinates of the two projected images are similar, the actual size and position of an object in contact with the floor, as represented by the image, are nearly the same. If the rest of the images not including the floor area are eliminated, the object image can be expressed with an air view. Equation (3) represents the size and position of an object on the projected plane $H(x, y)$.

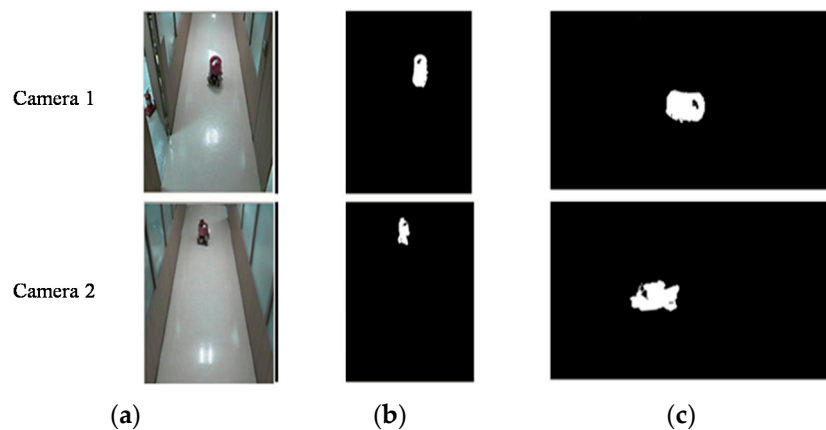


Figure 5. Projected images of two surveillance cameras by homography. (a) Original images; (b) Binary images; (c) Projected images.

$$H(x, y) = \begin{cases} 1, & \text{if } [I_1^H(x, y) \& I_2^H(x, y)] = 1 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where $I_1^H(x, y)$ and $I_2^H(x, y)$ are the projected images of cameras 1 and 2, respectively.

The common area between cameras 1 and 2, as shown in Figure 5c, can be presented as shown in Figure 6. The area is the object image on the two-dimensional map based on homography.

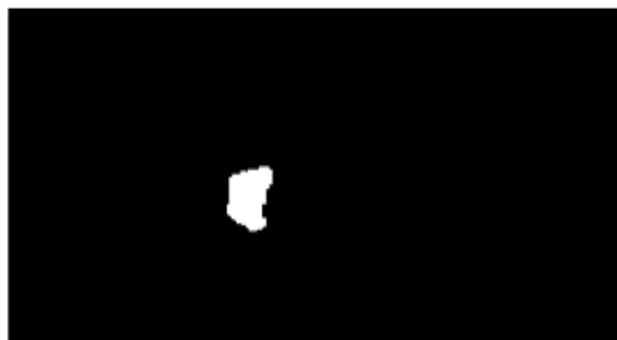


Figure 6. Common area of the projected images of cameras 1 and 2 in Figure 4 (3).

Now, the size and center position of the object image on the two-dimensional map can be calculated. Generally, labeling or a contour technique is used to detect the area shape of an object in the image. These techniques are suitable for detecting the area shape from an image, such as Figure 6, which is a binary image. We apply the contour technique to rapidly determine its size and center position. Then, the moment of the area shape is calculated to obtain its center point and area.

The moment is used to measure the size of the area shape. We obtain the size and center position of the object area from the edge information via the abovementioned process.

To compare the object area with its real position, the actual floor image is transformed onto a projected plane and the object area is detected by contour processing, as shown in Figure 7.

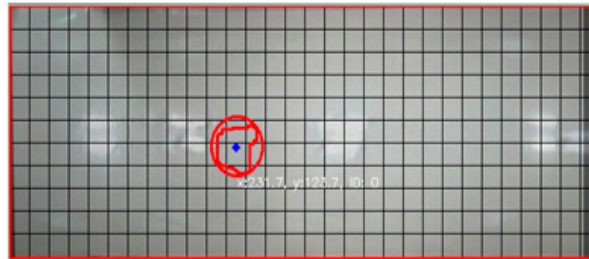


Figure 7. Two-dimensional map combined with the object area and the projected floor image.

2.3. Calibration of the Detected Object Area

When the size and center position of the object area obtained from contour processing are compared to the actual values of the object, considerable errors are revealed. We then measure the difference between the position of the real and visually detected object using a measurement grid, as shown in Figure 8.

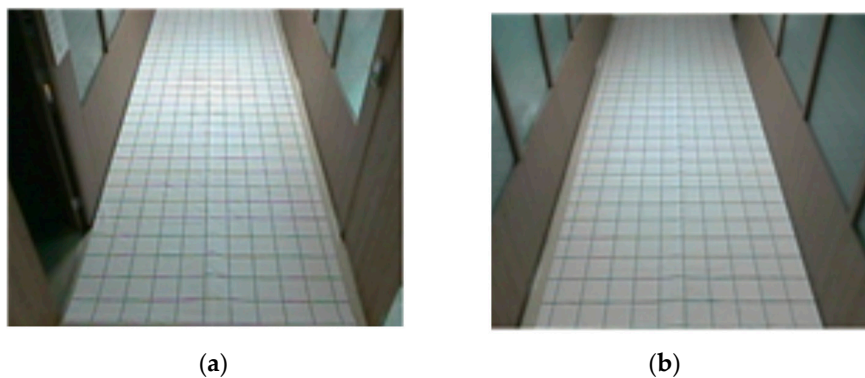


Figure 8. Grid placard for measuring the position error of the visually detected object area (a) camera 1; (b) camera 2.

In a calibration experiment, we used a cylindrical object (diameter, 20 cm). Figure 9 shows position errors between the real (black line) and visually detected center position (red line) of the object.

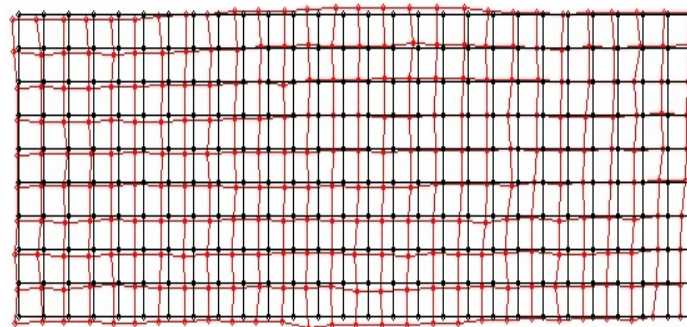


Figure 9. First experimental results of position errors between the real and the visually detected center position of the object.

There is considerable difference between the real and visually detected center position of the object, as shown in Figure 9. We consider that the error is caused by scale changes in the image projection and the installation error of the measurement grid. After the initial measurement to detect the object area, error compensation was performed using homography. When the visually detected grids were mapped to the floor image in an air view, we obtained the position with an error bound of 7.1 cm on the two-dimensional map. Figure 10 shows a representation of the error-compensated results of the object area, as detected by homography.

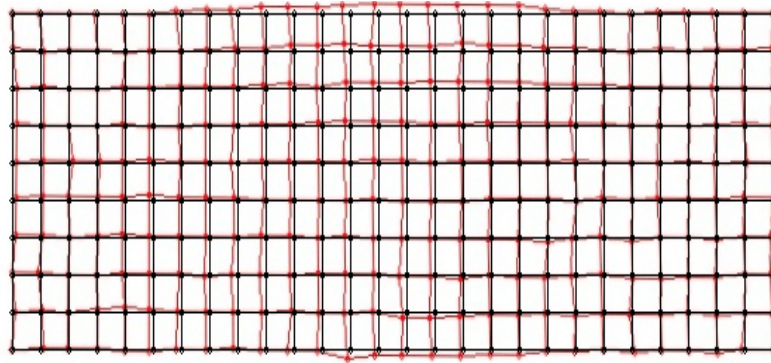


Figure 10. Error compensation results for the object area detected by homography.

3. Topology Map Building and Optimal Path

Path planning is required for a robot to safely move without colliding into any object placed on the two dimensional map (Section 2). Herein, we employ the thinning algorithm [22]. The thinning algorithm leaves a single pixel in the center after continuous elimination of contour in random areas. Thus, a path by which a robot can avoid obstacles and move safely is generated by the thinning algorithm.

After generating a moving path with the thinning algorithm, as shown in Figures 11 and 12 a movement indicator is required for a robot. Thus, a topological map is generated to create the path indicator. However, the algorithm can generate a path that is difficult for a robot to move along; thus, eliminating information about such paths by checking the area around nodes is required. To make robots move through nodes, an area around each node that is larger than that of the robot should be examined to eliminate nodes and edges that cannot be traversed by the robot. Note that a robot cannot pass if there is an object in the search field around a node and if the search field exists beyond the boundary of an image around a node.

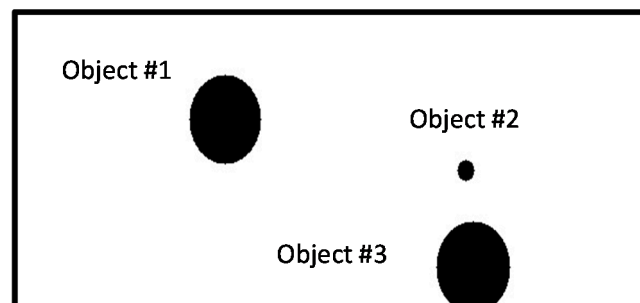


Figure 11. Example image of objects on a two-dimensional map before applying the thinning algorithm.

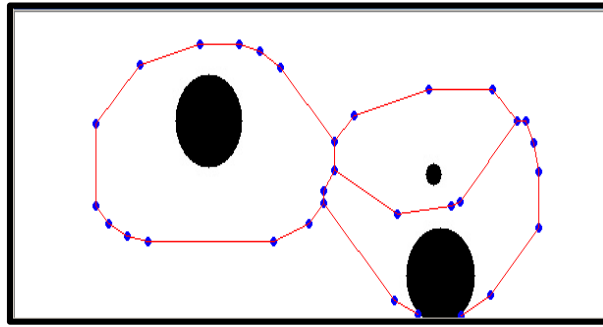


Figure 12. Topology map after applying the thinning algorithm.

Figure 13 shows an image formed after removing a searched path by which a robot cannot pass. A path by which a robot could pass was generated with a topology map by applying the thinning algorithm.

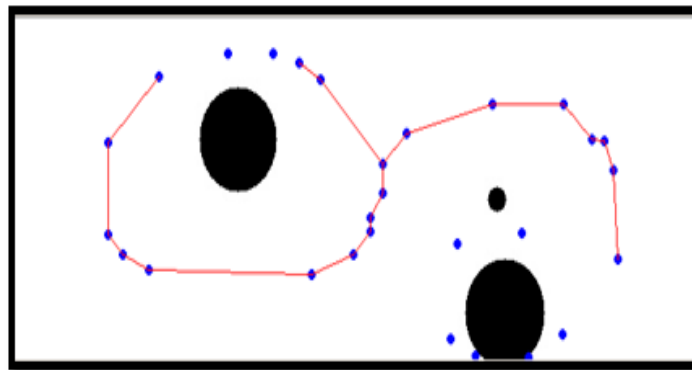


Figure 13. Modified driving path via searching path.

The A* algorithm is a graph exploring algorithm that calculates an optimal driving path with a given starting point and goal [23]. It uses a heuristic estimate on each node to estimate the shortest route to the target node with minimal calculation.

Figure 14 shows the shortest robot path estimated by the A* algorithm. A robot moves along nodes on the estimated path.

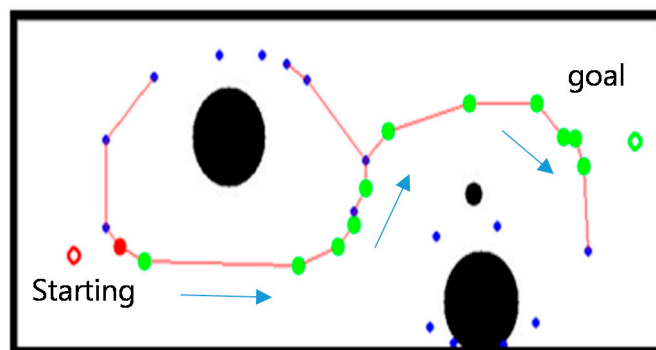


Figure 14. Shortest robot path by A* algorithm.

4. Experimental Results

We performed a series of experiments to demonstrate the effectiveness of the proposed two-dimensional-map-based localization method using indoor surveillance cameras. The width and length of the floor viewed by the two neighboring cameras were 2.2 and 6 m, respectively. The detected two-dimensional map by homography represents the area of the floor viewed by the two cameras in an air view. We used a self-developed mobile robot with an omnidirectional wheel in the experiment. The surveillance camera had a resolution of 320×240 pixels and three RGB (Red Green Blue) channels.

The accuracy of the two dimensional map with the proposed method was experimentally obtained. Each position error in Figure 15 has two x - and y -axis components in a plane. To represent two error components as a single parameter at each position, we suggest the position error estimation shown in Figure 16. Here, x_{real} and y_{real} mean the real position, and x_{measur} and y_{measur} represent the visually detected object position on the two dimensional map. x_e and y_e are obtained using Equation (4).

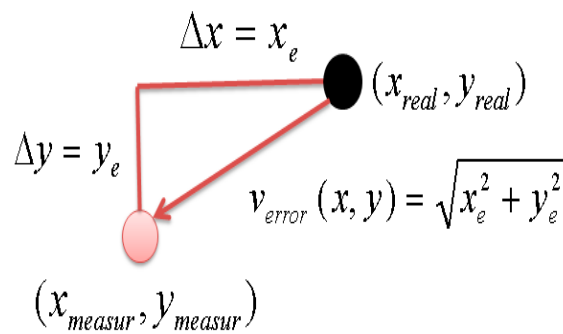


Figure 15. Position error estimate.

$$x_e = x_{real} - x_{measur}, \quad y_e = y_{real} - y_{measur} \quad (4)$$

Position error v_{error} is composed of x_e and y_e in Equation (4) and is expressed by Equation (5). The magnitude of position error estimate v_{error} is obtained using Equation (6).

$$v_{error}(x, y) = [x_e, y_e] \quad (5)$$

$$|v_{error}(x, y)| = \sqrt{x_e^2 + y_e^2} \quad (6)$$

We define the position error estimate $E(x, y)$ as the absolute value of v_{error} using Equation (7).

$$E(x, y) = |v_{error}(x, y)| \quad (7)$$

Here, $E(x, y)$ is the error plane of the difference between the real and visually detected object position. To examine the effectiveness of the proposed error compensation by homography (Section 2), we compared $E(x, y)$ before and after error compensation. The position error estimate before error compensation is shown in Figure 16. The x - y plane of Figure 16 is the x - y plane of the two-dimensional map, and the z plane is the value of $E(x, y)$. The maximum and average values of $E(x, y)$ are 11.5 and 6.7 cm, respectively. After error compensation by homography, the maximum and average values of $E(x, y)$ are 7.1 and 2.6 cm, respectively, as shown in Figure 17. The maximum position error was decreased by 38% through the proposed error compensation method. The accuracy of the two-dimensional map obtained by two ceiling surveillance cameras is 7.1 cm, which is sufficient for mobile robot localization.

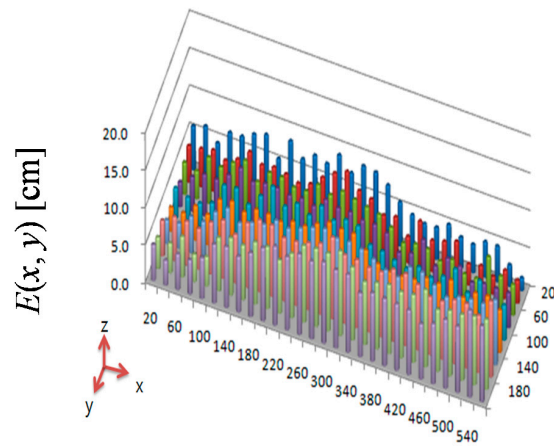


Figure 16. Position error estimate before error compensation.

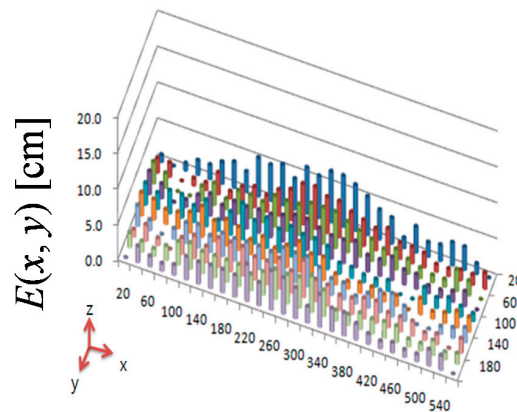


Figure 17. Position error estimate after error compensation.

Figure 18 shows two images from the two neighboring surveillance cameras. There are several objects on the floor. A mobile robot was controlled to move from one position to the opposite position using the proposed localization based on the two-dimensional map described in Section 2.2. We used the A* algorithm as the path planning method for the mobile robot. The objects on the floor were detected by homography as the object area in the projected plane, and the robot's moving path was planned considering the object area in the two-dimensional map. The experimental results of the robot's path control are shown in Figure 19. The error bounds between the planned and actual movement path of the robot was ± 5 cm. This means that the proposed localization method may be effective for indoor mobile robots.

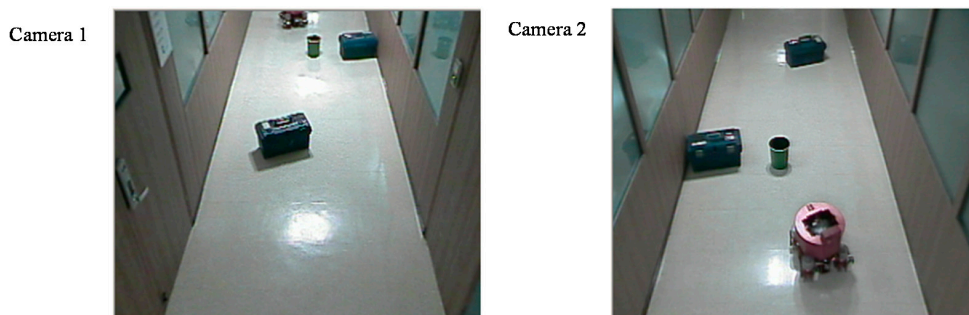


Figure 18. Objects in the experimental environment.

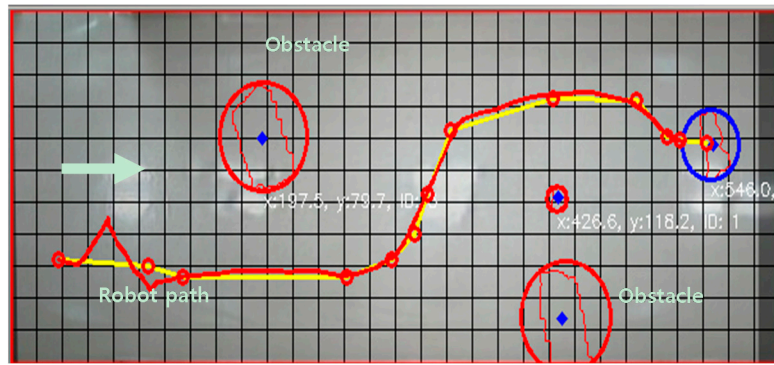


Figure 19. Experimental results of the robot's path control using the proposed localization method.

To show that the proposed method can be applied for complex indoor environments, an experiment was carried out at a T-shaped indoor environment. As shown in Figure 20a, three surveillance cameras were used to build a two-dimensional map. Figure 20b–d show images from the three cameras.

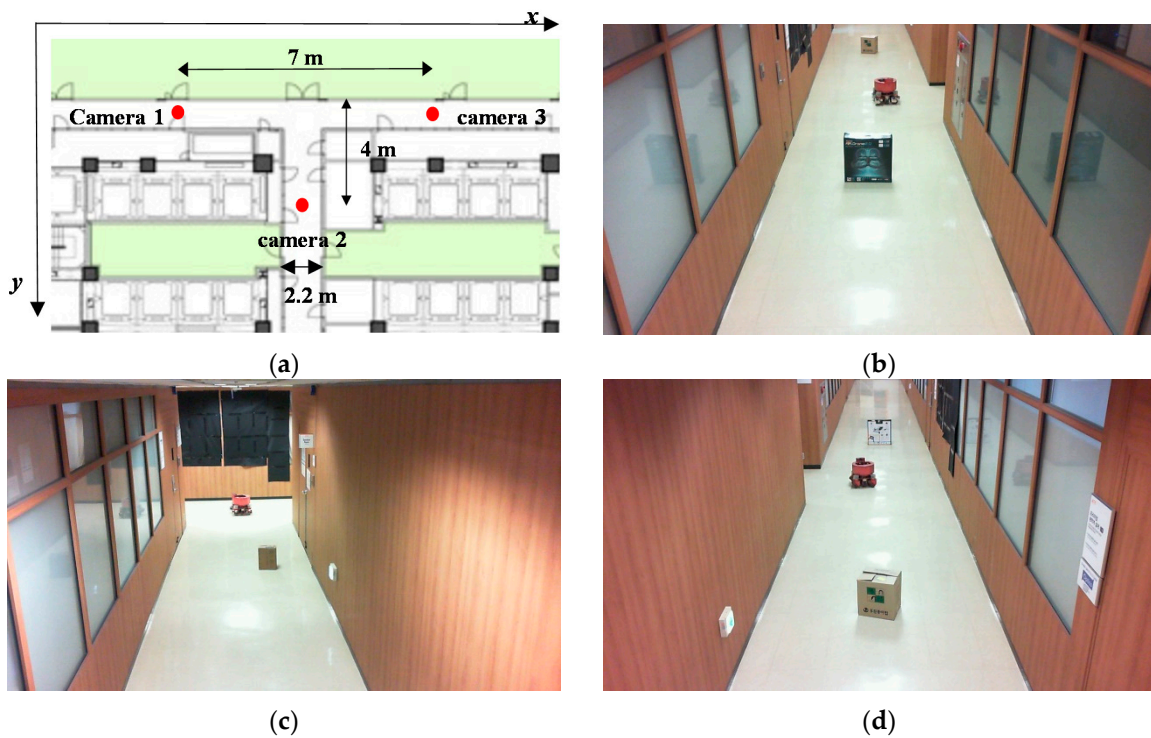


Figure 20. (a) Map of our experimental environment; (b) Image of camera 1; (c) Image of camera 2; (d) Image of camera 3.

Figure 21 shows the two-dimensional map using the proposed method. The accuracy of the two-dimensional map by three ceiling surveillance cameras is 7 cm, which is satisfactory for mobile robot localization.

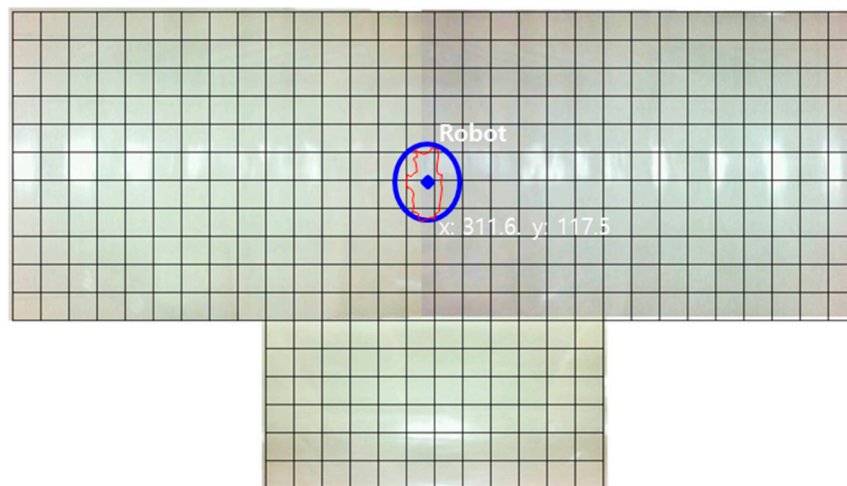


Figure 21. Experimental results of the two-dimensional map using the proposed localization method.

5. Conclusions

We have proposed a new vision-based approach for mobile-robot localization in an indoor environment using multiple remote ceiling-mounted cameras. The proposed approach uses a two-dimensional mapping technique between camera and ground-image plane coordinate systems. We used homography to transform the image planes. Two camera-image planes were combined into a single ground-image plane with an air view, which resulted in a two-dimensional map. The position error bound of the developed two-dimensional map was within 7.1 cm. We performed a series of experiments to demonstrate the effectiveness of the proposed two-dimensional-map-based localization method. Among several obstacles fixed on the floor, the mobile robot successfully maneuvered to its destination position using only the two-dimensional map without the help of any other sensor. In future, we plan to extend the proposed method to the localization of multiple mobile robots in an indoor environment.

Acknowledgments: This work was supported by the National Research Fund of Korea under Grant NRF 20151D1A1A01061271.

Author Contributions: Young Im Cho proposed the research topic. Jae Hong Shim developed the algorithms and performed the experiments; Jae Hong Shim and Young Im Cho analyzed the experimental results and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hopper, A.; Harter, A.; Blackie, T. The active badge system. In Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, Amsterdam, the Netherlands, 24–29 April 1993; pp. 533–534.
2. Ni, L.M.; Lau, Y.C.; Patil, A.P. Indoor location sensing using active RFID. *Wirel. Netw.* **2004**, *10*, 701–710. [[CrossRef](#)]
3. Subramanian, V.; Burks, T.F.; Arroyo, A.A. Development of machine vision and laser radar based autonomous vehicle guidance systems for citrus grove navigation. *Comput. Electr. Agric.* **2006**, *53*, 130–143. [[CrossRef](#)]
4. Barawid, O.C.; Mizushima, H.; Ishii, K. Development of an autonomous navigation system using a two-dimensional laser scanner in an orchard application. *Biosyst. Eng.* **2007**, *96*, 139–149. [[CrossRef](#)]
5. Guerrieri, J.R.; Francis, M.H.; Wilson, P.F.; Kos, T.; Miller, L.E.; Bryner, N.P.; Stroup, D.W.; Klein-Berndt, L. RFID-assisted indoor localization and communication for first responders. In Proceedings of the First European Conference on Antennas and Propagation, Nice, France, 6–10 November 2006; pp. 1–6.

6. Bahl, P.; Padmanabhan, V. RADAR: An in-building RF based user location and tracking system. In Proceedings of the Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, Tel Aviv, Israel, 26–30 March 2000.
7. Lin, T.H.; Chu, H.H.; You, C.W. Energy-efficient boundary detection for RF-based localization system. *IEEE Trans. Mob. Comput.* **2009**, *8*, 29–40.
8. Patridge, K.; Arnstein, L.; Borriello, G.; Whitted, T. Fast intrabody signalling. In Proceedings of the Demonstration at Wireless and Mobile Computer Systems and Applications, Monterey, CA, USA, 7–8 December 2000.
9. Hills, A. Wireless Andrew. *IEEE Spectr.* **1999**, *36*, 49–53. [[CrossRef](#)]
10. Sungho, H. Ultrasonic interference reduction technique in indoor location sensing system. *J. Korea Acad. Ind. Cooper. Soc.* **2012**, *13*, 364–369.
11. Kim, J.B.; Jun, H.S. Vision-based location positioning using augmented reality for indoor navigation. *IEEE Trans. Consum. Electron.* **2008**, *54*, 954–962. [[CrossRef](#)]
12. Choek, A.D.; Yue, L. A novel light-sensor-based information transmission system for indoor positioning and navigation. *IEEE Trans. Instrum. Meas.* **2011**, *60*, 290–299. [[CrossRef](#)]
13. Osugi, K.; Miyauchi, K.; Furui, N.; Miyakoshi, H. Development of the scanning laser radar for ACC system. *JSAE Rev.* **1999**, *20*, 549–554. [[CrossRef](#)]
14. Nakahira, K.; Kodama, T.; Morita, S.; Okuma, S. Distance measurement by an ultrasonic system based on a digital polarity correlator. *IEEE Trans. Instrum. Meas.* **2001**, *50*, 1478–1752. [[CrossRef](#)]
15. Tsai, C.Y.; Song, K.T. Dynamic visual tracking control of a mobile robot with image noise and occlusion robustness. *Image Vis. Comput.* **2007**, *27*, 1007–1022. [[CrossRef](#)]
16. Purarjay, C.; Ray, J. External cameras & a mobile robot: A collaborative surveillance system. In Proceedings of the Australasian Conference on Robotics and Automation, Sydney, Australia, 2–4 December 2009; pp. 1–10.
17. Kuscua, E.; Rababaah, A.R. Mobile robot localization via efficient calibration technique of a fixed remote camera. *J. Sci. Inform.* **2012**, *2*, 23–32.
18. Li, I.H.; Chen, M.H.C.; Wang, W.Y.; Su, S.F.; Lai, T.W. Mobile robot self-localization system using single webcam distance measurement technology in indoor environment. *Sensors* **2014**, *14*, 2089–2109. [[CrossRef](#)] [[PubMed](#)]
19. Philipp, A.; Henrik, C. Behaviour coordination for navigation in office environment. In Proceedings of the AIEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne, Switzerland, 30 September–4 October 2002; pp. 2298–2304.
20. Huuiying, C.; Kohsei, M.; Jun, O. Self-calibration of environmental camera for mobile robot navigation. *J. Robot. Auton. Syst.* **2007**, *55*, 177–190.
21. Jaehong, S.; Youngim, C. A mobile robot localization using external surveillance cameras at indoor. In Proceedings of the International Workshop on Communication for Human, Agents, Robots, Machines and Sensors 2015 (CHARMS2015), Belfort, France, 17–20 August 2015.
22. Gonzalez, R.; Woods, R. *Digital Image Processing*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2007; pp. 671–675.
23. Hart, P.; Nelson, N.; Raphael, B. A formal basis for the heuristic determination of minimum paths in graphics. *IEEE Trans. Syst. Sci. Cybern.* **1968**, *4*, 100–107. [[CrossRef](#)]

