Indoor Navigation for Visually Impaired Using AR Markers

Guojun Yang and Jafar Saniie

Department of Electrical and Computer Engineering

Illinois Institute of Technology

Chicago, Illinois, United States

Abstract—**Helping people with visual difficulties navigating indoor is challenging. Unlike most of outdoor navigation systems, indoor navigation devices are not able to receive satellite signals. Moreover, indoor navigation usually requires higher positioning accuracy when compared with outdoor applications. In this paper, we propose an indoor positioning system based on AR (augmented reality) markers and computer vision techniques. In the proposed system, devices with camera such as smart phone will be used to provide position information by reading AR markers placed in the environments, subsequently help users gain awareness of their locations. With the assistance of proposed system, visually impaired people may be able to navigate indoor by following the instruction generated by their smart phones. Method to register indoor environment with AR markers will also be discussed in this paper. To analyze the accuracy of the system, data collected in experiment will also be presented.**

I. INTRODUCTION

Without visual perception, it is hard for people to gain awareness of their surrounding environments. The lack of awareness often come along with inconveniences, sometime even dangers. There are attempts on building variety of systems and tools to assist the visually impaired people. Such as using ultrasonic sensors to detect cars while user crossing streets [1]. While navigating indoor, using ultrasonic is also possible [2]. There are also systems rely on computer vision techniques to help visually impaired people recognize objects such as road signs, cars, pedestrians and so on [3] [4]. What these solutions have in common is they all aiming at helping the visually impaired people by locating certain objects, so that they can interact with or avoid these objects. However, we proposed a method that helps visually impaired people, by allowing them have spatial awareness.

Spatial awareness is fundamental for navigation. There are some existing infrastructures helping people with visual difficulties to navigate in urban area both indoor and outdoor. For example, braille, the language system representing common characters using raised dots [5]. Braille can be found on the bottoms in elevators or hand rails of train subway stations. By touching braille, visually impaired people are able to *read* information contained, so they can get directions just as people reading overhead signs. While the braille is one effective and widely adopted solution, which gives visually impaired people

sense of direction, there is one major disadvantage. Before getting any direction, a visually impaired person must find brailles. Unlike normal people searching overhead signs visually, visually impaired people can only search braille within the reach of their hands. It is challenging due to the limited searching range, especially for places that one has never visited. To assist visually impaired people more effectively, a system needs not only providing the location of user, but also providing direction to destination, regardless its user's location.

In this paper, we propose a novel solution that is able find one's position using AR makers as reference. While perform positioning, there are two major issues need to be resolve. One is to find an object or point that has a fixed or known position. The other is to find the position of the object to be positioned in relative to the reference object or point. For example, while using GPS, satellites are performing as reference objects, since they orbiting earth, while through triangulation, a GPS device can be located relative to satellites, thus the device can be located. Similar concepts are also adopted by other positioning applications, such as method uses wireless signals indoor to achieve indoor navigation [6] [7]. Given these examples, we developed our system aiming at finding the absolute position of AR markers and estimating the relative position between device and AR markers.

In the following sections, the algorithm and techniques used for calculate pose and position of camera in respect to AR markers will be introduced. Then the result from experiment will be presented and discussed. Finally, the indoor positioning system and potential application will be presented.

II. POSITIONING USING AR MARKER

In this section, two major topics will be discussed. One is the method used to calculate the relative position between camera and AR marker. The other is the procedure of registering AR markers in an environment.

A. Calculate Position in Relative to AR Markers

It is a challenging task to estimate pose and position of camera when given images or videos only [8]. But with the assist of AR markers, estimating camera pose and position become easier. AR marker is one kind square-based fiducial marker implemented for estimating the camera poses and position for augmented reality applications. The AR markers

used in these experiments, were generated using ArUco library, which adopted MILP (Mixed Integer Linear Programming) algorithm to encode identification numbers (IDs) into the AR marker. Generating AR markers using MILP assures AR markers has proper hamming distance with AR marker that has different ID. More importantly, AR markers generated using MILP algorithm are rotation variant. Given the fact each AR marker has a square shape, to distinguish its rotation in respect to the real-world coordinate system, each AR marker has hamming distance with itself after being rotated by 90, 180 or 270 degree [9]. With such feature, it enables algorithms to estimate the pose of the camera in respect to the AR marker.

The position and pose of camera in respect to real-world coordinate system can be model as 6-DOF (Dimension of Freedom) movement. 6-DOF includes three dimension of rotation and three dimension of translation. As shown in Figure 1, a 3D space is defined using a three-dimension plane. Any object can be rotated or moved along axis x, y and z.
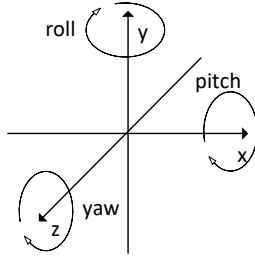

Figure 1. Illustration of 6 DOF movement

While the AR markers are designed for camera pose and position estimation, such estimation still relies on projective geometry. Projective geometry models taking photos as similar projection. Such process is usually described as equation 1, X is the vector describes the position of an object in real-world coordinate. R and t are 3 by 3 rotation matrix and 3 by 1 translation vector respectively. R describes rotation of the camera relative to the real-world coordinate, while t describes spatial translation of the camera in respect to the origin of real-world coordinate. While X is 3 by 1 vector represents a point in real-world coordinate. K is 3 by 3 camera intrinsic matrix, which is defined by the character of the camera, these parameters constant once the camera found [10]. Finally, x is a 2 by 1 vector of the position of the object on image plane.

$$x = K(RX + t) \tag{1}$$

Estimating the pose and position of the camera can be achieved by finding the value of R and t. When calculating the pose of the camera, all four corners of the AR marker must be detected in the image, meaning multiple corresponding x and X are known [11]. Moreover, since there are four corners need to be detected in image, yielding four equations, with 2 by 1 vector. With such information, it is possible to solve rotation matrix, denoted by R. As to the translation vector t, due to the natural of similarity, t vector can only be calculated up to scale. Because similar objects can create identical projection on image plane if they been placed at distances proportion to size. To solve this problem, sizes of AR markers need to be fixed, yet such information need to be provided to the algorithm.

While navigating indoor, surrounding environments are unpredictable, occlusion is unavoidable. In such scenario, using only one AR marker would not be able to provide reliable position of the camera constantly. Therefore, multiple AR markers are required to minimize the impact due to occlusion. To provide consistent position and pose information of the camera while using different AR markers, the system must be forfeit following requirements:

- All markers should be generated by the same dictionary.
- All AR markers should have the same size.
- Each AR marker should have different ID.
- At least one AR marker need to be visible for positioning.
- All AR markers must be registered in the same coordinate.

To elaborate, using AR markers generated by the same dictionary means all markers have the same code word length and parity bits, so that they can be detected through the same process. As discussed previously in this section, due to the similarity issue, the size of AR marker need to be fixed for solving t vector. Using AR markers with different IDs help the system to differentiate markers placed at different locations. Occlusion may occur at certain location, it can be solved by increasing the number of markers in environment. Amount these requirements, the most challenging one is register all AR markers in one consistent coordinate system. To make the registration process easier, we proposed a novel process to finish the registration.

B. Incremental Registration of AR Markers

The incremental registration process help finding the position and pose of AR markers through an incremental process. As shown in Figure 2, to find the position and pose of an unregistered AR marker, camera must has visual on at least of one AR marker that has known pose and position relative to real-world coordinate, in this example, AR marker 1 in Figure 2. Then move the camera to gain visual on an AR marker with unknown position and pose, in this case AR marker 2. As mentioned in previous section, the program is able to find out the poses and positions of 2 AR markers in respect to the camera. With such information, and the pose and position of first AR marker, the pose and position of AR marker 2 in respect to real-world coordinate system can be calculated. When camera move again and capture AR marker 3, the position and pose of AR marker 3 can be calculated using AR marker 2.
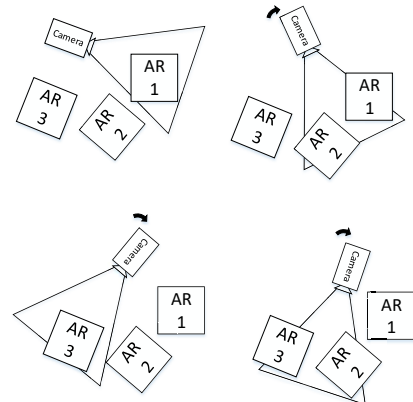

Figure 2. Illustration of incremental registration

By such incremental process, any AR marker can be registered, as long as one registered AR marker is captured in the same frame. To model such process, we use R_1 and R_2 to represent the pose (rotation matrix) of AR marker 1 and AR marker 2 in respect to the real-world coordinate system, t_1 and t_2 as the position (translation vector) of these markers. While registering AR marker 2 using a registered marker 1 as a reference, the rotation and translation of camera in respect to these markers can be obtained as R'_1, R'_2, t'_1 and t'_2 respectively. To find out the pose (rotation) as R_2 and position (translation) as t_2 of the AR marker 2, we should solve the pose R_c and position t_c of camera relative to the real-world coordinate. Then the pose and position of the second marker can be calculated as shown in equation 2.

$$R_c = R_1(R'_1)^{-1}$$
$$t_c = t_1 - R_1(R'_1)^{-1}t'_1 \tag{2}$$

Knowing camera pose R_c and position t_c and its rotation R'_2 and translation t'_2 relative to the second AR marker, the pose and position of the second AR marker can be calculated as shown in equation (3.

$$R_2 = R_cR'_2$$
$$t_2 = t_c + R_1(R'_1)^{-1}t'_2 \tag{3}$$

By adopting the incremental registration process, the pose and position of any AR marker can be computed efficiently, as long as another registered AR marker is visible. The pose and position of the camera can also be calculated by equation (3).

Such process can make the registration process easier and faster. That means, when putting the AR markers in to an environment for the first time, one can simple put the AR markers at desired positions regardless their poses and positions. After measuring the position and pose of the first AR marker in respect to real-world coordinate system, the first reference marker has been created, then the rest of the markers can be registered one by one using incremental registration. Eventually, all AR markers will be registered in a database, which includes the ID number, pose and position of each marker.

III. ACCURACY ANALYSIS

To verify the accuracy of positioning while using AR markers, a series of experiments were conducted. Measuring the relative pose and position of 2 objects with 6-DOF is challenging. To minimize the error created by measurement, during these experiments, we print AR markers on the same piece of paper, thus we have the liberty to specify the angel of rotation of an AR marker relative to another using document editing software with higher precision.

A. Experiment Setup

As mentioned earlier, to minimize the measuring error, two different AR markers were print on the same piece of paper, one with ID 88 the other 66. The center of the marker with ID 88 is always considered as the origin of real-world coordinate. As shown in Figure 3, the position and pose of ID 66 marker is generated in respect to the pose and relation of the marker with ID 88. The real-world coordinate system is defined as such, the origin of the coordinate locates at the center of first marker (ID

88), the x axis and y axis belongs to the surface of marker and perpendicular to upper and right edges of the AR marker, while z axis is the normal vector of the surface, as shown in Figure 3. The second marker is place at location $(78, -118, 0)$ in real-world coordinate system, with $-\frac{2\pi}{3}$ rad of rotation along z axis.
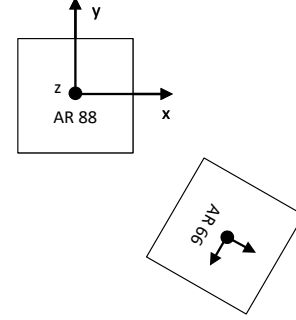

Figure 3. AR markers used for error analysis

Knowing the true pose and position of the second AR marker, the pose and position calculated by the system can be evaluated.

During the experiment, the paper with 2 AR markers were place in front of the camera, AR markers were printed with dimension 85mm by 85mm, the camera was recording with resolution 640 by 480 pixel. A video clip was recorded while the paper with markers was constantly moving in front of the camera. Using the video, the algorithm is able to find pose position information of AR marker with ID 66 in respect to the AR marker with ID 88 which used as real-world coordinate system.

B. Error Analyze

In those video clips, we randomly picked 576 frames with both markers appeared and calculate the location of the marker with ID 66. As shown in Figure 4, the point cloud (blue dots) of calculated positions of AR marker with ID 66. The red dot in the figure is the true location of the AR marker.
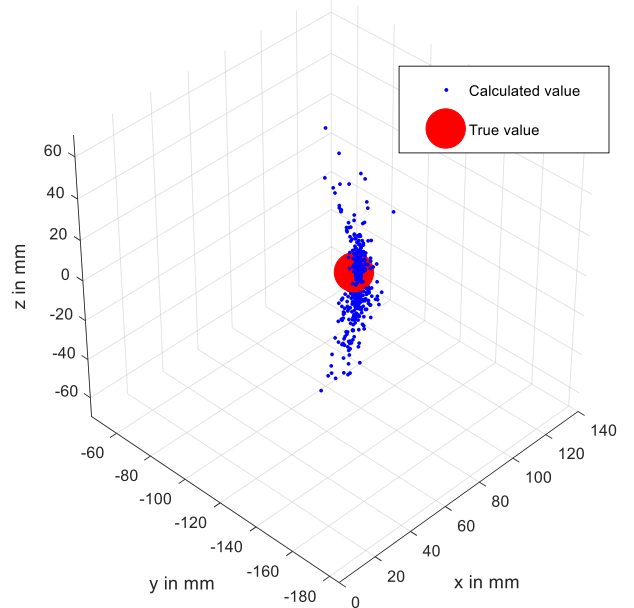

Figure 4. Calculated position of marker compared with true value

To further analyze the calculated position with true value, we split the position information by the projection on each axis. As

shown in Figure 5, positions calculated by each sample (drawn as blue line) are compared with true position on each axis (drawn as red line).
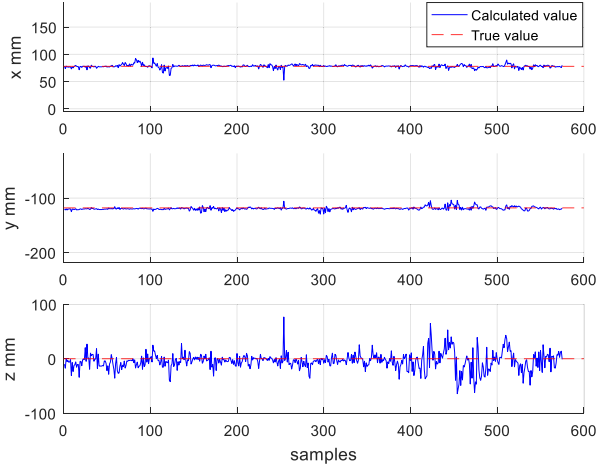

Figure 5. The calculate position compared with true value

Similar process was conducted to analyze the error in pose calculation, results are shown in Figure 6. Since the 2 markers can only be printed on the same piece of paper, therefore, it can only be rotated alone z axis, in other word yaw.
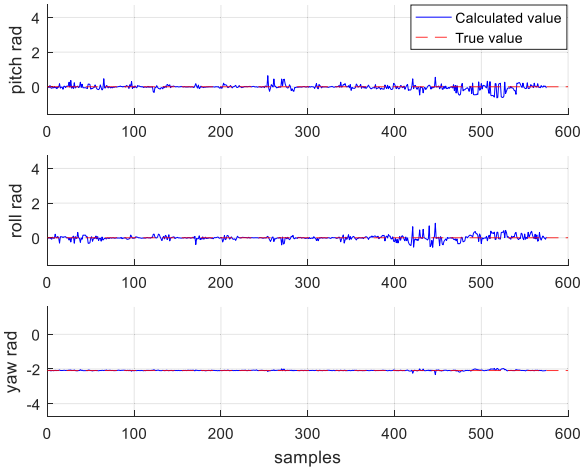

Figure 6. The calculate pose compared with true value

Results obtained in the experiment demonstrated the capability of such system to conduct indoor navigation. The data analyzed was collected without any filtering or post processing, yet the error is lower than GPS which is within meters. Moreover, most of data was generated by the camera recording at least resolution possible. The data set of position estimation in z direction generates most of the noise. This error can be reduced with larger AR marker or use images with higher resolution as the input of the algorithm.

IV. NAVIGATION INDOOR

Despite the fact, that with AR markers and proper computer vision techniques it's possible to compute the pose and position of an AR marker, to assist visual impaired persons navigating indoor, a systematic design is still needed. The system need to combine the computer vision technique, a human computer interface and proper protocols. To be more specific, the system need to use computer vision technique to locate the AR markers, then figure out where the user is, and instruct the user accordingly.

More importantly, while using AR markers as references, in order to specify what the AR markers is related to, a protocol need to be created. For example, when using AR marker to assist visually impaired navigating in a shopping mall, there are 2 major ways to apply the AR markers to the environment. The first one is the more intuitive and easier one. Appling AR markers directly to places such as stores and restrooms. As shown in Figure 7.
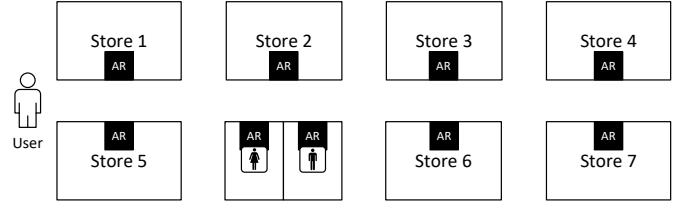

Figure 7. The markers apply to locations

Once AR markers applied to locations, the advantage is obvious, no tedious registration process is needed when creating the database of the location and ID number of each marker. The user can specify the destination and the system find ID of the AR marker accordingly. However, such method requires one unique ID for each location, when apply such method in larger indoor environment, the total ID pool will not be able to provide sufficient ID numbers.

Another approach for registration is put every location and AR marker in to a unified coordinate system, as shown in Figure 8.
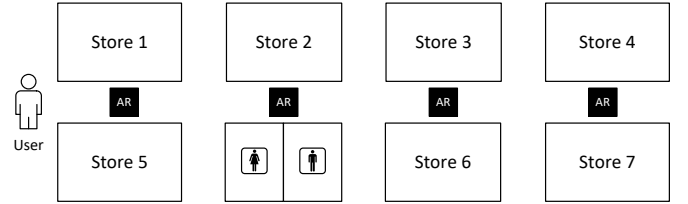

Figure 8. Markers applied as reference points

When all AR markers applied to indoor environment as shown in Figure 8, the system is able to find the location of user in respect to AR markers, then relative location to the destination, since all markers and locations are registered in a same coordinate system. By doing so, the amount of AR markers can be reduced significantly. Such method can also reduce losing sense of location due to occlusion, by detecting alternative markers. Another advantage of using AR markers as reference points is it can achieve turn by turn navigation in indoor environments. As shown in Figure 9, turn by turn indoor navigation can be achieved once all AR markers and locations are registered.
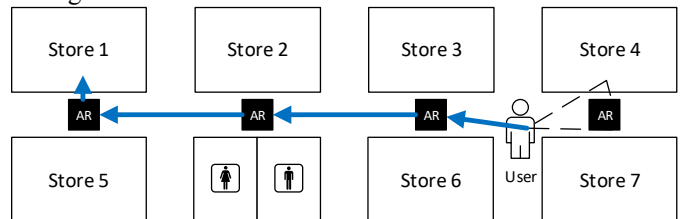
Figure 9. Use AR marker as to achieve turn by turn navigation

In Figure 9, an example of using AR markers to achieve indoor navigation is demonstrated, user start as any location with a smart phone that can detect an AR marker. Then the system find the location of the AR marker as well as the destination, in this case store 1. A series of markers that in the path will be found, markers lie under the blue arrow. Finally, the system will provide voice instruction to assist the user to navigate to the destination.

V. Conclusion

In this paper, we presented a system that developed to help visually impaired people to navigate in indoor environments. The system relies on AR markers to estimate the pose and position of device in respect to real-world coordinate system. With the incremental registration process we proposed, any environment can be easily registered with AR markers using cellphones and cameras. Despite the fact such indoor navigation technique is proposed to assist visually impaired person, similar techniques can also be applied to robotic applications, where the device have no GPS system.

References

[1] S. Hashino and S. Yamada, "An Ultrasonic Blind Guidance System for Street Crossing," in *Computers Helping People with Special Needs: 12th International Conference*, Vienna, 2010.

[2] B. Ilias, S. A. A. Shukor, A. H. Adom, M. F. Ibrahim and S. Yaacob, "A novel indoor mobile robot mapping with USB-16 ultrasonic sensor bank and NWA optimization algorithm," in *2016 IEEE Symposium on Computer Applications Industrial Electronics (ISCAIE)*, 2016.

[3] T. Gonnot and J. Saniie, "Intergrated Machine Vision and Communication System for Blind Navigation and Guidance," in *2016 IEEE International Conference on Electro Infornmation Technology*, 2016.

[4] T. Gonnot and J. Saniie, "Image Sensing System for Navigaing Visually Impaired People," in *2013 IEEE SENSORS*, Baltimore, 2013.

[5] A.-S. Abdul Malik S, E.-Z. Ali, A.-S. Yousef, A.-H. Khaled and G. Abdu, "Designing Braille Copier Based on Image Processing Techniques," *International Journal of Soft Computing and Engineering (USCE) ISSN,* pp. 2231-2307.

[6] V. Kulyukin, C. Gharpure, J. Nicholson and S. Pavithran, "RFID in robot-assisted indoor navigation for the visually impaired," in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, 2004.

[7] B. Kim, M. Kwak, J. Lee and T. T. Kwan, "A multi-pronged approach for indoor positioning with WiFi, magnetic and cellular signals," in *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2014.

[8] G. Yang, Z. Zhou, T. Gonnot and J. Saniie, "Design flow of motion based single camera 3D mapping," in *2015 IEEE International Conference on Electro/Information Technology (EIT)*, 2015.

[9] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas and R. Medina-Carnicer, "Generation of Fiducial Marker Dictionaries Using Mixed Integer Linear Programming," *Pattern Recognition,* vol. 51, pp. 481-491, 2016.

[10] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 22, pp. 1330-1334, 2000.

[11] L. Quan and Z. Lan, "Linear N-point camera pose determination," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 21, no. 8, pp. 774-780, 1999.