Development of Wide-Area Tracking System for Augmented Reality

Hirotake Ishii[1], Hidenori Fujino[1], Bian Zhiqiang[1], Tomoki Sekiyama[1], Toshinori Nakai[1], and Hiroshi Shimoda[1]

[1]Graduate School of Energy Science, Kyoto University,
611-0011 Kyoto, Japan
{hirotake, fujino, bianzq, sekiyama, nakai, shimoda}@uji.energy.kyoto-u.ac.jp

Abstract. In this study, two types of marker-based tracking systems for Augmented Reality were developed: a tracking system using line markers and a tracking system using circular markers. Both markers were designed for use inside buildings such as nuclear power plants, which are very complicated and in which it is difficult to paste many large markers. To enlarge the area in which tracking is available using only a limited number of markers, a hybrid tracking method and a two-layer tracking method were also developed. The experimental evaluation shows that both methods can increase the available area compared to legacy methods such as single-camera tracking and simple square markers.

Keywords: Augmented reality, Marker-based tracking, Wide area tracking

1   Introduction

A promising application of Augmented Reality (AR) is the maintenance support of large plants such as nuclear power plants and chemical plants. The safety and efficiency of the maintenance work can be improved by applying AR, for example, by navigating workers toward their work place and indicating dangerous locations [1]. To apply AR to such maintenance support tasks, some improvements of tracking technology that measures the user's viewing position and orientation in real time are necessary because existing tracking technologies are affected easily by metal, magnetic fields, and sounds, and therefore cannot be used in the plants. The authors, therefore, designed two types of fiducial markers and developed algorithms to recognize them on camera images using image-processing techniques and those results are used for tracking [2]. Using the newly designed markers and algorithms, it became possible to execute tracking inside the plant buildings both stably and accurately. Nevertheless, a persistent problem is that many markers must be pasted and their positions must be measured in advance to execute tracking in a wide area. Preparation of the markers is too burdensome for practical use. In this study, therefore, a hybrid tracking method and a two-layer tracking method were developed to enlarge the area in which tracking can be done using only small number of markers.

2 Tracking Method Using Line Markers

2.1 Design of Line Marker and Tracking Algorithm

A typical marker tracking method is ARToolKit [3]. However, ARToolKit can recognize a marker only when the distance between the marker and the camera is very short. We must use very large markers when we want to use ARToolKit on the condition that the distance between the markers and the camera is very long. Unfortunately, it is difficult to paste large square markers inside buildings such as nuclear power plants because many instruments are installed inside such buildings: few large spaces can be used for the markers. However, many pipes with thin rectangular surfaces exist in nuclear power plants and chemical plants. The authors considered that these surfaces are useful for pasting the markers and designed line-shaped markers as shown in Fig. 1. The line marker is a combination of black elements: each element corresponds to one bit. The square element signifies "0". The double-sized rectangle element denotes "1". The total number of elements is 11. Therefore, one line marker can express an 11-bit code. Of its 11 bits, 7 express the ID, whereas the remaining 4 bits are the Hamming code. Using this coding method, 128 kinds of line markers can be made with arbitrary one-bit error correction. The shape of the line marker is almost a single *line*. Consequently, it is easy to paste them on the pipes that are ubiquitous in the plants. Details of how to recognize the markers on images captured by a camera are described in the literature [2].
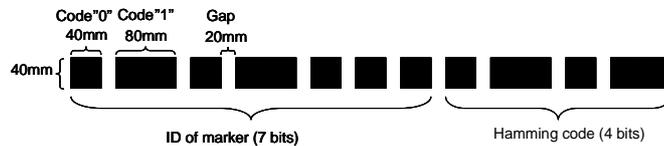

Fig. 1 Line marker design.

To obtain a unique solution of the relative position and orientation between the camera and the markers, at least four feature points must be recognized on the camera image. These feature points must not be on the same line and the accuracy depends on the distances between the feature points. Therefore, the authors decided to obtain two feature points from each line marker (two terminals of the line marker), so that at least two line markers must be recognized on the camera image simultaneously.

The P3P solution algorithm and nonlinear algorithm are adopted to calculate the relative position and orientation from n feature points and minimize the error between the estimated position of the four feature points calculated using the tracking result and the actual position on the camera image.

2.2 Performance Evaluation of Line Marker

A laboratory experiment was conducted to evaluate the maximum detection distance and the accuracy of the developed methods. Six line-code markers of about 50 cm

were pasted in a room (10 m × 20 m), as shown in Fig. 2. Then the camera was moved in the room with 1 m steps. The estimated and actual positions and orientations of the camera were compared.

Figure 3 shows the estimated camera position error. The maximum detection distance at which the markers were recognized without any failure was about 11 m and the maximum distance with some recognition failure was about 17 m under the condition that an XGA black-and-white camera with 6.37 mm focal length lens was used. About 96% of the tracking was executed with less than 30 cm position errors when the markers were recognized without any failure.
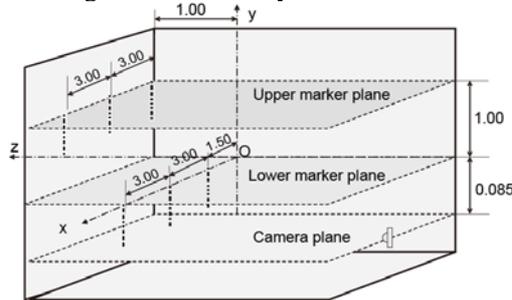

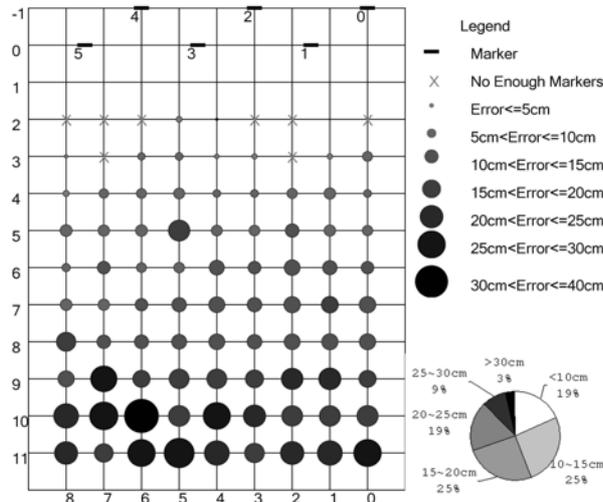**Fig. 2**　Experimental Setup (Unit: m).

Fig. 3　Estimated camera position error.

2.3　Development of Hybrid Tracking Method

The tracking method described in 2.1 and 2.2 is useful at a longer distance than a legacy tracking method like ARToolKit; moreover, it uses very slim markers that can be pasted easily inside a complicated structure. However, to execute tracking, at least two line markers must be captured by a camera (in the case of ARToolKit, one

marker is sufficient for tracking). Therefore, the number of line markers that must be pasted in the working environment is much larger than the case of ARToolKit. One candidate to solve this problem might be the use of a wide-angle camera that can capture wide angle images that include more markers. But the resolution of the images captured with a wide angle camera becomes very low and the distortion of the images also becomes very large. This might decrease the maximum detection distance and the tracking accuracy.

In this study, therefore, a hybrid tracking method using a multi-camera unit and gyro-sensor was developed to enlarge the area in which tracking is available with a limited number of markers.

Tracking with a Multi-camera Unit. In this study, a multi-camera unit with three cameras which view angle is about 40 degrees was developed as shown in Fig. 4. These three cameras are located with about 40 degrees relative angle and can capture three-times-wider angle images than a single camera. Of course, using the multi-camera unit, the users can obtain a three-times-wider angle of view than in the case of a single camera without increasing the number of markers pasted in the environment. However, tracking can not be executed by a tracking algorithm developed only for a single camera when two line markers are captured separately by two cameras. In this study, therefore, a new tracking algorithm for multi-camera units was developed.
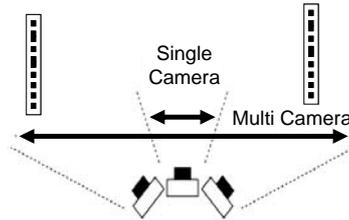

Fig. 4   Extension of Viewing Angle using a Multi-Camera Unit.

In this study, a condition by which all line markers must be mutually parallel was introduced to simplify the tracking algorithm considerably. Many pipes in the plants are parallel. Therefore, the condition will not spoil the applicability of the line marker.

Details of the tracking algorithm for the multi-camera unit are as follows:

Assume that three-dimensional positions of the pasted markers, the relative position and orientation among three cameras and the inertial parameters of the cameras are measured in advance. The transformation equation for mapping a point $\mathbf{x} = (x, y, z, 1)^T$ in the camera 1 coordinate system to a point in the screen coordinate system of camera j is the following.

$$\begin{pmatrix} u \\ v \\ 0 \\ 1 \end{pmatrix} = \mathbf{PT_j x} = \begin{pmatrix} t_{j11} & t_{j12} & t_{j13} & t_{j14} \\ t_{j21} & t_{j22} & t_{j23} & t_{j24} \\ t_{j31} & t_{j32} & t_{j33} & t_{j34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{1}$$

Therein, $\mathbf{T}_j$ (j=1,2,3) is a transformation matrix from the camera 1 coordinate system to the camera j coordinate system, and $\mathbf{P}$ is a projection matrix of the cameras.

1. Capture the line markers by three cameras and recognize them. Select two line markers which the distance is longest. Adopt the tracking algorithm for a single camera if the selected markers are in the same image. Assume that one of the line markers is captured by camera j and that the other line marker is captured by camera k if the selected markers are captured by different cameras.

2. Set the terminals of the line marker captured by camera j as f_1 and f_2; set the terminals of the line marker captured by camera k as f_3 and f_4. where f_1 and f_3 are the same side of the line markers.

3. Find two lines, respectively, through f_1 and f_2, f_3 and f_4. Equations of the lines in each camera coordinate are

$$a_1u + b_1v + c_1 = 0, a_2u + b_2v + c_2 = 0. \tag{2}$$

4. Substituting u and v obtained from eq. (1) into eq. (2) produces

$$
\begin{aligned}
&(a_1t_{j11} + b_1t_{j21})x + &&(a_1t_{k11} + b_1t_{k21})x + \\
&(a_1t_{j12} + b_1t_{j22})y + &&(a_1t_{k12} + b_1t_{k22})y + \\
&(a_1t_{j13} + b_1t_{j23})z = A &&(a_1t_{k13} + b_1t_{k23})z = B
\end{aligned}
\tag{3}
$$

5. Equation (3) is two planes in the camera 1 coordinate; each plane includes a line marker.

6. Set normal vectors of the planes $\mathbf{v_1}$ and $\mathbf{v_2}$; then the direction of the line markers in the camera 1 coordinate is obtainable using eq. (4) because the two line markers are parallel.

$$\mathbf{d} = \frac{\mathbf{v_1} \times \mathbf{v_2}}{|\mathbf{v_1} \times \mathbf{v_2}|} \tag{4}$$

7. The vector from one terminal to the other of the line marker captured by camera j can be expressed as

$$\mathbf{m_j} = \mathbf{R_j}\mathbf{d}l = \left(m_{j1}, m_{j2}, m_{j3}\right)^T, \tag{5}$$

where the length of the line marker captured by camera j is l and the rotation matrix of transformation matrix $\mathbf{T}_j$ is $\mathbf{R}_j$.

8. Set the positions of the terminals of the line marker captured by camera j in the screen coordinate system $\left(u_{f1}, v_{f1}\right)$ and $\left(u_{f2}, v_{f2}\right)$. Terminals of the line marker can be expressed as $\left(\alpha u_{f1}, \alpha v_{f1}, \alpha\right)^T$ and $\left(\beta u_{f2}, \beta v_{f2}, \beta\right)^T$. Then the following equation is obtainable.

$$\begin{pmatrix} \alpha u_{f1} \\ \alpha v_{f1} \\ \alpha \end{pmatrix} + \begin{pmatrix} m_{j1} \\ m_{j2} \\ m_{j3} \end{pmatrix} = \begin{pmatrix} \beta u_{f2} \\ \beta v_{f2} \\ \beta \end{pmatrix} \tag{6}$$

9. Obtain α and β from eq. (6). Substituting them into eq. (7) produces the positions of terminals of the line marker in the screen coordinate of camera 1.

$$\mathbf{f}_1 = \left(\mathbf{PT_j} \right)^{-1} \begin{pmatrix} \alpha u_{f1} \\ \alpha v_{f1} \\ \alpha \end{pmatrix} \quad \mathbf{f}_2 = \left(\mathbf{PT_j} \right)^{-1} \begin{pmatrix} \beta u_{f2} \\ \beta v_{f2} \\ \beta \end{pmatrix} \tag{7}$$

10. Repeating 6, 7 and 8 for camera k produces positions of the terminals of the line marker captured by camera k in the screen coordinate of camera 1.

The four terminals of the two line markers captured by the three cameras in the screen coordinate of camera 1 are obtainable using the algorithm given above. Then the same algorithm can be adopted for the case of a single camera.

Tracking with a gyro-sensor. Using the multi-camera unit described above, the number of the line markers that must be pasted in the environment can be decreased at a certain level. But when a part of the line marker is hidden by a user's body or tools used for maintenance, the tracking will not be able to continue. In this study, therefore, a hybrid tracking method that combines a gyro-sensor and the line marker has been developed. From the gyro-sensor, only the orientation is obtainable at all times, but the output is not so accurate and the cumulative error will increase as time passes. Therefore, the gyro-sensor is useful only for a very short period of time.

The process using the gyro-sensor and the line marker is divisible into three parts.

A case in which more than two markers can be captured by a camera. In this case, the position and the orientation of the camera are calculated based on the two line markers and an output from the gyro-sensor is recorded. Then the amount of the error of the tracking result based on the two line markers is estimated. A weighted average of the tracking result based on the two line markers and the output from the gyro-sensor is calculated if the amount of the error is larger than a threshold value. The result is used as the final tracking result. Otherwise, the tracking result based on the two line markers is used directly as the final tracking result. The accumulated error of the gyro-sensor is corrected if the amount of the error is smaller than a threshold value.

A case in which only one marker can be captured by a camera. In this case, the position of the final result of the previous frame is set as the position of a temporary tracking result. The output from the gyro-sensor is also set as the orientation of the temporal tracking result. Then the position of the line marker on the camera image is estimated using the temporary tracking result. The temporary tracking result is adjusted nonlinearly to minimize the difference between the estimated position and

the actual position of the marker on the image. The adjusted result is used as the final tracking result.

A case in which no marker can be captured by a camera. In this case, the output from the gyro-sensor is used as the orientation of the final tracking result; the position of the final result of the previous frame is used as the position of the final tracking result. Therefore, if the camera moves in parallel, the error of the final tracking result will be the amount of the parallel movement.

2.4　Performance Evaluation of Hybrid Tracking

To evaluate the effectiveness of introducing the hybrid tracking method, an evaluation experiment was conducted in a room of Fugen, which is a real but retired nuclear power plant in Japan. Figure 5 shows that 15 line markers were pasted in a room (8 m × 9.5 m) and the multi-camera unit with a gyro-sensor was moved along the two routes, as shown in Fig. 5. Figure 6 shows a comparison between the case of the hybrid tracking and the case of a single camera. In this case, an assumption is made that the output of the gyro-sensor is reliable for 500 ms after no markers can be captured by the multi-camera unit.

Fig. 5　Marker layout for the evaluation experiment.

Fig. 6　Comparison between hybrid tracking and single camera.

　Figure 6 shows that the areas for which the tracking are available with the hybrid tracking method are 3 times (Route 1) and 17 times (Route 2) larger than when using the single camera. This result confirmed that the hybrid tracking method can enlarge the available area without increasing the number of markers.

3 Tracking Method Using Circular Markers

3.1 Basic Idea of Circular Marker Design

The edges of the square on the image become jagged, as shown in Fig. 7, if a square is captured by a low resolution camera from a long distance. This jagged edge affects the accuracy of the tracking strongly because the tracking calculation for square markers like ARToolKit is based on the intersections of the four lines that the jagged edges form. In contrast, even if a circle is captured from a long distance, a center of the circle can be recognized accurately, as shown in Fig. 8. Therefore if a marker consists only of circles and the tracking calculation is based on the center of the circles, the accuracy of the tracking will not decrease even if the marker is captured from a long distance.
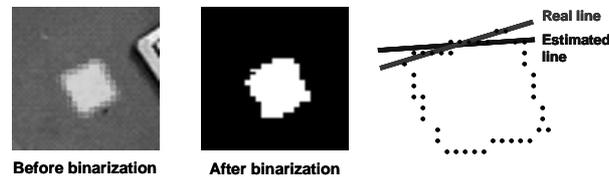
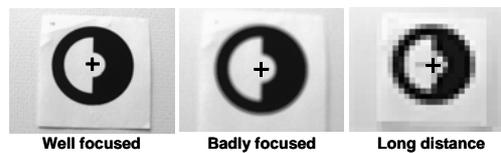**Fig. 7**   A square captured from a long distance and estimated lines.

Fig. 8   Recognition of circle centers in severe conditions.

Another idea to design a new marker is the relationship between a number of markers captured by a camera and the distance between the camera and the markers. Figure 9 shows that the captured markers become very numerous and the size of the marker on the image becomes very small when the distance between the camera and the markers is long. It therefore becomes difficult to obtain plural feature points from one marker because the marker size on the image is too small, but plural feature points are obtainable from plural markers. Moreover, the distances among the feature points can be long because markers can be pasted so as not to be so near with each other. On the other hand, when the distance between the camera and the markers is short, the number of the captured markers becomes small and the size of the markers on the image becomes very large. In this case, it becomes possible to obtain plural feature points from one marker because the marker size on the image is sufficiently large.

Based on the idea described above, the authors have designed circular markers as shown in Fig. 10. This marker consists of one black outer circle (with thickness of 30% of the marker radius), one white center circle (the radius is 30% of the marker radius), and a middle circle (with thickness of 40% of the marker radius), which

consists of 10 black or white fans that represent a binary code and four small circles (the radius is 24% of the marker radius) at the corner of the marker.
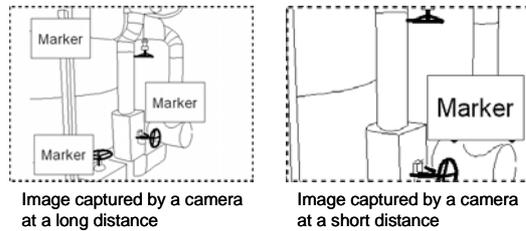

Image captured by a camera at a long distance

Image captured by a camera at a short distance

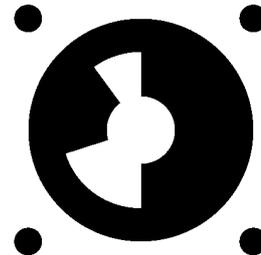**Fig. 9.** Relationship between a number of captured markers and the camera distance.

Fig. 10. Design of a circular marker.

The black outer circle and the white center circle are used for determining the threshold to analyze the binary code of the middle circle. The markers that can be used simultaneously are 99. The four small circles at the corners can be recognized when the marker is captured at a short distance and the centers of these small circles are used as four feature points to execute the tracking using a P4P solution. Thereby, even if only one or two markers are captured by a camera because the distance between the camera and the markers is too short, the tracking can be continued if the marker size on the image is sufficiently large. Therefore the marker shown in Fig. 10 is useful in two-layer modes: long distance mode with plural markers, and short distance mode with single marker.

3.2 Performance Evaluation of Circular Marker

To estimate the maximum detection distance of the circle marker, one circular marker was pasted on a wall and the maximum distance was measured under the condition that the tracking system can correctly recognize the code of the circular marker. The circular marker diameter was changed from 3 cm to 10 cm by 1-cm steps. For comparison, one square marker (ARToolKit) was pasted on a wall and its size was changed from 3 cm to 10 by 1-cm steps. 128 markers with binarized patterns were registered to the ARToolKit system in advance. The result shown in Fig. 11 verifies that the maximum detection distance of the circular marker is about twice that of the square marker.
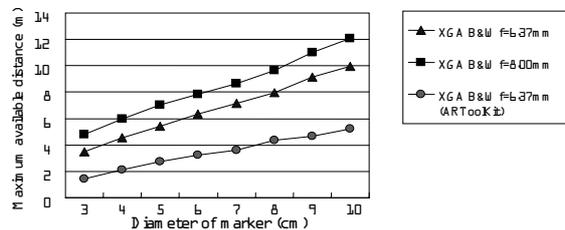

Fig. 11. Comparison of maximum detection distance (a circular and a square marker).

To confirm that the circular marker is useful for two-layer tracking, a simple experiment was conducted. In the experiment, five circular markers of 8 cm diameter are pasted on a wall and the camera (Dragonfly II XGA black-and-white) was moved from a short distance (about 30 cm) to a long distance (about 8 m). This experiment verified that the tracking can be executed properly at all distances. In addition, the mode change between the tracking using single marker and the tracking using plural markers were unnoticeable to the extent that the calibration of the camera lens distortion was done properly.

Evaluation of the tracking accuracy using circular markers has not been done yet, but the tracking accuracy depends on the layout of the feature points obtained from the markers. Therefore, when the circular markers are pasted in a condition that the centers of the circular markers are on the terminals of the line markers, the tracking accuracy will be nearly equivalent to that of the case of the line markers. Concrete evaluation of the accuracy using the circular markers will be conducted as future work.

4 Summary and Future Works

In this study, a hybrid tracking method and a two-layer tracking method were developed to enlarge the area in which tracking is available using only a limited number of markers. These tracking methods were developed separately and can not be used simultaneously. Future work will integrate these methods into a single tracking method to cover short distances and long distances seamlessly.

References

1. N. Navab: Developing Killer Apps for Industrial Augmented Reality: Computer Graphics and Applications, IEEE, Vol.24, No.3, pp.16-22, 2004.
2. H. Ishii, H. Fujino, Z. Bian, T. Sekiyama, H. Shimoda, H. Yoshikawa: Development of Marker-based Tracking Methods for Augmented Reality Applied to NPP Maintenance Work Support and its Experimental Evaluation, 5th Intl. Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies, pp.973-980, 2006.
3. H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto, K. Tachibana: Virtual Object Manipulation on a Table-Top AR Environment, International Symposium on Augmented Reality, pp.111-119, 2000.