An Initialization Tool for Installing Visual Markers in Wearable Augmented Reality

Yusuke Nakazato, Masayuki Kanbara, and Naokazu Yokoya

Nara Institute of Science and Technology,
8916–5 Takayama, Ikoma, Nara, 630–0192, Japan

Abstract. It is necessary to precisely measure pose (position and orientation) of a user in order to realize an augmented reality (AR) system with a wearable computer. One of major methods for measuring user's pose in AR is visual marker-based approach which calculates them by recognizing markers pasted up on the ceilings or walls. The method needs 3D pose information of visual markers in advance. However, much cost is necessary to calibrate visual markers pasted up on the ceiling in a wide environment. In this paper, an initialization tool for installing visual markers in wearable AR is proposed. The administrator is assisted in installing visual markers in a wide environment by the proposed tool. The tool calibrates alignment of visual markers which exist in the real environment with high accuracy by recognizing them in the images captured by a high-resolution still camera. Additionally, the tool assists the administrator in repairing the incorrect pattern of marker using a wearable AR system.

1 Introduction

Recently augmented reality (AR) system using a wearable computer has received a great deal of attention as a method for displaying location-based information in the real world[1]. To realize a wearable AR system, the exact pose (position and orientation) of the user is required. Especially in indoor environments, since a GPS cannot be used, various localization methods have been investigated[2,3,4,5]. User's pose is sometimes estimated by recognizing visual markers pasted up on the ceilings or walls whose poses are known[6,7,8]. We have proposed a localization method with an infrared camera and invisible visual markers consisting of translucent retro-reflectors as shown in Fig.1[9]. The method can realize indoor localization without impairing the scenery in a wide environment. In the visual marker-based approach such as in [9], the alignment information of visual markers, which is 3D poses of markers located in real environment, is needed in advance.

In this paper, we focus on initialization methods of visual markers for wearable AR system. The initialization of visual markers means calibration of their alignment in advance. The simplest method is calibration by hand using a ruler and others, although the accuracy of calibration results is low. Another method
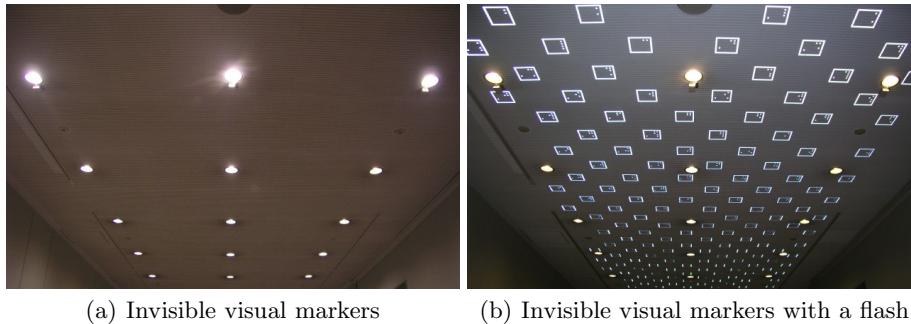
(a) Invisible visual markers   (b) Invisible visual markers with a flash

Fig. 1. Positioning infrastructures with invisible visual markers placed on the ceiling. Though it is difficult to perceive the markers, these can be captured clearly with a flash.

using a laser rangefinder can calibrate the poses of visual markers with high accuracy. However the laser rangefinder is an expensive instrumentation in general. Furthermore, these methods need much labor to calibrate the poses of visual markers in a wide environment. On the other hand, the calibration methods using images of visual markers have been proposed[10,11,12]. Baratoff, et al have proposed a semiautomatic image-based calibration method requiring only minimal interaction[10]. Maeda, et al have calibrated all visual markers by capturing markers which include one reference marker automatically[11]. These methods can work online, however the calibration results of markers' poses do not have high accuracy, because the pose of marker is usually estimated based on propagation with some low resolution images. When a large number of visual markers are installed as positioning infrastructures for user's localization, the calibration of visual markers with high accuracy is generally required. In this case, the offline calibration tool for wearable AR system is useful.

We focus on high accuracy calibration of visual markers' poses used as positioning infrastructures. This paper proposes an initialization tool which assists on administrator in installing visual markers. First, the administrator sets up visual markers on the ceilings or walls and then captures images of visual markers by a high-resolution still camera. Secondly, the proposed tool generates alignment information of visual markers by recognizing visual markers on the images captured by the administrator. Finally, the tool assists the administrator in repairing duplicate patterns of visual markers using a wearable AR system. Because such patterns of visual markers placed in the neighborhood sometimes cause failure of user localization. However, it is difficult to confirm such incorrect pattern from a large number of visual markers by human beings. In particular, when the invisible visual markers as shown in Fig.1 are used, the task is more difficult because the markers are invisible. Therefore, the tool provides the administrator with images containing indication of repairing markers using a wearable AR system.

This paper is structured as follows. Section 2 describes an initialization tool for visual marker-based localization. In Section 3, experiments of infrastructure
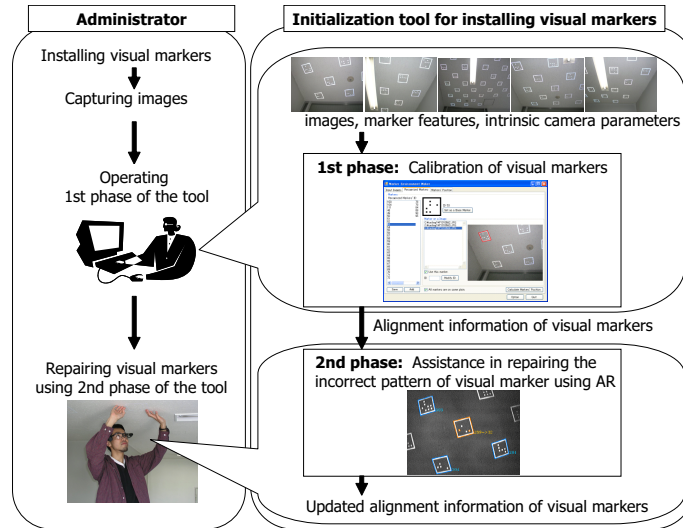
Fig. 2. Flow diagram of initialization tool of visual markers

initialization with the proposed tool are described. Finally, Section 4 gives conclusion and future work.

2   Initialization Tool for Installing Visual Markers

In this section, the proposed initialization tool is explained in detail. The overview of the proposed tool is illustrated in Fig.2. First, visual markers are installed as positioning infrastructures and then high-resolution images of visual markers are captured by the administrator. Secondly, the tool calibrates poses of visual markers by using captured images and outputs calibration results as alignment information. Finally, using the calibration results, the tool provides the administrator with augmented images in order to repair incorrect patterns of visual markers.

This tool consists of two phases: the first phase is concerned with calibration of visual markers and the second phase is to assist in repairing incorrect markers using a wearable AR system. In the first phase, poses of visual markers are calibrated from high-resolution images of markers. The administrator inputs (1) intrinsic camera parameters (focal length, pixel size, center of image, radial distortion factor coefficient), (2) features of visual marker (frame length and width, pattern size), and (3) images of visual markers. The tool estimates poses of visual markers, and outputs this information as alignment information of visual markers. In the second phase, the tool assists the administrator in repairing incorrect patterns of markers using alignment information of visual markers generated in the first phase. The tool provides the administrator with overlaid images containing indication of repairing visual marker and then outputs updated alignment

information of visual markers. Sections 2.1 and 2.2 explain about the first and second phases, respectively.

2.1 First Phase: Calibration of Visual Markers

In this section, the first phase of initialization tool is described.

Overview of First Phase of Initialization Tool. In the first phase, the alignment information of visual markers is generated. This phase uses three forms as shown in Figs. 3,4, and 5: an input, a confirmation, and an output forms. Fig.3 illustrates the input form. The administrator inputs intrinsic camera parameters, visual marker features, and images of visual markers. Then the administrator
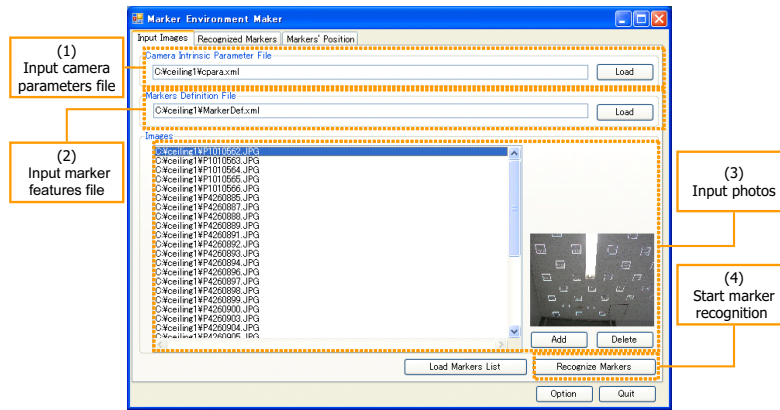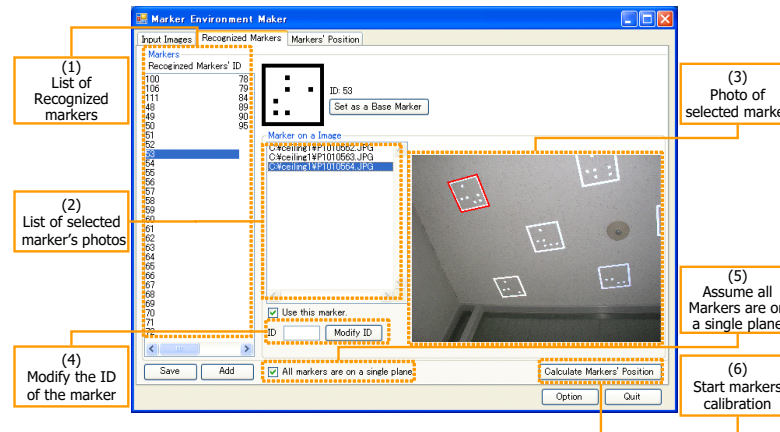

Fig. 3. Input form in the first phase

Fig. 4. Confirmation form in the first phase
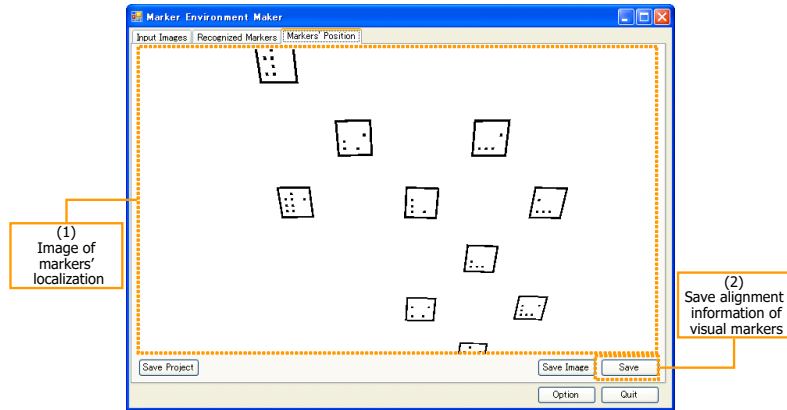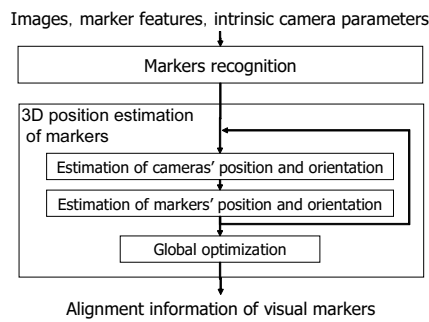
Fig. 5. Output form in the first phase

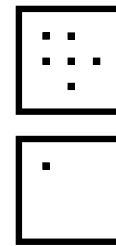**Fig. 6.** Flow diagram of visual marker calibration

Fig. 7. Instances of markers

dictates to start recognition of visual markers, and switches to the confirmation form. In the confirmation form shown in Fig.4, the administrator confirms result of marker recognition. If misrecognized or duplicate visual markers exist, the administrator modifies this field(Fig.4. (4)). Furthermore, if it is supposed that all visual markers are on a single plane, the administrator sets a flag on this field(Fig.4. (5)). Then the administrator dictates to start calibration of visual markers, and switches to the output form. Fig.5 illustrates the output form. The form shows an image on which the calibrated visual markers are drawn and then the tool outputs the alignment information of visual markers. This information is used in the second phase.

Visual Markers Calibration. Fig.6 shows the flow diagram of the process of visual markers calibration in the first phase. First, visual markers in each image in the input form are recognized. Secondly, one recognized visual marker is set the 3D pose as a reference marker. Poses of cameras capturing the reference

visual marker are estimated. Then, position and orientation of visual markers are estimated by estimated poses of the cameras. Finally, the poses of the visual markers are optimized all over the images. The details of these processes are described below.

Marker Recognition. We employ invisible visual markers[9] which consist of grid points showing their IDs as shown in Fig.7. In the process of marker recognition, edges are detected from the input image using canny operator [13], then contours of marker are retrieved. The area of original image corresponding to the contour is binarized using a threshold which is determined by the discriminant analysis method[14]. Four vertices of a square as a visual marker are detected from the area, and its ID is recognized. The four vertices and centroids of grid points in the marker are used as feature points in the process of 3D pose estimation of markers described below.

3D Pose Estimation of Markers. First, one recognized visual marker is given its 3D pose as a reference marker; vote that the reference marker determines the origin of the world coordinate system. Next, using a Sato's method[15], pose of camera capturing the reference marker are estimated. Then, poses of other markers captured by camera known 3D position and orientation are estimated. Using the estimated poses of markers, unestimated poses of cameras are estimated. These processes of pose estimation of camera and visual markers are iterated until all the markers' poses are determined. In the iteration process, visual markers captured by multiple images, which has a large reprojection error, are considered as those with duplicate patterns. The reprojection error is defined as the sum of squared distances between detected positions of marker's feature points in the image and projected positions of marker's feature points in the world coordinate system. The poses of duplicate visual markers are estimated individually. These visual markers with duplicate patterns are candidates for repairing in the second phase. Finally, 3D poses of cameras and visual markers are globally optimized so as to minimize the sum of reprojection errors all over the images.

2.2   Second Phase: Assistance in Repairing Incorrect Markers Using AR

In the second phase, our tool assists the administrator in repairing incorrect patterns of markers using a wearable AR system. The administrator equips a wearable computer, an HMD and an IR camera with IR-LEDs in order to capture the invisible visual markers as shown in Fig.8. The tool uses alignment information of visual markers generated in the first phase, and provides the administrator with images containing indication of repairing markers as shown in Fig.8. The administrator's view images are overlaid with IDs and contours of recognized markers. If incorrect visual markers such as duplicate or lacked patterns of markers exist in the administrator's view, the tool provides the administrator images with the repaired pattern of marker. In case of Fig.8, the administrator has modified ID No. 289 to No. 32. When the pattern of marker is repaired by the administrator, the tool updates alignment information of visual markers.
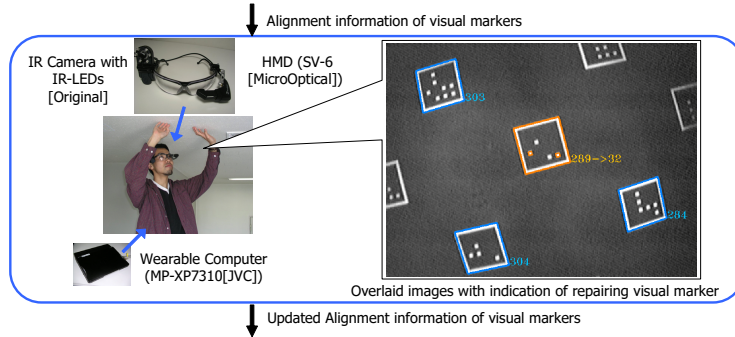
Fig. 8. Overview of the second phase

3   Experiments

First, we evaluated the accuracy of estimated position of visual markers in the
first phase. In this experiment, invisible visual markers were set up on the ceil-
ing. The size of the marker was 16cm. These markers are arranged at intervals
of 60cm on the ceiling plane. The 20 images captured by a digital still camera
"OLYMPUS C-5050ZOOM" were used as shown in Fig.9. The size of these im-
ages was 2560×1920 pixels. The intrinsic camera parameters were estimated
by Tsai's method [16] in advance. Since the visual markers were set up along
pattern of ceilings, the grand truths of the visual marker poses were calibrated
manually.   When it is assumed that all markers are not located on the single
plane (no co-planar constraint), the projected points to 2D plane of estimated
positions of visual markers' four vertices are shown in Fig.10(a). Fig.10(b) shows

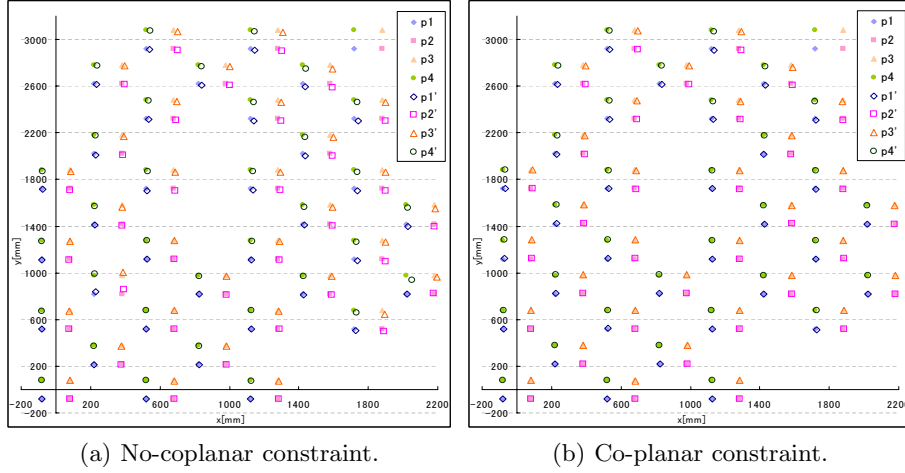**Fig. 9.** Examples of input images (2560×1920 pixels)

(a) No-coplanar constraint.   (b) Co-planar constraint.

Fig. 10. Results of marker calibration (p*n*:true position, p*n*':estimated position)

Table 1. Average and standard deviation of marker calibration errors

	no co-planar constraint			co-planar constraint		
	x	y	z	x	y	z
average of absolute error[mm]	9.9	11.6	16.9	6.0	3.7	0.0
standard deviation of error[mm]	7.7	9.7	13.6	5.2	4.6	0.0

estimated markers' positions when it is assumed that all markers are located on the single plane (co-planar constraint). In these figures, p1, p2, p3 and p4 show true positions of visual marker's four vertices, and p1', p2', p3' and p4' show corresponding estimated positions. The average of absolute errors of markers' positions and standard deviation of errors are shown in Table 1. In this experiment, we used a note PC (Pentium M 1.6GHz, Memory 512MB). When the no co-planar constraint is assumed, it took about 173 seconds for calibration of the markers, and reprojection error was about 1.5 pixels. When the co-planar constraint is assumed, the processing time of marker calibration was about 75 seconds, and reprojection error was about 1.7 pixels. From Table 1, we confirmed that the average error of estimated position was about 6mm and 17mm in case of co-planar constraint and no co-planar constraint, respectively. In user localization, when we use a camera whose angle is 90 degrees and resolution is 640 × 480 pixels and then the user captures the visual markers on the ceiling 1.2m away, the marker's position error of 6mm corresponds to the error of about 1.6 pixels on the image. We can conclude that the accuracy obtained is sufficient for user's localization. Most visual markers were captured in about 3 images; however some markers were captured only one time. These markers caused some errors.
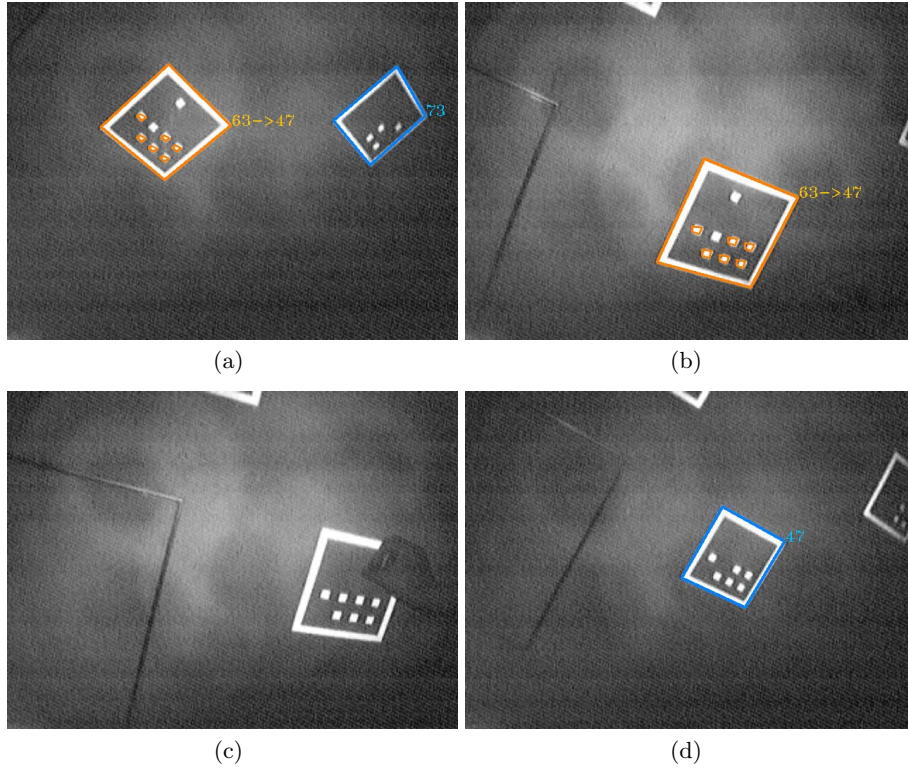
Fig. 11. Augmented images. These images are overlaid with IDs and repaired pattern of marker. The administrator modified the duplicate pattern of No.63 to No.47 according to the indications given in these images.

In the experiment of the second phase, we confirmed that proposed tool was able to assist in repairing the patterns of markers. Fig.11 shows augmented images with which the administrator is provided repairing the marker. From the result of the first phase, it was understood that No.63 is duplicated. Thus the administrator modified the pattern to No.47 according to the indications given in augmented images.

4   Conclusions

This paper has proposed an initialization tool for installing visual markers in wearable AR. In order to construct an environment for visual marker-based localization, the tool calibrates pose of visual markers by using images of markers captured by high-resolution still camera. Furthermore, the tool assists the administrator in repairing incorrect patterns of markers by presenting augmented images. In experiments, we have confirmed that the proposed tool is able to estimate poses of markers with the high accuracy of about 6 mm when the co-planar

constraint is assumed. We have also confirmed that the tool is able to assist in repairing incorrect patterns of markers using a wearable AR system. In future work, we should carry out experiments of infrastructure initialization of visual markers in a wide area. In order to calibrate poses of visual markers which have some errors with more high accuracy, we will indicate the additional positions at which the images should be taken using wearable AR system.

Acknowledgments

References

1. Höller, T., Feiner, S., Terauchi, T., Rashid, G., Hallaway, D.: Exploring mars: Developing indoor and outdoor user interfaces to a mobile augmented reality system. Computers and Graphics **23**(6) (1999) 779–785
2. Kourogi, M., Kurata, T.: Personal positioning based on walking locomotion analysis with self-contained sensors and wearable camera. Proc. 2nd IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 03) (2003) 103–112
3. Tenmoku, R., Kanbara, M., Yokoya, N.: A wearable augmented reality system using positioning infrastructures and a pedometer. Proc. 7th IEEE Int. Symp. on Wearable Computers (ISWC'03) (2003) 110–117
4. Vacchetti, L., Lepetit, V., Fua, P.: Combining edge and texture information for real-time accurate 3D camera tracking. Proc. 3rd IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 04) (2004) 48–57
5. Oe, M., Sato, T., Yokoya, N.: Estimating camera position and posture by using feature landmark database. Proc. 14th Scandinavian Conf. on Image Analysis (SCIA2005) (2005) 171–181
6. Thomas, B., Close, B., Donoghue, J., Squires, J., De Bondi, P., Piekarski, W.: First person indoor/outdoor augmented reality application: Arquake. Personal and Ubiquitous Computing **6**(1) (2002) 75–86
7. Kalkusch, M., Lidy, T., Lnapp, M., Reitmayr, G., Kaufmann, H., Schmalstieg, D.: Structured visual markers for indoor pathfinding. Proc. 1st IEEE Int. Augmented Reality Toolkit Workshop (ART'02) (2002)
8. Naimark, L., Foxlin, E.: Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker. Proc. 1st IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR2002) (2002) 27–36
9. Nakazato, Y., Kanbara, M., Yokoya, N.: A localization system using invisible retro-reflective markers. Proc. IAPR Conf. on Machine Vision Applications (MVA2005) (2005) 140–143
10. Baratoff, G., Neubeck, A., Regenbrecht, H.: Interactive multi-marker calibration for augmented reality applications. Proc. 1st IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR2002) (2002) 107–116
11. Maeda, M., Habara, T., Machida, T., Ogawa, T., Kiyokawa, K., Takemura, H.: Indoor localization methods for wearable mixed reality. Proc. 2nd CREST Workshop on Advanced Computing and Communicating Techniques for Wearable Information Playing (2003) 62–65

12. Zauner, J., M. Haller, A.: Authoring of mixed reality applications including multi-marker calibration for mobile devices. Proc. 10th Eurographics Symp. Virtual Environments (EGVE2004) (2004) 87–90
13. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Analysis and Machine Intelligence **8**(6) (1986) 679–698
14. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Systems, Man, and Cybernetics **SMC-9**(1) (1979) 63–66
15. Sato, T., Kanbara, M., Yokoya, N.: 3-D modeling of an outdoor scene from multiple image sequences by estimating camera motion parameters. Proc. 13th Scandinavian Conf. on Image Analysis(SCIA2003) (2003) 717–724
16. Tsai, R.Y.: An efficient and accurate camera calibration technique for 3D machine vision. Proc. IEEE Conf. on Computer Vision and Pattern Recognition (1986) 364–374