Precise Geometric Registration by Blur Estimation for Vision-based Augmented Reality

Bunyo OKUMURA * Masayuki KANBARA * Naokazu YOKOYA *

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

ABSTRACT

This paper proposes an accurate geometric registration method by estimating blur effects from a degraded image with image markers for augmented reality. A small and inexpensive camera used in augmented reality systems sometimes captures degraded images because its focus and/or iris are fixed. This degradation of a captured image affects the accuracy of the detected positions of feature points in the image. The proposed method improves the accuracy of the estimated camera position and posture by estimating blur effects from the captured image, and by correcting the detected positions of feature points through the results. The effectiveness of the method is confirmed through experiments of corner estimation from simulated images and extrinsic camera parameter estimation from real images.

Keywords: augmented reality, geometric registration, blur estimation, point spread function

Index Terms: H.5.1 [Information interfaces and presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

1 INTRODUCTION

The geometric registration problem is important for augmented reality for overlaying virtual objects on a real scene, without seams between the real and virtual worlds. In video see-through augmented reality systems, the geometric registration problem has been solved by estimating the position and posture of the camera used to capture the real scene. In particular, square shaped [1, 2, 3] or circular [4] image markers have been used in order to estimate the position and posture of the camera.

A number of studies have been conducted recently on augmented reality with a small and inexpensive camera [2, 5, 6]. It is difficult for these cameras to capture a fine image under various situations because their focus and/or iris are fixed. Degradations are observed in captured images due to defocus and motion blur, which harm the detection accuracy of geometric registration as the image markers cannot be detected correctly. It should be noted that defocus and motion blur effects can be reduced by improving the depth-of-field of camera, and by using a short exposure time. However, it is still difficult to guarantee the complete elimination of these effects.

In the computer vision field, there are two typical approaches for improving the accuracy and robustness of estimating the position and posture of the camera. One method combines a camera with other sensors. In [5], a robust estimation of position and posture of a camera was achieved using a gyro sensor. In this method, the position and posture are estimated by fitting a wire frame model with edges in a captured image. The area for fitting edges is limited by the size of motion blur, which is estimated using the gyro sensor for eliminating the mismatch of edges and thus improving the robustness of estimation. In [6] too, a robust estimation of position and posture was achieved by combining several well-known methods: (1) absolute position and posture estimation based on model-based tracking; (2) relative position and posture estimation by tracking feature points; and (3) posture estimation by using a gyro sensor. The proposed method selects the estimation method depending on the success of each.

The other approach is an improvement method without other sensors, which has been cited in [7] for the accuracy of the estimated position of a camera. This method estimates corners from intersections of edges, which are determined by fitting a blurred edge function to pixels around edges. In addition, some studies have evaluated the marker detection techniques for their accuracy and robustness. In [8], some registration systems dealing with the accuracy of the estimated position of a camera and the robustness of marker detection have been evaluated. Another method [9] has evaluated the accuracy of registration systems using an industrial robot arm.

This paper proposes a method for improving the accuracy in geometric registration by estimating blur effects from a captured image. By treating both defocus blur and motion blur, our method detects markers from a captured image and estimates blur effects from the captured marker image. Finally, the position and posture of a camera are accurately estimated by considering blur effects on feature points of markers.

This paper is structured as follows. Section 2 describes a blur model of an image, and a method for estimating the camera position and posture from image markers. Section 3 describes the experimental results, and finally, Section 4 presents the conclusion and future work.

2 IMPROVEMENT OF POSITION AND POSTURE ESTIMATION BASED ON BLUR ESTIMATION

The proposed method improves the accuracy of the estimated position and posture of a camera with defocus and motion blur. Feature points are detected by matching a captured image with a template that shows a corner of a marker with blur effects estimated from a circular edge in the marker.

Figure 1 shows a flow diagram of the proposed method. First, the real scene is captured using a camera, and image markers of known color and shape are detected (Figure 1 (A)). In the second step, the size of blur is estimated by fitting a function defined by a model of the point spread function (PSF) to intensities in an edge region of the marker, and the parameters of PSF are acquired by integrating the estimated size of blur in various directions (Figure 1 (B)). Next, a template image is generated using estimated PSF parameters and the marker shape (Figure 1 (C)). Then, a corner position of the marker is estimated by minimizing the sum of squared distances (SSD) between the pixel values of the template and captured images (Figure 1 (D)). Processes (C) and (D) are repeated until the SSD is sufficiently small. Finally, the camera position and

*E-mail: {bunyo-o, kanbara, yokoya}@is.naist.jp

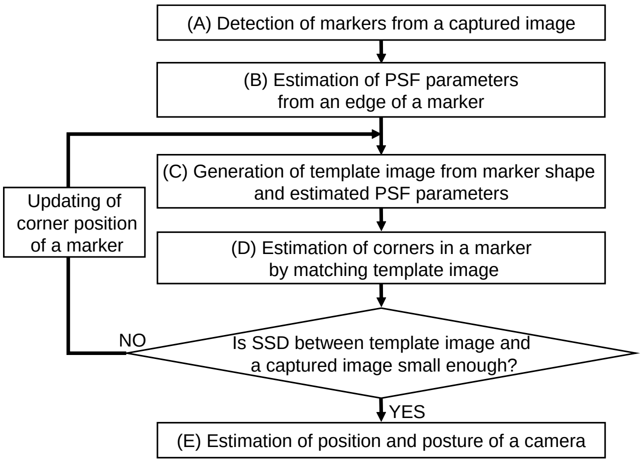


Figure 1: Flow diagram of the proposed method.

posture in the marker coordinate system are estimated by minimizing a re-projection error of detected feature points (Figure 1 (E)). In this paper, we mainly describe the blur model of a camera and the details of processes (C) and (D), because processes (A), (B), and (E) are based on methods described in [7].

2.1 Blur model in our approach

In general, the blur effect caused by the camera includes two major factors, namely, defocusing and motion blurring. The blurred image is generated by convolution of an ideal image with the PSF. In this study, we use the PSF that expresses defocus and motion blurs with three parameters, proposed in [7].

In the proposed method, the following approximated PSF is employed.

$$P(x,y;r,l,\theta) = \begin{cases} \frac{1}{\pi((r+l)^2+r^2)} & ; \quad \left(\frac{x'}{r+l}\right)^2 + \left(\frac{y'}{r}\right)^2 \leq 1 \\ 0 & ; \quad \text{otherwise} \end{cases}, \quad (1)$$

where

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (2)$$

Here, x and y represent the position of a target pixel in the image. The radius of defocusing blur, length of uniform motion, and the direction of motion are denoted by r, l, and θ, respectively, and are considered as parameters of the PSF of blur. The above equations refer to an elliptical spread of light on an image plane. The direction of the major axis of the ellipse shows the motion direction of the camera. From this approximation of PSF, the profile of a blurred edge can be described by a formula, thus reducing the calculation cost of estimating the PSF parameters.

2.2 Generating a template image from marker shape and estimated PSF parameters: Process (C)

A template image is generated from the position of corners of a marker that is estimated in process (A), and from PSF parameters that are estimated in process (B). Eq. (3) is used to generate a template image $R(x,y;r,l,\theta)$ with PSF parameters.

$$R(x,y;r,l,\theta) = \sum_{s=-w}^{w}\sum_{t=-w}^{w} M(x+s,y+t)\cdot PSF(s,t;r,l,\theta), \quad (3)$$

where

$$M(x,y) = \begin{cases} \hat{i}_{black} & ; \quad a_jx+b_jy+c_j \geq 0 \quad (j=0..3) \\ \hat{i}_{white} & ; \quad \text{otherwise} \end{cases}. \quad (4)$$

Here, x and y show coordinates of a pixel in the image. The size of window for convoluting the PSF with a marker is denoted by w, and w equals $r+l$. $\hat{i}_{white}$ and $\hat{i}_{black}$ show the intensities of white and black parts of the marker, respectively. Note that parameters a_j, b_j, and c_j satisfy $a_jx+b_jy+c_j \geq 0$ $(j=0..3)$, when a point (x,y) is inside the marker region. These parameters are calculated from the estimated corner position (x_i,y_i) in process (A). The indexes of corners and edges are denoted by i and j, respectively, and are counted on the top left of the marker in a clockwise manner.

2.3 Estimation of corners of marker by matching template image: Process (D)

Corner positions of the marker are estimated by matching a captured image with a template generated from the marker shape and estimated PSF parameters. We apply the sum of squared differences between the input image and the template image as a measurement. The position $(\tilde{x}_i,\tilde{y}_i)$ of a corner is estimated by minimizing the error function defined by Eq. (5). Here, the marker image with blur effects is generated only around a corner of the marker, because only the positions of the corners are needed for the estimation of the camera position and posture. Our method uses x_i and y_i as initial values of $\tilde{x}_i$ and $\tilde{y}_i$.

$$E_{SSD,i}(\tilde{x}_i,\tilde{y}_i) = \sum_{s=-W}^{+W}\sum_{t=-W}^{+W}\{I(x_i+s,y_i+t)-R(\tilde{x}_i,\tilde{y}_i)\}^2. \quad (5)$$

If the error value $E_{SSD,i}$ is larger than the threshold th_{SSD}, our method updates the shape of the marker by refining parameters a_j, b_j, and c_j from the estimated $\tilde{x}_i$ and $\tilde{y}_i$.

3 Experiments

To confirm the effectiveness of the proposed method, we carried out two quantitative evaluation experiments. In the first experiment, we evaluated the detection accuracy for the positions of corners in a marker by simulation. In the second experiment, we evaluated the accuracy of the estimated depth from the marker to the camera from a real scene with a robot arm. In these experiments, we compared the proposed method with the following two techniques:

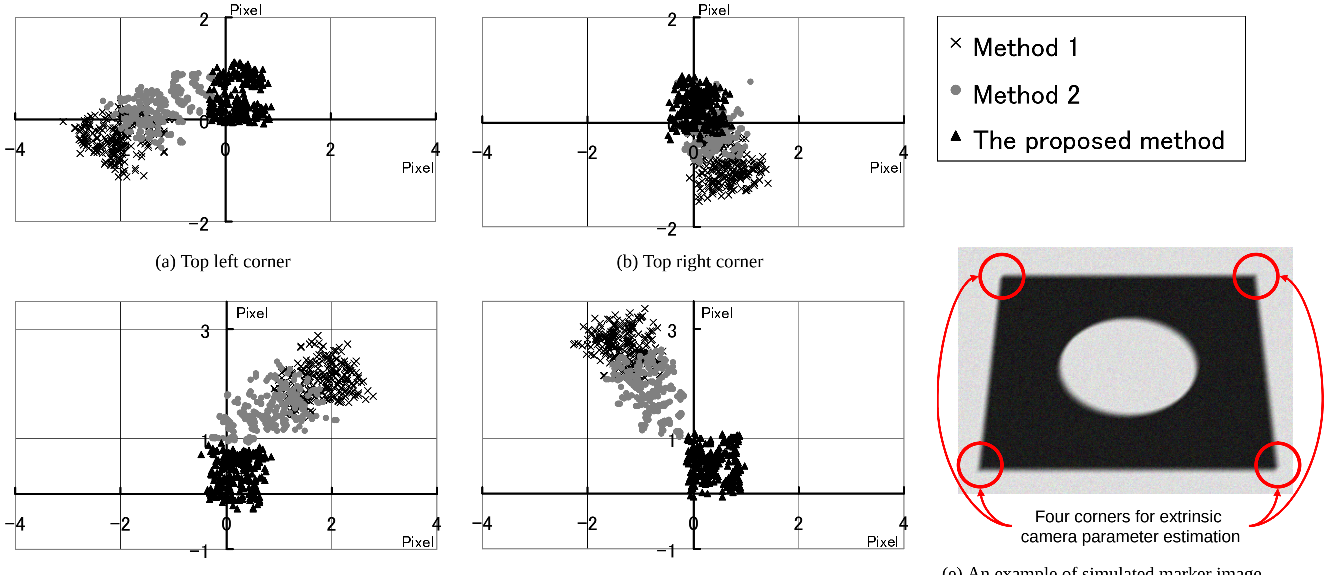
Method 1: Marker-based registration system called ARToolkit [1]. This method detects corners of a marker by binarizing an image with fixed threshold, and labeling it. It should be noted that the estimated positions of corners are equivalent to the positions estimated in process (A).

Method 2: Marker-based registration system by considering blur effects [7]. This method estimates edges of a marker before fitting a straight line to an edge. Then, marker corners are estimated from intersections of straight lines.

3.1 Evaluation of estimated positions of corners by simulation

In this experiment, we evaluated the positions of corners estimated using three methods by simulation: method 1, method 2, and the proposed method.

A marker for evaluation is placed 0.4 m away from a camera. We simulated both defocus and motion blur. For defocus, the focus of the camera is changed to three values: 0.15 m, 0.2 m, and 0.4 m. Defocus effects appear in those situations, and the sizes of blur are



(a) Top left corner (b) Top right corner (c) Bottom left corner (d) Bottom right corner (e) An example of simulated marker image

Figure 2: Errors in positions of estimated corners and an example of simulated image

Table 1: Average and standard deviation of errors in position of estimated corners

Estimation methods	Corner 0		Corner 1		Corner 2		Corner 3	
	Ave.	Std.	Ave.	Std.	Ave.	Std.	Ave.	Std.
Method 1	2.08	0.45	1.03	0.38	2.97	0.35	2.70	0.41
Method 2	1.43	2.34	0.71	2.48	2.22	3.47	1.91	2.87
The proposed method	0.59	0.31	0.45	0.20	0.72	0.28	0.56	0.25

(pixel)

4, 2.5, and 0.5 pixels, respectively. For motion blur, the sizes are changed to 0, 2, and 4 pixels, and the directions are also changed to 0, 45, and 90 degrees. Furthermore, we added Gaussian noise in order to simulate noise effects of the camera. The standard deviation of the noise is 3 in 256 gradation of intensity.

Figure 2 shows errors in estimated positions of corners by previously described methods and an example of simulated image of a marker. Table 1 shows the average distance and the standard deviation of the errors. The maximum value of average errors of four corners estimated by method 1 is larger compared with the others. The blur effects make corners of the marker round, and thus the marker is not extracted correctly, and the estimated positions of the corner are not precise. In method 2, the maximum error of the estimated corners is reduced by eliminating blur effects. However, the maximum standard deviation of errors is 3.47 pixels with method 2, and this is larger than that by method 1, owing to the errors of function fitting on blurred edges. In contrast, the average and standard deviation of the proposed method are smaller than the others. The maximum value of average errors of the positions estimated by the method is 0.72 pixels and the maximum value of standard deviation of errors is 0.31 pixels. This implies that the proposed method can detect a marker's corner accurately and robustly even if the blur effects exist. Method 1 works at 30 frames per second, method 2 at 15 frames per second, and the proposed method at 5 frames per second with a desktop computer (CPU: Pentium D 3.0GHz).

3.2 Evaluation of the estimated depth

In this experiment, we evaluated the estimated depth from the camera to the marker by three methods using blurred images. In a preliminary experiment, it was clear that the estimated position of a camera is affected in the direction of an optical axis of the camera through blur effects, because blur effects mainly affect the region size of the marker. For this reason, we evaluated the depth from the center of the marker to the camera.

Figure 3 shows the experimental environment with a robot arm. A USB camera (ARGO Lu-135c, Resolution: 1024×768 pixels) is placed on a robot arm. Markers are placed 0.45 m away from the camera. The true depth from the marker to the camera is measured using a total station (TOPCON GPT-9005A). We used only one marker to estimate the position and posture of the camera in this experiment, because the blur effect critically affects the accuracy of the estimated position and posture of the camera, when only one marker is used.

In order to generate blur effects, the camera is focused to 0.5 m, and moved by a robot arm at 300 mm per second. The size of the defocus blur on the image is about 2 pixels. In addition, we changed the exposure time and gain of the camera to change the size of the motion blur. The exposure time is changed to 15, 30, and 45 ms. The sizes of motion blur in the image are about 0.9, 1.8, and 2.8 pixels, respectively. Figure 4 shows the captured images in various exposure time.

Figure 5 illustrates the errors of the estimated depth of the camera. Each sequence shows variations in the exposure time and estimation method. Table 2 shows average and standard deviation of errors, which show the difference between the estimated and the true depth. In the results of method 1, the estimated depth is larger than the true depth. In addition, the results are changed when the size of motion blur is changed, thus proving that method 1 is not robust with regard to blur effects. In results obtained using method 2, the estimated depths are slightly better in comparison to results by method 1. However, the results are affected by the size of motion blur. On the other hand, we can confirm that the results by the proposed method are not affected by the size of motion blur, because average errors of the results are almost equal. The proposed method can improve the accuracy of the estimated position of the camera.

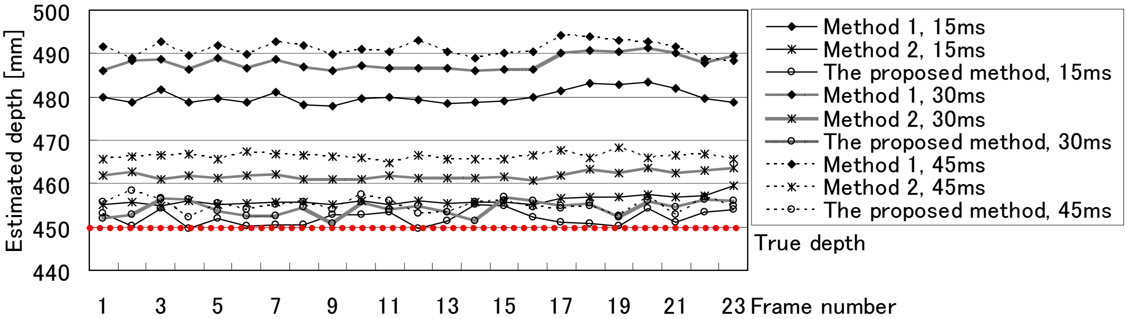


Figure 5: Estimated depth at various exposure times with three methods

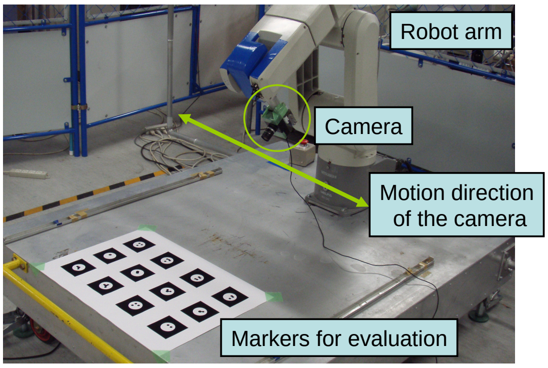


Figure 3: Experimental environment with robot arm

(a) Exposure time is 15 ms (b) Exposure time is 30 ms (c) Exposure time is 45 ms

Figure 4: Examples of captured images

3.3 Discussion

The experiments show that the proposed method improves the accuracy of the estimated position of the camera, and other methods are not robust for blur effects.

Method 1 cannot estimate the position of the camera accurately under both small as well as large motion blurs. This is caused by the irrelevant threshold in binarization, and by the corner detection algorithm, which assumes that corners of a marker in the captured image are almost sharp. This is predictable because this method assumes that there are neither defocus nor motion blurs in a captured image.

Method 2 is better in comparison with method 1, although both are affected by blur effects. The estimated corners are easily perturbed by errors in the angle of a fitted line to an edge of marker, and errors are increased when blur effects become large. Thus, the standard deviation of error in method 2 is larger than the others, as shown in Table 1, and the estimated depth is affected by blur effects.

On the other hand, the proposed method estimates corners directly so that the average and standard deviation of error is small as shown in Table 1, and the estimated depth is not highly affected by blur effects. For this reason, the proposed method is effective for such images as hardly defocus and motion blurred images.

4 Conclusion and Future Studies

This study proposed a method for improving the accuracy of estimating the position and posture of a camera. It is based on template matching of the simulated marker image with blur effects. In experiments, we have proven the effectiveness of this method by showing simulated as well as real images.

For further improvement, the computational cost should be reduced for augmented reality using multi-scale matching. In addition, the accuracy of estimated position and posture should be evaluated, for multiple markers in a scene.

Table 2: Average and standard deviation of errors of estimated depth in millimeter

Estimation methods	15ms		30ms		45ms	
	Ave.	Std.	Ave.	Std.	Ave.	Std.
Method 1	30.01	1.63	37.90	1.75	41.08	1.72
Method 2	6.09	1.04	11.90	0.88	16.29	0.78
The proposed method	2.02	1.75	4.29	1.81	5.04	1.63

References

[1] H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto and K. Tachibana: "Virtual Object Manipulation on a Table-top AR Environment," Proc. of IEEE/ACM Int. Symp. on Augmented Reality (ISAR 2000), pp. 111–119, 2000.

[2] M. Mohring, C. Lessig and O. Bimber: "Video See-Through AR on Consumer Cell Phones," Proc. of IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2004), pp. 252–253, 2004.

[3] Y. Nakazato, M. Kanbara and N. Yokoya: "A Localization System Using Invisible Retro-reflective Markers," Proc. of IAPR Conf. on Machine Vision Applications (MVA2005), pp. 140–143, 2005.

[4] L. Naimark and E. Foxlin: "Circular Data Matrix Fiducial System and Robust Image Processing for a Wearable Vision-Inertial Self-Tracker," Proc. of IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2002), pp. 27–36, 2002.

[5] G. Klein and T. Drummond: "Tightly Integrated Sensor Fusion for Robust Visual Tracking," Image and Vision Computing, Vol. 22, No. 10, pp. 769–776, 2004.

[6] G. Reitmayr and T. Drummond: "Going out: Robust Model-based Tracking for Outdoor Augmented Reality," Proc. of IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2006), pp. 109–118, 2006.

[7] B. Okumura, M. Kanbara and N. Yokoya: "Augmented Reality Based on Estimation of Defocusing and Motion Blurring from Captured Images," Proc. of IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2006), pp. 219–225, 2006.

[8] X. Zhang, S. Fronz and N. Navab: "Visual Marker Detection and Decoding in AR Systems: A Comparative Study," Proc. of IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2002), pp. 97–106, 2002.

[9] K. Sato, K. Takemoto, S. Uchiyama and H. Yamamoto: "A Registration Evaluation System Using an Industrial Robot," Proc. of IEEE/ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2006), pp. 79–87, 2006.