IPSJ Transactions on Computer Vision and Applications Vol.6 78–82 (July 2014)

[DOI: 10.2197/ipsjtcva.6.78]

Express Paper

Geometrical and Temporal Calibration of Multiple Cameras Using Blinking Calibration Patterns

Hirotake Yamazoe[1,a)]

Received: March 14, 2014, Accepted: April 24, 2014, Released: July 25, 2014

Abstract: In this paper, we propose a method to achieve positions and poses of multiple cameras and temporal synchronization among them by using blinking calibration patterns. In the proposed method, calibration patterns are shown on tablet PCs or monitors, and are observed by multiple cameras. By observing several frames from the cameras, we can obtain the camera positions, poses and frame correspondences among cameras. The proposed calibration patterns are based on pseudo random volumes (PRV), a 3D extension of pseudo random sequences. Using PRV, we can achieve the proposed method. We believe our method is useful not only for multiple camera systems but also for AR applications for multiple users.

Keywords: geometrical and temporal calibration, multiple cameras, augmented reality, mobile phones

1. Introduction

These days, augmented reality (AR)-based applications are employed in various situations. Many AR applications for iPhone and Android are registered. Furthermore, various AR systems have been introduced in shopping malls and museums.

However, although there are an enormous number of AR applications and systems, most do not suppose simultaneous use by multiple users. In this paper, we consider AR systems that can be used by multiple users simultaneously. Specifically, we aim at AR environments in which multiple users can see synchronized AR contents (**Fig. 1**).

Many researchers have proposed AR algorithms thus far. These algorithms are divided into two approaches: marker-based approaches [1], [2], [3] and markerless approaches [4] (using natural features etc.). Here, we focus on the marker-based approach.

In marker-based approaches [1], [2], AR markers are basically printed and static. Although relationships between marker and camera are obtained based on marker observations, temporal information or frame correspondences cannot be obtained. There are some examples that consider AR markers shown on a monitor [3], but AR markers themselves are static and do not consider temporal information.

In this paper, we propose a geometrical and temporal calibration method based on blinking calibration patterns. Our method can obtain not only positions and poses of cameras but also frame correspondences among camera observations. By using our blinking calibration patterns as an AR marker, we can achieve an AR environment in which multiple users can see synchronized AR contents.

Our blinking calibration patterns are based on pseudo random volumes (PRV), a 3D extension of pseudo random sequences [7]. In PRV, when a certain subvolume observation is obtained, the positions of the subvolume in the PRV can be uniquely determined. Using this property, we can achieve the proposed method.

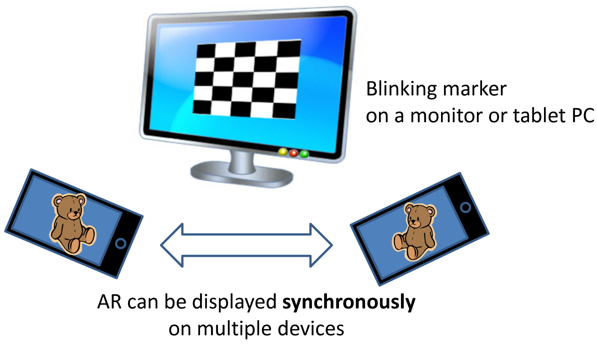

Fig. 1 Concept of method.

Recently, not only wearable devices with cameras and displays such as smart phones and tablet PCs but also head-mounted devices such as Google Glass [5] and Epson Moverio [6] have become more popular. We assume AR systems that can be used by multiple users at the same time will increase in various settings such as shopping malls. The proposed method is essential technology for achieving such AR applications.

In the next section, we introduce pseudo random volumes (PRV) and their properties and construction. In Section 3, we focus on the algorithm of the method. Section 4 gives experimental results and Section 5 concludes the paper.

2. Pseudo Random Volume

Before explaining pseudo random volumes (PRV), we will describe pseudo random sequences (PRS) and pseudo random arrays (PRA).

Pseudo random sequences (a.k.a. m-sequences, maximum

1 Osaka School of International Public Policy, Osaka University, Toyonaka, Osaka 560–0043, Japan

a) yamazoe@osipp.osaka-u.ac.jp

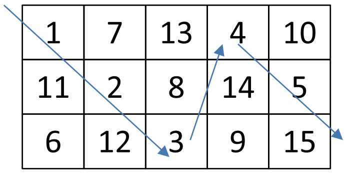

Fig. 2 Construction of pseudo random array (3×5).

length sequences) are binary sequences of cycle length $n = 2^m - 1$, constructed from m-bit shift registers. When the shift register is based on a primitive polynomial $h(x)$ of degree m, PRS can be generated [8]. For example, in the case of m = 4, one of the primitive polynomials is $h(x) = x^4 + x + 1$, and PRS $s = s_1, \ldots, s_n$ are generated as the following equation, $s_{i+4} = s_{i+1} + s_i$ (i = 1, ..., n − 4). An example of a generated PRS is as follows.

$$s_i = \{1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, \ldots\} \tag{1}$$

PRS have several properties [8]. Here, we introduce the following two properties relevant to this paper.

1) the window property: any consecutive m-bit sequences extracted from a certain PRS are different.

2) the ratio of 0 and 1: in any m-bit PRS, 0' and 1' appear $2^{m-1} - 1$ and 2^{m-1} times, respectively.

Pseudo random arrays are two-dimensional arrays of size $n_1 \times n_2$ constructed from PRS of length n ($n = 2^{k_1 k_2} - 1$). Here, $n_1 = 2^{k_1} - 1$ and $n_2 = n/n_1$ must be relatively prime. Since PRA also have the property that any possible $k_1 \times k_2$ subarrays are different, we can find the positions of the subarrays from $k_1 \times k_2$ subarray observation [9].

For example, in the case of $n = 2^{2 \cdot 2} - 1 = 15$, we can generate a 3×5 PRA ($n_1 = 2^2 - 1 = 3$ and $n_2 = 15/3 = 5$). Here, PRA can be constructed by writing the PRS down the main diagonal and continuing from the opposite side when an edge is reached as shown in **Fig. 2**. The PRA made from the above PRS of Eq. (1) is as follows:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{2}$$

Pseudo random volumes (PRV) are three-dimensional extensions of PRS and PRA. As with PRS and PRA, PRV also have the property that any possible $k_1 \times k_2 \times k_3$ subvolumes are different. By using this property, we can achieve the blinking calibration patterns for geometrical and temporal calibration of multiple cameras.

Next, we will explain the construction of the PRV [7]. PRV of size $n_1 \times n_2 \times n_3$ are constructed from PRS of length n ($n = 2^{k_1 k_2 k_3} - 1$), where n_1, n_2 and n_3 need to be relatively prime. Here, let us consider a $3 \times 5 \times 17$ PRV. We first make a PRS of length $n = 2^8 - 1 = 255 = 3 * 5 * 17$ from a primitive polynomial of degree 8 (e.g. $h(x) = x^8 + x^6 + x^5 + x^1 + 1$). Then, we fill in a $3 \times 5 \times 17$ volume using the generated PRS, by writing the PRS down the main diagonal and continuing from the opposite side when an edge is reached as shown in **Fig. 3**.

As described in the following sections, a marker and cameras need to share the PRV in the proposed method. However in practice, the whole PRV pattern does not need to be shared; just the

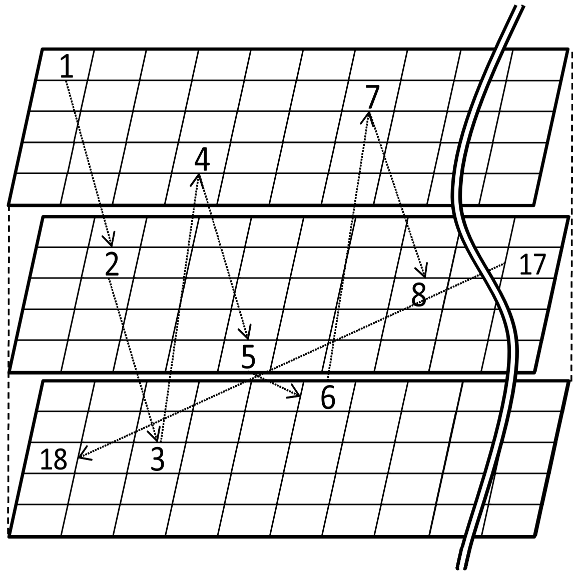

Fig. 3 Construction of pseudo random volume ($3 \times 5 \times 17$).

Table 1 Possible and appropriate PRVs.

m	$n = 2^m - 1$	Possible PRVs
8	255	$3 \times 5 \times 17$
10	1023	$3 \times 11 \times 31$
12	4095	$5 \times 9 \times 91$, $5 \times 7 \times 117$, $7 \times 9 \times 63$, $7 \times 13 \times 45$, etc.
16	65535	$3 \times 5 \times 4369$

parameters n, n_1, n_2, n_3 and the primitive polynomial $h(x)$ are sufficient because PRV can be uniquely generated when n, n_1, n_2, n_3 and $h(x)$ are given. **Table 1** shows examples of possible PRVs appropriate for our purpose.

3. Algorithm

3.1 System Overview

As shown in Fig. 1, the proposed blinking calibration patterns are shown on a monitor or tablet PC and multiple cameras are observing the markers. From the camera observations, the positions and poses of the cameras can be obtained. In addition, frame correspondences among camera observation sequences are achieved.

Figure 4 shows the process flow of the proposed method. First, the marker pattern constructed from PRV is displayed on a monitor or tablet PC. Next, the marker is observed by multiple cameras, and marker patterns are extracted from the observed images. The frame correspondences among image sequences are obtained from obtained marker patterns by using the PRV's properties. Then, positions and poses of the cameras are estimated.

3.2 Marker Detection

First, markers are detected from the observed image sequences. The procedures for the marker detection are based on Ref. [3]. Here, we briefly describe marker detection.

We first extract edges from observed images. Here, we assume the marker is covering the dominant area in the images. Next, we estimate two vanishing points from extracted edges. Let $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ be the vanishing points, and $\boldsymbol{l}_i$ be the edges corresponding to the vanishing points $\boldsymbol{v}_i$. We can obtain the following equation:

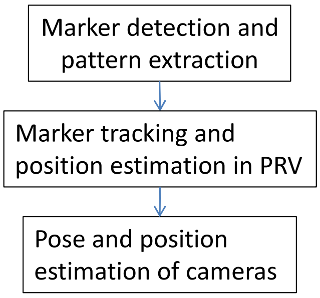

Fig. 4 Process flow.

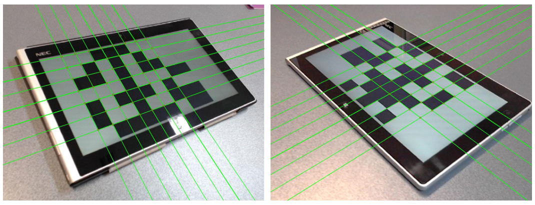

Fig. 5 Examples of marker detection and grid extraction.

$$\forall i : \boldsymbol{v}_k \cdot \boldsymbol{l}_i = 0. \tag{3}$$

The vanishing points $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ can be calculated as the minimum eigenvector of eigenvalue decomposition of the following matrix $\boldsymbol{C}$:

$$\boldsymbol{C} = (\boldsymbol{l}_0, \ldots, \boldsymbol{l}_N)(\boldsymbol{l}_0, \ldots, \boldsymbol{l}_N)^t. \tag{4}$$

Since the obtained edges include false edges (not derived from marker patterns), we employ RANSAC for estimating vanishing points $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$.

Then, positions of grids are estimated using the fact that correct edges (grids) should be equally spaced as follows:

$$\boldsymbol{l}_i = \hat{\boldsymbol{l}}_{base} + (ki + q)\hat{\boldsymbol{h}}, \tag{5}$$

where $\hat{\boldsymbol{h}} = \boldsymbol{v}_1 \times \boldsymbol{v}_2$. $\hat{\boldsymbol{l}}_{base}$ is an arbitrary line through the vanishing point, k and q are parameters for determining intervals and positions of the grids. **Figure 5** shows examples of the grid detection results.

3.3 Pattern Extraction

Next, marker patterns are extracted from estimated grid positions. Here, $M \times N$ grid is obtained where $g_{i,j}$ are intersection 2D positions. $v_{i,j}$ are averaged pixel values of area $a_{i,j}$ (**Fig. 6**). $v_{i,j}$ are grouped into several groups by using X-means method [10]. Depending on the observation timing, sometimes we can take images of the marker in a switching timing. In such a case, a number of groups of $v_{i,j}$ will become more than three. These images are excluded for the following process. Let $p_{i,j}$ be extracted patterns and

$$p_{i,j} = \begin{cases} 1 & (white) \\ 0 & (black) \\ -1 & (other) \end{cases} \tag{6}$$

where "other" includes when i and j are out of range. **Figure 7** shows pattern extraction results.

3.4 Marker Tracking and Position Estimation in PRV

Here, let us consider the images of frame t and $t+1$ and grids

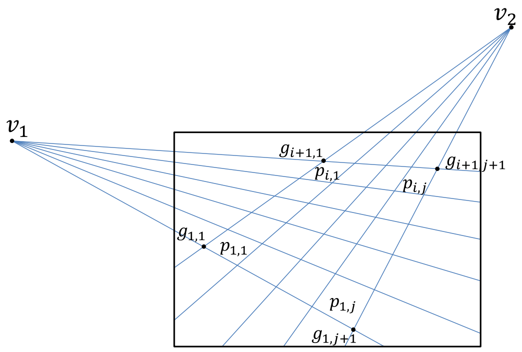

Fig. 6 Grid positions and pattern extraction.

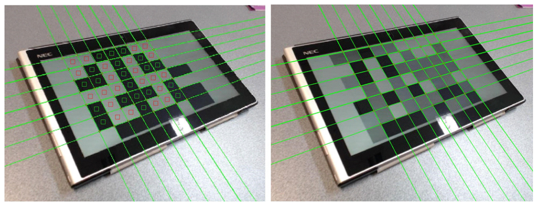

Fig. 7 Examples of pattern detection (left) and failure case (right).

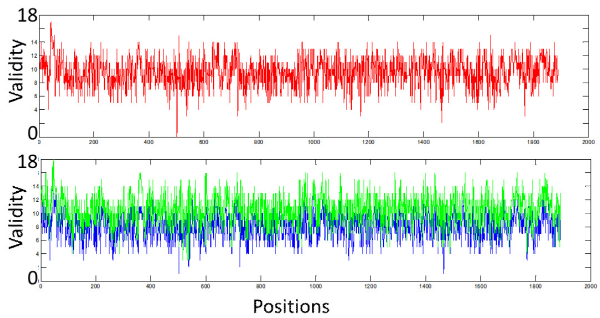

Fig. 8 Results of correspondence validity.

of sizes $M_t \times N_t$ and $M_{t+1} \times N_{t+1}$ are detected, respectively. Assuming the difference of observation time is small, we find initial grid correspondence between $\boldsymbol{G}^{(t)} = \{g_{1,1}, \ldots, g_{M_t,N_t t}\}$ and $\boldsymbol{G}^{(t+1)}$, the grids obtained in frame t and $t+1$ as follows:

$$\Delta i, \Delta j = \arg\min_{\Delta i, \Delta j} \sum_{i,j} d(i, j, \Delta i, \Delta j, t), \tag{7}$$

$$d = \begin{cases} (g_{i,j} - g_{i+\Delta i, j+\Delta j})^2 & (g_{i,j}, g_{i+\Delta i, j+\Delta j} > 0) \\ 0 & (else) \end{cases}. \tag{8}$$

Next, we evaluate validity of grid correspondence based on the window property of PRV and continuity of observation timing. In PRV, we can uniquely determine the positions of subvolumes from observations of subvolumes. Let us consider the case in which more large subvolumes are obtained. In such a case, if false correspondences exist in the subvolume, the tracked subvolume does not match any subvolume in the PRV. Thus, the validity of the grid correspondence $\boldsymbol{V}$ is calculated as follows:

$$\boldsymbol{V}_{\alpha,\beta,\gamma} = \sum_{i,j,k} \left(\boldsymbol{S}_{i,j,k} - \boldsymbol{P}_{\alpha+i,\beta+j,\gamma+k} \right), \tag{9}$$

where $\boldsymbol{P}$ are the patterns of the PRV, $\boldsymbol{S}$ are the patterns of the subarray. α, β, γ are positions in the PRV.

Figure 8 shows the calculated validity of $7 \times 9 \times 65$ PRV and $3 \times 4 \times 3$ subvolumes. The horizontal axis denotes the positions in

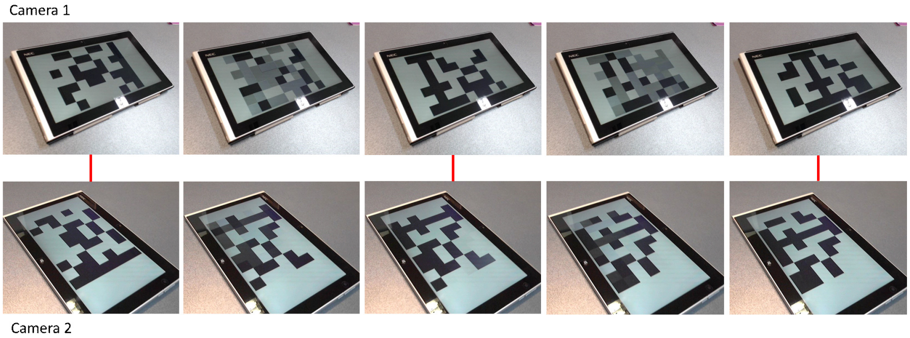

Fig. 12 Experimental results (red lines denote found correspondences).

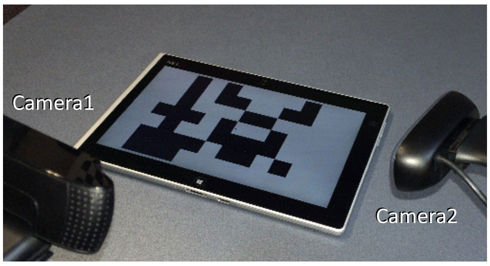

Fig. 9 Experimental setup.

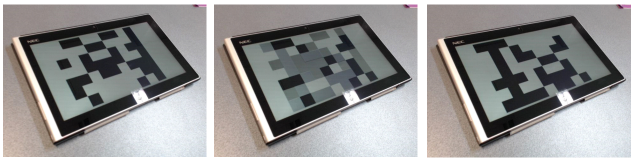

Fig. 10 Examples of observing images (camera 1).

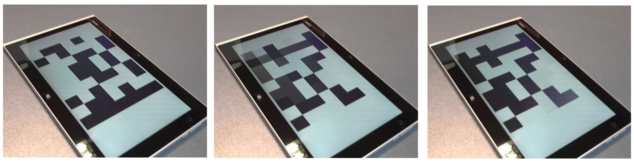

Fig. 11 Examples of observing images (camera 2).

the PRV, the vertical one denotes the validity (small means high validity). The red line shows correct subvolume. Blue and green lines are subvolumes that include false correspondences (1 pixel). By searching minimum $V_{\alpha,\beta,\gamma}$, we can find positions of subarray in the PRV, and achieve frame correspondences among cameras.

Here, we considered subarrays that are one grid larger than the necessary size ($2 \times 3 \times 2$). Actually, since we can obtain further larger observations, the validity can be estimated more robustly.

3.5 Position and Pose Estimation of Cameras

From the obtained information (positions of marker grid corners) correspondence with the marker plane, we can estimate the positions and poses of cameras by using homography relations [11].

4. Experiment

To evaluate the effectiveness of the proposed method, we performed the following experiment.

Figure 9 shows the experimental environment. A 10.1 inch Tablet PC is used to show the marker. Two USB cameras observe the marker. Geometrical and temporal calibration of two cameras are performed using observed image sequences.

As a PRV for the marker patterns, we constructed a PRV of size $7 \times 9 \times 65$ from the PRS ($m = 12$, $n = 2^m - 1 = 4{,}095$, and primitive polynomial $h(x)$ is $x^{12} + x^7 + x^4 + x^3 + 1$). Sixty-five marker patterns of size 7×9 for the blinking patterns are generated from the constructed PRV. These patterns are shown on the tablet PC in 10 frames/second. The period of the blinking patterns is 6.5 [sec]. **Figures 10** and **11** show examples of the observation of the displayed markers.

Figure 12 shows the experimental results. As can be seen, we can find the frame correspondences by using our method.

5. Conclusion

We proposed a method to estimate positions and poses of multiple cameras and to achieve frame correspondence among cameras by using blinking calibration patterns. The blinking calibration patterns that are constructed based on pseudo random volumes (PRV) are shown in a display or a tablet PC and all cameras observe the marker. Using PRV's properties, we can achieve geometrical and temporal calibration among multiple cameras. We confirmed the effectiveness of our system through experiments.

Future works include the implementation of an AR system based on our method and estimation algorithm for sub-frame synchronization among cameras by using the transition frames. We will also investigate extension of the marker patterns to non-binary (multi-value) markers for robust marker detection and tracking.

Acknowledgments This research was partly supported by JSPS KAKENHI Grant Number 25730114.

References

[1] Kato, H. and Billinghurst, M.: Marker tracking and hmd calibration for a video-based augmented reality conferencing system, *Proc. 2nd IEEE and ACM International Workshop on Augmented Reality* (*IWAR 99*) (1999).

[2] Fiala, M.: ARTag, a fiducial marker system using digital techniques, *Proc. Computer Vision and Pattern Recognition 2005*, pp.590–596 (2005).

[3] Herout, A., Szentandrasi, I., Zacharias, M., Dubska, M. and Kajan, R.: Five Shades of Grey for Fast and Reliable Camera Pose Estimation, *Proc. Computer Vision and Pattern Recognition 2013*, pp.1384–1390 (2013).

[4] Klein, G. and Murray, D.: Parallel Tracking and Mapping for Small AR Workspaces, *Proc. International Symposium on Mixed and Augmented Reality* (*ISMAR'07*) (2007).

[5] Google Glass, available from ⟨http://www.google.com/glass/start/⟩.

[6] Epson Moverio, available from ⟨http://www.epson.jp/products/moverio/⟩.
[7] Green, D.H.: Structural properties of pseudorandom arrays and volumes and their related sequences, *IEE Proc.*, Vol.132, Pt.E, No.3, pp.133–145 (1985).
[8] MacWilliams, F.J. and Sloane, N.J.A.: Pseudo-random sequences and arrays, *Proc. IEEE*, Vol.64, No.12, pp.1715–1730 (1976).
[9] Lloyd, S.A. and Burns, J.: Finding the position of a subarray in a pseudo random array, HP Labs Technical reports, HPL-91-159 (1991).
[10] Pelleg, D. and Moore, A.W.: X-means: Extending K-means with Efficient Estimation of the Number of Clusters, *Proc. International Conference on Machine Learning '00*, pp.727–734 (2000).
[11] Weng, J., Ahuja, N. and Huang, T.S.: Motion and structure from point correspondences with error estimation: Planar surfaces, *IEEE Trans. Signal Processing*, Vol.39, No.12, pp.2691–2717 (1991).

(Communicated by *Takeshi Masuda*)