Short Paper

A Correspondence Search Technique for Geometric Correction of Projected Images

（投影像の幾何補正のための画像対応付け手法）

Toru Takahashi†, Takafumi Aoki (member)†, Koichi Ito (member)†, Satoshi Kondo††

Abstract This paper presents a technique for correspondence search between projector and camera images using SIFT (Scale-Invariant Feature Transform) and POC (Phase-Only Correlation). The correspondence search is a key task for camera-based geometric correction of projected imagery. In conventional methods, the correspondence is obtained by projecting the structured light on a projector screen. This paper proposes a geometric correction technique where we use only one snapshot of ordinary video sequences to find projector-camera correspondence. We adopt (i) SIFT-based feature matching to estimate an approximate homography between projector and camera images with a limited number of corresponding points and (ii) POC-based dense sub-pixel correspondence search to adjust precise position of corresponding points. Through a set of experiments, we demonstrate that the proposed technique achieves high-accuracy geometric correction, even if significant brightness change is observed between the original video content to be projected and the real projector image captured by a camera.

Key words: projectors, image correction, sub-pixel image registration, phase-only correlation, SIFT

1. Introduction

Recently, large display systems using projectors and cameras have been widely reported[1]. To achieve seamless projected imagery, these systems correct each projector's image using information from the camera-captured images. In general, a system consisting of controllable projection units and cameras is called *projector-camera systems* (PROCAMS), and is widely studied to combine projection technology with computer vision[2]. Camera-based image correction[2)3)] is one of the applications of PROCAMS.

The conventional image correction methods use a set of structured light patterns to detect deformation of projected image due to the 3D structure of projector screen[2]. This approach is not suitable for real-time applications, since precise image correction requires projection of 20–30 patterns in advance. Addressing this problem, we propose an image-based technique of geometric deformation analysis without projecting structured light patterns, where we employ a snapshot of ordinary video sequence and an advanced image correspondence technique. The proposed technique enables the system to estimate deformation parameters in a background process of video projection. The key to success with this approach is the image correspondence algorithm that is robust against image intensity variation. Our algorithm combines SIFT[4] and POC[5] to achieve intensity-invariant dense correspondence search between projector and camera images.

2. Geometric correction of projected images

2.1 Basic requirements

The purpose of geometric correction is to find a set of parameters for image transformation between projector and camera images. If a projector screen is flat and lens distortion can be ignored, we can represent the transformation as one homography. In this case, we can estimate the transformation parameters with small number of corresponding points.

In practical situation, however, non-linear transformation model is required in many cases, and hence dense image correspondence is essential. In this paper,

Received September 6, 2010; Revised March 8, 2011; Accepted April 5, 2011
† Graduate School of Information Sciences, Tohoku University
(6–6–05, Aramaki Aza Aoba, Aoba–ku, Sendai–shi, 980–8579 Japan)
†† Panasonic Corporation
(1006 Kadoma, Kadoma–shi, 571–8501 Japan)

we assume that a projector screen is planar with some non-linear deformation. This model covers a wide variety of projector screen (e.g., curved walls and curtains) as well as various image deformation. For nonlinear geometric correction of projected images, B-spline approximation[6)] and piecewise linear approximation[1)] can be used. If the accurate and dense correspondence between projector and camera images can be obtained, piecewise linear approximation would be a reasonable approach from the viewpoint of computational cost. In this paper, we focus on a technique for correspondence search between projector and camera images. Therefore, we employ the piecewise linear approximation method to correct projected image deformation. In order to avoid misalignment in every image block, dense correspondence search which is highly robust against image intensity variation is indispensable.

2.2 Correspondence search with structured light patterns

A conventional approach of dense correspondence search is to use multiple structured light patterns, where we employ 20–30 patterns to be projected sequentially. As a typical approach, we adopt a hybrid method using a set of gray-code patterns and an equally-spaced grid-dot pattern in our experiments described in Section 4, where we use the gray-code patterns to partition projected imagery into small sub-regions, and find pixel correspondence between projector and camera images by using the grid-dot pattern.

3. Proposed correspondence search technique

Our proposed technique uses SIFT and POC to find corresponding points between projector and camera images with one snapshot of a video sequence. SIFT-based feature matching is robust against global image transformation, e.g., projective transformation, although the sparse corresponding points cannot be used for non-linear geometric correction. On the other hand, POC-based correspondence search technique can obtain dense corresponding points, although this technique cannot work well for the images having global image transformation. To achieve accurate nonlinear geometric correction of projected images, we first adopt SIFT-based feature matching to estimate an approximate homography between projector and camera images, and then adopt POC-based dense sub-pixel correspondence search to adjust local non-linear relationship (**Fig. 1**). The proposed algorithm can be summarized as follows.

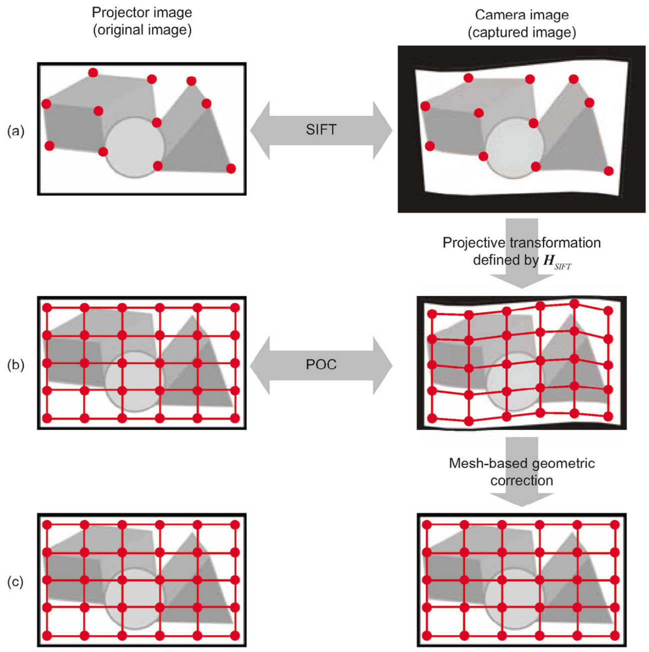

Fig. 1 Proposed correspondence search technique: (a) homography estimation by SIFT, (b) dense correspondence search by POC, and (c) mesh-based nonlinear geometric correction.

3.1 Homography estimation using SIFT

We first use SIFT[4)] to estimate a homography $\boldsymbol{H}_{SIFT}$ between camera and projector image coordinates. The estimated homography $\boldsymbol{H}_{SIFT}$ gives a first-order approximation for non-linear geometric transformation between camera and projector image planes. A major advantage of using SIFT for the first step of correspondence search is that its feature points are robust against global image transformation as well as illumination variation.

3.2 Dense correspondence search using POC for precise geometric correction

When the projector screen has a complex curvature and the projected imagery involves non-linear deformation, more precise geometric correction will be essential. SIFT-based approach produces only a limited number of corresponding points, and is not suitable for finding dense correspondence for non-linear geometric correction. This is the major motivation of introducing advanced area-based image correspondence technique.

We first warp the camera image by the projective transformation defined by $\boldsymbol{H}_{SIFT}$. Through this normalization, the appearance of camera image becomes close to that of the original image (i.e., projector image). Next step is to correct the nonlinear geometric deformation that cannot be compensated by a simple homography. To do this, we employ a mesh-based piecewise-linear image transformation (as illustrated in

Fig. 1 (c)). The key to success with this approach is an accurate area-based dense correspondence algorithm that is robust against photometric distortion due to projector-camera transfer function. The hierarchical correspondence search algorithm based on Phase-Only Correlation (POC)[5] seems to be one of the best approaches for this purpose, since the algorithm produces stable corresponding points even when significant intensity variation is included in the images.

The overall procedure of projector-camera correspondence can be summarized as follows (**Fig. 1**):

Step 1: Reduce the size of projector image (i.e., original image) and that of camera image by $1/S_p$ and $1/S_c$, respectively, in order to reduce computation time. ($S_p = 2$ and $S_c = 4$ in our experiments.)

Step 2: Estimate the homography $\boldsymbol{H}_{SIFT}$ from the reduced camera image to the reduced projector image using SIFT feature points (**Fig. 1** (a)).

Step 3: Warp the camera image by the homography $\boldsymbol{H}_{SIFT}$ (i.e., projective transformation).

Step 4: For a set of equally spaced reference points in the original projector image, find the corresponding points in the warped camera image using POC-based hierarchical block matching[5] (**Fig. 1** (b)). In our experiments, the number of layers for coarse-to-fine search is three and the block size is 64×64.

Step 5: Transform the coordinates of the obtained corresponding points by $\boldsymbol{H}_{SIFT}^{-1}$, which defines the coordinate transformation from the warped camera image to the original camera image.

Our experimental observation clearly shows that the proposed approach is highly robust against geometric and photometric distortion frequently appeared in practical projector-camera applications. The obtained correspondence can be used for precise mesh-based geometric correction (**Fig. 1** (c)).

4. Experiments

In our experiments, we evaluate the performance of our proposed technique in terms of error in correspondence, which is defined as an average distance (in pixel) between the corresponding points obtained by the proposed algorithm and the accurate corresponding points obtained by the active method using structured-light projection.

We carefully compare the three different approaches of projector-camera correspondence: (i) a method using only SIFT, (ii) a method combining SIFT and the area-based block matching using SSD (Sum of Squared Differences) with a sub-pixel estimation technique using parabola fitting, and (iii) the proposed method.

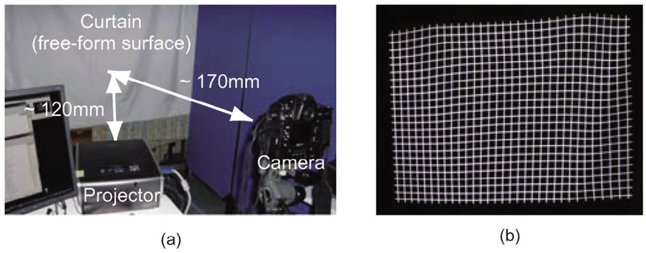

Fig. 2 Experimental setup: (a) projector-camera system for our experiments, and (b) image deformation on the projector screen.

We use "Canon X700" projector ($1,024 \times 768$ pixels) and "Nikon D3" camera ($4,256 \times 2,832$ pixels) with its lens "Ai AF Nikkor 50mm F1.4D" to build a projector-camera system shown in **Fig. 2** (a). The range of coordinates for correspondence search in the projector image is $24 \leqq x \leqq 1,000$ in horizontal and $24 \leqq y \leqq 744$ in vertical direction with reference points aligned at 16-pixel interval on the projector image. The projector screen used in our experiments are a white flat plate made of wood and a white curtain with a free-form surface as shown in **Fig. 2** (b). The images used for correspondence search are from three video sequences: *Buildings along the canal*, *Church*, and *Character pattern*. From every sequence, we extract with a spacing of 30 frames. In the each frame, we set the reference points ($62 \times 46 = 2,853$ points) and search corresponding points on the camera images using above three different approaches.

Figure 3 (a) plots the distribution of corresponding points with respect to the error in pixel coordinates, where the points whose error is larger than 5 pixels are regarded as outliers. Note that the method (i) uses only the homography $\boldsymbol{H}_{SIFT}$ to evaluate the accuracy of the correspondence, that is, we calculate corresponding points coordinates for all the reference points using the homography $\boldsymbol{H}_{SIFT}^{-1}$. The SIFT-only method (i) can produce only a limited number of corresponding points and is not suitable for non-linear deformation correction. The proposed approach (SIFT+POC method (iii)) shows significantly higher performance compared with the SIFT+SSD method (ii). For all the video sequences, average error is about 1–2 pixels.

The pictures in **Fig. 3** (b) show the original projector images, (c) show camera images without deformation correction, (d) show camera images warped by $\boldsymbol{H}_{SIFT}$ and (e) show camera images after geometric correction

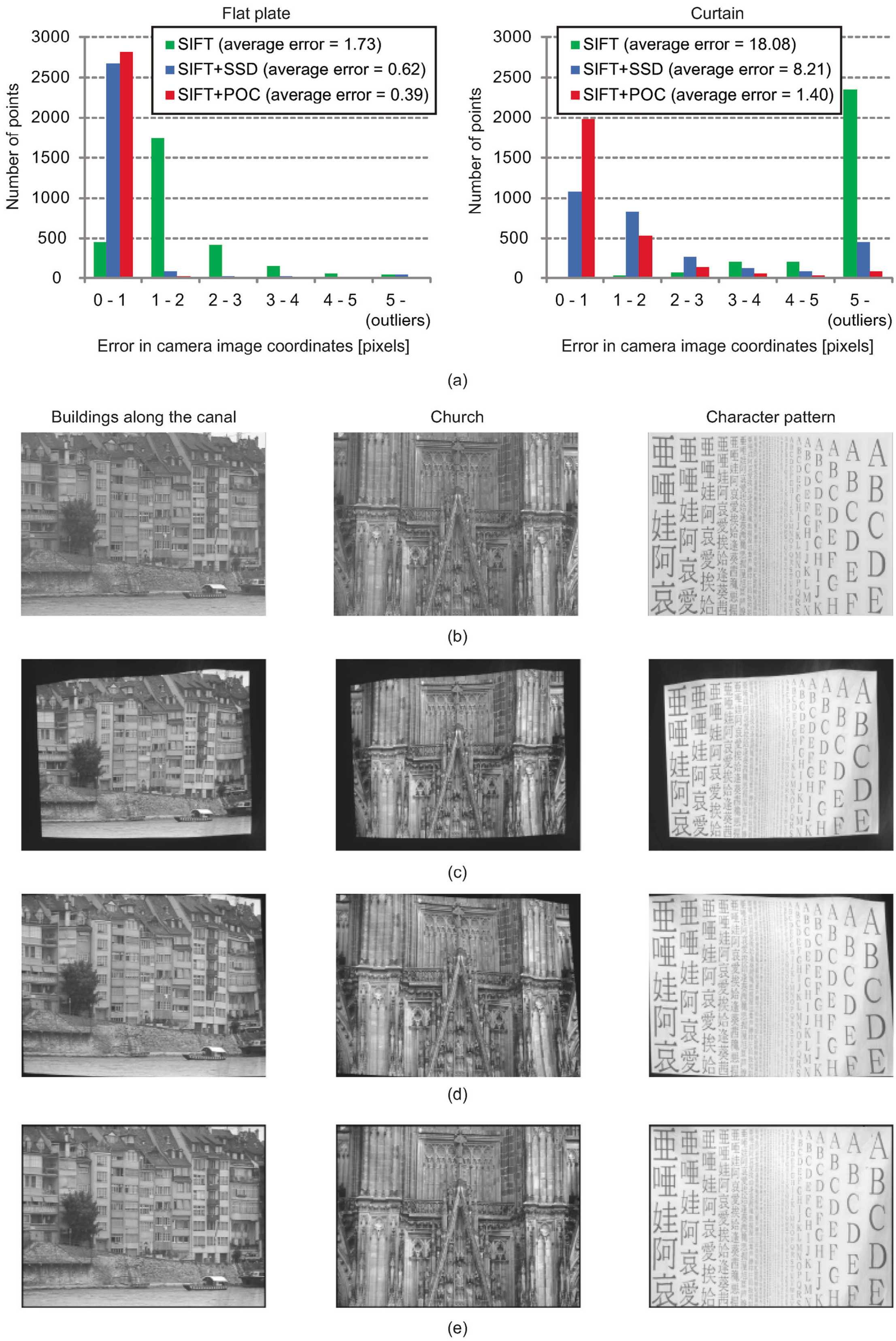

Fig. 3 Experimental results: (a) distribution of corresponding points with respect to the correspondence error in pixel, (b) original projector images, (c) camera images, (d) camera images warped with $\boldsymbol{H}_{SIFT}$, and (e) camera images after geometric correction by the proposed technique.

by the proposed method. As is observed in these results, if the projected imagery has enough texture, the proposed method achieves highly accurate projector-camera correspondence, whose accuracy is comparable to the active method using structured light illumination. In the experiment, we implement the proposed method with MATLAB and C. It takes about 12 seconds on a computer with Intel(R) Core(TM)2 Extreme X9650 (3.0 GHz) processor to obtain the correspondence. The processing time would be reduced by using GPU (Graphics Processing Unit) implementation of SIFT[7] and POC.

5. Conclusion

This paper presents a technique for correspondence search between projector and camera images using SIFT and POC. We adopt SIFT feature matching to estimate an approximate homography between projector and camera images and adopt POC to find dense local relationship. Our experimental observation clearly shows that the proposed combination (SIFT+POC) can achieve highly robust, accurate and dense matching of images even if significant geometric and photometric distortion is included. This property seems to be useful in many computer vision applications not limited to projector-camera systems.

〔References〕

1) H. Fuchs, A. Majumder, and M. Brown, Practical Multi-Projector Display Design, A K Peters, Ltd., (2007)
2) O. Bimber and R. Raskar, Spatial Augmented Reality Merging Real and Virtual Worlds, A K Peters, Ltd., (2005)
3) O. Bimber, D. Iwai, G. Wetzstein, and A. Grundhöfer, "The visual computing of projector-camera systems", Computer Graphics Forum, **27**, 8, pp.2219–2245 (Sept. 2008)
4) D. Lowe, "Distinctive image features from scale-invariant keypoints," Int'l J. Computer Vision, **60**, pp.91–110 (Nov. 2004)
5) K. Takita, M.A. Muquit, T. Aoki, and T. Higuchi, "A subpixel correspondence search technique for computer vision applications," IEICE Trans. on Fundamentals, **E87-A**, 8, pp.1913–1923 (Aug. 2004)
6) J. Kybic, P. Thevenaz, A. Nirkko, and M. Unser, "Unwarping of unidirectionally distorted EPI images," IEEE Trans. Medical Imaging, **19**, 2, pp.80–93 (Feb. 2000)
7) "SiftGPU: A GPU Implementation of Scale Invariant Feature Transform (SIFT)", http://www.cs.unc.edu/~ccwu/siftgpu/.


Toru Takahashi received the B.E. degree in information engineering, and the M.S. degree in information sciences from Tohoku University, Sendai, Japan, in 2006 and 2008, respectively. He is currently working toward the Ph.D. degree of the Graduate School of Information Sciences at Tohoku University. His research interest includes signal and image processing, and computer vision.


Takafumi Aoki received the B.E., M.E., and D.E. degrees in electronic engineering from Tohoku University, Sendai, Japan, in 1988, 1990, and 1992, respectively. He is currently a Professor of the Graduate School of Information Sciences at Tohoku University. For 1997.1999, he also joined the PRESTO project, Japan Science and Technology Corporation (JST). His research interests include theoretical aspects of computation, VLSI systems, multiple-valued logic, digital signal processing, image sensing, computer vision, biometric authentication, and secure embedded systems.


Koichi Ito received the B.E. degree in electronic engineering, and the M.S. and Ph.D. degree in information sciences from Tohoku University, Sendai, Japan, in 2000, 2002 and 2005, respectively. He is currently an Assistant Professor of the Graduate School of Information Sciences at Tohoku University. From 2004 to 2005, he was a Research Fellow of the Japan Society for the Promotion of Science. His research interests include signal and image processing, and biometric authentication.


Satoshi Kondo received his B.S., M.S. and Ph.D. degrees at Osaka Prefecture University in 1990, 1992 and 2004, respectively. He joined Matsushita Electric Industrial Co., Ltd. (currently Panasonic Corporation) in 1992. His research interests are in the fields of video coding and image/video signal processing. He has been also a participant in international standardization activities in IEEE1394 Trade Association, ISO/IEC MPEG & ITU-T VCEG and DICOM.