21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Calibration-Free Projector-Camera System for Spatial Augmented Reality on Planar Surfaces

Takayuki Nakamura, François de Sorbier, Sandy Martedi, Hideo Saito
Graduate School of Science and Technology, Keio University, Yokohama, Japan
{*takayuki,fdesorbi,sandy,saito*}*@hvrl.ics.keio.ac.jp*

Abstract

Spatial augmented reality extends augmented reality by projecting virtual data directly on a target surface, but requires to calibrate a projector-camera system. This paper introduces a free-calibration projector-camera system for spatial augmented reality with a planar surface. A pattern is projected on the target surface that can be freely moved. The main difficulty is to make the projected image fitting the target surface. This is often achieved by calibrating the projector according to the camera's pose, but the calibration process is offline and restricts the configuration of the setup. Our proposed method allows to independently move the projector, the camera or the surface while the projected image continuously fit the target surface. The experimental results demonstrate the efficiency of our calibration-free spatial augmented reality approach when applied on a moving planar surface.

1. Introduction

Augmented Reality (AR) augments the real world with virtual information. This process often mixes data generated on a computer with a live video stream. A target surface is detected and tracked in order to correctly display a virtual object onto the target surface. Thanks to the latest researches and development advances, it is now possible to add a projector that directly displays the virtual information. It is called Spatial Augmented Reality (SAR) [6] and is practical because users no longer need to carry a visualization device. The main difficulty of this approach is to modify the projected image to make it fit the target surface. It is often achieved by calibrating the projector-camera system which is offline and restricts the configuration of the setup.

The goal of our research is to project a texture on a moving textured planar surface with an uncalibrated projector-camera system like in Fig 1. The target planar surface can be freely moved while the projected image continuously fits the surface even if the projector or the camera are shifted. It means that the surface and the projected image have to be tracked. However, the projection hinders the tracking, since it badly modifies the appearance of the surface and prevents the use of standard trackers based on feature detection.

We propose to address this problem by adapting the Lucas-Kanade algorithm, a well known planar surface tracking method that minimizes the difference between a captured image and a target image. Our algorithm can simultaneously track both a rich planar surface texture and a projection texture. We then warp back both textures to the template coordinates and minimize the sum of the difference between each warped textures and the template. The target surface, the camera and the projector can then be freely moved or their parameters can be modified since no calibration is required.

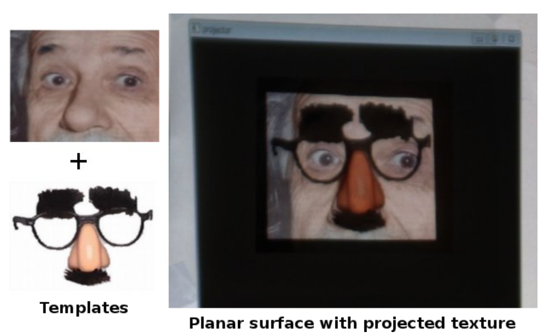

Figure 1. Goal of our method.

After introducing some related works, we will explain our approach in Section 3. Finally, we present our experimental results in Section 4 with an evaluation of the accuracy of our method.

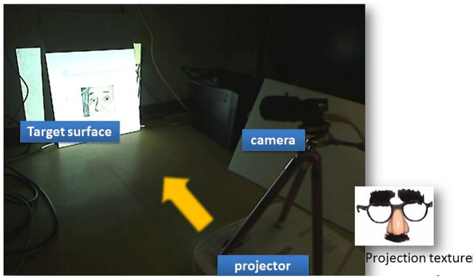

Figure 2. Overview of our setup.

2. Related works

In spatial augmented reality, the projected image needs to be transformed for fitting the target surface in real-time. Besides the tracking of the moving target object, we need to modify the projected image. However, common tracking techniques based on feature points [4, 11, 9] or gradient [5, 2] cannot be applied because of the lighting environment induced by the projector that hinders the features points.

Several solutions are focusing on imperceptibly embed markers like [7, 14] or sensing devices [3, 12]. But they reduce the dynamic range of the projector or require to set sensors on the object. [13, 8] proposed to add physical markers to track the surface, but the pattern should not be projected onto these markers. Leung *et al.* [10] proposed to track the borders of the planar surface. Audet *et al.* [1] introduced a directly alignment method of a projection image with a textured surface by minimizing the error function of a gradient based tracking.

However, all these methods use a calibrated projector-camera pair and implies that they might fail if components of the system are separately moved or if its parameters like the focal length are changed.

3. Our Method

3.1. The geometric Model

Our projector-based Lucas-Kanade algorithm defines two homographies H_s and H_p for warping the surface texture and the projector texture. These warping functions are respectively defined as $W(x;s), W(x;p)$ where s and p are homography's parameters and $\mathbf{x} = (x, y)$ is the coordinate of the image. The geometric model that defines the relation between the templates T_s and T_p, and the input image I warped back onto the template's coordinate system can be written as follows:

$$T_s(x) = I(W(\mathbf{x}; s)), \tag{1}$$

$$T_p(x) = I(W(\mathbf{x}; p)). \tag{2}$$

Also, in terms of coordinates, Eq.(1) and Eq.(2) show that a pixel $\mathbf{x_c}$ from the camera image is transformed by each homography H_s and H_p into a pixel $\mathbf{x_t}$ in the template image.

3.2. The energy function

The projection error is minimized by using the Lucas-Kanade algorithm which is a robustness direct alignment method against noise. This algorithm minimizes the sum of squared error between a template T and an input image I warped back into the coordinate frame of the template. Baker and Matthews have proposed the inverse compositional Lucas-Kanade [2] that switches the role of the template T and the input image I. This methods also allows to pre-compute several parts like the Hessian which improves the computational time during the incremental warp. Since we need to handle a surface texture and a projector texture, we propose to create an energy function E depending on the projector's energy function E_p and the surface's energy function E_s. E is then defined as follows:

$$E = E_s + E_p, \tag{3}$$

with

$$E_s = \sum_{\mathbf{x}} [T(W(\mathbf{x}; \Delta s)) - I(W(\mathbf{x}; s))]^2, \tag{4}$$

$$E_p = \sum_{\mathbf{x}} [T(W(\mathbf{x}; \Delta p)) - I(W(\mathbf{x}; p))]^2, \tag{5}$$

The template image T is then defined as the mix of the template image for the surface T_s and the template image for the projector T_p. Then $T = T_s \times T_p$.

3.3. Minimization

The goal of our method is to minimize Eq.(3) with respect to Δs and Δp, and to update the warping functions as follows:

$$W(\mathbf{x}; s) \leftarrow W(\mathbf{x}; s) \circ W(\mathbf{x}; \Delta s)^{-1}, \tag{6}$$

$$W(\mathbf{x}; p) \leftarrow W(\mathbf{x}; p) \circ W(\mathbf{x}; \Delta p)^{-1}. \tag{7}$$

Δs and Δp can be derived from Eq.(3) as follows:

$$\Delta s = He_s^{-1} \sum_{\mathbf{x}} [\nabla T_s \frac{\partial W}{\partial s}]^T [I(W(\mathbf{x}; s)) - T(\mathbf{x})] \tag{8}$$

$$\Delta p = He_p^{-1} \sum_{\mathbf{x}} [\nabla T_p \frac{\partial W}{\partial p}]^T [I(W(\mathbf{x}; p)) - T(\mathbf{x})] \tag{9}$$

with the Hessian matrices He_s:

$$He_s = \sum_{\mathbf{x}} [\nabla T_s \frac{\partial W}{\partial s}]^T [\nabla T_s \frac{\partial W}{\partial s}] \tag{10}$$

and He_p:

$$He_p = \sum_{\mathbf{x}} [\nabla T_p \frac{\partial W}{\partial p}]^T [\nabla T_p \frac{\partial W}{\partial p}] \tag{11}$$

and where ∇T_p and ∇T_s are the gradient images of respectively the projected and surface surfaces. $\nabla T_s \frac{\partial W}{\partial s}$ and $\nabla T_p \frac{\partial W}{\partial p}$ are the steepest descent images. Details about the derivation can be found in [2].

The Hessian matrices He_s,He_p and the steepest descent images $\nabla T_s \frac{\partial W}{\partial s}$ and $\nabla T_p \frac{\partial W}{\partial p}$ can be precomputed because they do not depend on the parameters s, p. Δs and Δp are also respectively independent of the parameter p and s in the Eq.(8) and Eq.(9), so each parameter can be independently computed. Since the coordinates of the corners of the homography s and p are directly employed as the warp parameter in our algorithm described in the Alg. 1, the computation of the warp inversion becomes:

$$\begin{aligned} s &= s - \Delta s \\ p &= p - \Delta p \end{aligned} \tag{12}$$

Algorithm 1 Our projector-based LK algorithm.

Require: T, s, p
  Precompute: $\nabla T_s \frac{\delta W}{\delta s}, \nabla T_p \frac{\delta W}{\delta p}, Hessians\ He_s\ He_p$
  repeat
    $Warp\ I\ with\ W(\mathbf{x}; s)\ to\ compute\ I(W(\mathbf{x}; s))$
    $Compute\ I(W(\mathbf{x}; s)) - T$
    $Compute\ \Delta s\ with\ equation\ 8$
    $s \leftarrow s - \Delta s$
    $Warp\ I\ with\ W(\mathbf{x}; p)\ to\ compute\ I(W(\mathbf{x}; p))$
    $Compute\ I(W(\mathbf{x}; p)) - T$
    $Compute\ \Delta p\ with\ equation\ 9$
    $p \leftarrow p - \Delta p$
  until $residual\ image < threshold$
  return: s, p

3.4. Initialization and update of the homographies

Before running the Lucas-Kanade algorithm, we need to initialize the homographies H_p and H_s. First, we manually define the corners' position of the surface texture in the camera image and compute the homography H_s that transforms it to the template space. Similarly, we manually detect the projected texture in the camera image and compute the homography H_p. The projected texture can then fit the surface texture by computing and applying the homography H:

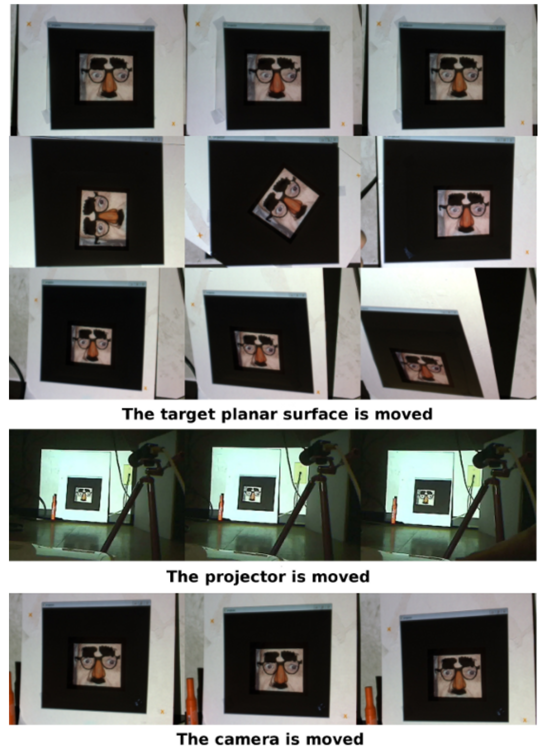

Figure 3. Results of our experiments.

$$H = H_p \times H_s^{-1} \tag{13}$$

Our projector Lucas-Kanade algorithm estimates the homographies H_s and H_p that respectively transforms projected texture defined in the template space to surface texture and the projected texture in the camera space. At the end of this minimization process, the projected texture defined in the template space is transformed by applying the homography H updated as follows:

$$H \leftarrow H_p \times H_s^{-1} \times H, \tag{14}$$

After updating the homography H, the projected pattern will fit the surface texture until either the projector, the camera or the target planar surface are moved.

4. Results

Our system is composed of a LCD projector EPSON EB-X8 and a PGR Flea3 camera with a 640×480 resolution that are set to form an uncalibrated projector-camera system like in Fig.2. The surface texture contains a human face with enough image gradient for

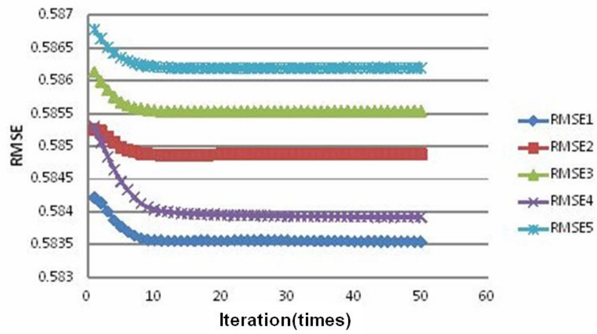

Figure 4. Result of the RMSE for 5 frames.

the tracking. Under this configuration, we achieved 3 frames per second on a Intel Core2 Duo 2.80GHz. Our method has been tested under two situations. First, we translated and rotated the textured target surface in every direction. Second, we moved the projector and camera. Results are presented in Fig. 3 with three different experiments.

The accuracy of our tracking was evaluated with the Root Mean Square Error. RMSE is computed in the region of interest as follows:

$$RMSE = \sqrt{(I - H_p^{-1}T_p - H_s^{-1}{T_s}^2/pix)} \quad (15)$$

where $RMSE$ is the Root Mean Square Error and pix is the number of pixels in the region of interest. The plot of the $RMSE$ is presented in Fig.4. The convergence of the $RMSE$S, computed for five non-consecutive frames, indicates that the homographies are correctly estimated.

5. Conclusions

In this paper we have presented a method to project a pattern on a moving textured planar surface using an uncalibrated projector-camera system. The target planar surface can be freely moved while the projected image will constantly fit the surface even if the projector or the camera are shifted. Our results have shown that the surface and the projector images can be both tracked but not in real-time. We are then currently trying to converting our method for GPGPU to speed-up the process since it can be parallelized.

6. Acknowledgments

We would like to thanks Yuji Oyamada, Julien Pilet, Ismaël Daribo and Bruce Thomas for their support.

References

[1] S. Audet, M. Okutomi, and M. Tanaka. Direct image alignment of projector-camera systems with planar surfaces. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 303–310, 2010.

[2] S. Baker and I. Matthews. Lucas-kanade 20 years on: A unifying framework. *International Journal of Computer Vision*, 56(3):221–255, 2004.

[3] D. Bandyopadhyay, R. Raskar, and H. Fuchs. Dynamic shader lamps: Painting on movable objects. In *International Symposium on Augmented Reality*, pages 207–216, 2001.

[4] H. Bay, T. Tuytelaars, and L. Van Gool. Surf: Speeded up robust features. *European Conference on Computer Vision*, pages 404–417, 2006.

[5] S. Benhimane and E. Malis. Real-time image-based tracking of planes using efficient second-order minimization. In *International Conference on Intelligent Robots and Systems*, volume 1, pages 943–948, 2004.

[6] O. Bimber and R. Raskar. *Spatial augmented reality*. AK Peters, 2005.

[7] D. Cotting, M. Naef, M. Gross, and H. Fuchs. Embedding imperceptible patterns into projected images for simultaneous acquisition and display. In *International Symposium on Mixed and Augmented Reality*, pages 100–109, 2004.

[8] S. Gupta and C. Jaynes. Active pursuit tracking in a projector-camera system with application to augmented reality. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, page 111, 2005.

[9] V. Lepetit, P. Lagger, and P. Fua. Randomized trees for real-time keypoint recognition. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, volume 2, pages 775–781, 2005.

[10] M. Leung, K. Lee, K. Wong, and M. Chang. A projector-based movable hand-held display system. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 1109–1114, 2009.

[11] D. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[12] D. Molyneaux, H. Gellersen, G. Kortuem, and B. Schiele. Cooperative augmentation of smart objects with projector-camera systems. *UbiComp 2007: Ubiquitous Computing*, pages 501–518, 2007.

[13] R. Raskar, J. Van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines. ilamps: geometrically aware and self-configuring projectors. In *ACM SIGGRAPH 2005 Courses*, page 5, 2005.

[14] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 179–188, 1998.