Projector View Synthesis and Virtual Texturing

T. Molinier[1], D. Fofi[1], Joaquim Salvi[2], Y. Fougerolle[1], P. Gorria[1]

[1] Laboratory Le2i UMR CNRS 5158 University of Burgundy, France
E-mail: thierry.molinier@bourgogne.fr
[2] Computer Vision and Robotics Group,University of Girona, Spain

ABSTRACT

We propose a technique to project a synthesized texture onto 3D textureless real objects without contact, which can be applied for example, in medical or archaeological applications. Our system uses two cameras and one projector. By projecting a suitable set of light patterns onto the surface of a 3D real object and by capturing images with one or several cameras, a large number of correspondences can be determined and the 3D geometry can be estimated. Since we deal with textureless objects, traditional corner detection algorithms can not be applied, but our approach successfully determines both 3D geometry and the appropriate texture to be projected onto the 3D model.

Keywords: Structured light, projective texture, augmented reality

1. INTRODUCTION

Surface reconstruction is one of the most investigated topics in computer vision due to its wide field of application and has found numerous application as well as range sensoring, industrial inspection of manufactured parts, reverse engineering, object recognition, and 3D map building. Amongst all the ranging techniques[1], stereovision consists in determining 3D geometry of a scene observed from at least two view points. Triangulation is possible if cameras have been calibrated beforehand[2,3]. However, difficulties in finding the correspondences may arise, even when taking into account epipolar constraints. Coded structured light consists in replacing one of the cameras by a device that projects a light pattern onto the measured surface. The projection of texture onto real objects, that converts a two-dimensional texture or image into a three-dimensional texture by projecting it along one or several directions onto the objects in the scene, has rarely been considered both in computer graphics and computer vision communities, though many potential applications exist, such as virtual reality, augmented reality, clothing designing, virtual museums and movie making. Usually, a texture is only projected onto real objects as a way to alleviate the correspondence problem in surface acquisition by the projection of coded patterns or pseudo random sequences. Few papers dealing with 3D adapted texture (texture that are transformed in order to fit the real 3D object they are projected on) have been published. A double usage of the projector not only allows us to project a pattern to extract points of correspondence, but also to project a texture for augmented reality: both usages are respectively called active and passive usages. Texturing has been widely used to render objects with a synthetic but realistic skin, for many and various applications, such as the modeling of virtual scenarios or the combination of real and synthetic objects in augmented reality. Debevec *et al.*[3] presented a technique to synthesize a new view from one camera with one projector with the VDTM (View Dependant Texture Method). Debevec *et al.* generated novel views of a scene with approximately known geometry, using a sparse set of original views. Unfortunately, this method is only appropriate for models that can be broken into polygonal patches. Raskar *et al.* showed the possibilities of texturing a static object[5], such as the Taj Mahal temple. Another application is the dynamic texture of an object with a pseudo pencil[6]. Here, shape of objects is previously acquired by using any passive or active technique, then the texture is projected by using a DLP projector. The image texture in the projector is synthesized so that once projected, it adequately adapts to the shape of the objects. As a result, one can illuminate the scene, which can be applied in several applications, such as virtual painting (car industry or sculptures in virtual museums). Salvi *et al*[7] presented a review of the techniques related to structured light, where the projector is used to project a coded pattern from which points of interest can be extracted.

The contribution of this paper is the presentation of a complete automatic method to adequately project a texture onto a real 3D object without human intervention, by projecting structured light and by estimating the pose of the real object in the virtual projector frame.

The rest of the paper is as follows: the second section describes the calibration steps, the third section is dedicated to the technique used for extracting the points of correspondence. The fourth section presents our results. We then present our conclusion and future work.

2. SYSTEM PRESENTATION

In this section we present in detail the three steps of our method, that are the sensor calibration, the points of correspondence estimation and the registration of our 3D reconstruction with a scanned model.

We have designed a sensor that is composed of two cameras and one digital projector. This system ensures flexibility and allows both scene capture and its display. Indeed, it is possible to use both cameras to establish correspondences and to triangulate the scene or, alternately, to use the projector and one camera as a structured light system (in the case of blank or poor-textured surfaces). Although the trifocal tensor[8] describes the epipolar geometry between our three views, its estimation is not as robust against noise as the estimation of the fundamental matrix. We therefore estimate the three fundamental matrices for each couple of views.

Raskar *et al.* defined the concept of Imperceptible Structured Light (ISL)[9]. In their early work, they proposed to project, at high frequency rate, a binary pattern followed by its complement so that the resulting pattern forms a uniform lighting. The psychophysical principle of the ISL is based on the optical illusion of flickering during the pattern projection and its complement: below the CFF (Critical Flicker Frequency), human eye perceives the flickering of the two images; beyond this critical frequency, human eye only perceives the fusion of the pattern and its spatial complement, that is a uniform light. In other words, *pattern and complement are visually integrated over time; the result is the appearance of a flat field, or 'white' light*[9]. The encoded images are visible only to cameras synchronized with the projector and exposed for a short time interval. Cotting *et al.*[10] improved the concept of ISL by embedding the structured pattern within a projected image and presented a method with non-intrusive structured light that allowed to simultaneously project a structured light and a texturing light.

2.1. Calibration

In order to precisely project the virtual texture onto the object surface, the hard calibration of the system is required. The intrinsic and extrinsic parameters of the projector must be known to adequately project the texture onto the 3D object. Calibration provides the two camera matrices $P_{c1} \approx K_{c1}[R_{c1} T_{c1}]$, $P_{c2} \approx K_{c2}[R_{c2} T_{c2}]$, and the projector matrix $P_p \approx K_p[R_p T_p]$ represent the pose of the camera (or projector) with respect to the world coordinate frame.

A plane-based calibration method could advantageously be used to calibrate the projector as suggested by Cotting[10]. Here, in order to make the process as simple as possible, we proceed in two steps:

- The cameras are first calibrated by using a plane-based method;
- A pattern is projected onto a blank planar object and captured by the two calibrated cameras. Illuminated points are reconstructed by triangulation in order to ease the correspondence between 2D (projected) and 3D (illuminated) points; Faugeras' method[11] is used to compute intrinsic and extrinsic parameters of the projector.

2.2 The correspondence problem

Extracting points of correspondence can be done by mixing two methods. The epipolar geometry describes the relationship between points of correspondence between two images: the point of correspondence on the second image belongs to the epipolar line of the point on the first image. This relation allows to reduce the research region on the image and the structured pattern allows to extract the point of correspondence on the epipolar line. To improve accuracy, epipolar geometry is estimated by re-using the points previously computed (low noise, no outlier).

2.2.1 Epipolar geometry

The fundamental matrix **F** is computed from the points of correspondence between two images, m_1 and m_2, using:

$$m_1^t \cdot F \cdot m_2 = 0 \,, \qquad (1)$$

The fundamental matrix F is computed for each combination of two views among three. We begin from the view of the projector (structured pattern) because it is easier to recognize the element of the pattern. Armangue and Salvi[12] survey up to nineteen of the most commonly used methods in fundamental matrix estimation. Experimental results show that: (a) linear methods are quite efficient if the points are well located in the image and the corresponding problem is previously solved; (b) iterative methods can cope with some gaussian noise in the localization of points, but are inefficient in the presence of outliers; (c) robust methods can cope with both discrepancy in the localization of points and false matchings. We use LMedS method,that is the most appropriate method for outliers detection and removal, and that still obtains the best results when a low computing time is not required.

2.2.2 Pattern correlation

The point of correspondence must belong to the epipolar line. So the point of correspondence is searched within a band centered around this line. With the coded pattern, the point position on the epipolar line can be estimated. The structured pattern (see Figure 1) is computed so that each element of a cell of five by five pixels is unique. This pattern is built in three steps :

- The first cell, in the left up corner, is a black cell of 5x5 pixels.
- A column of 5 pixels is selected randomly between the combination of all possible columns, then fills in the pattern. If each cell of five by five pixels is unique, another column is built, otherwise the column is changed.
- A line of 5 pixels is added, using a similar method. Then each new pixel of the line of the pattern must be filled in one by one. This step is repeated until the pattern is completed.


Figure 1 : Pattern with two unique and superimposed cells

Each cell of the pattern is unique, so if we can identify the same cell on two images, two points of correspondence are indeed identified. We seek for this cell within the region of interest defined by the epipolar geometry. There exist several methods to estimate the correlation between the cell of the point of the pattern and those on the region of interest[13] based on energy filter. These filters are composed of a set of masks that define some textural properties of the image. We directly use the comparison between the two cells (equation 2) of the two images, A and B . Generally, the point such as we find the maximum of correlation must be the point of correspondence.

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\sum_m (A_{mn} - \bar{A})^2 \sum_n (B_{mn} - \bar{B})^2}}, \qquad (2)$$

where $\bar{A}$ and $\bar{B}$ are the mean of the images A and B .

2.2.3 The outliers

Outliers are a critical problem. We have to choose parameters that eliminate most of the outliers to improve the next step of the method. Any error in the registration step will result in a bad visual result and the virtual texture will not fit the real object. Thresholds must be determined to solve this problem. In practice we set a threshold to 0.8. The second parameter is the epipolar distance that is the distance between a point and its epipolar line. If this distance is higher than a threshold of one pixel (common precision of the LMedS method), the point is considered as an outlier.

2.3 Registration

Once the points of correspondence between the views are estimated, the triangulation can be done through the calibration of the system. It appears that the necessary one shot technique of structured light we used does not allow us to estimate enough 3D points to adequately texture the model. To improve the rendering, we use a scanned set of points of the real model. We register our 3D reconstruction with the scanned model. Once the registration is realized, the last step is to synthesize the view of the projector. We now know both the pose of the real object in the projector frame, as well as the extrinsic and intrinsic parameters of the projector. The projection of the cloud of points is then realized following the equation:

$$\boldsymbol{m} = \boldsymbol{K}_p \cdot [\boldsymbol{R}_p \boldsymbol{t}_p] \cdot \boldsymbol{M} , \qquad (3)$$

where m and M are the 2D and 3D points.

3. RESULTS

a) Textureless model b) Model with real texture

c) Model with virtual teture d) Another virtual texture

Figure 2 : Model with or without texture

The Figure 2 shows the results of our method: Figure 2.b presents the original model (with real and permanent texture), Figure 2.a presents a textureless model. Figures 2.c and 2.d show the textureless model with a projected virtual texture. Figure 2.d shows a virtual texture similar to the real one: mouth, nose as well as details on the hears texture are correctly located on the model. Once the pose in the projector frame is estimated, another texture can be projected as illustrated in Figure 2.c.

4. CONCLUSIONS AND FUTURE WORK

We have presented an automatic method to adequately project a virtual texture onto a real textureless 3D object. Hard calibration, points of correspondence extraction, and registration are the three main steps of our method. The correspondence problem can not be solved in our case by using common algorithms. We overcome this problem with structured light by creating a new pattern for our purpose. This method allows us to adequately and automatically project virtual textures onto real 3D object.

Current work focus on improving the extraction of the point of correspondence to reach a sub-pixel precision on point position. A precise and automatic capture of a cell of the projected structured pattern, as well as another color pattern are actually tested and show preliminary promising results. Future work will focus on virtual oriented texture: for example one may think about projecting brick or wood pattern, or any other arbitrary texture, onto not yet painted buildings to see how they really fit their environment.

1. R. Jarvis, “Range sensing for compute vision, advances in image communications, elsevier sciences,” 1993.
2. Olivier Faugeras, *Three-dimensional computer vision:a geometric viewpoint*, MIT Press, Cambridge, MA, USA, 1993.
3. J. Salvi, X. Armangue, and J. Batlle, “A comparative review of camera calibrating methods with accuracy evaluation,” 2002.
4. Paul Debevec, Yizhou Yu, and George Boshokov, “Efficient view-dependent image-based rendering with projective texture-mapping,” Tech. Rep. CSD-98-1003, 20, 1998.
5. Ramesh Raskar, Greg Welch, Kok-Lim Low, and Deepak Bandyopadhyay, “Shader lamps: Animating real objects with image-based illumination,” in *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, London, UK, 2001, pp. 89–102, Springer-Verlag.
6. “Dynamic shader lamps: Painting on movable objects,” in *ISAR ’01: Proceedings of the IEEE and ACM International Symposium on Augmented Reality (ISAR’01)*, Washington, DC, USA, 2001, p. 207, IEEE Computer Society.
7. J. Batlle. J. Salvi, J. Pags, “Pattern codification strategies in structured light systems,” in *Pattern Recognition*, April 2004, pp. 37(4), pp 827–849.
8. C. Ressl, *Geometry, Constraints and Computation of the Trifocal Tensor*, Ph.D. thesis, Institut fr Photogrammetrie und Fernerkundung, Germany.
9. M. Cutts A. Lake L. Stesin R. Raskar, G. Welch and H. Fuchs, “The office of the future: A unified approach to image-based modeling and spatially immersive displays,” *Computer Graphics*, vol. 32, pp. 179–188, 1998.
10. Gross M. Cotting D., Naef M. and Fuchs H., “Embedding imperceptible patterns into projected images for simultaneous acquisition and display,” in *ISMAR ’04: Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2004, pp. 100–109.
11. O. D. Faugeras and G. Toscani., “The calibration problem for stereo,” in *In Proc. Conf. Computer Vision and Pattern Recognition, Miami, FL,*, June 1986, pp. 15–20.
12. X. Armangue and J. Salvi, “Overall view regarding fundamental matrix estimation, image and vision computing,” February 2003, vol. 21, pp. 205–220.
13. Rafael Garcia Campos, *A proposal to estimate the motion of an underwater vehicle through visual mosaicking*, Ph.D. thesis, IA Electrnica, Informtica i Automtica, University of girona, Spain, 2001