VIDEO MOSAICING FOR CURVED SURFACE BY 3-D RECONSTRUCTION USING FEATURE POINTS

Akihiko IKETANI[1] *Tomokazu SATO*[1, 2] *Sei IKEDA*[2] *Masayuki KANBARA*[1, 2]
Noboru NAKAJIMA[1] *Naokazu YOKOYA*[1, 2]

[1] Internet System Laboratories, NEC Corporation
[2] Graduate School of Information Science, Nara Institute of Science and Technology

1. INTRODUCTION

Video mosaicing methods have been widely used for capturing high resolution document images with low resolution video cameras [1, 2]. In these methods, given a video where partial images of the target document are captured in each frame, pairwise registrations between two successive images are estimated, then, by warping all the images to a reference frame (in general, the first frame), a high resolution document image called a mosaic image is synthesized. These methods, however, assume that the target is a plane, thus will not work for documents with curved surface. An example of a document with a curved surface is shown in Figure 1.

We propose a novel video mosaicing method for curved surface based on 3-D reconstruction. With the proposed method, the plane mosaic image of the target document is generated, even if the document has a curved surface. We have developed a demonstration system based on this method.

In the following sections, we briefly describe the proposed method, and then introduce our demonstration system.

Figure1. A target document with curved surface.

2. VIDEO MOSAICING FOR CURVED SURFACE

In the proposed method, first, the shape of the target document is estimated by 3-D reconstruction, and then a mosaic image is generated by mapping the texture of the target onto a plane surface. The assumptions made on the target are that the curve lies only along one axis, and the curvature changes smoothly along this axis. It is also assumed that the intrinsic camera parameter is already known, and does not change throughout the video capturing. The flow of the proposed method is as follows.

(1) 3-D Reconstruction by Feature Tracking

Image features in the input video are automatically tracked, and 3-D positions of the feature and the extrinsic camera parameters are estimated from the tracking result [3]. Here, the initial 3-D positions of image features are calculated, assuming that the image plane in the initial frame is parallel to the target surface. Using this assumption, 3-D reconstruction can be carried out automatically, without any need for manually giving the initial positions of the image features.

(2) Target Shape Estimation by Surface Fitting

For each feature point obtained by (1), a quadratic surface is locally fitted to features lying within a certain distance from itself, and the minimum principal curvature direction of the target is estimated. Next, 2-D coordinate $(\hat{x}, \hat{y})$ is calculated for each feature point by projecting its 3-D position (x, y, z) to a plane perpendicular to the minimum principal curvature direction. Finally, the target shape parameter $(a_0, \cdots, a_m)$ is obtained by fitting the following polynomial equation to the projected coordinates $(\hat{x}, \hat{y})$.

$$\hat{y} = f(\hat{x}) = \sum_{i=0}^{m} a_i \hat{x}^i \ .$$

Here, the optimal number of order m is automatically determined by geometric AIC [4].

(3) Detection of Re-appearing Features and Global Optimization

For each feature point, the image pattern around the feature position is projected onto the fitted surface to eliminate the distortion in the observed pattern induced by the curved shape of the target or the posture of the camera. Next, for every feature pair whose distance is less than a given threshold, the similarity of the patterns is examined by normalized cross correlation, and pairs whose patterns are similar enough are linked as re-appearing features. Finally, 3-D reconstruction result is globally optimized by minimizing the sum of re-projection errors for every feature.

(4) Mosaic Image Generation

In this process, a mosaic image is generated by mapping the texture on the curved surface onto a mosaic image plane. The relation between the 2-D coordinate on the mosaic image plane (u, v) and the 3-D coordinate on the fitted polynomial surface $(\hat{x}, \hat{y}, \hat{z})$ is given as follows:

$$u = \int_0^{\hat{x}} \sqrt{1 + \{\frac{d}{dx} f(x)\}^2}\, dx, \quad v = \hat{z}.$$

Using the equation, the texture on the curved surface is expanded on a mosaic image plane.

3. DEMONSTRATION SYSTEM

We have developed a demonstration system which consists of a desktop PC (Pentium-4 3.2GHz, Memory 2GB) and a calibrated IEEE1394 CCD camera (Aplux C104T).

Experiments were done using this system on a curved document shown in Figure 1. As shown is Figure 2, the document was captured as 640×480 images of 200 frames at 15 fps. The proposed method was applied to this video to generate a plane mosaic image. The estimated shape and the generated plane mosaic image is shown in Figure 3 and 4, respectively. As can be seen, our system is capable of generating a plane mosaic image for a curved document.

The processing time was approximately 20 seconds for 3-D reconstruction by feature tracking, and 58 seconds for shape estimation, global optimization, and mosaic image generation.

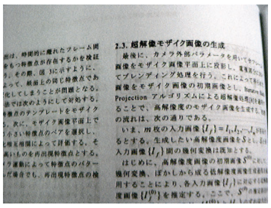

Figure 2. Sampled frame of input image sequence.

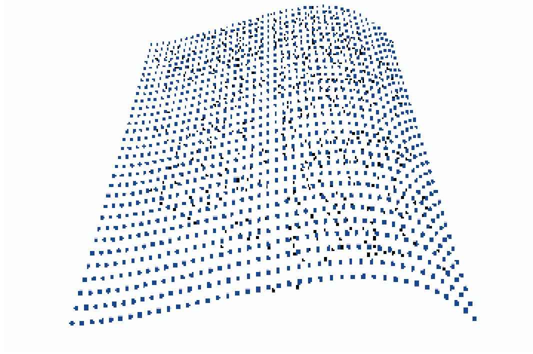

Figure 3. Estimated target shape.

Figure 4. Generated plane mosaic image.

4. CONCLUSION

A novel video mosaicing method for curved documents based on 3-D reconstruction was proposed. The proposed method is capable of generating a plane mosaic image by first estimating the shape of the target, and then mapping the texture on the curved surface onto the mosaic image plane. We have developed a demonstration system based on this method. An experiment on a curved document shows the feasibility of our system. Our future work is to quantitatively evaluate our method, and to extend this method to deal with more complex shaped target.

REFERENCES

[1] Szeliski, “Image Mosaicing for Tele-Reality Applications,” Proc. IEEE Workshop on Applications of Computer Vision, pp. 230–236, 1994.

[2] Capel and Zisserman, “Automated Mosaicing with Super-resolution Zoom,” Proc. CVPR, pp.885-891, 1998.

[3] Sato, Kanbara, Yokoya and Takemura, “Dense 3-D Reconstruction of an Outdoor Scene by Hundreds-baseline Stereo Using a Hand-held Video Camera,” Int. J. on Computer Vision, vol. 47, no. 1-3, pp. 119-129, 2002.

[4] Kanatani, “Geometric Information Criterion for Model Selection,” Int. J. on Computer Vision, vol. 26, no. 3, pp. 171-189, 1998.