ISMR 2001 in Yokohama, Japan, pp.182-183 (2001)

Panorama-based annotation overlay with cylindrical transform

Masakatsu Kourogi† Takeshi Kurata‡ Katsuhiko Sakaue‡ Yoichi Muraoka†

†Graduate School of Science and Engineering, Waseda University,
3–4–1 Okubo, Shinjuku, Tokyo, 169–8555, JAPAN
Phone: +81–3–3209–5198, E-mail: {kourogi,muraoka}@muraoka.info.waseda.ac.jp
‡Electrotechnical Laboratory, 1–1–4 Umezono, Tsukuba, Ibaraki, 305–8568, JAPAN
Phone: +81–298–61–5789, E-mail: {kurata,sakaue}@etl.go.jp

Abstract

In this paper, we propose an improved method of annotation overlay on live video which is one of the most promising applications in augmented reality and wearable computer. In our previous work, we proposed the panorama-based method of annotation overlay which aligns an input video frame and a panoramic image using affine transform. However, affine transform is generally not capable of image registration between a frame and a panorama. We improved the previous method so that it can estimate projective transform parameters without severely increasing computational cost. We confirmed that improved method could estimate image registration parameters under conditions that hindered the previous method. Its computational cost increased only by 10–20% and its real-time processing with software implementation was achieved.
keywords : annotation overlay, panoramic image, wearable vision system

1 Introduction

Annotation overlay on a live video is an essential feature of augmented reality (AR), since it makes possible various kinds of such applications as augmented memory, touring assistance, and amusements[1].

In our previous work[3], we proposed a method for annotation overlay which used a set of panoramic images captured at various points of environment and annotations manually attached to them, as a prior knowledge. The method overlaid the annotations according to the image alignment using affine transform parameters between the input video frame and the panoramic images.

However, it is generally impossible to align video frames with a panoramic image using affine transform. Therefore, the previous work is applicable only if the user's view angle is not apart from the horizontal direction.

To handle the limitations, we propose an improved method of image alignment between frames and panoramas. The method firstly transforms the frames to cylindrical plane using the multiple assumptions about view angles of elevation of the frame. Next, the method finds affine parameters between the transformed frames and the panorama for each assumption, and selects the best result that gives the smallest mean square error of image alignment.

Experimental results show that proposed method stably finds the parameters while the previous method has difficulties to estimate affine parameters for image alignment. The computational cost required for the method increases only 10–20% and its real-time processing can be achieved with our software implementation.

2 Image alignment with cylindrical transform

If the frame is transformed to cylindrical plane, it is possible to align the transformed frame with a panoramic image using affine transform. The proposed method assumes that the center of the frame is mapped to one of a set of several points on the panorama, which are chosen so that they can cover the entire panorama or cover the neighborhood of the previous result.

Let the view angle of elevation of the frame be α, the focal length of the panorama be f_P, and one of the chosen point on the panorama be (x_P, y_P), the following equation can be obtained.

$$\alpha = \tan^{-1}\left(\frac{y_P}{f_P}\right). \quad (1)$$

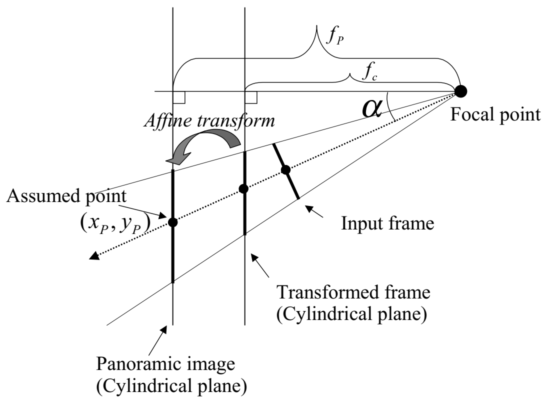

Figure 1: Cylindrical transform.

The point (x, y) on the frame is firstly mapped to (x', y') by the vertical rotational transform using the view angle α.

$$x' = \frac{f_c x}{y \sin\alpha + f_c \cos\alpha} \quad (2)$$

$$y' = \frac{f_c(y\cos\alpha - f_c \sin\alpha)}{y \sin\alpha + f_c \cos\alpha}, \quad (3)$$

where f_c is the focal length of the camera.

Next, the mapped point (x', y') is transformed to (x'', y'') on the cylindrical plane whose radius is equal to the focal length f_c using the following equation,

$$x'' = f_c \tan^{-1}\frac{x'}{f_c} \quad (4)$$

$$y'' = f_c \cos\left(\tan^{-1}\frac{y'}{f_c}\right). \quad (5)$$

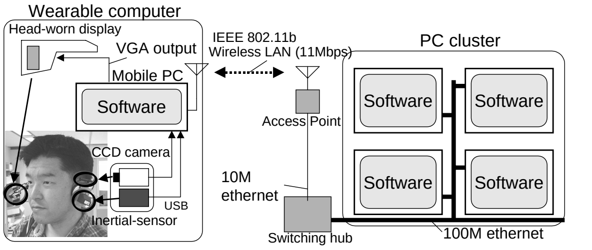

Figure 2: Diagram of our wearable vision system.

Once the frame is transformed to the cylindrical plane, we estimate affine transform parameters between the transformed frame and the panorama.

The diagram of the proposed method can be summarized as follows.

1. Select points on the panoramic image to which the center of the current frame is supposed to be mapped so that they can cover the entire portion of the panorama or the neighborhood of the previous result.
2. (if not at first) Cancel the rotational component around the optical axis using the affine parameters between transformed frame and the panoramic image.
3. Compute view angles α of the current frame from the corresponded points selected in stage 1.
4. Transform the frame to cylindrical plane using the computed angle α.
5. Estimate affine parameters for image alignment between the transformed frame and the panoramic image.
6. Compute measure square error of image brightness using the estimated parameters and select the best result as an output.

3 Implementation

We implemented the proposed method on our developing *wearable vision system*[3], consisting of a wearable computer (OS: Windows 98, CPU: Mobile PentiumIII-500MHz) equipped with a head-worn display (MicroOptical, Clip-on display), a CCD camera (Toshiba, IK-SM43H), a 3-DOF motion sensor with gyro-sensors, accelerometers and compasses (Tokin, MS3D-U7) which claims to have resolution of ± 1.0 degree at 125Hz and a wireless LAN card complied with 11-Mbps IEEE 802.11b, and a remote PC cluster consisting of 5 PCs (OS: Linux-2.2.14 SMP supported, CPU: Dual PentiumIII-500MHz) as shown in Figure 2. This implementation also works for other wearable vision applications [4].

4 Experiments

We evaluate the proposed method by both the off-line and on-line experiments described below. In the off-line experiment, we capture the video frames (size: 320×240, 100 frames) and a panoramic image (size: 2560×480), and we manually give the correspondence of 10 points from the frames to the panorama. We evaluate the method by the error of image alignment by using the ground-truth correspondence data. The error of image alignment by the proposed method and by the previous method (affine transform) are shown in Figure 4, respectively. The results show that the error of image alignment is improved by 3–10 pixel. The improvement is eminent when the view angle is apart from the horizontal direction (frame number: 50–80).

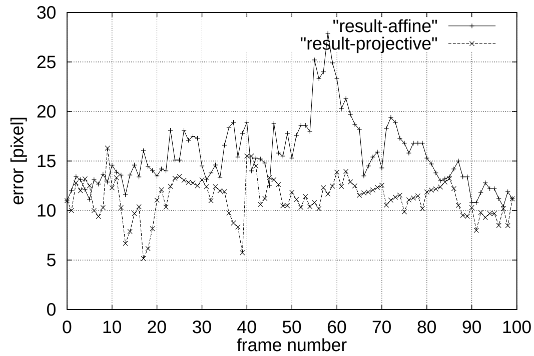

Figure 3: Error of image alignment.

In the on-line experiment, we use the software implementation of the method described in Section 3. The computational cost is increased by only 10–15% and its real-time processing (10–15 frames/s) can be achieved. The results show that the proposed method stably provides video frames overlaid with the annotations. Some of the output examples are shown in Figure 4.

Figure 4: Output frames overlaid with annotations.

Acknowledgements

This work was conducted as a part of the Real World Computing (RWC) Program.

References

[1] J. Rekimoto and K. Nagao, "The world through the computer: Computer augmented interaction with real world environments", in *Proc. of UIST '95*, pp. 29–36, 1995.

[2] T. Starner et al., "Augmented reality through wearable computing," in *Presence*: Teleoperators and Virtual Environments, vol. 6, no. 4, pp. 386–398, 1997.

[3] M. Kourogi et al., "A real-time panorama-based technique for annotation overlay on video frames," in *Proc. of ICME'2000*, TA. 2.05, 2000.

[4] T. Kurata et al., "The Hand-mouse: A Human Interface Suitable for Augmented Reality Environments Enabled by Visual Wearables," in *Proc. of ISMR'2001* (To appear), 2001.