TUPM 6.8

IMAGE MOSAICING WITH AUTOMATIC SCENE SEGMENTATION FOR VIDEO INDEXING

Yun-Hee Choi, Yeong Kyeong Seong and Tae-Sun Choi
Dept. of Mechatronics, Kwangju Institute of Science and Technology (K-JIST)
1 Oryong-dong, Buk-gu, Gwangju 500-712, Rupublic of Korea, dals@sipl.kjist.ac.kr

Abstract

In this paper, an image mosaicing algorithm for video indexing is presented with efficient and robust automatic scene segmentation. Phase correlation and motion compensation are used for image mosaicing and scene segmentation. Simulation results show that the proposed algorithm is fast, robust and efficient for scene-based video indexing.

I. Introduction

As the mount of multimedia data is increasing, browsing of videos in the multimedia database becomes important. Conventional frame-based video representation is not appropriate for efficient access to information of interest. Therefore, Scene-based video representations using image mosaicing for video indexing have been proposed recently [1]. Shot boundary detection is indispensable for the image mosaicing as one video sequence is composed of many different subsequences. In general, the first step for scene analysis is temporal video segmentation. Many shot boundary detection algorithms have also been proposed in last decades. Most of shot boundary detection algorithms can be classified into histogram-based, motion-based and MPEG compressed domain-based methods. Histogram-based methods find the difference between histograms of consecutive frames, and large frame differences are considered as possible shot changes. Most of MPEG compression domain-based methods utilized DCT coefficients.

Conventionally, video segmentation and image mosaicing have been developed independently. The separate works for image mosaicing and video segmentation are not an efficient way if there exist some novel methods for image mosaicing with simultaneous video segmentation. In this paper, an image mosaicing algorithm is proposed with efficient and robust automatic scene segmentation. Phase correlation and motion compensation are used for image mosaicing and scene segmentation.

II. Global Motion Estimation and Mosaic Image

1. Global Motion Estimation by Phase Correlation

The phase correlation is based on the Fourier shift theorem, *i.e.*, a shift in the spatial domain is equivalent to a phase shift in the frequency domain.

Let an image $f_2(x,y)$ be a shifted replica of image $f_1(x,y)$ displaced by (x_0, y_0), *i.e.*,

$$f_2(x,y) = f_1(x - x_0, y - y_0), \tag{1}$$

Then, the cross-power spectrum of two images $f_1(x,y)$, $f_2(x,y)$ is defined as:

$$\frac{F_1(\omega_x,\omega_y)F_2^*(\omega_x,\omega_y)}{|F_1(\omega_x,\omega_y)F_2(\omega_x,\omega_y)|} = e^{j(\omega_x x_0 + \omega_y y_0)}. \tag{2}$$

It is obvious from equation (2) that the inverse Fourier transform of equation (2) gives displacement vector (x_0, y_0) of spatial domain.

2. Mosaic Construction

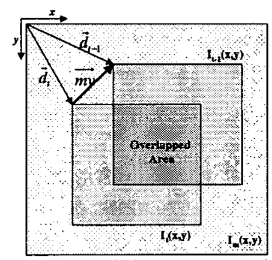

Figure 1. Mosaic image construction

The mosaic image construction procedure is depicted in Figure 1. If the image I_{t-1} is located at $\vec{d}_{t-1}$ and global motion vector is $\overrightarrow{mv}$, $\vec{d}_t$ and $\vec{d}_{t-1}$ are related as:

$$\vec{d}_t = \vec{d}_{t-1} - \overrightarrow{mv}. \tag{3}$$

In case of static mosaic [2], all frame locations, which compose one scene, must be identified on the mosaic coordinates to construct a static mosaic image at a time. When one or both of the components of $\vec{d}_t = (d_{tx}, d_{ty})$ from equation (3) is negative, all $\vec{d}_{t-1}, \vec{d}_{t-2}, \cdots, \vec{d}_0$ must be recalculated, because the origin of mosaic plane is changed. The updated $\vec{d}_{t-1}, \vec{d}_{t-2}, \cdots, \vec{d}_0$ are calculated as

$$\vec{d}_{new\_l} = \vec{d}_l - neg(\vec{d}_t), \text{ where } l = t, t-1, t-2, \ldots, 0, \tag{4}$$

and

$$neg(\vec{v}) = \begin{cases} (0,0), & v_x \ge 0, v_y \ge 0 \\ (v_x,0), & v_x < 0, v_y \ge 0 \\ (0,v_y), & v_x \ge 0, v_y < 0 \\ (v_x,v_y), & v_x < 0, v_y < 0 \end{cases}$$

0-7803-7300-6/02 $17.00 © 2002 IEEE

The total mosaic image size at time t is

$$MW_t = d_{(t-1)x} - mv_x + MW_{t-1}$$
$$MH_t = d_{(t-1)y} - mv_y + MH_{t-1}, \quad (5)$$

where, MW_t, MH_t are mosaic image width and height at time t respectively.

For the dynamic mosaicing [2], a previous mosaic image must be aligned to the current mosaic image. The previous mosaic image position on the current mosaic plane must be defined.

$$dm_x = \begin{cases} mv_x - d_{(t-1)x}, & d_{(t-1)x} < mv_x \\ 0, & \text{otherwise} \end{cases}$$
$$dm_y = \begin{cases} mv_y - d_{(t-1)y}, & d_{(t-1)y} < mv_y \\ 0, & \text{otherwise} \end{cases}, \quad (6)$$

where, dm_x and dm_y are x, y coordinates of a previous mosaic image in the current mosaic plane respectively.

III. Cut Detection

When peak value of correlation surface is less than a certain threshold value σ_1, shot boundary is suspected. In this case shot boundary is declared according to the following criteria.

$$\eta = \frac{1}{A} \sum_{(x,y)\in A} [I_t(x,y) - I_{t-1}(x,y)]^2, \quad (7)$$

where, A denote the overlapped area in Figure 1.

If η is larger than a certain threshold value σ_2, shot boundary is occurred between the video frames I_{t-1} and I_t.

IV. Simulation Results and Discussion

For the simulation of proposed algorithm, a test sequence of 340 frames was made. The sequence is composed of 50 frames of coast guard sequence (standard test sequence with moving objects: 51 frame to 100), 150 frames of airplane sequence (taken by author) and 140 frames of scenery sequence (taken by author).

Figure 2-(a) shows the static mosaic image of coast guard sequence. It can be seen that the moving objects (two boats) are blurred out because the pixels on the moving object area are averaged with different intensity values as object moves. Figure 2-(b) shows the dynamic mosaic image of same sequence. Each mosaic image is updated with the most recent frame, thus, moving objects do not blur out in the dynamic mosaic image as seen in Figure 2-(c). Figure 2-(c) and 2-(d) show the next two synopsis images of successive subsequences, represented by a dynamic mosaic image.

Test sequence has two abrupt scene change points. One is between frame 50 and 51, the other is between frame 200 and 201. Figure 3-(a) shows the peak value distribution of phase correlation surface of each frame. It can be seen from the Figure 3-(a) that frame 51 and 201 are the candidate for the possible cut points. It is obvious from the Figure 3-(b) that frame 51 and 201 are shot boundary points. The threshold values used in simulations are σ_1=0.03 [4] and σ_2=3500.

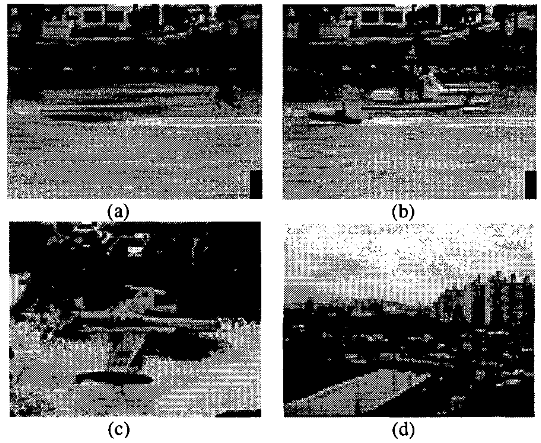

Figure 2: (a) Static mosaic image of coast guard (51 to 100 frames), (b) Dynamic mosaic image of coast guard, (c) Dynamic mosaic image of airplane, (d) Dynamic mosaic image of scenery.

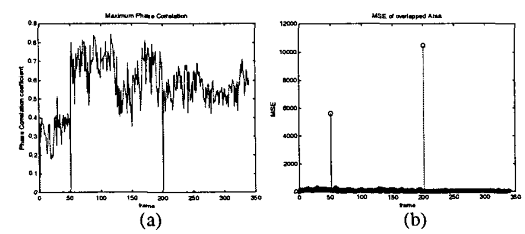

Figure 3: (a) Maximum values of phase correlation surface, (b) MSE of overlapped area.

V. Conclusions

In this paper, a new algorithm for image mosaicing was proposed with automatic scene segmentation. A phase correlation is usually used for image registration. Global motion-compensated cut detection algorithm was used along with phase correlation peaks. Simulation results showed that the proposed algorithm is fast and robust.

VI. References

[1] M. Irani and P. Anandan, "Video indexing based on mosaic representations," *Proceeding of the IEEE*, vol. 86, no.5, pp. 905-921, May 1998.

[2] M. Irani, P. Anandan, and S. Hsu, "Mosaic Based Representations of Videos Sequences and Their Applications," *Proc. of IEEE International Conference on Computer Vision*, pp. 605-611, June 1995.

[3] Q. S. Chen, M. Defrise, and Deconinck, "Symmetric phase-only mached filtering of Fourier-Mellin transform for image registration and recognition," *IEEE Trans. On PAMI*, vol. 16, no. 12, pp. 1156-1168, December 1994.

[4] B. S. Reddy and B. N. Chatterji, "An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration," *IEEE Trans. On Image Processing*, vol. 5, no. 8, pp. 1266-1271, August 1996.