International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:2, 2007

Automated Feature Points Management for Video Mosaic Construction

Jing Li , Quan Pan, Stan. Z. Li , Tao Yang

***Abstract*—** A novel algorithm for construct a seamless video mosaic of the entire panorama continuously by automatically analyzing and managing feature points, including management of quantity and quality, from the sequence is presented. Since a video contains significant redundancy, so that not all consecutive video images are required to create a mosaic. Only some key images need to be selected. Meanwhile, feature-based methods for mosaicing rely on correction of feature points' correspondence deeply, and if the key images have large frame interval, the mosaic will often be interrupted by the scarcity of corresponding feature points. A unique character of the method is its ability to handle all the problems above in video mosaicing. Experiments have been performed under various conditions, the results show that our method could achieve fast and accurate video mosaic construction.

***Keywords*—**video mosaic, feature points management, homography estimation.

I. Introduction

It is well-known that the imaging areas of current CCD devices are way smaller than that of traditional film cameras with consequent view size reduction. Meanwhile, an observer watching this video of CCD must pay constant attention, as objects of interest move rapidly in and out of the camera field of view. The video also lacks a larger visual context—the observer has difficulty perceiving the relative locations of objects seen at one point in time to objects seen moments before. Then, video mosaicing is called for to produce undistorted large view images by means of digital cameras [1]. It is an active area of research and the techniques have various applications such as satellite photographs, video surveillance, stabilization, compression, virtual environments, virtual travels, 3D world scene medical imaging, etc [2,3,4].

Numerous techniques have been approved for image mosaicing which can be classified broadly into: direct methods[5,6] and feature-based methods[7,8,9,10].Direct methods use information from all pixels and discover parameter set often through an iterative process to minimize the sum of squared difference(SSD). However, these methods require good guesses for the parameters of the transform be given as initial values to the program. If they are not correspond to physical movements of the camera, it is very difficult to evaluate the these parameters value. On the other hand, feature-based methods assume that feature correspondences between image pairs are available, and utilize these correspondences to find transforms which register the image pairs. A major difficulty of these methods is the acquisition and tracking of image features. Good features are often hand-selected, and reliability of feature tracking is often a problem due to image noise and occlusion [11].

This paper describes a feature-based mosaic construction algorithm through managing the feature points during the video to achieve highly accurate results while maintaining low computational costs. The algorithm uses a feature selector and tracker previously to obtain corresponding pairs between images. The management procedure has been signed to supervise these feature points to enhance the speed and accuracy of the mosaicing process.

The remainder of this paper is organized as follows. Section 2 gives an outline of the algorithms. Section 3 describes the feature management in detail including quantity management and quality management. In Section 4, we show real image examples to demonstrate that the management of feature points allows robust image mosaicing. Section 5 contains the conclusion and future extensions.

II. Outline of the Algorithm

The block diagram of the algorithm is shown in Figure 1. The aim of the method is to automatically, fast and robustly create a panoramic image from a video sequence through feature points selecting, tracking and managing. There are four modules for the whole flow: 1) Feature points selection, 2) Feature points tracking, 3) Feature points quantity management, and 4) Feature points quality management. These four modules and some conditions are interacted to build up the whole flow.

The proposed feature-based mosaicing method consists of four steps: (i) Suppose we detect N feature points on the initial frame of the video and set this frame as the reference frame; (ii) Tracking these N feature points frame-to-frame until condition One is broken, and set the current frame as a key frame for mosaicing; (iii) Send the reference image, the key image and their corresponding feature points into the quality feature points management module to obtain the most accurate mosaic, according to the condition Two; (iv) Updating the key frame to the new reference image and rereselect N feature points on it.

Jing Li is with the College of Automatic Control, Northwestern Polytechnical University. (phone: 029-88495954-802; fax: 029-88494352; e-mail:jinglinwpu@163.com).

Quan Pan is with the College of Automatic Control, Northwestern Polytechnical University (e-mail: quanpan@nwpu.edu.cn).

Stan.Z.Li is with National Lab of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences (e-mail: szli@nlpr.ia.ac.cn).

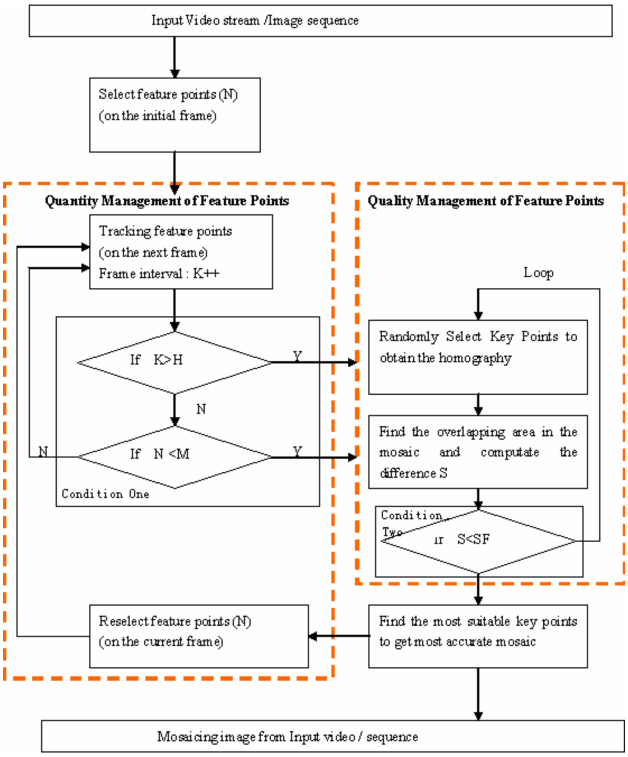


Fig 1. The block diagram of the algorithm.

III. AUTOMATIC FEATURE POINTS MANAGEMENT

This section briefly describes the procedure for managing feature points during the mosaicing. In the feature-based approach, by using corresponding feature points between images, we obtain the translation matrix for mosaicing. However, this kind of method depends mainly on the correction of the corresponding of these feature points. Once the condition is not good, it will influence the mosaicing result greatly. Thus, many researchers have done a lot work on better feature correspondence and have well performance. However, in many situations, we often meet problems as following, and only relying on good feature correspondence can not solve these problems very well.

Problem Statement:

（1） Since a video contains significant redundancy, so that not all consecutive video images are required to create a mosaic. In order to meet the real time demand in many systems, only some key images are selected to create a mosaic. Some methods will give a frame interval such as twenty or thirty to decide the key images. However, in many situations, this parameter is not suitable. How to decide these key images automatically is a different problem;

（2） Feature-based methods always have the assumption that the corresponding points are tracked correctly enough, however, if error corresponding points appearance, or the result is not very precise, it will deeply influence the performance of mosaic. Furthermore, given that all correspondences are correct enough, choose different pairs of feature points may result in different accurate of mosaic. So how to choose the most suitable pairs of feature points is also a problem.

To solve thses problems above, in this paper, we pay more attention to concerning the quantity and quality of feature points during mosaicing. Through managing these feature points, we get the useful feedback to control the whole processing, obtain the mosaic more fast, robust and accurate. In the following subsections, we will describe in detail how to manage these feature points.

A. Quantity management

Existing feature based methods often filter the sequence images to extract feature points and try to find matches between sets of points, then get the homography of the images. The mosaic performance is depending on the corresponding correction very deeply. So in order to increase the corresponding correction rate, track feature points frame-to-frame is the best method. While on the other hand, as we all know, video frames are typically 30 frames a second then this long sequence contains significant redundancy. In order to meet the real time constrained for the whole system, as the first step, we take some measures to identify key frames which includes enough and effective information for later mosaicing.

To solve this problem, we propose a quantity management for feature points. Here, through managing the number of matched feature points frame-to-frame, we decide which two frames are the key frames.

1. Extracting N feature points on the initial frame F1 of the video;
2. Tracking these N points frame-to-frame until M points have been matched, and the current frame is F2. M is the threshold for reselecting new feature points;
3. Once the F2 is reached, we decide the F1 and F2 as the key frames which show the enough information including the sequence F1~F2.
4. Using these two key frames F1、F2 and their matched feature points M to calculate the transform matrix for mosaicing construction.

In this paper , we select feature points according to the gradient value , meanwhile in order to ensure the accurate of the mosaic construction, we input F1、F2 and their feature points M into the quality management module which be described in next section.

Furthermore, considering that if some strong feature points occur in a video and they are almost matched very well during many frames. The step two above will be very difficult to reach, and thus the interval between F2 and F1 are very large. In order to ensure that the key frames remain enough information, we should give a upper limitation H for the frame interval. If the frame interval between the two key frames is larger than the H, we should select these two frames as the key frames. Here, we also proposed a general idea to identify the upper limitation H in different situation.

As noted earlier, video frames are typically 30 frames a second then this long sequence contains significant redundancy. The general idea is shown in Figure 2.

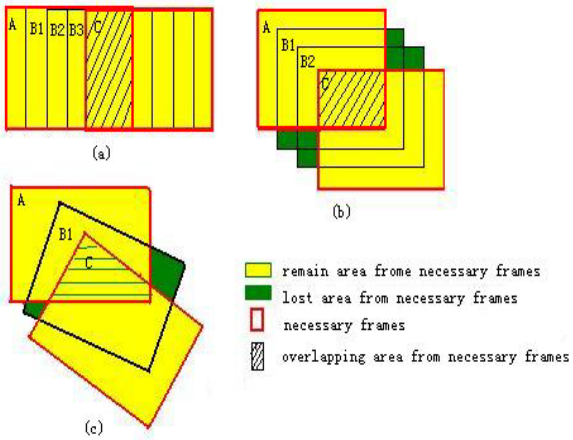


Fig 2. The key frames selection

Given there are some successive frames from a video: $A, B_1 \dots B_n, C$, the yellow block shows remain area from key frames, and the green block means lost area. The shadow represents the overlapping area from key frames. Figure 2 (a) (b) (c) shows three different moving model of camera.

The general idea we proposed is that the overlapping area between two key frames should remain enough information for this scene , meanwhile the lost information must be smaller than a limit. Equation (1) and (2) are used to define the conditions.

$$Overlap\_Area > P_{overlap} \tag{1}$$

$$Lost\_Area < P_{lost} \tag{2}$$

Where $P_{overlap}$ is the threshold which is used to make sure that the two key frames' overlap area is enough large. This parameter is often set to 30%. And at the same time, the lost information should not be too large to influence the final mosaic a lot. Therefore, another parameter P_{lost} is also very useful. This threshold P_{lost} is often been set to 10%.

Suppose we decide the frame A and C as the key frames, then the overlap_Area and Lost_Area are defined as follows:

$$Overlap\_Area = \frac{Area(A \cap C)}{Area(oneframe)} \tag{3}$$

$$Lost\_Area = \frac{Area(A \cup B_1 \cup B_n \cup C) - Area(A \cup C)}{Area(oneframe)} \tag{4}$$

If these two conditions (1) and (2) are both met, we can continue track on next frame . Once one of the two conditions has been broken , that means, the frame interval between A and C are the upper limitation H. This parameter H can be calculated off line if the camera's moving model maintains the same .

Figure 2 (a) (b) (c) show three moving model . And as the moving speed change rapidly, the upper limitation H decreases, which means the key frames interval decreases. So the original step two can be modulated as following:

2. Tracking these N points frame-to-frame until M points have been matched or the frame interval reaches the upper limitation H, thus the current frame is F2. M is the threshold for reselecting new feature points;

B. Quality management

In the previous section A, a quantity management for feature points is proposed. Using this management method, we can select key frames through the feedback of the matched points number and also introduce an idea to ensure that these key frames can remain enough information. In this section some evaluate measures to manage the quality of these matched points will be developed to get most accurate mosaic.

According to the section A, the key frames F1 and F2 and their matched feature points M are send into quality management module. In order to calculate the most accurate transform matrix, we iterate the following computation with $s_f = \infty$ as initial values, and suppose the two key frames are on the same plane, so we adopt the affine transform model to mosaic the images and then we need two pairs every time.

1. Randomly sample two pairs $(u_1, v_1), (u_2, v_2)$ among the candidate pairs M;
   where $u_{1,2}$ are the feature points on reference image F1, and $v_{1,2}$ are the corresponding feature points on current image F2;
2. Compute the homography T determined by these two pairs $(u_1, v_1), (u_2, v_2)$;
3. Finding the overlap area O_F1, O_F2 in mosaicing image by using this matrix T, Where O_F1, O_F2 are from the two original images F1 and F2 separately.
4. Calculate the difference between O_F1 and O_F2. Here , we use the simple function as following:

$$difference(X, Y) = \frac{1}{W} \sum \|X - Y\|^2 \tag{5}$$

   where W is the sum number of pixels 。
   Therefore, $s = difference(O\_F1, O\_F2)$
5. If $s < s_f$, update s_f: $s_f \leftarrow s$

Repeat the above computations until the s_f reaches its minimum. Then regarding the two pairs as the most suitable key points for mosaicing.

This quality management of feature points can do moasicing well even if there are some wrong matched pairs appearance .What's more, we find that even all these matched points are correct and accurate, different combination of these points can result in different mosaic. By using this method to manage, we can get the most accurate mosaic under the current correspondence feature points.

IV. EXAMPLESNITAL RESULTS

An automated video mosacing system based on the presented algorithm is developed. The video image size is 320x240 (24 bits per pixel) captured by Sony EVI-D100 at 25fps. The system is tested with complex outdoor video sequences and achieves satisfied results. The follows represent the results.

Figure 3 shows the decrease curve of feature points number, usually, the threshold of tracking point number is shown in Figure 3 with a broken line. Once the tracked point number is less than the threshold, a reselection step of quantity management module is active.

Figure 4 displays the different mosacing images with various points selection. Figure 4.(a) shows the relative differences map. Our quality management module correctly chose the best mosaic results (Figure 4.a, No.3) with the lowest difference.

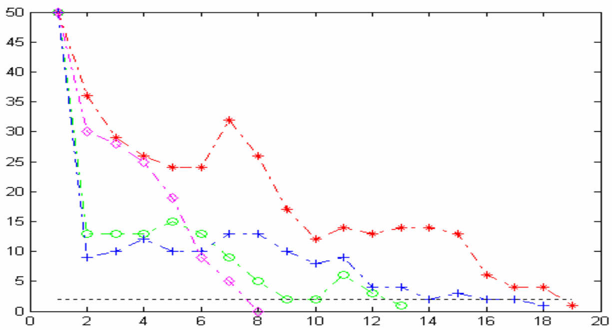


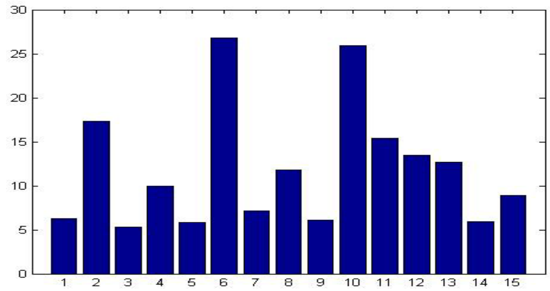
Fig 3. The decrease curve of feature points number

V. CONCLUSION

We have described a novel and robust feature-based method for mosaic construction. Through managing the quantity and quality of feature points, we can select key frames and from which we obtain the most accurate mosaic. Compare with other mosaic methods, we use the most efficient and effective feature points to automatically get the large view mosaic image for many applications such as surveillance etc.

ACKNOWLEDGMENT

The work is sponsored by the foundation of *National Laboratory of Pattern Recognition, Institute of Automation Chinese Academy of Sciences.*



(a)

(b)

Fig 4. Mosacing results under various feature points selection. (a) Difference map. (b) Mosacing results.

REFERENCES

[1] R. Kumar, H. Sawhney, S. Samarasekera, S. Hsu, H. Tao, Y. Guo, K. Hanna, A. Pope, R. Wildes, D. Hirvonen, M. Hansen, and P. Burt, "Aerial Video Surveillance and Exploitation," *IEEE proceedings*, vol. 89, no. 10, pp. 1518-1539, October 2001.

[2] Y. Kanazawa, K. Kanatani, " Stabilizing image mosaicing by model selection," *Second Workshop on 3D Structure from Multiple Images of Large-Scale Environments and Applications to Virtual and Augmented Reality (SMILE 2000),* July 1, 2, 2000, Dublin, Ireland, pp. 10-17.

[3] R. Szeliski, H.-Y. Shum ," Creating full view panoramic image mosaics and environment maps," *Proceedings of the SIGGRAPH'97*, Los Angeles, CA August, pp:251–258.1997

[4] I. Zoghlami, O. Faugeras, R. Deriche, "Using geometric corners to build a 2D mosaic from a set of images", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Puerto Rico June, pp: 420–425.1997

[5] Shum, H. & Szeliski, R. "Construction and refinement of panoramic mosaics with global and local alignment, " *IEEE Int'l Conf. Computer Vision,* pp. 953-958.1998

[6] Szeliski, R. "Image Mosaicing for Tele-Reality Applications," *Digital Equipment Corporation, Cambridge*, USA.1994

[7] Hartley,,R.&Zisserman,A. "Multiple View Geometry, " Cambridge University Press, UK, 2000

[8] Zoghlami, I. & Faugeras,O. & Deriche,R. "Using geometric corners to build a 2d mosaic from a set of images", *Computer Vision and Pattern Recognition,* pp 421-425. 1997

[9] Capel,D&Zisserman,A. "Automated mosaicing with super-resolution zoom", *Computer Vision and Pattern Recognition*, 23-25 Jun ,1998, pp: 885-891

[10] C.Tat-Jen and R.Clipolla, "A Statistical Framework for Long-Range Feature Matching in Uncalibrated Image Mosaicing" , *Proc. Of IEEE Comp.Soc. Conference on Computer Vision and Pattern Recognition,* Santa Barbara, CA, June.23-25, pp.442-447,1998,

[11] Yihong Gong, Guido Proietti, David LaRose, "A Robust Image Mosaicing Technique Capable of Creating Integrated Panoramas", *International Conference on Information Visualization*, London, England , July 14 - 16, pp: 1999