Automatic Panorama Recognition and Stitching Based on Graph Structure

Zhao Yi-Li
College of Computer and Information
Southwest Forestry University
Kunming, China
e-mail: ylzhao@swfu.edu.cn

Xia Yan
College of Material and Engineering
Southwest Forestry University
Kunming, China
e-mail: panovr@foxmail.com

***Abstract*—An automatic panorama recognition and stitching method based on graph structure was proposed to solve the multiple images recognition and matching problem. With multiple unordered images of user input, the method can automatically finding overlapping portion between images, and stitch them. First, the MOPS feature points were detected from the input images, and using kd-tree nearest neighbor search to perform fast feature matching between images. The motion model between any two images can be established by RANSAC algorithm based on the correspondence of feature points, and robust verification by a probabilistic algorithm. The automatic recognition of multiple panoramic images problem can be solved by building undirected connected graph corresponding to image matching relationship. Finally, a recursive algorithm was used to do depth-first traversal of the established undirected connected graphs, and multi-band blending algorithm was used to eliminate stitching seam. The experiments showed that multiple unordered images can be matched and recognized automatically, and panoramas can also be stitched automatically and seamlessly.**

Keywords-feature detection; feature matching; image recognition; image matching; image stitching

I. INTRODUCTION

The image matching method is generally divided into two types, direct matching or feature-based matching. Direct matching method attempts to use the pixel value of the image and the registration is performed by iterative method [1] [2]. The feature-based approach tries to extract different types of characteristics, such as line features or points feature from the image, and uses neighborhood information and these features to do feature matching [3] [4].

In the feature-based approach, currently the most popular used are invariant feature-based methods. Such methods commonly need to calculate the neighborhood of the point features and corresponding feature descriptors are used to complete the feature retrieval and matching. Work in this area was first proposed by Schmidt and Mohr [5], their method extends Harris corner by Gaussian derivation, and forms a rotation invariant descriptor. Lowe's method is aim to add scale invariance [6]. Some other researchers also designed affine transformation invariant feature descriptor [7] [8]. Currently the most common feature point detection operators include Harris corner detection operator, the DOG detection operator and the maximally stable region method [9]. Repeatability of feature points and descriptor matching performance evaluation has also made good progress in literatures [10] [11].

Invariant feature-based methods have been successfully applied to many areas, including the object recognition [6], structure from the motion [12] as well as a panoramic image stitching [13]. In spite of a lot of progress has been made for image matching, there is still space worthy of study, especially in the existing literature lack detailed discussions on the following two problems: (1) given unordered multiple images as input, how to effectively discriminating whether there is overlap portion and the size of the overlapping area between two images; (2) how to classified different images belonging to the same scene automatically and synthesis of the corresponding panoramic images.

A graph based automatic panorama recognition and stitching method was proposed, which can classify and recognize multiple unordered input images. First the MOPS features were detected on the input images, then using kd-tree to perform fast matching between feature points. Initial matching pairs of feature points can be obtained according to the ratio of the distance between the nearest feature point distance and the second nearest neighbor feature points. According to the correspondence of image feature points, the motion model between any two images can be computed by RANSAC algorithm, and using probability and statistics strategy to do robust validation. Multi-image matching can be modeled as constructing undirected connected graphs between different image nodes, and multiple image alignment can be modeled as depth-first traversal on the established connected graphs.

II. FEATURE DETECTION AND MATCHING

In order to determine the overlapping area and motion model between any two images, the method first extracts MOPS feature from input images. The MOPS by Matthew Brown [14] is a relatively lightweight scale invariant feature detector compared with the SIFT [6], and has advantage of faster detection speed. The MOPS algorithm extended Harris algorithm with rotation and scale invariance. If aiming to cylindrical panoramic image or spherical panorama stitching, it first needs to use the reverse mapping to warp input images to cylindrical or spherical coordinate, and use bilinear interpolation during the conversion to avoid aliasing. Then the feature points in each image are also converted to the corresponding cylindrical or spherical coordinate through forward mapping, and then matched. When matching feature points between different images, it is need to perform nearest neighbor search in the feature space. Using kd-tree based nearest neighbor search algorithm can reduce search time complexity.

Algorithm 1 – fast feature matching based on k-d tree

(a) Constructing kd-tree for each image's feature point set;

(b) Traversing every feature points of each image, initial image index $i = 0$, and feature point index $n = 0$. For the n^{th} feature point of image I, find the nearest neighbor nn_1 and second nearest neighbor nn_2 to all other images, and their Euclidean distance d_1 and d_2. If the value of d_1 / d_2 is less than 0.6, nn_1 is considered as the best match point;

(c) Once all image feature points have been traversed, it is also need to validate the feature matching results. Assume the feature point n_i of image I has matching index n_{ij} with image J, then checks whether the matching index of the feature point n_{ij} of image J with image I equals to n_i, if not, then the match is considered wrong.

III. Multi-image Matching Based on Graph Structure

According to the correspondence between the feature points, the matching between any two images of the input images set can be established and modeled as undirected graph structure. The model takes each input image as one node in the graph, if the matching between two images meets a given criterion, and then there is a connection line between these two nodes. Multi-image matching problem can be solved by finding all undirected connected graphs in the structure.

Algorithm 2 - multi-image matching based on undirected graph structure

(a) Traversing every image of input, if there have matching feature points between current image and the second image, the feature points set of the second image is added to current image;

(b) If the number of matching feature points between two images is greater than a given threshold value, that indicates a motion model exist between these two images, and the index value of the second image is added to the model matching set of the current image;

(c) Using RANSAC algorithm [15] to estimate the motion model between these matched two images, and excluding the outliers;

(d) For each potential match between two images, using statistics-based strategy [13] to validate matching's robustness.

IV. Panorama Recognizing

Once pairwise matching has been established, panoramic image sequence can be recognized automatically according to the connected set of the matching images, while rejecting the "noise" images which don't match with any other image. The recognizing process can be modeled as depth-first traversal to connected graphs established in part 3.

Algorithm 3 - automatic panorama recognition

(a) Check whether there are images need to be stitched, if there is one, select this image I_{from} as the start of the new panoramic image I_{result} and labeled this image as "stitched", if not, then the algorithm will exit;

(b) Assume there are N images in the matching list of image I_{from}, let index $s = 0$, and select image I_{to} with index equals to s in the list. If image I_{to} has not been stitched, invoke algorithm 4 to stitch I_{from} and I_{to};

(c) Let $s = s + 1$, if $s < N$ then return to step (b), otherwise output I_{result}, and return to step (a).

Algorithm 4 - automatic panorama stitching

(a) Adjusting the bounding box based on the matching between I_{to} and I_{result} as shown in Figure 1. In figure 1, each row represents one stitching step, and assume stitching start from IMG1, and bounding box is shown in dotted lines;

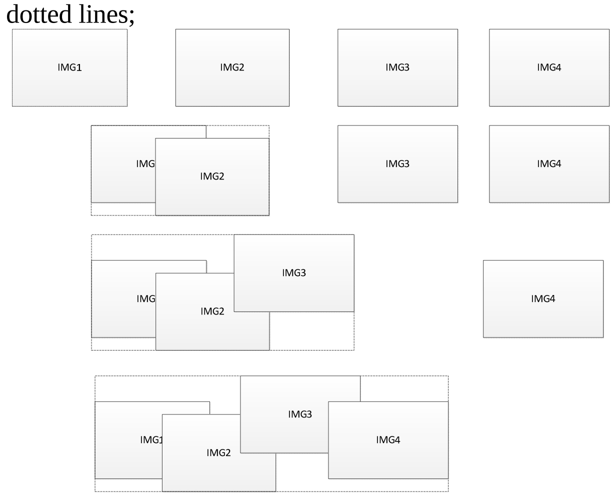

Figure 1. Bounding box dynamic adjusting.

(b) Adjusting the stitched images' position in I_{result} according to location of the bounding box, and generate a new result image I'_{result}, old image I_{result} is copied to the new location, and let $I_{result} = I'_{result}$;

(c) Using multi-band blending algorithm [16] to synthesize I_{to} and I_{result}, and recording the location and index of image I_{to} in image I_{result}, marking image I_{to} as "stitched".

(d) Fetching each element in turn from the matching list of image I_{to}, recursive calls Algorithm 4.

V. EXPERIMENTAL RESULTS

The upper part of figure 2 is 8 images of different scenes captured with a Canon IXUS 980IS digital camera by hand, the lower part is 4 panoramic images generated by the method. As the figure shown, the method can automatically recognize panorama sequence and stitched them. In figure 2 there is one image doesn't match with other images, so it is outputted as a separate "stitching" panorama.

Figure 2. Automatic panoramas recognizing and stitching

VI. CONCLUSION

This paper presents a new automatic panoramic image sequence recognizing and stitching method based on graph structure. The method uses MOPS and the probability model to perform image matching and validation for unordered image sets. The panoramic image of the unordered image set can be automatically recognized without user input. Because of using the multi-band blending strategy, even if the brightness differences exist between images due to illumination change, smooth transition can be formed in the overlapping area between images, while maintaining high-frequency details. Further research topics include using OpenCL or CUDA parallel computing framework to do the computation in GPU, as well as to explore other image fusion methods.

REFERENCES

[1] B. Lucas, T. Kanade, "An iterative image registration technique with an application to stereo vision," In: 7th International Joint Conference on Artificial Intelligence, pp. 674-679, IEEE Press, 1981.

[2] R. James, P. Anandan, J. Keith, H. Rajesh, "Hierachical model-based motion estimation," In: 2nd European Conference on Computer Vision, pp. 237-252. Springer-Verlag Press, 1992.

[3] W. Förstner, "A feature-based correspondence algorithm for image matching," In: International Archives of Photogrammetry & Remote Sensing, pp. 150-166, 1986.

[4] C. Harris, "Geometry from visual motion," Active Vision, MIT Press, Cambridge (1992).

[5] C. Schmid, R. Mohr, "Local grayvalue invariants for image retrieval," In: IEEE Transac-tions on Pattern Analysis and Machine Intelligence, vol.19, pp. 530–535, 1997.

[6] D. David, Distinctive image features from scale-invariant keypoints. In: International Journal of Computer Vision, vol.60, pp. 91–110, 2004.

[7] A. Baumberg, "Reliable feature matching across widely separated views," In: International Conference on Computer Vision and Pattern Recognition, pp. 774–781, IEEE Press, 2000.

[8] T. Tuytelaars, L. Van Gool, "Wide baseline stereo matching based on Local, affinely invariant regions," In: 11th British Machine Vision Conference, pp. 412-422, Bristol(2000).

[9] J. Matas, O. Chum, M. Urban, T. Pajdla, "Robust wide baseline stereo from maxi-mally stable extremal regions," In: 13th British Machine Vision Conference, pp. 384-393, Cardiff(2002).

[10] C. Schmid, R. Mohr, C. Bauckhage, "Evaluation of interest point detectors," In: Interna-tional Journal of Computer Vision, vol.37, pp. 151-172, 2000.

[11] K. Mikolajczyk, C. Schmid, "A Performance Evaluation of Local Descriptors," In: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.27, pp. 1615-1630, 2005.

[12] F. Schaffalitzky, A. Zisserman, "Multi-view Matching for Unordered Image Sets," In: 7th European Conference on Computer Vision, pp. 414-431, Springer-Verlag, Berlin (2002).

[13] M. Brown, D. Lowe, "Recognising panoramas," In: 9th International Conference on Computer Vision, pp. 1218-1225. IEEE Press, Nice(2003).

[14] M. Brown, R. Szeliski, S. Winder, "Multi-Image Matching using Multi-Scale Oriented Patches," In: International Conference on Computer Vision and Pattern Recognition, pp. 510-517, IEEE Press, San Diego(2005).

[15] A. Martin, C. Robert, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," In: Communication of the ACM, vol.24, pp. 381-39. NY(1981).

[16] J. Peter, H. Edward, "A Multiresolution Spline with Application to Image Mosaics," In: ACM Transactions on Graphics, vol.2, pp. 217-236, NY(1983).