An Improved RANSAC homography Algorithm for Feature Based Image Mosaic

FULI WU
Zhejiang University of Technology
College of Information Engineering
No.6 District Zhaohui, 310032 Hangzhou
China

XIANYONG FANG
Anhui of University
Key Lab. of Computing & Signal Processing
230039 Hefei
China

Abstract: In this paper, we proposed a new feature based image mosaic algorithm. The improved RANSAC homography algorithm based on the modified media flow filter, to detect wrong matches for improving the stability of the normal RANSAC homography algorithm. The method improved the local registration between neighboring images. Experiments and Statistical Analysis show that our mosaic method is robust.

Key–Words: Feature based mosaic, RANSAC, Homography, Media flow filter

1 Introduction

Image mosaic tries to composite several narrow-angle images into a wide-angle image and is widely used in aerial and satellite photographing [1], virtual touring and exhibition [2-3], photo edition [4], etc. Szeliski [5] reviewed the principles and advances of image mosaic. As described in [5], there are two types of method: direct method and feature based method, and the latter becomes more robust than the former with the advances of feature detection.

Feature based methods [6-9] mosaic the images by first automatically detecting and matching the features in the source images, and then warping these images together. Normally it consists of three steps: feature detection and matching, local and global registration, and image composition. Feature detection and matching aims to detect features and then match them. Local and global registration starts from these feature matches, locally registers the neighboring images and then globally adjusts accumulated registration error so that multiple images can be finely registered. Image composition blends all images together into a final mosaic. For more details on the current state of feature based mosaic, please refer to [5]. Registration is still not finely solved with existing techniques and our focus in this paper is also on how to improve the stability of registration for feature based mosaic.

Many papers on wide baseline matching [10-12], object recognition [13-14] and image/video retrieval [15-17] consider how to improve the stability of feature matching. In these works, feature matching is improved by spatial consistency which means the match features of each feature and its every neighboring feature should have the same spatial arrangement. Sivic et al. [15] used each region match in the neighborhood of each feature match to vote this feature match. The sum of votes of the whole frame decides the rank of the frame and match without vote is rejected. Ferrari et al. [10, 16] iteratively applied a expansion and contraction scheme to add new matches and remove wrong matches while expansion is fulfilled based on the similarity of affine transformations between neighboring region matches and contraction is reached by the sidedness constraint which bases on the fact that, to a triple of region matches, the center of a first region should be on the same side of the directed line going from the center of a second region to the center of a third region. The median flow filter [18] is also used to remove wrong matches, which compares the length and anger of each match vector with the median length and anger of its several neighboring match vectors respectively and selects the one whose length and anger below the thresholds.

But on the image mosaic side, there are few researches considering eliminating wrong matches before robust registration. As far as we know, only Cho et al. [6] applied the median flow filter to remove wrong matches before registration for image mosaic.

For image mosaic, to locally register the neighboring images, 8-parameter homography can be applied to accurately model the mapping between views under general image condition. RANSAC [19] is a commonly accepted way to refine the homography between images [7-9] because RANSAC can return the final inliers when getting the final homography. Except RANSAC, LMedS (the Least Median of Squares) [20] is also used for homography calculation, but it can not return the final inliers.

Even with RANSAC, simply applying RANSAC is still hard to robustly recover the homography because there are still some wrong matches. Therefore we can combine RANSAC with the spatial consistency requirement and then calculate a robust homography. Particularly, the median flow filter can be included into the normal RANSAC homography calculation process because the median flow filter fits to RANSAC homography algorithm naturally in two aspects: first, the median flow filter selects n matches for match verification, and RANSAC need select four matches in the initialization step of each iteration of homography calculation and these four matches affect the quality of RANSAC; second, to image mosaic, images are usually captured without apparent change of the optical center and the focal length, the angle and direction of the vector of each feature match should be very similar as the median flow filter requires. So the median flow filter can be applied to the random selection step of each RANSAC iteration to get the correct matches, and thus get an accurate homography. But the median flow filter need first select k nearest neighbors and then select n vectors as the tightest group among the nearest neighbors, i.e., select multiple times, it should be modified to consider all the matches selected in the initialization step of each iteration before being applied to the normal RANSAC algorithm because there are only four matches in the initialization step and comparison can be processed among all of them.

With above discussions, a new robust feature based image mosaic method can be proposed. First, features are localized with SIFT [21], described by PCA-SIFT [22] and matched with priority search in k-d tree structure [23]. Then the RANSAC homography algorithm based on the modified median flow filter is applied to locally register neighboring images. After registration, image composition can be fulfilled by multi band blending [24], featuring algorithm [2] or gradient domain fusion [25] and then a final mosaic can be generated.

In the following paragraphs, the local will be discussed separately in detail in the Section 2. Then experimental results will demonstrate the stability and efficiency of our method. The first step, feature detection and matching, and the last step, image composition, will not be discussed here and interested readers can refer to related papers.

2 The Improved RANSAC Homography for Local Registration

After feature detection and matching, the source images need to be registered together. RANSAC algorithm can be applied to get the homography of each image pair. Four initial putative feature matches are selected in the random selection step of each iteration in RANSAC [19], and a correct homography can be got after one iteration if they are the real inliers. However each feature will have more than one nearest neighbor after feature matching because of the similarity of the local patch. Therefore perhaps wrong match will be selected as the initial putative match in one iteration according to the random property of RANSAC. If selected, the weak homography of this iteration will be definitely generated. This homography will be likely to generate the largest number of matches and then the final homography will be weak. Therefore the stability of the random selection step needs to be improved.

2.1 The Modified Median Flow Filter

As discussed above, the result of RANSAC is affected by the initial putative matches who are affected by the random selection step of each RANSAC iteration. Therefore we need filters to improve the stability of the random selection step. Smith et al.[18] used a median flow filter to remove the wrong feature matches. The median flow filter consists of two parts: angle filter and length filter. Each match is considered as a motion (or 'flow') vector and each vector compares with its k nearest neighbors in each filter. First the angle filter checks whether the direction of each vector is beyond the median direction threshold of the n tightest vectors. If not beyond, the vector will be retained as an inlier. If beyond, and then if there exists a vector which is below a certain length threshold, the length filter is applied. This time the length of each vector is checked so that the length of each final vector is less than the median length threshold of the n tightest vectors. In the random selection step of each RANSAC iteration, there are only four feature matches can be selected, so we do not consider the k nearest neighbor vectors and the n tightest group. In this step, we use all the four feature matches and compare all of them. Figure 1 and Figure 2 explains how this modification works.

Figure 1 shows the angle filter which is the first step to check the four feature matches in the modified median flow filter. The four green lines show that the directions of four vectors composed by the four correct corresponding feature matches. α stands for the median direction of these four correct vectors, while

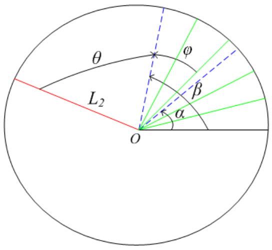

Figure 1: The median flow angle filter.

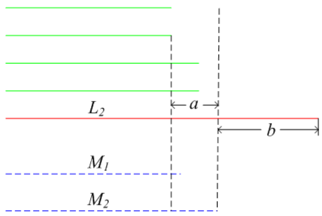

Figure 2: The median flow length filter.

β stands for the median direction if there is a wrong vector L among the four vectors. The direction difference between each correct vector and α is not so large. However, when L exists, θ, which is the direction difference between L and β, is obviously larger than φ which is the direction difference between the correct vector and β. If θ is larger than a threshold, these four matches will not pass the angle filter.

To the four feature matches that do not pass the angle filter, a second check on length is needed. This is because perhaps there are short motion vectors which can not be determined with great certainty especially when noise is strong. Figure 2 explains the length filter. If all the four corresponding vectors (shown in green lines) are correct, we can get a median line M_1. But if there is a wrong vector L, then we will get a media line M_2. The length difference between each correct vector and M_1 is not so large. But when β exists, b, which is the length difference between L and M_2, is obviously larger than a which is the length difference between the correct vector and M_2. So if b is larger than a threshold, these four feature pairs will not pass the length filter and thus can not pass the median flow filter.

In the random selection step of RANSAC, if the four randomly selected feature matches can not pass the median flow filter, they can not be selected as the initial feature matches to calculate a homography and this random selection step should try again. If they can pass the median flow filter, a homography can be calculated and the next step of the RANSAC iteration can execute.

2.2 The Modified RANSAC Homography Algorithm

The improved homography algorithm can be summarized after previous discussion. First each RANSAC iteration works in the following four steps:

- Select a random sample of four feature matches and then applying the angle filter. If pass, skip to step 3. If not, go to the next step.
- Apply the length filter. If pass, go to the next step. If not, go to step 1.
- Compute the homography G.
- Compute the number of inliers consistent with G by a distance threshold.

Then G with the largest number of inliers is selected after many iterations and the final homography H can be recalculated with the inliers consistent with the selected G. For details on homography calculation in the RANSAC and the distance threshold function in step 4, please refer to Hartley et al. [19].

3 Experiments

Figure 3 and 4 compare the normal RANSAC homography algorithm with the improved RANSAC homography algorithm. To the normal algorithm, usually only about 35 inliers are returned. Figure 3 shows the typical registration result after the normal RANSAC homography calculation and, as showed in the bottom of Figure 3, clearly there are ghosting effects in the rectangular areas. But after applying our improved RANSAC homography algorithm, usually there are more than 350 inliers returned and the homography can be accurately returned. Figure 4 shows the typical registration result after the improved RANSAC homography calculation and, as showed in the bottom of Figure 4, obviously there is no ghosting in the rectangular areas.

More comparisons between the improved RANSAC homography algorithm and the normal RANSAC homography algorithm are undertaken. 40 image pairs whose sizes range from 250*300 to 640*480 are first captured with a hand held camera. Then the features of them are detected with SIFT, described by PCA SIFT and matched with priority search. After feature detection and matching, the normal RANSAC and the improved RANSAC run 10 times respectively for each pair of images. In the improved RANSAC, 5 and 3 are always set as angle threshold and length threshold as Smith [18]. We show the statistical comparison between the normal RANSAC homography algorithm vs. the improved RANSAC homography algorithm. The histograms

Figure 3: The normal homography algorithm and the red rectangular areas.

Figure 4: The improved homography algorithm and the red rectangular areas.

of Figure 5 and Figure 6 shows the average inliers returned for 10 example image pairs. Clearly the improved homography algorithm returns about 2 times of the average inliers of the normal homography algorithm.

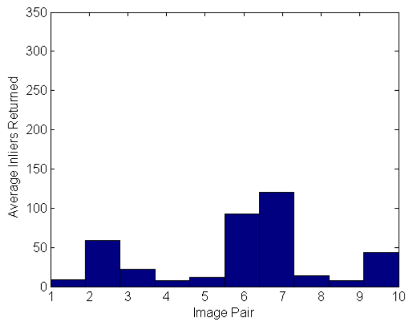

Figure 5: The average inliers returned after normal homography algorithm with 10 sample image pairs.

4 Conclusion

In this paper, we put forward a robust feature based image mosaic method. First features are detected with SIFT, described by PCA-SIFT and matched with priority search. Then images are locally registered with the improved RANSAC homography algorithm. Finally the images are fused to be one image with a blending method. In the improved RANSAC homography algorithm, a modified median flow filter is introduced for robustly selecting initial feature matches and thus improving the accuracy of the final homography. Experiments show that our feature based image mosaic method is highly effective.

Acknowledgements: We would like to express our thanks to Zhigeng Pan at Zhejiang University and Bin Luo at Anhui University for much helpful advice. The work is under the supports of Nature Science Foundation of Zhejiang Province (Grant No. Y105095) and Science Fund for Higher Education of Anhui Province (Grant No. KJ2007B005)

References:

[1] F.H.Moffitt and E.M.Mikhail, *Photogrammetry*. Harper & Row, New York, 1980.

[2] Richard Szeliski, Video Mosaics for Virtual Environments, *IEEE Computer Graphics & Applications* 16(2)1996, pp. 22–30.

[3] Zhigeng Pan, Xianyong Fang, Jiaoying Shi and etc. EasyTour: a New Image-Based Virtual Tour System, *In Proceedings of ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry*, 2004, pp. 467-471.

[4] Shenchang Eric Chen, QucikTime VR-an Image-Based Approach to Virtual Environment

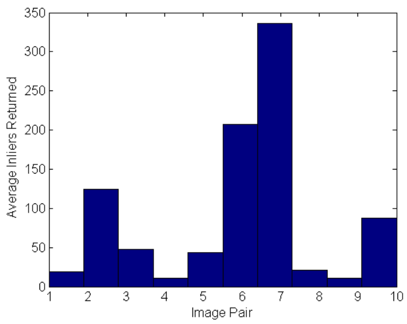

Figure 6: The average inliers returned after improved homography algorithm with 10 sample image pair.

Navigation, *Computer Graphics*,1995, pp. 29-38.

[5] Richard Szeliski, Image Alignment and Stitching: A Tutorial, *Technical Report*, MSR-TR-2004-92, Microsoft Research, 2004.

[6] Soo-Hyun CHO, Yun-Koo CHUNG and Jae Yeon LEE, Automatic Image Mosaic System Using Image Feature Detection and Taylor Series, *In Proceedings of the 7th International Conference on Digital Image Computing: Techniques and Applications*, Sydney, Australia, 2003, pp. 549-556.

[7] Matthew Brown and David Lowe, Recognising Panorama, *In Proceedings of the 9th International Conference on Computer Vision*, Nice, France, 2003, pp. 1218-1225.

[8] Philip H. S. Torr and Andrew Zisserman. Feature Based Methods for Structure and Motion Estimation, *In Proceedings of the International Workshop on Vision Algorithms: Theory and Practice*,Springer-Verlag,London, United Kingdom, 1999, pp. 298-373 .

[9] David Capel and Andrew Zisserman. Automated Mosaicing with Super-Resolution Zoom, *In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Santa Barbara, United States, 1998, pp. 885-891.

[10] Vittorio Ferrari, Tinne Tuytelaars and Luc Van Gool, Wide-Baseline Multiple view Correspondences, *In Proceedings of IEEE Computer Society Conference on Computer Vision Pattern Recognition*, Madison, USA, 2003, pp. 718-725.

[11] Jir Matas, Ondrej Chum, Martin Urban and etc, Robust Wide Baseline Stereo from Maximally Stable Extremal Regions, *In Proceedings of British Machine Vision Conference*, Cardiff, UK, 2002, pp. 384-393.

[12] Tinne Tuytelaars and Luc Van Gool, Wide Baseline Stereo Matching Based on Local, Affinely Invariant Regions, *In Proceedings of British Machine Vision Conference*, Bristol, UK, 2000 , pp. 412-425.

[13] Vittorio Ferrari, Tinne Tuytelaars and Luc Van Gool, Integrating Multiple Model Views for Object Recognition, *In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, USA, 2004, pp. 105-112.

[14] Stepan Obdrzalek and Jin Matas, Object Recognition Using Local Affine Frames on Distinguished Regions, *In Proceedings of British Machine Vision Conference*, Cardiff, UK, 2002, pp. 113-122.

[15] Josef Sivic and Andrew Zisserman, Video Google: a Text Retrieval Approach to Object Matching in Videos, *In Proceedings of International Conference on Computer Vision*, Nice, France, 2003, pp. 1470-1477.

[16] Vittorio Ferrari, Tinne Tuytelaars and Luc Van Gool, Retrieving Objects from Videos Based on Affine Regions, *In Proceedings of European Signal Processing Conference*, Vienna, Austria, 2004, pp. 128-131.

[17] Josef Sivic and Andrew Zisserman, Video Data Mining Using Configurations of Viewpoint Invariant Region, *In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, USA, 2004, pp. 488-495.

[18] P. Smith, D. Sinclair, R. Cipolla and etc, Effective Corner Matching, *In Proceedings of the Ninth British Machine Vision Conference*, Southampton, United Kingdom, September 1998.

[19] Richard Hartley, Andrew Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, United Kingdom, 2000.

[20] Rachid Deriche, Zhengyou Zhang, Quang-Tuan Luon and etc, Robust Recovery of the Epipolar Geometry for an Uncalibrated Stereo Rig, *In Proceedings of European Conference on Computer Vision*, Stockhom, Sweden, 1994, pp. 567-576.

[21] David Lowe, Distinctive Image Features from Scale-Invariant Interest Points, *International Journal of Computer Vision*, 60(2), 2004, pp. 91-110.

[22] Yan Ke and Rahul Sukthankar, PCA-SIFT: A More Distinctive Representation for Local Image Descriptors, *In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, United States, 2004, pp. 511-517.

[23] Sunil Arya and David Mount, Approximate Nearest Neighbor Queries in Fixed Dimensions, *In Proceedings of the 4th Annual ACM-SIAM Symposium on Discrete Algorithms*, Austin, United States, 1993, pp. 271-280.

[24] Peter J. Burt and Edward H. Adelson, A Multiresolution Spline With Application to Image Mosaics, *ACM Transactions on Graphics*, 2, 1983, pp. 217-236.

[25] Anat Levin, Assaf Zomet, Shmuel Peleg and etc, Seamless Image Stitching in the Gradient Domain, *In Proceedings of 8th European Conference on Computer Vision*,2, Prague, Czech Republic, 2004, pp. 377-389.