The Journal of **Macro**Trends in Technology and Innovation

MACROJOURNALS

A HYBRID METHOD FOR FINDING CORRESPONDENCES FROM UN-CALIBRATED PAIR OF HIGH RESOLUTION IMAGES

J. Zraqou, W. Alkhadour, A. Tayyar and S. Masaadeh
Computer Multimedia Systems, Isra University, Amman, Jordan

Abstract

Finding correspondences automatically from a pair of un-calibrated images is a challenging phase in 3D reconstruction. Two proposed methods based on the existing SIFT and SURF key point detector and descriptor are introduced in this work to find correspondences automatically from high resolution pair of images for aerial scenes. The characteristics of the data set involved in this research in terms of size and the presence of affine transformation add more challenges in finding the set of desirable matches.

Keywords: *Finding correspondences, un-calibrated images, fundamental matrix, 3D reconstruction*

1. Introduction

Finding correspondences from a pair of un-calibrated images for reconstructing 3D data for a scene is renowned problem in 3D reconstruction field. Several approaches have been presented by scholars to tackle this problem. However the introduced methods tackle different situations. Some algorithms were employed to find correspondences from two images taken from small baseline. The methods in such cases are based on comparing two regions from both images using different similarity measurements such as correlation. However these methods fail to find correspondences from pairs of images taken from widely different angles. Several studies have been introduced to work with pair of images with affine transformation.

An approach for finding correspondences from multispectral images was introduced in [1]. The algorithm starts by defining edges from both images using canny algorithm, then the edges are mapped into Hough-space, finally matching between images is conducted. Multi-image matching algorithm via fast alternating minimization is proposed in [2]. The method is built based on a global optimization approach to jointly matching a set of images. A near-optimal joint object matching via convex relaxation method is introduced in [3]. The method aims to find correspondences between multiple objects that have only partial similarities. A consistent partial map collection is encoded into a 0-1 semidefinite matrix, then a spectral technique followed by a parameter-free convex program called MatchLift are used to identify the correspondences. A multi-graph matching via affinity optimization with graduated consistency regularization is presented in [4]. Two correlation steps algorithm was applied in this study; first the local matching affinity across pairs of graphs was achieved followed by the global matching consistency. A robust feature correspondence approach for matching objects in a set of images is introduced in [5]. Both inliers and outliers are extracted, then a partial permutation matrix (PPM) was optimized for each image, where then the (PPM) was used to find correspondences.

Surveying the literature shows that the SIFT keypoint detector and descriptor [6] is a powerful method to find matches between two pair of images. However the SIFT is slow and requires large amount of memory, one of the best algorithms is the SURF method that outperforms the SIFT method in terms of speed and accuracy proposed in [7]. The need for lower cost computational algorithm motivates researchers to introduce different replacements of the SIFT.

The aim of this work is to develop an automatic 3D reconstruction system from high resolution aerial images. The images were taken from widely separated angles and hence include an affine transformation which produces local deformations that cannot be estimated using intensity cross correlation methods, thus an affine model is required to find corresponding points between images [8] [9].

This paper introduces the authors proposed methods to find a set of reliable correspondences employed to generate a dense disparity map. Experimental results show that the proposed methods successes in achieving the author objective in automatically finding a set of desired correspondences and creating dense disparity map and finally creating the 3D map.

2. Data set

A set of aerial images taken for rural, urban and architectural scenes is used in this study.

2.1 Rural Images

A light aircraft flying freely was used to capture images for archaeological sites. To produce images with prominent features, most images were taken in sunshine. However the shadow affects some parts of the image producing poor contrast which adds more challenge to the process of finding correspondences.

The set of rural images concerned in this study includes images of hill forts in the Northumberland, Top O’ Noth image pair scanned at 2674×2664 pixels, Humbelton image pair 2266×2222 pixels, Hop image pair 6144×4606 pixels, Black_castle Newlands and Parkburn images, both image pairs are scanned at 6144×4606 pixels.

2.2 Urban Images

The urban scene tested in this work includes images for Whitby Abbey images 1400x1390 pixels. The presence of occlusion at edges of the roof and shadow adds more complication on the process of finding correspondences which is more noticeable than what it is in the rural images.

2.3 Architectural Scenes

The architectural image interested in this work is an image for Bradford University Catholic Chaplaincy. The structure of the architectural images which is in most often includes prominent features adds advantages over the aerial images. On the other hand the presence of occlusion and the periodicity increases the challenges in finding correspondences.

3. THE PROPOSED METHOD

Several investigations are conducted to find correspondences from the data set interested in this work including Harris method introduced in [10] and KLT method presented in [11]. However both Harris and KLT algorithms fail to find the desired correspondences. Previous studies in this field prove that the SIFT algorithm which is invariant to lighting and perspective transformations is practical and powerful in producing reliable matches from pairs of images taken from wide baseline [12]. Experimental results conducted by the author show that the SIFT algorithm is capable to return a set of matches concerned in this study. However a large number of matches is produced by SIFT, in addition the SIFT is very slow and needs a large space in memory. Since the SIFT is slow it was worth to test the SURF algorithm. Experimental results reveal that SURF successes in returning the correspondences from the interested data set. However both SIFT and SURF algorithms would not have run on the full resolution images interested in this work. Methods for tackling the limitations of the SIFT is presented below in addition a method is proposed to take advantage of the accuracy of the SIFT and the speed of the SURF.

3.1. Cropped Filter SIFT (CFSIFT) Algorithm

The First proposed algorithm in this study aims to speed up SIFT and reduce memory requirements.

3.1.1. Reducing Number of Matches

The number of matches produced by SIFT is very large, reducing the number of matches can be achieved by reducing the threshold value. However reducing the threshold value will yield a set of matches which is not distributed widely over the whole image. Since the focus of this study is to produce a set of matches spread widely over the whole image which will be used later in generating dense disparity map. A technique for reducing the number of matches is designed without affecting the robustness or matches distribution. The image is divided into a grid of 100 cells and one match is selected from each cell, the algorithm is called filtered SIFT (FSIFT) and illustrated in figure 1.

3.1.2. Speed and Memory Utilization

First of all the resolution for each image is reduced by a specific factor (*f)*, the degraded resolution images are used to find matches using FSIFT technique, The obtained matches are considered the seed for finding new matches in the higher resolution images by multiplying the coordinates of the seed matches in the lower resolution images by the inverse of the factor (*f)*. The next stage is cropping a square region of size 100x100 around each match in the higher resolution image and then running SIFT to obtain the matches in these regions. The algorithm is called Cropped Filter SIFT (CFSIFT) and illustrated in figure 2.

3.1.3. Refining the Set of Matches

The matches obtained from CFSIFT is used to estimate the fundamental matrix computed using the modified eight point algorithm introduced in [13] , where the quality of this matrix Q_f is highly dependent on the quality of the matches. The quality of the matrix can be improved by eliminating less robust matches which can be measured by perpendicular distance from the point to the epipolar line Q_b. Performing the refinement process of the matches can be achieved manually but it is time consuming and might limit the distribution of matches to specific regions in the image. An automatic refinement technique is introduced in this study to overcome this problem. The algorithm is called MQB and summarized in the following steps. First, the left image is divided into blocks, the number of blocks N_b is set to 10, 20, and 30 based on experimental results. Second step includes searching for the point with minimum Q_b in each block and saving this point along with its coordinates in matrix of matches points (MatMP). Final step is removing all other matches not included in (MatMP).

3.2. Hybrid Method of SIFT&SURF

The SURF method investigated in this study proves to be practical in finding matches and faster than SIFT. However the SURF and the SIFT cannot be applied directly to the full resolution images of size 6144x4606 pixels interested on this study based on our experiments. The SURF was run on lower resolution version extracted from the higher resolution one and then the cropping technique described in the CFSIFT were applied in the same manner but instead of SIFT the SURF method was used. Experimental results show that the quality of the matches in most cases is not acceptable which indicates the presence of a lot of false matches. Based on the results of experiments which shows that the SIFT is more accurate in finding the correspondences and that the SURF is faster than SIFT, it was worth to take advantages from both by combining SURF and CFSIFTF to generate new method called SIFTSURF. In the SIFTSURF the initial matches were found from lower resolution image using SURF algorithm and then CFSIFT were applied to find the final set of matches.

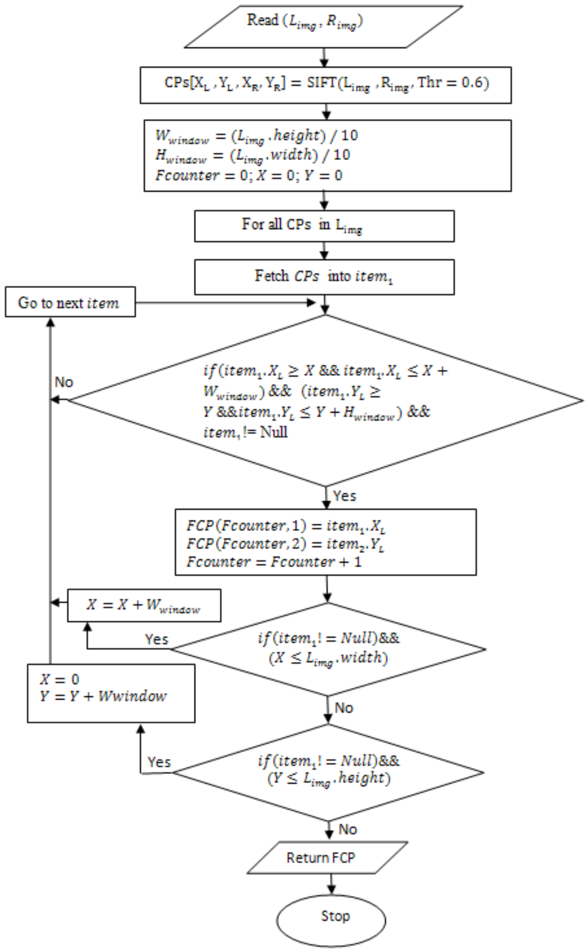

Fig. 1. Flow chart of the FSIFT technique**.**

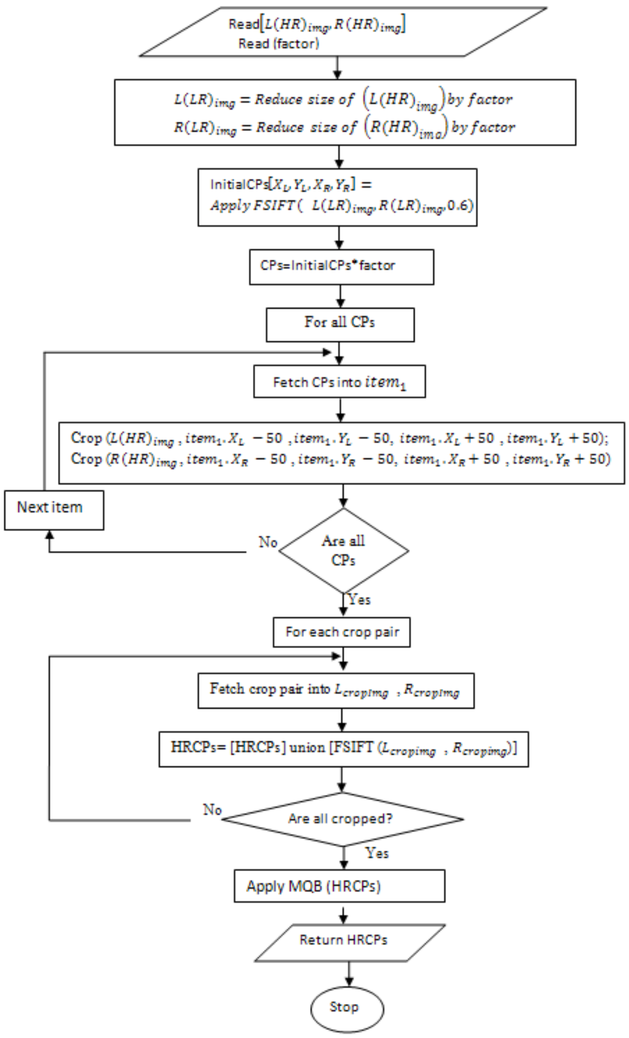

Fig. 2. Flow chart of the proposed CFSIFT method.

4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed methods several experiments were conducted on the interested data set using the proposed methods CFSIFT and SIFTSURF. Experimental results show that both proposed methods meet the aim of this study in producing automatically a set of matches with good quality factor and widely distribution on the images. However these results show that CFSIFT works better than SIFTSURF method on some of the data set in terms of quality factor and distribution of matches while on other images it is preferable to use SIFTSURF method to produce the matches.

The matches were refined using the proposed MQB algorithm which proves to be faster and more efficient than refining manually to discard the outliners or the less reliable matches in some cases. Some samples of the experimental results conducted in this study are shown in figures 3, 4 and 5. A comparison between original SIFT and both proposed methods in terms of number of matches, quality factor (Q_f) and status of distribution of matches were performed and presented in table1.

Table 1. Comparison between SIFT, CFSIFT, and SIFTSURF based on time consuming.

Original Image	Size of LR image	SIFT	CFSIFT	SIFTSURF
		Time (sec)	Time (sec)	Time (sec)
Top O' Noth 2674x2664	1338x1331	Null	340.86	43.30
Black-castle 6144x4606	3072x2303	Null	Null	58.38
Black-castle 6144x4606	1536x1152	Null	81.82	21.33
Parkburn 6144x4606	3072x2303	Null	Null	45.21
Parkburn 6144x4606	1536x1152	Null	104.43	72.26
Hop 6144x4606	3072x2303	Null	Null	100.01
Hop 6144x4606	1536x1152	Null	232.43	69.76
Humbelton 2260x2222	1130x1111	9362.7	238.52	88.33

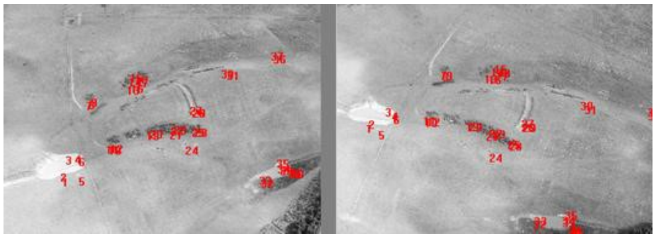

Fig. 3. Park Burn HR pair of images 3072x2303 pixels. Final matches found with CFSIFT in HR images based on initial matches found from LR images 768x576 with FSIFT. Refinement is done using MQB algorithm, *Qf* = 0.29.

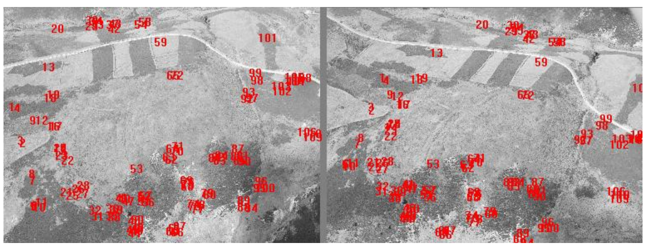

Fig. 4. Hope HR original image pair 3072x2303 pixels, final matches found with CFSIFT in HR images based on initial matches found from LR images 768x576 with FSIFT. Refinement is done using MQB algorithm Qf = 0.22.

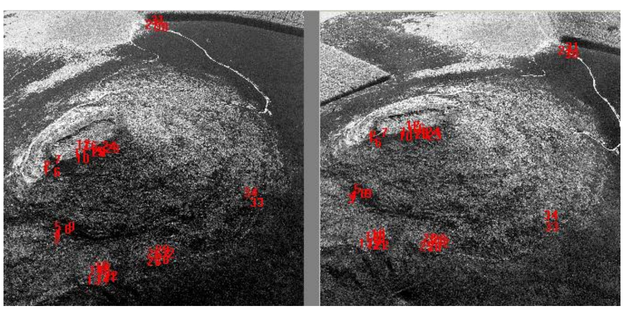

Fig. 5. Tap O‟ Noth HR image pair 1338x1331 pixels, final matches is found with CFSIFT in HR images based on Initial matches found from LR images 335x333 with FSIFT. Refinement is done using MQB algorithm Q_f = 0.49.

Finally the matches produced and refined automatically by the proposed methods are used to estimate the fundamental matrix which later is used to rectify the images along with a reference plane determined by three points [14]. The disparity maps for the data set can be generated using correlation based approach. A disparity map for one image in the data set is shown in figure 6.

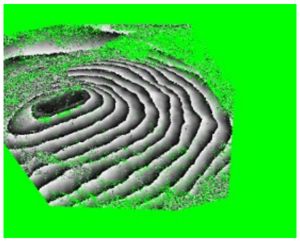

Fig. 6. Tap O' Noth disparity map generated using final matches shown in fig.5.

5. CONCLUSIONS AND FUTURE WORK

This paper introduces a hybrid method for finding correspondences automatically from un-calibrated large high resolution images for aerial scenes. The proposed CFSIFT and SIFTSURF methods tackle the limitation of existing methods in yielding the desirable set of matches.

Currently, the set of produced matches using the proposed method is used to estimate the fundamental matrix which then be used to rectify the images along with a reference plane determined by three points of the set of matches. The focus of the future work will be on presenting a new local rectification algorithm which works on rectifying each region from the image separately and then applying transformation to move from local rectified images to global one. The local rectification would improve the accuracy of rectification which in result improves the density and accuracy of the disparity map.

References

[1] Azzopardi, George, Petkov, Nicolai (Eds.), "Computer Analysis of Images and Patterns", 16th International Conference, CAIP 2015.

[2] X. Zhou, M. Zhu, and K. Daniilidis, "Multi-Image Matching via Fast Alternating Minimization", International Conference on Computer Vision (ICCV), 2015

[3] Y. Chen, L. Guibas, and Q. Huang,"Near-optimal joint object matching via convex relaxation", In International Conference on Machine Learning, 2014.

[4] J. Yan, M. Cho, H. Zha, X. Yang, and S. Chu. "Multi-graph matching via affinity optimization with graduated consistency regularization". IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015.

[5] K. Jia, T.H. Chan, Z. Zeng, S. Gao, G. Wang,T. Zhang, and Y. Ma," ROML: A Robust Feature Correspondence Approach for Matching Objects in A Set of Images", International Journal of Computer Vision, Springer US, 2015.

[6] D.G. Lowe, "Distinctive image features from scale-invariant keypoints". International journal of computer vision, pp. 91-110, 2004.

[7] H. Bay, T. Tuytelaars, , and L. Van Gool, "Surf: Speeded up robust features". Lecture notes in computer science, 3951: 404, 2006.

[8] M. Urban, J. Matas, O.m. Chum and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions". Image and Vision Computing, 22(10): 761-767,2004.

[9] A. Baumberg, "Reliable feature matching across widely separated views". IEEE Computer Society, 2000.

[10] C. Harris, and M. Stephens," A combined corner and edge detector", Proceedings of The Fourth Alvey Vision Conference, pp. 147-151, 1988.

[11] http://www.ces.clemson.edu/~stb/klt/, Last accessed on April 2010.

[12] E. Tola, V. Lepetit, and P. Fua," A fast local descriptor for dense matching", Citeseer, 2008.

[13] R.I. Hartley, "A linear method for reconstruction from lines and points". Published by the IEEE Computer Society, 1995.

[14] A.H, Alzhrani," 3-D Reconstrucion by Correlating Un-calibrated Images". University of Bradford, England, UK, 2003.