An Efficient Approach Towards Image Stitching in Aerial Images

R.M.N.B. Rathnayake
Department of Electrical and Computer Engineering
Faculty of Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
namal.b@sliit.lk

L. Seneviratne
Department of Electrical and Computer Engineering
Faculty of Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
lasantha.s@sliit.lk

***Abstract*— Aerial images taken by Unmanned Aerial Vehicles (UAV) contain rich geographic information. However, the higher the altitude the cost for image capturing increases and image quality or the visibility decreases. To overcome this issue, people use low altitude UAV's to capture images in bulk quantities and stitching algorithms were heavily used for merging these images. This paper presents a novel technique to automate this stitching process. This new technique is implemented based on well-known algorithms such as Speeded-up Robust Features (SURF), Principle Component Analysis (PCA) and Random Sample Consensus (RANSAC). The result of the proposed method was compared with some existing algorithms and the results are presented in this paper.**

Keywords--- Areal Images, Image Stitching, PCA, RANSAC, and SURF.

I. Introduction

With the rapid growth in UAVs, techniques in aerial imaging have widely improved. State of the art UAVs are equipped with a large number of sensors, actuators, and sophisticated controllers and are very stable even in harsh environments. Therefore, these UAVs are widely used in aerial image capturing. The cameras attached to UAV's continuously capture and transmit images to base stations. These images are then merged to gather and generates a large high-quality image. This margin can be achieved by using a number of existing algorithms, such as SUM, AVERAGE, and CRREJECT etc. The literature clearly shows that stitching algorithms produce promising results when it comes to aerial imaging [1].

Image Stitching has become an important topic in computer vision in recent years. It is used in various fields and especially in 2D, 3D aerial image mapping [2], Digital Videography [3], Medical image analysis [4] etc. In stitching, Image Registration, Calibration, and Blending are the three main sections that are widely discussed.

Image Registration is a process that aligns images into a single image. The ultimate goal is to find the transformation of images and perform image mapping. Rigid, Affine, Polynomial, B-spline, Thin-plate Spline, Radial basis, wavelets, and Fourier series are the common transformation parameterizations [5]. Minimizing the effects of distortion, exposure differences between images, vignetting, camera response and chromatic aberration are called image calibration. Image features are very important when it comes to image stitching. Image calibration is used to obtain the absolute position of the features & Image blending is the final step in image stitching. The illumination differences and misalignment in stitching makes the blending process more challenging [1].

II. Related Work

Stitching images is a hot topic in the field of image processing. Many scholars have proposed different methods in stitching. Image Registration is the key part of image stitching and dominant based registration and feature-based registration are the two main methods used in image registration [5]. However, feature-based registration is the most commonly used method in image registration. Feature-based techniques are low in complexity and high robustness. In practice, Image registration can be achieved with local and global approaches as given in [6].

In [7], authors have proposed a method for image stitching using SIFT features. For further improvements in accuracy, RANSAC algorithm was embedded with SIFT. This method overcomes the translation effects and improves accuracy. However, the SIFT algorithm still shows some issues in view angles of input images [8]. In practice, the SURF algorithm performs much better than the SIFT algorithm. Authors in [9] have proposed a binary classifier that extracts features using the SURF algorithm. Additionally, they also have used RELIEF-F algorithm for image registration. In [10] an adaptive uniform method for image stitching is proposed using SURF features. The proposed algorithm utilizes the radius to remove the unwanted interest points while selecting the best uniform distribution radius.

A fast and robust scene matching algorithm is proposed in [11] using SVD-SURF. They have built a SURF scale space based on the singular value features in real time. Using RANSAC algorithm, they have removed the outlier matches to obtain better results. In [12], a multi-view image registration method was proposed using SURF feature descriptor. Features were matched using Sum of Square Difference (SSD) algorithm. Then RANSAC was used to remove false matching points. Furthermore, they have also done a comparison between SURF and Harris algorithms. The results of the above method have shown that SURF is robust and has a higher precision in detecting features.

978-1-5386-4417-1/18/$31.00 ©2018 IEEE

III. METHODOLOGY

The proposed approach consists of a number of modules as shown in Figure 1. Here, the SURF feature extraction module considers two input images and extracts the corresponding features of each image. Then features are passed to the next module to apply RANSAC. In this module, the false matchings are eliminated and resulting features are transferred to the next module. This module calculates the H (Hessian) matrix. As the next step of the algorithm image feature points are transferred to the coordinate system of the reference image. Finally, image fusion is executed to stitch images. As shown in Figure 01, SURF features are extracted from the input images. The images are sorted in a systematic way to generate an order sequence. In general, SURF has two main tasks, namely, feature point detection and feature point description. Hessian matrix (Eq. 1) below was used for feature point detection.

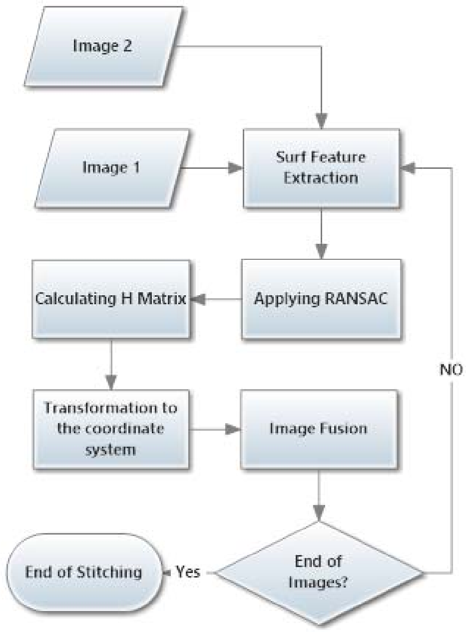

Figure 1: Workflow of the algorithm

$$H = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{bmatrix} \quad (1)$$

Where L_{xx}, L_{xy}, L_{xy} and L_{yy} are the convolution of the second derivative of Gaussian of the input image and the corresponded feature point [13]. After constructing the integral image and interest points, the locations are selected based on the Hessian matrix. Once all the interest points are calculated, they are localized in the scale image space.

In general, SURF descriptor is scale invariant and rotationally invariant. In SURF, the descriptors are constructed using square regions around the feature points. Then the regions are divided into 4×4 sections. At regularly spaced sample points, Haar wavelets are extracted. Then using the above-constructed regions a 4-dimensional vector is formed. Main steps in the SURF algorithm is shown in Figure 02.

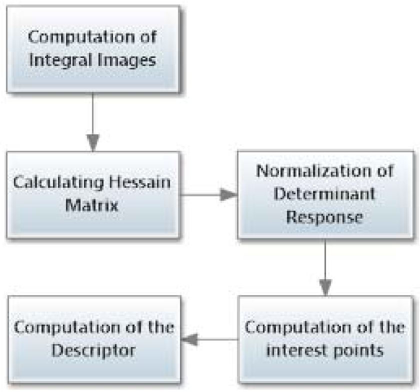

Figure 2: Stages of the SURF algorithm

A. Sorting Images

After obtaining features from the corresponding aerial images, image sorting is performed on the feature vectors as described below. Mainly there are three steps in image sorting.

1. The two images have to be separated into two bins as in Figure 3 & 4. The overlap area of image 1 over image 2 has to be in the right-hand side of image 2 according to the proposed algorithm.

Figure 3: Image 1

Figure 4: Image 2

2. The next step involves determining the order or the sequence of input images. Once features are extracted by SURF, the key points are matched using Eq.2. This process is called the nearest neighbor matching.

$$D = \sum_{i=1}^{64} (X_2[i] - X_1[i])^2 \quad (2)$$

Where D is the distance and X_1 and X_2 are the image 1 & 2 respectively, i is the conceded feature point. Based on D calculation, minimum distance D_{min} and second minimum distance D_{Sec} is calculated. Based on the relationship of D_{min} and D_{Sec}, the threshold value R was calculated as in Eq.3. Depending on the threshold value R the matching was decided. More accurate results can be obtained by having a larger value for R. For aerial images, having a value of 0.7 for R gives 72.2% accuracy in matching. Therefore, this figure was used through this research to find the best matching points i.e. a feature point will be considered as a matching point when R is calculated above 0.7.

$$R = \frac{D_{min}}{D_{sec} } \quad (3)$$

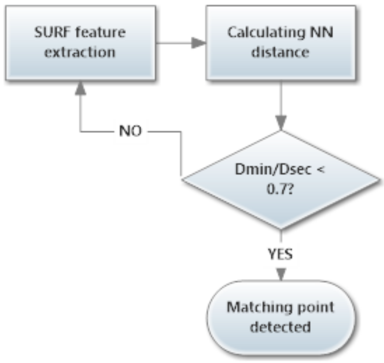

Figure 3: Flowchart for the process - Image Matching

The complete image sorting process can be concluded as follows. Using SURF descriptor, features are extracted for each and every image. PCA is used to reduce the number of feature points to 25. Then the nearest neighbor method is applied to match the corresponding feature points. A feature point will be considered as a matching point when R is calculated above 0.7.

B. Image Stitching

Once the sorting is over images are aligned in a sequence order. Then the first two images are stitched to gather. The resulting image and the third image is stitched afterword. Until the end of the image sequence the stitching is performed. Eq.4 is used to calculate the homography matrixes of the images.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

Where projection matrix H,

$$H = \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & h_8 \end{bmatrix}$$

Let x' *and* y' be the feature point coordinates of image 1 and, x and y be the feature point coordinates of image 2. The H matrix is the multiplier that represents the rotation and the scaling factors of image 2. By simplifying Eq. 4,

$$x' = \frac{h_0 x + h_1 x + \ h_2}{h_6 x + h_7 x + 1} \quad (5)$$

$$y' = \frac{h_3 y + h_4 y + \ h_5}{h_6 x + h_7 x + 1} \quad (6)$$

By solving homography matrixes, the images can be finally stitched to each other. As in Eq.4, it is clear that the transformation relation between the two images can be obtained using the corresponding match of feature pairs. By solving each homography matrixes the images can be stitched with each other.

RANSAC algorithm is used to match the points more accurately [14]. There are three main steps in the RANSAC algorithm [15]. Calculating H matrixes, divide the data into inliers and computing the optimal model parameters to extract homography matrix are the steps involved in RANSAC.

C. Fusion of images

Image fusion involves transforming the matching image into the coordinate system with reference to the reference image coordinate.

IV. Experiments and Results

Experiments are conducted for two different feature descriptors, namely SIFT and SURF. Visual Studio 2015 Community version along with OpenCV 2.4.13 was used as the main platform for experiments. Inria Aerial Image labeling dataset [16] was used in all the experiments.

An example set of images used with the proposed algorithm is given below in Figure 04.

Figure 4: Multiple Images

Results when using SIFT algorithm and SURF algorithm are shown below. Figure 5 shows the feature point mapping for SIFT features and Figure 6 shows the final mapped image. Similarly, Figure 7 shows the feature point mapping for SURF features and Figure 8 shows the final mapped image.

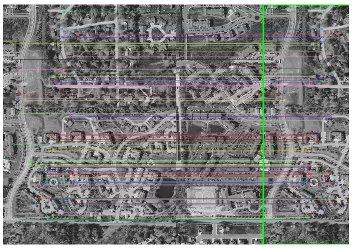

Figure 5: SIFT feature matching

Figure 6: Stitched image (SIFT)

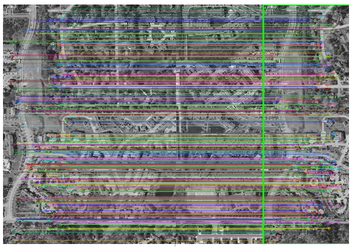

Figure 7: SURF Feature Matching

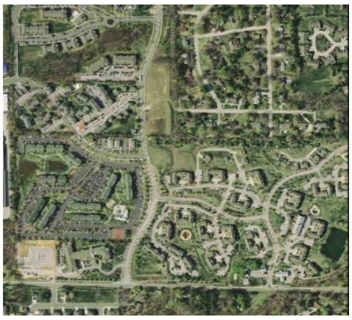

Figure 8: Stitched Image (SURF)

The above figures show that a large number of feature points can be mapped precisely with SURF features compared to SIFT. It clearly indicates that with low-quality images, SURF will be the best feature descriptor to be used in image mapping.

A. *Entropy-based results*

The entropy of an image indicates how dispersedly/compactly the color is distributed in RGB color space [17]. After stitching with SIFT and SURF features, Entropy was calculated to check the quality of the final stitched image. Eq 7 was used to measure the Entropy of images.

$$E = \sum_{i=1}^{L-1} P_i \log_2 P_i \tag{7}$$

Where P_i is the probability of occurrence of a color. Higher the Entropy, a quality image will be the output. The time taken for stitching also recorded and compared with SIFT and SURF. Table 1 shows the Entropy and the time taken for processing both SIFT and SURF features.

Using 120 images, 30 stitched images are constructed. Then the entropy is calculated for each stitched image. Figure 9 shows sample images taken to stitching process.

Figure 9: Sample Images before stitching

Table 1 and Table 2 presents the final results for each SIFT and SURF based algorithms. In each tables, 'No' indicates the resulting stitched image number (30 each). 'No. F.P' stands for Number of feature points and 'G.M.' is the number of good matching points. Information Entropy is denoted by 'Info. Ent'.

Table 1: SIFT Based Image-stitching Results.

No	Size (kb)	No. F.P	G.M	Info. Ent	Time
1	62.5	2533	1222	8.623	2.13
2	63.5	2554	1200	8.600	2.14
3	61.3	2530	1195	8.633	2.11
4	60.5	2553	1235	8.602	2.12
5	61.8	2545	1250	8.624	2.19
6	80.5	2504	1309	8.701	2.21
7	75.1	2552	1300	8.700	2.21
8	136.5	2533	1363	8.712	2.22
9	180.4	2553	1356	8.722	2.20
10	93.1	2553	1352	8.772	2.19
11	105.2	2550	1320	8.709	2.22
12	121.1	2553	1333	8.777	2.23
13	221	2535	1363	8.862	2.20
14	233.2	2504	1378	8.855	2.25
15	265.4	2532	1355	8.889	2.26
16	750.2	2543	1368	8.812	2.29
17	526.1	2556	1348	8.871	2.20
18	369.6	2533	1355	8.811	2.21
19	425.1	2530	1325	8.865	2.31
20	652.4	2553	1328	8.856	2.41
21	662.1	2545	1329	8.849	2.45
22	721.0	2503	1330	8.846	2.51
23	689.3	2532	1338	8.855	2.94
24	821.1	2533	1340	8.859	2.99
25	635.1	2533	1340	8.860	2.91
26	668.8	2530	1345	8.853	2.89
27	1203.1	2553	1359	8.866	3.01
28	1825.3	2545	1358	8.865	3.00
29	1500.1	2503	1360	8.870	2.98
30	1689.3	2530	1368	8.886	3.01

Table 2: SURF based image-stitching Results.

No	Size (kb)	No. F.P	G.M	Info. Ent	Time
1	62.5	3533	1528	9.024	2.39
2	63.5	3354	1530	9.033	2.38
3	61.3	3530	1595	9.045	2.38
4	60.5	3553	1535	9.049	2.39
5	61.8	3545	1550	9.050	2.40
6	80.5	3504	1609	9.051	2.41
7	75.1	3552	1600	9.053	2.44
8	136.5	3533	1663	9.070	2.41
9	180.4	3553	1656	9.098	2.40
10	93.1	3553	1752	9.102	2.42
11	105.2	3550	1720	9.110	2.40
12	121.1	3553	1733	9.122	2.43
13	221	3535	1763	9.125	2.41
14	233.2	3304	1778	9.129	2.40
15	265.4	3532	1755	9.135	2.45
16	750.2	3543	1868	9.150	2.44
17	526.1	3556	1848	9.169	2.45
18	369.6	3333	1855	9.175	2.46
19	425.1	3530	1825	9.199	2.46
20	652.4	3553	1828	9.290	2.49
21	662.1	3545	1829	9.350	2.48
22	721.0	3503	1830	9.351	2.49
23	689.3	3532	1938	9.382	2.45
24	821.1	3533	2040	9.401	2.48
25	635.1	3533	2240	9.420	2.47
26	668.8	3330	2445	9.511	2.49
27	1203.1	3553	2459	9.521	2.50
28	1825.3	3545	2558	9.549	2.51
29	1500.1	3503	2560	9.550	2.50
30	1689.3	3530	2627	9.592	2.51

Table 1 and Table 2 clearly shows that SURF outperforms SIFT in both good matches and Entropy. It is clear that SURF produces a large number of feature points in general.

V. CONCLUSION

This paper proposed an effective and efficient technique for aerial image stitching. The suggested algorithm was compared with the well-known SIFT features and results are reported. The proposed algorithm based on SURF, PCA and RANSAC give promising results in image stitching against the traditional SIFT base features. Experimental results clearly show that SURF is always better when it comes to feature extraction. It always generates a large number of feature points compared to the SIFT. In addition, the amount of good matches in SURF is also greater than the SIFT. When it comes to image quality, the Entropy is fairly greater in SURF based algorithms. Since the time taken for stitching is promising with SURF, the possibility of applying this algorithm in real-time applications will be considered in the future.

VI. REFERENCES

[1] N. Zhao and X. Zheng, "Multi-band blending of aerial images using GPU acceleration," in *10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Shanghai, China, 2018.

[2] T. Bergen and T. Wittenberg, "Stitching and Surface Reconstruction From Endoscopic Image Sequences: A Review of Applications and Methods," *Journal of Biomedical and Health Informatics,* vol. 20, no. 1, pp. 304 - 321, 2014.

[3] Y. Zhao , Z. Lu and Y. Liu, "Video image stitching based on moving object detection and motion prediction compensation," in *3rd International Congress on Image and Signal Processing*, Yantai, China, 2010.

[4] Y.Wang and M. Wang, "Research on stitching technique of medical infrared images," in *International Conference on Computer Application and System Modeling (ICCASM 2010)*, Taiyuan, China, 2010.

[5] J. Rühaak ,L. König,M. Hallmann,N. Papenberg,S. Heldmann, H. Schumacher and B. Fischer, "A fully parallel algorithm for multimodal image registration using normalized gradient fields," in *IEEE 10th International Symposium on Biomedical Imaging*, San Francisco, CA, USA, 2013.

[6] G. Gerganov , A. Papucharov , I. Kawrakow and K. Mitev, "Portal image registration using the phase correlation method," in *IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*, Seoul, South Korea, 2014.

[7] G. Zhao and Q. Song, "SIFT Image Stitching Technology Based on Music Score Scanning Recognition System," in *9th International Symposium on Computational Intelligence and Design (ISCID)*, Hangzhou, China, 2016.

[8] Z. Jiang, J. Wu, D. Cui , T. Liu and X. Tong, "Stitching method for distorted image based on SIFT feature matching," in *8th International Conference on Computing and Networking Technology (INC, ICCIS and ICMIC)*, Gueongju, South Korea, 2012.

[9] L.Zhu, Y. Wang ,B. Zhao and X. Zhang, "A Fast Image Stitching Algorithm Based on Improved SURF," in *Tenth International Conference on Computational Intelligence and Security*, Kunming, China, 2015.

[10] Z. Wang , F. Yan and Y. Zheng, "An adaptive uniform distribution surf for image stitching," in *6th International Congress on Image and Signal Processing (CISP)*, Hangzhou, China, 2014.

[11] Y. Li , Q .Pan , C. Zhao , H. Liu and J. Zhang, "SVD-SURF based fast and robust scene matching algorithm," in *Proceedings of the 32nd Chinese Control Conference*, Xi'an, China, 2013.

[12] M. S. Patel , N. M. Patel and M. S. Holia, "Feature based multi-view image registration using SURF," in *International Symposium on Advanced Computing and Communication (ISACC)*, Silchar, India, 2016.

[13] Woodruff, Caleb, "https://courses.cs.washington.edu," Computer Vision, CSE 576, 1 3 2013. [Online]. Available: https://courses.cs.washington.edu/courses/cse576/13sp/projects/project1/artifacts/woodrc/index.htm. [Accessed 16 10 2018].

[14] G. Shi ,X. Xu and Y. Dai, "SIFT Feature Point Matching Based on Improved RANSAC Algorithm," in *5th International Conference on Intelligent Human-Machine Systems and Cybernetics*, Hangzhou, China, 2013.

[15] M. Zhao , H. Chen , T. Song and S. Deng, "Research on image matching based on improved RANSAC-SIFT algorithm," in *16th International Conference on Optical Communications and Networks (ICOCN)*, Wuzhen, China, 2017.

[16] E. Maggiori, Y. Tarabalka, G. Charpiat and P. Alliez., "Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017.

[17] M.Sandholze., "researchgate," 1 1 2014. [Online]. Available:https://www.researchgate.net/post/How_to_measure_the_level_of_camouflage. [Accessed 18 10 2018].

[18] B Zitová, J Flusser, "Image registration methods: a survey," *Image and Vision computing,* vol. 21, no. 11, pp. 977-1000, 2003.