International Journal of Trend in Research and Development, Volume 4(3), ISSN: 2394-9333
www.ijtrd.com

# Geometric Correction of Image Acquired by UAV Using Non-parametric Approach

[1]Thet Su Win and [2]Sao Hone Pha
[1,2]Department of Electronics Engineering, Yangon Technological University, Yangon, Myanmar

***Abstract***— UAV photography system is implementation of efficient data acquisition techniques, high resolution texture, low cost and using in commercial application. In research paper, we got several images from UAV (unmanned aerial vehicle) and produced image mosaic. Firstly, this paper proposes image mosaic method based of SIFT (scale invariance feature transform) feature to detect key points, scaling, rotation and matching between two images in the image mosaic. RANSAC (random sample consensus) method is used to find homography, transformation and adjust colour for rgb or grey scale for grey image with SIFT (scale invariance feature transform) matching location and then produced mosaic image from combination of SIFT (scale invariance feature transform) and RANSAC (random sample consensus). The results from experiment based on four pairs of images with 70% overlapped captures by using the camera of UAV show that our method has many feature points and matching points for image mosaic. In this paper, the combination of SIFT and RANSAC algorithm is used to produce mosaic image and then this image is used in geometric correction.

Finally, this paper proposes before and after geometric correction of mosaic image because raw digital images contain geometric distortion and cannot be used directly as a map. In general, there are two approaches for the geometric correction. The parametric approach is model-based while the non-parametric one makes use of ground control points (GCPs). This paper contains the geometric correction of UAV image which is non-parametric approach.

In this paper, geometric correction is considered two conditions. These conditions are that distributions of GCPs (Ground Control Points) are considered with distinct location and indistinct location based on geometric metric correction accuracy that is Root Mean Square Error (RMSE).

***Keywords*** *– UAV; SIFT; RANSAC; mosaic; geometric correction; non-parametric; RMSE.*

## I. INTRODUCTION

UAV system is high flexibility, large scale and high accuracy. UAV photography is widely used in agricultural monitoring, disaster monitoring, documentation of archaeological sites and cultural heritage. UAV photography is limited in flying high, overlap percentage, power, time, weather and small scale area. Single image cannot cover the entire target area [7]. So, image mosaic becomes a key technology to solve this problem [8]. If we compare unmanned aerial vehicles (UAVs) and manned aircrafts, we will see UAVs have many advantages these are 1) UAVs do not need a qualified pilot on board; 2) UAVs can enter dangerous zone; 3) It is easy to implement high-risk and high-tech missions; 4) UAVs can stay in the air for up to 30 hours, performing a precise, repetitive raster scan of a region, day-after-day, night- after-night in complete darkness or in fog; 5) Lower cost of the platforms; 6) High resolution images and precise position data [1]. Therefore, UAV technique becomes very popular in commercial application and UAV based photography is useful in many area.

In general, geometric correction is considered as a two-step procedure. First, distortions are considered as systematic, or predictable. Second, those distortions are considered as essentially random, or unpredictable. Systematic distortions corrected by using formulas derived as modelling the sources of the distortions mathematically. By analysing well-distributed ground control points (GCPs) occurring in an image, random distortions and residual unknown systematic distortions are corrected.

## II. OVERVIEWS AND METHODOLOGY

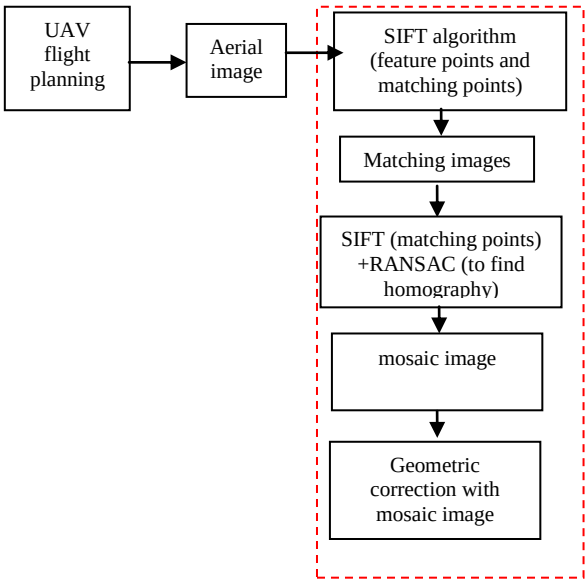


Fig. 1. Block diagram of geometric correction with mosaic image

### *A. SIFT (Scale Invariance Feature) Algorithm*

In 1999, SIFT operator was first put forward by David G. Lowe at British Columbia university, and it was summarized and improved in 2004.

SIFT operator looking for extreme value point in scale space, and extract position, dimension and rotation invariants [9]. The major steps in the computation of SIFT algorithm are as follow.

1) Scale Space Construction
2) Key Point Localization
3) Orientation Assignment
4) Key Point Descriptor and Matching

*1) Scale Space Construction:* Scale space construction is construction of Gaussian and Difference-of-Gaussian pyramids [6]. The first step is to realize image location Maintaining the Integrity of the Specifications coordinates and scale that can be repeatedly assigned under pose variation of the object of interest. Finding locations that are invariant to scale is

performed by scale function that searches for stable features across difference scales. The scale space convolution kernel of choice is the Gaussian used to define the scale space function of an input image according to

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2}\exp[-(x^2 + y^2)/2\sigma^2)] \quad (2)$$

Where, $(x, y)$ represents point coordinates; $\sigma$ is scale space, $G(x, y, \sigma)$ is Gaussian function; $L(x, y, \sigma)$ is scale space; $I(x, y)$ is input image; "$*$" represents convolution operation. The different scale space will be established with the change of parameter $\sigma$.

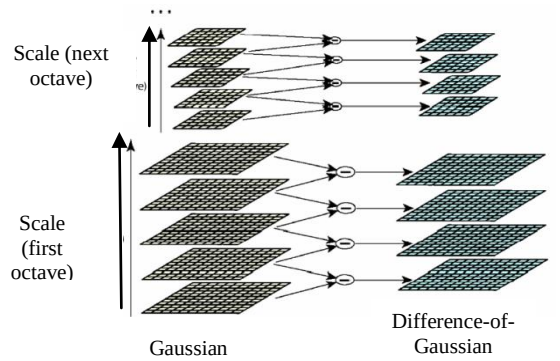


Fig. 2. Gaussian and Difference-of-Gaussian (DOG)

To detect stable key point locations in scale space, Difference-of-Gaussian (DoG) function convoluted with the image $D(x, y, \sigma)$ is computed from the difference of two nearby scales separated by a constant multiplicative factor k as in

$$D(x, y, \sigma) = \big(G(x, y, k\sigma) - G(x, y, \sigma)\big) * I(x, y) \quad (3)$$

$$= \mathrm{L}(x, y, k\sigma) - L(x, y, \sigma) \quad (4)$$

The Difference-of-Gaussian (DoG) function is a close approximation to the scale normalized Laplacian of Gaussian $\sigma^2\nabla^2 G$ [6]. Fig. 2 shows Gaussian and Difference-of-Gaussian. Fig. 3 shows one interval of local extrema computation which uses 3 levels of Difference-of-Gaussian (DoG) functions [6].

The scale space is divided into octaves. An octave represents a series of filter response maps obtained by convolving the same input image with a filter of increasing size. In total, an octave encompasses a scaling factor of 2.

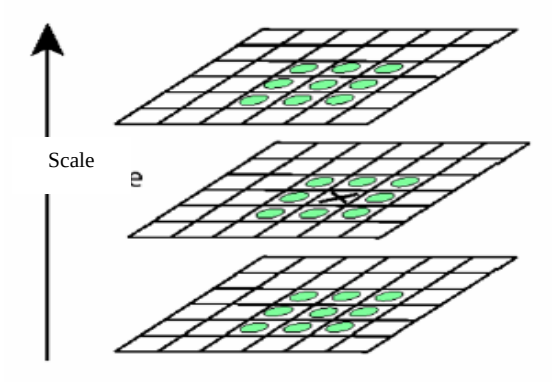


Fig. 3. One interval of local extrema detection

The construction of the scale space starts with 9×9 filter, which calculates the blob response of the image for the smallest scale. Then, filters with sizes 15×15, 21×21 and 27×27 are applied. Fig. 4 shows graphic representation of the filter side lengths for three different octaves.

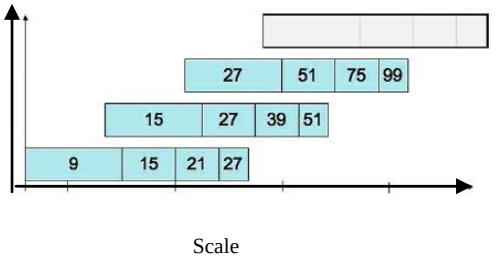


Fig. 4. Graphic representation of the filter side lengths for three different octaves

For each new octave, the filter size increase is doubled (going from 6 to 12 to 24-48). At the same time the sampling intervals for the location of the interest point can be doubled as well for new octave.

The filter sizes for the second octaves are 15, 27, 39 and 51. A third octave is computed with the filter size 27, 51, 75 and 99 and if the original image size is still larger than the corresponding filter sizes, the scale space analysis is performed for a fourth octave, using the filter sizes 51, 99, 147 and 195 [2].

*2) Key Point Localization*: Key point candidates are chosen from the extrema in the scale space, and key points are selected based on measures of their stability [6].

3) *Orientation Assignment*: Orientations are assigned to each key point based on histograms of gradient directions computed in 16×16 window [6]. The scale of the key point is used to select the Gaussian smoothed image, L, with the closet scale, so that all computations are performed in a scale-invariant manner [10]. For each image sample, L(x,y), at this scale, the gradient magnitude, m(x,y), and orientation, ɵ(x,y), is pre computed using pixel differences:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (5)$$

$$\theta(x, y) = tan^{-1}\frac{(L(x,y+1) - L(x,y-1))}{(L(x+1,y) - L(x-1,y))} \quad (6)$$

where, $m(x, y)$ is the gradient magnitude and $\theta(x, y)$ is the orientation.

An orientation histogram is formed from the gradient orientations of sample points within a region around the key point.

Each sample added to the histogram is weighted by its gradient magnitude and by a Gaussian-weighted circular window with a $\sigma$ that is 1.5 times that of the scale of the key point [10].

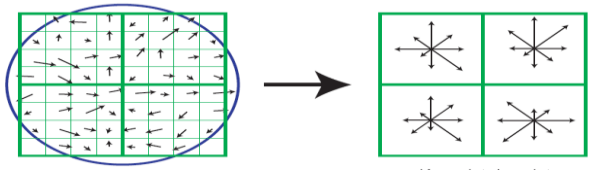


Fig. 5. Example of image gradients and key point descriptor

*4) Key Point Descriptor and Matching*: Representation in 128-dimensional vector and the best candidate match is found by its nearest neighbour [6]. The key point descriptor is shown

on the right side of Fig. 5. It allows for significant shift in gradient positions by creating orientation histograms over 4×4 sample regions. The Fig. 5 shows eight directions for each orientation histogram, with the length of each arrow corresponding to the magnitude of the histogram.

The descriptor is formed from a vector containing the values of all the orientation histogram entries, corresponding to the lengths of the arrows on the right side of Fig. 3. The Fig. 3 shows a 2×2 array of orientation histograms. SIFT descriptor has 4×4 gradient window and histogram of 4×4 samples per window in 8 directions. Therefore, SIFT descriptor has 4×4×8=128 dimension feature vector [10].

Key point matching is that Lowe used a modification of the k-d tree algorithm called the BBF (Best-bin-first) search method that can identify the nearest neighbours with high probability using only a limited amount of computation. The BBF (Best-bin-first) algorithm uses a modified search ordering for the k-d tree algorithm so that bins in feature space are searched in the order of their closest distance from the query location. The best candidate match for each key point is found by identifying its nearest neighbour in the database of key points from training images. The nearest neighbours are defined as the key points with minimum Euclidean distance from the given descriptor vector.

The probability that a match is correct can be determined by taking the ratio of distance from the closest neighbour to the distance from the closest neighbour to the distance of the second closest [5].

Find a feature point from the first image, then two points from the second image, between which Euclidean distance is the nearest. If the ratio that the smallest distance is divided by the secondary one is smaller than the threshold, this point will match the nearest one successfully [12].

***B. RANSAC (Random Sample Consensus) Algorithm***

RANSAC (Random Sample Consensus) is an iterative method to estimate parameters of a mathematical model from a set of observed data that contains outliers, when outliers are to be accorded no influence on the values of the estimates. Therefore, it also can be interpreted as an outlier detection method. It is a non-deterministic algorithm in the sense that it produces a reasonable result only with a certain probability, with this probability increasing as more iteration are allowed. The algorithm was first published by Fischler and Bolles at SRI international in 1981. They used RANSAC (Random Sample Consensus) to solve the Location Determination Problem (LDP), where the goal is to determine the points in the space that project onto an image into a set of landmarks with known locations [4].

Parameters are that like described in the beginning of the algorithm flow, RANSAC (Random Sample Consensus) needs some predefined parameters to run - size of sample subset, error tolerance threshold, minimum consensus threshold and number of iterations. In addition, it is important to estimate the proportion of inliers in the data set, in order to calculate some of these parameters. [11].

Proportions of inliers are used in the calculations of algorithm parameters and even it can affect the algorithm complexity implicitly.

Hence, it is beneficial to have some information about the outlier rate in the data set before running RANSAC (Random Sample Consensus) [11].Size of sample subset is the number of samples that are randomly chosen by RANSAC (Random Sample Consensus) to initial model at each iteration.

It is directly related with the model that is intended to fit the data set. RANSAC (Random Sample Consensus) uses the minimum number of samples needed to define the model as the sample subset size. Example of in order to fit a linear model it chooses 2 data samples or to fit a circle–shaped model it selects 3 data samples as 3 points would be sufficient to define a circle [11].

Error tolerance threshold is used by RANSAC in order to determine if a data sample agrees with a model or not. The samples under this threshold would then form that consensus for that model, which would be the inliers of the data set if the correct model is found. Hence, it should be chosen according to the Gaussian error in the inliers [11].

Minimum consensus threshold is the minimum number of samples that would be accepted as a valid consensus to generate a final model for that iteration. As RANSAC (Random Sample Consensus) tries to capture the inliers with constituting consensuses, the number of samples in a valid consensus is directly related with the number of inliers in the data set. Hence, RANSAC (Random Sample Consensus) uses a threshold value that is equal to or a bit smaller tolerated than the number of inliers in order to accept consensus as valid [11].

Numbers of iterations are that exhaustive deterministic algorithms would try every possible sample subset in order to find the best one. Hence, instead of a deterministic way, RANSAC (Random Sample Consensus) chooses sample subset randomly. But, it is also important to determine the number of these random choices in order to obtain a high probability [11]. Homography estimation using RANSAC (Random Sample Consensus) is that two feature points are matched when their difference is small, although mismatching may be created in this method. The RANSAC algorithm is used to correct these errors. It divides all points to outliers and inliers. The outlier will be rejected by RANSAC (Random Sample Consensus) algorithm and inliers will be used to estimate model and evaluate the model by error rate [8].

The procedure of RANSAC (Random Sample Consensus) is

1. Select randomly the minimum number of points required to determine the model parameters.
2. Solve for the parameters of the model.
3. Determine how many points from the set of all points fix with a predefined tolerance.
4. If the fraction of a number of inliers over the total number points in the set exceeds a predefined threshold, re estimate the model parameter using all the identified inliers and terminate.
5. Otherwise, repeat steps 1 through 4 (maximum of N times).

The number of iterations, N, is chosen high enough to ensure that the probability p (usually set to 0.99) that at least one of the sets of random samples does not include an outlier.

Let u represent the probability that any selected data point is an inlier and v=1-u the probability of observing an outlier. *N* iterations of the minimum number of point denoted *m* are required, where

$$1 - p = (1 - u^m)^N \tag{7}$$

And thus with some manipulation,

$$N = \frac{\log(1-p)}{\log(1-(1-v)^m)} \tag{8}$$

Extensions of RANSAC (Random Sample Consensus) include using a Maximum Likelihood frame work and importance sampling [3].

### *C. Geometric Correction Method*

In geometric correction, minimum number of ground control point (GCP) is used the following equation.

$$\frac{((t+1)(t+2))}{2} \quad (9)$$

Where, t is the transformation of order [13].

Resampling process is consisted nearest neighbor method, bilinear interpolation method, cubic convolution and bicubic Spline Interpolation. Rectification is the process of transforming the data from one grid system into another grid system using a geometric transformation [13].

Polynomial transformation and triangle-based methods are various rectification techniques. Since the pixels of the new grid may not align with the pixels of the original grid, the pixels must be resampled. Resampling is the process of extrapolating data values for the pixels on the new grid from the values of the source pixels [13].

RMS error is the distance between the input (source) location of a GCP and the retransformed location for the same GCP. In other words, it is the difference between the desired output coordinate for a GCP and the actual output coordinate for the same point, when the point is transformed with the geometric transformation. RMS error is calculated with a distance equation:

$$RMS\ error = \sqrt{(x_r - x_i)^2 + (y_r - y_i)^2} \quad (10)$$

Where, $x_i$ and $y_i$ are the input source coordinates and $x_r$ and $y_r$ are the retransformed coordinates [14].

## III. EXPERIMENTAL RESULT

Fig. 6. Testing of Phanton-3 UAV

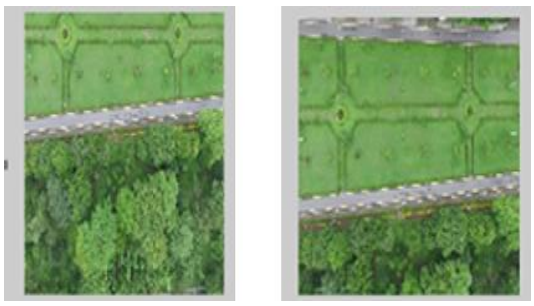

Image 1 Image 2

Fig. 7. Images 1 and 2 from UAV photography

In research paper, UAV is designed with the specifications as Phanton-3 UAV, overlap 70%, altitude 60m, speed 4.83m/s, flying time 5 minutes 03 second and camera 12MG pixel. Flying area, Yangon Technological University, Myanmar area, is used 8 images in 59*180 m. In Fig. 6, Phanton-3 UAV is used in research.

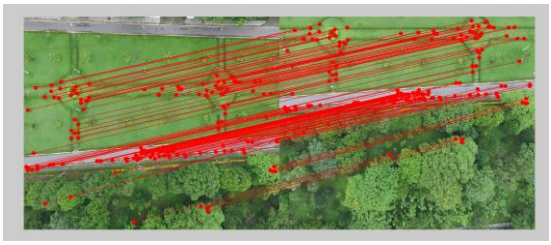

Fig. 8. Matching points of Fig. 7

Image 3 Image 4

Fig. 9. Images 3 and 4 from UAV photography

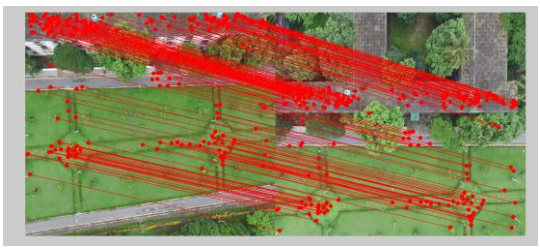

Fig. 10. Matching points of Fig. 9

Image 5 Image 6

Fig. 11. Images 5 and 6 from UAV photography

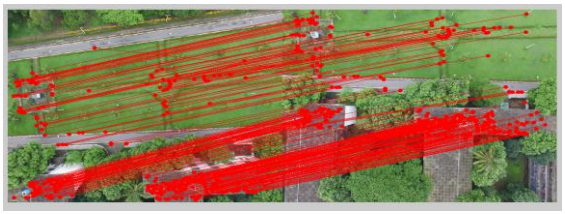

Fig. 12. Matching points of Fig. 11

Image 7 Image 8

Fig. 13. Images 7 and 8 from UAV photography

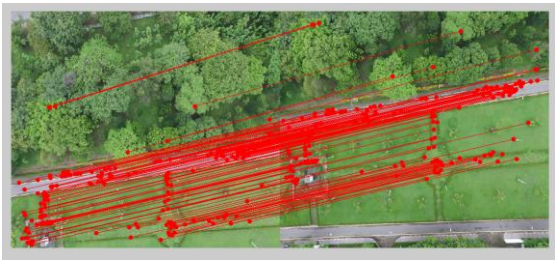

Fig. 14. Matching points of Fig. 13

Fig. 15. Mosaic image with SIFT + RANSAC algorithm for Fig. 7, Fig. 9. Fig 11 and Fig. 13

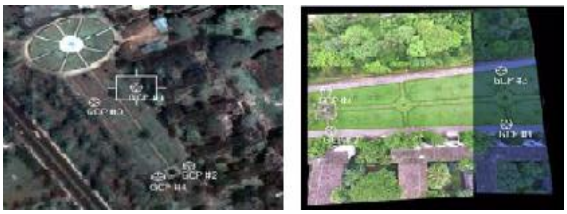

Fig. 16. Condition 1, GCPs with distinct location on reference image and distorted image with non-parametric approach

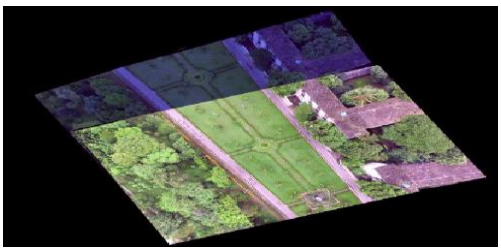

Fig. 17. After geometric correction of mosaic image for Fig. 15

Condition 1 (GCPs with distinct location) with non-parametric approach

Fig. 18. Condition 2, GCPs with indistinct location on reference image and distorted image with non-parametric approach

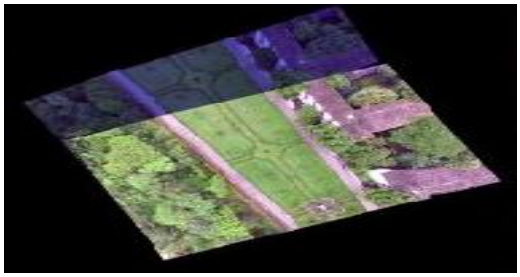

Fig. 19. After geometric correction of mosaic image for Fig. 15
Condition 2 (GCPs with indistinct location) with non-parametric approach

GCPs with distinct location and indistinct location based on geometric correction accuracy that is Root Mean Square (RMS) Error are shown in TABLE.1. Polynomial order 1 is used because of small area.

Table 1: Root mean square (RMS) error result for condition 1 and condition 2.

| GCPs Conditions | 1 | 2 |
|---|---|---|
| Polynomial Order | 1 | 1 |
| Minimum Requirement of GCPs | 3 | 3 |
| Total Number of GCPs Used | 4 | 4 |
| Control points Error | 0.0384 | 0.8964 |
| RMS Error | 0.152 | 3.381 |

In experimental result, combination of SIFT (Scale Invariance Feature Transform) algorithm and RANSAC (Random Sample Consensus) algorithm are used to produce mosaic image.

Testing images from UAV photography are shown in Fig. 7, Fig. 9, Fig. 11 and Fig.13. The matching points of Fig. 9 are shown in Fig. 10.

The matching points of Fig. 11 are shown in Fig.12. And then the matching points of Fig. 13 are shown in Fig. 14. Before geometric correction of the mosaic image of Fig. 7, Fig. 9, Fig.11 and Fig.13 is shown in Fig.15.

Condition 1, GCPs with distinct location on reference image from Geo Eye image and distorted image from UAV mosaic image are shown in Fig.16 with non-parametric approach and condition 2, GCPs with indistinct location on reference image from Geo Eye image and distorted image from UAV mosaic image are shown in Fig.18 non-parametric approach.

GCPs with distinct location of after geometric correction of the mosaic image of Fig. 15 is shown in Fig. 17 as condition 1 and GCPs with indistinct location of after geometric correction of the mosaic image of Fig. 15 is shown in Fig.19 as condition 2 which has UTM coordinate using ERDAS IMAGINE software.

## CONCLUSIONS

2877 key points on the left and 4951 key points on the right are found from Fig. 7, 3107 key points on the left and 4928 key points on the right from Fig. 9. 184 matching points are found in Fig. 8 and 263 matching points are found in Fig. 10.

And then processing time are found 24.662138 seconds, 28.611765 seconds, of 33.816515 seconds and 26.276118 seconds to produce mosaic for Fig.7, Fig. 9, Fig.11 and Fig.13.

3893 key points on the left and 5662 key points on the right from Fig. 11, 5726 key points on the left and 3592 key points on the right from Fig. 13, 350 matching points in Fig. 12 and 209 matching points in Fig. 14 are found.

Fig. 15, Fig. 17 and Fig.19 are shown the comparison of before and after geometric correction. The result of geometric correction accuracy that is root mean square (RMS) Error is shown in TABLE. 1.

First condition is that GCPs are located with distinct location (easily and accurately identifiable) such as road intersection in both images.

Second condition is that GCPs are located with indistinct location (difficult identifiable) in both images.

Finally, the result based on geometric metric correction accuracy is root mean square (RMS) Error. In condition 1, root mean square (RMS) error is low in TABLE. 1 where GCPs distribution is distinct locations are at road intersection. In condition 2, root mean square (RMS) error is high in TABLE. 1where GCPs distribution is indistinct locations that are not easily identifiable.

In conclusion, based on these result, mosaic image are extracted by SIFT (Scale Invariance Feature Transform) algorithm with many feature points and matching points and RANSAC (Random Sample Consensus) algorithm with finding homography, transformation and adjust colour for rgb or grey scale for grey image.

In geometric correction result, the accuracy of geometric correction depends on the GCPs location of distinct and indistinct conditions with non-parametric approach.

In further extension, researchers' can do comparison with other methods for key point detection, matching, error percentage and geometric correction with parametric approach for accuracy.

### *Acknowledgment*

We would like to express my gratitude to Dr. Myint Thein, Rector of the Yangon Technological University. We would like to express our gratitude to our supervisor Dr. Sao Hone Pha, Professor, for the continuous support of our Ph.D study and related research, for his patience, generous advice, inspiring guidance, motivation and immense knowledge. His guidance helped us in all the time of research and writing of this paper. We would like to express our gratitude to Dr. Yu Yu Lwin, Head of the Department of Electronics Engineering in Yangon Technological University, for the useful comments and suggestions throughout this research. We would like to appreciate to all of our teachers in the Department of Electronics Engineering in Yangon Technological University and everyone who supported us throughout this research.

### *References*

[1] Cheng Xing, Jinling Wang and Yaming Xu, "A Robust Method for Mosaicking Sequence Images Obtained from UAV", *IEEE*, Univ of Wuhan, China and Univ of New South Wales, Sydney, Australia, 2010.

[2] Herbert Bay, Andreas Ess, Tine Tuytelaars and Luc Van Gool, "Speed-Up Robust Features (SURF)", *Computer Vision and Image Understanding* 110. pp 346-359. 2008.

[3] Konstantinos G Derpanis, *Overviews of the RANSAC Algorithm*, May. 2010.

[4] (2016) RANSAC Wikipedia website [Online]. Available: https://www.google.com/,https://en.wikipedia.org/wiki/ (Random- sample-consensus).

[5] (2016) SIFT Wikipedia website [Online]. Available: https://www.google.com/,https://en.wikipedia.org/wiki/ (Scale- Invariance-Feature-Transform).

[6] Naotoshi Seo and David A. Schug, "Image Matching using Scale Invariant Feature Transform (SIFT)", *ENEE631 Digital Image and Video Processing, Final Project*, Univ. of Maryland.

[7] Yinjiang Jia, Zhongbin Su, Qi Zhang, Yu Zhang1 Yunhao Gu and Zhongqiu Chen, "Research on UAV Remote Sensing Image Mosaic Method Based on SIFT", *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 8, no-11, pp. 365-374, 2015.

[8] Bahareh Kalantar Ghorashi Harandi, Shattri Bin Mansor and Helmi Zulhaidi M. Shafril, *Image Mosaic Methods for UAV Sequences Inages*, Univ. of Putra Malaysia.

[9] Ming Lia, Deren Li and Dengke Fan, " A study on Automotive UAV Image Mosaic Method for Paroxysmal Disaster", *International Achieves of the Photogrammetry, Remote Sensing and Information Sciences*, vol. xxx1x-B6,pp. 123-128, Sep.2012.

[10] David G. Lowe, *Distinctive Image Features from Scale-Invariant Key points*, International Journal of Computer Vision, Jan.2004.

[11] *RANSAC tutorial*, Apr. 2006.

[12] LI Chang-chun, Zhang Guang-sheng, LEI Tian-jie and GONG A-du, "Quick Image-Processing Method of UAV without Control Points Data in Earthquake Disaster Area", *Trans. Nonferrous Met. Soc. China*, pp. 523-528, Nov.2011.

[13] Available: https://www.google.com/,https://en.wikipedia.org/wiki/(Geometric Correction Method).