PERFORMANCE ANALYSIS OF SELECTED FEATURE DESCRIPTORS USED FOR AUTOMATIC IMAGE REGISTRATION

AJAYI Oluibukun Gbenga*
Department of Surveying and Geoinformatics, Federal University of Technology, Minna, Nigeria, ogbajayi@gmail.com, gbenga.ajayi@futminna.edu.ng

KEY WORDS: Feature Descriptors, Image Registration, Conjugate Points, Corresponding features, Scale Invariance.

ABSTRACT:

Automatic detection and extraction of corresponding features is very crucial in the development of an automatic image registration algorithm. Different feature descriptors have been developed and implemented in image registration and other disciplines. These descriptors affect the speed of feature extraction and the measure of extracted conjugate features, which affects the processing speed and overall accuracy of the registration scheme. This article is aimed at reviewing the performance of most-widely implemented feature descriptors in an automatic image registration scheme. Ten (10) descriptors were selected and analysed under seven (7) conditions viz: Invariance to rotation, scale and zoom, their robustness, repeatability, localization and efficiency using UAV acquired images. The analysis shows that though four (4) descriptors performed better than the other Six (6), no single feature descriptor can be affirmed to be the best, as different descriptors perform differently under different conditions. The Modified Harris and Stephen Corner Detector (MHCD) proved to be invariant to scale and zoom while it is excellent in robustness, repeatability, localization and efficiency, but it is variant to rotation. Also, the Scale Invariant feature Transform (SIFT), Speeded Up Robust Features (SURF) and the Maximally Stable Extremal Region (MSER) algorithms proved to be invariant to scale, zoom and rotation, and very good in terms of repeatability, localization and efficiency, though MSER proved to be not as robust as SIFT and SURF. The implication of the findings of this research is that the choice of feature descriptors must be informed by the imaging conditions of the image registration analysts.

1. INTRODUCTION

This article is aimed at providing an empirical review of the strength and weaknesses of the most implemented feature descriptors as used in automatic registration of overlapping images. The analysed descriptors are the Scale Invariant Feature Transform (SIFT), the Speeded Up Robust Features (SURF), Modified Harris and Stephens Corner Detector (MHCD), the Maximally Stable Extremal Regions (MSER), and the Features from Accelerated Segment Test (FAST). Others are Smallest Uni-value Segment Assimilating Nucleus (SUSAN), Fast Retina Key point (FREAK), Hessian, Difference of Gaussian and the Hessian-Laplace algorithms. The review first provided a broad overview of feature detection and extraction, before providing a summary of the characteristics of some of the feature descriptors. It further analysed the qualities of the selected descriptors under seven (7) conditions which are Invariance to rotation, scale and zoom, their robustness, repeatability, localization and efficiency using UAV acquired images. Finally, details of the procedures of implementing the three descriptors adjudged to outperform others were provided and experimental findings of the performance evaluation were presented.

1.1 Feature detection and extraction

In image processing, images are generally represented by the features that can be extracted from them. These features are broadly categorised into two, namely, the global features and local features while the extraction of these image features can also be categorised into both high-level features and low-level features (Nixon and Aguado, 2008).

The global feature representation depicts the image as one multi-dimensional feature vector which describes the whole image. More specifically, the global feature representation approach produces one single vector with values that measure various part of the image such as tone, texture, pattern, shape (Hassaballah *et al.*, 2016). Though global feature representation is generally fast, simple to compute and requires small amount of memory, they are also notably limited. Specifically, they are variant to transformations and are very sensitive to occlusion and blurs. In local feature representation, images are distinctively represented based on their local structures using local features which are also known as key points or interest points, and can be described as specific and unique patterns that are distinct from the pixels within its neighbourhood (Tuytelaars and Mikolajczyk, 2007) and are generally associated with one or more properties of the image (Li *et al.*, 2015). They are points with a well-defined position in the image space, unambiguous mathematical description, and they are stable under perturbations such as variations in brightness (Mubarak, 1997). Examples of such features are regions, edges, and corners. When compared to global feature representation, the local features are notable for superior performance, distinctiveness and better stability (Jégou *et al.*, 2012) though they require significant amount of memory because many local features can be found on a single image. The advantages of local feature representation make it more suitable for object recognition and image matching (Hassaballah *et al.*, 2016).

Ideally, local features are expected to have the following qualities or characteristics: distinctiveness, locality,

* O. G. Ajayi; ogbajayi@gmail.com

accuracy, quantity, efficiency, repeatability, invariance and robustness which attests to their less sensitivity to noise or blurs (Ehab and Murad, 2017). These qualities are also expected to be inherent in the formulation of feature detection and extraction algorithms which are the algorithms that detect and extract these features and prepares them for further applications in image registration. They are also referred to as feature descriptors which are described also as the methods that are used in the computation of abstractions of the information on an image pair, which is used in making informed decisions of the identity of every image point on an image, whether there is an image feature of a particular type or not.

Feature-based descriptors are broadly categorised into:

1. Spatial relations (Tuytelaars, 2006).
2. Edge based (differentiation based) descriptors such as Canny and Sobel.
3. Corner based (gradient based) descriptors such as Harris and Stephens descriptors and its derivatives.
4. Corner based (template based) descriptors such as Features from Accelerated Segment Test (FAST), Smallest Uni-value Segment Assimilating Nucleus (SUSAN) and Binary Robust Independent Elementary Features (BRIEF) which belongs to the family of binary descriptors.
5. Corner based (contour based) descriptors such as hyperbola fitting.
6. Invariant descriptors (Kazhdan *et al.*, 2003; Tuytelaars, 2006) such as Scale Invariant Feature Transform (SIFT) algorithm developed by Lowe (2004), and Brown and Lowe (2007), Speeded Up Robust Features (SURF) descriptor proposed by Bay *et al.* (2006, 2008).
7. Blob (interest region) and salient regions such as the Maximally Stable Extremal Region (MSER) algorithm developed by Matas *et al.* (2004).
8. Blob (Key point) such as Fast Retina Key point (FREAK), and Binary Robust Invariant Scalable Key point (BRISK) and the Accelerated Binary Robust Invariant Scalable Key point (ABRISK) which detects binary features only. They are designed basically for tracking and not for providing solutions to image classification problem.

1.2 Characteristics of selected feature descriptors

The major characteristics of some of the selected feature descriptors are presented in Table 1 while the result of the performance evaluation of the Ten (10) descriptors is presented in Table 2. The performance evaluation shows that while some of the descriptors are invariant to the trio of scale, rotation and zoom (SIFT, SURF MSER, etc), others are only invariant to either of them. The analysis also shows that only MHCD is excellent in terms of robustness, repeatability, efficiency, and localization. Other algorithms are also very good under these four (4) conditions except for FAST and FREAK.

1.3 Implementing MHCD, SIFT and SURF

The basic steps involved in the implementation of the MHCD, SIFT and SURF are discussed under this section and each of the algorithms are discussed in the following subsections:

1.3.1 Modified Harris corner detection (MHCD) algorithm: While MHCD is partially invariant to affine intensity change, it is non-invariant to spatial scale. The activity diagram depicting the algorithmic stages of implementing the Modified Harris Corner detection (MHCD) algorithm is presented in Figure 1 while the step-by-step procedure of the algorithm's implementation are as follows:

Step 1. Computation of horizontal and vertical derivatives of the stereo image.
Step 2. Computation of three images corresponding to the three terms in matrix M.
Step 3. Convolving these three images with a large Gaussian window.
Step 4. Computation of scalar corner response using one of the corner response measure.
Step 5. Finding local maxima above some predefined threshold as detected interest points.
Step 6. Computation of SURF descriptor around detected interest points.
Step 7. Matching the corresponding points based on the descriptor difference.
Step 8. Filtering out the outliers from matched points using RANSAC algorithm.

1.3.2 Scale invariant feature transform algorithm (SIFT): The SIFT descriptor is a vector of 128 values, each between [0 - 1]. Its feature point is associated with location, orientation and scale (Lowe, 2004). It is invariant to image rotation, scale, intensity change, and to moderate affine transformations. Figure 2 presents the activity diagram showing the implementation stages of the SIFT algorithm while the step by step procedure are as described in the following steps:

Step 1. Detection of key points: Locally distinct points over different image pyramid levels were detected by:

a. Applying Gaussian smoothing,
b. Using Difference-of-Gaussians (DoG) to find extrema (over smoothing scales),
c. Maxima suppression at edges.

Step 2. Computation of SIFT descriptor which transformed image content into features that are invariant to scaling, image translation, and rotation by:

i. Computing image gradients in local 16x16 area at the selected scale,
ii. Creation of an array of orientation histograms; 8 orientations $\times 4 \times 4$ histogram array of 128 dimensions (yields best result).

Step 3. Matching of the corresponding points based on the descriptor difference.
Step 4. Filtering out outliers from matched points using RANSAC algorithm.

1.3.3 Speeded up robust feature (SURF) detection and extraction algorithm: Figure 3 presents the activity diagram for the implementation of SURF algorithm while the following procedural steps of the algorithm's implementation are as follows:

Step 1. Creation of an integral image,
Step 2. Extraction of key points by:
a. Creating approximation of Hessian matrix.
b. Calculating responses of kernel used.

c. Finding local maxima across scale space.

Step 3. Determination of the SURF descriptor size to be used.

Step 4. Obtaining the dominant orientation.

Step 5. Extraction of the SURF descriptor.

Step 6. Matching the corresponding points based on the descriptor difference.

Step 7. Filtering out the outliers from matched points using RANSAC algorithm.

Table 1: Major characteristics of some of the widely-used feature detection and extraction algorithms

S/N	Feature Descriptor	Characteristics
1	Modified Harris Corner Detector (MHCD)	MHCD is rotationally invariant. It can perform optimally in the absence of scale difference.
2	Smallest Uni-value Segment Assimilating Nucleus (SUSAN)	It is a corner detector with a mask that calculates the intensity differences to detect or find the corners. It is scale variant (not invariant to scale).
3	Features from Accelerated Segment Test (FAST)	It uses Bresenham circle of radius 3 (circle of 16 pixels) to classify whether a candidate is actually a corner. It is invariant to scale and rotation with great improvement in the execution or processing time in the absence of noise.
4	Speeded Up Robust Features (SURF)	It is basically time economical when compared to other models but at the expense of accuracy and extracted corresponding features.
5	Scale Invariant Feature Transform (SIFT)	It is invariant to rotation, affine transformation changes and illumination. It performs optimally in feature extraction but with a slow execution time.
6	Principal Component Analysis (PCA)-SIFT	It reduced SIFT's execution time for matching (executes faster) but was proved to be less effective in feature detection compared to SIFT.

Table 2: Weighted analysis of the qualities of selected feature descriptors

S/N	Features Detector	Invariance			Characteristics/ Qualities			
		Scale	Rotation	Zoom	Robustness	Repeatability	Localization	Efficiency
1	SIFT	†	†	†	¶¶¶	¶¶	¶¶¶	¶¶
2	SURF	†	†	†	¶¶¶	¶¶¶	¶¶	¶¶
3	MHCD	X	†	†	¶¶¶	¶¶¶	¶¶¶	¶¶¶
4	MSER	†	†	†	¶	¶¶¶	¶¶¶	¶¶¶
5	FAST	X	†	X	X	X	¶¶¶	¶¶¶
6	FREAK	†	†	X	¶¶¶	¶¶¶	X	X
7	SUSAN	X	†	X	¶¶	¶¶¶	¶¶	¶¶¶
8	Hessian	†	†	X	¶¶	¶¶	¶¶	¶
9	DoG	†	†	X	¶	¶¶	¶¶	¶¶
10	Hessian-Laplace	†	†	†	¶¶¶	¶¶¶	¶¶¶	¶

Key: Where † means Yes, X means No or None, ¶ means good, ¶¶ means better and ¶¶¶ means best.

Based on the performance evaluation result as shown in Table 2, the SIFT, SURF and MHCD proved to exhibit more qualities in image registration. These algorithms are all known to be invariant to zoom, noise, scale, rotation and illumination (Krishna and Varghese, 2015). Hence, detailed algorithmic procedure of implementing these three (3) selected algorithms are provided in Figures 1 – 3. In order to achieve this, the mathematical description of the three (3) algorithms was first highlighted as presented in subsections 1.3.1, 1.3.2 and 1.3.3 for MHCD, SIFT and SURF respectively, and attempt was made to implement them following the procedures described in the process flow or activity diagrams (Figure 1-3) while their transformation homography was formulated using Random Sampling Consensus (RANSAC) algorithm because it only makes use of the required minimum number of input data set possible for the generation of candidate solutions, before proceeding to the enlargement of these data set with consistent data points in its estimation of model parameters (Ajayi, 2014; Fischler and Bolles, 1981) and also because of its ability to effectively cope with large percentage of outliers or mismatches in the input data set. It was also used for the exclusion of outliers from the matched points. The activity diagrams were composed within the Microsoft Enterprise Architecture software environment. The implementation phase was divided into input, processing and output stages for the three algorithms.

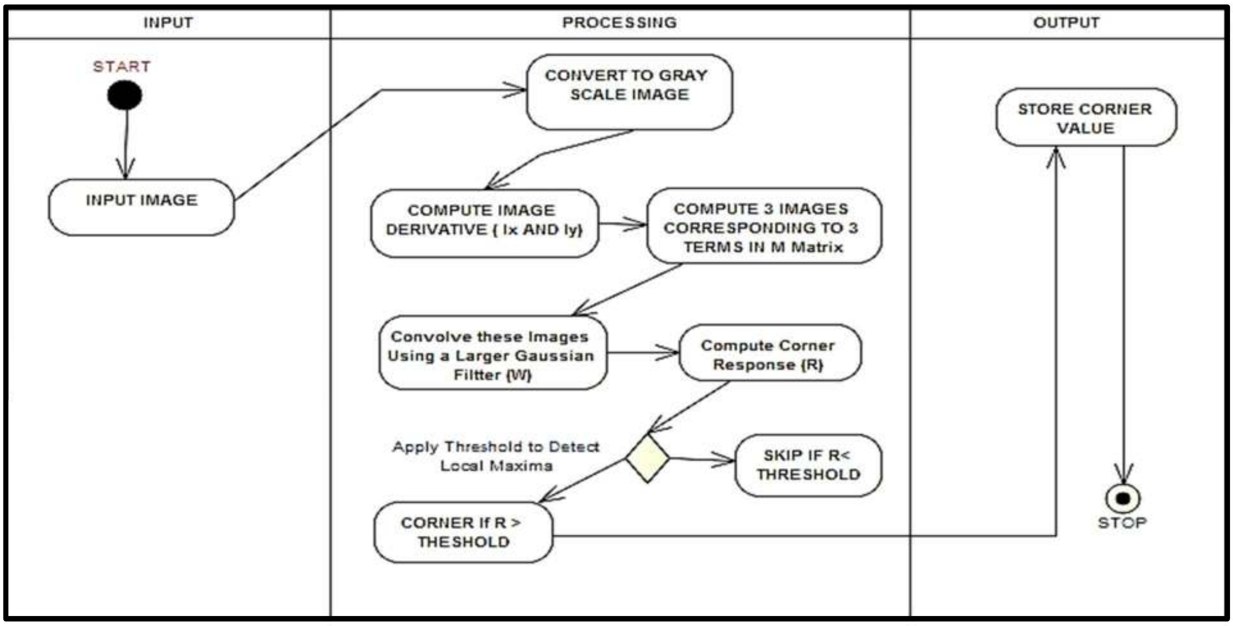

Figure 1: Activity diagram of Harris corner detection algorithm

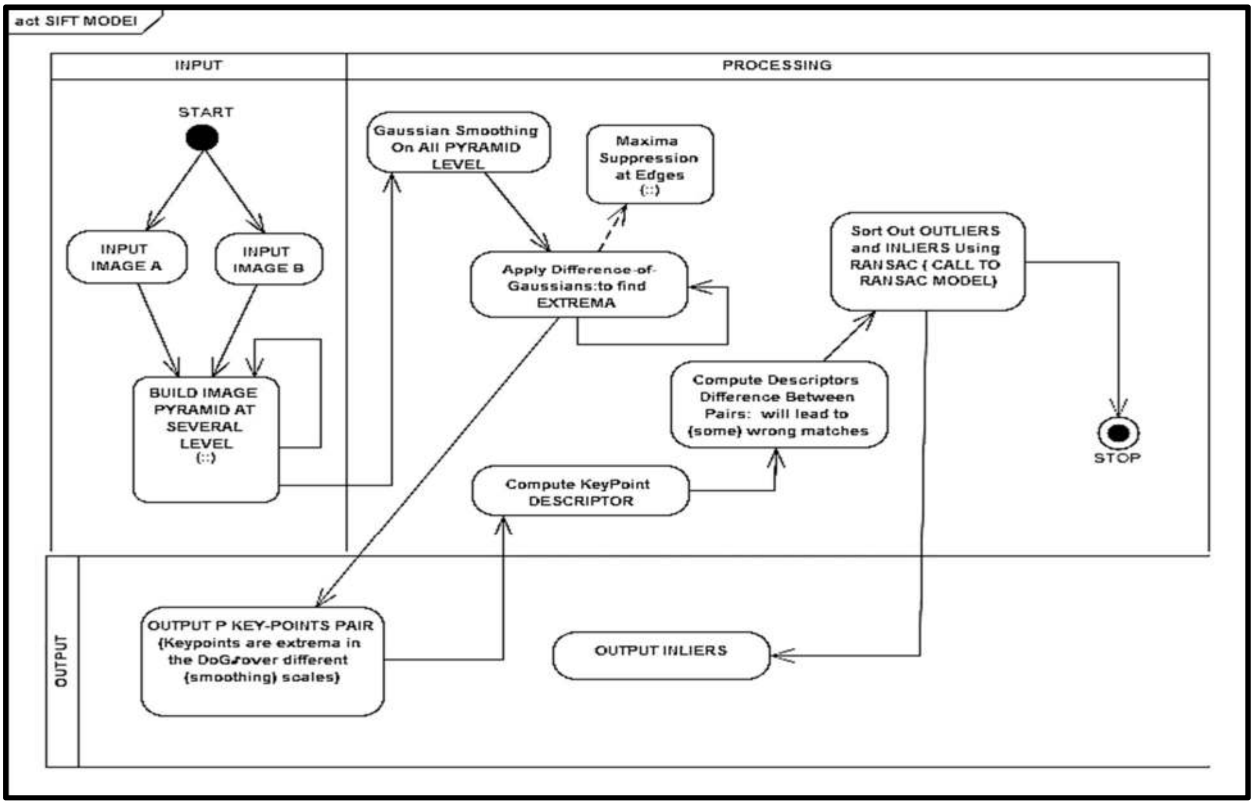

Figure 2: Activity diagram of Scale Invariant Feature Transform (SIFT) detector

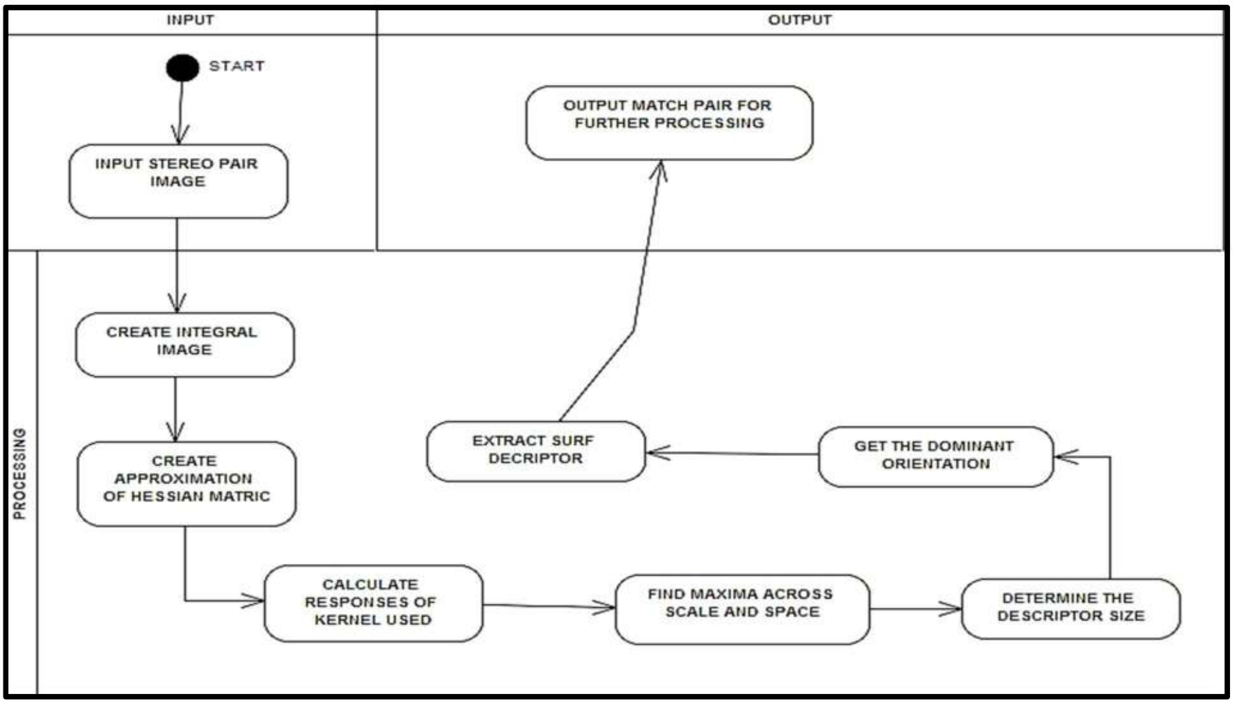

Figure 3: Activity diagram for Speeded Up Robust Feature (SURF) detector

2. EXPERIMENTAL ANALYSIS

Apart from the result of the weighted analysis of the ten (10) selected feature descriptors presented in Table 2, The three (3) feature descriptors discussed in subsection 1.3 were also implemented in an experimental design of an image registration scheme using UAV acquired overlapping image pairs of 80% overlap, presented in Figures 4a and 4b. While Figure 4a presents the base or reference image, Figure 4b presents the sensed or floating image. The size of each of the image pair is 3000 x 4000 pixels and it covers part of the Main campus of the Federal University of Technology, Minna, Nigeria. The result shows that the three feature descriptors proved to be indeed invariant to rotation as observed from the parameter vectors recorded in their estimated homography which shows a rotation angle that is equal to zero. The efficiency of the three feature descriptors was also tested with respect to their processing speed and the number of automatically extracted features or point correspondences. The result of this analysis is presented in Table 3, while the inliers of the automatically extracted conjugate points using the three descriptors are presented in Figures 5a, 5b and 5c for MHCD, SURF and SIFT respectively.

Table 3: Results of the selected three feature descriptors' speed and number of extracted point correspondences.

S/N	Feature Descriptor	Processing Run Time (Milli Seconds)	Number of Automatically extracted point correspondences
1	MHCD	6649	172
2	SIFT	10646	1067
3	SURF	13109	671

From the experimental result (Table 3), it was discovered that the SIFT algorithm proved to be more robust than the MHCD and the SURF algorithms in the automatic detection and extraction of point correspondences. It automatically extracted 1067point correspondences which is approximately 6.20 times more than the point correspondences automatically extracted by the MHCD algorithm (172) and 1.59 times more than the point correspondences automatically extracted by the SURF algorithm (671). This observation also agreed with the findings of Vivek and Kanchan (2014) and Panchal *et al*., (2013) which submitted that the SIFT model is very powerful in the automatic extraction of corresponding features. Also, though SIFT extracted the highest number of corresponding points, it proved to be very slow in processing or registering the images because it expended more processing run time when compared to the other implemented algorithms. The MHCD outperformed SIFT and SURF in terms of speed. It proved to be 1.60 times faster than SIFT and approximately 2 times faster than SURF. This is also consistent with the findings of Juan and Gwun (2009), and El-gayar *et al.* (2013).

3. CONCLUSIONS

The basic characteristics of the selected Ten (10) feature descriptors have been reviewed in this article. Also, an evaluation of the performance of these descriptors was also carried out under seven different conditions. The analysis shows that each of the descriptors have different qualities which makes them suitable for different image registration conditions. From the selected feature descriptors, MHCD, SIFT and SURF were further discussed in details with emphasis on their algorithmic implementation procedures while an experimental analysis was also conducted using these three algorithms on UAV acquired overlapping images. The result of the experiment shows that the three feature descriptors are indeed invariant to zoom, noise, scale, rotation and illumination. It also shows that while MHCD is very fast, it automatically extracts the least number of key points when compared to the three feature descriptors, while the SIFT automatically extracts the highest number of key points, though it expends more processing time. Finally, the choice of feature descriptor for an image registration task should be based on the peculiarities of the imaging conditions as no single feature descriptor can be acclaimed to be significantly better than others.

4. ACKNOWLEDGEMENT

This project was gratefully funded in part by the research grant awarded to the author by the Aubrey Barker Fund (ABF), UK and the Surveyors Council of Nigeria (SURCON).

Figure 4a: UAV acquired image of part of FUTMinna, Main Campus (reference image)

Figure 4b: UAV acquired image of part of FUTMinna, Main Campus (floating image)

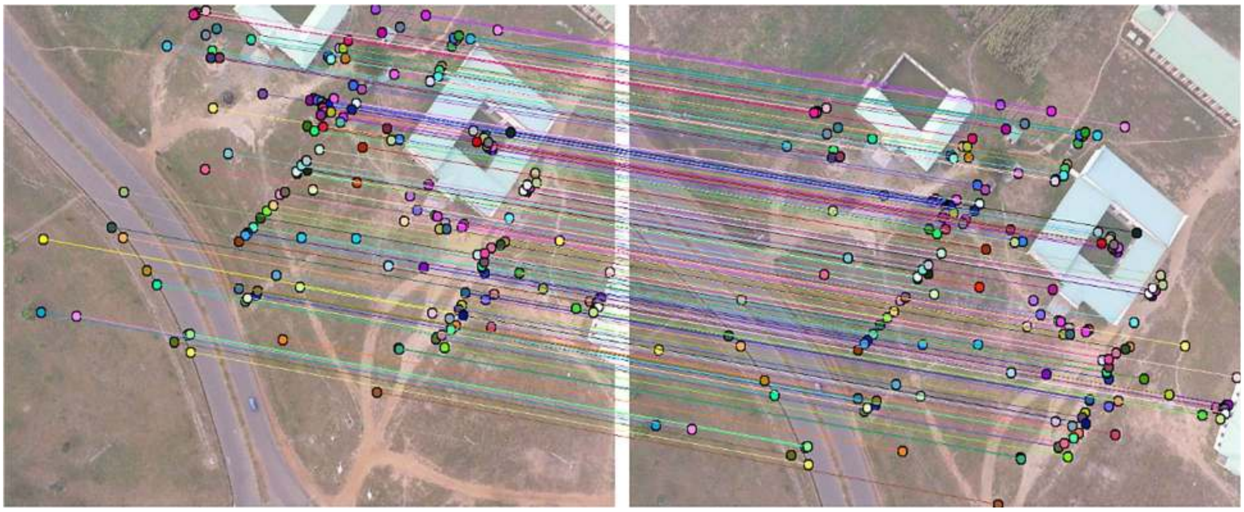

Figure 5a: Matched inliers using Modified Harries Corner Detector (MHCD)

Figure 5b: Matched inliers using Speeded Up Robust Features Algorithm (SURF)

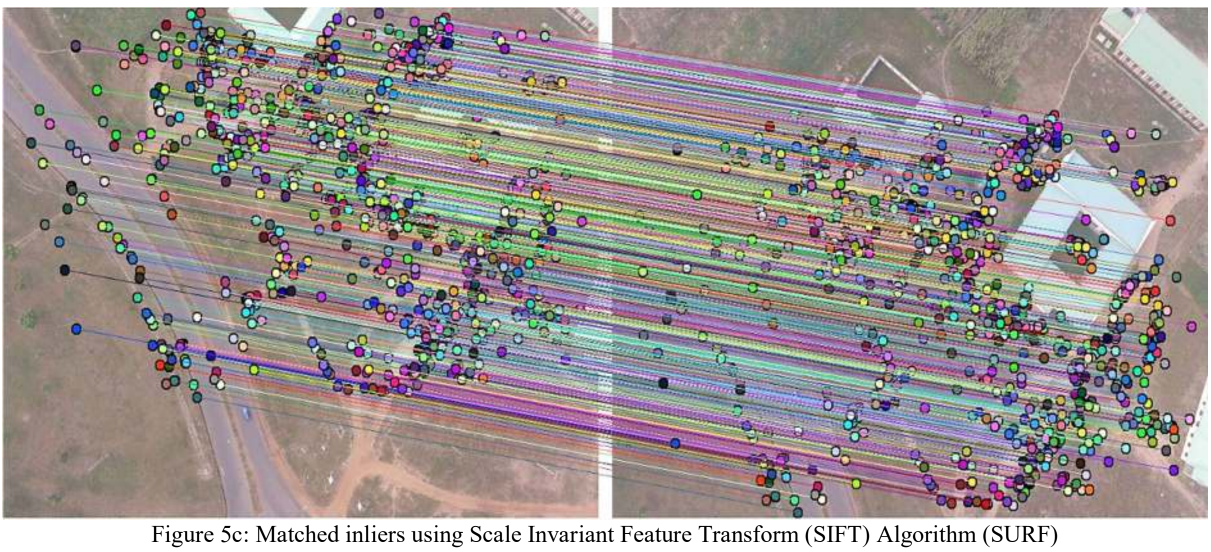

Figure 5c: Matched inliers using Scale Invariant Feature Transform (SIFT) Algorithm (SURF)

5. REFERENCES

Ajayi, O. G., 2019. Development of an integrated automatic image registration scheme. PhD Thesis submitted to the Postgraduate School, Federal University of Technology, Minna, Nigeria.

Ajayi, O. G., 2014. A MATLAB Program for the automatic registration of overlapping images. Unpublished MSc thesis submitted to the Department of Surveying and Geoinformatics, Faculty of Engineering, University of Lagos, Nigeria.

Bay, H., Ess, A., Tuytelaars, T., & Van, G. L. 2008., Speeded Up Robust Features (SURF). *Computer Vision and Image Understanding,* 110 (3), 346-359.

Bay, H., Tuytelaars, T., & Van Gool, L., 2006. SURF: Speeded Up Robust Features. In A. Leonardis, H. Bischof, and A. Pinz (Eds.), *Proceedings of the Ninth European Conference on Computer Vision* (pp. 404-417), Part I, LNCS 3951. Berlin, Germany: Springer-Verlag Berlin Heidelberg.

Brown, M., & Lowe, D. G., 2007. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision,* 74(1), 59–73.

Ehab, S., & Murad, Q., 2017. Recent advances in features extraction and description algorithms: a comprehensive survey. *18th International Conference on Industrial Technology (ICIT),* Toronto, Canada.

El-gayar, M. M., Soliman, H., & Meky, N., 2013. A comparative study of image low level feature extraction algorithms. *Egyptian Informatics Journal,* 14(2), 175-181.

Fischler, M. & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM,* 24(6), 381–395.

Hassaballah, M., Abdelmgeid, A. A., & Alshazly, H. A., 2016. Image features detection, description and matching. In A.I. Awad & M. Hassaballah (Eds.), *Image Feature Detectors and Descriptors, Studies in Computational Intelligence 630* (pp. 11-45). Switzerland: Springer International Publishing.

Jégou, H., Perronnin, F., Douze, M., Sánchez, J., Pérez, P., & Schmid, C., 2012. Aggregating local descriptors into a compact code. *IEEE Transaction on Pattern Analysis and Machine Intelligence,* 34(9), 1704–1716.

Juan, L., & Gwun, O., 2009. A comparison of SIFT, PCA-SIFT, and SURF. *International Journal of Image Processing,* 3(4): 143-152.

Kazhdan, M., Thomas, F., & Szymon, R., 2003. Rotation invariant spherical harmonic representation of 3D shape descriptors. *Eurographics Symposium on Geometry Processing*. Edited by Kobbelt, L., Schröder, P., and Hoppe, H.

Krishna, S., & Varghese, A., 2015. Feature based automatic multiview image registration. *International Journal of Computer Science and Software Engineering,* 4(11), 308-314.

Li, Y., Wang, S., Tian, Q., & Ding, X., 2015. A survey of recent advances in visual feature detection. *Neurocomputing,* 149, 736–751.

Lowe, D. G., 2004. Distinctive image features from scale invariant key points. *International Journal of Computer Vision,* 60(2), 91-110.

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

Mubarak, S., 1997. *Fundamentals of computer vision.* Orlando: Computer Science Department, University of Central Florida, 1-133.

Nixon, M. S., & Aguado, A. S., 2008. *Feature extraction and image processing.* Oxford, OX: Elsevier Press.

Tuytelaars, T. 2006., *Local Invariant Features: What? Why? When? How?* ECCV Tutorial Delivered on May 7th, 2006. Retrieved from Kuleuven Website: http://homes.esat.kuleuven.be/~tuytelaa/ECCV06tutorial.html

Tuytelaars, T., & Mikolajczyk, K., 2007. Local invariant feature detectors: a survey. *Foundation and Trends in Computer Graphics and Vision,* 3(3), 177–280, doi: 10.1561/0600000017.

Xiaohui, W., Kehe, W and Shengzhuang, W., 2013. Research on Panoramic Image registration Approach based on Spherical Model. International Journal of Signal Processing, Image processing and pattern Recognition. 6(6), 297-308.