Scientific Journal of Impact Factor(SJIF): 3.134

e-ISSN(O): 2348-4470
p-ISSN(P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 2,Issue 6, June -2015

Automatic Image Registration and Mosaicking using SURF and SIFT Feature Detection Algorithm

Bhavin B Patel[1], Mr. Mohammedsayeemuddin Shaikh [2], Prof. Sanjay D Joshi [3]

[1]*Electronics and Communication, LJIET, Ahmedabad*
[2]*Electronics and Communication, LJIET, Ahmedabad,*
[3]*Electronics and Communication, Vishwakarma Government Engineering College, Chandkheda,*

Abstract — *The area of image processing, search for the "points of interest" is a biggest issue. For the same many methods and algorithms have been proposed. It is very essential step in evaluation process. Image registration and mosaicking is an important operation in remote sensing applications that basically involves the identification of many control points in the images. Imaging mosaicking is being done on such that images taken by normal camera can be used to create a larger field of view using an image mosaicking. There are many existing methods for matching interest points and most of them are related to the parameters of the detectors. In this paper, we have present a multiple steps in which various image processing techniques are used. Speed-up Robust Features (SURF) and Scale Invariant Feature Transform (SIFT) can be mounted alternatively to detect and define features on images. This paper also address the issues of registering multi-spectral and multi-temporal images. The effectiveness of the proposed techniques has been studied by registering partially overlapping mosaic images.*

Keywords- *Image registration, Image mosaicking, Speed-up Robust Features (SURF), Scale- Invariant Feature Transform (SIFT), Multi-spectral and Multi-temporal Images*

I. INTRODUCTION

Vision is the most advanced of our sensors, so it is not surprising that images play the single most important role in human perception [1]. In parallel with space application, digital image processing techniques began in late 1960s and early 1970s to be used in medical imaging, remote earth observations and astronomy [1]. It is very important part of computer vision and digital image processing the feature detection, feature extraction and matching technology, and has been widely use in the object detection, 3D reconstruction and image registration. The detail information of individual image senses in temporal and spatial domain can be combined to produce unsegment panorama using images of smaller field of view.

Image registration is the process of overlaying two or more images of the same scene taken at different time intervals, from different angles, and by different cameras. It does geometrically alignment of two images the reference image and the sensed image [2]. Image registration is a crucial step in all image analysis tasks in which the final information is gained from the combination of various data sources.

Imaging mosaicking is the project such that images taken by any type of camera can be used to create a larger area of view using an image mosaicking program [3]. The program is basically a tools which help you to find out the corresponding camera angles that you used to take the image.

Feature point extraction and characterization are related to repeatability creation that evaluates the noise of feature point detectors, stability and robustness under the image transformations [6]. If matches are not properly found, a misleading transformation function is produced and most probably yield a completely wrong result. The problem of image matching consists of identification for two or more images at same scene.

The work presented in this paper focus on feature matching based on points of feature based on images of the same scene with same resolutions. The main goal of this work is to detect a feature points and compare many feature point detectors like SIFT, SURF in terms of repeatability. We accomplished a practical comparison of feature detectors. A numbers of experiments were performed to evaluate feature point detectors.

II. INTEREST POINTS DETECTORS

1) SIFT

The Scale Invariant Feature Transform (SIFT) algorithm was proposed by Lowe in the year 1999. SIFT [4] is a feature detection algorithm which detects feature in an image that identifies similar objects in other images. It produces key point descriptors which are the image features. For a set of input images SIFT extracts features. SIFT algorithm is both rotational invariant and scale invariant. SIFT is very popular for object detection in images with high resolution.
SIFT has computational phases which includes:

- **Scale-space extreme detection:** This is the first phase which identifies the potential interest points. It searches over all scales and image locations by using a difference-of-Gaussian (DoG) function. Here, the middle point is compared with its neighborhood points to detect utmost points.

- **Key-point Localization:** For all the interest points so found in phase one, location and scale is determined. Key-points are selected based on their stability. A stable key point should be resistant to image distortion. This is done by using Taylor series Expansion, the extreme points and location are carefully determined by using the following equation:

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \tag{1}$$

- **Orientation Assignment:** SIFT algorithm computes the direction of gradients around the stable key-points. One or more orientation are assigned to each key-point based on local image gradient directions.
  Compute Gradient for each blurred image

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

$$\theta(x,y) = \tan^{-1} \frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1)\,y} \tag{3}$$

The seed point can be formed by aligning the direction along with the amplitude of pixels. The seed point can be formed by aligning the unidirectional gradients followed by the normalization. Wavelet response in the vertical direction. Each sub-regions responses are summed-up along with the absolute value of response and each sub-region vector will form the four-dimensionality: For each sub feature point, format the 4×(4×4)=64 dimensional description vector, then normalize the vector.

2) SURF

Speeded-Up Robust Features (SURF) uses a different and far more intensive filter approximation when compared with DoG approach in SIFT. Created by Bay et al compute the Key-Points in an image, in this method the Box Filter or Mean Filter responses are used [5]. Here Box Filters are computed using Integral Images. For a given image, the number of features extracted using DoG approach in SIFT and simple box feature approach in SURF are comparable.

SURF feature extraction algorithm also has 4 stages – Scale Space Analysis, Key Point Localization and Orientation Assignment and Key Point Descriptor generation. These stages are explained in below.

STEP: 1 Scale Space Analysis

In this step, the Scale Space is built and analyzed for possible extrema locations across all scales. Here, every layer is computed by applying a Box Filter of different sizes on input image. Filter sizes increase both within and across the Octaves. Box Filters referred here as approximation of second order Gaussian Derivatives in x, y, xy directions represented as D_{xx}, D_{yy}, and D_{xy}. The starting filter size of 9x9 which is applied at lowest scale of image and with every next interval, filter size increases until last interval in fourth Octave. The convolution of these box filters with input image can be efficiently computed using Integral Image.

The responses D_{xx}, D_{yy}, and D_{xy} computed using Integral Image and Box Filter are then used to compute, what are known as approximated Normalized determinant of Hessians, proposed by Bay et al and given by equation,

$$\det(H\ approx) = D_{xx} \times D_{yy} - (0.9 \times D_{xy})^2 \tag{4}$$

The determinant of Hessians or interest points, are computed for all positions of input image, for all scales and filter sizes.

STEP 2: Key – Point Localization

Key Point Localization is a two-step process, in first step only those interest points are chosen which are scale and rotation invariant or in other words strong interest points and in second step, chosen points are localized across scales.

Excluding or filter out weak interest points is carried out at two levels. In the first, all interest points are passed through a threshold test. In this test, the interest points which are above threshold value are considered and the rest are discarded. The second level is known as Non – Maximum Suppression. It is carried out across three layers with different scales. The last step in Key-Point Localization is to interpolate nearby data to determine the position and scale of interest point to a sub-pixel accuracy.

STEP 3: Orientation Assignment

In this step, every interest point which has passed the previous tests are assigned with reproducible orientation information to achieve invariance to image rotation. The value orientation is very important in computing the final interest point description, which is expressed in sixty four floating point values. Assigning an orientation detail to every interest point is carried out in two steps. In the first step, a circular region of radius 6 x scale around each interest point is considered and within this region the Haar wavelet responses of size 4 x scale in x and in y directions are computed. Responses so obtained are weighted with a Gaussian centered around an interest point and plotted as vector points along x and y coordinates. In step two, a window of size $\frac{\pi}{3}$ is rotated around an interest point and the points which are covered within the window are summed-up. The most dominant result of such summing actions is considered as dominant orientation of the interest point, which is used in calculating the description of the interest point.

Step 4: Key-Point Descriptor Generation

The last step in Key-Point generation is to give a description to all localized Key Points with orientation information. The Key-Point description is expressed in 64 values. In this step, a square region which is divided into sixteen sub-squares is considered around the center of every interest point. This square is aligned along the orientation computed in previous step. Every sub-square is sampled at twenty five (5 x 5) regularly spaced points. Like in the previous step, Haar-x and Haar-y wavelet responses are com- puted at every 25 points within a sub-square. These responses are then applied with Gaussian weights. From every sub-square region, four vectors - two in x (dx, |dx|) and two in y (dy, |dy|) co-ordinates are computed. The summation of all the four values from all the 25 samples gives rise to four vectors $v = (\sum dx ; \sum |dx| \sum dy \sum |dy|)$ from one sub-square. So, from the entire square region, which has 16 sub-squares there are 16 x 4 = 64 values which forms the description of an interest point.

III. PERFORMANCE EVALUATION

There are various quality metrics on the basis of which the image quality can be evaluated. Subjective quality metric depends on the human observer whereas the objective metric depends on the computation. Various quality metrics that have been determined to evaluate the image quality are PSNR, MSE (mean square error), and MAXERR (maximum absolute error).

MSE: - The MSE is the cumulative squared error between the compressed and the original image. The mathematical formulae for the two are

$$MSE = \frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N}[I(x,y) - I'(x,y)]^2 \tag{5}$$

Where I(x, y) is the Image 1, I'(x, y) is the Image 2 and M, N are the dimensions of the images.

PSNR: - It is defined as the peak signal to noise ratio in decibels. PSNR is used to measure the quality of reconstruction. It is calculated between references to processed image. If R is the measure of the input image data type, then the PSNR is given by:-

$$PSNR = 10log_{10}(\frac{R^2}{MSE}) \tag{6}$$

Thus the PSNR value tends to infinity as two MSE tends to zero, which means that the larger PSNR value corresponds to higher image quality.

IV. EXPERIMENTAL SETUP

To test and verify the performance of SIFT and SURF algorithm, Performance is mainly compared to the matching rate and matching time of feature point extraction with SIFT and SURF algorithm.

Experimental environment: Operating system: Windows 8 Pro, Processor: Intel(R) Core(TM) i3-2330U CPU @ 2.20 GHz, 4.00 GB RAM, Development environment: MATLAB R2013a – 64-bit (win64).

Different types of images were tested in order to demonstrate the feasibility of the system. Some preliminary results are presented in this section. The images have captured by the camera DSC-W830. The technical specification of above camera is as given:- Maximum aperture- 5.3125, Focal length – 33mm, Exposure time - 0.02 sec. The images have been captured at same location, same time and different Focal Length. The size of images are 640x 480 pixels.

Figure 1: Input Images

V. EXPERIMENTAL RESULTS

Figure 2: Output SURF method

Figure 3: Output SURF method

TABLE 1:different types of analiysis using SIFT method

Match Image	Image-1 Key Point	Image-2 Key Point	Outlier Point	Inlier Point	Success (%)	PSNR (dB)	MSE
1	2578	2179	483	424	87.78468	15.79505	1712.282
2	2179	2034	386	341	88.34197	16.76409	1369.846
3	2034	2259	784	680	86.73469	18.15869	993.6004
4	2259	2513	713	637	89.34081	17.09691	1268.79
5	2513	2818	924	795	86.03896	16.94726	1313.27
6	2818	2610	1314	1102	83.86606	16.5779	1429.85
7	2610	2347	954	822	86.16352	17.32758	1203.158

TABLE 2:different types of analiysis using SURF method

Match Image	Image-1 Key Point	Image-2 Key Point	Outlier Point	Inlier Point	Success (%)	PSNR (dB)	MSE
1	652	613	184	160	86.95652	15.79505	1712.282
2	613	435	64	60	93.75	16.76409	1369.846
3	435	535	194	190	97.93814	18.15869	993.6004
4	535	601	216	209	96.75926	17.09691	1268.79
5	601	678	215	208	96.74419	16.94726	1313.27
6	678	629	316	292	92.40506	16.5779	1429.85
7	629	522	210	200	95.2381	17.32758	1203.158

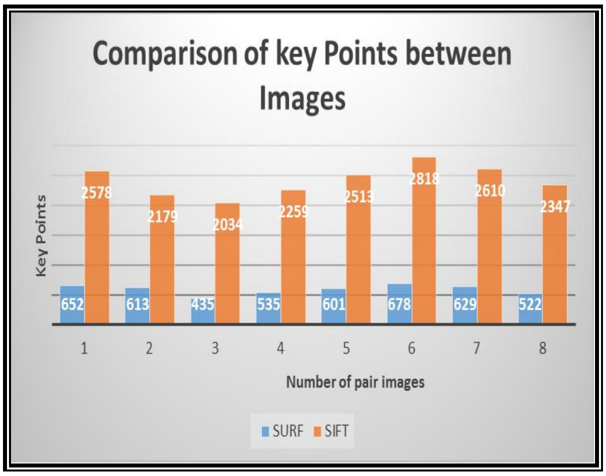

Figure 4: comparision of key Points

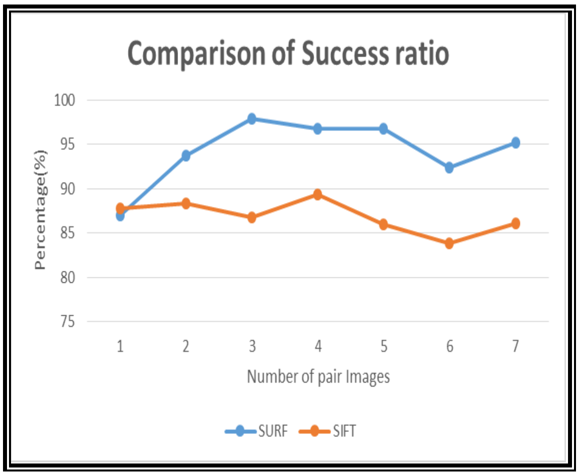

Figure 5:comparision of Success ratio

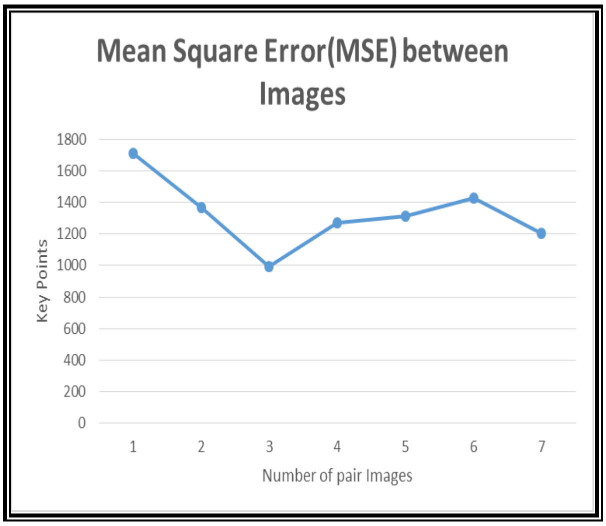

Figure 6: MSE between Images

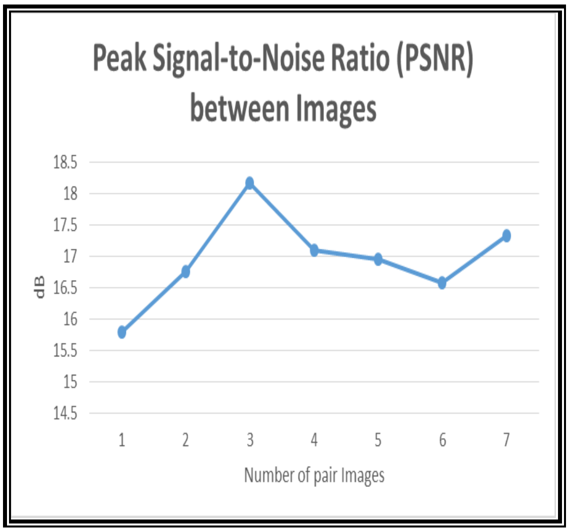

Figure 7: PSNR between Images

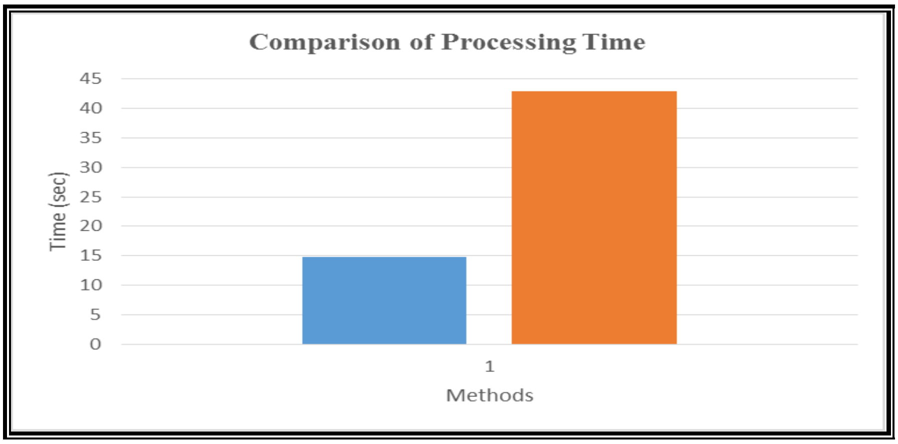

Figure 8: Comparision of Processing Time

VI. CONCLUSION

By observing the result it can be concluded that if the feature points are not selected properly then the output image will be degraded. Due to the difference in viewpoint of the sensors, there is a great difficulty in the image mosaicking. Still problems are faced for more than two images and projective transformation.

In Test Case the images are taken from cellphone which is of 8MP resolution. The processing time using SURF method is 14.81255 Second and Processing time using SIFT method is 42.8437 Second. The success ratio lies 75 to 90 (%) in SURF and it lies 85 to 90(%) in SIFT. Thus, in SIFT the success ratio is good and remains stable.

Thus, the parameters of image registration and mosaicking depend on the types of image and the viewpoint of sensor of image acquisition device.

Future Scope

- Error occurred in mosaicking can be minimized.
- As it is to be made automatic, the algorithm should be good enough to solve problem occurred during mosaicking.
- The system can be extended to create a large field of view using normal camera, the result image can also be used for texture mapping of a 3D environment such that users can view the surrounding scene with real images.

REFERENCES

[1] Gonzales, Rafael C., and Richard E. Woods. "Digital Image Processing, 2-nd Edition." (2002).
[2] L. G. Brown, “A survey of image registration techniques,” ACM Computing Surveys, vol. 24, ISSN: 093-18470, (1992), pp. 325–376.
[3] Fedorov, Dmitry, Leila MG Fonseca, Charles Kenney, and B. S. Manjunath, “System for automatic registration of remote sensing images”, In Geoscience and Remote Sensing Symposium, IEEE International, vol. 4, Print ISBN: 0-7803-7536-X, (2002), pp. 2459-2461.
[4] D.G. Lowe, “Object recognition from local scale-invariant features”, The International Conference on “Computer Vision”, Print ISBN: 0-7695-0164-8, (1999), pp. 1150 -1157.
[5] Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool, “Surf: Speeded up robust features”, In Computer vision–ECCV, Springer Berlin Heidelberg, ISSN: 2162-9900, (2006), pp. 404-417.
[6] Sunanda Gupta, S.K. Chakarvarti, and Zaheerudin, “Image Registration Methods: A Short Review”, American Journal of Algorithms and Computing, ISSN: 2162-9900, (2013), pp. 39-49.
[7] Bentoutou, Youcef, Nasreddine Taleb, A. Bounoua, Kidiyo Kpalma, and Joseph Ronsin, “Feature based registration of satellite images”, 15th International Conference on “Digital Signal Processing”, E-ISBN: 1-4244-0882-2, (2007), pp. 419-422.

