JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Implementing SIFT for Remote Sensing Image Registrations

[1]Aparna Joshi, [2]Asst. Prof. Lalita Randive

[1,2] Department of Computer Science & Engineering, MIT College, Aurangabad

Abstract : This paper presents Image matching and registration method that is invariant to scale, rotation, translation and illumination changes. The method is named as Scale Invariant Feature Transform (SIFT). This algorithm will detect and describe image features such as contours, points, corners etc. SIFT descriptors are the characteristic signature of the feature. The features calculated from the image to be registered should be distinctive and then it can be matched. It can be useful in object recognition, image mosaicing, 3 D reconstruction and video tracking. The simulation results shows that this algorithm works well in all types of cases having scale and rotation difference, it also register the object having occlusion and clutter background. Due to good invariance of scale, rotation, illumination, SIFT (Scale Invariant Feature Transform) descriptor is commonly used in image matching. The steps of extracting SIFT feature are analyzed in detail, and SIFT Key-point location is optimized.

***IndexTerms* – SIFT, Key Descriptors, Image Processing**

I. Introduction

SIFT (Scale Invariant Feature Transform) calculation is fruitful in include coordinating with research regions. The elements extricated by SIFT are invariant to picture scaling and turn, and somewhat invariant to change in brightening and 3D camera perspective. Pictures with the bigger contrasts can be coordinating by the consistent highlights. Yet, the time intricacy of the calculation is somewhat high. Simultaneously, constant handling and coordinating with qualities precision of the calculation are relied upon to be improved. In picture handling field, this issue is as yet not totally tackled. A few researchers have investigated a few techniques to enhance this calculation. A worked on calculation depicted by Li Liu [1] utilized a component point with just 12 aspects dependent on a round window to work on the effectiveness of coordinating, however power against clamor is more fragile than unique calculation. Shen Qian [2] further developed the calculation speed by powerfully adjusting examining step when registering the inclination histogram of the locale around the central issue area. Zhu Daixian [3] proposes a strategy, which utilized straight blend of city-block distance and chessboard distance instead of Euclidean distance in coordinating with process. The age interaction of SIFT is examined right off the bat, then, at that point, The Slant distance is utilized to gauge similarity of element descriptors. Coordinating with calculation time is diminished and the precision of picture coordinating is likewise improved.

Picture enlistment is the most common way of adjusting at least two pictures of a similar scene caught from various points/view focuses, at various occasions and by various camera. It is significant stage in all picture examination undertakings in which last data is acquired from the enlisted picture. It comprise change which adjust a picture with the goal that it coordinates with another. Many sorts of changes might be viewed as like Fourier change, cosine change, Hough change, wavelet change and relative change and some more. In this paper, we are keen on mathematical relative changes like scaling, turn, interpretation. Picture enrollment mathematically adjusts two pictures – reference picture and detected picture. One of the pictures is kept unaltered called as reference picture and other can be disfigured called as detected picture. Enrolling at least two pictures is helpful in different fields like PC vision – shape recuperation, programmed change identification, programmed target acknowledgment, clinical imaging – checking mending treatment, observing of the cancer, finding, remote detecting – movement following, mosaicing, combination and metrology – study of estimation. Most of the picture enlistment techniques comprises of element identification, highlight coordinating, change model assessment lastly picture re-examining and change. The elements are particular components of the pictures that are first removed freely for two pictures to enlist and afterward correspondence between them is set up. Change model assessment will gauge coordinated with elements and boundaries of planning capacities are processed and afterward addition of detected picture. This region is exceptionally difficult in view of boisterous climate, brightening changes, constraint of sensor limit and some more. Heartiness is vital property for achievement in above conditions.

II. Related Work

Many component based enlistment strategies are accessible in writing. Creators of various papers attempt to find and match picture elements like focuses, lines, corners and articles between reference picture and detected picture. Reference [4] shows SIFT calculation for picture enrollment. It recognizes highlights that are invariant to pivot, scaling, interpretation and light.

M – Estimator based picture enrollment calculation introduced by K. V. Arya [5] endeavor to coordinate with format picture with the objective picture. This strategy heartily enrolls pictures within the sight of clamor and impediment (falsely created) up to 60%. It has been displayed in that the proposed calculation performs effectively though other connection based techniques produce the inaccurate enlistment. This permits one to get assessment systems more strong to wrong information. For example, in the component based methodology, they are frequently bogus element correspondence. This for the most part comes from a helpless extraction of the highlights or from a terrible coordinating of the elements. M – assessor strategy isn't reasonable when two info pictures are having high distinction in pivot and scale.

Meisen Pan [6] proposed fluffy based picture enlistment in which Fuzzy C-Means (FCM) grouping is utilized and separated the directions of the pixels in the picture into two bunches to fit a straight line, and afterward determined the incline of the line to register the turn point of the picture. Fluffy strategy gives great outcome to mono modular and multi modular (with various sensors) pictures.

In paper [7], common data based strategies have been proposed. The pictures are caught at various points with similar sensor and pictures are skewed by little change, so this technique gives quick and serious outcomes. However, constraint of this enlistment is powerlessness to enroll the unique pictures having distinctive data.

III. PROPOSED SYSTEM

Image matching and registration requires coordinating between detected picture and reference picture, scale contrast, change of point and interpretation between two pictures. Scale Invariant Feature Transform (SIFT) technique will observe separate interest focuses in both the pictures, dispose of focuses which gives equivocal data and afterward coordinating with solid interest focuses utilizing SIFT descriptors lastly adjust both the pictures. Descriptor are called trademark mark of the component.

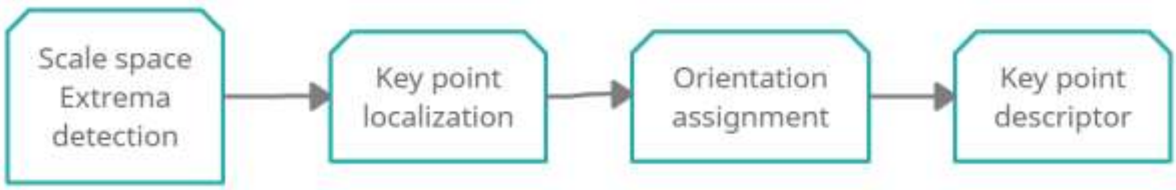


Figure 1.0 Basic Flow Diagram

In above block outline displayed in fig. 1, ventures for removing central issues are referenced. In sync 1, calculation tracks down likely areas of central issues for tracking down top/extrama highlights. Stage 2 finds precisely finding the element central issues by disposing of ill defined data central issues. Stage 3 allocates direction points to the interest focuses. Lastly portrays the central issue as a high dimensional vector of size 128.

Constructing the Scale Space

We need to identify the most distinct features in a given image while ignoring any noise. Additionally, we need to ensure that the features are not scale-dependent. These are critical concepts so let's talk about them one-by-one. We use the Gaussian Blurring technique to reduce the noise in an image. So, for every pixel in an image, the Gaussian Blur calculates a value based on its neighboring pixels. Below is an example of image before and after applying the Gaussian Blur. As you can see, the texture and minor details are removed from the image and only the relevant information like the shape and edges remain:

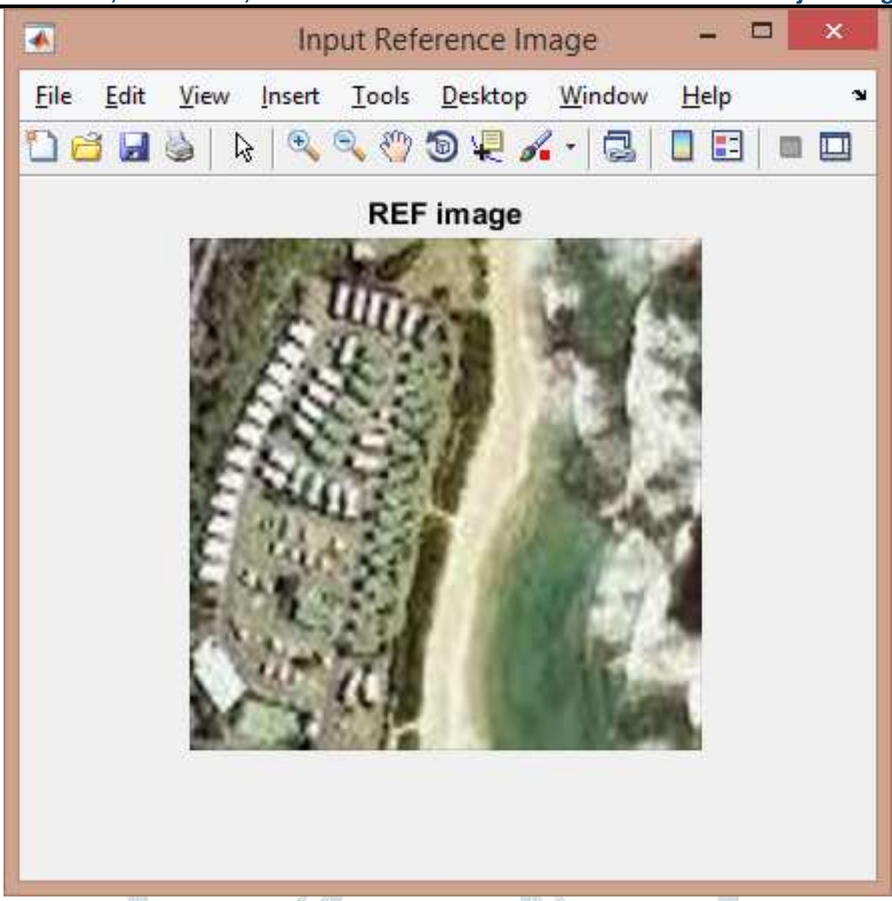


Figure 1.0 Gaussian Blur successfully removed the noise from the image.

Scale space is an assortment of pictures having various scales, created from a solitary picture. Consequently, these haze pictures are made for a very long time. To make another arrangement of pictures of various scales, we will take the first picture and decrease the scale considerably. For each new picture, we will make obscure variants as we saw previously.

We have the first picture of size (275, 183) and a scaled picture of aspect (138, 92). For both the pictures, two haze pictures are made:

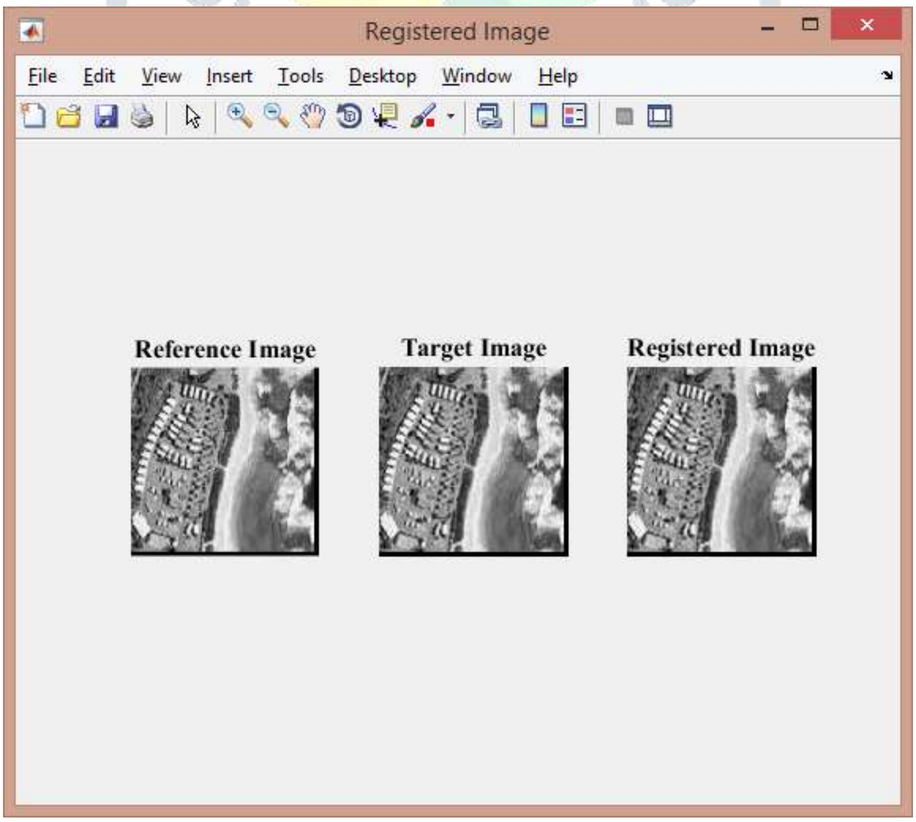


Figure Octave Images I

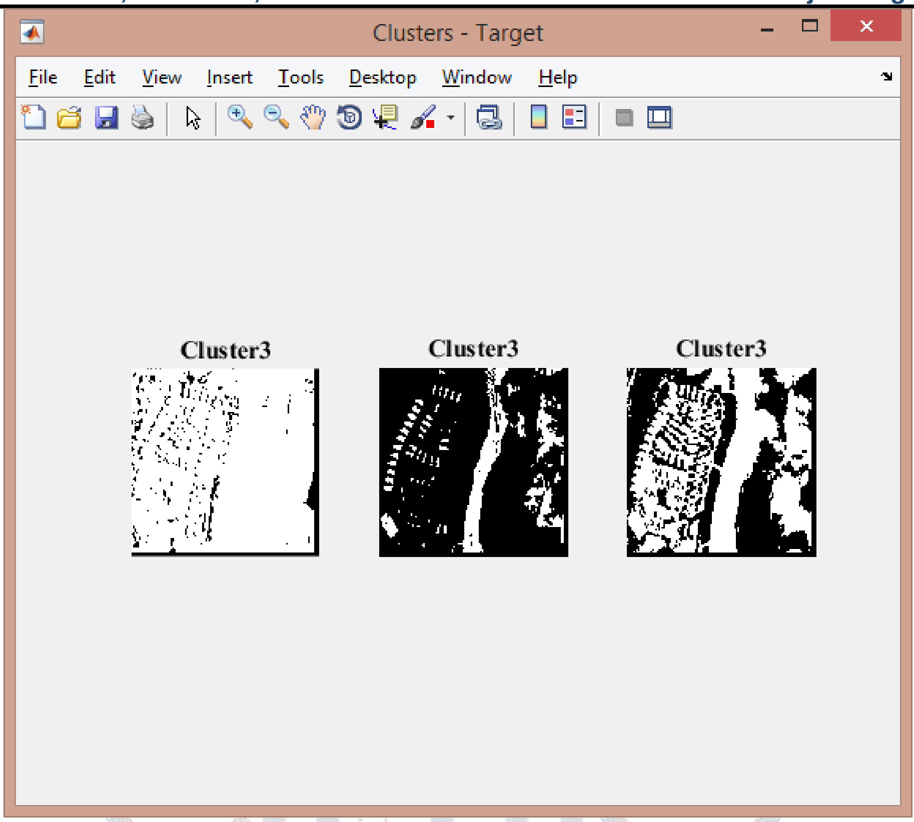


Figure Octave Images II

SIFT, or Scale Invariant Feature Transform, is a feature detection algorithm in Computer Vision. SIFT helps locate the local features in an image, commonly known as the '*keypoints*' of the image. These keypoints are scale & rotation invariant that can be used for various computer vision applications, like image matching, object detection, scene detection, etc.

- **Constructing a Scale Space:** To make sure that features are scale-independent
- **Keypoint Localisation:** Identifying the suitable features or keypoints
- **Orientation Assignment:** Ensure the keypoints are rotation invariant
- **Keypoint Descriptor:** Assign a unique fingerprint to each keypoint

Constructing the Scale Space

We need to identify the most distinct features in a given image while ignoring any noise. Additionally, we need to ensure that the features are not scale-dependent. These are critical concepts so let's talk about them one-by-one.

We use the Gaussian Blurring technique to reduce the noise in an image.

So, for every pixel in an image, the Gaussian Blur calculates a value based on its neighboring pixels. Below is an example of image before and after applying the Gaussian Blur. As you can see, the texture and minor details are removed from the image and only the relevant information like the shape and edges remain:

Gaussian Blur successfully removed the noise from the images and we have highlighted the important features of the image. Now, we need to ensure that these features must not be scale-dependent. This means we will be searching for these features on multiple scales, by creating a 'scale space'.

Scale space is a collection of images having different scales, generated from a single image.

Hence, these blur images are created for multiple scales. To create a new set of images of different scales, we will take the original image and reduce the scale by half. For each new image, we will create blur versions as we saw above.

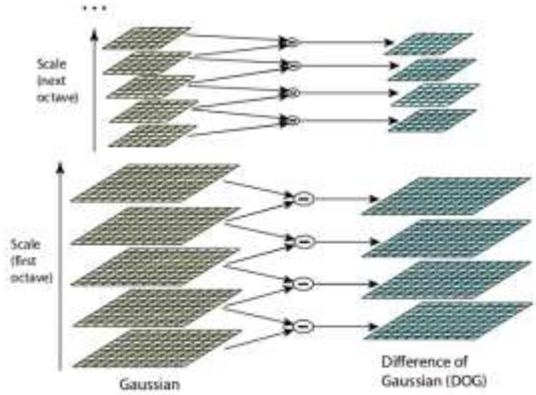
Here is an example to understand it in a better manner. We have the original image of size (275, 183) and a scaled image of dimension (138, 92). For both the images, two blur images are created:

Difference of Gaussian

So far we have created images of multiple scales (often represented by σ) and used Gaussian blur for each of them to reduce the noise in the image. Next, we will try to enhance the features using a technique called Difference of Gaussians or DoG.

Difference of Gaussian is a feature enhancement algorithm that involves the subtraction of one blurred version of an original image from another, less blurred version of the original.

DoG creates another set of images, for each octave, by subtracting every image from the previous image in the same scale. Here is a visual explanation of how DoG is implemented:



The octaves are now represented in a vertical form for a clearer view.

Keypoint Localization

Once the images have been created, the next step is to find the important keypoints from the image that can be used for feature matching. The idea is to find the local maxima and minima for the images. This part is divided into two steps:

Find the local maxima and minima

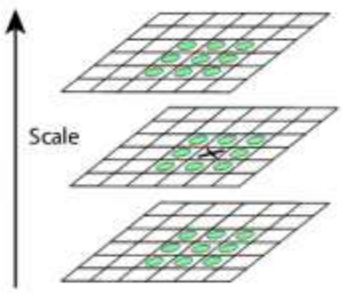
Remove low contrast keypoints (keypoint selection)

Local Maxima and Local Minima

To locate the local maxima and minima, we go through every pixel in the image and compare it with its neighboring pixels.

When I say 'neighboring', this not only includes the surrounding pixels of that image (in which the pixel lies), but also the nine pixels for the previous and next image in the octave.

This means that every pixel value is compared with 26 other pixel values to find whether it is the local maxima/minima. For example, in the below diagram, we have three images from the first octave. The pixel marked *x* is compared with the neighboring pixels (in green) and is selected as a keypoint if it is the highest or lowest among the neighbors:



We now have potential keypoints that represent the images and are scale-invariant. We will apply the last check over the selected keypoints to ensure that these are the most accurate keypoints to represent the image.

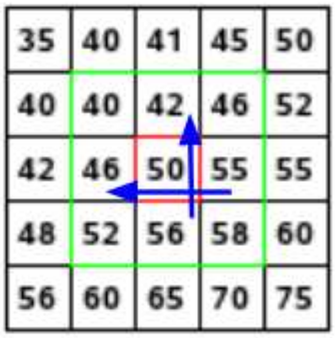
Keypoint Selection

Hence, we will eliminate the keypoints that have low contrast, or lie very close to the edge. To deal with the low contrast keypoints, a second-order Taylor expansion is computed for each keypoint. If the resulting value is less than 0.03 (in magnitude), we reject the keypoint. So what do we do about the remaining keypoints? Well, we perform a check to identify the poorly located keypoints. These are the keypoints that are close to the edge and have a high edge response but may not be robust to a small amount of noise. A second-order Hessian matrix is used to identify such keypoints. You can go through the math behind this here.

Orientation Assignment

At this stage, we have a set of stable keypoints for the images. We will now assign an orientation to each of these keypoints so that they are invariant to rotation. We can again divide this step into two smaller steps:

- Calculate the magnitude and orientation
- Create a histogram for magnitude and orientation
- Calculate Magnitude and Orientation
- Consider the sample image shown below:



To calculate the gradients in x and y directions by taking the difference between 55 & 46 and 56 & 42. This comes out to be Gx = 9 and Gy = 14 respectively.

Once we have the gradients, we can find the magnitude and orientation using the following formulas:

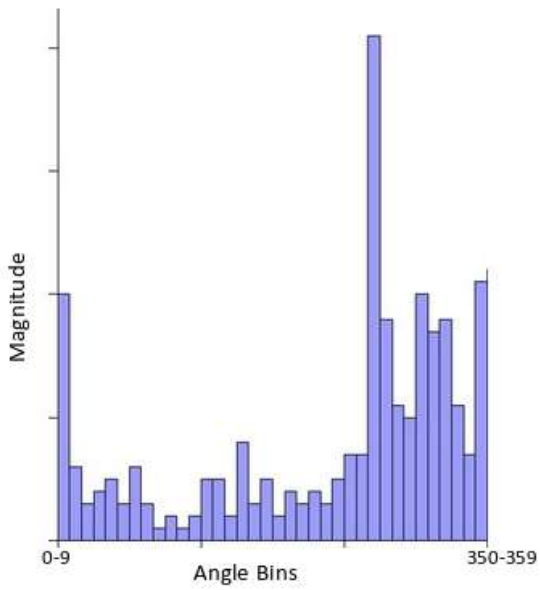
$$\text{Magnitude} = \sqrt{[(Gx)2+(Gy)2]} = 16.64$$

$$\Phi = \text{atan}(Gy / Gx) = \text{atan}(1.55) = 57.17$$

The magnitude represents the intensity of the pixel and the orientation gives the direction for the same.
We can now create a histogram given that we have these magnitude and orientation values for the pixels.

Creating a Histogram for Magnitude and Orientation.

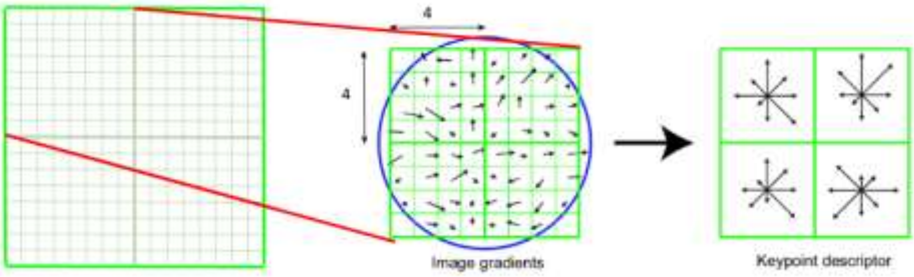
On the x-axis, we will have bins for angle values, like 0-9, 10 – 19, 20-29, up to 360. Since our angle value is 57, it will fall in the 6th bin. The 6th bin value will be in proportion to the magnitude of the pixel, i.e. 16.64. We will do this for all the pixels around the key point.



This histogram would peak at some point. The bin at which we see the peak will be the orientation for the keypoint. Additionally, if there is another significant peak (seen between 80 – 100%), then another keypoint is generated with the magnitude and scale the same as the keypoint used to generate the histogram. And the angle or orientation will be equal to the new bin that has the peak.

Keypoint Descriptor

This is the final step for SIFT. So far, we have stable keypoints that are scale-invariant and rotation invariant. In this section, we will use the neighboring pixels, their orientations, and magnitude, to generate a unique fingerprint for this keypoint called a 'descriptor'.
Additionally, since we use the surrounding pixels, the descriptors will be partially invariant to illumination or brightness of the images.
We will first take a 16×16 neighborhood around the keypoint. This 16×16 block is further divided into 4×4 sub-blocks and for each of these sub-blocks, we generate the histogram using magnitude and orientation.



At this stage, the bin size is increased and we take only 8 bins (not 36). Each of these arrows represents the 8 bins and the length of the arrows define the magnitude. So, we will have a total of 128 bin values for every key point.

IV. EXPERIMENT SETUP

The above algorithm is achieved with program language MATLAB in this paper. An example of two images matching, two images matching time used Lowe's method is 2950ms, but the time used in the proposed method in this article is 1020 ms.

The SIFT key points described in this paper are particularly useful due to their distinctiveness, which enables the correct match for a key point to be selected from a large database of other key point. Fig. 5 shows registration of two images (set 1) having only rotation difference and some part of information is different in the two input images because of camera angle change. We had applied SIFT algorithm so we get key points of two images that is 638 and 508 respectively in sensed and reference image. After applying algorithm, we get only 218 interest points (indicated with blue color) which can be used to register the images.

The best candidate match for each key-point is found by identifying its nearest neighbor in the database of key-points from training images. The nearest neighbor is usually defined as the key-point with the minimum Euclidean distance. Calculating Euclidean distance [6], it needs n times multiplication computing and one times square root computing, so the time of matching is large. Distance Transform image of Chamfer Distance is similar with Euclidean distance, and the computational of Chamfer Distance is much smaller than that of Euclidean distance. The Chamfer distance is comparability measurement of SIFT feature matching in this article. Distance Transform. The effect of part image matching is showed in Fig.8, which screened in the different zoom scale, lighting conditions and rotate angle and Binocular vision matching.

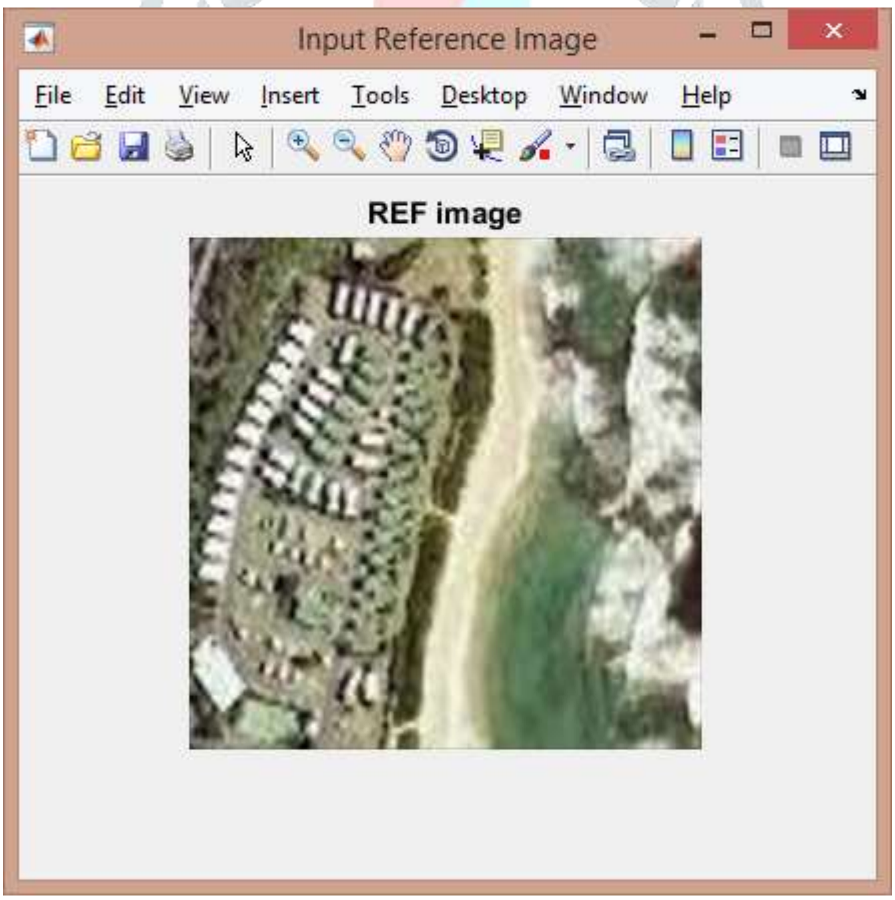
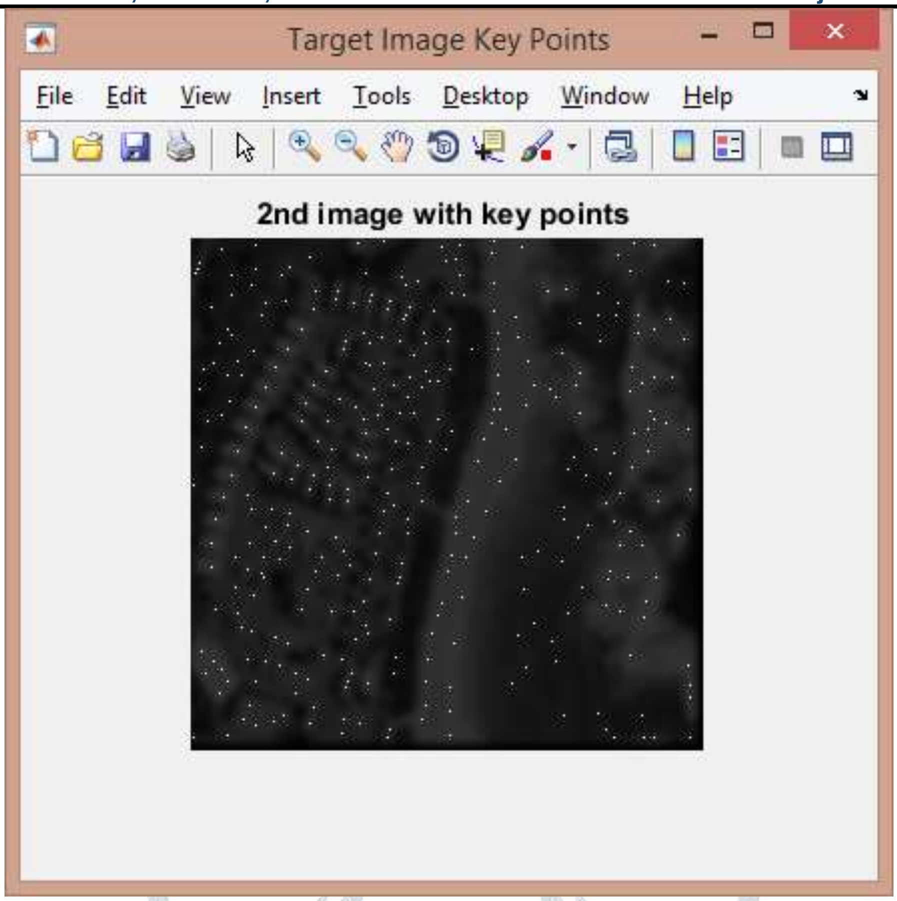


Figure Reference Image



Key Points Descriptor

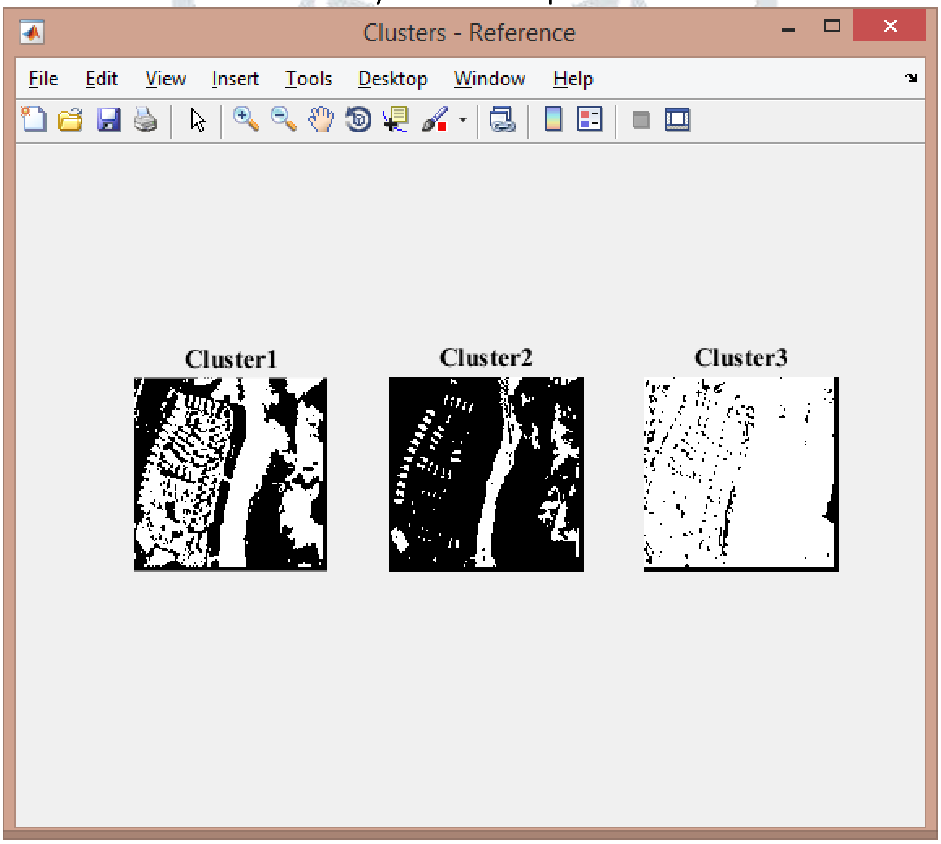


Figure 3.0 Clustering of Image

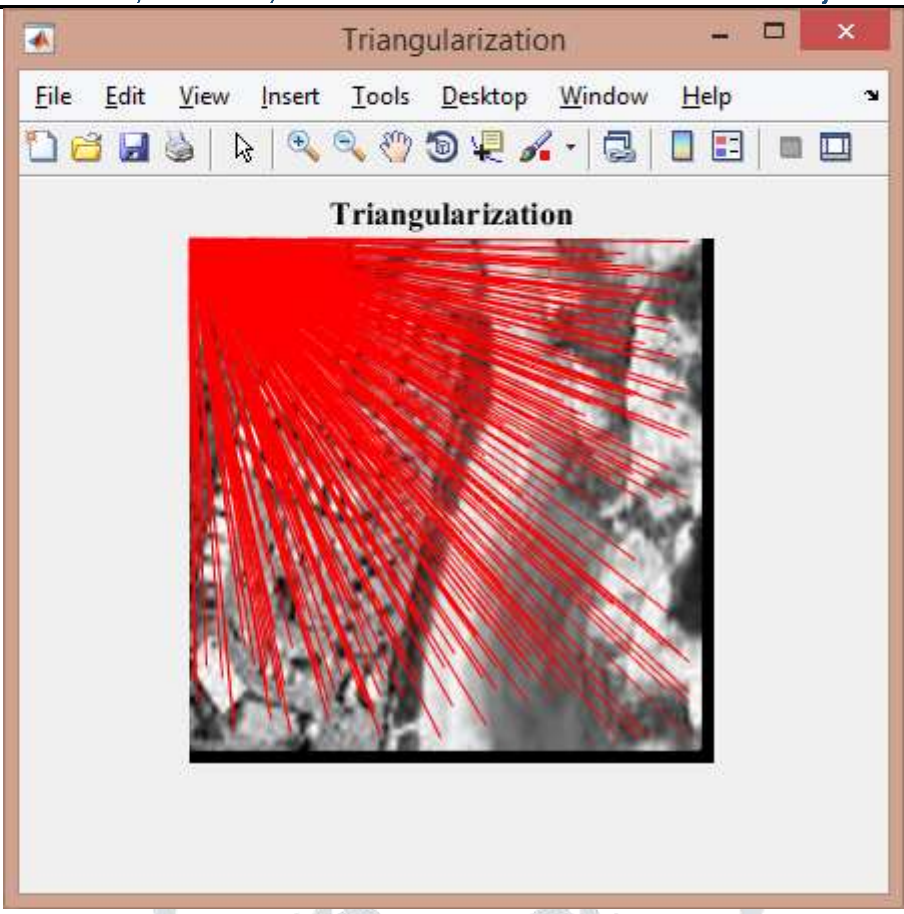


Figure 4.0 Triangularization

Table Key Index Parameters and Average readings

Accuracy	94.87 %
RMSE	2.37
Response Time	3.32 milliseconds

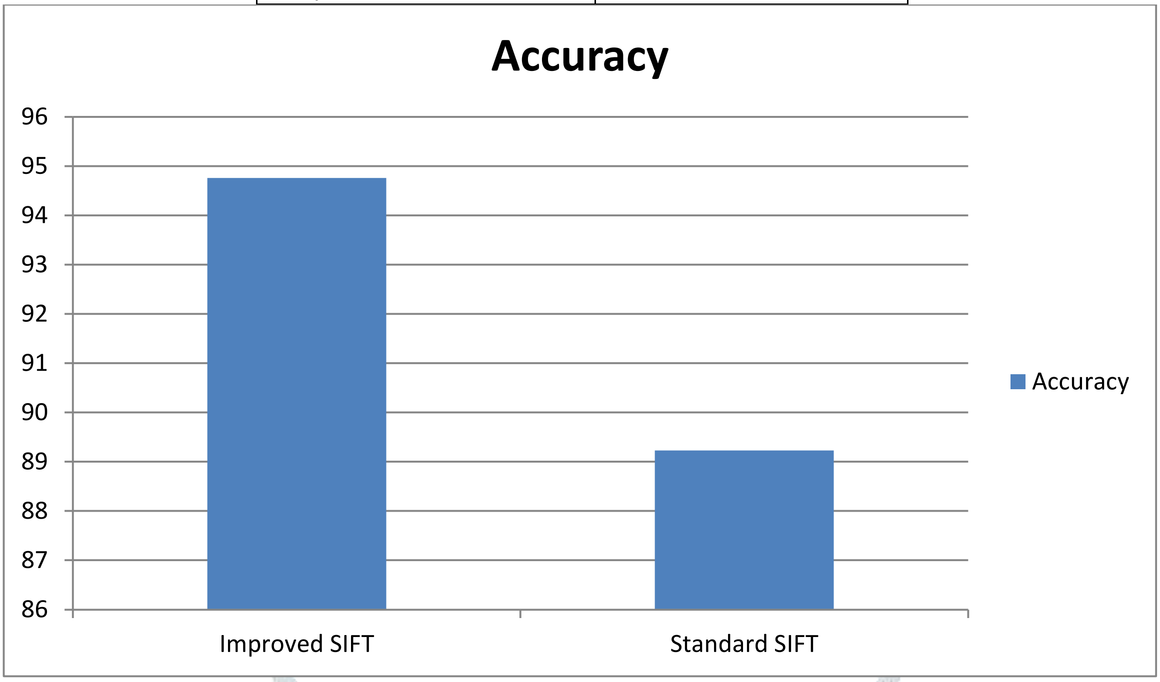


Figure 5.0 Comparison with existing method

In order to achieve orientation invariance, the coordinate of the descriptor and the gradient orientations are rotated relative to the key-point orientation. Select a 16×16 rectangular block pixels around a key-point, and divided it into 16 4×4 sub-regions. the gradient magnitude and orientation of each 4×4 sub-regions pixels are computed using pixel differences, These samples are then accumulated into orientation histograms summarizing the contents over 4×4 sub-regions, These are weighted by a Gaussian window, indicated by the overlaid circle. each orientation histogram is distributed into 0,45,90,135,180,225,270,315 angle eight direction projection, the length of each angle histogram projection is figure with arrow corresponding to the sum of the gradient magnitudes near that direction within the region.

V. CONCLUSION

Table 4.1 SIFT image matching algorithm is analyzed profoundly in this paper. The analysis of feature extraction can be used in the intelligent computing of this algorithm. Key descriptors in place of Euclidean distance reduces the time of the algorithm in key-point matching. Real-time quality and stability of algorithm are enhanced. We had tried this algorithm on 30 different set of images having different rotation, scaling, translation and illumination changes. The proposed method provides perfect results. Time taken by this algorithm is also very less. As number of key points will increase, algorithm requires more time to discard indistinct key points. This algorithm finds many distinct features from even small objects.

REFERENCES

[1] Liu Li, Peng Fuyuan, Zhao Kun, "Simplified SIFT algorithm for fast image matching," Infrared and Laser Engineering, vol.37, pp.181–184,Jan. 2008.
[2] Qian Shen, Zhu Jianying, " Improved SIFT-based bi-directional image matching algorithm," Mechanical Science and Technology for Aerospace Engineering, Vol.26,no.9,pp. 1179-1182, Sep. 2007.
[3] Zhu Daixian, Wang Xiaohua, "A method of Improving SIFT algorithm matching efficiency," in Proc.CISP' 09, 2009, vol.2, p.891.
[4] Lowe D.G. "Instinctive image features from scale-invariant key-points," International Journal of Computer Vision. 2004, p. 91~110.
[5] Lindeberg.T, "Scale-space theory: A basic tool for analyzing structures at different scales," Journal of Applied Statistics, Vol.21,no.2,pp. 224–270, Feb.1994.
[6] Hua Luo-geng. An introduction of higher mathematics, other issues, Beijing, China: Science Press, 1984
[7] David G. Lowe, "Distinctive image features from Scale Invariant Keypoint," International Journal of Comupter Vision, pp. 1-28, January 2004.
[8] Vedrana Andersen, Lars Pellarin and Renee Anderson, "Scale Invarinat Feature Transform (SIFT): Performance and Application", pp. 1-14, June 2006.

[9] Vinividyadharan, SubuSurendran "Automatic image Registration using SIFT - NCC," Special Issue of International Journal of Computer Applications on Advanced Computing and Communication Technologies for HPC Applications, pp. 29 – 32, June 2012.

[10] Mahesh and Subramanyan M. V., "Automatic feature based image registration using SIFT algorithm," IEEE-20180, ICCCNT'12, July 2012.

[11] K.V.Arya, P. Gupta, P.K. Kalra, P. Mitra, " Image registration using robust M – estimators" Pattern recognition letters, Science Direct, pp.:1957 – 1968, May 2007.

[12] Meisen Pan, Jingtian Tang, Qi Xiong, " Medical image registration using fuzzy theory" , Computer methods in biomechanics and biomedical engineering, Taylor & Francis, pp: 721- 734.

[13] Mehfuza Holia, Heena kher, V.K.Thakar, " Registration of license plate images captured at different angles using mutual information" 2011 IEEE International conference on computational intelligence and computing research, pp.: 989 – 997, December 2011.