Comparative Analysis of Feature Detectors for Automatic Satellite Image Registration

Ami S. Shete¹, Sandip R. Panchal²

PG Student, Department of Electronics & Communication Engineering, SVIT, Vasad, India¹
Asst. Professor, Department of Electronics & Communication Engineering, SVIT, Vasad, India²

ABSTRACT: Image registration is the first step towards using the satellite images for any purpose. It is the fundamental task used to match two or more partially overlapping multi view, multi modal or multi temporal images and stitches these images into one image comprising the whole scene. Automatic satellite image registration is a challenging task of overlaying two images for geometric conformity aligning common features by establishing a transformation model using distinguishable feature points collected simultaneously in reference image and the sensed images in a completely unassisted manner. Accuracy of registered image depends on accurate feature detection and matching. This paper aims to present a comparative analysis of feature detection techniques like Scale invariant Feature Transform (SIFT), Speeded Up Robust Features (SURF), Harris Corner Detector and Improved version of Harris Corner Detector used for automatic satellite image registration.

KEYWORDS: Feature detection, SIFT, SURF, Harris Corner Detector, Improved Harris Corner Detector

I. INTRODUCTION

Image registration is the process of geometrically aligning two or more images of the same scene taken by different sensors, at different times, and/or from different viewpoints. It geometrically aligns two images—the reference image and the sensed image [1]. The image with respect to which the alignment is carried out is called the reference image. The image which is aligned is called the sensed image. The transformed sensed image which aligns with the reference is called registered image [2].

The main idea behind the any image registration process is that the sensed image undergoes the registration process and its pixel coordinates are converted into the reference image pixel coordinates. In this way we get the transformed sensed image. Then this transformed sensed image is super imposed on the reference image in visually plausible way. Once we have super imposed both the images then we have a larger 2D view of the scene or highly informative single output image [3].

Image registration is the foundation of applications, such as image fusion, medical image processing, remote sensing and three dimensional (3D) image reconstructions [4].

There are two types of image registration named as area based image registration and feature based image registration. The area based methods are used when distinctive and important information is provided by pixel intensity [1]. They use some statistical information to measure the degree of similarity of the whole image [5]. The feature based methods are used when important information is given by the image features like point, edge, corners, and contours [6].

Compared to area based methods, feature based methods are more widely applied in remote sensing application due to their advantages. The area based methods find correspondences in the image space whereas the feature based methods find correspondences in the feature space which represents information at higher and abstract level. If there is complex distortion between the images to be aligned, then the computational complexity or the search space of the area based methods increases nonlinearly with the transformation complexity. The feature-based methods can overcome this drawback as their search space is proportional to the number of features detected from the images. Sometimes the selected features are invariant to the changes of the image’s geometric and radiometric conditions, presence of noise, and the changes in the target scene. Therefore, this type of method is suitable for the situations where multisensor analysis is...
demanded or illumination changes are expected [7]. Feature-based methods are capable of registering the images with distinctive features, such as map and photograph, as well as those with complex distortions [8]. One of the main advantages of feature based methods is that they are fast and robust to noises, significant radiometric differences, and complex geometric distortions [9].

In remote sensing applications, the conventional image registration is generally carried out. Conventional image registration techniques involve manual selection of control points (CPs) which are used to estimate the geometric transformation model that establishes a mapping between reference and sensed image. Manual registration is not feasible in the cases where large amount of data is to be processed. Also the conventional method needs an expert with a special skill to select the individual CPs precisely for estimating the transformation model which is a laborious activity. Thus, automated techniques that require little or no operator supervision is needed. An Automatic Image Registration (AIR) technique can solve the drawbacks of conventional methods and it is a highly desirable requirement of the remote sensing world to deal with large volumes of satellite data available for quick and accurate registration [10], [11], [12]. The main concept of automatic image registration for satellite images is to obtain acute set of CPs and then apply the transformation model which is most suitable to the pair of images to be registered [13]. Accuracy of registered image depends on accurate feature detection and matching.

This paper focuses on feature detection step and shows comparative analysis of feature based methods such as Scale Invariant Feature Transform (SIFT), Speeded Up Robust Feature (SURF), Harris corner detector and Improved Harris corner detector.

The paper is organized as follows. The section II contains the related work, the section III describes the overview of the feature detection methods, the section IV discusses the experiments and results and finally conclusion is specified in section V.

II. RELATED WORK

Scale Invariant Feature Transform (SIFT) detector detects the local extrema in scale space as candidate feature by approximating the Laplacian with difference of Gaussian filter [14] [15]. It allows distinctive invariant feature extraction from images and it can be applied to perform reliable matching between sensed and reference image presenting a substantial range of affine distortion, addition of noise, changes in illumination and change in 3-D viewpoint [13]. SIFT detector improves the detection stability even if the input is noisy. The detected features are highly distinctive. It has proven to be good for multi-angle imagery [16].

It is invariant to scale, illumination changes, rotation and is preferable to certain extent in 3D camera viewpoint [16] [14]. It can be ineffective in finding CPs for high view angles in areas with low elevation differences [16]. It suffers from high complexity while extracting feature points and it also lacks in feature points and distribution quality [17].

Speeded Up Robust Feature (SURF) is a scale invariant feature detector that uses integral images and Hessain matrix for very fast computation of detectors. Using a set of box filters, Hessain matrix is roughly approximated and no smoothing is applied when going from one scale to another. SURF approximates second order derivatives. SURF is five times faster than Difference of Gaussian [18]. It has low accuracy and processing time is slightly improved as compared to SIFT [17]. It detects less number of features as compared to SIFT [19]. It is fast detector but it is not stable to rotation and illumination changes [20].

The Harris corner detector is based on the local auto correlation function of the signal. This detector determines whether a point shows significant change in all the directions to designate it as a corner point. It has invariant to rotation, translation and illumination change. It is most informative and most repetitive detector [18]. The drawback of this detector is that it is not invariant to large scale change [21]. Modifications are performed on the standard Harris corner detector which gives better performance in several conditions. Harris-Laplace and Harris-Affine are scale and affine invariant versions of standard Harris corner detector [18].

Improved Harris corner detector uses only first-order derivatives and is one of the most stable and robust corner detectors [12]. It gives better result than Harris corner detector in case of rotation, scaling, illumination changes and viewpoint change. The interest points are largely independent of the imaging conditions and are geometrically stable for this detector [21]. In Harris corner detector, the corner response function involves the use of constant parameter \( k \).
A. SIFT Detector:

Scale Invariant Feature Transform (SIFT) detector detects the local extrema in scale space as candidate feature by approximating the Laplacian with difference of Gaussian filter [14] [15]. It allows distinctive invariant feature extraction from images and it can be applied to perform reliable matching between sensed and reference image presenting a substantial range of affine distortion, addition of noise, changes in illumination and change in 3-D viewpoint [13].

The SIFT algorithm has four main steps: (1) Scale Space Extrema Detection, (2) Key point Localization, (3) Orientation Assignment and (4) Description Generation.

The first stage is to identify location and scales of key points using scale space extrema in the DoG (Difference-of-Gaussian) functions with different values of σ, the DoG function is convolved of image in scale space separated by a constant factor k as in the following equation.

\[
D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma) \times I(x, y)
\]

where, \( G \) is the Gaussian function and \( I \) is the image. Now the Gaussian images are subtracted to produce a DoG, after that the Gaussian image subsample by factor 2 and produce DoG for sampled image. A pixel compared of \( 3 \times 3 \) neighborhood to detect the local maxima and minima of \( D(x, y, \sigma) \). In the key point localization step, key point candidates are localized and refined by eliminating the key points where they rejected the low contrast points. In the orientation assignment step, the orientation of key point is obtained based on local image gradient. In description generation stage is to compute the local image descriptor for each key point based on image gradient magnitude and orientation at each sample point in a region centered at key point; these samples building 3D histogram of gradient location and orientation; with \( 4 \times 4 \) array location grid and 8 orientation bins in each sample. That is 128-element dimension of key point descriptor.

B. SURF Detector:

Speeded Up Robust Feature (SURF) is a scale invariant feature detector that uses integral images and Hessain matrix for very fast computation of detectors. Using a set of box filters, Hessain matrix is roughly approximated and no smoothing is applied when going from one scale to another. SURF approximates second order derivatives. SURF is five times faster than Difference of Gaussian [18].

In image \( I \), \( x = (x, y) \) is the given point, the Hessian matrix \( H(x, \sigma) \) in \( x \) at scale \( \sigma \), it can be defined as

\[
H(x, \sigma) = \begin{bmatrix}
Lxx(x, \sigma) & Lxy(x, \sigma) \\
Lxy(x, \sigma) & Lyy(x, \sigma)
\end{bmatrix}
\]

where \( Lxx(x, \sigma) \) is the convolution result of the second order derivative of Gaussian filter \( \frac{\partial^2}{\partial x^2} g(\sigma) \) with the image \( I \) in point \( x \), and similarity for \( Lxy(x, \sigma) \) and \( Lyy(x, \sigma) \).

SURF creates a “stack” without 2:1 down sampling for higher levels in the pyramid resulting in images of the same resolution. Due to the use of integral images, SURF filters the stack using a box filter approximation of second-order Gaussian partial derivatives, since integral images allow the computation of rectangular box filters in near constant time.

C. Harris Corner Detector:

The Harris corner detector is based on the local auto correlation function of the signal. This detector determines whether a point shows significant change in all the directions to designate it as a corner point. A matrix related to the autocorrelation function that takes into account the first derivatives of the signal on a window is computed. The
eigenvectors of this matrix are the principal curvatures of the autocorrelation function. Two significant Eigen values indicate the presence of an interest point as shown in Figure 2.2 [22].

Harris detector detects the L-junctions and points with the higher curvature along with the corner points [18]. Here we find the second moment matrix which requires finding the gradients of an image which is not sensitive to noise and computationally expensive.

Harris corner detector gives a mathematical approach for determining which case holds when we shift the window over an image in different directions.

- **Flat region**: No change in intensity in all directions
- **Edge**: No change in intensity along the edge direction
- **Corner**: Significant change in intensity in all directions

Suppose the current pixel intensity is \( I(x,y) \). The window may be step or Gaussian and is denoted by \( w(x,y) \) then the change in intensity for the shift \((u,v)\) is the sum of square difference between the shifted intensity and the current intensity as given below.

\[
E(u,v) = \sum_{x,y} w(x,y) [ I(x+u,y+v) - I(x,y) ]^2
\]  
(3)

For the constant intensity region the intensity difference in above equation is small or zero and therefore \( E(u,v) \) will be small or zero. Hence ultimately we want to find the point for which \( E(u,v) \) is large.

The Taylor series for 2D functions is given by

\[
f(x + u, y + v) = f(x, y) + \frac{uf_x(x,y) + vf_y(x,y)}{1!} + \text{Second partial derivative terms} + \cdots + \text{higher order terms.}
\]  
(4)

The first order approximation of above equation is given by

\[
f(x + u,y + v) \cong f(x,y) + \frac{uf_x(x,y)+vf_y(x,y)}{1!}
\]  
(5)

Hence substituting above equation in equation (3) will give

\[
E(u,v) = \sum_{x,y} w(x,y) [ I(x+u,y+v) - I(x,y) ]^2
\]  
(6)

\[
E(u,v) = \sum_{x,y} w(x,y) [ I(x,y) + uI_x + vI_y - I(x,y) ]^2
\]  
(7)
International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

\[ E(u, v) = \sum_{x,y} w(x, y) [u^2 I_x^2 + 2uv I_x I_y + v^2 I_y^2] \]  

(8)

Rewrite as a matrix equation

\[ E(u, v) = \sum_{x,y} w(x, y) [ I_x^2 \quad I_x I_y \quad I_y^2 ] [u, v]^T \]

(9)

\[ E(u, v) = [u, v] \left( \sum_{x,y} w(x, y) [ I_x^2 \quad I_x I_y \quad I_y^2 ] \right) [u, v]^T \]

(10)

Hence for small shift of \([u, v]\) we have

\[ E(u, v) \cong [u, v] M [u, v]^T \]  

(11)

Where \(M\) is 2x2 matrix computed from the image derivatives and it is called as Harris matrix. \(M\) contains the products of the image derivatives \(I_x\) and \(I_y\).

\[ M = \sum_{x,y} w(x, y) [ I_x^2 \quad I_x I_y \quad I_y^2 ] \]

(12)

Chris Harris & Mike Stephens [22] have defined the corner response as

\[ R = \det M - k(\text{Trace} M)^2 \]  

(13)

Here, \(\det M = \lambda_1 \lambda_2\) and \(\text{Trace} M = \lambda_1 + \lambda_2\). Hence \(R\) is completely depending on two Eigen values of matrix \(M\) that are \(\lambda_1\) and \(\lambda_2\). Here \(k\) is the constant having the value between 0.04 to 0.06.

- If the \(\lambda_1\) and \(\lambda_2\) both are small then \(|R|\) will be smaller and that indicates the flat region in the image.
- If the \(\lambda_1\) or \(\lambda_2\) is large then \(R\) will be negative with large magnitude and that indicates the edge in the image.
- If the \(\lambda_1\) and \(\lambda_2\) both are large then \(R\) will be larger and that indicates the corner in the image.

Once the \(R\) is calculated then apply threshold to get the corner points. Hence after thresholding we have only those points which have larger value of \(R\) (larger then threshold value) and this points represents the corner in the image. Once the corner points are obtained we perform the non-maxima suppression using 3x3 mask around each corner point by checking the 8 neighborhood pixel values. If any of the neighbor pixel value is higher than the corner strength then the current corner is replaced by that neighbor pixel.

**Algorithm:**

- Compute the image derivatives \(I_x\) and \(I_y\)
- Compute the product of the derivatives at every pixels
  - \(I_x^2 = I_x \ast I_x\)
  - \(I_y^2 = I_y \ast I_y\)
  - \(I_x I_y = I_x \ast I_y\)
- Compute the sums of the product of derivatives at each pixel
- Define \(M\) for each pixel \((x, y)\) shown in equation
- Compute Harris corner response \(R = \det M - k(\text{Trace} M)^2\) as equation (13)
- Apply threshold on value of \(R\) and perform non maxima suppression
D. Improved Harris Corner Detector:

Harris corner detector uses only first-order derivatives and is well known as one of the most stable and robust corner detectors in image processing [12]. In Harris corner detector, the corner response function involves the use of constant parameter k [22]. For better result, k=0.04 should be used [23]. Now in improved Harris corner detector, the corner response function is modified and is made independent of constant parameter k [24].

In improved Harris corner detector, the value of $\sigma=1$ is taken. The corner response is calculated as

$$R = \frac{\det M}{\text{Trace} M}$$

(14)

Finally, in the matrix, the point that meeting the $R(x, y)$ is greater than the threshold value and $R(x, y)$ is the local maximum value of a certain neighborhood is considered corner.

IV. EXPERIMENTS & RESULTS

A. Evaluation measurement:

The repeatability rate is defined as the number of points repeated between two images w.r.t. total number of detected points. To measure the number of repeated points, Normalised Cross Correlation is used. The repeatability rate is given as

$$R_{1,2} = \frac{C(I_1, I_2)}{\min(m_1, m_2)}$$

(15)

where, $C(I_1, I_2)$ is the number of corresponding couples, $m_1, m_2$ are the number of points detected in image 1 and 2 respectively.

B. Processing Time:

The time evaluation is a relative result, which only shows the tendency of the three methods’ time cost.

To verify the effectiveness of the algorithms, two images are taken. Image 1 is original image and the image 2 is rotated, scaled or blurred image. The experiments are performed on Intel(R) Core(TM) i3-2350M CPU @ 2.30 GHz processor and 4 GB RAM with windows 7 as an operating system. Features are detected in both images using SIFT, SURF, Harris corner detector and Improved Harris corner detector algorithm.

![Fig. 2. (a) Original Image (b) Gray scaled version of original image](image_url)

Table 1 shows the detected features in image 1 and image 2 which is 10° rotated, 0.5 scaled and blurred using 5x5 average mask. Table 2 shows the comparison of different feature detection algorithms.
TABLE: 1 Detected features of different feature detection algorithms

<table>
<thead>
<tr>
<th>Method</th>
<th>SIFT</th>
<th>SURF</th>
<th>Harris Corner Detector</th>
<th>Improved Harris Corner Detector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image 1</td>
<td><img src="image1_sift" alt="Image 1 SIFT" /></td>
<td><img src="image1_surf" alt="Image 1 SURF" /></td>
<td><img src="image1_harris" alt="Image 1 Harris" /></td>
<td><img src="image1_improved_harris" alt="Image 1 Improved Harris" /></td>
</tr>
<tr>
<td>Image 2 Rotation (10°)</td>
<td><img src="image2_sift_rotation" alt="Image 2 SIFT Rotation" /></td>
<td><img src="image2_surf_rotation" alt="Image 2 SURF Rotation" /></td>
<td><img src="image2_harris_rotation" alt="Image 2 Harris Rotation" /></td>
<td><img src="image2_improved_harris_rotation" alt="Image 2 Improved Harris Rotation" /></td>
</tr>
<tr>
<td>Image 2 Scaling (0.5)</td>
<td><img src="image2_sift_scaling" alt="Image 2 SIFT Scaling" /></td>
<td><img src="image2_surf_scaling" alt="Image 2 SURF Scaling" /></td>
<td><img src="image2_harris_scaling" alt="Image 2 Harris Scaling" /></td>
<td><img src="image2_improved_harris_scaling" alt="Image 2 Improved Harris Scaling" /></td>
</tr>
<tr>
<td>Image 2 Blur (Avg 5x5 mask)</td>
<td><img src="image2_sift_blur" alt="Image 2 SIFT Blur" /></td>
<td><img src="image2_surf_blur" alt="Image 2 SURF Blur" /></td>
<td><img src="image2_harris_blur" alt="Image 2 Harris Blur" /></td>
<td><img src="image2_improved_harris_blur" alt="Image 2 Improved Harris Blur" /></td>
</tr>
</tbody>
</table>

TABLE: 2 Comparison of different feature detection algorithms

<table>
<thead>
<tr>
<th>Method</th>
<th>No of features Detected</th>
<th>Processing time (in sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Image 1</td>
<td>Image 2 Rotation (10°)</td>
</tr>
<tr>
<td>SIFT</td>
<td>227</td>
<td>285</td>
</tr>
<tr>
<td>SURF</td>
<td>340</td>
<td>360</td>
</tr>
<tr>
<td>Harris Corner Detector</td>
<td>418</td>
<td>339</td>
</tr>
<tr>
<td>Improved Harris Corner Detector</td>
<td>564</td>
<td>430</td>
</tr>
</tbody>
</table>
Fig. 3. Repeatability comparison for (a) Rotation (b) Scaling (c) Blurring and Comparison of (d) Processing time

V. CONCLUSION

This paper has evaluated four feature detection methods for automatic satellite image registration. Based on the experimental results, it is found that the Improved Harris Corner Detector detects more features and has better repeatability rate for rotated and blurred images but suffers with scaled images. SURF gives better result for scaled images. The processing time for Improved Harris Corner Detector is the least in all the cases than other detectors. So it gives better results in less time.

REFERENCES

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016


BIOGRAPHY

Prof. Sandip R. Panchal has received his M.E. (Automatic Control & Robotics) degree from Faculty of Technology & Engineering, M S University of Baroda, Vadodara, Gujarat, India. He is pursuing Ph.D from CHARUSAT, Changa, Gujarat, India. His research interest includes Image Registration, Image Processing and Control System. He has 14 years of teaching experience. He has published number of papers in National/International Journals and Conferences. He has published a book titled Circuits & Networks, 2/e (Atul Prakashan, Year of Publication: 2010-11). Prof. S R Panchal has professional membership of ISTE (Indian Society for Technical Education), IETE (Institution of Electronics & Telecommunication Engineering), IE (Institute of Engineers, India).