A Detailed Analysis on Feature Extraction Techniques of Panoramic Image Stitching Algorithm

Shivangi Pandey
Department of Electrical & Electronics Engineering
National Institute of Technical Teachers' Training & Research
Bhopal, India

Anjali Potnis
Department of Electrical & Electronics Engineering
National Institute of Technical Teachers' Training & Research
Bhopal, India

Madhuram Mishra
Department of Electrical & Electronics Engineering
National Institute of Technical Teachers' Training & Research
Bhopal, India

Abstract—**Image stitching is a technique which is used for attaining a high resolution panoramic image. In this technique, distinct aesthetic images that are imaged from different view and angles are combined together to produce a panoramic image. In the field of computer graphics, photographic and computer vision, Image stitching techniques are considered as current research areas. For obtaining a stitched image it becomes mandatory that one should have the knowledge of geometric relations among multiple image co-ordinate system [1].First, image stitching will be done based on feature key point matches. Final image with seam will be blended with image blending technique. Hence in this paper we are going to address multiple distinct techniques like some invariant features as Scale Invariant Feature Transform and Speeded up Robust Transform and Corner techniques as Harris Corner Detection Technique that are useful in sorting out the issues related with stitching of images.**

Keywords-SIFT; SURF; RANSAC; Harris Corner Detector; Panoramic Image.

I. INTRODUCTION

Image stitching is an artificial combination of series of distinct picturesque images. For stitching of an image some isolated transformations are required to perform through warping and merging operations. By merging the overlapping fields of warped image, we can easily achieve a tantamount image of same objects, which covers the visible area of scene completely. [2] Shortly after the photographic process was developed in 1839, the use of photographs was demonstrated on topographical mapping. In introductory years when the technology was not so urbanized on that time for creating the panorama 'n' number of cameras were required to capture the images which were positioned at poles apart locations and at divergent angles. As we know that for creating the a panoramic image, one should have the best knowledge about the geometrical relationships among the co-ordinate system of the images that are going to be stitched, but as in introductory period the geometric relations were not taken into consideration hence a good panoramic image was difficult to created. For

producing the panoramic image some steps are required that are shown as below diagrammatically.

The below drawn figure shows the distinct steps that are to be taken into consideration for generation of panoramic image.
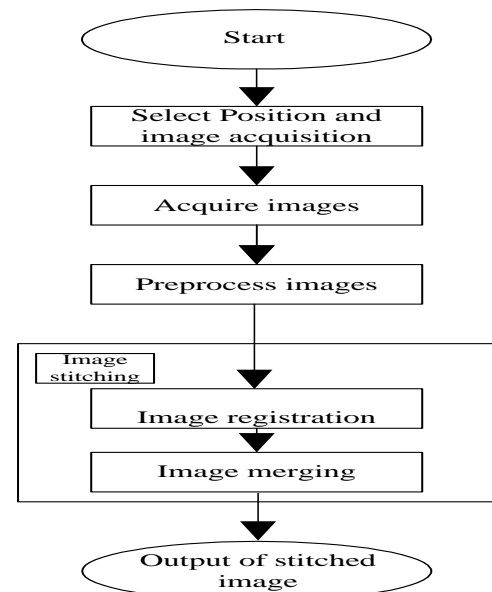

Figure 1. Flow Chart for Generating Panoramic Image

As the images are acquired that are to be stitched, some preprocessing is required like projection of the images need to be made on surface that can be any mathematical surface like spherical and cylindrical. The combination of image registration and image merging helps in further processing of image stitching process.

The main work of Image registration is to correlate two or more than two images that belong with the same scene. Here one image is considered as referenced image. The other images that are about to stitch undergoes for geometrical metamorphosis. Many apprehensions are investigated which

are responsible for wrong calibration of multiple images that are captured like if we are capturing the image from a dynamic platform or maybe there are some distortions in lens or sensors.
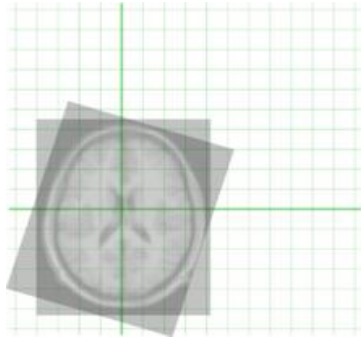

Figure 2.   Registration of Two MRI Images of Brain

There are various application are available of image registration. The flow chart evolving panoramic image is shown below.
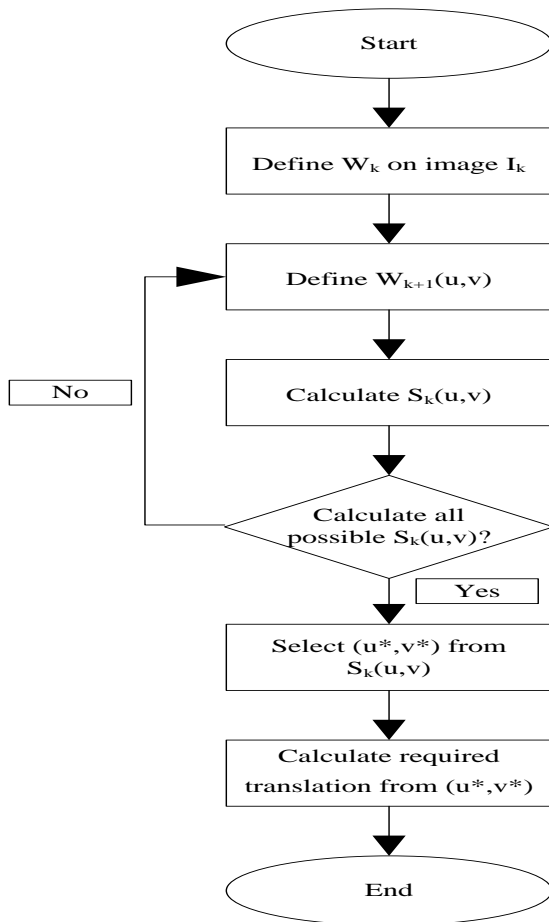

Figure 3.   Flow Chart of Image Registration Technique for Generating Panoramic Image

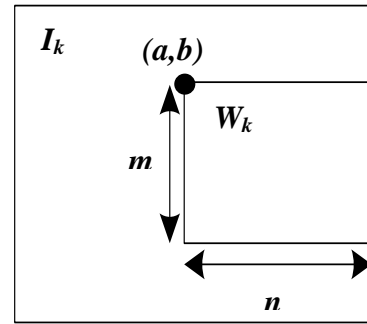Where W_k gives a $m \times n$ window that is defined on I_k.

Figure 4. W_k is the $m \times n$ Window at location (a,b) in I_k.

By averaging the intersection of red, green and blue channels of kth images in sequence of applied input image an image will be obtained given as I_k.

The variable k can be from I to the total no. of images in sequence of images.

$W_{k+1}(u,v)$ is an another window with the position (u,v).

$S_k(u,v)$ is called as similarity measure with the position (u,v) and its value can be given as,

$$S_k(u,v) = \sum_{i=1} m \sum_{j=1} n |W_k(i,j) - W_{k+1(u,v)}(i,j)| \qquad (1)$$

Again the value of similarity measure is being calculated which is shown as $S_k(u^*,v^*)$ at the optimal matching position (u^*,v^*) and its value is given as,

$$S_k(u^*,v^*) = min_{1 \leq i \leq H-M, 1 \leq j \leq L-M}^{\{S_k(i,j)\}} \qquad (2)$$

As the process of registration is being completed now we need to merge the images for producing a panoramic image hence process of merging is called image merging.

II.   APPROACHES OF IMAGE STITCHING ALGORITHM

There are various approaches used for stitching of image, here we are going to discuss only two of them.

(i)   Direct Approach: Direct techniques helps in reducing the summation of accurate differences between imbricating pixels of an image. In this method, each pixel value is compared with each other. These approaches are having conglomerate characteristics. In this approach, juxtaposition is made among all the pixel intensity values of images that are going to stitch. There are also some benefits of using this method given as follows.

a)   As in this method, each pixel value is compared with each other; hence contribution of each pixel is also measured.

b)   Direct techniques aims at using the existing data in image coalition optimally.

(ii) Feature Based Approach: The compilation of image feature (extrema) points can be performed by measuring the

whole features of the existing image with the available images. Various distinct steps are used in feature based techniques for extraction of features, registration and blending. For feature extraction of several images we basically use number of techniques like Scale Invariant Feature Transform (SIFT), Speeded up Robust Feature (SURF) and Harris Corner Detection.

As Scale Invariant Feature Transform (SIFT) technique has sturdiness but it is having low calculation speed hence it is not fine for real time applications. Speeded up Robust Feature (SURF) is more exceptional than SIFT and it produces better computational cost. For improving the computational cost an integral image is used. One more technique is introduced which is used for feature extraction is known as Harris Corner Detector. Harris Corner Detector is not invariant to scale changes and cross correlation [4].

III. MODEL OF IMAGE STITCHING TECHNIQUE

In this section we will discuss about the model of image stitching technique. Some vital steps are to be held on for this.
(i)   Image acquisition
(ii) Feature Detection and Matching
(iii) Image Matching RANSAC Translation Estimation
(iv) Global Alignment
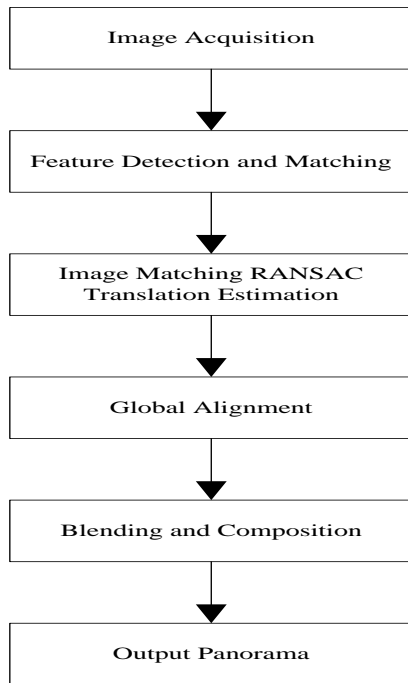(v)   Blending and Composition

Figure 5.   Image stitching model

(i) Image Acquisition: Image acquisition is the primary stride that pacts with capturing an input image for engendering diverse separate panoramic images. The acquisition action can be performed in various ways like acquisition by camera

rotations, acquisition by camera translation and acquisition by hand held camera.
a)   Acquisition by camera rotation: In this method a tripod is used and it is set at a particular chosen location. Its location remains constant throughout the whole acquisition process.
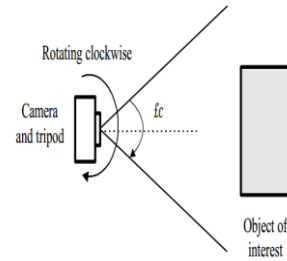

Figure 6.   Acquisition by Camera Rotation using Tripod and Camera

The geometry of overlapping images is as shown in figure given below
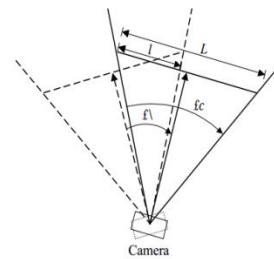

Figure 7.   Geometry of overlapping images

The ratio of the width of the overlapping region to the width of the image can be calculated as,

$$\frac{l}{L} = \frac{\frac{L}{2}\left[1 - \frac{\tan(\alpha - \frac{\theta}{2})}{\tan\frac{\theta}{2}}\right]}{L} \qquad (3)$$

This can also be written as, $\frac{l}{L} = \frac{1}{2}\left[1 - \frac{\tan(\alpha - \frac{\theta}{2})}{\tan(\frac{\theta}{2})}\right]$ $\qquad$ (4)

Where L = width of acquired image
l = the width of overlapping region between adjacent images.
There are a number of advantages are illustrated of using the camera rotation for image acquisition some are listed as below,
(i)   In this method camera remains constant at single position for capturing a number of images.
(ii) Measurement requirement is less and this can be executed smoothly.
b)   Acquisition by Camera Translation: In this proposed method, camera position does not remain same it keeps on moving parallel in direction of imaging plane. For this purpose camera is positioned onto a sliding plate. For capturing the image directly, the camera and sliding plates are implanted in front of object and it captures the images until a succession of image does not envelop the entire range.
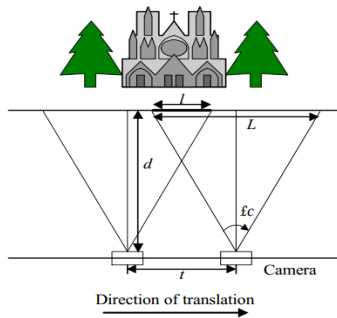
Figure 8.   Geometry for image acquisition for camera translation

Where l = the width of overlapping region between adjacent images.
L = length of acquired image.
t = camera translation
d = object of interest
The ratio of the overlapping image to the whole image can be calculated as,

$$\frac{l}{L} = 1 - \frac{t}{2dtan(\frac{\theta}{2})} \qquad (5)$$

c)   Acquisition using Hand Held Camera: Working with the handheld camera is very simple as just take the camera and start capturing the images either by staying at the same place only by rotating your body or moving along the directions of imaging plane. Bu the images captured by this method is difficult to stitch because of unnecessary camera rotations for acquiring the images.

(ii) Feature Detection and Matching: The base of image stitching model lies under the two main steps as disclosure of feature and feature matching. It would be favorable for analyzing some important features of the images than going through the whole image. Multiple feature extraction techniques are used like Scale Invariant Feature Transform (SIFT), Speeded up Robust Feature (SURF) and Harris Corner Detection.
Some leverage related with this approach is as shown below.

a)   This algorithm produces more robustness in opposition to scene movement.
b)   These algorithms are probably faster.

(iii) RANSAC for Homography: Homography basically deals with relating the two images that are having the same planar surface in space. Homography is adopted for rectification of images, registration of images and analyzing the camera translation rotation and movement between two images. Homography is a 3×3 matrix M given as below,

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \qquad (6)$$

Homography relates the pixel co-ordinates in two images if $x' = Mx$

Where x and x' are the points given as $x = (u, v, 1)$ in one image and $x' = (u', v', 1)$ in another image.

In this section, the parameter of homography is determined using Random Sample Consensus (RANSAC) algorithm. RANSAC loop involves selecting four feature pairs (at random); compute Homography H (exact); compute inliers, keep largest set of inliers, and finally it recomputed least-squares H estimate on all of the inliers [3]. RANSAC algorithm finds its goal by choosing a random subset of original data iteratively. RANSAC is used for adjusting the model existence of data outliners. Here we are going to discuss a fitting problem with parameter 'x'. Following are the assumptions that are to taken into consideration for determining parameters.

a)   'N' number of data items is used for determining the parameters.
b)   Total number of available data item is M.
c)   If in any situation algorithm fails to find out a good fit test then the probability if one exists is denoted as P_{fail}.

The algorithm for RANSAC is given presented by using a flow chart.
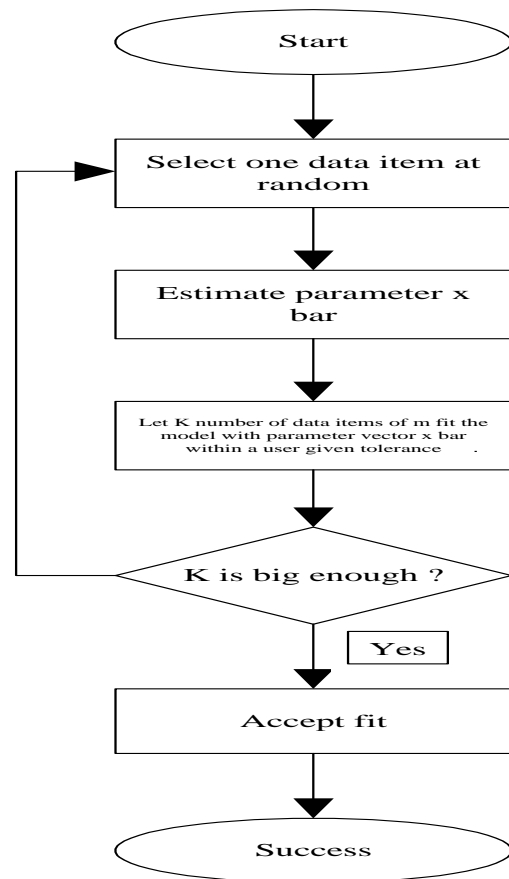

Figure 9.   Flow Chart for RANSAC Algorithm

(iv) Global Alignment: in case that we are having the numerous images of same scene and want to combine all of them in an exact 3D reconstruction so for this purpose bundle adjustment is contemplate as the best technique. The aim of this step is to find a globally consistent set of alignment parameters that minimize the miss-registration between all pairs of images. Initial estimates of the 3D location of features in the scene must first be computed, as well as estimates of the camera locations [3]. Then, bundle adjustment applies an iterative algorithm to compute optimal values for the 3D reconstruction of the scene and camera positions, by minimizing the log-likelihood of the overall feature projection errors using a least-squares algorithm [6]. If we are given an unordered set of images to register, we need to discover which images go together to form one or more panoramas [5].

(v) Blending and Composition: The fundamental objective of blending as well as composition steps is to administer an attractive panoramic image.

IV. EXPLANATION OF FEATURE EXTRACTION TECHNIQUES

In field of image processing various feature extraction techniques are used some are given as below

(i) Scale Invariant Feature Transform (SIFT)

(ii) Speeded up Robust Feature (SURF)

(iii) Harris Corner Detection

(i) Scale Invariant Feature Transform (SIFT): This feature extraction technique was developed by D.G. Lowe in 2004. SIFT multi-scale feature relies on the Gaussian function to scale the image transformation into a single multi-scale image space, on which stable feature points can be extracted [7].The below given equation is used in determining the scale space of an image,

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \qquad (7)$$

Where, σ represents the scale space factor.
$I(x,y)$Represent applied input image.
$G(x,y,\sigma)$Represent 2D Gaussian convolutional cord.

The value of $G(x,y,\sigma)$ is as given,
$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \qquad (8)$$

For efficiently detecting the key points in scale space, difference of Gaussian (DOG) of scale space is calculated as,
$$D(x,y,\sigma) = \big(G(x,y,K\sigma) - G(x,y,\sigma)\big) * I(x,y) \qquad (9)$$
This also can be written as,
$$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \qquad (10)$$
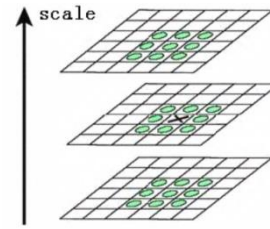

Figure 10. Local Maximum Detection in DoG Scale Space

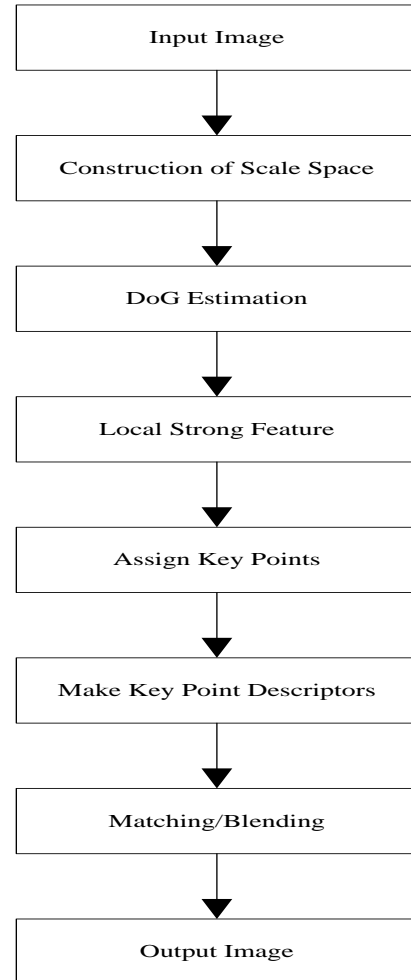Scale Invariant Feature Transform can be given diagrammatically hence it is shown as below.

Figure 11. Flow Chart for SIFT

In SIFT, Key points are first extracted by searching over all scales and image locations, then the descriptors defined on the key point neighborhood are computed, through to compare the Euclidean distance of their descriptors to extract the initial feature points pair, then eliminate spurious feature points pair by applying RANSAC, finally transform the input image with the correct mapping model for image fusion and complete image stitching [8].

(ii) Speeded up Robust Feature (SURF): In the list of feature extraction techniques, Speeded up Robust Feature (SURF) is one of them. It is very famous algorithm. It was given by Herbert Bay et. al. in 2006. It can be used in various tasks like object recognition or 3D reconstruction. As SURF algorithm provides better result than SIFT and it is several times faster than SIFT. For detection of interest points, SURF uses an integer approximation of determinant of Hessian Blob Detector. Square shaped filter are used as an approximation.

If we are using the integral image, then square shaped filters provides the best result.

$$S(x,y) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i,j) \qquad (11)$$

The SURF detector is based on the determinant of the Hessian matrix [9]. Let we have a point $X = (x,y)$ in an image I, then Hassian matrix $H(X,\sigma)$ at scale σ in X can be calculated as,

$$H(X,\sigma) = \begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \qquad (12)$$

Where, $L_{xx}(x,\sigma)$ is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial \chi^2} g(\sigma)$ with image I in point X, and similarly for $L_{xy}(X,\sigma)$ and $L_{yy}(X,\sigma)$. Speeded up Robust Feature can also be shown diagrammatically in form of flow chart which is given as below,
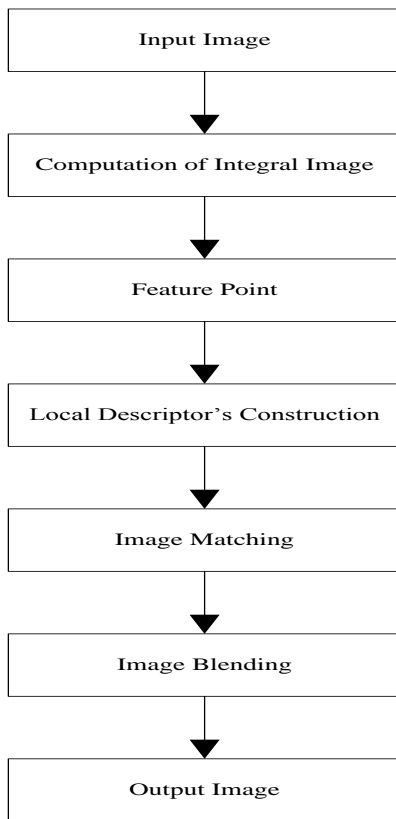

Figure 12. Flow chart of SURF

SURF algorithm has great advantages in the feature point extraction, main direction identification as well as feature vector alignment [10].

(iii) Harris Corner Detection: The Harris Corner Detection is a point feature extracting algorithm. This algorithm was provided by Chris Harris and MJ Stephens since 1988. The main reason for developing Harris Corner detector was to build up a local detecting corner feature in image. As this algorithm demands for high computation, apart from this it is broadly adopted. The algorithm for Harris Corner Detector is as given below.

1. Compute x and y derivaties of image
$$I_x = G_\sigma^x * I \;,\; I_y = G_\sigma^y * I \qquad (12)$$
2. Compute products of derivatives at every pixel
$$I_{x2} = I_x.I_X, I_{y2} = I_y.I_y \;, I_{xy} = I_x.I_y \qquad (13)$$
3. Compute the sums of the products of derivatives at each pixel.
$$S_{x2} = G_{\sigma 1} * I_{x2}, S_{y2} = G_{\sigma 1}.I_{y2} \;, S_{xy} = G_{\sigma 1}.I_{xy} \qquad (14)$$
4. Define at each pixel (x, y) the matrix,
$$H(x,y) = \begin{bmatrix} S_{x2}(x,y) & S_{xy}(x,y) \\ S_{xy}(x,y) & S_{y2}(x,y) \end{bmatrix} \qquad (15)$$
5. Compute the response of detector at each pixel
$$R = Det(H) - K\big(Trace(H)\big)^2 \qquad (16)$$
6. Threshold on value of R. compute non max suppression.

The Harris Corner Detection technique is proposed to extract the corners which need not to set the threshold by manual and is insensitive to isolated point, noise and edge [11].

V. CONCLUSION

In the field of computer vision image stitching is painstaking to be the best research area. It deals with distinct feature extraction algorithms. We have analyzed various algorithms with their merits and demerits as Scale Invariant Feature Transform (SIFT) is robust but it is not good for real time applications and having low calculation speed, Speeded up Robust Feature (SURF) has mediocre accuracy but it is slowest than any other algorithms and Harris Corner Detector algorithm is having the poor accuracy but it provides the good computational cost. Hence in future some other measures are to be taken into consideration for removing their demerits as well and also the video stitching is to be done to provide the dynamic panorama.

REFERENCES

[1] Taherim S. Shaikh, Archana B. Patankar, "Multiple Feature Extraction Techniques in Image Stitching", *International Journal of Computer Applications (0975 – 8887)*Volume 123 – No.15, August 2015.

[2] Mittali and Jyoti Rani, "Detailed Survey on Various Image Stitching Techniques", *International Journal of Computer & IT,* ISSN No. : 2320-8074.

[3] Pranoti Kale and K.R. Singh, "A Technical Analysis of Image Stitching Algoritm", *International Journal of Computer Science and Information Technologies*, Vol. 6 (1), 2015, 284-288, ISSN: 0975-9646.

[4] Ebtsam Ade, Mohammed Elmogy and Hazem Elbakry, "Image Stitching based on Feature Extraction Techniques: A Survey", *International Journal of Computer Applications* (0975-8887) Volume 99– No.6, August 2014.

[5] Xianyong Fang, "Feature Based Stitching of a Clear/Blurred Image Pair", *International Conference on Multimedia and Signal Processing,*2011, 978-0-7695-4356-7/11

[6] Russol Abdelfatah and Dr. Haitham Omer, "Automatic Seamless of Image Stitching", *Computer Engineering and Intelligent Systems,*ISSN 2222-1719,Vol.4, No.11, 2013.

[7] Meiqun Jiang and Jingxin Hong, "A SIFT based Method for Image Mosaic", *3rd International Conference on Advanced Computer Theory and Engineering (ICACTE),2010.*

[8] Yang zhan-long Guo bao-long, "IMAGE MOSAIC BASED ON SIFT", *International Conference on Intelligent Information Hiding and Multimedia Signal Processing,2008* 978-0-7695-3278-3/08.

[9] Jingxin Hong*, Wu Lin, Hao Zhang and Lin Li, "Image mosaic based on SURF feature matching", *The 1st International Conference on Information Science and Engineering (ICISE2009)"*978-0-7695-3887-7/09.

[10] Deepak Kumar Jain, Gaurav Saxena and Vineet Kumar Singh, "Image mosaicing using corner techniques", International Conference on Communication Systems and Network Technologies, 2012. 978-0-7695-4692-6/12.

AUTHORS PROFILE

Shivangi Pandey received her B.Tech degree in Electronics & Communication Engineering from Uttar Pradesh Technical University, Lucknow in 2014 and currently pursuing M.Tech in Digital Communication Engineering from NITTTR, Bhopal. Her area of interest includes Digital Image Processing, Digital Signal Processing and Digital Communication.

Dr.Anjali Potnis is Professor at Department of Electrical & Electronics Engineering, National Institute of Technical Teachers' Training & Research Bhopal. She has got a total of 16 years of teaching experience. She has published many research papers. Her area of interest includes Digital Image Processing and Digital Signal Processing.

Madhuram Mishra received his B.E degree in Electronics & Communication Engineering from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal and currently pursuing M.Tech in Digital Communication Engineering from NITTTR, Bhopal. His area of interest includes Digital Image Processing and Digital Communication.