International Journal of Engineering and Applied Sciences

www.eaas-journal.org

# An EFFICIENT ALGORITHM FOR IMAGE STITCHING BASED ON SCALE-INVARIANT FEATURE TRANSFORM

**V S Veena Devi[1], S.M. Prasad[1] and H. Gangadhara Bhat[2]**

[1]Department of Electronics and Communication (E&C), St. Joseph Engineering College, INDIA
[2]Department of Marine Geology, Mangalore University, INDIA
vvvadithya@yahoo.com

**Abstract:**

**The** panoramic image stitching is used in many applications. In this article, we have investigated the Key points present in the images by using SIFT algorithm. The extracted features are termed as Key points, the SIFT algorithm takes an image and transforms it into a collection of local feature vectors. Each of these feature vectors is distinctive and invariant to any scaling, rotation or translation of the image. In the original implementation, these features were used to find distinctive objects in different images and the transformation was extended to match faces in images.This article describes an approach to utilize these extracted features for image stitching. We have formulated stitching as a multi-image matching problem, and have used invariant local features to find matches for all of the images. There are several algorithms available for panorama creation and we have mainly focused to extract the Key points present in the image by using SIFT algorithm which can be further extended to generate the panorama image. The developed application for panorama image creation is simulated in **MATLAB.**

Keywords: Electronics and communication, panoramic image stitching, SIFT algorithm, MATLAB

## I. INTRODUCTION

The Virtual Reality (VR) technology has several applications. Because of the limitation of geometry-based rendering, Image-based Rendering (IBR) is becoming more and more important. The panoramic image stitching is a technique to merge a sequence of images with limited overlapping area into one blended picture.

The automatic construction of image stitching is an active research area in VR which is an emerging area in the field of simulation technology. The satisfying image stitching results are crucial in the construction of virtual environment, which means a natural transition from one image to another, both in structure and intensity within and possibly beyond the overlapping region.

The method for automatic matching fall broadly into two categories: direct and feature-based. The direct methods attempt to iteratively estimate the camera parameters by minimizing an error function based on the intensity difference in the area of overlap. The direct methods have the advantage since it uses all of the available data and hence can provide very accurate registration, but they depend on the illumination.

On the other hand, feature-based methods begin by establishing correspondences between points, lines or other geometrical entities. Recently, there has been great progress in the use of invariant features for object recognition and matching. These features can be matched more reliably than traditional methods. Since the feature-based methods have higher accuracy, some image stitching methods have been introduced and implemented. In this article, some satisfying stitching results have been obtained using the SIFT features-based method.

## II. SIFT ALGORITHM DETAILS

It is a key to match features in the images stitching. So the result of image stitching will be good if the interest points are found correctly. Among the local descriptors compared, SIFT features generally perform the best [1] [4] [7-11]. Because of the unreliability of many algorithm such as the algorithm based on area and based on feature pattern [2], the algorithm based on SIFT features is of importance. The algorithm of image stitching based on SIFT features uses the invariable local features to select interest points and then calculate the homograph applying these matching points. The images with the same viewpoint, but in different directions can be related by Homograph. The details of the algorithm are:

(1) select a referenced image.
(2) Find the feature to be matched with the neighbouring images.

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

(3) Calculate the homography H of the two images.
(4) Apply H to warp and project the image 2 (in the same coordinate system) as the image, and process the image 2 and stitch them seamlessly.

### *1.1 Feature matching of images*

Because of collections of large images, where the geometric relationship between them is unknown, it is difficult to take into account the salient features of the images. In our analysis, we have used SIFT features and have obtained very good results [4]. The SIFT provides a set of features of an object that are not affected by the complications present in other methods, such as object scaling and rotation. To aid the extraction of these features the SIFT algorithm is applied in a four stage filtering approach.

1.1.1 Scale-Space extremum detection

This stage of the filtering attempts to identify those locations and scales which can be identified from different view of the same object. This can be efficiently achieved using a "scale space" function which under certain assumptions can be approximated to a Gaussian function. The scale space is defined by the function:

$$L(x,y,\sigma)=G(x,y,\sigma)*I(x,y) \quad (1)$$

Where * is the convolution operator, G(x, y, $\sigma$) is a variable-scale Gaussian and I(x,y) is the input image. Various techniques can be employed to detect stable key point locations in the scale-space. The difference of Gaussians is one such technique to locate scale-space extremum, (D(x, y, $\sigma$)) by computing the difference between two images, one with scale k times the other. The D(x, y, $\sigma$) is given by:

$$D(x,y,\sigma)=L(x,y,k\sigma)-L(x,y,\sigma) \quad (2)$$

To detect the local maxima and minima of D(x, y, $\sigma$) each point is compared with its 8 neighbors at the same scale, and its 9 neighbors up and down the scale. If the obtained value is a minimum or maximum for all these points the point is an extremum.

1.1.2 Key point Localization

This stage attempts to eliminate more points from the list of key points by finding those that have low contrast or are poorly localized on an edge. This is achieved by calculating the Laplacian, value for each key point found in stage 1. The location of extremum of z is given by:

$$z=-\frac{\partial^2 D^{-1}}{\partial x^2}\frac{\partial D}{\partial x} \quad (3)$$

If the function value at z is below the threshold value then the point has to be excluded.

1.1.3 Orientation Assignment

This stage aims to assign a consistent orientation to the key points based on local image properties. The key point descriptor can be represented relative to the orientation, which is achieved by the invariance under rotation. The approach is to find an orientation using the keypoint scale to select the Gaussian smoothed image L, and compute the magnitude and the orientation. An orientation histogram is formed from gradient orientations of sample points. Finally, using the histogram, orientation to the key points can be assigned.

1.1.3 Key point Descriptor

The local gradient data generated is used to create key point descriptors. The gradient orientations of sample points is rotated to match the orientation of the key point which is then weighted by a Gaussian. The above data is used to create a set of histograms over a window centered on the key point. The point descriptor typically uses a set of 16 histograms, aligned in a 4x4 grid, each with 8 orientation bins.

The obtained vectors are known as SIFT keys and are used in a nearest-neighbours approach to identify possible objects in an image. The collections of keys which corresponds to the model are identified. when three or more keys agree on the model parameters this model is evident in the image with high probability. Due to the large number of SIFT keys in an image of an object, substantial levels of occlusion are possible while the image is still recognized by this technique.

In our work, the SIFT features are extracted from a pair of adjacent images. The features are matched using *k-d* tree and the results are shown in **Figs.1-3.**

## II Homography calculating

As the above described, given two images taken from the same viewpoint but in different directions, the relationship of the two images can be described by a planar projective motion model which is Homography metrics denoted H. This projective transformation warps an image into another one using xi'~H . The calculation of H correctly is crucial in the images matching. In our analysis, we have used direct linearity transformation (DLT) method [3] [5].

The *H* is a 3 x 3 matrix and hence has 9 elements in total, and we need at least four pair of matching points to calculate

February 2013. Vol. 2, No.2 ISSN 2305-8269

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

H. Assuming that given four pairs of interest points from image2 to image1 given by the transformation, xi$\leftrightarrow$xi' where xi is from image 2 and xi' from image 1. The transformation is xi '~ H xi where ~ denotes equality up to scale, there exist a nonzero gene between them. Hence, the relationship is xi'* H xi = 0. The H can be calculated. Denoted the j row of H as $h^{jT}$ , we have,

$$H X_i = \begin{pmatrix} h^{1T} X_i \\ h^{2T} X_i \\ h^{3T} X_i \end{pmatrix}$$

Denoting,

$$X_i' = (x_i', y_i', w_i')^T$$

$$X_i' \times H X_i = \begin{pmatrix} y_i' h^{3T} X_i - w_i' h^{2T} X_i \\ w_i' h^{1T} X_i - x_i' h^{3T} X_i \\ x_i' h^{2T} X_i - y_i' h^{1T} X_i \end{pmatrix} \qquad (4)$$

Since j = 1,2,3, $h^{jT} X_i = X_i^T h^j$ , the 9 elements of H are given by:

$$\begin{bmatrix} 0^T & -w_i' X_i^T & y_i' X_i^T \\ w_i' X_i^T & 0^T & -x_i' X_i^T \\ -y_i' X_i^T & x_i' X_i^T & 0^T \end{bmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = 0 \qquad (5)$$

Where

$$H = \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}$$

It should be noted that:
(1)Ai h =0 is a linear equation in h. The elements of matrix Ai are the quadratic polynomial of given points.

(2) Although there are three equations in the equation above, only two of the equations are linearly independent hence when calculating H, the equation can be simplified as shown below:

$$\begin{bmatrix} 0^T & -w_i' X_i^T & y_i' X_i^T \\ w_i' X_i^T & 0^T & -x_i' X_i^T \end{bmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = 0 \qquad (6)$$

(3) the above equation is also true for every homogeneous coordinate, so we can use WI' =1, and the (xi, yi') is the coordinate obtained from the image.

## III Experimental Results of Image stitching

After calculating H, we transform the image 2 into the same coordinate system as image 1 according to the equations introduced in section 2.2 [5]. Due to transformation, lots of information about the image can be lost.

In our work, after the transformation the coordinates of the image 2 are recorded and the two images are stitched directly. Using the above method, complex algorithms and unnecessary calculation can be avoided and higher graphics resolutions can be obtained. The illumination of the two original images is different, but the stitching result is not affected because the SIFT features are insensitive to illumination, oriental and scale. The results are shown in **Fig.4 and Fig.5**.

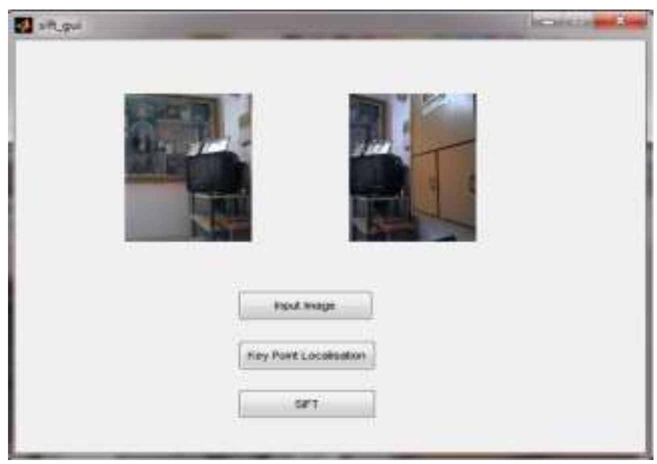

***Figure1***: GUI window displayed when code is executed

February 2013. Vol. 2, No.2 ISSN 2305-8269

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

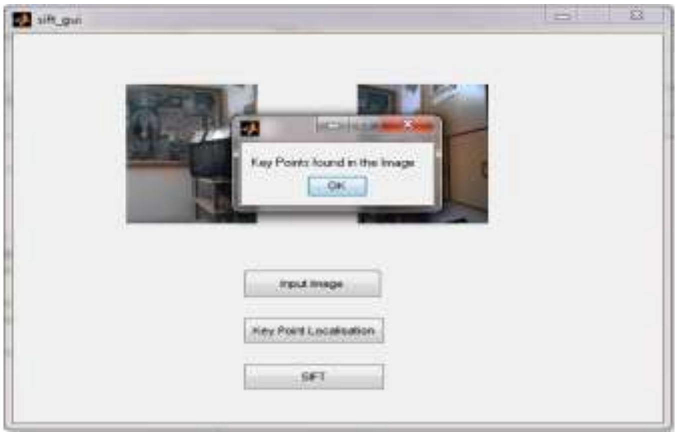


***Figure 2:*** *GUI window, dialog box displayed on it.*

***Figure 3:*** *Key points detected in the images*

***Figure 4:*** *Key points in two images are found and mapped*

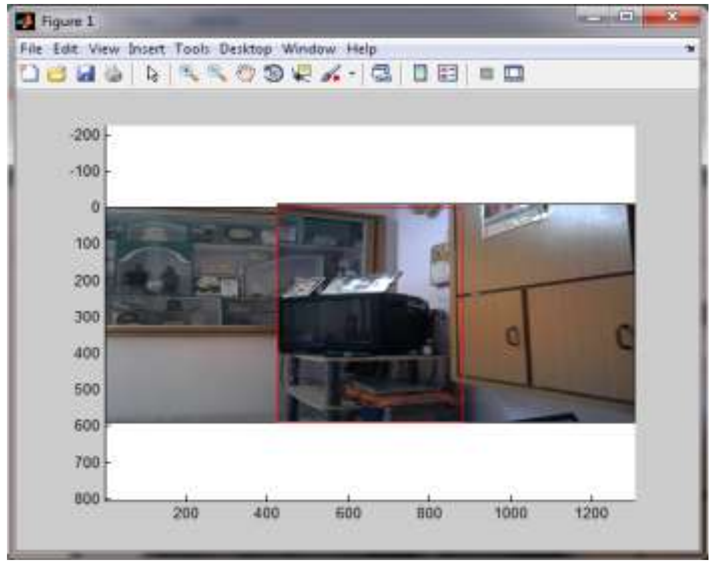


***Figure 5:*** *Panorama image created from two images*

## III. Conclusion

This work carried out in this paper, considers the detection and extraction of Key points by using SIFT algorithm. From our analysis it was found that SIFT Key point features are highly distinctive and invariant to image scaling and rotation. It provides correct matching in images when subjected to noise, viewpoint and illumination changes. The created panorama image can be used for many applications like face detection, Digital sign boards in Shopping Mall, Railway Station and so on.

From our work, we conclude that the SIFT features to transform images allows to stitch them in an accurate way. The SIFT feature ensures smooth transformation between Images with different illumination and orientation and it can also overcome the difficulty of matching in vertical direction. High accuracy and better effect are obtained from the method based on SIFT features in image stitching. Since we have stitched the two images by using the coordinates of the image after transformation instead of interpolation, unnecessary calculation has been avoided. Since the carried out work has tedious calculations the speed of stitching images is a bit slow. The aim of the future work is to find an efficient way to simplify the algorithm.

**References:**

[1]H.Y. Shum and R. Szelishi: "Panoramic Image Mosaics", Technical Report, Microsoft Research (1998).

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

[2]G. Kang: "Virtual Environment Touring based on image", master thesis, Nanjing Institute of Technology (2004).

[3]R. Hartley A. Zisserman: "Multiple View Geometry in Computer Vision", AnHui University Press (2002).

[4]H.Y. Shum and R. Szelishi: "Construction of Panoramic Image Mosaics with Global and local Alignment", International Journal of Computer Vision, 36(2) 101-130 (2000).

[5]Hui Zhao: "Matching Technology based on SIFT Features", Shan Dong University, Institute of Information Science and Engineering, (2006).

[6] D. Lowe: "Distinctive image features from scale-invariant key points", Int. Journal of Computer Vision, 60(2):91–110, (2004).

[7]Hsieh, Jun-Wei , "Fast stitching algorithm for moving object detection and mosaic construction", *Image and Vision Computing*, 22(4), 291-306 (2004).

[8]S. Mann and R. Picard, "Video orbits of the projective group: A simple approach to featureless estimation of parameters." *IEEE Trans. Image Processing*, 6,1281-1295, 1997.

[9]W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery: "Numerical Recipes in C: the Art of Scientific Computing", *Cambridge University Press*, Cambridge, 1993.

[10]Brown, L.G: "A survey of image registration techniques", *ACM Comput. Surv.* 24 (4), 325-376 (1992).

[11]Reddy, B.S and Chatterji, B.N: "An FFT-based technique for translation, rotation, and scale-invariant image registration". *IEEE Trans. Image Process.* 5(8),1266-1271 (1996).