

A Combined Approach of Harris-SIFT Feature Detection for Image Mosaicing

Monika B. Varma, Prof. Kinjal Mistree
C.E Department,
Chotubhai Gopalbhai Patel Institute of Technology
Uka Tarsadia University
Gujarat, India.

Abstract—This Image Mosaicing is a process of assembling multiple overlapping images of the same scene into a large image. The output of the image Mosaicing operation is the union of the two or more overlapping input images. This technology is widely used in photography, digital video, motion analysis, medical image processing, remote sensing image processing, document image processing and other fields.

This paper describes a combined approach of Harris-SIFT feature detection for Image Mosaicing. Firstly, feature points are detected by using Harris corner detector, then after SIFT descriptor is computed to store feature vector for each detected keypoints and then feature matching is applied. The RANSAC Homography algorithm is used to detect wrong matches for improving the stability of the algorithm and for estimating the transformation model. Finally Image Blending procedure is used for smooth and seamless panorama.

Combined approach of Harris corner detector and stable invariant SIFT descriptor is reliable in extracting the translation, scale and rotation invariant features, less affected by illumination issues and increases the speed of the feature detection process which also affects the matching accuracy.

Keywords— Feature Detector, SIFT descriptor, RANSAC

I. INTRODUCTION

Image mosaic technology is one of the gradual developing digital image processing technologies applied to various fields. The field of view of human eye is restricted to $200 \times 135^\circ$, in order to increase the field of view $360 \times 180^\circ$ panorama is used. Image Mosaicing or Image Stitching is a technique to create panorama, where images are assembled with some common field of view. In many fields there is always a need of full view images, the most traditional application is the construction of large aerial and satellite photographs from collections of images, document image analysis, medical imaging[5], field of photography, 3D rebuilding of objects etc.

An Image Mosaicing is a synthetic composition generated from a sequence of images and it can be obtained by understanding the geometric relation between the images. The geometric relations are the coordinate transformations that relate the different image coordinate systems. By applying the

appropriate transformations via a warping operation and merging the overlapping regions of warped Images, it is possible to construct a single image indistinguishable from a single large image of the same object, covering the entire visible area of the scene. The basis of the Image Mosaicing technique is to find the common part in the overlap input images and then smoothly blend them to produce final Panorama.

The procedure of the Image Mosaicing Consists of two basic Concept: 1) Image Registration and 2) Image Blending and Warping[6], as shown in the Fig 1.

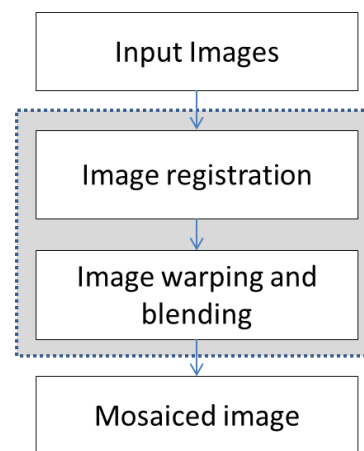


Fig.1 Basic Block Diagram of Image Mosaicing

A. Image Acquisition

First, In this stage overlapping image of the same scene is captured from any mobile, digital cameras to create a full mosaiced image. The input images can be of any type ex: image outdoor scene images, aerial images, satellite images, medical images, document images etc.

B. Image Registration

The core of a Mosaicing procedure is the Image Registration of the overlapping parts of the images we want to merge together. Image Registration refers to the geometric alignment of a set of images. The set may consist of two or more digital images taken of a single scene at different times, from different sensors, or from different viewpoints. This

means, the best transformation is computed between the input two images and then make the geometric alignment of those images. One of the images is referred to as the references or source and the second image is referred to as the target or sensed.

The majority of the registration methods consist of the following four steps: Feature Detection, Feature Matching, Transform model estimation, Image Resampling and transformation[1].

The feature detection is the first step, if the feature are detected accurately and speedy then further process accurate in accurate way. The Feature can be salient and distinctive objects (closed-boundary regions, edges, contours, line intersections, corners, etc.) are manually or, preferably, automatically detected. There are the two main approaches for Feature detection: Area based methods and Feature based methods. Second approach of feature detection is based on the extraction of the salient structures features like edges, corners, region etc. Corners are important local features in the images. They are the points having high curvature and lie in the junction of the different brightness region of the images[8]. They are the distinct and have a higher chance of being detected in different views of the same scene.

The corner based feature detector are Harris, SUSAN etc. which are descriptor less. Harris corner detector[3] finds the feature points at fixed scale[8]. This detectors takes the less time. Here, feature detectors are used to find areas of interest in the images. A feature descriptor is used to build a description of a local feature. Disadvantage of Harris corner detector is that it is not scale invariant and though it is rotational invariant it is not that much reliable[9]. There are some scale and rotation invariant feature detector or descriptors for ex: SIFT, SURF, BRIEF etc. whose descriptor can be used for further feature matching.

After feature has been detected from the both reference and sensed images, they can be matched by means of the image intensity values in their close neighborhoods, feature spatial distribution and other symbolic feature description. Again feature matching is divided into further two categories: 1) Area-based and 2) feature based[1].

After the feature correspondence has been established the mapping function is constructed. It should transform the sensed image to overlay it over the reference image. Homography is the mapping between two spaces which often used to represent the correspondence between two images of the same scene . It's widely useful for images where multiple images are taken from a rotating camera having a fixed camera center ultimately warped together to produce a panoramic view. In homography undesired corners which do not belong to the overlapping are removed[12]. Homography calculation can be performed by Least Mean Square algorithm, RANSAC etc. RANSAC (Random Sample Consensus)[7] is a common and robust accepted way to refine the homography between images because RANSAC can return the final inliers when getting the final homography[11]. Least Squares method[6] is used for homography calculations, but it cannot return final inliers.

The mapping functions constructed during the previous step are used to transform the sensed image and thus to register the images.

C. Image Blending

The next step, following registration, is image integration. It is a process of overlaying images together on a bigger canvas. The images are placed appropriately on the bigger canvas using registration transformation to get final mosaic. Images aligned , even after undergoing geometric corrections, require further processing to eliminate distortions and discontinuities. Since the parts are recorded under different conditions, including weather, lighting and noise they may have different gray level characteristics. This may cause seams to be apparent between two different parts. The seams can be very noticeable, and they often interfere with the perception of the details of the picture. Blending is applied to make seamless stitching. Two popular methods of blending the images are, One is alpha blending or feathering also called linear blending and another is Gaussian pyramid.

II. METHODOLOGY

In the process of Image Mosaicing, a variety of different situations may be encountered, such as the translation of the image, image rotation, and different scales image mosaic. Therefore, the requirement to the applicability of algorithm is also increasing.

The most two popular algorithms used for the feature detection are Harris corer detector and SIFT, both are having advantages and disadvantages.

Harris corner detector is not Scale invariant and though it is rotation invariant it is not that much reliable. SIFT which is a stable method to detect feature points but the computation in the procedure of image pyramid construction, key point location determination by extreme value detection and the computation cost is too much time consuming. Therefore instead of applying them individually, it is profitable to use them together to take the advantage of both and to overcome the disadvantage of each other. hence computing time can be shorter.

SIFT descriptor operation is performed only to those features which are extracted by Harris algorithm; hence the computing cost will be reduced. Thus here by applying this approach, i.e. by using Harris corner detector and SIFT descriptor in combined manner we get the scale and rotation invariant features, less affected by illumination issues. The idea of applying Harris algorithm before SIFT descriptor reduces the computational time of SIFT feature detector algorithm for detecting the features in the image and increases the feature matching capability of features detected by Harris detector and the image mosaic can be accelerated. Below fig 2 shows over all working of proposed approach and each of the blocks are explained in details below.

A. Harris Corner Detector

Firstly the two images are taken with some overlapping regions so the information for stitching both images can be taken from that overlapping region.

Corners in images represent a lot of important information. Corners are important local features in images. In a variety of image features, corners are not affected by illumination and having the property of rotational invariance[8]. For the feature detection Harris corner detector is applied. It is based on local auto-correlation function of a signal where local auto-correlation function measures the local changes of the signal with patches shifted by a small amount in different directions and having the fast computation speed[4,6].

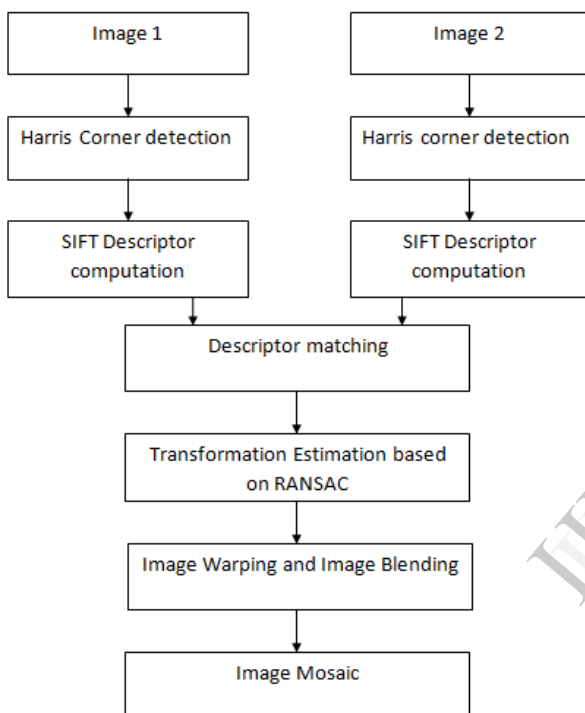


Fig.2 Block Diagram of Proposed approach

For this reason, Harris algorithm is used for feature detection in proposed work. The steps to implement Harris corner detection are:

Input : Gray scale image $I(x,y)$

Output : Interest (corner) points

Step: 1 Apply image smoothing using averaging filter

Step: 2 Compute magnitude of the x and y gradients at each pixel

Step: 3 Construct C in a window around each pixel, where C is autocorrelation matrix. I_x and I_y are derivative of image I.

$$C = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad (1)$$

Step: 4 calculate cornerness(R) measure using below eq. 2

$$R = Det(C) - k * (TC)^2 \quad (2)$$

Det is determinant of the matrix C. T is the trace of the matrix. R is for the value of the corresponding pixel of interest $-\lambda_1; \lambda_2$ is proportional to the principle curvatures of partial autocorrelation function. Therefore, by judging the value of λ_1 and λ_2 to determine the slow changes of areas, corners and edge. The changes are the three cases:

1. If $\lambda_1 \approx 0$ and $\lambda_2 \approx 0$, then this pixel (x,y) has no feature interest.
2. If $\lambda_1 \approx 0$ and λ_2 has some large positive value, then an edge is found.
3. If λ_1 and λ_2 have large positive values, then corner is found [15].

B. Feature Descriptor

In feature descriptor, SIFT descriptor is used to represent feature extracted from Harris corner detection. SIFT Descriptor is the part of the SIFT algorithm. Descriptor calculates orientation and direction around each keypoint and stores them in a feature vector. This part of SIFT detector makes SIFT more stable, rotation invariant and scale invariant.

The feature descriptors detected by SIFT claim to be capable of distinguishing each and every image in the dataset from one another with the cost involved in its operations. Computation of orientation is to achieve scale and rotation invariance. The Descriptor estimates a dominant orientation at each detected keypoint. Once the local orientation and scale of a keypoint have been estimated, a scaled and oriented patch around the detected point can be extracted and used to form a feature descriptor[13].

After the features are detected from the Harris detector the SIFT descriptor is implemented to generate the feature vector for each corner point which are extracted from the Harris corner detector. The descriptor is useful for feature matching from both overlapping images.

SIFT feature description mainly includes two steps:

1. Determine direction parameter and Orientation assignment of feature points

The idea is to collect gradient directions and magnitudes around each keypoint. The figures 3 show the most prominent orientation(s) in that region and then assign this orientation(s) to the keypoint. Any later calculations are done relative to this orientation. This ensures rotation invariance. The size of the "orientation collection region" around the keypoint depends on its scale. The bigger the scale, the bigger collection region is. Gradient magnitudes and orientations are calculated using these formulae 3 and 4[2]:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (3)$$

$$\theta(x,y) = \tan^{-1} \left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)} \right) \quad (4)$$

The magnitude and orientation is calculated for all pixels around the keypoint. Then, a histogram is created for this.

2. Use graphic information around the feature point to construct 128 – dimensional feature vector.

This step insures invariance to image location, scale and rotation. Now we create a fingerprint for each keypoint. This is to identify a keypoint. We want to generate a very unique fingerprint for the keypoint. It should be easy to calculate. We also want it to be relatively lenient when it is being compared against other keypoints. To do this, a 16×16 window is taken around the keypoint. This 16×16 window is broken into sixteen 4×4 windows.

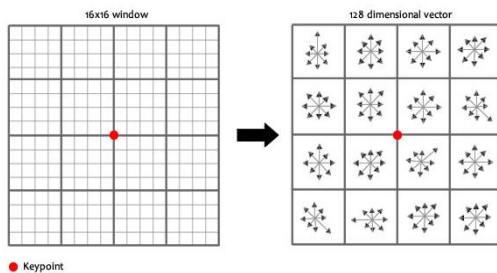


Fig.3 Key point Descriptor[14]

Within each 4×4 window, gradient magnitudes and orientations are calculated. These orientations are put into 8 bin histogram.

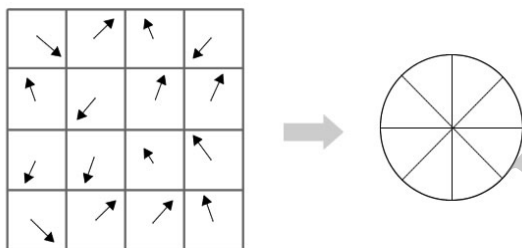


Fig.4 8 bin histogram[14]

Any gradient orientation in the range 0 - 44 degrees add to the first bin. 45-89 add to the next bin and so on. The amount of the Bin depends upon the magnitude of the gradient. For all 16 pixels, have to “compile” 16 totally random orientations into 8 predetermined bins. Do this for all sixteen 4×4 regions, 4x4x8 = 128 numbers. These 128 numbers form the “feature vector”. This keypoint is uniquely identified by this feature vector. At end of this procedure, feature vector is extracted of size 128 for the detected keypoints of the image. Matrix is generated of size 128*total no. of keypoints.

C. Feature matching

The Feature point matching is used to define the reference point in an image, and to search for a matching point in another image. After Feature extraction and descriptor generation the goal of the matching stage is to find geometrically consistent feature matches between all images. The Sum of Squared Difference (SSD) method is implemented because its gives good score to very ambiguous matches.

$$SDD(f_1, f_2) = \sum_{i=0}^{N-1} (f_{1i} - f_{2i})^2 \quad (5)$$

Where f_1 is the feature vector of the 1st image and the f_2 is the feature of the 2nd image.

For each corner, finding the 2 nearest neighbors (using SSD from above as the distance) in the second image.

If the ratio ($\frac{SSD \text{ of } 1st \text{ match}}{SSD \text{ of } 2nd \text{ match}}$) < thresh

(Thresh = 0.5)

Then the 1st match will be considered else the 2nd match of features is considered from two overlapping images. Therefore at the end of this procedure we get the matching pairs of the features from two input images.

D. RANSAC

Homography is important to calculate where multiple images are taken from a rotating camera center ultimately warped together to produce a panorama. After matching the features from the input images, we have to find the transformation of the images which are going to be stitched by using corresponding corner points, but there are some false matches to skew the transform. Not all the matches computed in the previous step of feature point matching will be correct[10].

Among all the matches found, we have to consider only the inliers means the point which fit the model and not to consider the outliers which do not fit the model. The method to remove outlier and estimating transformation is implemented by RANSAC (Random Sample Consensus).

RANSAC estimation works as a two-stage process:

1. Classify data points as outliers or inliers
2. Fit model to inliers while ignoring outliers

Steps for the Computation of the RANSAC are:

- a. Selecting four feature pairs (at random), p_i, p_i^1 .
- b. Compute homography H.
- c. Compute inliers where $SSD(p_i^1, Hp_i) < \text{thresh}$.
- d. Keeping largest set of inliers.
- e. Re-compute least-squares H estimate on all of the inliers.

E. Image Blending

Alpha Image blending technique is applied in this proposed to generate the final smooth and seamless mosaiced.

III. EXPERIMENTAL RESULTS

The combined proposed approach can work for rotation and scaled images as well as for all image type e.g. jpg, bmp, png etc. which shows the better performance.

A. Image Mosaicing for scaled image



Fig.5 Input image with scale transformation



Fig.6 Corner detection using Harris



Fig.7 Inliers (Yellow mark) By RANSAC



Fig.8 Mosaiced Image

The In Fig.5 two different images are captured, both having some overlapping region, in which left side image is scaled

image and right side image is reference image. After that corner features are detected by Harris corner detector method on both the images which is shown in the Fig.6. Then after SIFT feature descriptor is computed on the detected feature points. SSD approach is used to match this feature detector and then after RANSAC is used to filter out unwanted matches this represents in the Fig.7. The final output is shows in the Fig. 8

B. Image Mosaicing for document Image

The proposed approach works on the document images also.

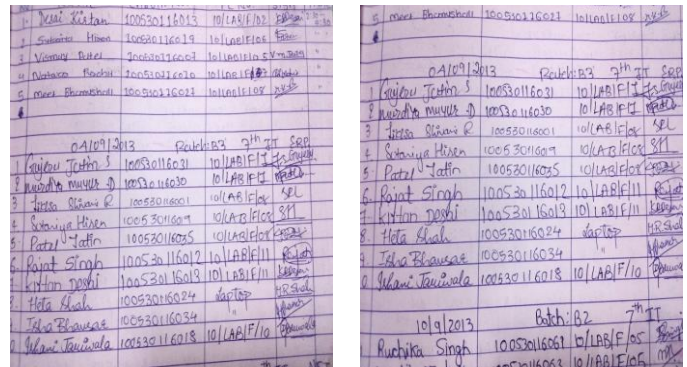


Fig.9 Input Document Images

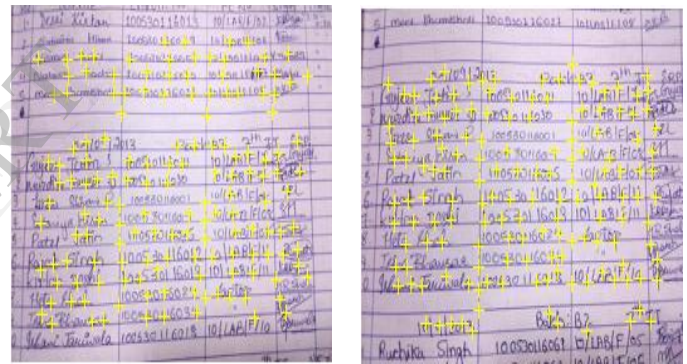


Fig.10 Corner detection using Harris

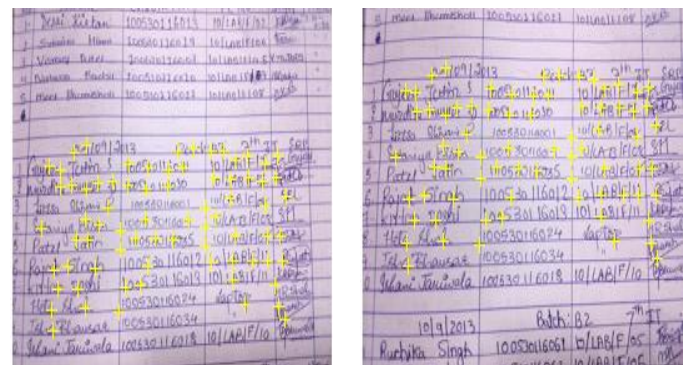


Fig.11 Inliers (Yellow mark) By RANSAC

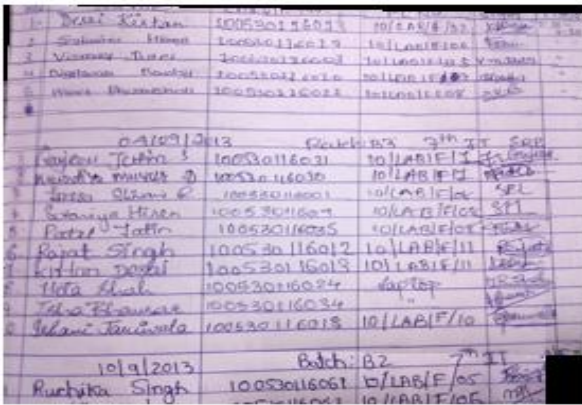


Fig.12 Mosaiced Document Image

IV. CONCLUSION

Image Mosaicing have a wide spread applications in the photography, medical image processing, document image processing etc. In feature based Image Registration, point as a feature is used for outdoor and indoor scene images, aerial images and even in the document images as the extraction of feature points is more feasible, because they remain unchanged regardless of illumination, translation, rotation, resolution and etc. SIFT and Harris are two important feature point detector algorithms. This approach describes a robust feature based image mosaic algorithm. Firstly, feature points are detected by using Harris corner detector, then after SIFT descriptor is computed to store feature vector for each detected keypoints and then feature matching is applied. The RANSAC Homography algorithm is used to detect wrong matches for improving the stability of the algorithm and for estimating the transformation model. Finally Image Blending procedure is used for smooth and seamless panorama.

Experiment results suggest that proposed method can serve as a reliable approach for image Mosaicing which having rotated or scaled input images and probably of any type.

REFERENCES

- [1] Zitova, Barbara, and Jan Flusser. "Image registration methods: a survey." *Image and vision computing* 21.11 (2003) Elsevier: 977-1000.
- [2] Lowe, David G. "Distinctive image features from scale-invariant keypoints" *International journal of computer vision* 60.2 (2004): 91-110.
- [3] Chris Harris & Mike Step hens. "A COMBINED CORNER AND EDGE DETECTOR" Plessey Research Roke Manor, United Kingdom,1988.
- [4] Bheda, Dilipsinh, Asst Prof Mahasweta Joshi, and Asst Prof Vikram Agrawal. "A Study on Features Extraction Techniques for Image Mosaicing."
- [5] Jain, Ms Parul M., and Vijaya K. Shandliya. "A Review Paper on Various Approaches for Image Mosaicing?"
- [6] Kang, Peng, and Hongbing Ma. "An automatic airborne image mosaicing method based on the SIFT feature matching." *Multimedia Technology (ICMT), 2011 International Conference on.* IEEE, 2011.
- [7] Fischler, Martin A., and Robert C. Bolles. "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography." *Communications of the ACM* 24.6 (1981): 381-395.
- [8] Chen, Jie, et al. "The Comparison and Application of Corner Detection Algorithms." *Journal of Multimedia* 4.6 (2009).
- [9] Qiu, Pengrui, Ying Liang, and Hui Rong. "Image Mosaics Algorithm Based on SIFT Feature Point Matching and Transformation Parameters Automatically Recognizing." *Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering.* Atlantis Press, 2013.
- [10] Guandong, Gao, and Jia Kebin. "A new image mosaics algorithm based on feature points matching." *Innovative Computing, Information and Control, 2007. ICICIC'07. Second International Conference on.* IEEE, 2007
- [11] Patel, Udaykumar B., and Hardik N. Mewada. "Review of Image Mosaic Based on Feature."
- [12] Joshi, Hemlata, and Mr KhomLal Sinha. "Image Mosaicing using Harris, SIFT Feature Detection Algorithm."
- [13] Szeliski, Richard. *Computer vision: algorithms and applications.* Springer, 2010.
- [14] AISHACK. Retrieved 1 April, 2013, [Online] <http://www.aishack.in/2010/05/sift-scale-invariant-feature-transform/>
- [15] Jain, Deepak Kumar, Gaurav Saxena, and Vineet Kumar Singh. "Image Mosaicing Using Corner Techniques." *Communication Systems and Network Technologies (CSNT), 2012 International Conference on.* IEEE, 2012.