ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 2, Issue 11, November 2013

Image Mosaicing using Harris, SIFT Feature Detection Algorithm

Hemlata Joshi[1]
M. Tech. Student, Department of Computer Science and Engg. CSIT, Durg (CG) INDIA
Mr.KhomLal Sinha[2]
Assistant Professor, Department of Computer Science and Engg. CSIT, Durg (CG) INDIA

***Abstract*— Image Mosaicing has wide applications in the field of 3D image reconstruction, medical imaging, military automatic target recognition, video conferencing, and data from satellites. An Image Mosaicing is a method of assembling multiple overlapping images of the same scene into a larger image. The input images are stitched together using image mosaicing algorithms i.e. Harris and SIFT (Scale Invariant Feature Transform). In this paper, a comparative study is done among the common stitching algorithms: Harris corner detection algorithm, SIFT algorithm. The principles of these algorithms are described in terms of computing time and stitching precision. Harris algorithm detects more features and is only rotationally invariant. SIFT is both rotationally and scale invariant, although it is slow.**

***Index Terms*— Mosaicing, Panorama, Registration, SIFT, Warping.**

I. Introduction

Image Mosaicing is a technique of developing a panoramic image which has a wide application in the field of 3D image construction, medical image, military automatic target recognition, data from satellites, etc. Image mosaicing is a important branch of computer vision areas which is done by using various stitching algorithms. Image mosaicing technique is divided into three types of image registration methods [1]. These image registration methods have the following classes which includes; Region-based matching: algorithms that use image pixel values directly e.g. Correlation methods [2]. Frequency-based matching: algorithms that use the frequency domains e.g. FFT-based methods [3]. Matching based on features: algorithms that use low level features such as edges and corners e.g. Feature-based methods [4]. Algorithms that use high-level features such as identified parts of image objects, relation between image features e.g. Graph-theoretic methods [4]. These image mosaicing algorithms have a good robustness when they are viewed in terms of feature point extraction e.g. corner points or critical points, some invariant moments and contours. Such stitching algorithms have well distinctiveness and repeatability.

II. Harris corner detection

Harris corner detection algorithm [5] was proposed by Harris C and Stephens MJ in the year 1988. Harris algorithm is a point feature extracting algorithm based on Moravec algorithm. It is an algorithm based on still image used for combined corner and edge detector. Reasonable amount of corner features are extracted which gives a better quantitative measurement by using a stable operator. A local detecting window in image is designed. The average variation in intensity is determined by shifting the window by a small amount in different direction. The centre point of the window is extracted as corner point. The point can be recognized easily by looking at the intensity values within a small window. Shifting the window in any direction gives a large change in appearance. Harris corner detector is used for detecting corners. On shifting the window if it's a flat region than it will show no change of intensity in all direction. If an edge region is found than it will show no change of intensity along the edge direction. But if a corner is found than there will be a significant change of intensity in all directions. Harris corner detector gives a mathematical approach for determining the region is flat, edge or corner. Harris corner technique detects more features and it is rotational invariant and scale variant.

Harris gave a formula that can be calculated out the change of the pixel values in any direction, rather than fixed in the 8 directions.

$$M = G(\tilde{s}) \otimes \begin{pmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{pmatrix} \tag{1}$$

$$R = \Delta(M) - k \bullet T^2(M) \tag{2}$$

gx, gy represent the first order gray gradient of horizontal and vertical $G(\tilde{s})$ is for Gaussian model. T is for Matrix trace. R is for the value of the corresponding pixel of interest – λ_1; λ_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: When the two curvatures are small (Fig 1) the local autocorrelation

All Rights Reserved © 2013 IJSETR

function is very flat, changes are not dramatic; A small curvature, and the other is large, this means partial autocorrelation function changes little in one direction, changes in the vertical direction is large means it's the edge; when the two curvatures are large, indicating the local autocorrelation function has a peak, it's the corner.

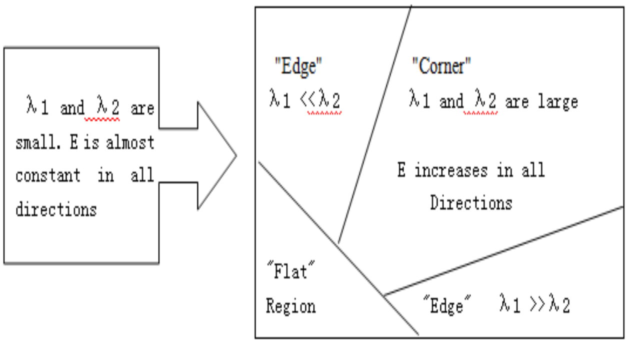


Fig1. Harris operator

Harris corner algorithm fully utilizes the characteristics of the corner; search speed has improved greatly than block matching. But the calculation is still larger. The total amount of calculation (1):

$O(n^2) + O(2n) + O(b^2n) + O(n^2) = O(2n^2 + 2n + b^2n)$

For the change of intensity for the shift [u, v]:

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

Where w(x, y) is a window function, I(x + u, y + v) is the shifted intensity and I(x, y) is the intensity of the individual pixel. Harris corner algorithm is given below as:

(1) For each pixel (x, y) in the image calculate the autocorrelation matrix M as:

$$M = \sum_{x,y} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \qquad (4)$$

(2) For each pixel of image has Gaussian filtering, get new matrix M, and discrete two dimensional zero-mean Gaussian function as :-

$$\text{Gauss} = \exp(-u^2 + v^2)/2\delta^2 \qquad (5)$$

(3) Calculating the corners measure for each pixel (x, y) we get;

$$R = \{I_x^2 x I_y^2 - (I_x^2 I_y^2)\} - k\,\{I_x^2 + I_y^2\}^2 \qquad (6)$$

(4) Choose the local maximum point. Harris method considers that the feature points are the pixel value which corresponding with the local maximum interest point.

(5) Set the threshold T, detect the corners points.

III. SIFT ALGORITHM

The Scale Invariant Feature Transform (SIFT) algorithm was proposed by Lowe in the year 1999. SIFT [6] is a feature detection algorithm which detects feature in an image that identifies similar objects in other images. It produces key point descriptors which are the image features. For a set of input frames SIFT extracts features. Image matching is done using Best Bin First (BBF) algorithm for estimating initial matching points between input frames. To remove the undesired corners which do not belong to the overlapped area, RANSAC algorithm is used. It removes the false matches in the image pairs. Reprojection of frames is done by defining its size, length, width. Stitching is done finally to obtain a final output mosaic image. In stitching, each pixel in every frame of the scene is checked whether it belongs to the warped second frame. If so then the pixel is assigned the value of the corresponding pixel from the first frame. SIFT algorithm is both rotational invariant and scale invariant. SIFT is very popular for object detection in images with high resolution. It is a robust algorithm for image comparisons though it is slow. The running time of a SIFT algorithm is large as it takes more time to compare two images. SIFT has four computational phases which includes:

(1) Scale-space extrema detection [7]: - This is the first phase which identifies the potential interest points. It searches over all scales and image locations by using a difference-of-Gaussian function. Here, the middle point is compared with its neighborhood points to detect utmost points.

(2) Key-point localization: - For all the interest points so found in phase one, location and scale is determined. Key-points are selected based on their stability. A stable key-point should be resistant to image distortion. This is done by using Taylor expansion, the extreme points and location are carefully determined by using the following equation:

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \qquad (7)$$

(3) Orientation Assignment: - SIFT algorithm computes the direction of gradients around the stable key-points. One or more orientation are assigned to each key-point based on local image gradient directions. By the help of key-point neighborhoods, the gradient m(x, y) and the direction ϴ (x, y) are estimated for an image L(x, y). The gradient and direction can be formulated as:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \qquad (8)$$

$$\theta(x,y) = \arctan\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \qquad (9)$$

Taking the gradient value and characteristic into consideration, each sample points is added to the histogram. The direction for the feature points are estimated from the maximum peak values from the histogram.

(4) Feature vectors are generated [8] which is shown in fig.2. The arrow in each cell stands for gradient

All Rights Reserved © 2013 IJSETR

direction along with the amplitude of pixels. The seed point can be formed by aligning the unidirectional gradients followed by the normalization.

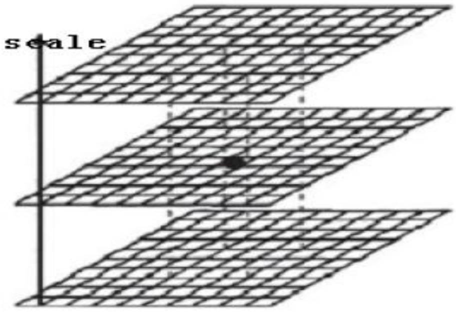


Fig2. Local extremum in DOG scale-space

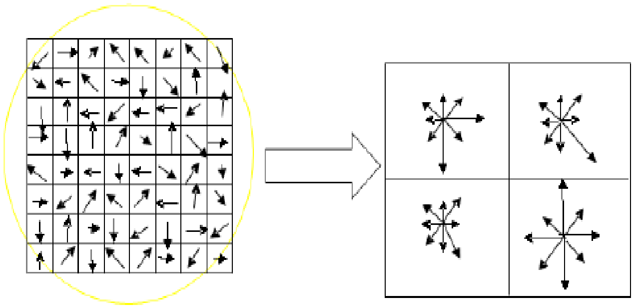


Fig3 Feature vector generation

wavelet response in the vertical direction. Each sub-regions responses are summed-up along with the absolute value of response and each sub-region vector will form the four-dimensionality: For each sub feature point, format the 4×(4×4)=64 dimensional description vector, then normalize the vector.

IV. Performance Evaluation

There are various quality metrics on the basis of which the image quality can be evaluated. Subjective quality metric depends on the human observer whereas the objective metric depends on the computation. Various quality metrics that have been determined to evaluate the image quality are PSNR, FSIM, Mutual Information, Enhancement performance measure (EME), Normalized Absolute Error.

1. PSNR: - It is defined as the peak signal to noise ratio in decibels. PSNR is used to measure the quality of reconstruction. It is calculated between references to processed image. If R is the measure of the input image data type, then the PSNR is given by :-

$$PSNR = 10\log_{10}\left(\frac{R^2}{MSE}\right) \tag{12}$$

Thus the PSNR value tends to infinity as two MSE tends to zero, which means that the larger PSNR value corresponds to higher image quality.

2. FSIM :- Feature similarity index is based on the concept of phase congruency and the gradient magnitude which is extracted as a low level image feature for the calculation of the local similarity indesx.
   For Combined similarity

$$S_L(x) = [S_{PC}(x)][S_G(x)] \tag{13}$$

and $$PC_M(x) = \max(PC_1(x), PC_2(x)) \tag{14}$$

$$FSIM = \frac{\sum_{x\in\Omega} S_L(x).PC_M(x)}{\sum_{x\in\Omega}.PC_M(x)} \tag{15}$$

Where, Ω means the whole image spatial domain.

3. Mutual Information :- Given two images *M(i, j)* and *N(i, j).* Mutual information is defined as amount of uncertainity that decreases about B when A is given. It measures asymmetry between two images as well as fluctuation from its mean value. MI for two images *M(i, j)* and *N(i, j)* can be expressed as;

$$\text{MI} = \text{H (M)} + \text{H (N)} - \text{H (M, N)} \tag{16}$$

Where H (M) is the entropy of images *M (i, j),* H (N) is the entropy of images *M (i, j),* H (N) is the entropy of image *N (I, j)* and H (M, N) is the joint entropy of image *M (i, j)* and *N (i, j).*

4. Enhancement Performance Measure (EME):-
   EME is used to measure the enhancement quality of the algorithm used for processing. Given an image X (n, m), let it be split into k1k2 blocks $W_{k,\ l}$(i, j) of sizes l_1xl_2. Then EME in terms of entropy is given by :-

$$EME = \min_{\phi\in\{\phi\}}(EME(\phi))$$
$$= \min_{\phi\in\{\phi\}}\left(\frac{1}{k_1k_2}\sum\sum 20\log\frac{I^w_{\max k,l}(\phi) - I^w_{\min k,l}(\phi)}{I^w_{\max k,l}(\phi) + I^w_{\min k,l}(\phi)}\right) \tag{17}$$

Where $I^w_{\max k,l}(\phi)$ and $I^w_{\min k,l}(\phi)$ are maximum and minimum of image X (n1, n2).

5. Normalized absolute error: - The lower the value of NAE, the better is the fused image. It is defined as:-

$$NAE = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(|A_{ij} - B_{ij}|)}{\sum_{i=1}^{m}\sum_{j=1}^{n}(A_{ij})} \tag{18}$$

All Rights Reserved © 2013 IJSETR

Table.1 Performance Analysis

Algorithm/ Parameters	**SIFT**	**Harris**
PSNR (dB)	41.36	39.869
EME	8.55	7.929
FSIM	0.619	0.612
MI	1.28	1.134
NAE	0.175	0.216
SSIM	.9834	1.145

V. RESULTS AND DISCUSSION

Two test images were captured by the camera DSC-ws-70. The technical specification of the above camera is as given:-

Maximum aperture - 2.75
Focal length - 14mm
Exposure Time - 0.02sec

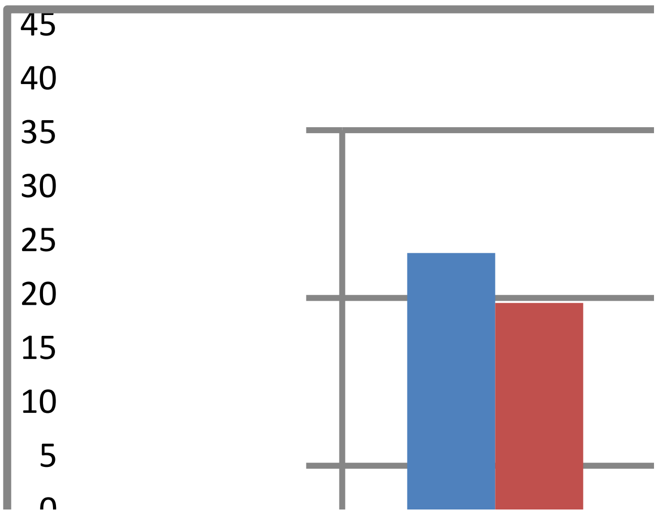
The two images have been captured at an angle of 10 degree of rotation between them to test for the rotational variance. In addition, there is illumination variation between the two images. The input image is first processed by the Harris algorithm, and then by the SIFT algorithm to generate the input panorama image. The panoramic images generated from these algorithms are passed through the blending process. SIFT algorithm is having a good Scale and rotational invariance property whereas Harris is superior algorithm for rotational image stitching, but it is not good with illumination variation. Thus, the panoramic image gets boosted up in terms of the common features present in both the images as well as complementary features gets compensated. The resultant image for Harris, and SIFT algorithms are compared in terms of PSNR (peak signal-to-noise ratio), MI (Mutual information), NAE (Normalized Absolute Error), FSIM (Feature similarity index measure), SSIM (Structural similarity index measure) and EME (Enhanced performance measure) as shown in Table .1.

Fig-4 Input Image 1

Fig-5 Input Image 2

Fig-6 Harris Response

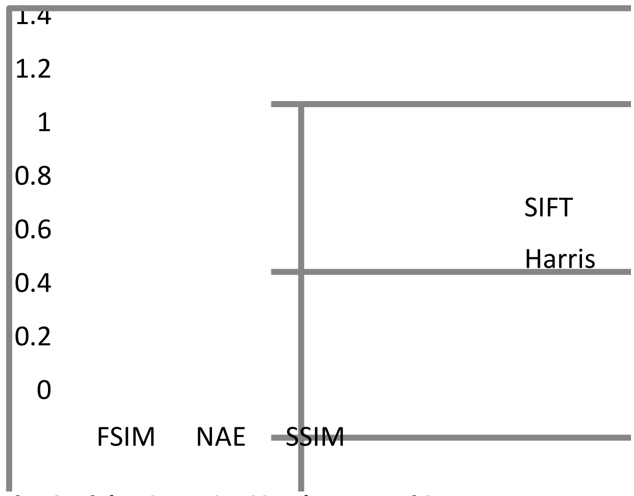
Fig-7 SIFT Response



Graph-1 Graph for PSNR, EME, MI for Harris and SIFT Responses (Blue-SIFT, Red-Harris)

All Rights Reserved © 2013 IJSETR

ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 2, Issue 11, November 2013



Graph-2 Graph for FSIM, NAE, SSIM for Harris and SIFT Responses (Blue-SIFT, RED-Harris)

VI. CONCLUSION

In this paper a comparative study is done among the image mosaicing algorithms; Harris, SIFT. The Harris algorithm is rotational invariant and it detects considerable more amounts of features with a simple calculation. SIFT has the property of affine, rotational invariant and scale invariant. The performance evaluation of proposed technique is done in terms of PSNR (peak signal-to-noise ratio), MI (Mutual Information), NAE (Normalized Absolute Error), FSIM (Feature similarity index measure), SSIM (Structural similarity index measure), and EME (Enhancement performance measure). The comparative study shows superior results for SIFT as compared to Harris and algorithm.

REFERENCES

[1] Qin Zhao, Tao Zhou, Qin Shu. The discussion based on the feature points of image registration technology. Infrared Technology. 2006(06).
[2] D. I. Barnea ;and H. F. Silverman, "A class of algorithms for fast digital registration," IEEE Trans. Comput, vol.C-21,pp.179-186,1972.
[3] C. D. Kuglin and D .C. Hines," The phase correlation image alignment method",in Proc. IEEE Int. Conf. Cybernet. Society, New York, NY, pp 163-165,1975.
[4] Lisa G. Brown. A survey of image registration techniques. ACM Computing Surveys, 24(4) ; pp 325-376, December 1992.
[5] Jain, D.K.; Saxena, G .;Singh, V.K., "Image mosaicing using corner technique" in International Conference on Communication System and Network Technologies ,pp 79-84,2012.
[6] D.G. Lowe, "Object recognition from local scale-invariant features", Proc. of the International Conference on Computer Vision, pp. 1150 - 1157 Sept. 1999.
[7] C. Harris, "Geometry from visual motion" A. Blake and A. Yuille, editors, Active Vision, pages 263–284. MIT Press, 1992.
[8] A.P. Witkin, "Scale-space filtering", International Joint Conference on Artificial Intelligence, pp.1019-1022, 1983
[9] Yu Wang, Yong-tian Wang, "The image matching algorithm based on SIFI and Wavelet transform", Journal of Beijing Institute of Technology. Vol.5, 2009.
[10] S. Krishnamoorthy, K. Soman, "Implement and Comparative Study of Image Fusion Algorithm", Intl. J. Computer Applications , vol.9, pp.25-34, 2010.
[11] D. Ghosh, S. Park, N. Kaabouch, W. Semke, " Quantum Evaluation Of Image Mosaicing In Multiple Scene Categories", *IEEE Conference on Electro/Information Technology*, pp. 1-6, 2012.
[12] P. Azzari, L.D. Stefano, M. Stefano, "An Evaluation Methodology For Image Mosaicing Algorithms", Advanced Concept for Intelligent Vision Systems, Springer, vol.5259, pp. 89-100, 2008.
[13] Lei Yang; Xiaoyu Wu; Jun Zhai; Hui Li, "A Research Of Feature-Based Image Mosaic Algorithm", Image and Signal Processing (CISP), pp.846 – 849, 2011.

Authors

[1] Hemlata Joshi

She received her B.E. degree in Computer Science and Engineering from Raipur Institute of Technology under Chhattisgarh Swami Vivekanand Technical University, Chhattisgarh India in 2010 and M. Tech. Computer Science and Engineering pursuing from Chhatrapati Shivaji institute of technology Durg under Chhattisgarh Swami Vivekanand Technical University, Bhilai Chhattisgarh India .Her research area of interests includes Digital Image Processing and Computer Network.

[2] Mr.KhomLal Sinha

He received his B.E. degree in Information Technology in year 2005 and M. Tech. in year 2012 (Computer Sc.) from Chhatrapati Shivaji Institute of Technology (CSIT), Durg, Chhattisgarh, India. His interests are Digital Image Processing, Operating Systems and Data Mining. Also he is having Life Membership of Indian Society of Technical Education, India (ISTE) and Institutional Member of Computer Society of India (CSI).

All Rights Reserved © 2013 IJSETR