Video Mosaicing Using SURF

Patel Foram N.[1], Neha Parmar[2]
[1]ME Student, Computer Engineering Department,
[2]Asst. Prof Computer Science & Engineering Department,
Parul Institute of Technology, Limda, Vadodara, India.

***Abstract*— Video Mosaicing is well thought-out as an active research area in computer vision and computer graphics. Video mosaicing is define as a video processing technique in which multiple images-frames are combined into a composite image that covers, more seamless and a larger view than the field of view of the compact camera. There are two main types of techniques used for creating mosaics: direct methods and featurebased methods. The greatest advantages of feature-based methods are Can Handle large disparities, Convergence, more accurate. For reliable performance direct methods rely on feature based initialization. We compared the performance of our proposed system with other fast feature-based techniques. We will do a comparative result between all those detectors according to similarity transformation problem, SURF algorithm for feature extraction, Image blending and wrapping are used.**

***Index Terms*— mosaicing, feature based method, SURF, similarity transform.**

I. INTRODUCTION

Mosaic Construction is an active area of research in computer vision and it has various application such as satellite photographs, video surveillance, stabilization, compression,virtual environments, virtual travels and 3D world scene medical imaging.[19]

Video mosaicing is a video processing technique in which multiple images are merged into a composite image that covers a larger, more seamless view than the field of view of the camera. The task of finding point correspondences between two images of the same scene or object is part of many computer vision applications. [3]

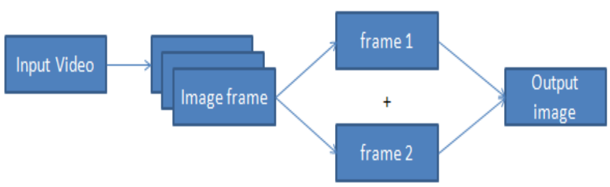


fig 1. general process for mosaic of video

In this paper, the creation of mosaic using SURF (speeded up robust feature) , which is fast than other state of art algorithm like SIFT and used for feature detection and matching and also used similarity transform which is helps to create smooth mosaicing. Also present the blending and wrapping for mosaicing blnding mix up the color so it seem good.

Feature detection is the process where we automatically examine an image to extract features, that are unique to the objects in the image, in such a manner that we are able to detect an object based on its features in different images. This detection should ideally be possible when the image shows the object with different transformations, mainly scale and rotation, or when parts of the object are occluded.

The processes can be divided in to 3 over all steps.

Detection Automatically identify interesting features, interest points this must be done robustly. The same feature should always be detected irregardless of viewpoint. **Description** Each interest point should have a unique description that does not depend on the features scale and rotation. **Matching** Given and input image, determine which objects it contains, and possibly a transformation of the object, based on predetermined interest points.[20][21]

II. LITERATURE SURVEY EASE OF USE

A. SIFT algorithm (SCALE INVARIANT FEATURE TRANSFORM)[7]

SIFT algorithm is method to extract features points and describe feature points, which is more robust to scale, rotation and change in illumination. There are four steps to implement the SIFT algorithm is as follow:

1. Scale-space extrema detection
2. Feature point localization
3. Orientation assignments
4. Feature point descriptor

(1) Scale-space Extreme Detection:

The first step to identify potential interest points that are invariant to scale and orientation, searches over scale space using a Difference of Gaussian (DoG) function is used. With the help of the convolution of variable-scale Gaussian $G(x, y, \sigma)$ with an input image $I(x,y)$ The scale space of image is defined as a function $L(x, y, \sigma)$ is produced:

$$\mathbf{L(x,y,\sigma)= G(x,y,\sigma)* I(x,y)} \ldots\ldots\ldots(1)$$

$$\mathbf{G(x,y,\sigma)=12\pi\sigma 2e-x2+ y22\sigma 2} \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

Using scale-space extrema in difference-of-Gaussian function convolved with the image, $D(x, y, \sigma)$which can be computed from the difference of two nearby scales separated by a constant multiplicative factor k for efficiently detect a stable key-point locations in scale space :

$$\mathbf{D(x,y,\sigma)= (G(x,y,k\sigma)- G(x,y,\sigma))*I(x,y) = L(x,y,k\sigma)- L(x,y,\sigma)} \ldots\ldots\ldots\ldots(3)$$

2) Feature Point Localization: In this step, The scale and location of each candidate point are determined and the feature points are chosen based on measures of stability and this information allows points to be discarded that have low contrast (and are therefore sensitive to noise) or are poorly localized along an edge.

3) Orientation Assignment: In this step, Based on local image gradient directions One or more orientations are assigned to each feature point location. For each image sample at this scale L(x, y), the orientation (,) and gradient magnitude m(x, y) are pre-computed using pixel differences:

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

$$\boldsymbol{\theta(x,y)}=\mathbf{tan-1((L(x,y+1)-L(x,y-1))(L(x+1,y)-L(x-1,y)))} \ldots\ldots(4)$$

4) Feature Point Descriptor: As shown on the left of figure 2.4.1, feature descriptor is created by computing the gradient magnitude and orientation at each image sample point in a region around the feature point location. These are weighted by a Gaussian windows, indicated by the overlaid circle. These samples are then accumulated into orientation histograms summarizing the contents over 4x4 sub-regions, with 8 orientation bins. So here each feature point has a 128-bins feature as shown on the right, with the length of each arrow corresponding to the sum of the gradient magnitudes near that direction within the region.

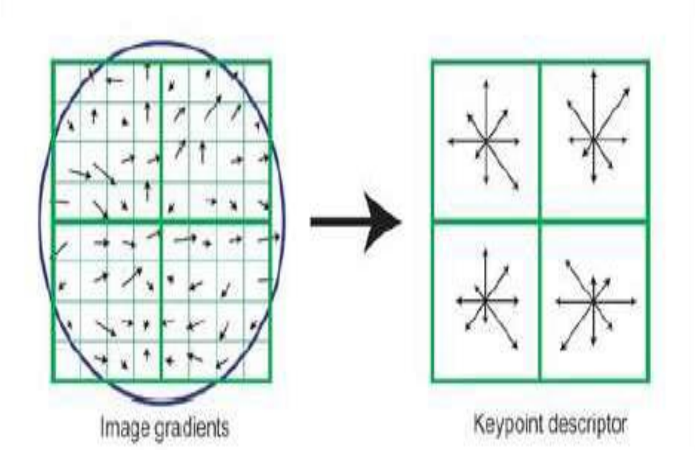


Fig 2: Feature descriptor creation

B. FAST[17]

From SUSAN with respect to a bimodal segmentation goal the FAST algorithm are derived. However, FAST algorithm relies on a connected set of pixels in a circular pattern to find out a corner. Either number possibly chosen, referred to as FAST9 and FAST10, the connected area size is commonly 9 or 10 out of a possible 16. FAST algorithm is known to be fast to match and efficient to calculate and accuracy is also relatively good. FAST can be measured a relative of the local binary pattern LBP. It may create many more edge detections at the given scale than a scale-space method such as used in SIFT because FAST is not a scale-space detector. As shown in Figure 2.4.2 , To find out if a pixel is less than or greater than the center pixel , FAST uses binary comparison with each pixel in a circular pattern against the center pixel using a threshold. The resultant descriptor is stored as a closest bit vector in order from 0 to 15. Also, it is possible to retrofit FAST and accumulate the bit vector in a rotational-invariant representation, as demonstrated by the RILBP descriptor, due to the circular nature of the pixel compare pattern.

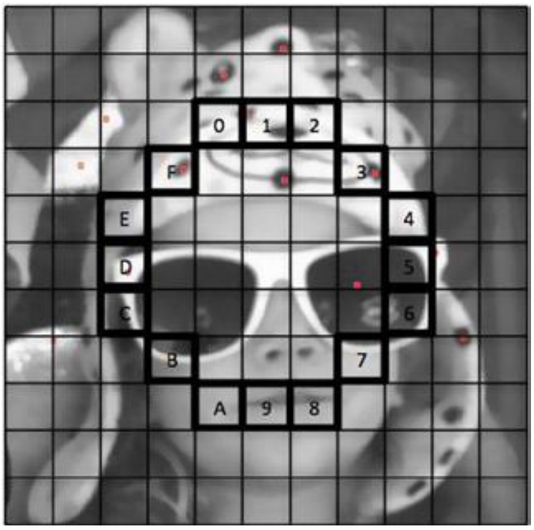


Fig 3 . The FAST detector algorithm

C. Harris feature[17]

The Harris feature or Harris-Stephens corner detector family provides improvements over the Moravic algorithm. To find the direction of fastest and lowest change for feature orientation, using a covariance matrix of local directional derivatives is the goal of the Harris algorithm. The directional derivative values are compared with a scoring factor to categorize which are edges, which features are corners, and which are likely noise. Depending on the formula of the Harris algorithm, the Harris algorithm can provide limited intensity invariance, high rotational invariance, and in some of the formulations of the algorithm, scale invariance is provided such as the Harris-Laplace technique using scale space . As a compute-efficient manner, many Harris family algorithms can be implemented. Note that basic proposal behind the Harris corner detector is corners have ill-defined gradient, since two edges converge at the corner, but near the corner the gradient can be detected with two different values with respect to x and y.

III. SEARCH CRITERIA

In this section the proposed algorithm flow chart and algorithm are there.

A. Mosaicing methods[7]

Mosaicing methods can be classified broadly into direct method and feature based method. Direct Method uses information from all pixels. It iteratively updates an estimate of homography so that a particular cost function is minimized. Sometimes Phase-Correlation is used to estimate the a few parameters of the homography. In Feature Based Method a few corresponding points are selected on the two images and homography is estimated using these reliable points only. Feature Based Methods are in general more accurate. It can handle large disparities. Direct methods, may not converge to the optimal solution is the presence of local minima. For reliable performance direct methods rely on feature based initialization. Feature based methods mosaic the images by first automatically detecting and matching the features in the source images, and then warping these images together. Normally it consists of three steps: feature detection and matching, local and global registration, and image composition.

Feature detection and matching aims to detect features and then match them. Local and global registration starts from these feature matches, locally registers the neighboring images and then globally adjusts accumulated registration error so that multiple images can be finely registered. Image composition blends all images together into a final mosaic. Direct methods attempt to iteratively estimate the camera parameters by minimizing an error function based on the intensity differences in the area of overlap. But this type of methods needs initialization, either by correlation or by manually setting some corresponding points. It is hard for the user to manually set the corresponding points correctly especially when the photographed scene does not have planar faces while Feature Based Methods mosaic the images by detecting the features in the images automatically, matching these features, and then creating the final mosaic image by warping other images related to one base image. Direct methods are useful for mosaicing large overlapping regions, small translations and rotations. Feature based methods can usually handle small overlapping regions and in general tend to be more accurate but computationally intensive.

B. SURF (speeded up robust feature)[20][21]

SURF is used to find the features like blob in the image, blob features can be found at corners of objects and also at where the reflection of light on specular surface is maximum. SURF is invariant to common image transformations, rotation, scale change, illumination change and small change in viewpoint.

The algorithm has three main parts Interest point detection, Interest point description, Matching.

Interest point detection uses very basic hessian-matrix approximation and integral images which reduces the computation time. Integral image is an intermediate representation for image and contains the sum of gray scale pixel values of image.

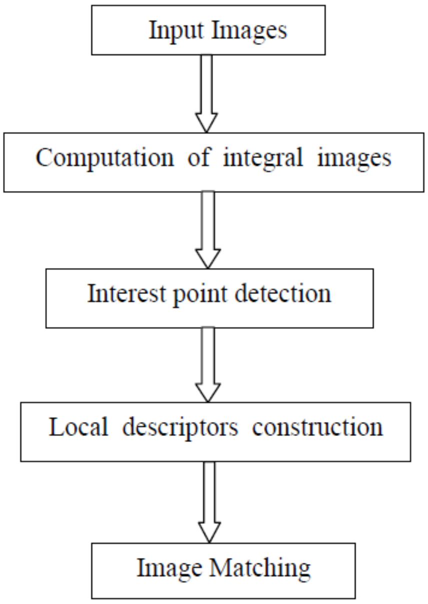


Fig 4. SURF feature detection algorithm

SURF uses a blob detector based on the Hessian matrix to find points of interest. It is the convolution of the Gaussian second order derivative with the image. The determinant of the Hessian matrix is used as a measure of local change around the point and points are chosen where this determinant is maximal.

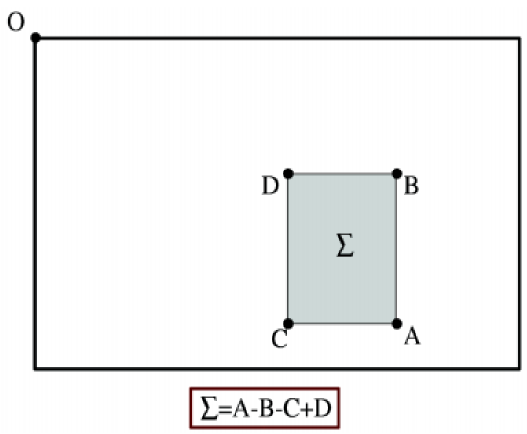


Fig 5. Integral image

Given a point p=(x, y) in an image I, the Hessian matrix H(p, σ) at point and scale σ, is defined as follows:

$$H = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{bmatrix}$$

Where, *Lxx* are the second-order derivatives of the grayscale image.

Interest point description is to provide a unique and robust description of a feature, a descriptor can be generated based on the area surrounding a interest point.

C. Similarity transform[23]

The ratio of magnification is define as A transformation that preserves the angles and changes all distances in the same ratio. A transformation that preserves ratios of distances can be also called a similarity transformation. A similarity transformation used for transforms images into similar images. Similarities are commonly referred to as similarity transformations, the ratio of magnification."

The term "similarity transformation" is used either to refer to a geometric similarity. similarity transformation is a conformal mapping whose transformation matrix can be printed in the form

$$A' \equiv B A B^{-1}$$

where A and A' are similar matrices.
Translation , rotation , scaling supported by Similarity transformation.
Examples of similarities include the following.
1. Central dilation: A lines to parallel lines are transformed that isn't merely a translation.
2. Geometric contraction: Scale is reduced in transformation.
3. Dilation: Whose length is fixed multiple of the length of the original line, a transformation taking each line to parallel line.
4. Expansion: Scale is increased in transformation.
5. Isometry: Preserves the distances in transformation.
6. Reflection: All points are exchanged with their corresponding reflections in infinite plane mirror in this transformation.
7. Rotation: Preserves angles and distances in transformation.
8. Improper rotation: Reflection through the origin ,combined with a rotation.
9. Translation: With no rotation or distortion , transformation consisting of constant offset."

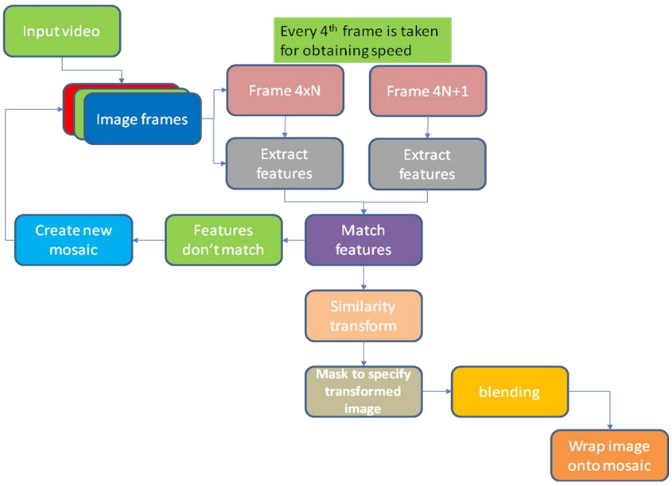
D. Blending[10][11]

To avoid the seams The final step is to blending the pixel colors in the overlapped regions. To utilize feathering is a simplest available form, to blend the overlapping pixels which uses weighted averaging colors values. Image blending is technique, To obtain a smooth transition between images by removing these seams and creating a blended image by determining how pixel in an overlapping area should be presented, which modifies the image gray levels in the surrounding area of a boundary."

E. Wrapping[10][11]

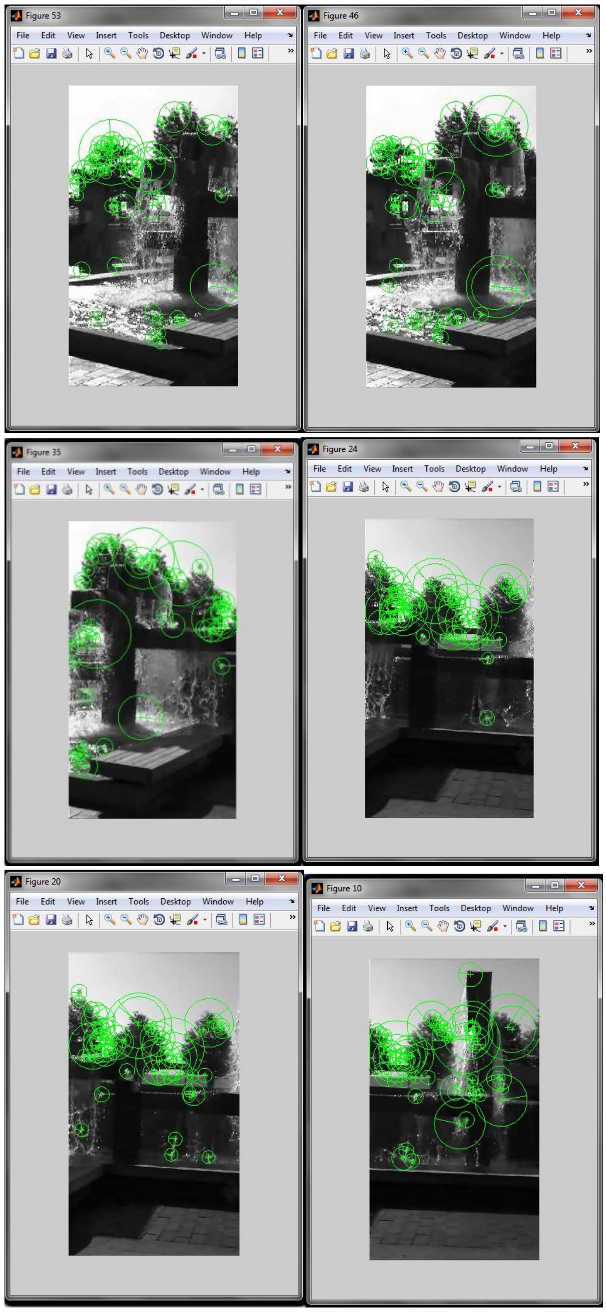
Image warping is“a method of digitally manipulate an image such that any shapes portrayed in image have been notably distorted. Defined by one of them known as reference image basically we can simply warp all the input images to plane. For correcting image distortion as well as for creative purposes also used wrapping. The same method are equally applicable for video. By using the geometric transformation, the two images that will form the mosaic are warped. While image can be transformed in various ways, pure warping means that Without changing the colors points are mapped to points. It can be mathematically based on any function from (part of) plane
to plane. It can be reconstructed if function is put in the original.”
For generation of an image for any kind of distortion there are two methods.
Forward-mapping: A given mapping as of sources to images is straightforwardly applied
Reverse-mapping : for a given”mapping from sources to images, the source is invent from the image
Optical “flow estimation techniques are used for determine the category of wrapping which takes place between consecutive images.

F. flow diagram



IV. EXPERIMENTAL ANALYSIS

Here , Experiment is carried out in this section. Fig 6 shows the extracted features in the frames. Here for creating mosaic from video, video is converted into a frames and fig 6 shows some of them frame in which features are extracted which shows by green blob feature.



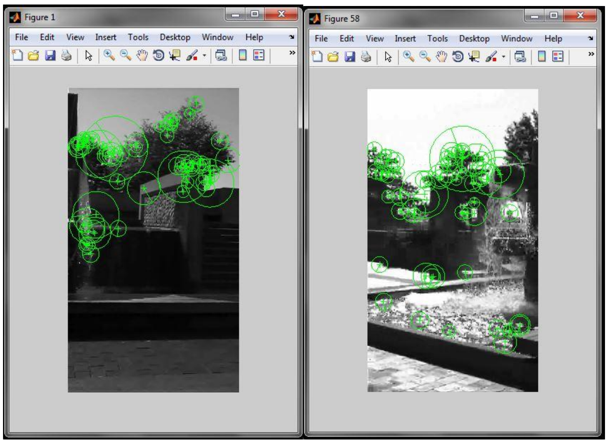
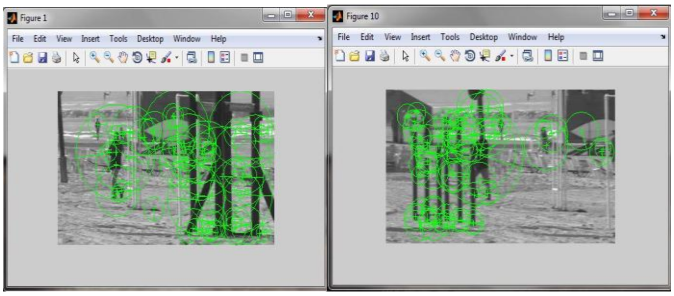


Fig 6 extacted feature in frames(here some of the frames shows from video and green blob are detected surf features)

The mosaics in fig 7 which made from video and good in quality and comparatively less noise mosaic created.

We have implemented proposed method in MATLAB and tested it by extracted strongest ~60 features. Here video is high quality and it is comparatively fast.

fig 7 Mosaic from video



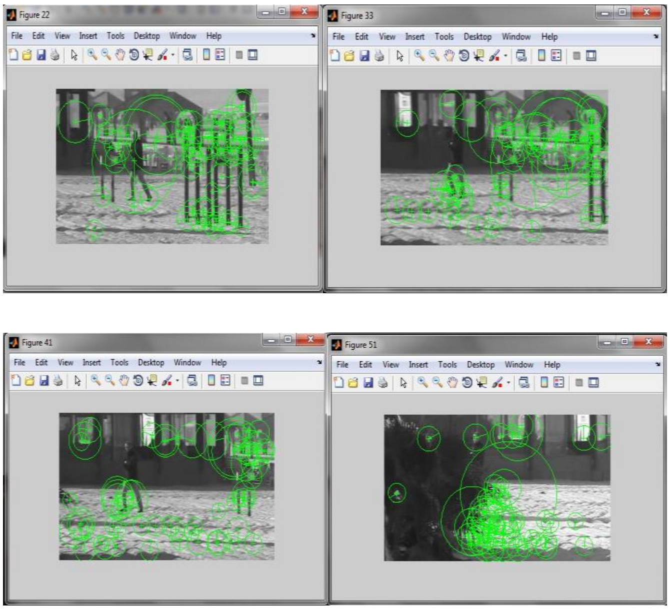


Fig 8 extracted features on another video frames

Fig 8 shows on another video example in which feature extracted frames are shownand the fig 9 show this frames mosaic construction.

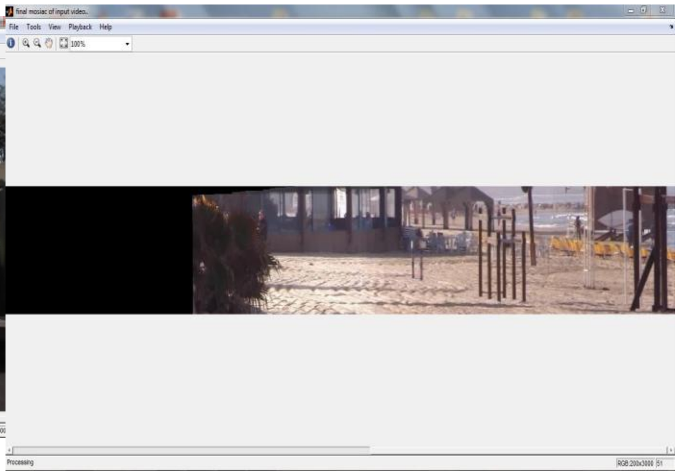


Fig 9 mosaic from another video

V. CONCLUSION

Video Mosaicing techniques are widely used in creating panoramic or mosaic of video. In this thesis, some of the popular algorithm are explores. SURF algorithm is used with similarity transform, blending and wrapping and creating fast and accurate way of video Mosaicing and results show the accuracy same. The presented generation of video mosaicing is good in quality than the previous approaches. SURF helps to feature extract and match faster than other state of art algorithm and similarity transform transform frame and reconstruction region as compare to other. Blending used for mixup color for smooth mosaic.

ACKNOWLEDGMENT

The author like to thank faculties of CSE department of PIT, Family and friends for their details review, valuable comments, constructive suggestion and valuable support.

REFERENCES

[1] Marc Vivet, Shmuel Peleg, Xavier Binefa.,"Real-Time Stereo Mosaicing using Feature Tracking", 2014 IEEE.

[2] Frank Liu, Rob Hess, Alan Fern.,"Efficiently Constructing Mosaics from Video Collections", 2015,IEEE.

[3] Ken-ichi Okumura, Sushil Raut, Qingyi Gu, Tadayoshi Aoyama, Takeshi Takaki, and Idaku Ishii "Real-time Feature-based Video Mosaicing at 500 fps", 2013,IEEE.

[4] Edgardo Molina, Zhigang Zhu.," Persistent Aerial Video Registration and Fast Multi-View Mosaicing", 2014,IEEE.

[5] ZOU Xiao-chun, HE Ming-yi, ZHAO Xin-bo, FENG Yan ,"A Robust Mosaic Panorama Technique for Video",2010,IEEE.

[6] Tom Botterill, Steven Mills, Richard Green," Real-time aerial image mosaicing", 978-1-4244-9631-0/10/$26.00 2010 IEEE.

[7] Ms. Parul M.Jain, Prof. Vijaya K.Shandliya,2013 "A Review Paper on Various Approaches for Image Mosaicing"2005,IEEE

[8] Utsav Shah, Darshana Mistry, Yatin Patel "Survey of Feature Points Detection and Matching using SURF, SIFT and PCA-SIFT" JETIR (ISSN-2349-5162)

[9] Lei Yang, Jun Zhai, Xiaoyu Wu, Hui Li."A Research of Feature-based Image Mosaic Algorithm", 2011 4th International Congress on Image and Signal Processing

[10] Dr. techn. Bernhard Rinner, Prof. Gian Luca Foresti "Orthorecti_ed Mosaicking of Images from Small-scale Unmanned Aerial Vehicles" March 2013

[11] Vimal Singh Bind, Prof. Umesh Chandra Pati ,"Robust Techniques for Feature-based Image Mosaicing", 211EC3304,2013.

[12] Jack Sparrow, Pirates of the Caribbean "Interest Point Detector and
Feature Descriptor Survey" Book

[13] Tom Botterill, Steven Mills, Richard Green "Real-time aerial image mosaicing" 2010,IEEE

[14] Roberto Marzotto, Andrea Fusiello, Vittorio Murino "High Resolution Video Mosaicing with Global Alignment" 2004,IEEE

[15] Shmuel Peleg, Moshe Ben-Ezra, Yael Pritch "Omnistereo: Panoramic Stereo Imaging" 2001,IEEE

[16] Utsav Shah, Darshana Mistry, Yatin Patel " Survey of Feature Points Detection and Matching using SURF, SIFT and PCA-SIFT "JETIR

[17] BOOK- "Feature detection and matching"

[18] Chiou-Ting Hsu, Tzu-Hung Cheng, Rob A. Beukers and Jyh-Kuen Horng"FEATURE-BASED VIDEO MOSAIC"IEEE 2000

[19] jing li,quan pan,yao yang,stan.z.li , "Automated feature points management for video mosaic construction", IEEE 2015.

[20]Herbert Bay , Andreas Ess , Tinne Tuytelaars , and Luc Van Gool ," *Speeded-Up Robust Features (SURF)* "2008.

[21] Jacob Toft Pedersen, "*Study group SURF: Feature detection & description*",2011.

Sites:

[22] http://www.engineersgarage.com/articles/image-processing-tutorial-applications

[23] http://in.mathworks.com/discovery/similarity-transformation.html

[24] www.mathworks.com/help/matlab/

[25] https://en.wikipedia.org/wiki/MATLAB