© April 2016 | IJIRT | Volume 2 Issue 11 | ISSN: 2349-6002

A Review for Video Mosaicing Using Feature Tracking

Foram N. Patel[1], Neha Parmar[2]
[1]*Department of Computer Engineering, Parul Institute of Technology, Vadodara*
[2]*Asst. Prof., Department of Computer Engineering, Parul Institute of Technology, Vadodara*

***Abstract*- In recent years, videos/images have become extremely important in our daily life information. Mosaic construction is an active area of research in computer vision. Mosaic is one of the most important tasks in a variety of video processing and analysis applications, by means of which a synthetic representation of successive video frames is constructed. The application of mosaic include virtual reality environments, video compression, video indexing and representation. 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. Video mosaicing has been used to generate panoramic pictures in many applications such as tracking, surveillance and augmented reality. Most video mosaicing methods are SIFT feature, SURF feature, Harris corner, MSER Features and so on. This paper encompasses a description of some feature extraction methods.**

***Index Terms*- video mosaicing, sift feature ,surf feature, affine transform, wrapping.**

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.[9]

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.[1]

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.[10][11]

II. FEATURE EXTRACTION METHODS

A SURF (SPEEDED UP ROBUST FEATURE)[10][11]

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.

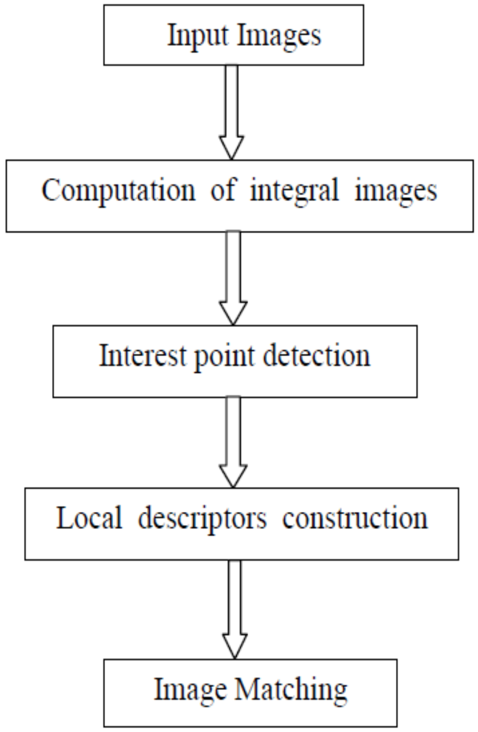

Fig SURF feature detection algorithm[13]

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.

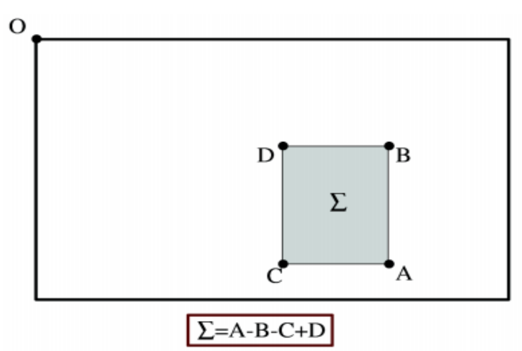

Fig. Integral image

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, L_{xx} 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.

B SIFT (SCALE INVARIANT FEATURE TRANSFORM)[4]

Scale Invariant Feature Transform termed as SIFT is used to identify locations and scales that can be repeatedly assigned under different views of the same object There are four steps to implement the SIFT algorithm Scale-space extreme detection, Feature point localization, Orientation assignments, Feature point descriptor.

Scale-space Extrema Detection searches over scale space using a Difference of Gaussian (DoG) function to identify potential interest points that are invariant to scale and orientation. The scale space of an image is defined as a function L (x , y , σ) which is produced from the convolution of a variable-scale Gaussian G(x , y , σ) with an input image I(x,y):

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

$$\mathbf{G}(\mathbf{x},\mathbf{y},\boldsymbol{\sigma}) = \mathbf{12\pi\sigma 2}\boldsymbol{e}\mathbf{-}\boldsymbol{x}\mathbf{2+}\boldsymbol{y}\mathbf{22}\boldsymbol{\sigma}\mathbf{2} \text{..........}(\mathbf{2})$$

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

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

Feature Point Localization use location and the scale of each candidate point are determined and the feature points are selected based on measures of stability this information allows points to be rejected that have low contrast (and are therefore sensitive to noise) or are poorly localized along an edge.

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

pixel differences:

m(x ,y)=√(L(x+1,y)−L(x−1,y))2+(L(x,y+1)−L(x,y−1))2

$\theta(x,y)$= tan−1((L(x,y+1)−L(x,y−1))(L(x+1,y)−L(x−1,y))) …….(4)

Feature Point Descriptor is created by first 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 window, 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 each feature point has a 128-element 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.

C. FAST Algorithm [6]

FAST is a corner detector algorithm. Trajkovic and Hedley founded this algorithm in 1998.The detection of corner was prioritized over edges in FAST as corners were found to be the good features to be matched because it shows a two dimensional intensity change, and thus well distinguished from the neighboring points. According to Trajkovic and Hedley the corner detector should satisfy the following criteria:-

1. The detected positions should be consistent, insensitive to the variation of noise, and they should not move when multiple images are acquired of the same scene.
2. Accuracy; Corners should be detected as close as possible to the correct positions.
3. Speed; the corner detector should be fast enough.

FAST incremented the computational speed required in the detection of corners. This corner detector uses a corner response function (CRF) that gives a numerical value for the corner strength based on the image intensities in the local neighborhood.CRF was computed over the image and corners which were treated as local maxima of the CRF.A multi-grid technique is used to improve the computational speed of the algorithm and also for the suppression of false corners being detected. FAST is an accurate and fast algorithm that yields good localization (positional accuracy) and high point reliability.

III. AFFINE 2D TRANSFORM[12]

Affine transformation is a linear mapping method that preserves points, straight lines, and planes. Sets of parallel lines remain parallel after an affine transformation. The affine transformation technique is typically used to correct for geometric distortions or deformations that occur with non-ideal camera angles. For example, satellite imagery uses affine transformations to correct for wide angle lens distortion, panorama stitching, and image registration. Transforming and fusing the images to a large, flat coordinate system is desirable to eliminate distortion. This enables easier interactions and calculations that don’t require accounting for image distortion. Different affine transformations are translation, scale, shear, and rotation.

IV. IMAGE WARPING[9]

Image warping is the process of digitally manipulating an image such that any shapes portrayed in the image have been significantly distorted. Basically we can simply warp all the input images to a plane defined by one of them known as reference image. Warping can also be used for correcting image distortion as well as for creative purposes. The same techniques are equally applicable to video. The two images that will form the mosaic are warped, by using the geometric transformation. While an image can be transformed in various ways, pure warping means that points are mapped to points without changing the colors. It can be mathematically based on any function from the (part of) plane to the plane. If the function is put in the original then it can be reconstructed.

There are two methods for generation of an image for any type of distortion.

1.Forward-mapping: a given mapping from sources to images is directly applied

2.Reverse-mapping : for a given mapping from sources to images, the source is found from the image

To determine the type of warping which takes place between consecutive images, optical flow estimation techniques are used.

V. IMAGE BLENDING[9]

Image blending is the technique, which modifies the image gray levels in the vicinity of a boundary 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.

VI. CONCLUSIONS

Video Mosaicing techniques are widely used in creating panoramic or mosaic of videos. In this paper, some of the popular algorithms have been explores. The FAST algorithm is both rotation and scale invariant with optimized execution time. But, its performance is poor when noise is present. SIFT algorithm is rotation as well as scale invariant and more effective in presence of noise. It has highly distinctive features However, it suffers from illumination variation.

REFERENCES

[1] 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. .

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

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

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

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

[6] 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

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

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

[9] Jing li, Quan Pan, Yao Yang, Stan.Z.Li,"*Automated Feature Points Management For Video Mosaic Construction*",IEEE,2005.

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

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

[12] http://in.mathworks.com/discovery/affine-transformation.html

[13]https://en.wikipedia.org/wiki/Speeded_up_robust_features