International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 3 Issue: 6 3514 - 3519

Image Stitching Using Speeded Up Robust Features

Mr. Venugopala Gowda D, (MTech)
Electronics and Communication Engg Dept.,
Malnad college of Engineering, VTU-Belagavi,
Hassan,Karnataka,INDIA.
Email: venudgowda@gmail.com

Mrs..Padmajadevi G, M.Tech
AssosiateProfesor,Electronics and Communication Engg Dept,.
Malnad college of Engineering, VTU-Belagavi,
Hassan,Karnataka,INDIA.
Email: padmajadevig@gmail.com

***Abstract*:** Images are an integral part of our daily lives. Image stitching is the process to generate one large panoramic image from collective relative image sequence without overlapping. Stitched images are used in applications such as Defence documentary photos, glob map, Inside or outs of the gallery, Space research images, sculpture art etc., the stitched images multi-node movies and other applications associated with modelling the 3D environment using images acquired from the real world. The existing method and algorithms are don't have enough accuracy to make larger panoramic image. On-time collective of image sequence more easier compare to pre-collective images stitching. The better method for image stitching is **SURF** (Speeded Up Robust Features) algorithm. This robust method help out to stitches relative images much faster than SIFT (Scale Invariant Features Transform) algorithm. Most image processing techniques involve treating the image as a two-dimensional signal and applying standard signal processing techniques to it. Specifically, image stitching presents different stages to render two or more overlapping images into a both horizontal as well as vertical moves stitched image, from the detection of features to blending in a final panoramic image.

***Keywords**: SIFT (Scale Invariant Features Transform), SURF (Speeded Up Robust Features), RANSAC (Random Sample Consensus).*

*****

1. Introduction

This paper main goal is to make that will stitch two or more images together to create one larger image. Taken a series of images from a single point in space, but with little different orientations, it is possible to map the images into a common reference frame and create a perfectly arranged larger needy photograph with a wider vision; this is normally referred to as panoramic image stitching.

Image mosaic is mainly composed of image registration and fusion. Image registration means to identify overlapping areas of the image to be spliced and then determine the transformation relationship. Image fusion refers to stitch the images to form a smooth and seamless image. And image registration is the core technology of all the stitching process. Image mosaic or stitching technology is one of the gradual improving digital image processing technologies applied to various fields. With the help of digital system, it can do matching work automatically, and from a seamless, high definition, wide sight angle image through aligning a series of spatial overlapping single image. In this project, I observe that a successful image stitching algorithm should not only create a smooth variation within the overlapped region but also preserve the following properties. The stitched image should not break existing or the overlapped region due to structure misalignment, causing obvious ghosting artefact, Intensity adjustment. Normally human eyes are sensitive to large intensity change not to the smaller changes. Unbalanced contrast beyond the overlapped region of a stitched image can be perceptually magnified. There will be uncommon colour transition from left to right reveals the unmatched intensities inherent in the input images. The context information of objects in the input images should be taken into account during the stitching process.

2. SURF (Speeded Up Robust Features)

As for the different methods of stitching, image registration fall broadly into three methods: gray information based, transform domain based and feature based. Among them, the feature-based image stitching technology is widely used because it has the quality of high time efficiency, maximum matching accuracy and good robustness.Point feature is an important feature of the image in a various image features[1]; it has the benefits of rotational invariance, not varying with changes in light conditions and high speed. The common feature points are Harris corner detection, SIFT (Scale- Invariant Feature Transform) and SURF (Speeded Up Robust Features). For all above features has compared the commonly used local invariant features and found SURF feature detection is more effective than other feature detection. So in this project, we propose a fast stitching method based on SURF. It can be majorly divided into four steps: feature points extraction, feature points matching, determining the transformation relationship and image fusion.

3. Scale Invariant Feature Algorithm

The SIFT algorithm (Scale Invariant Feature Transform) proposed by Lowe is an approach for extracting distinctive invariant features from images. It has been successfully applied to a variety of computer vision problems based on feature matching providedobject recognition, face recognition, pose estimation, image retrieval and many others. However, in real world applications there is still a need for improvement of the algorithm's robustness with respect to the correct matching of SIFT features. In this paper, an improvement of the original SIFT algorithm providing more reliable feature matching for the purpose of object recognition is proposed. The main idea is to divide the features extracted from both the test and the model object image into several sub-groups before they are matched. The features are divided into several sub groups considering the features arising from different octaves, that is from different frequency domains. Following are the major stages of computation used to generate the set of image features:

1. *Scale-space extreme detection*: The first stage of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation.
2. *Key-point localization*: At each candidate location, a detailed model is fit to determine location and scale. Key-points are selected based on measures of their stability.
3. *Orientation assignment*: One or more orientations are assigned to each key-point location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.
4. *Key-point descriptor*: The local image gradients are measured at the selected scale in the region around each key-point. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

This approach has been named the Scale Invariant Feature Transform (SIFT), as it transforms image data into scale invariant coordinates relative to local features.

Drawback of SIFT algorithm

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.

4. System design and implementation

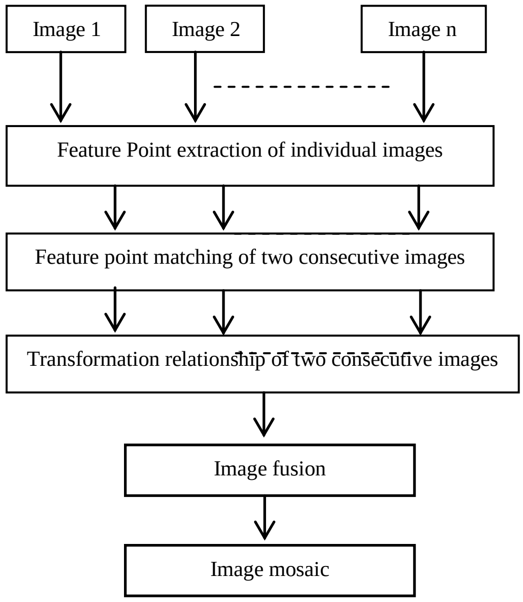

Fig.1. Image stitching flow based on point feature

5. SURF Feature point detection

SURF is a feature point extraction algorithm and it is three times faster than commonly SIFT algorithm and the overall performance is much better than SIFT algorithm. SURF feature point extractions include feature detection and feature points description. Generally SURF involves three steps: establishing integral image, building scale-space image and positioning feature points. The entry of an integral image $I \sum (x)$ at a location x = $(x, y)^T$represents the sum of all pixels in the input image I within a rectangular region formed by the origin and x.

$$I(x) = \sum_{i=0}^{i\le x} \sum_{j=0}^{j\le y} (i,j) \qquad (1)$$

The main reason obtain of integral image is that it is used to accelerate the convolution between original images and box filters with different sizes in the process of SURF feature detection. When building scale space, we conduct the image convolution by leaving the original image size unchanged and changing the size of the box, which form an image scale space. Generally every four boxes filter are called an order (Octave). In the first Octave, theadjacent box filter size at 6 pixels, the second Octave differ at 12 pixels, the third Octave differ at 24 pixels, and so on. The size of the first box in each octave is same as the second box in the forward octave. We can locate the feature points based on the established image scale space. We detect the extreme points of the image using fast Hessian matrix at each level image scale space. Taken a point $X = (x, y)$ in the space, the Hessian matrix in point X at scale s is defined as follows:

ISSN: 2321-8169
3514 - 3519

$$H(X,s)=\begin{bmatrix} L_{xx}(X,s)L_{xy}(X,s) \\ L_{xy}(X,s)\ L_{yy}(X,s) \end{bmatrix} \qquad (2)$$

Where L_{xx}(X,s),L_{yy}(X,s) and L_{xy} (X,s)is the convolution of the Gaussian second order derivative with the image and point *X*. To reduce the amount of calculation, Bay and someone else made an approximation, that is replacing L_{xx} ,L_{xy} and $L_{yy}$with D_{xx} , D_{xy} and D_{yy} ,which represent the convolution of the box filter with the original image.This yields:

$$Det\ (H_{aprox})=D_{xx}D_{yy}-w(Dxy)^2 \qquad (3)$$

Where *w* is a weighting factor is 0.9 in general. We set a threshold value of the calculation result. We judge the extreme points only when it is greater than the threshold, then the pixel will be compared with eight nearby pixels of the same scale and each nine pixels of the adjacent upper and lower scale, we will obtain the local maximum point which are marked as feature points.

6. Description of feature points

In order to make the descriptor to be rotation invariant, we first get the orientation of the feature points. So we first calculate the Haar wavelet responses in *x* and *y* direction within a circular neighbourhood of radius 6*s* around the feature points and the responses are given different Gaussian weighting coefficients which make the contribution is greater if it is closer to the feature points. Then a local direction vector is formed by adding the Haar response in *x* and *y* directions within 60°, the direction of the longest vector is selected as the main direction of the feature point by traversing the entire circular area. Then we construct a square region centred on the interest point and the size of this window is 20*s*. The region is split up regularly into smaller 4 x 4 square sub-regions. This saves important spatial information. For each sub-region, we compute Haar wavelet responses at 5 x 5 regularly spaced sample points. Conveniently, we call *dx* the Haar wavelet response in horizontal direction and *d*y the Haar wavelet response in vertical direction. To increase the robustness towards geometric deformations and localisation errors, the responses *dx* and *d*y are first weighted with a Gaussian centred at the interest point. So we get a 4D vector V = ($\sum d_x$, $|\sum d_x|$,$|\sum d_y|$,$|\sum d_y|$). Attaching this for all 4 x 4 sub-regions, then result in a descriptor vector of length 64. The wavelet responses are invariant to a change in illumination. Invariance to contrast (a scale factor) is gained by changing the descriptor into a unit vector, as shown in Fig.2.

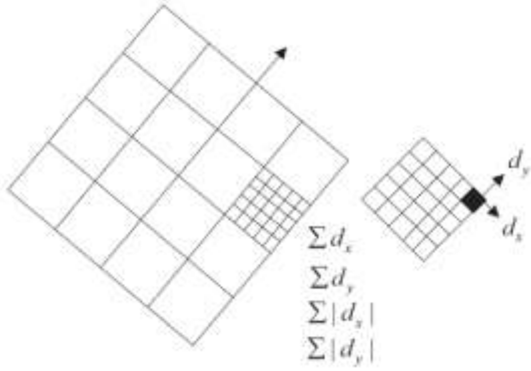

Fig.2. Composition of Feature descriptors

7. Feature Matching

Matching feature points can be attributed to a similarity search problem by distance function in high-dimensional space. Although exhaustive search can search the correct match point, it consumes large. As for common K-D tree, the search efficiency will decline for high-dimensional data and it will be not used anymore while the efficiency is close to exhaustivesearch. Taking the cost of matching time into account, we use BBF (Best-Bin-First) feature matching algorithm for the above generated feature vectors. When looking for matching feature points, the nearest neighbour searching algorithm focus only on the node position which has been stored while ignoring inquired node position. BBF algorithm by means of a priority queue priority queue, the search process is as follows:

• Search start from the root of three until a leaf node and find the first Euclidean nearest neighbour feature points.

• Unselected nodes (actually a pointer pointing to node) are pushed into the priority queue, and next time the search will begin from the sub-tree which has the smallest key value(the key value is the absolute dimension of the target feature points and unselected branches of dim dimension when one node decline).

• In this paper, repeat the process until the queue is empty or repeated 50 times so far, and ultimately find coarse matching pairs.

8. Transformation Relationship

Before the image fusion, the initial matrix transformation matrix *H* of the image must be first calculated. Image transformation matrix has eight parameters, these parameters representing the picture of the scale variable rotation, horizontal and vertical displacement. Projection transformation formula obtained is shown below.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}=\begin{bmatrix} h_o & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & h_8 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (4)$$

ISSN: 2321-8169
3514 - 3519

Where H in the formula is the parameter of the imagetransformation matrix. (x, y) and (x', y') are respectivelymatching the coordinates of feature points.There may be mismatch in the matching pairs obtained bythe Nearestneighbor matching method. To calculate the relationship between images, we combined with RANSAC and the least squares method in this paper, the basic step is asfollows [1]:

a) Repeat sampling with n times, three groups of thecorresponding point are randomly taken out to composite asample then calculate the matrix H. The assumed corresponding distance d of each group were calculated thencompared with the limits. The points consistent with H will be
selected as the inliers.

b) Select a set of points that contains the maximumnumber of inliers (when the numbers of the points are equal,choose the smallest standard deviation of the set of points).

c) Recalculate H with the selected set of matchingpoint, using the least squares method to minimize the error. Sobefore the final solution, we first remove the most pointswhich do not meet the most satisfied relationship andmismatch effects and finally obtain the solution that satisfiesthe majority of the matching points.

9. Image Fusion

There will be obvious splice seam if stitching directlyafter registration due to different lighting conditions whenshooting. We must use some methods of image fusion toeliminate the seam. The commonly used image fusion methodare direct average method, the weighted average method andmulti-resolution spline method. In this paper, we use gradatingin-and-out amalgamation algorithm. The weight of theoverlap region pixel is related to the distance between pixeland overlapping areas. The pixel relationship between imagemosaic and two images needed to stitch is as follows:

$$f(x,y)=\begin{cases} I_1(x,y) & (x,y)\in I_1 \\ d_1I_1(x,y)+d_2I_2(x,y) & (x,y)\in I_1\cap I_2 \\ I_2(x,y) & (x,y)\in I_2 \end{cases} \quad (5)$$

Where d_1,$d_2$represent the weight which is related to thewidth of the overlap region, we select d = 1/*width* and$d$ + d =1. In the overlapping area, 1 d change from 1 to 0,while 2 d change from 0 to 1. Finally, we achieve a smoothtransition diagram, as shown in Fig.

PB (i,j) = (1-w) * PA (i,j) + w * PB (i,j) (6)

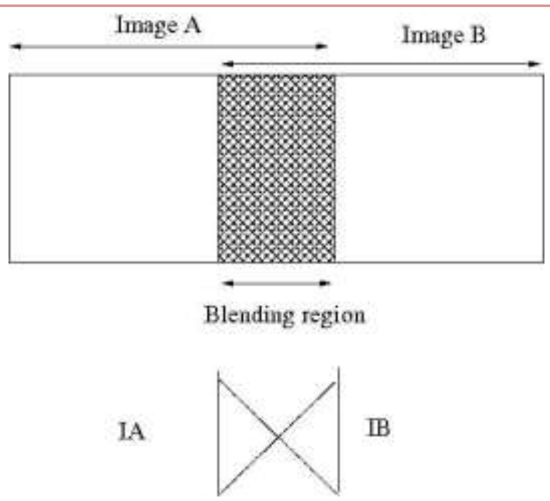

Fig.3. Change in weights

10. RanSaC Translation

RanSaC algorithm is general purpose algorithm that can be used to calculate full homography in the presense of outliers. The use of cylindrical warping has the advantage that only the translation motion has to calculated on the warped image. Also for the translation estimation, only one feature is sufficient. Ransac estimation counts the inliners based on some tolerance value (d<eps) which depends on the noise present in the images. Since our image capturing was high quality, only two pixel tolerance was sufficient to get good estimation [2]. (We found that that in our images only 5-10% outliers).

Image Blending:

When different images are stitched together, for various reasons (changed lighting conditions, vignette effects) the adjacent pixel intensities differ enough to produce artifacts as shown in the following pictures. To remove these artifacts, we experimented with two algorithms Feathering Image pyramids [8].

11. Result

Step 1 – Acquiring Images

Collect the Image sequence with vertical or horizontal move. These were taken with regular and common size and pixels. The images are usually gives unequal intensity and angular position. These non-regular formats of the images will be the challenging task to my project. Good calibrated camera will makes better result.

Step 2 - Reviewed the successive image

To make the panorama, start by registering successive image pairs using the following procedure:

1. Find and match features points between Image-1 and Image-2.

2. Estimate the geometric transformation, successive image and that maps those.
3. Compute the transformation that maps into the panorama image.
4. Find key-points and descriptors of successive images by using SURF algorithm.

Step 3 - Initialize the Resulted Image

Create an initial empty Image into which all the images are mapped. Here we stitch one image at a time onto the cumulative result.

Step 4 - Create the Panorama

Blending the complete images together to improve overall accuracy and to get full view panoramas.

Images	SIFT	SURF
krishna	234 sec	64 sec
Bicycle	222sec	54 sec
Building	256 sec	66 sec
Lena	188 sec	50 sec

Table.1. Time taken Time taken to get key

Sample images: *Example-1*

Image-1

Image-2

Feature points of two images

Resulted Image 1 and 2

12. Conclusion

I implemented and checked out of stitching two images, successfully by the correlated method the result shown below. The stitched images are the sample piece of stored pictures those have the 15 to 30 percent common position, but the same algorithm will not suite for the current situation. So in my project I preferred do with feature points and descriptor points matching by SURF and RANSAC algorithms. The results will be quite nice. While running, the program might display some intermediate results, like the set of *tentative correspondences* (pairs of matched points between two images) and inliers (estimated in RANSAC). Inliers are always subset of tentative correspondences. A sample result is shown before. The experimental results show that our method computes features with higher precision and can realize smooth transition in images of same scenes. The same algorithm can also best situate for collected images (not to live merge images) and as well in video extracted image blending. In feature work this will be worth to build 3D images its printing application.

References:

[1] Fei Lei and Wenxue Wang ‘A Fast Method for Image Mosaic Based on SURF’ (IEEE 978-1-4799-4315) September 2014.

[2] V S Veena Devi, S.M. Prasad and H. Gangadhara Bhat2 ‘An efficient algorithm for image based on SIFT’ February 2013.

ISSN: 2321-8169
3514 - 3519

[3] XIE Kai, GUO Heng-ye, Zhang Tian-wen. A Survey of Image Mosaics Technology[J]. ACTA ELECTRONICA SINICA, 2004, 4(4):630-634.

[4] Yajuan Zhang, Research on Image and Video Mosaic Based on SURF, 2013.

[5] Johannes Bauer, NikoSunderhauf, Peter Protzel, "Comparing several implementations of two recently published feature detectors," In Proc. of the International Conference on Intelligent and Autonomous Systems, IAV, Toulouse, France, 2007.

[6] Herbert Bay, "Speeded-Up Robust Features (SURF)", Computer Vision and Image Understanding 110 (2008) 346–359.

[7] Shi Lu, Su Gang, Han Fei. "An Improved Algorithm Based on Sift and Surf Image Stitching Algorithm", Computer Application and Software,Vol.30. No.6, Jun.2013.

[8] [8].Richard Szeliski. 'Video mosaics for virtual environments. Computer Graphica and pplications',IEEE,Vol.16,No.2,pp.22-30,1996.