

Panoramic Image Stitching based on Feature Extraction and Correlation

Arya Mary K J ¹, Dr. Priya S ²

PG Student, Department of Computer Engineering, Model Engineering College, Ernakulam, Kerala, India¹

Professor, Department of Computer Engineering, Model Engineering College, Ernakulam, Kerala, India²

ABSTRACT: Image stitching is an active research area in computer vision and computer graphics. Image stitching is combining two or more images of the same scene to produce a high resolution image which is called as panoramic image or mosaic. Image stitching techniques can be categorized into two general approaches: direct and feature based techniques. Feature based technique is used to determine a relationship between the images through different features descriptors from the processed images. Feature extraction, Feature Matching, Outlier Elimination, Stitching and Blending are the five steps needed to generate panorama. In this work different feature descriptors such as SIFT, SURF, ASIFT and different outlier elimination techniques are used to create panorama. Rather than feature based technique, Correlation is also used to create panorama. Panoramas created by using different methods are evaluated based on the quality of the output mosaic and the time needed to generate the panorama

KEYWORDS: Stitching, Panorama, Feature extraction, Feature matching, Outlier Elimination, Blending.

I. INTRODUCTION

Image stitching is the process of stitching different photographic images with overlapping fields of view of the corresponding images to produce a segmented panorama or high-resolution Mosaic so it is also known as image mosaicing. The word panorama [1] means all round view. It derived from the Greek words 'pan' and 'horama'. where 'pan' means everything and 'horama' means to view. The most common method of image stitching require exact overlaps between consecutive images and identical exposures to produce seamless result or panorama. Rather than using image stitching in computer graphics and computer vision applications, there are some digital cameras apps can stitch their photos internally While this type of stitching has too many disadvantage such as orientation and transform problem.

The quality of output can be measured by the similarity of the stitched image to each of the input image in the data set, features extracted from the images and the taken for the execution. It can also be measured by the basis of visibility of the seam between the images in the panorama [2]. The applications of the image stitching is considered as the challenging field for image processing experts due to the development of the mobile apps that create panorama. It has huge number of applications in the field of video conferencing, video matting by cloning, video stabilization, video summarisation, 3D image reconstruction, satellite imaging, video compression, and several medical applications also. An interesting application of image stitching is the ability to summarize and compress videos taken with a panning camera that can reduce the size of the transmission.

For videos, the applications such as video compression and video indexing. While early approaches uses affine motion models, so construction of panorama is restricted to long focal lengths. Therefore Video stitching can be define as the generalization of multiple-image stitching. The presence of large amounts of independent motion in the scene, camera zoom parameters, and the desire to visualize dynamic events create additional challenges to image stitching. The important application of the video stitching created panorama is animating the cartoons. Video can also provide an interesting source of



Fig 1. Panorama Creation

II. MAIN STEPS IN STITCHING

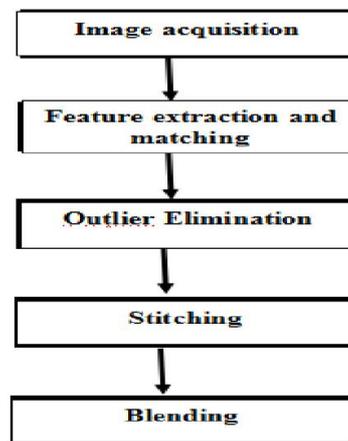


Fig 2. Main steps in Stitching

The first step in the creation of a panoramic image is to select the positions for the acquisition of images. In this step, a decision needs to be made on the type of resultant panoramic images. According to the required panoramic images, different image acquisition methods may be used to acquire the series of images. After the images have been acquired, some processing might need to be applied to the images before they can be stitched. Feature points are extracted from the processed image by using different feature extractors. Based on the feature points, the matching between the consecutive images are estimated by feature matching technique.

Outliers are the deviation from the original data. After feature matching, outliers from the images are removed by using Outlier elimination. The output of the outlier elimination is fed into stitching technique which is used to stitch the images together to create the mosaic. Blending is the technique used to create seamless panorama.

III. LITERATURE SURVEY

Image stitching is divided into two: Direct technique and feature based technique.

Direct techniques [3] minimize the sum of absolute differences between overlapping pixels or use any other available cost functions. They are not invariant to image scale and rotation. The main advantage of direct methods is that they make optimal use of the available information about image alignment. They measure the information that can be extracted from every pixel in the image. The disadvantage of direct techniques is they have only a limited range of convergence. Different

A. FEATURE EXTRACTION

It establish the correspondences between points, lines, edges, corners, or other geometric entities. Characteristics of robust detectors include invariant to image noise, scale invariant, translation invariant, and rotation transformations. Using the image features obtained from the corresponding images taken from different viewpoints, image stitching is achieved. The two ways to identify the matching region from the input images are Block matching and feature-point matching. Block

matching algorithms initially calculate the correlation between the regularized blocks generated in sequential consecutive images in the set. It can be done either by Normalized Cross-Correlation or by phase correlation using an Fast Fourier Transform. Such methods involve a series of complex calculations and also very sensitive to the slight difference between the images. Feature based methods [4] extract different features from each image and matches these features to establish a global correspondence between all the images. Feature descriptors are used to extract the features points from the given images for matching it with other images.

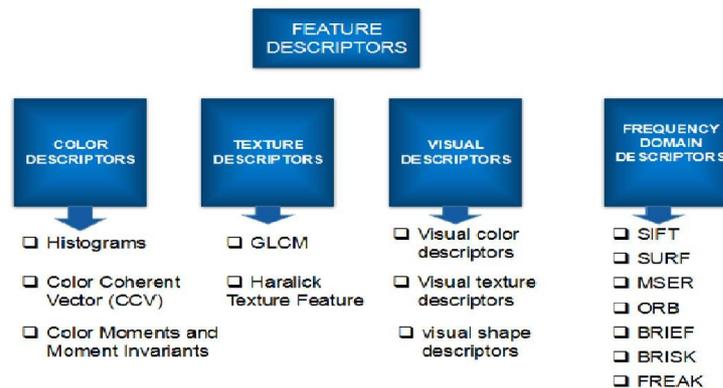


Fig 3: Different feature descriptors

A.1 COLOR DESCRIPTORS

Histogram is defined as the distribution of the number of pixels in the image. The number of elements in a histogram is related to the number of bits in each pixel of the same image.

Color Coherent Vector each histogram bin is partitioned into two types: coherent and incoherent. Coherent type contains pixel values of the images that belongs to a large informally colored region of the image. Otherwise it is incoherent type. Color histogram does not contain the spatial information of pixels thus similar color distribution for different images results. Color moments uses all the generalized color moments up to the second degree of the histogram bin and the first order results in nine combinations.

Color moment invariant can be generated from generalized color moment. It gives a total 24 color moment invariant.

A. 2 TEXTURE DESCRIPTORS

GLCM (gray level co-occurrence matrix) is used for the motion estimation of images, it can be used to extract second order statistical texture features from the images. The four features such as Angular second moment, correlation, Inverse difference moment and Entropy are computed in GLCM texture extraction.

Haralick Texture Feature captures the information about the patterns that created in patterns of texture of the image. Haralick features are calculated by using co-occurrence matrix, which is computationally expensive in real time application.

A. 3 VISUAL DESCRIPTORS

Visual color descriptors are used to extract Color features of the image that are not robust to the changes in the background colors and are independent of orientation and size of image. It can be used for distinguish still images and vide

Visual texture descriptors extract textures which contains the visual patterns of the set of images that have the properties of homogeneity of texture or not, that result from the presence of multiple colors or intensities in the image. It contains important structural information of surfaces of the objects in the image and their relationship to the surrounding environment.

Visual shape descriptors extracts the shape of image objects that provides a useful hint for similarity matching between the different objects. For image retrieval the shape descriptor wants to be invariant to scaling, rotation and translation transformations.

A. 4 FREQUENCY DOMAIN DESCRIPTORS

SIFT technique has four computational steps for extracting key points which is described in section 4.1. It reduces number of keypoints which helps in increasing the efficiency and also the robustness of the technique. Keypoints are rejected when they had a low contrast or if they were located on an edge.

SURF is a fast and robust algorithm used for local, similarity invariant representation and comparison of the feature points. Keypoints are found by using the approximation of the Hessian matrix of a given image point so it called Fast- Hessian

Department of Electrical & Electronics Engineering, College of Engineering Perumon, Kollam, Kerala – 691601, India

Detector. The main advantage of the SURF approach is its fast computation, provides real-time applications such as tracking and object recognition.

FAST algorithm is an interest point detector or feature extractor for the use in real time frame rate applications. The FAST detector compares pixels within a circle of fixed radius around a point.

ORB technique is a fast binary descriptor based on Binary Robust Independent Elementary Features key point descriptor. The binary based features have several advantages over the vector based features, it can be faster to compute, less store size, more efficient to compare the feature points and requires verylow memory.

Maximally Stable Extremal Regions (MSER) is used for the blob detection for the feature extraction. The algorithm for MSER is used to extracts co-variant regions from an image. Thus it provides a stable connected component of some gray level sets of the image.

BRIEF is a recent feature descriptor that uses simple binary tests between pixels in the smoothed image patch of the original image . Its performance is similar to SIFT detector in many ways, including robustness to lighting, blur, and the perspective distortion. Its disadvantage is very sensitive to in-plane rotation.

Fast Retina Key point (FREAK) finds a cascade of binary strings by efficiently comparing pixel intensities of image over a retinal sampling pattern of the image.

B. FEATURE MATCHING

Feature Matching [5] is used to match two images based on the features extracted from the image set. Relationship between the different images is calculated by Feature matching. SIFT and SURF matching techniques is used to find the match between the images.

Feature matching techniques are Normalized cross correlation, Randomized k-d trees and Best bin first

Randomized k-d trees is using multiple search-trees to enhance independence search efficiency. It will decline for high dimensional data and it will be not used for the efficiency is close to exhaustive search.

K nearest neighbor and k-d trees based on Nearest neighbour searching algorithm and it focus only on the node position which has been stored while ignoring inquired node position.

Best bin first algorithm based on the priority queue. Search start from the root of three subtree until a leaf node then find the first Euclidean nearest neighbour feature points. Unselected nodes are pushed into the priority queue, and next time the search will begin from the sub-tree which has the smallest key value. After estimating the features from the image.

C. OUTLIER ELIMINATION

After estimating the features from the image. The following steps are carried out:

- a. Identify the outliers where outliers are the deviation from the original data or the original pixels in the image.
- b. Eliminate the outliers by using different outlier elimination techniques.

Outlier elimination	Methods
RANSAC	"RANdom SAMple Consensus". It is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers . It is a non deterministic algorithm in the sense that it produces a reasonable result only with a certain probability, with this probability increasing as more iterations are allowed.
Least Square Method	Least Square means that the overall solution minimizes the sum of the squares of the errors made in the results of every single image. It only considers the residuals in the dependent pixels.
Multi- Model Fitting	The best fit in the least squares minimizes the sum of squared residuals, a residual being the difference between an observed value and the fitted value provided by the model. This will calculate recall and precision with better accuracy. No. of features will also be increased. this will increase the reliability and accuracy of stitched image

Table I. Different Outlier Elimination

D. STITCHING

Once the source pixels have been mapped into the final composite surface for the stitching, the next step is to blend them in order to create an attractive final panorama. If all of the images are in perfect registration and identically exposed, this is an easy problem for stitching the images. There are many blending methods used in image stitching. Weighted average is the simplest approach, in which the pixel values in the blended regions are based on the weighted average from the two overlapping images. Sometimes it doesn't work well. But if all the images were taken at the same time and using high quality tripods, then simple algorithm produces excellent results. Smoothing algorithm is used to blend the images to compensate for exposure differences and other mis-alignment. Smoothing algorithm which performs simply and fast. It is used to remove the seam occurring in between the images which produce better result.

Direct average method is a simplest approach, in which the pixel values in the blended regions are the average from the two overlapping images. It doesn't work well.

Gradating in and out amalgamation is based on the weight of the overlap region pixel is related to the distance between pixel and overlapping areas in the images. The pixel relationship between image mosaic and two images needed to stitch. Blending process is divided into Alpha blending and Poisson blending. Poisson blending is based on solving the Poisson equation using boundary information. Alpha blending is based on pixels with mixed colors, a measurement is needed to weight the influence of different elements.

E. PROJECTIVE LAYOUTS

Projective Layout	Methods
Rectilinear Projection	Lines that are straight in reality are shown as straight regardless of their directions on the image. Wide views around 120. It can exhibit severe distortion near the image borders.
Cylindrical Projection	Panoramas in this projection are meant to be viewed as though the image is wrapped into a cylinder and viewed from within. Cylindrical which have less distortion near the poles of the panosphere.
Spherical	Stitched image shows a 360 horizontal by 180 vertical field of view i.e. the whole sphere. Panoramas in this projection are meant to be viewed as though the image is wrapped into a sphere and viewed from within.
Panini	Specialized projections which may have more aesthetically pleasing advantages over normal cartography projections.
Stereographic	Stereographic projection or fisheye projection can be used to form a little planet panorama by pointing the virtual camera straight down and setting the field of view large enough to show the whole ground and some of the areas above it.

Projective layouts [6] are the different ways in which the image can be arrange in the mosaic or panorama

F. DIFFERENT TYPES OF PANORAMA

Brown and Lowe used the SIFT feature descriptor to implement a feature-based image stitching for creating the panorama. The first step is to extract using SIFT and match SIFT features between all of the images in set. Then, they use RANSAC outlier elimination to find a set of inliers that are compatible with a Homography between the images. After that, it applied to probabilistic model to verify the match between the images. They use bundle adjustment to solve for all of the camera parameters jointly to create the panorama. Finally they have applied the multi-band blending technique to combine the image to create final result.

Eden et al presented a technique to stitch different images by a two step graph cut approach to produce final result. The first step is to fix the positions of moving objects in the scene and the second step fill the entire dynamic range.

Lowe proposed a improvement of the SIFT based panorama shows how to solve for a photometric parameter, namely the overall gain between images, with an error function defined over all images. The error function is the sum of gain normalized intensity errors for all overlapping pixels.

Deepak Jain proposed a corner technique for image mosaicing. Here, he had used three step methods. In first step, he take two images and find corner of both these images. In second step, he removed the false corner from both images. Finally, he used Homography to find matched corner and get mosaic image.

Yanfang concerned on the problem of automatic image stitching which mainly applies to the image sequence even those including noise images. He used a method based on invariant features to realize fully automatic image stitching, in which it includes two main parts: image matching and image blending. In order to define and get the best possible stitching, several

formal cost functions for the evaluation of the stitching quality. Cost functions are used to find the similarity to the input images and the visibility of the seam in the basis of the gradient domain, minimizing the disturbing edges along the seam.

Vimal Singh Bind presented a technique for feature based image mosaicing using image fusion where the input images are stitched together using the popular stitching algorithms. To extract the best features m from the stitching results, the blending process is done by means of Discrete Wavelet Transform (DWT) using the maximum selection rule for both approximate as well as detail-components. The robustness and quality of the above mosaicing techniques are tested by means of three-dimensional rotational images.

Russol Abdelfatah presented a technique to implement image stitching by adopting feature-based alignment algorithm and blending algorithm to produce a high quality image. The images used to generate the panorama are captured in a fixed linear spatial interval. Feature extractio and image matching are based on Harris corner detectors method as the feature detection and neighboring pairs alignment using RANSAC algorithm. Linear blending is applied to remove the transition between the aligned images.

Guofeng Zhang introduced a novel image stitching technique, uses meshbased framework to optimize alignment and regularity in 2D. ASIFT is used for feature extraction and matching. Local homography for identifying the inliers and outliers. Eneury function is used for stitching the images.

There are various other techniques for creating the panorama. Each and every technique has its own merits and demerits.

IV. PROPOSED WORK

It is the comparison of different types of panorama creation methods by evaluating the quality of the output image and time needed to generate the output. Comparisons are Panorama based on SIFT [7], Panorama based on SURF, Panorama based on ASIFT [8] and local homography [9] and Panorama based on correlation.

A. PANORAMA BASED ON SIFT

First step of creating the panorama is extracting features from the images. Feature extraction is based on the SIFT descriptor. Scale Invariant Feature Transform (SIFT) consists of four stages as below:

1. Scale-space computation: It is the first stage of feature extraction which searches over all scales and image locations. It is implemented by using a difference-of-Gaussian function which is used to find potential interest points that are invariant to scale and orientation in the image.
2. Keypoint localization: At each key location, a detailed model is build to determine location and scale of the keypoint. Keypoints are selected based on measures of their stability.
3. Orientation assignment: One or more orientations are assigned to each keypoint in the image 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 the result will be invariant to the scale, orientation and location transformation.
4. Keypoint descriptor: The local image gradients of the keypoints are measured at the selected scale in the region around each keypoint in the image. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination of the image.

After estimating the features, feature matching is carried out on the basis of the SIFT matching technique. Eliminate the the outliers from the image by using the ransac method which is described above. Weighted average method is used to stitch the images together and alpha blending to used to remove the seam in between the images. The time needed to generate the output is 1.20 min.

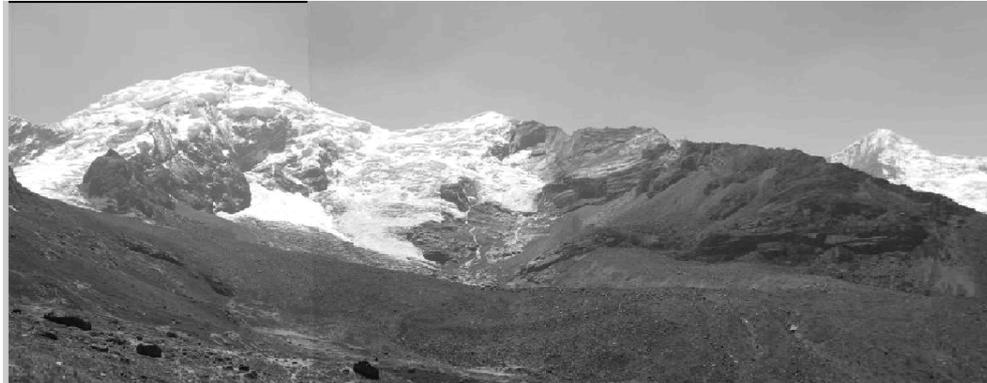


Fig 4. Mosaic based on SIFT

B. PANORAMA BASED ON SURF

Panorama based on SURF is similar to the SIFT, the difference is that here SURF is used as the feature descriptor and matching. A basic second order Hessian matrix approximation is used for feature point detection. The time needed to generate the output is 40 sec. The following steps are used in SURF:

1. In the construction of scale image pyramid in SURF algorithm, the scale space is divided into octaves, and there are 4 scale levels in each octave.
2. Each octave represents a series of filter response maps obtained by convolving the same input image with a filter of increasing size.
3. The minimum scale difference between subsequent scales depends on the length of the positive or negative lobes of the partial second order derivative in the direction of derivation.
4. Do nonmaximum suppression in a 3x3x3 neighbourhood to get the steady feature points and the scale of values.

a. Panorama based on ASIFT

ASIFT is similar to the SIFT. But more features can be extracted by using ASIFT more than SIFT. Feature extraction and matching is based on the ASIFT.

Affine Scale Invariant Feature Transform feature descriptors consists of following stages:

1. Rotate the images using Bilinear interpolation and apply anti-aliasing filter.
2. Tilt the images by subsampling the images in the vertical direction.
3. Calculate scale, rotation and translation-invariant features using SIFT.
4. Remove the keypoints close to boundary and normalize the image.

After extracting the features, ASIFT feature matching is used to find the matching between the images. The outlier elimination is based on local homography. Local homography is based on the homography of the image created. The steps of the local homography are:

1. Define the neighboring sets of the features.
2. Estimate the homography.
3. Inlier detection based on the threshold.
4. Different images are stitched together by Weighted average method and alpha blending to remove the seam in between the images. The time needed to generate the output is 1.30 min.

a. Panorama based on Correlation

There is no feature extraction in Panorama based on Correlation. It is based on the relation between the images. If any part of the images are same then based on the weighted average it is stitched together. The time taken for the creation of the panorama is 20sec. But the quality of the output produced is less. Comparing more than two is difficult in the correlation. It is variant to rotation, scaling and other transformation. But for small number of accurate preprocessed images, correlation is efficient to use because it takes only few amount of time to generate the output panorama with the desired image.

The output is generated by the execution of the matlab code in the MATLAB R2013a.



Fig 5. Mosaic based on Correlation

Various image stitching methods	Feature extraction	Feature matching	Outlier Elimination	Stitching	Time taken	Output quality
Panorama based on SIFT	SIFT	SIFT	Ransac	Weighted average	80sec	Medium
Panorama based on SURF	SURF	SURF	Ransac	Weighted average	40sec	Less
Panorama based on ASIFT	ASIFT	ASIFT	Local Homography	Weighted average	90sec	High
Panorama based on Correlation	Correlation	Correlation	No Outlier Elimination	Weighted average	20sec	Very low

Table III. Various Image stitching methods

V. CHALLENGES

- Very large number of images is needed for efficient creation of panorama
- Large amount of images may lead to high processing time, since each image needs some processing.
- Another recurring problem in creating panorama is the elimination of visible seams.

VI. CONCLUSION

Image stitching is the process of creating a high resolution panorama by combining the different images together. Images are overlapped with wide field of view to form a segmented panorama. It has a large amount of different algorithms for features detection and description. The choice of the feature detector depends on the problem. Feature descriptors consumes more time but produce high quality output. Correlation method takes only less time but the quality of the output is not good. Panorama based on ASIFT is the best method for creating panorama in the feature based techniques but for stitching two images it consumes around two minutes. When all the input images are preprocessed and the alignment of the images are corrected then Panorama based on the correlation can be used to generate the panorama.

REFERENCES

- [1] Ward, G. (2006). Hiding seams in high dynamic range panoramas. In R. W. Fleming, and S. Kim (Ed.), APGV. 153, p. 150. ACM.
- [2] Image stitching using matlab Shweta Mishra ,Poorva Chaudhari , Shalaka Khandale nformation Tech. Department of computer science. [3] A Detailed Survey on Various Image Stitching Techniques; Mittali et al., International Journal of Computer and IT.
- [4] A Survey on Image Feature Descriptors, Rekhil M Kumar, Dept. Of Computer And Information Science College Of Engineering, Poonjar
- [5] Faraj Alhwarin, C. W. (2008). Improved SIFT-Features Matching for Object Recognition. In E. Gelenbe, S. Abramsky, and V. Sassone (Ed.), BCS Int. Acad. Conf. (pp. 178-190). British Computer Society.
- [6] Comprehensive Study And Review Of Image Mosaicing Methods; Mrs. Hetal M. Patel, Asst. Prof. Pinal. J. Patel , Asst. Prof. Mr. Sandip G. Patel; IJCERT
- [7] Yanfang Li, Y. W. (2008). Automatic Image Stitching Using SIFT.
- [8] G. Yu and J.-M. Morel, ASIFT: An Algorithm for Fully Affine Invariant Comparison, Image Processing On Line, vol. 1, 2011.
- [9] Multi-Viewpoint Panorama Construction with Wide-Baseline Images Guofeng Zhang, Member, IEEE, YiHe, Weifeng Chen, Jiaya Jia, IEEE transaction on image processing 2015.