

# Image Stitching based on Feature Extraction Techniques: A Survey

Ebtsam Adel  
Information System Dept.  
Faculty of Computers and  
Information,  
Mansoura University, Egypt

Mohammed Elmogy  
Information Technology Dept.  
Faculty of Computers and  
Information,  
Mansoura University, Egypt

Hazem Elbakry  
Information System Dept  
Faculty of Computers and  
Information,  
Mansoura University, Egypt

## ABSTRACT

Image stitching (Mosaicing) is considered as an active research area in computer vision and computer graphics. Image stitching is concerned with combining two or more images of the same scene into one high resolution image which is called panoramic image. Image stitching techniques can be categorized into two general approaches: direct and feature based techniques. Direct techniques compare all the pixel intensities of the images with each other, whereas feature based techniques aim to determine a relationship between the images through distinct features extracted from the processed images. The last approach has the advantage of being more robust against scene movement, faster, and has the ability to automatically discover the overlapping relationships among an unordered set of images. The purpose of this paper is to present a survey about the feature based image stitching. The main components of image stitching will be described. A framework of a complete image stitching system based on feature based approaches will be introduced. Finally, the current challenges of image stitching will be discussed.

## Keywords

Image stitching/mosaicing, panoramic image, features based detection, SIFT, SURF, image blending.

## 1. INTRODUCTION

Image/photo stitching is the process of combining multiple photographic images with overlapping fields of view to produce a segmented panorama or high-resolution image. It is also known as image mosaicing. Most common approaches of image stitching require exact overlaps between images and identical exposures to produce seamless results. In addition of using image stitching in computer vision and computer graphics applications, there are some digital cameras can stitch their photos internally [1].

On the other hand, the human visual system has a field of view of around 135 x 200 degrees, but a typical camera has a field of view of only 35 x 50 degrees. Therefore, panoramic image mosaicing works by taking lots of pictures from an ordinary camera and stitching them together to form a composite image with a much larger field of view [2]. The quality of image stitching is measured by the similarity of the stitched image to each of the input images. It also can be measured by the visibility of the seam between the stitched images [3].

The use of image stitching in real time applications is considered as a challenging field for image processing experts. It has wide applications in the field of video conferencing, video matting, video stabilization, 3D image reconstruction, video summarization, video compression, satellite imaging, and several medical applications. An

interesting application of image stitching is the ability to summarize and compress videos taken with a panning camera.

For videos, the applications of image stitching were extended to additional applications, such as video compression and video indexing. While these early approaches used affine motion models, they were therefore restricted to long focal lengths. Video stitching is in many ways a straightforward generalization of multiple-image stitching. The potential presence of large amounts of independent motion, camera zoom, and the desire to visualize dynamic events impose additional challenges to image stitching [3].

Videos can also be used to create animated panoramic video textures. In which, different portions of a panoramic scene are animated with independently moving video loops, or to shine “video flashlights” onto a composite mosaic of a scene. Video can also provide an interesting source of content for creating panoramas taken from moving cameras.

Medical image stitching has many applications in clinical diagnosis, such as diagnosis of cardiac, retinal, pelvic, renal, abdomen, liver, tissue and other disorders [38]. Advances in computer science have led to reliable and efficient image processing methods useful in medical diagnosis, treatment planning, and medical research. In clinical diagnosis, integration of useful data obtained from separate images is often desired.

The images need to be geometrically aligned for better observation. This procedure of mapping points from one image to corresponding points in another image is called image registration [38]. The reference and the referred images could be different because they were taken at different times. We can use different devices and the images can be captured from different angles in order to have 2D or 3D perspective (multi temporal) [4].

Another important application of panoramic image stitching is the localization systems which had presented by Arth et al [5]. They proposed a highly accurate real-time outdoor localization system that can work on current mobile devices. The most important characteristics of their approach are its high degree of accuracy and its low computational demands that make it suitable to run on off-the-shelf mobile phones. They managed to overcome the problem of narrow field of view (FOV) of current cameras on mobile devices by employing a technique for image capturing. Their approach can be used intuitively and in a very straight forward way, as the user is simply asked to capture the environment as a visually pleasing panoramic snapshot.

Image stitching applications also include video summarization. A large number of cameras record video around the clock, producing huge volumes. Processing these

huge chunks of videos demands plenty of resources like time, man power, and hardware storage etc. Video summarization plays an important role in this context. It helps in efficient storage, quick browsing, and retrieval of large collection of video data without losing important aspects [6].

The remainder of the paper is organized as follows. An overview of the main components of the image stitching is elucidated in section 2. Section 3 introduces the general approaches of the image stitching. Section 4 discusses a complete model of the image stitching. Section 5 contains an overview of the current effort in image stitching field. In Section 6, some challenges of image stitching will be discussed. Finally, section 7 presents the conclusion and the future work.

## 2. THE MAIN COMPONENTS OF IMAGE STITCHING

The image stitching can be divided into three main components: calibration, image registration, and blending [8], as shown in Fig. 1. The goal of camera calibration is to produce an estimate of the extrinsic and intrinsic camera parameters. During the image registration, multi-images are compared to find the translations that can be used for the alignment of images. After registration, these images are merged (blended) together to form a single image. In the following subsections, these main components will be discussed briefly.

### 2.1. Calibration

Image calibration aims to minimize differences between an ideal lens model and the camera-lens combination that was used. These differences are resulted from optical defects such as distortions and exposure differences between images [8]. Intrinsic and extrinsic camera parameters are recovered in order to reconstruct the 3D structure of a scene from the pixel coordinates of its image points. Extrinsic camera parameters define the location and orientation of the camera reference frame with respect to a known world reference frame. Intrinsic camera parameters link the pixel coordinates of an image point with the corresponding coordinates in the camera reference frame [9].

### 2.2. Registration

Image registration is the core of a mosaicing procedure. Its purpose is to create geometric correspondence between images. Therefore, we can compare images and apply other steps appropriately [20]. Image registration is defined as the process of aligning two or more images which are captured from different point of perspectives. Image blending is processed to make the transition from one image to another image smoother. So, the joint between two images can be removed.

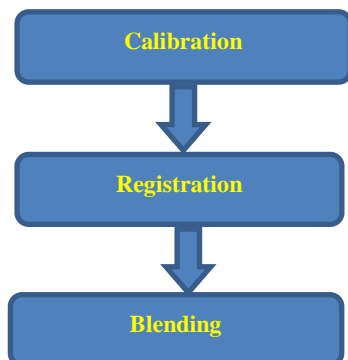


Fig. 1: The main components of image stitching

## 2.3. Blending

Blending is applied across the stitch so that the stitching would be seamless. There are two popular ways of blending the images [10]. One is called alpha “feathering” blending, which takes weighted average of two images. The cases that alpha blending works extremely well is when image pixels are well aligned to each other and the only difference between two images is the overall intensity shift. Another popular approach is Gaussian pyramid [20]. This method essentially merges the images at different frequency bands and filters them accordingly. The lower the frequency band, the more it blurs the boundary. Gaussian pyramid blurs the boundary while preserving the pixels away from the boundary.

## 3. IMAGE STITCHING APPROACHES

The direct and feature based techniques are considered as the main approaches for image stitching. The direct techniques work by directly minimizing pixel to pixel dissimilarities. Otherwise, the feature based techniques work by extracting a sparse set of features and then matching these to each other [11].

### 3.1. Direct Techniques

The direct technique depends on comparing all the pixel intensities of the images with each other. Direct techniques minimize the sum of absolute differences between overlapping pixels or use any other available cost functions. These methods are computationally complex as they compare each pixel window to others. They are not invariant to image scale and rotation. There are many techniques for solving image mosaicing problems using direct methods such as Fourier analysis techniques [12] and also Simon Baker have proposed a unifying framework to fine optimization of cost or objective functions [39].

The main advantage of direct methods is that they make optimal use of the information available in image alignment. They measure the contribution of every pixel in the image. However, the biggest disadvantage of direct techniques is that they have a limited range of convergence [11].

### 3.2. Feature based Techniques

The simplest way to find all corresponding feature points in an image pair is to compare all features in one image against all features in the other using one of the local descriptors. Unfortunately, this is quadratic in the expected number of features, which makes it impractical for some applications [11]. For image stitching based on feature based techniques, feature extraction, registration, and blending are different steps required for doing image stitching.

Feature based methods begin by establishing correspondences between points, lines, edges, corners, or other geometric entities. Characteristics of robust detectors include invariance to image noise, scale invariance, translation invariance, and rotation transformations. There are many feature detector techniques, such as Harris [18], SIFT [13], SURF [14], FAST [15], PCA-SIFT [16] and ORB [19] techniques.

The well-known SIFT (Scale-Invariant Feature Transform) [13] technique is very robust, but the computation time is not feasible for real-time applications. Harris corner detector [18] is used to detect the features. It uses a normalized cross-correlation of intensity values to match them. Harris corner is not invariant to scale changes and cross correlation. Whereas, SURF (Speeded Up Robust Features) [14] improves the computation time of SIFT by using an integral image for fast local gradient computations on an image. Recently, binary feature descriptors have received more attention. These

descriptors are described with a binary string such as ORB [19] technique which is trying to keep track of the popularity with SIFT and SURF. It's extremely fast operation, while sacrificing very little on performance accuracy. ORB is scale and rotation invariant, robust to noise and affine transformations. The ORB algorithm is actually a combination of the FAST (Features from Accelerated Segment Test) keypoint detection, and the BRIEF (Binary Robust Independent Elementary Features) keypoint descriptor algorithm modified to handle oriented keypoints. Feature based methods have the advantage of being more robust against scene movement. They are potentially faster and they have the ability to recognize panoramas by automatically discover the adjacency (overlap) relationships among an unordered set of images. These features make them ideally suited for fully automated stitching of panoramas taken by casual users [3].

#### 4. IMAGE STITCHING MODEL BASED ON FEATURES BASED TECHNIQUES

In this section, a complete image stitching model based on feature based techniques will be discussed. As shown in Fig. 2, the image stitching model consists of five stages: images acquisition, features detection and matching, RANSAC estimation, global alignment, and image blending. In the following subsections, the main stages of feature based image stitching will be described in detail.

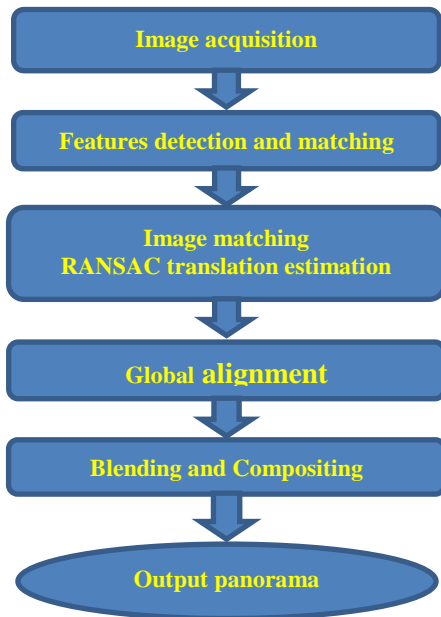


Fig. 2: The block diagram of general panoramic image stitching model based feature based approaches

##### A. Image Acquisition

The first stage of any vision system is the image acquisition stage. Image acquisition can be broadly defined as the action of retrieving an image from some sources. Typically, images can be acquired for panoramic imaging by three different methods, as shown in Fig. 3. These methods are translating a camera parallel to the scene, rotating a camera about its vertical axis by keeping optical center fixed, or by a handheld camera [20].

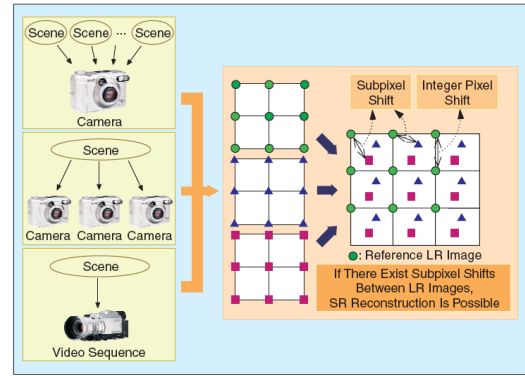


Fig. 3: The different image acquisition methods [21]

##### B. Features Detection and Matching

The second step in image stitching process is the features detection which is considered as the main image stitching stage. Features can be defined as the elements in the two or more input images to be matched. It relies on the idea that instead of looking at the image as a whole, it could be advantageous to select some special points in the image and perform a local analysis on these ones.

Feature detection forms an important part of many computer vision algorithms. Online image processing algorithms need real-time performance. Thus the speed at which features are detected is crucial in many applications, such as visual SLAM (Simultaneous localization and mapping), image registration [4], 3D reconstruction, and video stabilization which are needed to match corresponding image features between multiple views. The detected corners or feature points need to be described unambiguously so that the correspondence between multiple views can be computed reliably. Real-time processing requires the feature detection, description, and matching to be as fast as possible.

To provide a better feature matching for image pairs, corners are matched to give quantitative measurement. Corners are good features to match. The features of corners are more stable features over changes of viewpoint. The other most important feature of corner is that if there is a corner in an image than its neighborhood will show an abrupt change in intensity.

On the other hand, local feature descriptors describe a pixel (or a position) in an image through its local content. They are supposed to be robust to small deformations or localization errors, and give us the possibility to find the corresponding pixel locations in images which capture the same amount of information about the spatial intensity patterns under different conditions [24].

There are many requirements of a local feature detector, such as it should invariant to translation, rotation, scale, affine transformation, presence of noise, and blur. It must be robust to occlusion, clutter, and illumination changes. It should also be repetitive. Finally, there should be enough points to represent the image with time efficient.

There are many features descriptors such as SIFT [22], SURF [14], HOG [17], GLOH [23], PCA-SIFT [16], Pyramidal HOG (PHOG), and Pyramidal Histogram Of visual Words (PHOW).

In the following subsections, we briefly explain the SIFT, SURF, and ORB descriptors which offer scale and rotation invariant properties.

## 1. Scale Invariant Feature Transform (SIFT) Technique

The SIFT technique is one of the most robust and widely used image matching algorithm based on local features. It ensures a good mosaic image and a reliable result. SIFT is a feature detection and description technique. SIFT produces key point descriptors which describes the image features [22].

SIFT technique has four computational steps for extracting key points: scale-space peak selection, key-point localization, orientation assignment, and defining key-point descriptors.

To each image, it builds image pyramid by generating progressively blurred out images and it subtracts neighbor images to get the difference of Gaussian (DOG) pyramid. Then, it detects the extreme for DOG pyramid. The number of keypoints was reduced to help in increasing efficiency and also the robustness of the technique. Keypoints are rejected if they had a low contrast or if they were located on an edge. The following step is orientation assignment which uses orientation histogram to statistics the gradient orientation with sampling the center neighborhood of the key points. The last step is to describe the key points.

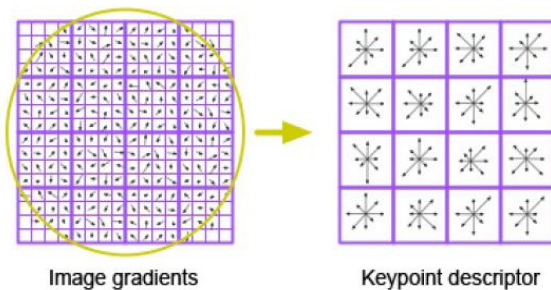


Fig. 4: The SIFT descriptor generation [2].

A keypoint descriptor is created by first computing the gradient magnitude and orientation at each image sample point in a region around the keypoint location, as shown in Fig. 4. These are weighted by a Gaussian window which indicated by the overlaid circle. These samples are then accumulated into orientation histograms summarizing the contents over 4x4 sub regions. The length of each arrow corresponding to the sum of the gradient magnitudes near that direction within the region.

SIFT technique suffer from some disadvantages. First, it is based on histogram which expresses grads variations in every scale and feature direction. Second, the complex of time is very difficult [35].

## 2. Speeded Up Robust Features (SURF) Technique

SURF is a fast and robust algorithm developed by Bay [14] for local, similarity invariant representation and comparison. The SURF approach can be divided into three main steps. First, keypoints are selected at distinctive locations in the image, such as corners, blobs, and T-junctions. Next, the neighborhood of every keypoint is represented by a feature vector. This descriptor has to be distinctive. At the same time, it should be robust to noise, detection errors, and geometric and photometric deformations. Finally, the descriptor vectors are matched among the different images. Keypoints are found by using a so-called Fast-Hessian Detector that is based on the approximation of the Hessian matrix of a given image point. The responses to Haar wavelets are used for orientation assignment before the keypoint descriptor is formed from the

wavelet responses in a certain surrounding to the keypoint. Therefore, the SURF constructs a circular region around the detected keypoints.

The main advantage of the SURF approach lies in its fast computation, enabling real-time applications such as tracking and object recognition. It speeded-up the SIFT's detection process by keeping in view of the quality of the detected points. It gives more focus on speeding-up the matching step. The Hessian matrix is used along with descriptors low dimensionality to significantly increase the matching speed [24], but it is poor at handling viewpoint and illumination changes.

## 3. Features from Accelerated Segment Test (FAST) Technique

The FAST technique introduced by *Rosten Drummond* [15] for identifying interest points in an image. The reason behind the work of FAST algorithm was to develop an interest point detector for use in real time frame rate applications.

The FAST detector compares pixels only on a circle of fixed radius around a point. A point is classified as a corner only if one can find a large set of pixels on a circle of fixed radius around the point are all significantly brighter or darker than the central point. On the other hand the major limitation of FAST detector is that multiple features are detected adjacent to one another [15].

As shown in Fig. 5, the FAST algorithm considers a circle of 16 pixels around the corner candidate  $p$ . An interest point is indicated when a set of  $n$  contiguous pixels in the circle are all brighter than the candidate pixel  $I_p$  plus a threshold  $T$ , or all darker than  $I_p \leq t$ .

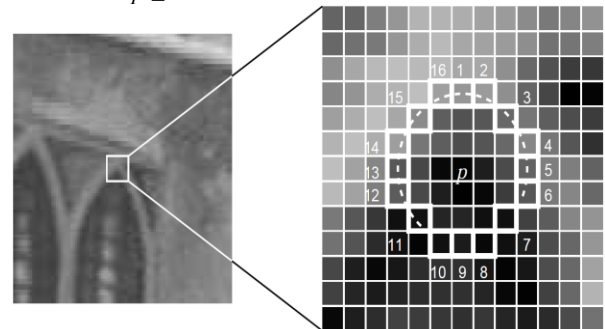


Fig 5: The corner detection by using FAST algorithm [15].

In the first step, a center pixel  $p$  is labeled as “corner” if there exist at least  $n$  consecutive “circle pixels”. If at least three of the four pixel values -  $I_1, I_5, I_9, I_{13}$  are not above or below  $I_p + T$ , then  $p$  is not an interest point (corner). In this case, reject the pixel  $p$  as a possible interest point. If at least three of the pixels are above or below  $I_p + T$ , then check for all 16 pixels and check if 12 contiguous pixels fall in the criterion. Repeat the procedure for all the pixels in the image [15].

## 4. Oriented FAST and Rotated BRIEF (ORB) Technique

ORB technique is a very fast binary descriptor based on Binary Robust Independent Elementary Features (BRIEF) [19] key point descriptor. The binary based features have several advantages over vector based features. They can be faster to compute, more compact to store, more efficient to compare, and requires very low memory. It can be demonstrated through experiments how ORB is at two orders of magnitude faster than SIFT, while performing as well in many situations. The efficiency is tested on several real-world

applications, including object detection and patch-tracking on a smart phone.

The closest system to ORB proposes a multi-scale Harris keypoint and oriented patch descriptor. ORB builds on the well-known FAST keypoint detector and the recently developed BRIEF descriptor. Both these techniques are attractive because of their good performance and low cost. ORB includes the addition of a fast and accurate orientation component. ORB also uses learning method for de-correlating BRIEF features under rotational invariance, leading to better performance in nearest-neighbor applications [19].

### C. Homography Using RANSAC

After we have the information of feature matching of all pictures, we can use this useful information to do image matching. In image matching step, we are going to find out which picture is a neighbor of another picture, and find the correctly feature matching set we need for next step of all feature matching set [25].

RANSAC (RANDOM SAMPLE Consensus) is a non-deterministic algorithm, because it doesn't ensure to return acceptable results. It is used to estimate parameters for Homography of a mathematical model from a set of observed data which contains outliers iteratively.

RANSAC loop involves selecting four feature pairs (at random); compute Homography  $H$  (exact); compute inliers, keep largest set of inliers, and finally it re-compute least-squares  $H$  estimate on all of the inliers.

### D. Global Alignment

The most relevant technique is bundle adjustment, which is a photogrammetric technique to combine multiple images of the same scene into an accurate 3D reconstruction. The aim of this step is to find a globally consistent set of alignment parameters that minimize the mis-registration between all pairs of images. Initial estimates of the 3D location of features in the scene must first be computed, as well as estimates of the camera locations. Then, bundle adjustment applies an iterative algorithm to compute optimal values for the 3D reconstruction of the scene and camera positions, by minimizing the log-likelihood of the overall feature projection errors using a least-squares algorithm [26].

In order to do this, we need to extend the pairwise matching criteria to a global energy function that involves all of the per-image pose parameters. Once we have computed the global alignment, we need to perform local adjustments such as parallax removal to reduce double images and blurring due to local mis-registration. Finally, if we are given an unordered set of images to register, we need to discover which images go together to form one or more panoramas [3].

### E. Compositing

Once we have registered all of the input images with respect to each other, we need to decide how to produce the final stitched (mosaic) image. This involves selecting a final compositing surface, e.g., flat, cylindrical. Finally, we must decide how to blend them in order to create an attractive looking panorama.

The first step to be made is how to represent the final image. If only a few images are stitched together, a natural approach is to select one of the images as the reference and to then warp all of the other images into the reference coordinate system. The resulting composite is sometimes called a flat panorama. Since the projection onto the final surface is still a perspective projection, hence straight lines remain straight (which is often a desirable attribute).

There are many different projective layouts on which image mosaicing can be used, such as rectilinear projection, where

the stitched image is viewed on a two-dimensional plane intersecting the panosphere in a single point. Lines that are straight in reality are shown as straight regardless of their directions on the image. Wide views - around  $120^\circ$  or so - start to exhibit severe distortion near the image borders. One case of rectilinear projection is the use of cube faces with cubic mapping for panorama viewing. Panorama is mapped to six squares. Each cube face showing 90 by 90 degree area of the panorama [20]. It also shows the cylindrical projection where the stitched image shows a  $360^\circ$  horizontal field of view and a limited vertical field of view. Panoramas in this projection are meant to be viewed as though the image is wrapped into a cylinder and viewed from within.



Fig. 6: The cylinder mosaicing projection.

To build a cylindrical panorama, a sequence of images is taken by a camera mounted on a leveled tripod. If the camera focal length or field of view is known, each perspective image can be warped into cylindrical coordinates. Two types of cylindrical warping are forward warping and inverse warping. In forward warping, the source image is mapped onto cylindrical surface, but it can have holes in the destination image (because some pixels may never get mapped there). Therefore, we use inverse mapping where each pixel in the destination image is mapped to the source image. Since the mapping is unlikely to be exactly on the pixel values, bilinear interpolation is used to calculate the colors at the destination pixels [27].

Once the source pixels have been mapped onto the final composite surface, the second step is to blend them in order to create an attractive looking panorama. If all of the images are in perfect registration and identically exposed, this is an easy problem (any pixel combination will do). There are many different pixels blending methods used in image stitching, such as feathering image blending, gradient domain and Image Pyramid blending [11].

Featuring image blending is a technique used in computer graphics software to smooth or blur the edges of a feature; it is the simplest approach, in which the pixel values in the blended regions are weighted average from the two overlapping images. Sometimes this simple approach doesn't work well (for example in the presence of exposure differences). But if all the images were taken at the same time and using high quality tripods, therefore, this simple algorithm produces excellent results. An alternative approach to multi-band image blending is to perform the operations in the *gradient domain*. Here, instead of working with the initial color values, the image gradients from each source image are copied; in a second pass, an image that best matches these gradients is reconstructed. Copying gradients directly from the source images after seam placement is just one approach to gradient domain blending. Another important approach of image blending is *Image Pyramid* blending; the image

pyramid is actually a representation of the image by a set of the different frequency-band images (i.e. Hierarchical representation of an image at different resolution). Image pyramid provides many useful properties for many applications, such as noise reduction, image analysis, image enhancement, etc.



**Fig. 7: The pixel blending process.**

Laplacian pyramid is an algorithm using Gaussian to blend the image while keeping the significant feature in the meantime. It downsizes the image into different levels (sizes) with Gaussian. Later, it expands the Gaussian in to the lower lever and subtracts from the image in that lever to acquire the Laplacian image. This Laplacian Pyramid is the true useful member of the image pyramid. Each layer of this pyramid is the band-pass image. We can now do some things to the specific frequency just like in the frequency domain. We also see that even after its frequencies are shown, the local features of the image are still there [36].

## 5. LITERATURE REVIEW OF THE FEATURE-BASED IMAGE STITCHING

In the last two decades, there are many researchers implemented and proposed some image stitching systems. For example, Levin and Weiss [29] introduced several formal cost functions for the evaluation of the quality of stitching. Their approach is demonstrated in various applications, including generation of panoramic images, object blending, and removal of compression artifacts. The aim of a stitching algorithm is to produce a visually plausible mosaic with two desirable properties. First, the mosaic should be as similar as possible to the input images, both geometrically and photometrically. Second, the seam between the stitched images should be invisible. While these requirements are widely acceptable for visual examination of a stitching result, their definition as quality criteria was either limited or implicit in previous approaches. Authors presented several cost functions for these requirements, and define the mosaic image as their optimum. The stitching quality in the seam region is measured in the gradient domain. The mosaic image should contain a minimal amount of seam artifacts.

Brown and Lowe [32] used the SIFT algorithm to implement a feature-based image stitching system. The first step in the panoramic recognition algorithm is to extract and match SIFT features between all of the images. SIFT features are located at scale-space maxima/minima of a difference of Gaussian function, and then the objective of second step “image matching” is to find all matching (overlapping) images, it is only necessary to match each image to a small number of neighboring images in order to get a good solution for the image geometry. Then, they used RANSAC to select a set of inliers that are compatible with a Homography between the images. After that, they applied a probabilistic model to verify

the match; then they used bundle adjustment to solve for all of the camera parameters jointly; then they used bundle adjustment to solve for all of the camera parameters jointly; finally they have applied the multi-band blending strategy. The idea behind multi-band blending is to blend low frequencies over a large spatial range and high frequencies over a short range. This can be performed over multiple frequency bands using a Laplacian Pyramid.

Eden et al. [33] presented a technique to automatically stitch multiple images at varying orientations and exposures to create a composite panorama that preserves the angular extent and dynamic range of the inputs. The proposed method allows for large exposure differences, large scene motion or other mis-registrations between frames and requires no extra camera hardware. To do this, they introduced a two-step graph cut approach. The purpose of the first step is to fix the positions of moving objects in the scene. In the second step, they fill in the entire available dynamic range.

Lowe [28] extended his previous by introducing gain compensation and automatic straightening steps. They showed 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 [30] 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 [13] 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. As the noises images have large differences between the other images, when using SIFT features to realize correct and robust matching, it supplies a probabilistic model to verify the panorama image sequence. The stitching quality is measured visually by the similarity of the stitched image to each of the input images, and by the visibility of the seam between the stitched images. In order to define and get the best possible stitching, several formal cost functions for the evaluation of the stitching quality are introduced in this paper. In these cost functions the similarity to the input images and the visibility of the seam are defined in the gradient domain, minimizing the disturbing edges along the seam. A good image stitching will optimize these cost functions, overcoming both photometric inconsistencies and geometric misalignments between the stitched images.

Vimal Singh Bind [31] 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. The performance evaluation of proposed technique is done in terms of PSNR (peak signal-to-noise ratio), FSIM as Quality Measure for Combined similarity, MI (Mutual Information), EME (Enhancement performance measure), NAE (Normalized Absolute Error) and SD.

Russol Abdelfatah [34] presented a technique to implement image stitching by adopting feature-based alignment algorithm and blending algorithm to produce a high quality

image. The used images to create panorama are captured in a fixed linear spatial interval. The processing method involves feature extraction, image matching based on Harris corner detectors method as the feature detection and neighboring pairs alignment using RANSAC (RANdom Sample Consensus) algorithm. Linear blending is applied to remove the transition between the aligned images.

## 6. CHALLENGES OF IMAGE STITCHING

There are many challenges in image stitching such as:

- Noisy image data or data with uncertainties: An image is often corrupted by noise in its acquisition and transmission, the cost of extracting features is minimized by taking a cascade filtering approach.
- Very larger images collection need for efficient indexing: large amount of images may lead to high processing time, since each image needs some processing.
- The main challenge on image stitching is the using of handled camera which may lead to presence of parallax (a shift in apparent object position while observed from different angles of view), small scene motions such as waving tree branches, and large-scale scene motions such as people moving in and out of pictures. This problem can be handled by bundle adjustment.
- Another recurring problem in creating photo-mosaics is the elimination of visible seams, for which a variety of techniques have been developed over the years [3].

## 7. CONCLUSION

Image mosaicing is considered as an active research area in the fields of computer vision and computer graphics. It has a large amount of different algorithms for features detection and description. The choice of the feature detector depends on the problem. In this paper, we have offered a comprehensive study on features-based image stitching such as SIFT algorithm which is rotation, scale invariant as well as more effective in presence of noise. It has highly distinctive features. However, it needs high computational time; the SURF algorithm proves superior in terms of execution time and illumination invariance property; The ORB algorithm is rotation and scale invariant with improved execution time but its performance is poor in presence of noise. In future we want to compare between the algorithms we have studied and other feature based image stitching algorithms, Also Stitching videos together to create dynamic panoramas, and stitching videos and images in the presence of large amounts of parallax.

## 8. REFERENCES

- [1] Ward, G. (2006). Hiding seams in high dynamic range panoramas. In R. W. Fleming, & S. Kim (Ed.), *APGV*. 153, p. 150. ACM.
- [2] Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60, 91-110.
- [3] Szeliski, R. (2010). *Computer Vision: Algorithms and Applications* (1st Ed.). New York, NY, USA: Springer-Verlag New York, Inc.
- [4] Brown, L. G. (1992, Dec). A Survey of Image Registration Techniques. *ACM Comput. Surv.*, 24(4), 325-376.
- [5] Arth, C., Klopschitz, M., Reitmayr, G., & Schmalstieg, D. (2011). Real-time self-localization from panoramic images on mobile devices. *ISMAR* (pp. 37-46). IEEE.
- [6] Ajmal, M., Ashraf, M., Shakir, M., Abbas, Y., & Shah, F. (2012). Video Summarization: Techniques and Classification. In L. Bolc, R. Tadeusiewicz, L. Chmielewski, & K. Wojciechowski (Eds.), *Computer Vision and Graphics* (Vol. 7594, pp. 1-13). Springer Berlin Heidelberg.
- [7] WIKIPEDIA.(n.d.).[http://en.wikipedia.org/wiki/File:Rochester\\_NY.jpg](http://en.wikipedia.org/wiki/File:Rochester_NY.jpg).  
[http://en.wikipedia.org/wiki/File:Rochester\\_NY.jpg](http://en.wikipedia.org/wiki/File:Rochester_NY.jpg).
- [8] K.Shashank, N. G. (MarCh 2014). A Survey and Review over Image Alignment and Stitching Methods. *The International Journal of Electronics & Communication Technology* (IJECT), ISSN: 2230-7109 (Online).
- [9] Zhang, Z. (2000, Nov). A Flexible New Technique for Camera Calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(11), 1330-1334.
- [10] Deng, Y., & Zhang, T. (September 07, 2003). Generating Panorama Photos.
- [11] Szeliski, R. (December 10, 2006). Image Alignment and Stitching. Tech. rep.
- [12] Bergen, J. R., Anandan, P., Hanna, K. J., & Hingorani, R. (1992). Hierarchical Model-Based Motion Estimation. *Proceedings of the Second European Conference on Computer Vision* (pp. 237-252). London, UK, UK: Springer-Verlag.
- [13] Yanfang Li, Y. W. (2008). Automatic Image Stitching Using SIFT.
- [14] Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008, Jun). Speeded-Up Robust Features (SURF). *Comput. Vis. Image Underst.*, 110(3), 346-359.
- [15] Rosten, E., & Drummond, T. (2006). Machine Learning for High-speed Corner Detection. *Proceedings of the 9th European Conference on Computer Vision - Volume Part I* (pp. 430-443). Berlin, Heidelberg: Springer-Verlag.
- [16] Ke, Y., & Sukthankar, R. (2004). PCA-SIFT: A more distinctive representation for local image descriptors. (pp. 506-513).
- [17] Dalal, N., & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01* (pp. 886-893). Washington, DC, USA: IEEE Computer Society.
- [18] Harris, C., & Stephens, M. (1988). A combined corner and edge detector. In *Proc. of Fourth Alvey Vision Conference*, (pp. 147-151).
- [19] A.V.Kulkarni, J. V. (September-2013). Object recognition with ORB and its Implementation on FPGA. *International Journal of Advanced Computer Research*, 2277-7970.
- [20] Mrs. Hetal M. Patel, A. P. (November- 2012). Comprehensive Study and Review of Image Mosaicing Methods. *International Journal of Engineering Research & Technology* (IJERT), Vol. 1 Issue 9, ISSN: 2278-0181.

- [21] Park, S. C., Park, M. K., & Kang, M. G. (2003, may). Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, 20(3), 21-36.
- [22] Faraj Alhwarin, C. W. (2008). Improved SIFT-Features Matching for Object Recognition. In E. Gelenbe, S. Abramsky, & V. Sassone (Ed.), *BCS Int. Acad. Conf.* (pp. 178-190). British Computer Society.
- [23] Mikolajczyk, K., & Schmid, C. (2005, Oct). A Performance Evaluation of Local Descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(10), 1615-1630.
- [24] Oyallon, E. (February 25, 2013). An analysis and implementation of the SURF method, and its comparison to SIFT.
- [25] Fischler, M. A., & Bolles, R. C. (June 1981). "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography".
- [26] Mclauchlan, P. F., Jaenicke, A., & Xh, G. G. (2000). Image mosaicing using sequential bundle adjustment. In *Proc. BMVC*, (pp. 751-759).
- [27] Mrs. Hetal M.Patel, A. P. (May - 2013). Panoramic Image Mosaicing. *International Journal of Engineering Research & Technology*, 2278-0181.
- [28] Brown, M., & Lowe, D. G. (2007, Aug). Automatic Panoramic Image Stitching Using Invariant Features. *Int. J. Comput. Vision*, 74(1), 59-73.
- [29] Anat Levin, A. Z., & Weiss, Y. (2000). Seamless Image Stitching in the Gradient Domain. The Hebrew University of Jerusalem.
- [30] Deepak Jain, G. S. (2012). Image Mosaicing using corner techniques. *International Journal of Engineering Research & Technology (IJERT)*.
- [31] Vimal Singh Bind, P. R. (2013). A Robust Technique for Feature-based Image Mosaicing using Image Fusion.
- [32] Brown, M., & Lowe, D. G. (2003). Recognising Panoramas. *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2*, pp. 1218-. Washington, DC, USA: IEEE Computer Society.
- [33] Eden, A., Uyttendaele, M., & Szeliski, R. (2006). Seamless Image Stitching of Scenes with Large Motions and Exposure Differences. *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2* (pp. 2498-2505). Washington, DC, USA: IEEE Computer Society.
- [34] Russol Abdelfatah, H. O. (2013). Automatic Seamless of Image Stitching. *International knowledge shating platform*, 4, ISSN (Paper) 2222-1727.
- [35] Li Jin, W. Y. (2012). Image Mosaic Based on Simplified SIFT. *International Conference on Mechanical Engineering and Automation*, Vol.10.
- [36] Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J., & Ogden, J. M. (1984). Pyramid methods in image processing. *RCA Engineer*, 29(6), 33-41.
- [37] Burt, P. J., Edward, & Adelson, E. H. (1983). The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, 31, 532-540.
- [38] Medha V. Wyawahare, D. P., & Abhyankar, H. K. (September 2009). Image Registration Techniques: An overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol. 2.
- [39] Baker, S., & Matthews, I. (2004). Lucas-Kanade 20 Years On: A Unifying Framework. *International Journal of Computer Vision*, 56(3), 221-255.