03-134142-008 Ahsan Ali
03-134142-038 Kashif Aftab
03-134142-047 Muhammad Usman

Multi Camera Video Stitching Surveillance System

In partial fulfilment of the requirements for the degree of

Bachelor of Science in Computer Science

Supervisor: Asghar Ali Shah

Department of Computer Sciences
Bahria University, Lahore Campus

June 2018

C e r t i f i c a t e

We accept the work contained in the report titled

"MULTI CAMERA VIDEO STITCHING SURVEILLANCE SYTSEM",

written by

AHSAN ALI

KASHIF AFTAB

MUHAMMAD USMAN

as a confirmation to the required standard for the partial fulfilment of the degree of Bachelor of Science in Computer Science.

Approved by:

Supervisor: Mr. Asghar Ali Shah

__________________________________ (Signature)

June 4*th*, 2018

DECLARATION

We hereby declare that this project report is based on our original work except for citations and quotations which have been duly acknowledged. We also declare that it has not been previously and concurrently submitted for any other degree or award at Bahria University or other institutions.

Enrolment	Name	Signature
03-134142-008	AHSAN ALI	
03-134142-038	KASHIF AFTAB	
03-134142-047	MUHAMMAD USMAN	

Date : ____________________________

Specially dedicated to

my beloved grandmother, mother and father

(Ahsan Ali)

my beloved grandmother, mother and father

(Kashif Aftab)

my beloved grandmother, mother and father

(Muhammad Usman)

ACKNOWLEDGEMENTS

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express my gratitude to my research supervisor, Mr Asghar Ali Shah for his invaluable advice, guidance and his enormous patience throughout the development of the research.

In addition, we would also like to express my gratitude to our loving parent and friends who had helped and given me encouragement.

Ahsan Ali
Kashif Aftab
Muhammad Usman

MULTI CAMERA VIDEO STITCHING SURVEILLANCE SYTSEM

ABSTRACT

CCTV cameras are commonly used for the security issues. Pan-tilt-zoom (PTZ) cameras are mostly used for this purpose. To stitch two or more video streams from different cameras is much cheaper than PTZ solution. There are three stages of video stitching. Feature identification is the first stage of video stitching. To scale the invariant features like rotation, scaling and noise etc. Direct and feature base identification is the basic two types of feature identification. Shifting and warping the images purpose to identify that how there features are agree with each other is concern of direct base identification.

While feature identification rely of extracting the features and then perform matching among them on the base of features. Calibration is the second stage of the video stitching. The images are stitch in panoramic way in calibration depending upon alignment among them. Blending is the last stage of video stitching where numerous videos are display in single panoramic way. Any blending algorithm use to blend the pixels together and final view.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background

As there are many kind of processing like image processing, signal processing etc. in this process some inputs are given to process the image. There are different type of image processing such as digital and analogy image processing. By using the electrical source alteration of images is the concern of analogy image processing, like television image. The digital computers are used in digital image processing purpose to process the images. Image processing play a very important role in video stitching perspectives. Video stitching process totally depends upon the image processing. Video stitching is process to combine multiple videos to a single view. Today there are lot of cameras placed at every point but the main problem is that we need to observe each camera. To watch the numerous videos simultaneously is a lot of time consuming. There is a solution proposed for this problem that all videos can combine in single view. Video stitching technique give the best option to see all camera results in a single view. If we observe all cameras results simultaneously there is chance of miss observances. So due to this reason the interaction is becomes poor with the screen. So video stitching cover all area and give a full reliable panoramic view of multiple cameras. Now a days the panoramic feature has become a very common module in latest mobiles. But the panoramic trick in security cameras is not much popular. Many peoples try to work on it but they faced many problems such as low visual quality and required an expensive hardware. The demand of security cameras is rapidly increasing in security and monitoring perspectives. Pan-tilt- zoom (PTZ) cameras are used mostly for its better quality. But the major problem is that the PTZ monitoring software has to be purchased.

Market demand the cheaper and reliable solution rather than PTZ cameras. PTZ camera solution is costly way and especially its monitoring software's are required to buy [1]. To stitch two or more camera Video by using any free software is much better solution. To review the market demand and cheaper solution the video stitching technique is proposed for CCTV cameras.The smart phones panorama takes input images from single camera, it takes multiple images by same view and stitch them accordingly. Through our proposed solution society will familiar with the panoramic view of multiple security cameras. Which will reduce the human errors, reduce the time consumption and improve the human interaction, which is very reliable and cheap solution.

The given below will the desired functionality of our achievements:

- 360 view of whole room.
- Panoramic view of image
- Total view of all images that captured by camera.

1.2 Problem Statements

The previous work has various issues such as Low visual quality [6].Require expensive and specialized hardware [8, 9].Stitches only a few videos [4].Other issues such as Monitoring blind areas. Removing artifacts (ghosting like effects). A panoramic video stitching for surveillance system is very costly by using the PTZ (pen tilt zoom) cameras. Video based monitoring and assistance is necessary.

1.3 Aims and Objectives

Create a panoramic video of multi cameras to provide a reliable view to multi camera users. This system can be applied anywhere. Provide an extreme cheap solution for multi cameras with reducing human errors, time complexity and improve the human interaction. Improve the visual quality and reducing the numerous screens are also a primary objective.

1.4 Scope of Project

The video stitching is a desktop application that will use to provide a reliable view for the multiple security cameras. The organizations that are using multiple security cameras for their security issue they can get benefit through this application. By using this application they can get very clear and reliable view. Panorama feature in smart phones use to stitch the images, same like panorama feature this application will Stitch the videos of multiple cameras. Today security cameras show their result on the screen in the segments that is not a proper view, it seems not good .the video stitching will provide a proper single view of cameras.

CHAPTER 2

LITERATURE REVIEW

It has been a hot topic since last decade [2].Real time preview is also created [3]. Panoramic video of accepted standard at real time is also created [4] but it can combine only two cameras. Three video streams are also combined [5].Four video streams are also combined [6] but it does not focus on visual quality. Real time panorama video with good performance is also proposed [7] but it only works with two cameras and low resolution image. Good visual quality and high resolution image is also proposed [8, 9] but it require expensive and specialized hardware. The previous work has various issues such as Low visual quality [6].Require expensive and specialized hardware [8, 9].Other issues such as monitoring blind areas.

Removing artefacts (ghosting like effects). Feature extraction is first step, the matching is second step and the last step is blending to monitoring and assistance of the panoramic video. The video stitching technique is totally similar with the concepts of Image stitching which is pre-processing for video stitching. In [10] there are main steps of stitching the images into single view which are image calibration, image registration and image blending. The purpose of calibration is to minimize difference between ideal lens and camera lens. The difference between images can be resulted by optical defects such as distortion and exposure. In [11] represents that there are two types of camera parameters the extrinsic parameter and intrinsic parameter. In [12] image registration discussed as alignment of images taken by multiple points. Image registration focuses to create geometric representation among the images. In [13] image blending concerns preview images to form a single view. There are two ways of blending are discussed the alpha feathering and Gaussian pyramid. Alpha feathering work well when images pixel are well aligned and intensity represent the difference between two images. Gaussian pyramid work on the frequency of image it merge the image on the basis of frequency and then filter them accordingly. There are two approaches of image stitching. Direct technique basically compare the pixel intensities of images. It minimize the overlapping difference between images. The actual benefit of direct technique is that it provide more useful detail for image alignment. There is also disadvantage that there is limited coverage range. In feature base technique feature point in one image and other image which need to compare can be find by local descriptor. Feature base method start by establishing a corresponding connection between points, lines, edge and other geometric entities.

2.1 Feature Detection

There are many types of feature descriptors such as vector descriptor and binary descriptor. ORB detector is binary descriptors while Scale Invariant Feature Transform (SIFT) and Speed up Robust Feature are direction descriptors.

SIFT:

SIFT (scale invariant features transform) extracts the invariant features and partially illuminate the changes like scaling and rotation. A descriptor also called vector (dimension 128) is related with every feature. The feature matching among the image pairs is allowed by the descriptors (vector) to satisfy the geometric transformations. The piece filtering approach is also used for feature detection. The imaginary and tentative outcomes of lindebury shows that Gaussian function that provide possible weighbridge of different scale image space. It is the new method for scale-invariant key point's detection. As the Lundeberg algorithms motivations that key points are extract by SIFT as the minimum Difference of Gaussian method, that could be calculated by differentiation one or more neighbours gages that are disconnected to persistent multiplication element k. the computation cost is the main advantage of the LOW'S detector but it reduced when compare with this method. By using the SIFT the matching between key points is extracted. Describing indigenous appearance area nearby basic ideas is goal of this matching. SIFT descriptor links to the distribution base descriptor (DBD) class.

SURF:

It is the extension of SIFT based on multi-scale space theory but work with different method to extract the features. The fast approximation of HESSIAN matrix speed up its commutations [21].

FAST:

It is applicable feature detector for actual applications. It is known as high speeded detector. The circle of 16 pixels measured in this algorithm that are around the candidate corner p. It often happen when there is pixels that involved in circle are optimistic [14].

Harris:

Suggested by Harris and Stephens. It is junction indicator based on moravee method. Fluctuating the frame in a bit in multiple track is resulted to determining the variation in intensity. The window centre point is extracted as corner points [15].

ORB:

Rubble develop the ORB technique et al [16], which is combination of FAST and BRIEF descriptors that describe the input image features in binary string rather than vector [17]

OPENCV:

Source library for computer vision and image processing that contain many inbuilt functions for real-time image processing [18]. Video stitching is post processing of image stitching .image stitching consists of image acquisition that has three pre-processes the geometric transformation, visual quality alignment and temporal synchronization. The visual quality will improve the quality of the video which is much cheaper than PTZ camera solution [19]. Post processing like tracking and panning used to improve the cinematic colour, contrast and white balance. Such algorithms that provide best quality are preferred to increase the video quality. In [10] there basic three steps of image stitching is discussed such as image calibration, image registration and image blending.

2.2 Image Registration

In this stage of the stitching process, different algorithms are used to find features within the video. These features include corners, black dots on white backgrounds, the endings of branches and any location with significant 2D texture [2]. The features are also known as "interest points". There are two broad camps for image registration: direct based and feature based [3]. Feature based registration involves extracting distinctive features from each image, matching these features to establish a global correspondence and then estimating the geometric transformation between the images [4]. Direct based registration involves shifting and warping images to find out how much the pixels agree [4]. This requires an error function to be created which needs to be minimised. A search function is also needed to dictate the shifting process. Direct based registration is often preferred, since it makes use of all the data and provides a very accurate registration [3]. Both techniques, however, are not robust to image zoom, change in illumination or noise [3]. Schmid et al used two metrics to measure different interest point detection algorithms: repeatability and information content [2]. Repeatability is the condition that signifies that detection is independent of changes in the imaging conditions, such as the position of the camera relative to the scene and the illumination conditions. Information content is a measure of the distinctiveness of an interest point i.e. the distribution of the interest points. Algorithms that can find many interest points over the entire image will score highly in this metric. Schmid et al [2] have compared six different detectors and have found that the Imp Harris detector (an improved version of the Harris detector) scores highly for the repeatability metric as well as the information content metric. Thus, the algorithm to use for interest point detection is the ImpHarris algorithm. The next step is to extract invariant features based on Lowe's Scale Invariant Feature Transform (SIFT) [3]. Invariant features have the property that they are invariant to image scale and rotation, distortion, change in 3D viewpoint, noise and change in illumination [4]. Lowe describes the SIFT algorithm as an approach that transforms image data into scale-invariant coordinates relative to local features [4].

The above method of identifying features is recommended over the feature or direct based image registration, since invariant features can be found more repeatedly and matched more reliably [3]. Steedly et al mentions that the large numbers of interest points is a computational bottleneck in the process of image stitching [3]. They suggest a method to compress the number of interest points to improve computation speeds. The method involves recursively splitting up the image until the number of interest points within a “split” is not more than a certain number. An interest point is then placed at the centroid of the group of interest points. At this point, all overlapping images need to be matched together. Brown et al mentions that after every feature is extracted from all the images, they are matched to their respective nearest neighbours using a k-d tree [5].

2.3 Calibration

Once all the images are ordered, they will undergo bundle adjustment [3]. Bundle adjustment is the process of refining a visual reconstruction to produce jointly optimal 3D structure and viewing parameter estimates [6]. In a sense, bundle adjustment tries to minimise the differences between an ideal lens model and the camera lens used to take the video. Szeliski mentions that one way to register a large number of images is to add new images to the panorama, one at a time, aligning the most recent image with the previous ones already in the collection [7]. Brown et al uses the above approach, with the distinction that each image will be initialised with the same rotation and focal length as the image to which it best matches [5].

2.4 Blending

The last stage in video stitching is to stitch the video streams together to create one panoramic video stream. In this stage, all visual artefacts are minimised, so as to give an impression that the panoramic video was recorded using one camera. A compositing surface must first be chosen as well as a reference image. This decision depends on the positions of the cameras during the recording of the videos. If there are only a few videos that need to be stitched together, selecting any one of the videos to be a reference frame does not matter much as there will be minimal projection distortion. In this case, one of the videos will be a reference frame and the rest of the videos will be warped into the reference coordinate system. This is a flat panorama [7]. For larger fields of view involving many cameras, using a flat surface as the projection surface will severely stretch the pixels at the extremities of the video [7]. In this case, a cylindrical or a spherical surface would be a better choice. Szeliski mentions three approaches to project images onto surfaces. The first approach is to divide a viewing sphere into several large, overlapping regions with each region being a planar surface [8]. The second approach involves choosing a reference image, or a “base frame”. Each frame has its position computed relative to the base frame. The difference here is that a new base frame is chosen periodically for doing the alignment [8]. This method is useful when the camera undergoes large motions when taking images.

The last approach Szeliski mentions is to use a cylindrical viewing surface to represent the image mosaic [8]. Once the pixels have been mapped to a surface, there is still the issue of blending the pixels. The pixels can be averaged to form a final pixel, or a weighted average can be used, where the pixels in the centre are weighted more heavily than the pixels on the edges [7]. Brown et al used multi-band blending to minimise visual artefacts [5]. This method involves taking the area of overlap of the images and blending the pixels depending on the frequencies of the colours of the pixels. Low frequencies are blended over a large spatial range; high frequencies are blended over a short spatial range.

CHAPTER 3

DESIGN AND METHODOLOGY

3.1 Proposed Process

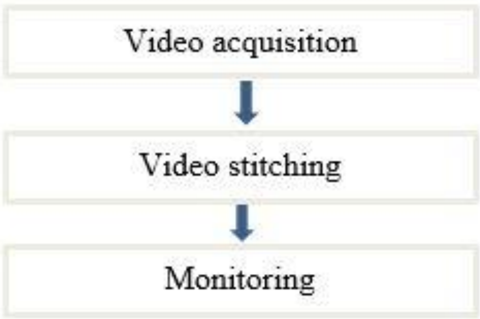

Figure 1: Proposed Model

In our proposed process the first step is video acquisition, the second stage is video stitching and blending is last stage. In first stage the video streams by multiple cameras will input to the system. Then the video will be stitch in panoramic view and then view on the screen by using blending algorithm.

3.2 Video Acquisition

The video acquisition stage is concern of input video streams. In this case the different dedicated cameras are used to capture the multiple video streams of same view. But the videos frequency, ratio, resolution and frame rate should be same. Every camera should follow the same standards. Given below is example of video acquisition:

Figure 2 : Camera Alignment for Video Acquisition

In above fig the basic video acquisition camera adjustment using two cameras is shown for video acquisition understanding purpose. It is basically representing the dedicated cameras physically adjustment. The position of cameras should be static. There should be no motion between cameras and the second main thing is the second cameras video should start capture where the first cameras FOV (field of view) ends. After the video is acquire the next major step is video stitching but during this process there are some pre-processing occurs which is given below.

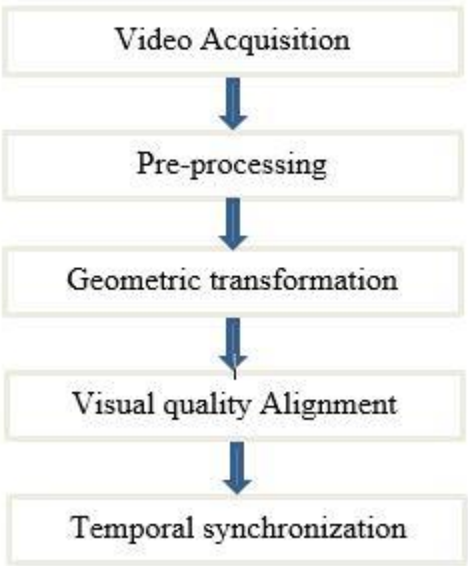

Figure 3 : Proposed Process

3.2.1 Pre-processing

Basically the aim of image pre-processing is an improvement of image data that finish the unwanted distortion and enhance some image feature so that some important steps apply for further processing. There are four category of image pre-processing methods that is used for the calculation of new pixel brightness.

Pixel brightness transformation.

- Geometric transformation.
- Pre-processing method that use a local neighbourhood of processed pixel.
- Image restoration that requires knowledge about the entire image.
- Image pre-processing method use the considerable redundancy in images. In this step neighbouring pixels of images essential have the similar brightness value corresponding to real images.

Video stitching pre-processing is consisting of transformation of images. Transformations focus rotation, scaling of the images. The next is to improve the visual quality, for this purpose many algorithms are available in computer vision. After the temporal synchronization is necessary that focus on frame to frame adjustments of the images. So, the mainly pre-processing has three steps.

1. Geometric transformation
2. Visual quality Alignment
3. Temporal synchronization

3.2.2 Geometric Transformation

The image processing is the pre-processing for video stitching. The homography help to detect the overlapping pairs of images. The optimal values of the elements in the 2D homography are estimated by minimizing the total transformation error. The homography approximation method proposed for the proper image transformation [20]. It convert one plane to another based on feature points matching intended by match pairs. SIFT [21], SURF [22], SFOP [23], EBR [24] commonly recycled feature finders. The SIFT is the best algorithm as compared with the other algorithms due to its best performance. After feature point extraction and approximation matching is needed for this purpose the matching algorithms used. The geometric transformation is achieved when the homography is properly computed. If the homography is not compute the matching algorithm will not work. Homography computation is also play important role for visual quality. The proper homography resulted better results.

3.2.3 Visual Quality Alignment

The visual quality in video stitching is achieved by the better alignment of the images. Essentially it is programmatic technique in which we adjust bundle. Resolution of this phase is to reduce miss-registration between all pairs of images. Least-squares algorithm [12] is used to get the optimal value of 3D reconstruction of image scene and camera position. Once completed the global alignment then there may be some errors due to local miss-registration like double images, images may be blur, so for these problem there is need to perform local adjustment like parallax removal. Different layouts use on which image stitching use identical rectilinear projection.
On this layout stitched image can be viewed on 2D dimensional level overlapping the panosphere in a single point. Different lines of images shows their direction on 2D plane. The disparity calculation algorithm is used to measure the disparity among the images and better for visual quality alignment.

3.2.4 Temporal Synchronization

Temporal synchronization achieved by to recover the temporal alignment. It determines the sequence of the frames which first frame is given to the next corresponding frame. The cameras should be rigidly joined together and there should be non-uniform motion between the cameras. This process consist assess the errors. For cylindrical panorama images captured with camera mounted on level of tripod. Each image can be warped on their cylindrical coordinates if the camera focal length known. There are two types of cylinder-shaped changing

- Forwarding changing
- Inverse changing

In forwarding changing image source is mapped onto the cylindrical apparent but there is problem that it have hovels in the end point of images since some pixels not mapped on the cylindrical surface. So that reason we use inverse warping because in this warping each pixels in end point image clearly mapped onto the basis image. The following bilinear exclamation second-hand to evaluate colours of pixels at end point [25]. Temporal synchronization refer further to the visual quality. After achieving the visual quality we combine the videos taken by multiple cameras in single panorama shape [26].

3.3 Video Stitching

After video acquisition and pre-processing the video stitching is the next step where the multiple videos are stitched together to form a panoramic view. The video stitching process totally works on image stitching. The image stitching process is consisting of four stages. The user will select input overlapping images first, which are to be stitched. Image merging and registration will provide the final results. Image registration consists of feature matching which is done by feature descriptor technique and creating feature detector and. combination of key point dataset created by SURF, SIFT, MSER feature detector algorithms will included in Feature key points. The proposed technique below is a panorama image stitching system that focusing on image registration part. We match the features of images in image registration part by using some feature detection technique and output will be generated. The image registration's output will provide you higher quality of image stitching i.e. your final image. Feature matching provide key point detectors and key point descriptors. We can transform or overlay 2 images to efficient manner Based on large number of matches. Feature abstraction and equivalent are Entire process of stitching. In this section we will identify the features and physically adjustment of the CCTV cameras. Here camera station, angle or field of view is the highly concentrated. No overlapping regions in cameras field of view. There should not be any area that uncovered by cameras FOV. Following are different category for image pre-processing. These methods are used for new pixel intensity calculation.

- Pixel intensity transformation.
- Linear transformation.
- Pre-processing method that use a local neighbourhood of processed pixel.

The video acquisition is process of getting the videos from multiple cameras and performs actions to next. In this section there are pre-process that are needed to perform for further process.

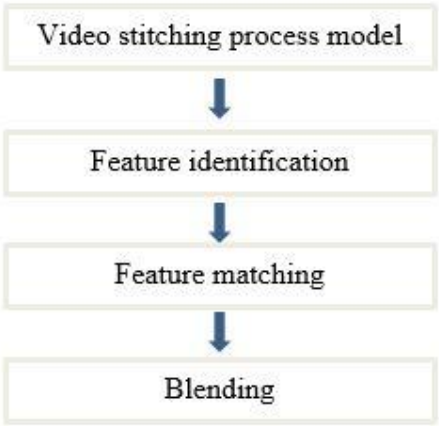

Figure 4 : Video Stitching Process

3.3.1 Feature Identification

This section consists of finding interesting features or interesting points in the overlapping part of the image. To recognize the features of the images is the first stage where we citation the features of the images and match the features by using matching algorithms. There are many feature indicator technique such as [27] Harris, Scale Invariant Feature Transformation (SIFT) [28], [29] Speeded up Robust Feature (SURF), ORB descriptor (oriented fast and rotated brief) [30] and the feature from accelerated segment (FAST). So additional algorithm which is SURF try to improve calculation stretch of Scale Invariant Feature Transformation via use about uncomplicated image to profligate limited ascent computation upon image. Harris corner detector most useful because it detect the interest point (common point) between images. These points have large intensity. When point detect then there is need to match them.

ORB feature detector is the combination of basically two detectors

1. Oriented fast
2. Rotated brief

ORB detector has the object with function that have no. of Optional Parameters. It select two points at a time and identify its features.

3.3.2 Feature Matching

After identify the features the next stage is matching. There points are match by normalized cross correlation which generate pairs. After that these pair need to be match if anyone is wrong or not. In [9] there is also Harris corner algorithm in which it apply multiple constraints to remove incorrectly match pair of images. Basically this algorithm work on frame by frame because it need to identify which image is on left or right side. In [16] there is solution of Harris corner that instead of using frame by frame stitching there is need to find the interest point of previous of image by using optical flow and area of overlapping. They found that there is 85% time consume in matching interest point of images. By using OPENCV after completing the feature detection then matching algorithm use for matching. The NORM-HAMMING distance algorithm use for matching which takes 3 or 4 points to produce brief descriptor then NORM-HAMMING2 distance is used.

3.3.3 Blending

Feature image blending technique is basically used in computer graphics. It is simple method in which pixel value in merged regions have average value over weighted overlapping regions. Due to presence of exposure this technique does not work well. Another approach is multi-band image blending. In this is image gradient form each source is copied and matches to another. If match then these gradients are reconstructed. Approach image pyramid is actually demonstration of image .it mean hierarchical representation of images at different resolution. It provide useful properties for different application like noise reduction, image exploration and image development etc. it is the last stage the viewer can enjoy the panoramic video of the multiple cameras. It is the last stage the viewer can enjoy the panoramic video of the multiple cameras.

The given below fig part (a) contains multiple inputs by numerous cameras while the part (b) is the final panoramic preview of input images. The part (b) preview is considered as blending.

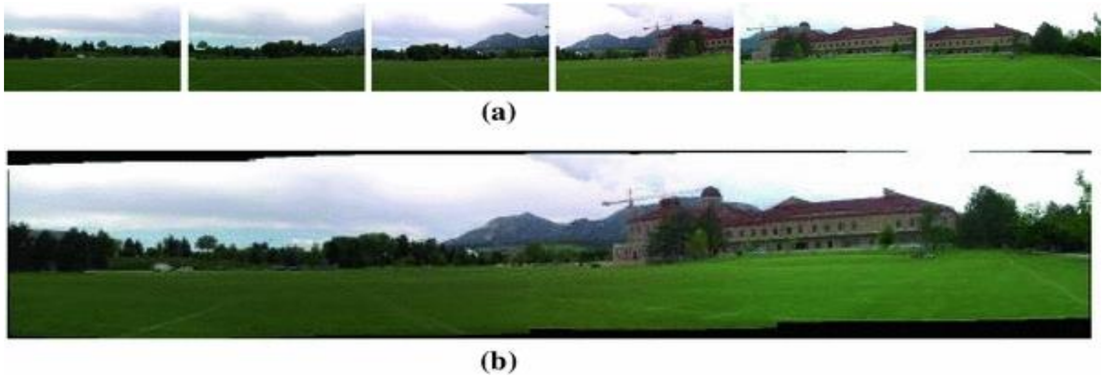

Figure 5 : Image Stitching Example

CHAPTER 4

DATA AND EXPERIMENTS (IMPLMENTATION)

4.1 Data Collection

We collect the relevant data about our project from the published thesis of different universities and research papers. We properly referenced to whom we gathered the data. The international research papers throughout helping us to achieve our goal. The knowledge about tool regarding our project is driven by our supervisor, internet and papers. Moreover the data sets are freely available for this purpose.

4.2 Experiment

We perform the pre-processing (image stitching) for our project. As we already mentioned that the image stitching is pre-processing for video stitching. We are using the tool OPENCV for this purpose. The OPENCV provide us the build in functions for image stitching. We perform the feature identification in OPENCV by using the ORB (oriented fast and rotated brief) detector and we achieve the image matching by using RANSAC algorithm a new paragraph should not begin on the last line of a page. A subsection title should not begin on the last line of a page.

4.3 Implementation

We are using the OPENCV tool for our project. Video stitching process contains the numerous stages. We are at the stage where we installed the OPENCV and perform the image stitching by using the different algorithms which is very special task for the project. There are many algorithms used for feature identification, matching and blending. We are using the best and latest techniques and algorithms for our project. The OpenCv provide many inbuilt functionality for image processing, especially its configuration with python is more efficient rather than C++. Python provide a complete pipeline for image stitching.

4.4 Camera Frame Rate

The cameras using in project will be the same capacity and nature. The every camera will capture the frames 25/second or 30/second.

Table 1 : Frame Rates of Dedicated Cameras

Camera	Frame rate/second
1	25
2	25
3	25
4	25
5	25
6	25

The nature of adjusted cameras should be same because it is very necessary for feature identification, matching and blending.

4.5 Physical Adjustment Camera Model

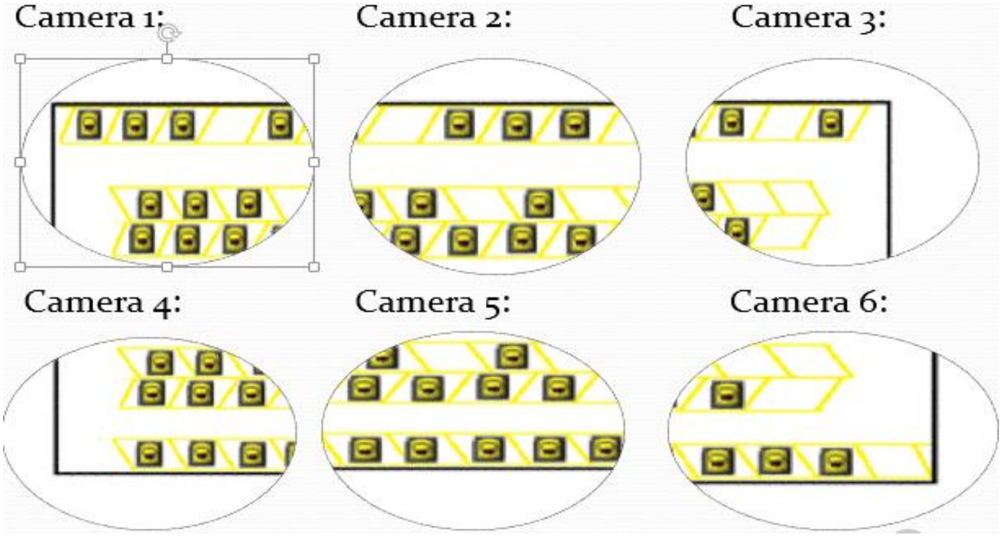

Figure 6 : Multi Cameras Video Capture

The above image representing the physically adjustment of cameras. The dedicated static cameras should be use and there should be non-uniform motion between the cameras. The cameras adjustment should be according to the cameras FOV (field of view). The camera 2 should start where the camera 1 FOV ends. The camera 3 should start capture where the camera 2 FOV ends and the same sense will use for all the cameras. The above image representing the many cases for video stitching those are given below.

Case 1: for upper cameras the camera 1 and camera 2 should be stitch and then stitched video will further stitch with camera 3.

Case 2: the camera 1 and camera 4 will stitch first and the resulted stitched video of 1 and 4 will further stitch with the resulted stitched video of 2 and 5 and so on the resulted video will further stitch with the resulted stitched video of 3 and 6.

Case 3: the upper cameras 1, 2 and 3 stitch first and the resulted stitched video will further stitch with the resulted stitched video of camera 4, 5, 6.

CHAPTER 5

OFFLINE VIDEO STITCHNG

5.1 Video Acquisition for Offline Video Stitching

The multi camera video stitching can be perform through off line videos. In this sense the logic will be same for off line videos and real time stitching. The only difference is that off line videos are taken not at real time while real time stitching perform live videos captured by the cameras at real time. In this section the dedicated cameras should be used and there should be no motion between the cameras. The given below is the physically cameras adjustment model.   For video acquisition all the cameras should follow the same standard and quality and ration. The frame rate should be same for every camera and time must be same. The cameras FOV(field of view) must be good.

Figure 7 : Video Capture Form Multi cameras

The above image represents the physically adjustment of cameras for off line video capturing. Every camera will follow the same standards. The video stitching process for off line video is given below where the video acquisition is the first stage.

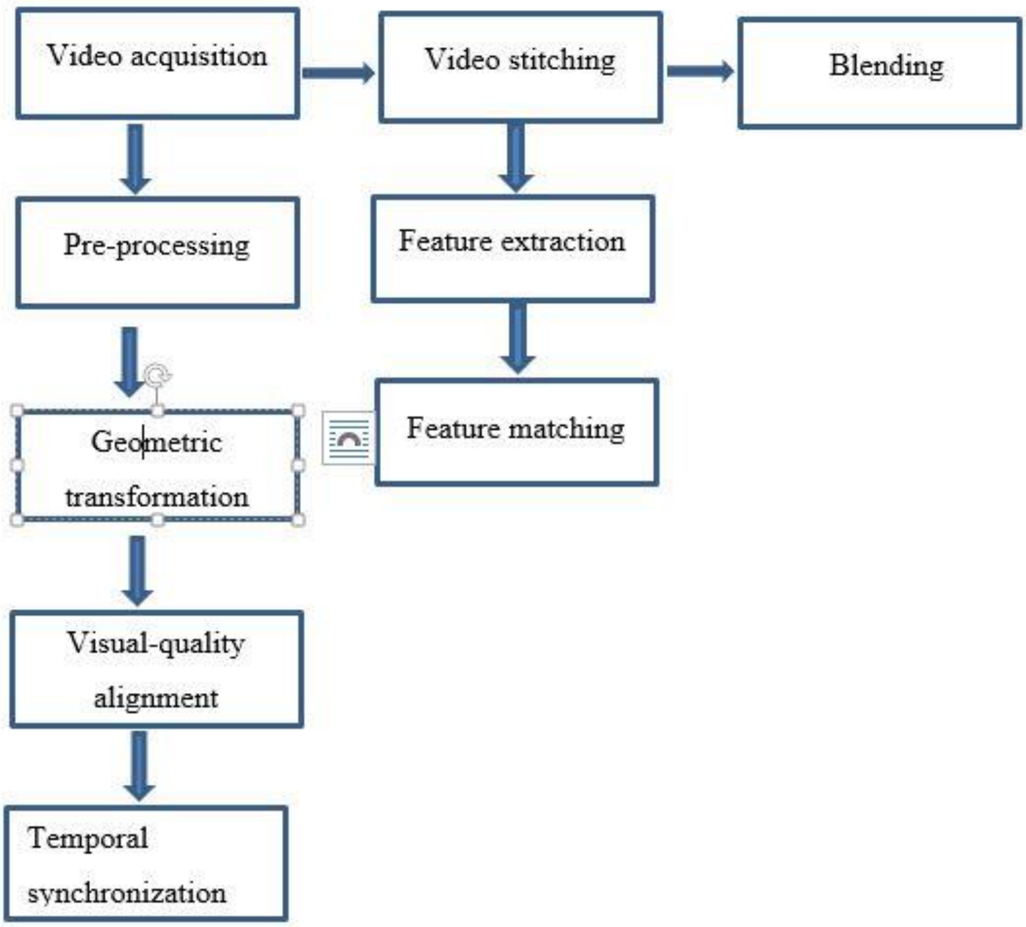

Figure 8 : Offline Video Stitching Processes Steps

5.2 Video Acquisition for Offline Video Stitching

The image processing is the pre-processing for video stitching. The homography help to detect the overlapping pairs of images. The optimal values of the elements in the 2D homography are estimated by minimizing the total transformation error. The homography approximation method proposed for the proper image transformation [20]. It convert one plane to another based on feature points matching intended by match pairs. SIFT [21], SURF [22], SFOP [23], EBR [24] commonly recycled feature finders. The SIFT is the best algorithm as compared with the other algorithms due to its best performance. After feature point extraction and approximation matching is needed for this purpose the matching algorithms used. The geometric transformation is achieved when the homography is properly computed. If the homography is not compute the matching algorithm will not work. Homography computation is also play important role for visual quality. The proper homography resulted better results.

5.2.1 Visual Quality Alignment

The visual quality in video stitching is achieved by the better alignment of the images. Essentially it is programmatic technique in which we adjust bundle. Resolution of this phase is to reduce miss-registration between all pairs of images. Least-squares algorithm [12] is used to get the optimal value of 3D reconstruction of image scene and camera position. Once completed the global alignment then there may be some errors due to local miss-registration like double images, images may be blur, so for these problem there is need to perform local adjustment like parallax removal. Different layouts use on which image stitching use identical rectilinear projection.
On this layout stitched image can be viewed on 2D dimensional level overlapping the panosphere in a single point. Different lines of images shows their direction on 2D plane. The disparity calculation algorithm is used to measure the disparity among the images and better for visual quality alignment.

5.2.2 Temporal Synchronization

After video acquisition and pre-processing the video stitching is the next step where the multiple videos are stitched together to form a panoramic view. The video stitching process totally works on image stitching. The image stitching process is consisting of four stages. The user will select input overlapping images first, which are to be stitched. Image merging and registration will provide the final results. Image registration consists of feature matching which is done by feature descriptor technique and creating feature detector and. combination of key point dataset created by SURF, SIFT, MSER feature detector algorithms will included in Feature key points. The proposed technique below is a panorama image stitching system that focusing on image registration part. We match the features of images in image registration part by using some feature detection technique and output will be generated. The image registration's output will provide you higher quality of image stitching i.e. your final image. Feature matching provide key point detectors and key point descriptors. We can transform or overlay 2 images to efficient manner Based on large number of matches. Feature abstraction and equivalent are Entire process of stitching.

In this section we will identify the features and physically adjustment of the CCTV cameras. Here camera station, angle or field of view is the highly concentrated. No overlapping regions in cameras field of view. There should not be any area that uncovered by cameras FOV. Following are different category for image pre-processing. These methods are used for new pixel intensity calculation.

- Pixel intensity transformation.
- Linear transformation.
- Pre-processing method that use a local neighbourhood of processed pixel.

The video acquisition is process of getting the videos from multiple cameras and performs actions to next. In this section there are pre-process that are needed to perform for further process. The video stitching process consists of feature extraction and feature matching. After extraction and matching of features the blending is the last stage where the video is represented on the screen.

5.3 Blending

Feature image blending technique is basically used in computer graphics. It is simple method in which pixel value in merged regions have average value over weighted overlapping regions. Due to presence of exposure this technique does not work well. Another approach is multi-band image blending. In this is image gradient form each source is copied and matches to another. If match then these gradients are reconstructed. Approach image pyramid is actually demonstration of image .it mean hierarchical representation of images at different resolution. It provide useful properties for different application like noise reduction, image exploration and image development etc. it is the last stage the viewer can enjoy the panoramic video of the multiple cameras. It is the last stage the viewer can enjoy the panoramic video of the multiple cameras.

CHAPTER 6

RESULTS AND DISCUSSIONS

6.1 Result

The ORB detector is proved very accurate for feature extraction rather than SIFT,SOFT and SAFT etc. it perform the better feature identification and its performance is the best, we are using the ORB for feature extraction. The RANSAC is proved best for matching rather than other techniques we are using RANSAC algorithm for matching purpose. And the last stage blending is required to touch, we will choose best technique for blending. There are 6 types of blending we studied. The multi band blending technique is best for blending.

Table 2 : Process and Technique

Process Name	Technique	Best Technique
Video Acquisition	Physically Adjustment of cameras	Adjustment of cameras accordingly camera's FOV(field of view)
Feature Identification	1.SIFT(Scale Invariant Feature Technique) 2. SURF(Speed up Robust Feature) 3.ORIENTED FAST(feature from accelerated segment test) 4.ROTATED BRIEF(binary robust independent elementary feature) 5.ORB(oriented fast and rotated brief)	ORB detector (oriented fast and rotated brief)
Feature matching	1.HARRIS CORNER PAIR algorithm 2.RANSAC(random sample consensus) 3.NORM-HAMMING ALGORITHM 4.NORM-HAMMING ALGORITHM2	RANSAC(random sample consensus)
Blending	1.Alpha Feathering, Gaussian pyramid 2.Multi band blending technique	Multi band blending technique

6.2 Discussion

There are different technique used in video stitching process. Video acquisition, feature identification, matching and blending are the main steps of video stitching process. There are many techniques used. The video acquisition focuses on the physically adjustment of cameras. SURF, SIFT and ORB descriptors are used for feature identification but best technique in ORB descriptor. It is the combination of two descriptors the oriented fast and rotated brief. This algorithm is best as compare to other algorithm due to its best performance. HARRIS corner pair, RANSAC, NORM-HAMMING algorithm and NORM-HAMMING2 are used or feature matching technique but the best one is RANSAC algorithm for its better performance.. In RANSAC basically find picture to its neighbour picture which match it. In feature matching the points are matched by normalized cross correlation which generate pairs. In blending process alpha featuring is Gaussian pyramid techniques are used. Basically blending technique applied on image to stitch all images seamless. In many research papers the SIFT and SOFT are mostly used for the feature identification.

CHAPTER 7

CONCLUSION AND RECOMMENDATIONS

7.1 Conclusion and Future Work

Even though image/video processing demands a lot of computing power, algorithms that do this have existed since the 1990s. Consequently, there are many algorithms to choose from. Image registration is the first step in video stitching. Lowe's Scale Invariant Feature Transform is the better algorithm in recognising interest points, as these points are invariant to orientation, scale, distortion and noise. The second step in video stitching is calibration. Images are added to the panorama, one at a time, which are then aligned to the images currently in the panorama already. The last step in video stitching is blending. Since the lecture recordings will involve many static cameras, a flat projection should be fine as a projection surface. With regards to the blending of pixels, any algorithm will do, since all recordings will be indoors. Multi-view video stitching is an important topic in the joint domain of computer vision, image and video processing, multimedia and computational photography. In this thesis I addressed several important sub-problems of multi-view video stitching, including: video stabilization, efficient alignment and stitching of high-resolution and long-duration panoramic videos, color correction for panoramic image/video stitching, image partial blur detection and classification, and blurred frame detection and repair. In Chapter 1, I first introduced the concepts and the standard procedure for panoramic video stitching . Then, I described my contributions to individual stages of this stitching procedure in the following chapters. The "video pre-processing" stage aims at removing motion jitter and repairing blurred frames existing in the individual mono-view videos because they not only decrease the quality of the videos but also may distort image information and cause later computer vision algorithms to fail. For motion jitter removal, the standard Kalman filtering based smoothing technique may generate abnormal values in the global motion chain because it does not consider the physical continuity of rotation angle values when successive values cross the boundary of Cartesian quadrants II and III. I thus proposed an approach called continuity aware Kalman filtering of rotation angles to solve this problem. For repairing the blurred frames in a video, the first task is to detect those frames. My blurred frame detection scheme is based on evaluating the percentage of motion blurred blocks in an input frame. I developed a learning-based image partial blur detection and classification approach for this purpose which combines various kinds of local blur features in an SVM-based learning framework. My blurred frame repair approach treats the blurred frame as a frame with “missing” global motion parameters in its temporal neighbourhood, and tries to interpolate this frame out from its immediate neighbouring frames after first completing the global motion chain using polynomial curve fitting. In 92 developing my blur classification technique I have created an easy to use labelling tool for acquiring human-labelled instances of the three blur classes (sharp/motion/defocus). In addition I will release my labelled blur examples as a public dataset for evaluation of blur detection methods. With the popularization of commodity HDTV cameras, the second stage “multi-view video alignment” and the fourth stage “panoramic video stitching” in the stitching procedure face huge input data and demand processing efficiency. Traditional alignment and stitching algorithms are incapable processing HDTV-level (e.g., 1920x1080 pixels) and long-duration (e.g. one hour long) multi-view videos on standard workstations in a timely manner. I thus proposed three approaches to solve this problem including constrained and multi-grid SIFT matching schemes, concatenated image projection and warping and min-space feathering of high-resolution images. These three approaches together can greatly reduce the computational time and memory requirements for panoramic video stitching, which makes it feasible to stitch long, high-resolution panoramic videos using standard workstations. In addition to the efficiency issues, I also proposed to use multiple frames instead of a single frame for multi-view video alignment to achieve better robustness to the uneven spatial distribution of SIFT features. My work on the “colour correction” stage a primary emphasis of this thesis. I first performed an extensive survey and performance evaluation of representative colour correction approaches in the context of two-view image stitching. One contribution of this evaluation work is that it unifies colour balancing approaches, colour transfer approaches and colour correction approaches, which were originally proposed in different research domains including computer vision, video processing, computer graphics and colour and imaging sciences. Analysing these approaches under the single framework of panoramic image stitching, broadens the horizon of the research. The other contributions include giving useful insights and conclusions about the relative performance of the selected approaches, and pointing out the remaining challenges and possible directions for future colour correction research. As a result of this worked I have created and made public a standard evaluation dataset for evaluation of colour correction algorithms. Based on the conclusions drawn by this evaluation work, I proposed a hybrid and scalable colour correction approach for general n-view image stitching and a two-view video colour correction approach for panoramic video stitching. In conclusion, in this thesis I described a sequence of approaches I proposed to address the technical or practical problems existing in the different stages of panoramic image/video stitching. The proposed approaches are novel, effective and efficient solutions to the problems they address.

REFERENCES

[1] “Marquard, Stephen”; “Matterhorn 2014 Unconference: Ideas for automated post-recording video Handling,” 2014. 19 March 2014.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] “Brown, Matthew, Lowe, David G”; Automatic panoramic image stitching using invariant features." International journal of computer vision 74, 1, 59-73, 2007

[3] “Baudisch, Patrick; Tan, Desney; Steedly, Drew; Rudolph, Eric ; Uyttendaele, Matt; Pal, Chris , Szeliski , Richard “; “Panoramic view finder: shooting panoramic pictures with the help of a real time preview”, In Proc. UIST, 2005

[4] “Abood, Saif Salah; Abdullah Karim, Abdulamir”; “Fast generation of video panorama”, International Journal of emerging Trends & Technology in Computer Science (IJEITCS), 3, 6, 80-84, 2014

[5] “Jiang, Wei; Gu, Jinwei”; “Video stitching with spatial-temporal content-preserving warping." In Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on, 42-48, 2015.

[6] “Tang, Wai-Kwan; Wong, Tien-Tsin; Heng, PhengAnn” ; “A system for real time panorama generation and display in tele-imersive applications”, Multimedia IEEE Transections, 7,2, 280-292, 2005

[7] “Adam, Michael; Jung, Christoph; Roth, Stefan; Brunnett, Guido”; ” Real-time Stereo-Image Stitching using GPU-based Belief Propagation”. In VMV, 215-224, 2009

[8] “Kulkarni, Arun V; Jagtap, J.S.;.Harpale, V.K;” “Object recognition with ORB and its Implementation on FPGA". International Journal of Advanced Computer Research, 3, 3, 164, 2013International Journal of Advanced Computer Research, 3, 3, 164, 2013

[9] “Kaiser, Rene; Thaler, Marcus; Kriechbaum, Andreas; Fassold, Hannes; Bailer, Werner ; Rosner, Jakub”; Real-time person tracking in “high-resolution panoramic video for automated broadcast production”, 21-29, 2011

[10] “Shashank, K; SivaChaitanya, N; Manikanta, G; Balaji, Ch, N, V; Murthy, V, V, S”; “A Survey and Review over Image Alignment and Stitching Methods”, International Journal of Electronics &Communication technology, 5, 50-52, 2014

[11] “Rosten, Edward; Drummond, Tom”; “Machine Learning for High-speed Corner Detection”, European Conference on Computer Vision, 1,430-443, 2006

[12] ”Hetal, Mrs; Patel, M”; “Comprehensive Study and Review of Image Mosaicing Methods”, International Journal of Engineering research & Technology (IJERT), 1, 1-9, 2013

[13] “Zhang, Zhen”; “A flexible new technique for camera calibration. IEEE Transactions on pattern analysis and machine intelligence, 22, 11, 1330-1334, 2000

[14] “Rublee, Ethan; Rabaud, Vincent; Konolige, Kurt; Bradski, Gary”; “ORB: An efficient alternative to sift or surf”, In Proc. Of International Conference on Computer Vision (ICCV), 2564-2571, 2011

[15] “Agarwal, Anubhav; Jawahar, C. V.; Narayanan, P. J. “; “A survey of planar homography estimation techniques”. Centre for Visual Information Technology, Tech. 2005.

[16] “Ramasamy, Karthik; AnnisFathima, V. Vaidehi “Panoramic view creation using invariant momentsand SURF features”,376-382, 2013

[17] “Deng, Yining; Zhang, Tong”; “Generating Panorama Photos”, International Society for Optics and Photonics, 270-279, 2003

[18] “Rublee, Ethan; Rabaud, Vincent; Konolige, Kurt; Bradski, Gary”; “An efficient alternative to SIFT or SURF”. In Computer Vision (ICCV), 2011 IEEE international conference on 2564-2571, 2011

[19] “Brown, Matthew; Lowe, David G”; “Automatic panoramic image stitching using invariant features”. International journal of computer vision, 59-73, 4, 2013.

[20] “Tennoe, Marius; Helgedagsrud, Espen; Naess, Mikkel; Alstad, Henrik Kjus; Stensland, Hakon Kvale; Gaddam, Vamsidhar Reddy; Johansen, Dag; Griwodz, Carsten; Halvorsen, Pal”; “Efficient implementation and processing of a real-time panorama video pipeline”,76-83, 2013

[21] “Kim, Chi-Woong;Hyun Seong, Sung;Hee-Jun ,Yoon.”; “Camera image stitching using feature points and geometrical features”. International Journal of Multimedia and Ubiquitous Engineering, 289-296, 2015

[22] “Motaz El-Saban, Mostafa Izz, Ayman ,Kaheel “;“Fast stitching of videos captured from freely moving devices by exploiting temporal redundancy”. In Image Processing (ICIP), Conference on1193-1196, 2010;

[23] “Fischler, Martin A; Bolles, Robert C”; “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography”, ACM 24, 6, 381-395, 1981

[24] “Karthik, R; AnnisFathima, A; Vaidehi, V”; “Panoramic view creation using invariant momentsand SURF features”. In Recent Trends in Information Technology (ICRTIT), 376-382, 2013

[25] “El-Saban, Motaz; Izz, Mostafa; Kaheel, Ayman”; “Fast stitching of videos captured from freely moving devices by exploiting temporal redundancy”, In Image Processing (ICIP), 17th IEEE International Conference on 1193-1196, 2010

[26] “Krystian,Mikolajczyk;Luc ,Van Gool, Tinne ;Tuytelaars, Cordelia Schmid, Andrew Zisserman, Jiri Matas, Frederik Schaffalitzky”, “A comparison of affine region detectors”, International journal of computer vision, 65, 1-2, 2005

APPENDICES

APPENDIX A: Definition, acronyms, and abbreviation

Definition, acronyms, and abbreviation.

Term	**Definition**
PTZ	Pan-tilt- zoom
FOV	Field of view
SIFT	Scale invariant feature transformation. This algorithm is used to find feature between images.
RANSAC	Random Sample Consensus. This algorithm is used to stitch the images if feature are properly matched.
CCTV	closed-circuit television
SURF	Speed up Robust Feature
FAST	feature from accelerated segment test
ORB	oriented fast and rotated brief
HARRIS CORNER	This algorithm is used to extract corner points between images.
Calibration	Difference between ideal lens and camera lens
Image Registration	Alignment of image captured from different points
Blending	Stitch the image to view clearly
Frame	A picture that contain area with pixels
Acquisition	Get image from camera

Detection and Matching	Find the feature between images and then match if they are same.
Homography	Identify which image area is near to other image
Pre-processing	Improvement of image data distortion and enhance image quality
Alpha feathering	Alpha feathering work well when images pixel are well aligned and intensity represent the difference between two images
Gaussian pyramid	Gaussian pyramid work on the frequency of image it merge the image on the basis of frequency and then filter them accordingly
Open CV	Source library for computer vision and image processing that contain many inbuilt functions for real-time image processing
Stitching	Stitch two or more videos is such a way that they display a panoramic view.