<div dir="rtl">

إقـــــرار

أنا الموقع أدناه مقدم الرسالة التي تحمل العنوان:

</div>

Free Viewpoint Video Based on Stitching Technique

<div dir="rtl">

أقر بأن ما اشتملت عليه هذه الرسالة إنما هو نتاج جهدي الخاص، باستثناء ما تمت الإشارة إليه حيثما ورد، وإن هذه الرسالة ككل أو أي جزء منها لم يقدم من قبل لنيل درجة أو لقب علمي أو بحثي لدى أي مؤسسة تعليمية أو بحثية أخرى.

</div>

DECLARATION

The work provided in this thesis, unless otherwise referenced, is the researcher's own work, and has not been submitted elsewhere for any other degree or qualification

Student's name: Rami O. M. Skaik
<div dir="rtl">اسم الطالب: رامي عثمان محمدحسن سكيك</div>

Signature:
<div dir="rtl">التوقيع:</div>

Date: 22/12/2014
<div dir="rtl">التاريخ: ٢٠١٤/١٢/٢٢</div>

Islamic University – Gaza –Palestine

Deanery of Graduate Studies

Faculty of Information Technology

Free Viewpoint Video Based on Stitching Technique

Prepared by:

Rami Othman M. H. Skaik

120100560

Supervisor

Dr. Ashraf Alattar

بسم الله الرحمن الرحيم

الجامعة الإسلامية – غزة
The Islamic University - Gaza

مكتب نائب الرئيس للبحث العلمي والدراسات العليا   هاتف داخلي: 1150

الرقم: ج س غ/35/........ Ref

التاريخ........2014/12/22م Date

نتيجة الحكم على أطروحة ماجستير

بناءً على موافقة شئون البحث العلمي والدراسات العليا بالجامعة الإسلامية بغزة على تشكيل لجنة الحكم على أطروحة الباحث/ رامي عثمان محمد حسن سكيك لنيل درجة الماجستير في كلية **تكنولوجيا المعلومات** برنامج تكنولوجيا المعلومات وموضوعها:

فيديو متعدد المنظور باستخدام تقنية التخييط للفيديو
Free-Viewpoint Video (FVV) Based on Stitching Technique

وبعد المناقشة التي تمت اليوم الاثنين 30 صفر 1436هـ، الموافق 2014/12/22م الساعة الحادية عشرة صباحاً، اجتمعت لجنة الحكم على الأطروحة والمكونة من:

د. أشرف محمد العطار	مشرفاً ورئيساً
د. إياد محمد الأغا	مناقشاً داخلياً
د. ناجي شكري الظاظا	مناقشاً خارجيًا

وبعد المداولة أوصت اللجنة بمنح الباحث درجة الماجستير في كلية **تكنولوجيا المعلومات**/ برنامج تكنولوجيا المعلومات.

واللجنة إذ تمنحه هذه الدرجة فإنها توصيه بتقوى الله ولزوم طاعته وأن يسخر علمه في خدمة دينه ووطنه.

والله ولي التوفيق،،،

مساعد نائب الرئيس للبحث العلمي والدراسات العليا

أ.د. فؤاد علي العاجز

Deduction

To my beloved father

To my beloved mother

To my dearest wife

To my beloved brother and sisters

To my best friends

Acknowledgment

First and above all, I praise God, the almighty for providing me this opportunity and granting me the capability to proceed my studies successfully. This thesis appears in its current form due to the assistance and guidance of several people. I would, therefore, like to offer my sincere thanks to all of them.

I would like to express my deepest gratitude to my supervisor, Dr. Ashraf Alattar, for his excellent guidance, caring, patience, and for providing me with an excellent atmosphere to do the research.

I would like to thank my father for his financial and moral support during the years of my study. I also extend my thanks to my beloved mother for her love and prayers, and to my lovely wife for her support and patience during the years of my study. Furthermore, I cannot forget to express thanks to my big brother and my sisters, and I'm grateful to every member in my great family. Without my family I would not have been able to achieve anything.

I would like to take this opportunity to express my deepest gratitude to the academic staff of Information Technology Program at the Islamic University-Gaza. Specially for Dr. Tawfiq Barhoom, Dr. Iyad Alagha, Mr. Raed Rasheed and Mr. Ahmed El-aff.

Special thanks to all my friends who have directly or indirectly contributed to my success in completing this thesis, special mention for Mr. Omar El-Radie.

Finally, I offer my regards and blessings to all of those who supported me in any respect during the completion of the research, special mention for Zakaria Saqallah.

Table Of Contents

List of Figures

List of Tables

List of Abbreviations

MVV: Multi Viewpoint Video

FVV: Free Viewpoint Video

OpenCV: Open Computer Vision

3-D Video: Three-dimensional video

FTV: Free-viewpoint Television

MVC: Multi viewpoint coding

GPU : graphics processing unit

SIFT: Scale-Invariant Feature Transform

(PCA)–SIFT : Principal Component Analysis of Scale-Invariant Feature Transform

SURF : Speeded Up Robust Features

Abstract

*I*mage stitching is a technique used for creating one panoramic scene from multiple images. It is used in panoramic photography and video where the viewer can only scroll horizontally and vertically across the scene. However, stitching has not been used for creating free-viewpoint videos (FVV) where viewers can change their viewing points freely and smoothly while playing the video.

current research, implemented FVV playing system using image stitching, this system allows users to enjoy the capability of moving their viewpoint freely and smoothly.

To develop this system, user should capture MVV from different viewpoints and with appropriate region area for each pair of cameras then the system stitch the overlapped video to create stitched video/videos to display it in FVV playing system with applying freely and smoothly switching and interpolation of viewpoints over video playback.

Current research evaluated the performance of video playing system based on system idea, system accuracy, smoothness, and user satisfaction. The results of evaluation have been very positive in most aspects.

الملخص

تقنية "تخييط الصور" تقنية تستخدم في تكوين مشهد بانوراما من عدة صور. فهي مستخدمة في انشاء صور البانوراما او فيديو البانوراما بحيث يمكن للمستخدم التحرك في المشهد بشكل افقي او عمودي. لكن تقنية تخييط الصور لم تستخدم في انشاء الفيديو متعدد الوجهة الذي يسمح للمشاهد بتغيير وجهة الفيديو بحرية وبسلاسة (بدون أي تشويش أو تقطيع) خلال تشغيل الفيديو .

في هذا البحث ، سنقوم بعمل نظام تشغيل للفيديو متعدد الوجهة باستخدام تقنية تخييط الصور بحيث يسمح للمستخدم بالتمتع بمشاهدة الفيديو وتغيير وجهة الفيديو بشكل حر وسلس.

قمنا بعمل تقييم لجودة نظام تشغيل الفيديو بحيث قمنا بتقييم جودة عملية تخييط الصور و سرعتها ومدى رضى المستخدم بهذا النظام وحصلنا على نتائج ايجابية ومرضية.

Chapter (1)

Introduction

This chapter introduces current research by describing the fundamental concepts and technologies which represent the basis for understanding the thesis work presented, in addition to some details for terms. Finally, the chapter is concluded with an outline of the research problem, its significance and objectives. Furthermore, it presents the scope and some of the limitations that faced in this research.

Sometimes the image is not enough to show all details of a given scene as wants. Video can show more details that just images; also can show environmental changes, and interaction between the different objects of the scene.

Traditional video is a two-dimensional (2D) medium and only provides a passive way for viewers to observe the scene. In this case, if you need to show the video as a TV program show, the producer will control the viewpoint of the scene from any camera displaying. This means that the traditional media presentation is based only on one video stream. This video stream can be represented with one camera or clipped together with a content from multiple cameras. The red line of the traditional video that the user receives only one video stream.

With the continuous increase in computing power, technology, and the advances in multimedia technologies such as stereoscopic display, holography and 3D-video, it has become increasingly possible to break the viewer out of this controlled 2D Box into a more realistic 3D multimedia view. One possible technology to use is multi-view video (MVV).

MVV includes multi-viewpoint video sequence captured by several cameras at the same time, capturing the same scene at different camera locations.

If we acquire multi-view videos from multiple cameras, it is possible to generate scenes at arbitrary view positions.

Switching views seems to be the most desired function, and this can be considered as a key function for a multi-view video service. Also, the content for a multi-view video service should be entertaining and action-oriented to fully exploit the use of multiple videos, such as sports, music videos/concerts, live entertainment and reality- TV are desirable content for multi-view video.

In MVV the main motivation is to get entertained, to be able to better judge situations in the event and to get the feeling that you are present in the event.

FVV is one of the technologies based on viewing multi-view video, allows specification of the viewpoint at the point of rendering rather than the point of recording, and it breaks the restriction of rendering traditional fixed-viewpoint video of an event. The only viewpoint available for playback is the camera that recorded the event. FVV is technique of combining multiple video sources to generate a novel video from a virtual viewpoint.[1]

The trick-point in FVV is how to change the controlling of viewpoint of the scene from producer-controlling to user-controlling by allowing the user to change the video viewpoints from any viewpoint to another freely. It offers arbitrary viewpoints of dynamic scenes and thus provides a more realistic video for the users.

In our media world, taking panoramic pictures has become a common scenarios and it is included in the most of smartphones and tablets native camera applications. Panorama stitching applications work by taking multiple images, algorithmically matching features between images and, then, blending them together.

Image stitching used in panorama is the process of taking multiple images and "stitching" them together to form a single, seamless image. Video stitching uses this process to merge the frames of separate videos, resulting in one stitched video.

Most manufacturers use their own internal methods, which are very fast to stitching. There are also a few open source alternatives.

Stitching videos together has many applications in areas such as aviation, security, and other disciplines where a wide field-of-view is helpful. It can also be useful to the average computer user for personal needs like video lecture, to be able to view a large classroom or lecture hall, video chat, and etc.

Video and image stitching is an obvious solution to increase the amount of information an image/video could transmit to the user. Image stitching is already a well-studied problem with several good solutions (even commercial programs like Photoshop are capable of performing a really high quality image stitching), but there are not a lot of solutions for video stitching especially with the degree of freedom (during the capture), speed and performance of stitching video processes, and equipment capabilities needed to capture videos and other.

Next chapter gives literature review and reviews some related work that have attempted to solve the problem.

1.1 Problem Statement

Image stitching is primarily used in constructing still panoramic scenes and fixed viewpoint panoramic videos. However, it has not been used in free viewpoint videos (FVV) to construct video scenes where viewers can change their viewpoint freely during play time. The problem, therefore, is how to implement Image stitching in FVV to achieve real-time and smooth free viewpoint video playing.

1.2 Objectives

1.2.1 Main Objective

To develop a video playing system that uses stitching technique for free view-point video that allows user to stitch overlapped videos and display it with free and smooth move between any view-point in FVV.

1.2.2 Specific Objectives

To achieve the main objective, the following set of specific objectives shall be completed:

- To develop FVV video playing system based on stitching technique.
- To evaluate the usability and functionality of the FVV system.

1.3 Significance of the research

In traditional videos, the director of the film is the controller of the video viewpoint if the scene is filmed from multiple cameras, so the user displays the scene after producing and editing the video without any control over the scene. Current research allows users to real-time interactively control and freely change the viewpoint and view-angle of his/her video, using the stitching technique. By using the stitching technique, all the information needed to generate the viewpoint (actual and virtual) is available and can be displayed in the FVV playing system.

The user can also use this system to generate a stitched video file from multiple videos that are overlapped, and it gives the ability to choose the way to display the video: frame by frame if the user has only frames, or display video if he has generated video.

Overall, current research attempts to provide a better user experience in viewing and playing video in 3D vision.

1.4 Scope and Limitations

Current research is important for the people who are interested in the multimedia field especially in 3D video and FVV to help them show the video from different viewpoints in order to show more details of the scene like tourism scenes, sports, video conferences or games. Thus, anyone interested in media production can use this system.

In addition, Current research supports the people interested in panorama films, and the people producing commercial video products to create 3d video presentation.

1.4.1 Limitations:

During working in this research specially in implementation phase , we faced some limitation:

- Capturing multi-viewpoint video needs very complex equipment (camera controller, tripod, high resolution cameras, and etc.
- OpenCV libraries implemented to stitching images not for video stitching.
- The stitching technique takes long time.
- Image processing needs equipment with high quality to give a good performance results.
- OpenCV library with visual studio do not support GUI components.

Chapter (2)

Literature review

This chapter, given background of current research. Firstly, it describes multi-view video technique and its technologies. Then, it moves to describing Free-view point video (FVV) and its applications and FVV system components. It also describes the image/video stitching techniques. Also, different related works are studied and investigated. The related works are introduced and analyzed with respect to research problem in order to show how far these works address the requirements of current research problem. We divided this related works to three categories; works related to image stitching technology, works related to video stitching technique, and works related to FVV systems and technology. For each category, we present some related works and discuss the difference between these works and current research.

2.1 Background

This section provides a state of the art review of the available technologies in MVV (free-viewpoint television (FVT), immersive teleconference, omnidirectional video, and 3DTV), also it describes FVV technology and its application and it's system components. Finally, it presents Stitching technique with two parts image and video stitching.

2.1.1   Multi-viewpoint Video (MVV)

MVV is a collection of multiple videos, capturing the same scene at different camera locations. If we acquire multi-view videos from multiple cameras, it is possible to generate scenes at arbitrary view positions.

It's a new technique which can provide the user with the benefits of added realism, selective viewing, and improved scene understanding. It means that users

can change their viewpoints freely within the limited range of captured videos and can feel the visible depth with view interaction[2].

2.1.1.1 MVV Technologies

The multi-view video is the key technology for various applications and technologies as shown in Figure 2-1, including free viewpoint video (FVV), free-viewpoint television (FVT), immersive teleconference, omnidirectional video, and 3DTV.
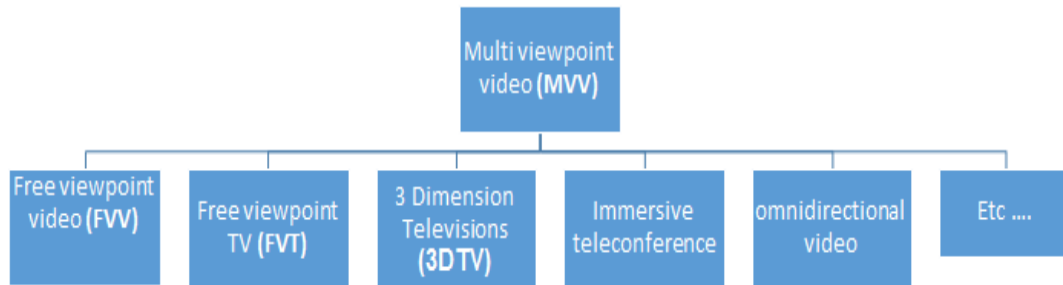
Figure 2-1 Multi-view video applications

The next section we present the most promising techniques that use multiple video streams and some of these techniques were illustrated in Figure 2-1.

3-D Video

The most common number of views for multi-view video (MVV) is two, which can be used to experience 3-dimensional video (3DV). The ability to perceive depth scenes from a 2-D representation format is a quite old technique.

3D video based on stereoscopic principle is based upon providing the user two separated images; one for the right eye and one for the left. These images are captured from slightly different viewing positions.

Another requirement is that the user must use polarized glasses which filters out the images to each eye. This principle generates the illusion of depth or three-dimensional imaging.

There are some drawbacks for traditional stereoscopic imaging. First, it has low acceptance is the usage of glasses, which can be perceived as exhausting and as a constraint. Second, it lacks interactivity, and viewers get the same static 3-D image if they move with respect to the display.

Omnidirectional Video

In this technology the scene represents with multiple cameras in a way that mostly of the spherical field is captured in high resolution. This enables the user, with the proper software, to zoom and rotate around the cameras. This can provides the user with the feeling of being a part of the scene. However, in contrast to free viewpoint video, the user is not able change the position of the viewpoint interactively. The viewpoint may change but this requires that the cameras have been moved during capturing [3].

Omnidirectional camera can be used to create panoramic art in real time, without needing for post processing, and it will typically give much better quality products[4].

An example of an omnidirectional video is shown in Figure 2-2.

Figure 2-2 Omnidirectional camera and a corresponding spherical view.[5]

Free-viewpoint TV (FTV)

FTV is an innovative visual media that enables us to view a 3D scene by freely changing viewpoint. It is easy to achieve the free viewpoint for virtual scenes made by computer graphics. Achieving such a function for real scenes will bring an epochal change in the history of visual media.

FTV is the ultimate 3DTV with infinite number of views, and it's ranked at the top of visual media devices.

The international standardization of FTV has been conducting in MPEG. The first phase of FTV was MVC (Multi-view Video Coding), and the second phase is 3DV (3D Video).

There are many process configurations in FTV system including color correction, MVC encoder and decoder, and generation view and display. These processes have illustrated in Figure 2-3.

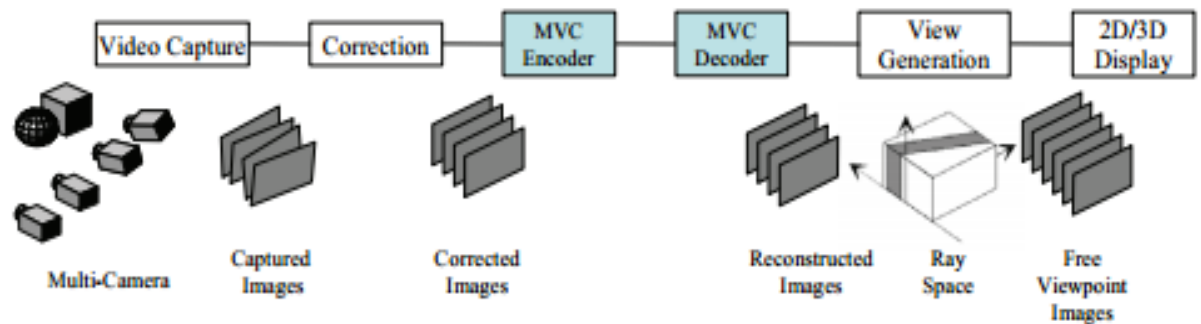Figure 2-3 Basic configuration of FTV system [5]

FTV can find many applications in the fields of broadcast, communication, game, entertainment, advertising, exhibition, education, medicine, art, archives, security, surveillance, and so on[5].

Immersive Teleconference

An immersive video is basically a video recording of a real world scene, where the view, in every direction, is recorded at the same time.

During playback, the viewer can control the viewing direction, up, down, and sideways. Generally, the only area that can't be viewed is the one towards the camera support. The material is recorded as data, when played back through a software player, allows the user to control the viewing direction and playback speed[6].

In immersive teleconference, the imagery is meant to be a piecewise collection of images, together, forming a panorama or a wide-area image. This technique is illustrated in Figure 2-4.
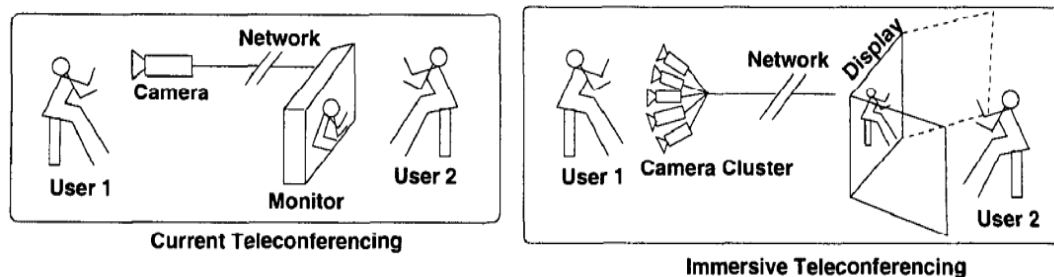
Figure 2-4 Difference between current and immersive teleconferencing [6]

2.1.1.2   MVV Presentation

Current research focuses on the view generation and presentation of the MVV scene. Traditional video representation is, in many cases, good enough, but for special interactive or entertainment-oriented applications, it has some limitations, as in a live lecture where the user might like to control the viewpoint. The fixed viewpoint makes the user's interaction limited or not present, and it puts the user in a passive position. Users can only watch the event from a video sequence that is not user-selected.

An extension of single-view video is a multi-view video presentation which is generated from multiple cameras and recorded synchronously with different viewpoints. The viewer receives multiple video streams and can, then, enjoy the video from his/her desired view by easily switching between the views. User's interaction is considered a key function in this kind of multi-view video system[7].

The system structure of current research is based on the idea of how the system can generate a video playing system to present MVV as FVV system by allowing the user to control his viewpoint when playing back the video.

For more details about FVV, the next section gives a description of FVV, presents the applications of FVV, and shows the components needed to generate an FVV system.

2.1.2 Free Viewpoint Video (FVV)

In traditional videos and films, events are recorded from a single fixed viewpoint. These confine the viewer to the fixed linear format dictated by the director and at the two-dimensional (2D) viewing experience. Free-viewpoint video breaks this restriction by providing three-dimensional (3D) content with interactive control of the viewpoint in visualization[8].

Free viewpoint video (FVV) is one of the applications that is used for multi-view video representation. Multiple cameras are used to capture a scene. With techniques from computer vision, these synchronized video streams can be transformed into a data representation that allows the user to freely choose or change his/her viewpoint (i.e., viewing position and viewing direction).

FVV is multiple video streams providing freedom of viewpoint in the video. It provides the user with realistic impressions by means of high interactivity and photorealistic image quality[9].

With FVV the focus of attention can be controlled by the viewer rather than the director. This means that each viewer can be observing a different viewpoint[10].

Multi-view video signal used in FVV, captured by cameras arranged in parallel, convergent, or any angle at distinct viewpoints, provides the information of multiple views of a scene.

There are two main technical areas involved in FVV: multi-view video coding and view generation. Multi-view video coding is used to code the video signals captured by two or more cameras. View generation technology is used to generate image views for display, which may not be captured by a camera.

Current research focused on the second area – view generation – using stitching technology, which will be described in more details in this next Section 2.1.3. Also, current research developed a system to display the FVV. More details about the system will be presented in Chapter 3.

2.1.2.1   FVV Applications

Free viewpoint is based on the user interactively changing viewpoint and view direction within the range covered by the shooting cameras. Such scenario can appear in the following applications[11]:

a) Entertainment–concerts, sports, multi-user games, movies, drama, news

b) Education–cultural archives, manuals with videos, instructions of playing sports, medical surgeries, and so on.

c) Sightseeing–zoos, aquariums, botanical gardens, museums, etc.

d) Surveillance–traffic intersections, underground parking, banks, etc.

e) Archive–space archives, living national treasures, traditional entertainment, etc.

f) Art/Content–creation of new type of media art and digital content.

2.1.2.2   FVV System Components:

FVV system contains multiple components, which are almost similar to FTV system components. In this section, the components are divided into three stages as shown in Figure 2-6. We briefly described the process under each stage and what stage current research was focused on.
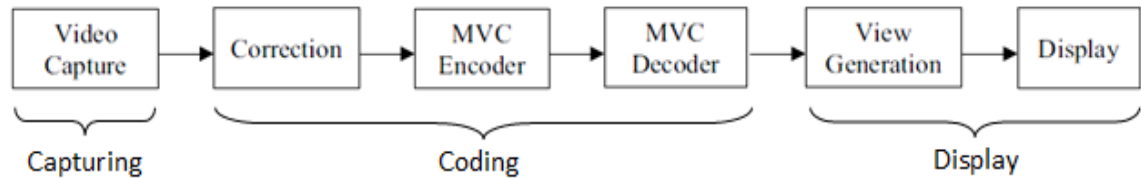
Figure 2-5 Basic components of FVV system [11]

Capturing stage:

In the capturing stage, there are many different ways to capture frames for FVV system. These ways differ according to their purposes. The frames can be capture by the same camera, but in different places, or by several cameras in different positions and angles, so that all the cameras capture the same scene at the same time.

The first way can be used for creating still like creating 3D objects. The second can be used to create free viewpoint video. For example, it can create a dynamic video that allows the user to freely change his viewpoint when viewing the video.

Current research, implemented a video playing system to display the video based on FVV. For the purpose of our research, we used the second way to capture the video using multiple video cameras to capture the scene from multiple viewpoints at the same time.

Coding stage:

The captured images contain the misalignment and luminance differences of the cameras [12]. Therefore, in the coding stage shown in Figure 2-5, we need a correction process to rectify the misalignment and normalize the colors before compressing and encoding.

Moreover, some processes may be done after correction and before coding such as describing the depth value for each frame, combining the camera

calibration parameters, finding the geometry of the object, and finding a color value. In addition, there are methods such as shape-from-silhouettes method[13]that is based on the image-based visual hulls algorithm to extract the depth information. For each pixel, there are some attributes that can be computed and stored as an option such as a surface normal vector, a splat size, and other attributes.

MVC encoder is used in the coding stage to exploit both spatial and temporal redundancy for compression. Since all cameras capture the same scene from different viewpoints, inter-view redundancy is present[14]. Thus, there are a lot of algorithms and MVC standers to solve this coding problems.

The coding method for FVV should address requirements like scalability, random access, and interactivity. In addition, the codecs must enable encoding of all attributes i.e., color, depth, optional normal and splat size. A codec should also be able to decode multiple multi-attributed video frames in real-time[13].

The final part of coding stage is MVC decoder, it provides reconstructed video data, which is used in the view generation process. It includes multi-view video elementary information, video resource management information, timing information, and camera parameters information.

The video resource management information may be used for managing the picture's memory in an efficient way and for generating predictive images for the MVC decoder[11].

Display stage:

The output images from the MVC decoder are used for FVV view generation. This process interpolates images from different views. View generation is one of the most time consuming parts at the user's side in free

viewpoint video (FVV) system. It focuses on the requirement and capability of user side.

View generation should be fast, so that real-time navigation can be achieved. Additionally, the quality of synthesized view should also be high enough to support the realistic impression.

A novel view from an arbitrary position and direction can be generated by appropriately combing image pixels from the existing views[15].

Current research focus on the third technical stage (Display), which focuses on achieving high-quality view generation scene from multiple views, and then display it for free interpolation. View generation module will render the captured scene in real-time using the stitching technique to create the final view.

In current research, generating the view can be done using offline subsystem for performance issue, then it can be displayed in the final system, allowing the user to change his viewpoint when playing back the video.

2.1.3   Stitching Techniques

In this section, we present and discuss the techniques of stitching. These techniques are divided into two categories; image and video stitching, which we will present in more details.

2.1.3.1   Image Stitching

Image stitching or photo stitching is the process of combining multiple photographic images with overlapping fields of view to produce a segmented panorama or high-resolution image[16].

The techniques of stitching images into a large photo-mosaic are some of the oldest and most used in the field of computer vision. These techniques have a

wide range of application; from interactive panoramic viewing of images, architectural walk-through, multi-node movies, and producing high resolution maps collected from satellite imageries, to creating ultra-wide-angle panoramas for the casual user equipped with only a digital camera. Image stitching technology can be used in parking issues, computer vision, medical imaging, military automatic target recognition, and compiling and analyzing images as well as data from satellites [17].

Image stitching is divided into four main steps[12]:

1. Choose the mathematical model relating pixel coordinates in one image to pixel coordinates in another image
2. Align the images by estimating the relationship between various pairs
3. Choose a final compositing surface for warping the aligned images
4. Seamlessly cut and blend overlapping image

Panoramic image stitching is the process performed to generate one panoramic image from a series of smaller, overlapping images.

The main steps for producing a panoramic image are illustrated in Figure 2-6[18]:
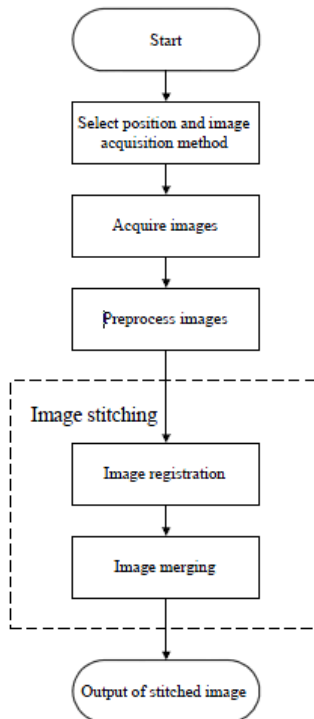
Figure 2-6 The flowchart of producing a panoramic image [19]

In Figure 2-6, the first step in the generation of a panoramic image is to select the position and acquisition of images. In this step, a decision needs to be made on the type of resultant panoramic images. According to the required panoramic images, different image acquisition methods are used to acquire the series of images.

After acquiring the images, some processing might be required to the images before they are stitched. For instance, the images might need to be projected onto a surface, which can be a mathematical surface such as a cylindrical, spherical, or planar surface. Distortions caused by the camera lenses also need to be corrected before the images are processed further.

The process of image stitching is divided into two steps; image registration and image merging. During image registration, portions of adjacent images are compared in order to find the translations which align the images.

Once the overlapping images are registered, they need to be merged together to form a single panoramic image. The process of image merging is performed to make the transition between adjacent images visually undetectable.

A panoramic image is generated after the images have been stitched. By generating panoramic images with image stitching, the images can be acquired using a relatively inexpensive camera, and the angle of view covered by the panoramic image can be determined by the user. The stitched image can also be of higher resolutions than a panoramic image acquired by a panoramic camera.

2.1.3.2 Video Stitching

Video stitching is the process to merge the input image sequences into one high resolution panoramic video with wide view field. It is based on image stitching technology.

The video stitching process includes the following four steps, as shown in Figure2-7:
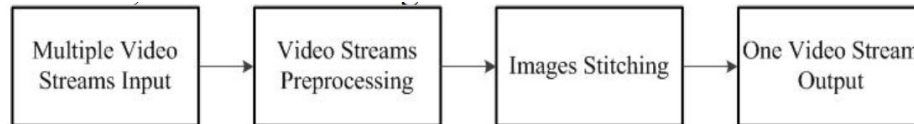
Figure 2-7 Video stitching procedure[17]

In video stream preprocessing stage, the distortion lens of wide-angle shot can be corrected. Some smoothing processes are also carried out in this stage. The image stitching technology is the key part of the whole process. Image stitching refers to an image processing technology which is used to align a group of overlapping image sequences, producing an integrated high-quality image including a wide view of all image information[17].

Video Stitching has a wide range of application such as visual surveillance, human--machine interaction, video processing and editing, and remote video conference.

Also, there are many methods used in the video stitching these methods can be categorized from two standpoints according to the application cases of video

stitching. The first standpoint is to apply video stitching approaches by manipulating the source image sequence from a single camera or multiple cameras. The second standpoint is to utilize video stitching systems for offline or online processing[19].

To achieve a good panoramic video from two or more separate video streams, this video should be seamless. This means there should be no traces of the placement or existence of the input videos because a good panoramic video in real-time requires that the output video should have the same frame rate as the input videos. Additionally, there should be no noticeable differences in exposure or lighting in each stitched frame of the output video. In other words, there should be a unified exposure. Finally, a "good panoramic video" should be stitched as if the videos are in-sync. This means that each frame of the output video should be comprised of frames from the input videos that were taken at the exact same point in time [20].

2.1.3.3   stitching algorithms

To generate stitched frame from two pair of frames, There are two classifications of stitching algorithms: direct (pixel-based) and feature-based alignment algorithms. The Feature-based approach has advantage because it is more robust against scene movement, and it is potentiality faster. Moreover, it can automatically discover the overlap relationship among images[21].

There are a lot of algorithms used in feature-based like Scale-Invariant Feature Transform (SIFT), Principal Component Analysis (PCA)–SIFT, and SURF. SIFT is a successful approach to extract distinctive invariant features from images that can be invariant [22]. Principal Component Analysis (PCA)–SIFT reduces the execution time of SIFT matching, but it was proved to be less effective than SIFT in extracting the feature points[23]. It is also more effective in rotation and illumination changes. SURF-algorithm[24] is based on the same principles and

steps of SIFT, but it utilizes a different scheme. It provides better results, it is faster[21][23], and it gives high performance results.

In current research "Stitcher" class in OpenCV was used to generate stitching frames. This class used SURF algorithm for stitching, so it is very useful to use it in current research. Also, This class makes many processes to stitch multi images in one stitched image. These processes are illustrated in the Figure 2-8.
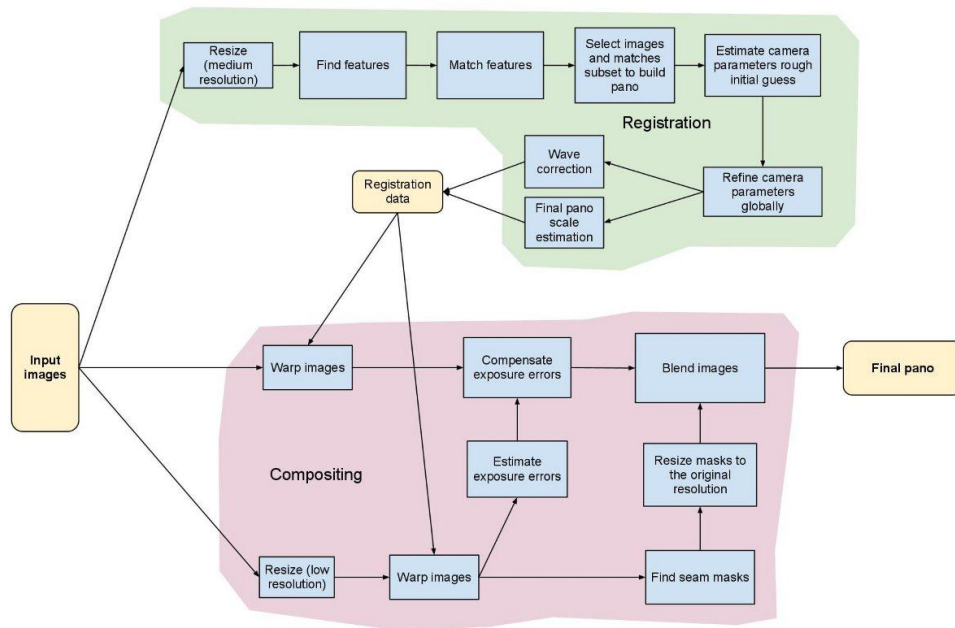

Figure 2-8 Stitching module pipeline implemented in the Stitcher class [25]

2.2 Related works

2.2.1 Image Stitching:

This section presents some related works that used image stitching in panorama field. We present and discuss the output of these works and define the difference between this works and current research.

M. Brown *et al.* [26] described a system they built for stitching still images that used the image stitching technique called feature matching. It gives an effective solution for the problem of automatic image stitching. Automatic image stitching implies that, unlike many current solutions, there is no initialization or user input required before stitching. In attempts to produce outputs that are more accurate as well as aesthetically pleasing, they viewed their problem as a multi-image matching problem. In doing so, their solution is insensitive to ordering, orientation, scale, and illumination of the input images.

In addition, they achieved outputs that eliminated noise through a technique called multi-band blending, which produces panoramas of a good quality.

Figure 2-9 Images registered [26]

Figure 2-10 Result With gain compensation and multi-band blending [26]

To solve their image matching problem, Brown and Lowe used feature matching and bundle adjustment. Once the images are matched, they implement automatic panorama straightening, gain compensation, and the aforementioned multi-band blending.
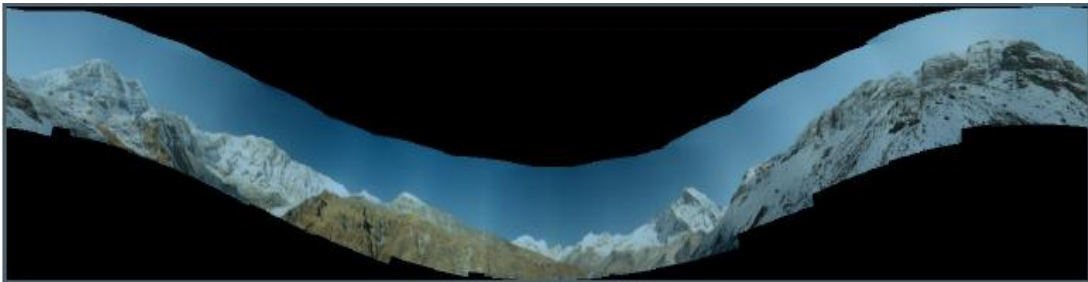
(a)

(b)

Figure 2-11 image result of straightening (a) with automatic straightening and

(b) without automatic straightening [26]

The system they built performed seamless automatic image stitching, but took 15 minutes to generate the output panorama although the work done by Brown and Lowe is inspiring in the quality of the output.

Current research created real-time video system, so we need something faster than this solution if we notice that the output video has matching mistakes,

seams, etc. Moreover, current system has a new functionality to the user to freely change viewpoints of the scene during watching the video.

A. Agarwala et al. [27] proposed a system for producing multi-viewpoint panorama for long scene. The user's required input goes beyond capturing the photographs themselves in order to identify the dominant plane of the photographed scene, then, the system will compute a panorama automatically using Markov Random Field optimization.
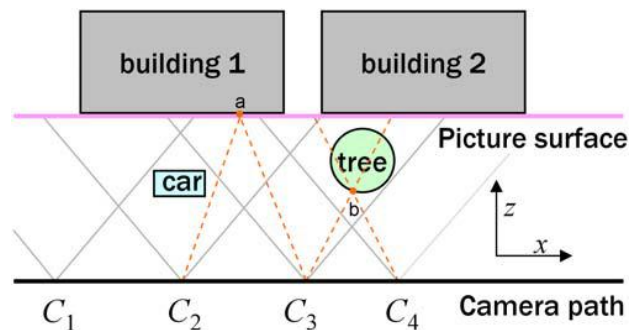
Figure 2-12 plan view (xz slice) of a hypothetical scene captured with four photographs [27]

For creating a multi-viewpoint panorama from a sequence of still photographs, the system has three main phases. First, preprocessing stage: taking the source images and removing radial distortion, recovering the camera projection matrices, and compensating for exposure variation.

In the next stage, the user defines the picture surface on which the panorama is formed. The source photographs are then projected onto this surface.

Finally, the system selects a viewpoint for each pixel in the output panorama using an optimization approach. The user can optionally choose to interactively refine this result by drawing strokes that express various types of constraints, which are used during additional iterations of the optimization.

Figure 2-13 multi-viewpoint panorama composed from 107 photographs [27]

The output is a single panorama that visualizes the entire extent of the scene captured in the input photographs and resembles what a human would see when walking along the street.
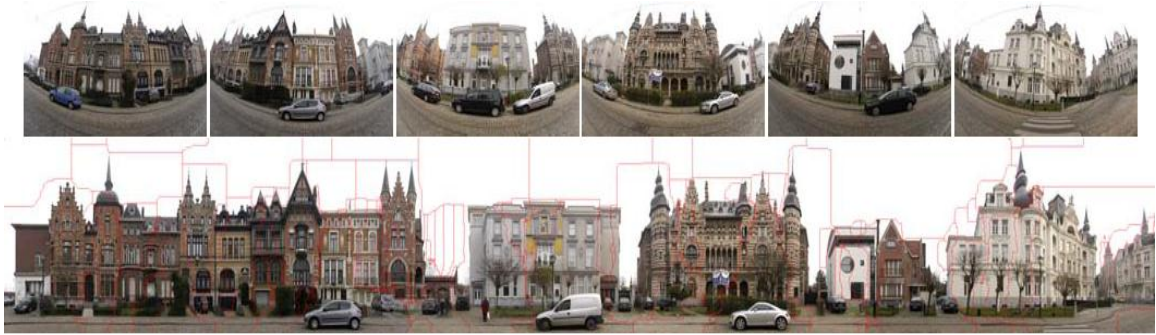
There are some problems in this work as it is not able to produce effective multi-viewpoint panoramas for every scene like suburban scenes.

This work used image stitching technique to create long scene. The output is a single panorama photo that contain all images, but this system does not support video viewing and doesn't allow user to move from one viewpoint to another.

Current research used image stitching method to generate view step, but for creating free viewpoint video and not for creating only image panorama scene.

M. Adam et al. [28] Produced system that stitches multiple High-definition videos together at a respectable (more or less real-time) frame rate. To do so, they applied a new algorithm for the video stitching that encompasses many different techniques for accuracy. They can reach a very efficient output level using the multiple cores the graphics processing unit (GPU) has to offer by the use of parallel computing, but it relied heavily on the capabilities of a distinct GPU hardware exploiting the nVIDIA CUDA Framework. By using the GPU splitting up the various tasks of the algorithm for analyzing and altering the incoming video stream, they state the GPU approach is about 40-50 times faster than the not optimized, single core. The use of the GPU enabled speed without compromising the quality of the output video.

Current research, developed an FVV system for the average user, so it doesn't need an expensive equipment. Their research requires the GPU to function in real-time, but current research cannot adopt their stitching approach. However, current research, make FVV playing system, and their research does not cover this functionality. If we want the system to be more efficient and make the stitching function very fast, we can use it in the future, but in this case, the user should have GPU.

2.2.2 Video Stitching:

This section presents some related works that utilized video stitching technique in its system. These works are based on mutable camera videos stitched together for different field. We present and discuss this systems and define the difference between these works and current research.

S. Segv et al.[29] conducted an experiment to stitch three videos together acquired by three cameras placed in a horizontal plane on the top of the service vehicle in a way that the fields of view of the lateral cameras overlap with the field of view of the middle camera.
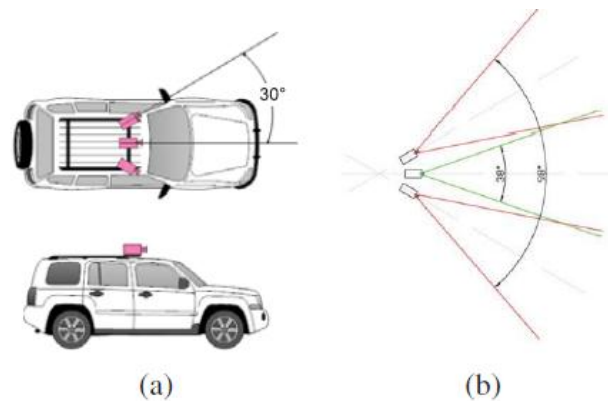

Figure 2-14 The EuroRAP image acquisition platform is shown on the left (a), The angular configuration of the three cameras is shown on the right (b). [29]

In this experiment, the required mappings are estimated by aligning the common parts of the three views in correspondence with the video frames. They have

obtained encouraging results which would clearly be quite useful in assisting road safety inspection procedures.

In this experiment, the system used stitching-video technique to stitch three videos together to create one complete scene without giving the ability to the user to change his viewpoint or to freely move between these viewpoints.

In current video playing system ,we used stitching video, but with the ability to freely change and move between the viewpoints.

Ying Zhang et al[17]**.** proposed a safety parking system based on video stitching technology, through which the driver can see the surrounding environment. Panorama parking aided systems installed four wide-angle camera around the car, collecting multiple video images at the same time for processing to make 360 degree top view of a vehicle around the car body. The image is displayed on the screen for driver, allowing the driver to see the vehicle's surrounding obstacles clearly, helping the driver park safely as shown in Figure 2-15.

Figure 2-15 Top view for ideal panoramic result parking[17]

In this research, the researchers suppose that the core of the image stitching technology is image registration and image fusion technology. The quality of the selected algorithm in these two processes will affect the entire mosaic effect.

In this experiment, they used Speeded Up Robust Features (SURF) registration algorithm and linear fusion algorithm.

Figure 2-16 experimental results and analysis[17] (a) Input 4 original images.

(b) final image after registration and re-matching processes

As we see in this research, the proposed system used image stitching technology. The image will be displayed on the screen for driver, but without allowing driver to change the viewpoint. Current research used image stitching technique in FVV, allowing the user to change his viewpoint to be more realistic.

H. Zhang, et al[19]**.** In this research, the researchers proposed an approach designed to merge videos from different viewpoints into one high resolution video in real time, which conducts borderline adjusting based on the foreground information. This research addresses the problem of dynamic scenes, where foreground objects often cause broken objects like artifacts in panoramic video. Accordingly, they propose a method to achieve smooth video stitching.

The target of this paper is online multi-camera video stitching for dynamic environments.

This approach focuses on minimizing the quality depredation of foreground objects because the quality of foreground objects is critical in succeeding applications such as object tracking and face recognition.

Extensive experiments on indoor and outdoor scenarios are conducted to evaluate the effectiveness of the proposed multi-camera video stitching method.

The proposed system contains three parts: image acquisition, geometry calibration, and video frame stitching as shown in Figure 2-17:
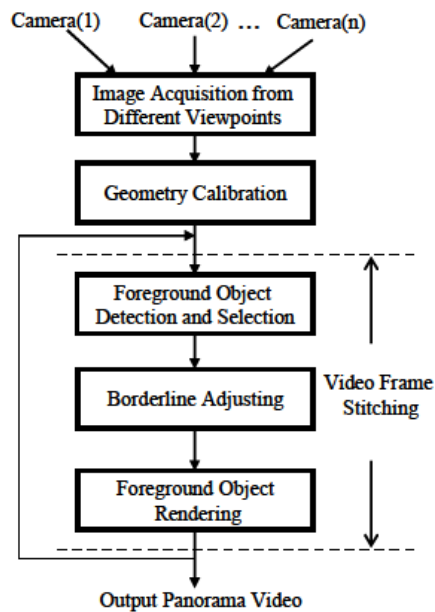
Figure 2-17 Process of the proposed multi-camera video stitching approach[19]

The experiment results of this research show that this approach obtains more than 81% correct rate for dynamic scenes, compared with 25%-50% correct rate on the condition of not utilizing foreground information, and it achieves processing speed of 10~13 frames per second.

In this research, the proposed system uses a method for video stitching for dynamic scenes, but it is a fixed viewpoint video. My proposed system, however, supports multi view video and allows the user to freely change his/her view point.

2.2.3 Free View point:

Lou et al[30]. presents system architecture for real-time capturing, processing, and interactive delivery of multi-view video. In this system, users can interactively select their desired viewing directions and enjoy many exciting visual experiences such as view switching, frozen moment, and view sweeping in real time.
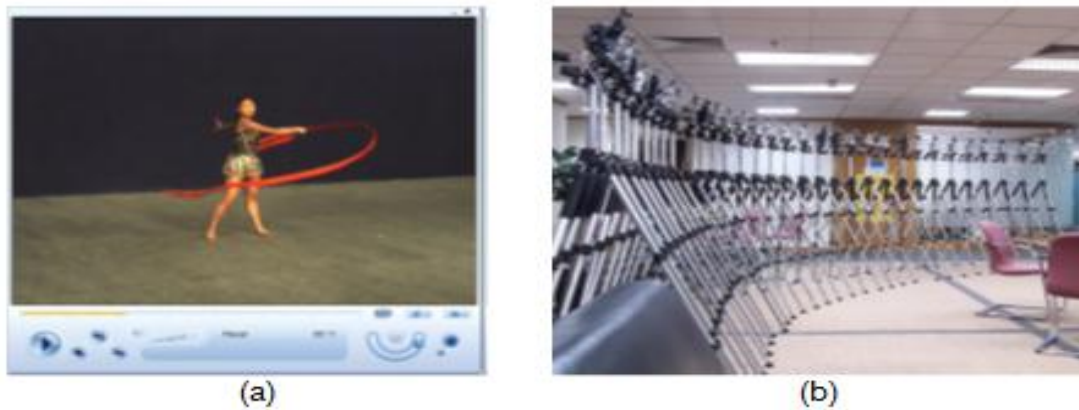

Figure 2-18 User interface of a remote client shown on the left (a)

A deployment of authors system shown on the right (b) [30]

In this research, four main contributions were achieved: first, a pattern free calibration algorithm is proposed to dramatically reduce the workload for calibration in a large environment space. Second, a computational complexity scalable video encoder is proposed for real-time video capturing with the ability of adjusting the encoding speed according to the computing resources, and thus, better complexity adaptation and system scalability are achieved. Third, the multi-view video contents are re-organized and re-encoded into a frozen moment stream and a view sweeping stream. Hence, the most exciting viewing experience can be provided with low bandwidth consumption. Fourth, a streaming mode and a broadcast mode are proposed for delivering a multi-view video in different application scenarios while unique features of multi-view video are still preserved.

In this research, N video streams, captured by the system, are transmitted respectively from the control PCs to the server. The server will then process these N streams and provide real-time interactive multi-view video service to the end users.

Current FVV system is offline processing, which means it does not need a server for streaming multi view video.

Another different point is that the system of current research uses stitching technique to generate displayed views, but this system does not use it.

J. Starck, et al[31]*.* presents a system to synthesize free-viewpoint video from multiple- view video streams together with a time-varying geometric proxy for a scene. The technique is implemented on the GPU for video-rate view synthesis. The research focuses on the multiple camera acquisition systems and the computer vision algorithms required to recover 3D scene geometry and perform virtual view synthesis, either in real-time or as an off-line post-process.

Figure 2-19 Image-based rendering[31]

Free-viewpoint visualization is achieved by rendering a 3D scene model to a virtual viewpoint with the appearance sampled from adjacent camera images.

The input of the render technique is a set of real world video images and a 3D proxy for the scene geometry along with the camera calibration defining the projective transformation from the scene coordinate system to each camera. A virtual view is synthesized by compositing the appearance sampled from the

cameras closest to the virtual camera. At the point of view synthesis, the algorithm first selects a subset of cameras to use in rendering. The surface visibility is then computed in each camera to prevent the appearance of sampling in the presence of occlusion.

In this system, researchers use MVV scenes to create 3D object and system for controlling viewpoints of this object in real time. The system of current research used MVV to capture multiple video file and display it with controlling the viewpoints in playback, so The system of current research produce 3D view for video scene.

Chapter (3)

Methodology

This chapter, presents current research methodology. It discuss the idea of current research and how understood and collected research information. Then, this chapter discuss in details the proposed method of The system of current research and illustrate the steps in Figure 3-1. Finally present the evaluation method of current research.

Current research uses the image stitching technique to develop video playing system for FVV technology. This allows the user to change his/her viewpoint in real-time while watching video scenes. The tricky point in FVV is changing the control of viewpoint in video from producer control to user control.

In traditional videos the scene are filmed from multi camera shots in different positions and different viewpoints, then the producer composites these shots to produce one scene from multiple viewpoints, but the user watches this video without the ability to change the viewpoint or direction during playback, but in FVV technology, the video is filmed from multi cameras, but the controlling of display viewpoint will be at runtime based on the user's wish.

To allow the user to smoothly and freely change the viewpoints in real-time, we used image stitching technique. By using stitching in FVV, all information needed to generate any virtual viewpoint over watching video will be given, so it allows the user to freely and smoothly change his/her viewpoint without any distortion or shredder over interpolation.

In the next section, we discuss in details the steps of current system.

3.1   Proposed Method:

In this section, we discuss the method's steps for generating FVV playing system based on image stitching technique. These steps are divided into four steps that are capturing videos, generating stitching frames, generating stitching frames from video, and displaying FVV playing system: first step is preprocessing step before using current system. It is used to capture appropriate video files using on current system. The other three steps for displaying FVV begin from generating stitching frames from video files, then generating stitched video for all scene, and finally, displaying the video in FVV system player. These steps are illustrated in Figure 3-1 and discussed in details in this section.
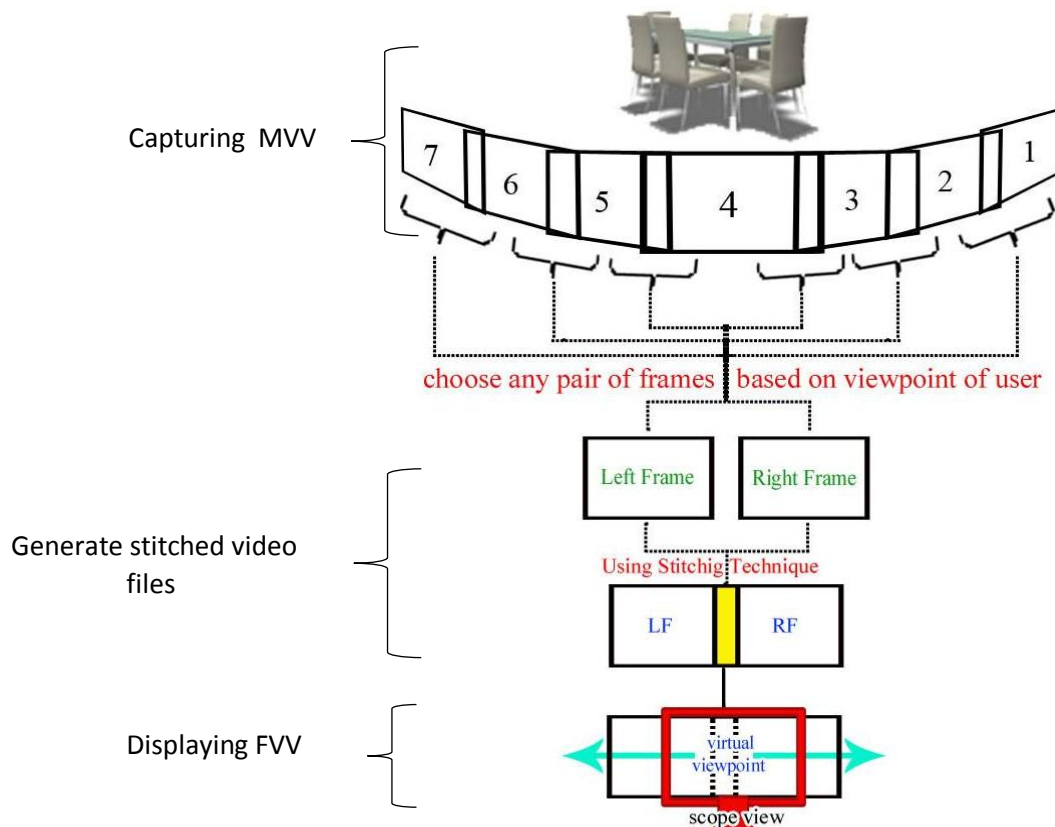

Figure 3-1 proposed video playing system for FVV based on Stitching technique

3.1.1 First step - Capture multi view video (Preprocessing):

Before using current system, the user should film a number of different viewpoint videos for a specific scene. The cameras must be synchronized and fully calibrated prior to the capture session and placed in different viewpoints and different angles. They, also, must contain overlapped regions between camera shot frames, as we see in Figure 3-1. This overlapped region is needed for the stitching phase.

Current system, receives sequences of frames for each video of multi-view video frame by frame. In our example illustrated in Figure 3-1, there are seven overlapped cameras that give seven video files, so the system received seven sequences of frames for each video.

3.1.2 Second step – Generate stitched video files:

In this step, the user gives the system an order to generate a stitched video for video files received by system in step one. In this step, virtual viewpoints that are not captured by cameras will be generated. The virtual viewpoints will be generated using two actual viewpoint frames (left frame "LF" and right frame "RF") as we see in Figure 3-1.

There are two ways that can be used in current research to generate video file. First, the system generate multiple stitched-files from each neighbor pair of video frames received in the system (video 1 with 2, video 2 with 3, and so on).

The second way is to generate one stitched-video file for all videos captured – not for only two pairs – this is based on number of video files. On one hand, it may take more time in stitching phase than the first way, and it may cause a crash in memory if there is not enough memory space in the PC. On the other hand, this

way gives better performance in display step because the system will display only one file.

(In our example in the Figure 3-1, we have 7 videos, so according to the first way, the system will generate (n-1~7-1=6) stitched-video files where n is the number of video files received, but in the second way, we get only one stitched-video file).

3.1.3 Third step: Display the stitched video with FVV:

In this step, the user gives an order to the system to display the stitched-video generated in second step. Since there are two ways to generate the stitched-video file/files, we have two ways to display it based on the generating way. If the system generated multiple stitched-videos for each pair, the system will display the appropriate stitched-video based on the viewpoint of the user, but if the system generated one stitched-video for all video files, the system will display this stitched video only, and in display window, only a part of this stitched-video will be visible to the user based on his viewpoint.

Finally, after displaying one of viewpoints of the stitched-video file/files in the system, the user can now freely and smoothly interpolate between any viewpoint using scroll-bare available in video playing system. When the user moves the scroll-bar left and right, the system will display the part of stitched-video, and the viewpoint will change according to the scroll-bare movement.

In the next chapter, will present research implemented experiments and describe the results in details.

3.2 Evaluation :

Evaluation is an important component of refining programs and documenting impacts. Evaluation aids the profession as a whole and assists Extension faculty in meeting promotion requirements. There are two type of evaluation: Qualitative and Quantitative. quantitative methodologies will not provide practical and in-depth information often needed for program improvement, However, Qualitative methods are commonly used in evaluations in order to explore specific facets of programs and to give voice to participants' experiences [32]. So current research used qualitative evaluation to evaluate the FVV system based on usability principles. Usability is the extent to which users can use a computer system to achieve specified goals effectively and efficiently while promoting feelings of satisfaction in a given context of use.

Current research used questionnaires to evaluate the usability. "A questionnaire is a method for the elicitation, and recording, and collecting of information" [33].

Because of, there are no appropriate previous works can compare it with current research system we used usability, functionality, and user experience evaluation with appropriate questionnaire to caver research objectives.

Usability evaluation is itself a process that entails many activities depending on the method employed. Common activities include [33]:

- Capture - collecting usability data, such as task completion time, errors, guideline violations, and subjective ratings;
- Analysis - interpreting usability data to identify problems in the interface; and
- Critique - suggesting solutions or improvements to mitigate problems.

Measurement Instruments and factors:

The instruments current research used to do in the evaluation of the free viewpoint concept:

- Observation of user behavior observation during the test
- A detailed user questionnaire

the factors that could affect the user experience and his/her degree of satisfaction, these were grouped as follows:

- User personal and situational factors
- Overall enjoyment: How does he/she feel about the product
- Usability: easy to use and learn, simple to understand and control, appealing, coherent and quick.
- Other aspects not taken into account: user's suggestions about aspects that we did not think about a priori.

Chapter (4)

Implementation and Results

This chapter, presents the implementation of current system. It also describe the tools used in the system, programing language, and libraries needed. Then, it present and discuss, in details, the functions of our project and how the user can use it. For each function, this chapter present the internal processes, how it is done, and an example for output results.

4.1 Tools and system architecture :

Current research, implemented an FVV system using visual studio environment with OpenCV, which stands for Open Source Computer Vision. It is a framework written in C++, Java, Python, and Matlab that allows a high-level analysis and modification of images and videos. There are many built-in classes in OpenCV that facilitate image processing in areas such as facial detection, facial recognition, motion tracking, object identification, and image stitching.

Current research, used these built-in classes, mainly the image stitching class named "stitcher class", as a black box since current research focuses more on the application of image stitching rather than the actual algorithms involved. Current research used OpenCV written by C++ language that integrated with visual studio software.

For using OpenCV with visual studio, first, we included OpenCV libraries with visual studio program to allow using the classes needed in this project.

In implementation of current research *Stitcher* class used with OpenCV library. By using this class, it is possible to configure/remove some steps, i.e. adjust the stitching pipeline according to the particular needs.

Stitcher class is used to stitch two or more image together, but in current system, we need to stitch two or more video files, not image files. Therefore, in the

implementation code, we must capture the frames of videos frame by frame in sequence order from all video files that need to be stitched. Then we stitch them to create stitched frames.

4.1.1 System architecture (functions)

The system of current research has many functions based on multi view video and can be used for many purposes in video production and playing.

Based on the proposed method presented in the first chapter section 1.8, current system includes the following functions:

- Display captured videos
- Stitching videos to stitched frames
- Creating stitched video files from stitched frames
- Displaying FVV playing system

These functions are discussed in the next section with more details.

4.2 Execution sequence

In this section we discuss in details current system units, functions, and the sequence steps of how the system is used. We also present experimental results and screens of the system.

4.2.1 Display captured videos:

To begin with, the system displays video files that were captured from the user. These videos are multi viewpoint videos in different perspectives. They must be overlapped videos for stitching purpose, as illustrated in Figure4-1:
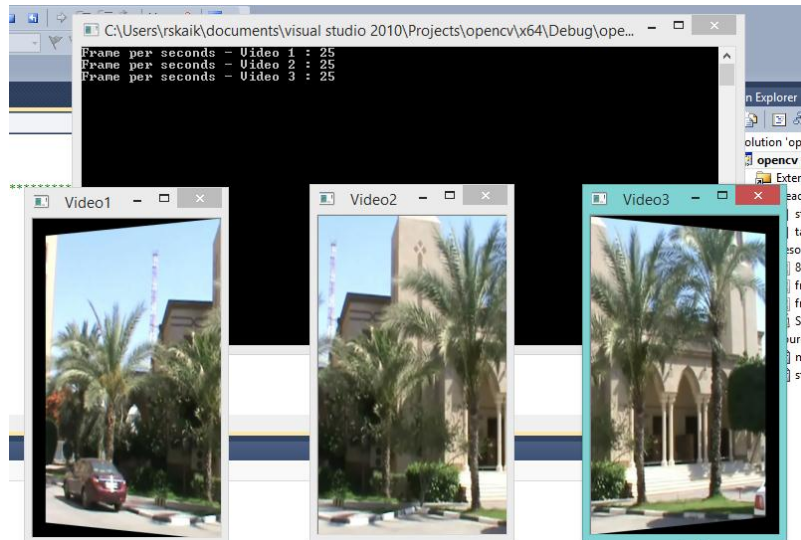
Figure 4-1 : three different perspective and overlapped video files displayed by the system

User can interact with the system by using keyboard shortcuts for the following:

S: to create Stitched frames for all video files.

P: to create stitched frames for each two pair of files in sequence order.

V: to create one Video file for all stitched frames created.

D: to Display video created as FVV window that allows the user to change his viewpoint.

Q: to Quit any process (stitching, creating video, FVV window ...etc.)

4.2.2 Video Stitching process:

As presented previously, the proposed FVV system is based on image stitching technique. For this purpose, we used Stitcher class in OpenCV written in C++ language.

There are two ways to stitch the frames of the videos in current system:

1. Stitching all Captured-video files to one Stitched-video file.
2. Stitching each pair of Captured-video files to one Stitched-video file.

Now, we give more details about these two ways:

In the first way, we stitch all sequence frames of videos captured to one sequence stitched frame, which means if we have a number of video files N, the system will first capture the first frame from (video1, video2, video3... videoN) then stitch it to create the first stitched file. Then, it captures the second frame from all N video files and stitches it to create a second stitched file and so on.

This way can be implemented in current system when the user presses "S" key in the keyboard.

The output of this process is: one sequence of stitched frames for all video files captured by user.

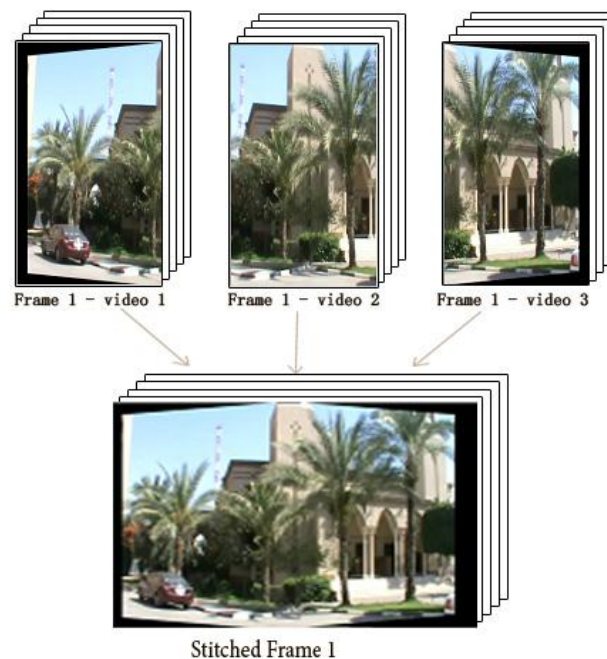Figure 4-2 illustrates the first way of stitching sequence frames of videos:

Frame 1 – video 1   Frame 1 – video 2   Frame 1 – video 3

Stitched Frame 1

Figure 4-2 1st way of stitching process-stitching all video files frames to one stitched-video frame

The second way is stitching the video files pair by pair sequentially, which means if we have N video files, the system will capture the first frame from first pair of video files (frame1 of video1 and frame1 of video2) and create the first stitched

frame. After that, it will capture the second frame of this pair of the video files (frame2 of video1 and frame2 of video2), create a second stitched frame, and so on for each two ordered frames in these two video files.

After finishing the first pair, this process will be repeated for each pair of video files available; video1 with video2, video2 with video3, video3 with video4 … videoN-1 with video N, in which N is the number of video files captured by user.

This way can be implemented in current system when the user presses "P" key in the keyboard.

The output of this process is: multiple (N-1) of sequence of stitched frames for each pair of video files captured by user.

In our experiment, we used three video files, so we have two pair of sequence frames as shown in figure 4-3, witch illustrates the second way of stitching sequence frames of videos:
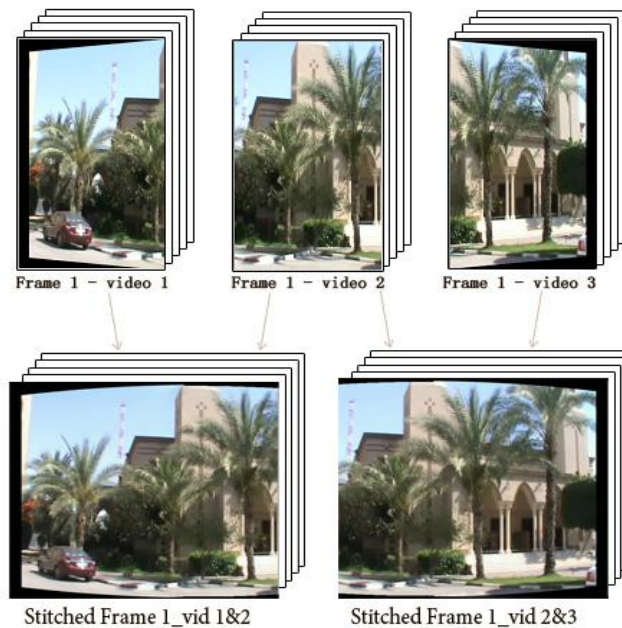

Figure 4-3 : 2nd way of stitching process-stitching each pair of frames to one stitched frame

Now, choosing the way of stitching (first or second) should be based on some factors:

- Size of video files captured by user.
- Number of video files captured by user.
- Size of overlapped region between each pair of user's video files.
- The capabilities of hardware (memory, CPU, cache, etc.) in user PC or laptop used for stitching.

Therefore, if the size of the video files is large, stitching process will need more power of hardware PC to create a stitched frame. It will also take long time to create a stitched frame. Accordingly, in this case, if we do not have enough capabilities hardware, the second way of stitching is preferred.

Furthermore, if the number of the user's video files is large, stitching process needs a hardware with high capabilities to do the stitching process, if not available, the program will break and crash. Hence, if we do not have a hardware with high capabilities, the second way is preferred.

In our implementation and testing process of proposed system, we noticed, after doing a number of experiments, that the size of overlapped region affects the stitching process. If the region is very small, the stitching process will not occur because one of the internal processes of stitching stage looks for the features between the images, and this will be difficult if we have a small overlapped region, so we need enough overlapped region to do stitching process.

As presented above, the factor that was effected the most in the stitching process time is the quality and capabilities of hardware. If the capabilities of hardware are high, the stitching time is reduced and vice versa. Therefore, we need to choose a high performance hardware PC if we need high performance results in the stitching process.

4.2.3 Create stitched video file:

After creating the sequence/sequences of stitched frames, the system allows the user to create one stitched-video file for each stitched sequence frame. The number of stitched video files is based on the way used in the stitching process. That means that if the user chooses the first way, one sequence of stitched frames, the system will create one stitched video file as shown in Figure 4-4(a), but if the user chooses the second way, N number of sequences of stitched frames, the system will create a number of stitched video file (N-1) as shown in figure 4-4 (b).

Figure 4-4 video file/s generated from sequence/s of stitched frames

a: one stitched video file generated by using first way of stitching

b: stitched video files generated by using second way of stitching

In our implementation code we can specify video properties like: frame per sec FPS, width, height, type of compression... etc.

To create video file in the system the user must press "V" key in the keyboarded, then the system done the process and save video file/s in specific location on user hard disk.

So, the output of this process is: one or multiple (N-1) stitched video files created in the user hard disk.

4.2.4 Display video as FVV:

The final process to achieve the objective of current research, is displaying the created stitched video file/files and allows user to freely and smoothly change his viewpoint over this video. To achieve this goal, we implemented a window in the system that contains stitched video and track-bar for changing the viewpoint.

In this stage, the system displays one viewpoint of scene, and the user moves the track-bar left and right to change his viewpoint as he/she wants.

This process is also based on stitching process and video creation process, and there are two ways to display the FVV based on the number of stitched video available. The first way is: if we have one stitched video file, the system displays the viewpoint region from the original stitched video, and when the user moves the track-bar to change his viewpoint, the system calculates the new region – number of pixels – needed to display in the available window.

The Second way is: when we have multiple video files (pair by pair stitching process), the system will display the region of one of the video files. Then, when the user moves the track bar in the window, the system will calculate the new region. If the new region comes out of the boundaries of the current video, the system will display the second file based on its order and so on.

The system gives ability to change from viewpoint to another with smooth interpolation. That is because our information -pixels-, needed to generate the scene, is always available because we use stitching technique in which all information for all viewpoints is available and does not need any prediction. Also, the system calculation of the new region is very fast, and that helps interpolate from one viewpoint to another smoothly, and this is objective in current research.

To display the final FVV in the system, the user must press "D" key in the keyboard, then the system will display window to view the video and support track-bar to change the viewpoint of the video.

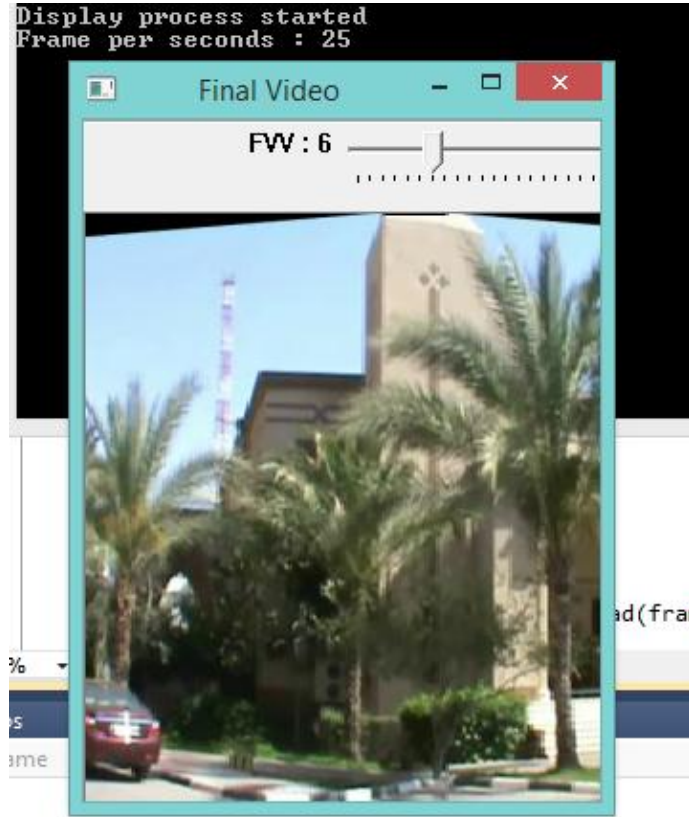So, the output of this process is: final FVV playing system as shown in Figure 4-5.

Figure 4-5 FVV system displays stitched video with track-bar for controlling interpolation between viewpoints

Chapter (5)

Testing and Evaluation

This chapter, describe test process of current system's units in more details. Also, it is discuss and evaluate the final system from multiple aspect:

- Stitching accuracy: which means testing the disruption in overlapped region.
- Stitching speed: which means testing the time needed to stitch frames
- Playback smoothness: which means testing the smooth interpolation between viewpoints over display FVV
- User experience: which means measuring the user satisfaction about the system using user survey and finding the results.

For each aspect, current research discuss all steps needed to evaluate this aspect.

A problem for evaluation of free viewpoint video is that for virtual intermediate views there are no original views to compare with

 In following, it discuss in details all steps needed to test each aspect.

5.1 Accuracy of stitching process:

In the system, as we discussed in implementation chapter, stitcher class in OpenCV was used to stitch multiple video files into one stitched video file.

To guarantee that the stitching technique used in the system gives our aspired results, current research test many different experiments, which can be listed as:

- Stitching two images.
- Stitching more than two images – three or more –
- Stitching two video files
- Stitching more than two video files – three or more –

For each case, current research tested and evaluated the results of the system.

5.1.1 Stitching two images:

To begin with, in the first case current research used two images as shown in Figure 5-1. These images have properties as shown in the following table:

File type: Jpg	Width :800	High:600	Resolution: 72

In the Figure 5-1, current research used *stitcher* class in OpenCV to stitch two overlapped images (Left and right image) with the same perspective view to create one image (result image). in result-image there is no distortions in the boundaries of overlapped areas as we see in the middle of the result-image.

(a) Left image   (b) Right image
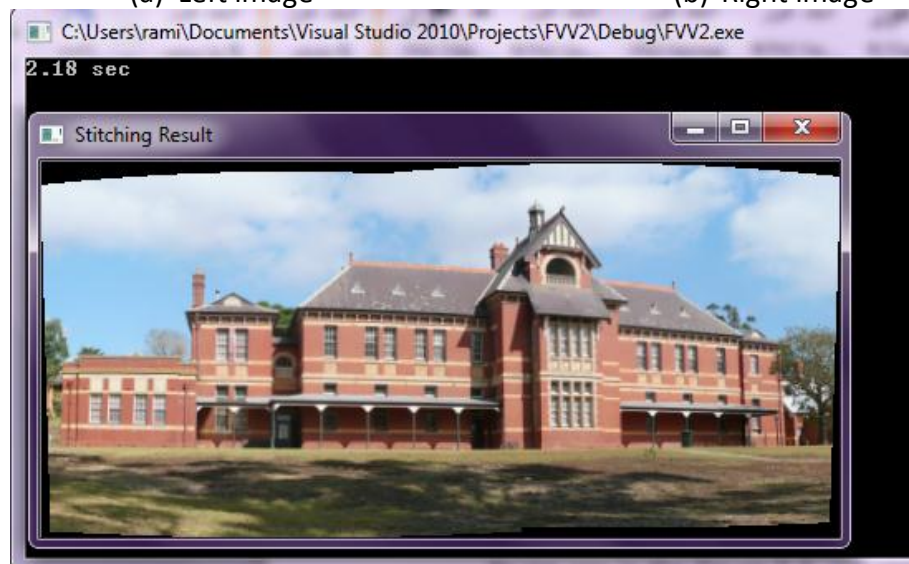
(c) Result image

Figure 5-1 stitching two images process with same perspective

To be more closer to objective of the system of FVV, the same previous experiment repeated by using the left image (a) as it is, but we changed the perspective of the right image (b) to be captured in a different viewpoint as seen in Figure 5-2

(a) Left image   (b) Right image- different viewpoint
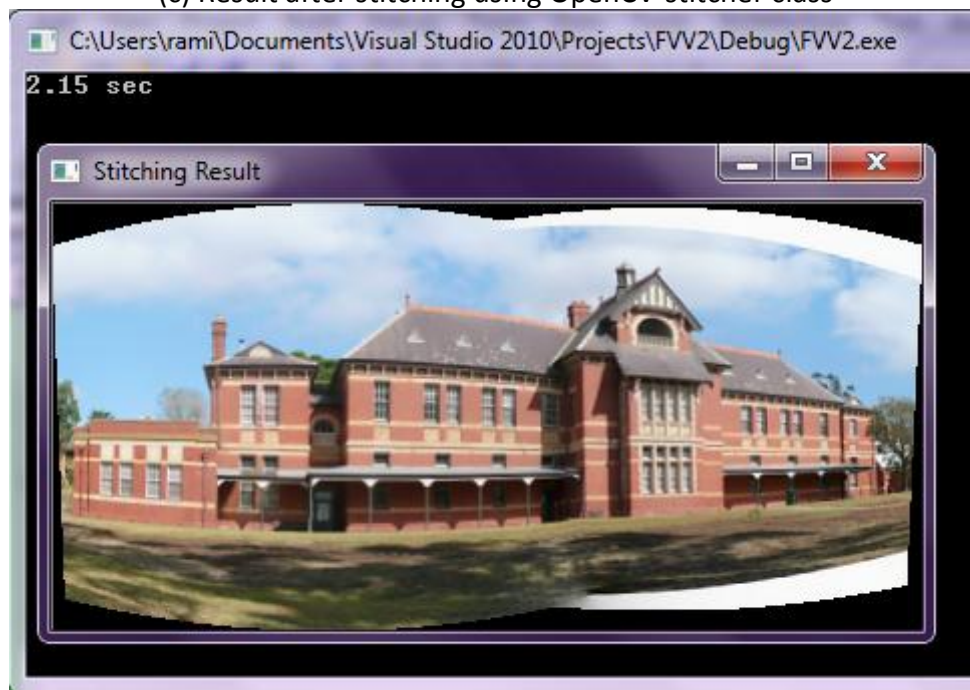(c) Result after stitching using OpenCV-stitcher class

Figure 5-2 stitching two images process with different perspective view

As we see in the two Figures 5-1 and 5-2, the results of stitching two images (same perspective or different perspective) gives images without any distortions. This means that the stitching process works accurately with two images.

Now, because the system used stitching for more than two images, in the next section we tested same experiments by using multiple images.

5.1.2 Stitching more than two images:

In this case, we tried to stitch more than two images to evaluate the ability of the system to make a full image from all the camera shots for free transition between viewpoints in our scene. Wherefore, we used four images in the stitching process of the system. These images have size 230*350 pixel.

In Figure 5-3, the system stitched four images and created one stitched image for all of them. It take 13.95 second in stitching. This time is long[26], but it give excellent stitching results, and that is our interest in this stage of the system.

This can be done for images with same perspective or different perspective images as we seen in previous section.

Figure 5-3 stitching four images

5.1.3 Stitching two video files

After testing the quality of the system in stitching images, we tested stitching for two video files. This requires video files with overlapped area. Then, the system captures the frames of these video file -frame by frame- and stitches it with corresponding frames in another video file.

First, we tested stitching two video files with the same perspective as seen in Figure 5-4.

These videos have the size of 300*240px, 25 fps, and avi type.

According to previous experiments in stitching images, stitching two video files with a different perspective will give the same accurate results as we see in Figure 5-2.
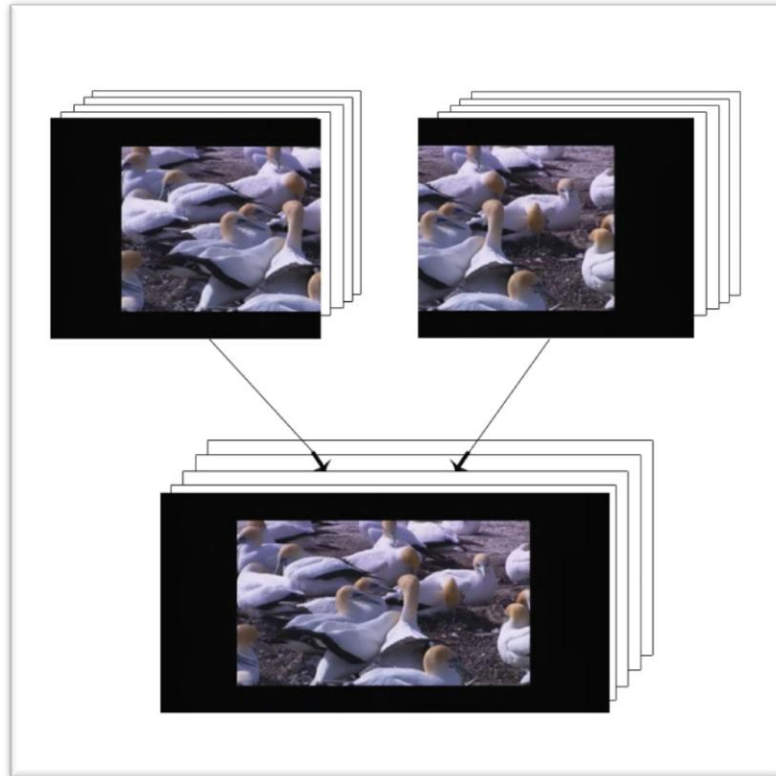
Figure 5-4 stitching process for two overlapped video files

As we can see in Figure 5-4, the system stitched the two videos accurately and gave frames with no distortions.

Now, we need to test the system with more than two video files and with different perspective.

5.1.4 Stitching more than two video files – three or more –

Finally, current research tested the system to stitching multiple video files with different perspective. For this case, the research used one video file and divided it to three videos with different perspective and used it in the system to create one stitched-video file.

Each of the three videos is an avi video and has the size of 340*570px.

Figure 5-5 illustrates how we can stitched three video files with different perspective and presents the time needed to stitch three frames from each video to create one stitched video.

As we can see in the results in Figure 5-5, the frames generated in the stitching process have no distortions.
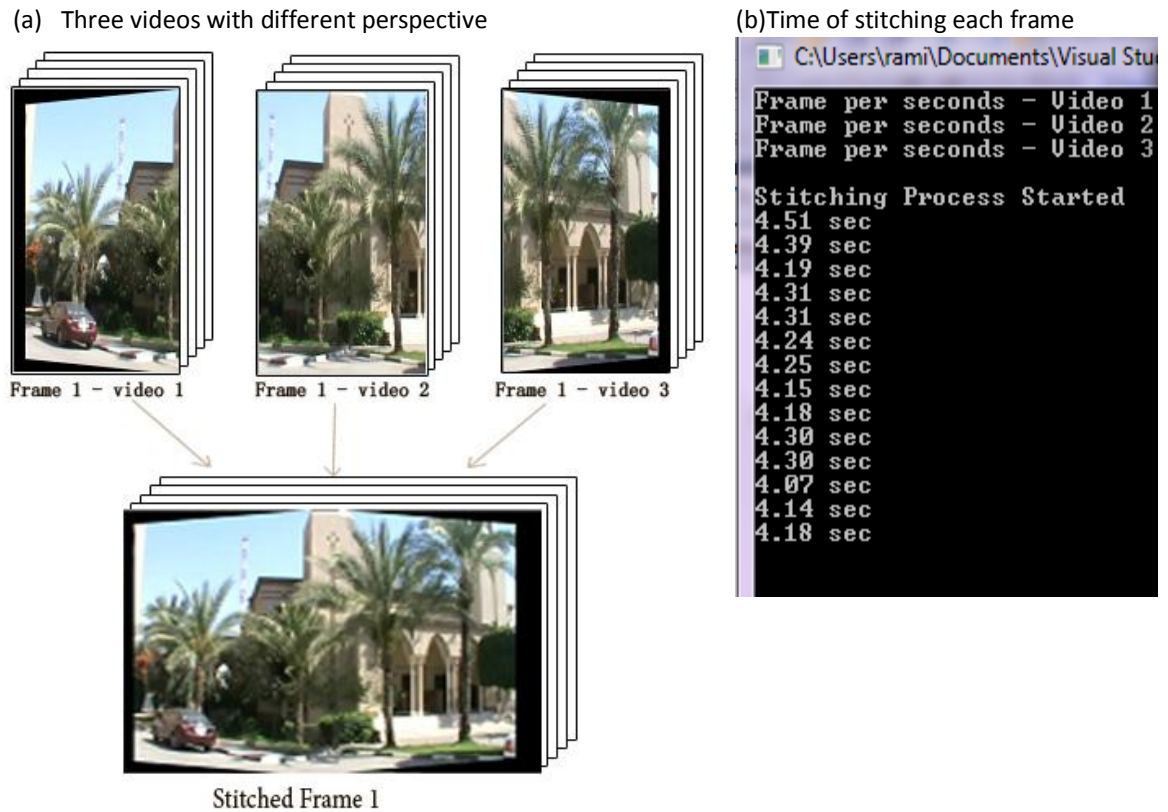
(a) Three videos with different perspective   (b)Time of stitching each frame

Figure 5-5 stitching three videos with different perspective views

5.2   Speed of stitching:

This section, presents the stitching time results for the previous experiments presented in the previous section. Then, it discuss the factors that affected the results.

To begin with, as we can see in Figures 5-1 to 5-5 in section 5.1, the stitching time of the experiments is as following:

Process	File Type	Size	Time /sec
stitching two image with the same perspective	Jpg	800*600	2.18
stitching two image with a different perspective	Jpg	800*600	2.15
stitching four images	Jpg	230*350	13.9
stitching two videos with the same perspective	avi	300*240	1.5-2 / frame
stitching three videos with a different perspective	avi	340*570	4

Table 1 Stitching time of the experiments

In fact, these results are not fixed, and they differ if we test the same experiment multi time because there are some factors affecting these results like:

- Power capabilities of user PC or laptop
- Number of processes running with the stitching process in user PC
- Frame properties and overlapped areas of these frames in user videos
- The video size used for stitching
- The video number used in stitching

All these factors will affect the results of stitching time. Therefore, the results are not stable for each experiment in the testing phase. For instance, in the testing phase, if we calculate the stitching time for the two-video-file experiments, we calculate this process again in a different time for the same video files, and the results will be different. This is based on the state of our PC in this time, how many processes are running in the PC in this time, what is the state of cache memory in this time, and so on. These all affect the results of stitching process.

In the experiments, current research used three video files. The average of stitching time for each frame is 2-4 sec. It may be very long time for stitching video, especially for large video files, but if we have more powerful equipment, we would get better time results.

In general, as we searched and experimented: image stitching technique takes a long time, but it also gives accurate results.

5.3 Smooth interpolation during display FVV

The most important thing in free viewpoint video system is video display and allowing the user to change his/her viewpoint. Furthermore, the most critical point in FVV is when user changes his/her viewpoints, and the system has to interpolate to next viewpoint without any distortions. We named this: smooth interpolation between viewpoints in FVV system.

In traditional FVV applications there are problem to evaluate of free viewpoint video ,that is for virtual intermediate views there are no original views to compare with[34]. But in current research FVV system based on stitching technique all information need to generate views are available.

Current research, to evaluate the quality and smooth interpolation during display free viewpoint video only the visual impression is appropriate[35]. In current research the system gives the ability to freely and smoothly interpolate in stitched video without distortions. This is because The system is based on the stitching technique, and this technique gives all information about the scene in order to display it without the need of any predictions as in another panorama techniques and systems like in[30]..

Current research used a track-bar (Figure 4-6) in the system. The user moves this track-bar lift and right for interpolating between different viewpoints. When the user moves the track-bar, the value of this bar changes, and based on this value,

the system will calculate the region that will be displayed on the screen of the system along with the new viewpoint displayed with a smooth move.

5.4 User experience

In the traditional video, the scene is filmed by multiple cameras, then the director composes shots to make the final scene. Hence, the director controls the viewpoints of the video displayed to the viewer, and the viewer can only watch the scene without any control over this video. However, FVV system gives the user the ability to control the video viewpoints and to freely interpolate over viewpoints during video playback.

Chapter 3 presented that current research based on questionnaire to evaluate user experience. The detailed user questionnaire comprised a list of questions that were aimed at measuring the usability of the product, the quality of the user experience and also some general information about the respondent.

When designing the questionnaire we took into account the main usability principles.

Questions concerning **usability** address the following concepts:

- Functional correctness
- Ease of use and learning

Questions concerning the "**user experience**" address the following
- Fun, enjoyable
- Appealing presentation: this increases user's attention
- Attention
- Motivation: whether the user is interested in the content
- Overall degree of satisfaction

In current research we need to measure the key factors that contribute to the success or failure of the application. This requires considering whether the concept could be improved, and if so, how. Also, we need to know how users felt

about the concept in terms of general interest, satisfaction, and the economic value, and whether they are willing to assign to it.

Accordingly, the evaluation had the following goals:

1. To find out the level of **enjoyment** obtained by the end-users and to investigate what were the concept's weakest points, and what functional enhancements were needed to improve the user experience.

2. To assess the degree of **usability** of the use cases in terms of ease of use, efficiency of stitching, speed of response, and ease of interpolating in different viewpoints of the video.

3. To find out how the service provided by the application is **economically** valued by the test users.

Moreover, there are other aspects not taken into account: users will probably have suggestions about aspects that we did not think about previously, so the facilitators listed any additional suggestions and comments offered by the users.

Apart from these factors we are also interested in the **spontaneous evaluation** that comes up from the users in a subjective and unrestricted way. This may eventually give us feedback on **unexpected aspects** concerning our future system that we did not take into account a priori.

Test facilitators assisted in gathering this information by observing the user attitude, verbal comments and gestures during the tests.

To evaluate the system and measure the previous factors, we search about available questionnaire formats and founded a lot of them like:

- Software Usability Measurement Inventory (SUMI) [35].
- System Usability Scale (SUS) [36].
- Questionnaire for User Interaction Satisfaction (QUIS) [37].
- "Usefulness, Satisfaction, and Ease of use "(USE) [38].

All of these questionnaires templates and other included questions for general applications. However, current research need some specific questions to measure the functionality of the system.

So, to evaluate the system and measure the factors, we asked the end users to fill in a user-questionnaire.

User-questionnaire was adapted from USE[38], QUIS[37]and other questions that important to evaluate specific factors in the FVV system. This questionnaire includes two categories of questions:

First, for respondents personal information: that include age , gender, job ,and some specific information questions as following :

- Degree of interesting in media playing system
- Degree of enjoinment with 3d videos
- Degree of using panorama systems
- Degree of how he/she was follow-up the new technology

Second, for measuring the FVV functionality issues, and it divided into two classes:

1- Statements about the system's idea: that measure if the idea is novel, beneficial, useful, and enjoyment
2- Statements about the system's accuracy: that measure the quality of stitching results, speed, and smoothness.

For each part of statements in questionnaire we calculate Cronbach's alpha to measure the reliability of our statements, also we measure the scale statistics using Mean and Slandered deviation (Std. Deviation) and results we presented it in Table 2.

Statements	Number of items	Cronbach's alpha	Mean	Std. Deviation
Statements about the system's idea	6	.782	27.59	2.671
Statements about the system's accuracy	4	.906	18.11	2.413

Table 2 results of measuring user-questionnaire

For seeing the form of user-questionnaire, check appendix (A).

5.4.1   Results

This section presents the results of the home user surveys. We presented the system to 37 users in the information technology field; teachers and students.

For each question, the respondents were asked to rate the statement from 1 to 5, where 1 means that they strongly disagree with the statement, and 5 means that they strongly agree.

This section, presents the results for two categories (personal information and FVV functionality issues) and will discuss it in details below:

1- Results for **personal information** show that:

   o 76% of responses are greatly interested in media playing system,

   o 86% of those who answered greatly enjoy with 3D videos,

   o 43% those who answered greatly used panorama systems, and 31% of the users don't use it a lot,

   o And 85% of those who answered is up-to-date with technology.

For more details of responses results see appendix (B).

2- Results for **FVV functionality system issues**: (statements about the system's idea and accuracy) that presented in the following table (table 3), and discussed it below .
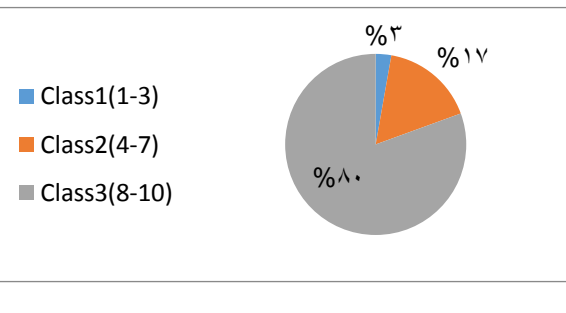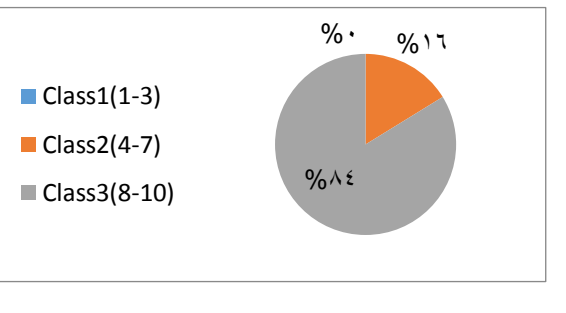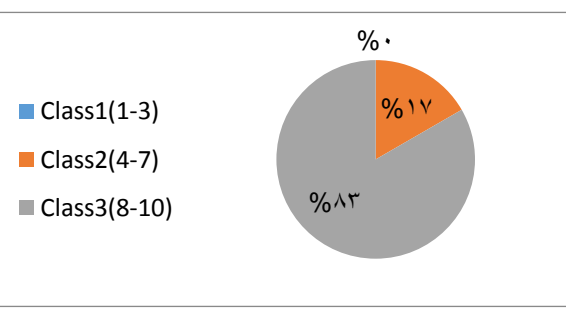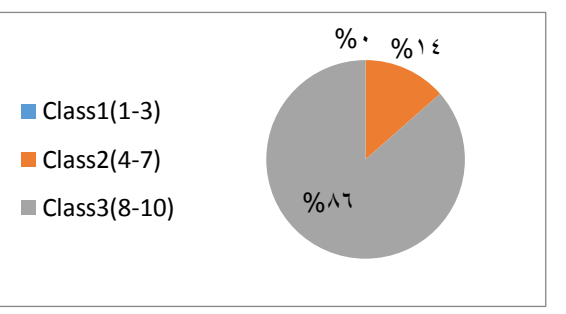
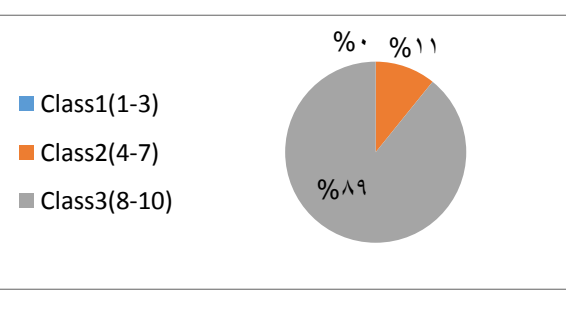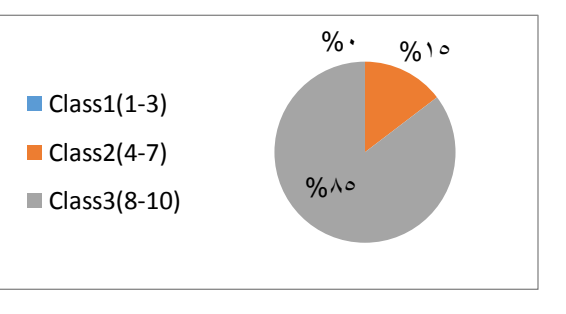## A. Outputs of testing system idea:

Q2.1 : "I think idea of project is a novel idea" :	Q2.2 : "I think idea of system is beneficial and useful":
%٣ %١٧ ■ Class1(1-3) ■ Class2(4-7) ■ Class3(8-10) %٨٠	%٠ %١٦ ■ Class1(1-3) ■ Class2(4-7) ■ Class3(8-10) %٨٤
Q 2.3 "I think the video playing systems would be great using FVV systems":	**Q 2.3 "It is very useful to control the viewpoints over watching the video "**
%٠ ■ Class1(1-3) ■ Class2(4-7) %١٧ ■ Class3(8-10) %٨٣	%٠ %١٤ ■ Class1(1-3) ■ Class2(4-7) ■ Class3(8-10) %٨٦
Q2.5 : " I think such a system would increase my enjoyment of watching video " :	**Q2.6 : " I will use this type of system if it is available ":**
%٠ %١١ ■ Class1(1-3) ■ Class2(4-7) ■ Class3(8-10) %٨٩	%٠ %١٥ ■ Class1(1-3) ■ Class2(4-7) ■ Class3(8-10) %٨٥

Table 3 testing results for system's idea
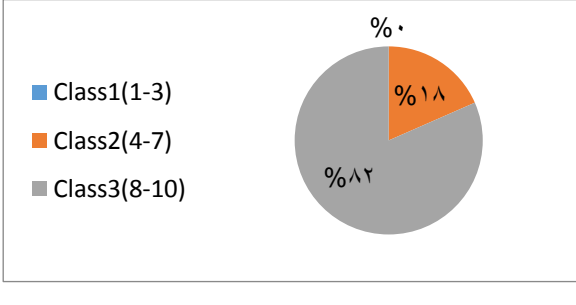
B. Outputs of testing system accuracy:

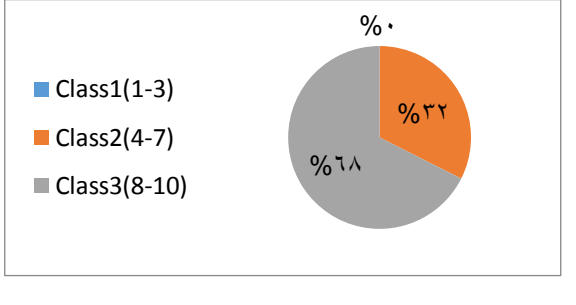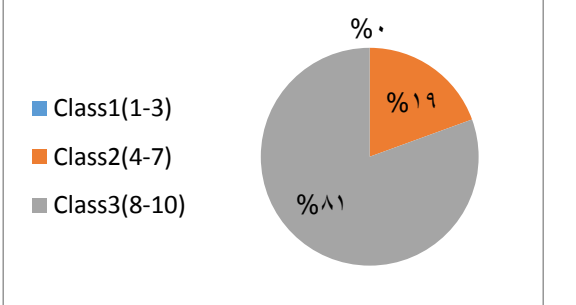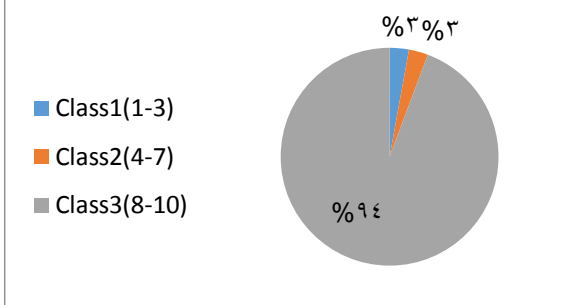| Q3.1:"The system has got a great picture quality" :	Q3.2: "The system has a good timing when stitching frames from multiple video files ":
Q3.3:"The system has a good timing when creating a stitched video from frames " :	Q3.2: "The system has a good reaction timing in interpolation viewpoints ":

Table 4 testing results for the system's accuracy

Few additional comments from the respondents were also listed in order to better measure their impressions of the concept. The following additional comments were received from respondents:

Some users would like to make the controlling of interpolation of viewpoints by using mouse over the video area itself, not by using scroll.

Some users would like to develop this idea in the spot area.

5.4.2 Outputs discussion:

The results presented in Section 5.4.1 and Appendix (C) divided into three aspects: respondents information, system idea, and system accuracy. In this section we present the results for each aspect and discuss it in details.

A. Discuss part of "Respondents information":

All samples of respondents were selected have irrelevant with multimedia field. In the results most of them are interested in media playing system and enjoy with 3d video, but only 43% of them used panorama system before, Due to lack of availability of Panorama programs for most of them.

B. Discuss part of " idea of the system ":

Additionally, throw the results of evaluating the idea of the system, most of respondents (80≈85%) consider that the idea of the system is novel, beneficial and useful idea. The idea is "Novel", that because this type of projects are not widespread in Arab countries, and the idea is "beneficial and useful", that because it is makes video more realistic for viewers.

Also the majority of respondents (83≈89%) encouraged using FVV in video playing systems and consider it is very useful when control the perspective viewpoints when watching video, also consider that the idea will increasing enjoyment of watching videos. This gives us a positive satisfaction of the idea and encouraged us to continue forward develop the system.

C. Discuss part of " Accuracy of the system ":

There are three important aspects current research focused on this part of testing : quality of stitched-frames, speed of stitching process, and smoothly interpolation in FVV playing system.

As we see in Table 3, the most of respondents (82%) consider that the system gives good quality picture in stitching process, that because there are no noticeable distortions in the overlapped regions in stitched frames.

For smoothly interpolation in the system, also most of respondents (81%) consider that the FVV playing system gives smoothly and freely interpolation between viewpoints in the video. And this point is very important in the objective in current work.

Last point in this part, discuss the results of stitching process speed. In this point, just (68%) of respondents consider that the system is very fast in stitching process, but there are (32%) of respondents consider that it has low speed in stitching, that because stitching technique we used has entire processes to give the results, as we presented it before in Figure 4-1 in Chapter 4, and if we have more processing power we can reduce the time of stitching process in the system. So, we need to increase the performance of the system in stitching phase, one of solutions is to divide the system into two parts, one for stitching phase and second for playback free viewpoint video.

5.4.3 Conclusions and user recommendations to the project:

The results of the evaluation have been very positive in most aspects. The respondents involved in the test emphasized the high level of interest and positive reactions of nearly all the people who took part in the test.

The main indicators of this positive perception are that:

- 85% of the respondents will use the system if it is available
- 89% of the respondents consider that the system is increasing the enjoyment of watching video.
- 82% of the respondents consider that the system gives good-quality picture.
- 94% of the respondents consider that the interpolation between viewpoints in FVV is smooth and has a good reaction timing.

Chapter (6)

Discussion and Conclusion

This chapter, describe and discuss, in details, the limitations we faced in our system, and what is the different between the proposed system and current end system. We also discuss the alternative solutions of these limitations.

Through the implementation phase of the project, we faced some technical difficulties which caused some limitations in the end system.

These limitations fall in three aspects:

1. MVV Capture.
2. Graphical user interface GUI.
3. Processing power.

In the following section, we explain and discuss what was intended to be done, the problems and limitations we faced, and its alternative solutions.

6.1 MVV capture:

In free view point video, if we need to create multi view video, we need to use multiple cameras together to capture MVV scene. Also, we need a controller for all cameras to synchronize the filming process to give the frames in the exact time of using it in the stitching process, in order to give the right overlapped region between each adjacent frames in the cameras.

Furthermore, the cameras used in the system must give the same properties of image (resolution, size, quality, etc.) to make it easy to find features in the stitching process.

The equipment needed in the system must have special capabilities, and it is very expensive.

With modest possibilities, current research used one video file in the system to test the experiments. This video was divided initially to three video files, and we changed the perspective of these files to be captured from multiple viewpoints. By using one divided video, we avoided a lot of difficulties that may be faced in the modest system:

- The cost of equipment (camera, tripods, control-system … etc.)
- The asynchronies frames in the captured videos.
- The quality and size difference between images in different video files.

Finally, the objective of the capturing MVV for the experiment was done by capturing one video file and dividing it into different perspective files to be captured from different viewpoints, and this is what we need to complete testing the system.

6.2 Graphical user interface GUI:

In the system, as we presented in previous chapter, current research used shortcuts of keyboard for the interaction between the user and this system. That is because the implementation is based on OpenCV library with Visual Studio environment.

GUI is supported in OpenCV using Visual Studio confined in:

- Creating a window and controlling its properties.
- Displaying the video or image in this window.
- Creating a track bar in the window.
- Mouse operation over window.

Nonetheless, it does not create some important elements needed to make interactive interface for media player system. We need some components like button, radio button, checkbox, etc. And all of these components are not supported yet in OpenCV with Visual Studio.

After we searched for advanced GUI that can be used with OpenCV, we found the QT creator program. Some features, like creating buttons in OpenCV, were added to it. However, it is very difficult to convert the current system that is based on visual studio to QT creator with OpenCV because:

- Configuration steps to installing QT creator are very difficult.
- Linking OpenCV library with QT creator needs some external programs like CMake, which is a tool to create project files (for your chosen IDE) from the OpenCV source files.
- There are differences when writing the code of the system using visual studio C++ with OpenCV and converting it to QT creator code because QT/OpenCV functions differ from the functions used in Visual studio.

We read many tutorials and online articles talking about QT configurations and how to link it with OpenCV, and we tried to set it up, but it was not working correctly.

In the system, we replaced the action button needed for GUI with shortcut keys on the keyboard that have the same functions as we presented in details in implementation section.

6.3 Processing power:

The stitching process takes a long time to do because there are many internal processes, so it is need equipment with professional capabilities to decrease the time needed to do its processes. Also, video files affect stitching if they are large files, and this requires more time for stitching.

In research experiment, the Laptop with processor Core i7, 4G RAM was used. It made modest results in the stitching process, but if we need more efficient result we need to use more efficient capabilities.

6.4   Conclusion and Future Works:

With the modest capabilities of the system, and in spite of all limitations that faced the implementation, current research achieved most of the functionality needed to approve the proposed idea for using stitching technique in FVV playing system. The results of the system are positive in creating a good-quality image in stitching video phase and in smoothly and freely interpolating between viewpoints in FVV playing system. And finally, we won the user satisfaction as shown in chapter 5.

As seen in the previous section, if we have a better power processing PC, current research can get a better performance results in stitching time to create stitched video frames.

In the future work, for more interaction between user and the system, we intend to add GUI to the system that includes controlling components. Also, for speed limitation in stitching process we can divide the system into two subsystems: the first is an offline system for stitching and creating stitched frames, and the second is for displaying FVV.

References

[1]   J. Kilner, J. Starck, A. Hilton, and O. Grau, "Dual-Mode Deformable Models for Free-Viewpoint Video of Sports Events," in *Sixth International Conference on 3-D Digital Imaging and Modeling (3DIM 2007)*, 2007, pp. 177–184.

[2]   S. Yoon, E. Lee, S. Kim, Y. Ho, K. Yun, S. Cho, and N. Hur, "Coding of Layered Depth Images Representing Multiple Viewpoint Video," *Gwangju Inst. Sci. Technol.*, vol. SS3–2, pp. 1 –4, 2006.

[3]   S. K. Nayar, "Catadioptric omnidirectional camera," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 482–488.

[4]   en.wikipedia.org, "Omnidirectional_camera," 2014. [Online]. Available: http://en.wikipedia.org/wiki/Omnidirectional_camera.

[5]   M. Tanimoto, "Overview of FTV (free-viewpoint television)," in *2009 IEEE International Conference on Multimedia and Expo*, 2009, vol. 4, pp. 1552–1553.

[6]   A. Majumder, W. B. Seales, M. Gopi, and H. Fuchs, "Immersive teleconferencing," in *Proceedings of the seventh ACM international conference on Multimedia (Part 1) - MULTIMEDIA '99*, 1999, pp. 169–178.

[7]   S. Shimizu, H. Kimata, K. Kamikura, and Y. Yashima, "FREE-VIEWPOINT SCALABLE MULTI-VIEW VIDEO CODING USING PANORAMIC MOSAIC DEPTH MAPS," *16th Eur. Signal Process. Conf.*, no. Eusipco, 2008.

[8]   J. Starck, J. Kilner, and A. Hilton, "Free-viewpoint video renderer," *J. Graph. tool*, pp. 1–15, 2008.

[9]   "MULTI-VIEW VIDEO CODING USING REFERENCE PICTURE SELECTION FOR FREE-VIEWPOINT VIDEO COMMUNICATION Hideaki Kimata , Masaki Kitahara , Kazuto Kamikura and Yoshiyuki Yashima NTT Cyber Space Laboratories , NTT Corporation," *Pict. Coding Symp. (PCS 2004*, pp. 4–7.

[10]   R. Szeliski and R. Szeliski, *Image Alignment and Stitching*. 2004, p. 89.

[11]   Y. Ho and K. Oh, "Overview of Multi-view Video Coding," in *2007 14th International Workshop on Systems, Signals and Image Processing and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services*, 2007, vol. 3164, no. d, pp. 5–12.

[12]   A. T. H. Au, "Development of Multiview Image / Video Stitching Systems for Mobile Devices by Andrew Au," School of Engineering Sience, 2013.

[13]   S. Würmlin, E. Lamboray, M. Waschbüsch, P. Kaufmann, A. Smolic, and M. Gross, "Image-space Free-viewpoint Video," *Fraunhofer Inst. Telecommun.*, no. November, 2005.

[14]   A. Vetro, S. Yea, M. Zwicker, W. Matusik, and P. Hanspeter, "OVERVIEW OF MULTIVIEW VIDEO CODING AND ANTI-ALIASING FOR 3D DISPLAYS Mitsubishi Electric Research Labs Cambridge , USA Department of Computer Science and Engineering University of California , San Diego , USA," *Mitsubishi Electr. Res. labs*, pp. 17–20, 2007.

[15]   G. Jiang, L. Fan, M. E. I. Yu, and F. Shao, "New View Generation Method for Free-Viewpoint Video System," *WSEAS Trans. Comput.*, vol. 7, no. 6, pp. 589–598, 2008.

[16]   E. Adel, M. Elmogy, and H. Elbakry, "Image Stitching based on Feature Extraction Techniques: A Survey," *Int. J. Comput. Appl.*, vol. 99, no. 6, pp. 1–8, Aug. 2014.

[17]   Y. Z. L. Y. Z. Wang, "Research on Video Image Stitching Technology Based on SURF," *Comput. Intell. Des.*, vol. 2, no. Oct., pp. 335–338, 2012.

[18]   C. Chen, "Image Stitching - Comparisons and New Techniques," *CITR*, no. October, 1998.

[19]   H. Zhang, W. Zeng, and X. Chen, "Foreground Based Borderline Adjusting for Real Time Multi-camera Video Stitching," in *2009 Fifth International Conference on Image and Graphics*, 2009, pp. 355–360.

[20]   D. Appello and D. Erickson, "Multi-Cam Video Stitching : Panoramic Video Option During a Skype Call by," Dickinson, 2014.

[21]   M. Brown and D. G. Lowe, "Recognising panoramas," in *Proceedings Ninth IEEE International Conference on Computer Vision*, 2003, pp. 1218–1225 vol.2.

[22]   L. Juan, "A Comparison of SIFT , PCA-SIFT and SURF," *Int. J. Image Process.*, vol. 3, no. 4, pp. 143–152, 2009.

[23]   D. Bheda, A. Prof, M. Joshi, and V. Agrawal, "A Study on Features Extraction Techniques for Image Mosaicing," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 2, no. 3, pp. 3432–3437, 2014.

[24]   H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Comput. Vis. Image Underst.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.

[25]   "Stitching Pipeline — OpenCV 3.0.0-dev documentation," 2007. [Online]. Available: http://docs.opencv.org/trunk/modules/stitching/doc/introduction.html#bl07. [Accessed: 08-Jan-2015].

[26]   M. Brown and D. G. Lowe, "Automatic Panoramic Image Stitching using Invariant Features," *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59–73, Dec. 2006.

[27]   A. Agarwala, M. Cohen, and D. Salesin, "Photographing Long Scenes with Multi-Viewpoint Panoramas," *ACM Trans. Graph.*, 2006.

[28]   M. Adam, C. Jung, S. Roth, and G. Brunnett, "Real-time Stereo-Image Stitching using GPU-based Belief Propagation," *Work. Vision, Model. Vis.*, no. November, pp. 215–224, 2009.

[29]   I. D. S. Segvic, M. Sevrovic, G. Kos, V. Stanisavljevic, "Pr el i m i n ary e xp e ri men ts i n m u l ti-vi e w vid eo stit ch i n g," *34th Int. Conv.*, pp. 892–896, 2011.

[30]   J. Lou, H. Cai, and J. Li, "A Real-Time Interactive Multi-View Video System," *ACM Multimed.*, pp. 161–170, 2005.

[31]   J. Starck, J. Kilner, and A. Hilton, "Free-viewpoint video renderer," pp. 1–15, 2008.

[32]   J. M. J. Higginbotham, "Qualitative program evaluation methods," vol. 16, no. 1540, 2011.

[33]   J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *Proceedings of the SIGCHI conference on Human factors in computing systems Empowering people - CHI '90*, 1990, pp. 249–256.

[34]   A. Smolic, K. Mueller, P. Merkle, T. Rein, M. Kautzner, P. Eisert, and T. Wiegand, "Free viewpoint video extraction, representation, coding, and rendering," in *2004 International Conference on Image Processing, 2004. ICIP '04.*, 2004, vol. 5, pp. 3287–3290.

[35]   V. C. P. Erik, "Questionnaire based usability testing," *Conf. proceeding Eur. Softw. Qual. Week*, 1998.

[36]   John Brooke, "SUS - A quick and dirty usability scale," *Redhatch Consult.*, 1996.

[37]   Human-Computer Interaction Lab, "Questionnaire for User Interaction Satisfaction (QUIS)," *University of Maryland at College Park*, 2000. .

[38]   A. . Lund, "USE Questionnaire: Usefulness, Satisfaction, and Ease of use," *STC Usability SIG Newsletter*, 2001. [Online]. Available: http://hcibib.org/perlman/question.cgi?form=USE. [Accessed: 08-Jan-2015].

Appendix (A)

## User survey and questionnaire form

Based on the idea of project described in next page, please respond to each statement by rating how it applies to you, where:

1 = It doesn't apply to me at all.
5 = It applies exactly to me.

Free viewpoint video playing system using stitching technique					
PERSONAL INFORMATION :					
Section 1 — **Statements about the _survey respondents_ :**					
Question	**Degree**				
	1	2	3	4	5
1.1 I am interested in media playing systems					
1.2 I enjoy 3d videos					
1.3 I used panorama systems before					
1.4 I consider myself as a general "early adaptor " when it comes to technology					
FVV FUNCTIONALITY ISSUSE:					
Section 2 — **Statements about _system idea_ :**					
Question	**Degree**				
	1	2	3	4	5
2.1 I think idea of project is novel idea					
2.2 I think idea of system is beneficial and useful					
2.3 I think video playing systems would be great using FVV systems					
2.4 It very useful when controlling the viewpoints when watching video					
2.5 I think such a system would increase my enjoyment of watching video					
2.6 I will use this type of system if it available					
Section 3 — **Statements about _system accuracy_ :**					
3.1 System has got a great picture quality					
3.2 System has a good time when stitched frames from multiple video files					
3.3 System has a good time when create stitched video from frames					
3.4 System has a good reaction time in interpolation viewpoints					
Section 4 — **Additional notes :**					

Free viewpoint video playing system using stitching technique
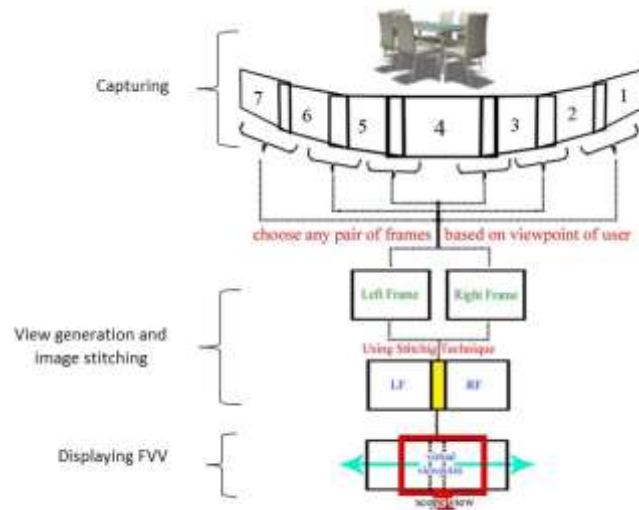
First, our system displays video files that were captured by the user. These videos are multi viewpoint in different perspectives. In our system, we used one video and divided it into three videos with different perspective. This idea can be used for more than three videos to view the scene more like a 3D view.

To interact with our system, you must use keyboard shortcut as following:

S: to create *Stitched* frames for all video files at once all together.

P: to create stitched frames for each *Pair* of files in sequence order.

V: to create one *video* file for all stitched frames created.

D: to Display video created as FVV window allowing user to change his viewpoint.

Q: to *Quit* from any process (stitching, creating video, FVV window, etc.)

Appendix (B)

Results for all questionnaire survey

Each question has scale from 1 to 5, next table presents the number of users for each scale of each question:

| Q#	1	2	3	4	5
1.1	0	1	4	9	23
1.2	0	0	1	10	26
1.3	7	3	6	11	12
1.4	1	2	6	12	17
2.1	0	1	2	10	23
2.2	0	0	3	14	20
2.3	0	0	3	8	25
2.4	0	0	2	8	27
2.5	0	0	3	6	28
2.6	0	0	1	9	31
3.1	0	0	2	13	23
3.2	0	1	4	16	16
3.3	0	0	3	10	23
3.4	1	0	0	3	31

There are 14 statements divided in to three categories. The first 4 statements (1.1 – 1.4) take about user information. The second 6 statements (2.1 – 2.6) are about the idea of system, and the last 4 statements (3.1 -3.4) are about the accuracy of our system. For each question, there are 5 degrees. The table presents the number of users who check each degree for each question.

Appendix (C)

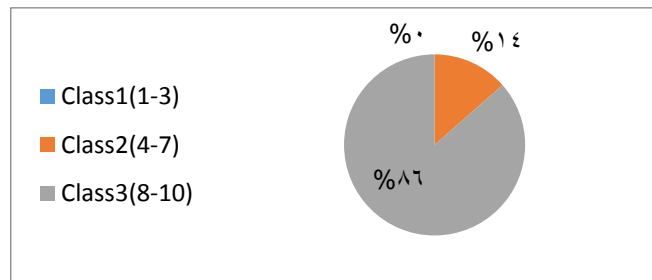## User questionnaire results of category one

"Statements about the survey respondents"
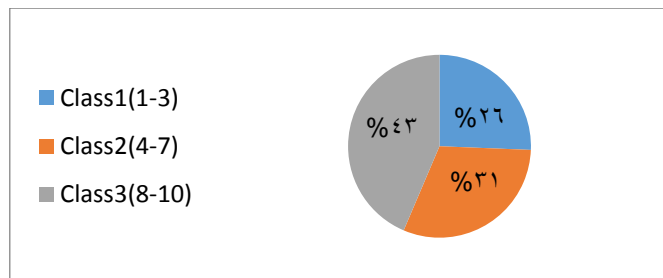
1.1 I am interested in media playing system
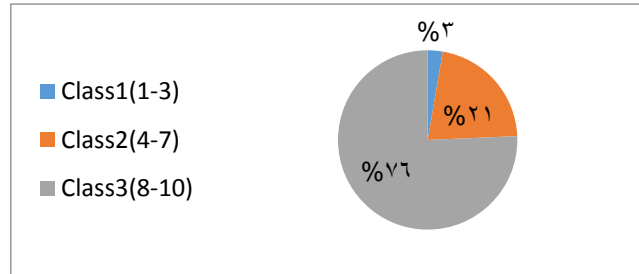
Statement: 76% of those who answered is greatly interested in media playing system.
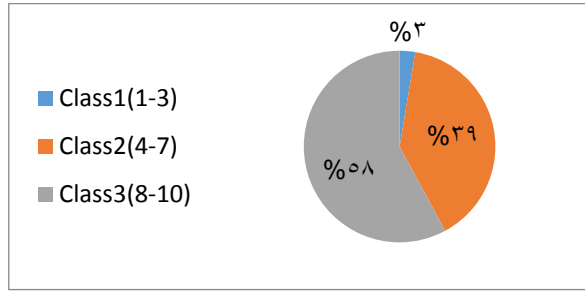
1.2 I enjoy 3d videos

Statement: 86% of those who answered greatly enjoy with 3D videos

1.3 I used panorama systems before

Statement: only 43% those who answered greatly used panorama systems,
And 31% of the users don't use it a lot

1.4 I consider myself a general "early adaptor" when it comes to technology

Statement: only 58% of those who answered is up-to-date with technology