Using Shape Representation to Design Panorama Video System

Abeer Salim Jamil*, Ph.D,(Lecturer) Abdul Monem S. Rahma**, Ph.D(Prof)
Nada Hussein M. Ali***, Ph.D(Lecturer)

Abstract

The panoramic scene generation is a research active area. The importance of panorama generation is increasing the view angle. Using video to generate the panoramic video, a single camera cannot meet the objective to generate the panoramic view in suitable angle. To solve this problem, two or more video sequences are required to produce the panoramic video. This will give a scene wider than that with a single camera. When the number of cameras is increased, the panoramic video view angle will increase.

The goal of this research is to generate a new form of video called panoramic video using two or more videos. The panoramic video is generated using two proposed methods. First, the system of panoramic generation depends on the capturing of samples by using normal cameras and the operation will be done through capturing samples for fixed place at the same time. Secondly, the estimation of motion between sample transformation considers on the challenge in the field which depends on the feature-based method. Feature- based method must be strong enough to extract the fixed features in different circumstances like: distortion, light change, noise and blurring by different transformations (scaling, rotation, translation... etc.) and also the use of the shape representation (region - based method) to resolve the problem of overlapping region. The experimental results show that the system can generate panoramic video from two cameras in different places.

Keywords: Panoramic Video, Feature Based Registration, Shape Representation, Panorama Construction, Overlap Region

* Al-Mansour University College
** University of Technology
*** University of Baghdad

1- Introduction

The photographs obtained from a camera have a limited field of view. There may be data on a scene not appeared in one image. Then by combining several images with overlapping fields of view called panorama [1]. The field of vision is extended angular of the observable world that can be seen at any moment. It can provide full image on location or an area that cannot fit in a single shot.

In the last few years, there has been a growing increase in the use of digital video due to the expansion in both the use of multimedia technology and Internet technology [2]. A video is a wealthy source of information. The video provides visual information about the scenes [3]..

Video has become a part of our everyday life, think of television Broadcast for example. To capture the events in the real world around us must use most effective way which is video. Also, video the most dramatic medium which combines the sounds with photo-realistic images. Combining video advantages & computers will expand a field of the information that a computer could process [2].

The video is composed of spatial and temporal components. These ingredients supply the compact description of the video data. Video data can be converted from a serial frame-based representation, that in which this common scene information is distributed in many forms, into one, common scene-based representation to which each frame can be related directly. This kind representation, will provide immediate & direct access to scene information, like a static location and dynamic objects moving [4].

This makes the video very eye catching and interesting and encourage us to see video from a different point of generating a new form of video called panoramic video. Although image stitching has been a focus of interest in the computer vision community, much less attention has been paid to the closely related topic of panoramic video. The goal of video stitching is to generate a panoramic video from multiple overlapping video streams captured by a number of cameras, see figure (1). When capturing a number of videos, the cameras can be fixed while they are capturing the videos, but their relative geometry is unchanged during the capturing procedure [4].Panoramic video can be used for many purposes such as in classroom lectures or video

conferencing or even as a virtual tour of hotels, museums or exotic travel destinations .

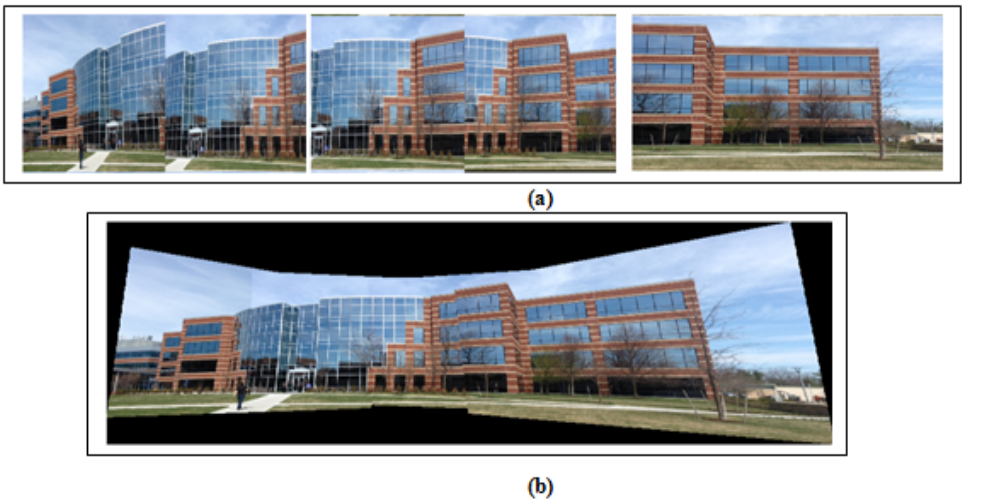


Figure (1): (a) Sequence of Frames and (b): The Result of Panorama Generate

2- Relate Works

For many years creating the panoramic image from the off set of video sequence or images has been widely studied. Basing on the different image stitching technologies, the methods for panorama generation can be classified mainly into three classes (Capture images, panoramas construct from the set of images taken from multiple viewpoints and last class construct the panoramic image from the video sequence which is taken by the camera with motion).The following is a review of different methods used for panoramic image and panoramic video generation:-

1- Hyun-Dae K., Yong-In Y., and Jong-Soo Ch**.** In 2003 [5] present a method for generating panoramic image. This method begins with determining the overlapping area between a sequence of images by using the method of phase correlation,then applying histogram equalization on images and selecting four seed points to estimate

the transformation between images. Detection of the corresponding point based on using a block matching algorithm (BMA) in the overlapped area of the image.

2- Abdulameer A. Kareem and Rana F. Ghani in 2010 [6] present a panoramic image construction algorithm based on three steps: image registration, focus sensing for the registered images, and image fusion that stitch the images at the overlapping location. To find the location of overlapping among the original images, this work using the probability of similarity which is obtained through XORing the edge pixels of the images to be stitched.
3- D'Orazio T., Leo M. and Mosca N. in 2010 [7] present a mosaic based approach for enlarged view the soccer video production that can be supplied to the audience as a complementary view for the greater enjoyment of the relevant events, such as counter attack ,offside or goal, which spread out all over playing zone.
4- Kwon O. and Yeong H. in 2010 [8] present a proposed method of the stitching of image depend on features color-invariant for the video automated panoramic. The SIFT produced algorithm is outstanding performance in the feature-point matching that is invariant to the rotation, scaling of images, and changes in illumination luminance.

3- Panorama Classifications

It is useful to classify Panorama by looking at the type of display. The Panorama is classified into static and dynamic panoramic, as shown in figure (2)[9].

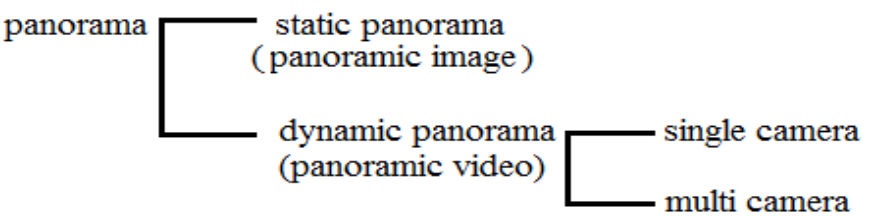
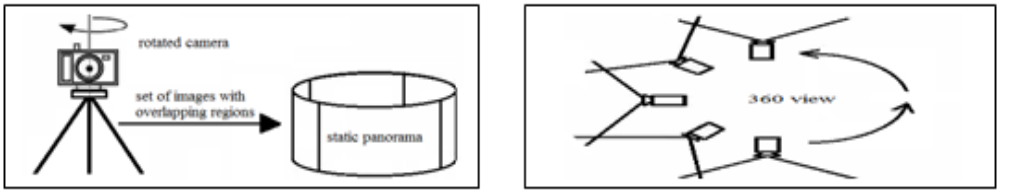


Figure (2): Classification of Panoramic

With static Panorama, the user is provided with a single still image of every direction of view, as shown in figure (3(a))[10]. Still Panorama is

typically created by taking a group of images of the scene [6].On panoramic video,a view is formed by digital video by either using a single camera system , or multi camera system to obtain the video [11]. Where, the digital video cameras are aligned in a circle form, and point out in different directions from the same viewpoint and the same angle and the translation between them (figure (3(b))). The Panoramic video is able to provide the audience with a video sequence, instead of a single still image, in any direction[12].



Figure(3): (a) Static Panorama and (b) Multi Camera System

4- Feature- Based Method

The feature can be defined as an interesting part image, and for the many computer vision algorithms, it is using as a starting point.This algorithm considers at a long time as an a good algorithm in feature detector [12]. The Features which are stable under noise and fixed to changes in the geometry are most useful [9]. The feature being useful by producing the same scene images values while the different values should produce the images of different scenes. The method of feature based depending on features extraction in images. Large regions (lakes, forests, fields), lines (coastlines, region surround, rivers, streets) or points (line intersections, region corners, points on curves with high curvature) it is consider as a feature [12]. In between the many kinds of the features that can be extracted from the image, extraction method of feature point required to satisfy the invariant to scaling, translation and rotation, were desirable in the process of the matching [12,13].

5- The Feature - Based Registration Method

By the transformation of geometric a two neighboring image can be related to each other. The transformation geometric is the relationship between two images such that one image has a point given on it compatible to the only one point in another. With recovering the transformation of geometric, images could be combined together & created the panorama [10].

The panoramic viewing will be constructed by multiple sample alignment with region overlapping. The operation of alignment needs a detection of images similar regions. A proposed feature based method will determine the similarities between images[10].

Constructing a panoramic video needs, taking a number of videos so as covering full space of viewing. So, those videos are composited & aligned into a full video of panoramic by using an accurate method of registration. The process of registration is a very important for viewing panoramic of generation. Registration of frame is an operation of matching two or more frames & estimating transformations of geometric between them, to align frames.

Step one is to capture two videos by using two normal cameras. Step two is the noise removal from frames. Then, each frame important feature to be extracted using the feature based algorithm. It needs every extracted feature to compute the descriptor which determines its invariability under different conditions. A region feature based method called (grid based method) will be used to calculate the feature descriptor. Then, the feature descriptors for both frames are matched by using a distance metric [12,14].

6- Panorama Construction

Construction of the panorama will be performed in a two steps: alignment & composition of an image. Alignment of the image based on the estimation of motion, while the composition of images uses the alignment of image in order to join all of the images into a single image[15].

Alignment using the feature- based algorithm consists of the important image parts like corners, points,lines,edges.Just the neighboring frames are searching for the matching of features [15].There are smaller feature groups for matching. The image transformation can be estimated and geometrically aligned when the matching features are gotten[16]. The figure(4) explains the alignment operation. One time the overlapping of images has been registered, they must combine together to make a single panoramic image.

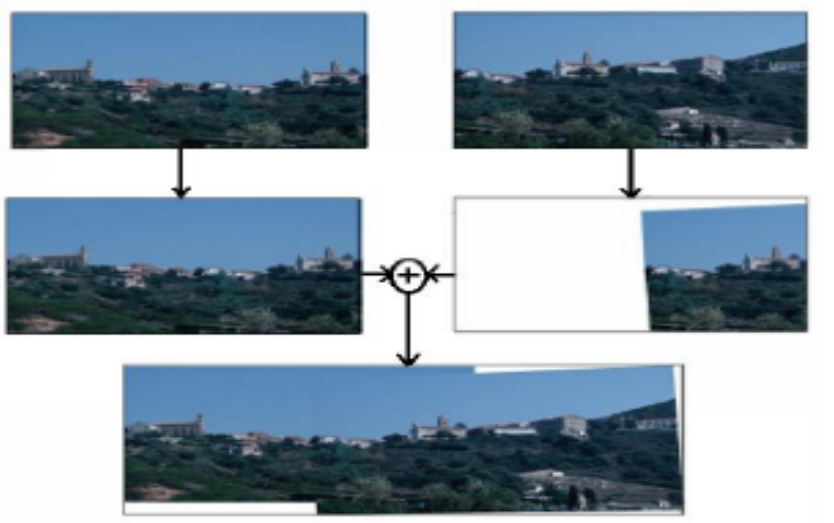


Figure (4) : Motion Estimation

The technique of mosaic construction (construction of panorama), is another method to show the whole scene. Works on view generation of panoramic could be classified into two main categories [13]:

1- Mosaic construction from static scenes.
2- Mosaic construction from dynamic scenes under two situations:
- Dynamic scenes captured with a static camera.
- Dynamic Scenes captured with a moving camera.

7- Image Registration Steps

The registration is an important operation to generate the video of the panoramic. Today registration of video plays a very important role in many applications. Consists of the estimation of motion for the video compression, change surveillance & detection, the creation of panorama.

The registration of video is the process of the aligning ensembles of the pixels over many time instants. The registration of video algorithms & applications could be categorized according to how they represent & extract the following entities [13].

1. Models of transformation between video frames.
2. Models of appearance of surfaces and objects.

Methods of registration consist of four steps [17]:

1- The detection of feature: Salient & distinctive features are automatically or manually detected.

2- Matching of feature: Detected the correspondence between features on 1st frame & 2nd frame is extracted.

3- Estimation transforms model: The mapping function parameters, aligning first frame with second frame, were estimated. Parameters of mapping function were computed by the correspondence of extract feature.

4- Transformation of frame:Second frame will be transformed by the function of mapping .

8- Grid - Based Shape Representation and Description Techniques

In the region-based techniques, shape region within the entire pixels have been taken into consideration to get the representation of shape ,instead of only using information of boundary as in the methods of contour base. Method of the common region based methods using the a describtion of moment in the description of the shapes. Another method of region based includes shape matrix, grid method, convex

divided into structural methods & global methods ,depend on whether they are separate shapes into subparts or not [18].

Shape is treated as a whole in global methods, the result is represented as the numeric vector of feature that could be used to describe the shape. The similarity of shape can be easily measured by a distance of metric between their vectors of feature [18,19].

The descriptor of grid shape is proposed and has been used by Lu and Sajjanhar [18]. Essentially, the grid of cells was overlaid on the shape, then the a grid was scanned from the left to the right & up to down. A bitmap is the result . Cells which covered by the shape are assigned one & those hasn't been covered by the shape is assigned zero. Then, a shape could be represented as the binary vector of feature.A binary hamming distance used to measure the similarity between 2 shapes. A descriptor advantages of a grid are its simplicity in the conformance to the intuition, representation & an agreement with the method of the shape coding in MPEG-4[18,20].

9- Transformation Estimation

The estimation of motion is the process of motion vectors determining which describe the transformation from one 2D image to another; typically from neighboring frames in a video sequence [15,16]. The vectors of motion may belong to the whole image (global motion estimation) or certain parts, such as arbitrary shaped patches, rectangular blocks or even per pixel. The finding motion vectors methods can be classified into feature based methods (indirect) and pixel based methods (direct). The example of the direct method is the method of block matching. While, indirect methods using features, like lines, points, & match corresponding the features between the frames [20].

❖ **Nearest Neighbor Matching Algorithm**

This algorithm using the distance ratio between the neighbor nearest feature points to the 2^{nd} neighbor nearest feature points to make matching feature points. Assuming (I) is a feature point on image 1 & (J) represent a feature point of nearest neighbor in

the image 2. If a nearest distance rao to 2 nd nearest distance was less than the certain threshold, then these pairs of points registration was matched,as shown in the following equaon[14] :

$$D_{(i,j)} = \sqrt{\sum_{i=1}^{n} (DP1_{(Ii)} - DP2_{(Ji)})^2} \qquad \text{.....(1)}$$

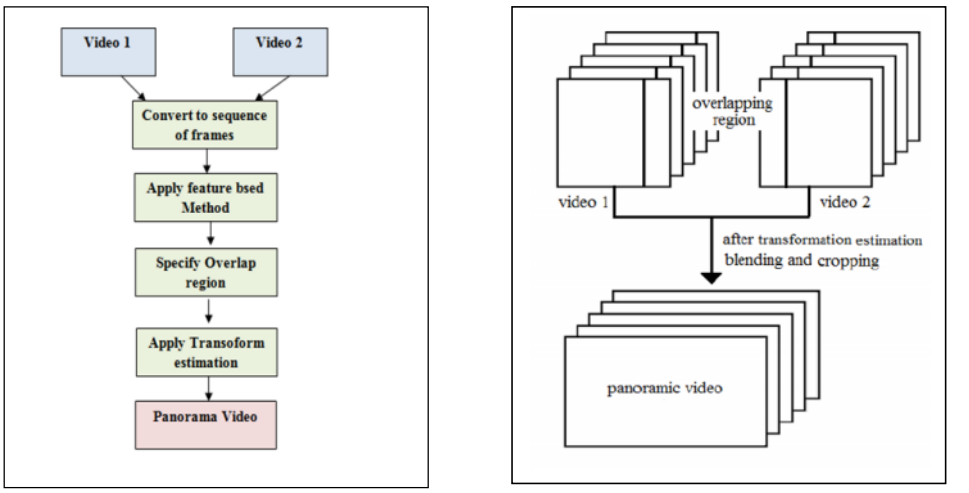
Where :

$DP1(Ii)$:→ Vector describing point(I) in the image1.
$DP2(Ji)$:→ Vector describing point (J) in the image2.

n :→ Is the number of the descriptor elements.
$D(i, j)$:→ Is distance between ($DP1(Ii)$) and ($DP2(Ji)$).

10-The Proposed of Panoramic Video System

The first main objective of the system of panoramic generation depends on the capturing of samples by using normal cameras and the operation will be done through capturing samples for fixed place at the same time. Secondly, the estimation motion between samples transformation considers on the challenge in the field which depends on the feature-based method.Feature- based method must be strong enough to extract the fixed features in different circumstances like: distortion,light change,noise and blurring by different transformations (scaling,rotation ,translation ... etc.), see figure (5).



(a) **(b)**

(c)

Figure (5): Panoramic System :(a) Structure of system , (b) Panorama Constructor and (c) Overlap Rigon Between two Frames

The steps of the test proposed method to generate panoramic video are illustrated in figure (6).

Input : Two video

Output : G enerte Panoramic video

Step 1: Begin

Step 2: Divided the two video into two sequence of frames.

Step 3: Analyze the frame for each video, then extract information useful for reducing the a mount of frame data and representing it in a frame of features by extract basic feature.

Step4: Apply descriptor , after extract features. Descriptor is constructed for each feature and using shape representation .

Step 5: Find matching the description between the test frames by using nearest neighbor method.

Step 6: After complete the operation of matching between two video, the motion between the frames must be estimated in accurate from to avoid the stitching errors in the result panoramic scene.

Step 7: Display the result (panoramic scene).

Step 8: End.

Figure (6): Generate Panoramic Algorithm

11- Experimental Results and Discussion

In this paper, the experimental test is used (AVI) video format with the frequency of 30 frames per second on video(AVI video format used). The samples were taken using Nikon Camera in different places like university and theme park. The captured video must have an overlap region to satisfy the condition of the generation the panoramic scene. All video samples are implemented in different transforming and lightening circumstances. To validate the system performance,the assumed method is applied on two videos(university and theme park), the first step converts the two videos into a sequence of frames, as shown in figure (7) and figure(8).

(a)

(b)

Figure (7): University Video(frames (1,55,80,100)) (a) Camera 1 and (b) Camera 2

(a)

(b)

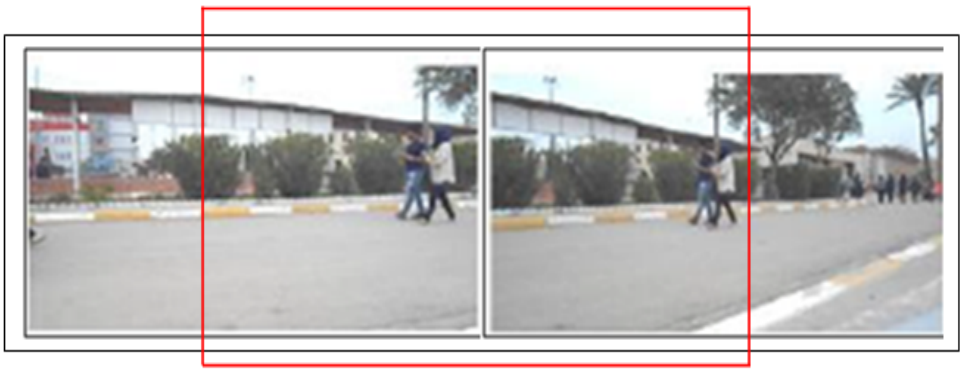
Figure (8): Theme Park Video (Frames (1,40,90)), (a) Camera 1 and (b) Camera 2

After converting the two videos to the sequence of frames ,then extract the feature and specify the overlap region and apply transforms estimation, the results represent a figure (9), figure (10).

(a)

(b)

Figure (9) : (a) Overlap Region and (b) Result of Panorama Video of Frames (1,55,80,100)



(a)

(b)

Figure (10): (a) Overlap Region and (b) Result of Panorama Video of Frames (1,40,90)

12- Conclusions

This work represents a framework for generating panoramic video by using shape representation and depending on video taken by two cameras, then convert to sequence of frames and the important feature for each frame are extracted by feature- based method, this feature needs to calculate the descriptor, the feature descriptor is computed by using region -based algorithm, after that find matches between the

frames, then the panoramic video generates after detecting the overlap region between the frames of video by applying shape representation.

In this work, the panorama technologies provide users with the ability to experience an environment and then. Pick the direction of view in the scene. The most important point in the work is that the videos are taken using normal cameras, in different transformations and light conditions. and taken from different viewpoints. Also, in this work the panorama is generated overlapping region between samples while the previous work generates the panorama from multiple images or single camera.

13- Reference

[1] Brown M. and Lowe D., ***"Recognizing Panoramas"***, IEEE, In *Proc. ICCV'03*, France, Vol. 2, Pp: 1218–1225, 2003.

[2] Alan C. Bovik, ***"The Essential Guide to Video Processing"***, Elsevier inc., 2009.

[3] Irani M., Anandan P., ***"Video Indexing based on Mosaic Representations"***,Proceedings of the IEEE, 1998.

[4] Agarwala A., Zhang K., Pal C., Agarwala M., Cohen M., Curless B., Salesin D.& Szeliski R.,***"Panoramic Video Textures"***, University of Washington, Microsoft Research and University of Massachusetts, (2005).

[5] Hyun-Dae K., Yong-In Y., and Jong-Soo Ch**.,** ***"An efficient method to build panoramic image mosaics"*****,** Elsevier, pattern recognition letters, Vol. 24, issue 14, 2003.

[6] Abdulameer A. Kareem and Rana F. Ghani, ***"Panoramic Image Construction Using Edge Pixels Connectivity"*****,** first conference for computer science, computer science department, university of technology, 2010.

[7] D'Orazio T., Leo M. and Mosca N., ***"Panoramic Video Generation by Multi View Data Synthesis"***, International Conference on Pattern Recognition, Italy, 2010.

[8] Kwon O. and Yeong H., ***"Panoramic Video using Scale-Invariant Feature Transform with Embedded Color-Invariant Values"***, IEEE Transactions on Consumer Electronics, Vol. 56, No. 2, 2010.

[9] Imran A., James C., ***"A Panoramic Video System"***,

Department of Computer Science, University of the Western Cape, South Africa, 2010.

[10] Szeliski, R., Shum H., ***"Creating Full View Panoramic Image Mosaics and Environment Maps"***, ACM SIGGRAPH, 251-258, 1997.

[11] Wei X. and Jane M., ***"Panoramic video stitching from commodity HDTV cameras"***, Springer, Multimedia Systems, Vol. 19, Issue 5, Pp 407-426, 2013.

[12] Byeong-Ho K., ***"A Review on Image and Video processing"***, International Journal of Multimedia and Ubiquitous Engineering, Vol. 2, No. 2, 2007.

[13] Barbara Z. and Flusser J., ***"Image Registration Methods: A Survey"***, Elsevier, Image and Vision Computing, 2003.

[14] Abddul S. R, Kawther A. S.," ***Feature Descriptor Based on Normalized Corners and Moment Invariant for Panoramic Scene Generation",*** (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 5, No. 7, 2014.

[15] Pratt K. William, ***"Digital Image Processing"***, Third Edition, John Wiley & Sons, Inc., ISBN: 0-471-37407-5, 2001.

[16] Xunyu P., ***"Motion Panorama Construction from Streaming Video for Power Constructed Mobile Multimedia Environment",*** Master Thesis, Georgia University, 2004.

[17] Philip H. and Zisserman A., ***"Feature Based Methods for Structure and Motion Estimation"***, ICCV Workshop on Vision Algorithms, Pp: 278-294, 1999.

[18] Changsoo J., Hyung-Min P., ***"Optimized Hierarchical Block Matching for Fast and Accurate Image Registration"***, Elsevier, Signal Processing: Image Communication Vol. 28, Pp:779–791, 2013.

[19] NagaRaju N., Satya T., and Swamy Ch., ***"Image Registration Using Scale Invariant Feature Transform"***, International Journal of Scientific Engineering and Technology, Vol. 2, No.7, Pp : 675-680, 2013.

[20] A. Sajjanhar , Region G.J. Lu, , ***"Region-based shape representation and similarity measure suitable for content-based image retrieval"***,Multimedia Syst. 7 (2) (1999) 165–174.

استخدام طريقة تمثيل الشكل لبناء نظام فديو بانورامي

م.د. عبير سالم جميل*،أ.د. عبد المنعم صالح رحمة**،م.د. ندى حسين محمد علي***

المستخلص

لقد اصبح توليد المشهد البانورامي مجالا فعالا للبحث . ان اهمية توليد البانوراما هو لزيادة نطاق الرؤيا للمشهد باستخدام الفديو لتوليد فديو بانورامي. ان استخدام كامرا واحدة غير كافية لتوليد نطاق رؤيا مناسب. لحل هذه المشكلة يتم استخدام كامرتين او اكثر لتوليد فديو بانوراما ذات نطاق رؤيا واسع افضل مما لو تم استخدام كامرا واحدة، حيث انه كلما زاد عدد الكامرات زادت القابلية على توليد نطاق الرؤيا .

ان الهدف الاساسي من هذا البحث هو انشاء نموذج فديو جديد يدعى فديو بانورامي باستخدام اكثر من فديو ، لهذا الغرض تم اقتراح طريقتين للانشاء . الاولى هي القيام بعملية التقاط عينات لمشاهد معينة بواسطة كامرات عادية . الثانية هي القيام بحساب درجة التحول ما بين العينات التي تم تسجيلها وتسجيل درجة الحركة وذلك من خلال الاعتماد على خوارزميات الخصائص كأساس

(**Feature Based**)لاستقطاع المعلومات المهمة ويعتبر التحدي الرئيس في هذا البحث، وكذالك تمثيل الشكل **(Shape Representation)** وطريقة المنطقة كأساس(**Region- Based Method)** لحل مشكلة منطقة التداخل **(overlap Region)** مابين العينات التي تم التقاطها من خلال الكامرات وذالك للحصول على فديو بنورامي صحيح باستخدام كامرتين في اماكن مختلفة .

الكلمات المفتاحية: فديو بنورامي، تمثيل الشكل، منطقة التداخل ، توليد البانوراما

***كلية المنصور الجامعة**

****الجامعة التكنولوجية**

*****كلية العلوم، جامعة بغداد**