Eng. &Tech.Journal, Vol. 33,Part (B), No.1, 2015

Real-Time Panoramic Video Construction using Harris Corners Detector

Dr.Abdulameer A. Kareem
Computer Science Department, University of Technology/Baghdad.
Email:ameer.aldelphi@yahoo.com
Saif Salah Abood
Computer Science Department, University of Technology/Baghdad.
Email:forsaif_82@yahoo.com

Received on: 14/9/2014 & Accepted on: 8/1/2015

ABSTRACT

This paper discusses the possibility of stitching two video streams to create a panoramic view in real-time, The challenge with constructing Panorama Video in real-time is the time-consuming in stitching images.

Image alignment didn't need to detect all interest point, only few well distributed point is enough to calculate good projection matrix, and the great similarity between the frames in video stream provides temporal information which can be used to reduce the number of selected points

In this paper the stitching images method which using Harris Corners Detector has been customized to be appropriate to meet the requirements of viewing a video in real time, through reduce the number of Interest Points in each stitching and employ the Parallel Processing to separate the calculation of transform matrix from the process of blending.

Reducing the number of Interest Points has been done by fragmentation of images into numerous regions and getting a feedback information from the stitching of the previous frame to limitate the search space based on the area of overlapping.

Results shows that (i) panoramas generated from the proposed algorithm feature a smooth transition in image overlapping areas and satisfy human visual requirements; and (ii) the preview speed of the generated panorama satisfies the real-time requirement that are commonly accepted in video panorama stitching.

بناء بانوراما فيديو في الوقت الحقيقي باستخدام كاشف الزوايا هاريس

الخلاصة

يناقش هذا البحث إمكانية دمج تيارين من الفيديو لإنشاء عرض بانورامي واحد في الوقت الحقيقي، إن التحدي في بناء بانوراما الفيديو في الوقت الحقيقي هو طول الوقت الذي يستغرقه دمج الصور.

إن دمج الصورة مع بعضها لا يحتاج الكشف عن جميع نقاط الاهتمام الموجودة في الصورة، فان عدد قليل من النقاط الموزعة بشكل جيد يكفي لحساب مصفوفة الإسقاط، وان التشابه الكبير بين إطارات الفيديو يوفر معلومات زمنية يمكن استخدامها للحد من عدد نقاط الاهتمام المكتشفة.

https://doi.org/10.30684/etj.33.1B.13
2412-0758/University of Technology-Iraq, Baghdad, Iraq

في هذا البحث يتم تعديل طريقة دمج الصور التي تستخدم كاشف الزوايا هاريس (Harris) لتكون مناسبة لتلبية متطلبات عرض شريط فيديو في الوقت الحقيقي، وذلك من خلال تقليل عدد نقاط الاهتمام في كل عملية دمج واستخدام المعالجة المتوازية بدل المعالجة المتسلسلة لفصل عملية حساب مصفوفة التحويل عن عملية الدمج .

إن تقليل عدد نقاط الاهتمام يتم من خلال تجزئة الصور إلى عدد من المناطق ومن خلال الحصول على معلومات التغذية الراجعة من خياطة الإطار السابق لتقليص مساحة البحث استناداً إلى مساحة التداخل بين الصور.

النتائج تظهر أن اولاً: البانوراما المتولدة من الخوارزمية المقترحة تمتاز بالانتقال السلس في مناطق تداخل الصور وتلبي المتطلبات البصرية للإنسان. ثانياً: إن سرعة عرض البانوراما تلبي متطلبات الوقت الحقيقي التي تكون مقبولة عموماً في دمج بانوراما فيديو.

INTRODUCTION

Panorama means constructing of high resolution images with large field of view. The terminology "Panoramic Video" sometime used to refer to the construction of a static panoramic image from one video stream, but in this work will be used to refer to the stream of panoramic frames constructed from two or more video streams.

Image stitching is based on finding the similarity between adjacent images and determination of overlapping regions. Image stitching (Image Registration or Image Alignment) algorithms can be classified into three different categories:

(i) Intensity-based algorithms which need a large amount of computation and in case of do not give good result rotation and scaling [1] [2].

(ii) Frequency-domain-based algorithms are faster with small translation, rotation, and scaling, but it fail with small overlapping regions [3] [4].

(iii) Feature-based algorithms, which reduced the computational complexity by processed a small amount of information (points, lines, or edges) to align images, and they are robust to changes in image intensity [5] [6] .

Therefore, most of the researches has been focused on feature-based algorithms. one of the most commonly used feature-based algorithms is the one which has been shown in Figure (1), in which Harris Corners Detector [7][8] detect the common points (Interest Points) between the two images, these points (Corners) have a large intensity variations between the directions around.

then these points are matched by Normalized-Cross-Correlation (NCC) [9] to generate a candidate pairs. It found that some of these pairs are wrong and must be eliminated, so the Random Sample Consensus algorithm (RANSAC) [10] has been used, which estimate a mathematical model representing the largest number of these pairs and remove all the pairs that do not fit with this model. After that Projection one of the two images overlaid on top of the other and calculate the Transform Matrix to transform this image, finally the two images blended together [11].

This algorithm is classified as a feature-based algorithm and it success with small overlapping areas but with small difference in rotation, and scaling.

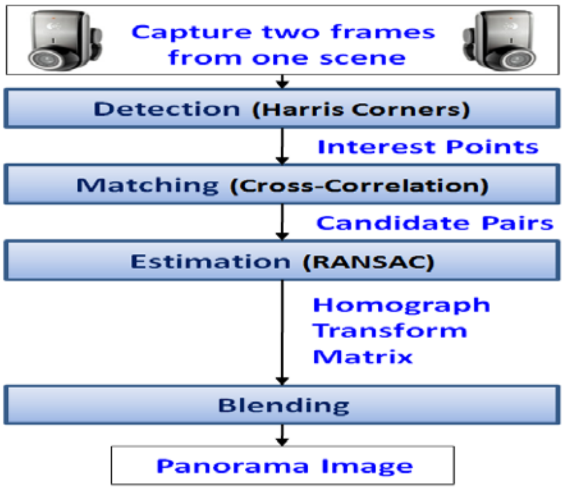


Figure (1): The structure of the traditional images stitching system

In this work ,the above traditional image stitching system has been improved through three main orientations , first one, the Fragmentation of images into regions can reduce the number of Interest Points and thus reduce the time of matching and correlation, while maintaining the distribution and avoid clustering of the Interest Points in one side of the frame. As well as Feedback information from stitching previous frame can reduce the number of Interest Points by limiting the search space based on previous Area of overlapping.

In addition, the technique of threading has been used to separate the process of calculating the Homograph Transformation Matrix which takes a few seconds than the process of Transforming the frames and Blending them, to reach a rate acceptable according to the requirements of Display Video in real-time

The rest of this paper is organized as: Section 2 “Related Work” briefly discusses related work; Section 3 “Concepts of the Proposed Video Stitching Algorithm” describes three basic concept of the solution; Section 4 “Structure of the Proposed Video Stitching Algorithm” Introduces the Proposed Algorithm; Section 5 “Performance Evaluation and Results Discussion” describes experimental results; and finally Section 6 “Conclusions”.

Related Work

When studying the subject of video and image stitching , it is clear that there are many works relating image stitching, on the other side very little works has been done on the video stitching, especially considering the case when the capturing devices are allowed to move freely, as opposed to the case where the cameras are fixed.

The authors in [5][6], proposed a video stitching system combine multiple video feeds from ordinary cameras or webcams into a single panoramic video. And to reduce the computation satisfy real-time requirements, the parameters for stitching are calculated only once, during the system initialization. So their system requires an initialization phase that is not real-time and is needed whenever the cameras or webcams positions change.

In [11] the authors presented a new improved algorithm, self-adaptive selection of Harris corners by fragment the image into regions and select corners according to the normalized variance of region grayscales. they apply multiple constraints, e.g., their midpoints, distances, and slopes, on every two candidate pairs to remove incorrectly matched pairs. Their algorithm, solved the video stitching problem on a frame-by-frame basis, and does not fully exploit the correlation between video frames, and requires to identify which of the two images on the right and which is on the left, and matching only the right-third of the left image with the left-third of the right image. finally they fixed the total number of points that are matching in each stitching.

The authors in [12] proposed a system for stitching videos streamed previously recorded by freely moving mobile phones cameras. However, their proposed system treats the video stitching problem essentially as an image stitching problem on individual frames independent of each other.

In [13] the authors exploits temporal information to avoid solving the stitching problem fully on a frame by frame basis. by employ previous frames stitching results such as tracking interest points using optical flow and using areas of overlap to limit the search space for interest points.

In all these related works it was found that the most time-consuming steps is the computation of detection and matching large number of Interest Points, it takes more than 85% of the time in stitching two frames. While only four well distributed Interest Points are enough to calculate the Homography Transform Matrix. Hence, most of the study approaches in this paper are aiming at detection and matching less number of Interest Point.

Concepts of the Proposed Video Stitching Algorithm

The performance of the proposed video stitching algorithm is based on three basic concepts:

Fragmentation

With Harris corners detector, the corners tend to cluster around regions with richer texture, whereas fewer corners will be selected in regions with less texture information. That mean , selected corners are not evenly distributed. Therefore, each frame will be fragmented into ($n \times n$) regions where n is the number of fragments in the one side of frame, and select one corner from each region by sorted all candidate corners in their values, The number of all selected corners will not be more than the number of fragments ($n \times n$), this number of fragments will be modified depending on the feedback from stitching the previous frame, (Eq. (1)) where i is the current frame stitching.

$$n_{i+1} = \begin{cases} n_i - 1\,, \text{if detected corners} > 100\ \mathit{and}\ n > 10 \\ n_i + 1\,, \text{if detected corners} < 50\ \mathit{and}\ n < 20 \\ n_i \qquad ,\mathit{else} \end{cases} \qquad \ldots.(1)$$

If the number of detected corners in the previous stitching are few, then the fragments number will be increased and reduced its size, and vice versa.

This fragmentation will enable to reduce the time of matching and correlation , while maintaining the distribution and avoid clustering of the corners in one side of the frame.

Exploit the Area of Overlap from Previous Frame

The area of overlap from frames (i) can potentially limit the search space for Interest Points in frames (i+1). The stitching algorithm performs the same steps as in the case of the first frame pairs but instead of trying to detect Interest Points in the whole frame (i+1), they are detected only in the area of overlap which found in frames (i), plus some buffer region in the frames (i+1) (best value experimentally determined to be a 7 pixel band). The efficiency of the buffer based approach performance is inversely proportional to the size of the overlap area between frames (n-1) and (n).

Parallel Processing

Parallel processing is the simultaneous use of more than one CPU or processor core to execute a program or multiple computational threads. Theoretically, parallel processing makes programs run faster because there are more (CPUs or cores) running it. In practice, it is often difficult to divide a program in such a way that separate CPUs or cores can execute different portions without interfering with each other, all new CPU models have multi-core processor chips with Hyper-threading Technology.

This work need one thread for each camera to capture the video stream, and two threads for two spirited loops for stitching operations, and more one or two thread for monitoring the result and record the output video.

Chart (1) shows the time line of the parallel processing in T_0 CPU(1,2) begin capturing frames from the two cameras, in T_1 finish capturing first frames (A_1,B_1) and begin calculate Transform Matrixes, in T_2 finish calculate first Matrix (M_1) and CPU(4) begin blending A_n with B_n*M_1, in T_3 finish calculate Matrix (M_2) and begin blending A_{2n} with B_{2n}*M_2.

		T_0	T_1		T_2			T_3	
Time Line									
CPU (1) Capture Cam 1 frame	A_1	A_2	…	A_n	A_{n+1}	…	A_{2n}	A_{2n+1}	…
CPU (2) Capture Cam 2 frame	B_2	B_2	…	B_n	B_{n+1}	…	B_{2n}	B_{2n+1}	…
CPU (3) Calculate Transform Matrix		M_1			M2			M_3	
CPU (4) Blending					$A_n+B_n*M_1$	$A_{n+1}+B_{n+1}*M_1$	…	$A_{2n}+B_{2n}*M_2$	…

Chart (1): The time line of the parallel processing

Structure of the Proposed Video Stitching Algorithm

As previously mentioned the proposed algorithm have two spirited loops for stitching operations, First loop contain three steps, Detection , Matching and Estimation & Projecting, which is the first three main steps in image stitching algorithms, and the second loop contain Blending step, which is the forth main steps in image stitching algorithms. As in the Figure (2).

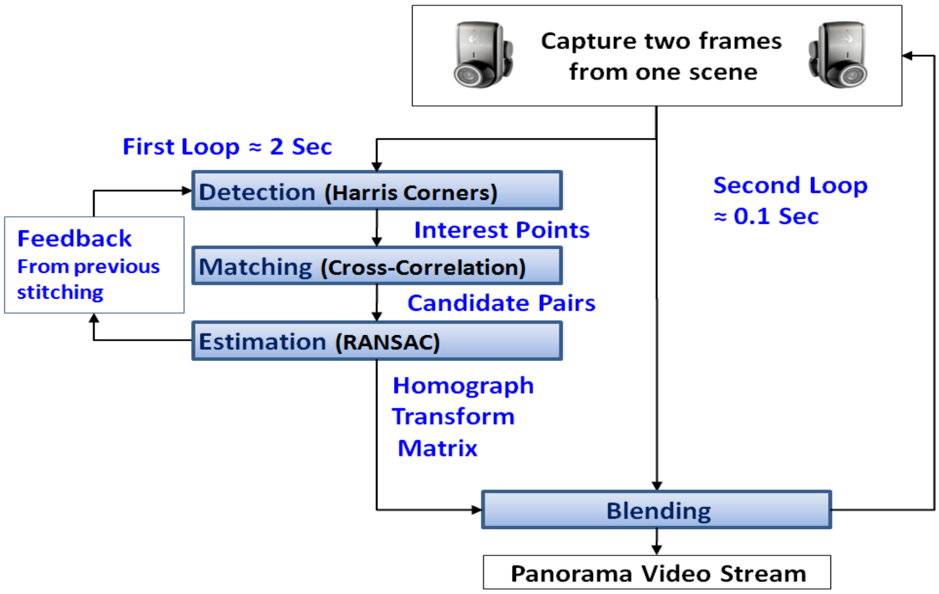


Figure (2): The structure of the proposed image stitching system

Feedback contains three key information: the result of previous stitching (success or failure), the overlap area boundaries, and the number of the Interest Points in the previous frame.

Inability of the algorithm to stitch frame pairs can be due to either: (i) there is no enough overlap between frames (ii) enough correlation pairs between frames, can't be found (iii) the algorithm computes an incorrect geometric alignment.

In the Blending step the degradation method has been used, by linear gradient Transparency from the center of one image to the other

Performance Evaluation and Results Discussion

Experiments were conducted on an Intel® Core™2 Duo CPU T6500@ 2.10GHZ, with 4.0 GB RAM, NVIDIA GeForce G 105M video card and two Logitech C905 webcam HD720p , USB 2.0, Ultra-smooth, light correction, Auto Focus with Carl Zeiss optics and 2-megapixel resolution for video and photos up to 8-megapixels. The application was built using C#.Net framework 4.0 and Accord.NET 2013 library. The details (dimension H×W, size of two image, overlapping between two image, distribution of corners, horizontal line, point of view and scene depth) of images pairs that have been used to test the algorithm are shown in Table (1) and Figure (3).

Table (1): Details of the testing images pairs

Images	Dimension	Size	Overlapping	Corners	Horizontal	POV	Scene
Lake	640×640 pixel	73KB 82KB	486 pixel = 76 %	clustering	similar	similar	too far
House	640×640 pixel	121KB 132KB	235 pixel = 36 %	distribution	different	different	close indoor
School	1024×768 pixel	151KB 157KB	747 pixel = 73 %	clustering	different	similar	nearby outdoor
Nature	1670×1024 pixel	846KB 914KB	775 pixel = 46 %	distribution	similar	similar	far

Lake:


House:


School:


Nature: 

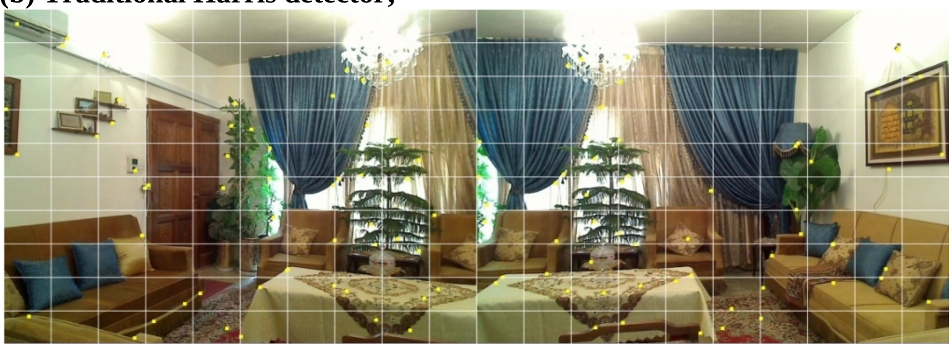
Figure 3. Images pairs which used to test the algorithm:

Evaluation of Fragmentation

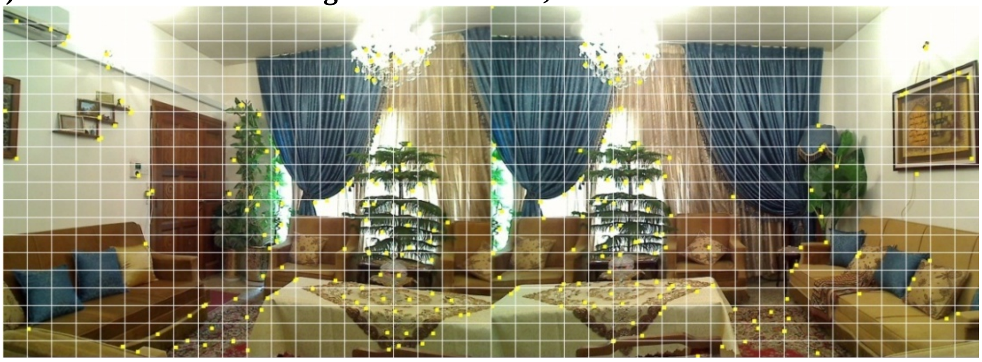
The experimental results on the (House) image pair shown in Figure (4). The result of applying the traditional Harris detector on the original image, Figure (4.a), is shown in Figure (4.b) give 351×361 corners, and the result with fragmentation 10×10 in Figure (4.c) give 55×57 corners, also with fragmentation 20×20 in Figure (4.d) give 131×132 corners.

(a) Original image pair;

(b) Traditional Harris detector;



(c) Harris detector with fragmentation n=10;



(d) Harris detector with fragmentation n=20;

Figure (4). Fragmentation on House image.

To be more specific, with increasing the value of *n*, the number of corners will be increased and the result will be closer to that of the traditional Harris detector. On the other hand, if the value of *n* is small, the number of corners will be decreased, knowing that the execution time has not changed Table (2) content the results of Applying the Fragmentation on the four images pairs with different value of *n* (order by the number of corners), and Chart (2) show the number of corners VS the number of Fragment.

Table (2): Details of testing Fragmentation on four images pairs

Images	corners with Traditional Harris	corners with Fragmentation *n*=20	ratio of reduced corners	corners with Fragmentation *n*=15	ratio of reduced corners	corners with Fragmentation *n*=10	ratio of reduced corners
Lake	77×132	40×58	53 %	30×40	66 %	18×22	80 %
House	351×361	131×132	63 %	98×88	73 %	55×57	84 %
School	524×519	151×125	73 %	111×91	80 %	68×56	88 %
Nature	2622×4404	295×324	91 %	177×194	94 %	88×93	97 %

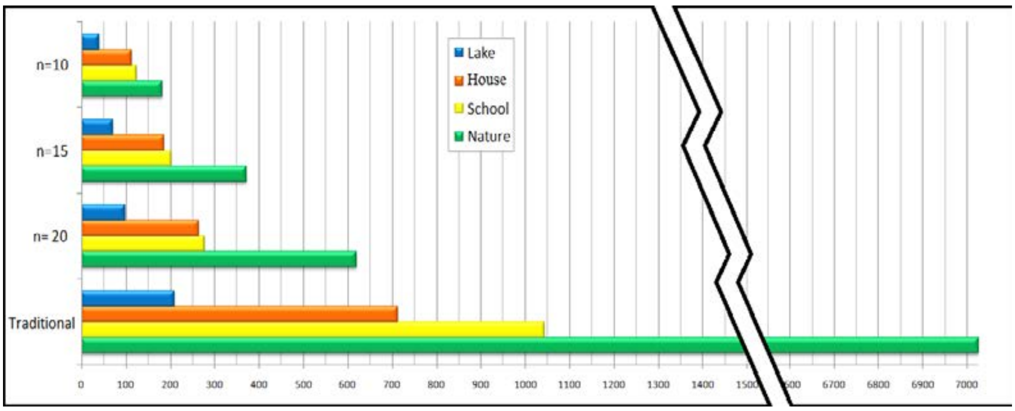


Chart (2): Details of Fragmentation VS number of corners

It is clear from the previous, that benefit of fragmentation increases with the large-size and high-resolution images.

Evaluation of Limitation the Search Space

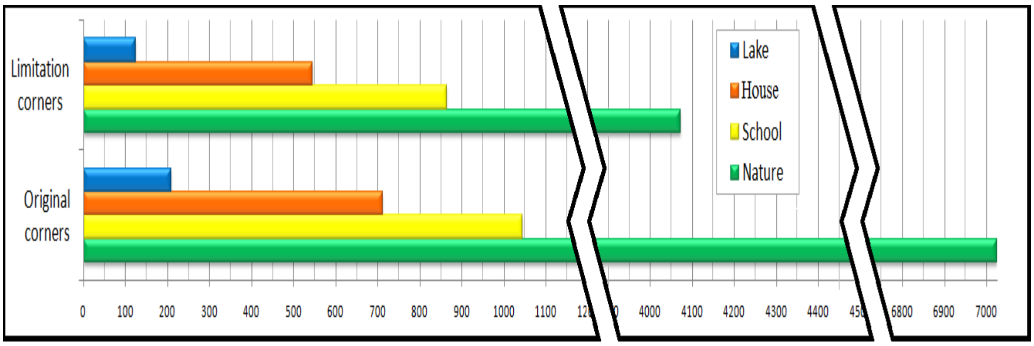
The results of limitation the search space based on the area of overlap in the previous frame on the image pair (House) (without using the Fragmentation) is shown in Figure (5). Where the number of the selected corners become 249×295 corners instead of 351×361 corner in traditional Harris detector. As well as the time of the search became 0.34 seconds instead of 0.82 seconds.

Figure (5). Limitation the search space on House image

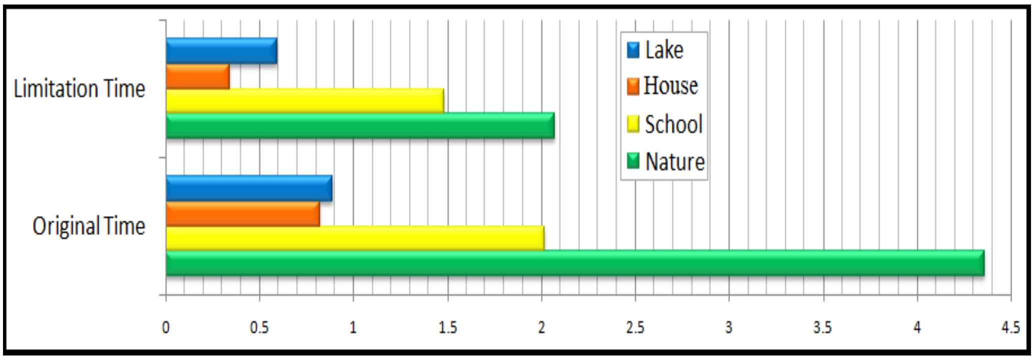
Table (3) and Chart (3) are illustrate the change in selected corners limitation the search space on the four images pairs, and how the usefulness of limitation is directly proportional to the ratio of overlap between images

Table (3): Details of testing Limitation on four images pairs

Images	Overlapping	original time	limitation time	reducing time	original corners	limitation corners	reducing corners
Lake	76 %	0.89	0.59	33 %	77x132	62x61	41 %
House	36 %	0.82	0.34	58 %	351x361	249x295	23 %
School	73 %	2.02	1.48	26 %	524x519	432x432	17 %
Nature	46 %	4.36	2.07	52 %	2622x4404	2044x2029	42 %



(A): Change in selected Corners



(B): Change in Time

Chart(3): Details of testing Limitation on four images pairs

Evaluation of Parallel Processing

It is clear that the very little time required for the blending operation, makes the use of parallel processing provides the possibility of merging video frames in real time (camera frame rate), but with a delay in responding to any movement of the scene or camera. This delay has been reduced significantly when using the proposed algorithm, and the results will show in detail in the final Result.

Final Results

The implementation of both fragmentation and limitation on the pair (House) will decrease the total execution time of all stitching operations from 0.95sec. to 0.36sec. , without any change in the quality of the resulting panorama. Figure (6) show the four stitching operations on the image pair (House) with using the proposed algorithm. Table (4) show details of outputs and execution time of each stitching operations in proposed algorithm compared with the outputs and execution time in the traditional image stitching.

(a) Harris Detection with fragmentation & Limitation;

(b) Matched by (Normalized-Cross-Correlation);

(c) Estimation & Projec ting by RANSAC;

(d) Blending ;

Figure (6). stitching the image pair (House) using the proposed algorithm:

Table 4: Details of outputs and execution time of each stitching operation in proposed algorithm on the four images pairs

Images	Algorithm	Point selection by Harris Detector	Matched by (NCC)	Estimation & Projecting by (RANSAC)	Total Time	percentage of reducing time
Lake	Traditional	77×132 Corners 0,79 sec.	48 pairs 0.03 sec.	42 pairs 0.04 sec.	0.86 sec.	
	New	26×30 Corners 0.61 sec.	18 pairs 0.01 sec.	17 pairs 0.01 sec.	0.63 sec.	26 %
House	Traditional	351×361 Corners 0.79 sec.	124 pairs 0.15 sec.	75 pairs 0.01 sec.	0.95 sec.	
	New	64×69 Corners 0.34 sec.	18 pairs 0.01 sec.	11 pairs 0.01 sec.	0.36 sec.	62 %
School	Traditional	524×519 Corners 2.01 sec.	149 pairs 0.34 sec.	75 pairs 0.01 sec.	2.36 sec.	
	New	101×87 Corners 1.46 sec.	31 pairs 0.03 sec.	22 pairs 0.01 sec.	1.5 sec.	36 %
Nature	Traditional	2622×4404 Corners 4.14 sec.	1943 pairs 13.68 sec.	1908 pairs 0.02 sec.	17.84 sec.	
	New	181×175 Corners 2.01 sec.	96 pairs 0.03 sec.	96 pairs 0.01 sec.	2.05 sec.	88 %

DISCUSSION

From comparing the percentage of reducing time between the execution time of the traditional algorithm and the execution time of the proposed algorithm with the percentage of overlapping area in each image pairs, find that the percentage of reducing time (i) directly proportional to the size and resolution of the image, In other words, percentage of reducing time increases with larger and high resolution images, because these image have a lot of interest points that can be discarded, (ii) an inverse proportional to the area of the overlapping , In other words, percentage of reducing time increases with small overlap area, because the search in a smaller space needs less time.

CONCLUSIONS

A new algorithm has been presented to handle challenges in real-time video panoramic Construction, which is time-consuming of stitching processes. the contribution can be summarized by improving images stitching algorithm to reduce the number of Interest Points in each stitching process through: (i) Employ Fragmentation to avoid the points clustering as much as possible, (ii) Employ Feedback to update the Fragmentation and limitate the search space by exploit the previous overlapping area. (iii) Employ Parallel Processing which has been resulted in more efficient panorama stitching process. Experimental results shows that (i) panoramas generated from the proposed algorithm feature a smooth transition in image overlapping areas and satisfy human visual requirements; and (ii) the preview speed of the generated panorama satisfies the real-time requirement that are commonly accepted in video panorama stitching.

REFERENCES

[1]. Anat Levin, Assaf Zomet, Shmuel Peleg, Yair Weiss "Seamless Image Stitching in the Gradient Domain ", Computer Vision - ECCV 2004, Volume 3024, pp 377-389

[2]. Chalechale A, Naghdy G, Mertins ," A Sketch-based image matching using angular partitioning systems", IEEE Trans on Man and Cybernetics and Part A 35: 28–41. , 2005.

[3]. He, Y., Wang, L. ,"Image stitching algorithm based on feature block and wavelet transform", Computer Engineering and Design 31, 2010.

[4]. Sivic J, Zisserman A ," Video google: A text retrieval approach to object matching in videos". In: Proceedings of the 9th IEEE International Conference on Computer Vision. IEEE Computer Society Washington DC and USA: Computer Vision-Volume 2, 27–38, 2003.

[5]. M. Zheng, X. Chen and L. Guo, “Stitching Video from Webcams”, Advances in Visual Computing Volume 5359, 2008, pp 420-429

[6]. W. Tang, T. Wong and P. Heng, "A System for Real-time Panorama Generation and Display in Tele-immersive Applications ", Multimedia, IEEE Transactions on (Volume:7 , Issue: 2) April 2005.

[7]. C. Harris, M. Stephens, A Combined corner and edge detector, in: Proceedings on Fourth Alvey Vision Conference, Manchester, UK,1988, pp.147–151.

[8]. Harris CG, Stephen M "A combined corner and edge detector", 4th Alvey vision conference. Manchester, 147–151(1998).

[9]. Feng Zhao, Qingming Huang and Wen Gao, “Image Matching By Normalized Cross-Correlation ” 142440469X/ ©2006 IEEE II 729 ICASSP 2006

[10] .M. Fischler and R. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography”, Communications of the ACM, 24(6), 1981.

[11]. Minchen Zhu, Binghan Liu ,Weizhi Wang , Jingshan Huang , " Efficient Video Panoramic Image Stitching Based on an Improved Selection of Harris Corners and a Multiple-Constraint Corner Matching", PLOS ONE, December 2013 | Volume 8 | Issue 12 | e81182

[12]. A. Kaheel, M. El-Saban, M. Refaat, and M. Izz, “Mobicast - A system for collaborative event casting using mobile phones”, ACM MUM , 2009.

[13]. M. El-Saban, M. Izz and A. Kaheel, “Fast stitching of videos captured from freely moving devices by exploiting temporal redundancy”, ICIP 2010.

[14]. Rana F.Ghani & Abdulameer A. Kareem "Multi Focus Image Fusion Using Statistical Approaches " Eng. & Tech. Journal, Vol28, No.2,2010