Al-Nahrain University, College of Engineering Journal (NUCEJ) Vol.17, No.2, 2014, pp.221-228

Panoramic Image Generation Algorithm based on Hu's Moment Invariants

Kawther Abbas Sallal
Computer Science Dept, Al-Mustansiriyah University, Iraq,

Abdul_monem Saleh Rahma
Computer Science Dept., University of Technology, Iraq.

Abstract

The construction of large, high-resolution image is an active area of research in the fields of computer vision, image processing, and computer graphics. A Panorama is the process of combining multiple images with overlapping fields of view to produce a panorama. It is possible to produce a complete view of an area or location that cannot fit in a single shot. In this paper a high performance method for generating panoramic image is introduced. The process to generate a panoramic view can be divided into three main components: image acquisition, image registration, and merging. Geometric moment invariant produces a set of feature vectors that are invariant under shifting, scaling and rotation. The technique is widely used to extract the global features for pattern recognition due to its discrimination power and robustness. In this paper, moment invariant is used to determine the locations of merging the images to produce the panorama. The final step is adjusting the colors of the merged images. The results of experiments conducted on images taken by camera and test images loaded from the internet. The results show that the proposed algorithm is fast and efficient.

Keywords: traditional moments, image registration, panoramic image.

1- Introduction

The construction of large, high-resolution image is an active area of research in the fields of computer vision, image processing, and computer graphics. Our contribution is introducing a new direction for developing the multimedia that used now and devising new Medias more arousing and attracting for viewers. Some researchers worked on designing the panoramic image, but this work aims to develop an algorithm to generate this new type of images and developing toward producing films of the same type.

Panoramic photography has received a growing interest. Among its numerous applications, such techniques may be used to efficiently represent video sequences, in terms of compression, enhancement, visualization etc. Recent years have seen a growing interest in the mapping and visualization of the world's cities and sights. For example, systems such as Google Street View and Bing Maps enable users to browse street level imagery by presenting a panorama-based visualization of video captured at street level from a moving vehicle [1]. Constructing full view panoramas requires taking many regular photographic or video images in order to cover the whole viewing space. These images must then be aligned and composited into complete panoramic images using an image stitching algorithm.

Many advanced creation methods have been developed in recent years. [2] Proposed A method based on the scale invariance feature transform (i.e. SIFT) algorithm to stitch images captured by the turning video cameras together to form panoramic images. Based on the SIFT features and the retrofitted KDT structure, the BBF searching strategy is employed to match feature points. Then, in a post-processing pass, the author removes the mismatching feature points. Photos captured by a surveillance camera are taken as the input to test the proposed method. A very popular approach is the panoramic image mosaic concept proposed by Shum and Szeliski [3] and its variants. In these approaches a rotation matrix is associated with each input image and to reduce registration errors a local as well as a global registration step is initiated. Irani et al. [4], [5] proposed mosaic based compression. A static panorama background is first constructed out of the video sequence and then each video frame is compressed using the static panorama background as a reference. Furthermore, it detects and indexes the motion objects and provides content-based video indexing. They don't deal with on-demand transmission. Kuglin and Hines [6] proposed another method called phase correlation**,** which gives good results when the camera motion is very large. This method relies on the fact that it translation in the spatial domain corresponds to a phase shift in the frequency domain. Although the results are quite accurate, phase correlation requires a memory capacity that grows with the log of the image area, which makes it slow when handling larger images. QuickTime VR [7] is a method that does not involve image registration. It combines a set of overlapping images take

from a fixed location but at different viewing orientations to form a cylindrical panoramic image. During the rendering stage, novel views of arbitrary orientations can be synthesized from the panoramic image. This method is simple and efficient, but the user's viewpoint is fixed.
Dani and Chaudhuri [8] method is presented to deal with the case of rotational shift. Their general method consists of three stages. In the first stage, features in the image are computed. In the second stage, feature points in the reference image, often referred to as control points, are corresponded with feature points in the data image. In the last stage, a spatial mapping is determined using these matched feature points. Reassembling of one image onto the other is performed by applying the spatial mapping and interpolation. This method works well for up to 15 degrees of rotation using angles between edge points.

This work presents a method to register a group of images to generate the panoramic image by using moment invariants, which is to measure the shape characteristics. Through the use of nonlinear combination of geometric moments obtain one group of scale invariance, translation invariance and rotation invariance of moment invariants. The reminder of this paper is organized as follows. Section 2 the general system. Section 3 included the features extraction from image blocks. Section 4 explains the registration operation. Section 5 included the matching operation and compositing of the images. Section 6 and 7 explain the experimental results and conclusions.

1. The General System

As illustrated in fig. (1) the first step is capturing two or more images using normal cameras. Then the images are divided into segments. After that, the similar segments between the samples will be determined. The final step is using the determined segments in reconstruction step to generate the panoramic scene.

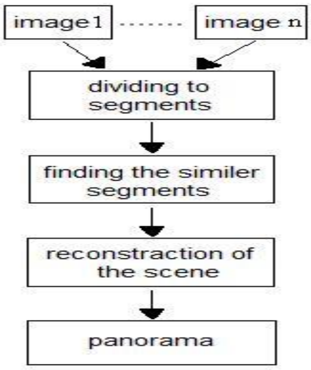

Figure (1): The General System

The first step for making the panorama is to capture the desired images which are suitable for panorama stitching. The desired images means that successive photos need to have roughly the same camera settings, enough overlap with each other and known camera parameters. A two hand held camera provides us with a 14 mega-pixel resolution is used. Since the photos are taken by the user. To satisfy the overlapping requirement, the two cameras are fixed on a slate with some distance to obtain the overlapping between the captured images as shown in fig. (2).

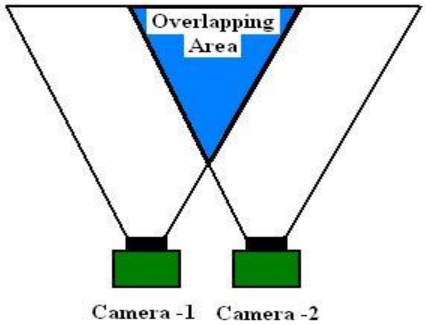

Figure (2): Image Acquisition using Two Camera

2. Image Blocks Feature Extraction Based on Traditional Moment

Moment invariants are important shape descriptors in computer vision. There are two types of shape descriptors: contour-based shape descriptors and region-based shape descriptors. Regular moment invariants are one of the most popular and widely used contour-based shape descriptors derived by Hu [9, 10]. It was derived from the theory of algebraic invariant.

A 2-D continuous function $f(x, y)$ of order $(p+ q)$ is defined as [9]:

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)dxdy \quad \ldots. (1)$$

For p, q = 0, 1, 2…

The uniqueness theorem states that if $f(x, y)$ is piecewise continuous and has non zero values only in a finite part of xy plane, moments of all order exist and the moment sequence (m_{pq}) is uniquely determined by $f(x,y)$. Conversely, (m_{pq}) uniquely determines $f(x, y)$. The central moments can be expressed as [9]:

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\bar{x})^p (y-\bar{y})^q f(x,y)dxdy \ldots(2)$$

where

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \text{and} \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

For a digital image, Equation (3.2) becomes

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y). \quad \text{..... (3)}$$

$$\mu_{00} = m_{00}\ , \quad \mu_{10} = 0\ , \quad \mu_{01} = 0$$

The normalized central moments, denoted η_{pq}, are defined as

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \quad \text{.... (4)}$$

where

$$\gamma = \frac{p+q}{2} + 1 \quad \text{... (5)}$$

A set of seven invariant moments can be derived from the second and third moments:

$$\phi_1 = \eta_{20} + \eta_{02} \quad \text{... (6)}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad \text{.... (7)}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad \text{....(8)}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad \text{...(9)}$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad \text{...(10)}$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad \text{...(11)}$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad \text{.....(12)}$$

The seven invariant moments, which are invariant to translation, scaling, mirroring and rotation, composed of the linear combination of the second-order and third-order central moments. Because seven moment invariants are relatively large, in order to facilitate comparison making use of logarithmic methods. At the same time, taking into account the possible negative moment invariants situation, you have to get absolute value before getting logarithm. The example in fig. (4) shows the moment values for the lena image with different situations. Image 1 is the original image. Image 2 is scaling with a scale factor 0.9. Iimage3 is scaling with scale factor 0.4. Image4 is scaling 0.9 and rotation with 30 degree. Image 5 is rotation with 90 degree.

Figure (3): different situation for image, from left to right (original, scaled 0.9, scaled 0.4, scaling 0.9 and rotation 30, rotation 90)

Table1: The Moment Values for Lena Images

	M1	M2	M3	M4	M5	M6	M7
Img1	2.876	8.175	11.932	10.947	23.768	15.068	22.387
Img2	2.876	8.176	11.946	10.946	23.753	15.069	22.393
Img3	2.876	8.178	11.961	10.946	23.748	15.069	22.399
Img4	2.876	8.175	11.933	10.947	23.768	15.068	22.387
Img5	2.876	8.175	11.932	10.947	23.768	15.068	22.387

3- Image Registration

Image Registration is the process of matching two or more images of the same scene. This requires estimating optimal geometric transformations to align the images with respect to a common reference. Image registration is of crucial importance in all processing and analysis tasks based on the combination of data from sets of images. Image registration algorithms can be divided into two major categories: feature-based methods and area-based methods.

Feature-based methods try to find relevant image features, known as control points, such as corners, point-like structures, line intersections, line ending points or high-curvature points that can be matched between two or more images. Once a sufficiently large number of control points have been matched by correspondence on two images, a suitable geometric transformation can be computed and applied to align them. Area-based methods, also known as featureless, correlation-based or template matching methods, work by finding correspondences between regions of the images without considering any salient features. Some of these algorithms are based on cross correlation in the spatial or frequency domain, or in maximization of mutual information.

Correlation can be estimated locally, for example, for squared regions distributed over a regular lattice, or globally for the whole image. If two images can be correlated, then the registration process continues as in feature-based methods: the parameters of a geometric transformation that maximizes cross correlation are estimated, and the images are aligned accordingly. Another important criterion for classification of image registration algorithms considers the applied geometrical transformations. According the order of complexity, these transformations can be rigid, affine, piecewise affine and non-rigid or elastic. Rigid registration models are linear and only allow for translations, rotations and uniform scale changes without any distortion. Affine models are also linear and support overall distortions represented as shears and stretches. Piecewise affine and elastic models are nonlinear and allow for arbitrary local and global distortions.

The image registration task is one of the complexes and challenging problems of image analysis, where the extreme diversity of images and working scenarios make impossible for any image registration algorithm to be suitable for all applications. Area-based algorithms are necessary in those cases where registration control points cannot be determined without high uncertainty. Two good examples are registration of planetary images and most medical images. Feature-based registration algorithms are appropriate for images with high detail contents, where enough features can be easily and accurately detected.

The basic steps of the proposed algorithm is shown in fig. (3). The basic idea is extract the features of the first image by computing the moment values for all image blocks. Select some blocks (at least three blocks not located on one straighten and located in the region of overlap) from the second image. These blocks are invariant for all the next images. The moment values are computed again for the selected blocks and then compared with the moment values of the first image to find the match. If there is a match the images are composed together and the colors of them are adjusted.

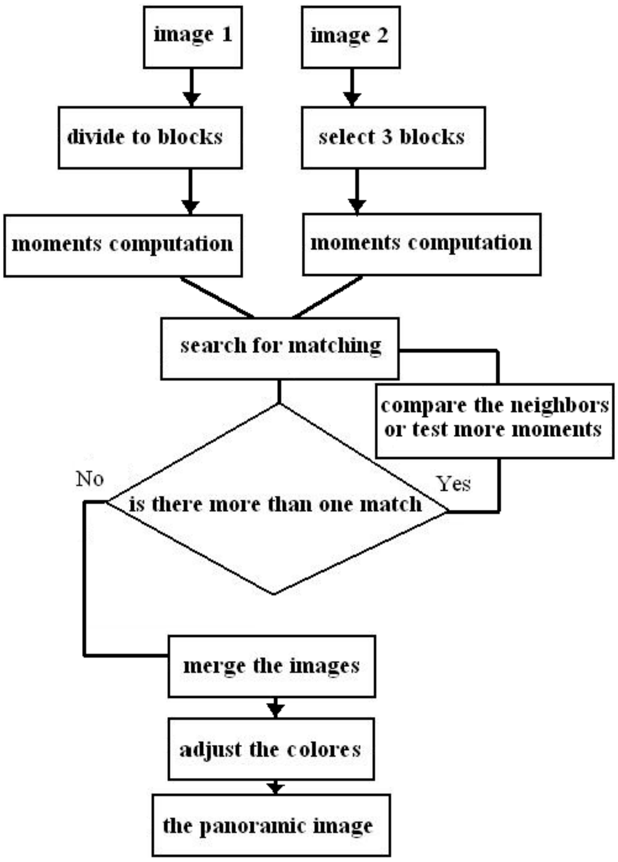

Figure (4): the proposed system

The basic steps of the proposed panoramic image generation algorithm are illustrated below:

Panoramic Image Generation Algorithm
Input: 1… N of images. **Output:** panoramic image
Step 1: take the first image and divide it to non overlapped blocks. **Step 2:** for i = 1 to no. of blocks compute the moments for each block. **Step 3:** from the next image select (at least) three points not located on one straighten and compute the moments for it. **Step 4:** compare the moments values to find the match. **Step 6:** if there is one match go to step 7, else If there is two moment values similar to the moment values of any one of the selected blocks then: - compare its neighbors or take more moments or select another blocks. - Select the block that the moment values are similar to the moment values of the block in the second image. **Step7:** merge the images. **Step8:** adjust the color difference of the merged images. **End**

4. Features Matching and Image Compositing

For each pair of potentially matching images we have a set of feature matches that are geometrically consistent. Where, the absolute difference between the features of the two images is computed to determine the matched moment values. Similar moment values may be appears and this problem can cause false compositing. Therefore, three ways are determined to find the exactly match, the first is comparing the moment values of the neighboring blocks. The second is, taking more moment values for the blocks. The third is selecting another blocks. Two spatially neighboring images are related to each other by a homography transformation.

Homography is the relationship between two images such that any given point in one image corresponds to one and only one point in the other. By recovering the homography transformation, images can be merged together and ultimately the panorama can be created. Images aligned after undergoing geometric corrections most likely require further processing to eliminate remaining distortions and discontinuities.

Alignment of images may be imperfect due to registration errors resulting from incompatible model assumptions, dynamic scenes, etc. Furthermore, in most cases images that need to be mosaiced are not exposed evenly due to changing lighting conditions, automatic controls of cameras, printing/scanning devices, etc. These unwanted effects can be alleviated during the compositing process. The main problem in image compositing is the problem of determining how the pixels in an overlapping area should be represented. Finding the best separation border between overlapping images has the potential to eliminate remaining geometric distortions.

The other problem is eliminating the discontinuities of the merged images. There are two popular ways of blending the images. One is called alpha blending, which takes weighted average of two images. The cases that alpha blending works extremely well is when image pixels are well aligned to each other and the only difference between two images is the overall intensity shift. Alpha blending will merge two images seamlessly. However, if the images are not aligned well, the disagreements will show in the blended image.

Another popular approach is Gaussian pyramid [11]. This method essentially merges the images at different frequency bands and filters them accordingly. The lower the frequency band, the more it blurs the boundary. Gaussian pyramid blurs the boundary while preserving the pixels away from the boundary. It does not work well, however, if the two images are at significantly different intensity levels. The transition is not as smooth as alpha blending for this case.

The test images used in this work is taken by a moving camera or by two camera positioned in parallel way as shown in fig. (2). According this way of acquiring images, we proposed two methods, the first is by dividing the overlapping region for the images to columns and then taking the in between columns for the two images as illustrated in fig.(5).
The second is by smoothing the overlapping region between two images eliminating the discontinuities of the merged images.

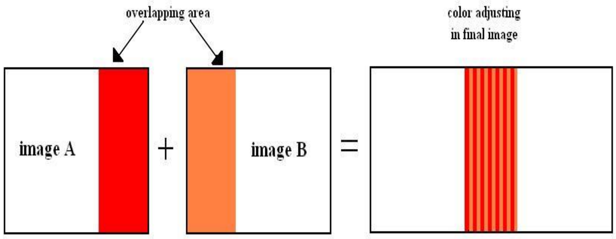

Figure (5): The Color Adjusting of the Stitched Images

5. Experimental Results

We have explained the results of our algorithm mainly on pictures taken by two horizontally normal cameras and on pictures taken from the internet. The test images are converted to grayscale images with size 512x512. Fig. (6) and fig. (7) explain the implementation of the proposed method. The extracted features are shown in fig.6 and the differences between the moments value of each selected block from the second image with the moments values of the first image blocks are shown in fig.(7). First we convert the RGB image to grayscale image and then determine the region of overlapping to apply the moment invariants. Dividing the region of overlapping to blocks and then applying the moment invariants gives us fast and better results than applying the moment invariants on the complete image. After determining the matched blocks, the exact position of them in the complete image will be determined according the following equation:

$$(x',y')=((((x*bs)+i)-m),(((y*bs)+i)-n)) \quad \dots(1)$$

Where x’ and y’ represent the exactly coordinates of the matched block in the first image, x and y represent the coordinates of the extracted block in the overlapping region, bs is the block size, i and j are the high and width of the image and m and n are the high and width of the overlapping region.

Figure (6): The Extracted Matched Blocks between Two Images in the Overlapping Region.

The difference between the moments values of the selected blocks from one image with moments values of all blocks of the other image is shown in fig.(7)

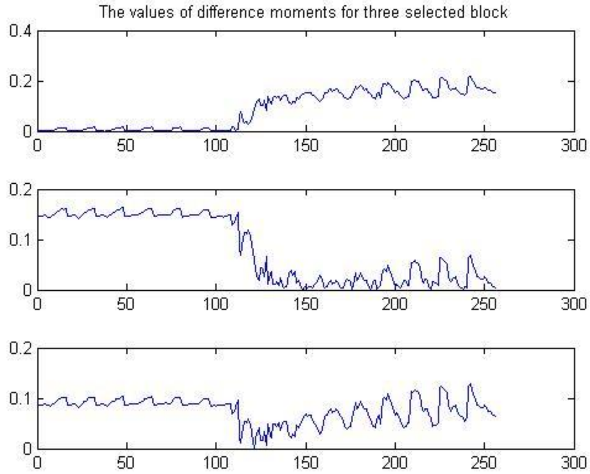

Figure (7): The Difference between the Moment Values of the Three Selected Blocks with the Moment Values of All Blocks of the other Image.

Figure (8): The Final Result of Proposed Algorithm

Another results are explained in fig. (9) below.

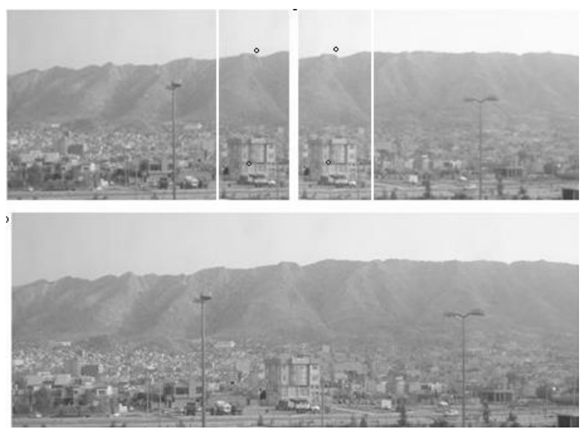

Figure (9): The Final Result of Proposed Algorithm

The proposed algorithm spends approximately 1.05 seconds for fusing 2 images. The time changes according to the number, size and the nature of images.

By comparing the execution time for the proposed method and other methods like the most famous methods SIFT and phase correlation, its found that the SIFT take about 2.45 seconds. While execution by the phase correlation method expends 1.24 seconds.

6. Conclusions

Through the use of nonlinear combination of geometric moments obtain one group of scale invariance, translation invariance and rotation invariance of moment invariants. The traditional moments method has a lower computational cost and also limited to the affine or any simpler model, where the projective moment invariants are not known to exist. Intensity in the image does not change arbitrarily, but there may be a change in overall contrast due to changes in illumination or camera parameters. The results show that the proposed algorithm is efficient and speed.

The limitation of the proposed algorithm are, it not work for higher order alignment models, such as perspective. But, we believe there is a way to extend the proposed approach further to deal with this issue.

References

[1] Zheng E., Raguram R., Fite-Georgel P., "Efficient generation of multi-perspective panoramas", IEEE, International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, 2011.

[2] Lin L., Nan G.,"Algorithm for Sequence Image Automatic Mosaic based on SIFT Feature", WASE International Conference on Information Engineering, 2010.

[3] Szeliski R., "Video mosaics for virtual environments", IEEE Computer Graphics and Applications, pages 22–30, March 1996.

[4] Irani M., Hsu S., and Anandan P., "Video compression using mosaic representations", in Signal Processing: Image Communication, vol. 7, pp. 529–552, 1995.

[5] Irani M. and Anandan P., "Video indexing based on mosaic representations", in Proceedings of the IEEE, vol. 86, no. 5, pp. 905–921, May 1998.

[6] Kuglin C.D. and Hines D. C," the phase correlation image alignment method", in conference on cybernetics and society, pp.163-165, New York, 1975.

[7] Chen E., "QuickTime VR: An image-based approach to virtual environment navigation," in Proc. ACM SIGGRAPH, pp. 29–38, 1995.

[8] Dani P. and Chadhuri S., "automated assembling of images: image montage preparation", pattern recognition, 1995.

[9] Hu M.K., "Visual pattern recognition by moments Invariants", IRE Trans. Information Theory, 8: 179- 87, 1962.

[10] Rizon M. and et al.,"Object Detection using Geometric Invariant Moment", American Journal of Applied Sciences 2 (6): 1876-1878, ISSN 1546-9239, 2006.

[11] Adelson E. and et al., "Pyramid methods in image processing," RCA Engineer, 29-6, 1984.

خوارزمية توليد الصورة البانورامية باستخدام ثوابت عزوم Hu's

كوثر عباس صلال
الجامعة المستنصرية /قسم علوم الحاسبات

أ.د.عبد المنعم صالح ابو طبيخ
الجامعة التكنولوجية /قسم علوم الحاسبات

الخلاصة :

ان بناء صورة كبيرة و ذات دقة عالية تعتبر مساحة نشطة للبحث في مجالات الرؤية ألحاسوبية ومعالجة ألصور والرسم بالكمبيوتر. البانوراما هي عملية الجمع بين عدة صور مع وجود منطقة تداخل بهدف إنتاج البانوراما. حيث من الممكن انتاج مساحة كاملة من منطقة أو موقع معين لا يمكن أن نراه في صورة واحدة. في هذا البحث تم عرض طريقة لتوليد الصورة البانورامية. تتألف عملية توليد المشهد البانورامي من ثلاثة خطوات اساسية هي: اكتساب ألصور وتسجيل الصورة والدمج. العزوم الهندسية الثابتة تقدم مجموعة من الصفات التي تكون ثابتة تحت ظروف الحركة والتدوير والتحجيم. تستخدم هذه التقنية على نطاق واسع في استخراج صفات فريدة من أجل التعرف على الأنماط نظرا لقوة التمييز والمتانة التي تتمتع بها. في هذا البحث، يتم استخدام العزوم الثابتة لتحديد مواقع دمج الصور لإنتاج البانوراما. والخطوة الأخيرة هي ضبط ألوان الصور المدمجة. نتائج التجارب أجريت على صور تم التقاطها بواسطة الكاميرا وصور تم تحميلها من شبكة الانترنت. وبينت النتائج أن الخوارزمية المقترحة سريعة وفعالة.

الكلمات الدالة :العزوم , تسجيل الصورة , الصورة البانورامية