

# Feature Descriptor Based on Normalized Corners and Moment Invariant for Panoramic Scene Generation

Kawther Abbas Sallal  
Computer Science Department  
Babylon University  
Babylon, Iraq

Abdul-Monem Saleh Rahma  
Computer Science Department  
University of Technology  
Baghdad, Iraq

**Abstract**—Panorama generation systems aim at creating a wide-view image by aligning and stitching a sequence of images. The technology is extensively used in many fields such as virtual reality, medical image analysis, and geological engineering. This research is concerned with combining multiple images with a region of overlap to produce a wide field of view by the detection of feature points for images with different camera motion in an efficient and fast way. Feature extraction and description are important and critical steps in panorama construction. This study presents techniques of corner detection, moment invariant and random sampling to locate the important features and built storing descriptors in the images under noise, transformation, lighting, little viewpoint changes, blurring and compression circumstances. Corner detection and normalization are used to extract features in the image, while the descriptors are built by moment invariant in an efficient way. Finally, the matching and motion estimation is implemented based on the random sampling method. The results of experiments conducted on images and video sequences taken by handheld camera and images taken from the internet. The results show that the proposed algorithm generates panoramic image and panoramic video of good quality in a fast and efficient way.

**Keywords**—Feature extraction; feature description; motion estimation; registration; panoramic scene

## I. INTRODUCTION

Panoramic view construction is one of the most computer vision applications that have a great attention recently. The technology of panoramic view is developed rapidly and becomes a kind of popular visual technology, because the visual panorama technology can bring people a new real visualization of the scene and interactive experience [1][2]. It aims at creating a wide view image by aligning and stitching a sequence of images that having a significant overlap. The technology is extensively used in many fields such as virtual reality, medical image analysis, mapping, visualization, and geological engineering [2].

Image registration operation is very important for panoramic view generation. Image Registration is the process of matching two or more images of the scene. This requires the estimation geometric transformations to align the images with respect to a common reference. Image registration is of great 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 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 sufficient number of 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 correlation based or template matching methods, work by finding correspondences between regions of the images without considering any features. Some of these algorithms are based on cross correlation in the spatial or frequency domain, or on mutual information. The correlation can be estimated locally, for example, for squared regions distributed over a regular lattice, or globally for the whole image. If two images are correlated, then the registration process continues by finding the parameters of a geometric transformation that maximizes cross correlation, and the images are aligned accordingly.

The image registration process is one of the most complex and challenging problems of image analysis, where the extreme diversity of images and working scenarios make impossible for any registration algorithm to be suitable for all applications.

## II. RELATED WORK

Many methods have been presented in recent years. Szeliski in [3] looks at one way to use video as a new source of high-resolution. Video is a low-resolution medium that compares poorly with computer displays and scanned imagery. It also suffers, as do all input imaging devices, from a limited field of view. He present algorithms that align frames of video and composite scenes of increasing complexity beginning with simple planar scenes and progressing to panoramic scenes and, finally, to scenes with depth variation. His approach directly minimizes the discrepancy in intensities between pairs of images after applying the recovered transformation.

In [4] the researchers used an algorithm based on Correlation. For Automatic Image Registration Applications, the features like edges are detected by using Sobel Edge Detection Algorithm. For matching the features, first Segmenting the image file in terms of different blocks and then applying the Hierarchical matching to create a pyramid of blocks. Finally, applying correlation based matching starting from the top level of the pyramid. Otherwise take a suitable

pixel block size say about 32 x 32 pixel block from right image and search for the exact location of that 32 x 32 pixel block in the left image. [5] Proposed the video serial images registration based on a feature based method algorithm. A matching method based on the improved algorithm can get better image fusion and image registration which is introduced for attaining more precise aggregate of matching points. In order to estimate the fundamental matrix which describes the whole geometry accurately and robustly, an improved SVD decomposition with weighted normalized fundamental matrix calculating method is proposed. [6] Describes a simple method for taking two videos and creating a panoramic video. Because using of hand held camera, it is difficult to obtain the case of stability. Therefore to get the optimization case, the researchers assume there is no change in acquiring circumstances of video. They assumed that the transformation between the two frames is the translation case. The translation between the two frames is estimated based on the matching of moment values for selected points in the overlapping regions. Also, they assume to use motion estimation techniques like three step search for the stitched frames to produce a compressed panoramic view. The limitation of this algorithm is, it works with the optimal case because this method is not excellent when the distorted transformations occurred or the gray values difference between the two images is found. [7] Presents an approach for the panoramic view generation. First, salient features are robustly detected from the input images by a robust algorithm called Scale Invariant Feature Transform (SIFT). SIFT features are invariant to translation, rotation, image scaling and partially invariant to viewpoint, illumination changes and image noise. These features are matched between the successive images and hence image transformation is estimated. Then, the image blending technique blends the images together to get a panoramic view without visible edge seam. [8] Proposes the feature based image fusion approach. The fusion image system includes features point detection, feature point descriptor extraction and matching. A RANSAC algorithm is applied to eliminate the number of mismatches and obtain a transformation matrix between the images. The input image is transformed with the correct mapping model for image stitching. In this paper, feature points are detected using steerable filters and Harris, and compared with traditional Harris, KLT, and FAST corner detectors.

Our contribution is introducing a new direction for developing the used multimedia and devising new Media more arousing and attracting for viewers. An efficient method for invariant feature detection is introduced based on mixing local and global feature extraction methods. Some researchers worked on designing the panoramic image, but this work aims to develop an algorithm to produce the panorama in a simple and accurate way and making it as a seed in generating films of that type. The remainder of this paper is organized as follows. Section 2 describes the general system. Section 3 includes the methodology. Section 4 explains the steps for features detection and descriptors construction. Section 5 explains the feature matching operation and motion estimation. Section 6 discusses the experimental results and section 7 explains the conclusions.

### III. METHODOLOGY

The panoramic scene is constructed by alignment of multiple images with overlapping region. The alignment operation requires detection of similar regions in images. The similarities between images are determined by an efficient feature extraction method. The first step is capturing two or more images using handheld camera. The second step is enhancing the image details by using median filter. Then, the important features in each image are extracted using a local feature extraction method. The extracted features are normalized. Normalization is important step if the illumination invariance is required. Each extracted feature needs to calculate a descriptor that determines its invariability under different circumstances. The feature descriptor is calculated using a famous region feature extraction method called moment invariant. After that, the feature descriptors for both images are matched using a distance metric. The matching result is refined using second matching step based on random sampling method and the motion between the two images is estimated accordingly. The final step is using the estimated motion in the construction of the panoramic scene. The complete steps of the proposed system are shown in fig. 1 below. A hand held 4300S Nikon camera with a 16 mega-pixel resolution is used. Since the images and video sequences are taken by the user. The captured images and video sequences must have region of overlap to satisfy the condition of generating the panoramic view.

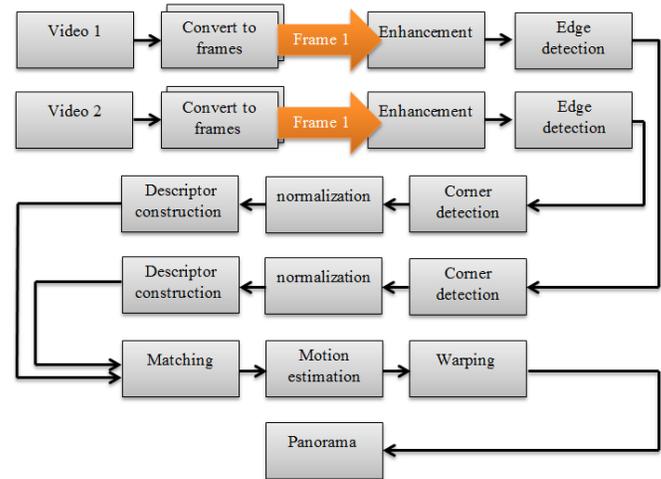


Fig. 1. The proposed system

### IV. IMAGE PREPROCESSING

It is desirable to perform some kind of noise reduction on an image before the analysis process. Because the edge detection will be used later in feature extraction process, it is important to enhance the image details first. The median filter is a nonlinear digital filtering technique, often used to remove noise and used here to enhance the edge details. Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise. The main idea of the median filter is to run through the image pixel by pixel, replacing each pixel with the median of neighboring pixels. The pattern of neighbors is called the "window", which slides, pixel by pixel,

over the entire image. Fig. 2 shows the result of applying median filter with window 3 x 3. In this work, the image is filtered with median filter first, then the filtered image is subtracted from the original image to obtain the details image. Then, the details image is summed with the filtered image to get the final filtered image.



Fig. 2. Implementation of the details enhancement method

#### V. THE IMPROVED CORNER DETECTOR

The harris autocorrelation detector [9] is a development detector of the moravec's detector. The 'corner' is a location in the image where the local autocorrelation function has a distinct peak. Corner point detection has found its application in various computer vision tasks. In this work, improved corner detector method is proposed to extract corner information as the first step of the proposed algorithm.

This method not only solved the problem of the discrete shifts, but also it deals with the issue of directions with the advantage of the autocorrelation function and increased the accuracy of localization. Feature point extract by the Harris operator has a rotation and translation invariability and has a good robustness against noise and change of parameters during acquisition of image. The improved method consists of the following steps:

1) Enhance the input image details using median filter. Then convert the RGB image to gray color image.

2) To get rid of the noise, the image is smoothed with Gaussian filter with sigma  $\sigma$ . larger  $\sigma$  increasing the smoothing.

$$G_{x,y} = \frac{1}{2\pi\sigma^2} \exp^{-(x^2+y^2)/2\sigma^2} \quad (1)$$

Where  $x$  is the distance from the origin in the horizontal axis,  $y$  is the distance from the origin in the vertical axis, and  $\sigma$  is the standard deviation of the Gaussian distribution.

3) Apply prewitt edge detection algorithm on the smoothed image by convolving two masks, horizontally and vertically to obtain the first derivatives in  $x$ -direction ( $f_x$ ) and  $y$ -direction ( $f_y$ ) as following:

$f(i-1,j-1)$	$f(i,j-1)$	$f(i+1,j-1)$
$f(i-1,j)$	$f(i,j)$	$f(i+1,j)$
$f(i-1,j+1)$	$f(i,j+1)$	$f(i+1,j+1)$

Figure (4-13) pixel neighbors

The observation pixel is  $(i,j)$ .

$$f_x = f(i+1,j-1) + f(i+1,j) + f(i+1,j+1) - f(i-1,j-1) - f(i-1,j) - f(i-1,j+1) \quad (2)$$

$$f_y = f(i+1,j+1) + f(i,j+1) + f(i-1,j+1) - f(i-1,j-1) - f(i,j-1) - f(i+1,j-1) \quad (3)$$

4) Three values must be obtained from the result above. These values are called the second order moment. The first value is the square of gradient in  $x$ -direction ( $f_x^2$ ). The second value is the square of gradient in  $y$ -direction ( $f_y^2$ ). The third value is the multiplication of gradient in  $x$ -direction and gradient in  $y$ -direction ( $f_x f_y$ ).

5) The resulted images from the previous step is smoothed again to keep as possible only the true corners.

6) Compute the corner value ( $R$ ) using the following equation:

$$f(x, y, \delta 1) = f(x, y) * G(\delta 1) \quad (4)$$

$$R = \frac{G(\delta 2) f_x^2(x, y, \delta 1) * G(\delta 2) f_y^2(x, y, \delta 1) - (G(\delta 2) f_x f_y(x, y, \delta 1))^2}{G(\delta 2) f_x^2(x, y, \delta 1) + G(\delta 2) f_y^2(x, y, \delta 1) + \beta} \quad (5)$$

Where  $\delta_1$  and  $\delta_2$  is the standard deviation of the Gaussian distribution ( $G$ ),  $f_x(x,y)$  and  $f_y(x,y)$  are the partial derivatives of  $f(x,y)$  in  $x$  and  $y$  directions, and  $\beta$  is a constant value.

7) Apply the non-maxima suppression to extract the maximum value within predefined neighbors. The non-maxima suppression is like a filter which only lets the value pass if it is the maximum of its neighbors.

$$R_i > R_j \quad \forall j \in N_i \quad (6)$$

Where  $R_i$  and  $R_j$  are the corner values of pixel  $i$  and its neighbor pixel  $j$ .  $N_i$  is the  $N$  neighbors of pixel  $i$  in a predefined window.

8) Compute the mean and standard deviation for the result  $ME = \sum_r \sum_c \frac{R(r,c)}{N}$  (7)

$$std = \sqrt{\frac{1}{N} \sum (R_{(r,c)} - ME)^2} \quad (8)$$

9) Obtain the strongest or the more stability corners by normalization of the corners using the following equation:

$$NC(r,c) = R(r,c) - ME / std \quad \geq T \quad (9)$$

Where,  $NC$  is the normalized corner,  $R$  is the corner response value,  $ME$  is the mean value,  $std$  is the standard deviation value and  $T$  is the threshold.

10) Extract the final corners if the result of the above equation is above the predefined threshold, else ignore the corner. The result is shown in fig.3.

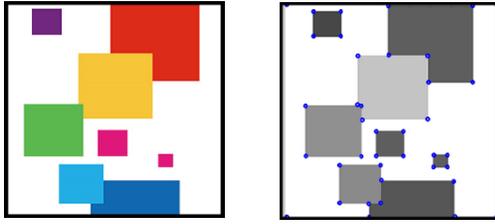


Fig. 3. corner feature extraction

## VI. DESCRIPTOR CONSTRUCTION

One of the main challenges lies in classifying and recognizing objects from different views and lighting conditions. Depictions of natural scenes typically do not maintain their viewpoint, having rotational, perspective, projective and zoom changes between images of the same object. Interest points have to focus on the same locations of an object, no matter from which point of view they are shown. When provided with stable interest points under these circumstances, local descriptors become more effective than using random features of an object. In the previous step, the important features are gotten but it is necessary to identify each feature point. Therefore, an efficient descriptor will be generated for each detected feature point. For each feature, an  $N \times N$  window centered on this feature is determined. Then, the window is divided to  $n \times n$  blocks where  $N > n$ . For each block the average of geometrical moments is calculated. The result is a descriptor of  $n \times n$  matrix for each feature point.

Moment invariants are the most popular and widely used shape descriptors in computer vision derived by Hu. A 2-D function  $f(x, y)$  of the order  $(p + q)$  is defined as [10]:

$$\mu_{pq} = \sum_x \sum_y (x)^p (y)^q f(x, y) \quad (10)$$

For  $p, q = 0, 1, 2, \dots$

The uniqueness theorem states that if  $f(x, y)$  is piecewise continuous and has non zero values only in a finite part of  $x$ - $y$  plane, moments of all orders 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 [10][11][12]:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (11)$$

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

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

The normalized central moments, denoted  $\eta_{pq}$ , are defined as

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (12)$$

$$\text{Where } \gamma = \frac{p+q}{2} + 1$$

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

$$\phi_1 = \eta_{20} + \eta_{02} \quad (13)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (14)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (15)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (16)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (17)$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (18)$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (19)$$

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 of the seven moment invariants is relatively large, and to simplify comparison, making use of logarithmic methods. At the same time, taking into account the possible negative moment invariants situation, you have to take the absolute value before getting logarithm.

## VII. MOTION ESTIMATION FROM CORRESPONDENCES

Images can be in different transformations that can be resulted during camera acquiring. According the complexity, these transformations are rigid, affine, non-rigid. Rigid registration models are linear and only allow for translation, rotation and scale changes without any distortion. Affine transform is also linear and support overall distortions besides shearing and stretching. Non-rigid models are nonlinear and allow for arbitrary local and global distortions. These transformations effect on matching operation. Therefore, it is important to find a way to determine the true matches. The true matches can be determined if the points fit with a predefined model. The matching is done by computing the Euclidian distance between two descriptors depending on the second nearest neighbor technique. The matching result contains a number of error matches which effect on the motion estimation results. Therefore, a random sampling method is used to get the invariance matches and estimating the motion between the images.

### 1) Nearest Neighbor Matching

The Nearest Neighbor algorithm uses the ratio of distance between nearest neighbor feature points to that of second nearest neighbor feature points to match feature points. Using the ratio of nearest neighbor to second nearest neighbor to match feature points can obtain a good result, because a correct registration will have a more obvious shortest distance of nearest neighbor than that of misregistration, which will achieve a stable registration. Assume (i) as the feature point in image 1, and (j) as the feature point of nearest neighbor in image 2. If the ratio of the nearest distance to second-nearest distance (a) is less than a certain threshold, then this pair of registration points is matched, as described below:

$$R = D(i, j) / D(i, a), \quad (20)$$

If  $R < T$  then (i and j) is matched

Where  $D(i, j)$  is the distance between point (i) in the first image and point (j) in the second image and (a) is the second nearest neighbor point as the following equation:

$$D(i, j) = \sqrt{(i - j)^2} \quad (21)$$

The points (i) and (j) is matched if the value of R is lower than the predefined threshold T. If T is decreased, the number of registration points will be reduced but more stable. The mismatched points can be regarded as outliers, which are the data that do not conform to the model. These outliers can rigorously disturb the estimated motion, and consequently should be identified.

## 2) Random Sampling Technique

It's important to determine a set of invariance matches, which are the data whose distribution can be explained by some set of model parameters from the presented correspondences so that the transformation can be estimated in an optimal manner. In the computer vision field, any two images of the same planar surfaces are related by a transformation. This is very important in computing the camera movement, like rotation and translation and other transformation between two images. In mathematical definition the homogeneous coordinates are used, because matrix multiplication cannot be used directly to perform the division required by the perspective projection.

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (22)$$

Where  $p' = Tp$ .

Each point correspondence generates two linear equations for the elements of T (dividing by the third component to remove the unknown scale factor).

$$x' = \frac{t_{11}x + t_{12}y + t_{13}}{t_{31}x + t_{32}y + t_{33}} \quad (23)$$

$$y' = \frac{t_{21}x + t_{22}y + t_{23}}{t_{31}x + t_{32}y + t_{33}} \quad (24)$$

And multiplying out

$$x'(t_{31}x + t_{32}y + t_{33}) = t_{11}x + t_{12}y + t_{13} \quad (25)$$

$$y'(t_{31}x + t_{32}y + t_{33}) = t_{21}x + t_{22}y + t_{23} \quad (26)$$

Then,  $n \geq 4$  points generate  $2n$  linear equations, which are sufficient to solve T. The above equation can be rearranged as:

$$\begin{pmatrix} x & y & 1 & 0 & 0 & 0 & -x'x & -x'y & -x' \\ 0 & 0 & 0 & x & y & 1 & -y'x & -y'y & -y' \end{pmatrix} T = 0 \quad (27)$$

Where,

$T = (t_{11}, t_{12}, t_{13}, t_{21}, t_{22}, t_{23}, t_{31}, t_{32}, t_{33})^T$  is the matrix T written as a vector.

The random sampling algorithm [13] is suggested in this work to be applied on the initial matches to determine the invariance matches. This algorithm was first introduced by Fischler and Bolles as a method to estimate the model's parameters in the presence of large amounts of variance matches. It has been widely used in the computer vision and image processing for many different purposes.

This algorithm includes two steps that are repeated in an iterative fashion. First, a set of points is randomly selected from the input dataset and the model parameters are computed using only the elements of this set as opposed to least squares, where the parameters are estimated using all the data available. In the second step the algorithm checks which elements of the full data set are consistent with the model instantiated with the parameters estimated in the first step. The algorithm ends when the probability of finding better set points is below a certain threshold [14]. In this work four points are randomly selected from the set of candidate matches to compute the transformation. Then, select all the pairs which agree to the transformation. A pair (p; p') is considered to agree to a T, if:

$$\text{Dist}(T, p; p') < \epsilon, \quad (28)$$

For some threshold  $\epsilon$  (represents the amount of error) and Dist is the Euclidean distance between two points. The third step is repeating the previous two steps until enough pairs are consistent with the computed transformation. The results from this step are the group of invariance matching features and the motion matrix that the second image is wrapped accordingly to generate the panoramic view.

## VIII. GAUSSIAN MASK COLOR ADJUSTING

Due to various reasons, including the light, the geometry of the camera and other reasons, the overlapping regions of frames are almost never the same. When mosaicking the two frames, the resulted frame contains a distinctive seam. A seam is the artificial edge produced by the intensity differences of pixels immediately next to where the images are joined. Therefore, to avoid the intensity disparity on the mosaicking line, a color adjusting step must be done. To adjust the color of the resulted frames, the frames are combined with a background (using Gaussian function) to create the appearance of partial transparency. It is often useful to render image elements in separate passes, and then combine the resulting multiple frames into a single frame, final frame in this process called the composite frame.

The first step in color adjusting process is determining the suitable distance in image for blending and this depends on the computed motion. Then, the distance value is used to build two masks for the images. The masks are filtered with Gaussian function to introduce the transparent background. Then, each image will be multiplied with its mask to produce the final color adjusted image.

## IX. EXPERIMENTAL RESULTS AND DISCUSSION

To evaluate the performance of the proposed algorithm, the experiments include images and video sequences taken by handheld camera in different situations. For each image, after details enhancement with median filter, the initial corners and the normalized corners are extracted depending on its mean and standard deviation and shown in home image in fig. 4. The value of the used threshold must be sufficient to extract enough number of corners. After extracting the final corner points in the two images, the descriptors for them are created by taking a 25\*25 window around each point, divide the window to blocks of 5\*5 each and finding the average value of moment invariant for them.

Then, the matching operation is done by finding the nearest neighbor. A very important factor here is the matching threshold because it must be chosen to get as more matching points as possible. The result of this step illustrated in fig. 5. The next step is eliminating the mismatch points to obtain only the invariance matches that represent the key feature points by using a random sampling algorithm.

The number of iterations is an important factor here to obtain more matches between the images. The result of this step is shown in fig. 6. Finally, the panoramic scene is constructed according to the estimated motion from the previous step. The projected image is shown in fig. 7.

The constructed panorama before and after color adjusting is explained in fig. 8. Second test of the proposed method are shown in fig. 9. Fig.10 explains an example on panoramic video generation using the proposed method. For example, five frames are taken from each video.



Fig. 4. The detected and the normalized corners

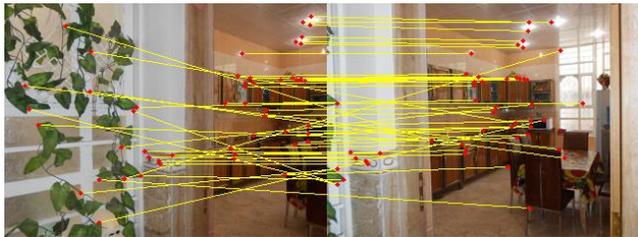


Fig. 5. The initial matches



Fig. 6. The invariance matches

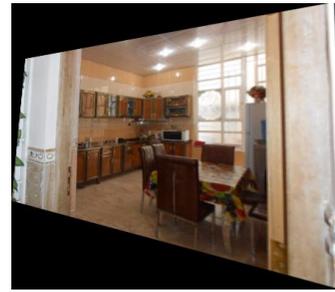


Fig. 7. Transformed image according the estimated T

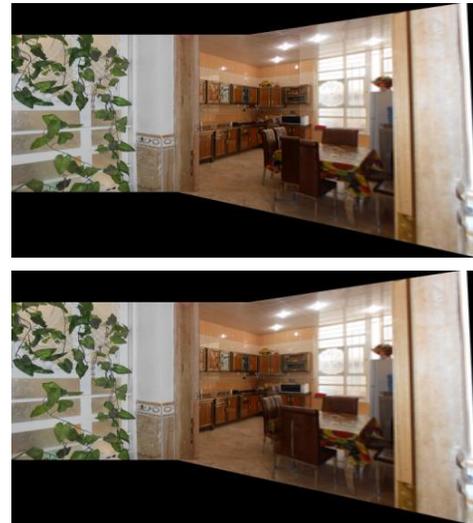


Fig. 8. Panoramic image before and after color adjusting



Fig. 9. Panoramic image from ice age movie

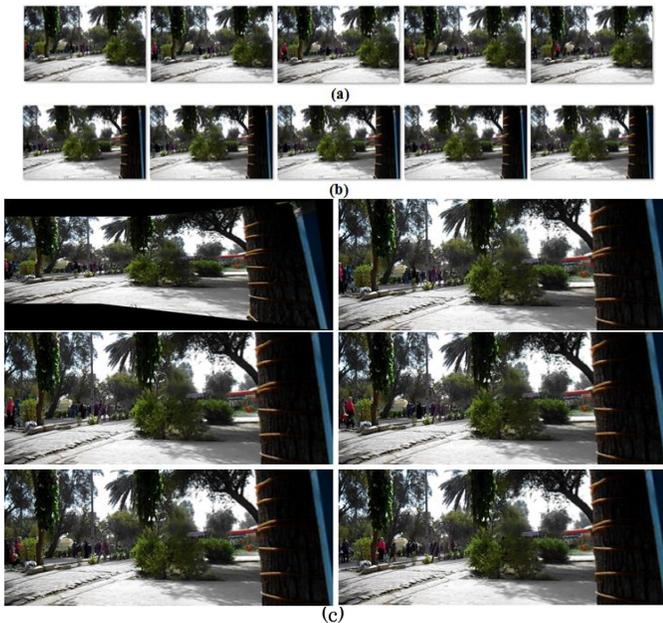


Fig. 10. Panoramic video for Al\_Zawraa city. frames(1,9,19,26,100) in: (a) video 1 and (b) video 2 and (c) panoramic frames.

As shown in the figures and in experiments, the accuracy of the proposed algorithm is coming from the using of the random sampling algorithm because it follows the object's motion to estimate the matched feature points under different circumstances like transformations, lighten, noise, view point. For, panoramic video generation, the proposed method was applied on the first frame only. The other frames will be accompanied according to the estimated motion directly for fast execution.

## X. CONCLUSIONS

Panoramic scene generation is an important topic because it needs a fusion of image processing, graphics and computer vision techniques. Many researchers deal with generating panoramic image depending on the SIFT method but in this work, a panoramic image using a proposed feature extraction method is generated. The extracted features and constructed descriptors in the overlapping region of the images are based on corner points, geometrical moments and random sampling for feature matching and motion estimation from initial matches. The mixing of local feature extraction method represented by improved corner detector and region feature extraction method represented by geometrical moments and random sampling for feature filtering and motion estimation

gives us a fast, efficient, accurate method and less complexity than the famous SIFT method. Also, the proposed method is applied on two video sequences and gives very good result according to many factors like in time and in quality of the resulted panoramic video. The execution time changed due to the number of extracted features and the number of iterations used in random sampling method. For images the execution time between 6-35 sec. while for videos it is between 380-260 sec. the quality of image before and after color adjusting is measured by PSNR and for all images and video sequence samples between 33-43dbi.

## REFERENCES

- [1] E. Zheng, R. Raguram, and P. F. George, "Efficient generation of multi-perspective panoramas", IEEE, International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, 2011.
- [2] Yun Liu, Jiru Zhang, Jian Gao, "Panoramic Technique in the Video Monitoring System and Implementation", TELKOMNIKA, Vol.11, No.1, pp. 91-96, ISSN: 2087-278X, 2013.
- [3] R. Szeliski, "Video mosaics for virtual environments", IEEE Computer Graphics and Applications, 1996, pp: 22-30.
- [4] R. Asha, K. Asha and M. Manjunath, "Image Mosaicing and Registration", International Journal of Computer Science Issues, Vol. 10, Issue 2, No 1, March 2013.
- [5] W. Bing and W. Xiaoli, "Video Serial Images Registration Based on FBM Algorithm", Research Journal of Applied Sciences, Engineering and Technology 5(17), 2013.
- [6] A. Kawther and S. Abdul Monem, "Generation of Video Panorama System", International Journal of Computer Applications, USA vol.73 issue 5, pp: 20-26, July 2013.
- [7] L. Lin and G. Nan, "Algorithm for Sequence Image Automatic Mosaic based on SIFT Feature", WASE International Conference on Information Engineering, 2010.
- [8] M. V. Subramanyam and Mahesh, "Feature Based Image Mosaic Using Steerable Filters and Harris Corner Detector", I.J. Image, Graphics and Signal Processing, 6, 2013, pp:9-15.
- [9] C. Harris and M. Stephens, "A combined corner and edge Detector", In Alvey Vision Conference. 1988, Pp. 147-151.
- [10] M.K. Hu, "Visual pattern recognition by moments Invariants", IRE Trans. Information Theory, Vol 8. 1962, Pp.: 179- 87.
- [11] R. Mohamed, Y. Haniza, S. Puteh, Y. Ali, S. Abdul R., R. Mohd, Y. Sazali, D. Hazri and M. Karthigayan, "Object Detection using Geometric Invariant Moment", American Journal of Applied Sciences. Vol 3 no.6. 2006, Pp.: 1876-1878.
- [12] J. Bei and L.Chen, "Map Matching Algorithms Based on Invariant Moments", Journal of Computational Information Systems Vol. 7 (16). 2011, Pp.: 5668-5673.
- [13] 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, vol. 24. No. 6. 1981, Pp.: 381-395.
- [14] M. Zuliani, "Computational Methods for Automatic Image Registration", PHD thesis. University of California, USA, 2006.