مجلة

كلية التــــراث الجامعة

مجلة علمية محكمة

متعددة التخصصات نصف سنوية

العدد الثاني والثلاثون

12 كانون الثاني 2022

ISSN 2074-5621

رقم الايداع في دار الكتب والوثائق 719 لسنة 2011

مجلة كلية التراث الجامعة معترف بها من قبل وزارة التعليم العالي والبحث العلمي بكتابها المرقم (ب 3059/4) والمؤرخ في (4/7 /2014)

A review of Features Detection Methods for recognition the location of a specific frame from Video Stream

مراجعة لطرق اكتشاف الصفات للتعرف على موقع إطار معين في السلسه الفيديويه

Alaa Noori Mazher **Dr. Ekhlas Falih Naser** **Enas Tariq Kuthear**

Abstract

The recognition of an image is vastly employed in various application zones like recognition of a shape, recognition of eye and recognition of a gesture. Among present algorithms for this objective, vision_based and statistical algorithms are further efficient solutions for being precise, versatile and inexpensive. Vision_based algorithms can mostly be classified as feature_based algorithms and appearance_based algorithms. In this paper, we can be surveyed algorithms for recognition the location of a specific frame from video stream based on both vision_based and statistical algorithms. To compare the performances of the algorithms, we conducted a series of experiments on four types of algorithms such as Harris Corner Detector, moment's invariants, Fast Retina Key-points (FREAK) Detector and histogram matching. The comparison process relied on two important measures: time it takes to locate the image in the video stream and the accuracy. Experimental results on a sample of videos of different sizes showed that the time taken to detect the location of the image within the video stream is high when using the FREAK detection algorithm, while the time is very little when using the moment's invariants algorithm. As for the detection accuracy, it is high when using the FREAK detection algorithm, while the very little time when using the moment's invariants algorithm.

Keywords: Invariant Moments, Harris detector, Image Recognition, FREAK, Histogram

الخلاصه

يتم استخدام التعرف على الصورة بشكل كبير في مناطق التطبيق المختلفة مثل تمييز الشكل تمييز العين تمييز على البصمة . من بين الخوارزميات الحالية لهذا الهدف ، تعد الخوارزميات القائمة على الرؤية والخوارزميات الإحصائية حلولًا فعالة أخرى لكونها دقيقة ومتعددة الاستخدامات وغير مكلفة. يمكن تصنيف الخوارزميات القائمة على الرؤية في الغالب على أنها خوارزميات تستند إلى السمات وخوارزميات تستند إلى المظهر. في هذا البحث ، يمكننا مراجعة الخوارزميات لتمييز موقع إطار معين في السلسه الفيديويه بناءً على كل من الخوارزميات القائمة على الرؤية والخوارزميات الإحصائية. لمقارنة أداء الخوارزميات ، أجرينا سلسلة من التجارب على أربعة أنواع من الخوارزميات مثل كاشف الزاويه هاريس ، العزوم الثابته ، كاشف نقاط شبكية العين السريع والمدرج التكراري. اعتمدت عملية المقارنة على مقياسين مهمين: الوقت المستغرق لتحديد موقع الصورة في تدفق الفيديو

والدقة. أظهرت النتائج التجريبية على عينة من مقاطع الفيديو ذات الأحجام المختلفة أن الوقت المستغرق لاكتشاف موقع الصورة داخل السلسه الفيديويه يكون عالي عند استخدام خوارزمية اكتشاف نقاط شبكية العين السريعه ، في حين أن الوقت قليل جدًا عند استخدام خوارزمية العزوم الثابته . بالنسبة لدقة الكشف ، فهي عالية عند استخدام خوارزمية خوارزمية اكتشاف نقاط شبكية العين السريعه ، في حين أن الوقت الضئيل للغاية عند استخدام خوارزمية العزوم الثابته.

1. Introduction

The detection of Features is a substantial early issue of vision. Preceding works for the detection of feature include the employ of grey-level statistics and the revelation of corners and edges. Methods which are established on detecting corners and edges are mostly beneficial in applications like facilities of airport, images for matching map, and analysis of aerial images of urban scenes and so on.

Algorithms employed the statistics of grey-level are viable to a wider diversity of images like vegetation and desert scenes that may or may not hold any structures of man-made [1]. Features can be defined as locations in the image which are interesting perceptually. It can describe an algorithm for detection of feature within an image via 2 attributes (1) Robustness and (2) Generality. Given that the salient features' nature of vary from application to another application, it can be desired to select general an algorithm. In the situation of structured objects like features could be locations and corners within significant changes of curvature [2].Corners within images appears substantial information to describe the features of an object, which play irreplaceable and crucial part in processing of an image and computer vision. Numerous computer vision areas depend on the detection of corners successfully, including matching of stereo, 3D reconstruction, tracking of an object and recognition of an object etc. [3]. However, they yet haven't a strict mathematically definition for corner. The corners can be defined as the points with depressed self-similarity or position where the intensity's variations are high in the whole directions [4].Recognition of Pattern deals with understanding objects which are depend on patterns, the objects might be sounds, images and texts. Recognition of the images has numerous applications and methods in real life that deal with the Recognition of an image. Any person can travel to other country requires scan the eye to fetch print of an eye for ensuring that person is pliable for leaving legally that country. Therefore, there are numerous implementations which are looking substantial in our real life within numerous domains including learning, security matters and business [**5**]. Another research employed PCA to obtain the recognition efficiently and to recognize (21) invariant moments within color images.

2. Survey of Feature Detection

There are Nemours methodologies for recognition of a specific frame. This paper employed four major algorithms. These algorithms can be illustrated as:-

2.1 Harris Corner Detection algorithm

The second algorithm can be used in this paper was Harris algorithm for corner detection. This algorithm can be recognized by calculating the gradient of each pixel's. The pixel can be taken as a corner if the absolute gradient values in two directions are both mighty. Harris algorithm for detection the corner can be calculated as illustrated in Eq.1 [6]:

$$\boldsymbol{Reult} = \mathbf{detm}(\boldsymbol{G}) - \boldsymbol{ktr}^2\,(\boldsymbol{G}) \qquad \dots 1$$

$$\boldsymbol{G}(\boldsymbol{a},\boldsymbol{b}) = \begin{bmatrix} \boldsymbol{Img}_{\boldsymbol{m}}^{2}\,(\boldsymbol{a},\boldsymbol{b}) & \boldsymbol{Img}_{\boldsymbol{mn}}(\boldsymbol{a},\boldsymbol{b}) \\ \boldsymbol{Img}_{\boldsymbol{mn}}(\boldsymbol{a},\boldsymbol{b}) & \boldsymbol{Img}_{\boldsymbol{m}}^{2}\,(\boldsymbol{a},\boldsymbol{b}) \end{bmatrix} \qquad \dots 2$$

$$\boldsymbol{Img}_{\boldsymbol{m}}^{2}\,(\boldsymbol{a},\boldsymbol{b}) = \boldsymbol{A}^2 \otimes \boldsymbol{GAUS}(\boldsymbol{a},\boldsymbol{b}),$$

$$\boldsymbol{Img}_{\boldsymbol{n}}^{2}\,(\boldsymbol{a},\boldsymbol{b}) = \boldsymbol{B}^2 \otimes \boldsymbol{GAUS}(\boldsymbol{a},\boldsymbol{b}),$$

$$\boldsymbol{Img}_{\boldsymbol{mn}}(\boldsymbol{a},\boldsymbol{b}) = \boldsymbol{AB} \otimes \boldsymbol{GAUS}(\boldsymbol{a},\boldsymbol{b}),$$

$$\boldsymbol{GAUS}(\boldsymbol{a},\boldsymbol{b}) = \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\pi}}\boldsymbol{e}^{\frac{-\boldsymbol{a}^2+\boldsymbol{b}^2}{\mathbf{2}}} \qquad \dots 3$$

Where $Img_m(a,b)$ and $Img_n(a,b)$ are gray values partial derivatives in coordinate m and n at pixel (a,b), and $Img_{mn}(a,b)$ represent the partial derivative jumbled of second-order; k represent practical value; $GAUS(a,b)$ refer to a Gaussian function; A and B are the directional differentials at first-order, which can compute by convolving gray values and operators of difference in coordinate m and n. A function of Gaussian that used to minimize the noise affect due to first-order of directional differentials which are critical to noise. Pick the pixel's point as a corner if R exceeds the values of threshold [6].

- **2.2 Moments Invariant**

Moment invariants have been widely applied to image pattern recognition in a variety of applications due to its invariant features on image translation, scaling and rotation. The moments are strictly invariant for the continuous function. However, in practical applications images are discrete. Consequently, the moment invariants may change over image geometric transformation [7]. Suppose (F (x, y)) defined a 2-Dim image in a spatial model. A Geometric moment of arrangement *(p + q)* is illustrated in Eq.4.

$$m_{p,q} = \sum_{x=0}^{N-1}\sum_{y=0}^{M-1} x^p y^q * pixel(x,y) \qquad (4)$$

where N represent the width of an image, M represent the height of an image, for(p , q = 0 , 1 , 2 , ...).The central of moments are defined via Eq.5 [7]

$$x_c = m_{1,0} / m_{0,0}$$
$$y_c = m_{0,1} / m_{0,0} \qquad (5)$$

Where xc and yc offers at Eq. 5 defined the region's center of an object. Hence the *Central moments,* of order up to 3.The normalized moments' central, indicate $\eta_{p,q}$ as illustrated in Eq.6.

$$\eta_{p,q} = \mu_{p,q} / \mu^{\gamma}{}_{0,0} \qquad (6)$$

Where

$$\gamma = p + q/2 \qquad (7)$$

For Eq.8, (p + q = 2 , 3 ,…,p*q) . A group of seven *transformations moments invariant* constructed vial 2nd order and 3rd order moments via Eq.8 [8].

$$\left.\begin{aligned}
&\phi 1 = \eta_{2,0} + \eta_{0,2} \\
&\phi 2 = (\eta_{2,0} + \eta_{0,2})^2 + 4\eta_{1,1} \\
&\phi 3 = (\eta_{3,0} - 3\eta_{1,2})^2 + (3\eta_{2,1} - \eta_{0,3})^2 \\
&\phi 4 = (\eta_{3,0} + 3\eta_{1,2})^2 + (3\eta_{2,1} + \eta_{0,3})^2 \\
&\phi 5 = (\eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + 3\eta_{1,2})[(\eta_{3,0} + 3\eta_{1,2})^2 \\
&- 3(\eta_{2,1} + \eta_{0,3})^2] + (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3}) \\
&[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2] \\
&\phi 6 = (\eta_{2,0} + \eta_{0,2})[(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} - \eta_{0,3})^2] \\
&+ 4\eta_{1,1}(\eta_{3,0} + \eta_{1,2})(\eta_{2,1} - \eta_{0,3}) \\
&\phi 7 = (3\eta_{2,1} - \eta_{0,3})(\eta_{3,0} + \eta_{1,2})[(\eta_{3,0} + \eta_{1,2})^2 \\
&- 3(\eta_{2,1} + \eta_{0,3})^2] + (3\eta_{1,2} - \eta_{3,0})(\eta_{2,1} + \eta_{0,3}) \\
&[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} - \eta_{0,3})^2]
\end{aligned}\right\} \qquad (8)$$

This group of normalized moments central is constant to rotation, translation and scale alterations for an image [8].

2.3 **Features Description uses Fast Retina FREAK**

FREAK is a dual descriptor which calculated relies with brightness rapprochement produced trials for the quantum of sampling positions through the key_point [9]. FREAK depended on the following stages: -

A. Pattern of Sampling

Pattern of sampling adopted by a descriptor of FREAK biologically inspired within a manner of eye's retinal.The points of the sample constructs the foundation to compute descriptor of FREAK that coordinated in a sampling style like mentioned within Fig.1. N-points sample present about given key_point smoothed a kernel of Gaussian before descriptor's calculating. A kernel's bulk is assorted within respect of the position of sampling points to simulate a behavior of a human's retina in an identity to a human visual system [9]. Sampling points of FREAK descriptor illustrated by means of Eq. (9).

$$P_i = P(x_i, y_i) = L_{ri}(x_i, y_i) \qquad (9)$$

Where

$$L_{ri}(x, y) = I(x, y) * G_{ri}(x, y, \sigma_{ri}) \qquad (10)$$

In Eq. (4) [10], I(x, y) the input pixels of an image, G_{ri} (x, y,σ_{ri}) explains the kernel of the Gaussian in the i-th receptive domain from i = 1 to N and L_{ri}(x, y) explains the input image's version which can be smoothed. The i-th point's sampling P_i conformable to the i-th receptive domain center (ri) and can be known within the predefined coordinates (x_i, y_i) of the model's sampling where i = 1 to N [10].

B. Building of a Descriptor

Constructing of FREAK's descriptor depend on the rapprochement intensity amongst diverse pairs of smoothed sampling points like centers within receptive fields defineded via the following. Suppose a pair of sampling points Pa = (Pi, Pj), where i # j, i and j ∈ {1 to N} as illustrated in Fig. 1. The approach of FREAK can define a binary encoded comparison of intensity, s(Pa) on the pair as illustrated in Eq.(11) [10].

$$S(P_a) = \begin{cases} 1, & \text{if } P_i > P_j \\ 0, & otherwise \end{cases} \qquad (11)$$

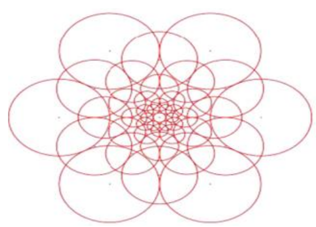


Figure 1: FREAK sampling pattern [10].

The offered comparison can form the basis to build the descriptor of FREAK F as illustrated in Eq. (12) [10].

$$F = \sum_{0 \le a < N} 2^a \quad S(P_a) \qquad (12)$$

The FREAK sampling manner can enable numerous pair wise comparisons which are leading to much great descriptor. Due to numerous pairs that may not be useful to describe image content, the writers implement an algorithm for training within the manner of sampling that illustrated within Fig. 3 to determine beneficial pairs to build a descriptor. The FREAK trained shape defends 512**.** Sampling pairs that require to examine for bit-string computing could be displayed within Eq. (12) [11].Fig.2 shows the elected 512 pairs which are distributed into (4) clusters, every cluster consist of (128) pairs. Since pattern orientation over the global gradient, symmetric manner can be taken in these clusters.

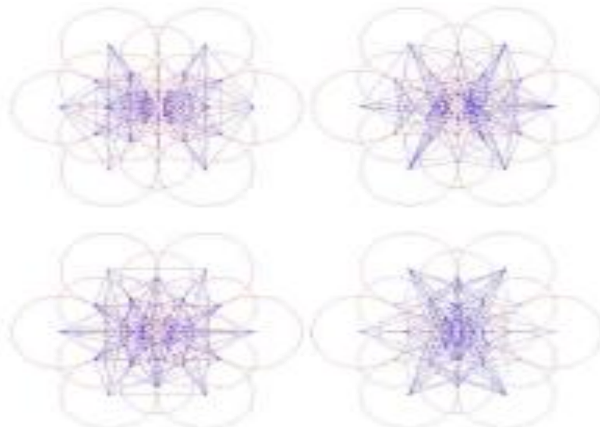


Figure 2: Explain 4-binary experiments clusters to form FREAK [10].

C. Normalization of Orientation

An orientation of FREAK could be predestined by employing (45) elected sampling-pairs which were symmetrically arranged within consideration to the center of sampling pattern like Fig.3.

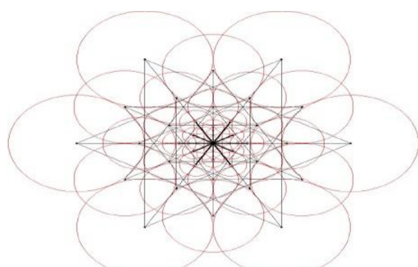


Figure 3: Pairs can be selected to calculate the key-point orientation [10].

2.4 Histogram Overview

A histogram is a kind of graph that has broad implementations in statistics. Histograms supply a visual translation of numerical data by referencing the number of data points that Located in values' domains. These domains of values are called bins or classes. The data frequency that incidence in every class is drawn by the employ of a bar. The upper that bar is the maximal data values of the frequency in that bin [11].

2.4.1 Histogram Clustering Descriptor

A cluster histogram can be employed to show each cluster's features in this part. The x-axis of histogram shows the norm values into which the convenient measurements. These measurements are showing the clusters for local features of images which are aid to abstract great sets of features data. Individual points of data are not showed and every cluster show how many features are existed at each cluster [11]. Y-Axis: Y-axis is the norm that displays you the number of times or frequency for the values of local features within each cluster.

2.4.2 Histogram Similarity

Matching speed of feature is performed by a unique step of indexing depend on the value of the Manhattan of clustering values. Compute a distance that would be traveled through a Manhattan distance function to obtain with one point of data to another if a grid-like track is followed. The distance of Manhattan among two components is the summation of differences of their corresponding items [12].The distance's formula among a point *X*=(*X1*, *X2*, etc.) with a point *Y*=(*Y1*, *Y2*, etc.) can be computed using Eq.13:

$$d = \sum_{i=0}^{n} |x_i - y_i| \quad \text{... (13)}$$

3. The Performance of Proposed Methodology

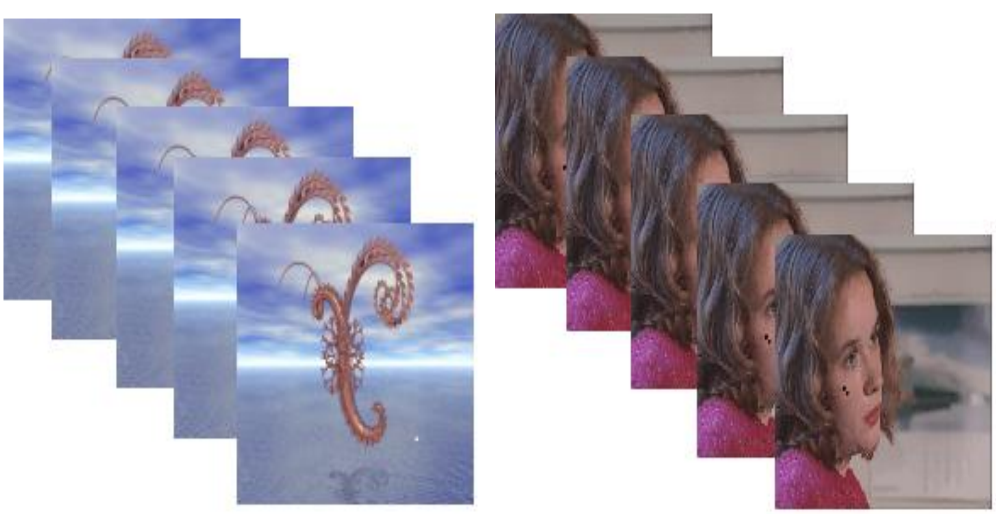
The performance of the recognition system is measured by employing the measurement of precision. The precision gauges the system's capability for recovering only the pertinent images [13]. Eq. (14) [13] is applied to calculate the accuracy of the recognition adequacy:

$$P = \frac{P_T}{P_T + P_F} \qquad (14)$$

Where P_T explains how many images are correctly acquired from the data-sets of an image while P_F explains how many images are incorrectly acquired from the datasets of an image.

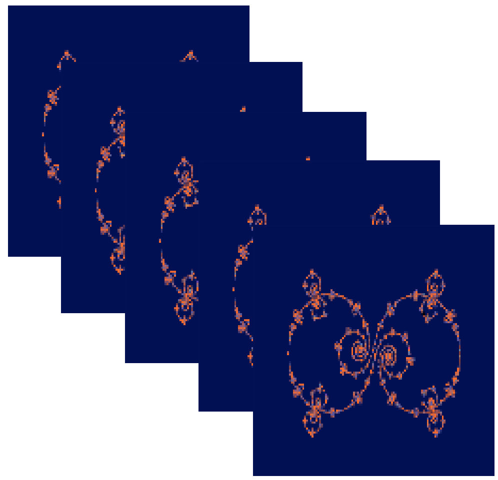
3. Empirical Results

The empirical results of methodologies were discussed and displayed in this section. The methodologies were implemented using C# programming language. Twenty kinds of data bases are utilized to evaluate the accuracy of the methodologies. The images within a data base are colored with size of 200 × 200 pixels. The results of the comparisons showing the time taken to detect the location of the image in the video sequence can be shown in Table 1 for a sample of the videos, while Table 2 shows the accuracy of the image detection.Fig.4 shows the frames of the video samples that were used in this paper.



a) b)

c)



d)

Fig.4: samples of videos frames (a) Air horse video ,b)Girl Video c)Bicycle Video ,d)Tracery Video

Table1: Computational time in seconds' when using Harris Corner Detection, Moments Invariant, FREAK and Histogram methods on four video frames for recognition the location of a specific frame from Video Stream.

	Time consuming			
Method name	**Air horse**	**Girl**	**Bicycle**	**Tracery**
Harris Corner Detector	11.48	2.94	1.73	7.44
Moments Invariants	2.17	1.01	0.38	1.50
FREAK Detector	21.20	5.77	3.19	19.47
Histogram Matching	5.86	2.01	1.58	3.48

Table2: The accuracy of image detection when using Harris Corner Detection, Moments Invariant, FREAK and Histogram methods on four video frames for recognition the location of a specific frame from Video Stream.

	Accuracy			
Method name	**Air horse**	**Girl**	**Bicycle**	**Tracery**
Harris Corner Detector	92%	90%	95%	93%

Moments Invariants	88%	85%	91%	90%
FREAK Detector	98%	97%	98%	98%
Histogram Matching	90%	89%	94%	96%

4. Conclusions

This paper introduced a review of the features detection methods for recognition the location of a specific frame from Video Stream. The methods included two types, the first type employed vision based methods such as FREAK and Harris detectors and the second type employed statistical methods such as histogram and invariants moments. The comparison process relied on two important measures: time it takes to locate the image in the video stream and the accuracy. Experimental results on a sample of videos of different sizes showed that the time taken to detect the location of the image within the video stream is high when using the FREAK detection algorithm, while the time is very little when using the moment's invariants algorithm. As for the detection accuracy, it is high when using the FREAK detection algorithm, while the very little time when using the moment's invariants algorithm. As shown in Table 1, the time taken for recognition the location of a frame from tracery video stream was 1.50 when using the Moments Invariants, while the time taken is 19.47 when using the FREAK. As shown in Table 2, the accuracy taken for recognition the location of a frame from also tracery Video Stream was 90% when using the Moments Invariants, while the time taken is 98% when using the FREAK.

References

1. G. Pooja, P. Achala and C. Umesh, "Comparison of Different Feature Detection Techniques for Image Mosaicing ", National Institute of Technology Rourkela, Odisha, India. ACCENTS Transactions on Image Processing and Computer Vision ISSN (online) 2455-4707. Vol.1, Issue1, pp.1-7, November, 2015.
2. K. Wamidh, K. Shaker,M. Zahoor ,M. Aydam and H. Bahaa," Feature Extraction Methods: A Review", Journal of Physics: Conference Series, IOP Publishing, 1591012028,2020,pp.1-10,doi:10.1088/1742-6596/1591/1/012028.
3. D. Ambar, A. Kar, and B.N. Chatleri, “A new approach to corner matching from image sequence using fuzzy similarity index,” Pattern Recognition Letters. Vol. 32, Issue 5,pp.712-720, 2017.
4. W. Junqing and Z. Weichuan," A Survey of Corner Detection Methods ", 2nd International Conference on Electrical Engineering and Automation (ICEEA), Advances in Engineering Research, Vol. 139,pp. 214-219 ,2018.
5. A. Mohammad Arafah and M. Qusay," Efficient Image Recognition Technique Using Invariant Moments and Principle Component Analysis", Journal of Data Analysis and Information Processing, Vol. 5,pp.1-10, 2017.
6. A.A. Karim and E. F. Nasser, "Improvement of Corner Detection Algorithms (Harris, FAST and SUSAN) Based on Reduction of Features Space and Complexity Time," Engineering & Technology Journal, Vol. 35, No. 2,Part B., pp.112-118, 2017.

7. F. Ekhlas and T.Enas," Retrieve an Image and its location from Video stream based on Hybrid Features Extraction methods ", Al-Turath University College Journal,Vol.29,pp.77-93, 2021.
8. F. Ekhlas," Human Face Recognition is Dependent on Computing the Similarity and Difference of the Seven Moments Values as a Face Features", Eng.&Tech. Journal ,Vol.30, No.11, pp.1843-1860, 2012.
9. F. Ekhlas, " Word Retrieval based on FREAK Descriptor to Identify the Image of the English Letter that Corresponds to the First Letter of the Word", Engineering and Technology Journal ,Vol. 38, Part B , No. 03, Pages 150-160, 2020.
10. F. Ekhlas and N. Alaa," Using FREAK descriptor to classify plasma influence in Mice sperm", Karbala International Journal of Modern Science, Vol.6, Issue1, pp.36-43,2020.
11. A. Abdulameer and E. Ekhlas ,"Image Retrieval from Video Streams Databases using Similarity of Clustering Histogram", Al-Mansour Journal, No.29 , pp.1-22,2018.
12. F. Ekhlas, "Compare between Histogram Similarity and Histogram Differencing For More Brief Key Frames Extraction from Video Stream", Journal of Physics: Conference Series. FISCAS 2021, IOP Publishing, 1897 (2021) 012022, pp.1-9,2021. doi:10.1088/1742-6596/1897/1/012022.
13. M. Alkhawlani, M.Elmogy, and H. Elbakry, "Content-based Image Retrieval using local Features descriptors and Bag-of-Visual Words," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 9, pp.212-219, 2015.