Applying Genetic Algorithm for Optimizing the User Query in Image and Video Retrieval

Ehsan Lotfi
Department of Computer Engineering, Torbat-e-Jam Branch, Islamic Azad University, Torbat-e-Jam, Iran
Email: esilotf@gmail.com

Received August 2011 Revised July 2012 Accepted April12013

ABSTRACT:
In an information retrieval system, the query can be made by user sketch. The new method presented here, optimizes the user sketch and applies the optimized query to retrieval the information. This optimization may be used in Content-Based Image Retrieval (CBIR) and Content-Based Video Retrieval (CBVR) which is based on trajectory extraction. To optimize the retrieval process, one stage of retrieval is performed by the user sketch. The retrieval criterion is based on the proposed distance metric from the user query. Retrieved answers are considered as the primary population for evolutionary optimization. The optimized query may be achieved through reproducing and minimizing the proposed measurement by using Genetic algorithm (GA). The optimized query could then be used for the retrieval of concepts from a given Data Base (DB). The proposed algorithms are evaluated for trajectory retrieval from urban traffic surveillance video and image retrieval from a DB. Practical implementations have demonstrated the high efficiency of this system in trajectory retrieval and image indexing.

KEYWORDS: Trajectory retrieval, image retrieval, sketch query, optimization, genetic algorithm.

1. INTRODUCTION

In multimedia applications, the videos and the images include various meaningful messages conveyed to its audience. For example, in an image database, the contents such as flowers, buildings, cars, and the very concepts, are conveyed when fans watch the images. And for another example, by watching a traffic surveillance video, the audience understands concepts such as turning left, going the wrong way, and eventually the observance and infractions of traffic laws. Actually, static objects in an image, activities and events in a video can make the semantics. The fact that a human understands the concepts of an image or a video while a machine cannot is normally referred to as the semantic gap between man and machine [2]. If computers were capable of extracting the meanings contained in a video or image, such meanings could then be used for automatic indexing and the retrieval of the concepts from a given database. In this paper, our aim is to propose a novel algorithm extracting the semantics from image and video datasets and because of the dramatic increasing of multimedia data volumes in recent years, this system can be useful in retrieval process and the indexing and retrieval of information can be handled automatically.

As noted earlier there is a semantic gap between the user and retrieval systems. The user queries the system with his/her ideas and the system presents it findings for such ideas as an answer. There are different manners for query modeling including: 1. Query by keyword [3], 2. Sketch-based query [4] 3. Example-based query[5]and 4. Semantic-based query [6]. Methods 2 and 3 are usually used in CBIR and CBVR [7]; In such methods media is modeled by features such as color, texture, etc. and special relationships between objects and movements. In CBIR suitable features are extracted and related to high level concepts and meanings, while the system does not know such meanings. In fact the system represents and models the video contents in such a way that is efficient for content retrieval such as in [8], [9] and [1]. In [1] we proposed a method based on textual query for video retrieval. The number of concepts in that work is very small while many concepts can be retrieved by using the sketch query. Approaches introduced for video retrieval are mostly based on trajectory retrieval in recent years. A group of such methods use the semantic description of object motion to retrieve concepts from videos, such as those reported in [6], [11], [12], [13] and [14]. [6]offers a structure that uses sketch and keyword queries to retrieve traffic trajectories. The method presented here optimized the sketch query and then retrieved the trajectories by using optimized query. And also**,** this method may be used for image retrieval systems. The following section includes the detail of proposed method in trajectory and image retrieval. Then, the experimental results are presented in Section 3. In the experimental results section, we examine the proposed method on the two data sets. The first data set includes the traffic surveillance video information and the second dataset includes the image datasets.

2. PROCEDURE STRUCTURE

In sketch-based retrieval, it has been observed that different user sketches for one query present different results. This is because the sketch is not optimized among the possible answers. The system presented in this paper optimizes the user sketch before implementation. In this system, the optimization is performed by using Genetic Algorithm (GA) and is applied for both image and trajectory retrieval purpose. GA is a random optimization method [16] that widely used in various applications [16], [17].

In fact**,** optimizing the sketch query can be used in CBIR and CBVR systems. The proposed optimization was tested on two methods; first, trajectory retrieval from traffic surveillance video and second, image retrieval from database. In trajectory retrieval schema, once the spatial vectors of motion are extracted, a user sketch is used to extract the trajectories. Retrieved trajectories are considered as the primary population. Then GA by using a proposed fitness function optimizes the user sketch. And finally optimized sketch is used to extract the answers. In Subsection 2.1, the trajectory retrieval system based on the query optimization is presented and in Subsection 2.2, this manner is proposed to retrieval a flower picture from an image dataset.

2.1. Trajectory Retrieval

Extraction of trajectories is a very critical part of the traffic behavior retrieval process. The overall structure is such that first the dynamic background is separated from the static one and all motion features are extracted. The feature vector of each trajectory may be of any size in this stage. In the next stage, the feature vectors of all trajectories are normalized and turned into vectors with known lengths. Then, trajectory classification is performed according to normalized features. Fig. 1 demonstrates the stages of motion extraction.

We used consequential frame differences to separate the background and dynamic foreground of a given video. The scene may have dimensions proportionate to the frame or smaller. The processing of two consecutive frames gives the following feature vector for each moving block.

$$Fv\,(i)=\{x,y,r,g,b,Vx,Vy\} \qquad (1)$$

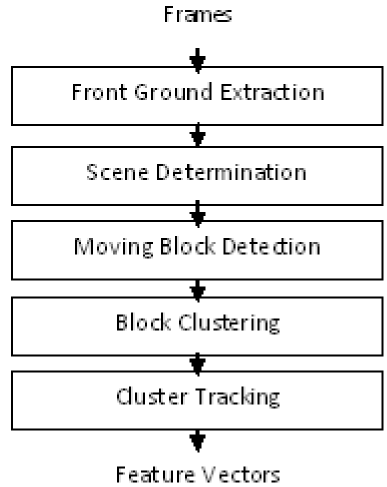


Fig. 1.Trajectory extraction

Y is the block beginning coordinate; R, G, B are average block pixel color values; Vy, Vx are the velocity (speed) resulting from block motion calculated by equation V = dx/dt. Once the features are extracted, clustering occurs through Fuzzy-K means. After training, the cluster center is formed by the blocks pointing to the displacement of a group of close blocks with similar displacement in the same direction. The following algorithm is used to follow the cluster center motion:

- Take cluster *i* of frame *t*
- Use the following equations to estimate the cluster center point in frame *i+t:*
  - $X = x_i + dx_i$
  - $Y = y_i + dy_i$
- Search the set of all cluster centers in frame *i*+*t* and find the closest cluster (*j*) to vector {X,Y, dx_i, dy_i}
- Add (x_i, y_i) to Fs set
- Let *i* = *j* and *t* = *t*+1
- If the cluster center has not exited the scene, proceed to the first

To normalize the resulting vectors, we take *n* samples from each spatial vector and extract them by finding the means on different parts of the trajectory.

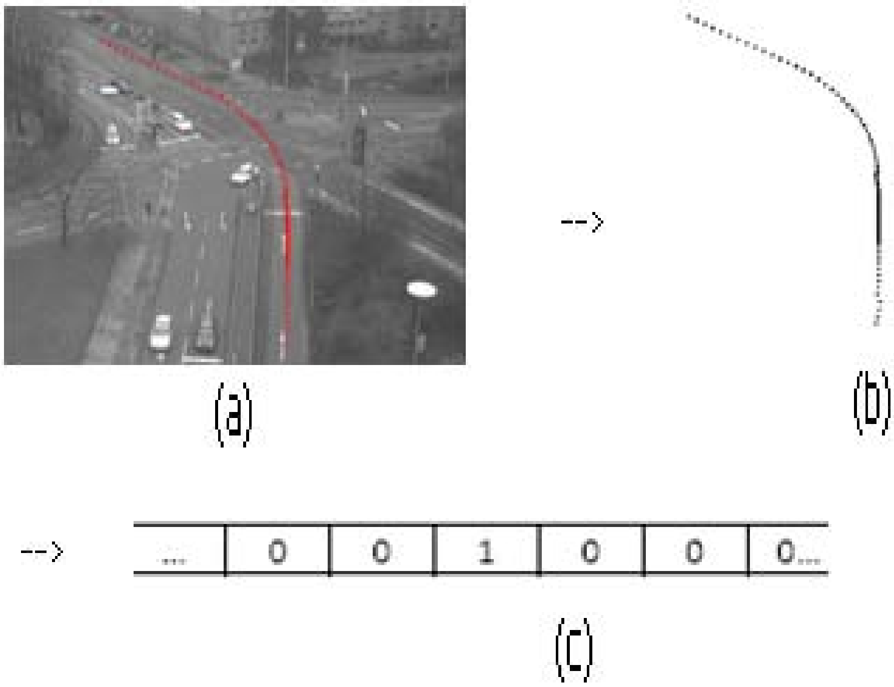


Fig. 2. Trajectory representation process; a) extracted trajectory, b) binary representation, c) bit stream as genotype representation

These normalized vectors may be used to retrieve the trajectories by sketch. In sketch-based retrieval, it has been observed that different user sketches for one query present different results. This is because the sketch is not optimized among the available trajectories.

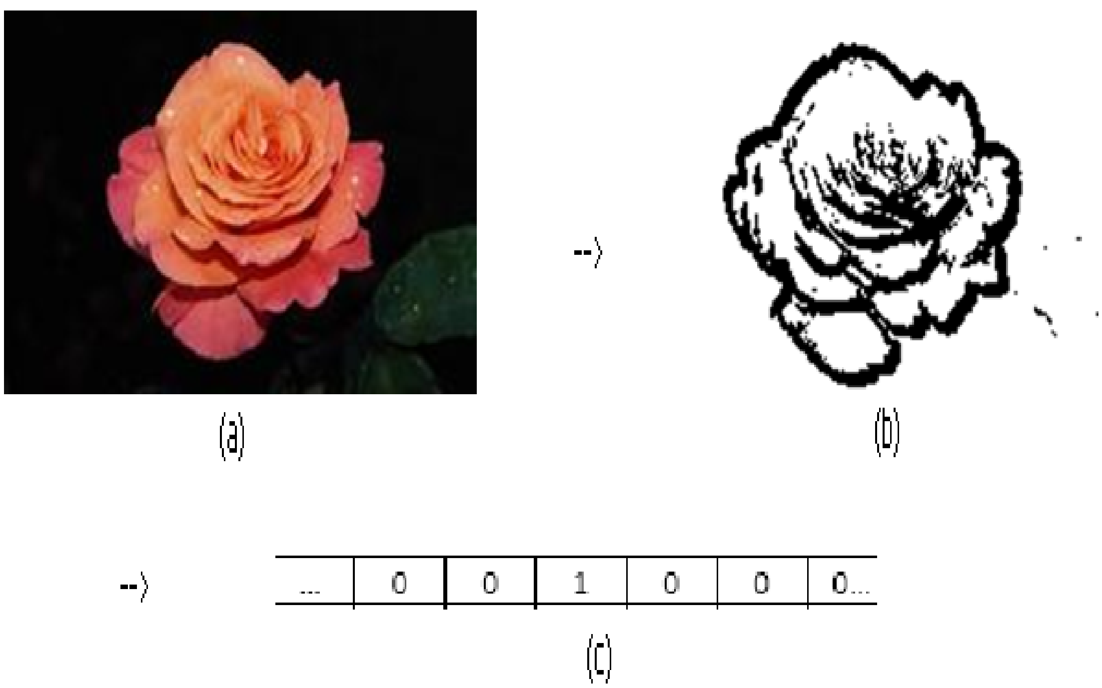
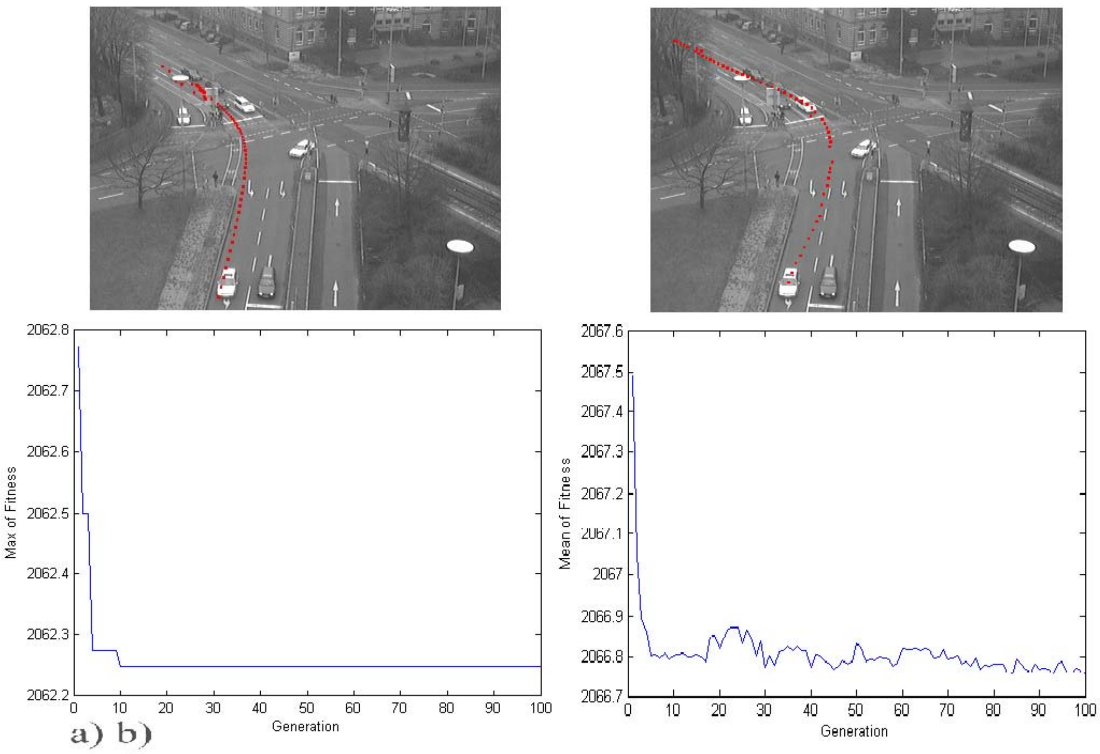


Fig. 3. Image representation process; a) main image, b) binary representation of filtered image c) bit stream as genotype representation

(a) (b)

Fig. 4. Trajectory extraction



a) b)
c) d)

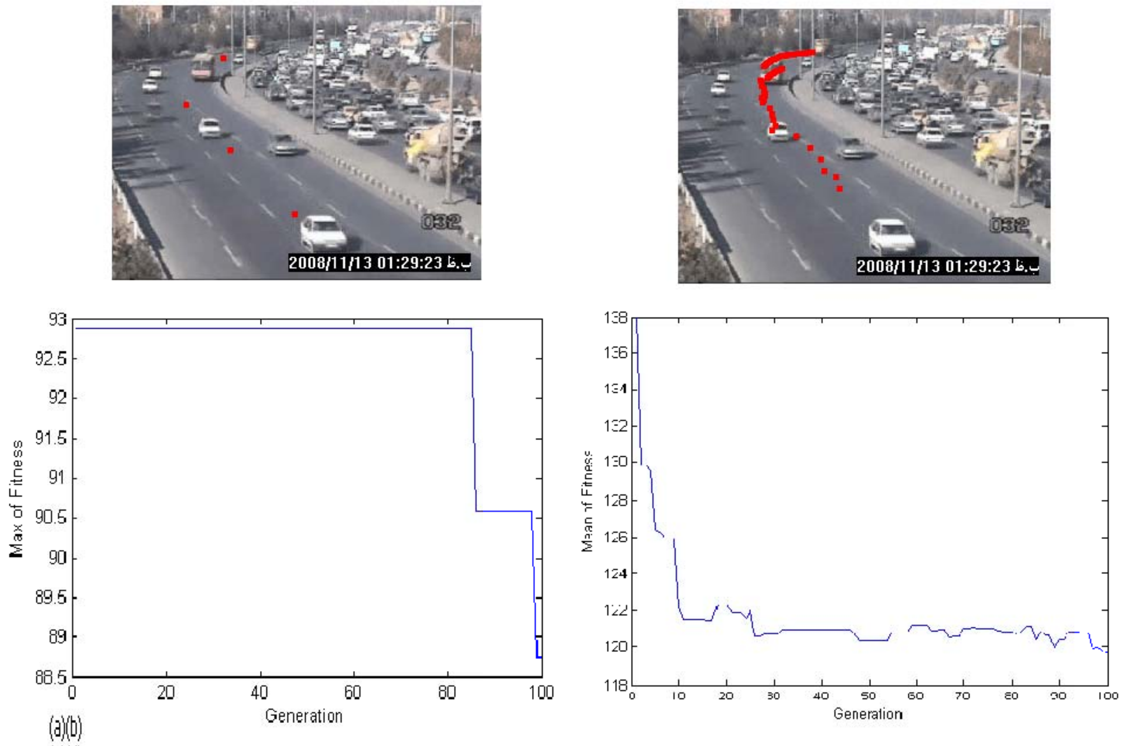
Fig. 5. Optimizing the user query by using GA (example 1); a) user query b) optimized query c) minimum fitness d) mean of fitness

To optimize the retrieval process, one stage of retrieval is performed by the user sketch. The retrieval criterion is the Euclidian distance from the user query. Retrieved trajectories are considered as the primary population. Fig. 2 presents the process of encoding and bit string genotype representation. Evaluating a trajectory needs sampling of *n* (number of query points) points on it. So phenotype representation is:

$$\boldsymbol{Traj_i} = \{(x_1, y_1),(x_2, y_2),\ldots,(x_n, y_n)\}$$
$$X_i = \{x_1, x_2, \ldots, x_n\},\ Y_i = \{y, y_2, \ldots, y_n\} \tag{2}$$

The following fitness function is now used to find the best answer:

$$\mathbf{Fitness}(\mathbf{Traj_i}) = \frac{|X_i - X_{query}| + |Y_i - Y_{query}|}{2} \tag{3}$$



(a)(b)
(c)(d)

Fig. 6. Optimizing the user query by using GA (example 2); a) user query b) optimized query c) minimum fitness d) mean of fitness

The optimized query may be achieved through reproducing and minimizing the said function by using GA. And finally optimized sketch is used to retrieve the trajectories.

2.2. Image Retrieval

In the first step of image retrieval, all images in the DB are filtered by using kirsch filter, and then, the system ranks the images by using a sketch query. Retrieved images are considered as the primary population for evolutionary optimization. Fig. 3 presents the result of filtering as a binary image as phenotype representation (Fig. 3.b), and bit stream genotype representation (Fig. 3.c).

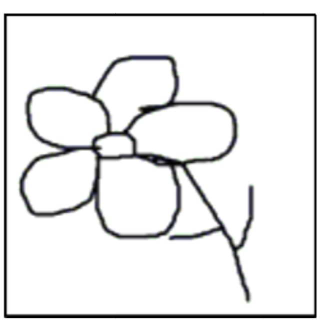


Fig. 7. User sketch query

The ranking and the evaluating measurement is the standard deviation of distance of all “1’ points in a binary image from the sketch query. The optimized query may be achieved through reproducing and minimizing the said measurement by using GA. And finally optimized sketch is used to retrieve the images.

3. EVALUATION

Two following subsection present the result of the trajectory retrieval and the image indexing.

3.1. Trajectory Retrieval Result

In order to test and assess (prefer) the offered algorithms in trajectory retrieval, video data was collected from two sources. One group of data was provided by the “Mashhad Traffic Surveillance and Control Center”, Mahhsad, Iran and the other group was acquired from the “Institut fur Algorithmen und Kognitive Systeme” web site. Overall, more than 3,000 frames from at least three surveillance stations were used in the database. Processing was entirely done on fixed views of crossroads. Processing occurred on each five frames, when the frame size was 768 × 576 and the block size was 15×15 and when the frame size was 320 × 240 and block size 5 × 5. Fig. 4.b presents all trajectories extracted from crossroad Fig. 4.a. And it has a total of 37 trajectories.

Fig. 8. Result of retrieval process by using user sketch

Fig. 9. Result of retrievalprocess by using optimized sketch

Table 1. Retrieval Result by using the Optimized Query

User Query	Retrieval accuracy	accuracy after optimizing the query
	43%	75%
	75%	93%

Fig. 5 and 6 shows the result of the user query optimization by GA and includes the main query, its optimized version, and its fitness information for different generations. Table 1 provides the effects of optimization on retrieval results. In the table accuracy formula is:

Accuracy = 100* (Correct Trajectories)/(Total + (False Trajectories)) (4)

3.2. Image Retrieval Result

To evaluate the proposed algorithm in image retrieval, 92 images was collected from Oliva Dataset and Wang Dataset. Collected data include more than eleven image categories.

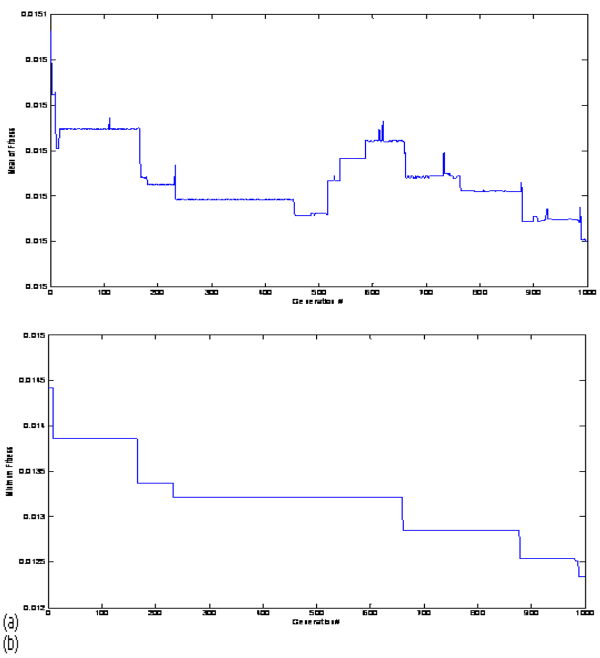


Fig. 10. Fitness information during optimization of flower image query by GA; a) mean of fitness, b) minimum fitness

For evaluating optimization method in image retrieval, consider the following scenario; the user queries the system with a sketch presented in Fig. 7. The top nine result of retrieval by using this orginal sketch is presented in Fig. 8 and top nine result of retrieval process using optimized query is shown in Fig. 9. According to the Fig, 8 and 9 the optimize query shows better results with respect to the original query. Fig. 10 shows the result of the user query optimization in image retrieval by GA and includes its fitness information (mean and minimum) for different generations.

4. CONCLUSION

The method we have proposed here applies GA to optimize the user query in order to retrieve the trajectories from traffic surveillance video and the images from given data base. The retrieval criterion in trajectory retrieval is the Euclidian distance from the sketch query. And the ranking and the evaluating measurement in image retrieval is the standard deviation of distance of all “1” points in a binary image from the sketch query. The optimized query achieved through reproducing and minimizing the said measurements by using GA. Also, a comparative improvement in retrieval by using optimized user sketches was observed. And experimental results show a high accuracy in the retrieval process.

REFERENCES

[1] E. Lotfi, H. R. Pourreza, **“Path Normalization for Traffic Surveillance Video Retrieval,”** *Majlesi Journal of Multimedia Processing*, Vol 1, No 2 September 2011.

[2] G. M. C. Snoek, B. Huurnink, and L. Hollink, **“Adding semantics to detectors for video retrieval,”** *IEEE Transaction on Multimedia*, Vol. 9, No. 5, pp. 975-985, 2007.

[3] Y.-L. Kang, **“Soccer video event detection with visual keywords,”** *ICICS-CM*, Singapore, pp. 15-18, December 2003.

[4] J. P. Collomosse, G. McNeill and Y. Qian, **“Storyboard sketches for content based video retrieval,”** *In Proc. Intl. Conf on Computer Vision (ICCV)*, Sep. 2009.

[5] Z. Q. Liu, L. T. Bruton, J. C. Bezdek, J. M. Keller, S. Dance, N. R. Bartley, and C. Zhang, **“Dynamic image sequence analysis using fuzzy measures,”** *IEEE Tran. on Syst., Man and Cyber.*, Vol. 31, No. 4, pp. 557-571, 2001.

[6] W. Hu, D. Xie, Zh. Fu, W. Zeng, and S. Maybank, **“Semantic-Based surveillance video retrieval,”** *IEEE Transaction on Image Processing*, Vol. 16, No. 4, pp. 1168-1181, 2007.

[7] G. Vasumathi Narayanan, “**A Survey of Content-Based Video Retrieval”,** *Journal of Computer Science*, Vol. 4 No. 6, pp. 1549-3636, 2008.

[8] G. Piriou, P. Bouthemy, and J.-F. Yao, **“Recognition of dynamic video contents with global probabilistic models of visual motion”,** *IEEE Tran. on Image Processing*, 2006, Vol. 15, No. 11, pp. 3418-3431, 2006.

[9] A. D. Doulamis, N. D. Doulamis Stefanos, **"A fuzzy video content representation for video summarization and content-based retrieval,"** *Signal Processing*, Vol. 80, No. 6, pp. 1049-1067, 2000.

[10] E. Lotfi, **"Trajectory Clustering and Behaviour Retrieval from Traffic Surveillance Videos,"** *Majlesi Journal of Multimedia Processing*, Vol. 1, No. 2, June 2011.

[11] M. Haag and H.-H. Nagel, **"Incremental recognition of traffic situations from video image sequences,"** Image Vis. Comput., Vol. 18, No. 2, pp. 137–153, 2000.

[12] Z. Q. Liu, L. T. Bruton, J. C. Bezdek, J. M. Keller, S. Dance, N. R. Bartley, and C. Zhang, **"Dynamic image sequence analysis using fuzzy measures,"** *IEEE Tran. on Syst., Man and Cyber.*, Vol. 31, No. 4, pp. 557-571, 2001.

[13] F. Safara, and S. Naderi, **"Fuzzy decision maker for knowledge discovery from image archives,"** 3rd *International Conference on Geometric Modeling and Imaging*, pp. 126-129, July 2008.

[14] H. Fashandi, A. M. Eftekhari-Moghadam, **"An image mining approach for clustering traffic behaviors based on knowledgediscovery of image databases,"** *IEEE Int. Conf. on Computational Intelligence for Measurement Systems and Applications*, pp. 203-207, July 2005.

[15] E. Lotfi, **"A Novel Hybrid System Based on Fractal Coding for Soccer Retrieval from Video Database,"** *Majlesi Journal of Electrical Engineering*, Vol. 6, No. 1 2012.

[16] A. Golmakani, Kh. Mafinejad, and A. Kouzani, **"The parasitic-aware design and optimization of CMOS distributed amplifier using multi objective genetic algorithm,"** *Majlesi Journal of Electrical Engineering*, Vol. 2, No. 2, pp. 1-8, 2009.

[17] H. M. Shirazi, **"An intelligent intrusion detection system using genetic algorithms and features selection,"** *Majlesi Journal of Electrical Engineering*, Vol. 4, No. 1, 2010.