ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

A Novel Optimum Real-Time Color Reduction on FPGA Based on Swarm Intelligence

Khairy Assar[1]

Computer Science Department, Arab East Colleges, Riyadh, KSA

Computers and Systems Engineering, Al-Azhar University, Cairo, Egypt[1]

Abstract: Data clustering is a popular approach for automatically finding set of objects into a specific number of clusters. Clustering is largely used in many fields including text mining, information retrieval and pattern grouping. Particle Swarm Optimization (PSO) is a population-based optimization algorithm modelled after the simulation of social behaviour of bird flocks and widely used for optimize problem solving. In clustering problem PSO gives optimal solution but takes long time (so called iterations) to find the optimum solution. The hybrid PSO and K-means algorithm is developed to automatically detect the cluster centers of geometrical structure data sets. The proposed algorithm gives the benefits for each of two-merged algorithms. K-means is fast algorithm, PSO optimize the solution. The implementation of the hybrid K-means PSO structure is realized in hardware. The clustering based on hybrid K-means PSO architecture is described by different technique for hardware description (i.e. block diagram) and implemented on field programmable gate array (FPGA). Its feasibility is verified by experiments. Results show that the proposed architecture implemented on the FPGA has a good clustering technique especially for testing with color reduction for true color video.

Keywords: Clustering, K-means, Color image reduction, Particle Swarm Optimization (PSO), field programmable gate array (FPGA), and Real Time Video Color Reduction.

I. INTRODUCTION

Clustering involves dividing a set of objects into a specified number of clusters [1][2][3]. Clustering is widely used in unsupervised document organization, pattern recognition, image processing, segmentation, vector quantization and information retrieval. Clustering algorithms can be divided into four main classes: partitioning methods, hierarchical methods, density-based clustering and grid-based clustering. An extensive survey on clustering techniques is described in [1][2][3][4]. Several papers have highlighted the efficiency of approaches inspired from the nature. In particular, a variety of algorithms inspired from the swarm intelligence observed from the ants behaviours such food hunting and nest building or from simulation of social behaviour of bird flocks, both have been introduced for solving several combinatorial optimization problems.

The major problem of K-means algorithms is that it is sensitive to the selection of the initial cluster and may converge to the local optima. Therefore, the initial selection of the cluster centroids decides the main processing of K-means and the clustering result of the dataset. The same initial centroids in a dataset will always generate the same clustering result. On other hand, the optimization algorithms, like ACO (Ant Colony Optimization) or PSO, both can run for clustering and give good result but the drawback is from the requested time consumed (number of iterations).Recently, hybridization between normal cluster technique (k-means, Self Organized Maps) and optimized techniques (ACO and PSO) give the best feature of two terms of hybridization, i.e. the optimum clustering result reached fast.

In this paper, we present a hybrid PSO and K-means document clustering algorithm that performs fast clustering and can avoid being trapped in a local optimal solution as well. The hardware implementation of the proposed hybrid PSO and K-means is studied and presented using field programmable gate array (FPGA). The designed structure mainly concerns color reduction for true color images [4][5][6].

Three approaches can be used to implement proposed hybrid PSO and K-means systems exist: microprocessor-based (software), dedicated ASIC and FPGA-based solutions. Software offer maximum flexibility to build different systems. However, PSO systems are highly parallel and microprocessor-based solutions perform poorly compared to their hardware counterparts. Dedicated ASIC is the best solution for achieving good performance, but such an approach dramatically reduces the adaptability of the system. Finally, FPGA-based systems provide both higher performance for parallel computation than software solutions and enhanced flexibility compared to ASIC thanks to their dynamic partial reconfiguration feature. They thus constitute the best candidate for evolving hardware. Moreover, their run-time reconfiguration features can be used to reduce execution time by hardwiring computationally intensive parts of the algorithm.

II. CLUSTERING

The problem of clustering is to partition a data set consisting of N points into K distinct set of clusters such that the data points within a cluster are more ‘similar’ among them than to data points in other cluster. The term similar, when applied to clusters, means closer by some similarity measure. The Euclidean and Manhattan distance between two dots in the document vector space can be

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

used to compute how similar the two documents represented by the two dots are. In order to manipulate equivalent threshold distances, considering that the distance ranges will vary according to the dimension number. Manhattan distance will be implemented to reduce mathematical operations. This algorithm uses the normalized Manhattan distance as the similarity metric of patterns m_p and m_j in the vector space. Equation 1 is the distance measurement formula.

$$d(m_p, m_j) = \sum_{k=1}^{dm} [m_{pk} - m_{jk}] / d_m \quad \text{----- (1)}$$

Where m_p and m_j are two vectors, d_m(adjust subscript) is the dimension number of the vector space. The vector dimension is dm. The two terms of proposed hybrid Kmeans-PSO will be studied in this section.

2.1. K-means Clustering Algorithm

K-means is simple unsupervised learning algorithms that solve the well-known clustering problem [7][8]. The procedure classifies a given data set through a certain number of clusters [8][9]. A prior knowledge of number of clusters is must for K-means clustering algorithm. The k-means clustering algorithm partitions data points into k clusters. The better choice of centroids is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. The new values for centroids are computed. A loop is generated for previous two steps. The loop is repeated until the change of centroids values is limited.

The K-means clustering algorithm as follows:
1. Initialize the centroids using random samples.
2. Decide membership of the patterns in one of the K-clusters according to the minimum distance from cluster center criteria.
3. Calculate new Cj centers as:

$$C_j = \frac{1}{n_j} \sum_{\forall d_j \in S_j} d_j \quad \text{----- (2)}$$

Where n_j is the number of data items in the j^{th} (adjust subscript) cluster

4. Repeat steps 2 and 3 until there is no change in cluster centers.

2.2. PSO Clustering Algorithm

The objective of the PSO clustering algorithm is to find out the proper centroids of clusters for minimizing the intra-cluster distance as well as maximizing the distance between clusters [4][5]. In the PSO clustering algorithm, the multi-dimension pattern vector space is modelled as the problem space in PSO. Each term in the pattern represents one dimension of the problem space. Each pattern vector can be represented as a dot in the problem space. A single particle in the swarm represents one possible solution for clustering the document collection. Therefore, the swarm represents a number of candidate clustering solutions for the document collection. Each particle maintains a matrix $X_i = (C_1, C_2, \ldots, C_i, .., C_k)$, where C_i represents the i^{th} cluster centroid vector and k is the number of clusters. At each iteration, the particle adjusts the centroid vectors' positions in the vector space according to its own experience and that of its neighbour particles. The average distance between a cluster centroid and a document is used as the fitness value to evaluate each particle has represented solution. The fitness value is measured by below equation

$$f = \frac{\sum_{i=1}^{N_c} \left\{ \frac{\sum_{j=0}^{P_i} d(O_i, m_{ij})}{P_i} \right\}}{N_c} \quad \text{---- (3)}$$

Where m_{ij} is the j^{th} vector, which belongs to cluster i. O_i is centroid vector of i^{th} cluster.), d is the distance between m_{ij} and cluster centroid O_i. P_i stand for the pattern number, which belong to cluster C_i. N_c stands for the cluster number.

2.3. Hybrid PSO and K-means Clustering Algorithm

K-means is faster than the PSO to find the final color-map (centroids). However, the k-means is less accurate than PSO, in addition that the problem of local minimum. PSO normally solve these problems but with long time (number of iterations). Merging between two algorithms is presented by many researcher [10][11]. The order of PSO relative to K-means is the main difference among these researches. In [12] of hybrid PSO k-means structure, k centroids assigned randomly to each particle of the swarm and each particle run number of iteration of k-means internal and then all particle update its centroids by PSO updating functions (velocity and position). This iteration is repeated until termination criteria occurred. In [13], the PSO iteration run first and the best one will be used as initial value for k-means iteration. On other words, some researches concern on pre-clustering approach and the other concern on post-clustering approach.

III. Proposed Model

The proposed hybrid technique also is based on k-means and PSO. Our concern is hardware implantation using FPGA, on other words, increase compactness and speed in presence of sufficient accuracy. The algorithm seems as merging of k-means and PSO algorithms together in the same time. Only one of the particles is work k-means and PSO while other particle works only by PSO algorithm. In second step, determine the nearest cluster for each particle centroids by Calculate the distance using Euclidean distance. The comparison between item of data set and all centroids for each particle perform in pipeline manner a will be discuses later. Also, in the same step an accumulation for the real data for each cluster for particle number one only which work as k-means clustering. While the other particle work with PSO algorithm which no need for this accumulation but only need summation of divination (distance between the original value of data set item and the center of nearest centroid).

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

3.1. The Flowchart of Proposed Hybrid Model

The algorithm starts with running k-means for low number iterations, the output result from K-means will feed as input one particle of the swarm while the other particles are initialized randomly. The flowchart of hybrid PSO–Kmeans is shown in Figure 4.

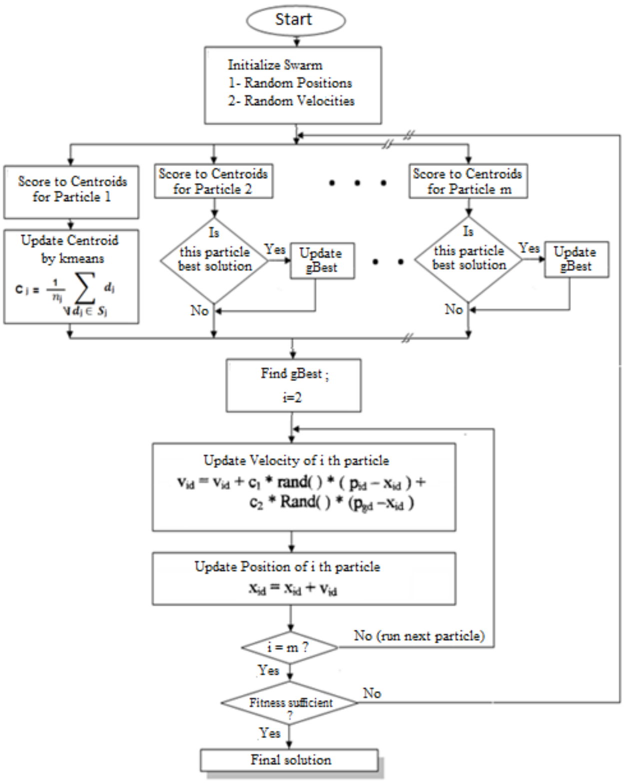

Fig.1 flowchart of hybrid parallel PSO – Kmeans

3.2. The Abstract Architecture of Proposed Hybrid Model

The abstracted architecture for each technique is built. The techniques differ in order between PSO and Kmeans. There are two techniques; the first technique runs the kmeans for each particle at each iteration before running PSO for all particle of the swarm. The second technique uses the PSO for all particles of the swarm and the final result is fed to Kmeans to optimize the best solution found by PSO. The proposed algorithm runs both of PSO and kmeans simultaneously. One particle run kmeans while the whole particle work as PSO and run updating equations of PSO algorithm. The proposed algorithm gives the same optimization result of first technique while the complexity and number of iterations are reduced. In addition the speed of optimization is increased. Also, the problems from kmeans are limited. The timing analysis for the proposed algorithm is studied. The algorithm starts with running k-means for low number iterations, the output result from K-means will feed as input one particle of the swarm while the other particles are initialized randomly. The abstract architecture of hybrid PSO-Kmeans is shown in Figure 4. Figure 5, shows the flowchart of hybrid PSO–Kmeans.

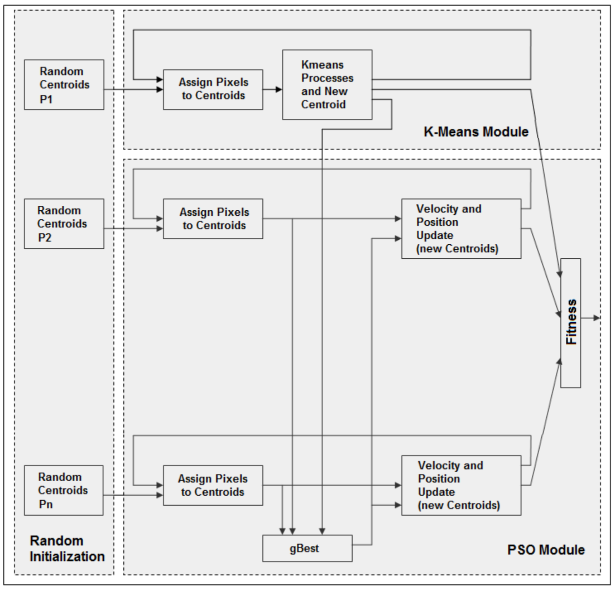

Fig. 2 abstract architecture of hybrid PSO-Kmeans

IV. REAL TIME SYSTEMS

The main constrains for any real time video processing is time constrain. The standard of video signal is 25 or 30 frame per second for PAL or NTSC standards. Therefore, the main aim of the proposed algorithm is to guarantee this constrain. For this purpose, the proposed algorithm works in time constraint technique. Therefore, the number of iteration for the algorithm is constant and the algorithm does not depend on satisfaction of level of optimization.

4.1. Modification for proposed algorithm for video processing

The proposed algorithm guarantees a level of parallelism as listed before. As long as the changes between two cascades frames in video stream are small, also the change of centroids for cascaded frames is also small. So, the algorithm is modified by cancelling the first step of the algorithm (in which we assign random initialization for all particles of the swarm at each frame starting).
The modification guarantees two advantages:

- The time required for the swarm to reach to sufficient level of optimization which depends on centroids of the particles is very low, and then it spend the time of frame for optimization only.
- The number of iteration for next frame is seemed like double or triple number of the first frame.

The result of this modification for more optimization will be shown later.

4.2. Timing Analysis

The time in video processing is very important. In video with 25 frames per second the time allowed for processing one frame is 1/25 second (40 mSec). So, the proposed system and then the proposed algorithm must grantee this constraint, which will be seen in block diagram in the following chapter. Briefly the system store one frame into first system memory which has two ports (Dual Port Block Impeded RAM) in the same time the color reduction subsystem works and stores the reconstructed image into

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

the second dual port RAM. On other hand another subsystem works as viewer for this frame into viewing device (VGA screen). So, as seen the algorithm of color reduction can take full time of one frame (i.e. 40 mSec). Many of video source work in interlaced manner, so the system can start working directly after odd line ware received. The processing time can be expanded to 1.5 frame time. The algorithm has some parts that work in parallel processing (pure parallel and pipeline) plus some other parts that work in serial processing.

The Parallel Part:
All particles work at the same time of comparing the same pixel with its own first palette.
The Pipeline Part:
The next pixel pass through all particles (first item of palette) while the previous pixel handled by the all particle but the second item of palette.
The Serial Part:
- Calculating the new centroids of K-means Particle (only one particle).
- Calculating the new position (new centroids) for all particles.

The time required for completing the parallel and pipeline parts depends only on the number of pixels and not depended on the number of clusters.

$T_{parallel} = ((C-1) * T) + (N * T)$

Where:
N = number of pixels
T = time required for processing one pixel (about 2 system clock, first clock for subtract pixel value with the centroid, second clock adding the differences)
C = number of clusters

While N is very large comparing to C so, the first term can be negligible. And $T_{parallel}$ become

$T_{parallel} = T * N = 2N$

The time of serial part :

$T_{serial}= T_{kmeans} + T_{pso}$

$T_{kmeans} = C * T_{div}$

(T_{div}: is the time for division operation for allocating new centroid of kmeans =~ 20 system clock)

V. CLUSTER IMPLEMENTATION ON FPGA

The proposed algorithm was implemented on FPGA for the purpose of verification. It is implemented with Parallel and pipeline technique. The overall system works as real time video processing. It receives a video signal in PAL or NTSC standard video decoding and out the final result also in real time after image reconstruction. The system receives one video frame, and decodes it to seem as raw binary image and storing it into Dual Port memory into the FPGA chip. The internal Block Memory RAM can be configured into many different configurations. The suitable configuration for our system is built, which is Dual port of eight-bit interface. The overall system structure of FPGA based real time parallel hybrid PSO and Kmeans for color reduction is shown in figure 5.24.

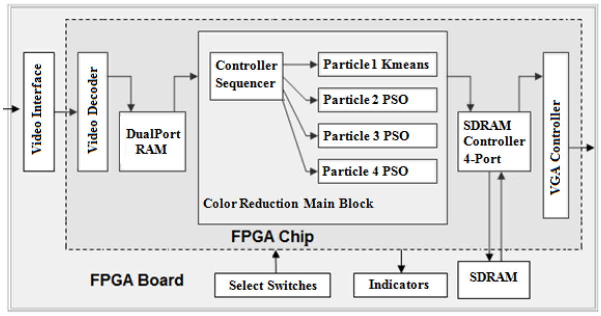

Fig. 3 overall system structure for Hybrid PSO Kmeans module implemented on FPGA

The system use many different components out of the board like Video interface chip, SDRAM, Switches, LEDs, Seven Segment Display, and VGA analog interface chip.

5.1. *Video Acquisition*

The FPGA board is equipped with an analog devices ADV7181 TV decoder chip. The ADV7181 is an integrated video decoder that automatically detects and converts a standard analog baseband television signal (NTSC, PAL, and SECAM) into 4:2:2 component video data compatible with 16-bit/8-bit CCIR601/CCIR656. The ADV7181 is compatible with a broad range of video devices, including DVD players, tape-based sources, broadcast sources, and security/surveillance cameras. The registers in the TV decoder can be programmed by a serial I2C bus, which is connected to the FPGA.

5.2. *Random Number Generation*

PSO behaviour depends on the random number and the way to generate. The random number cannot be replaced by constant values that all particles of the swarm move in the same owned direction or all moves in the same position. The generation of random number in hardware is performed by Liner Feedback Shift Register (LFSR). The construction of LFSR can take many forms. Figure 4 shows the block diagram for LFSR with two outputs. The system needs multiple of random number in PSO calculation each LFSR generates two outputs numbers.

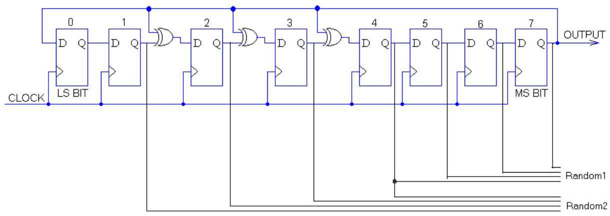

Fig. 4 the block diagram for LFSR with two outputs.

VI. SIMULATION AND RESULTS

Many of test phases were done for verifying system components. First, kmeans structure is built and tested with true color images. The reconstructed low color images are feedback to PC and viewed. The system feeds back to PC software to calculate the Mean Error (summation of Manhattan distance per each centroid) after each iteration of processing, and the number of pixels per each centroid. The second structure is built based on PSO.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

The PSO structure also built to feedback all possible results (centroids, number of pixel per centroids, image after reconstruction, and mean error per each particle). Finally, the proposed structure is built. A PC software built to verify the result and calculate the MSE and also for comparison with other structures. Figure 5 shows how the effect of adding one kmeans with swarms of PSO.

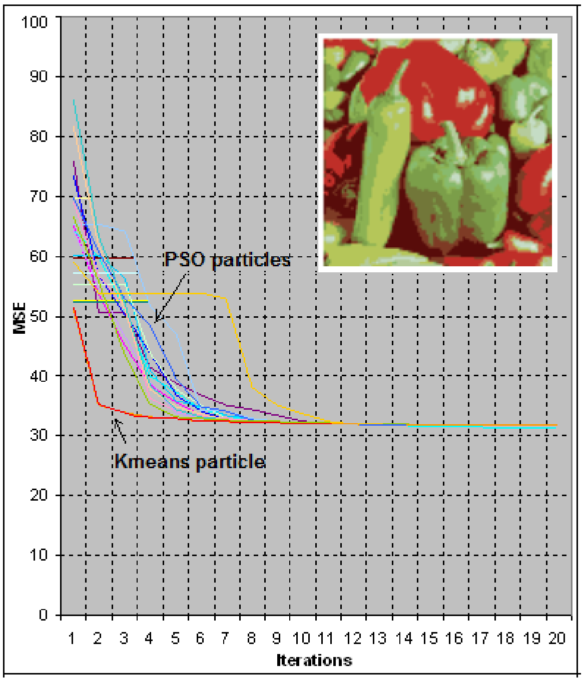

Fig. 5 the behaviour of proposed parallel PSO and Kmeans with Pepper true color image

The MSE of kmean is best and fast reached in ealier iterations, while the other particles give optimum result with recent iterations.

Other type of simulation was presented to test each block of FPGA. Figure 6 shows the timing diagram for LFSR and two random numbers generation.

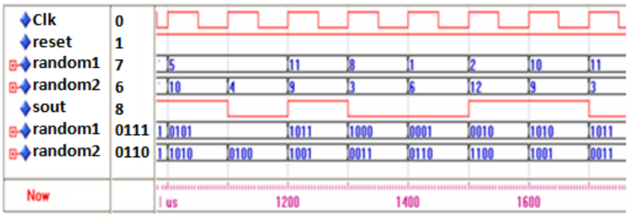

Fig. 6 the timing diagram for LFSR and two random numbers generation

Many structured have to be implemented until reaching final real time system.

- **The First system structure** is PC Simulator which is a PC based with GUI interface. This structure is based on the proposed algorithm but in sequential form. It helps us to readout the result of proposed algorithm and making the comparison with the output of other algorithms. In addition, it helps in tuning PSO parameters.
- **The Second structure** is FPGA Based that is built especially for testing algorithm after building on FPGA. This structure read an image from PC via serial interface, and run the color reduction algorithm in parallel form and returns the reconstructed image to PC via serial interface.
- **The Third structure** is the recommended structure, which is an FPGA based and work with real time video. In which a video signal was decoded into digital form, processing this frame, and again view the output reconstructed frames into VGA monitor.

The final structuer was implemented on FPGA board and run with real time video. The real time videofeed to system from satalitte video receiver. The system can handle two types of video system (NTSC and PAL). The frames was processed in FPGA board and viewed by monitor with VGA input port. Figure 7 view a picture for the system work in real time.

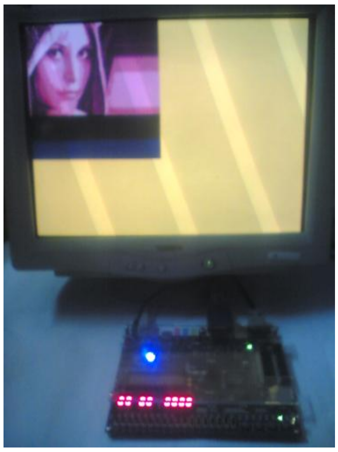

Fig. 7 experimental environment for Hybrid PSO-Kmeans based color reduction on FPGA

VII. CONCLUSION

In this paper, the implementation of optimized color reduction using FPGA was proposed. PSO is used to find an optimum solution. Only simple operations are required to adjust the underlying parameters. We adapt and create some techniques to speed up the process. The experimental result demonstrates the implementation success of the proposed color reduction with the PSO algorithm using FPGA. Some of the features of the color reduction with the PSO algorithm can be summarized as follows,Color reduction is realized based on digital circuits, Hardware implementation of the PSO, and Hardware implementation can take advantage of parallelism to reduce the processing time.

In the future, as the FPGA capacity and speeding increase, more parallelism can take place to expand the size of image to be clustered. Also, hybrid techniques for PSO with other clustering techniques can be merged for more clustering optimality.

ACKNOWLEDGMENT

The authors would like to acknowledge the director of Arab East Collages for helping to publish this work.

REFERENCES

[1] A. K. Jain and R. C. Dubes. "Algorithms for Clustering Data, Prentice Hall", 1988.

[2] T. Warren Liao, "Clustering of time series data—a survey", Pattern Recognition, Volume 38, Issue 11, November 2005, Pages 1857-1874

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

[3] Jain, A. K. And Flynn, P. J. Image segmentation using clustering. In Advances in Image Understanding: A Festschrift for Azriel, (1996).
[4] Wang, Z.S., Li, L.C. and Li, B. "Reactive power optimization based on particle swarm optimization and simulated annealing cooperative algorithm", Journal of Shandong University (Engineering Science), 2008, 38(6),15-20
[5] Maurizio Filippone, Francesco Camastra, Francesco Masulli, Stefano Rovetta, "A survey of kernel and spectral methods for clustering", Pattern Recognition, Volume 41, Issue 1, January 2008, Pages 176-190
[6] Jie He, Hui Guo,"A Modified Particle Swarm Optimization Algorithm", TELKOMNIKA, Vol. 11, No. 10, October 2013, pp. 6209 ~ 6215
[7] HaiTao,ThomasS.Huang, "COLOR IMAGE EDGE DETECTION USING CLUSTER ANALYSIS", Beckman Institute, University of Ilinoisat Urbana- Champaign Urbana,IL61801,USA,
[8] B. Bhanu and B.A. Rarvin, "Segmentation of natural scene", Pattern Recognition, vol. 20, pp. 487-496, 1987.
[9] G.B. Coleman and H.C. Andrews, "Image segmentation by clustering", Proc. IEEE, vol. 67, pp. 773-785, 1979.
[10] M. Law , M. A. T. Figueiredo , A. K. Jain . "Simultaneous Feature Selection and Clustering Using Mixture Models", IEEE Transactions of Pattern Analysis and Machine Intelligence. vol.26, no. 9, pp. 1154-1166, September 2004.
[11] Chen, T.Q. and Y. Lu, "Color Image Segmentation--An Innovative Approach", Pattern Recognition, 2001. 35: p. 395-405.
[12] Omran, M.; Salman, A. &Engelbrecht, A.P. (2002). "Image Classification Using Particle Swarm Optimization", 2002.
[13] X. Cui and T. E. Potok, "Document Clustering Analysis based on Hybrid PSO+K-means Algorithm", The Journal of Computer Science, 2005, volume1(3), pp 27 – 33.

BIOGRAPHY

Khairy Assar
Assistant Professor, Arab East Colleges, Riyadh, KSA.
Faculty of Engineering, Al-Azhar University, Cairo, Egypt.