

Object Recognition Based on Image Segmentation and Clustering

¹S. Thilagamani and ²N. Shanthi

¹Department of Information Technology,
M. Kumarasamy College of Engineering,
Thalavapalayam, 639113, Karur

²Department of Information Technology,
K.S.R College of Technology, Thiruchengode, Namakkal

Abstract: Problem statement: This study deals with object recognition based on image segmentation and clustering. Acquiring prior information of an image is done via two separate processes. **Approach:** The first process deals with detecting object parts of an image and integration of detected parts into several clusters. All these cluster centers form the visual words. The second process deals with over segmenting the image into super pixels and formation of larger sub region using Mid-level clustering algorithm, since it incorporates various information to decide the homogeneity of a sub region. **Results:** The outcome of the two processes are used for the similarity graph representation for object segmentation as proposed. In order to model the relationship between the shape and color or texture matrix representation has been used. Mask map ensures that the probability of each super pixel to harp inside an object. **Conclusion:** The basic whim is to integrate all the priors into an uniform framework. Thus the ORBISC can handle size, color, texture and pose variations better than those methods that focus on the objects only.

Key words: Super pixel, oversegmentation, similarity dependence graph, euclidean distance measure, object recognition, super pixels, object segmentation, image editing, graph representation, euclidean distance formula

INTRODUCTION

Object segmentation is one of the fundamental problems in computer vision. Its goal is to segment an image into foreground and background, with the foreground solely containing objects of a class and background contains different attributes like color, texture. For this two algorithms are used namely supervised and unsupervised clustering algorithms. According to user feedback, grouping is done in supervised clustering. Object recognition system in practical needs to handle larger number of classes and objects, it is necessary that the learning does not involves any user feedback. This makes unsupervised clustering more appear in practical system. In order to increase the robustness of the system, Similarity Dependence Graph (SDG) is used. As it integrates object recognition and object segmentation into an commixed process.

The inputs to the similarity dependence graph are the outcomes of the two separate processes. The first process involves detecting the object parts of an image

and integration of detected parts into several clusters. The second process deals with over segmenting the image into super pixels. The vertices of the similarity graph represent the super pixel and object parts of an image. These vertices are connected by directed and undirected edges. The directed edge represents the dependence between the entities (for recognition) whereas the undirected edge represents the similarity between the entities (for segmentation). The matrix that can be formed using the graph is known as mask. In mask map, there is difficulty in directly applying shape priors to super pixels and color or texture priors to object parts, because object parts are square while super pixels are irregular (Sheng and Qi-Cong, 2009). In order to overcome these difficulty SDG is used.

Related works: Clustering is a process of organizing the objects into groups based on its attributes. Clustering techniques can be classified into supervised that demands human interaction to decide the clustering criteria and the unsupervised clustering that decides the clustering criteria by itself. Supervised

Corresponding Author: S. Thilagamani, Department of Information Technology, M. Kumarasamy College of Engineering Thalavapalayam, 639113, Karur

includes hierarchical approaches such as relevance feedback techniques and unsupervised includes density based clustering methods. An image can be grouped based on keyword (metadata) or its content (description). In keyword based clustering, a keyword is a form of font which describes about the image keyword and its different features. The similar featured images are grouped to form a cluster by assigning value to each feature.

In content based clustering, a content refers to shapes, textures or any other information that can be inherited from the image itself. The tools, techniques and algorithms that are used, originate from fields such as statistics, pattern recognition, signal processing. Clustering based on the optimization of an overall measure is a fundamental approach explored since the early days of pattern recognition.

These clustering techniques are done to perform image segmentation. Where image segmentation refers to the process of partitioning a digital image into multiple segments (based on pixels). It is a critical and essential component of image analysis system. The main process is to represent the image in clear way. Real world image segmentation problems (Shirakawa and Nagao, 2009) actually have multiple objectives such as minimize overall deviation, maximize connectivity, minimize the features or minimize the error rate of the classifier.

Image segmentation is a multiple objective problem (Saha and Bandyopadhyay, 2010). It involves several processes such as pattern representation, feature selection, feature extraction and pattern proximity. Considering all these objectives is a difficult problem, causing a gap between nature of images. To bridge this gap multi-objective optimization approach is an appropriate method (Guliashki *et al.*, 2009). The objective of image segmentation is to cluster pixels into salient image regions i.e., regions corresponding to individual surfaces, objects or natural parts of objects.

MATERIALS AND METHODS

A segmentation might be used for object recognition, image compression, image editing. The quality of the segmentation depends upon the digital image. In the case of simple images the segmentation process is clear and effective due to small pixels variations, whereas in the case of complex images, the utility for subsequent processing becomes questionable. In order to make the system practically implementable there is a need for unsupervised object segmentation, which does not demand any human interaction.

Images can be represented graphically either discrete or continuous, a discrete one is based on pixels

and a continuous one is based on points in a plane. Graph based representations are mainly used in image analysis to represent irregular structures. Graphs were used in segmentation, shape matching, video action description, technical drawings, fingerprint recognition,. By using graphical representation we can conveniently achieve simultaneous segmentation and recognition by integrating both top-down and bottom-up information into a unified process.

Implementation: Superpixel formation: An image can be grouped based upon its properties such as brightness, color, pixel value. By converting pixels into superpixel groups reduces the computational cost and complexity (Achanta *et al.*, 2010). The main aim of the superpixel formation is to achieve oversegmentation. Superpixel should be local, coherent and preserve most of the information necessary for segmentation. The value of the superpixel is the average of the all the pixel values. Each superpixel should be unique and provide accuracy improvements. The pixel difference within a superpixel should be minimum whereas the difference between two different superpixels should be maximum.

It is computationally efficient in reducing the complexity of images from hundreds of thousands of pixels to only a few hundred superpixels. It is also represents pair wise constraints between units efficiently, while only for adjacent pixels on the pixel-grid, can now model much longer-range interactions between superpixels. The superpixels are perceptually meaningful whereas such superpixel is a perceptually consistent unit, i.e. all pixels in a superpixel are most likely uniform in, say, color and texture. It is near-complete because superpixels are results of an oversegmentation, most structures in the image are conserved. There is very little loss in moving from the pixel-grid to the super pixel map. The Novel Recursive Clustering Algorithm is used to over segment the image into super pixels.

NRCA: Novel Recursive Clustering Algorithm is a gradient ascent based algorithm. It refines the cluster center value at each iteration. Our algorithm takes as input a desired number of approximately equally-sized superpixels K . For an image with N pixels, the approximate size of each super-pixel is therefore $N=K$ pixels. For roughly equally sized superpixels there would be a superpixel center at every grid interval $S = \text{Sqrt}(N/K)$.

At the onset of our algorithm, we choose K superpixel cluster centers $C_k = [R_k; G_k; B_k]^T$ with $k = [1, K]$ at regular grid intervals S . Since the spatial extent of any superpixel is approximately S^2 , we can safely

assume that pixels that are associated with this cluster center lie within a $2S \times 2S$ area around the superpixel center on the xy plane. This becomes the search area for the pixels nearest to each cluster center. Euclidean distances in CIELAB color space are perceptually meaningful for small distances (Eq. 1). If spatial pixel distances exceed this perceptual color distance limit, then they begin to outweigh pixel color similarities (resulting in superpixels that do not respect region boundaries, only proximity in the image plane).

Algorithm I:

Input: Image of $n \times n$ resolution

- (1): For $i=0$ to no. of pixels
 - If the i pixels is not visited already then process it
 - (2): Create cluster and add i to it
 - (3): Assign window size
 - (4): Check adjacent of the i for the similarity with i
 - (5): Add all the adjacent similar pixels into i clusters
 - (6): Mark i and other pixels in cluster as visited
 - (7): End For
 - (8): Draw boundary for each and every clusters
- Output: Display the results of superpixel image

Over segmentation: Over segmentation is the process by which the objects being segmented from the background are themselves segmented or fractured into subcomponents. It also refers to when the image is intentionally broken up into hundreds or thousands of small segments, more than necessary to perform segmentation into the objects in the scene. Typically they are uniform in size and commonly referred to as super pixels. It increases the chances of extracting important boundaries. It does so at the cost of creating many insignificant boundaries. In this case, prefiltering techniques should be used in an attempt to eliminate noise, improve inter object definition or smooth image textures, all of which might cause segmentation difficulties. If these techniques are not sufficient, grouping process is used following over segmentation in order to reassemble the objects into singular image events.

The effect of using larger segments is to increase the area of support, which usually improves the reliability and accuracy of pixel correspondence. Using segments, correct matches are possible even in the presence of noise, intensity bias, or slight deviations. The segment size needs to be at a trade-off point where the amount of information within a segment is sufficient for matching without compromising the characterization of the true disparity distribution.

If a segment is too small, it is difficult for it to unambiguously find the correct pixel correspondence. As a result, some mechanism for using information

from neighboring segments is typically required to reduce the ambiguity. The use of over-segmentation strikes a good balance between providing segments that contain enough information for matching and reducing the risk of a segment spanning multiple objects (Shirakawa and Nagao, 2009; Guliashki *et al.*, 2009). The use of over-segmentation also reduces the computational complexity of the algorithm, since disparities only need to be estimated per-segment rather than per-pixel. Given the smaller size of the segments, more information needs to be shared between segments to find correct correspondences than other segmentation approaches. However, more confidence can be placed in simple matching functions than with single pixel approaches.

Algorithm II:

Input: Superpixel image of $n \times n$ resolution

- (1): For each cluster C_i
 - For all C_j adjacent to C_i
 - If (distance (C_i, C_j) < threshold)
 - Then change values of C_j to C_i
 - (2): End For
- Output: Display oversegmented image

Graph formation: In graph based algorithms each superpixel is treated as a node in a graph and edge weight between two nodes are set proportional to the similarity between the two superpixels (Yu *et al.*, 2002; Felzenszwalb and Huttenlocher, 2004). Let $G=(V,E)$ be an undirected graph with vertices $v_i \in V$, represents a set of object parts and edges $v_i, v_j \in E$, represents the similarities between the superpixels. Each edge $(v_i, v_j) \in E$, has a weight value $dis((v_i, v_j))$ proportional to similarities between v_i and v_j and also it is non negative measure (Taghouti and Mami, 2010; Thilagamani and Shanthi, 2011a; 2011b). The dissimilarity is calculated based on the color, motion, location, or some other attributes.

The relationship can be identified using the following entities: Conditional dependency matrix and similarity dependency matrix.

Conditional dependency matrix: Let P be the matrix and V be the vertex with several nodes and they denoted as $V=\{v_1, v_2, \dots, v_n\}$ where n represents total number of nodes. The object parts are used as vertices in conditional dependency matrix. It is referred as asymmetric matrix because the node v_1 satisfies v_2 whereas v_2 doesn't satisfies v_1 in terms of conditional measures:

$$v_{ij} \neq v_{ji}$$

In matrix P, the term p_{ij} defines the conditional measure of v_i and v_j :

$$P=[p_{ij}]_{n \times n}$$

In case of conditional dependency matrix, more number of trained images are necessary, this increases the time complexity for an image.

Similarity dependency matrix: In this dependency matrix, the edge value is based on the (dis)similarity entities. Let E be the edge with several superpixel nodes. It is denoted as $E=\{v_1, v_2, \dots, v_n\}$ whereas n represents the total number of superpixel nodes. Euclidean distance formula has been used to find distance between the two superpixel nodes (Sayeed *et al.*, 2009; Abas and Ono, 2010; Odeh *et al.*, 2009; Jusoff, 2010; Al-Haddad *et al.*, 2009; Al-Saqer *et al.*, 2010; Nazif and Lee, 2010; Sleit *et al.*, 2009; Moghaddasi *et al.*, 2009). This formula is simple and easy to calculate distance between the nodes. It is represented as dis:

$$Dis = \sqrt{(r_i - r_j)^2 + (g_i - g_j)^2}$$

Hence the similarity matrix is symmetric as the distance between v_i and v_j and v_j and v_i are equal:

$$Dis(v_i, v_j) = Dis(v_j, v_i)$$

In this edge is represented as:

$$E=[a_{ij}]_{n \times n}$$

where, a_{ij} represents the similarity between v_i and v_j .

Algorithm III:

Input: Oversegmented image of $n \times n$ resolution

- (1): For each cluster C_i
 - For $j=0$ to all cluster
 - Calculate the distance between C_i and C_j
- (2): Store the result in the two dimensional array
- (3): End For
- (4): End For

Output: Graph representation of an image.

Object identification: Mask is a small matrix, whose values are called weights. Each mask has an origin which is usually one of its positions. The origin of symmetric masks is usually their center pixel position. For non-symmetric masks, any pixel location may be chosen as the origin. Mask map increases the

probability of Object Parts and Super Pixel to happen inside an Object. Its basic notion is to integrate all the priors into a unified framework. In this work, the segmentation task only needs the probability of Super Pixels, because the Object Parts cannot be waived, as it carries shape Priors (Al Rahedi and Atoum, 2009; Al-Salami, 2010; Harishchander *et al.*, 2010; Maalla *et al.*, 2009).

The mask map is first segmented, in which the two nearby regions, having the closest intensities are merged, as long as the intensities lie below the certain threshold value. Using similar criteria, threshold value T is selected and made to grow beginning with the regions with the intensity greater than $(1+T)/2$ and merge the next adjacent region with the highest intensity until all the intensities of adjacent regions fall below T .

Algorithm IV:

Input: Superpixel image of $n \times n$ resolution

- (1): C_i = first cluster
- (2): For all C_j
 - If (distance(C_i, C_j) < threshold)
 - Mask it to other color
- (3): End For

Output: Segmented object of an image.

Flowchart:

RESULTS

In this we represent the experimental results of using similarity dependence graph model for unsupervised object identification. In order to achieve object identification, an input image is shown in Fig. 1 which should be of resolution 440×380 . Depending upon the similarity among the pixels, the input image is segmented to form a super-pixelated image as shown in Fig. 2. The similar super-pixels are grouped to form an over-segmented image that is shown in Fig. 3. The similarity dependency graph as shown in Fig. 4 is formed by taking an over-segmented image as input in which edge value is based on the similarity entities. The edge value E is the collection of super-pixel nodes and the distance between super-pixel nodes are calculated using Euclidean distance formula as shown:

$$Dis = \sqrt{(r_i - r_j)^2 + (g_i - g_j)^2}$$

The similarity dependency graph is called so, because the distance between V_i and V_j is equal to the distance between V_j and V_i . From the above graph, mask map is learned and object is identified as shown in Fig. 5.



Fig. 1: Input image

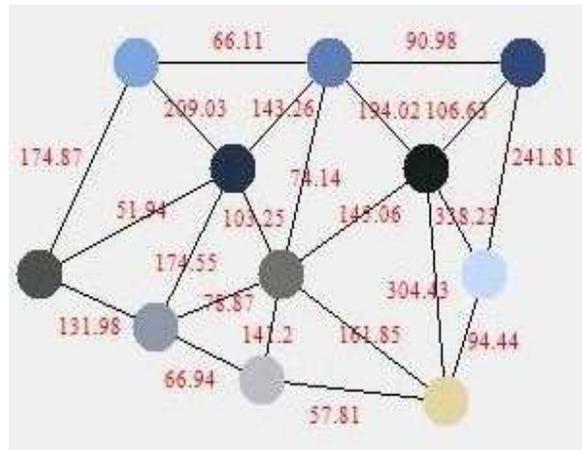


Fig. 4: Graph formation



Fig. 2: Superpixel formation



Fig. 5: Object identification



Fig. 3: Oversegmentation

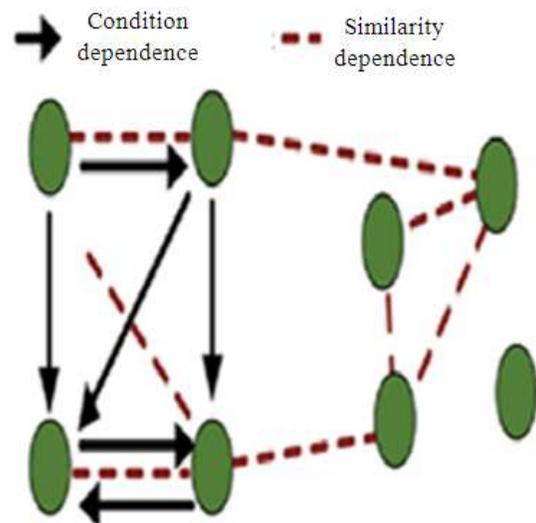


Fig. 6: Similarity dependency matrix

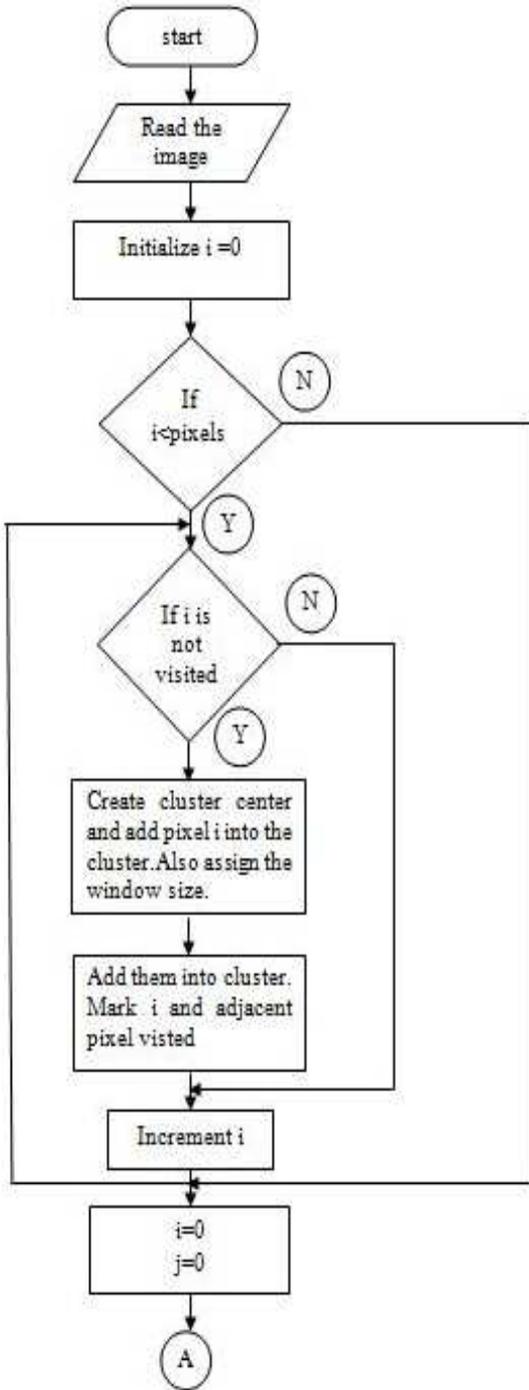


Fig. 7: Flowchart-cluster formation

For object identification the mask map is first segmented, in which the two nearby regions, having the closest intensities are merged, as long as the intensities lies, below the certain threshold value (Fig. 6-9).

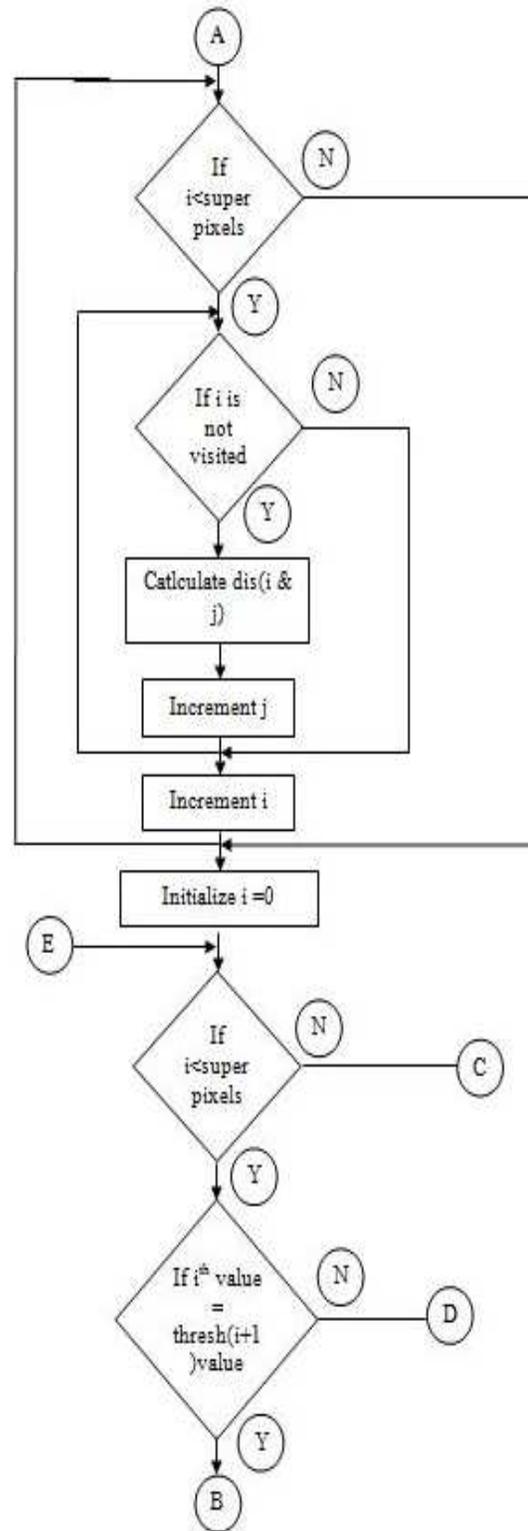


Fig. 8: Flowchart-superpixel comparison

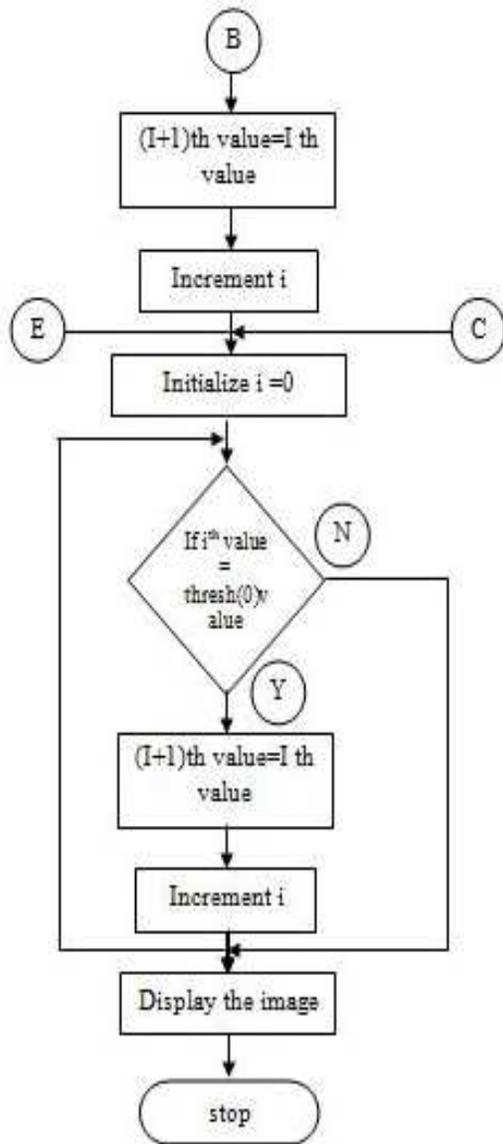


Fig. 9: Flowchart-image display

DISCUSSION

The Novel Recursive Clustering Algorithm is used to over segment the image into super pixels. single object is identified from the given input image.

CONCLUSION

In this study we proposed an approach for identifying the object from the input image. Object identification involves super pixel formation, over segmentation, graphical representation. As this study

only identifies a single object from an image, it can be enhanced to identify multiple objects in the future.

REFERENCES

- Abas, K.H. and O. Ono, 2010. Implementation of multi-centroid moment invariants in thermal-based face identification system. *Am. J. Applied Sci.*, 7: 283-289. DOI: 10.3844/ajassp.2010.283.289
- Achanta, R., A. Shaji and K. Smith, 2010. Aurelien Lucchi, Pascal Fua and Sabine Susstrunk, SLIC Superpixels, EPFL Technical Report no. 149300.
- Al Rahedi, N.T. and J. Atoum, 2009. Solving the traveling salesman problem using new operators in genetic algorithms. *Am. J. Applied Sci.*, 6: 1586-1590. DOI: 10.3844/ajassp.2009.1586.1590
- Al-Haddad, S.A.R., S.A. Samad, A. Hussain, K.A. Ishak and A.O.A. Noor, 2009. Robust speech recognition using fusion techniques and adaptive filtering. *Am. J. Applied Sci.*, 6: 290-295. DOI: 10.3844/ajassp.2009.290.295
- AL-Salami, N.M.A., 2010. Evolutionary algorithm definition. *Am. J. Eng. Applied Sci.*, 2: 789-795. DOI: 10.3844/ajeassp.2009.789.795
- Al-Saqer, S.M., P. Weckler, J. Solie, M. Stone and A. Wayadande, 2010. Identification of pecan weevils through image processing. *Am. J. Agric. Biol. Sci.*, 6: 69-79. DOI: 10.3844/ajabssp.2011.69.79
- Guliashki, V., H. Toshev and C. Korsemov, 2009. Survey of Evolutionary Algorithms used in multiobjective optimization. *Problems of Engineering Cybernetics and Robotics*, Bulgarian Academy of Science, pp: 42-54.
- Harishchander, A., D.A. Anand and S. Senapati, 2010. Analysis of resistance to human immunodeficiency virus protease inhibitors using molecular mechanics and machine learning strategies. *Am. Med. J.*, 1: 126-132. DOI: 10.3844/amjsp.2009.126.132
- Jusoff, K., 2010. A new approach in individual tree counting of nyatoh (*Palaquium* spp.) using an airborne spectroradiometer. *Am. J. Applied Sci.*, 7: 486-492. DOI: 10.3844/ajassp.2010.486.492
- Maalla, A., C. Wei and H.J. Taha, 2009. Optimal power multicast problem in wireless mesh networks by using a hybrid particle swarm optimization. *Am. J. Applied Sci.*, 6: 1758-1762. DOI: 10.3844/ajassp.2009.1758.1762
- Moghaddasi, R., A. Bakhshi and M.D. Kakhki, 2009. Analyzing the effects of water and agriculture policy strategies: An Iranian experience. *Am. J. Agric. Biol. Sci.*, 4: 206-214. DOI: 10.3844/ajabssp.2009.206.214

- Nazif, H. and L.S. Lee, 2010. Optimized crossover genetic algorithm for vehicle routing problem with time windows. *Am. J. Applied Sci.*, 7: 95-101. DOI: 10.3844/ajassp.2010.95.101
- Odeh, S., R. Faqeh, L.A. Eid and N. Shamasneh, 2009. Vision-based obstacle avoidance of mobile robot using quantized spatial model. *Am. J. Eng. Applied Sci.*, 2: 611-619. DOI: 10.3844/ajeassp.2009.611.619
- Saha, S. and S. Bandyopadhyay, 2010. A symmetry based multiobjective clustering technique for automatic evolution of clusters. *Patt. Recog.*, 43: 738-751. DOI: 10.1016/j.patcog.2009.07.004
- Sayed, S., N.S. Kamel and R. Besar, 2009. A novel approach to dynamic signature verification using sensor-based data glove. *Am. J. Applied Sci.*, 6: 233-240. DOI: 10.3844/ajassp.2009.233.240
- Sheng, X. and P. Qi-Cong, 2009. A view-based approach to three-dimensional object perception. *Inform. Technol. J.*, 8: 1189-1196. DOI: 10.3923/itj.2009.1189.1196
- Shirakawa, S. and T. Nagao, 2009. Evolutionary image segmentation based on multiobjective clustering. *Proceedings of the IEEE Congress on Evolutionary Computation*, May 18-21, IEEE Xplore Press, Trondheim, Norway, pp: 2466-2473. DOI: 10.1109/CEC.2009.4983250
- Sleit, A., W. AlMobaideen, M. Qatawneh and H. Saadeh, 2009. Efficient processing for binary submatrix matching. *Am. J. Applied Sci.*, 6: 78-88. DOI: 10.3844/ajassp.2009.78.88
- Taghouti, H. and A. Mami, 2010. How to find wave-scattering parameters from the causal bond graph model of a high frequency filter. *Am. J. Applied Sci.*, 7: 702-710. DOI: 10.3844/ajassp.2010.702.710
- Thilagamani, S. and N. Shanthi, 2011a. A novel recursive clustering algorithm for image over segmentation. *Eur. J. Sci. Res.*, 52: 430-436.
- Thilagamani, S. and N. Shanthi, 2011b. A survey on image segmentation through clustering. *Int. J. Res. Rev. Inform. Sci.*, 1: 14-17.