

An Evolutionary and Graph-Based Method for Image Segmentation

Alessia Amelio^{1,2} and Clara Pizzuti¹

¹ Institute for High Performance Computing and Networking,
National Research Council of Italy, CNR-ICAR,
Via P. Bucci 41C, 87036 Rende (CS), Italy
pizzuti@icar.cnr.it

² DEIS, Università della Calabria
Via P. Bucci 41C, 87036 Rende (CS), Italy
aamelio@deis.unical.it

Abstract. A graph-based approach for image segmentation that employs genetic algorithms is proposed. An image is modeled as a weighted undirected graph, where nodes correspond to pixels, and edges connect similar pixels. A fitness function, that extends the normalized cut criterion, is employed, and a new concept of nearest neighbor, that takes into account not only the spatial location of a pixel, but also the affinity with the other pixels contained in the neighborhood, is defined. Because of the locus-based representation of individuals, the method is able to partition images without the need to set the number of segments beforehand. As experimental results show, our approach is able to segment images in a number of regions that well adhere to the human visual perception.

1 Introduction

Image segmentation is an important problem in pattern recognition that aims at partitioning an image into uniform regions [4]. More formally, the problem can be stated as follows: let R be an image constituted by a set of pixels. Segmenting the image R consists in subdividing R into a finite number of non-overlapping and connected regions $R_1 \dots R_s$ such that

$$R = \bigcup_{i=1}^s R_i, R_i \cap R_j = \emptyset, i \neq j$$

A homogeneity measure must be defined over pixels that takes into account characteristics such as intensity, color, or texture. Pixels belonging to the same region are similar on the base of the homogeneity measure adopted, while adjacent regions are significantly dissimilar with respect to the same features.

1.1 Related work

The image segmentation problem has been intensively investigated with the use of several computational techniques, and many different methods have been proposed. A broad classification divides the existing methods in two main categories [20]: boundary

detection-based approaches and region clustering-based approaches. The former approaches search for closed boundary contours by detecting pixels that sensibly change in intensity. Boundaries of objects are obtained by linking such pixels in contours. The main limitation of these approaches is that a threshold value must be set in order to produce a continuous contour [16,8]. Region cluster-based methods group similar closed pixels into clusters. Many of these approaches use Fuzzy C-means [2] or the K-means method, such as [1,13]. A drawback of these methods is that the number of clusters must be predetermined, which implies that a user should know in advance the region number of the image to segment. In order to overcome these limitations, methods based on representing an image as a graph have been introduced. One of the earliest graph-based methods dates back over 40 years and it is based on the minimum spanning tree (MST) of a graph [21]. Zahn's method gives a weight to edges on the base of differences between pixel intensities, and breaks large edges by fixing a threshold. Improvements on the policy of edge breaking were proposed by Urquhart in [18]. Wu and Leahy [19] presented a method based on the minimization of the concept of cut, which is the weight of edges connecting two regions. To avoid unnatural cuts of small groups of isolated nodes, Shi and Malik [17] introduced a new measure of dissimilarity between two groups named *normalized cut*. More recently, Felzenszwalb and Huttenlocher [5] defined a measure of evidence of the boundary between two regions by considering both the differences of intensity across the boundary and among neighboring pixels within a region.

1.2 Evolutionary-Based Related Work

In the last years much effort has been done in the definition of effective evolutionary-based approaches for solving complex problems of computer vision. In particular, evolutionary techniques have been successfully applied to the image segmentation problem, that is of prior importance for facing more complex higher level problems such as *Object Recognition*. A survey on the application of genetic algorithms for image enhancement and segmentation can be found in [15]. Many of the approaches use a representation of the image based either on the cluster centers or on the label of the cluster a pixel is assigned to. A color image segmentation algorithm based on evolutionary approach has been proposed by Halder and Pathak in [7]. Each individual is a sequence of cluster centers and the cost function is the inverse of the sum of Euclidean distances of each point from their respective cluster centers. In order to determine the most appropriate number k of clusters, the algorithm is repeated for values of k equals to 2 until k_{max} . The choice of the best k is done by computing a cluster validity index based on inter and intra distances between clusters, while the value of k_{max} must be given as input to the algorithm. Jiao in [9] proposed an evolutionary image texture classification algorithm where the individuals are the cluster representatives, and the total distance between the pixels and the corresponding centroids is optimized. The distance for a couple of pixels is the value of the shortest path between them in the undirected weighted graph representing the image. In the same paper the author defines a Memetic Image Segmentation approach, where a genetic algorithm is applied on a set of regions previously extracted from a watershed segmentation in order to refine or merge the partitions into clusters. In this case each gene of a chromosome is the cluster label of the corresponding pixel.

The association of the regions with the clusters is evolved by optimizing the total distance between the pixels and the corresponding centroids. In the former approach the number of clusters must be fixed a priori, while in the latter an approximate initial number is obtained by using the watershed segmentation, and then a final local search procedure merges regions to obtain the optimal number of clusters. Lai and Chang [10] proposed a hierarchical structure of the chromosome, composed by control genes, representing a partitioning in regions, and parametric genes, containing the representative gray levels of each region. The goal is to optimize a fitness function that is the sum of the distances between the gray level of each pixel and the representative gray level of its region. The number of control genes, as stated by the authors, is a *soft* estimate of the upper bound of the number of regions. Merzougui et al. [12] proposed an evolutionary based image segmentation technique where the individuals are the components of the cluster centers and the fitness is the mean distance between the pixels and the centroids. In order to determine the optimal number of clusters, a criterion based on separability and compactness of the clusters is first applied. Di Gesù and Bosco [6] introduced an image segmentation algorithm where each chromosome represents the position and the region label where the pixel is located. The fitness function is defined on the similarity value and the spatial proximity between a pixel (chromosome) and the mean gray value of its corresponding region.

Because of the representation adopted, one of the main problems of the described approaches is the determination of the number of regions. Though different criteria are used to fix this number beforehand, the genetic algorithm cannot change this number while executing. The method we propose in the following dynamically computes the number of regions that optimizes the fitness function.

1.3 Contributions

In this paper we present a new graph-based algorithm, named *GeNCut* (*Genetic NCut*), to solve the image segmentation problem by using an evolutionary approach. In particular, we represent an image as a weighted undirected graph, then a genetic algorithm optimizing a fitness function is executed in order to find an optimal partitioning of the graph and, consequently, a good segmentation of the image. The fitness function is an extension of the normalized cut concept of Shi and Malik [17] that allows for a simultaneous k -way partitioning of the image without the need of fixing the number k of divisions beforehand, which is typical of many image segmentation approaches. In fact, because of the locus-based representation of individuals adopted, k is automatically determined by the optimal value of the objective function. Experiments on images of different difficulty show that *GeNCut* outperforms the method of Shi and Malik by partitioning natural and human scenes in meaningful objects.

The paper is organized as follow. In the next section the problem of image segmentation is defined together with its formalization as a graph partitioning problem and a description of the homogeneity measure adopted. Section 3 introduces the concept of normalized cut and the fitness function used by *GeNCut*. Section 4 explains the genetic representation and operators employed. Section 5 presents the experimental results. Finally, section 6 summarizes the approach presented and outlines future work.

2 Problem definition

An image R can be represented as a weighted undirected graph $G = (V, E, w)$, where V is the set of the nodes, E is the set of edges in the graph, and $w : E \rightarrow \mathcal{R}$ is a function that assigns a value to graph edges. Each node corresponds to a pixel in the image, and a graph edge (i, j) connects two pixels i and j , provided that these two pixels satisfy some property suitably defined that takes into account both pixel characteristics and spatial distance. The weight $w(i, j)$ associated with a graph edge (i, j) represents the likelihood that pixels i and j belong to the same image region and provides a similarity value between i and j . The higher the value of $w(i, j)$, the more likely the two pixels are members of the same region. Let W be the adjacency weight matrix of the graph G . Thus W_{ij} contains the weight $w(i, j)$ if the nodes i and j are connected, zero otherwise. Depending on the method adopted to compute the weights, any two pixels may or may not be connected. In our approach we employed the *Intervening Contour* method described in [11,3]. In this framework, given a generic pixel, the magnitude of the orientation energy at that pixel is considered. If the maximum image edge magnitude along a straight line connecting the two pixels i and j in the image plan is large, then a deep change and, consequently, an intervening contour is present, indicating that the two pixels don't belong to the same segment. Hence, the weight $w(i, j)$ between these pixels will be low. On the other hand, if the image edge magnitude is sufficiently weak, this usually happens in a region that is flat in brightness, the affinity between the two pixels will be very high. More formally, the weight $w(i, j)$ between the pixels i and j is computed as:

$$w(i, j) = \begin{cases} e^{-\max_{x \in \text{line}(i, j)} \|Edge(x)\|^2 / 2a^2} & \text{if } \|X(i) - X(j)\|_2 < r, i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

where $a = (\max_{y \in I} \|Edge(y)\|) \times \sigma$, $Edge(x)$ is the image edge strength at position x , I is the image plan, $\text{line}(i, j)$ is a straight line between i and j , $X(i)$ is the spatial location of the pixel i , r is a distance threshold and σ is the image edge variance. In order to compute the weight between the pixels i and j , image edges across various scales are considered.

3 Objective Function

In the last few years many different criteria have been defined to partition a graph representing an image into non-overlapping connected components. Shi and Malik [17] introduced the dissimilarity measure *normalized cut* to divide a graph into two sub-graphs, that revealed successful for image segmentation. The concept of normalized cut is an extension of the notion of *cut* proposed by Wu and Leahy [19] that avoids the bias for partitioning in small sets of nodes. Given a partition of a graph G in two disjoint sets of nodes A and B , the cut between A and B is defined as

$$\text{cut}(A, B) = \sum_{i \in A, j \in B} w(i, j)$$

In [17] the authors pointed out that the cut value diminishes when small sets of isolated nodes are generated. Thus a disassociation measure, that takes into account the total edge weight connecting two partitions, has been introduced. Let

$$assoc(A, V) = \sum_{i \in A, t \in V} w(i, t)$$

be the total connection from nodes in A to all the nodes in V , then the normalized cut is defined as

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)}$$

Shi and Malik formalize the problem of minimizing the normalized cut as a generalized eigenvalue problem and compute an optimal partition by using the eigenvector with the second smallest eigenvalue. Two extensions of the approach to k -way partitioning are also proposed. The former recursively partitions the groups obtained in the previous step by checking the values of the eigenvectors, the latter exploits the top n eigenvectors and the clustering algorithm K-means. A main limitation of this method is that the number k of desired partitions must be fixed beforehand.

We now introduce an extension of the concept of normalized cut that is used as criterion to partition a graph in a generic number k of regions. Note that, the value of k in our approach must not be fixed in advance, but it is determined by the optimal value of the objective function. Let $G = (V, E, w)$ be the graph representing an image, W its adjacency matrix, and $P = \{S_1, \dots, S_k\}$ a partition of G in k clusters.

For a generic cluster $S \in P$, let

$$c_s = \sum_{i \in S, j \notin S} W_{ij} \quad m_s = \sum_{i \in S, j \in S} W_{ij} \quad m = \sum_{i \in V, j \in V} W_{ij}$$

be respectively the sum of weights of edges on the boundary of S , the sum of weights of edges inside S , and the total graph weight sum. The *weighted normalized cut* $WNCut$ measures for each cluster $S \in P$ the fraction of total edge weight connections to all the nodes in the graph

$$WNCut = \sum_{s=1}^k \frac{c_s}{m_s + c_s} + \frac{c_s}{(m - m_s) + c_s}$$

Note that c_s corresponds to $cut(A, B)$ where $B = V - A$. Because of the affinity measure w defined in the previous section, and the relationship between cut and $assoc$ formalized in [17], more uniform regions can be obtained with low cut values between the subgraphs representing the regions and the rest of the graph. This implies that low values of $WNCut$ should be preferred.

4 Genetic Representation and Operators

The genetic algorithm uses the locus-based adjacency representation proposed in [14]. In this graph-based representation an individual of the population consists of N genes

g_1, \dots, g_N and each gene can assume allele values j in the range $\{1, \dots, N\}$. Genes and alleles represent nodes of the graph $G = (V, E, w)$ modelling an image, and a value j assigned to the i th gene is interpreted as a link between the pixels i and j . This means that in the clustering solution found i and j will belong to the same region.

The initialization process assigns to each node i one of its neighbors j . This guarantees a division of the graph in connected groups of nodes. The kind of crossover operator adopted is uniform crossover. Given two parents, a random binary vector is created. Uniform crossover then selects the genes where the vector is a 0 from the first parent, and the genes where the vector is a 1 from the second parent, and combines the genes to form the child. The mutation operator, analogously to the initialization process, randomly assigns to each node i one of its neighbors.

The genetic operators need to determine the neighbors of each node. In our approach we introduced the concept of neighbors of a node by taking into account not only the spatial closeness, but also the pixel affinity. More in details, given a generic node i in the graph, let $w_{max}^h = \{w^1, \dots, w^h \mid w^1 \geq, \dots, \geq w^h\}$ be the first h highest weights of row i in the weight adjacency matrix W .

The h nearest neighbors of i , denoted as nn_i^h , are then defined as $nn_i^h = \{j \mid w(i, j) \in w_{max}^h\}$.

nn_i^h is thus the set of those pixels that are no more than r pixels apart from i , and that have maximum similarity with i . It is worth to note that, even if h is fixed to 1, the number of nearest neighbors of i could be sufficiently large if many of its spatial neighbors have the same maximum weight w_{max}^h . This definition of nearest neighbors guarantees to choose the most similar neighbors during the initialization process, and to bias the effects of the mutation operator towards the most similar neighbors, thus it contributes to improve the results of the method.

5 Experimental Results

In this section we present the results of *GeNCut* on five images with details of increasing complexity, and compare the performances of our algorithm in partitioning natural and human scenes in meaningful objects with the segmentations obtained by the algorithm of Shi and Malik [17] (in the following referred as *NCut*) on the same images. The *GeNCut* algorithm has been written in MATLAB 7.14 R2012a, using the Genetic Algorithms and Direct Search Toolbox 2. In order to set parameter values, a trial and error procedure has been employed and then the parameter values giving good results for the benchmark images have been selected. Thus we set crossover rate to 0.9, mutation rate to 0.2, elite reproduction 10% of the population size, roulette selection function. The population size was 100, the number of generations 50. The value h of nearest neighbors to consider has been fixed to either 1 or 2. As already pointed out, this does not mean that the number of neighbors is 1 or 2, but that the first (and second) most similar neighbors are taken into account for the initialization and mutation operators. The fitness function, however, is computed on the overall weight matrix. For all the data sets, the statistical significance of the results produced by *GeNCut* has been checked by performing a t-test at the 5% significance level. The p-values returned are very small, thus the significance level is very high since the probability that

a segmentation computed by *GeNCut* could be obtained by chance is very low. The version of the *NCut* software that we used is written in MATLAB and it is available at <http://www.cis.upenn.edu/~jshi/software/>. The weight matrix of each image is the same for both methods, and, as already described in section 2, it is based on the Intervening Contour framework by fixing $r = 10$, number of scales 3, number of orientations 4 and $\sigma = 0.1$. Since *NCut* needs the number k of clusters, we executed the algorithm by using two different inputs. The first sets the number k of segments to the same number of clusters found by *GeNCut*, the second one is a higher value. In the following, for each image, we compare the segmentation results of *GeNCut* and *NCut* by depicting the contours of the regions obtained by the two approaches. For a more clear visualization, we show two images. The first image reports the boundary lines of the segmentation obtained on the original color image, the second one delineates the same contours without the image background.

Fig. 1 shows the execution of *GeNCut* and *NCut* on an image of a melanoma. In particular, *Fig. 1(a)* is the original image, *Fig. 1(b)* and *Fig. 1(c)* display the segmentation obtained by *GeNCut* with and without the background image resp., while *Fig. 1(d)* and *Fig. 1(e)* are the results of *NCut* when the number of segments is fixed to two, and *Fig. 6(a)* when this number is 5. *Fig. 1(b)(c)* show that *GeNCut* is able to find the right partitioning and correctly discriminates between the melanoma and the skin, although we don't set a priori the number of segments. *Fig. 1(d)(e)* and *Fig. 6(a)* point out that if *NCut* receives the true number of segments, it is able to find the correct partitioning, otherwise an over-segmentation is obtained. In general, however, given an input image, it is hard to know a priori the true number of partitions.

The next experiment represents a more complex scenario, due to the presence of irregular shapes (clouds) around a spherical object (moon) (*Fig. 2(a)*). *Fig. 2(b)(c)* illustrate

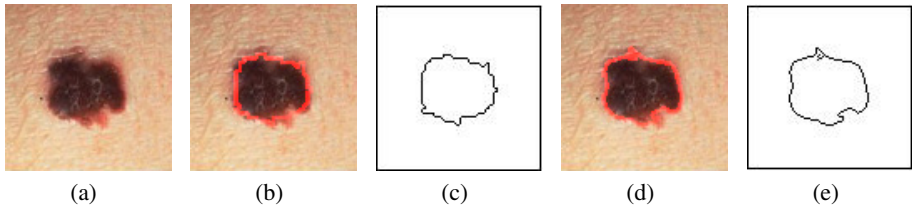


Fig. 1. (a) The original image representing a melanoma, (b) the segmentation result on the original image of *GeNCut* with $h=2$, (d) *NCut* with $k = 2$ and (c-e) the corresponding contours

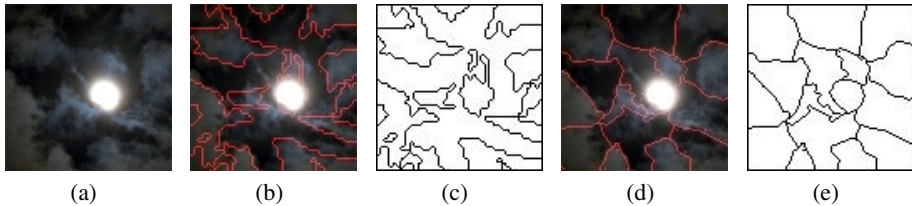


Fig. 2. (a) The original moon image, (b) the segmentation results on the original image using *GeNCut* with $h=1$, (d) *NCut* with $k = 16$ and (c-e) the corresponding contours

the results obtained by using the *GeNCut* approach, while *Fig. 2(d)(e)* and *Fig. 6(b)* the *NCut* method when the number of segments is set to 16 and 22, respectively. Although the halo makes difficult to segment the moon, by using our algorithm we are able to perform a segmentation that is more flexible in capturing the real shape of the clouds. *NCut*, instead, realizes a more flat partitioning with an equal number of segments that is not able to distinguish and capture some inner parts of the original image. *Fig. 3(a)* and *Fig. 4(a)* show two different kinds of landscapes: a natural picture and a snatch from an X-SAR image of the Vesuvius volcano (Italy), acquired by the Spaceborne Imaging Radar-C/X-Band Synthetic Aperture Radar (SIR- C/X-SAR) aboard the Space Shuttle Endeavour in 1994. For both the images, our algorithm is able to discover the meaningful objects, *Fig. 3(b)(c)* and *Fig. 4(b)(c)*, respectively, while a poor segmentation of the major components like in *Fig. 3(d)(e)* and *Fig. 6(c)* is obtained, despite the setting of the same number of segments, naturally extracted from our technique. The satellite image in *Fig. 4(a)* is a scene where it is quite difficult to differentiate the meaningful objects due to the details of the terrain. However, as it can be observed in *Fig. 4(b)(c)*, *GeNCut* is able to separate the volcano area from the landscape and to distinguish the building barely visible at the bottom right corner and the area of the deep sea. All these significative features are not visible in the segmentation results of the *NCut* approach, *Fig. 4(d)(e)*, even if we increase the number of partitions, *Fig. 6(d)*. Finally, we used *GeNCut* to segment a human face image (*Fig. 5(a)*). In this case the two approaches are comparable. Although more details are discovered from *NCut* in correspondence of the eyes, it over-segments the face (*Fig. 5(d)(e)* and *Fig. 6(e)*). On the other hand, *GeNCut* obtains a uniform and natural segmentation of the face (*Fig. 5(b)(c)*) that is able to capture also the shape of the nose, although it appears linked to the eyes, probably due to the similar gray intensities along the contours of the nose and the contours of the eyes.

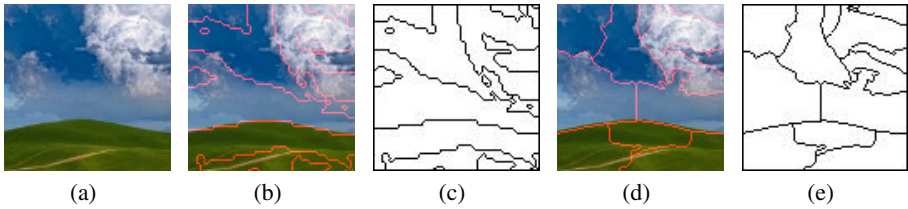


Fig. 3. (a) The original image, (b) the segmentation results on the original image using *GeNCut* with $h=1$, (d) *NCut* with $k = 12$ and (c-e) the corresponding contours



Fig. 4. (a) An X-SAR image of the Vesuvius volcano, (b) the segmentation results on the original image using *GeNCut* with $h=1$, (d) *NCut* with $k = 8$, and (c-e) the corresponding contours

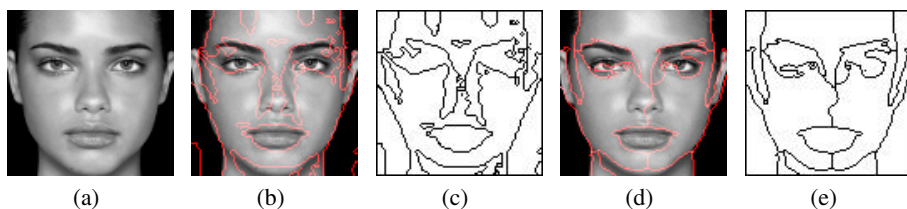


Fig. 5. (a) The original face image, (b) the segmentation results on the original image using *GeN-Cut* with $h=2$, (d) *NCut* with $k=12$ and (c-e) the corresponding contours

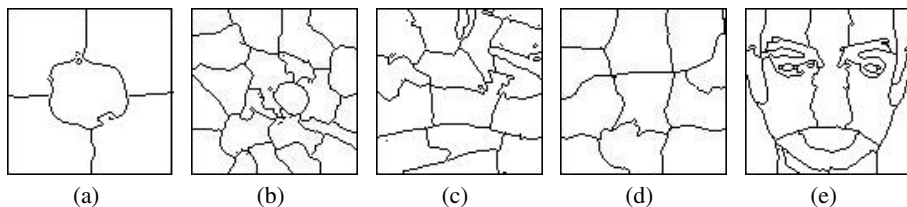


Fig. 6. The segmentation results representing the contours using *NCut* with (a) $k=5$, (b) $k=22$, (c) $k=20$, (d) $k=12$, (e) $k=20$

6 Conclusions

The paper presented a graph-based approach to image segmentation that employs genetic algorithms. A fitness function, that extends the normalized cut criterion introduced in [17], is proposed, and a new concept of nearest neighbor, that takes into account not only the spatial location of a pixel, but also the affinity with the other pixels contained in the neighborhood, is defined. The locus-based representation of individuals, together with the fitness function adopted, revealed particularly apt to deal with images modeled as graphs. In fact, as experimental results showed, our approach is able to segment images in a number of regions that well adhere to the human visual perception. Future work will extend the segmentation process to partition more complex texture images including medical X-ray and ultrasound images, by combining the contour information considered in our approach with textural features of the regions.

Acknowledgements. This work has been partially supported by the project *MERIT : MEDical Research in Italy*, funded by MIUR.

References

1. Chen, C.W., Luo, J., Parker, K.J.: Image segmentation via adaptive k-means clustering and knowledge-based morphological operations with biomedical applications. *IEEE Transactions on Image Processing* 7(12), 1673–1683 (1998)
2. Chen, S., Zhang, K.: Robust image segmentation using fcm with spatial constraints based on a new kernel-induced distance measure. *IEEE Transactions on Systems Man and Cybernetics B* 34, 1907–1916 (2004)

3. Cour, T., Bénézit, F., Shi, J.: Spectral segmentation with multiscale graph decomposition. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)), pp. 1124–1131 (2005)
4. Duda, R.O., Hart, P.E.: Pattern Classification and Scene Analysis. John Wiley & Sons, New York (1973)
5. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. *International Journal of Computer Vision* 59(2), 167–181 (2004)
6. Di Gesù, V., Lo Bosco, G.: Image Segmentation Based on Genetic Algorithms Combination. In: Roli, F., Vitulano, S. (eds.) *ICIAP 2005*. LNCS, vol. 3617, pp. 352–359. Springer, Heidelberg (2005)
7. Halder, A., Pathak, N.: An evolutionary dynamic clustering based colour image segmentation. *International Journal of Image Processing* 4, 549–556 (2011)
8. Helderbrand, J.D.: One pixel-wide closed boundary identification. *IEEE Transactions on Image Processing* 5(5), 780–783 (1996)
9. Jiao, L.: Evolutionary-based image segmentation methods. *Image Segmentation* (10), 180–224 (2011)
10. Lai, C.-C., Chang, C.-Y.: A hierarchical evolutionary algorithm for automatic medical image segmentation. *Expert Syst. Appl.* 36(1), 248–259 (2009)
11. Leung, T., Malik, J.: Contour Continuity in Region Based Image Segmentation. In: Burkhart, H.-J., Neumann, B. (eds.) *ECCV 1998*. LNCS, vol. 1406, pp. 544–559. Springer, Heidelberg (1998)
12. Merzougui, M., Allaoui, A.E., Nasri, M., Hitmy, M.E., Ouariachi, H.: Evolutionary image segmentation by pixel classification and the evolutionary Xie and Beni criterion - application to quality control. *International Journal of Computational Intelligence and Information Security* 2(8), 4–13 (2011)
13. Pappas, T.N.: An adaptive clustering algorithms for image segmentation. *IEEE Transactions on Signal Processing* 40(4), 901–914 (1992)
14. Park, Y.J., Song, M.S.: A genetic algorithm for clustering problems. In: *Proc. of 3rd Annual Conference on Genetic Algorithms*, pp. 2–9 (1989)
15. Paulinas, M., Uinskas, A.: A survey of genetic algorithms applications for image enhancement and segmentation. *Information Technology And Control, Kaunas, Technologija* 36(3), 278–284 (2007)
16. Sahoo, P.K., Soltani, S., Wong, A.K.C.: A survey of thresholding techniques. *CVGIP* 41, 233–260 (1988)
17. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(8), 888–905 (2000)
18. Urquhart, R.: Graph theoretical clustering based on limited neighborhood sets. *Pattern Recognition* 15(3), 173–187 (1982)
19. Wu, Z., Leahy, R.: An optimal graph theoretic approach to data clustering: Theory and applications to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(11), 1101–1113 (1993)
20. Xu, Y., Oltman, V., Uberbacher, E.C.: A segmentation algorithm for noisy images: Design and evaluation. *Pattern Recognition Letters* 19, 1213–1224 (1998)
21. Zahn, C.T.: Graph theoretical methods for detecting and describing gestalt clusters. *IEEE Transactions on Computers* 20(1), 68–86 (1971)