Evolutionary Image Segmentation

Primo Zingaretti, Antonella Carbonaro, Paolo Puliti

Istituto di Informatica, Facoltà di Ingegneria, Università di Ancona
Via Brecce Bianche, I-60131 Ancona (Italy)
e-mail: {zinga, carbonar, puliti}@inform.unian.it

Abstract - We describe an approach to image segmentation based on a two-layer module that is executed until a good segmentation is achieved, providing an evolution of previous segmentation results at each execution. The first layer performs a global segmentation of an image of decreasing area at each evolution by adopting a genetic algorithm learning technique to select segmentation parameters that give better results. The second layer provides the input to the next evolution by selecting the segmented regions that need further optimisation. A main goal of our system is to perform the segmentation without using neither ground-truth information nor human judgement. Thus, edge detection is performed to assess the performance of region segmentation and to guide the evolution of segmentation. Experimental results are consistent with what is observed visually.

1. Introduction

Image segmentation is the process of classifying an image into a set of disjoint regions whose characteristics such as intensity, colour, texture, etc. are similar. Image segmentation is a very important process because it is typically the first task of any automatic image understanding process, and all subsequent steps, such as feature extraction and classification, object detection and recognition, depend heavily on its results. The segmentation process is a complicated task because most of available segmentation techniques use numerous control parameters that must be adjusted to obtain an optimal performance. In practice, we do not know general automatic systems whose parameters provide optimal performance for all images; even the choice of an ad hoc setting based on the actual algorithm or *a priori* domain knowledge may result in an unsatisfactory segmentation. Many different segmentation algorithms have been developed [6, 9]. Most of them are based on thresholding techniques, which usually are further classified in global, local, region or edge based techniques. A performance evaluation of the different algorithms proposed for a given application is inherently difficult to obtain, due to the lack of appropriate measures for judging the quality of segmentation results. None of the segmentation quality measures suggested in the literature [6, 14] has achieved widespread acceptance as a universal measure. At present most segmentation results are evaluated visually and qualitatively.

Image segmentation is formulated in this paper as an optimisation problem. A genetic algorithm (GA) [4] is used to efficiently search, in the hyperspace of segmentation parameters, the set that maximises the segmentation quality criteria. GAs have been recently applied in machine vision problems (for example, image

segmentation [2, 3], edge detection [1] and target recognition [8]) mainly because they can locate an approximate global maximum in a search space in a way not dependent on the particular application domain and without using detailed knowledge about the processing technique. The goal of our system is not to build a new segmentation algorithm for a certain image class, but, rather, to develop a most general possible technique that uses neither ground-truth information nor human judgement in the evaluation of segmentation results.

The paper is organised as follows. Section 2 describes the evolutionary image segmentation process. Experimental results are presented in Section 3, while some conclusions and future research plans in Section 4.

2. Image Segmentation Using a Genetic Algorithm

In this Section we describe the architecture and some implementation aspects of the developed system.

2.1. System architecture

The proposed segmentation system is based on a module consisting of two layers: *global segmentation layer* and *single-region evaluation layer*. The module is executed until a good segmentation is achieved, providing an evolution of previous segmentation results at each execution. Our strategy is based on the hypothesis that, by modifying processing parameters, it is possible to obtain a good segmentation at least for some regions. This strategy extends the space search, but the difficulty in finding a good solution among many alternatives should be overcome by the GA. Fig. 1 shows the global architecture and the information flow for the approach described in this paper.

The global segmentation layer adopts a genetic learning technique to produce a new classification of pixels into regions. The GA analyses the characteristics of the input image and uses this information to select an appropriate parameter combination to segment the image. A fitness is then evaluated and the GA cycles until a segmentation result of acceptable quality is reached, that is, the stopping criterion of the GA is based on a fitness threshold T_l that is independent from the adopted processing parameters. In this way, the choice of the number of thresholds (over-under segmentation) is not a problem any more.

The single-region evaluation layer is mainly based on the local evaluation of fitness criteria for each region resulting from the global segmentation. The aim of this layer is to select those regions that need further optimisation. The set of all these regions will constitute the object-of-interest image (OOI) for the next evolution.

The global segmentation layer performs a segmentation of the whole image at the first application of the module and of the generated OOIs at the successive evolutions. First a subset of individuals (seed population, SP) that has image features more similar to those ones of currently processed OOI is extracted from a long term population (LTP) characterised by N individuals (strings). The similarity measure is
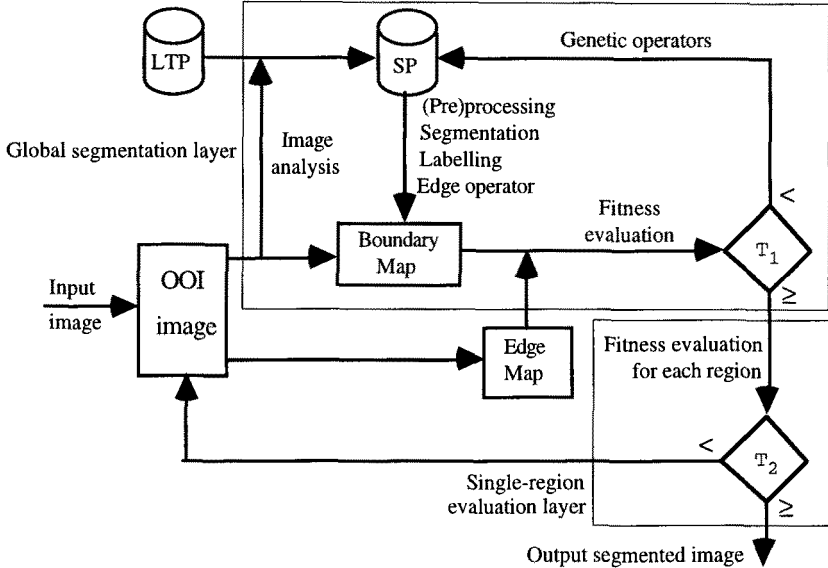
Fig. 1. System architecture.

based on the Euclidean distance normalised with respect to the range of actual feature values, and weighted on the importance of each feature, that is:

$$dist_{ab} = \frac{1}{NF} \sum_{i=1}^{NF} w_i \left| \frac{F_{ia} - F_{ib}}{\max(F_i) - \min(F_i)} \right|$$

where: $dist_{ab}$ is the similarity measure between the two individuals a and b; NF is the number of features represented in the image feature part of the string; w_i is the weight assigned to each feature; F_{ix} is the value of feature i for the individual x; $\max(F_i)$ and $\min(F_i)$ are maximum and minimum values of feature i in the current global population. If the features have the same importance, as in the present application, $w_i = 1$ for all i. Then, image segmentation is performed using the processing parameters of the individuals extracted from the seed population and the genetic operators are applied on the processing parameters of these individuals until the global fitness does not exceed T_1 (see Fig. 1).

Our segmentation system is based on the multithresholding of the intensity histogram [12]. Histogram-based thresholding is computationally simpler than other existing algorithms and does not need a priori information of the image, but using global features it may be not optimal for all objects of the input image. The parameters of the global segmentation layer include the (pre)processing of the input image, which may be based on global features, such as histogram transformations, or may concern local operations, such as spatial filtering and noise removal. Both these operations should be considered as "global processing" because the processing is applied to the whole OOI, even if it enhances some local features of the input image. A real "local processing" is introduced by means of successive evolutions. In fact, at

each evolution the second layer should exclude from the input OOI those regions with a fitness value greater than a given threshold T_2. Then in the next evolution the global processing will be performed on the histogram of a smaller OOI, that is, in a more local manner. In this sense we say that our approach uses global techniques in a local context, and then it is able to adapt to varying image characteristics.

The aim of this layer is to reduce the total computational cost of the system by decreasing the number of fitness evaluations, which are computational expensive. As our approach requires the segmentation evaluation of each region of the segmented image, the splitting of the system in two layers avoids to compute a high number of fitness evaluations at each segmentation performed by the GA. In fact, for each segmentation performed by the GA in the first layer, the fitness is evaluated only one time by considering all the regions as a whole, while in the second layer the fitness is computed for each region resulting from the global segmentation. Thus, in our system the total number of fitness computations is given by: $N + R$, where N is the number of segmentations performed by the GA and R is the number of regions in the final segmented image. On the contrary, a system with a unique layer should be compelled to evaluate about $N*R$ fitness evaluations, hypothesising that the regions after each segmentation are R on the average.

2.2. System implementation

In this Section we describe in detail the three parts that constitute the knowledge structure of each individual of the GA population: image features (used to discriminate the images presented to the system), segmentation parameters (processing alternatives) and fitness (used to quantitatively evaluate the segmentation effectiveness).

2.2.1 Image features
The system uses two categories of features: global features of the image and features of the image histogram. We have considered as global features the normalised abscissa and ordinate of the intensity centroid, the ratio between the inertial major axes, the angles of the inertial major axes, and the seven invariant moments proposed by Hu [7]. From the image histogram of grey level intensities we have computed the number of peaks, the grey levels and the normalised frequencies of the maximum peak and the minimum valley, the distance between the two valleys with the greater number of pixels and the normalised maximum frequency of these regions, the amplitude of the image spectrum. In addition, we have considered the mean, standard deviation, skewness, kurtosis, and entropy of the histogram to measure its shape.

2.2.2. Segmentation parameters
The segmentation parameter part of each individual of the GA population is constituted by: (pre)processing operations (such as averaging, sharpening and median filtering, noise removal and introduction of local characteristics in the image intensity histogram by using gradient and laplacian operators [11, 13]), number of histogram averaging (to modify the number of thresholds by which the image will be segmented), object size filter (to reduce the computational cost of the labelling process

by excluding small regions), and coefficient of small object removal (to determine the number of erosions and dilations performed to fill small holes inside regions).

2.2.3 Fitness

We perform edge detection to assess the performance of region segmentation and then to guide the evolution of segmentation. In this sense the contribution of our paper may be viewed as a new method for the integration of region-based and edge-based segmentation techniques [5, 10]. In this context, segmentation results suffer from three kinds of errors: 1) *false boundary* - a region boundary is not an edge and there are no edges nearby (for example, if our uniformity criterion is to keep intensity approximately constant over a given area and the light intensity varies linearly within a region then artificial boundaries need to be set even if there is no clear line where a transition occurs); 2) *imprecise boundary* - a region boundary does not coincide with an edge (for example, highlights may displace the real region boundary, at least in some of its parts); 3) *missing boundary* - there exist edges in the image with no region boundaries near them. It should be pointed out that if we reduce errors of the last kind by performing an over-segmentation (for example, by allowing small valleys in the histogram or by gathering thresholds according to different criteria) the probability of errors of the first kind will increase.

We have chosen as fitness criterion the correspondence between boundaries of regions resulting from segmentation (*boundary map*) and edges obtained from the Roberts edge operator (*edge map*). This correspondence is difficult to find out in real images even in presence of a good segmentation, so we have introduced a matching tolerance, dependent on image and region size, to overcome errors in the extraction of significative edges. The achievement of a good segmentation is represented by a fitness higher than T_2. For the efficiency of our two-layer approach the threshold T_1, adopted in the global segmentation layer, should be high enough to permit the existence of at least one region exceeding T_2, but also not too high to avoid a great number of GA iterations. In any case $T_1 < T_2$.

In particular, we compute two types of correspondences: F_1, an exact correspondence between boundaries and edges, and F_2, a fuzzy correspondence. In particular, imprecise boundaries occur whenever F_1 is low and F_2 high. These values are computed using the following two formula:

$$F_1 = \frac{P_1}{\min(N_1, N_2)} \qquad F_2 = \left(P_1 + \sum_{i=0}^{d-1} \frac{hist(i) \cdot (d-i)}{d} - \frac{E_1 + E_2}{2} \right) \bigg/ N_1$$

where: P_1 is the number of pixels resulting from a logical *'and'* between boundary and edge maps; N_1 is the total number of pixels in the boundary map; N_2 is the total number of pixels in the image obtained from the binarization of the edge map at the threshold that minimises the difference between N_1 and the pixels resulting from the binarization itself; d is the tolerance in the matching distance; $hist(i)$ is the number of boundary pixels that have a distance i from the nearest edge pixel in the corresponding edge map; E_1 is the number of boundary pixels that have a distance greater than d from any edge pixel in the corresponding edge map; E_2 is the number of edge pixels that have a distance greater than d from any boundary pixel in the corresponding boundary map.

The fitness function, F, also used in the single-region evaluation layer, is obtained by weighting in the same way the two correspondence measures:

$$F = \frac{1}{2}F_1 + \frac{1}{2}F_2$$

A final global fitness F_{tot} is computed as a weighted average of the areas and the fitness values of each resulting region for the objective evaluation of the output segmented image:

$$F_{tot} = \sum_i F_i \cdot \frac{area_i}{area}$$

3. Experimental Results

The work is in progress. Further (pre)processing techniques need to be coded in the segmentation parameter part of an individual and many other images need to be processed to construct an efficient LTP. Due to these limitations the convergence of the GA in some images is not so quickly as it should be.

The system has been tested on real images using an LTP consisting of 100 individuals characterised by processing parameters that yielded visually good segmentation results on a set of sample images. As a processing example, elaborations performed on the standard image "camera", shown in Fig. 2, are described in the following. The processing parameters adopted by the first evolution of the global segmentation layer, consisting of 5 iterations of the GA, were: a noise removal coefficient of 1 as pre-processing operation, 40 histogram averagings, and a minimum object size of 50 pixels. The selected thresholds yielded the 38 regions shown in Fig. 3. The global fitness was 0.806 ($F_1 = 0.708$ and $F_2 = 0.904$, $T_1 = 0.750$). The single-region evaluation layer extracted the coat-region, white region R reproduced in Fig. 4, as the only one region with a fitness value ($F = 0.907$) greater than T_2. Normally the system continues by considering the complement $\overline{R}$ of R as the OOI for the next evolution. In this case, as R constituted a large part of the initial image, two separate processings of the evolutionary system were executed for each of the two OOIs, R and $\overline{R}$, and a combination (add operation) of the results was performed at the end. After other 9 evolutions for the processing of R (in order they were extracted: tower and near buildings, sky, leftmost grassland, rightmost grassland, etc.) and only one evolution for $\overline{R}$, the final segmented image shown in Fig. 5 was obtained. This image is constituted of 39 regions plus a set of unclassified small regions (black holes representing only 4.38% of the whole image). The global fitness for the objective evaluation of the segmentation result was $F_{tot} = 0.829$. Comparisons with other segmentation algorithms, as above mentioned, are difficult to perform. Anyway, our final segmentation result was an improvement with respect to the global segmentation of Fig. 3, at least from a visual comparison.

The computational cost of the GA, synthesised by the number of evolutions and of iterations for each evolution, and the segmentation results, in terms of number of regions, percentage of unclassified regions (holes) and global fitness, have been analysed also for other standard grey level images. Different behaviours for these

Fig. 2. The grey level standard image "camera" 256x256x8 pixels.

Fig. 3. The image obtained from the global segmentation layer.

Fig. 4. The OOIs generated by the first evolution of the system.

Fig. 5. Final result of the segmentation process.

parameters have been noticed. In particular, some images required a great number of iterations of the first layer and many holes resulted in some other segmentations. In both these cases better results were obtained by lowering the T_1 threshold. The T_2 threshold is also a critical factor. While a low T_2 impairs the segmentation results, a high T_2 heavily increments the computational cost because in many evolutions no region resulting from the first layer can be considered as well segmented, that is, no real evolution occurs. Finally, in other images, for example in the standard "Lena" image, we have to introduce a compactness factor in the fitness function to avoid the inclusion of regions of high complexity that overcome the object size filter.

4. Conclusions

In this work we have described an automatic two-layer approach to image

segmentation that, using an adaptive methodology, globally searches for a good segmentation, and then performs subsequent local segmentation refinements without using neither ground-truth information nor human judgement. We have outlined a GA-based system using many of standard segmentation techniques to select optimal image processing parameters.

Experimental results exhibit a promising performance for the proposed method. In fact, the final segmented images are consistent with what is observed visually. Great improvements appear to be possible without requiring substantial modifications of the structure. The integration of other (pre)processing alternatives, such as histogram modification techniques, need only to be included in the parameter part of the GA string, apart from their implementation. In addition, future works will focus on other evaluation techniques, for example by introducing perceptually motivated features, and on improvements regarding refinements of extracted regions.

References

[1] S.M.Bhandarkar, Y.Zhang, W.D.Potter, An Edge Detection Technique Using GA-Based Optimisation, *Pattern Recognition*, 27(9), 1159-1180, (1994).

[2] B.Bhanu, S.Lee, J.Ming, Adaptive Image Segmentation using a genetic algorithm, *IEEE Trans. on SMC*, 25(12), 1543-1567, (1995).

[3] D.N.Chun, H.S.Yang, Robust image segmentation using genetic algorithm with a fuzzy measure, *Pattern Recognition*, 29(7), 1195-1211, (1996).

[4] D.E.Goldberg, *Genetic Algorithms in Search, Optimisation, and Machine Learning*, Addison-Wesley, Reading, MA, (1989).

[5] J.F.Haddon and J.F.Boyce, Image Segmentation by unifying region and boundary information, *IEEE Trans. on PAMI*, 12(10); 929-948, (1990).

[6] R.M.Haralick, L.G.Shapiro, Image segmentation techniques, *Computer Vision, Graphics, and Image Processing*, 29, 100-132, (1985).

[7] M.K.Hu, Visual problem recognition by moment invariants, *IRE Trans. Inf. Theory*, 8, 179-187, (1962).

[8] A. J. Katz, P. R. Thrift, Generating Image Filters for Target Recognition by Genetic Learning, *IEEE Trans. on PAMI*, 16(9); 906-910, (1994).

[9] N.R.Pal and S.K.Pal, A review on image segmentation techniques, *Pattern Recognition*, 26, 1277-1294, (1993).

[10] J.T.Pavlidis and Y.T.Liow, Integrating region growing and edge detection, *IEEE Trans. on PAMI*, 12(3); 225-233, (1990).

[11] G.Tascini, P.Puliti, P.Zingaretti, Region Detection in grey-level images, in *"Progress in Image Analysis and Processing"*, Cantoni, Cordella, Levialdi, Sanniti di Baja Eds., World Scientific Publ. Co., Singapore, 106-110, (1990).

[12] S.Wang and R.M.Haralick, Automatic multithreshold selection, *Computer Vision, Graphics, and Image Processing*, 25, 46-67, (1984).

[13] J.S. Weszka and A.Rosenfeld, Histogram modification for threshold selection, *IEEE Trans. on SMC*, 9(1), 38-52, (1979).

[14] Y.J.Zhang, A survey on evaluation methods for image segmentation, *Pattern Recognition*, 29(8), 1335-1346, (1996).