FUZZY CLUSTERING OF DIGITAL IMAGES BY EXPLOITING DENSITOMETRIC AND TOPOLOGICAL INFORMATION

M. Mari, C. Garcia, and S. Dellepiane
Department of Biophysical and Electronic Engineering (DIBE)
University of Genoa
via Opera Pia, 11a, 16145 Genova, Italy
Tel. +39 10 3532754; fax: +39 10 3532134
e-mail: silvana@dibe.unige.it

ABSTRACT

Topological features are very seldom exploited in image processing, also due to the complexity of their extraction. When topological features are used, densitometric information are usually not considered at the same time. The simultaneous exploitation of both kinds of features, as proposed in this paper, allows a more appropriate automatic processing of digital images. A novel image segmentation approach is presented (based on fuzzy clustering) that exploits topological and densitometric image features. The novelty of such image segmentation consists mainly in using easy and fast computation methods to improve the handling of any digital image, whenever automatic segmentation or data reduction processing is required.

1 INTRODUCTION

Considering the success of intensity-based and contrast-based methods for image processing and segmentation, as well as the success of the most advanced techniques that fully exploit contextual information [1], novel methods have been devised, based on topological information and on uncertainty-handling approaches.

Morphological Watersheds [2] is quite a new and promising approach to segmentation of real images; it mainly uses topological image characteristics, focusing on the geometrical and topological aspects of image regions. This approach performs a topographical (geometrical) interpretation of grey levels, usually applied to the image gradient. Therefore, it can be useful for image segmentation but, in its most basic form, it tends to lose fine structures.

Moreover, fuzzy methods for the management of uncertainties in data and models have been extensively applied at the numerical level of image processing [3].

More recently, a modified definition of fuzzy connectedness has been proposed [4] that has proved very useful and effective in the segmentation of a single object, starting from a seed point [5]. The so-called *intensity connectedness* or *χ-connectedness* measure, which integrates densitometric and topological information, allows one to extract an object area through the extraction of the χ-connected component associated with a given seed point.

This paper is a preliminary step toward a novel segmentation of a digital image through the description of its connected components. Starting from the approach presented in [4], which relates topological and densitometric properties with respect to only one seed in an image, an improved version aimed at the processing and segmentation of a whole image is proposed.

A multiseed approach is presented that allows one to describe a whole image by considering the most significant connected components.

Applications to real images with different characteristics demonstrate the main properties of the proposed multiseed approach.

2 FUZZY CLUSTERING

The segmentation problem is stated more rigorously, as the definition of a segment to be extracted coincides with the definition of a fuzzy χ-connected component. The detection of the most significant connected components in an image is then further reduced to the problem of finding the most significant seed points, and the final result consists in a fuzzy clustering of the image.

For an automatic fuzzy clustering, a fuzzy partition [8] of the image is obtained as the final processing result, where the clusters are defined by the automatically selected seeds.

This problem can be compared with the classical problem of seed selection in image segmentation, long investigated [6] but with only partial results. The properties of χ–connectedness make this measure a very useful tool for the search for an appropriate set of seeds. At the same time, it is very helpful in avoiding undesired effects due to noise, or in neglecting insignificant details that are filtered out when their topological and densitometric relationships are poor, as compared with significant image structures.

2.1 Intensity Connectedness

In a more formal way, as described in [4], an intensity connectedness map, $C_{\chi^a} = \{c_{\chi^a}(p)\}$, is easily extracted from an image field, $I = \{i(p)\}$, defined for each point p belonging to the lattice, at level i. With reference to a given seed point a, the χ*–field* χ^a is obtained, that is a fuzzy field defined (on an 8-bit image) as:

$$\forall p \in I, \qquad \chi^a(p) = 1 - \left| \frac{i(p)}{255} - \frac{i(a)}{255} \right|.$$

The intensity connectedness map represents the membership of each image point in the object pointed by the seed, according to its topological and densitometric relationships. Some interesting properties characterize the map [4], which is also a connected image, as defined in [7]. In addition, the searched object is associated with the fuzzy connected component Γ^a, defined as the fuzzy set:

$$\forall p \in I, \qquad \Gamma^a(p) = \max_{paths(p,a)} \left[\min_{z \in path(p,a)} \chi^a(z) \right],$$

where the maximum is applied to all the possible paths linking the point p to the seed a, and the minimum takes the minimum value along that path.

χ-connectedness, like original fuzzy-connectedness, is a nonlinear topological measure. Its value is a function of the local field behaviour as well as of a past memory, as connectivity can never increase along a path.

2.2 Multiseed Approach

As a consequence, for the segmentation of a whole image, we can also exploit the properties of the intensity connectedness measure and define a fuzzy *Partition Image PI* . Its elements p have a fuzzy membership $\mu(p)$ in the object identified by the seed a_i, expressed by

$$\mu(p) = \bigcup_i \left\{ c_{\chi^{ai}}(p) \right\}.$$

We can also define the operator *domain* as

$$dom_a(p) = 1 - \left| c_{\chi^a}(p) - \chi^a(p) \right| \qquad (1),$$

giving low values to the elements not well represented by the seed a .

Then we can define a fuzzy *Domain Image DI,* assigning to each element p the fuzzy value

$$DI(p) = \bigcup_i dom_{a_i}(p) \ .$$

This article aims to find a complete set of seeds whose related connectivity maps can describe all the connected components contained in an image: care is also exercised to avoid an explosion of the number of seed points, without adding new information. To this end, we have to find a set of seeds such that $DI(p) = 1$ for all the elements p belonging to I , with fast comparisons based on the use of the domain operator, as shown in (1), for the already extracted seed points.

Moreover, in order to be complete, a set of seeds must satisfy the condition that each point in the image should have a high membership value. This holds when $PI = 1$, with the application of the *or* operator to each image site of the fuzzy Partition Image elements.

The use of such simple operators, like the domain operator, based on the difference between the χ-connectedness and intensity maps associated with the seed a, and like the fuzzy *or* operator, applied to both the Domain Image and the Partition Image , makes this new approach very simple and fast.

An iterative procedure is then applied, starting from a seed and looking for the other seeds in the minima of the Domain Image and of the Partition Image. Both the Partition and Domain Images are updated whenever a new seed is selected. The starting seed can be selected at random, or by exploiting some statistical information concerning the histogram, such as the peaks or the extreme values. The stopping condition is based on the histogram of the Partition Image: when the elements of PI are all above a reasonably high value, it means that a high degree of certainty has been reached in the process.

Experimental results show that this is a good choice for a stopping criterion.

Fuzzy clustering is obtained by assigning the label l to the set of pixels for which the maximum in PI is given by the seed l .

3 APPLICATIONS AND RESULTS

As a qualitative example of the choice of the stopping criterion, in Figure 3.1 a real image is presented, acquired with a fixed camera and showing a simplified scene.

Results of the process at an intermediate step (Figure 3.2) and at the conclusive step (Figure 3.3) of the iterative procedure are shown.

Figure 3.1 - Digital image (first seed indicated).

Figure 3.2 - Fuzzy clusters at an intermediate step.

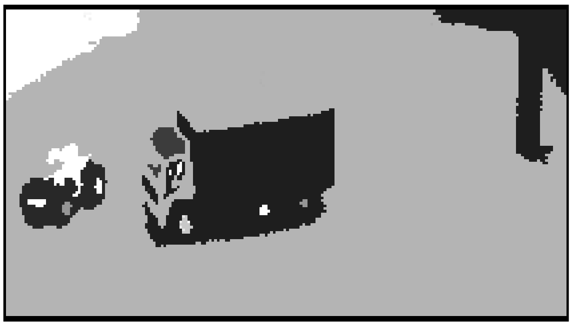

Figure 3.3 - Fuzzy clusters: final result.

By selecting, as the first seed, a point belonging to the histogram maximum, as depicted in Figure 3.1, the background can be located and the related connected component can be extracted. From now on, by alternating the detection of the minimum value in the updated Domain Image and in the updated Partition Image, new seeds can be iteratively detected. At an intermediate step of the procedure, the best contrasted regions are already detected (see Figure 3.2), but some particulars, like the wheels and part of the little tractor, are not distinct from the other regions.

In the final clustering result displayed in Figure 3.3, all connected components have been extracted, with a high degree of certainty, according to the stopping criterion.

About the sensitivity of the result to the choice of the starting seed, the result of a medical application is shown. The mean values and standard deviations of *PI* are given to indicate the independence of the final result of the position of the starting seed. The values in Table 3.1 have been obtained from a set of Magnetic Resonance images of the head [9].

An example of such images is shown in Figure 3.4, and the clustering result can be seen in Figure 3.5, for an automatic choice of the seed.

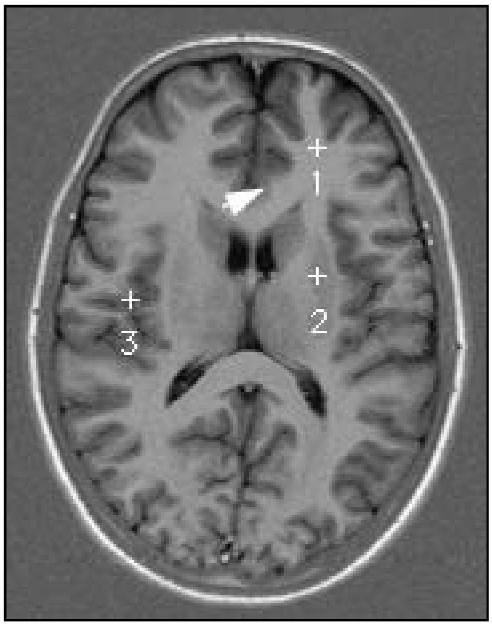

Figure 3.4 - An MR test image: automatic seed indicated by the arrow; numbers indicate manual seeds.

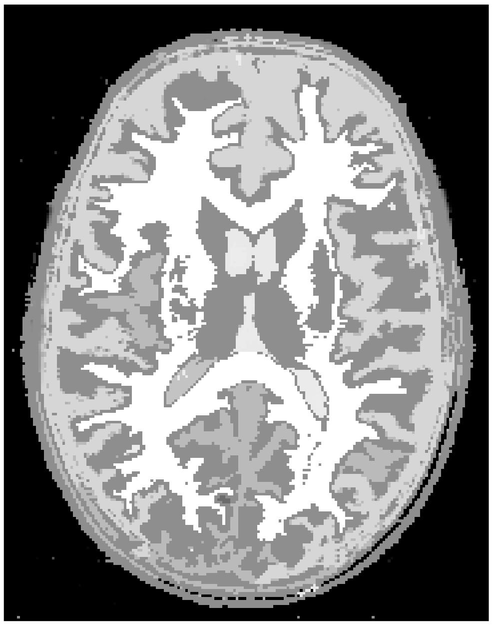

Figure 3.5 - Clustering result from Figure 3.4; automatic seed.

Values	Autom. Seed	Manual		
		Seed 1	Seed 2	Seed 3
PI mean	0.90	0.91	0.89	0.90
PI st.dev.	0.5	0.6	0.4	0.5

Table 3.1 - Mean and standard deviation values of the final PI for automatic and manual seed selections (number of seeds corresponds to Figure 3.4).

In table 3.1 the similar values of the mean of *PI* and of its standard deviation for the automatic and the different manual choices of the seeds show that the procedure is quite insensitive to the choice of the starting seed.

As the computation time for the extraction of each connectedness map on an 8 bit, 256x256 pixel image is about 0.3 sec on a SPARC-10 workstation, few seconds are required for the entire processing session.

4 CONCLUSION

The presented multiseed approach to fuzzy clustering exploits densitometric and topological information in an image. It allows the extraction of all connected components in an image, without thresholds or fixed parameters. The result is quite insensitive to the position of the starting seed, which can be chosen interactively or fully automatically.

The importance of the proposed approach also lies in its possible applications for the segmentation of a whole image, for compression and for database retrieval. The novelty of such image segmentation, which takes into account densitometric and topological information at the same time, consists mainly in using easy and fast computation methods to improve the handling of any digital image, whenever automatic segmentation or data reduction processing is required.

5 REFERENCES

[1] S. Geman and D. Geman, “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images,” IEEE Trans. on PAMI, vol. PAMI-6, No. 6, Nov. 1984, pp. 721-741.

[2] L. Vincent and P. Soille, ”Watersheds in digital spaces: an efficient algorithm based on immersion simulations,” IEEE Trans. on PAMI, vol. 13, No. 6, June 1991, pp. 583-598.

[3] M. Zehana, D. Saint-Joan and J. Desachy, "Fuzzy set technique for spatial problem solving,", Proc. IGARSS '94.

[4] S. Dellepiane and F. Fontana, "Extraction of intensity connectedness for image processing," Pattern Recognition Letters, vol. 16, March 1995, pp.313-324 .

[5] S. Dellepiane, F. Fontana, and G. Vernazza, "Non-linear image labelling for multivalued segmentation," IEEE Trans. on Image Processing, March 1996, vol.5, no. 3, pp. 429-446.

[6] L. O'Gorman and A.C. Sanderson, "The Converging Squares Algorithm: an efficient method for locating peaks in multidimensions," IEEE Trans.on PAMI, vol. 6, No. 3, May 1984, pp. 280-288.

[7] A. Rosenfeld, “On connectivity properties of grayscale pictures,” Pattern Recognition, vol. 16, No.1, 1983, pp.47-50.

[8] Bezdek J.K., "Pattern Recognition With Fuzzy Objective Function Algorithms," Plenum Press, New York, 1981.

[9] COVIRA Project A2003 of the AIM programme of the European Community.