International Journal of Pattern Recognition
and Artificial Intelligence
Vol. 22, No. 5 (2008) 829–842

CASE-BASED-REASONING FOR IMAGE SEGMENTATION

MARIA FRUCCI[*,†], PETRA PERNER[‡]
and GABRIELLA SANNITI DI BAJA[*,§]

[*]*Institute of Cybernetics "E.Caianiello"*
CNR, Pozzuoli (Naples) Italy

[‡]*Institute of Computer Vision and Applied Computer Science*
Leipzig, Germany
[†]*m.frucci@cib.na.cnr.it*
[‡]*p.perner@ibai-institut.de*
[§]*g.sannitidibaja@cib.na.cnr.it*

This paper proposes to use case-based-reasoning for grey-level image segmentation. Different approaches to image segmentation have been proposed in the literature. The selection of the segmentation approach and the assignment of the values to the parameters involved in the selected algorithm depend on image domain and on the specific application. Case-based-reasoning seems a promising way to make the above selection automatic. In this paper, we describe the results of a preliminary study done in this respect. In particular, we refer to the automatic selection of the values of the parameters for a new watershed image segmentation algorithm.

Keywords: Segmentation; watershed transformation; case-based-reasoning.

1. Introduction

Image segmentation is a key step in many applications in pattern recognition, computer vision and image understanding to allow further image content exploitation in an efficient way. The result of segmentation is a partition of a digital image into a number of regions, which are homogeneous according to some criteria and constitute components of either the foreground or the background.

A number of different approaches to segmentation have been suggested in the literature (see, e.g. the surveys in Refs. 5, 7, 16 and 27). The selection of the best approach depends on image domain as well as on the specific application. Moreover, once a segmentation scheme has been selected, the values for the parameters involved in the selected segmentation algorithm have to be tuned to the specific case.

Histogram thresholding is suitable for images where grey-levels constitute well-defined peaks, separated by not too broad and flat valleys. However, it does not take into account spatial information, while often, the grey-level distribution along the image is such that the same grey-level can characterize pixels that a human observer perceives as belonging to the foreground or to the background, depending on the local context. A survey of about 40 thresholding techniques, together with the evaluation of their performance, can be found in a paper by Sezgin and Sankur.[32] There, thresholding methods are categorized in terms of the information used to achieve the goal, such as histogram shape, measurement space clustering, entropy, object attributes, spatial correlation and local grey-level surface.

Feature space clustering can be regarded as a multidimensional extension of the concept of thresholding. A number of features, such as multispectral information, mean/variation of grey-level, texture, color, have been used for clustering schemes (see, e.g. Refs. 10, 22 and 34). Feature space clustering can be successfully used if each perceived region of the image constitutes an individual cluster in the feature space. This requires a careful selection of the proper features, which depends on image domain.

Edge detection can be used to extract the contours of the objects of interest in the image by taking into account the difference in contrast between adjacent regions. However, edge detection does not work well if the image is not well contrasted, or in the presence of ill-defined edges or too many edges. Better results are obtained by resorting to a combination of region-based and edge-detection-based methods (see, e.g. Refs. 14, 17, 18, 20, 28 and 38). One of the schemes using region and edge information is the watershed segmentation, which has received wide attention in the literature. A number of papers improve the basic watershed algorithm[2] and give possible solutions to the over-segmentation problem (see, e.g. Refs. 3, 8, 11, 19, 30 and 36); other papers deal with different applications of this segmentation scheme (see, e.g. Refs. 1, 4, 15, 29, 33 and 37). The basic idea of watershed segmentation is to identify, in the gradient image of the input image, a suitable set of *seeds* from which to perform a growing process guided by a homogeneity criterion. The seeds are mostly detected as the sets of pixels with locally minimal grey-level (called *regional minima*). The growing process groups to each seed all pixels of the gradient image that satisfy a certain homogeneity in grey-level with the seed and are closer to that seed more than to any other seed.

Other schemes involve a membership function to represent the degree of some properties and originate fuzzy segmentation (see, e.g. Refs. 13, 21, 31 and 35). A drawback is the generally high computational cost. For example, neural network techniques can be used to perform classification of regions, but the training phase is long and the results may be biased by the initialization phase.

Each of the above approaches involves the selection of suitable values for some parameters, e.g. the threshold value for histogram thresholding, or the parameters of the edge detector operator when segmentation is done by edge detection and so on. Thus, given an image domain and a specific application, the user should not only decide which segmentation scheme to use among various possible schemes, but also which are the values to be assigned to the corresponding parameters, so as to obtain the best segmentation results. We think that case-based-reasoning, CBR for short, can be a promising way to automatically select both the segmentation scheme and the corresponding parameters for a specific segmentation task.

The use of CBR has been already proposed in the past for some segmentation schemes. In the framework of histogram thresholding, based on a rule set the histogram can be properly smoothed so that the right number of peaks can be detected; case-based-reasoning was used in Ref. 23 to ensure the incremental learning of the rule set with the proper parameters. In Ref. 24, CBR was used to optimize image segmentation at the low-level stage of the process, i.e. by taking into account image acquisition conditions and image quality. In Refs. 25 and 26, CBR and dissimilarity classification methods were considered.

Our hypothesis is that the use of CBR can be exploited for image segmentation to select the best scheme for the current segmentation task among a number of possible schemes, and to derive automatically the values of the corresponding parameters. This hypothesis is based on the assumption that two images with similar characteristics should be segmented equally well by using the same scheme and the same values for the segmentation parameters. Of course, to reach this ambitious goal, a (large) case-base should be available. Each case should consist of the description of the characteristics of an image, which can be seen as the prototype for a class of similar images, together with the indication of the segmentation scheme and the values of the corresponding segmentation parameters, producing the best segmentation for the prototype. Then, whenever an input image is presented to the segmentation system, the similarity between the current image and each prototype should be evaluated and the case including the most similar prototype should be selected to derive the best segmentation scheme and the corresponding values of the parameters to process the current image. The criteria adopted to decide on image similarity obviously have a crucial role.

In this paper, we limit ourselves to use CBR in the framework of watershed segmentation and illustrate in more detail a preliminary study presented in Ref. 9. In particular, we improve the performance of the watershed segmentation algorithm[8] and start to use CBR to automatically identify the values for the watershed segmentation parameters. We first build a case-base including only images for which watershed segmentation produces satisfactory results. The segmentation algorithm runs on each image a number of times with different assignments of the values of the parameters, so producing different segmentations. The best segmentation is found for each image by comparing the result of the segmentation algorithm with a ground truth. The images are then grouped into classes of similar images and for each class a prototype is described in terms of statistical features. The values of the segmentation parameters producing an average best segmentation for all images in a class are associated to the prototype as its solution. Each case in the case-base consists of the description of the prototype and the values of the corresponding segmentation parameters. Then, given an input image, we use CBR to identify in the case-base the most similar prototype and the solution associated to the prototype is used to run the segmentation algorithm on the input image. Our work is still in a preliminary stage and we need to increase the number of images used to build the case-base to evaluate the performance of the system. However, the results obtained so far are promising. In the future, we will also improve the criteria adopted to check image similarity and will include in the segmentation system more than one segmentation scheme.

2. CBR for Image Segmentation

Segmentation can be interpreted as a classification problem, where the current input image I is compared to the images in a data-base to identify the image K with the highest similarity degree and, hence, select the segmentation criteria adopted for K also to segment I.

During a learning phase, the classifier learns the mapping function between the image features and the segmentation parameters involved in the selected algorithm. If learning is done by using a large set of images, a general model for the segmentation problem can be built. Of course, the initial set of images, though large, does not generally include the prototypes of all possible classes of images. Thus, the segmentation model should be adjusted to fit new data by means of a suitable case-base maintenance process (not yet included in our system), when the current image I does not suitably match any image in the initial set. We point out that a general model does not guarantee that the best segmentation is obtained for each image, but guarantees an average best fit over the data-base.

In the case-base, a case consists of a description of the prototype of a class of similar images, coupled with the best solution to its segmentation (i.e. the values of the parameters producing the best result). In our system, the description of the prototype is given in terms of the statistical features characterizing the whole image. These features are used for indexing the case-base and for retrieval of a set of cases close to the current problem, based on a proper similarity measure. The CBR system for image segmentation is schematically shown in Fig. 1, where the boxes with a grey background correspond to parts not yet included in the current version of the system.

Image similarity has a key role. It will be used both to build the case-base, by grouping similar images into classes, and to compare any input image to the prototypes of the classes in order to detect automatically the proper values for the segmentation parameters. Image similarity is here evaluated by taking into account the following statistical features, suggested in Ref. 6, that can be extracted from the grey-level image: mean, variance, skewness, kurtosis, variation coefficients, energy, entropy, centroid. The features are described in Table 1. The first order histogram $H(g)$ is equal to $N(g)/S$, where g is the grey-level, $N(g)$ is the number of pixels with grey-level g and S is the total number of pixels.

The similarity between two images A and B is computed on the basis of the above features by the distance between A and B, as follows:

$$\mathrm{dist}_{AB} = \frac{1}{k}\sum_{i=1}^{K} w_i \left| \frac{C_{iA} - C_{i\,\min}}{C_{i\,\max} - C_{i\,\min}} - \frac{C_{iB} - C_{i\,\min}}{C_{i\,\max} - C_{i\,\min}} \right|$$

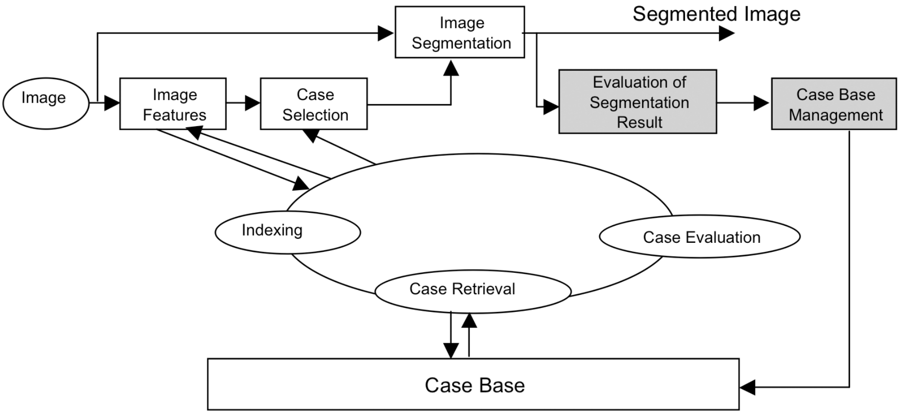


Fig. 1. The CBR system.

Table 1. Statistical features.

Feature Name	Calculation	Feature Name	Calculation
Mean	$\bar{g} = \sum_g g \cdot H(g)$	Variance	$\delta_g^2 = \sum_g (g - \bar{g})^2 H(g)$
Skewness	$g_s = \frac{1}{\delta_g^3} \sum_g (g - \bar{g})^3 H(g)$	Kurtosis	$g_k = \frac{1}{\delta_g^4} \sum_g (g - \bar{g})^4 H(g) - 3$
Variation Coefficient	$v = \frac{\delta}{\bar{g}}$	Entropy	$g_E = -\sum_g H(g) \log_2 H(g)$
Centroid_x	$\bar{x} = \frac{\sum_x \sum_y x f(x,y)}{\sum_x \sum_y f(x,y)} = \frac{\sum_x \sum_y x f(x,y)}{\bar{g} S}$	Centroid_y	$\bar{y} = \frac{\sum_x \sum_y y f(x,y)}{\sum_x \sum_y f(x,y)} = \frac{\sum_x \sum_y y f(x,y)}{\bar{g} S}$

where k is the number of features, C_{iA} and C_{iB} are the values of the ith feature of A and B, $C_{i\min}$ and $C_{i\max}$ are the minimum and maximum value of the ith feature of all images in the data-base, and w_i weights the ith feature, with $w_1 + w_2 + \cdots + w_k = 1$. In this paper, the weights w_i assume equal values.

3. The Segmentation Algorithm

Our segmentation algorithm is based on the watershed transformation. The gradient image obtained from an input grey-level image can be interpreted as a 3D binary landscape, where the grey-level of each pixel of the 2D image is interpreted as its height in the 3D landscape. Pixels with grey-levels locally higher than those of the neighboring pixels of the 2D image are mapped into peaks of mountains and pixels with grey-levels locally lower than those of the neighboring pixels are mapped into pits of valleys. Let us suppose that the pits are pierced and that the landscape is immersed in water. Then, as soon as the water level will reach the pits, the landscape will start to be flooded. The valleys that will be flooded first are those whose pits are the lowest ones, since these pits are reached first by the increasing level of the water. A dam has to be built wherever waters from different basins are going to meet, in order to prevent water to spread from a catchment's basin into the close ones. When the whole landscape has been covered by water, the top lines of the dams constitute the watershed lines. These lines surround, in the 2D grey-level image, all the regions into which the image has been partitioned. In practice, the regional minima in the gradient image (i.e. the pits in the 3D landscape) have to be identified as seeds for the growing process, which will incorporate in each region the pixels whose grey-level satisfy an homogeneity criterion and are closer to that seed than to any other seed.

It is well known that if all the regional minima detected in the gradient image are used as seeds for the growing process, the image is fragmented into a too large number of homogeneous regions, not all perceptually significant. This problem is generally known as over-segmentation. An example is shown in Fig. 2, where an input image (taken from the image data-base http://www.eecs.berkeley.edu/Research/Projects/CS/vi-sion/bsds/ of EECS, University of California, Berkeley, USA), the gradient image and the watershed partition obtained by using all the 1519 seeds detected in the gradient image are shown from left to right.

In Ref. 8, a segmentation algorithm based on the watershed transformation has been suggested, which significantly reduces over-segmentation. The goal was reached by introducing a criterion to evaluate region significance in watershed partitioned images and, accordingly, to filter out the irrelevant seeds. Region significance was defined in terms of two parameters, respectively measuring the relative depth, D_{XY}, of a catchment basin X with respect to an adjacent basin Y and the difference, Δ_{XY}, between the regional minima of X and Y. Two thresholds, d_t and δ_t, computed automatically by using statistics on the initial watershed partition of the grey-level image, were used and a region X was classified as significant with respect to the adjacent region Y if $D_{XY} > d_t$ or $\Delta_{XY} > \delta_t$. Seeds corresponding to nonsignificant regions were filtered out and the watershed transformation was applied again starting from the surviving seeds. Seed removal and watershed transformation were iterated until all regions in the partition were significant.

Fig. 2. (Left) A grey-level image, (middle) the gradient image and (right) the watershed partition.

Here, we improve the above algorithm by replacing the crisp rule, based on the OR of the two conditions on relative depth and difference between the regional minima, with a new rule where both conditions $D_{XY} > d_t$ and $\Delta_{XY} > \delta_t$ contribute to the decision on region significance. When the rule suggested in Ref. 8 is used, it is enough that any of the two conditions is satisfied to classify a region as significant, independently of image domain.

However, we think that for images having different characteristics the two conditions should be weighted in such a way to let them contribute differently to the decision on region significance. Thus, we introduce two weights a and b and a threshold T and use the following rule to decide on region significance:

$$1/2(a\Delta_{XY}/\delta_t + bD_{XY}/d_t) > T.$$

For each image used to build the case-base, the segmentation algorithm runs a number of times with different combinations of values of the parameters a, b, and T. The obtained results are compared to a ground truth to select the best segmentation as the one minimizing dissimilarity with respect to the expected segmentation. In this way, we identify the best combination of values for the segmentation parameters.

For example, for the image in Fig. 2 left, Table 2 summarizes some results obtained with different combinations of a, b, and T(where $a + b = 2$).

For the image in Fig. 2 (left), the best segmentation has been obtained by selecting $a = 0.75$, $b = 1.25$, and $T = 0.8$, i.e. by weighting more D_{XY} than Δ_{XY}. Only 72 seeds, and hence only 72 regions, are detected and over-segmentation is really very limited. The resulting segmentation is shown in Fig. 3 (right).

Rather good results are also obtained for the image in Fig. 2 by still selecting $a = 0.75$ and $b = 1.25$, but using different values for T. As expected, the number of regions decreases when the value of T increases. Slightly under-segmented results

Table 2. Some selections of a, b, and T for the image in Fig. 2 (left).

a	b	T	# Regions	Evaluation
1.	1.	0.7	72	Under-segmented: part of the head is merged to the background
1.	1.	0.6	85	A good result with small over-segmentation
1.	1.	0.5	93	A good result with small over-segmentation
0.75	1.25	0.9	64	Under-segmented: part of the head is merged to the background
0.75	1.25	0.8	72	A very good result. The best one.
0.75	1.25	0.7	83	A good result with some over-segmentation (the wing and the body consist both of two parts)
0.75	1.25	0.6	99	Over-segmented
1.5	0.5	0.4	70	Under-segmented: part of the head is merged to the background
1.5	0.5	0.3	86	A good result with small over-segmentation
1.5	0.5	0.2	115	Over-segmented

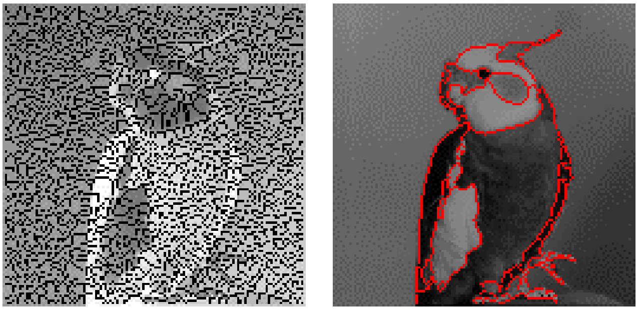


Fig. 3. (Left) Watershed partition and (right) over-segmentation reduction by the proposed segmentation algorithm.

are obtained for $T \geq 0.9$, and slightly over-segmented results are obtained for $T \leq 0.7$. Reasonably good results were also obtained with different selections of the weights. For example, for $a = 1$, $b = 1$, the value $T = 0.6$ produces a result that is only slightly over-segmented.

The fact that good results are obtained when running the algorithm with different combinations of a, b, and T is important to build the case-base. In fact, each case will include the description of the prototype of a class of similar images together with the values of the segmentation parameters expected to produce the best segmentation. Thus, when the combination of values for a, b, and T producing the best segmentation is not the same for all images in a class, we can still identify a common solution for the case, which produces on the average the best (or at least a good) segmentation for all images in the class.

4. Building the Case-Base for Image Segmentation

To build our case-base, we analyzed images from different domains (neural cells, faces, buildings, flowers and animals). The statistical features shown in Table 1 were used to describe the images. Then, clustering based on the normalized city-block metric and the average linkage method[12] was applied to separate different cases and to form groups of similar cases. As an example, the results produced by clustering accomplished on a small set of 15 images are shown in Fig. 4.

Our expectation is that images, for which we got the best segmentation by using the same values of the parameters, would cluster into groups of similar images. By following this idea, we have to cut the dendrogram in Fig. 4 by a value of the cophenetic-similarity value such that the first aggregation, where images having different similarity measures meet, does not form a group.

When the values of the segmentation parameters experimentally found to produce the best segmentation results of all images in a cluster are identical, these values are selected as the solution and are recorded in the corresponding case together with the description of the prototype of the cluster. When different best values are

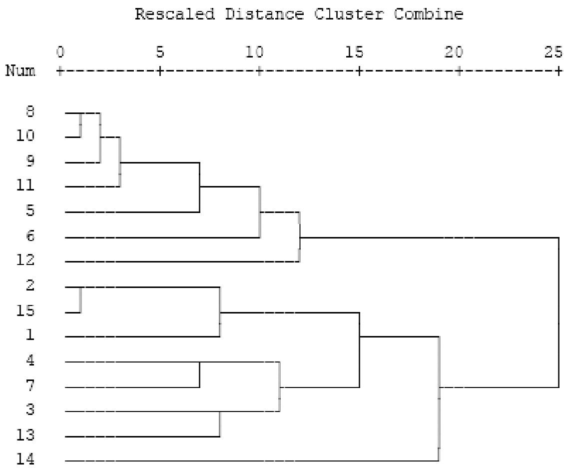


Fig. 4. Dendrogram based on statistical features.

found for the segmentation parameters of images in the same cluster, the solution is the set of values producing on the average the best segmentation results for the images in the cluster.

When a new input image A is taken into account, its similarity to the images in the case-base is measured. The description of the image A is compared to the descriptions of the prototypes and the distances are ranked according to increasing distance values. The prototype ranked first, i.e. the one from which the current input image has the minimal distance, is used to identify the class to which the input image is expected to belong. The solution associated to that class is used to run the segmentation algorithm on the current input image. If no image reasonably similar to A is found, then A has to be added to the case-base as a new case. To this purpose, the best values of a, b, and T have to be found experimentally.

As an example, in Fig. 5 (left), the best segmentation obtained by using different combinations of the values of the parameters is shown for a new test

Fig. 5. (Left) Segmentation obtained by running the algorithm with different combinations of a, b and T and (right) segmentation obtained with the solution suggested by using CBR.

image. Also in this case, the input image has been taken from the image data-base (http://www.eecs.berkeley.edu/Research/Projects/CS/vi-sion/bsds/) of EECS, University of California, Berkeley, USA. The values of the parameters, experimentally found, are $a = 1.5$, $b = 0.5$, and $T = 0.4$. Only 382 regions constitute the partition, while the watershed partition, obtained by taking into account all the regional minima, originated 6370 regions. If we use CBR and compare this image to the prototypes in the case-base, the most similar image results to be the image in Fig. 2 (left) CBR would, hence, suggest for this image the solution $a = 0.75$, $b = 1.25$, and $T = 0.8$. In Fig. 5 (right), the segmentation into 454 regions obtained by selecting this solution is shown. We note that, though the image is slightly over-segmented, the result obtained by using CBR is acceptable.

5. Concluding Remarks

CBR can be used to solve all aspects of image segmentation, from choosing the appropriate image segmentation method/parameters for the actual image up to the evaluation of the results, and to provide feedback to the system for performance improvement. In this paper, we have described how CBR can be applied to watershed-based image segmentation by controlling the merging process.

The results obtained so far are promising, but we are aware that further work is necessary. More research has to be done on the definition of the proper image description. Image description based on statistical features might properly cover the information about image quality, but this is possibly not enough for watershed-based image segmentation. We believe that texture features could be profitably used in conjunction with statistical features.

Acknowledgments

This work has been partially supported by the Italian National Research Council, CNR, in the framework of the Short Term Mobility Program 2006.

References

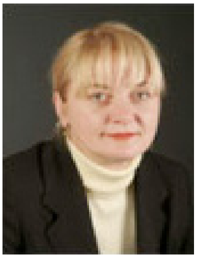
1. J. Barraud, The use of watershed segmentation and GIS software for textural analysis of thin sections, *J. Volcanol. Geoth. Res.* **154**(1–2) (2006) 17–33.
2. S. Beucher and C. Lantuéjoul, Use of watersheds in contour detection, *Proc. Int. Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation*, Rennes, France (1979), pp. 12–21.
3. A. Bleau and L. J. Leon, Watershed-based segmentation and region merging, *Comput. Vis. Imag. Underst.* **77**(3) (2000) 317–370.
4. J. E. Cates, R. T. Whitaker and G. M. Jones, Case study: an evaluation of user-assisted hierarchical watershed segmentation, *Med. Imag. Anal.* **9**(6) (2005) 566–578.
5. H. D. Cheng, X. H. Jiang, Y. Sun and J. Wang, Color image segmentation: advances and prospects, *Patt. Recogn.* **34** (2001) 2259–2281.
6. H. Dreyer and W. Sauer, *Prozeßanalyse* (Verlag Technik, Berlin, 1982).

7. J. Freixenet, X. Muñoz, D. Raba, J. Martí and X. Cufí, Yet another survey on image segmentation: region and boundary information integration, *Proc. 7th ECCV*, Lecture Notes in Computer Science, Vol. 2352 (Springer, 2002), pp. 408–422.
8. M. Frucci, Oversegmentation reduction by flooding regions and digging watershed lines, *Int. J. Patt. Recogn. Artif. Intell.* **20**(1) (2006) 15–38.
9. M. Frucci, P. Perner and G. Sanniti di Baja, Watershed segmentation via case-based reasoning, in *Brain, Vision and Artificial Intelligence*, eds. F. Mele, G. Ramella, S. Santillo and F. Ventriglia, Lecture Notes in Computer Science, Vol. 4729 (Springer-Verlag, Berlin, 2007), pp. 244–253.
10. K. Hammouche, M. Diaf and J.-G. Postaire, A clustering method based on multi-dimensional texture analysis, *Patt. Recogn.* **39**(7) (2006) 1265–1277.
11. S. E. Hernandez, K. E. Barner and Y. Yuan, Region merging using homogeneity and edge integrity for watershed-based image segmentation, *Opt. Engin.* **44**(1) (2005) 017004.
12. A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data* (Prentice-Hall, Inc., 1988).
13. G. C. Karmakar and L. S. Dooley, A generic fuzzy rule based image segmentation algorithm, *Patt. Recog. Lett.* **23**(10) (2002) 1215–1227.
14. C. D. Kermad and K. Chehdi, Automatic image segmentation system through iterative edge–region co-operation, *Imag. Vis. Comput.* **20**(8) (2002) 541–555.
15. M. M. J. Letteboer, O. F. Olsen, E. B. Dam, P. W. A. Willems, M. A. Viergever and W. J. Niessen, Segmentation of tumors in magnetic resonance brain images using an interactive multiscale watershed algorithm1, *Acad. Radiol.* **11**(10) (2004) 1125–1138.
16. L. Lucchese and S. K. Mitra, Color image segmentation: a state-of-the-art survey, *Image Processing, Vision, and Pattern Recognition, Proc. Indian Nat. Sci. Acad. (INSA-A)*, New Delhi, India **67**(2) (2001) 207–221.
17. M. Mueller, K. Segl and H. Kaufmann, Edge- and region-based segmentation technique for the extraction of large, man-made objects in high-resolution satellite imagery, *Patt. Recogn.* **37**(8) (2004) 1619–1628.
18. X. Muñoz, J. Freixenet, X. Cufí and J. Martí, Strategies for image segmentation combining region and boundary information, *Patt. Recogn. Lett.* **24**(1–3) (2003) 375–392.
19. V. Osma-Ruiz, J. I. Godino-Liorente, N. Sáenz-Lechón and P. Gómez-Vilda, An improved watershed algorithm based on efficient computation of shortest paths, *Patt. Recogn.* **40**(3) (2007) 1078–1090.
20. X. M. Pardo, M. J. Carreira, A. Mosquera and D. Cabello, A snake for CT image segmentation integrating region and edge information, *Imag. Vis. Comput.* **19**(7) (2001) 461–475.
21. L. Patino, Fuzzy relations applied to minimize over segmentation in watershed algorithms, *Patt. Recogn. Lett.* **26**(6) (2005) 819–828.
22. E. J. Pauwels and G. Frederix, Finding salient regions in images: nonparametric clustering for image segmentation and grouping, *Comput. Vis. Imag. Underst.* **75**(1–2) (1999) 73–85.
23. P. Perner, An architecture for a CBR image segmentation system, *J. Engin. Appl. Artif. Intell.* **12**(6) (1999) 749–759.
24. P. Perner, CBR ultra sonic image interpretation, *Advances in Case-Based Reasoning*, eds. E. Blanzieri and L. Portinale, Lecture Notes in Artificial Intelligence, Vol. 1898 (Springer Verlag, 2000), pp. 479–481.

25. P. Perner, Are case-based reasoning and dissimilarity-based classification two sides of the same coin? *J. Engin. Appl. Artif. Intell.* **15**(3) (2002) 205–216.
26. P. Perner, H. Perner and B. Müller, Similarity guided learning of the case description and improvement of the system performance in an image classification system, in *Advances in Case-Based Reasoning*, eds. S. Craw and A. Preece, ECCBR2002, Lecture Notes in Artificial Intelligence, Vol. 2416 (Springer Verlag, 2002), pp. 604–612.
27. D. L. Pham, C. Xu and J. L. Prince, Current methods in medical image segmentation, *Ann. Rev. Biomed. Engin.* **2** (2000) 315–337.
28. M. I. Rajab, M. S. Woolfson and S. P. Morgan, Application of region-based segmentation and neural network edge detection to skin lesions, *Comput. Med. Imag. Graph.* **28**(1–2) (2004) 61–68.
29. R. Rodríguez, T. E. Alarcón and O. Pacheco, A new strategy to obtain robust markers for blood vessels segmentation by using the watersheds method, *Comput. Biol. Med.* **35**(8) (2005) 665–686.
30. J. B. T. M. Roerdink and A. Meijster, The watershed transform: definitions, algorithms and parallelization strategies, *Fund. Inform.* **41** (2001) 187–228.
31. P. K. Saha, J. K. Udupa and D. Odhner, Scale-based fuzzy connected image segmentation: theory, algorithms, and validation, *Comput. Vis. Imag. Underst.* **77**(2) (2000) 145–174.
32. M. Sezgin and B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, *J. Electr. Imag.* **13**(1) (2004) 146–165.
33. P. S. U. Adiga and B. B. Chaudhuri, An efficient method based on watershed and rule-based merging for segmentation of 3-D histo-pathological images, *Patt. Recogn.* **34**(7) (2001) 1449–1458.
34. R. P. Velthuizen, L. O. Hall, L. P. Clarke and M. L. Silbiger, An investigation of mountain method clustering for large data sets, *Patt. Recogn.* **30**(7) (1997) 1121–1135.
35. Q. Wang, Z. Chi and R. Zhao, Image thresholding by maximizing the index of nonfuzziness of the 2-D grayscale histogram, *Comput. Vis. Imag. Underst.* **85**(2) (2002) 100–116.
36. J. Weickert, Efficient image segmentation using partial differential equations and morphology, *Patt. Recogn.* **34**(9) (2001) 1813–1824.
37. J. Yan, B. Zhao, L. Wang, A. Zelenetz and L. H. Schwartz, Marker-controlled watershed for lymphoma segmentation in sequential CT images, *Med. Phys.* **33**(7) (2006) 2452–2460.
38. Y. Zhou, J. Starkey and L. Mansinha, Segmentation of petrographic images by integrating edge detection and region growing, *Comput. Geosci.* **30**(8) (2004) 817–831.

Maria Frucci received the doctoral degree cum laude in physics from the University of Naples, Italy, in 1983. In the period 1984–1987, she worked as a senior researcher at the Centre for Informatics and Industrial Automation Research, Portici, Naples, Italy. Particularly, she was active in the area of natural language, expert systems and shape analysis. In 1988, she moved to the Institute of Cybernetics "E. Caianiello" of the National Research Council of Italy, Naples, where she has been working mainly in the field of image analysis. Since 1999, she has a teaching contract with the University of Naples.

Her main interests are in digital topology and geometry, properties of medial axes, thinning algorithms, shape analysis, description and segmentation.



Petra Perner received the diploma degree in electrical engineering in 1981 and the Ph.D. in computer science in 1985. In 1995, she founded the Institute of Computer Vision and Applied Computer Sciences in Leipzig, Germany, of which she is the Director.

Dr. Perner has been the project leader and principal investigator of various national and international research projects for which she received several research awards.

Dr. Perner was the principal founder and chairwoman of the IAPR Technical Committee "Machine Learning and Data Mining", is a IAPR Fellow and is the initiator and organizer of three distinct conferences in Machine Learning, Data Mining and High-Content Analysis of Signals and Images (MLDM, ICDM and MDA).

She has published numerous scientific publications and patents and is often requested as a plenary speaker in distinct research fields as well as across disciplines.

Her interests are in basic and applied research, in the fields of computer vision, data mining, machine learning, case-based reasoning and image databases.

Gabriella Sanniti di Baja received the doctoral degree cum laude in physics from the University of Naples, Italy, in 1973. In January 2002, she received the Ph.D. Honoris Causa from the Uppsala University, Sweden. Since 1973, she has been working in the field of image processing and pattern recognition at the Institute of Cybernetics "E. Caianiello" of the National Research Council of Italy, Naples, where she is currently Director of Research.

She has published more than 130 papers in international journals and conference proceedings and is Editor-in-Chief of *Pattern Recognition Letters* (Special Issues). She has been a member of the Executive Committee of the International Association for Pattern Recognition (IAPR) in the period 1994–2004, being IAPR President in the period 2000–2002. She is IAPR Fellow.

Her main research activities concern 2D and 3D shape representation, decomposition and description.