Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

Pattern Recognition Letters 24 (2003) 1545–1554

Pattern Recognition Letters

www.elsevier.com/locate/patrec

A multichannel watershed-based algorithm for supervised texture segmentation

Norberto Malpica, Juan E. Ortuño, Andrés Santos *

Departamento de Ingeniería Electrónica, Universidad Politécnica de Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain

Received 27 April 2002; received in revised form 15 November 2002

Abstract

Segmentation of image regions based on their texture is a standard problem in image analysis. Once a set of texture features is selected, several algorithms can be applied to segment the image into regions. This paper presents an extension of the watershed algorithm using a vector gradient and multivariate region merging methods. The algorithm uses a set of texture images, and it only depends on an adjustable parameter. Results are presented on a standard set of synthetic images and on textured medical ones, using different texture parameters and merging criteria.

Keywords: Texture segmentation; Watershed algorithm; Multichannel information

1. Introduction

Texture is an important feature in the visual perception of natural images and its description and regional analysis are basic stages of many image analysis applications, that include the identification of regions with a homogeneous texture, or image segmentation based on texture. This task is usually achieved by region classification: every pixel in the image is assigned to a class based on different parameters previously selected, computed on a window around the pixel. Much work has focused on finding texture parameters adequate for classification such as cooccurrence matrices (Haralick, 1979), histogram based features, wavelets (Unser, 1995), Gabor filtering (Weldon et al., 1996), Gaussian Markov random fields (Chellappa and Chatterjee, 1985) and recently signed gray-level differences and local binary patterns (Ojala et al., 2001). A comparative review of filter based methods can be found in (Randen and Husoy, 1999). Classification is usually carried out, for example, by multivariate coefficient thresholding (Corneloup et al., 1996), supervised classifiers (Randen and Husoy, 1999) or relaxation algorithms (Muzzolini et al., 1993). Ojala and Pietikäinen (1999) used a split and merge approach.

Texture classification usually involves the processing of several features to adequately characterize each region texture, and this multichannel nature of texture information creates difficulties in the applicability of most segmentation tools, usually defined for intensity-based single value images.

* Corresponding author. Fax: +34-913367323.
E-mail address: andres@die.upm.es (A. Santos).

0167-8655/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0167-8655(02)00393-8

However, several authors have adopted mathematical morphology tools for segmenting textures. One of the main gray-level segmentation tools, the watershed, allows for an intuitive initial image splitting, dependent on the selection of significant initial image minima, which can be accomplished using morphological algorithms (dynamics, waterfall). Jones (1994) used a watershed like algorithm for segmentation applied to a specifically designed texture difference map. In (Marcotegui et al., 1995) a texture parameter, signal to noise ratio, was used to improve the merging process starting from a flat zone splitting of the image. The use of the watershed algorithm for color images can be seen as a three channel extension of the algorithm. Shafarenko et al. (1997) presented a watershed algorithm for color images, based on a gradient in color space which, due to the multichannel nature of color images, can be considered as an antecedent of texture multichannel segmentation. Recently, Hill et al. (2002) have proposed a watershed based segmentation for texture analysis, using wavelet filter banks and the addition of each channel gradient. They avoid oversegmentation by selecting a marker for each region.

This paper presents a straightforward extension of the watershed algorithm to make use of information in all texture channels. In the technique proposed, several texture channels are computed from the original image. A vector gradient is used to compute the edges of the multichannel image. After an automatic selection of significant minima, a watershed transform is applied. Finally, a multivariate region merging step is carried out to obtain the final segmentation. The steps of the algorithm are shown in Fig. 1.

The complete segmentation procedure consists on the following steps:

- *Texture analysis:* A set of n channels is obtained from the original image by computing n texture features on every pixel. This procedure uses the most discriminant texture features that have previously been selected by statistical analysis.
- *Gradient computation:* Considering each point of the image as an n-valued vector, a gradient image of all the channels is obtained by computing the gradient of the vector field.

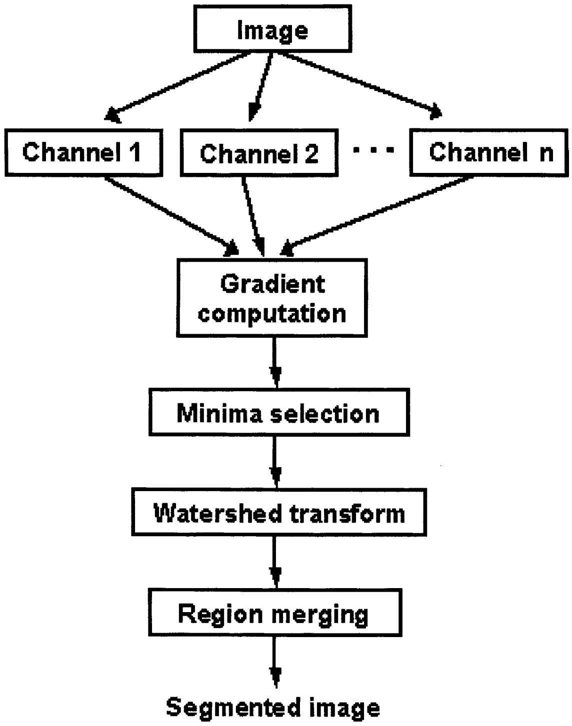

Fig. 1. Steps of the segmentation algorithm.

- *Minima selection:* According to some measure of local minima importance, some minima are selected from which the watershed process will begin. The topographic surface will then only be pierced in those points to start the immersion process.
- *Watershed segmentation:* Starting from the selected minima, the image is flooded by extending their zone of influence in higher gray levels. When two regions come into contact a watershed line is erected.
- *Region merging:* Watershed regions are iteratively merged, according to a similarity criterion, to obtain the final segmentation.

The whole procedure is automatic except for a minima selection threshold that will be described in Section 4.

The rest of the paper is structured as follows. Section 2 defines the vector gradient used. Section 3 reviews the watershed transform and the main segmentation steps. Section 4 describes the minima selection step. Section 5 describes the multivariate region merging approach. Section 6 presents experimental results on real texture images and a conclusion is presented in Section 7.

2. Edge detection: vector-valued edges

To apply edge based segmentation algorithms to multichannel data, a gradient of the multichannel image must be defined. Instead of separately computing the scalar gradient for each channel, some authors have proposed solutions using tensor notation. This framework was first suggested in (DiZenzo, 1986) and extended in (Cumani, 1991). Lee (1991) also proposed to use the vector gradient to detect boundaries in multidimensional images (specially color images), proving that, when the attribute components are highly correlated, vector gradients are less sensitive to noise than the sum of the squares of scalar gradients in each channel. As this is also the case with multichannel texture images, this will be the approach used in this paper.

A multichannel image can be seen as a vector field $f : S \to R^m$, defined on a subset S of R^n. Let f_k denote the kth component of f (if f is a three color image, then f_1, f_2, and f_3 might represent the red, green and blue components of the image). It can be proved Lee (1991) that the first-order Taylor expansion takes the form

$$f(x+a) = f(x) + [f'(x)](a) + \|a\|e(x,a)$$

where $e(x,a) \to 0$ as $a \to 0$ and $f'(x)$ is now an $m \times n$ Jacobian matrix $D(x)$:

$$f'(x) = D(x) = \begin{bmatrix} D_1f_1(x) & D_2f_1(x) & \cdots & D_nf_1(x) \\ D_1f_2(x) & D_2f_2(x) & \cdots & D_nf_2(x) \\ \vdots & \vdots & & \vdots \\ D_1f_m(x) & D_2f_m(x) & \cdots & D_nf_m(x) \end{bmatrix}$$

being D_jf_k the first partial derivative of the kth component of f with respect to the jth component of x.

It can be proved (Cumani, 1991) that the extrema of $D(x)$ are obtained in the directions of its eigenvectors $\theta_\pm$ and the values attained there are the corresponding eigenvalues $\lambda_\pm$. A first approximation of edges for vector-valued images, should be a function $g = f(\lambda_+, \lambda_-)$. We will use the simple form $g = \lambda_+ - \lambda$ in this work.

3. The watershed transform

The watershed transform is the main segmentation technique in mathematical morphology. Since its introduction in (Beucher and Meyer, 1993) and (Vincent and Soille, 1991), it has been widely studied in several image segmentation problems. It has been applied successfully to 2D (Haris et al., 1998) as well as to 3D images (Sijbers et al., 1997). It is a connected (it divides the image into sets of connected pixels) and complete (it assigns every pixel to one of the classes) method.

The watershed concept comes from the field of topography: in a topographic surface, the *watersheds* are the lines dividing two catchment basins. Each basin is associated to a local minimum. If we were to drop water on the surface, every drop would fall into a single catchment basin and would follow a downward path until it reached a minimum of the surface. Other way of visualizing the concept is by analogy to immersion. Starting from every minimum, the surface is progressively flooded until water coming from two different minima meet. At this point a watershed line is erected. Efficient algorithms have been proposed to simulate the immersion process (Vincent and Soille, 1991; Beucher and Meyer, 1993).

Homogeneous regions in a gray-level image can be segmented by applying the algorithm to the gradient image, where borders between objects correspond to high gray levels (peaks) and the inside of objects correspond to valleys. We have used the algorithm by Vincent and Soille (Vincent and Soille, 1991), which is a recursive simulation of the immersion process:

- All local minima are assigned a different label, which is propagated to all adjacent pixels at a certain level h.
- All levels from the minimum to the maximum gray level are processed. At every level h:

- Every pixel at a certain level with an already labeled neighbor receives its label. Labels are propagated from these to all pixels in that level.
- Pixels having two or more neighbors with different labels are labeled as watershed points.
- Remaining pixels are considered as new minima, and are assigned a new label.

In all watershed algorithms, both eight and four neighbour connectivity can be used. We have used four neighbour connectivity in the experiments presented in this work.

As the transform starts from every minimum, the result is an oversegmented image if a minima selection procedure or a merging step are not used. Both methods are applied in this work.

4. Minima selection

Several algorithms have been proposed for minima selection. The simplest way is interactive selection by the user. A second approach is the automatic selection of minima using a priori knowledge of the image (Beucher and Meyer, 1990). Another approach is to hierarchically order all minima and to select only those above a threshold. The *dynamics* (Najman and Schmitt, 1996) and waterfall (Beucher, 1994) algorithms fall into this category.

We have used *dynamics* as the criterion for minima reduction, as it provides an intuitive selection scheme controlled by a single parameter. The concept is easily visualized using the immersion simulation. The deepness of a basin would be the level the water would reach, coming in through the minimum of the basin, before the water would overflow into a neighbor basin. That is, the height from the minimum to the lowest point in the watershed line of the basin. The *dynamics* of a basin is a similar concept, but referring to a neighbor basin with a lower minimum than the actual one.

It can be computed using morphological reconstruction (Vincent, 1993). Given two gray-scale images f and g, where g is smaller than or equal to f in all points, the morphological reconstruction by erosion of f over g is the result of the following iterative procedure as follows:

$$f^{k+1}(p) = \max(g(p), f^k(p)\Theta B)$$

where Θ denotes the gray-level erosion of the function using a structuring element B.

It can be proved that if the morphological reconstruction (Vincent, 1993) of the image g from an image $f = g + d$ is computed, the resulting function will have a watershed transform in which the regions with dynamics lower than d, have been eliminated by joining them to the neighbor region with the lower minimum gray-level value. The parameter d serves as a minima selection threshold.

The *dynamics* algorithm is applied to the gradient of the image. We can expect the regions that are due to noise to be surrounded by a smaller gradient, and therefore, to have a smaller *dynamics* than regions formed by objects in the image. The value of the *dynamics* is automatically selected depending on the dynamic range of the image. Fig. 2 shows the results of the final watershed after minima selection using two different values of *dynamics*, before the merging step.

5. Region merging

To obtain the final segmentation, regions resulting from the watershed need to be merged to reduce the number of regions. This section explains the region merging criteria and methods employed. All regions resulting from the watershed transform are ordered into a region adjacency graph (RAG) in which nodes correspond to regions and arcs to the frontier between two regions. The aim of the merging step is to reduce the RAG until the desired number of nodes remains. Several merging criteria have been proposed in the literature: mean, normalized mean (Haris et al., 1998), T test (Sijbers et al., 1997), minimum description length (Maes et al., 1995). Here we extend three of these criteria to our multichannel framework:

Mean difference: The difference on the mean value of all the channels in the two regions is computed as the sum of the squared roots of the individual differences:

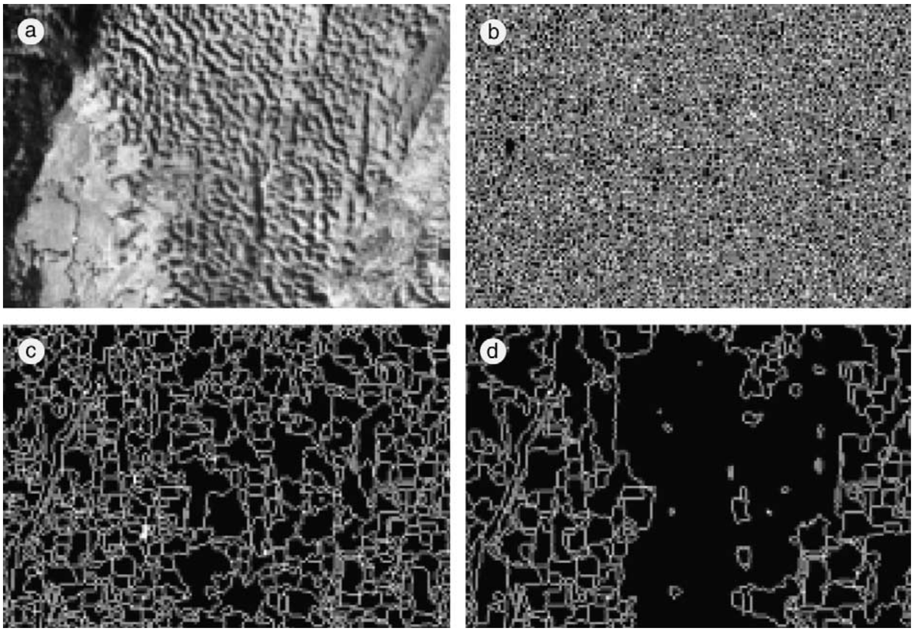

Fig. 2. (a) Satellite image showing different terrain textures. (b) Watershed of the gradient using 8 cooccurrence feature channels. (c) Watershed with minima selection with *dynamics* $d = 8$. (d) Watershed with minima selection with *dynamics* $d = 12$. Final merging step is not performed to show the effect of different *dynamics* values.

$$\mathrm{Diff} = \sum_{j=1}^{n} \sqrt{\mu_{2j} - \mu_{1j}}$$

where n is the number of channels and μ_{ij} is the mean value of channel j in region i.

Weighted mean criterion: In (Haris et al., 1998) a dissimilarity function for gray-level images was proposed, by normalizing the difference of means between regions by the number of points in them. For several channels the following dissimilarity function is used:

$$\mathrm{Diff} = \frac{n_1 n_2}{n_1 + n_2} \sum_{j=1}^{n} \sqrt{\mu_{2j} - \mu_{1j}}$$

being n_1 and n_2 are the number of pixels in regions 1 and 2, respectively.

Hotelling T^2 test: Assuming a multivariate normal distribution of gray levels in a region, we can apply Hotelling's T^2 test to compare the means of two populations. The applied statistic is given by:

$$T^2 = (\mu_1 - \mu_2)^T \left[\sqrt{\frac{1}{n_1} + \frac{1}{n_2} S} \right]^{-1} (\mu_1 - \mu_2)$$

where μ_1 and μ_2 are the vector means of each region, respectively, and S is a pooled covariance matrix of the form:

$$S = \sqrt{\frac{n_1 \Sigma_1 + n_2 \Sigma_2}{n_1 + n_2}}$$

being Σ_1 and Σ_2 the covariance matrices of each region.

When the number of regions is high, a full search for the most similar neighboring regions can be a computationally intensive task. We can make the assumption that very small regions are due to noise in the image. A faster sub-optimum merging scheme consists in joining in every iteration the smallest region with the most similar neighbor. We have used a combined scheme as in (Sijbers et al., 1997). The merging process starts by merging the smallest region in every iteration, until a certain number of regions is obtained. This process assumes desired regions are of a certain size, and very small regions are due to noise. From there on, a complete search for the two most similar regions is carried out in every iteration.

This allows a considerable reduction in computing time, while maintaining a correct result.

6. Results

The proposed methodology has been tested with a standard set of synthetic texture images that was originally proposed by Randen and Husoy (1999) to review filtering-based texture analysis methods. The images are available on the internet at http://www.ux.his.no/~tranden/data.html.

The first three images were used to evaluate the merging strategies and the effect of changing the *dynamics* parameter. Histogram and cooccurrence matrix features (Haralick, 1979) were used to classify the different textures. The most discriminant features were selected using a stepwise F test and the vector gradient was computed using the channels obtained. The watershed segmentation was applied with the three different merging criteria and different values of the dynamics, using a vector gradient image computed from the same set of features. The percentage of correctly segmented pixels was used as the quality factor. Fig. 3 shows the results of applying the different methods on the segmentation of the three images.

As mentioned in the methods section, in order to decrease computing time, a combined merging approach was also evaluated. In Table 1, the result of using the full and combined search strategies are compared. Table 1 also shows computing time of each approach on a Pentium III at 500 MHz.

From these experiments, it can be seen that merging with mean value provides worse results, while weighted mean and Hotelling T^2 both have similar performance. The influence of *dynamics* is not critical as far as it is low enough. Only when important minima are lost, the classification performance decreases.

Results in Table 1 also show that a great improvement in merging speed can be achieved with the mixed-merging approach, without losing segmentation accuracy.

Examples of segmentation results are presented below. Minima pre-selection was performed in all cases using a *dynamics* of 8% of the dynamic range of the image and merging with Hotelling T^2 test.

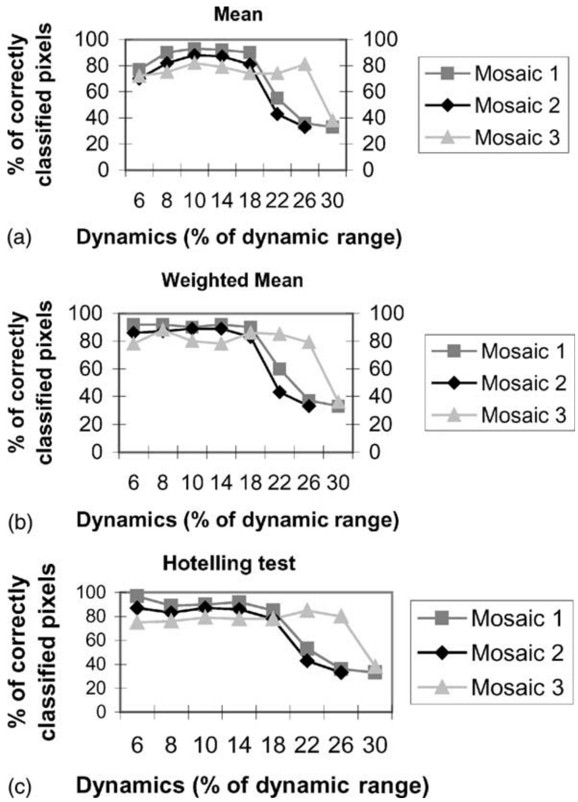

Fig. 3. Results of segmentation of mosaics #1, #2 and #3 with different values of the dynamics parameter. Merging based on the mean (a), the weighted mean (b) and Hotelling's test (c).

Table 1
Segmentation results and computing times of full vs. partial search merging schemes

	Full search results (%)	Full search times (s)	Partial search results (%)	Partial search times (s)
Mean	21.44	90	21.85	17
Weighted mean	7.76	51	9.72	19
Hotelling T^2 test	9.88	26	10.67	7

In the first example, images 10 and 11 from the test set (two-texture images) are segmented. An initial training for parameter selection was carried out, in which 12 regions of interest (whose sizes were 16×16 pixels) from each texture were selected, and cooccurrence matrix (Haralick, 1979) as well as the histogram parameters were computed for each of them. A stepwise F test was performed to select the most discriminant parameters. For these texture channels, the vector gradient and the watershed were computed. Region merging using the combined method described above was performed until only two regions were left. The result is shown in Fig. 4.

Another experiment has been carried out to analyze the effect of the number of texture channels. Fig. 5 shows the segmentation of mosaic #1 (five-texture image) also cooccurrence matrix texture features, all computed on a 16×16 window. The three results shown were obtained using the first 4, 6 and 10 channels obtained in the stepwise F test feature selection.

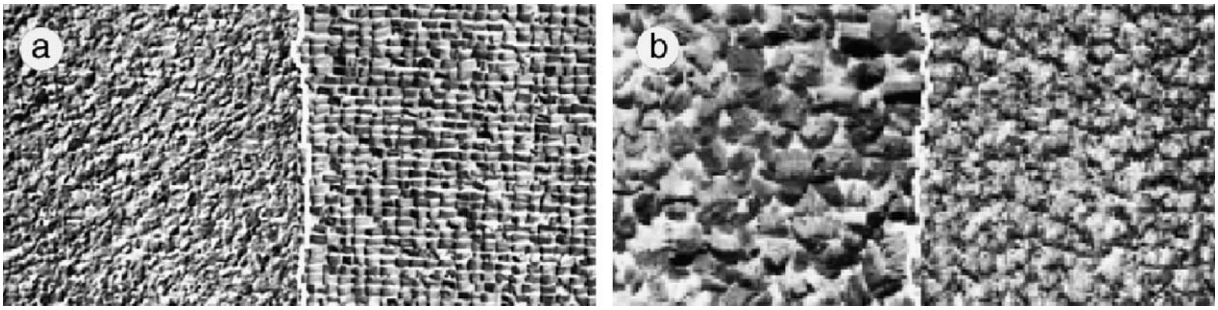

Fig. 4. Results of segmentation of test images #10 and #11.

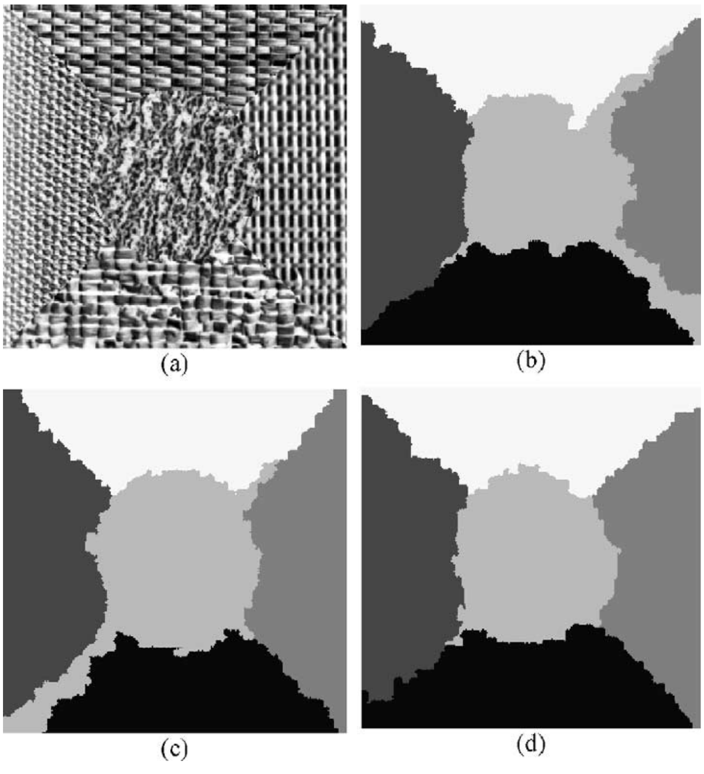

Fig. 5. Results of mosaic #1 in test set. Original image (a) and result using 4 channels (b), 6 channels (c) and 10 channels (d).

Table 2
Classification errors for the proposed algorithm and results published in (Randen and Husoy, 1999) using cooccurrence parameters and best results with any parameter

	Im 1	Im 2	Im 3	Im 4	Im 5	Im 6	Im 7	Im 8	Im 9	Im 10	Im 11	Im 12
Proposed algorithm	7.1	10.7	12.4	11.6	14.9	20.0	18.6	12.0	15.3	1.2	1.1	1.7
Cooccurrence in (Randen and Husoy, 1999)	9.9	27.0	26.1	51.1	35.7	49.6	55.4	35.3	49.1	1.9	4.8	3.3
Best result in (Randen and Husoy, 1999)	7.2	18.9	20.6	16.8	17.2	34.7	41.7	32.3	27.8	0.7	0.2	2.5

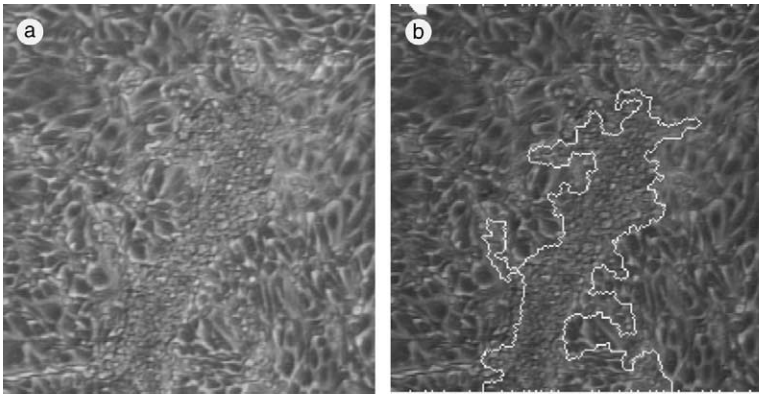

Fig. 6. Segmentation of dead cells in a cell culture image. (a) Original image and (b) image with border of region superimposed.

Table 2 shows the percentage of incorrectly classified pixels on the whole test set. For comparison purposes, the best results and the results on similar features (cooccurrence matrix based) published in (Randen and Husoy, 1999) are also shown. It can be seen that the proposed multi-channel algorithm provides better classification results in almost every case. It should be noted, however, that both results are not directly comparable, as Randen and Husoy (1999) uses a learning vector quantizer supervised classifier.

We have also tested our algorithm on real images. In (Santos et al., 2001) a method was proposed to discriminate regions of dead and of live cells in epithelial cell cultures based on visual texture differences. Live cells are bigger, while dead cells loose their original shape. Segmentation was carried out by discriminant analysis of fixed-size (32×32 pixels) regions. Each square region was classified into one of the classes. We have tested the watershed-based segmentation using eight-texture parameters based on the histogram and the cooccurrence matrix, computed on a window of size 9×9 pixels. Results of watershed segmentation and merging are shown in Fig. 6. The multi-channel approach allowed to reduce the window size, providing a segmentation with better spatial resolution.

7. Conclusions

We have proposed an algorithm based on the watershed transform which allows a connected segmentation approach using several texture channels. A preselection of minima and final merging is performed, allowing for very efficient segmentations. Of the three multivariate merging criteria tested, the best results were obtained using Hotelling's T^2 test. Merging of the smallest region until an intermediate number of regions is obtained gave the same segmentation accuracy while allowing a meaningful decrease in computing time.

The quality of the segmentation depends on the classification efficiency of the texture features used. Although we have tested the method using histogram and cooccurrence matrix parameters only, the method can be applied with any feature extraction algorithm. The segmentation resolution depends on the window size used to compute the texture feature maps. The optimum size can be determined in the feature selection phase.

Apart from the final number of regions, the only parameter of the process is the value of the dynamics for minima selection. It can be chosen automatically as a function of the dynamic range of the image, and results have shown to be robust with respect to this selection.

Acknowledgements

This work has been partially founded by European Commission (COST B11) and Comunidad de Madrid (III PRICIT). Cell images were kindly provided by Hospital G.U. Gregorio Marañón (Madrid, Spain). The authors wish also to thank Dr. Manuel Desco for fruitful discussions on statistical analysis.

References

Beucher, 1994. Watershed, hierarchical segmentation and waterfall algorithm. In: Dougherty, E. (Ed.), Mathematical Morphology and its Applications to Image Processing. Kluwer, Boston.

Beucher, S., Meyer, F., 1990. Morphological segmentation. J. Vis. Comm. Image Rep. 1 (1), 21–45.

Beucher, S., Meyer, F., 1993. The morphological approach to segmentation: the watershed transformation. In: Dougherty, E. (Ed.), Mathematical Morphology in Image Processing. Marcel Dekker, New York.

Chellappa, R., Chatterjee, S., 1985. Classification of textures using gaussian Markov random fields. IEEE Trans. Acoustics, Speech Signal Process. 33 (4), 959–963.

Corneloup, G., Moysan, J., Magnin, I.E., 1996. BSCAN image segmentation by thresholding using cooccurrence matrix analysis. Pattern Recognit. 29 (2), 281–296.

Cumani, A., 1991. Edge detection in multispectral images. Comput. Vision Graphics Image Process.: Graphical Models Image Process. 53, 40–51.

DiZenzo, S., 1986. A note on the gradient of a multi-image. Comput. Vision Graphics Image Process. 33, 116–125.

Haralick, R., 1979. Statistical and structural approaches to texture. Proc. IEEE 67 (5), 786–804.

Haris, K., Efstratiadis, S., Maglaveras, N., Katsaggelos, A., 1998. Hybrid image segmentation using watersheds and fast region merging. IEEE Trans. Image Process. 7 (12), 1684–1699.

Hill, P.R., Canagarajah, C.N., Hull, D.R., 2002. Texture gradient based watershed segmentation. IEEE Internat. Conf. Acoustics, Speech, Signal Process. 4, 3381–3384.

Jones, G., 1994. Image segmentation using texture boundary detection. Pattern Recognition Lett. 15, 533–541.

Lee, H., 1991. Detecting boundaries in a vector field. IEEE Transactions on Image Processing 39 (5), 1181–1194.

Maes, F., Vandermeulen, D., Suetens, P., Marchal, G., 1995. Automatic image partitioning for generic object segmentation in medical images. In: 14th International Conference on Information Processing in Medical Imaging, pp. 215–226.

Marcotegui, B., Crespo, J., Meyer, F., 1995. Morphological segmentation using texture and coding cost. In: IEEE Workshop on Nonlinear Signal and Image Processing, pp. 246–249.

Muzzolini, R., Yang, Y., Pierson, R.A., 1993. A multiresolution texture segmentation approach with application to diagnostic ultrasound images. IEEE Transactions on Medical Imaging 12 (1), 108–123.

Najman, L., Schmitt, M., 1996. Geodesic saliency of watershed contours and hierarchical segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (12), 1166–1173.

Ojala, T., Pietikäinen, M., 1999. Unsupervised texture segmentation using feature distributions. Pattern Recognition 32, 477–486.

Ojala, T., Valkealahti, K., Oja, E., Pietikinen, M., 2001. Texture discrimination with multidimensional distributions of signed gray-level differences. Pattern Recognition 34, 727–739.

Randen, T., Husoy, J.H., 1999. Filtering for texture classification: A comparative study. IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (4), 291–310.

Santos, A., Ramiro, C., Desco, M., Malpica, N., Tejedor, A., Torres, A., Ledesma-Carbayo, M.J., Castilla, M., Garcia-Barreno, P., 2001. Automatic detection of cellular necrosis in epithelial cell cultures. In: Proc. SPIE Medical Imaging 2001: Image Process., Vol. 4322, pp. 1836–1844.

Shafarenko, L., Petrou, M., Kittler, J., 1997. Automatic watershed segmentation of randomly textured color images. IEEE Trans. Image Process. 6 (11), 1530–1544.

Sijbers, J., Scheunders, P., Verhoye, M., Linden, A.V.D., Dyck, D.V., Raman, E., 1997. Watershed segmentation of 3D MR data for volume quantization. Magnetic Resonance Imaging 15 (6), 679–688.

Unser, M., 1995. Texture classification and segmentation using wavelet frames. IEEE Trans. Image Process. 4, 1549–1560.

Vincent, L., 1993. Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. IEEE Trans. Image Process. 2 (2), 176–201.

Vincent, L., Soille, P., 1991. Watersheds in digital space: An efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Machine Intell. 13 (6), 583–598.

Weldon, T.P., Higgins, W.E., Dunn, D.F., 1996. Efficient Gabor filter design for texture segmentation. Pattern Recognit. 29 (12), 2005–2015.