Transactions on
Case-Based Reasoning
Vol.2, No 1 (2008) 41 - 53
© ISSN: 1867-366X (Journal),
ISBN: 978-3-940501-08-0,
IBaI Publishing ISSN 1864-9734

ibai Publishing
www.ibai-publishing.org

A Study on the Case Image Description for Learning the Model of the Watershed Segmentation

Anja Attig and Petra Perner

Institute of Computer Vision and Applied Computer Sciences, IBAI, Leipzig, Germany
pperner@ibai-institut.de,
www.ibai-institut.de

Abstract. Many image analysis methods need a lot of parameters that have to be adjusted to the particular image in order to achieve the best results. Therefore, methods for parameter learning are required that can assist a system developer in building a model. This task is usually called meta-learning. One problem in meta-learning is to describe the properties of the input so that it can be properly mapped to the parameters. In this paper, we consider this task for image segmentation based on the watershed transformation. We use Case-Based Reasoning to control the parameter selection process. Our previous investigation on the theoretical and implementation aspects of the watershed transformation allowed us to draw conclusions for suitable image descriptions. Four different descriptions have been considered based on: statistical and texture features; marginal distributions of columns, rows, and diagonals; similarity between the regional minima; and central moments. The two descriptions based on statistical and texture features and on central moments resulted to be the best ones for segmentation based on watershed transformation. They can best separate the cases into groups having the same segmentation parameters and work nicely also for rotated and rescaled images.

1 Introduction

The aim of image processing is to develop methods for automatically extracting from an image or a video the desired information. The developed system should assist a user in processing or understanding the content of a complex signal, such as an image. Usually, an image consists of thousands of pixels. This information can hardly be quantitatively analyzed by the user. In fact, some problems related to the subjective factor or to the tiredness of the user arise, which may influence reproducibility. Therefore, an automatic procedure for analyzing an image is necessary.

Although in some cases it might make sense to process a single image and to adjust the parameters of the image processing algorithm to this single image manually, mostly the automation of the image analysis makes only sense if the developed methods have to be applied to more than one single image. This is still an open problem in image processing. The parameters involved in the selected processing method have to be adjusted to the specific image. It is often hardly possible to select the parameters for a class of images in such a way that the best result can be ensured for all images of the class. Therefore, methods for parameter learning are required that can assist a system developer in building a model [1] for the image processing task.

While the meta-learning task has been extensively studied for classifier selection it has not been studied so extensively for parameter learning. Soares et. al [2] studied parameter selection for the identification of the kernel width of a support-vector machine, while Perner [3] studied parameter selection for image segmentation.

The meta-learning problem for parameter selection can be formalized as following: For a given signal that is characterized by specific signal properties A and domain properties B find the parameters P of the processing algorithm that ensure the best quality of the resulting output signal/information:

$$f : A \cup B \rightarrow P \tag{1}$$

Meta-data for images may consist of image-related meta-data (gray-level statistics) and non-image related meta-data (sensor, object data) [4]. In general the processing of meta-data from signals and images should not require too heavy processing and should allow characterizing the properties of the signal that influence the signal processing algorithm.

The mapping function f can be realized by any classification algorithm but the incremental behavior of Case-Based Reasoning (CBR) fits best to many data/signal processing problems, where the signal-class cannot be characterized ad-hoc since the data about the signal appear incrementally. The right similarity metric that allows mapping data to parameter-groups and, as consequence, allows obtaining good output results should be more extensively studied. Performance measures that allow to judge the achieved output and to automatically criticize the systems performance are another important problem [5].

Abstraction of cases to learn domain theory would allow better understanding the behavior of many signal processing algorithms that are not anymore to be described by means of standard system theory [6].

The aim of our research is to develop methods that allow us to learn a model for the desired task from cases without heavy human interaction (see Fig. 1). The specific emphasis of this work is to develop a methodology for finding the right image description for the case that group similar images in terms of parameters within the same group and map the case to the right parameters in question.

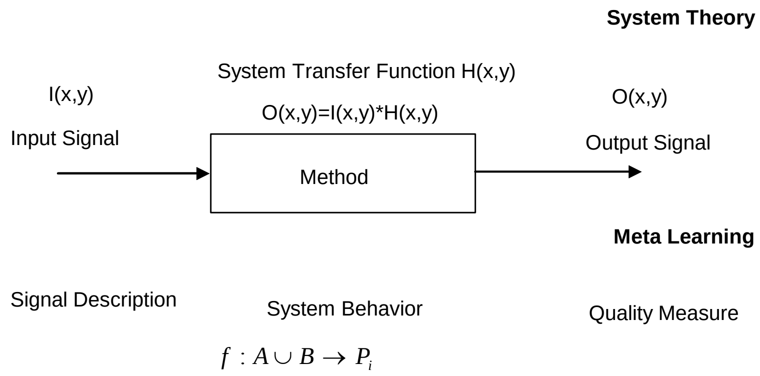


Fig. 1. Problem description in modeling.

We recently investigated [7] the theoretical and implementation aspects of the watershed transformation, which allowed us to draw conclusions for suitable image descriptions. Four different image descriptions have been considered, respectively based on: statistical and texture features; marginal distributions of columns, rows, and diagonals; similarity between regional minima; and central moments.

The basic idea for meta-learning with case-based reasoning for image processing tasks is also introduced in the same Section. In Section 2, we briefly describe the segmentation based on watershed transformation and on the use of Case-Based Reasoning, and point out that the behavior of the watershed transformation may influence the result of the segmentation when the same image is processed after rotation or scaling. The test images and the corresponding best segmentation parameters are given in Section 3, where also the problems concerning the evaluation of the results are briefly addressed. Out of the four different image descriptions that we have considered, the most promising two descriptions, namely those based on statistical and texture features, and on central moments, are described in Section 4. Results are given in Section 5 and the conclusions are given in Section 6.

2 Watershed Transformation based on CBR

Many segmentation algorithms based on watershed transformation have been developed (for a survey, see e.g., [8]). As for the basic watershed transform algorithm, we implemented it according to the approach suggested by Vincent and Soille [9]. For the oversegmentation reduction process we followed the approach by Frucci [10] and transformed the crisp rules into a CBR-approach [11], [12] that made the whole process more flexible.

The conventional watershed algorithm usually produces over-segmentation, which is reduced by combining an iterative computation of the watershed transform with processes called digging and flooding [10]. Flooding merges a non-significant basin to adjacent basins by suitably increasing the gray level of the bottom of the non significant basin. In this way, when the watershed transform is newly computed, no regional minimum is found within the non-significant basin. Digging merges a partially significant basin to specific adjacent regions. The basin A, regarded as partially non-significant, is merged with an adjacent basin B, by digging a canal into the watershed line separating A and B, to prevent that a regional minimum is found in A when applying again the watershed transformation. As a result of merging, the number of local minima found at each iteration diminishes. Flooding and digging and watershed computation are iterated until only significant basins are left.

In order to determine if a basin *X* has to be merged to a basin *Y*, Frucci et al. [11] perform the following check:

$$\frac{1}{2}\left(a \cdot \frac{SA_{XY}}{At} + b \cdot \frac{D_{XY}}{Dt}\right) \geq T \text{ , with } a,b,T \geq 0 \tag{2}$$

where *At* and *Dt* are threshold values (for their automatic computation see [10]), a, b and T are constants, SA_{XY} is a similarity parameter that is the difference between the regional minima of *X* and *Y,* and D_{XY} is the relative depth of the basin *X* with respect to the adjacent basin *Y* (for more details see [10], [11]).

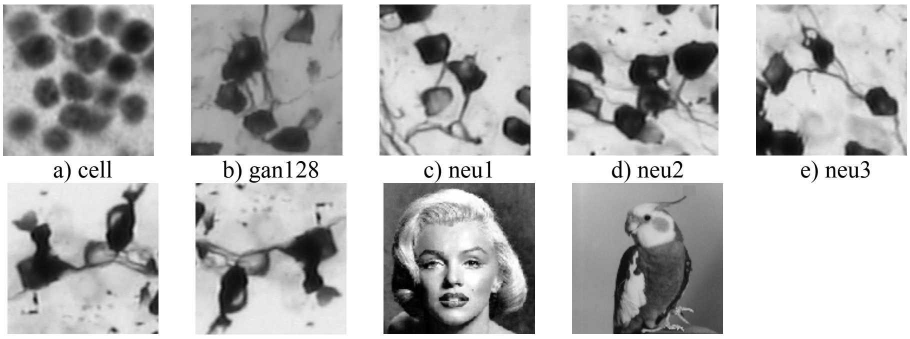
We are interested in analyzing the image properties in order to detect the proper values for the constants *a*, *b* and *T*. The constants *a* and *b* control the influence of the similarity parameter and the relative depth. *T* can be regarded as a threshold. If *T* is 0 and $a,b \geq 0$, then we obviously get the same segmentation like the one produced by the conventional Vincent-Soille algorithm [9]. If we choose in our CBR-based watershed algorithm the parameters *a,b =2* and *T*=1 we obtain often similar results as those obtained by using the crisp rule-based algorithm described in [10].

It is essential to study the true behavior of the used Watershed algorithm and implementation in order to build a general image segmentation model. The behavioral aspects of the Vincent-Soille algorithm [9] for the watershed transformation that influence the segmentation results have been studied in detail in [7]. In summary, we can say that some of the Watershed algorithms are not invariant for image rotation and scaling due to the dependence of the Watershed algorithm from visitin order of the pixels. The detailed theoretical description and demonstration of this can be found in [7]. Since the basic Watershed algorithm is rotation and scaling dependent, the Watershed algorithm based on Case-Based Reasoning is also dependent on rotation and scaling. The best values of a, b and T can be different for two images after scaling or rotation. Other problems are likely to arise, since the watershed lines may be missing or may be too thick. Thus, we have to find an image description that can take into account the behavior of the algorithm.

Making a compromise between computational cost and quality of the obtained segmentation results, we have finally opted for the Vincent-Soille algorithm as basis for CBR-based watershed algorithm. In future work, we will carry out further tests on the different behavior of the basic watershed algorithms. The hope is that the choice of the basic algorithm can be included as parameter into a CBR-based watershed algorithm.

3 Test Images and Parameters for Watershed Segmentation

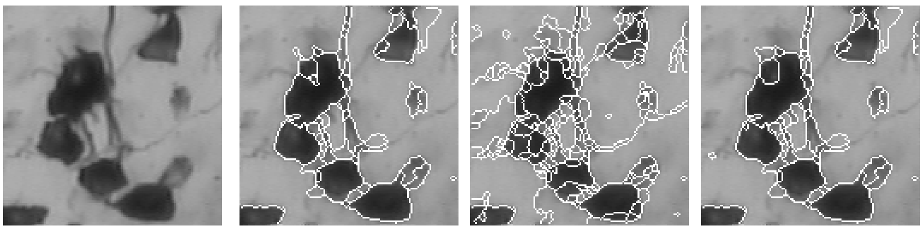
For our study we used nine images of different types (biological images, faces and animals, see Fig. 2a-j). The images {neu1,neu2,neu3,neu4,neu4_r180} shown in Fig. 2 c-h belong to the same class, except that neu4_r180 is the 180 degree rotated image of neu4 and neu4 is at slightly larger scale with respect to the other four images.



f) neu4 h) neu4_r180 i) monroe j) parrot

Fig. 2. Images used for the study

The parameters for the watershed segmentation have been obtained by running the algorithm a number of times, adjusting the parameters until the result has the best segmentation quality. For example, in Fig. 3 three different segmentation results are shown for the image gan128, obtained with different selections of the parameters. N=67 basins are obtained in the best segmentation.



a) gan128 b) (0.75, 0.75, 1) N=67 c) (0.75, 2, 1) N=211 d) (1.75,0.75,1) N= 73

Fig. 3. Influence of the selected parameter-set to the segmentation result for image gan128.

The parameters corresponding to the best segmentation results for the nine test images in Fig. 2 are shown in Table 1.

Note, that in Table 1 only two of the neuron images have the same parameters (neu1 and neu3). Actually, neu4 contains cells slightly larger than those in the other neuron images and cannot be segmented by using the same parameters of the other images, because the Vincent-Soille algorithm is not invariant with respect to scaling. Moreover, image neu2 differs from neu1 and neu3 by the presence of a larger number of cells, which justifies a different set of parameters.

Image	a	b	T
monroe	0,5	2	1
gan128	0,75	0,75	1
parrot	0,75	2	1
cell	1	1	1
neu1	2	1	1
neu2	0,25	1	1
neu3	2	1	1
neu4	0,75	2	1
neu4_r180	1	0,5	1

Table 1. Segmentation parameters a, b, and T for the test images shown in Fig. 2

A problem related to the determination of the segmentation parameters for the CBR-based watershed algorithm is how to judge the best segmentation quality. Evaluation done by humans is subjective and can result in differently segmented versions of the same input image. For example, two humans asked to select the preferred segmentation between three possible results (shown in Fig. 3b-d) gave opposite answers (see Fig. 3b-c). In turn, the best segmentation automatically obtained by comparing the watershed lines of the segmented image to the edge image, generated by the Prewitt-Operator [13], of the input based on similarity procedure described in [11] is shown in Fig. 3d. An automatic evaluation of the segmentation results is really necessary, but even an automatic evaluation may have some weakness.

The automatically computed binarized gradient image is only an auxiliary method for getting the true segmented image (gold standard) to which we have to compare. The gold standard could be obtained by manually labeling the regions of the input image, which is not appropriate.

4 Elicitation of Image Descriptions and Assessment of Similarity for Watershed Transform

The aim of the image description is to find out among a set of images the group of images that need the same processing parameters to achieve the best segmentation results. To give an example, the images neu1 and neu3 should be grouped together in one group based on the best parameters a, b, and T (see Table 1) and the images neu4 and parrot should be grouped together in another group.

We consider different image descriptions in our study that should allow us to group images based on the image features and by doing this to learn a model for image segmentation by samples.

Cases are composed normally of

- non-image information
- features specifying image characteristics, and
- parameters for solution (image segmentation parameters).

Non-image information is different depending on the application. In our study we use images from different domains, such as biological images, faces and animals. Our aim is to describe image similarity only by general image features. Hence our cases are composed of

- features specifying image characteristics, and
- parameters a, b, T for segmentation.

Images, which are classified as being similar based on the image features, should be segmented with the same parameters of the segmentation algorithm, which should produce the best segmentation for any of them.

To understand which are the features to be used, we resort to hierarchical clustering. This gives us a graphical representation of the different image groups. Single linkage is used to show outlier while the distance between two classes is defined as the minimal.

The question is: What are the right image features that allow us to map the images to the proper image segmentation parameters for the watershed transformation?

The image description should reflect the behavioral approach of the watershed transformation with respect to the particular image characteristics. Therefore, we studied the theoretical details and the implementation limits of the watershed transformation in [7] to get insights into this question. Based on this work we decided to test four image descriptions based on:

- Statistical and Texture Features,
- Marginal Distribution for Columns, Rows, and Diagonals,
- Similarity between the Regional Minima, and
- Central Moments.

The details of this study are given in [7]. The two most promising descriptions, i.e., those based on Statistical and Texture Features, and Central Moments are discussed in this paper.

4.1 Image Description based on Statistical and Texture Features

According to Perner [3], who used this description for meta-learning the parameters for a CBR-based image segmentation model, we used statistical features (centroid, energy, entropy, kurtosis, mean, skewness, variance and variation coefficient) and textures feature (energy, correlation, homogeneity, contrast) for case description. The input image is the gradient image of the original image, since the watershed transformation works on that image. First results on this image description are reported in Frucci et al. [11]. The texture features have been chosen to describe the particular distribution of the regional minima in an image, while the statistical features describe the signal characteristics.

Like in Perner [3], the distance between two images A and B is computed as:

$$dist_{AB} = \frac{1}{k}\sum_{i=1}^{k}\omega_i\left|\frac{C_{iA}-C_{iB}}{C_{i\max}-C_{i\min}}\right|, \tag{3}$$

where k is the number of features in the data base, C_{imax} and C_{imin} are the maximum and minimum value of the ith feature for all images in the data base, C_{iA} (C_{iB}) is the value of the ith feature for image A (B) and the ω_i are weights. For the weights it is:

$$\sum_{i=1}^{k} \omega_i = 1. \tag{4}$$

where $\omega_i = 1/k \ \forall i \in [1,\ldots,k]$.

Results are reported in Fig.4.

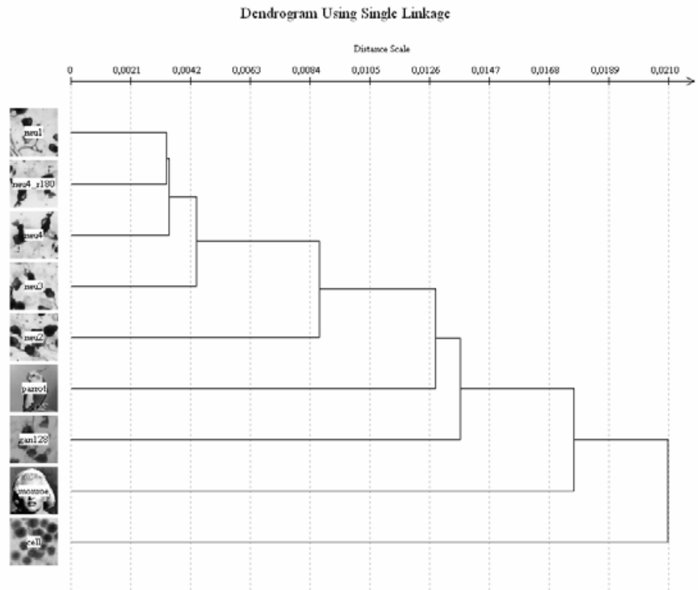


Fig. 4. Dendrogram for image description based on texture and statistical features.

If we virtually cut the dendrogram by the cophenetic similarity of 0.0043 we obtain the groups G1={neu4, neu4_r180, neu1, neu3}, G2={neu2}, G3={parrot}, G4={gan128}, G5={monroe}, and G6={cell}.

The images neu4 and the image neu4_r_180 (which is the 180 degree rotated version of neu4) are grouped into the same group, although they have completely different segmentation parameters. We obtained the best result for neu4 with the parameter-set a=0.75, b= 2 and T=1, while for neu4_r180 the best segmentation was obtained with the parameter-set a=1, b=0.5 and T=1. By using the latter parameter-set for neu4, we would get an undersegmented result. Overall we observe that not all images having the same image segmentation parameters are grouped into one group such as neu4 and parrot.

To sort out rotated images from the group that includes the un-rotated images we have to give more emphasis to the feature centroid, because this features is only one which is not invariant for rotations. Therefore, we divide the image features set into three groups: texture features, centroid, and the remaining statistical features. Each group gets a total weight ω_g of 1/3. The local weights ω_{gi} in each group are computed as follow

$$\omega_g = \sum_{i=1}^{l} \omega_{gi} = \sum_{i=1}^{l} \frac{1}{3*l} = \frac{1}{3}, \quad (5)$$

where l is the number of features in the group.

In the resulting dendrogram the images neu4 and neu4_r180 are clustered in different groups (see Fig. 5).

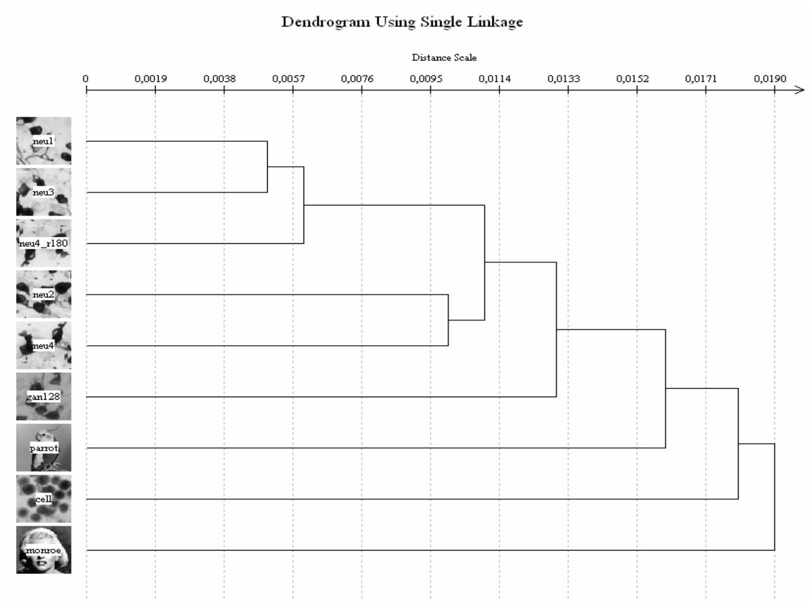


Fig. 5. Dendrogram for image description based on weighted texture, centroid and statistical features.

If we virtually cut the dendrogram by a cophenetic similarity value of 0,0057 then we obtain the following groups G1={neu1, neu3}, G2={neu4_r180}, G3={neu2}, G4={neu4}, G5={gan128}, G6={parrot}, G7={cell}, and G8={monroe}. Except for the images neu4 and parrot, these groups seem to reflect better the relationship between the image description and the parameter-set. The proper weighting of the features can improve the grouping of the images for which we need to work out a strategy in a further study.

4.2 Image Description based on Central Moments

An image can be interpreted as a two-dimensional density function. So we can compute the geometric moments:

$$M_{pq} = \int\int x^p y^q g(x,y) \quad p,q = 0,1,2,\ldots \tag{6}$$

with continuous image function *g(x,y)*.

In the case of digital images, we can replace the integrals by sums:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y) \quad p,q = 0,1,2,\ldots \tag{7}$$

where *f(x,y)* is the discrete function of the gray levels.

The seven moment invariants from Hu [14] have the property to be invariant under translation, rotation and scale. Since our implementation of algorithm [11] is not invariant with respect to rotation and scale, the invariant moments are unsuitable for our image description.

We can consider only the central moments, which are translation invariant:

$$m_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q f(x,y) \quad p,q = 0,1,2,\ldots \tag{8}$$

where

$$x_c = \frac{\sum_x \sum_y x f(x,y)}{\sum_x \sum_y f(x,y)} \text{ and } y_c = \frac{\sum_x \sum_y y f(x,y)}{\sum_x \sum_y f(x,y)}. \tag{9}$$

For our study, we use central moments m_{pq} with *p* and *q* between 0 and 3 and as input image the binarized gradient image obtained with the thresholding algorithm of Otsu [15]. To determine the similarity between two images we use the normalized city-block distance in equation (3), where $\omega_i = 1/k \ \forall i \in [1,\ldots,k]$. Fig. 6 shows the dendrogram of this test.

If we virtually cut the dendrogram by the cophenetic similarity measure of 0.0129, we obtain the following groups G1={Monroe, parrot}, G2={neu3, neu4}, G3={gan128}, G4={neu4_r180}, G5={neu2}, G6={neu1}, and G7={cell}. Compared to the groups obtained by using statistical and texture features for image description, we now obtain that neu4_r180 gets separated from neu4. Unfortunately, neu3 is assigned to the same group as neu4.

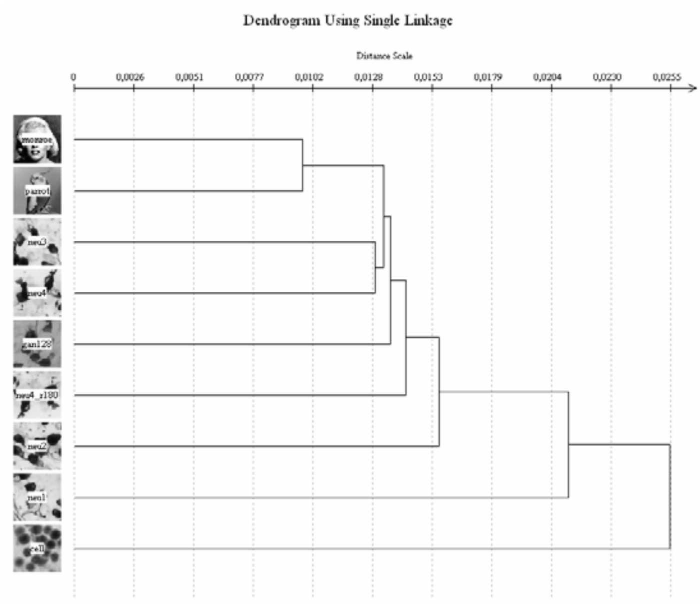
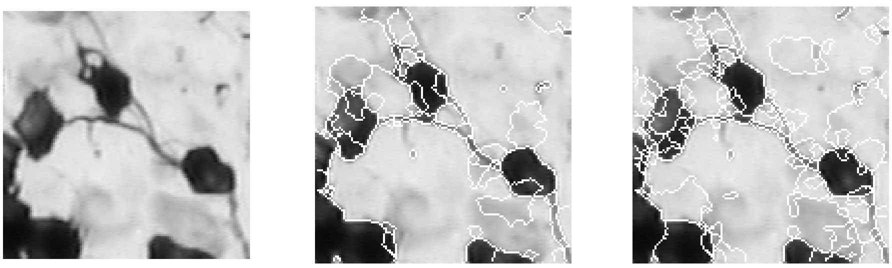


Fig. 6. Dendrogram for CBR based on Central Moments (on binary gradient image).

Fig. 7 shows the segmentation results for neu3 by using the parameters in Table 1 (Fig. 7b), and the parameters selected for neu4 (Fig. 7c). Fig. 7c is more oversegmented then Fig. 7b. Since oversegmentation can be reduced by thresholding short edges, the result suggested by using the Central Moments can be regarded as acceptable.



(2,1,1) N=98 (0.75,2,1) N=180

Fig. 7. The influence of the selected parameter-set on the segmentation result and on the number N of basins for image neu3.

5 Discussion

Out of the four image descriptions based on statistical and texture features; marginal distributions of columns, rows, and diagonals; similarity between regional minima; and central moments, the descriptions working better for the watershed transformation are those based on statistical and texture features (STDescript) and on Central Moments (CMDescript). We have also considered the weighted STDescript. While the STDescript groups images that can be rotated or scaled versions of each other, the weighted STDescript and the CMDescript do not. The obtained groups for the three descriptions are the following:

STDescript
G1={neu4, neu4_r180, neu1, neu3}, G2={neu2}, G3={parrot}, G4={gan128}, G5={monroe}, and G6={cell}

weighted STDescript
G1={neu1, neu3}, G2={neu4_r180}, G3={neu2}, G4={neu4}, G5={gan128}, G6={parrot}, G7={cell}, and G8={monroe}

CMDescript
G1={Monroe, parrot}, G2={neu3, neu4}, G3={gan128}, G4={neu4_r180}, G5={neu2}, G6={neu1}, and G7={cell}.

The groups obtained by means of the weighted STDescript seem to reflect better the relationship between the image characteristics and the segmentation parameters than the groups obtained by the other descriptions. The computation time of the image descriptions is more or less the same. Thus, we can say that the weighted statistical and texture features description is the best image description that we have found so far during our study.

6 Conclusions

The aim of our work was to find an image description that characterizes each particular image with respect to the behavior of the image segmentation method. The image description should allow retrieving the best possible segmentation parameters. We first studied the theoretical and implementation aspects of the watershed transformation in order to draw conclusions for the image description.

The watershed transformation produces different results if the image is rotated or rescaled. The particular implementation of the algorithm puts constraints on the behavior of the algorithm. As result of our study we concluded that we need an image description that describes the distribution of the regional minima and that is not invariant against rotation and scaling.

We studied four different image descriptions, respectively based on: statistical and texture features; marginal distributions of columns, rows, and diagonals; similarity of regional minima; and central moments.

Two of the above four image descriptions did not lead to any success. The description based on statistical and texture features is useful, but is invariant under rotation and scaling. The best image descriptions is the description based on weighted statistical and texture features and the description based on central moments. This image description seems to well represent the relationship between the image characteristics of the particular image and the segmentation parameters. Cases having the same segmentation parameters could be grouped based on the image description into the same group. This will make generalization over these groups of cases possible, which is expected to lead to a complete image segmentation model. However, how to automatically choose the proper weight values for the statistical and texture features is still an open problem.

References

[1] Perner, P.: Why Case-Based Reasoning is Attractive for Image Interpretation. In: Aha, D., Watson, I. (Eds.), Case-Bases Reasoning Research and Developments. LNAI, Vol. 2080, 27-44, Springer Verlag, Heidelberg, 2001.

[2] Soares, C., Brazdil, P.B., Kuba,P.: A meta-learning method to select the kernel width in support vector regression. Machine Learning, 54, 195–209, 2004.

[3] Perner, P.: An Architecture for a CBR Image Segmentation System. Engineering Applications of Artificial Intelligence, Vol. 12 (6), 749-759, 1999.

[4] Perner, P.: Case-Based Reasoning for Image Analysis and Interpretation. In: Chen, C. and Wang, P.S.P. (Eds.): Handbook on Pattern Recognition and Computer Vision, 3rd Edition, 95-114, World Scientific Publisher, 2005.

[5] Perner, P.: Case-based reasoning and the statistical challenges. Journal Quality and Reliability Engineering International, Vol. 24 (6), 705-720, 2008.

[6] Wunsch, G.: Systemtheorie. Akademische Verlagsgesellschaft, Leipzig, 1975.

[7] Attig, A., Perner, P.: A Study on the Case Image Description for Learning the Model of the Watershed Segmenation. IBaI Report, May 2009, ISSN 1431-2360.

[8] Roerdink, J.B.T.M., Meijster, A.: The watershed transform: definitions, algorithms and parallelization strategies. Fundamenta Informaticae, 41, 187-228, 2001.

[9] Vincent, L., Soille, P.: Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans. Patt. Anal. Mach. Intell.,13, 6, 583-598, 1991.

[10] Frucci. M.: Oversegmentation Reduction by Flooding Regions and Digging Watershed Lines. International Journal of Pattern Recognition and Artificial Intelligence World Scientific, Singapore, 20, 1, 15-38, 2006.

[11] Frucci, M., Perner, P., Sanniti di Baja, G: Case-based Reasoning for Image Segmentation by Watershed Transformation. In: Perner, P. (Eds.): Case-Based Reasoning on Images and Signals, SCI, Vol. 73, 319-353,Springer, Berlin, 2008.

[12] Frucci, M., Perner, P., Sanniti di Baja, G.: Case-based Reasoning for Image Segmentation. International Journal of Pattern Recognition and Artificial Intelligence, Vol. 22 (5), 1-14, 2008.

[13] Prewitt, J.M.S: Object enhancement and extraction. In. Lipkin, B.S., Rosenfeld, A. (Eds.): Picture Processing and Psychopictorics, Academic Press, 75-149, New Yoek, 1970.

[14] Hu, M.K.: Visual pattern recognition by moment invariants. IEEE Trans. Information Theory, Vol. 8, 179-187, 1962.

[15] Otsu, N.: A threshold selection method from gray level histograms. IEEE Trans. Syst. Man Cybern. SMC-9, 62–66,1979.