

Multivariate Watershed Segmentation of Compositional Data

Michael Hanselmann¹, Ullrich Köthe¹, Bernhard Y. Renard¹, Marc Kirchner¹,
Ron M. A. Heeren², Fred A. Hamprecht^{1*},**

¹ Heidelberg Collaboratory for Image Processing, University of Heidelberg, Germany

² FOM-Institute for Atomic and Molecular Physics, Amsterdam, The Netherlands

Abstract. Watershed segmentation of spectral images is typically achieved by first transforming the high-dimensional input data into a scalar boundary indicator map which is used to derive the watersheds. We propose to combine a Random Forest classifier with the watershed transform and introduce three novel methods to obtain scalar boundary indicator maps from class probability maps. We further introduce the multivariate watershed as a generalization of the classic watershed approach.

1 Introduction

The watershed transform [1] is a region-based segmentation algorithm for gray-scale images and is a popular method in image segmentation [2]. Figuratively speaking, the gray-valued boundary indicator image is considered as a height map which is flooded with water. Whenever two water basins that originate from different local minima meet, a watershed is constructed. Various definitions for the continuous and discrete case have been given, most of which operate on scalar-valued input. Typically the gradient is used as a boundary indicator, featuring high values at border locations and low values in homogeneous areas.

High-dimensional data is typically transformed into a scalar boundary map, such that the conventional watershed can be applied. A direct generalization for color images is the color gradient. Alternatively, the watershed transform is individually performed on each color band and the obtained segmentation results are integrated into a single segmentation map [3]. Some authors have suggested to use color models that feature an intensity channel [4, 3]. However, most of these approaches are based on the assumption that color channels are highly correlated. This is typically not the case for more complex data, e.g. in remote sensing [5], the analysis of medical data [6] and data acquired in imaging mass spectrometry experiments [7]. In fluorescence microscopy, experimentalists often deliberately select dyes that highlight different, uncorrelated structures [8].

To analyze such kind of data, authors have used (weighted) channel-wise gradients [9–11] or the metric-based gradient [11]. Noyel [11] proposed the sum

* corresponding author, fred.hamprecht@iwr.uni-heidelberg.de

** We thank Erika R. Amstalden and Kristine Glunde for providing data.

or supremum of channel-wise morphological gradients that are computed as difference between channel dilation and erosion. Malpica [12] and Karvelis [13] construct the Jacobian matrix J of partial derivatives (in direction of the two spatial axes) in each pixel and calculate the Eigenvalues of $J'J$. The difference between the two Eigenvalues is used as a boundary indicator. Zhang [14] uses the spectral angle between a pixel’s underlying spectrum and a reference spectrum to obtain a scalar boundary indicator. Angulo proposed a stochastic watershed algorithm [15] which was later extended to multispectral images [16]. Soille [17] combines a histogram-based clustering with shape priors and pixel-wise gradients which are used as input for the classic watershed transform. Authors have also suggested dimensionality reduction [11, 16, 6] to reduce the effect of noise.

In many segmentation scenarios, more than two classes are of interest (e.g. background, healthy cells, diseased cells, ...). Often, prior knowledge on the composition of the data exists which can be used to assign class probabilities to pixels. This way, the dimensionality of the data is significantly reduced. Such data, where the probability vector in each pixel sums to one, is termed *compositional* [18]. In this paper, we propose three novel methods to obtain a scalar boundary indicator based on such compositional data, and we introduce the multivariate watershed which is a multi-class generalization of the classic watershed transform. We quantitatively compare the presented methods to three algorithms that previously showed good performance [11, 12]: Noyel’s [11] sum/supremum of channel-wise morphological gradients and Malpica’s [12] tensor approach.

2 Watershed segmentation of compositional data

The discrete and sequential watershed methods on scalar-valued input can be grouped into watershed by immersion [19] and watershed by topographic distance [20]. In this paper we argue in terms of the latter, however our results also carry over to the former setting. Watershed by topographic distance is based on gradient descent. An efficient algorithm proceeds as follows: For each pixel we store a reference to the neighboring pixel with the maximum slope

$$LS(u) = \max_{v \in \text{neighbors}(u) \cup \{u\}} \frac{f_u - f_v}{d(u, v)} \quad (1)$$

where f_u is the gray value at pixel u and $d(u, v)$ is the distance between pixels u and v . Minima pixels point towards themselves and a distinct label is assigned to each local minimum. Starting from the minima, the constructed paths are traversed in reverse order and all pixels are labeled with the corresponding label. In this implementation, the resulting watersheds lie *between* pixels (inter-pixel boundaries) and the algorithm has linear complexity in the number of pixels.

In the following, we work with spectral data in two spatial dimensions. Note, however, that the methods presented below are also applicable to other high-dimensional data. Let $S = \{(\mathbf{x}^1, y^1), \dots, (\mathbf{x}^K, y^K)\}$ be a set of available M -dimensional training samples, i.e. K spectra \mathbf{x}^i with M channels and corresponding class labels $y^i \in L_1, \dots, L_D$, e.g. “cancerous”, “healthy tissue” and

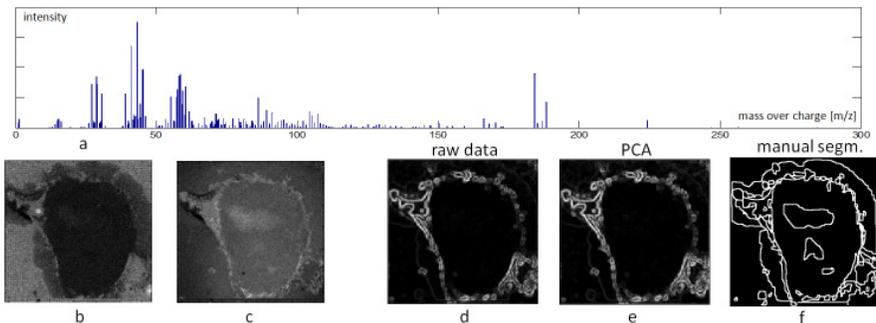


Fig. 1. An excerpt of a typical mass spectrum (a) and 2 channels of the IMS dataset (b,c). Only parts of the images are correlated. A boundary map is obtained with the sum of channel-wise morphological gradients on the raw data (d) or the first few principal components (e), revealing only few boundaries of the manual segmentation (f).

“blood vessels” in medical applications or “faulty” or “intact” in quality control. The Random Forest classifier [21] is a state-of-the-art supervised learning method that features high classification performance [22], allows fast training, can cope with a large number of input variables and is relatively robust to its hyperparameters. We train a Random Forest on S and classify the whole data set comprising N data points to obtain the posterior probability for each of the D classes in each pixel. The resulting data set has D dimensions in each pixel and is compositional since each probability vector sums to one.

We use data from an imaging mass spectrometry [7] (IMS) experiment of human breast cancer grown in mice [23] to illustrate the performance of the different methods. IMS allows the detailed analysis of spatial distributions of molecules in organic samples. In each pixel a mass spectrum was acquired that comprises several thousand mostly uncorrelated channels (see fig. 1a for an example). Chemical staining of an adjacent slice was used to identify 5 different tissue classes, to label parts of the data and obtain an approximate manual segmentation. Random Forest classification hence yields a compositional data set with 5 dimensions in each pixel.

Noyel’s sum of morphological gradients [11] was used to calculate boundary indicator maps directly from the raw input and the first principal components of the raw input (fig. 1d+e) as well as from the probability maps (fig. 2). Visual inspection suggests that the latter contains more information. Therefore, we used probability maps as input for all following boundary map computations.

3 Multivariate gradients

We next present three novel methods that create scalar boundary maps from such compositional input data as well as the multivariate watershed which is a generalization of the classic watershed definition.

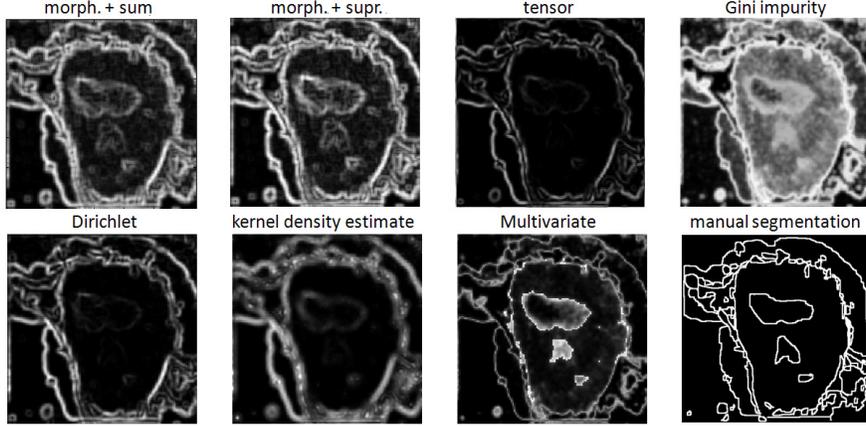


Fig. 2. The probability map-based boundary indicators obtained with Noyel’s sum/supremum of channel-wise morphological gradients, Malpica’s tensor and the methods introduced here. The result of an approximate manual labeling is also shown.

Gini impurity. The underlying idea for the Gini impurity watershed is that the class impurity in the classification results can be used to identify borders between different regions. Gini impurity [21] essentially is a measure of vector sparseness. For probability distribution f_u at pixel location u it is defined by

$$GI(f_u) = 1 - \sum_{j=1}^D f_u(j)^2 \quad (2)$$

where $f_u(j)$ is the j -th value in f_u . The minimum degree of impurity is obtained if one of the $f_u(j)$ equals one. We calculate the Gini impurity index for each pixel and perform the conventional watershed segmentation on the obtained scalar boundary map. Slight smoothing of the probability maps with a channel-wise Gaussian filter (zero mean, unit variance) prior to calculation of the impurity indices preserves the sum-constraint and ensures that the border between two adjacent points with (different) pure components has indeed higher impurity.

Dirichlet boundary indicator. The observed probability vectors at pixel u and its neighbors $O(u)$ can be interpreted as realizations of a (single) Dirichlet distribution [24] (which can be seen as the multivariate generalization of the beta distribution). Its D -dimensional realizations sum to one as do the class probabilities for each pixel. The Dirichlet distribution is parameterized by a vector $\alpha = (\alpha_1, \dots, \alpha_D)$ where $\alpha_j, j = 1, \dots, D$ is positive, i.e. $f \sim Dir$ with

$$f(x_1, \dots, x_D | \alpha_1, \dots, \alpha_D) = \frac{\Gamma\left(\sum_{j=1}^D \alpha_j\right)}{\prod_{j=1}^D \Gamma(\alpha_j)} \prod_{j=1}^D x_j^{\alpha_j - 1} \quad (3)$$

for all $x_j > 0$ with $\sum_j^D x_j = 1$ and Γ is the Gamma function. For given observations $O(u)$ its optimal parameters $\hat{\alpha}(u)$ can be estimated by maximizing the

log-likelihood of the data [24]. This maximum likelihood estimation is performed in a neighborhood of each pixel. The shape of the distribution is determined by the parameters $\hat{\alpha}_j$. At pixel locations within homogeneous regions of the spectral image their sum is high and we obtain highly peaked distributions with low variances. In contrast, in the vicinity of borders the sum of the $\hat{\alpha}_j$ is low and the distributions are broad. Thus, we propose to use the inverted precision – defined as 1 divided by the sum of all $\hat{\alpha}_j(u)$ – as boundary indicator at pixel location u . The resulting boundary indicator map is used as input for the classic watershed.

Kernel density estimate. In this method, in each pixel a kernel density estimation [25] is performed. Here we use Gaussian kernels k , but other choices are also possible. The density *KDE* at a pixel u is calculated from

$$KDE(u) = \frac{c}{N\sigma_{spat}^2\sigma_{prob}^D} \sum_{i=1}^N k_{spat} \left(\left\| \frac{g_u - g_{u^i}}{\sigma_{spat}} \right\|^2 \right) k_{prob} \left(\left\| \frac{f_u - f_{u^i}}{\sigma_{prob}} \right\|^2 \right) \quad (4)$$

where u^i is the i -th pixel in the data set, g_u is the two-dimensional vector of spatial components corresponding to pixel u , and c is a normalization constant. Contributions are weighted by the distance in space (term with k_{spat}) as well as by distance in composition (term with k_{prob}). σ_{spat} and σ_{prob} are the corresponding (Gaussian) kernel bandwidths. The inverse density is used as boundary indicator map for the classic watershed. The density estimation formula in eq. 4 is well known from the literature. It constitutes a link of the watershed algorithm with the mean shift procedure [26] and bilateral filtering [27].

Multivariate watershed. In the classic watershed algorithm, only one type of basin is known. For compositional data, we generalize this by defining one basin type per class, i.e. basin types B_1, \dots, B_D . Each pixel is assigned to the class whose posterior probability dominates. In absence of prior knowledge, dominance is established by a simple “winner takes all” rule (corresponding to a fair arbitrator); but certain classes *can* be favored if desired by introducing bias (corresponding to a bribed arbitrator). We first consider the two-class case (cf. fig. 3): For pixel u the probabilities for classes 1 and 2 are equal to $f_u(1)$ and $f_u(2) = 1 - f_u(1)$. Maximum likelihood estimation (MLE) for the assignment of pixels to basin types corresponds to introducing a threshold $T = 0.5$ where u is assigned to basin type 1, i.e. $w_T(u) = B_1$, if $f_u(1) \leq T$, and to B_2 otherwise. By moving T , different risks can be assigned to the two classes. The basin assignment is then obtained in a risk-weighted maximum a-posteriori decision.

When compositional data of higher dimension is analyzed, the points live on a simplex. In the 3D case, the MLE corresponds to using the threshold $T = [1/3; 1/3; 1/3]$ and the perpendiculars to the lines connecting the corners of the simplex to assign points (see fig. 7). Moving T on the simplex again controls the emphasis of the different classes. If the threshold point is close to one corner of the simplex, the corresponding class will be less influential in the segmentation. In the D -dimensional case, T has D components and the set of pixels that are assigned to basin B_k is given by $Z_k := \{u \mid w_T(u) = B_k\}$, i.e.

$$Z_k = \left\{ u \mid 1 - \frac{f_u(k)}{f_u(k) + f_u(j)} < 1 - \frac{T(k)}{T(k) + T(j)} \forall j \in \{1, \dots, D\} \setminus \{k\} \right\}. \quad (5)$$

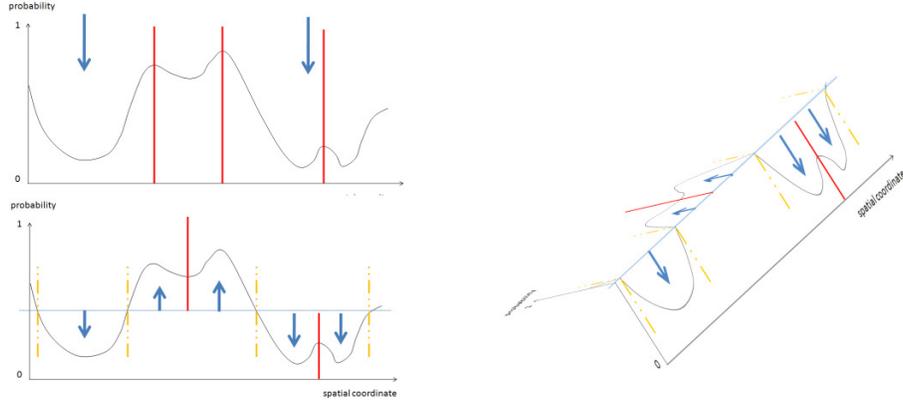


Fig. 3. Visualization of the multivariate watershed for $D = 2$. The classic approach (top left) only has one basin type. Figuratively, rain falls from above and fills the basins. Whenever two basins meet, a watershed is constructed (vertical lines). In the multivariate case (bottom left), a threshold is introduced that controls the assignment of each point to one of the two basin types. Figuratively, the rain now falls from a higher dimension onto a surface that has been folded at the threshold value (right). New watersheds emerge for the parts above the threshold and between basins of different type. Classic watershed emerges as a special case in which the threshold has been set to the maximum intensity.

For each of the D sets of points, the Euclidean distance between each point in the set and the respective basin (i.e. corner of the simplex) is calculated. On the resulting scalar map, a watershed-like segmentation is performed. It differs from the conventional watershed transform in the following way. The basin assignments define areas of influence for each basin type. Water is never allowed to cross borders between different influence zones. This can be included in the conventional watershed algorithm by changing its distance function $d(u, v)$ such that points that have been assigned to different basin types have infinite distance: $d(u, v) = \infty \forall v \in neighbors(u) : w_T(v) \neq w_T(u)$. The threshold T is a parameter which can be set by the user. Its default value is $[1/D; \dots; 1/D] \in \mathbb{R}^D$.

4 Data and Experiments

We used the topographic distance version of the watershed algorithm and a neighborhood system with 8-connectivity. For the real data studied at the end of this section, no exact ground truth is available; for a quantitative evaluation, we hence resort to simulated data which is described next.

Simulation. First, three spectra from a real imaging mass spectrometry measurement were taken and defined as “pure” spectra (see fig. 1a for an example). Then, different mixture maps were generated that contain pure areas as well as impure ones (fig. 4). The “observed” data was created by mixing the pure spectra according to the mixture maps. IMS measurements typically consist of

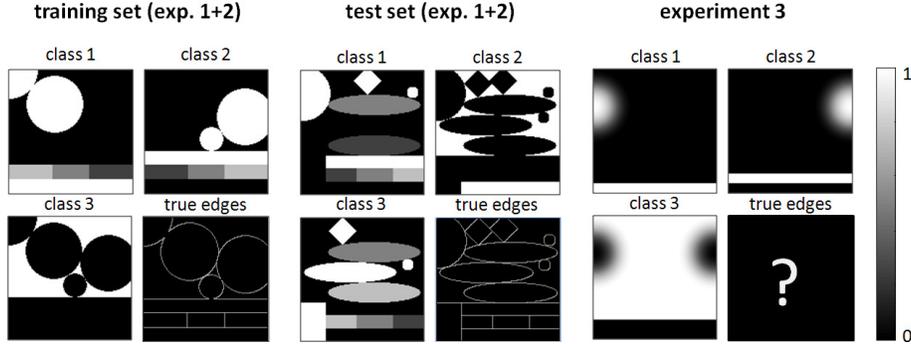


Fig. 4. The ground truth mixture maps for the simulated experiments. In all experiments we used three ground truth spectra and mixed them according to the mixture maps shown above. White areas correspond to a pure concentration, black indicates that this class is absent at the respective location. The correct boundaries - if unambiguous - are also given.

ion counts. A Poisson noise model was used to simulate instabilities in the data acquisition process. Samples from the noisy data were used to train a Random Forest classifier with 250 trees and 100 samples per class that were selected as described below. After training, the whole dataset was classified and a set of probability maps was obtained. A total of three experiments was performed: In the first experiment, the mixture map also contained impure regions, but training was performed with samples from pure regions only (3 classes). For experiment 2 the same setting as in experiment 1 was used, but each mixture area was considered an individual class and the classifier was trained with samples from each of them (6 classes). Finally, the last experiment demonstrates the influence of the threshold of the multivariate watershed on the obtained segmentation.

Real data. We also applied the methods to the imaging mass spectrometry data set described at the end of section 2. This spectral image comprises a total of 4000 mass over charge channels, two of which are shown in fig. 1.

Postprocessing. The segmentation maps obtained with the different watershed methods typically contain oversegmentation. We use watershed dynamics [28, 29] to amend this problem. For each edge a dynamics value is calculated and edges that correspond to a dynamics value below a given threshold are removed, i.e. the respective basins are merged.

Evaluation criteria. The true edges as well as the watersheds obtained with the topographic distance watershed algorithm lie between pixels of the grid. We use inter-pixel edges represented on a half-integer grid to quantitatively compare the estimated edge map E_{est} with the ground truth edge map E_{gt} with the Baddeley distance [30] which is calculated for all pixels u :

$$dist_{Bad}(E_{gt}, E_{est}) = \left[\frac{1}{N} \sum_u (DT(u, E_{gt}) - DT(u, E_{est}))^q \right]^{1/q}. \quad (6)$$

$DT(u, E)$ is the closest distance between pixel u and any of the edges in the edge map E [31]. In our study, q was set to 2 and we only considered distances up to 5 pixels, i.e. pixels that were more than 5 pixels away from any edge in both of the edge maps were ignored. The Baddeley distance penalizes both over- and undersegmentation. Since the optimal dynamics threshold value is different for the various watershed algorithms under consideration, for each method the best dynamics threshold with respect to the Baddeley distance was chosen.

Besides using the Baddeley distance, we quantify the segmentation quality by means of sensitivity and specificity. The former measures which percentage of the true edges are identified by a method, the latter how many background points are wrongly classified as edges. Since the edges in the obtained segmentation maps E_{est_i} can be displaced from their positions in the ground truth edge map E_{gt} by a few pixels, we first match them to the ground truth edges. To this aim, we employ the stable marriage algorithm [32] that uniquely assigns each pixel in E_{est_i} to a close pixel in E_{gt} as long as the maximum distance is below a given threshold (here 2 pixels). Edge pixels in E_{gt} without a “partner” in E_{est_i} are considered false negatives (FN, indicating undersegmentation), edge pixels in E_{est_i} without a partner in E_{gt} are false positives (FP, oversegmentation), pairs are considered true positives (TP) and the rest are true negatives (TN). We then calculate $sensitivity = \frac{TP}{TP+FN}$ and $specificity = \frac{TN}{FP+TN}$. An estimate for the test error is obtained by testing the watershed methods on a independent test image set where we keep the parameter settings. Training and test set differ in geometry and class mixtures (cf. fig. 4).

In case of the kernel density estimation watershed, σ_{spat} and σ_{prob} have to be specified. We calculated the segmentations for a variety of different parameter settings out of a given range and the best settings in the training set were used in the comparison. In experiment 1, the best choice was $\sigma_{spat} = 2.0$ and $\sigma_{prob} = 1.0$, in experiment 2 $\sigma_{spat} = \sigma_{prob} = 1.0$. Similarly, different structuring elements can be used for the calculation of the morphological gradient. We experimented with discs of varying sizes and found a radius of 1 to be most adequate.

5 Discussion

We now discuss the outcome of the three experiments described in section 4.

Experiment 1. Regarding the training set, the results of Malpica’s tensor and especially the Dirichlet approach are very close to the ground truth (cf. fig. 5 and tab. 1). The Dirichlet boundary indicator reliably finds edges between different pure mixtures and impure mixtures and results in straight boundary lines (cf. fig. 6a). The optimal dynamics threshold for the kernel density-based method leads to some oversegmentation in the lower right part of the image, but the remaining part of the image is well segmented. The Gini impurity watershed and the multivariate watershed accurately identify boundaries between pure mixture areas, but have some problems with boundaries separating impure regions (cf. fig. 6b). This task is indeed difficult for all methods since the classifier output shows a relatively low gradient in these areas. The Gini impurity watershed itself

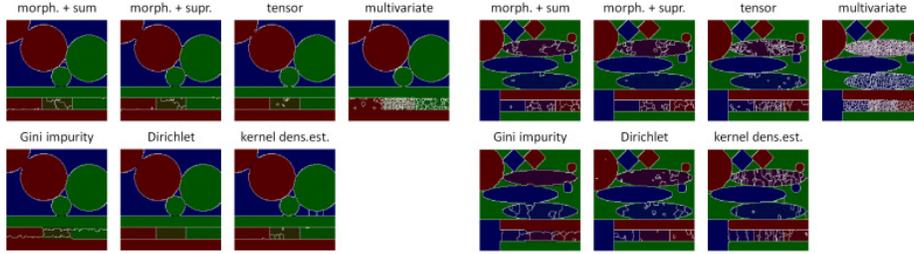


Fig. 5. Experiment 1: Segmentation results on training (left) and test set (right) after training of the classifier with samples from pure mixture regions only.

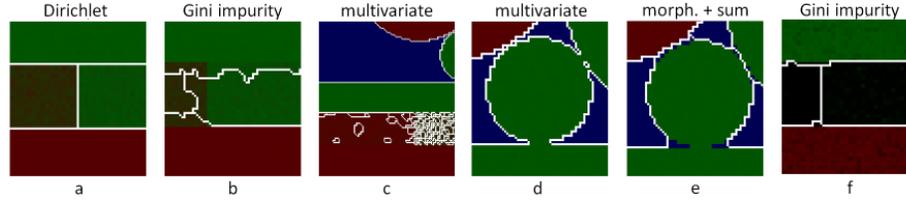


Fig. 6. Experiment 1: Zoomed-in areas of the segmentation results. See text for details.

cannot detect edges between pure and highly impure regions since highly impure mixtures have a high Gini boundary indicator and are therefore interpreted as boundaries instead of regions. However, in the postprocessing step, some of these boundaries are removed by merging basins and the real edges can be recovered.

In the multivariate watershed the two regions in the lower left of fig. 6c are both assigned to the same basin type (class) and thus, the boundary pixels are reduced to jumps in the distance function used to construct the boundary indicator maps. The oversegmentation of the 50% to 50% mixture area results from the fact that depending on the Poisson noise, the pixels in this area are randomly assigned to classes 1 and 2. Here, some smoothing prior to the watershed segmentation could improve results. In contrast to all other methods including the summation of morphological gradients, the multivariate watershed is able to reconstruct the contours in areas of narrow bends (cf. fig. 6d+e).

On the test set, the morphological gradients, the Dirichlet boundary indicator and the Gini impurity perform best. The latter leads to the best distance value, partly because it is least oversegmented. However, some of the boundaries are less precise, but displaced by a few pixels.

Experiment 2. One can argue that the 3 mixture areas (with 25%, 50% and 75% contributions) show clear spatial extent and, thus, constitute classes in their own right. To account for this effect, the classifier was trained with samples from 6 classes. Table 1 shows that in this scenario both the Gini impurity and the multivariate watershed achieve very good results. For example, the edges between pure and impure areas are now much better identified by the Gini impurity watershed (cf. fig. 6f). The multivariate watershed even results in a

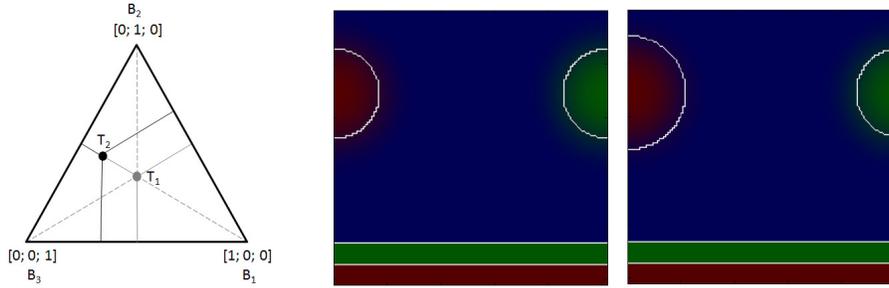


Fig. 7. Experiment 3: The influence of the threshold selection on the segmentation result of the multivariate watershed. Changing the threshold $T_1 = [1/3; 1/3; 1/3]$ to $T_2 = [1/7; 3/7; 3/7]$ in the simplex (left) leads to an accentuation of class 1 (red). The boundaries between classes 2 and 3 remain unchanged.

method	threshold	Experiment 1			Experiment 2		
		distance	sensitivity	specificity	distance	sensitivity	specificity
morph./sum	0.0502	2.4958	0.9396	0.9864	0.2193	0.9408	0.9995
morph./supr.	0.0752	2.4835	0.9423	0.9855	0.3150	0.9253	0.9997
tensor	0.0019	3.1687	0.9710	0.9702	0.2000	0.9648	0.9996
Gini impurity	0.0576	2.3467	0.8734	0.9906	0.3755	0.9199	0.9999
Dirichlet	0.0033	2.5908	0.9311	0.9882	0.2955	0.9210	0.9999
kernel dens.e.	0.0013	3.2484	0.9206	0.9757	0.4054	0.9001	0.9998
multivariate	0.0118	4.7705	0.9965	0.8917	0.0	1.0	1.0

Table 1. The results obtained on the test set show that there is no clear winner. However, the methods introduced here compete well with existing ones.

perfect reconstruction of all boundaries and thus in the best sensitivity and specificity values. The kernel density and Dirichlet boundary indicators again compete well with the existing methods.

Experiment 3. Fig. 7 shows the effect that the threshold of the multivariate watershed has on the segmentation result. The threshold was varied so that the proportions of classes 2 and 3 did not change, but the first class was emphasized with respect to the other classes. It can be seen from the two segmentation results that the border between classes 2 and 3 remains unchanged whereas the border between classes 1 and 3 is shifted in favor of class 1. By setting the threshold the user can control the class weights. Thus, the multivariate watershed provides the user with more control over the segmentation result than other methods like the watershed based on morphological gradients.

Real data. Fig. 8 shows the results of the different methods after post-processing with watershed dynamics. The optimal parameters have been tuned manually. Judging from the approximate manual segmentation, all approaches lead to reasonable results. However, some oversegmentation remains that cannot be removed with the concept of dynamics.

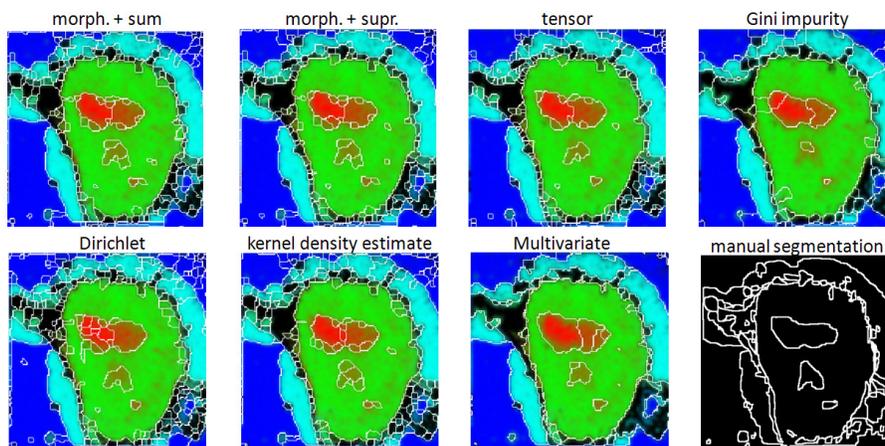


Fig. 8. Segmentation results on the real data (manually selected parameters).

6 Conclusion

We have introduced four watershed-based methods for segmenting multivariate compositional data: the Gini impurity watershed, the Dirichlet boundary indicator watershed, the kernel density estimate based watershed and the multivariate watershed. The former three approaches use novel techniques to obtain a scalar boundary indicator map which is used as input for the classic watershed transform. The latter one generalizes the classic definition of the watershed to multispectral compositional data. In our experiments on simulated data, no overall best performing method could be identified. However, the methods introduced in this paper have been shown to compete well with existing methods and are superior in some scenarios.

References

1. H. Digabel, C. Lantuéjoul. Iterative algorithms. *Proc. 2nd Europ. Symp. Quant. Anal. of Microstr. in Material Science, Biology and Medicine*, pages 85–99, 1978.
2. J.B.T.M. Roerdink and A. Meijster. The watershed transform: Definitions, algorithms and parallelization strategies. *Fundamenta Informaticae*, 41:187–228, 2001.
3. O. Lezoray. Supervised automatic histogram clustering and watershed segmentation. *Image Anal. Stereol.*, 22:113–120, 2003.
4. C.Y. Chang, W.S. Shie, J.H. Wang. Color image segmentation via fuzzy feature tuning and feature adjustment. *Conf. on Syst., Man & Cybernetics*, 4:6, 2002.
5. A.B. Hugué, M.C. de Andrade, R.L. Carceroni, and A.A. Araujo. Color-based watershed segmentation of low-altitude aerial images. *17. Brazilian Symp. on Comp. Graphics and Image Proc.*, 17-20:138–145, 2004.
6. G. Noyel, J. Angulo, D. Jeulin, D. Balvay, and C.-A. Cuenod. Filtering, segmentation and region classification by hyperspectral mathematical morphology of DCE-MRI series for angiogenesis imaging. *Intern. Symp. on Biomedical Imaging: From Nano to Macro*, pages 1517–1520, 2008.

7. L. A. McDonnell and R. M. A. Heeren. Imaging mass spectrometry. *Mass Spectrometry Reviews*, 26:606–643, 2006.
8. J. Moffat et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell*, 124(6):1283–1298, 2006.
9. P. Scheunders. Multivalued image segmentation based on first fundamental form. *Proc. of the Intern. Conf. on Image Anal. and Processing*, pages 185–190, 2001.
10. P. Li and X. Xiao. Evaluation of multiscale morphological segmentation of multispectral imagery for land cover classification. *Proc. of the IEEE Geoscience and Remote Sensing Symposium*, 4:2676–2679, 2004.
11. G. Noyel, J. Angulo, and D. Jeulin. Morphological segmentation of hyperspectral images. *Image Anal. Stereol.*, 26:101–109, 2007.
12. N. Malpica, J.E. Ortuno, and A. Santos. A multichannel watershed-based algorithm for supervised texture segmentation. *Patt. Rec. Letters*, 24:1545–1554, 2003.
13. P.S. Karvelis, D.I. Fotiadis, A. Tzallas, and I. Georgiou. Region based segmentation and classification of multispectral chromosome images. *IEEE Trans. on Medical Imaging*, 27:697–708, 2008.
14. Y. Zhang, X. Feng, X. Le. Segmentation on multispectral remote sensing image using watershed transformation. *Congr. on Image & Sign. Proc.*, 4:773–777, 2008.
15. J. Angulo and D. Jeulin. Stochastic watershed segmentation. *Intern. Symp. on Mathematical Morphology*, 8:265–276, 2007.
16. G. Noyel, J. Angulo, and D. Jeulin. Random germs and stochastic watershed for unsupervised multispectral image segmentation. *LNCS*, 4694/2008:17–24, 2008.
17. P. Soille. *Morphological Image Analysis*. Springer, 1999.
18. J. Aitchison. *The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability*. Chapman and Hall, 1986.
19. L. Vincent and P. Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. on PAMI*, 13:583–598, 1991.
20. F. Meyer. Topographic distance and watershed lines. *Sig. Proc.*, 38:113–125, 1994.
21. L. Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.
22. P.J. Ulintz, J. Zhu, Z.S. Qin, and P.C. Andrews. Improved classification of mass spectrometry database search results using newer machine learning approaches. *Molecular and Cellular Proteomics*, 5:497–509, 2006.
23. M. Hanselmann et al. Concise representation of mass spectrometry images by probabilistic latent semantic analysis. *Anal. Chem.*, 80(24):9649–9658, 2008.
24. G. Casella and R.L. Berger. *Statistical Inference*. Duxbury Advanced Series, 2002.
25. F. Godtliebsen, J.S. Marron, and P. Chaudhuri. Significance in scale space for bivariate density estimation. *Journal of Comp. and Graph. Stat.*, 11:1–21, 2002.
26. D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *Trans. on Patt. Anal. and Mach. Intelligence*, 24(5):603–619, 2002.
27. C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. *Proc. of the Conf. on Computer Vision*, pages 836–846, 1998.
28. M. Grimaud. New measure of contrast: the dynamics. *Image Algebra and Morphological Image Processing III*, pages 292–305, 1992.
29. L. Brun, M. Mokhtari, and F. Meyer. Hierarchical watersheds within the combinatorial pyramid framework. *Proc. of DGCI 2005*, 2005.
30. A.J. Baddeley. An error metric for binary images. *Robust Computer Vision*, pages 59–78, 1992.
31. A. Rosenfeld and J. Pfaltz. Sequential operations in digital picture processing. *Robust Computer Vision*, 13(4):471–494, 1966.
32. D. Gale and L.S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69:9–14, 1962.