Unsupervised Classification of X-Ray Mapping Images of Polished Sections

Klaus Baggesen Hilger, Allan Aasbjerg Nielsen, Jens Michael Carstensen
IMM, Department of Mathematical Modelling
Technical University of Denmark, Building 321
DK-2800 Kgs. Lyngby, Denmark
http://www.imm.dtu.dk/~kbh, kbh@imm.dtu.dk

Abstract

X-ray mapping images of polished sections are classified using two unsupervised clustering algorithms. The methods applied are the k-means algorithm and an extended spectral fuzzy c-means algorithm. The extentions include new types of memberships that are related to the contextual information. In addition to the traditional spectral membership we apply a spatial membership and a parental membership. The parental membership is introduced by implementing the algorithm in a scale-space representation. Both spectral and spatial information is carried across levels, enhancing speed and visual impression of the segmentation of image data.

1 Introduction

This paper deals with unsupervised classification of multichannel scanning electron microscope (SEM) energy dispersive spectroscopy (EDS) image data from polished sections, also known as x-ray mapping imagery. As a case study we segment a multivariate image containing 176 rows and 256 columns. The image consists of 10 channels that represent the elements Al, C, Fe, Mg, Na, O, P, S, Si, and Ca, see Figure 1. As the data are counts and thus ideally follow a Poisson distribution as a variance-stabilising measure all numbers are squarerooted before the analysis.

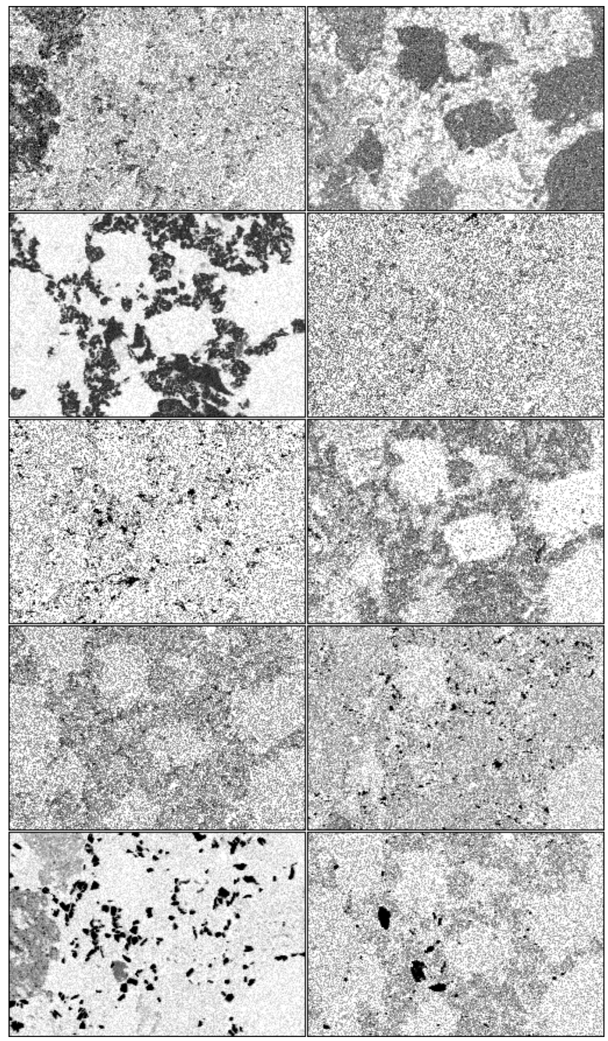

Figure 1: The channels 1-10 (row wise) representing Al, C, Fe, Mg, Na, O, P, S, Si, and Ca. Dark regions represent high counts. Each image is stretched linearly between its mean ±3 standard deviations.

2 The Extended Fuzzy Algorithm

Segmentation of multivariate data into a desired number of clusters is often done by means of the k-means algorithm. If we wish to consider a (fuzzy) cluster membership degree, the c-means algorithm applies, [2]. In [7] a spatial element is added. [3, 8] add a multi-resolution aspect.

Given N P-dimensional observations that we wish to classify into C classes using the fuzzy c-means (FCM) algorithm. Let $\boldsymbol{U}$ be an $N \times C$ matrix with the elemenets u_{ic} that describe the membership for observation i to class c, and let $\boldsymbol{R}$ be a $P \times C$ matrix that contains the cluster centres, i.e., the centroids of the classes. The spectral FCM algorithm minimises the within class sum of squared errors functional $J(\boldsymbol{U}, \boldsymbol{R})$ under conditions:

$$u_{ic} \in [0,1] \quad \forall\, i = 1,\dots,N; c = 1,\dots,C \tag{1}$$

$$\sum_{c=1}^{C} u_{ic} = 1 \qquad \forall\, i = 1,\dots,N \tag{2}$$

$$\sum_{i=1}^{N} u_{ic} > 0 \qquad \forall\, c = 1,\dots,C \tag{3}$$

Condition 1 implies that the memberships are allowed to be partial and can take any value between 0 and 1. Condition 2 implies that the sum across all classes of the memberships of an individual is unity. Condition 3 indicates that there must be at least one individual with some degree of membership to a given class. Hence, empty classes are not allowed for.
Any $\boldsymbol{U}$ that satisfies conditions 1, 2 and 3 is referred to as a fuzzy partitioning of N individuals into C classes.

Under the above conditions we shall minimise a functional proposed by [2]:

$$J(\boldsymbol{U}, \boldsymbol{R}) = \sum_{i=1}^{N} \sum_{c=1}^{C} u_{ic}^{m} d_{ic}^{2}. \tag{4}$$

Here $m \geq 1$ is a fixed parameter that determines the degree of fuzziness of the final solution, that is the degree of overlap between groups. With $m = 1$ the solution is a hard partition, equivalent to that obtained by the k-means algorithm. The degree of fuzziness increases with m. The higher the value of m, the lower the membership degrees of observations which are far from all cluster centres. As m approaches infinity the solution approaches its highest degree of fuzziness, with $u_{ic} = 1/C$ for every pair of i and c. The squared distance between the ith observation and the cth cluster center is denoted d_{ic}^2.

2.1 Spectral FCM

The spectral fuzzy c-means algorithm

1. assigns values to P-dimensional feature vectors for C cluster centres, $\tilde{\boldsymbol{r}}_c,\ c = 1,\dots,C$;
2. assigns to each observation $i = 1,\dots,N$ calculates membership weights for clusters $c = 1,\dots,C$

$$u_{spec,ic} = \frac{1/d_{ic}^{2/(m-1)}}{\sum_{j=1}^{C} 1/d_{ij}^{2/(m-1)}} \tag{5}$$

   where d_{ic} is the (Euclidean) spectral distance from the running observation to each cluster centre $d_{ic}^2 = (\boldsymbol{r}_i - \tilde{\boldsymbol{r}}_c)^T(\boldsymbol{r}_i - \tilde{\boldsymbol{r}}_c)$, and $m > 1$. Alternertively one can apply the Mahalanobis distance as in [4];
3. calculates new cluster centres from

$$\tilde{\boldsymbol{r}}_c = \frac{\sum_{i=1}^{N} u_{ic}^{m} \cdot \boldsymbol{r}_i}{\sum_{i=1}^{N} u_{ic}^{m}}. \tag{6}$$

Steps 2 and 3 are iterated until the largest change in cluster membership becomes small or zero.

2.2 Including Spatial Information

In this section two new memberships, the spatial and the parental, are introduced that include spatial context information. Before calculating the new cluster centres we merge the spectral, the spatial, and the parental memberships into joint memberships which are applied as weights for the spectral observations.

We define spatial membership by

$$u_{spat,ic} = \frac{1}{Z}\exp(-\beta E_{\mathcal{N}}) \tag{7}$$

where

$$E_{\mathcal{N}} = \frac{1}{|\mathcal{N}|} \sum_{j \in \mathcal{N}} (1 - u_{jc}) \tag{8}$$

corresponds to a Markov random field (MRF) energy function, $\beta \geq 0$ is a weighting parameter, and Z is a normalising constant. The sum over $\mathcal{N}$ indicates a sum over the neighbourhood of an observation and $|\mathcal{N}|$ is the number of neighbours in $\mathcal{N}$. With $\beta = 0$ no spatial context information is included. The spatial membership to a class is large if the observations in the neighbourhood have large memberships to the same class and small if the neighbours tend to belong to other classes. The clique potential is parameterized by a user defined parameter, β. In each iteration we estimate the spatial membship degrees from the spectral memberships.

Using a combination of blurring and subsampling a multiresolution scale-space pyramid is constructed. We apply a Gaussian blurring kernel and perform subsampling such that a parent pixel at level j has four children at level $j-1$. We use reflection in order to handle border effects. Level 0 corresponds to the level of highest resolution, and the number of levels is set by the user.

Applying the FCM algorithm to the scale-pyramid we start at the top level performing a segmentation. The resulting cluster centres are passed down to the next level in the pyramid as initial cluster centers. Passing the memberships found at a higher level down through the pyramid introduces additional spatial awareness into the algorithm. We introduce the additional membership as an external field that corresponds to an a priori knowledge of the memberships of the given observation to the different classes. The membership of a parent of an observation to a class c is denoted $u_{parent,c}$. The joint spectral-spatial-parental membership can now be calculated as

$$u_{ic} = \frac{u_{spec,ic} \cdot u_{spat,ic} \cdot u_{parent,ic}}{\sum_{j=1}^{C} u_{spec,ij} \cdot u_{spat,ij} \cdot u_{parent,ij}}. \tag{9}$$

3 Results

The k-means result is generated by the SAS fastclust procedure, [5].

Applying the c-means algorithm to the data with spectral information only reveals that although the hard classification looks sensible, the membership degrees are all very low (not shown). Hence, the confusion concerning the segmentation is relatively high. Adding spatial and parental context leads to more distinct membership degrees, i.e., membership degrees closer to 0 and 1 indicating a better classification.

The algorithm is applied using $m = 2$, $\beta = 1$ in a scale pyramid consisting of four levels. The parental membership is introduced from level two and down, see Figures 2 and 3.

Comparing the hard result for the k-means and the extended c-means algorithms we notice that the noise-like structures have been avoided for the latter. The memberships have been thresholded at 0.5 introducing a reject class (the white regions in the right image of Figure 3). The reject class dominates boundary regions between classes and a few smaller regions. Comparing these regions with Figure 1 we find that the reject class is dominated by a single element, Ca.

The classes have been compared in order to verify that they are true classes. This is done for both the k-means and the c-means results by applying a Wishart distribution based test to check whether any two classes simultaneously have equal means and dispersions, [1].

In Figures 4 and 5 descriptors are calculated for the different classes. The images contain the class centres determined by the k-means and the extended c-means algorithms. The figures are to be compared to Figures 1 and 3. The spectra in Figure 5 are calculated using the memberships from Figure 2 without thresholding.

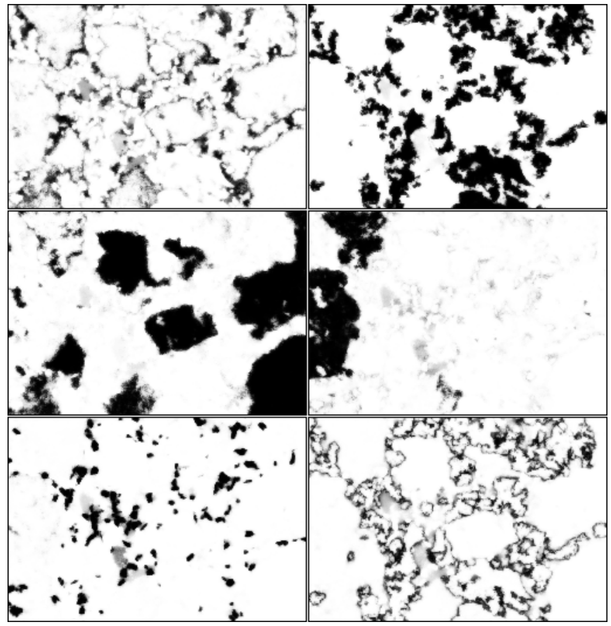

Figure 2: Segmentation based on spectral, spatial, and parental memberships. The six cluster membership-degrees (row wise). Dark regions represent high memberships.

4 Conclusions

Two types of unsupervised classification have been applied to SEM/EDS or x-ray mapping images, namely k-means and (fuzzy) c-means (FCM) classification. The FCM algorithm has been extended by incorporating a multi-scale representation of the image data, partially for speed up and partially for carrying spatial information across scale levels. Also, a spatial element at each scale level has been included. Simultaneous inspection of plots of class means (Figure 5) and Figures 1 and 2 provide support for the applications expert in interpreting the contents of the classes found. Comparisons between results from k-means and c-means analyses

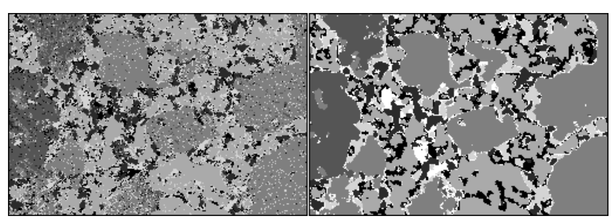

Figure 3: Hard result obtained from the membership-degrees. White pixels correspond to the reject class. The image to the left is obtained based on k-means with spectral features. The image to the right is obtained based on c-means using additional spatial and parental features.

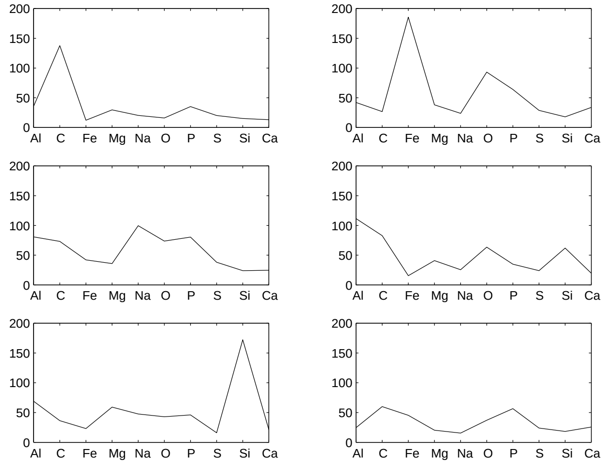

Figure 4: Class descriptors obtained applying the k-means algorithm. The class signatures $\tilde{\boldsymbol{r}}_c, c = 1, \ldots, 6$ row wise.

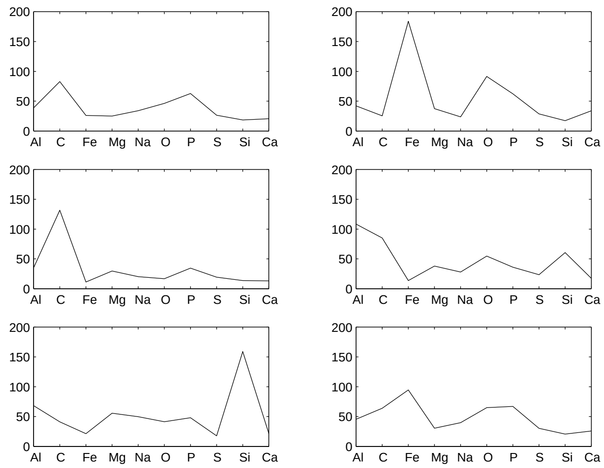

Figure 5: Class descriptors obtained applying the c-means algorithm. The class signatures $\tilde{\boldsymbol{r}}_c, c = 1, \ldots, 6$ row wise.

show that although the class means exhibit some similarities the visual impression of the c-means result when using the multi-scale approach with spatial information being carried across scale levels is much pleasing indicating large same-class regions. Also, the degrees of membership are much closer to 0 and 1 when compared to a purely spectral c-means classification indicating a better result. In spite of the problem with classifying the regions rich in Ca this type of analysis seems a good exploratory tool to obtain insight into the discriminatory power of the data. It thus constitutes a good preprocessor for a more thorough supervised analysis (in which one could explicitly introduce the Ca-rich regions as a class).

Acknowledgement

The data used come from the COMB project, the Industrial Centre for Surface Microscopy, Microanalysis and Image Analysis headed by Dr. Leif Højslet Christensen, [6]. The COMB project is funded by Erhvervsfremme Styrelsen, the Danish Agency for Trade and Industry.

References

[1] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis.* John Wiley, New York, second edition, 1984. 675 pp.

[2] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms.* Plenum Press, 1981.

[3] C. A. Bouman and M. Shapiro. A multiscale random field model for Bayesian image segmentation. *IEEE Transactions on Image Processing*, 3(2):162–177, 1994.

[4] Y. Chtioui, D. Bertrand, D. Barba, and Y. Dattee. Application of fuzzy c-means clustering for seed discrimination by artificial vision. *Chemometrics and Intelligent Laboratory Systems*, 38:75–87, 1997.

[5] SAS Institute Inc. *SAS/STAT User's Guide, ver. 6.* SAS Institute, 1990.

[6] Internet http://www.teknologisk.dk/251. COMB, the Industrial Centre for Surfacemicroscopy, Microanalysis and Image Analysis.

[7] R. Wiemker. Unsupervised fuzzy classification of multispectral imagery using spatial-spectral features. In I. Balderjahn, R. Mathar, and M. Schader, editors, *Data Highways and Information Flooding, A Challenge for Classification and Data Analysis.* Springer, 1997.

[8] R. Wilson and C.-T. Li. Hidden multiresolution random fields and their application to image segmentation. In *Proceedings of the 10th International Conference on Image Analysis and Processing (ICIAP'99)*, pages 346–351, Venice, Italy, 27-29 September 1999.