

Sensitivity Study of a Semi-automatic Supervised Classifier Applied to Minerals from X-Ray Mapping Images

Rasmus Larsen^a and Allan Aasbjerg Nielsen^a and Harald Flesche^b

^a Department of Mathematical Modelling
Technical University of Denmark, Building 321
DK-2800 Kgs. Lyngby, Denmark
{rl,aa}@imm.dtu.dk

^b Norsk Hydro Research Centre
N-5020 Bergen, Norway
Harald.Flesche@hydro.com

This paper addresses the problem of assessing the robustness with respect to change in parameters of an integrated training and classification routine for minerals commonly encountered in siliciclastic or carbonate rocks. Twelve chemical elements are mapped from thin sections by energy dispersive spectroscopy (EDS) in a scanning electron microscope (SEM). Extensions to traditional multivariate statistical methods are applied to perform the classification. Training sets are grown from one or a few seed points by a method that ensures spatial and spectral closeness of observations. Spectral closeness is obtained by excluding observations that have high Mahalanobis distances to the training class mean. Spatial closeness is obtained by requiring connectivity. The marginal effects of changes in the parameters that are input to the seed growing algorithm are evaluated. Initially, the seed is expanded to a small area in order to allow for the estimation of a variance-covariance matrix. This expansion is controlled by upper limits for the spatial and Euclidean spectral distances from the seed point. Second, after this initial expansion the growing of the training set is controlled by an upper limit for the Mahalanobis distance to the current estimate of the class centre. Also, the estimates of class centres and covariance matrices may be continuously updated or the initial estimates may be used. Finally, the effect of the operator's choice of seed among a number of potential seeding points is evaluated. After training, a standard quadratic classifier is applied. The performance for each parameter setting is measured by the overall misclassification rate on an independently generated validation set. The classification method is presently used as a routine petrographical analysis method at Norsk Hydro Research Centre.

1. Introduction

Mineral classification and quantification is traditionally done using point counting of thin sections or by x-ray diffraction (XRD). The first of these methods is very time consuming and requires a trained petrographer; the latter does not give any spatial information about the samples being analysed. Point counting also has an element of subjectivity in that a more skilled petrographer will be better at recognising rare minerals, separating cement from detrital grains, etc.

A third method is to do x-ray mapping or energy dispersive spectroscopy (EDS) in a scanning electron microscope (SEM). Here, an x-ray spec-

trum is acquired for each pixel. By means of this spectrum the mineral present in each pixel can be identified using a manual or an automatic classification method. Spatial information about the mineral composition can thereby be obtained by an objective and reproducible method. Earlier work in this field [9,13,2] used classification methods that range from lookup tables to maximum likelihood classification. Earlier, long image acquisition times made the use of EDS images for mineral classification difficult. New equipment enables acquisition of a 256×256 pixels image with 12 elements mapped in 36 minutes. Accelerating voltage, current, dwell time and sensor parameters are adjusted for an acceptable trade-

off between data noise level and image acquisition time. The chosen configuration results in 40% deadtime in the EDS detector. Typically the pixel size is $2.4\mu\text{m}\times 2.4\mu\text{m}$.

A method for classification of EDS images based on a semi-automatic training algorithm is described in [10,4]. This paper aims at evaluating the robustness of this semi-automatic training algorithm to parameter settings and operator influence.

2. Data: minerals and elements

The purpose of the analysis is to evaluate mineralogical composition of sedimentary rocks from thin sections of core samples from different wells and fields.

As the aim is to use the method for standard studies of sedimentary rocks, it is important to cover the most frequently occurring minerals. In Table 1 all the minerals in the model are shown. In some cases more than one class is needed to describe a mineral. This can be due to natural variation in the chemical composition of the mineral, such as in the biotites, the chlorites, the garnets and the siderites. There is also a case, for illite and muscovite, where it is known in advance that different minerals have the same chemical composition, and they will therefore be overlapping in the EDS measurement. In addition to minerals it is of course also important to have a class for porosity. It is possible to map porosity as the samples are impregnated with epoxy. The

Table 1

Mineral classes in the model

Albite	Chlorite 2	Gypsum	Quartz
Ankerite	Chlorite 3	Illite/Musc.	Rutile
Apatite	Dolomite	Ilmenite	Siderite 1
Barite	Fe-calcite	Kaolin	Siderite 2
Biotite 1	Garnet 1	K-feldspar	Titanite
Biotite 2	Garnet 2	Monazite	Tourmaline
Calcite	Garnet 3	Porosity	Zinblend
Chlorite 1	Glauconite	Pyrite	Zircon

input data are not continuous x-ray spectra but mappings of 12 regions of interest in the spectra covering wavelengths for selected elements. The mapped elements reflect the major components

Table 2

Mapped elements

Al	C	Ca
Fe	K	Mg
Mn	Na(+Zn)	P(+Zr)
S	Si	Ti(+Ba)

in the minerals. It is normally the K_α line that is mapped, but in some instances this is superimposed by another element's L_α line. This is the case for P and Zr, Ti and Ba, and Na and Zn. There have not been any problems with this duality in the data, rather it potentially increases the discriminatory power between more minerals. The set of mapped elements are shown in Table 2.

3. Methods

3.1. The seed algorithm

Good supervised classification is contingent on good training sets. In order to obtain a properly performing classification algorithm the training classes should represent statistically well separated classes. This is not always the case for training sets drawn by human operators. This is partly due to the human inability to get an overview of multidimensional spaces. Another problem with training sets drawn by humans is inconsistency. Training sets need to be extracted in a consistent way across time and classes irrespective of operator and shape of image structures.

Therefore we propose a semi-automatic algorithm for generation of a set of training classes from a series of seeding points, i.e. for each class the operator need only supply one (or more) points in the image that belongs to that particular class. This algorithm was previously described in [10,4].

From these points training classes are grown. This ensures at least these two important points: spatial and spectral closeness.

Spatial closeness is ensured by demanding that all the pixels in one training class are connected with the seeding point. This is a very useful condition, because it is the nature of most relevant phenomena to appear as objects. The connectivity may be defined in terms of first- or second-

order neighbours etc. This definition has the advantage of being able to allow for small gaps in the training sets by specifying that pixels should be higher order neighbours. This is useful for classes that occur as clusters of smaller objects in the image, and also in the case of classes that occur as thin strings or layers.

Spectral closeness is achieved by making restrictions on the distance to the current mean value of the class while growing the training set. Here, two types of distance are considered. The first distance is the Euclidean spectral distance

$$D_E^2 = (\mathbf{x} - \boldsymbol{\mu}_i^*)^T (\mathbf{x} - \boldsymbol{\mu}_i^*) \quad (1)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_p)^T$ is the value observed in a pixel, and $\boldsymbol{\mu}_i^*$ is the current estimate of the class i mean. The application of Euclidean distance to seed growing is suggested in [3]. The second distance is the Mahalanobis distance

$$D_M^2 = (\mathbf{x} - \boldsymbol{\mu}_i^*)^T (\boldsymbol{\Sigma}_i^*)^{-1} (\mathbf{x} - \boldsymbol{\mu}_i^*). \quad (2)$$

Here the distance is scaled by the inverse of the current estimate of the covariance matrix of training class i , $\boldsymbol{\Sigma}_i^*$.

For the Euclidean distance an upper limit for the distance should be supplied by the user. For the Mahalanobis distance we can utilise that these distances are approximately χ^2 -distributed with p degrees of freedom, i.e. we include only pixels that have a Mahalanobis distance to the current estimates of the class means less than a preset quantile of the corresponding χ^2 -distribution.

Since we often do not have any prior knowledge of the scales of variation of the phenomena at hand the Mahalanobis distance is often preferred. However, here a problem is encountered. If the seeding is started by a single pixel we cannot get an immediate estimate of the covariance matrix. Therefore our algorithm starts with growing the training set to a small preset maximum spatial radius using the Euclidean spectral distance method with a preset maximum spectral distance. From the small number of pixels thus included estimates of the mean vector and the covariance matrix are obtained. These estimates are first used to exclude possible outliers in the current set, and second, the training set is grown

using the Mahalanobis distance method from this point on.

Finally, there is a choice to update the estimates of the mean value and the covariance matrix continuously as pixels are included in the training set, as opposed to basing the Mahalanobis growing on the initial estimates obtained from the Euclidean expansion.

Use of this method gives us training classes that are in good correspondence with normal distributions with the estimated means and covariances.

3.2. Classification

We will apply a standard supervised classifier. We presume that noise may be described by a multivariate Gaussian distribution possibly after some transformation. The Gaussian distribution is by far the most commonly adopted density model for continuous image features. This is partly supported by the central limit theorem, and partly due to the resulting nice analytical expressions in various kinds of analyses.

Suppose that a pixel is an observation from one of the populations $\pi_1, \pi_2, \dots, \pi_k$. The classification of the observation depends on the vector of features of that pixel, which we will denote $\mathbf{X} = (X_1, X_2, \dots, X_p)^T$. Assuming a Gaussian feature model, we find the class conditional density function of class π_i to be

$$f_i(\mathbf{x}) = P(\mathbf{X} = \mathbf{x} \mid C = \pi_i) = \frac{1}{\sqrt{2\pi^p}} \frac{1}{\sqrt{\det \boldsymbol{\Sigma}_i}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i)\right) \quad (3)$$

where C is the class variable, and $i = 1, 2, \dots, k$.

Furthermore, let us assume knowledge of the prior distribution of the classes, i.e. the prior probabilities, $P(C = \pi_i) = p_i$. This distribution determines the probability with which an arbitrary feature vector has been generated from a particular population.

The combined knowledge of the class conditional density and the prior distribution allows us to write the unconditional feature vector density function; this is a compound distribution

$$h(\mathbf{x}) = \sum_{i=1}^k p_i f_i(\mathbf{x}). \quad (4)$$

The posterior distribution for the class variable is

$$k(\pi_i | \mathbf{x}) = \frac{p_i f_i(\mathbf{x})}{h(\mathbf{x})}. \quad (5)$$

The Bayes solution of a classification problem is to choose the action that minimises the posterior expected loss. In the case of equal losses, i.e. the same loss is assigned to all misclassifications, the Bayes solution consists of choosing the class that has the highest posterior probability.

In the case of class dependent covariances the posterior probability, after taking logarithms and excluding terms that are common to all classes is given by

$$S_i = \log p_i - \frac{1}{2} \log(\det \Sigma_i) - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i), \quad (6)$$

i.e., a quadratic function of the observations. Therefore this is denoted quadratic classification. If this function is maximum for $i = j$ we assign the pixel to population π_j .

For a general statistical reference, see [1].

3.3. Evaluation of classification

The evaluation of the classification resulting from a particular training set is made using a validation set. This validation set is the same as used in [10,4]. To assure independence samples for the validation set are preferably chosen from other wells (or fields) than samples used for the training set. For some rarer minerals it has not been possible to find independent samples and no validation therefore exists.

4. Results

In [10,4] classifications of a SEM EDS dataset consisting of the elements in Table 2 were carried out using four different classifiers with the purpose of discriminating between the mineralogical classes listed in Table 1. The classifiers used were a quadratic classifier as described above, a contextual quadratic classifier, [6,11,7,5], a hierarchical quadratic classifier, [12,8], and an extension of the hierarchical classifier, [4].

The training and validation sets were grown using the seed algorithm described above from seed points identified by an experienced geologist. In [10,4] it is shown that some of the mineralogical

classes could not be discriminated between using the set of features at hand. This goes for the three garnets (garnet 1, garnet 2, garnet 3, cf. Table 1) and for the two calcites (Fe-calcite and calcite, cf. Table 1). These classes were therefore combined into two superclasses after classification. The validation set included samples from all minerals except biotite 2, chlorite 3, siderite 2, and zinblend. Furthermore biotite 1 was removed from the validation set as it was concluded by geologists that the sample chosen for validation was altered so that it was closer to the chlorites in its chemical composition. Taking the combination into superclasses and the exclusion of the validation sample for biotite 1 into account the misclassification rates for the training set (resulting in the so-called resubstitution rate) and the validation set as well as the computing time relative to the classical quadratic classifier are shown in Table 3.

Table 3
Misclassification rates and computing times for the four classifiers tested in [10,4]

	Training Set	Validation Set	Processing Time
Quadratic	0.25%	1.06%	1
Contextual			
Quadratic	0.33%	0.65%	5.48
Hierarchical	0.25%	1.09%	0.47
Extended			
Hierarchical	0.25%	1.13%	3.49

The performances of the four classifiers are more or less the same. The classical quadratic classifier was chosen for its low processing time. Although the hierarchical classifier had a lower processing time it was not used because it was demonstrated that the classification results were dependent of the order in which the populations were presented to the classifier.

The parameters of the seed growing are shown in Table 4.

4.1. Euclidean seed growing

The parameters for the initial step of the seed algorithm are varied one at a time, and the results are summarised below. For selected changes

Table 4

Seed growing parameters for the classification performed in [10,4], nominal values

Parameter	Value
Euclidean distance threshold	30*
Range for Euclidean growing	5
Mahalanobis distance threshold	$\chi_{0.99}^2(p)$
Update $(\boldsymbol{\mu}_i^*, \boldsymbol{\Sigma}_i^*)$	no update

*For classes barite, K-feldspar, zinblende, and zircon the Euclidean distance is set to 50 in order to include a sufficient number of pixels.

in the Euclidean spectral distance threshold the misclassification rates are shown in Table 5, and for the spatial range parameter the results are shown in Table 6.

Table 5

Misclassification rate as a function of selected changes in the Euclidean spectral distance threshold

Euclidean Distance	Training Set	Validation Set
0	0.25%	1.06%
+5	0.33%	0.85%
+10	0.52%	1.03%
+50	1.75%	1.06%
+100	4.03%	1.11%

When increasing the Euclidean spectral distance threshold the misclassification rate for the validation set is fairly constant whereas the rates for the training set increases from 0.25% at the nominal setting to 4.03% at nominal setting +100 as shown in Table 5. The latter increase is mainly explained by misclassification rates of 57.5% for monazite (classified mainly as porosity), 19.2% for chlorite 2 (classified mainly as porosity and chlorite 1) and 13.1% for chlorite 1 (classified mainly as chlorite 2).

The variation of the spatial range for the Euclidean seed growing results in almost no change in misclassification rates (Table 6). For the validation set the rates are constant, and for the training set there is a slight increase with the quantile from the nominal range 5 and up. For range 4 the rate is slightly higher. This may well be caused by estimation variation.

Table 6

Misclassification rate as a function of selected changes in the spatial range for the Euclidean seed growing

Range	Training Set	Validation Set
-1	1.13%	1.04%
0	0.25%	1.06%
+1	0.51%	1.11%
+2	0.60%	1.07%
+3	0.66%	1.07%

4.2. Mahalanobis seed growing

Table 7 shows the misclassification rate as a function of selected values of the χ^2 -quantile. We see that the misclassification rate for the validation only varies a little, whereas the rates for the training set, though small, steadily increase with the quantile. This is simply because larger quantiles increase the overlap between classes in feature space, and thus result in higher misclassification.

Table 7

Misclassification rate as a function of selected values of the χ^2 -quantile

Quantile	Training Set	Validation Set
0.950	0.05%	0.94%
0.975	0.11%	0.83%
0.990	0.25%	1.06%
0.995	0.32%	0.83%
0.999	0.58%	0.91%

With respect to continuously updating the parameters of the distributions when performing the Mahalanobis seed growing we see from Table 8 that a change is induced in the misclassification rates. The reason for this is that some classes, mainly those that occur in the images as clay or cement (e.g. thin layers of a class coating another class that occurs as grains) or are otherwise spatially dispersed grow beyond their spatial boundaries.

4.3. Seed point sensitivity

Finally, we examine how sensitive the classification is to the particular choice of seeding point. For each class a new seeding point is picked at random from the originally generated training set

Table 8

Misclassification rate as a function of using update vs. no update of the distribution parameters

Update option	Training Set	Validation Set
No update	0.25%	1.06%
Update	5.48%	2.00%

within a 5×5 neighbourhood of the original seeding point. This procedure is repeated 4 times. The results in terms of misclassification rates are shown in Table 9.

We see that the misclassification rates for the training set remain very similar when varying the

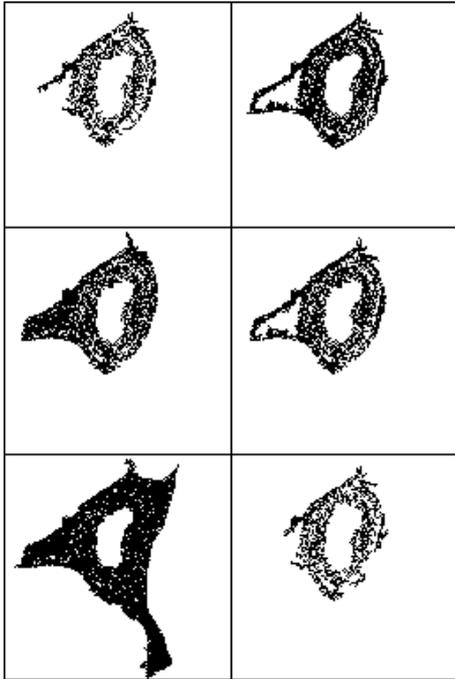


Figure 1. Example of training sets for the class chlorite 2; upper left: original training set; upper right: using range “+3” for Euclidean seed growing; middle left: using Euclidean distance threshold of “+100”; middle right: using quantile “0.999” for Mahalanobis seed growing; lower left: using update option for Mahalanobis seed growing; lower right: using new a seed point

Table 9

Misclassification rates as a function of using different seeding points.

Seed point set	Training Set	Validation Set
original	0.25%	1.06%
set 1	0.35%	0.99%
set 2	0.28%	1.04%
set 3	0.26%	0.83%
set 4	0.27%	2.18%

seeding point. With the exception of the last new training set this is also the case for the misclassification rates for the validation set. For the fourth new training set the increase in misclassification rate for the validation set is almost exclusively explained by a misclassification rate of 29.9% for kaolin (as chlorite 2 and illite/muscovite). Kaolin is the mineral that has the highest inhomogeneity of the 5×5 neighbourhood from where the new seed point was chosen (9 out of 24 of the neighbours of the original seed point were rejected as kaolin pixels in the original seed growing).

5. Discussion

In Figure 1 examples of the seed growing of chlorite 2 are shown. Corresponding elements are shown in Figures 2 and 3. We see from Figures 1, 2 and 3 that the region included into the training set for the extreme value of range, “+3”, as well as that resulting from using the extreme value of quantile for the Mahalanobis growing, “0.999”, does not seem to include pixels that are spectrally different from those included in the original training set. This is in good correspondence with the fairly small changes in misclassification rates shown in Tables 6 and 7. However, when using the updating option for the Mahalanobis seed growing, it is evident that regions that are spectrally different (in this case porosity) are included in the new training set. When comparing with the misclassification rates in Table 8 we see that these are significantly higher than for the original training set. Also, when applying the extreme value of the Euclidean distance threshold, “+100”, we see that a smaller region of porosity is included in the training set here. This is also reflected in the misclassification rates of the train-

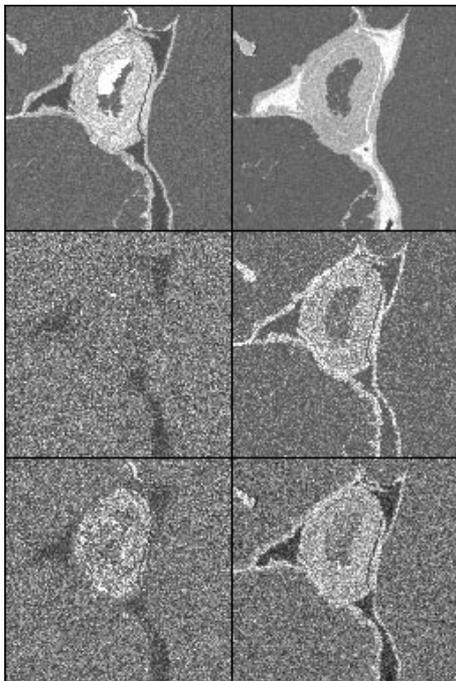


Figure 2. Elements corresponding to the training set example shown in Figure 1; row-wise: Al, C, Ca, Fe, K and Mg

ing set. But it does not seem to have an impact on the misclassification rate of the validation set. Finally, the insensitivity of the method to changes in the seeding point is illustrated by the training set resulting from a new seed. No spectrally different pixels seem to be included here. This is confirmed by the low misclassification rates seen in Table 9.

6. Conclusions

We have conducted a sensitivity study with respect to the parameters of the generation of training set of a semi-automatic classifier for classification of minerals from x-ray mapping images. The marginal change in misclassification rates as a function of selected changes in five parameters is evaluated.

The initial seed growing based on the Euclidean distance measure seems independent of the maxi-

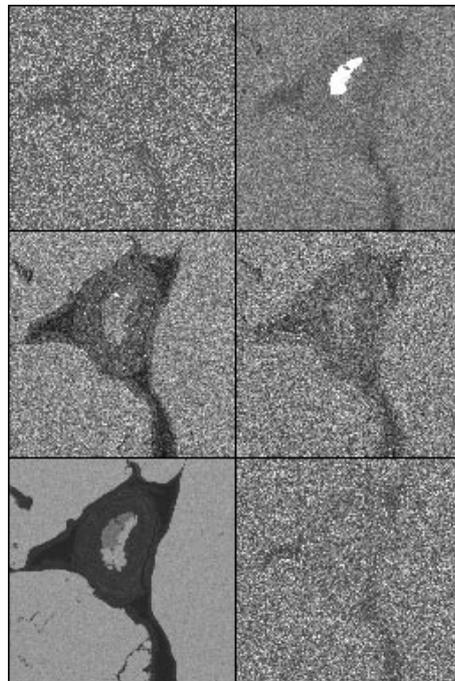


Figure 3. Elements corresponding to the training set example shown in Figure 1; row-wise: Mn, Na, P, S, Si and Ti

mum spatial range whereas some sensitivity with respect to the Euclidean distance threshold is seen. The increased misclassification rates are connected with classes where the seeding point is placed in inhomogeneous regions. This cannot always be avoided, particularly in the cases of clay classes (e.g. classes that occur as thin layers or are otherwise spatially dispersed).

For the Mahalanobis distance seed growing we see independence of the choice of quantile used for the Mahalanobis distance threshold. However, it is evident that using a continuous update of the class parameters during seed growing has a negative effect on the misclassification rates.

Finally, with respect to the interaction of the operator with the classifier through the identification of seed points, we see from randomly choosing new seed points in the vicinity of the original seed points that increased misclassification

rates may occur. Again, the problem occurs with classes that are spatially dispersed (as opposed to classes occurring as grains).

In conclusion, we have found the seed growing to be very robust and insensitive to the parameters of the algorithm. Only in the case of spatially dispersed classes a tuning of the Euclidean distance threshold is necessary for the initial seed growing. This is not a problem for classes that occur as grains. The algorithm is insensitive with respect to misclassification rate to all other parameters. Also, it is demonstrated that the update option should not be used for the Mahalanobis seed growing.

Future work should include the definition of a procedure for the initial seed growing (without parameter tuning). Also, a more thorough study examining the sensitivity to choosing seed points from different grains, from different wells, from different fields should enter into consideration. Finally, a study on sensitivity to simultaneous change in all parameters as opposed to this marginal sensitivity study would be interesting.

Acknowledgement

We thank Johannes Rykkje, Norsk Hydro Research Centre, and Mogens Ramm, Norsk Hydro Technology and Competence, for valuable contributions. We thank Johan Doré, Department of Mathematical Modelling, for coding part of the seed algorithm.

REFERENCES

1. Anderson, T. W. *An Introduction to Statistical Analysis*, second edition, Wiley, 1984.
2. Clelland, W. D. and Fens, T. W. Automated rock characterization with SEM/Image-analysis technique. *SPE Formation Evaluation*, 437–443, 1991.
3. ERDAS Inc. *ERDAS Version 7.4*. 1990.
4. Flesche, H., Nielsen, A. A., and Larsen, R. Supervised mineral classification with semi-automatic training and validation set generation in scanning electron microscope energy dispersive spectroscopy images of thin sections. Submitted, 1998.
5. Hjort, N. L., Mohn E., and Storvik, G. Contextual classification of remotely sensed data, based on an auto-correlated model. In H. V. Sæbø, K. Bråten, N. L. Hjort, B. Llewellyn, and E. Mohn, editors, *Contextual classification of remotely sensed data: Statistical methods and development of a system*. Technical report No. 768, Norwegian Computing Center, 1985.
6. Hjort, N. L. and Mohn, E. A Comparison of some Contextual Methods in Remote Sensing Classification. *Proceedings of the 18th International Symposium on Remote Sensing of Environment*. Paris, France, October 1984.
7. Hjort, N. L. Estimating parameters in neighbourhood based classifiers for remotely sensed data, using unclassified vectors. In H. V. Sæbø, K. Bråten, N. L. Hjort, B. Llewellyn, and E. Mohn, editors, *Contextual classification of remotely sensed data: Statistical methods and development of a system*. Technical report No. 768, Norwegian Computing Center, 1985.
8. Jia, X. and Richards, J. Feature reduction using a supervised hierarchical classifier. In *The 8th Australasian Remote Sensing Conference*, Canberra, Australia, March 1996.
9. Minnis, M. M. An automatic point-counting method for mineralogical assessment. *AAPG Bulletin*, 68(6):744–752, 1984.
10. Nielsen, A. A., Flesche, H., and Larsen, R. Semi-automatic supervised classification of minerals from X-ray mapping images. *Proceedings of the Fourth Annual Conference of the International Association of Mathematical Geology (IAMG'98)*, pp. 473–478, Ischia, Italy, October 1998.
11. Owen, A. A neighbourhood-based classifier for LANDSAT data. *The Canadian Journal of Statistics*, 12:191–200, 1984.
12. Safarian, S. R. and Landgrebe, D. A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(3):660–674, May–June 1991.
13. Tovey, N. K. and Krinsley, D. H. Mineralogical mapping of scanning electron micrographs. *Sedimentary Geology*, 75:109–123, 1991.