Seabed Classification from backscatter sonar data using statistical methods

Ragnar Bang Huseby, Otto Milvang, Anne Solberg, Katrine Weisteen
Image Processing group
Norwegian Computing Center
Oslo, Norway

Abstract

The development of reliable methods for automatic seabed classification enjoys widespread interest at the present time. In this paper, statistical methods for seabed classification from backscatter sonar data are investigated. The aim of classification is to divide the sea bottom into smaller regions and assign each region to one of several sediment types.

Statistical classification from raw data consists of two steps; extracting a vector of feature components from raw sonar data for each region, and assigning for each feature vector a class. In this study, raw backscatter data from the Simrad EM 1000 Multibeam Echo Sounder are used.

Several combinations of a total of 50 different features are examined systematically with respect to performance of classification. The features are based on the backscatter strength, the backscatter probability density function, the spectral distribution, and texture.

We consider classification rules which are derived from the Bayes decision rule, and involve probability models of the features. The attention is focused on the k-Nearest-Neighbor classifier and a classifier based on the multivariate normal distribution. The results show that it is possible to differentiate between seabeds of various sediment types.

1 Introduction

The present work is a part of a project which started in December 1990 with the aim of developing methods for automatic seabed characterization. Participating companies are Simrad Subsea A/S, Norwegian Computing Center (NR), and SINTEF SI (former Center for Industrial Research (SI)). From June 1992, the project was merged with the ESMAC project.

In this paper, we concentrate on statistical methods for classification of feature vectors extracted from backscatter sonar data. As is true in most fields that deal with measuring and interpreting physical events, probability considerations become important in seabed characterization because of the randomness under which data are generated. Statistical pattern recognition techniques have previously been applied to side scan sonar data [2].

The performance of a classifier depends highly on the features which are used. In [9], more than 50 different methods for feature extraction from sonar data were evaluated for this project. From the examination we know that at least 20 of the features examined may be used in seabed classification. However, only a small subset of features should be used simultaneously. Including too many features in the classification (using a feature vector with a high dimension compared to the number of training samples) results in high error rates for the classifier. This is due to a well-known phenomenon in statistics; the "curse of dimensionality" [8]. Based on the previous results [9] we have only examined classifiers using the three features: *0.8 Quantile, Pace* D_{f_2}, *and GLCM Contrast,* where *0.8 Quantile* is the 0.8th quantile of the vector of backscatter values from an area of the seabed, *Pace* D_{f_2} is a feature related to the spectral distribution of the backscatter signal, and *GLCM Contrast* is a measure related to the spatial co-occurence matrix of the backscatter values. For more details on these features, see [9] and the references given in that paper.

The classification methods of this paper are based on supervised learning, that is, preliminary knowledge about the classes is required. A set of feature vectors or pixels from each seabed type to be classified is to be collected and separated into a design set, on which the classifiers are to be trained. Due to the fact that there is a large number of nuances of seabed types, one cannot expect that all possible nuances are represented in the training set. Thus assignment of a pixel to one of the seabed types in the training set does not necessarily mean that the area corresponding to the pixel is of that type. Instead the classification of a pixel should be interpreted as the seabed type which is most similar to the seabed of that area.

In section 2, we describe classifiers based on the multivariate normal distribution and the k-Nearest-Neighbor classifier. Both noncontextual and contextual methods are considered. We also present procedures for detection of outliers and classification of pixels as a mixture of two different classes. In section 3, we describe the data and the results. A discussion of the results and comments to the methods are given in section 4.

The results show that it is possible to differentiate between seabeds of various sediment types. However, the degree of discrimination is not known because precise ground truth over large areas has not been available for the test sets. Nevertheless the classification results can provide broad guidance. In addition, our methods show a high degree of consistency when classifying multiple passes over the same area. We conclude that the methods can be used in a prototype system for seabed characterization.

2 Classification rules

2.1 Traditional noncontextual classification

In noncontextual classification [11], each pixel is classified on the basis of the data for only that pixel. The methods of this paper are based on the Bayes classification theory [4]. Thus each pixel belongs to one of K classes with prior probabilities, $\pi_1, \ldots, \pi_K$. Pixels or feature vectors from class k are distributed according to the density f_k. The Bayes decision rule [4], assigns to a pixel class k where k maximizes $P(C = k|X)$, and

$$P(C = k|X) = \frac{\pi_k f_k(X)}{\sum_{l=1}^{K} \pi_l f_l(X)}$$

is the posterior probability of class k given the feature vector X.

In this section, we assume that the class densities are multivariate normal. Thus

$$f_k(x) = (2\pi)^{-\frac{d}{2}}(\det(\Sigma_k))^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu_k)'\Sigma_k^{-1}(x-\mu_k)},$$

where d is the dimension of the feature vector space, μ_k is the mean vector, Σ_k is the covariance matrix, and $'$ denotes transpose. The parameters μ_k and Σ_k are unknown and will be replaced by estimates in computation of $P(C = k|X)$.

Training of the classifier consists of estimating μ_k and Σ_k for $k = 1, \ldots, K$. Let $X_1^{(k)}, \ldots, X_{N_k}^{(k)}$ be feature vectors which are known to be of class k. The parameters are estimated by $\hat{\mu}_k$ and $\hat{\Sigma}_k$ where

$$\hat{\mu}_k = \frac{1}{N_k} \sum_{i=1}^{N_k} X_i^{(k)}$$

and

$$\hat{\Sigma}_k = \frac{1}{N_k - 1} \sum_{i=1}^{N_k} (X_i^{(k)} - \hat{\mu}_k)(X_i^{(k)} - \hat{\mu}_k)'$$

Thus the training procedure is very simple because there are explicit formulas for the estimates.

Classification of the unknown feature vectors is now a simple task. It can be shown [4] that the classification rule becomes: assign to a pixel class k where k minimizes

$$\ln(\det(\hat{\Sigma}_k)) + (X - \hat{\mu}_k)'\hat{\Sigma}_k^{-1}(X - \hat{\mu}_k).$$

It is important to know to which extent one can trust the classification results. If the pixel is assigned to class k, a reasonable measure of uncertainty is the estimated posterior probability of class k, $P(C = k|X)$.

2.2 k-NN classification

A traditional non-parametric classifier is the k-Nearest Neighbor (k-NN) classifier [4]. As one would expect from the name, this rule classifies pixel X by assigning it the label most frequently represented among the k nearest samples from a training set (with respect to e.g. Euclidian distance). In other words, a decision is made by examining the labels on the k nearest neighbors and taking a vote. The k-NN rule is related to the Bayes decision rule [7].

The classifier is based on the Euclidian distance between the samples. The Euclidian distance is not invariant to the scaling of the features. To combine features on different scales, a normalization of each feature is required to assure zero mean and unit variance for each feature. The only parameter in the model is k. Preliminary experiments have shown that $k = 5$ is the best choice for our test data. Therefore, $k = 5$ is used in the following experiments.

2.3 Classification based on mixels

In statistical image classification, each pixel is classified to *one* of the classes in the training set. Often, however, this is not an adequate model of reality – the signals detected in one pixel may be derived from two or more different classes. This situation arises in two different settings.

- In the border zone between regions corresponding to different classes.
- When the spatial resolution is lower than the size of some of the objects in the image.

Both of these settings are found in seafloor classification based on EM 1000 data. Bottom samples indicate that the seafloor model should allow mixtures of different sediment types, and not just a number of predefined distinct classes. Failure to take the presence of mixed pixels into account may lead to misclassification.

In this section, we present a model for statistical image classification, which allows each pixel to be classified as a mixture of *two* different classes. We use the term *mixel* to denote a pixel consisting of two different classes. The classes which are allowed to form mixtures must be specified by the user.

The model for mixture classification In traditional statistical classification, $f_k(x)$ represents the class-conditional probability density of the feature vector x. For the mixture model, let $f_{k,l}(x)$ be the density for feature vector x. $f_{k,l}$ is modelled as a mixture of two densities, f_k and f_l:

$$f_{k,l}(x) = af_k(x) + (1-a)f_l(x),$$

where a is the mixture coefficient ($0 \leq a \leq 1$). For seafloor classification, f_k is assumed to be multivariate normal. The class-dependent parameters for the multivariate normal distribution ($\hat{\mu}_k$ and $\hat{\Sigma}_k$) are estimated from training data. Traditional statistical classification assigns a pixel into the class that maximizes $f_k(X)$. For a model involving mixtures, the mixture coefficient a must be estimated for each pixel for each two-class mixture so as to maximize $f_{k,l}(X)$. If $a = 1$, then the pixel consists of only class k, and if $a = 0$, the pixel consists of only class l ($0 \leq a \leq 1$). The set of allowed mixtures is to be user-specified.

2.4 Detecting outliers

Occasionally, the preprocessing machinery may have included "alien objects" and a feature vector may have been incorrectly evaluated, etc. Thus there is a need for a procedure committed to the detection of "incredible vectors", thereby avoiding incorrect forced classification.

One way in which to formalize the problem is to test the hypothesis

$$H_0 : f \in \{f_1, \ldots, f_K\},$$

where f denotes the density from which the observed candidate vector X is drawn, and $f_1, \ldots, f_K$ are the densities of the classes we want to assign the pixels [7]. If H_0 is rejected, X is defined as outlier.

If we assume that f belongs to a set of multivariate normal densities where the covariance matrix is restricted to range over a set of positive definite symmetric matrices where the minimum of the determinant is nonzero, it can be shown [7] that the *likelihood ratio* test [1] becomes: Declare X as outlier if

$$f_k(X) \leq (2\pi)^{-\frac{d}{2}}(\det(\hat{\Sigma}_{k_0}))^{-\frac{1}{2}}e^{-\frac{1}{2}\gamma_{d;1-\epsilon}}, k = 1, \ldots, K.$$

Here ϵ is the *level of significance of the test*, d is the dimension of the feature vector space, $\gamma_{d;1-\epsilon}$ is the $100(1-\epsilon)$ -th percentile of the *chi-square distribution with d degrees of freedom*, and k_0 is the class with largest $\det(\hat{\Sigma}_k)$. If $\epsilon = 0.01$, for example, in the long run 1% (or less) of the pixels belonging to one of the classes $1, \ldots, K$, will be wrongly declared as outliers.

It is also possible to construct outlier tests without any parametric assumptions on the class densities [7]. One may define X to be an outlier if all K *Mahalanobis distances* are sufficiently large,

$$(X - \hat{\mu}_k)'\hat{\Sigma}_k^{-1}(X - \hat{\mu}_k) \geq \frac{d}{\epsilon}, k = 1, \ldots, K.$$

However, this and related procedures yield very mild bounds compared to the method based on multivariate normal distribution and are therefore less powerful.

2.5 Contextual classification

In contextual classification ([3], [5], [6], [10], [11]) the classification rule is based on a stochastic model for the behavior of the classes in the scene and the behavior of the feature vectors given the underlying classes. For the classification of a pixel, the feature vectors of neighboring pixels are taken into account. Then a better classification should be obtained. The reason for this is that the contextual methods utilize correlation between neighboring pixels wheras the noncontextual methods overlook it.

As a contextual model, we use Haslett's [6] noniterative model. To explain the model, we need to define some notation. Let the $M \times N$ image consist of MN pixels or feature vectors $X_{1,1}, \ldots, X_{M,N}$, where $X_{i,j} = (X_{i,j}(1), \ldots, X_{i,j}(d))$, and d is the number of features. The scene consists of K classes, $k \in \{1, K\}$. The class of pixel (i,j) is denoted by $\mathcal{C}_{ij}$. Let $P(X(i,j) \mid \mathcal{C}_{ij} = k)$ denote the conditional probability density of $X_{i,j}$ given $\mathcal{C}_{ij} = k$. Define the neighborhood of pixel (i,j) as $\mathcal{D}_{ij} = \{(i-1,j), (i,j-1), (i+1,j), (i,j+1)\}$. The a posteriori probabilities can be written as

$$P(\mathcal{C}_{ij} = k \mid X_{i,j}, \mathcal{D}_{ij}) \propto \pi_k P(X_{i,j} \mid \mathcal{C}_{ij} = k) R_k(\mathcal{D}_{ij}),$$

where the term $R_k(\mathcal{D}_{ij})$ represents the contextual information given by

$$R_k(\mathcal{D}_{ij}) = \sum_{a,b,c,e} [\, g(a,b,c,e \mid k) \cdot h(X_{i,j-1}, X_{i-1,j}, X_{i,j+1}, X_{i+1,j} \mid X_{i,j}, k, a, b, c, e)\,].$$

Here, $g(a,b,c,e \mid k)$ is the probability of a particular configuration (a,b,c,e) of classes in the neighborhood of pixel (i,j), given that pixel (i,j) is labeled class k, and $h(\cdot)$ is the joint probability density of the feature vectors given the feature vector $X_{i,j}$ and the classes of the neighboring pixels. We will assume conditional independence between the neighboring feature vectors, a situation where h is reduced to be simply the product of the corresponding densities. In Haslett's model,

$$g(a,b,c,e \mid k) = P(\mathcal{C}_{i,j-1} = a \mid \mathcal{C}_{ij} = k) \cdot P(\mathcal{C}_{i-1,j} = b \mid \mathcal{C}_{ij} = k) \cdot P(\mathcal{C}_{i,j+1} = c \mid \mathcal{C}_{ij} = k) P(\mathcal{C}_{i+1,j} = e \mid \mathcal{C}_{ij} = k).$$

Then

$$P(\mathcal{C}_{ij} = k \mid X_{i,j}, \mathcal{D}_{ij}) = const\ \pi_k P(X_{i,j} \mid \mathcal{C}_{ij} = k) Z_k(\mathcal{D}_{ij}),$$

where

$$Z_k(\mathcal{D}_{ij}) = T_k(X_{i,j-1}) T_k(X_{i-1,j}) T_k(X_{i,j+1}) T_k(X_{i+1,j}),$$

$$T_k(X_{i,j}) = \sum_{m=1}^{K} P(\mathcal{C}_{i'j'} = m \mid \mathcal{C}_{ij} = k) P(X_{i,j} \mid \mathcal{C}_{ij} = k),$$

and

$$i'j' \in \mathcal{D}_{ij}.$$

This model reduces to the usual noncontextual maximum likelihood model if the transition probabilities $P(\mathcal{C}_{i'j'} = m \mid \mathcal{C}_{ij} = k)$ are equal ($P(\mathcal{C}_{i'j'} = m \mid \mathcal{C}_{ij} = k) = 1/K$). Then, the contextual factor, Z_k, is 1. The largest contextual effect is achieved when $P(\mathcal{C}_{i'j'} = m \mid \mathcal{C}_{ij} = k) = 1$, $m = k$, and 0 otherwise. We use the default value of the transition probabilities, $P(\mathcal{C}_{i'j'} = m \mid \mathcal{C}_{ij} = k) = 0.9$ when $m = k$, and equal probabilities are used elsewhere.

3 Data and Results

3.1 Description of data sets

The data used in this project were recorded by a Simrad EM 1000 sonar. Three series of data were available.

The first set (Set 1) was recorded at different locations in Oslofjorden. By examination of the data we selected five homogeneous regions of various types. These regions define five classes: *Seabed type 1, Seabed type 2, Seabed type 3, Seabed type 4, and Seabed type 5*, where Seabed type 1 is the hardest and Seabed type 5 is the softest. The classes may correspond to *rock, sand, silt, clay*, and *mud*, respectively, but we stress that this is not confirmed.

The second data set (Set 2) covers a cruise from Horten, around Bastø and up to Mølen in Oslofjorden. This cruise was supplied by ground examination at ten locations along the route. The seabed consists mainly of a mixture of clay, silt and sand, but there are also some areas of mud and some spots of hard bottom.

The third data set (Set 3) was recorded in an area located around Nidingen on the west coast of Sweden. The area covers different seabed types and will in the near future be supplied with ground examination.

A ping covers a sector of 150° or about 7.5 times the depth, and the sampling rate is about 6-7 samples/meter. A ping is divided into 60 beams, each covering 2.5°. In this paper, a pixel corresponds to a feature vector which is extracted from sonar data. The data come from a region covered by several neighboring beams (beam no 2 - 4, beam no 5 - 24, beam no 37 - 56 or beam no 57-59) from each of 20 consecutive pings.

For training of the classifiers, we used pixels from Set 1. The training set consisted of 80, 84, 44, 72, and 84 pixels from Seabed type 1, 2, 3, 4, and 5, respectively. The classifiers were tested on Set 2 and Set 3.

In the following, two seabed types corresponding to two consecutive numbers will be denoted as adjacent classes. This is reasonable because such classes are similar compared to other pairs of classes in the training set.

3.2 Comparison of the classification methods

For the classifiers based on the multivariate normal distribution, noncontextual (NNC) and contextual (NC), we have used the outlier criterion of $\epsilon = 0.01$. This implied that 300 of a total of 4387 pixels from Set 2 and 460 of a total of 17056 pixels from Set 3 were declared as outliers.

From *Table 1* we see that the normal distribution noncontextual classifier (NNC) and the k-NN classifier gave similar results. For both data sets, more than 80% (excluding outliers) of the pixels were assigned to the same class, while almost all pixels were classified to the same or an adjacent class. As can be seen from *Table 2*, also the NNC-classifier and the NC-classifier gave very similar results. Therefore, we have only visualized the results of the NNC-classifier.

	Set 2	Set 3
Same	80.28 %	81.06 %
Same or adjacent class	97.72 %	99.73 %
Total (exclusive outliers)	4087	16596

Table 1: *Comparison of the normal distribution noncontextual classifier (NNC) and the k-NN classifier. The table shows the number of pixels assigned to the same class and to the same or an adjacent class*

Classification plots in UTM coordinates are shown in *Figure 1* and *Figure 2*.

To a large extent, there is consistency between multiple

	Set 2	Set 3
Same class	84.46 %	89.62 %
Same or adjacent class	98.75 %	99.97 %
Total (exclusive outliers)	4087	16596

Table 2: *Comparison of the noncontextual (NNC) and the contextual (NC) classifiers. The numbers show how many pixels which were assigned to the same and the same or an adjacent class*

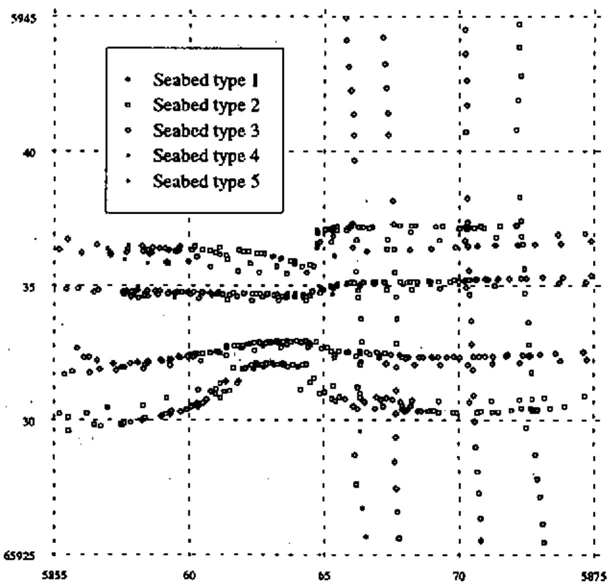

Figure 1: *Classification of a subset of Set 1. Most of the pixels were assigned to Seabed type 2 or Seabed type 3. A few pixels were classified as Seabed type 1 or Seabed type 4.*

passes over the same area. For set 2, where some bottom samples were available, the results were in reasonable accordance with the ground truth. Various seabeds consisting of clay, silt, sand and gravel were classified as *Seabed type 2* or *Seabed type 3*, while a region consisting of mud was classified as *Seabed type 5*.

3.3 Detection of outliers

In order to investigate the powerfulness of the outlier detection procedure, we have trained the NNC-classifier on training sets, from where one of the classes has been excluded. The results are given in *Table 3*. When a class was removed, many pixels assigned to that class were assigned to an adjacent class. However, pixels of *Seabed type 1*, *Seabed type 2*, and *Seabed type 5* are frequently declared as unknown when these types are not present in

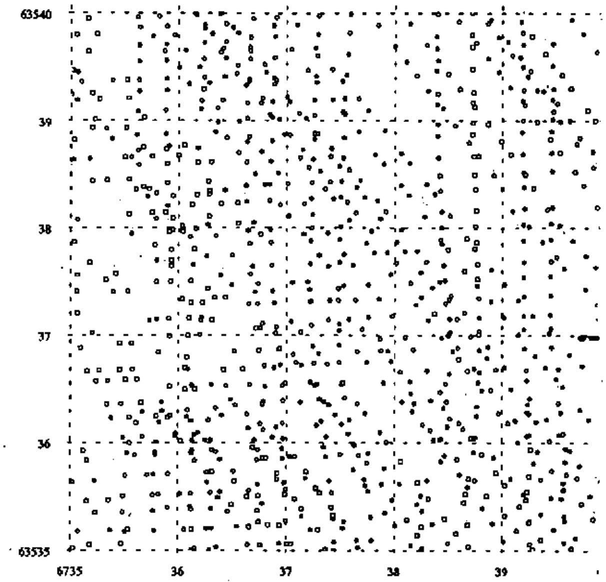

Figure 2: *Classification of a subset of Set 2. The pi were assigned to Seabed type 1 or Seabed type 2.*

the training set.

Seabed type	Set 2 Outlier	Total	Set 3 Outlier	Total
1	36.3 %	273	41.2 %	7663
2	40.6 %	1267	28.4 %	5354
3	12.7 %	1358	11.7 %	2146
4	5.8 %	445	1.6 %	1394
5	82.1 %	744	41.0 %	39

Table 3: *The ability to detect regions of unknown sea type. The second and the fourth column contain the nun of pixels assigned to the different classes when the trai set consisted of all five classes, for Set 2 and Set 3, res tively. The numbers in the first and the third column are corresponding percentages of pixels which were decl as outliers when the class corresponding to the row excluded from the training set.*

3.4 Classification of mixels

We have also investigated classification which all each pixel to be classified as a mixture of *two* diffe classes. The following mixtures have been considered:

1. Seabed type 1 and Seabed type 2
2. Seabed type 2 and Seabed type 3

3. Seabed type 3 and Seabed type 4
4. Seabed type 4 and Seabed type 5

This corresponds to allowing mixtures between a class and the adjacent classes. Overall, the mixel model gives similar results compared to the standard normal model. However, some transitions between regions of different seabed types are smoother. The results are best visualized using colors, and therefore, they are not shown in this paper.

4 Discussion

We have investigated the performance of the the k-NN classifier and classifiers based on the multivariate normal distribution for the purpose of seabed classification. The methods gave similar results when classifying the available data sets into five different seabed types.

The degree of consistency when classifying multiple passes over the same area was high. The results were in reasonable accordance with ground truth. Consequently, it is possible to differentiate between seabeds of various sediment types. However, the degree of discrimination is not known because precise ground truth over large areas has not been available for the test sets.

We have observed that the gain of using contextual classification compared to noncontextual methods is not significant. This is probably due to the fact that the pixels are computed from a large number of backscatter values and an averaging of neighboring data has taken place in the feature extraction step.

If we should select either the k-NN classifier or a classifier based on the multivariate normal distribution, we would select the latter. This is because the normality assumption yields a more powerful outlier procedure. We have demonstrated the usefulness of this procedure by detecting seabed of unknown type when excluding one of the classes from the training set. In the case of normality, we have also investigated a method for mixel classification which provides a more nuanced classification.

For a more accurate characterization, in terms of quantitative physical properties of the seabed, more research is required, and a mathematical seafloor scattering model should be established. Nevertheless the presented methods can provide broad guidance. Therefore, we conclude that the methods can be used in a prototype system for seabed characterization.

Acknowledgments

We are grateful to Trym Eggen, Simrad Subsea A/S and to Geoffrey Shippey, Chalmers University, Göteborg for interesting and fruitful discussions.

This work has been supported by the Royal Norwegian Council for Scientific and Industrial Research (NTNF) and Nordic Industry Fund. The project is a part of the NTNF research program for Image Analysis and Pattern Recognition. It is also a subproject under the ESMAC (Environmental Seabed Mapping and Characterization) program.

References

[1] P.J.Bickel and K.A.Doksum, *Mathematical Statistics, Basic Ideas and Selected Topics*, Holden-Day, 1977.

[2] M.F. Czarnecki, *An application of pattern recognition techniques to side scan sonar data*, Oceans '79, San Diego, pp. 112-119, 1979.

[3] R.C. Dubes and A.K. Jain, *Random field models in image analysis*, Journal of Applied Statistics, pp. 131-164, 1989.

[4] R.O.Duda and P.E.Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, 1973.

[5] S. Geman and G. Geman, *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*, IEEE Trans. Pattern Analysis and Machine Intelligence, 6, pp. 721-741, 1984.

[6] J. Haslett, *Maximum likelihood discriminant analysis on the plane using a Markovian model of spatial context*. Pattern Recognition, 18, pp. 287-296, 1985.

[7] N. L. Hjort. "Notes on the theory of statistical symbol recognition.", Technical Report 778, Norwegian Computing Center, 1986

[8] L.N. Kanal and B. Chandrasekran, *On dimensionality and sample size in statistical pattern classification*, Pattern Recognition, 3, pp. 225-234, 1971.

[9] O. Milvang, R.B. Huseby, A. Solberg and K. Weisteen, *Feature extraction from backscatter sonar data*, Acoustic Classification and Mapping of the Seabed, Bath 1993.

[10] A. Owen and P. Switzer, *A neighbourhood based classifier for Landsat data*, Techn. Report, Dep. of stat., Stanford University, 1982.

[11] H.V. Sæbø, K. Bråten, N.L. Hjort, B. Llewellyn, E. Mohn, *Contextual classification of remotely sensed data: Statistical methods and development of a system*. Norwegian Computing Center. Report no. 768. 100pp, 1985.