



A sensor fusion method for detection of surface laid land mines

DANIEL WESTBERG, GUSTAV TOLT, CHRISTINA GRÖNWALL

Daniel Westberg, Gustav Tolt, Christina Grönwall

A sensor fusion method for detection of surface laid land mines

Titel	Sensorfusion för detektion av landminor
Title	A sensor fusion method for detection of surface laid land mines
Rapportnr/Report no	FOI-R--2488--SE
Rapporttyp Report Type	Vetenskaplig rapport
Sidor/Pages	72 p
Månad/Month	Januari
Utgivningsår/Year	2008
ISSN	ISSN 1650-1942
Kund/Customer	FM
Forskningsområde Programme area	4. Sensorer och signaturanpassning
Delområde Subcategory	42 Sensorer
Projektnr/Project no	E3084
Godkänd av/Approved by	
FOI, Totalförsvarets Forskningsinstitut	FOI , Swedish Defence Research Agency
Avdelningen för Sensorsystem	Sensor Systems
Box 1165	P.O.Box 1165
581 11 Linköping	SE-581 11 Linköping

Sammanfattning

Landminor är ett stort problem både under och efter krigstid. De metoder som används för att detektera minor har inte ändrats mycket sedan 1940-talet. Forskning med mål att utvärdera olika elektro-optiska sensorer och metoder som kan användas för att skapa mer effektiv mindetektion genomförs på FOI. Försök som har gjorts med data från bland annat laserradar och IR-sensorer har gett intressanta resultat.

I den här rapporten utvärderades olika fenomen och egenskaper i laserradar- och IR-data. De testade egenskaperna var intensitet, IR, ytlikhet och höjd. En metod som segmenterar intressanta objekt och bakgrundsdata utformades och implementerades. Metoden använde sig av expectation-maximization-skattning och ett minimum message length-kriterium. Ett scatter separability-kriterium användes för att bestämma kvalitén på de olika egenskaperna och på den resulterande segmenteringen.

Data insamlad under en mätkampanj av FOI användes för att testa metoden. Resultaten visade bland annat att ytlikhetsmättet gav en bra segmentering för stora objekt med släta ytor, men var sämre för små objekt med skrovliga ytor. Vid jämförelse med en manuellt skapad målmask visade det sig att metoden klarade av att välja ut egenskaper som i många fall gav en godkänd segmentering.

Arbetet är även publicerad som examensarbete: D. Westberg, "A sensor fusion method for detection of surface laid land mines", Master Thesis, LITH-ISY-EX—07/4021—SE, Linköpings Universitet, Linköping, Sweden.

Nyckelord: Mindetektion, Gaussian mixtures, segmentering, expectation-maximization, minimum message length, scatter separability, infraröd, laserradar

Summary

Land mines are a huge problem both during the conflict and a long time afterwards. Methods used to detect mines have not changed much since the 1940's. Research aiming to evaluate output from different electro-optical sensors and develop methods for more efficient mine detection is performed at FOI. Early experiments with laser radar sensors show promising results, as do analysis of data from infrared sensors.

In this report, an evaluation is made of features found in laser radar and in infrared sensor data. The tested features are intensity, infrared, a surfaceness feature extracted from the laser radar data and height above an estimated ground plane. A method for segmenting interesting objects from background data using the expectation-maximization algorithm and the minimum message length criterion is designed and implemented. A scatter separability criterion is utilized to determine the quality of the features and the resulting segmentation.

The method is tested on real data from a field trial performed by FOI. The results show that the surfaceness feature supports the segmentation of larger object with smooth surfaces but gives no contribution to small object with irregular surfaces. The method generally produces a decent result of selecting contributing features for different neighbourhoods of a scene. A comparison with a manually created target mask of the neighbourhood and the segmented components show that in most cases a high percentage separation of mine data and background data is possible.

This work is also published as: D. Westberg, "A sensor fusion method for detection of surface laid land mines", Master Thesis, LITH-ISY-EX—07/4021—SE, Linköpings Universitet, Linköping, Sweden.

Keywords: Mine detection, Gaussian mixtures, segmentation, expectation-maximization, minimum message length criterion, scatter separability criterion, infrared, laser radar

Contents

1	Introduction	1
1.1	Background	1
1.1.1	Description of the problem	1
1.1.2	Description of the task	2
1.2	Previous work	2
1.3	Outline	3
2	Sensors and sensor data	5
2.1	3D scanning laser radar	5
2.2	IR System	6
2.3	Sensor data	6
2.3.1	3D data	7
2.3.2	Intensity	8
2.3.3	IR Data	10
2.3.4	Features	10
3	Sensor fusion method for mine detection	11
3.1	Mixture models	12
3.2	The EM algorithm	12
3.3	Initialization of the EM algorithm	13
3.4	Minimum Message Length criterion	14
3.5	Scatter Separability criterion	15
3.6	Scatter Separability normalization	17
3.7	The complete method	17
3.8	Simple segmentation	20
4	Results	21
4.1	Feature elimination and selection priority	21
4.2	Detection of a medium sized mine	23
4.3	Detection of a small mine	28
4.4	Detection of a large object	32
4.5	Result from non-mine data	35
4.6	Summary	39

5	Test on large area	41
5.1	Detection of multiple objects in a large area	41
6	Discussion	45
6.1	Summary	45
6.2	Conclusions	46
6.3	Future directions	46
	Bibliography	49
A	Description of the data sets	51
B	Result plots (real data)	53
B.1	Set m1	53
B.2	Set m2	54
B.3	Set m3	55
B.4	Set m4	56
B.5	Set m5	57
B.6	Set m6	58
B.7	Set m7	59
B.8	Set m8	60
B.9	Set m9	61
B.10	Set m10	62
B.11	Set m11	63
B.12	Set m12	64
B.13	Set m13	65
B.14	Set m14	66

Chapter 1

Introduction

1.1 Background

This report is commissioned by the Swedish Defense Research Agency (FOI), Linköping, Sweden. It is a part of the Multi Optical Mine Detection System project, MOMS [7], at the division of Sensor Systems. The MOMS project strives to evaluate if a realization of an electro-optical (EO) multi sensor system for mine detection is achievable, and, if possible, produce specifications for a demonstrator system.

1.1.1 Description of the problem

Land mines are a huge problem in time of conflict, limiting mobility of forces and material. After a conflict, land mines that are left oppose a great threat for individuals who reside in affected areas. In the 1940's advances was made in electronics which led to the possibility of developing portable metal detectors. The current methods for mine clearance are essentially unchanged since then [11]. The time needed to clear land varies enormously depending on local conditions, but the quantity of mines hardly affects the clearing time of an area. There are quick and safe ways of disposing mines once they are found and identified. It is finding them that takes time and is difficult. Land mines are usually very simple devices and readily manufactured anywhere. Two common types are: anti-vehicle or anti-tank (AT) mines, and anti-personnel (AP) mines. AT mines are comparatively larger, often laid in unsealed roads or potholes, and detonate when a vehicle drives over them. They are typically activated by force, magnetic influence, or remote control. AP mines are much smaller and are usually activated by force or tripwires [11].

In recent research at FOI, a method for segmenting surface laid mines placed on a gravel road using data from laser radar has been investigated and shown some promising results. This method relies on a fusion between intensity and height data received from a laser radar sensor. Intensity is usually a good feature for separating mines from background data, but not sufficient. There may be other non-mine objects with intensity that differs from the background, and the

intensity measurement can differ due to the conditions. On the gravel road, which is a relatively flat surface, height above the ground plane is a feature that helps the separation. When the vegetation surrounding the mine is more complex, the height feature worsens the separation of the mine from the background, which motivates a search for other features. The scene in Figure 1.1 contains 11 mines of different model, even though none of the mines are buried they are not that easy to find.



Figure 1.1. A scene containing 11 mines of different models and an ammunition box.

1.1.2 Description of the task

In this report, 3D data received from the laser radar have been examined and processed to find features relevant for mine detection in various vegetations. These features can vary with the vegetation. Data from an infrared (IR) sensor have been synchronized with the 3D laser radar data and included as an additional feature. The features have been evaluated to determine the combination that gives a robust anomaly detection. A signal processing method has been proposed and implemented. The method was evaluated with laser radar data from real scenes. Only surface laid mines are considered in this report.

1.2 Previous work

Assuming that data retrieved from a laser radar, 3D data (voxel) and intensity, can be described with a two kernel Gaussian mixture. Expectation Maximization (EM) can be used to estimate the mixture parameters. Bayesian hypothesis testing for two classes can then be applied to segment object and background [6].

In a natural forest environment, fabricated objects such as vehicles will typically represent the largest structured objects in the laser radar data. A structured

object can be recognized by searching for surfaces that can be considered flat in local neighborhoods of sufficient size. This property can be exploited for detecting of vehicles, using principal component analysis, PCA, of data partitions [2].

Multi- and hyper-spectral data can be used to find objects that do not fit in a spectral model of the background, for example man-made objects in a natural environment [1].

Cluster analysis is the process of finding “natural” groupings by grouping “similar” (based on some similarity measure) objects together. Some of the features used to find this “natural” grouping may be redundant, some may be irrelevant, and some can even misguide clustering results. In addition, reducing the number of features increases comprehensibility and avoids the problem that some unsupervised learning algorithms break down with high dimensional data. Issues involving the developing of automated feature subset selection algorithms for unsupervised learning are explored in [4, 5, 9].

1.3 Outline

Chapter 2 displays a short summary of the specifications of the sensors used to collect the raw data. Information of features received directly or estimated from the raw data is also described in this chapter. Chapter 3 covers all the steps of the method implemented, and displays some examples. In Chapter 4 result are found from four typical data sets, results from sets not covered in this chapter can be found in Appendix B. Chapter 5 contains example of a large set. A conclusion of the work and some thoughts and directions for future work is found in Chapter 6.

Chapter 2

Sensors and sensor data

In this chapter a short presentation of the laser radar sensor system and the IR sensor system is given. We take a look at generated data and features that can be used directly or extracted from this data.

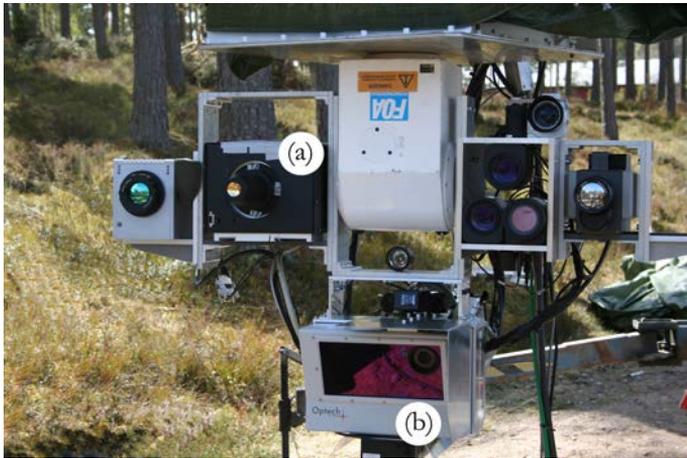


Figure 2.1. Sensors mounted on the skylift used at the Eksjö trials. (a) The infrared sensor (MultiMIR) (b) The laser radar sensor system (ILRIS-3D).

2.1 3D scanning laser radar

The laser radar sensor system used to retrieve the (raw) data in this report was an ILRIS-3D manufactured by Optech Inc. in Canada, see Figure 2.1. The laser is of the eye safe class 1 type, and has the operating wavelength of $1.5 \mu\text{m}$. Specifications for the laser radar are found in Table 2.1. The ILRIS-3D creates a point cloud of geometric samples and intensity value from the surface of the scanned area.

Table 2.1. Specifications for the ILRIS-3D

Parameter	Data
Wavelength	1.5 μm
Maximum range Accuracy	350 m (4% reflectance); 800 m (20% reflectance) X-Y @ 100 m \pm 10 mm Z @ 50 m \pm 10 mm, Z @ 100 m \pm 10 mm
Field of view	40° (\pm 20°, programmable, horizontal and vertical)
Divergence	0.2 mrad
Range Resolution	1 cm
Angle Resolution	0.2 mrad
Sampling Frequency	2000 points/s
Working Temperature	0°C – +40°C
Size	30 x 30 x 20 cm
Weight	12 kg

2.2 IR System

The MultiMIR is a multi spectral infrared sensor with a sensitivity in the spectral area of 1.5 – 5.2 μm , see Table 2.2. The sensor was delivered from AEG Infrarot-Module (AIM) and has gone through several software modifications at FOI in Linköping. It is based on a cooled MCT detector (78K) and a staring focal plane array with 384 \times 288 pixels. The sensor is based on a spinning filter wheel containing four filters which makes the collection of four different bands possible.

Table 2.2. Specifications for the MultiMIR

Parameter	Data
Spectral bands	
<i>SWIR</i>	1.5 – 1.8 μm and 2.1 – 2.5 μm
<i>MWIR</i>	3.5 – 4.0 μm and 4.5 – 5.2 μm
Temperature resolution(NETD)	< 25mK at 300K
Sampling Frequency	100 images/s

2.3 Sensor data

The data was collected at a field trial in Eksjö [8]. The scenes were scanned from a sky lift. Three different areas were scanned, a forest area with dense undergrowth, a gravel road, and an area where vegetation had recovered from having been completely removed some years ago (Figure 2.2). The board that is present in all three areas is a reference board used for calibration purposes.



Figure 2.2. The three mine scenes at the Eksjö trial: (A) Forest with dense undergrowth. (B) Gravel road. (C) Recovered vegetation

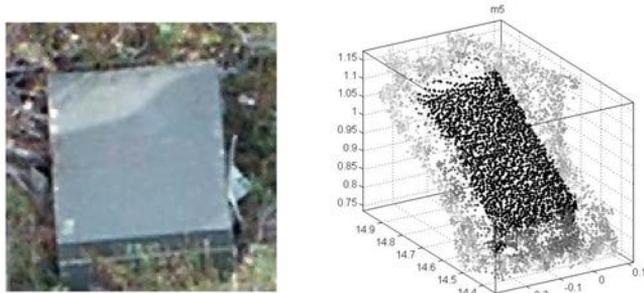


Figure 2.3. Picture of sample area and a plot of the 3D point cloud representation. The rectangular shape in the picture is an ammunition box.

2.3.1 3D data

The raw 3D data comes from the ILRIS-3D sensor. In each sample of the scanned image, a 3D position value is received and an intensity value. Due to self-occlusion (we cannot see through objects), no sample are received from areas behind or under objects. It would be possible to get a more complete representation of the scene by combining multiple images from various positions, this method will not be used in this report and it might not be applicable scenario for mine detection. The 3D point cloud representation of the ammunition box can be seen in Figure 2.3.

Height

On flat surfaces it has been shown that height can be used to improve the segmentation [6]. When the background becomes more complex by increased vegetation the height above the ground plane will not always contribute to the segmentation. Figure 2.4 shows the height values for the ammunition box. The height feature is measured as the distance in each sample from an approximation of the ground plane (using PCA) from the 3D data.

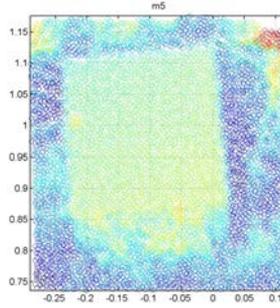


Figure 2.4. Height image from the ammunition box, grass surrounding the ammunition box is almost of the same height as the box. The variation in height is lower over the box, so it can still be visually distinguished from the background.

Surfaceness

As mentioned in [2], manufactured objects will typically represent a large structured objects in laser radar data. In cluttered environments, we expect such an object to be more structured than the surrounding environment.

A measure of “structure” is obtained through local fitting of a set of 2D surfaces to a point and its neighbors. Each point is thus assigned a value that represents how well a surface fits to the data surrounding this point. It was found that by only considering the residual between the fitted surface and the points did not perform adequately, as points in the (unstructured) background could get significant values by chance. However, it was seen that the estimated normal direction of points in the background typically varied considerably. Hence, the *surface smoothness value* (S) here defined in terms of residual distance between the points and the surface *and* the normal direction similarity. For a particular point p , S is written as

$$S_p = \sum_{i \in N} n_i s_i, \quad (2.1)$$

where N defines the points in the neighborhood, n_i denotes the similarity (a scalar product) between the estimated normal at point p and the normal at the corresponding point on the surface, and s_i denotes the proximity between point p and the surface, that equals 1 when the distance is zero, decreases for increasing distances and equals 0 beyond a distance threshold. In this way, only points that lie very close to the surface and have normal directions similar to that of the fitted surface contribute significantly to the surface smoothness value. Figure 2.5 shows the surface similarity values for the ammunition box.

2.3.2 Intensity

Intensity is an important feature for finding objects that differ from the background, but not sufficient. Military objects, like mines, are often painted with a color that returns a low intensity value. The ammunition box is a good example of this, see Figure 2.6.

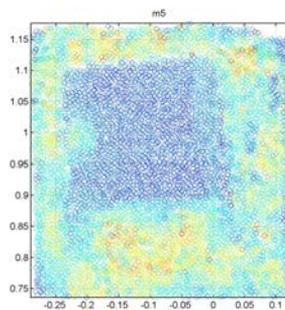


Figure 2.5. Surface similarity image from the ammunition box. The flatness rate is highest in the middle of the large flat surface of the box and gets slightly lower at the edge of the box.

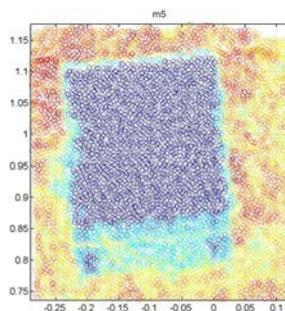


Figure 2.6. Intensity image from the ammunition box. The paint on the ammunition box gives a lower intensity value than the surrounding vegetation.

2.3.3 IR Data

The IR data gathered with the MultiMIR sensor must be matched with the data from the laser radar sensor [8]. The IR data also have a lower resolution than the laser radar data so an interpolation has to be done to cover all the laser radar samples. A visual inspection was performed to make sure the data at least covers the same area, but an offset of a few centimeters is possible. Figure 2.7 shows the interpolated IR data for the ammunition box, a fuzziness can be seen at the border of the box due to the interpolation.

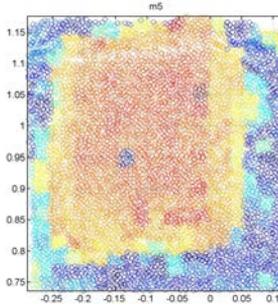


Figure 2.7. IR image for the ammunition box. The data from the IR sensor is not of the same resolution as the laser radar data, an interpolation was done to cover the empty samples. The fuzzy edges of the ammunition are artifacts from the interpolation.

2.3.4 Features

From the three scenes in Figure 2.2, smaller data sets were created. Each set covered a ground area of about $0,4 \times 0,4m^2$, with close to 6500 laser radar samples per set. A feature vector set was created for every set, consisting of the four features (mentioned in Section 2.3.1-2.3.3): intensity, IR, surfaceness and height. Height is still part of the feature set used, but its priority is low compared to the other features in the set, further discussed in Section 4.1. In summary, each sample point is associated with a d -dimensional feature vector ($d = 4$ in this case), the complete set could be viewed as a $n \times d$ feature vector set, where d is the total number of features and n is the sample total of the area.

Chapter 3

Sensor fusion method for mine detection

Several problems have to be overcome in the process of developing an automatic method for mine detection. The data is multi-faceted, many phenomena are registered, and hence methods that can combine these phenomena are wanted. Assuming that the data are samples from mixtures of Gaussian distributions reduces the problem partly to a “missing parameter problem”. The parameters of the Gaussian mixture models can be estimated with the Expectation Maximization algorithm, EM. The EM algorithm is initialization sensitive so an effort has been made to estimate “good” starting values. In previous research, [6], it has been assumed that there are only two natural groupings of the examined data (object of interest and background) and the number of Gaussian components in the EM algorithm was set to two. This assumption might not always be true, for example, there might be more than one interesting object in the data that does not naturally group with the other objects or the background might be better described with more than one component thereby reducing the number of miss-classified samples. In this chapter, the theories behind the parts of the method are exposed along with a description of how the problematic steps in the method have been handled. Starting with the assumption that the underlying distribution of the data is Gaussian, a description of the Gaussian mixture model is given in Section 3.1. In Section 3.2 we take a look at the different steps in the EM algorithm. A description of the Minimum Message Length (MML) criterion and how it is used to select the number of Gauss mixtures (model order) is found in Section 3.4. Finally the Scatter Separability (SS) criterion, a way of evaluating the features sets, is presented in Section 3.5 and a description of the simple segmentation applied after the method is presented in Section 3.8.

3.1 Mixture models

A mixture model consists of a sum of independent variables. In this case, these variables are Gaussian distributions also known as normal distributions. Each Gaussian can be described with two parameters, location and scale, more commonly known as mean and standard deviation. The reason for using this distribution is simply that many physical phenomena can be well approximated with it; it is well known and commonly used in many other fields. Distributions that are more suited for mine data may exist, but no effort has been made to evaluate this in this report.

The probability density function for a Gaussian mixture can then be written as:

$$P(\mathbf{y}|\Theta) = \sum_{m=1}^k \alpha_m p(\mathbf{y}|\theta_m) \quad (3.1)$$

$$\theta_m \equiv [\mu_m, \sigma_m] \quad (3.2)$$

$$\Theta \equiv [\theta_1, \dots, \theta_m, \alpha_1, \dots, \alpha_m] \quad (3.3)$$

$$p(\mathbf{y}|\theta_m) \in N(\mu_m, \Sigma_m) \quad (3.4)$$

$$\alpha_m \geq 0, \quad m = 1, \dots, k, \quad \sum_{m=1}^k \alpha_m = 1 \quad (3.5)$$

where $\mathbf{y} = [y_1, \dots, y_n]^T$ is the given feature vector with n samples, k is the number of Gaussian components, $p(\cdot|\cdot)$ is the Gaussian probability function, α_m is the relative weight between each Gaussian with the constraints in (3.5), θ_m contains the mean, μ_m , and covariance matrix, Σ_m , parameters for each component m .

3.2 The EM algorithm

The EM algorithm is a common choice for estimating mixture parameters. It is an iterative algorithm that computes the maximum likelihood estimates of the parameters. $Y = \{\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(d)}\}$ is a feature set consisting of d feature vectors each with n samples. Y is seen as an incomplete data set, missing is a set of d labels, $Z = \{\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(d)}\}$ where $\mathbf{z}^{(i)} = [z_1^{(i)}, \dots, z_n^{(i)}]^T$ is a n -dimensional vector associated with the n samples. Label $z_j^{(i)} = m$ indicates that sample j in feature vector i belongs to component m , $m = 1, \dots, k$. If we had the complete data set, $X = \{Y, Z\}$, then Θ could easily be estimated. The EM algorithm works in two alternating steps, the expectation step (E-step) and the maximization step (M-step). The E-step produces a sequence of estimates of Θ , from estimates of the unobserved Z ($W \equiv E[Z|Y, \hat{\Theta}(t)]$) conditioned on the observations using values from the last M-step, $\hat{\Theta}(t)$. $W = [\mathbf{w}^{(1)}, \dots, \mathbf{w}^{(d)}]$ and $\mathbf{w}^{(i)} = [w_1^{(i)}, \dots, w_n^{(i)}]^T$

are estimated from

$$w_j^{(i)} \equiv Pr \left[z_j^{(i)} = m | y^{(i)}, \hat{\Theta}(t) \right] \quad (3.6)$$

$$= \frac{\hat{\alpha}_m(t) p(\mathbf{y}^{(i)} | \hat{\Theta}_m(t))}{\sum_{j=1}^k \hat{\alpha}_j(t) p(\mathbf{y}^{(i)} | \hat{\Theta}_j(t))} \quad (3.7)$$

this plugs in to the Q function,

$$Q(\Theta, \hat{\Theta}(t)) \equiv E \left[\log p(Y, Z | \Theta) | Y, \hat{\Theta}(t) \right] \quad (3.8)$$

$$= \log p(Y, W | \Theta). \quad (3.9)$$

The M-step then updates the parameters according to,

$$\hat{\Theta}(t+1) = \arg \max_{\Theta} Q(\Theta, \hat{\Theta}(t)).$$

The EM algorithm will alternate between these two steps until a convergence criterion is met, in this case until the estimation of the loglikelihood (3.9) does not improve with more than 10^{-6} in each step, or when the maximum number of iterations is reached (40 iterations). More information on the EM-algorithm can be found in [5, 3].

3.3 Initialization of the EM algorithm

The EM algorithm relies heavily on good initialization. If too many components are placed in one region of the space, and too few in another, the algorithm will not be able to move components across low likelihood regions. In addition, when the assumed number of components is larger than the true number of components, at least one of the weights (α) may approach zero which affects the covariance matrix, that then may become close to singular.

From a smoothed 1D histogram of a feature vector in set Y , the location of the top five local maxima are chosen as mean, μ_m , start values. If five maxima cannot be found, the ones found are used. The weights, α_m , are distributed equally between the start values and the standard deviation, σ_m is given a small constant value. Figure 3.1 shows the histogram, with the five start values marked, for the Intensity feature from a set containing background data and a AP land mine of model M18A1 (Claymore).

Results from two 1D EM-MML model estimations are used to create the start values for a 2D EM-MML model estimation. Figure 3.2(a) shows the first 2D EM-MML model estimation received from the start values created by 1D EM-MML estimations of intensity and IR features from the scene mentioned earlier. The 1D EM-MML model estimation for the intensity, initialized with values from histogram displayed in Figure 3.1, returns a mixture model consisting of three Gaussians. The 1D EM-MML model estimation for the 2^{nd} feature (IR) also returns three Gaussians. This gives the nine (3×3) start values seen in Figure 3.2(a).

Figure 3.2(b)-(c) shows a selection of models estimated for a complete run of the method and Figure 3.2(c) is the chosen model since it has the lowest MML value. Expanding to higher dimension is simply done by combining one of the remaining features 1D EM-MML model estimation to the current multi dimensional estimation (in this case the model in Figure 3.2(c) because it has the lowest MML value) and using this as initialization values.

No real effort has been made in the initializing step to avoid the over-fitting problem that leads to a singular covariance matrix (see Figure 3.2(a)), except setting the upper limit of mixtures to five in the 1D case, but in the MML part of the method components with small weights or small covariance are removed first.

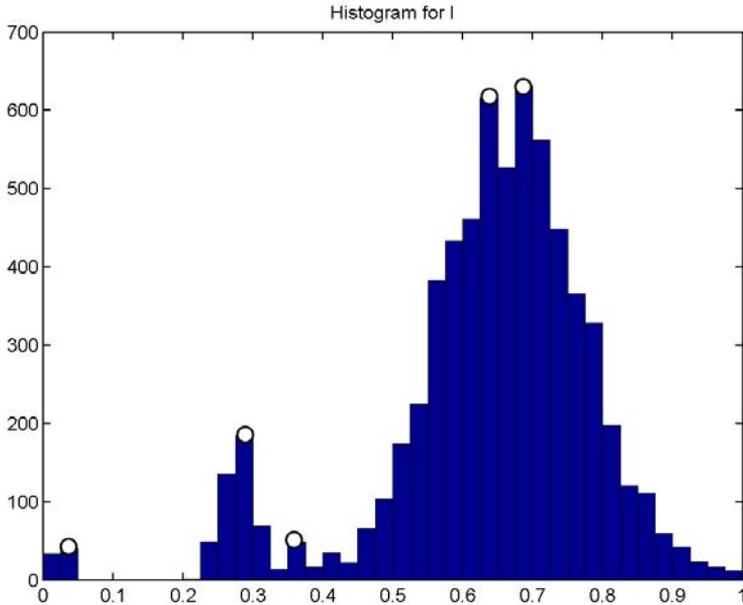


Figure 3.1. A 1D-histogram with the top five local maximums marked with circles.

3.4 Minimum Message Length criterion

When estimating the parameters with the EM algorithm it is required that the number of mixtures is known, which is not the case for unsupervised methods. There are several different methods that estimate the number of mixtures that best describes the data.

An information theory approach has been used to estimate the number of mixtures. The underlying idea is that if a short code can be built for the data, the data generation model is good, this is described more in depth in [5]. According to Shannon's theory the shortest code for data set Y measured in bits is $\lceil -\log p(Y|\Theta) \rceil$, where $\lceil a \rceil$ denotes "the smallest integer no less than a ." Since

$-p(Y|\Theta) \gg 1$ for moderately large data sets, the $[\cdot]$ operator can usually be ignored. In the unsupervised case when Θ have to be estimated the total code length can be written in two parts,

$$L(\Theta, Y) = L(\Theta) + L(Y|\Theta). \quad (3.10)$$

The estimated parameter is the one minimizing $L(Y|\Theta)$. For a deeper understanding of minimum encoding length criterias, see [9]. Making some approximations and simplifications to (3.10) the MML criterion formulated in [5] is obtained as,

$$L(\Theta, Y) = \frac{D}{2} \sum_{m:\alpha_m>0} \log\left(\frac{n\alpha_m}{12}\right) + \frac{k}{2} \log \frac{n}{12} + \frac{k(D+1)}{2} - \log p(Y|\Theta), \quad (3.11)$$

D is the parameter total for the distribution, k the total number of components and n the sample total. This is the MML criterion that will be used in this paper to select the model for the feature sets.

3.5 Scatter Separability criterion

The number of mixtures, k , are highly dependent on the input features, Y . The features are added iteratively in this algorithm, starting with one feature and expanding with one feature at a time. There are two major things to be addressed here. First, in what order should the features be added? Second, does adding another feature improve the result? A Scatter Separability (SS) criterion, proposed in [4], is used to evaluate both these issues. The selection of feature order was done by a experimental evaluation, an SS score was calculated for each feature for several different sample images of mines, the feature with highest score in multiple sets was given the highest priority, the second highest next highest priority, and so on. Since this is an averaging over the features of many sets, the feature order may not be optimal for all the sets, but should give a good result over all. This evaluation could of course be done for each set individually, but then a better criteria for "good features" is needed. The scatter separability criterion is a measure of how separated the clusters are and how compact each cluster is. If the clusters are well separated, they should give a better segmentation. If the samples in a cluster are compact, the probability is high that the distribution estimation is correct. The within-class scatter matrix, S_w , and the between-class scatter matrix, S_b , are defined as,

$$S_w = \sum_{m=1}^k \alpha_m E \left\{ (Y - \mu_m)(Y - \mu_m)^T | \omega_m \right\} = \sum_{m=1}^k \alpha_m \Sigma_m \quad (3.12)$$

$$S_b = \sum_{m=1}^k \alpha_m (\mu_m - M_o)(\mu_m - M_o)^T, \quad (3.13)$$

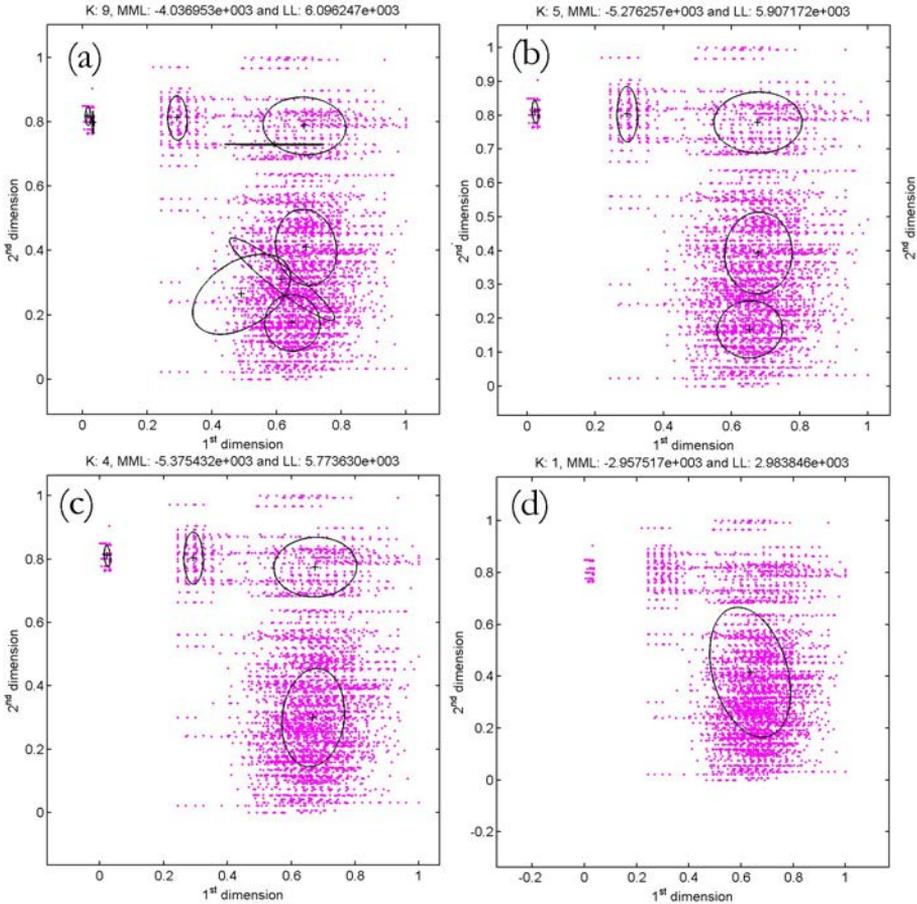


Figure 3.2. EM-MML estimations from a set containing a Claymore AP mine, intensity is the 1^{st} dimension and IR is the 2^{nd} dimension. (a) 9 mixtures, this is the first estimation, initialized by the data from the 1D estimation from both features. One of the components has a close to singular covarians matrix. (b) 5 mixtures, the component with the singular covariants matrix has been removed and some of the other components have been merged. (c) 4 mixtures, the merge of the two components furthest down in the last image can clearly be seen (lowest MML). (d) 1 mixture, the method tries to fit all samples in one component, since the background has most samples the center of the Gaussian will be located closer to the location of the background samples.

where

$$M_o = E\{Y\} = \sum_{m=1}^k \alpha_m \mu_m \quad (3.14)$$

is the total sample mean, α_m (weight) is the probability that an instance belongs to mixture ω_m , Y is a d -dimensional feature vector representing the data, k is the number of mixtures, μ_m is the sample mean vector of mixture ω_m and Σ_m is the covariance matrix of mixture ω_m . The SS criterion is defined as,

$$SS = \text{trace}(S_w^{-1}S_b). \quad (3.15)$$

A high value of SS is wanted, because this equals a maximization of the between class scatter matrix, S_b , and a minimization of the within-class scatter matrix, S_w .

3.6 Scatter Separability normalization

The scatter separability criterion generally gives a higher value when you increase the feature dimension, because of this a comparison between the results from a feature set and the same set with an additional feature cannot be done directly. In [4], an approach is taken to normalize the scatter separability values with respect to dimension. The proposed normalization function is,

$$\text{normalizedValue}(s_j, C_j) = SS(s_j, C_j) \cdot SS(s_{j+1}, C_j), \quad (3.16)$$

where s_j is a subset of $Y = [y_1, \dots, y_d]$ ($s_j = [y_1, \dots, y_j]$), where $j = 2, \dots, d-1$, C_j is the number of mixtures given by the MML criterion after the EM estimation of s_j , $SS(s_{j+1}, C_j)$ is the SS value for set s_{j+1} when the number of components are C_j . When the results from different subsets is identical, $C_j = C_{j+1}$, then (3.16) will be,

$$\text{normalizedValue}(s_j, C_j) = \text{normalizedValue}(s_{j+1}, C_{j+1}). \quad (3.17)$$

If the normalized values of sets s_j and s_{j+1} are equal, then adding this feature to s_j will not improve the separability. This applies if the normalized value of s_j is larger than the normalized value of s_{j+1} . If the normalized value of set s_{j+1} is larger than the normalized value of s_j the separability is better in the higher dimension set.

3.7 The complete method

The EM-MML method is initialized with k mixtures, received from the histogram evaluation covered in Section 3.3. The EM model estimation is made for each $m = k, \dots, 1$. A threshold is set that removes weights (α 's) that are too small between each estimation, to avoid the problem with a singular covariance matrix. If none of the weights (α 's) are close to zero, the two mixtures with the shortest Euclidean distance between them will be selected and the mixture of those two with

the smallest weight (α) will be removed. After one component has been removed the remaining are used as initialization for the next EM model estimation, this continues until $m = 1$. An MML value is calculated for each model m . In Figure 3.2 some steps of the EM-MML process can be viewed for a selection of components. Comparing Figure 3.2(a) with 3.2(b), the removal of mixtures with small weights (α) can be noted, in Figure 3.2(c) the Euclidean distance rule has been utilized and Figure 3.2(d) shows the last model estimated. The model with the lowest MML value is the model who finally is chosen (Figure 3.2(c) in this example). As mentioned in Section 3.3, the parameters for the selected model is used to create the initialization values for a higher dimension feature set, if there is an improvement between these sets is evaluated with a comparison of the normalized SS value for both sets. This method is similar to the one used in [4].

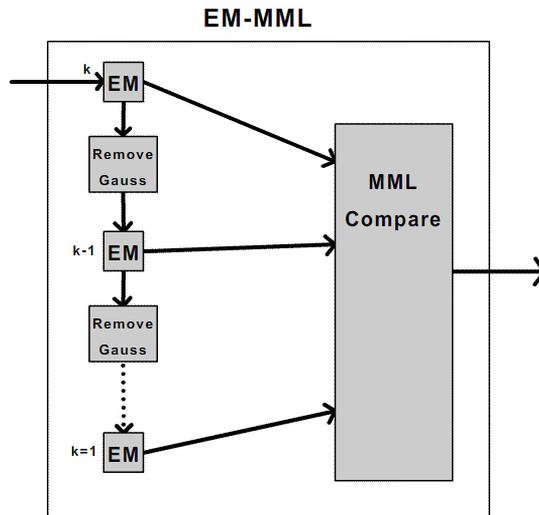


Figure 3.3. Overview of the EM-MML algorithm. Input to the first EM is the initialization parameters talked about in Section 3.3. Output is the EM estimation with the largest MML value.

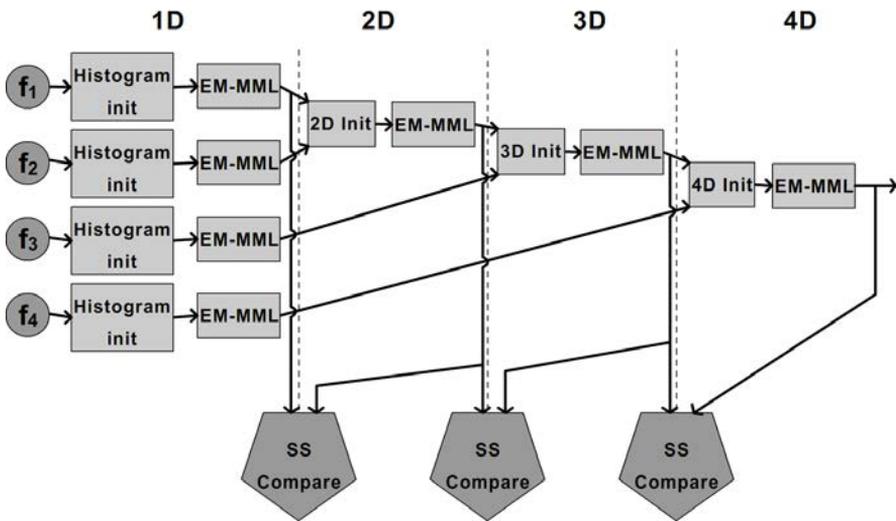


Figure 3.4. A simple overview of the method. f_1 , f_2 , f_3 and f_4 are the input feature vectors. A scatter separability comparison is done over the dimensions to help decide if the recently added feature contributes to more separation.

3.8 Simple segmentation

When all the parameters for the Gaussian mixtures have been estimated in the EM-MML method it is possible to calculate a probability density function (pdf) for each Gaussian from the samples with the estimated parameters. The segmentation is done by calculating the pdf of each Gaussian for a sample and label the sample to belong to the Gaussian with the highest probability density in that sample point. In Figure 3.5 the resulting Gaussian components from the EM-MML estimation of a Claymore mine is shown and in the same figure the result from the simple segmentation is shown. Classification statistics are shown in Table 3.1.

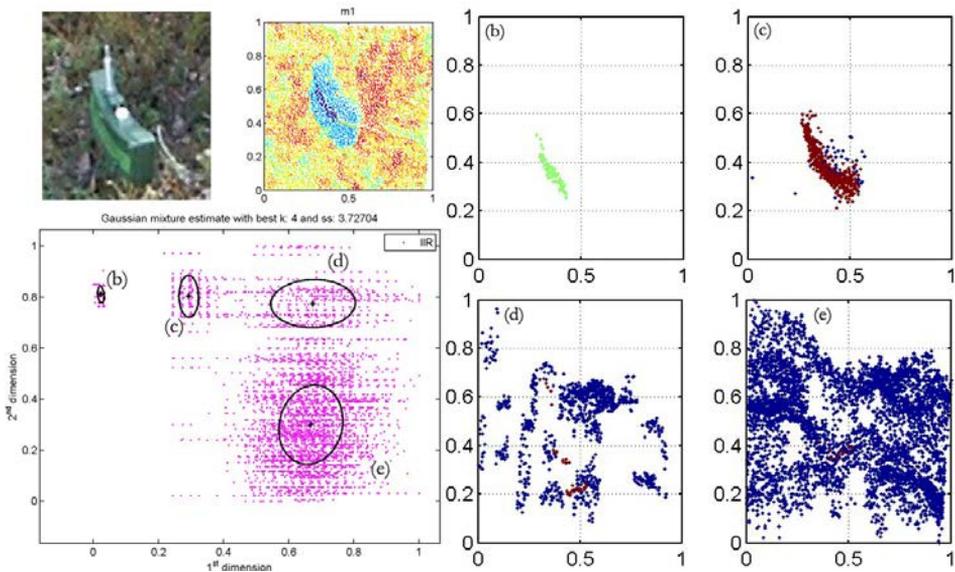


Figure 3.5. In the top left corner there is a photograph and a point cloud representation, laser radar data, of a Claymore mine. The picture in the bottom left corner is the Gaussian components received from the EM-MML estimation and to the right the result from the simple segmentation. The letter in each image corresponds to the Gaussians in the bottom left picture.

Table 3.1. Statistics for each component in Figure 3.5 (Claymore mine). True detection, false detection and missed samples are shown. Component b contains samples from the tape on the mine. Most mine samples are in component c. Component d and e consist mainly of background data.

Index	True	False	Miss
component b	75	0	511
component c	467	46	119
component d	33	977	553
component e	11	4869	575

Chapter 4

Results

The method in Chapter 3, was tested on 14 different data sets, where 8 sets contained mines of various kind and 6 sets only contained rocks, sticks and various forms of vegetation, see Appendix A for more information on each set. Each data set consists of a feature set, where each feature contains about 6500 samples from a $0,4 \times 0,4m^2$ square of the ground.

A target mask has been made for all the sets containing a mine. This was done manually, by placing a polygon around the samples that belong to the mine. The quality of the mask varies from each set, vegetation and a small target area sometimes makes it difficult to make a perfect separation. However, as long as most of the mine is marked as a target, a comparison between the mask and the output will give some indication of how “well” the segmentation went. The first part of this chapter covers the initial feature evaluation used to decide which features to include and in what order they are added to create the higher dimension feature sets.

Results from four typical data sets are presented in this chapter. These sets consist of one medium sized mine, one small mine that is hard to detect, one large object and one non-mine data set. Figures showing the result from the EM-MML estimation for each set are presented, and figures showing each component received from the segmentation with the feature set selected by the method. Each component plot has a corresponding table containing the data from the comparison with the target mask. Plots of the results from the remaining sets, that are not covered in this chapter, are found in Appendix B.

4.1 Feature elimination and selection priority

It is important to find features of “good” quality, but it is also crucial to eliminate features that give an insignificant contribution, a negative contribution or no contribution at all to the segmentation. Positional features, x and y , is an example of features that was dropped due to being of a degenerating nature. Nevertheless, they might be of interest in the post processing stage, more on this in Chapter 6. The initial evaluation of which features to keep and which to eliminate was done by

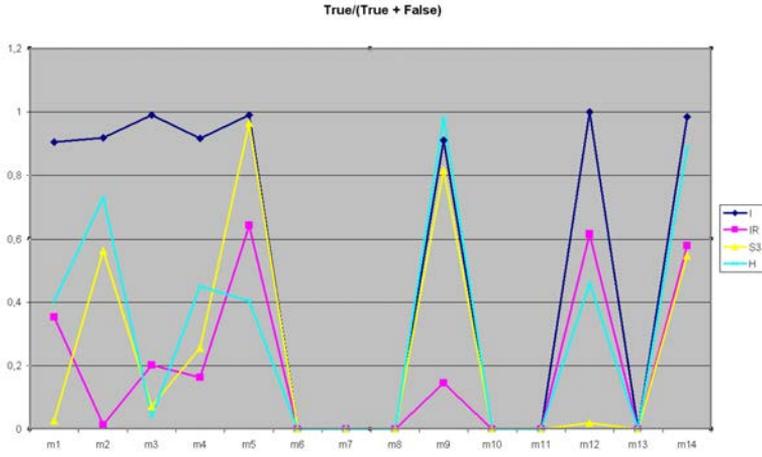


Figure 4.1. The $True/(True + False)$ ratios for each feature (intensity (I), infrared (IR), surfiness (S3) and height (H)) for all sets. True is the number of samples in the component that also are in the target mask, False is all the samples included in the component that is not in the target mask.

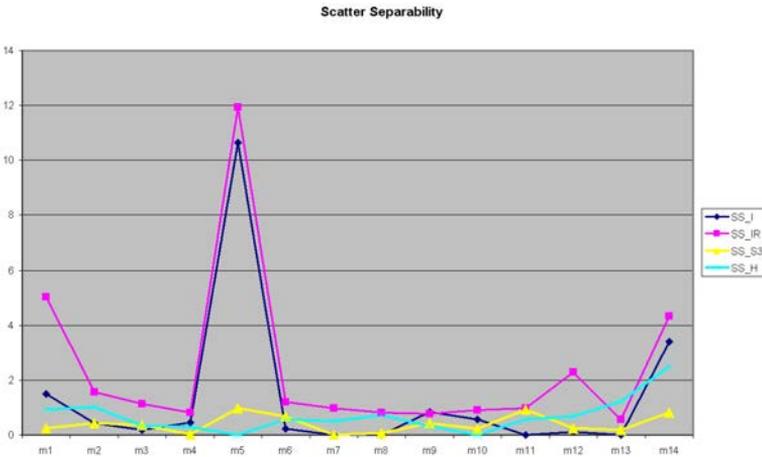


Figure 4.2. The scatter separability value for all features (intensity (I), infrared (IR), surfiness (S3) and height (H)) and sets.

running the method with one feature at a time and doing a visual inspection after the segmentation and a comparison with the target mask. The features remaining after the initial elimination are intensity, IR, surfaceness and height, all covered more deeply in Chapter 2. On these features an individual EM-MML estimation and segmentation was done again to decide the priority, or order to add features, to create higher dimension feature sets. In Figure 4.1 results from the comparison with the target mask and the estimations and segmentation can be seen for all the sets. Intensity is the feature that gives the best separation on its own for most of the sets, determining the feature second best for overall separation is harder. Looking at the scatter separability value for each feature, see Figure 4.2, and calculating a priority as described in Section 3.5 gives Table 4.1. IR gives the highest scatter separability value over all, intensity second, surfaceness third and height is last, which endorse earlier statements about height being a degenerating feature in complex environments. A compromise was made between the comparison with the target mask and the scatter separability values when deciding the final priority order, since intensity clearly gives the best segmentation on its own most times it was chosen as the “base” feature, then each feature was added in the priority order in Table 4.1. An evaluation as mentioned in Section 3.5 is performed each time a feature is added giving an indication if adding another feature improves the separability.

Table 4.1. Feature priority order calculated from all sets containing mines or objects of interest.

Priority	Feature	Score
1	IR	31
2	Intensity	20
3	surfaceness	16
4	Height	13

4.2 Detection of a medium sized mine

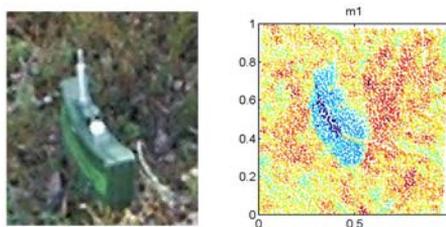


Figure 4.3. Image and point cloud representation of set m1, which contains a medium sized AP mine of the Claymore type.

The mine in Figure 4.3 is of model M18A1, also known as “Claymore” (set m1 in Appendix A). It is a directional fragmentation, high explosive AP land mine which may be electrically or non-electrically initiated [10]. The histogram

with selected start values, marked with circles, and the result from the EM-MML estimation are displayed in Figure 4.4(a)-(h). The scatter separability value is highest for IR and intensity, which once again indicates that they are “good” base features. The surfaceness feature receives the lowest scatter separability value of the four, probably because a larger “flat” surface is needed for “good” surface detection.

A visual inspection of Figure 4.5 indicates that the Gaussian components generated by the EM-MML algorithm, with intensity and IR as input features gives a decent fit. The mine data separates into the two components in the top left corner. The reason for this is that there is a small tape on the mine, which has a different intensity value than the rest of the mine. The background also separates into two components. The result from the segmentation of the set with each component plotted separately can be seen in Figure 4.6.

Table 4.2 shows statistics of how each component match the target mask. Component b is the small tape on the mine mentioned earlier, component c is the rest of the mine, while components d and e consist mainly of background data.

Table 4.2. Statistics for each component in Figure 4.6 (Claymore mine). True detection, false detection and missed samples are shown. Component b contains samples from the tape on the mine. Most mine samples are in component c. Component d and e consist mainly of background data.

Index	True	False	Miss
component b	75	0	511
component c	467	46	119
component d	33	977	553
component e	11	4869	575

Adding the surfaceness feature creates a 3D feature space, the projections from the results of EM-MML algorithm can be seen in Figure 4.7.

As in the 2D case, four components are chosen for the model, the normalizing equation (3.16) selects the 2D set over the 3D in this case. This was indicated in the low scatter separability value for the surfaceness.

A visual inspection of the plots in Figure 4.7 verifies the conclusion that adding this feature does not improve the separability of the components, the fit is still decent, but the weight for the tape on the mine becomes too small and is not separated from the rest of the mine data.

The component plots received from the estimation of the 3D feature set are shown in Figure 4.8. They show that most of the target is in one component, one could argue that this also is a plausible result, since the tape and the mine data is merged to one component. Statistics for each component is shown in Table 4.3.

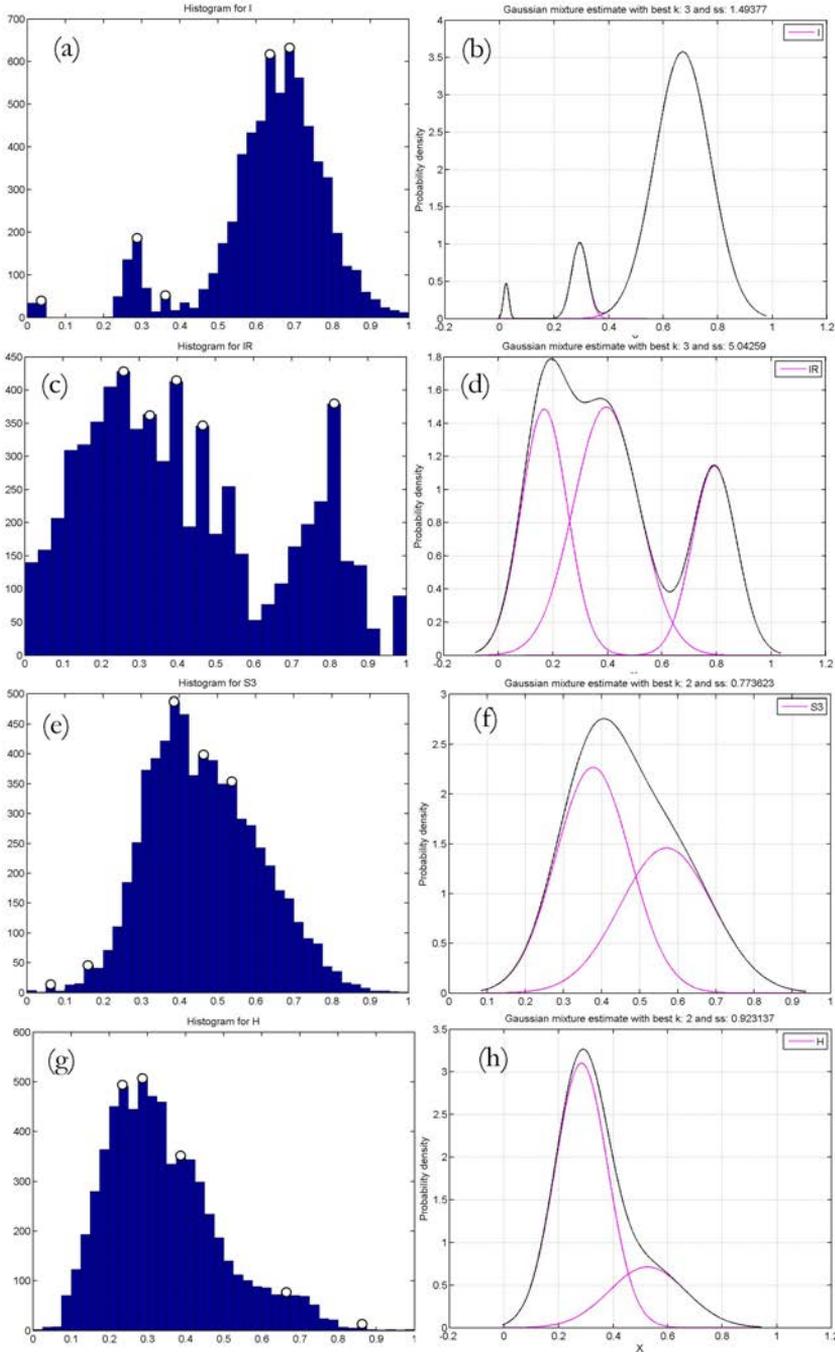


Figure 4.4. Histograms (left) and EM estimations (right) for the medium sized Claymore mine. (a)-(b) intensity, (c)-(d) IR, (e)-(f) surfaceness, (g)-(h) height.

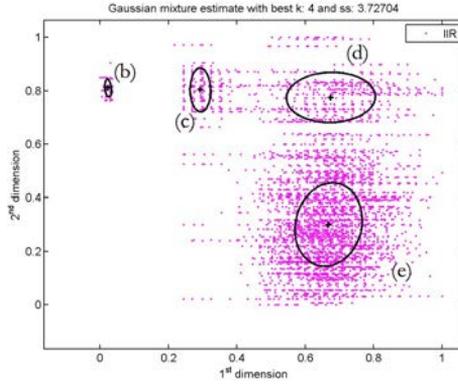


Figure 4.5. The mixture model with the lowest MML value after EM estimation for a medium sized Claymore mine with intensity, 1st dimension, and IR, 2nd dimension, as input features. In Figure 3.2, on Page 16, several steps in the mixture model process can be viewed

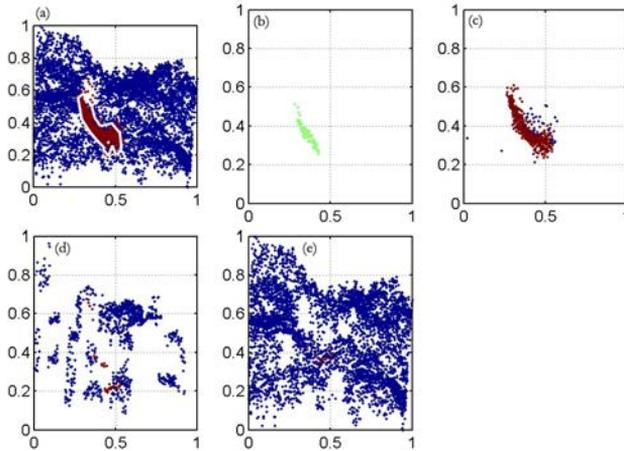


Figure 4.6. Segmentation from the Claymore mine data set: (a) complete data set, the mine data is gray. (b)-(c) Mostly mine data, (b) contains samples from the tape on the mine and (c) the rest of the mine data. (d)-(e) Ground data and some samples of mine data.

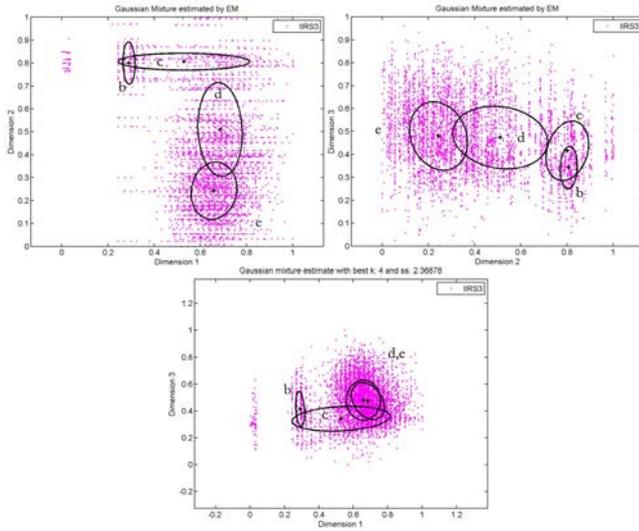


Figure 4.7. Projections of the result from the EM-MML estimation of a Claymore mine with intensity, 1st dimension, IR, 2nd dimension and surfaceness, 3rd dimension, as input features.

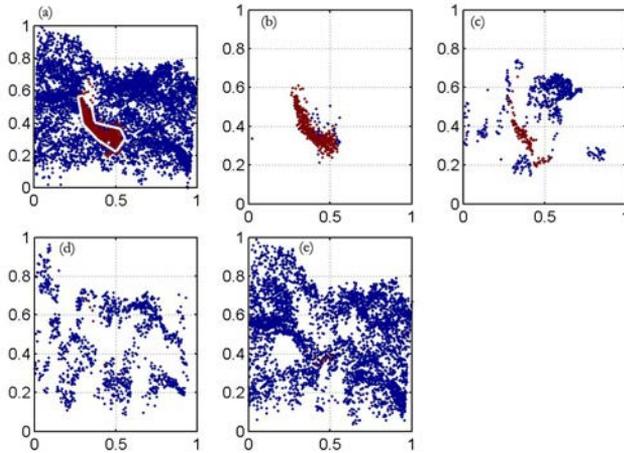


Figure 4.8. Segmentation from the 3D feature set for the Claymore mine, (a)Data set, mine data gray (b) Mine data, merge of the tape and the mine data. (c) Background data and some mine samples. (d)-(e) Mostly background data.

Table 4.3. Statistics for each component in Figure 4.8 of the Claymore mine. Corresponding Gaussians can be seen in Figure 4.7.

Index	True	False	Miss
component b	455	38	131
component c	114	567	472
component d	5	1105	581
component e	11	4148	575

4.3 Detection of a small mine

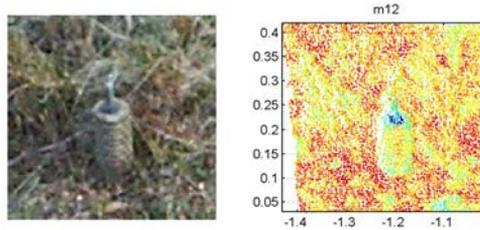


Figure 4.9. Photograph of a Russian AP mine and its point cloud representation.

The mine in Figure 4.9 is of model OZM-72, a Russian AP mine developed from earlier versions of the same type in order to achieve greater efficiency, mine m12 in Appendix A. Detonation is normally accomplished from a tripwire fuse, but command-detonated, tension-release, or simple pressure fuses could also be used [10].

The point cloud representation in Figure 4.9 gives an indication that the intensity feature will not be of much help. In Figure 4.10(a)-(b) the histogram and EM-MML estimation for the intensity verifies this. The estimation from the IR feature indicates a possible object, see Figure 4.10(d). The scatter separability value for the surfaceness is very small, but more than one component is found, Figure 4.10(f). The low value is probably a combination of noise in the data being of the same size as the object and that it is small and has an irregularly shaped surface. The height feature gives a slightly better separation, Figure 4.10(h).

Starting with the intensity as a base in this case is a bad choice, since it only has one component it will not improve the segmentation, Figure 4.10(b), but removal of a “bad” features has not been implemented so it is still used as a start feature. Adding the IR feature will improve the segmentation, the same goes for the surfaceness and height features even though their scatter separability values are low. The 4D feature set gives the best segmentation, projections from the EM-MML estimation can be seen in Figure 4.11.

Visually determining the quality of the estimation is hard in this case. Three of the Gaussians are very close and maybe a merge would have been appropriate. The two remaining Gaussians are very scattered, but since the quality of the input

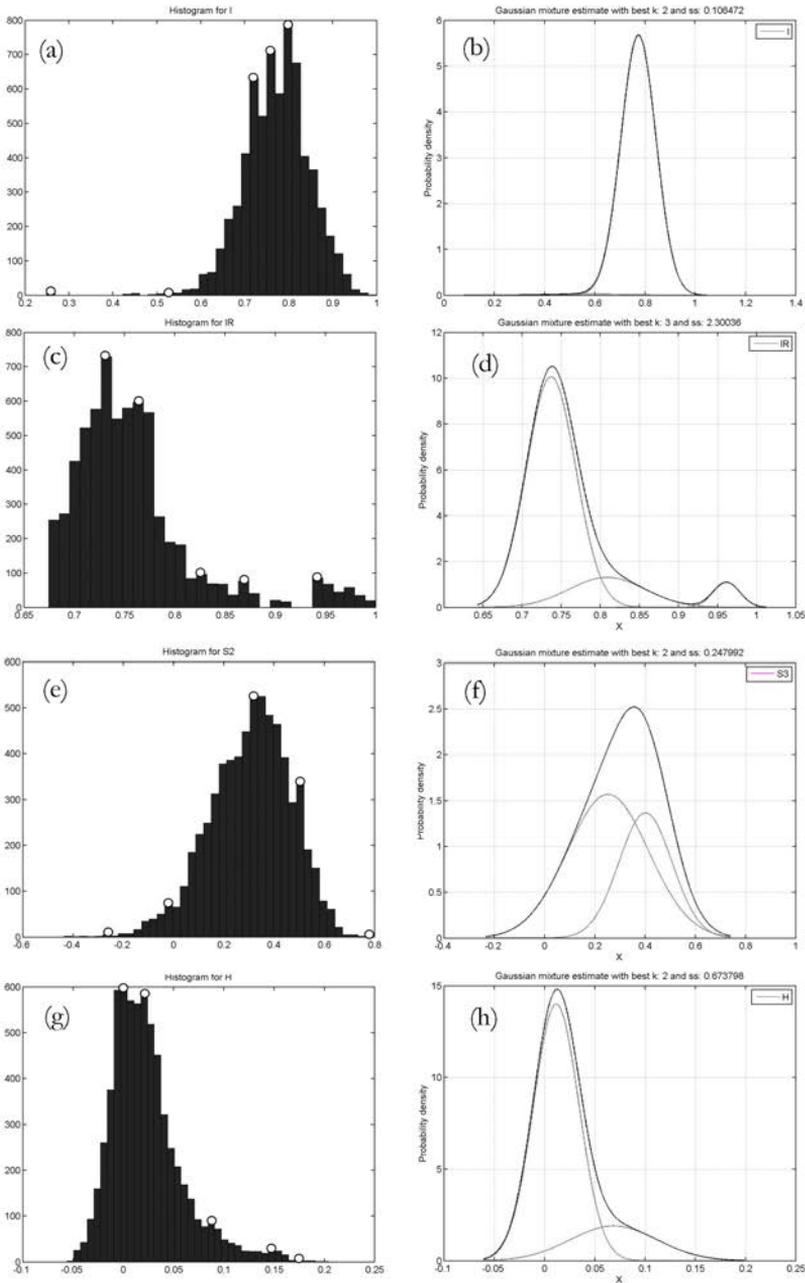


Figure 4.10. Histograms (left) and EM estimations (right) for a small irregular shaped Russian AP mine (a)-(b) intensity, (c)-(d) IR, (e)-(f) surfaceness, (g)-(h) height.

features were not that high, better results are not expected.

Component b in Table 4.4 has over a hundred true samples and only four false, this could be seen as a good result if so many samples had not been missed. The number of true samples is also high in component c, but there are many falsely marked samples in that component too.

Table 4.4. Statistics for each component in Figure 4.12 of the a small irregular shaped Russian AP mine.

Index	True	False	Miss
component b	129	4	250
component c	191	116	188
component d	23	490	356
component e	15	3816	364
component f	14	1725	365

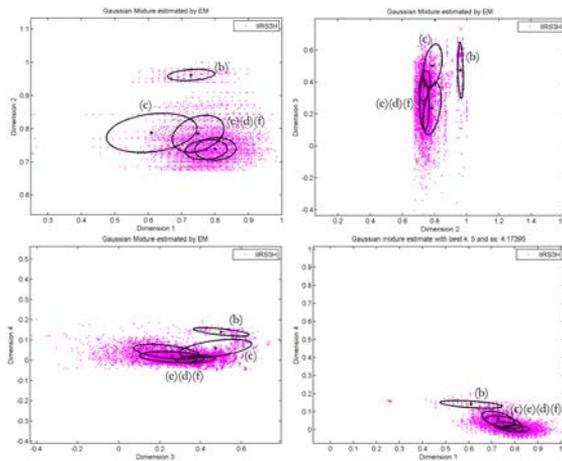


Figure 4.11. 4D EM-MML estimation projections for the set containing a Russian AP mine, intensity, 1^{st} dimension, IR, 2^{nd} dimension, surfaceness, 3^{rd} dimension and height, 4^{th} dimension.

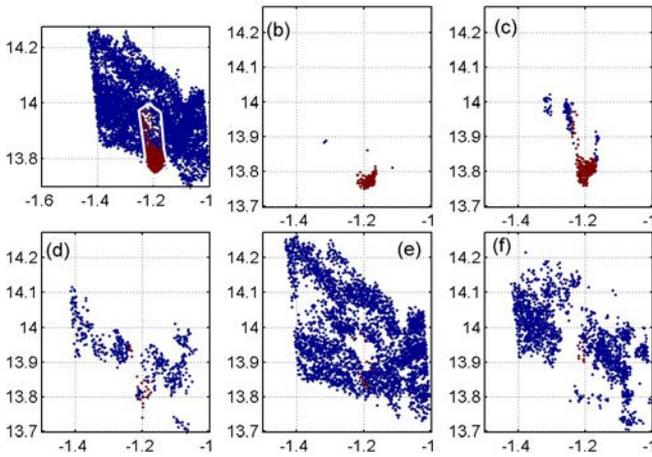


Figure 4.12. Components from 4D feature set of the Russian AP mine, (b) and (c) contain most of the mine samples, but (c) includes many false samples too, see Table 4.4. (d), (e) and (f) consist mostly of background data.

4.4 Detection of a large object

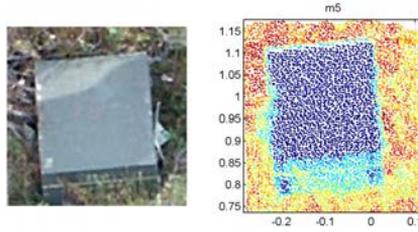


Figure 4.13. Picture of the object in set m5 (ammunition box) and point cloud representation of the area.

The target in this set is an ammunition box, see Figure 4.13. It is a large square box that takes up most of the samples in the set. Histograms and EM-MML estimation of each feature of the set can be viewed in Figure 4.14(a)-(h). Most grass surrounding the box have the same height as the box, but a few straws rise above the edge of the box, see the spike the histogram in Figure 4.14(g). These straws are not enough to make a contribution to an extra component in the EM-MML estimation as can be seen in Figure 4.14(h). Since only one component can be found in the height feature its scatter separability value is zero and it will not contribute to a better segmentation.

The box has a big flat surface, this makes it a good candidate for the surfaceness feature. Figure 4.14(f) shows the two components. The scatter separability value for the surfaceness is fairly low compared to the intensity and the IR, this displays one of the weaknesses in the scatter separability criterion when used this way. Only two different areas are found with the surfaceness, the box and the background, this leads to only two components, which actually is a good result, but the scatter separability criterion usually gives a higher value for more components. All features, except for height, contribute to the separation. The projections from the 3D feature set is shown in Figure 4.15. Visually inspecting the figure, it seems to be a good fit once again, maybe one of the Gaussians is a bit small and some seems to be close, but since it is only projections, they cannot be judged individually. Many samples from the edge of the mine ends up in the components in Figure 4.16(b)-(c), many of these may be faulty samples in the target mask. A variation of the intensity or IR at the edges or sides of the box is also likely. Figure 4.16 consist mainly of mine samples, Table 4.5 shows only one False hit. The number of missed samples is high, but as mentioned, this could be because of false target samples in the target mask.

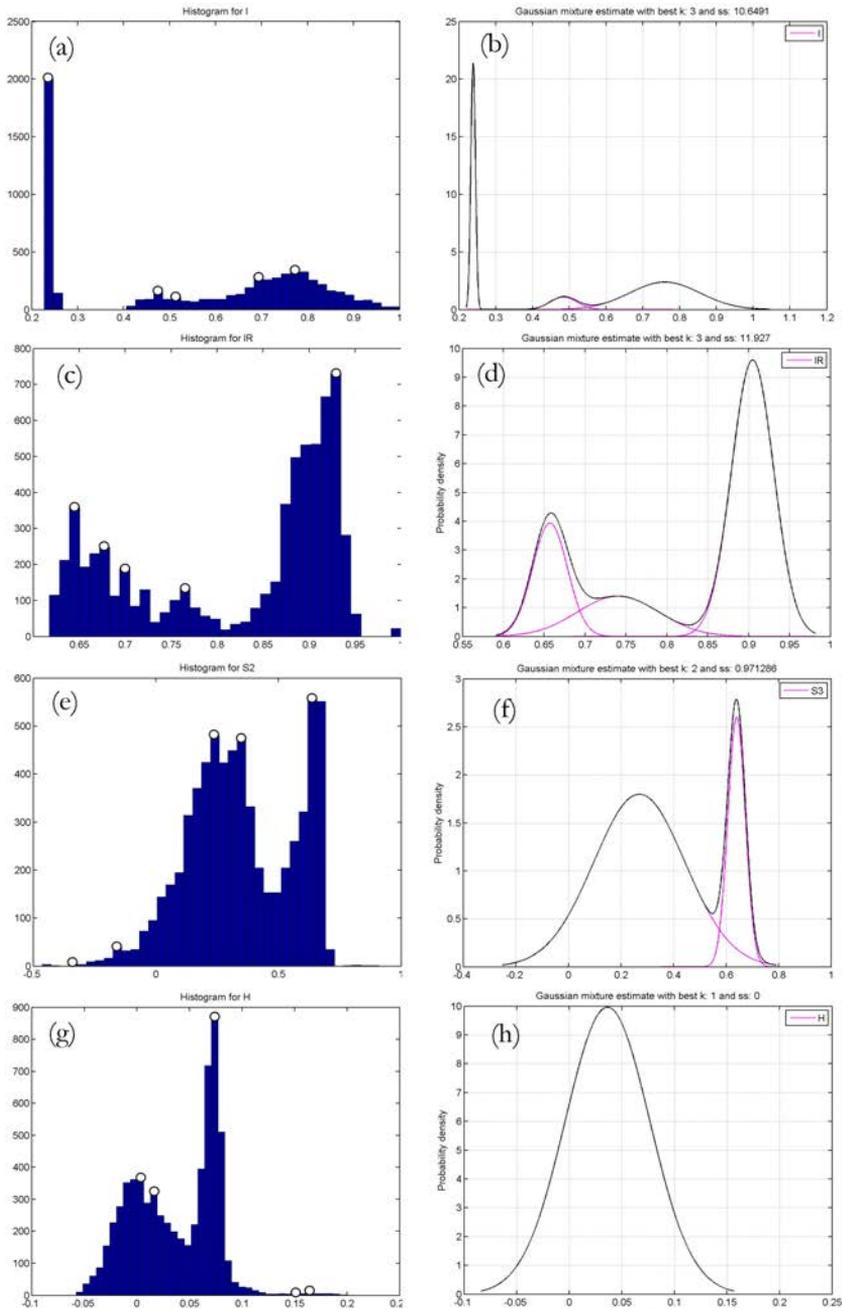


Figure 4.14. Histograms (left) and EM estimations (right)for set 5, an ammunition box, (a)-(b) intensity, (c)-(d) IR, (e)-(f) surfaceness, (g)-(h) height.

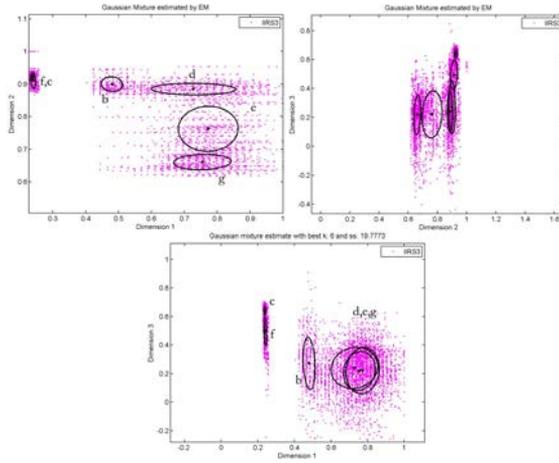


Figure 4.15. Result from 3D EM-MML estimation of the set containing an ammunition box, intensity, 1^{st} dimension, IR, 2^{nd} dimension and surfaceness, 3^{rd} dimension.

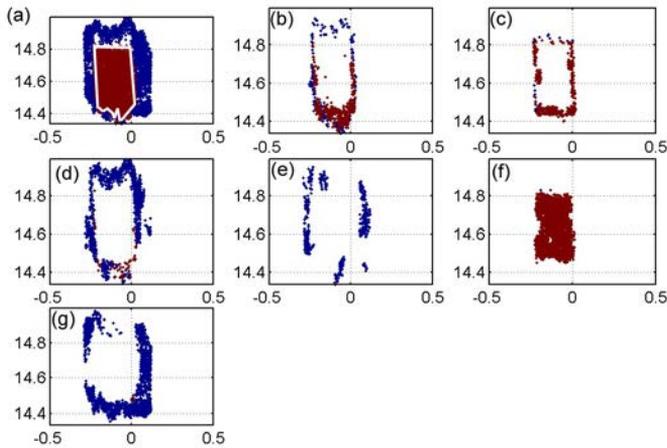


Figure 4.16. Segmentation of data set from the set containing an ammunition box (a) Complete data set, mine data is gray. (b)-(c) Data from the edge of the mine. (f) Most of the mine data. (d)-(e) and (g), mostly background data.

Table 4.5. Statistics for each component of the ammunition box in Figure 4.16.

Index	True	False	Miss
component b	465	230	2187
component c	420	21	2232
component d	57	1215	2595
component e	1	527	2651
component f	1682	1	970
component g	4	4148	2648

4.5 Result from non-mine data

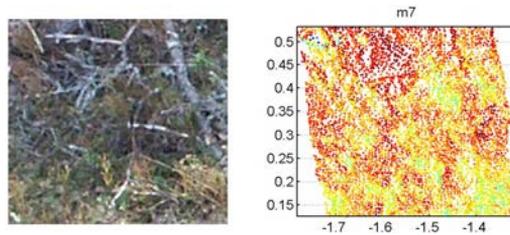


Figure 4.17. Photograph (left) and point cloud representation (right) of a data set without a mine.

The set in Figure 4.17 contains no mine data, the set consist mainly of grass and some of branches and twigs. The intensity values does not vary much in the vegetation, see Figure 4.18(a), the EM-MML estimation for the intensity, Figure 4.18(b), returns one component as the best fit, the same result is retrieved from the surfaceness, Figure 4.18(f). Both the IR and height features returns two components, Figure 4.18(d) and Figure 4.18(h), with very low scatter separability values, which indicates that even if there might be an object in this set the probability of a good segmentation is really low.

There are no mine data in this set so a comparison with a target mask is not possible. Table 4.6 displays the scatter separability values for each feature and for the higher dimension feature sets. Adding IR to the intensity feature improves the separability, which should not be a surprise since the SS value for the intensity is zero. Intensity is not a good base feature in this case since it only consists of one component. The surfaceness feature does not improve the separability for the same reason as the intensity, only one component is found so the separability is zero. The lack of improvement can be seen in Table 4.6, the 2D set has the same number of components as the 3D set. The height set actually improves the separability in this case. Two components are found in both the IR and Height data, both with low SS value. This lead to that the maximum number of found components for the complete feature set cannot be larger than four. The last row in Table 4.6 shows the components and SS for an estimation with the complete

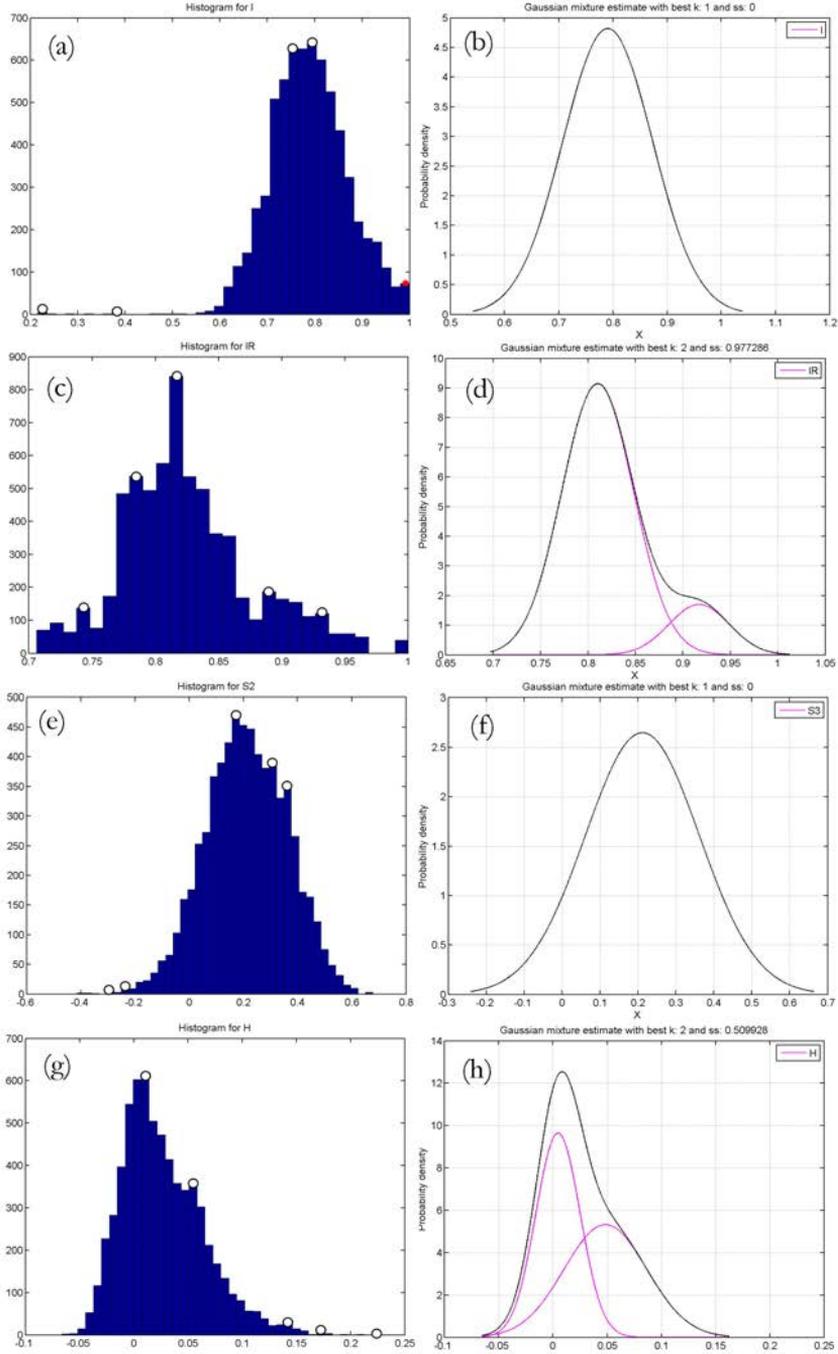


Figure 4.18. Histograms (left) and EM estimations (right) for a data set without a mine, (a)-(b) intensity, (c)-(d) IR, (e)-(f) surfaceness, (g)-(h) height.

set, the maximum number of four components are found. Reservations should be made here since this probably means that the phenomena registered in the IR and Height does not describe the same attributes.

Table 4.6. Number of estimated components and scatter separability statistics for all features from the set without a mine and for the higher dimension features sets created from these features.

Features	Components	SS
Intensity	1	0.00
IR	2	0.97
surfaceness	1	0.00
Height	2	0.51
Intensity + IR	2	0.15
Intensity + IR + surfaceness	2	0.20
Intensity + IR + surfaceness + Height	4	1.44

Figure 4.19 shows the plots of the projections from the EM-MML estimation from the 4D features set. All the Gaussians are grouped considerably tight which was indicated in the low SS value. Judging the projections visually, we might consider that one or maybe two components would be a sufficient model for the data. The IR and height features had low SS values and since only these two features contributed to the separation this in some way supports the idea of fewer components.

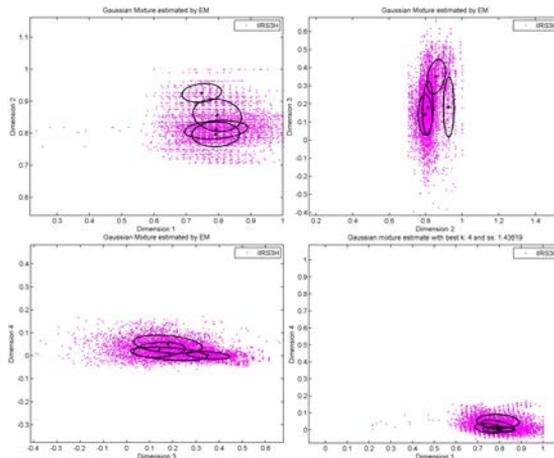


Figure 4.19. 4D EM-MML estimation for set with no mine data, intensity, 1st dimension, IR, 2nd dimension, surfaceness, 3rd dimension and height, 4th dimension.

The resulting components are shown in Figure 4.20. With some imagination, one could argue that there might be an object in the top left component. This

could be the result from a branch or some other difference in the vegetation. The other components do not show anything interesting visually. The low SS values for all the features and feature set combinations in this run should certainly alarm the user that either there is no interesting object in this set or that the quality of the features used are too poor.

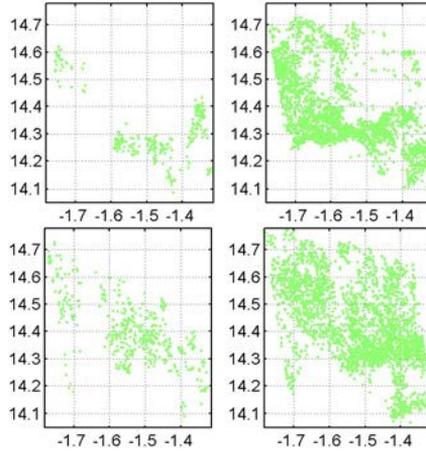


Figure 4.20. Components for the data set without a mine. No clear object is found in the visual inspection except for the top left component that may have traces of branches or some other debris.

4.6 Summary

A target mask was created manually for all the sets. The method has been applied to each set and the result from each step has been compared to the target mask. In Figure 4.21 the highest $True/(True + False)$ ratio received from each set and which features that produced this result can be seen. This graph indicates that if the “right” feature set is chosen, we get a good segmentation measured according to the $True/(True + False)$ ratio in most of the sets. Set m6, m7, m8, m10, m11 and m13 do not have a $True/(True + False)$ ratio since there are no mines in these sets.

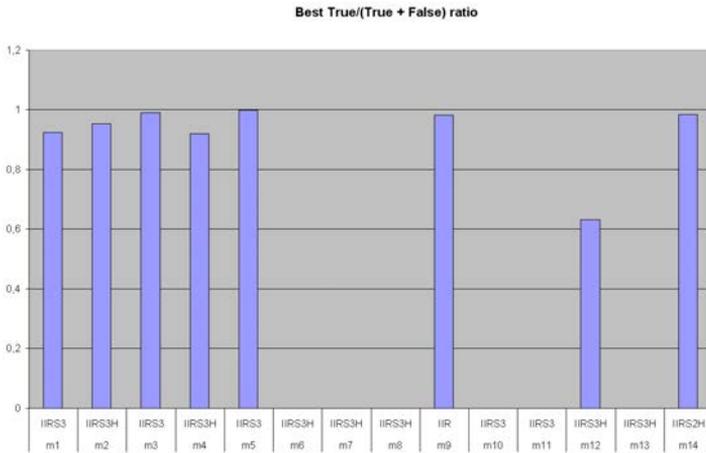


Figure 4.21. Best possible ratio comparing with target mask. In set m5 (ammunition box) only one sample of 1683 is false in the mine component. Set m12 (Russian AP mine) is hard to segment even if the optimal feature set is found. Set m6, m7, m8, m10 and m11 do not have a $True/(True + False)$ ratio since there are no mine data in these sets.

A scatter separability value was also calculated for each iteration and a comparison was made between higher and lower dimensions through a normalization function. Figure 4.22 displays the scatter separability value for the features that gave the highest separation for all the sets. The scatter separability value is not normalized in this graph, so a comparison between the sets when dimension differs is not meaningful.

The results displayed in Figure 4.23 show the highest possible $True/(True + False)$ ratio and the same ratio for features selected with the scatter separability criterion. In three of the sets the scatter separability criterion manages to select the features that give the best result and even in the other sets the features selected give a decent result.

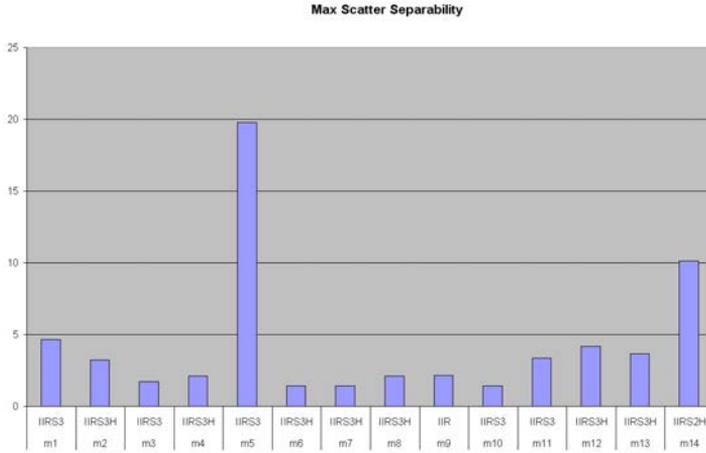


Figure 4.22. Maximum Scatter Separability value for all sets. The high SS value on set m5 (ammunition box) indicates a good separability for this set and as was seen in Figure 4.21 this is true.

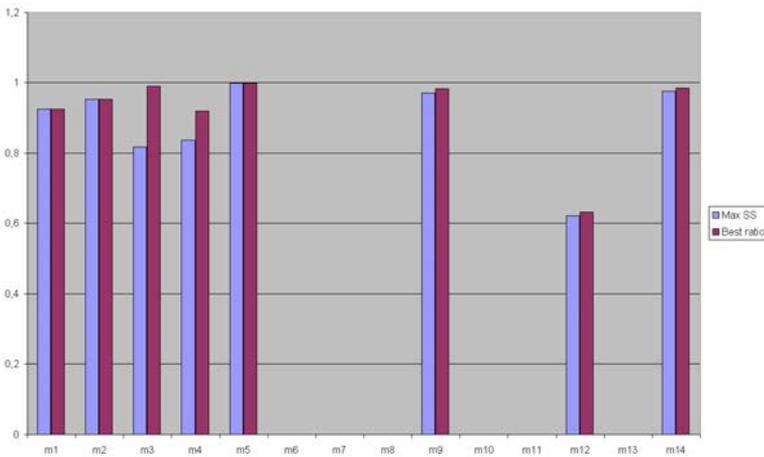


Figure 4.23. Best possible ratio compared with ratio for the feature set selected with the method (the set that gives the highest scatter separability value). Set m6, m7, m8, m10, m11 and m13 do not have a $True/(True + False)$ ratio since there are no mine data in these sets.

Chapter 5

Test on large area

The area of the data sets used to evaluate the method is approximately $0.4 \times 0.4m^2$, which is a small area considering the size of the whole scene. There are several reasons for examining a small neighborhood of the scene instead of the complete. Smaller sets decrease the computational burden in each run, which gives a shorter execution time. Looking at a large area we also have the risk of detecting too many interesting objects which might lead to multi mine components. If this actually is a problem depends on how the result is going to be used.

5.1 Detection of multiple objects in a large area

One test on a larger set was done, to see how the method performed when more than one mine was present. An image of this set is shown in Figure 5.1. The set contains two mines of different model and an ammunition box and corresponds to an area of about $2 \times 2m^2$.

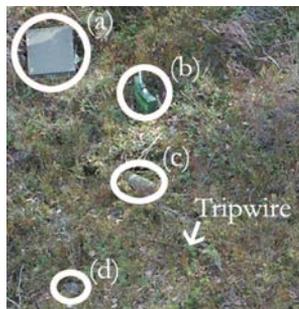


Figure 5.1. Picture of a set that covers a large area, several objects of interest are present in the set, all marked with a circle. There is also a tripwire from an PMR-2A mine going across the lower part of the picture. (a) Ammunition box (b) MRUD (Claymore) (c) Bursting shell (d) PMA-3.

The resulting components received with only intensity as input feature are

shown in Figure 5.2. In component (b) and (c) samples from each mine can be found, as mentioned earlier intensity is a good feature but in the right part of both components a large area of faulty samples can be seen.

Adding one feature at a time, evaluating the scatter separability criterion and the normalization between dimensions in each step, we end up with the complete feature set. The components received can be seen in Figure 5.3. Eight components are chosen, this is one effect of using a larger set, the variation in the data increase so more Gaussians are needed in the fitting process. The area with faulty samples in the 1D estimation is not present in the component containing most of the mine data (components 1) but the smallest mine and the bursting shell are not included at all. There could be several reasons for this but the noise factor is probably the major reason. The small mine is of the same size or smaller than the noise variation level which makes it impossible to find with the surfaceness criteria. Since the mine also is small the chance that the height feature would give a good result is very low. The resolution of the IR data is not as good as it could be, if we are lucky a small object will receive a different value from its surroundings but since many values are calculated by interpolation too much trust cannot be given to this either. So what we are left with in this case if the noise level cannot be lowered is the intensity feature. The reoccurring question that also has to be answered is “what is a good result?”. Is it better to be sure that all the mine samples are in one component, even though that might lead to many false samples? Alternatively, is it better to minimize the faulty samples, knowing this will reduce the number of true samples too?

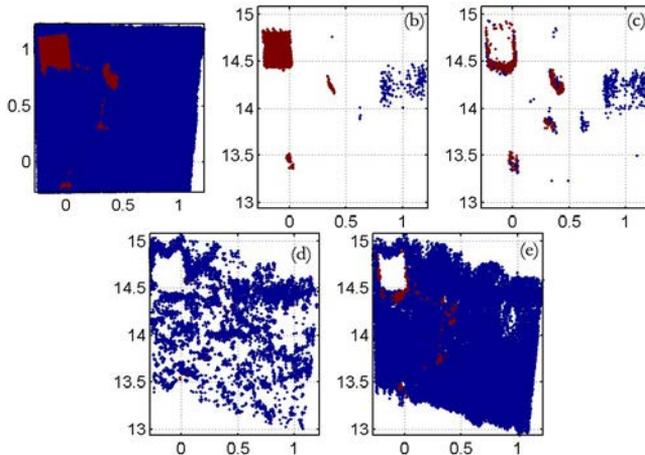


Figure 5.2. Components received from the method with only intensity as input feature. Most of the mine data is located in component (b) and (c) but an area with miss-labeled sample points is visible in the right region of both components.

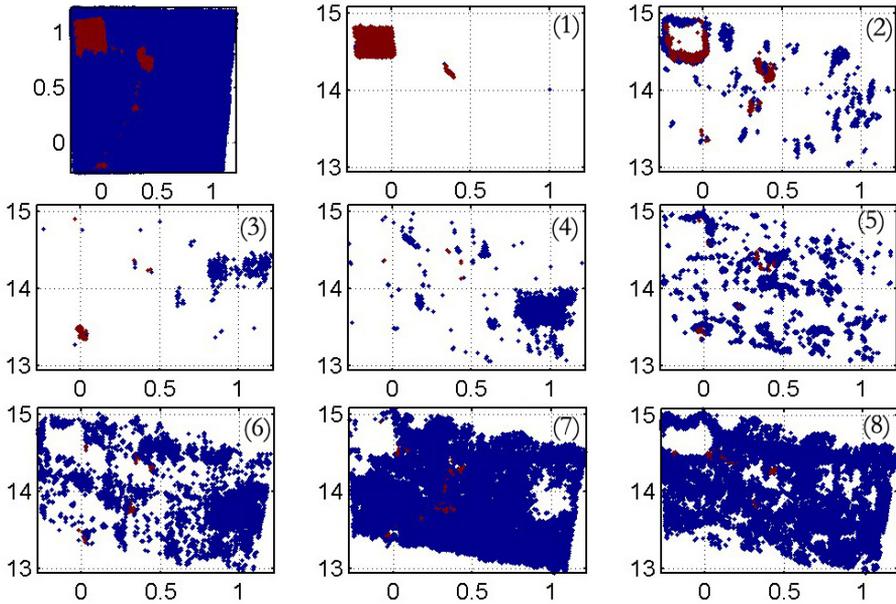


Figure 5.3. The scatter separability criterion in combination with the normalizing scheme selects all features as the best feature set. This image shows all components for the complete feature set. (1) contains samples from the large ammunition box and the medium sized Claymore mine but no trace of the small PMA-3 mine or the bursting shell. The count of false samples is very low with the price of a high miss rate as seen in Table 5.1. Close to a quarter of the true object samples are placed in component (2) but this component also contains many false samples. (3) and (4) contain some traces of the targets but also a lot of background area. (5)-(8) contains mostly background data.

Table 5.1. Statistics for each component in Figure 5.3 of the large area. Component 1 contains only 12 faulty samples and 2171 true samples.

Index	True	False	Miss
component 1	2171	12	1672
component 2	1034	1576	2809
component 3	74	355	3769
component 4	6	2285	3837
component 5	175	3855	3668
component 6	60	6999	3783
component 7	219	46545	3624
component 8	76	24091	3767

Chapter 6

Discussion

6.1 Summary

In this report, an evaluation has been performed of features received directly from the laser radar sensor, 3D positional features (x , y and height) and intensity. Different criteria for determining if the origin of 3D samples is from a surface of a manufactured object was also explored resulting in the surfaceness criterion. Anomaly images received from the infrared sensor was also associated to the laser radar data and used as an additional feature.

The underlying source of the data retrieved from the sensors was assumed to be a mixture of Gaussians. The expectation-maximization algorithm was used to estimate the parameters of the mixtures. The “true” number of mixtures is not known when the source consist of real data so a minimum message length criterion was utilized to determine which one of the expectation-maximization mixture parameter estimations to use as a representation of the original data. After the number of mixtures and their parameters have ben estimated, the samples can be labeled as belonging to one of these mixtures thus giving a simple segmentation of the set.

A scatter separability criterion was used to determine the priority and quality of the different features. The scatter separability criterion is designed in a way that a mixture with compact well-separated Gaussians receives a high value. A normalization of the scatter separability value was also used to determine if a higher dimension feature set improves the result. The method has been tested on real data collected at a field trial performed by FOI. For the sets that contained mine data or any other object of interest a target mask was created. This mask was then used to evaluate the resulting segmentation.

The method manages to select the features that give the best result for three of the real data sets and features that give a decent result in the remaining sets. The scatter separability value is generally low when there is no object of interest in the set. A larger set takes more time to process and smaller objects are hard to find, but indication is shown that there is a possibility to segment medium sized and large objects.

6.2 Conclusions

Anomaly detection with laser radar and IR sensor gives very interesting results. Objects in a complex environment can be segmented with a plausible result using common algorithms. Considering that the noise level in the laser radar data used sometimes is of the same size as the object in question, it is remarkable that the method does as well as it does in most of the cases. The surfaceness criteria is very sensitive to the noise level and a usage of a laser radar with higher precision will most likely improve the surfaceness detection for smaller objects.

6.3 Future directions

One thing that this report gives no clear answer too is the question: “What is a good result?” The score used to evaluate the components received from the method is based on how many true and false samples (according to the target mask) that are found in each component. Therefore selecting a good target mask becomes an important step in the evaluating processes. If this is a good way to measure component quality could certainly be argued. There is always the possibility that we miss more target samples than we find, but still have low number of mis-labeled samples. One way to determine if the result is good is to specify the application. An argument that might support maximizing the *true* samples is that there is an existing method developed at FOI that can calculate the size of an object from sample point. Of course many points are preferred, but it is also obvious that it is important that a majority of these points are on the actual object of which the size is to be estimated. Another scenario could be that we want a system that can detect that there are mines in a scene with no interest of location or maybe we want to know how many candidates there are, if any? Alternatively, just reduce anomalies for some other process. These scenarios might or might not require a re-evaluation of how to rate the feature quality and its result. Some time should be dedicated to specify the level the detection and its application.

The distribution that the data is collected from is assumed Gaussian, this means we are trying to describe the “real world” data with mixtures of Gaussians. There could be a distribution that represents the “real world” data better, but that is left for future work.

Characteristics of different features may vary with different conditions; the measured IR radiation varies with the temperature, which for example could lead to a different reading depending on if it is night or day. The measured intensity may also vary with different weather conditions, for example rain or high humidity. A preselection process of base features depending on the vegetation, weather, temperature or hour of the day would be interesting, but has not been handled in this report, but should be considered in future works.

No spatial feature information (x and y) was used in the clustering, as mentioned in the beginning of Chapter 3, these features were removed in an early stage of the feature selection process. Using the spatial information directly as features in the method did not give a satisfying result. There is a possibility of still using this information in some way, maybe in a pre-processing step or a post-processing

step. An example of a possible post-process could be an evaluation of how close in the spatial domain that the samples in a component are located and how compact they are and from this estimate the possibility of some samples being faulty or more general decide if it is even possible that the data actually represents a mine.

A new laser radar sensor system with higher precision, with a lower noise level, would certainly increase the chances of success with a surfaceness criteria, since the algorithm in this report was able to find large flat surfaces, but smaller surfaces disappeared in the noise.

More time has to be dedicated for finding a stable unsupervised way to estimate the quality of the features before usage and a possibility to exclude or include features based on this quality measurement. The scatter separability criterion used in this report was considered to be a candidate for such quality measurement, simply because if the separability is low it should be harder to segment correctly and if it is high it should be easier. The criterion gives high values if all the components are well separated, it is not taken into consideration if the component belongs to background or object. We are interested in how well we can separate the object component from the background components, so a modification of the scatter separability criterion might be a good alternative. This modification could be done in several ways. One possibility would be to go through each component treating it as the object component, alternatively identifying the object component in some other way, and comparing the separability from the union of the remaining components.

The choice of minimum message length as a criterion for choosing the number of Gaussians in the mixtures could certainly be argued, there are a variety of other methods that have shown reasonable results. The minimum message length criterion showed pleasing results on synthetic data and a visual inspection of the result from the real data came close to what was expected in most cases. In conclusion, the minimum message length does a decent job of selecting model order, but it is not guaranteed to be the best criterion in all conditions.

Bibliography

- [1] Jörgen Ahlberg and Ingemar Renhorn. Multi-and hyperspectral target and anomaly detection. Technical Report FOI-R-1526-SE, FOI, Sensor Technology, Linköping, Sweden, 2004.
- [2] Pierre Andersson and Gustav Tolt. Detection of vehicles in a forest environment using local surface flatness estimation in 3-D laser radar data. In *Proceedings SSBA 2007*, pages 65–72, March14–15, 2007.
- [3] Richard O. Duda and Peter E. Hart. *Pattern Classification, 2nd ed.* Wiley-Interscience, 2001. ISBN 0-471-05669-3.
- [4] Jennifer G. Dy and Carla E. Brodley. Feature selection for unsupervised learning. *Journal of Machine Learning Research*, 5:845–889, August 2004.
- [5] Mario A.T. Figueiredo and Anil K. Jain. Unsupervised learning of finite mixture models. *IEEE Transactions on pattern analysis and machine intelligence*, 24(3):381–396, 2002.
- [6] Christina Grönwall and Anna Linderhed. Statistical approaches to mine detection using optical sensors. In *Proceedings SSBA 2007*, pages 73–76, March14–15, 2007.
- [7] D. Letalick, G. Tolt, S. Sjökvist, S. Nyberg, C. Grönwall, P. Andersson, A. Linderhed, G. Forsell, H. Larsson, and M. Uppsäll. Multi-optical mine detection: results from a field trial. In *Proceedings SPIE*, volume 6217, 2006. 10.1117/12.665517.
- [8] Dietmar Letalick, Tomas Chevalier, Håkan Larsson, Claes Nelsson, Sten Nyberg, Stefan Sjökvist, Ove Steinvall, and Gustav Tolt. MOMS - analysis and evaluation of experimental data. Technical Report FOI-R-2012-SE, FOI, Sensor Technology, Linköping, Sweden, 2006.
- [9] J. Oliver, R. Baxter, and C. Wallace. Unsupervised learning using MML. *Proceedings 13th International Conference Machine Learning*, pages 364–372, 1996.
- [10] S. Sjökvist, S. Abrahamson, P. Andersson, G. Forsell, C. Grönwall, D. Letalick, A. Linderhed, D. Menning, I. Renhorn, M. Severin, T. Chevalier, H. Larsson,

- S. Nyberg, O. Steinvall, G. Tolt, and M. Uppsäll. MOMS multi optical mine detection system - initial report. Technical Report FOI-R-1721-SE, FOI, Sensor Technology, Linköping, Sweden, 2005.
- [11] James Trevelyan. Landmines - problems and solutions. *Asia-Pacific Magazine*, May 1998.

Appendix A

Description of the data sets

Table A.1. Names used for the sets in this rapport and the name of the mine in the specific set and a short description of the mine. A closer look at set m1, m12, m5 and m7 can be found in Chapter 4. Result plots from the remaining sets can be found in Appendix B.

Set	Mine	Description
m1	MRUD (Claymore)	The MRUD is convex rectangle shaped, plastic bodied, directional type anti-personnel (AP) mine designed to wound or kill by fragmentation.
m2	Improvised explosive device (IED)	An IED is a bomb constructed and deployed in ways other than in conventional military action. They may be partially comprised of conventional military explosives, such as an artillery round, attached to a detonating mechanism.
m3	PMA-3	The PMA-3 is a small, circular, plastic bodied AP mine which is designed to wound or kill by blast effect.
m4	PMR-2A	The PMR-2A is a cylindrical, cast iron bodied, stake mounted AP mine which is designed to wound or kill by fragmentation.
m5	Ammunition box	The ammunitions box is rectangle shaped. This could be an IED.
m6	No mine in this set	This set consist of grass and small bushes.
m7	No mine in this set	This set consist of grass, small twigs and branches.
m8	No mine in this set	This set consist of grass, twigs and an area with dirt and small stones.
m9	PMA-2	The PMA-3 is a small, circular, plastic bodied AP mine which is designed to wound or kill by blast effect.

Continued on next page

Set	Mine	Description
m10	No mine in this set	This set consist of a flat gravel road with small rocks.
m11	No mine in this set	This set consist of grass and a big stone.
m12	PMR-2A	The PMR-2A is a cylindrical, cast iron bodied, stake mounted AP mine which is designed to wound or kill by fragmentation.
m13	No mine in this set	This set consist of long grass and small bushes.
m14	AT2	The AT2 is a cylindrical, plastic bodied anti tank (AT) mine which is designed to damage or destroy vehicles by a penetrating effect.

Appendix B

Result plots (real data)

B.1 Set m1

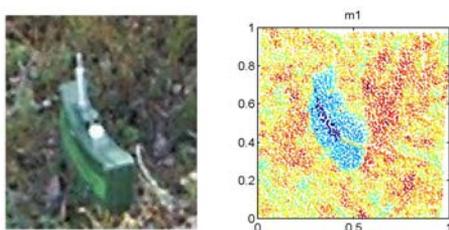


Figure B.1. Image of set m1, Claymore mine, and point cloud representation of the set.

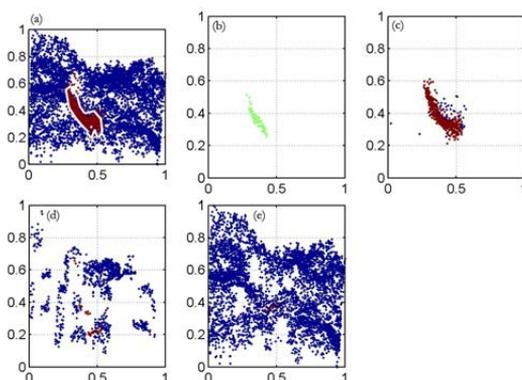


Figure B.2. Components received from intensity and IR features for set m1.

B.2 Set m2

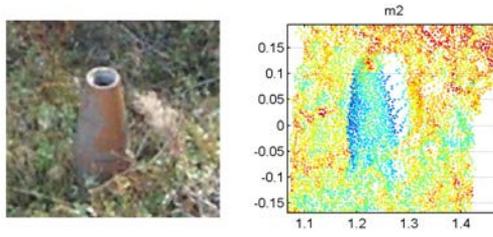


Figure B.3. Image of set m2 and point cloud representation.

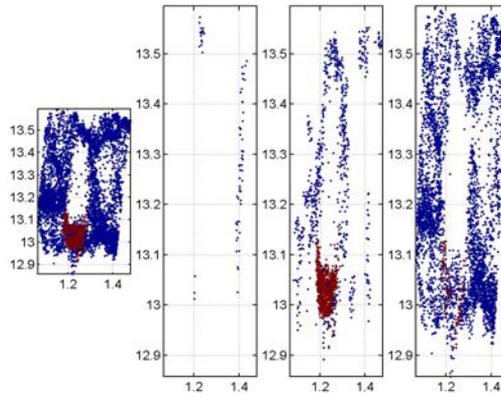


Figure B.4. Components received from intensity, IR, surface similarity and height features for set m2.

B.3 Set m3

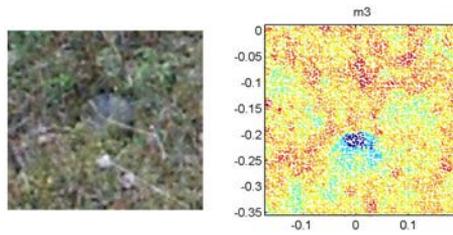


Figure B.5. Image of set m3 and point cloud representation.

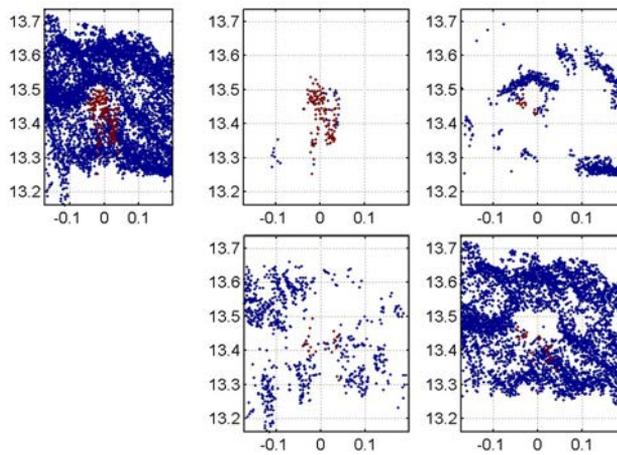


Figure B.6. Components received from intensity, IR, surface similarity and height features for set m3.

B.4 Set m4

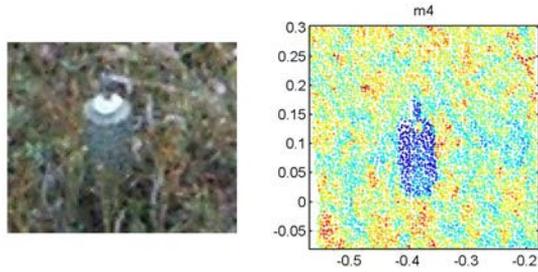


Figure B.7. Image of set m4 and point cloud representation.

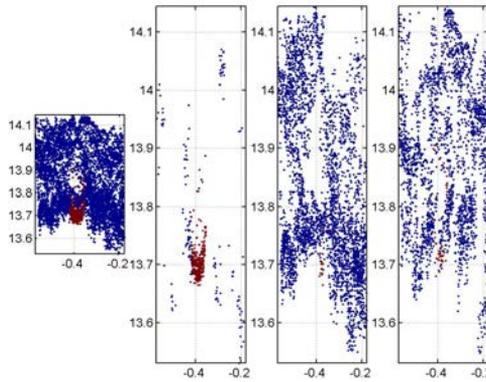


Figure B.8. Components retrieved from intensity, IR, surface similarity and height for set m4.

B.5 Set m5

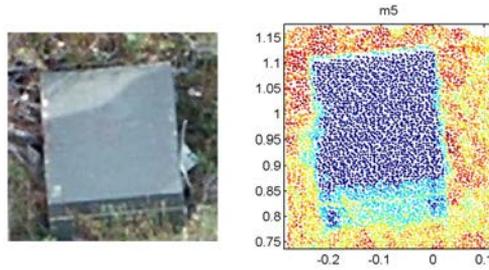


Figure B.9. Image of set m5 and point cloud representation.

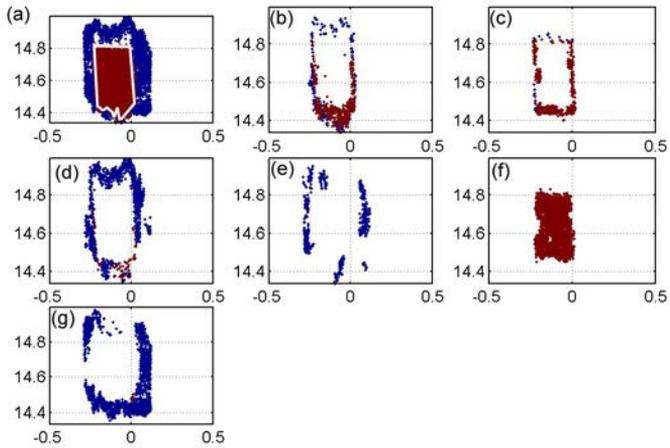


Figure B.10. Components received from intensity, IR, surface similarity and height features for set m5.

B.6 Set m6

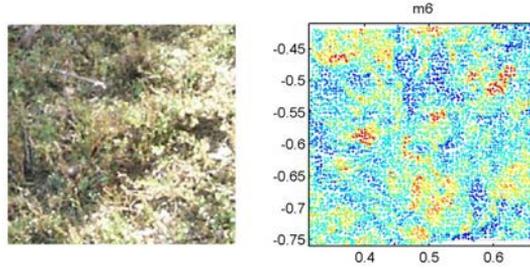


Figure B.11. Image of set m6 and point cloud representation.

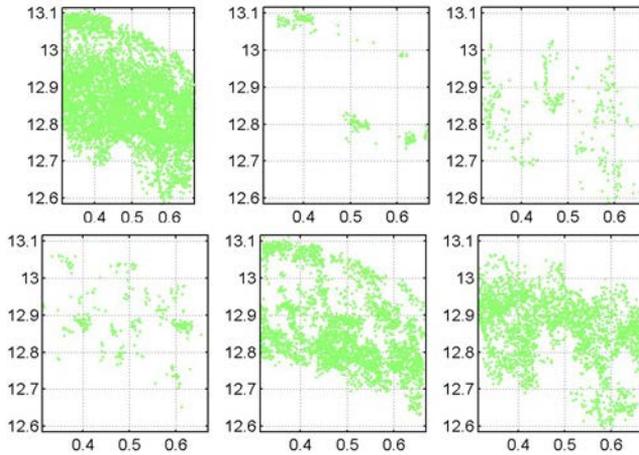


Figure B.12. Components recieved from intensity, IR, surface similarity and height features for set m6.

B.7 Set m7

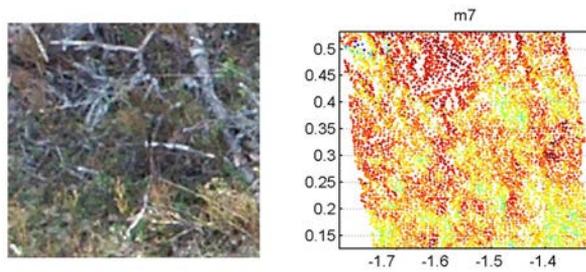


Figure B.13. Image of set m7 and point cloud representation.

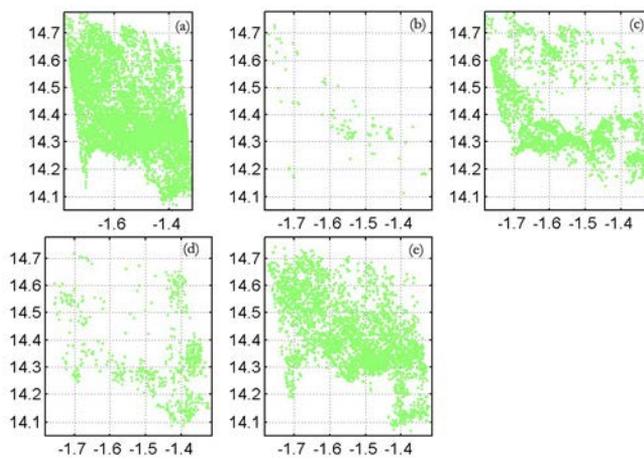


Figure B.14. Components retrieved from intensity, IR, surface similarity and height features for set m7.

B.8 Set m8

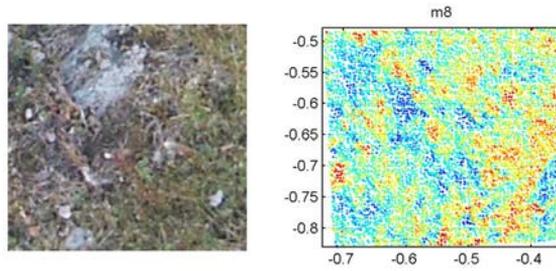


Figure B.15. Image of set m8 and point cloud representation.

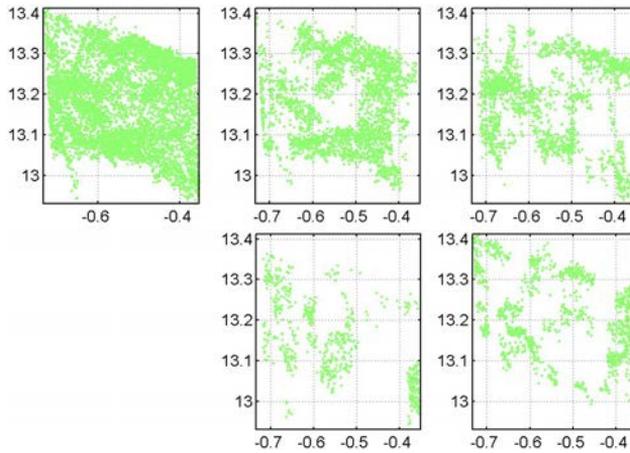


Figure B.16. Components received from intensity, IR, surface similarity and height features for set m8.

B.9 Set m9

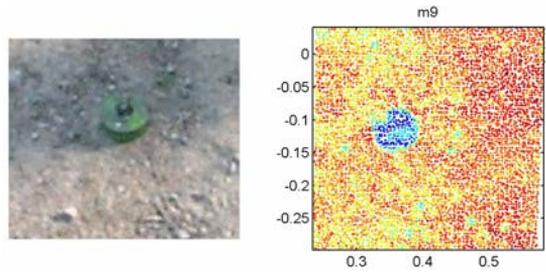


Figure B.17. Image of set m9 and point cloud representation.

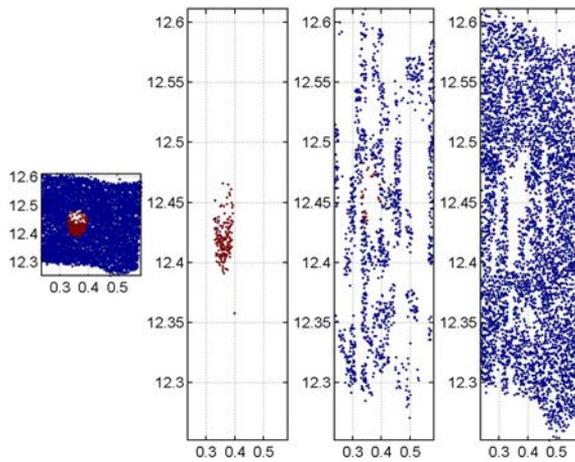


Figure B.18. Components received from intensity and IR features for set m9.

B.10 Set m10

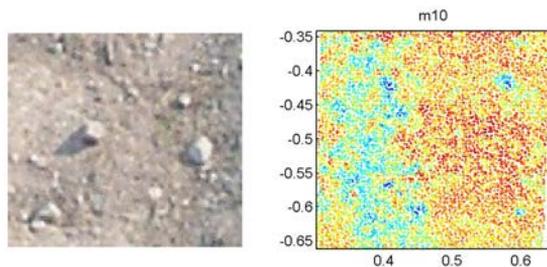


Figure B.19. Image of set m10 and point cloud representation.

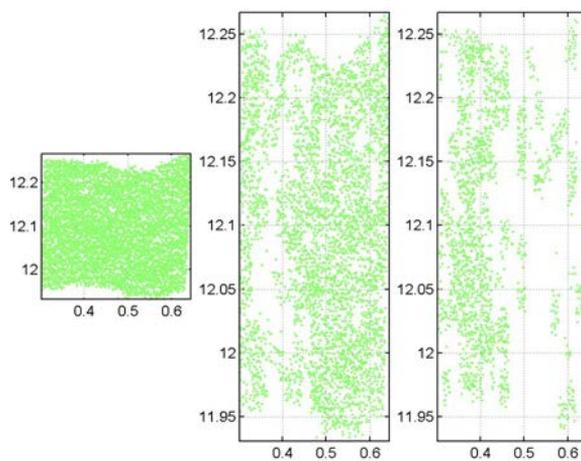


Figure B.20. Components received from intensity and IR features for m10.

B.11 Set m11

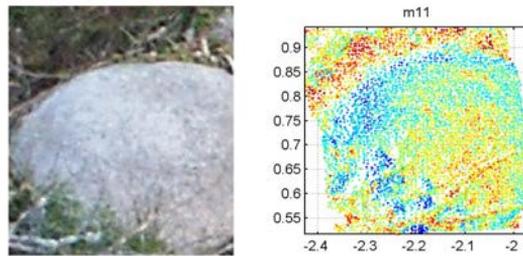


Figure B.21. Image of set m11 and point cloud representation.

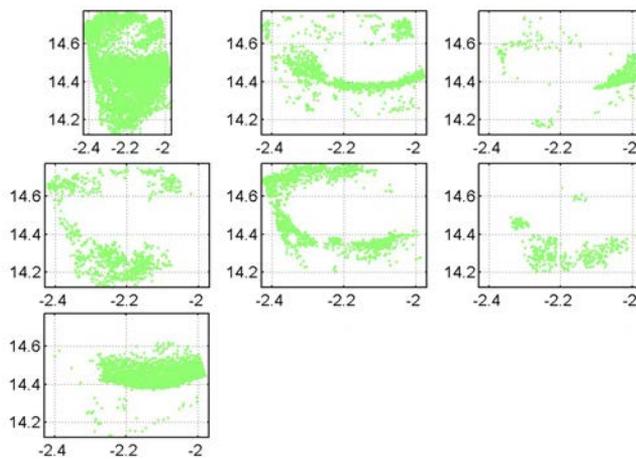


Figure B.22. Components retrieved from intensity, IR, surface similarity and height features for set m11.

B.12 Set m12

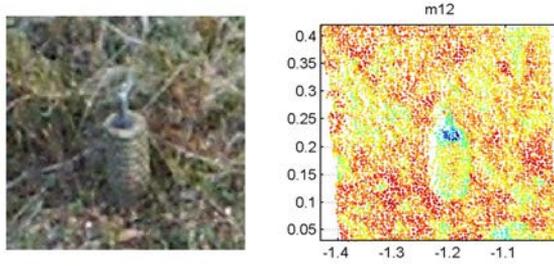


Figure B.23. Image of set m12 and point cloud representation.

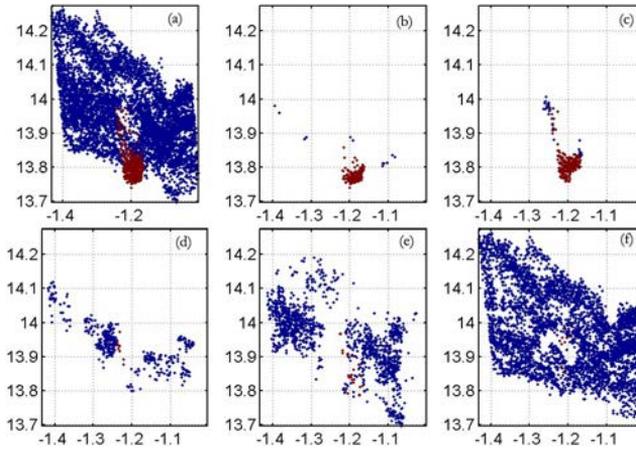


Figure B.24. Components recieved from intensity, IR, surface similarity and height features for set m12.

B.13 Set m13

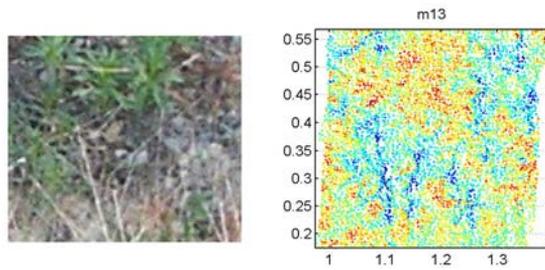


Figure B.25. Image of set m13 and point cloud representation.

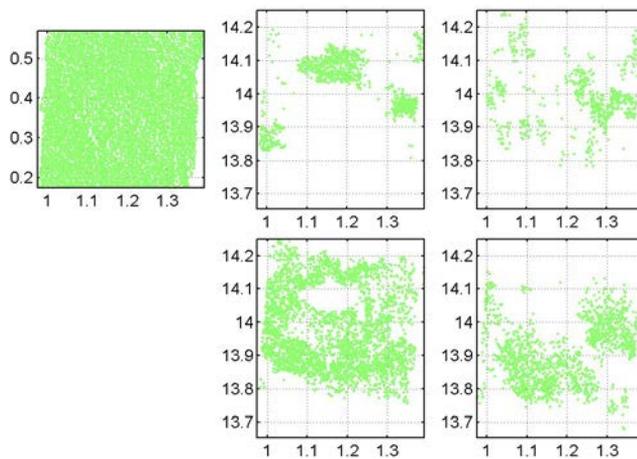


Figure B.26. Components received from intensity, IR, surface similarity and height features for set m13.

B.14 Set m14

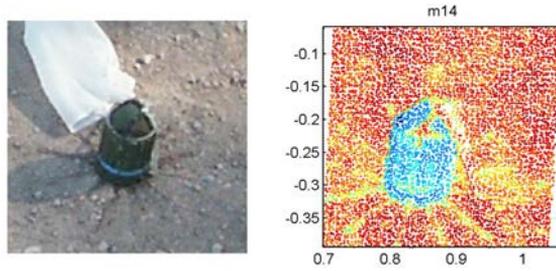


Figure B.27. Image of set m14 and point cloud representation.

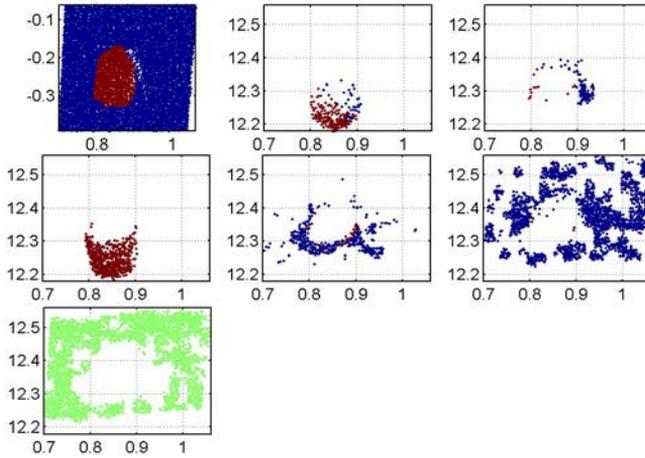


Figure B.28. Components received from intensity, IR, surface similarity and height features for set m14.

FOI, Swedish Defence Research Agency, is a mainly assignment-funded agency under the Ministry of Defence. The core activities are research, method and technology development, as well as studies conducted in the interests of Swedish defence and the safety and security of society. The organisation employs approximately 1000 personnel of whom about 800 are scientists. This makes FOI Sweden's largest research institute. FOI gives its customers access to leading-edge expertise in a large number of fields such as security policy studies, defence and security related analyses, the assessment of various types of threat, systems for control and management of crises, protection against and management of hazardous substances, IT security and the potential offered by new sensors.



FOI
Defence Research Agency
Sensor Systems
P.O.Box 1165
SE-581 11 Linköping

Phone: +46 13 37 80 00
Fax: +46 13 37 81 00

www.foi.se