

Comparison of vehicle-mounted forward-looking polarimetric infrared and downward-looking infrared sensors for landmine detection

Frank Cremer^{abc}, John G. M. Schavemaker^a, Wim de Jong^a and Klamer Schutte^a

^aTNO Physics and Electronics Laboratory, P.O. Box 96864, 2509 JG, The Hague, The Netherlands

^bPattern Recognition Group, Delft University of Technology, Delft, The Netherlands

^cSection of Applied Geophysics, Delft University of Technology, Delft, The Netherlands

ABSTRACT

This paper gives a comparison of two vehicle-mounted infrared systems for landmine detection. The first system is a downward looking standard infrared camera using processing methods developed within the EU project LOTUS. The second system is using a forward-looking polarimetric infrared camera. Feature-based classification is used for this system. With these systems data have been acquired simultaneously of different test lanes from a moving platform. The performance of each system is evaluated using a leave-one-out method. On the training set the polarimetric infrared system performs better especially for low false alarm rates. On the independent evaluation set the differences are much smaller. On the ferruginous soil test lane the down-ward looking system performs better at certain points whereas on the grass test lane the forward-looking system performs better at certain points.

Keywords: polarimetric infrared, infrared, landmine detection

1. INTRODUCTION

In the field of landmine detection the research on sensor systems focuses on three main topics. Firstly, the development of new sensors. Secondly the improvement of existing sensors and thirdly the integration of these sensors into a sensor-fusion system. This paper deals with the second topic; the enhancement of an existing sensor to contribute to the detection performance of a multi-sensor landmine system.

Within the European Union project LOTUS a multi-sensor vehicle-mounted platform was constructed.¹ This platform consists of a metal detector from Foerster (Germany), a ground penetrating radar from Emrad (UK) and an down-ward looking infrared (IR) camera. The processing algorithms for the infrared camera were developed by TNO.²

In a project funded by the Netherlands Ministry of Defence, a polarimetric IR sensor was developed.³ This forward-looking sensor provides information about the surface roughness of objects. Man-made objects, like landmines, are distinguished from natural background, since the surface of these objects is much smoother than the surface of the natural background. Therefore the IR polarimetric setup has the potential of an increased detection performance over a normal IR camera as used within the LOTUS project.

These two different systems consisting of the sensor, the application (forward looking or downward looking) and the processing are compared based on the detection performance. This detection performance is determined on a data set that was acquired simultaneously.

In this paper, first the measurement setup is discussed in Section 2, followed by a description of the measurements and pre-processing in Section 3. The detection and classification algorithms are discussed in Section 4, followed by training and evaluation procedures in Section 5. The detection results on the test lanes are discussed in Section 6. Finally, a discussion and conclusions about the results are given in Section 7.

Further author information:

F.C.: E-mail: FCremer@etro.vub.ac.be Phone: +32 (2) 629 2787 Fax: +32 (2) 629 2883

K.S.: E-mail: Schutte@fel.tno.nl Phone: +31 70 374 0469 Fax: +31 70 374 0654



Figure 1. The infrared and polarimetric infrared camera systems mounted together on a measurement platform.

2. MEASUREMENT SETUP

The measurement setup consists of two parts. The first is the infrared camera and acquisition system as used in the LOTUS project. The second part is the proprietary-designed polarimetric setup. These two systems have been mounted on a measurement platform, see Figure 1. In this figure, the polarimetric setup is mounted above the infrared camera. To the left of the infrared camera an additional Duncan Tech camera is mounted, but that camera falls outside the scope of this paper.

2.1. Infrared camera setup

The infrared camera used in this experiment is a 3-5 μm (MWIR) Radiance HS camera. This camera runs at a frame rate of 25 Hz. The integration time is set to 1 ms. The camera is mounted downward-looking on the measurement platform at a height of 1.6 m above the ground. On this camera a 13 mm lens is mounted, giving a field of view (FOV) of 36.4°. The camera is connected to an acquisition computer that grabs the image both digitally and using analog connection. The acquisition computer receives triggers from the laser distance meter for every 3 cm displacement. This trigger signal is divided by 14, so that an image is acquired every 0.42 m.

2.2. Polarimetric setup

The same polarimetric setup has been used in previous experiments.^{3,4} It consists of a wire grid polarisation filter that synchronously rotates with the frame sync of the camera. The camera is also a Radiance HS. The frame rate of the camera is set at 100 Hz and the integration time is set at 1 ms. During a full rotation 60 images are acquired. The camera is mounted on the measurement platform at a height of 2.0 m in a forward-looking direction. The orientation of the camera is 65° off-normal, i.e. looking 25° downwards.

The lens used in this experiment is a 50 mm lens. This lens gives a field of view of 9.1°. The field of view on the ground has a width between 0.65 m and 0.91 m. The depth of the field of view is 1.83 m and it starts at 3.53 m from the platform.

The polarimetric system is operated in a free-running mode in which image sequences are acquired continuously. The 3 cm trigger pulses from the laser distance meter are recorded along with the images. The exact position of each image in the sequence is estimated using interpolation of these trigger pulses.

3. MEASUREMENTS AND PRE-PROCESSING

The measurements were performed at the TNO-FEL test facility for landmine detection systems⁵ at 8 Augustus 2002. The measurement bridge was used as the measurement platform. The first measurement was performed on the sand test lane between 14:30 and 15:00. At that time the test lane was under direct sun light. Furthermore all test lanes (including a re-run of the sand test lane) have been measured between 16:30 and 18:45. At that interval the tube of the measurement bridge was blocking direct sun light onto the test lane while performing the measurements. Before the measurements the test lane was still in direct sun light. Except for a few thin clouds the sky was clear.

3.1. IR camera pre-processing

For the downward-looking infrared camera pre-processing methods are not necessary. This is because the image non-uniformity is already corrected by the camera and the image itself is a single shot image. The integration time was 1 ms and the velocity of the vehicle at maximum 0.2 m/s; so the movement during the integration time is 0.2 mm, which is a fraction of a pixel.

3.2. Polarimetric IR pre-processing

Calibration and motion correction is essential for polarimetric measurements. Each image is a summation of at least three components: part of the camera body reflected by the filter, emitted radiation from the filter itself and the scene image.³ The scene image is obtained by performing a two-point calibration procedure similar to a non-uniformity correction as described in an aforementioned paper.³

This calibration procedure would be sufficient for a static scene and a static setup, because pixel correspondence is guaranteed for every image in the sequence*. However, in the measurements described here the polarimetric setup is mounted on a moving platform and thus the pixel correspondence between the different images of a sequence does not exist. Every sequence consists of 60 images. These images are acquired with a frame-rate of 100 Hz, giving a time delay between the first and last image of 0.6 s. During this time delay the vehicle moves 0.12 m, which is obviously more than several pixels. Vibrations are another source of motion, since the polarimetric camera setup is directly mounted on the platform. Even though the platform is on a flat track (the measurement bridge), there always are some vibrations. Mounting the polarimetric setup on a stabilised platform, decreases vibrations, but never eliminate them fully.

Finally there is a third source of motion, or apparent motion to be precise. The filter in front of the camera bends the light rays and the direction depends on the orientation of the filter. As a consequence the 60 images in an image sequence appear to be moving with respect to each other. This apparent motion will be corrected for in our pre-processing method, without an extra step.

To overcome the problem of motion in the image sequence, we have devised a pre-processing method.⁶ Basically, there are two types of motion: one is motion in world coordinates (forward motion) and the other is in image coordinates (vibration and bending of rays by the filter). These two types of motion need to be addressed separately, since in the forward-looking measurement method the two coordinate systems are not the same, as they (almost) would be with a downward-looking setup. In Fig. 2 an overview of the pre-processing steps is given.

The first step of the pre-processing is the calibration of the images. For each angle a cold and a warm calibration is performed. These two calibration sequences give the gain and offset for every pixel and every angle. The offset correction removes the reflected image, the emission from the filter and the offset of each pixel. The gain factor removes differences in transmission of the filter and sensitivity of each pixel.

The second step performs the transformation of the image into world coordinates. This is necessary since the forward motion of the platform causes a shift between the images in a sequence in world coordinates and not in camera coordinates. This anti-perspective transformation uses the geometry of the setup (height and orientation of the polarimetric camera) as input to generate a plan view of the image.

The shift between the images (in world coordinates) in each sequence is corrected using a simple translation in the third step. The amount of translation is determined from the laser distance meter. This meter gives a position for every 3 cm of movement. The exact position of every image in the sequence is obtained using interpolation of the 3 cm samples of the

*This only under the circumstance that the filter does not bend the incoming rays in a way that depends on the orientation of the filter.

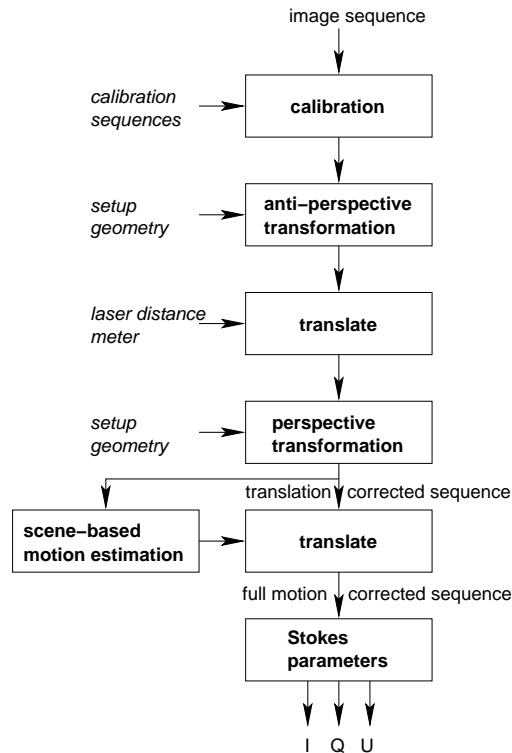


Figure 2. The pre-processing steps involved to transform an uncalibrated polarimetric image sequence into three calibrated full motion corrected Stokes images.

laser distance meter. A perspective transformation with the same geometric parameters is used to transform the translated image back into camera coordinates; this is the fourth step.

With the forward motion of the sequence corrected for, only the motion in camera coordinates remains. The vibrations and motion induced by the filter both reside in the camera coordinates and thus cannot be separated. A proprietary scene-based motion estimation is used to estimate the motion with respect to the first frame. The obtained translation vector is used to translate the image. The scene-based motion estimation and the translation is the fifth step of the pre-processing method.

The final step is to calculate the three independent Stokes images I , Q and U from the full motion corrected and calibrated sequence. The calculation involves a weighted integration over the sequence.³

3.3. Mosaicing

The pre-processing methods for the IR camera and the polarimetric IR setup both give a set of (polarimetric) images. These individual images will be input for the detection algorithm, as discussed in Section 4. For representation to the user it is possible to build a so-called mosaic of these image sets. The individual images are each placed in the mosaic image from the position based on the distance indicated by the laser distance meter and by using scene-based motion estimation. In Figure 3, the mosaics for the clay test lane are shown.

4. DETECTION ALGORITHMS AND CLASSIFICATION METHODS

For the conventional IR data and the polarimetric IR data different detection and classification approaches are followed. The conventional IR data is processed using the earlier described LOTUS IR processing.^{2, 7} This algorithm performs a normalisation of the image (to remove global variations) and performs some additional morphology operations to select objects that are conforming to size and some shape criterions. Alternative methods have been tested on a similar data set.⁸

For the polarimetric IR the algorithm is split into two phases. First there is a object detection step in which possible landmine objects are identified and secondly these objects are classified using two different algorithms: Naive Bayes

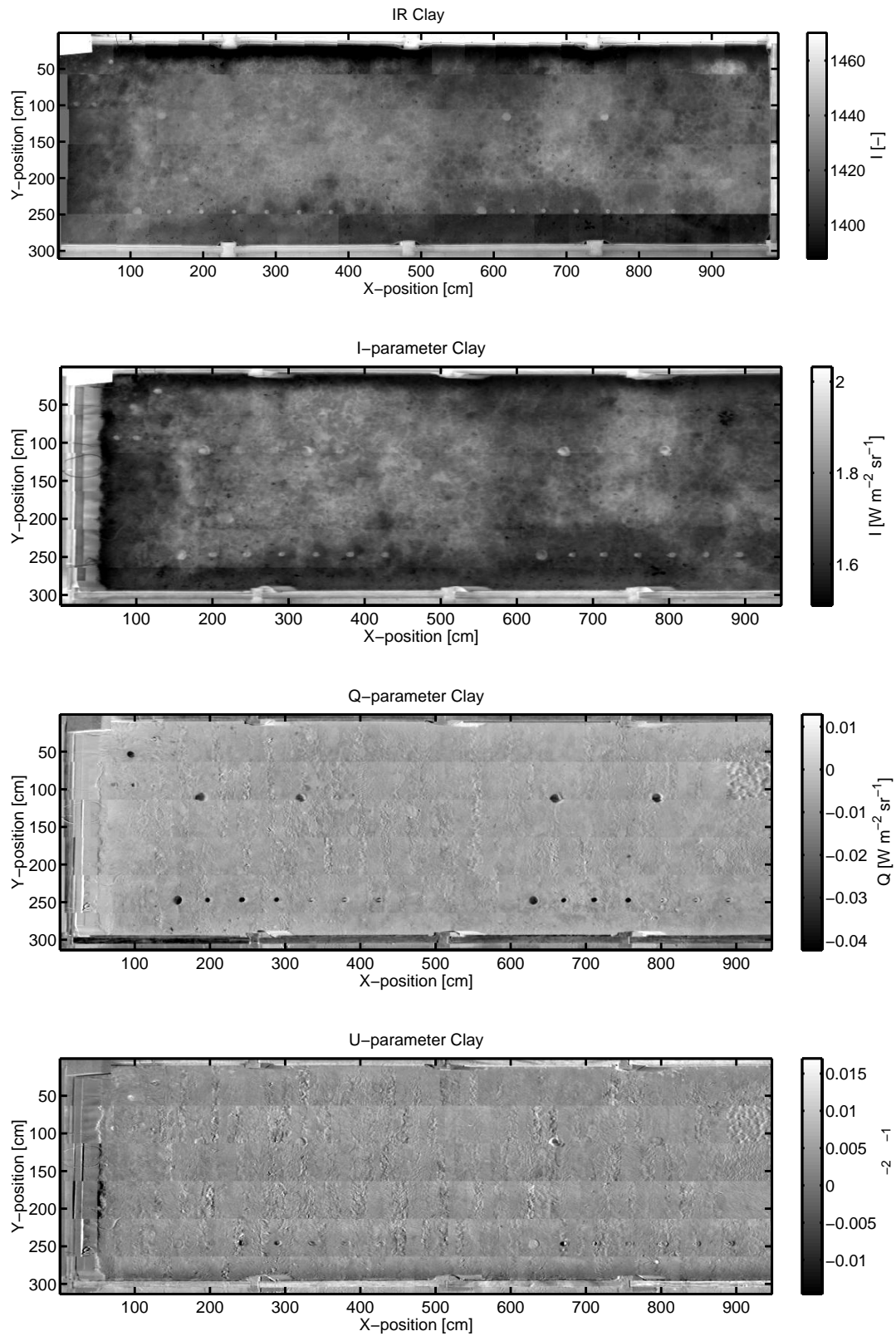


Figure 3. The mosaic of the clay test lane with the IR camera and the polarimetric IR setup.

and LVQ-dist. The object detection step is described in Section 4.1. The two different object classification methods are discussed in Sections 4.2 and 4.3.

4.1. Object detection and feature extraction

The polarimetric IR pre-processing results into 3 independent Stokes images.⁶ One of the 3 Stokes images, the intensity image I , is input for the region selection algorithm. This region selection algorithm selects regions using a tophat filter. For the selected regions the features are measured in the feature extraction process. The following features of each object are measured and are available for feature-level sensor fusion:

- mean values of I , Q and U ,
- contrasts between the object and the background (about half a square meter around the object) for the values of I , Q , U ,
- the area of the object,
- the length of the major axis,
- the fraction of the area and the convex area of the object,
- the fraction of the minor axis and major axis.

4.2. Naive Bayes

One of the parametric classifiers is the so-called plug-in rules classifier. This classifier makes an assumption about the shape of the class distributions and subsequently estimates the parameters based on the training set. For the sake of simplicity, we propose to use the following assumption: the class distribution is Gaussian (with non-equal variance) and independent for each feature. The parameters for this classifier are the mean and standard deviation for each parameter and each class.

For a two class problem, the likelihood ratio $\Lambda(\bar{x})$ for observation \bar{x} is given by:

$$\Lambda(\bar{x}) = \frac{p(\bar{x}|H_1)}{p(\bar{x}|H_0)}, \quad (1)$$

with $p(\bar{x}|H_0)$ and $p(\bar{x}|H_1)$ the probability densities of \bar{x} given the class H_0 and H_1 respectively. When a threshold is applied to the likelihood ratio, an optimal Bayes decision is made (i.e. lowest linear cost), where the threshold depends on the *a priori* probabilities and cost matrix.

It is assumed that the features are independent. This assumption may not be true for all cases, so the classifier is naive, hence its name Naive Bayes classifier. Under this assumption of independent features, the likelihood ratio can be rewritten as:

$$\Lambda(\bar{x}) = \frac{\prod_{i=1}^{N_f} p(x_i|H_1)}{\prod_{i=1}^{N_f} p(x_i|H_0)}, \quad (2)$$

with N_f the number of features. The class distribution $p(x_i|H_j)$ per feature for class j is assumed to be Gaussian with a mean m_{ji} and a variance σ_{ji}^2 . Substituting the Gaussian probability density function in Eq. 2 gives:

$$\Lambda(\bar{x}) = \frac{\prod_{i=1}^{N_f} \frac{1}{\sqrt{\sigma_{1i}^2}} e^{-\frac{(x_i - m_{1i})^2}{\sigma_{1i}^2}}}{\prod_{i=1}^{N_f} \frac{1}{\sqrt{\sigma_{0i}^2}} e^{-\frac{(x_i - m_{0i})^2}{\sigma_{0i}^2}}} \quad (3)$$

The normalisation term in front of the exponent does not make a difference for the decision, since it is independent of the feature value. Furthermore taking the logarithm of the likelihood does not make a difference either, since the logarithm is a monotonous increasing function[†]. These steps simplify the likelihood ratio into:

$$\Lambda'(\bar{x}) = \sum_{i=1}^{N_f} \left(\frac{(x_i - m_{0i})^2}{\sigma_{0i}^2} - \frac{(x_i - m_{1i})^2}{\sigma_{1i}^2} \right) \quad (4)$$

[†]If $x > y$ then $\log(x) > \log(y)$, $\forall x > 0, \forall y > 0$.

The mean and standard deviations are calculated using:

$$m_{ji} = \frac{1}{N_{sj}} \sum_{i=1}^{N_{sj}} x_i, \quad \forall x_i \in H_j \quad (5)$$

$$\sigma_{ji}^2 = \frac{1}{N_{sj} - 1} \sum_{i=1}^{N_{sj}} (x_i - m_{ji})^2, \quad \forall x_i \in H_j, \quad (6)$$

with N_{sj} the number of samples in class j . Together with the final threshold, the means and standard deviations are the parameters of the rules plug in method.

4.3. LVQ-dist

One of the non-parametric classifiers is the K-Nearest neighbour classifier.⁹ The operation of this classifier is rather simple:

- The 'training' parameters consist of the training data set itself.
- An unknown vector is classified based on some distance measure between this vector and all vectors in the training set. The k nearest neighbours are selected, with k being an odd number. The majority of vectors from one class determine the appointed class.

There are two disadvantages of this method. First, there is no final threshold, making it difficult to select the right operating point (false alarm rate and detection rate) for the classifier. Second, the number of training samples may be large and therefore evaluation may be calculation intensive.

The distance measure provides information about how close a sample is to either one class or the other. For the Naive Bayes classifier, the likelihood ratio is also dependent on the weighted distance difference. Now assuming all variances to be equal (i.e. $\sigma_{0i} = \sigma_{1i} = \sigma, \forall i \in N_f$), then Equation 4 can be used as classification measure, with \bar{m}_0 and \bar{m}_1 replaced by the two nearest neighbours from classes 0 and 1:

$$\Lambda(\bar{x}) = |\bar{x} - \bar{x}_{0w}| - |\bar{x} - \bar{x}_{1w}|, \quad (7)$$

with \bar{x}_{0w} and \bar{x}_{1w} the k nearest neighbours of \bar{x} of class 0 and 1 respectively, and $|\bar{x}|$ the distance measure of vector \bar{x} .

The first nearest neighbour may not be so suitable for measuring the distance, as the generalisation may not be right. The distance to the first nearest neighbour is highly dependent on chance.

The large number of vectors in the training set can be reduced in an number of different ways. The method we chose is to use a method similar to LVQ.¹⁰ A fixed number of vectors from the training set is chosen as the initial value of the set. Then all vectors from both classes are applied in a random order. If the nearest vector in the set is of the same class of the vector under consideration, then this vector is shifted towards the applied vector, otherwise it is shifted away from it.

The combination of using LVQ to represent the data sets and the distance as calculated in Equation 7 gives a new type of classifier that we denote as LVQ-dist. This classifier does not require independent features.

5. TRAINING AND EVALUATION

To compare the detection performance of the two sensors and the different detection methods to each other some form of ranking is necessary in which one detection result is acknowledged as a better result than another detection result. Without this performance measure it is impossible to select the best method. For a fair and realistic comparison only results on the same independent evaluation sets are considered. This means that the measured data set has to be divided into a training set for tuning of parameters and an evaluation set where these parameters are tested on.

There are two primary evaluation criteria. These two primary criteria are performance criteria. One criterion is related to the number of detected landmines and the other is related to the area(s) that have incorrectly been marked as landmines. The number of detected landmines (or percentage thereof) is a well-defined criterion. However the other criterion, the effectiveness given the detection rate, can be expressed in different ways: as false alarm area, as number of false alarms

or something in between: SCOOP.¹¹ SCOOP poses a limit on the size of each false alarm and is therefore more suitable for determining the effectiveness of the detection algorithm. The idea behind SCOOP is that for each false alarm the same action has to be taken: examination of the area by prodding and/or digging; taking up about the same time for each false alarm. The latter is only true if the false alarm sizes are more or less the same. Since processing of SCOOP is quite time consuming, a different approach is chosen here that leads to comparable results. For each detection a limit is placed on the size of the object, comparable to the maximum size of the landmines. This size limit is thought to be equal to the SCOOP size. Subsequently, evaluation is simply counting the number of detected objects (since their size is already limited, they do not need to be split up).

The selected operating point, i.e. combination of false-alarm rate and detection rate, for the detection system depends on two factors. These factors are: the operational scenario and the use of this system in a sensor-fusion system. Therefore the operation point cannot be determined beforehand. That means that the performance evaluation should be made on a range of detection and false-alarm rates, the so-called Receiver Operator Characteristics (ROC) curve. By comparing these ROC curves for the various methods the overall performance is considered.

In case of equal performance, secondary criteria may be used to select the detection algorithm. These secondary criteria are the (computational) complexity of the detection algorithm and the difficulties involved for obtaining the parameters on the train set.

For dividing the data set for each test lane into a training and evaluation set a leave-one-out method¹² has been used. In a leave-one-out evaluation method the classifier is trained on all but one sample and tested on the remaining sample. This process is repeated until all samples have been part of the evaluation set. The training set results are averaged over all possible training sets and the evaluation set results are summed up over all evaluation sets.

The sample size we used is not a single object, but one landmine and its surrounding area.⁶ Furthermore leave-one-out by itself does not have a way to generate ROC curves. A modified approach that uses a range of cost functions solves this problem.⁶

6. RESULTS

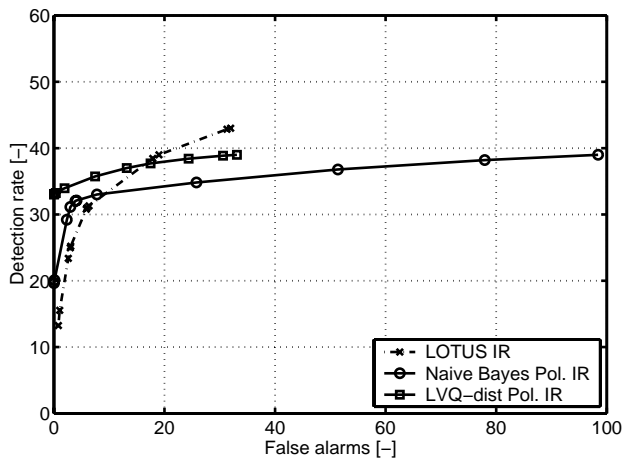
The training and evaluation set results are presented in Figure 4. It is clear that on the training set both object-based classification methods using the polarimetric IR data performs better than the LOTUS IR processing using the downward-looking IR camera data. From the object-based classification methods the LVQ-dist has the highest performance on the training set.

However, on the evaluation set the results are different. On the ferruginous soil the LOTUS IR performs better on most points than the polarimetric IR. On the grass test lane the LOTUS IR performs on most points less than the polarimetric IR.

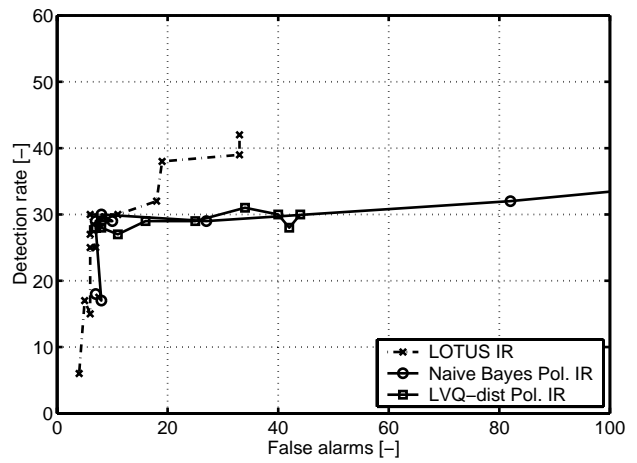
The large decrease in performance can be explained by the number of parameters. The object-based classification methods have several parameters to estimate the conditional probability functions. In the evaluation of the LOTUS IR processing two thresholds are used. Methods with more parameters are more likely to be overtrained than methods with less as is shown by the results here.

7. CONCLUSIONS

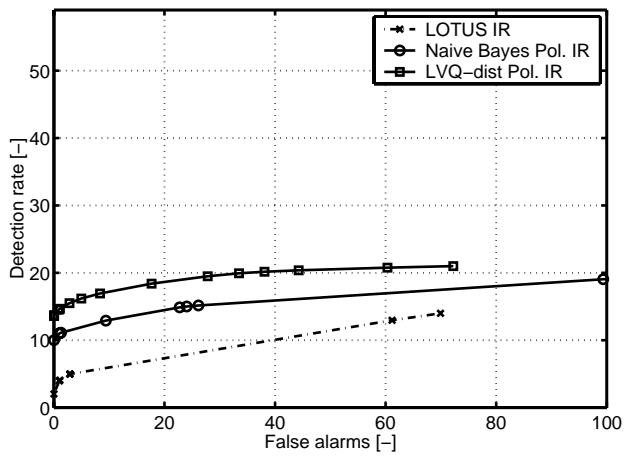
- Simultaneous measurements have been performed with polarimetric and conventional IR cameras on a moving platform.
- Pre-processing routines are shown for a polarimetric IR camera that employs a rotating polarisation filter.
- Object-based detection algorithms have been applied on the polarimetric IR data.
- Polarimetric IR performs better than conventional IR on the training set for test lanes with various soil/vegetation types.
- On an independent evaluation set the differences in performance between polarimetric and conventional IR are not significant. There are points where the polarimetric IR performs better than the LOTUS IR and vice versa.
- The object classification is sensitive to overtraining and procedures on how to provide more stable results should be investigated.



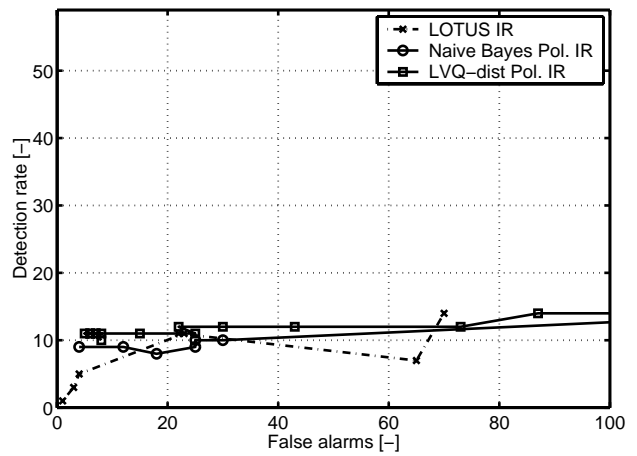
(a) Ferruginous (training set)



(b) Ferruginous (evaluation set)



(c) Grass (training set)



(d) Grass (evaluation set)

Figure 4. The training and evaluation results for the grass and ferruginous test lanes. The ferruginous soil test lane contains 60 landmines present the grass test lane contains 59 landmines.

ACKNOWLEDGEMENTS

The work presented in this paper is partly sponsored by the Netherlands Ministry of Defense.

This work is partly funded by the European Union as ESPRIT project LOTUS, number 29812.

REFERENCES

1. J. G. M. Schavemaker, E. den Breejen, K. W. Benoist, K. Schutte, P. Tettelaar, M. de Bijl, and P. J. Fritz, "LOTUS field demonstration of integrated multisensor mine-detection system in Bosnia," in *Proc. SPIE Vol. 5089, Detection and Remediation Technologies for Mines and Minelike Targets VIII*, R. S. Harmon, J. T. Broach, and J. H. Holloway, eds., (Orlando (FL), USA), Apr. 2003.
2. J. G. M. Schavemaker, F. Cremer, K. Schutte, and E. den Breejen, "Infrared processing and sensor fusion for anti-personnel land-mine detection," in *Proceedings of IEEE Student Branch Eindhoven: Symposium Imaging*, pp. 61–71, (Eindhoven, the Netherlands), May 2000.
3. F. Cremer, W. de Jong, and K. Schutte, "Infrared polarisation measurements and modelling applied to surface laid anti-personnel landmines," *Optical Engineering* **41**, pp. 1021–1032, May 2002.
4. F. Cremer, W. de Jong, K. Schutte, J. T. Johnson, and B. A. Baertlein, "Surface mine signature modeling for passive polarimetric IR," in *Proc. SPIE Vol. 4742, Detection and Remediation Technologies for Mines and Minelike Targets VII*, J. T. Broach, R. S. Harmon, and G. J. Dobeck, eds., pp. 51–62, (Orlando (FL), USA), Apr. 2002.
5. W. de Jong, H. A. Lensen, and Y. H. L. Janssen, "Sophisticated test facility to detect land mines," in *Proc. SPIE Vol. 3710, Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey and J. F. Harvey, eds., pp. 1409–1418, (Orlando (FL), USA), Apr. 1999.
6. F. Cremer, W. de Jong, and K. Schutte, "Processing of polarimetric infrared images for landmine detection," in *2nd International Workshop on Advanced Ground Penetrating Radar (IWAGPR)*, (Delft, The Netherlands), May 2003.
7. J. G. M. Schavemaker, E. den Breejen, F. Cremer, K. Schutte, and K. W. Benoist, "Depth fusion for anti-personnel landmine detection," in *Proc. SPIE Vol. 4394, Detection and Remediation Technologies for Mines and Minelike Targets VI*, A. C. Dubey, J. F. Harvey, J. T. Broach, and V. George, eds., pp. 1071–1081, (Orlando (FL), USA), Apr. 2001.
8. W. A. C. M. Messelink, K. Schutte, A. M. Vossepoel, F. Cremer, J. G. M. Schavemaker, and E. den Breejen, "Feature-based detection of landmines in infrared images," in *Proc. SPIE Vol. 4742, Detection and Remediation Technologies for Mines and Minelike Targets VII*, J. T. Broach, R. S. Harmon, and G. J. Dobeck, eds., pp. 108–119, (Orlando (FL), USA), Apr. 2002.
9. A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on pattern analysis and machine intelligence* **22**, pp. 4–37, Jan. 2000.
10. J. Laaksonen and E. Oja, "Classification with learning k-nearest neighbors," in *Proceedings of ICNN'96*, pp. 1480–1483, (Washington DC, USA), June 1996.
11. E. den Breejen, K. Schutte, and F. Cremer, "Sensor fusion for anti personnel landmine detection, a case study," in *Proc. SPIE Vol. 3710, Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey and J. F. Harvey, eds., pp. 1235–1245, (Orlando (FL), USA), Apr. 1999.
12. S. M. Weiss and C. A. Kulikowski, *Computer systems that learn*, Morgan Kaufmann Publishers, 1991.