Processing of polarimetric infrared images for landmine detection

Frank Cremer*†‡, Wim de Jong* and Klamer Schutte*

*TNO Physics and Electronics Laboratory,
The Hague, The Netherlands
Email: Cremer@fel.tno.nl

†Pattern Recognition Group, Delft University of Technology,
Delft, The Netherlands

‡Section of Applied Geophysics, Delft University of Technology,
Delft, The Netherlands

Abstract—Infrared (IR) cameras are often used in a vehicle-based multi-sensor platform for landmine detection. Additional to thermal contrasts, an IR polarimetric sensor also measures surface properties and therefore has the potential of increased detection performance. We have developed a polarimetric IR setup, which has to be used in a forward-looking manner.

This paper describes all the steps to reach detection performance. The first step is the acquisition of the polarimetric IR image data. The next step is the pre-processing to (re)construct polarimetric images. A subsequent segmentation step is made to identify objects. Features, like intensity, reflectivity and shape, of these objects are measured. For independent performance analyses, the data set is divided into a training and evaluation section. A classifier is trained on the training section and evaluated on the classification section. The detection performance of the forward-looking IR camera is shown in receiver operator characteristics (ROC) curves.

I. INTRODUCTION

The single-sensor detection performance of current sensors is not sufficient to meet the requirements for landmine detection. To reach a detection rate that is high enough, too many false positives (false alarms) are generated. For improved detection, commonly a sensor suit consisting of a ground penetrating radar, a metal detector and an infrared camera is used. Data from these sensors is fused using various sensor-fusion methods [1]. The aim of this paper is to show the processing algorithms for images measured with a polarimetric infrared camera. In images that result from this processing mines can be detected and features can be measured for fusion with other sensors. In another paper [2] sensor fusion of polarimetric IR data and data obtained by the Video Impulse Radar [3] is discussed.

A complete overview of the pre-processing, detection and classification algorithms is shown in Fig. 1. The first step is the acquisition of the raw images with the polarimetric infrared camera as discussed in Sec. II. The second step, the pre-processing, converts the raw images from the moving camera into three independent and calibrated Stokes [4] images, see Sec. III. In these images regions of interest are detected using tophat filtering, see Sec. IV. Features of these regions of interest in the Stokes images are measured and subsequently classified in Sec. V. How the performance of
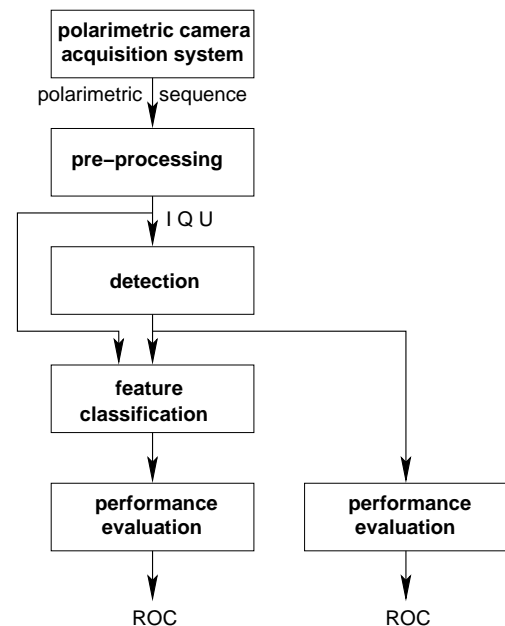

Fig. 1. Overview of the complete processing chain from measurement to performance evaluation for polarimetric images obtained from a moving platform.

both the detections and the classified features is evaluated is discussed in Sec. VI. The results of the detection classification experiments are shown in Sec. VII. Finally the conclusions are drawn in Sec. VIII.

II. POLARIMETRIC INFRARED MEASUREMENTS

Polarimetric measurements have been performed at the TNO-FEL test facility for landmine detection systems [5]. In this paper only results obtained at the sand lane are described. The reason for this choice is the fusion with GPR data [2] that was only available for this sand lane. The sand lane is a 10 m long box with a width of 3 m and a depth of 1.5 m that is filled with filtered sand. In this box different types of landmines have been placed at different depths.

The polarimetric infrared setup consists of a rotating wire grid polarisation filter in front of the lens of the IR camera.

Fig. 2. The infrared polarimetric measurement setup (with the TNO logo) as mounted on top of the measurement platform, that moves along the measurement bridge. Also visible are two downward looking cameras, that are mounted below the polarimetric setup.

The filter rotates synchronously with the frame sync of the camera [6], [7]. The frame rate of the camera is set at 100 Hz and the integration time is set at 1 ms. During a full rotation of the filter 60 images are acquired. The camera is mounted on the measurement platform at a height of 2.0 m in a forward-looking direction, see Fig. 2. This platform moves with a constant speed of about 0.2 m/s. 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 at the bottom and 0.91 m at the top. 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. Using a laser distance meter the distance travelled by the platform along the measurement bridge is measured and recorded. Every 3 cm the laser distance meter gives a trigger pulse. These trigger pulses are recorded along with the images. The exact position of each image in the sequence is estimated using interpolation of the trigger pulses.

III. 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 [6]. The scene image is obtained by performing a two-point calibration procedure similar to a non-uniformity correction as described in aforementioned paper.

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[1]. 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 will always be some vibrations. Mounting the polarimetric setup on a stabilised platform, will decrease the 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 [8]. 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. 3 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

---

[1]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.
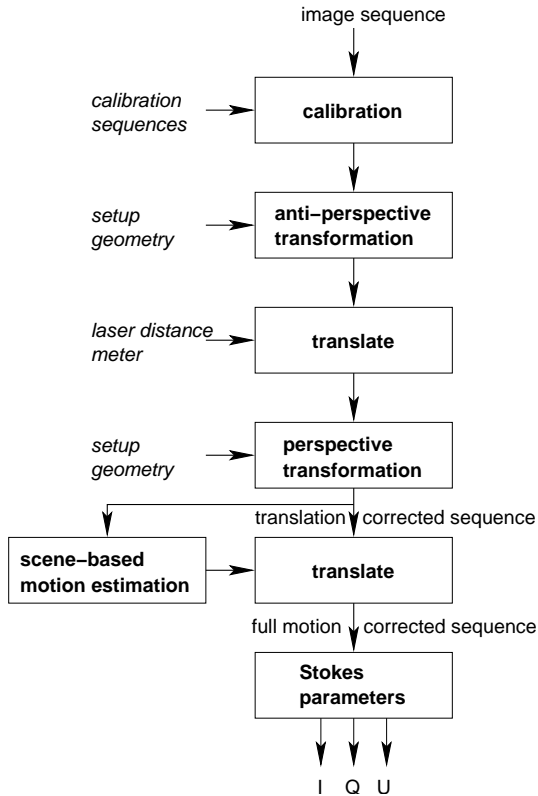
Fig. 3. The pre-processing steps involved to transform an uncalibrated polarimetric image sequence into three calibrated full motion corrected Stokes images.

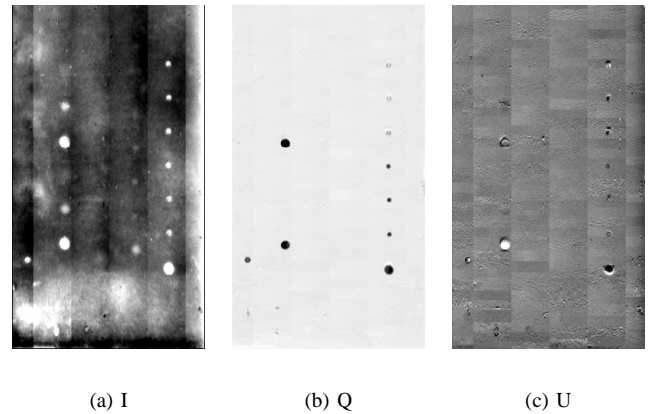(a) I   (b) Q   (c) U

Fig. 4. The mosaics of the three Stokes parameters for part of the sand test lane. The intensity scale is different for the three images. During the measurement the platform moved along the 'vertical' direction. Between the measurements the complete measurement bridge was moved in the 'horizontal' direction in 6 steps.

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 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 Stokes images I, Q and U from the full motion corrected and calibrated sequence. The calculation involves a weighted integration over the sequence [6].

For user representation, images of all positions have been used to generate a mosaic of the full test lane, see Figure 4. In the I-image of the mosaic all surface-laid landmines (the most right column of white spots in image (a)), most of the flush buried landmines (the second column counted from right) and

some of the deep-buried landmines are visible. All the surface-laid landmines clearly stand out in the Q-image (black spots). Finally in the U-image some of the sides of the surface-laid landmines are visible.

IV. DETECTION ALGORITHM

Tophat filtering has been used as a filtering step for detection [9]. Tophat filtering is a morphological transform that is able to detect local maxima in an image within the region of a structuring element. A disc-shaped structuring element is used. The size of this disc and the threshold are optimised using a random search method [10] for the full test lane. Two optimisation criteria are used.

The first optimisation criterion is the area under the Receiver Operator Characteristics (ROC) curve [11] between 0 and 50 false alarms. This criterion ensures optimal performance for moderate false alarm rates and finds mainly the surface and the flush-buried landmines.

The second criterion is the maximum number of detections for a maximum of 500 false alarms. All detections are input for the feature classification procedure. By bounding the detection to 500 false alarms the workload of the following classification procedure is limited. However, the classification procedure can only reduce the false alarm rate and not increase the detection rate, since landmines missed in the detection procedure are not evaluated in the feature classification procedure.

The detector works on individual image sequences instead of directly on the mosaic as presented in Fig. 4. Within the mosaic there are always boundary effects at places where two different images are combined. These boundaries adversely affect the detection and therefore the individual image sequences are used.

Multiple observations of the same object appear, since the data from the polarimetric camera is recorded at full frame-rate. Within one strip of measurements an object appears in around 10 image sequences. Detections of the same object

I Q U   detections

feature measurement

feature list

labeling ← ground truth

labeled features

feature selection   feature list

selected labeled features

classification   parameters

classified selected   labeled features

thresholding   TH
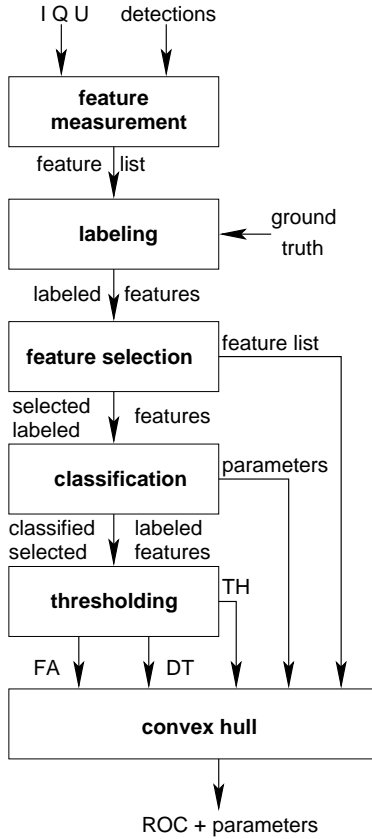
FA   DT

convex hull

ROC + parameters

Fig. 5. The forward processing approach for training the feature classifier.

from different image sequences are clustered. Aside from a threshold on the detector output (or the feature classifier) a threshold is also applied to the number of detections in one cluster.

V. FEATURE CLASSIFICATION

Using more relevant information in the classification process, improves the detection performance under the condition that the probability density function of this information is known [12]. One way to include this information is by means of features of regions of interest (ROI). Examples of features are size, shape, intensity. However, which features are used is discussed in Sec. VII. This section focuses on a general feature-based classification process.

An overview of the feature-based classification process is given in Fig. 5. The first step, once detections are made, is the measurement of features. This measurement of features is performed based on the results of the detection step and the original data.

After the features are measured in the original images, they are labelled according to the ground truth. This is necessary to train the classifier later on. Since *a priori* it is unknown which feature or which feature combination performs best, all feature combinations are tried for one by one. In the feature selection step only one such combination is forwarded to the classifier.

The task of the classifier is two-fold. Firstly the parameters of the classifier must be set based on the labelled features. In many cases this involves the estimation of the probability density function of both classes (landmine or false alarm). The second step is to classify all features, that means to add a class-confidence value to it.

By applying a threshold to the class-confidence value, the final decision is made about to which class each feature belongs. For each threshold the number of correctly detected landmines and the number of false alarms is determined.

Of all possible points in ROC space only the points that are on the convex hull of the ROC curve are of value [13]. Intermediate points are obtained by randomly switching classifiers between two neighbouring points. With each convex point of the feature list, all the classification parameters and thresholds are stored. This allows to select these working points for the evaluation set. For evaluation, the same process as in Figure 5 is used. However, the feature list, classification parameters and threshold are now inputs to the process instead of output.

For classification two different methods are evaluated. First, there is the Naive Bayes classifier, also known as Bayes plug-in classifier [12]. This classifier assumes independent features with Gaussian probability density functions. The equivalent likelihood ratio $\Lambda'(\bar{x})$ for the unknown feature vector $\bar{x}$ is given by:

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

with N_f the number of features and m_{0i} and m_{1i}, σ_{0i} and σ_{1i} the mean value and the standard deviations for feature i and classes 0 or 1.

The second classifier, is a novel classifier that we call 'LVQ-dist' classifier. This classifier uses distances in the first nearest neighbour and a learning vector quantisation (LVQ) [14] algorithm to reduce the training data in both classes. After training the equivalent likelihood ratio is given by:

$$\Lambda'(\bar{x}) = ||\bar{x} - \bar{x}_{0w}|| - ||\bar{x} - \bar{x}_{1w}||, \quad (2)$$

with $\bar{x}_{0w}$ and $\bar{x}_{1w}$ the closest vector in class 0 and 1 of the training set respectively.

The major difference between the two classifiers is the way the probability distribution function (PDF) is modelled. The Naive Bayes assumes independence of features and a Gaussian shaped PDF. When features are not independent the solution provided by Naive Bayes is suboptimal. The LVQ-dist classifier can model arbitrary PDFs to a certain level and therefore be more specific. It is expected that this method provides good results on the training set, because of this advantage.However, this method may perform worse on an independent evaluation set, since it is too specifically tuned on the training set.

VI. PERFORMANCE EVALUATION

In previous work the evaluation of the detection results was performed using the SCOOP algorithm [15]. This algorithm
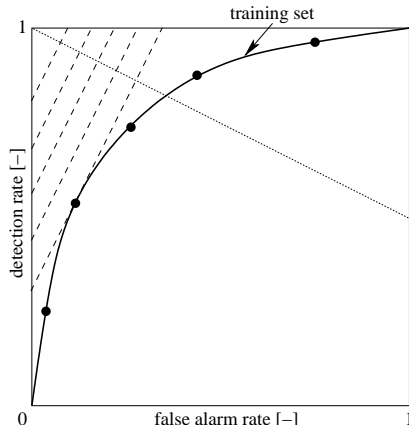
Fig. 6. Selecting one cost function in the ROC of the training set.

weighs the number of false alarms against the size. The rational for this is that it takes more time to investigate a false alarm with a large area than one with a small area.

By limiting the number of false alarms in the detection process and giving out only point detections, sizes are not of influence. Each detection is of limited size. To compare results presented here, each false alarm can be seen as one SCOOP false alarm.

The data set is limited with 60 landmines and around 20 m^2. Although, this data set is larger than in previous sensor-fusion experiments [1], leave-one-out [16] has been used to evaluate the performance on an independent data set. The advantage of leave-one-out is that as much training data is used as possible.

The disadvantage of leave-one-out is that optimisation of a complete ROC curve is not straight forward. One approach is to remove landmines in a specific order, so called populations [1]. Only the 100% detection point on the training set was selected and step by step the training set was reduced. Since there are different orders in which landmines can be removed from the training set, the evaluation process is not simplified.

In this paper we introduce a novel leave-one-out optimisation procedure that is more intuitive. Instead of selecting only the 100% detection point of the training set, more points are used. Each ROC point on the training set is optimal for a specific range of linear cost functions. Since this range of cost functions may vary from one training set to the other, the procedure cannot start with the ROC points itself. In these linear cost functions the cost of a missed landmine is weighted against the cost of a false alarm. Instead of starting from ROC points, several linear cost functions are defined that span the ROC curve. Now for each cost function and each training set result, the optimal working point is selected, see Fig. 6. This point is used to evaluate the performance on the evaluation set. The evaluation set consists of the landmine plus area around it that is left out of the training set.

The results of the different training and evaluation sets can be added up together for each specific cost function. The rational behind this is that the same optimal working point

with respect to this cost function is selected. This approach is similar to the situation where the ROC is generated using a single threshold. Results from training and evaluation sets using the same threshold may also be added up together. In essence with the cost function a virtual threshold is constructed.

VII. RESULTS

The following features of each object are measured and made available for classification:

- mean value of I, Q and U,
- contrast between I, Q, U and the background,
- 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

The first six features give information about the intensities in the three polarimetric images. The last two features give information about the shape of the object. The second last feature gives a measure for how regular shaped the object is (the fraction approaches unity, for objects like ellipses and circles). The last feature gives a measure for how square the object is. A perfect circular object has a unity value for both features.

In Fig. 7, the results of both feature classifiers and the tophat detectors are shown for both the training and evaluation set. The maximum number of detected landmines is 46. The area as covered by the polarimetric IR measurements consists of 20.4 m^2.

The best classifier on the training set is the LVQ-dist classifier. The second best classifier is the Naive Bayes classifier. This means that including features (both polarimetric and normal features) helps to improve the detection results on the training set.

On the evaluation set, the feature-based classifiers have a larger decrease in performance compared to the training set than the two tophat detectors. This may be due to over training, where the settings of these classifiers (and the features) are chosen so that they are specific for the training set (and thus have an optimal performance), but fail on the independent evaluation set.

However, for a number of points the feature-based classifiers are still significantly better than the tophat detectors on the evaluation set. Especially there is one point of Naive Bayes with 39 detections and only 32 false alarms. This point may be very suitable for sensor-fusion where the GPR sensor may be able to give confirmation of these points.

VIII. CONCLUSIONS

In this paper we have shown working pre-processing algorithms necessary for a polarimetric IR setup consisting of a rotating filter and a standard MWIR camera. Using a laser distance meter, motion estimation and correction functions, (polarimetric) Stokes images are reconstructed.

The Stokes images are input for our detection and classification algorithms. Classification is performed on features measured in the Stokes images at locations indicated by a

(a) Training set
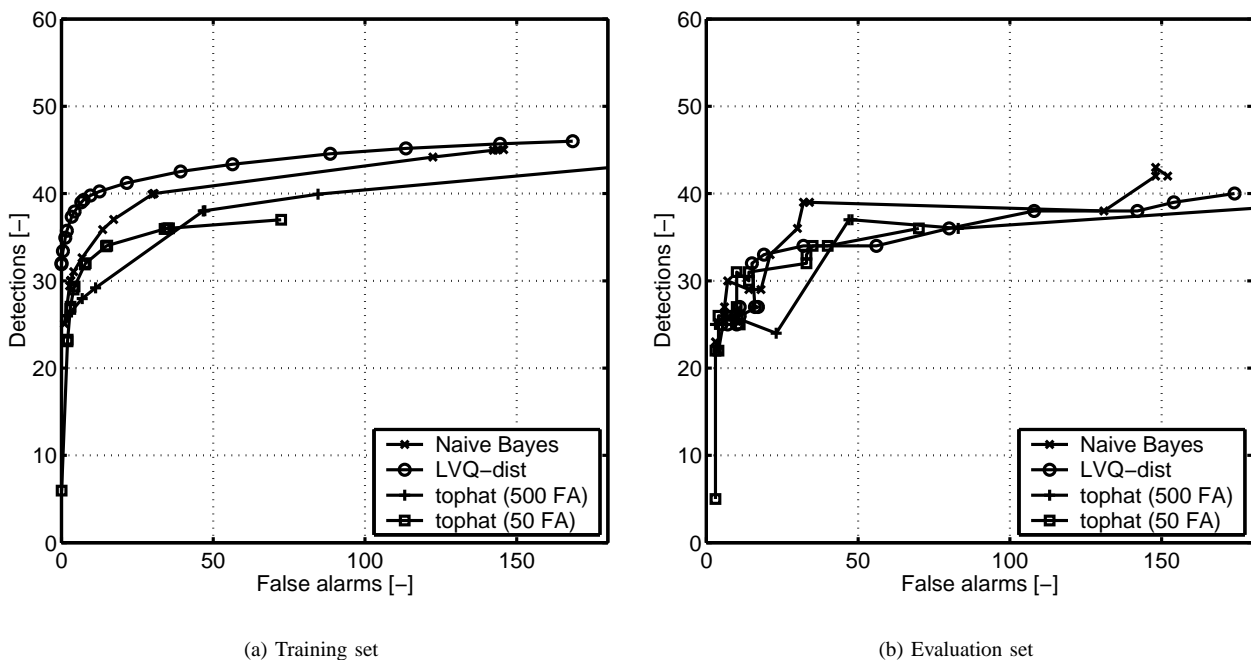

(b) Evaluation set

Fig. 7.   The results on the training set and the evaluation set for the various methods.

tophat detector. The best combination of features is determined by exhaustive search.

The feature-based classifiers with polarimetric and shape features give an improvement on the training set. The performance on the independent evaluation set is lower for all classifiers. Some ROC points of the feature-based classifiers are significantly better than the ROC of the tophat detector.

Feature-based classification using shape and polarimetric features improves the detection performance of landmines in polarimetric IR images.

REFERENCES

[1] F. Cremer, K. Schutte, J. G. M. Schavemaker, and E. den Breejen, "A comparison of decision-level sensor-fusion methods for anti-personnel landmine detection," *Information Fusion*, vol. 2, no. 3, pp. 187–208, Sept. 2001.

[2] F. Cremer, W. de Jong, and K. Schutte, "Fusion of polarimetric infrared features and GPR features for landmine detection," in *2nd International Workshop on Advanced Ground Penetrating Radar (IWAGPR)*, Delft, The Netherlands, May 2003.

[3] A. G. Yarovoy, L. P. Ligthart, A. Schukin, and I. Kaploun, "Polarimetric video impulse radar for landmine detection," *Subsurface Sensing Technologies and Applications*, vol. 3, no. 4, pp. 271–293, Oct. 2002.

[4] E. Hecht, *Optics*, 2nd ed.   Addison-Wesley publishing company, Reading (MA), USA, 1987.

[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., Orlando (FL), USA, Apr. 1999, pp. 1409–1418.

[6] F. Cremer, W. de Jong, and K. Schutte, "Infrared polarisation measurements and modelling applied to surface laid anti-personnel landmines," *Optical Engineering*, vol. 41, no. 5, pp. 1021–1032, May 2002.

[7] 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., Orlando (FL), USA, Apr. 2002, pp. 51–62.

[8] F. Cremer, J. G. M. Schavemaker, W. de Jong, and K. Schutte, "Comparison of vehicle-mounted forward-looking polarimetric infrared and downward-looking infrared sensors for landmine detection," in *Proc. SPIE Vol. 5089, Detection and Remediation Technologies for Mines and Minelike Targets VIII*, R. S. Harmon, J. T. Broach, and J. John H. Holloway, Eds., Orlando (FL), USA, Apr. 2003.

[9] 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., Orlando (FL), USA, Apr. 2002, pp. 108–119.

[10] U. Seiffert and B. Michaelis, "Directed random search for multiple layer perceptron training," in *Neural Networks for Signal Processing XI*, D. J. Miller, Ed.   Piscataway (NJ), USA: IEEE, 2001, pp. 193–202.

[11] N. A. Macmillan and C. D. Creelman, *Detection theory: a user's guide*. Cambridge university press, Cambridge (UK), 1991.

[12] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 1, pp. 4–37, Jan. 2000.

[13] F. Provost and T. Fawcett, "Robust classification for imprecise environments," *Machine Learning*, vol. 42, no. 3, pp. 203–231, Mar. 2001.

[14] J. Laaksonen and E. Oja, "Classification with learning k-nearest neighbors," in *Proceedings of ICNN'96*, Washington DC, USA, June 1996, pp. 1480–1483.

[15] 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., Orlando (FL), USA, Apr. 1999, pp. 1235–1245.

[16] S. M. Weiss and C. A. Kulikowski, *Computer systems that learn*. Morgan Kaufmann Publishers, 1991.