Usage of polarisation features of landmines for improved automatic detection

Wim de Jong[a], Frank Cremer[abc], Klamer Schutte[a] and Jesper Storm[d]

[a]TNO Physics and Electronics Laboratory, P.O. Box 96864, 2509 JG, The Hague, The Netherlands
[b]Pattern Recognition Group, Delft University of Technology, Delft, The Netherlands
[c]Section of Applied Geophysics, Delft University of Technology, Delft, The Netherlands
[d]Danish Defence Research Establishment, P.O. Box 2715, DK-2100 Copenhagen, Denmark

ABSTRACT

In this paper the landmine detection performance of an infrared and a visual light camera both equiped with a polarisation filter are compared with the detection performance of these cameras without polarisation filters. Sequences of images have been recorded with in front of these cameras a rotating polarisation filter.

Due to optical distortion of the rotating filter the sequence contains apparent motion. This motion is estimated from the images and corrected using bilinear interpolation. From the motion corrected sequences both the unpolarised intensity and the polarisation parameters are estimated.

The detection performance is evaluated with Receiver Operator Characteristics curves. It is shown that, for higher detection rates, classifiers based on polarisation features perform better than classifiers based on the intensity cue only. The linear combination of both has the best overall performance for the scenes considered. The infrared camera outperforms the visual light camera in the difficult scene.

Keywords: polarisation, infrared, visual, landmine detection

1. INTRODUCTION

One of the sensors often used for the detection of abandoned landmines is the infrared camera. However, since the detection requirements are very high,[1] the detection performance has to be increased. A possibility to increase the detection performance of cameras is by using polarisation features. With the polarisation technique (parts of) mines can be discriminated from a cluttered background, since man made objects have different polarisation features from natural objects. This effect holds both in the visible and the IR, where IR is not suffering from low light level conditions. Using the polarisation filter with an IR camera especially enhances the IR performance in situations with very low thermal contrast, since the IR polarisation contrast is independent of temperature contrast. This enhanced performance may be used to improve a subsequent sensor-fusion process[2] under these conditions.

To test the hypothesis of enhanced performance we conducted several measurements with a VIS and MWIR camera, both equipped with a rotating polarisation filter. This paper will show if and by how much these filters improve the detection performance. The performance increase of the IR camera will be compared with the performance of a visual light camera, with and without polariser.

Due to optical imperfections in the rotating filter a specific spot in the field of view of the cameras will appear on different pixels in subsequent images of a sequence. This so called motion reduces the accuracy of the estimation of polarisation parameters from the image sequence. To avoid this effect, the motion between the images is estimated and corrected before the polarisation parameters are determined.

After this motion correction, we estimated the Stokes-Müller[3] polarisation parameters. Some of the polarisation parameters are used to discriminate between objects and background. The detection results are presented in a Receiver Operator Characteristic (ROC) curve and compared with detection results of the unpolarised images.

Further author information: (Send correspondence to Wim de Jong)
Wim de Jong: E-mail: W.deJong@fel.tno.nl Phone: +31 70 374 0438 Fax: +31 70 374 0654
Jesper Storm: E-mail: jes@ddre.dk

In *Detection and Remediation Technologies for Mines and Minelike Targets V*,
Abinash C. Dubey, James F. Harvey, J. Thomas Broach, Regina E. Dugan, Editors,
Proceedings of SPIE Vol. 4038 (2000) • 0277-786X/00/$15.00

2. POLARISATION

In general, man-made objects have smooth surfaces. These smooth surfaces will reflect and/or emit polarised visual and infrared radiation, where the degree of polarisation depends on the surface properties and the angle between the surface normal and the line of sight. Natural occurring objects, however, have generally rougher surfaces and they will be reflecting or emitting radiation which is less polarised.

The measured intensity I_p as a function of the angle φ, where φ is the angle between the principal axis of the polarisation filter and the horizon, is given by[3]:

$$I_p(\varphi) = I + Q\cos(2\varphi) + U\sin(2\varphi), \tag{1}$$

where $\varphi = 0$ represents the situation that all horizontal polarised light waves pass the linear polariser.

If the intensity is measured for half a period (or $k\pi, k \in \{1, 2, \ldots\}$), then the Stokes-Müller polarisation parameters $\tilde{I}$, $\tilde{Q}$ and $\tilde{U}$ are estimated by:

$$\begin{aligned} \tilde{I} &= \frac{1}{N}\sum_{i=1}^{N} I_p(\varphi_i) \\ \tilde{Q} &= \frac{2}{N}\sum_{i=1}^{N} I_p(\varphi_i)\cos(2\varphi_i) \\ \tilde{U} &= \frac{2}{N}\sum_{i=1}^{N} I_p(\varphi_i)\sin(2\varphi_i), \end{aligned} \tag{2}$$

with N the number of frames, i the frame number and $\varphi_i = \frac{ik\pi}{N}$ the angle of the linear polariser.

3. MEASUREMENTS

The experiments have been performed in November 1999 at the TNO outdoor test facility.[4] During these measurements the weather was quite unstable with rain, hailstorms and a few sunny periods. The thermal contrast between mines and background was relatively low (2 to 3 degrees C) since there was hardly any heating of the autumn sun and a contrast reduction due to the rain.

3.1. Scenes and mines

The mines have been laid out in 5 different areas: on a lawn, in a forest area, in scrubs, on sand and on an unpaved road. In most of these areas the top of the mines were fully visible by the camera. The environment in which the mines were placed was different in all scenes and contained different kinds of natural clutter, like plants, leafs and branches. In automatic detection of mines using unpolarised images this clutter raises a lot of false alarms. In this paper we will present results from the sand and from the forest area.

In all the scenes, different kind of mines have been used. The mine set consisted of a metal anti tank (AT) mine with the original casing and a plastic AT mine with a camouflage paint. Several anti personnel (AP) mines have been used. Small plastic AP mines with original casings, medium sized AP mines with a rubber top, a wooden mine and two different types of the butterfly mine.[4]

3.2. Measurement setup

The measurement setup consisted of two camera systems, VIS and MWIR, both mounted on a tripod. The MWIR is a Santa Barbara focal plane array in the 3-5 μm wavelength band. This focal plane array contains 256 by 256 pixels and has a total dynamic range of 100 dB, an instantaneous dynamic range of 66 dB and a frame rate of 42 Hz. The images of the MWIR are digitally recorded (16 bits). The VIS camera is a Philips LDH 0703, with 752 by 582 pixels, a signal to noise ratio of 59 dB and a frame rate of 25 Hz. The images of the VIS are grabbed from the analog output of the camera (10 bits).

Both cameras were equipped with a rotating polarisation filter. The IR filter consisted of a wire grid on a KRS-5 substrate and the VIS polarisation filter was a polaroid sheet.

The filter was rotating at a constant speed. A trigger pulse from the rotation setup started the acquisition of the frame grabber in the computer. Since the camera is running at a fixed frame rate, the time between the frames and thus the rotation of the polarisation filter between the different frames is fixed. A complete sequence consists of 48 images for the VIS and 100 images for the MWIR. The sequences contain at least one full rotation of the filter, which had a rotation time of just under 2 seconds.

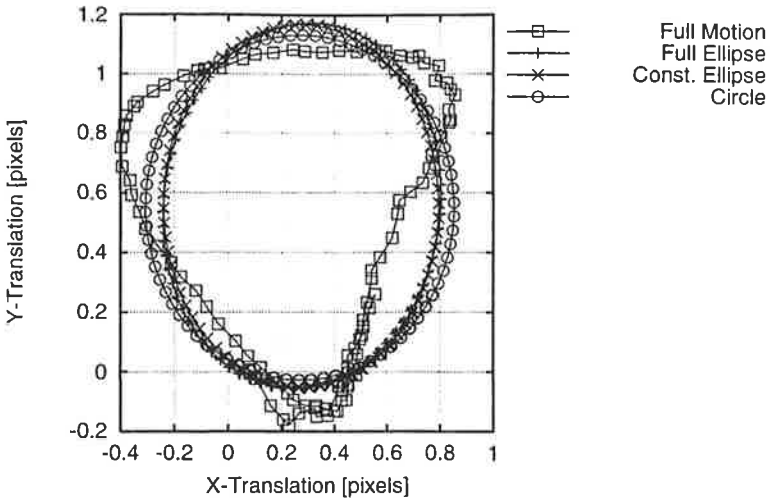


Figure 1. Estimated translation of the images in the MWIR forest sequence, relative to the first image.

4. ANALYSIS

4.1. Motion estimation and correction

If each pixel contains independent Gaussian noise with variance σ^2 then the estimation of I, Q and U also contains Gaussian noise. The variance of this noise is $\frac{\sigma^2}{N}$ for the noise on $\tilde{I}$ and $\frac{2\sigma^2}{N}$ for the noise on $\tilde{Q}$ and $\tilde{U}$. An estimate of the variance of the Gaussian noise $\tilde{\sigma}^2$ on every pixel is given by:

$$\tilde{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(\tilde{I} + \tilde{Q}\cos(2\varphi) + \tilde{U}\sin(2\varphi) - I_p(\varphi_i)\right)^2 \tag{3}$$

This estimated variance gives an indication of the motion induced by optical distortion of the rotating polariser in the sequence. When there is motion in the sequence, the edges between surfaces with different polarisation parameters will have a higher variance than the surfaces themselves. For the rest of the paper these edges will be identified by their higher than average variance. The surfaces are defined as non edges.

To correct for this motion, only translations in the horizontal and vertical directions are considered. After estimation of this translation[5] of each image relative to the first image, the images are translated back using bilinear interpolation.

The performance of the motion correction and estimation is measured by estimating the standard deviation or Root Mean Square (RMS) of the residual noise in the image: the root of Eq. 3. Two different parts of the image are evaluated separately: the edges and the surfaces of objects.

In the first line of Table 1 the residual noise for the uncorrected forest MWIR sequence is given. It is clear that there is motion, since the edges have a much higher residual noise than the surfaces.

Correction method	edges	surfaces	overall
Uncorrected	34.4	6.4	6.9
Full correction	7.2	5.5	5.5
Unconstrained ellipsoid	7.6	5.5	5.6
Constrained ellipsoid	7.6	5.5	5.6
Circular	7.3	5.6	5.6

Table 1. The average residual noise (RMS) in the MWIR forest sequence for the different sections of the image after correcting for motion induced by the rotation of the polarisation filter.

Correction method	edges	surfaces	overall
Uncorrected	32.9	14.8	15.4
Full correction	21.4	8.8	9.2
Unconstrained ellipsoid	21.6	9.0	9.4
Constrained ellipsoid	21.5	9.1	9.5
Circular	21.2	8.8	9.2
Sub image correction	14.2	8.2	8.4

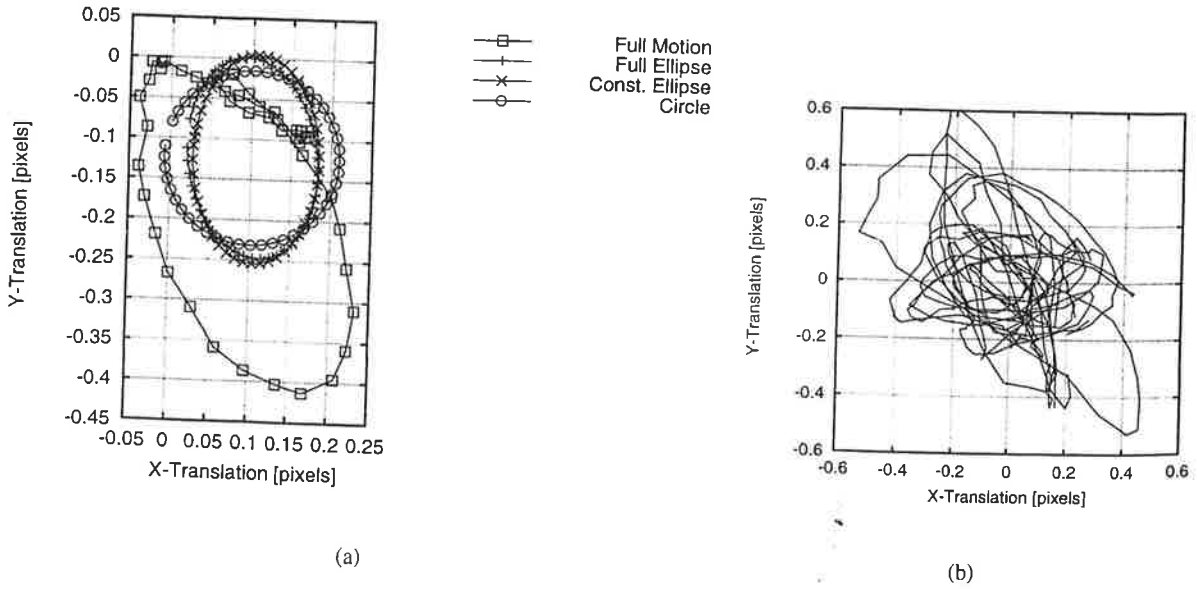
Table 2. The average residual noise (RMS) in the VIS forest sequence for the different sections of the image after correcting for motion induced by the rotation of the polarisation filter. For the sub image correction the image is divided into 9 sub images.

We first attempted to estimate the translation without any restrictions. The results are given by the full motion plot in Figure 1. The Fourier spectrum of this estimated translation shows a large component of twice the rotation frequency. This component may be caused by the polarised pixels, since the intensity variation of these pixels has this double frequency. However, if we constrain the translation estimation to ellipsoid, ellipsoids with the principle axis in the horizontal and vertical direction, or a circle, the residual noise is similar to the full correction, see the numbers in Table 1. After correction, the residual noise of the edges of the MWIR sequence has the same magnitude for the edges as for the surfaces, which indicates that the motion estimation is working.

The most simple motion model is the circular motion. With this motion estimation and correction the residual noise is one of the lowest and in the remaining of this paper, we will be using this motion model for correction and estimation for the MWIR camera.

For the VIS sequence, the residual noise for the full and restricted motion models are given in Table 2 and Figure 2(a). The residual noise for the edges still is substantially higher than the residual noise of the surfaces. This means that there is still motion present in the corrected images.

This residual noise can be caused by a curvature of the moving polaroid filter, leading to more complex motion than only translations. The image is divided into 9 sub images of equal sizes to investigate this supposition. A full motion estimation and correction is performed on these sub images. The motion is different for each sub image, see Figure 2(b). Subsequently, the sub images are joined. This procedure leads to a significant decrease of the residual noise. However, the difference between the edges and the surfaces is still significant, so a more complex motion model should be used to get further improvement.



(a) (b)

Figure 2. Estimated translation of the images in the VIS forest sequence, relative to the first image for the whole image (a) and the chaotic translations for the image divided into 9 sub images (b).

4.2. Performance evaluation

A classifier for mine detection assigns every pixel to one of the two classes: mines or background. Based on this classification, four situations occur:

1. correct classification of a mine (i.e. a detection),
2. correct classification of background,
3. wrong classification of a mine (i.e. a false alarm)
4. wrong classification of background (i.e. a missed detection).

The detection rate per frame is defined as the number of detected mines divided by the total number of mines. A mine is detected if one of the pixels is classified as a mine. For a complete evaluation of the classifiers based on the polarisation parameters, the detection rate as function of the false alarm rate will be presented in the form of an ROC curve.[6] The definition of the false alarm rate is not straight forward: there are no identifiable background objects. Three possible definitions exists:

1. fraction of the detected background pixels,
2. number of clustered background objects,
3. fraction of detected background grid cells (an area containing pixels) of a grid laid out over the image, for example SCOOP processing.[7]

The first definition does not take into account the spatial distribution of the pixels of the grid. The second definition is unbounded and does not take into account the size of the objects. The third definition takes into account both the distribution and the object size, for a well chosen grid cell size. This third definition will be used. Every grid cell with at least one detection in its pixels is counted as a detected grid cell.

This third type of false alarm calculation, follows possible operational scenarios more closely than the other two. Whenever there is a single detection it is most likely that the mine clearance personnel will take a larger area into consideration than this single detection. Thus an area around this detection will be prodded and therefore the fraction of detected background pixels is not a good measure for the workload of the mine clearance personnel. The same holds for the second definition: the size of a detection influences the amount of work, but it is not weighed in this definition.

The remaining of the image, that does not contain mines is divided into rectangles with a foot print of roughly 20 by 20 cm^2. The false alarm rate is defined as the fraction of these background rectangles with at least one detection.

4.3. Classifiers

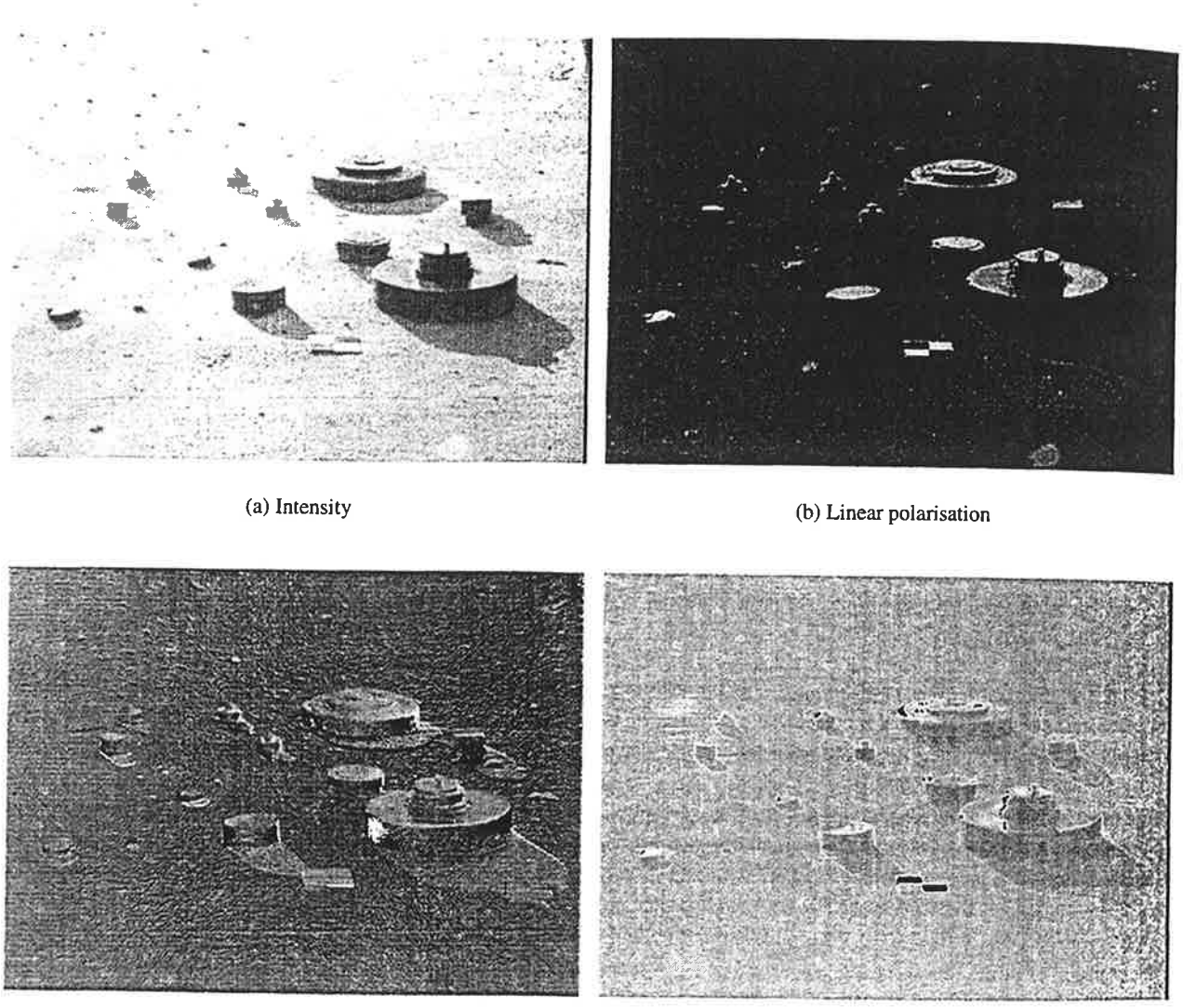
The mines in the MWIR sequences have a higher apparent temperature than their surrounding background. The intensity cue is a measure for the temperature and can therefore be used in a classifier: mines can be discriminated from background based on the intensity.

The mines in the VIS sequence have a different shade of grey than their background. That's the reason why they are visible*. The intensity of the background can be higher or lower than the intensity of the mines, so a single threshold will not suffice and double thresholds are necessary in the VIS sequence.

In this paper, classifiers based on the intensity cue (I) and the polarisation features are evaluated. The threshold of every classifier is set to a whole range of values. For each threshold the false alarm rate and the detection rate (i.e. a point of the ROC) is calculated. Only those points of the ROC are presented that are on the upper half of the convex hull through all points. These points are connected by lines, since every point on this line can be reached by combining two classifiers as defined by the adjacent points.[8]

For the classifiers based on polarisation features, the linear polarisation (LP) amplitude $LP = \sqrt{Q^2 + U^2}$ and the degree of linear polarisation (DoLP) $DoLP = LP/(LP+I)$ are used. Another classifier could be based on the angle of the polarisation found, but such a classifier is not yet considered by this paper.

*The mines can also visible due to different texture, but this feature is not used here.



(a) Intensity

(b) Linear polarisation

(c) Polarisation angle

(d) Residual noise (logarithm)

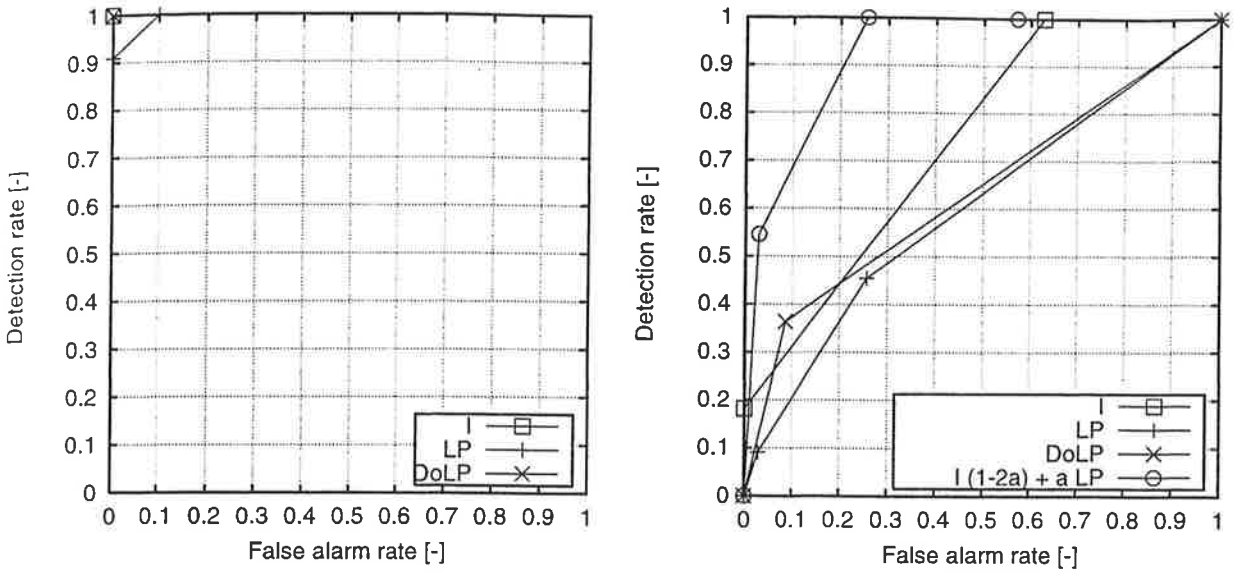
Figure 3. Images of VIS camera with test mines on a clean sand background after motion correction. Note that the angle image is not calibrated.

Combining intensity cues and polarisation cues may enhance the discriminating features of the mines in relation to their background. A simple form of combination is the linear combination: $I(1-2a)+aLP$, with a as a factor in the range between 0 and 1. For a range of scaling factors the detection results are calculated and the best results are included in the ROC. This implies that when one moves from one point of the ROC to the other, not only the threshold varies but also the scaling factor a. For the VIS camera, the absolute value of the difference between the pixel intensity and its mean is used as its intensity cue.

5. RESULTS

5.1. Sand test lane results

The visual images of the sand test lane after motion correction are shown in Figure 3. This scene is rather easy, since all mines are clearly visible in 3(a). The image of the linear polarisation, Figure 3(b), shows the top of the mines. The angle as shown in Figure 3(c) gives an indication of the orientation of the surfaces. The motion correction is not perfect since there are still visible edges present in the residual noise image Figure 3(d). These edges are so small compared to the size of the mines that



(a) VIS

(b) MWIR

Figure 4. ROC curves for both the VIS and the MWIR sequences of the sand test lane, for the intensity (I), the linear polarisation (LP) and the degree of linear polarisation (DoLP) and the linear combination of LP and I.

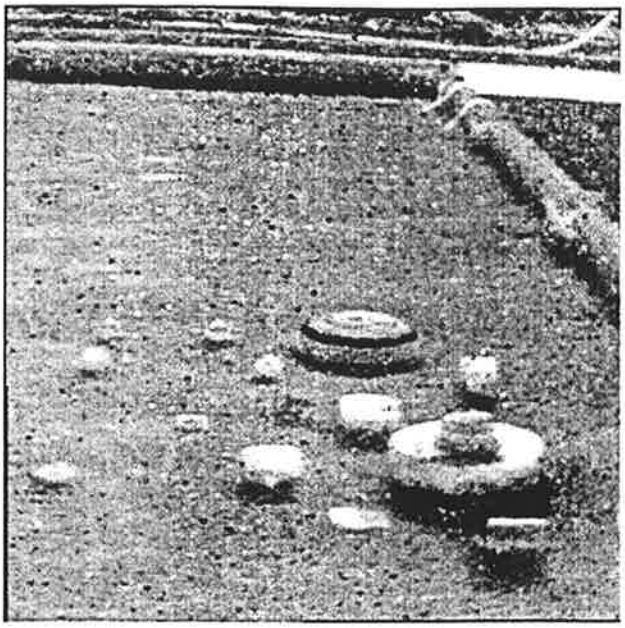
their influence on the detection performance is expected to be limited. Besides these edges, there is also a correlation with the intensity image visible, which needs to be studied.

The detection results on the images in this scene are given in Figure 4(a); only the I, LP and DoLP are shown. A mine is detected when one of its pixels has a detection. The detection results are perfect for the classifiers based on the intensity cue and the degree of linear polarisation, since 100% detection rate is achieved without any false alarms. On the intensity cue a double threshold is used centered around the mean pixel value. The classifier based on the linear polarisation cannot detect 100% of the mines without a false alarm. This is due to some spots in the background with a high LP, see Figure 3(b).

The images of the infrared camera are shown in Figure 5. Due to the low thermal contrast, the quality of the images is not optimal and they contain more noise than the images of the VIS camera. The edge of the test lane which is visible in these images is ignored to make the scene similar to the VIS scene. In the linear polarisation (Figure 5(b)) and the angle image (Figure 5(c)), the mines are less visible than in the VIS scene. The residual noise (Figure 5(d)), however, does not contain edges, so the motion correction is performing better on this sequence. This is more relevant for this scene, since due to the more limited resolution (256 by 256 pixels), motion has a larger impact.

The detection results of the MWIR forest sequence are given in Figure 4(b). As may be expected from the image quality, the detection performance is lower than for the VIS sequence. The intensity cue performs here also better than the classifier based on the linear polarisation. The linear combination of the intensity and the linear polarisation performs better than each single classifier.

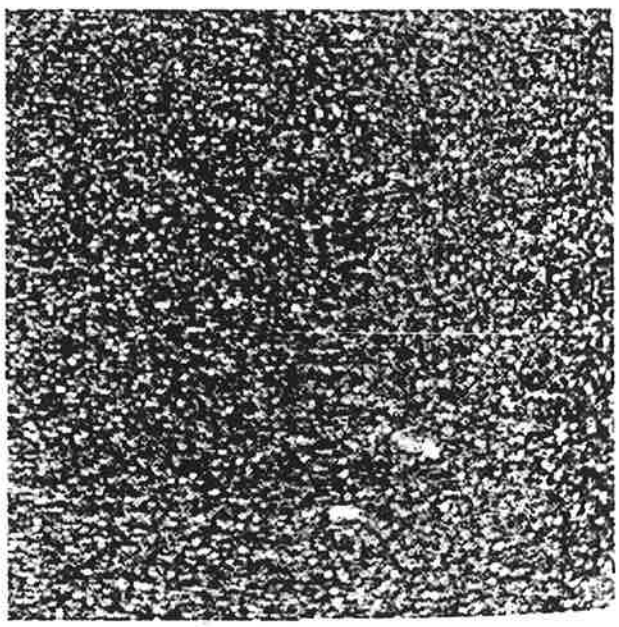
For this sequence, it is evident that an MWIR camera is not necessary and may even perform worse than the VIS camera. Nonetheless, the fusion of the intensity cue and the linear polarisation clearly improves the detection performance of the infrared camera: the circles in Figure 7(b).



(a) Intensity

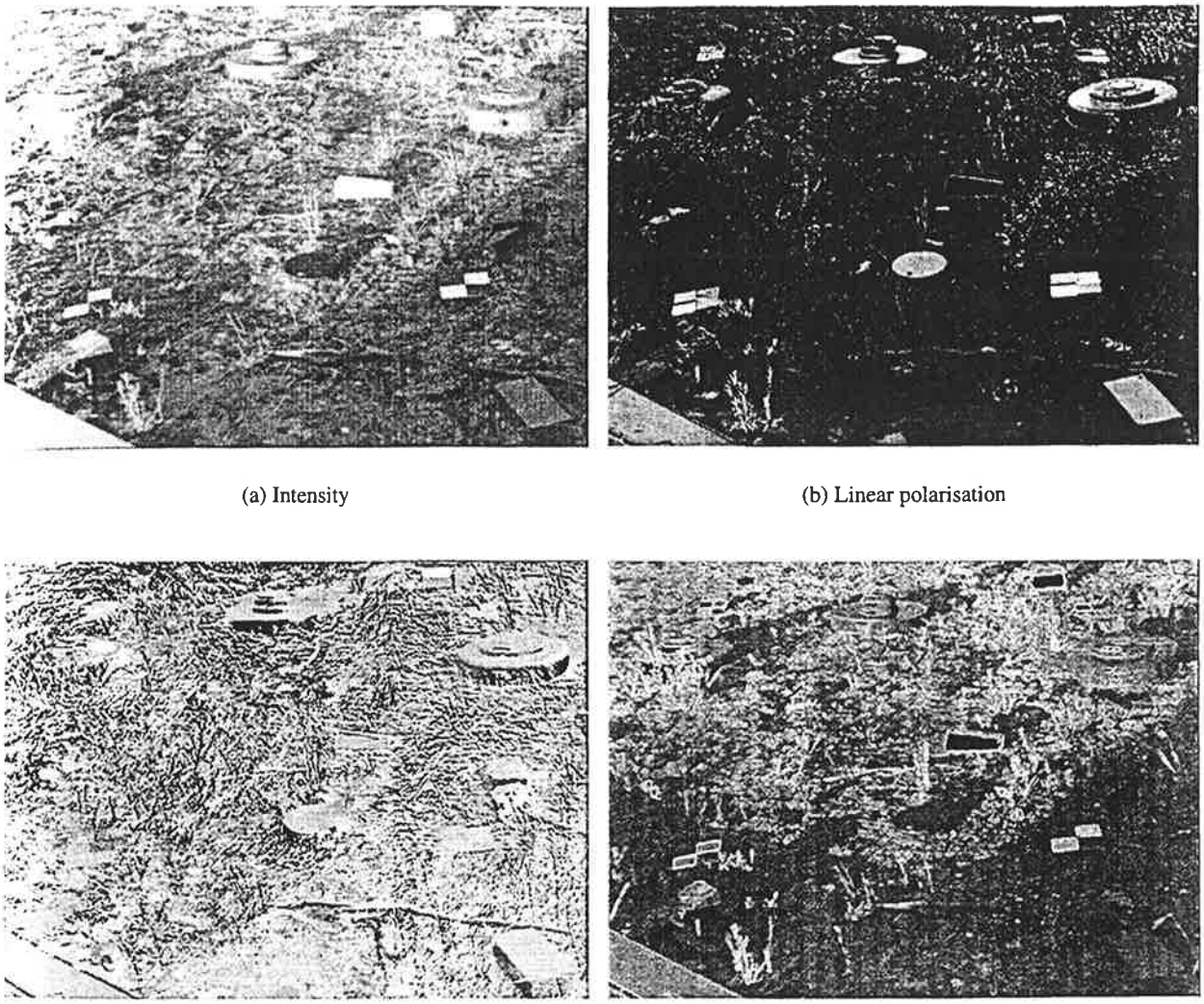
(b) Linear polarisation

(c) Polarisation angle



(d) Residual noise (logarithm)

Figure 5. Images of MWIR sand sequence after motion correction.



(a) Intensity

(b) Linear polarisation

(c) Polarisation angle

(d) Residual noise (logarithm)

Figure 6. Images of VIS forest sequence after motion correction.

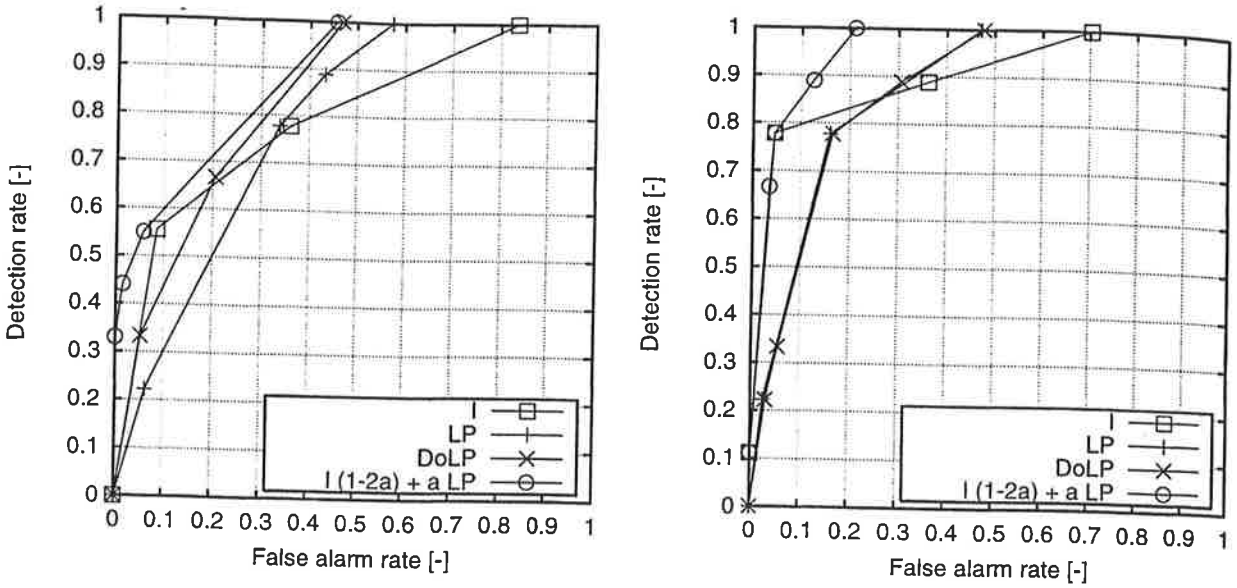
5.2. Forest area

This area represents a more difficult scenario than the sand lane. The scene contains fallen leaves and branches of trees. Most of the test mines here have been laid here over a year ago and have not been disturbed since. The images were taken under relative good weather conditions: the last hail shower was almost 5 hours ago and the sky was clear throughout the morning.

In Figure 6, the polarisation features for the VIS camera are shown. In the image representing the residual noise there are still edges visible, indicating that the motion estimation and correction is not perfect.

In Figure 7(a), the detection results based on the intensity cue and the polarisation features are shown for the MWIR sequence. For this sequence, classifiers based on polarisation features perform better than the classifier based on the intensity for high detection rates.

The optimal combination of the double threshold in the intensity and the linear polarisation gives the best detection results for the VIS. The double threshold is the absolute difference between the intensity and the average intensity. As before, the optimal weighing factors a and threshold differ for each point of the ROC.



(a) VIS

(b) MWIR

Figure 7. ROC curves for the forest sequence, for the intensity (I), the linear polarisation (LP) and the degree of linear polarisation (DoLP) and the linear combination I and LP.

In Figure 8, the estimated intensity, the linear polarisation and the polarisation angle are shown for the MWIR camera acquired at the forest area. The residual noise does not show edges, so the motion correction seems to work correctly for this MWIR sequence.

In Figure 7(b), ROC curves for the MWIR sequence are shown. It is clear that for lower false alarm rates, the intensity cue can make a good discrimination between objects and background, however for high detection rates the linear polarisation has a lower false alarm rate than the classifier based on the intensity.

The optimal linear combination of the intensity cue and the linear polarisation performs better than each individual classifier. The weigh factor a is optimal chosen for each threshold.

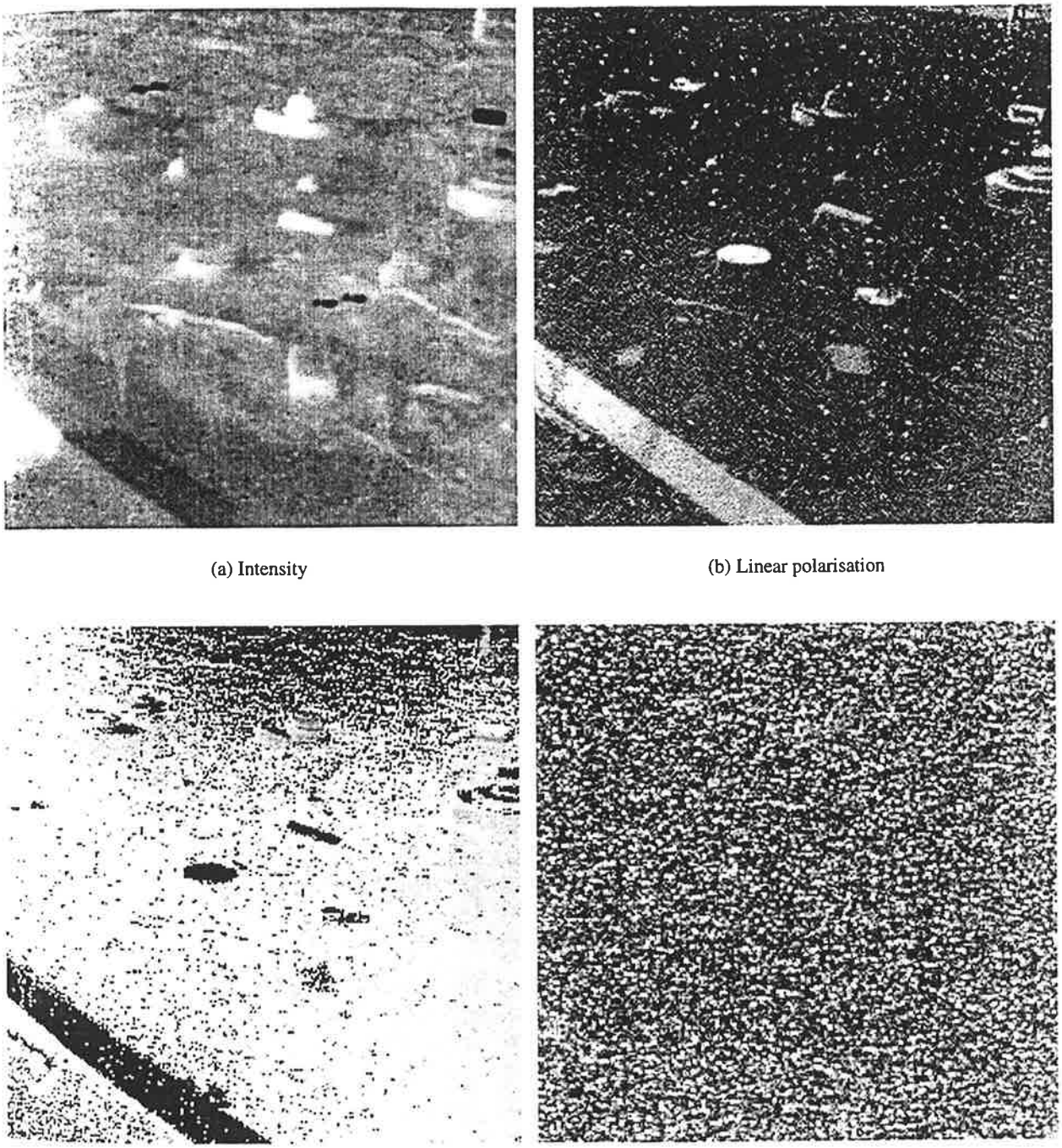
The detection performance of the VIS camera on this sequence is lower than the detection performance of the MWIR camera. This may be caused by the more discriminating intensity cue of the MWIR, due to differences in thermal properties between the mines and their background.

6. CONCLUSIONS

There is a need for better detection performance for the detection of abandoned landmines. One way to increase the detection performance is by using (infrared) cameras with polarisation filters. We have evaluated the detection performance of an VIS and an MWIR camera with a rotating polarisation filter in front of the lens. A problem is that such a filter induces motion in the sequence due to optical distortions. The motion is estimated and corrected almost completely.

The advantage of this setup with rotating polarisation filter compared to just a single camera is that more information about pixels is available. Furthermore the noise on the images is reduced due to the fact that more images are recorded. For high detection rates it is shown that the polarisation features (LP and DoLP) alone perform better than detection based on the intensity cue alone. The optimal combination or fusion of the intensity and the linear polarisation gives the best detection performance.

In the images of the sand sequence with an almost homogeneous background the VIS camera performs much better than the MWIR camera. In the forest area the performance of the VIS camera is lower than the performance of the MWIR camera, due to the presence of a lot of natural clutter. For this scene, the intensity cue of the MWIR camera performs better than the intensity cue of the VIS camera. Fusion with the linear polarisation gives even an better result.



(a) Intensity

(b) Linear polarisation

(c) Polarisation angle

(d) Residual noise (logarithm)

Figure 8. Images of MWIR forest sequence after motion correction.

7. FUTURE WORK

The MWIR camera used here is suffering from noise at low thermal contrasts. By using a camera from a newer generation with lower signal to noise ratio, the detection performance should be further increased. The difference in polarisation features of the MWIR and the long wave IR wavelength bands is also a topic that needs research.

Further improvement of the detection performance is possible by using the angular information and by applying spatial processing on the detection results. The angle of the principal axis of the polariser gives information about the orientation of objects. Upto now a single pixel can lead to a false alarm. By clustering pixels into objects the false alarm rate may be decreased.

More scenes will have to be evaluated to make a better comparison between the normal cameras and the cameras equipped with polarisation filters. The influence of different weather and lighting conditions on the detection performance must also be studied.

In this paper, the optimal threshold and fusion factor a were acquired and evaluated on the same data set. For an operational system, these parameters have to be set beforehand based on the performance of data from a training set. The consequences on the detection performance using separate data sets for training and evaluation are to be determined.

REFERENCES

1. A. Craib, "Standards for humanitarian mince clearance operations," in *International Conference on Mince Clearance Technology*, (Copenhagen, Denmark), July 1996.
2. F. Cremer, J. G. M. Schavemaker, E. den Breejen, and K. Schutte, "Towards an operational sensor fusion system for anti-personnel landmine detection," in *Proc. SPIE Vol. 4038, Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey and J. F. Harvey, eds., (Orlando (FL), USA), Apr. 2000.
3. E. Hecht, *Optics*, Addison-Wesley publishing company, Reading (MA), USA, second ed., 1987.
4. 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. 1409–1419, (Orlando (FL), USA), Apr. 1999.
5. L. J. van Vliet and C. L. L. Hendriks, "Improving spatial resolution in exchange of temporal resolution in aliased image sequences," in *SCIA'99, Proc. 11th Scandinavian Conference on Image Analysis*, B. Ersboll and P. Johansen, eds., pp. 493–499, (Kangerlussuaq, Greenland), June 1999.
6. N. A. Macmillan and C. D. Creelman, *Detection theory: a user's guide*, Cambridge university press, Cambridge (UK), 1991.
7. 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.
8. M. J. J. Scott, M. Niranjan, D. G. Melvin, and R. W. Prager, "Maximum realisable performance: a principled method for enhancing performance by using multiple classifiers," Tech. Rep. CUED/F-INFENG/TR. 320, Cambridge Univeristy Engineering Department, Cambridge, UK, Apr. 1998.

PROCEEDINGS OF SPIE

SPIE—The International Society for Optical Engineering

Detection and Remediation Technologies for Mines and Minelike Targets V

Abinash C. Dubey
James F. Harvey
J. Thomas Broach
Regina E. Dugan
Chairs/Editors

24–28 April 2000
Orlando, USA

Volume 4038
Part One of Two Parts