Towards an operational sensor-fusion system for anti-personnel landmine detection

Frank Cremer[abc], Klamer Schutte[a], John G.M. Schavemaker[a] and Eric den Breejen[a]

[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

ABSTRACT

To acquire detection performance required for an operational system for the detection of anti-personnel landmines, it is necessary to use multiple sensors and sensor-fusion techniques. This paper describes five decision-level sensor-fusion techniques and their common optimisation method.

The performance of the sensor-fusion techniques is evaluated by means of Receiver Operator Characteristics curves. These techniques are tested on an outdoor test facility. Three of four test lanes of this facility are used as training set and the fourth is used as evaluation set.

The detection performance of naive Bayes, Dempster-Shafer, voting and linear discriminant are very similar on both the training and the evaluation set. This is probably caused by the flexibility of the sensor-fusion techniques resulting into similar optimal solutions independent of the fusion technique.

Keywords: Sensor-fusion, naive Bayes, Dempster-Shafer, voting, linear discriminant, landmine detection.

1. INTRODUCTION

The ESPRIT project LOTUS strives towards a vehicle mounted sensor-fusion system for the detection of anti-personnel (AP) landmines. The sensor suite of this system consists of a ground penetrating radar (GPR), a metal detector (MD) and a thermal infrared (IR) camera. The detection performance of each single sensor is not high enough to be used in an operational scenario. Using sensor-fusion methods, the detection performance can be improved.

To test the sensors and sensor-fusion performance, measurements have been taken at the TNO outdoor test facility.[1] The single-sensor results are presented per grid cell (2.5 cm by 2.5 cm) as well as the detection performance.

The confidence levels are combined by different decision-level sensor-fusion methods: best sensor, naive Bayes, Dempster-Shafer, voting and linear discriminant. Four of these methods were evaluated earlier in LOTUS' predecessor, the GEODE project[2-4] on data determined on a smaller test site. In this paper, the optimisation parameters will be acquired on three test lanes and evaluated on the fourth. This differs from the leave-one-out evaluation[3] and training and evaluation on the same set.[4]

2. MEASUREMENT SETUP

2.1. Test lanes

The first LOTUS trial took place from May 17th through 21st 1999 at the TNO test site, situated in the dunes of the Hague, the Netherlands. This test site consists of 6 test lanes, each with a different soil type: sand, clay, peat, ferruginous, forest and road. Each test lane measures 3 by 10 m^2 and has a controlled water table. The test lanes are free of vegetation and the soil of the first four test lanes is filtered to remove metal objects. In the test lanes, a mixture of anti-personal (AP) mines and anti-tank (AT) mines are buried or laid on the surface. The sensors are attached to a metal free X-Y table for taking stand-off measurements.

In Table 1 an overview is given of the mine types in the six test lanes. In total there are 265 mines consisting of 216 APs and 49 ATs. The mine types AP5 and AT3 are metal free. The others (73%) contain at least some metal, see Table 2. The first four test lanes contain roughly similar number of mines of each type. However, there are more mines present of mine type AP5 and AP7 on the first four test lanes.

Further author information: (Send correspondence to Frank Cremer)
Frank Cremer: E-mail: Cremer@fel.tno.nl Phone: +31 70 374 0795 Fax: +31 70 374 0654

lane	AP type							AT type			FA	other
	1	2	3	4	5	6	7	1	2	3		
sand	5	4	5	5	9 (6)	5	7 (5)	3 (2)	3 (2)	4 (2)	9	3
clay	5	4	4	4	10 (7)	5	8 (6)	3 (2)	3 (2)	4 (2)	9	3
peat	4	4	4	4	11 (8)	4	9 (7)	3 (2)	3 (2)	4 (2)	9	3
ferruginous	5	4	4	5	10 (7)	5	9 (7)	3 (2)	3 (2)	4 (2)	9	3
forest	6	6	-	6	9	6	2	-	-	1	9	3
road	5	3	5	-	8	-	2	-	3	6	11	3
total	30	25	22	24	47	25	37	12	15	23	56	18
total sensor-fusion evaluation	19	16	16	18	28	25	25	8	8	8	32	8

Table 1. The number of mines and false-alarm (FA) targets sorted on mine type and soil type. AP5 and AT3 are non-metal mines. The other objects are sounding tubes for water table control and thermocouples for temperature measurements. The mines in the first four lanes are used in the sensor-fusion evaluation; with between brackets the number of mines, if not all mines were used.

Type	casing material	metal content
AP1	wood	medium
AP2	plastic	high
AP3	PVC	medium
AP4	PVC	medium
AP5	plastic	non
AP6	ABS	low
AP7	PE	medium
AT1	metal	high
AT2	PVC	low
AT3	no casing	non

Table 2. The different types of test mines that have been used in the test lanes. Non metal mines contain 0 g metal, low metal mines upto 10 g metal, medium metal mines 10 to 20 g metal and high metal mines more than 20 g.

The burial depth is distributed as shown in Table 3. Depth -1 means that the top of the mine is just below the surface of the soil. Most of the mines (78%) are buried, but only 31% is buried deeply (5 cm and below for APs and 30 cm for ATs). This scenario is made more difficult than can be expected in a real mine field. In a real mine field, we expect more mines on the surface and relative more metal containing mines.

lane	AP depth [cm]				AT depth [cm]			FA depth [cm]				
	-10	-5	-1	5	-30	-1	10	-30	-10	-1	5	10
sand	8 (7)	6	21 (19)	5 (3)	3	4 (3)	3 (0)	1	2	3	2 (1)	1
clay	8 (7)	5	21 (19)	6 (4)	3	4 (3)	3 (0)	1	2	3	2 (1)	1
peat	8 (7)	4	21 (19)	7 (5)	3	4 (3)	3 (0)	1	2	3	2 (1)	1
ferruginous	8 (7)	6	21 (19)	5 (3)	3	4 (3)	3 (0)	1	2	3	2 (1)	1
forest	11	-	12	12	-	-	1	-	3	4	2	-
road	5	-	9	9	2	3	3	1	2	5	2	1
total	48	21	105	35	14	20	15	5	13	21	12	5
total sensor-fusion evaluation	28	21	76	15	12	12	0	4	8	12	4	4

Table 3. The depth of the mines and false-alarm targets as measured from the top of the targets to the surface. The mines in the first four lanes are used in the sensor-fusion process; with between brackets the number of mines, if not all mines were used.
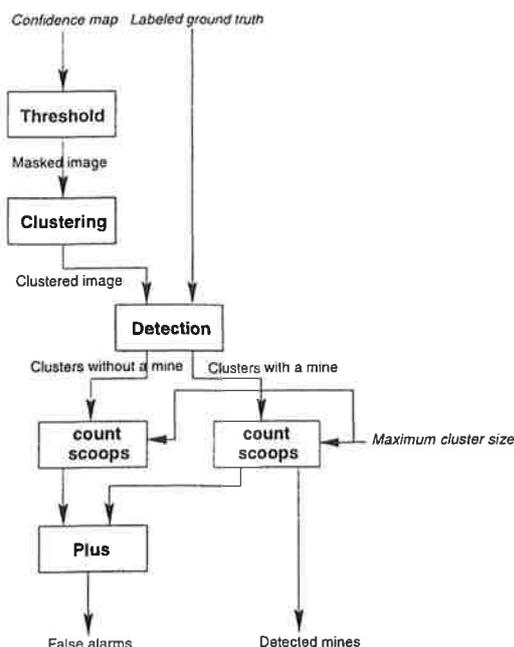
Figure 1. *The flowchart for the SCOOP evaluation method.*

2.2. Sensors

The GPR by Emrad has a bandwidth matched to the expected radar cross-section of anti-personnel mines. Radar data was collected on 3.75×7.5 cm^2 grid cells. The spacing of the antenna elements in the radar is approximately 15 cm. Successive scans were taken with the antennas shifted in such a way that the test bed was measured with the required density.

The Foerster Minex 2FD array is a dual-frequency, continuous-wave, metal detector. It is difficult to detect small metal objects which are close in distance to a large metal object with this detector, because the weak signals of the small objects are superimposed on the strong signal of the large object. In such areas where a strong signal of a large metal object is present, the detector is to some degree blind to other smaller metal objects. To compensate for this, areas with a strong signal are set to a non-zero confidence level.

Images of the IAI/TAMAN infrared camera in the wavelength band 3-5 μm were recorded and pre-processed by TNO. The mines have a higher temperature (and thus a higher intensity in the image) than their surrounding. The intensity is used as a confidence level. Spatial processing is applied to reduce the amount and size of the false-alarms.

3. PERFORMANCE EVALUATION

For a comparison of performance, Receiver Operator Characteristics (ROC) curves will be used. In a ROC curve, the detection rate is plotted versus the false-alarm rate. The detection rate is defined as the fraction of the detected mines. The number of false-alarms per unit area will be calculated using the SCOOP[4] evaluation method.

The flowchart of the SCOOP evaluation method is shown in Figure 1. The detections of grid cell values above the threshold are clustered. Every cluster is treated differently whether or not it contains one or more mines. If a cluster does not contain any mines it is counted as many false-alarms as it contains scoops. Clusters with one or more mines that are larger than the product of the number of mines and the scoop size also contain false-alarms. The number of false-alarms is set equal to the cluster size divided by the scoop size minus the number of mines. The SCOOP size is set to an area of 20 by 20 cm^2.

The SCOOP evaluation method results in a percentage of detected mines as a function of the number of false-alarms per unit area. This kind of performance measure is requested by the demining experts as an indication of the reduction in workload of the human operator and mine-clearance personnel. It takes into account that it takes mine clearance personnel more time to reject a false-alarm occupying a large area than a false-alarm occupying small area.
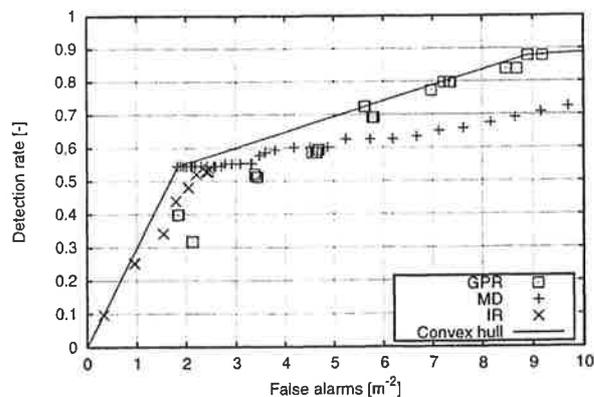
Figure 2. The performance of the GPR, MD and IR sensors for a range of thresholds and the convex hull of the best sensor on the "not sand" training set.

3.1. Optimisation method

Each of the fusion methods with a fixed parameter setting (i.e. operating point) can be seen as a classifier: it assigns a class (mine or background) to each input. By varying the parameter settings, different classifiers are constructed. These classifiers differ in number of false-alarms per unit area and detection rates. The maximum realisable ROC of these classifiers is defined as the upper half of the convex hull of all possible detection rates and false-alarm rates of the classifiers. Interpolation between points of the ROC can be done by random switching between the classifiers which generate the neighbouring ROC points.[5-7]

An exhaustive search over all parameter settings is carried out to find the optimal classifiers. Note that this is different from more traditional ROCs in which only the final threshold is varied. In the presented ROCs, if one moves from one point to the next, not only the threshold but also other parameters may be different.

In Figure 2, the performance of the three sensors is shown, by varying the threshold on the confidence level. The best sensor method, see Section 5.1, selects the operating points of the sensor which are on the (upper half) of the convex hull of all points.

For naive Bayes, Dempster-Shafer and linear discriminant the mapping parameters (one per sensor), see Section 5, are varied in 11 steps and the threshold is set to the minimum value to detect all mines. This gives a total number of 1331 parameter settings, besides 101 final thresholds. For voting we used 32768 parameter settings and only four final thresholds. The best sensor implementation does an optimal search for the thresholds, so it does not have a discretisation of parameter settings.

3.2. Leave-lane-out evaluation

To create an independent training and evaluation set, there are a number of options. In previous work,[3] we have used leave-one-out for this purpose on the GEODE data set. This method is more suitable for small data sets with fewer than 30 examples.[8] The GEODE data set contains only 26 mines in 25 m². The disadvantage of this method is that it only works for 100% detection on the training set. To acquire points below 100% detection, mines have to be removed in a specific order.

Since the LOTUS data set is larger and there are four test lanes with almost similar mine patterns, our approach is different. The training set consists of three of these four test lanes and the evaluation set will be the remaining test lane. This evaluation method is called leave-lane-out evaluation. The soil type is different for each of these four test lanes, so this evaluation method test how generic the parameter settings are, as they are evaluated on a fourth (untrained) soil type. As such, the performance of this evaluation method is negatively biased. It can be seen as a more realistic performance, since it will be impossible to train for every variation of mine type, depth or soil type.

4. SINGLE SENSOR RESULTS

The raw and pre-processed sensor data for the peat test lane is shown in Figure 3, alongside with its ground truth. The GPR data presented here is a projection onto the surface of the reflected energy after background subtraction. For the metal detector the logarithm of the sum of one and the absolute signal is shown. The larger the signal, the more metal is present (or the more closer it is to the detector). When the array moves over a metal object the signal, there is a zero crossing in the signal. For the infrared a mosaic of the images is shown: the apparent temperature increases with intensity.

(a) GPR	(b) MD	(c) IR	(d) Ground truth
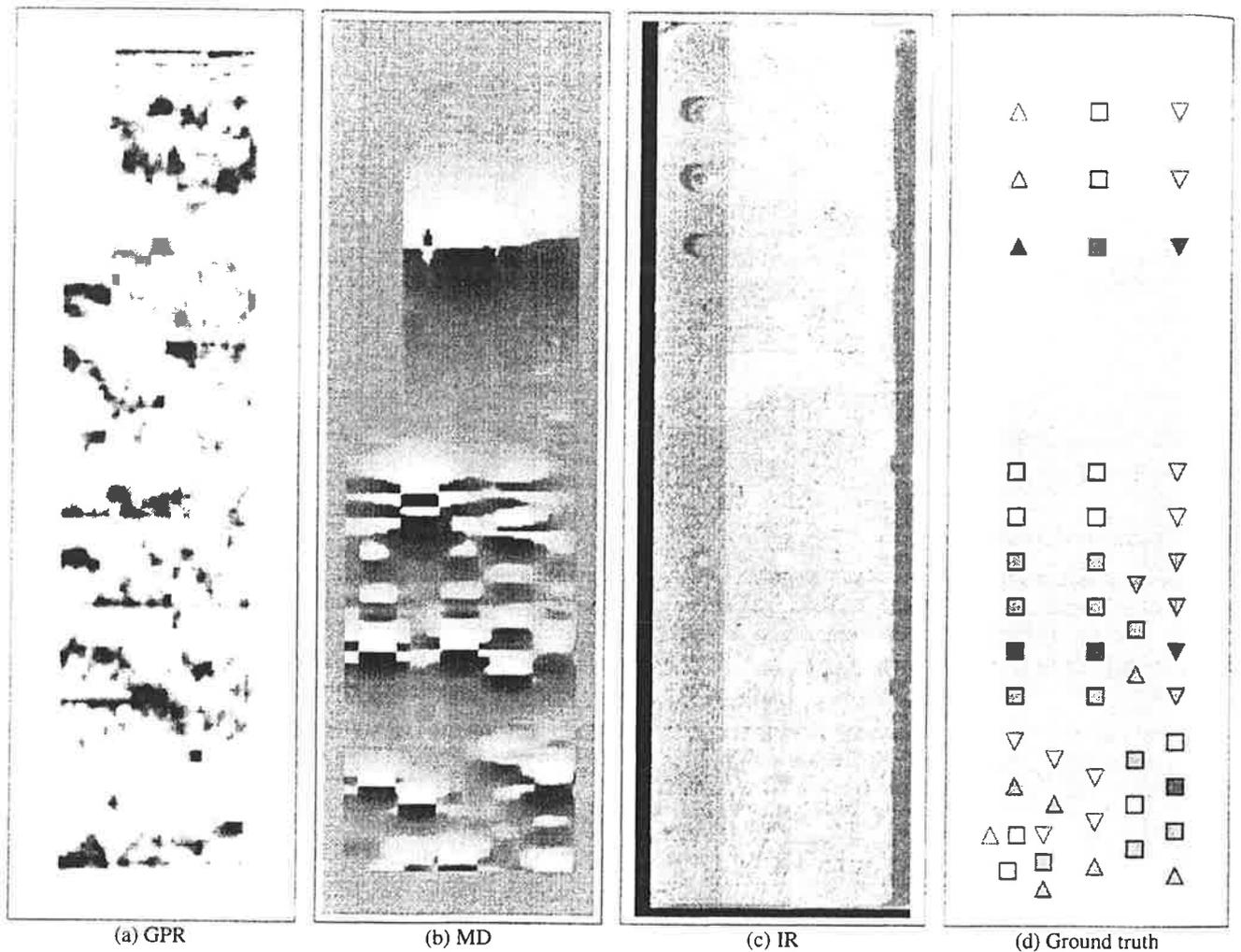
Figure 3. The sensors measurements and the ground truth for the peat test-lane. The grey area in the ground truth represents the part of the test lane that is measured by all three sensors. The triangles pointing up are mines on the surface; the squares are mines just below the surface and the triangles pointing down are buried mines. The metal content is shown in shades of gray, darker shades represents more metal.

The sampled and projected energy of the GPR is used as confidence level, which is the input of the fusion process.

For the MD, the zero crossings are detected and given a high confidence level. Since the processing of the MD cannot detect zero crossings superimposed on a large signal, the areas with a large signal are assigned a lower confidence (depending on the signal strength). The performance on the ferruginous test lane is not significantly lower compared to the other test lanes due to the use of two frequencies in the detector array.

The local contrast of the IR images is determined and used as an input for the spatial processing. As outcome of this spatial processing, areas are assigned a confidence level.

Based on these confidence levels, the detection performance of each sensor on each test lane is easily determined by applying a threshold and evaluated with the SCOOP evaluation method. These single-sensor results are shown in Figure 4. The detection results for the first four test lanes (sand, clay, peat and ferruginous) are similar.

The overall detection performance of the GPR is worst for the clay and the road test lane. The road test lane contains a lot of stones, which probably give a reflection. Difference in densities can also cause the lower detection rate in the clay test lane. The variations of the GPR performance are generally smaller than the variation of the other sensors.
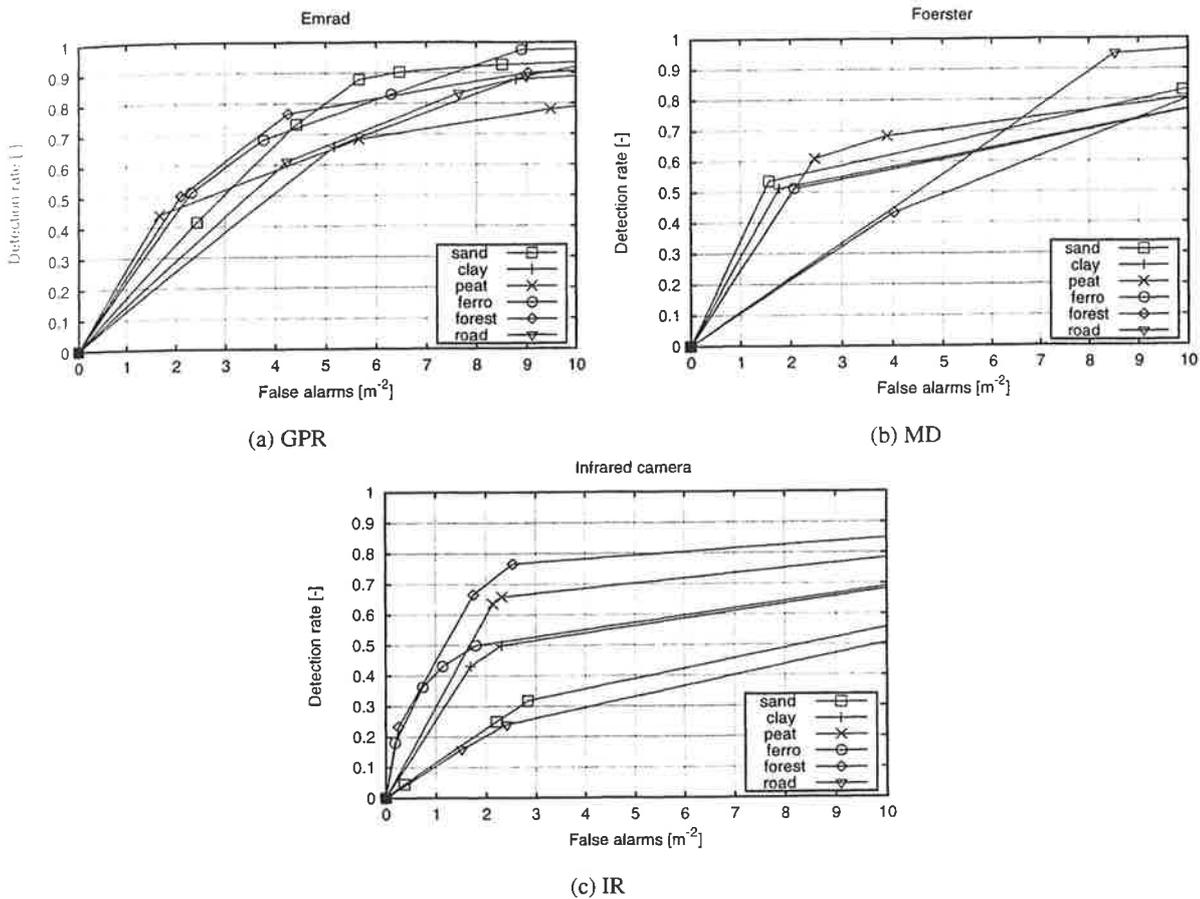
Figure 4. The detection results of the GPR, MD and IR on all test lanes.

The MD has difficulties with the forest and road test lane. The number of false-alarms is much higher than in the other lanes. This may be caused by the larger number of metal objects present in the these lanes. These metal objects are also the cause of the high detection rate (at the cost of high number of false-alarms) for the road; the mines are detected due to false-alarms in the vicinity. The operating point at around 2 false-alarms per square meter and 50% detection rate is rather stable for the first four test lanes.

The IR camera also shows bad performance on the road test lane. In this test lane a lot of clutter is present. Rocks and stones are not easily distinguished from mines, since they also have a positive temperature contrast. The detection performance on the sand test lane is also low compared to other lanes. This may be caused by the auto-gain of the IR camera influencing the detection on this test lane. Also difference in moisture of the soil can impact the detection performance.

The detection performance of the road is much lower for all three sensors. Since the road and forest test lanes also have a different layout than the other four test lanes, these two will not be considered in the rest of this paper.

5. SENSOR FUSION METHODS

The general layout of the five sensor-fusion methods is shown in Figure 5. The input of each sensor-fusion method is a confidence level per grid cell. A confidence level is a number in the range from zero to one. It is related to probabilities in the sense that a higher confidence level implies a higher probability. However, this relation is not necessarily linear and may even differ per sensor.
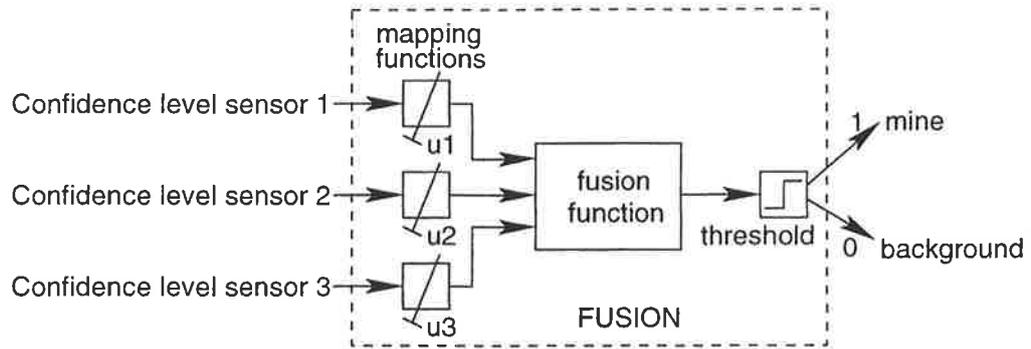
Figure 5. The generic decision-level sensor-fusion layout.

The output of the fusion process is one for a detection and zero for no detection per grid cell. Each of the methods scales the influence of each of the sensors in a different way. This mapping requires one parameter (u_1, u_2, u_3) per sensor. This mapping may remove the differences in definitions of the confidence levels.

The mapped inputs are combined in a fusion function to acquire a single value per grid cell. The mapping functions and the fusion function are given in Table 4. For the GEODE data set we also used the fuzzy probabilities method.[2] For most parameter settings, the discriminant functions of fuzzy probabilities are (almost) linear functions, so in this paper the linear discriminant is used instead.

method	mapping function	fusion function
best sensor	none	selection
naive Bayes	linear scaling around 0.5	product
Linear discriminant	linear scaling	summation
Dempster-Shafer	uncertainty level	Dempster's rule of combination
Voting	threshold	summation

Table 4. The different functions for scaling the input and combining these into a single (fused) result.

5.1. Best sensor

This method can be seen as the minimal sensor fusion method. It is the most simple form of sensor fusion: it just selects the best sensor. Which sensor is the best depends not only on the test lane, but also on the required detection rate (or number of false-alarms). To be able to apply this method, it still requires all three sensors on the training set, but for a given operating point only one of the sensors is necessary for the evaluation set.

5.2. Naive Bayes

The naive Bayes assumes that the confidence levels scale with the likelihood ratio per sensor. The likelihood ratio is the quotient of the conditional probability densities for both classes (mines and background). Based on this likelihood ratio the optimal Bayes decision can be made.[9] If these conditional probabilities are sensor independent, then the joint likelihood ratio is the product of these individual likelihood ratios. Since the confidence levels most likely do not linear scale with the likelihood ratios, a mapping factor is included for each sensor.

The naive Bayes fusion function $f(c_1, c_2, c_3)$ and the mapping are given by:

$$f(c_1, c_2, c_3) = \left((1 - u_1)c_1 + \frac{1}{2}u_1\right) \cdot \left((1 - u_2)c_2 + \frac{1}{2}u_2\right) \cdot \left((1 - u_3)c_3 + \frac{1}{2}u_3\right). \tag{1}$$

with c_1, c_2, c_3 the confidence levels and u_1, u_2, u_3 the mapping parameters.

5.3. Linear discriminant

The linear discriminant is another (naive) implementation of the Bayesian optimal decision. If confidence levels are interpreted as the logarithm of the likelihood ratio per sensor and these likelihood ratios are independent, then the summation gives the optimal decision boundary. Since the confidence levels are most likely not proportional to the likelihood ratio, a scaling factor is used. A linear classifier is also optimal when the sensors values have equal normal distributions.[9]

The linear discriminant fusion function and the mapping are given by:

$$f(c_1, c_2, c_3) = u_1 c_1 + u_2 c_2 + u_3 c_3. \tag{2}$$

with c_1, c_2, c_3 the confidence levels and u_1, u_2, u_3 the mapping parameters.

5.4. Dempster-Shafer

For application of the Dempster-Shafer theory[10] to sensor fusion, three inputs per sensor are needed: the probability mass assigned to a mine $m(M)$, the probability mass assigned to background $m(\bar{M})$, and the unassigned probability mass $m(M \cup \bar{M})$. The sum of these masses always equals one, so there are only two independent masses ($m(M)$ and $m(\bar{M})$). The mass $m(M)$ represents a belief in a mine, the mass $m(\bar{M})$ represents the belief in background, and the mass $m(M \cup \bar{M})$ reflects the uncertainty of the sensor. Each sensor produces one confidence level at each sample location, which must be mapped onto the three required probability masses. This is done by using the uncertainty as a optimisation parameter.

The confidence levels for sensor i are mapped onto probability masses, using:

$$m_i(M) = (1 - u_i)c_i \tag{3}$$
$$m_i(\bar{M}) = (1 - u_i)(1 - c_i) \tag{4}$$
$$m_i(M \cup \bar{M}) = u_i. \tag{5}$$

with u_i the mapping parameter and c_i the confidence level for sensor i. The probability masses for sensor 1 and 2 are combined using Dempster's rule of combination:

$$m_{1,2}(M) = m_1(M)m_2(M) + m_1(M)m_2(M \cup \bar{M}) + m_2(M)m_1(M \cup \bar{M}) \tag{6}$$
$$m_{1,2}(\bar{M}) = m_1(\bar{M})m_2(\bar{M}) + m_1(\bar{M})m_2(M \cup \bar{M}) + m_2(\bar{M})m_1(M \cup \bar{M}) \tag{7}$$
$$m_{1,2}(M \cup \bar{M}) = m_1(M \cup \bar{M})m_2(M \cup \bar{M}), \tag{8}$$

with $m_{1,2}(M)$ the combined probability mass assigned to a mine, $m_{1,2}(\bar{M})$ the combined probability mass assigned to a background and $m_{1,2}(M \cup \bar{M})$ the combined uncertainty mass. The probability masses of third sensor are combined with the other two by reapplying this rule for the combined probability masses for sensor 1 and 2 and the probability masses for sensor 3. The output of the Dempster-Shafer fusion function is given by the three combined probability mass assigned to a mine plus half the uncertainty:

$$f(c_1, c_2, c_3) = m_{1,2,3}(M) + \frac{1}{2}m_{1,2,3}(M \cup \bar{M}). \tag{9}$$

5.5. Voting

Voting fusion is another decision-level fusion method, see also Dasarathy[11] and Klein.[12] Our voting fusion is described by 4 thresholds: one for each sensor and one for the required number of votes. A vote is given by a sensor if the confidence level of this sensor is larger than the threshold. The votes are summed and the final threshold select between one out of three ("or" voting), two out of three or three out three ("and" voting) votes. Similar voting has also been applied to other sensor data acquired at the same test lanes, see Breuers[13] and Schwering.[14]

The voting fusion function and the mapping are given by:

$$f(c_1, c_2, c_3) = \text{Threshold}(c_1, u_1) + \text{Threshold}(c_2, u_2) + \text{Threshold}(c_3, u_3), \tag{10}$$

with c_1, c_2, c_3 the confidence levels and u_1, u_2, u_3 the mapping parameters. The threshold function is defined as:

$$\text{Threshold}(a, b) = \left\{ \begin{array}{l} 1, \text{if } a \geq b \\ 0, \text{otherwise} \end{array} \right. \tag{11}$$

6. SENSOR-FUSION RESULTS

The detection results per evaluation lane are given in Figure 6. The training set contains the three lanes and the evaluation set contains the remaining, fourth, lane. In the training set, there are 123 mines on 51 m^2. The sensor-fusion methods are optimised on the training set and only operating points on the upper half of the convex hull are presented in the ROC curves.

The parameters for the mapping function and the final threshold of the operating points of the training set are used to acquire the detection performance on the evaluation set. These resulting points of the ROC of the evaluation set are not necessarily located on a convex hull. The evaluation set contains 41 mines on 17 m^2.

6.1. Training set

The detection results on the training set are almost the same, independent of which sensor-fusion method is used, besides the best sensor method. A remarkable operating point is the point at 50% to 60% detection. At this point only the MD is used in all fusion methods. This point selects all zero crossings and thus all metal objects in the test lane. Outside this so called MD operating point all methods outperform the best sensor. It is not possible to rank the methods, since the ROCs are almost always overlapping.

This similar performance indicates that all of these four methods (naive Bayes, Dempster-Shafer, voting and linear discriminant) are so flexible that in the optimisation process similar performance is reached. We cannot prove that this performance is the optimal performance, but if four different methods reach this performance, it seems likely that this represents the highest possible with only four parameters.

6.2. Evaluation set

The detection results of the evaluation sets show more variation than the results of the training sets. This is not surprising since the number of mines is a factor three less compared to the training set. This means that for every mine missed, the detection rate drops by 2.5% versus 0.8% on the training set.

The lowest detection performance is found on the sand lane. For the individual sensors, we saw that the IR camera was not performing too well on this test lane. Since it performed well on the other three test lanes, it has influence on the selected operating points of the training set. Once evaluated on the sand test lane, it is apparent that these results will deteriorate. The best sensor does not suffer from this deterioration, since the IR camera is not selected in the training set. In fact, the best sensor for high detection rates (above 80%) is the GPR and its performance is better on the sand test lane than on the other test lanes.

On the clay test lane the four methods improve over the best sensor for a part of the operating points. The differences between these four methods are smaller than their variation, so it is not possible to rank these methods.

On the peat test lane the improvement of the four methods over the best sensor is even more evident. The performance of the three sensors is more similar on this test lane and apparently the sensor-fusion methods function better when this is the case. For high detection rates the GPR performs less than at other test lanes leading to a lower performance of the best sensor method.

On the ferruginous test lane, the four fusion methods perform similar to the best sensor. The differences between the five methods are so small that it impossible to rank them.
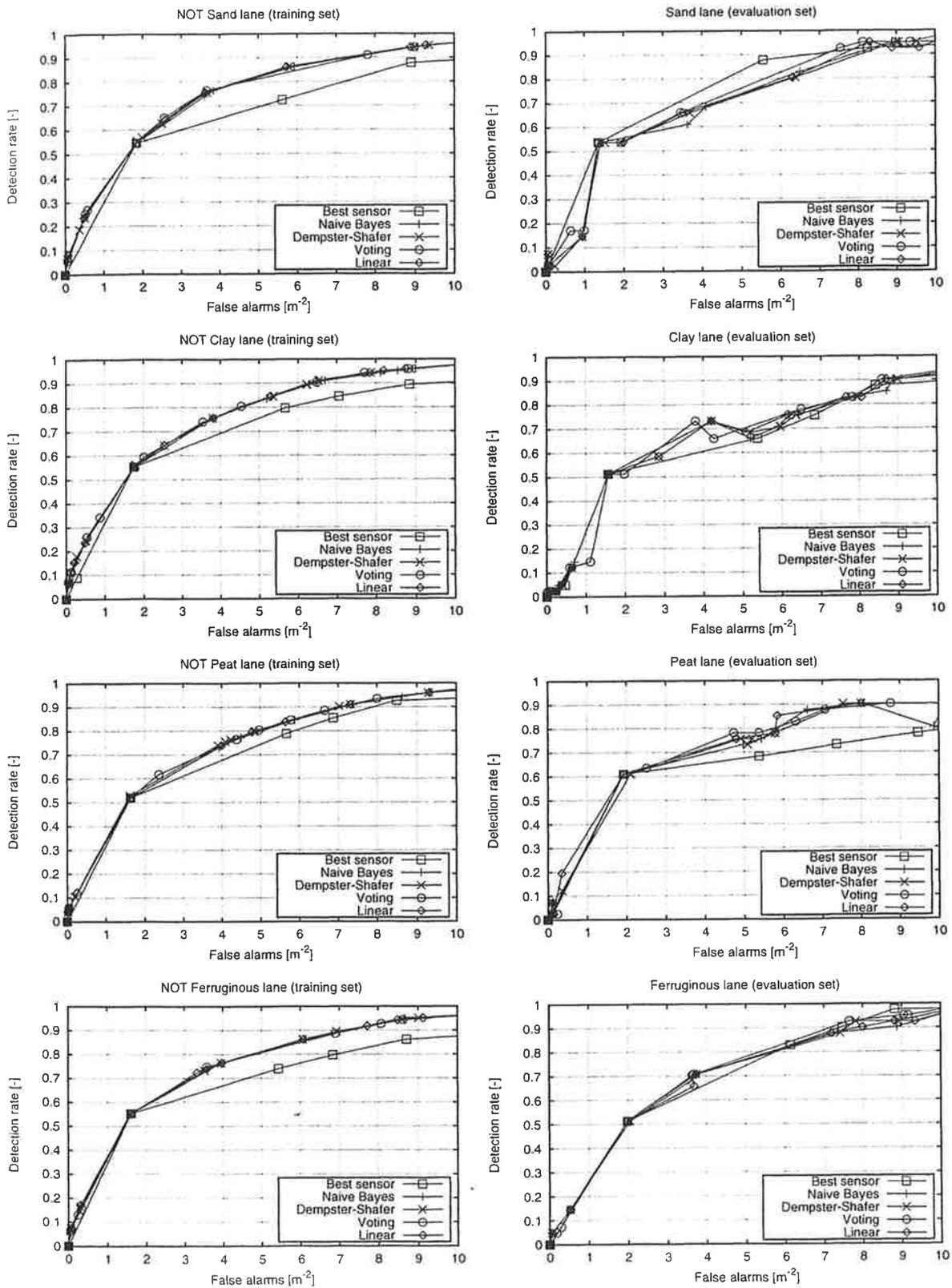
Figure 6. The training and evaluation set results for the four test lanes.

7. CONCLUSIONS AND DISCUSSION

The main conclusion is that the detection performance of the sensor fusion techniques is lower than we would expect from an operational automated multi-sensor mine-detection system. This is explained by the following:

- difficult scenario: there are relatively few easy to detect surface mines, metal mines and anti-tank mines,

- alignment: the alignment between the sensors does not seem to be optimal.

However, with regard to the sensor-fusion techniques, we conclude that on the training sets used here, the differences in performance between four of these techniques (naive Bayes, Dempster-Shafer, voting and linear discriminant) are practically non-existent. This means that no matter what technique is used in the optimisation procedure each technique will reach the nearly the same detection performance. This leads to the following question: Is there a general, four-parameter, fusion technique that has better performance on this data set than the performance seen in this paper? Since the performances of our four-parameter fusion techniques are so similar, this may be an indication that the performance is close to the optimal performance of a general four-parameter fusion technique.

For the evaluation set, the differences between the techniques are larger. Given the variation on the curves, it is still impossible to rank four of the fusion techniques: naive Bayes, Dempster-Shafer, voting and linear discriminant. The performance of these four sensor-fusion techniques is thus considered similar for the evaluation set. For a considerable number of cases, these four sensor-fusion techniques perform better than the best sensor technique.

The differences in performance between the training set and the evaluation set is lower than in our previous work.[3] This is due to the larger data set used and because there are more similar cases present in the training and evaluation sets. The only real difference between the training and evaluation set is the soil type. It is expected that the differences will be further reduced if the evaluation set contained samples of the same soil type as was trained on. In the ultimate case of an infinite large and equivalent training and evaluation set, the training and evaluation set results will be the same.

8. FUTURE WORK

The sensor-fusion results presented in Section 6 are not considered acceptable for an operational multi-sensor system for AP land-mine detection. Therefore, future work should focus on improvement of sensors and their processing and on sensor fusion. For example, improvement of sensors can be achieved by using different imaging techniques. One can think of polarisation of infrared and visual imagery[15] or the utilisation of an hyper-spectral imager.

Besides physical improvement of the sensors also the sensor processing should be altered to reach operational goals. On a theoretical level, we are, for instance, thinking of a more strict definition of a confidence level. The current definition does not have a statistical foundation. This is the only option when you have a small data set like GEODE. For larger data sets, more statistics are available and these should be used in the fusion techniques.

When considering sensor-fusion techniques there are also some options to explore. An obvious example is the use of spatial relationships between grid cells. It is clear that for most of the sensors there is correlation between the confidence levels of neighboring grid cells which could be used for further false-alarm reduction. This could be achieved by applying sensor-fusion on a local neighborhood of grid cells, much like two-dimensional filtering. Besides the applied decision-level sensor fusion, an alternative or extension could be feature-level fusion in which features (morphological, depth, etc.) of the mines measured by different sensors can form the input for a classifier.

The results of Section 6 show few differences in performance for the general four-parameter methods. The less strict definition of a confidence level and the allowed scaling of the input confidence levels makes the choice for a fusion function rather arbitrary. Sensor-fusion methods (or classifiers) which allow for more specialisation by using more free parameters, e.g. rule-based fusion and artificial neural networks, could possibly better represent the input training data.

ACKNOWLEDGMENTS

REFERENCES

1. 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.

2. F. Cremer, E. den Breejen, and K. Schutte, "Sensor data fusion for anti-personnel land-mine detection," in *Proceedings of EuroFusion98. International Conference on Data Fusion*, M. Bedworth and J. O'Brien, eds., pp. 55–60, (Great Malvern, UK), Oct. 1998.

3. F. Cremer, J. G. M. Schavemaker, E. den Breejen, and K. Schutte, "Detection of anti-personnel land-mines using sensor-fusion techniques," in *Proceedings of EuroFusion99. International Conference on Data Fusion*, T. Windeatt and J. O'Brien, eds., pp. 159–166, (Stratford Upon Avon, UK), Oct. 1999.

4. 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.

5. 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.

6. F. Provost and T. Fawcett, "Robust classification for imprecise environments," *Machine Learning* , 2000. To be published.

7. F. Provost and T. Fawcett, "Robust classification for imprecise environments," Tech. Rep. IS 99-011, Department of Information Systems, Stern School of Business, New York (US), 1999.

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

9. K. Fukunaga, *Introduction to statistical pattern recognition*, Academic press, Inc., Boston, USA, 2 ed., 1990.

10. A. P. Dempster, "Upper and lower probabilities induced by a multi-valued mapping," *Annals of Mathematical Statistics* **38**, pp. 325–339, 1967.

11. B. V. Dasarathy, "Decision fusion benefits assessment in a three-sensor suite framework," *SPIE Optical Engineering* **37**, pp. 354–369, Feb. 1998.

12. L. A. Klein, "A boolean algebra approach to multiple sensor voting fusion," *IEEE transactions on Aerospace and Electronic systems* **29**, pp. 317–327, Apr. 1998.

13. M. G. J. Breuers, P. B. W. Schwering, and S. P. van den Broek, "Sensor fusion algorithms for the detection of 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. 1160–1166, (Orlando (FL), USA), Apr. 1999.

14. P. B. W. Schwering, M. G. J. Breuers, and S. P. van den Broek, "Sensor fusion algorithms for the detection of land mines," pp. 365–371, (Quebec, Canada), Oct. 1998.

15. W. de Jong, F. Cremer, K. Schutte, and J. Storm, "Usage of polarisation features of landmines for improved automatic 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.

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