JST-TECH-MINE06-004

Test and Evaluation of Japanese GPR-EMI Dual Sensor Systems at Benkovac Test Site in Croatia

Jun Ishikawa and Mitsuru Kiyota
Japan Science and Technology Agency (JST), Japan

Nikola Pavković
Croatian Mine Action Centre - Center for Testing, Development and Training (HCR-CTRO), Croatia

Katsuhisa Furuta
Tokyo Denki University (TDU), Japan

TABLE OF CONTENTS

Test and Evaluation of Japanese GPR-EMI dual sensor systems at Benkovac Test Site in Croatia

Jun Ishikawa and Mitsuru Kiyota
Japan Science and Technology Agency (JST), Japan

Nikola Pavković
Croatian Mine Action Centre - Center for Testing, Development and Training (HCR-CTRO), Croatia

Katsuhisa Furuta
Tokyo Denki University (TDU), Japan

Abstract – This article presents an experimental design and evaluation results of trials that were carried out from 1 February to 9 March 2006 at Benkovac test site in Croatia. The objective of the Croatia-Japan joint trials is to confirm performance of dual sensor systems, which use both ground penetrating radar (GPR) and electromagnetic inductive (EMI) sensor, in comparison with existing EMI sensors, i.e., metal detectors (MDs) and to provide reliable data as a basis for future work. Increasing probability of detection (PD) and decreasing false alarm rate (FAR) will contribute to improve working efficiency in humanitarian demining. Therefore, by analyzing the data from which general principles can be established on the relative value of the different technologies, the trials aim at evaluating differences in performance between dual sensors and MDs, especially in terms of discrimination of landmines from metal fragments and extension of detectable range in the depth direction. Devices to be evaluated here are four prototypes of anti-personnel landmine detection systems developed under a project of the Japan Science and Technology Agency (JST), the supervising authority of which is the Ministry of Education, Culture, Sports, Science and Technology (MEXT). The prototypes that provide operators with subsurface images make no explicit alarm and final decision whether or not a shadow in the image is a real landmine is left to the operator. This is similar to the way that medical doctors find cancer by reading CT images. Since operators' pre-knowledge of the locations of buried targets significantly influences the test results in these kinds of systems, three test lanes have been designed to be suitable for blind tests. The results showed that the dual sensor systems improve PD for minimum-metal landmines such as a PMA-2 buried in mineralized soil and have a potential to discriminate landmines from metal fragments. On the other hand, it has been learned that reducing operation time is the most important problem to be solved for practical use.

1. Introduction

Japanese Research teams from universities and industries, which are funded by the Japan Science and Technology Agency (JST), have been developed the GPR+EMI dual sensor systems since October 2002 under the program of “Research and Development of Sensing Technology, Access and Control Technology to Support Humanitarian Demining of Anti-personnel Mines.” To evaluate the prototypes, a series of trials were carried out from 8 February to 11 March 2005 in Sakaide City, Japan[1][2]. The concept of the developed systems is to make no explicit alarm and to dedicate itself to provide operators with clear subsurface images (Figure 1). Therefore, decision-making using the subsurface images is entirely left to operators’ subjectivity. Since operators’ pre-knowledge of the locations of buried targets significantly influences the detection results for these kinds of systems, all the test lanes in Japan were designed to be suitable for blind tests. Evaluation results of the trial showed that probability of detection for targets in deeper levels than 10cm can be improved by combining GPR with an EMI sensor.

After the trials in Japan, the prototypes have been improved to be more robust, simple and cost-effective, and the next step of the project has been to take field tests to evaluate these features in Croatia, which is a well-experienced country in test and evaluation for humanitarian demining equipment. This article shows evaluation results of the Croatia-Japan joint test and evaluation for anti-personnel landmine detection systems using GPR+EMI dual sensors at the test site Benkovac of Croatian Mine Action Centre - Center for Testing, Development and Training (HCR-CTRO) in Croatia.

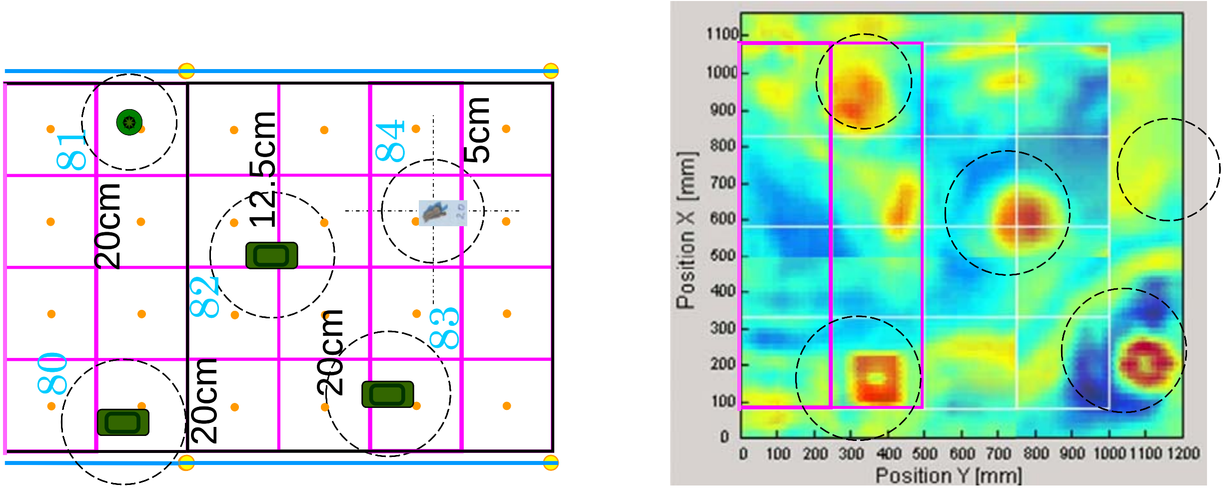


Figure 1. Examples of detection image acquired by a stepped-frequency SAR-GPR[6] mounted on MHV[5] during the Croatia trials. The left figure shows locations of targets with depth (one PMA-2s, three PMA-A1s and one metal fragment) and the right image is a wrapped image composed of several slices of different depth.

2. Test and Evaluation Overview

The objective of the test and evaluation is to confirm performance of GPR+EMI dual sensor systems in comparison with existing metal detectors (MDs) and to provide reliable data as a basis for future work. By using the data from which general principles can be established on the relative value of different equipment and techniques, the trial aims at clearing differences of performance between dual sensors and MDs, especially in terms of discriminating landmines from fragments and expanding detectable range in the depth direction. Improvement of the performance will contribute to increasing probability of detection (PD) and decreasing false alarm rate (FAR). The trial was conducted from 1 February to 9 March 2006 at the test site of Croatian Mine Action Centre - Center for Testing, Development and Training (HCR-CTRO) in Benkovac, Croatia (Table 1).

Table 1. Trial schedule: Character “B” and numbers "1, 3 and 7" show which test lane was used on that day (see Sections 2.2 and 2.3 for details of the test lanes and devices to be tested).

Trial Schedule in Polygon Benkovac		30-Jan							6-Feb							13-Feb							20-Feb							27-Feb							6-Mar						
		Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun
		30	31	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	1	2	3	4	5	6	7	8	9	10	11	12
Whole	Unpacking/Check			🟥	🟥	🟥																																					
Benchmarking Using MD	by Japanese Tester								7	1	3								🟧																								
	by Croatian Deminers																		🟧				🟩																				
	Mr. Benkovic (A)																		🟧					1	3/1	7																	
	Mr. Kukovec (B)																		🟧					3	1/7	B																	
Gryphon Hirose	Preparation								🟩	🟩	🟩	🟩							p																								
	Test												1	7		3	3	1	o	7	B																						
	Packing/Withdrawal																		h				🟥	🟥																			
MHV #2 Arai/Nonami	Preparation								🟩	🟩	🟩	🟩							s																								
	Test												7	3			1	7	k	1	3		B																				
	Packing/Withdrawal																		r				🟥	🟥																			
ALIS Sato	Unpacking/Check											🟥	🟥						o																								
	Preparation															🟩	🟩	🟩	W	🟩																							
	Test																		🟧		7		3	7	7/3	1				1	3	7	B										
MHV #1 Sato/Nonami	Preparation																		🟧				🟩	🟩	🟩	🟩																	
	Test																		🟧								1	7		3	7	X	1	3	B								
Whole	Packing/Withdrawal																																		🟥		🟥	🟥	🟥	🟥			

2.1. Test site Benkovac

The test site Benkovac is well-known to have been used in the International Test and Evaluation Programme (ITEP) project 2.1.1.2 “Reliability Model for Test and Evaluation of Metal Detectors[3]” in accordance with the CEN workshop agreement (CWA) 14747[4]. As written in [3], there are three types of soils available in the Benkovac test site, that is, (a) red bauxite with neutral stones, (b) red bauxite, and (c) neutral clay (Table 2). These 3 types of soils are referred respectively as (a) uncooperative and heterogeneous (lane 7), (b) uncooperative and homogeneous (lane 1), and (c) cooperative and homogeneous (lane 3) as shown in Figure 2. Regarding the weather, it was harsh, raining and snowing in the later half of the trials. Measurements of the soil moisture sometimes reached more than 40% (Figure 3).

Table 2. Soils in Benkovac cited from Table 4 on page 19 in [Meuller]: Ground reference height of a MD and susceptibility measurements.

Soil Types in Benkovac Trials	Ground Reference Height [cm]	Susceptibility at 958 Hz [10^{-5} SI]	Susceptibility difference at 465 and 4650 Hz [10^{-5} SI]
Red bauxite with neutral stones (uncooperative and heterogeneous)	19.7 ± 2.5	190 ± 36	35.4
Red bauxite (uncooperative and homogeneous)	18.8 ± 0.9	154 ± 13	25.5
Neutral clay (cooperative and homogeneous)	no signal	13 ± 2	0.6

(a) Uncooperative and heterogeneous (lane 7)

(b) Uncooperative and homogeneous (lane 1)

(c) Cooperative and homogeneous (lane 3)

Figure 2. Three types of soils at Benkovac test site

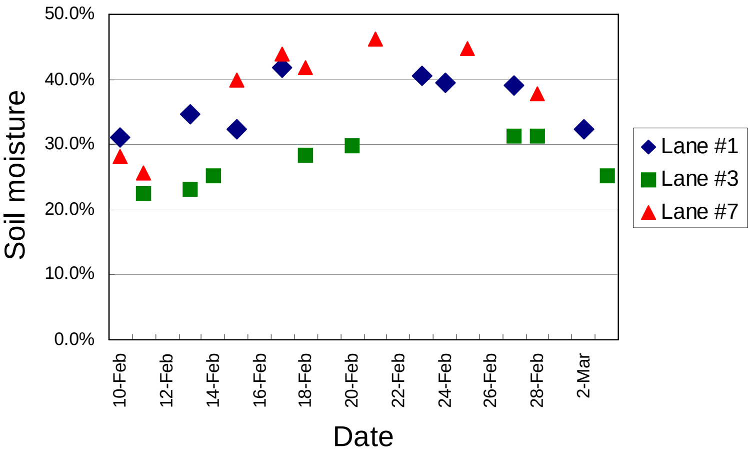


Figure 3. Soil moisture measurements through trials.

2.2. Four devices to be evaluated

Four sensor systems were evaluated in the trials. One of those is Mine Hunter Vehicle (MHV), the vehicle and manipulator part of which have been developed by a research team of Prof. Nonami, Chiba University[5]. MHV can interchangeably mount 2 GPR sensors in addition to a commercial-off-the-shelf MD. One is a stepped frequency SAR-GPR developed by Prof. Sato's team of Tohoku University[6] referred as **MHV#1** in the following part (Figure 4, left). Stepped frequency radar determines distance to a target by constructing a synthetic range profile, which is a time domain approximation derived from the frequency response of a combination of stepped frequency signals via inverse fast Fourier transform (IFFT). The major advantage of stepped frequency methods is that the spectrum bandwidth can be easily tuned to set the parameters to be optimum according to environment conditions such soil moisture. The other is an impulse GPR, LAMDAR-III, developed by Prof. Arai's project of University of Electro-Communications[7] referred as **MHV#2** in the following part (Figure 4, right). This kind of GPR operates by transmitting a very narrow pulse of electromagnetic wave (less than 1 nanosecond), the advantage of which is that the measurement time required to generate one range profile is very short.

The 3rd system to be evaluated is **Gryphon** (Figure 5, left), which can be remotely controlled to access to minefields. The robotic buggy has been developed by Prof. Hirose's team of Tokyo Institute of Technology[8]. The manipulator that is mounted on the buggy has been designed so as to cancel reaction force induced by sensor scanning. The sensor part of Gryphon is a GPR+EMI dual sensor named Advanced Landmine Imaging System (**ALIS**), which can be also used as a hand-held detector[9]. ALIS has been developed by the above mentioned Prof. Sato's team and took a field trial in Afghanistan in December 2004. The hand-held type ALIS is the 4th system to be tested here (Figure 5, right).

Figure 4. MHV#1 (left) and MHV#2 (right).

Figure 5. Gryphon (left) and ALIS (right).

3. Test and Evaluation Plan

3.1. Experimental design

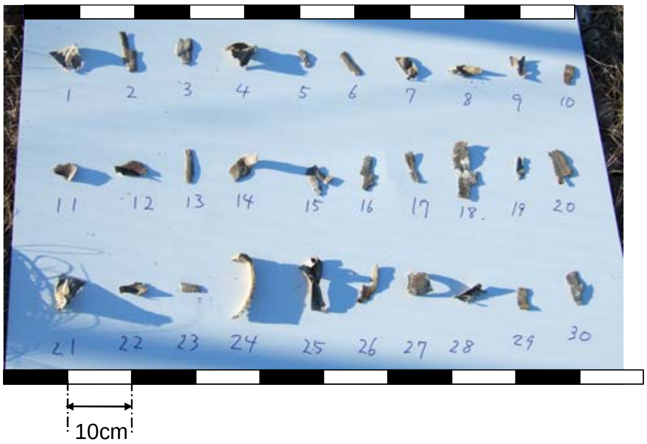
Through the trials, influences of 3 factors on probability of detection (PD) are evaluated by analysis of variance (ANOVA), that is, target types that consist of landmines and metal fragments, target depth and soil types as follows:

- Target type: PMA-1A, PMA-2, ITOP I_0 and Free-formed metal fragment (Figure 6),
- Target depth: 5.0cm, 12.5cm and 20.0cm (see Figure 7 for the definition), and
- Soil type: uncooperative and heterogeneous (Lane #7), uncooperative and homogeneous (Lane #1) and cooperative and homogeneous (Lane #3).

Due to the limitation of time for the trials and the number of landmines used, it is impossible to test all the combinations of levels (4 levels for target type, 3 levels of target depths and 3 levels of soil conditions). To impartially collect unbiased data for statistical analysis under this limitation, an orthogonal experimental design based on L_{18} ($2^1\times3^7$) described in Table 3 is used. Columns A and B of the L_{18} ($2^1\times3^7$) array in Table 3 are combined to generate a new 6-level column A via a multi-level method, and the levels 5 and 6 in the new column A are replaced the levels 1 and 2 respectively to reduce the number of levels from 6 to 4, resulting in Table 4. According to the modified L_{18} array in Table 4, factors “target type,” “target depth” and “soil condition” are respectively allocated to the columns A, B and C, and then a combination of levels in every factor is derived as depicted in Table 5. The number of target used in each level is 7. Note that that the number of experimental runs is reduced from 36 to 18. Accuracy of statistical estimation for the target types PMA-1A and PMA-2 (the levels 1 and 2 in the column A) is higher than that of ITOP I_0 and Free-formed metal fragment (the levels 3 and 4 in the column A) because the number of experimental runs for the levels 1 and 2 is twice as many as that for the levels 3 and 4.

Targets classified into soil types of “uncooperative and heterogeneous,” “uncooperative and homogeneous” and “cooperative and homogeneous” are buried at random locations respectively in Lanes 7, 1 and 3 at specified depths defined in Figure 7. Levels 1, 2 and 3 of Factor C in Table 4 are matched “uncooperative and heterogeneous,” “uncooperative and homogeneous” and “cooperative and homogeneous” in order, but the lane numbering 7, 1 and 3 originates in the lane number of the Benkovac test site (see Annex 1 for the lane layout). Eventually, burying targets was done on 8-9 December 2005 so that the targets could be left as it is for 2 months as well as targets in calibration areas which contain all the combinations of levels (Figure 8).

(a) PMA-1A (b)PMA-2 (c)ITOP I_0 (a 12.7-millimeter vertical aluminum tube)



(d) Metal fragments

Figure 6. Four kinds of target used in the trial.

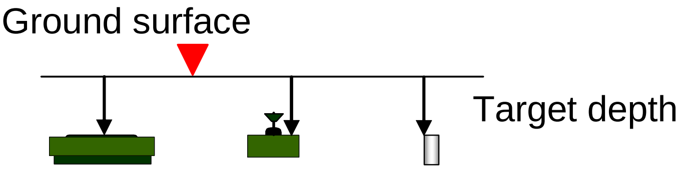


Figure 7. Definition of target depth: from left to right for PMA-1A, PMA-2 and both ITOP I_0 and fragment. Note that the depth of PMA-2 is measured form the ground surface to the top of the case, not the trigger.

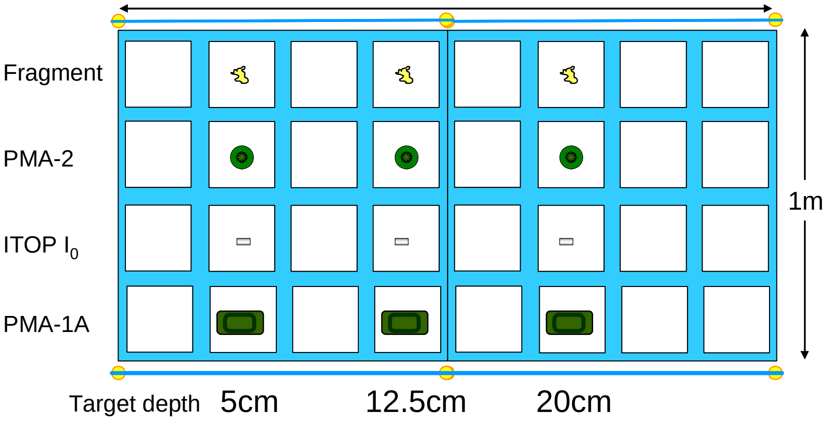


Figure 8. Calibration area.

Table 3. Original $L_{18}(2^1x3^7)$ orthogonal array.

No.	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
1	1	1	1	1	1	1	1	1
2	1	1	2	2	2	2	2	2
3	1	1	3	3	3	3	3	3
4	1	2	1	1	2	2	3	3
5	1	2	2	2	3	3	1	1
6	1	2	3	3	1	1	2	2
7	1	3	1	2	1	3	2	3
8	1	3	2	3	2	1	3	1
9	1	3	3	1	3	2	1	2
10	2	1	1	3	3	2	2	1
11	2	1	2	1	1	3	3	2
12	2	1	3	2	2	1	1	3
13	2	2	1	2	3	1	3	2
14	2	2	2	3	1	2	1	3
15	2	2	3	1	2	3	2	1
16	2	3	1	3	2	3	1	2
17	2	3	2	1	3	1	2	3
18	2	3	3	2	1	2	3	1

Table 4. Modified Orthogonal array using multi-level made a new column A from the columns A and B in Table 3.

No.	**A**	**B**	**C**	**D**	**E**	**F**	**G**
1	1	1	1	1	1	1	1
2	1	2	2	2	2	2	2
3	1	3	3	3	3	3	3
4	2	1	1	2	2	3	3
5	2	2	2	3	3	1	1
6	2	3	3	1	1	2	2
7	3	1	2	1	3	2	3
8	3	2	3	2	1	3	1
9	3	3	1	3	2	1	2
10	4	1	3	3	2	2	1
11	4	2	1	1	3	3	2
12	4	3	2	2	1	1	3
13	1	1	2	3	1	3	2
14	1	2	3	1	2	1	3
15	1	3	1	2	3	2	1
16	2	1	3	2	3	1	2
17	2	2	1	3	1	2	3
18	2	3	2	1	2	3	1

Table 5. Combination results of levels of each factor via $L_{18}(2^1x3^7)$ experimental design.

No.	Target type	Target depth	Lane # (Soil type)
1	PMA-1A	5.0cm	7
2	PMA-1A	12.5cm	1
3	PMA-1A	20.0cm	3
4	PMA-2	5.0cm	7
5	PMA-2	12.5cm	1
6	PMA-2	20.0cm	3
7	ITOP-I0	5.0cm	1
8	ITOP-I0	12.5cm	3
9	ITOP-I0	20.0cm	7
10	Fragment	5.0cm	3
11	Fragment	12.5cm	7
12	Fragment	20.0cm	1
13	PMA-1A	5.0cm	1
14	PMA-1A	12.5cm	3
15	PMA-1A	20.0cm	7
16	PMA-2	5.0cm	3
17	PMA-2	12.5cm	7
18	PMA-2	20.0cm	1

3.2. Trial procedures

Two testees of each system took blind tests of 3 lanes, i.e., #1, #3 and #7. All the testees declared detected anomalies by putting tags on the ground where the targets are considered to be buried as shown in Figure 9. As described in Table 6, the tags show confidence rating of the testee and the final decision whether the declared anomaly is a target (landmine/fragment) or clutter. The concrete trial procedures are as follows:

1. Before the test starts, the tester records volumetric water content that is measured using a time domain reflectometry (TDR) meter (Figure 10) at the location depicted in Figure 11.
2. The tester records the start time.
3. The testee does close-in detection using a sensor system cooperatively with vehicle operators.
4. After putting a transparent template on the ground, the testee declares all the detected anomalies every 1m×1m area by putting tags at the positions where the targets are considered to be buried. As explained in Table 6, the tags show confidence ratings of the testee and the final decision whether the declared anomaly is a target (landmine/fragment) or clutter. Target depth is also shown by another tag on which the depth is written by hand.
5. The tester takes a digital photograph to record the declaration.
6. After the declaration by the testee finishes, the tester records the end time and volumetric water content. The time required for close-in detection is calculated from the start and end times, subtracting break time (if any).

According to this procedure, 2 testees for every system independently were tested according to the schedule described in Table 1. Taking tests from Lanes 1 through 3, the testee finishes all the 18 experimental runs of the experimental design described in Table 5.

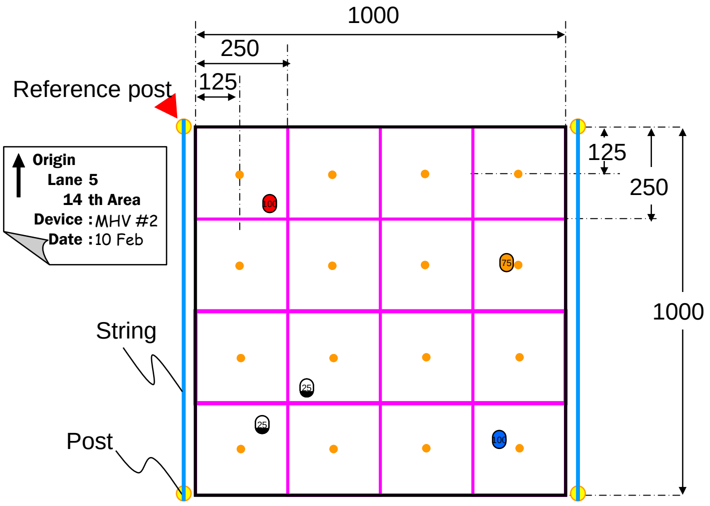


Figure 9. Declared position recording using a template and tags: Unit of length is millimeter.

Figure 10. Time Domain Reflectometry (TDR) meter for volumetic water content measurement.

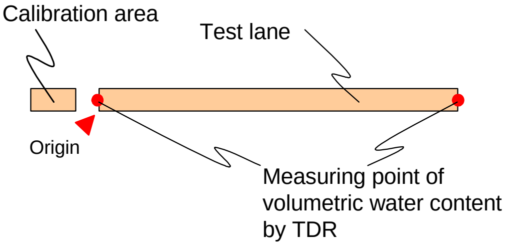


Figure 11. Measurement point of volumetic water content.

Table 6. Definition of confidence rating and tag that indicates declared location.

Definition of confidence rating	I'm 100% sure that there is nothing here.	It seems that there might be something here.	I'm almost sure that there is something here.	I would classify the detected object as a landmine or fragments.	I confidently classify the detected object as a landmine or fragments.
Final decision	**I declare that it is a clutter.**		**I declare that it is a landmine.**		
Confidence rating and tag color	N/A	25 Yellow	50 Pink	75 Orange	100 Red
Final decision	**I declare that it is a clutter.**		**I declare that it is a fragment.**		
Confidence rating and tag color	N/A	25 White/Black	50 White	75 Green	100 Blue

3.3. Evaluation method

Based on the photographs and hearing results acquired in the procedures explained in Section 3.2, the tester determines whether the declared locations can be considered to be from the intended targets according to the concept of halo radius depicted in [4]. Note that decoys that consist of 7 ITOP I_0 and 7 free-formed metal fragments for each lane are used as intended targets so that a discriminating ability between landmines and decoys can be evaluated.

False alarm rate (FAR) is evaluated based on the Poisson model, assuming that each sensor system being tested has its own fixed rate of clutter false alarms per unit area for the regions where neither landmine nor decoy is laid underground.

Eventually, the tester classifies the declared positions into 4 categories as described below.

- True positive (TP): the case that the testee declared it as a target (landmine/fragment) and this is true. In the case of fragment, only when detecting a decoy it is classified into TP (Figure 12, right).
- False positive (FP): the case that the testee declared it as a target (landmine/fragment) and this is not true. This is a false alarm due to clutter.
- True negative (TN): the case that the testee declared it as clutter and this is true.
- False negative (FN): the case that the testee declared it as clutter and this is not true. This is missing a target.

To compare the performance of GPR+EMI dual sensors with that of existing MDs, a benchmarking was conducted by two Croatian deminers, who do not know the target positions. The design of the experiment, the training of deminers and the monitoring has been organized by the Federal Institute for Materials Research and Testing (BAM). The deminers claimed that they cannot distinguish a metal fragment from a landmine only based on the audio signal of the MD. Therefore, only two levels of confidence were used in this trial. When the deminers heard an audio signal, they marked its location with 100% confidence if they would investigate that location with a prodder. If they would not investigate it, believing it comes from the soil or other source of noise, they marked the location with 25% confidence. In the following section, when results of Japanese testees are compared with those of Croatian deminers, tags with higher or equal to 50% confidence were counted as declaration, and the color of tags was not taken into consideration. For example, both a red tag on a metal fragment and a blue tag on a landmine were considered to be true positives, and this is hereinafter referred to as the normal detection criterion.

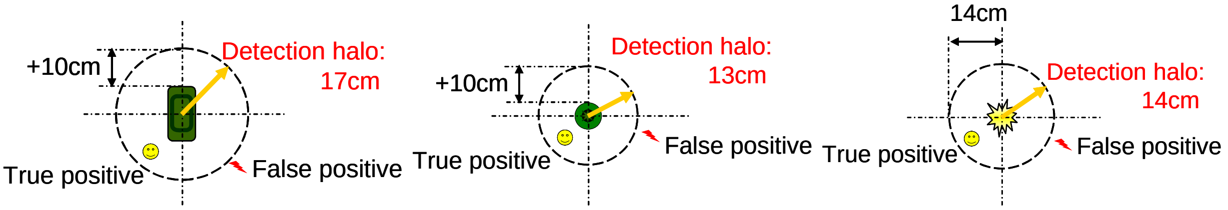


Figure 12. Detection criteria[4]: from left to right for PMA-1A, PMA-2 and both ITOP I_0 and fragment.

3.4. Use of Blind Test Lane

The Benkovac test site has 39 1m×47m blind test lanes where landmines were buried at 5-27cm deep more than 4years ago and have been left as it is to keep the natural minefield situation. On the day shown in the schedule of Table 1 marked by "B," each dual sensor takes tests using one of the 39 blind test lanes. The HCR-CTRO has chosen Lane #14 lane and will evaluate the detection results. The lane searched by the dual sensor was also searched the by a HCR deminer with a metal detector. Vegetation on the selected lane has been cleared due to the limited ability for the developed systems to adapt to rough terrain.

4. Experimental Results

According to the experimental design proposed above, data from ten testees (two each from every system) have been acquired. The comprehensive results of probability of detection (PD) and false alarm rate (FAR) are shown in Table A.1 in Annex 2. This section discusses how the data are analyzed.

4.1. Analysis of Variance (ANOVA)

Data of the two Croatian deminers and two Japanese testees who attained higher probability of detection (PD) than the other Japanese (i.e., one of ALIS and one of Gryphon) were separately analyzed by ANOVA to see effects of factors on PD. This is for confirming that each factor in an experimental design has been designed well enough to analyze the influences on PD.

Tables 7 and 8 show ANOVA results for Croatian deminers and Japanese testees, respectively, and Figures 13 and 14 show plots of factor effects with 95-percent confidence intervals. In Tables 7 and 8, factors, the null hypothesis of which has been rejected at the level of significance of 0.05/0.01, are indicated by * (0.05) /** (0.01). For those factors, there have been significant differences in PD between the levels, and it can be said that it is meaningful to discuss how those factors influence PD and that the test lanes were well-designed to evaluate the sensor systems. It has been shown that there is a strong dependence of PD on target depth. Regarding soil type (factor C), the ANOVA showed that there was no significant difference in PD among 3 kinds of soils. However, there has been a difference observed between Croatian and Japanese testees about how the soil type affects on PD. Japanese dual sensors has been less affected by uncooperative soil in Lane #1 than MDs. On the other hand, it was difficult for Japanese testees to find ITOP I0 because ITOP I0 had almost all no reflection for GPR. In the next section, PD are concretely discussed.

Table 7. Result of ANOVA for two Croatia deminers

Source of Variation	Degree of freedom	Sum of squares	Mean of squares	Computed F statistic
A: Target type	3	0.385	0.128	3.20 *
B: Target depth	2	0.821	0.410	10.23 **
C: Soil type	2	0.205	0.103	2.56
Error	28	1.124	0.040	-
Total	35	2.535		

Table 8. Result of ANOVA for two Japanese testees from ALIS and Gryphon.

Source of Variation	Degree of freedom	Sum of squares	Mean of squares	Computed F statistic
A: Target type	3	0.540	0.180	7.27 **
B: Target depth	2	0.926	0.463	18.70 **
C: Soil type	2	0.144	0.072	2.91
Error	28	0.693	0.025	-
Total	35	2.304		

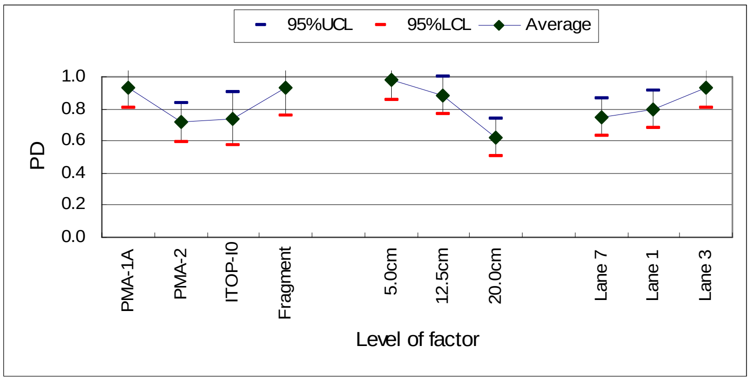


Figure 13. Effects of factors on Croatian deminers' PD.

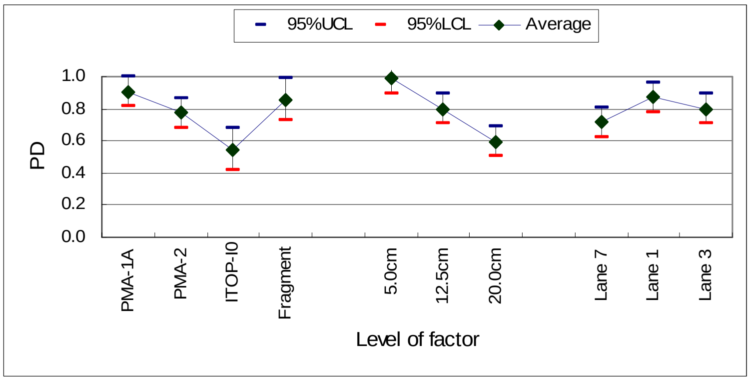


Figure 14. Effects of factors on the two Japanese testees.

4.2. Probability of Detection

Figure 15 shows probability of detection (PD) of 5 testees for 18 experimental runs with the normal detection criterion, where ITOP I_0 and metal fragments treated as targets intended to be detected. One testee from each system, who attained higher PD than the other, has been chosen. It has been shown that GPR+EMI systems attained higher PD than a deminer (Deminer 1) for deeply-buried PMA-2 in mineralized uncooperative soil. On the other hand, the deminer can very precisely determine the location of ITOP I_0, which is very small and has no recognizable shape by GPR.

A testee of ALIS (ALIS 1) attained 81.7% of average PD, which matches those of two deminers, who attained 84.0% and 81.0%. Figure 16 shows their average PD for each level of factor. Superiority of ALIS to deminers was observed in the levels of PMA-2 and Lane #1 (uncooperative and homogeneous soil). On the other hand, deminers' work is very fast and they took only about 5min for $1m^2$ detection while ALIS took 30-40min and the other 3 vehicle-mounted systems 15-20min.

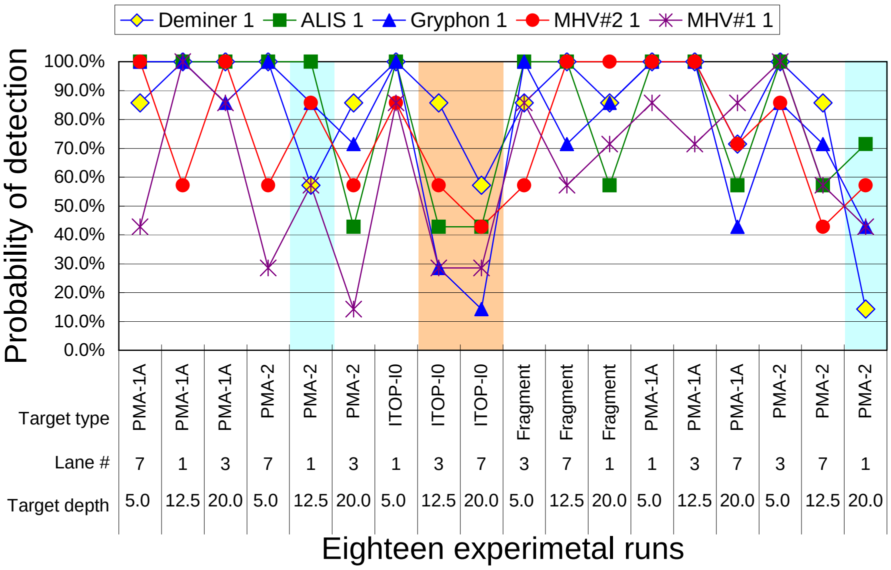


Figure 15. PD for 18 experimental runs: one testee from each system has been selected.

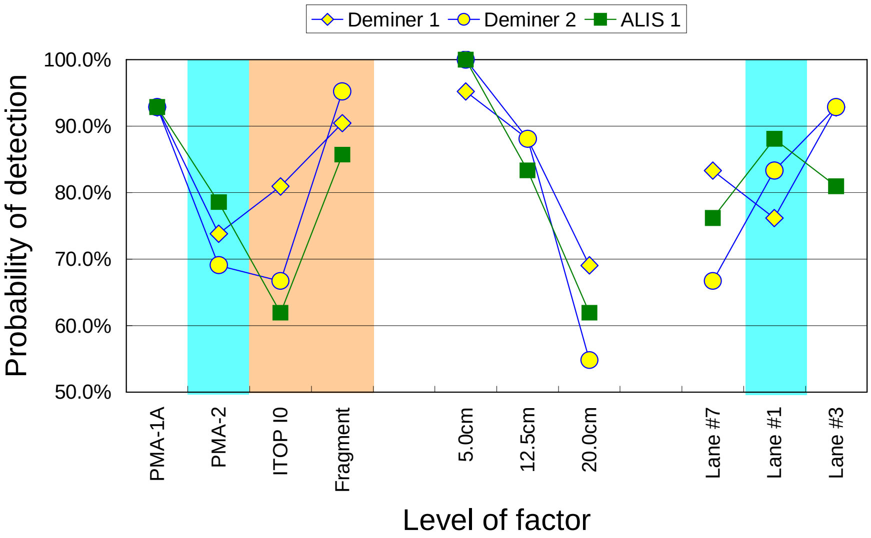


Figure 16. Average PD for each level of factors.

Data of ALIS have been analyzed by the colored tags explained in Table 6 to see if the dual sensor can tell landmines from metal fragment. In this case, declarations by wrong-colored tags were treated as false positives. For example, a declaration for a landmine with a blue tag means having missed the landmine. Lines with diamonds in Figure 17 show this strict criterion case for ALIS. Compared with the case where the PD was calculated by the normal detection criterion (line with squares in Figure 17) , it was shown that ALIS found 50% of metal fragments as metal fragments discriminating from landmines although a small degradation in PD for landmines was observed. This means that dual sensors have a possibility of discrimination of landmines from metal fragments.

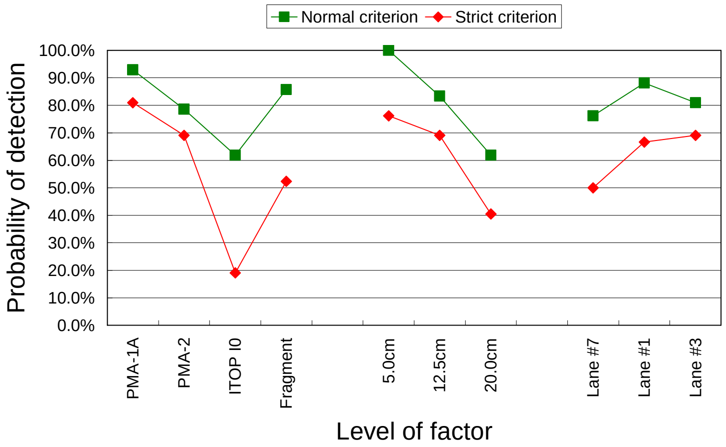


Figure 17. Comparison of average PD of ALIS to see an ability of metal fragment discrimination.

4.3. False Alarm Rate (FAR)

Figure 18 shows false alarm ratio (FAR) separately calculated from each soil type (i.e., each lane) for the same 5 testees as in Figure 15 with the normal detection criterion. Compared with the deminer's results, the GPR+EMI sensors tended to have less influence of soil type on FAR, especially for Gryphon and MHV#2.

The FAR in Figure 18 is essentially based on detection results reported by testees, which are classified into four categories: true positive, false positive, true negative and false negative. However, the classification based on a testee's discrimination threshold is a one-sided view, and the number of true positives and the number of false positives change as the threshold is varied. An ROC curve shows us the relationship between the true positive and false positive for a variety of different thresholds and is useful to see the qualification of sensors, taking into account tradeoff between PD and false alarm rate. Figures 19 (a) to (d) show ROC curves of 5 testees. It can be clearly seen that Deminer 1 has not swerved from his conviction and his FAR was affected by soil types. On the other hand, Japanese testees were indecisive, and further analysis should be needed to see if the indecision is due to the nature of testees or the system performance itself.

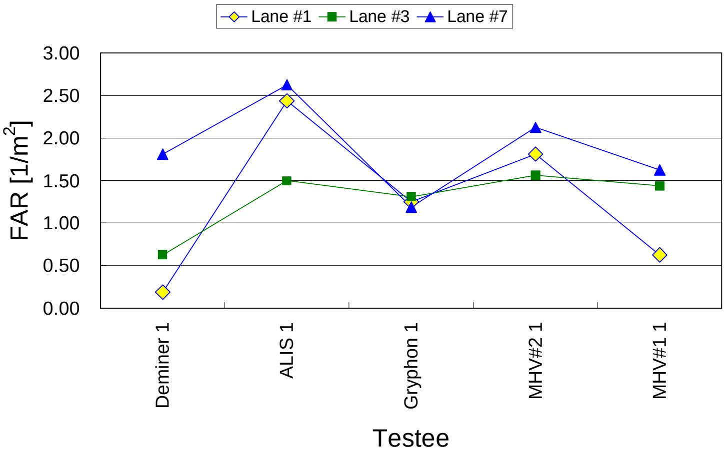
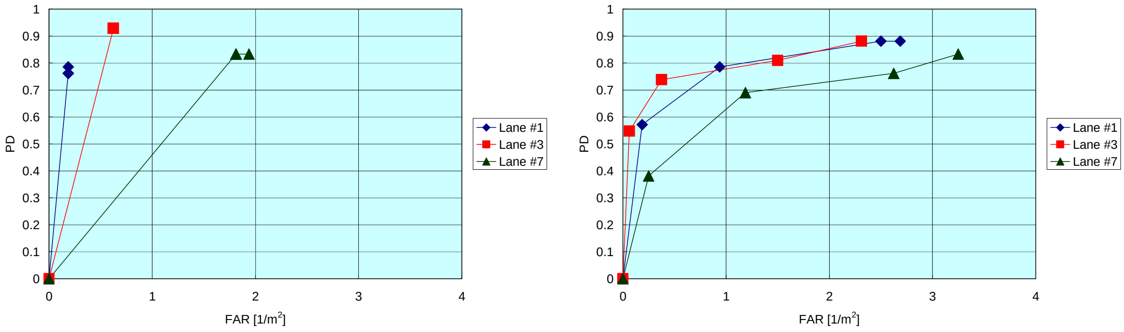
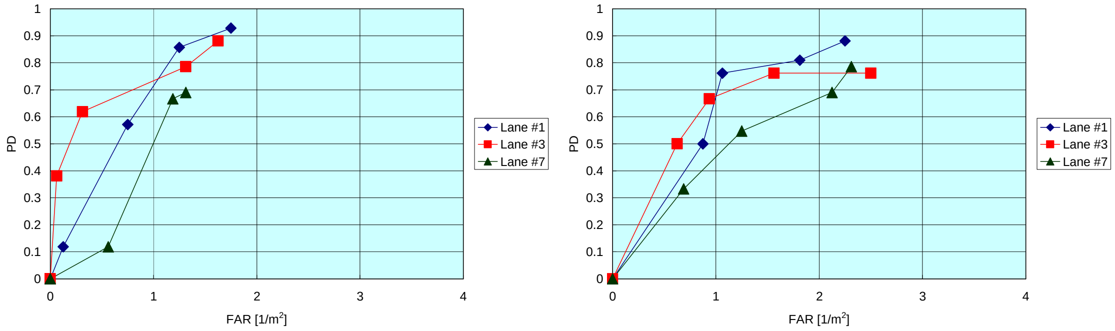


Figure 18. FAR for 5 testees: the same testee from each system has been selected.



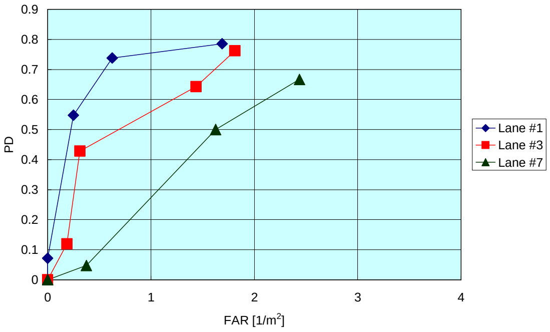
(a) Deminer 1

(b) ALIS 1



(C) Gryphon 1

(d) MHV#2 1



(e) MHV#1 1

Figure 19. ROC curves for 5 testees.

4.4. Human Factor Consideration

As discussed the former section, one of the most important future works in system evaluation is how quantitatively to evaluate human factors. For metal detectors (MDs), there are some methods to see the best performance of devices. For example, another benchmarking of MD has been conducted by a Japanese tester and a tester from HCR-CTRO, who know the exact positions of targets, checking if there is a MD response occurred just above every buried target. Figure 20 shows the result compared with two deminers' result, and the figure tells us the deminers have achieved almost all the same PD as the best performance. Thus, for MD evaluation, we can discuss the effect of human factors on PD by comparing their result by the best performance. However, for GRP+EMI dual sensor systems that are based on image analysis, so far, there is no effective method to see the best performance. Therefore, establishing a way to measure the best performance of dual sensor system is one of our future works.

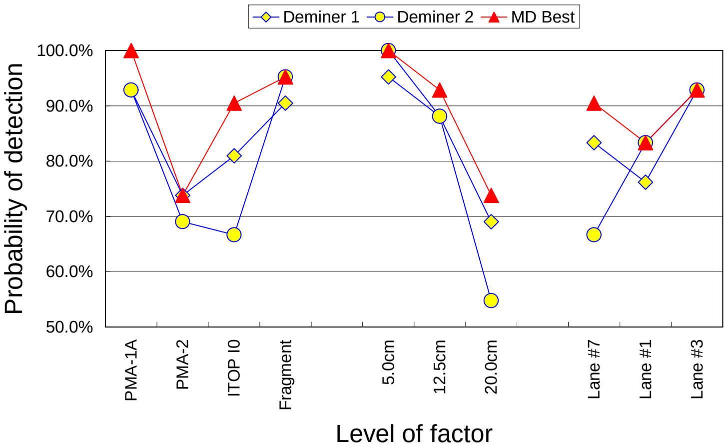


Figure 20. MD best performance compared with deminers' result

5. Comments on Evaluation Results from CROMAC-CTDT

As a coordinator of the trial, CROMAC-CTDT, the director of which is a co-author of this report, has made comments regarding evaluation results as follows:

The objective of the trial was to confirm dual sensor system performances, that is, increasing the probability of detection at depths deeper than 10 cm and decreasing the false alarm rate, and to elaborate reliable dada as a base for future R&D activities. Through the trial it has been shown that

a) in relation to metal detectors, the tested systems have proved the advantages in the increase of detection depth and in the decrease of false alarm rate,

b) the Japanese dual sensor systems can be operationally used at the present time although there are some pre-actions required and limitations such as heavy vegetation,

c) by combining use of demining machines and GPR systems, it would be possible to increase the safety and the process of demining itself, and

d) CROMAC-CTDT can offer services based on the capacities and knowledge to all interested parties for all types of tests and trials and for elaboration of such necessary documentations as a standard operating procedure (SOP).

Consequently, possible applications of dual sensors in humanitarian demining could be as the second method in demining following demining machine, as the second method in technical survey after mechanical demining, which clears vegetation, quality assurance during demining and quality control of demining. Potential system users will be demining organizations, NGO, police, army, and national mine action centers.

6. Conclusions

Through the trials, many lessons have been learned such that PD for small targets in mineralized (uncooperative) soil can be improved by using GPR. The results showed that the dual sensor systems improve PD for minimum-metal landmines such as a PMA-2 buried in mineralized soil and that the systems have a potential to discriminate landmines from metal fragments. On the other hand, it has been learned that reducing operation time is the most important problem to be solved for practical use as well as evaluation of human factors.

These results were fed back to the testees for further improvement. The next step is to put promising systems into practical field trials that will be conducted by a third party after the modification.

ACKNOWLEDGEMENT

The authors would like to thank all the project members especially for the trial participants from Tohoku University, Chiba University, Tokyo Institute of Technology, University of Electro-Communications, Fuji Heavy Industries Ltd., TAU GIKEN Co. Ltd., Yamate Corporation, HiBot Corporation.

The authors also would like to thank Dr. Christina Müler and her colleagues of the Federal Institute for Materials Research and Testing (BAM) for their comments in statistical analysis and coordination of blind-tests by deminers.

REFERENCES

[1] Ishikawa, J., Kiyota, M. and Furuta, K., “Experimental design for test and evaluation of anti-personnel landmine detection based on vehicle-mounted GPR systems,” *Proc. of SPIE Vol. 5794, Detection and Remediation Technologies for Mines and Minelike Targets* X, pp. 929-940, 2005.

[2] Ishikawa, J., Kiyota, M. and Furuta, K., “Evaluation of Test Results of GPR-based Anti-personnel Landmine Detection Systems Mounted on Robotic Vehicles,” *Proc. of the IARP International workshop on Robotics and Mechanical Assistance in Humanitarian Demining* (*HUDEM2005*), pp. 39-44, 2005.

[3] Mueller, C., et al., “Reliability Model for Test and Evaluation of Metal Detectors,” *ITEP Project 2.1.1.2 Final Report*, http://www.itep.ws/pdf/Itep_no2.1.1.2_reliability_model.pdf, 2004.

[4] CEN Workshop Agreement, *Humanitarian Mine Action - Test and Evaluation- Metal Detectors*, CWA 14747, http://www.itep.ws/pdf/CWA_metal_detectors.pdf, 2003.

[5] Nonami, K. and Aoyama, H., “Research and Development of Mine Hunter Vehicle for Humanitarian Demining,” *Proc. of the HUDEM2005*, pp.76-81, 2005.

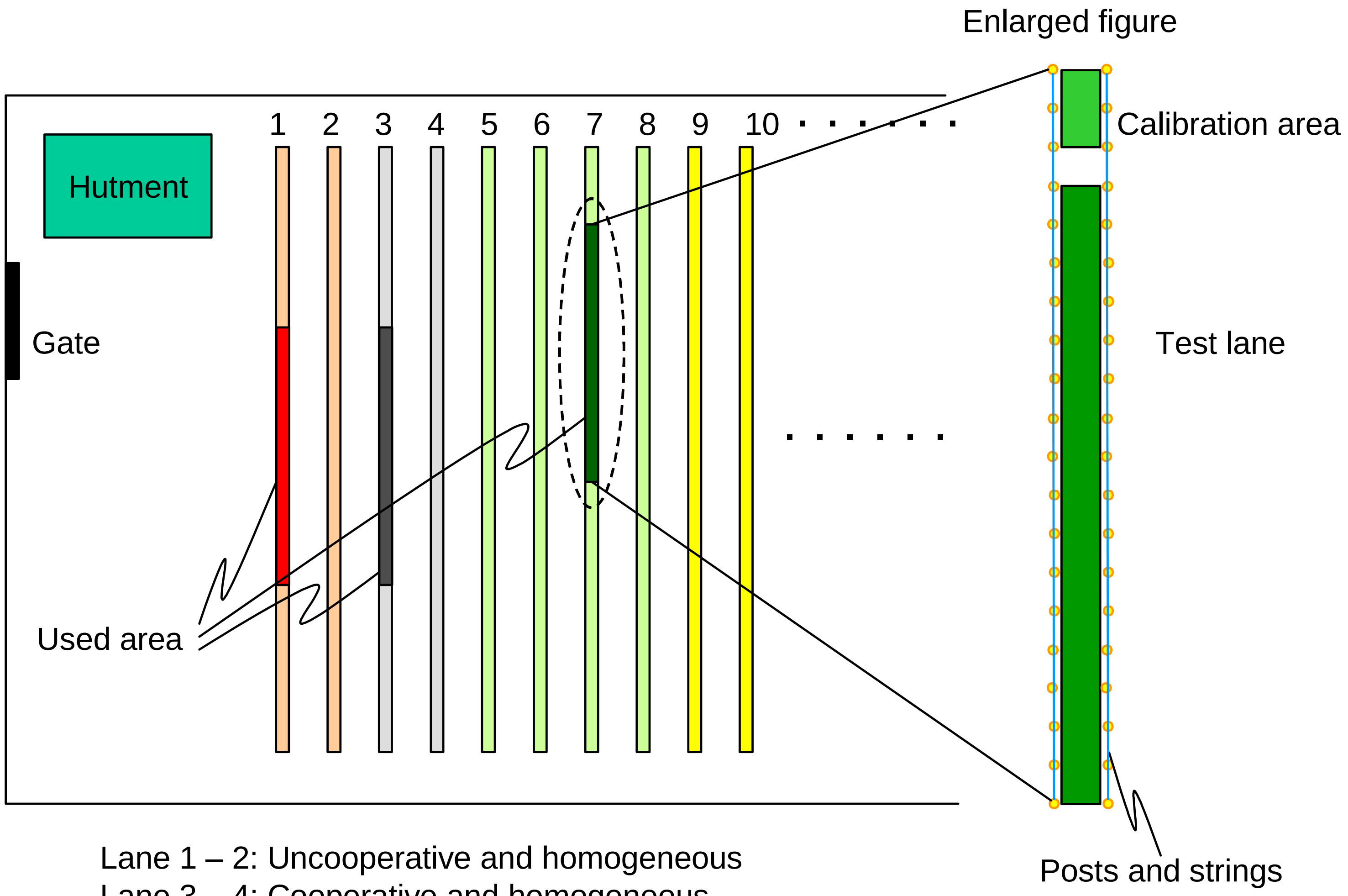
[6] Sato, M., Hamada, Y., Feng, X., Kong, F., Zeng, Z. and Fang, G., “GPR using an array antenna for landmine detection,” *Near Surface Geophysics*, Vol. 2, pp. 3-9, February, 2004.

[7] Shrestha, S. M. and Arai, I., “High Resolution Image Reconstruction by GPR using MUSIC and SAR Processing Method for Landmine Detection,” *Proc. of the 2003 IEEE International Geoscience and Remote Sensing Symposium* (*IGRASS2003*), pp. 505-508, 2003.

[8] Fukushima, E. F., et al., “Teleoperated Buggy Vehicle and Weight Balanced Arm for Mechanization of Mine Detection and Clearance Tasks,” *Proc. of the HUDEM2005*, pp. 58-63, 2005.

[9] Sato, M., Fujiwara, J., Feng, X., Zhou, Z. and Kobayashi, T., “Development of a hand-held GPR MD sensor system (ALIS),” Proc. of SPIE Vol. 5794, Detection and Remediation Technologies for Mines and Minelike Targets X, pp. 1000-1007, 2005.

Annex 1. Lane layout of the test site Benkovac



Lane 1 – 2: Uncooperative and homogeneous
Lane 3 – 4: Cooperative and homogeneous
Lane 5 – 8: Uncooperative and heterogeneous
Lane 9 – : Blind test lanes

Annex 2. Comprehensive results of probability of detection (PD) and false alarm rate (FAR)

Table. A.1. Probability of detection (PD) and false alarm rate (FAR) with the normal detection criterion of ten testees.

No.	A: Target type	B: Target depth	C: Soil type: Lane #	Target Quantity	MD Best	Deminer 1	Deminer 2	ALIS 1	ALIS 2	Gryphon 1	Gryphon 2	MHV#2 1	MHV#2 2	MHV#1 1	MHV#1 2
1	PMA-1A	5.0cm	7	7	100.0%	85.7%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	85.7%	42.9%	100.0%
2	PMA-1A	12.5cm	1	7	100.0%	100.0%	100.0%	100.0%	71.4%	100.0%	85.7%	57.1%	71.4%	100.0%	85.7%
3	PMA-1A	20.0cm	3	7	100.0%	100.0%	100.0%	100.0%	85.7%	85.7%	71.4%	100.0%	100.0%	85.7%	42.9%
4	PMA-2	5.0cm	7	7	100.0%	100.0%	100.0%	100.0%	57.1%	100.0%	42.9%	57.1%	85.7%	28.6%	100.0%
5	PMA-2	12.5cm	1	7	85.7%	57.1%	85.7%	100.0%	42.9%	85.7%	71.4%	85.7%	71.4%	57.1%	42.9%
6	PMA-2	20.0cm	3	7	57.1%	85.7%	57.1%	42.9%	28.6%	71.4%	14.3%	57.1%	57.1%	14.3%	14.3%
7	ITOP-I0	5.0cm	1	7	100.0%	100.0%	100.0%	100.0%	42.9%	100.0%	100.0%	85.7%	85.7%	85.7%	57.1%
8	ITOP-I0	12.5cm	3	7	100.0%	85.7%	100.0%	42.9%	28.6%	28.6%	42.9%	57.1%	85.7%	28.6%	28.6%
9	ITOP-I0	20.0cm	7	7	71.4%	57.1%	0.0%	42.9%	28.6%	14.3%	28.6%	42.9%	28.6%	28.6%	14.3%
10	Fragment	5.0cm	3	7	100.0%	85.7%	100.0%	100.0%	71.4%	100.0%	100.0%	57.1%	57.1%	85.7%	85.7%
11	Fragment	12.5cm	7	7	100.0%	100.0%	85.7%	100.0%	57.1%	71.4%	71.4%	100.0%	57.1%	57.1%	85.7%
12	Fragment	20.0cm	1	7	85.7%	85.7%	100.0%	57.1%	42.9%	85.7%	28.6%	100.0%	85.7%	71.4%	71.4%
13	PMA-1A	5.0cm	1	7	100.0%	100.0%	100.0%	100.0%	85.7%	100.0%	100.0%	100.0%	85.7%	85.7%	85.7%
14	PMA-1A	12.5cm	3	7	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	85.7%	100.0%	85.7%	71.4%	100.0%
15	PMA-1A	20.0cm	7	7	100.0%	71.4%	57.1%	57.1%	57.1%	42.9%	42.9%	71.4%	57.1%	85.7%	71.4%
16	PMA-2	5.0cm	3	7	100.0%	100.0%	100.0%	100.0%	42.9%	85.7%	85.7%	85.7%	85.7%	100.0%	57.1%
17	PMA-2	12.5cm	7	7	71.4%	85.7%	57.1%	57.1%	57.1%	71.4%	14.3%	42.9%	42.9%	57.1%	42.9%
18	PMA-2	20.0cm	1	7	28.6%	14.3%	14.3%	71.4%	0.0%	42.9%	28.6%	57.1%	42.9%	42.9%	14.3%
				Average	**88.9%**	**84.1%**	**81.0%**	**81.7%**	**55.6%**	**77.0%**	**61.9%**	**75.4%**	**70.6%**	**62.7%**	**61.1%**

FAR [1/m^2]

Lane #1			0.19	0.44	2.50	2.31	1.25	0.81	1.81	2.75	0.63	1.63
Lane #3			0.63	0.13	1.50	2.19	1.31	1.31	1.56	3.31	1.44	1.94
Lane #7			1.81	0.69	2.63	3.00	1.19	0.81	2.13	2.69	1.63	1.25
	Average		**0.88**	**0.42**	**2.21**	**2.50**	**1.25**	**0.98**	**1.83**	**2.92**	**1.23**	**1.60**