Evaluating a Pattern-Based Visual Support Approach for Humanitarian Landmine Clearance

Lahiru G. Jayatilaka[*], **Luca F. Bertuccelli**[†], **James Staszewski**[‡], and **Krzysztof Z. Gajos**[*]

[*]Harvard University
Cambridge, MA 02138, USA
{lahiru, kgajos}@seas.harvard.edu

[†]MIT
Cambridge, MA 02139, USA
lucab@mit.edu

[‡]Carnegie-Mellon University
Pittsburgh, PA 15213, USA
jjs@cmu.edu

ABSTRACT

Unexploded landmines have severe post-conflict humanitarian repercussions: landmines cost lives, limbs and land. For deminers engaged in humanitarian landmine clearance, metal detectors remain the primary detection tool as more sophisticated technologies fail to get adopted due to restrictive cost, low reliability, and limited robustness. Metal detectors are, however, of limited effectiveness, as modern landmines contain only minimal amounts of metal, making them difficult to distinguish from the ubiquitous but harmless metallic clutter littering post-combat areas. We seek to improve the safety and efficiency of the demining process by developing support tools that will enable deminers to make better decisions using feedback from existing metal detectors. To this end, in this paper we propose and evaluate a novel, pattern-based visual support approach inspired by the documented strategies employed by expert deminers. In our laboratory study, participants provided with a prototype of our support tool were 80% less likely to mistake a mine for harmless clutter. A follow-up study demonstrates the potential of our pattern-based approach to enable peer decision-making support during landmine clearance. Lastly, we identify several design opportunities for further improving deminers' decision making capabilities.

Author Keywords

humanitarian landmine clearance, decision support, visual support, demining

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces

General Terms

Human Factors

INTRODUCTION AND MOTIVATION

A landmine is a passive explosive device hidden beneath the topsoil. During armed conflict, it is used to deny access to military positions or strategic resources. Unexploded landmines can persist for decades after a conflict's end, indiscriminately killing and maiming innocent civilians. As of 2010, 66 internationally recognized states were confirmed or suspected to be affected by mines, and in 2009 alone, landmines were responsible for over 1,300 casualties [17].[1] Beyond this loss of life, post-conflict societies must face the challenges presented by survivors of landmine incidents; these individuals may require amputation of limbs, suffer a loss of hearing or vision, and experience post-traumatic stress disorder [26, 16]. The number of affected individuals is significant—a 2009 report states that Cambodia alone has over 40,000 explosive remnants of war and mine survivors [13]. Also, significant resources are needed by each survivor. For example, a child with a lower limb amputation may need up to 15–20 artificial limbs in a lifetime [26]. However, the extensive resources required for adequate rehabilitation and reintegration of victims are often in limited supply in the developing countries where mine incidents predominantly occur [13, 16]. In addition to causing these human casualties, landmines severely inhibit the economic growth of developing countries by rendering vast tracts of farmland useless, injuring livestock, and reducing the size of working populations [1, 16].

Humanitarian demining aims to mitigate these problems by removing the threat of landmines and explosive remnants of war from civilian life [10, 11]. A crucial component of the humanitarian demining process is the physical clearance of landmines, which typically involves a human operator using audio feedback from a metal detector to find mines and employing hand-held digging tools to unearth them. Presently, mistakes in probing for suspected targets are among the main causes of accidental detonation, while operator inability to distinguish low-metal landmines from harmless clutter such as cans or shrapnel is the main source of inefficiency [11, 16]. There is a clear need for improvements to the mine clearance practice [6], yet more advanced detection and neutralization technologies designed to increase safety and efficiency (e.g., [8, 12, 4, 5]) have failed to be widely adopted due to their cost, complexity, low reliability, bulk, and high power requirements [25, 16, 11].

An alternative approach for improving safety and efficiency is to focus on improving the decision-making abilities of

---

[1]Casualty statistics are based on incomplete information and are strict underestimates

deminers operating existing technology. A recent study [22] revealed that expert deminers employ a sweeping strategy that allows them to mentally visualize the shape and size of a *metallic signature*—an area where a metal detector responds with acoustic feedback to the presence of a buried metallic object. By referring to previously encountered signatures, the expert deminer can more accurately determine if the buried object under investigation is a landmine or a piece of harmless clutter. Explicitly teaching this strategy to novice deminers has been shown to significantly improve their ability to correctly classify buried objects [22].

Informed by these findings, we hypothesized that the ability to *explicitly* visualize the edge points of metallic signatures as visual patterns on a computer screen would further improve the performance of novice deminers. To evaluate this hypothesis, we conducted a laboratory evaluation of a *pattern-based* visual support approach that enables deminers to build patterns on a display screen by serially collecting boundary points belonging to metallic signatures. In a study with 44 participants, we found that our visual support approach resulted in an 80% reduction in the number of missed mines and it reduced mine localization errors by 30%. In a follow-up study, we presented the visual patterns collected in the first experiment to a new set of 13 evaluators. The results suggest that individuals not directly involved in the demining process can successfully interpret the patterns, implying that our approach may allow deminers to benefit from peer support. Lastly, results from both studies contributed several design insights for further improving deminers' decision making performance. We intend to use the results of these studies to develop a cheap, portable, and robust visual support device that can be deployed as an add-on to existing metal detectors and that can contribute to improving the safety and efficiency of humanitarian landmine clearance.

In the remainder of this paper, we provide further pertinent background on humanitarian mine clearance and briefly synthesize previous work in the area. We then describe our main experiment and the follow-up study, and conclude with our interpretation of the findings.

BACKGROUND ON HUMANITARIAN DEMINING
Humanitarian mine clearance procedures require the identification and removal of all mine and other explosive hazards from a given area to a specified depth [11]. In practice, mine clearance takes place in environments as diverse as the deserts of Egypt, the mountains of Croatia and the tropical forests of Cambodia, with most programs operating on limited resources [16]. Given the procedural demands and the practical realities, mine clearance technology must be cheap, reliable and robust. Because of the difficulty of deploying technology within these constraints, a human with a metal detector has remained the primary method of mine clearance since the 1950s [6, 9]. Even though safer methods such as machine [4] or animal clearance [5] are available, their use is not widespread because machines are expensive to maintain and are constrained by terrain, while animals are difficult to train and are practically suited only for specific clearance scenarios [11].

Humanitarian deminers are usually recruited from local populations and trained over a two to four week period [7]. These deminers work according to strict operating procedures, which aim to preserve their safety and promote the efficacy of the mine clearance process [11, 7]. When using a metal detector, the primary tasks of a deminer are to carefully segment the ground in a marked lane with a meter-long stick, cut vegetation to clear the ground for detection, sweep with a detector, and investigate the ground carefully and methodically using a prodder or excavator [7, 16].

As of 2005, the average clearance rate for a deminer was 15–20 square meters a day [7]. Factors such as vegetation, topography, dehydration, monotony, and protective equipment have all been identified to impact performance of deminers. However, as mentioned previously, metallic clutter is the biggest source of inefficiency: the increased use of plastics in modern landmines has necessitated metal detectors that are highly sensitive, which has increased the rate of false alarms from non-hazardous metallic items [11, 16]. Demining organizations have little choice but to stipulate that deminers treat every signal as a threat, since there is presently no reliable way to make the mine versus clutter distinction based on audio feedback [16, 8, 7]. This strategy consumes a significant amount of time and resources [16]. However, advances in detection technology have demonstrated that procedural changes are possible in order to improve efficiency. In evaluation studies of the PSS-14 dual sensor detector, which combines a metal detector with a ground-penetrating radar (GPR), deminers used rapid clutter excavation techniques whenever they had compelling evidence that the buried object was not a mine. This enhanced efficiency of landmine clearance without detracting from deminer safety [2]. Despite its potential to improve the demining process, the PSS-14 remains unavailable to most organizations due to its high cost [11].

PREVIOUS WORK
The pattern-based visual support approach we propose and evaluate in this work is informed by findings on expert deminer performance by Staszewski and colleagues [22]. Their work shows that expert deminers systematically sweep the area near a potential threat to find and remember *edge points* of the metallic signature. That is, they systematically find and remember points where the metal detector response starts and stops. These experts then refer to past experiences to draw inferences from the spatial pattern outlined by the edge points held in their "mind's eye". These behavioral findings have been incorporated into US military deminer training programs [3] and have been shown to improve novice performance with not only metal detectors but also dual sensor detectors such as the PSS-14 described above [22]. With the PSS-14, inferences made from metallic signatures are used to guide investigation with the GPR. Our work is driven by the hypothesis that augmenting deminers' internal representations of metallic signatures with explicit visualizations of the edge point patterns will further increase the effectiveness of this procedure, especially for novice deminers.

Several visual feedback systems for metal detector–based demining are in development, but we are not aware of any published results contributing to our understanding of how these systems impact deminers' decision making capabilities. Kruger and Ewald have developed a laboratory setup that uses an ultrasonic positioning system and detector-mounted palmtop computer to present detector feedback as 2D intensity-graded regions [15]. Tohoku University has developed and field-tested the Advance Landmine Detection System (ALIS) which is an add-on system to a metal detector [20]. ALIS consists of a GPR sensing unit strapped to the operator's back, a palmtop computer (for computation and display) that is extended over the deminer's shoulder, and a camera attached to the detector handle for positioning. Metal detector feedback is presented as 2D intensity-graded regions, while the GPR signal is visualized as intensity fields for specific depths. Field evaluations have suggested that ALIS has potential for adoption, but to the best of our knowledge no data is available on how this system affects deminers' performance.

The most salient distinction between prior work and our approach is that we propose a discrete visualization that emphasizes information about the shape and size of the electromagnetic induction pattern induced by the buried object or objects instead of visualizing continuous signal intensity information at each location. As previously explained, our approach visualizes the information that expert deminers have been shown to rely on to make classification decisions. We also note that our approach is complementary to the graded approaches used by others and future work may demonstrate that a combination of the two is appropriate in some situations.

Visual support tools have also previously been used to support the training of deminers. The Sweep Monitoring System (SMS) [19, 14] uses remotely mounted stereo cameras to visually track the deminer's detector head in a simulated minefield to provide real-time auditory and visual feedback. Audio messages like "too fast" or "too slow" are provided to the trainee deminer, while visualizations of sweep coverage, detector speed and detector height are presented to the training supervisor. SMS is currently in use at U.S. army training centers, deployed after formative laboratory evaluations that demonstrated that visual feedback on sweeping performance improved the performance of demining trainees [14].

INTENDED DEVICE
The experiments presented in this paper are motivated by the desire to design and build a visual decision support device that can be used to improve the current humanitarian landmine clearance practice. To address the practical requirements of cost, portability, versatility, robustness, and power requirements, we envision that our solution will be centered around a commercial mobile device. The device will be engineered so that it can be mounted on the handles of most existing metal detectors. Its screen will present visual feedback to the deminer, while its built-in camera will be used to track the detector head's position with respect to the ground (an approach that is feasible with existing technology [18]).

Several commercially-available mobile devices are marketed in "ruggedized" versions which function robustly in adverse environmental conditions.

We note that the metallic signatures edge points could be visualized using other, non-electronic, approaches. Indeed, in training, edge points are often marked with physical markers, such as poker chips. However, such marking is physically cumbersome and is time consuming for the operator, and thus necessitates the assistance of a trainer or training buddy. This adds to training costs. For these reasons and because of the safety concerns, this approach is not useful in actual demining. Another possibility is to use spray paint to mark the edge points. A prototype of such a device has even been designed [24]. Compared to our approach, the paint-based approach requires a more time-consuming and more permanent modification to the metal detector, and will probably result in lower precision of the edge point placement. It also lacks a mechanism to completely erase markings, which is a concern for training facilities. The cost advantage of the paint-based solution is also unlikely to be substantial, while the electronic approach will support novel functionalities such as the storage and retrieval of previously-recorded patterns.

EXPERIMENT 1
In this section, we describe a laboratory experiment to evaluate our pattern-based visual support approach. Our participants, who were extensively trained to behave like novice deminers, carried out simulated mine detection tasks using a hobby metal detector. All of our participants used the method developed by Staszewski and colleagues [22] (described earlier) to locate edge points of the metallic signatures belonging to hidden objects. Half of the participants were provided with an experimental implementation of our visual support system that allowed them to view the edge point patterns on a computer screen. The other half of the participants had to rely solely on their mental representations of these patterns, as is the case in current demining practice.

Hypotheses
This experiment was driven by the hypothesis that pattern-based visual support would improve deminers' decision making abilities in classifying and locating buried objects. Specifically, we formulated the following hypotheses:

H.1 Visual support will result in the reduction of false negatives (missed mines);

H.2 Visual support will result in the reduction of false positives (mines declared where only clutter is present);

H.3 Visual support will result in smaller localization errors.

Participants
49 university students (31 male, 18 female) volunteered for this experiment. Their ages ranged from 18 to 24 (M=21). All participants were right-handed and either had normal or corrected-to-normal vision. Participants received a monetary reward of $10–$30 based on performance.
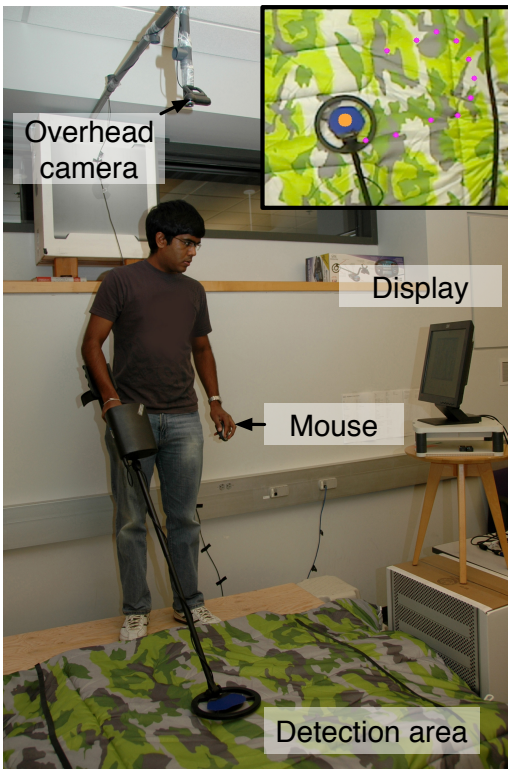
Figure 1. Experimental setup. The inset shows the contents of the visual support display: a live image of the ground with recorded edge points overlaid on top of the image.

(a) Mine with clutter task. The red oval was added later to help the reader notice the circular signature of a mine simulant.

(b) Clutter task

Figure 2. Two of the detection tasks used in our study together with screen shots of the our visual support systems showing sample patterns generated by our participants.
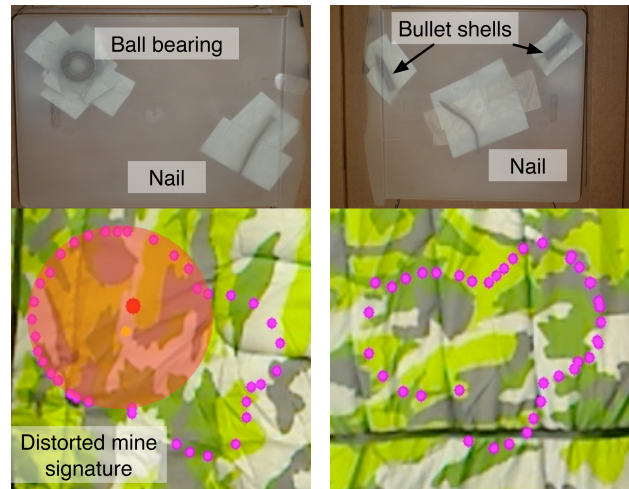
Apparatus

Physical setup

The physical setup (illustrated in Figure 1) consisted of a 1.5m by 2m detection area under which simulated mine detection tasks were concealed, a display mounted a little below face level to the left of the detection area, and an overhead camera. The detection area was raised 0.5m above the floor to account for metallic interference from the indoor concrete floor, and was covered with a camouflage fabric to approximate visual cues and distractors found in actual demining environments. The overhead camera was suspended 1.85m above the detection area with its optical axis perpendicular to the plane of the detection surface.

Equipment and the Visual Support System Prototype

A lightweight hobby metal detector (Bounty Hunter Pioneer 505) was used as a simulated landmine metal detector. The detector provided binary response—beeping or no beeping—in the presence/absence of metal.

Our custom-built prototype visual decision support system relied on the overhead camera (Logitech Pro 9000) to track the position of the metal detector head. Participants used the right button of a wireless mouse (Logitech 450), held in the non-dominant hand, to signal that the detector head was over an edge point (detector feedback on-off point) of the metallic signature. The on-screen display (presented on a 17" IBM Think Vision) showed a live-feed from the overhead camera with the positions of the reported edge points and currently tracked detector head position overlaid on top of the image. (Figure 1). Participants used the middle mouse button to declare the suspected location of the mine for those tasks for which they decided that a mine was present, and the left button to erase previously recorded mine or edge points. To maintain consistency, participants in both conditions were instructed to use the mouse to indicate that the detector was over an edge point or over a mine but only half of the participants were provided with the visual feedback.

For analysis, we recorded the complete trajectories traced by the metal detector head in addition to the positions explicitly marked by the participants. The software was written in Visual C++ using Open CV 2.0.

This implementation of our visual support approach captures the core aspects of the experience we want to provide for deminers while requiring only a modest engineering investment.

Simulated Mine Detection Tasks

We used metal ball bearings to simulate anti-personnel landmines, because with our experimental setup they generated signatures similar in form to those belonging to typical landmines [22, 3]. Metallic items collected from military training grounds such as barb wire, bullet casings and nails, along with household metallic items such as screwdrivers and wires, were used to simulate battlefield clutter. To ensure that identical detection tasks could be reproduced for each participant, we affixed combinations of mine simulants and clutter to plastic trays (10 x 24 inches) and carefully marked the positions where each tray should be placed during the experiment. The trays (illustrated in Figure 2) were placed (one per experimental condition) underneath the cam-

ouflage fabric while the participant was behind a partition, thus ensuring that the participants were unaware of the location and composition of objects in each task. Half of the trays contained one mine simulant plus 2–3 items of clutter (*Mine with clutter* tasks); the other half of the trays had only clutter items (*Clutter only* tasks). Four of these trays were set aside for the training and the remaining 6 were used in the testing phase of the experiment. We designed the Mine with clutter tasks such that the clutter significantly interfered with the metallic signature of the mine while still leaving some of the mine's characteristic circular signature visible (Figure 2(a))—this design choice was ecologically motivated given the prevalence of clutter on minefields.

Procedure

Training

Training and practice lasted about 1.5 hours on average. Participants were taught about detector sweeping styles, how to operate the metal detector, the difference between mines and clutter, how to identify edge points of metallic signatures, how to interpret the resulting patterns, and how to use the visual decision support system. Participants were trained to associate a fragment of a circular arc embedded in the overall pattern as indicating mine presence, with the center of the circle corresponding to the approximate location of the mine (see Figure 2(a)).

To avoid systematic differences in training across the experimental conditions, all participants were trained to build and interpret metallic signatures both with and without the visual support (only at the conclusion of the training were they told which condition they were assigned to). For both conditions, following Staszewski's protocol [3], participants were taught to move their detector head in small semi-circular motions between neighboring edge points, clicking the mouse button whenever they found one. When training with the visual support enabled, participants were specifically instructed to build patterns while focusing on detector head motion and only refer to the screen once they had finished construction.

During the practice session, participants had to perform detection tasks on four practice tasks (2 Mine with clutter, 2 Clutter). These tasks were similar in nature to the ones used for testing, and were presented in random cycles of four with visual support provided on every other trial. Subjects had to determine mine presence or absence for each task and localize a mine if they thought one was present. Participants were encouraged to think aloud about geometric features of the pattern that indicated mine presence or absence before making a decision. For each task, participants were provided with feedback on the validity of the reasoning applied. For training purposes, after completing a task without visual support participants were shown on screen the edge points they had collected. Participants had to correctly classify all 4 consecutive tasks in one full cycle before moving onto the testing phase. This ensured that all participants had a similar skill level at the beginning of the testing phase and that their performance had stabilized and was unlikely to change significantly during the experiment. Participants who failed to reach this proficiency criterion after 60 minutes of practice

were discharged (with base compensation of $10) without completing the testing phase of the study.

Testing

The 6 detection tasks (3 Mine with clutter and 3 Clutter only) used for testing were randomly ordered and counterbalanced for each participant to reduce ordering effects. To induce a consistent behavior among our participants, we instructed them to complete each task within 120 seconds, and we provided them with auditory notifications at regular intervals. This amount of time was sufficient to comfortably scan the area, establish the location of the edge points, and make a decision. Participants were tested under a monetary incentive structure intended to induce ecologically valid conservative behavior [23], where a false negative (a missed mine) yielded a penalty of $3 while a false positive (incorrectly classifying clutter as a mine) yielded a penalty of $1. Additionally, a reward of $1 was given for localizations within 6 inches of the underlying target position.

Each testing task commenced with participants sweeping the detector in densely-spaced side-to-side motions across the cell to locate the general area of feedback (area where the simulated mine detection task was hidden). Once feedback was encountered, participants engaged in pattern construction and interpretation. With visual support present, participants could consult the display to make a decision and possibly refine the pattern. The task concluded when the participants made and verbally announced their final decision. If a mine was thought to be present, participants indicated its suspected location using the middle mouse button during the course of the task. Participants had a mandatory break of two minutes at the end of each task to rest and to complete a post task questionnaire. At the end of the testing phase, participants completed an additional questionnaire that assessed overall ease and confidence. Results of testing performance were only divulged at the end of the experiment.

Design and Analysis

The experiment was a mixed between- and within-subject design with the following factors and levels:

- *Visual support* {Provided, Not provided}

- *Task type* {Mine with clutter, Clutter only}

- *Task ID* $\{1 \ldots 6\}$ – 3 tasks containing mine with clutter, and 3 tasks containing clutter only

- *Trial* $\{1 \ldots 6\}$

Visual Support was a between-subject factor, while Task type was a within-subject factor. The 44 participants whose data were included in the analysis (see the Results section for a discussion of exclusions) completed $44 \times 6 = 264$ trials. Of these, 180 were completed by male participants and 84 by female. Half of the trials contained a mine simulant with clutter and half only clutter. Genders were balanced between the two Visual support conditions. Task ID and Trial were both within-subject factors; their effects were analyzed only to verify uniform difficulty of tasks and the effectiveness of the training program at mitigating learning effects.

Our dependent measures were:

- **Classification errors** which included both the *false positives*, i.e., declaring a mine where one was not present, and *false negatives*—missing a mine. Because the two types of classification errors have very different implications for the demining practice, in all our analyses we include separate results for both.

- **Localization errors** which indicate the distance between the mine location indicated by the user and the actual location of the mine simulant. Localization errors were computed only for tasks where participants correctly detected a mine when one was actually present.

After each trial, we also collected 6 subjective measures on a Likert scale from 1 (strongly disagree) to 5 (strongly agree) relating to the ease of making classification and localization decisions, participants' confidence in their decisions, and the ease of creating and interpreting the patterns. Additionally, at the very end of the study, we asked participants to report (on a 5-point Likert scale) their overall confidence in 3 different aspects of their performance.

We used non-parametric tests in the analyses because the error data followed a highly skewed distribution and the Likert scale responses are more correctly treated as ordinal data. For the two-level factors we used Wilcoxon signed rank test for within-subject factors and Wilcoxon rank-sum test [27] (equivalent to the Mann-Whitney U test) for between-subject factors. We used the Friedman test to analyze the effects of Task ID. To test for learning effects across Trials, we performed a contrast analysis with all the pairwise comparisons using the Wilcoxon signed rank test.

To guard against Type I errors, we applied the Holm's sequentially-rejective Bonferroni procedure [21] to all analyses except the three planned comparisons that corresponded to our stated hypotheses and the two analyses verifying the task design and learning effects (Preliminaries section).

Results
Adjustment of Data
Three out of the 49 volunteers did not pass the exit criterion of our training program and were thus discharged without completing the study. Data from one participant were excluded because of an experimenter's error in following the protocol, and another subject's data were lost due to software malfunction. Consequently, data from 44 participants (30 male, 14 female) were considered in the analysis.

Preliminaries
In the design of this experiment, we strove to maintain uniform difficulty among all the tasks of the same type (i.e., all the Mine with clutter tasks and all the Clutter only tasks). We also designed our training procedure to minimize learning effects during the testing phase. Indeed, we observed no significant effect of Task ID on classification errors for either task type. Similarly, a contrast analysis showed no learning effect on classification performance.

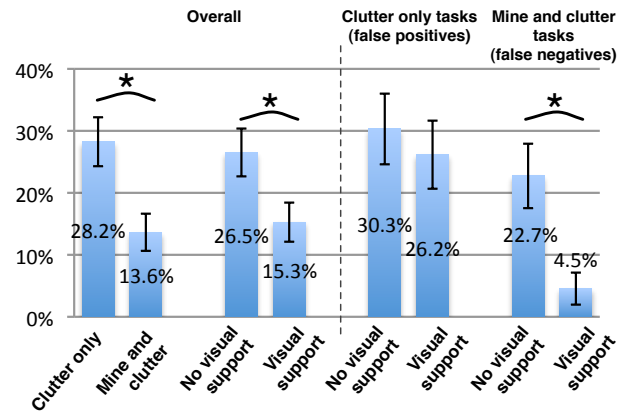### Average Classification Errors

Figure 3. Average classification error rates. Error bars correspond to the standard error and the ⋆ symbols denote statistically significant differences.

Classification Errors
We observed a significant main effect of Visual support on classification errors ($z = 2.36$, $p < 0.05$): with visual support provided, participants made an average of 15.3% classification errors, compared to 26.5% without visual support. Visual support resulted in significant reduction in false negatives ($z = 2.87$, $p < .01$): participants missed on average 22.7% mines without visual support but only 4.6% when the visual support was provided. In other words, Visual support resulted in an 80% reduction in the number of missed mines. Hypothesis H.1 was thus supported. Hypothesis H.2, however, was not supported: Visual support had no significant effect on false positives ($z = 0.41$, n.s.).

We also observed a main effect of Task type ($z = 130.0$, $p < .01$): on average, participants made 28.2% classification errors on tasks that contained only clutter compared to 13.6% on tasks that involved a mine.

Localization Errors
We analyzed the effects on localization errors for those tasks where the participants correctly classified the task as containing a landmine. We observed a significant main effect of visualization support on localization error ($z = 4.25$, $p < .0001$): the availability of visual support resulted in the reduction of the average localization error from 30.0 pixels to 20.7 pixels (approximately from 3 inches to 2 inches). Hypothesis H.3 was thus supported.

Subjective Results
Participants reported that it was significantly easier to decide whether there was a mine present when the visual support was provided (M=3.35 vs. 2.84 without visual support, $z = 3.53$, $p < .001$). For this measure there also was evidence of interaction between Visual support and Task type: participants found that visual support made the decision easier for tasks containing a mine ($z = -4.73$, $p < .0001$) but not those containing clutter only ($z = 0.15$, n.s.).

	Visual support		
	not provided	provided	* = significant
Ease of classifiaction	2.84	3.35	*
Confidence in classification	3.37	3.64	
Ease of pattern construction and interpretation	3.28	3.80	*
Noticed good symmetry	2.77	3.60	*
Ease of localization	3.18	3.48	
Confidence in localization	3.43	3.67	

Table 1. Summary of subjective results. Responses were collected on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree).

With the visual support present, participants also found it easier to construct and interpret the edge point patterns (M=3.80) than without visual support (M=3.28, $z = 4.15$, $p < .0001$). As with the previous measure, we again see evidence of interaction between Visual support and Task type: the presence of visual support resulted in increased ease of pattern creation and interpretation for tasks involving a mine ($z = -4.37$, $p < .0001$) but not for those that involved only clutter ($z = 1.52$, n.s.).

Finally, participants reported noticing significantly better partial circular symmetry in the constructed patterns when the visual support was provided (M=3.60 vs. M=2.77 without visual support, $z = 4.98$, $p < .0001$).

We did not observe significant effects of Visual support on any of the other subjective measures.

Patterns and Pattern Creation Strategies
Next we take a qualitative look at the sensor head trajectories traced by four of our participants as well as the patterns they created. We will use the term "petals" to refer to the continuous arcs traced with a metal detector between two adjacent edge points.

To evaluate whether pattern creation strategies influenced a deminer's success at correctly classifying targets, we first rank-ordered the participants (separately for each of the two Visual support conditions) based on their classification error rates (weighing false negatives a little higher than false positives) and their localization errors. The sensor head trajectories and edge points created by the highest and lowest performers in each of the two visual support conditions are shown in Figure 4. The most striking difference is that the lowest performers—both with and without visual support—returned the sensor head to the same positions multiple times and had a large number of overlapping petals (Figures 4(c) and 4(d)). Additionally, the number and density of the edge points varied considerably between the high and low performers. Lower performing participants under-sampled the boundary (as in the Mine with clutter task in Figure 4(d), where the participant marked a relatively small number of edge points) or over-sampled in some areas (such as the bottom left corner of the boundary of that same pattern). Note the dense and regular spacing of the edge points collected by the high performers (Figures 4(a) and 4(b)) when compared to the patterns of lower performers shown in Figures 4(c) and 4(d).

EXPERIMENT 2
To gain further insight into the factors affecting deminers' success at classifying and locating targets using our pattern-based visual support approach, we conducted a second experiment for which we recruited a new set of 13 participants to interpret the patterns created by a subset of the participants from the first study. The new participants only received brief instruction on interpreting the patterns but they were given no practical training on the actual demining procedures. From now on, whenever the term participant may be ambiguous, we will refer to the participants from the first experiment as the (original) *deminers* and those from the second experiment as *evaluators.*

Participants
13 university students (7 male, 6 female) volunteered for this experiment. Their ages ranged from 18 to 22 (M=20). All participants either had normal or corrected-to-normal vision. Participants received a monetary reward of $10–$16 based on performance.

Apparatus
We created new custom software to display patterns and elicit participants' classification and localization feedback. Participants viewed the information on a 17" IBM Think Vision monitor.
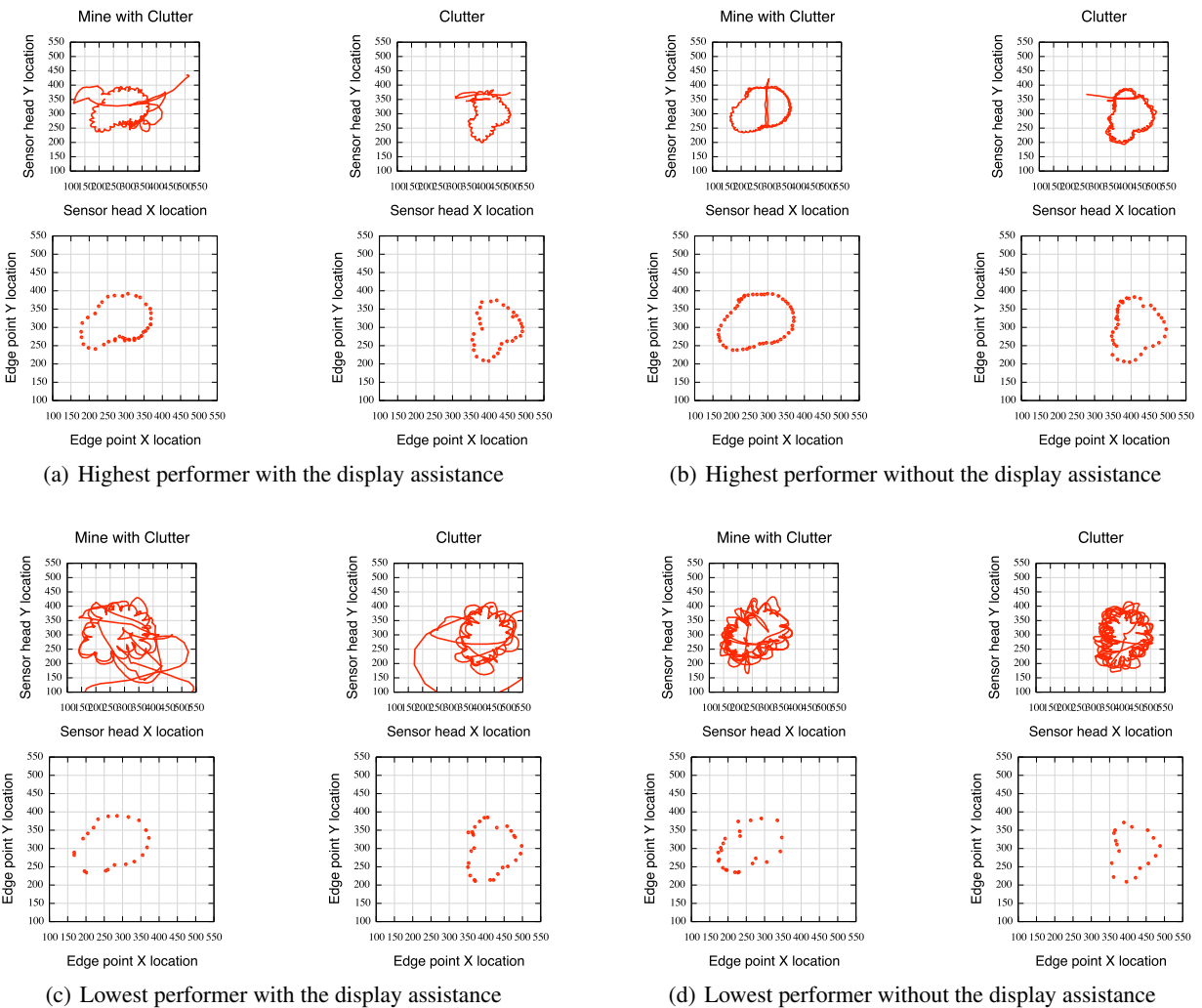
Selection and Preparation of Patterns
In this study, we presented the patterns generated by participants in the first experiment on a neutral background. We used patterns generated on training tasks in the first study for training in the current experiment. Similarly, we used patterns from the testing phase of the first experiment for the testing phase of this study. We randomly selected 24 patterns from the training phrase (equally split between the two task types). For the testing phase, we first sorted the original deminers according to their performance using the same procedure as the one described in the earlier section on Patterns and Pattern Creation Strategies. We then randomly picked, separately for the two Visual support conditions, 2 participants in the top performance quartile (High Performers) and 2 in the low performance quartile (Low Performers) for a total of $2 \times 2 \times 2 = 8$ original deminers. These original deminers had each generated 6 patterns (3 for tasks with Mine with clutter, 3 for tasks with just Clutter). This provided us with a testing set of $8 \times 6 = 48$ unique patterns. Even though all the patterns were unique, they were all generated based on the same smaller set of tasks (4 training, 6 testing). To increase diversity of the patterns experienced by the new participants, we randomly rotated and translated each of the patterns.

Procedure
Training
Training and practice lasted 30 minutes on average. Participants were first briefed on the challenges of mine detection with a metal detector, the difference between mines and clutter, the usefulness of metallic signatures, and the use of the experimental software. Next, participants used our software

(a) Highest performer with the display assistance

(b) Highest performer without the display assistance

(c) Lowest performer with the display assistance

(d) Lowest performer without the display assistance

Figure 4. Qualitative comparison of sensor head trajectories (top plots in each subfigure) and the patterns (bottom plots) created by four of our participants.

to carry out practice classification and localization tasks on randomly selected patterns drawn from the training set. Selected patterns were checked to ensure that no two consecutive patterns were from the same task. If a pattern was incorrectly classified the participant received immediate notification of the error. As in the previous experiment, participants were asked to think aloud as they made their decisions and they were provided with extensive feedback on their reasoning. Participants had to correctly classify 6 consecutive patterns before moving on to the testing phase. To ensure a minimum amount of training among participants, all had to experience a minimum of 10 patterns during the training phase even if they were immediately successful.

Testing
The 48 patterns used in the testing phase were ordered randomly for each participant to reduce ordering effects. As in the training phase, we ensured that no two consecutive patterns were generated based on the same task. To mir-

ror the time pressure from Experiment 1, participants were instructed to spend no more than 15 seconds on each pattern (this bound was determined based on average perceived time deminers with visual support took to make decisions once they had constructed visual patterns). As in the first experiment, the monetary incentive structure was designed such that missing a mine incurred a larger penalty (15¢) than declaring a mine where none was present (5¢). Additionally, a reward of 5¢ was given for each localization within 20 pixels of the true mine position. Results of performance during the testing phase were only divulged at the end of the experiment.

Design and analysis
The experiment was a within-subjects design with the following factors and levels:

- *Visual support (in the original study)* {Provided, Not provided}

- *Performance level of the original demiiner* {High performer, Low performer}

- *Original classification correctness* {Correct, Incorrect}; this factor indicates if a pattern was correctly classified by the original demiiner who created it in Experiment 1

- *Task type* {Mine with clutter, Clutter only}

As in the first experiment, our dependent measures were the Classification error and the Localization error. We used the same non-parametric tests to analyze our data as in Experiment 1 and we used the Holm's sequentially-rejective Bonferroni procedure [21] to guard against Type I errors.

Results
The evaluators in this study made on average fewer errors (14%) than the 8 original demiiners whose patterns were used in the second experiment (19%), but the difference was not statistically significant. There was also no significant difference in the number of classification errors the evaluators made on patterns that were originally created with the Visual support, compared to those constructed without Visual support. We did observe, however, a significant main effect of Task type on classification error ($z = -31.5$, $p = .01$): on average, participants had a false positive rate (detecting a mine when there was none) of 6.4% while they missed a mine 21.5% of the time. This is in contrast to the results from the first experiment, where the demiiners' false positive rate was substantially higher than their false negative rate (28.2% vs. 13.6%).

We observed a significant main effect of the Performance level of the original demiiner ($z = 191.5$, $p < .001$): participants made an average of 10% classification errors on patterns from high performers, compared to 18% on patterns from low performers. We also observed a significant main effect of Original correctness ($z = 286.000$, $p < .0001$): participants made an average of 10.7% classification errors on patterns that were correctly classified by the original demiiner, and an average of 24.7% errors on patterns incorrectly classified in Experiment 1. These results provide further evidence that quality of created patterns impact the likelihood of correct classification.

We observed no significant effects of any of the factors on the localization errors.

DISCUSSION AND DESIGN INSIGHTS
The results of the first experiment support our hypothesis that visual support helps to reduce false negatives: participants who worked with the visual support missed 80% fewer mines than those who worked without it. Visual support also resulted in a 30% reduction in the magnitude of mine localization errors. Our hypothesis that visual support would reduce false positives (cases where a demiiner declares a mine where none is present) was, however, not supported. While not an immediate threat to demiiners' safety, false positives can be a significant source of inefficiency in metal detector–based landmine clearance. In order to help mitigate decision errors in general and the false positives in particular, we sought to gain insights into the sources of these errors.

The qualitative analysis of the sweeping trajectories indicated a consistent difference in behavior between the most successful participants and those who made the most classification errors: the successful participants created the patterns in a single pass along the perimeter of the metallic signature and sampled the edge points at small, regular intervals. The least successful participants, meanwhile, re-traced the outline two or more times and sampled the edge points at irregular intervals, with subsequent passes occasionally filling gaps left earlier and occasionally duplicating earlier work. These observations suggest that classification errors by demiiners can be partially attributed to the quality of the patterns they created. They also point to a specific behavior that results in low-quality patterns. Results from the second experiment—which show that evaluators made significantly more errors interpreting patterns that were originally misclassified and those generated by the lowest performing original demiiners—provide further evidence that poorly-constructed patterns contribute to the classification errors.

These findings suggest two ways of further reducing classification errors. First, the discussion of successful and unsuccessful sweeping strategies can be incorporated more prominently into the training program. Second, a decision support system can make demiiners aware when their sweeping strategy is likely to result in unreliable patterns. Errors in pattern construction might be reduced if the process of edge point capture was automated.

The results of our second experiment also demonstrate that people other than those who created the patterns can successfully interpret them. With current approaches, collaborative decision making is not possible because the person operating the metal detector cannot externalize and share the pattern she has observed. In contrast, a system that visually captures these patterns would enable demiiners to easily consult peers and supervisors. Because it is not uncommon for more than one person to be assigned to a single demining cell (with only one operating the metal detector) [7], such peer assistance is often readily available.

Finally, we note that our work demonstrated that explicitly visualizing the edge points of a metallic signature improves demiiners' decision making capabilities. To evaluate the resulting patterns, however, demiiners still have to refer to mental representations of previously encountered patterns in order to make a classification decision. We hypothesize that demiiners' classification performance can be further improved if the device can enable the storage and retrieval of previously encountered patterns for subsequent side-by-side comparisons.

The major limitation of our work has been the limited ecological validity of our experimental setup. Our results are, however, still informative and they will serve to inform our design of a higher fidelity outdoor prototype, which we in-

tend to evaluate at an actual deminer training facility with real deminer trainees.

CONCLUSION

In this paper, we made three main contributions. First, we proposed and evaluated a novel visual support approach to improve the decision making capabilities of deminers. Our approach enables metal detector operators to explicitly visualize—instead of having to imagine—the geometry of the electromagnetic induction patterns produced by buried metallic objects. In our first experiment, participants working with our visual support missed 80% fewer mines and reduced their mine localization errors by 30% compared to those who relied on mental representations alone. Contrary to our expectations, however, the visual support did not result in the reduction of false positives: both groups of participants were equally likely to misclassify a piece of harmless clutter as a landmine.

Second, our results allowed us to identify some likely sources of deminer classification errors. These findings allowed us to propose several design suggestions that are likely to help reduce deminers' classification errors.

Lastly, the results of our second experiment demonstrate that people other than those who originally constructed the pattern can successfully interpret these patterns. This opens the possibility of new, collaborative, decision making procedures for landmine clearance.

We intend to implement this approach on top of an existing mobile hardware platform to produce a cheap and robust visual support device that can be easily mounted on top of most existing metal detectors. Because such a device will be simple, inexpensive, robust, portable, and able to function in the field without recharging for hours at a time, we hope that it will gain adoption in the humanitarian demining community.

REFERENCES

1. Andersion, N., da Sousa, C., and Paredes, S. Social cost of land mines in four countries: Afghanistan, Bosnia, Cambodia, and Mozambique. *British Medical Journal 311*, 7007 (1995), 718.

2. Boshoff, C., and Cresci, R. The HALO Trust and HSTAMIDS. *Journal of Mine Action 12*, 1 (2008).

3. Davison, A., and Staszewski, J. Handheld mine detection based on expert skill: Reference guide for training plan design and training site development. Aberdeen Proving Ground, MD: Army Research Laboratory Human Research and Engineering Directorate, 2002.

4. Geneva International Centre for Humanitarian Demining (GICHD). *A Study of Mechanical Application in Demining*. GICHD, Geneva, 2004.

5. Geneva International Centre for Humanitarian Demining (GICHD). *Mine detection dogs: Operations*. GICHD, Geneva, 2005.

6. Geneva International Centre for Humanitarian Demining (GICHD). *A study of manual mine clearance, Book 1*. GICHD, Geneva, 2005.

7. Geneva International Centre for Humanitarian Demining (GICHD). *A study of manual mine clearance, Book 2*. GICHD, Geneva, 2005.

8. Geneva International Centre for Humanitarian Demining (GICHD). *Guidebook on Detection Technologies and Systems for Humanitarian Demining*. GICHD, Geneva, 2006.

9. Geneva International Centre for Humanitarian Demining (GICHD). *Detectors and Personal Protective Equipment*. GICHD, Geneva, 2009.

10. Geneva International Centre for Humanitarian Demining (GICHD). *A Guide to International Mine Action Standards*. GICHD, Geneva, 2009.

11. Geneva International Centre for Humanitarian Demining (GICHD). *A Guide to Mine Action*. GICHD, Geneva, 2010.

12. Habib, M. Humanitarian demining: The problem, difficulties, priorities, demining technology and the challenge for robotics. In *Humanitarian Demining Innovative Solutions and the Challenges of Technology*, M. Habib, Ed. I-Tech Education and Publishing, Croatia, 2008, ch. 1, 1–56.

13. Handicap International. *Voices from the Ground*. Handicap International, Brussels, 2009.

14. Herman, H., McMahill, J. D., and Kantor, G. Enhanced operator interface for hand-held landmine detector. *Proceedigns of SPIE 4394*, 1 (2001), 844–851.

15. Kruger, H., and Ewald, H. Handheld metal detector with online visualisation and classification for the humanitarian mine clearance. In *Sensors, 2008 IEEE* (2008), 415–418.

16. Landmine and Cluster Munition Monitor. *Landmine Monitor Report 2009, Executive Summary*. 2009.

17. Landmine and Cluster Munition Monitor. *Landmine Monitor Report 2010, Executive Summary*. 2010.

18. Lee, T., and Soatto, S. Tracking and recognition on a hand-held. Demonstration at IEEE CVPR'10, 2010.

19. National Robotics Engineering Center. Sweep Monitoring System. http://www.rec.ri.cmu.edu/projects/sweep_monitoring/. Last accessed on January 13, 2011.

20. Sato, M., Fujiwara, J., Kido, T., and Takahashi, K. ALIS evaluation tests in croatia. *Proceedings of SPIE 7303*, 1 (2009).

21. Shaffer, J. P. Multiple hypothesis-testing. *ANNUAL REVIEW OF PSYCHOLOGY 46* (1995), 561–584.

22. Staszewski, J. Spatial thinking and the design of landmine detection training. In *Applied spatial cognition: from research to cognitive technology*, G. A. Allen, Ed. Mahwah, NJ, Erlbaum Associates, 2006, 231–265.

23. Stevens, S. *Handbook of experimental psychology*. John Wiley, New York, 1951.

24. Trevelyan, J. Paint marker device. http://school.mech.uwa.edu.au/~jamest/demining/tech/markers/minem3.html, 1999. Last accessed on January 11, 2011.

25. Trevelyan, J. Research challenges. In *Humanitarian Demining Innovative Solutions and the Challenges of Technology*, M. Habib, Ed. I-Tech Education and Publishing, Croatia, 2008, ch. 2, 57–68.

26. Walsh, N. E., and Walsh, W. S. Rehabilitation of landmine victims : the ultimate challenge. *Bulletin of the World Health Organization 81* (09 2003), 665 – 670.

27. Wilcoxon, F. Individual comparisons by ranking methods. *Biometrics Bulletin 1*, 6 (1945), 80–83.