

Aggregation of Choquet integrals in GPR and EMI for handheld platform-based explosive hazard detection

Ryan E. Smith^a, Derek T. Anderson^a, John E. Ball^a, Alina Zare^b, and Brendan Alvey^c

^aElectrical and Computer Engineering, Mississippi State University, Mississippi State, MS

^bElectrical and Computer Engineering, University of Florida, Gainesville, FL

^cElectrical and Computer Engineering, University of Missouri, MO

ABSTRACT

Substantial interest resides in identifying sensors, algorithms and fusion theories to detect buried explosive hazards. This is a significant research effort because it impacts the safety and lives of civilians and soldiers alike. Herein, we explore the fusion of different algorithms within and across *ground penetrating radar* (GPR) and *electromagnetic induction* (EMI) sensors on a U.S. Army NVESD furnished *experimental handheld demonstrator* (EHHD) platform. Fusion is not trivial as these sensors have different sampling rates, resolutions, they observe different spatial areas and they span different portions of the electromagnetic spectrum. Herein, we investigate and compare the use of different approaches for co-registration and decision-level fusion using the *Choquet integral* (ChI). With respect to the ChI, we explore the impact of using a “global” (single) aggregation strategy (operator) versus tailoring different ChIs to subsets of algorithms and sensors. *Receiver operating characteristic* (ROC) curve results are shown for data from a U.S. Army test site containing multiple target and clutter types, burial depths and times of day.

Keywords: Explosive hazard detection, fusion, electromagnetic induction, ground penetrating radar

1. INTRODUCTION

Explosive hazards remain a great threat to soldiers and civilians alike. *Explosive hazard detection* (EHD) technologies include, to name a few, *ground penetrating radar* (GPR),^{1–8} which uses radio frequency pulses to scan below the ground for hazards, *electromagnetic induction* (EMI),⁹ which uses magnetic fields to sense metallic objects. EHD is an active research field, and is certainly not a solved problem. EHD is a difficult task for several reasons. First, there are many target types that need to be detected, ranging from smaller anti-personnel to larger anti-tank mines. These mines may be composed of high amounts of metal, low amounts of metal, or plastic. Second, false returns caused by non-mine objects (clutter) in the soil and natural variations in the soil’s electromagnetic properties need to be mitigated as much as possible, while not masking real targets. The GPR sensors can detect non-metallic hazards, but do not perform as well for objects too deep in the soil (due to the natural decay of the EM signal) or objects too close to the ground (due to the large ground return signal). EMI sensors, on the other hand, are better at detecting metallic objects. No single sensor works well across a large range of EHD targets.

Herein, we investigate a small, lightweight multi-sensor handheld EHD platform that can be swept by a soldier. This platform is called the *experimental hand held demonstrator* (EHHD) and it is provided by the US Army RDECOM CERDEC NVESD. The goal of this system is to automatically analyze sensor data and provide a warning if anything suspicious is found. Too many *false alarms* (FAs) will render the device useless, whereas too few of detections is lethal. In order to provide robust detection, we take advantage of the strengths of GPR, EMI, and precise positioning (GPS and IMU). First, the individual sensors are processed by their respective prescreeener algorithms, whose outputs represent the target confidence with respect to that sensors *field of view* (FOV) at a given sample location. Note, the sensors typically have different FOVs, due in part to the underlying sensor and its placement. Next, using precise positioning, the raw sensor data and/or prescreeener outputs are re-sampled onto a common underlying grid (a process we refer to herein as “direct linking”) so measurements in one sensor and/or algorithm correspond to a known measurement in the other sensor(s). Herein, seven different approaches are explored for grid re-sampling. The first four methods are based on *k-nearest neighbor* (kNN), the

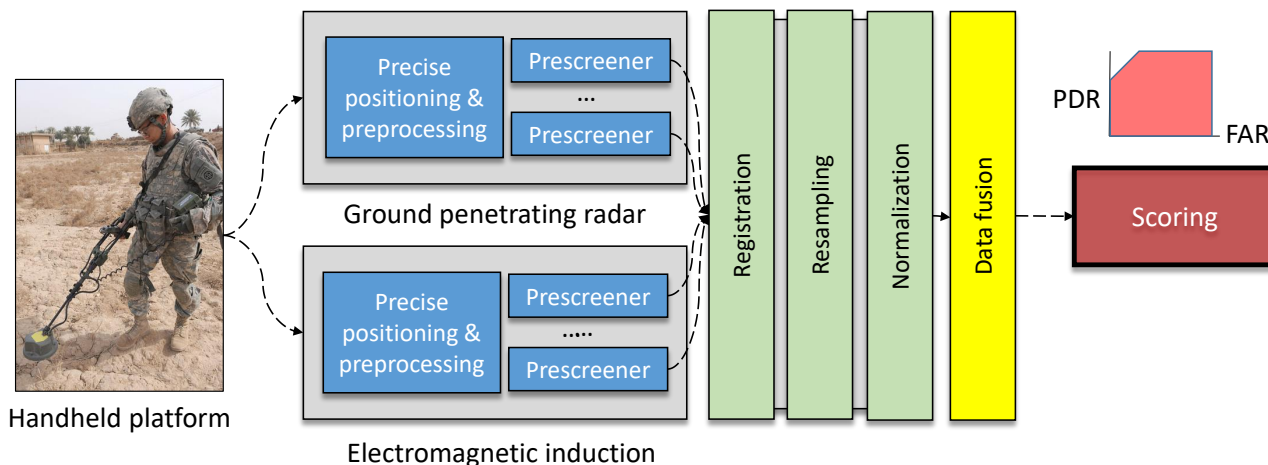


Figure 1. Illustration of explosive hazard detection using a multi-sensor handheld platform (Section 2). Our approach begins with individual sensor data collection and processing (Section 1), followed by multi-sensor registration, resampling and normalization (Section 3). Last, the information within and across these sensors are fused (Section 4). The focus of this paper is registration, re-sampling, normalization and fusion (green and yellow boxes).

fifth uses a Gaussian to capture spatial uncertainty and the last two use spline interpolation.¹⁰ With respect to fusion, we consider a “fixed” (single operator) *Choquet integral* (ChI) aggregation as well as the aggregation of ChIs for different subsets of sensors and/or algorithms within and across sensors.

The contributions of this paper are as follows. First, we explore seven methods for direct linking of EMI and GPR data relative to precise positioning. Second, we consider an extension of the ChI via tailoring multiple ChIs to different subsets of algorithms and sensors to better model the aggregation needs of our different inputs. As we show, selection of direct linking and allowing for the aggregation of multiple ChIs has a noteworthy performance impact.

The remainder of this article is organized as follows. In Section 2 our EHHD platform is discussed. Section 3 is direct linking, Section 4 is fusion within and across sensors via the ChI and the aggregation of ChIs, and Section 5 is experimental validation. Last, Figure (1) illustrates the primary steps in our proposed approach.

2. EXPERIMENTAL HANDHELD DEMONSTRATOR (EHHD)

As stated in Section 1, the EHHD platform used here is provided by the U.S. Army RDECOM CERDEC NVESD and it has GPR, EMI and precise positioning (IMU and GPR) sensors. The following subsections quickly summarize relevant information about the existing sensors and prescreeners.

2.1 Ground Penetrating Radar

A GPR system must be able to accomplish a number of tasks, such as detect metal and plastic hazards, detect *anti-tank* (AT) and *anti-personnel* (AP) hazards, work in different soils, obtain a satisfactory *positive detection rate* (PDR) relative to a specified *false alarm rate* (FAR), detect hazards at varying depths, work in different environments and environmental conditions, operate a given number of days per week and hours per day, be simple to use, etc. As the reader can see, EHD is not trivial. In general, a GPR system works by transmitting *electromagnetic* (EM) waves into the ground and it samples the backscattered echoes, which occur as a response to electrical parameter contrasts in the air and ground. GPR is ideal as it detects echoes from relative permittivity, relative permeability and conductivity. In theory, this enables GPR to locate both metal and non-metal hazards, provided there is adequate contrast with the surrounding soil. Herein, we focus on the EHHD handheld GPR system, which uses a single transmitter (Tx antenna) and one or more receivers (Rx antennas).

At time step t , the GPR system records its location, e.g., easting and northing coordinate according to *Universal Transverse Mercator* (UTM), altitude, and a sample (the signal). This signal is referred to hereafter as an A-Scan. For a GPR system pointed at the ground, samples represent time/depth. As multiple A-Scans are

collected in a sweep, we can put them together and form a B-scan. Without precise positioning, the variable rate at which a user sweeps the ground results in a variable resolution/sampled B-scan. This complicates a number of things such as the appearance of targets and our ability to extract features and detect them. Herein, we use precise positioning to re-sample a signal to some near constant sample rate, e.g., every 1cm on the ground. If an *area of interest* (AOI) is swept multiple times in different ways, i.e., directions or patterns, a subject of research in itself, then we get multiple B-scans and can use precise positioning to construct a dense or sparse C-scan. In this article, we restrict our analysis to the processing of just B-scans. In part, this is due to the nature of a handheld system and the NVESD EHHD data that we have access to at the moment. Figure (2) illustrates the A-Scan, B-Scan and C-Scan concepts.

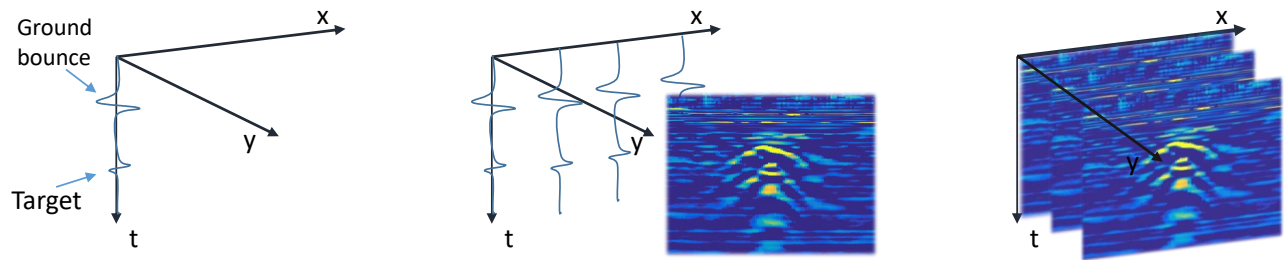


Figure 2. Illustration of a GPR A-Scan (left), B-Scan (middle), and C-Scan (right).

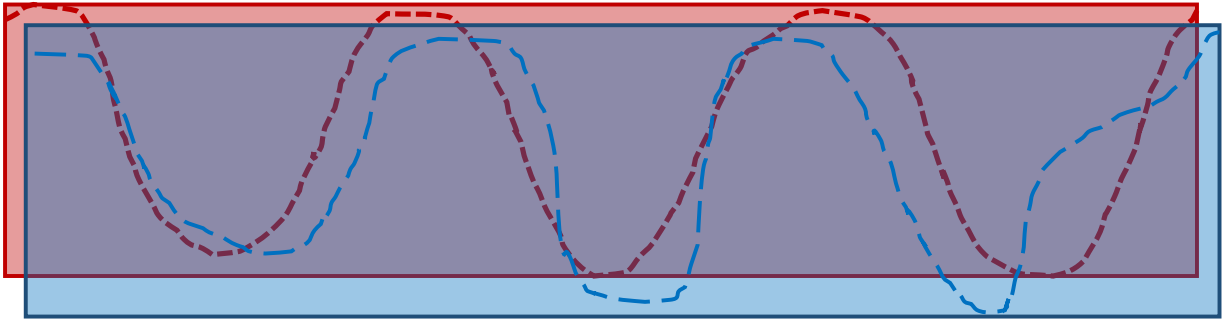
Previously, we put forth a GPR technique called *background adaptive division* (BAD).⁷ BAD is an enhancement method that can be used in the design of a prescanner and/or feature extraction. BAD uses past radar samples to estimate the average intensity versus depth. It then uses this information to enhance current scans. The primary advantage of BAD is that it produces high-quality enhancements at a very low computational complexity, which allows it to run in real-time. BAD does not require data ground alignment prior to processing. Moreover, BAD does not require the sensor to maintain a constant distance between the hand-held unit and the ground surface, which makes it more attractive from a hand-held sensor viewpoint.

Previously, we put forth two additional techniques for denoising and enhancing target returns from GPR data based on *Curvelets—reverse reconstruction* (RRE) and *wedge selection* (WS).⁸ These techniques make use of the Curvelet transform to address challenges in GPR such as ground return, noise and contrast of targets in the signal. We built a prescanner that works on the Curvelet denoised returns.⁸ This prescanner consists of the aggregation of a bank of matched filters, size contrast filtering and clustering for alarm linking.

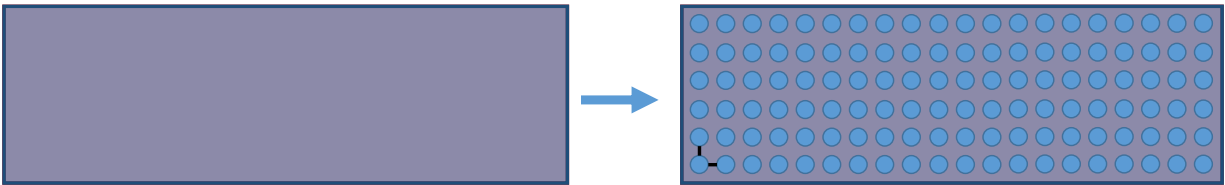
2.2 EMI

The EHHD *wideband EMI* (WEMI) sensor used here emits energy, via a transmit coil, in the form of a time varying electromagnetic field. The field causes elements below the sensor to induce their own electromagnetic fields, which is then picked up by one or more receive coils. The EHHD sensor operates at twenty-one logarithmically spaced frequencies and the result is a complex number-valued vector that represents the measured response at each operating frequency. As in the case of GPR, the WEMI sensor is on a platform with GPS and IMU sensors to record the position of the sensor during data collection.

We quickly summarize our WEMI prescanner.⁹ First, the data is filtered in the down-track direction by convolving it with a zero-mean sine filter. The data is then normalized. The complex number vector is extended to a forty-two dimensional vector, where the first half of values are the real numbers and the remainders are the imaginary responses. The real mean is subtracted and each sample is divided by its L2 norm. Last, the global *adaptive coherence estimator* (ACE) is used for alarm confidence generation.⁹ ACE essentially computes the inner product between a test point and a target signature in a whitened coordinate space. In global ACE, the inner product is computed in a coordinate space that is whitened according to the global data mean and covariance (i.e., the background is assumed to be Gaussian distributed according to μ and Σ which represent the global data mean and covariance, respectively.) In the global ACE implementation, target signature are constructed from the *discrete spectrum of relaxation frequencies* (DSRF) model. One hundred dictionary elements are generated



(a) Simplified illustration of sampling with respect to two sensors (red and blue)



(b) Imputation of a uniformly sampled grid with respect to the intersection of sensor coverage

Figure 3. Illustrations showing (a) hypothetical sweeps (locations and FOVs) for two sensors and (b) re-sampling of their intersected area via a uniformly sampled grid. Sensor one (blue) is at a lower rate and sensor two (red) has a higher sampling. Purple is the area they jointly observe. Light blue and red are regions seen by only sensor one and two. Note, their sampled locations sometimes intersect and sometimes do not.

using relaxation frequencies logarithmically spaced from 45Hz to 670Khz to model metallic objects. The shift is set to zero, as the real mean is relatively consistent amongst buried objects, and thus is removed without losing much useful information. Appropriate background estimation is very important for ACEs performance as a detection statistic. Once the background model is estimated using the global data mean and covariance, the ACE statistic is computed using each element in the DSRF target dictionary. The maximum ACE statistic across all the elements is stored as the points final confidence. Last, hits from global ACE are linked herein using the mean shift clustering algorithm.¹¹

While Alvey et al. explored a number of different detectors in,⁹ e.g., *joint orthogonal matching pursuits* (JOMP) and *Woodbury identity updated ACE* (WACE), we focus on just global ACE herein based on its reported performance. In future work, we will likely investigate the fusion of global ACE, WACE and JOMP.

3. MULTI-SENSOR CO-REGISTRATION AND RE-SAMPLING

The term “multi-sensor fusion” is rather broad and somewhat ambiguous as numerous *theories* are involved,¹² e.g., data imperfection, outliers, conflicting data, data modality, data correlation, data alignment/registration, data association, dimensionality, to name a few. Our EHHD platform is currently geared towards research. As such, the two sensors are not co-located. Instead, different pre-programmed carts are used to collect their data. As a result, we have different sweeps that may or may not observe the same location. Furthermore, the collection times of these platforms were different and while precise positioning helps there will inevitably be registration and linking error. Regardless, even if the sensors were co-located on a single platform they would still have different FOVs, sampling rates, resolutions, etc.—and therefore be a degenerative case of the challenge we investigate herein. Figure (3) is an illustration of our particular multi-sensor challenge.

Herein, each sensors FOV is taken into consideration. As the sensor moves, we take the union of those FOVs. Next, we intersect these unioned areas across sensors. For S sensors, this creates at most $2^S - 1$ (potentially non-contiguous) partitions. In Figure (3)(a), there are two sensors ($S = 2$), thus there are 3 areas (red, blue and purple). Relative to our data collection and sweep patterns, things are simplified (computation and indexing versus memory storage) if we compute the convex hull followed by the rotating calipers algorithm.¹³ This is done within each sensor and we intersect those results. Once the minimum bounding box is identified we discretize it

to a desired uniform distance sample rate (illustrated in Figure (3)(b)). In the following subsections we discuss different methods for imputing values in this sampled grid relative to our sensors—specifically relative to the outputs of our different prescreeners.

3.1 Re-Sampling: Closest, Nearest Neighbor and Neighborhood Surfaces

We started by exploring four simple techniques; closest grid point, *nearest neighbor* (NN), bounded NN and bounded neighborhood aggregation. We are aware that these are simple re-sampling strategies. However, if the sensors sampling rate, sensors FOV, sweep pattern and coverage, and other related platform factors are good enough for the given application, then we might be able to justify the performance of such a technique relative to its computational benefits in a real-time system. In the first method, coined closest grid point, each sample from our sensor is assigned to its nearest (Euclidean distance) grid point. However, a shortcoming of this approach is when the sampling is not high enough we induce a sparse grid (i.e., most grid locations have value zero) that does not help us with co-locating two (or more) sensor data for fusion. In the case that two samples map to the same grid point, we simply take the maximum of those values (again, we are operating on confidence values herein). In the NN approach, each location in the re-sampled grid is assigned a value corresponding to the closest (Euclidean distance) sampled point from our sensor. Whereas this helps us now assign a value to each location in the grid (its not sparse like the first method), one shortcoming is that it can be extreme in the sense that values are propagated potentially further than desired (is safe/reasonable to assume). The third method, which we call bounded NN, remedies this challenge by bounding the maximum distance that we can search for neighbors. Locations are now assigned a value of zero if there is no sensor data in its neighborhood. In the final approach we search in a halo of radius r for all samples and the maximum value is taken. If there are no samples then a value of zero is assigned.

3.2 Re-Sampling: Gaussian Propagation

A drawback of the methods in Section 3.1 is they do not take into account spatial uncertainty per se. That is, each sensor observation is relative to a specific location and the sensors FOV. As we move away from the sampled location the signal, and our subsequent decisions, could very likely change. In this subsection, we use a Gaussian distribution with height equal to the assigned confidence value and a standard deviation equal to a user defined amount. Thus, sampled points centrally located on or around our observation have more confidence and as we move away from that point our confidence decays. In general, we could model the FOV of our sensor and use that to specify sigma. However, for our initial work the cart was at a fixed height and we selected a single sigma proportional to the approximate FOV of the different sensors. Algorithm 1 summarizes the Gaussian propagation method for a single sensor.

Algorithm 1 Gaussian-based confidence surface re-sampling

1: Input $X = \{x_1, \dots, x_k\}$	▷ Set of positions and associated confidences
2: Input σ	▷ User defined standard deviation
3: Initialize S and R to all zeros	▷ Re-sampled set of grid points
4: for All points i in X do	
5: for All points j in S do	
6: $S(j) = X(i)N(j, i, \sigma)$	▷ Gaussian at location j with respect to mean i , confidence $X(i)$ and σ
7: end for	
8: $R = R \cup S$	
9: end for	
10: Output R	

3.3 Re-Sampling: Spline Interpolation

Whereas Section 3.2 tries to model and exploit spatial uncertainty, it does so by populating the re-resampled grid on a per-*element* (aka sample from our sensor) basis. In this section, we consider a more holistic approach, in the sense that all points are simultaneously used to generate the grid. In particular, we do this by using splines. However, numerous *types* of splines exists, see¹⁴ for an introduction and overview and Figure (5) shows four

different examples. In general, splines are piecewise defined polynomial functions that vary in terms of smoothness assumptions (different desirable n th order derivative properties). Well-known and heavily utilized splines include cubic splines, Bezier curves, B-splines and different specific instantiations, e.g., *non-uniform rational B-splines* (NURBs). Splines are engineered and used extensively in fields like computer graphics. Relative to our current article, each sensor provides a set of sample locations and confidences (output of prescreeners). These are our known *control points*. From this point specification, the objective is to calculate a spline that governs the surface that these points were drawn from. However, depending on spline, parameters of a spline function, and other factors, we have flexibility in specifying if the resultant surface must intersect all of the control points or instead use them as guidance in creating a surface with certain desirable properties (e.g., continuity of derivatives). Figure (4) illustrates a simple linear spline and a n th order spline.



Figure 4. Illustration of a simple linear (left) and n th order spline (right). Blue dots are our specified control points.

Our first approach was to take the output of each prescreener and a C^2 continuous biharmonic spline was built (specific implementation in Matlab is the `griddata` function with method 'v4' applied). After the spline is computed we use it to calculate the value of each sample in the grid. Figure (6)(a) shows the result.

Figure (6)(a) is the challenge that we discovered. If we just use the sensor samples then the spline solves relative to them and maintaining its continuity conditions. Whereas this sounds correct, the result is often an *over inflation* of confidences in regions where little-to-no sensor observations exist (illustrated in Figure (7)). Furthermore, in a prescreener we often threshold its result based on *receiver operating characteristic* (ROC) curve analysis. Typically, prescreeners are the input to more complicated feature extraction, pattern recognition/machine learning and data fusion algorithms. By reducing the number of locations to look at, we can obtain more efficient systems versus running these complicated and *resource* intensive operations every sample location and it sometimes also helps reduce the overall FAR. So, using a threshold means that we throw away low confidence samples. Ultimately, this means that we have lowered our sensors sampling by discarding samples in low confidence regions. This further amplifies the over inflation phenomenon. As a result, we explored two remedies. The first option is to not apply a threshold to a prescreener. However, depending on the sensors sample rate this can result in a lot of data is computationally challenging to store and compute our spline with respect to. The second option we explored is based on the fact that some prescreeners do not assign a value to each sensor sample. Instead, some prescreeners go through a series of operations that result in modified/new locations that were not truly observed by the sensor, e.g., possibly due to an operation like clustering to reduce FAs and/or improve localization. In many cases we are therefore forced to work with a reduced and/or altered hit list. As such, we identify all re-sampled grid points that are outside a user specified halo, e.g., the FOV of our sensor, of prescreener hits. For sake of discussion, we refer to this set of non-data supported grid samples *grounding points*. We calculate the spline with respect to our prescreener outputs and zero values at the grounding points (see Figure (8)). Figure (6)(b) shows the result of computing our spline with respect to both prescreener and grounding points.

In closing, we end with a few examples of surface re-sampling for GPR and EMI (Figures (9) and (10)).

4. DECISION-LEVEL FUSION

It is well-known that EMI and GPR detect different buried objects and it is of benefit to combine (aka fuse) their information (see Figure (11)). The general hope is that the fusion of these two sensors, with respect to precise positioning, is better than the utility of the sensors by themselves. However, fusion can take place at different so-called "levels", e.g., signal/spectrum, feature, algorithm or decision. Herein, we focus on decision level fusion as the objective is to combine the evidence (outputs) from multiple prescreeners. In the following subsection, we

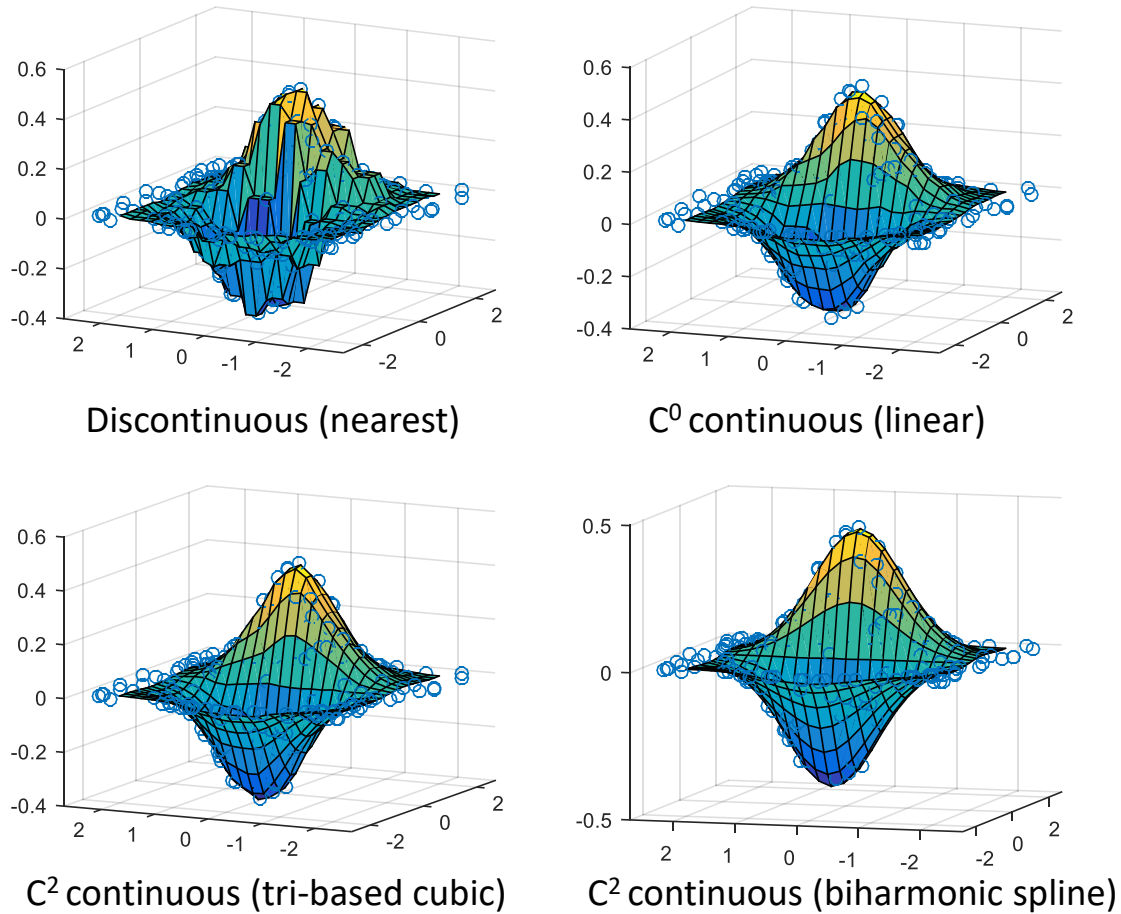


Figure 5. Synthetic example of nearest, linear, tri-based cubic and biharmonic spline techniques. Induced surfaces are shown relative to the specified set of control points (blue circles).

outline a procedure for using the *Choquet integral* (ChI) for prescreener alarm fusion. Furthermore, we outline a procedure for fusing multiple ChIs, where each ChI is tailored to a different subset of sensors and/or algorithms.

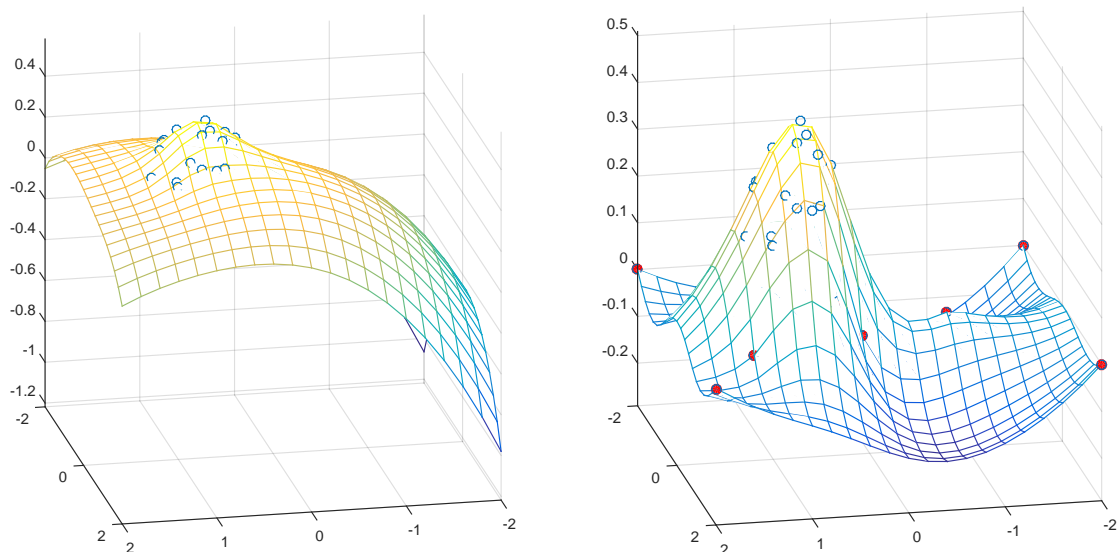
4.1 Choquet Integral (ChI)

The ChI is a flexible and parametric aggregation operator. The ChI is defined with respect to the FM. Let $X = \{x_1, x_2, \dots, x_N\}$ be a set of N inputs from sources like experts, algorithms and/or sensors. A FM is a monotonic set-valued function defined on the power set of X , 2^X , as $\mu : 2^X \rightarrow \mathbb{R}^+$ that satisfies the following two properties: (i) Boundary condition, $\mu(\emptyset) = 0$, and (ii) monotonicity, if $A, B \subseteq X$ and $A \subseteq B$, $\mu(A) \leq \mu(B)$. Often an additional constraint is imposed on the FM to limit the upper bound to 1, i.e., $\mu(X) = 1$. Throughout this paper, without loss of generality, we consider this condition for simplicity and convenience.

Let $h(x_i)$ (or h_i) be the information (beliefs, evidence, sensor measurements, etc.) provided by input i . The discrete (i.e., finite X) ChI is

$$\int_C h \circ \mu = C_\mu(h) = \sum_{i=1}^n h(x_{\pi(i)}) [\mu(A_i) - \mu(A_{i-1})], \quad (1)$$

where π is a permutation of X , such that $h(x_{\pi(1)}) \geq h(x_{\pi(2)}) \geq \dots \geq h(x_{\pi(n)})$, $A_i = \{x_{\pi(1)}, \dots, x_{\pi(i)}\}$, and $g(A_0) = 0$. Based on selection of μ , the ChI turns into a specific aggregation operator. For example, when



(a) Spline for prescreeener hit set

(b) Spline for prescreeener and grounded points

Figure 6. Example of computing a continuous biharmonic spline with respect to just prescreeener hits (left) versus prescreeener hits (blue) and zero values at grounded (red) grid locations (right).

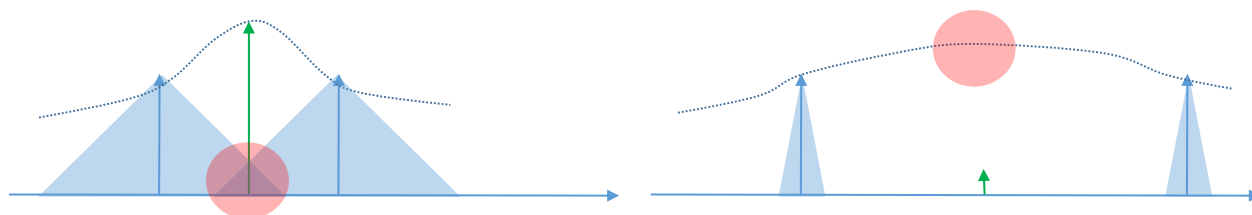


Figure 7. Illustration of the over inflation phenomenon encountered during spline surface construction. The left figure shows the spline (black line) for two samples (vertical blue arrows with shaded blue triangular FOVs). The true target location and confidence is shown by the green arrow. It has not been directly observed by the sweep of the sensor but it has been partially seen due to the sweep and sensor FOV (shown in red). This illustrates that the spline has the potential to help us connect these two off target samples and better resolve the location (peak) of the target. The right figure shows two samples that are not close to each other, with respect to the sensors FOV, and the potential for the spline to, if not grounded, created over inflated values (red circle) in regions where the prescreeener said there is low-to-no confidence or no known observations exist (green arrow indicates the true confidence).

$\mu(A) = 0, \forall A \in 2^X \setminus X$, the ChI is equivalent to the minimum operator. When $\mu(A) = 1, \forall A \in 2^X \setminus \emptyset$, we obtain the maximum operator. More generally, when $\mu(A) = \mu(B) \forall A, B \in 2^X$ such that $|A| = |B|$, we recover the familiar class of *linear combinations of order statistics* (LCOS), e.g., min, max, soft min and max, mean, median, trimmed statistics, etc.

Before we delve into detail about how to extend the ChI, we first quickly review where the FM comes from. The FM can be specified by an expert, however since there are 2^N terms, this becomes intractable quickly. If one has training data, then a number of methods can be used to learn the FM (see¹⁵). However, another popular solution is to work with what we can obtain and to approximate the rest (under a set of assumptions). Often, the measure on the singletons, $\mu(x_i)$, are referred to as densities. If the densities are known or can be learned, then an “imputation function” can be used to fill in the remainder of the FM. A famous imputation function is Sugeno’s λ -FM.¹⁶ The Sugeno λ -FM finds a unique solution for λ based on the densities and Sugeno’s famous characteristic polynomial function. Other well-known imputation functions are the S-Decomposable FM, of which S equal to the maximum operator is a pessimistic propagation of values up the FM lattice. The point

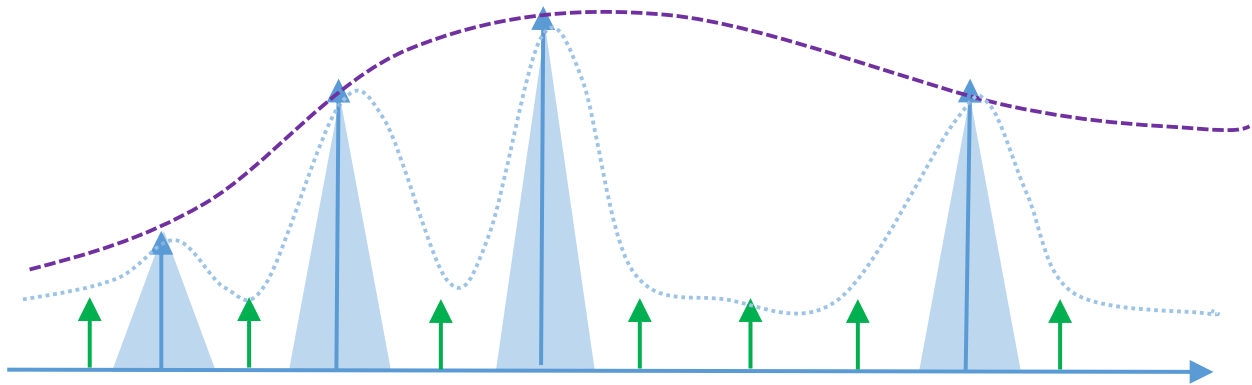


Figure 8. Illustration of grounding in spline surface production. The samples and their FOVs are shown in blue, green is grounded points, purple is the non-grounded spline result and blue line is the grounded spline result.

is, a FM can be specified by a human, learned from data or imputed.

4.2 Arithmetic Combinations and Compositions of ChIs

Whereas the ChI is a powerful and flexible parametric function that has been applied successfully to numerous applications, it is limited in mathematical expressiveness. In particular, the ChI is not well equipped to tailor different aggregation philosophies, e.g., optimistic (union like), mean-like (expected value) or pessimistic (intersection like), to different subsets of inputs. Herein, we investigate the application of different ChIs to different subsets of inputs. Furthermore, we also consider the ChI of these ChIs and arithmetic combinations of such results. Figure (12) illustrates this concept according to its corresponding underlying tree form.

Herein, we have a limited number of inputs (prescreener algorithms). As such, we manually explore different aggregations of ChIs to see if there is any merit in this data fusion concept. In future work we will investigate methods to learn such a solution from data.

5. EXPERIMENTS

In this section, we investigate the fusion of the following four prescreeners; Curvelet filtering with mean shift clustering on GPR (A_1), BAD filtering with *Fishers linear discriminate analysis* (FLDA) features on GPR (A_2), global ACE with mean shift clustering on EMI (A_3) and global ACE with connected components on EMI (A_4). Each algorithm produces an alarm set, easting and northing world space locations with associated algorithm confidence. These confidences are not on the same scale. For example, A_1 values are a result matched filtering, size contrast filtering with Bhattacharya divergence and clustering. The confidence value in BAD is generated via FLDA. Ideally, in decision-level fusion, evidence is on a similar scale, both numerically and semantically. Numerically refers to the fact that they are on the same interval, e.g., $[0, 1]$. Semantically, which is much harder, refers to the fact that pairs of values in each are associated with the other (follow similar trends). Herein, we use training data to identify the mean and standard deviation for each prescreener algorithm. The alarms are then scaled such that the mean is now located at 0.5 and 0 and 1 are associated with three standard deviations. Outliers are clamped to the $[0, 1]$ interval endpoints. Whereas this ensures similar numeric scale, it does not ensure similar semantics (a focus of future work).

For sake of completeness, we quickly summarize this experiment and data set. A total of six runs were collected. Each run is a different lane with various metallic and non-metallic explosive hazards of varying size, materials and burial depths. In total, there are seventy targets in the six runs. The data sets were collected independently for the two sensors and then registered to created the input data. Two scripted platforms with precise positioning were used to scan the lanes. As such, the exact sweep pattern and region of coverage of the two sensors are not the same. In order to fuse these two sensors, we re-sampled the prescreeners at a sample

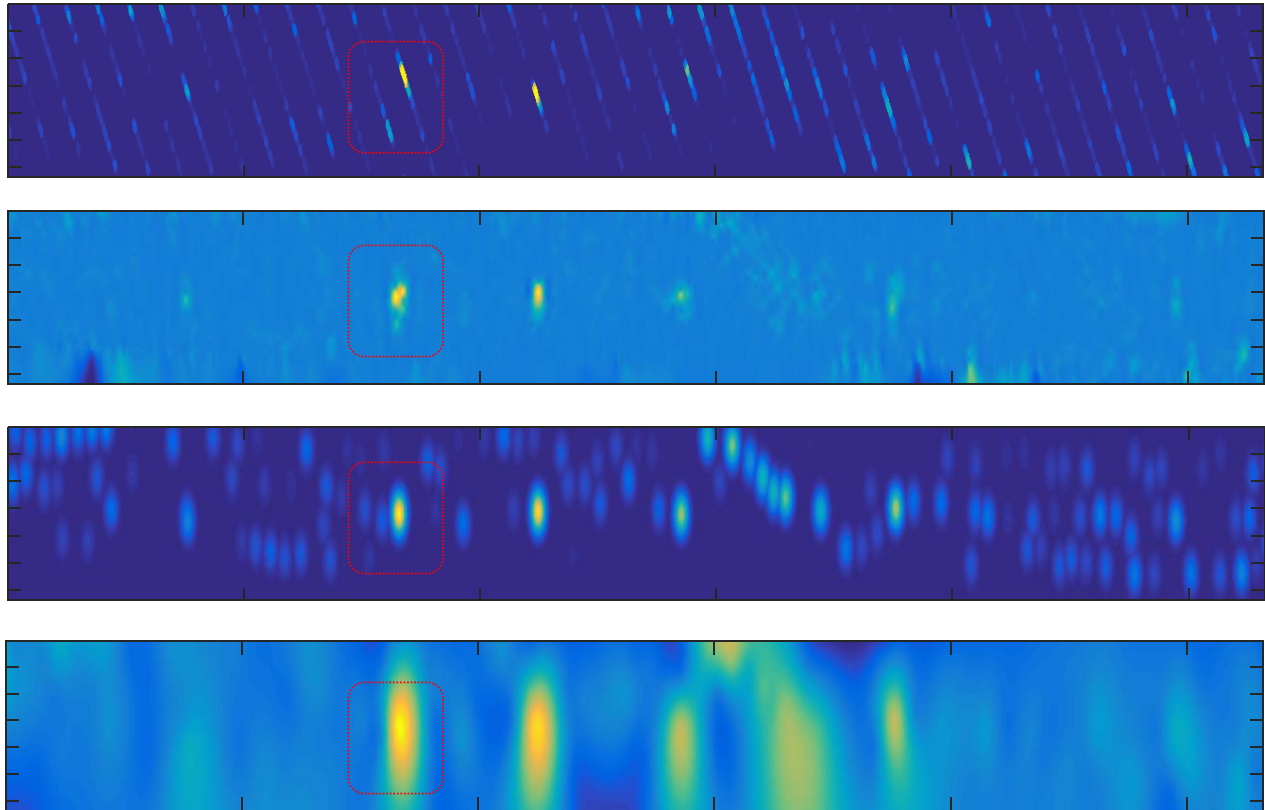


Figure 9. Grid re-sampling examples on GPR. The left of each image is the start and the right is the end of the sweep. The road was swept in the cross-track direction, the cart advanced (down-track) and then swept again in the opposite direction (cross-track). This pattern was repeated. The top row (image) is closest re-sampling for the Curvelet prescreener (no mean shift clustering). The second row is spline (no grounding) interpolation for the full set of Curvelet prescreener confidences. The third row is Gaussian re-sampling for a threshold applied to the Curvelet prescreener (mean shift clustering applied). Last (bottom) row is spline (no grounding) on the threshold applied Curvelet prescreener hits. Note, beyond the obvious issues due to not grounding the spline, the closest (top row) contains an instance in which the target was not directly on the path of a sweep but between sweeps (red dashed rectangle). The second row, spline, combined these two detections into a single point/alarm. The third row shows that the mean shift clustering combined those two detections as well.

distance rate of 1cm. At each point on this common sampled surface, we perform fusion. From that geo-spatial result, a simple extrema finder is used to generate a final set of alarms to score with. There is not enough data to run a comprehensive set of cross validation experiments. In part, this is due to the complexity and cost of collecting such data using two different sensors under development at a U.S. Army test site. Each collection takes one to more weeks of teams of experts running hours each day. In the future, if/when more data is collected, we will run additional experiments and build more comprehensive results.

In Figures (13) and (14), we report *receiver operating characteristic* (ROC) curve performance. We report PDR bars, where each FAR is treated as a binomial experiment and the PDR is treated as probability of success. The 95% binomial confidence interval for each PDR is generated and its curve it plotted. In general, a ROC shows the PDR/FAR behavior of an algorithm with respect to some parameter. Here, that parameter is the output of the prescreener, ChI or aggregation of ChIs. The x-axis is the number of mistakes, described in terms of FAs per meter squared. The y-axis is the PDR. Ideally, we would like a 0 FAR and 1 PDR curve. However, in reality, as we change our threshold—the binary decision associated with calling anything above that value a target—we obtain a different PDR relative to some FAR. In some cases we are given a maximum FAR and we

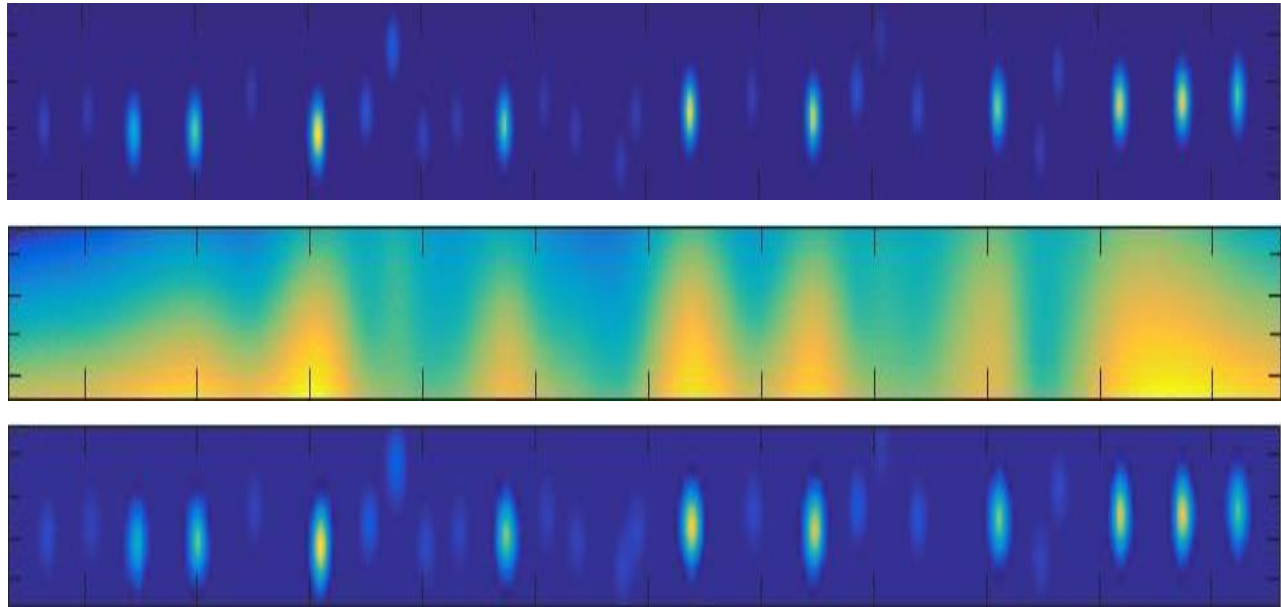


Figure 10. Grid re-sampling examples on EMI. The road was swept in the same fashion as described in Figure (9). The top row (image) is global ACE post mean shift clustering. The second row is the spline (no grounding) of global ACE with mean shift. The last (bottom) row is a spline applied to grounded global ACE with mean shift.

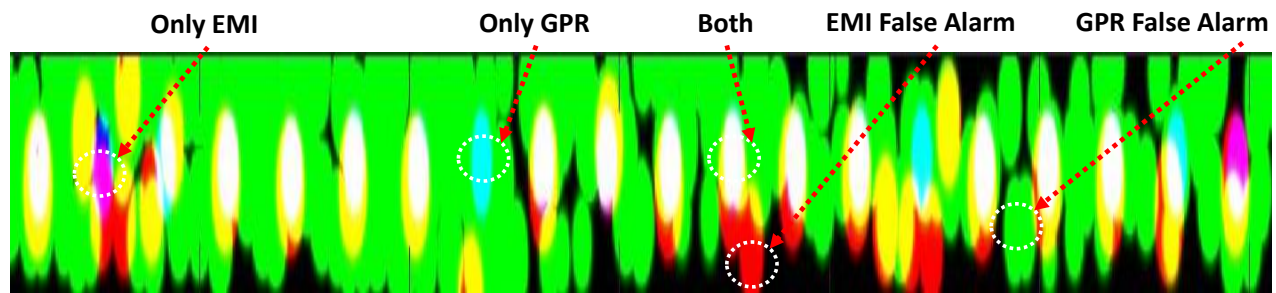


Figure 11. Example of GPR and EMI prescreener hits for a lane from our data set. Circles are drawn with respect to the FOV of the sensors and the result is shown as an RGB image; where red is EMI, green is GPR and blue is the ground truth. Thus, white is regions of agreement, magenta is EMI only and cyan is GPR only.

are asked what is the best PDR. In other instances, we are asked for a certain PDR and the question is what FAR is associated with that value? Typically, one looks at a ROC and searches for “runs”. We want the ROC to climb fast, i.e., make few mistakes to get more detections. However, when we find long runs, horizontally flat or linearly increasing regions in the ROC, that means we have to accept a lot more FARs to get more detections or we are accepting as many mistakes as detections.

Figure (13) shows that EMI currently outperforms GPR by a good amount. We cannot show—it is too visually complex to co-plot—each individual prescreener and different aggregation operators on combinations of these algorithms. Instead, we report the top performing GPR algorithm, the top performing EMI algorithm, and the best manually identified aggregation of ChIs (relative to those two algorithms). We can see that fusion improves the PDR at each FAR (makes fewer mistakes) and it leads to an overall higher PDR (leads to more detections).

In Figure (14), we report three different ChIs—the minimum (intersection like), maximum (union like) and average (expected value). The reason for showing the intersection to union is an attempt to cover the “fusion

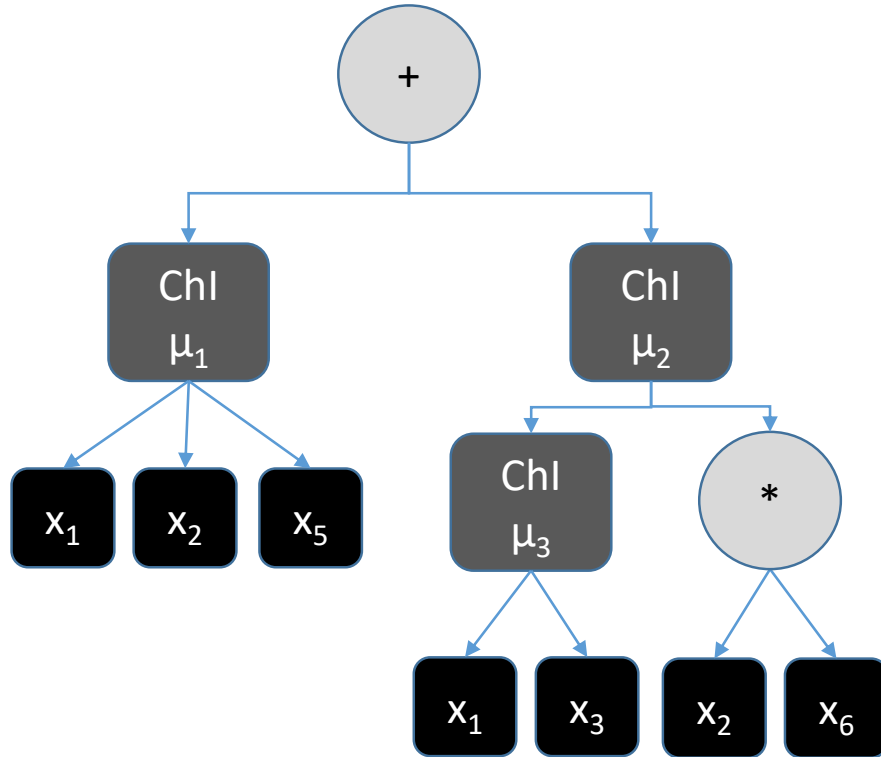


Figure 12. Example of the aggregation of three different ChIs for $N = 6$ inputs. This particular aggregation can be expressed as $C_{\mu_1}(\{h_1, h_2, h_5\}) + C_{\mu_2}(\{C_{\mu_3}(h_1, h_3), h_2 * h_6\})$.

spectrum” (optimistic to mean-like to pessimistic). Last, in Figure (14) we also report the aggregation of ChIs for all four inputs (A_1 to A_4). Overall, we see that a single ChI provides little-to-no benefit over just EMI. However, the aggregation of ChIs improves the PDR at each FAR and achieves an overall higher PDR. At that, the aggregation of ChIs for all four prescreeners produces the best ROC.

Overall, these ROCs show that the aggregation of ChIs has benefit and the fusion of GPR and EMI is advantageous. Furthermore, the best result involves different ChIs on the sensors and another ChI across sensors—which supports our claim that sometimes different combinations of algorithms and/or sensors will demand different aggregations. Last, it is worth noting that many of our targets are difficult-to-impossible to “see”. This is due to various reasons, i.e., the ability for GPR and EMI to detect such an object, its size, material properties, soil, the GPR and EMI systems (transmitted signals to antennas and range resolution), etc. The reason we state this is to express the concern that 100% is not realistic—given the current state of our EHHD platform. Therefore, in analyzing the ROCs we focus specifically on some early PDRs versus overall ROC curve trends that increase the PD relative to long runs in FAR.

As a final experiment, Figure (15) shows the performance of the different re-sampling approaches. Note, once again there are too many subsets of inputs (prescreeners) and aggregation of ChIs to consider. Herein, for tractability sake we fix all system parameters except re-sampling algorithm. The trends reported in Figure (15) are however consistent with what we have observed via our different experiments. What we see is the following. First, as we suspected, NN, bounded NN and bounded neighborhood with max are the worst performers. Next is the ungrounded spline, which we suspected due to the inflation phenomenon. Next best is closest. We suspected that closest would do well for this data in general because the sampling cross-track (not down-track) was fair and alignment appeared to be within an acceptable tolerance relative to our underlying grid sampling distance (1cm). The Gaussian approach is the next best. We suspected it would do well because the sensors did not always overlap and the Gaussian propagation of values allows us to link measurements between sensors relative

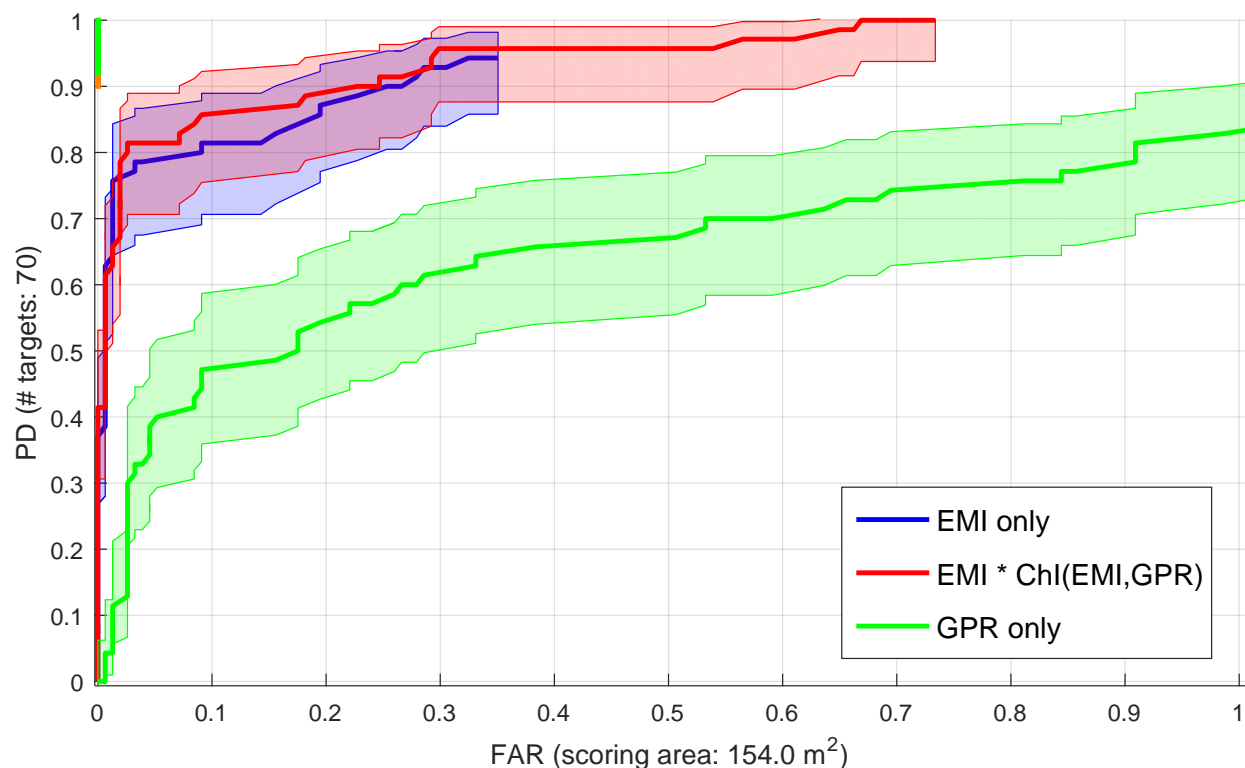


Figure 13. Comparison of best GPR prescreener (Curvelet with mean shift clustering) and best EMI prescreener (global ACE with mean shift clustering). The red ROC, $h_1 + C_{\mu=\max}(\{h_1, h_2\})$ (where h_1 is EMI and h_2 is GPR), is the best ChI fusion discovered manually relative to these two prescreeners.

to their FOVs. The best performer, as expected, was the grounded spline. However, the grounded spline and Gaussian often perform similarly. As we conducted manual visual analysis of the re-sampled grids, we did see instances where the spline linked alarms, better resolving a single improved alarm location versus two offset lower valued confidences in the Gaussian. Overall, these findings are in line with what we expected given the parameters of the system and our data collection.

In closing, whereas the spline did best overall we would like to note that the Gaussian is simpler to implement. For the Gaussian we do not technically have to build a surface. Instead, we can manually link and fuse on the fly with respect to our sensor data. That is, we identify all hits from the other sensors, with respect to sigma, and the confidences used in fusion are the Gaussian discounted versions of those other confidences. This is a lot less memory and a lot quicker, versus solving the spline.

6. CONCLUSION AND FUTURE WORK

Herein, we focused on the fusion of EMI and GPR for handheld EHD using an U.S. Army provided EHHD platform. Specifically, we explored seven different re-sampling techniques and spline interpolation with grounding showed the most promise with respect to our setup. In addition, the best fusion result was not one of the individuals prescreeners nor a single ChI solution. Instead, the top performer is the aggregation of different ChIs within and across sensors. In the context of EHD, these results are interesting and very encouraging.

In future work, we will look to learn, versus specify, the aggregation of ChIs. We will also investigate methods for semantically, versus just numerically, the prescreener outputs for decision level fusion. Next, we will further analyze and compare the performance versus resource usage of different re-sampling algorithms relative to the goal of creating a low price real-time operational handheld platform.

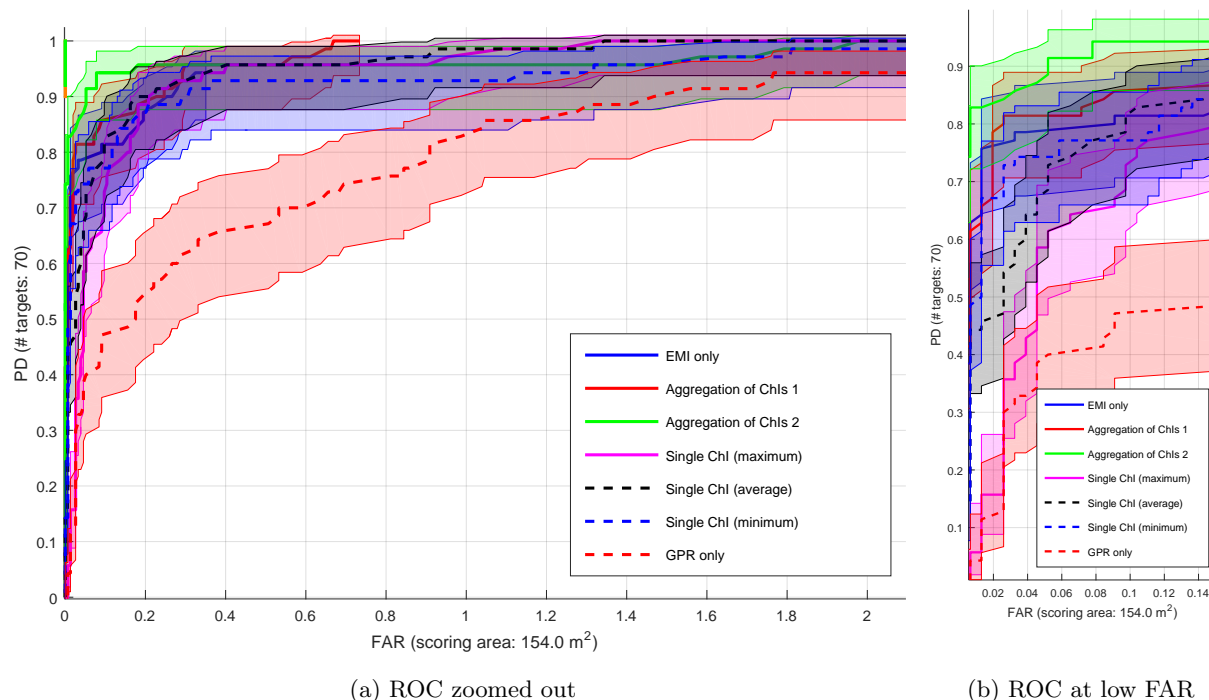


Figure 14. Comparison of best individual GPR and EMI prescreeners (same ones reported in Figure (13)), maximum, minimum and average ChIs, aggregation of ChIs reported in Figure (13) (red) and the aggregation of ChI fusion of all four prescreeners (A_1 , A_2 , A_3 and A_4) (green). The latter, which is the overall winner, is $C_{\mu_1=\max}(\{C_{\mu_2=\max}(\{h_1, h_2\}), \{C_{\mu_3=\max}(\{h_3, h_4\})\})$ (h_1 and h_2 are GPR and h_3 and h_4 are EMI prescreeners).

ACKNOWLEDGMENTS

This material was supported by Army Research Office grant W911NF-14-1-0673 in support of the US Army RDECOM CERDEC NVESD.

REFERENCES

- [1] F. Cremer, K. Schutte, J. G. Schavemaker, and E. den Breejen, "A comparison of decision-level sensor-fusion methods for anti-personnel landmine detection," *Information fusion* **2**(3), pp. 187–208, 2001.
- [2] D. J. Daniels, "A review of gpr for landmine detection," *Sensing and Imaging: An international journal* **7**(3), pp. 90–123, 2006.
- [3] T. C. Havens, K. Stone, D. T. Anderson, J. M. Keller, K. Ho, T. T. Ton, D. C. Wong, and M. Soumekh, "Multiple kernel learning for explosive hazard detection in forward-looking ground-penetrating radar," in *SPIE Defense, Security, and Sensing*, pp. 83571D–83571D, International Society for Optics and Photonics, 2012.
- [4] T. C. Havens, K. Stone, J. M. Keller, and K. Ho, "Sensor-fused detection of explosive hazards," in *SPIE Defense, Security, and Sensing*, pp. 73032A–73032A, International Society for Optics and Photonics, 2009.
- [5] K. Ho and P. D. Gader, "A linear prediction land mine detection algorithm for hand held ground penetrating radar," *IEEE Transactions on Geoscience and Remote Sensing* **40**(6), pp. 1374–1384, 2002.
- [6] K. Ho, L. M. Collins, L. G. Huettel, and P. D. Gader, "Discrimination mode processing for emi and gpr sensors for hand-held land mine detection," *IEEE Transactions on Geoscience and Remote Sensing* **42**(1), pp. 249–263, 2004.
- [7] M. A. Lee, D. T. Anderson, J. E. Ball, and J. L. White, "Background adaptive division filtering for hand-held ground penetrating radar," in *SPIE Defense+ Security*, **9823**, pp. 98230I–98230I–11, 2016.

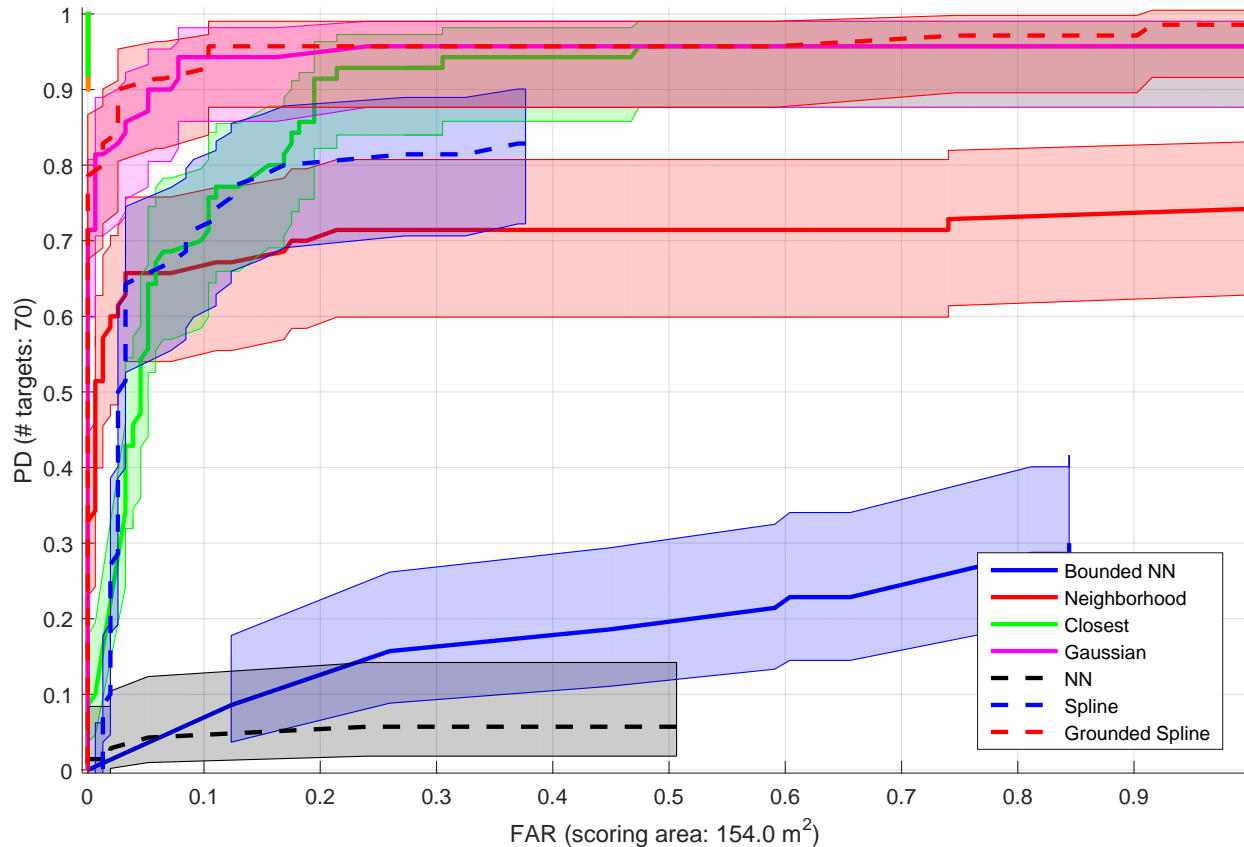


Figure 15. Comparison of seven proposed re-sampling methods relative to the best aggregation of ChIs solution (winner shown in Figure (14)).

- [8] J. L. White, D. T. Anderson, J. E. Ball, and B. Parker, "Curvelet filter based prescreener for explosive hazard detection in hand-held ground penetrating radar," in *SPIE Defense+ Security*, pp. 98230H–98230H, International Society for Optics and Photonics, 2016.
- [9] B. Alvey, A. Zare, M. Cook, and D. K. Ho, "Adaptive coherence estimator (ace) for explosive hazard detection using wideband electromagnetic induction (wemi).," *CoRR* **abs/1603.06140**, 2016.
- [10] C. De Boor, C. De Boor, E.-U. Mathématicien, C. De Boor, and C. De Boor, *A practical guide to splines*, vol. 27, Springer-Verlag New York, 1978.
- [11] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, pp. 603–619, May 2002.
- [12] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Information Fusion* **14**(1), pp. 28 – 44, 2013.
- [13] G. Toussaint, "Solving geometric problems with the rotating calipers," 1983.
- [14] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The theory of splines and their applications*, 1967.
- [15] D. Anderson, S. Price, and T. Havens, "Regularization-based learning of the choquet integral," in *Fuzzy Systems (FUZZ-IEEE), 2014 IEEE International Conference on*, pp. 2519–2526, July 2014.
- [16] M. Sugeno, "Theory of fuzzy integrals and its applications," *Ph.D. thesis Tokyo Institute of Technology*, 1974.