SPATIAL LATENCY REDUCTION IN GPR PROCESSING USING STOCHASTIC SAMPLING

Peter Torrione and Leslie Collins

Duke University ECE Departiment
Box 90291
Durham NC, 27701

1. INTRODUCTION

Ground penetrating radar (GPR) provides a complimentary phenomenology to electro-magnetic induction based landmine detection, and recent advances in radar fabrication and related signal processing have significantly improved landmine classification performance using GPR sensors (e.g. [1, 2, 3, 4, 5, 6, 7, 8]).

However, many of the algorithms that achieve the most robust performance for target identification in GPR data make use of significant portions of data that are spatially located "past" the center of a potential landmine response - i.e. the algorithms are not strictly causal with regards to the spatial data used to make declarations at a particular location [1, 2, 3, 4, 5, 6]. When collecting data with an overpass-capable vehicle, non-causality is not a significant issue as long as declarations can be made before a threat marking system passes over the locations of interest. However, in non-overpass situations, algorithm latency can have a significant effect on the ability of a vehicle to stop prior to encountering a buried explosive threat, especially given variable vehicle speeds and human reaction times.

In this work we propose a technique for reducing the spatial latency of pre-existing landmine classification algorithms using stochastic sampling. We show that the proposed approach can maintain the performance capabilities of the original classification algorithms while significantly reducing the average spatial latency required to make a declaration.

2. LATENCY CONSIDERATIONS AND SCOPE OF THIS WORK

This work focusses on the reduction of algorithmic "spatial latency" in vehicular mounted antenna array systems. We define "spatial latency" as the number of radar responses that are required to make an alarm declaration and which are collected after the radar has passed over the center of a possible threat. Note that in addition to the spatial latency of an algorithm, additional temporal latency is required to 1) process the resulting feature sets, 2) alert the user, and 3) stop the vehicle. Temporal sources of latency are not considered in this work.

A schematic diagram of the situation encountered for vehicular mounted radars can be seen in Fig. 1. In Fig. 1, a vehicle moving from left to right encounters a landmine response at spatial location P3, when the vehicle wheels are at point P1. Due to the spatial latency of the algorithms under consideration, the user can not be notified of the presence of the target until the radar antennae has reached point P4, when the wheels will be at a point between P1 and P3. Once a decision can be made, depending on the vehicle speed, there is a limited amount of time for the user to notice the alarm and stop the vehicle before the vehicle wheels reach P3. The goal of this work is to develop techniques to reduce the algorithm-specific required distance between the target, P3, and point P4, thus increasing the distance between the vehicle wheels and point P3 when a declaration can be made, and providing the user more time to stop the vehicle.

3. STOCHASTIC SAMPLING TO REDUCE LATENCY

Reducing the amount of data required to make a decision has been studied previously, in particular [9] provides an efficient technique for calculating a "sequential probability ratio test" under the conditions that the distributions of data under the hypotheses are known, and the data being collected is independent and identically distributed (i.i.d.). Under these conditions,

This work was supported under a grant from the Army's Night Vision And Electronic Sensors Directorate.

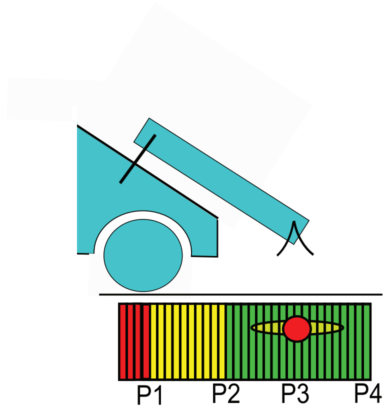

Fig. 1. Schematic diagram of sensor platform wheel and radar (light blue), landmine (yellow oval), and landmine trigger (red circle) with effects of latency. The sensor platform is moving from left to right over the landmine. The area below the radar is divided into equal-distance "bins" corresponding to the data collected. Shading below the ground illustrates the spatial latency of an algorithm about a target center (green, P2 to P4), the amount of space/time left for the system to make a declaration, alert the user, and stop (yellow, P1 to P2), and the area after which a blast event will have occurred (red, left of P1).

techniques for optimal sampling and decision making to achieve a given probability of detection or false alarm rate are known and tractable.

Unfortunately, in the case of multiple spatial measurements from landmine responses, the distributions of the data under the various hypotheses are not well known, and the data are not i.i.d. Further, for many advanced algorithms, the classification procedures can be extremely complex [1, 2] and training algorithms may include various heuristic rules which are not easily generalizable to new training sets with reduced spatial latency. As a result, in this work, we consider the case when a classification algorithm is provided to the user as a "black box", and no information regarding the algorithm (other than the features utilized) is available.

Consider a set of N feature vectors x_j from a particular target or non-target location,

$$X_N = \{x_1, x_2, ..., x_N\}.$$

Here N represents the maximum number of down-track samples we can collect before we are required to make a decision, and our goal is to make a decision using

$$X_M = \{x_1, x_2, ..., x_M\}$$

where $M \leq N$. Assume that we have an algorithm $\mathcal{A}(X)$ such that

$$p(H_1|X_N) \approx \mathcal{A}(X_N), \tag{1}$$

i.e. a developer has provided an algorithm $\mathcal{A}$ whose output approximates the posterior probability of a target given all the data allowable. Examples of such algorithms from the landmine detection literature include EHD based approaches [1], HMM based approaches [3], and the FOWA algorithm [2]. Let us also assume that although it is difficult to calculate $p(H_1|X_M)$ for $M < N$, it is possible to calculate or estimate the conditional distribution $p(X_{N\setminus M}|X_M)$ where

$$X_{N\setminus M} = \{x_{M+1}, x_{M+2}, ..., x_N\}$$

.

Under these assumptions, one can estimate posterior probability of H_1 using

$$p(H_1|X_M) = \int_{X_{N\setminus M}} p(H_1|X_M, X_{N\setminus M}) p(X_{N\setminus M}|X_M) dX_{N\setminus M}. \tag{2}$$

This can be approximated using $\mathcal{A}(X)$, and the fact that $X_N = \{X_M, X_{N\setminus M}\}$,

$$p(H_1|X_M) \approx \int_{X_{N\setminus M}} \mathcal{A}(X_N) p(X_{N\setminus M}|X_M) dX_{N\setminus M}. \tag{3}$$

Since a functional form of $\mathcal{A}(X_N)$ will not in general be known, this can be further approximated using monte-carlo integraton,

$$p(H_1|X_M) \approx \frac{1}{N_s}\sum_{j=1}^{N_s} \hat{p}_j(H_1|X_M) \tag{4}$$

$$\hat{p}_j(H_1|X_M) = \mathcal{A}(\{X_M, \hat{X}_{j,N\setminus M}\}), \tag{5}$$

where $\hat{X}_{j,N\setminus M}$ represents one of N_s samples from $p(X_{N\setminus M}|X_M)$.

In addition to calculating an approximation to $p(H_1|X_M)$ using the average of the $\hat{p}_j(H_1|X_M)$, one can also consider each of the N_s samples of $\hat{p}_j(H_1|X_M)$ as a draw from the posterior distribution on $p(H_1|X_M)$. This enables an algorithm to not only estimate the probability of a target being present given a limited sub-set of data, but also to estimate the variance on the posterior estimate and determine whether it will be effective to continue sampling using both the mean and variance of the current estimates $\hat{p}_j(H_1|X_M)$.

Previously we assumed that $p(X_{N\setminus M}|X_M)$ was available. In actuality, estimating $p(X_{N\setminus M}|X_M)$ from training points can pose a difficult problem. In this work we assume that all elements of X_N are jointly normally distributed, and sample from $p(X_{N\setminus M}|X_M)$ using this assumption, and we estimate parameters for the joint distribution using training data. Future work will explore modifications to this density estimation approach, keeping in mind both computational concerns and classification performance metrics.

4. RESULTS AND CONCLUSIONS

In this work we applied the stochastic sampling algorithm described above to a version of the EHD algorithm [1] where the classification algorithm specified in [1] was replaced with a relevance vector machine (RVM) [10, 11]. The EHD algorithm feature extraction makes use of 7 spatial bins collected around a potential target location. Thus, data collected after the 3rd spatial EHD bin adds to the non-causality of the algorithm and reduces the amount of time the user has to respond to an alarm.

Fig. 2 illustrates a performance comparison for the proposed stochastic sampling algorithm at various latencies, with M varying from 1 to 7 down-track bins (only results for 1, 3, 5, and 7 bins are shown for clarity). The data used in this work was collected at 3 U.S. test facilities over the course of 3 years, with varying soil and weather conditions. The target populations under consideration are weighted towards difficult low-metal anti-tank landmines.

For each value of M, and each alarm, 100 $\hat{p}_j(H_1|X_M)$ values were generated using stochastic sampling. For each alarm, and each value of M, a decision can be made to re-sample with a larger value of M or to make a declaration. Declarations are made when 99% of the mass of the distribution of $\hat{p}_j(H_1|X_M)$ resides on one side of a decision threshold (0.5 in this case). As a result of this processing, different numbers of alarms were flagged at different values of M; this corresponds to an adaptive latency which changes depending on the amount of information present in X_M. The average latency for the data shown in Fig. 2 is 2.6 down-track bins, which is a significant reduction from the 7 bins used in the static EHD/RVM classifier.

Continuing work will address the performance of the proposed algorithm in aggregate, and also consider varying the parameters used to determine when decisions should be made. Further, we have so far only briefly discussed the issue of estimating $p(X_{N\setminus M}|X_M)$, which can pose a significant challenge. Our work will explore various approaches to estimating the density, and compare and contrast performance for varying estimation procedures.

5. REFERENCES

[1] H. Frigui and P. Gader, "Detection and discrimination of land mines in ground-penetrating radar based on edge histogram descriptors and a possibilistic k-nearest neighbor classifier," *Fuzzy Systems, IEEE Transactions on*, vol. 17, no. 1, pp. 185–199, Feb. 2009.

[2] Paul Gader, Wen-Hsiung Lee, and Joseph N. Wilson, "Detecting landmines with ground penetrating radar using feature based rules, order statistics and adaptive whitening," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 11, pp. 2522–2534, 2004.

[3] Paul D. Gader, Miroslaw Mystkowski, and Yunxin Zhao, "Landmine detection with ground penetrating radar using hidden Markov models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 6, pp. 1231–1244, June 2001.

[4] K.C. Ho and Paul Gader, "A linear prediction land mine detection algorithm for hand held ground penetrating radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 6, pp. 1374–1384, 2002.

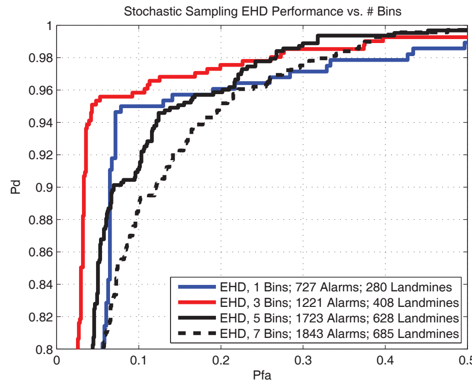

Fig. 2. Classification performance for the stochastic sampling algorithm applied to EHD features for varying numbers of down-track feature extraction bins ($1 \leq M \leq 7$). For each value of M, every alarm is considered, and decisions are made when the probability of changing a decision given the next observation is less than 0.001. As a result the different ROC curves in this plot contain different numbers of landmine and false alarm responses, and care must be taken when comparing across the ROC curves.

[5] Peter Torrione and Leslie Collins, "Texture features for antitank landmine detection using ground penetrating radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 2374–2382, 2007.

[6] Peter Torrione, Chandra Throckmorton, and Leslie Collins, "Performance of an adaptive feature-based processor for a wideband ground penetrating radar system," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 2, pp. 644–658, 2006.

[7] Joseph N. Wilson, Paul Gader, Wen-Hsiung Lee, Hichem Frigui, and K.C. Ho, "A large-scale systematic evaluation of algorithms using ground-penetrating radar for landmine detection and discrimination," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 2560–2572, 2007.

[8] Xiaoyin Xu, E. L. Miller, C. M. Rappaport, and G. D. Sower, "Statistical method to detect subsurface objects using array ground-penetrating radar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 4, pp. 963–976, April 2002.

[9] Abraham Wald, "Sequential tests of statistical hypotheses," *Annals of Mathematical Statistics*, vol. 16, pp. 117186, 1945.

[10] Michael E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.

[11] Mario A.T. Figueiredo, "Adaptive sparseness for supervised learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1150–1159, September 2003.