Extending Ground Penetrating Radar Imaging Capabilities Through Signal Processing

Roger Tilley, Hamid Sadjadpour, Farid Dowla
Department of Electrical Engineering University of California
Santa Cruz, CA, 95064, USA
Rtvax@soe.ucsc.edu; Hamid@soe.ucsc.edu; Dowla@soe.ucsc.edu

Abstract – Ground Penetrating Radar (GPR) Imaging is, and has been, a challenge to perfect on deeply buried targets. Developing a method to choose the best single frequency to image an object at depth has taken a back seat to developing better compositing techniques. Several methods have been explored to define a weighting structure to be applied to data taken at many frequencies to solve this problem. None of the methods have explored using an optimization problem solver. In this paper, we explore using the Expectation-Maximization Algorithm (EM), an optimization problem solver, to define the weights used to combine multiple frequency GPR scans of the same target area. Our method exploits the Gaussian Mixture Model feature of the EM algorithm, and shows promise as a measured improvement toward producing a sharper image at depth.

Keywords: Ground Penetrating Radar, Expectation-Maximization, Gaussian Mixture Model, Maximum Likelihood Parameter Estimation, Finite Difference Time Domain Method, GprMax

1. Introduction

Ground Penetrating Radar (GPR) signal scans are used to illuminate various terrain types and objects at various depths. The frequency for the best illumination is different for each depth one wishes to scan. In general, higher frequencies image objects and terrain closer to the surface in great detail while lower frequencies image objects deeper with less fidelity. Developing a way to combine high and low frequencies could have the effect of increasing the resolution of the combined signal to a lower depth. Determining how to weight each frequency signal to be combined for an optimal result, poses as an optimization problem to solve. In the literature three authors examined methods to solve this problem with varying success, all with a very similar approach. Methods by Dougherty[1], Booth[2] and Bancroft[3] all discussed ways to weight each signal used in the sum of each frequency. Absent from these works are optimization problem solvers. We have chosen to investigate the possibility of using the data mixture feature of the Expectation and Maximization (EM) Algorithm to develop optimal weights.

In this paper, we describe, briefly the methods of Dougherty, Booth and Bancroft and their relative successes. We cover in depth the EM Algorithm data mixture feature as an optimization problem solver. The paper is organized as follows. In section 2 we describe the Dougherty process in developing signal weights. In section 3 we discuss the methods examined by Booth. In section 4, we discuss the methods proposed by Bancroft. In section 5, we describe the EM Algorithm process as it applies to mixture models. In section 6 we present results and in section 7, draw some conclusions.

2. Dougherty approach

Dougherty first aligned each trace by the direct arrival pulse in each trace, and then removed DC shifts, the low frequency "wow" component, followed by scaling each trace by the L2 norm of the direct arrival pulse. The traces were summed and the resultant direct arrival estimate was then subtracted from each trace removing the direct arrival signal. An exponential gain recovery function was applied to each trace. Equal weighting was applied to each trace with the direct arrival signal removed as each frequency was summed. Dougherty demonstrated some clarity of shallow reflections due to direct wave removal. He achieved resolution and continuity of reflection enhancement by summing the frequencies.

Spectral bandwidth was increased as well. However, the resultant signal was overwhelmed by the lower frequencies in the summation.

3. Booth approach

Booth examines three methods of achieving an increased bandwidth and thus a more approximate delta function through composite GPR traces. The simple summation of Dougherty was examined, as one method. A second method, examined a scaled summation approach where the maximum value of each frequency spectra is determined and the spectrums equalized. The values used to equalize the spectra provide the signal weighting prior to summation. A third method, involved shifting traces such that the main peaks of the direct arrival pulses were aligned with the dominant peak then the scaled summation of method two was applied. Method three provided the best result for increased spectral bandwidth, thus the best delta function, and GPR resolution. As a final discussion, another weighting method was examined where the weighting factors were obtained from a least squares analysis, Optimal Spectral Whitening (OSW), that attempts to match the summed result to a defined optimal amplitude spectrum. The defined optimal spectrum determines over what set of frequencies the frequency data sets would be enhanced. A time-varying Fourier transformation of each data set must be performed prior to implementing the least squares analysis.

4. Bancroft approach

Bancroft discusses the findings of Dougherty and the methods described by Booth, while defining other methods to determine the weighting factors. One method uses a ramped summation method where the higher frequency data is suppressed by the same amount the lower frequency data is enhanced over the two-way transit time of a GPR scan. Bancroft discussed this double ramped summation technique using linear or Butterworth function ramps. Another method discussed by Bancroft is one he names the Amplitude Envelope Equalization (AEE) technique. With Automatic Gain Control (AGC) applied to each frequency data set, a multiplier is calculated as the ratio of the average envelope value of the lowest frequency and the average envelope value of a nominal frequency data set over the portion of time the multiplier is to be applied to the data set. The amplitude envelope is the calculated absolute value of the Hilbert transform of a single trace. All scans at one frequency have envelopes calculated then averaged. Determining the portion of time a ratio is applied is calculated by finding the minimum value of the log of the amplitude envelope (envelope computed without AGC applied); indicating the time where the suppression of that particular frequency data begins. This point is defined as data too attenuated to provide useful information.

Bancroft also suggests an alternative subjective method to determine the weighting through visual inspection of each frequency data set; but there is much less clarity as to how this is done and how experienced the reviewer must be.

5. Expectation Maximization Algorithm

The EM Algorithm, (Expectation- Maximization Algorithm) is often used to group like items contained in complex mixtures. Another use is to solve incomplete data problems by performing Maximum Likelihood (ML) Parameter estimation. An offshoot use for the EM algorithm is determining the membership weights of points in a cluster within a finite Gaussian mixture model [4][9]. It is this feature that will be exploited to combine several frequency scans into a composite wave. The entire data set can be represented by other mathematical distributions but we used Gaussian because it is often used when the distribution for the real-valued random variables is unknown.

We can define a finite mixture model $f(x; \theta)$ of K components as mixtures of Gaussian functions:

$$f(\underline{x}; \theta) = \sum_{k=1}^{K} \alpha_k p_k(\underline{x}| \theta_k) \tag{1}$$

Where:

- $p_k(\underline{x}|\theta_k)$ are K mixture components with a distribution defined over $p(\underline{x}|\theta_k)$ with parameters
  - $\theta_k = \{\underline{\mu_k}, C_k\}$ (mean, covariance)
  - $p_k(\underline{x}|\theta_k) = \frac{1}{(2\pi)^{d/2} |C_k|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu_k})^T C_k^{-1}(\underline{x}-\underline{\mu_k})}$
  - α_k are the mixture weights, where $\sum_{k=1}^{K} \alpha_k = 1$

- $\{\underline{x}_i, \ldots\ldots\ldots, \underline{x}_n\}$ Data set for a mixture component in d dimensional space

In each iteration of the EM Algorithm there are 2 steps, the Expectation step (E-step) and the Maximization step (M-step). In this case the E-Step computes the conditional expectation of the group membership weights $(w_{ik}'s)$ for $\underline{x}_i's$, adding unobservable data given θ_k. The M-Step computes new parameter values $\left(\alpha_k, \underline{\mu}_k, C_k\right)$ to maximize the finite mixture model using the membership weights. The E-Step and M-Step are repeated until stopping criteria is reached (convergence). Convergence is signaled by the log-likelihood of $f(\underline{x}; \theta)$ not appearing to change substantially from one iteration to the next.

E-Step –

$$w_{ik} = \frac{p_k(\underline{x}_i|\theta_k)*\alpha_k}{\sum_{m=1}^{K} p_m(\underline{x}_i|\theta_m)*\alpha_m}, \tag{2}$$

for $1 \le k \le K, \ 1 \le i \le N$; with constraint $\sum_{k=1}^{K} w_{ik} = 1$

M-Step –

$$N_k = \sum_{i=1}^{N} w_{ik} \tag{3}$$

$$\alpha_k^{new} = \frac{N_k}{N}, \text{ for } 1 \le k \le K \tag{4}$$

$$\underline{\mu}_k^{new} = \left(\frac{1}{N_k}\right)\sum_{i=1}^{N} w_{ik} * \underline{x}_i, \text{ for } 1 \le k \le K \tag{5}$$

$$C_k^{new} = \left(\frac{1}{N_k}\right)\sum_{i=1}^{N} w_{ik} * \left(\underline{x}_i - \underline{\mu}_k^{new}\right)\left(\underline{x}_i - \underline{\mu}_k^{new}\right)^{T} \tag{6}$$

Convergence (log likelihood of $f(\underline{x}; \theta)$) –

$$\text{Log } l(\vartheta) = \sum_{i=1}^{N} \log f(\underline{x}_i; \theta) = \sum_{i=1}^{N}\left(\log \sum_{k=1}^{K} \alpha_k p_k(\underline{x}_i|\theta_k)\right) \tag{7}$$

These equations that make up the EM Algorithm were implemented in MATLAB. The variable 'k' and 'x' represent the different scanning frequencies and GPR trace scan respectively. Each trace, at a frequency and data position, are analyzed and combined using the EM Algorithm before moving on to the next position.

6. Processing GPR Scan Examples

6.1. Test Case 1 (TC1)

As an example of the EM Algorithm capability, a series of six sine waves (50, 150, 250, 350, 450 and 550Hz) were constructed, all with the same magnitude. These sine waves were combined using the EM Algorithm noted above. The weights derived for combining these sine waves resulted in a square wave, not quite complete but well on the way. Fig. 1 shows the difference between the desired square wave and the EM Algorithm calculated sum [11]. As a test case, two fictional areas were defined using using a Finite Difference Time Domain (FDTD) [5][6][7] modelling software package. A proprietary package in development similar in operation to the popular GprMax software package by A. Giannopoulos [8] and GprMax were used to model the defined areas [10]. One area, consisted of a Transmitter (Tx) and Receiver (Rx) suspended 5 meters above the ground in air with a target (perfect electrical conductor) buried 10 meters below ground in a moist sand medium with relative permittivity (ε_r) of 9.0 (with a velocity in the medium of 0.1 m/ns) and an electrical conductivity of 0.001 mS/m. The transmitter and receiver were moved along the length of the defined space as shown in

Fig. 2, for a total of 36 scans at 0.25 meters per step. The Tx starts at 0.5 meters ending at 9.5 meters along the scan axis. The Rx is spaced 0.25 meters away from the Tx, starting at 0.75 meters ending at 9.75 meters. The space is 10 meters in length by 25 meters in depth with 5 meters of space above ground in a medium of air (free-space ε_r of 1, with a velocity in the medium of 0.3 m/ns).

For the EM algorithm, the result is shown in Fig. 3. Processing consisted of first removing the direct arrival signal which usually presents the largest magnitude in a trace. The resultant traces were scaled to have the same maximum value, then processed by the EM algorithm. Gpr scans were run at 20, 30, 50, 100, 500 and 900 MHz. The target depth is correctly identified at approximately 240ns (two-way travel time). The reflection trace is very broad along the scan axis length and does not distinctly depict the edges of the original target. We suspect this is for two reasons; one, the defined area is more like a bore hole with the length less than half the depth and two, the reflection is being dominated by the lower frequency scans.
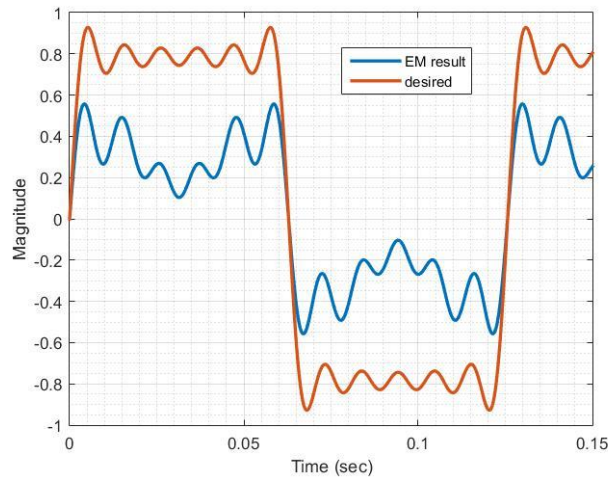
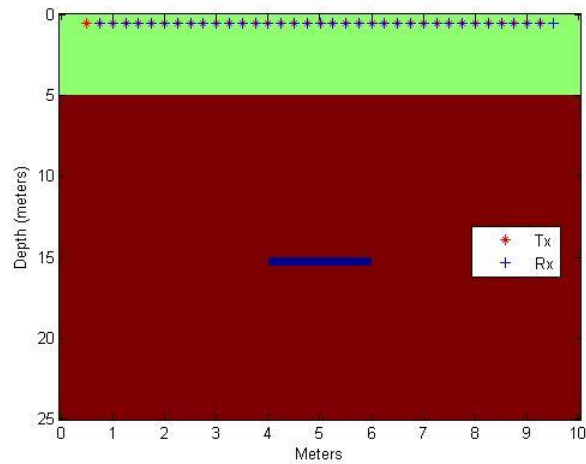Fig. 1: EM Algorithm result; combination of sinewaves.

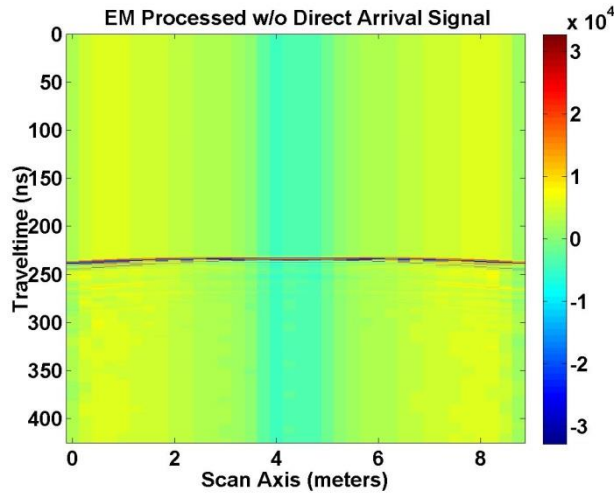Fig. 2: EM Algorithm Test Case Area 1 (TC1).

Fig. 3: EM Algorithm combination of GPR scans of TC1 with 6 Frequencies.

Applying the Dougherty approach to the first test area results in the ouput shown in Fig 4. The target depth is correctly identified but the depth is indication is slightly less crisp than the EM algorithm case. The EM case shows a thinner line at depth. The lower frequencies in the sum appear to account for the delta depth issue. The Dougherty approach depicts better edge detection. The width of the target is better defined though still wider than the defined area but less than the EM Algorithm depiction; Fig 4 verses Fig 3 (EM response). As in the EM method the addition of the lower frequency signals tend to make the depth indication less crisp and the edge detection less accurate. Part of the Direct arrival/ground bounce signal is visible due to the method used by Dougherty to remove them.
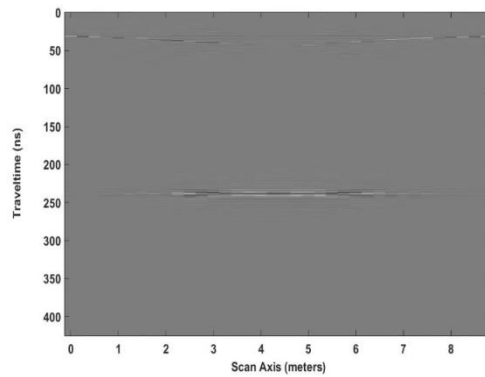

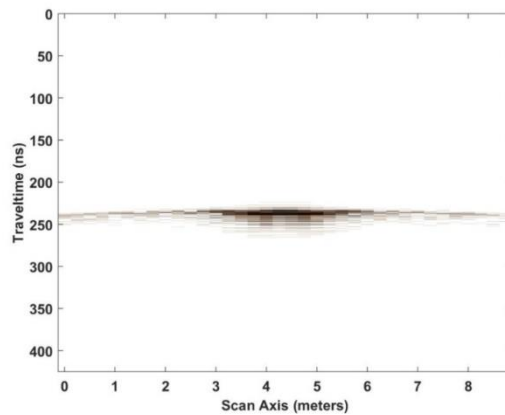Fig. 4: Dougherty standard response for TC1.

Fig. 5: Booth response for TC1.

For the same test area, applying the Booth OSW approach results in the output shown in Fig 5. Again the target depth is correctly identified but like the Dougherty method the depth indication is quite broad. A larger delta time depth indication than the EM and Dougherty methods, exists (Fig 3 and Fig 4). Like the EM method the width in scan axis length is large; edge detection is not well defined. The thickest part of the trace indicates the test target.

Applying the AEE method of Bancroft results in the ouptut shown in Fig 6. Again the target depth is correctly identified. The depth indication is like that of the EM method, sharp but slightly broader in depth. This is about the same as the Dougherty method, and smaller than the Booth method. Target edge detection is more like Booth where the thickest part of the GPR result indicates the test target. Again, like the EM method the GPR reflection covers a wide area (0 to 9 meters) in scan axis length. How much of this response is do to the bore hole effect of the target area is unknown at this time.
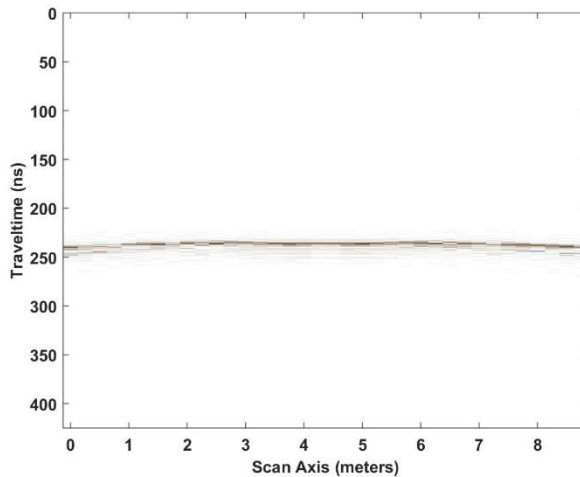

Fig. 6: Output fromAmplitude Envelope Equaization method of Bancroft for TC1.

For this test case, a few problems with the test area definition poses a question. A question as to the effect of the bore hole definiton test area, on the outcome of the scan response. Of concern, are whether a less than crisp edge detection or the addition of large magnitude lower frequencies, making the depth indicators broad, indicative of this definition. By exercising a more complicated test area with broader scan area we attempt to address these concerns.

6.2. Test Case 2 (TC2)

The second defined area consists of an area 30 meters in length and 25 meters in depth with little or no space above ground, (0.15 meters), for the Tx and Rx used. They are swept along the scan axis length starting at 0.5 meters (Tx) and ending at 24.85 meters with spacing between the Tx and Rx the same as before (0.25 meters), Fig. 7. The number of GPR scans is 145. The electrical conductivity of the ground is the same as before but the relative permittivity (ε_r) is 3.0 for dry sand. Buried in the ground at 8 different levels (4.565m, 6.065m, 8.565m, 10.065m, 12.815m, 14.065m, 16.565m and 18.065m) are sheets of corrugated aluminum, modelled as perfect electrical conductors for ease of computation. Each sheet is approximately 2 meters in length and 0.1 meters thick. The GPR scanning frequencies are the same as before. The result for the EM method, shown in Fig. 8, identifies 8 targets at very close to the correct depth (approximately 50ns, 70ns, 100ns, 116ns, 148ns, 160ns, 190ns and 208ns for two-way travel time at a velocity in the medium of 0.1732 m/ns for the defined relative permittivity) with edges depicted reliably but with less fidelity as one descends in depth. Fig. 9 displays the individual GPR traces instead of the image response.
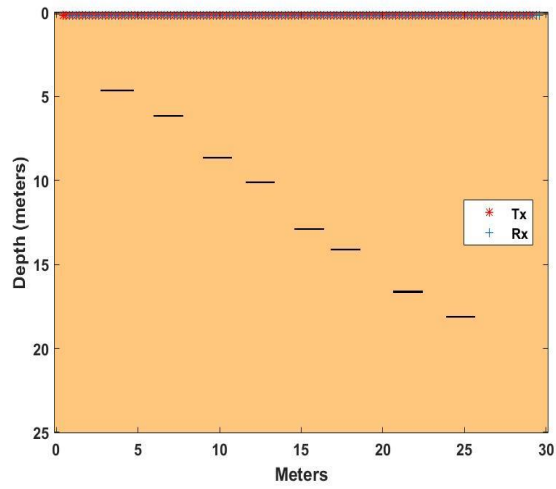
Fig. 7: EM Algorithm Test Case Area 2 (TC2), (8) 2 meter long plates, 0.1meter thick.
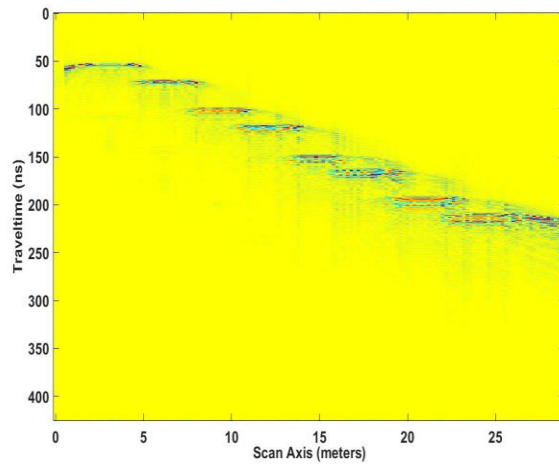

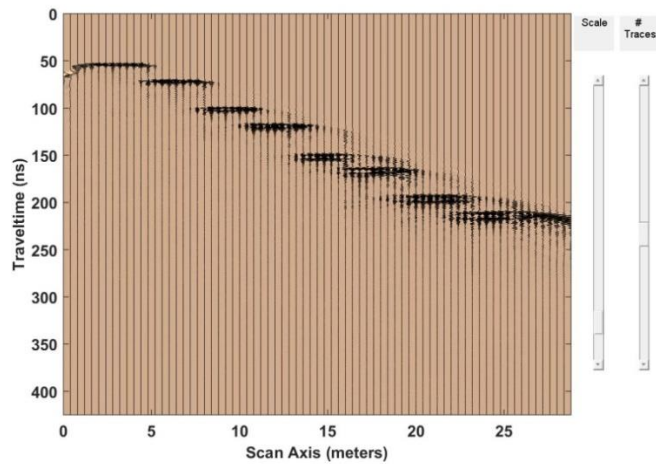Fig. 8: EM Algorithm GPR scan result for TC2, (8) plates shown.

Fig. 9: EM Algorithm GPR scan result showing individual traces for TC2.

Applying the Dougherty method to this second test area produces the GPR response shown in Fig 10. An item to note here is that not all plates are depicted. Only 5 and barely 6 of the 8 are designated. Where the plates end in width is tolerably detectable; edges are noted but not clearly. The result is poorer than the EM processed response of Fig 8.
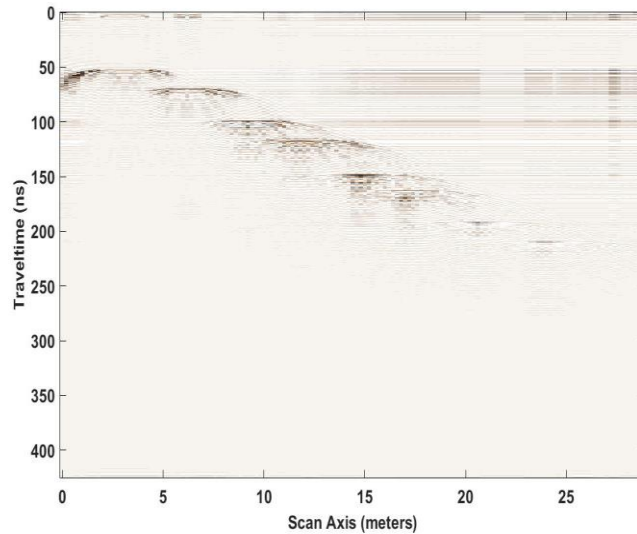

Fig. 10: Dougherty standard response for TC2.

Applying the Booth OSW method to this second test area produces GPR response shown in Fig 11. The ground bounce (shown as a straight line at approximately 50ns in the Fig 11) is still present in the image because the mute feature had to balance between removing the direct arrival/ground bounce signal and not removing the reflection of the first plate at a depth of approximately 50ns two-way travel time. Again only a few plates are detectable. Easily shown are 4 plates and barely 5 and 6 of the 8 plates in the test area. Comparing the result to the EM method, the Booth method falls short at depth. Edge detection is poorer than the EM method but comparable to the Dougherty method.
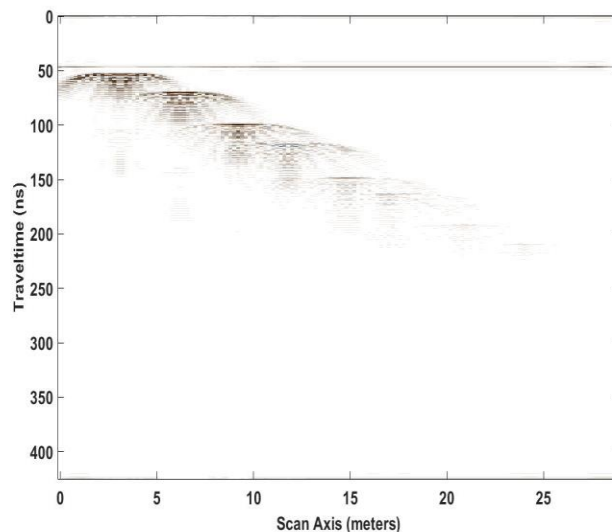

Fig. 11: Booth response for TC2.

Employing the AEE Bancroft method on test area 2, produced the output response depicted in Fig 12. Like Booth and Dougherty before, not all buried plates are illuminated. Of the 8 plates, 4 are definitely depicted with a possibility of 3

more. Added under plates at 50ns and 75ns in depth are "ghost" plates at 100ns and 150ns. There were no targets buried at these two points. Edge detection is better than Booth and Dougherty and on par with the EM algorithm.
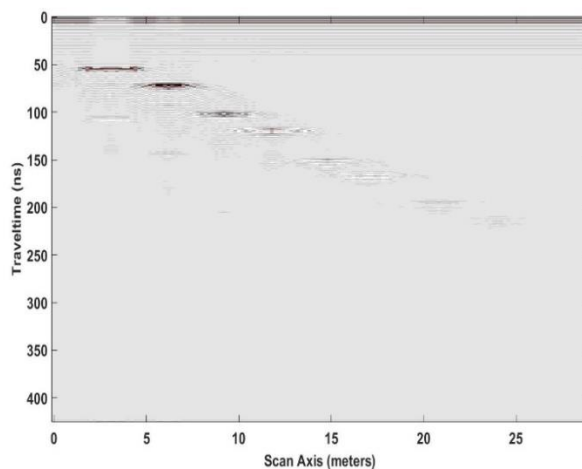

Fig. 12: Bancroft response for TC2.

For this test case, the EM algorithm has performed the best in terms of revealing the 8 buried plates. Edge detection is still not great but tolerable for the EM method for both test cases and the Bancroft method for the second test case. The larger scan area does address the bore hole effect question of wider scan axis length reducing wide reflection traces. The dimensions of the scan area do affect the width of the scan axis reflections. The ability to achieve crisp edge detection has not changed much however.

7. Conclusion

In this paper we have examined the use of an optimization problem solver to define the best set of weights for compositing. The Gaussian Mixture Model feature of the Expectation-Maximization Algorithm was the technique used to combine several GPR scans, at various frequencies. The goal of studying this method was to determine how to produce the best image at the surface as well at depth. We compared the result of two well defined test areas with 3 other methods, to evaluate the ability of the EM method to compete with them. The 3 methods explored were developed by Dougherty, Booth and Bancroft, none of which were optimization problem solvers. Each had limited success but the EM algorithm faired the best in recognizing images at depths down to 10 meters in moist sand and dry sand media. Though actual scans were not used, simulated scan software proved to be commensurate with actual data; verified by reference [10], a previous paper. A Tx and Rx were moved across the length of a defined space as if it were an actual GPR radar scanning device. We feel confident that the analysis is accurate. The study did uncover problem areas in need of future study. The edge detection ability, how to reliably remove the direct wave/ground bounce without removing the response from buried targets and how to best align trace starting points across frequencies, as examples. The latter appears to help reduce the thickness in depth of GPR scan results. We have shown that the EM method using Gaussian Mixture Models is a useful tool for compositing. We believe, that more study using more complicated models as well as actual scans is warranted.

Acknowledgements

References

[1] M. E. Dougherty, P. Michaels, J. R. Pelton, L. M. Liberty, "Enhancement of Ground Penetrating Radar Data Through Signal Processing," *Symposium on the Application of Geophysics to Engineering and Environmental Problems,* 1994, pp. 1021-1028, DOI 10.4133/1.2922053.

[2] A. D. Booth, A. L. Endres, T. Murray, "Spectral Bandwidth Enhancement of GPR Profiling Data Using Multiple-Frequency Compositing," *Journal of Applied Geophysics*, 2009, vol. 67, pp. 88-97, DOI 10.1016/j.jappgeo.2008.09.015.

[3] S. W. Bancroft, "Optimizing the Imaging of Multiple Frequency GPR Datasets using composite Radargrams: An Example from Santa Rosa island, Florida," Ph.D. dissertation, University of South Florida, 2010.

[4] P. Smyth, "The EM Algorithm for Gaussian Mixtures, Probabilistic Learning: Theory and Algorithms, CS274A," University of California, Irvine, Department of Computer Science, Lecture Note 4.

[5] A. P. Annan, "Electromagnetic Principles of Ground Penetrating Radar," *in Ground Penetrating Radar: Theory and Applications*, Amsterdam: Elsevier, 2009, pp. 1-40, ISBN: 978-0-444-53348-7.

[6] A. Tavlove, "Review of the formulation and Applications of the Finite-Difference Time-Domain Method for Numerical Modeling of Electromagnetic-Wave Interactions with Arbitrary Structures," *Wave Motion*, 1988, vol. 10, pp. 547-582, DOI 10.1016/0165-2125(88)90012-1.

[7] N. Blindow, D. Eisenburger, B. Illich, H. Petzold, and T. Richter, "Ground Penetrating Radar," in *Environmental Geology*, Ed. Springer Berlin Heidelberg, 2008, pp. 283-235, DOI 10.1007/978-3-540-74671-3_10.

[8] A. Giannopoulos, "Modelling Ground Penetrating Radar by GprMax," *Construction and Building Materials*, 2005, vol. 19, pp. 755-762, DOI 10.1016/j.conbuildmat.2005.06.007.

[9] J. J. Verbeek, N. Vlassis, and B. Kröse, "Efficient Greedy Learning of Gaussian Mixtures," in *Proceedings of the 13th Belgian-Dutch Conference on Artificial Intelligence* (BNAIC'01), 2001. pp. 251-258, INRIA-00321510.

[10] R. Tilley, F. Dowla, F. Nekoogar, and H. Sadjadpour, "GPR Imaging for Deeply Buried Objects: A comparative Study based on FDTD models and Field Experiments," Selected Papers Presented in *Proceedings of the MODSIM World 2011 Conference and Expo*, 2012, pp. 45-51, (NASA/CP-2012-217326), (SEE 20130008625).

[11] R. Tilley, H. Sadjadpour, and F. Dowla, "Combining Ground Penetrating Radar Scans of Differing Frequencies Through Signal Processing," *in Proceedings of The Ninth International Conference on Advanced Geographic Information Systems, Applications, and Services*, GEOProcessing, 2017.