Duke UNIVERSITY

Application of Markov Random Fields to Landmine Discrimination in Ground Penetrating Radar Data

Peter Torrione & Leslie Collins

Duke University

QMDNS; May 21, 2008

This work was supported under a grant from the US Army RDECOM CERDEC Night Vision and Electronic Sensors Directorate & ARO.

Alternative Phenomenologies

- Many other phenomenologies for landmine detection have been suggested
  - **Electromagnetic induction (EMI)**
  - Infrared techniques [Lopez, 2004]
  - Seismic & Acoustic-seismic coupling [Sabatier, 2001. Scott, 2001]
  - Ground penetrating radar (GPR)
  - Many others [MacDonald, 2003]
- Note:
  - Due to differences in:
    - Landmine types
    - Percent clearance requirements
    - Other operational requirements
  - No "silver bullet" landmine detection phenomenology
- Sensor fusion is an active area of research [Collins, 2002. Ho, 2004.]

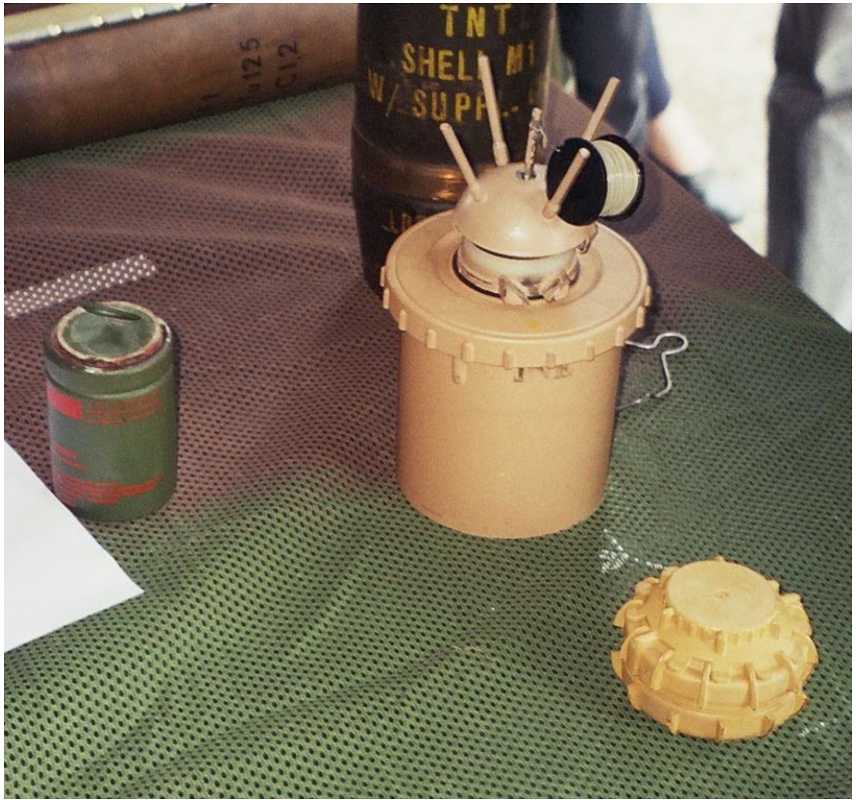


Image by David Monniaux available under Creative Commons License

Motivation & Goal

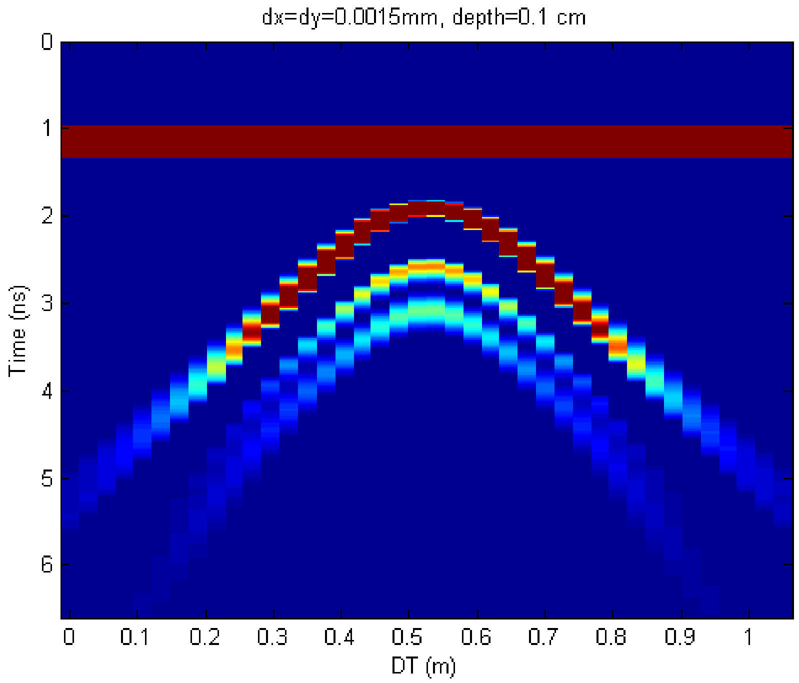
- Significant diverse research on landmine detection in time-domain GPR data
  - Ground tracking and removal [Gu, 2002. Abrahams, 2001. Larsson, 2004. Guangyou, 2001]
  - Pre-screening [Carevic, 1999. Zoubir, 2002. Kempen, 2001. Karlsen 2001]
  - Feature extraction [Kleinman,1993. Carevic,1997. Frigui, 2004. Gader, 2004. Ho, 2004]
  - Image segmentation [Verdenskaya, 2006. Bhuiyan, 2006. Shihab, 2003]
  - Etc…
- Many proposed techniques are implicitly based on different underlying models of received time-domain data
  - Makes direct motivation and comparison of algorithms difficult without expert modifications
- Propose an underlying statistical model for GPR responses that incorporates spatial variations in response heights and response gains
  - Can formalize development of pre-screener algorithms based on underlying models
    - Under what conditions will adaptive algorithms perform well?
    - Are other algorithms also applicable?
  - Can provide forward *generative* model of large data sets
    - Given parameters, can simulate roads
    - Can not model responses from mines, etc.

Outline

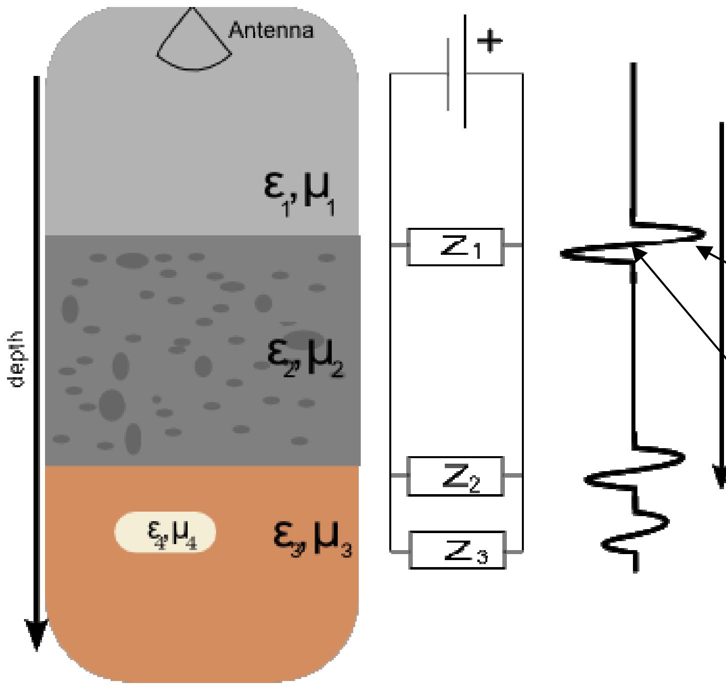
- Consider various modeling techniques for GPR data
  - Computational concerns – FDTD, transmission lines
  - Applicability under fielded (unknown soil property) scenarios
- Incorporating statistical parameterization of transmission line models
  - Markov Random Fields (MRF)
    - Gaussian Markov random fields (GMRF)
  - Application of MRFs to parameters of interest in transmission-line model
- Implications of proposed statistical model for pre-screener development
  - Adaptive maximum likelihood solution for GMRF parameters in GPR data time-slices
  - Adaptive discriminative algorithms for dual GMRF under both hypotheses
- Results & Conclusions / Future work

Modeling of GPR Returns

- Finite difference time-domain (FDTD) models provide state of the art modeling of GPR responses
  - Highly generalizable
  - Computationally expensive
- Require:
  - Accurate knowledge of soil and anomaly properties
  - Locations of discontinuities
  - Etc
- Inversion / fielded application of FDTD models is difficult

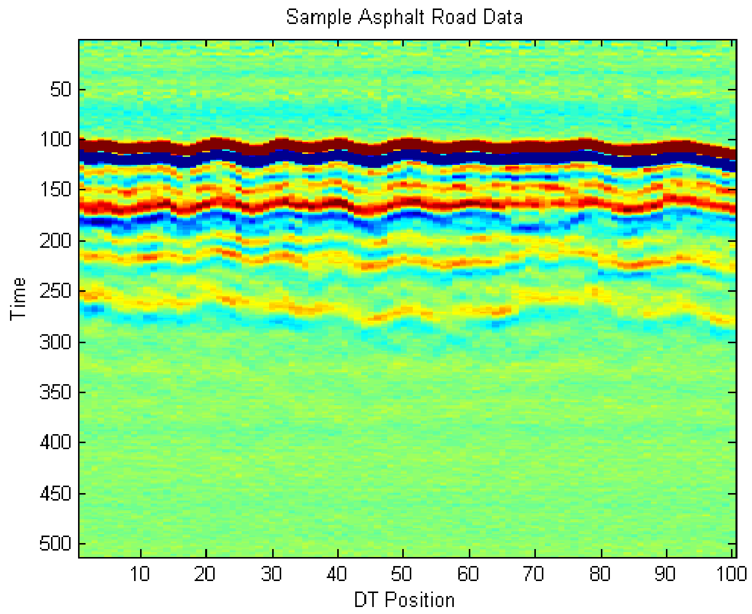


Basic Transmission Line Model



- Significant simplification of GPR responses
  - Treats dielectric discontinuities in soils as impedance mismatches on a transmission line
- Received signal is a sum of time-delayed pulses
  - Response depends on: time of arrival, gain on received pulses

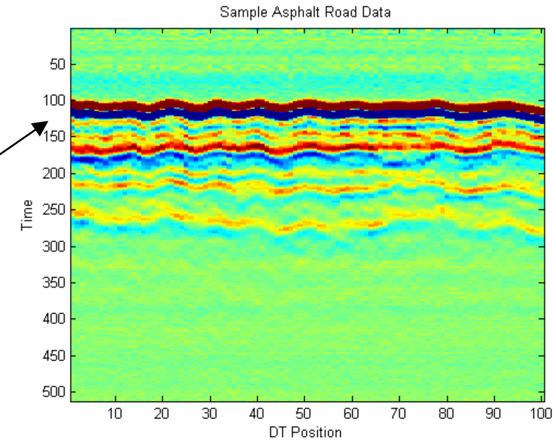
Restrictions of Transmission Line-Based Modeling

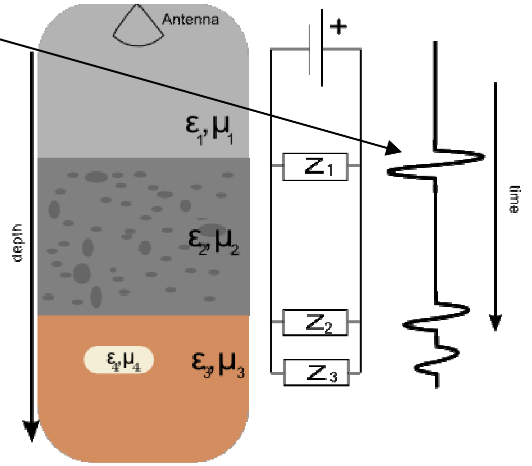


- Transmission line models assume:
  - Planar waves
  - Planar interfaces
  - Homogeneous transmission media
  - Etc.
- Obviously these assumptions are violated in fielded scenarios
- Question:
  - *Can a statistical model over parameters (time of arrival, gain) mitigate these violated assumptions?*

GMRF Modeling of TOA and Gain

- For simplicity; focus on modeling of air/ground interface
  - Other subtleties for sub-surface layers
- Estimating TOA is straightforward; model as GMRF
- Model received gain as combination of deterministic & stochastic part





Modeling of Received Gain

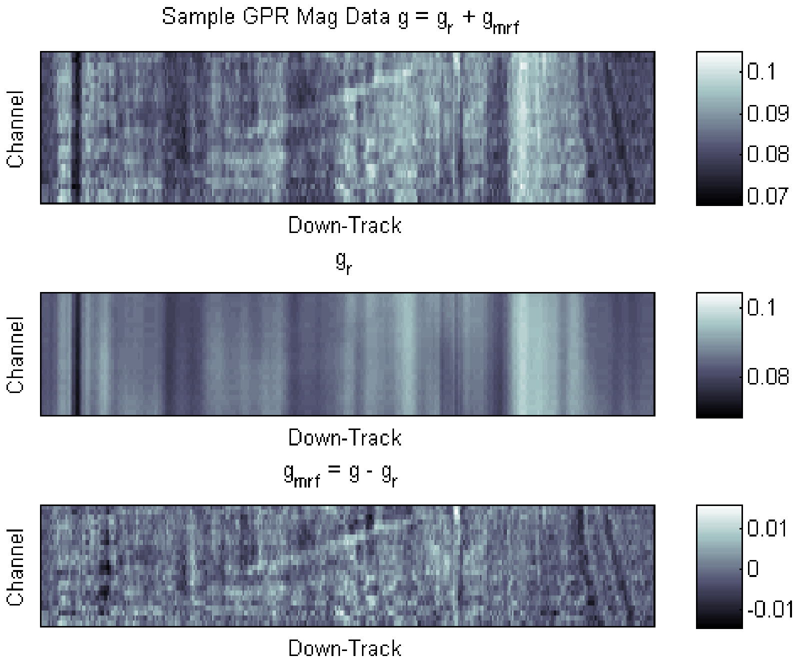
- Model received gain as combination of deterministic part (spreading loss)

$$g_r = A + B\frac{1}{t_0}$$

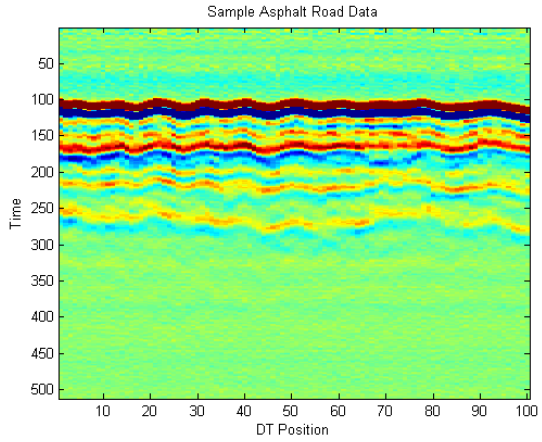
- Stochastic part (soil roughness, dielectric properties, etc)

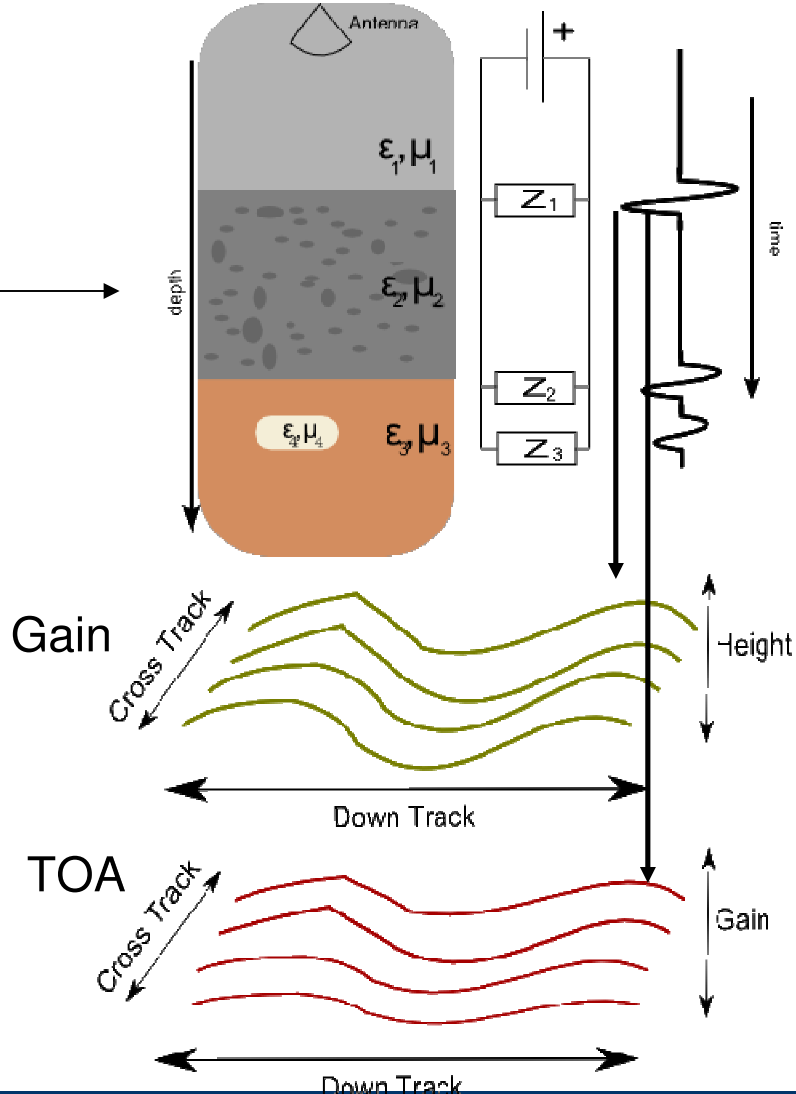
$$g = g_r + g_{mrf}$$

- Image on right shows original measured gain, deterministic gain, MRF gain



Proposed Statistical Model

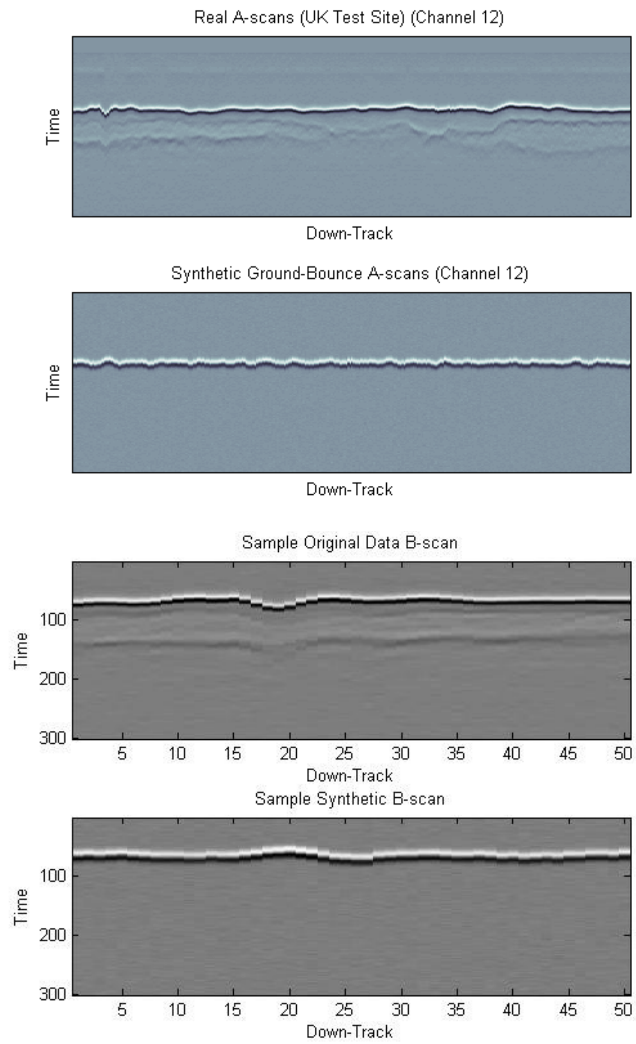




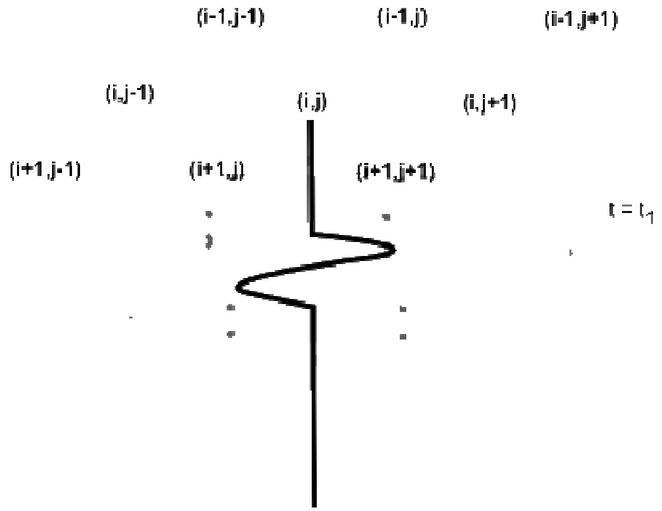
- Combination of simple A-scan transmission line modeling & spatial statistical modeling of underlying gain & time of arrival (TOA)
- By applying spatial statistical models over A-scan parameters ➔ computationally tractable 3-D volume model for GPR data

Sample Generative Model Application

- Images on right show original data (top images), synthetic data (bottom images)
  - Top figure shows ~500 scans
  - Bottom figure shows 50 scans
- Synthetic data only models initial ground bounce response
  - Both height and gain terms are modeled stochastically using Markov random fields
  - MRF parameters trained using data from UK testing site
- Generative model may be useful in its own right for simulating responses over soils with varying parameters, simulating large data sets, etc.
  - Modeling sub-surface structure is a little more complicated; requires parameter estimation techniques, statistics for appearance / disappearance of sub-surface responses



Implications of Transmission Line MRF Modeling of Soils For Pre-Screening



- Consider distribution of data in a time-slice

$$A_{i,j}(t_m) = g_{i,j} f(t_m - t_{0_{i,j}})$$

$$p(A_{i,j}(t_m)) = p(g_{i,j} f(t_m - t_{0_{i,j}}))$$

$$p(A_{i,j}(t_m)) = p(g_{t_{0_{i,j}}} f(t_m - t_{0_{i,j}})) + p(g_{mrf_{i,j}} f(t_m - t_{0_{i,j}}))$$

- ➔ Data in time slice also MRF, although not closed form;
  - Assume GMRF

Target Detection Using GMRF For Data Under H_0

- Desire LRT:

$$\lambda(x) = \frac{p(x|H_1)p(H_1)}{p(x|H_0)p(H_0)}$$

- Assume data under H_1 is ~ improper uniform; data under H_0 is ~ GMRF

$$p(x(n)|\mathbf{x}_{N_n}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x(n)-\sum_{n' \in N_n} \beta_{n'} x(n'))^2}{2\sigma^2}}$$

- Need parameters for GMRF!
- Consistent parameter estimation equations [Kashyap, 1983]

$$\beta_c = [\textstyle\sum_{s\in\Omega} \mathbf{x}(N(s))\mathbf{x}^T(N(s))]^{-1} \textstyle\sum_{s\in\Omega} \mathbf{x}(N(s))x(s)$$

MPLE MRF Modeling ➔ Weiner Hopf?

$$p(x|\mathbf{w},\mathbf{x}_N) = \frac{1}{\sqrt{2\pi}\sigma}\exp^{\frac{(x-\mathbf{w}^T\mathbf{x}_N)^2}{2\sigma^2}}$$

$$p(\mathbf{x}|\mathbf{w}) \approx \prod_s p(x_s|\mathbf{w},\mathbf{x}_{N_s})$$

$$\max_{\mathbf{w}} \mathbf{E}_{x,\mathbf{x}_N}(\log(p(x|\mathbf{w},\mathbf{x}_N)))$$

$$\max_{\mathbf{w}} \mathbf{E}_{x,\mathbf{x}_N} \log\frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{2\sigma^2}(x-\mathbf{w}^T\mathbf{x}_N)^2$$

$$\max_{\mathbf{w}} \log\frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{2\sigma^2}\mathbf{E}(x^2) - \mathbf{w}^T\mathbf{R}\mathbf{w} \ - \ 2\mathbf{w}^T\rho$$

$$\frac{d}{d\mathbf{w}} = 0 = 2\mathbf{R}\mathbf{w} \ - \ 2\rho$$

➔ $$\mathbf{w} = \mathbf{R}^{-1}\rho$$

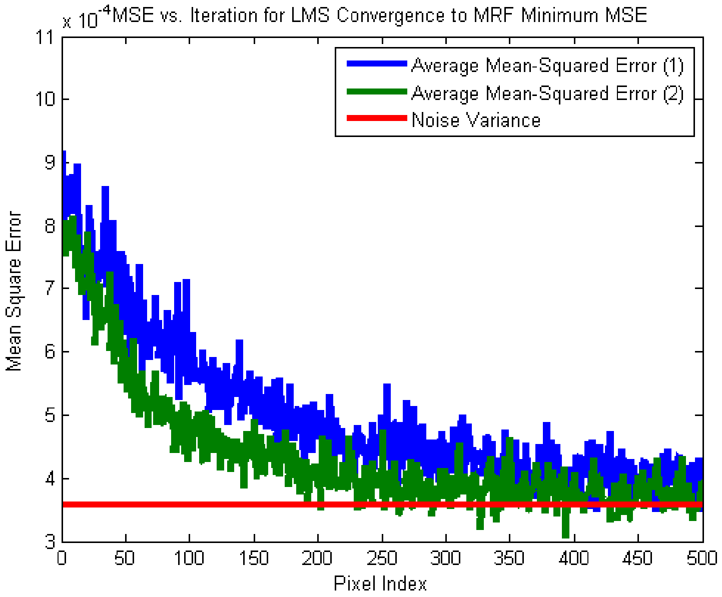
- Kashyap et al. result is very similar to Weiner-Hopf equations
- Turns out, can directly motivate Weiner-Hopf from maximum pseudo-likelihood form of distributions

Motivating Adaptive Pre-Screening

- Last slides illustrated how pseudo-likelihood GMRF leads to Weiner-Hopf
- Similar arguments (removing expected values) show that ML estimates of non-stationary GMRF parameters yield LMS update equations
- *This provides a model-based motivation of the application of AR based signal processing to pre-screening in GPR data*

$$\frac{d}{d\beta} = -2x(n)d(n) + 2\mathbf{x}_N\mathbf{x}_N^T\hat{\beta}_n$$

$$\hat{\beta}_{n+1} = \hat{\beta}_n + \mu\mathbf{x}_N(x(n) - \mathbf{x}_N^T\hat{\beta}_n)$$



Discriminative Learning in GMRF Models

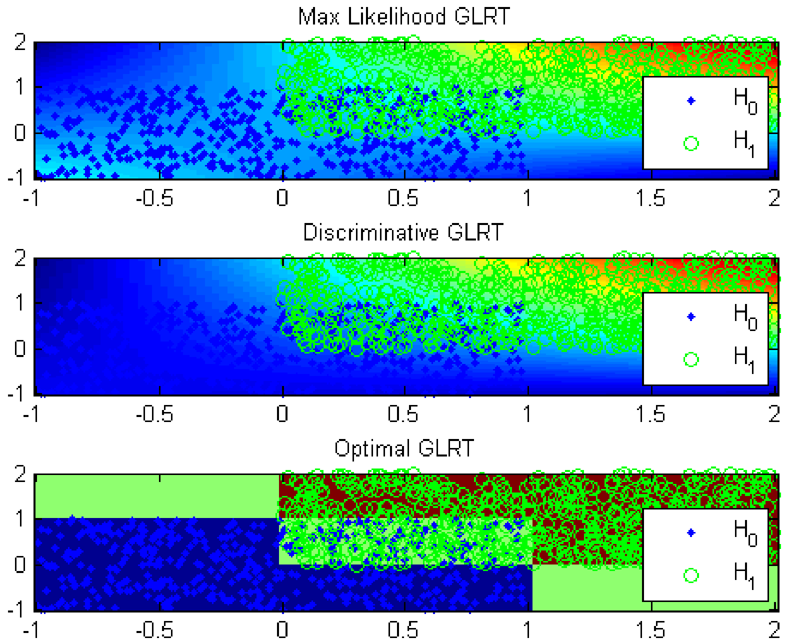
- Previously H_1 ~ improper uniform
- Alternatively, Consider if data under H_1 is also ~ GMRF
- Can directly solve for *discriminative* parameters

$$p(y_i|x_i,\theta) = \frac{p(x_i,y_i|\theta)}{\sum_k p(x_i,c_k|\theta)}$$

- Turns out, for many models the form of the discriminative logistic function is *linear* in the weights

$$p(H_1|\mathbf{X}) = \sigma(\mathbf{w}^T\mathbf{x})$$

- GMRF Models do not lead to linear logistic discriminative models



Solving For Adaptive Discriminative GMRF/GMRF Update Equations

$$p(x_i|\mathbf{x}_{N_i}, \theta_1, H_1) = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{(\theta_1^T \mathbf{x}_{N_i} - x_i)^2}{2\sigma_1^2}}$$

$$a_{gmrf} = \log \frac{p(H_1)}{p(H_0)} + \log \frac{\sigma_0}{\sigma_1} - \frac{(\theta_1^T \mathbf{x}_{N_i} - x_i)^2}{2\sigma_1^2} + \frac{(\theta_0^T \mathbf{x}_{N_i} - x_i)^2}{2\sigma_0^2}$$

$$\frac{da_{gmrf}}{d\theta_1} = -\frac{(\theta_1^T \mathbf{x}_{N_i} - x_i)\mathbf{x}_{N_i}}{\sigma_1^2}$$

- Turns out
  - Given: $\Theta_1, \Theta_2, \sigma_1, \sigma_2$
  - Given: x_i, y_i

New GMRF update equations

$$\theta_1 = \theta_1 + -\frac{(\theta_1^T \mathbf{x}_{N_i} - x_i)\mathbf{x}_{N_i}}{\sigma_1^2}(y_i - \sigma(a)) * \mu$$

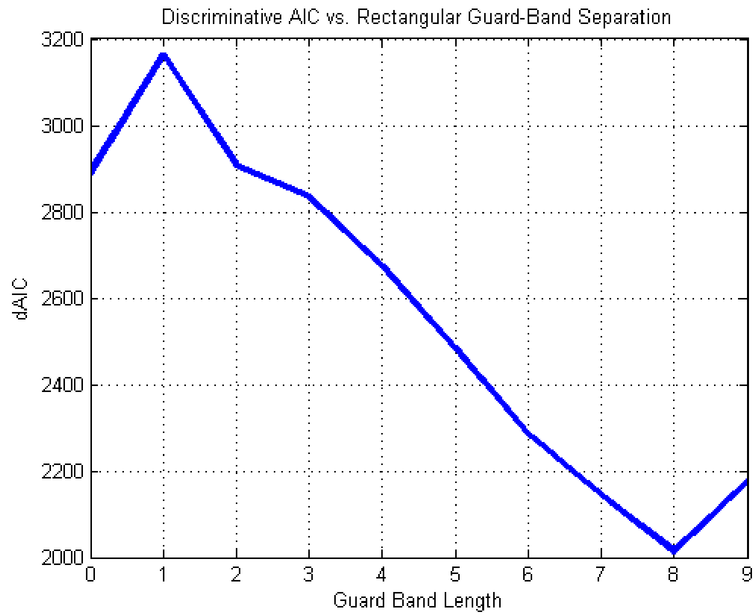
$$\theta_2 = \theta_2 + \frac{(\theta_2^T \mathbf{x}_{N_i} - x_i)\mathbf{x}_{N_i}}{\sigma_2^2}(y_i - \sigma(a)) * \mu$$

$$\sigma_1 = \sigma_1 + (\frac{-1}{\sigma_1} + \frac{(\theta_1^T * \mathbf{x}_{N_i} - x_i)^2}{\sigma_1^3}(y_i - \sigma(a))) * \mu$$

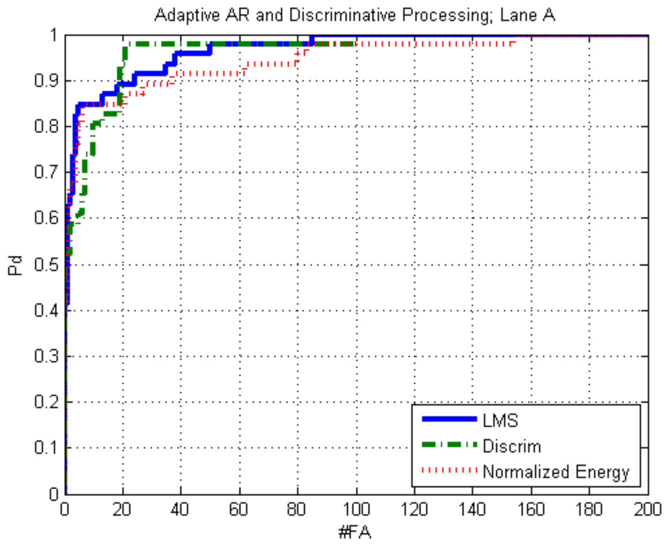
$$\sigma_2 = \sigma_2 + (\frac{1}{\sigma_2} - \frac{(\theta_2^T * \mathbf{x}_{N_i} - x_i)^2}{\sigma_2^3}(y_i - \sigma(a))) * \mu$$

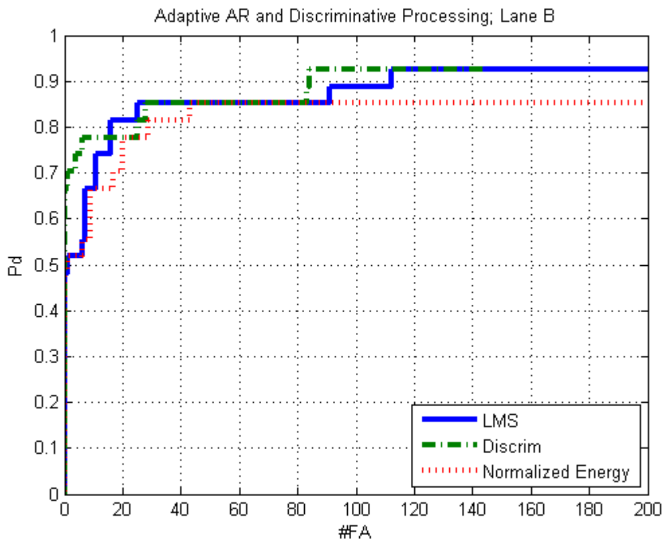
Advantages of Discriminative Classification

- Modeling data under H_1 as GMRF has several implicit advantages
  - Provides natural estimation of discriminative Akaike Information Criteria
  - Probabilistic outputs from each time-slice allow principled depth-bin fusion
    - Inclusion of prior information regarding target depths
- Can be computationally complex, however



Pre-Screener ROC Curves

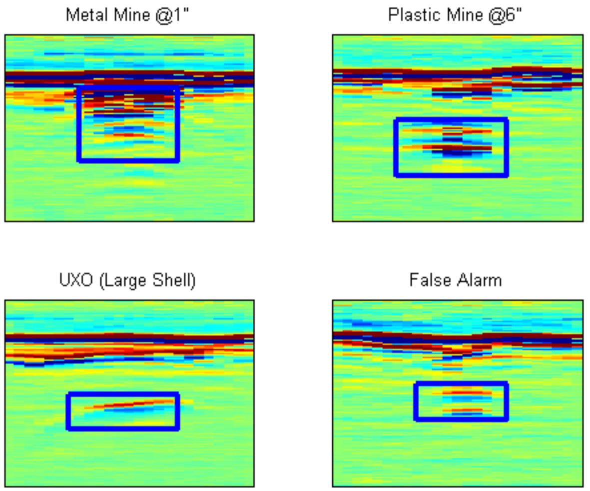


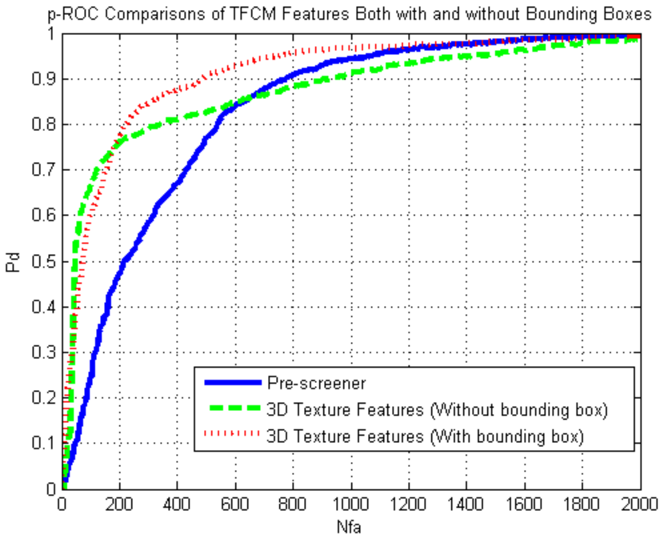


- Results show sample ROC curves for energy (red-dotted), LMS (blue), discriminative (green-dashed)
  - Note, no pre-processing/post-processing of outputs.
  - ROCs not indicative of system performance, provide algorithm comparison only
- Discriminative algorithm provides slight performance improvements
  - Underlying H_1 model (GMRF) may be overly simplistic

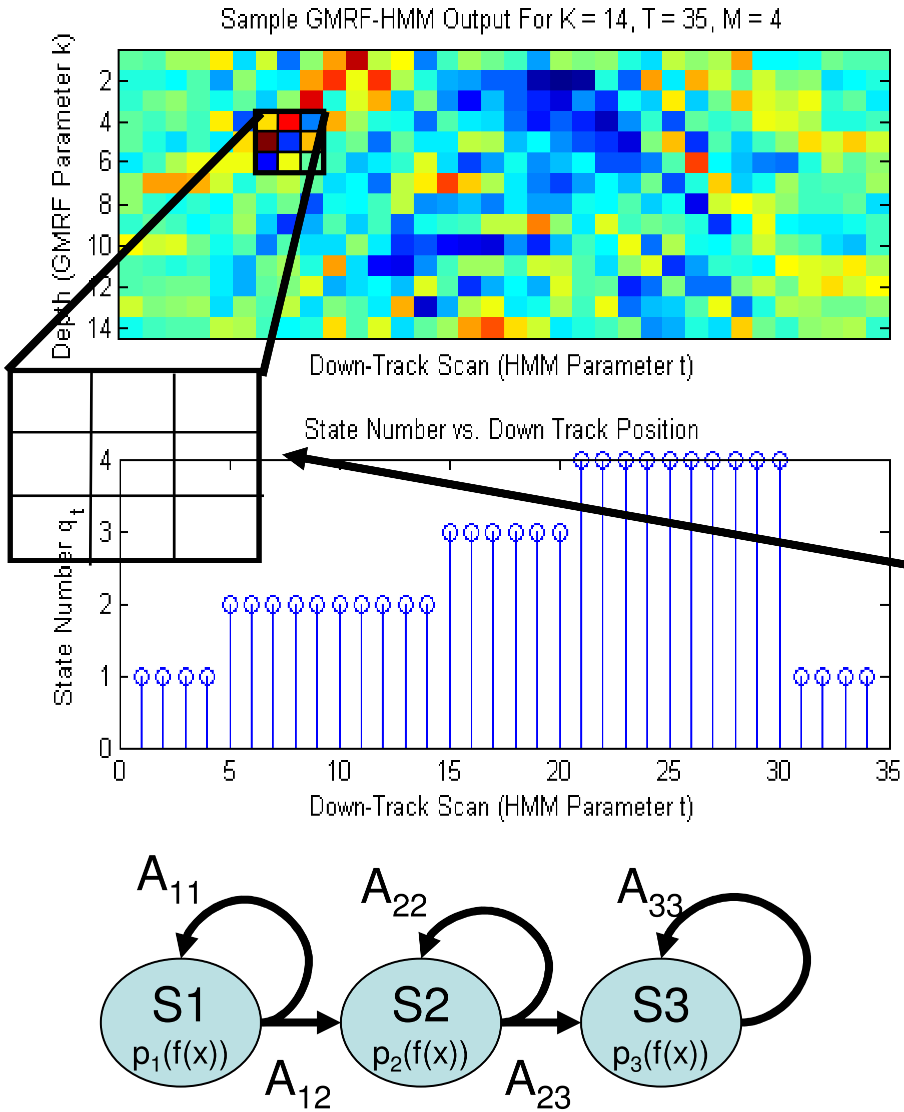
Other MRF Applications (Image Segmentation)

- Image segmentation for target localization
  - Improve extracted feature SNR, computational complexity
- Shown to improve performance for target identification against AP, AT, IED responses





GMRF-HMM For Landmine Detection



- Similar to [Gader, 2001] consider locally stationary distributions of target responses
- Idea: *Directly model received data as GMRF*
  - No need for ad-hoc feature extraction
  - Requires neighborhood system N
  - Can we simultaneously learn parameters of GMRF (features) and underlying states?

$$p_{s_n}(x_n|x) = p_{s_n}(x_n|x_{N_n}) =$$

$$p_{s_n}(x_n|x_{N_n}) = \mathrm{GMRF}(\theta_{s_n}, \sigma_{s_n})$$

Conclusions & Future Work

- Developing a generative model for GPR responses based on spatial stochastic parameterization of the transmission line model
  - Enables generation of data from sample data; eliminates need to estimate soil electromagnetic properties directly
- Proposed model
  - Provides direct motivation for application of AR approaches to pre-screening
  - Motivates application of discriminative approaches to pre-screening when distribution under H_1 is known
    - Current GMRF distribution appears to be overly simplistic
- Future work:
  - Incorporate model implications to:
    - Ground tracking, image segmentation, feature extraction

Acknowledgements

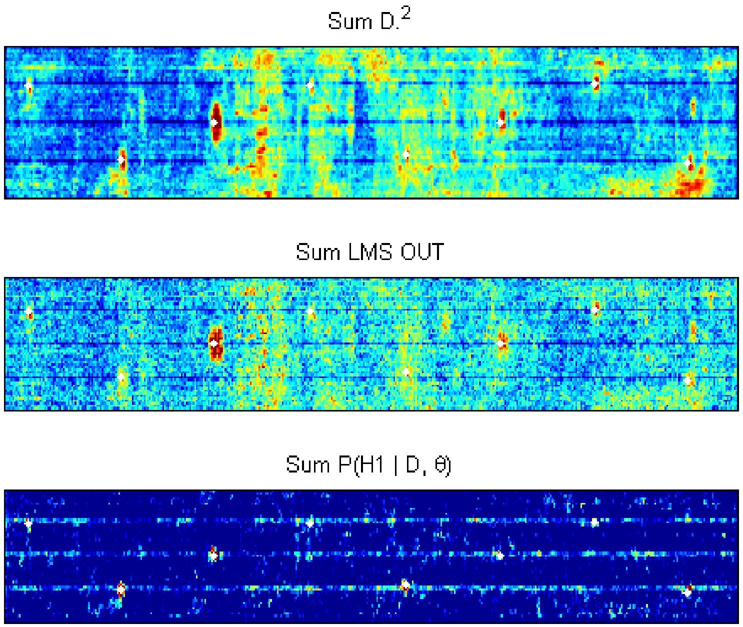
- This work was supported under a grant from the US Army RDECOM CERDEC Night Vision and Electronic Sensors Directorate & ARO.
- The authors would like to thank their colleagues at NVESD, UFL, UM, UL, NIITEK, BAE, and IDA.

Backup

Adaptive Training Issues

- Haven't incorporated the p(H1), p(H0) terms in adaptive updates; these will need to be set
  - Should not be learned adaptively?
- Issues in adaptively training discriminative models when we may only see data from H0 – the parameters under H1 will be driven to unrealistic values since model will do "well" when everything is considered H0
  - Solution: Consider library of mine signatures; stochastically select from these and for every H0 sample, train the model also with a random set of mine data

Image Depth-Bin Fused Decision Statistics



- Top image: Energy
- Middle image: LMS Outputs
- Bottom image: p(H1|D,M)

Global Model

$$p(\mathbf{Y}|\mathbf{X}, M) = \prod_{n=1}^{N} p(H_1|M, X_n)^{y_n}(1 - p(H_1|M, X_n))^{1-y_n}$$

$$\log(p(\mathbf{Y}|\mathbf{X}, M)) = \sum_{n=1}^{N} y_n \log(p(H_1|M, X_n)) + (1-y_n)\log(1-p(H_1|M, X_n))$$

$$p(H_1|\mathbf{X}) = \sigma(a)$$

- Differentiating:

$$\frac{d}{d\theta_1} = \sum \frac{da_{gmrf}}{d\theta_1}(y_n - \sigma(a))$$

$$\frac{d}{d\theta_1} = \sum -\frac{(\theta_1^T \mathbf{x}_{N_i} - x_i)\mathbf{x}_{N_i}}{\sigma_1^2}(y_n - \sigma(a))$$