# Soft Computing for Sensor and Algorithm Fusion

**Jim Keller**

**Electrical and Computer Engineering**

**University of Missouri- Columbia**

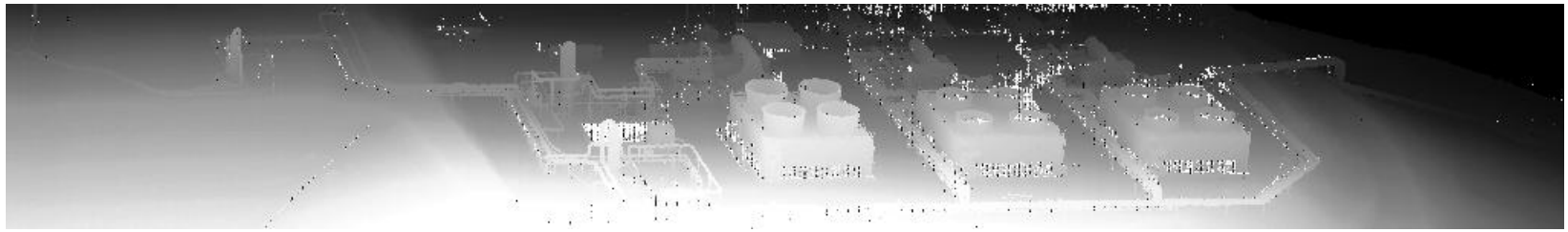
**KellerJ@missouri.edu**

**http://www.missouri.edu/~kellerj**

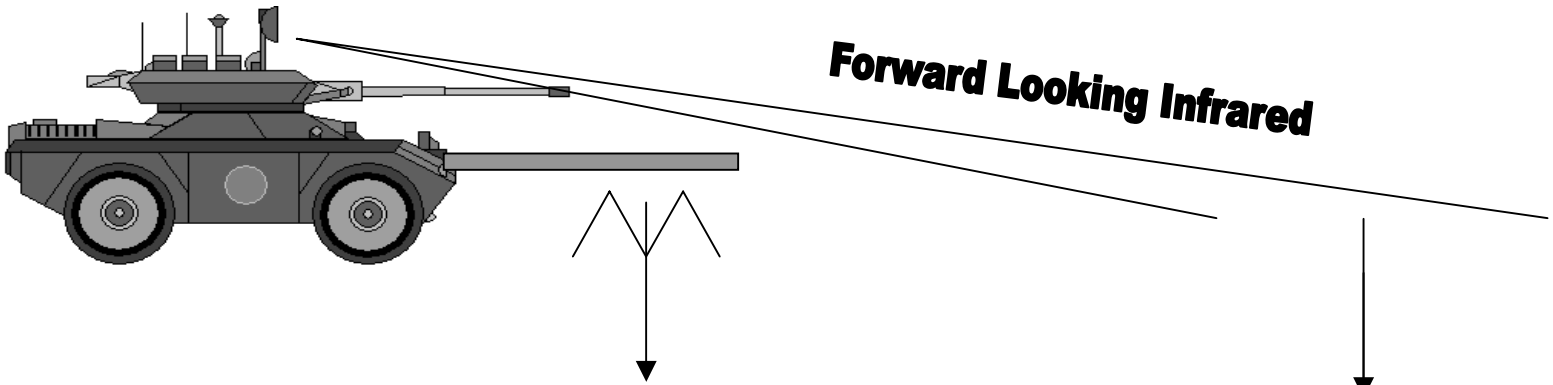
**Work done with many colleagues and students at the University of Missouri and elsewhere**

# Introduction

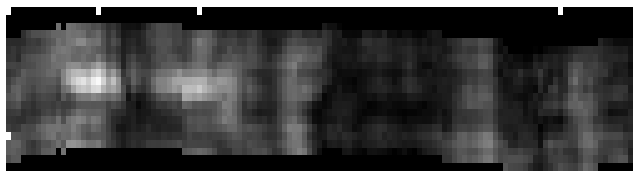
- **Complex problem domains like Landmine Detection, ATR, Handwriting Recognition, etc**
  - **Need multiple features, algorithms, sensors to resolve ambiguity**
- **Uncertainty abounds in these situations**
- **We Need Models and Calculi to Manage Uncertainty in Computational Systems for Algorithm/Sensor Fusion**
- **Here, we discuss Fuzzy Set Theory and Fuzzy Logic as the underlying fusion technology**

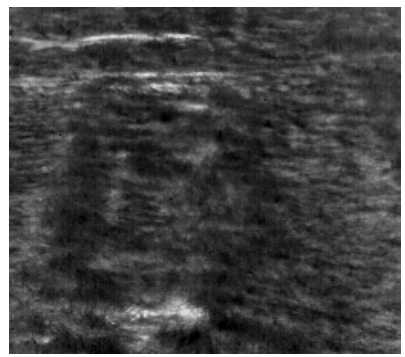


LADAR Range Image




Ground Penetrating Radar





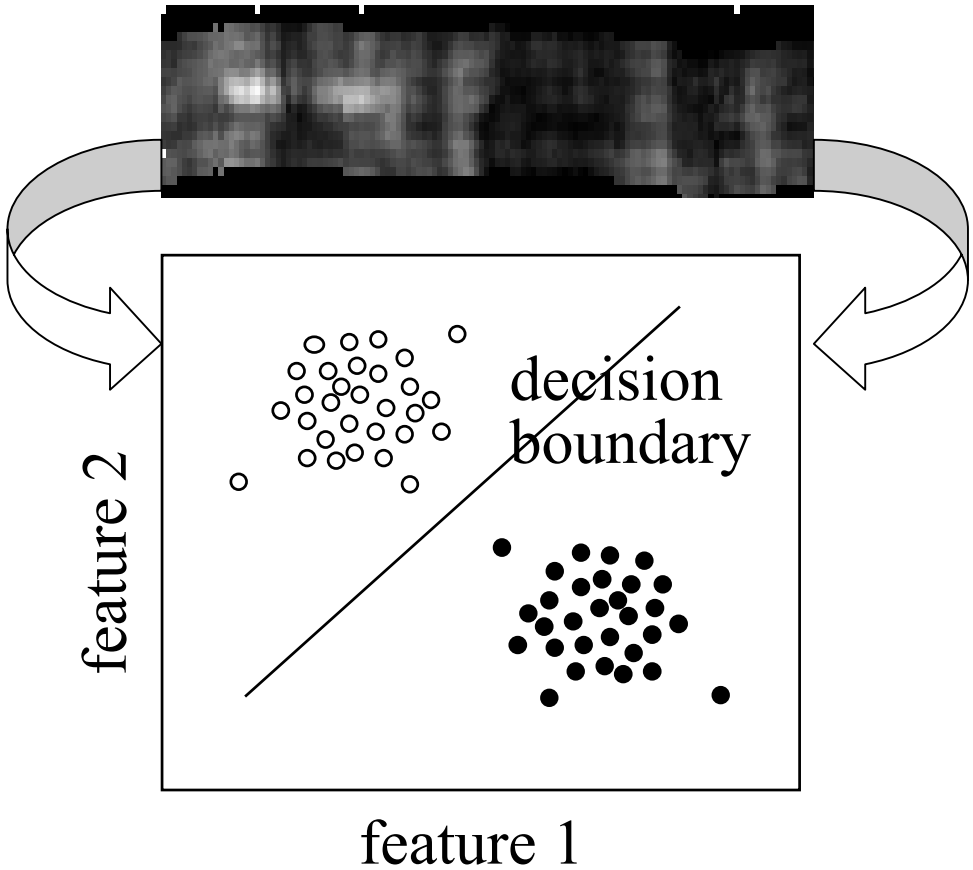
Humanitarian Demining

# Pattern Recognition

- **Find Meaningful Associations Among Representations of Objects**
  - **Assign Class Labels to New (Unknown) Samples**
- **Supervised Learning**
  - **Have Labeled Training Data to "Learn"the Assignment Function**
    - **Bayes Decision Theory,**
    - **MLP Neural Networks, …**
- **Unsupervised Learning**
  - **Find "Natural" Groups Within a Complex DataSet**
    - **Clustering,**
    - **SOFM, etc**

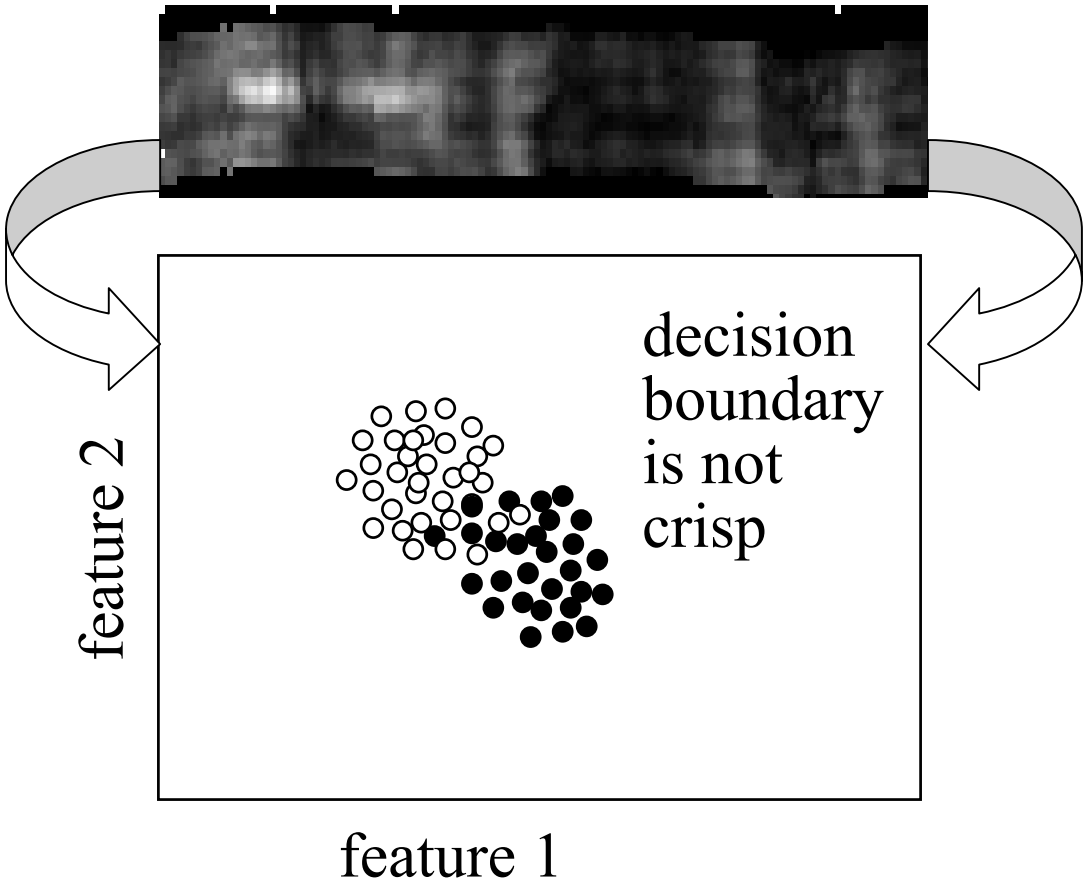
# How We'd Like the "World" to Be

**Nice Crisp Boundaries between classes**



# More Likely How It Is!

**Need to Model and Manage the Uncertainty in the Data**

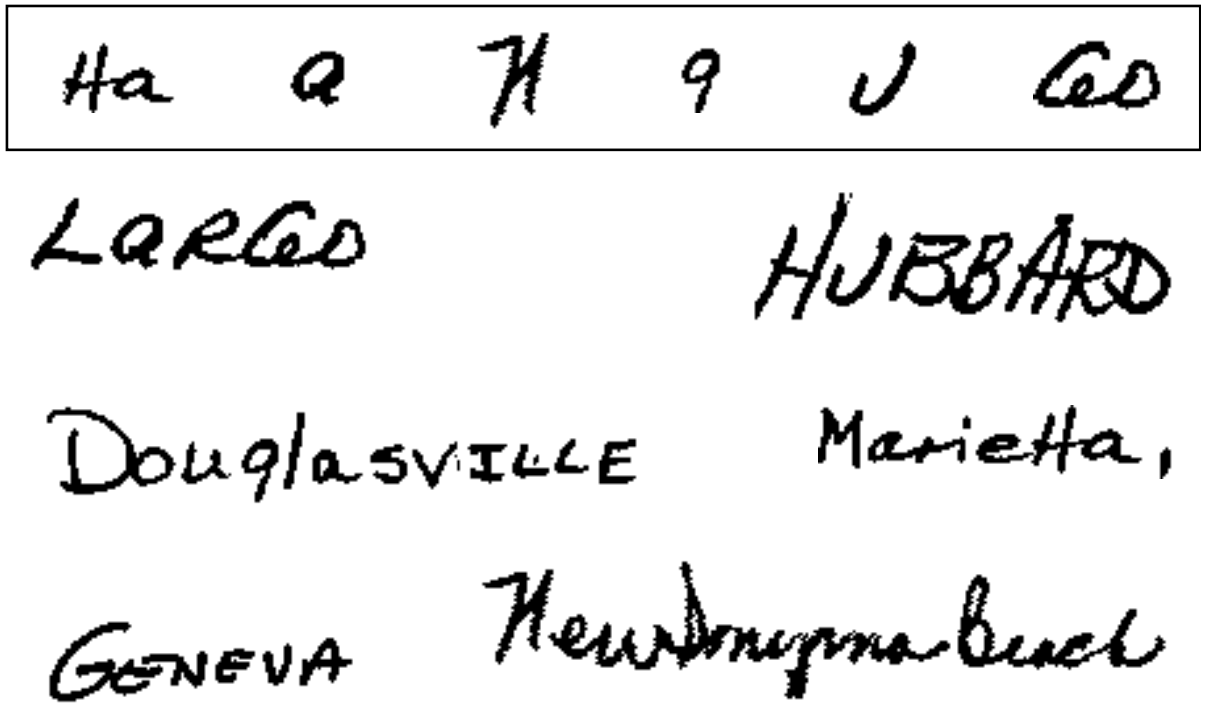


# Class Membership Ambiguity

- To What Chacter Classes Do the Following Cursive Examples Belong?

- Can We Make a Precise Assignment Here for Use in a Handwritten Word Recognition System?

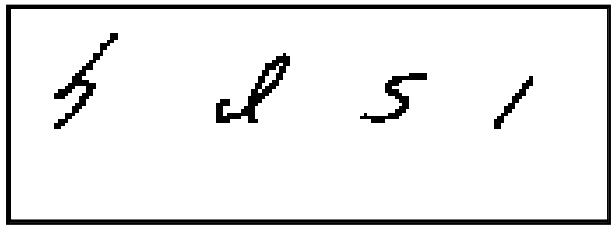
# Class Membership Ambiguity



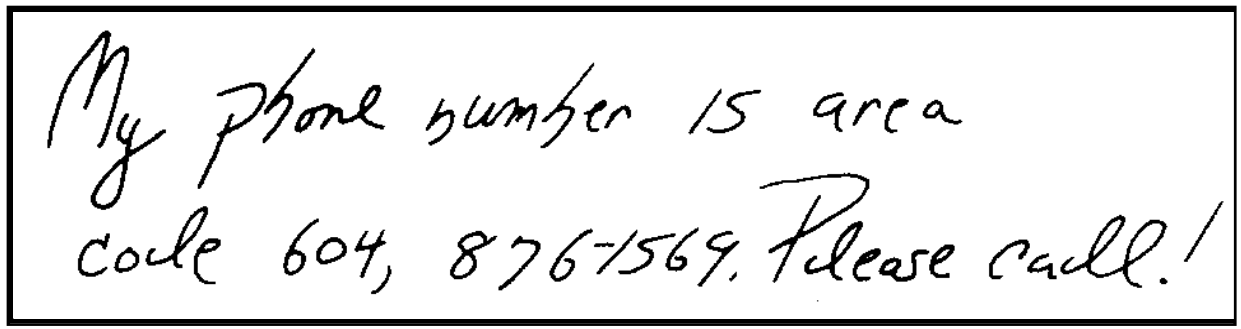

- **Context is Needed to Resolve Ambiguity**

# Another Look at Handwriting Recognition

**Pixel Patterns**



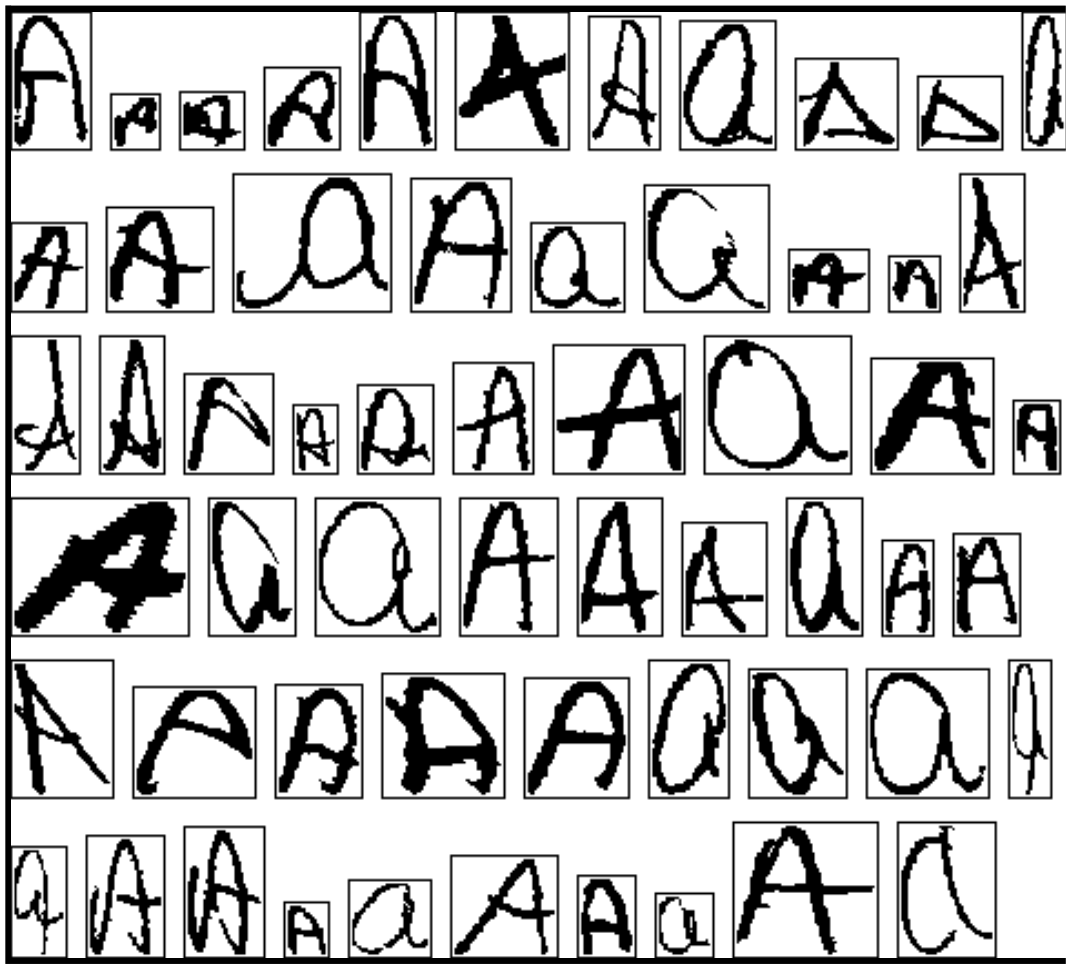
**Assignment is not Unique**




**Need to Allow Multiple Hypotheses for Character Class Assignment**

**Principle of Least Committment**

# Uncertainty in Class "A" Definition



**All are perfectly good A's**

# Fuzzy Set Theoretic Solutions Support the Two Principles of David Marr for the Design of Intelligent (Vision) Algorithms:

## Principle of Least Commitment

**Don't do something that may later have to be undone**

**"If the Principle of Least Commitment has to be disobeyed, one is either doing something wrong or something very difficult"**

# Marr's Second Principle

## Principle of Graceful Degradation

**Degrading the data will not prevent the delivery of at least some of the answer**

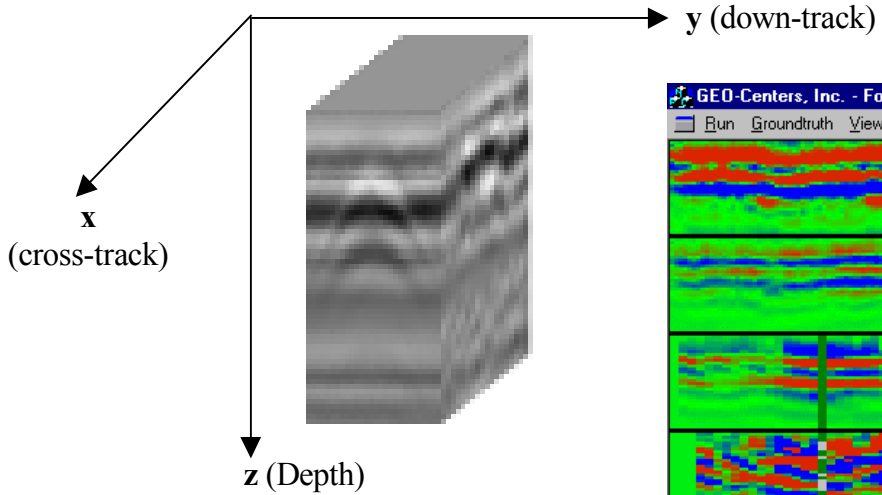
**Algorithms Should Be Robust**

**Processes should possess some degree of Continuity**

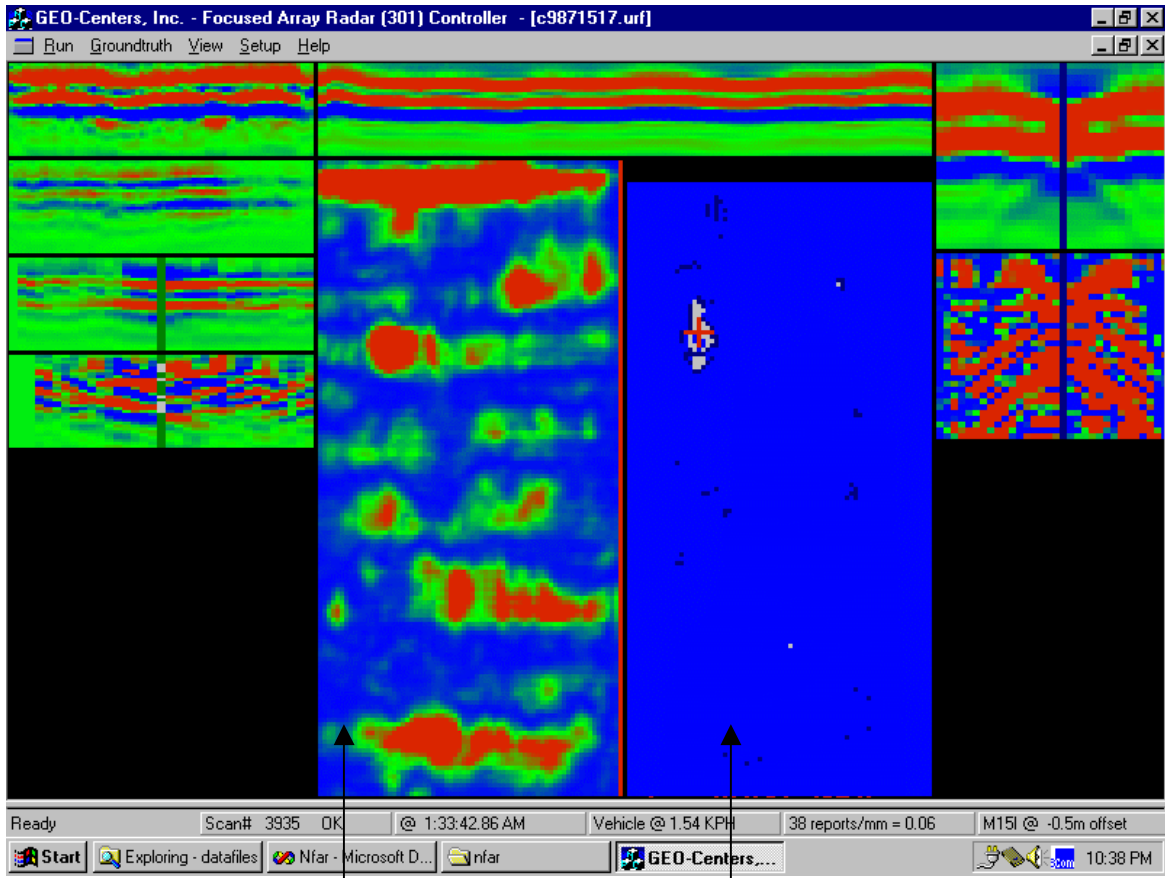
# Fuzzy Confidence in Pattern Recognition

- Many (most) pattern recognition algorithms can be "softened"
- Advantage comes from combination of multiple features
- Example in Land Mine Detection using 3-D volumetric GPR data
  - Fuzzy prototypes for mines/background developed
    - Based on idea that a hyperbola in radar should exist for idea mine
  - Confidence assigned in 3 space, aggregated and projected to surface

# Simple Fuzzy Set Approach to Landmine Detection



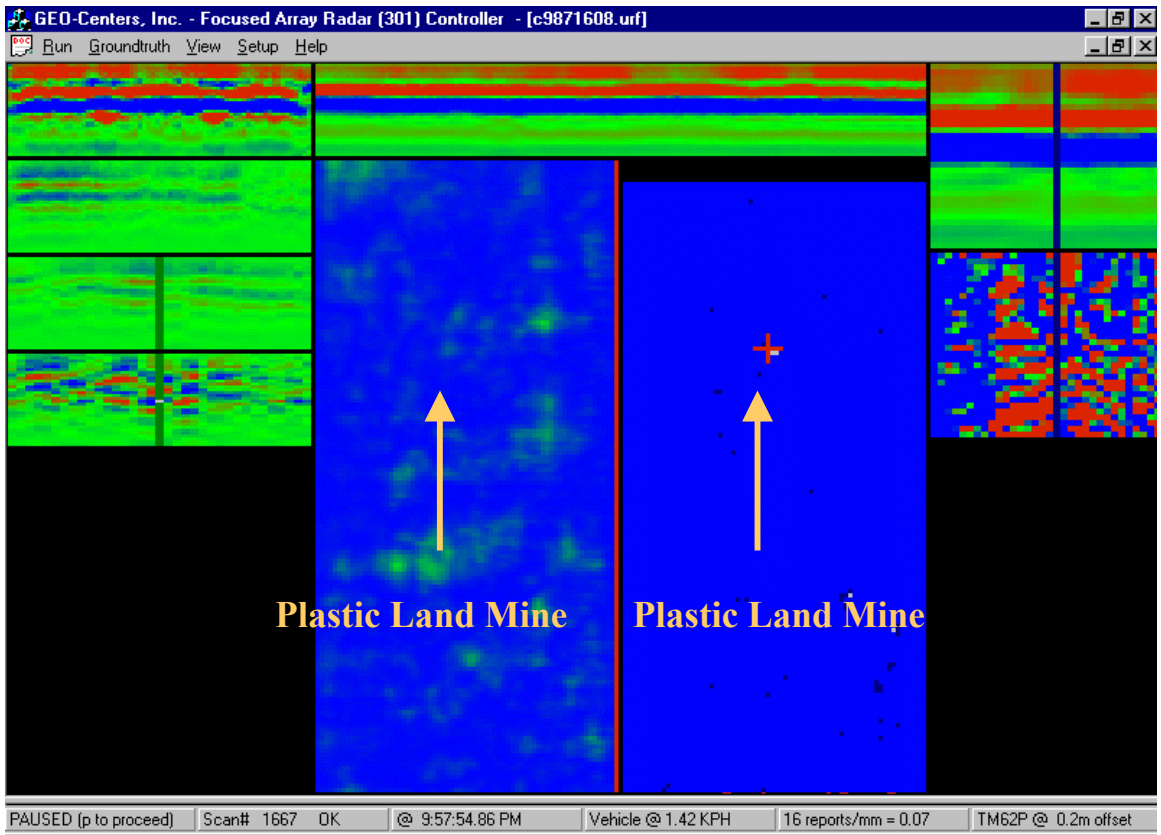

## DECREASE False Alarms




**Three dimensional Ground Penetrating Radar**

# Simple Fuzzy Set Approach to Landmine Detection

## Increased Detections (Verified through Blind Tests)




**OLD ALGORITHM** **MU ALGORITHM**

# Information Fusion

- **Wald gave a definition of information fusion**
  - **"Information Fusion is a formal framework in which are expressed means and tools for the alliance of data originating from different sources".**

- **The *different* sources of information**
  - **Different Sensor Systems**
  - **Different classifiers**
    - **Statistical**
    - **Deterministic**
    - **Fuzzy**
  - **Different features for the same classifier**
  - **Human Intelligence**

# What Fusion methods will we discuss?

- **Fuzzy Integrals**
  - **Non-linear combination of information source confidence with (possibly subjective) estimates of worth of subsets of the information sources**
  - **Will consider numeric and linguistic (higher order) fusion**
- **Fuzzy Logic**
  - **Rule-based approaches, attempting to model human decision making**

# Fuzzy Integrals

- Sources of information in a set X (sensors, features, algorithms, etc.)
- Worth of sources comes from a Fuzzy Measure: $g: 2^X \rightarrow [0,1]$
  - $g(\phi) = 0$ and $g(X) = 1$
  - $g(A) \leq g(B)$ if $A \subseteq B$
  - If $\{A_i\}$ is an increasing sequence of subsets of X, then

$$\lim_{i \to \infty} g(A_i) = g\left( \bigcup_{i=1}^{\infty} A_i \right)$$

# Fuzzy Integrals

- **A fuzzy measure g is called a Sugeno measure ($g_\lambda$-fuzzy measure) if additionally:**

  **For all $A, B \subseteq X$ with $A \cap B = \phi$,**

$$g(A \cup B) = g(A) + g(B) + \lambda \cdot g(A) \cdot g(B) \text{ for some } \lambda > -1$$

- **For any Sugeno fuzzy measure $\lambda$ can be uniquely determined for a finite set X by solving**

$$1 + \lambda = \prod_{i=1}^{n} \left(1 + \lambda g^i\right)$$

- **where $X = \{x_1, \ldots, x_n\}$ and $g^i = g(\{x_i\})$ interpreted as the (possibly subjective) importance of the single information source $x_i$ in determining the evaluation of a class hypothesis**

# Where do the fuzzy measures come from?

- **Heuristic assignment**
  - **easier for λ-measures: Only need to specify "worth" of each source**
- **Direct training of the densities**
  - **Reward/punishment approach**
  - **Genetic algorithms**
- **Optimize entire measure for Choquet integral via Quadratic Programming – more on this later**

# The Sugeno Fuzzy Integral

- Let X be a finite set of information sources, h:X→[0,1] a partial evaluation function, and g:$2^X$→[0,1] a fuzzy measure

- The Sugeno fuzzy integral of h with respect to g is

$$\int_S h \circ g = \sup_{\alpha \in [0,1]} \{\min(\alpha, g(H_\alpha)\}$$

where $H_\alpha = \{x | h(x) \geq \alpha\}$

- The Best Pessimistic Agreement between evidence and the worth of the evidence

# The Sugeno Fuzzy Integral

- Since X is finite, re-order X so that $h(x_1) \geq h(x_2) \geq \cdots \geq h(x_n)$
- Then the Sugeno Fuzzy Integral is

$$\int_S h \circ g = \bigvee_{i=1}^{n} \left( h(x_i) \wedge g(H_i) \right)$$

where $H_i = \{x_1, \cdots, x_i\}$

- The measures of these n subsets can be found recursively for λ-measures
- Problem: Not a true extension of Lebesgue integral (if measure is additive)

# Choquet Fuzzy Integral

- Let $X=\{x_1,\ldots,x_n\}$, $g$ be a fuzzy measure and $h$: $X \rightarrow [0, 1]$

- Finite Choquet Integral

$$\int_c h \circ g = \sum_{i=1}^{n} g(X_i)\left[h(x_i) - h(x_{i+1})\right]$$

or

$$\int_c h \circ g = \sum_{i=1}^{n} h(x_i)\left[g(X_i) - g(X_{i-1})\right]$$

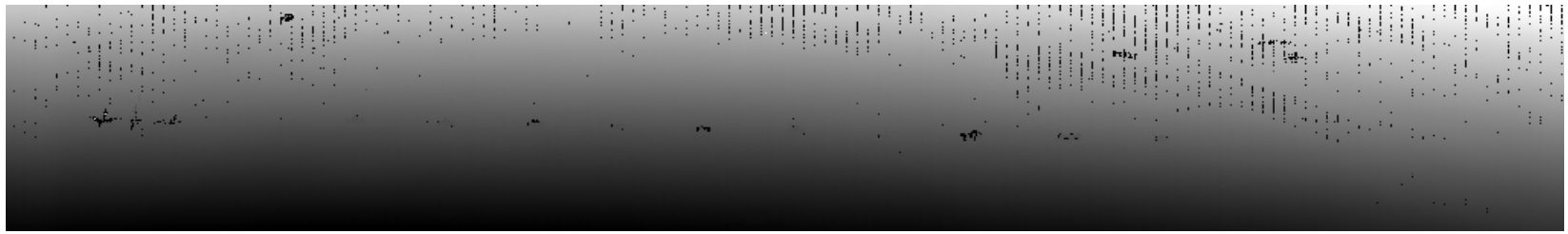
- where $h(x_{n+1}) = 0$, $g(X_0) = 0$, $h(x_1) \geq h(x_2) \geq \ldots \geq h(x_n)$ and $X_i = \{x_1, \ldots, x_i\}$

- Looks like a linear combination, but depends on the sort

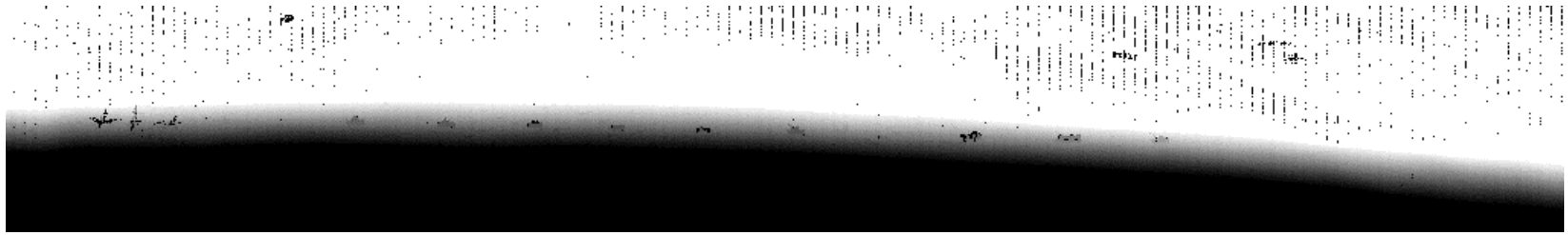
# Fuzzy Integrals for Image Processing

- **Both Sugeno and Choquet fuzzy integrals used for non-linear image filtering**
- **Can implement morphological filters, all linear and order statistic filters, all linear combination of order statistic filters, etc.**
- **Used instead of means and variances for "size-contrast" filters**
- **Used to segmentation and to fuse multiple detectors**

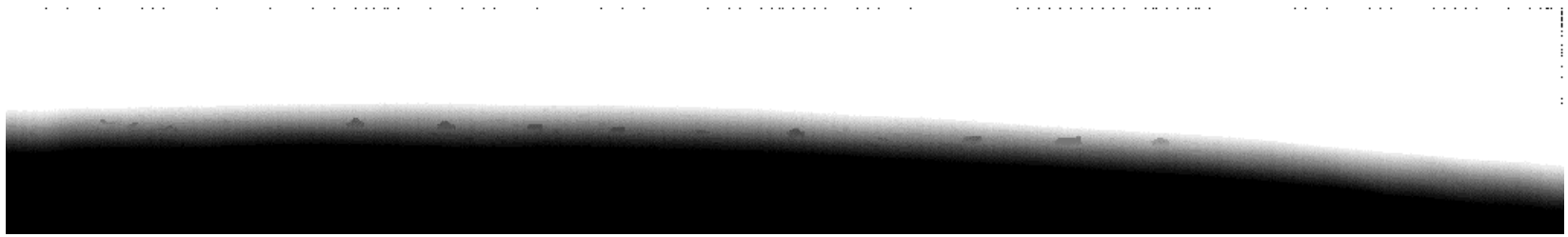
# Choquet Integrals in LADAR ATR



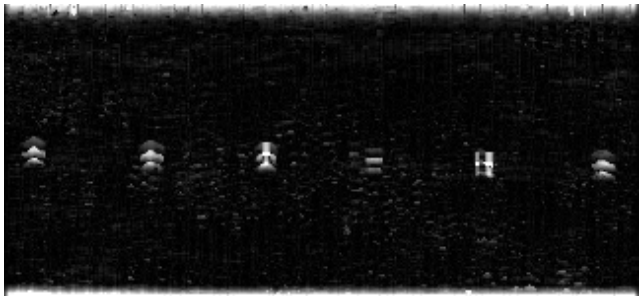
Original



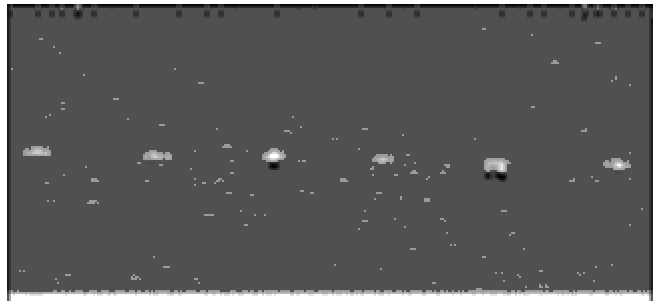
Scaled (for Viewing)



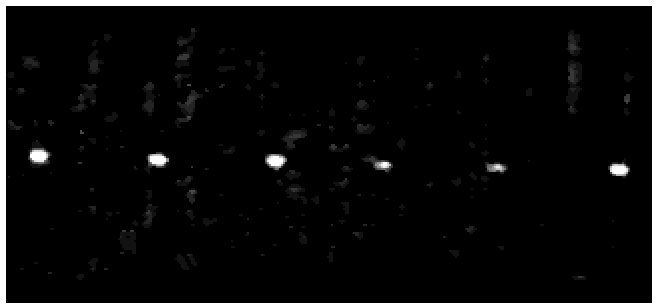
Preprocessed by Choquet Filter



Detector 1

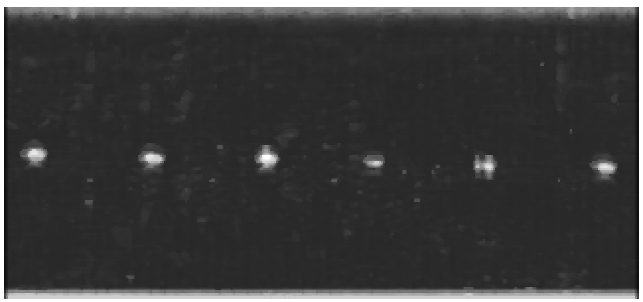


Detector 2



Detector 3

Keep detections up but lower false alarms

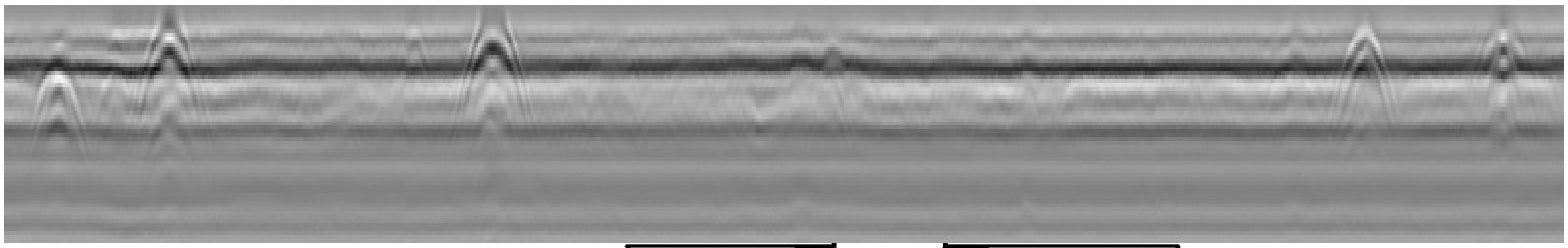


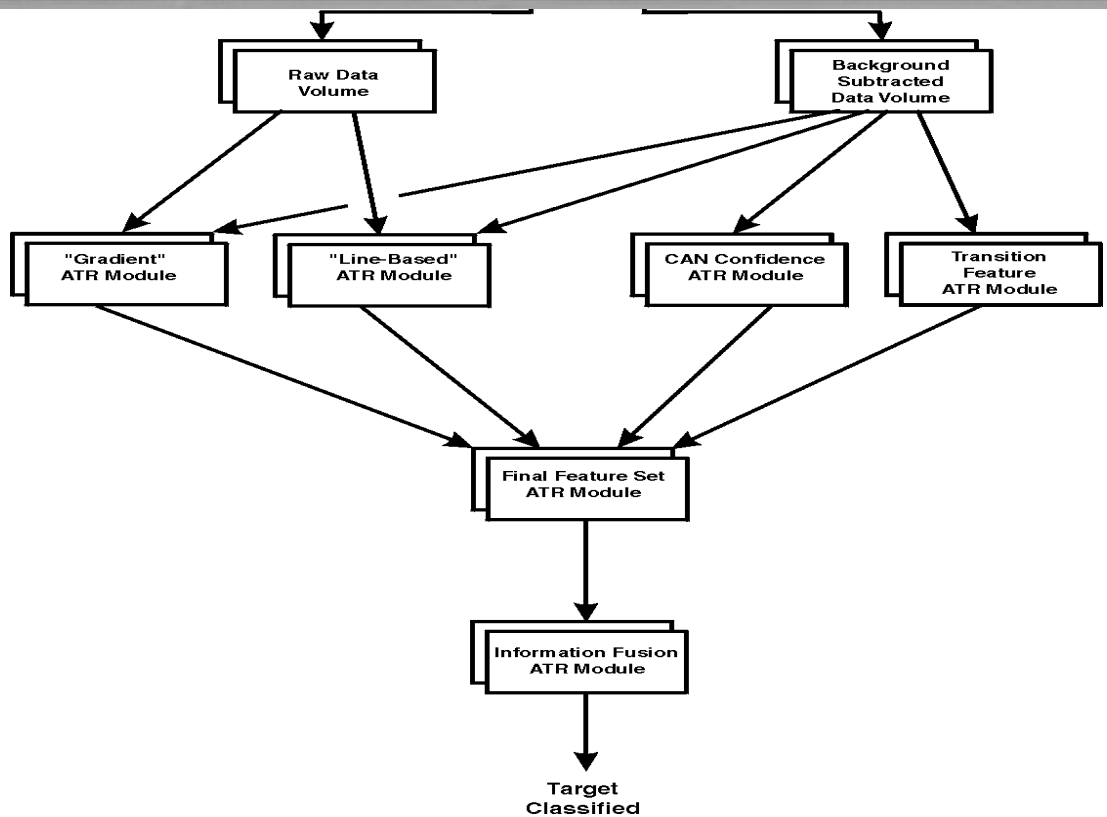
Choquet Fusion

Thresholded and Dilated

# Information Fusion for Demining

## Fusing Outputs of Multiple Algorithms on GPR





# Numeric Fusion on MU Signature Library

## Improvement given by Algorithm Fusion with Choquet and Sugeno Fuzzy Integrals

**Reduction in False Alarms over Line-Based ATR (One Basic ATR)**

| DETRANGE | LINE | CHOQ | SUG | CHOQ_RED | SUG_RED |
|---|---|---|---|---|---|
| 206-207 | 82 | 97 | ND | -18% | NA |
| 187-193 | 42 | 27 | 27 | 36% | 36% |
| 183-186 | 40 | 22 | 24 | 45% | 40% |
| 181-183 | 32 | 22 | 24 | 32% | 26% |
| AVERAGE REDUCTION IN FALSE ALARM | | | | 24% | 34% |

ND = No detections in that range

# Numeric Fusion on MU Signature Library

## Reduction in False Alarms over Gradient-Based ATR

(Another Basic Detection Algorithm)

| DETRANGE | GRAD | CHOQ | SUG | CHOQ_RED | SUG_RED |
|---|---|---|---|---|---|
| 206-207 | 82 | 97 | ND | -18% | NA |
| 201-203 | 64 | 67 | 42 | -4% | 35% |
| 198-199 | 57 | 45 | 36 | 22% | 37% |
| 195-197 | 54 | 37 | ND | 31% | NA |
| 191-193 | 52 | ND | 27 | NA | 48% |
| 187-189 | 50 | 27 | ND | 46% | NA |
| 181-183 | 50 | 22 | 18 | 56% | 64% |
| AVERAGE REDUCTION IN FALSE ALARMS | | | | 22% | 46% |

ND = No detections in that range

# Minimum Classification Error for Information Fusion

## Hot off the Press!

Mendez-Vazquez, A., Gader, P., Keller, J., and Chamberlin, K., "Minimum classification error training for Choquet integrals with applications to landmine detection",

*IEEE Transactions on Fuzzy Systems*, in Press, 2007

# Choquet Integral for Fusion

- **Potential Problem when used for classification**
  - **Fusion of evidence for each class is treated independently**
  - **Measures are learned for each class**
    - **Provide target values for each intgral**
  - **Class-Integral values are not directly comparable**
  - **Similar to problem with HMMs**

- **Like with HMMs, use discriminative training**
  - **Train all integrals simultaneously**

# The Choquet Integral

- **For this development, use the alternate formulation**
- **Let the original sources of information $\{x_1, \ldots, x_n\}$ be reordered into $\{x_{(1)}, \ldots, x_{(n)}\}$ such that**

$$0 \le f(x_{(1)}) \le \cdots \le f(x_{(n)})$$

- **Let $A_{(i)}$ be defined as** $A_{(i)} = \{x_{(i)}, \ldots, x_{(n)}\}$

- **Then**

$$C_g(f) = \sum_{i=1}^{n} g(A_{(i)}) \cdot (f(x_{(i)}) - f(x_{(i-1)}))$$

# Least Squared Error (LSE)

- **Error Function depends on knowing desired outputs**

$$E^2 = \frac{1}{2}\left( \sum_{x \in \omega_1} (C_g(f(x)) - \alpha_1)^2 + \sum_{x \in \omega_2} (C_g(f(x)) - \alpha_2)^2 \right)$$

- **Use Quadratic Programming to find the measure**
- **Or use Gradient Descent for Sugeno λ-measures**
  - **Need some of the derivatives to follow**

# Minimum Classification Error

- **In minimum classification error (MCE):**
  - **We do not consider cost functions that used fixed desired outputs**
  - **W depend on a function of differences of Choquet integral in the dissimilarity measure**

$$d_i(x) = -C_{g_i}(f_i(x)) + \max_{j \neq i}\{C_{g_j}(f_j(x))\}$$

*Note that under this function, smaller values of* $d_i(x)$ *give higher confidence that* $\omega$ *belongs to the class i.*

# Minimum classification error (Cont)

- **Use a smooth, monotonically increasing function that is differentiable almost in everywhere**

$$l_i(x) = \begin{cases} \dfrac{1}{1+e^{(-\alpha d_i(x))}}, & d_i(x) > 0 \\ 0 & , d_i(x) \le 0 \end{cases}$$

# Final Cost function

- **The final cost function to be minimized looks like**

$$E = \sum_{x \in C_1} l_1(x) + \ldots + \sum_{x \in C_n} l_n(x)$$

- **Use Gradient Descent**

# Derivative of the cost function

- **Differentiating the cost function with respect to the densities we have**

$$\frac{\partial E}{\partial g_i^j} = \sum_{x \in C_1} l_1(x)(1 - l_1(x)) \frac{\partial d_1(x)}{\partial g_1^j} + \ldots + \sum_{x \in C_n} l_n(x)(1 - l_n(x)) \frac{\partial d_n(x)}{\partial g_n^j}$$

**where $g_i^j$ represent the *j*th density for class *i*.**

- The term $\frac{\partial d_n(x)}{\partial g_n^j}$ is equal to

$$\frac{\partial d_n(x)}{\partial g_n^j} = \begin{cases} \frac{\partial C_{g_k}(f_k(x))}{\partial g_i^j} & \text{if } \mathrm{k} = \mathrm{i} \\ \frac{\partial C_{g_i}(f_i(x))}{\partial g_i^j} & \text{if } \mathrm{k} \neq \mathrm{i} \text{ and } C_{g_i}(f_i(x)) = \max_{s \neq k}\{C_{g_s}(f_s(x))\} \\ 0 & \text{if } \mathrm{k} \neq \mathrm{i} \text{ and } C_{g_i}(f_i(x)) \neq \max_{s \neq k}\{C_{g_s}(f_s(x))\} \end{cases}$$

# Problem: What is $\dfrac{\partial C_{g_k}(f_k(x))}{\partial g_i^j}$ ?

- **This is simply**

$$\frac{\partial C_g(f)}{\partial g_j} = \sum_{i=1}^{n} \frac{\partial g(A_{(i)})}{\partial g_j}\left(f(x_{(i)}) - f(x_{(i-1)})\right)$$

- **Hence, we still need to know**

$$\frac{\partial g(A_{(i)})}{\partial g_j}$$

# The derivative of $\dfrac{\partial g\left(A_{(i)}\right)}{\partial g_j}$

- We know that

$$g(A_{(i)}) = g\left(\left\{x_{(i)}\right\} \cup A_{(i+1)}\right) = g_{(i)} + g(A_{(i+1)}) + \lambda g_{(i)} g(A_{(i+1)})$$

- Thus

$$\frac{\partial g\left(A_{(i)}\right)}{\partial g_j} = \begin{cases} 1 + \lambda g\left(A_{(i)}\right) + g_{(i)} g\left(A_{(i+1)}\right)\dfrac{\partial \lambda}{\partial g_j} + \left(1 + \lambda g_{(i)}\right)\dfrac{\partial g\left(A_{(i+1)}\right)}{\partial g_j}, & \text{if } (\mathrm{i}) \neq n, (i) = j \\ g_{(i)} g\left(A_{(i+1)}\right)\dfrac{\partial \lambda}{\partial g_j} + \left(1 + \lambda g_{(i)}\right)\dfrac{\partial g\left(A_{(i+1)}\right)}{\partial g_j}, & \text{if } (\mathrm{i}) \neq n, (i) \neq j \\ 1 & \text{if } (\mathrm{i}) = n,\ j = n \\ 0 & \text{if } (\mathrm{i}) \neq n,\ j \neq n \end{cases}$$

## The derivative $\frac{\partial \lambda}{\partial g_j}$

- **Using the fact that the Sugeno λ-measure has the property**

$$(\lambda + 1) = \prod_{i=1}^{n} (1 + \lambda g_i)$$

- **We have (using implicit differentiation)**

$$\frac{\partial \lambda}{\partial g_j} = \frac{\lambda^2 + \lambda}{\left(1 + g_j \lambda \left[1 - (\lambda + 1) \sum_{i=1}^{n} \frac{g_i}{1 + g_j \lambda}\right]\right)}$$

# Gradient Descent for the MCE

- **With the previous results, we implement a Gradient Descent algorithm for the MCE**

$$\vec{g} = \vec{g} + \alpha \nabla E$$

- **Does not need complex formulations**
  - **Required for non-linear optimization**

# Experiments in Landmine Detection

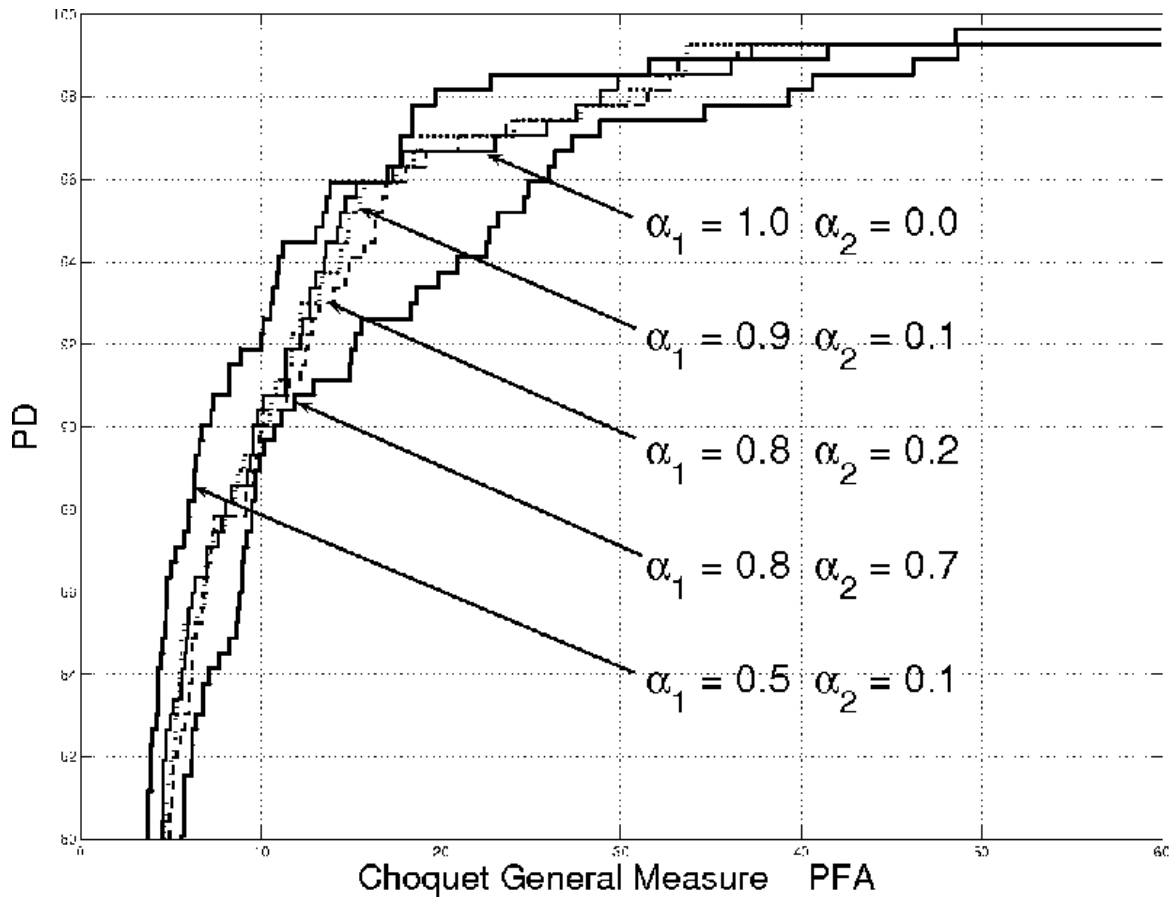
- **The Least Squared Error(LSE) and MCE training methods were applied to a two-class algorithm fusion problem in landmine detection**
  - **The landmine detection problem involved processing Ground Penetrating Radar (GPR) sensor returns**

- **The data set contained 2422 8-dimensional samples of confidence values from each of the eight detection algorithms. The data set contained 271 mines samples and 2151 non-mine samples.**

# Experimental Results

## Least squared error sensitive to desired outputs

- ROC curves showing the sensitivity to desired outputs for **LSE** in the full measure. $\alpha_1$ and $\alpha_2$ represent the desired outputs for mine and non-mines.



# Advantages of MCE

- MCE is not sensitive to desired outputs, as is LSE
- MCE is, in general, better than the individual detectors, with a range of improvement between 0.44% and 65.07%
- On this data, improvement of MCE over LSE training
  - Ranges between 11.06% and 37.51%
- It is possible for the Sugeno $\lambda$-measure and the general measure trained with LSE to be as good as the one trained by MCE. For this to happen, it is necessary to have a set of correct desired outputs!

# Experiments as a Classifier

- The MCE training was also applied to the Iris and Breast Cancer data with results in
- K. Xu, Z. Wang, P.-A. Heng, and K.-S. Leung, "Classification by nonlinear integral projections," IEEE Transactions on Fuzzy Systems, ol. 11, no. 2, pp. 187–2001, 2003

| *Method* | *Iris Data(%)* | *Breast Cancer Data(%)* |
|---|---|---|
| Linear | 2 | 29 |
| Quadratic | 2.7 | 34.4 |
| Nearestneighbor | 4 | 34 |
| Bayes independent | 6.7 | 28.2 |
| Bayes quadratic | 16 | 34.4 |
| Neuronal net | 3.3 | 28.5 |
| PVM rule | 4 | 22.9 |
| QUAD | 3.3 | 31.5 |
| CLMS | 4 | 27.1 |
| HLMS | 4.7 | 22.6 |
| WCIPP | 4 | 26.2 |
| **MCE** | **4** | **22.73** |

- **But, sometimes numbers aren't enough**
  - **Confidence is High**
  - **IR sensor is Weak**
  - **Mine location is Around (1.5, 3.7)**
- **Need Higher Order Constructs**
  - **Type II fuzzy sets and linguistic vectors**
  - **Fuzzy logic rules**

Story from the Local Newspaper

Doctors Make Decisions from This type of Information

Shouldn't Computers?

# Facing up to flu season

As the American Lung Association kicks off a month-long flu campaign this week, predictions are the flu season will be especially severe this year. How to tell what ails you:

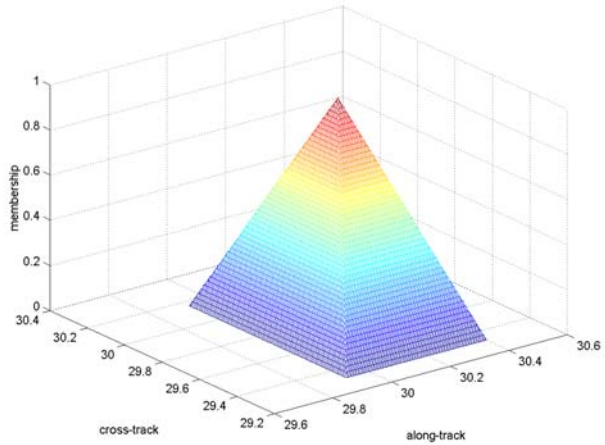
| SYMPTOMS | COLD | FLU | PNEUMONIA |
|---|---|---|---|
| FEVER | rare | 100.4°-104°, sudden onset, lasts 3-4 days | may or may not be high |
| ACHES/PAINS | slight | usual, often severe | occasional |
| FATIGUE/ WEAKNESS | quite mild | extreme | may occur |
| RUNNY/ STUFFY NOSE | common | sometimes | not characteristic |
| SNEEZING | usual | sometimes | not characteristic |
| SORE THROAT | common | sometimes | not characteristic |
| CHEST PAIN, COUGH | mild to moderate | can become severe | frequent/ may be severe |

Source: American Lung Association

*Stories by Leslie Ansley*

# Linguistic Sensor Fusion

- **What if Entities to be Fused Don't Lend Themselves to Numeric Representations?**
  - **Position Estimates - Commonly Use Halos or Windows**
  - **Confidence May Not be Known Exactly**
    **- varies among classifiers/sensors**
- **Generalize These Values to Fuzzy Sets**
  - **Linguistic Vectors**
    - **Generalize halos and confidence intervals**



# Linguistic Choquet Fuzzy Integral

- So, How Do We Fuse Linguistic Vectors of Sensor Output?
- Both Position Confidence and Detection Confidence
- Extend the Choquet Integral to These Vectors
- Based on Extension Principle and Decomposition Theorem of Fuzzy Set Theory
  - Plus interval arithmetic if we use fuzzy numbers - nice fuzzy subsets of reals

# Background

- **Extension Principle (How to extend functions from numbers to fuzzy sets)**

$$\mu_Z(y) = \sup_{x_1 \otimes x_2 = y} \min(\mu_A(x_1), \mu_B(x_2))$$

- **For fuzzy numbers (normal convex fuzzy subsets of the reals)**
  - **Use interval arithmetic and the decomposition theory:**

$$A \otimes B = \bigcup_{\alpha \in [0,1]} [A \otimes B]_\alpha$$

$$= \bigcup_{\alpha \in [0,1]} [A]_\alpha \otimes [B]_\alpha$$

- **Linguistic Vectors**
  - **Vectors of Non-interactive Fuzzy Numbers**

# Linguistic Choquet Fuzzy Integral

- **Extend regular Choquet integral to a linguistic Choquet fuzzy integral [Grabisch94b]:**
- **Let *h*: X→ℑ([0, 1]) where ℑ([0, 1]) is a fuzzy power set of [0, 1], and**

$$[h_i]_\alpha = \left[\underline{[h_i]_\alpha}, \overline{[h_i]_\alpha}\right] \quad \text{for } 1 \le i \le n \text{ and } 0 \le \alpha \le 1.$$

- **Then**

$$\int_c h \circ g = \bigcup_{\alpha \in [0,1]} \left[\int_c h \circ g\right]_\alpha$$

**where**

$$\left[\int_c h \circ g\right]_\alpha = \left[\int_c \underline{[h]_\alpha} \circ g, \int_c \overline{[h]_\alpha} \circ g\right]$$

# Linguistic Choquet Fuzzy Integral

- **How to Compute Densities? Two properties:**
  - **Property 1. The more uncertain a sensor is, the less important it is for the fusion**
  - **Property 2. The further away the sensor's detection location is from the true location in training data, the less important that sensor is for the fusion**

# Linguistic Choquet Fuzzy Integral

- **How to Compute Densities? Two properties:**

**Property 1. The more uncertain a sensor is, the less important it is for the fusion**

– **Use U-uncertainty measure[Klir95] to compute density of fuzzy number $A_j$ in a particular dimension.**

$$g_u^j = 1 - \frac{\sum_{i=2}^{p} (\alpha_i - \alpha_{i+1}) \log_2 \left| [A_j]_{\alpha_i} \right|}{\log_2 \left( \left| \sup(A_j) \right| \right)}$$

# Linguistic Choquet Fuzzy Integral

- **Property 2.** **The further away the sensor's detection location is from the true location in training data, the less important that sensor is for the fusion**
  - **Know that** $\left\{ [A_j]_{\alpha_1} \subseteq [A_j]_{\alpha_2} \subseteq \cdots \subseteq [A_j]_{\alpha_p} \right\}$
  - **The distance between** $[A_j]_{\alpha_i}$ **and** $[\beta]_{\alpha_i}$ **is**

$$d_{A_j}(\alpha_i) = \max\left( \left| \underline{[A_j]_{\alpha_i}} - \beta \right|, \left| \overline{[A_j]_{\alpha_i}} - \beta \right| \right)$$

  - **where** $\beta$ **represents the actual location in a particular dimension**

# Linguistic Choquet Fuzzy Integral

- **The densities of fuzzy number *j* in a particular dimension relating to the distance are:**

$$g_d^j = \sum_{i=1}^{p} (\alpha_i - \alpha_{i+1}) P\left( [A_j]_{\alpha_i} \right) \qquad \text{(Generalized Expected Value)}$$

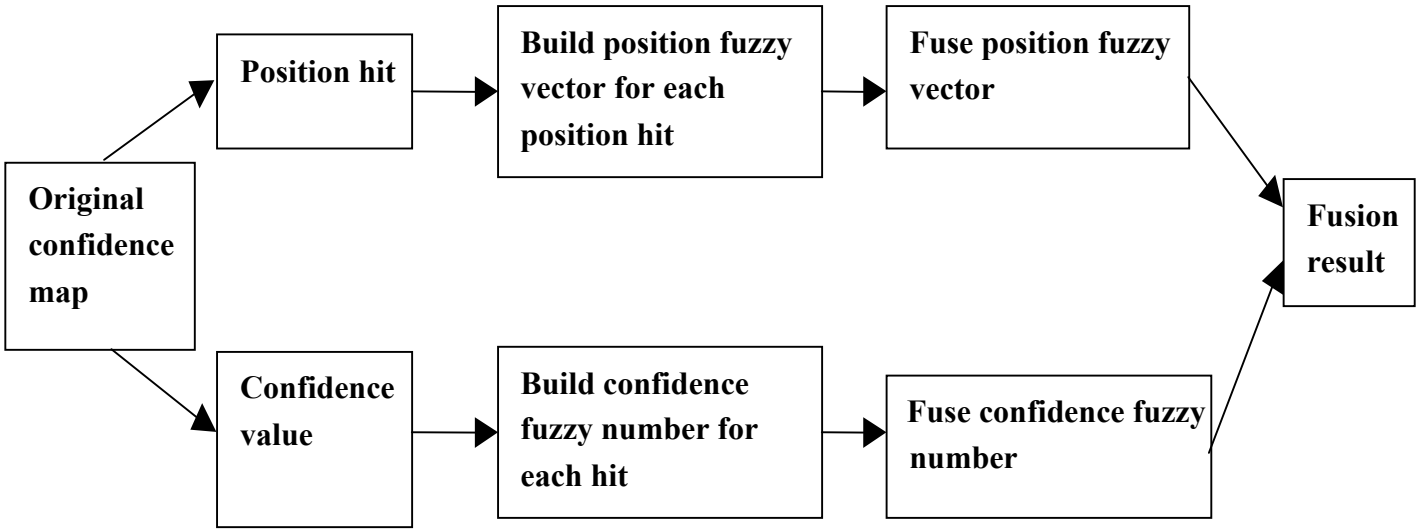
**where**

$$P\left( [A_j]_{\alpha_i} \right) = 1 - \frac{d_{A_j}(\alpha_i)}{\sum_k d_{A_k}(\alpha_i)}$$

- **The final densities of fuzzy number $A_j$ in a particular dimension relating to both properties are:**

$$g^j = g_u^j \times g_d^j$$

# Linguistic Fusion of Real Algorithm Results

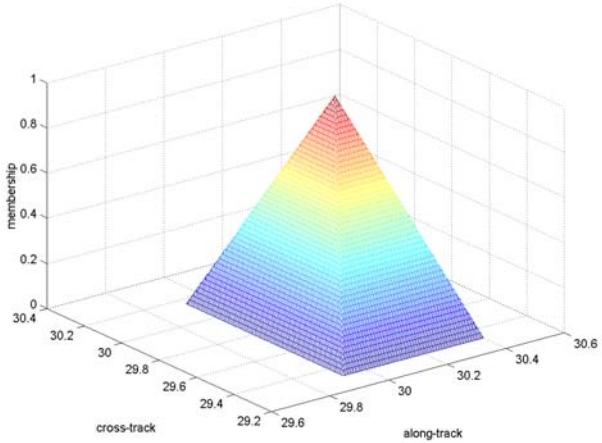


# Linguistic Fusion Experiment

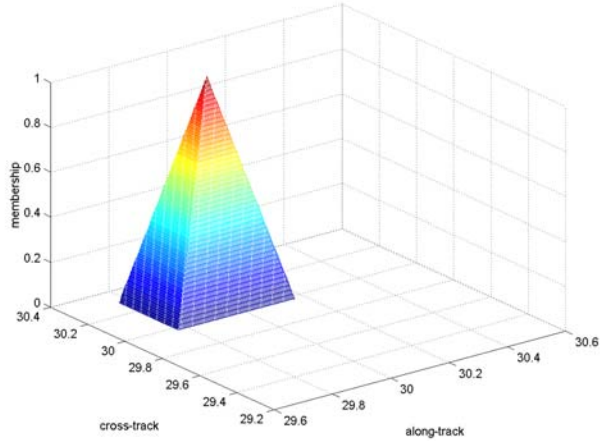
- Fuse the information from the Fuzzy ATR, the HMM and the CMSNN
  - Generated from GPR data at Aberdeen site
- Threshold confidences from three algorithms
- Use the centroid of each connected component to be the hit position ($x$,$y$)
- Fuzzify each dimension of each hit location:
  - 5x5 window with center at ($x$,$y$), average column-wise to get vector in along-track and average row-wise to get vector in cross-track
  - Build normal convex fuzzy sets for these vectors
- Build the mask with the center of mask at the current hit position
- Fuse locations inside the mask

# Example of Position Fusion

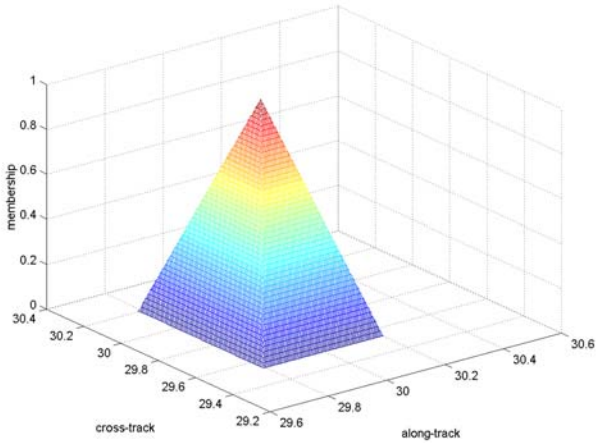
Sensor 1

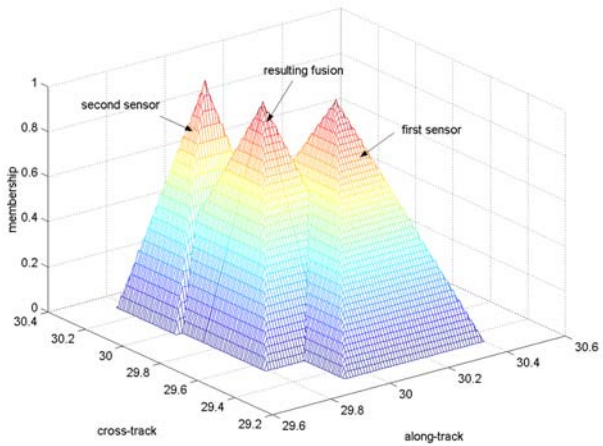



Sensor 2



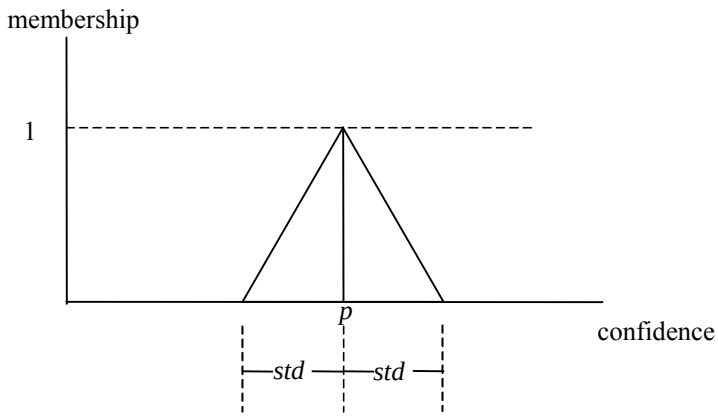

Fusion result



# Linguistic Fusion

- **Build the confidence fuzzy number**
  - **Pick the maximum confidence *p* in corresponding 5x5 window**
  - **Compute standard deviation *std* of the confidence in that neighborhood**

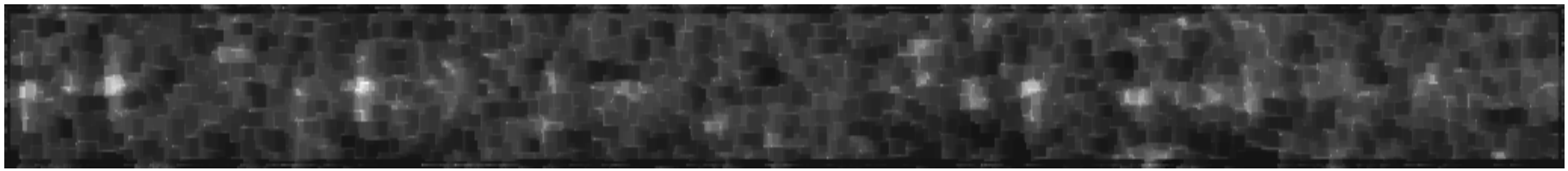



- **Fuse confidences corresponding to the position linguistic vectors that are combined at a particular "position"**

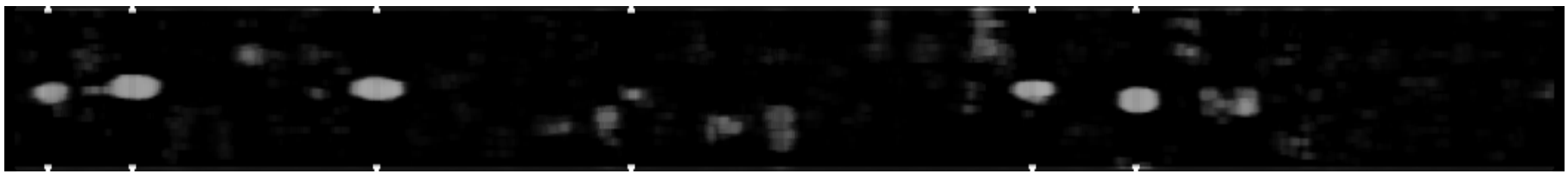
# Linguistic Fusion

- **Original confidence map from three algorithms on GPR data on Mine Lane collected at a US site**

**Fuzzy ATR**



**HMM**



**CMSNN**

# Linguistic Fusion

- **After Thresholding: The Fuzzy ATR has 6 detections and 27 false alarms, The HMM has 5 detections and 10 false alarms, and The CMSNN has 6 detections and 6 false alarms**

**Fuzzy ATR**

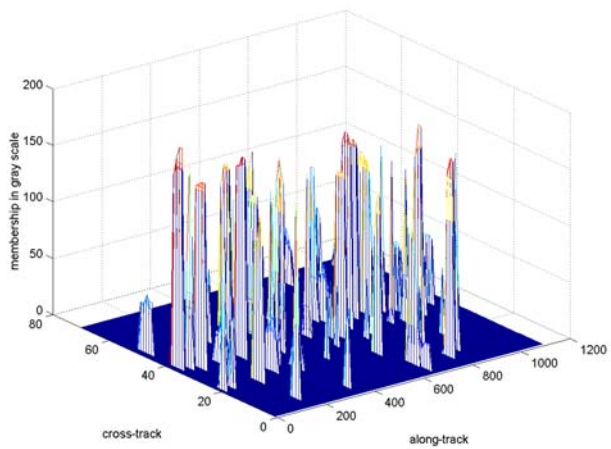
**HMM**

**CMSNN**

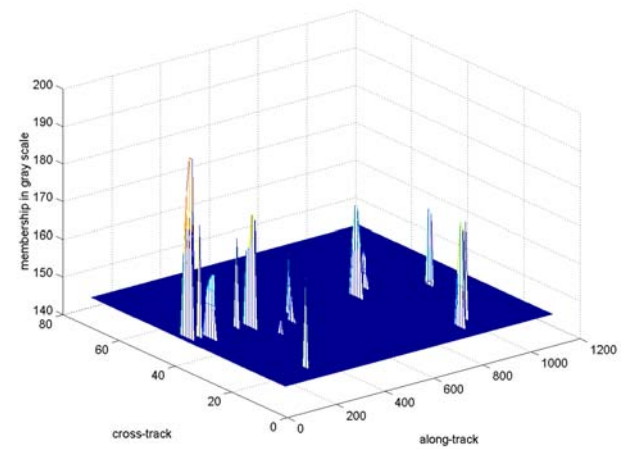
# Linguistic Fusion

- **After Fusion**

**After Thresholding: 6 detections and 2 false alarms**



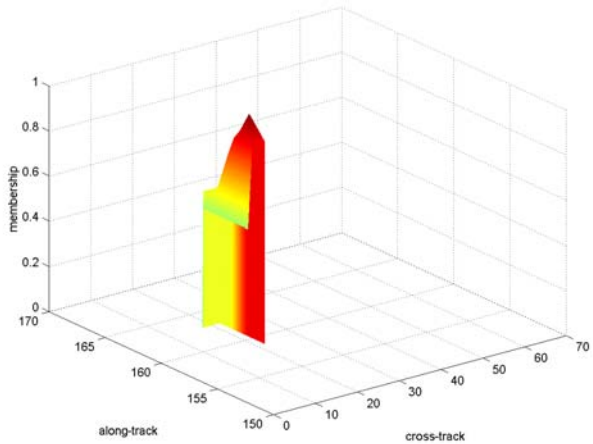

**Before thresholding**



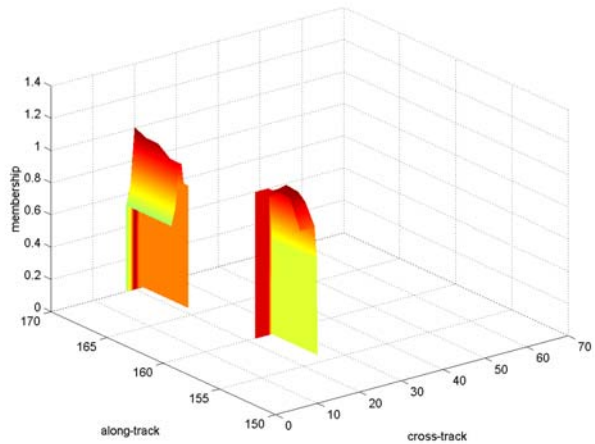

**After thresholding**

# Linguistic Fusion

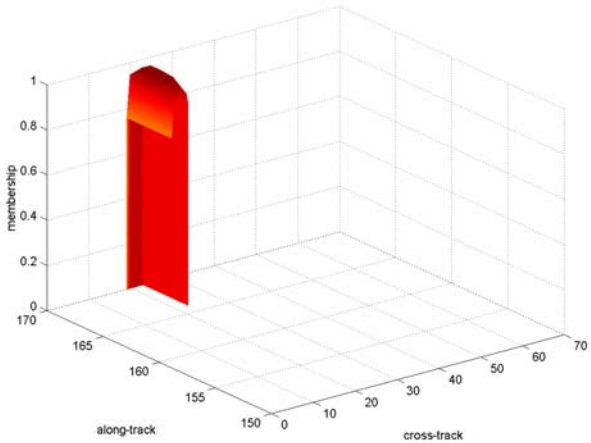
- **Example of false alarm eliminated by fusion**



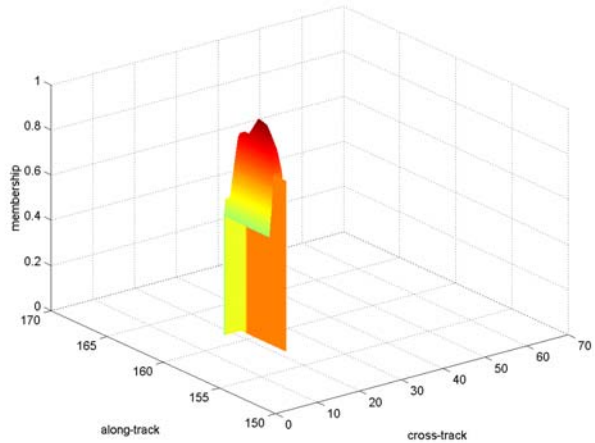

**Position from Fuzzy ATR**




**Position from HMM**

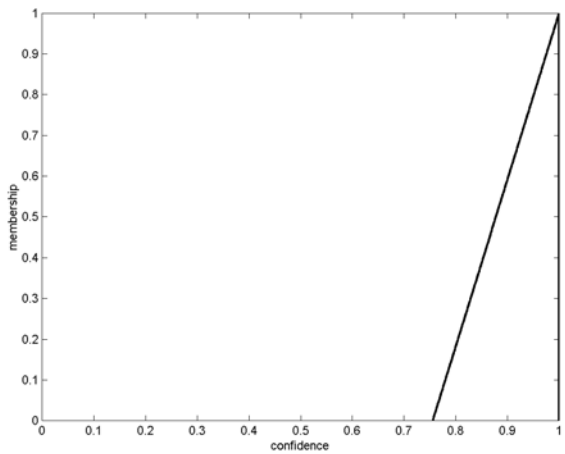



**Position from CMSNN**

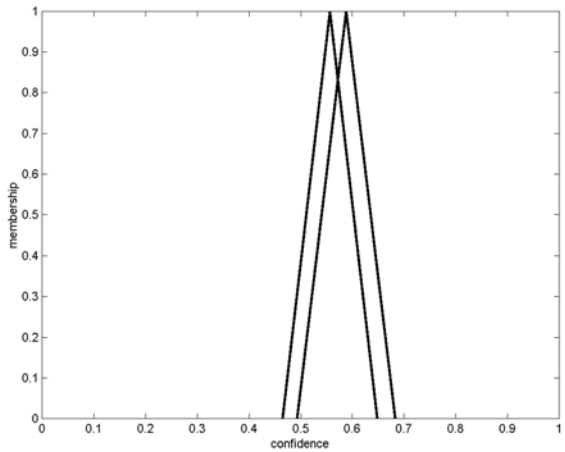



**Position fusion result**

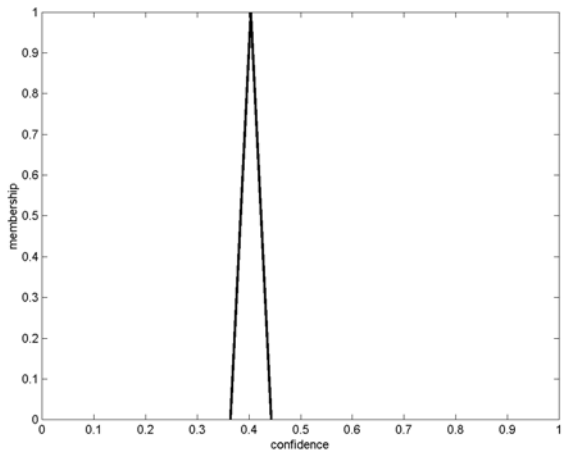
# Linguistic Fusion



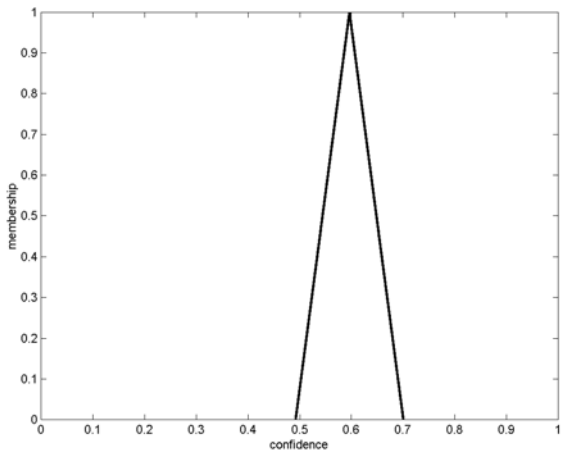

Confidence from Fuzzy ATR




Confidence from HMM

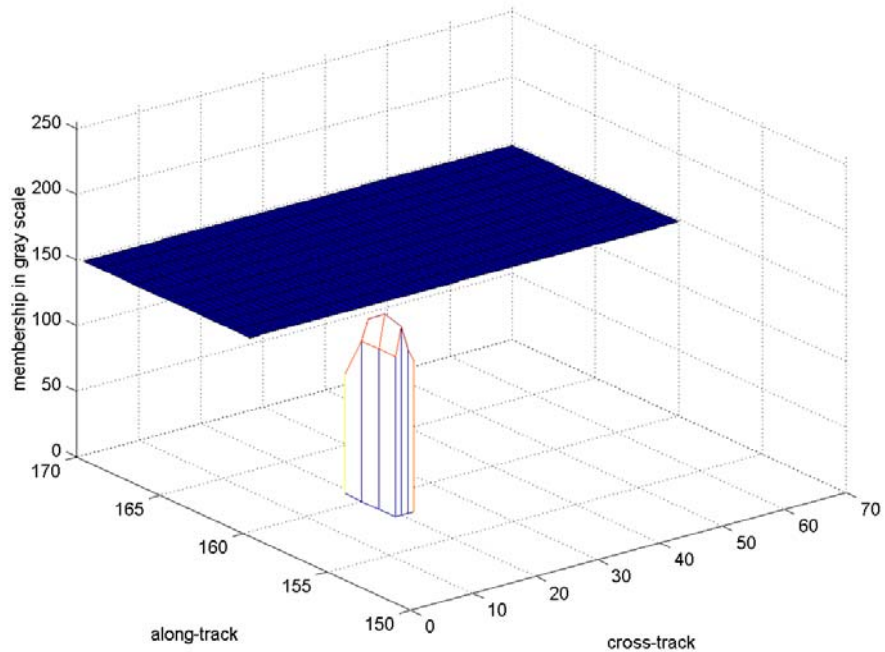



Confidence from CMSNN




Confidence fusion result

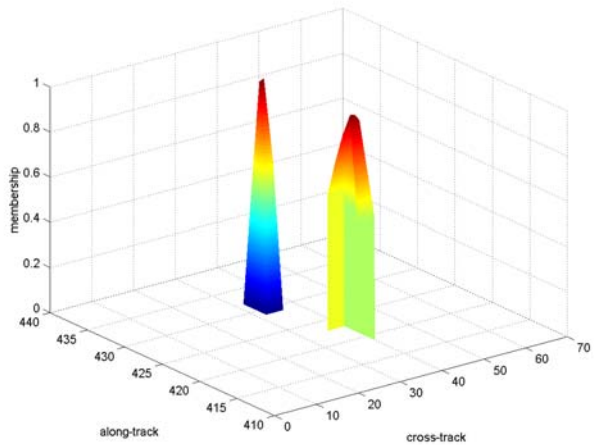
# Linguistic Fusion



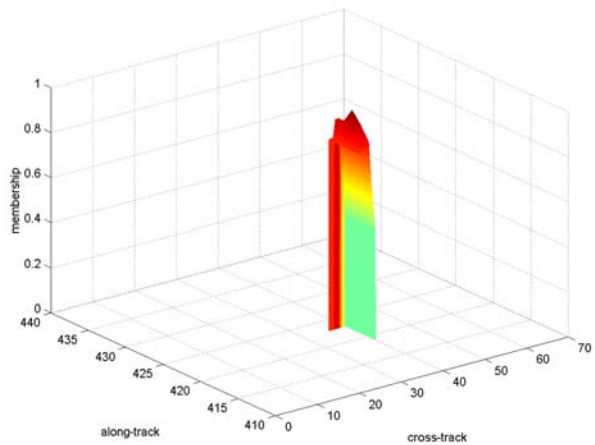

**After multiplying the peak of confidence fusion with the membership of position fusion**

# Linguistic Fusion

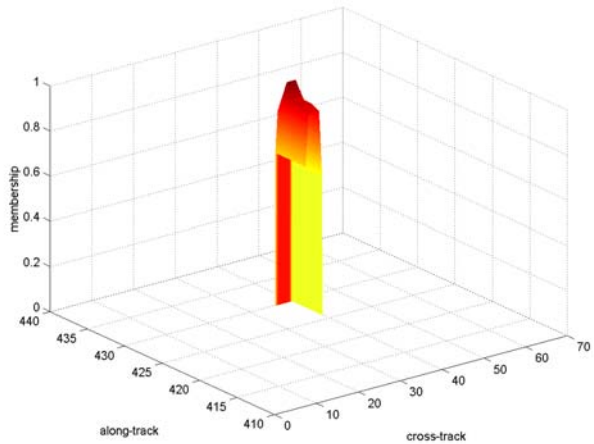
- **Example of mine that fusion can detect while one algorithm cannot**




**Position from Fuzzy ATR**

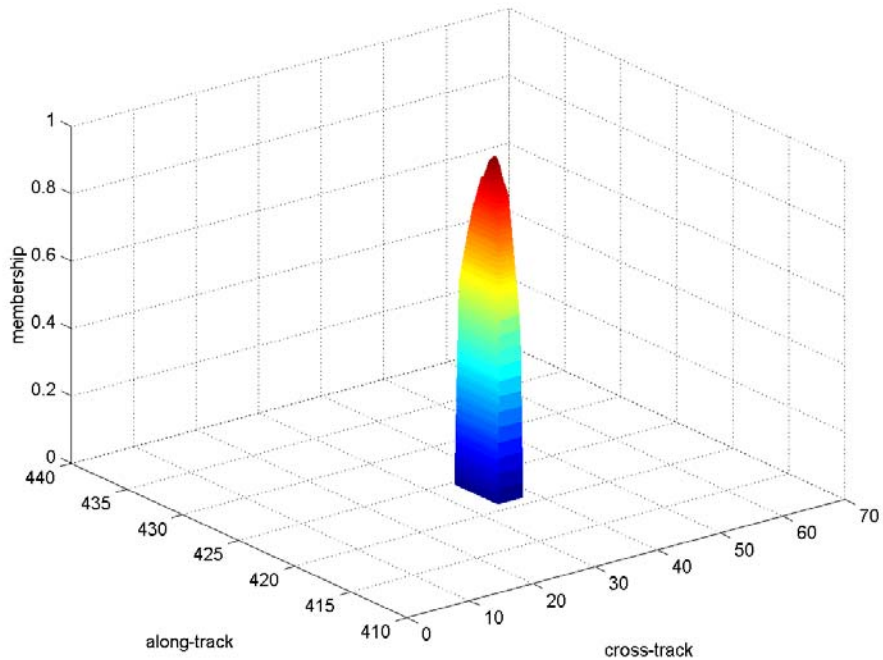



**Position from HMM hit location at (38,419). Considered as noise by HMM when scoring because there are only 2 pixels above threshold.**



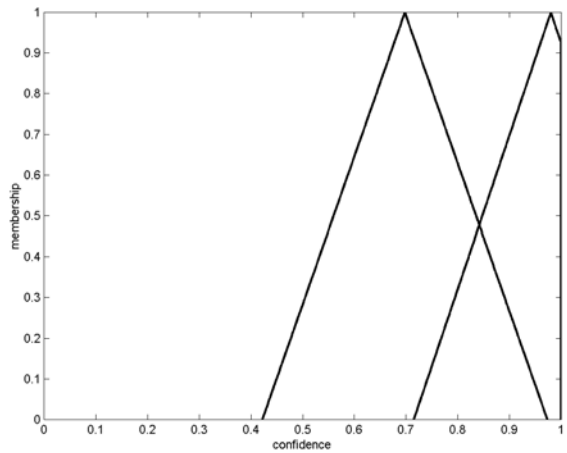

**Position from CMSNN**

# Linguistic Fusion

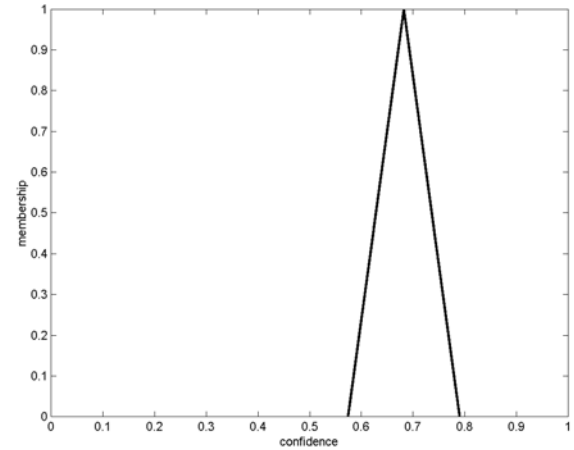



**Position fusion result**

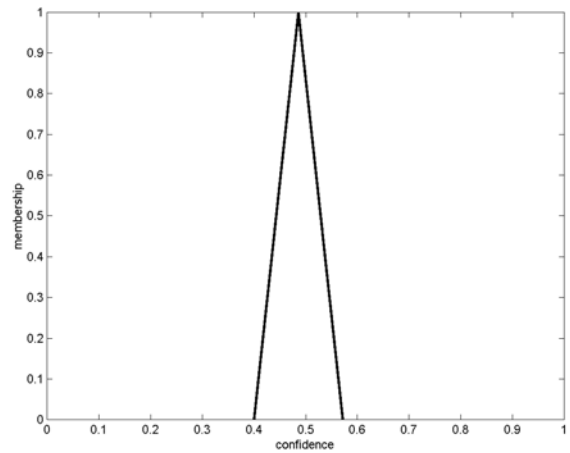
# Linguistic Fusion



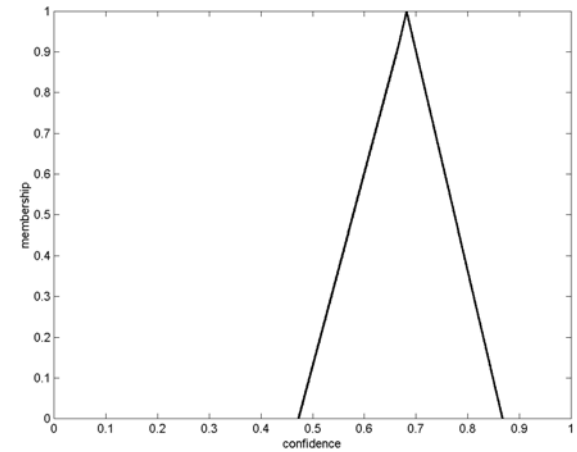

Confidence from Fuzzy ATR




Confidence from HMM

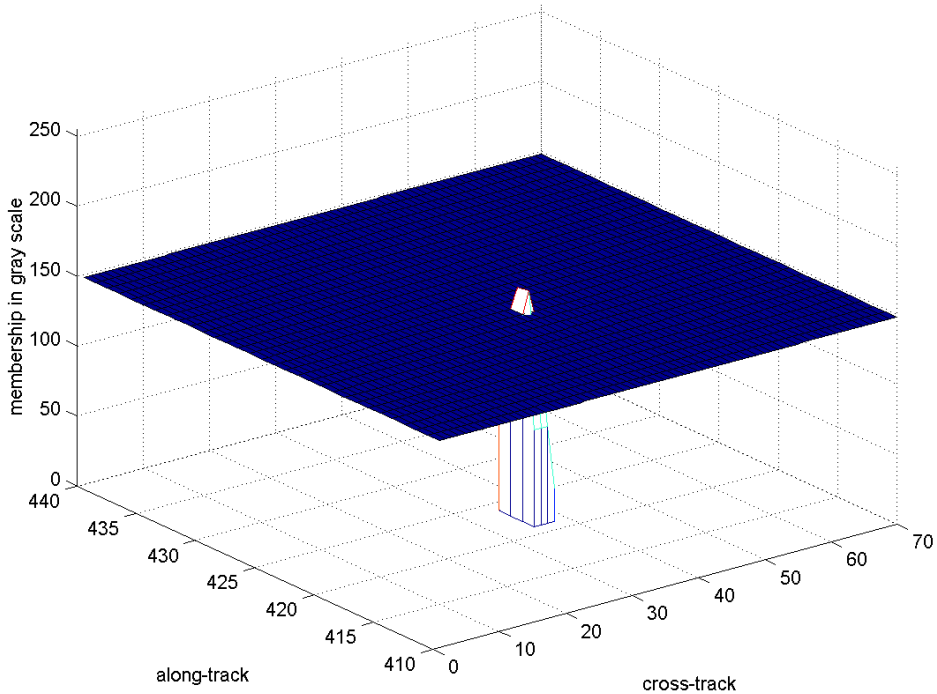



Confidence from CMSNN




Confidence fusion result

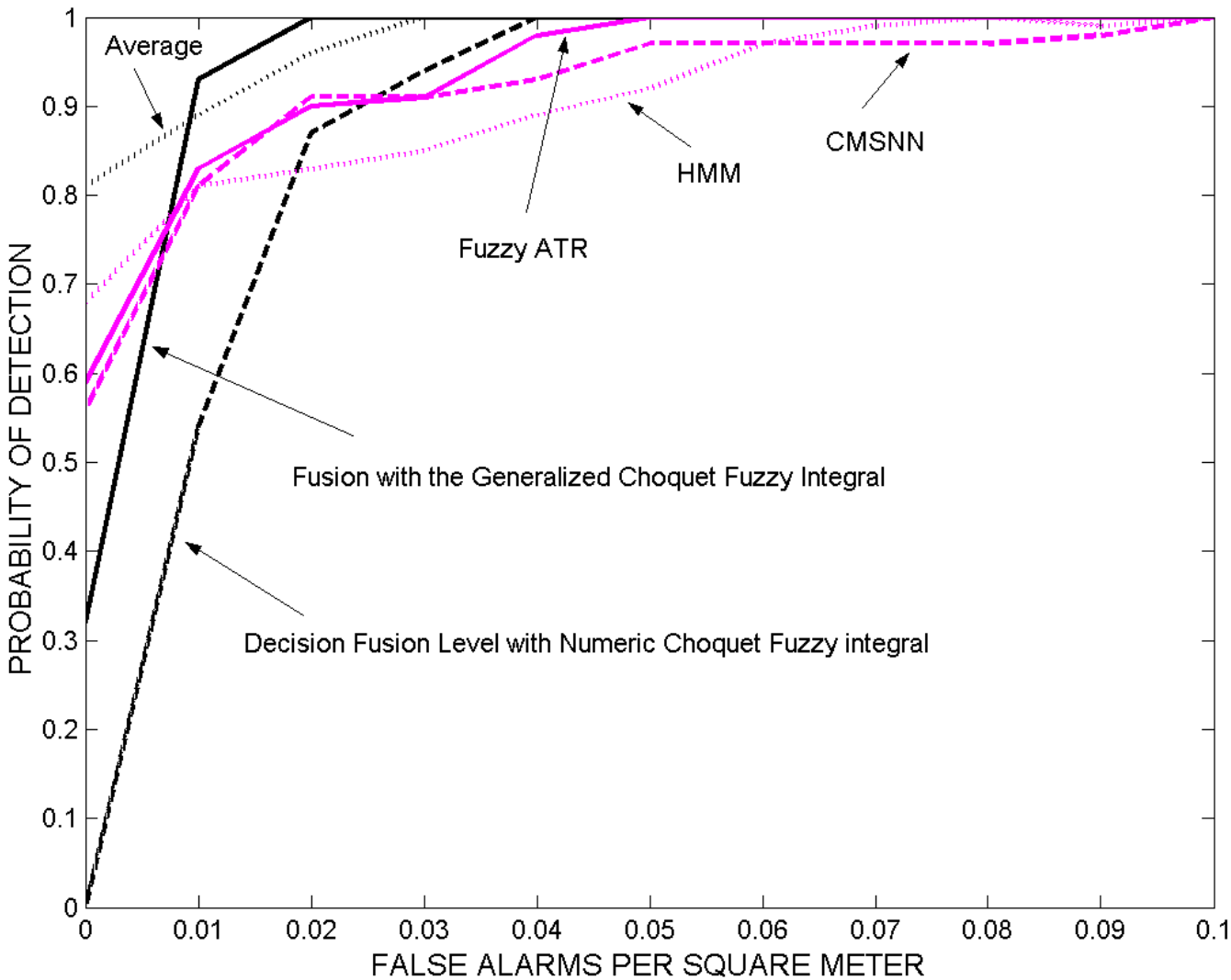
# Linguistic Fusion




**After multiplying the peak of confidence fusion with the membership of position fusion**

# Linguistic Fusion - The Payoff

- **Run on ten lanes (16 passes) from US Army temperate site**



# Predictive Sensor Fusion

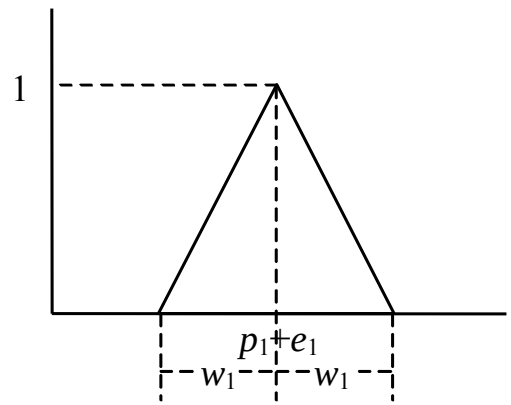
- **Two Questions:**
  - **Can we predict the value added by fusing multiple sensor system outputs together?**
  - **Can we specify the needed characteristics of a new sensor system to add to an existing suite to a gain a desired improvement in performance?**
- **Approach**
  - **Use Monte Carlo simulations based on actual data from 2 GPRs to study effect of adding a third senor**

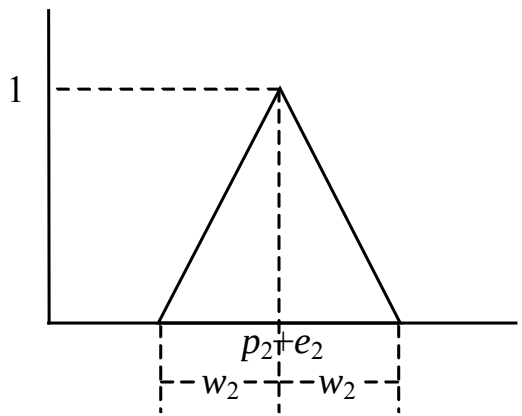
# Experimental Setup

- **Summary results from [Rotondo98] on the position accuracy of GPR detection systems for two contractors during the Advanced Technology Demonstrations at Aberdeen**
  - **We chose the on-road information generated at Aberdeen**
  - **The data is in terms of the position error over the lanes modeled by a gaussian density function.**
  - **Hence, we have the expected performance of these two sensor/algorithm suites in terms of mean and standard deviation of position error in two dimensions.**

# Predictive Fusion Experiment

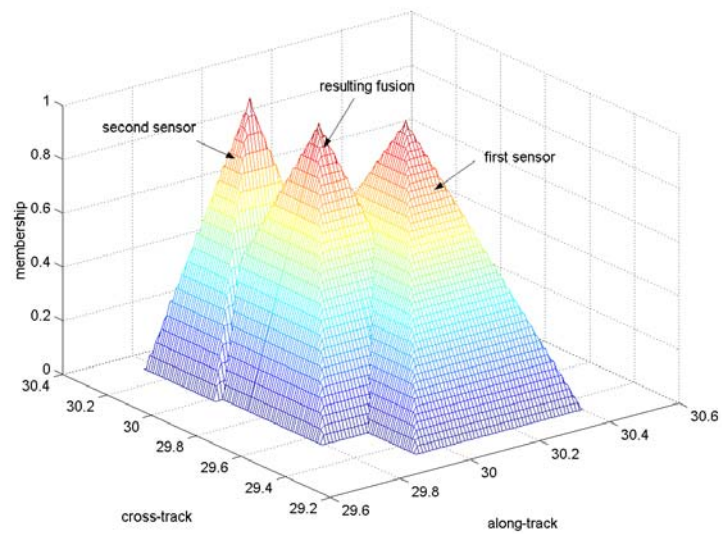
- **Chose position resolution from GPR on on-road generated at Aberdeen**
- **Generate error position ($e_1$,$e_2$) in along-track and cross-track**
- **Assume the actual location is ($p_1$,$p_2$). The detection location is at ($p_1$+$e_1$,$p_2$+$e_2$)**
- **Fuzzify each dimension of each location using Guassian distributions, e.g., N(0.1,$\sigma_1$) or N(0.21, $\sigma_1$) etc.**



# Predictive Fusion Experiment

- An example of the linguistic position fusion of the 2 sensor systems



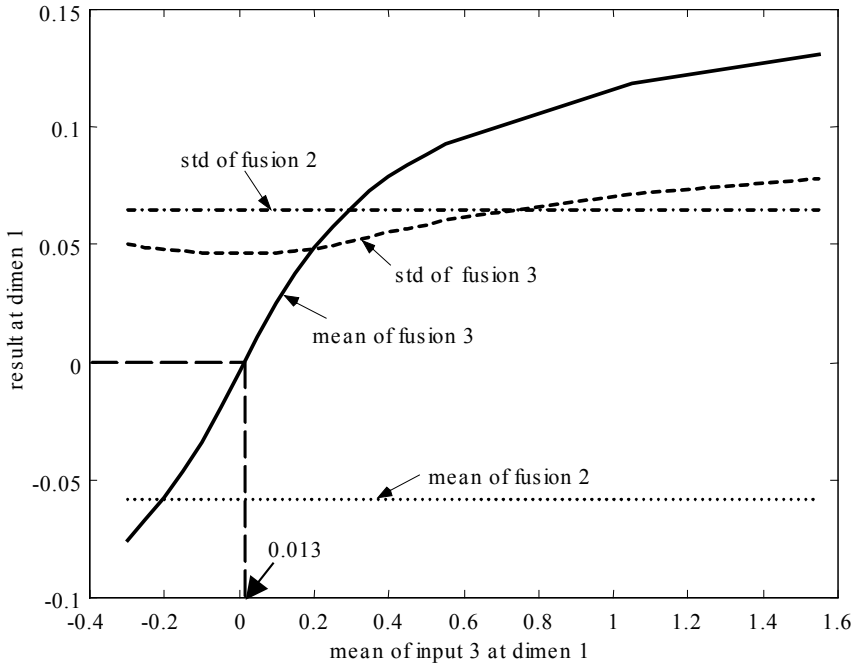

- Defuzzy each fuzzy number by picking the midpoint of the core. Also, compute the radius of the support to represent the width of that fuzzy number
- Generate 2000 trials according to a given Guassian distribution.
- Collect the error and width, then compute mean and standard deviation

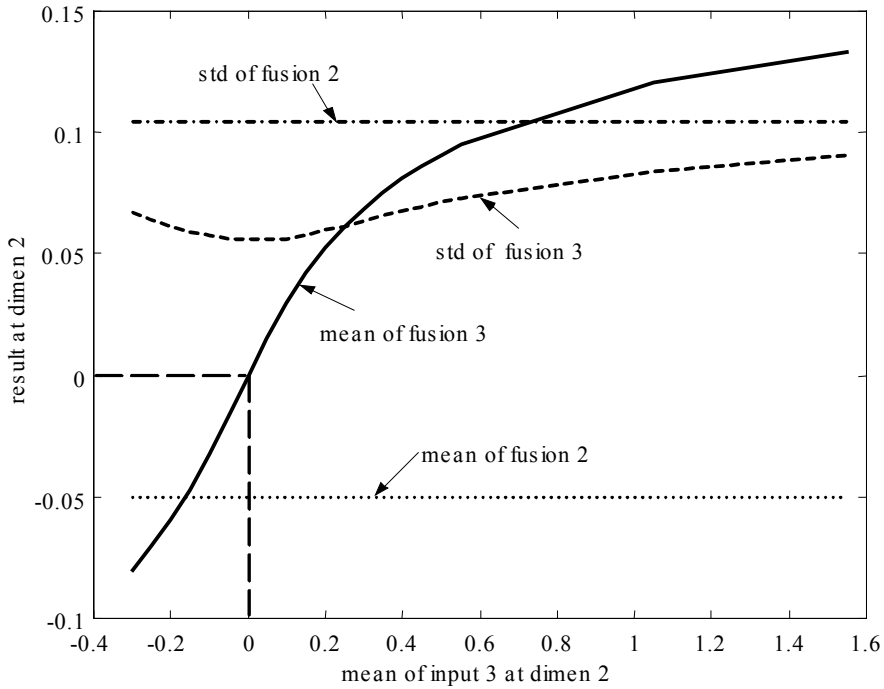
# Adding a "Third" System

- **Changing error mean of hypothetical sensor system 3**

| Sensor | Along-track | | Cross-track | |
|---|---|---|---|---|
| | Position resolution (mean(m),std (m)) | Width (mean(m),std (m)) | Position resolution (mean(m),std (m)) | Width (mean(m),std( m)) |
| 1 | (0.1,0.12) | (0.12,0) | (-0.06,0.21) | (0.21,0) |
| 2 | (-0.1,0.1) | (0.1,0) | (0.04,0.08) | (0.08,0) |
| 3 | ($x$,0.1) | (0.1,0) | ($x$,0.1) | (0.1,0) |

# Predictive Fusion Results



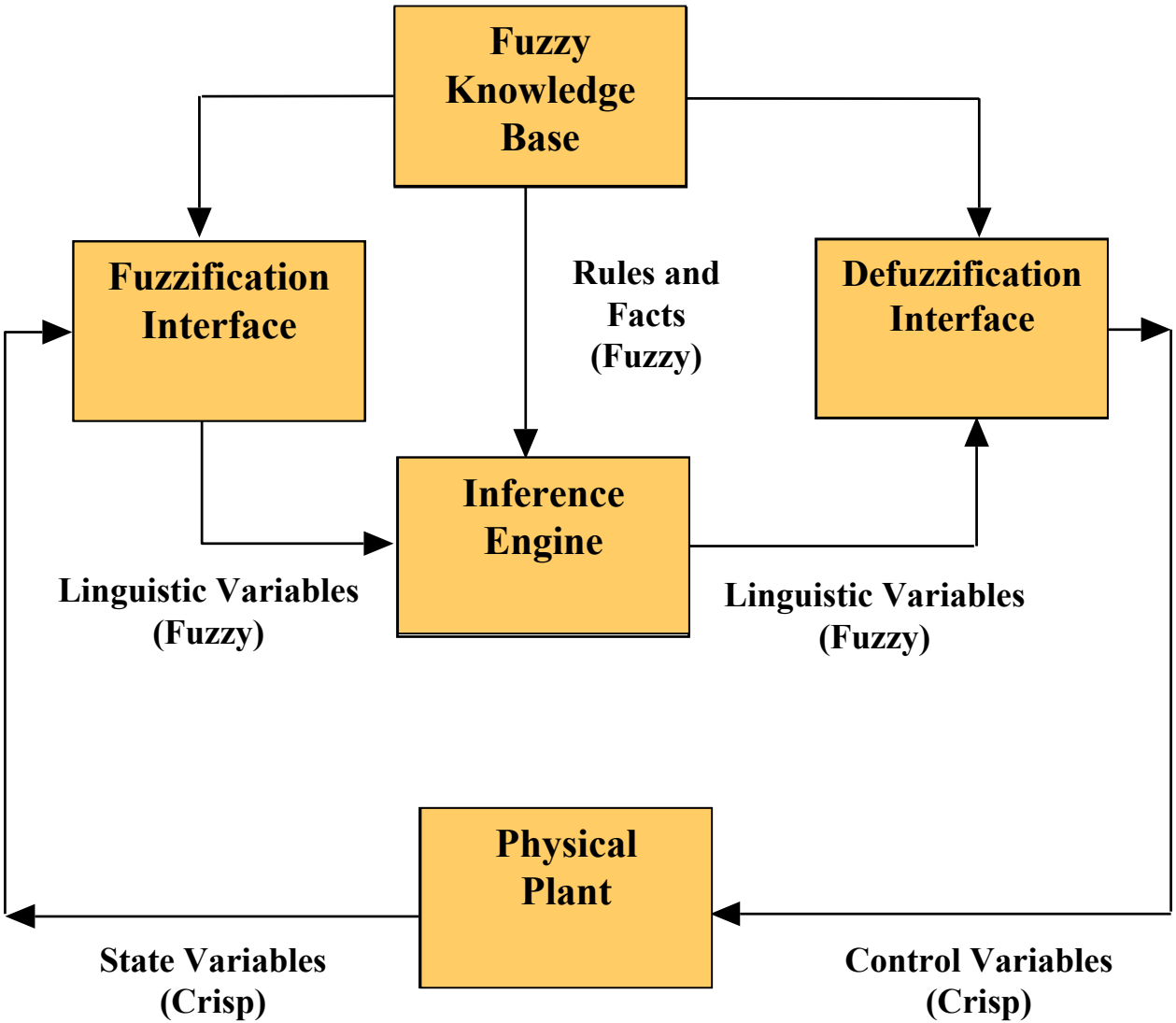



# Fuzzy Logic Systems

- **Rule-based approaches to combine evidence**
- **Rules used by "intelligent" systems may not contain**

  **PRECISE PREDICATES**
- **Propositions modeled via fuzzy set theory**
  - **Simple propositions, conjunctions, disjunctions, implication**
- **Rules of inference generalized to accommodate extensions from binary to continuous**

# Basic Configuration of a Fuzzy Logic Control System



# Knowledge Base contains the Translations of Linguistic Propositions into Fuzzy Logic Constructs

- **Simple Propositions** **"X is A"**

- **Compound Propositions**

  **Conjunctions:** **"X is A AND Y is B"**

  **Disjunctions:** **"X is A OR Y is B"**

  **Implications(Rules):** **"IF X is A THEN Y is B"**

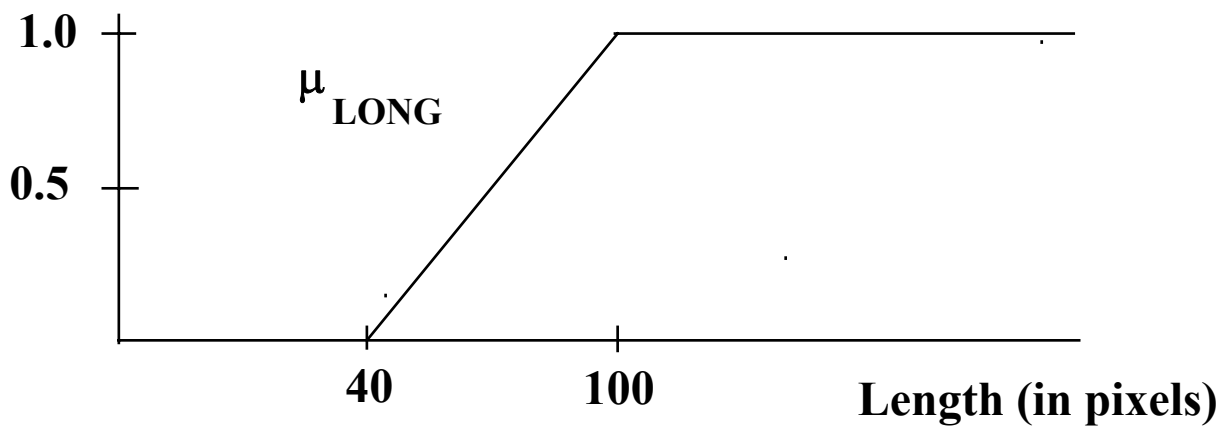
- **Plus, current knowledge of system parameters and choices**

# Knowledge Base

- **Propositions are modeled by Possibility Distributions**

**"Region is LONG" induces a fuzzy variable -LENGTH whose value is LONG**

**LONG is then described by a fuzzy set over an appropriate domain of discourse**



# Knowledge Base
# Compound Propositions

**X is NOT A : Replace membership function with that of $A^c$**

$$\mu_X(u) = \mu_{A^c}(u) \quad \{= 1 - \mu_A(u)\}$$

**Conjunctions and Disjunctions (Induces fuzzy set on pair (X,Y)):**

**Use your favorite Intersection and Union Operator**

**X is A AND Y is B:** $\mu_{(X,Y)}(u,v) = \mu_A(u) \wedge \mu_B(v)$

**X is A OR Y is B:** $\mu_{(X,Y)}(u,v) = \mu_A(u) \vee \mu_B(v)$

# Knowledge Base
# Implication

**Implication is more difficult: IF X is A THEN Y is B**

**From Classical Logic: "A implies B" equivalent to "Not A OR B"**

**Direct Translation:**

$$\mu_{(X,Y)}(u,v) = \mu_{A^c}(u) \vee \mu_B(v)$$

$$= (1 - \mu_A(u)) \vee \mu_B(v)$$

$$= (1 - \mu_A(u) + \mu_B(v)) \wedge 1$$

**compatible with Lukasiewicz multivalued logic**

# Knowledge Base Implication

**<u>There are many other translation schemes</u>**

**Most Common for Control:**

**Correlation Min:** $\mu_{(X,Y)}(u,v) = \mu_A(u) \wedge \mu_B(v)$

**Correlation Product:** $\mu_{(X,Y)}(u,v) = \mu_A(u) \times \mu_B(v)$

**Why? Because they are Simple and Fast!**

# Fuzzy Logic Inference

- **So, Now You've Got**

  **P1: IF X is A Then Y is B**

  **P2: X is A′**

- **Want to Derive Value for Y: (Y is B′)**

  **Generalized Modus Ponens**

  **Actually, Many Ways to Do This**

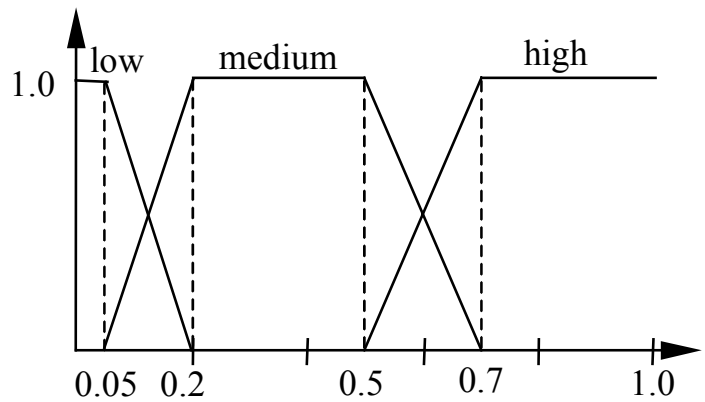
- **Interesting to Note that if all Domains are finite, this looks like Matrix-Vector Multiplication with Sum replaced by Union and Product Replaced by Intersection**

# Ingredients of a Fuzzy Logic System

- **Linguistic Variables: Angle, Δangle, GPR Confidence, Aspect Ratio, etc.**

  **- Problem Dependent**

- **Membership Functions to Model the Meaning of Variable Values**




**Typical Membership Functions**

# More Ingredients

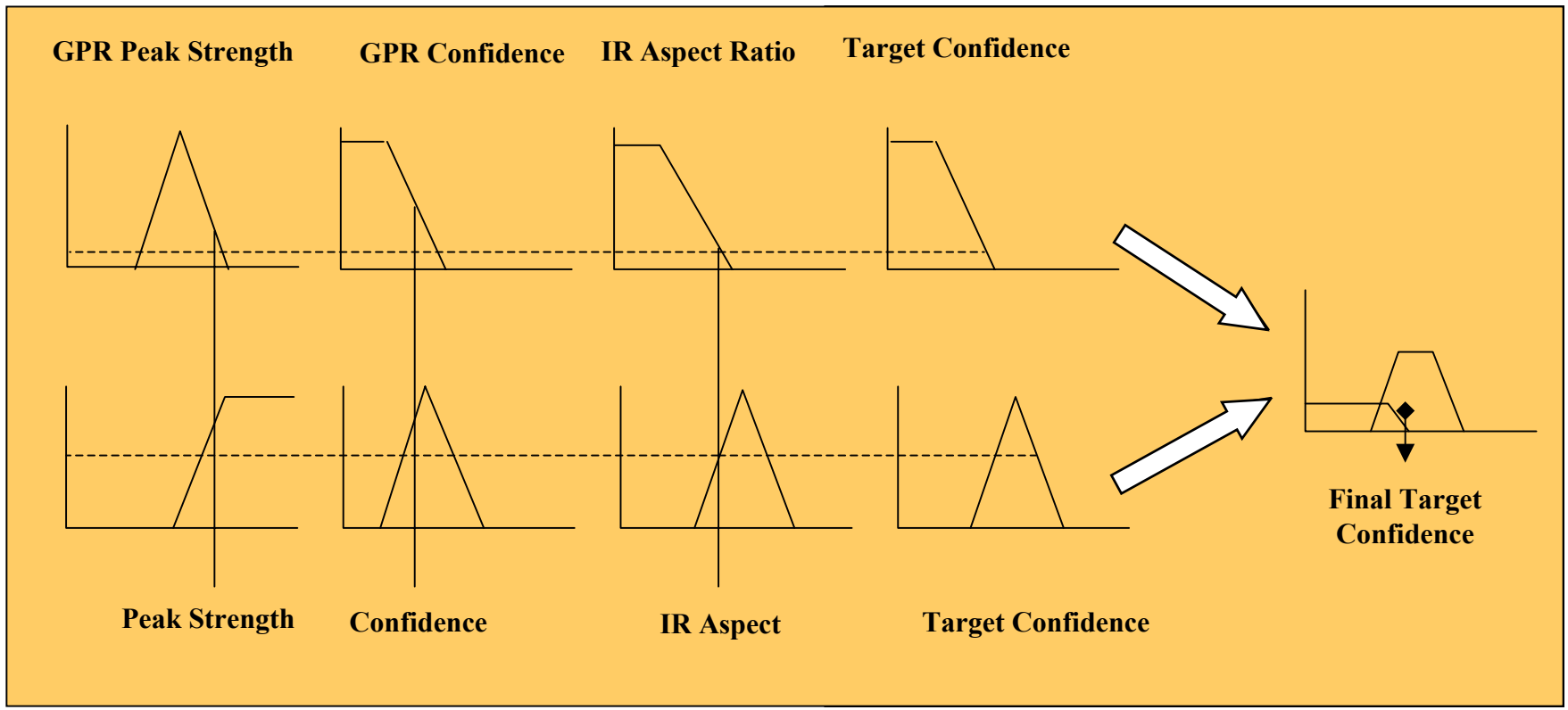
- **Add a cup full of rules (the more the merrier)**

**IF** GPR peak strength is *MEDIUM* AND
GPR confidence is *LOW* AND
ASPECT RATIO of the IR is *NOT_MINE_LIKE*
**THEN**
Overall_Target_Confidence is *LOW*

**IF** GPR peak strength is *HIGH* AND
GPR confidence is *MEDIUM* AND
ASPECT RATIO of the IR is *MEDIUM_MINE_LIKE*
**THEN**
Overall_Target_Confidence is *RATHER HIGH*

# How's Inference Really Done?

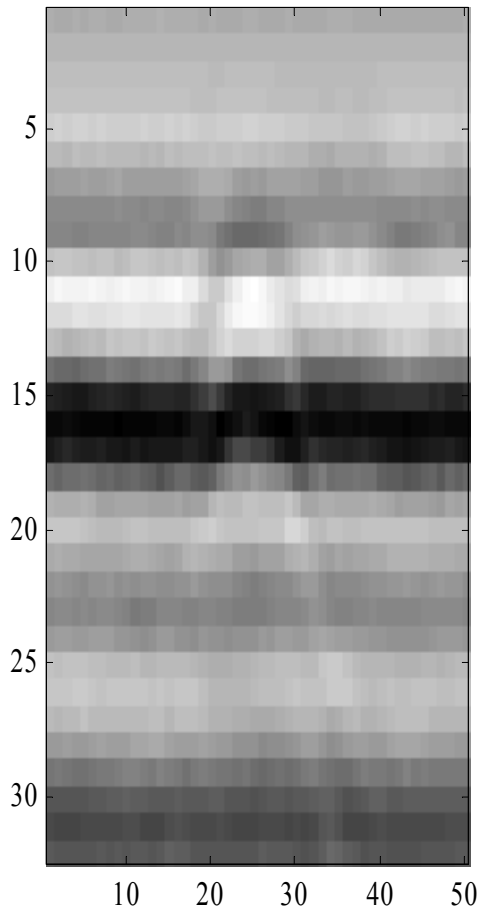
- **All Rules Fire to Some Degree - Produce Output Fuzzy Sets**
- **Output Fuzzy Sets Are "Aggregated" - Usually Added; Sometimes Maxed**
- **Output Value is Computed From All Outputs - Defuzzification (UGH!!!!)**



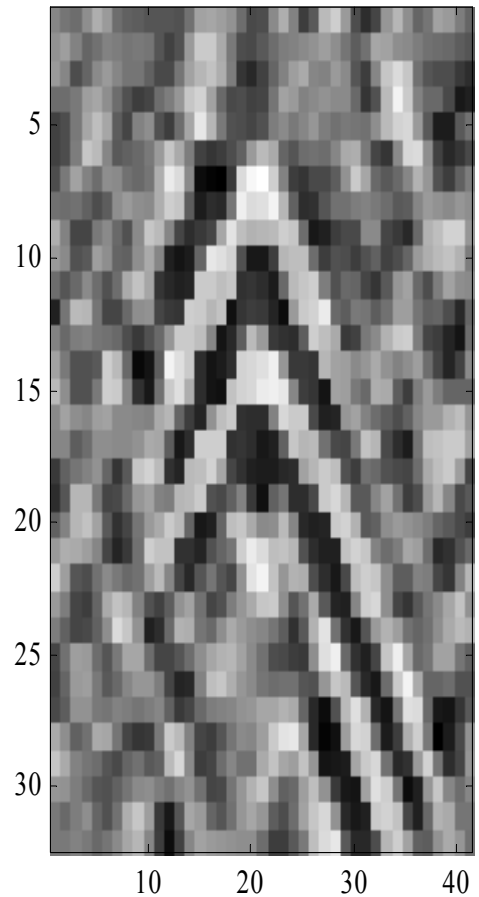
# Example - Describing a GPR Signature

## Plastic Mine

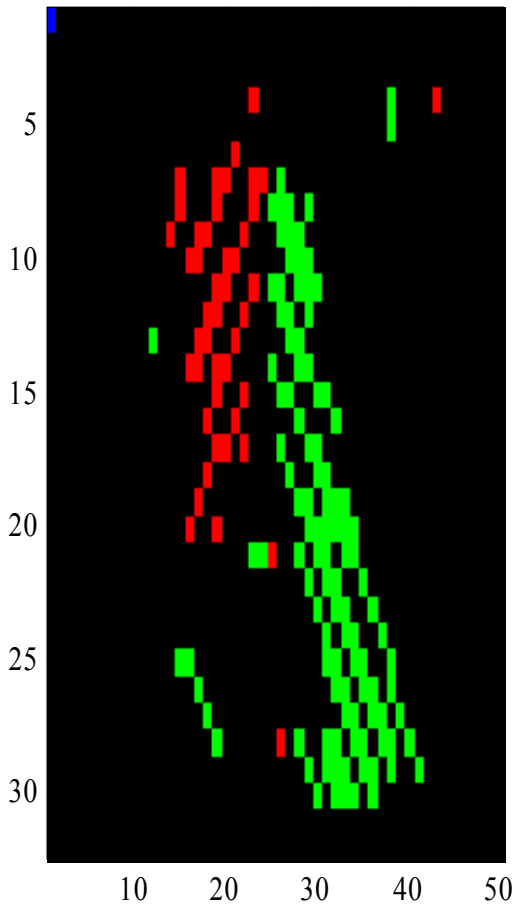
RAW GPR




ENERGY




LEADING & TRAILING EDGES

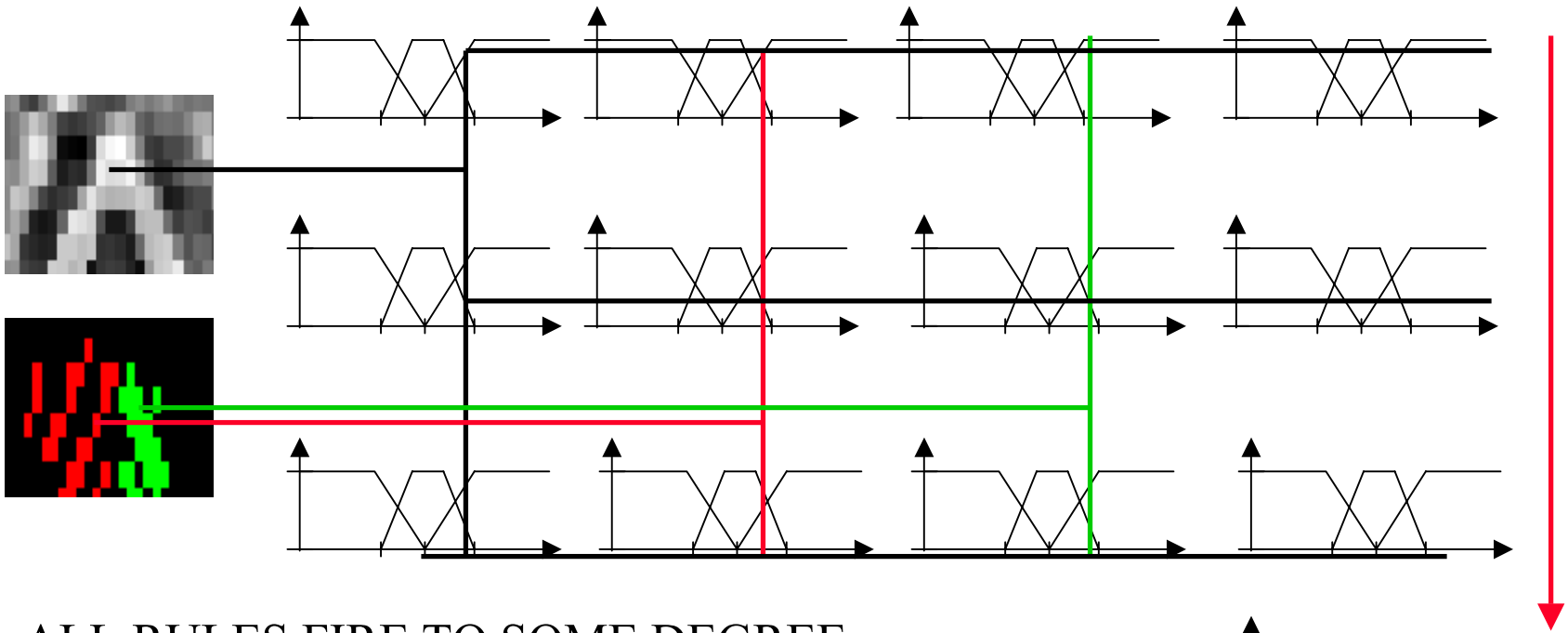


# Example - Describing a GPR Signature

| | | | | |
|---|---|---|---|---|
| **R1:** | **If** | **Energy is** | **HIGH** | **AND** |
| | | **Leading Edge is** | **STRONG** | **AND** |
| | | **Trailing Edge is** | **STRONG** | |
| | **THEN** | | | |
| | | **Mine Confidence is** | **HIGH** | |
| **R2:** | **If** | **Energy is** | **HIGH** | **AND** |
| | | **Leading Edge is** | **MEDIUM** | **AND** |
| | | **Trailing Edge is** | **STRONG** | |
| | **THEN** | | | |
| | | **Mine Confidence is** | **MEDIUM** | |
| **R3:** | **If** | **Energy is** | **HIGH** | **AND** |
| | | **Leading Edge is** | **WEAK** | **AND** |
| | | **Trailing Edge is** | **WEAK** | |
| | **THEN** | | | |
| | | **Mine Confidence is** | **LOW** | |

# Example - Describing a GPR Signature (Mine)

ENERGY LEADING EDGE TRAILING EDGE

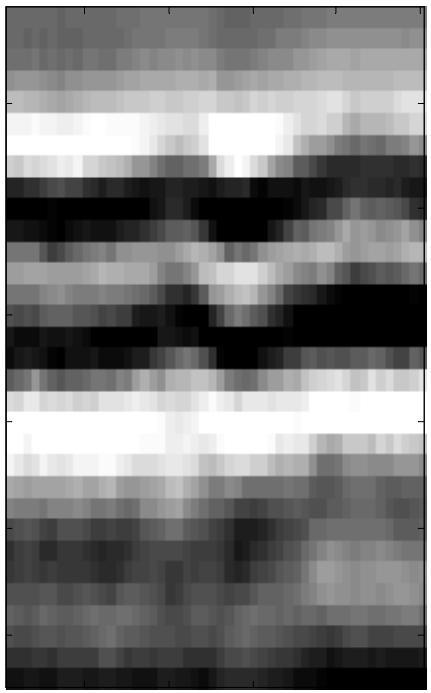


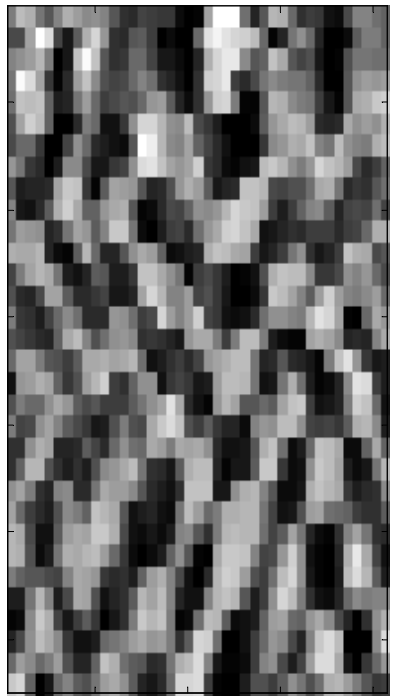
ALL RULES FIRE TO SOME DEGREE

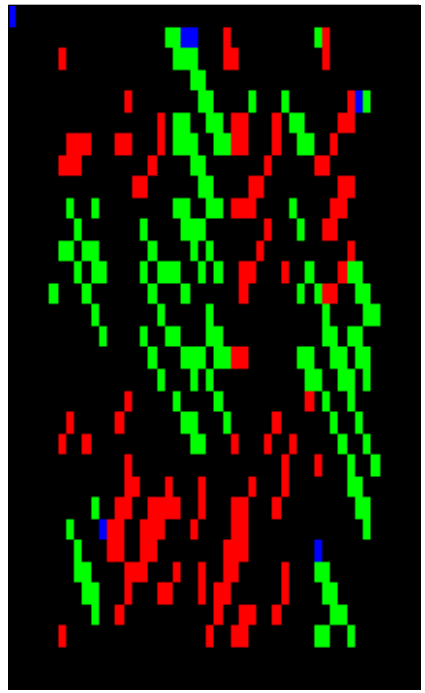
CAN BE IMPLEMENTED QUICKLY WITH LUT

# Example - Describing a GPR Signature

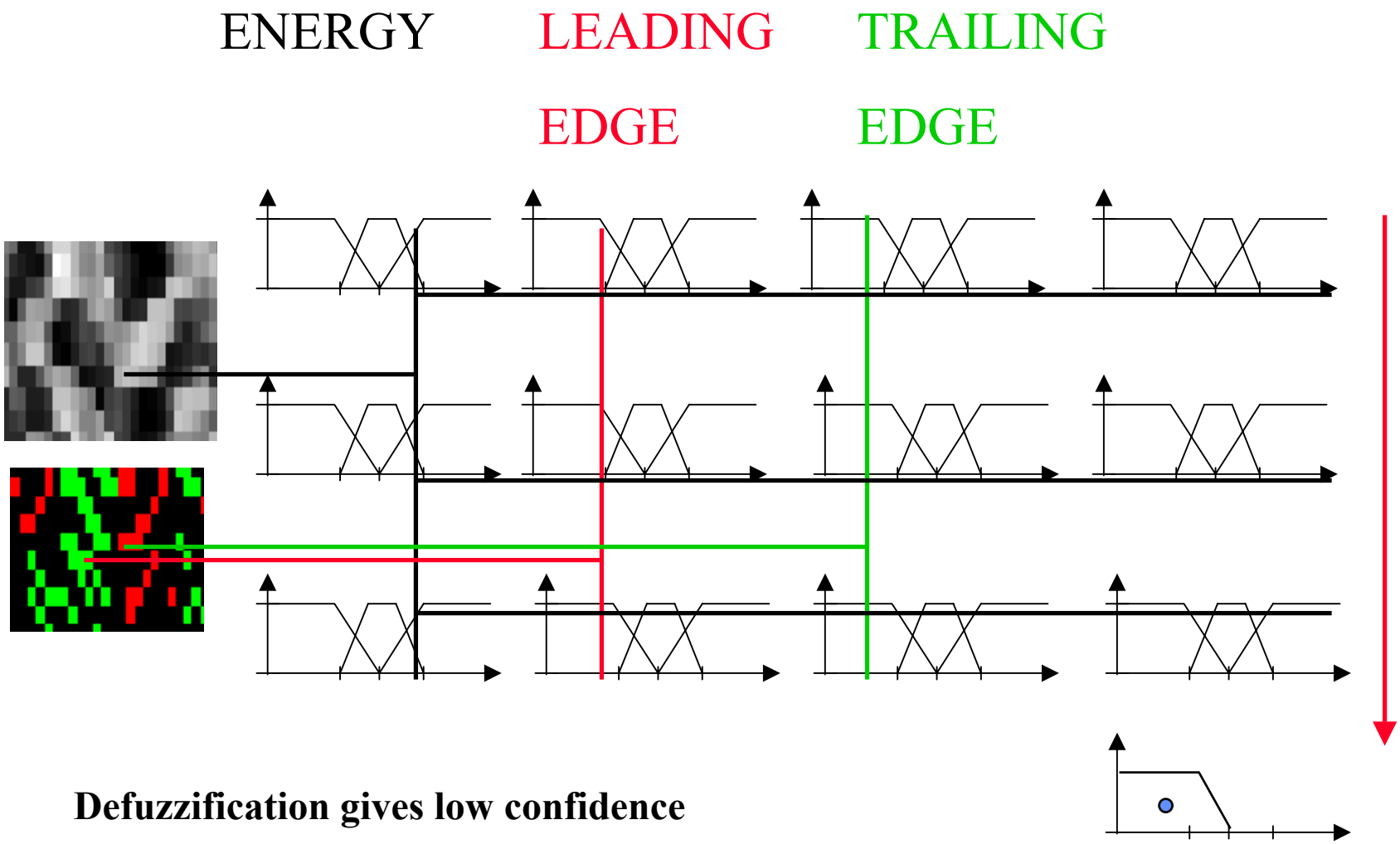
## Clutter







# Example - Describing a GPR Signature (Clutter)




**Defuzzification gives low confidence**

# Fuzzy Rule Bases for Image Processing

**Rules use Pixel Neighborhood Characteristics to "select" strength of different enhancement/smoothing operators in combination**

**Original (with Noise)**

**Median Filter**

**Fuzzy Rule Base**

**Which one do you like best?**

Lena

Noisy

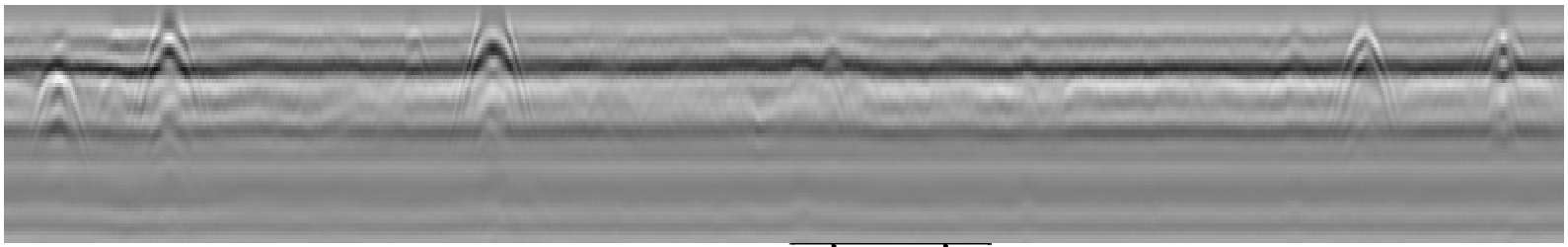
Median

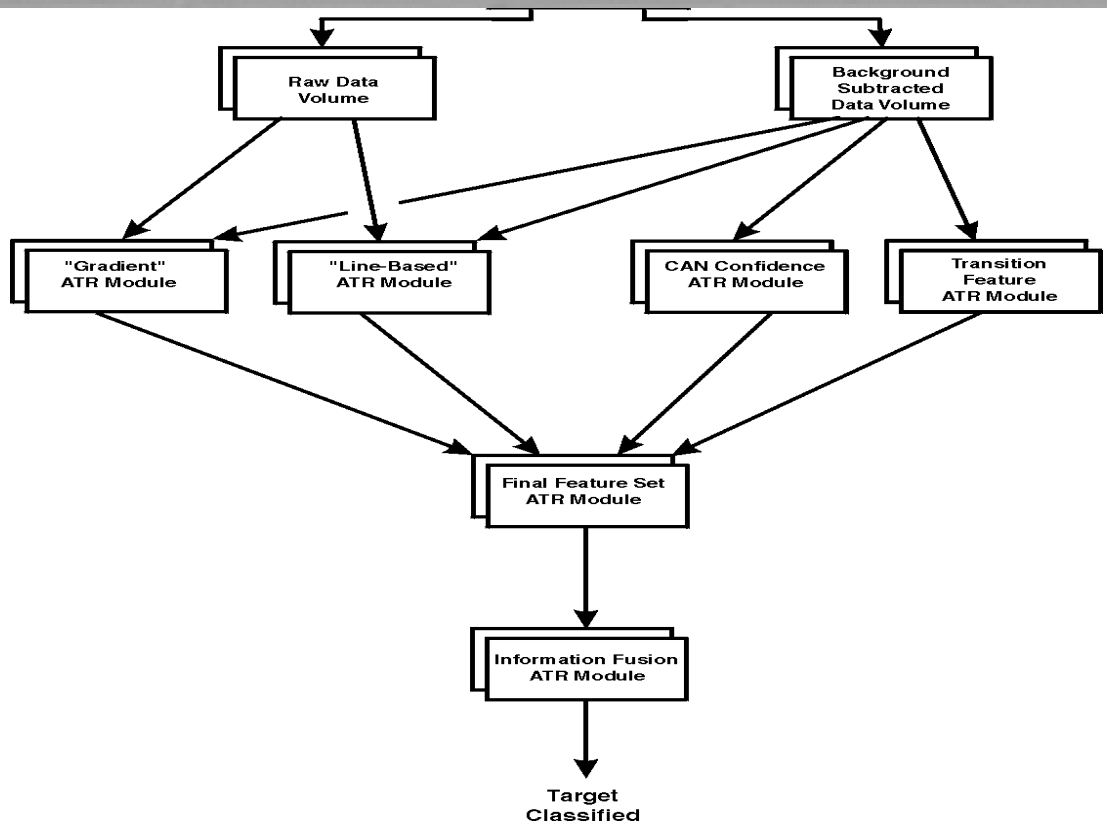
Saint-Marc

Fuzzy Rules

# Information Fusion

Fusing Outputs of Multiple Algorithms on GPR
Bruce Nelson of Geo-Centers Made the Rule Base





# Multiple Algorithm Fusion Results for Blind Tests

- **Fuzzy Integrals and Fuzzy Logic Rule Base**
  - **Developed on Signature Library**
  - **Fused confidence outputs from several sources**
  - **(3 primary detectors and 2 secondary algorithms)**
- **Taken to the field with Geo-Centers Mobile Landmine Detection System**
- **Target Reports submitted to US Army**

# Information Fusion - Rule Base

## Joint with Bruce Nelson - GeoCenters

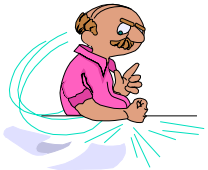
| | GRAD ATR | LINE ATR | CAN CONF | TRANS CONF | BLOB LENGTH | SYSTEM CONF |
|---|---|---|---|---|---|---|
| 1 | L | L | L | L | | L |
| 2 | L | L | L | | | L |
| 3 | L | L | | L | | L |
| 4 | L | | L | L | | L |
| 5 | | L | L | L | | L |
| 6 | L | H | M | M | | ML |
| 7 | M | H | M | L | | ML |
| 8 | | | L | M | | ML |
| 9 | | | L | H | | ML |
| 10 | | | L | L | L | ML |
| 11 | | L | | L | L | ML |
| 12 | L | | | L | L | ML |
| 13 | | L | L | | L | ML |
| 14 | L | | L | | L | ML |
| 15 | L | L | | | L | ML |
| 16 | | | M | H | | M |
| 17 | | | M | M | | M |
| 18 | | | L | L | | ML |
| 19 | M | M | H | L | | M |
| 20 | M | H | M | L | | M |
| 21 | H | M | M | L | | M |
| 22 | | | M | L | | MH |
| 23 | H | H | H | | | MH |
| 24 | H | H | | H | | MH |
| 25 | H | | H | H | | MH |
| 26 | | H | H | H | | MH |
| 27 | | | H | L | | MH |
| 28 | | | H | M | | MH |
| 29 | | H | H | L | | MH |
| 30 | H | | H | L | | MH |
| 31 | H | H | | L | | H |
| 32 | H | H | H | H | | H |
| 33 | M | H | L | M | | ML |
| | | | | | | |

# Blind Test Results at U.S. Army Site

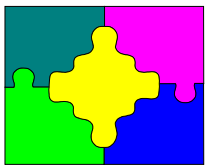
## Blind Test Results (Scored by U.S. Army)

| Soil Type | Method | Fuzzy Rules | | Choquet Integal | | Sugeno Integral | |
|---|---|---|---|---|---|---|---|
| | | PD | FAR | PD | FAR | PD | FAR |
| DIRT | AVERAGE | 75% | 0.05 | 74% | 0.06 | 71% | 0.06 |
| DIRT | MEDIAN | 73% | 0.05 | 75% | 0.06 | 74% | 0.06 |
| GRAVEL | AVERAGE | 94% | 0.03 | 93% | 0.01 | 93% | 0.02 |
| GRAVEL | MEDIAN | 93% | 0.03 | 91% | 0.01 | 92% | 0.02 |
| OFFRD | AVERAGE | 83% | 0.03 | 77% | 0.02 | 80% | 0.02 |
| OFFRD | MEDIAN | 82% | 0.02 | 77% | 0.02 | 82% | 0.02 |
| ALL | AVERAGE | 83% | 0.04 | 80% | 0.04 | 80% | 0.04 |
| ALL | MEDIAN | 86% | 0.03 | 78% | 0.02 | 82% | 0.02 |

# Conclusions



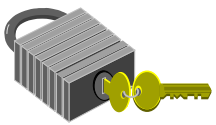
**Fuzzy Set Theory and Fuzzy Logic are serious tools for complex, ill-defined decision making problems like fusion**



**Not competitive, but complementary to traditional methods**

**Many Open Questions**

**Good for research activities**



**Keep an open mind -- Need a Big Bag of Tools**

**Have Fun!**