Probabilistic Approach to Model Extraction from Training Data

Michael D. DeVore
(mdd2@cis.wustl.edu)

Joseph A. O'Sullivan
(jao@ee.wustl.edu)

Sushil Anand
(sushil@ee.wustl.edu)

Natalia A. Schmid
(nar@ee.wustl.edu)

Electronic Systems and Signals Research Laboratory

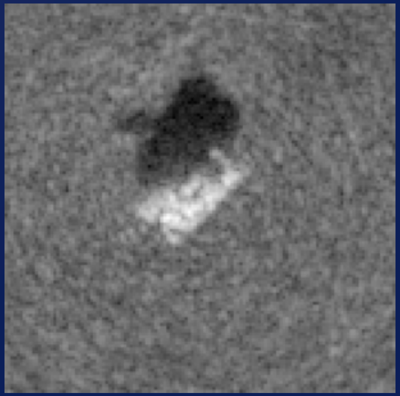
Washington University in St. Louis
School of Engineering and Applied Science

Supported by: ARO Center for Imaging Science DAAH 04-95-10494
ONR MURI N00014-98-1-06-06
Boeing Foundation

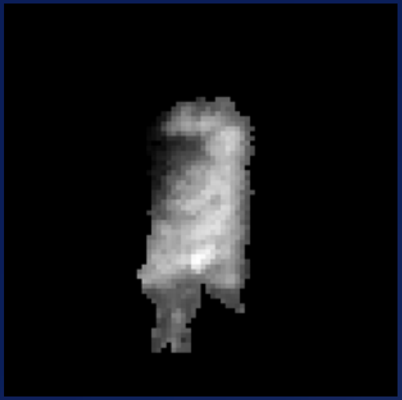
Motivation

Combine sensor output prediction with training data

ATR from CAD Models



Template Based ATR



Model Extraction

Outline

1. Conditionally Gaussian Model for SAR imagery
2. Likelihood Based Approach to Recognition
3. Rejecting Confuser Targets
4. Target Model Estimation & Segmentation
5. Successively-Refinable Sensor Models
6. Example

MSTAR Dataset

Class	Vehicle	Serial No.	Training Set		Testing Set	
			Depression	Images	Depression	Images
BMP-2	#1	9563	17°	233	15°	195
	#2	9566		231		196
	#3	c21		233		196
BRDM-2	#1	E-71	17°	298	15°	263
BTR-70	#1	c71	17°	233	15°	196
T-72	#1	132	17°	232	15°	196
	#2	812		231		195
	#3	s7		228		191

Model Training and Recognition Assessment

Class	Vehicle	Serial No.	Depression	Images
2S1	#1	b01	15°	274
BTR-60	#1	k10yt7532	15°	195
D7	#1	92v13015	15°	274
T-62	#1	A51	15°	273

Confuser Rejection Assessment

Conditionally Gaussian Model

Model each pixel *i* as independent, zero mean, complex conditionally Gaussian

$$p_{\boldsymbol{R}|\Theta,A,C^2}\left(\boldsymbol{r}\middle|\theta,a,c^2\right)=\prod_i \frac{1}{\pi c^2 \sigma_i^2(\theta,a)} e^{-\frac{|r_i|^2}{c^2\sigma_i^2(\theta,a)}}$$

Where: $\sigma_i^2(\theta,a)$ = variance function over pose and class
c^2 = constant over all pixels to account for power fluctuation

To classify an image of an unknown target, we seek the maximum over a, θ, and c^2

Likelihood Based Recognition

Presume a null-hypothesis H_0: $r_i \sim \mathrm{CN}(0,\xi^2)$ which models the scenario where no target is present in the image*

Accommodate image segmentation $I(a,\theta)$ by computing likelihood relative to a null-hypothesis

$$\left[\hat{a},\hat{\theta}\right] = \underset{[a,\theta]}{\operatorname{argmax}} \prod_{i\in I(a,\theta)} \left\{ \frac{1}{\pi \hat{c}^2 \sigma_i^2(\theta,a)} \exp\left(-\frac{|r_i|^2}{\hat{c}^2 \sigma_i^2(\theta,a)} \right) \Big/ \frac{1}{\pi \xi^2} \exp\left(-\frac{|r_i|^2}{\xi^2} \right) \right\}$$

where ξ^2 is a small constant variance, and

$$\hat{c}^2 = \frac{1}{|I(\theta,a)|} \sum_{i\in I(\theta,a)} \frac{|r_i|^2}{\sigma_i^2(\theta,a)}$$

*Schmid & O'Sullivan "Thresholding Method for Reduction of Dimensionality"

Confuser Rejection

Relative entropy to reject target classes not in the database

$$D(p\|q) = \int p(x)\log\left(\frac{p(x)}{q(x)}\right)dx$$

- **One-sample estimate of variance from observation: $|r_i|^2$**
- **Relative entropy quantifies difference from training estimate**
- **Reject observation if the empirical relative entropy is too large**

$$\frac{1}{\left|I\left(\hat{\theta},\hat{a}\right)\right|}\sum_{i\in I\left(\hat{\theta},\hat{a}\right)} D\left(p\left(\bullet;|r_i|^2\right)\middle\|\, p\left(\bullet;\hat{c}^2\hat{\sigma}_i^2\left(\hat{\theta},\hat{a}\right)\right)\right) > \gamma$$

Confuser rejection rate is a function of γ

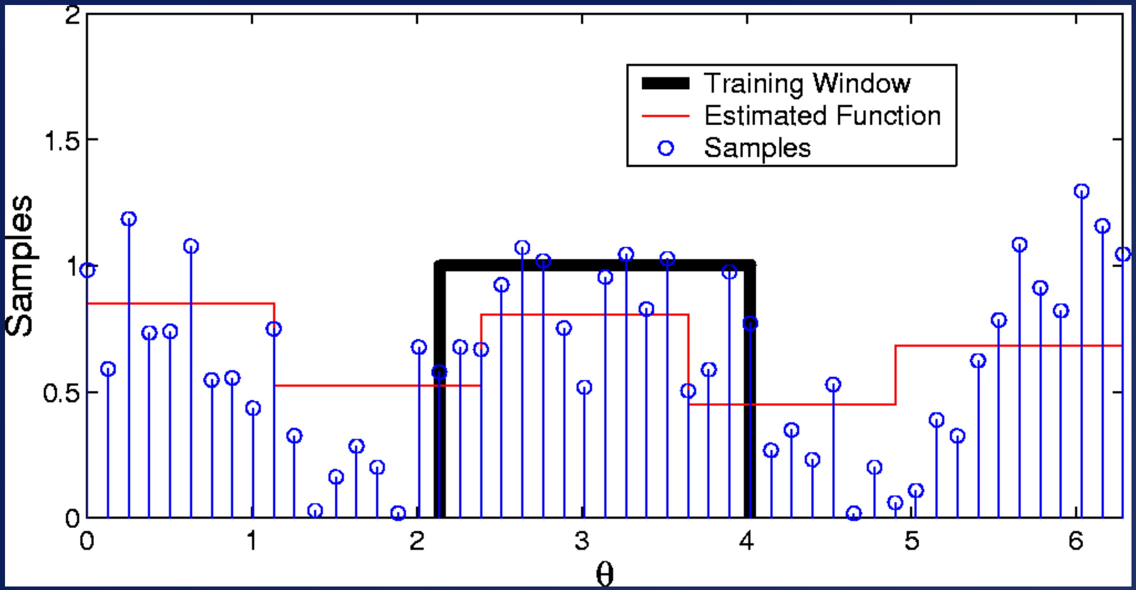
Target Model Estimation

Given N registered training images q_j of a target with pose ϕ_j, estimate variances over N_w windows of width d

$$\hat{\sigma}^2(\theta_k, a) = \frac{1}{n_k} \sum_{\{j:\phi_j \in w_k\}} \left| q_j \right|^2$$

where

$$w_k = \left[\frac{2\pi k}{N_w} - \frac{d}{2}, \frac{2\pi k}{N_w} + \frac{d}{2} \right)$$



Target Model Segmentation

Quantify pixel information relative to the same null-hypothesis used for target recognition

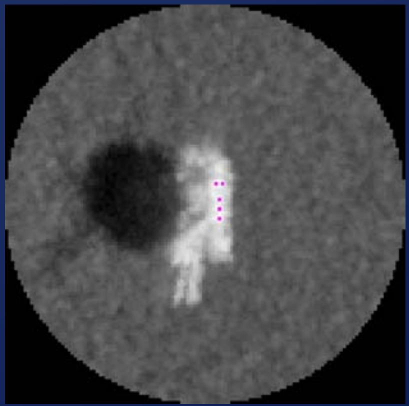
For each target retain pixels *i* that are informative relative to the null-hypothesis:

$$S_a = \left\{ i : \frac{1}{N_w} \sum_k D\left(p\left(\bullet; \hat{\sigma}_i^2(\theta_k, a)\right) \middle\| p\left(\bullet; \xi^2\right) \right) \geq \tau \right\}$$

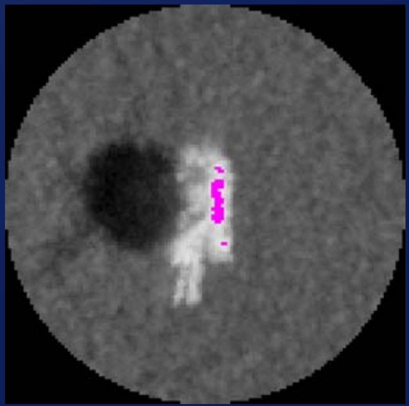
Segmentation of target models, not of images

Target Model Segmentation

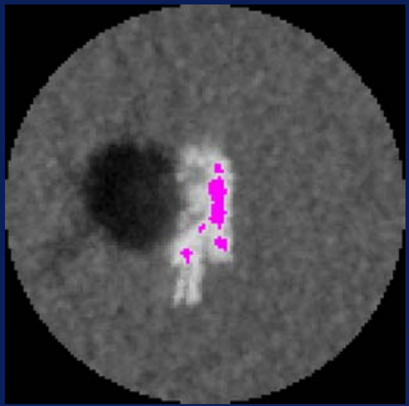
Yields an ordering of pixels by their empirical information relative to null-hypothesis.



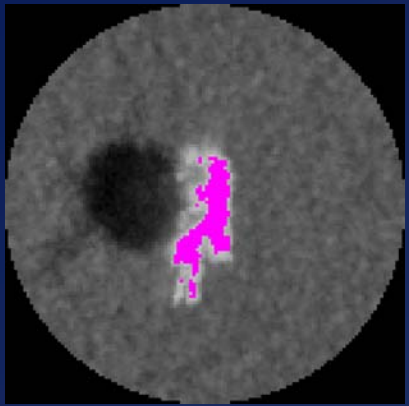
Top 5



Top 50



Top 100



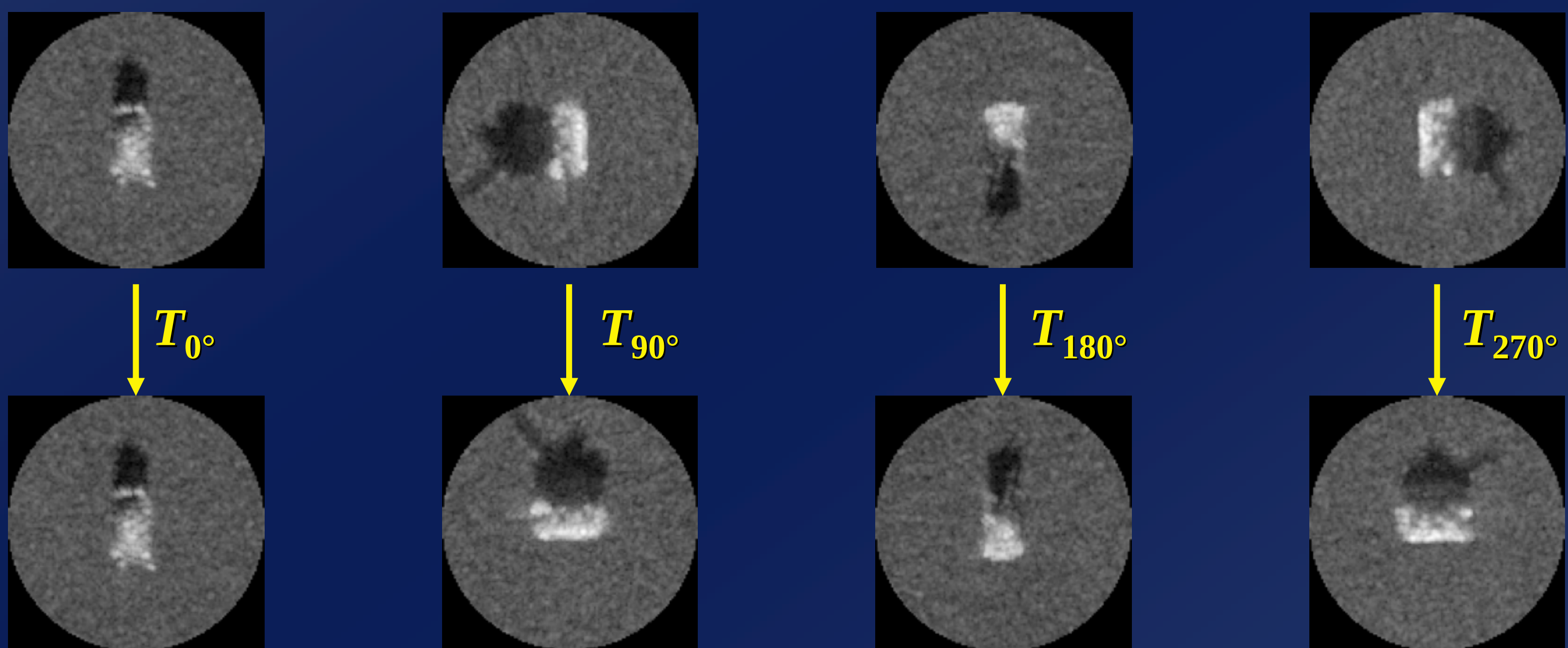
Top 300

For null-hypothesis ξ^2=0.0028 - approximate background variance - pixels on illuminated side of target are deemed most informative.

Successively-Refinable Sensor Models

Variance estimate for an unregistered image r with pose θ formed by appropriately transforming the estimate from the closest w_k

Registered Variance Images



Transformed Estimates

Image segmentation from target model, $I(\theta,a) = T_\theta S_a$

Successively-Refinable Sensor Models

- Represent model parameter functions, i.e. $\sigma_i^2(\theta,a)$, as a collection of hierarchical approximations
- Divide azimuth into N_d non-overlapping intervals of width d

$$\tilde{\sigma}_{d,i}^2(\theta_k,a)=\frac{1}{d}\int_{\frac{2\pi k}{N_d}-\frac{d}{2}}^{\frac{2\pi k}{N_d}+\frac{d}{2}}\sigma_i^2(\theta,a)\,d\theta$$

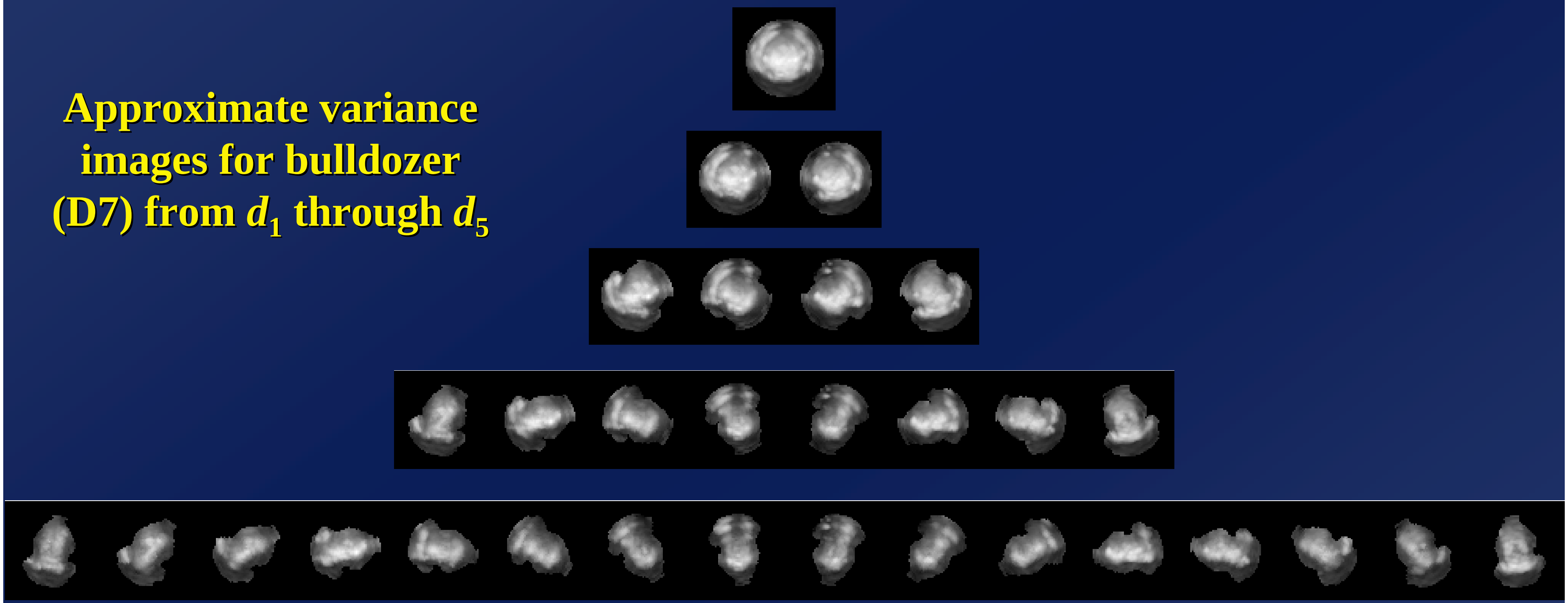
Approximations d and $d/2$ are hierarchically related:

$$\tilde{\sigma}_{d,i}^2(\theta_k,a)=\frac{1}{2}\left[\tilde{\sigma}_{\frac{d}{2},i}^2(\theta_{2k},a)+\tilde{\sigma}_{\frac{d}{2},i}^2(\theta_{2k+1},a)\right]$$

Successively-Refinable Sensor Models

Consider decreasing interval widths $d_1 = 2\pi$, $d_2 = \pi$, …, $d_m = 2\pi/2^{m-1}$

Approximate variance images for bulldozer (D7) from d_1 through d_5

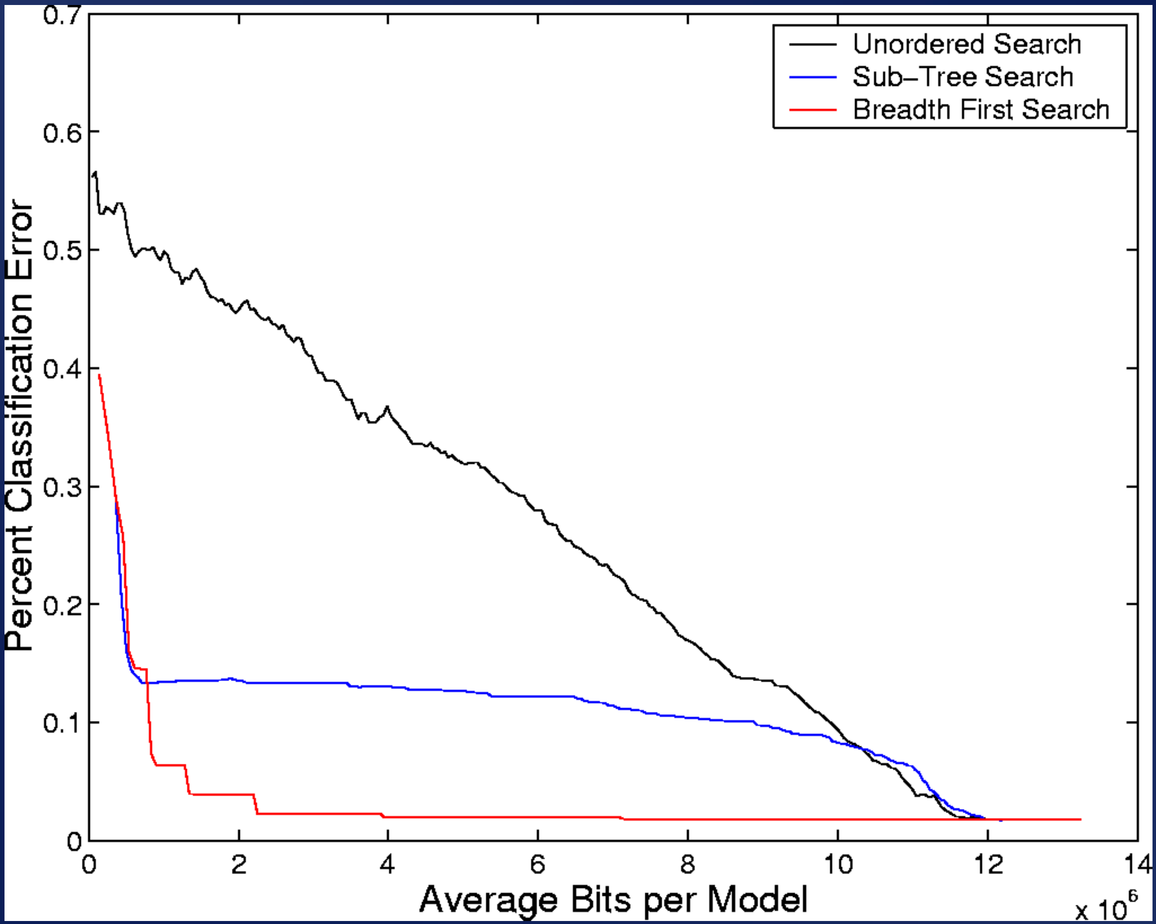


⋮

Search over θ_k in level i ordered by the most likely pose at level i-1

Four-Class Example

Classification error rate as a function of the number of bits transmitted between the database and processor

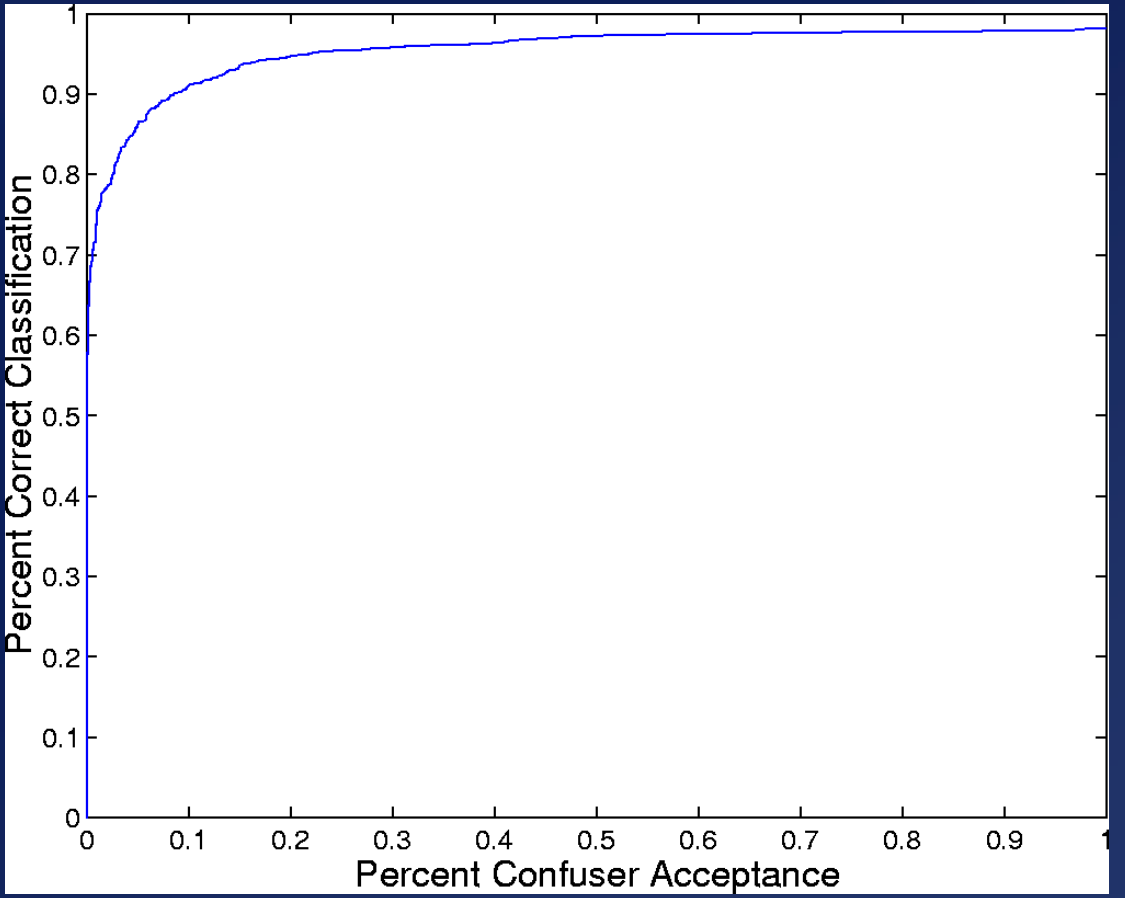


- Classification depends on extent of search
- Eventually, search covers all possibilities
- Breadth-first search quickly finds good combinations of (θ,a)
- Small overhead present with ordered searches

Confuser Rejection Example

Confuser rejection rate a function of threshold γ and approximation d

- Rejection characteristics at approximation d_9
- Error if target misclassified or falsely rejected
- For $\gamma = 1.09$, $P_{cc} = 91\%$ with 10% confuser acceptance



Conclusions

- Method for modeling target reflectivity statistics from sample images
- Target models used to estimate conditional sensor output statistics
- Image distribution estimation, not image prediction
- Successively-refinable sensor models yield successively-refinable decisions
- Breadth-first search of resolution tree quickly finds good parameter combinations