Theoretical Issues in ATR System Design:

Approximation Error,
Estimation Error,
Model Complexity, and
Computational Complexity

Joseph A. O'Sullivan and Michael D. DeVore

Electronic Systems and Signals Research Laboratory

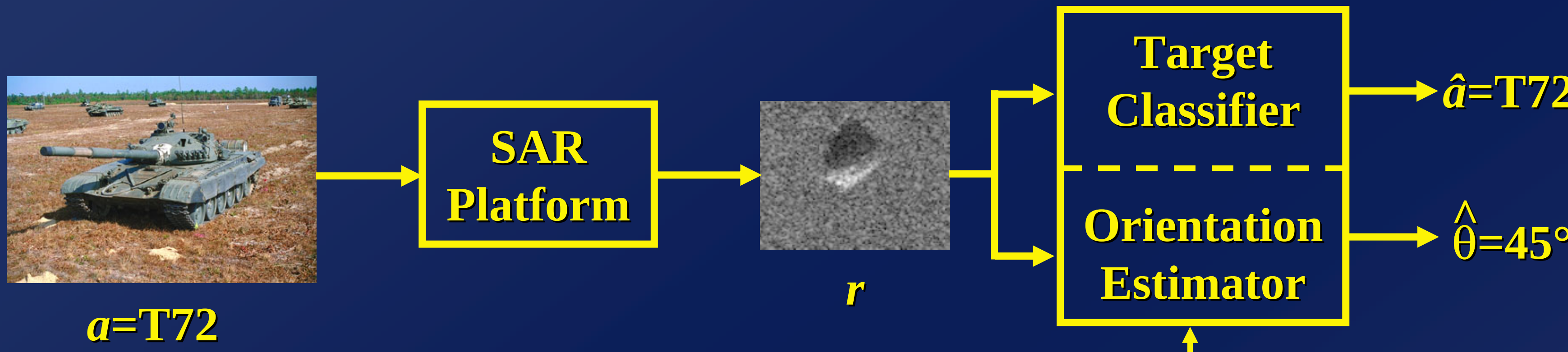
Department of Electrical Engineering

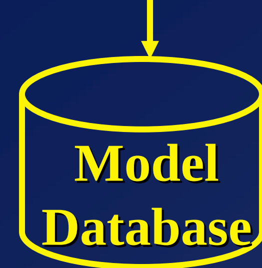
Washington University in St. Louis

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

ATR Theory and Performance

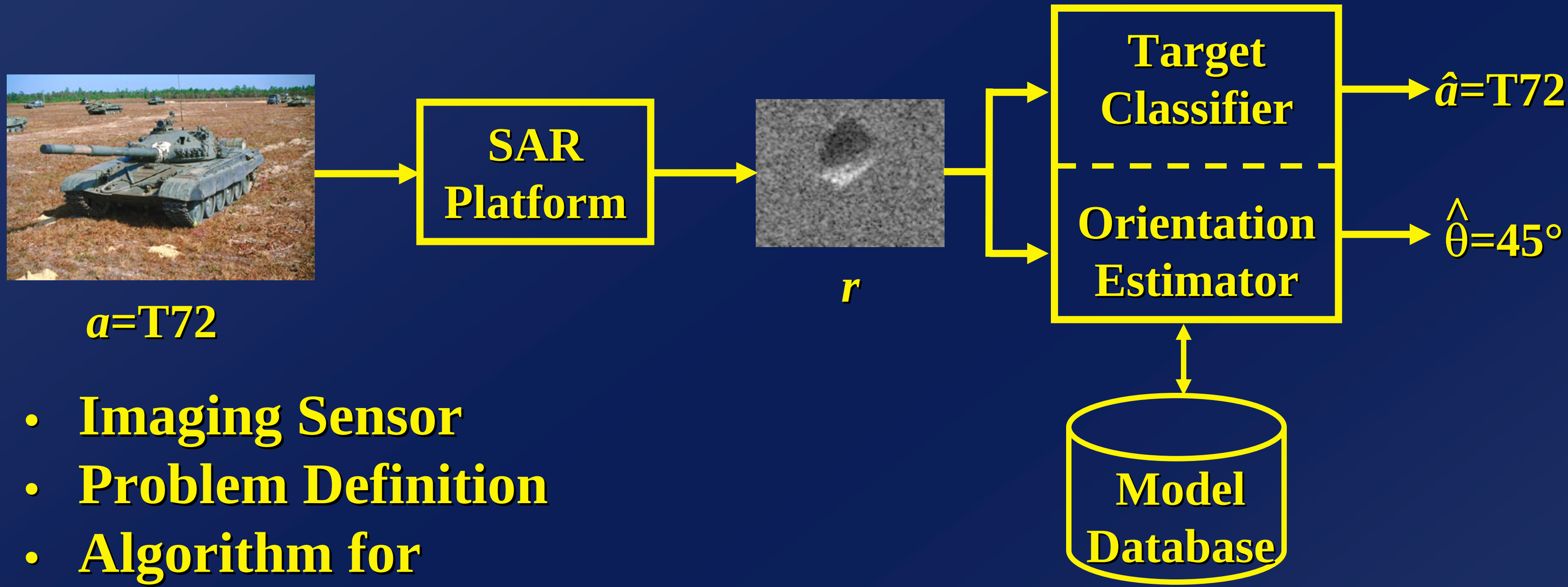




Outline

- ATR Systems of Interest
- Training and Testing Paradigm
  – Approximation Error
  – Estimation Error
- Model Complexity
- Computational Complexity
- Conclusions

ATR Systems of Interest



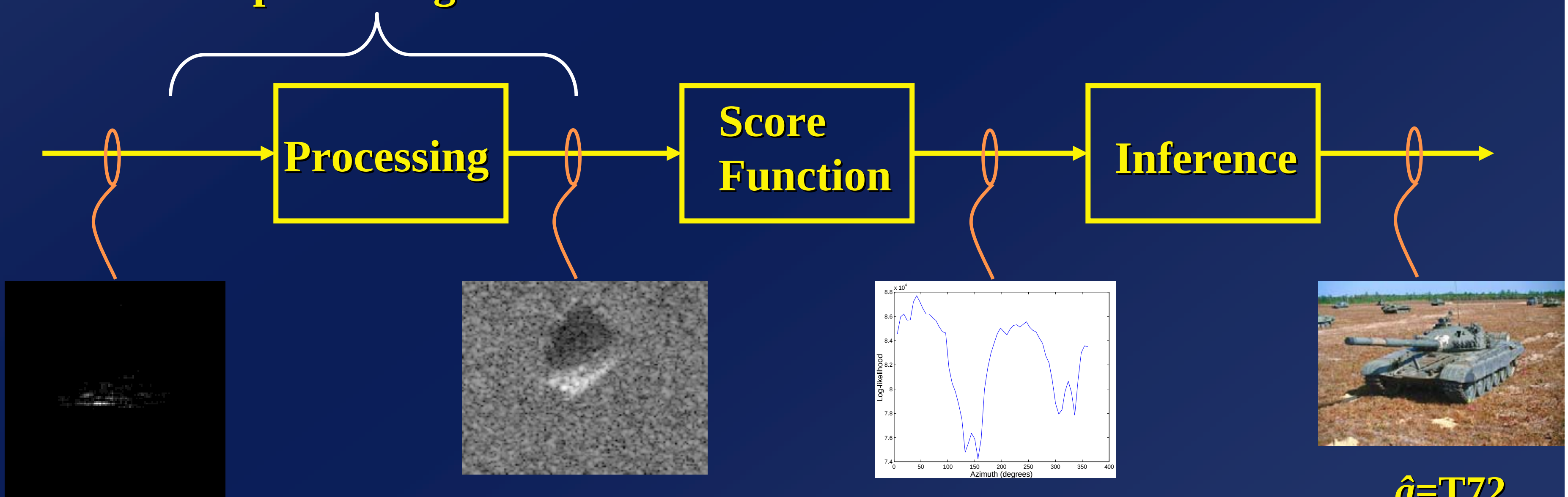
- Imaging Sensor
- Problem Definition
- Algorithm for
  - Classification
  - Parameter estimation
- System Resource Constraints
  - Database size
  - Processor speed
  - Communication speeds
  - Architecture

Parameters:

- Pose
- Velocity
- "Features"

ATR System Design: Sensor Models

- Scene and Imaging Sensor Physics
  – Reflectivity, scattering, emission phenomena
  – Electromagnetic wave propagation
  – Antenna-amplifier or optics-focal plane array
- Sensor Signal Processing
  – Point spread function, intrinsic parameters
  – Noise model, reflectivity model
  – Fixed processing



Raw HRR Data

SAR Image

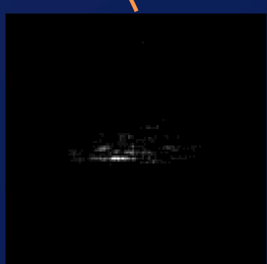
Score function

$\hat{a}$=T72

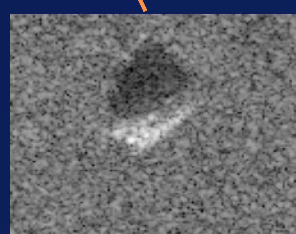
ATR System Design: Inference

- **Likelihood-Based Approach**
  - Derive likelihood function from scene and sensor physics
  - Search likelihood function over target type and pose
  - Potential difficulties for clutter, new scenarios
- **Fixed Processing**
  - Constrain System Impementation
  - Tune parameters
  - Suboptimal for known parameters
- **Feature-Based Approaches**
  - Extract features (e.g. scattering centers
  - Score features (often likelihoods)

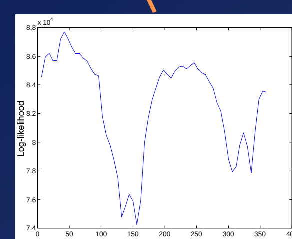
Processing → Score Function → Inference →



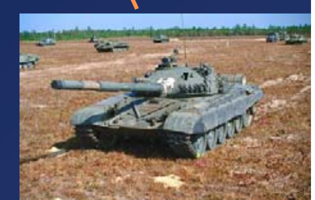
Raw HRR Data



SAR Image



Score function



$\hat{a}$=T72

ATR System Design: Training Paradigm

Training Data

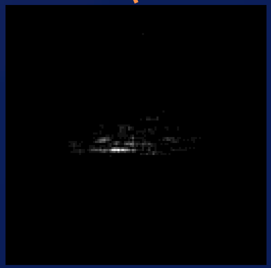
Scene and Sensor Physics

Parameter Extraction

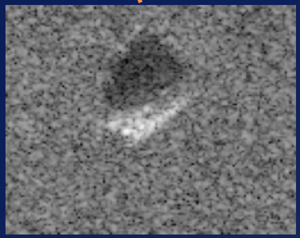
Processing

Score Function

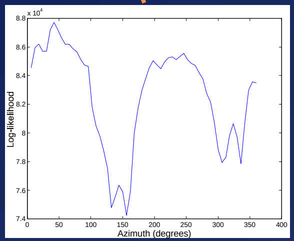
Inference



Raw HRR Data



SAR Image



Score function

$\hat{a}$=T72

ATR System Design: Training Paradigm

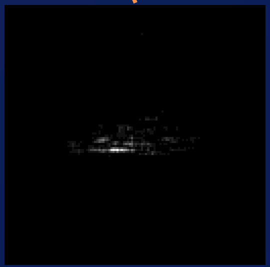
Training Data

Scene and Sensor Physics

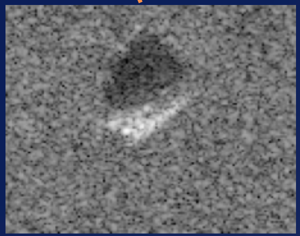
Function Estimation

Processing → $L(r|a,\theta)$ → Inference

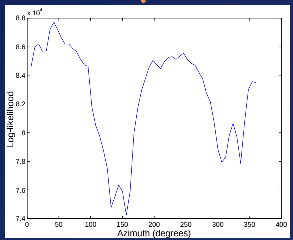
- Likelihood functions parameterized by functions
- Training
  - Function estimation
- Inference
  - Hypothesis testing
  - Parameter estimation



Raw HRR Data



SAR Image



Log-likelihood function

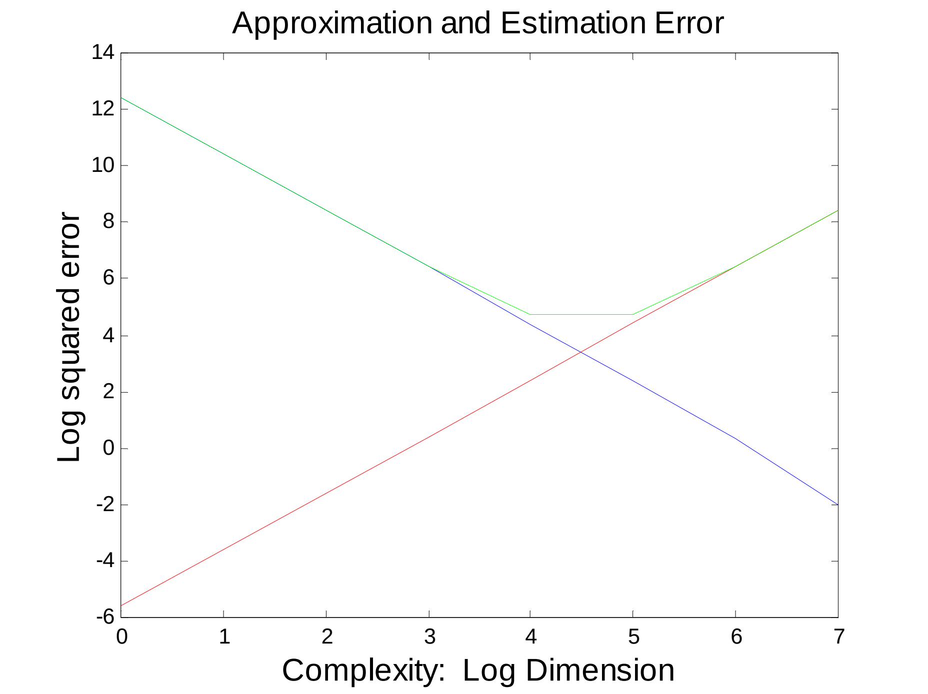
$\hat{a}$=T72

Advantages of Likelihood-Based Approach

- Likelihood Formulation Provides
  - Framework for optimal system design
  - Formalism for
    - Predicting performance
    - Deriving performance bounds
  - Benchmarks for practical system design
  - Fundamental approach to model verification
  - Approach for design for robustness
  - Information, Estimation, Detection, and Statistical Pattern Recognition Theories

Issues in Function Estimation

- Statistical Tradeoffs:
  - Approximation error—Estimation error
  - Bias—Variance
  - Overtraining
- Learning Theory Basis
- Current Information-Theoretic View:
  - Complexity regularization
  - MDL Basis

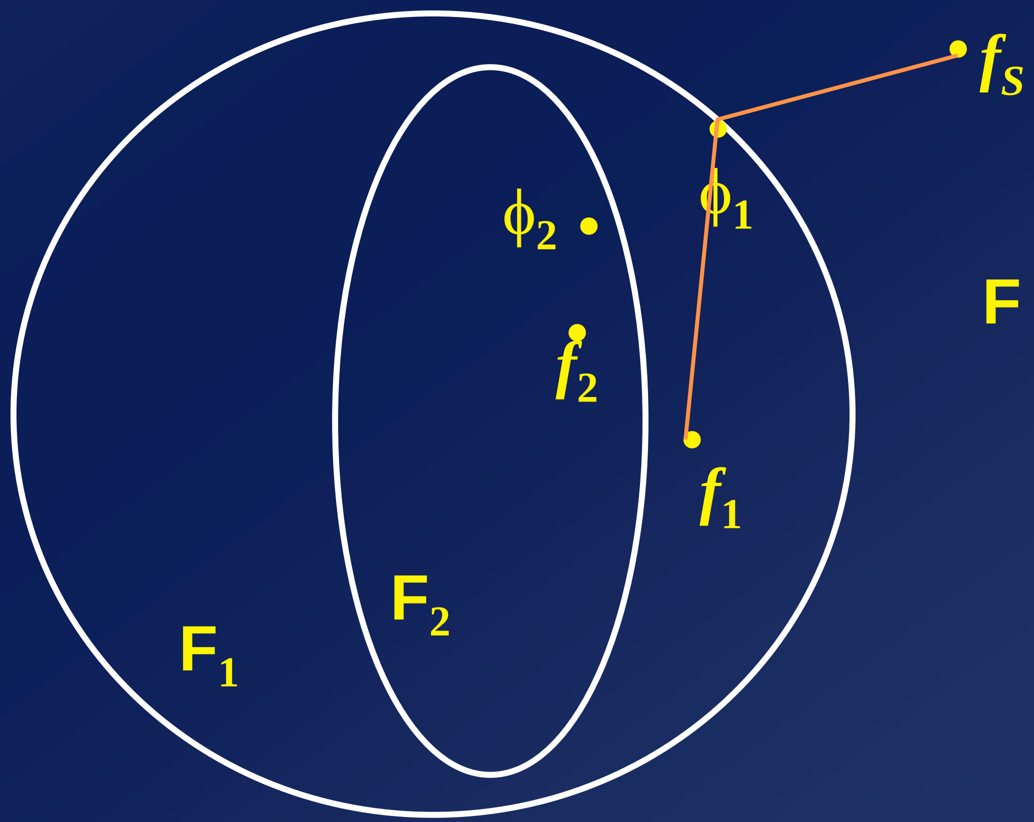


Moulin, Yu, Barron, Rissanen

- LLR(f) + η Complexity(f)

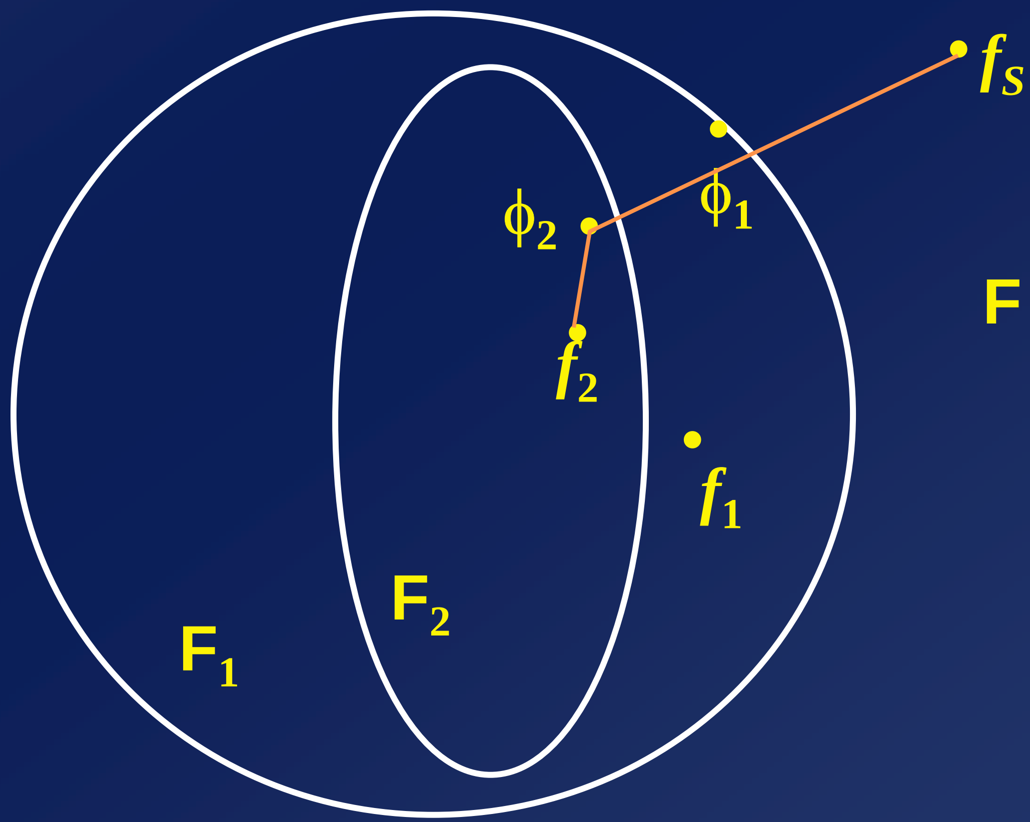
Log integrated squared error
ISE=App. Error + Est. Error
Individual errors exponential in dimension

Regularization for Function Estimation



Tikhonov
Grenander's Sieves
Prior Likelihoods
Constraint Sets
Penalty Functionals
Complexity Regularization

Regularization for Function Estimation



Tikhonov
Grenander's Sieves
Prior Likelihoods
Constraint Sets
Penalty Functionals
Complexity Regularization

Sources of Approximation Error

- Constraints on Implementations
  - Processor, architecture, bandwidth, memory, time
  - Adaptability
- Incorrect target or sensor model assumptions
- Unmodeled or incorrectly modeled clutter
- Finite dimensional parameterizations
- Incorrect model of sensor or target dynamics

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**

Recognition by maximizing the log-likelihood ratio*

$$\left[\hat{a},\hat{\theta},\hat{c}^2\right]=\underset{\left[a,\theta,c^2\right]}{\operatorname{argmax}}\ln\left[\prod_i \frac{p\left(\boldsymbol{r}|a,\theta,c^2\right)}{p\left(\boldsymbol{r}\middle|\xi^2\right)}I_i(a,\theta)\right]$$

Where: ξ^2 **= average clutter variance**
I_i **= mask function**

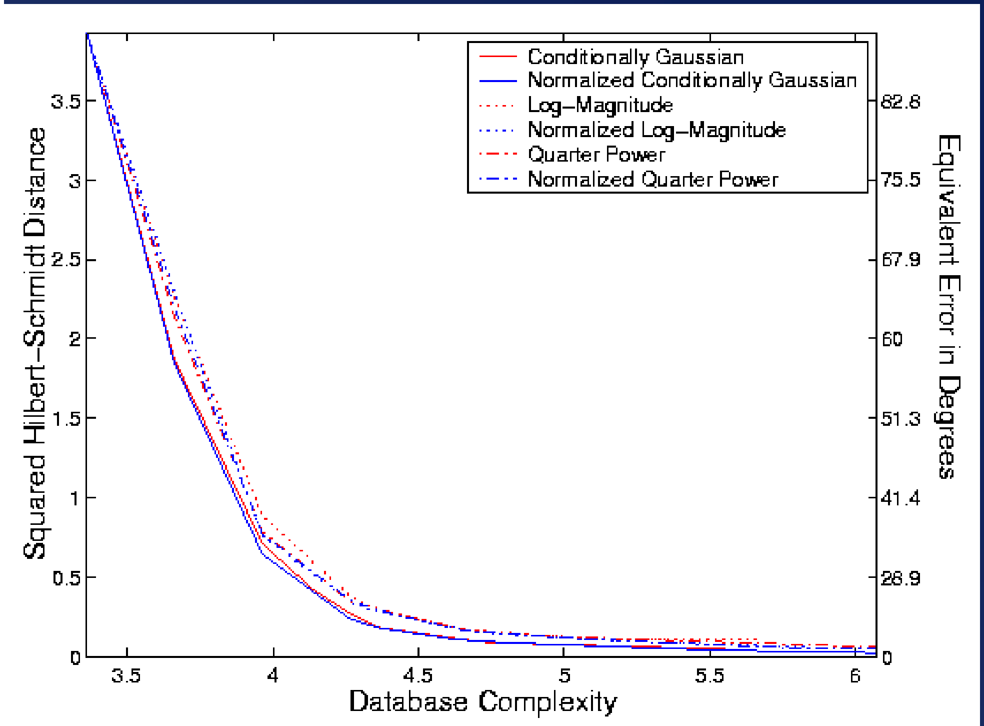
*Schmid & O'Sullivan "Thresholding Method for Reduction of Dimensionality"

ATR Performance and Complexity

Comparison in terms of:

- **Performance achievable at a given complexity**
- **Complexity required to achieve a given performance**

Information Theory Basis: Rate-Distortion Theory → Rate-Recognition Theory



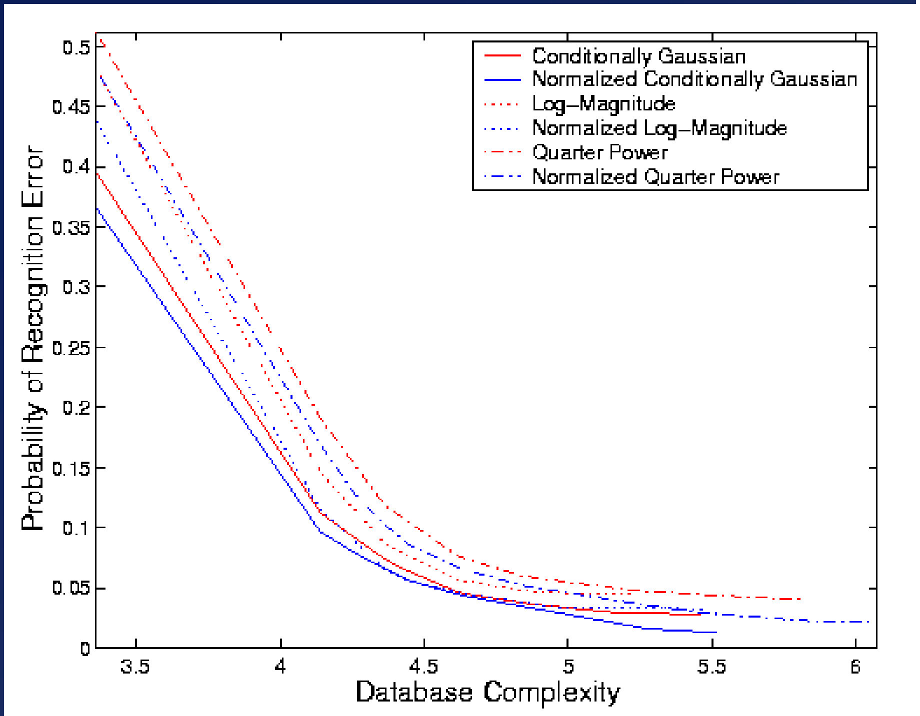


Image Segmentation → Target Extraction Information-Theoretic Approach

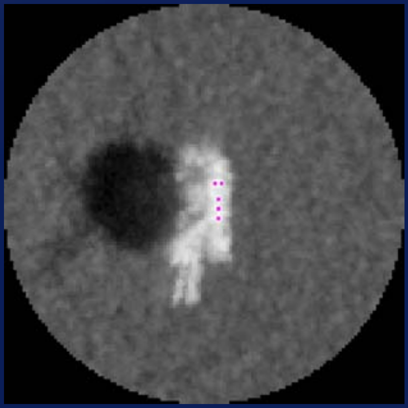
- **Hypothesis Test:**
  - pixels on target vs. on clutter
- **Pixelwise measure of information for discrimination**

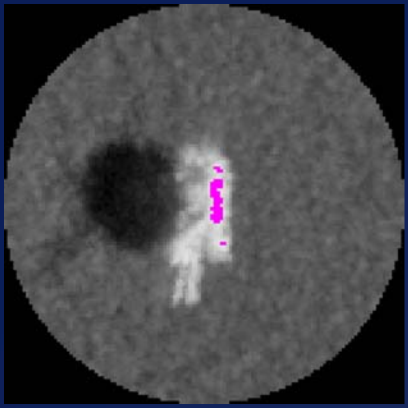
  $D(p_i||p_0)$

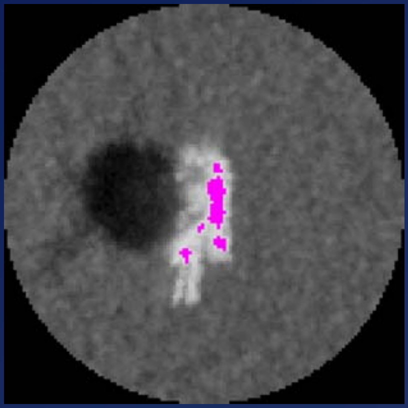
  Conditionally Gaussian

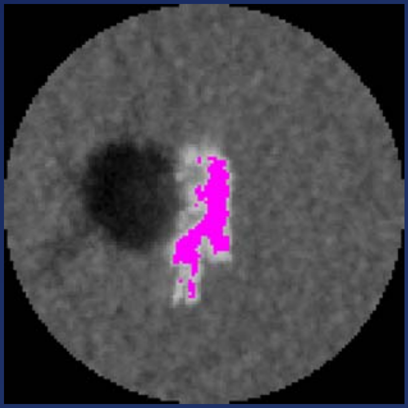
$$\left(\frac{|r_i|^2}{\xi^2}\right) - 1 - \ln\left(\frac{|r_i|^2}{\xi^2}\right)$$

- **Segmentation Complexity**
  - Likelihoods on snakes (contours)
  - Complexity of region


Top 5


Top 50


Top 100


Top 300

System Design Issues: Dynamically Reconfigurable Algorithms

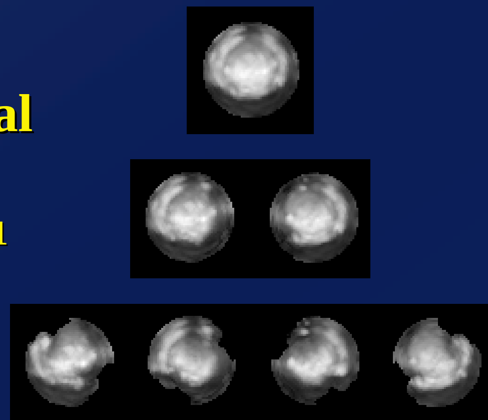
- Information Theory Contributions
  - Successively Refinable Models
    - » Effros, Cover and Equitz, Rimoldi
    - » J. Shapiro, Said and Pearlman
    - » R. DeVore, A. Cohen,
    - » I. Daubechies, D. Donoho, …
  - Successively Refinable Recognition
    - » Rate-distortion → Rate-recognition
    - » Log-time, log-space ←→ Rate

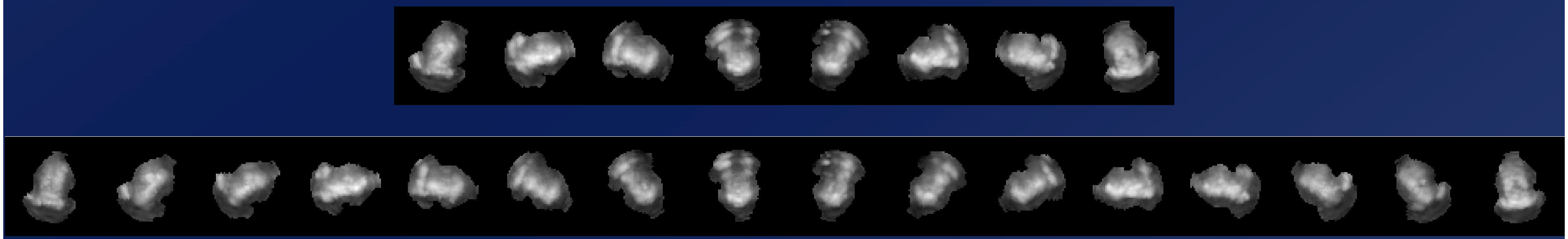
Successively-Refinable Sensor Models

Divide azimuth into N_d non-overlapping intervals of width d

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

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



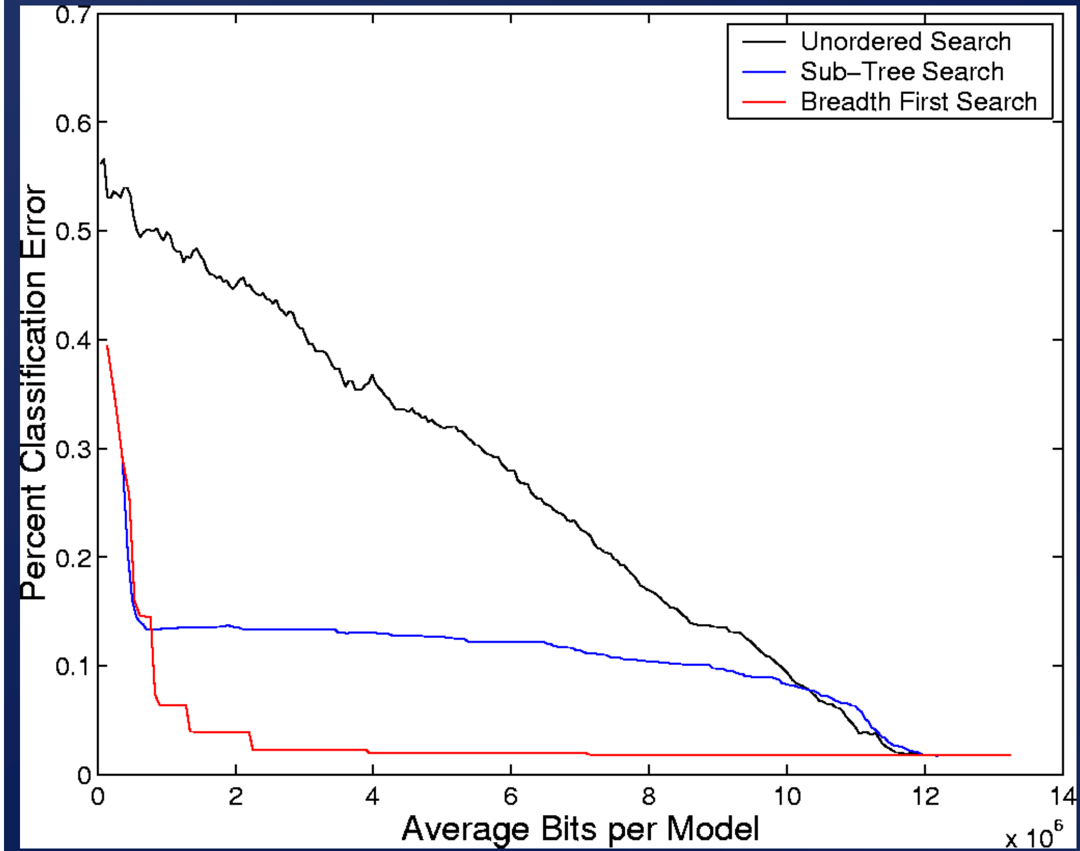


⋮

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

Four-Class Example

Successively-refinable sensor models yield successively-refinable decisions

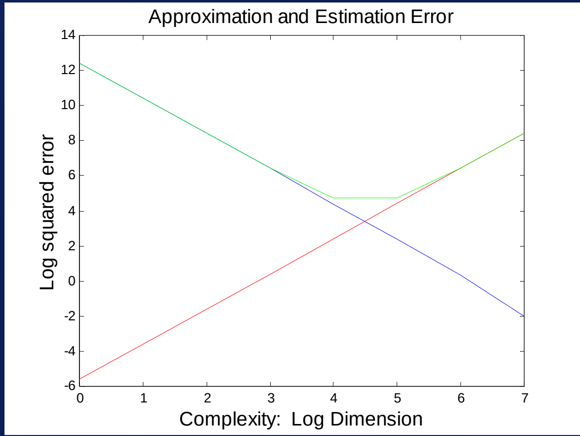


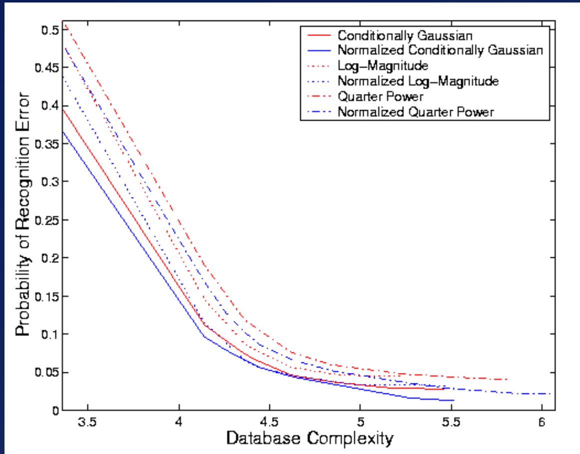
Classification error as a function of number of bits passed between the database and processor

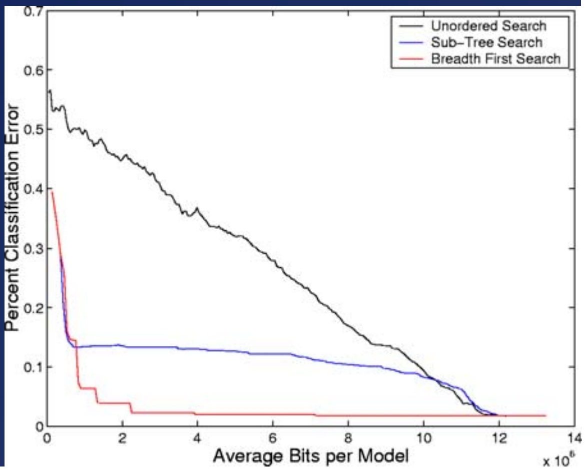
- Eventually, search covers all possibilities
- Breadth-first search quickly finds good combinations of (θ,a)
- Method for modeling target reflectivity statistics from sample images
- Target models used to estimate conditional sensor output statistics

Conclusions

- ATR Theory
  - Approximation Error
  - Estimation Error
  - Model Complexity
  - Computational Complexity
- Model-Based Approaches: Known Distributions
- Successive Refinement
- Implementation Considerations







References

•M. D. DeVore, "Recognition Performance from Synthetic Aperture Radar Imagery Subject to System Resource Constraints," D.Sc. Dissertation, Department of Electrical Engineering, Washington University, St. Louis, MO 2001.

•M. D. DeVore and J. A. O'Sullivan, "A Performance Complexity Study of Several Approaches to Automatic Target Recognition from Synthetic Aperture Radar Images," *IEEE Transactions on Aerospace and Electronic Systems*, to appear, 2001.

•M. D. DeVore, J. A. O'Sullivan, R. D. Chamberlain, and M. A. Franklin, "Relationships Between Computational System Performance and Recognition System Performance," Automatic Target Recognition XI, *Proceedings SPIE*, April 2001.

•M. D. DeVore and J. A. O'Sullivan, "Statistical Assessment of Model Fit for Synthetic Aperture Radar Data," Algorithms for Synthetic Aperture Radar, *Proc. SPIE*, April 2001

•M. D. DeVore and J. A. O'Sullivan, "Probabilistic Approach to Model Extraction from Training Data," Algorithms for Synthetic Aperture Radar, *Proc. SPIE*, April 2001

•M. D. DeVore, N. A. Schmid, and J. A. O'Sullivan, "Analytical and Experimental Performance-Complexity Tradeoffs in ATR," *Proceedings of the 34th Asilomar Conference on Signals, Systems, and Computers*, 2000.

•J. A. O'Sullivan, R. E. Blahut, and D. L. Snyder, "Information Theoretic Image Formation," Invited *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2094-2123, October 1998.

•N. A. Schmid and J. A. O'Sullivan, "Thresholding Method for Reduction of Dimensionality," to appear *IEEE Transactions on Information Theory*, 2001.

•N. A. Schmid, "Performance Analysis in Stochastic Recognition and Authentication." D.Sc. dissert., Department of Electrical Engineering, Washington University, St. Louis, MO, 2000.