ONR Mine Warfare Autonomy
Virtual Program Review
September 7, 2017

Information-driven Guidance and Control for Adaptive Target Detection and Classification

Silvia Ferrari
Pingping Zhu and Bo Fu
Mechanical and Aerospace Engineering
Cornell University

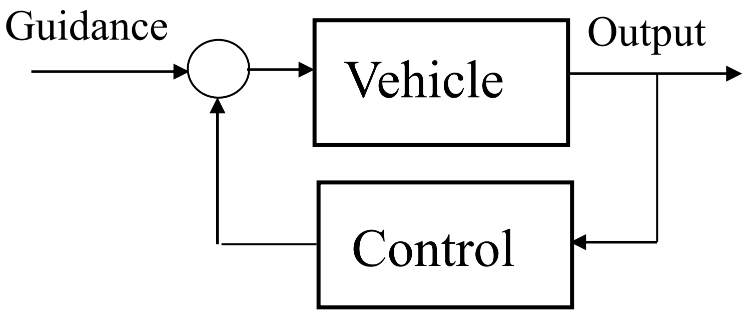
Outline

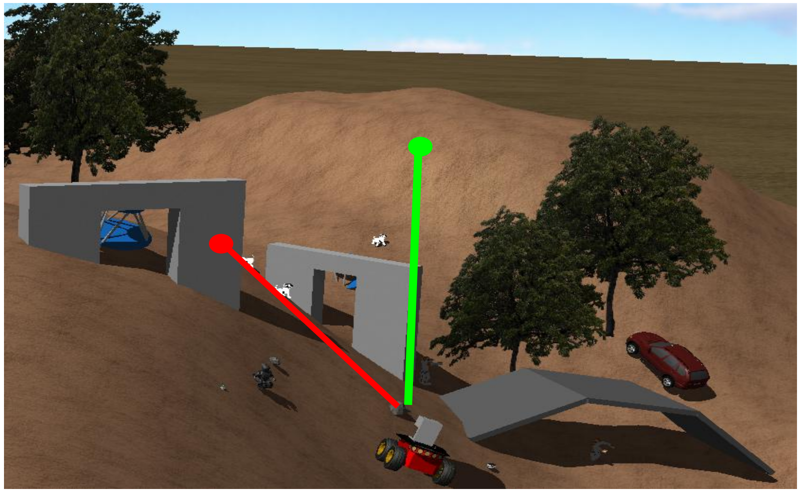
- Introduction
- MCM Motivation
- Directional Information Gain
- UUV-sonar Imaging Frames of Reference
- UUV-sonar Feature Extraction
- CNN-SVM ATR
- UUV-sonar Bayesian Modeling
- UUV-sonar Information Value Learning
- Conclusions and Q&A

Vehicle Path Planning and Control

Traditional paradigm:

Proprioceptive and exteroceptive sensor (output) used as feedback to vehicle in support of vehicle navigation objectives.

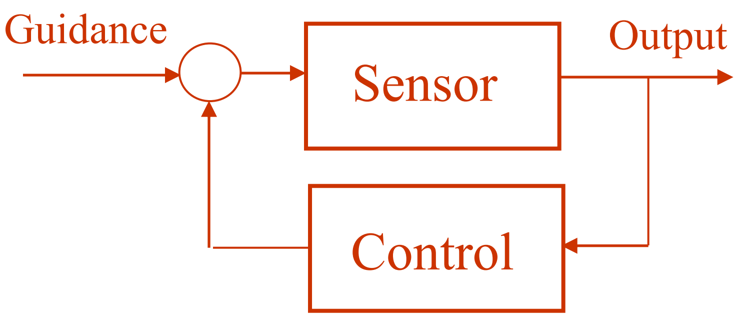


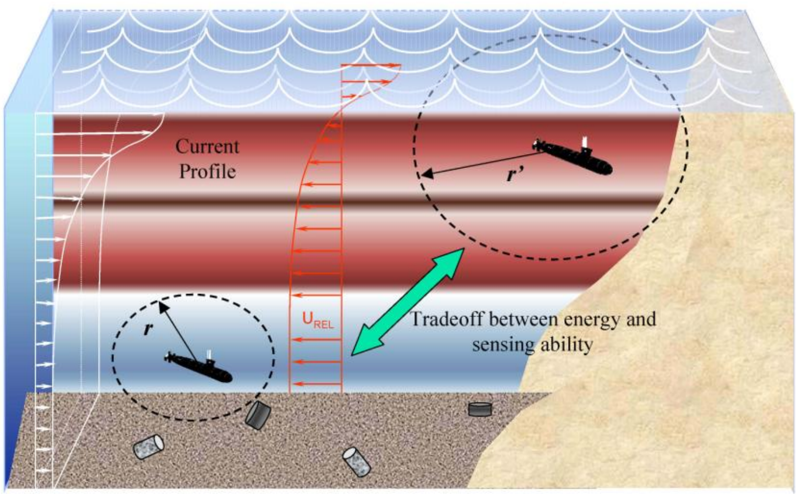


Sensor Navigation and Control

New paradigm:

Vehicle is used to gather information (output) to support **sensing objectives**, such as target acquisition, or DCLT.



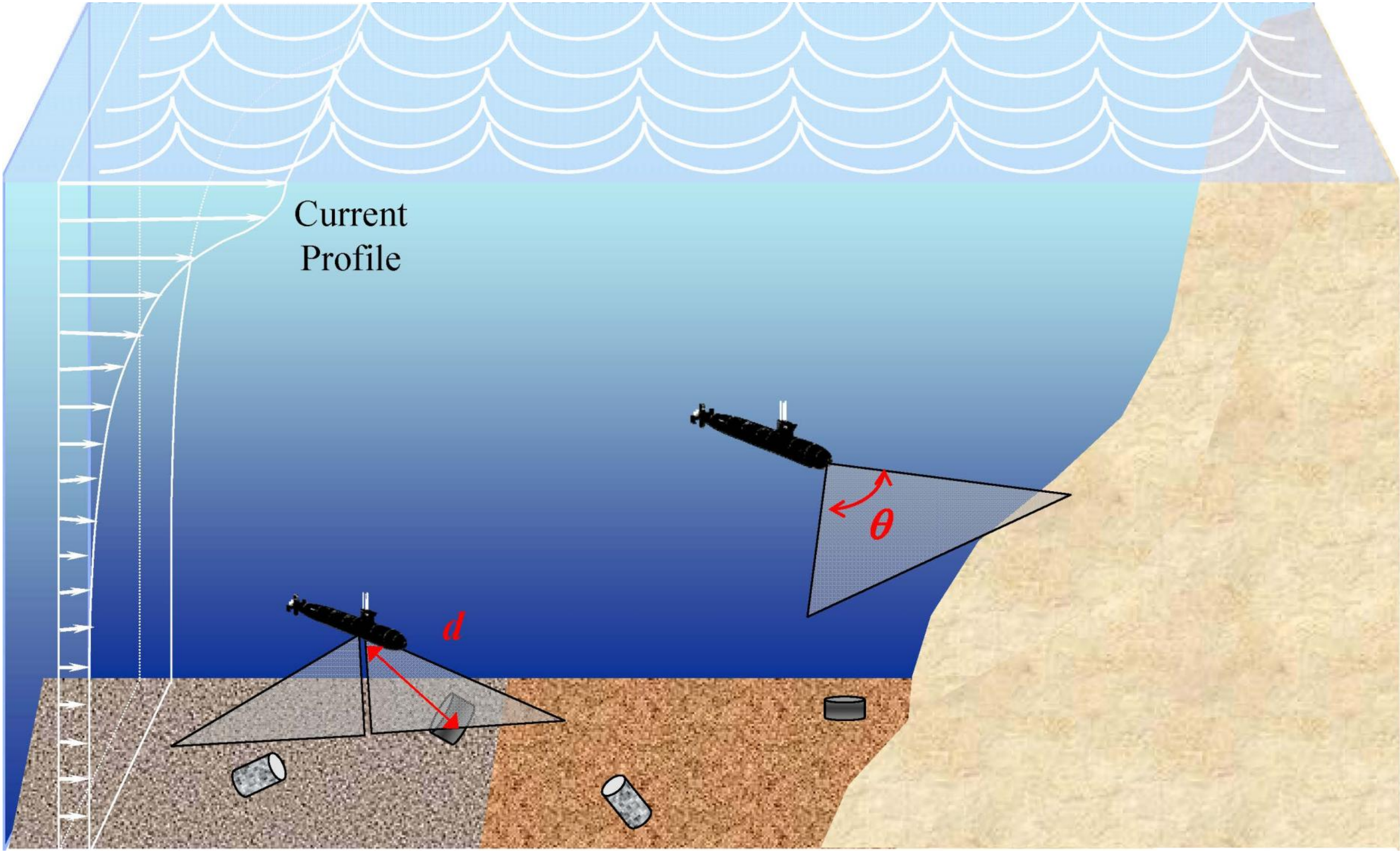


Research challenges:

- **Represent sensor objectives in closed-form**
  Computational geometry; information theory
- **Environmental and target feedback (output)**
  Significant uncertainties; Bayesian updates
- **Information-driven guidance and control**
  Couplings between sensor measurements and vehicle dynamics

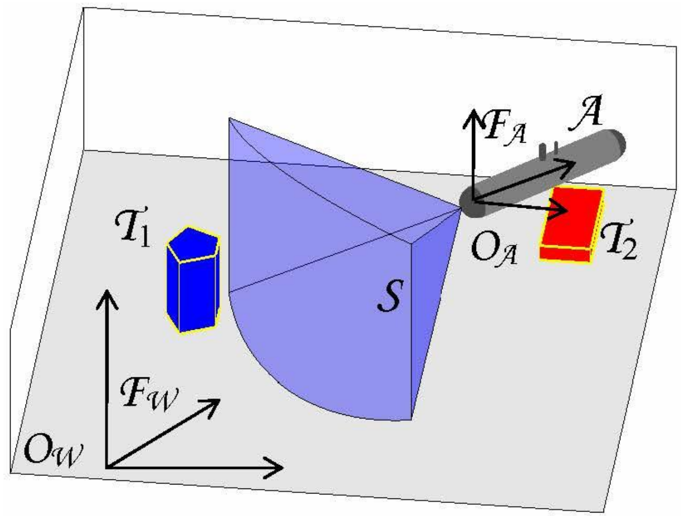
MCM Motivation and Application

Classification-driven Path Planning



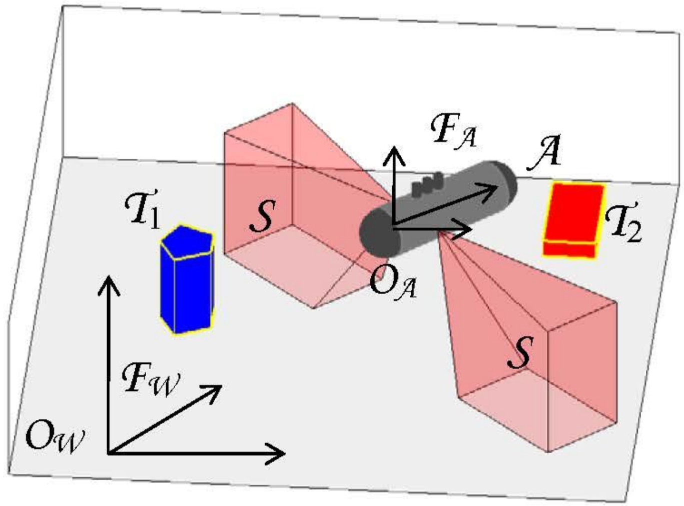
Influence of UUV Position and Orientation

Forward-looking Sonar:



(a)

Side-scan Sonar:



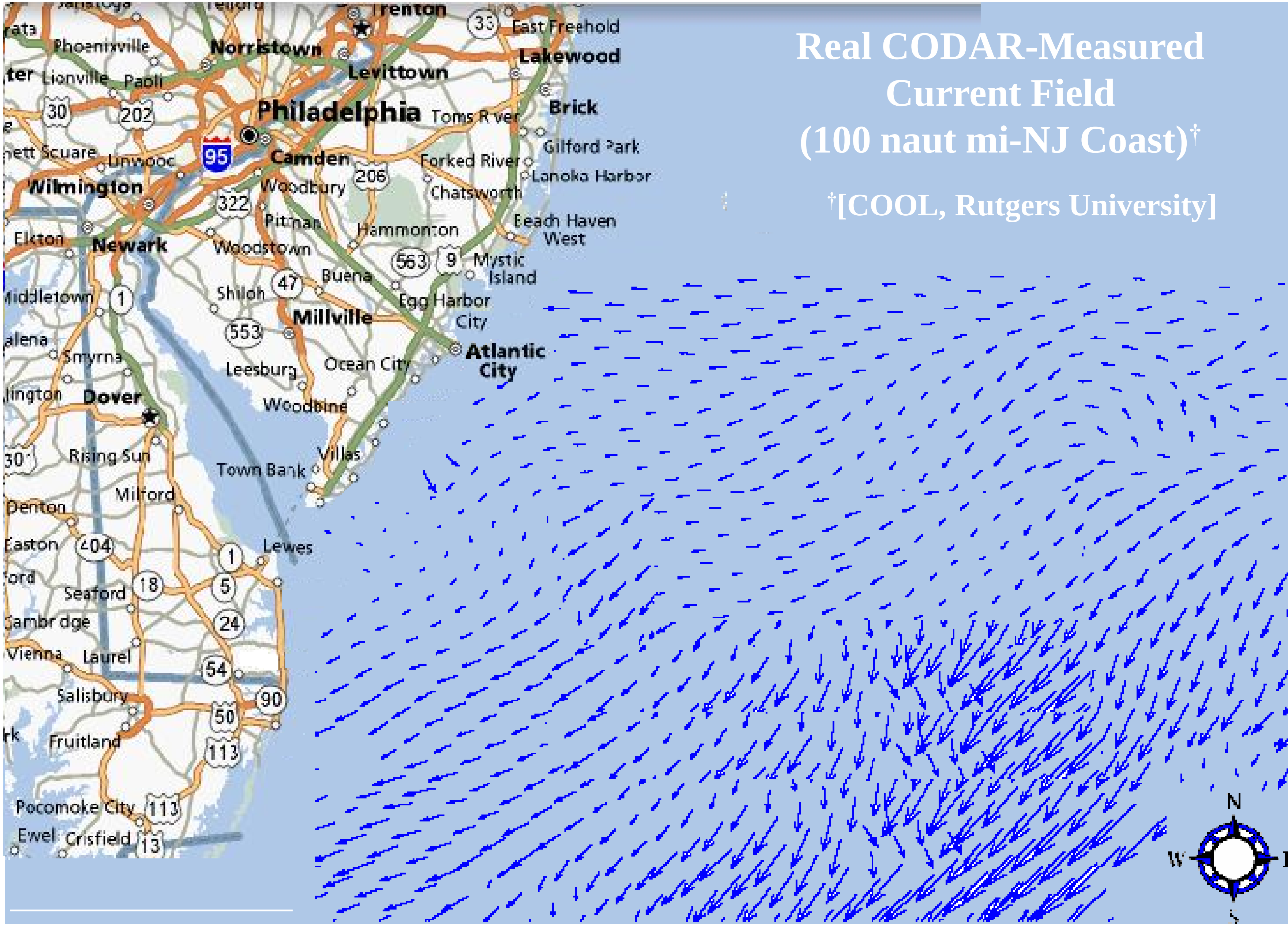
(b)

The sensor FOV, denoted by $\mathcal{S} \subset \Re^3$, is defined as a compact subset of the workspace, $\mathcal{W}$, in which the robot can obtain sonar measurements.

Motivation: Sonar is directional and the information gain depends on the sonar FOV geometry, position, and orientation relative to the target.

LISC

Surveillance Region and Oceanic Currents



Real CODAR-Measured Current Field
(100 naut mi-NJ Coast)†

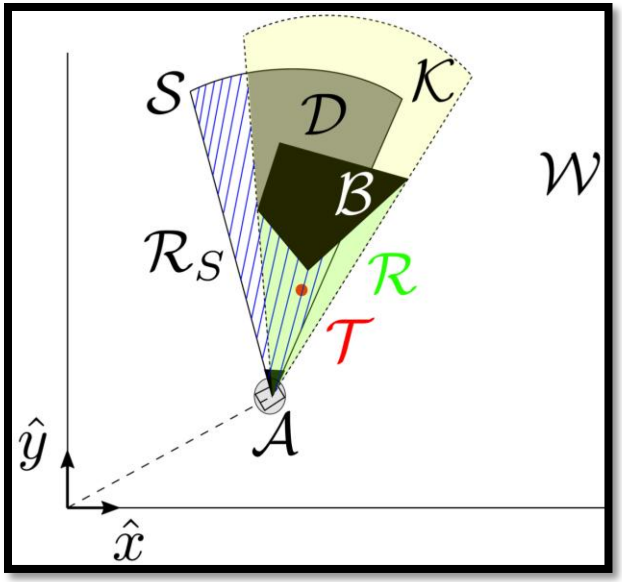
†[COOL, Rutgers University]

Directional Information Gain

Visibility Problem

- For a polygonal obstacle $\mathcal{B}$ in a workspace $\mathcal{W}$, consider a sensor on $\mathcal{A}$ observing a target $\mathcal{T}$.

- Determine obstacle-generated cone $\mathcal{K}$. Then, for a sensor FOV, S, we seek to find a shadow region, $\mathcal{D}$, such that the visibility region is $\mathcal{R} \cup \mathcal{R}s$



Solution:

- Construct a polygon with sensor at point P:
  $S = \{P, V_1, \ldots, V_N\}$, where $V_1, \ldots, V_N$ are the obstacle vertices.
- Compute the convex hull of S:

$$CS = conv(S) = \left\{ \sum_{i=1}^{|S|} k_i x_i \mid (\forall i : k_i \geq 0 \wedge \sum_{i=1}^{|S|} k_i = 1) \right\}$$

- Visibility region for one target:

$$\mathcal{R} = CS \setminus \mathcal{B}$$

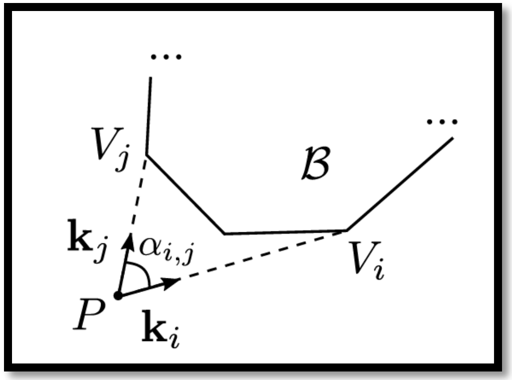
$$\mathcal{R}_S = (S \setminus \mathcal{K}) \cup ((S \cup \mathcal{K}) \setminus (\mathcal{D} \cup \mathcal{B}))$$

Visibility Region in Closed Form

Obstacle cone in closed-form:

- Construct unit vectors $\mathbf{k}_i$, $\mathbf{k}_j$
- Find boundary unit vectors $\mathbf{k}_1$ and $\mathbf{k}_2$:

$$\mathbf{k}_1 = \frac{\overline{PV}_p}{\left\|\overline{PV}_p\right\|} \qquad \mathbf{k}_2 = \frac{\overline{PV}_q}{\left\|\overline{PV}_q\right\|}$$



- Compute obstacle cone $\mathcal{K}$:

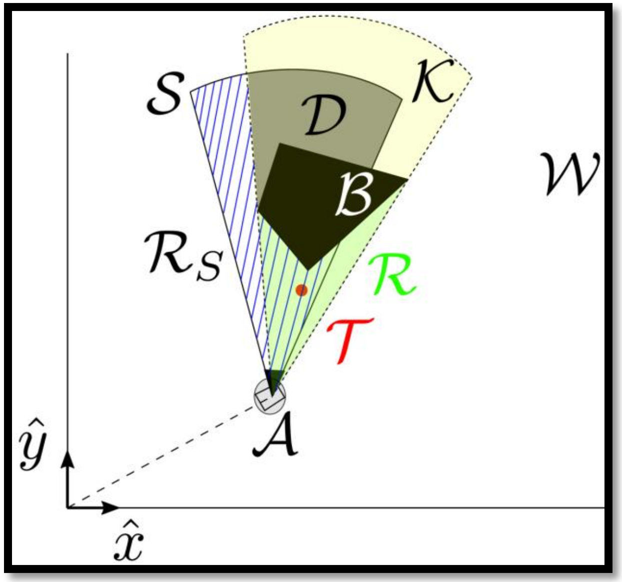
$$\mathcal{K} = cone(\hat{\mathbf{k}}_1, \hat{\mathbf{k}}_2) = \{\mathbf{x} \mid \mathbf{x} = \mathbf{r} + c_1\hat{\mathbf{k}}_1 + c_2\hat{\mathbf{k}}_2, c_1, c_2 \geq 0\}$$

$$\mathcal{D} = (S \cap \mathcal{K}) \setminus (CS)$$

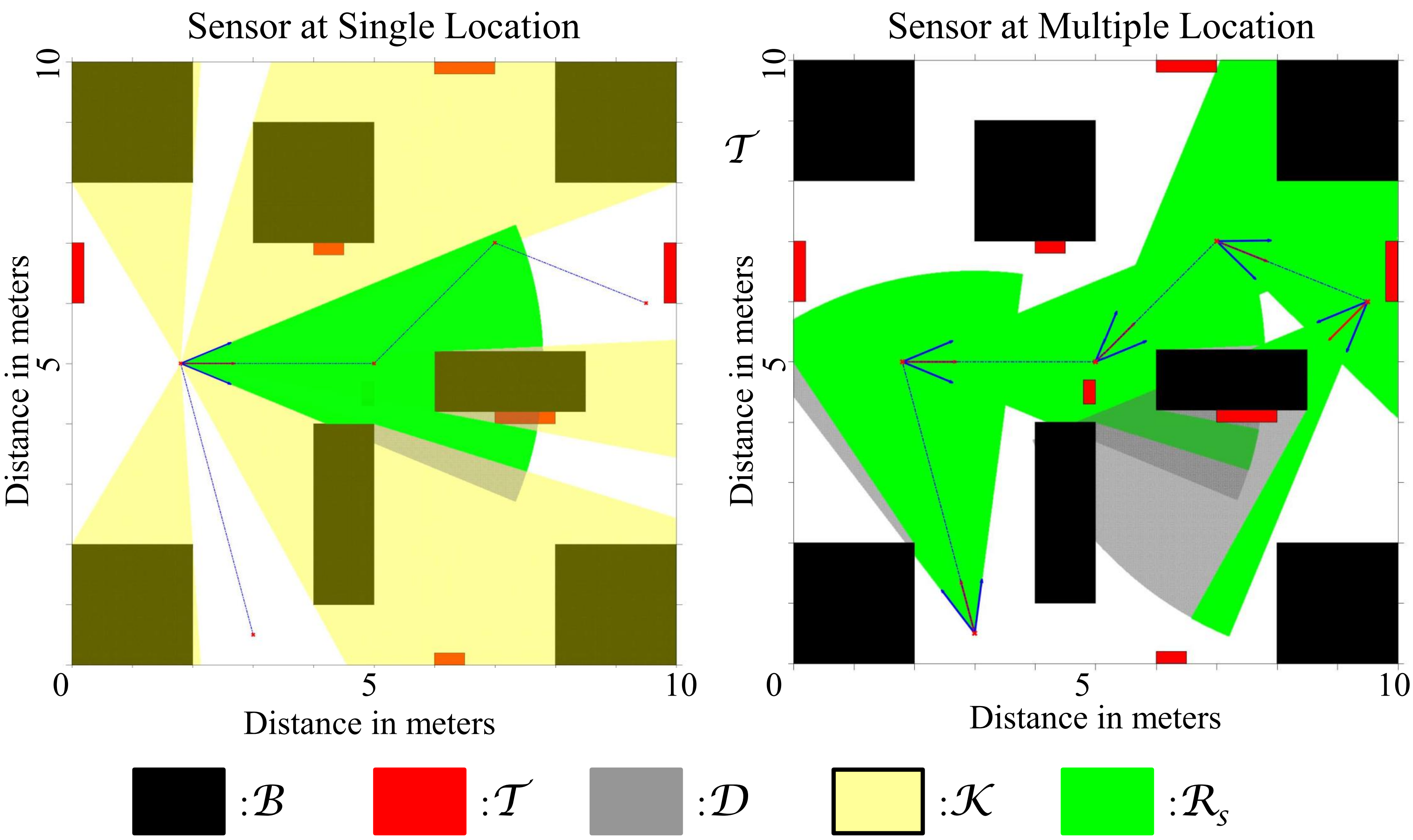
- Visibility region for multiple targets (i) and obstacles (j):

$$\mathcal{R}_{(i)} = CS_{(i)} \setminus \mathcal{B}_{(i)}; \qquad \mathcal{D}_{(i)} = (S \cup \mathcal{K}_{(i)}) \setminus CS_{(i)}$$

$$\begin{aligned} \mathcal{R}_S &= \cup_i \, \mathcal{R}_{S(i)} \\ &= \cup_i \{(S \setminus \mathcal{K}_{(i)}) \cup ((S \cup \mathcal{K}_{(i)}) \setminus (\mathcal{D}_{(i)} \cup \mathcal{B}_{(i)})\} \end{aligned}$$

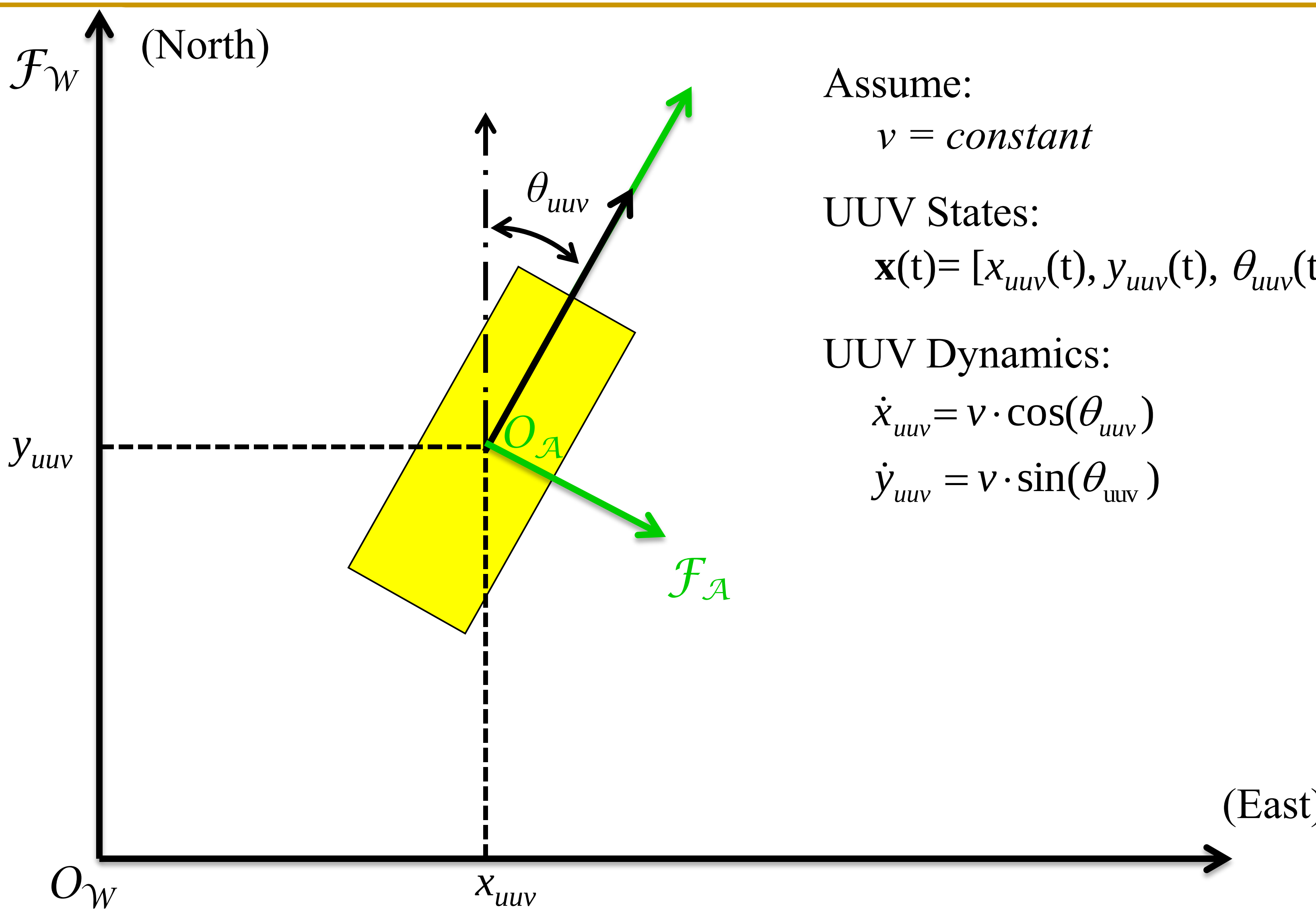


Example: Art Gallery Problem



UUV-Sonar Directional Information Gain

UUV Kinematics Frames of Reference



Assume:

$v = constant$

UUV States:

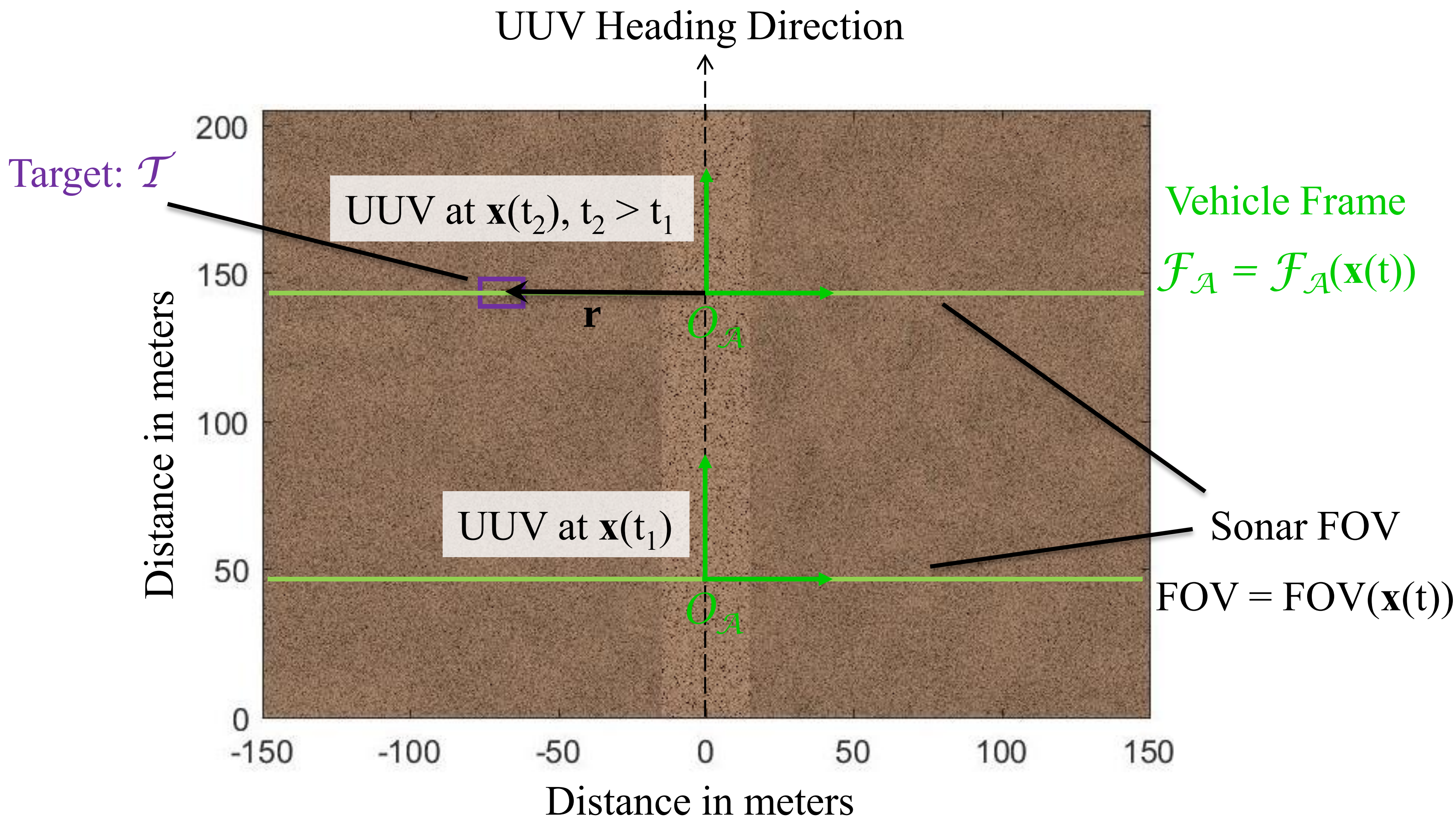
$\mathbf{x}(\mathrm{t}) = [x_{uuv}(\mathrm{t}), y_{uuv}(\mathrm{t}), \theta_{uuv}(\mathrm{t})]^T$

UUV Dynamics:

$$\dot{x}_{uuv} = v \cdot \cos(\theta_{uuv})$$

$$\dot{y}_{uuv} = v \cdot \sin(\theta_{uuv})$$

UUV Frames Relative to Sonar Image



Vehicle Frame and Image Frame

• Consider the image segmentation as the target geometry $\mathcal{T}$. Then, the actual target geometry, orientation, and shape are viewed as hidden random features.

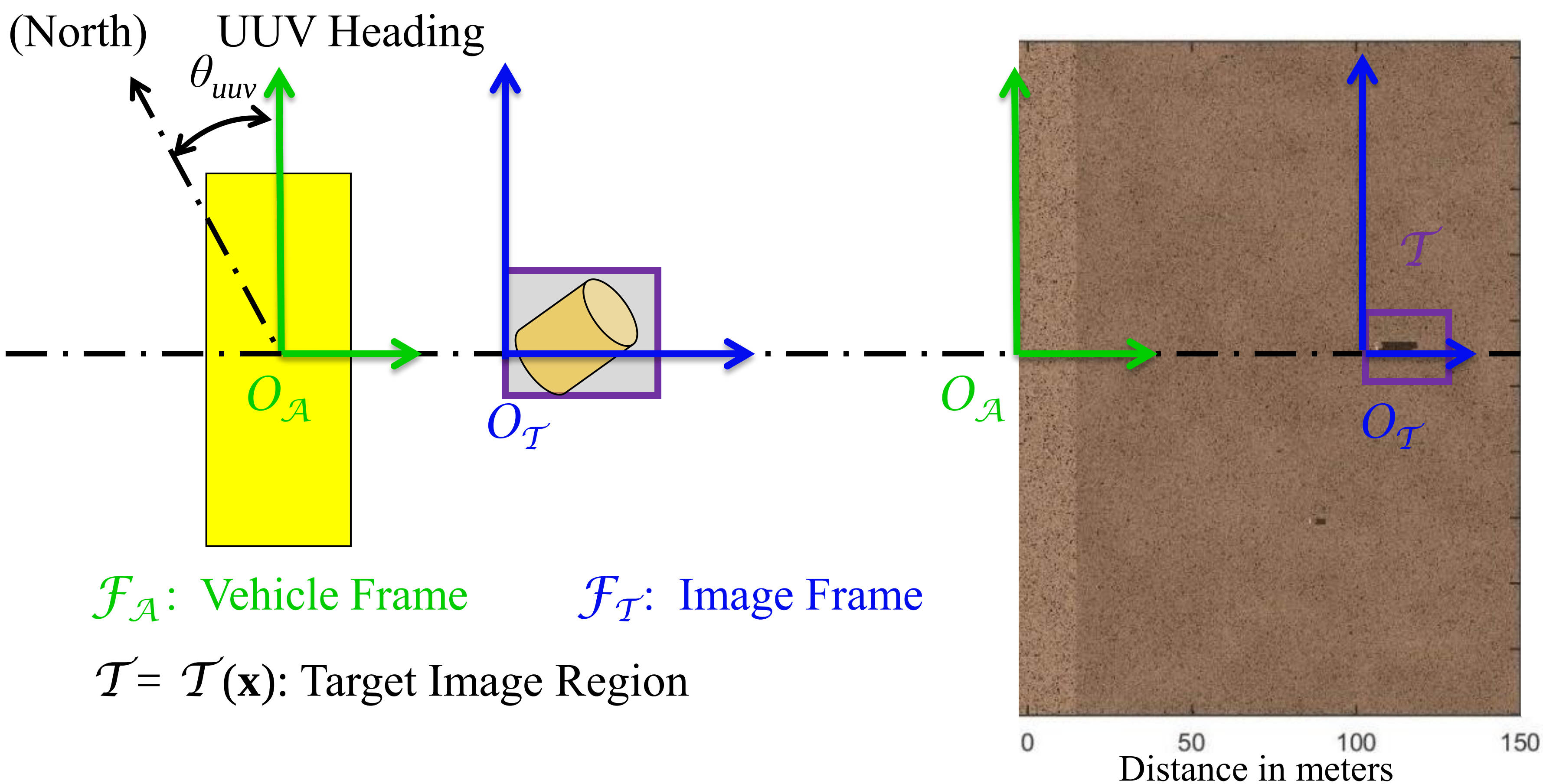


Image Frame and Target Frame

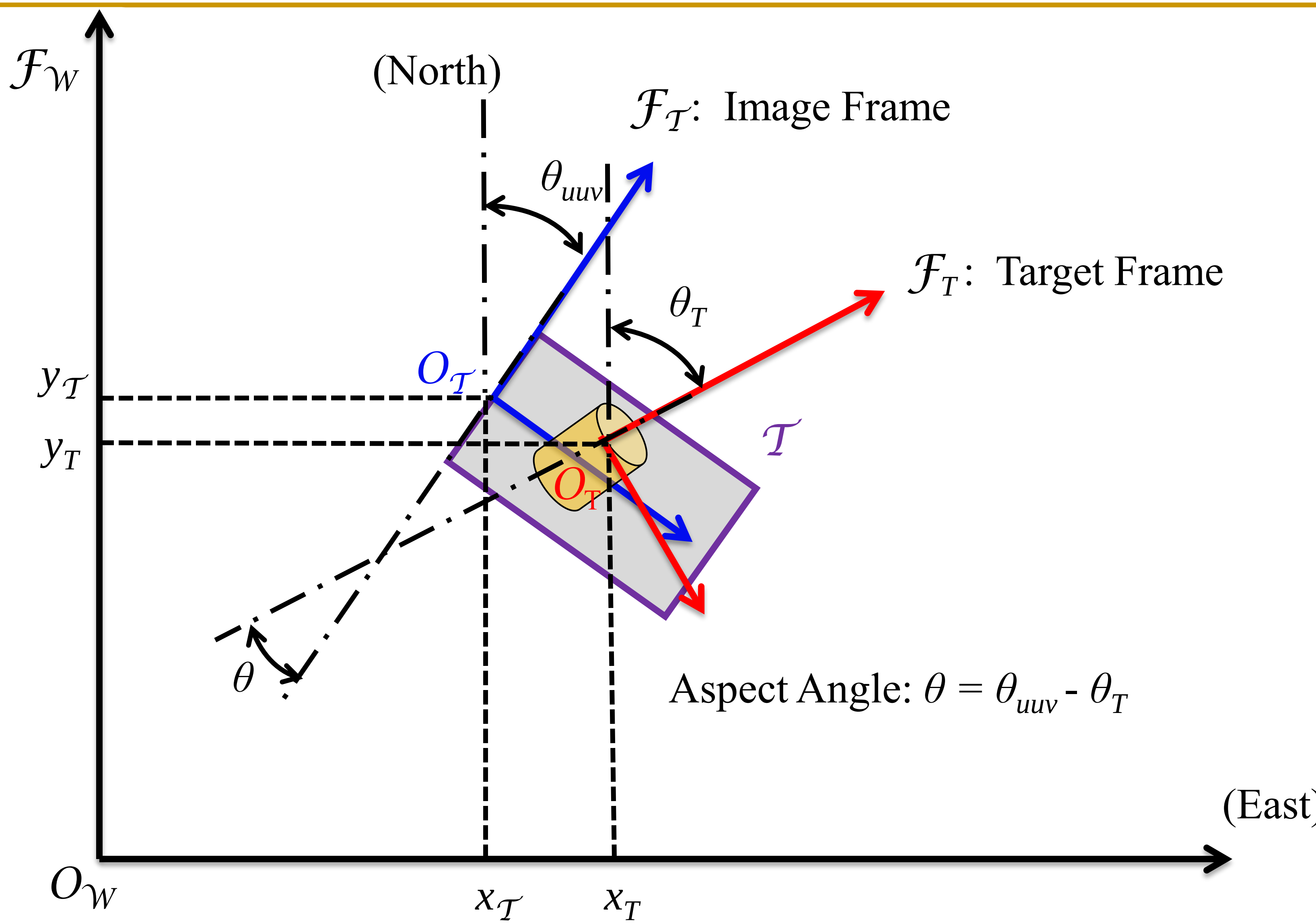
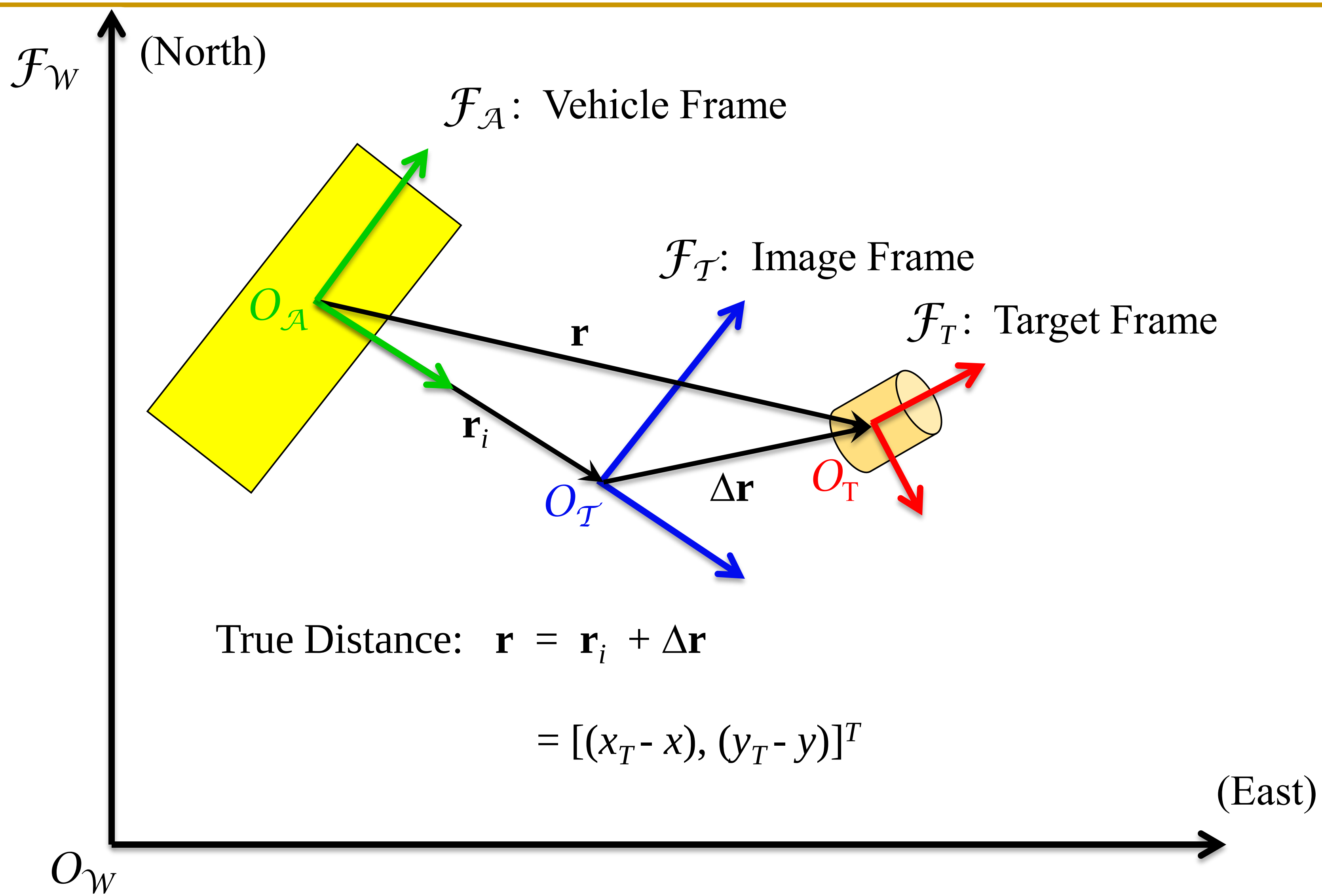
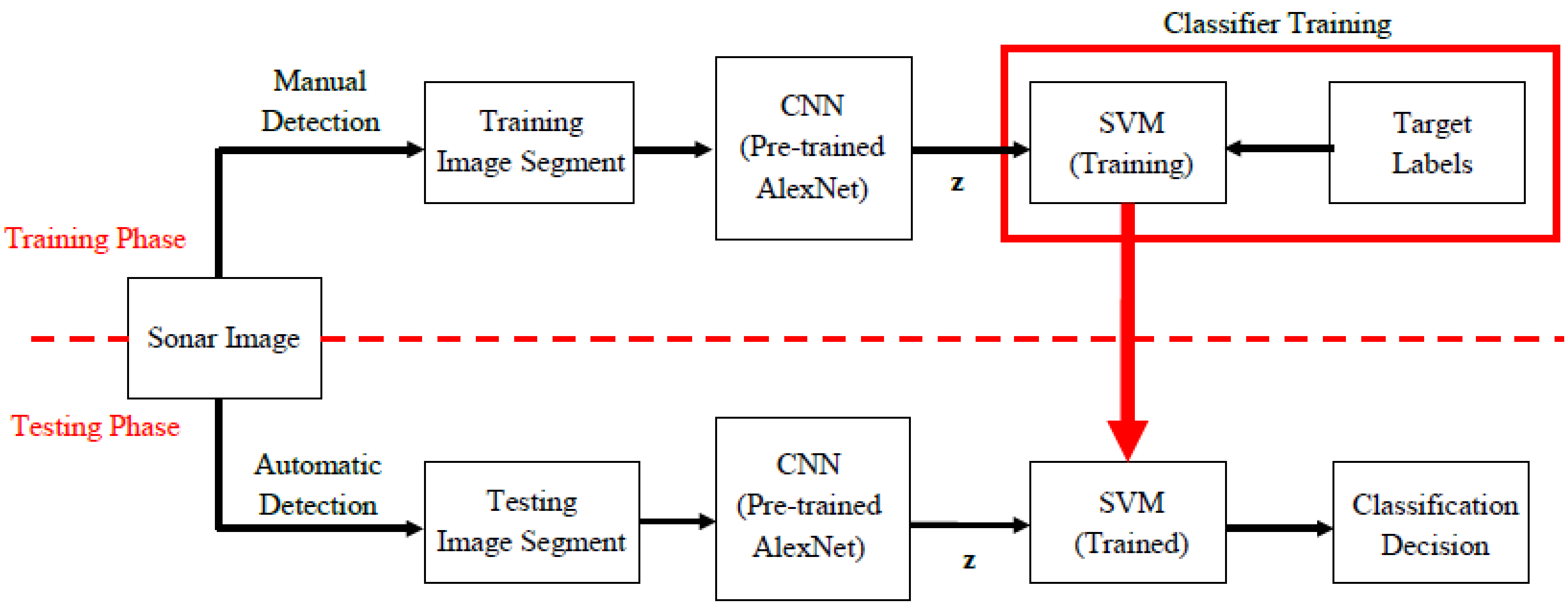


Image and Target Distance Vectors



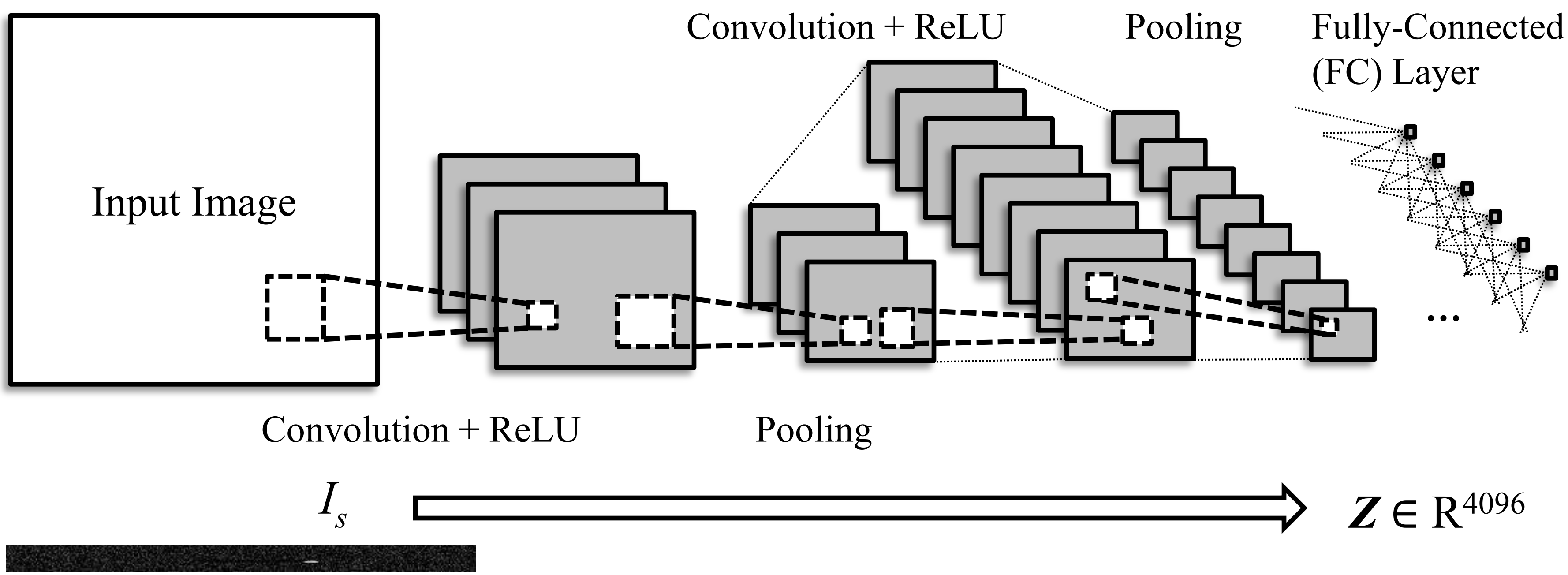
UUV-Sonar Feature Extraction

Automatic Target Recognition an Detection



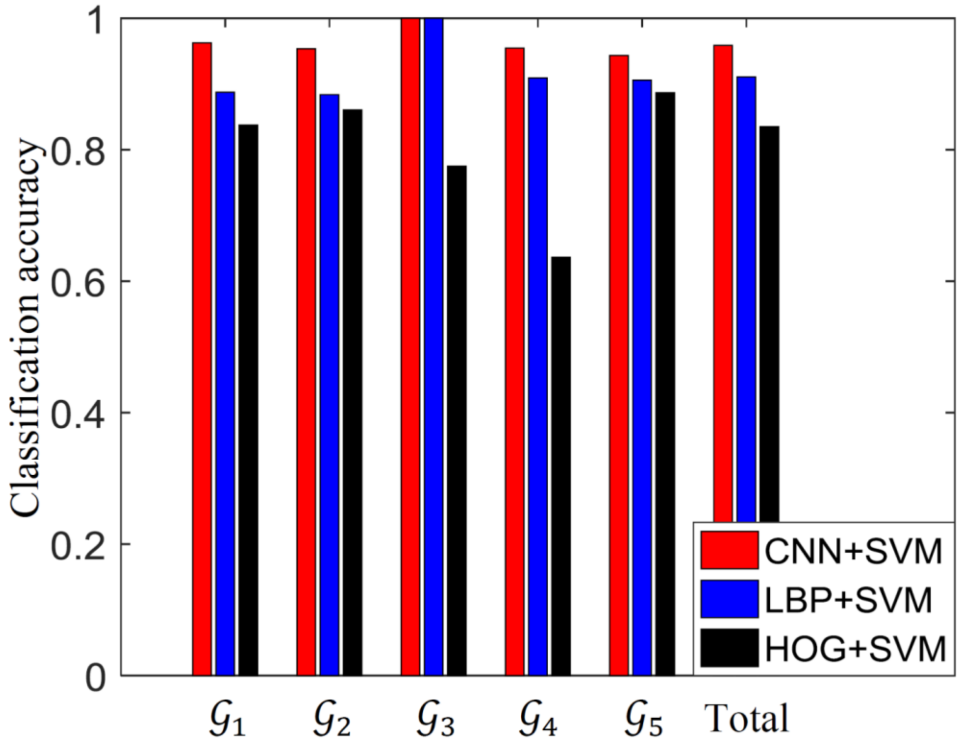
- **Training Phase:** Train SVM classifier with Sonar Image I_s and true classification Y_T
- **Testing Phase:** Matched filter for image segmentation and CNN+SVM for target classification

CNN for Sonar Image Features Extraction



- Automatic segmentation is performed via matched filter
- Matched filter provides better performance than Markov random fields (MRFs)
- Deep learning features extracted from segmentation by Pre-trained AlexNet
- AlexNet CNN provides better performance than other features extraction techniques such as HOG and LBP.

CNN Performance: Classification Accuracy



Legend:

- TP: True Positive
- FP: False Positive
- TN: True Negative
- FN: False Negative
- $\mathcal{G}_i$: Image Group i

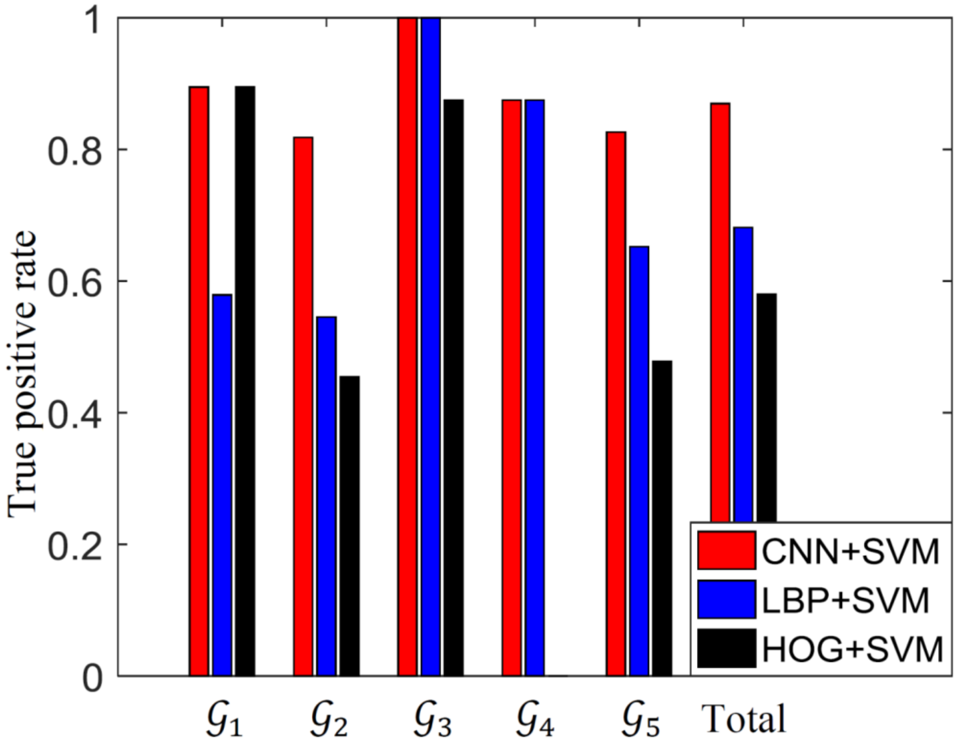
Classification Accuracy:

$$CA = \frac{n_{TP} + n_{TN}}{n_{TP} + n_{FN} + n_{FP} + n_{TN}}$$

Probability of detection (Matched Filter performance): 88.31%

* Zhu, P., Issacs, J., Fu, B., Ferrari, S., "Deep Learning Feature Extraction for Target Recognition and Classification in Underwater Sonar Images" 56th IEEE Conference on Decision and Control, Melbourne, Australia (Accepted).

CNN Performance: True Positive Rate



Legend:

- TP: True Positive
- FP: False Positive
- TN: True Negative
- FN: False Negative
- $\mathcal{G}_i$: Image Group i

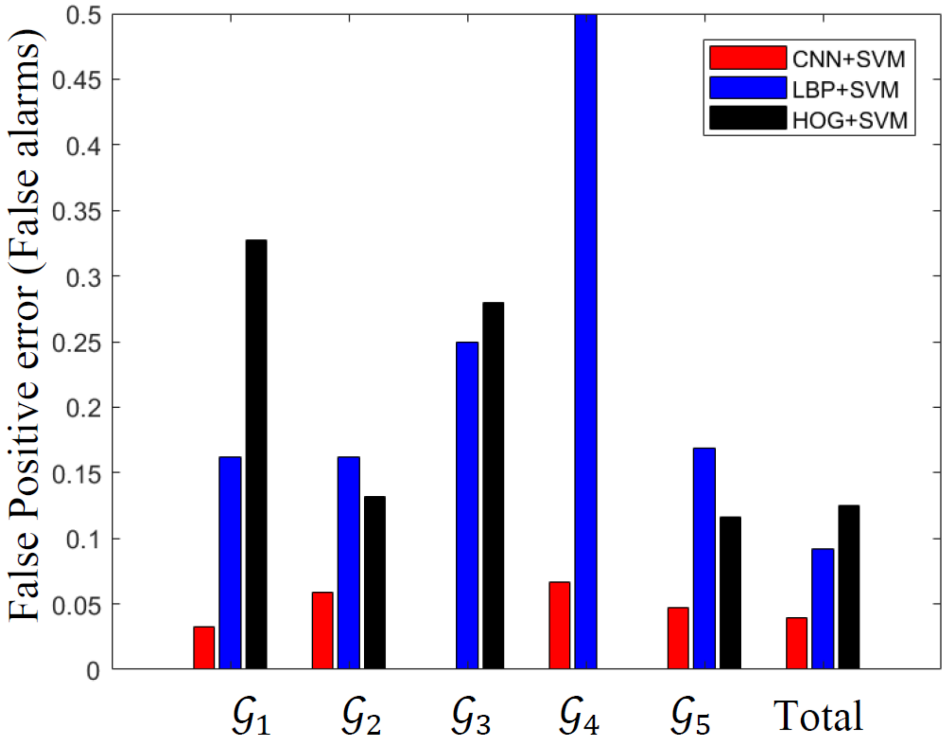
True Positive Rate (TPR):

$$TPR = \frac{n_{TP}}{n_{TP} + n_{FP}}$$

Probability of detection (Matched Filter performance): 88.31%

* Zhu, P., Issacs, J., Fu, B., Ferrari, S., "Deep Learning Feature Extraction for Target Recognition and Classification in Underwater Sonar Images" 56th IEEE Conference on Decision and Control, Melbourne, Australia (Accepted).

CNN Performance: False Alarms



Legend:

- TP: True Positive
- FP: False Positive
- TN: True Negative
- FN: False Negative
- $\mathcal{G}_i$: Image Group i

False Positive Rate (FPR):

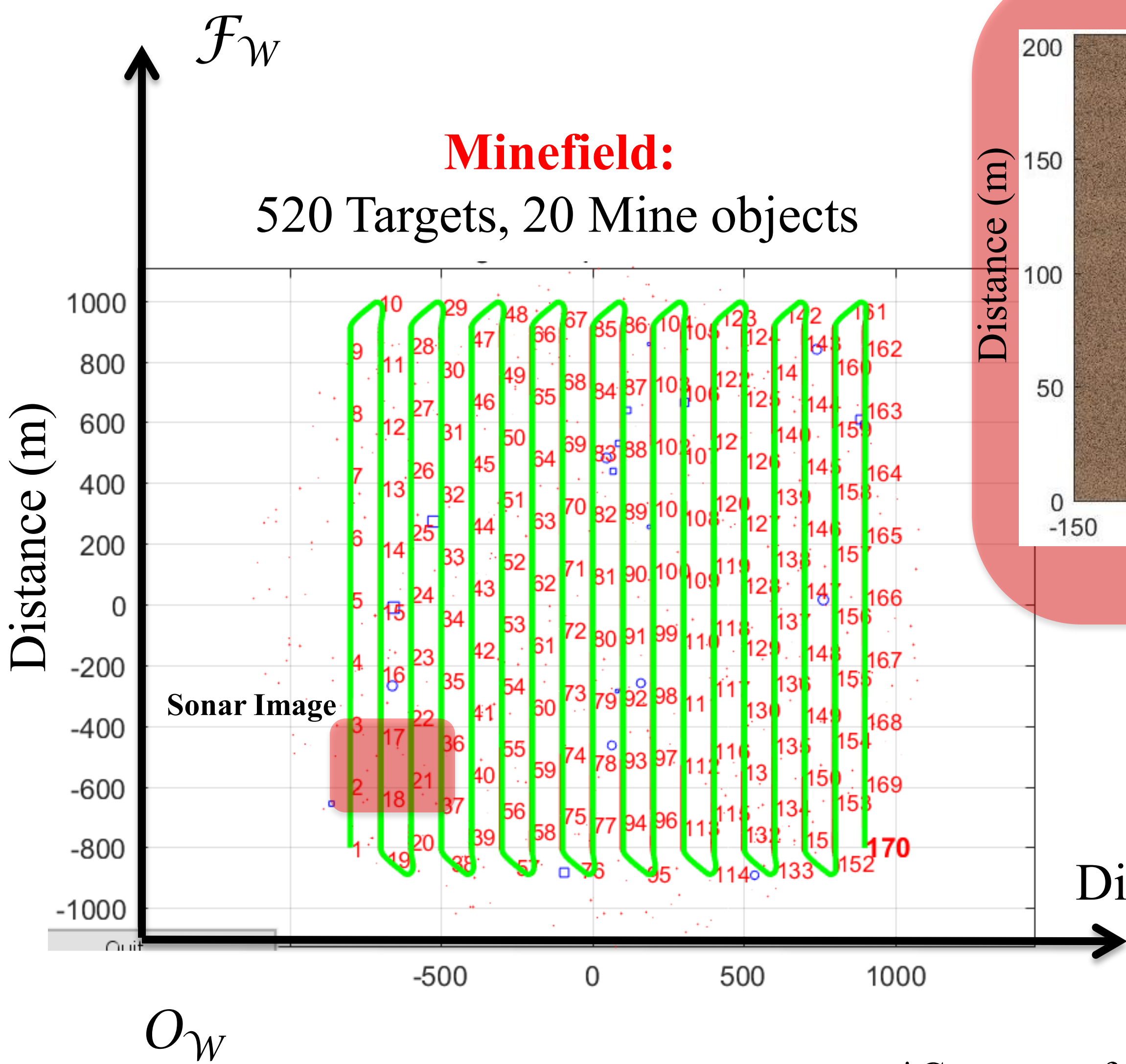
$$FPR = \frac{n_{FP}}{n_{FP} + n_{TN}}$$

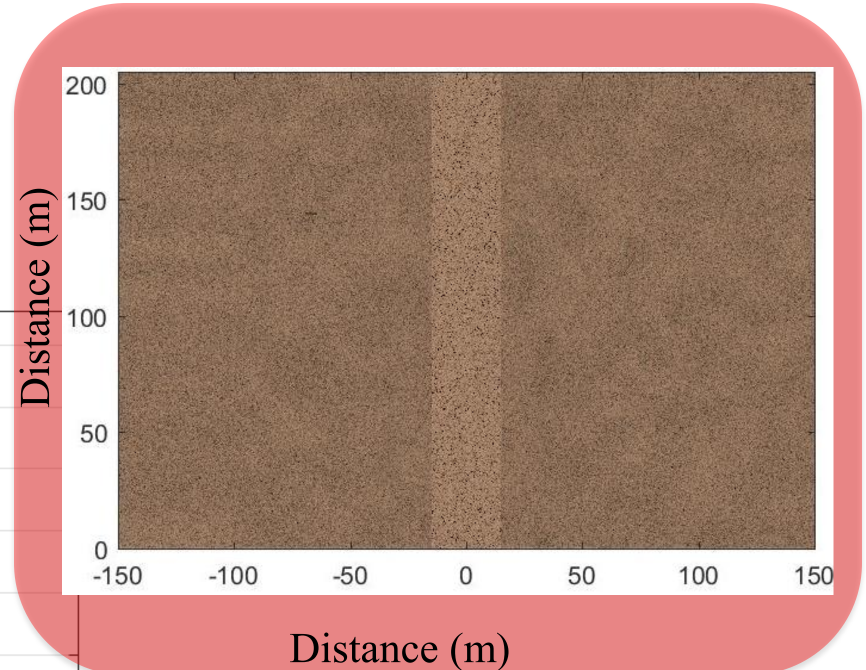
Probability of detection (Matched Filter performance): 88.31%

* Zhu, P., Issacs, J., Fu, B., Ferrari, S., "Deep Learning Feature Extraction for Target Recognition and Classification in Underwater Sonar Images" 56th IEEE Conference on Decision and Control, Melbourne, Australia (Accepted).

UUV-Sonar Information Value

UUV-Sonar Information Value





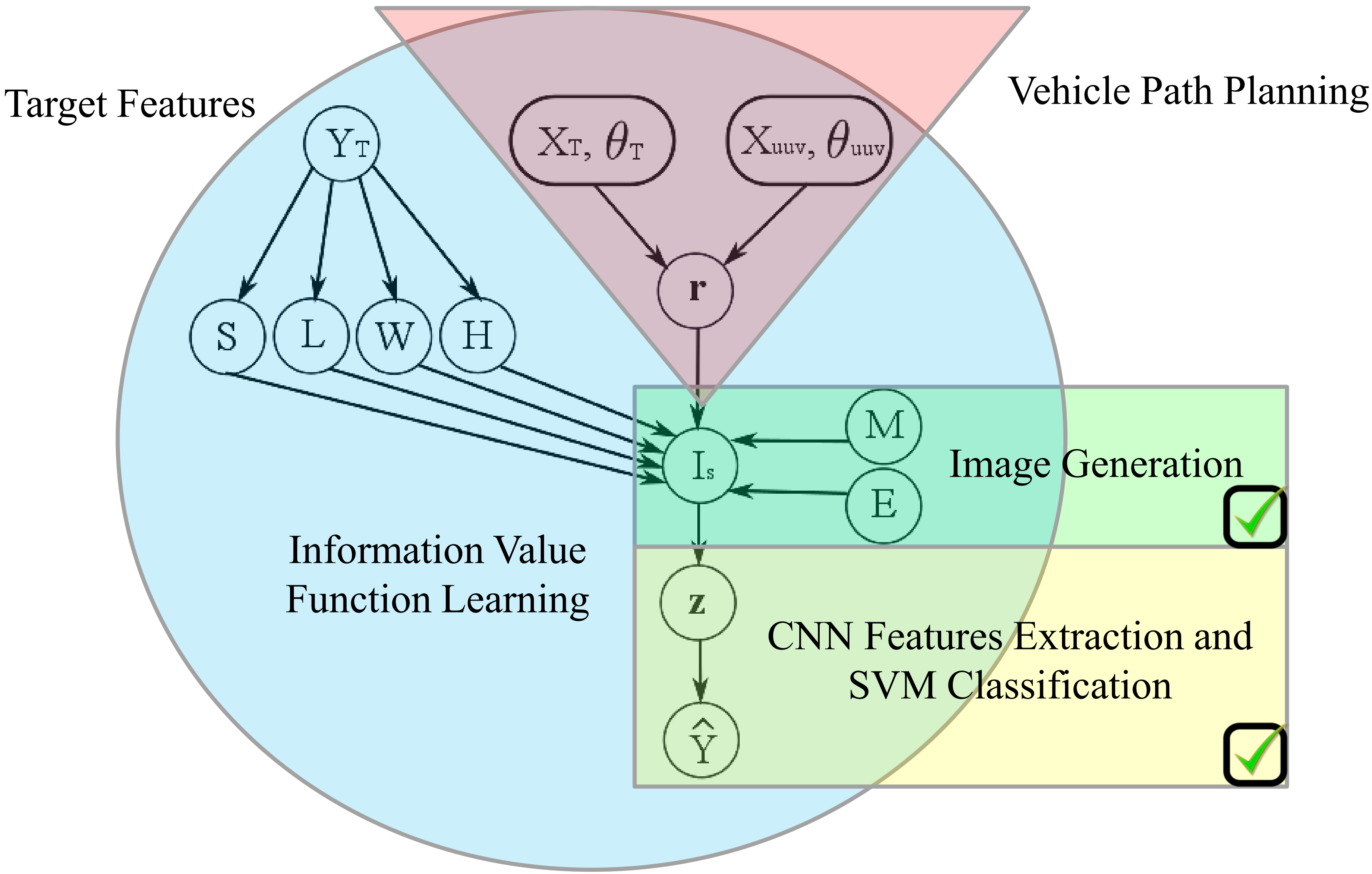
Sonar Image

*Courtesy of Jason Isaacs, NSWC Panama City, FL

Key Variable Description

	Name	Description	Name	Description
	X_{uuv}	Vehicle Position	X_T	Target Position
	θ_{uuv}	Vehicle Orientation	θ_T	Target Orientation
$\mathbf{r} = [\, d, \boldsymbol{\theta} \,]$	d	Relative Distance	θ	Relative Orientation
	Y_T	True Classification	$\hat{Y}$	Estimated Classification
Features $\mathbf{F} = \{S,H,W,L\}$	S	Object Shape	L	Object Length
	W	Object Width	H	Object Height
	I_s	Segmented Image	$\mathbf{z}$	CNN output Features
Assume Constant	M	Sensor Mode	E	Noise level

Key Variables and Causal Relationships

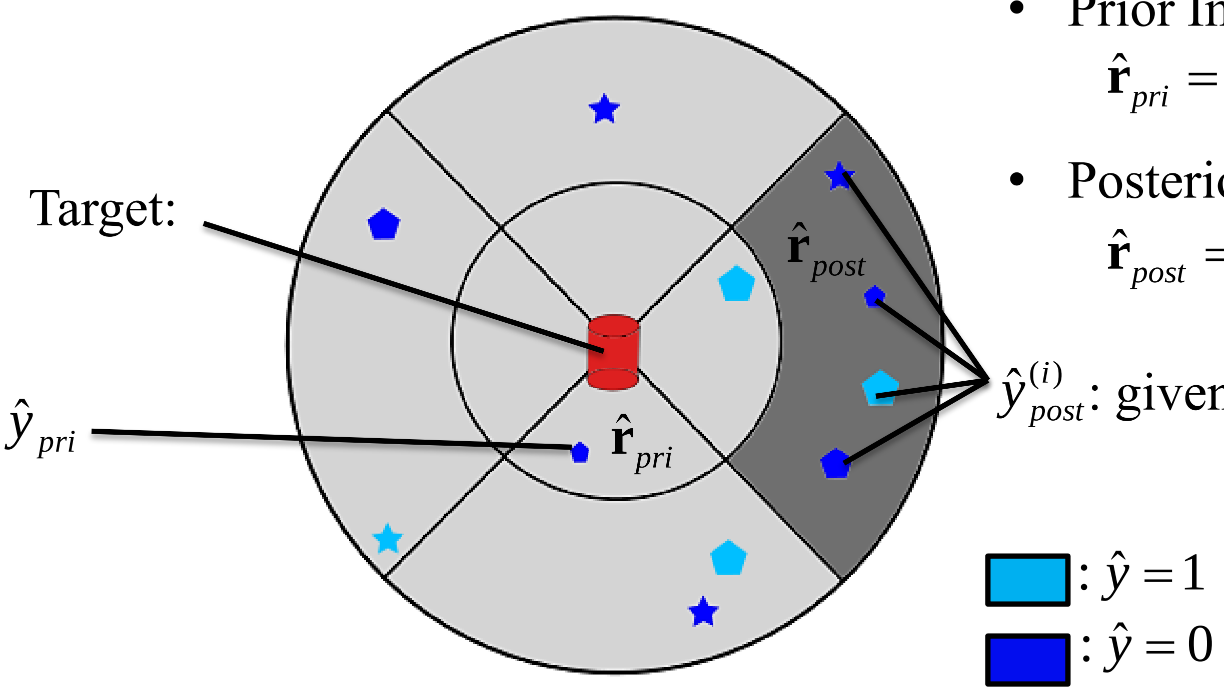


Information Value Function Learning

- Assume information gain can be modeled as Expected Entropy Reduction (EER)

$$EER(\hat{y}_{pri}, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}, \hat{\mathbf{r}}_{post}) = H_{pri}(\hat{y}_{pri}, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}) - \frac{1}{N}\sum_{i=1}^{N} H_{post}(\hat{y}_{pri}, \hat{y}_{post}^{(i)}, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}, \hat{\mathbf{r}}_{post})$$

- Example:

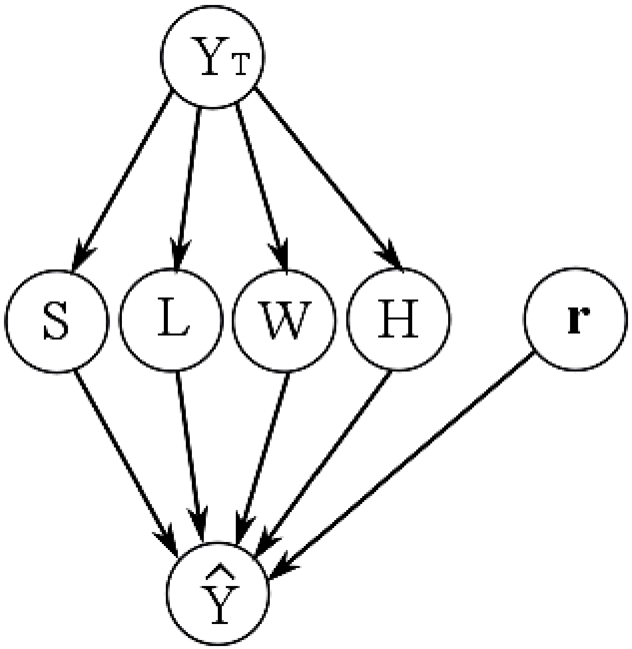


- Prior Image location
  $\hat{\mathbf{r}}_{pri} = \{[0,75), [5\pi/4, 7\pi/4,)\}$
- Posterior Image location
  $\hat{\mathbf{r}}_{post} = \{[75,150], [-\pi/4, \pi/4,)\}$

Discretization:

$\mathbf{r}$ = [d, θ]; d = {[0,75), [75,150]}; θ = {[-π/4, π /4), [π /4, 3 π /4), [3 π /4, 5 π /4), [5 π /4, 7 π /4]}

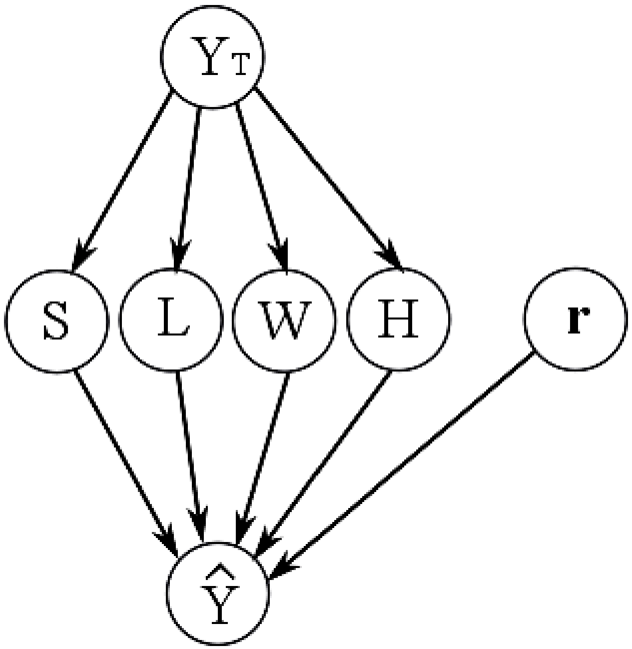
Information Value Function Learning



Bayesian Network Model,
Features **F** = {S,H,W,L}

- Continuous probabilities (PDF's) are difficult to learn, random variables are first discretized.
- Conditional probability table (CPT) is estimated from dataset:
  $P(\hat{Y} \mid Y_T, S, L, W, H, \mathbf{r})$
- Entropy *H* is calculated based on estimated conditional probability
- Information Value Function (EER) is learned from the dataset images and Features.

Bayesian Network Model



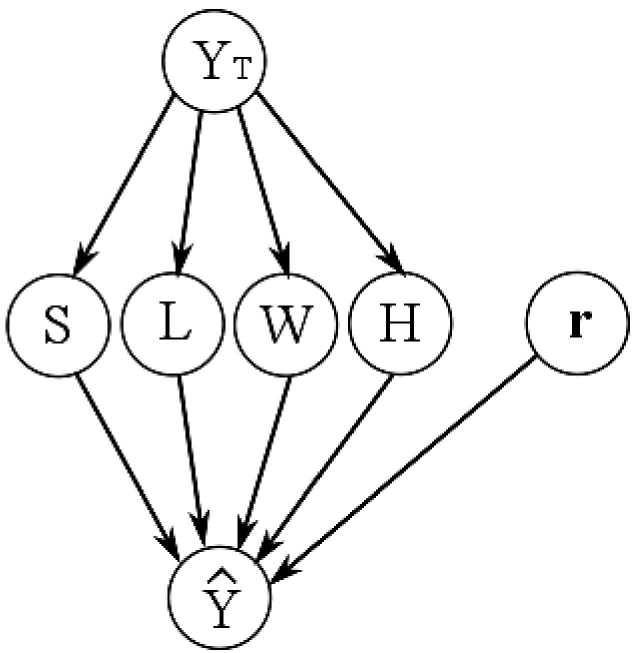
Bayesian Network Model,
Features **F** = {S,H,W,L}

- Training data set: $\mathcal{D} = \{(s_i, l_i, w_i, h_i, d_i, \theta_i, y_{T,i}, \hat{y}_i)\}_{i=1}^{n}$
  - Random variables: $S, L, W, H, D, \Theta, Y_T$, and $\hat{Y}$
  - Realization variables: $s_i, l_i, w_i, h_i, d_i, \theta_i, y_T$, and $\hat{y}$

- Probability parameters learned from dataset:
  - Prior: $P(Y_T \mid \hat{\mathbf{F}}, \hat{\mathbf{r}})$
  - Likelihood: $P(\hat{y} \mid Y_T, \hat{\mathbf{F}}, \hat{\mathbf{r}})$

S = {Cylindrical, Rectangular};
V = $(LWH)^{1/3}$ = {[0, 0.14), [0.14, 0.3), [0.3, 1.1), [1.1, 1.7]}
r = [d, □]; d = {[0,75), [75,150]}; □ = {[-□/4, □/4), [□/4, 3□/4), [3□/4, 5□/4), [5□/4, 7□/4]

BN Parameters Learning



Bayesian Network Model,
Features $\mathbf{F} = \{S,H,W,L\}$

i :	Variable Node index
j:	Parent configuration index
k:	Value index of variable
r_i:	Number of values of variable
N_{ijk}:	Count number in data set

- Bayesian parameter estimation:
  - Parameters are random variables η_{ijk}
  - Multinomial likelihood of training data set:

$$P(\mathcal{D}|\eta) = \prod_n \prod_{ijk} \eta_{ikj}^{N_{ikj}}$$

  - Dirichlet Prior of the parameter:

$$P(\eta|\boldsymbol{\alpha}) = \frac{\Gamma(\sum_k \alpha_{ijk})}{\prod_k \Gamma(\alpha_{ijk})} \prod_{k=1}^{r_i} \eta_{ikj}^{(\alpha_{ijk}-1)}$$

$$= Dirichlet(\eta_{ij}; \alpha_{ij1}, \dots, \alpha_{ijr_i})$$

    - Hyper parameter: $\boldsymbol{\alpha} = [\alpha_{ij1}, \dots, \alpha_{ijr_i}]$
  - Posterior of the parameter:

$$P(\eta|\mathcal{D},\boldsymbol{\alpha}) \propto \prod_{ijk} \eta_{ikj}^{(N_{ijk}+\alpha_{ijk}-1)}$$

  - MAP estimate:

$$\hat{\eta}_{ijk}^{MAP} = \frac{N_{ijk}+\alpha_{ijk}}{\sum_{j'}(N_{ij'k}+\alpha_{ij'k})}$$

Posterior Probability Learning

- Posterior probabilities calculated via Bayes rule:
  - Given image I_{pri}

$$P(Y_T = 1 \mid \hat{y}_{pri}, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}) = \frac{P(Y_T = 1 \mid \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}) P(\hat{y}_{pri} \mid Y_T = 1, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri})}{P(\hat{y}_{pri} \mid \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri})}$$

  - Given image I_{pri} and I_{post}

$$P(Y_T = 1 \mid \hat{y}_{pri}, \hat{y}_{post}, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}, \hat{\mathbf{r}}_{post}) =$$

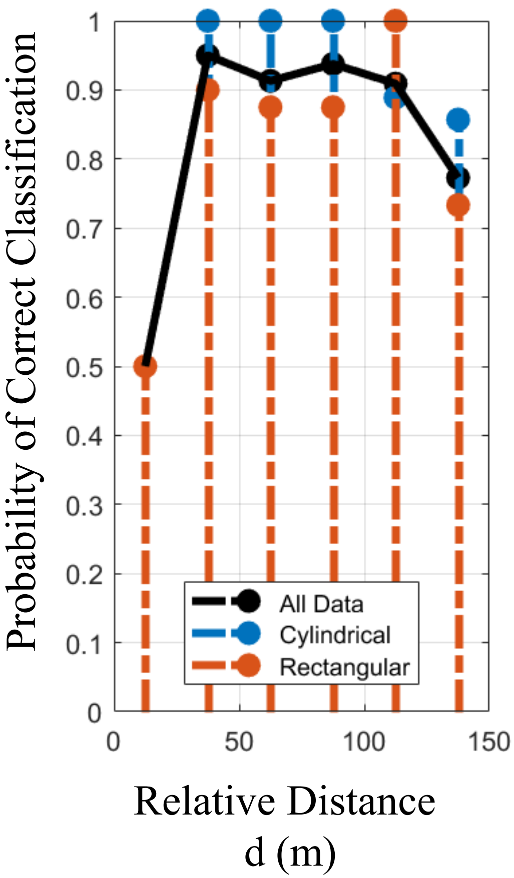
$$\frac{P(Y_T = 1 \mid \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}, \hat{\mathbf{r}}_{post}) P(\hat{y}_{pri} \mid Y_T = 1, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}) P(\hat{y}_{post} \mid Y_T = 1, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{post})}{P(\hat{y}_{pri}, \hat{y}_{post} \mid \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}, \hat{\mathbf{r}}_{post})}$$

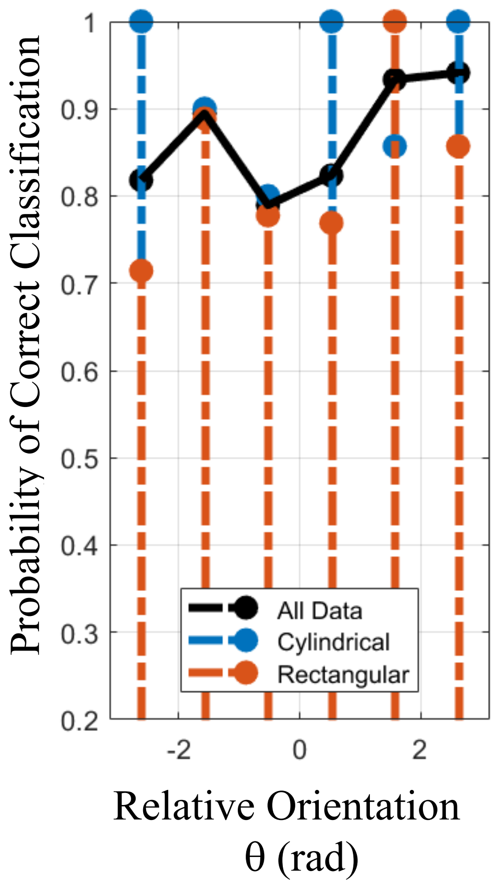
  - Instantiations can also be calculated:

$$P(\hat{y}_{pri} \mid \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}) \text{ and } P(\hat{y}_{pri}, \hat{y}_{post} \mid \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}, \hat{\mathbf{r}}_{post})$$

Results

Influence of UUV position on Classification



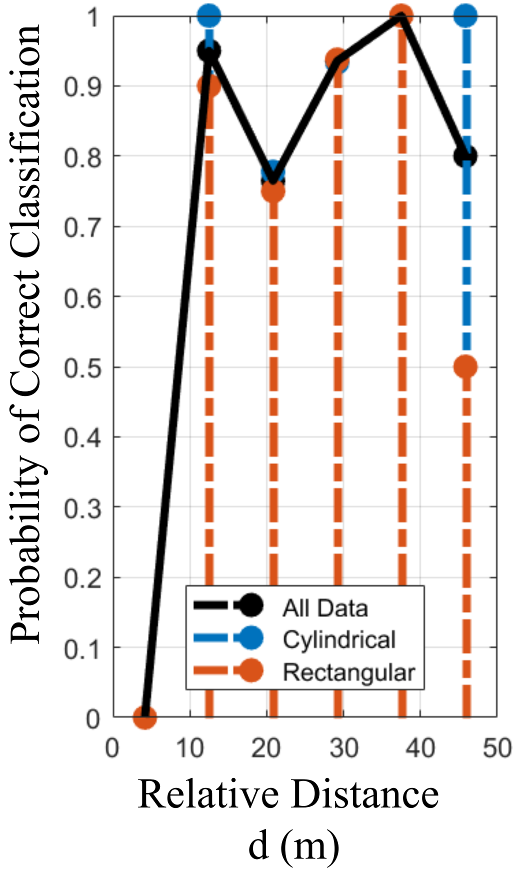


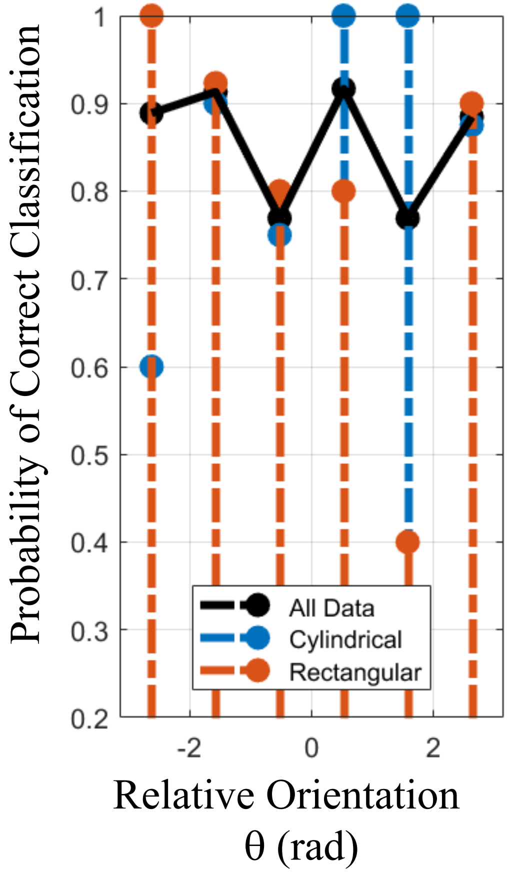
Dataset A: Low Frequency

Probability of correct classification (total):

- Cylindrical object 91.11%
- Rectangular object 83.02%

Influence of UUV position on Classification



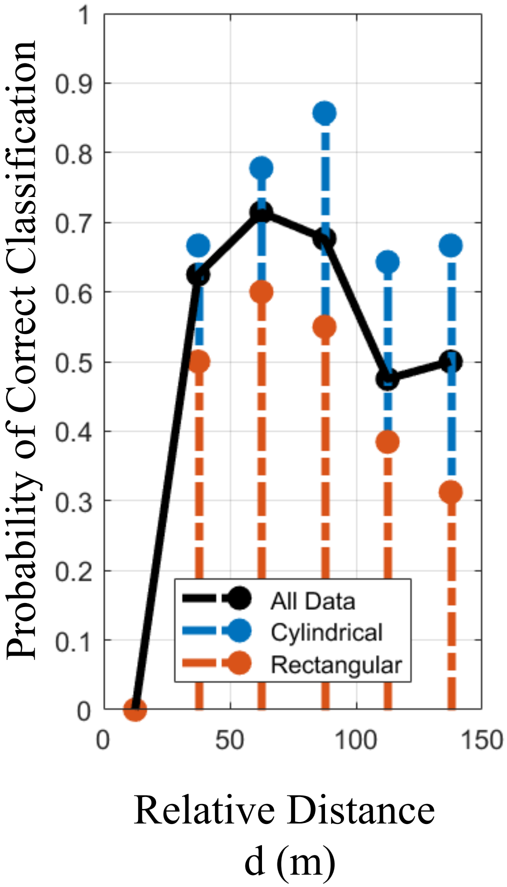


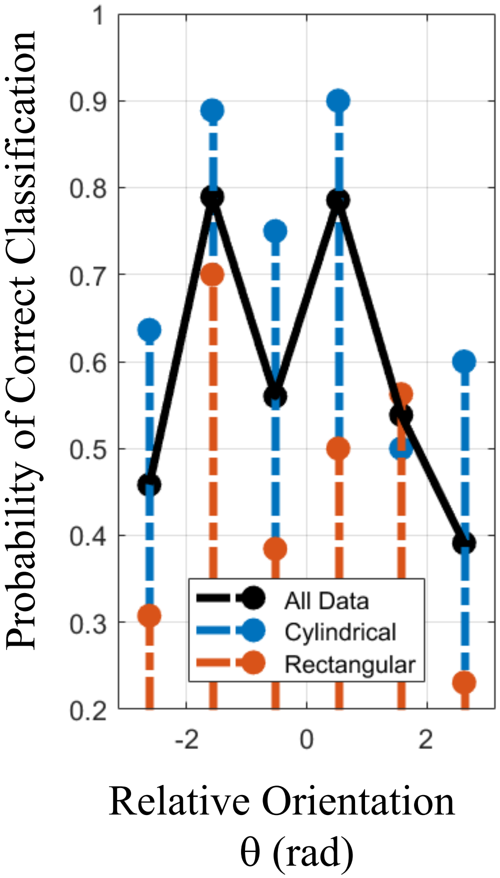
Dataset A: High Frequency

Probability of correct classification (total):

- Cylindrical object 87.04%
- Rectangular object 86.27%

Influence of UUV position on Classification



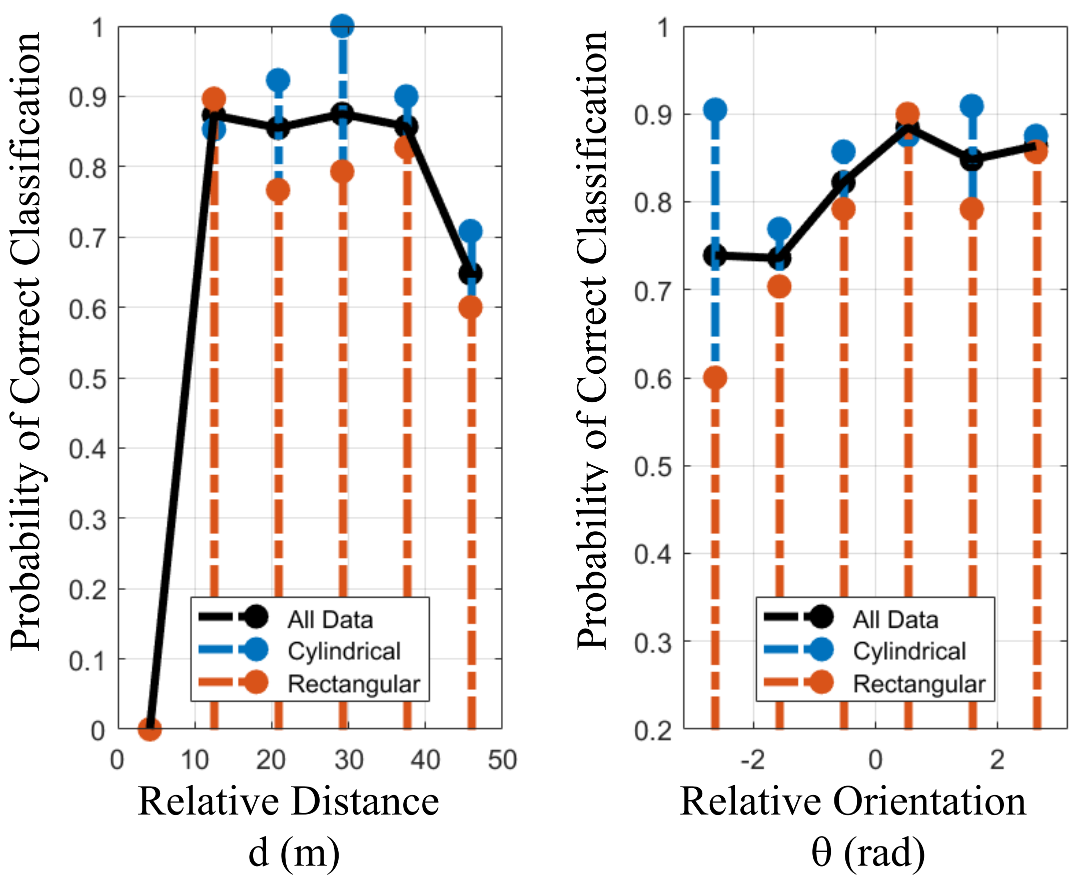


Dataset B: Low Frequency

Probability of correct classification (total):

- Cylindrical object 70.97%
- Rectangular object 43.48%

Influence of UUV position on Classification



Dataset B: High Frequency

Probability of correct classification (total):

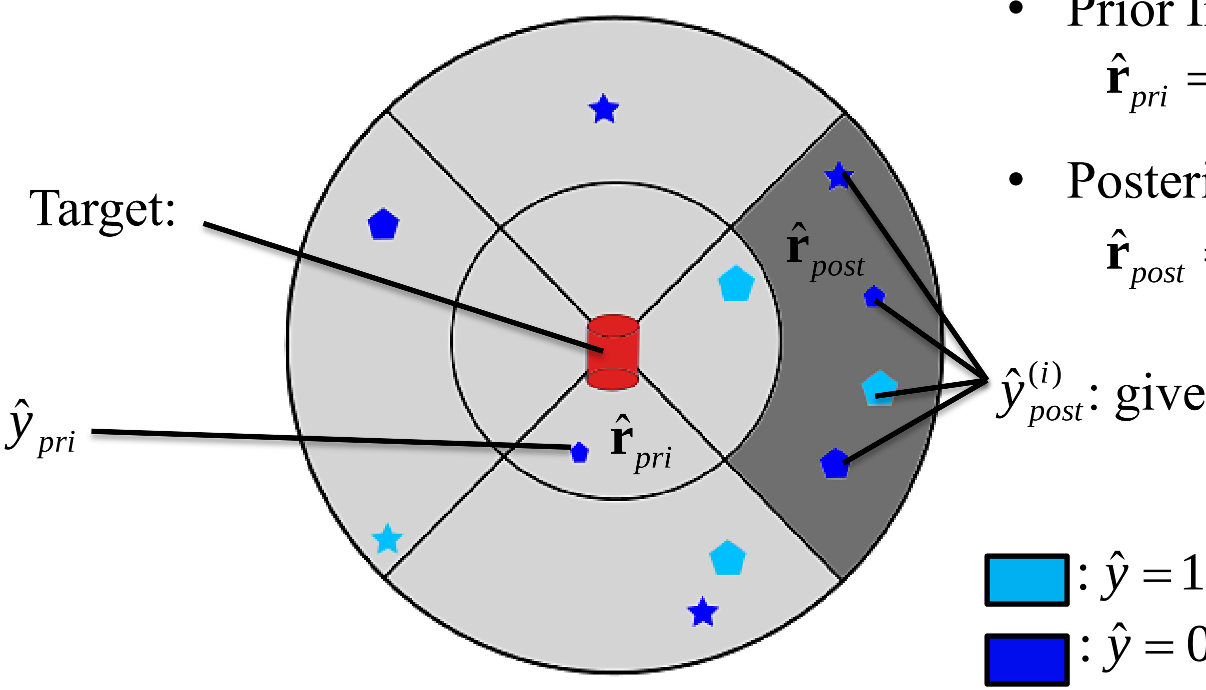
- Rectangular object 86.23%
- Cylindrical object 77.03%

Information Value Function Learning

- Assume information gain can be modeled as Expected Entropy Reduction (EER)

$$EER(\hat{y}_{pri}, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}, \hat{\mathbf{r}}_{post}) = H_{pri}(\hat{y}_{pri}, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}) - \frac{1}{N}\sum_{i=1}^{N} H_{post}(\hat{y}_{pri}, \hat{y}_{post}^{(i)}, \hat{\mathbf{F}}, \hat{\mathbf{r}}_{pri}, \hat{\mathbf{r}}_{post})$$

- Example:



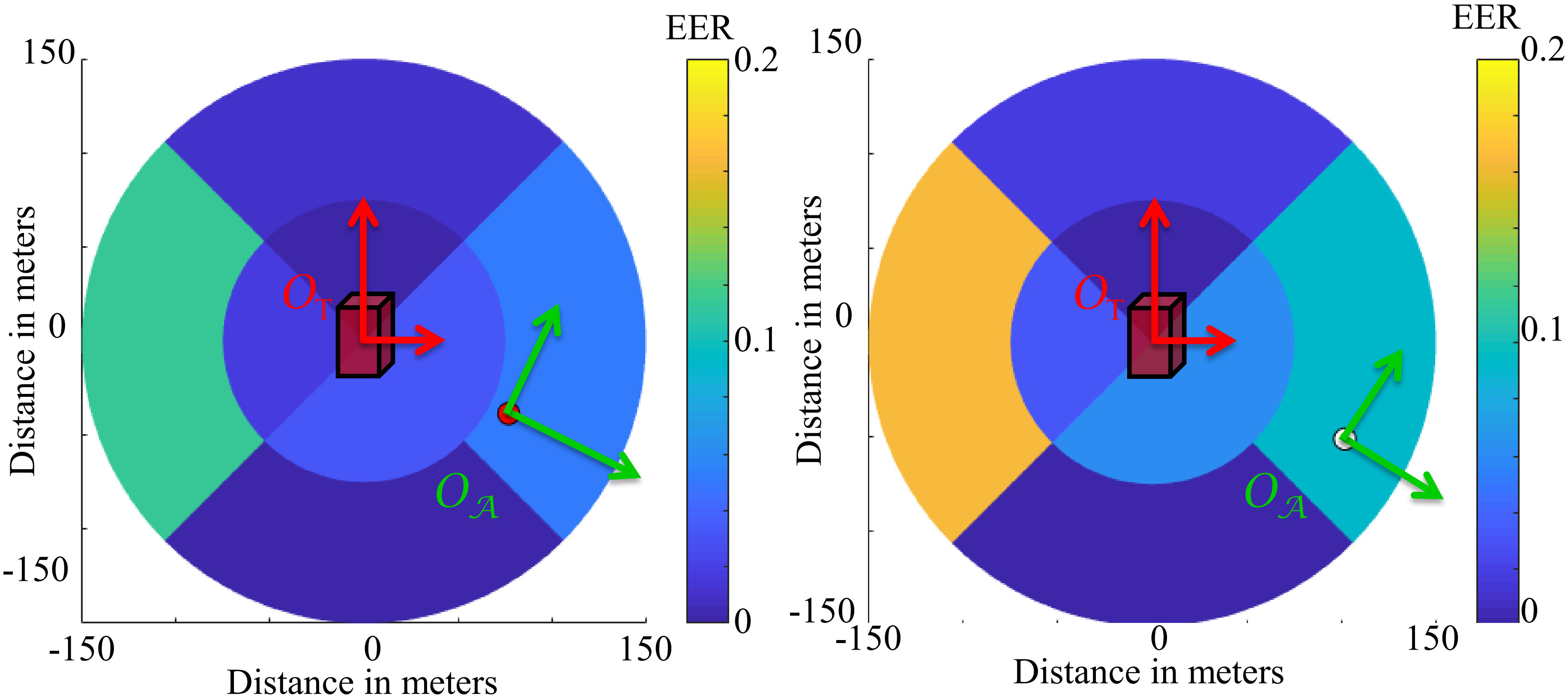
- Prior Image location
  $\hat{\mathbf{r}}_{pri} = \{[0,75),[5\pi/4, 7\pi/4,)\}$
- Posterior Image location
  $\hat{\mathbf{r}}_{post} = \{[75,150],[-\pi/4, \pi/4,)\}$

Discretization:

r = [d, θ]; d = {[0,75), [75,150]}; θ = {[-π/4, π /4), [π /4, 3 π /4), [3 π /4, 5 π /4), [5 π /4, 7 π /4]}

Results: Learned Information

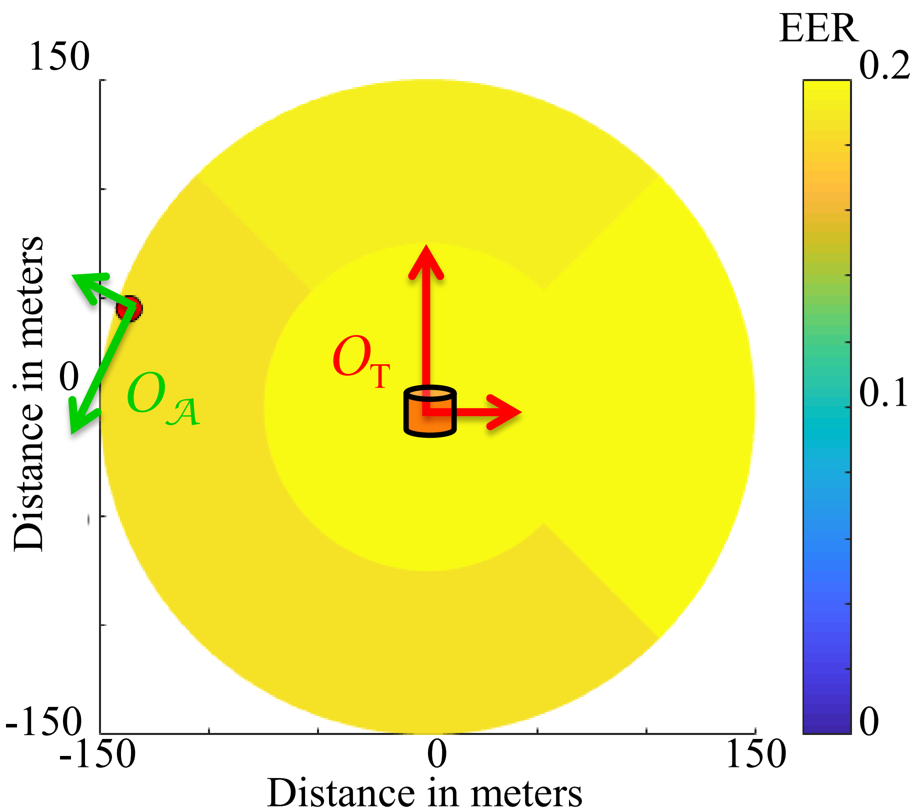
- **Target #2:** $y_T = 1$, s = Rectangular, $l = 1.25$m, $w = 0.61$m, $h = 0.47$ m
- Correct classification $(\hat{y}_{pri} = 1)$
- Incorrect classification $(\hat{y}_{pri} = 0)$

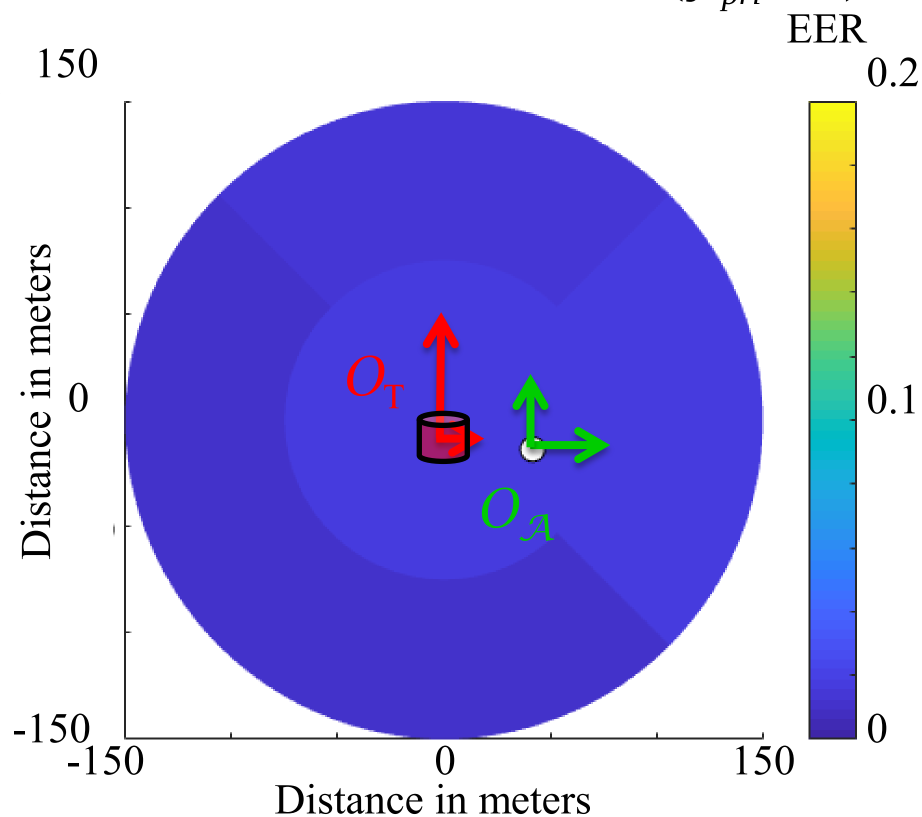


$\mathcal{F}_T$: Target Frame $\mathcal{F}_{\mathcal{A}}$: Vehicle Frame

Results: Learned Information Gain

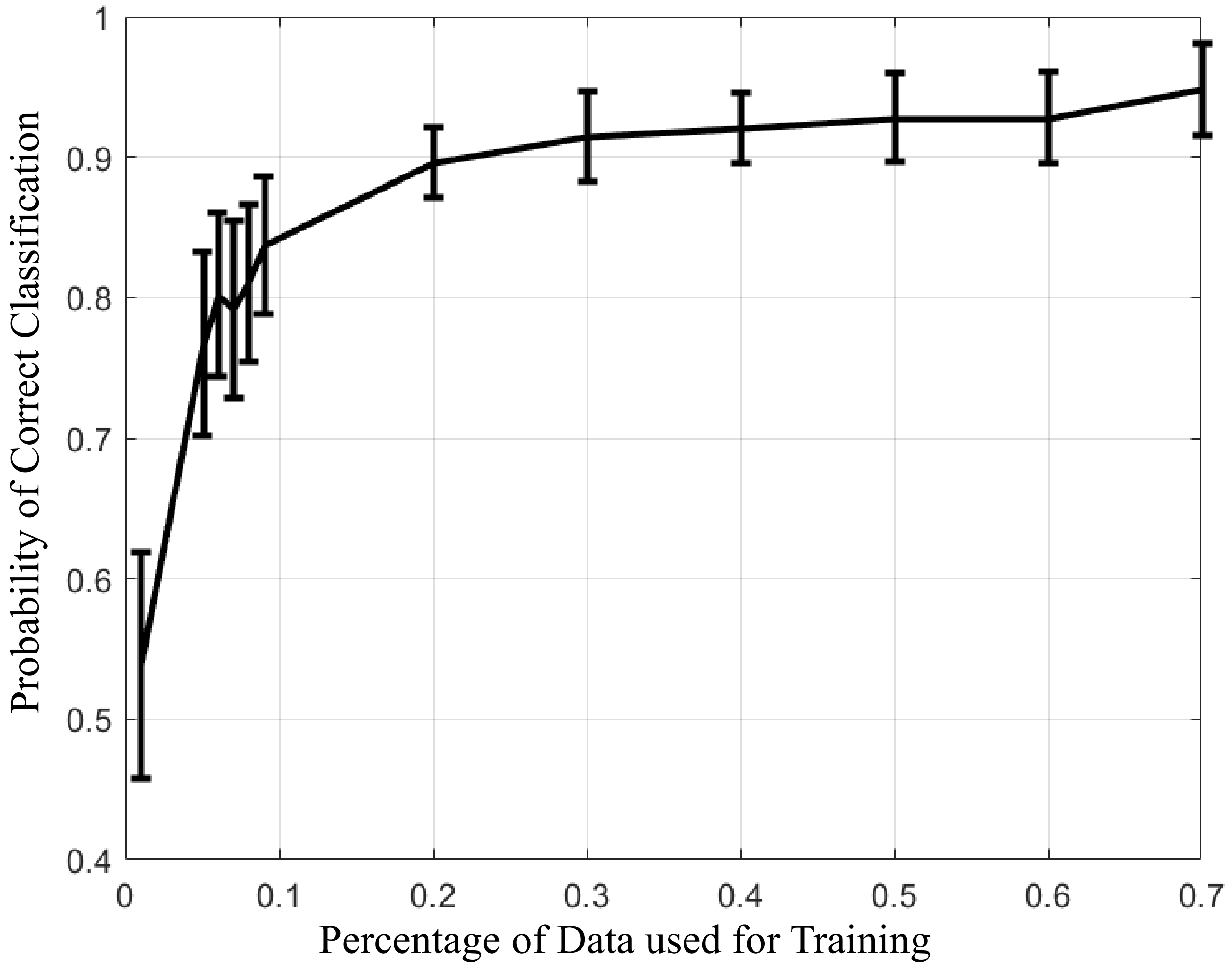
- **Target #230:** $y_T = 0$, s = Circular, l = 0.05m, w = 0.20 m, h = 0.27 m
- Correct classification $(\hat{y}_{pri} = 0)$
- Incorrect classification $(\hat{y}_{pri} = 1)$



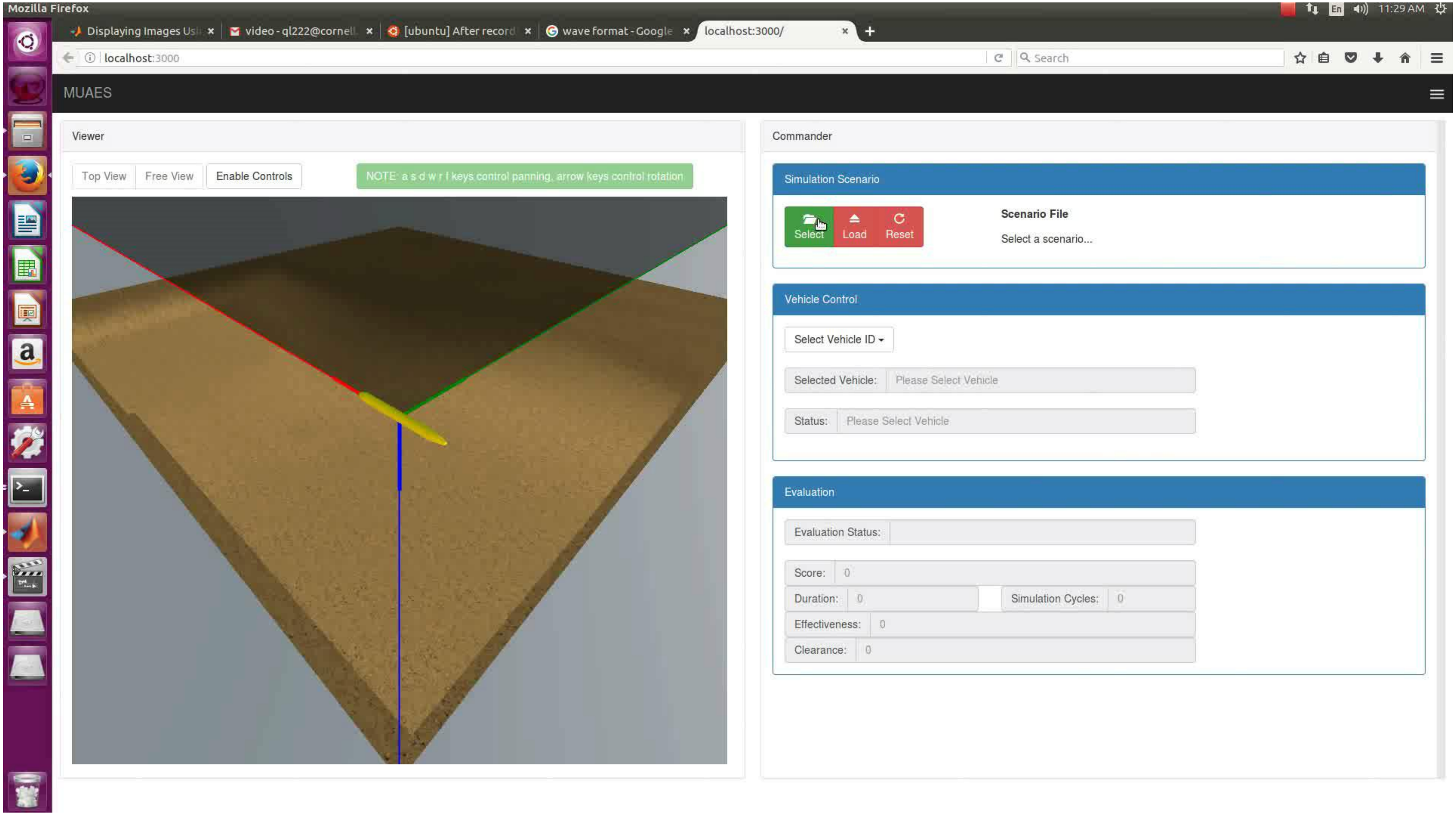


$\mathcal{F}_T$: Target Frame $\mathcal{F}_{\mathcal{A}}$: Vehicle Frame

CNN-SVM Learning with limited data



MCM Unmanned Autonomy Evaluation Simulator (MUAES)



Conclusions

MCM information-driven path planning and control:

- Directional information gain
- Automatic CNN-SVM target recognition and detection
- New reference frames and problem formulation
- Bayesian model of UUV-sonar sensing problem
- Information gain function learning
- MUAES visualization and testing

Ongoing and Future Work:

- Information-driven MCM acquisition and classification
- MCM path planning and control
- Validation in MUAES-Sonar Imaging

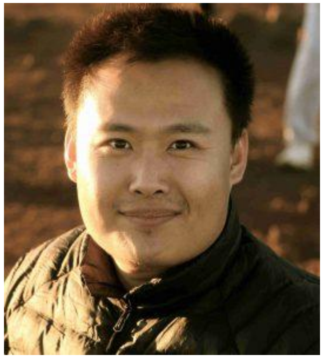
Acknowledgments

Collaborators:

- **Jason Isaacs**, Ph.D.
  Naval Surface Warfare Center
  Panama City, FL
- **Drew Lucas**, Ph.D.
  Naval Surface Warfare Center
  Panama City, FL
- **Matt Bays**, Ph.D.
  Naval Surface Warfare Center
  Panama City, FL

- **Pingping Zhu**, Ph.D.
  LISC, MEMS, Duke

Pingping Zhu



- **Bo Fu**, Ph.D.
  LISC, MAE, Cornell

This work was funded by ONR Code 321OE, MCM Autonomy Program