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

The sensor FOV, denoted by $S \subset \mathbb{R}^3$, 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.
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 $P$:
  $$S = \{P, V_1, ..., V_N\}, \text{ where } V_1, ..., V_N \text{ are the obstacle vertices.}$$
- Compute the convex hull of $S$:
  $$CS = \text{conv}(S) = \left\{ \sum_{i=1}^{\lvert S \rvert} k_i x_i \mid (\forall i : k_i \geq 0 \land \sum_{i=1}^{\lvert S \rvert} k_i = 1) \right\}$$
- 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 $k_i$, $k_j$
• Find boundary unit vectors $k_1$ and $k_2$:
  \[
  k_1 = \frac{PV_p}{\|PV_p\|} \quad k_2 = \frac{PV_q}{\|PV_q\|}
  \]
• Compute obstacle cone $\mathcal{K}$:
  \[
  \mathcal{K} = cone(\hat{k}_1, \hat{k}_2) = \{x \mid x = r + c_1\hat{k}_1 + c_2\hat{k}_2, c_1, c_2 \geq 0\}
  \]
  \[
  \mathcal{D} = (S \cap \mathcal{K}) \backslash (CS)
  \]

• Visibility region for multiple targets ($i$) and obstacles ($j$):
  \[
  R_{(i)} = CS_{(i)} \backslash B_{(i)}; \quad D_{(i)} = (S \cup \mathcal{K}_{(i)}) \backslash CS_{(i)}
  \]
  \[
  R_S = \bigcup_i R_{S(i)} = \bigcup_i \{ (S \backslash \mathcal{K}_{(i)}) \cup ((S \cup \mathcal{K}_{(i)}) \backslash (D_{(i)} \cup B_{(i)})) \} 
  \]
Example: Art Gallery Problem

Sensor at Single Location

Sensor at Multiple Location

Distance in meters

Distance in meters

Distance in meters

Distance in meters

\(\mathcal{B}\)

\(\mathcal{T}\)

\(\mathcal{D}\)

\(\mathcal{K}\)

\(\mathcal{R}_s\)
UUV-Sonar Directional Information Gain
UUV Kinematics Frames of Reference

Assume:
\[ v = constant \]

UUV States:
\[ \mathbf{x}(t) = [x_{uuv}(t), y_{uuv}(t), \theta_{uuv}(t)]^T \]

UUV Dynamics:
\[ \dot{x}_{uuv} = v \cdot \cos(\theta_{uuv}) \]
\[ \dot{y}_{uuv} = v \cdot \sin(\theta_{uuv}) \]
Target: $T$

UUV Frames Relative to Sonar Image

- UUV at $x(t_2)$, $t_2 > t_1$
- UUV at $x(t_1)$

Distance in meters

Vehicle Frame

$\mathcal{F}_A = \mathcal{F}_A(x(t))$

Sonar FOV

$\text{FOV} = \text{FOV}(x(t))$

UUV Heading Direction

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

$F_A$: Vehicle Frame  $F_T$: Image Frame

$\mathcal{T} = \mathcal{T}(x)$: Target Image Region
Image Frame and Target Frame

\[ F_W \]: Target Frame
\[ F_T \]: Image Frame

Aspect Angle: \( \theta = \theta_{uuv} - \theta_T \)
Image and Target Distance Vectors

\[ \mathbf{r} = \mathbf{r}_i + \Delta \mathbf{r} \]

\[ = [(x_T - x), (y_T - y)]^T \]
UUV-Sonar Feature Extraction
• **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
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.
Legend:
- TP: True Positive
- FP: False Positive
- TN: True Negative
- FN: False Negative
- $G_i$: 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%

CNN Performance: True Positive Rate

Legend:
- TP: True Positive
- FP: False Positive
- TN: True Negative
- FN: False Negative
- $G_i$: Image Group $i$

True Positive Rate (TPR):

\[
TPR = \frac{n_{TP}}{n_{TP} + n_{FP}}
\]

Probability of detection (Matched Filter performance): 88.31%

**CNN Performance: False Alarms**

Legend:
- TP: True Positive
- FP: False Positive
- TN: True Negative
- FN: False Negative
- $G_i$: Image Group $i$

**False Positive Rate (FPR):**

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

Probability of detection (Matched Filter performance): 88.31%

UUV-Sonar Information Value
UUV-Sonar Information Value

Minefield:
520 Targets, 20 Mine objects

*Courtesy of Jason Isaacs, NSWC Panama City, FL
### Key Variable Description

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_{\text{uuv}}$</td>
<td>Vehicle Position</td>
<td>$X_T$</td>
<td>Target Position</td>
</tr>
<tr>
<td>$\theta_{\text{uuv}}$</td>
<td>Vehicle Orientation</td>
<td>$\theta_T$</td>
<td>Target Orientation</td>
</tr>
<tr>
<td>$d$</td>
<td>Relative Distance</td>
<td>$\theta$</td>
<td>Relative Orientation</td>
</tr>
<tr>
<td>$Y_T$</td>
<td>True Classification</td>
<td>$\hat{Y}$</td>
<td>Estimated Classification</td>
</tr>
<tr>
<td>$S$</td>
<td>Object Shape</td>
<td>$L$</td>
<td>Object Length</td>
</tr>
<tr>
<td>$W$</td>
<td>Object Width</td>
<td>$H$</td>
<td>Object Height</td>
</tr>
<tr>
<td>$I_s$</td>
<td>Segmented Image</td>
<td>$z$</td>
<td>CNN output Features</td>
</tr>
<tr>
<td>$M$</td>
<td>Sensor Mode</td>
<td>$E$</td>
<td>Noise level</td>
</tr>
</tbody>
</table>

Assume Constant

$\mathbf{r} = [d, \theta]$
Key Variables and Causal Relationships
Information Value Function Learning

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

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

• Example:

  Prior Image location
  \( \hat{r}_{pri} = \{[0, 75), [5\pi/4, 7\pi/4), \} \)

  Posterior Image location
  \( \hat{r}_{post} = \{[75, 150], [-\pi/4, \pi/4), \} \)

\( \hat{y}^{(i)}_{post} \): given in training data

Discretization:
\( r = [d, \theta]; d = \{[0, 75), [75, 150]\}; \theta = \{[-\pi/4, \pi/4), [\pi/4, 3\pi/4), [3\pi/4, 5\pi/4), [5\pi/4, 7\pi/4]\} \)
Continuous probabilities (PDF’s) are difficult to learn, random variables are first discretized.

Conditional probability table (CPT) is estimated from dataset:

\[ P(\hat{Y} | Y_T, S, L, W, H, r) \]

Entropy \( H \) is calculated based on estimated conditional probability

Information Value Function (EER) is learned from the dataset images and Features.
Bayesian Network Model

- 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: \( s_i, l_i, w_i, h_i, d_i, \theta_i, y_{T,i}, \) and \( \hat{y} \)

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

\[
\begin{align*}
S &= \{\text{Cylindrical, Rectangular}\}; \\
V &= (LWH)^{1/3} = \{[0, 0.14), [0.14, 0.3), [0.3, 1.1), [1.1, 1.7]\} \\
r &= [d, \square]; d = \{[0,75), [75,150]\}; \square = \{[-\pi/4, -\pi/4), [-\pi/4, 3\pi/4), [3\pi/4, 5\pi/4), [5\pi/4, 7\pi/4]\}
\end{align*}
\]
Bayesian Network Model,
Features \( F = \{S, H, W, L\} \)

- Bayesian parameter estimation:
  - Parameters are random variables \( \eta_{ijk} \)
  - Multinomial likelihood of training data set:
    \[
    P(D | \eta) = \prod_i \prod_j \prod_k n_{ikj} \eta_{ikj}^{N_{ikj}}
    \]
  - Dirichlet Prior of the parameter:
    \[
    P(\eta | \alpha) = \frac{\Gamma(\sum_k \alpha_{ijk}) \prod_i r_i (\alpha_{ijk} - 1)}{\prod_k \Gamma(\alpha_{ijk}) \prod_{k=1}^{r_i} \eta_{ikj}^{(\alpha_{ijk} - 1)}} = \text{Dirichlet}(\eta_{ij}; \alpha_{ij1}, \ldots, \alpha_{ijr_i})
    \]
  - Hyper parameter: \( \alpha = [\alpha_{ij1}, \ldots, \alpha_{ijr_i}] \)
  - Posterior of the parameter:
    \[
    P(\eta | D, \alpha) \propto \prod_{ijk} \eta_{ikj}^{(N_{ijk} + \alpha_{ijk} - 1)}
    \]
  - MAP estimate:
    \[
    \hat{\eta}_{ijk}^{MAP} = \frac{N_{ijk} + \alpha_{ijk}}{\sum_j (N_{ijk} + \alpha_{ijk})}
    \]
Posterior Probability Learning

- Posterior probabilities calculated via Bayes rule:
  - Given image $I_{\text{pri}}$
    \[
    P(Y_T = 1 \mid \hat{y}_{\text{pri}}, \hat{F}, \hat{r}_{\text{pri}}) = \frac{P(Y_T = 1 \mid \hat{F}, \hat{r}_{\text{pri}})P(\hat{y}_{\text{pri}} \mid Y_T = 1, \hat{F}, \hat{r}_{\text{pri}})}{P(\hat{y}_{\text{pri}} \mid \hat{F}, \hat{r}_{\text{pri}})}
    \]
  - Given image $I_{\text{pri}}$ and $I_{\text{post}}$
    \[
    P(Y_T = 1 \mid \hat{y}_{\text{pri}}, \hat{y}_{\text{post}}, \hat{F}, \hat{r}_{\text{pri}}, \hat{r}_{\text{post}}) = \frac{P(Y_T = 1 \mid \hat{F}, \hat{r}_{\text{pri}}, \hat{r}_{\text{post}})P(\hat{y}_{\text{pri}} \mid Y_T = 1, \hat{F}, \hat{r}_{\text{pri}})P(\hat{y}_{\text{post}} \mid Y_T = 1, \hat{F}, \hat{r}_{\text{post}})}{P(\hat{y}_{\text{pri}}, \hat{y}_{\text{post}} \mid \hat{F}, \hat{r}_{\text{pri}}, \hat{r}_{\text{post}})}
    \]
  - Instantiations can also be calculated:
    \[
    P(\hat{y}_{\text{pri}} \mid \hat{F}, \hat{r}_{\text{pri}}) \text{ and } P(\hat{y}_{\text{pri}}, \hat{y}_{\text{post}} \mid \hat{F}, \hat{r}_{\text{pri}}, \hat{r}_{\text{post}})
    \]
Results
Influence of UUV position on Classification

Relative Distance $d$ (m)

Probability of Correct Classification

Relative Orientation $\theta$ (rad)

Dataset A: Low Frequency

Probability of correct classification (total):
- Cylindrical object 91.11%
- Rectangular object 83.02%
Influence of UUV position on Classification

Probability of correct classification (total):
- Cylindrical object 87.04%
- Rectangular object 86.27%

Dataset A: High Frequency

Relative Distance $d$ (m)

Relative Orientation $\theta$ (rad)
Influence of UUV position on Classification

Relative Distance $d$ (m)

Probability of Correct Classification

Relative Orientation $\theta$ (rad)

Dataset B: Low Frequency

Probability of correct classification (total):
- Cylindrical object 70.97%
- Rectangular object 43.48%
Influence of UUV position on Classification

Probability of correct classification (total):
- Rectangular object 86.23%
- Cylindrical object 77.03%

**Dataset B: High Frequency**
Information Value Function Learning

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

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

• Example:

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

\(\hat{y}^{(i)}_{post}\): given in training data

**Discretization:**
\(r = [d, \theta]; d = \{[0,75), [75,150]\}; \theta = \{[-\pi/4, \pi/4), [\pi/4, 3\pi/4), [3\pi/4, 5\pi/4), [5\pi/4, 7\pi/4]\}\)
Results: Learned Information

• **Target #2**: $y_T = 1$, $s =$ Rectangular, $l = 1.25\text{m}$, $w = 0.61\text{m}$, $h = 0.47\text{m}$

• Correct classification ($\hat{y}_{pri} = 1$)

• Incorrect classification ($\hat{y}_{pri} = 0$)

$\mathcal{F}_T$: Target Frame \hspace{1cm} $\mathcal{F}_A$: Vehicle Frame
Results: Learned Information Gain

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

$\mathcal{F}_T$: Target Frame  \hspace{1cm} \mathcal{F}_A$: Vehicle Frame
CNN-SVM Learning with limited data

![Graph showing the probability of correct classification versus the percentage of data used for training. The graph illustrates a positive correlation between the amount of training data and the accuracy of classification.]
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

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

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