Decentralized and Autonomous Data Fusion Service (DADFS) for Heterogeneous Unmanned Vehicles (UVs)

April 24, 2015

Daniel H. Wagner Associates, Inc.

Dr. W. Reynolds Monach
Vice President

Dr. C. Allen Butler
President & CEO

JHU/APL

Mr. David Scheidt
Principal Investigator

Daniel H. Wagner Associates, Inc.

- Consulting services in Operations Research, Mathematics, and Software Development since 1963
- Technical staff of 25 includes 12 PhDs in mathematical sciences
- Offices in Exton PA (HQ), Hampton VA and Vienna VA
- Primary client base – DoD and financial community

Daniel H. Wagner Associates, Inc.

- Missile Defense Agency: Tracking, Registration, and Data Fusion
- Air Force: AWACS Multi-Sensor Fusion; Tracking ground targets
- Army: Combat Identification for IBCS
- Department of Homeland Security (DHS) : Field tested Data Fusion systems for the Mexican (ground) and Canadian (water) borders
- Navy: Torpedo Defense, Data Fusion, Mission Planning, Mine Warfare, Submarine Warfare, Unmanned Vehicles
- NASA: Random Number Generation on GPGPUs
- Financial Community: Retirement Spending Planner, Mean Variance Optimization Library, Statistical Arbitrage, Optimal Coupon selection, Portfolio Optimization

Decentralized and Autonomous Data Fusion Service (DADFS) for Heterogeneous Unmanned Vehicles (UVs)

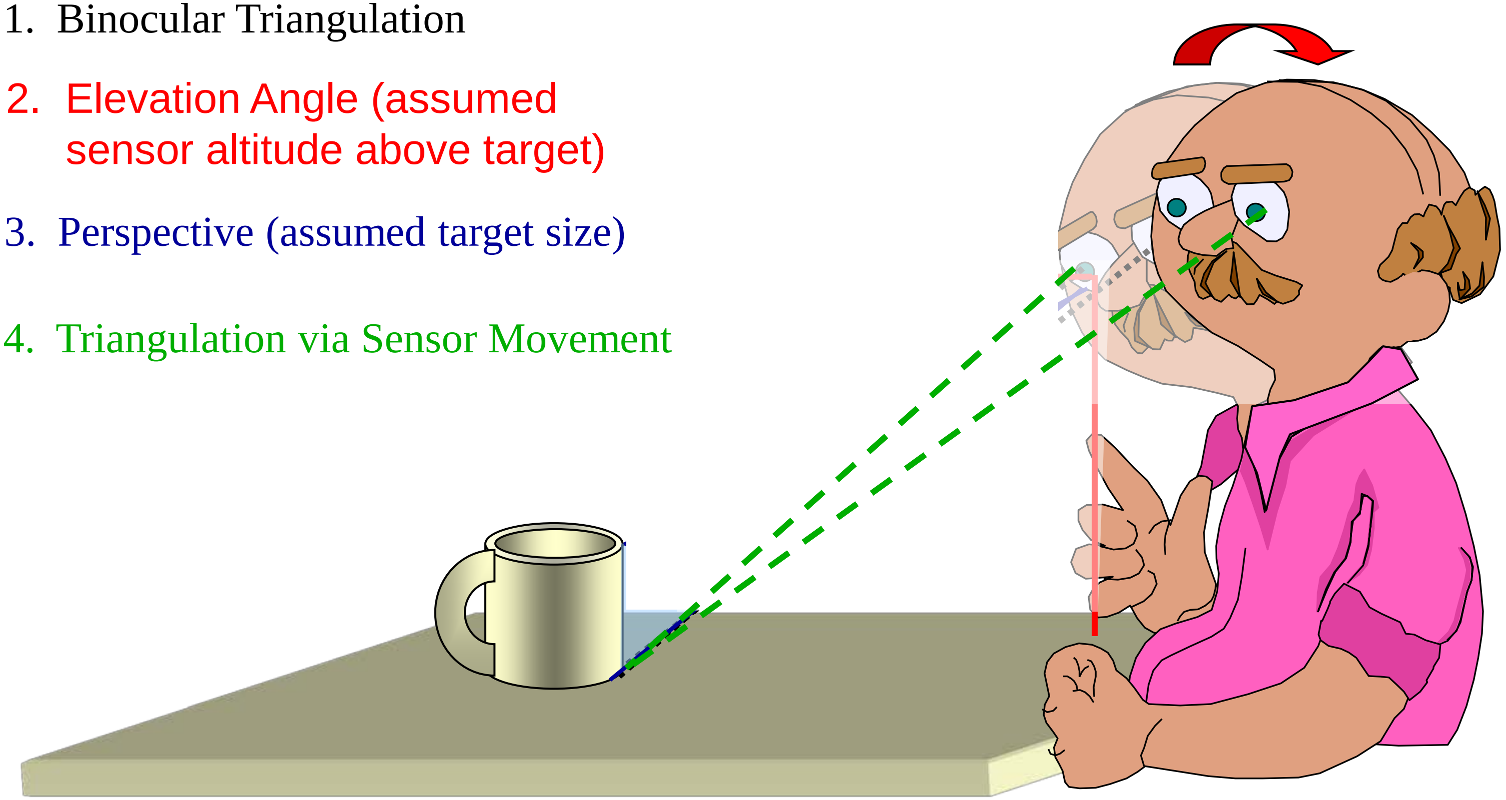
What is Data Fusion?

- Analogous to human cognitive process for drawing inferences about the real world from five senses: sight, sound, smell, taste, touch
- Different types of information comes from different sensors (e.g., acoustic, electromagnetic, infrared, electro-optical)
- Need to determine which sensor measurements derive from the same real world object

- **Ultimate Goal:**
  Achieve usable knowledge of your surroundings
  - Common Operational Picture (COP)
  - Situational Awareness (SA) picture

Sensor Fusion for Improved Estimation of Situation

1. Binocular Triangulation
2. Elevation Angle (assumed sensor altitude above target)
3. Perspective (assumed target size)
4. Triangulation via Sensor Movement



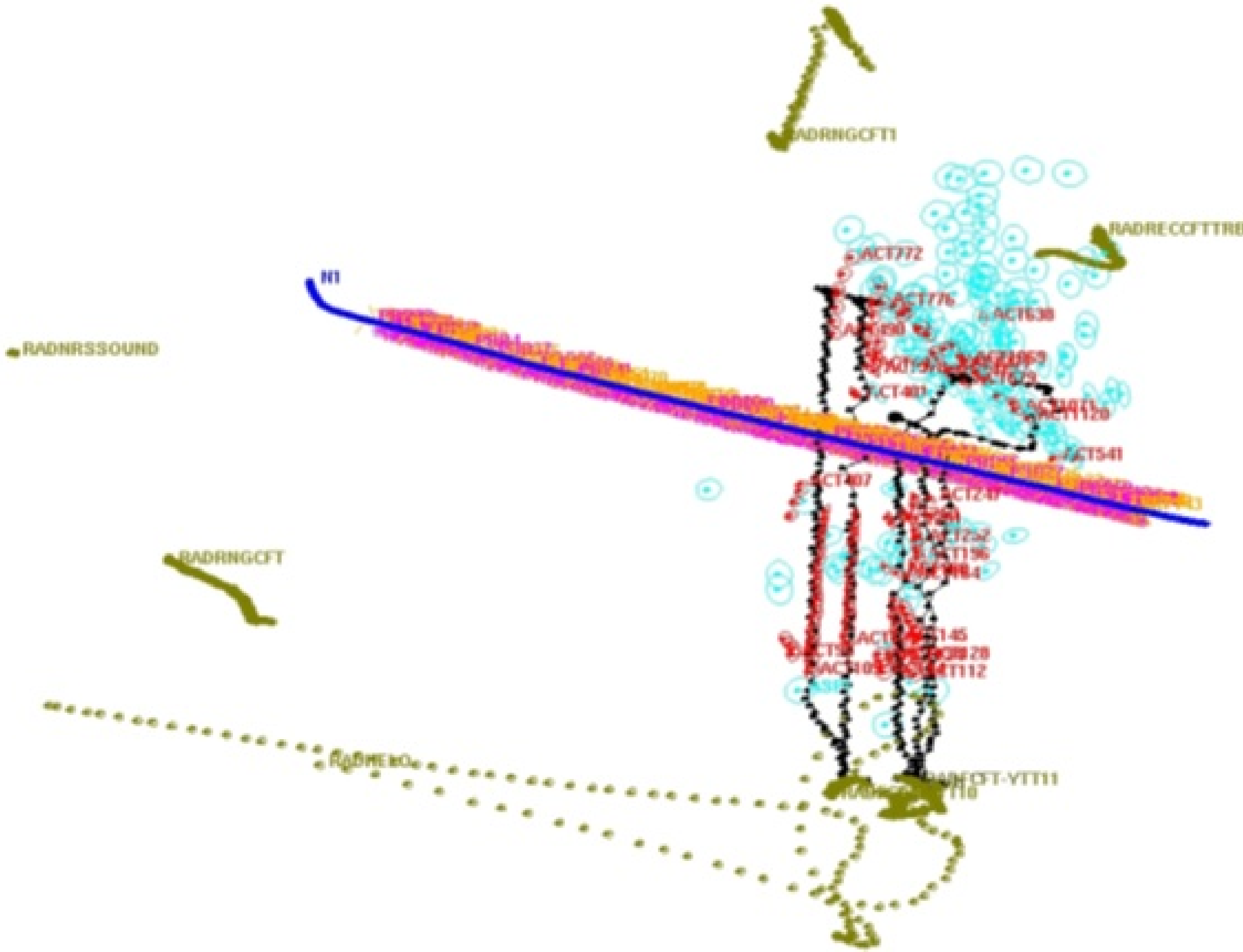
How do I know whether I can reach the cup?

Improve Estimation by Combining Multiple Types of Information

Source: Alan Steinberg

Daniel H. Wagner Associates

Lots of Objects; Lots of Measurements



Challenges to be Addressed

- Cleaning up clutter (radar returns from ocean waves)
- Associating Measurements
  - From each sensor (over time)
  - From all sensors on a single platform
  - Across platforms
- Dealing with uncertainty
  - Kinematic (position/velocity)
  - Non-kinematic (identity)
- Registration of sensors
  - Aboard a single platform
  - Between platforms
- Measuring Performance
  - Kinematic Accuracy
  - ID Estimation

And More Challenges

- Optimizing Information Flow between platforms
  - Intermittent Communications
  - Low Bandwidth
- Maintaining Accurate Picture; Common among platforms
  - Avoiding duplication of information (correlated information)
- Dealing with Nonlinearities
  - Nonlinear Platform Motion
  - Nonlinear Measurements

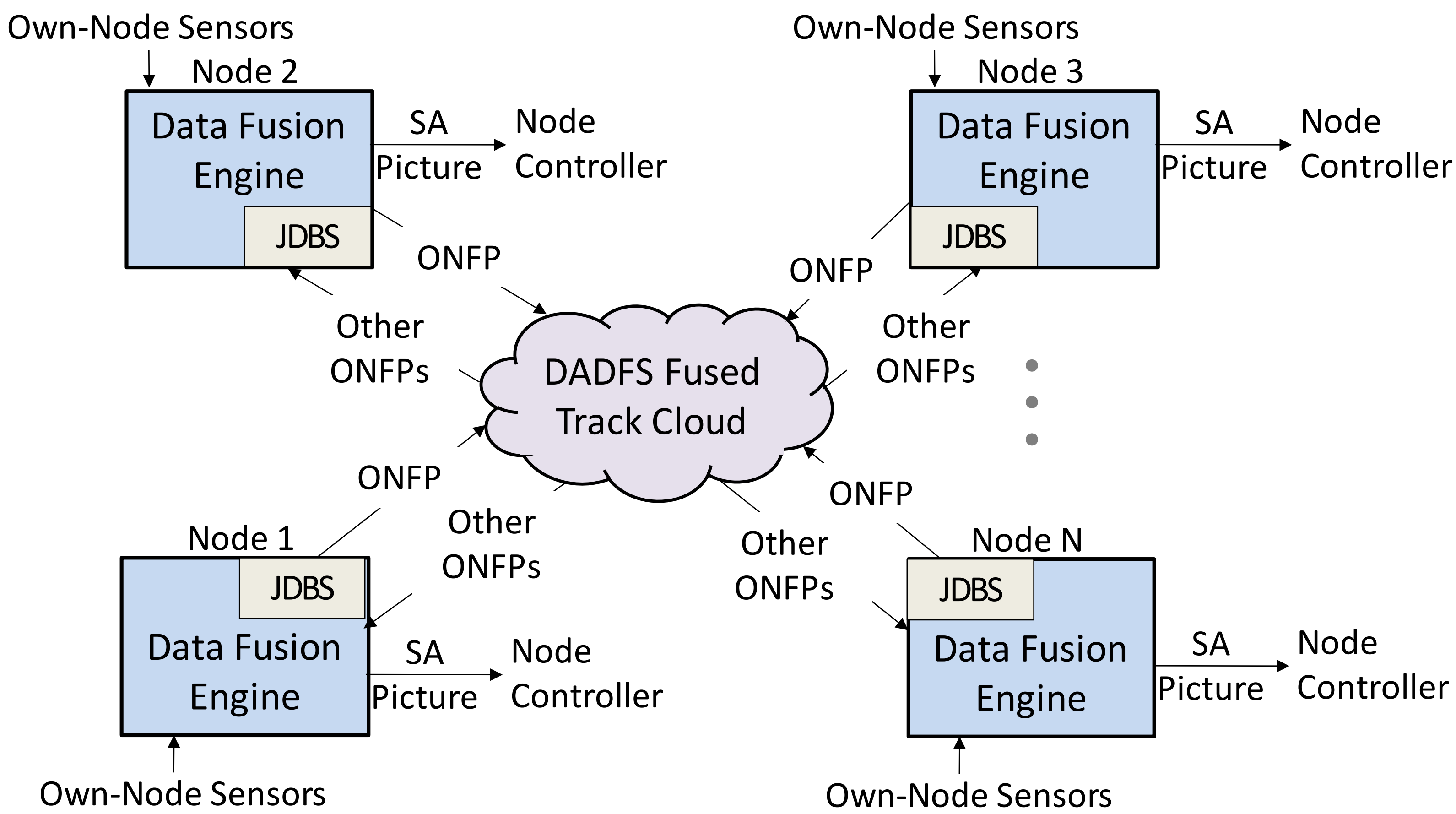
Multiple Analytical Techniques

- Stochastic Differential Equations (target motion)
- Kalman Filtering (and variants for nonlinearities)
- Measuring Distances between Distributions
- Bayesian Networks
  - Identity Estimation
  - Behavior Prediction
- Discrete Optimization (Hungarian Algorithm)
- Discrete and Continuous Simulation

Military Perspective

- **Commanders and Platforms need *more accurate, less cluttered, automatically generated Common Operational Picture (COP)/ Situational Awareness (SA) picture to*:**
  - Improve the quality of the decisions made
  - Improve the speed at which decisions are made
  - Reduce operator workload

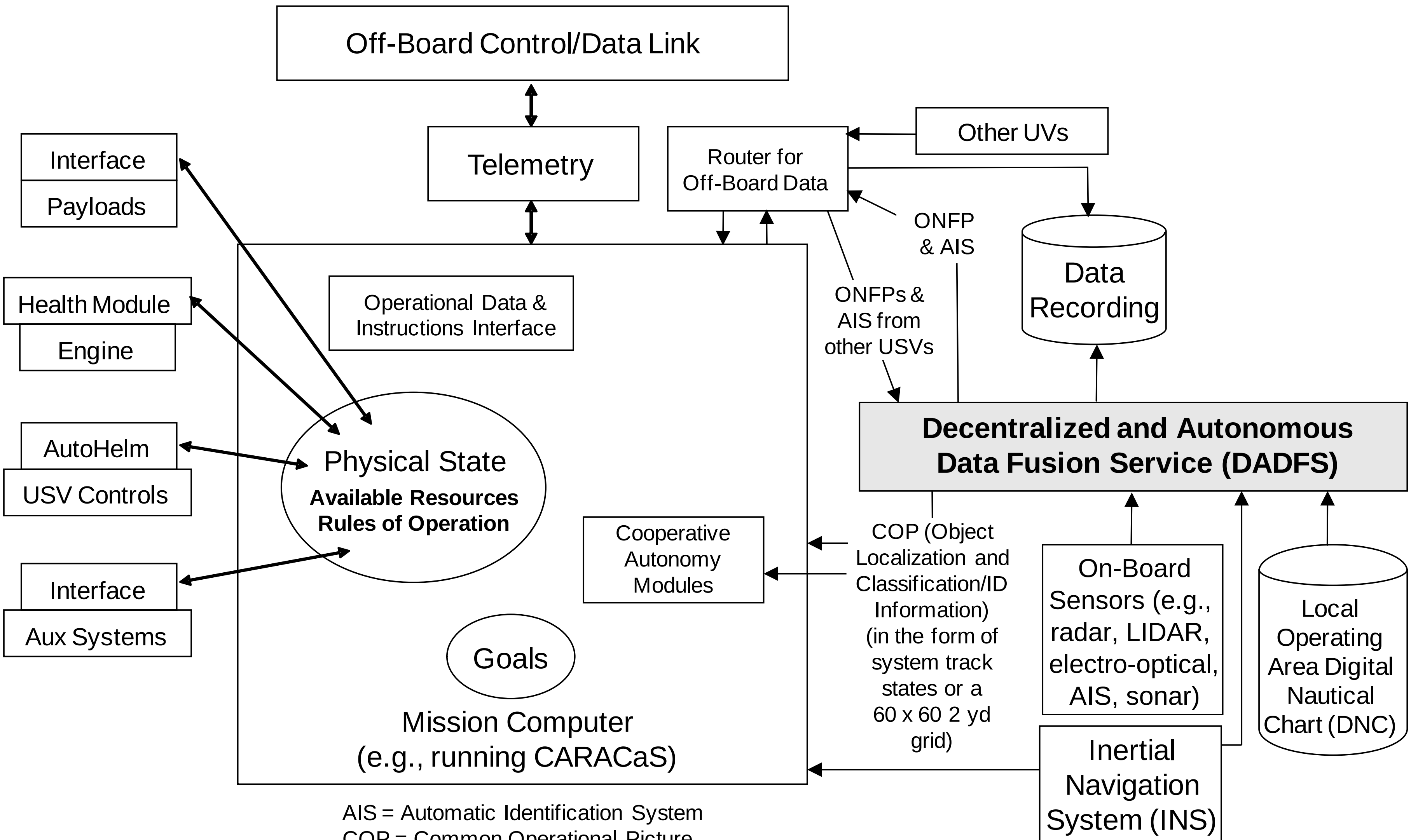
Decentralized and Autonomous Data Fusion Service (DADFS) for Heterogeneous Unmanned Vehicles (UVs)



DADFS = Decentralized and Autonomous Data Fusion Service
JDBS = JHU/APL Distributed Blackboard System
ONFP = Own-Node Fused Picture
SA = Situational Awareness

Daniel H. Wagner Associates

DADFS for Heterogeneous UVs Data and Process Flow on One AMN Type USV (including CARACaS and third party Cooperative Autonomy Modules)



AIS = Automatic Identification System
COP = Common Operational Picture
ONFP = Own-Node Fused Picture

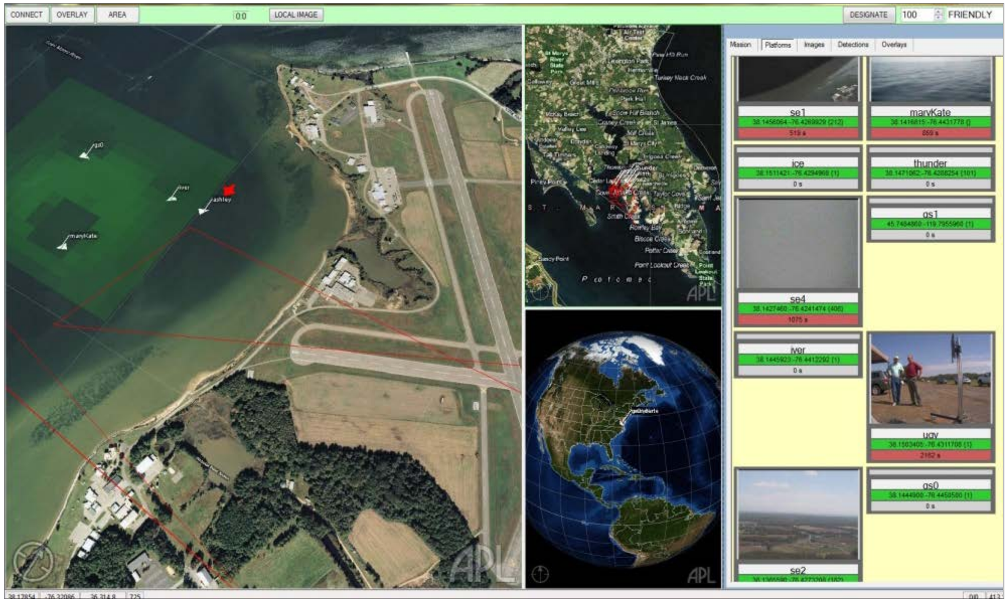
Distribution Statement A: Approved for public release; distribution is unlimited.

Daniel H. Wagner Associates

DADFS APL Unmanned Vehicles and Command and Control Stations

User interface is an application on an existing handheld device

Images are not stored "Google Earth" Data but Real-time and Near Real-time ISR Data including Automated Target Recognition and Multi-Sensor Data Fusion

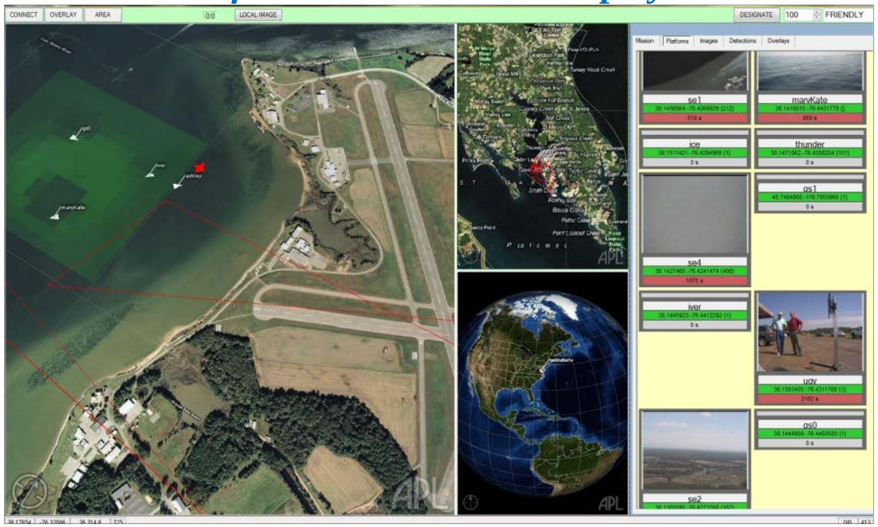


Daniel H. Wagner Associates

OPISR 11 Technical Accomplishments

Webster Field, 21 – 29 Sep 11

View from Distributed C2 Display



View from Boeing Scan Eagle

View from
UGS – via UAV Comms-Chain

View from
Segway UGV

View from
USV (SKB)

View from
Procerus Unicorn

Daniel H. Wagner Associates

OPISR 11 Technical Accomplishments

Webster Field, 21 – 29 Sep 11

View from Distributed C2 Display

RED DOT IS AN MLO (Mine Line Object) DETECTION FROM THE UUV RELAYED BY UAV LINK TO LAND C2

View from Boeing Scan Eagle

WE BELIEVE THIS IS OUR PLATFORM

OUR CHASE BOAT FROM THE SE – YOU CAN SEE OUR UUV CRATE IN BACK

ALL OF THIS IMAGERY WAS AVAILABLE ON BOTH LAND AND WATER C2 DISPLAYS DIRECTLY OR AS COMMS CHAINS FORMED

THIS SENSOR WAS DISTANT AND ONLY AVAILABLE VIA AIRBORNE LINK

View from UGS – via UAV Comms-Chain

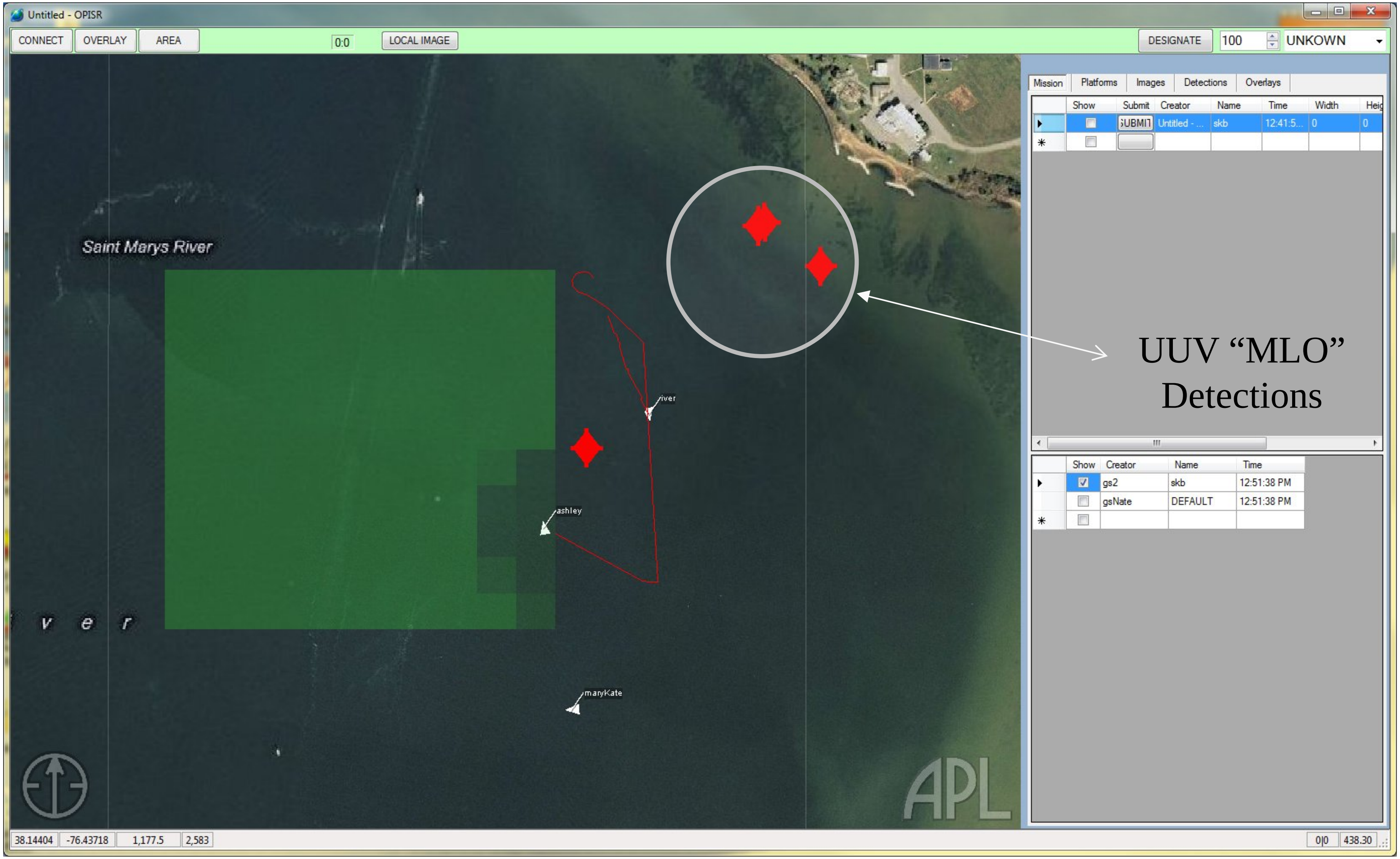
View from Segway UGV

View from USV (SKB)

View from Procerus Unicorn

UUV MLO
Contacts -> USV -> UAV -> User

Daniel H. Wagner Associates



Distribution Statement A: Approved for public release; distribution is unlimited.

Primary Mission Benefits

- **More accurate & less cluttered Common Operational Picture (COP)/ Situational Awareness (SA) picture**
  - Number of objects correct (reduced clutter)
  - Object position and classification correct
  - Provides ad hoc decentralized ISR for multiple users
- **Reduced risk**
  - Timely and accurate alerts concerning potential threats
- **Significantly reduced decision timelines**
- **Better utilization of scarce resources**
  - Operators
  - Platforms and sensors
  - Bandwidth
- **Highly automated**
  - Reduced operator workload

Higher Mission Success Rate Using Fewer Resources

Show SWARM Video

Daniel H. Wagner Associates

Multiple Analytical Techniques

- Stochastic Differential Equations (target motion)
- Kalman Filtering (and variants for nonlinearities)
- Measuring Distances between Distributions
- Bayesian Networks
  - Identity Estimation
  - Behavior Prediction
- Discrete Optimization (Hungarian Algorithm)
- Discrete and Continuous Simulation

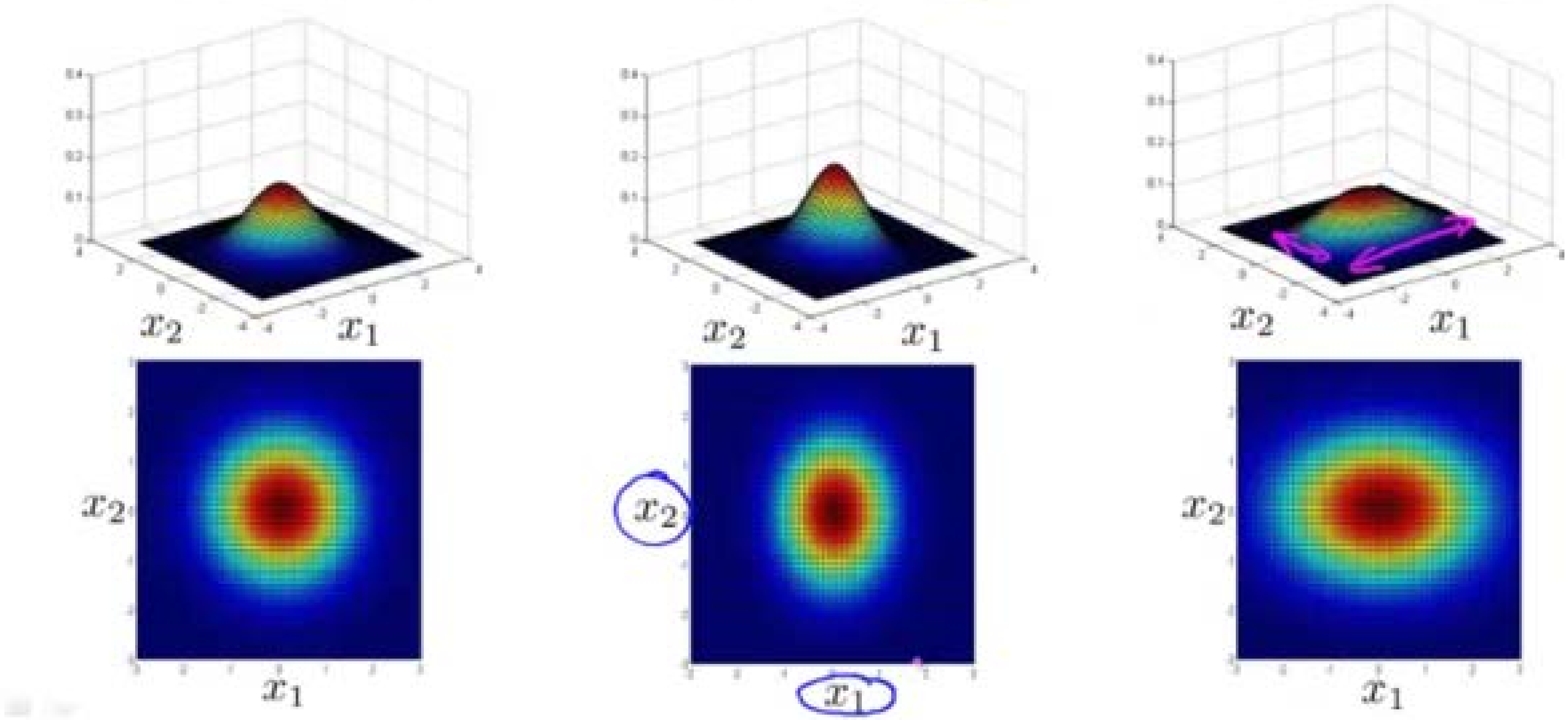
Stochastic Differential Equation

A commonly used SDE in target tracking is the Integrated Ornstein Uhlenbeck (IOU) Motion Model

$$dX = Vdt$$
$$dV = -\beta V dt + \sigma \, dW_t$$

Where X is 1, 2 or 3-D position, V is the corresponding velocity and dW_t is the Weiner differential.

$$\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Sigma = \begin{bmatrix} 0.6 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$



Daniel H. Wagner Associates

Multiple Analytical Techniques

- Stochastic Differential Equations (target motion)
- Kalman Filtering (and variants for nonlinearities)
- Measuring Distances between Distributions
- Bayesian Networks
  - Identity Estimation
  - Behavior Prediction
- Discrete Optimization (Hungarian Algorithm)
- Discrete and Continuous Simulation

Data Association – Problem Definition

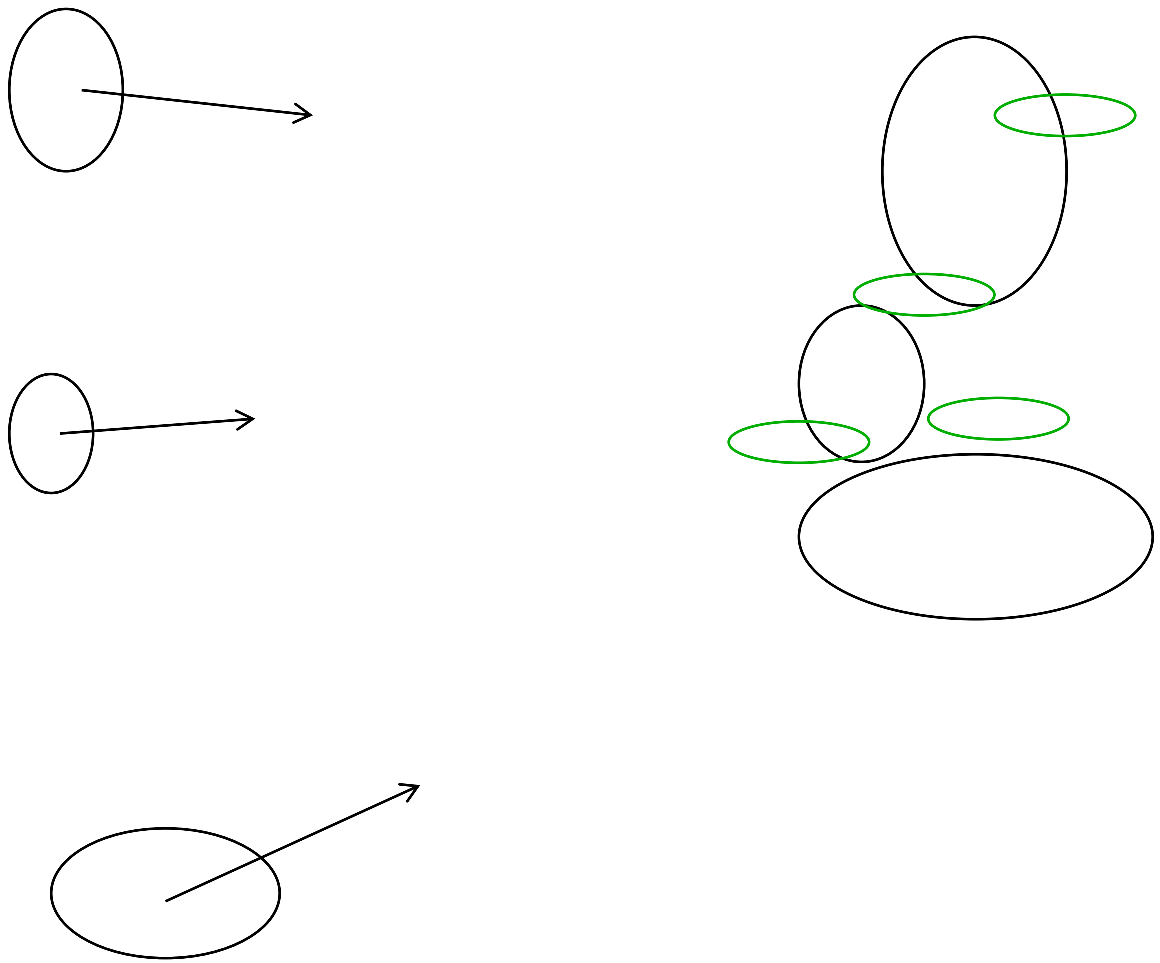
- Assume database has a set of tracks
- Ready to process new set of observations (e.g., one radar scan, one GMTI Frame)
- Which observations are from same target?
- Which are clutter/false alarms/noise?

Daniel H. Wagner Associates

Data Association – Example

Tracks at time T

Predicted Locations



Four New Observations

Data Association
The Assignment Problem

- Classical assignment problem
  - Workers to jobs, frequencies to telecommunication firms, etc.
- Assigning Observations to Tracks
- Entries in assignment matrix are “distances” d_{ij}

	obs 1	obs 2	obs 3	obs 4
Track 1	12	x	7	x
Track 2	x	9	8	12
Track 3	8	6	x	x

Data Association
The Assignment Problem

- Optimal Solution: Set of pairings (i,j) such that Σd_{ij} is minimized
- Suboptimal Solution: Nearest Neighbor Association
  - Find minimum distance observation-to-track pair and make the indicated assignment (Greedy Algorithm)
- Optimal Solution Methods:

  Munkres Algorithm
  Ford-Fulkerson Algorithm
  Hungarian Algorithm
  Auction Algorithm

Data Association – PD & FAR

- Need to account for
  - (1) Tracks which may not have an observation (Probability of detection < 1)
  - (2) Observations not associating with any track (Probability of new track/false alarm)
- Add a “not detected” column to association matrix for (1)
- Add a “false alarm” row to association matrix for (2)
- Question: What do we use for a “distance” in the entries of the association matrix?

Daniel H. Wagner Associates

Data Association

Actual Association Cost (distance) used in most tracking applications

$$\text{cost} = \begin{cases} -\ln(\beta_{NT} + \beta_{FA}) & \text{observation is false alarm/new target} \\ \dfrac{d_{ij}^2}{2} - \ln\left[\dfrac{P_D}{(2\pi)^{M/2}\,|S|^{1/2}}\right] & \text{observation to track} \\ -\ln(1 - P_D) & \text{no observation to track} \end{cases}$$

$(d_{ij})^2$ = Mahalanobis distance between track i and observation j

β_{NT} = probability of observing a new target per incremental area (NT density)

β_{FA} = probability of observing a false alarm per incremental area (FA density)

P_D = probability of detecting target

M = dimension of observation space

S = residual covariance (HPH^T+R)

Data Association

Where does this association "cost" come from?

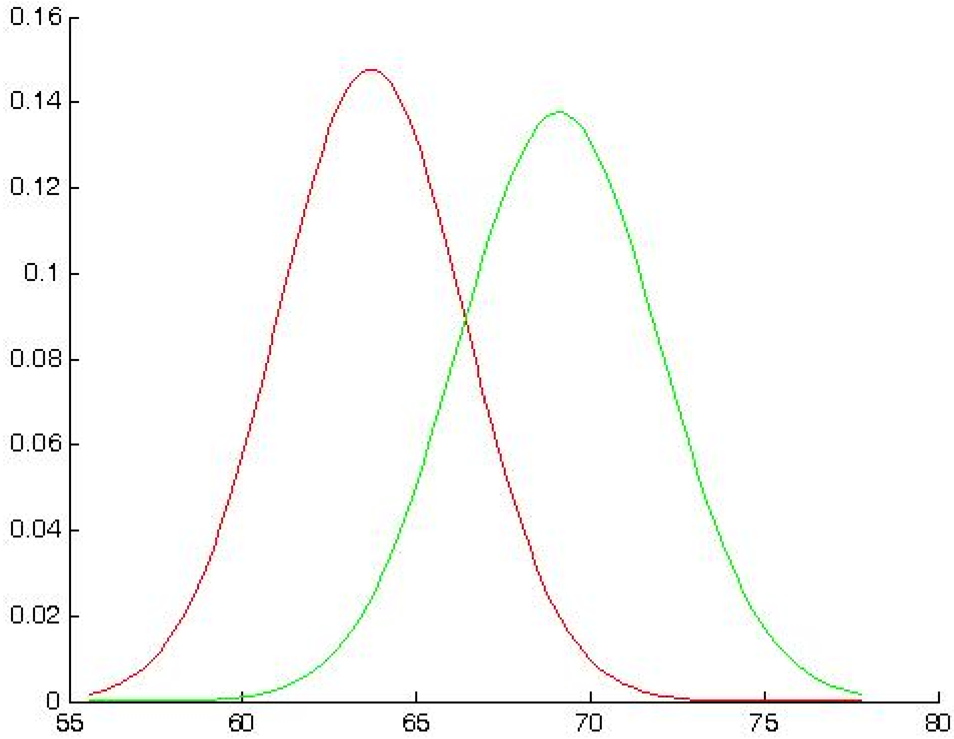
Association Likelihood

- Association Likelihood – a relative measure of how likely it is observation ‘z’ came from (should be associated with) track ‘j’
- L(obs = z | track j), which is read the likelihood of observing the value z given that (or under the hypothesis that) track j produced this observation

Association Likelihood

Distribution of heights for male and female adults in the United States is approximately normal. In inches, the two pdfs are:

$p_M(x) \sim N(x, 69.1, 2.9)$ $p_F(x) \sim N(x, 63.7, 2.7)$



Association Likelihood

Given that we are told a person (U.S. adult) is 65 inches tall, is it more likely that they are male or female?

To answer this question we evaluate each density function at this point observation.

$$L(obs=65 \mid Male) = p_M(65) = 0.051$$

$$L(obs=65 \mid Female) = p_F(65) = 0.131$$

Conclusion: It is roughly two and a half times as likely that this observation (the height) was from the female population as from the male population.

Association Likelihood

Let's add errors!

Assume that the observed height may be off by up to an inch.

The observed height will be correct with probability 0.5 and equally likely to be off by an inch plus or minus.

Now given that the observed height was 67 inches, the likelihood that the observed adult is from the male population is given by the *weighted average* of the density values.

$$
\begin{aligned}
L\ (obs=67 \mid Male) &= Pr\ (obs=67 \mid true=66)*p_M(66) + \\
&\quad\ Pr\ (obs=67 \mid true=67)*p_M(67) + \\
&\quad\ Pr\ (obs=67 \mid true=68)*p_M(68) \\
&= 0.25 * (0.078) + 0.5 * (0.106) + 0.25 * (0.128) \\
&= 0.1045
\end{aligned}
$$

$$L\ (obs=67 \mid Female) = 0.071$$

Association Likelihood

Now suppose the error in our observation is continuous, not discrete.

The error is described by a probability density function $p_e(\ )$, having mean zero.

Now the "weighted average" needs to be calculated over all possible states and so the sum becomes ???

$$\mathrm{L}\ (\mathrm{obs}{=}67 \mid \mathrm{Male}) = \int_{-\infty}^{\infty} \Pr\ (\mathrm{obs}{=}67 \mid x)\,\mathrm{p}_{\mathrm{M}}\left(x\right) dx \quad \text{, with}$$

$$\Pr\ (\mathrm{obs}{=}67 \mid x) = \mathrm{p}_{\mathrm{e}}\left(x-67\right)$$

Association Likelihood

Back to our height example: Suppose the error distribution is normally distributed with mean 0 and standard deviation 0.25 inches. If the observed height is 67.3, what are the likelihoods for the different populations?

$$\mathrm{L}\ (\mathrm{obs}=67.3 \mid \mathrm{Male}) = \int_{-\infty}^{\infty} \Pr\ (\mathrm{obs}=67.3 \mid x)\mathrm{p_M}\left(x\right)dx$$

$$= \int_{-\infty}^{\infty} \left(1/\sqrt{2\pi}\,(0.25) \right) e^{\frac{-(x-67.3)^2}{2(0.25)^2}} \left(1/\sqrt{2\pi}\,(2.9) \right) e^{\frac{-(x-69.1)^2}{2(2.9)^2}} dx$$

Association Likelihood

The integral looks very messy, but we are lucky because there is a nice closed form solution!

In general, if $p_1(x) \sim N\{x: \mu_1, \sigma_1\}$ and $p_2(x) \sim N\{x: \mu_2, \sigma_2\}$ then the integral of the product of the density functions is the value of the density function having mean 0 and standard deviation $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$ evaluated at the difference of the means $\mu_1 - \mu_2$.

$$\int_{-\infty}^{\infty} p_1(x) p_2(x) dx = \left(1 / \sqrt{2\pi} (\sigma) \right) e^{\frac{-(\mu_1 - \mu_2)^2}{2\sigma^2}}$$

Association Likelihood

The likelihoods for the two populations are:

$$L\ (\text{obs}=67.3 \mid \text{Male}) = \left(1 \Big/ \sqrt{2\pi}\,(2.91) \right) e^{\frac{-(67.3-69.1)^2}{2(2.91)^2}} = 0.113$$

$$L\ (\text{obs}=67.3 \mid \text{Female}) = \left(1 \Big/ \sqrt{2\pi}\,(2.71) \right) e^{\frac{-(67.3-63.7)^2}{2(2.71)^2}} = 0.061$$

Time left for more detail?

Association Likelihood

For the Gaussian Tracking problem the association likelihood is

$$\frac{1}{\left|2\pi S\right|^{\frac{1}{2}}} e^{\frac{-\left(Z-H\tilde{X}\right)^{\mathrm{T}} S^{-1}\left(Z-H\tilde{X}\right)}{2}}$$

where

Z is the new observation,

H is the observation matrix,

$\tilde{\mathrm{X}}$ is the extrapolated state,

and S is the residual error covariance

$$\mathrm{S} = \mathrm{H}\,\tilde{\mathrm{P}}\,\mathrm{H}^{\mathrm{T}} + \mathrm{R} \quad .$$

Association Likelihood

Now consider the likelihood of a hypothesis, where the hypothesis consists of a set of tracks that are associated with new observations, a set of tracks that were not detected (did not associate with a new observation) and a set of unassociated observations (new targets or false alarms). For example, given three tracks in the database, T1, T2, T3 and four new observations, O1, O2, O3, O4, one association hypothesis is: {T1<->O2; T2 undetected; T3<->O4; O1 unassoc; O3 unassoc}

To accurately represent this combined event, three additional factors are required:

1) P_D – the sensor's probability of detection

2) β_{NT} – the probability density for new tracks (previously undetected target)

3) β_{FA} – the probability density for false alarms (or sensor noise)

Association Likelihood

The likelihood for the hypothesis is simply the product of the likelihoods for each of the three types of events within the hypothesis:

1) track-to-observation association : P_D * assoc_likelihood
2) track with no observation : $(1-P_D)$
3) unassociated observation : $\beta_{FA} + \beta_{NT}$

where the assoc_likelihood term in the first case is the total likelihood described previously.

Association Likelihood

- The best hypothesis is the one that is the most likely, i.e. the maximum likelihood hypothesis. In order to cast the problem in a more tractable form, the calculations are done in negative natural logarithm space.
- We refer to the negative logarithm of the association likelihood as an *association cost*. Thus, the goal of maximizing the likelihood becomes one of minimizing the cost.
- Finally, note that all of the products that are inherent in working with likelihoods become sums in “cost space”.
- When we convert the likelihoods shown in the previous slide into costs, we obtain the following slide.

Data Association (Repeat)

Actual Association Cost used in most tracking applications

$$\text{cost} = \begin{cases} -\ln(\beta_{NT} + \beta_{FA}) & \text{observation is false alarm/new target} \\ \dfrac{d_{ij}^2}{2} - \ln\left[\dfrac{P_D}{|2\pi S|^{1/2}}\right] & \text{observation to track} \\ -\ln(1 - P_D) & \text{no observation to track} \end{cases}$$

$(d_{ij})^2$ = Mahalanobis distance between track i and observation j

β_{NT} = probability of observing a new target per incremental area (NT density)

β_{FA} = probability of observing a false alarm per incremental area (FA density)

P_D = probability of detecting target

S = residual covariance ($HPH^T + R$)

Discrete Example of Association Calculation

Association Likelihood

Military Example with discrete population distributions and discrete observation conditional probabilities

- Suppose we have two targets, each with a probability distribution on ID
- Discrete Random Variable is X
- Three possible values - HOS, NEU, FRI
- Distribution on X is maintained as a table (one-D array)

Target 1

X	$p_1(X)$
HOS	0.4
NEU	0.4
FRI	0.2

Target 2

X	$p_2(X)$
HOS	0.8
NEU	0.1
FRI	0.1

Association Likelihood

- Observation 'A' is detection of certain RF emitter
  - 10% of Hostile targets carry this emitter
  - 80% of Neutral targets carry this emitter
  - 30% of Friendly targets carry this emitter

- This is a description of the observation likelihood function

$$Pr(obs='A' \mid X=HOS) = 0.1$$

$$Pr(obs='A' \mid X=NEU) = 0.8$$

$$Pr(obs='A' \mid X=FRI) = 0.3$$

Association Likelihood

$$L\,(\text{obs}='A' \mid \text{Target } i) = \sum_{x} \Pr\,(\text{obs}='A' \mid X = x) \bullet p_i(x)$$

$$L\,(\text{obs}=`A' \mid \text{Target } 1) = (0.4)(0.1)+(0.4)(0.8)+(0.2)(0.3)$$
$$= 0.42$$

$$L\,(\text{obs}=`A' \mid \text{Target } 2) = (0.8)(0.1)+(0.1)(0.8)+(0.1)(0.3)$$
$$= 0.2$$

Conclusion: The emitter observation is slightly more than twice as likely to have come from Target 1 as from Target 2.