Introduction to Sensor Fusion

Roland Hostettler

September 10, 2018

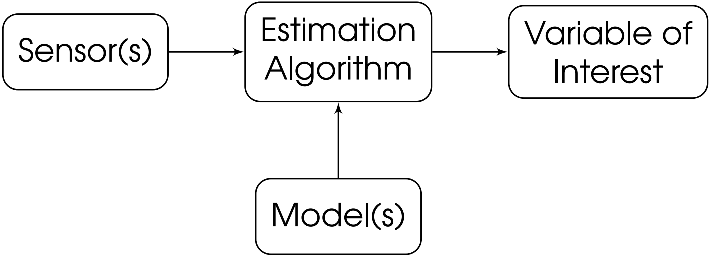
Intended Learning Outcomes

After this lecture, you will be able to:

- Identify the role of sensor fusion in society;
- Recognize and name the components of sensor fusion systems: sensors, models, estimation algorithms;
- Describe and identify the purpose of an optimality criterion and cost functions.

Sensor Fusion: Example Applications

The Components of Sensor Fusion



Variable of Interest

Definition

One or more unknown static parameter(s) or a time-varying state of a dynamic system of interest that can be measured directly or indirectly.

Notation

- A single (scalar) static parameter is denoted θ,
- a vector of K static parameters is denoted as $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 & \theta_2 & \dots & \theta_K \end{bmatrix}^\mathsf{T}$,
- a scalar time-varying state is denoted $x_n = x(t_n)$,
- a vector time-varying state is denoted $\boldsymbol{x}_n = \boldsymbol{x}(t_n)$.

Sensors (1/3)

Definition

A sensor is a device (physical or not) that provides a measurement of a variable of interest.

- May measure the variable directly or indirectly
- Measurement range and environmental conditions
- Is affected by noise, biases, and uncertainty
- May give measurements frequently or infrequently
- Scalar or vector measurements

Sensors (2/3)

Notation

- A scalar measurement is denoted y_n
- A vector measurement is denoted $\boldsymbol{y}_n$
- n is a measurement number, sensor id, time, etc.

Sensors (3/3)

Examples

Model

Definition

Describes how the variable of interest is observed by the sensor in a systematic way.

- May be very simple or very complex
- Takes noise, uncertainty, and other error sources into account
- Formulated using mathematics

A Basic Model

Measurement $=$ Function of Parameter(s) $+$ Noise

Mathematically:

$$y_n = g_n(\boldsymbol{\theta}) + r_n$$

Anatomy:

- measurement y_n is on the left hand side, and
- a function $g_n(\boldsymbol{\theta})$ of $\boldsymbol{\theta}$ and a *noise* term r_n on the right hand side.

This is called *sensor model*, *measurement model*, or *observation model*.

Measurement Noise

- Encodes thermal sensor noise, uncertainty, etc.
- r_n is modeled as a random variable, follows a probability density function (pdf)

$$r_n \sim p(r_n),$$

- For now, we assume zero-mean, independent random variables with variance $\sigma_{r,n}^2$

$$\begin{aligned}
\mathrm{E}\{r_n\} &= 0, \\
\mathrm{var}\{r_n\} &= \mathrm{E}\{r_n^2\} - (\mathrm{E}\{r_n\})^2 = \sigma_{r,n}^2, \\
\mathrm{Cov}\{r_m, r_n\} &= \mathrm{E}\{r_m r_n\} - \mathrm{E}\{r_m\}\,\mathrm{E}\{r_n\} = 0 \quad (m \neq n)
\end{aligned}$$

Vector Model

- Extending the basic model for *vector-valued* measurements:

$$\boldsymbol{y}_n = g_n(\boldsymbol{\theta}) + \boldsymbol{r}_n,$$

- $\boldsymbol{y}_n$ and $\boldsymbol{r}_n$ are d_y-dimensional column vectors
- $\boldsymbol{r}_n$ is a multivariate random variable with pdf

$$\boldsymbol{r}_n \sim p(\boldsymbol{r}_n)$$

- Assume zero-mean, independent random variables with covariance $\boldsymbol{R}_n$

$$\begin{aligned}
\mathrm{E}\{\boldsymbol{r}_n\} &= 0, \\
\mathrm{Cov}\{\boldsymbol{r}_n\} &= \mathrm{E}\{\boldsymbol{r}_n \boldsymbol{r}_n^\mathsf{T}\} - \mathrm{E}\{\boldsymbol{r}_n\}\,\mathrm{E}\{\boldsymbol{r}_n\}^\mathsf{T} = \boldsymbol{R}_n, \\
\mathrm{Cov}\{\boldsymbol{r}_m, \boldsymbol{r}_n\} &= \mathrm{E}\{\boldsymbol{r}_m \boldsymbol{r}_n^\mathsf{T}\} - \mathrm{E}\{\boldsymbol{r}_m\}\,\mathrm{E}\{\boldsymbol{r}_n\}^\mathsf{T} = 0 \quad (m \neq n)
\end{aligned}$$

Multiple Measurements

- Sensor fusion requires multiple sensors, repeated measurements, or both
- In the terminology of the measurement model, they can be regarded the same: $y_1, y_2, \dots, y_N$ or $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_N$
- We denote a set of measurements:
  - $y_{1:N} = \{y_1, y_2, \dots, y_N\}$ for the scalar case
  - $\boldsymbol{y}_{1:N} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_N\}$ for the vector case
- Examples: Sensor networks, sensor arrays, multi-view imaging, etc.

Measurement Stacking: Scalar Case

- Remember: $y_n = g_n(\boldsymbol{\theta}) + r_n$
- Given the measurements $y_{1:N}$, we can write:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} g_1(\boldsymbol{\theta}) \\ g_2(\boldsymbol{\theta}) \\ \vdots \\ g_N(\boldsymbol{\theta}) \end{bmatrix} + \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_N \end{bmatrix}$$

- *Compact* notation for *all measurements*:

$$\boldsymbol{y} = g(\boldsymbol{\theta}) + \boldsymbol{r}.$$

- Covariance for $\boldsymbol{r}$:

$$\mathrm{Cov}\{\boldsymbol{r}\} = \boldsymbol{R} = \begin{bmatrix} \sigma_{r,1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{r,2}^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \sigma_{r,N}^2 \end{bmatrix}$$

Measurement Stacking: Vector Case

- Vector case: $\boldsymbol{y}_n = g_n(\boldsymbol{\theta}) + \boldsymbol{r}_n$
- Stacked notation:

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_N \end{bmatrix} == \begin{bmatrix} g_1(\boldsymbol{\theta}) \\ g_2(\boldsymbol{\theta}) \\ \vdots \\ g_N(\boldsymbol{\theta}) \end{bmatrix} + \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_N \end{bmatrix}$$

- Hence,

$$\boldsymbol{y} = g(\boldsymbol{\theta}) + \boldsymbol{r}.$$

where $\mathrm{Cov}\{\boldsymbol{r}\} = \boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_1 & 0 & \dots & 0 \\ 0 & \boldsymbol{R}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \boldsymbol{R}_N \end{bmatrix}.$

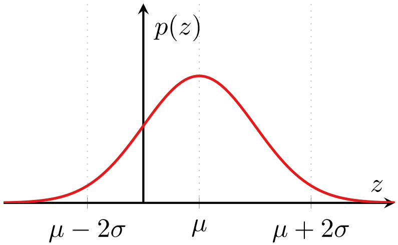
Gaussian Measurement Noise

- Noise is often assumed to be Gaussian, i.e.

$$p(\boldsymbol{r}) = \frac{1}{(2\pi)^{d_y N/2}|\boldsymbol{R}|^{1/2}} e^{-\frac{1}{2}\boldsymbol{r}^\mathsf{T}\boldsymbol{R}^{-1}\boldsymbol{r}}$$

- Compact notation $p(\boldsymbol{r}) = \mathcal{N}(\boldsymbol{r}; 0, \boldsymbol{R})$, where

$$\mathcal{N}(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{M/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{\mu})^\mathsf{T}\boldsymbol{\Sigma}^{-1}(\boldsymbol{z}-\boldsymbol{\mu})}$$



Estimation Algorithm

Definition

Combines the measurements from multiple sensors by using the corresponding models to estimate the parameters of interest in some optimal sense.

- Combining multiple measurements increases the precision (on average)
- Measurements from different sensors can easily be incorporated
- Can account for the uncertainty of different measurements

Cost Functions (1/2)

- An optimality criterion is required to develop an estimation algorithm
- We focus on algorithms that minimize a cost function of the error

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, J(\boldsymbol{\theta})$$

where
  - $\hat{\boldsymbol{\theta}}$ denotes the estimate of $\boldsymbol{\theta}$
  - $J(\boldsymbol{\theta})$ is the cost function
  - $\operatorname{argmin}_{\boldsymbol{\theta}} J(\boldsymbol{\theta})$ denotes "the argument $\boldsymbol{\theta}$ that minimizes $J(\boldsymbol{\theta})$"
- The error is given by the difference between the measurement and the output predicted by $\boldsymbol{\theta}$

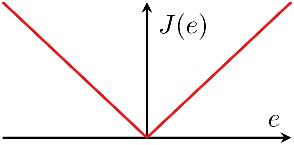
$$e_n = y_n - g_n(\boldsymbol{\theta})$$

Cost Functions (2/2)

Absolute Error

Penalizes all errors equally:

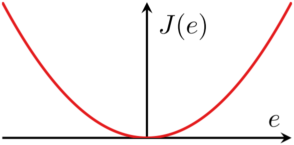
$$|e_n| = |y_n - g_n(\boldsymbol{\theta})|,$$



Quadratic Error

Penalizes large errors more than small ones:

$$e_n^2 = (y_n - g_n(\boldsymbol{\theta}))^2.$$



Least Squares (1/2)

- The quadratic cost is much more common
- Closely related to Gaussian measurement noise
- Minimizing the quadratic cost function is the least squares method
- Cost function for N scalar measurements $y_{1:N} = \{y_1, y_2, \ldots, y_N\}$

$$J_{\mathsf{LS}}(\boldsymbol{\theta}) = \sum_{n=1}^{N} e_n^2 = \sum_{n=1}^{N} (y_n - g_n(\boldsymbol{\theta}))^2$$

Least Squares (2/2)

- Quadratic error for vector measurements

$$e_n^2 = (\boldsymbol{y}_n - g_n(\boldsymbol{\theta}))^\mathsf{T}(\boldsymbol{y}_n - g_n(\boldsymbol{\theta}))$$

- Cost function for N vector measurements $\boldsymbol{y}_{1:N} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_N\}$

$$J_{\mathsf{LS}}(\boldsymbol{\theta}) = \sum_{n=1}^{M} (\boldsymbol{y}_n - g_n(\boldsymbol{\theta}))^\mathsf{T}(\boldsymbol{y}_n - g_n)$$

- Quadratic error and cost function for stacked (batch) notation

$$J_{\mathsf{LS}}(\boldsymbol{\theta}) = (\boldsymbol{y} - g(\boldsymbol{\theta}))^\mathsf{T}(\boldsymbol{y} - g(\boldsymbol{\theta}))$$

Weighted Least Squares (1/2)

- How to include confidence in sensor readings?
- Weighted least squares (WLS) cost function:

$$J_{\mathsf{WLS}}(\boldsymbol{\theta}) = \sum_{n=1}^{N} w_n (y_n - g_n(\boldsymbol{\theta}))^2.$$

  where $w_n > 0$ is a weighing factor for the nth measurement
- WLS vector cost function:

$$J_{\mathsf{WLS}}(\boldsymbol{\theta}) = \sum_{n=1}^{N} (\boldsymbol{y}_n - g_n(\boldsymbol{\theta}))^{\mathsf{T}} \boldsymbol{W}_n (\boldsymbol{y}_n - g_n(\boldsymbol{\theta})).$$

  where $\boldsymbol{W}_n$ is a positive-definite weighing matrix

Weighted Least Squares (2/2)

- WLS stacked cost function:

$$J_{\mathsf{WLS}}(\boldsymbol{\theta}) = (\boldsymbol{y} - g(\boldsymbol{\theta}))^{\mathsf{T}} \boldsymbol{W} (\boldsymbol{y} - g(\boldsymbol{\theta}))$$

where $\boldsymbol{W}$ is the positive-definite weighing matrix

$$\boldsymbol{W} = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & w_N \end{bmatrix} \text{ or } \boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_1 & 0 & \dots & 0 \\ 0 & \boldsymbol{W}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \boldsymbol{W}_N \end{bmatrix}$$

- Choice of w_n or $\boldsymbol{W}_n$ is in principle arbitrary
- In practice, good choices are

$$w_n = 1/\sigma_{r,n}^2 \text{ and } \boldsymbol{W}_n = \boldsymbol{R}_n^{-1},$$

Task: Sensor Fusion in Your Life

Task

With your neighbor, discuss and think about an example of information fusion in your daily life and identify the components. Prepare to describe your example in 3–4 sentences.

Time

5 minutes.

Summary

- Many problems in science, engineering, and society are sensor fusion problems
- Sensor fusion involves three components:
  1. Sensor: Measures a variable of interest, directly or indirectly
  2. Model: A mathematical formulation that relates the variables of interest to the measurements
  3. Estimation Algorithm: Combines the measurements and models to estimate the variables of interest
- The least squares method is a good starting point for deriving estimators