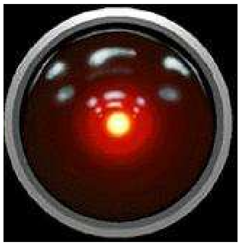


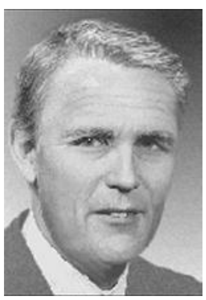
HCI/ComS 575X:
Computational Perception

Instructor: Alexander Stoytchev
http://www.cs.iastate.edu/~alex/classes/2007_Spring_575X/

The Kalman Filter (part 1)

March 5, 2007

Rudolf Emil Kalman



[http://www.cs.unc.edu/~welch/kalman/kalmanBiblio.html]

Definition

- A Kalman filter is simply an optimal recursive data processing algorithm

- Under some assumptions the Kalman filter is optimal with respect to virtually any criterion that makes sense.

Definition

"The Kalman filter incorporates all information that can be provided to it. It processes all available measurements, regardless of their precision, to estimate the current value of the variables of interest."

[Maybeck (1979)]

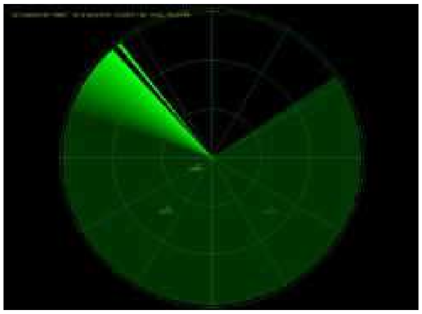
Why do we need a filter?

- No mathematical model of a real system is perfect

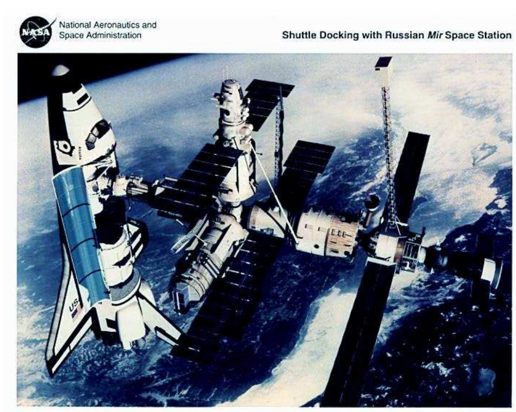
- Real world disturbances

- Imperfect Sensors

Application: Radar Tracking



Application: Lunar Landing



Application: Missile Tracking

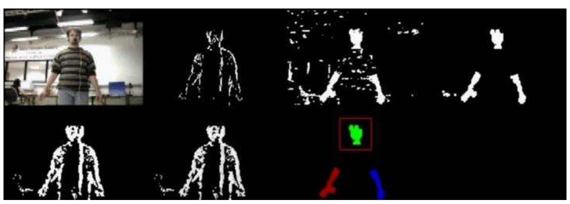
Application: Sailing

Application: Robot Navigation

Application: Other Tracking

Application: Head Tracking

Face & Hand Tracking



Brown and Hwang (1992)

“Introduction to Random Signals and Applied Kalman Filtering”

Ch 5: The Discrete Kalman Filter

Maybeck, Peter S. (1979)

Chapter 1 in ``Stochastic models, estimation, and control'',

Mathematics in Science and Engineering Series, Academic Press.

Arthur Gelb, Joseph Kasper, Raymond Nash, Charles Price, Arthur Sutherland (1974)

Applied Optimal Estimation

MIT Press.

A Simple Recursive Example

- Problem Statement:

Given the measurement sequence:
$z_1, z_2, \ldots, z_n$ find the mean

[Brown and Hwang (1992)]

First Approach

1. Make the first measurement z_1
   Store z_1 and estimate the mean as
   $\mu_1 = z_1$

2. Make the second measurement z_2
   Store z_1 along with z_2 and estimate the mean as

   $\mu_2 = (z_1 + z_2)/2$

[Brown and Hwang (1992)]

First Approach (cont'd)

3. Make the third measurement z_3
   Store z_3 along with z_1 and z_2 and estimate the mean as

   $\mu_3 = (z_1 + z_2 + z_3)/3$

[Brown and Hwang (1992)]

First Approach (cont'd)

n. Make the n-th measurement z_n
   Store z_n along with $z_1, z_2, \ldots, z_{n-1}$ and estimate the mean as

   $\mu_n = (z_1 + z_2 + \ldots + z_n)/n$

[Brown and Hwang (1992)]

Second Approach

1. Make the first measurement z_1
   Compute the mean estimate as

   $\mu_1 = z_1$

   Store μ_1 and discard z_1

[Brown and Hwang (1992)]

Second Approach (cont'd)

2. Make the second measurement z_2
   Compute the estimate of the mean as a weighted sum of the previous estimate μ_1 and the current measurement z_2:

   $\mu_2 = 1/2\ \mu_1 + 1/2\ z_2$

   Store μ_2 and discard z_2 and μ_1

[Brown and Hwang (1992)]

Second Approach (cont'd)

3. Make the third measurement z_3
   Compute the estimate of the mean as a weighted sum of the previous estimate μ_2 and the current measurement z_3:

   $$\mu_3 = 2/3\ \mu_2 + 1/3\ z_3$$

   Store μ_3 and discard z_3 and μ_2

[Brown and Hwang (1992)]

Second Approach (cont'd)

n. Make the n-th measurement z_n
   Compute the estimate of the mean as a weighted sum of the previous estimate μ_{n-1} and the current measurement z_n:

   $$\mu_n = (n-1)/n\ \mu_{n-1} + 1/n\ z_n$$

   Store μ_n and discard z_n and μ_{n-1}

[Brown and Hwang (1992)]

Comparison

Batch Method	Recursive Method
$\hat{x}_1 = z_1$	$\hat{x}_1 = z_1$
$\hat{x}_2 = \frac{z_1 + z_2}{2}$	$\hat{x}_2 = \frac{1}{2}\hat{x}_1 + \frac{1}{2}z_2$
$\hat{x}_3 = \frac{z_1 + z_2 + z_3}{3}$	$\hat{x}_3 = \frac{2}{3}\hat{x}_2 + \frac{1}{3}z_3$
$\hat{x}_n = \frac{z_1 + z_2 + \ldots + z_n}{n}$	$\hat{x}_n = \frac{n-1}{n}\hat{x}_{n-1} + \frac{1}{n}z_n$

Analysis

- The second procedure gives the same result as the first procedure.
- It uses the result for the previous step to help obtain an estimate at the current step.
- The difference is that it does not need to keep the sequence in memory.

[Brown and Hwang (1992)]

Second Approach (rewrite the general formula)

$$\mu_n = (n-1)/n\ \mu_{n-1} + 1/n\ z_n$$

Second Approach (rewrite the general formula)

$$\mu_n = (n-1)/n\ \mu_{n-1} + 1/n\ z_n$$

$$\mu_n = \mu_{n-1} \quad + \quad 1/n\ (z_n - \mu_{n-1})$$

Second Approach (rewrite the general formula)

$\mu_n = (n-1)/n\ \mu_{n-1} + 1/n\ z_n$

$\mu_n = \mu_{n-1} + 1/n\ (z_n - \mu_{n-1})$

- μ_{n-1}: Old Estimate
- $1/n$: Gain Factor
- $(z_n - \mu_{n-1})$: Difference Between New Reading and Old Estimate

Second Approach (rewrite the general formula)

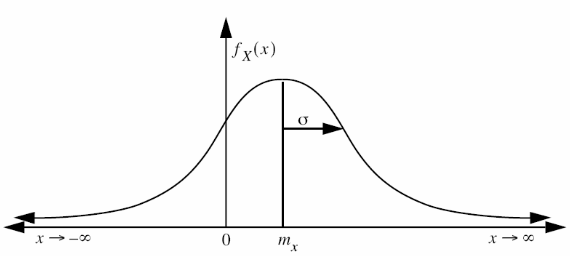
$$\begin{aligned} \hat{x}_n &= \left(\frac{n-1}{n}\right)\hat{x}_{n-1} + \left(\frac{1}{n}\right) z_n \\ &= \left(\frac{n}{n}\right)\hat{x}_{n-1} - \left(\frac{1}{n}\right)\hat{x}_{n-1} + \left(\frac{1}{n}\right) z_n \\ &= \hat{x}_{n-1} + \frac{1}{n}\left(z_n - \hat{x}_{n-1}\right) \end{aligned}$$

Gaussian Properties

The Gaussian Function

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

Gaussian pdf



Properties

- If $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$
- Then $Y \sim N(a\mu + b, a^2\sigma^2)$

pdf for $Y = aX + b$

$$f_Y(y) = \frac{1}{\sqrt{2\pi a^2\sigma^2}} e^{-\frac{1}{2}\frac{(y-(a\mu+b))^2}{a^2\sigma^2}}$$

Properties

Finally, if X_1 and X_2 are independent (see Section 2.5 below), $X_1 \sim N(\mu_1, \sigma_1^2)$, and $X_2 \sim N(\mu_2, \sigma_2^2)$, then

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2), \qquad (2.14)$$

and the density function becomes

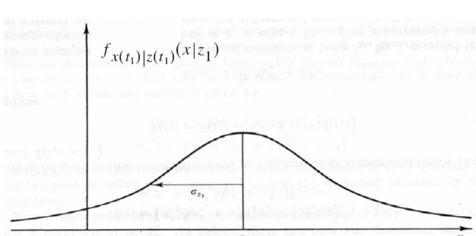
$$f_X(x_1 + x_2) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{1}{2}\frac{(x-(\mu_1+\mu_2))^2}{(\sigma_1^2+\sigma_2^2)}}. \qquad (2.15)$$

Summation and Subtraction

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$
$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2$$

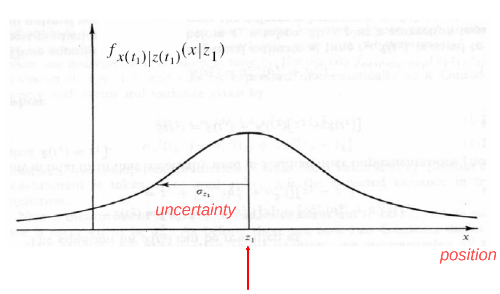
A simple example using diagrams

Conditional density of position based on measured value of z_1



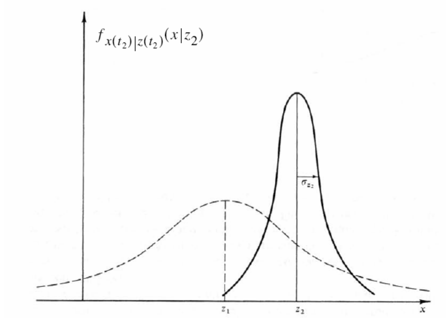
[Maybeck (1979)]

Conditional density of position based on measured value of z_1



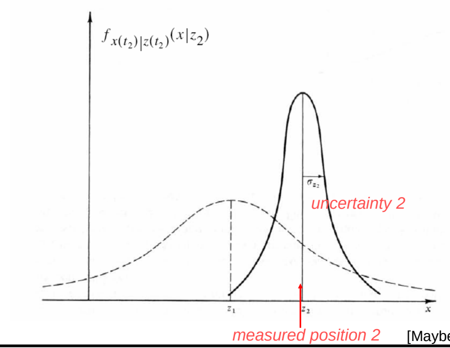
[Maybeck (1979)]

Conditional density of position based on measurement of z_2 alone



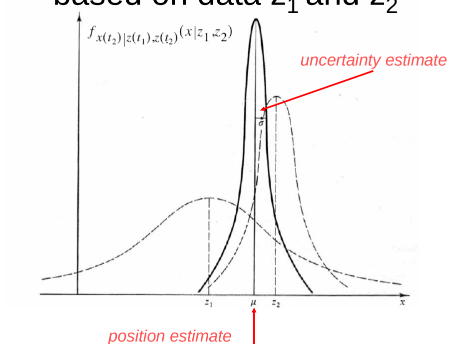
[Maybeck (1979)]

Conditional density of position based on measurement of z_2 alone



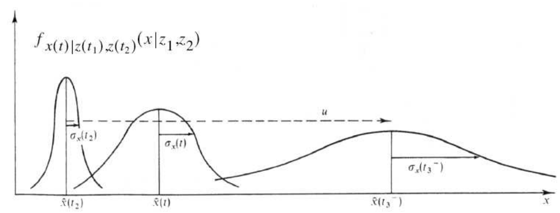
[Maybeck (1979)]

Conditional density of position based on data z_1 and z_2



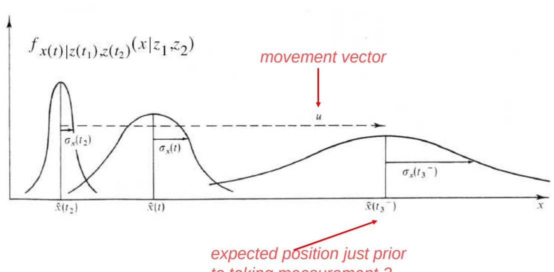
[Maybeck (1979)]

Propagation of the conditional density



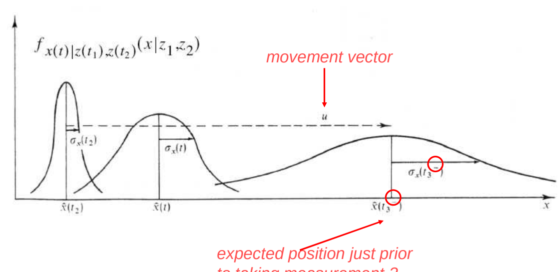
[Maybeck (1979)]

Propagation of the conditional density



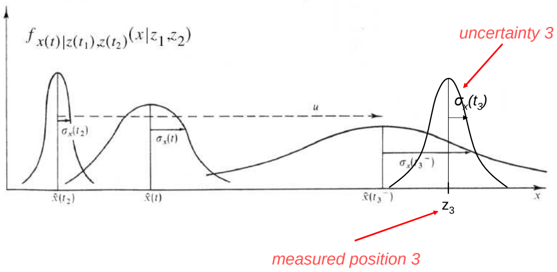
[Maybeck (1979)]

Propagation of the conditional density

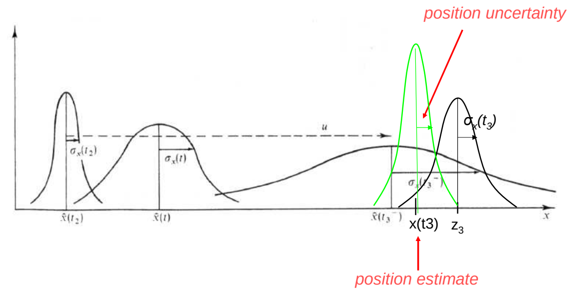


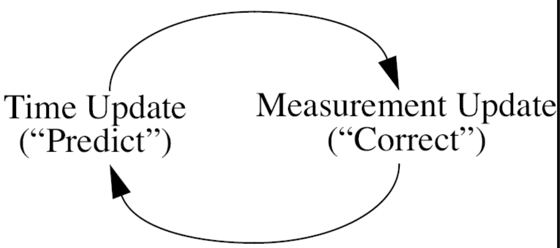
[Maybeck (1979)]

Propagation of the conditional density



Updating the conditional density after the third measurement

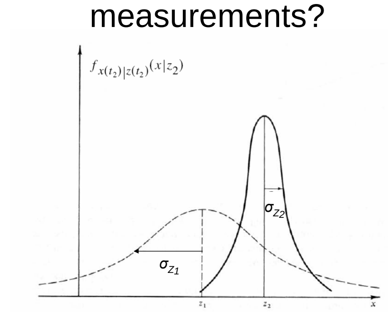




Questions?

Now let's do the same thing
…but this time we'll use math

How should we combine the two measurements?



[Maybeck (1979)]

Calculating the new mean

$\mu = $ Scaling Factor 1 $z_1 +$ Scaling Factor 2 z_2

Calculating the new mean

$\mu = $ Scaling Factor 1 $z_1 +$ Scaling Factor 2 z_2

$$\mu = [\sigma_{z_2}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_2$$

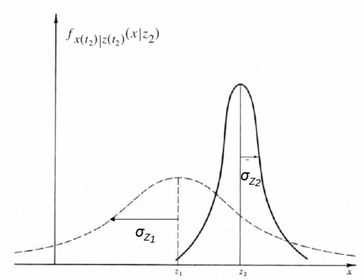
Calculating the new mean

$\mu = $ Scaling Factor 1 $z_1 +$ Scaling Factor 2 z_2

$$\mu = [\sigma_{z_2}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_2$$

Why is this not z_1?

Calculating the new variance



[Maybeck (1979)]

Calculating the new variance

$\sigma^2 = $ Scaling Factor 1 $\sigma_{z_1}^2 +$ Scaling Factor 2 $\sigma_{z_2}^2$

Calculating the new variance

$\sigma^2 = $ Scaling Factor 1 $\sigma_{z_1}^2 +$ Scaling Factor 2 $\sigma_{z_2}^2$

$\sigma_{z_2}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)$ $\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)$

Calculating the new variance

$$\sigma^2 = \underbrace{\boxed{\text{Scaling Factor 1}}}_{\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)}\sigma_{z_1}^2 + \underbrace{\boxed{\text{Scaling Factor 2}}}_{\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)}\sigma_{z_2}^2$$

$$\sigma^2 = [\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]\sigma_{z_1}^2 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]\sigma_{z_2}^2$$

Calculating the new variance

$$\sigma^2 = [\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]\sigma_{z_1}^2 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]\sigma_{z_2}^2$$

Calculating the new variance

$$\sigma^2 = [\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]\sigma_{z_1}^2 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]\sigma_{z_2}^2$$

$$\sigma^2 = \frac{\sigma_{z_2}^2\ \sigma_{z_1}^2\ +\ \sigma_{z_1}^2\ \sigma_{z_2}^2}{(\sigma_{z_1}^2+\sigma_{z_2}^2)} = \frac{2\ \sigma_{z_1}^2\ \sigma_{z_2}^2}{(\sigma_{z_1}^2+\sigma_{z_2}^2)}$$

Calculating the new variance

$$\sigma^2 = [\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]\sigma_{z_1}^2 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]\sigma_{z_2}^2$$

$$\sigma^2 = \frac{\sigma_{z_2}^2\ \sigma_{z_1}^2\ +\ \sigma_{z_1}^2\ \sigma_{z_2}^2}{(\sigma_{z_1}^2+\sigma_{z_2}^2)} = \frac{2\ \sigma_{z_1}^2\ \sigma_{z_2}^2}{(\sigma_{z_1}^2+\sigma_{z_2}^2)}$$

$$2/\sigma^2 = (1/\sigma_{z_1}^2)+(1/\sigma_{z_2}^2)$$

Why is this result different from the one given in the paper?

$$1/\sigma^2 = (1/\sigma_{z_1}^2)+(1/\sigma_{z_2}^2)$$

Remember the Gaussian Properties?

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2$$

Remember the Gaussian Properties?

- If $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$
- Then $Y \sim N(a\mu + b, a^2\sigma^2)$

This is a^2 not a

The scaling factors must be squared!

$$\sigma^2 = \underbrace{\text{Scaling Factor 1}}_{[\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]^2}\sigma_{z_1}^2 \quad \underbrace{\text{Scaling Factor 2}}_{[\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]^2}\sigma_{z_2}^2$$

The scaling factors must be squared!

$$\sigma^2 = \underbrace{\text{Scaling Factor 1}}_{[\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]^2}\sigma_{z_1}^2 + \underbrace{\text{Scaling Factor 2}}_{[\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]^2}\sigma_{z_2}^2$$

$$\sigma^2 = [\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]^2\sigma_{z_1}^2 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]^2\sigma_{z_2}^2$$

Therefore the new variance is

$$1/\sigma^2 = (1/\sigma_{z_1}^2) + (1/\sigma_{z_2}^2)$$

Try to derive this on your own.

Another Way to Express The New Position

$$\hat{x}(t_2) = [\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]z_2$$

$$= z_1 - z_1 + [\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]z_2$$

$$= z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)][z_2 - z_1]$$

[Maybeck (1979)]

Another Way to Express The New Position

$$\hat{x}(t_2) = [\sigma_{z_2}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)]z_2$$

$$= z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2+\sigma_{z_2}^2)][z_2 - z_1]$$

$$\hat{x}(t_2) = \hat{x}(t_1) + K(t_2)\,[z_2 - \hat{x}(t_1)]$$

[Maybeck (1979)]

Another Way to Express The New Position

$$\hat{x}(t_2) = \hat{x}(t_1) + K(t_2)[z_2 - \hat{x}(t_1)]$$

$$K(t_2) = \sigma_{z_1}^2 / (\sigma_{z_1}^2 + \sigma_{z_2}^2)$$

[Maybeck (1979)]

The equation for the variance can also be rewritten as

$$\sigma_x^2(t_2) = \sigma_x^2(t_1) - K(t_2)\sigma_x^2(t_1)$$

[Maybeck (1979)]

Adding Movement

$$dx/dt = u + w$$

[Maybeck (1979)]

Adding Movement

$$\hat{x}(t_3^-) = \hat{x}(t_2) + u[t_3 - t_2]$$

$$\sigma_x^2(t_3^-) = \sigma_x^2(t_2) + \sigma_w^2[t_3 - t_2]$$

[Maybeck (1979)]

Adding Movement

$$\hat{x}(t_3) = \hat{x}(t_3^-) + K(t_3)[z_3 - \hat{x}(t_3^-)]$$

$$\sigma_x^2(t_3) = \sigma_x^2(t_3^-) - K(t_3)\sigma_x^2(t_3^-)$$

$$K(t_3) = \sigma_x^2(t_3^-)/[\sigma_x^2(t_3^-) + \sigma_{z_3}^2]$$

[Maybeck (1979)]

Properties of K

- If the measurement noise is large K is small

$$K(t_3) = \sigma_x^2(t_3^-)/[\sigma_x^2(t_3^-) + \sigma_{z_3}^2]$$

$$\sigma_{z_3}^2 \to \infty \ , \ K(t_3) \to 0$$

[Maybeck (1979)]

THE END