Course 11

Tracking: Beyond 15 Minutes of Thought
Introductions

Danette Allen
Greg Welch
Gary Bishop
Why thinking about tracking is so fun

It’s a simple problem to state

It has a little of everything

- A little physics
- A little electronics
- A little math
- A little signal processing
- A little programming
Why don’t you just…

- Mount TV cameras on the walls?
- Use GPS?
- Use MEMS accelerometers?
- Use carbon nanotubes?
The 15 minute effect...
Goals

- To get you to the 15 minute point and beyond...
- To equip you to evaluate the various offerings and understand the strengths and weaknesses of each.
Tracking Technologies

(Danette Allen)
Many ways to slice!

Configuration
- Outside-in vs. Inside-out

Type of measurement
- Absolute vs. Relative
- Range vs. angle

Physical medium
- Five categories
Five (Six) Technologies

- Inertial
- Acoustic
- Magnetic
- Mechanical
- Optical

Radio (GPS) is the sixth...

- typically used outdoors
- not addressed in this course
Inertial Tracking

Passive

Newton’s 2nd Law of Motion

- $F = ma$ (linear)
- $\tau = I\alpha$ (rotational)

No physical limits on working volume

Accelerometers and gyroscopes

- Derivative measurements
Inertial Tracking

Accelerometers

• Measure force exerted on a mass since we cannot measure acceleration directly.

• Proof-mass and damped spring
  • Displacement proportional to acceleration

• Potentiometric and piezoelectric transducers
Inertial Tracking

Gyrosopes

- Inertia
  rigidity in space

- Precession
  a comparatively slow gyration of the rotation axis of a spinning body about another line intersecting it so as to describe a cone (Mirriam-Webster)

- Gimbal deflection

(Discovery, 2001)
Microgyroscope

MEMS
Tuning fork

Ring-laser and Fiber
  • Doppler effect
  • Beat frequency
Inertial Drift

Error accumulation due to integration
- Poor SNR at low frequencies
  - Inverse square weighting of noise

\[ \mathcal{E}_a(s) = \frac{1}{s} \frac{\hat{x}}{s} \frac{1}{s} \hat{x} \]

- Gravity vector misalignment
  - 1° tilt error over 10 seconds \(\Rightarrow\) 9m position error

Periodic recalibration
- hybrid systems typical

LaPlace Transform
\[ s = \sigma + j\omega \]
Time [s] to 0.1 [m] Error

Effect of Sensor Noise on Inertial-Only Performance (time to 0.1 meters error)

Rate gyro dynamic range (5 rad/s divided by noise) bits

Accelerometer dynamic range (3G divided by noise) bits
**Acoustic Tracking**

**The Geometry**
- The intersection of 2 spheres is a circle.
- The intersection of 3 spheres is 2 points.
  - One of the two points easily eliminated

**Speed of Sound**
- Varies with temperature and pressure
- \(~ 331 \text{ [m/s]}\) in air at \(0^\circ \text{C}\)
  - \(1 \text{ ft/ms} \Rightarrow \text{SLOW!!}\)

**Ultrasonic**
- \(40 \text{ [kHz]}\) typical
Acoustic Tracking Methods

Time of Flight

- Measures the time required for a sonic pulse or pattern to travel from a transmitter to a receiver.

\[ d [\text{m}] = v [\text{m/s}] \times t [\text{s}], \quad v = \text{speed of sound} \ (c) \]

- Absolute range measurement

Phase Coherence

- Measures phase difference between transmitted and received sound waves

- Relative to previous measurement
  - still absolute!!
Phase Coherence

Equations

• $A \cos(\omega t - \phi)$
• $c [\text{m/s}] = \lambda [\text{m}] * f [1/\text{s}]$
• $\delta [\text{m}] = \lambda [\text{m}] * (\phi / 2\pi)$

“Relative” Result

• Fractional wavelength
• Need previous range estimate
  • No integration!!!
Magnetic Tracking

Three mutually-orthogonal coils

\[
\begin{align*}
H_r &= \frac{M}{2\pi d^3} \cos \theta \\
H_\theta &= \frac{M}{2\pi d^3} \sin \theta
\end{align*}
\]

- Each transmitter coil activated serially
  - Three measurements apiece (three receiver coils)
  - Nine-element measurement for 6D position

AC vs. DC

- Ferromagnetic interference
Mechanical Tracking

Ground-based or body-based
Used primarily for motion capture
Provide angle and range measurements
  - gears
  - bend sensors

Elegant addition of force feedback
Optical Tracking

Provides angle measurements

• One 2D point defines a ray
• Two 2D points define a point for 3D position
• Additional 2D points required for orientation

Speed of light

• $2.998 \times 10^8$ m/s (1 ft/ns)
Active vs. Passive Targets

Typical detectors
- Video and CCD cameras
- Computer vision techniques

Passive targets
- Reflective materials, high contrast patterns
Active vs. Passive Targets

Typical detectors

• LEPDs

\[ I = I_0 \left( \frac{\sinh(\alpha(L - x))}{\sinh(\alpha L)} \right) \quad \text{or} \quad I \approx I_0 \left( \frac{L - x}{L} \right) \]

• Quad Cells

\[ x = \frac{(i_1 + i_2) - (i_3 + i_4)}{i_1 + i_2 + i_3 + i_4} \quad \quad y = \frac{(i_1 + i_4) - (i_2 + i_3)}{i_1 + i_2 + i_3 + i_4} \]

Active targets

• LEDs
Many ways to slice!

**Configuration**
- Outside-in vs. Inside-out

**Type of measurement**
- Absolute vs. Relative
- Range vs. angle

**Physical medium**
- Five categories
Sensor Configurations

Geometric arrangement of sensors and sources impacts:

- accuracy
- usability
- algorithms
for example CODA mpx30

3 1-D CCDs are stationary
LED targets move
Very interesting optics and sensing
CODA mpx30

- Measures angles in lab coordinate frame
- Angle determines a plane
- Intersecting 3 planes determines a point

“Flatland”
CODA mpx30

Such “outside-looking-in” systems

- measure position very well
- allow many small moving targets
- use multiple targets to get orientation
- trade off accuracy and working volume
- provide larger volume / more accuracy with more sensors
- use really simple math
HiBall

6 2-D sensors and 6 lenses in dodecahedron
1000’s of LEDs fixed on the ceiling
Calibration gives effectively 26 cameras
HiBall

- Measures angles in user coordinate frame
- Angles determine a constraint relating
  - position
  - orientation
  - view
  - led location

“Flatland”
HiBall

Such “inside-looking out” systems

- directly measure orientation
- allow large working volume with accuracy
- are larger than LED targets
- and thus harder to use for hands, feet, etc.
Arc Second Vulcan

Sources scan “planes of light” through space.
Sensors on target detect passing plane.
Arc Second Vulcan

- Time of plane passing converts to angle at the source
- Measures angles in world frame
- Thus like CODA and other “outside-looking-in” systems
- Direction of “looking” really isn’t the issue but coordinate frame of measurement
User and Sensor Uncertainty/Information

(Greg Welch)
Pose Uncertainty

- **Measurement uncertainty**
  - Pose estimates from *noisy* sensor measurements

- **User pose uncertainty**
  - Noisy and *temporally-discrete* measurements
  - Modeling user motion is difficult [Weber]
  - Modeling pose *uncertainty* is less difficult
Noise-Driven Processes

Random noise → Integration → Velocity uncertainty → Integration → Position uncertainty → User

“True” sensor measurement → Tracker

Random noise

Noisy measurement
Random Variables and Signals

- Map sample space $\rightarrow$ real numbers
  - For example, time to voltage

- Random Signals
  - For example, electrical signals
  - Continuous random variables
  - Probability over a *region* of sample space
  - Spatial (statistical) and temporal (spectral) aspects
Cumulative Distribution Function

\[ F_X(x) = P(-\infty, x] \]

1. \( F_X(x) \rightarrow 0 \) as \( x \rightarrow -\infty \)
2. \( F_X(x) \rightarrow 1 \) as \( x \rightarrow +\infty \)
3. \( F_X(x) \) is a non-decreasing function of \( x \)
Probability Density Function

\[ f_x(x) = \frac{d}{dx} F_X(x) \]

1. \( f_X(x) \) is a non-negative function
2. \( \int_{-\infty}^{\infty} f_X(x) \, dx = 1 \)
Probability (Continuous)

\[ P_x[a, b] = \int_a^b f_x(x) \, dx \]
Statistical Moments

\[ \mu_m = E[X^m] = \]

Continuous:

\[ \int_{-\infty}^{\infty} x^m f_X(x) \, dx \]

Discrete:

\[ \sum_{x} x^m p_X(x) \]
1st Moment or **Mean**

\[ \mu = E[X] = \]

Continuous:

\[ \int_{-\infty}^{\infty} x f_X(x) \, dx \]

Discrete:

\[ \sum_{x} x p_X(x) \]
Central Moments

\[ c_m = E \left[ (X - \mu)^m \right] = \]

Continuous:

\[ \int_{-\infty}^{\infty} (x - \mu)^m f_X(x) \, dx \]

Discrete:

\[ \sum_x (x - \mu)^m p_X(x) \]
2nd Central Moment or **Variance**

\[ V[X] = E[(X - \mu)^2] \]

\[ = E[X^2] - \mu^2 \]

"Mean of square minus square of mean"
Standard Deviation

\[ \sigma = \sqrt{V[X]} \]
The Gaussian/Normal Distribution is defined by the probability density function:

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

where \( \mu \) is the mean and \( \sigma^2 \) is the variance. The distribution is denoted as \( X \sim N(\mu, \sigma^2) \).
Autocorrelation (Time Domain)

$$R_X(\tau) = E[X(t)X(t + \tau)]$$
Spectral Density (Frequency Domain)

The Wiener-Khinchine relationship

\[ S_X(j\omega) = F[R_X(\tau)] = \int_{-\infty}^{\infty} R_X(\tau)e^{-j\omega \tau} \, d\tau \]
White Noise Process

\[ R_X(\tau) = \begin{cases} 
  \text{if } \tau = 0 \text{ then } A \\
  \text{else } 0 
\end{cases} \]

\[ S_X(j\omega) = A \]
Growth in Pose Uncertainty

\[ V[X] = \int_{0}^{dt} w \]

where

\[ w \sim N(0, q) \]

and “white.”
Control of Pose Uncertainty

Measurements $\Rightarrow$ pose information
Sensor Measurements

Acoustic Sdev (2x2): +2.00 +2.00 +1.00
Break

(15 Minutes)
Traditional Approaches

(Gary Bishop)
Traditional Solution Methods

- Simple problem: Determine pose given sensor readings.
- Linear algebra taught us about \( N \) equations in \( N \) unknowns
- Each equation is a constraint
- 3 DOF \( \rightarrow \) 3 constraints & 6 DOF \( \rightarrow \) 6 constraints
- Unfortunately often non-linear constraints often with multiple solutions
Range Tracker

Intersect 3 spheres

\[ (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r_0^2 \]
\[ (x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = r_1^2 \]
\[ (x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2 = r_2^2 \]

Unfortunately, the solution is ugly...
Simplify

a. put mike 0 at origin
b. put mike 1 out X axis 1 unit
c. put mike 2 out Y axis 1 unit
d. all 3 mikes in Z=0 plane

Usual coordinate transform to convert to “lab” coordinates
Simpler Range Equations

\[ x^2 + y^2 + z^2 = r_0^2 \]
\[ (x - 1)^2 + y^2 + z^2 = r_1^2 \]
\[ x^2 + (y - 1)^2 + z^2 = r_2^2 \]

Note ambiguity

\[ x = \frac{r_0^2 - r_1^2 + 1}{2} \]
\[ y = \frac{r_0^2 - r_2^2 + 1}{2} \]
\[ z = \pm \sqrt{r_0^2 - x^2 - y^2} \]
Optical with fixed 1D sensors

For example, 1D CCD with razor blade casting a shadow

- Calibrate to determine 3D plane equation from sensor reading (non-trivial)
- For each sensor reading, write a linear equation relating unknown x, y, z to plane
- Solve the system of equations for x, y, z
Solve...

\[ A_i x + B_i y + C_i z = D_i \]

\[ M = \begin{bmatrix}
A_1 & B_1 & C_1 \\
A_2 & B_2 & C_2 \\
A_3 & B_3 & C_3
\end{bmatrix} \]

\[ M \cdot \begin{bmatrix}
x \\
y \\
z
\end{bmatrix} = \begin{bmatrix}
D_1 \\
D_2 \\
D_3
\end{bmatrix} \]
Optical with fixed 2D sensors

For example, two video cameras looking at LEDs on the user.

- Could treat as four 1-D sensors
- OR
- Calibrate to get ray equation from u,v
- Rays won’t intersect!
- Minimize distance between them
Set up equations

Ray equations

\[ A_1 = C_1 + s_1 D_1 \]
\[ A_2 = C_2 + s_2 D_2 \]

Distance

\[ \|(C_2 + s_2 D_2) - (C_1 + s_1 D_1)\| \]

Minimum distance line must be perpendicular to both rays, so...

\[ [(C_2 + s_2 D_2) - (C_1 + s_1 D_1)] \cdot D_1 = 0 \]
\[ [(C_2 + s_2 D_2) - (C_1 + s_1 D_1)] \cdot D_2 = 0 \]
Solve

Distance out each ray to closest point

Halfway between

\[ s_1 = \frac{(B \cdot D_1) - (D_2 \cdot D_1)(B \cdot D_2)}{1 - (D_1 \cdot D_2)^2} \]

\[ s_2 = \frac{(D_1 \cdot D_2)(B \cdot D_1) - (B \cdot D_2)}{1 - (D_1 \cdot D_2)^2} \]

\[ \tilde{p} = \frac{(C_1 + s_1D_1) + (C_2 + s_2D_2)}{2} \]

B is the baseline
Stochastic Approaches

(Greg Welch)
Motivation

• **Take into account**
  • Stochastic nature of sensor signals
  • Varying amounts of sensor information
  • Model of user motion

• **Combine sensor/measurement information**
  • Combat (otherwise growing) pose uncertainty
  • Fuse information from heterogeneous sensors
State-Space Models

Begin with **difference equation** for process

\[ y_{k+1} = a_{0,k} y_k + \ldots + a_{n-1,k} y_{k-n+1} + u_k \]

Re-write as

\[
\begin{bmatrix}
    y_{k+1} \\
    y_k \\
    y_{k-1} \\
    \vdots \\
    y_{k-n+2}
\end{bmatrix}
= \begin{bmatrix}
    a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\
    1 & 0 & \cdots & 0 & 0 \\
    0 & 1 & \cdots & 0 & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
    y_k \\
    y_{k-1} \\
    y_{k-2} \\
    \vdots \\
    y_{k-n+1}
\end{bmatrix}
+ \begin{bmatrix}
    1 \\
    0 \\
    0 \\
    \vdots \\
    1
\end{bmatrix} u_k
\]
State-Space Models

\[ \bar{x}_{k+1} \equiv \begin{bmatrix} y_{k+1} \\ y_k \\ y_{k-1} \\ \vdots \\ y_{k-n+2} \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_k \\ y_{k-1} \\ y_{k-2} \\ \vdots \\ y_{k-n+1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u \]

\[ \bar{x}_i \quad G \]

\[ \bar{x}_{k+1} = A\bar{x}_k + Gu_k \]

\[ \bar{y}_k = H\bar{x}_k \]
Observer Design Problem

\[
\bar{x}_k = A\bar{x}_{k-1} + Gu_{k-1} \\
\bar{z}_k = H\bar{x}_k + \bar{v}_k
\]
Optimal Estimation

\[ c = \int_{0}^{T} \text{cost}(\bar{a}(t), \bar{b}(t), t) \, dt \]

Integral of Absolute Value of Error (IAE)

\[ \text{cost} = |\bar{a} - \bar{b}| \]

Integral of Square of Error (ISE)

\[ \text{cost} = (\bar{a} - \bar{b})^2 \]
The Kalman Filter

R.E. Kalman, 1960

- Recursive optimal estimator
  - Minimum variance of error
- Versatile & robust
  - Estimation
  - Sensor fusion
- Robotics, navigation, computer vision, economics, ...

  - Java-Based Learning Tool, books, papers, etc.
  - ACM SIGGRAPH 2001 tutorial (earlier today)
transition

uncertainty

\[
\begin{align*}
\bar{x}_k^- &= A\bar{x}_{k-1} \\
\bar{P}_k^- &= AP_{k-1}A^T + Q
\end{align*}
\]
\[
\bar{x}_k = \bar{x}_k^- + K(z_k - H\bar{x}_k^-)
\]
\[
P_k = (I - KH)P_k^-
\]
\[
K = P_k^-H^T(HP_k^-H^T + R)^{-1}
\]

“denominator”
(measurement space)
Hybrid Systems and Multi-Sensor Fusion

- Magnetic coil
- Accelerometer
- Gyroscope
- Magnetic coil
- Accelerometer
- Gyroscope
Incremental Estimation
A Single Constraint at a Time

\[ C = \text{constraints needed for a unique solution} \]
\[ M = \text{constraints used per estimate update} \]

\[ M > C \]
\[ M = C \]
\[ 1 < M < C \]
\[ M = 1 \]
Benefits of SCAAT Approach

*Purposefully using minimal constraints*

- Avoid *simultaneity assumption*
- Temporal improvements
- Simplicity and flexibility
- Online source/sensor *autocalibration*
- Can be applied to *virtually any tracking system*

  - Measurement model for each type of sensor
  - Dynamic model for user motion (possibly trivial)
Bishop, 1982

Diagram showing the relationship between change, speed, complexity, and levels of low and high.
Error Sources

(Greg Welch)
Error in Head Pose

• Hard to fool “mother nature”
  • Lifetime of visual experience and expectations
  • Visual-proprioceptive conflicts
  • Virtual-real misregistration

• What to do?
  • Some amount of error is unavoidable
  • Understand sources and seek to minimize
Error Classification

- **Pose estimate life cycle**
  - Noisy sensor measurement → Estimate → Transport → Transform → Display

- **Two primary classes of error**
  - Static (spatial)
  - Delay-induced (temporal)
Static Measurement Error

- **Static field distortion**
  - Repeatable error in the measurement data
  - “Bias” that might be corrected via calibration

- **Random noise or jitter**
  - Non-repeatable error
  - Random (electrical) noise such as described earlier
  - Often dependent on the current pose
Pose-Dependent Noise (Example)

Baseline noise $\xi_0$ and coefficients $a$, $b$, and $c$ were determined off line.

$$\sqrt{\xi_c} = \frac{\sqrt{\xi_0} d_b^2}{a \alpha_b^3 + b \alpha_b^2 + c \alpha_b + 1}$$
Delay-Induced Error

• **Measurement validity**
  • Good at sample time, then old (aging)
  • Finite, non-zero sample time
  • Old sample → misregistration

• **Motion prediction**
  • Measure where you *are*, but *want* where you *will be*
  • Later w/ Bishop
The Simultaneity Assumption

Moderate arm & wrist translation

\[
\frac{1}{2} \text{[s]} \cdot 3 \text{[m/s]} \cdot 20-80 \text{[ms]} \Rightarrow 1-10 \text{[cm]}
\]

Moderate head rotation

\[
\frac{1}{2} \text{[s]} \cdot 180 \text{[°/s]} \cdot 20-80 \text{[ms]} \Rightarrow 6-25 \text{[cm]}
\]

(at arm’s length)
First-Order Dynamic Error

\[
\varepsilon_{\text{dyn},\theta} = \dot{\theta} \Delta t
\]

\[
\varepsilon_{\text{dyn},x} = \dot{x} \Delta t
\]

Tracker + graphics pipeline latency

Instantaneous velocities
Synchronization Delay

A.k.a. *phase delay* or *rendezvous delay*
Pose Estimate Timeline

- Client Buffer
- Write Buffer
- Network
- Read Buffer
- Write Buffer
- Estimate
- Sample Sensor
- User Motion

Timeline:
- $t_m$
- $\Delta t_{crb}$
- $\tau_{net}$
- $\Delta t_{srb}$
- $\Delta t_e$
- $\Delta t_{ss}$

Time Scale: $t_{m'}$
Total Tracker Latency

\[ \Delta t_m = t_{m'} - t_m \]

\[ = \Delta t_{ss} + \Delta t_e + \tau_e + \Delta t_{srb} + \tau_{net} + \Delta t_{crb} \]

\[ = \frac{1}{2r_{ss}} + \frac{1}{2r_e} + \tau_e + \frac{1}{2r_{srb}} + \tau_{net} + \frac{1}{2r_{crb}} \]

- Sample the sensor(s)
- Estimate the cycle
- Server buffer synchronization
- Network transport
- Client buffer synchronization
- Estimate the pose
Total Tracker Error

\[ \varepsilon_\theta \approx \varepsilon_{\text{stat}, \theta} + \varepsilon_{\text{sa}, \theta} + \dot{\theta}(\Delta t_m + \Delta t_g) \]

\[ \varepsilon_x \approx \varepsilon_{\text{stat}, x} + \varepsilon_{\text{sa}, x} + \dot{x}(\Delta t_m + \Delta t_g) \]
Closing (Error Sources)

• Did I mention error magnification?
• Consider the technology
  • Understand its limitations
  • Stay within the envelope
• Prediction (next)
Motion Prediction

(Gary Bishop)
Motion Prediction

End-to-end delay
- hurts in VR / hurts worse in AR
- sources
  - time to measure pose
  - delay in communicating pose
  - application response to change
  - graphics update
  - display refresh
What to do about delay?

1. Monitor
2. Minimize
3. Mitigate

Latency is not *only* a tracker problem.
But mitigation is best handled at the tracker.
Can prediction help?

Blue $\rightarrow$ no prediction
Red $\rightarrow$ w/out inertial
Green $\rightarrow$ w/ inertial
Can prediction help?

![Graph showing angular error in degrees over time. The graph compares 'No prediction', 'Prediction without inertial', and 'Prediction with inertial'. The Swing motion dataset, 60 ms prediction distance is indicated.]
Limits to prediction

Prediction error grows quadratically with motion bandwidth and prediction interval.
Prediction ideas

- Extrapolate past behavior to the future
- The more history the better
- Correlations in the users coordinate frame
- Inertial sensors help
- Monitor $|\text{predicted} - \text{actual}|$ for tuning
- Use image shifting to reduce jitter
Conclusions

(Gary Bishop)
Final Thoughts

• **No silver bullet**
  • Tracking anywhere for any purpose is a dream

• **No free lunch, only tradeoffs**
  • Energy / Accuracy / Bandwidth / Latency / Noise

• **No end in sight**
  • Lots of possibilities for interesting work
  • ReActor not based on any of the principles described here
Resources

- [http://www.cs.unc.edu/~welch/kalman](http://www.cs.unc.edu/~welch/kalman)
- check out Course 8 notes
- Dozens of books on KF, here are a few
  - “Optimal Estimation with an ...” by Lewis
  - “Introduction to Random Signals...” by Brown
  - “Kalman Filtering Theory and Practice” by Grewal
- Beginnings of a tracking bibliography
Exhibits we’re going to check out

- 3rd Tech (cool demo)
- 5DT
- Ascension (ReActor is a new method)
- InterSense
- Measurand
- MetaMotion / PhoeniX Technologies / Vicon
- Polhemus