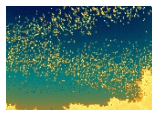
Tracking II

Tuesday, Dec 2

Kristen Grauman
UT-Austin



Announcements

- Reminder: check eGradebook to see all your scores
- Thursday: course recap and exam review
- **Pset 4 hardcopy turnin: two options**
  - Bring to class this Thursday (last class day), or
  - Anytime after Thursday's class, drop in drop box on Taylor first floor in front of undergrad advising office
    - write "CS378 Computer Vision" on top of your hardcopy

Outline

- Last time:
  - Using optical flow (dense motion estimates) to recognize activities
  - Tracking
    - Tracking as inference
    - Linear models of dynamics
    - Kalman filters
- Today:
  - Kalman filter recap, updates for n-d
  - Limitations of Kalman filtering
  - Other issues in tracking

Last time: Linear dynamic model

- Describe the *a priori* knowledge about
  - System dynamics model: represents evolution of state over time, with noise.

$$\mathbf{x}_t \sim N(\mathbf{D}\mathbf{x}_{t-1}; \mathbf{\Sigma}_d)$$

  - Measurement model: at every time step we get a noisy measurement of the state.

$$\mathbf{y}_t \sim N(\mathbf{M}\mathbf{x}_t; \mathbf{\Sigma}_m)$$

Last time: Kalman filter

Know corrected state from previous time step, and all measurements up to the current one → Predict distribution over next state.

Receive measurement

Know prediction of state, and next measurement → Update distribution over current state.

Time update ("Predict")

Measurement update ("Correct")

$P(X_t | y_0, \ldots, y_{t-1})$

Mean and std. dev. of predicted state:

μ_t^-, σ_t^-

Time advances: t++

$P(X_t | y_0, \ldots, y_t)$

Mean and std. dev. of corrected state:

μ_t^+, σ_t^+

1D Kalman filter: prediction vs. correction

$$\mu_t^+ = \frac{\mu_t^- \sigma_m^2 + m y_t (\sigma_t^-)^2}{\sigma_m^2 + m^2 (\sigma_t^-)^2} \qquad (\sigma_t^+)^2 = \frac{\sigma_m^2 (\sigma_t^-)^2}{\sigma_m^2 + m^2 (\sigma_t^-)^2}$$

- What if there is no prediction uncertainty $(\sigma_t^- = 0)$?

$$\mu_t^+ = \mu_t^- \qquad (\sigma_t^+)^2 = 0$$

The measurement is ignored!

- What if there is no measurement uncertainty $(\sigma_m = 0)$?

$$\mu_t^+ = \frac{y_t}{m} \qquad (\sigma_t^+)^2 = 0$$

The prediction is ignored!

Source: Lana Lazebnik

Kalman filter: General case (> 1dim)

What if state vectors have more than one dimension?

PREDICT

$$x_t^- = D_t x_{t-1}^+$$

$$\Sigma_t^- = D_t \Sigma_{t-1}^+ D_t^T + \Sigma_{d_t}$$

CORRECT

"residual"

$$K_t = \Sigma_t^- M_t^T \left(M_t \Sigma_t^- M_t^T + \Sigma_{m_t} \right)^{-1}$$

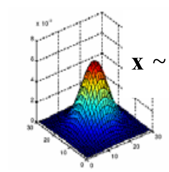
$$x_t^+ = x_t^- + K_t \boxed{\left(y_t - M_t x_t^- \right)}$$

$$\Sigma_t^+ = \left(I - K_t M_t \right) \Sigma_t^-$$

Less weight on residual as a priori estimate error covariance approaches 0.

Kalman filter: pros and cons

- Gaussian densities, linear dynamic model:
  - \+ Simple updates, compact and efficient
  - – But, restricted class of motions defined by linear model
  - – *Unimodal* distribution = only single hypothesis



$$\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

When is a single hypothesis too limiting?

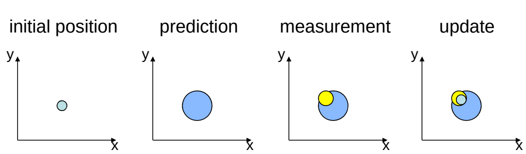
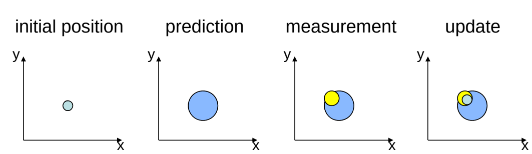


Figure from Thrun & Kosecka

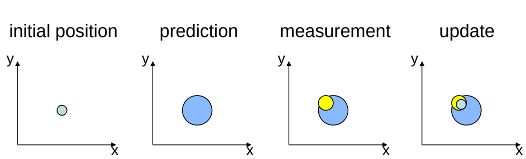
When is a single hypothesis too limiting?



Consider this example: say we are tracking the face on the right using a skin color blob to get our measurement.

Video from Jojic & Frey

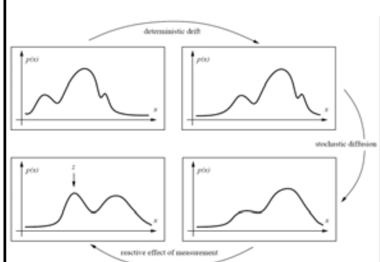
When is a single hypothesis too limiting?

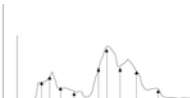


Consider this example: say we are tracking the face on the right using a skin color blob to get our measurement.

Video from Jojic & Frey

Alternative: particle-filtering and non-Gaussian densities



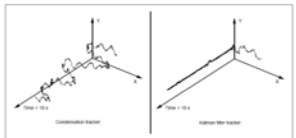


- Can represent distribution with set of weighted samples ("*particles*")
- Allows us to maintain **multiple hypotheses.**

For details: *CONDENSATION -- conditional density propagation for visual tracking*, by Michael Isard and Andrew Blake, Int. J. Computer Vision, 29, 1, 5--28, (1998)

Alternative: particle-filtering and non-Gaussian densities

Monitor is a distractor, multiple hypotheses necessary.



Kalman filter fails once it starts tracking the monitor.

http://www.robots.ox.ac.uk/~vdg/dynamics.html
Visual Dynamics Group, Dept. Engineering Science, University of Oxford, 1998

Tracking: issues

- Initialization
- Data association
- Multiple tracked objects
- Deformable and articulated objects
- Constructing accurate models of dynamics
- Drift

Tracking: issues

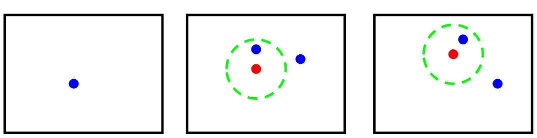
- Initialization
  - Often done manually
  - Background subtraction, detection can also be used
- Data association, multiple tracked objects
  - Occlusions

Data association

- We've assumed entire measurement (**y**) was cue of interest for the state
- But, there are typically uninformative measurements too–clutter.
- **Data association**: task of determining which measurements go with which tracks.

Data association

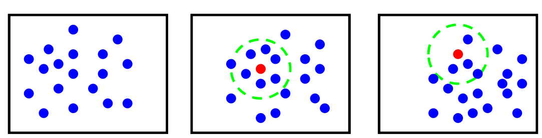
- Simple strategy: only pay attention to the measurement that is "closest" to the prediction



Source: Lana Lazebnik

Data association

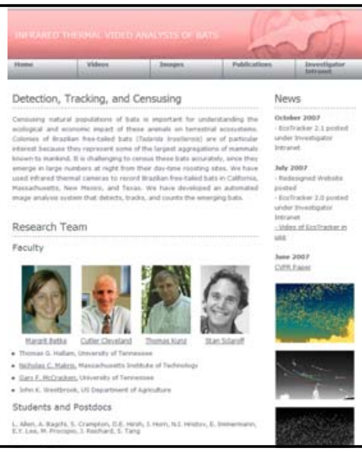
- Simple strategy: only pay attention to the measurement that is "closest" to the prediction



Doesn't always work…
Alternative: keep track of multiple hypotheses at once.

Source: Lana Lazebnik

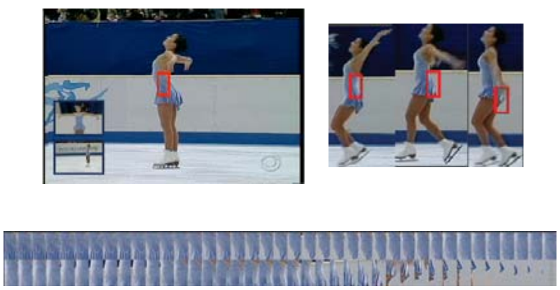
- http://www.cs.bu.edu/~betke/research/bats/



Tracking: issues

- Initialization
  - Often done manually
  - Background subtraction, detection can also be used
- Data association, multiple tracked objects
  - Occlusions
- Deformable and articulated objects
- Constructing accurate models of dynamics
  - e.g., parameters for a linear dynamics model
- Drift
  - Accumulation of errors over time

Drift



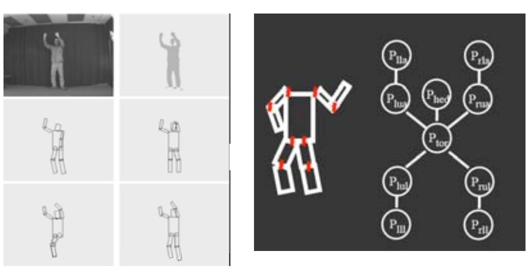
D. Ramanan, D. Forsyth, and A. Zisserman. Tracking People by Learning their Appearance. PAMI 2007.

Tracking people by learning their appearance

- Person model = appearance + structure (+ dynamics)
- Structure and dynamics are generic, appearance is person-specific
- Trying to acquire an appearance model "on the fly" can lead to drift
- Instead, can use the whole sequence to initialize the appearance model and then keep it fixed while tracking
- Given strong structure and appearance models, tracking can essentially be done by repeated detection (with some smoothing)

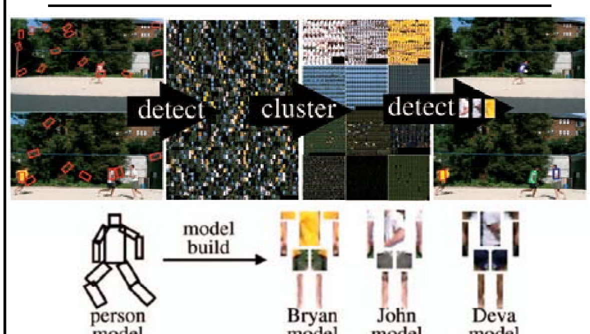
D. Ramanan, D. Forsyth, and A. Zisserman. Tracking People by Learning their Appearance. PAMI 2007.

Tracking people by learning their appearance



Use a **part-based model** to encode part appearance + relative geometry.

Bottom-up initialization: Clustering



D. Ramanan, D. Forsyth, and A. Zisserman. Tracking People by Learning their Appearance. PAMI 2007.

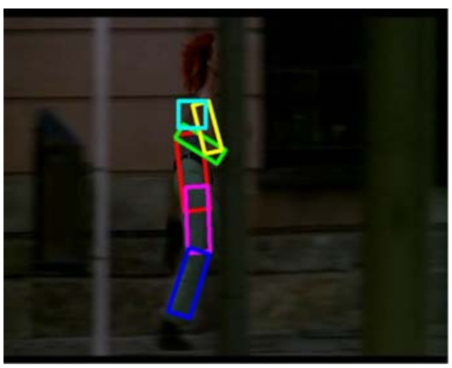
Top-down initialization: Exploit “easy” poses

D. Ramanan, D. Forsyth, and A. Zisserman. Tracking People by Learning their Appearance. PAMI 2007.

Example results

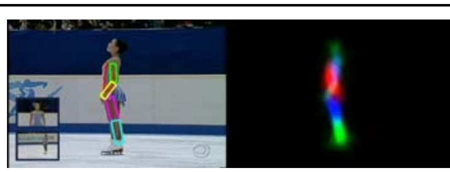
http://www.ics.uci.edu/~dramanan/papers/pose/index.html

Example results



Example results

Example results



http://www.ics.uci.edu/~dramanan/papers/pose/index.html

Tracking: summary

- Tracking as inference
  - Goal: estimate posterior of object position given measurement
  - Know where to look, can survive even with poor measurements
- Linear models of dynamics
  - Represent state evolution and measurement models
- Kalman filters
  - Recursive prediction/correction updates to refine measurement
  - Single hypothesis can be limiting → alternative models use non-Gaussian distributions
- Drift: as error accumulates we may gradually start tracking something else.
  - Tracking via detection one way to mitigate drift (though lose out on prediction help)