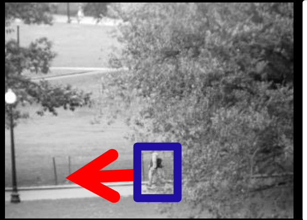
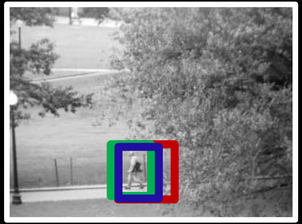
CS4495/6495
Introduction to Computer Vision

7B-L1 *Tracking as inference*

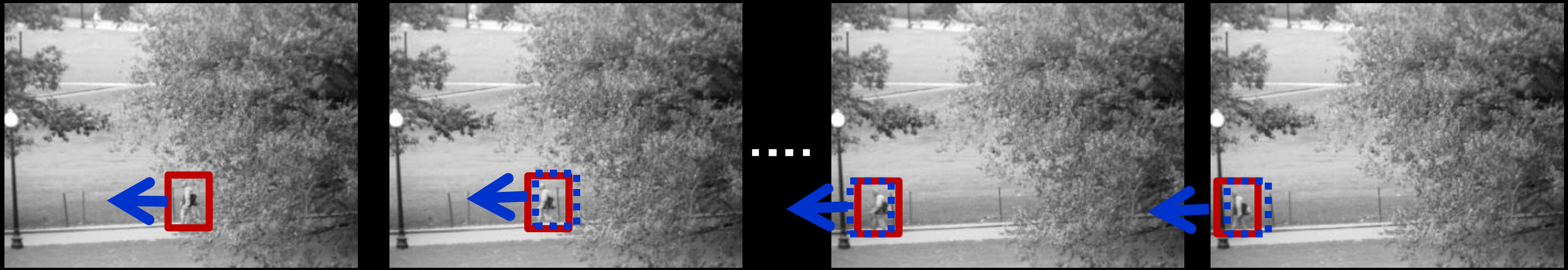


Time t



Time t+1

Detection vs. tracking



Tracking: We *predict* the new location of the object in the next frame using *estimated dynamics*. Then we *update* based upon measurements.

Tracking as inference

Hidden state (X): True parameters we care about

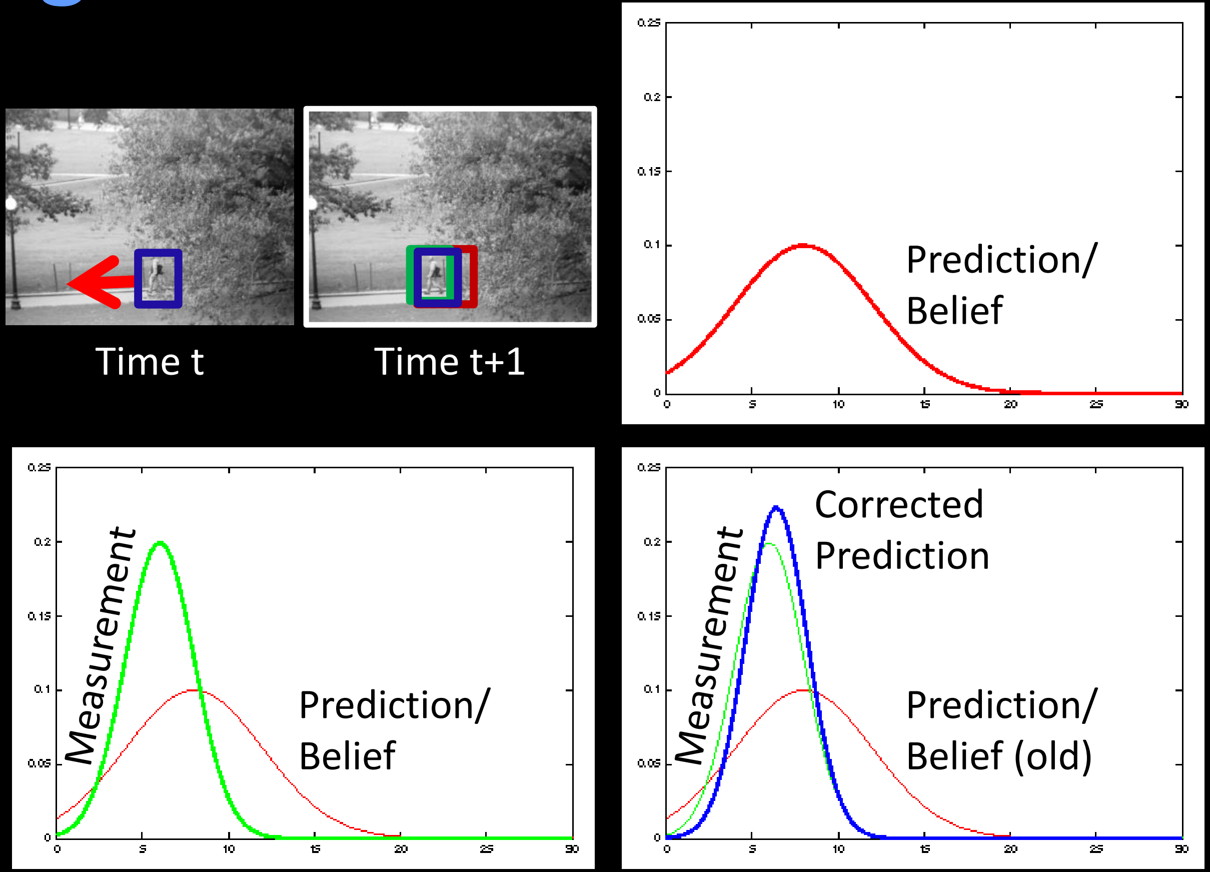
Measurement (Y): Noisy observation of underlying state

At each time step t, state changes (from X_{t-1} to X_t), and we get a new observation Y_t

Our goal: Recover most likely state X_t given

- All observations seen so far
- Knowledge about dynamics of state transitions

Tracking as inference: Intuition



Steps of tracking

Prediction: What is the next state of the object given past measurements?

$$P\left(X_t \middle| Y_0 = y_0, \ldots, Y_{t-1} = y_{t-1}\right)$$

Correction: Compute an updated estimate of the state from prediction and measurements

$$P\left(X_t \middle| Y_0 = y_0, \ldots, Y_{t-1} = y_{t-1}, \boxed{Y_t = y_t}\right)$$

*Tracking: The process of propagating this **posterior** distribution of state given measurements across time*

Simplifying assumptions

Only the immediate past matters

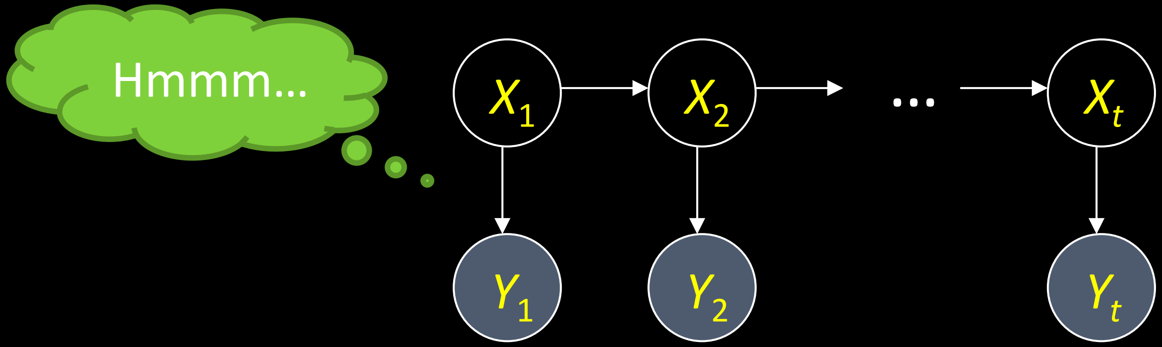
$$P\left(X_t \middle| X_0, \ldots, X_{t-1}\right) = \boxed{P\left(X_t \middle| X_{t-1}\right)}$$

dynamics model

Measurements depend only on the current state

$$P\left(Y_t \middle| X_0, Y_0 \ldots, X_{t-1}, Y_{t-1}, X_t\right) = \boxed{P\left(Y_t \middle| X_t\right)}$$

observation model



Tracking as induction

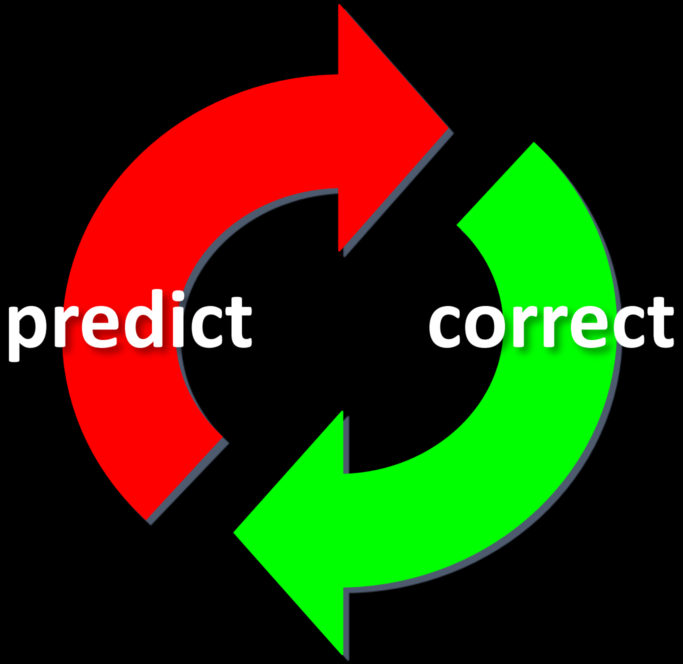
Base case:

- Assume we have some initial prior that predicts state in the absence of any evidence: $P(X_0)$
- At the first frame, *correct* this, given value of $Y_0 = y_0$

Tracking as induction

Given corrected estimate for frame t:

- Predict for frame $t + 1$
- Correct for frame $t + 1$



Tracking as induction

Base case:

- Assume we have some initial prior that predicts state in the absence of any evidence: $P(X_0)$
- At the first frame, *correct* this, given value of $Y_0 = y_0$

$$P(X_0 \mid Y_0 = y_0) = \frac{P(y_0 \mid X_0)P(X_0)}{P(y_0)} \propto P(y_0 \mid X_0)P(X_0)$$

Bayes rule

Prediction

Given: $P(X_{t-1}|y_0, \dots, y_{t-1})$

Guess: $P(X_t|y_0, \dots, y_{t-1})$

Prediction

Given: $P(X_{t-1}|y_0, \ldots, y_{t-1})$

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

$$= \int P\left(X_t, X_{t-1} \middle| y_0, \ldots, y_{t-1}\right) dX_{t-1}$$

Law of total probability - Marginalization

Prediction

Given: $P(X_{t-1}|y_0, \dots, y_{t-1})$

Guess: $P(X_t|y_0, \dots, y_{t-1})$

$$= \int P\left(X_t, X_{t-1} \middle| y_0, \dots, y_{t-1}\right) dX_{t-1}$$

$$= \int P\left(X_t \mid X_{t-1}, y_0, \dots, y_{t-1}\right) P\left(X_{t-1} \mid y_0, \dots, y_{t-1}\right) dX_{t-1}$$

Conditioning on X_{t-1} [recall $P(A, B) = P(A|B)P(B)$]

Prediction

Given: $P(X_{t-1}|y_0, \ldots, y_{t-1})$

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

$$= \int P\left(X_t, X_{t-1} \middle| y_0, \ldots, y_{t-1}\right) dX_{t-1}$$

$$= \int P\left(X_t \mid X_{t-1}, y_0, \ldots, y_{t-1}\right) P\left(X_{t-1} \mid y_0, \ldots, y_{t-1}\right) dX_{t-1}$$

$$= \int P\left(X_t \mid X_{t-1}\right) P\left(X_{t-1} \mid y_0, \ldots, y_{t-1}\right) dX_{t-1}$$

Independence assumption

Correction

Given predicted value $P(X_t|y_0, \dots, y_{t-1})$ and y_t

Compute $P(X_t|y_0, \dots, y_t)$

Correction

Given predicted value $P(X_t|y_0, \ldots, y_{t-1})$ and y_t

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

$$= \frac{P(y_t \mid X_t, y_0, \ldots, y_{t-1}) P(X_t \mid y_0, \ldots, y_{t-1})}{P(y_t \mid y_0, \ldots, y_{t-1})}$$

Bayes rule: $\mathrm{P(A|B)} = \dfrac{\mathrm{P(B|A)P(A)}}{\mathrm{P(B)}}$

Correction

Given predicted value $P(X_t|y_0, \ldots, y_{t-1})$ and y_t

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

$$= \frac{P(y_t \mid X_t, y_0, \ldots, y_{t-1}) P(X_t \mid y_0, \ldots, y_{t-1})}{P(y_t \mid y_0, \ldots, y_{t-1})}$$

$$= \frac{P(y_t \mid X_t) P(X_t \mid y_0, \ldots, y_{t-1})}{P(y_t \mid y_0, \ldots, y_{t-1})}$$

Independence assumption

(observation y_t depends only on state X_t)

Correction

Given predicted value $P(X_t|y_0, \dots, y_{t-1})$ and y_t

Compute $P(X_t|y_0, \dots, y_t)$

$$= \frac{P(y_t \mid X_t, y_0, \dots, y_{t-1}) P(X_t \mid y_0, \dots, y_{t-1})}{P(y_t \mid y_0, \dots, y_{t-1})}$$

$$= \frac{P(y_t \mid X_t) P(X_t \mid y_0, \dots, y_{t-1})}{P(y_t \mid y_0, \dots, y_{t-1})}$$

$$= \frac{P(y_t \mid X_t) P(X_t \mid y_0, \dots, y_{t-1})}{\int P(y_t \mid X_t) P(X_t \mid y_0, \dots, y_{t-1}) dX_t}$$

Really a normalization

Conditioning on X_t

Summary: Prediction and correction

Prediction:

$$P\left(X_t \mid y_0,\ldots,y_{t-1}\right) = \int \underbrace{P\left(X_t \mid X_{t-1}\right)}_{\text{dynamics model}} \underbrace{P\left(X_{t-1} \mid y_0,\ldots,y_{t-1}\right)}_{\text{corrected estimate from previous step}} dX_{t-1}$$

Summary: Prediction and correction

Correction:

$$P\left(X_t \mid y_0,\ldots,y_{t-1},y_t\right)=\frac{\overbrace{P\left(y_t \mid X_t\right)}^{\text{observation model}}\overbrace{P\left(X_t \mid y_0,\ldots,y_{t-1}\right)}^{\text{predicted estimate}}}{\int P\left(y_t \mid X_t\right)P\left(X_t \mid y_0,\ldots,y_{t-1}\right)dX_t}$$