Face/Head Tracking

Ying Wu
yingwu@ece.northwestern.edu

Electrical and Computer Engineering
Northwestern University, Evanston, IL
http://www.ece.northwestern.edu/~yingwu

Tracking Heads?

Courtesy of Y. Wu, 2001

- The task:

  Localize faces and track them in image sequences

- Challenges:

  Lighting, occlusion, rotation, etc.

Outline

- ✓ Motivation
- ✓ What is tracking?
- ✓ One solution (Birchfield_CVPR98)
- ✓ Other methods and open issues

Motivation

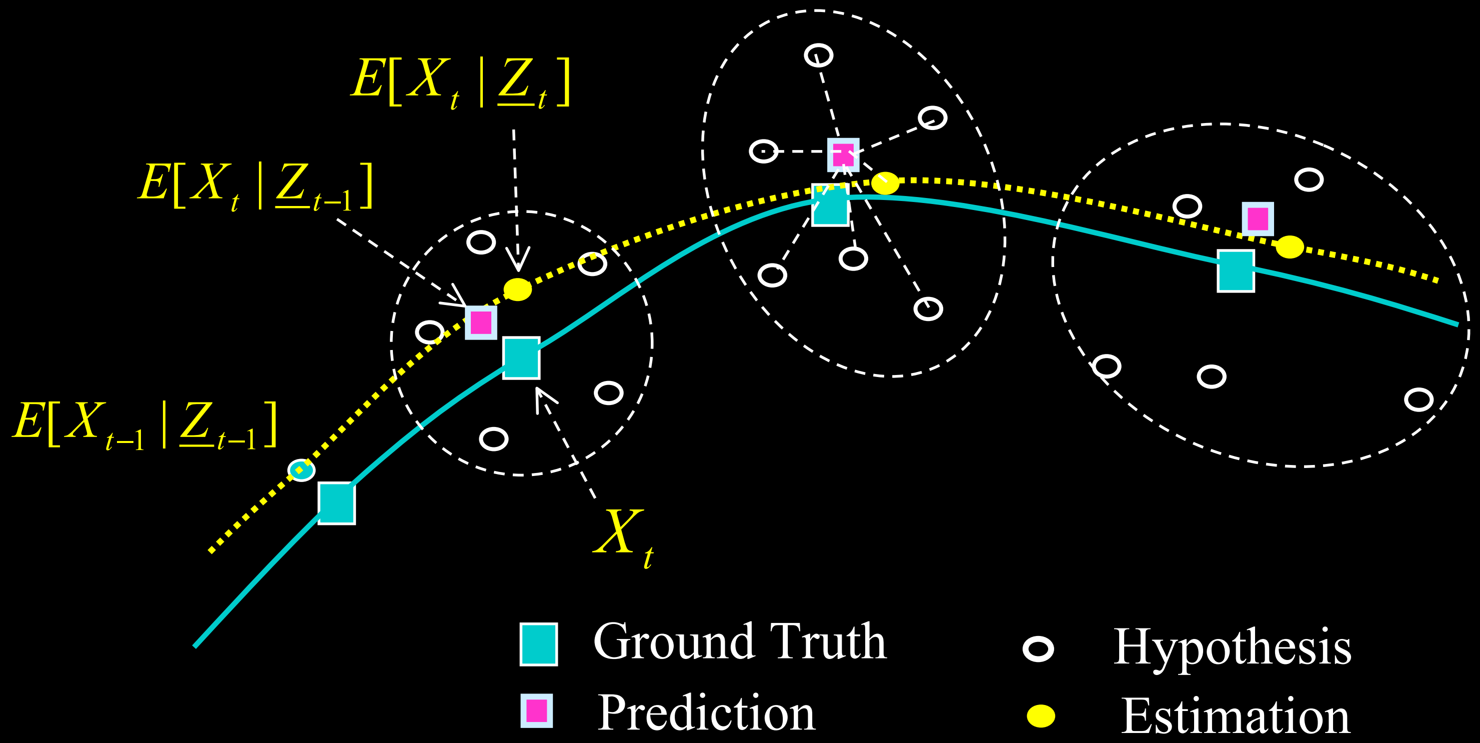
- Why tracking?
  - The complexity of face detection
    - ✓ scan all the pixel positions and several scales
  - The limitation of face detection
    - ✓ hard to handle out-of-plane rotation
  - Can we maintain the identity of the faces?
    - ✓ although face recognition is the ultimate solution for this, we may not need it, if not necessary
- Objectives
  - fast (frame-rate) face/head localization
  - handle 360^o out-of-plane rotation

Visual Tracking

Four Elements

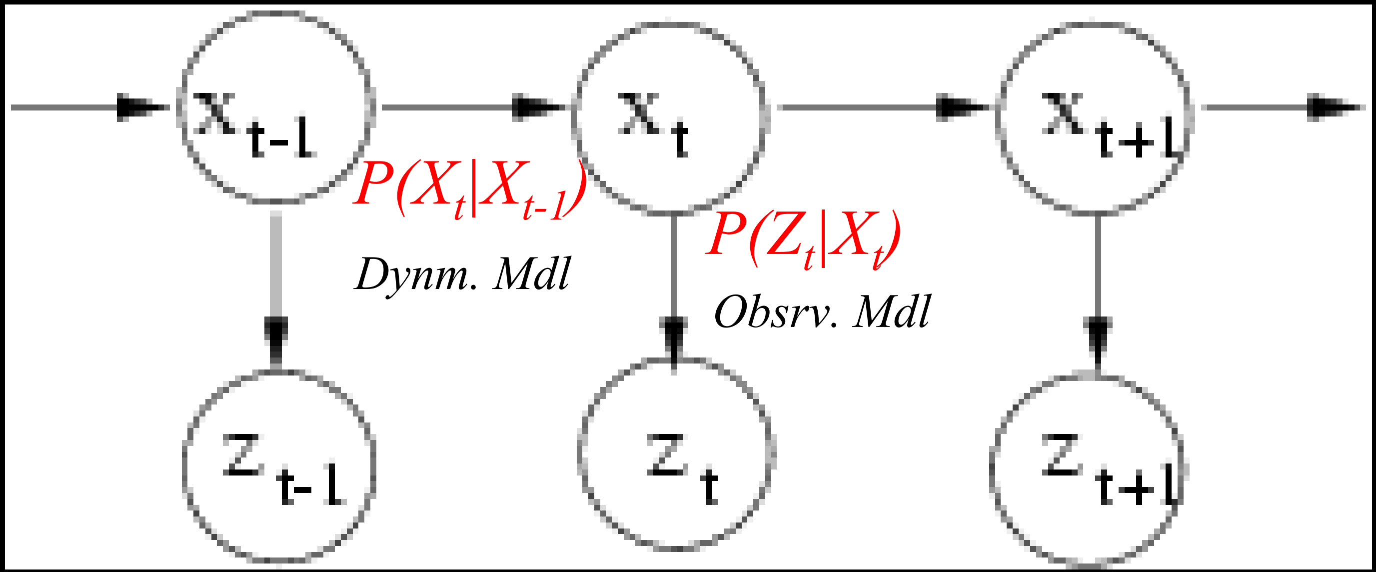
- Infer target states in video sequences
- Target states vs. image observations
- Visual cues and modalities
- Four elements
  - *Target representation* X
  - *Observation representation* Z
  - *Hypotheses measurement* $p(Z_t|X_t)$
  - *Hypotheses generating* $p(X_t|X_{t-1})$

Visual Tracking



$$(E[X_{t-1} \mid \underline{Z}_{t-1}, Z_t]) \Rightarrow E[X_t \mid \underline{Z}_t]$$

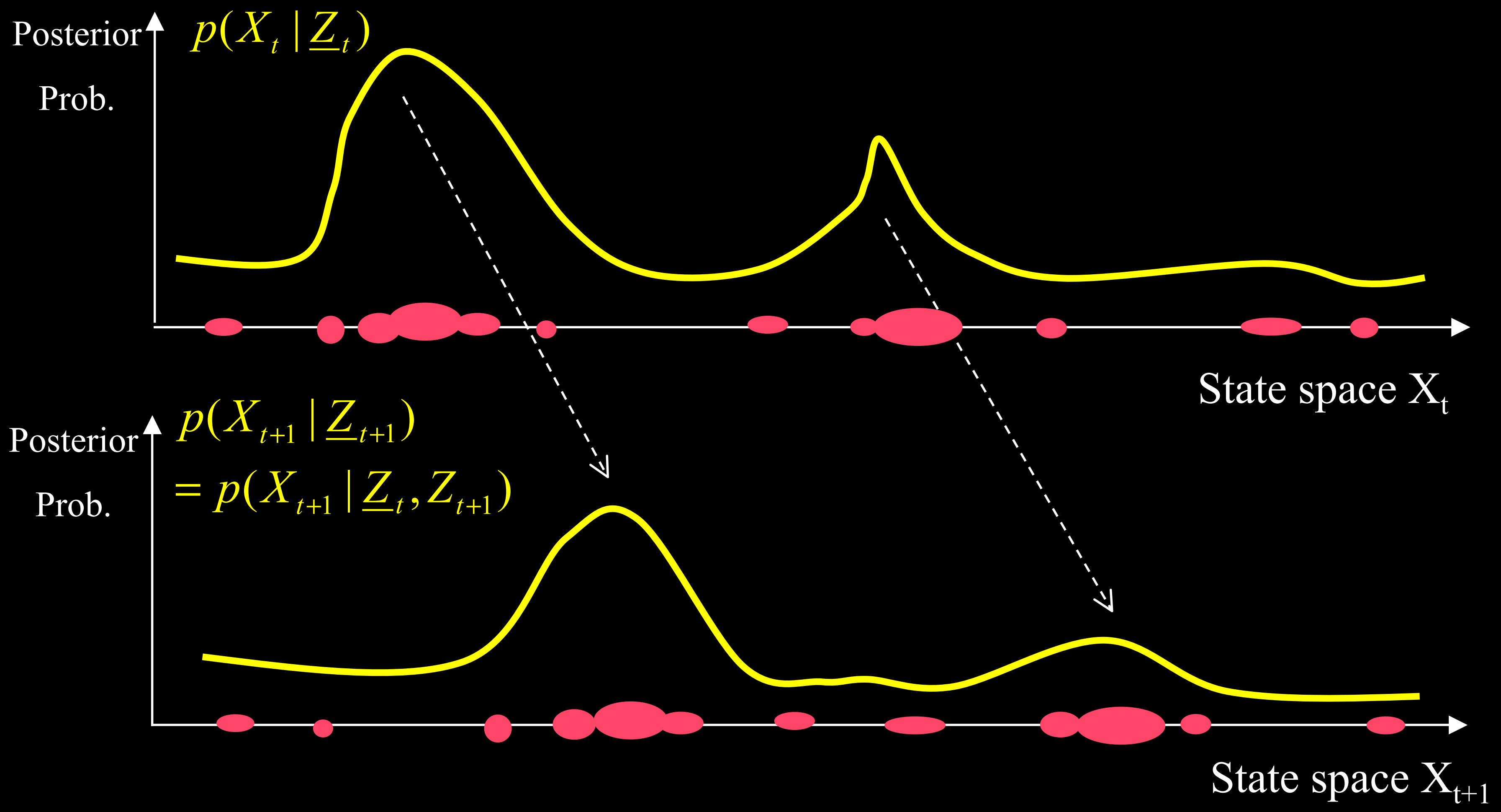
Formulating Visual Tracking



$$p(X_{t+1} \mid \underline{Z}_{t+1}) \propto p(Z_{t+1} \mid X_{t+1})\, p(X_{t+1} \mid \underline{Z}_t)$$

$$p(X_{t+1} \mid \underline{Z}_t) = \int p(X_{t+1} \mid X_t)\, p(X_t \mid \underline{Z}_t)\, dX_t$$

Tracking as Density Propagation



One Solution

(Birchfield_CVPR98)

- Framework
- Search strategy
- Edge cue
- Color cue

Framework

- s = (x,y,σ)
- Tracking is treated as a local search based on the prediction

$$\mathbf{s}^* = \arg\max_{\mathbf{s}_i \in S} \{\bar{\phi}_g(\mathbf{s}_i) + \bar{\phi}_c(\mathbf{s}_i)\},$$

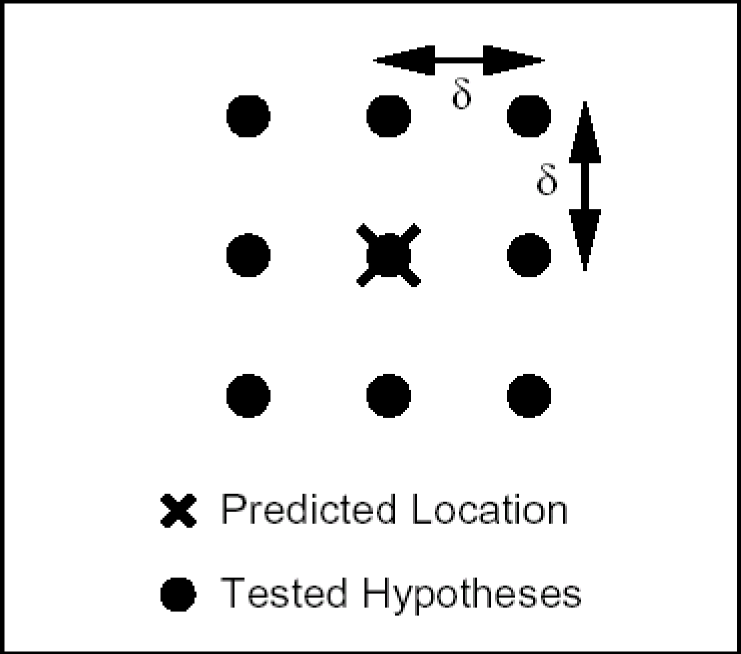
hypotheses

Edge matching score

color matching score

Search Strategy

- Local exhaustive search



δ is the search step size

- Do you have better ideas?

Edge Cue

- Method I

$$\phi_g(\mathbf{s}) = \frac{1}{N_\sigma} \sum_{i=1}^{N_\sigma} |\mathbf{g_s}(i)|.$$

The the magnitude of the gradient at perimeter pixel i of the ellipse s.

# of pixels on the perimeter of the ellipse

- Method II

$$\phi_g(\mathbf{s}) = \frac{1}{N_\sigma} \sum_{i=1}^{N_\sigma} |\mathbf{n}_\sigma(i) \cdot \mathbf{g_s}(i)|.$$

unit vector normal to the ellipse at pixel i.

- Which is better?

Normalization

$$\bar{\phi}_g(\mathbf{s}) = \frac{\phi_g(\mathbf{s}) - \min_{\mathbf{s}_i \in S} \phi_g(\mathbf{s}_i)}{\max_{\mathbf{s}_i \in S} \phi_g(\mathbf{s}_i) - \min_{\mathbf{s}_i \in S} \phi_g(\mathbf{s}_i)}.$$

- Why do we need normalization?
- How good is it?

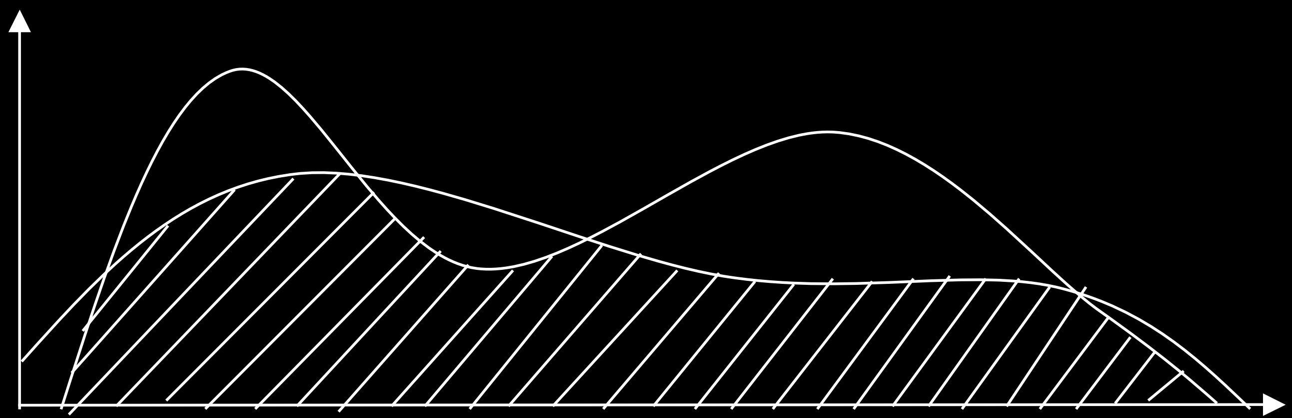
Color Cue

- Histogram intersection

$$\phi_c(\mathbf{s}) = \frac{\sum_{i=1}^{N} \min(I_{\mathbf{s}}(i), M(i))}{\sum_{i=1}^{N} I_{\mathbf{s}}(i)},$$

\# of bins

Model histogram



Color Cue

- Color space
  - B-G
  - G-R
  - R+G+B (why do we need that)
- 8 bins for B-G and G-R, 4 for R+G+B
- Training the model histogram
- Normalization

$$\bar{\phi}_c(\mathbf{s}) = \frac{\phi_c(\mathbf{s}) - \min_{\mathbf{s}_i \in S} \phi_c(\mathbf{s}_i)}{\max_{\mathbf{s}_i \in S} \phi_c(\mathbf{s}_i) - \min_{\mathbf{s}_i \in S} \phi_c(\mathbf{s}_i)}.$$

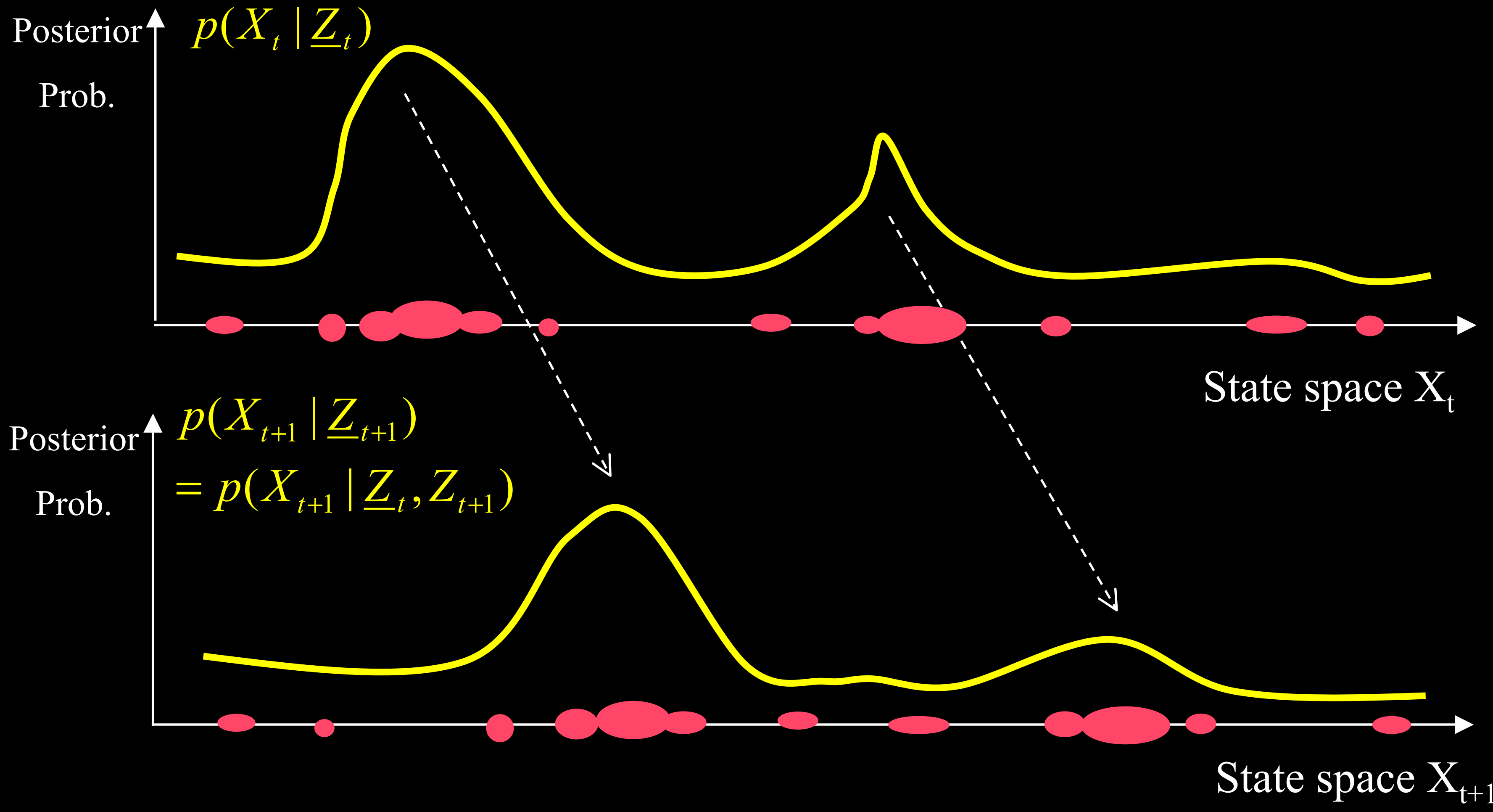
Comments

- Can the rotation be handled?
- Can the scaling issue be handled?
- Is the search strategy good enough?
- Is the color module good?
- Is the motion prediction enough?
- Is the combination of the two cues good?
- Can it handle occlusion?
- Can it cope with multiple faces
  - Coalesce
  - Switch ID

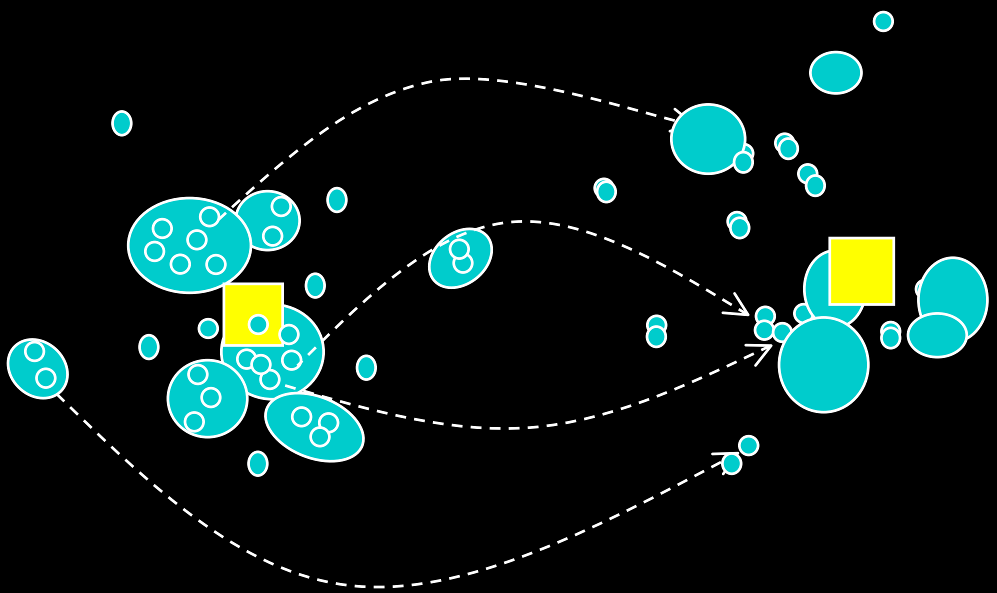
Other Solutions

- Condensation algorithm
- 3D head tracking

Tracking as Density Propagation



Sequential Monte Carlo

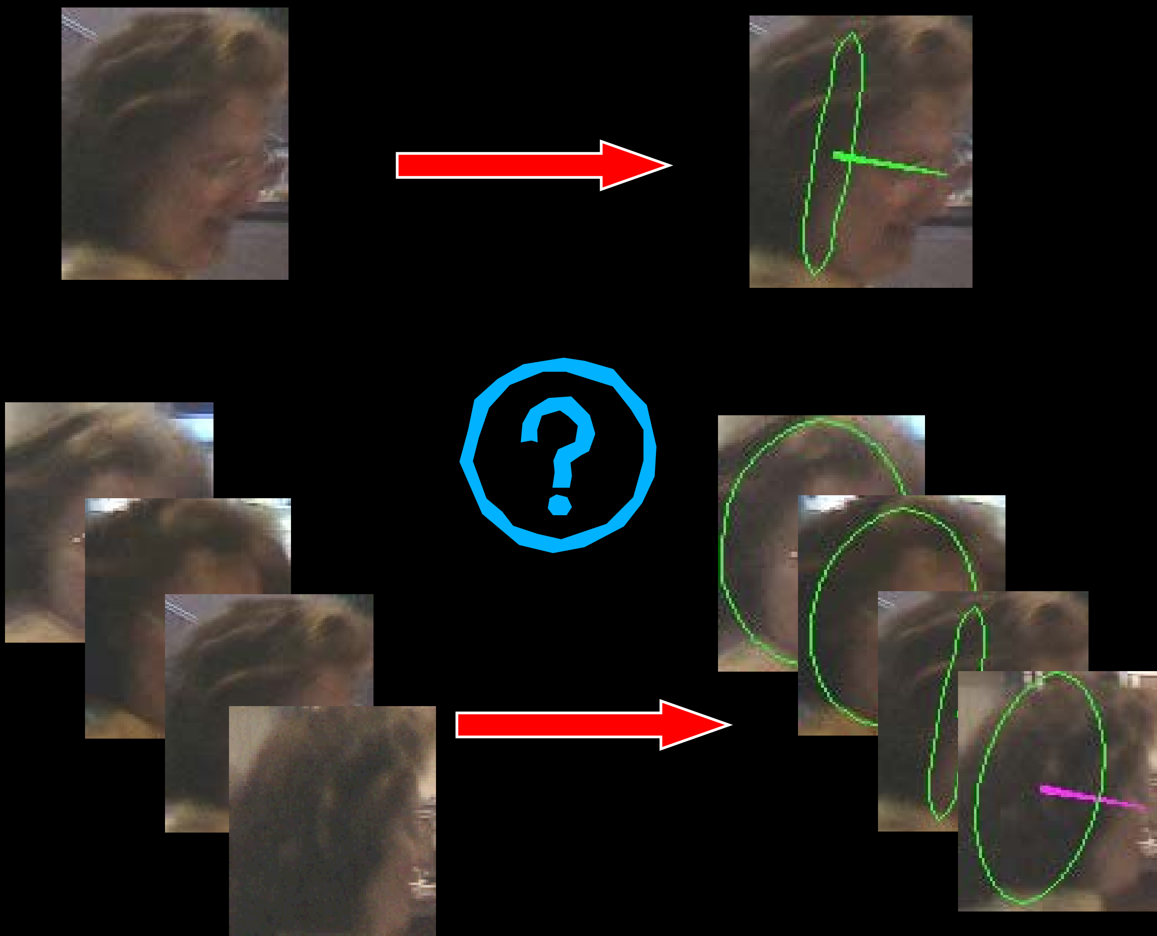


- $P(X_t|Z_t)$ is represented by a set of weighted samples
- Sample weights are determined by $P(Z_t^{(n)}|X_t^{(n)})$
- Hypotheses generating is controlled by $P(X_t|X_{t-1})$

Challenge to Condensation

- Curse of dimensionality
  - What to track?
    - ✓ *Positions, orientations*
    - ✓ *Shape deformation*
    - ✓ *Color appearance changing*
  - The dimensionality of X
  - The number of hypotheses grows exponentially

3D Face Tracking: The Problem



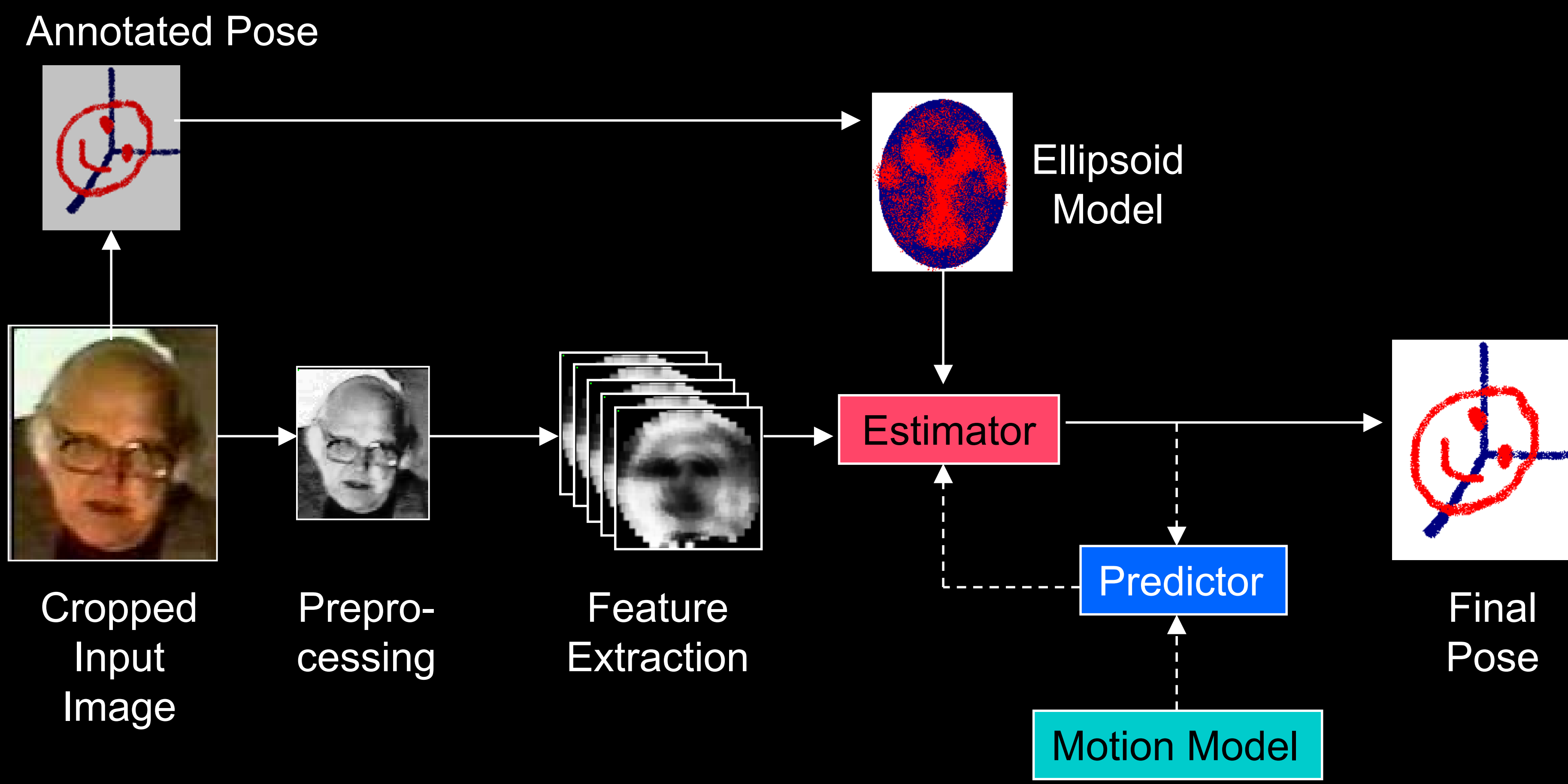
The goal:

Estimate and track 3D head poses

The challenges:

- ✓ Side view
- ✓ Back view
- ✓ Poor illumination
- ✓ Low resolution
- ✓ Different users

3D Face Tracking: A Solution



Courtesy of Y. Wu and K. Toyama, 2000

3D Face Tracking: some results