COM-711

SELECTED TOPICS IN COMPUTER VISION

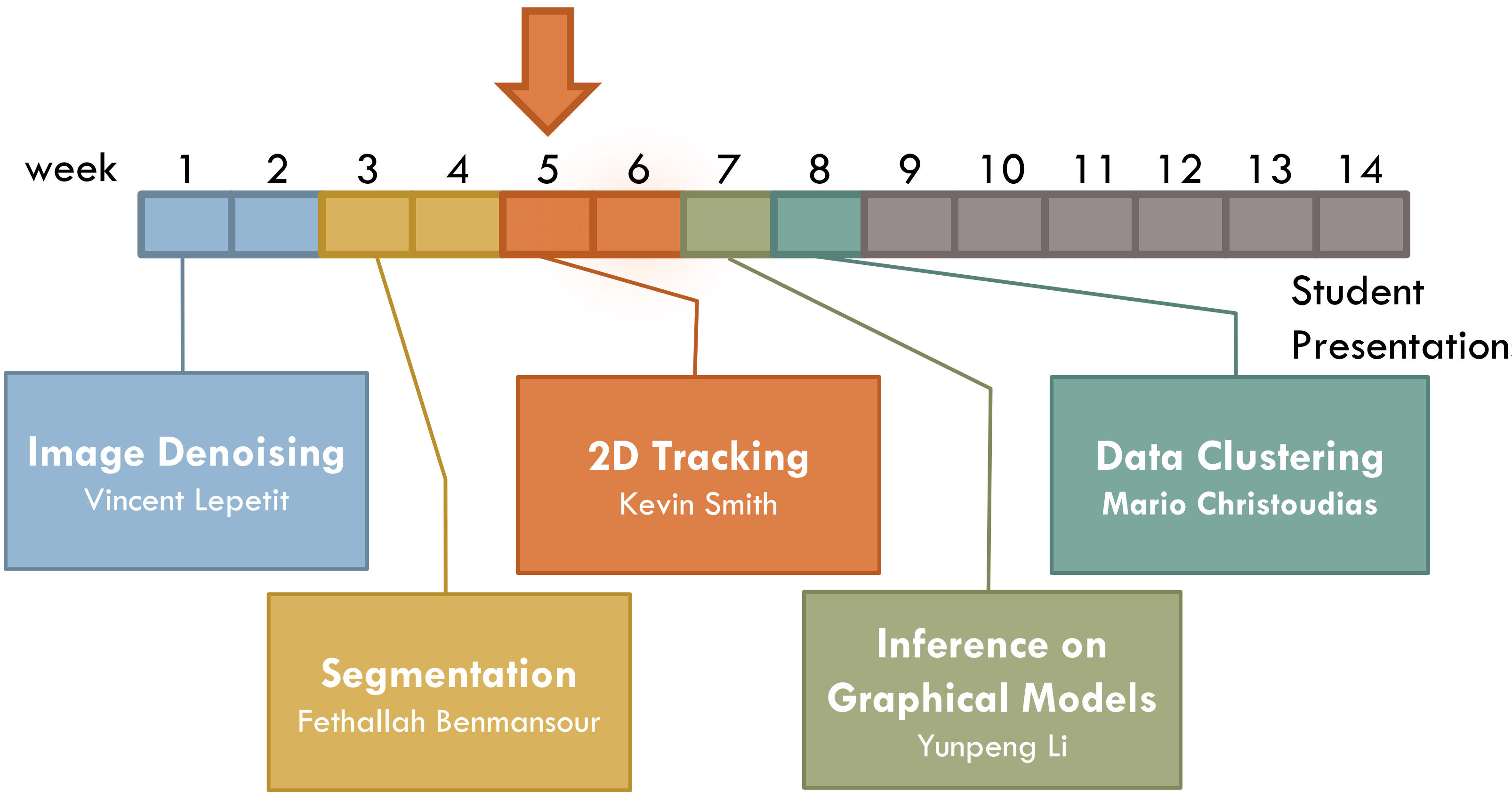
2D TRACKING PART 1/2

Oct 21, 2011

Kevin Smith kevin.smith@lmc.biol.ethz.ch

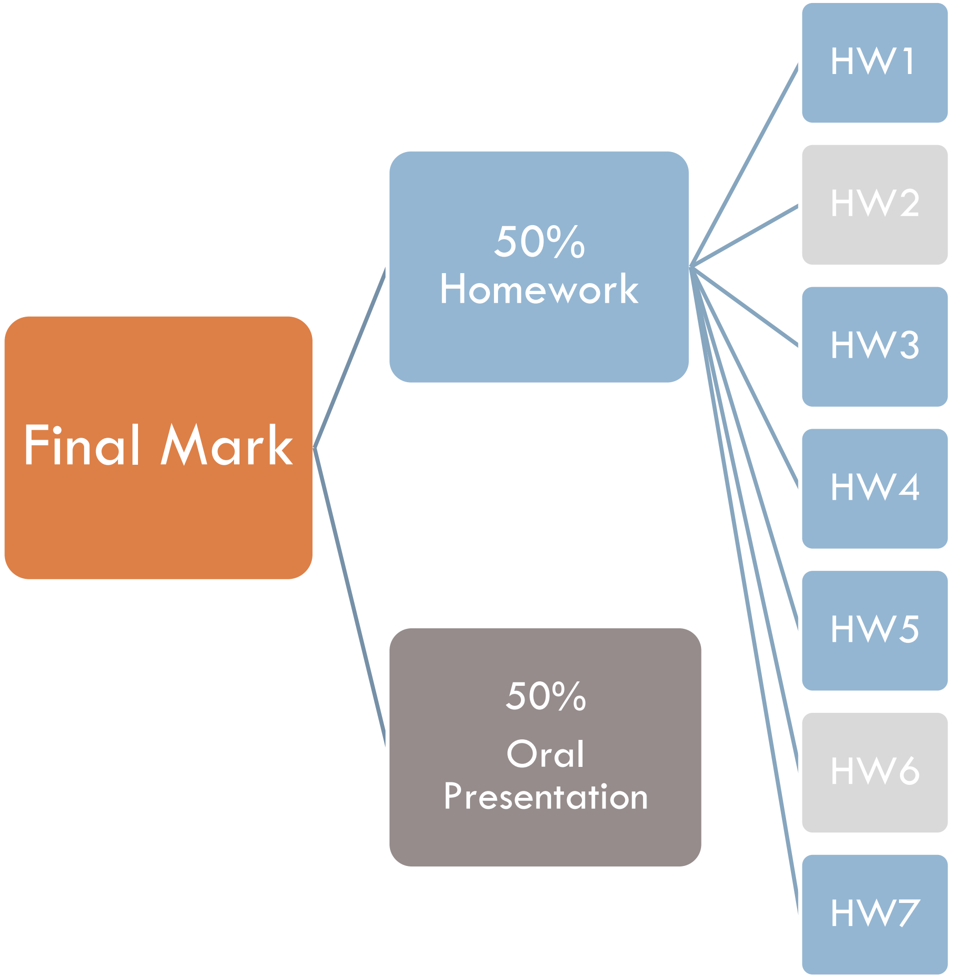
Course update

- Lectures



Course update

- Final mark
  - 50% homework
  - 50% presentation
- Homework mark
  - Considers only the best(*N*-2) scores from *N* total assignments



Oral presentations

- Each student **will present a published paper** on topics covered in the course to the rest of the class
  - Each student has approx **20 minutes** to speak (including questions)
  - A list of papers you may select from will be posted on the web site http://cvlab.epfl.ch/teaching/topics/index.php
  - Alternatively, you may propose a paper to present (subject to approval)
  - Instructors and other students will ask questions about the work
  - Presentations will be held during the last 6 weeks of the course. Time slots will be assigned on a first-come-first serve basis, after the list is posted. A web site will be made available to sign up with your selected paper and time slot

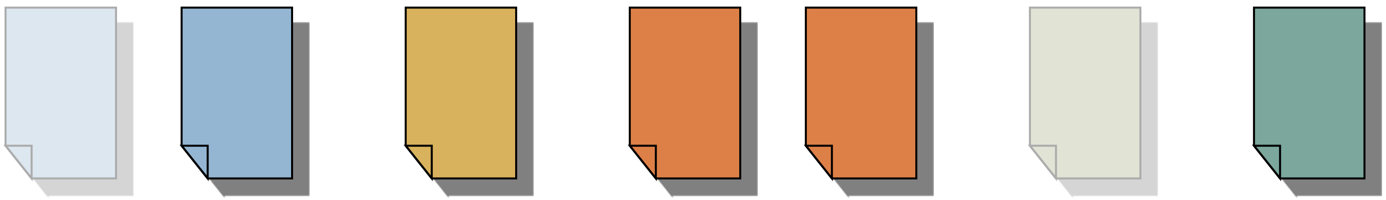
Oral presentations

- Hints for a good presentation

  - Goal: inform your classmates
  - Start strong
    - Skip the overview and introduction slides
    - Avoid outline slides
    - Start by showing the results/benefits of the work
  - Use formulas only when necessary (fewer is better)
    - Establish a notation the audience understands
    - Describe the intuition behind formulae
  - Concentrate on the novelty of the work
    - Summarize well-known methods to save time
  - Limit the number of slides (~ 1 slide / minute)
  - Anticipate questions
  - Practice!
    - In front of an audience, if possible

Course update

- Homework: 7 total assignments



- Only the best N-2 scores from N total assignments will be considered (you can "skip" two assignments)

- First 4 assignments available on course web site
http://cvlab.epfl.ch/teaching/topics/index.php

Homework

- 2D Tracking assignments
  - HW4: Derive the recursive Bayesian filtering equation
  - HW5: Implement a particle filter, use it to track objects in 3 video sequences

Outline

- Introduction to the tracking problem
- Recursive Bayesian filtering
- Batch probabilistic methods

Outline

Introduction to the tracking problem

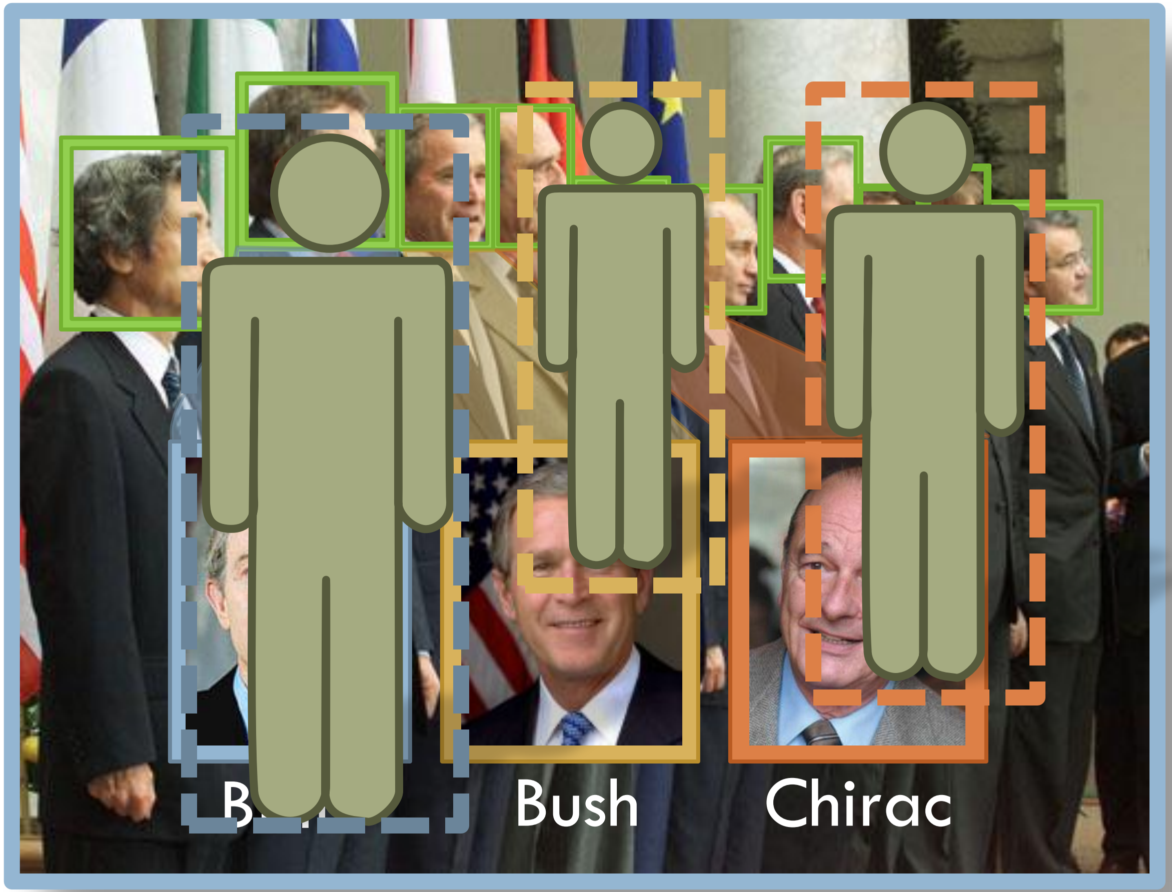
- What is tracking?
- Approaches & assumptions
- Tracking applications
- State of the art & challenges

Recursive Bayesian filtering

Batch probabilistic methods

What is tracking?

- detection
- recognition
- tracking

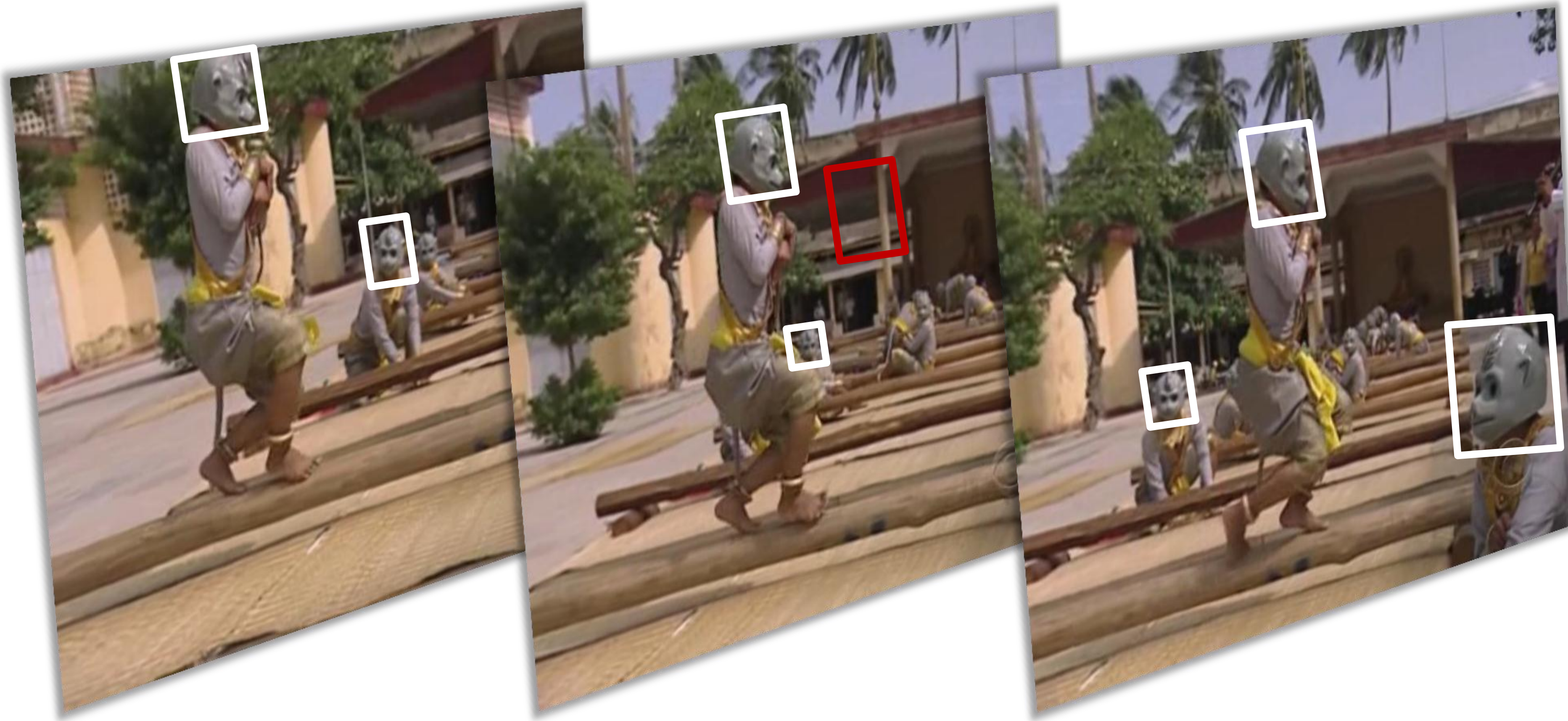


What is tracking?

- Definition: using **image measurements** and a **predictive dynamic model** to consistently estimate the state(s) X_t of one or more object(s) over the discrete time steps corresponding to video frames.

What is tracking?

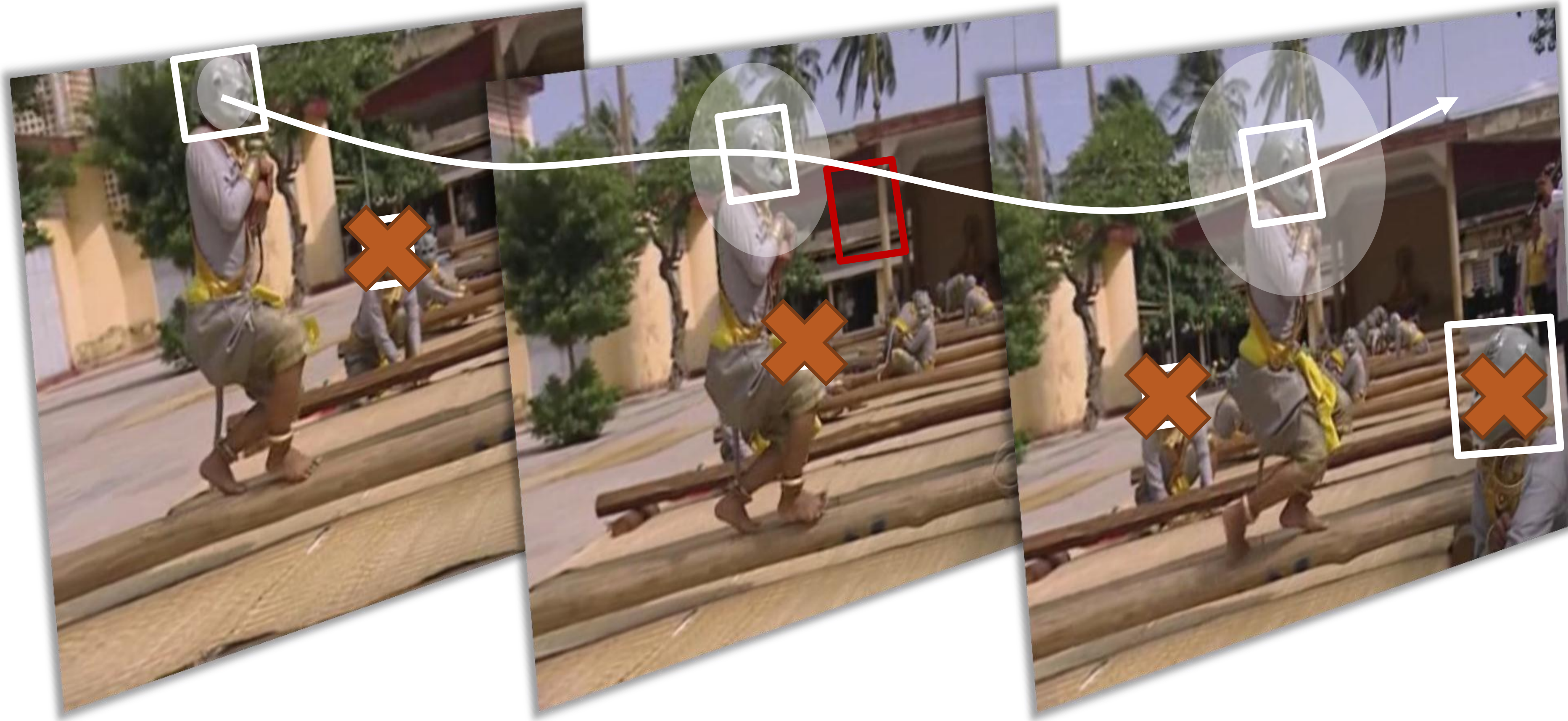
- Why not just do detection?
  - Estimate the state X at each time step
  - - inefficient
  - - data association problem



What is tracking?

- It's better to do tracking
  - Maintain an estimate of X over time, predict the future location

- + efficient, restricts search space
- + smoothes noisy measurements
- - requires knowledge about object behavior



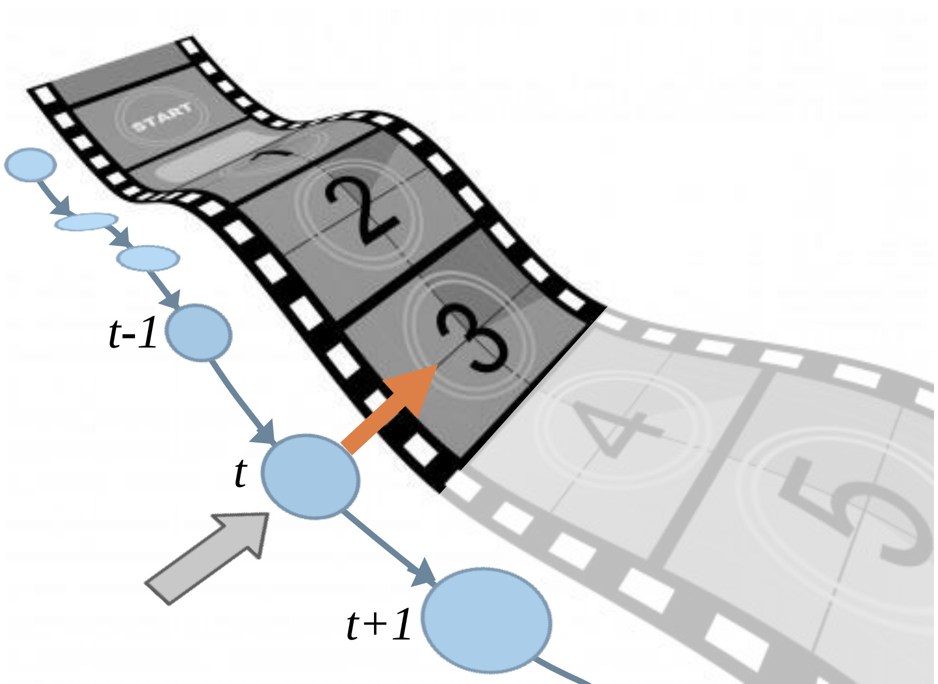
Tracking assumptions

- Smooth camera
  - No instant transitions between viewpoints
  - Any camera pose/parameter changes are gradual
- Object motion can be modeled
  - Linear models
  - Non-linear mod
- Likelihood of object presence at a location in the image can be modeled
  - Typically uses local image information

Approaches to tracking

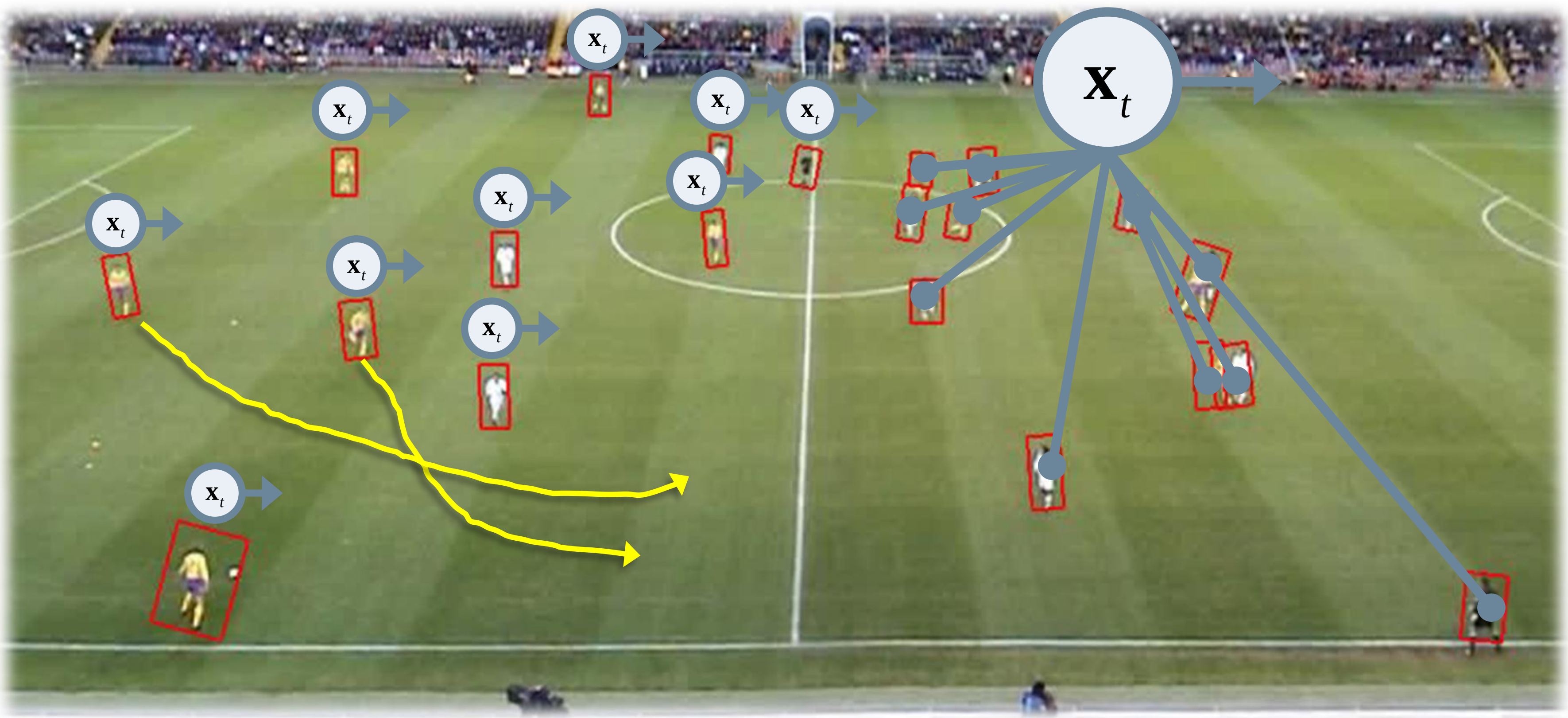
- **Sequential**
  - (recursive, online)
  - + Inexpensive → real-time
  - - no future information
  - - cannot revisit past errors

- **Batch Processing**
  - (offline)
  - - Expensive → not real-time*
  - + considers all information
  - + can correct past errors



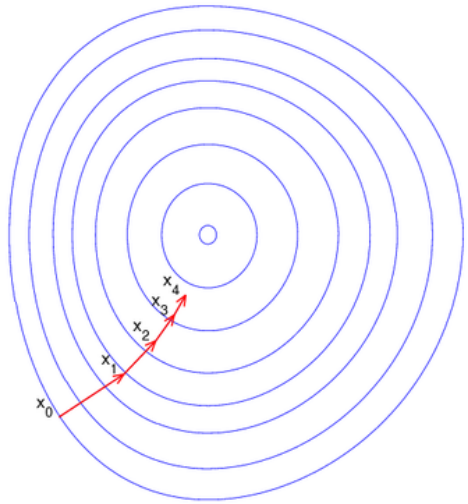
Approaches to tracking

- Parallel trackers
  - several single-object trackers
  - computationally less expensive
  - how to handle interaction, cross-overs?
- Joint state
  - single multi-object representation
  - computationally expensive
  - principled interaction models

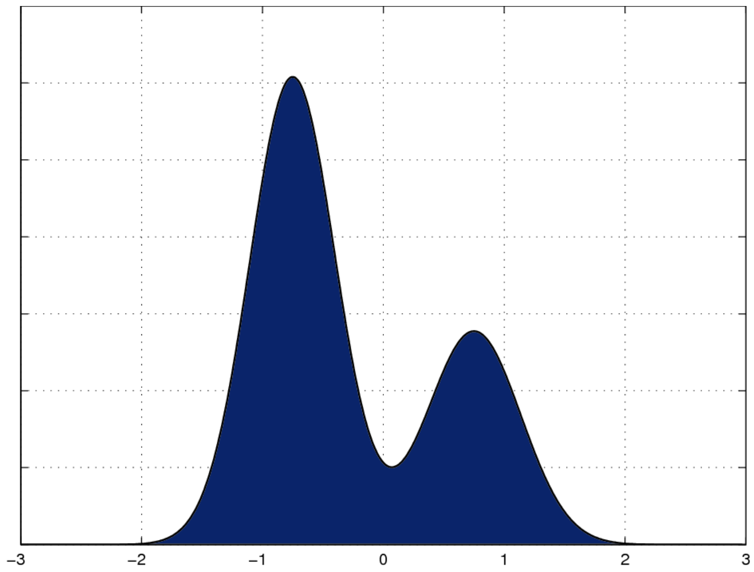


Approaches to tracking

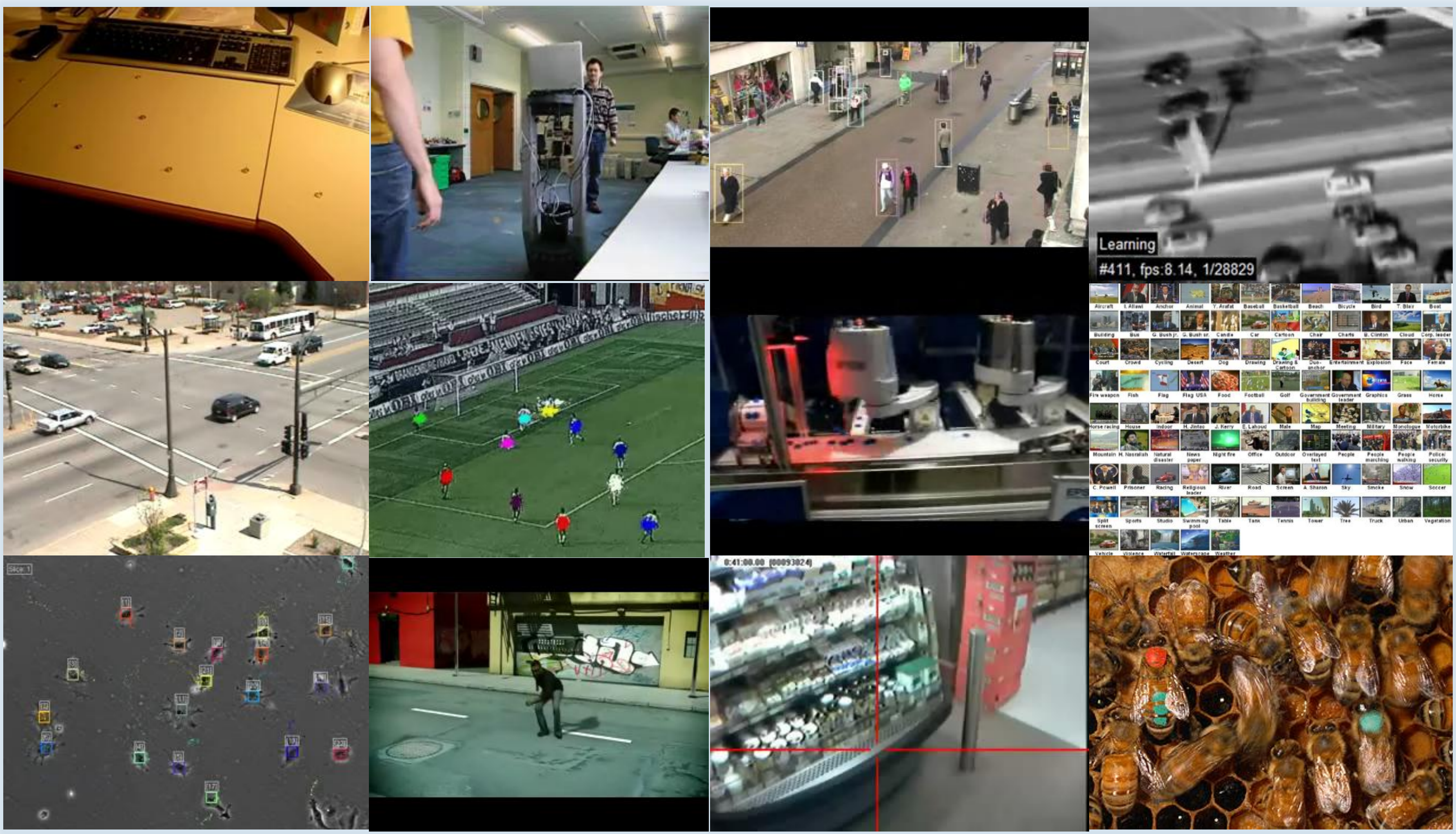
- Non-probabilistic
  - + quick convergence*
  - + efficient
  - - stuck in local minima
  - - does not model multiple objects



- Probabilistic
  - + flexible, principled
  - + multi-modal
  - - slower
  - - interpretation



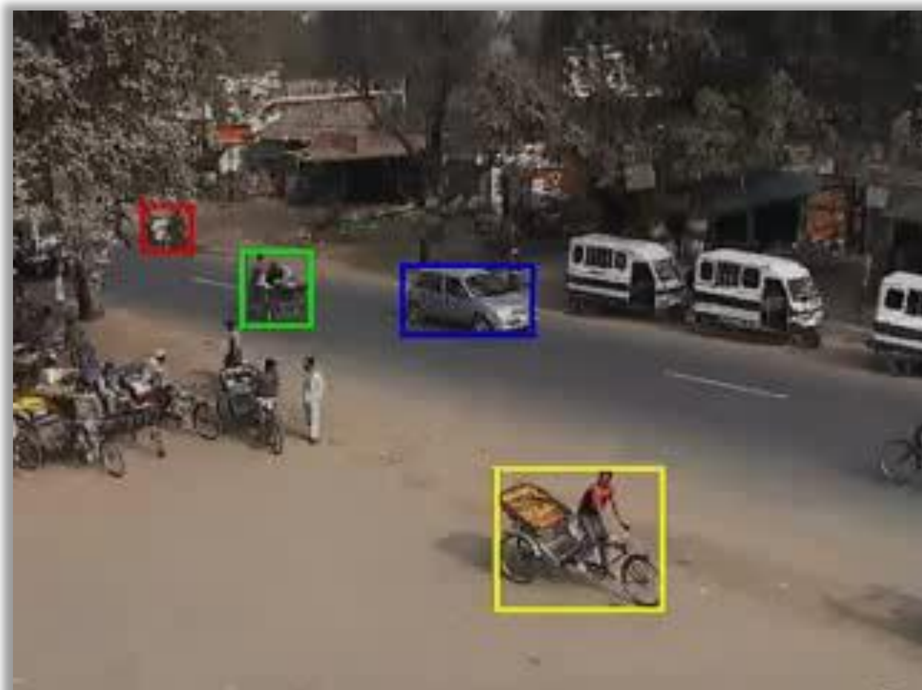
Tracking applications



Tracking applications

- Tracking is an essential step in many computer vision based applications

Detection + Feature Extraction → Tracking → Activity Recognition + Event Recognition → Behavior Analysis + Social Models



Prithwijit Guha , Amitabha Mukerjee and K. S. Venkatesh, Spatio-temporal Discovery: Appearance + Behavior = Agent, Computer Vision, Graphics and Image Processing 4338 516-527, 2007

Tracking applications

- Sports

P. Nillius, J. Sullivan, S. Carlsson , Multi-Target Tracking - Linking Identities using Bayesian Network Inference
Computer Vision and Pattern Recognition (CVPR), 2006

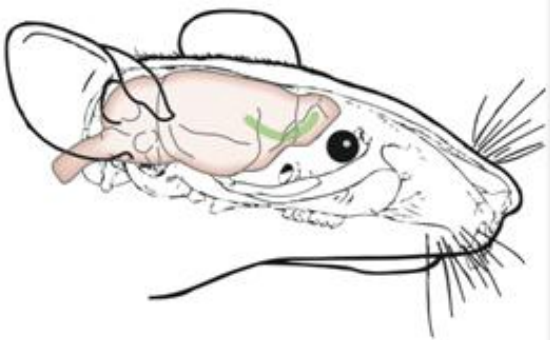
Tracking applications

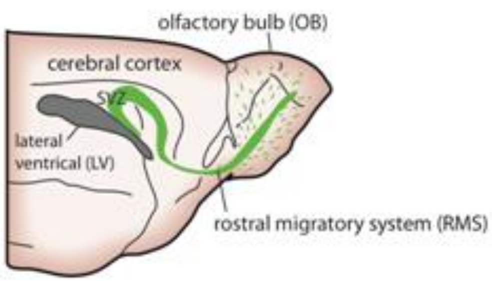
- Surveillance

K. Smith, P. Quelhas, and D. Gatica-Perez, Detecting Abandoned Luggage Items in a Public Space, Performance Evaluation of Tracking and Surveillance (PETS) Workshop at CVPR, New York, NY, June 18 2006

Tracking applications

- Biological research

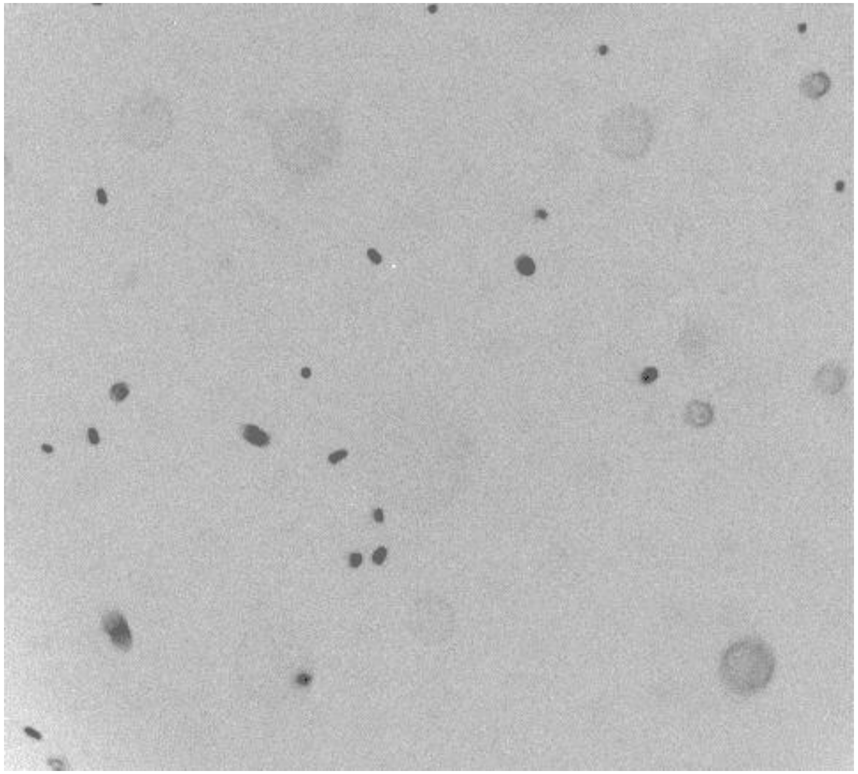




K. Smith, A. Carleton, and V. Lepetit, [General Constraints for Batch Multiple-Target Tracking Applied to Large-Scale Videomicroscopy](), Computer Vision and Pattern Recognition (CVPR), Anchorage, AK, June 2008

Tracking applications

- Biological Research
  - Goal: develop a method to "trap" Salmonella bacteria



P. Horvath, Q. Buhkari, 3D Tracking of point-like objects in 2D image sequences, LMC, ETHZ

What is the state of the art?

- Despite being classic computer vision problem, tracking is **largely unsolved**
  - Some limited successes
  - No general-purpose tracker
  - No standard data corpus for comparison
  - No standard evaluation methodology
  - Challenging problems remain

Obstacles to tracking

- appearance change
- occlusion
- distraction
- illumination change
- difficult motion
- multiple objects
- scale change
- efficient solution

Ruei-Sung Lin, David Ross, Jongwoo Lim, Ming-Hsuan Yang, Adaptive discriminative generative model and its Applications Neural Information Processing Systems Conference (NIPS), 2004

Obstacles to tracking

- appearance change
- occlusion
- distraction
- illumination change
- difficult motion
- multiple objects
- scale change
- efficient solution

Amit Adam, Ehud Rivlin and Ilan Shimshoni, Robust Fragments-based Tracking using the Integral Histogram (pdf).
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2006

Obstacles to tracking

- appearance change
- occlusion
- distraction
- illumination change
- difficult motion
- multiple objects
- scale change
- efficient solution

Michael Isard and Andrew Blake CONDENSATION -- conditional density propagation for visual tracking
International Journal of Computer Vision (IJCV), 29, 1, 5--28, (1998)

Obstacles to tracking

- appearance change
- occlusion
- distraction
- illumination change
- difficult motion
- multiple objects
- scale change
- efficient solution

Ruei-Sung Lin, David Ross, Jongwoo Lim, Ming-Hsuan Yang, Adaptive discriminative generative model and its Applications Neural Information Processing Systems Conference (NIPS), 2004

Obstacles to tracking

- appearance change
- occlusion
- distraction
- illumination change
- difficult motion
- multiple objects
- scale change
- efficient solution

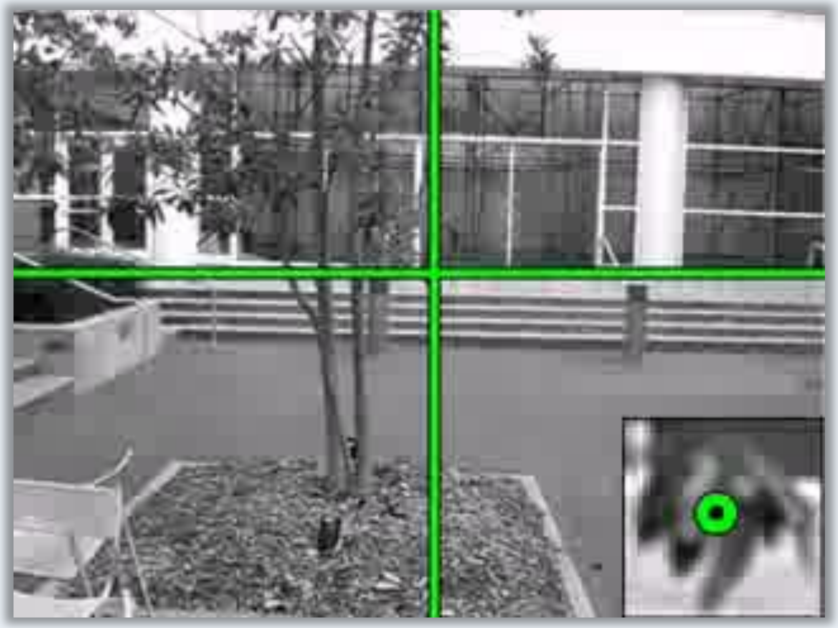
Michael Isard and Andrew Blake CONDENSATION -- conditional density propagation for visual tracking
International Journal of Computer Vision (IJCV), 29, 1, 5--28, (1998)

Obstacles to tracking

- appearance change
- occlusion
- distraction
- illumination change
- difficult motion
- multiple objects
- scale change
- efficient solution

Saad Ali and Mubarak Shah, Floor Fields for Tracking in High Density Crowd Scenes, European Conference on Computer Vision (ECCV), 2008.

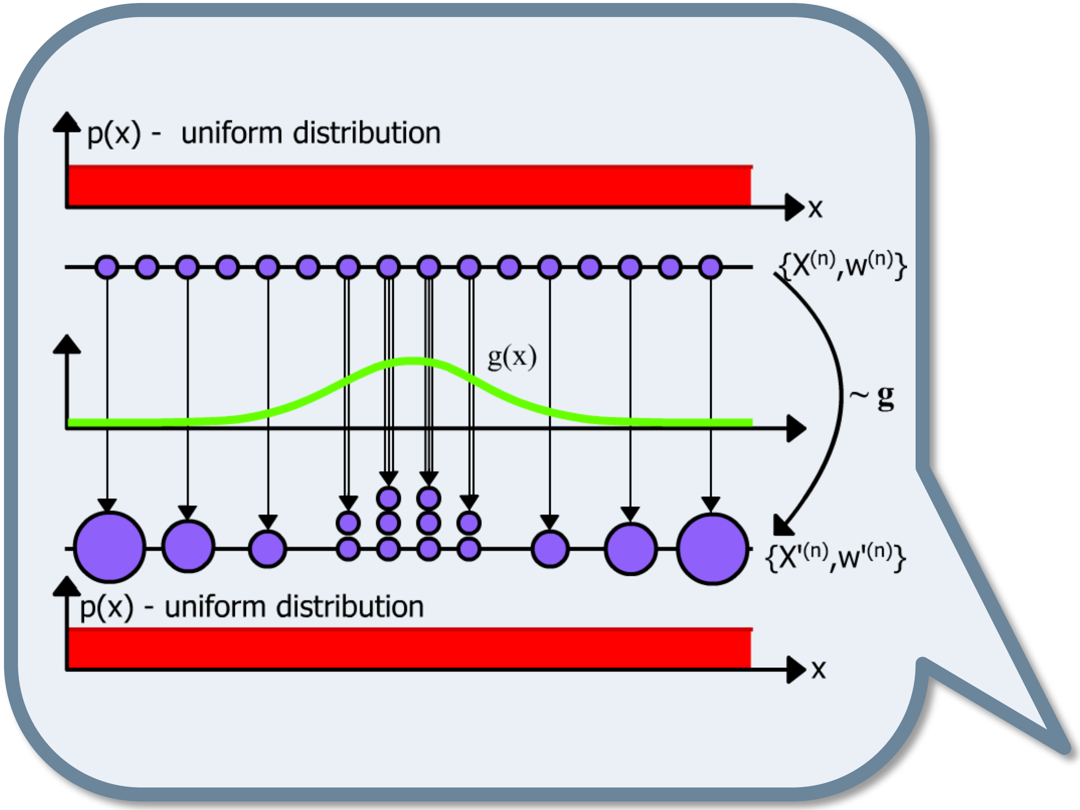
Obstacles to tracking



- appearance change
- occlusion
- distraction
- illumination change
- difficult motion
- multiple objects
- scale change
- efficient solution

Shawn Lankton, James Malcolm, Arie Nakhmani, and Allen Tannenbaum. Tracking Through Changes in Scale Proceedings of International Conference on Image Processing (ICIP), 2008.

Obstacles to tracking



- appearance change
- occlusion
- distraction
- illumination change
- difficult motion
- multiple objects
- scale change
- efficient solution

Outline

Introduction to the tracking problem

Recursive Bayesian filtering

- Background & formulation
- Kalman filter
- Particle filter

Batch probabilistic methods

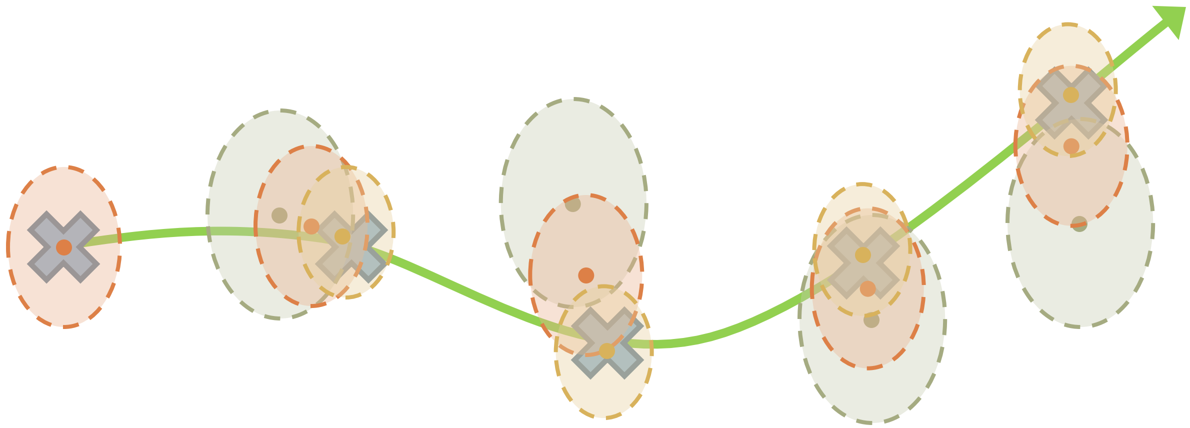
Recursive Bayesian filtering

- How is it characterized?
  - Sequential
  - Parallel trackers OR joint modeling of multiple objects
  - Probabilistic

- Popular examples
  - Kalman filter
  - Particle filter

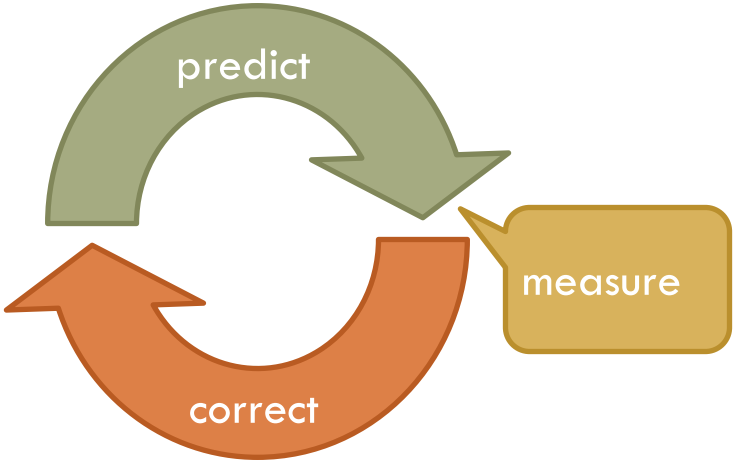
Recursive Bayesian filtering

- **Key idea 1:** Probability distributions represent our belief as to the state of the object

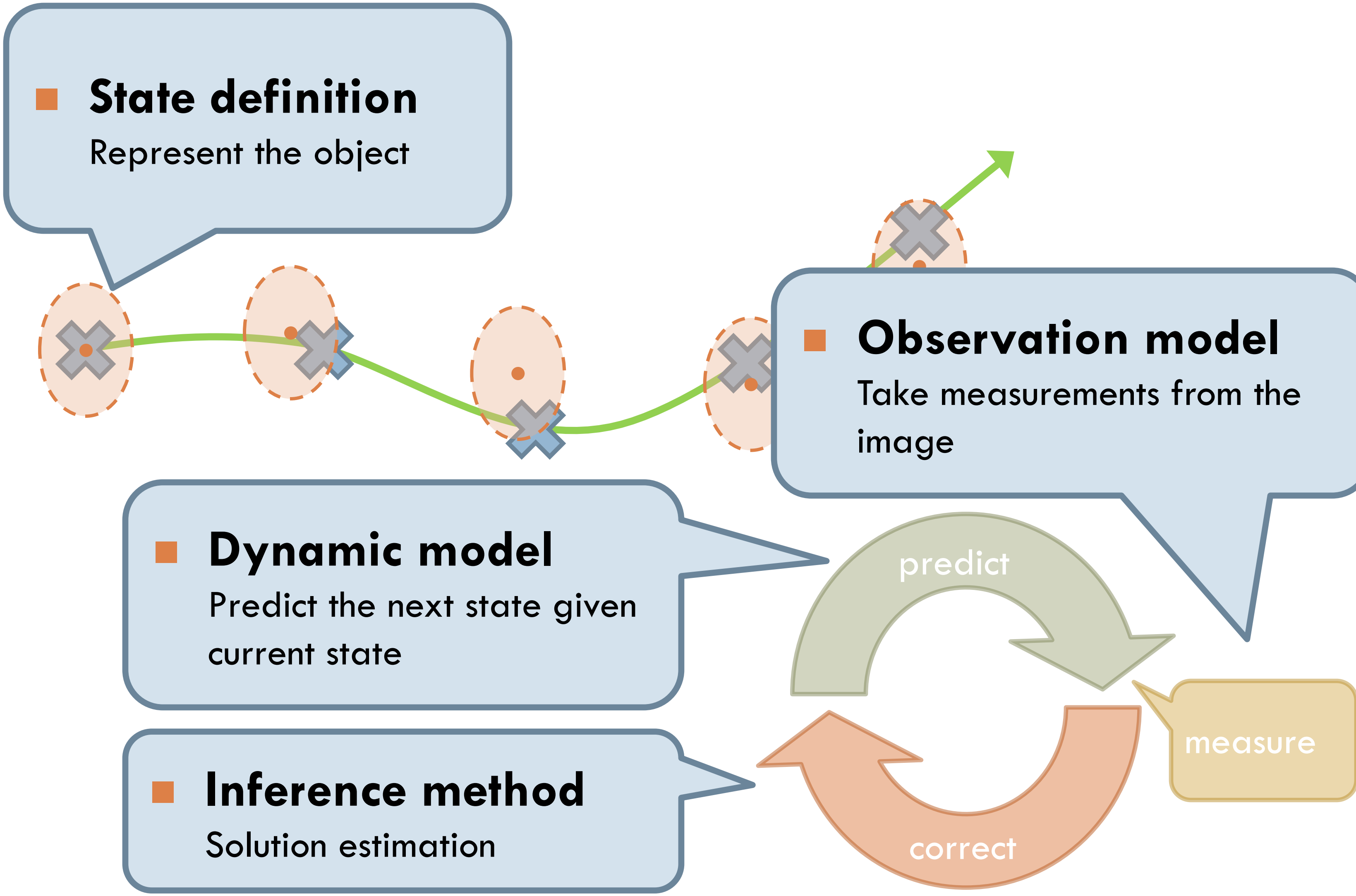


- **Key idea 2:** Recursive cycle
  1. **Predict** from motion model
  2. **Measurement** from image
  3. **Correct** the prediction

  …repeat



Recursive Bayesian filtering



Tracking ingredients

- State Definition

- Observation Model

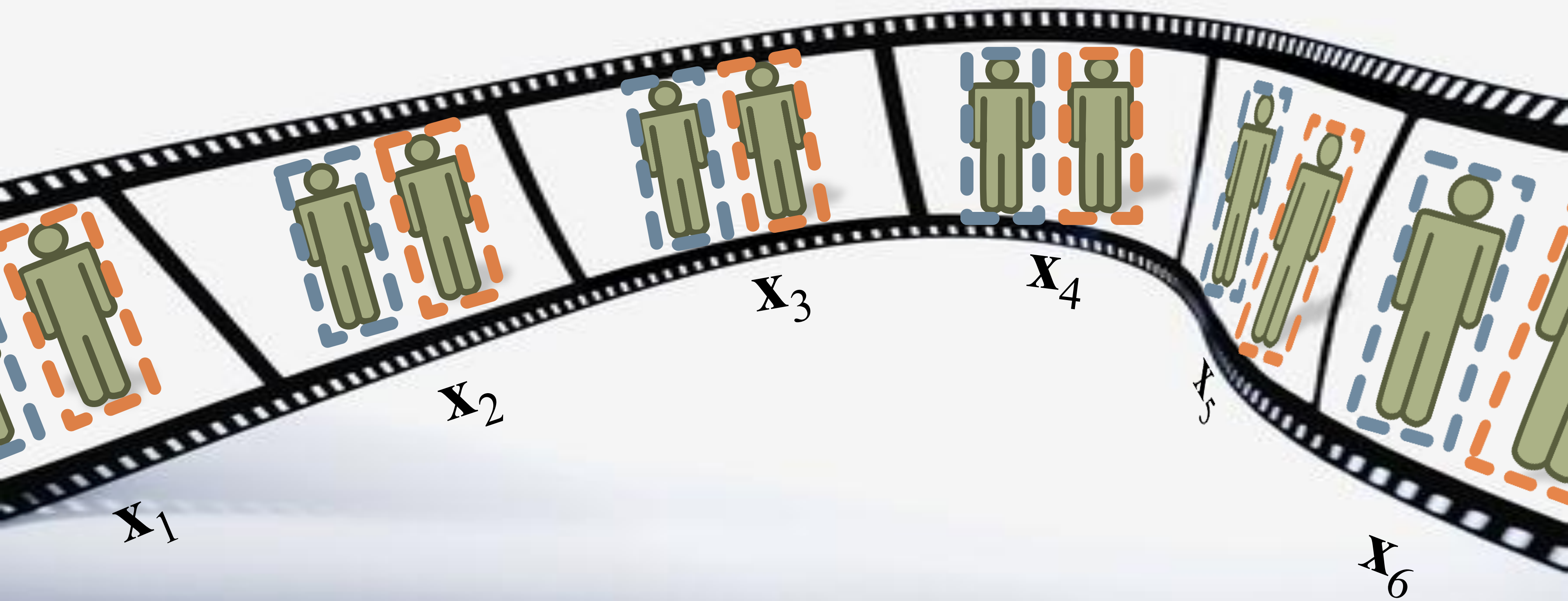
- Dynamic Model

- Inference

State definition

- Describes properties of the tracked object(s) at an instant in time
- Defines solution space

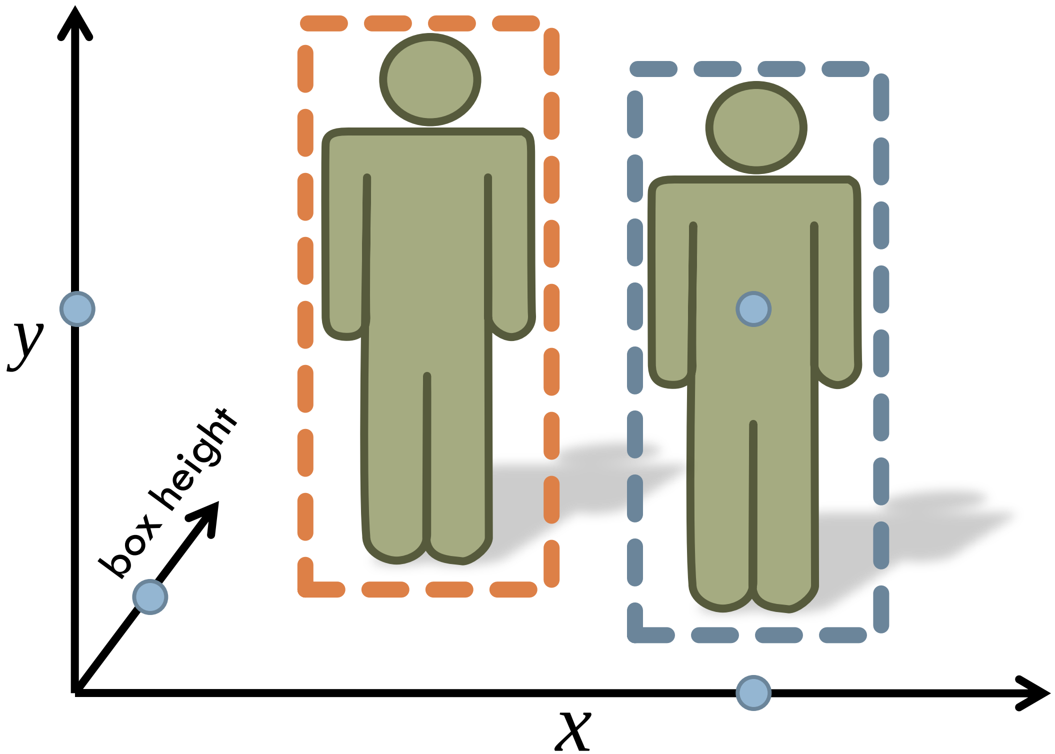
$$X_t = \{\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}, \mathbf{x}_t\}$$



State definition

- Decomposed for time step t, $\mathbf{X}_t$ can parameterize the object in many ways, often via:
  - location
  - velocity
  - size
  - shape
  - identity
  - switching model

$$\mathbf{x}_t = x$$
$$\mathbf{x}_t = (x, y)$$
$$\mathbf{x}_t = (x, y, h)$$
$$\mathbf{x}_t = \{\mathbf{x}_t^1, \mathbf{x}_t^2\}$$



State definition

- Object defined by a point
  - position
  - velocity
  - acceleration

$$\mathbf{x}_t = (x, y)$$

$$\mathbf{x}_t = (x, y, \dot{x}, \dot{y})$$

$$\mathbf{x}_t = (x, y, \dot{x}, \dot{y}, \ddot{x}, \ddot{y})$$

Saad Ali and Mubarak Shah, Floor Fields for Tracking in High Density Crowd Scenes, European Conference on Computer Vision (ECCV), 2008.

State definition

- Bounding box
  - position
  - height
  - aspect
  - velocity

$$\mathbf{x}_t = (x, y)$$

$$\mathbf{x}_t = (x, y, h, a)$$

$$\mathbf{x}_t = (x, y, \dot{x}, \dot{y}, h, a)$$

M. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, L. Van Gool, Robust Tracking-by-Detection using a Detector Confidence Particle Filter , International Conference on Computer Vision (ICCV), 2009

State definition

- Ellipse
  - location
  - eccentricity
  - major axis

$$\mathbf{x}_t = (x, y, m, e)$$

$$\mathbf{x}_t = (x, y, a, b)$$

Yuan LI, Chang Huang, Ram Nevatia *[Learning to Associate: Hybrid Boosted Multi-Target Tracker for Crowded Scene](), Computer Vision and Pattern Recognition (CVPR), June 2009*

State definition

- Active contour
  - b-splines
  - control points
  - spline length
  - H basis functions

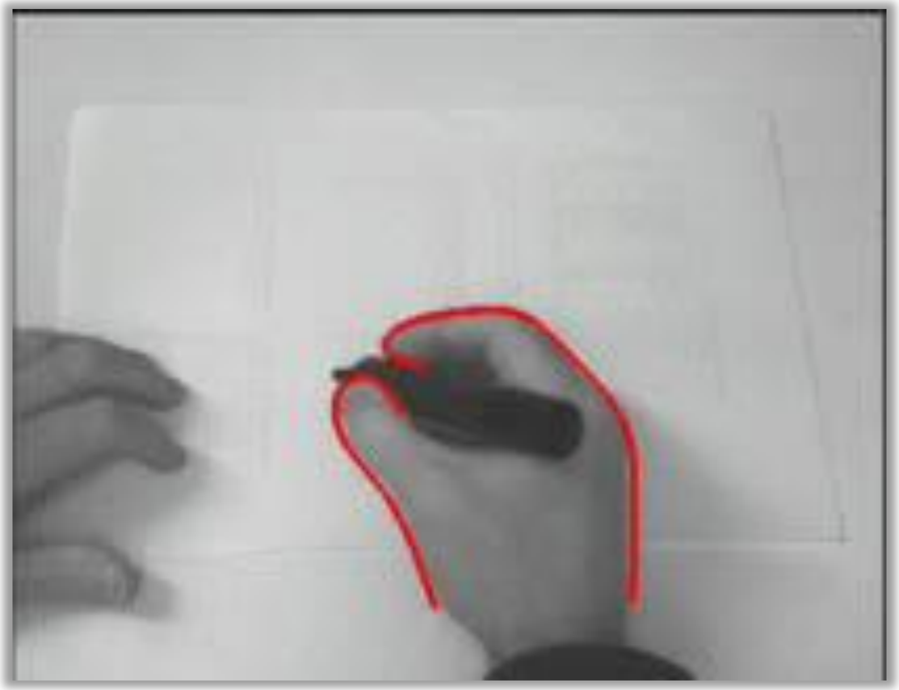
$$\mathbf{x}_t = \left(X(s), Y(s)\right)$$

$$X(s) = H(s)\mathrm{X},\ 0 \le s \le N$$

$$Y(s) = H(s)\mathrm{Y}$$

$$\mathrm{X} = \{x^1, x^s, \ldots, x^N\}$$

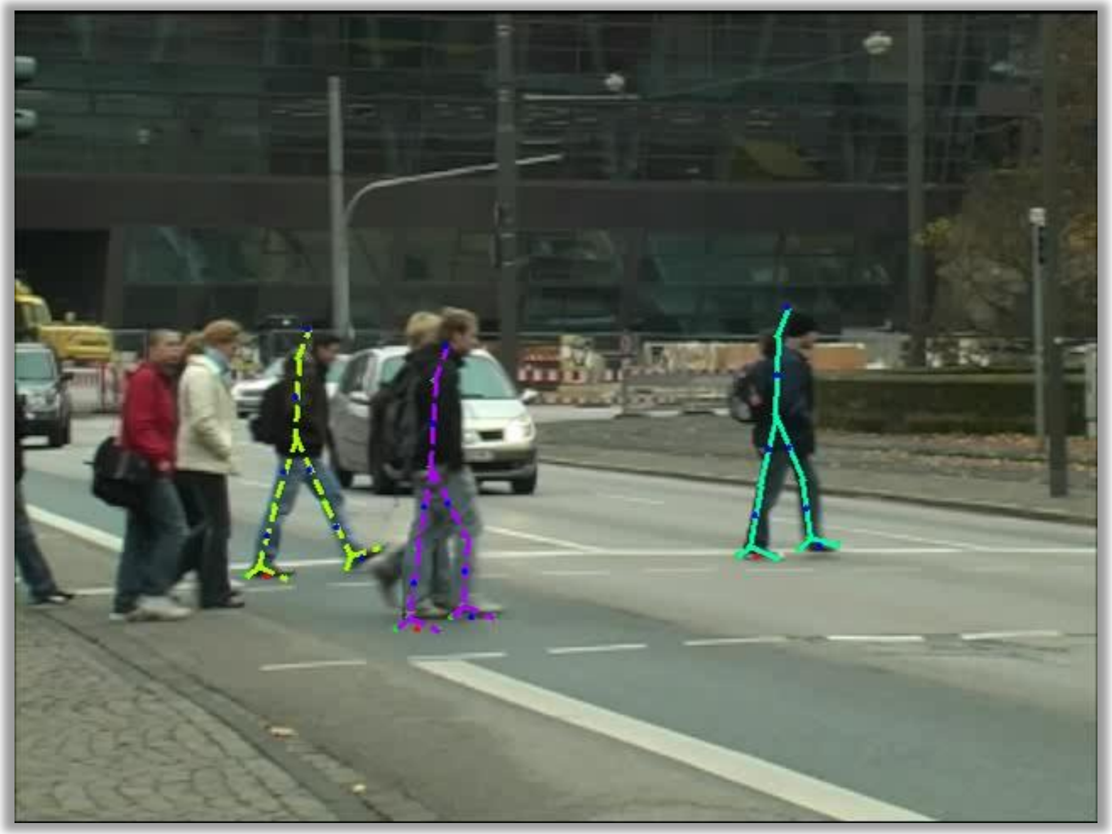
$$\mathrm{Y} = \{y^1, y^s, \ldots, y^N\}$$



Michael Isard and Andrew Blake [CONDENSATION -- conditional density propagation for visual tracking]
International Journal of Computer Vision (IJCV), 29, 1, 5--28, (1998)

State definition

- Articulated & Part-based Models
  - set of vertices
  - locations
  - scales
  - constraints



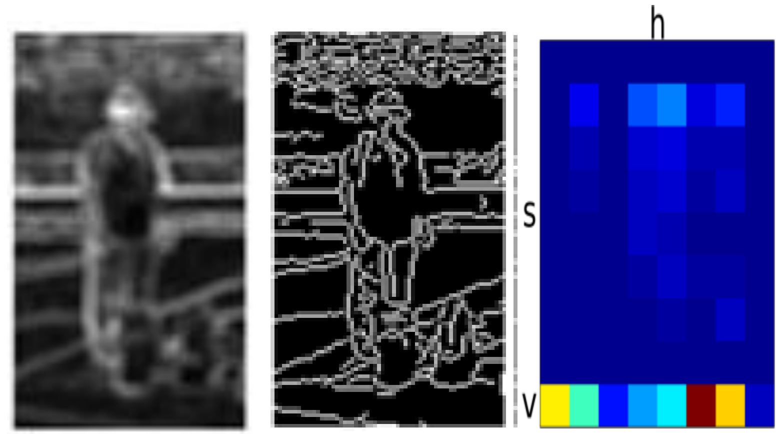
$$\mathbf{x}_t = \{v^1, v^i, \ldots, v^N\}$$
$$v^i = (x^i, y^i, s^i)$$

M. Andriluka, S. Roth, B. Schiele, People Tracking by Detection and People-Detection-by-Tracking, *Computer Vision and Pattern Recognition (CVPR'08), Anchorage, USA, June 2008*

Tracking ingredients

- State Definition

- Observation Model



- Dynamic Model

- Inference

Observation model

- Notation - observations

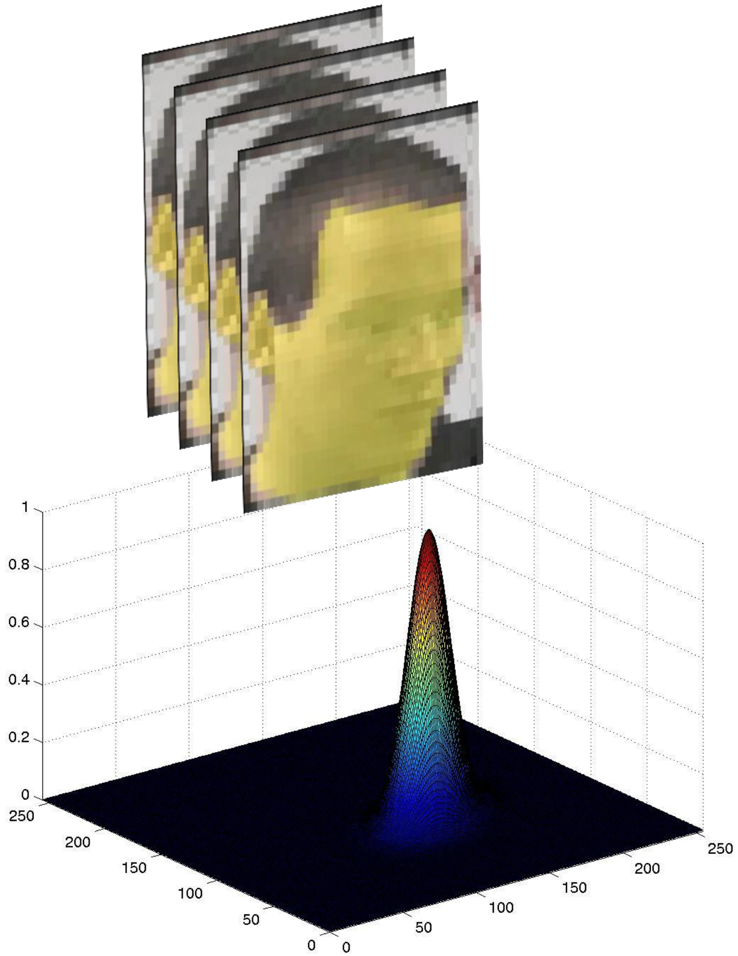
$$Z_t = \{\mathbf{z}_1, \ldots, \mathbf{z}_{t-1}, \mathbf{z}_t\}$$

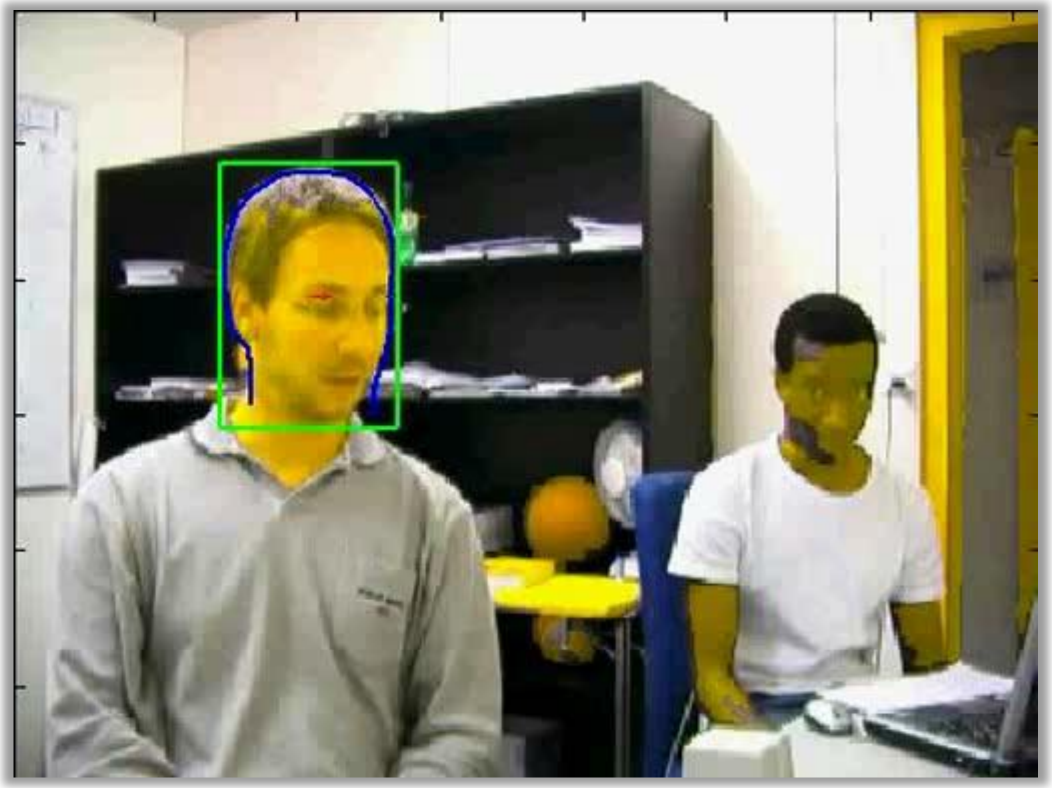
- Returns the likelihood that a state hypothesis gave rise to the observed image data

$$p(\mathbf{z}_t \mid \mathbf{x}_t)$$

Observation model

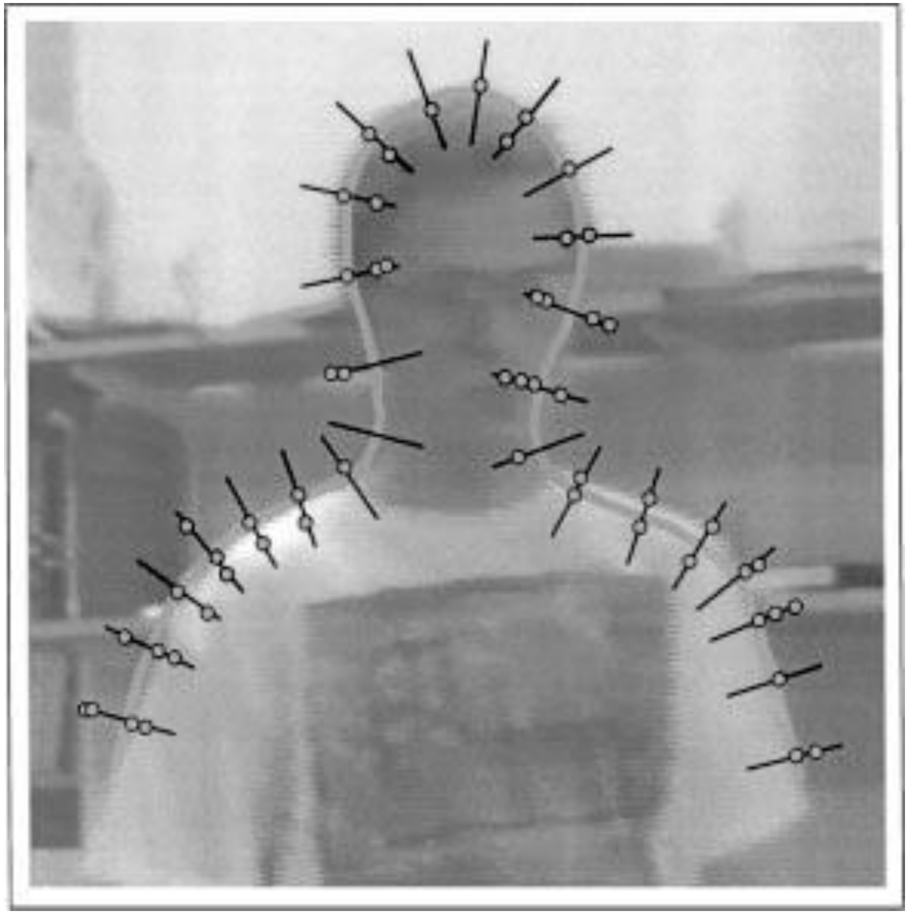
- Modeling skin color





Observation model

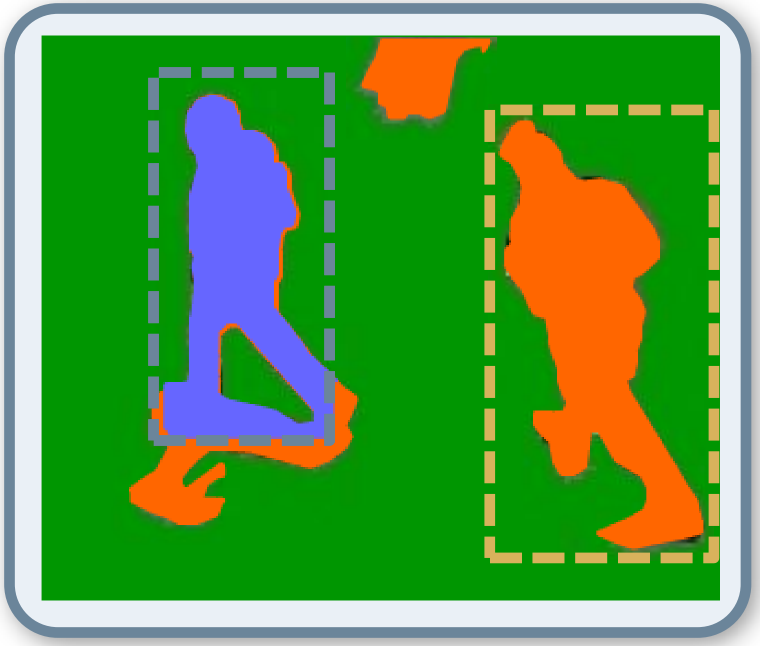
- Sum of measurements taken from lines perpendicular to a contour

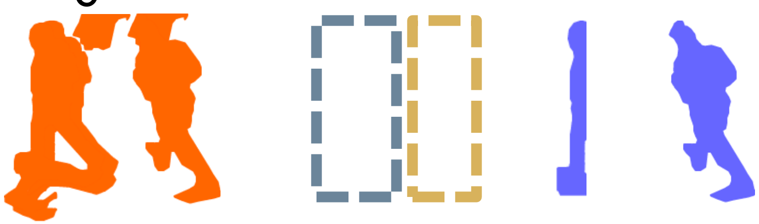


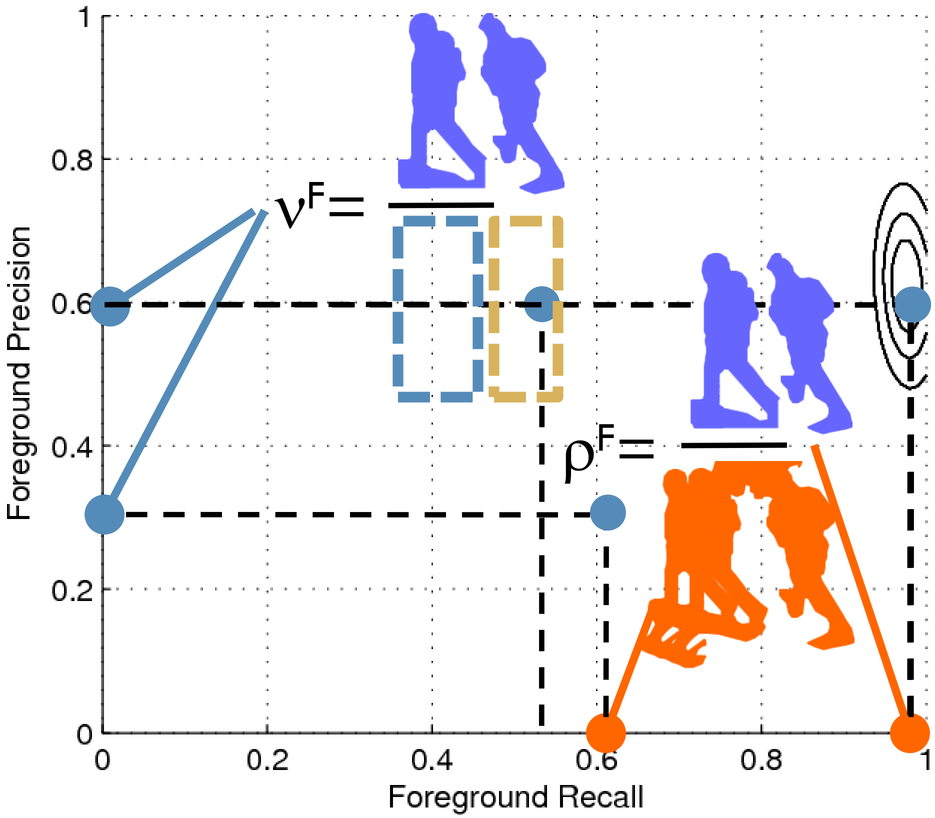
Michael Isard and Andrew Blake [CONDENSATION -- conditional density propagation for visual tracking]()
International Journal of Computer Vision (IJCV), 29, 1, 5--28, (1998)

Observation model

- Background/foreground silhouette modeling



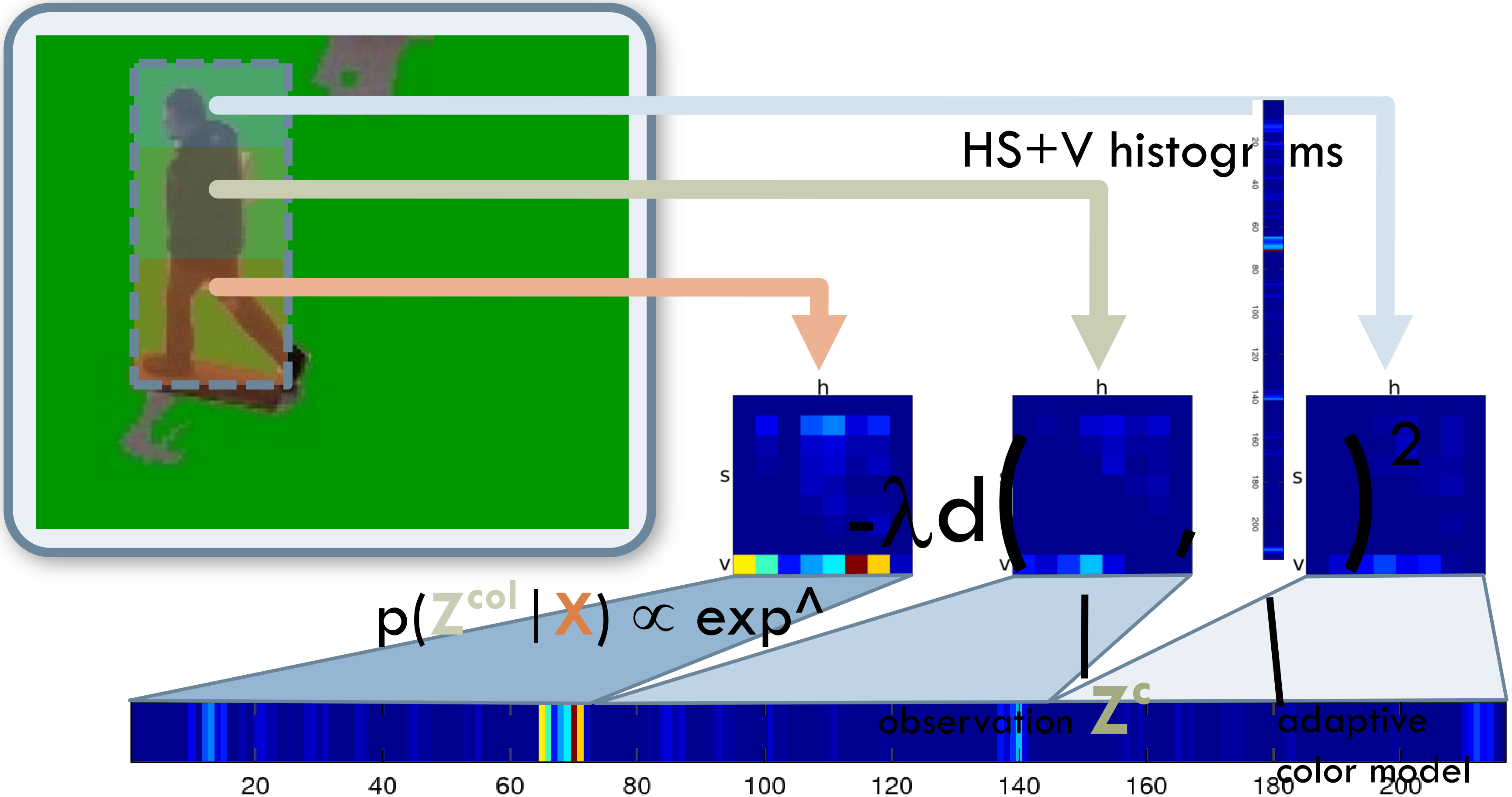




K. Smith, D. Gatica-Perez, and J.M. Odobez, Using Particles to Track Varying Numbers of Objects, CVPR, June 2005

Observation model

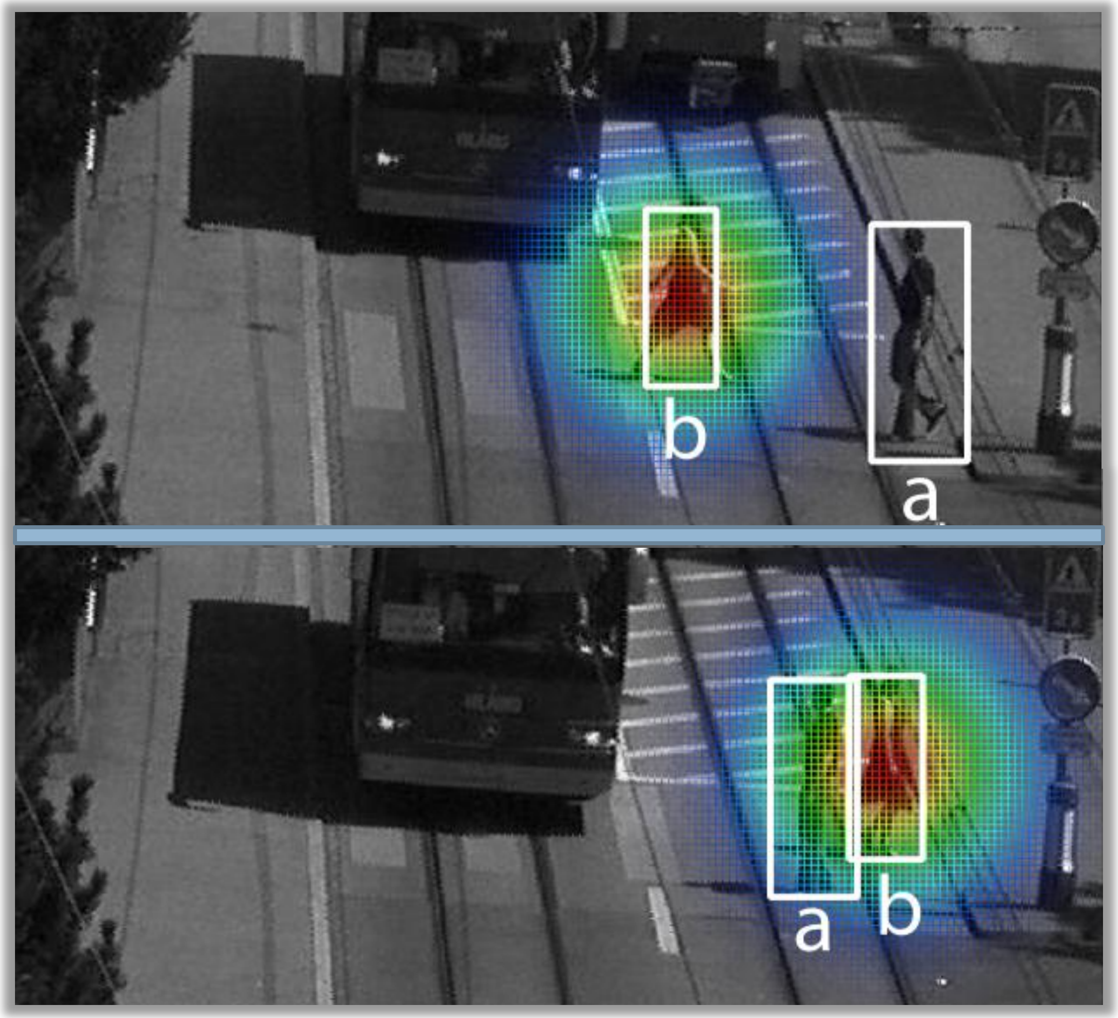
- Parts-based color model



P. Perez, C. Hue, J. Vermaak, and M. Ganget. Color-Based Probabilistic Tracking, in *ECCV,* May 2002

Observation model

- Detector confidence
  - HOG based sliding window detector



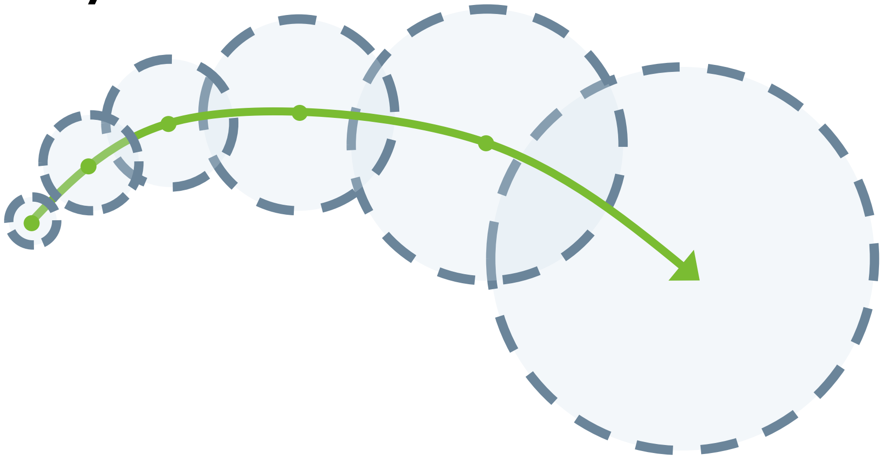
M. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, L. Van Gool, Robust Tracking-by-Detection using a Detector Confidence Particle Filter , International Conference on Computer Vision (ICCV), 2009

Tracking ingredients

- State Definition

- Observation Model

- Dynamic Model



- Inference

Dynamic model

- Current state is predicted from previous state

$$p(\mathbf{x}_t \mid \mathbf{x}_{t-1}) = N\left(\mathbf{F}_t \mathbf{x}_{t-1}, \Sigma_{F_t}\right)$$

$$\mathbf{x}_\mathrm{t} \sim N\left(\mathbf{F}_t \mathbf{x}_{t-1}, \Sigma_{F_t}\right)$$ to obtain samples

- Autoregressive linear dynamic model

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{w}_t$$

Predicted state time t

State transition model

Previous state time t-1

Noise term $\mathbf{w}_\mathrm{t} \sim N(0, \mathbf{Q}_t)$

Dynamic model

- 1 st order autoregressive
  - models position & velocity

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{w}_t$$

State vector

$$\mathbf{x}_t = \begin{pmatrix} x \\ y \\ \dot{x} \\ \dot{y} \end{pmatrix}$$

$$\begin{pmatrix} x_t \\ y_t \\ \dot{x}_t \\ \dot{y}_t \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \\ & & 1 & \\ & & & 1 \end{pmatrix}}_{\mathbf{F}_t} \begin{pmatrix} x_{t-1} \\ y_{t-1} \\ \dot{x}_{t-1} \\ \dot{y}_{t-1} \end{pmatrix} + \mathbf{w}_t$$

State transition model

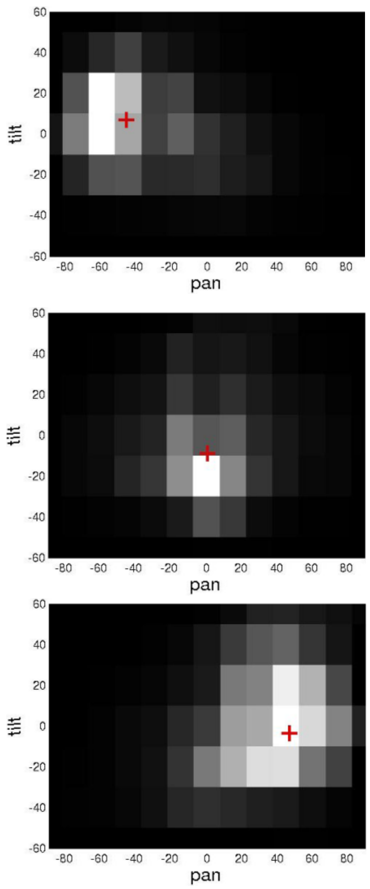
Previous state

Dynamic model

- Nonlinear dynamic models
  - Discrete state transitions

Discrete pose states

Transition probabilities



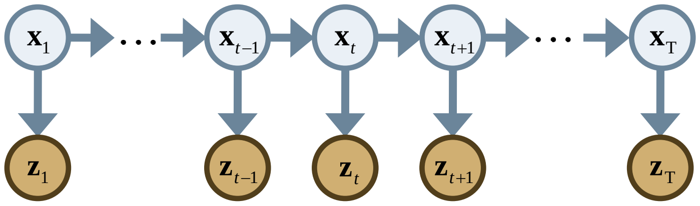
Tracking ingredients

- State Definition

- Observation Model

- Dynamic Model

- Inference



Recursive Bayesian filtering

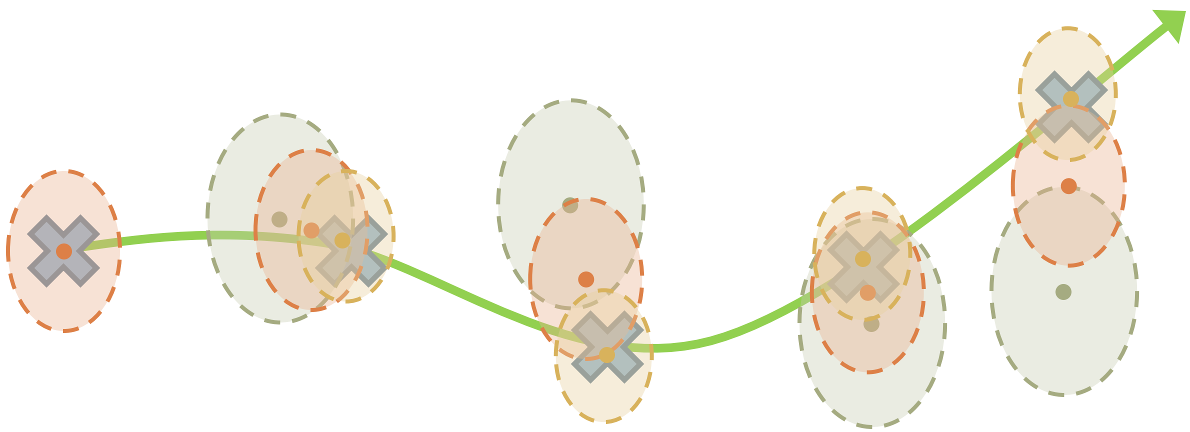
- Filtering equation:

$$\underbrace{p(\mathbf{x}_t \mid Z_t)}_{\text{posterior estimate}} \propto \underbrace{p(\mathbf{z}_t \mid \mathbf{x}_t)}_{\text{likelihood or observation model}} \int_{\mathbf{x}_{t-1}} \underbrace{p(\mathbf{x}_t \mid \mathbf{x}_{t-1})}_{\text{motion model or dynamic model}} \underbrace{p(\mathbf{x}_{t-1} \mid Z_{t-1})}_{\text{posterior estimated at t-1}}$$

- Definitions
  - State from 1 to time t: $X_t = \{\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}, \mathbf{x}_t\}$
  - Observations from 1 to time t: $Z_t = \{\mathbf{z}_1, \ldots, \mathbf{z}_{t-1}, \mathbf{z}_t\}$

Recursive Bayesian filtering

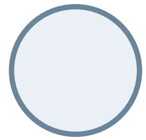
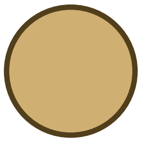
- Use probability distributions to model the tracking problem

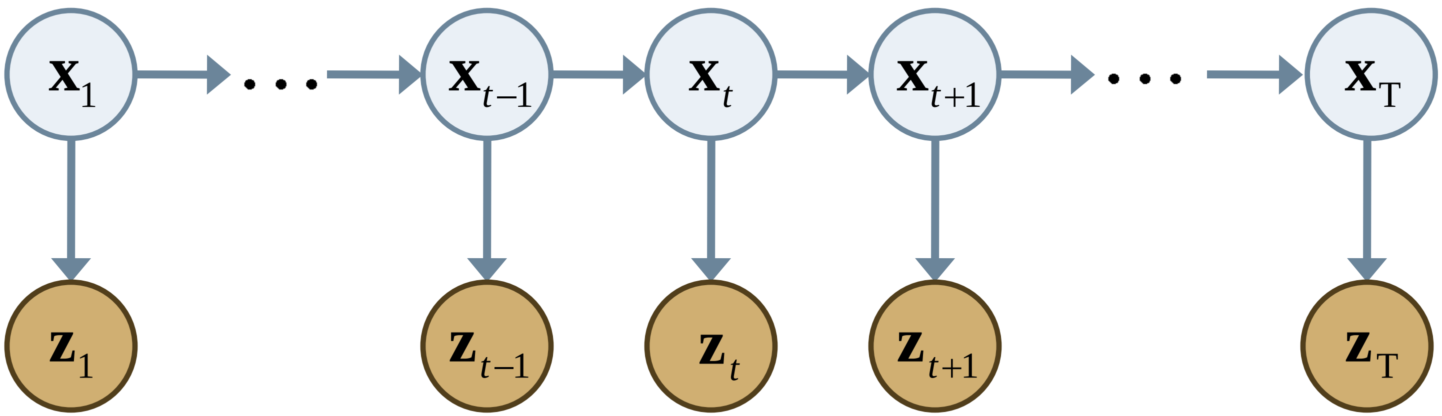


correct measure predict

$$\underbrace{p\left(\mathbf{x}_t \mid Z_t\right)}_{\text{posterior}} \propto \underbrace{p\left(\mathbf{z}_t \mid \mathbf{x}_t\right)}_{\text{likelihood}} \int_{\mathbf{x}_{t-1}} \underbrace{p\left(\mathbf{x}_t \mid \mathbf{x}_{t-1}\right)}_{\text{motion model}} \underbrace{p\left(\mathbf{x}_{t-1} \mid Z_{t-1}\right)}_{\text{posterior at } t\text{-}1}$$

Modeling the tracking problem

- Model the problem as a Hidden Markov Model (HMM)
  - Dependency 
  - Variables: hidden  , or observed 



- Assumptions
  - Dynamics form a Markov chain $p(\mathbf{x}_t \mid X_{t-1}) = p(\mathbf{x}_t \mid \mathbf{x}_{t-1})$
  - Independent observations $p(\mathbf{z}_t \mid \mathbf{x}_t, X_{t-1}, Z_{t-1}) = p(\mathbf{z}_t \mid \mathbf{x}_t)$

Recursive Bayesian filtering

Derivation setup

- Notation
  - State from 0 to time t: $X_t = \{\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}, \mathbf{x}_t\}$
  - Observations from 0 to time t: $Z_t = \{\mathbf{z}_1, \ldots, \mathbf{z}_{t-1}, \mathbf{z}_t\}$
- Assumptions
  - Dynamics form a Markov chain $p(\mathbf{x}_t \mid X_{t-1}) = p(\mathbf{x}_t \mid \mathbf{x}_{t-1})$
  - Independent observations $p(\mathbf{z}_t \mid \mathbf{x}_t, X_{t-1}, Z_{t-1}) = p(\mathbf{z}_t \mid \mathbf{x}_t)$

Useful probability relations

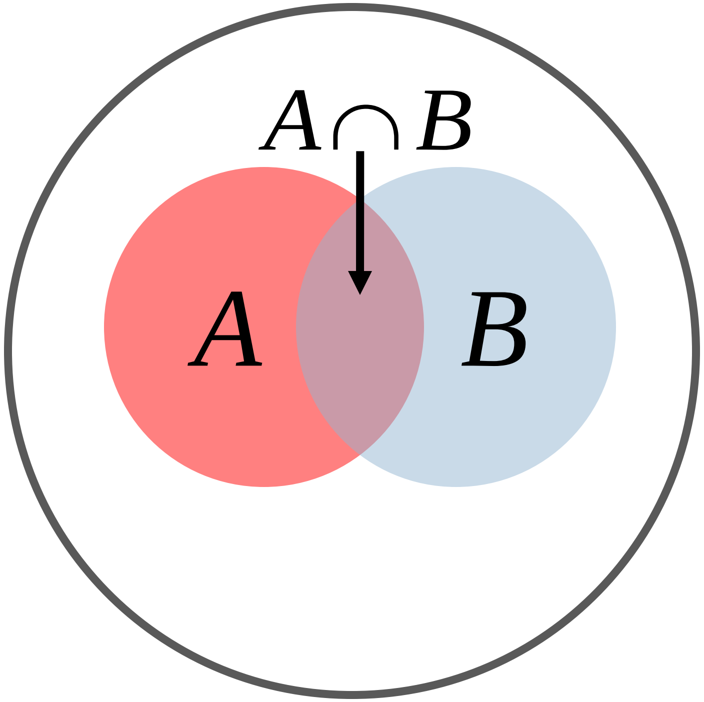
- Conditional probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A \cap B) = P(A, B) = P(A \text{ and } B)$$

$$P(A \cup B) = P(A \text{ or } B)$$



- Bayes theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- Marginal probability (discrete)

$$P(A) = \sum_B P(A, B) = \sum_B P(A|B)P(B)$$

Recursive Bayesian filtering

Derivation setup

starting from this relation

$$p\left(X_t, Z_t\right) = p\left(\mathbf{x}_t, \mathbf{z}_t, X_{t-1}, Z_{t-1}\right)$$

derive the recursive Bayesian filtering equation

$$p\left(\mathbf{x}_t \mid Z_t\right) \propto p\left(\mathbf{z}_t \mid \mathbf{x}_t\right) \int_{\mathbf{x}_{t-1}} p\left(\mathbf{x}_t \mid \mathbf{x}_{t-1}\right) p\left(\mathbf{x}_{t-1} \mid Z_{t-1}\right)$$

Recursive Bayesian filtering

- Derivation removed (homework assignment)

Recursive Bayesian filtering

- Derivation removed (homework assignment)

Homework

- HW4: Derive the recursive Bayesian filtering equation

COM-711 Homework 4

Kevin Smith

October 21, 2011

This assignment is due **December 2, 2011**. Turn in your work electronically to Kevin Smith at `kevin.smith@lmc.biol.ethz.ch`. You may submit your work as a PDF file or as a scanned handwritten document.

1 Derive the recursive Bayesian filtering distribution

In lecture 5, we discussed how to arrive at the formula underlying probabilistic tracking algorithms including the particle filter and the Kalman filter. Your task is to start from the joint relation

$$p(X_t, Z_t) = p(\mathbf{x}_t, \mathbf{z}_t, X_{t-1}, Z_{t-1}), \tag{1}$$

and derive the recursive Bayesian filtering distribution, which is given by

$$p(\mathbf{x}_t|Z_t) \propto p(\mathbf{z}_t|\mathbf{x}_t) \int_{\mathbf{x}_{t-1}} p(\mathbf{x}_t|\mathbf{x}_{t-1}) p(\mathbf{x}_{t-1}|Z_{t-1}), \tag{2}$$

where $\mathbf{x}_t$ is the state of the object(s) at time step t, $\mathbf{z}_t$ is the observation at time t, $X_t = \{\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}, \mathbf{x}_t\}$ is the set of all object(s) states, and $Z_t = \{\mathbf{z}_1, \ldots, \mathbf{z}_{t-1}, \mathbf{z}_t\}$ is the set of all current and previous observations. Keep in mind our assumptions: that the dynamic process is Markovian (the current state depends only on the previous state)

$$p(\mathbf{x}_t|X_{t-1}) = p(\mathbf{x}_t|\mathbf{x}_{t-1}), \tag{3}$$

and that the current observation is independent from the other observations,