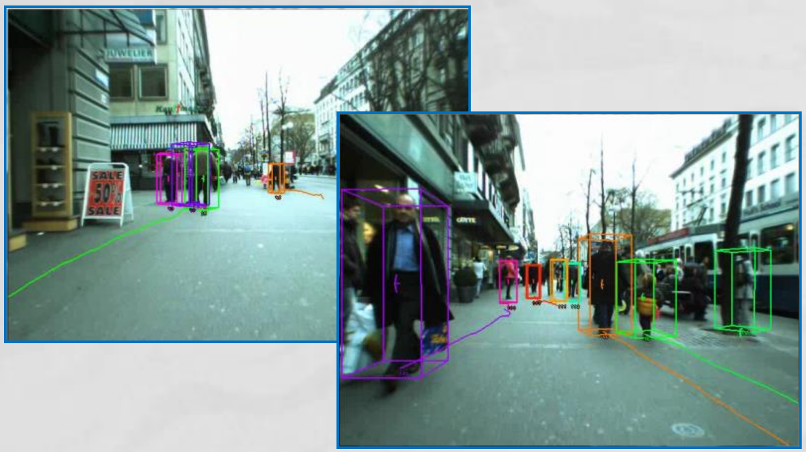
Object Tracking

Lecturer: Sang Hwa Lee

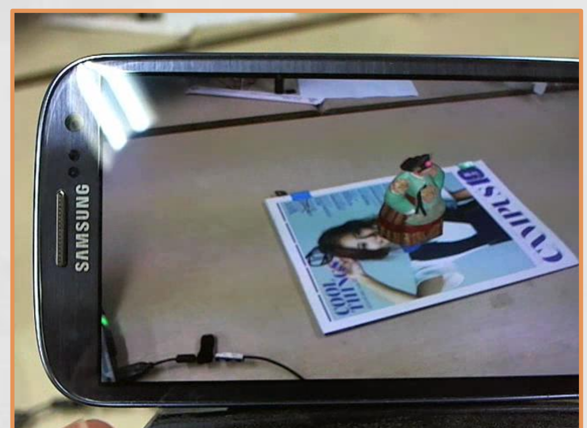
What is Tracking?

- Continuous identification and localization of moving object
  - ✓ estimating the trajectory with temporal consistency



Similar Subjects

- Object detection
  - Finding the object independently in the each frame
  - Not considering temporal consistency
- Camera Tracking
  - Estimation of camera motion and poses from image feature/marker matching
  - AR applications



Main Issues in Tracking

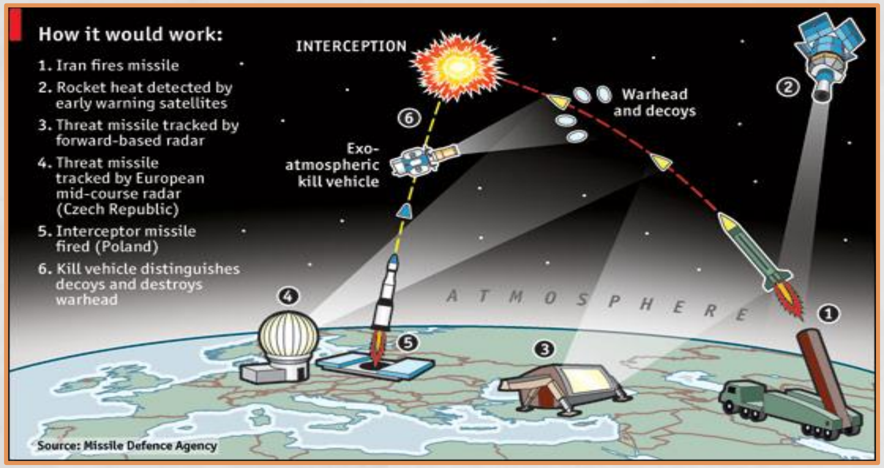
- Satets: What to define as object tracking
  - Position, size, shape, pose, . . .
- Likelihood model: How to measure the similarity
  - Template (block) matching, color histogram, shape, ...
- Prior model: How to model the motion trajectory
  - Prediction models
- Where to search for the object in the next frame
  - Full search, sampling, ...
  - Particle filtering

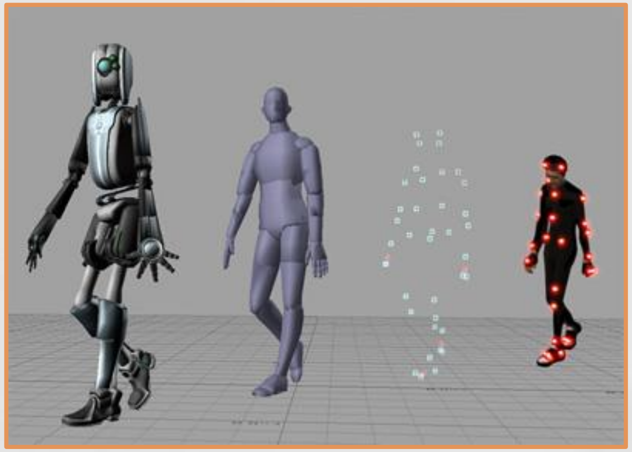
Three Categories

- Kalman filter
  - Classical and popular method
  - Linear prediction model of motion
- Particle filter
  - nonlinear and stochastic process
  - Condensation, adaptive sampling
- CAMSHIFT
  - mean shift process

Applications

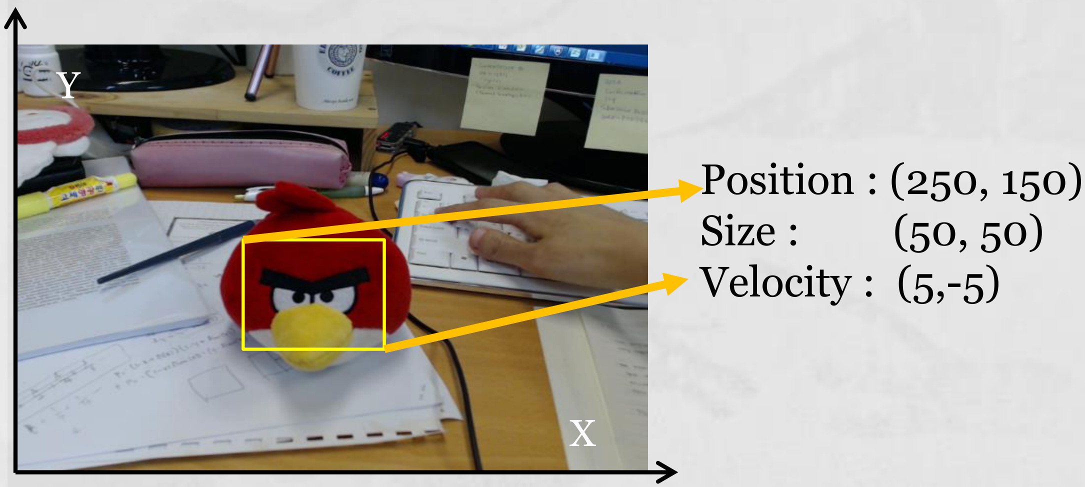
- Video surveillance
  - Human/car object tracking
- Military system
  - Missile targeting
- Entertainment
  - Video games and motion capture

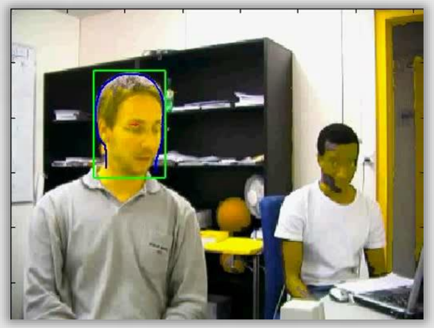




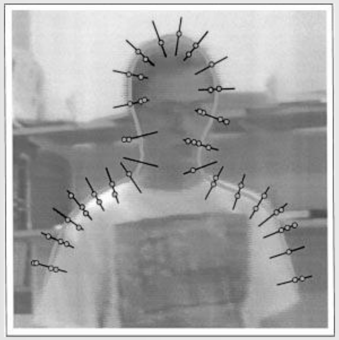
Basic Terminology (1)

- State (X)
  - Tracking results described in vector
  - Position, Size, Velocity, Acceleration etc.
- Observation (likelihood) (Y)
  - Visual cues of object to match
  - Appearance model of object – color, shape





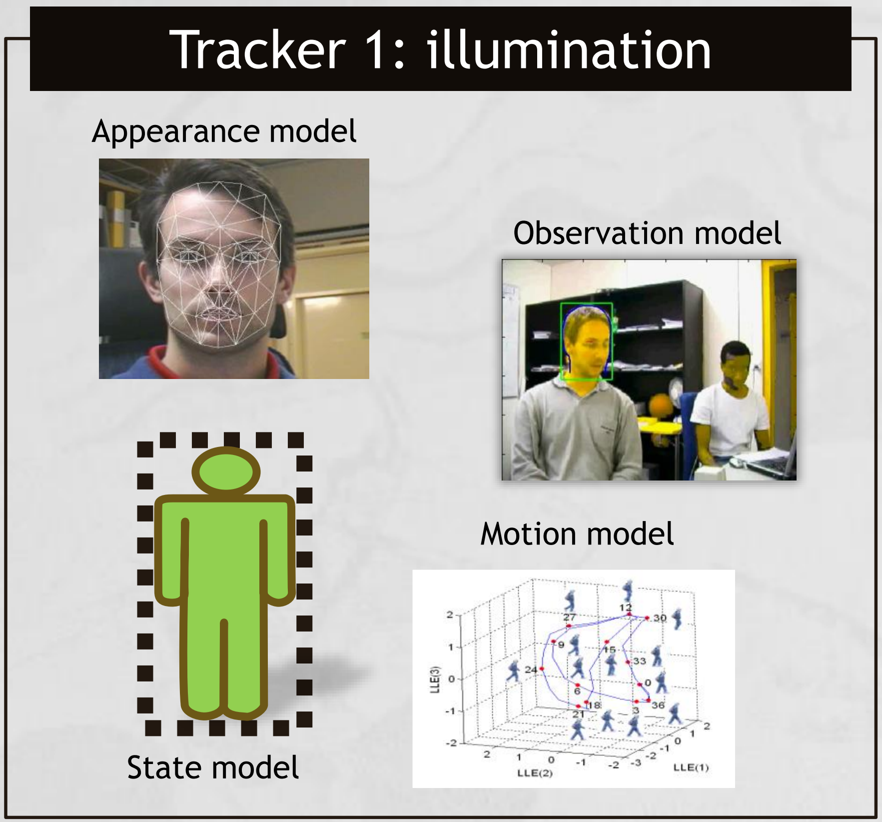
Color

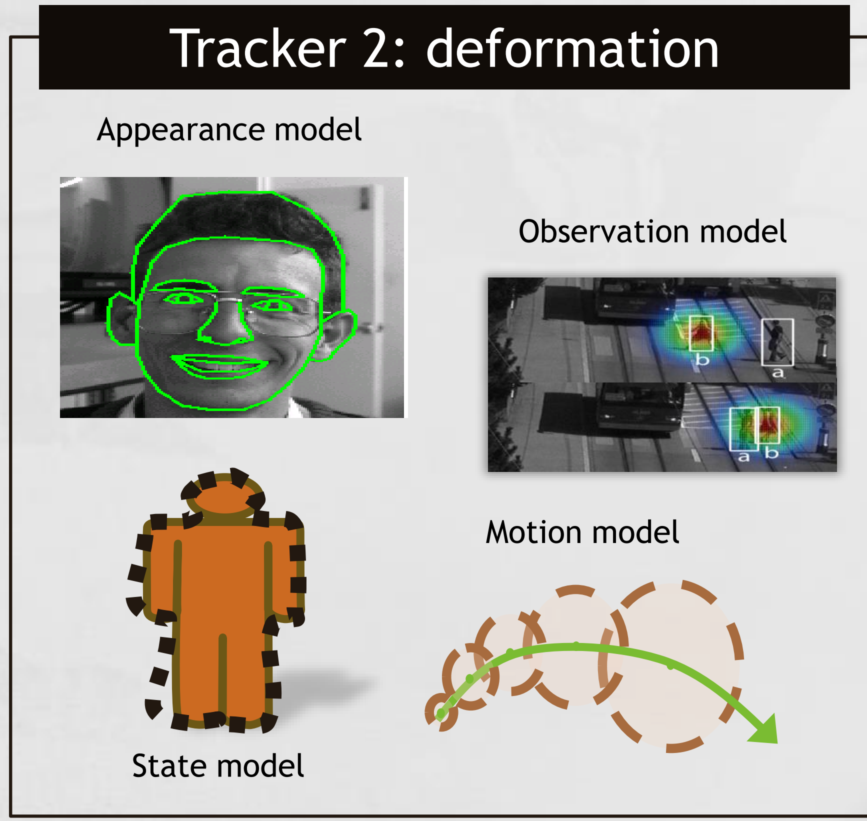


Contour

Basic Terminology (2)

- Tracker:
  - A complete module with observation and prior model
  - Multiple trackers are designed in a tracking system

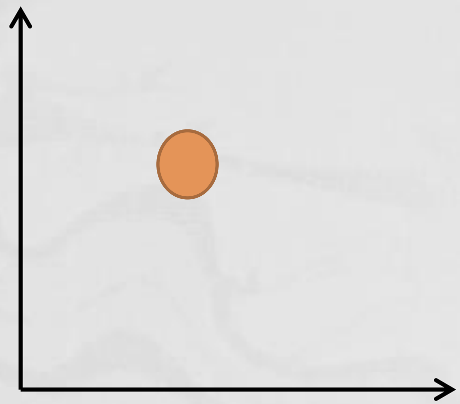


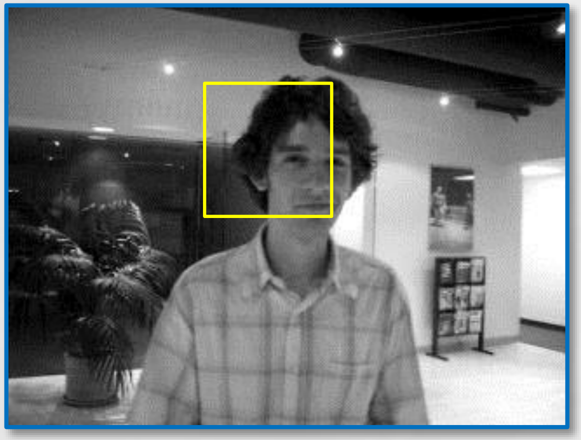


Basic Steps of Tracking (1)

Step 1. Detection

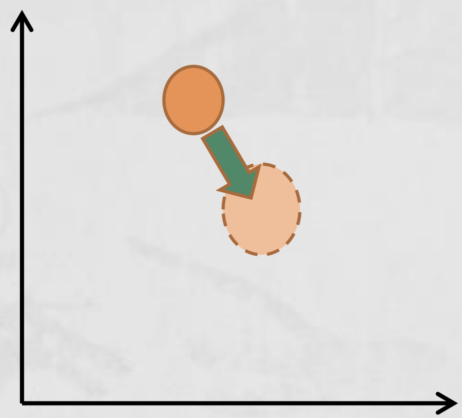
- Reference model

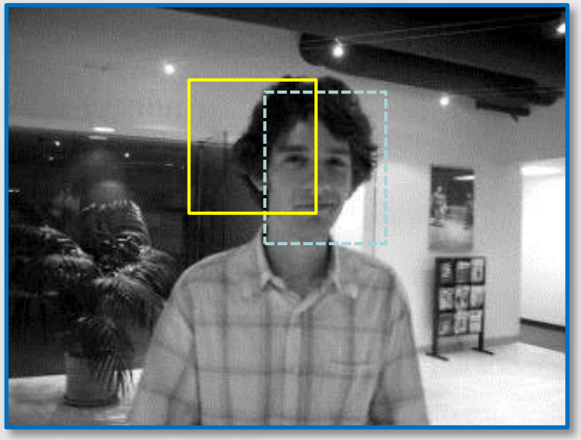




Step 2. Prediction

- Prior modeling
- Particles





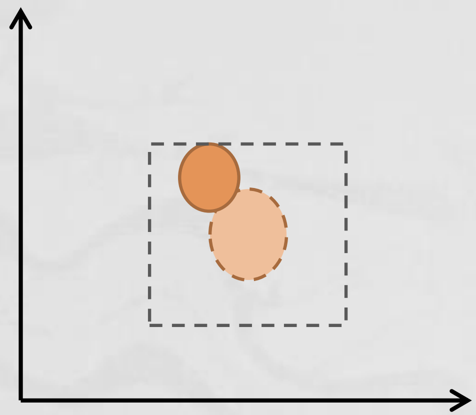
Basic Step of Tracking (2)

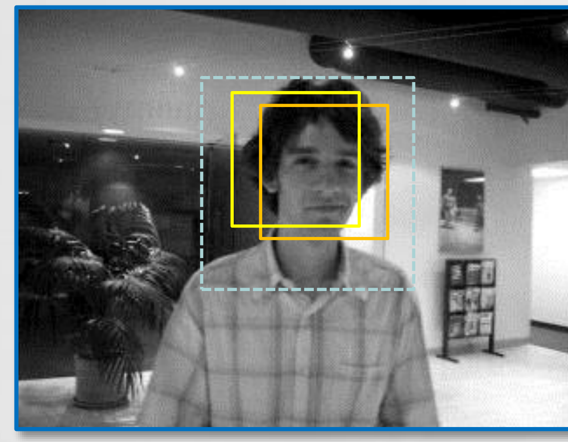
Step 3. Matching

-Observation model

-Likelihood model

-Maximization



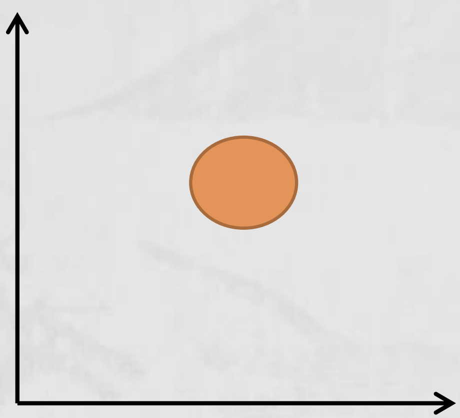


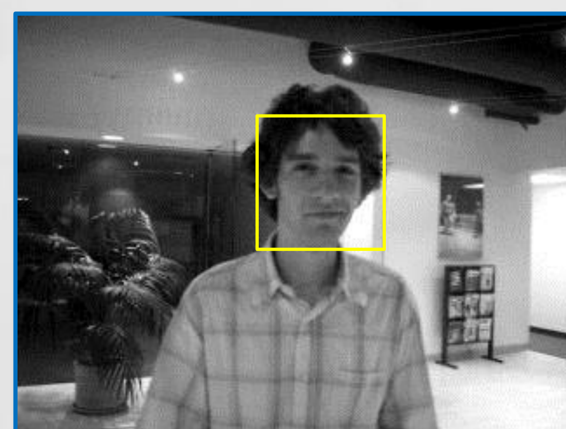
Step 4. Update

-States updates

-Reference model

-Expectation





Two Tracking Processes (1)

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

$$P\left(X_t \middle| X_{t-1}, X_{t-2}, \ldots. X_0\right)$$

- **Matching**: Compute the similarity (likelihood) between the reference observation model and the observation model of candidate states.

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

Two Tracking Processes (2)

- Considering only the neighboring past states
  - MRF chains

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

dynamics model

- Depending only on the current state
  - No update

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

observation model

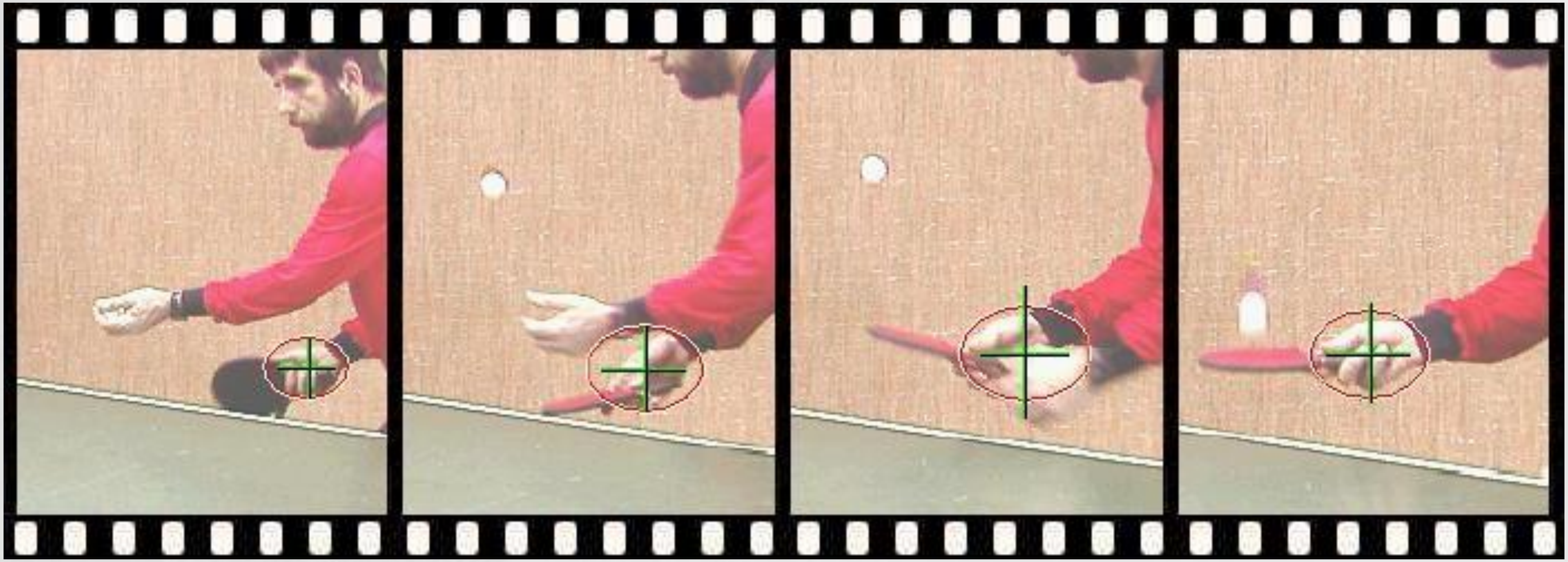
Mean Shift Tracking (CAMSHIFT)

Non-Rigid Object Tracking

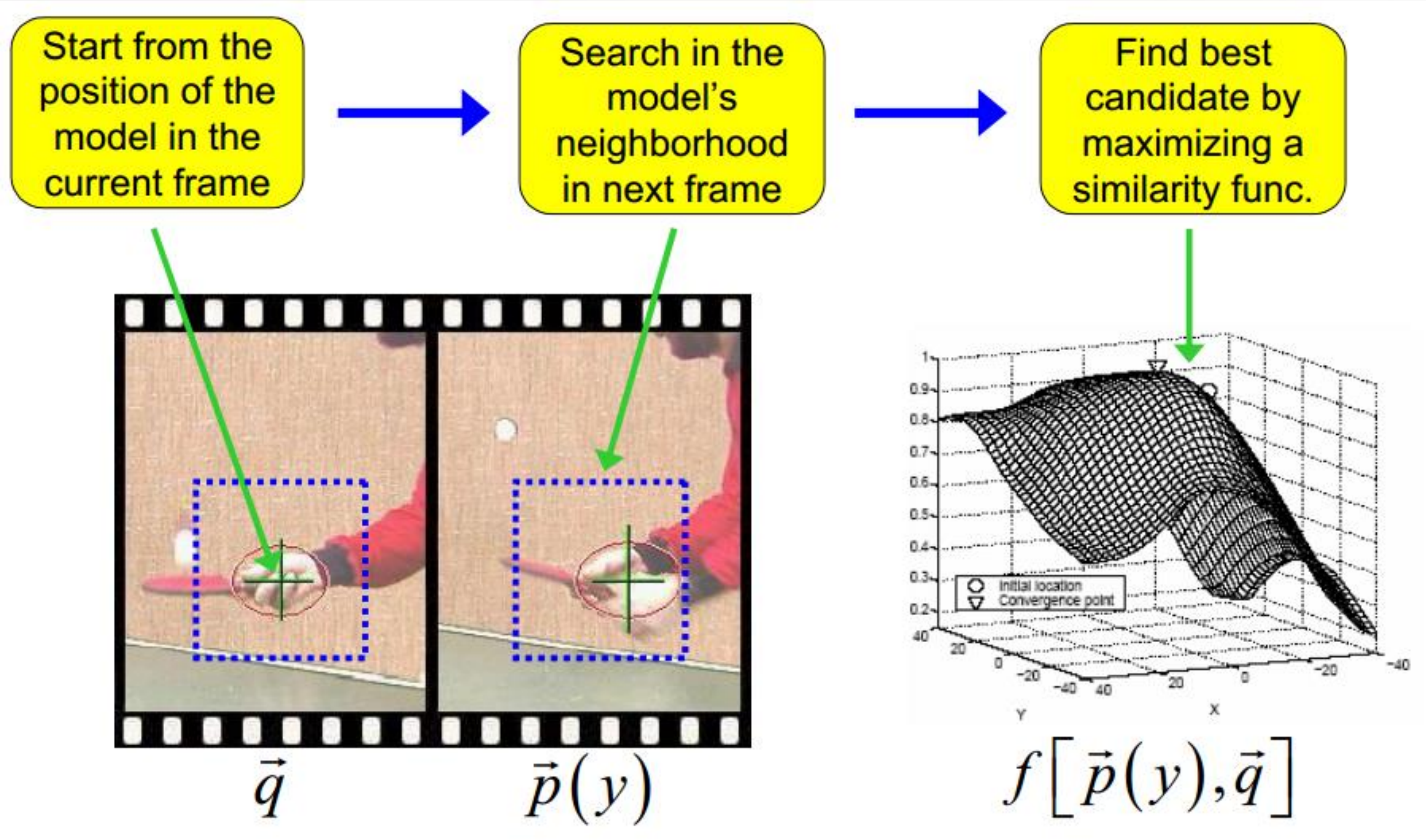
Block Matching is not proper to compare the similarity of non-rigid (deformable) objects.

$\Rightarrow$ Histogram matching

$\Rightarrow$ Mean shift processing



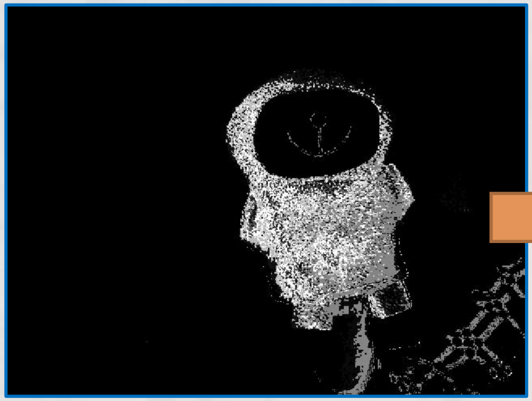
Mean Shift Tracking



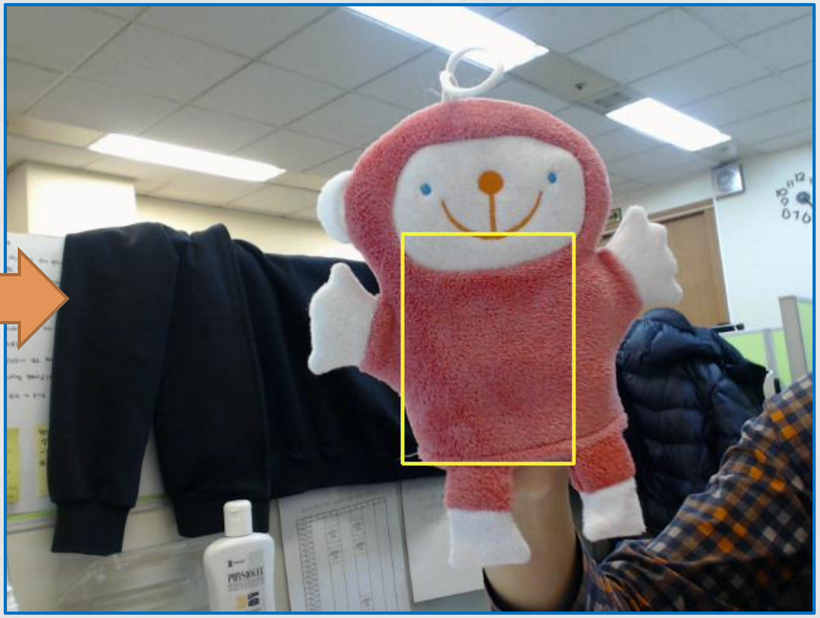
Back Projection

- Assign probability to each pixel using the object histogram
- Mean shift process for the pixels coordinates with the probability weights.

color histogram



back projection

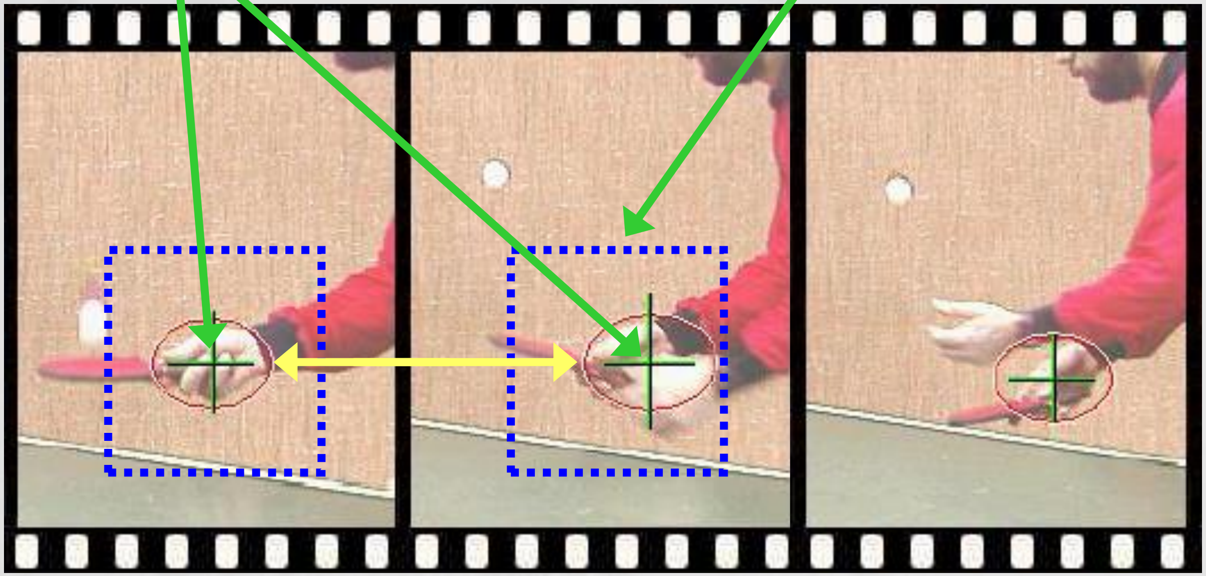


Mean shift process

Mean-Shift Object Tracking

General Framework: Target Localization

Start from the position of the model in the current frame → Search in the model's neighborhood in next frame → Find best candidate by maximizing a similarity func.



Model

Candidate

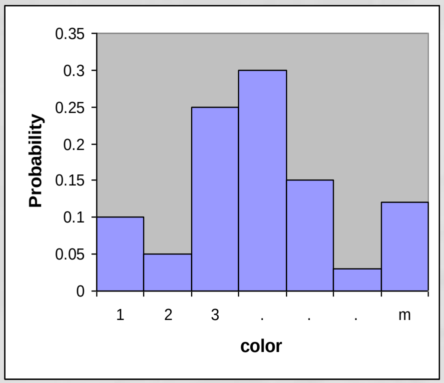
… Current frame → …

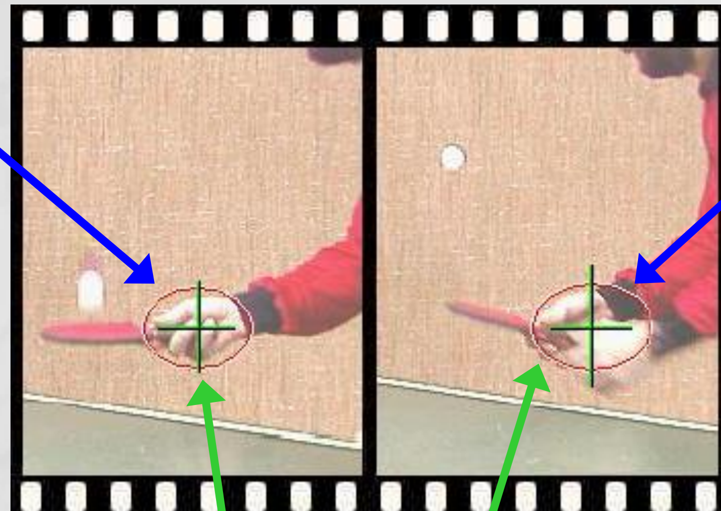
Repeat the same process in the next pair of frames

Mean-Shift Object Tracking

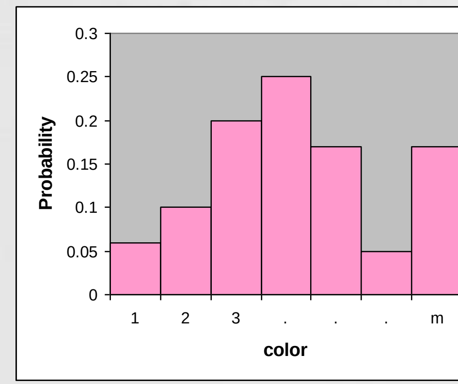
PDF Representation

Target Model
(centered at 0)





Target Candidate
(centered at y)



$$\vec{q} = \{q_u\}_{u=1..m} \qquad \sum_{u=1}^{m} q_u = 1$$

$$\vec{p}(y) = \{p_u(y)\}_{u=1..m} \qquad \sum_{u=1}^{m} p_u = 1$$

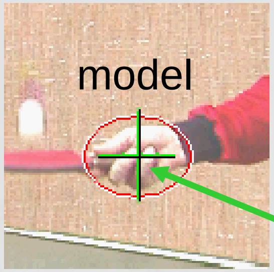
Similarity Function:

$$f(y) = f\left[\vec{q}, \vec{p}(y)\right]$$

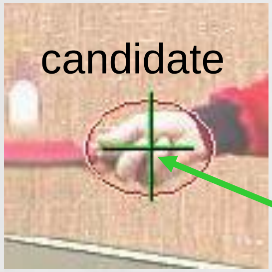
Mean-Shift Object Tracking

Finding the PDF of the target model

$\{x_i\}_{i=1..n}$ Target pixel locations



0



y

$k(x)$ A differentiable, isotropic, convex, monotonically decreasing kernel

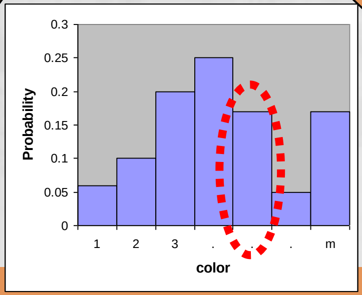
- Peripheral pixels are affected by occlusion and background interference

$b(x)$ The color bin index (1..*m*) of pixel *x*

Probability of feature u in model

$$q_u = C \sum_{b(x_i)=u} \underbrace{k\left(\left\|x_i\right\|^2\right)}$$

Normalization factor

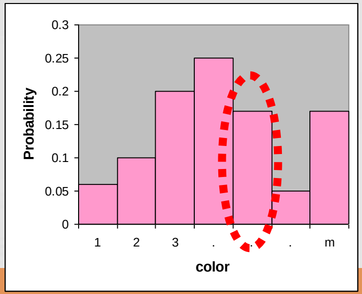


Pixel weight

Probability of feature u in candidate

$$p_u(y) = C_h \sum_{b(x_i)=u} \underbrace{k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)}$$

Normalization factor



Pixel weight

Mean-Shift Object Tracking

Similarity Function

Target model: $\vec{q} = \left(q_1, \ldots, q_m\right)$

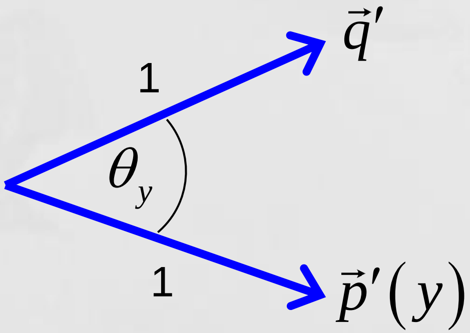
Target candidate: $\vec{p}(y) = \left(p_1(y), \ldots, p_m(y)\right)$

Similarity function: $f(y) = f\left[\vec{p}(y), \vec{q}\right] = ?$

The Bhattacharyya Coefficient

$$\vec{q}\,' = \left(\sqrt{q_1}, \ldots, \sqrt{q_m}\right)$$

$$\vec{p}\,'(y) = \left(\sqrt{p_1(y)}, \ldots, \sqrt{p_m(y)}\right)$$



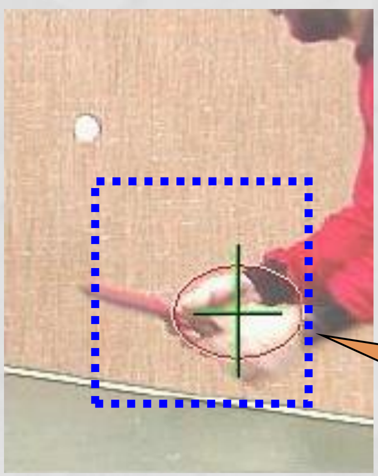
$$f(y) = \cos\theta_y = \frac{p'(y)^T q'}{\|p'(y)\| \cdot \|q'\|} = \sum_{u=1}^{m} \sqrt{p_u(y) q_u}$$

Mean-Shift Object Tracking

Maximizing the Similarity Function

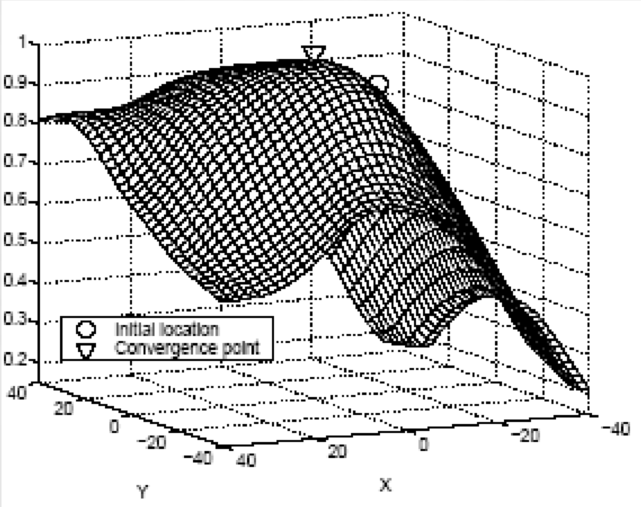
The mode of $$\frac{C_h}{2}\sum_{i=1}^{n} w_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)$$ = sought maximum

Important Assumption:



The target representation provides sufficient discrimination

One mode in the searched neighborhood



$$f\left[\vec{p}(y),\vec{q}\right]$$

Mean-Shift Object Tracking

Applying Mean-Shift

The mode of $\frac{C_h}{2}\sum_{i=1}^{n} w_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)$ = sought maximum

Original Mean-Shift: Find mode of $c\sum_{i=1}^{n} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)$ using $y_1 = \frac{\sum_{i=1}^{n} x_i g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}$

Extended Mean-Shift: Find mode of $c\sum_{i=1}^{n} \boxed{w_i} k\left(\left\|\frac{y-x_i}{h}\right\|^2\right)$ using $y_1 = \frac{\sum_{i=1}^{n} x_i \boxed{w_i} g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} \boxed{w_i} g\left(\left\|\frac{y_0-x_i}{h}\right\|^2\right)}$

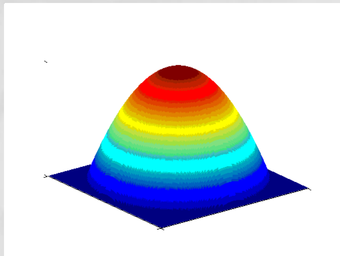
Mean-Shift Object Tracking

Choosing the Kernel

A special class of radially symmetric kernels:

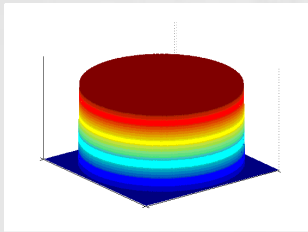
$$K(x) = ck\left(\|x\|^2\right)$$

Epanechnikov kernel:



$$k(x) = \begin{cases} 1-x & \text{if } \|x\| \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Uniform kernel:



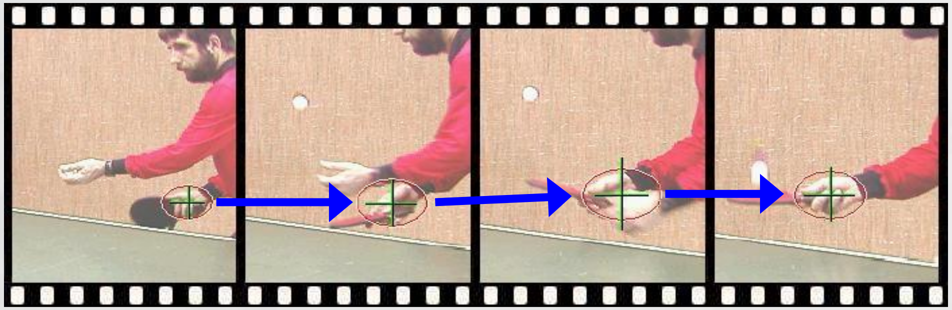
$$g(x) = -k(x) = \begin{cases} 1 & \text{if } \|x\| \le 1 \\ 0 & \text{otherwise} \end{cases}$$

$$y_1 = \frac{\sum_{i=1}^{n} x_i w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)} \longrightarrow y_1 = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i}$$

Mean-Shift Object Tracking
Adaptive Scale

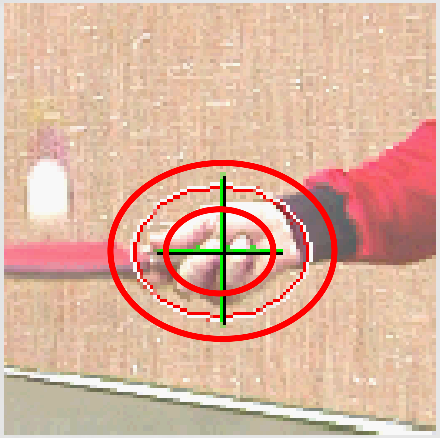
Problem:

The scale of the target changes in time → The scale (*h*) of the kernel must be adapted



Solution:

Run localization 3 times with different *h* → Choose *h* that achieves maximum similarity



Tracking Result

- PTZ control

Particle Filter Tracking

Particle Filter Algorithm

- 1. Select samples (particles) using dynamic (prediction) models
- 2. Calculate the likelihood and prior model for each state (particle)
- 3. Calculate the expected state from the calculated posterior probability
- 4. Update the state and reference model

Non-Gaussian Approximation

- Limitation of Kalman filter and its variations
  - Need a guarantee of linearity to use Kalman filter
  - Nonlinear dynamics create serious problem

- Potential solution
  - Particle Filter

Bayesian Filtering

- Estimation of Unknown Probability Distribution

State variable

Observation variable
(previous state +likelihood)

$$\underbrace{p(x_t|z_{1:t})}_{\text{posterior}} \propto \underbrace{p(z_t|x_t)}_{\text{likelihood}}\underbrace{p(x_t|z_{1:t-1})}_{\text{prior}}$$

$$= p(z_t|z_t) \int \underbrace{p(x_t|x_{t-1})}_{\text{Process transition probability}} p(x_{t-1}|z_{1:t-1}) dx_{t-1}$$

- Process model (dynamic model)

$$x_{t|t-1} = f(x_{t-1}, a_{t-1}) + u_{t-1}$$

Prior Model (1)

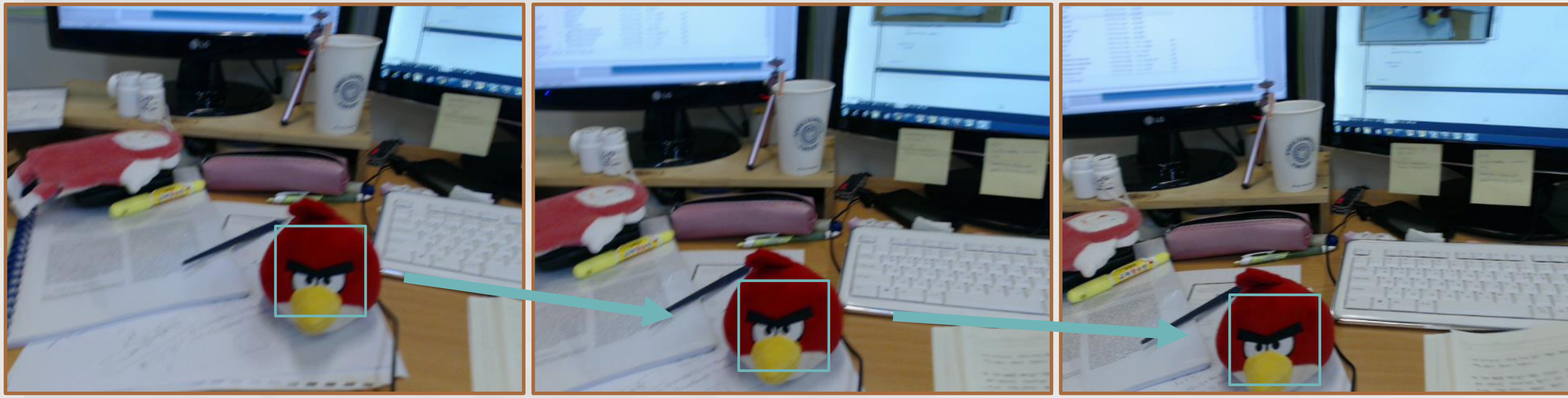
- Temporal transition models of state
  - Assume that object dynamics form a temporal Markov chain:

$$p(x_t|X_{t-1}) = p(x_t|x_{t-1})$$

- Dynamics of state
  - States are positiom, size, velocity,...

$$p(x_t|x_{t-1}) \propto \exp(-\frac{1}{2}(x_t - x_{t-1})^2)$$

Prior Model (2)



Tracking property is based on temporal coherence.

Dynamical Model :

1. Velocity is similar between frame to frame

$$x_t = x_{t-1} + v_{t-1} + N(0, v_{t-1})$$

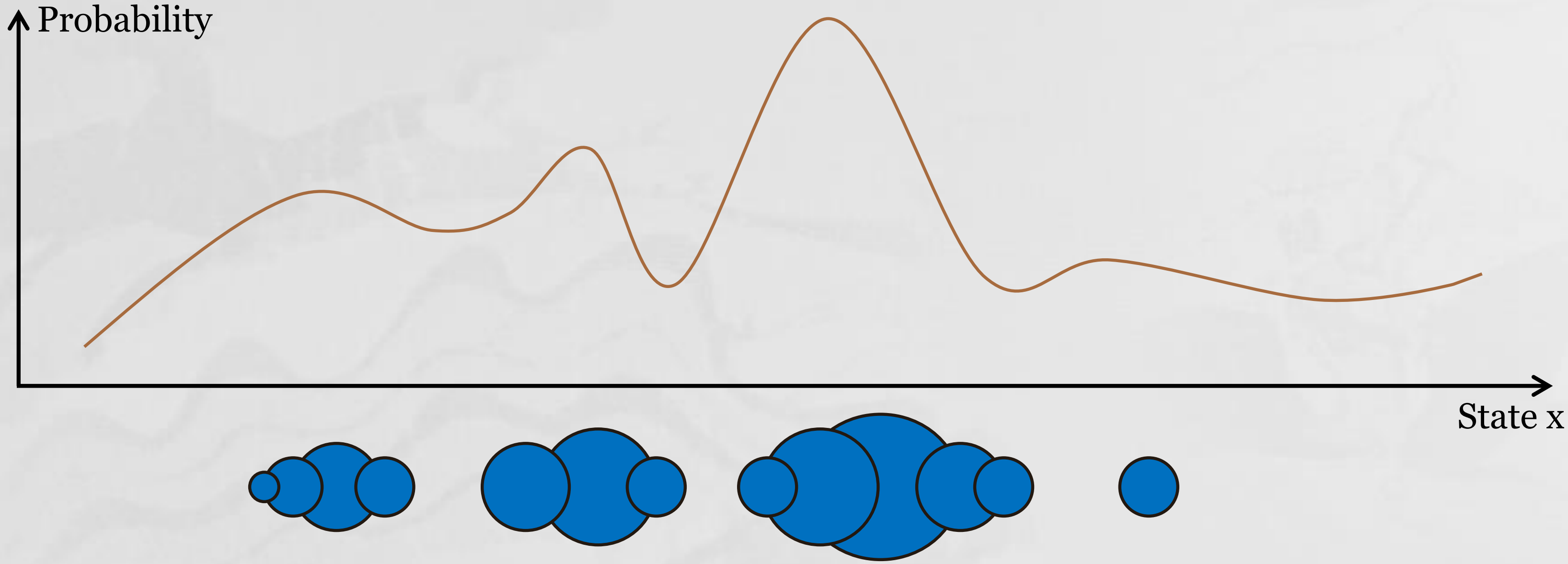
2. Size is similar between frame to frame

$$x_t = x_{t-1} + N(0, x_{t-1})$$

Monte Carlo Approximation

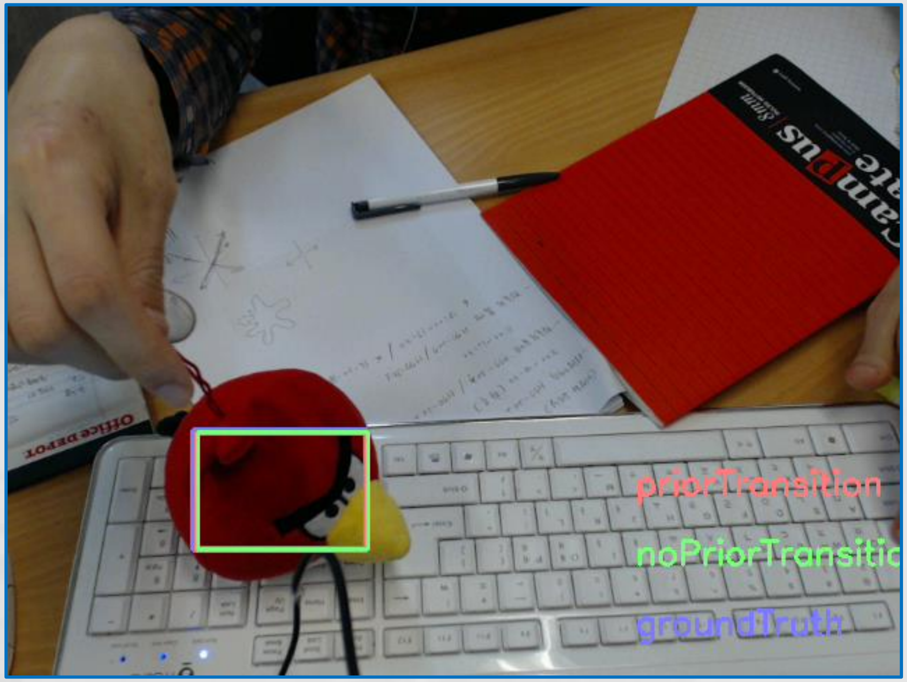
- Basic Idea
  - Sampling Based
  - The more we draw samples, the more accurate the estimation is.
  - Useful when analytical solution is unknown or hard to compute
- MC in sequential Bayesian filtering
  - Estimating unknown probability distribution using a set of samples
  - MC can easily be used to compute marginal posterior distribution $p(x|z)$
  - No need to any assumption

Factored Sampling



Randomly sampling from a prior density.
Each sample is assigned a weight π_i in proportion to the value of the observation density $p(z|x = s^{(n)})$

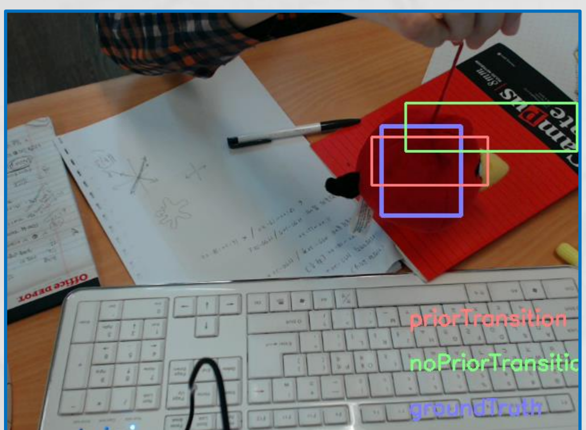
Tracking Example

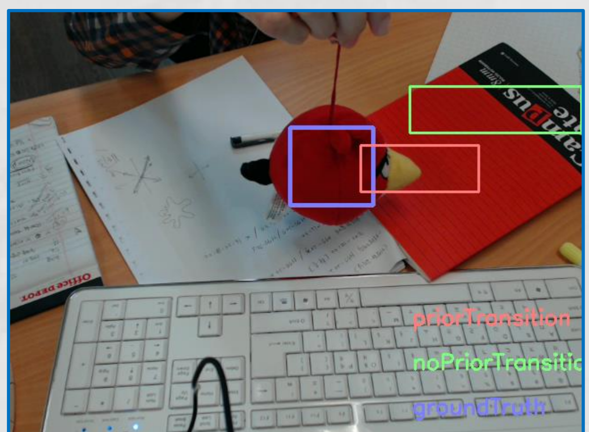


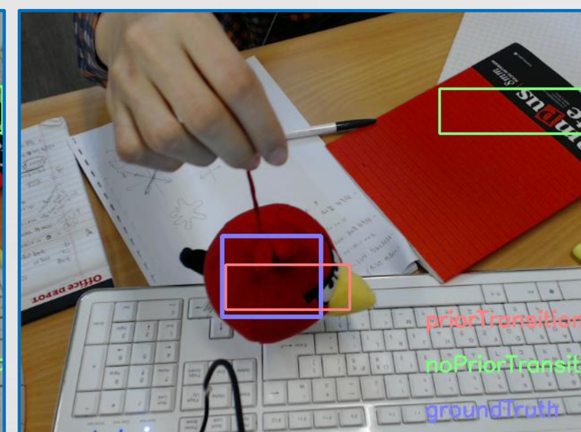
Prior

Only Likelihood

Ground truth







Tracking Comparison

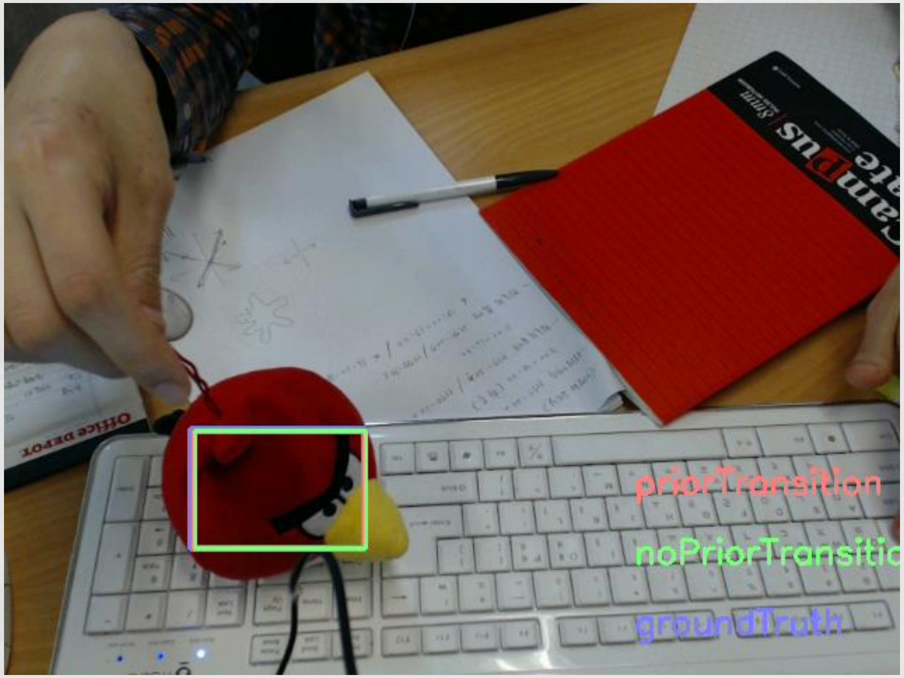
- Proposed method
- openCV CAMSHIFT

Some Comments (1)

- Proper prior model (transition model, dynamics, prediction model) improves the tracking performances.
- Factored sampling is more effective than randomly uniform sampling.
- Likelihood model (similarity measurement) is the most important factor in object tracking.
- In most cases, there are little differences in tracking performances among Kalman, particle, and mean shift filters.

Some Comments (2)

- Challenge problems
  - Similar objects: colors, shapes
  - Occlusions



Recent Trends in Tracking

- Multi-tracker and observation model bank
  - Multiple color spaces
  - Different number of histogram bins
  - various observation measurements
- Sampling based
  - Optimal tracker selection
  - Particle filtering
  - Likelihood model selection
  - Weighted sampling of matching
- Tradeoff between accuracy and speed