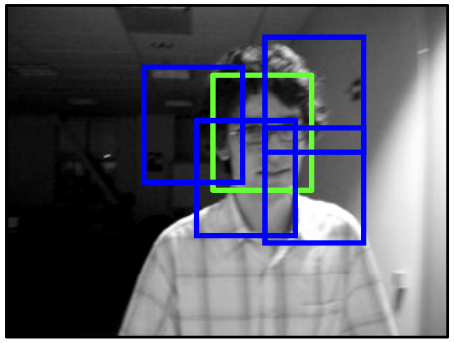
Multimedia Computing

Case Study:

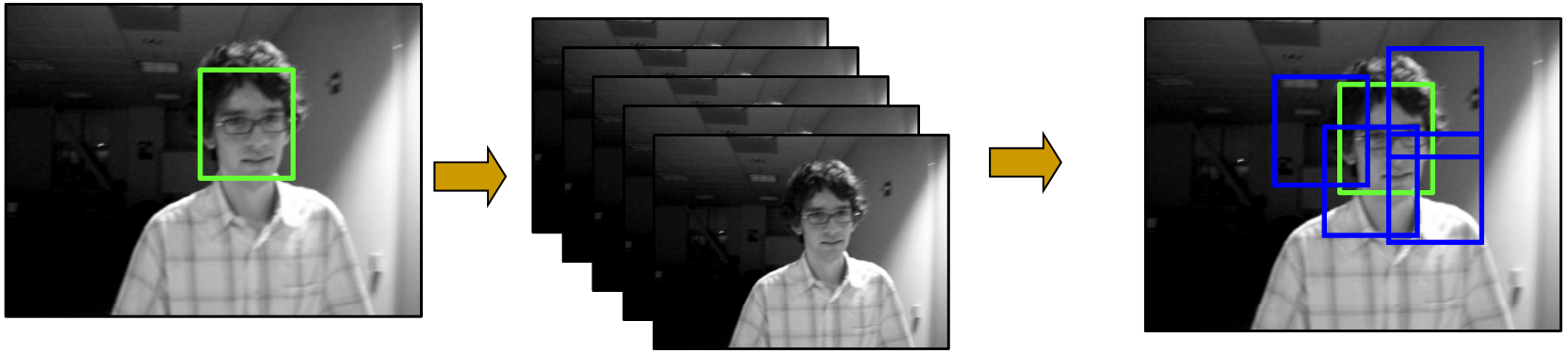
Object Tracking



General tracking system

- A tracking system usually has three parts:
  - Appearance model
  - Motion model
  - Search strategy

General procedures of tracking



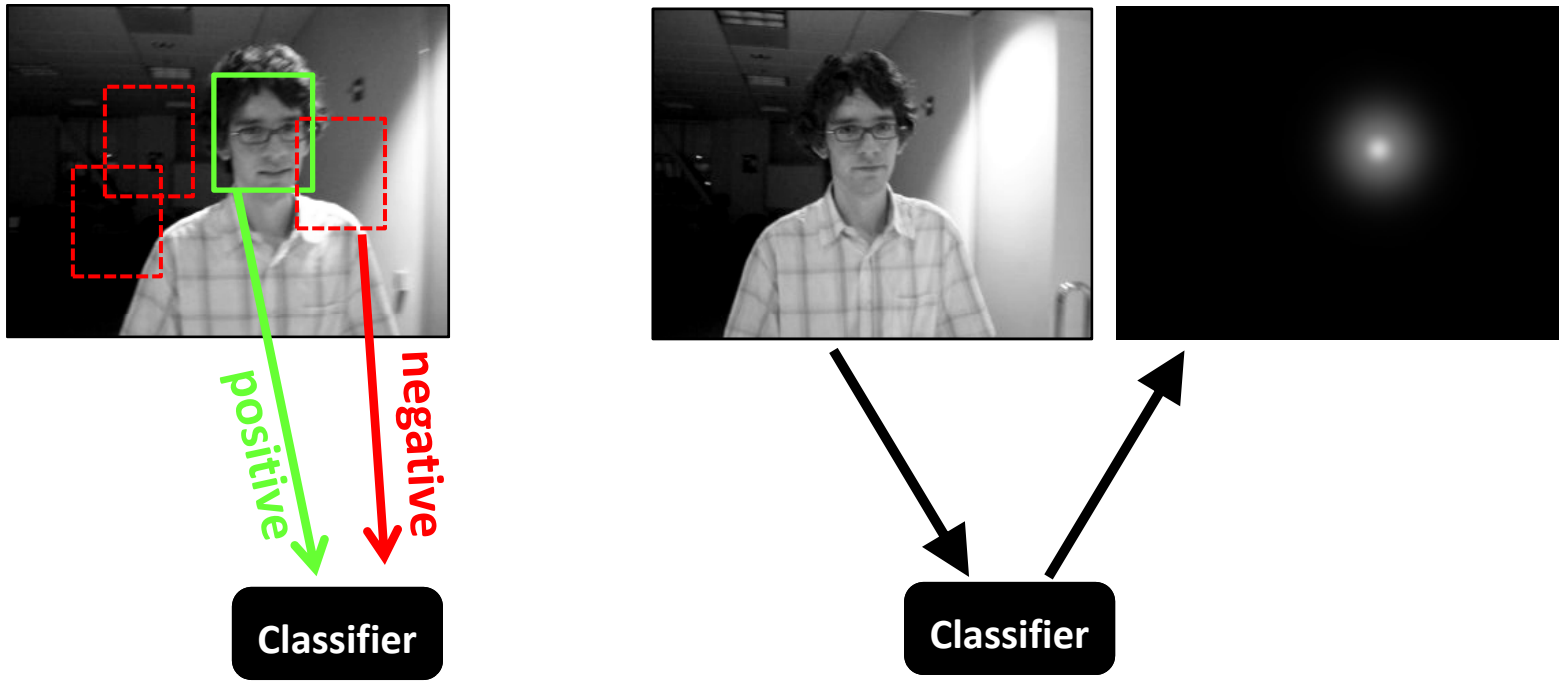
The object location in the first frame is labeled manually or detected automatically by some detection method

Training model: classifier or appearance model.

Search in current frame the object based on classifier response or appearance model.

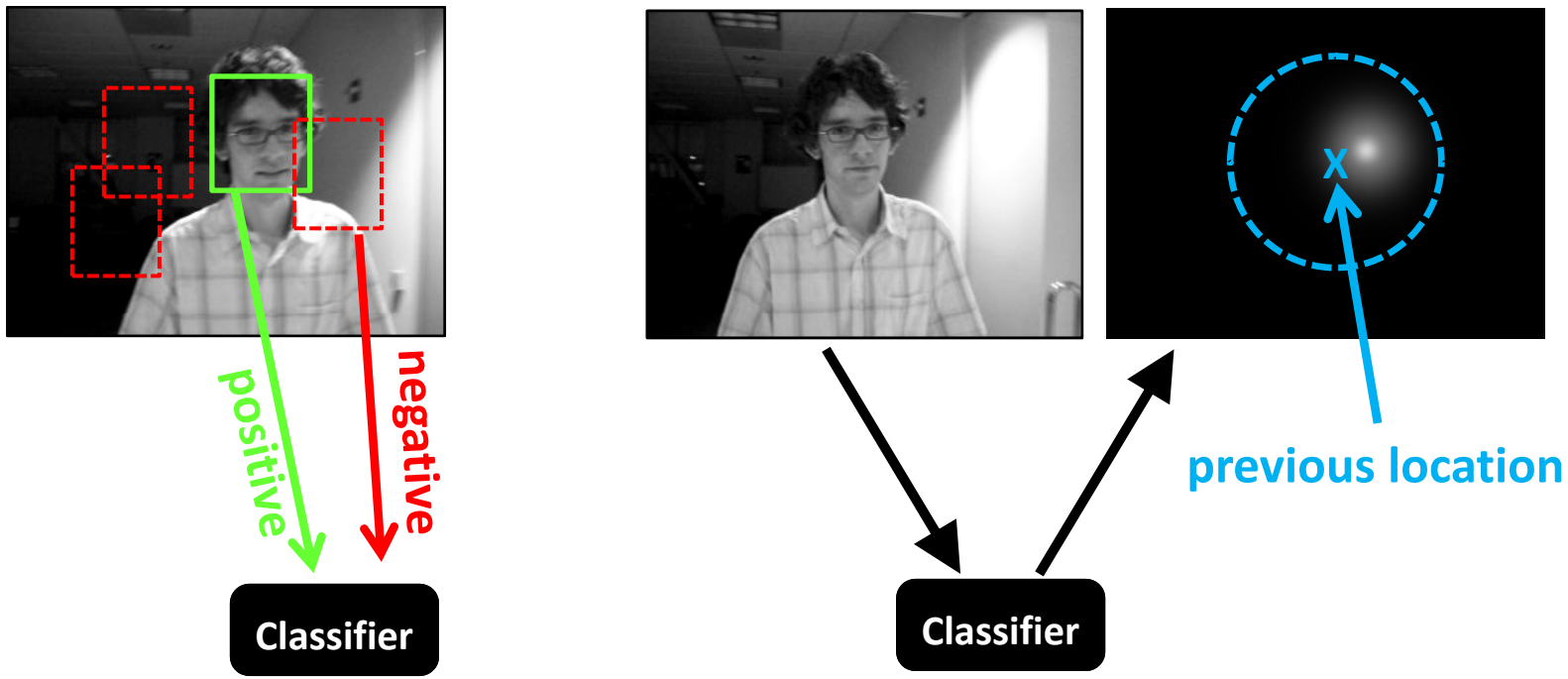
Procedures of tracking

- Evaluate classifier in a search window



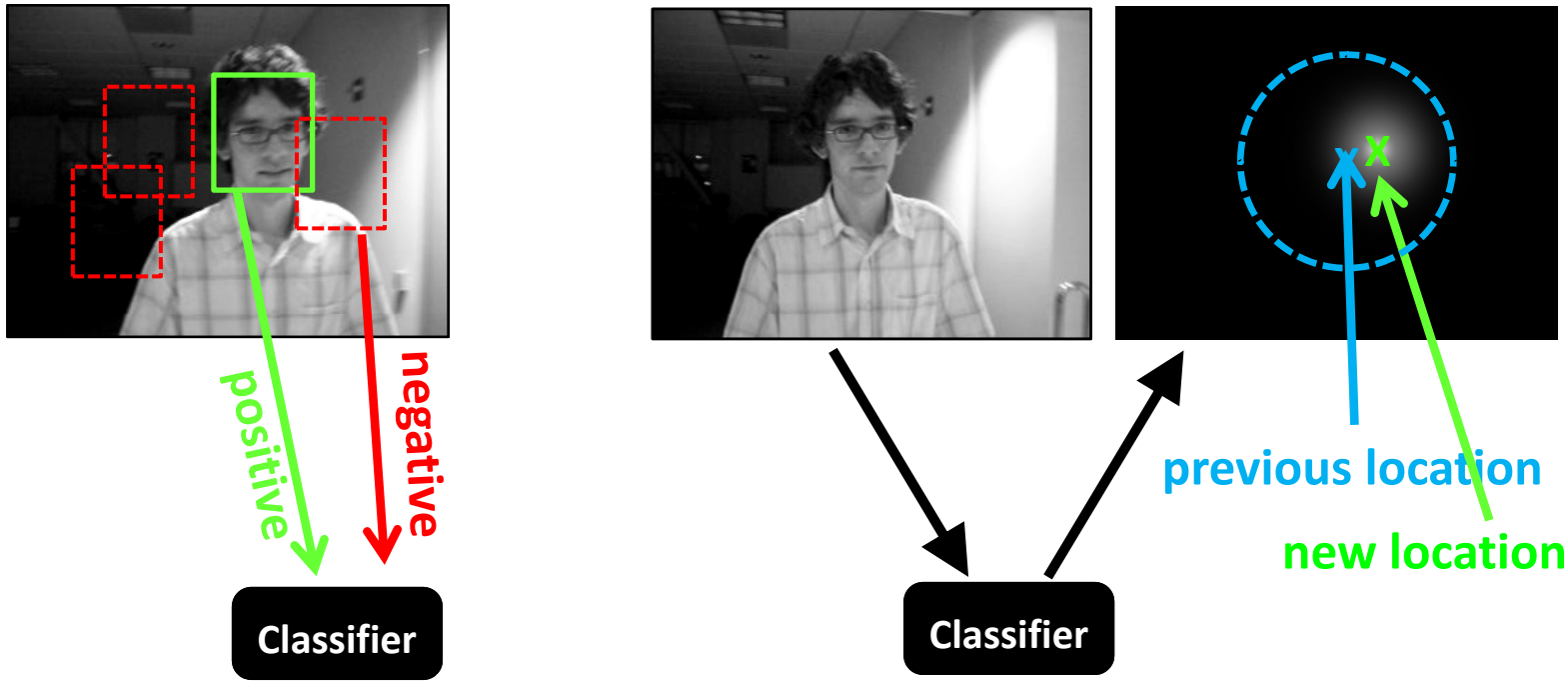
Procedures of tracking

- Evaluate classifier in a search window



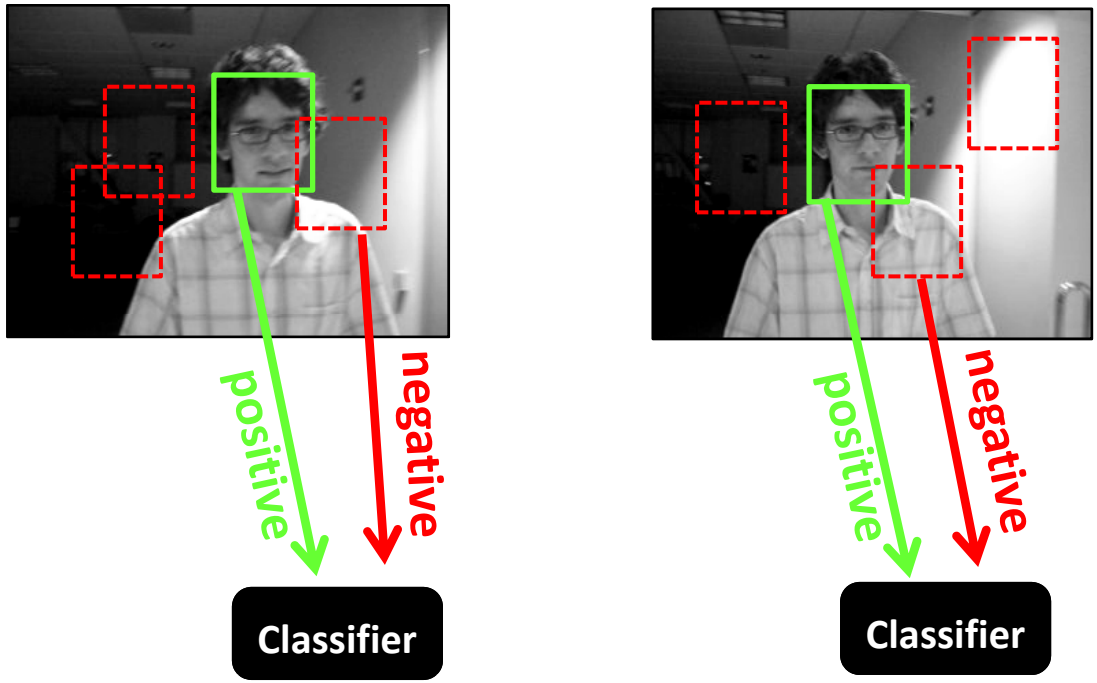
Procedures of tracking

- Find maximum response



Procedures of tracking

- Repeat…

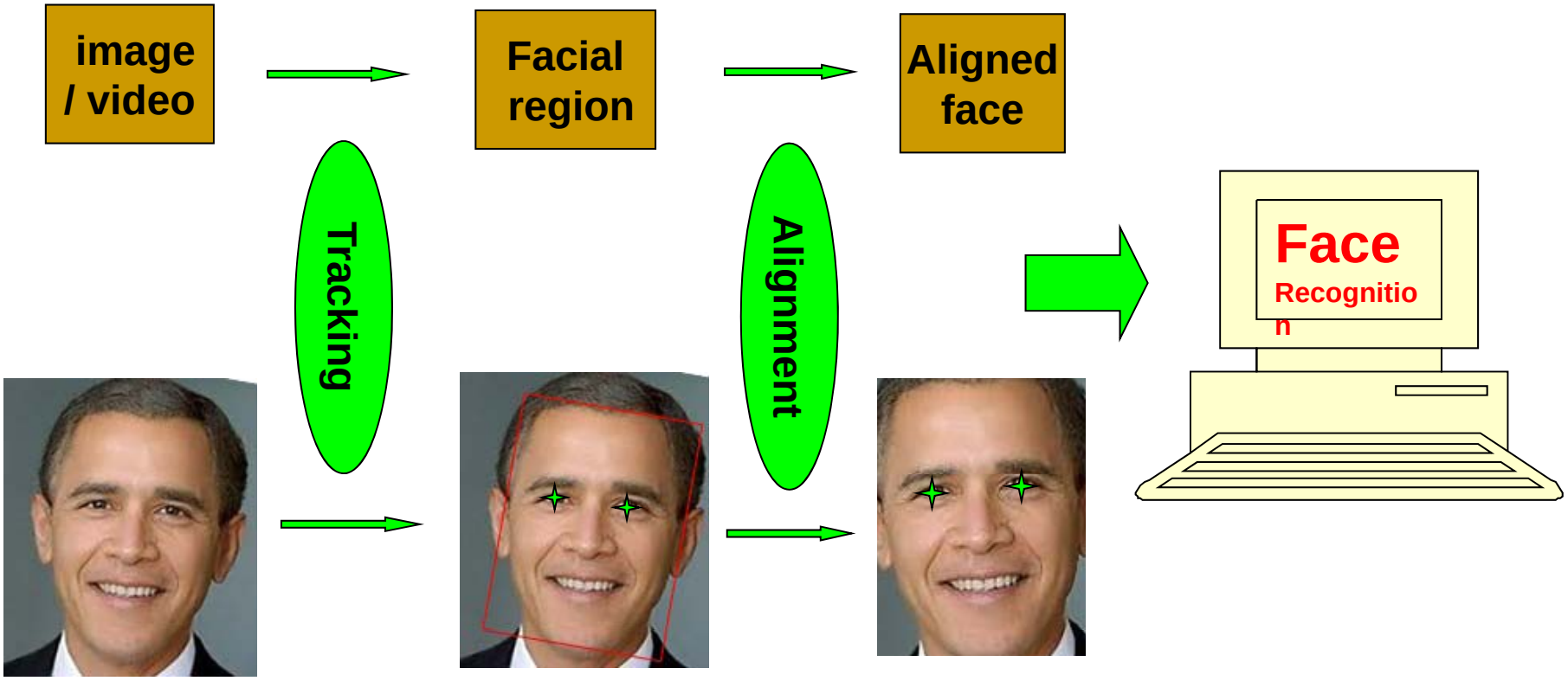


Examples

- Kinect : a motion sensing input system developed by Microsoft.

Kinect Tracking

Practical face recognition system



Face recognition system

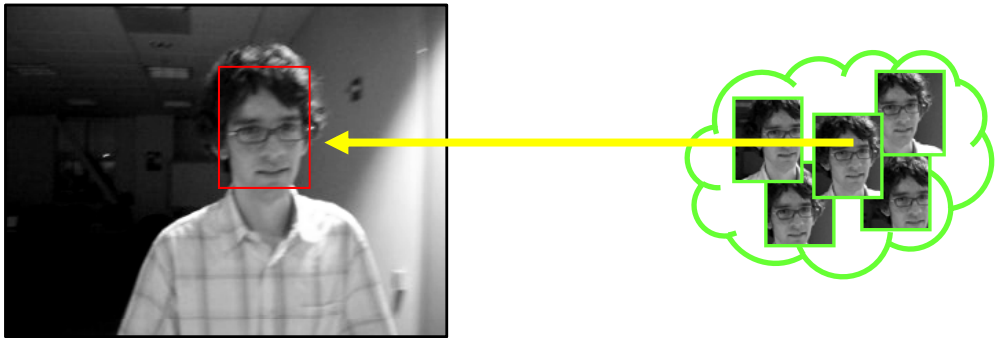
- Tracking face is a prerequisite.

Components of a tracking system

- Appearance model
- Motion model
- A search strategy to find the most likely location in the current frame

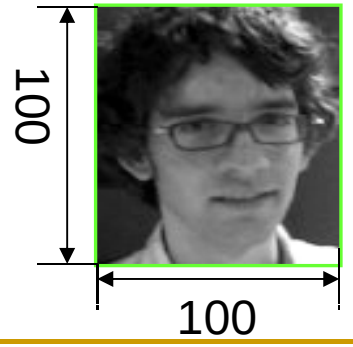
Appearance model

- The appearance model is defined as a model most similar to the object appearance.
  - For example: we can use the original image most similar to the object appearance as the appearance mode.



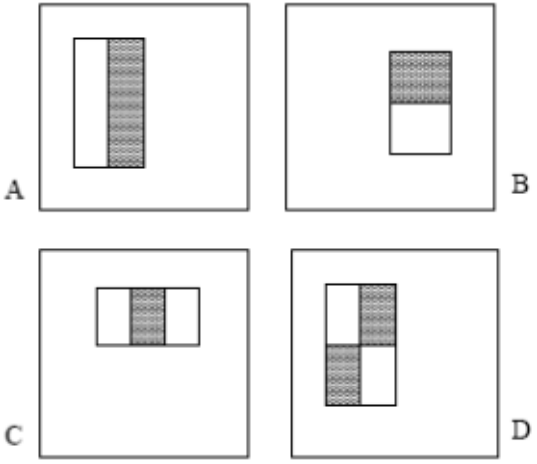
Appearance model

- The dimensionality of the appearance model can be too large.
  - For a 100x100 model, the dimensionality is 10000.
  - Such a large dimension makes the processing very time-consuming.



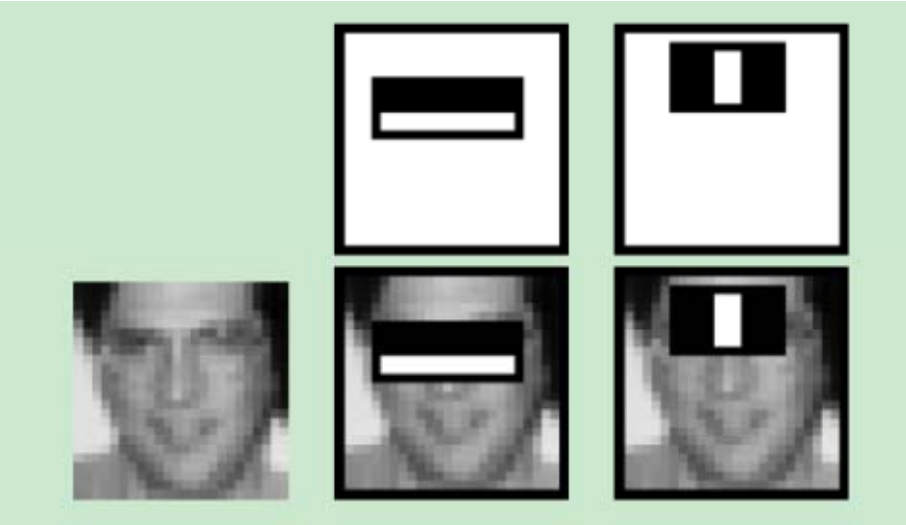
Trained appearance model

- We introduce a simple but very effective feature extraction method.
  - The Haar features



Four types of Haar features.

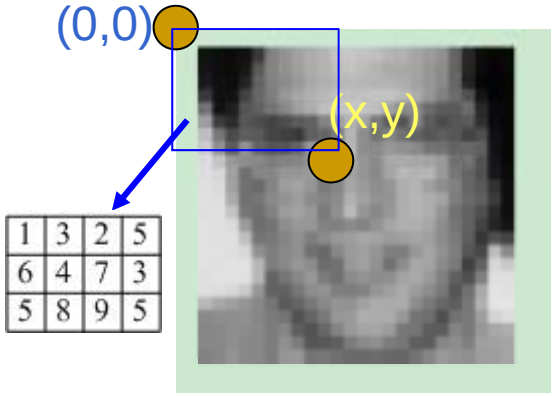
Haar features



- Each Haar feature is defined as the weighted sum of pixels in the rectangle.
  - The weight for the black area is -1 while for the white area is +1.

How to compute Haar features *efficiently* ?

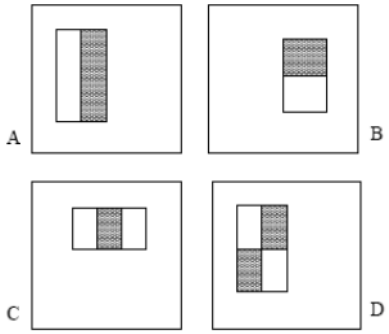
- Integral image method
- The value of integral image at point (x,y) is defined as the sum of pixels in the blue rectangle

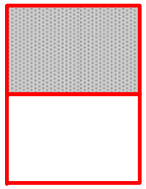


$$ii(x,y) = \int_0^x \int_0^y I(u,v)dudv$$

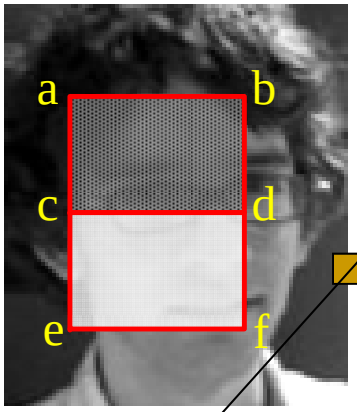
$$ii(x,y) = 1 + 3 + \ldots + 9 + 5 = 58$$

Example





B



1	3	4	2	3	5
4	5	3	1	2	5
2	4	3	2	5	2
3	7	3	1	5	1
4	8	3	3	2	3
5	4	4	2	3	4
7	7	7	3	1	2

Haar B The feature to be computed Image intensity

- We first compute the values of integral image at positions a, b, c, d, e, f

$ii(a) = 1, ii(b) = 1 + 3 + ...4 + 1 = 23$

$ii(c) = 1 + 4 + 2 + 3 = 10, ii(d) = 1 + 3 + ... + 3 + 1 = 48$

$ii(e) = 1 + 4 + 2 + 3 + 4 + 5 = 19, ii(f) = 1 + 3 + ... + 4 + 2 = 81$

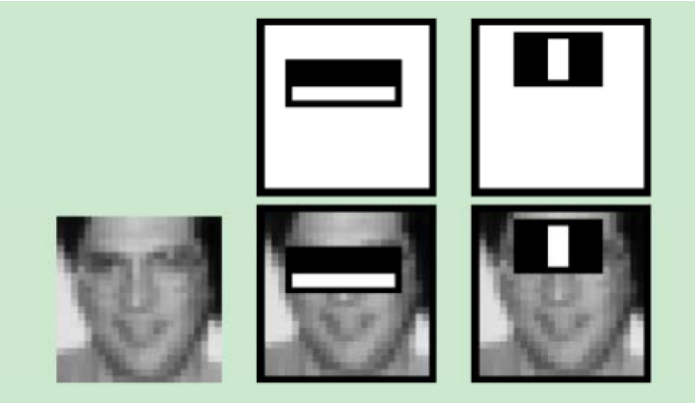
$$top = ii(a) + ii(d) - (ii(c) + ii(b)) = 5 + 48 - (10 + 23) = 20$$

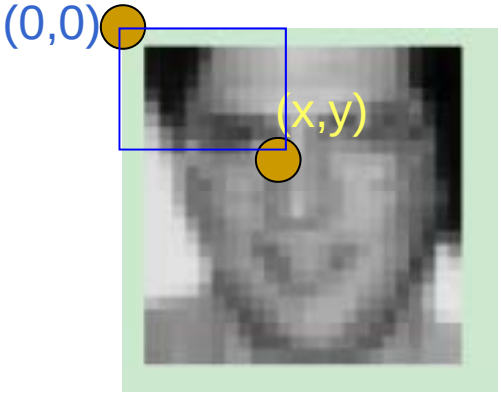
$$bottom = ii(c) + ii(f) - (ii(e) + ii(d)) = 10 + 81 - (19 + 48) = 24$$

Haar = bottom – top = 24 - 20 = 4

Advantages of Haar features

- Very efficient because the integral image can be computed fast.
- Using only a small number of Haar features can well represent the original image. For example, in face detection task, using only two types of Haar features can represent the face.





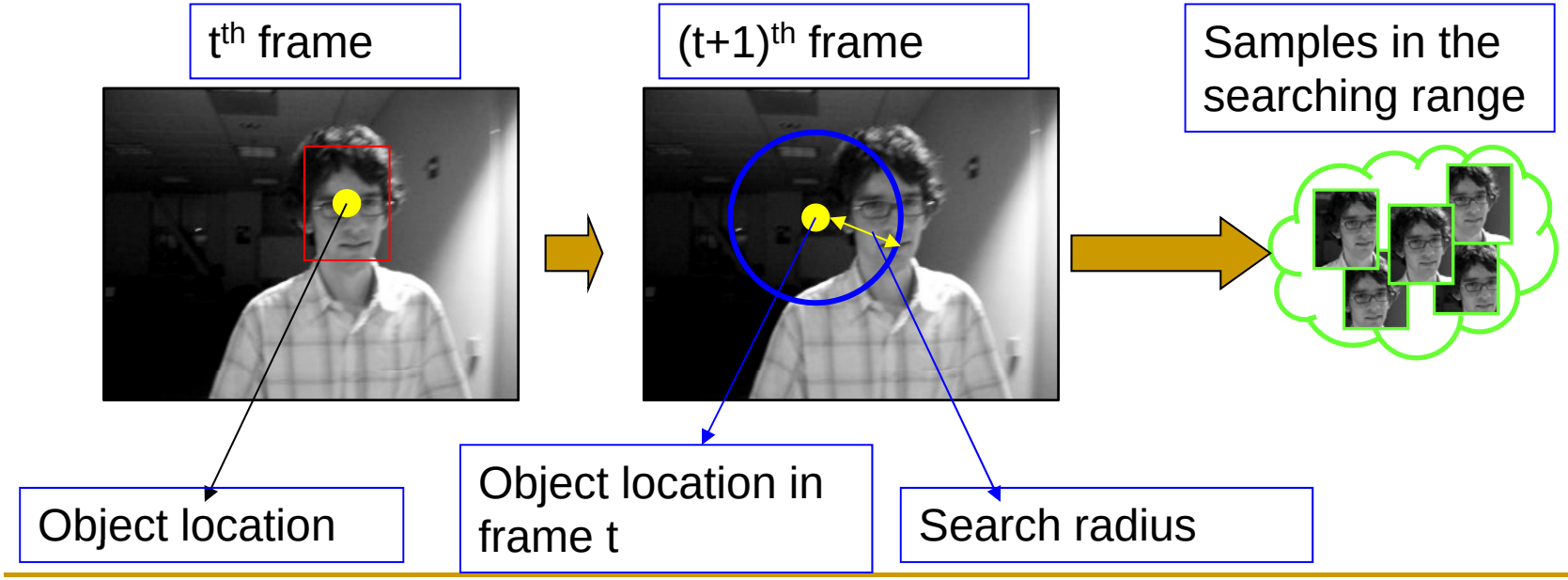
Advantages of Haar features

- Robust to occlusions because it is a type of local feature. For example, the Haar feature in the blue rectangle is localized which is not occluded by the book and cap.

Occlusions

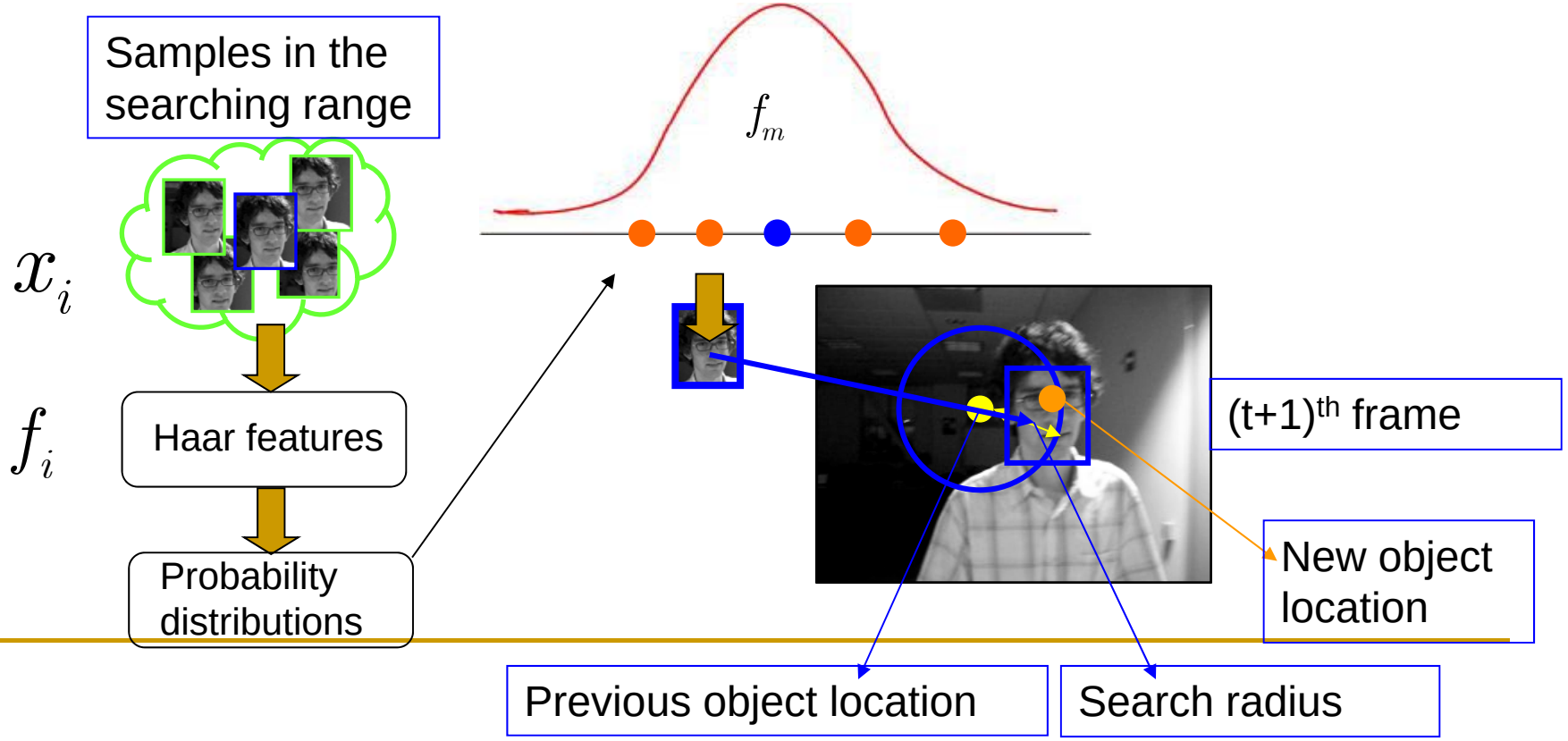
Motion model

- The motion model is very simple. Just move around the object location in the previous frame.



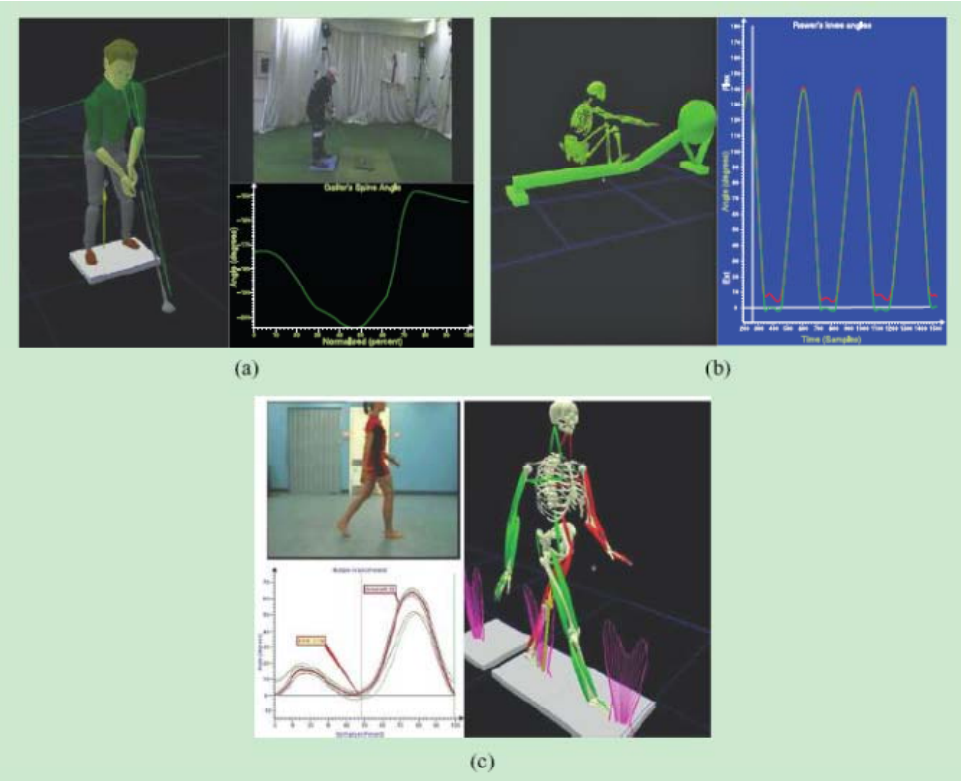
Search strategy

- Based on the previous frames, we can calculate a PDF of the Haar features of the object.
- In the current frame, for each sample in the search range, we extract its Haar features f_i and compute its probability $P(f_i)$; finally, the sample with the maximal probability is selected as the tracking result.



Applications

- Medical and sports



Example of video tracking for medical and sport analysis applications.

Motion capture is used to analyze the performance of (a) golfer, (b) rower, and (c) the gait of a patient.

Applications

- Surveillance

Examples of video tracking for intelligent retail applications.

Applications

- Robotics and unmanned vehicles

Example of object tracking from an Unmanned Aerial Vehicle.