CSCM77
Motion & Tracking

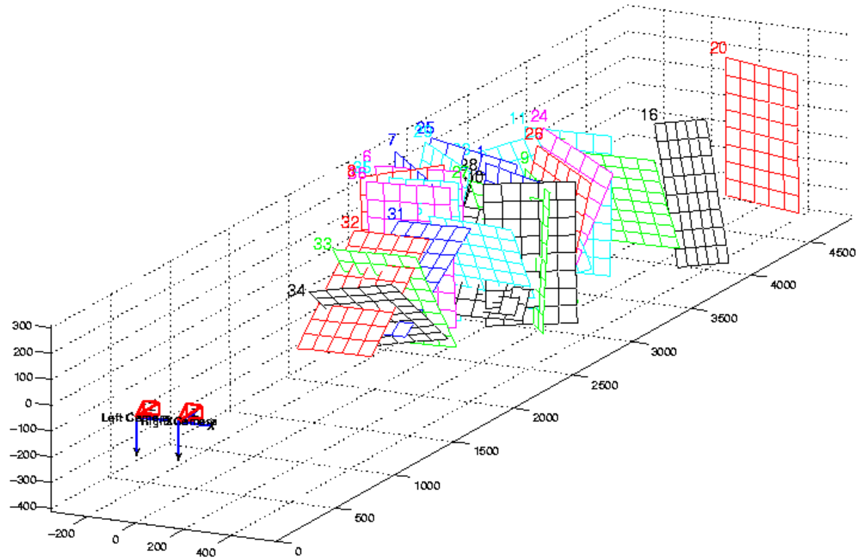
Prof. Xianghua Xie

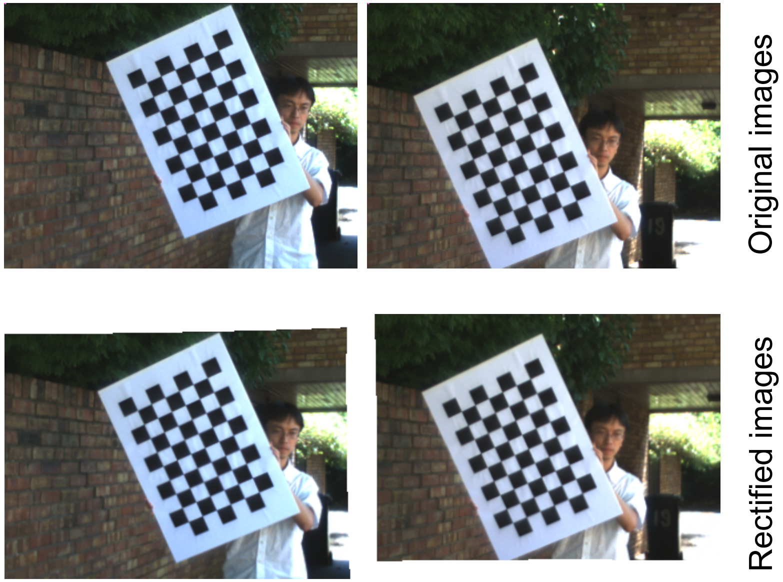
x.xie@swansea.ac.uk

http://csvision.swan.ac.uk

An example of depth estimation from stereo

- A quick example of depth estimation from stereo cameras
  - A narrow baseline (~16cm) setup
  - Real time performance
  - Images are rectified based on calibration
  - Matching is based on intensity values
  - Ordering and several constraints are used
  - Dynamic programming based optimisation





An example of depth estimation from stereo

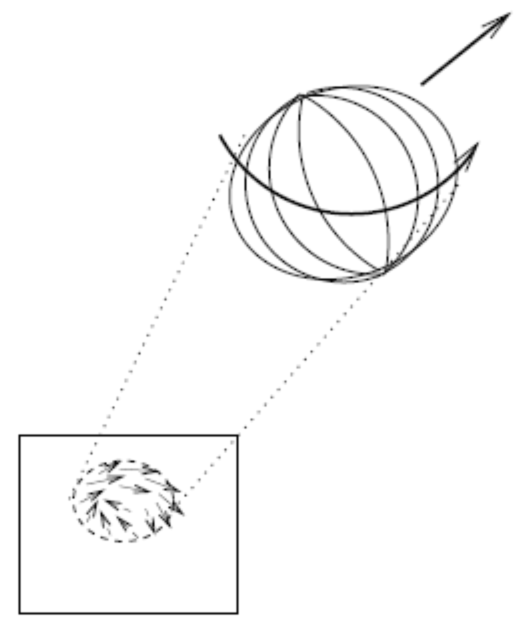
- A narrow baseline (~16cm) setup
- Real time performance
- Images are rectified based on calibration
- Matching is based on intensity values
- Ordering and several constraints are used
- Dynamic programming optimisation

Motion Analysis

- Video
  - Image sequence, snapshots of scene (frames), taken at regular intervals, e.g. 25 fps or 30 fps.

- Intensity variation between frames reflects relative motion of camera and objects in scene

- Motion analysis is a topic of estimating
  - Projected 2D motion field
  - 3D motion of objects
  - 2D and 3D trajectory (tracking)
  - …



Motion vs Stereo

- Motion analysis in many ways similar to stereo problem
  - From two (or more) views of scene, taken at different times or from different viewpoints, **reconstruct** after finding **correspondences**.

- However, there are important differences
  - Temporal sampling rate used to capture image sequence is usually high, which means the disparities between frames are small
  - In motion problems, disparities are not caused by a single 3D rigid transformation as in stereo;
    - Generally, both camera and objects move with distinct and possibly non-rigid motions
  - Usually, motion problems are considered harder

Apparent and True Motion

- In simple terms
  - Apparent motion is "perceived" motion field based on intensity variation across frames, also known as **optical flow**
  - True motion is projected actual motion field

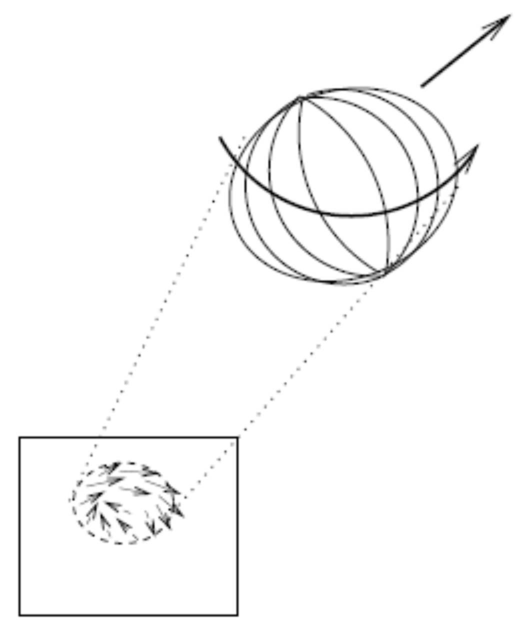
- Motion analysis involved here is to estimate 2D motion field from image sequence
  - Based on spatial and temporal variations of image intensity
  - In other words using intensity variations to estimation true motion

- Relationship between variation of image intensity (apparent motion or optical flow) and true motion field is not straightforward

Apparent and True Motion

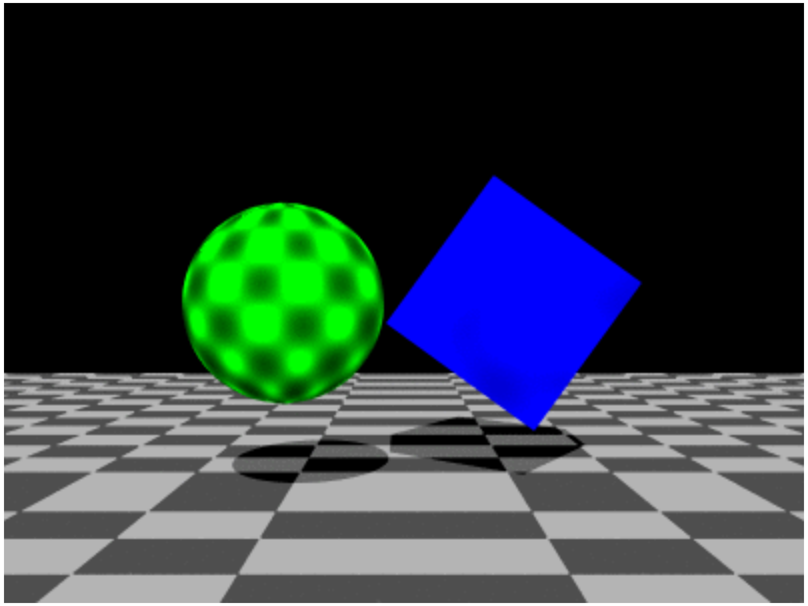
- Example difficult scenarios in establishing relationship between apparent and true motion fields
  - Zero apparent motion for non-zero motion field and vice versa
  - Sphere with constant colour surface rotating in diffuse lighting
  - Static scene and moving light sources

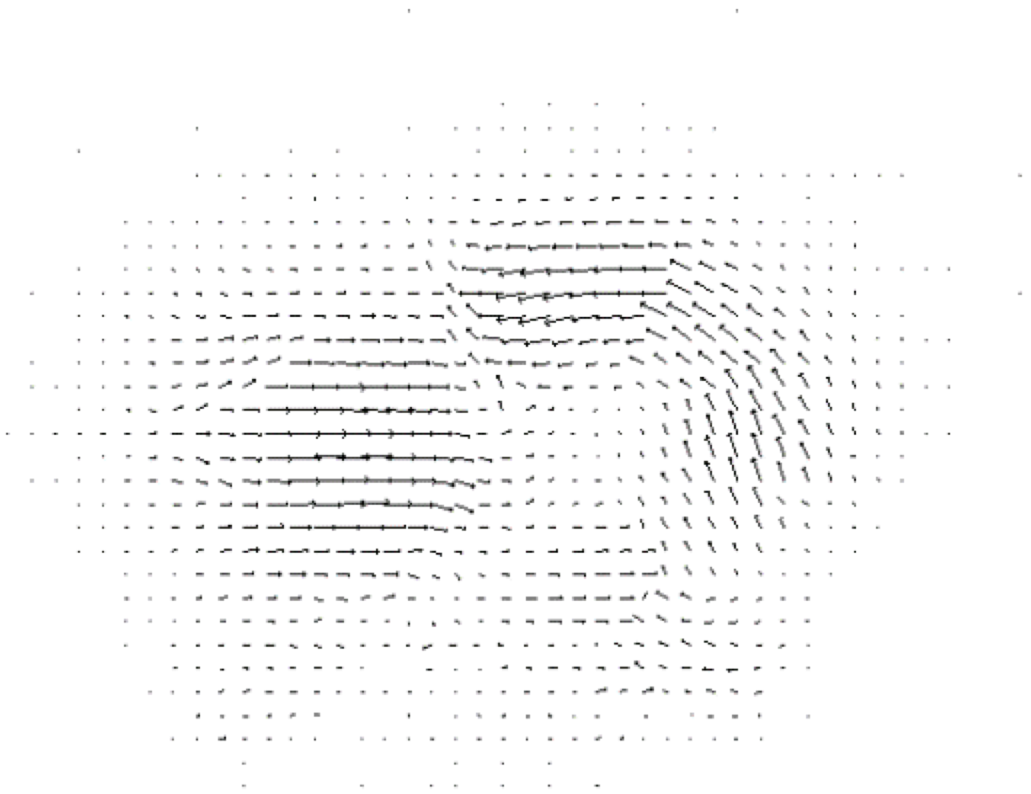
- In some cases, it is impossible to determine motion field without additional constraints



Apparent motion (Optical Flow)

- Example sequences
  - http://www.youtube.com/watch?v=ckVQrwYIjAs





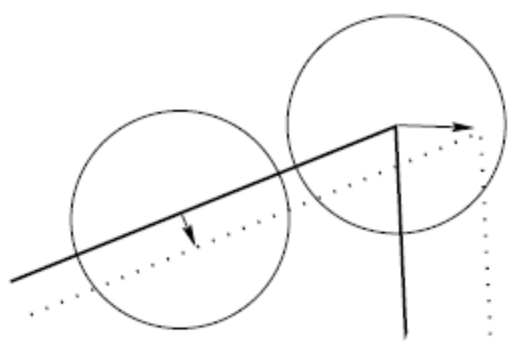
http://www.codeproject.com/KB/audio-video/Optical_Flow_Estimation.aspx

The Aperture Problem

- Limitation of optical flow estimation underlines a wider problem: the aperture problem

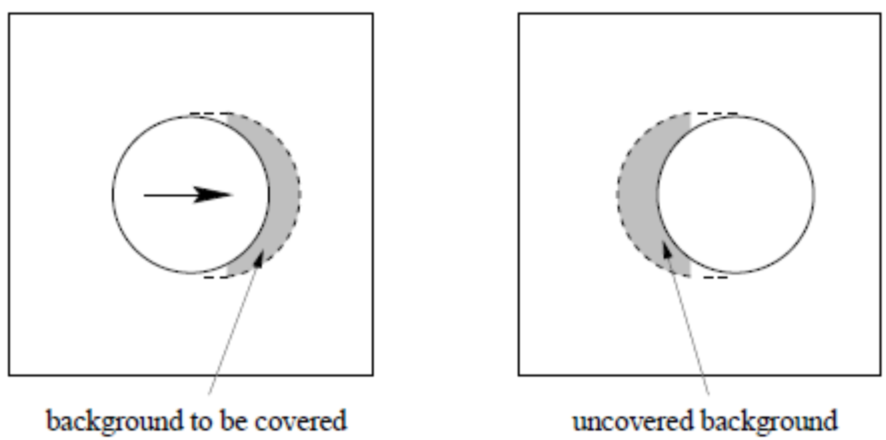
- Measurement window must contain sufficient spatial gradient variation in order to determine motion
  - E.g. corners and edges

- Requires integrating local motion information to obtain unambiguous estimation of true motion



The Occlusion Problem

- Objects moving results in covered and uncovered background
- No correspondences between frames in such areas
- Motion is undefined in such regions
  - Require additional constrains to infer the motion



Motion Analysis Example

- Distinguish independent motion from “moving” background due to camera ego-motion
- Based on reliable apparent motion analysis

Motion Tracking

- Tracking is a natural extension of motion estimation between adjacent frames

- Attempt to track the motion of elements over time

- Elements: objects, features, regions, points …

- Tracking motion over many frames, instead of adjacent frames
  - Improve performance towards ambiguities and uncertainties
  - Reduce noise interference (not just image noise but also motion noise)

Motion Tracking

- An example of tracking points and localising camera positions over time
  - Build up a map while performing tracking: simultaneous localisation and tracking (SLAM)

Videos: A. Davison

Motion Tracking

Key issues in designing motion tracking algorithm

- Select element for tracking
  - Decide what to track, e.g. features, regions and so on
- Select motion model
  - What assumption to use in trajectory estimation
- Estimate the trajectory parameters

Tracking Element

- Need to decide what image elements to track over time
- And how they are represented

- Features for tracking
  - Line segments, edge segments, corners, high entropy points …

- Regions for tracking
  - Homogeneous regions corresponding to object surfaces
  - E.g. track centroid or boundary of regions

- Object or 3D surface for tracking
  - More general but much more complex

Trajectory Model

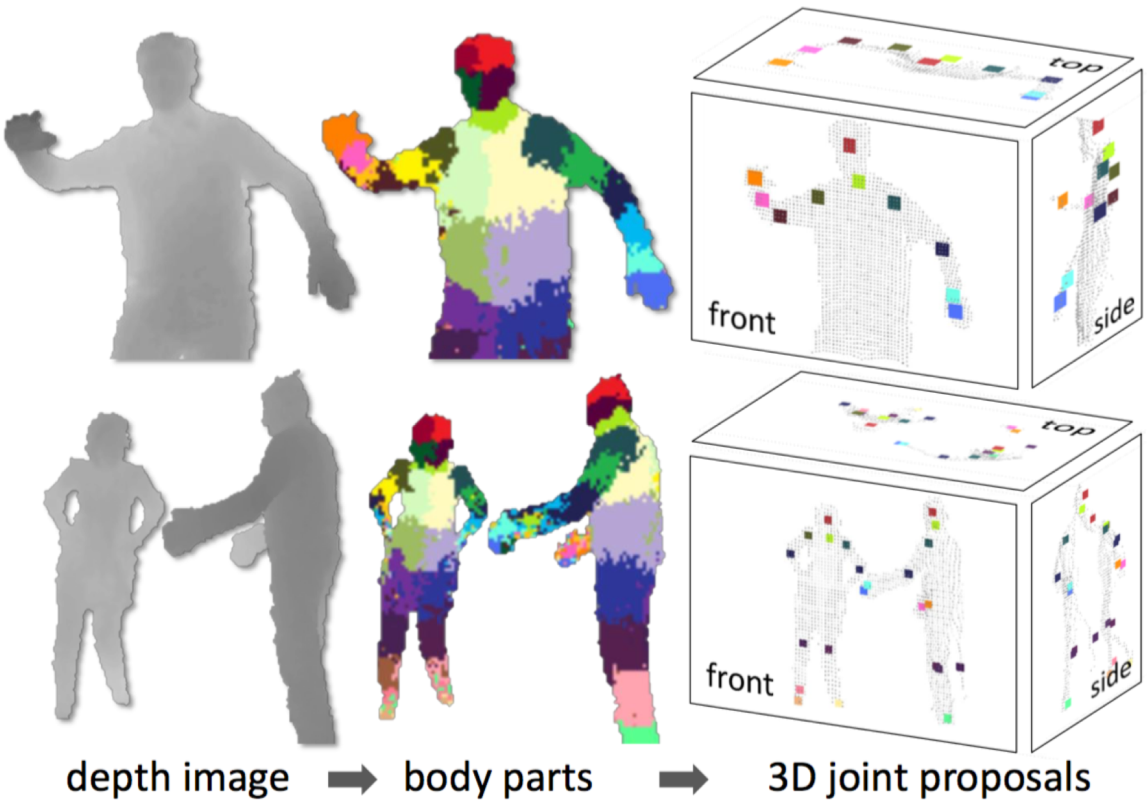
- How do we believe the motion should change over time?
  - Smooth variation?
  - Constant velocity?
  - Constant acceleration?
  - …
- Constant acceleration model:

$$
\begin{aligned}
x_k &= x_{k-1} + v_{k-1}\Delta t + a_{k-1}(\Delta t)^2/2 \\
v_k &= v_{k-1} + a_{k-1}\Delta t \\
a_k &= a_{k-1}
\end{aligned}
$$

- Challenges, e.g.
  - Variation of element attributes, e.g. shape
  - Motion dicontinuities

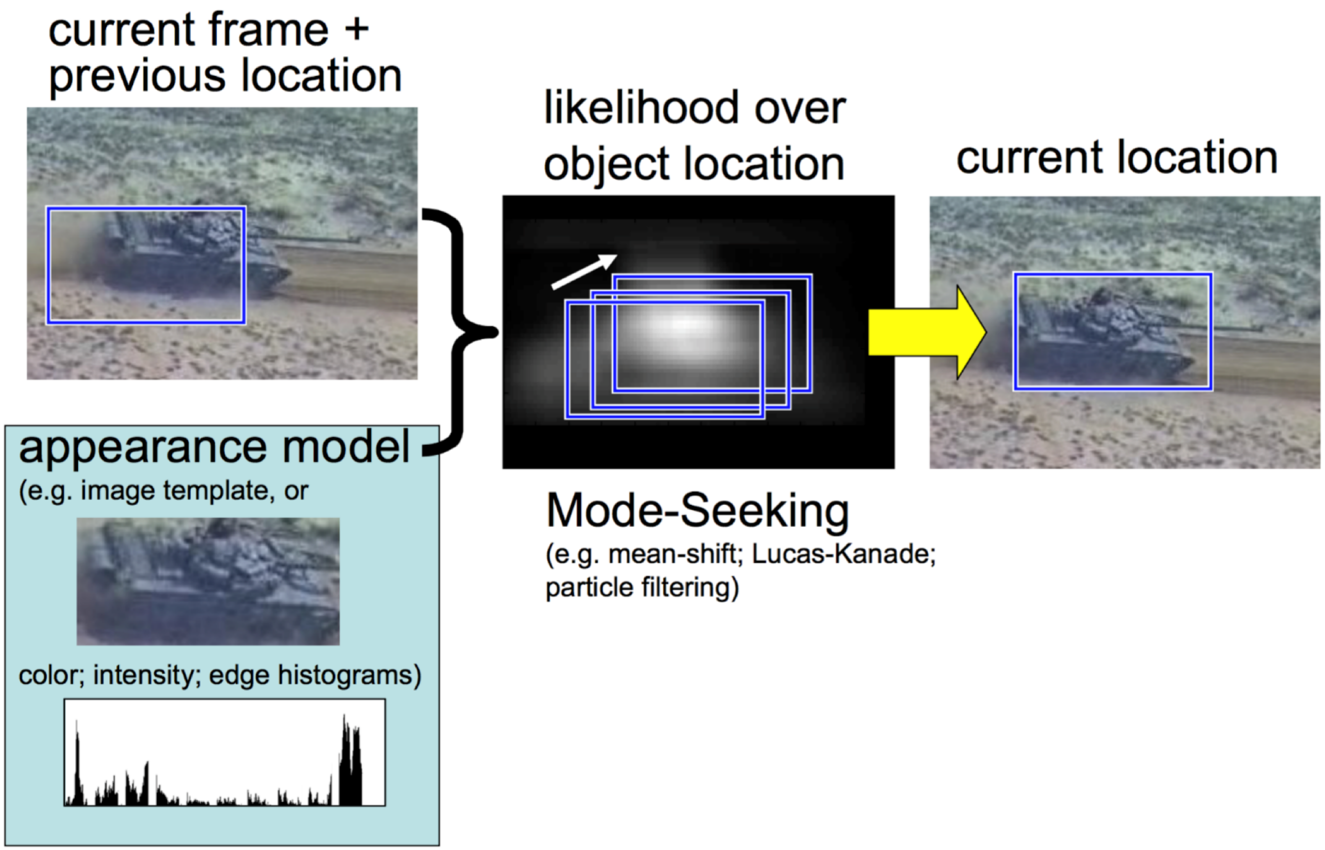
Tracking using Mean Shift

- In Kinect, Mean Shift algorithm to estimate the joint positions from classified body parts
- This "mode-seeking" technique can be used for tracking as well



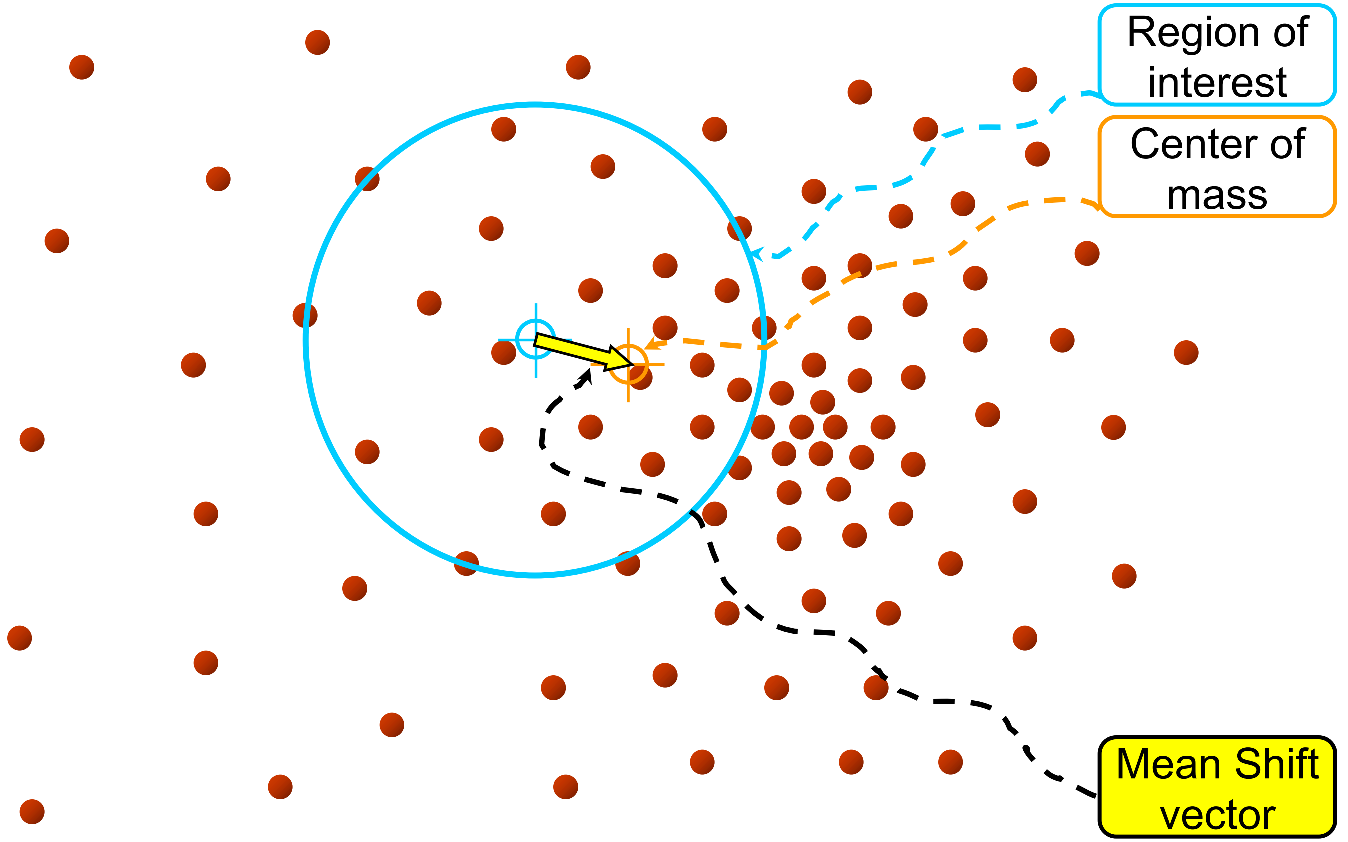
Appearance Based Tracking

- Mean-shift algorithm is also an efficient approach to tracking objects



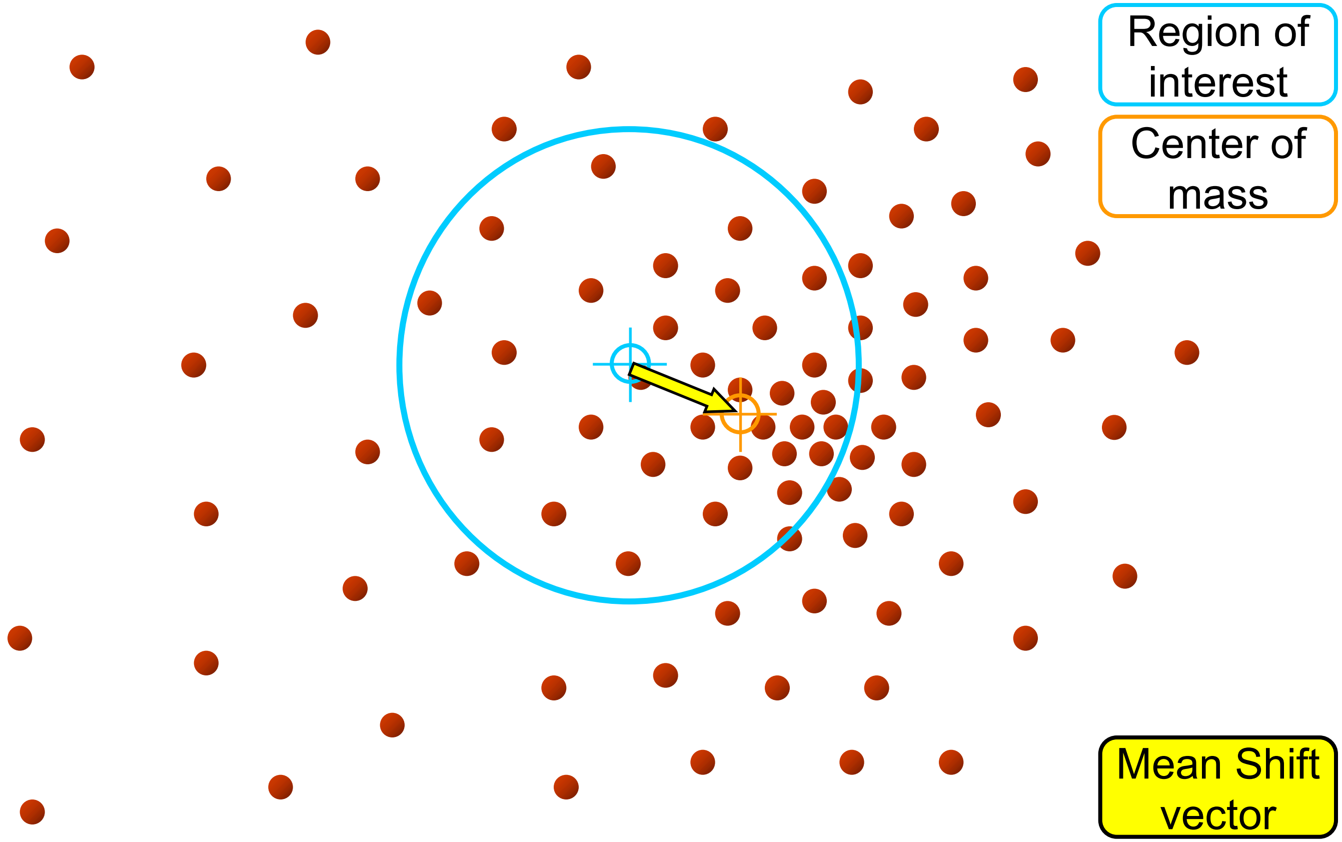
Mean Shift

- Objective: search the densest region for a given search window size
- Finding modes in a set of data samples



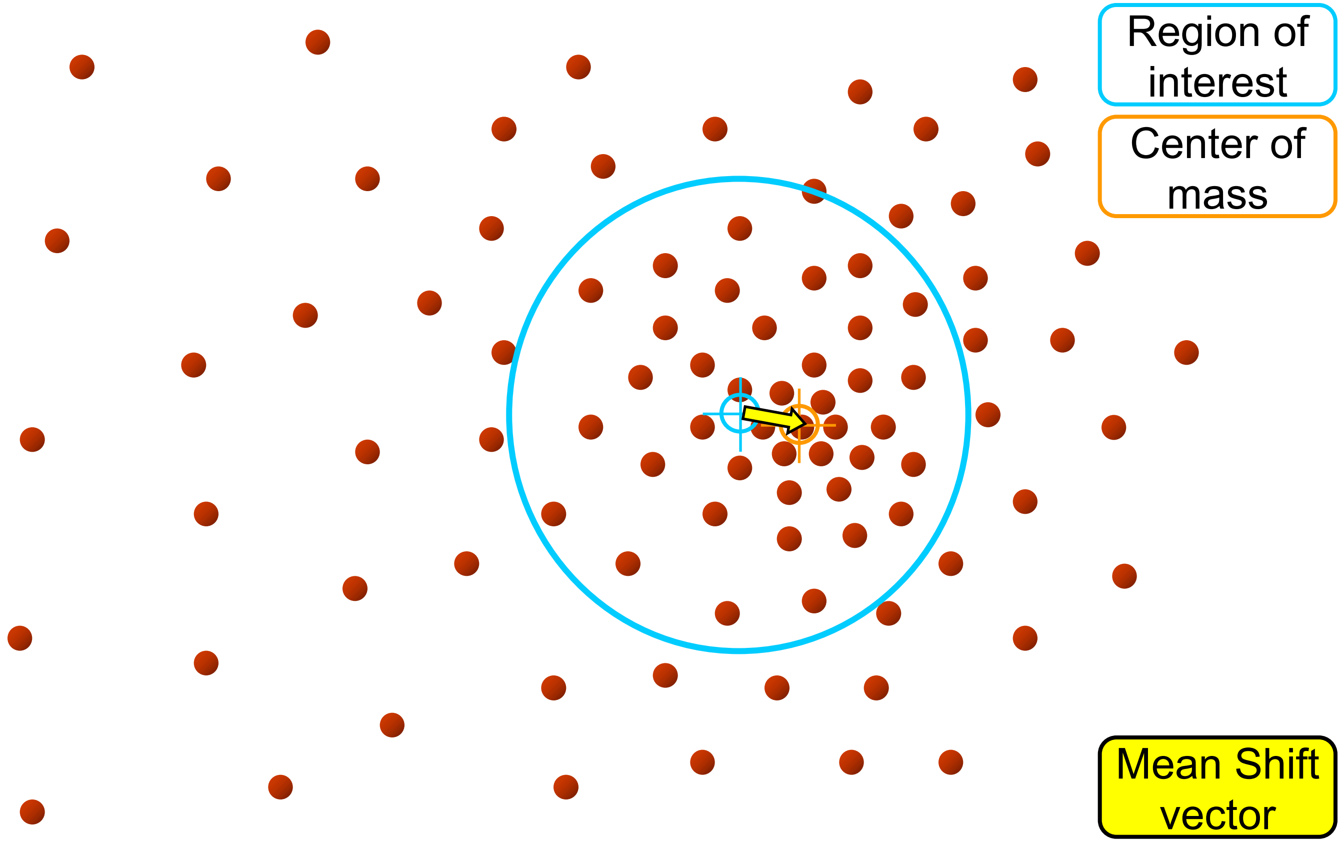
Mean Shift

- Objective: search the densest region for a given search window size
- Finding modes in a set of data samples



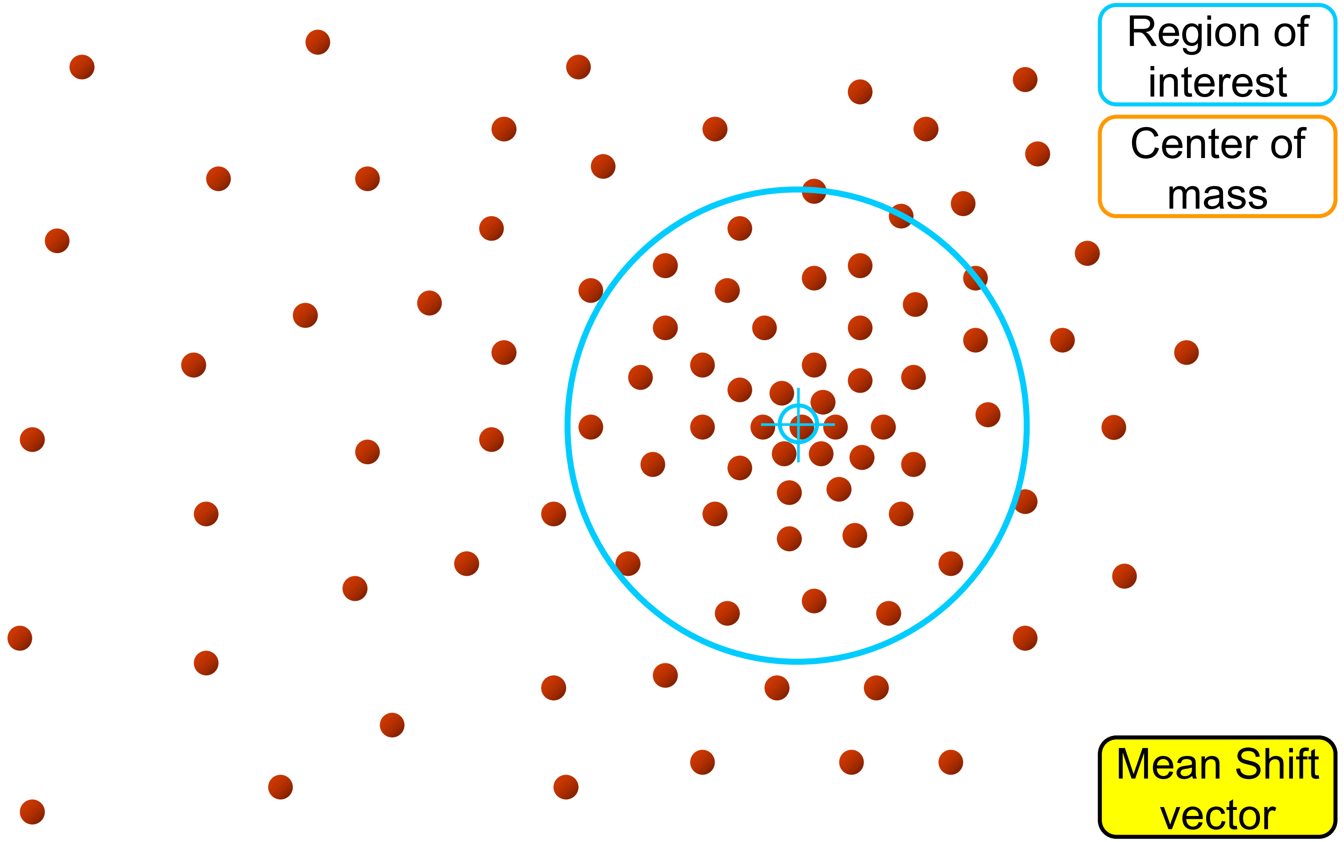
Mean Shift

- Objective: search the densest region for a given search window size
- Finding modes in a set of data samples



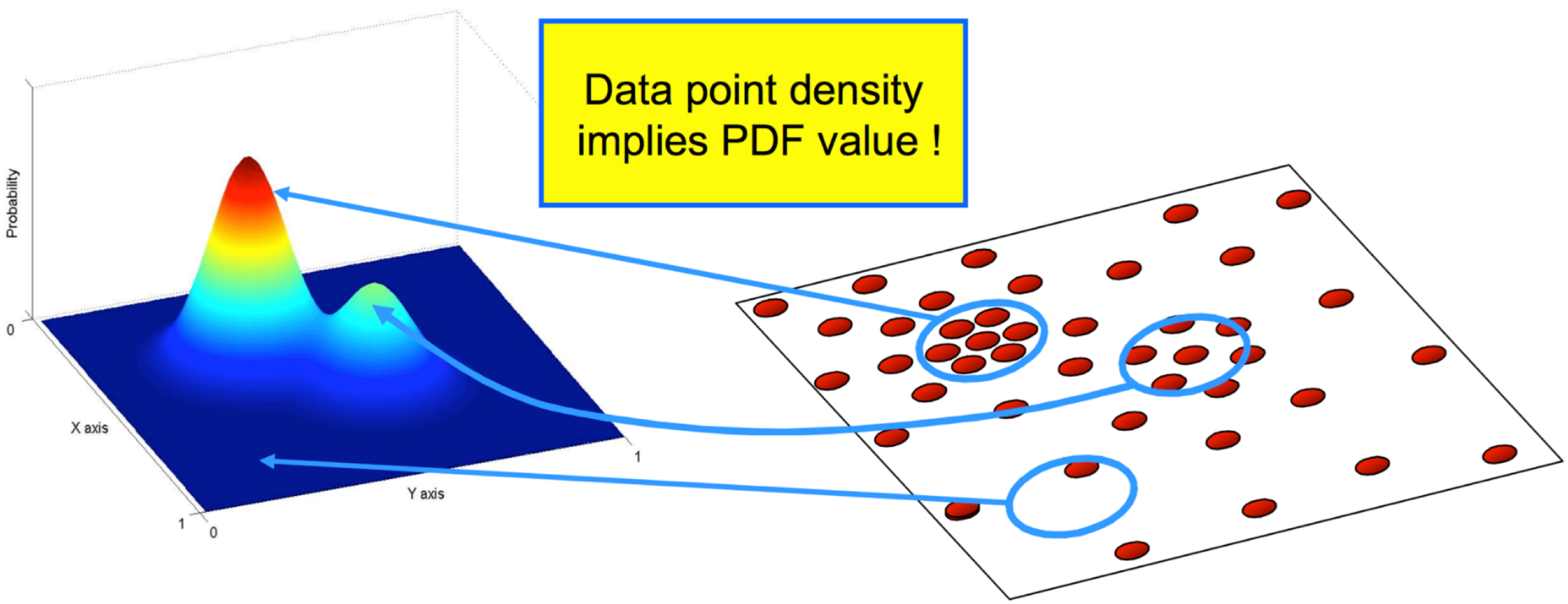
Mean Shift

- Objective: search the densest region for a given search window size
- Finding modes in a set of data samples



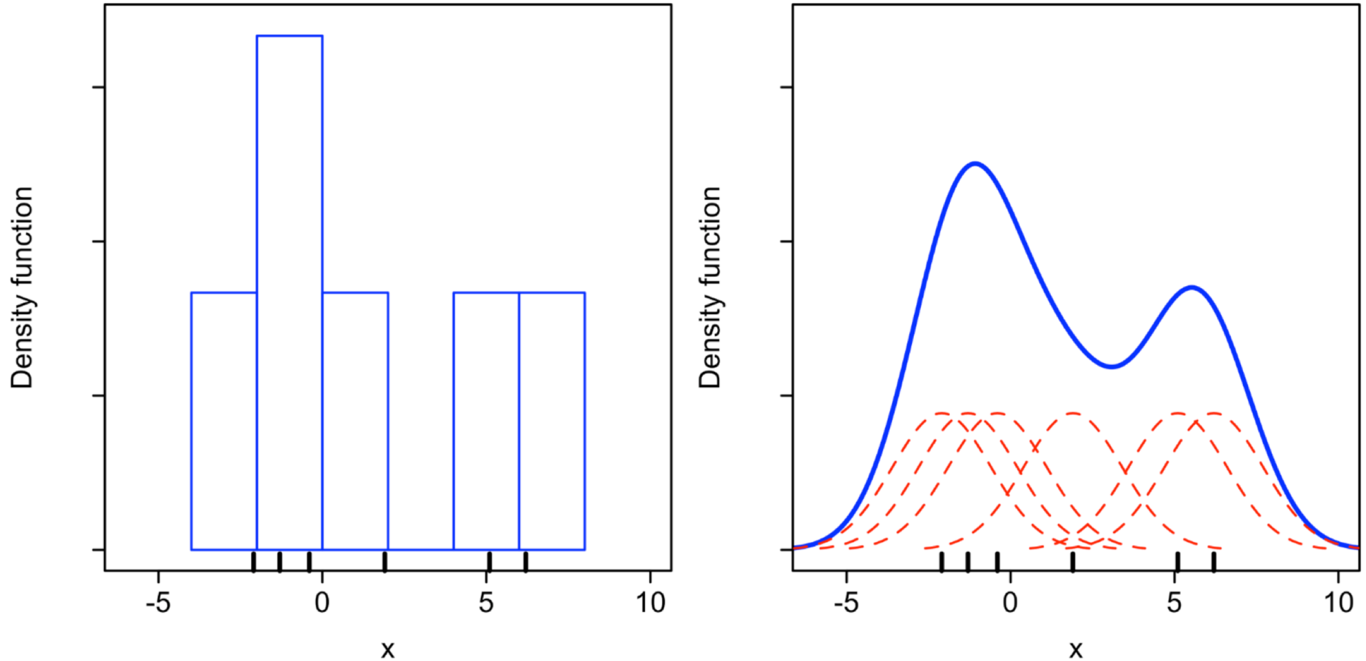
Mean Shift

- Assumption: the data points are sampled from an underlying probability distribution function (PDF)



Mean Shift

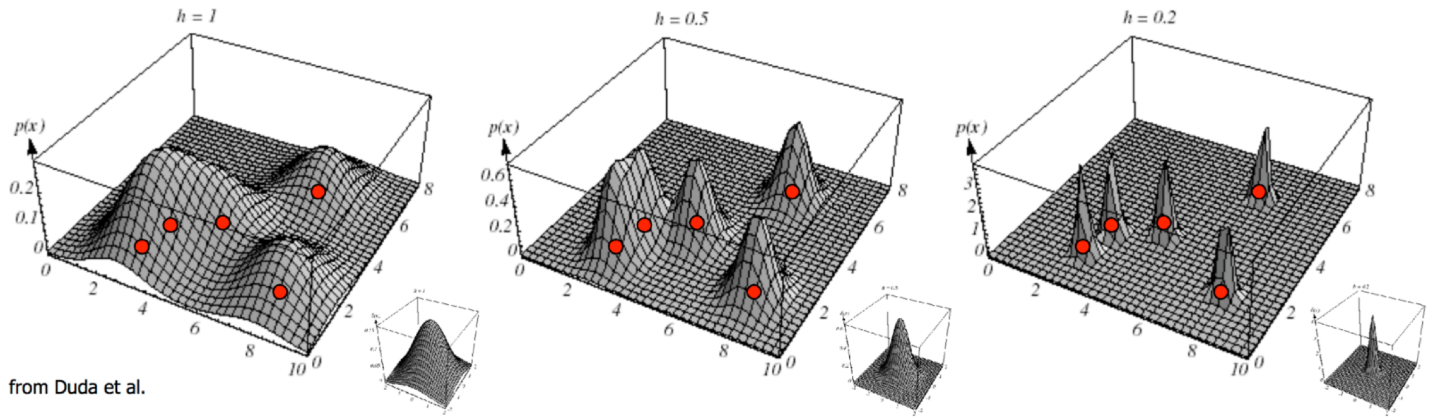
- Approximate PDF by estimating local density of points using kernel functions, such as Gaussian
  - Gaussian has a scale parameter: standard deviation (sigma)
  - Convolve the data points with the kernel function



Comparison of the histogram (left) and kernel density estimate (right) constructed using the same data. The 6 individual kernels are the red dashed curves, the kernel density estimate the blue curves. The data points are the rug plot on the horizontal axis.

Mean Shift

- Challenges:
  - Sensitive to bandwidth (scale of the kernel function)



  - May trapped to local minima; requires periodic perturbation
    - Periodically perturb the mean shift vector when seeking the modes