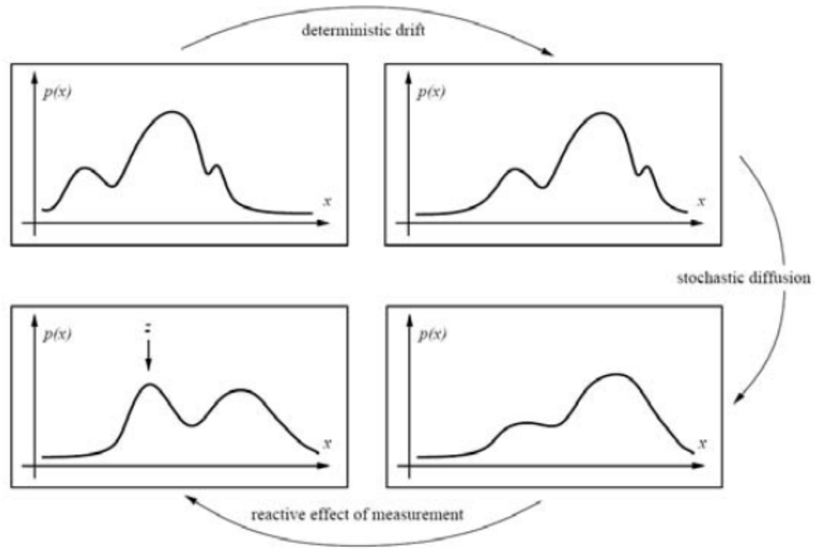


COMP 9517 Computer Vision

Tracking

Motion Tracking

- Tracking is the problem of generating an inference about the motion of an object given a sequence of images
  - What do we infer about an object's future position from a sequence of measurements?
- Applications
  - Motion capture
    - Control a cartoon
    - Modify the motion record to obtain slightly different motions
  - Recognition from motion
    - Determine the identity of the object
    - Tell what it is doing
  - Surveillance
    - Monitor activities and give a warning if it detects a problem case
  - Targeting
    - Decide what to shoot and hitting it

Motion Tracking

- When moving points are not tagged with unique texture or colour information, the characteristics of the motion itself must be used to collect points into trajectories
- Assumption:
  - The location of an object changes smoothly over time
  - The velocity (speed and direction) of an object changes smoothly over time
  - An object can be at only one location in space at a give time
  - Two objects cannot occupy the same location at the same time

Tracking with Dynamic Models

- Tracking can be considered as the problem of generating an (probabilistic) inference about the motion of an object given a sequence of images
- Tracking is properly thought of as an probabilistic inference problem
  - The moving object has internal state, which is measured at each frame
  - Measurements are combined to estimate the state of the object

Tracking with Dynamic Models

- State: the representation of an object at time (frame) t
  - Position
  - Transformation parameters
  - …
- Measurement: the observation
  - Colour
  - Texture
  - …

Tracking with Dynamic Models

- Given
  - X_i: the state of the object at the i_{th} frame
  - Y_i: the measurement obtained in the i_{th} frame
- There are three main problems:
  - Prediction: predict the state for the i_{th} frame by having seen a set of measurements y_0, y_1, …, y_{i-1}

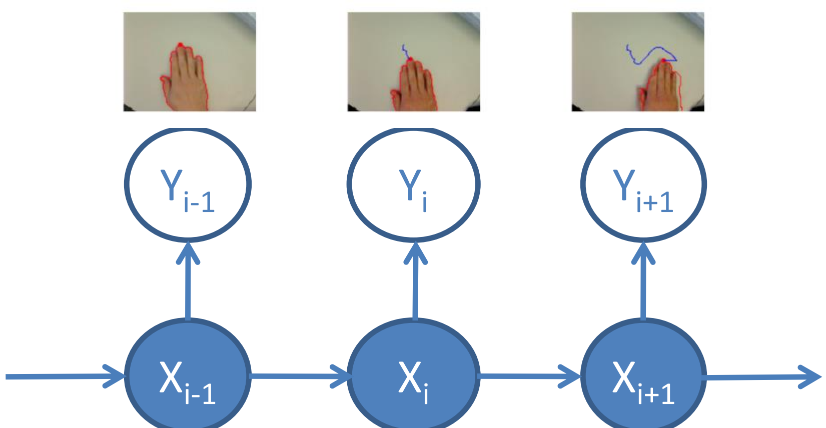
$$P(X_i \mid Y_0 = y_0, Y_1 = y_1, \ldots, Y_{i-1} = y_{i-1})$$

  - Data association: select the measurements that are related to the object's state
  - Correction: correct the predicted state by obtained relevant measurements y_i

$$P(X_i \mid Y_0 = y_0, Y_1 = y_1, \ldots, Y_i = y_i)$$

Tracking with Dynamic Models

- Independence Assumptions (Markov Assumption)
  - Only the immediate past matters
    - $p(x_i|x_{i-1},\ldots,x_1) = p(x_i|x_{i-1})$
  - Measurements depend only on the current state
    - $p(y_i|y_i, y_{i-1},\ldots,y_1,x_i) = \; p(y_i|x_i)$



  - These assumptions mean that a tracking problem has the structure of inference on a hidden Markov model

Tracking with Dynamic Models

- Tracking as Inference
  - Prediction

$$P(X_i \mid y_0, y_1, \ldots, y_{i-1}) = \int P(X_i, X_{i-1} \mid y_0, y_1, \ldots, y_{i-1}) dX_{i-1}$$

$$= \int P(X_i \mid X_{i-1}, \cancel{y_0, y_1, \ldots, y_{i-1}}) P(X_{i-1} \mid y_0, y_1, \ldots, y_{i-1}) dX_{i-1}$$

$$= \int P(X_i \mid X_{i-1}) P(X_{i-1} \mid y_0, y_1, \ldots, y_{i-1}) dX_{i-1}$$

  - Correction

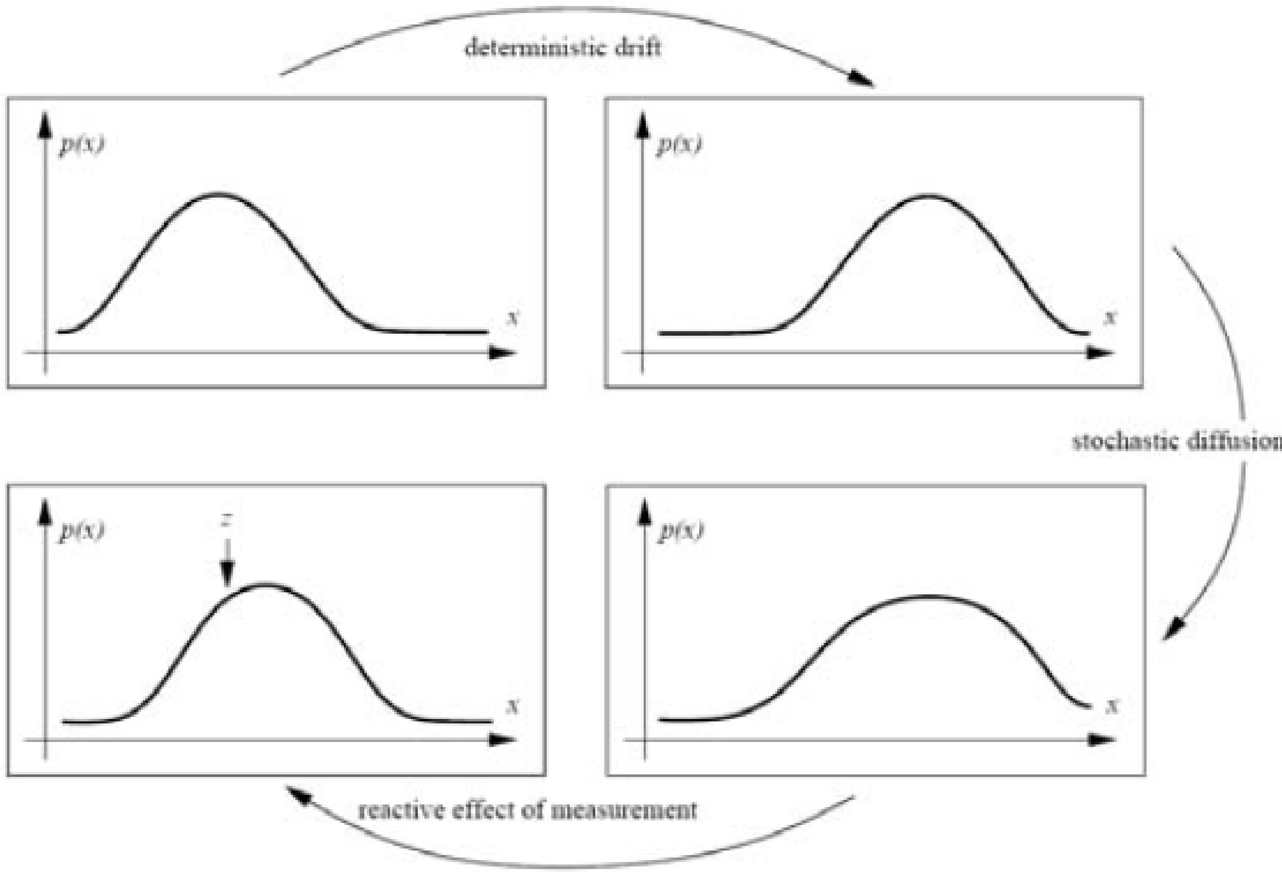
$$P(X_i \mid y_0, y_1, \ldots, y_i) = \frac{P(X_i, y_0, y_1, \ldots, y_i)}{P(y_0, y_1, \ldots, y_i)}$$

$$= \frac{P(y_i \mid X_i, \cancel{y_0, y_1, \ldots, y_{i-1}}) P(X_i \mid y_0, y_1, \ldots, y_{i-1}) P(y_0, y_1, \ldots, y_{i-1})}{P(y_0, y_1, \ldots, y_i)}$$

$$= P(y_i \mid X_i) P(X_i \mid y_0, y_1, \ldots, y_{i-1}) \frac{P(y_0, y_1, \ldots, y_{i-1})}{P(y_0, y_1, \ldots, y_i)}$$

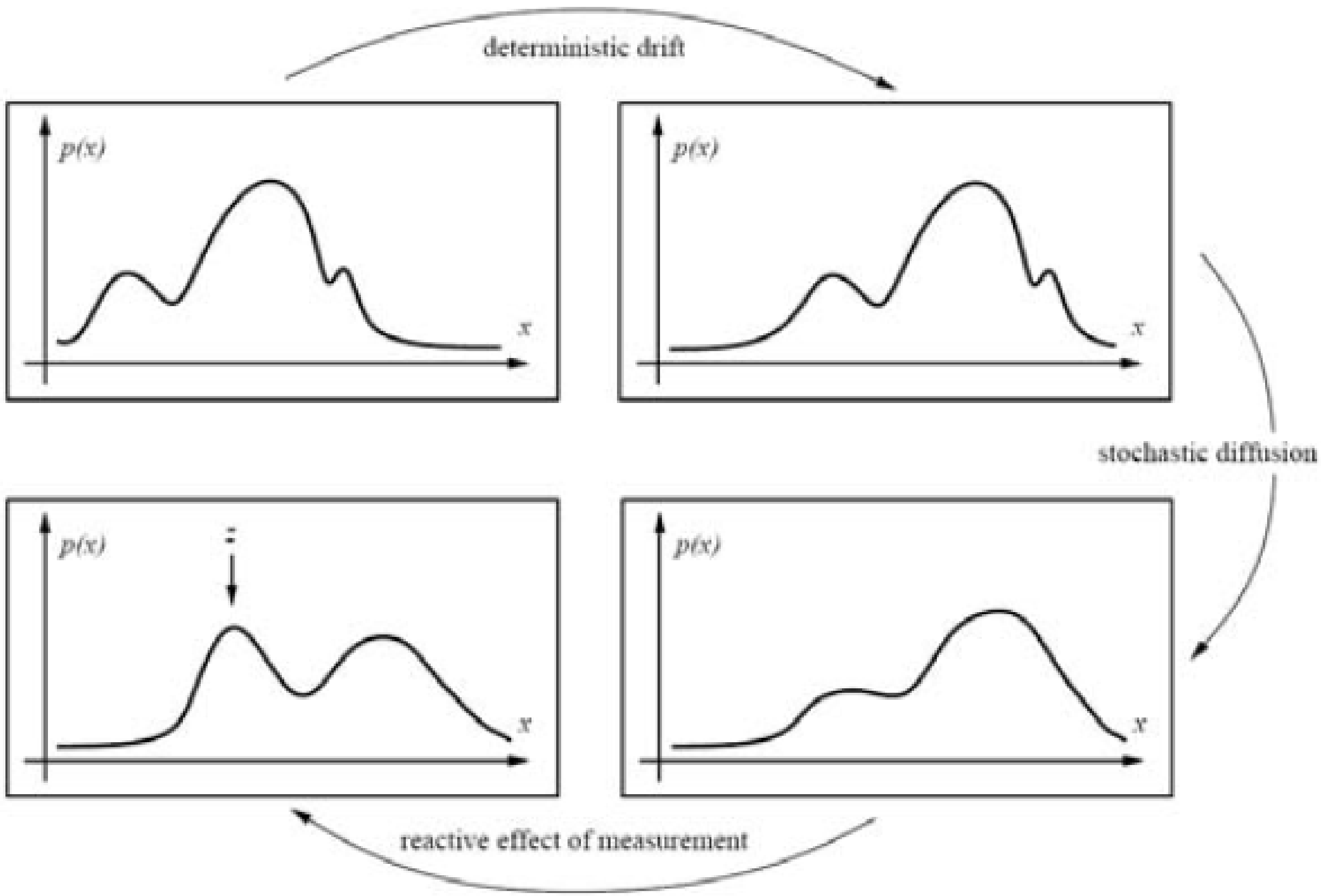
$$= \frac{P(y_i \mid X_i) P(X_i \mid y_0, y_1, \ldots, y_{i-1})}{\int P(y_i \mid X_i) P(X_i \mid y_0, y_1, \ldots, y_{i-1}) dX_i}$$

Probability Density Propagation



(a)

Probability Density Propagation



Tracking with Dynamic Models

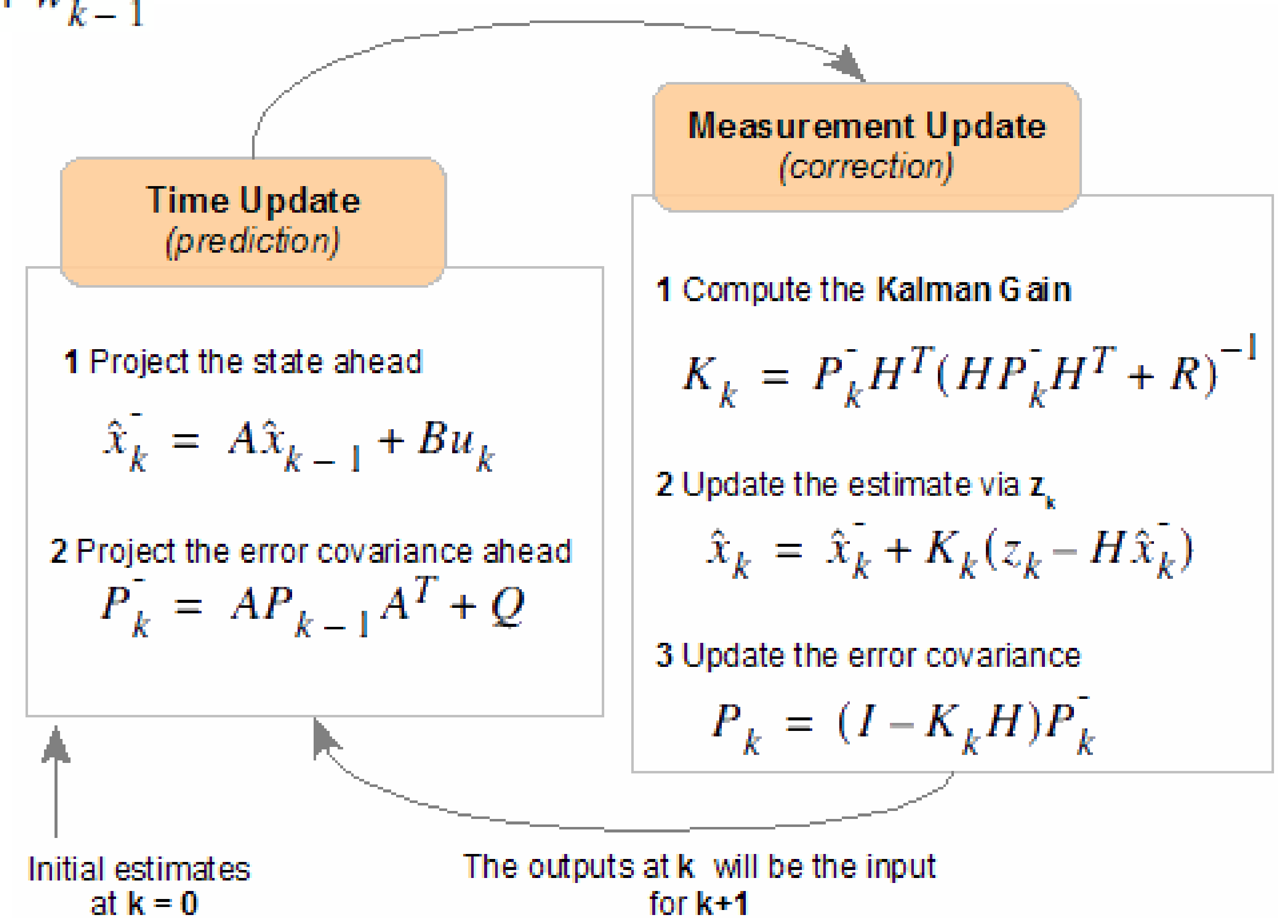
- Linear Dynamic Models
  - A random probability distribution with mean and covariance
    - $x \sim N(x; m, \Sigma)$
  - The state is advanced by multiplying it by some known matrix and then adding a normal random variable
    - $x_i \sim N(x_i; Ax_{i-1}, P_i)$, i.e. $x_i = Ax_{i-1} + w_i$
  - The measurement is obtained by multiplying the state by some matrix and then adding a normal random variable of zero mean and known covariance
    - $y_i \sim N(y_i; Hx_i, R_i)$, i.e. $y_i = Hx_i + v_i$

Tracking with Dynamic Models

- Kalman Filtering

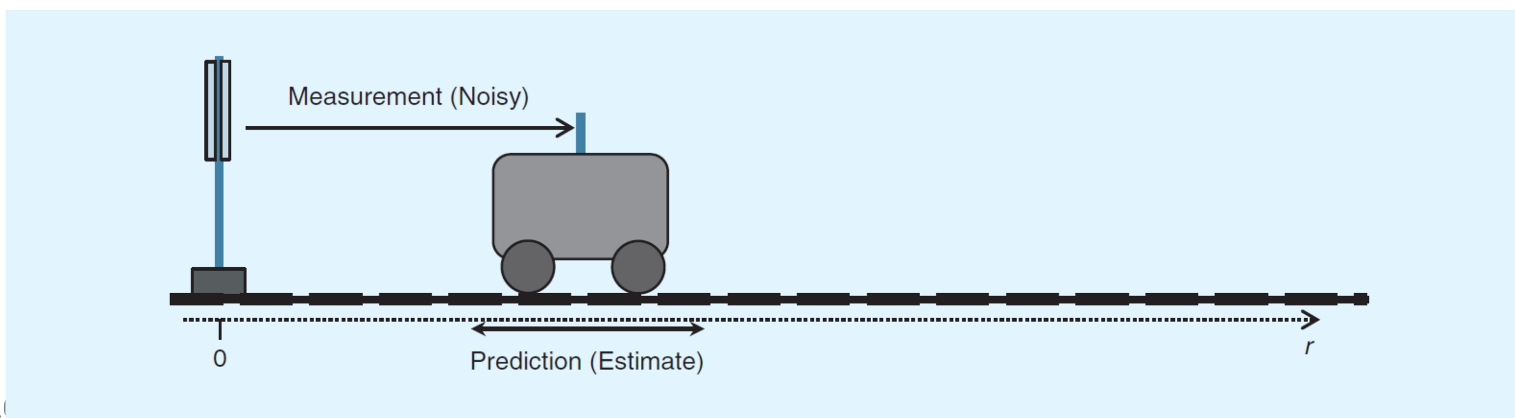
$$x_k = Ax_{k-1} + Bu_k + w_{k-1}$$

$$z_k = Hx_k + v_k$$



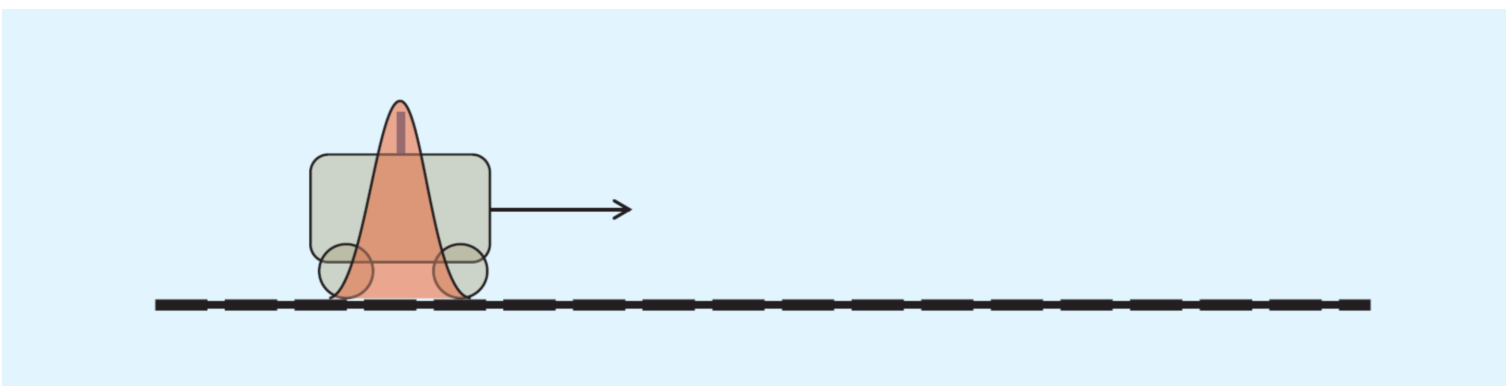
Tracking with Dynamic Models

- Kalman filtering – 1D example
  - $x_i = x_{i-1} + u + w_i$
    - $\Delta_i = x_i - x_{i-1} = u + w_i$
    - $w_i \sim N(w_i; 0, \sigma_w^2)$
  - $y_i = x_i + v_i$
    - $v_i \sim N(v_i; 0, \sigma_v^2)$



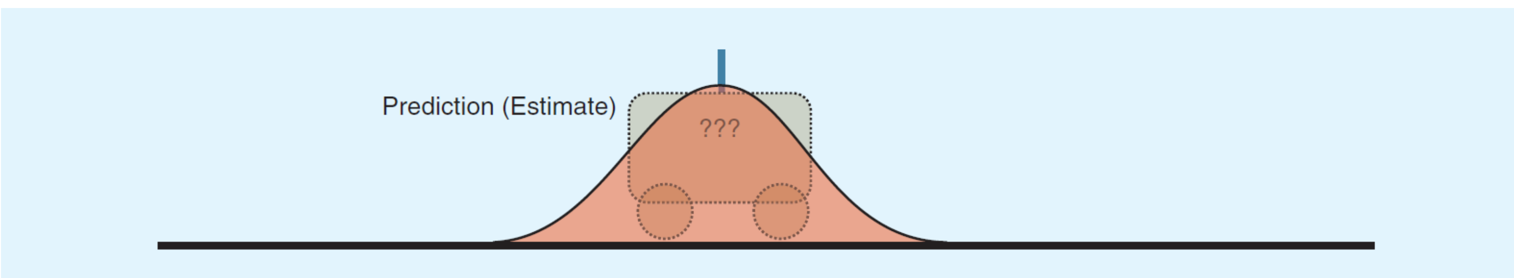
Tracking with Dynamic Models

- Kalman filtering - problem
  - Estimate $x_i | y_1, \dots, y_i$
    - $x_{i-1} | y_1, \dots, y_{i-1} \sim N(x_{i-1}; m_{i-1}, \sigma_{i-1}^2)$
    - $x_i = x_{i-1} + u + w_i$
    - $y_i = x_i + v_i$



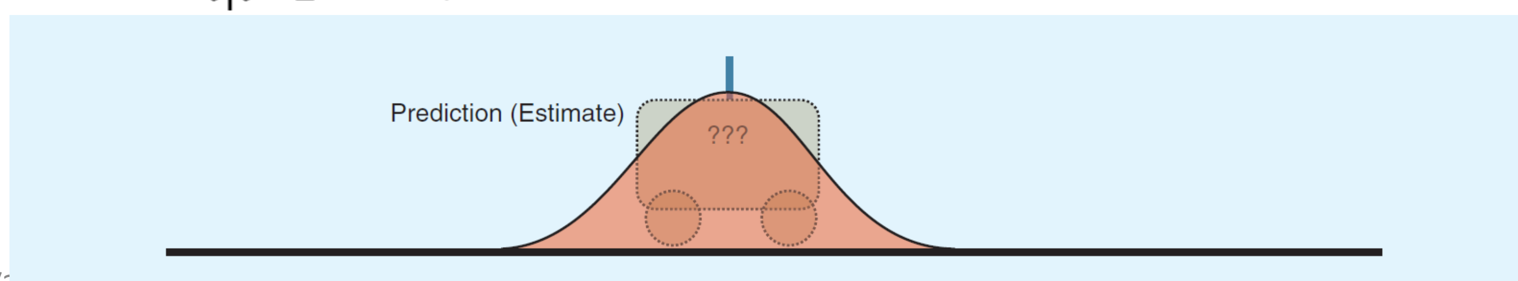
Tracking with Dynamic Models

- Kalman filtering - prediction
  - Estimate $x_i | y_1, \ldots, y_{i-1}$
    - $x_{i-1} | y_1, \ldots, y_{i-1} \sim N(x_{i-1}; m_{i-1}, \sigma_{i-1}^2)$
    - $x_i = x_{i-1} + u + w_i$
  - $p(x_i | y_1, \ldots, y_{i-1}) = \int p(x_i | x_{i-1}) p(x_{i-1} | y_1, \ldots, y_{i-1})\, dx_{i-1}$



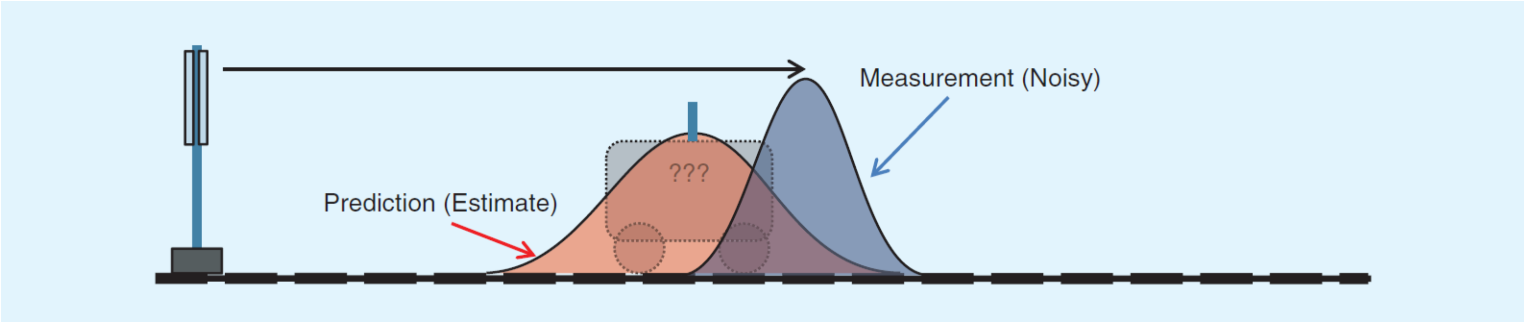
Tracking with Dynamic Models

- Kalman filtering - prediction
  - $\int p(x_i|x_{i-1})p(x_{i-1}|y_1, \dots, y_{i-1})\, dx_{i-1}$
    - $p(x_i|x_{i-1}) = N(x_i; x_{i-1} + u, \sigma_w^2)$
    - $p(x_{i-1}|y_1, \dots, y_{i-1}) = N(x_{i-1}; m_{i-1}, \sigma_{i-1}^2)$
  - $x_i|y_1, \dots, y_{i-1} \sim N(x_i; m_{i|i-1}, \sigma_{i|i-1}^2)$
    - $m_{i|i-1} = m_{i-1} + u$
    - $\sigma_{i|i-1}^2 = \sigma_{i-1}^2 + \sigma_w^2$



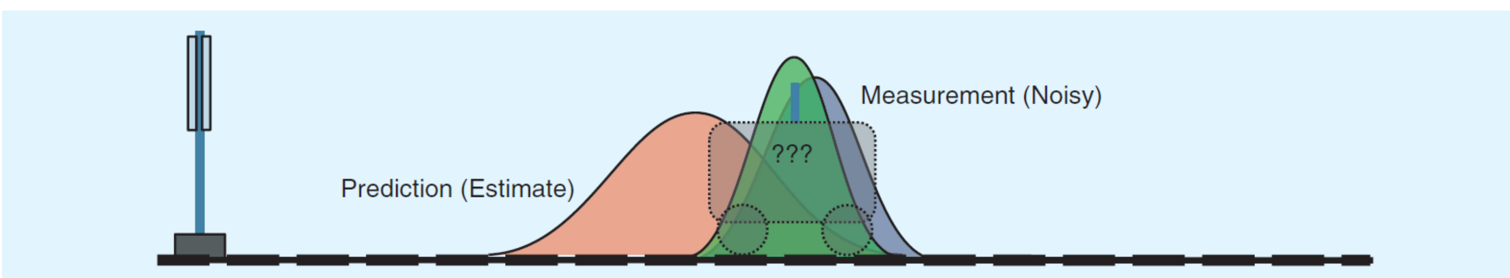
Tracking with Dynamic Models

- Kalman filtering-correction
  - Estimate $x_i | y_1, \ldots, y_i$
    - $x_i | y_1, \ldots, y_{i-1} \sim N(x_i; m_{i|i-1}, \sigma^2_{i|i-1})$
    - $y_i = x_i + v_i$
  - $p(x_i | y_1, \ldots, y_i) \propto p(y_i | x_i) p(x_i | y_1, \ldots, y_{i-1})$
    - $p(y_i | x_i) = N(x_i; y_i, \sigma_v^2)$
    - $p(x_i | y_1, \ldots, y_{i-1}) = N(x_i; m_{i|i-1}, \sigma^2_{i|i-1})$



Tracking with Dynamic Models

- Kalman filtering-correction
  - $x_i | y_1, \ldots, y_i \sim N(x_i; m_i, \sigma_i^2)$
    - $m_i = \frac{\sigma_v^2}{\sigma_{i|i-1}^2 + \sigma_v^2} m_{i|i-1} + \frac{\sigma_{i|i-1}^2}{\sigma_{i|i-1}^2 + \sigma_v^2} y_i$
    - $\sigma_i^2 = \frac{\sigma_{i|i-1}^2 \sigma_w^2}{\sigma_{i|i-1}^2 + \sigma_w^2}$

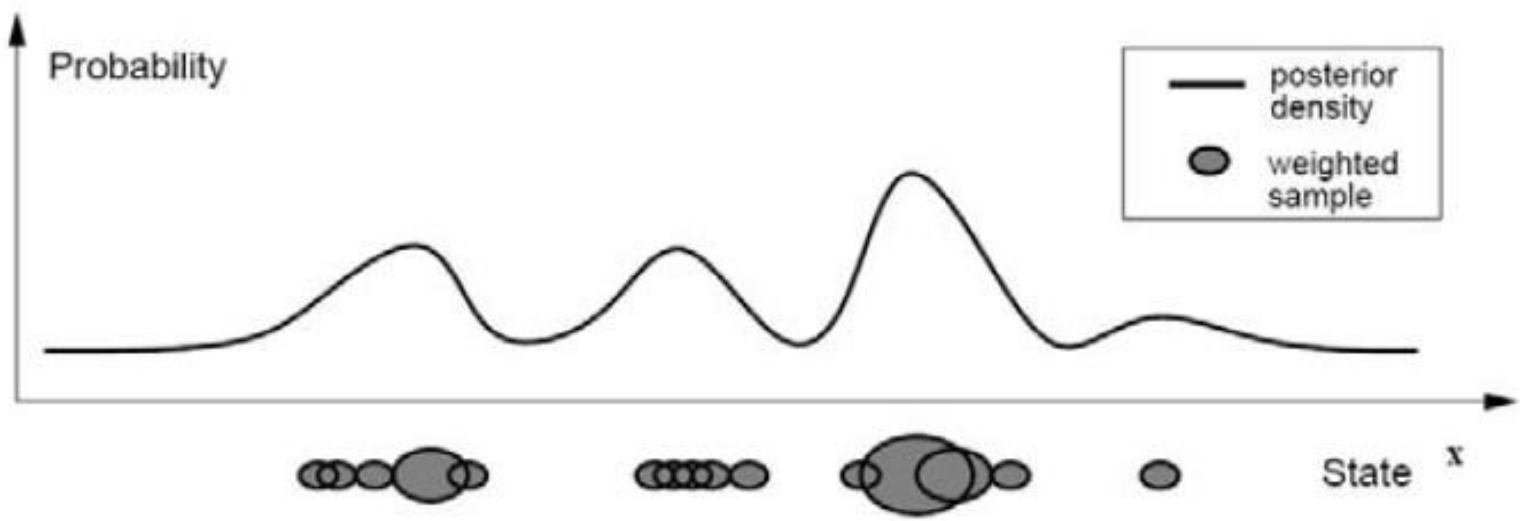


Tracking with Dynamic Models

- Particle Filtering
  - A non-linear dynamic model
  - Also known variously as :
    - Sequential Monte Carlo method,
    - bootstrap filtering,
    - the condensation algorithm,
    - interacting particle approximations and
    - survival of the fittest

Tracking with Dynamic Models

- Particle Filtering
  - The conditional state density at time *t* is represented by a set of sampling particles with weight (sampling probability).
  - The weight define the importance of a sample (observation frequency)

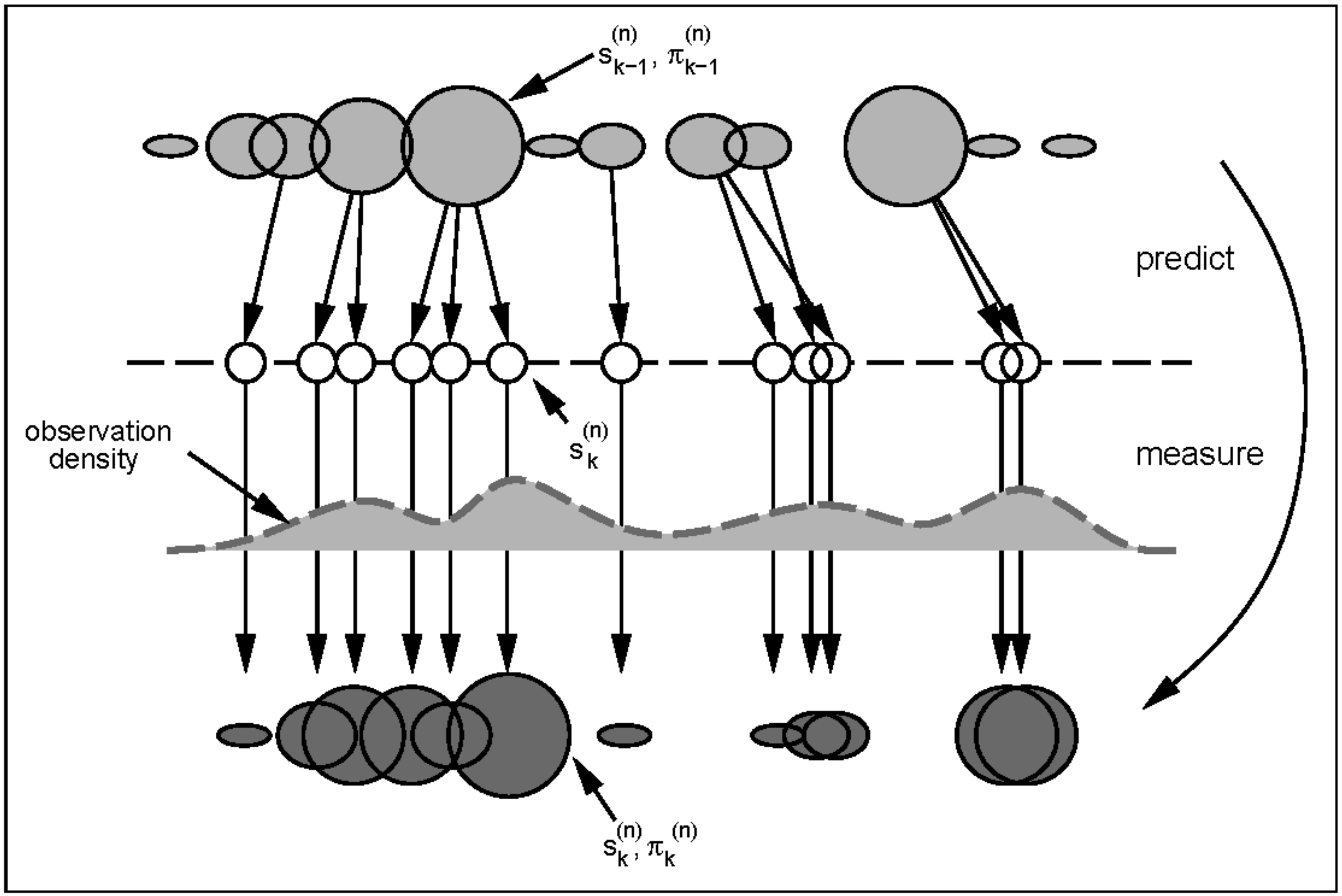


Tracking with Dynamic Models

- Particle Filtering
  - The common sampling scheme is importance sampling
    - Selection: select N random samples
    - Prediction: generate new samples with zero mean Gaussian Error and non-negative function
    - Correction: compute weights corresponding to the new samples using the measurement equation which can be modelled as a Gaussian density

Tracking with Dynamic Models

- Particle Filtering (A.Vlake and M.Isard, Active Contour)



Tracking with Dynamic Models

- Particle Filtering – the algorithm (SIS), A.Vlake and M.Isard

Iterate

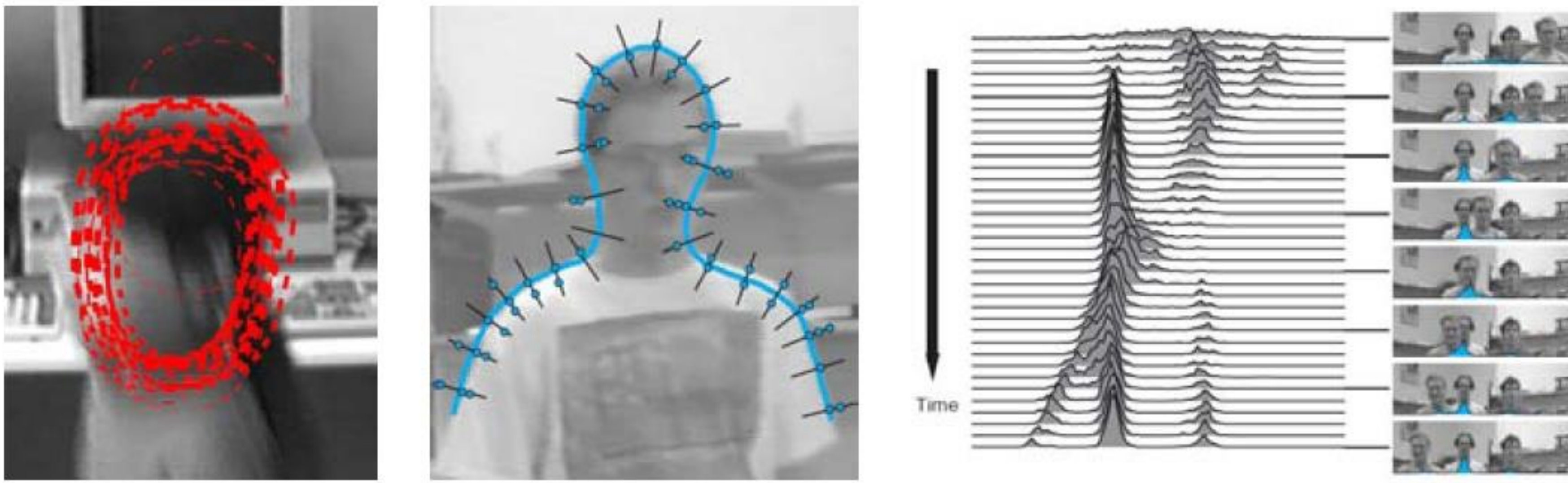
From the "old" sample set $\{\mathbf{s}_{k-1}^{(n)}, \pi_{k-1}^{(n)}, c_{k-1}^{(n)}, n = 1, \ldots, N\}$ at time-step t_{k-1}, construct a "new" sample set $\{\mathbf{s}_k^{(n)}, \pi_k^{(n)}, c_k^{(n)}, n = 1, \ldots, N\}$ for time t_k.

Construct the n^{th} of N new samples as follows:

1. **Select** a sample $\mathbf{s'}_k^{(n)}$ as follows:
   (a) generate a random number $r \in [0, 1]$, uniformly distributed.
   (b) find, by binary subdivision, the smallest j for which $c_{k-1}^{(j)} \geq r$
   (c) set $\mathbf{s'}_k^{(n)} = \mathbf{s}_{k-1}^{(j)}$
2. **Predict** by sampling from
$$p(\mathcal{X}_k | \mathcal{X}_{k-1} = \mathbf{s'}_k^{(n)})$$
to choose each $\mathbf{s}_k^{(n)}$. For instance, in the case that the dynamics are governed by a linear AR process, the new sample value may be generated as: $\mathbf{s}_k^{(n)} = A\,\mathbf{s'}_k^{(n)} + (I - A)\overline{\mathcal{X}} + B\mathbf{w}_k^{(n)}$ where $\mathbf{w}_k^{(n)}$ is a vector of standard normal random variates, and BB^T is the process noise covariance.
3. **Measure** and weight the new position in terms of the measured features $\mathbf{Z}_k$:
$$\pi_k^{(n)} = p(\mathbf{Z}_k | \mathcal{X}_k = \mathbf{s}_k^{(n)})$$
then normalise so that $\sum_n \pi_k^{(n)} = 1$ and store together with cumulative probability as $(\mathbf{s}_k^{(n)}, \pi_k^{(n)}, c_k^{(n)})$ where
$$\begin{aligned} c_k^{(0)} &= 0, \\ c_k^{(n)} &= c_k^{(n-1)} + \pi_k^{(n)} \quad \text{for} \quad n = 1, \ldots, N. \end{aligned}$$

Tracking with Dynamic Models

- Particle Filtering (A.Vlake and M.Isard, Active Contour)



(a) (b) (c)

Tracking can be complex

- **Loss of information** caused by projection of the 3D world on a 2D image,
- **Noise** in images,
- Complex object **motion**,
- **Non-rigid** or articulated nature of objects,
- Partial and full object **occlusions**,
- Complex object **shapes**,
- Scene **illumination** changes, and
- **Real-time** processing requirements.

References and Acknowledgements

- Shapiro and Stockman 2001
- Chapter 19 Forsyth and Ponce 2003
- Chapter 5 Szeliski 2010
- Images drawn from the above references unless otherwise mentioned