# 6.869

Computer Vision

Prof. Bill Freeman

Particle Filter Tracking
- Particle filtering

Readings: F&P Extra Chapter: "Particle Filtering"



## Schedule

- Tuesday, May 3:
  - Particle filters, tracking humans, Exam 2 out
- Thursday, May 5:
  - Tracking humans, and how to write conference papers & give talks, Exam 2 due
- Tuesday, May 10:
  - Motion microscopy, separating shading and paint ("fun things my group is doing")
- Thursday, May 12:
  - 5-10 min. student project presentations, projects due.



## 1D Kalman filter

**Dynamic Model:**

$$x_i \sim N(d_i x_{i-1}, \sigma_{d_i})$$

$$y_i \sim N(m_i x_i, \sigma_{m_i})$$

**Start Assumptions:** $\overline{x}_0^-$ and $\sigma_0^-$ are known

**Update Equations:** Prediction

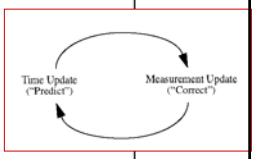
$$\overline{x}_i^- = d_i \overline{x}_{i-1}^+$$

$$\sigma_i^- = \sqrt{\sigma_{d_i}^2 + (d_i \sigma_{i-1}^+)^2}$$

**Update Equations:** Correction

$$x_i^+ = \left( \frac{\overline{x}_i^- \sigma_{m_i}^2 + m_i y_i (\sigma_i^-)^2}{\sigma_{m_i}^2 + m_i^2 (\sigma_i^-)^2} \right)$$

$$\sigma_i^+ = \sqrt{\left( \frac{\sigma_{m_i}^2 (\sigma_i^-)^2}{(\sigma_{m_i}^2 + m_i^2 (\sigma_i^-)^2)} \right)}$$






## Kalman filter for computing an on-line average

- What Kalman filter parameters and initial conditions should we pick so that the optimal estimate for x at each iteration is just the average of all the observations seen so far?



Kalman filter model

$$d_i = 1,\ m_i = 1,\ \sigma_{d_i} = 0,\ \sigma_{m_i} = 1$$

Initial conditions

$$\overline{x}_0^- = 0 \quad \sigma_0^- = \infty$$

| Iteration | 0 | 1 | 2 |
|---|---|---|---|
| $\overline{x}_i^-$ | 0 | $y_0$ | $\frac{y_0 + y_1}{2}$ |
| $\overline{x}_i^+$ | $y_0$ | $\frac{y_0 + y_1}{2}$ | $\frac{y_0 + y_1 + y_2}{3}$ |
| $\sigma_i^-$ | $\infty$ | 1 | $\frac{1}{\sqrt{2}}$ |
| $\sigma_i^+$ | 1 | $\frac{1}{\sqrt{2}}$ | $\frac{1}{\sqrt{3}}$ |



What happens if the x dynamics are given a non-zero variance?

$x_i \sim N(d_i x_{i-1}, \sigma_{d_i})$

$y_i \sim N(m_i x_i, \sigma_{m_i})$

$\overline{x}_0^-$ and $\sigma_0^-$ are known

Prediction

$\overline{x}_i^- = d_i \overline{x}_{i-1}^+$

$\sigma_i^- = \sqrt{\sigma_{d_i}^2 + (d_i \sigma_{i-1}^+)^2}$

Correction

$x_i^+ = \left( \frac{\overline{x}_i^- \sigma_{m_i}^2 + m_i y_i (\sigma_i^-)^2}{\sigma_{m_i}^2 + m_i^2 (\sigma_i^-)^2} \right)$

$\sigma_i^+ = \sqrt{\left( \frac{\sigma_{m_i}^2 (\sigma_i^-)^2}{(\sigma_{m_i}^2 + m_i^2 (\sigma_i^-)^2)} \right)}$

Kalman filter model

$d_i = 1,\ m_i = 1,\ \sigma_{d_i} = 1,\ \sigma_{m_i} = 1$

Initial conditions

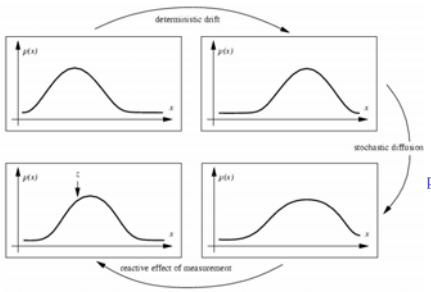
$\overline{x}_0^- = 0 \quad \sigma_0^- = \infty$

| Iteration | 0 | 1 | 2 |
|---|---|---|---|
| $\overline{x}_i^-$ | 0 | $y_0$ | $\frac{y_0 + 2y_1}{3}$ |
| $\overline{x}_i^+$ | $y_0$ | $\frac{y_0 + 2y_1}{3}$ | $\frac{y_0 + 2y_1 + 5y_2}{8}$ |
| $\sigma_i^-$ | $\infty$ | $\sqrt{2}$ | $\sqrt{\frac{5}{3}}$ |
| $\sigma_i^+$ | 1 | $\sqrt{\frac{2}{3}}$ | $\sqrt{\frac{5}{8}}$ |



## (KF) Distribution propagation

prediction from previous time frame



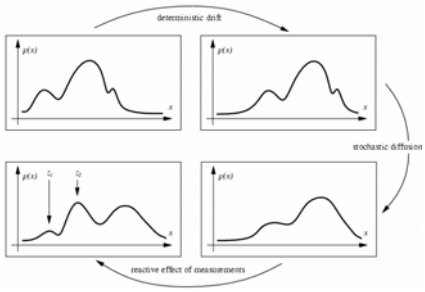

Noise added to that prediction

Make new measurement at next time frame



[Isard 1998]

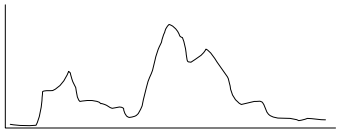
## Distribution propagation






[Isard 1998]

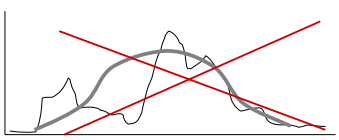
## Representing non-linear Distributions





## Representing non-linear Distributions

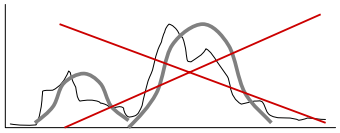
Unimodal parametric models fail to capture real-world densities…





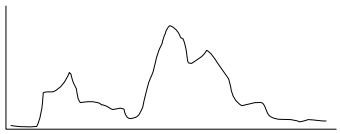
## Representing non-linear Distributions

Mixture models are appealing, but very hard to propagate analytically!



## Representing Distributions using Weighted Samples

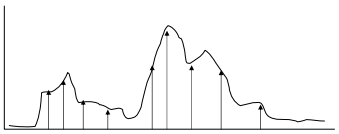
Rather than a parametric form, use a set of samples to represent a density:



13

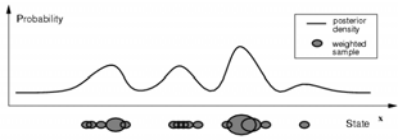
## Representing Distributions using Weighted Samples

Rather than a parametric form, use a set of samples to represent a density:



14

## Representing distributions using weighted samples, another picture




15

[Isard 1998]

## Sampled representation of a probability distribution

Represent a probability distribution

$$p_f(\boldsymbol{X}) = \frac{f(\boldsymbol{X})}{\int f(\boldsymbol{U})d\boldsymbol{U}}$$

by a set of $N$ weighted samples

$$\{(\boldsymbol{u}^i, w^i)\}$$

where $\boldsymbol{u}^i \sim s(\boldsymbol{u})$ and $w^i = f(\boldsymbol{u}^i)/s(\boldsymbol{u}^i)$.

You can also think of this as a sum of dirac delta functions, each of weight w:

$$p_f(x) = \sum_i w^i \delta(x - u^i)$$

16

## Marginalizing a sampled density

If we have a sampled representation of a joint density

$$\{((\boldsymbol{m}^i, \boldsymbol{n}^i), w^i)\}$$

and we wish to marginalize over one variable:

$$p_f(\boldsymbol{M}) = \int p_f(\boldsymbol{M}, \boldsymbol{N})d\boldsymbol{N}$$

we can simply ignore the corresponding components of the samples (!):

$$\begin{aligned}\int g(\boldsymbol{M})p_f(\boldsymbol{M})d\boldsymbol{M} &= \int g(\boldsymbol{M}) \int p_f(\boldsymbol{M}, \boldsymbol{N})d\boldsymbol{N}d\boldsymbol{M} \\ &= \int\int g(\boldsymbol{M})p_f(\boldsymbol{M}, \boldsymbol{N})d\boldsymbol{N}d\boldsymbol{M} \\ &\approx \frac{\sum_{i=1}^N g(\boldsymbol{m}^i)w^i}{\sum_{i=1}^N w^i}\end{aligned}$$

17

## Marginalizing a sampled density

Assume we have a sampled representation of a distribution

$$p_f(\boldsymbol{M}, \boldsymbol{N})$$

given by

$$\{((\boldsymbol{m}^i, \boldsymbol{n}^i), w^i)\}$$

Then

$$\{(\boldsymbol{m}^i, w^i)\}$$

is a representation of the marginal,

$$\int p_f(\boldsymbol{M}, \boldsymbol{N})d\boldsymbol{N}$$

18

## Sampled Bayes Rule

Transforming a Sampled Representation of a Prior into a Sampled Representation of a Posterior:

$$\int g(\boldsymbol{U})\underbrace{p(\boldsymbol{U}|\boldsymbol{V}=v_0)}_{\text{posterior}}d\boldsymbol{U} = \frac{1}{K}\int g(\boldsymbol{U})\underbrace{p(\boldsymbol{V}=\boldsymbol{v}_0|\boldsymbol{U})}_{\text{likelihood}}\underbrace{p(\boldsymbol{U})}_{\text{prior}}d\boldsymbol{U}$$

$$\approx \frac{1}{K}\frac{\sum_{i=1}^{N} g(\boldsymbol{u}^i)p(\boldsymbol{V}=\boldsymbol{v}_0|\boldsymbol{u}^i)w^i}{\sum_{i=1}^{N} w^i}$$

$$\approx \frac{\sum_{i=1}^{N} g(\boldsymbol{u}^i)p(\boldsymbol{V}=\boldsymbol{v}_0|\boldsymbol{u}^i)w^i}{\sum_{i=1}^{N} p(\boldsymbol{V}=\boldsymbol{v}_0|\boldsymbol{u}^i)w^i}$$

19

## Sampled Bayes rule

Assume we have a representation of $p(\boldsymbol{U})$ as

$$\{(\boldsymbol{u}^i, w^i)\}$$

Assume we have an observation $\boldsymbol{V}=\boldsymbol{v}_0$, and a likelihood model $p(\boldsymbol{V}|\boldsymbol{U})$.

The posterior, $p(\boldsymbol{U}|\boldsymbol{V}=\boldsymbol{v}_0)$ is represented by

$$\{(\boldsymbol{u}^i, w'^i)\}$$

where

$$w'^i = p(\boldsymbol{V}=\boldsymbol{v}_0|\boldsymbol{u}^i)w^i$$

20

## Sampled Prediction

$P(\boldsymbol{x}_i|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{i-1})$ = ?

$p(\boldsymbol{X}_{i-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{i-1})$

$\{(\boldsymbol{u}_{i-1}^k, w_{i-1}^k)\}$ → $\boldsymbol{x}_i = \boldsymbol{f}(\boldsymbol{x}_{i-1}) + \xi_i$ →

$\{((f(\boldsymbol{u}_{i-1}^k) + \xi_i^l, \boldsymbol{u}_{i-1}^k), w_{i-1}^k)\}$

$p(\boldsymbol{X}_i, \boldsymbol{X}_{i-1}|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{i-1})$

Drop elements to marginalize to get

$P(\boldsymbol{x}_i|\boldsymbol{y}_0,\ldots,\boldsymbol{y}_{i-1})$ ~=

$\{(f(\boldsymbol{u}_{i-1}^k) + \xi_i^l, w_{i-1}^k)\}$

## Sampled Correction (Bayes rule)

Prior → posterior

Reweight every sample with the likelihood of the observations, given that sample:

$$p(\boldsymbol{Y}_i = \boldsymbol{y}_i|\boldsymbol{X}_i = \boldsymbol{s}_i^{k,-})w_i^{k,-}$$

yielding a set of samples describing the probability distribution after the correction (update) step:

$$\left\{(\boldsymbol{s}_i^{k,-}, p(\boldsymbol{Y}_i = \boldsymbol{y}_i|\boldsymbol{X}_i = \boldsymbol{s}_i^{k,-})w_i^{k,-})\right\}$$

22

## Naïve PF Tracking

- Start with samples from something simple (Gaussian)
- Repeat
  - Correct

$$\left\{(\boldsymbol{s}_i^{k,-}, p(\boldsymbol{Y}_i = \boldsymbol{y}_i|\boldsymbol{X}_i = \boldsymbol{s}_i^{k,-})w_i^{k,-})\right\}$$

  - Predict

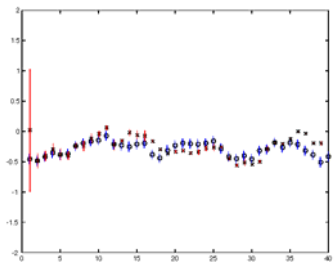
$$\{(f(\boldsymbol{u}_{i-1}^k) + \xi_i^l, w_{i-1}^k)\}$$

But doesn't work that well because of sample impoverishment…

23

## Sample impoverishment

Test with linear case:

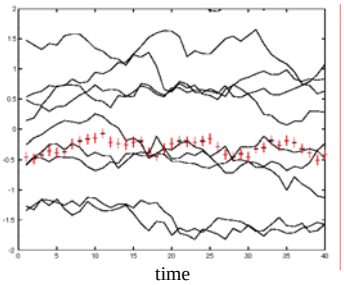


kf: x
pf: o

24

## Sample impoverishment

10 of the 100 particles, along with the true Kalman filter track, with variance:



time

25

## Resample the prior

In a sampled density representation, the frequency of samples can be traded off against weight:

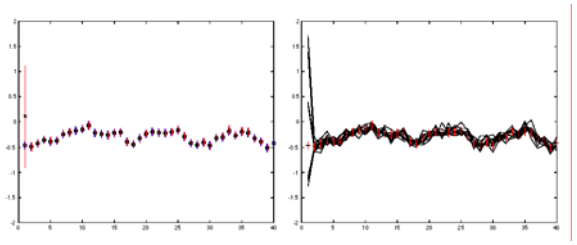
$$(\boldsymbol{s}_k, w_k) \longrightarrow \begin{matrix}(\boldsymbol{s}_k, 1)\\(\boldsymbol{s}_k, 1)\\(\boldsymbol{s}_k, 1)\\\vdots\end{matrix} N_k \text{ copies} \quad \text{s.t.} \quad \frac{N_k}{\sum_k N_k} = w_k$$

These new samples are a representation of the same density.

I.e., make N draws with replacement from the original set of samples, using the weights as the probability of drawing a sample.

26

## Resampling concentrates samples



27

**A practical particle filter with resampling**

**Initialization:** Represent $P(X_0)$ by a set of $N$ samples

$$\left\{\left(s_0^{k,-}, w_0^{k,-}\right)\right\}$$

where

$$s_0^{k,-} \sim P_s(S) \text{ and } w_0^{k,-} = P(s_0^{k,-})/P_s(S = s_0^{k,-})$$

Ideally, $P(X_0)$ has a simple form and $s_0^{k,-} \sim P(X_0)$ and $w_0^{k,-} = 1$.

**Prediction:** Represent $P(X_i|y_0, y_{i-1})$ by

$$\left\{\left(s_i^{k,-}, w_i^{k,-}\right)\right\}$$

where

$$s_i^{k,-} = f(s_{i-1}^{k,+}) + \xi_i^k \text{ and } w_{i-1}^{k,-} = w_{i-1}^{k,+} \text{ and } \xi_i^k \sim N(0, \Sigma_{d_i})$$

**Correction:** Represent $P(X_i|y_0, y_i)$ by

$$\left\{\left(s_i^{k,+}, w_i^{k,+}\right)\right\}$$

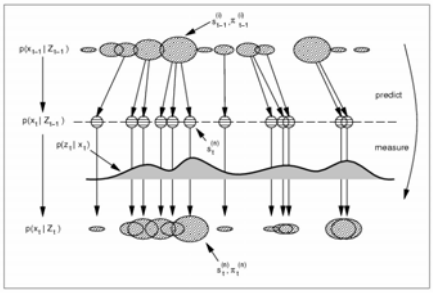
where

$$s_i^{k,+} = s_i^{k,-} \text{ and } w_i^{k,+} = P(Y_i = y_i|X_i = s_i^{k,-}) w_i^{k,-}$$

**Resampling:** Normalise the weights so that $\sum_i w_i^{k,+} = 1$ and compute the variance of the normalised weights. If this variance exceeds some threshold, then construct a new set of samples by drawing, with replacement, $N$ samples from the old set, using the weights as the probability that a sample will be drawn. The weight of each sample is now $1/N$.

28

## A variant (predict, then resample, then correct)




29

[Isard 1998]

## A variant (animation)

31
[Isard 1998]

## Applications

Tracking
- hands
- bodies
- leaves

32

## Contour tracking

33
[Isard 1998]

## Head tracking

34
[Isard 1998]

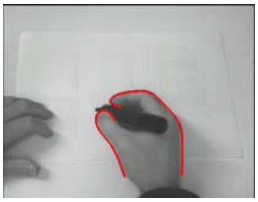
## Leaf tracking

35
[Isard 1998]

## Hand tracking

36
[Isard 1998]

## Mixed state tracking



37

[Isard 1998]

## Outline

- Sampling densities
- Particle filtering

[Figures from F&P except as noted]

38