Mean shift

Tomáš Svoboda, svoboda@cmp.felk.cvut.cz
Czech Technical University in Prague, Center for Machine Perception
http://cmp.felk.cvut.cz
Last update: April 7, 2014

Talk Outline

- appearance based tracking
- patch similarity using histogram
- tracking by mean shift
- experiments, discussion

1Please note that the lecture will be accompanied be several sketches and derivations on the blackboard and few live-interactive demos in Matlab

Meanshift segmentation of colours - color distribution

Figure from [2]

Meanshift segmentation of colours - color modes seeking

Figure from [2]
Mean shift segmentation - intensity and space

$u, v$ are here spatial pixel coordinates
different normalization for intensity and spatial coordinates

Multivariate kernel density estimator

Given $n$ data points $x_i$ in $d$-dimensional space $\mathbb{R}^d$.

$$f_{h,K}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

- looking for extremum of $f_{h,K}(x)$
- gradient $\nabla f_{h,K}(x) = 0$

Differentiating density estimator I

$$f_{h,K}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

$$\nabla f_{h,K}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} \nabla K\left(\frac{x - x_i}{h}\right)$$
Kernels and profiles

Normal kernel, \( K_N(x) = \exp\left(\frac{1}{2}\|x\|^2\right) \)

Epanechnikov kernel, \( K_E(x) = c(1 - \|x\|^2) \) if \( \|x\| \leq 1 \):

Kernel profile:
\( k_N(x) = \exp\left(-\frac{x}{2}\right) \), for \( x \geq 0 \).

\( k_E(x) = 1 - x \), for \( 0 \leq x \leq 1 \)

Differentiating density estimator II

\[ \nabla f_{h,K}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} \nabla K \left( \frac{x - x_i}{h} \right) \]

using profiles, instead of kernels

\[ K \left( \frac{x - x_i}{h} \right) = c_k k \left( \frac{\|x - x_i\|^2}{h^2} \right) \]

Detailed derivation/explanation on the board and in the talk-note.pdf.

Mean-shift iterations

Assuming a reasonable differentiable kernel \( K \), iterate till convergence:

\[ y_{k+1} = \frac{\sum_{i=1}^{n} g(\|y_k - x_i\|^2)}{\sum_{i=1}^{n} g(\|y_k - x_i\|^2)} \]

\( g \) is the derivative of kernel profile.
Mean shift segmentation - visionbook demo
chapter 7, [5], http://visionbook.felk.cvut.cz/downloads.html

\[ K(x) = c k_{E} \left( \frac{x^s}{h_s} \right)^2 + \frac{x^r}{h_r} \left( \frac{x^r}{h_r} \right)^2 \],

Appearance based tracking

Illustration from [1]

Histogram based representation
histogram difference

assume normalized histograms, i.e.
\[ \sum_{u=1}^{m} p_u = 1 \]

\[ d = \sqrt{1 - \rho[p,q]} \]

where \( \rho[p,q] \) is the Bhattacharyya coefficient

\[ \rho[p,q] = \sum_{u=1}^{m} \sqrt{p_u q_u} \]

Similarity measured by the Bhattacharyya coefficient

The object is the “4” plate and the model is histogram of image intensities.

\[ s(y) = \sum_{u=1}^{m} \sqrt{p_u(y)q_u} \]

where \( p(y) \) is the histogram of image patch at position \( y \) and \( q \) is the histogram of the template.
Problem: finding modes in probability density

- the complete enumeration of similarity surface can be costly.
- can we do it faster and more elegantly?

Mean-shift tracking - Bhattacharya coefficient

\[ s(y) = \sum_{u=1}^{m} \sqrt{p_u(y)q_u} \]

model, coordinates \( x_i \) centered at 0:

\[ q_u = C_n \sum_{i=1}^{n} k(||x_i^*||^2)\delta(b(x_i^*) - u) \]

target candidate centered at \( y \):

\[ p_u(y) = C_h \sum_{i=1}^{n_h} k\left( \frac{||y - x_i||}{h} \right)\delta(b(x_i) - u) \]

Detailed derivation/explanation on the board and in the talk-note.pdf.

Mean-shift tracking - ratio histogram

Ratio histogram:

\[ r_u = \min\left( \frac{q_u}{p_u}, 1 \right) \]

where \( q \) is the histogram of the target and \( p \) is the histogram of the current frame. \( w_i = r_b(x_i) \) (just binning)

Image intensities (or colors) are transformed into weights, \( w_i \), by back projection of the ratio histogram. Mean-shift iterations:

\[ y_{k+1} = \frac{\sum_{i=1}^{n} w_i x_i g(||y_k - x_i||^2)}{\sum_{i=1}^{n} w_i g(||y_k - x_i||^2)} \]
ms tracking - object and its model

Figure from [5]

ms tracking - iterations

maximizing the Bhattacharyya coefficient

References

Mean-shift originally from [3].


End