Mean shift [1]

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

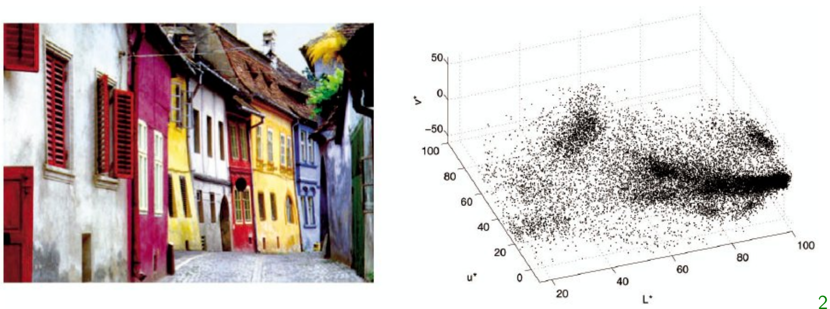
Last update: April 7, 2014

Talk Outline

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

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

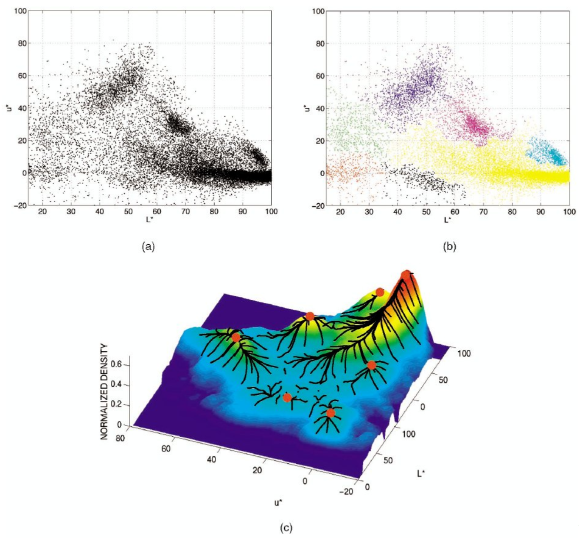
Meanshift segmentation of colours - color distribution



[2]

[2] Figure from [2]

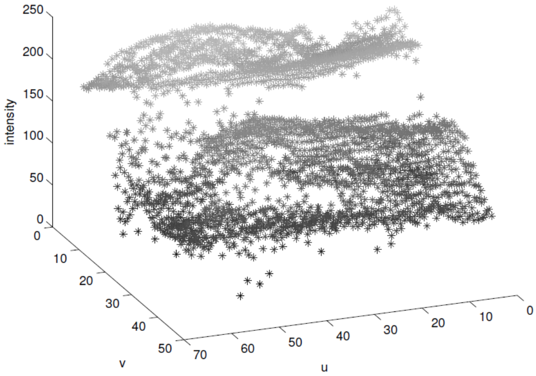
Meanshift segmentation of colours - color modes seeking



[3]

[3] Figure from [2]

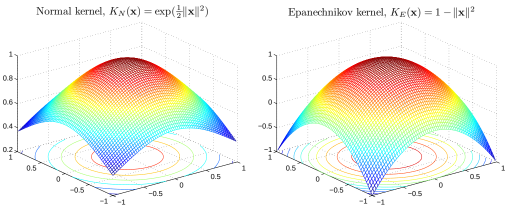
Mean shift segmentation - intensity and space

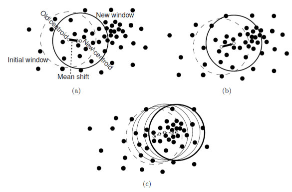


u, v are here spatial pixel coordinates

different normalization for intensity and spatial coordinates

Multivariate kernel density estimator





Given n data points $\mathbf{x}_i$ in d-dimensional space R^d.

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

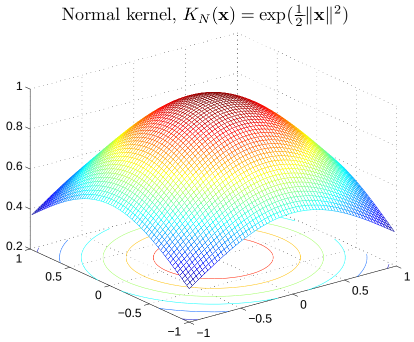
- looking for extremum of $f_{h,K}(\mathbf{x})$
- gradient $\nabla f_{h,K}(\mathbf{x}) = \mathbf{0}$

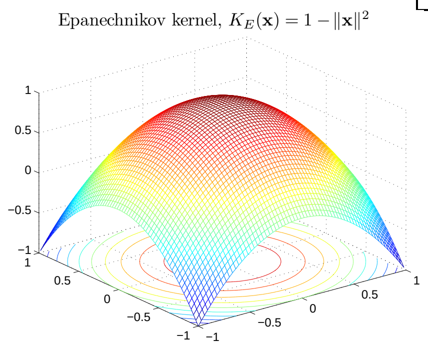
Differentiating density estimator I

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

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

Kernels and profiles





$$K_N(\mathbf{x}) = c \exp\left(\frac{1}{2}\|\mathbf{x}\|^2\right)$$

Kernel profile:

$$k_N(x) = \exp\left(-\frac{x}{2}\right), \text{ for } x \geq 0.$$

$$K_E(\mathbf{x}) = c(1 - \|\mathbf{x}\|^2) \text{ if } \|x\| \leq 1;$$

Kernel profile:

$$k_E(x) = 1 - x, \quad \text{for } 0 \leq x \leq 1$$

Differentiating density estimator II

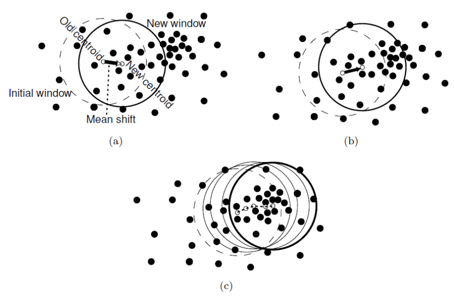
$$\nabla f_{h,K}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^{n} \nabla K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right)$$

using profiles, instead of kernels

$$K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) = c_k k\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)$$

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

Mean-shift iterations



Assuming a reasonable differentiable kernel K, iterate till convergence:

$$\mathbf{y}_{k+1} = \frac{\sum_{i=1}^{n} \mathbf{x}_i g(\|\mathbf{y}_k - \mathbf{x}_i\|^2)}{\sum_{i=1}^{n} g(\|\mathbf{y}_k - \mathbf{x}_i\|^2)}$$

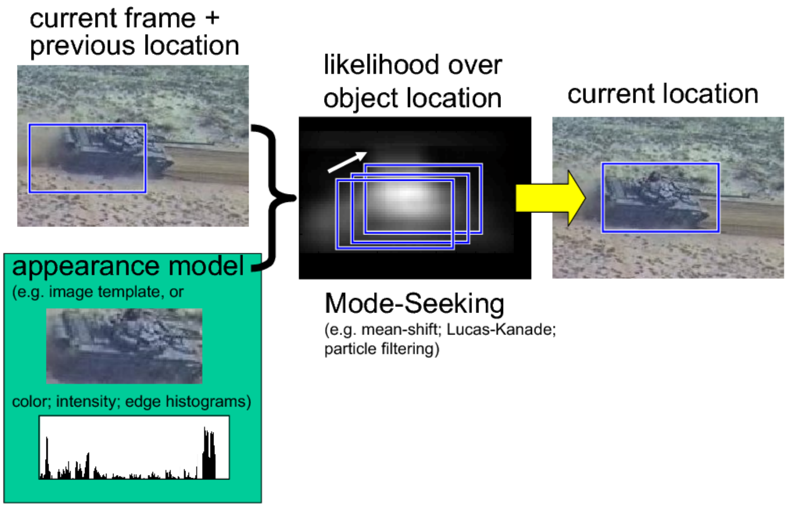
g is the derivative of kernel profile.

Mean shift segmentation - visionbook demo

chapter 7, [5], http://visionbook.felk.cvut.cz/downloads.html

$$K(\mathbf{x}) = c\, k_E \left(\left\| \frac{\mathbf{x}^s}{h_s} \right\|^2 + \left\| \frac{\mathbf{x}^r}{h_r} \right\|^2 \right) ,$$

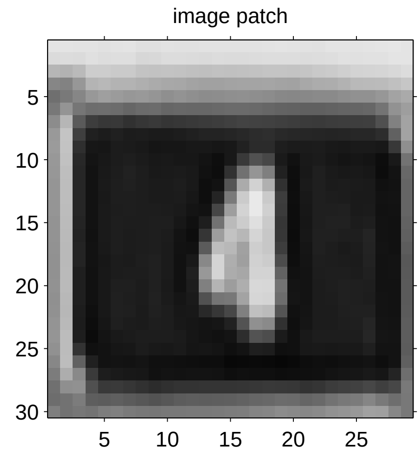
Appearance based tracking

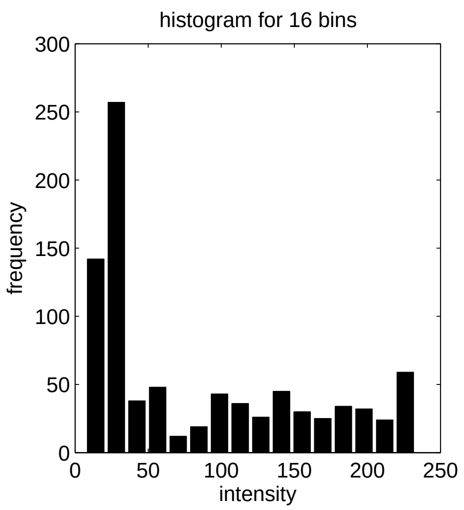


4

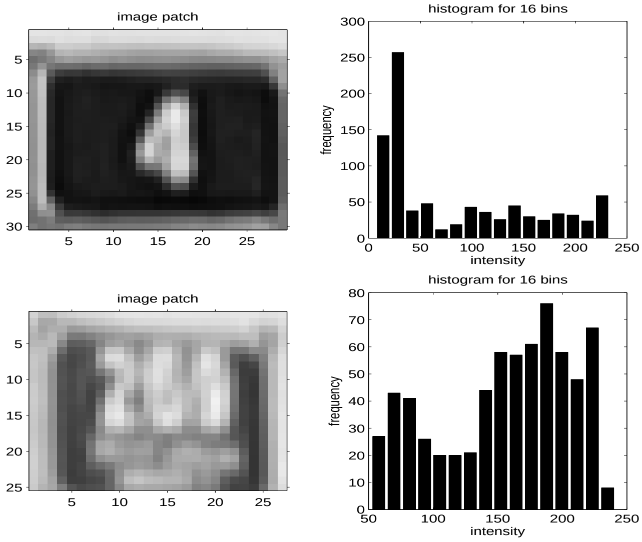
[4] illustration from [1]

Histogram based representation





Patch comparison



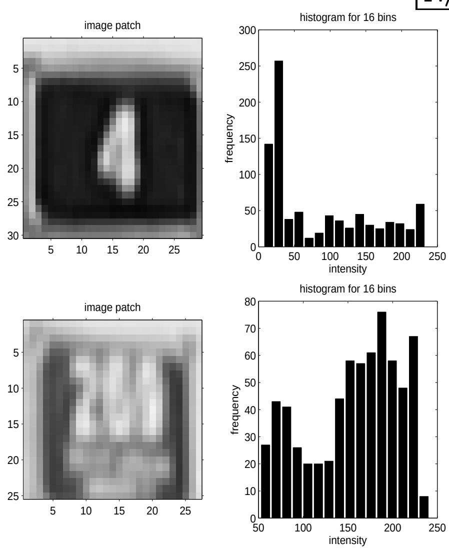
histogram difference

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

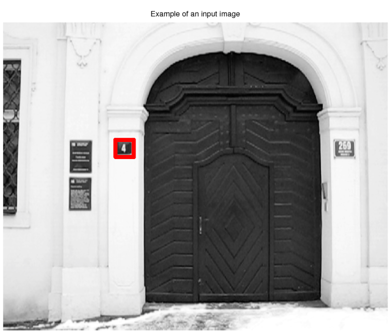
$$d = \sqrt{1 - \rho[p, q]}$$

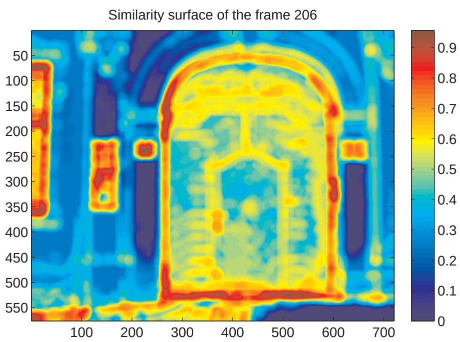
where $\rho[p, q]$ is the Bhattacharyya coefficient

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



Similarity measured by the Bhattacharyya coefficient



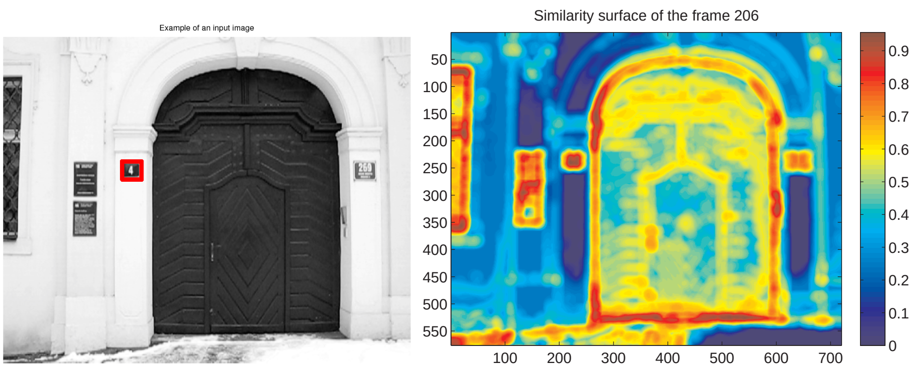


The object is the "4" plate and the model is histogram of image intensities.

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

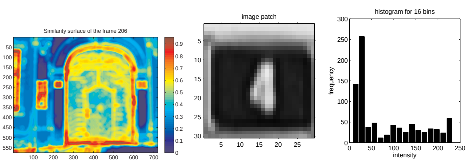
where $p(\mathbf{y})$ is the histogram of image patch at position $\mathbf{y}$ and q is the histogram of the template.

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 coeficient



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

model, coordinates $\mathbf{x}_i^*$ centered at $\mathbf{0}$:

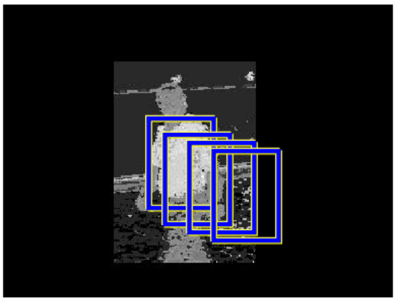
$$q_u = C \sum_{i=1}^{n} k(\|\mathbf{x}_i^*\|^2) \delta(b(\mathbf{x}_i^*) - u)$$

target candidate centered at $\mathbf{y}$:

$$p_u(\mathbf{y}) = C_h \sum_{i=1}^{n_h} k\left(\left\| \frac{\mathbf{y} - \mathbf{x}_i}{h} \right\|^2 \right) \delta(b(\mathbf{x}_i) - u)$$

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

Mean-shift tracking - ratio histogram



Ratio histogram:

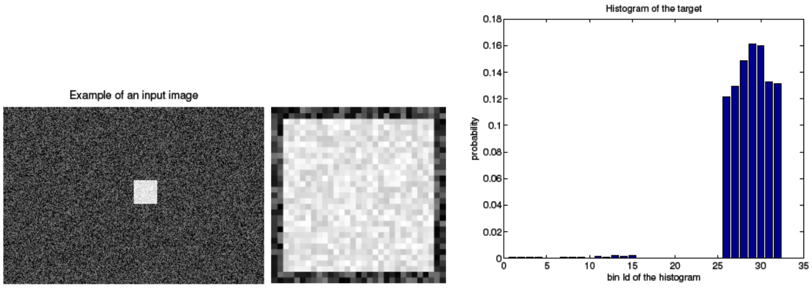
$$r_u = \min\left(\frac{q_u}{p_u}, 1 \right)$$

where q is the histogram of the target and p is the histogram of the current frame. $w_i = r_{b(\mathbf{x}_i)}$ (just binning)

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

$$\mathbf{y}_{k+1} = \frac{\sum_{i=1}^{n} w_i \mathbf{x}_i g(\|\mathbf{y}_k - \mathbf{x}_i\|^2)}{\sum_{i=1}^{n} w_i g(\|\mathbf{y}_k - \mathbf{x}_i\|^2)}$$

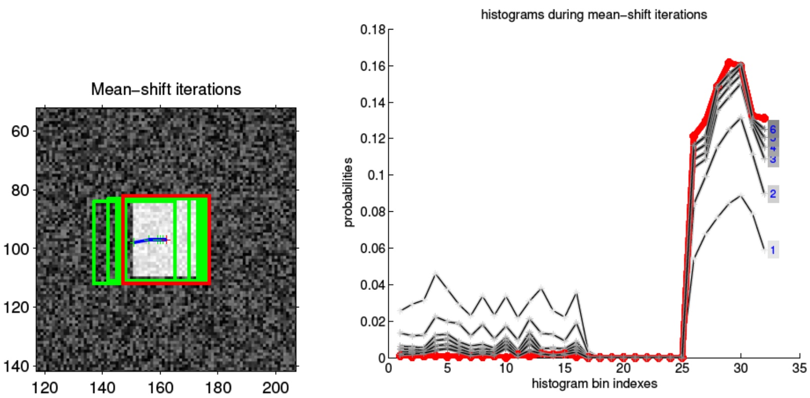
ms tracking - object and its model



(a) (b) (c) [5]

[5]Figure from [5]

ms tracking - iterations



(a) (b) [6]

maximizing the Bhattacharyya coefficient

[6]Figure from [5], chapter 16, http://visionbook.felk.cvut.cz/downloads.html

References

Mean-shift originally from [3].

[1] Robert Collins. CSE/EE486 Computer Vision I. slides, web page. http://www.cse.psu.edu/~rcollins/CSE486/. Robert kindly gave general permission to reuse the material.

[2] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Analysis, 24(5):603–619, May 2002.

[3] Keinosuke Fukunaga and Larry D. Hostetler. The estimation of the gradient of a density function, with appilcations in pattern recognition. IEEE Transactions on Information Theory, 21(1):32–40, January 1975.

[4] Milan Šonka, Václav Hlaváč, and Roger Boyle. Image Processing, Analysis and Machine Vision. Thomson, 3rd edition, 2007.

[5] Tomáš Svoboda, Jan Kybic, and Václav Hlaváč. Image Processing, Analysis and Machine Vision. A MATLAB Companion. Thomson, 2007. Accompanying www site http://visionbook.felk.cvut.cz.

End