Mean shift [1]

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

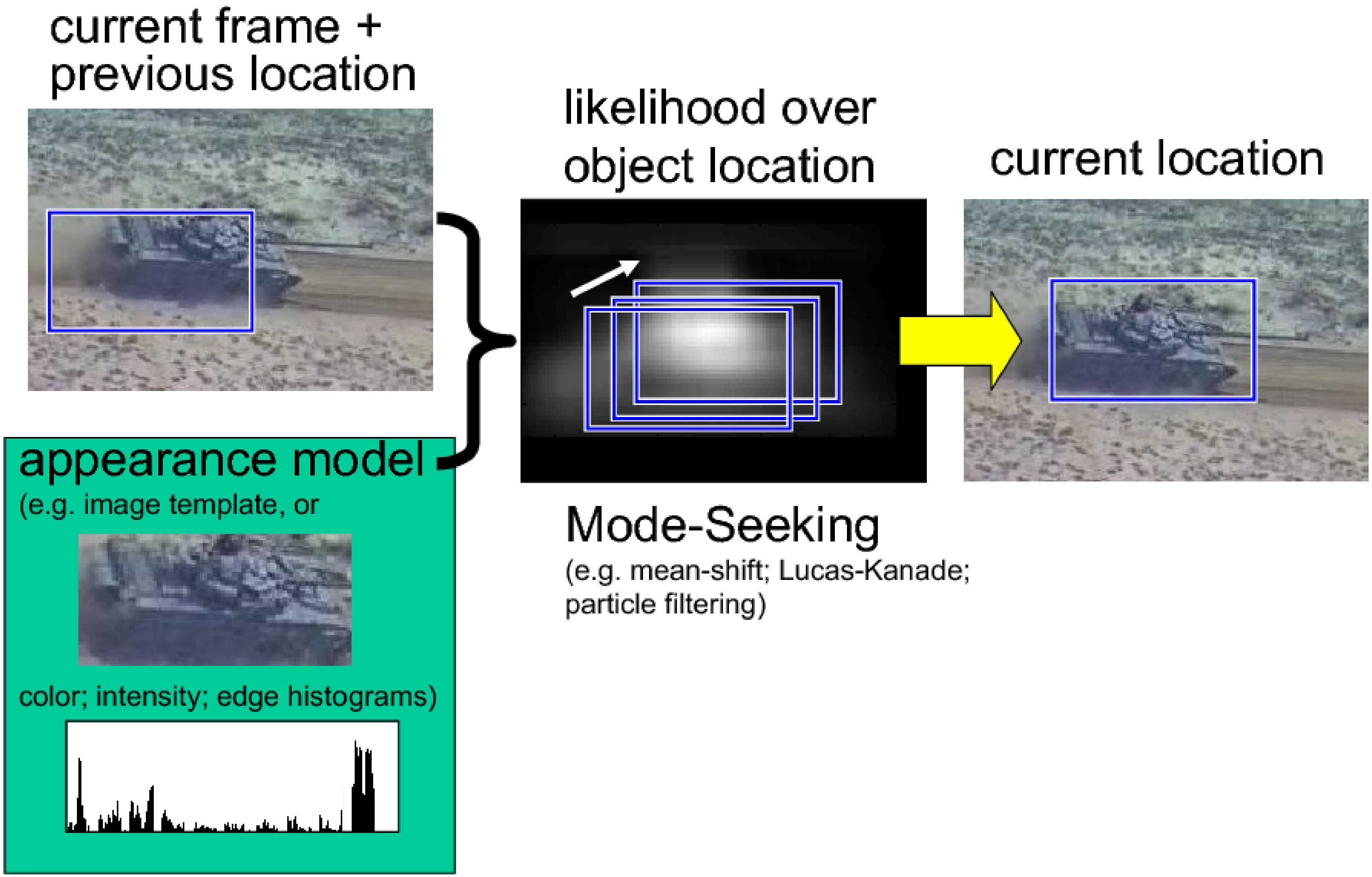
Last update: April 8, 2013

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

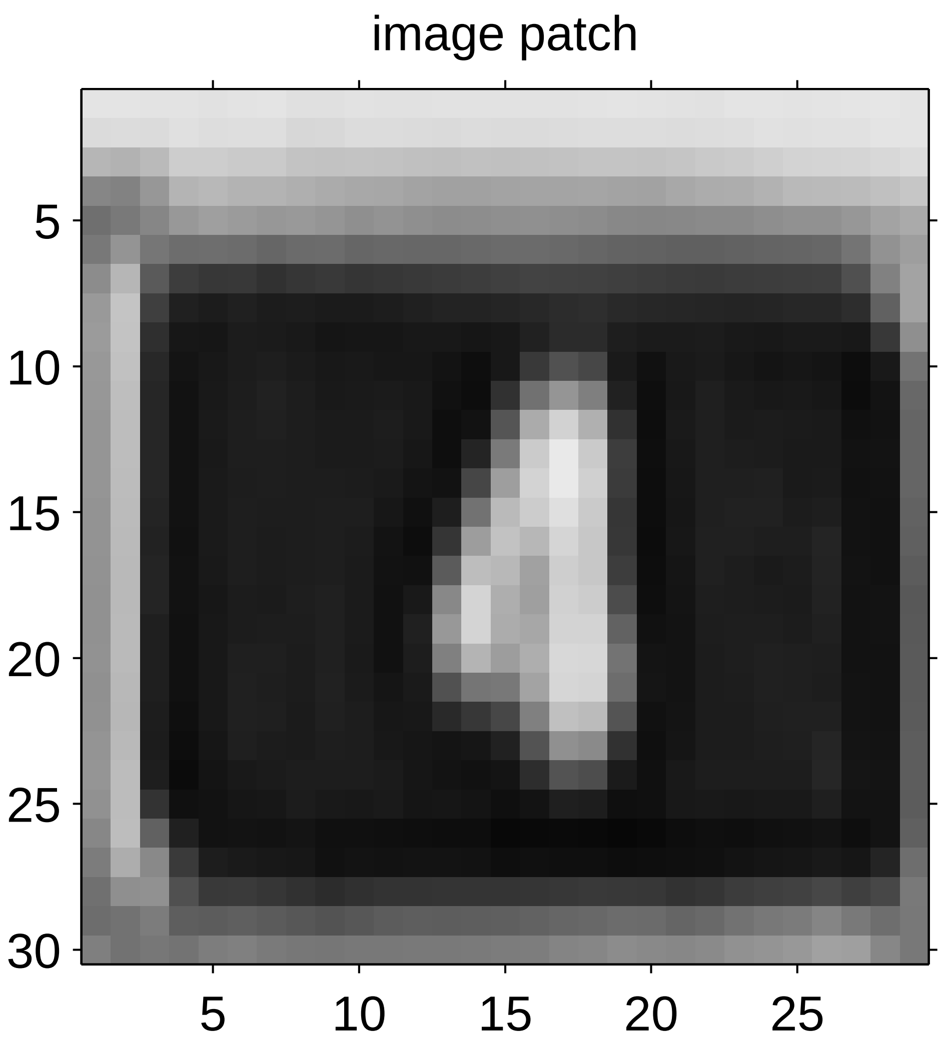
Appearance based tracking

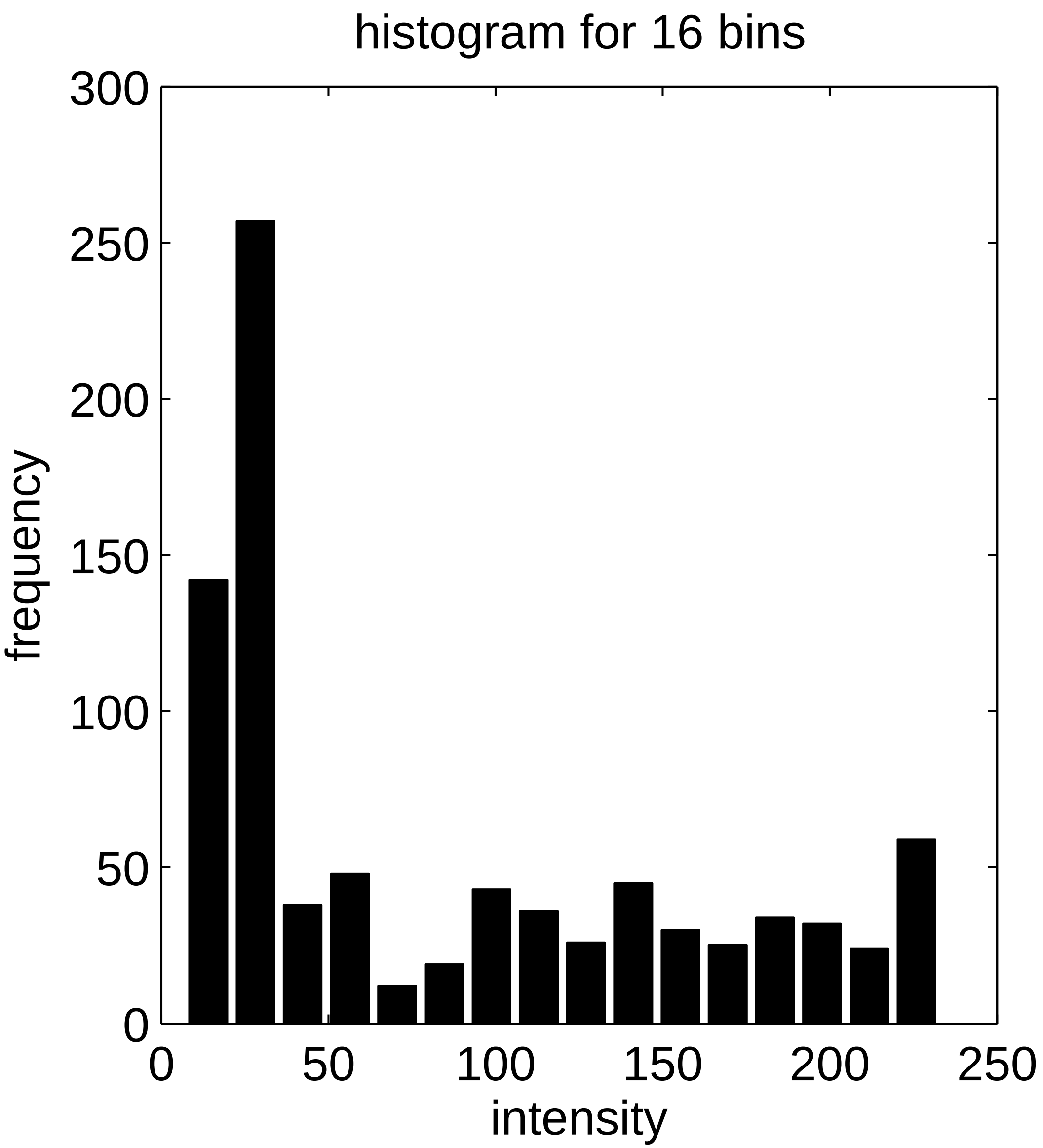


2

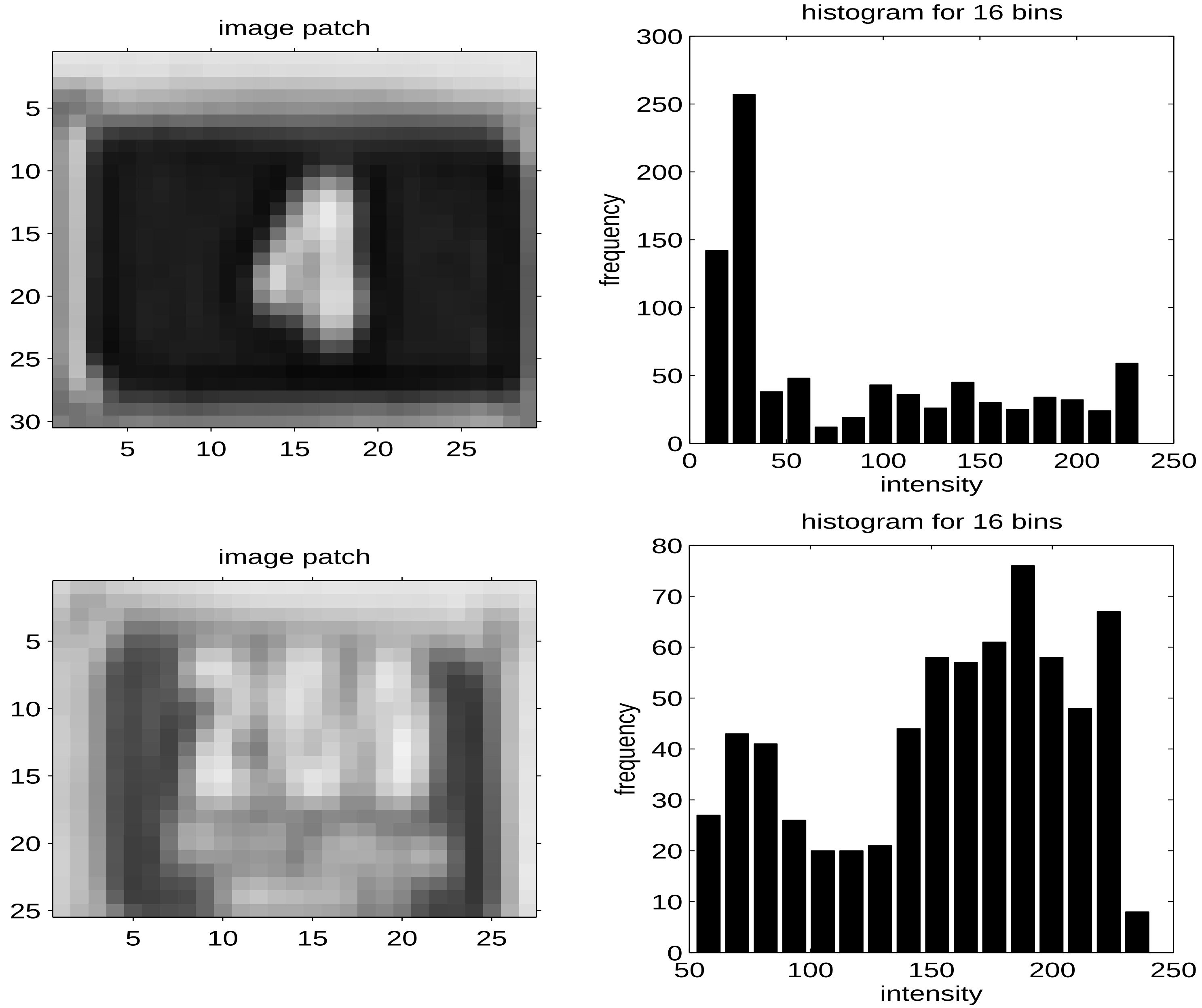
[2] illustration from [1]

Histogram based representation





Patch comparison



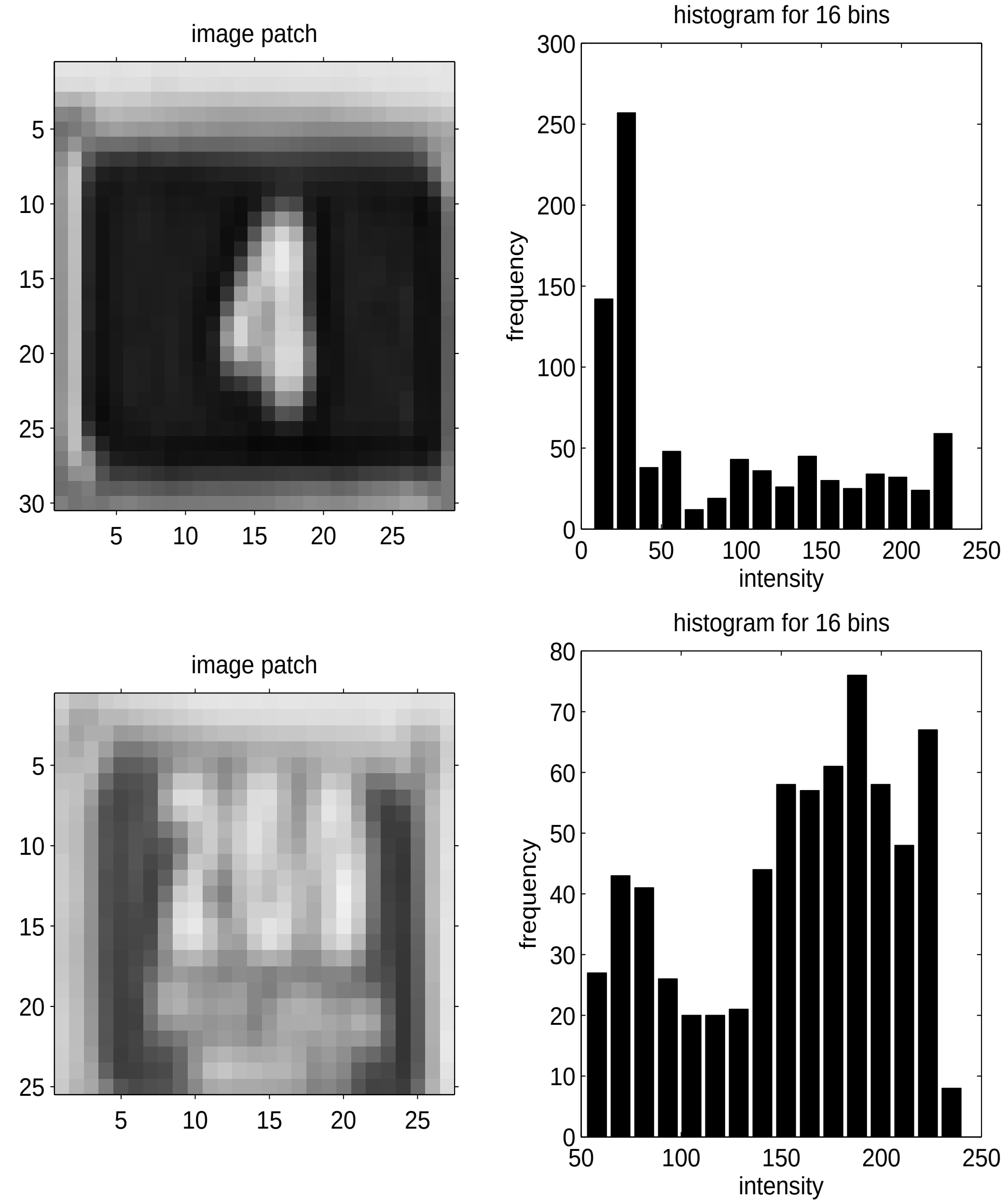
histogram difference

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

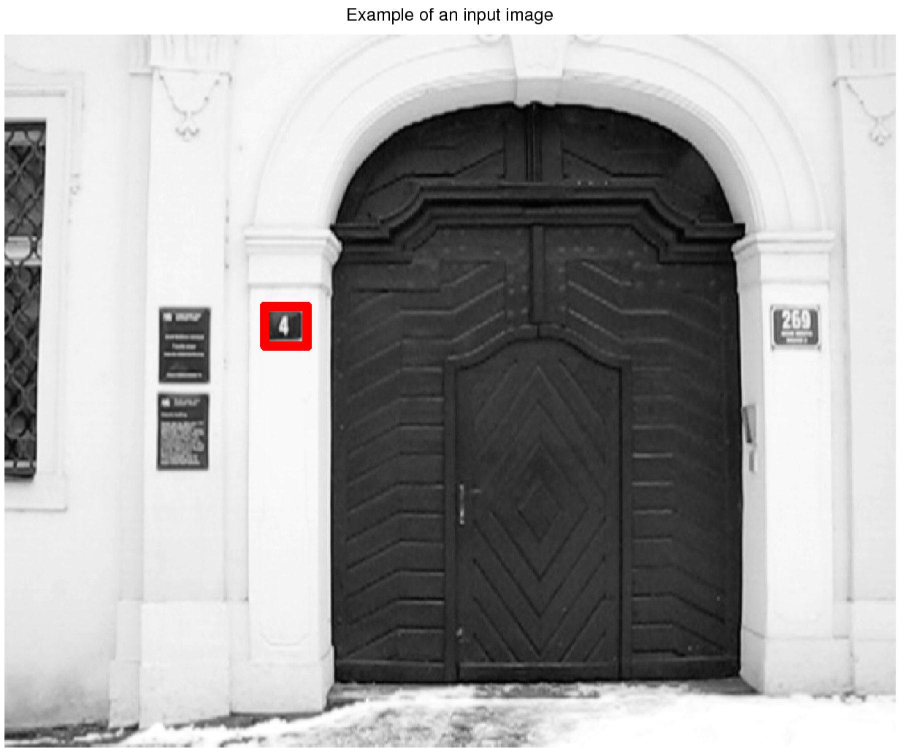
$$d = \sqrt{1 - \rho[p, q]}$$

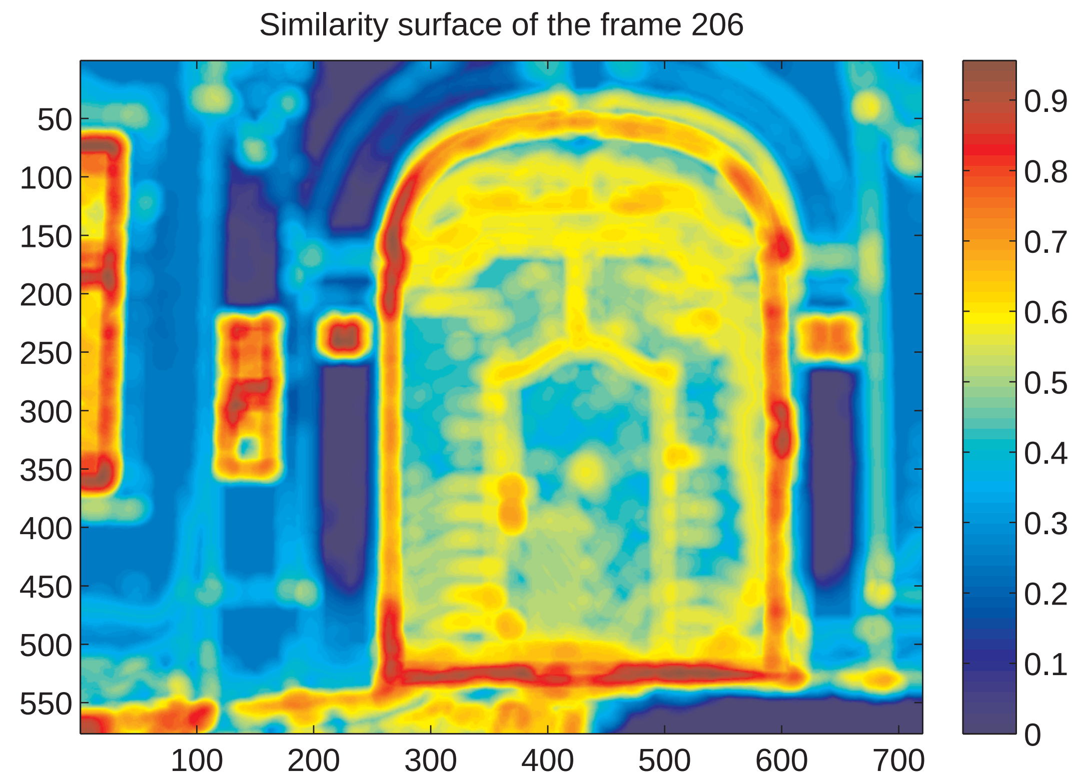
where $\rho[p, q]$ is the Bhattacharyya coefficient

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



Similarity measured by the Bhattacharyya coefficient





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

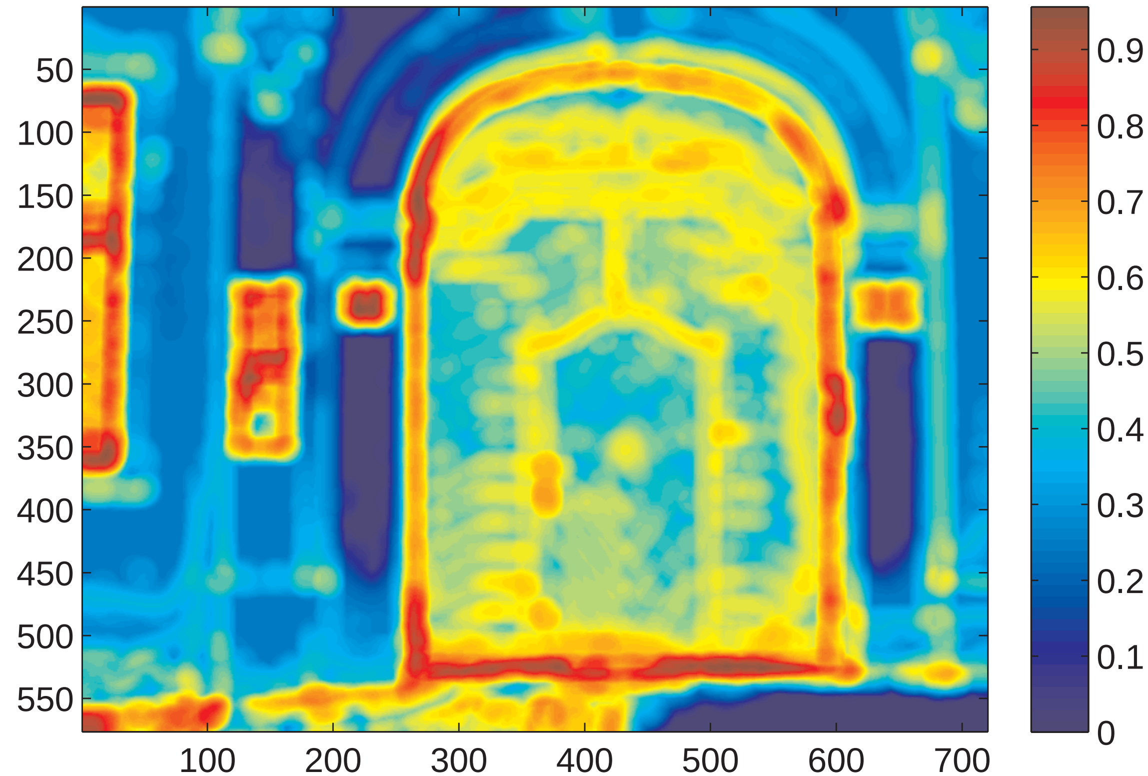
$$s(\mathbf{y}) = \sum_{u=1}^{m} \sqrt{p_u(\mathbf{y}) q_u}$$

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

Example of an input image

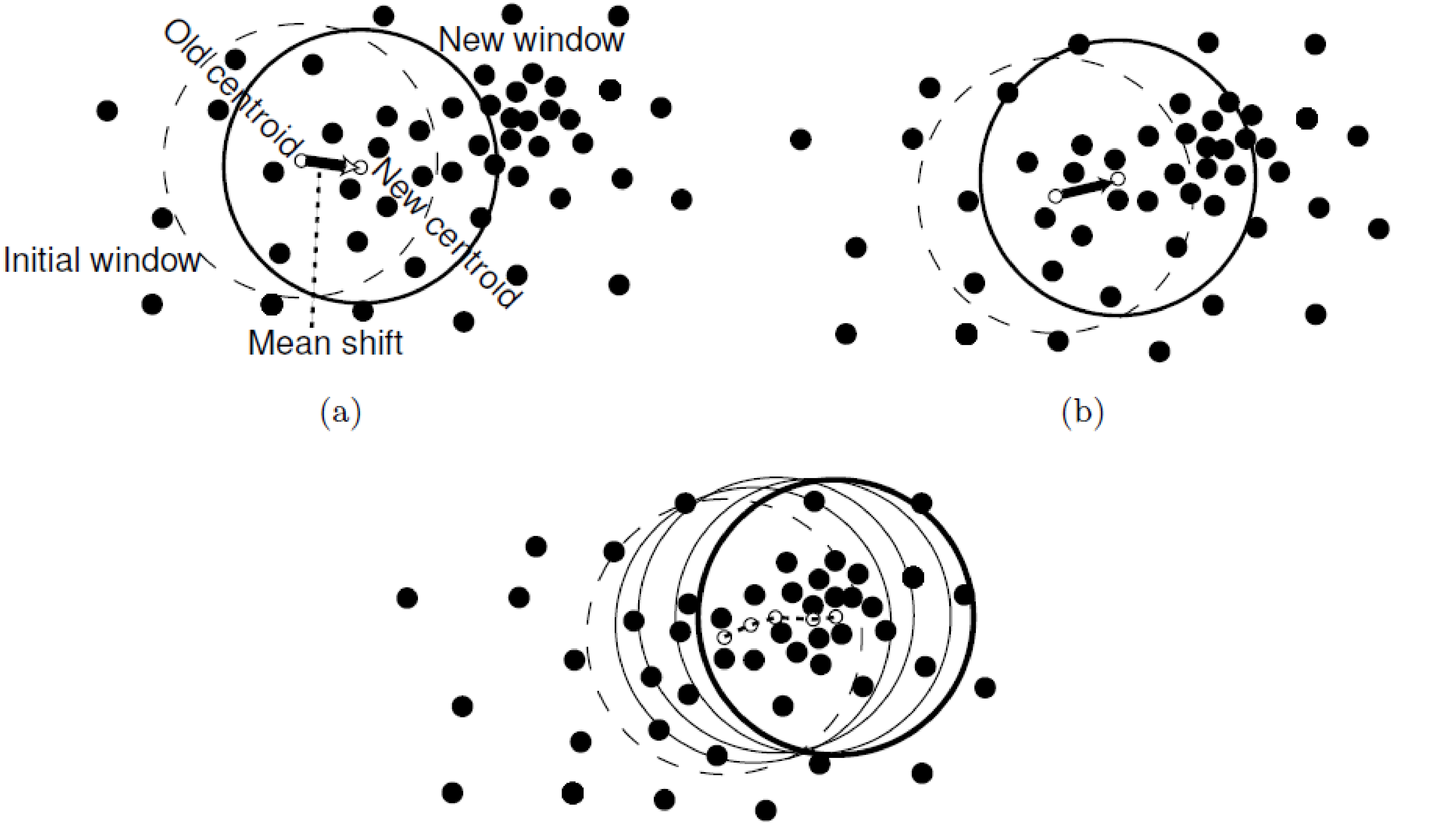
Similarity surface of the frame 206



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

Density Gradient Estimation

Mean shift procedure



(a)

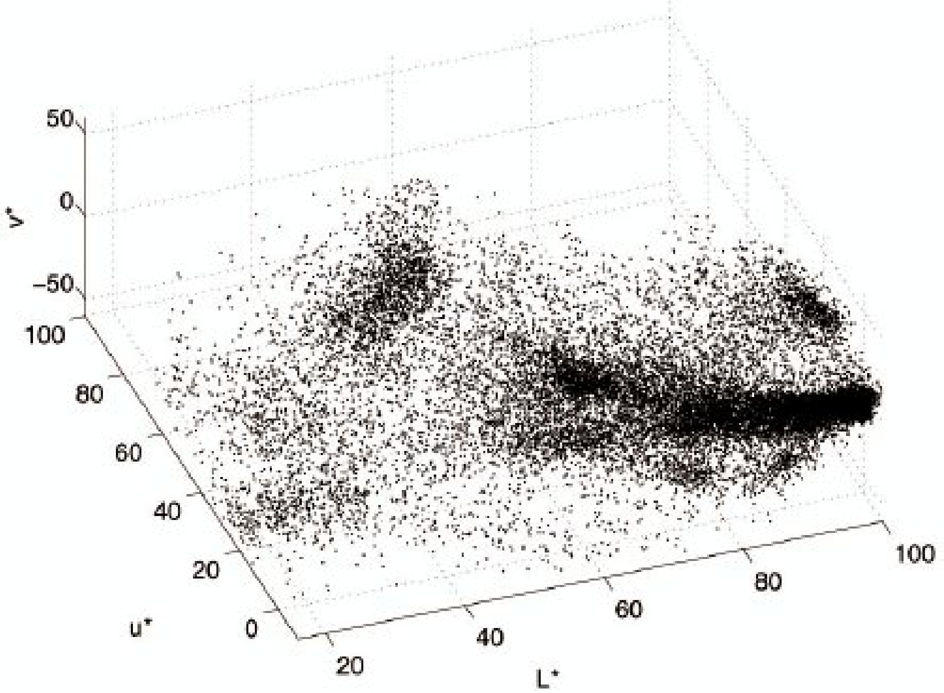
(b)

(c)

3

[3]Figure borrowed from [4]

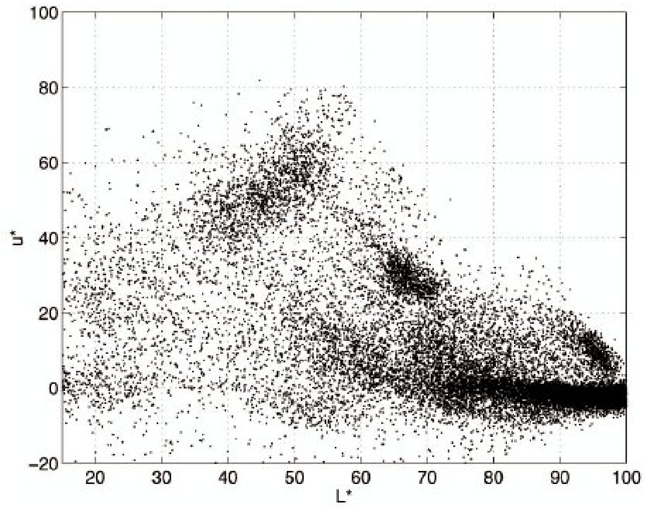
Meanshift segmentation of colours - color distribution



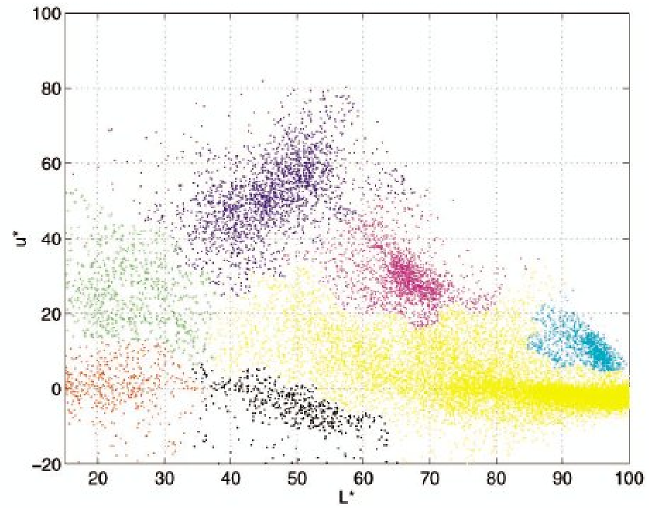
[4]

[4] Figure from [2]

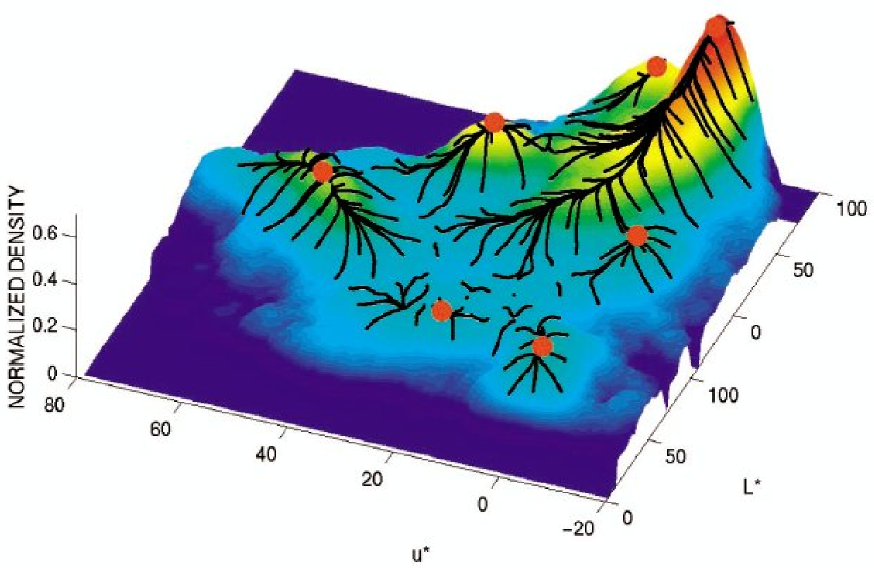
Meanshift segmentation of colours - color modes seeking



(a)



(b)

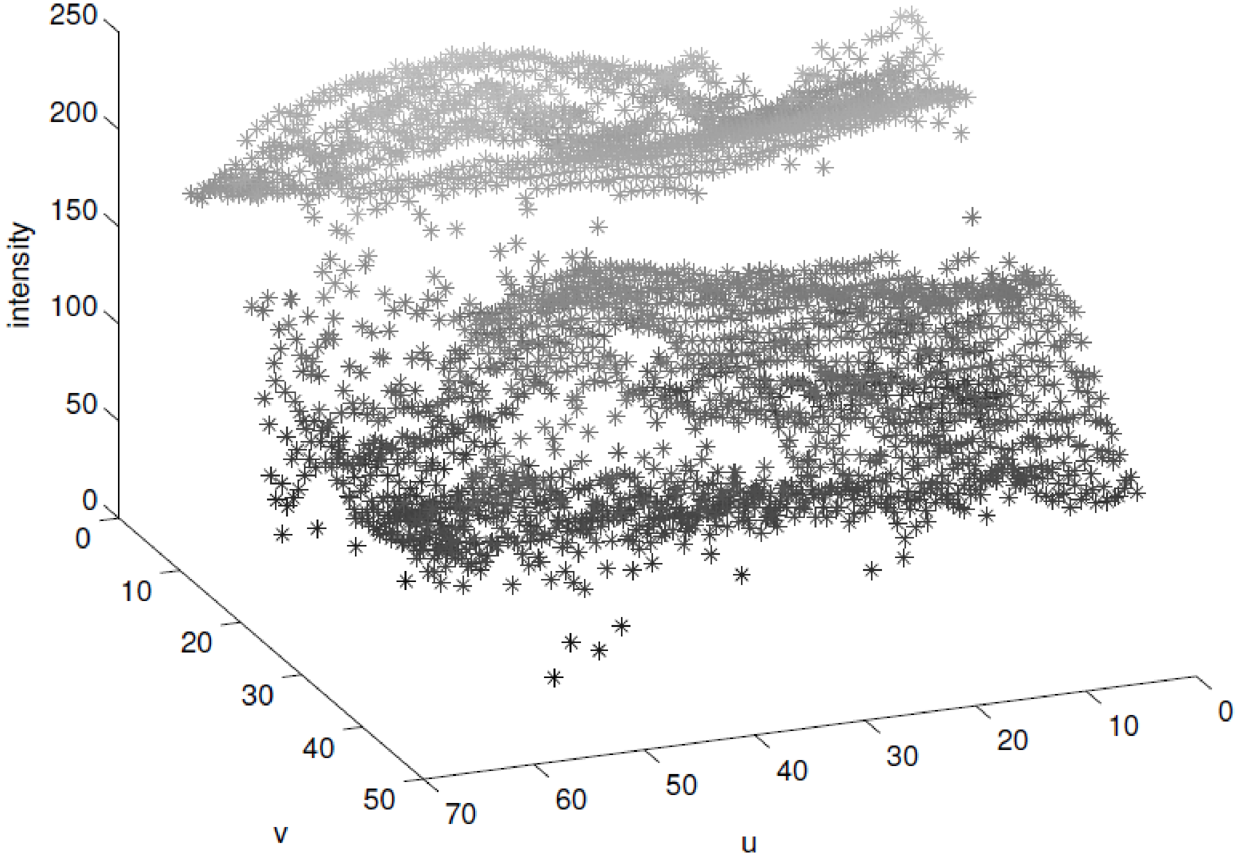


(c)

5

[5] Figure from [2]

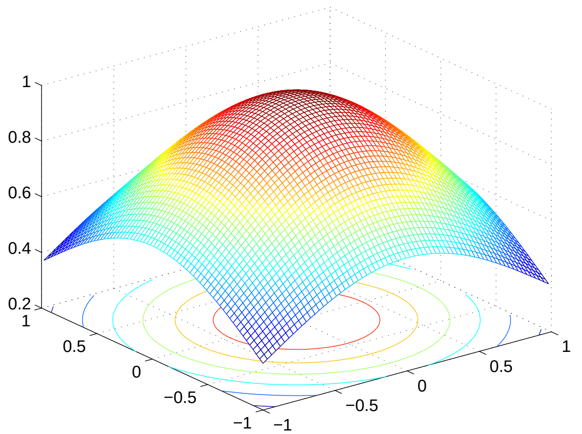
Mean shift segmentation - intensity and space

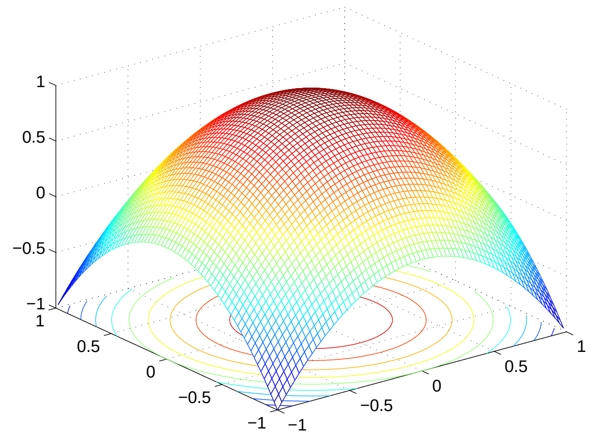


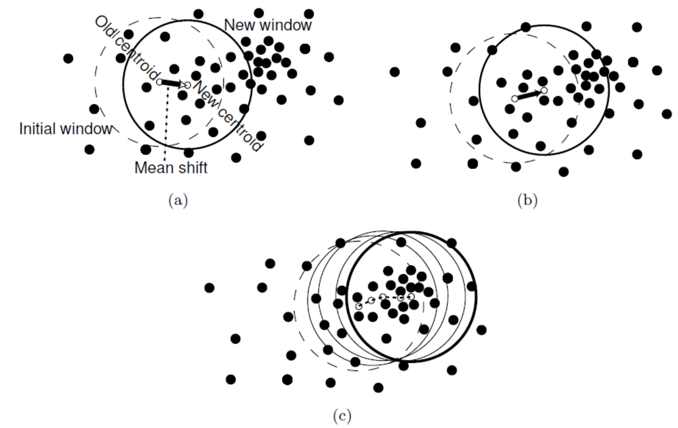
u, v are here spatial pixel coordinates

different normalization for intensity and spatial coordinates

Multivariate kernel density estimator



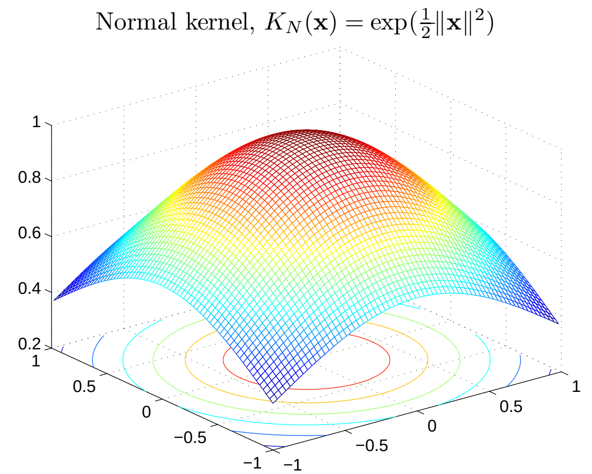


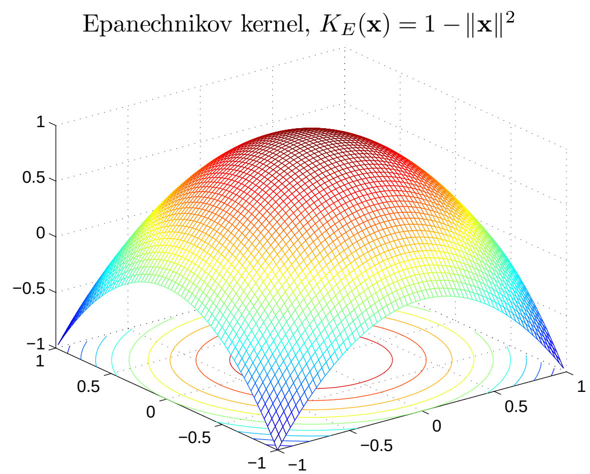


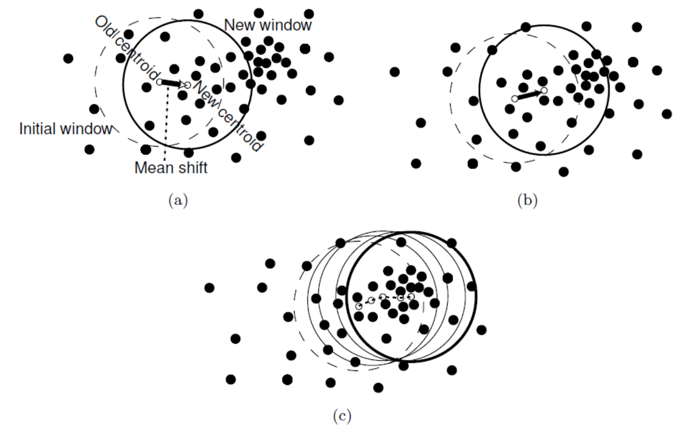
Given n data points $\mathbf{x}_i$ in d-dimensional space R^d.

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

Multivariate kernel density estimator





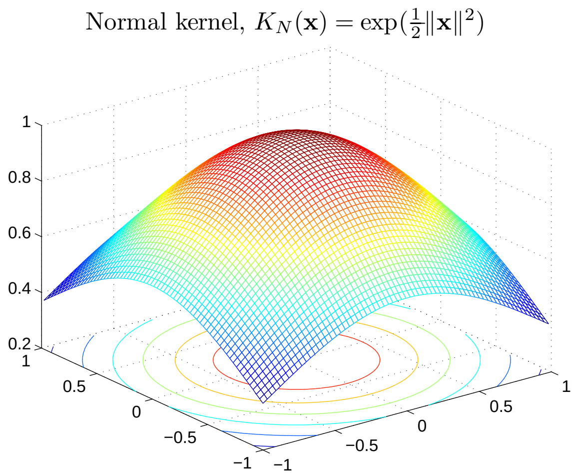


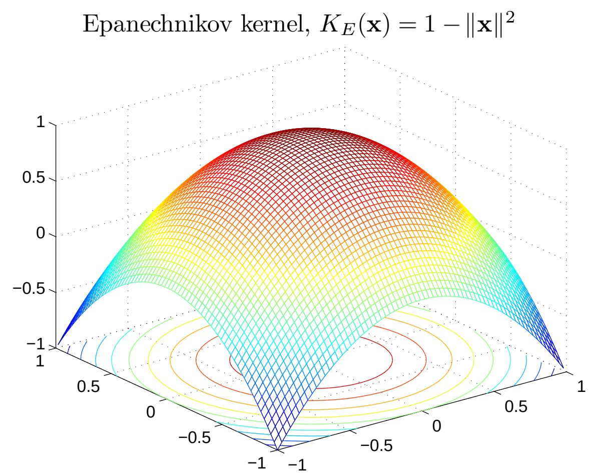
Given n data points $\mathbf{x}_i$ in d-dimensional space R^d.

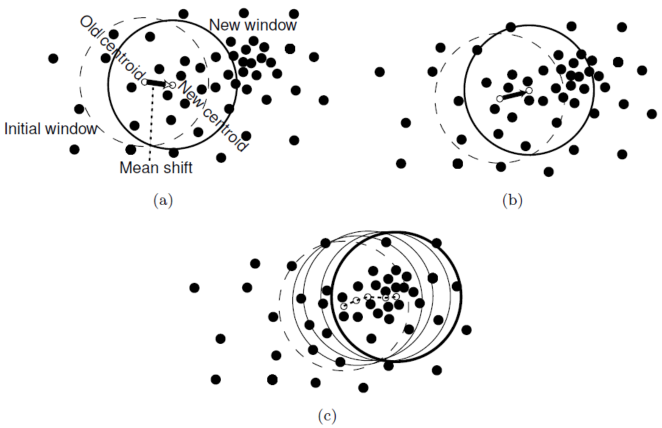
$$\tilde{f}_{h,K}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right)$$

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

Multivariate kernel density estimator







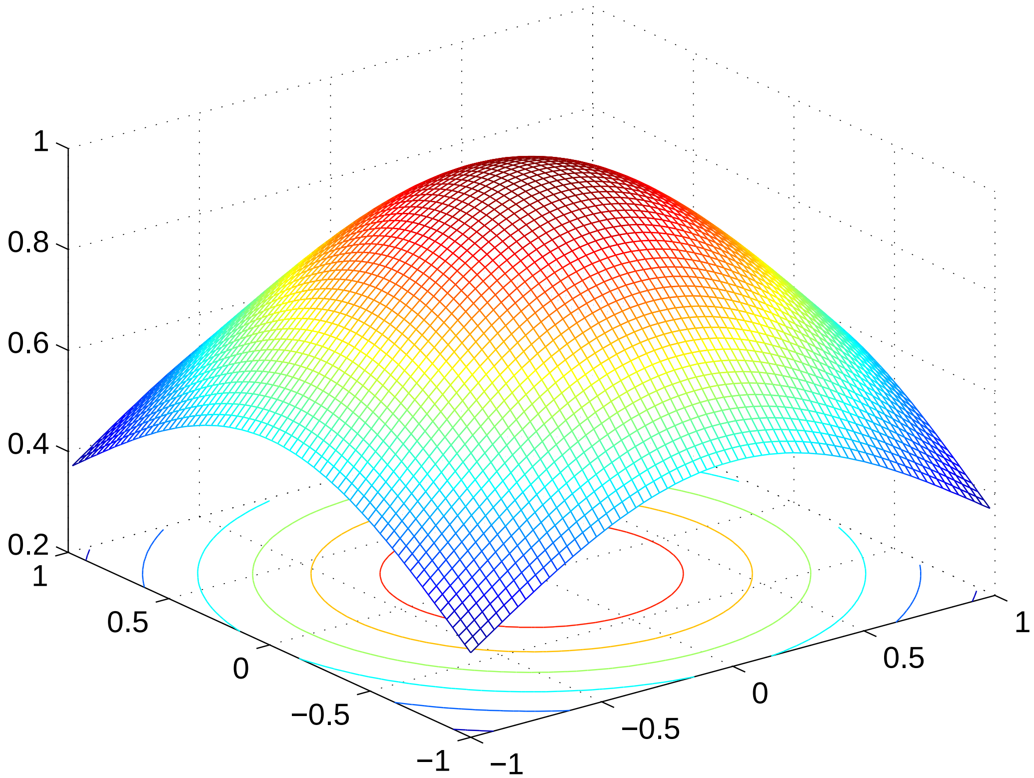
Given n data points $\mathbf{x}_i$ in d-dimensional space R^d.

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

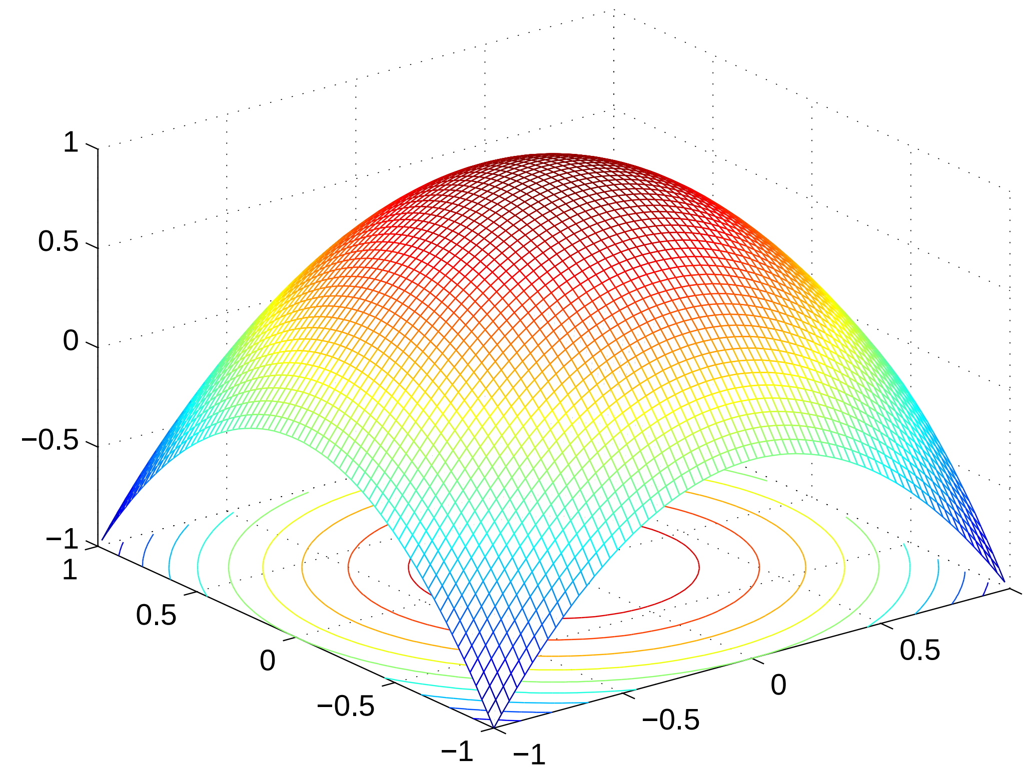
- looking for extremum of $f_{h,K}(\mathbf{x})$
- gradient $\nabla f_{h,K}(\mathbf{x}) = \mathbf{0}$

Kernels

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



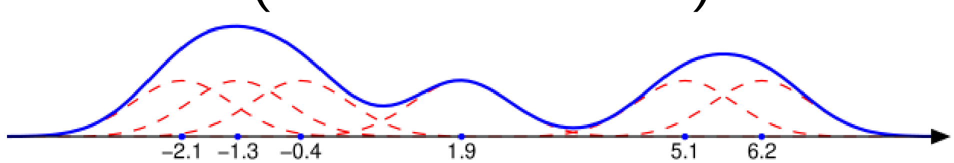
Epanechnikov kernel, $K_E(\mathbf{x}) = 1 - \|\mathbf{x}\|^2$



Can be seen as membership function.

Remind Kernel density estimation (Parzen method).

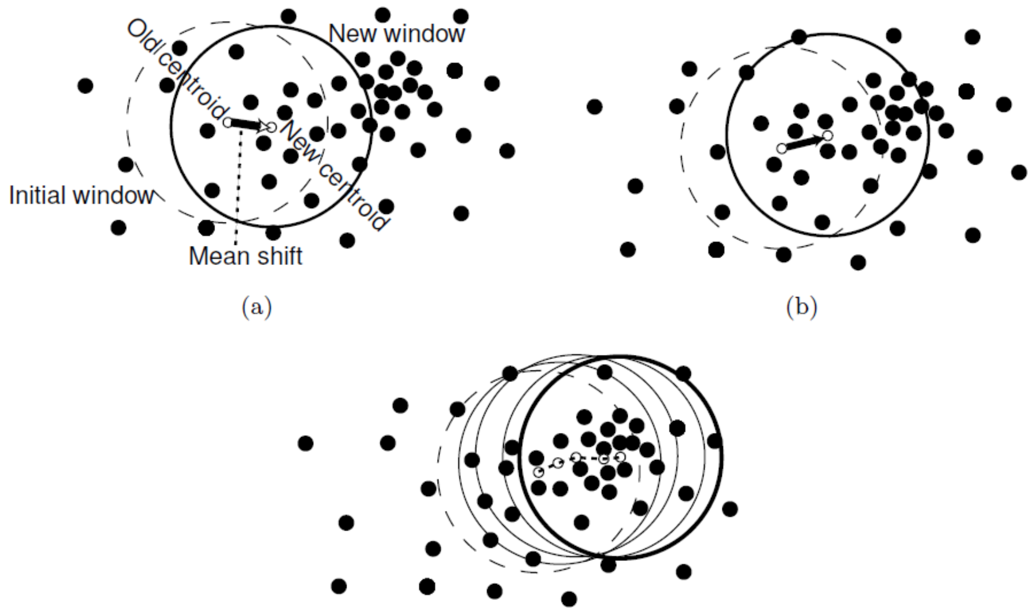
Remember convolution?



[6]

[6] Taken from http://en.wikipedia.org/wiki/Kernel_density_estimation

Mean-shift iterations



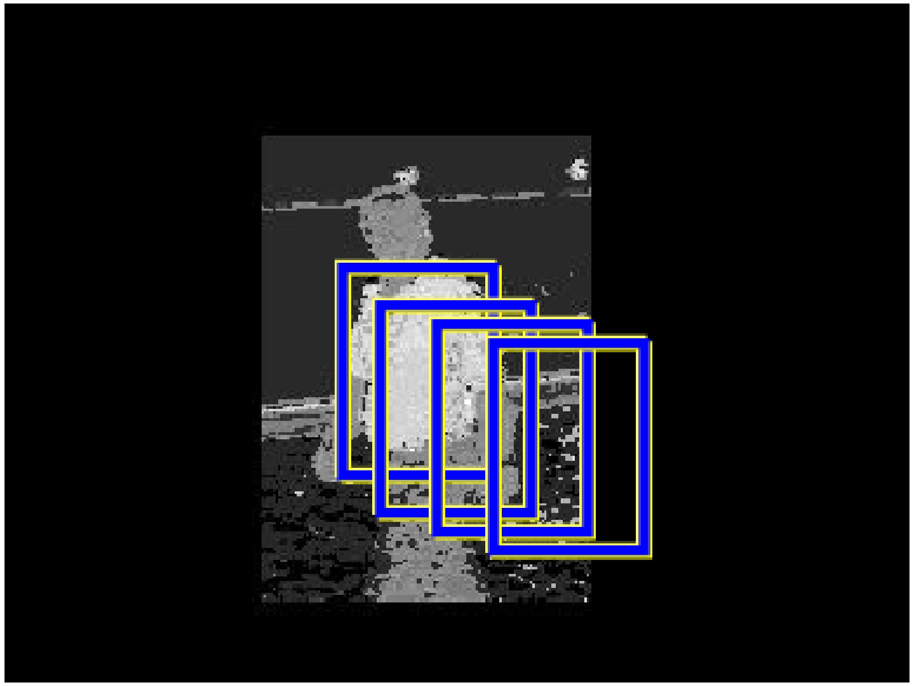
(a) (b) (c)

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 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:

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

Mean-shift tracking - Bhattacharya coeficient

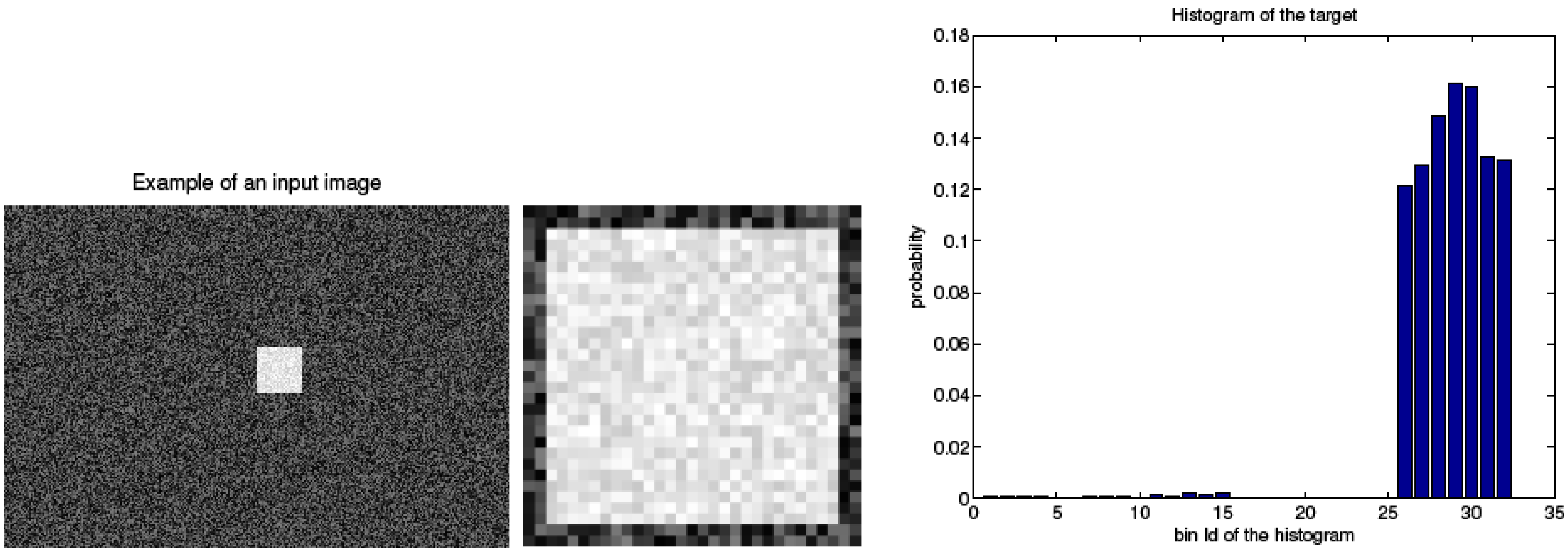
model, coordinates $\mathbf{x}_i^*$ centered at $\mathbf{0}$:

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

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

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

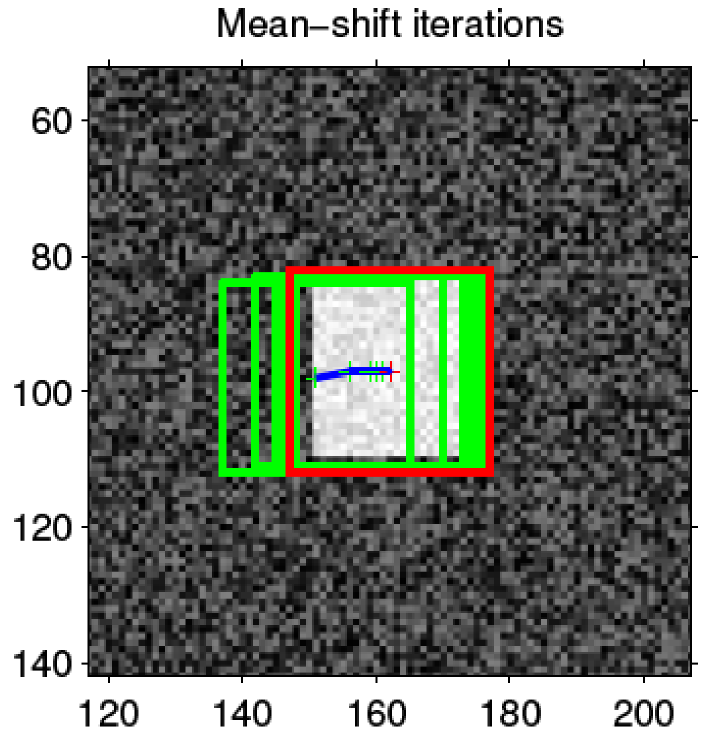
ms tracking - object and its model



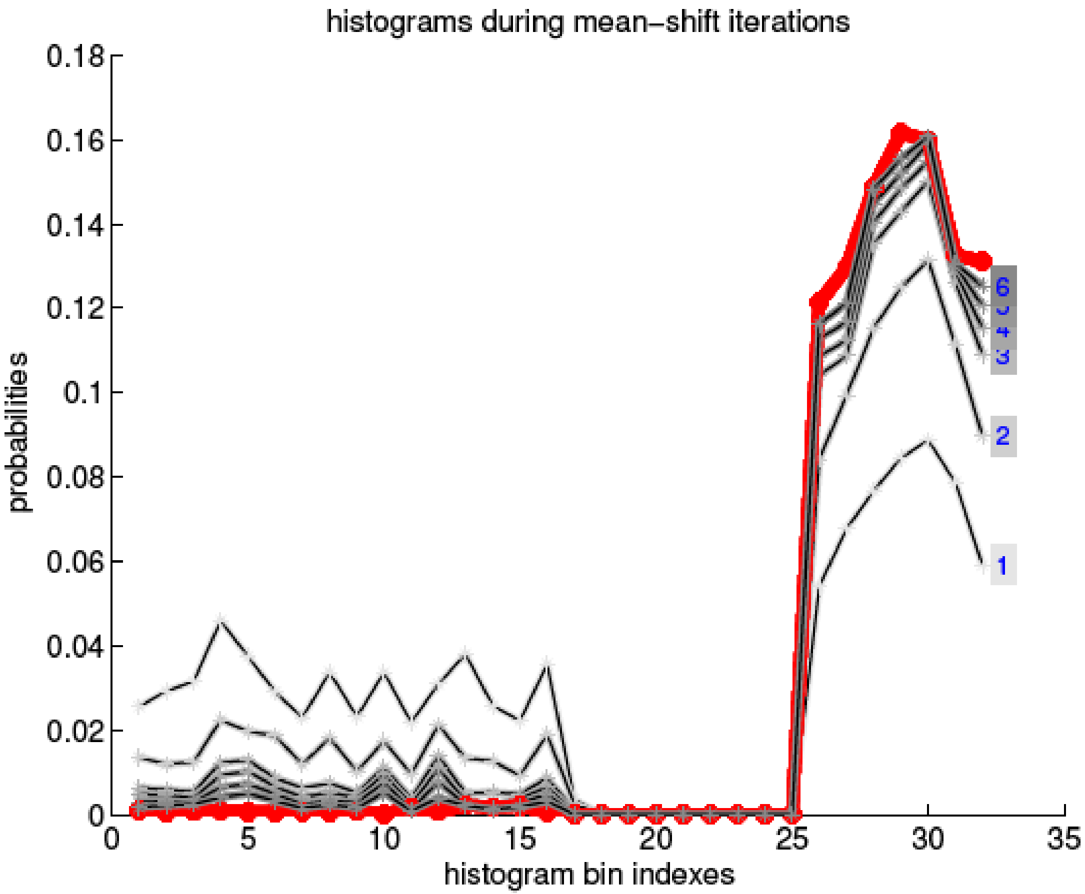
(a) (b) (c) [7]

[7]Figure from [5]

ms tracking - iterations



(a)



(b) [8]

[8] Figure from [5]

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.

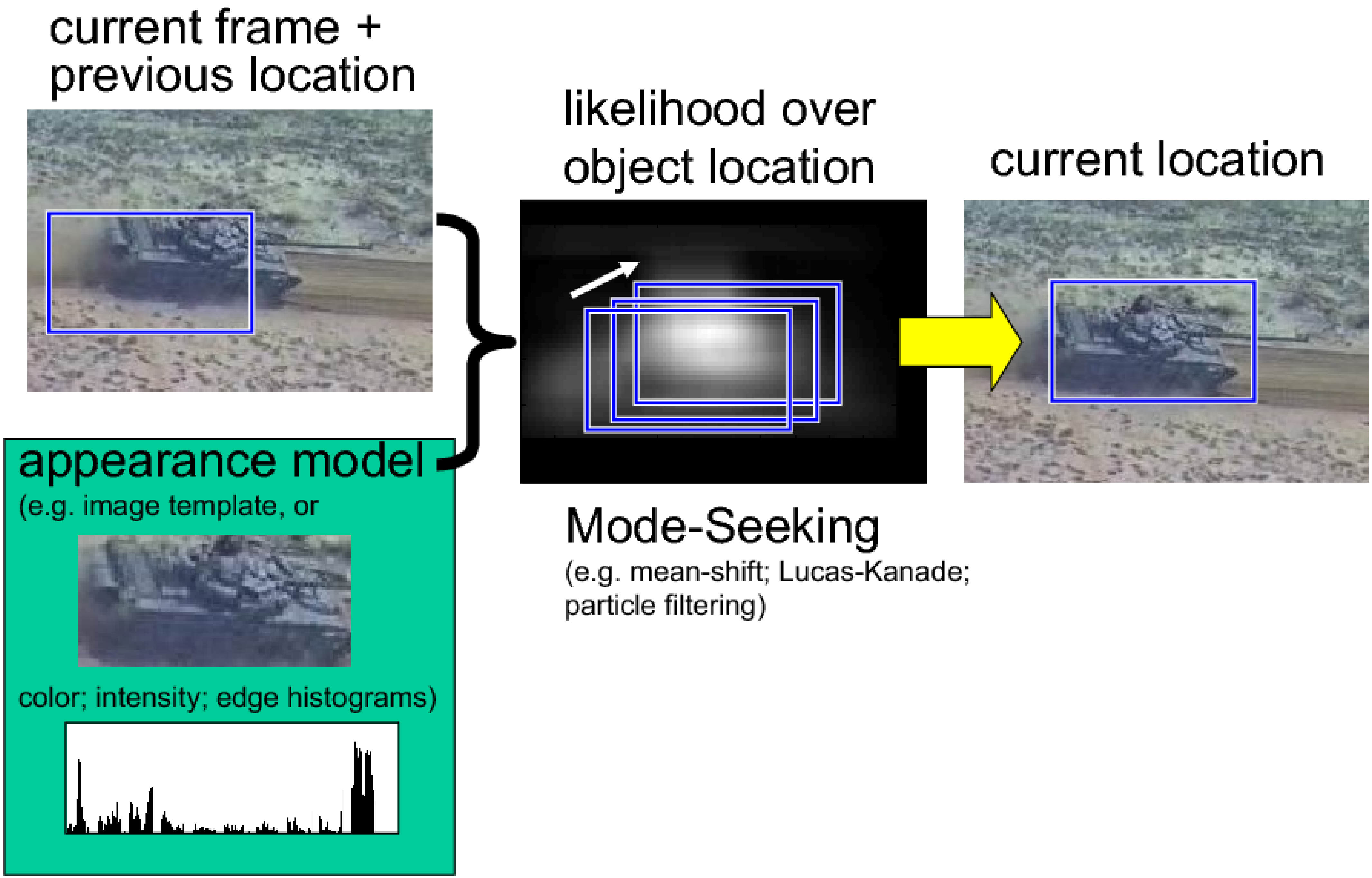
[2] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Analysis, 24(5):603–619, May 2002.

[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


current frame +
previous location
appearance model
(e.g. image template, or
color; intensity; edge histograms)
likelihood over
object location
Mode-Seeking
(e.g. mean-shift; Lucas-Kanade;
particle filtering)
current location

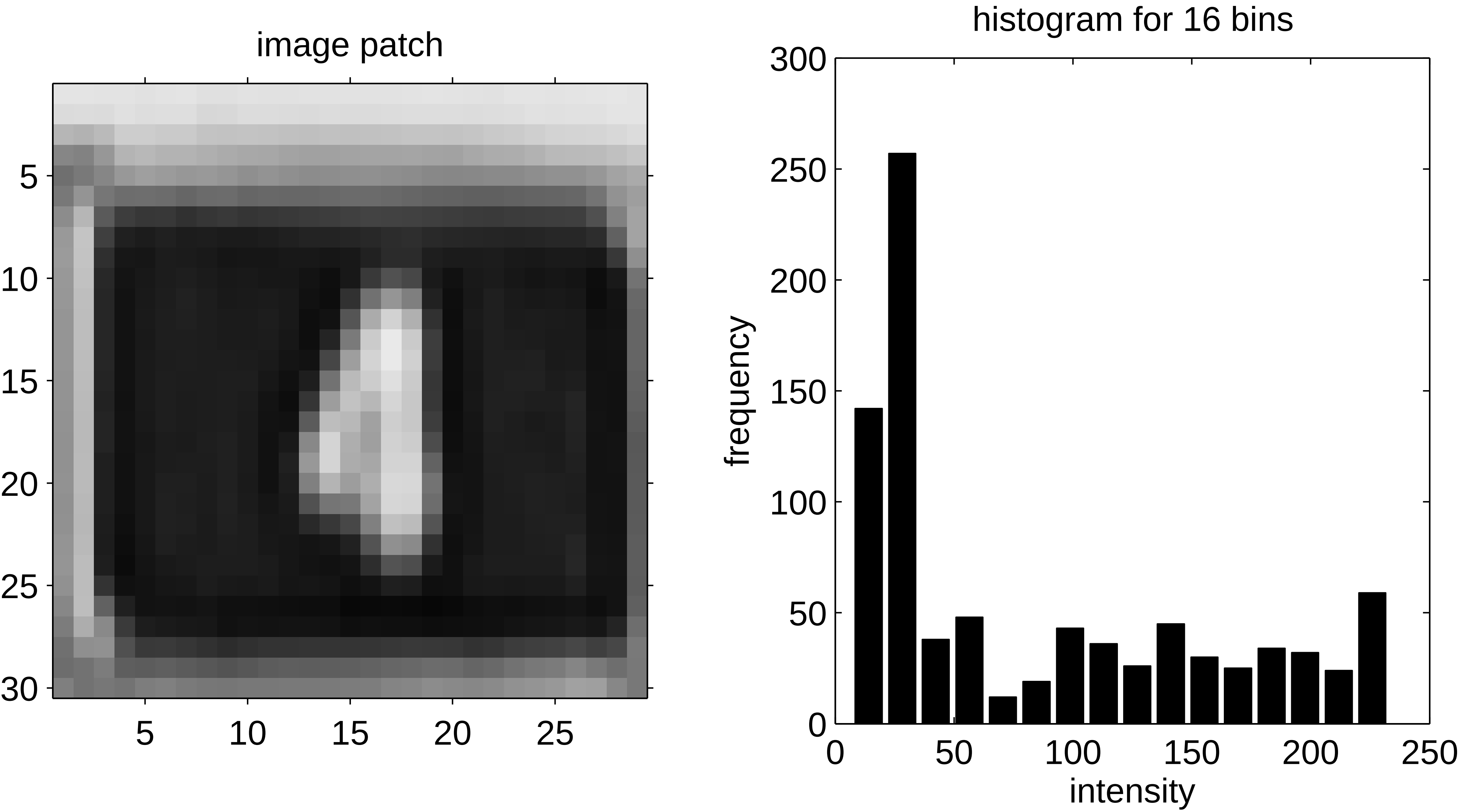


image patch
5
10
15
20
25
30
histogram for 16 bins
frequency
intensity
0
50
100
150
200
250
300

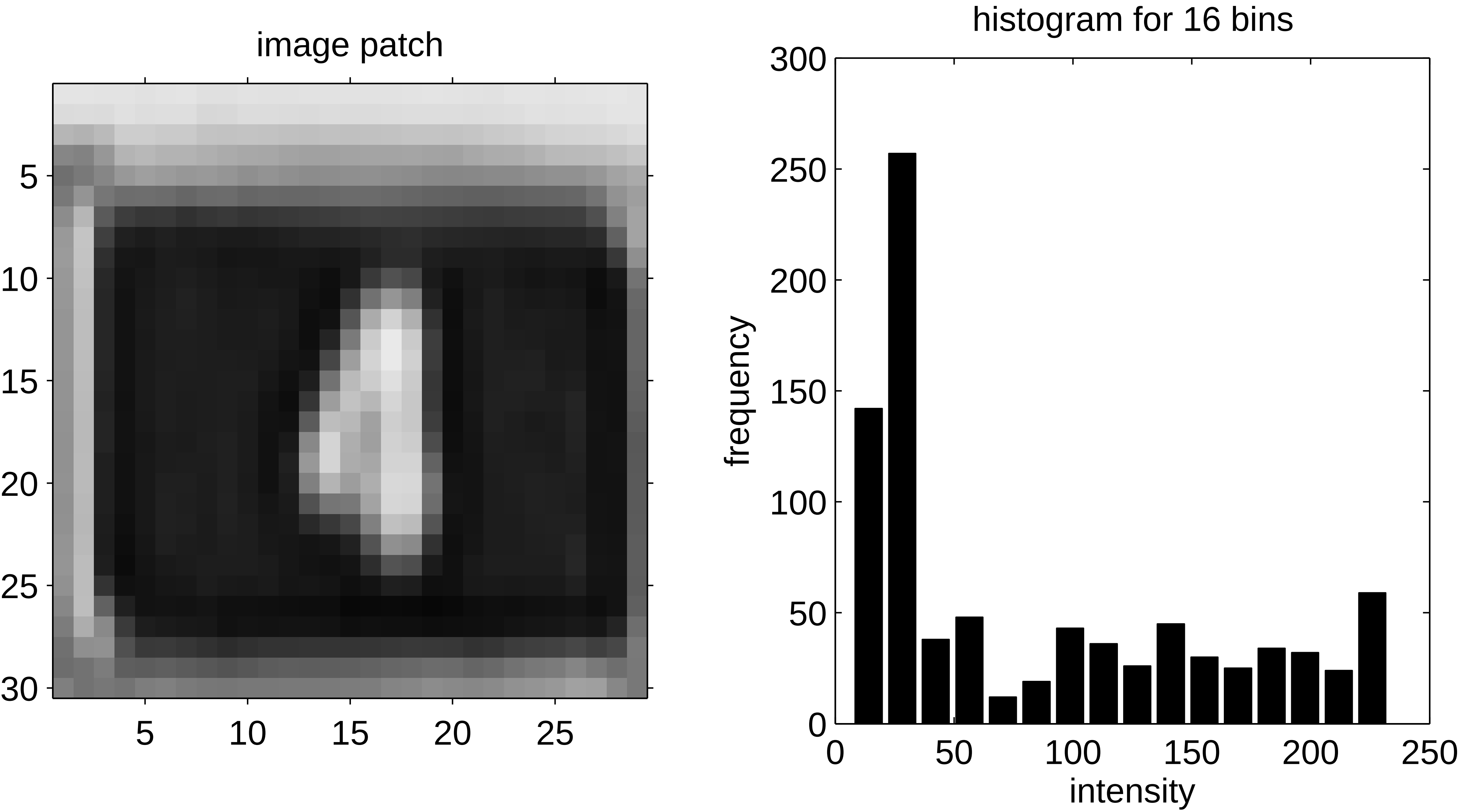


image patch
5
10
15
20
25
30
histogram for 16 bins
300
250
200
150
100
50
0
frequency
intensity

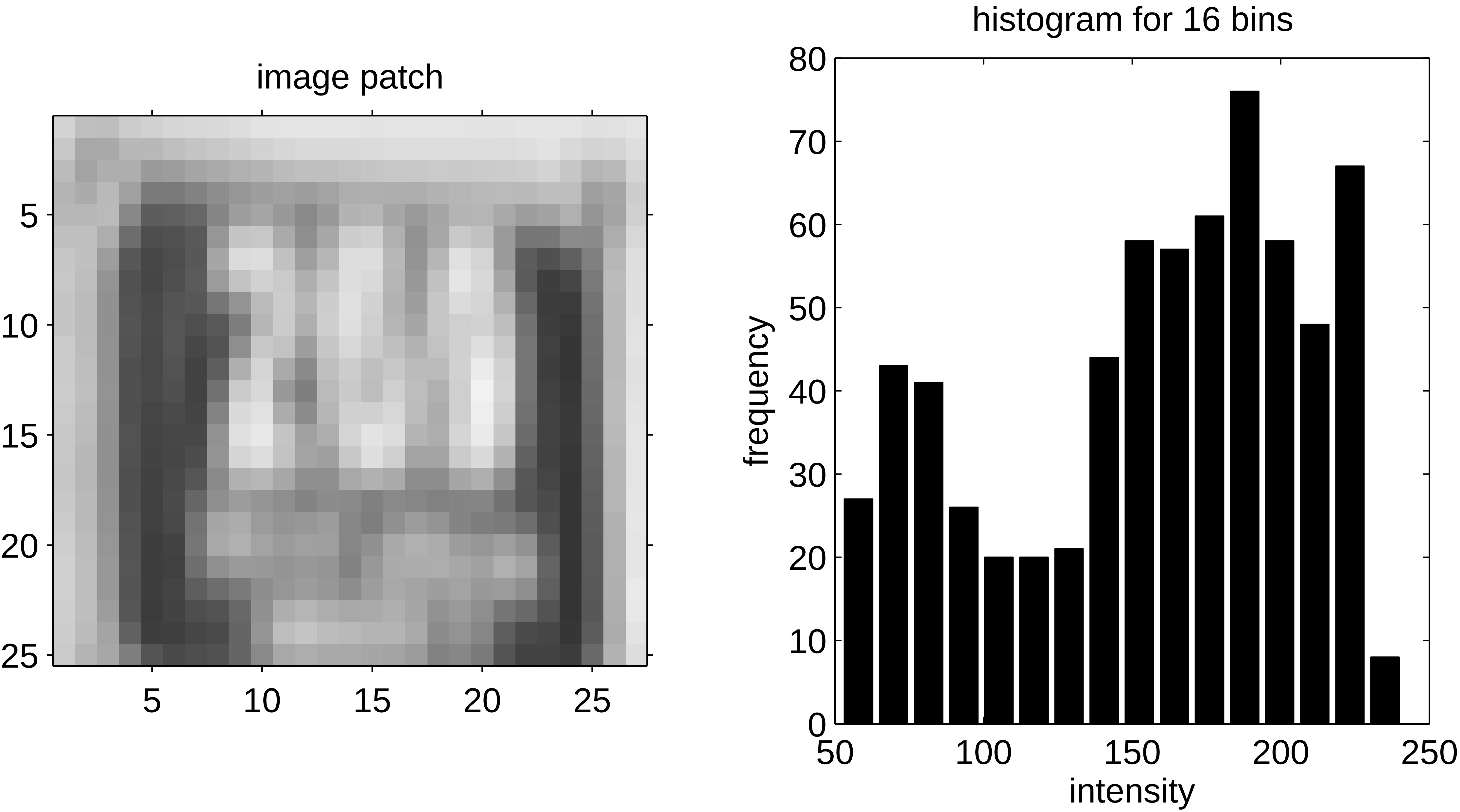


image patch
5
10
15
20
25
histogram for 16 bins
frequency
intensity
0
10
20
30
40
50
60
70
80
50
100
150
200
250

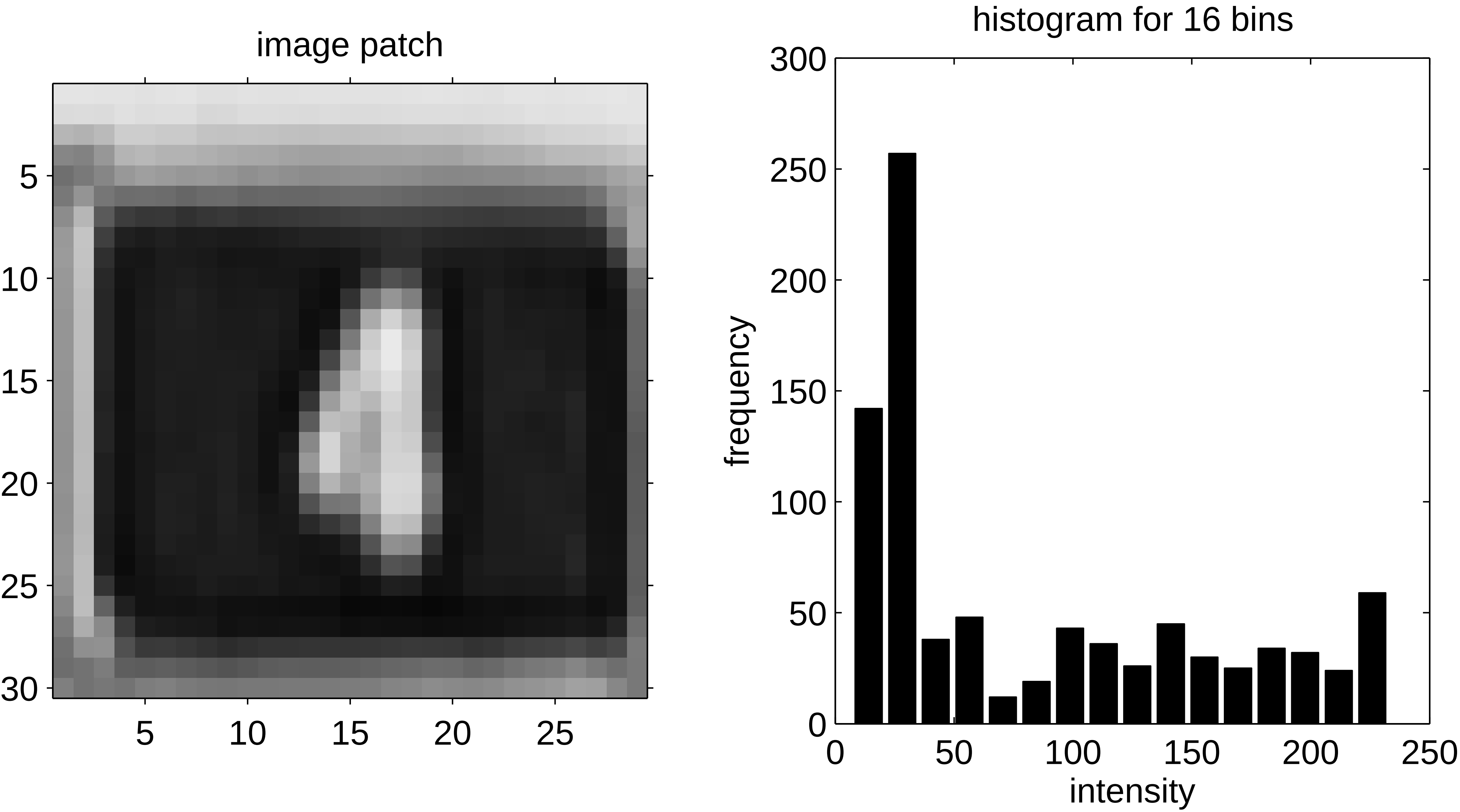


image patch
5
10
15
20
25
30
histogram for 16 bins
frequency
intensity
0
50
100
150
200
250
300

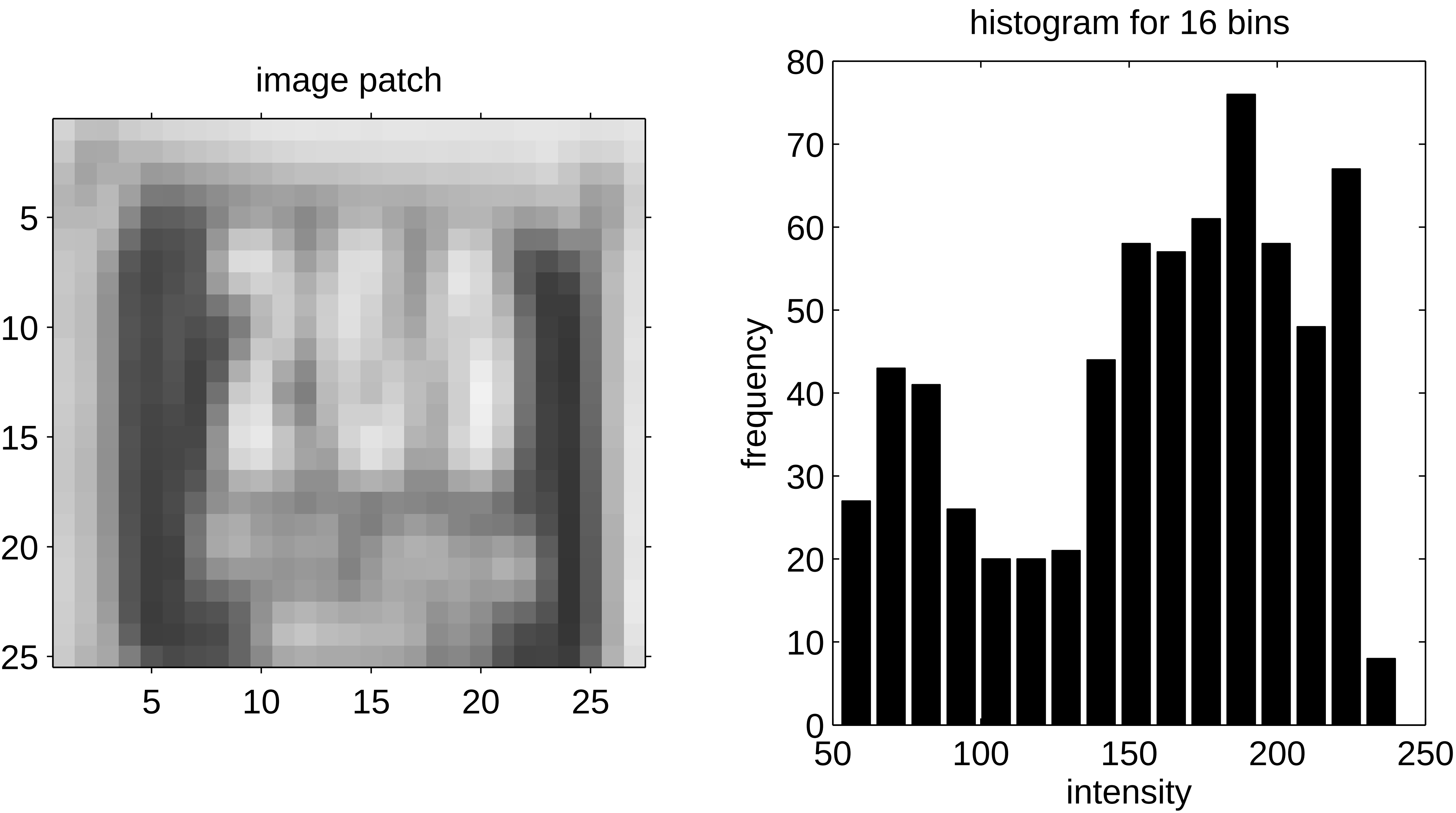
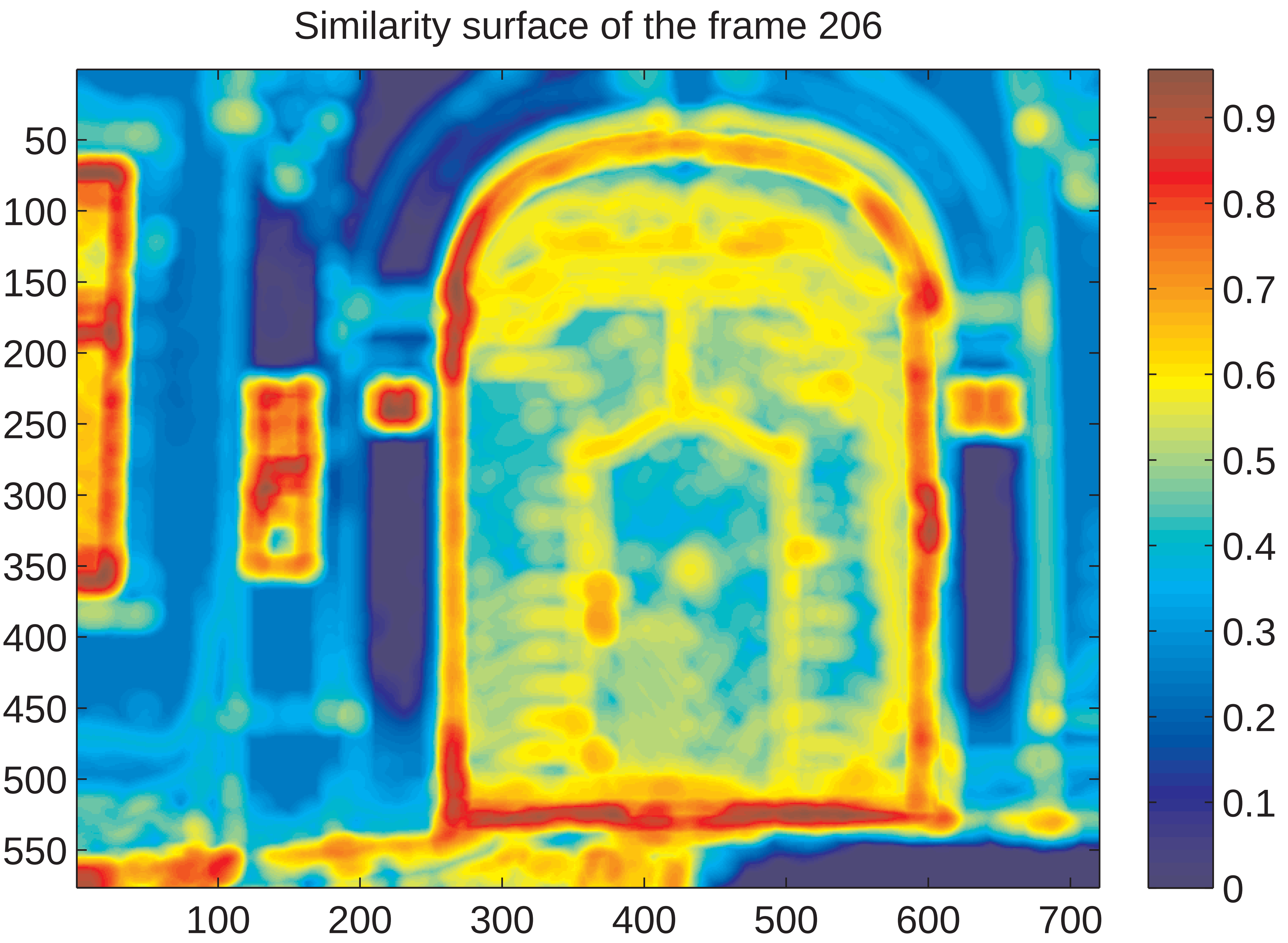
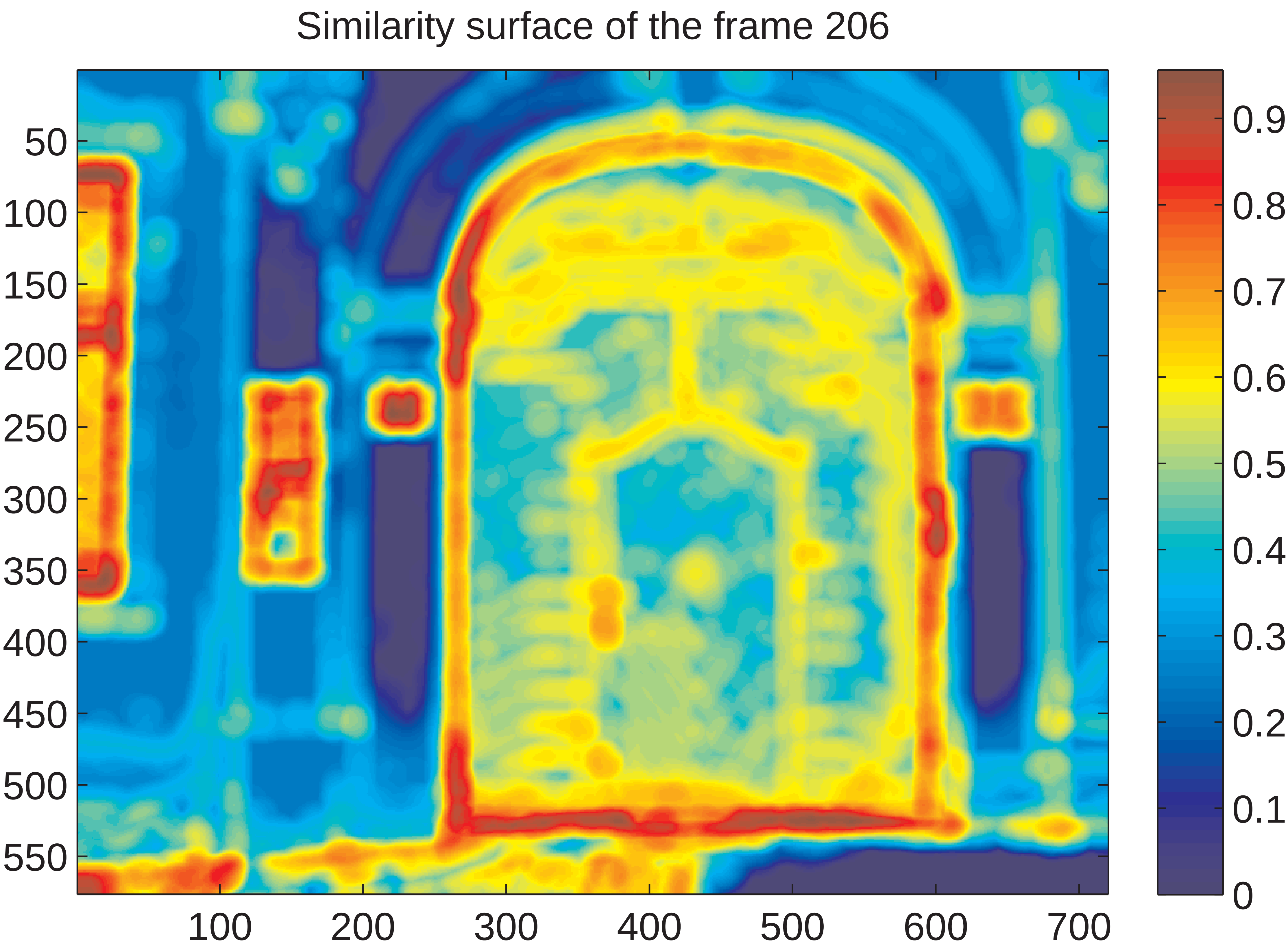


image patch
5
10
15
20
25
histogram for 16 bins
frequency
intensity
0
10
20
30
40
50
60
70
80
50
100
150
200
250

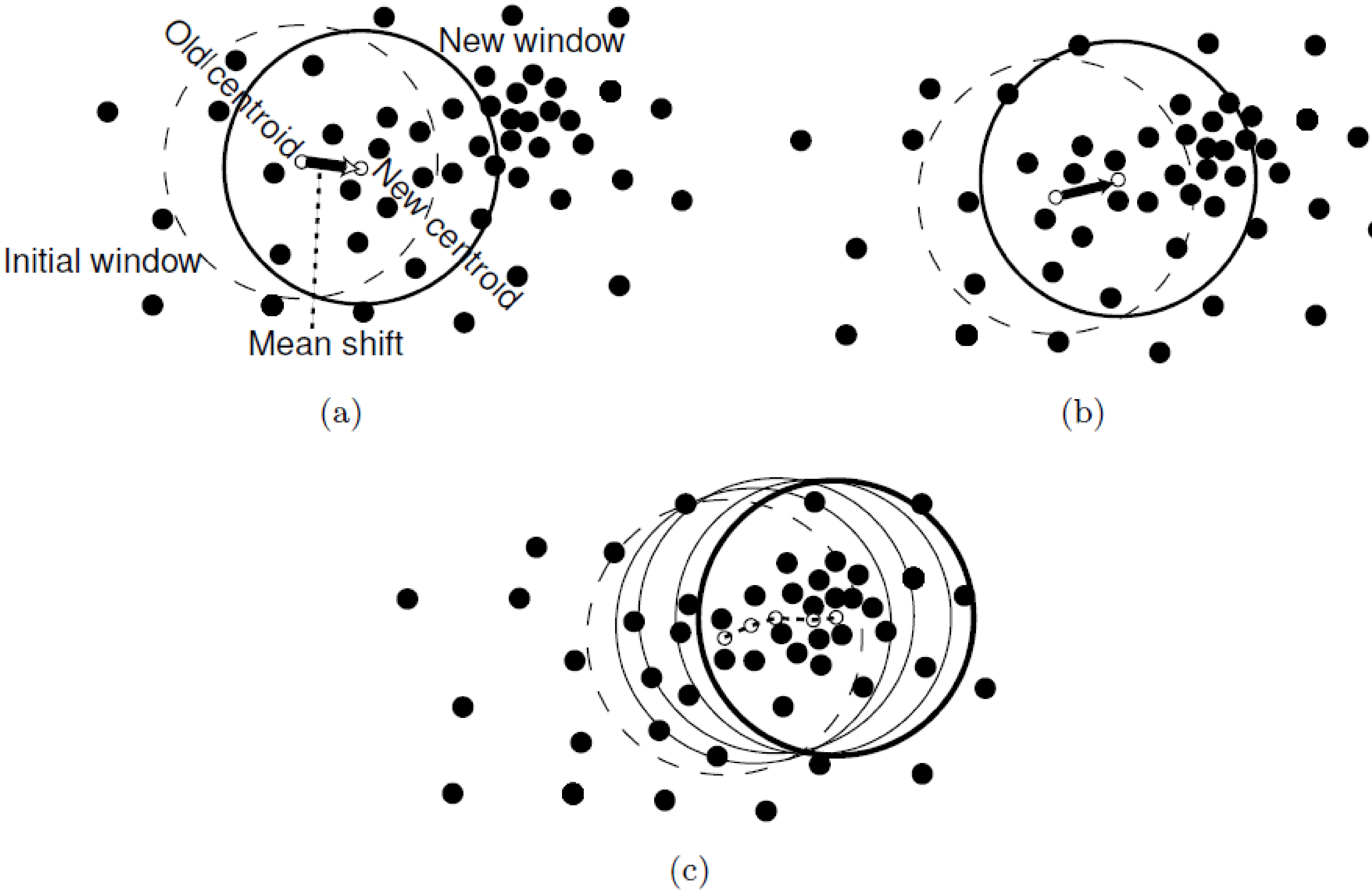
Example of an input image


Similarity surface of the frame 206
50
100
150
200
250
300
350
400
450
500
550
100
200
300
400
500
600
700
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

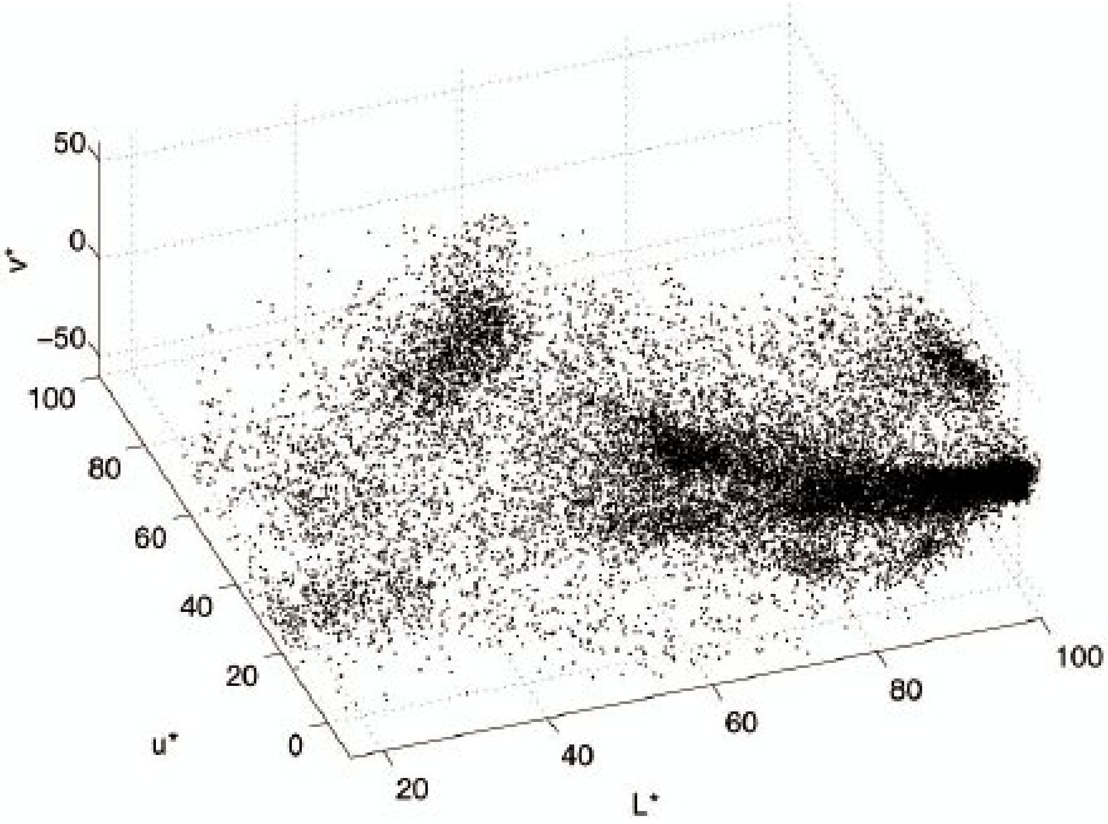
Example of an input image

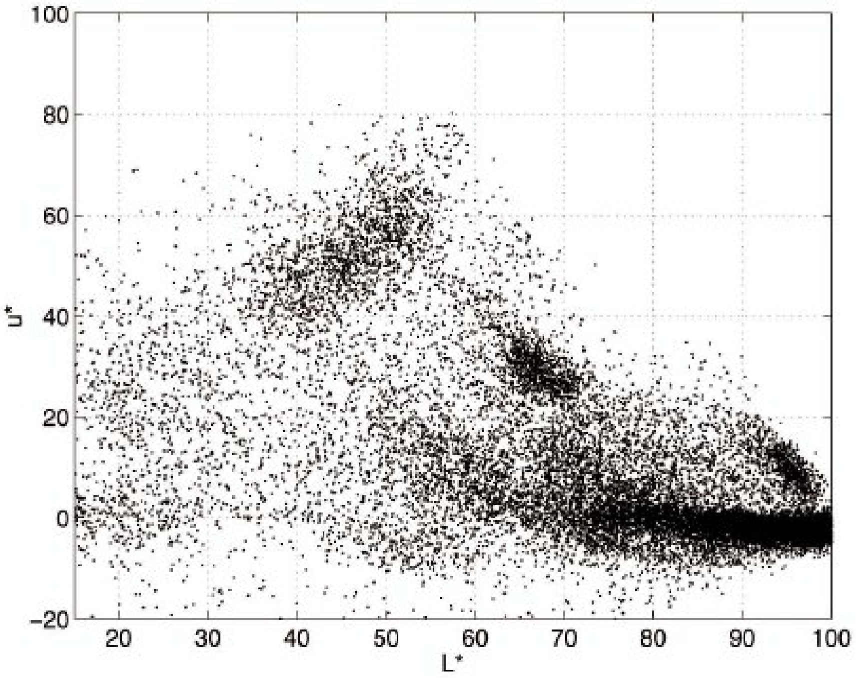


Similarity surface of the frame 206
50
100
150
200
250
300
350
400
450
500
550
100
200
300
400
500
600
700
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0



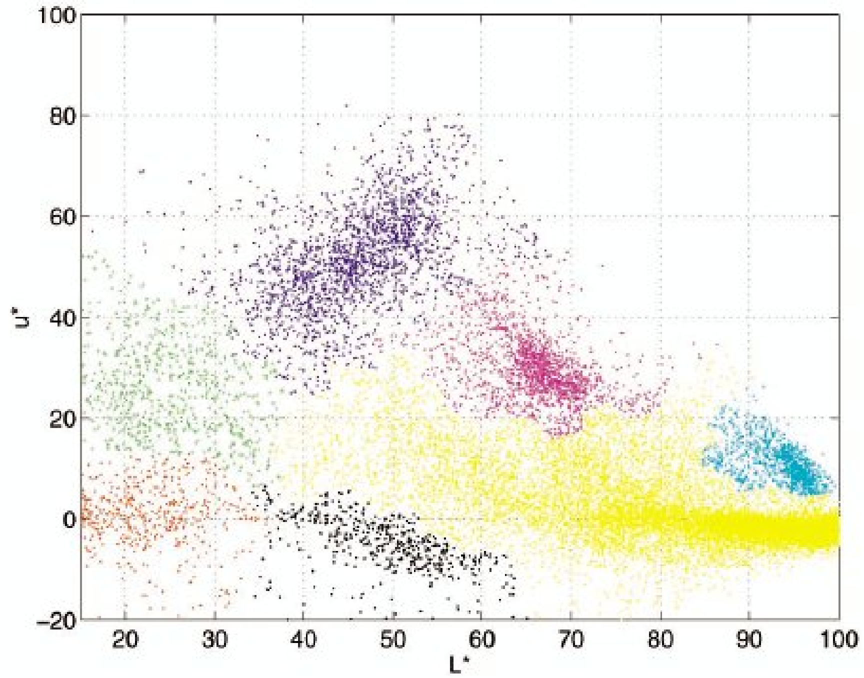
New window
Old centroid
New centroid
Initial window
Mean shift
(a)
(b)
(c)


50
0
-50
v*
100
80
60
40
20
0
u*
20
40
60
80
100
L*



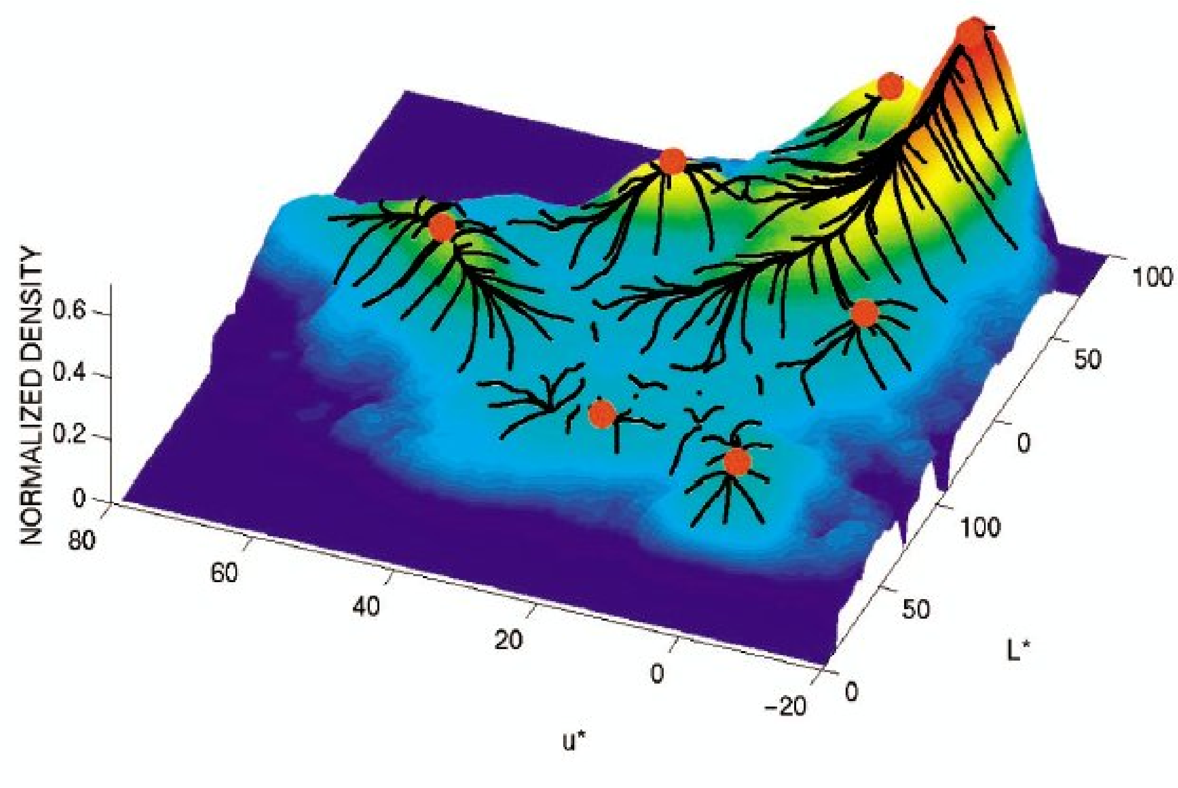
u*
L*
100
80
60
40
20
0
−20
20
30
40
50
60
70
80
90
100

(a)



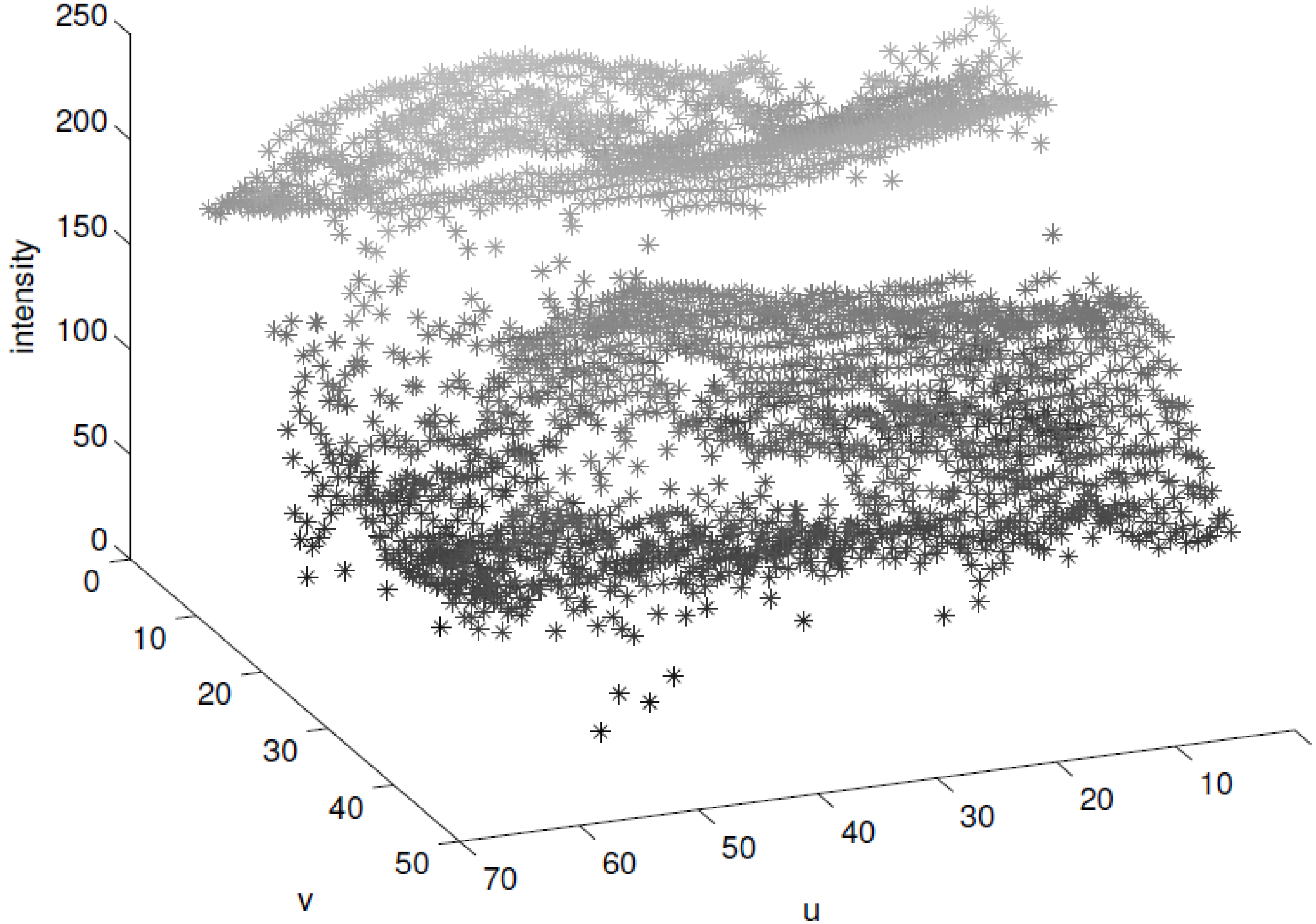
u*
L*
100
80
60
40
20
0
−20
20
30
40
50
60
70
80
90
100

(b)



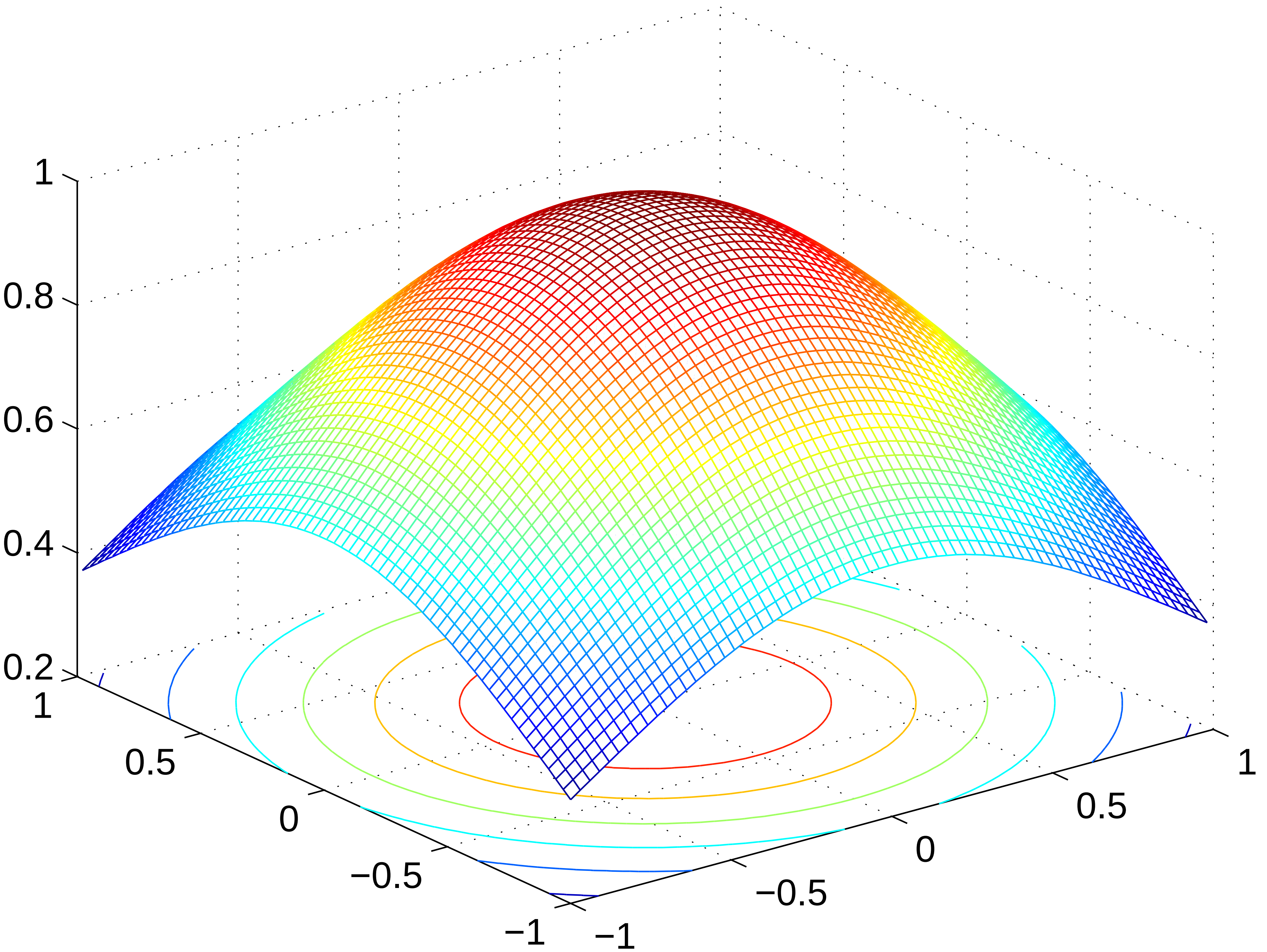
NORMALIZED DENSITY
0.6
0.4
0.2
0
80
60
40
20
0
−20
u*
100
50
0
100
50
0
L*

(c)

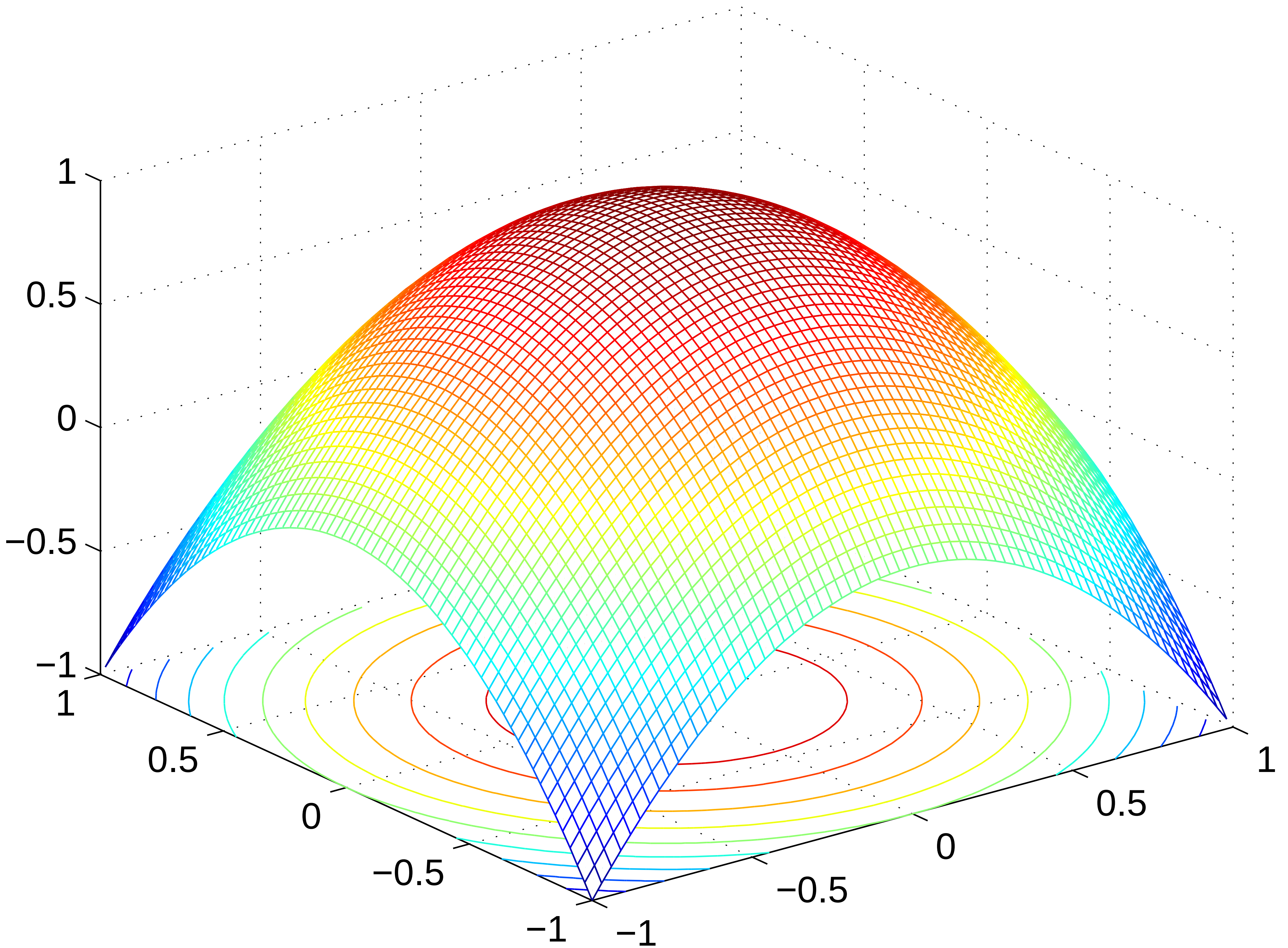


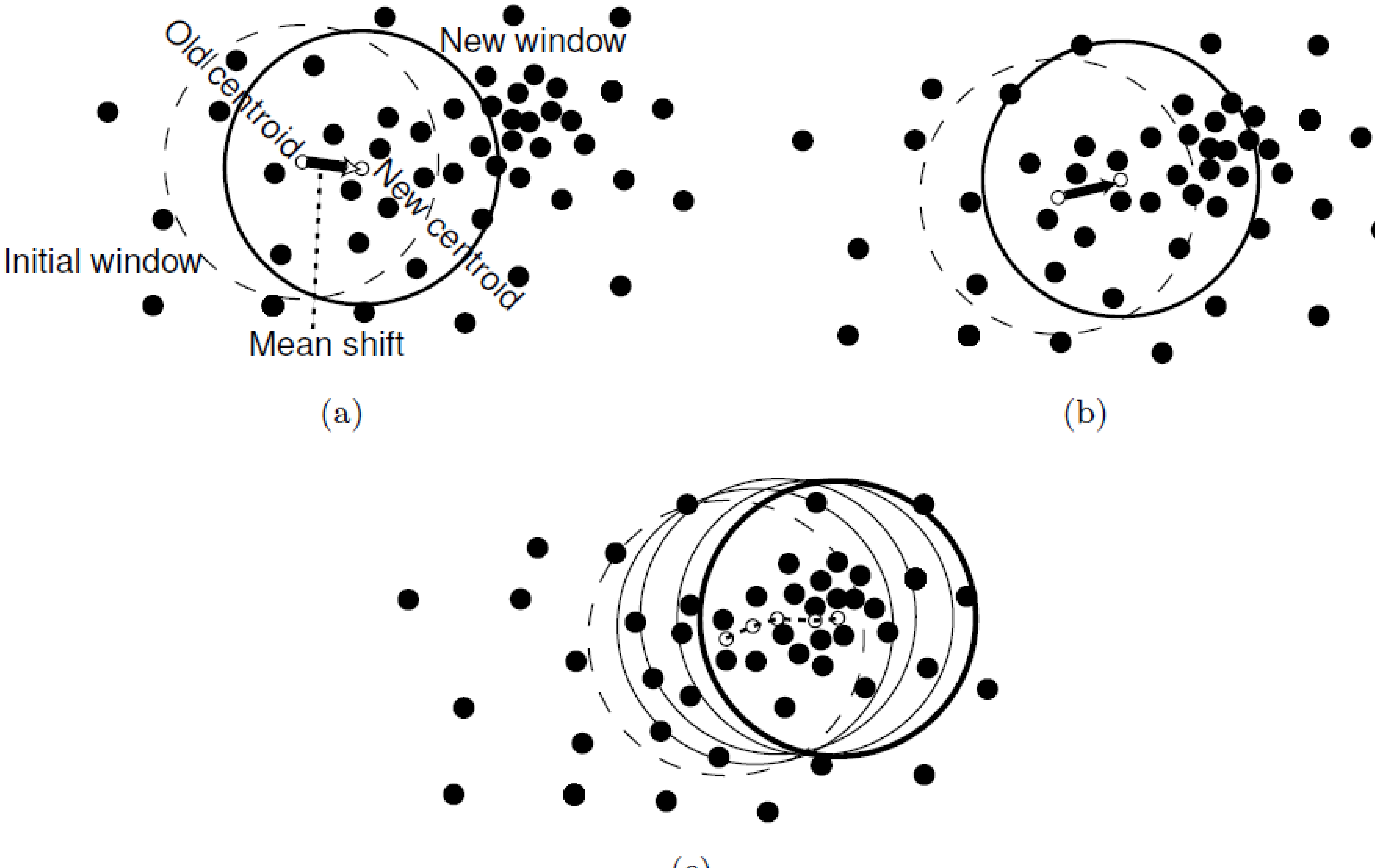
intensity
250
200
150
100
50
0
0
10
20
30
40
50
v
70
60
50
40
30
20
10
0
u

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

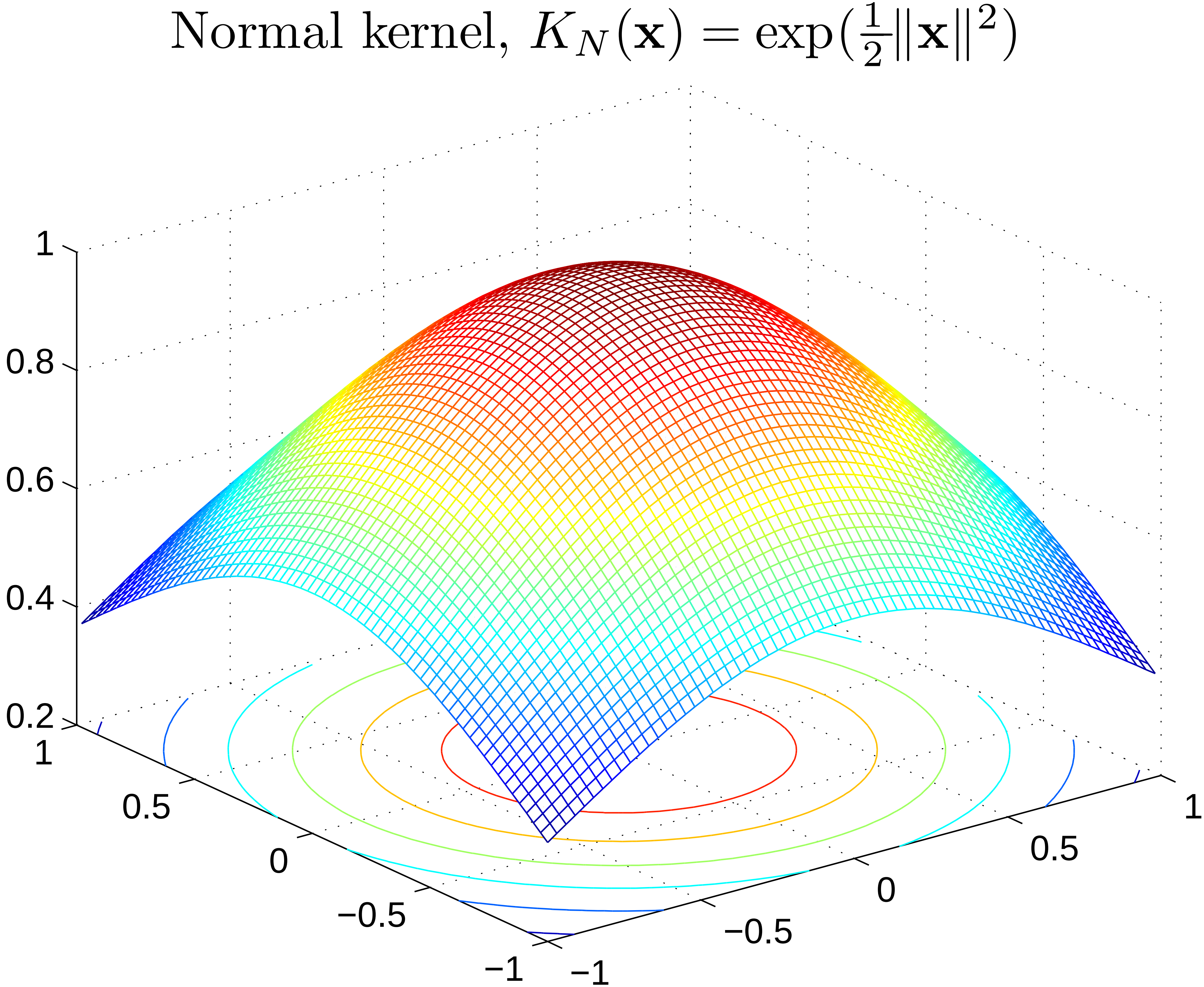


Epanechnikov kernel, $K_E(\mathbf{x}) = 1 - \|\mathbf{x}\|^2$



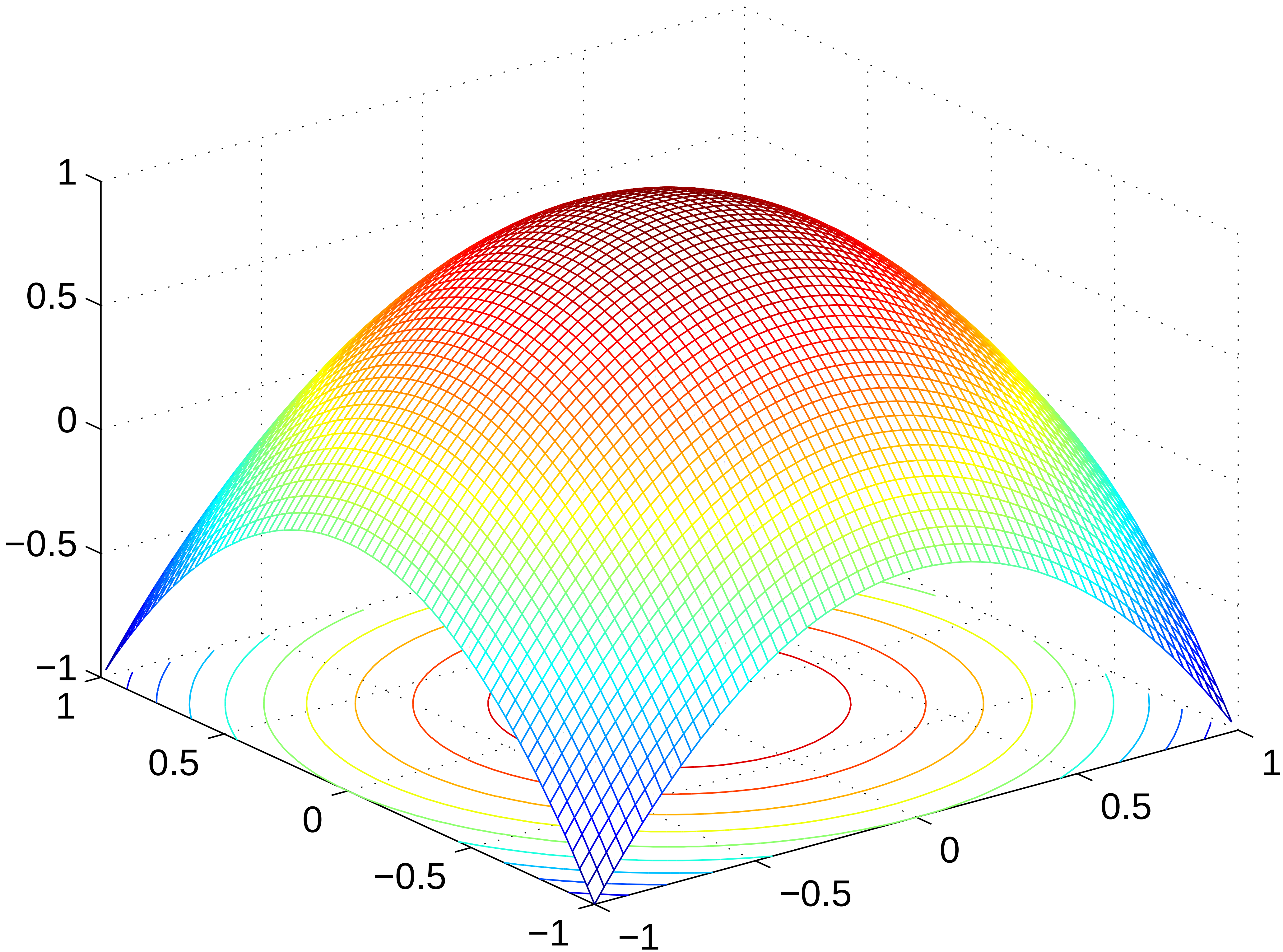


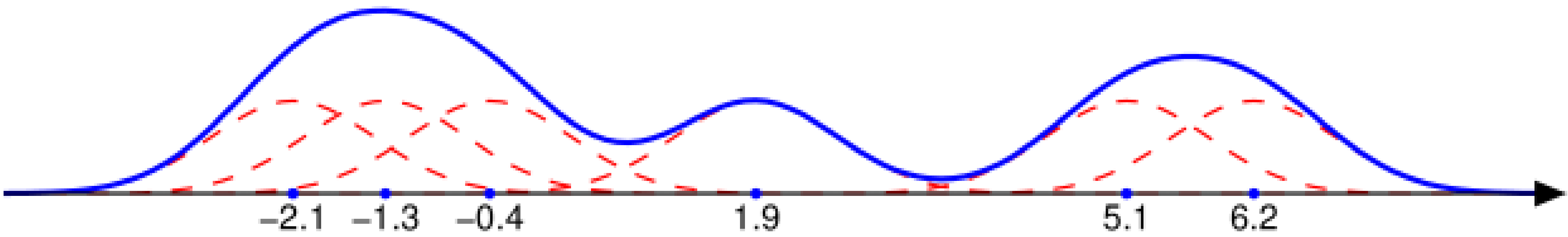
New window
Old centroid
New centroid
Initial window
Mean shift
(a)
(b)
(c)

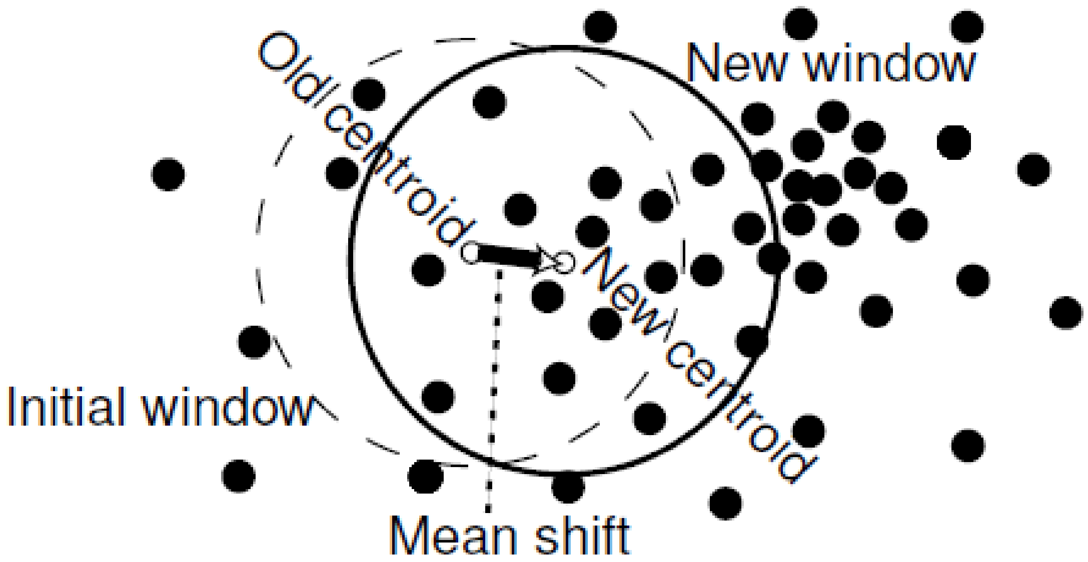


Normal kernel, $K_N(\mathbf{x}) = \exp(\frac{1}{2}\|\mathbf{x}\|^2)$
1
0.8
0.6
0.4
0.2
1
0.5
0
−0.5
−1
−1
−0.5
0
0.5
1

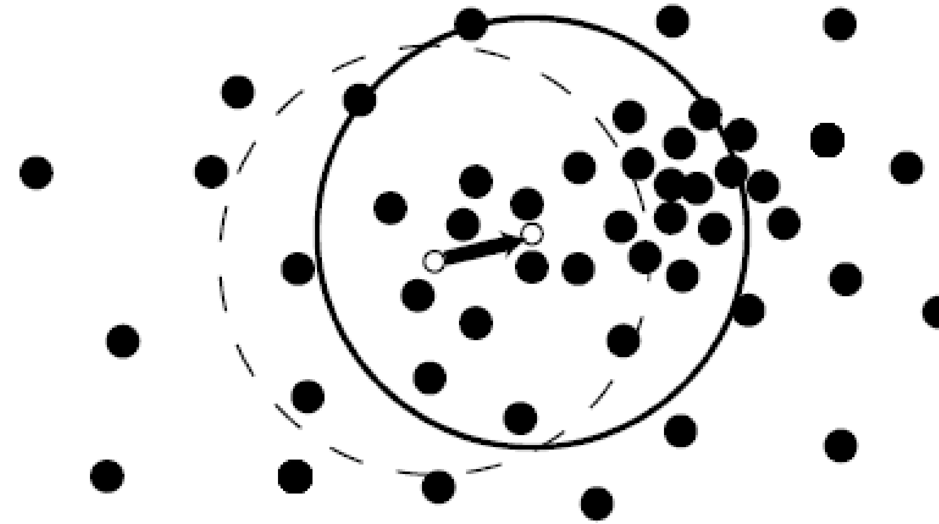
Epanechnikov kernel, $K_E(\mathbf{x}) = 1 - \|\mathbf{x}\|^2$



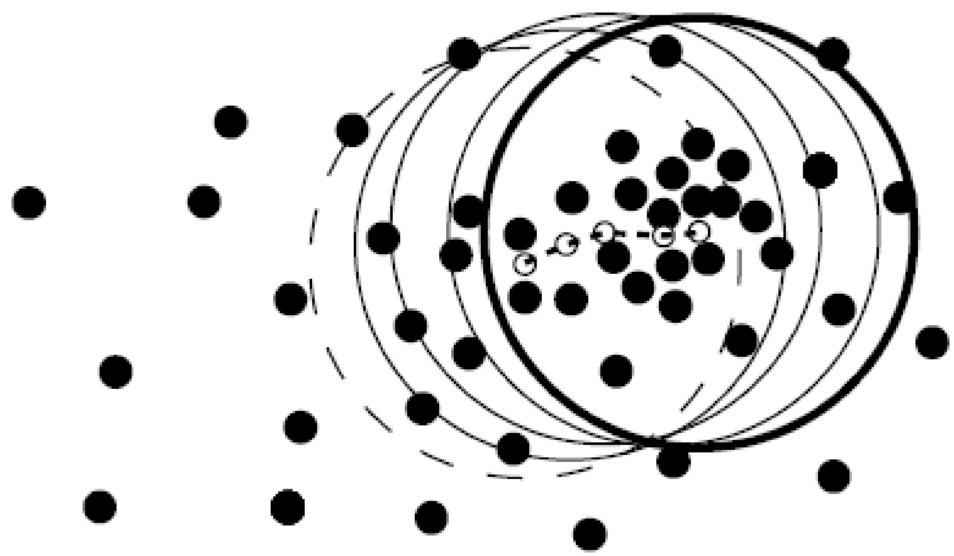

−2.1
−1.3
−0.4
1.9
5.1
6.2


New window
Old centroid
New centroid
Initial window
Mean shift

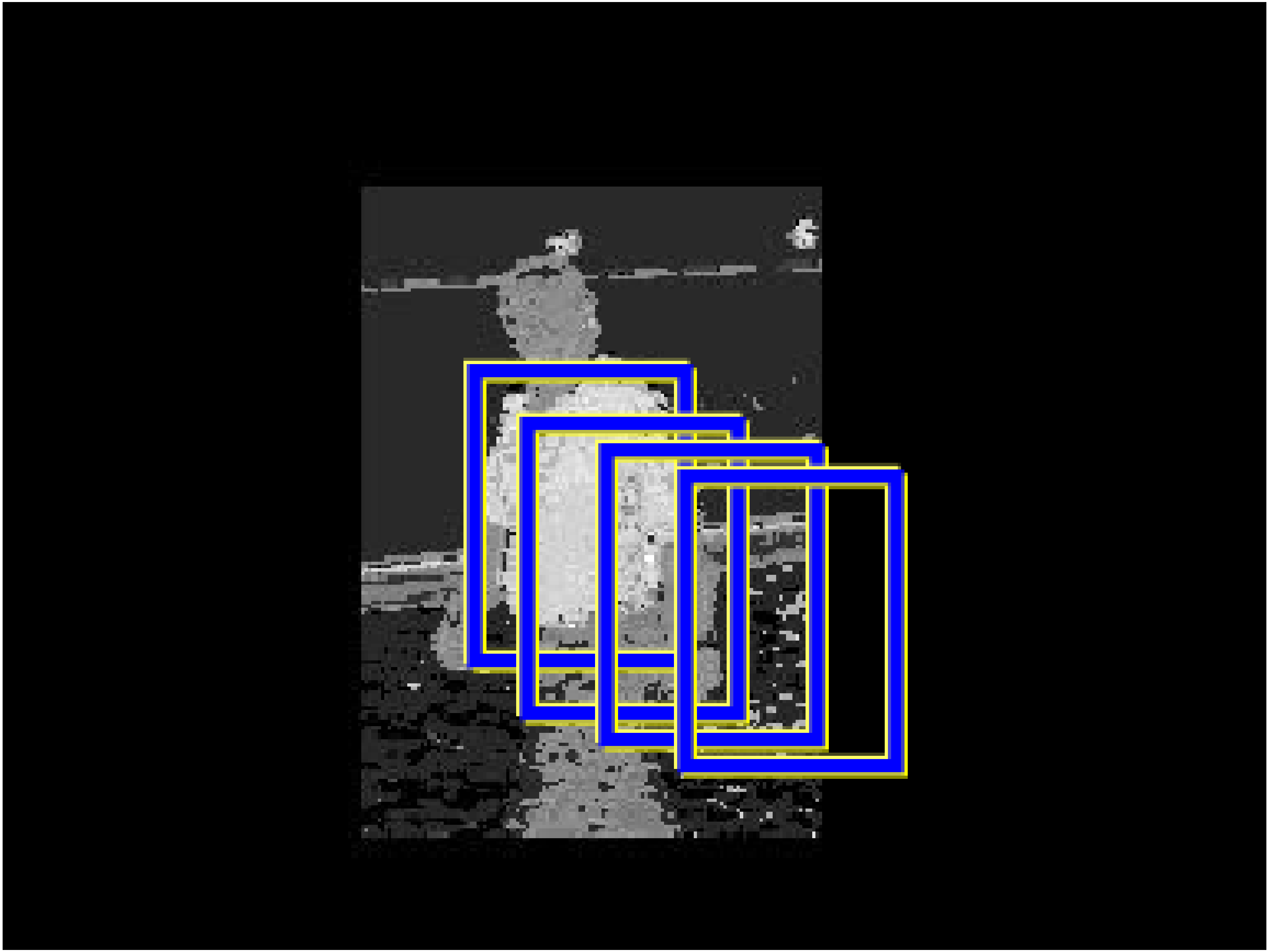
(a)

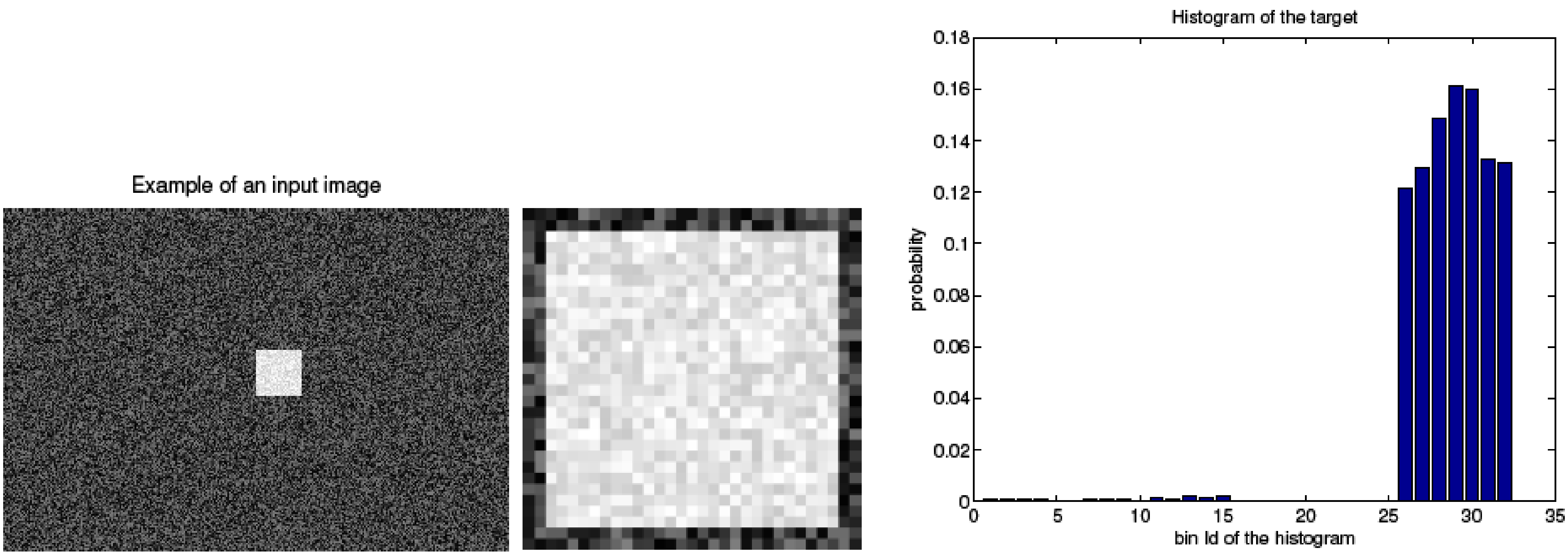


(b)

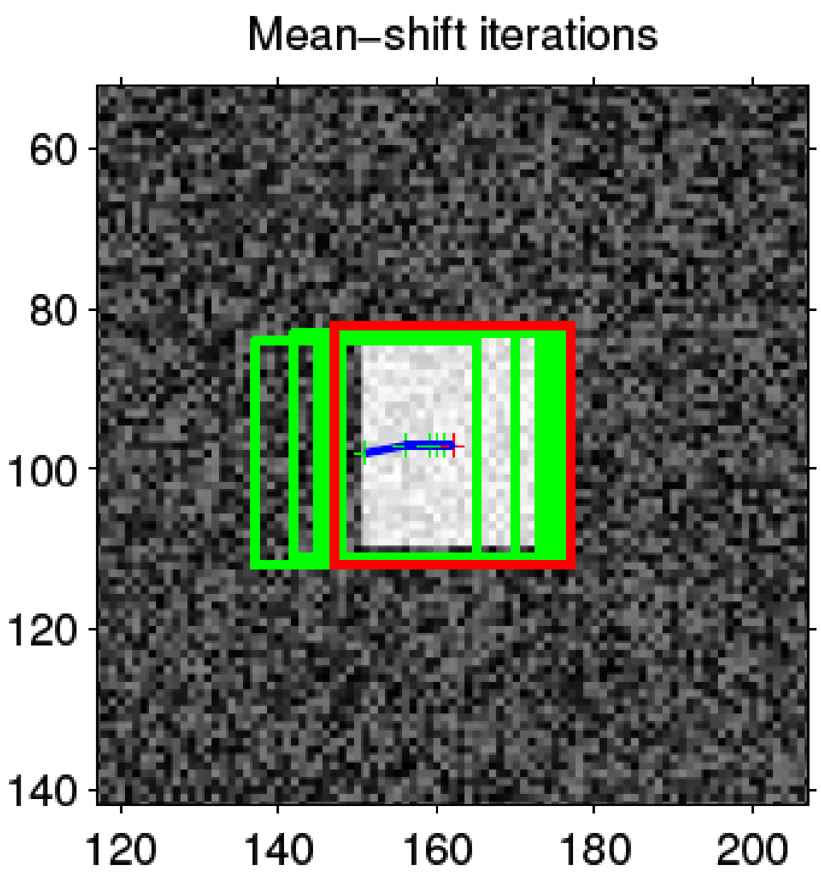


(c)

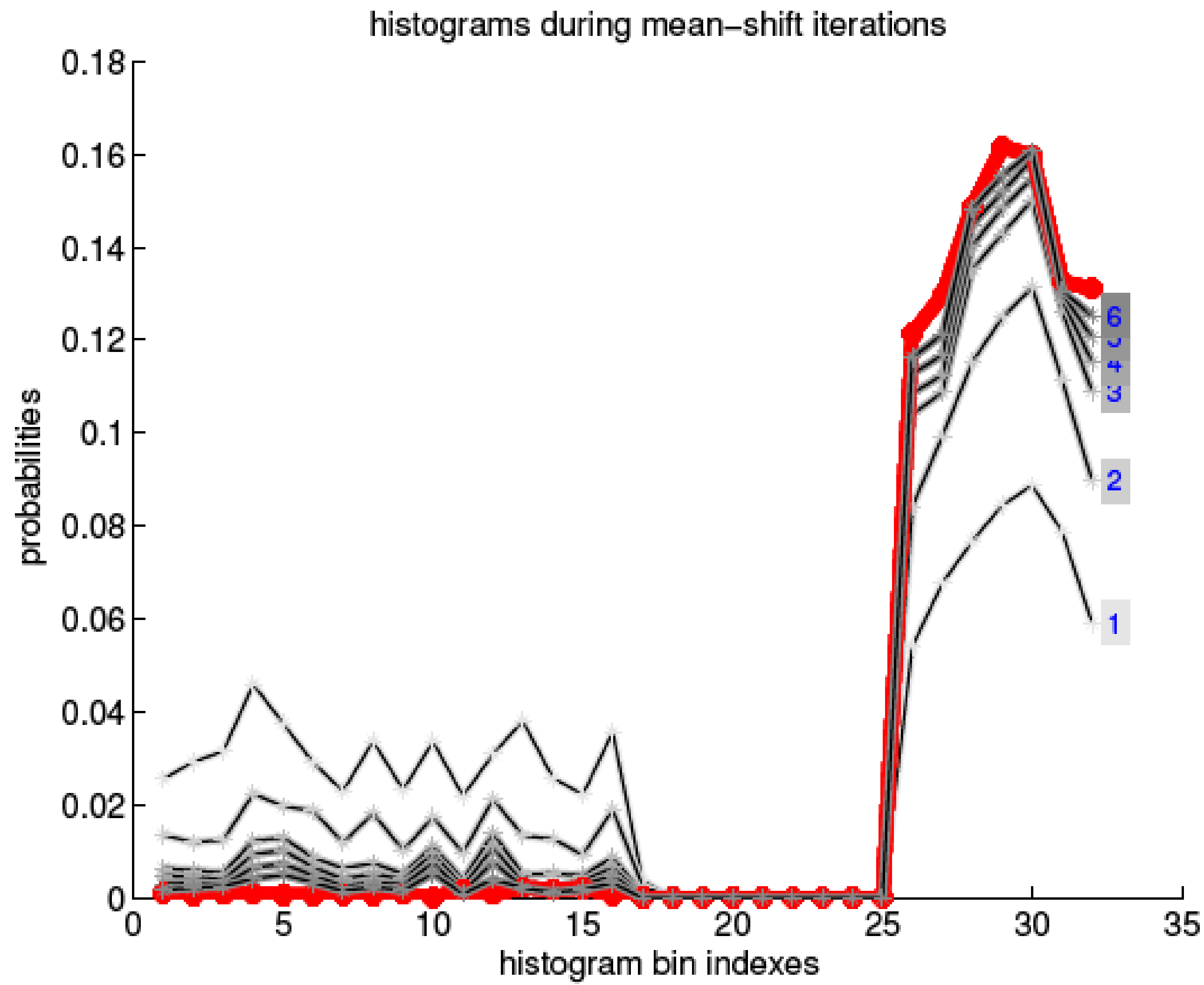



Example of an input image
Histogram of the target
probability
bin Id of the histogram
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18
0 5 10 15 20 25 30 35

(a) (b) (c)


Mean−shift iterations
60
80
100
120
140
120
140
160
180
200

(a)


histograms during mean−shift iterations
probabilities
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
20
25
30
35
histogram bin indexes
6
5
4
3
2
1

(b)