Digital image, basic concepts

Václav Hlaváč

Czech Technical University in Prague

Czech Institute of Informatics, Robotics and Cybernetics

160 00 Prague 6, Jugoslávských partyzánů 1580/3, Czech Republic

http://people.ciirc.cvut.cz/hlavac, vaclav.hlavac@cvut.cz

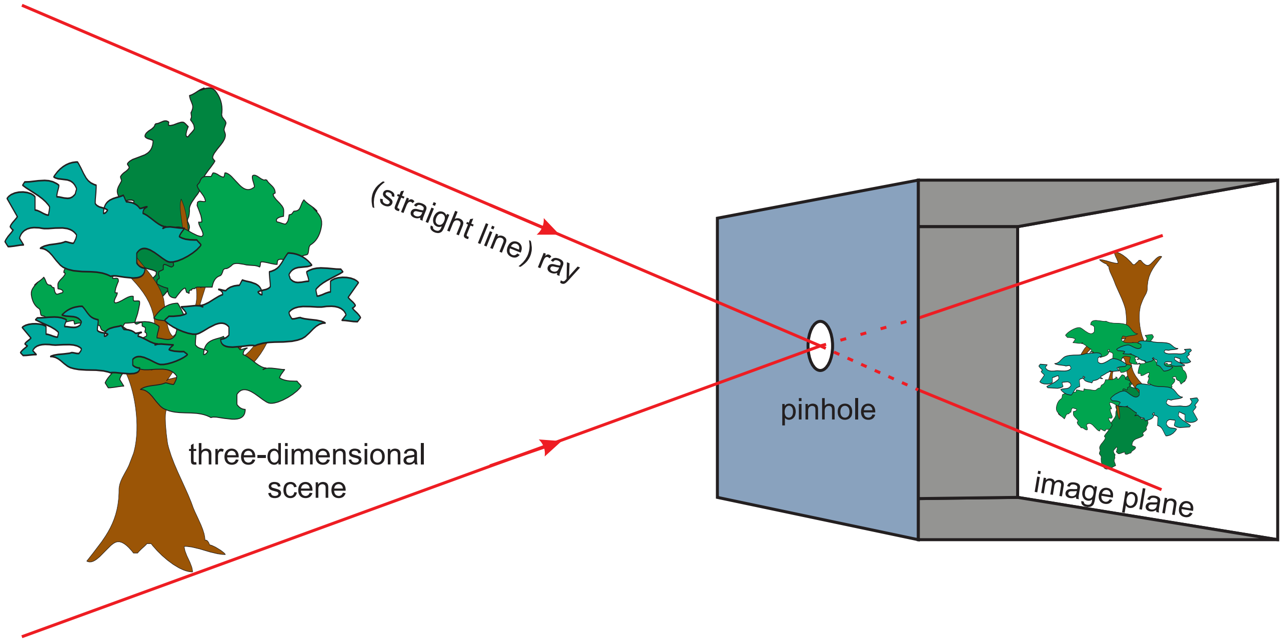
also Center for Machine Perception, http://cmp.felk.cvut.cz

Outline of the lecture:

- Image, perspective imaging.
- Image function $f(x, y)$.
- Image digitization: sampling $+$ quantization.
- Distance in the image, neighborhood.
- Contiguity relation, region, convex region.
- Distance transformation.
- The entire image properties: brightness histogram, brightness, contrast, sharpness.

Image

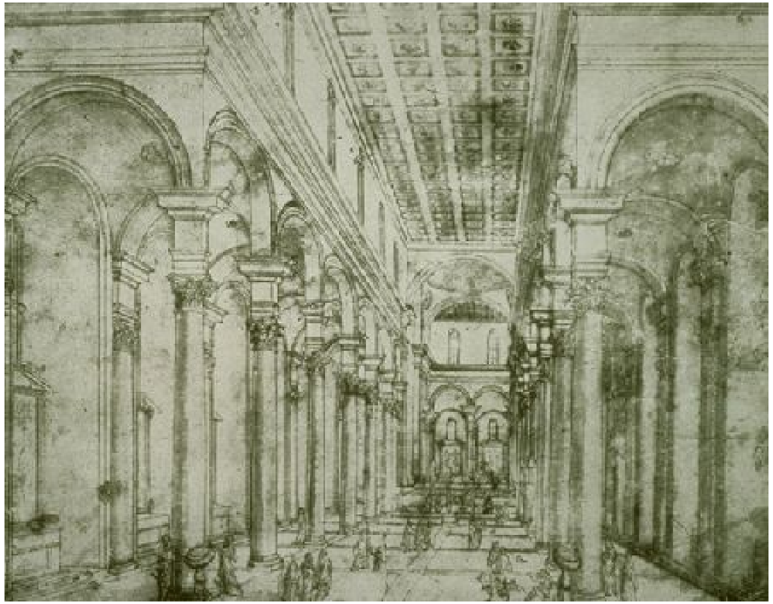
- The **image** is understood intuitively as the visual response on the retina or light sensitive chip in a camera, TV camera, . . .
- The image is often formed by a **perspective projection** corresponding to the intuitive pinhole model.



Convention:
We consider left to right direction of light.

Perspective projection started in Italian Renaissance painting

Filippo Brunelleschi created a perspective drawing to show his customers how the Church of Santo Spirito in Florence would look like.

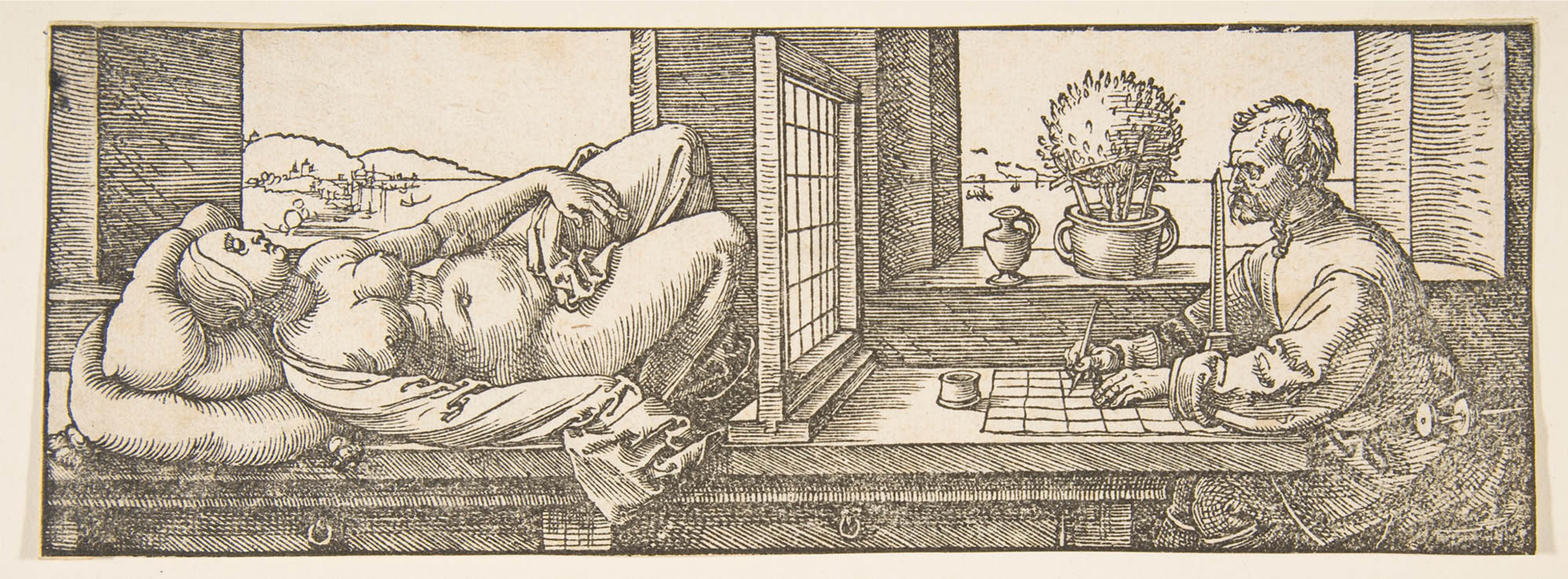


Drawing in $\approx$1420

View into the church built in 1434-82

Courtesy: pictures Khan Academy

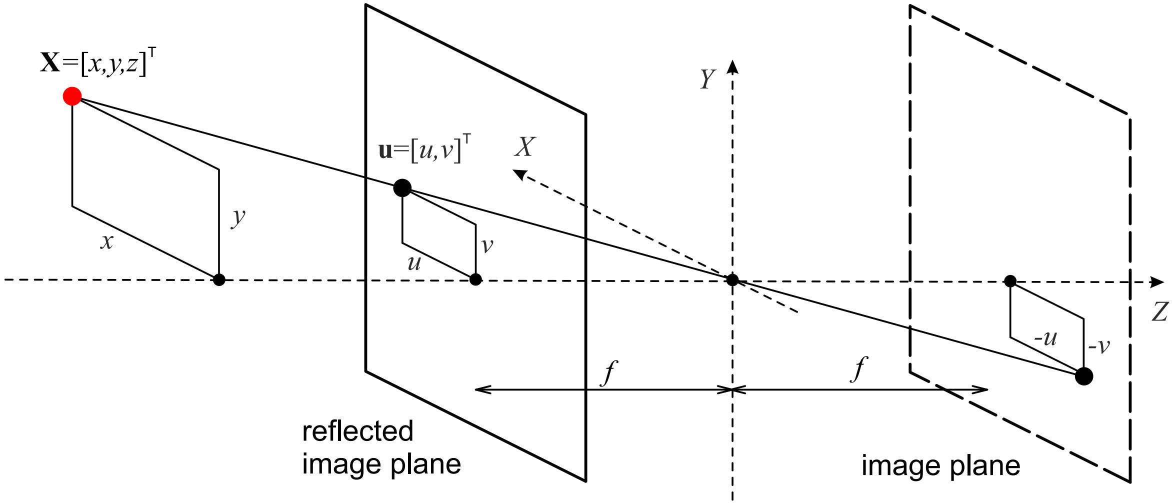
Example of the perspective drawing tool from the 16th century



Albrecht Dürer: Recumbent woman, wood engraving, 1525

Image function

- The image function is abstracted mathematically as $f(x, y)$, $f(x, y, t)$. It is the result of the perspective projection encompassing geometric aspects.



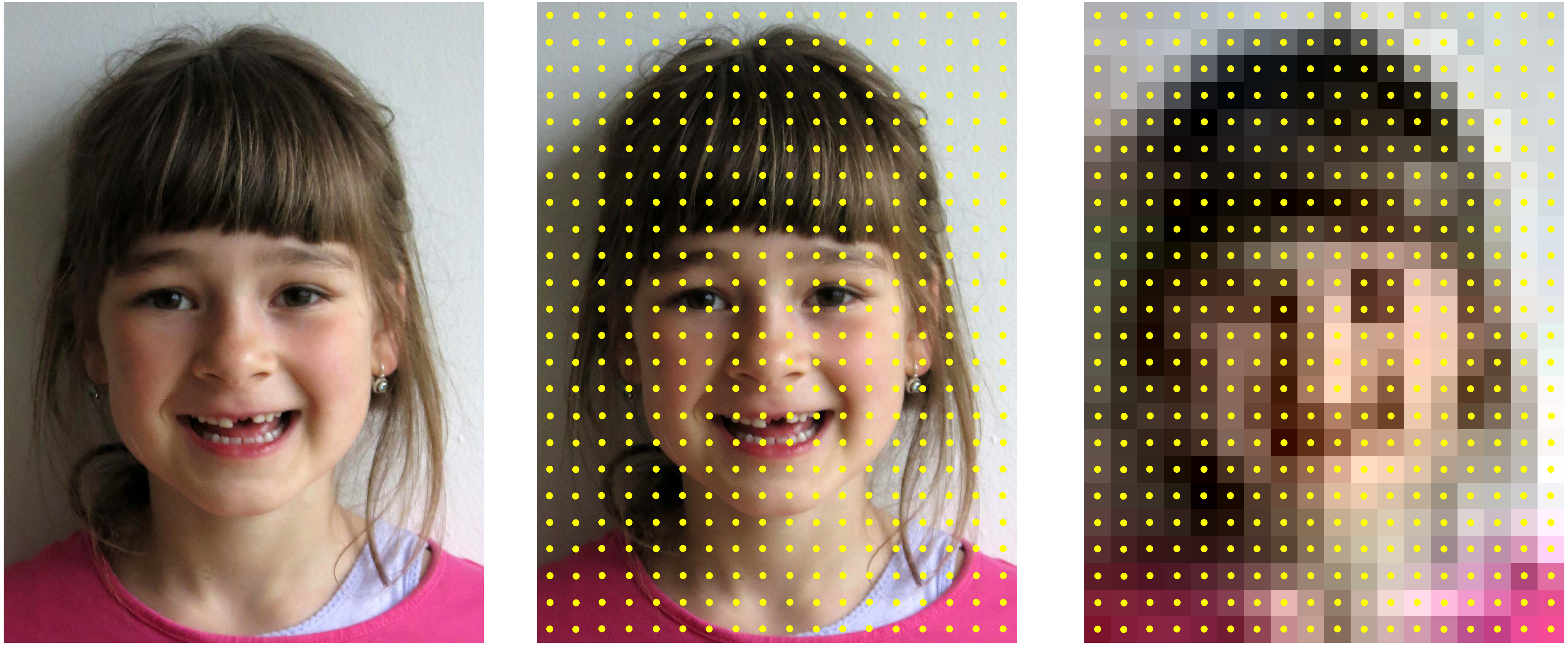
- Considering similar triangles: $u = \frac{x\,f}{z}$, $v = \frac{y\,f}{z}$. Instead of our derived 2D image function $f(u, v)$, it is usually denoted $f(x, y)$.
- The value of the image function matches color/intensity of a 3D point (a red dot in the figure above) in the scene, which is projected.

Continuous image and its mathematical representation

- (Continuous) image $=$ the input (understood intuitively), e.g., on the retina or captured by a TV camera.
- Let us assume a gray level image for simplicity.
- The continuous image function $f(x, y)$. *Later, after digitization, a matrix of picture elements, pixels.*
- (x, y) are spatial coordinates of a pixel.
- $f(x, y, t)$ in the case of an image sequence, t corresponds to time.
- $f(x, y)$ is the value of the image function usually proportional to brightness, optical density with transparent original, distance to the observer, temperature in termovision, etc.
- (Naturally) 2D images: A thin specimen in the optical microscope, an image of a letter (character) on a piece of paper, a fingerprint, one slice in the tomograph, etc.

A pixel corresponds to a sample

(not to a little square)



Digitization

- Digitization = sampling & quantization of the image function value (called also intensity).
  - Sampling selects samples from the continuous image function. The output is a finite number of are arranged in a discrete raster. The sample value is “continuous”, i.e. a real number.
  - Quantizing divides a real value of the sample into a finite number of values (also bins). It can be 256 gray-level values for gray-scale image.

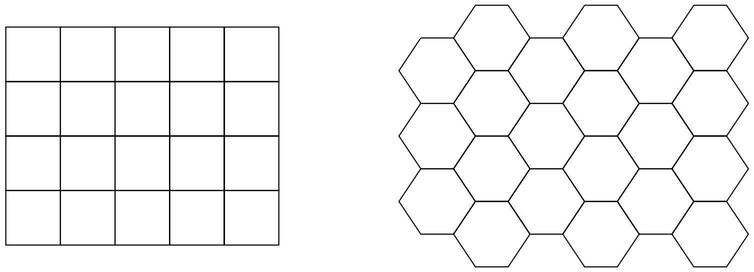
Example: *gray wedge* 6 *gray-level bins*

- Digital image is often represented as a matrix.
- Pixel = the acronym from the ‘picture element’.

Image sampling

Image sampling consists of two tasks:

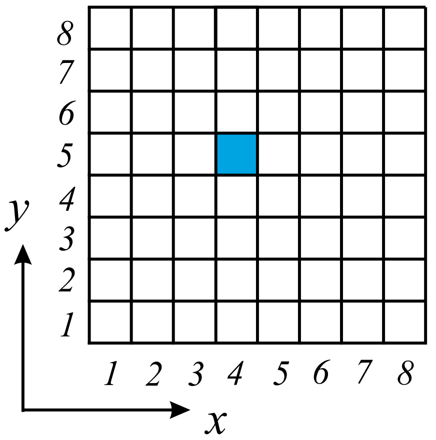
1. Sampling pattern (=arrangement of sampling points into a raster).

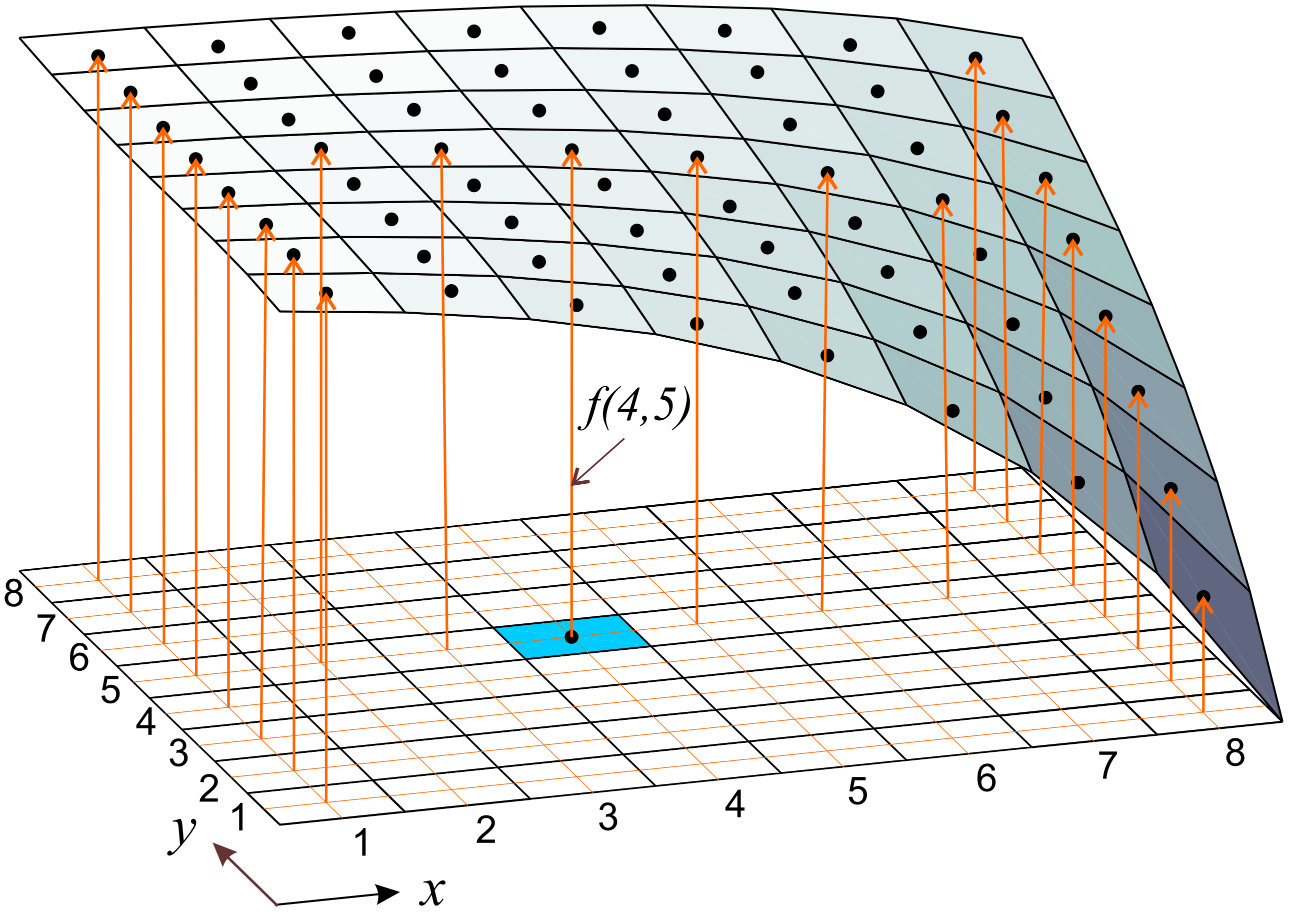


(a) (b)

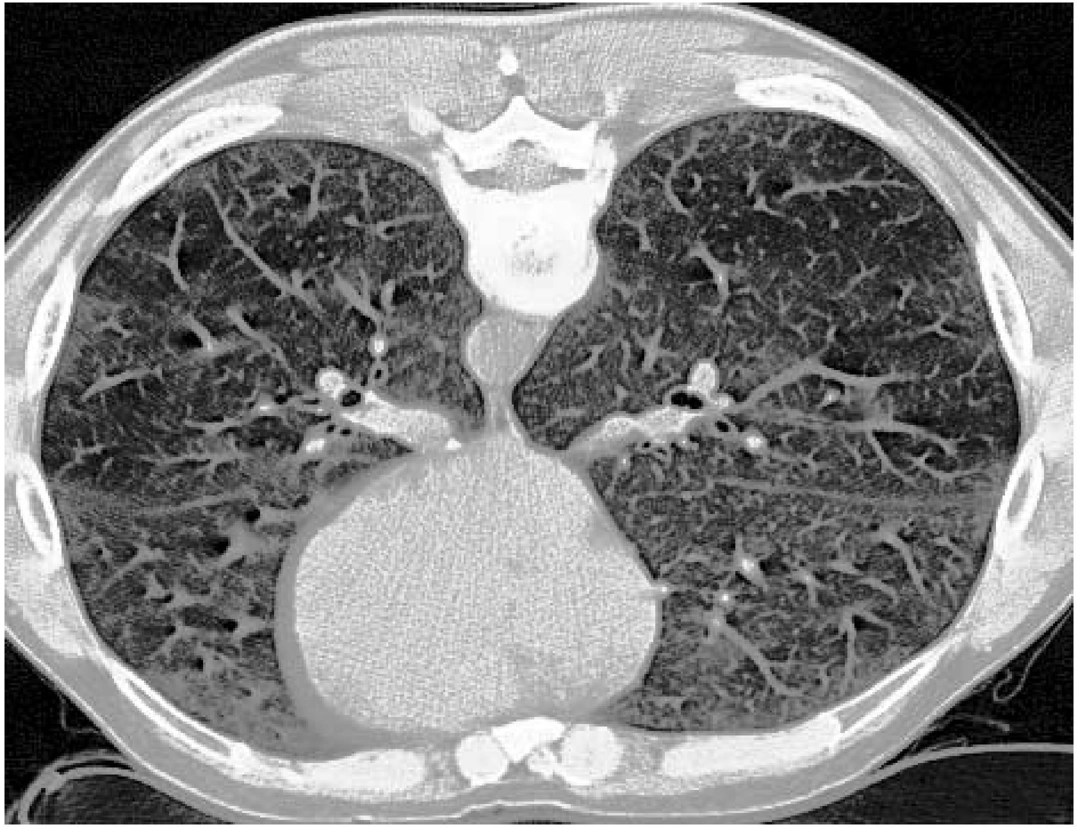
2. Sampling rate/distance (governed by Nyquist-Shannon sampling theorem).
   - The sampling frequency must be $> 2\times$ higher than the maximal frequency of interest; in the sense: which would be possible to reconstruct from the sampled signal. We will be able to derive the theorem after we explain Fourier transformation.
   - Informally: In images, the samples size (pixel size) has to be twice smaller than the smallest detail of interest.

Image sampling, illustration





Example of a digital image
a single slice from a X-ray tomograph



First image scanner, 1957

The SEAC Scanner
with control console in background

The first scanned image was the R. Kirsch's baby.

He used two thresholds and obtained three gray levels.

Russell Kirsch (*1929-†2020), SEAC and the start of image processing at the National Bureau of Standards. In: Annals of the history of computing, IEEE, vol. 20 (1998), p 7-13.

Image sampling, example 1

Original 256×256

128×128

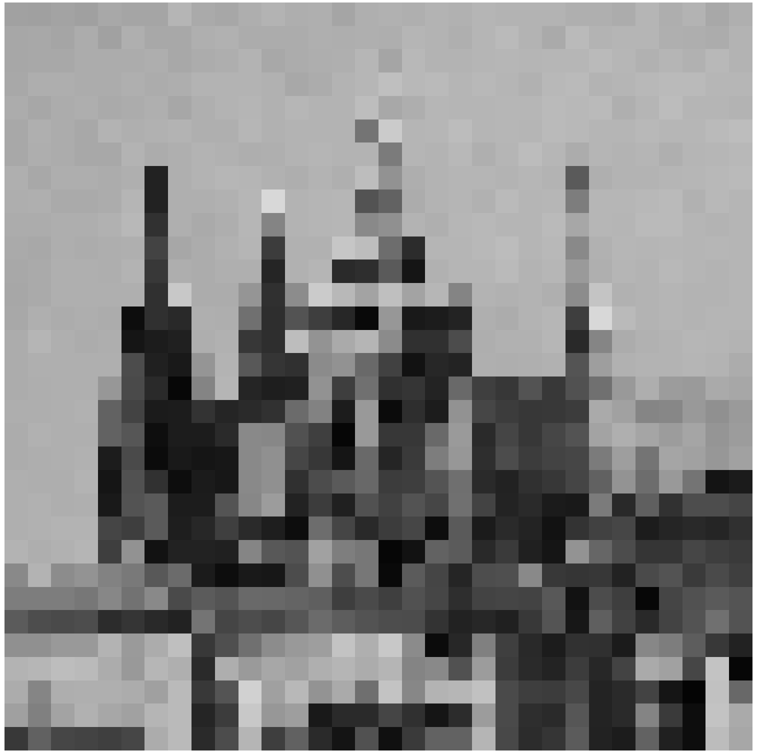
Image sampling, example 2

Original 256×256

64×64

Image sampling, example 3

Original 256×256



32×32

Image quantization, example 1

Original 256 gray levels

64 gray levels

Image quantization, example 2

Original 256 gray levels

16 gray levels

Image quantization, example 3

Original 256 gray levels

4 gray levels

Image quantization, example 4 (binary image)

Original 256 gray levels

2 gray levels

The distance, mathematically

Function D is called the **distance**, if and only if

$D(p,q) \geq 0$, specially $D(p,p) = 0$ (identity).

$D(p,q) = D(q,p)$, (symmetry).

$D(p,r) \leq D(p,q) + D(q,r)$, (triangular inequality).

Several distance definitions in the square grid

Euclidean distance

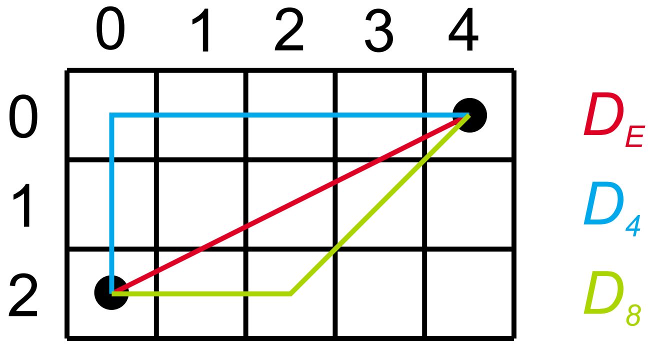
$$D_E((x,y),(h,k)) = \sqrt{(x-h)^2 + (y-k)^2}\,.$$

Manhattan distance (distance in a city with the rectangular street layout)

$$D_4((x,y),(h,k)) = \mid x-h \mid + \mid y-k \mid \,.$$

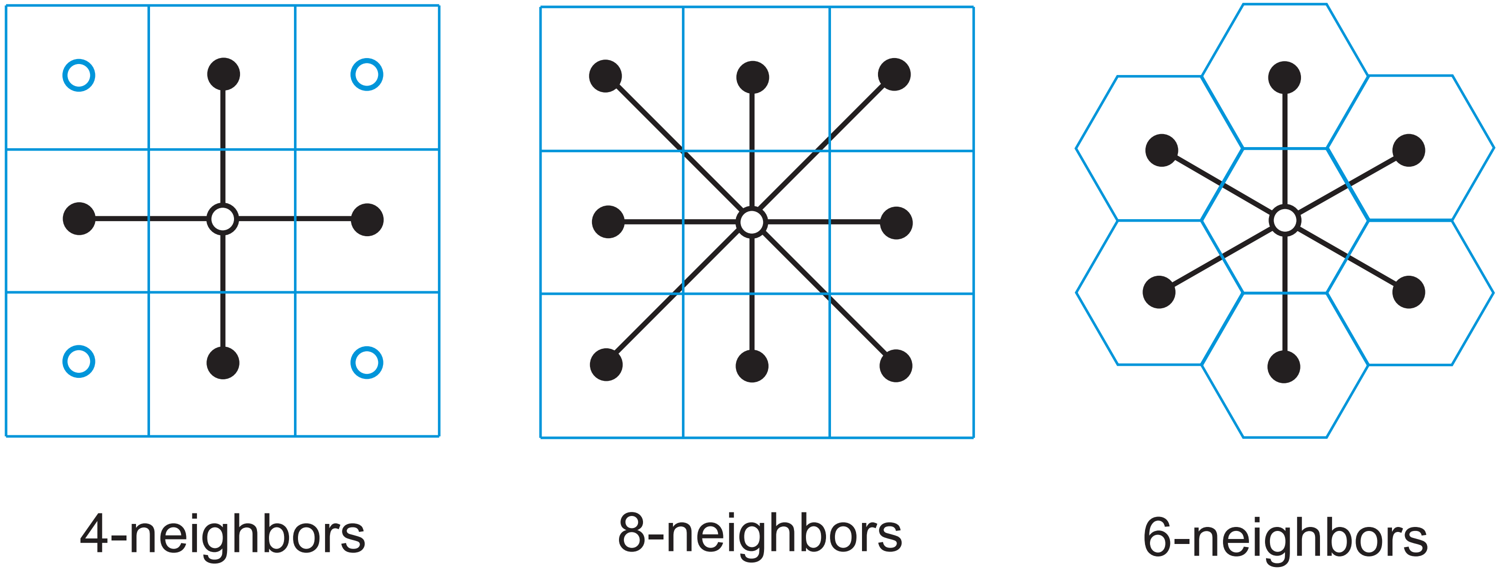
Chessboard distance (from the king point of view in chess)

$$D_8((x,y),(h,k)) = max\{\mid x-h \mid, \mid y-k \mid\}\,.$$

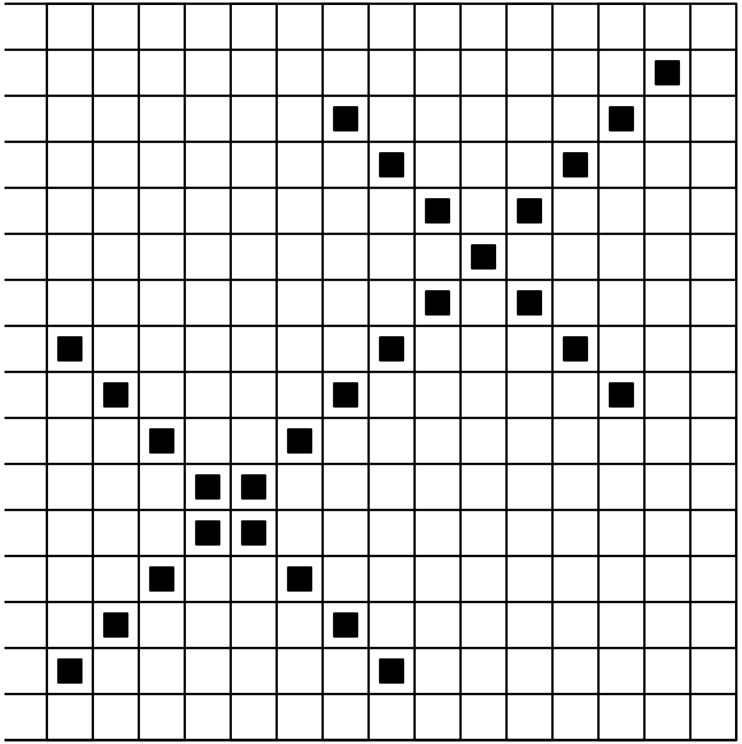


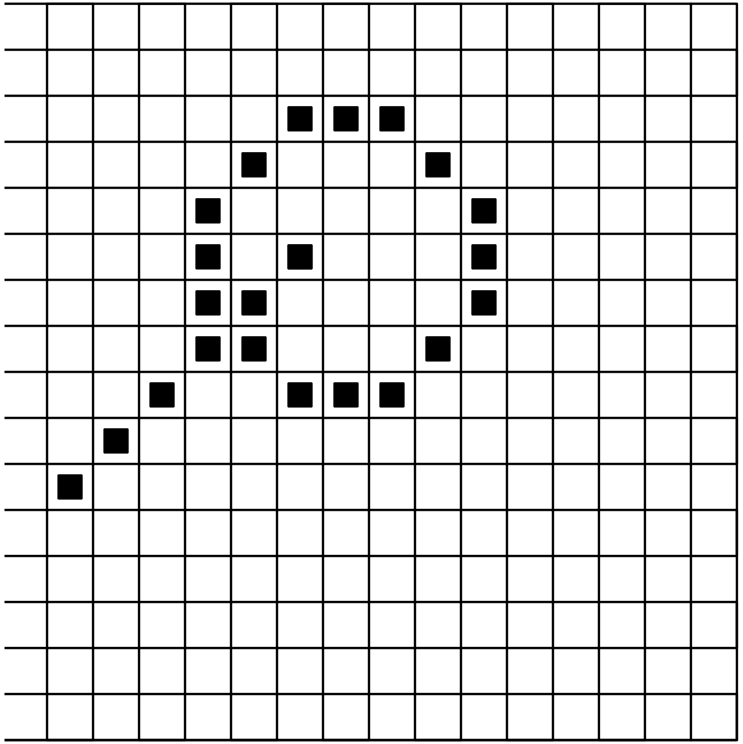
4-, 8-, and 6-neighborhoods

A set consisting of the pixel itself (shown in the middle as a hollow circle, called a representative pixel or a representative point) and its neighbors of distance 1 shown as filled in black circles.



Paradox of crossing line segments





Binary image & the relation ‘be contiguous’

black $\sim$ objects
white $\sim$ background

- *A note for curious. Japanees kanji symbol means ’near to here’.*
- Introduction of the concept ‘object’ allows to select those pixels on a grid which have some particular meaning (recall discussion about interpretation). Here, black pixels belong to the object – a character.
- Neighboring pixels are contiguous.
- Two pixels are contiguous if and only if there is a path consisting of contiguous pixels.

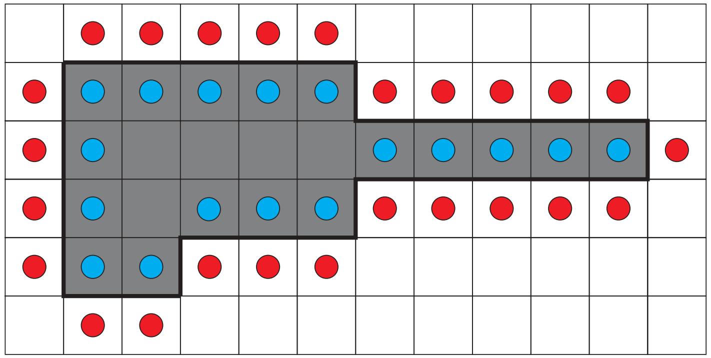
Region = contiguous set

- The relation 'x is contiguous to y' is
  - reflexive, $x \sim x$,
  - symmetric $x \sim y \implies y \sim x$ and
  - transitive $(x \sim y)$ & $(y \sim z)$

  $\implies x \sim z$. Thus it is an equivalence relation.
- Any equivalence relation decomposes a set into subsets called classes of equivalence. These are regions in our particular case of relation "to be contiguous".
- In the image below, different regions are labeled by different colors.

Region boundary

- The Region boundary (also border) R is the set of pixels is the set of pixels within the region that have one or more neighbors outside R.
- Theoretically, the the continuous image function $\Rightarrow$ infinitesimally thin boundary.
- In a digital image, the boundary has always a finite width. Consequently, it is necessary to distinguish inner and outer boundary.

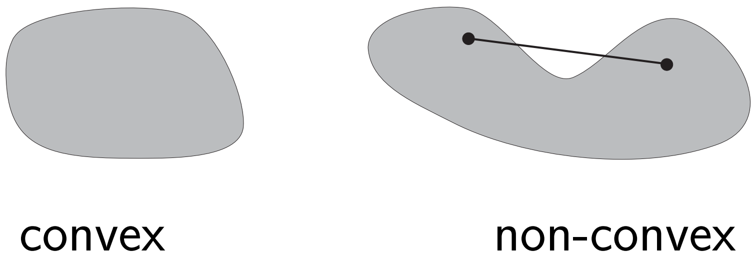


- Mind the terminology:

  Boundary (border) of a region $\times$ edge, i.e., gradient of the image function $\times$ edge element (edgel), i.e. a position with the significant magnitude of the gradient.

Convex set, convex hul

Convex set = any two points of it can be connected by a straight line which lies inside the set.



Convex hull, lake, bay.

Distance transform, DT

- Called also: distance function, chamfering algorithm (due to the analogy to woodcarving, in which material is removed layer by layer).
- Consider a binary input image, in which ones correspond to foreground (objects) and zeros to background.
- DT outputs a gray level image providing the distance from the foreground to the nearest non-zero pixel (one of objects) in the input image. DT assigns values of 0 to pixels belonging to foreground (objects)

input image

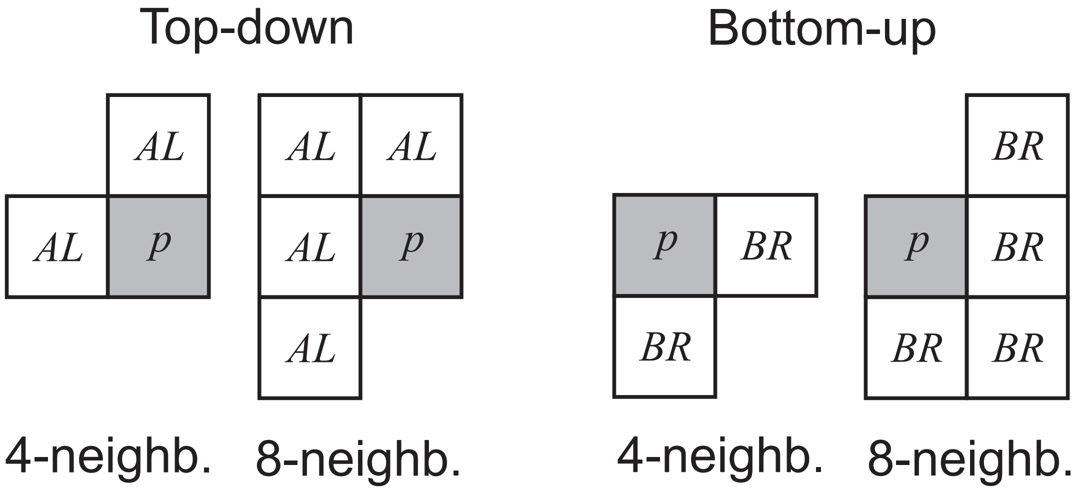
0	0	0	0	0	0	1	0
0	0	0	0	0	1	0	0
0	0	0	0	0	1	0	0
0	0	0	0	0	1	0	0
0	1	1	0	0	0	1	0
0	1	0	0	0	0	0	1
0	1	0	0	0	0	0	0
0	1	0	0	0	0	0	0

DT result

5	4	4	3	2	1	0	1
4	3	3	2	1	0	1	2
3	2	2	2	1	0	1	2
2	1	1	2	1	0	1	2
1	0	0	1	2	1	0	1
1	0	1	2	3	2	1	0
1	0	1	2	3	3	2	1
1	0	1	2	3	4	3	2

Distance transform algorithm informally

- There is a famous two-pass algorithm calculating DT by Rosenfeld, Pfaltz (1966) for distances D_4 and D_8.
- The idea is to traverse the image by a small local mask.
- The first pass starts from top-left corner of the image and moves row-wise horizontally left to right. The second pass goes from the bottom-right corner in the opposite bottom-up manner, right to left.

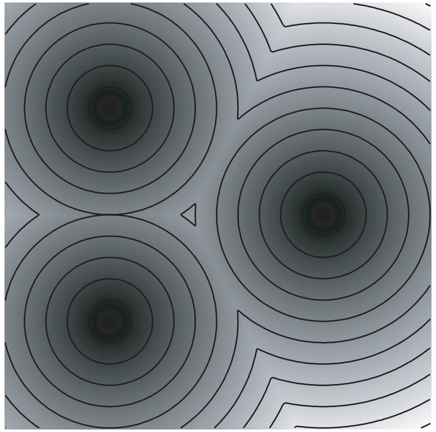


- The effectiveness of the algorithm comes from propagating the values of the previous image investigation in a 'wave-like' manner.

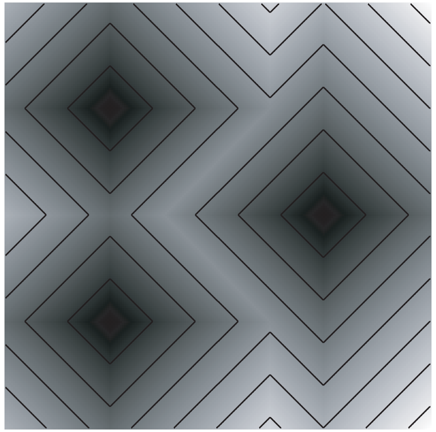
Distance transform algorithm

1. To calculate the distance transform for a subset S of an image of dimension $M \times N$ with respect to a distance metric D, where D is one of D_4 or D_8, construct an $M \times N$ array F with elements corresponding to the set S set to 0, and all other elements set to infinity.
2. Pass through the image row by row, from top to bottom and left to right. For each neighboring pixel above and to the left (illustrated by the set AL in the previous slide), assign
$$F(p) = \min_{q \epsilon AL} \big(F(p), D(p,q) + F(q)\big) .$$
3. Pass through the image row by row, from bottom to top and right to left. For each neighboring pixel below and to the right (illustrated by the set BR on the previous slide), assign
$$F(p) = \min_{q \epsilon BR} \big(F(p), D(p,q) + F(q)\big) .$$
4. The array F now holds a chamfer of the subset S.

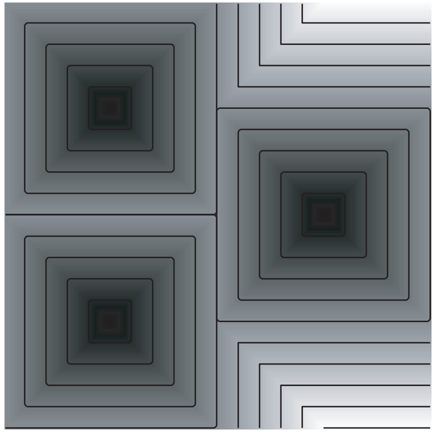
DT illustration for three distance definitions



Euclidean



D_4

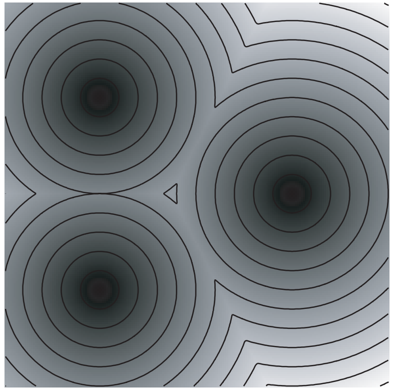


D_8

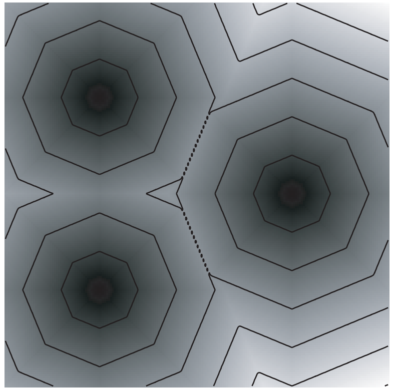
Quasieucledean distance

Eucledean DT cannot be easily computed in two passes only. The quasieucledean distance approximation is often used which can be obtained in two passes.

$$D_{\mathrm{QE}}\big((i,j),(h,k)\big) = \begin{cases} |i-h| + (\sqrt{2}-1)\,|j-k| & \text{for } |i-h| > |j-k|\,, \\ (\sqrt{2}-1)\,|i-h| + |j-k| & \text{otherwise.} \end{cases}$$

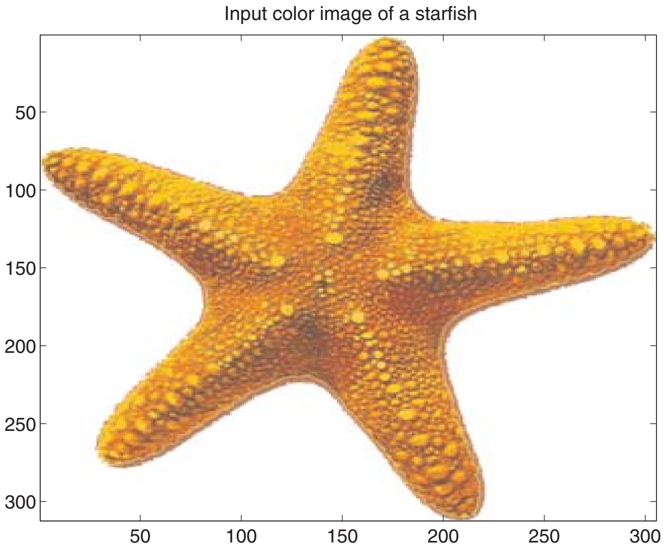


Euclidean

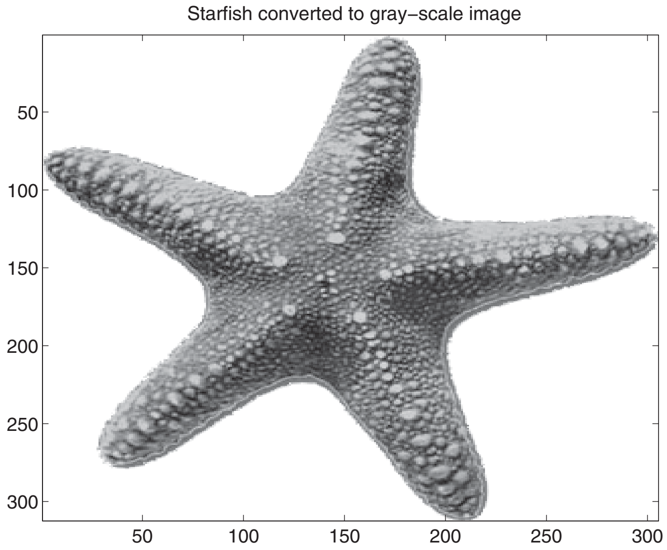


quasieuclidean

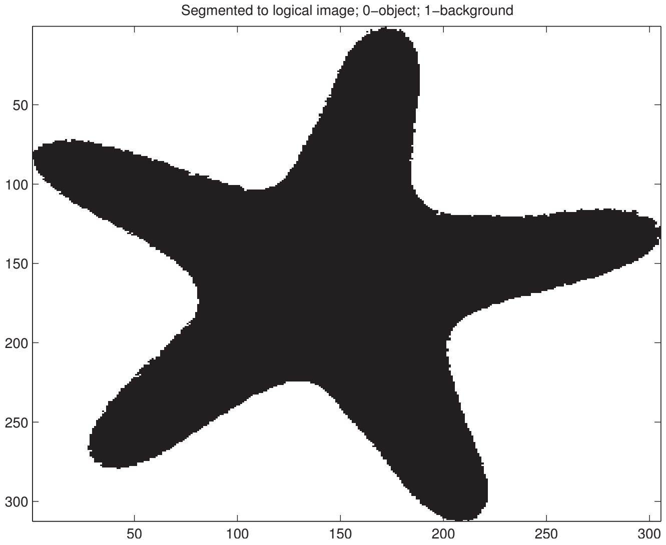
DT, starfish example, input image



color

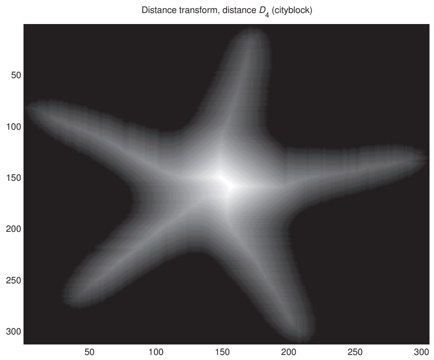


grayscale

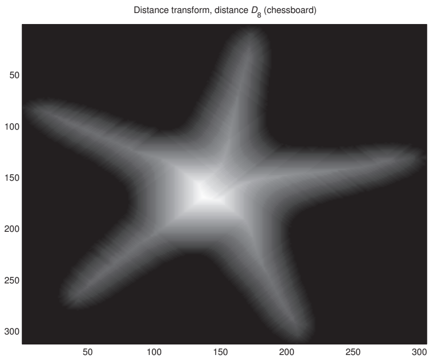


binary

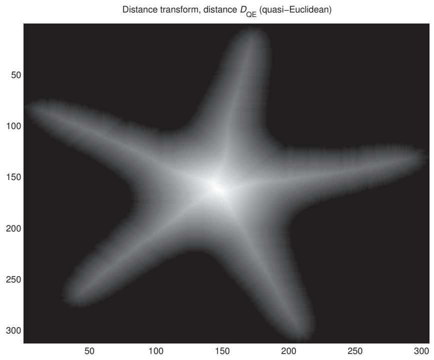
DT, starfish example, results



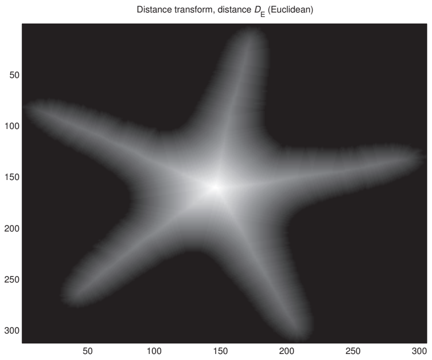
D4



D8



Quazieuclidean



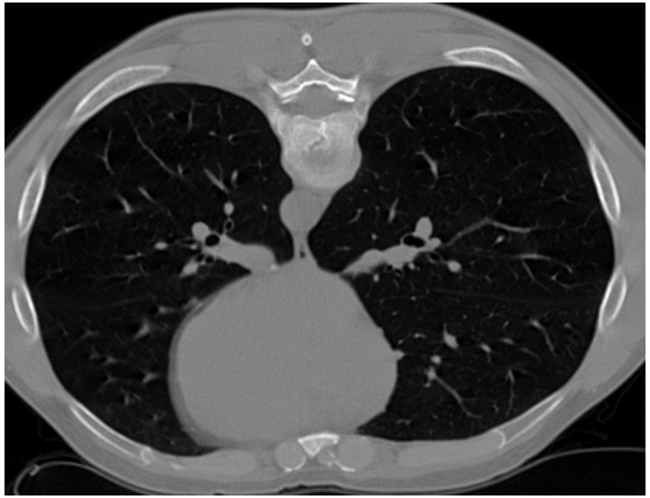
Euclidean

Image properties used in its assessment/enhancing

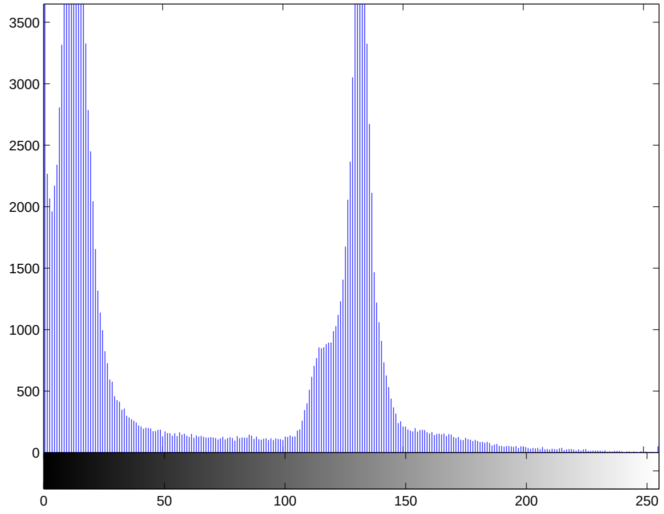
- When processing/enhancing an image by a human/machine, the image has to be assessed based on its properties.
- When a human observes the image, her/his image perception is influenced by a complex processing/interpretation in the brains and related illusions.
- We avoid such complexity pragmatically. The appropriateness of the image for human viewing is often simplified significantly to
  - one objective property – the image histogram, and
  - four subjective/objective properties: brightness, contrast, color saturation, and sharpness.

Image brightness histogram

- Let consider a gray scale image initially. We will use the human chest cross section in computer tomography image in the example below. Image histogram can be extended to color images too. Histograms are expressed independently in three color components, e.g. RGB.
- Histogram of brightness values serves as the probability density estimate of a phenomenon, that a pixel has a definite brightness.



input image



its brightness histogram

Image properties from a human perception perspective

- When a human observes the image, her/his image perception is influenced by a complex processing/interpretation of the percepts in the brains and related illusions.
- We avoid such complexity pragmatically. The appropriateness of the image for human viewing is often simplified significantly to four properties of the image, i.e. the
  - image brightness,
  - image contrast,
  - image saturation (for color images only),
  - sharpness.

Brightness, contrast, color saturation, sharpness in the image

- Image brightness characterizes the overall lightness or darkness of the image.
- Image contrast characterizes the difference (separation) in luminance/color between objects or regions. For instance, a snowy fox on a snowy background has a low contrast. A dark dog on a snowy background has a high contrast.
- Image color saturation is a similar property to the contrast. However, instead of increasing the separation between object/regions in gray scale representation, the separation in the color domain is considered.
- Image sharpness is defined as the contrast of the edges, i.e. in the direction of the image brightness gradient. When we increase the sharpness, we increase the contrast only near the edges, proportional to the size of the gradient. In parts of the image that change little, we change the value of the image function less.

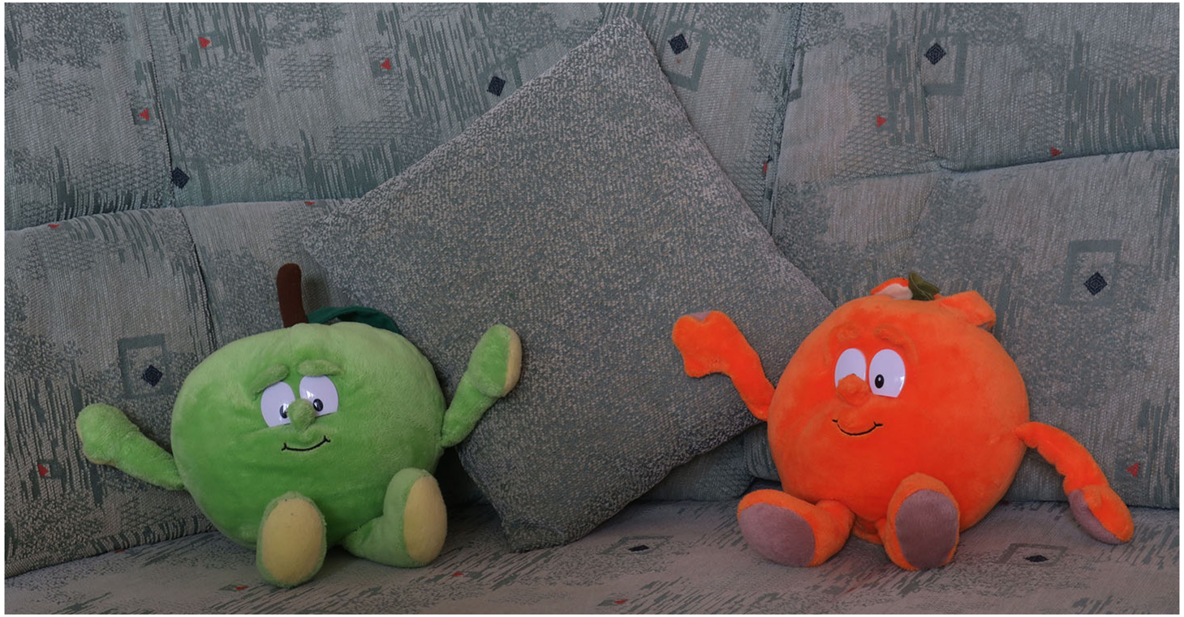
  To support imagination, we use the analogy of (image) $\leftrightarrow$ (landscape) or (brightness value) $\leftrightarrow$ (height in the landscape at the relevant location). Then sharpening will increase the slope of the slopes, more for steep slopes and less for gentle slopes.

Towards enhancing a single image

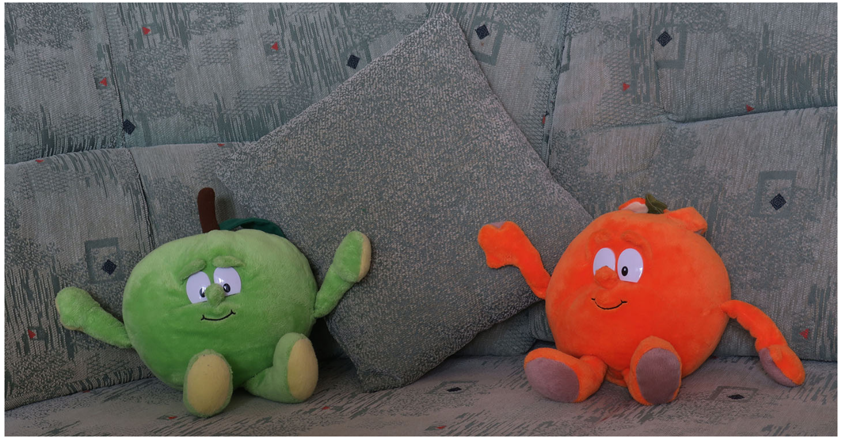
- Let consider a common practical task. Let have only a single captured digital image. A human observer is not satisfied with its appearance manifested by image brightness, contrast, color saturation or sharpness.
- The cause could be that the scene lighting was not appropriate, the image sensor dynamic range was not high enough, the objects were not distinguished from the background, etc.
- A human influences namely the brightness, contrast, color saturation or sharpness when processing the already captured digital image, e.g. in PhotoShop.
- Let us illustrate the need of the image assessment/enhancing on practical illustrative examples.

Introducing the image used in experiments

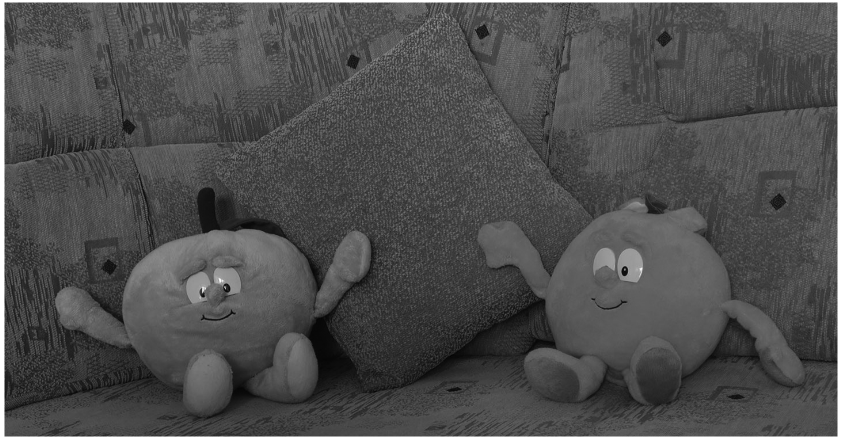
I captured the input color image of three objects on a light green sofa background. The scene shows deliberately one object (a square cushion) similar in color to the background and two objects, two plush toys, differing from the background in color. The first toy has different green color than the sofa and the second toy is orange.



We begin with gray level images, a conversion

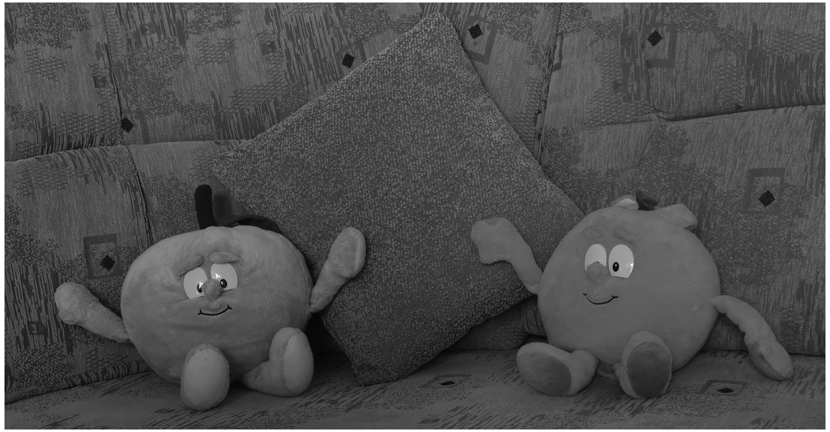


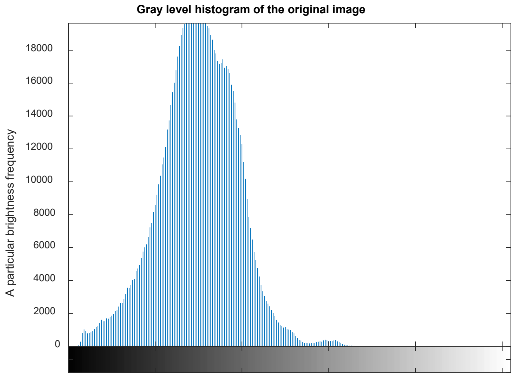
Original color image

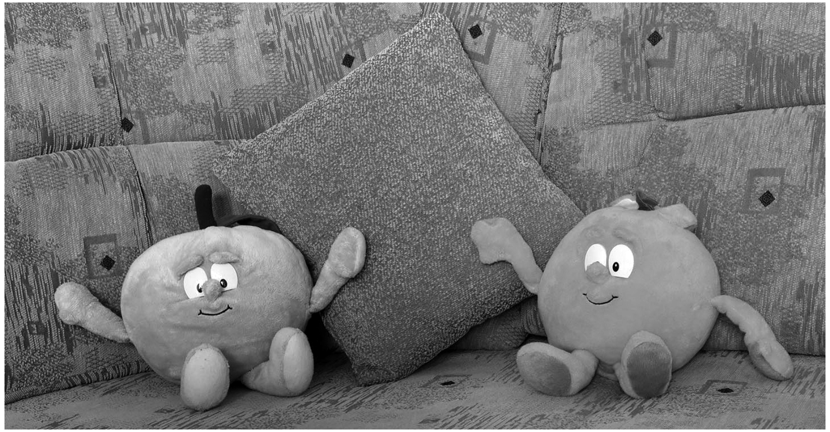


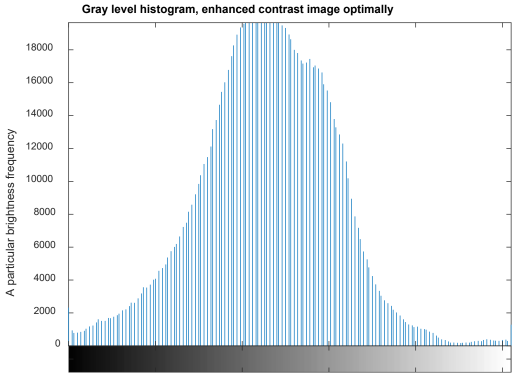
In the gray scale

Optimally increased contrast illustrated

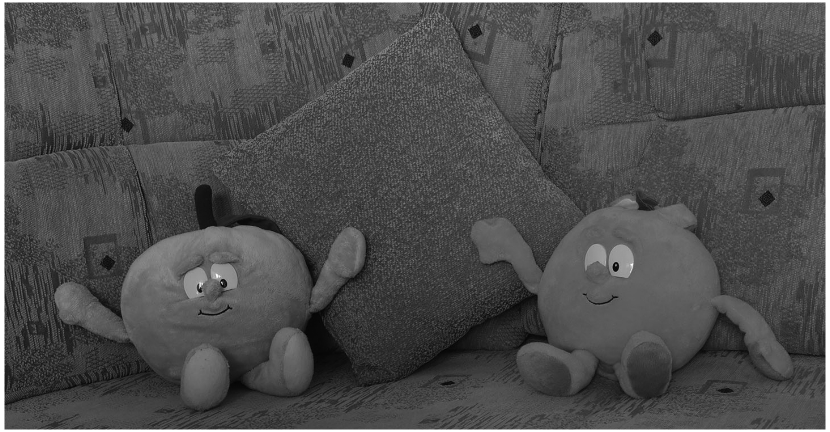


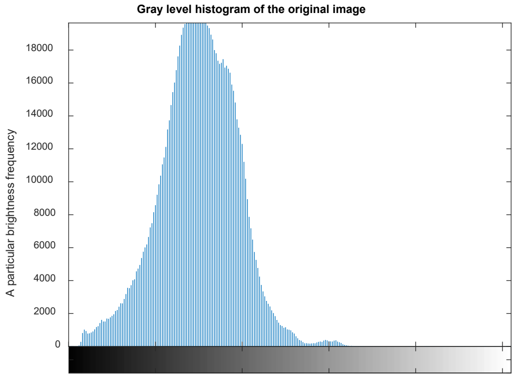


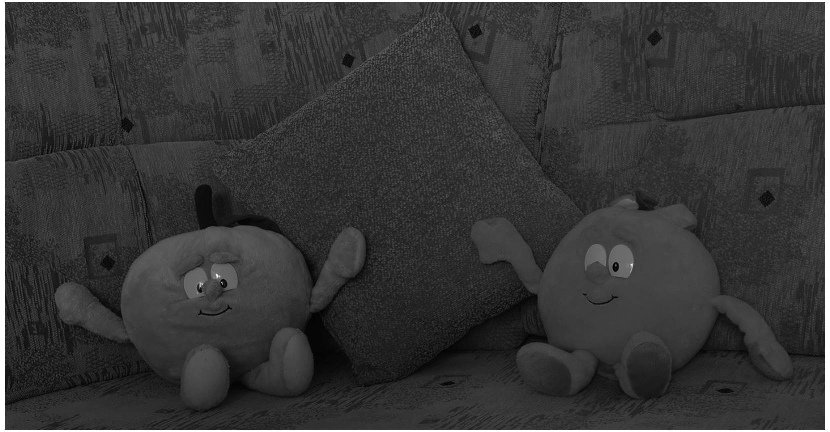


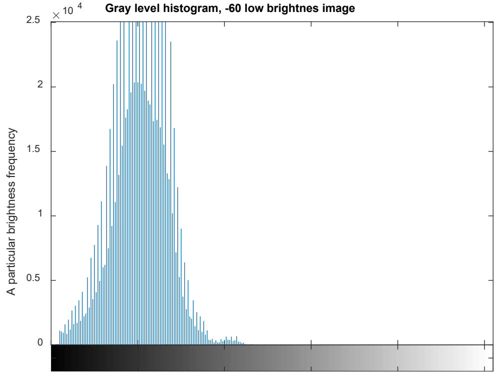


Low brightness illustrated

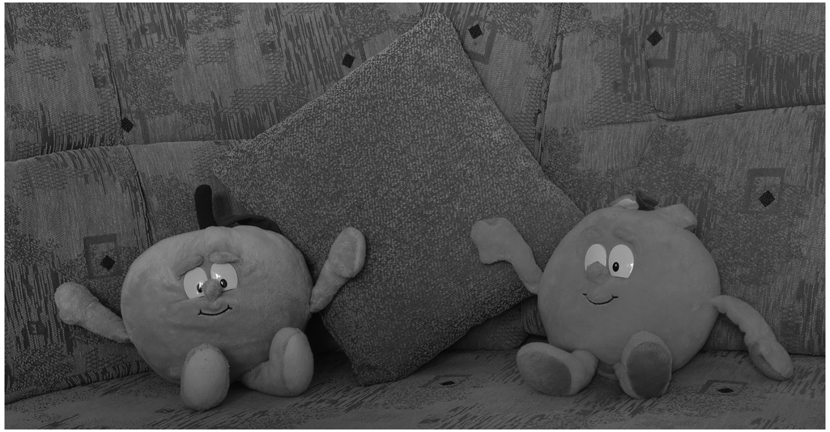


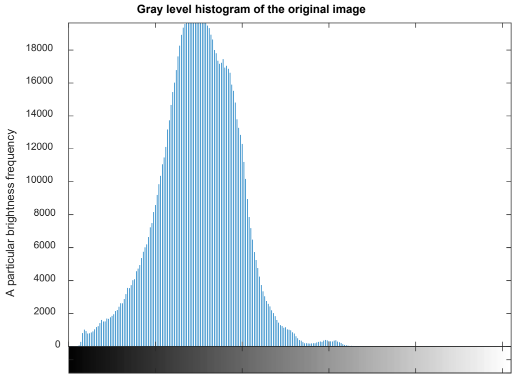


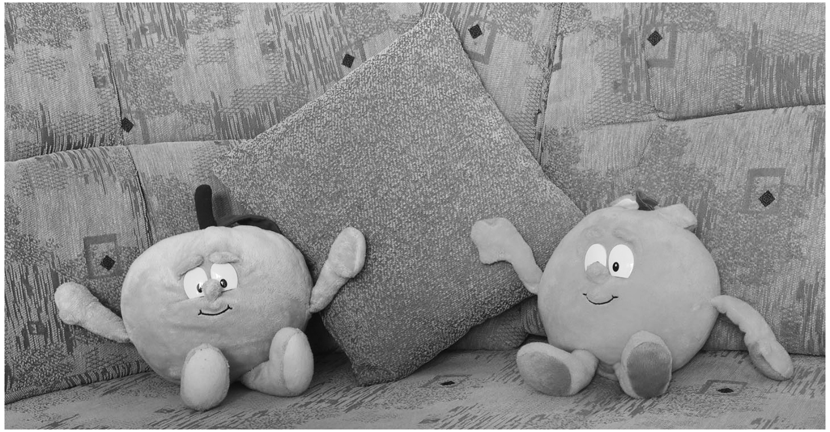


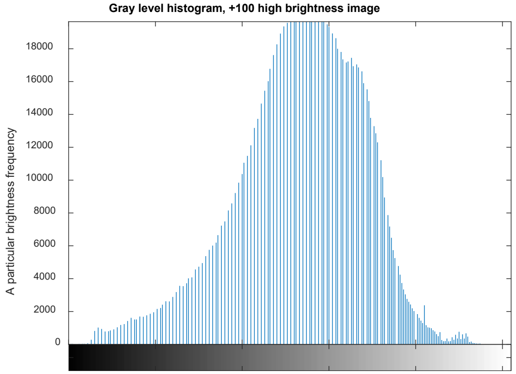


Increased brightness illustrated

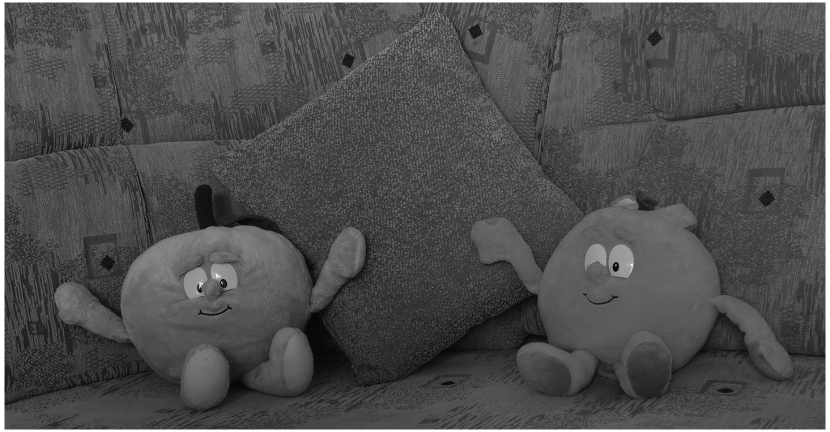


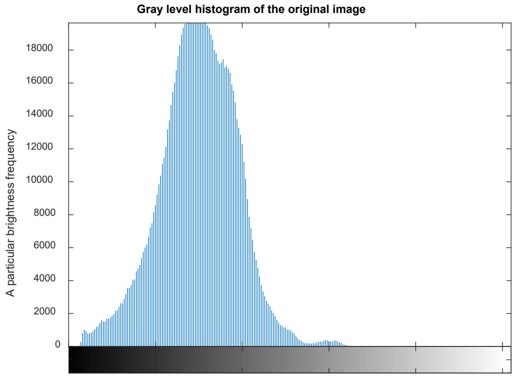


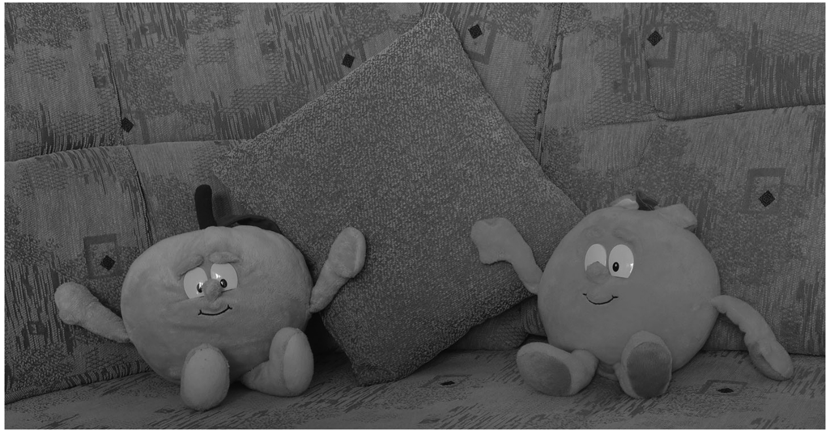


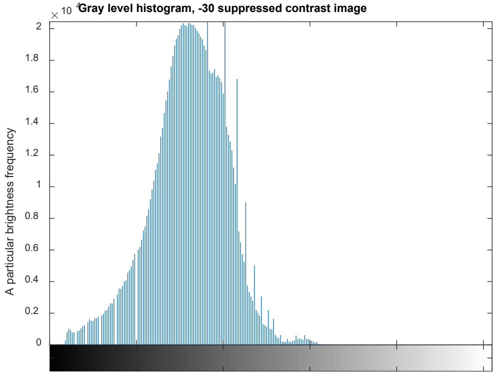


Suppressed contrast illustrated

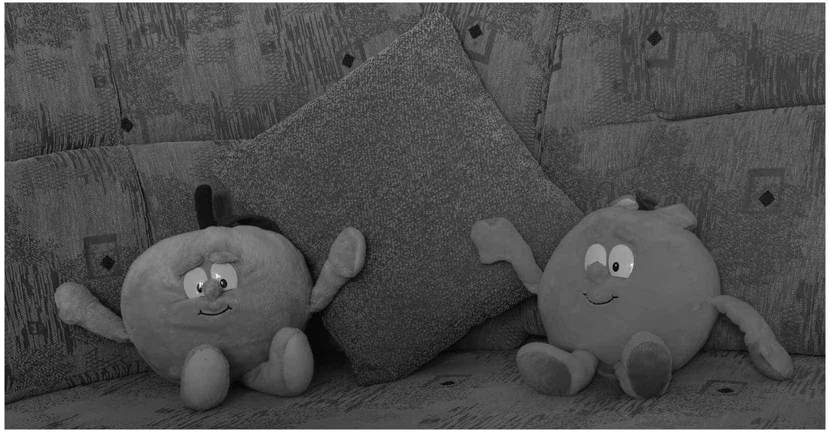


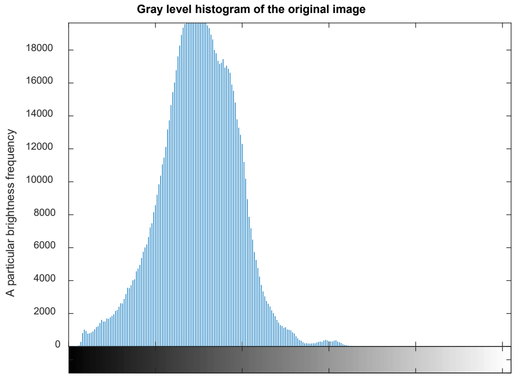


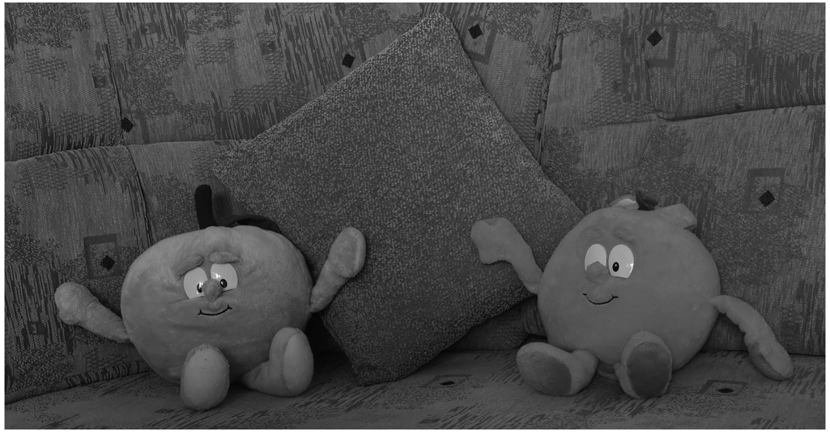


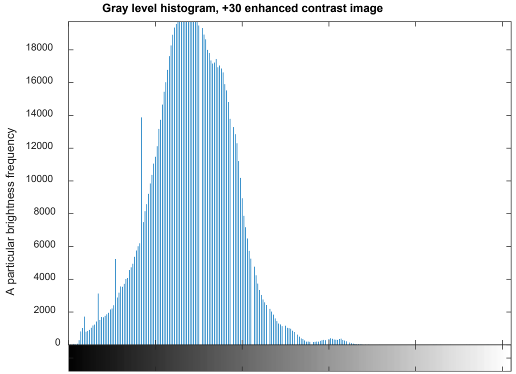


Enhanced contrast illustrated









Color saturation illustration

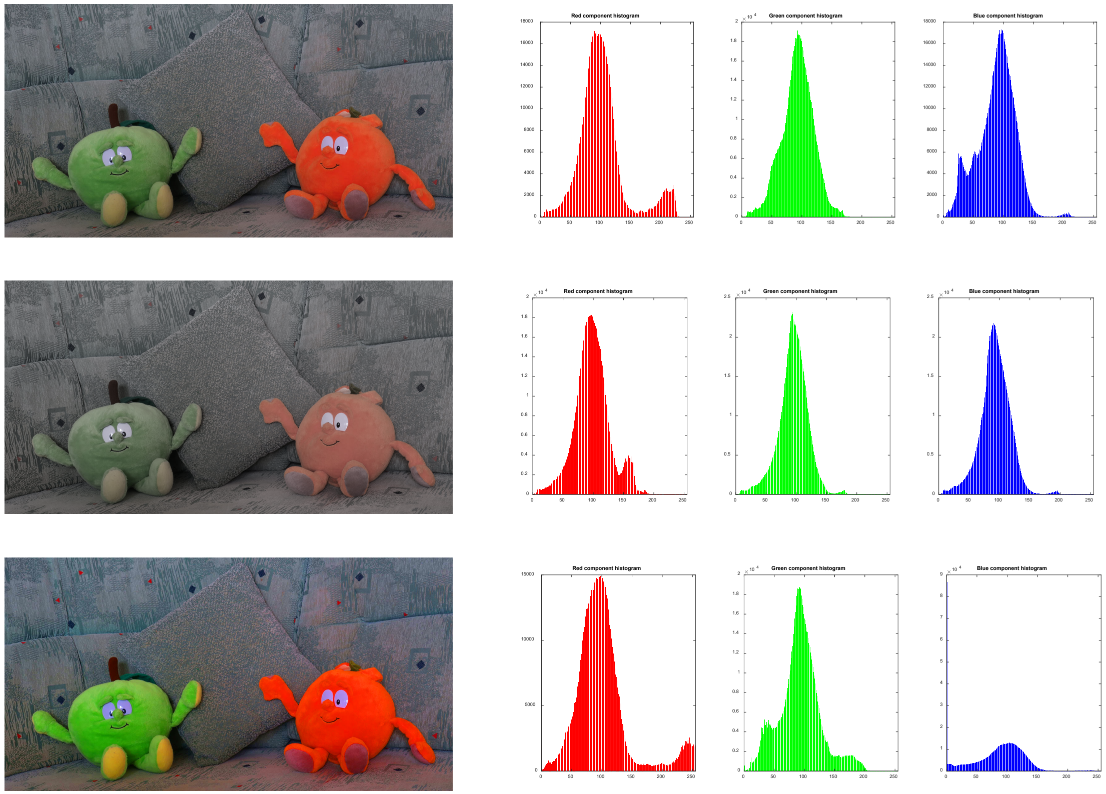
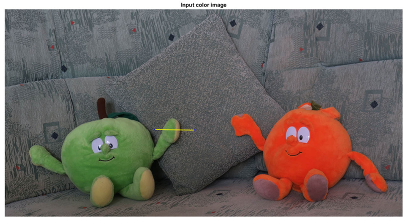
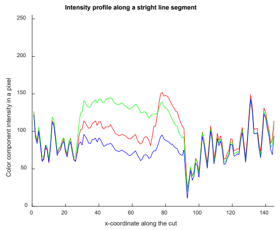


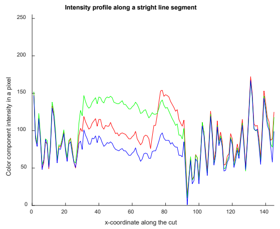
Image sharpness illustrated

Input image

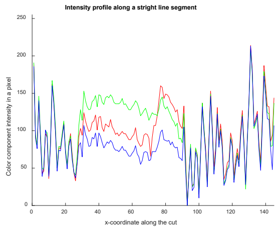


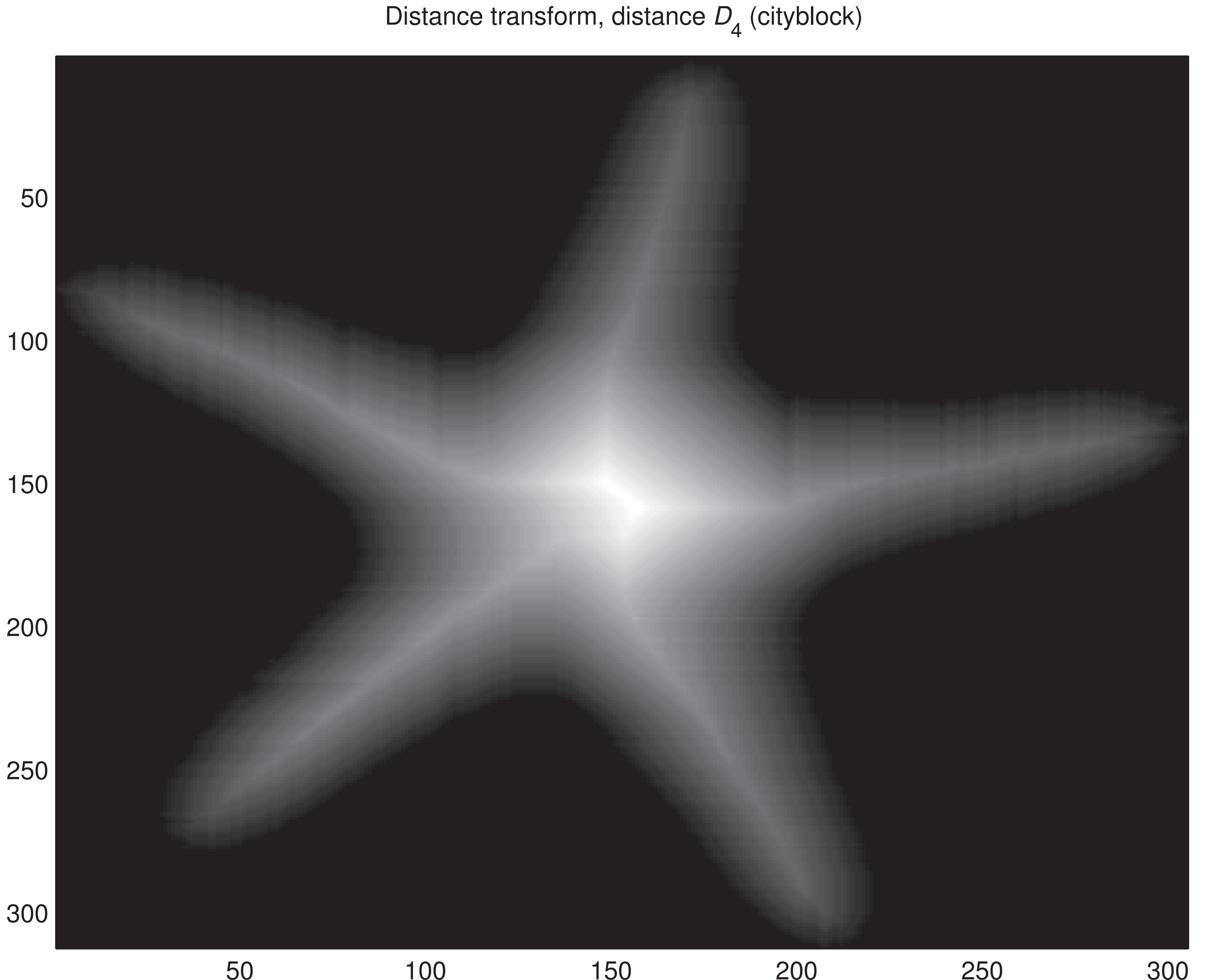


Moderate sharpening

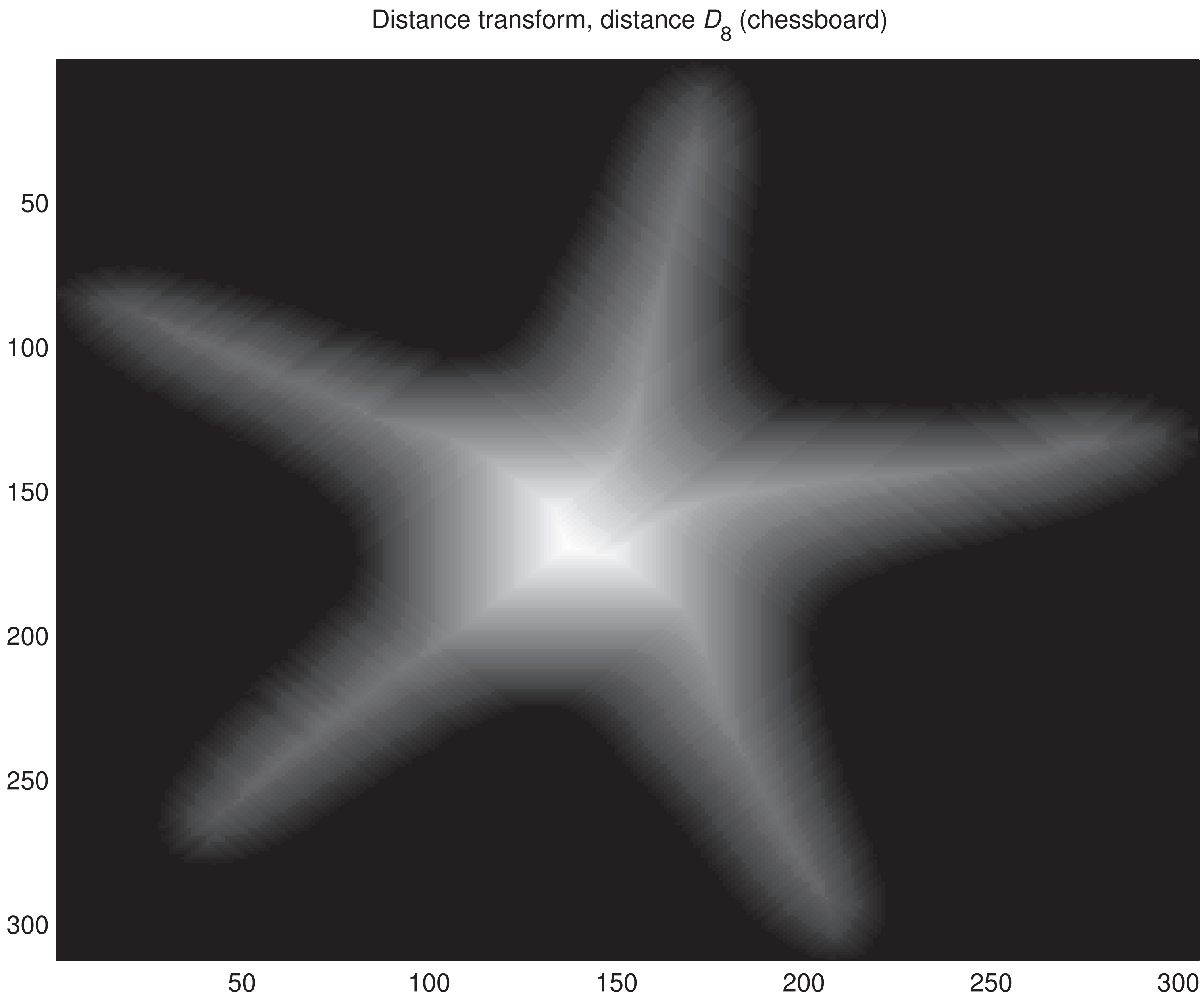


Stronger sharpening

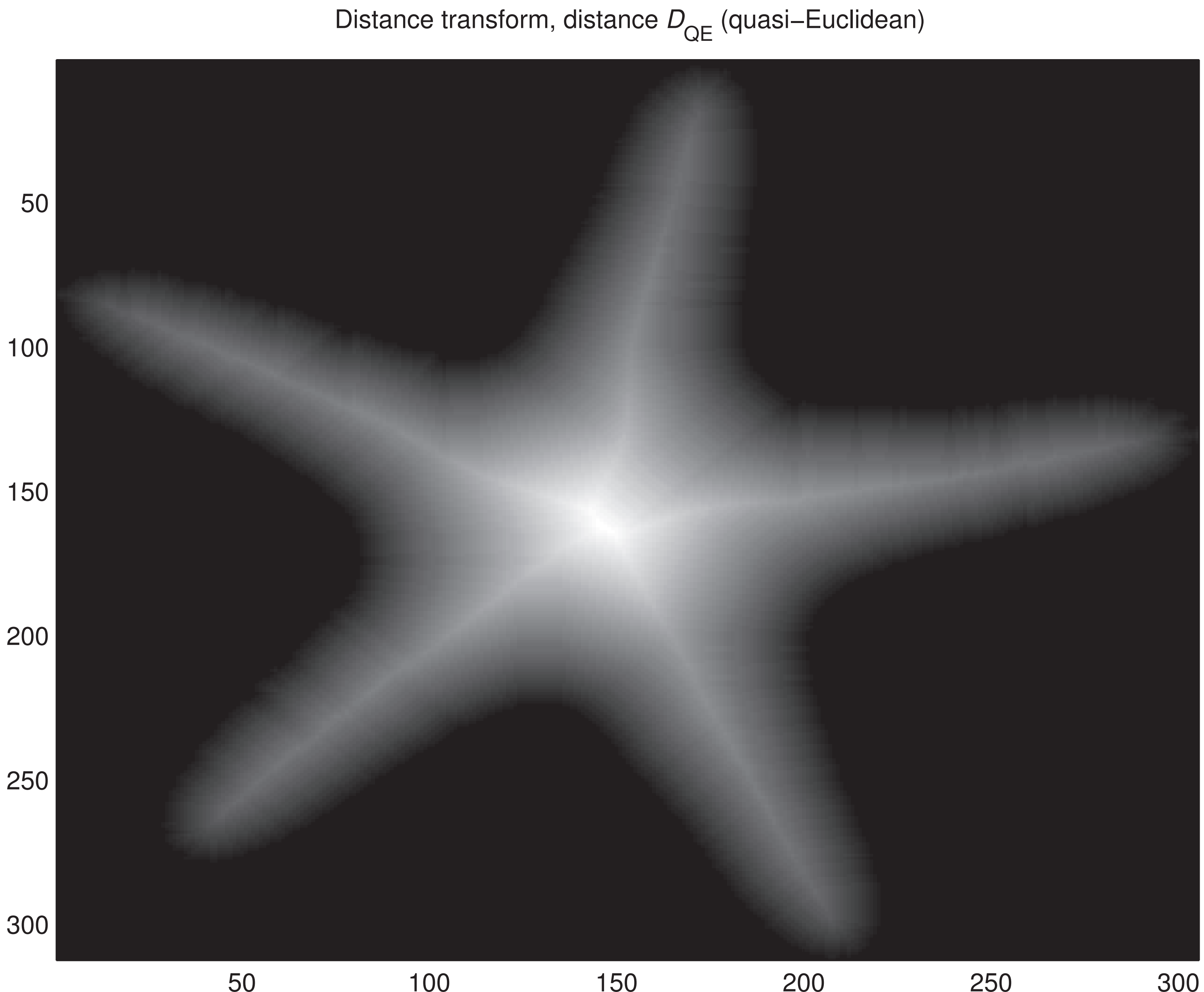


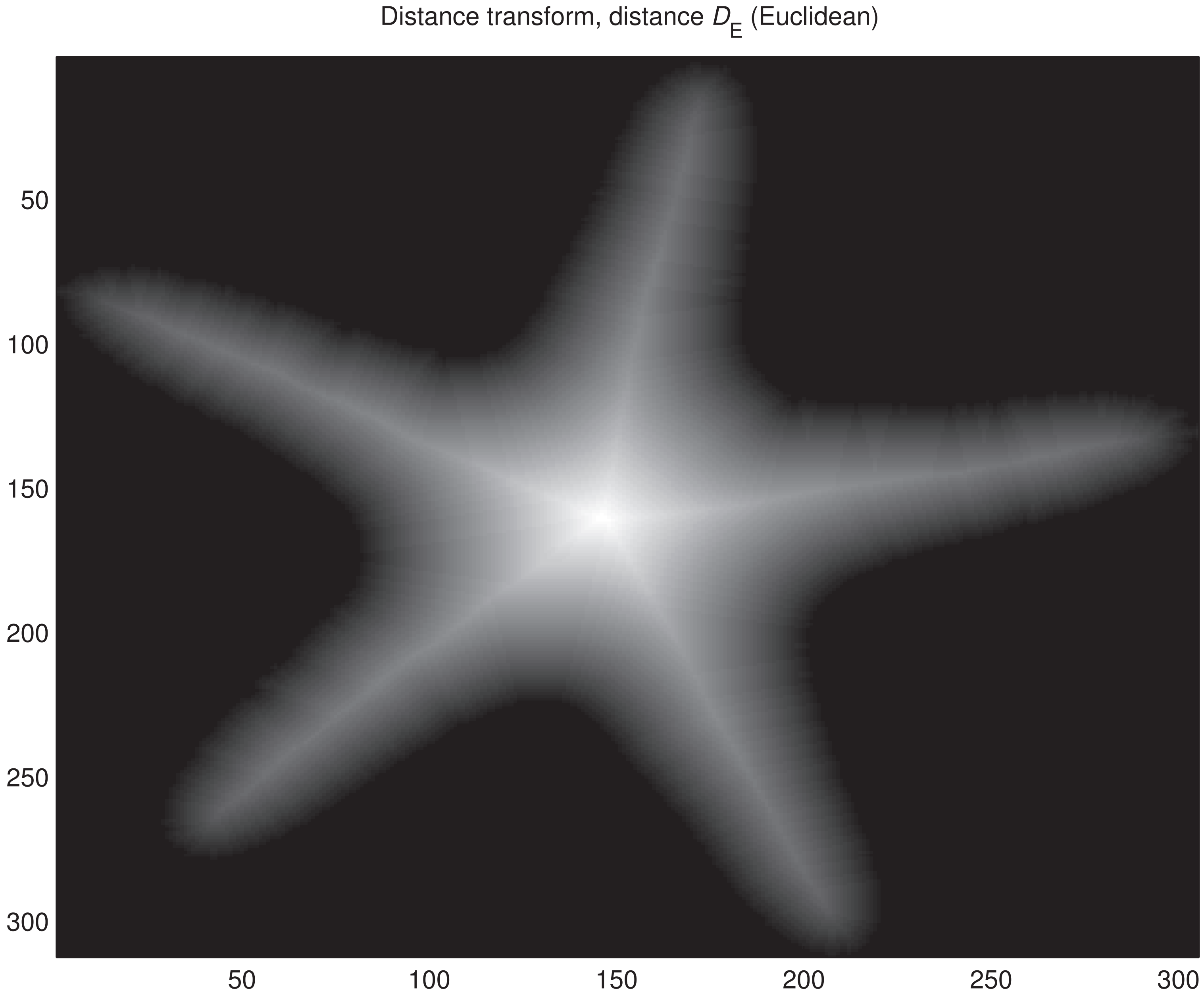


Distance transform, distance D_4 (cityblock)
50
100
150
200
250
300
50
100
150
200
250
300



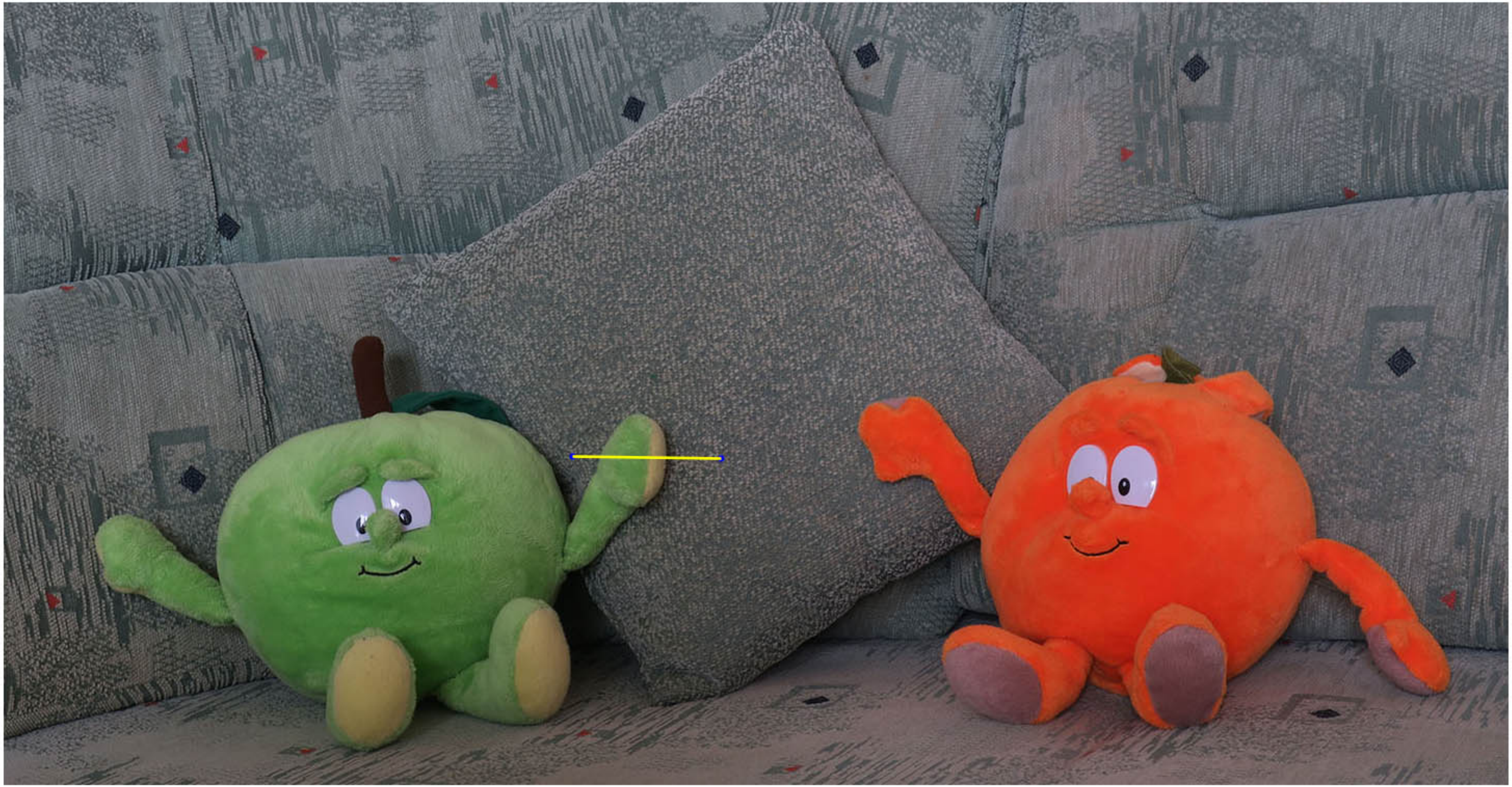
Distance transform, distance D_8 (chessboard)
50
100
150
200
250
300
50
100
150
200
250
300

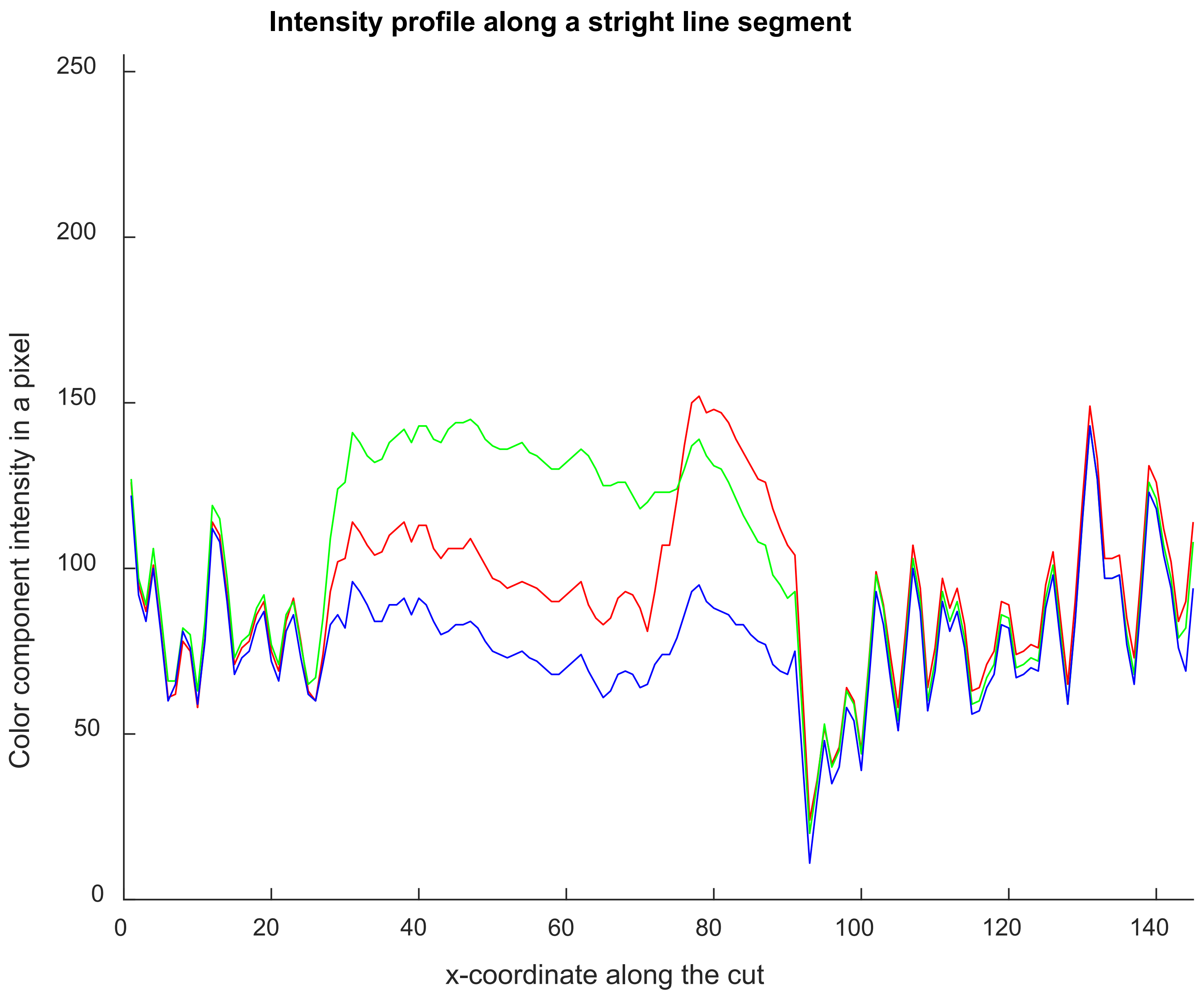

Distance transform, distance D_{QE} (quasi–Euclidean)
50
100
150
200
250
300
50
100
150
200
250
300



Distance transform, distance D_E (Euclidean)
50
100
150
200
250
300
50
100
150
200
250
300

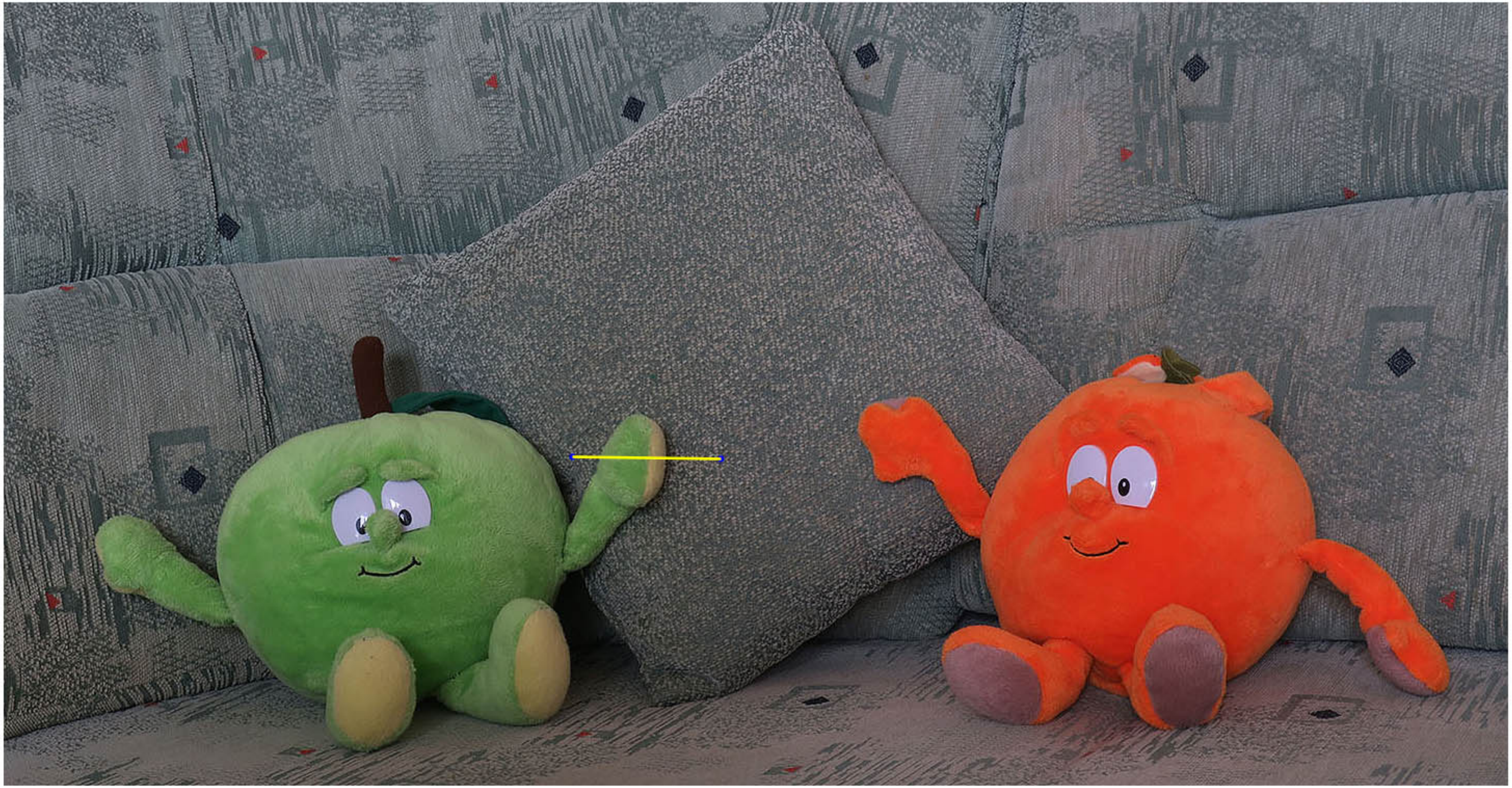
Input color image

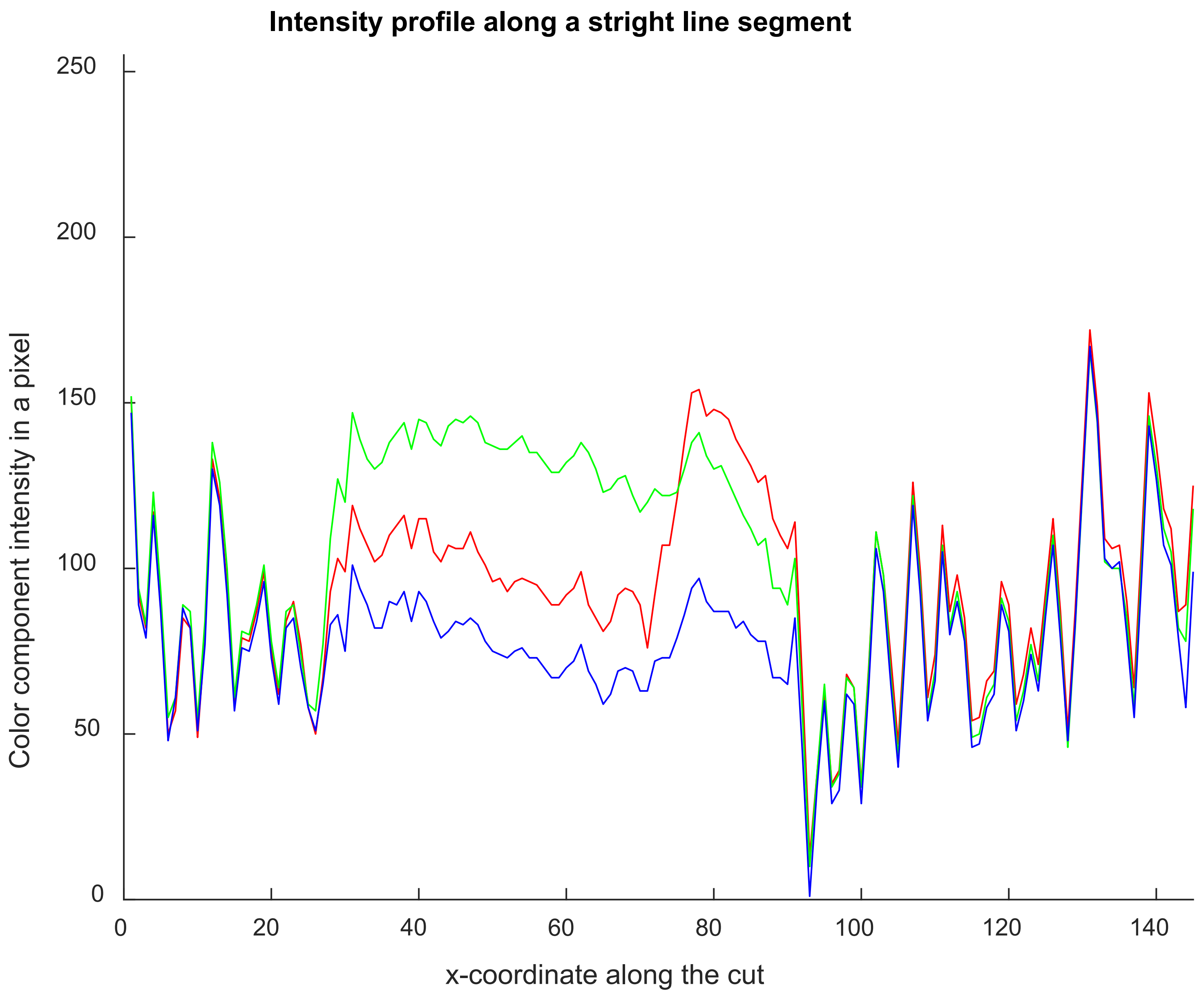




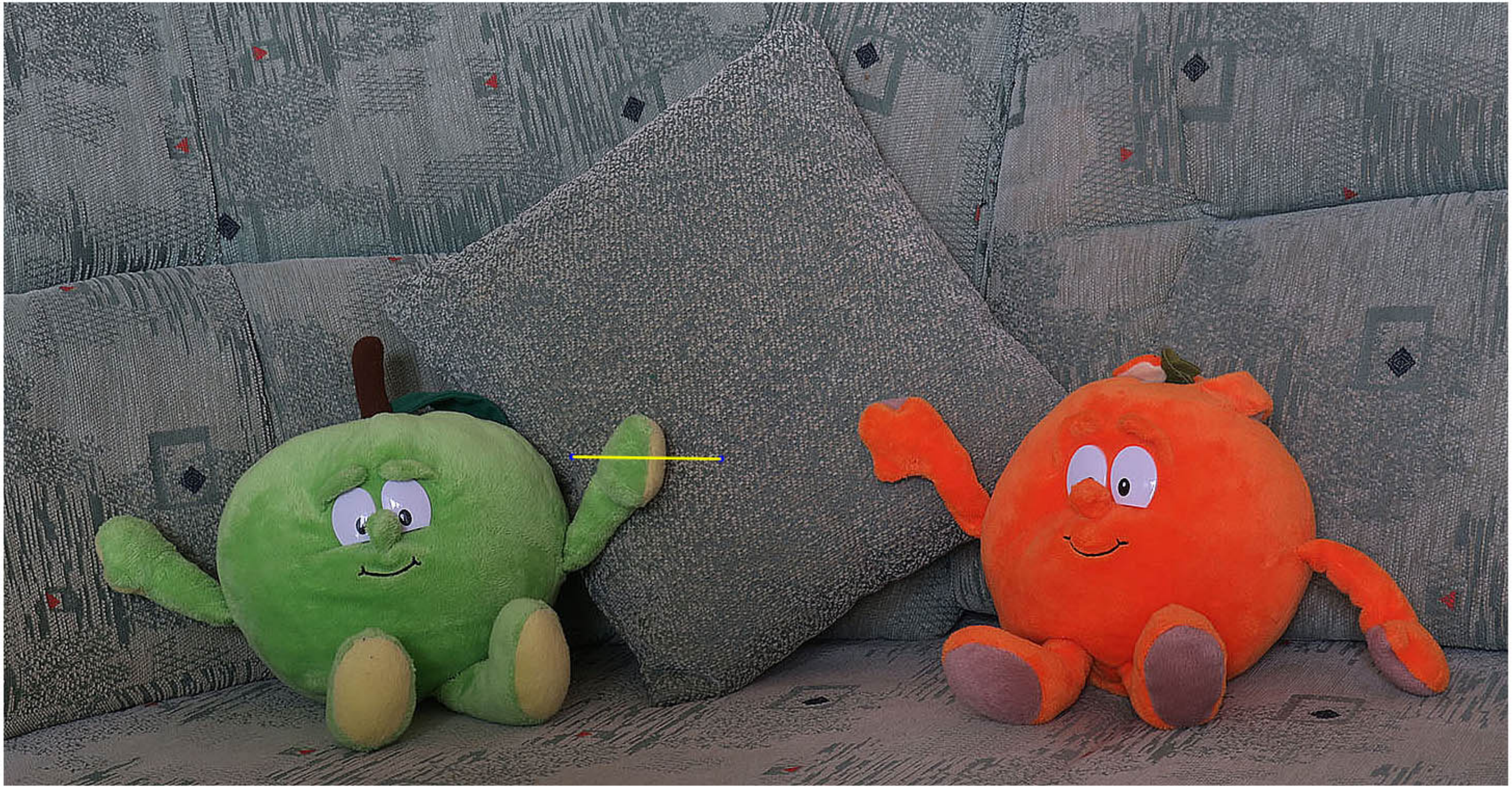
Intensity profile along a stright line segment
Color component intensity in a pixel
x-coordinate along the cut
0
50
100
150
200
250
0
20
40
60
80
100
120
140

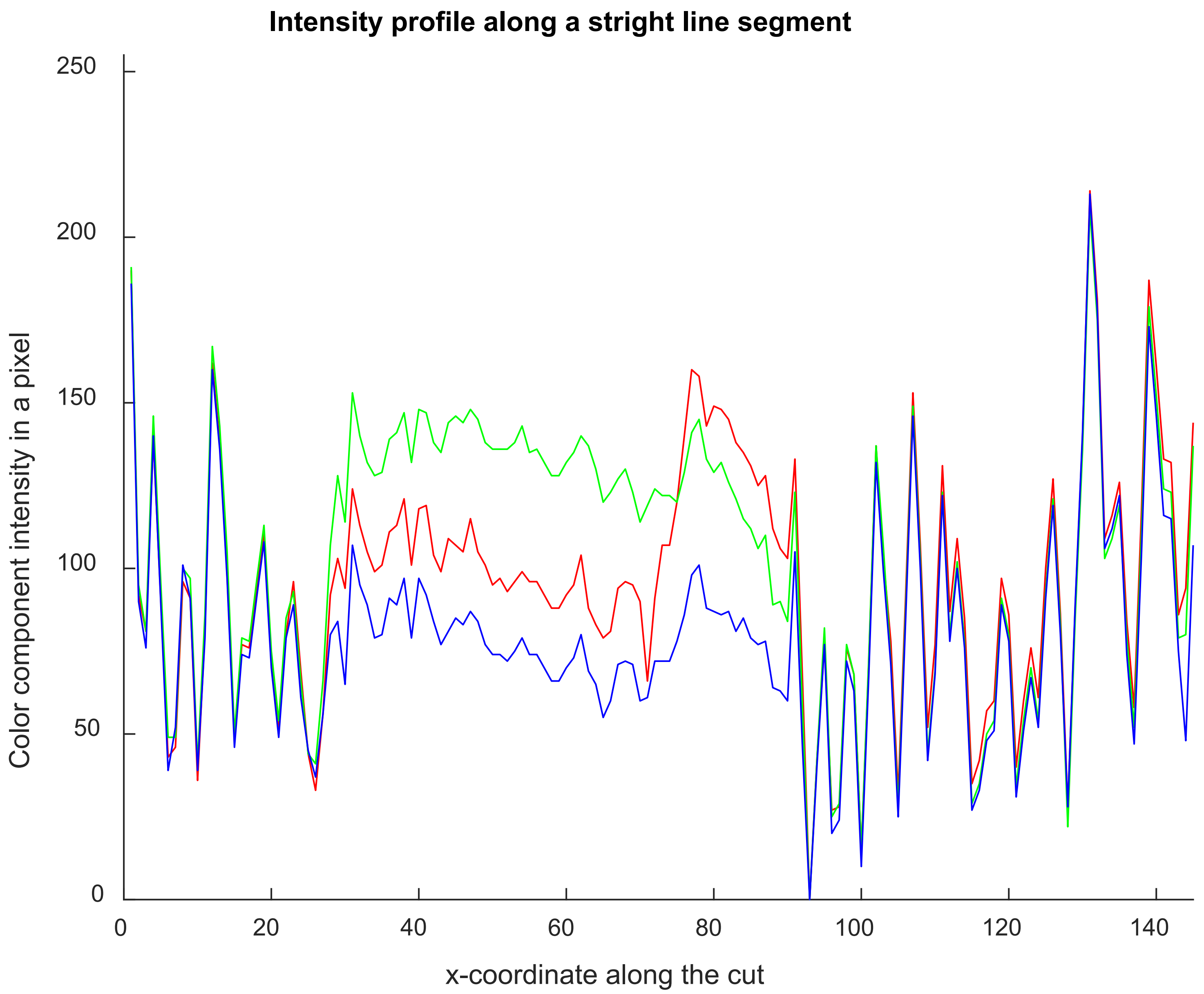
Sharpened image




Intensity profile along a stright line segment
Color component intensity in a pixel
x-coordinate along the cut
0
50
100
150
200
250
0
20
40
60
80
100
120
140

More sharpened image





Intensity profile along a stright line segment
Color component intensity in a pixel
x-coordinate along the cut
0
50
100
150
200
250
0
20
40
60
80
100
120
140