Exercise 1, outline solutions: Image formation

17 January 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

1. Draw parallels between a typical (still) camera and the following components of the eye:

    - **Lens → lens**
    - **Aperture → pupil**
    - **Shutter → no equivalent**
    - **Sensor → retina**

2. When the potential well in CCD becomes full?

    When the amount of charge an individual pixel can hold is reached or exceeded.

3. Is there a "full well" equivalent in the human vision?

    There is no exact equivalent as the eye has a number of mechanisms for adjusting the dynamic range. The closest phenomenon is when the retinal neurons firing frequency reaches the limit which leads to "bleaching". See e.g.

    https://en.wikipedia.org/wiki/Adaptation_%28eye%29

    http://www.telescope-optics.net/eye_intensity_response.htm#modes

4. Lists the key steps on the CCD pathway from photons to digital image values.

    - **Photons of light induce photoelectrons in a sensor (CCD)**
    - **Photoelectrons charge accumulates in the potential well**
    - **Charge is transferred to amplifier and converted into voltage**
    - **Voltage is measured and converted to numerical values by ADC**
    - **Numerical values are transferred to computer memory.**

5. Images can be generated from different forms of energy registered by a suitable sensor. Name three kinds of energy capable of generating images.

- **Electromagnetic waves**
- **Magnetic resonance**
- **Sound**
- **Particles**
- **Fields**

6. CCD sensors vary in physical size (e.g. 4.8 x 3.6 mm, 6.4 x 4.8 mm, etc) and the number of pixels (e.g 1.3 million, 2.1 million). Given two chips with the same sensor size but a different number of pixels,

- A - 1.3 million,
- B - 2.1 million,

which one would you choose for imaging in low-light conditions, and why.

A, because size of individual pixels is larger, which means that each pixel is capable of collecting more light.

Exercise 2, outline solutions: Digital image

18 January 2017

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

1. Name two key attributes of pixels in a monochrome image.

   - **Location (row, column)**
   - **Value**

2. Which of the two images of size 500 x 400 pixels will look better when displayed on a screen:

   A - spatial resolution 1mm per pixel, or B - spatial resolution 2mm per pixel?

   A, pixels in B will appear blocky.

3. This question is based on an example from http://www.nightskyimages.co.uk/sampling_rate.htm.

   "Imagine that you have a telescope with a focal length of 1,000mm and you wish to image a deep sky object using a digital camera with a sensor whose pixels are 5 micron square in size. The theoretical resolution of the system in arcseconds per pixel can be found by using the following calculation:-
   *Resolution = (CCD Pixel Size / Telescope Focal Length) * 206.265*
   *In our example, Resolution = (5 / 1000) * 206.265 = 1.03 arcseconds/pixel."*

   Is this sufficient to image the sky details of size 4 arcseconds/pixel? Why?

   Yes, according to the Nyquist criterion the sky detail (4 arcseconds/pixel) is more than twice the sampling frequency (2*1.03 = 2.06).

4. What radiometric resolution (in terms of bits per pixel) is sufficient for adequate perception of monochrome images?

   8 bits per pixel is considered adequate for general photography. For safety-critical applications much higher resolution is recommended. Please note that the number of shades of grey that can be resolved depends on many other factors such as brightness, local contrast, noise, etc. For discussion see e.g. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3043920/ and "How many shades of gray can the human eye see?" in https://en.wikipedia.org/wiki/Wikipedia:Reference_desk/Archives/Science/2012_March_3

5. A technical specification sheet of a camera provides the following information:

   - Full well capacity: 18,000 photoelectrons
   - Readout noise: 8 photoelectrons

   What is the camera's dynamic range in dB?

$$\mathrm{dyn}_{\mathrm{CCD}} = 20 \cdot \log\left(\frac{\text{fullwell capacity } [e^-]}{\text{readout noise } [e^-]}\right)[\mathrm{dB}]$$

   20*log(18,000 / 8) = 67 dB (answer rounded to the nearest integer)

Exercise 3: Colour, part 1

24 January 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

1. What is the wavelength range (approximately) of visible light? Specify units for the wavelengths.

   400 - 700 nm (nanometers).

2. Why digital cameras and displays use RGB primaries?

   To make them compatible with human visual perception.

3. What colour will you get by illuminating a yellow surface with a green light? Justify your answer.

   Green. Yellow surface absorbs blue and reflects red and green. There is no red component in the illuminating light, so only green is reflected.

4. What colour will you get by illuminating a magenta surface with a yellow light? Justify your answer.

   Red. Magenta surface absorbs green and reflects red and blue. Green component in the yellow light is absorbed and there is no blue component, so only red is reflected.

5. Describe the key steps in the process of generating colour image using the Bayer filter array in the sensor chip.

   Bayer array consists of alternating rows of red-green and green-blue filters. A sensors under a given filter detects photons only within a given range of wavelengths (R, G and B).

   The resulting digital image is a mosaic representing a mixture of R, G and B components of the image.

The final colour image is generated by interpolating (averaging) individual pixel colours over a small region centred on the pixel.

See the recommended readings for further details.

6. Given the following Bayer filter mosaic, compute the RGB colours for the four central pixels (marked with crosses below).

255	255	255	255
0	255	0	127
255	255	255	255
127	255	0	127

	R:255 G:255 B:0	R:255 G:223 B:0	
	R:255 G:255 B:32	R:255 G:204 B:0	

Exercise 4: Colour, part 2

25 January 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

1. Specify colour definitions for the following colours in the RGB and CMY colour spaces assuming the range of the primaries 0-255.

   - Black
   - White
   - Bright orange
   - Light pink

	RGB	**HSV**	**CMY**
Black	**[0,0,0]**	**[any, any, 0]**	**[255,255,255]**
White	**[255,255,255]**	**[any, any, 255]**	**[0,0,0]**
Orange	**[255,127,0]**	**[21 (=255/12) ,255,127]**	**[0,128,255]**
Pink	**[255,>127,>127]**	**[0,any,>127]**	**[0,<127,<127]**

Hint: Experiment using PowerPoint "object fill" feature.

2. Apple RGB colours are converted to CIE RGB values using the following (rounded-up) matrix (source: http://www.brucelindbloom.com/index.html?Eqn_RGB_XYZ_Matrix.html)

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.45 & 0.32 & 0.18 \\ 0.24 & 0.67 & 0.08 \\ 0.03 & 0.14 & 0.92 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Compute XYZ values for the RGB colour vector $\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 250 \\ 200 \\ 100 \end{bmatrix}$

$$\begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \\ \boldsymbol{Z} \end{bmatrix} = \begin{bmatrix} \mathbf{195} \\ \mathbf{202} \\ \mathbf{128} \end{bmatrix} \textbf{(rounded)}$$

3. Define a 9-long RGB colour map specifying red hue of decreasing saturation (from full to zero).

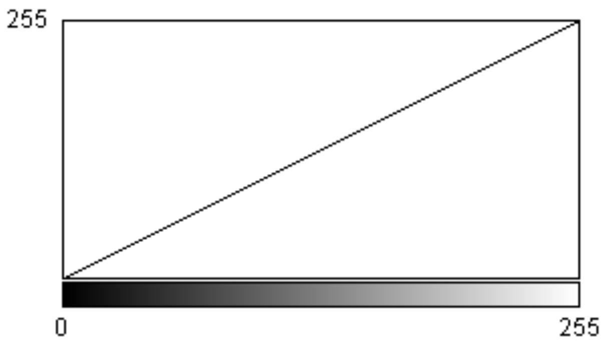
R	**255**	**255**	**255**	**255**	**255**	**255**	**255**	**255**	**255**
G	**0**	**32**	**64**	**92**	**128**	**160**	**192**	**224**	**255**
B	**0**	**32**	**64**	**92**	**128**	**160**	**192**	**224**	**255**

4. The LUT for a grey-scale image is shown below. Define the LUT to change the image into its negative (i.e. black becomes white, white becomes black, dark colours become bright and vice versa).

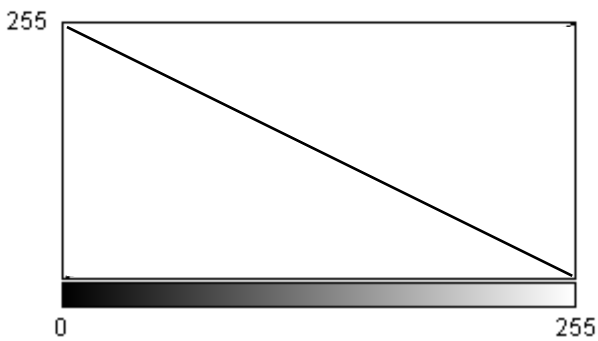
R	0	32	64	92	128	160	192	224	255
G	0	32	64	92	128	160	192	224	255
B	0	32	64	92	128	160	192	224	255

R	**255**	**224**	**192**	**160**	**128**	**92**	**64**	**32**	**0**
G	**255**	**224**	**192**	**160**	**128**	**92**	**64**	**32**	**0**
B	**255**	**224**	**192**	**160**	**128**	**92**	**64**	**32**	**0**

5. The LUT for a grey-scale image can be represented in the form of the mapping function graph showing the relationship between image values (x-axis) and shades of grey (y-axis), as below.



Draw a mapping function graph representing changing an image into its negative.

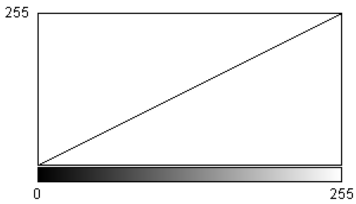


6. What is the gamut size for indexed colour images where each of the RGB primaries is represented by 4 bits per pixel and LUT has 256 entries?

Gamut is the total number of colours available for use, irrespective of the size of the LUT.

Each of the primaries can have 2^4 = 16 values, so for 3 primaries the gamut is 2^{3x4} = 4096.

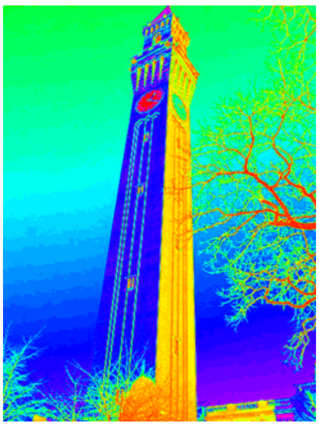
7. The image “Old Joe” and its mapping function are shown below.



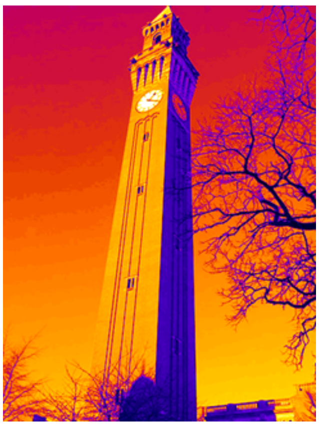
The image was displayed in false colours by changing its colour mapping function.

For the three false colour images A, B,C choose their correct mapping functions M1, M2, M3.

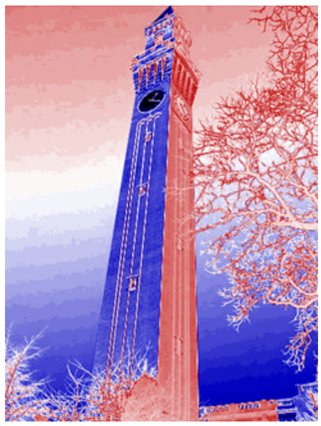
(This question can be difficult / impossible for people affected by “colour blindness”).



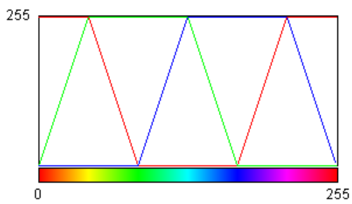
A



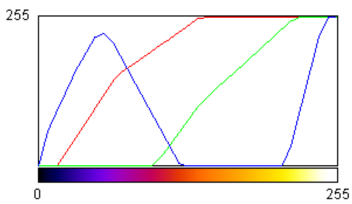
B



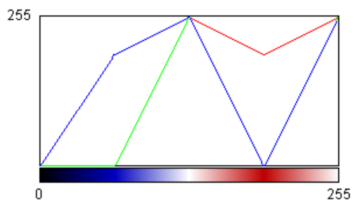
C



M3



M2



M1

A - M3, B - M2, C - M1

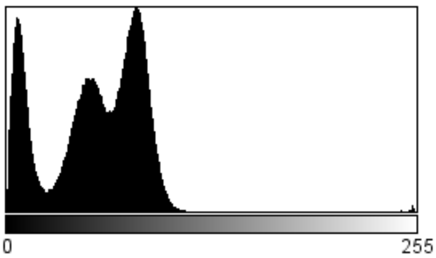
Exercise 5: Image enhancement, part 1

1 February 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

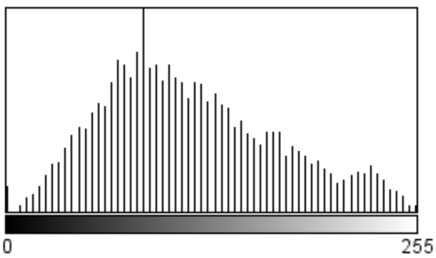
The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

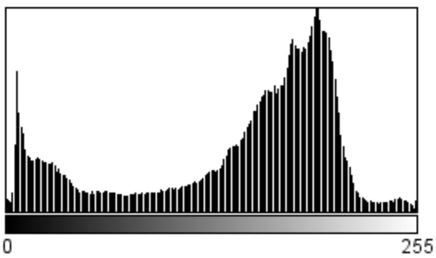
1. Given a histogram below, characterise its contrast (low - high?) and dynamic range (small - large?). Justify your answer.



Low contrast (most pixel values are between 0 and 255), high dynamic range (there are some pixels with values around 255, you can see a small peak on the right of the histogram).

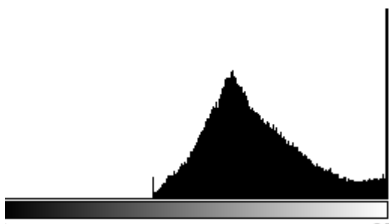
2. Given two histograms below, which one corresponds to an image with a better contrast? Why?





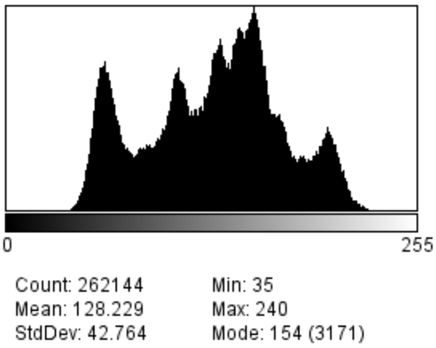
The one on the right because it has a significant number of dark and of bright pixels. In the one on the left pixel values are clustered around a dark grey colour.

3. What can you say about the image with histogram as below?



The image is *overexposed* because it has a large number of pixels with a single maximum value (a large peak on the right). Overexposed images cannot be easily corrected, whereas under-exposed can (e.g. by histogram stretch or gamma correction).

4. Given a histogram below suggest mathematical equation(s) that would maximally expand its dynamic range. The equations should include specific numerical parameters.

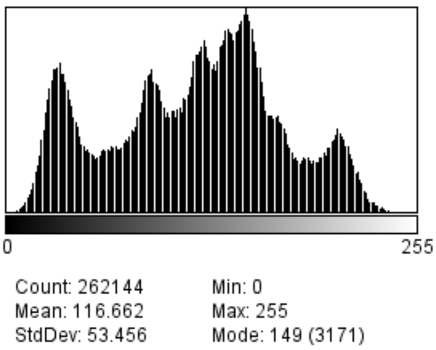


The image will benefit from shift and stretch:

- **Shift image values so that the minimum value is zero (currently 35, see the text under the histogram): I' = I - 35**
- **Stretch image values so that the maximum value is 255. The original max is 240. After the shift the max is 240 - 35 = 205. By multiplying image values by 255/205 = 1.24 the max will be 255 whereas min will remain zero: I'' = (255/205) · I'.**

The equation combining both operations is I'' = (255/205) · (I - 35).

After the transformation the resulting histogram is



5. Explain briefly why the mathematical operation(s) that you have specified above improve the dynamic range.

The image values span the full range of available grey levels.

6. How does the histogram equalisation works? In particular, *how* does it spread the image values across the whole range of possible image values?

Image values are re-distributed across the whole range (like in the shift-and-stretch), but in addition they are spaced in proportion to the number of pixels with a given value in the original. This visually enhances the dynamic range by lightening underexposed areas and darkening overexposed areas.

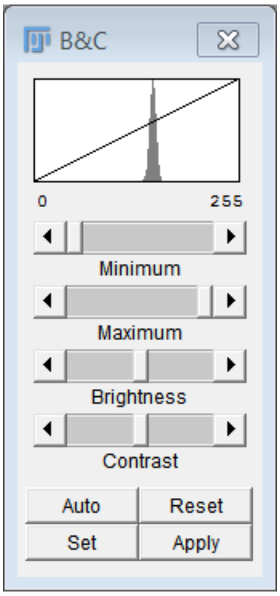
7. The equation specifying the gamma correction has the form

$$I'(x,y) = C \cdot I(x,y)^{\frac{1}{\gamma}}$$

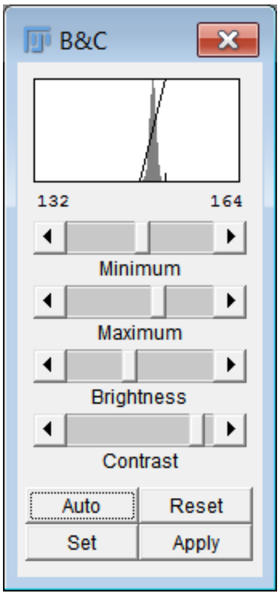
In the original image γ = 1.0. If you wanted to make dark regions lighter, would you make γ larger or smaller?

Smaller (greater than zero and smaller than one).

8. Image contrast can be improved by either changing image values (e.g. as in question 3 above), or by changing the LUT. Given an image and its current LUT, sketch the shape of the LUT that would improve the image visual appearance.



Original



Improved

In the improved image a narrow range of image values (132-164) is displayed with the full range of grey levels (0-255). In the original image only grey levels in the range 132-164 were displayed making it of a very low visual contrast.

Exercise 6: Image enhancement, part 2

7 February 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

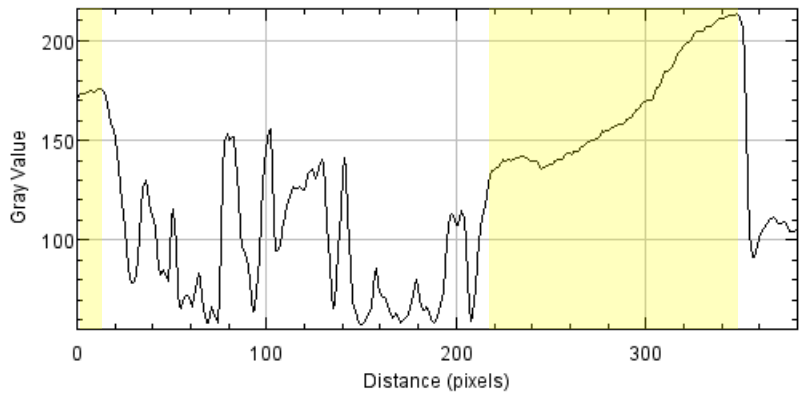
1. Why images acquired in low-light conditions often have noisy / gritty appearance? How can this problem be rectified by changing image acquisition parameters?

   Read an article at https://photographylife.com/photo-noise-reduction-tutorial.

   How can this problem be rectified by image processing?

   Noise can be reduced by low-pass filtering, but this may introduce blur at boundaries. We shall look at more advanced methods in the future lectures.

2. On the image profile below point out regions showing low frequency.



3. In terms of frequency manipulation, what does low-pass filtering do?

   It removes high-frequency components and lets through low-frequency components of an image.

4. Given the image fragment $\mathcal{I}$ below compute the result of its convolution with kernel $\mathcal{K}$ for the area highlighted grey in the image grid $\mathcal{J}$.

Image fragment $\mathcal{I}$

84	81	97	80	79
101	111	126	113	87
114	138	116	104	86
61	67	70	85	110
97	83	99	117	112

Convolution kernel $\mathcal{K}$

0.1	0.1	0.1
0.1	0.2	0.1
0.1	0.1	0.1

Image grid $\mathcal{J}$

Result

	108	**109**	**100**	
	104	**105**	**100**	
	91	**95**	**98**	

Example calculation for the top-left corner of the highlighted grid

0.1*84 + 0.1*81 + 0.1*97 + 0.1*101 + 0.2*111 + 0.1*126 + 114*0.1 + 138*0.1 + 116*0.1 =

8.4 + 8.1 + 9.7 + 10.1 + 11.1 + 12.6 + 11.4 + 13.8 + 11.6 =

107.9 → rounded to 108

(hint: use Excel).

5. Why is convolution referred to as a “linear transformation”?

Convolution is computed using multiplication and addition. A simple way to remember this is to think about the fact that these are mathematical operations which are sufficient to define the equation of a line:

$$y = m \cdot x + b$$

For a more advanced answer see e.g. http://mathinsight.org/linear_transformation_definition_euclidean

6. What are the numerical characteristics of coefficients of a low-pass filter kernel?

- **The values of kernel coefficients add up to 1 (applied to the area of constant value the filter returns this value).**
- **The coefficients take positive values from zero to less then 1 (the are fractional weights applied to image values with which they are multiplied).**
- **The coefficient in the kernel’s centre is greater or equal to the remaining coefficients (the same or larger weight is attached to the central pixel than the surrounding pixels; the closer the pixels under the kernel are to the centre the more influence they have on the computed final value).**

7. What is an unwanted effect of low-pass filtering (smoothing)?

Blur.

Can you explain this effect in terms of frequency processing?

Blurring occurs when previously sharp edges are “spread” over a certain distance. Sharp edges are high frequency features (change in image values occurs over a small distance), and low-frequency filters remove high-frequency features.

Exercise 7: Image enhancement, part 3

8 February 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

1. Name common causes of image blur.

   Poorly focused image
   Camera motion (e.g. image taken from a moving car)
   Light scatter (e.g. fog or smoke)

2. In what way blur affects image frequencies?

   It removes or decreases high frequency components.

3. In terms of frequency manipulation, what does high-pass filtering do?

   High pass filter reduces or suppresses low-frequency components and emphasises high frequency components.

4. Given the image fragment $\mathcal{I}$ below compute the result of its convolution with kernel $\mathcal{K}$ for the area highlighted grey in the image grid $\mathcal{J}$.

Image fragment $\mathcal{I}$

104	121	136	180	179
101	111	126	113	187
104	138	116	104	86
61	67	70	85	110
97	83	99	97	102

Convolution kernel $\mathcal{K}$

-1	-1	-1
-1	8	-1
-1	-1	-1

Image grid $\mathcal{J}$

	-58	-11	-210	
	348	114	-61	
	-232	-229	-104	

(hint: use Excel's function SUMPRODUCT).

5. What are the numerical characteristics of coefficients of a off-centre on-surround high-pass filter kernel?

   - **The values of kernel coefficients add up to 1 (applied to the area of constant value the filter returns zero, i.e. no value changes = zero frequency of change).**
   - **The coefficient in the kernel's centre is a whole number smaller than zero (it attaches high weight to low image values).**
   - **The other coefficients take positive values greater than zero (they respond to high image values).**

6. What is an unwanted effect of high-pass filtering?

   Noise enhancement.

   Can you explain this effect in terms of frequency processing?

   Noise is a high-frequency feature (i.e. rapid change of image values per unit of space). It consists of isolated pixels with values contrasting with their background. On-centre off-surround filter will assign highest values to bright noise pixels, off-centre on-surround will generate high values for dark noise pixels.

Exercise 8: Image enhancement, part 4

21 February 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

1. Explain the principle of difference filters. Write equations for (a) the vertical filter and (b) the left-diagonal filter.

   Difference filters subtract values of neighbouring pixels in a given direction (vertical, horizontal, right-diagonal and left-diagonal).
   (a) Vertical filter equation: I'(x,y) = I(x,y+1) - I(x,y)
   (b) Left-diagonal filter equation: I'(x,y) = I(x-1,y-1) = I(x,y).

2. What is the difference between an isotropic and an anisotropic edge detection filter? Give one example of each.

   Dictionary definition for isotropic: "having a physical property which has the same value when measured in different directions".

   An isotropic filter detect edges in all the directions simultaneously; example: Laplacian / high pass filter.

   An anisotropic filter detect edges in one particular direction; examples: gradient filters (Roberts, Sobel), difference filters.

3. What does this filter do? What image features give the strongest response (i.e. highest positive value)?

-1	-2	-1
0	0	0
1	2	1

   It detects horizontal edges (change of brightness in vertical direction). The strongest response is for transitions from a dark region lying above a light region, e.g.

4. Given the image fragment I below calculate the result of its convolution with kernel $\mathcal{K}$ for the area highlighted grey in the image grid $\mathcal{J}$.

Image fragment I

100	100	100	100	50
100	100	100	100	50
100	100	100	50	50
100	100	50	50	50
50	50	50	50	50

Convolution kernel $\mathcal{K}$

2	1	0
1	0	-1
0	-1	-2

Image grid $\mathcal{J}$

	0	100	200	
	100	200	200	
	200	200	100	

(hint: use Excel's function SUMPRODUCT).

5. Image fragment I below contains a bright line on a darker background. Would the filter shown in $\mathcal{K}$ detect the line (i.e. give high response when I is convolved with $\mathcal{K}$)?

Image fragment I

50	50	50	50	100
50	50	50	100	50
50	50	100	50	50
50	100	50	50	50
100	50	50	50	50

Convolution kernel $\mathcal{K}$

2	1	0
1	0	-1
0	-1	-2

No, it will give zero response because pixel values above and below the line are identical and when convolved with the filter they will cancel out. The result of convolution of I with $\mathcal{K}$ is shown in the image grid $\mathcal{J}$:

	-100	-100	0	
	-100	0	100	
	0	100	100	

6. Given the image fragment I below predict the result of applying a 3 x 3 median filter for the area highlighted grey in the image grid J. Confirm your prediction by carrying out the calculations.

Image fragment I

50	50	50	50	100
50	50	50	100	50
50	50	100	50	50
50	100	50	50	50
100	50	50	50	50

After median filtering all the values in the highlighted area will be 50. This is because the median of each 3x3 region is 50. The results is shown in the image grid J:

	50	50	50	
	50	50	50	
	50	50	50	

7. Briefly explain the principle and the method of the edge preserving smoothing.

Principle: The idea is that smoothing filter is NOT applied ACROSS a boundary thus avoiding is blurring.

Method: For each image pixel its neighbourhoods in 9 directions are examined in turn (in compass notation: N, NE, E, SE, S, SW, W, NW, and area centred around the given pixel). For each neighbourhood standard deviation is calculated. Standard deviation would be highest across a boundary and smallest in a sub-region lying on one side of the boundary. The value of the pixel is then replaced (in a new image) by the average of pixel values in the region with the smallest standard deviation. When the edge preserving filter is applied in an image region where boundaries are not present, it will carry out more-less normal smoothing.

Exercise 9: Image segmentation

22 February 2018

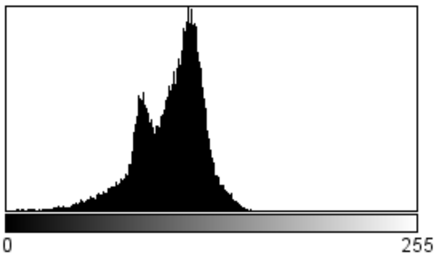
This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

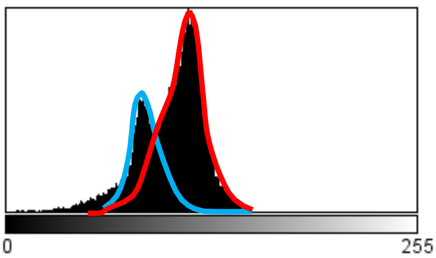
1. Why is it said that image segmentation by extracting object boundaries and by extracting regions is equivalent?

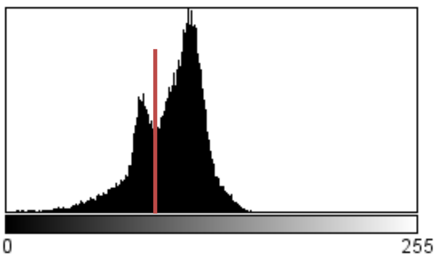
   If edges are extracted, image pixels within the edge boundaries can be identified.
   If regions are extracted, edges of the regions can be detected using edge detection filters.

2. Given the image histogram below, indicate the location of the threshold value you would choose to separate object(s) from the background. Justify your choice.



   Histogram shows the statistics of image values in a graphical format. Typically the populations of pixels belonging to background and foreground have normal distribution, and their distributions have different mean values. On a histogram each normally distributed population will have the shape of a Gaussian, and will cover a different range of grey level values. In most cases the distributions will overlap, so there separation between the two populations will not be complete. The best choice for the threshold is where the overlap is minimal, i.e. in the valley between the two Gaussians.





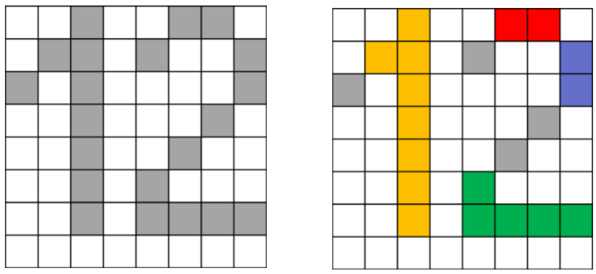
3. How is image segmentation computed for dark objects on a bright background?

```
Decide on a threshold value T
Create a new empty image I’
For every pixel in the original image I
    if I(x,y) < T
            set I’(x,y) to 1
    otherwise
            set I’(x,y) to 0.
```

4. One of the outputs of image segmentation can be a “label image”. What is represented by pixel values in the label image?

Pixel values correspond to object numbers. Background is zero. All the pixels corresponding to the same detected object have the same value. For example, if there are five objects, the maximum value in the label image will be five.

5. In the grid below mark all clusters of pixels that are 4-connected but not 8-connected.



Each cluster of 4-connected pixels is marked with a different colour. Grey pixels in the figure o the right are not 8- connected.

6. Otsu thresholding is one of the most commonly used methods for automatic threshold selection. Briefly explain its principle.

Otsu method is a binary thresholding method, i.e. image is segmented into two regions, normally “object(s)” and “background”. Threshold is chosen such that the variance <u>within</u> each region is <u>minimised</u> whereas variance <u>between</u> the two regions is <u>maximised</u>. (Variance is standard deviation squared.)

7. What is the key difference between global and local segmentation methods?

Global methods classify all pixels simultaneously, on the basis of the statistics of the whole population of pixels.
Local methods classify pixels or their groups sequentially, often taking into account the result of classification of the pixels already classified.
Results of global classification will always be the same (given the same criterion) whereas in local classification the results may depend on the choice of a starting point (e.g. in region growing).

Exercise 10: Post-processing

1 March 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

1. Erosion and dilations are the basic operations of mathematical morphology. Explain how they work and what effect they have on a binary image.

   Dilation (expansion)
   adding a "layer" of pixels to the periphery of objects
   the object will grow larger, close objects will be merged, holes will be closed.

   Erosion (shrinking)
   removing a "layer" of pixels all round an object
   the object will get thinner, if it is already thin it will break into several sections

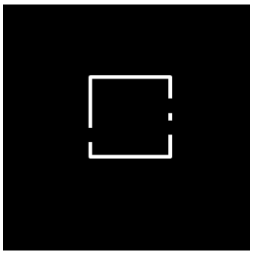
2. After thresholding, the image regions belonging to objects ended up having "holes" due to image noise. How would you remove the holes using mathematical morphology methods? Object pixels have value one, background pixels have value zero.

   Dilation would close the holes by adding a layer (or layers) of pixels to the object edges (including edges of the holes).

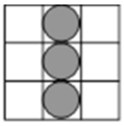
3. What is the mathematical morphology operation "opening"? What is it useful for?

   Opening is erosion followed by dilation. It is useful for smoothing peripheries and removing small features

4. Due to noise in the image, edge detection filter produced gaps in the edge outline. Can you repair the gaps (without upsetting the rest of the outline) by using mathematical morphology methods?

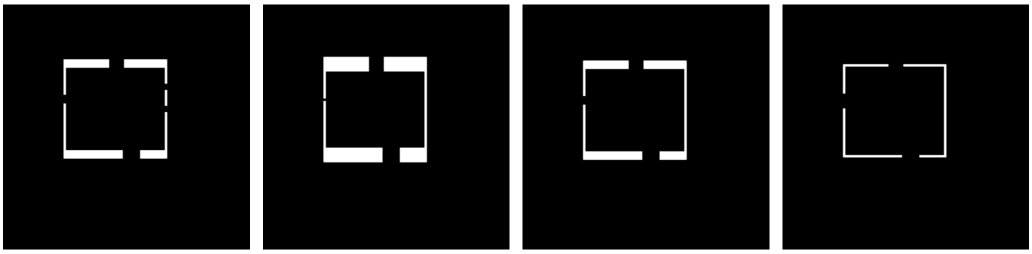


Closing with the vertical 3x3 structuring element will do the job.



Steps:

- **Apply dilation several times until the border is closed. As a side-effect the top and the bottom edges will expand and will become thicker.**
- **To compensate, and make all edges thin again, apply erosion with the same structuring element the same number of times.**



Dilation 1 Dilation 2 Erosion 1 Erosion 2

5. You were given an image of a car registration plate as below. Your task is to extract just the edges of the plate using only thresholding and the basic methods of mathematical morphology.

- **Generate a binary image by thresholding: assign 1-s to the plate and zeros elsewhere.**
- **To remove the lettering, dilate the binary image several times until all the letters and numbers disappear. The plate will be enlarged at the same time.**
- **To get it back to the original size, erode the dilated image the same number of times with the same structuring element.**

- **To get the edges of the plate, make a copy of the above image and erode it once. It will shrink. Subtract the eroded image from the above one.**

Dilation 1 Dilation 3 Erosion 3 times, one more erosion, subtraction

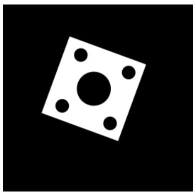
Exercise 11: Object properties: counting, measuring and localisation

7 March 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

1. The “hand-on-the-wall” algorithm is a simple method for outlining binary images. Is it suitable for outlining inner edges of objects with holes, such as the one below? Why?



No. This is because the holes are not accessible to the first part of the algorithm which scans the image line by line looking for the first “object” pixel.

2. After segmentation a labelling operation was applied to a binary image, generating a “label image”. What is represented by pixel values in the label image?

Pixel values correspond to object numbers. Background is zero. All the pixels corresponding to the same detected object have the same value. For example, if there are five objects, the maximum value in the label image will be five.

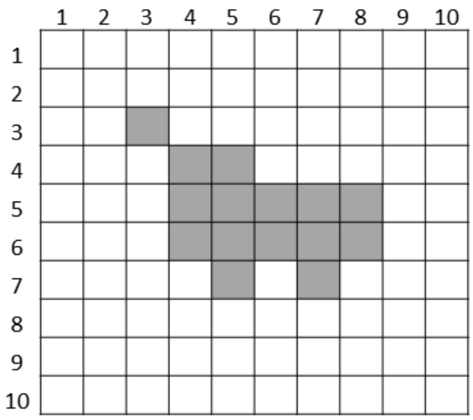
3. For a binary image fragment below enter the region ID for each pixel as it would be assigned after the ***first pass*** of the row-by-row labelling algorithm. Object pixels are indicated by a grey shade, background by white.

		1	1	1											
	2				3										
	2				3				4						
	2				3				4						
		5	5	5					4						
									4						
							6	6	4	4	4				
				7											
					8	8									
					8	8	8								
						8	8								
								9							

4. For the same image fragment enter now the region ID-s resulting from the ***second pass*** of the row-by-row labelling algorithm.

		1	1	1											
	1				1										
	1				1				4						
	1				1				4						
		1	1	1					4						
									4						
							4	4	4	4	4				
				7											
					7	7									
					7	7	7								
						7	7								
								7							

5. Compute the ***centre of mass*** for the object in the figure below. The numbers above and to the left of the grid are pixel coordinates.



Sum of pixel x-coordinates = 80
Sum of pixel y-coordinates = 84
Number of pixels in the object = 15
Centre of mass coordinates: (80/15, 84/15) = (5.3, 5.8)nearest pixel is (5,6)

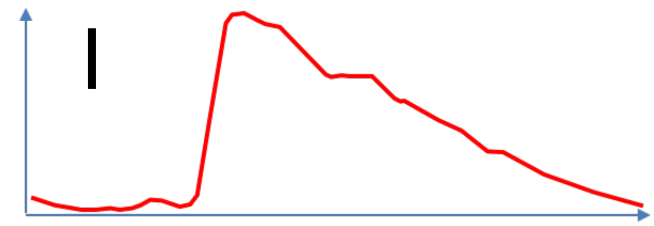
Exercise 12: Object properties: shape and texture

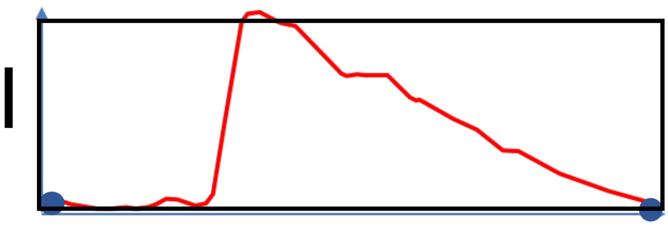
8 March 2018

This unassessed exercise gives you an opportunity to test your understanding of the material covered in the lectures and to encourage you to look at the materials suggested in further reading and exploration. You do not need to submit it, but you should write down the answers and keep them for future revisions.

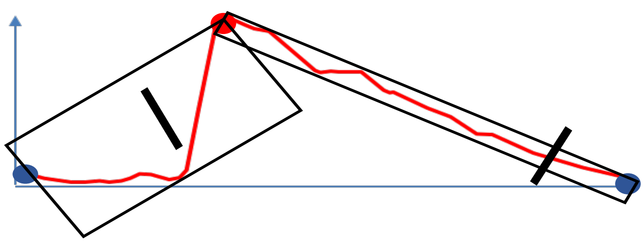
The exercise will not be marked, instead the outline solutions, answers and pointers to external sources will be provided one week after the exercise has been issued.

1. The curve below reproduces a red curve from slide 6, lecture 11. The minimum width of the bounding box (tolerance band) is shown as a vertical black bar. Show the steps leading to approximating this curve by a polyline according to the tolerance band.

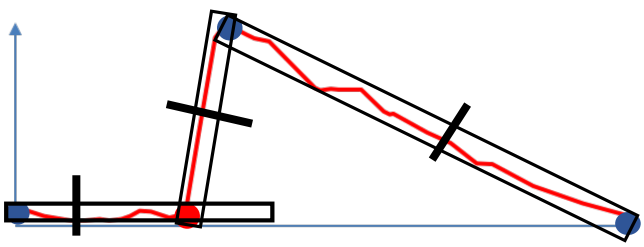




Max distance greater than the tolerance band. Keep dividing.



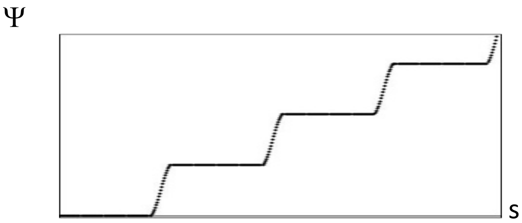
Max distance greater than the tolerance band for the first subdivision, keep dividing.
Max distance smaller than the tolerance band for the second subdivision, stop dividing.



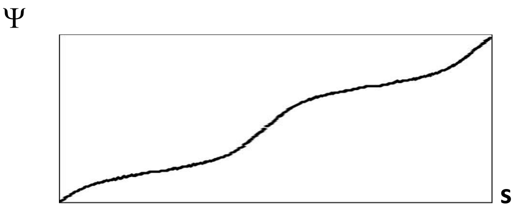
Max distance smaller than the tolerance band for all the subdivisions, stop dividing. Task complete

2. Sketch Ψ-s curve for

(a) Square



(b) Ellipse



3. Compute the co-ccurrence matrix at one pixel distance in **vertical** direction.

1	0	0	0
1	0	3	3
1	2	3	1
1	3	1	3

	0	**1**	**2**	**3**
0	**1**	**0**	**1**	**2**
1	**0**	**3**	**0**	**1**
2	**0**	**0**	**0**	**1**
3	**0**	**2**	**0**	**1**

4. What is the size of a co-occurrence matrix for an image with radiometric resolution of 8 bits per pixel?

The size of a co-occurrence matrix depends only on the number of grey levels, it does not depend on the number of pixels in the image.
For radiometric resolution 8 bits per pixel there are 256 grey levels.
The co-occurrence matrix has the size 256 x 256 (the count of pixels with every possible grey level (out of 256) co-occurring with a pixel with some other grey level (again, 256).