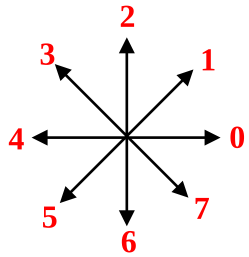
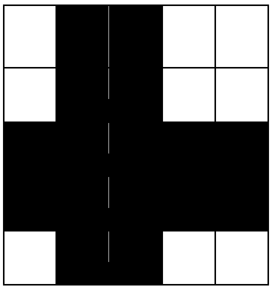
Some exercises from previous exams in INF4300/9305 Digital Image Analysis

with solution hints in red

Exercise 1, 2012: Chain Codes

You are given the 8-directional chain code and the object below, where black is object pixels.





a) Chain code the boundary of the object clockwise from the upper left pixel.

Answer: The absolute code starting at the upper left point and moving clockwise is: 067064543212.

b) Which technique will make the code invariant to the choice of start point? Demonstrate this by starting at lower right pixel of the object.

Answer: **a minimum circular shift of the clockwise absolute chain code**.
Demonstration: Minimum circular shift of absolute chain code starting at the top left point is: 064543212067.
The absolute code starting at the lower right and moving clockwise is 432120670645, and the minimum circular shift of this is: 064543212067. QED.

c) Which technique will make the code rotation invariant? Demonstrate this by rotating the object $\pi/2$ counterclockwise and start at one of the same object points as above.

Answer: **a minimum circular shift of the first difference (counterclockwise)**.
Demonstration: Starting at what was the upper left pixel, the absolute code is 201206765434.
The first difference is 611661777716, which gives a minimum shift of 116617777166.
The first difference of absolute code 067064543212 is 611661777716, and minimum shift of this is 116617777166, QED.

Exercise 2, 2012: GLCM

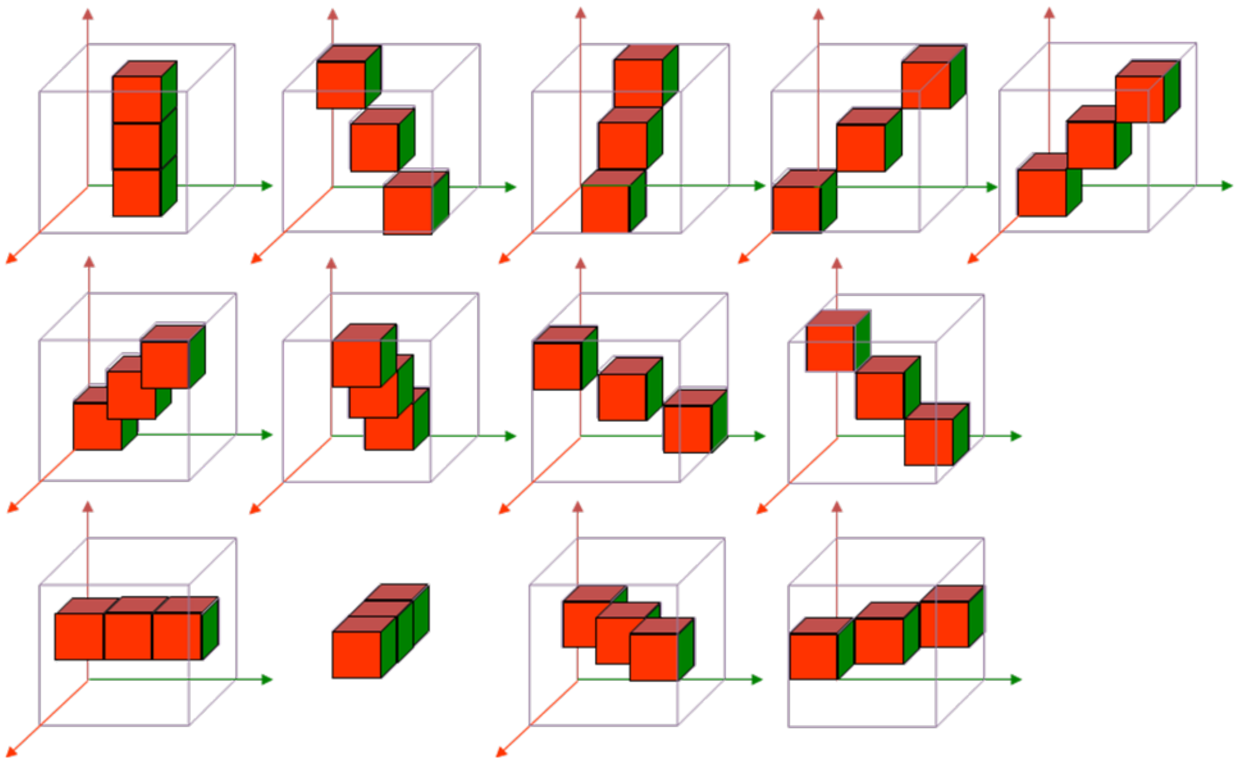
The Gray Level Cooccurrence Matrix method is a frequently used texture analysis method. In 2D gray level images we often use isotropic GLCMs based on symmetric matrices obtained for a given pixel distance in a limited number of directions.

a) How do you obtain a normalized symmetric Gray Level Cooccurrence matrix for a given (d,θ)?

Answer: Count the number of times that pixel pairs with gray levels zi and zj occurr in the image when moving a distance d in direction θ. This gives $P(i,j|d,\theta)$, which can be normalized by $c(i,j|d,\theta)=P(i,j|d\theta)/S$, where S is the sum of all elements in $P(i,j|d,\theta)$.
The normalized GLCM for the opposite direction, $c(d,\theta+\pi)$ is given by the transpose $c(d,\theta)^T$.
Adding the two results in a normalized symmetric matrix: $C(\theta) = [c(d,\theta) + c(d,\theta)^T]/2$.

b) In a 3D image, how many Gray Level Cooccurrence matrices do you need to accumulate to make an isotropic matrix describing the second order statistics covering all directions to the neareast neighbors of an image voxel (=volume pixel)?

Answer: The nearest neighbors of a voxel in 3D are contained in a 3x3x3 neighborhood, containing 27 voxels. So the central voxel has 26 neighbors. Half of these, 13, are in the opposite direction of some other neighbor.
So the answer is13, not counting the transposed matrices.
The 13 directions are illustrated below.



c) If you treat the GLCM as an image C(i,j), and compute the central moments $\mu_{1,0}$ and $\mu_{0,1}$, what are the values of these moments?

Answer: They are both zero. This can either be seen from the fact that the ordinary moments $m_{1,0}$ and $m_{0,1}$ give us the center of mass coordinates by

$$\bar{i} = \frac{m_{1,0}}{m_{0,0}}, \quad \bar{j} = \frac{m_{0,1}}{m_{0,0}}$$

Computing the central moments μ_{pq} is equivalent to first moving the object or the coordinate system so that the center of mass is in the origo, and then computing the ordinary moment, m_{pq}. And the m_{pq}'s must be zero when the distance to the center of mass is zero.

Alternatively, we can compute μ_{10} and μ_{01}:

$$\mu_{1,0} = \sum_i \sum_j (i - \bar{i}) C(i,j) = \sum_i \sum_j i C(i,j) - \bar{i} \sum_i \sum_j i C(i,j) = m_{1,0} - \bar{i} m_{0,0} = \underline{\underline{0}}$$

d) If you compute the three second order central moments of the normalized Gray Level Cooccurrence matrices P(i,j|d, θ, ω) from a number of texture classes ω, which parameters derived from these three moments would give you first and second order statistics that are useful in order to classify the textures?

Answer: The three second order moments $\mu_{0,2}$, $\mu_{2,0}$, and $\mu_{1,1}$ can be used to give you the orientation and semi-axes of an ellipse that is fitted to the distribution within the GLCMs.

The orientation of the fitted ellipse (we do not expect the equation to be reproduced!)

$$\theta = \frac{1}{2} \tan^{-1} \left[\frac{2\mu_{11}}{(\mu_{20} - \mu_{02})} \right], \quad where \;\; \theta \in \left[0, \pi/2\right] if \;\; \mu_{11} > 0, \;\; \theta \in \left[\pi/2, \pi\right] if \;\; \mu_{11} < 0$$

will be approximately the same for all matrices (along the diagonal), so this will probably be useless as a classification feature.

The major semi-axis will characterize the width of the range of gray-levels in the image, and is a first-order statistic, correlated to the variance obtained from the image histogram. This may be useful.

The minor axis will characterize the (lack of) correlation between pixel pairs at the specified (d,θ), and is thus a useful second order parameter (we do not expect the equation for the semi-axes to be reproduced!).

$$(\hat{a}, \hat{b}) = \sqrt{\frac{2\left[\mu_{20} + \mu_{02} \pm \sqrt{(\mu_{20} + \mu_{02})^2 + 4\mu_{11}^2}\right]}{\mu_{00}}}$$

e) Given a NxN pixel gray level image having G gray levels (0,…, G-1), explain how you can use a non-normalized Gray Level Cooccurrence Matrix to estimate the graylevel threshold that will give you a binary output image that contains a minimum number of transitions from 0 to 1 or 1 to 0 as you traverse the binary image in the horizontal direction.

Answer: A matrix element of the non-normalized cooccurrence matrix c(i,j|dx,dy) contains the number of times you go from gray level i to gray level j by stepping dx pixels in the x-direction and dy pixels in the y-direction. Using (dx,dy)=(1,0) we check the nearest neighbor horizontally. Going in the opposite directions we get the transposed matrix: $c(i,j|-1,0) = c(i,j|1,0)^T$. We can sum these two matrices into a single matrix C(i,j).

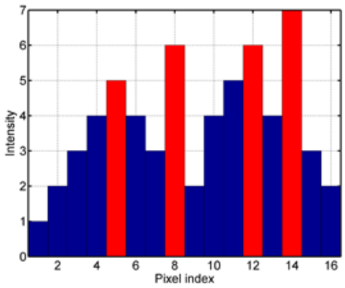
Putting a threshold at i=j=T, we find that the number of cases where one pixel in a pair of horizontal neighbors in the binary image resulting from a thresholding at T belongs to the background while the other belongs to the foreground:

$$O_{bf} + O_{fb} = \sum_{i=0}^{T} \sum_{j=T+1}^{G-1} C(i,j) + \sum_{i=T+1}^{G-1} \sum_{j=0}^{T} C(i,j)$$

Thus, we simply search for the value of T that minimizes the equation above.

Exercise 3, 2012: Watershed segmentation

a) What characterizes a pixel where the watershed algorithm will construct a dam? In the intensity profile below, please mark where the watershed algorithm will place a dam.



The pixel does not belong to only one catchment basin, and a dam must be constructed to prevent flooding.

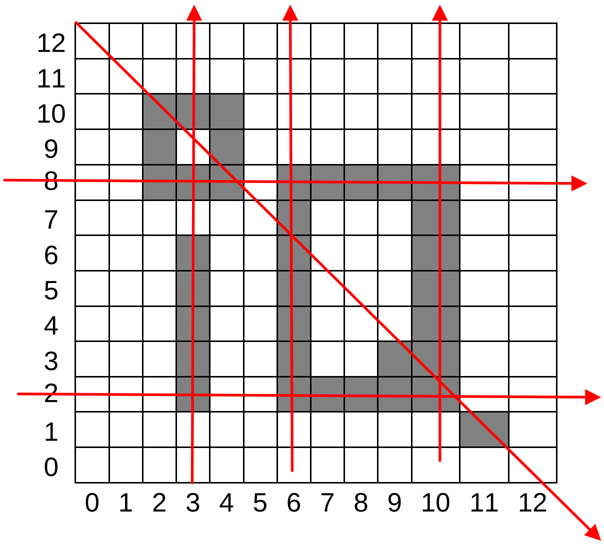
b) Assume that you are going to segment some round objects. You expect that the initial thresholding will produce results as illustrated below (white is foreground pixels.).

We want to separate the two components of this object into two approximately elliptical objects. Describe how this can be done with the help of the watershed algorithm.

Answer: compute the distance transform, i.e. the distance from every foreground pixel to the closest background pixel. Apply the watershed transform to the inverted distance image, and split the object along the watershed dam lines. The image below illustrates the solution.

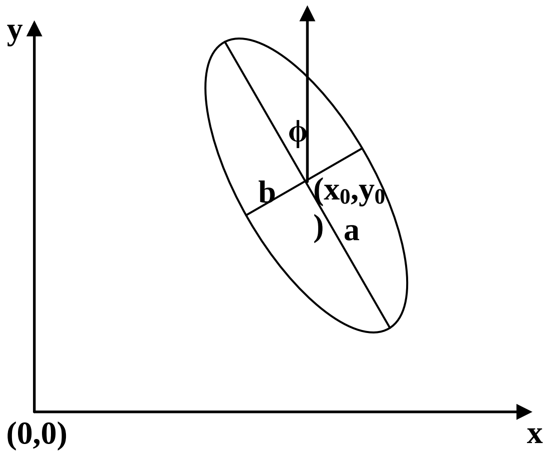
Exercise 6, 2012, PhD-students: The Hough Transform

a) Apply the Hough transform to the 12x12 pixel binary image below, using the normal representation, and mark the accumulator values of the SIX most prominent maxima in their proper (θ,ρ)-locations within the accumulator matrix given in the enclosure. We assume that the D_8 (*chess*) distance is used to measure distance from (x,y)=(0,0).



Answer: The six longest sets of co-linear pixels are markes by arrows in the image above.
The accumulator values of the maxima corresponding to these lines are found in their proper (θ,ρ)-locations the accumulator given in the enclosure sheet.
One should realize that the Hough transform considers co-linear pixels, not only consecutive pixels!

b) A fairly common Hough representation of an ellipse in the (x,y)-image plane is 5-dimensional: x_0 and y_0 to describe the centre of the ellipse, ϕ to describe the orientation (anticlockwise from the vertical y-axis), and the pair (a,b) to describe the length of the semi-axes of the ellipse.



Someone suggests using the normal (θ,ρ) - representation of the major axis, together with (a,b), in order to represent an ellipse.
Please indicate the parameters θ and ρ of the normal representation of the major axis.

c) If the major axis is rotated around (x_0,y_0), which curve will the major axis of the ellipse describe in the positive (θ,ρ)-plane? Please explain.

One should realize that the Hough transform is a "one-to-many" transform, i.e. each ***pixel*** in the image plane gives rise to curves in the parameter plane.

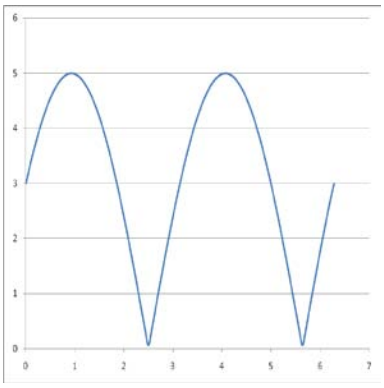
Using the "normal representation" of lines in the image plane gives rise to a set of sinusoidal curves in the (θ,ρ)-plane, each representing the family of lines that may pass through a a particular point in the (x,y)-plane (see page 734 in GW).

Correspondingly, a line in the image corresponds to a single point in parameter plane, as in exercise b) above. And if that line in the image plane is rotated around a single image point, the parameter point will move along a sinusoidal curve in the parameter plane.

So the simple answer is that as the major axis rotates around (x_0,y_0), it traces out a sinusoidal curve in the the (θ,ρ)-plane. But NOT because "a line in the image corresponds to a sinusoid in the Hough plane" – it doesn't!

A few points along the curve traced as the major axis is rotated are easy to identify:

- A vertical major axis of the ellipse corresponds to the point $(0, x_0)$ in the (θ,ρ)-plane.
- The maximum value of ρ is $\rho_{max} = (x_0^2+y_0^2)^{1/2}$ for $tg\theta = y_0/x_0$.
- A horizontal major axis corresponds to $(\pi/2, y_0)$.



One should realize that $\theta = \phi$.
And that $\cos\theta = \rho/(x_0+y_0\, tg\phi)$, from simple geometry.

We are in the positive (θ,ρ)-plane
So, the curve is given by

$$\rho = \left|\cos\phi\,\left(x_0 + y_0\, tg\phi\right)\right|$$

as illustrated to the right for $(x_0,y_0) = (3,4)$, which gives a maximum radius of 5.

d) Is this a unique representation of an ellipse?

No, it places an ellipse of the correct size and orientation on a line in the (x,y)-plane, but it does not determine WHERE on the line.

Enclosure 2, Exam, INF4300/INF9305, December 14, 2012

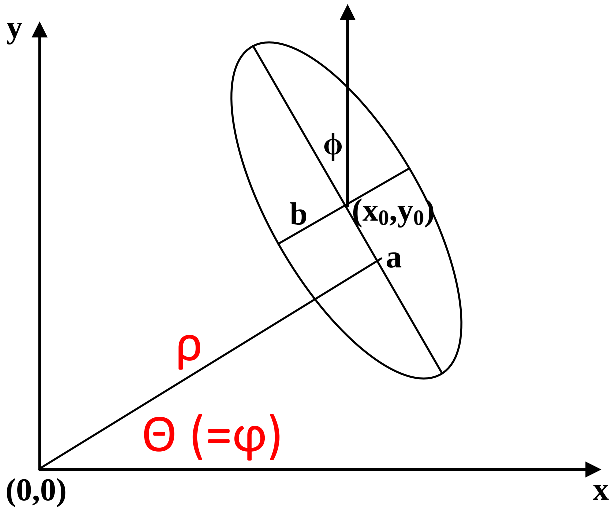
Exercise 6a, 2012: Hough Transform

Apply the Hough transform to the 12x12 pixel binary image, using the normal representation, and mark the accumulator values of the SIX most prominent maxima in their proper (θ,ρ)-locations within the accumulator matrix.

π/2			6						8				
π/4							6						
0				8			7				7		
	0	1	2	3	4	5	6	7	8	9	10	11	12

Exercise 6b, 2012: Hough Transform

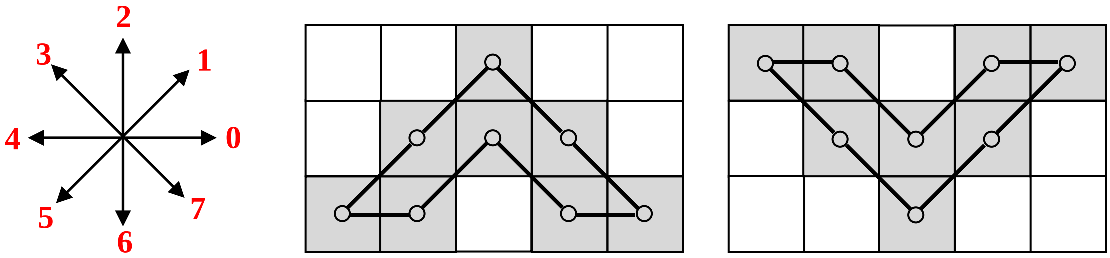
Please indicate the parameters θ and ρ of the normal representation of the major axis.



If you are a PhD-student, please tear out this page, and hand it in!

Exercise 1, 2011: Chain Codes

You are given the 8-directional chain code and the two objects below.



a) Chain code the boundary of the Λ-shaped object clockwise from the lower left pixel.

The absolute code starting at the lower left point and moving clockwise is: 1 1 7 7 4 3 5 4

b) Which technique, based on the 8-directional <u>absolute</u> chain code, can be used to make a description of the Λ-shaped object that is independent of the start point? Demonstrate this by starting at the top pixel of the object, instead of the lower left.

The simple answer is that **<u>a minimum circular shift of the clockwise absolute chain code</u>** gives a normalization for start point.
Demonstration: The absolute chain code starting at the top point is: 7 7 4 3 5 4 1 1.
The minimum circular shift of this is: 1 1 7 7 4 3 5 4, which is the same code as when we started at the lower left in a).

A **minimum circular shift of the counterclockwise first differences**
of the clockwise absolute chain code will give a normalization for rotation,
but that was not what we asked for here.

c) The V- shaped object is a rotation of the Λ-shaped object.
Which technique, based on the clockwise <u>relative</u> chain code, will give you the same description of the two objects, independent of the start point?
Demonstrate this by starting at the upper left pixel of the V-shaped object.

The simple answer is that **<u>a minimum circular shift of the clockwise relative chain code</u>** gives a normalization for start point, even when the objects are rotated.
Demonstration:
The clockwise relative code, starting at the upper left pixel of the V-shaped object, is:
7 1 4 1 7 2 0 2 . The minimum circular shift of this is: 0 2 7 1 4 1 7 2.
A different start point could be the bottom pixel of the V-ahaped object, corresponding to the top pixel of the Λ-shaped object.
A clockwise relative code, starting at the top pixel of the Λ-shaped object is: 0 2 7 1 4 1 7 2.
We do not need any circular shift to see that these codes are the same.

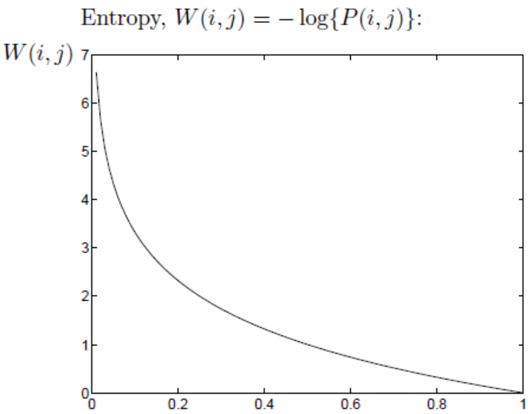
Exercise 3, 2011: GLCM

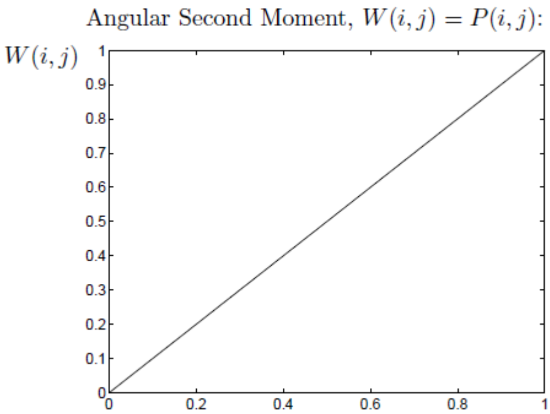
To the right you see a remote sensing image showing both steep slopes, deep valleys, flat grassland, numerous small streams, and some man-made constructions. Two subregions are incircled.

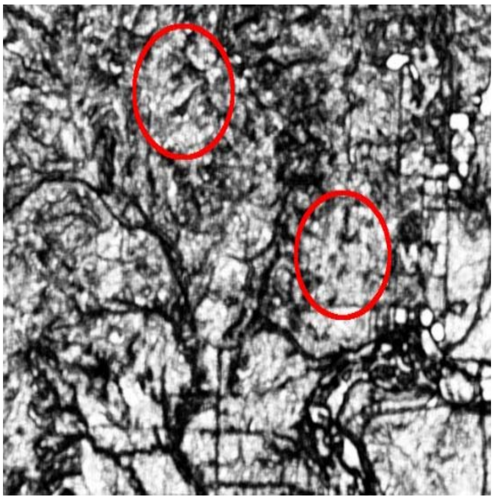
The two texture feature images at the bottom of the page were made using a moving window to compute a normalized GLCM and deriving two features that are using weight functions based on the value of the matrix elements, namely ENTROPY and ASM. The two equations are:

$$ENT = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) W_{ENT}(i,j), \quad P(i,j) > 0; \qquad ASM = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) W_{ASM}(i,j)$$

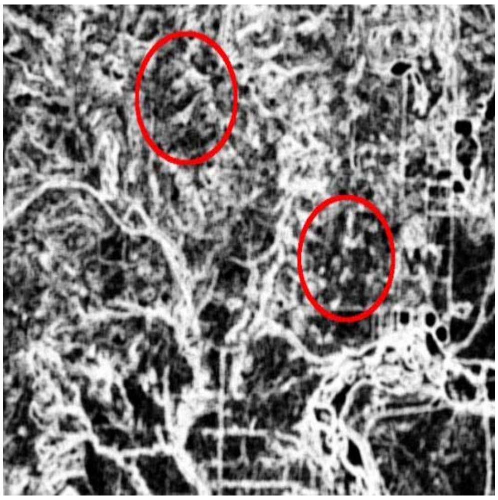
Where the weights are functions of the GLCM element probabilities, given by two curves:







Left feature image



Right feature image

a) A direction invariant normalized GLCM was computed with a pixel offset of 1. Explain the steps of this computation.

This means that we have accumulated Gray Level Cooccurrence Matrices from four directions, 0, 45, 90 and 135 degrees, all with a pixel distance of 1.
The remaining four matrices needed to cover all eight neighbors are obtained by using the transposes of the accumulated matrices.

We then add the matrices together, taking care that the matrices from the diagonal directions should have less weight (just a detail).

Finally, we normalize the matrix by dividing by the sum of all matrix elements, so that the normalized matrix elements sum to 1.

b) Based on what you observe in the circled regions and elsewhere in the two feature images, explain which is ENTROPY and which is ASM, and give some reasons.

Entropy to the right, ASM to the left, for a number of reasons:

A homogeneous (sub)image has only a few graylevels, so there are only a few entries in the GLCM having relatively high values, and the resulting entropy is low, while the sum of squares of the entries (ASM) is high.
An extremely inhomogeneous (sub)image will give a GLCM with a lot of low-value entries. This will give a high entropy, while the sum of squares (ASM) is low.

In the right hand feature image, homogeneous areas are darker than in the left hand image, including the right hand circle and the homogeneous structures to the lower right of it.
The rivers and creeks in the image – having high local contrast - stand out as high value areas in the right hand image, while they are dark in the left hand image.

c) There are also GLCM features where the weighting is a function of the element position. Two examples of this are Inertia (INT) and Inverse Difference Moment (IDM), given by

$$INT = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-j)^2 P(i,j) \; ; \qquad IDM = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{1}{1+(i-j)^2}P(i,j)$$

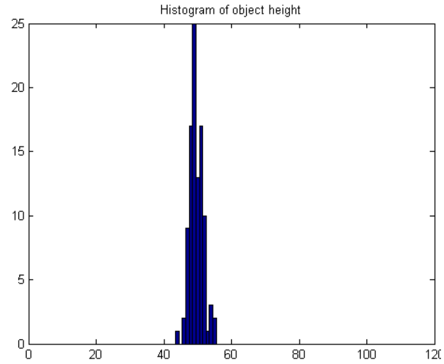
Discuss the two features in terms of local contrast and homogeneity.

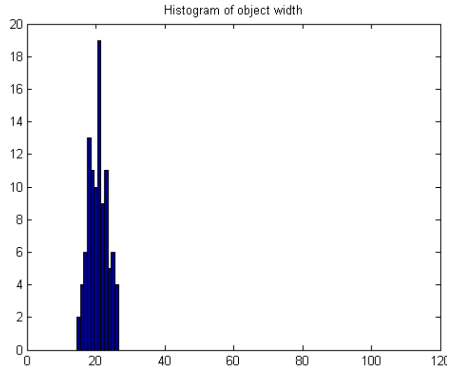
The INT weight function is zero along the diagonal (i=j), and increases rapidly away from the diagonal. Thus, it will favor GLCM elements away from the diagonal (i≠j), and give higher values for images with high local contrast.

IDM is influenced by the homogeneity of the image. The weight function has its maximum (W = 1) along the diagonal (i = j), and falls off away from the diagonal. The result is a low IDM value for inhomogeneous images (i ≠ j), and relatively higher values for homogeneous images (i≈ j).

Exercise 4, 2011: Morphology

A gray level image $f\,(x,y)$ contains 100 dark objects on a lighter background. The background has varying intensity. The dark objects have a size distribution as indicated in the figures below, showing the histograms of object width and height.





a) Which morphological operation would you use in order to estimate the varying intensity in the image?

To find dark objects on bright background: black top-hat (or bottom-hat).

b) What should the structuring element look like?

It should be bigger than the objects, so its width should be larger than (approx) 25 or height larger than (approx) 55

c) Let $g(x,y)$ be the image after a correction of the background intensity, but before segmentation by thresholding. Give an expression for $g(x,y)$ given the structuring element and the operation in a).

$(f \bullet b)-f$ in this case

d) Let $h(x,y)$ be the image after thresholding. Which morphological operation would you use to split connected objects in $h(x,y)$?

Opening (not just erosion)

e) Which morphological operation would you use to connect fragmented objects?

Closing (not just dilation)

f) Given a line in a gray level image. The line has pixel values
$f_1(x,y)$ = [1 2 3 4 5 4 3 2 1 2 3 3 3 3 3 3 3 3 3 3]
Compute $g_1(x,y)$, the result after a morphological opening by a the structuring element [1 1 1]. You can either give a graph, or give the values of $g_1(x,y)$.

Exercise 5, 2011: Moments and Hough Transform

Assume that you have thresholded a gray level image into a binary image ***b(x,y)*** containing a solid object (pixel value = 1) and background pixels (pixel value 0). We are going to use general discrete moments $\boldsymbol{m_{pq}}$ defined as

$$m_{pq} = \sum_x \sum_y x^p y^q b(x,y)$$

a) Describe a moment-based approach to find the centre of mass of the object.

We use first order moments to find the centre of mass

$$m_{10} = \sum_x \sum_y x\, b(x,y) = \bar{x} m_{00} \quad \Rightarrow \bar{x} = \frac{m_{10}}{m_{00}}$$

$$m_{01} = \sum_x \sum_y y\, b(x,y) = \bar{y} m_{00} \quad \Rightarrow \bar{y} = \frac{m_{01}}{m_{00}}$$

b) Assume that the object is rectangular and rotated, located somewhere in the image. Describe a moment-based approach to estimate the orientation of the object.

We use central moments defined by

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p \ (y - \bar{y})^q\, b(x,y)$$

And use the three second order central moments μ_{11}, μ_{20}, and μ_{02} to estimate the orientation of the object.

$$\theta = \frac{1}{2} \tan^{-1}\left[\frac{2\mu_{11}}{(\mu_{20} - \mu_{02})}\right]$$

c) The normal representation of a line segment in the (x,y)-plane is given as

$$\rho = x\cos\theta + y\sin\theta$$

Indicate the parameters $\boldsymbol{\rho}$ and $\boldsymbol{\theta}$ in the first sketch of the enclosed sheet.

See the sketch.

d) If we translate the origin to the centre of mass of the rectangle, apply a gradient detector to the binary image, and use the "normal representation" Hough Transform, we get peaks in the Hough space illustrated by ***red dots*** in the second sketch of the enclosure.

Assume that the rectangle has sides ***a*** and ***b*** (***a***>***b***), and has an orientation $\boldsymbol{\theta}$ relative to the x-axis in the image, where $-\pi/2 \leq \theta \leq \pi/2$.
Please fill in the proper parameters into the seven empty ovals in the second sketch in the enclosure, and give reasons for your choises.
Please remember to mark this enclosure with your candidate number and include it in your solution.

See the sketch.

e) What happens to the height of the peaks in the (ρ,θ)-plane as the rectangle becomes more like a square?

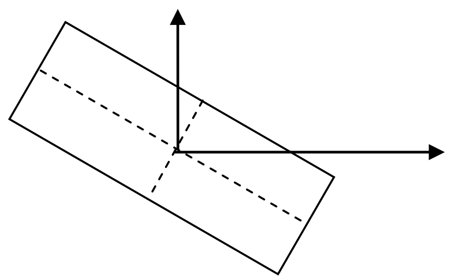
The sides of a square have the same length, so there will be the same number of pixels in the four line segments in the image. Hence, the four peaks in the parameter plane will have the same heigth.

The following two questions are intended for the PhD-students:

f) What is the (approximate) orientation of the rectangle relative to the x-axis of the image, given the positions of the peaks in the Hough space in the sketch?
Explain your reasoning, and make a simple drawing of the rectangle in the (x,y)-coordinate system.

The lines making up the rectangle are represented in Hough space by the "normal representation": the normal from the origin onto the line.
So $\boldsymbol{\theta}$ is the angle from the x-axis to this normal.

We observe that the angle $\boldsymbol{\theta_1}$ from the x-axis to the normal to the most distant sides – which must be the orientation of the rectangle – is negative and that its absolute value is about half the value of the other angle $\boldsymbol{\theta_2}$. Since $|\theta_1| + |\theta_2| = \pi/2$, and $|\theta_2| = 2|\theta_1|$,
<u>we must have $\theta_1 = \pi/6 = 60$ degrees.</u>



g) Explain what happens to the positions of the peaks in the $(\boldsymbol{\rho},\boldsymbol{\theta})$-plane
   a. If the rectangle is turned into a square of sides $\boldsymbol{s}$.
   b. As the square is rotated through 360 degrees in the (x,y)-plane.

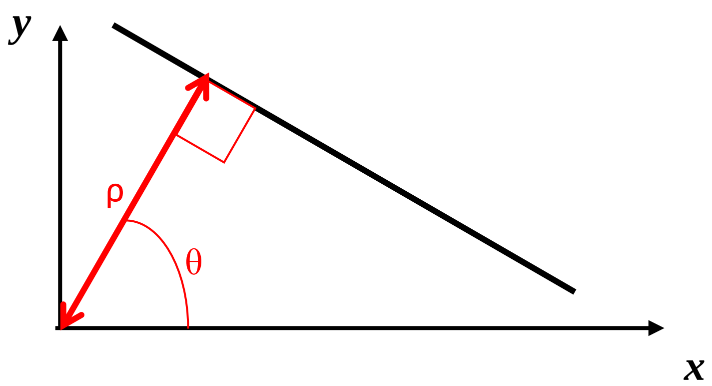
For a rectangle with sides $(\boldsymbol{a},\boldsymbol{b})$ centered at the origin, having an orientation $\boldsymbol{\alpha}$ relative to the x-axis, the distance from the origin to each of the long sides $(\boldsymbol{a})$ of the rectangle is $\boldsymbol{b/2}$, and the distance to the short sides $(\boldsymbol{b})$ are $\boldsymbol{a/2}$. When the rectangle is turned into a square, all four distances are the same, $\boldsymbol{s/2}$.
So the four peaks are then at the same distance from the $\boldsymbol{\theta}$-axis.

If the square is rotated, the distances from the origin to the four sides are not changed, but the direction of the shortest line from the origin to each side will change.
So the four peaks will slide along the θ-axis, keeping the difference at π/2.
When the orientation is along the x-axis, they will be at (-π/2 and 0), and when the orientation is perpendicular to the x-axis, they will be at (0 and π/2).

Enclosure to exam paper, INF4300, Wednesday December 13, 2011

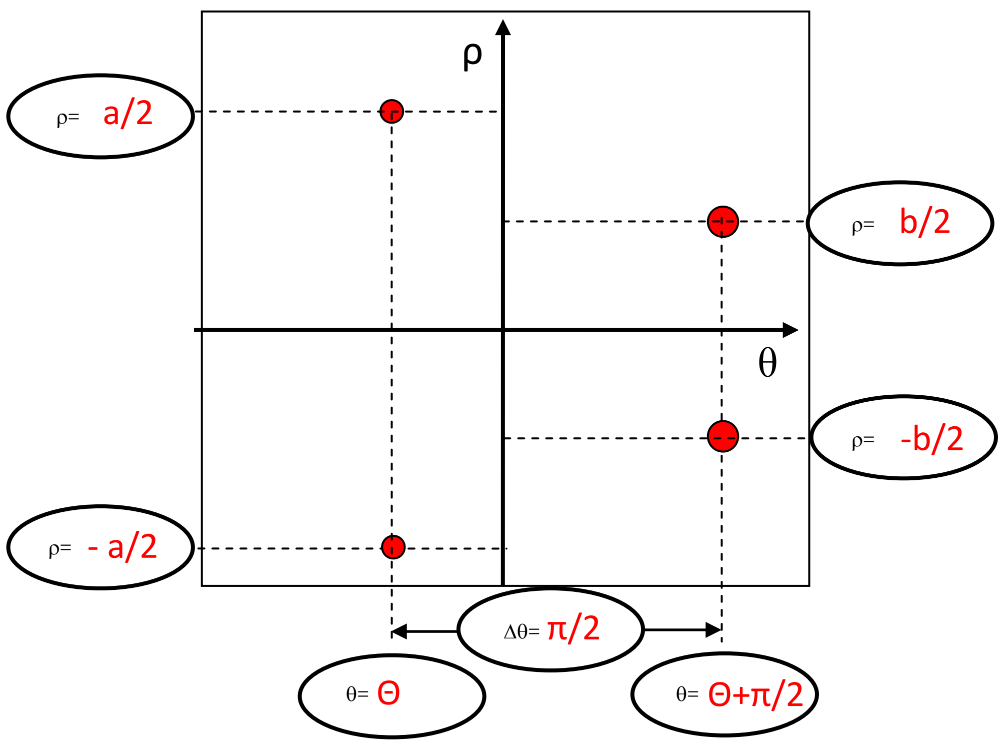
This sketch is part of Exercise 5 c.

Indicate the parameters ρ and θ of the normal representation of the line segment..



This sketch is part of Exercise 5 d.

Please fill in the proper parameters /values into the seven empty ovals



There are four line segments in the image, making a rectangle with sides a and b ($a>b$).
So there will be four peaks in Hough space, with different heights, indicated by different sizes of the red dots. The sides of length a (large dots) are at a distance $b/2$, while the sides of length b are at a distance $a/2$ from the origin.
The orientation of the rectangle is given by the normal to one of the shortest sides, so we must have θ and $\theta+\pi/2$ in the bottom ovals, and the difference between the directions of the normals to two orthogonal sides of a rectangle is always $\pi/2$