UNIVERSITY OF OSLO

Faculty of Mathematics and Natural Sciences

Exam:	**INF 4300 / INF 9305 – Digital image analysis**
Date:	**Tuesday December 13, 2011**
Exam hours:	**14.30-18.30 (4 hours)**
Number of pages:	**This is a 9 page solution**
Enclosures:	**1 sheet containing two sketches**
Allowed aid:	**Calculator**

- Read the entire exercise text before you start solving the exercises. Please check that the exam paper is complete. If you lack information in the exam text or think that some information is missing, you may make your own assumptions, as long as they are not contradictory to the "spirit" of the exercise. In such a case, you should make it clear what assumptions you have made.

- Please note that all parts of the exercises have equal weight. You should spend your time in such a manner that you get to answer all exercises shortly. If you get stuck on one question, move on to the next question.

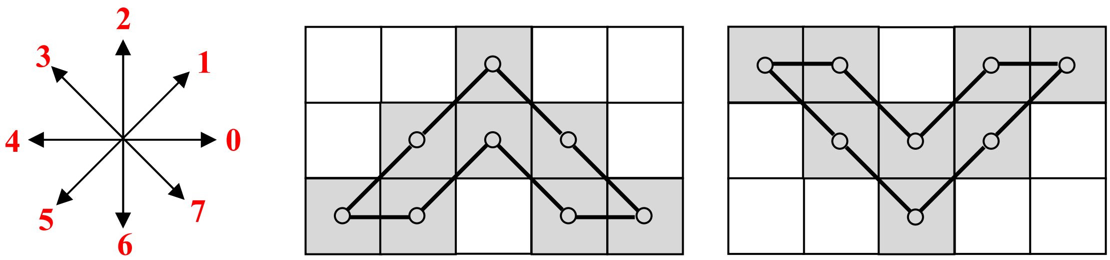
- Two of the questions are based on sketches enclosed as an extra sheet at the end of the exam text. Please give your solution on this sheet, mark it with your candidate number, and include it in your solution.

- Your answers should be **short**, typically 1-3 sentences and / or a sketch should be sufficient.

  Good luck!!

Exercise 1: 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 absolute 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 **a minimum circular shift of the clockwise absolute chain code** 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 relative 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 **a minimum circular shift of the clockwise relative chain code** 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 2: Classification

a) Given a classification problem with 2 classes.

$$\mu_1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \quad \Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mu_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \quad \Sigma_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Sketch the mean values and the covariance matrices in a plot.

b) Sketch the decision boundary for a multivariate normal distribution when the classes have equal *a priori* probabilities.

Draw a line between the mean values. The decision boundary will go through the midpoint (3/2, 1) and be normal to the line between the means.

c) What do we know about the value of the discriminant functions for the two classes at the decision boundary?

The two discriminant functions must be equal.

d) The discriminant function for a normal distribution with common covariance matrix is given by the expression

$$g_i(\mathbf{x}) = \mathbf{w}_i^t \mathbf{x} + w_{i0}$$

$$\text{where} \quad \mathbf{w}_i = \frac{1}{\sigma^2}\boldsymbol{\mu}_i \text{ and } w_{i0} = -\frac{1}{2\sigma^2}\boldsymbol{\mu}_i^t\boldsymbol{\mu}_i + \ln P(\omega_i)$$

Assume that the classes still have equal *a priori* probabilities and mean and covariance matrix as given in a). Given two features $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$, give an expression for the decision boundary as a function of $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$.

Find the boundary by solving, g1=g2, we get $x_2 = -3/4x_1+8.5/4$

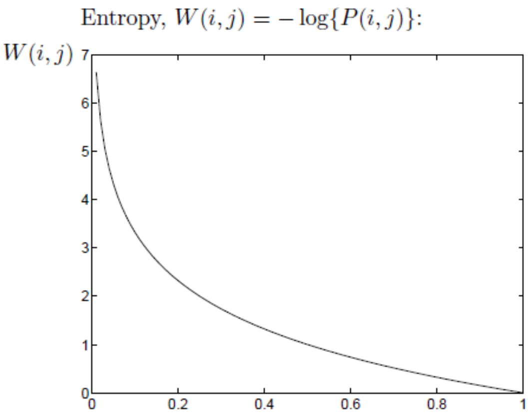
Exercise 3: GLCM

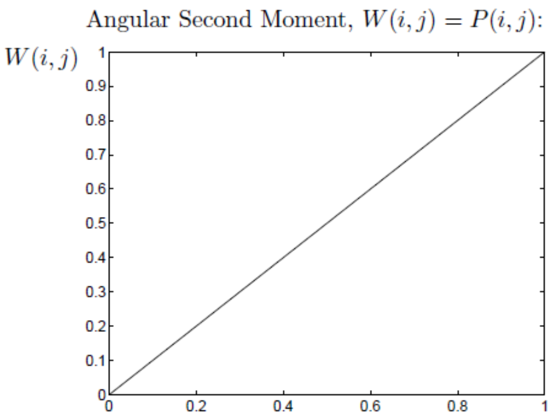
To the right you see a remote sensing image showing both steep slopes, deep valleys, flat grassland, numerous small streams, and some man-made constructions. Two subregions are incircled.

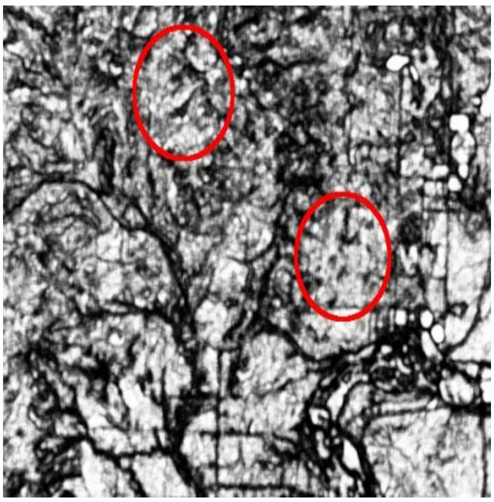
The two texture feature images at the bottom of the page were made using a moving window to compute a normalized GLCM and deriving two features that are using weight functions based on the <u>value</u> of the matrix elements, namely ENTROPY and ASM. The two equations are:

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

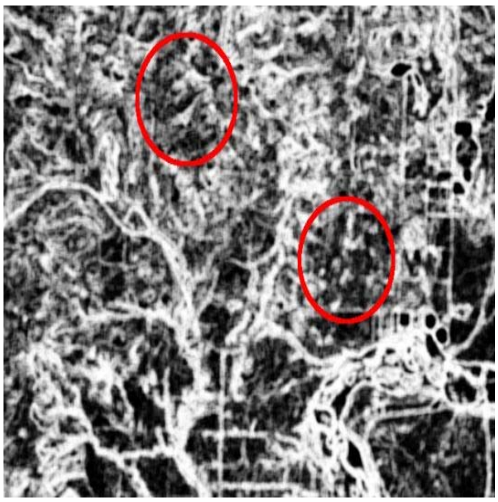
Where the weights are functions of the GLCM element probabilities, given by two curves:







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

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

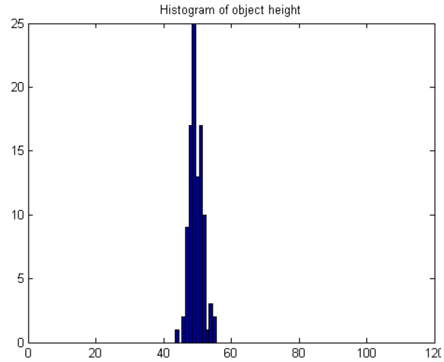
Discuss the two features in terms of local contrast and homogeneity.

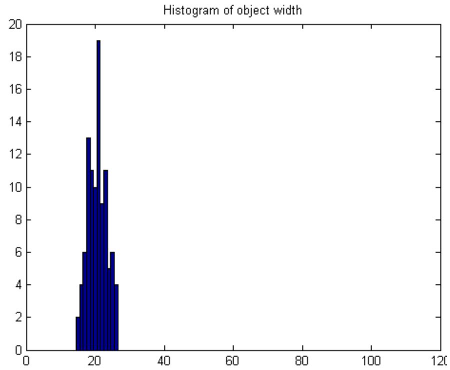
The INT weight function is zero along the diagonal (i=j), and increases rapidly away from the diagonal. Thus, it will favor GLCM elements away from the diagonal ($i \neq j$), and give higher values for images with high local contrast.

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 ($i \neq j$), and relatively higher values for homogeneous images ($i \approx j$).

Exercise 4: 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: 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 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 $-\pi/2 \le \theta \le \pi/2$.
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?

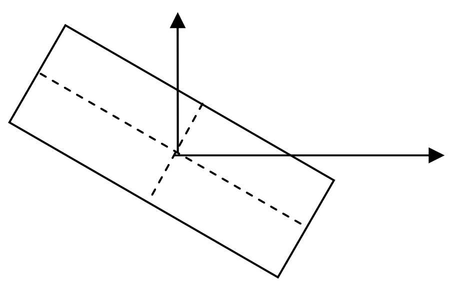
The sides of a square have the same length, so there will be the same number of pixels in the four line segments in the image. Hence, the four peaks in the parameter plane will have the same heigth.

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 θ is the angle from the x-axis to this normal.

We observe that the angle θ_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 θ_2. Since $|\theta_1| + |\theta_2| = \pi/2$, and $|\theta_2| = 2|\theta_1|$,
we must have $\theta_1 = \pi/6 = 60$ degrees.



g) Explain what happens to the positions of the peaks in the (ρ,θ)-plane
   a. If the rectangle is turned into a square of sides s.
   b. As the square is rotated through 360 degrees in the (x,y)-plane.

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

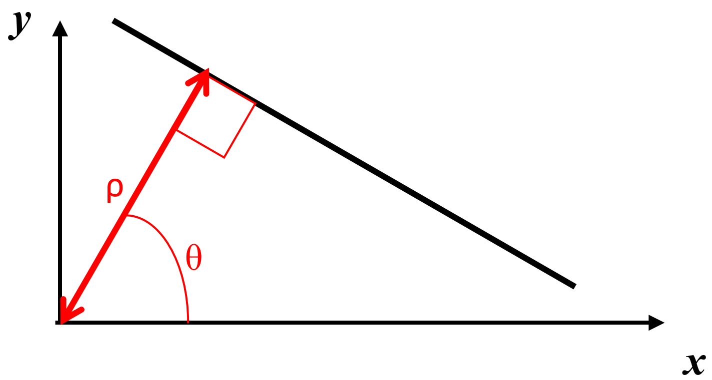
If the square is rotated, the distances from the origin to the four sides are not changed, but the direction of the shortest line from the origin to each side will change.
So the four peaks will slide along the θ-axis, keeping the difference at π/2.
When the orientation is along the x-axis, they will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

Candidate number:

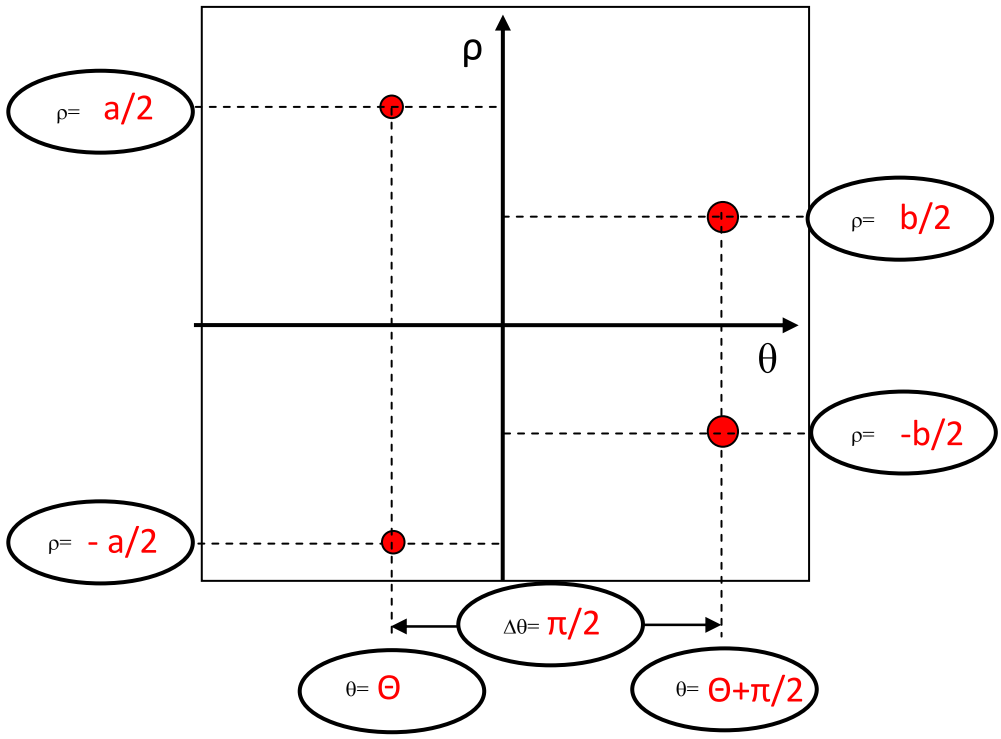
This sketch is part of Exercise 5 c.

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 ***θ+π/2*** in the bottom ovals, and the difference between the directions of the normals to two orthogonal sides of a rectangle is always ***π/2***.

Thank you for your attention !