# UNIVERSITY OF OSLO

## Faculty of Mathematics and Natural Sciences

| | |
|---|---|
| **Exam:** | **INF 4300 / INF 9305 – Digital image analysis** |
| **Date:** | **Thursday December 21, 2017** |
| **Exam hours:** | **09.00-13.00 (4 hours)** |
| **Number of pages:** | **8 pages of sketches to a solution** |
| **Enclosures:** | **1 page Appendix** |
| **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.

- 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.

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

*Good luck!!*

# Exercise 1: Representation of binary objects

Given a concave object outlined by its one pixel wide perimeter in a binary image.

a) What do we mean by *projections* of such an object?

*Answer: Most often this means the 1D horizontal and vertical projections of the pixels within the objec onto the (x,y) axes. May also mean projection onto a given axis.*

b) What do we mean by the *signature* of the object perimeter, and what is the relation between signature and projection?

*Answer: Signature is a 1D functional representation of the 2D boundary of an object. It can be represented in several ways. Simple choice: radius vs. angle, using the center of mass as origin. Radial projection in reference to centroid -> "signature".*

c) Describe and sketch the *convex hull* of such a concave object, and list some object features based on the *convex deficiency*.

*Answer: CH is the smallest convex set containing the object. Features: "CH solidity/convexity" = (object area)/(CH area), Number of and distribution of CH component areas.*

d) Describe an algorithm to obtain a simplifying polygon from the given object, so that the resulting vertexes are a subset of the original perimeter pixels.

*Answer: Draw line between initial breakpoints; boundary points farthest apart.*

1. *For each intermediate point:*
2. *If greatest distance from the line is larger than a given threshold, we have a new breakpoint.*
3. *Previous line segment replaced by two line segments, repeated 1-2 above for each of them.*

*The procedure is repeated until all contour points are within distance T from a corresponding line segment.*

a) *The resulting ordered set of breakpoints is then the set of vertices of a polygon approximating the original contour, and the set of breakpoints is a subset of the set of boundary points.*

e) Assume that you have a large set of binary images like the one produced in d), each containing only one object. The objects are all polygons, some concave, some convex. The number of polygon edges is unknown, but may vary. Explain the steps to obtain the Hough transform of each of these images, using the normal representation of a lines.

*Answer: Textbook stuff. Define a 2D accumulator matrix, determine resolution in r and θ. Find orientation for each edge pixel (using e.g. Sobel-operator). Run through all edge pixels and accumulate distance from origin (r) and angle to x-axis (θ).*

**For the PhD-students only :**

f) Based on the information given in e) above, please explain how a number of projections *less than or equal to* the number of peaks in the Hough domain can easily distinguish concave from convex objects. Why ≤ ?

*Answer: Why "less than or equal to": For convex polygons, the number of peaks in the HT is equal to the number of sides in the polygon, whereas for concave polygons, co-linear line segments may imply that the number of peaks in HT is less than the number of sides in the polygon. Besides, as you will see below, you may detect a concavity before having tested all HT peaks.*

*For each peak in HT, reconstruct the corresponding line in the image, and project the object onto an axis that is orthogonal to the line. For a convex polygon, the projection will always fall on one side of the line, whereas for a concave polygon, some projections (at least two of them) will fall on both sides of the line, and the test may be terminated before all HT peaks have been tested.*

# Exercise 2: Filtering and mathematical morphology

Assume that you have a binary image containing both large objects and small noise elements. The noise elements vary in shape, but are much smaller than the real objects. The latter have small details that are important for the subsequent image analysis.

a) What filters would you consider for removal of the noise elements, and what are the pros and cons of these filters?

*Answer: A simple lowpass filter of sufficient size would remove at least the smallest objects, but would not preserve the detailed shape of the bigger objects.*
*A median filter of sufficient size would also remove the smallest objects, but would not preserve details on the perimeter of the bigger objects.*
*Adaptive lowpass and median filters would be somewhat better.*

b) The basic morphological operators erosion and dilation, and the simplest combinations of these are often used for this task.
How would you use them, and what are the pros and cons?

*Answer: Erosion with a suitable structuring element would remove the smallest noise elements, but would also affect the size of the bigger objects, and would probably also remove the fine details on the objects. Dilation of the result of the erosion, known as* ***opening****, would regain the size of the objects, but not the details that were lost in the erosion.*

c) Please give one more advanced combination of morphological operators for this task, and discuss the pros and cons.

*Answer: Top-hat detects light objects on a dark background also called white top-hat.*
*Top-hat is given by (image minus its opening), given the structuring element for the opening.*
*Pro: Will handle all noise element shapes. Con: Needs a large structuring element.*
*The "Hit and Miss" operator could do a template matching of specific small elements, so elements of a given size, shape and orientation can be removed without affecting any other elements and objects. The con is that to remove all small elements, the process must be repeated using a list of templates.*

**For the PhD-students only :**

d) It is possible to remove the noise elements completely without altering the shape of the objects. Please explain how.

*Answer: This is covered by a group Exercise on shape representation:*

7. Compute region objects in matlab and create a region label image
to study how the regions correspond to numbers

```
b1 = bwlabel(bw);
cc=bwconncomp(b1);
ll = labelmatrix(cc);
regim = label2rgb(ll);
imshow(regim);
```

8. Compute simple region properties using regionprops in matlab.
Can you use region area to remove the frame and noisy segments?

*So by region analysis, each object and noise element is represented by a list of pixels.*
*A region property is area, and we can simply remove the shortest lists, and then regenerate the image from the object lists.*

# Exercise 3: Support vector machine classifiers

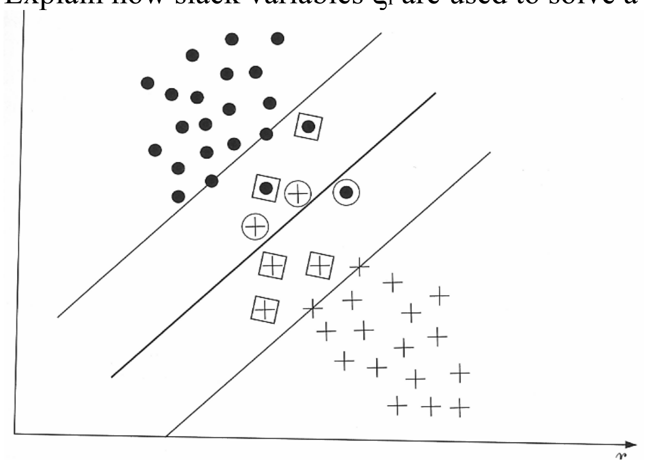
a) Consider a linear SVM with decision boundary $g(x) = w^T x + w_0$.
In SVM classification, explain why it is useful to assign class labels -1 and 1 for a binary classification problem.

*Answer: We scale g(x) such that g(x)=1 for one class and -1 for the other class, this is the value on the margin of the classifier.*

b) The basic SVM optimization problem is to minimize $J = \frac{1}{2} \|w\|^2$
What are the additional constraints for this optimization problem?
Ideally, you should answer both by math and explain what this expression means.

*Answer: With math:* $y_i(w^T x + w_0) >= 1$. *With words: all training pixels should be correctly classified.*

c) Explain how slack variables $\xi_i$ are used to solve a non-separable case like the one below:



*Answer: Points correctly classified have* $\xi_i=0$, while points inside the margin, but correctly classified, have $\xi_i < 1$. Points misclassified have $\xi_i >_i$

d) Discuss how likely a Gaussian classifier and an SVM classifier are to overfit to the training data.

*Answer: A Gaussian classifier has a restricted shape, so with complex noisy data it will not completely fit the data. An SVM without careful choice of C can easily overfit.*

e) Explain how an SVM can be used on a classification problem with M classes.

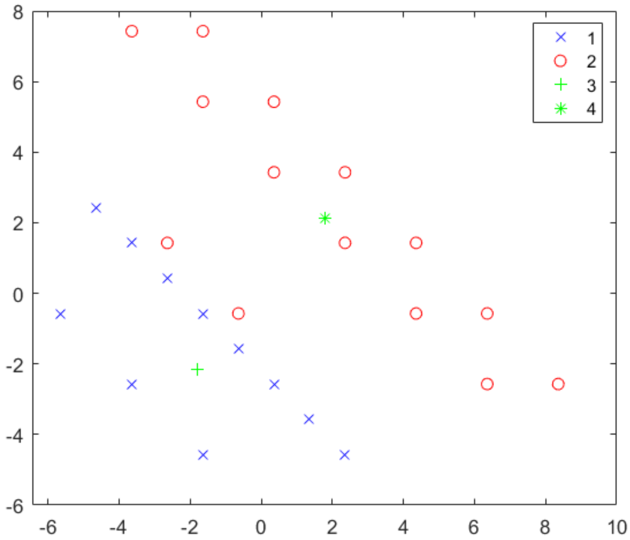
*Answer: Two approaches, all combinations of binary classifiers are normally used.*

f) Explain briefly how SVM parameters should be determined

*Answer: Briefly: grid search on valiation or cross-validation data.*

## Exercise 4: Feature selection, dimensionality reduction and classification

A two-class dataset is sketched below.




The mean values for the two classes are indicated in green.
In this exercise, the following values are given:

The mean for the entire data set is zero. $\mu 1 = \begin{bmatrix} -1.8 \\ -2.1 \end{bmatrix} \quad \mu 2 = \begin{bmatrix} 1.8 \\ 2.1 \end{bmatrix}$

Covariance $R = \begin{bmatrix} 12.4 & -2.4 \\ -2.4 & 13.1 \end{bmatrix}$

Eigenvalues of R : $\lambda_1 = 10.4 \quad \lambda_2 = 15.2$

Eigenvectors of R $v_1 : \begin{bmatrix} -0.76 \\ -0.65 \end{bmatrix} \quad v_2 : \begin{bmatrix} -0.65 \\ 0.76 \end{bmatrix}$

a) In Figure 1 the appendix, indicate the decision boundary for this problem if we use the two original features and a Gaussian classifier with equal diagonal covariance matrix.

*Answer: See appendix*

b) Based on this decision boundary, compute the confusion matrix for the training data.

*Answer: Confusion matrix:*

$$\begin{array}{cc} & \textit{True class} \\ \textit{Estimated class} & \begin{bmatrix} 11 & 0 \\ 2 & 12 \end{bmatrix} \end{array}$$

c) Principal component analysis projects the data based on a criterion function. What is this criterion function?

*Answer: Maximum variance (equal to minimum signal representation)*

d) How do we find the first principal component of the data?

*Answer: Find the eigenvector of the correlation matrix R corresponding to the largest eigenvalue*

e) Indicate in Figure 2 in the appendix the direction of the first principal component of the data.

*Answer: See appendix*

f) Consider sequential feature selection on this data set, with average Euclidean distance between the classes as the criterion function.
Set up the value of the criterion function for the two original features independently.

*Answer: J(first feature)=mu21-mu11 = 1.8-(-1.8) = 3.6*
*J(second feature) = mu22-mu12 = 2.1-(-2.1) = 4.2*

g) Select the best feature with this criterion function, and draw the decision boundary for this feature set on Figure 1.

*Answer: The criterion is maximum average class distance, so for one feature the best feature is feature 2, with a distance of 4.2. The boundary is midway between the class centers for feature 2, see plot. (boundary y=0) 8 samples are misclassified using feature 2. See fig. 1*

h) Perform a projection down onto the principal component by sketching the projection (no computation needed) on Figure 2.

*Answer: See plot*

i) Do a classification based on Euclidean distance of the first principal component and count the number of misclassified samples.

*Answer: 12 samples are misclassified based on the first principal component.*

# Exercise 5: Texture analysis

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, *d*, in a limited number of directions *θ*.

a) How do you obtain an isotropic symmetric GLCM for a given *(d,θ)*,
and how can you normalize it ?

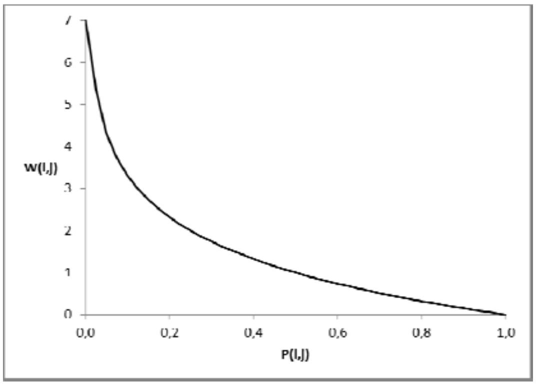
Answer: Count the number of times that pixel pairs with gray levels zi and zj occur in the image when moving a distance d in direction θ. This gives P(i,j|d,θ), which can be normalized by c(i,j|d,θ)=P(i,j|dθ)/S, where S is the sum of all elements in P(i,j|d,θ).
The normalized GLCM for the opposite direction,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

Adding the two results in a normalized symmetric matrix: $C(d,\theta) = [c(d,\theta) + c(d,\theta)^T]/2$.

b) Assume – without knowing what the image looks like – that you have obtained a normalized GLCM as described above. Then there are several possible formulae describing features that can be extracted from the GLCM. One of them is

$$E = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) \times log(P(i,j))$$

where the weight function $W(i,j) = -\log(P(i,j))$ is given in the graph below




Discuss, based on the figure above, what the output value E will be (high or low) for homogeneous and inhomogeneous images, respectively.

*Answer: Remember that since we are talking about a normalized GLCM, the P(i,j)-values sum to 1.*

- *In inhomogeneous scenes, a lot of relatively small P(i, j) values are present in the GLCM. So we are in the left part of the figure, making the Entropy value high.*
- *If all P(i, j) values are equal, Entropy will attain its maximum value of $\log_2(G)$ = bits per pixel.*
- *An extremely homogeneous image, on the other hand, will have only one gray level with P=1, so we are in the right hand part of the figure, getting an Entropy value of zero.*

**For the PhD-students only :**

c) Is the feature above invariant to a linear scaling of *the pixel values* in the image? Please explain, and also discuss another feature that behaves oppositely!

*Answer: No! We will still have a normalized histogram, but it will be distributed over a larger/smaller range of gray levels, so the entropy will be larger/smaller, since the weight function is not linear.*
*The ASM feature,*

$$ASM = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\{P(i,j)\}^2$$

*on the other hand, is invariant to linear scaling, since its weight function is linear.*

## ***Thank You for Your Attention!***

***Candidate number:***

## *Appendix, Exam INF4300/9305, Desember 21, 2017*

Figure 1

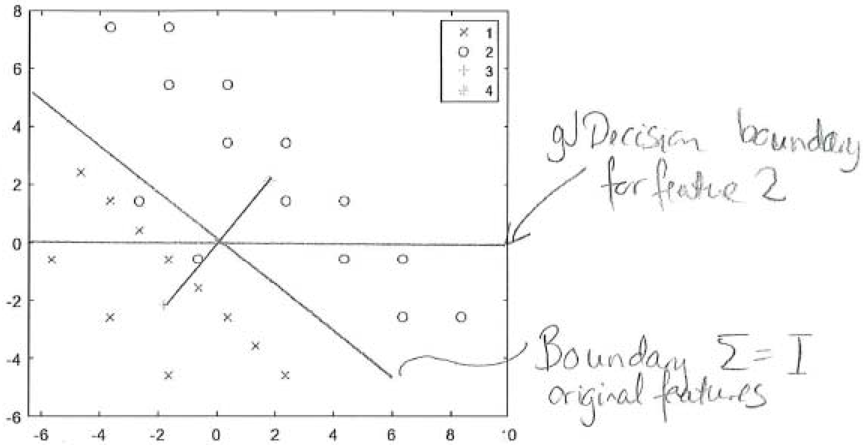



Figure 2