Exercises for
Pattern Analysis
Marco Bögel, Sebastian Käppler
Assignment 3, 05.05.2015

General Information:

Lecture (3 SWS):	Mo 08.30 – 10.00 (H16) and Tue 08.15 – 09.45 (H16)
Exercises (1 SWS):	Tue 12.15 – 13.15 (02.134-113) and Thu 8.30 – 9.30 (E1.12)
Certificate:	Oral exam at the end of the semester
Contact:	marco.boegel@fau.de sebastian.kaeppler@fau.de

Soft Clustering

Exercise 1 In this exercise, we extend hard clustering (K-means) examined in the last exercise to soft clustering (Fuzzy C-Means). Given a set of n unlabled samples $\mathcal{S} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ and K clusters, soft clustering determines the K cluster centers and estimates for each sample $\boldsymbol{x}_i$ a degree of membership $c_{ij} \in [0, 1]$ for the associated cluster j.

(a) Write down the underlying constrainted optimization problem for soft clustering.

(b) Formulate soft clustering as unconstrainted optimization problem using Lagrange multiplier.

(c) Derive a closed-form solution of the update formulas for the cluster centers $\boldsymbol{\mu}_j$ and the clustering matrix c_{ij} used in the soft clustering algorithm.

(d) **Matlab exercise:** Implement the soft clustering algorithm in Matlab to deal with multi-dimensional data and an arbitrary number of clusters.

(e) Now, we extend soft clustering such that the mutual distance between the cluster centers is maximized. Write down the objective function for soft clustering and derive the modified update formulas. Explain the benefits of this approach compared to conventional soft clustering for a simple example.

Hint: What happens for conventional soft clustering and the extended approach if the number of clusters was set to high?

Exercise 2 **Matlab exercise**

The problem of image segmentation can be formulated as clustering problem. This exercise examines several approaches for the special case of *image binarization* to segment color images into two segments (foreground and background). We apply our segmentation technique to microscopy images to separate cells from the background, see Fig. 1.

For interested students: Details on cell segmentation and more advanced algorithms are described in
F. Mualla et al., *Automatic Cell Detection in Bright-Field Microscope Images Using SIFT, Random Forests, and Hierarchical Clustering*. IEEE Transactions on Medical Imaging, vol. 32, no. 12, pp. 2274-2286, 2013

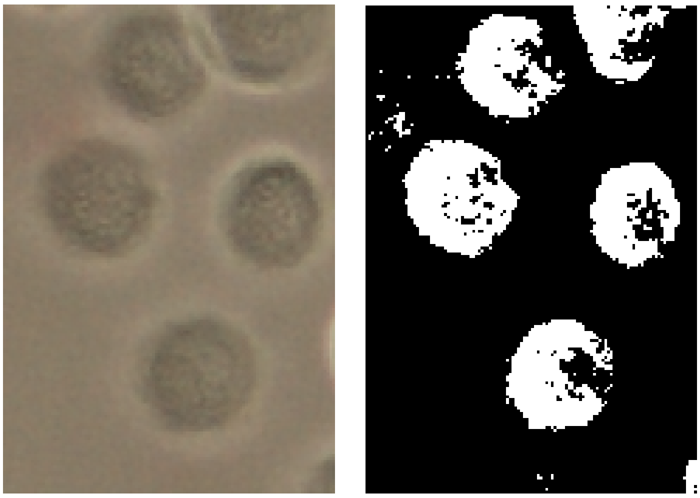

Figure 1: Insect cells (left) and example segmentation (right).

(a) Load the image `cell.png` for an evaluation of the different image segmentation algorithms compared in this exercise.

(b) Enhance the contrast of the image to discriminate cells from the background (Matlab: `imadjust`).

(c) We have to extract features from the given cell image that are utilized in the clustering algorithm. For this purpose, discriminative features can be obtained in the HSV (hue/saturation/value) color space of the image. Transform the given image from the RGB to HSV color space (Matlab: `rgb2hsv`) and use the H-, S-, and V-component for each pixel as feature for clustering. Test different combinations of features in your evaluation and find out a discriminative set of features for image binarization.

(d) Binarize the given cell image using different methods:

- Hard clustering using the K-Means algorithm based on the squared euclidean distance (re-use your code from the last exercise).
- Soft clustering using the Fuzzy C-Means approach. To obtain a segmentation of the cell image, the degree of cluster membership c_{ij} determined by the clustering algorithm has to be thresholded.

In each case, we use $K = 2$ clusters (background and cell). Visualize and compare the segmentation results obtained with the different algorithms.