Technische Universität Berlin
Faculty IV
Chair of Computer Vision and Remote Sensing
Automatic Image Analysis

Summerterm 2019

4th Homework: Expectation Maximization

July 4, 2019

Raphael Schymczyk
349690

Karel Smejkal
388638

David Luibrand
408183

Thomas Hartenstein
371926

1. Theory

Question	Answer			
	i)	ii)	iii)	iv)
1.			X	
2.			X	X
3.	X			

2. Implementation

We investigated the performance of the algorithm for different numbers of Principal Components and different numbers of GMM-Components. The observed accuracy is shown in the figure below.

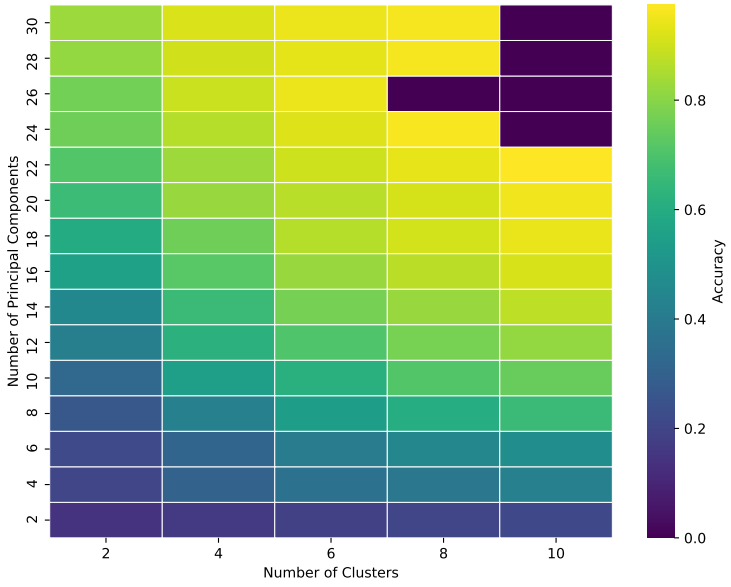

Figure 1: Accuracy of the EM-Algorithm measured as fraction of correctly classified samples.

The accuracy of the algorithm increases both with the number of principal components and the number of used GMM-Components. The highest accuracy (97,4 %) was observed for 22 principal components and 10 GMM-components. For very high number of principal components and GMM components the algorithm crashes (denoted in figure 1 by an accuracy of zero). For more than 12 GMM-Components, the algorithm crashed nearly every time.

In theory, with a higher number of principal components, the projection of the features keeps a higher amount of variance (and therefore information) of the original data. Since

the algorithm should profit from this additional information we would expect it to be able to generate predictions with a higher accuracy. The observed values of the accuracy agree with these theoretical consideration.

An explanation for increasing accuracy with a higher number of used GMM-Components is, that this allows the algorithm to estimate the probability distribution in a more flexible way. This leads to a more precise estimation of the posterior probabilities especially for 'unusual' observations.