Exercise 06 for MA-INF 2201 Computer Vision WS18/19
18.11.2018
Submission on 24.11.2018

1. **Investigating the EM Algorithm for Gaussian Mixture Models.**
   Assume a Gaussian Mixture Model where the λ_k follow a uniform distribution and all mixtures have unit variance, i.e. $\Sigma_k = \mathbf{I}$ (all covariance matrices are the unix matrix).

   (a) Give the explicit formula for r_{ik} in the E-Step.
   (2 points)

   (b) Simplify the above representation for r_{ik} by summarizing all terms independent of k into a constant C_i. If you consider the logarithm of the result, what do you observe?
   (2 points)

   (c) Sometimes, maximum approximation is used in the EM algorithm. That means, the mixture k with the highest probability for x_i is assigned all probability mass: $r_{ik} = \begin{cases} 1, \text{if } k = \arg\max_{k'}\{r_{ik'}\}, \\ 0, \text{otherwise.} \end{cases}$
   How does this affect the update of the mean μ_k in the M-Step? Taking your result from (b) into account, what does the EM algorithm for Gaussian Mixture Models break down to under the given assumptions (maximum approximation, uniform λ_k, $\Sigma_k = \mathbf{I}$)?
   (2 points)

2. **Implementation of the EM algorithm for Gaussian Mixture Models.**
   In this exercise, assume that all covariance matrices are diagonal.

   (a) In the template, implement the function `fit_single_gaussian` that fits a single Gaussian to the provided data.
   (1 point)

   (b) GMMs rely on a good initialization. One strategy is to start with a single Gaussian model, split it into two distributions (GMM with two mixtures) and train it using the EM algorithm. For training a GMM with four mixtures, both of the previous mixtures can be splitted again, and so on. Implement the `split` function, that doubles the number of components in the current Gaussian mixture model. Therefore, generate $2K$ components out of K components as follows:

   - Duplicate the K weights λ_k so you have $2K$ weights. Divide by two so that they sum up to one again.
   - For each mean μ_k, generate two new means: $\mu_{k_1} = \mu_k + \varepsilon \cdot \sigma_k$ and $\mu_{k_2} = \mu_k - \varepsilon \cdot \sigma_k$. Use $\varepsilon = 0.1$.
   - Duplicate the K diagonal covariance matrices so you have $2K$ diagonal covariance matrices.

   (2 points)

   (c) Implement the EM algorithm to train the GMM.
   (6 points)

(d) **Generating synthetic handwritten digits.** Load the USPS digit dataset and train a GMM for each of the ten digits. Each GMM should have 8 mixtures. Your training should start with a single Gaussian and build up the final 8-mixture GMM by splitting the smaller models as described above. Run the EM algorithm for 100 iterations each time. Implement the `sample` function that returns a random sample from the learned mixture distribution. Run your code to generate synthetic digits by drawing ten samples of each digit and save them in one big image (10×10 images, each 16×16 pixels, thus a 160×160 image). Please also upload this image to your solution.
(2 points)

(e) **Skin color segmentation.** The provided face image comes with a rectangular bounding box that contains some skin color pixels (foreground). All other pixels are background. Use your implementation to train a GMM with 8 components (start with single Gaussians and do 3 component splits) for both foreground and background pixels. Using the thresholding approach from the lecture, set every pixel in the image to zero that is below the threshold 0.1. Display the resulting image.
(3 points)