Computer Vision Final Term project

In this project, you need to implement PCA algorithm with numpy module only. Skeleton code written in Python and dataset are given. Please check out 'pca.py' attached. For the dataset, a part of MNIST image dataset (60k training images) is provided.

What you need to do:

1. Eigenface

Extract top 40 eigenfaces from the given 2-dim MNIST image dataset, and visualize them.

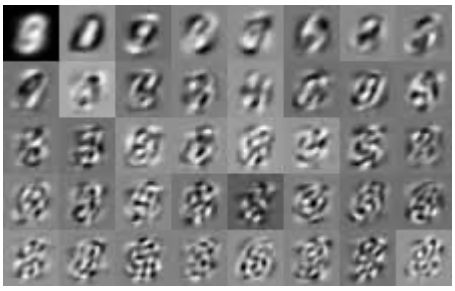

Fig 1. MNIST eigenface examples

Eigenfaces are 2-dimentaional version of eigenvectors, and utilize show_np_arr(...) function in the skeleton code for visualization. Note that, you need to implement PCA algorithm to obtain eigenfaces. Please refer to our lecture material for details.

2. Image approximation

Choose the top 'K' eigenfaces which can preserve 85% of information in the MNIST dataset (choose the smallest K. how?). Then, randomly select an image x from the MNIST dataset and approximate the image as $\hat{x}$ using the formula as follows:

$$\hat{x} = \mu + w_1u_1 + w_2u_2 + w_3u_3 + w_4u_4 + \dots$$

Where μ denotes the mean of the dataset, and u_i and $w_i$are eigenface and weight, respectively. Show the reconstructed image $\hat{x}$.

Finally, compare the mean squared errors (MSEs) between x and $\hat{x}$ by changing the number of eigenfaces used for approximation from 1 to K.

3. Fast approach

In general, it takes much time in computing eigenvectors when the image resolution is high. Implement a fast approach by changing the covariance matrix (see our lecture material), and show the achieved 'K' eigenfaces.

4. Submission

Submit the completed 'pca.py' file and a short report through our course website

(I have created a for the for the submission named with "Term Project").

Your report should include short description of your code, and figures and analysis about the results. Page limit is 2 pages. It should be written in English, and is a single pdf file (yourID.pdf). The deadline is 12/16 23:59, and cannot be extended (late submission is not allowed, please double check whether your results are correctly uploaded).