

KECE470: Pattern Recognition
School of Electrical Engineering, KOREA UNIVERSITY
(Project #1) (due via Blackboard, 5 pm, Mon 12 Oct 2020)
5-page report containing the code, results, discussions

1. Your task is to develop a code for PCA (hence construct eigen-digits) to project input digit images into eigen-digit space and find the matching digit index (0 to 9) given a test input digit image.

In this project you will develop a handwritten digit (from MNIST dataset) recognition code. Your system (hence a program code) reduces each digit image to a vector, then uses Principal Component Analysis (PCA) to find a linear subspace for 10 digits.

Upload, and show **any** sample images in Matlab or Python (MNIST dataset). The input data you will be using is divided into several sets. They can be downloaded from this link.

<http://yann.lecun.com/exdb/mnist/>

2. PCA is a technique by which we reduce the dimensionality of data points. Describe the steps in the PCA for **face recognition** (limit to half page length). Please see the Lecture notes for details.
3. Write a pseudo-code for 2 (limit to half page length)
4. Using the results from problem 1 to 3, construct a code that distinguishes 10 classes from 0 to 9 from a given dataset (MNIST) using PCA, so-called hand-writing recognition.
In your code, show how accurate the test image in terms of classification performance is. Also show how a set of k coefficients is obtained by a code implementation.

Helpful links:

<https://courses.cs.washington.edu/courses/cse455/09wi/projects/project4/web/project4.html>
<https://static1.squarespace.com/static/55133727e4b05ad39f9b9749/t/56ba9b20a3360ce1d09cb271/1455069985248/principal-component-analysis.pdf>
<https://medium.com/analytics-vidhya/principal-component-analysis-pca-with-code-on-mnist-dataset-da7de0d07c22>

T.A email: dhkim2@ispl.korea.ac.kr