Submission

Please submit answers to the questions below and any relevant code to ber@robots.ox.ac.uk by next Friday, May 19th before the end of the day.

Question 1

Read the following resource on *eigenfaces* here. This is an application of dimensionality reduction methods to face recognition.

Follow the given code and make your own implementation of *eigenfaces*, using the same given dataset. Then, answer the following questions:

1. Use a picture that is not in the dataset (e.g. a picture of yourself, a friend or a celebrity) and obtain its vector of coordinates on the *eigenface* basis. Then use this vector to reconstruct the original image. Is the reconstruction good? Discuss.

2. What are the advantages and disadvantages of this method? Mention as many as you can.

Question 2

In this question, you will visualise word representations using dimensionality reduction techniques. Use Word2vec (previously trained on a large vocabulary) to generate feature vectors for 10-word samples of three or four different domains, e.g. food, languages and hand tools.

Map the resulting high dimensional vectors onto a 2-dimensional manifold using PCA. Plot the output vectors on the same canvas. Additionally, compute the distance matrix of the Word2vec feature vectors using the Euclidean norm. Use multidimensional scaling to generate a 2D representation of the data.

1. Discuss any observations on the results, e.g. are the words clustered by semantic domain? Explain if the resulting plots met your prior expectations.

2. What similarities and differences did you find between both plots?