**HW1: Person Identification from Face Images**

**CSC 481**

**Total Point: 20**

**Due: Feb 25, 220**

------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

# Problem Description

In this assignment, you are going to identify persons from face images using K-Nearest Neighbors and Decision Tree classifier. You will be given a face dataset. You will have to extract features from the data, deploy the classifiers (using Matlab/Scikit-learn/your preferred programming language), and finally measure the performances (in terms of identification accuracy, precision, and recall rate) of the classifier.

# Definition: Identification

In face identification you have to identify a person from n number of persons. Initially template of n persons are created and stored. When a claim comes from a person p, the test pattern is matched with all stored templates. If the test pattern is found similar to one of n templates then, the test pattern is identified as that person. Otherwise, the system will say that the person does not exist in the system.

# Data Sets and Experiment Design

**Face Database**

We are using AR face database which is public and access is free. It contains face images of 130 persons. However, for our experiments I have selected images of 10 persons. Five of them are males and the other five are females. For each person we will be using four face images. We will use three images for training and the one image for testing.

**Data Format**

Each face image contains 22 markup points as shown in Figure 1. For each point, we have the (X, Y) coordinates (*i.e.,* abscissa and mantissa) which shows the spatial location of the markup point in the image. Therefore, each data file contains 22 coordinates corresponding 22 points in Figure 1.

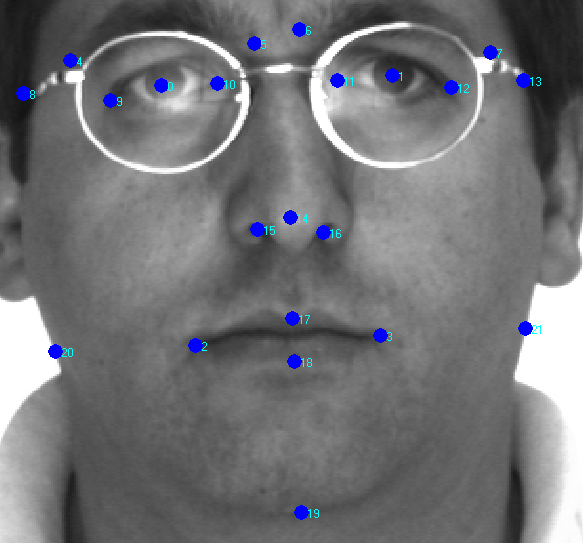


Figure 1: A face sample with 22 markup points

# Feature Definition and Extraction

**Defining Features**

I have identified seven features-

1. Eye length ratio: length of eye (maximum of two) over distance between points 8 and 13.
2. Eye distance ratio: distance between center of two eyes over distance between points 8 and 13.
3. Nose ratio: Distance between points 15 and 16 over distance between 20 and 21.
4. Lip size ratio: Distance between points 2 and 3 over distance between 17 and 18.
5. Lip length ratio: Distance between points 2 and 3 over distance between 20 and 21.
6. Eye-brow length ratio: Distance between points 4 and 5 (or distance between points 6 and 7 whichever is larger) over distance between 8 and 13.
7. Aggressive ratio: Distance between points 10 and 19 over distance between 20 and 21.

**Feature Extraction**

You will have to extract the above mention features. Compute the feature values as per the definition of the features given above. Use the Euclidian distance when you compute the distance between two points.

# Experimental Results and Analysis

1. You will have to prepare a confusion matrix and calculate precision, recall rate, and accuracy for each of the classifiers.
2. You need to analyze (discuss) the results.

# Appendices

1. Put the values of seven features.
2. Put the source code.

# Deliverables

You should upload the following items:

1. Report having: introduction, problem description, data sets and experiment design, defining features and feature extraction, result and analysis, conclusion, references, and appendices.
2. Source code.