Question 5: Assignment 4: CS 663, Fall 2024

Amitesh Shekhar
IIT Bombay
22b0014@iitb.ac.in

Anupam Rawat
IIT Bombay
22b3982@iitb.ac.in

Toshan Achintya Golla
IIT Bombay
22b2234@iitb.ac.in

October 23, 2024

1. What will happen if you test your system on images of people which were not part of the training set? (i.e. the last 8 people from the ORL database). What mechanism will you use to report the fact that there is no matching identity? Work this out carefully and explain briefly in your report. Write code to test whatever you propose on all the 32 remaining images (i.e. 8 people times 4 images per person), as also the entire test set containing 6 images each of the first 32 people. How many false positives/negatives did you get? [10 points]

Soln:

Since the system has no prior knowledge of these new 8 individuals, we propose a **threshold-based verification mechanism**. It involves setting a threshold value τ for the distance between the test image's eigencoefficients and those of the training images. If the minimum distance is above the threshold τ, we classify the test image as belonging to an unknown individual. This way, we can detect when the system is presented with an image of a person it has not seen before.

The threshold τ is computed by examining the **intra-class distance** for each individual in the training set. The final threshold is set as half the median of the maximum intra-class distances, making it robust against variations in the appearance of each person.

The steps for testing are as follows:

- Test on the last individuals, to decide if the person is known or unknown, using thresholding.
- Test on all 192 images from the first 32 individuals (6 images per person).

The following two metrices are used:

- **False Positives**: Cases where the system incorrectly accepts an unknown individual.
- **False Negatives**: Cases where the system incorrectly rejects a known individual.

The results for different values of k (the number of eigenvalues used for classification) are shown below:

k	Threshold τ	False Negative Rate (%)	False Positive Rate (%)
1	517836.51	0.00	100.00
2	724546.37	0.00	100.00
5	2817777.77	78.13	96.25
20	6985547.53	78.13	96.25
50	10374733.60	78.13	96.25
100	13887953.83	0.00	96.25

Table 1: False Positive and False Negative Rates for Different k Values

Observations:

- For smaller k values, the system fails to reject unknown individuals, leading to **100% false positive rates**.
- For larger k values ($k = 100$), the system performs better in rejecting unknown individuals, but still struggles with high false positive rates, especially with unknown people whose images are similar to those of known individuals.
- The best result is achieved for $k = 100$, where the false negative rate is minimized to 0%, though further tuning of the threshold is required to improve the false positive rate.

The corresponding code, is included in the code directory.