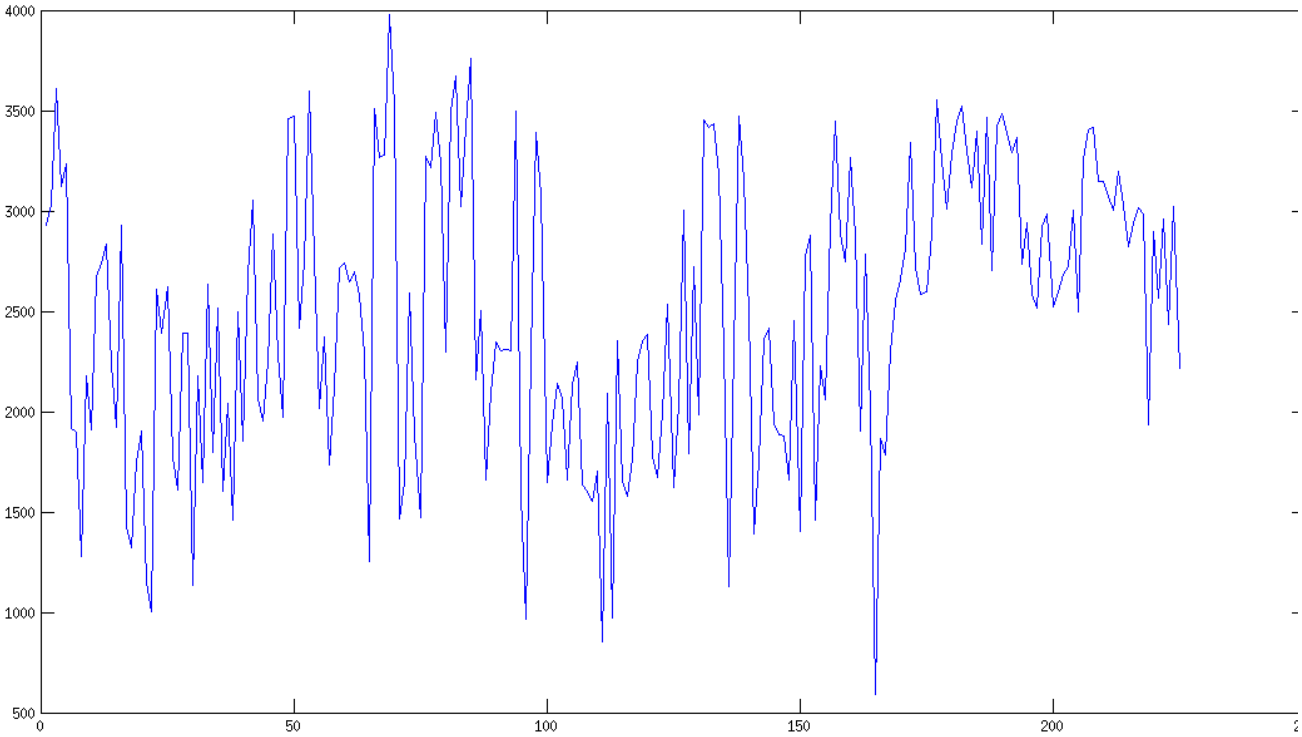
CS663 Assignment3 Question4 Report

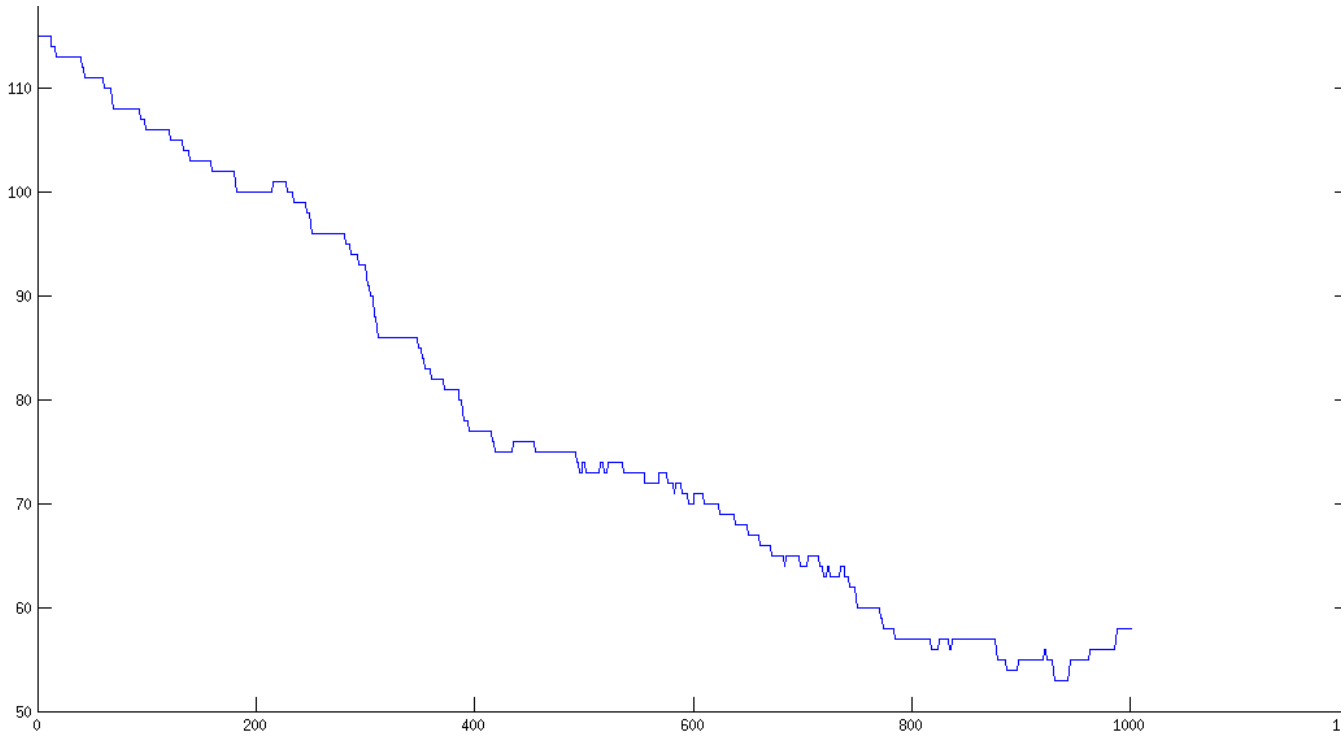
In case the designed face recognition system is tested using people not part of the testing set, the system will still report the 'closest' image as the identity of the person, which is wrong. This is rectified as follows:

A plot of the Euclidean 'distances' to the nearest component vector (the predicted identity of the person) is shown below. Persons 1-175 have images in the training set and 176-225 are new people.



The average distance for the new people is greater than the error for the people already present. So we set a threshold, and if the distance is more than this threshold, the person is reported as unrecognised.

This threshold needs to be set to minimize the number of wrong answers. So we seek to minimize (number of false positives + number of false negatives). We plot this number for threshold = average distance for recogized people to threshold=average distance for new people.



We notice that the minimum occurs at ~ 2938 (the plot is for values of threshold from 2000 to 3000). For this, we have 35 false negatives, 18 false positives, and a total of 53 wrong answers.