INTRODUCTION

Face Recognition is the process of identifying a face as a known or unknown individual. The framework uses a set of images (called the training or gallery set) to train a classifier and then match a set of new images (called the testing or validation set). The first step in face recognition is to define an appropriate representation of the data.

"Eigenfaces" [1] is the name given to the set of eigenvectors to represent a face in facial recognition. Principle Component Analysis (PCA) is used to reduce dimensionality of a vectorized image. This projection algorithm maintains variance while discarding unnecessary correlations among the original features. Since PCA is unsupervised, this algorithm also has the ability to learn to recognize new faces introduced in the validation set.

"Fisherfaces" [2] is a similar approach that uses Linear Discriminant Analysis (LDA) on a set of images. This method is usually regarded as ideal for face recognition because it seeks to maximize distance between classes. However, LDA is a supervised technique so it is difficult to learn new faces during the testing step.

OBJECTIVES

The main objective of this research is to implement a system to improve and optimize the process of face recognition on the MORPH-II database. Previous studies on face recognition using MORPH-II such as [3] show accuracy rates with room for improvement.

Much research has been done on the effect of age on the face recognition problem with MORPH-II but there is still a need to analyze the effects of gender and size of the database. Since research with face recognition based on race and gender on other datasets yield promising results [4], my goal was to implement some of these practices with the much larger and more challenging MORPH-II database. A simplified version of my proposed framework is outlined below.

EXPERIMENTAL SETUP

MORPH-II [5] is a large-scale, longitudinal database composed of 55,134 images of 13,617 distinct individuals. These images were collected over a period of 5 years and individuals’ ages range from 16 to 77 years. Number of photos for each individual varies from 1 to 53, with an average of around 4. Most face recognition research focuses on other smaller, controlled databases. The MORPH-II dataset provides a unique challenge for the problem of face recognition because of variations in age, expression, illumination, and unequal proportions for race and gender.

As a first step, all photos in the database were preprocessed. Each face image was detected and aligned with eye centers, and then cropped, resized to 70 X 60 pixels, and histogram equalized. The preprocessed, gray level images were used for face recognition.

Dimension Reduction using PCA (Eigenfaces) or LDA (Fisherfaces) is then performed on the set of images. After the dimension reduction technique, a classifier such as Support Vector Machine (SVM) or a Nearest Neighbor approach is trained on the gallery set. For the nearest neighbor approach, several distance measures are used such as Euclidean, CityBlock, Cosine, Bray-Curtis, Canberra, Correlation, and Mahalanobis.

As for subset schemes on the database, only subjects with at least 10 images or more were considered. The race and gender breakdown of this selection is shown below. Experiment 1 includes all 83 females and a random selection of 83 out of the 461 males. Experiment 2 includes all 544 subjects in this subset. For each subset scheme, 5 images are chosen at random for training the algorithm and 5 images (excluding the 5 training images) are chosen at random for the testing set for each distinct subject, resulting in 10 images total per person.

RESULTS

The next step is to try these algorithms on a larger, more difficult subset. The results of this subset are shown in the table below.

Accuracy rates dropped significantly for this second experiment (17.6%). Since this subset includes changes in both size and gender ratio, the next step is to analyze the effect of gender ratio and subset size on face recognition with the MORPH-II database.

To analyze the effect of gender and size on face recognition accuracy, two experiments were carried out to control for changes in both.

ACKNOWLEDGEMENTS

Special thanks to the National Science Foundation for this opportunity. Without the hard work of Troy Kling, Dr. Tracy Chen, and Dr. Yishi Wang, this research would not have been possible.