Subset Size and Gender Analysis on Face Recognition

Morgan Ferguson

Department of Mathematics and Statistics NSF REU in Statistical Data Mining

REU Site, DMS - 1659288

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 some 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 data 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.

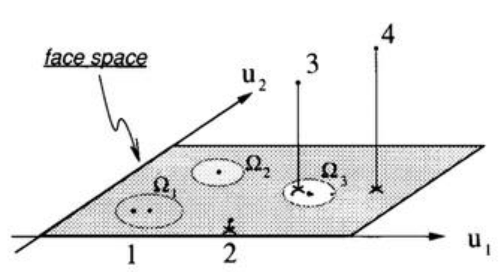


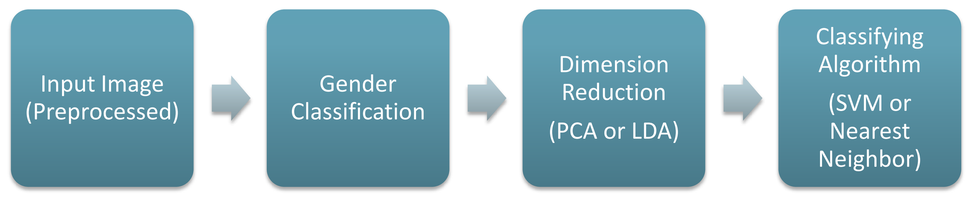
Figure 1: A simplified version of eigenvectors for a set of images, known as the face space [1]

"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

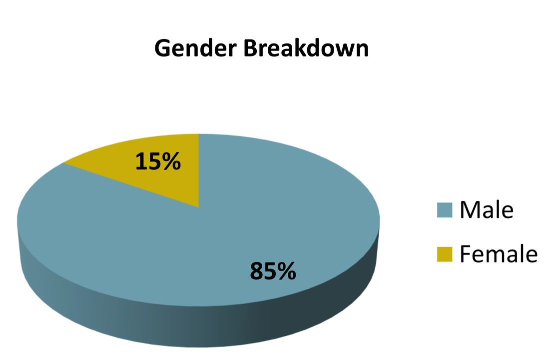
MORPH-II [5] is a large-scale, longitudinal database composed of 55,134 images of 13,617 distinct individuals. These images were collected over a period of 5 years and individuals' ages range from 16 to 77 years. Number of photos for each individual varies from 1 to 53, with an average of around 4. Most 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.

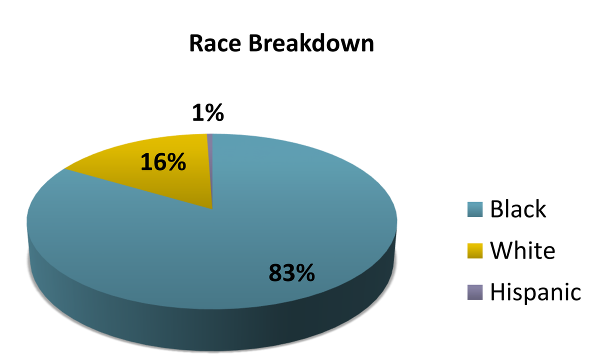
As a first step, all photos in the database were preprocessed. Each face image was detected and aligned with eye centers, and then cropped, resized to 70 X 60 pixels, and histogram equalized. The preprocessed, gray level images were used for face recognition.

Figure 2: Examples of images in the subset scheme before preprocessing (top row) and after preprocessing (bottom row)

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

Results on experiment 1 are shown in the table below.

		SVM-R	Euclidean	CityBlock	Cosine	BrayCurtis	Canberra
Eigenfaces (PCA)	Accuracy	**88.6%**	69.9%	78.9%	71.1%	79.5%	79.5%
	Run Time (sec)	211.3	7.4	3.9	10.0	6.8	9.6
Fisherfaces (LDA)	Accuracy	89.2%	92.8%	89.2%	**95.8%**	94.8%	83.1%
	Run Time (sec)	179.4	10.8	7.3	12.8	8.5	13.3

Table 3: Face Recognition on subset with 166 subjects with 10 images each and equal gender distribution (1 M : 1 F)

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

		Euclidean	CityBlock	Cosine	BrayCurtis	Canberra
Eigenfaces (PCA)	Accuracy (%)	54.0	63.1	56.0	**69.7**	66.0
	Run Time (sec)	139.2	78.8	268.5	121.6	223.9
Fisherfaces (LDA)	Accuracy (%)	62.9	55.9	**78.2**	71.7	51.9
	Run Time (sec)	163.7	115.3	281.0	146.1	255.8

Table 4: Face Recognition on subset with 544 subjects with 10 images each and unequal gender distribution (461 males, 83 females)

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.

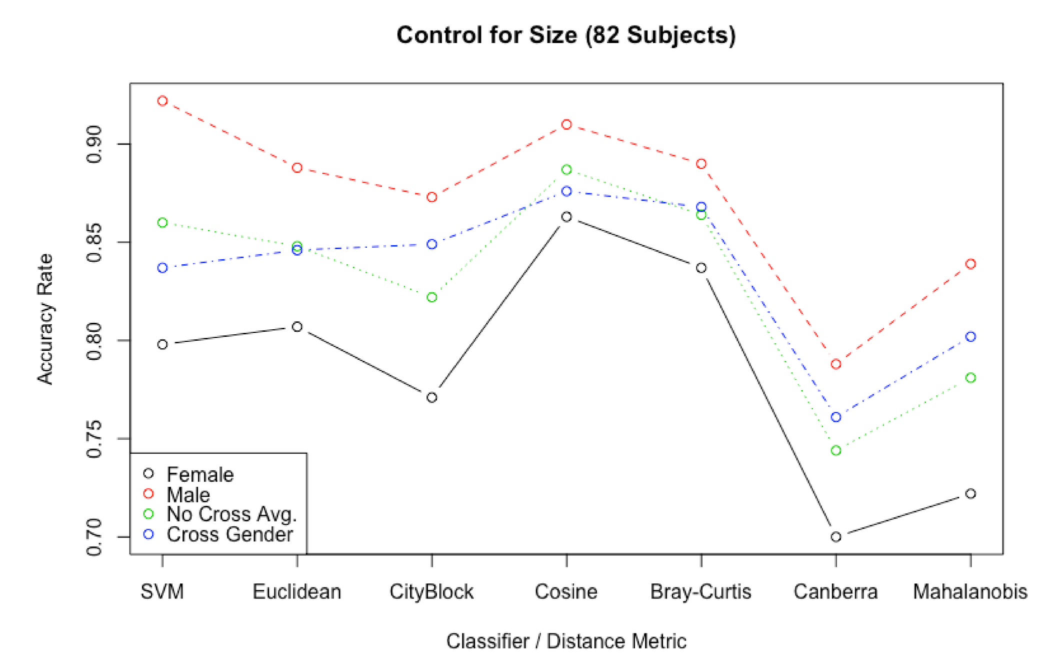


Figure 5: Face Recognition on subsets with 82 images each and varied gender ratios

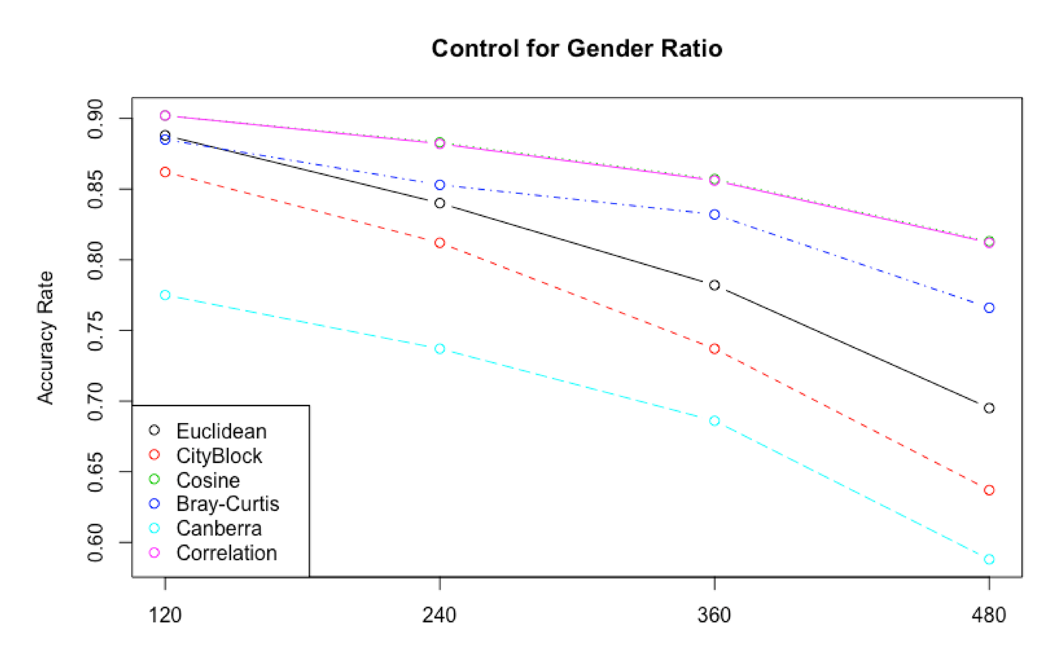


Figure 6: Face Recognition on subset with a gender distribution of 5 M: 1 F and varied subset sizes

CONCLUSION

The results on gender variations show that females perform much worse than males and even cross-gender subsets in face recognition. These findings were unexpected but could potentially be due to the variance across female images in the database with regards to makeup, pose, and jewelry changes that are not present or as prominent across the males in the database. However, the results on size variation showed a steady decrease in algorithm performance as subset size increases. It should also be noted that for increasing variations in subset size, run time for nearest neighbor increases in a quadratic fashion.

In conclusion, size has a major impact on both accuracy rates and running times for face recognition using the MORPH-II database. While the separation of gender alone might not make much difference on the performance, gender classification as a first step serves an important role in reducing the size of the database used for face recognition.

However, there is still much research to be done on this topic. Race impacts on face recognition within the MORPH-II database have not been analyzed. Since size has such a major impact on accuracy and computation time, race classification should be considered as an additional means to reducing subset size. With these specified classes, accuracy rates should be expected to further increase and running times drop. An updated flowchart of face recognition using the additional step of race classification is included below.

REFERENCES

[1] M. Turk, A. Pentland, "Eigenfaces for Recognition", J. Cognitive Neuroscience, vol. 3, no. 1, 1991.

[2] P. Belhumeur, J. Hespanha, and D. Kriegman, Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):711--720, 1997.

[3] G. Guo, G. Mu, K. Ricanek, "Cross-age face recognition on a very large database: The performance versus age intervals and improvement using soft biometric traits", Proc. ICPR, pp. 3392-3395, 2010.

[4] K. Veropoulos, G. Bebis, and M. Webster. Investigating the impact of face categorization on recognition performance. In International Symposium on Visual Computing (LNCS, vol. 3804), December, 2005.

[5] K. Ricanek, T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression", IEEE Conf. on AFGR, pp. 341-345, 2006.

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.