

CISC-820 Project 4: Dimensionality Reduction & Classification

Ayush Kumar Shah and Kai North

1 Objective

This report describes the use of Principle Component Analysis (PCA) for reducing the dimensions of face images by the extraction of eigenfaces as well as the reconstruction of reduced representation back to original dimensions. PCA works by finding better set of bases called as eigenfaces or principle components for representing the images and keeps only the important features (features with higher percentage of total variance in the images) to reduce the dimensions.

This report outlines the application of PCA in face images of 40 different subjects with 10 images each. It also analyzes what the eigenfaces represents visually and how the importance of various eigenfaces changes (See Section 2). Likewise, it shows how well a subset of images can be reconstructed as the number of eigenfaces used to represent the reduced image increases (See Section 3).

2 Using PCA to Extract Eigenfaces

2.1 How do the leading eigenfaces look like as an image?

The first 5 eigenfaces appear to resemble the outline of a human-face. In the first eigenface, you can see a circular object. Varying contrasts highlight facial features like hair, eyes, mouth and ears on this object. However, the image appears mostly fuzzy with defining facial characteristics being unrecognizable. In contrast, the fifth eigenface displays more facial structure. A chin, cheekbones, eyebrows and a hair-parting can be made out. This would suggest that the more eigenfaces used then the more detailed the image becomes as all the images are basically the linear combination of these eigenfaces.



Figure 1: First 5 Eigenfaces

2.2 How does the importance of eigenfaces decrease?

The first 50 eigenfaces out of 10,304 were found to be of very high importance since it represented 81.61% of the total variance in the 400 input images as shown in Figure 2. Adding more eigenfaces increase the overall importance (variance) but at a slow rate compared to the first 50 i.e. the importance of each eigenface starts decreasing. For instance, the first 100 eigenfaces represent 89.06% of the total variance, i.e. the second set of 50 eigenfaces account for just around 7.4% importance compared to 81.61% by the first 50.

The importance of each eigenface becomes less further and does not contribute much to the total importance as more eigenfaces are added. The first 400 eigenfaces represents 99.42% of the total variance. So, addition of more eigenfaces after 400 has nearly no contribution to the total variance. Therefore, the importance of an eigenface decreases almost quadratically up to 400 and then nearly stays constant at 0 up to 10,304 as shown in Figure 3.

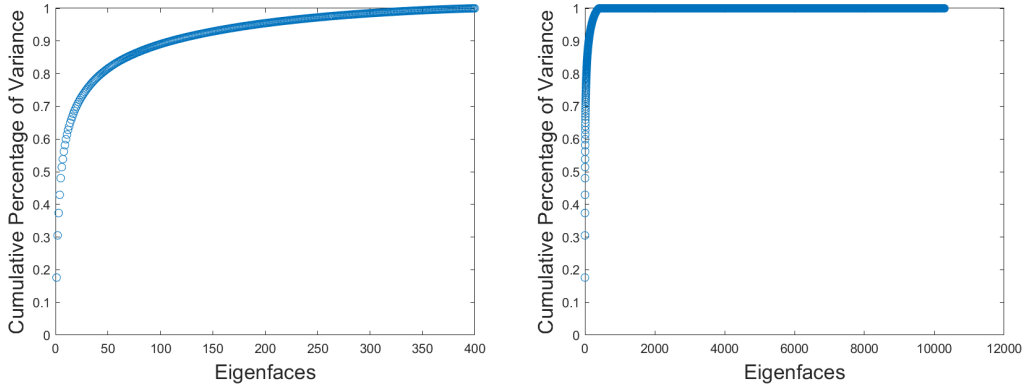


Figure 2: Eigenfaces and the Cumulative Percentage of Variances Captured. Ranging from 1 to 400 on the left and 1 to 10,304 on the right.

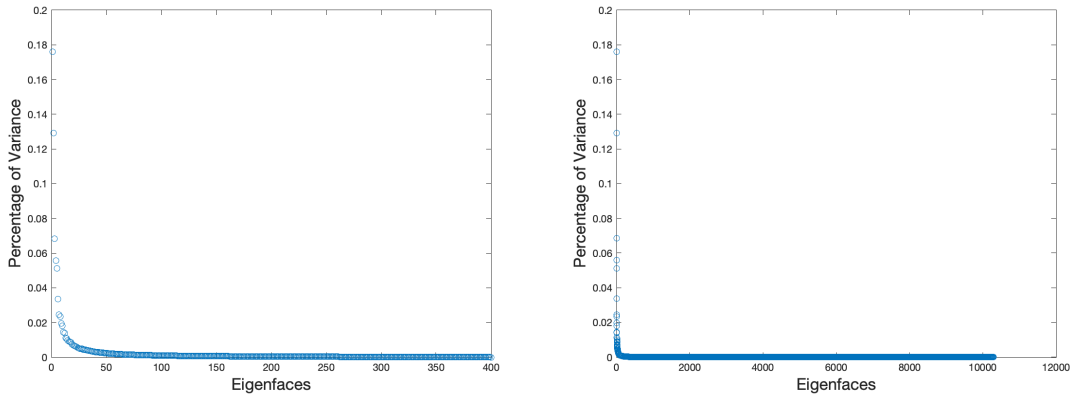


Figure 3: Eigenfaces and the Percentage of Variances Captured. Ranging from 1 to 400 on the left and 1 to 10,304 on the right.

3 Facial Image Reconstruction with PCA

3.1 What is the difference between reconstructed and original images as the number of eigenfaces used in reconstruction increases?

As the number of eigenfaces used in facial image reconstruction increases, the quality of the reconstructed image increases as well. Using a low number of eigenfaces results in a poor quality and somewhat “ghostly” reconstructed facial image as in Figure 4 (Turk & Pentland, 1991). Using a high number of eigenfaces alternatively produces a reconstructed facial image which is almost parallel to that of its original inputted images, such as those shown in Figure 6. This is because those made from the highest number of eigenfaces account for the most variance between the inputted images (See Figure 2). This is the case when using 400 eigenfaces as this was found to capture 99.99% of the variance of the inputted images.



Figure 4: Reconstructed Facial Images using 50 Eigenfaces



Figure 5: Reconstructed Facial Images using 200 Eigenfaces



Figure 6: Reconstructed Facial Images using 400 Eigenfaces

3.2 How many eigenfaces are required to recover an original face with reasonable errors?

We defined reasonable error through the use of both the metric Mean Squared Error (MSE) and by comparing visual differences between the original and reconstructed facial images. We found that reconstructed images consisting of 350 eigenfaces generated good quality images with identifiable facial characteristics. These images also had a MSE of 11.073. As a result, we concluded that any reconstructed facial image made from at least 350 eigenfaces and with an MSE value equal to or below 11.073, should be of a good enough quality so as to resemble their inputted original face with reasonable error.



Figure 7: Reconstructed Facial Images using 350 Eigenfaces

4 Conclusion

Thus, using PCA, we were able to reduce the dimensions of images from 10,304 to 350 with a very low reconstruction error. Moreover, using 400 dimensions has almost 0 reconstruction error. The ratio of reduction in size is great for various applications. Also, visualizing the eigenfaces as images helped us understand what it represents. For instance, the leading eigenfaces display fundamental facial features like chin, mouth, nose, ears, hair, etc and all the images are the linear combination of these eigenfaces.