



Dimensionality Reduction & Classification

Project 4 - CISC 820 Quantitative Foundations

Presented by:

Image Adhikari, Ester Chen



Dimensionality reduction

Using PCA to extract eigenfaces

How do the leading
eigenfaces look like
as an image?

1st Eigenface



2nd Eigenface



3rd Eigenface



4th Eigenface



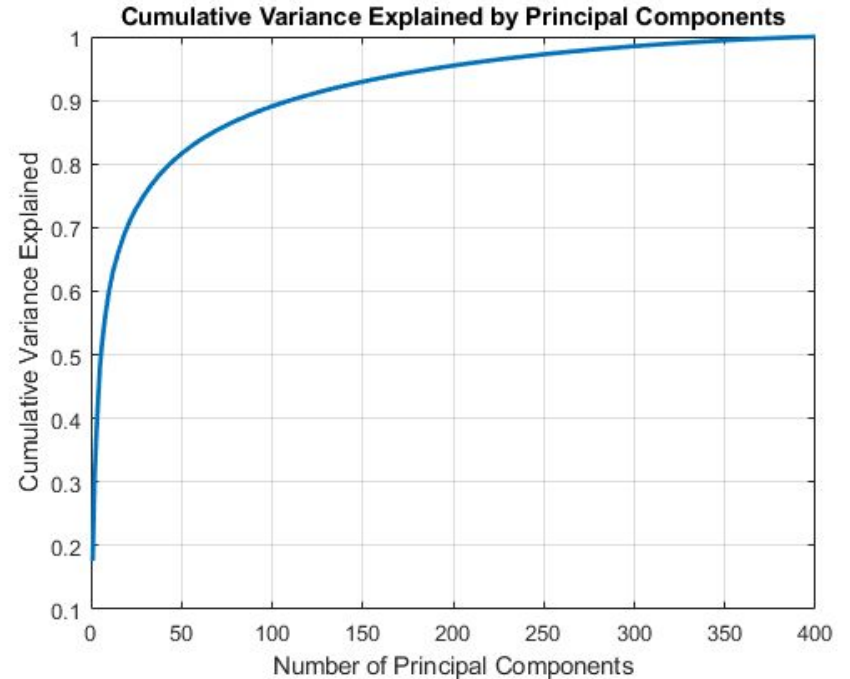
5th Eigenface



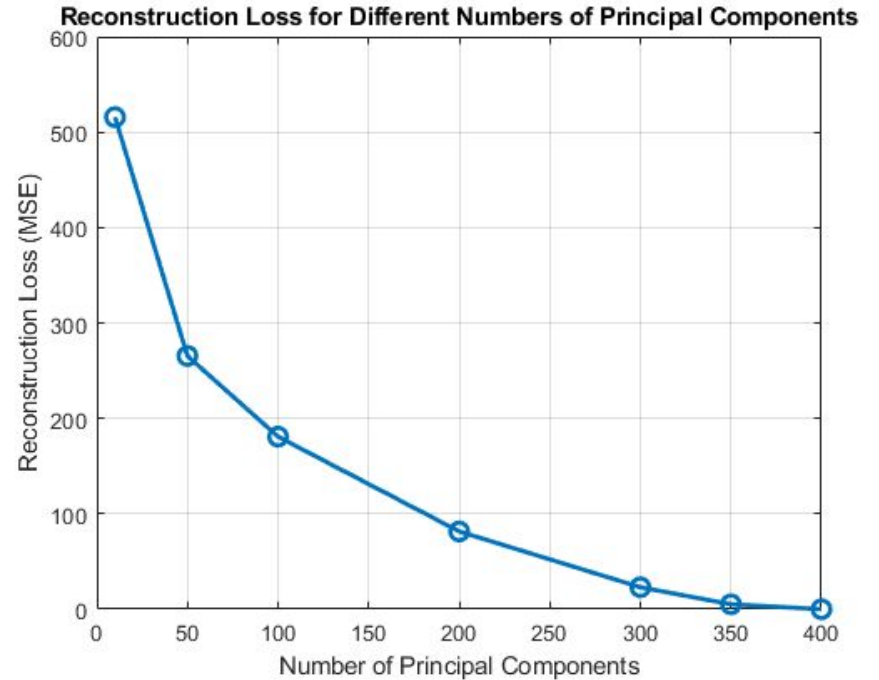
6th Eigenface



How does the
importance of the
eigenfaces decrease?



How does the
importance of the
eigenfaces decrease?





Face Reconstruction with PCA

Differences between reconstructed and original images using 50 Eigenfaces

Original Image (Subject 1)



Reconstructed: 50 PCs



Original Image (Subject 2)



Reconstructed: 50 PCs



Original Image (Subject 3)



Reconstructed: 50 PCs



Differences between reconstructed and original images using 100 Eigenfaces

Original Image (Subject 1)



Reconstructed: 100 PCs



Original Image (Subject 2)



Reconstructed: 100 PCs



Original Image (Subject 3)



Reconstructed: 100 PCs



Differences between reconstructed and original images using 150 Eigenfaces

Original Image (Subject 1)



Reconstructed: 150 PCs



Original Image (Subject 2)



Reconstructed: 150 PCs



Original Image (Subject 3)



Reconstructed: 150 PCs



Differences between reconstructed and original images using 200 Eigenfaces

Original Image (Subject 1)



Reconstructed: 200 PCs



Original Image (Subject 2)



Reconstructed: 200 PCs



Original Image (Subject 3)



Reconstructed: 200 PCs



Differences between reconstructed and original images using 250 Eigenfaces

Original Image (Subject 1)



Reconstructed: 250 PCs



Original Image (Subject 2)



Reconstructed: 250 PCs



Original Image (Subject 3)



Reconstructed: 250 PCs



Differences between reconstructed and original images using 300 Eigenfaces

Original Image (Subject 1)



Reconstructed: 300 PCs



Original Image (Subject 2)



Reconstructed: 300 PCs



Original Image (Subject 3)



Reconstructed: 300 PCs



Differences between reconstructed and original images using 350 Eigenfaces

Original Image (Subject 1)



Reconstructed: 350 PCs



Original Image (Subject 2)



Reconstructed: 350 PCs



Original Image (Subject 3)



Reconstructed: 350 PCs



Differences between reconstructed and original images using 400 Eigenfaces

Original Image (Subject 1)



Reconstructed: 400 PCs



Original Image (Subject 2)



Reconstructed: 400 PCs



Original Image (Subject 3)



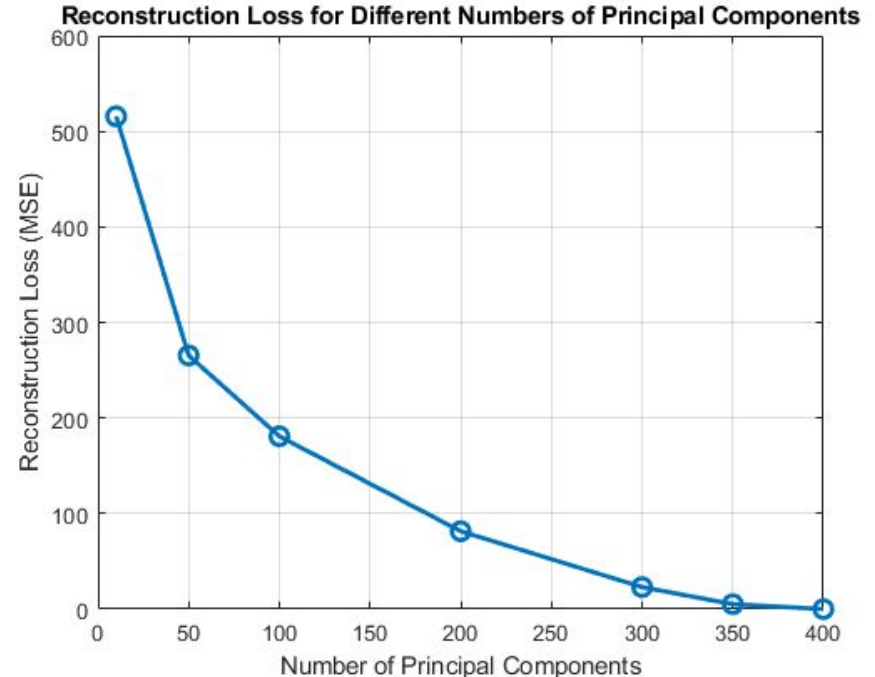
Reconstructed: 400 PCs



Number of eigenfaces required for reconstruction

Reasonable Error was defined by visual observation and Mean Squared Error (MSE):

- 350 Eigenfaces required
- MSE of 5.14





Classification (Face recognition and identification)

Training Data: Only face images.

Test Data: 120 face images, 30 non-face images.

Challenge: No exposure to non-face images during training.

PCA for Feature Extraction

- PCA calculated on training (face) images.
- Test data projected into PCA space.
- Normalization applied using training data parameters.



Classification: KNN and KMeans

Measured **distance to nearest neighbor** in PCA space.

Applied a threshold derived from validation data.

If distance > threshold → **Non-Face**.

K-means: Unsupervised Clustering

- Clusters: 2 (face and non-face).
- Test images assigned to clusters.
- Perfect separation achieved.



Key Takeaways

- **KNN**: Threshold-based distance → perfect classification.
- **K-means**: Unsupervised clustering → perfect separation.
- PCA transformation created distinct face/non-face separation.

Method		Task 1 Accuracy	Task 2 Accuracy
LR Classifier	Misclassified non-face images.	0.8	0.54167
KNN	Distance-based thresholding.	0.97	0.56667
K-Means	Two cluster separation.	1	-

Table 1: Different methods and Task Accuracies

Conclusion

- PCA and Reconstruction Quality

Increasing the number of **principal components** improves reconstruction quality.

Cumulative variance and **reconstruction loss** confirm this behavior

- PCA Reduces Dimensionality Effectively

Original dimension: **10,304**.

Reduced to **350–400 dimensions**.

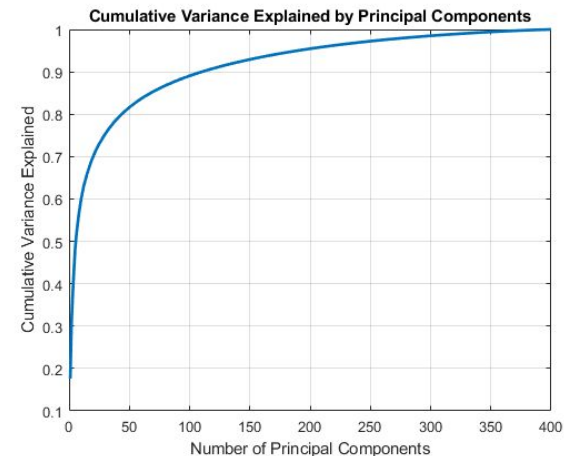
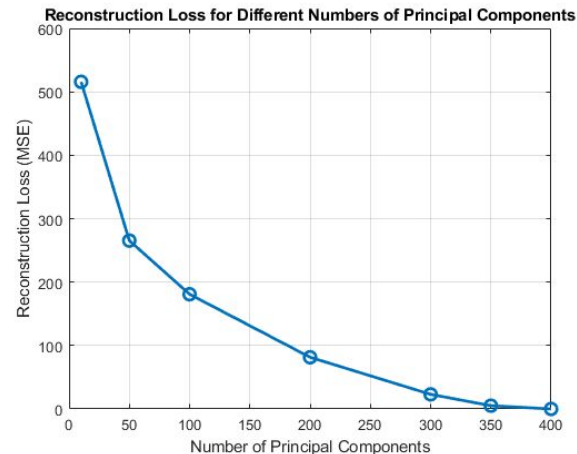
Top components capture the majority of variance.

At **350 dimensions**: **MSE is extremely low**.

At **400 dimensions**: **Reconstruction loss is effectively zero**.

- Eigenfaces as Features

- **Eigenfaces** represent fundamental facial features:
 - Chin, mouth, nose, ears, hair, etc.
- Faces are expressed as a **linear combination of eigenfaces**.
- **More components** → **More detailed reconstructions**.





Takeaways

1. PCA reduces high-dimensional data efficiently.
2. **Eigenfaces** capture the most important facial features.
3. **350–400 components** → high-quality reconstructions.
4. PCA achieves dimensionality reduction with **minimal loss**.

reconstruction_losses

400	4.2230455871106E-27
350	5.14066039119728
300	23.0254724432842
200	81.4994011544854
100	180.863873943162
50	265.414316904561
10	515.681925614058



**Thank you &
Questions**