2D Face Recognition

Part – 2

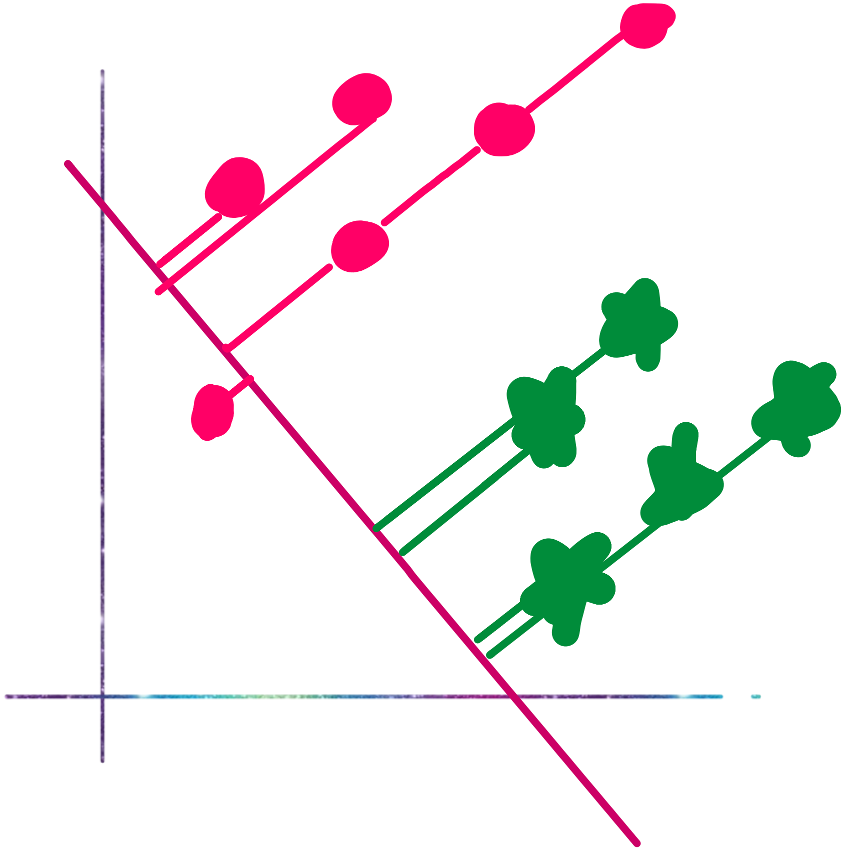
Ravitha N

Linear Discriminant Analysis

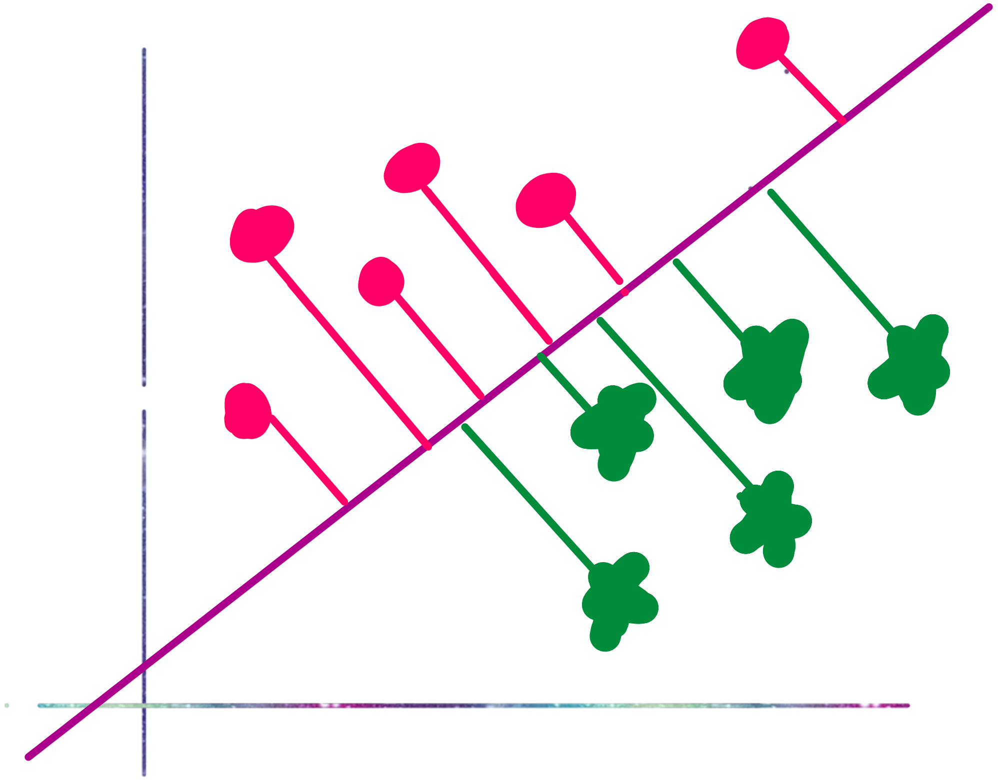
- LDA is used for both dimensionality reduction and classification
- In PCA, objective was to increase the scatter between the data points
- In LDA, it finds optimal vectors which minimize the scatter within class(within data points belonging to the same class) and maximize the scatter between the classes .

Interpretation

Scenario with LDA



Scenario with PCA



Data points belonging to the same class are clustered together

How to Find Optimal Vectors in LDA

- Computes Inter class Scatter Matrix

$$S_B = \sum_{i=1}^{k} N_i(\psi_i - \psi)(\psi_i - \psi)^t$$

K – number of classes

N_i - number of samples in each class

- Compute the within class scatter matrix

$$s_w = \sum_{i=1}^{k} \sum_{\Gamma_k \in C_i} (\Gamma_k - \psi_i)(\Gamma_k - \psi_i)^t$$

Find Optimal Vectors that maximize S_B and minimize s_w

$$W_{Opt} = \arg\max \left\{ \frac{w^t s_b w}{w^t s_w w} \right\}$$

These optimal vectors are known as Fisher Faces and are generalized eigen vectors of s_B and s_w

Questions

- How many optimal vectors will LDA produce?

  LDA produces k-1 vectors (where k indicate number of classes)

- Is there a formulation to identify the optimal vector from these matrices?

  Easiest way in obtaining these vector is $s_w^{-1}(\psi_1 - \psi_2)$

Pause Time

	LDA	**PCA**
Objective Criterion	Maximize Scatter between classes and minimize scatter within class	Maximize scatter between data points
# Number of Optimal Vectors	K-1 (K denotes number of classes)	M(Number of Images)
Procedure to find Optimal Vector	$s_w^{-1}(\psi_1 - \psi_2)$	Find Eigen Value for Covariance Matrix $(A - \lambda x) = 0$

Problem - LDA

Let x_1 represent data points belonging to Class 1

Let x_2 represent data points belonging to Class 2

$$x_1 = \{(4,1), (2,3), (2,4), (3,6), (4,4)\}$$
$$x_2 = \{(9,10), (6,8), (9,5), (8,7), (10,8)\}$$

# of features = 2

# of points in class 1 = 5

# of points in class 2 = 5

Step 1: Compute Mean for Classes

$$\psi_1 = \frac{1}{5}\begin{bmatrix} 4+2+2+3+4 \\ 1+3+4+6+4 \end{bmatrix}$$

$$= {}^1/_5 \begin{bmatrix} 15 \\ 18 \end{bmatrix} = \begin{bmatrix} 3 \\ 3.6 \end{bmatrix}$$

$$\psi_2 = \frac{1}{5}\begin{bmatrix} 9+6+9+8+10 \\ 10+8+5+7+8 \end{bmatrix}$$

$$= {}^1/_5 \begin{bmatrix} 42 \\ 38 \end{bmatrix} = \begin{bmatrix} 8.4 \\ 7.6 \end{bmatrix}$$

Step 2: Compute within class scatter matrix

$$s_w = \sum_{i=1}^{k} \sum_{\Gamma_k \in C_i} (\Gamma_k - \psi_i)(\Gamma_k - \psi_i)^t$$

$$s_w = s_1 + s_2$$

Find s_1

$$x_1 = \{(4,1), (2,3), (2,4), (3,6), (4,4)\}$$

$\psi_1 = \begin{bmatrix} 3 \\ 3.6 \end{bmatrix}$ $\phi_1 = \{(1, -2.6), (-1, -0.6), (-1, 0.4), (0, 2.4), (1, 0.4)\}$

S1 = $\frac{1}{5}\begin{bmatrix} 1 & -1 & -1 & 0 & 1 \\ -2 \cdot 6 & -0.6 & 0.4 & 2 \cdot 4 & 0.4 \end{bmatrix} \begin{bmatrix} 1 & -2.6 \\ -1 & -0.6 \\ -1 & 0.4 \\ 0 & 2.4 \\ 1 & 0.4 \end{bmatrix}$

= $\frac{1}{4}\begin{bmatrix} 4 & -2 \\ -2 & 13.2 \end{bmatrix}$

= $\begin{bmatrix} 1 & -0.5 \\ -0.5 & 3.3 \end{bmatrix}$

Find s_2

$$x_2 = \{(9,10),(6,8),(9,5),(8,7),(10,8)\}$$

ψ_2 = $\begin{bmatrix} 8.4 \\ 7.6 \end{bmatrix}$

$$\phi_2 = \{(0.6,2.4),(-2.4,0.4),(0.6,-2.6),(-0.4,-0.6),(1.6,0.4)\}$$

S2= $\frac{1}{5}\begin{bmatrix} 0.6 & -2.4 & 0.6 & -0.4 & 1.6 \\ 2.4 & 0.4 & -2.6 & -0.6 & 0.4 \end{bmatrix}\begin{bmatrix} 0.6 & 2.4 \\ -2.4 & 0.4 \\ 0.6 & -2.6 \\ -0.4 & -0.6 \\ 1.6 & 0.4 \end{bmatrix}$

= $\frac{1}{4}\begin{bmatrix} 9.2 & -0.2 \\ -0.2 & 13.2 \end{bmatrix}$

= $\begin{bmatrix} 2.3 & -0.05 \\ -0.05 & 3.3 \end{bmatrix}$

$$s_w = s_1 + s_2 \qquad s_1 = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 3.3 \end{bmatrix} \qquad s_2 = \begin{bmatrix} 2.3 & -0.05 \\ -0.05 & 3.3 \end{bmatrix}$$

$$\psi_1 = \begin{bmatrix} 3 \\ 3.6 \end{bmatrix}$$

$$\psi_2 = \begin{bmatrix} 8.4 \\ 7.6 \end{bmatrix}$$

$$\psi_1 - \psi_2 = \begin{bmatrix} -5.4 \\ -4 \end{bmatrix}$$

$$s_w = \begin{bmatrix} 3.3 & -0.55 \\ -0.55 & 6.6 \end{bmatrix}$$

Optimal Vector = $s_w^{-1}(\psi_1 - \psi_2)$

$$s_w^{-1} = \frac{1}{(3.3*6.6)-(0.55*.0.55)} \begin{bmatrix} 6.6 & 0.55 \\ 0.55 & 3.3 \end{bmatrix} = \frac{1}{21.48} \begin{bmatrix} 6.6 & 0.55 \\ 0.55 & 3.3 \end{bmatrix}$$

$$= \begin{bmatrix} 0.31 & 0.025 \\ 0.025 & 0.154 \end{bmatrix}$$

$$s_w^{-1}(\psi_1 - \psi_2) = \begin{bmatrix} 0.31 & 0.025 \\ 0.025 & 0.154 \end{bmatrix} \begin{bmatrix} -5.4 \\ -4 \end{bmatrix} = \begin{bmatrix} -1.77 \\ -0.751 \end{bmatrix}$$

Projection Vectors

$\psi = {}^{1}/_{2} \begin{bmatrix} 3+8.4 \\ 3.6+7.6 \end{bmatrix}$

$= \begin{matrix} 5.7 \\ 5.6 \end{matrix}$

Data Points in the Original Space	Mean Subtracted Data Points	Data Points in Projected Space
(4,1)	-1.7,-4.6	6.46
(2,3)	-3.7,-2.6	8.50
(2,4)	-3.7,-1.6	7.75
(3,6)	-2.7,0.4	4.47
(4,4)	-1.7,-1.6	4.21
(9,10)	3.3,4.4	-9.14
(6,8)	0.3,2.4	-2.33
(9,5)	3.3,-0.6	-5.39
(8,7)	2.3,1.4	-5.1
(10,8)	4.3,2.4	-9.41

Optimal Vector

$\begin{bmatrix} -1.77 \\ -0.751 \end{bmatrix}$

Find Class for a new datapoint

(8,1)

Subtract from $\psi = \begin{matrix} 5.7 \\ 5.6 \end{matrix}$

(2.3, -4.6)

- Transform Data through projection

$(2.3, -4.6)\begin{bmatrix} -1.77 \\ -0.751 \end{bmatrix}$ = -.6164

Point is closest to (-2.33) in the dataset

So, it belongs to class 2

Data Points in Projected Space	Distance w/rt to data point $\lvert \mathbf{d}_i - d_{new} \rvert$
6.46	7.07
8.50	9.11
7.75	8.37
4.47	5.09
4.21	4.83
-9.14	8.52
-2.33	1.71
-5.39	4.77
-5.1	4.48
-9.41	8.79

Pause Time

- Can LDA be used for Classification ?

Local Face Recognition Approaches

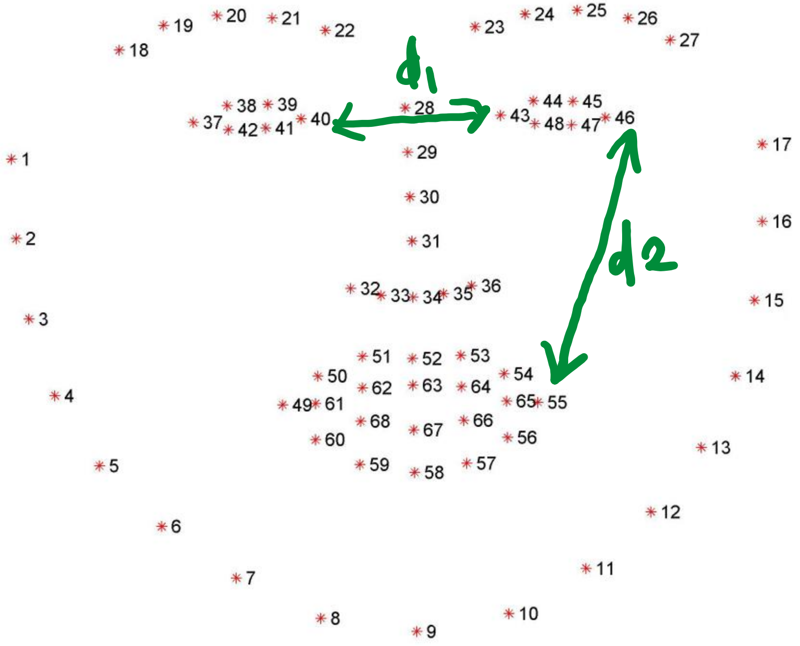
- Process face image as patches , interesting regions (or) key indicators

1. Geometric Techniques

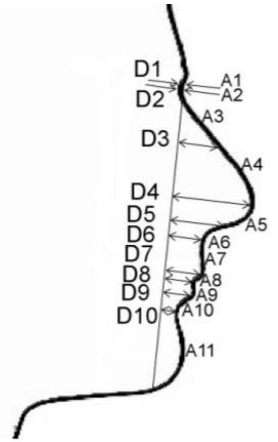
   Extract landmarks from face

   Use relative distance between landmarks

   Using both Frontal View and Profile View is said to Increase the accuracy of the system

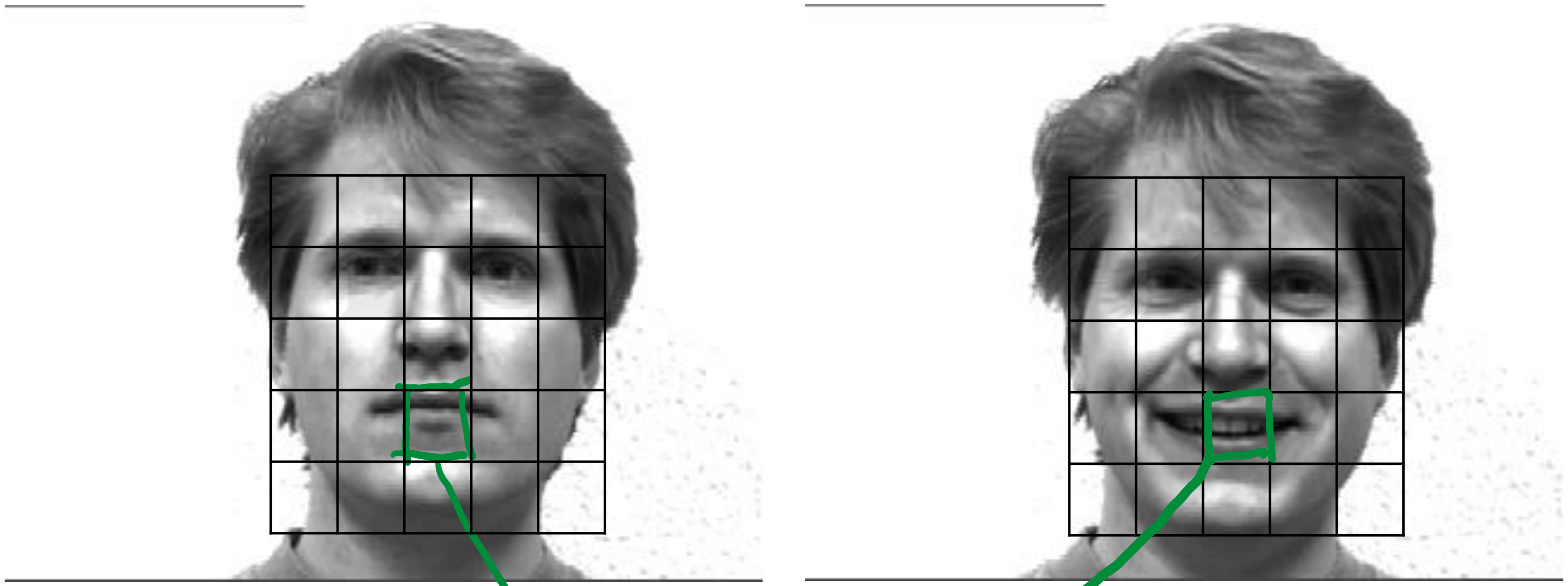


*Distance and the profile arcs
Considered for recognition from profile
View of the face[1]



[1]A scale space approach to face recognition from profiles

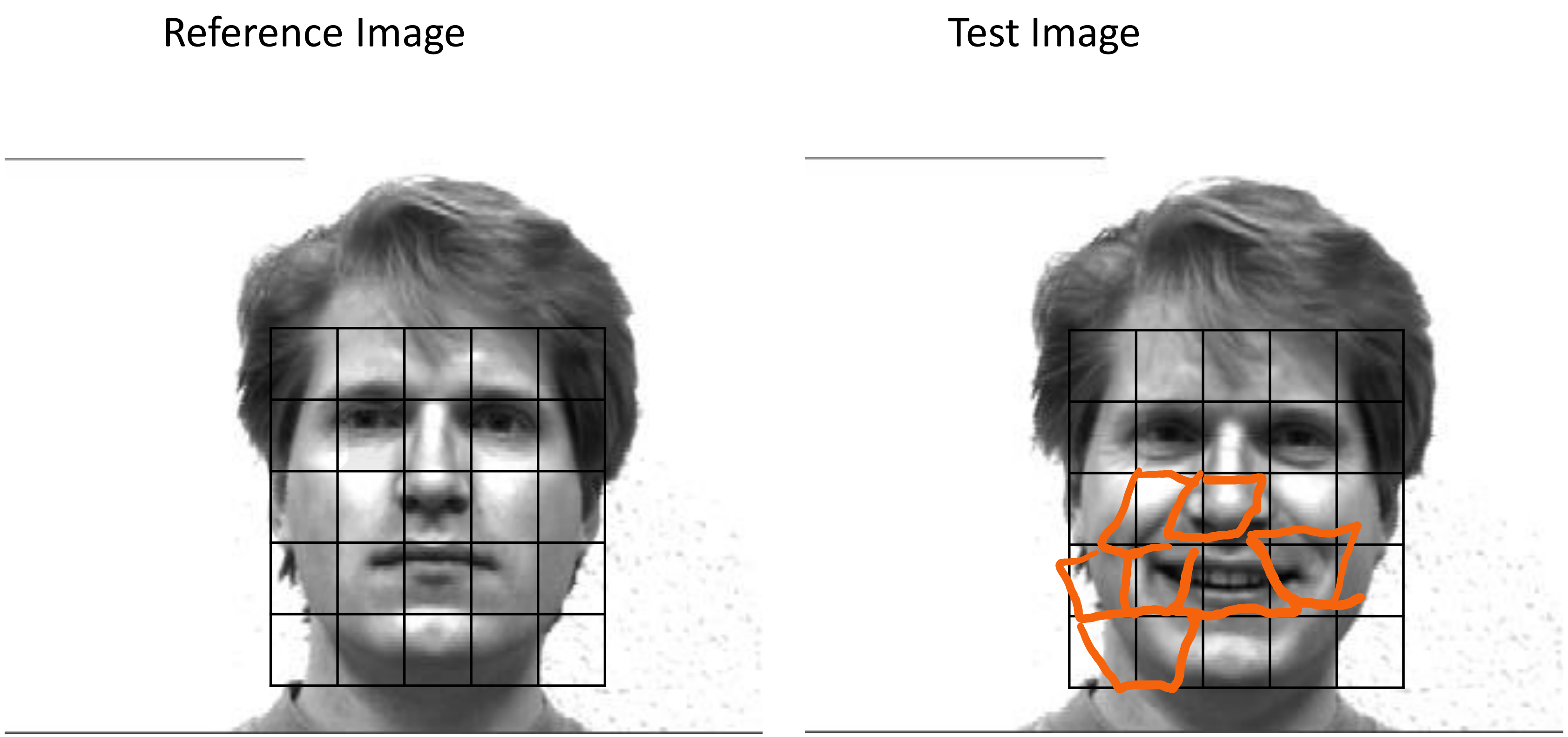
- Elastic Bunch Graph Matching Techniques (distortion invariant representation)



Grid of k nodes is superimposed on a face

And features are extracted from each grid

Will Features will be similar ??



During Matching face, nodes are moved until features corresponding to new position matches with reference node features

Food for thought

- How can the grid be modified ?

Useful Links & References

- http://www.sci.utah.edu/~shireen/pdfs/tutorials/Elhabian_LDA09.pdf
- **http://vision.ucsd.edu/~iskwak/ExtYaleDatabase/Yale%20Face%20Database.htm**
- **Amine Nait Ali, “Signal and Image Processing for Biometrics ”**