Face recognition From Single-image Per Person - A survey

One Sample Per Person Problem

Given a stored database of faces, the goal is to identify a Person from the database later in time in any different and unpredictable poses, lighting, etc. from just one image.

Advantages

1 → easy to collect samples
  ↳ Smart cards, Passport, Driver's license. etc.

2 → Save Storage Cost

3 → Save Computational Cost.

Objective:- How to improve Performance against extremely small sample size.

Face Recognition Methods :

Holistic methods :

↳ Identify face using whole face image as input.

Local methods :

→ Use local facial features for recognition.

Hybrid methods :

→ Use both → local & holistic features for recognition.

Holistic methods

→ PCA

→ Fisherface

→ $(PC)^2A$

→ 2DPCA

→ Noise model

→ Enlarge training set

→ New representation

SVD → Perturbation

→ Generate novel views

Parallel Deformation

→ Dis-advantages of Holistic methods

PCA

Preprocessing

- → Convert each image to same width & Height
- → Take training set of M images each with dimension → MXM
- → Create Matrix A

each column is a Vector representation of NXN image matrix

$$A = \left[\underbrace{\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \text{im-1} & \text{im-2} & \cdots\cdots & \text{im-M} \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}}_{M}\right]$$

NXN

$A \rightarrow (N^2 X M)$

→ Find mean face by taking average of each raw.

$$N^2 \left\{ \begin{bmatrix} \leftarrow \\ \vdots \\ \leftarrow \\ \vdots \\ \leftarrow \end{bmatrix} \right. \begin{array}{l} \dfrac{im_1(i,j) + im_2(i,j) \cdots}{M} \\ \cdots\cdots \vdots \cdots\cdots \\ \dfrac{im_1(N\times N) + im_2(N\times N) \cdots}{M} \end{array}$$

$$\underbrace{\quad}_{1}$$

→ Substract the mean Vector from each $N^2 \times 1$ Vector → so data become normalized → zero mean

→ Find Covariance Matrix

$AA^T \to N^2 \times N^2$

↳ Find eigenvectors →

total N^2 → tolarge

* Recovery of eigenvector from low-dimensional space to high-dimensional space.

Suppose $V_i \rightarrow$ eigenvector of

$A^TA \rightarrow (M \times 1)$

So,

$A^TAV_i = \mu_i V_i$

mutiply both side by A

$AA^TAV_i = \mu_i AV_i$

$C(AV_i) = (\mu_i A)V_i$

So $AV_i \rightarrow$ eigen vectors of C

$\underline{AV_i} \rightarrow (N^2 \times M) \times (M \times 1) \rightarrow \underline{N^2 \times 1}$

So by taking eigen vectors $V_1, V_2 \ldots V_M$ in lower space we can get the eigen vectors of AA^T

Higer space $\rightarrow$ $V_i = \sum_{k=1}^{M} A V_{1k} \leftarrow (N^2 \times M) \times (M \times 1) \rightarrow (N^2 \times 1)$

$l = 1, 2 \ldots M$

Recognition

→ accurate reconstruction of face is not needed.

total eigenvectors required for recognition << M
$\uparrow M'$

original Paper → M=16 then
M' = 7 ← top seven eigenvectors corresponding to heighest eigenvalues

→ each training image
↳ convert

$tr\text{-}1 \rightarrow tr_1$ - meanface
$\vdots \rightarrow tr_2 \rightarrow$ mean face
$tr \rightarrow m$ tr_m - mean face.

→ Compute weight vector for each training image.

$w_{11} = [e.g.V\text{-}1]^T [tr\text{-}1]$
$w_{12} \rightarrow [eg..V \rightarrow 2]^T [tr \rightarrow 1]$
$\vdots$
$w_{1M} \rightarrow [eg \rightarrow V \rightarrow M)^T [tr \rightarrow m]$

→ represent each image with weight vectors

$Tr_1 \rightarrow [W_{11}\ W_{12}\ ---\ W_{1M}]$

$Tr_M \rightarrow [W_{M1},\ W_{M2}\ ---\ W_{MM}]$

→ Test image → substract mean face → find weight vector

$[W_{11}, W_{12}, ...\ W_{1m}]$

→ Find Euclidian Distance of test image vector with each of the training image,

↳ select the image with minimum distance.

weight vectors

$$E^2 = \|\underset{\text{test}}{\Omega} - \underset{\text{each training}}{\Omega_k}\|^2$$

Fisherface

→ face classification

→ Assume → each face class → enough sample images are available for training.

→ Maximize between class scatters & minimize within class scatters

Steps

→ Suppose total number of classes are 4 with 2 images for training in each.

$$\underbrace{\begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} \begin{bmatrix} | \\ a_2 \\ | \end{bmatrix}} \quad \underbrace{\begin{bmatrix} | \\ b_1 \\ | \end{bmatrix} \begin{bmatrix} | \\ b_2 \\ | \end{bmatrix}} \quad \underbrace{\begin{bmatrix} | \\ c_1 \\ | \end{bmatrix} \begin{bmatrix} | \\ c_2 \\ | \end{bmatrix}} \quad \underbrace{\begin{bmatrix} | \\ d_1 \\ | \end{bmatrix} \begin{bmatrix} | \\ d_2 \\ | \end{bmatrix}}$$

→ Find avg. face.

$$\begin{bmatrix} \mu \end{bmatrix} = \frac{1}{M} \left[a_1 + a_2 + b_1 \cdots + d_2 \right]$$

→ Compute avg. face of each person e.g. here for 2 persons per class $\mu_1 = \frac{1}{2}\left[a_1 + a_2\right]$ same for μ_2, μ_3 & μ_4.

→ Substruct avg. face of each class from traing faces.

$$a_1' \rightarrow \begin{bmatrix} | \\ a_1 - \mu_1 \\ | \end{bmatrix} \quad \overset{a_2'}{\begin{bmatrix} | \\ a_2 - \mu_1 \\ | \end{bmatrix}} \quad \cdots \quad \overset{d_1'}{\begin{bmatrix} | \\ d_1 - \mu_4 \\ | \end{bmatrix}} \quad \overset{d_2'}{\begin{bmatrix} | \\ d_2 - \mu_4 \\ | \end{bmatrix}}$$

within class scatter matrix

$$S_1 = a_1' a_1'^T + a_2' a_2'^T$$

$$\vdots$$

$$S_4 = d_1' d_1'^T + d_2' d_2'^T$$

$$S_W = S_1 + S_2 + S_3 + S_4.$$

between class scatter matrix

$$S_B = 2(\mu_1 - \mu)(\mu_1 - \mu)^T + \cdots + 2(\mu_4 - \mu)(\mu_4 - \mu)^T$$

total images per class.

Objective → $W_{opt} = \underset{W}{\operatorname{argmax}} \dfrac{|W^T S_B W|}{|W^T S_W W|}$

$[W_1, W_2 \ldots W_m]$

If S_W is non singular

Columns of W are eigenvectors of $S_W^{-1} S_B$

If S_W is singular.

Apply PCA → reduce dimensionality.

$$W_{PCA} = \arg\max |W^T S_T W|$$

$$S_T = \sum_{k=1}^{M} (x_k - \mu)(x_k - \mu)^T$$

$$W_{fld} = \arg\max_W \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|}$$

$$\boxed{W_{opt} = W_{fld} W_{pca}}$$

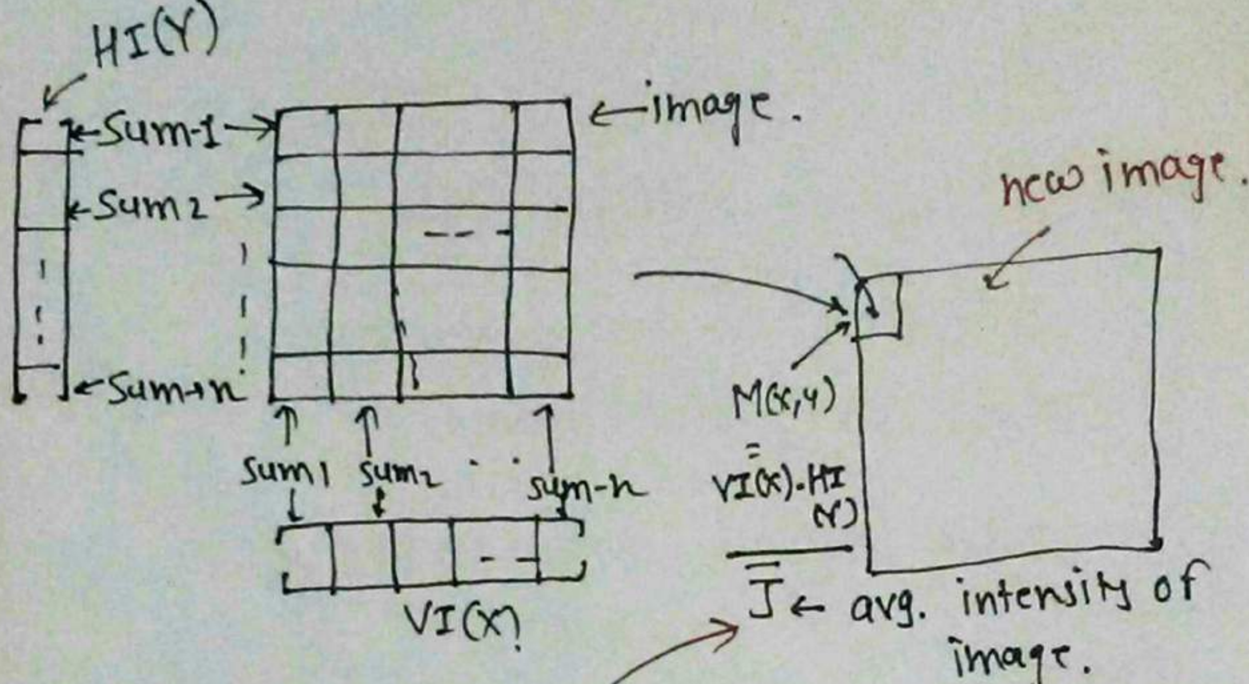
→ Classify new face by nearest neighbor

Disadvantage : → Expensive calculation time
→ Large storage.

→ fail when each ~~persons~~ ~~in~~ class contain one training face

→ It perform best in variation in lighting & facial expression.

$(PC)^2A$



Pre Processing

→ unimportant features → faded out and important features become more silent after Pre-Processing.

→ Results → improved when tested for one sample per person.

2DPCA

Idea

2D matrix → not 1D vector.

→ No need to transform Image Matrix to Vector.

Suppose each image size $m \times n$

& total images are M

then Covariance Matrix

$$G_t = \frac{1}{M} \sum_{j=1}^{M} (A_j - \bar{A})^T (A_j - \bar{A})$$

G_t: n×n matrix

A_j: training images

$\bar{A}$: mean face.

$$J(X) \rightarrow X^T G_t X$$

$J(X)$: generalized total scatter criterion

X: eigen vectors of size $n \times 1$

→ Maximize criteria

→ (Optimal Projection Axis) X_{opt}

X_{opt}: eigen vectors of G_t corresponding to largest e.g. value.

Advantages

→ Size of image covariance-matrix → much smaller.

↳ easy to evaluate

→ Less time to determine corresponding eigenvectors.

Disadvantages

Storage → PCA → $\omega_1, \omega_2 \ldots \omega_n$

2DPCA →

Principle Component Vectors.

How to reduce the Dimensions → not clear.

Noise model :

real world :

↳ Certificate photographs
→ currepted by various scratches, blur or dis-coloration.

Paper :

Authentication System to handle the Problem with one training image per Person.

Idea :

Synthesize multiple new face images which imitate the currepted images for recognition.

Imitation → Noise model

Contrast　Brightness　Gaussian Blur.

→ By changing the values of noise parameters → several corrupted images corresponding to one sample are imitated.

error rate → in 137 scanned Id-cards

↳ 1.32%

↓

indicates

method can significantly improve the similarity between corrupted images & training images.

For Better recognition ?

↳ enlarge the size. of training set.

Synthesizing Virtual samples

New representations for single training samples.

New Visual samples → not existed in Database.

↳ ROCA → Representational Oriented Component analysis.

↳ One sample per persons

Preprocessing :- apply linear & non-linear filters to → 150 representations

↳ OCA classifier is built on each representation

→ All OCA classifiers are combined to give weighted linear sum for final Deci

Image Perturbation

↓

Construct new Representation

→ For given image X → error range is set for Vertical & Horizontal localization Error.

A new sample is generated by changing Vertical & Or Horizontal Coordinate Value.

SVD Perturbation

face recognition → Single image Per Person.

$I \rightarrow$ Input image $\rightarrow M_1 \times M_2$

$P = U \Sigma^n V^T$

$\Sigma \rightarrow M_1 \times M_2 \rightarrow$ Diagonal Matrix

$U \rightarrow M_1 \times M_1 \rightarrow$ orthogonal matrix

$V \rightarrow M_2 \times M_2 \rightarrow$ orthogonal matrix

→ Combine I linearly with P to generate new image with following equation.

$$J = \frac{I + \alpha P}{1 + \alpha} \rightarrow 0 < \alpha < 1$$

→ Change the values of α to generate additional images.

J → works better by minor changes in expression, illumination & Occlusions.

Drawback

→ Generated Virtual images may be highly correlated.

→ New sample should not be considered as independent training images.

Generate Novel Views
goal

Virtual samples → Should
↳ occupy Different locations in the face space & represent specific variations of face images

Solution :

1. apply geometrical transformation
   → rotation
   → scale etc.
2. Synthesizing the face image under different pose & different illumination condition.

Technique :

Parallel Deformation

→ Generate novel view of single face image under different poses.

Need to learn transformation :

$\Delta X \rightarrow$ Difference between original face X & its reference face

$\Delta X_i \rightarrow$ Difference between images of other prototypical face to the same reference face

Assume. → linear class Assumption

$$\Delta X = \sum_{i=1}^{q} \alpha_i \Delta X_i$$

Obtain α_i by minimizing.

$$\left\| \Delta X - \sum_{i=1}^{q} \alpha_i \Delta X_i \right\|$$

→ one can use it to generate novel images of each single face.

Disadvantages :

→ Difficult to generate for diff. pose.

→ Novel image need → correspondance between real & reference images

Disadvantages of Holistic methods

→ One single feature Vector to represent each face image.

→ Sensitive to large appearence change due to
  - → expression
  - → illumination
  - → Pose & partial occlusion

Solution

↳ Local features

→ Original face → represented By set of low dimensional Feature Vectors.

→ Problem of Dimensionality can be removed.

→ Recognize face Based on its Parts

→