Face Recognition – Part I

"Fundamentals"

Dr. –Ing. Saquib Sarfraz

Prof. Rainer Stiefelhagen

25.11.2013

Outline

- Motivation: Why face recognition ?

- Technical Approaches
  - Feature-Based
  - Eigenfaces (PCA) and variants
  - Fisherfaces (LDA)
  - Local Appearance based

Visual Perception of Humans

- Where are people?
    - ➢ Person detection and tracking
- **Who are they?**
    - ➢ Person identification
        - ➢ **Mainly face recognition**
- What do they do?
    - ➢ Body posture tracking
    - ➢ Gesture & action recognition
- With whom do they interact, what is their intention?
    - ➢ Head pose estimation, tracking of gaze tracking & pointing direction
- How are they (affect, stress ...)
    - ➢ Facial expression analysis

Face Recognition for Human-Computer Interaction

- Person identification is need to build personalized human-computer interfaces
- Face recognition is a non-intrusive method to obtain the identity of a person
- Other techniques
  - Speaker identification
  - Finger print, iris recognition, …
- Also: many security & safety related applications

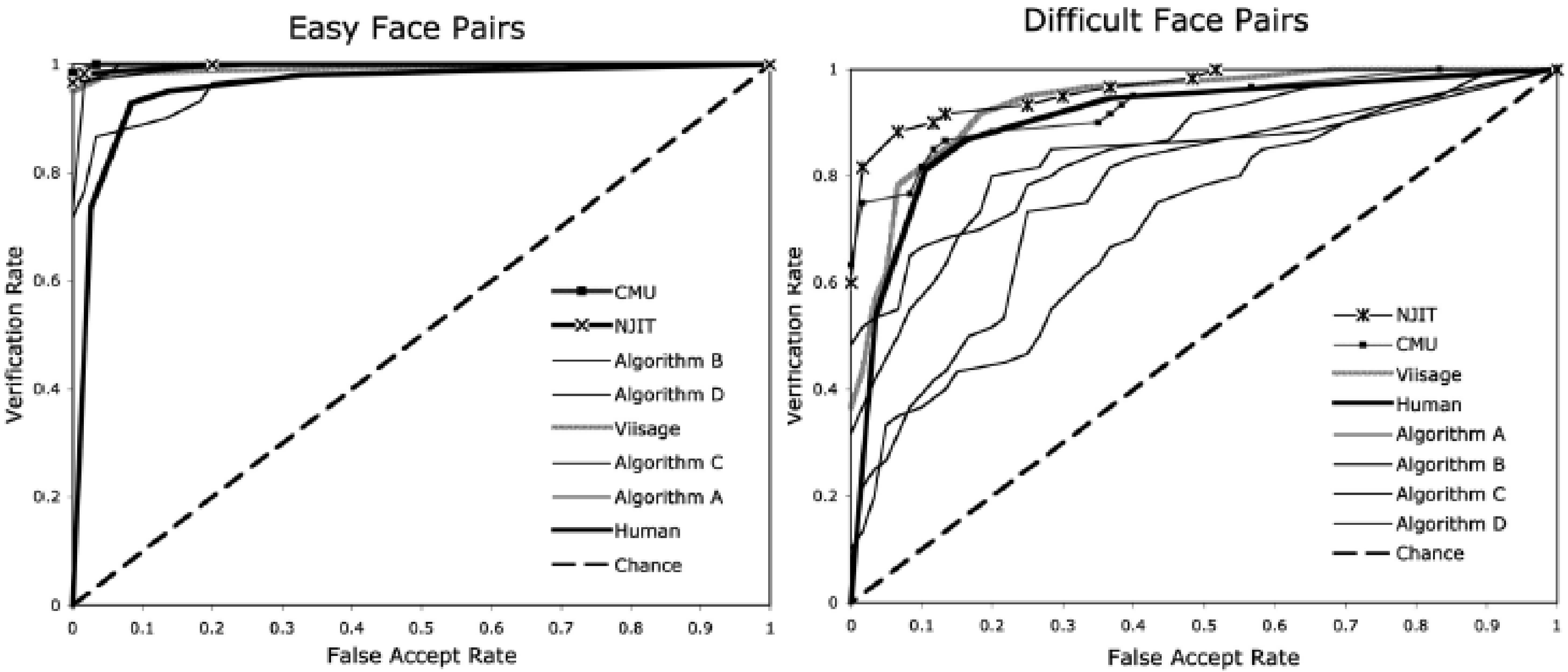
Applications of Face Recognition

Areas	Specific applications
Entertainment	Video game, virtual reality, training programs
	Human-robot-interaction, human-computer-interaction
Smart Cards	Drivers' licenses, entitlement programs
	Immigration, national ID, passports, voter registration
	Welfare fraud
Information security	TV Parental control, personal device logon, desktop logon
	Application security, database security, file encryption
	Intranet security, internet access, medical records
	Secure trading terminals
Law enforcement and surveillance	Advanced video surveillance, CCTV control
	Portal control, postevent analysis
	Shoplifting, suspect tracking and investigation

- Current market size of "face recognition for security domain" is 350 million USD, projected to exceed 1 billion USD in 2014
  - (*Source: Biometrics Market and Industry Report 2009-2014, International Biometric Group, Oct. 2008)

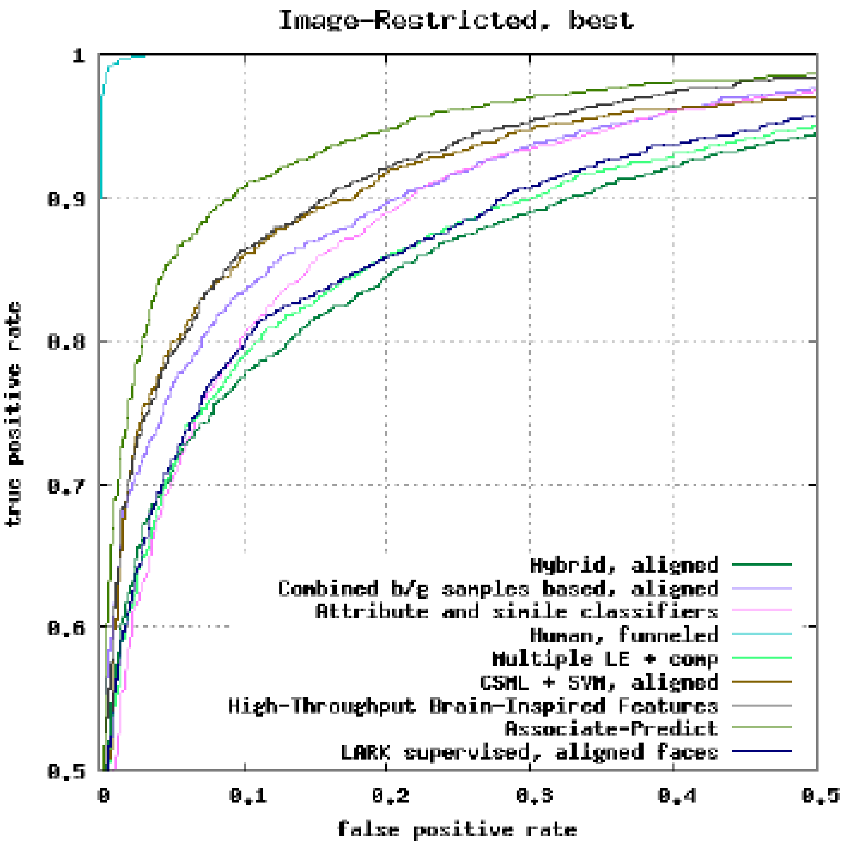
Which image pairs are from the same person?

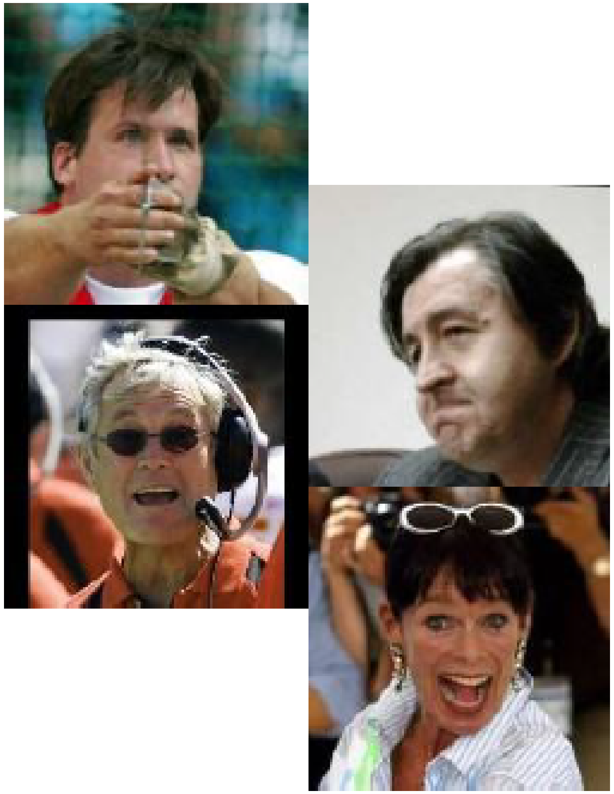
Face recognition : Human vs. Machine



- Alice J. O'Toole, et al. "Face Recognition Algorithms Surpass Humans Matching Faces over Changes in Illumination", PAMI, vol. 29, no. 9, 2007
  - Large scale face recognition evaluation : FRGC (Face Recognition Grand Challenge)
  - Machine performs better than human for easy face pairs

Faces in the wild : Human vs. Machine



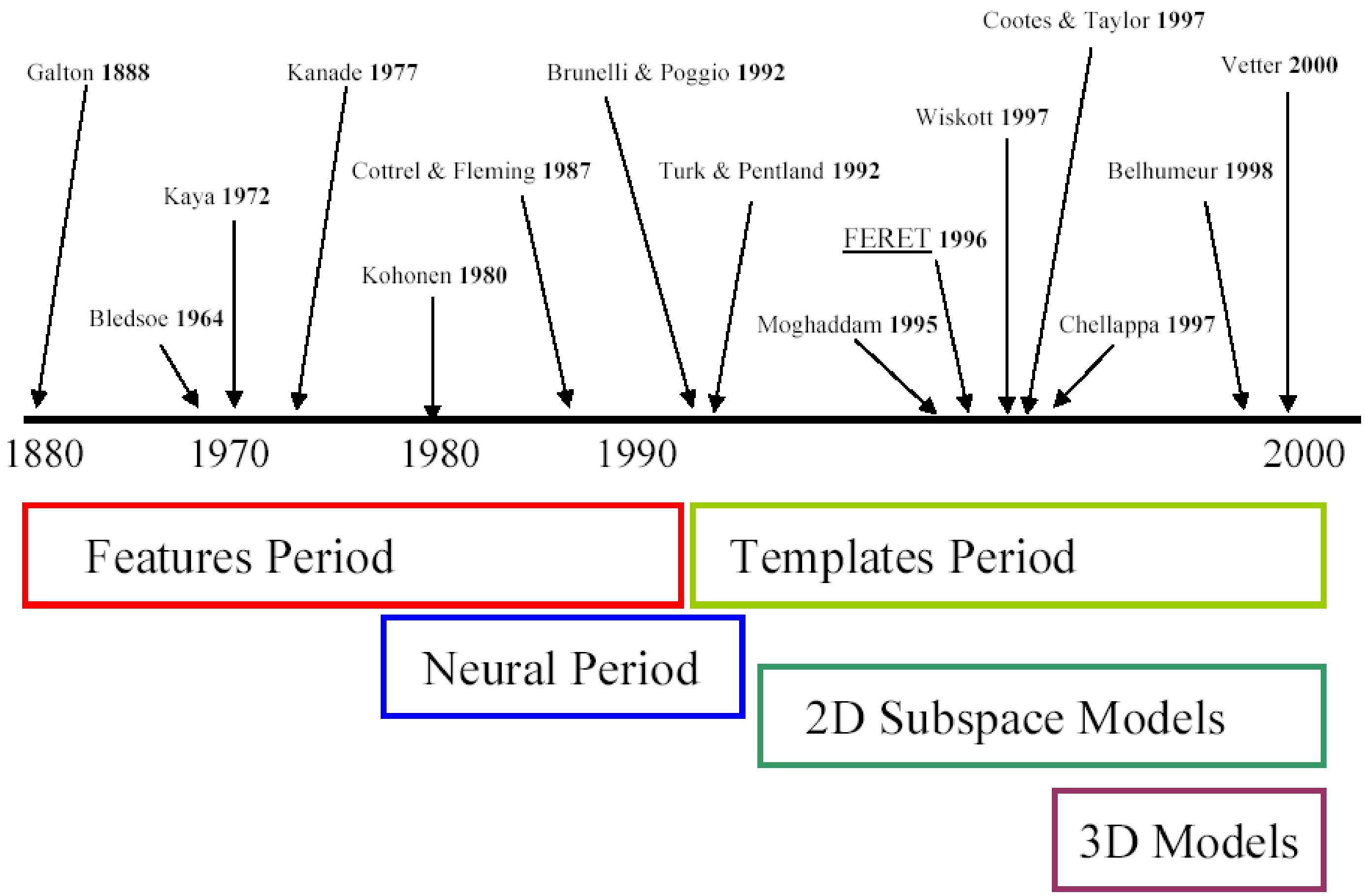


- **Source :** http://vis-www.cs.umass.edu/lfw/results.html
  - Matching face images from Internet
  - Human still performs far better than machine

Outline

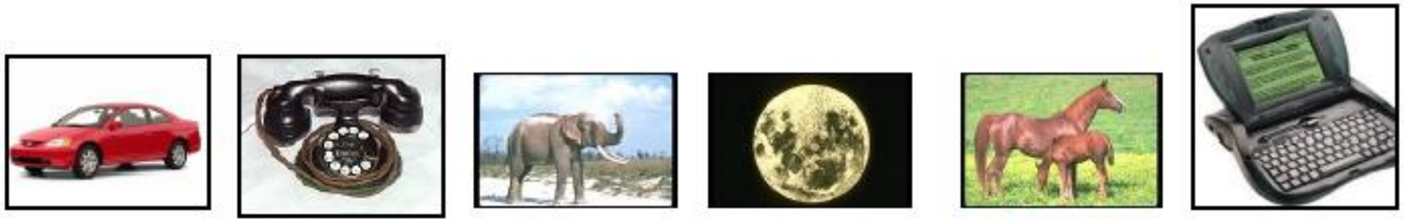
- Motivation
- Technical Approaches
  - Feature-Based Approaches
  - Appearance Based Approaches
    - Eigenfaces (PCA) and variants
    - Fisherfaces (LDA)

A Brief History (1900-2000)



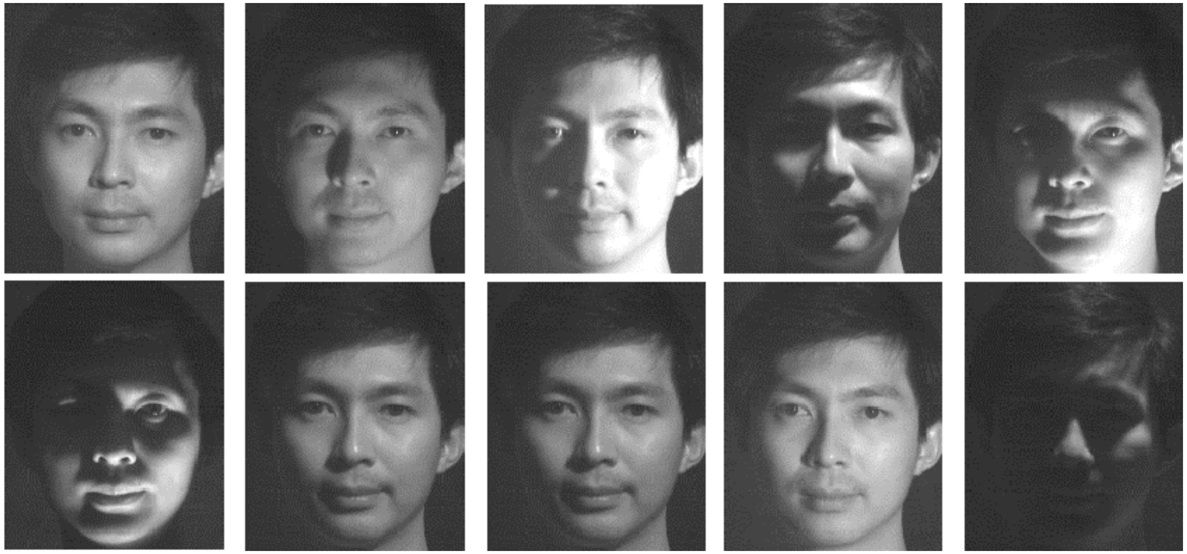
Object recognition perspective

- It is not general object recognition!



- It is a single-class object recognition task

Main problem



The variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity.

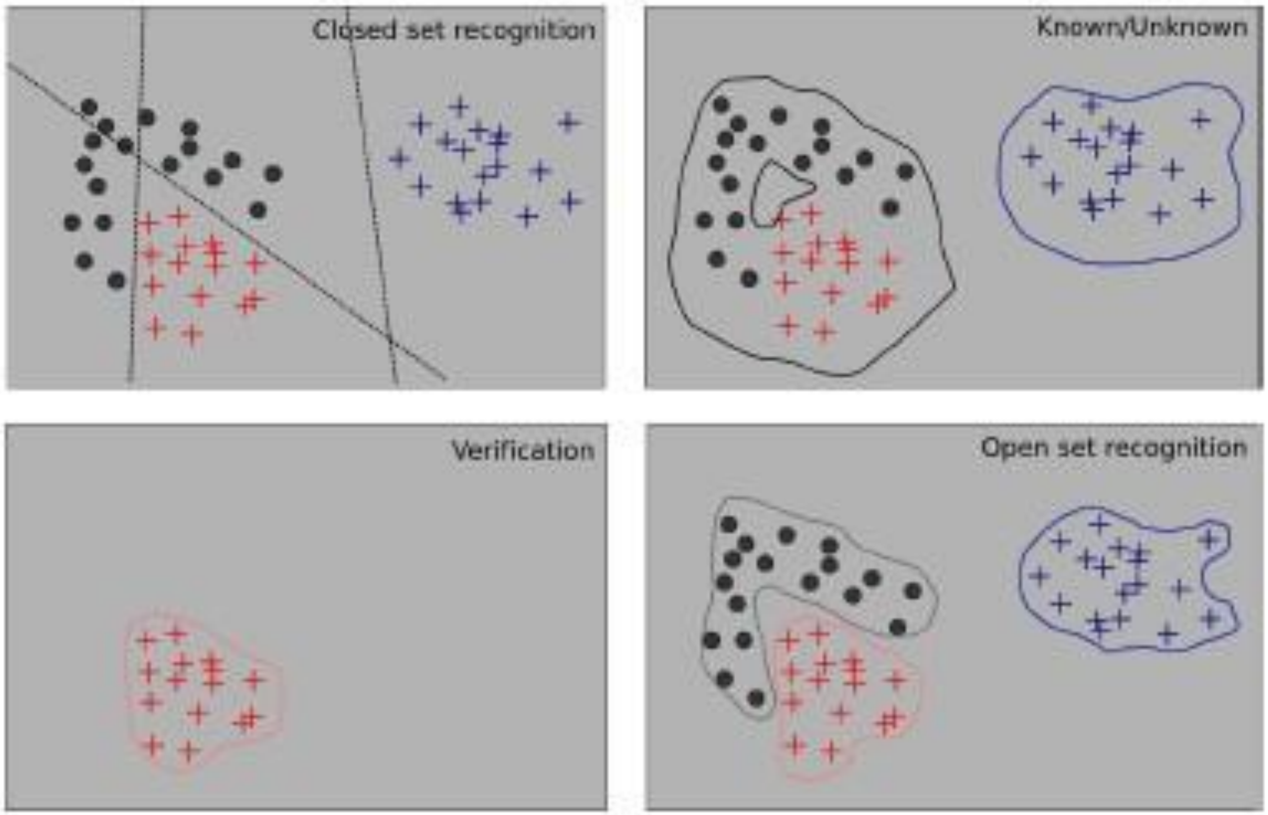
-- Moses, Adini, Ullman, ECCV'94

12

Face Recognition - Challenges

- Extrinsic variations of face images:
  - <u>Illumination variations</u>
  - <u>View-point variations (</u>frontal and non-frontal, …)
  - Imaging process, (low) resolution
  - Occlusions (Other objects or people, sun glasses, hats, beards. Make-up, etc.)
- Intrinsic variations of face image:
  - Facial expressions
  - Aging
- Acquiring the input image!
  - Face detection, tracking, normalization (often done by eye detection)
- All of these problems have to be dealt with in real environments !

Face Recognition Tasks



- Decision surfaces for three different people
  - From S. McKenna, S. Gong and Y. Raja. Face Recognition in Dynamic Scenes. Proceedings of British Machine Vision Conference, 1997

Closed Set vs. Open Set Identification

- **Closed-Set Identification:**
  - The system reports which person from the gallery is shown on the test image: Who is he?
  - Performance metric: Correct identification rate

- **Open-Set Identification:**
  - The system first decides whether the person on the test image is a known or unknown person. If he is a known person who he is?

    1. False accept: The invalid identity is accepted as one of the individuals in the database.

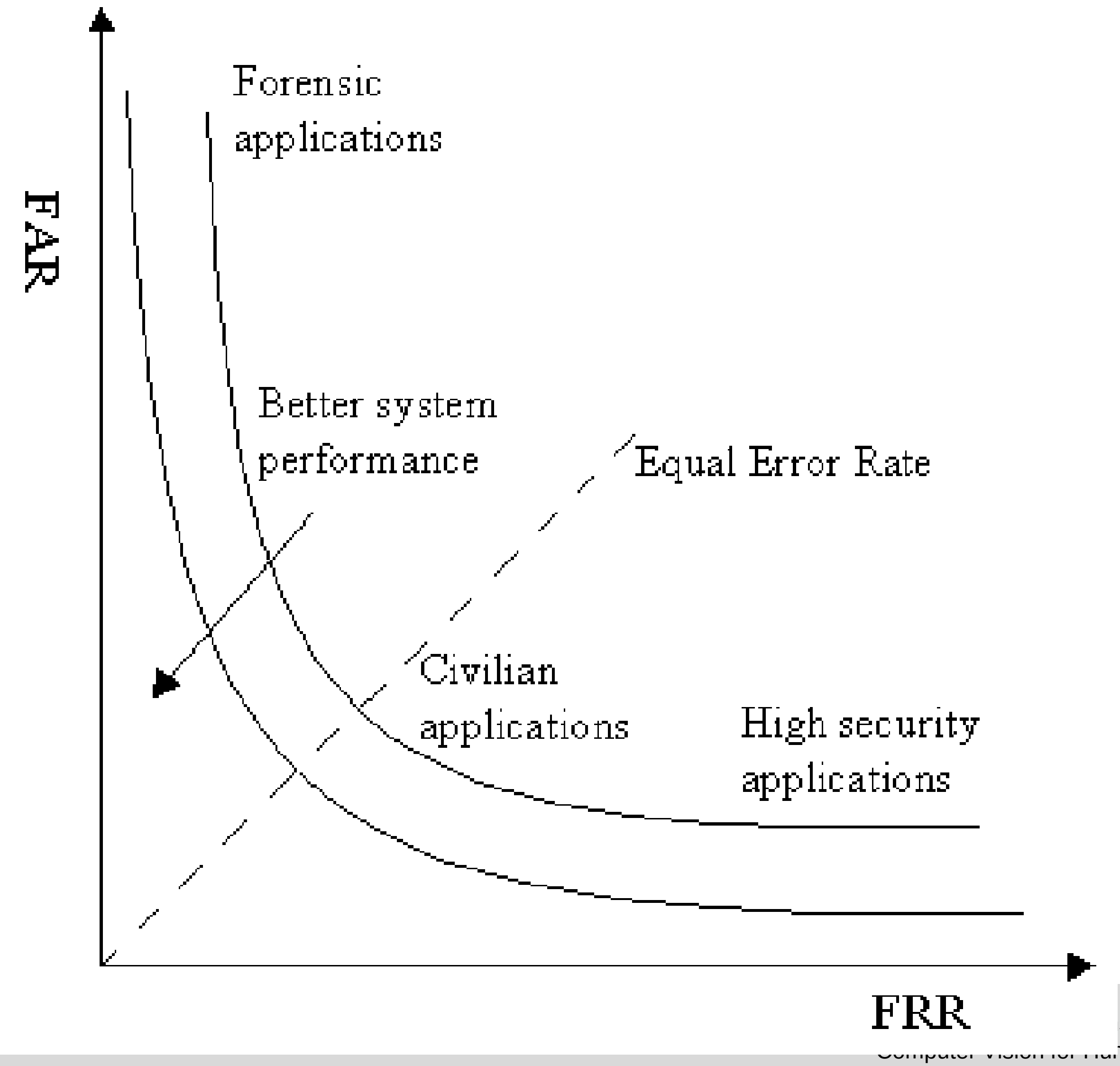
    2. False reject: An individual is rejected even though he/she is present in the database.

    3. False classify: An individual in the database is correctly accepted but misclassified as one of the other individuals in the training data

Authentication/Verification

- A person claims to be a particular member. The system decides if the test image and the training image is the same person: Is he who he claims he is?
- Performance metric: *false reject rate (FRR), false accept rate (FAR)*
- **False reject:** system rejects a valid identity;
- **False accept:** system incorrectly accepts an invalid identity.

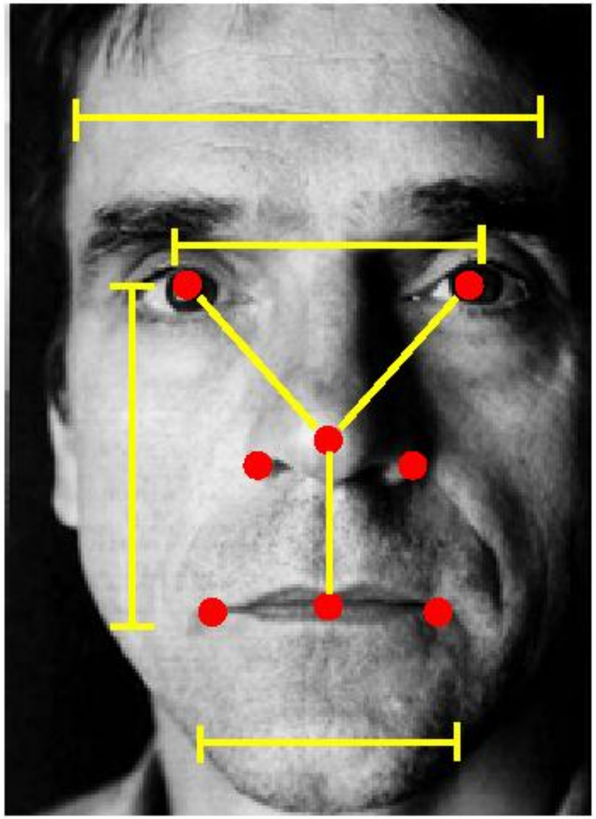
Receiver Operating Characteristics (ROC) Curve

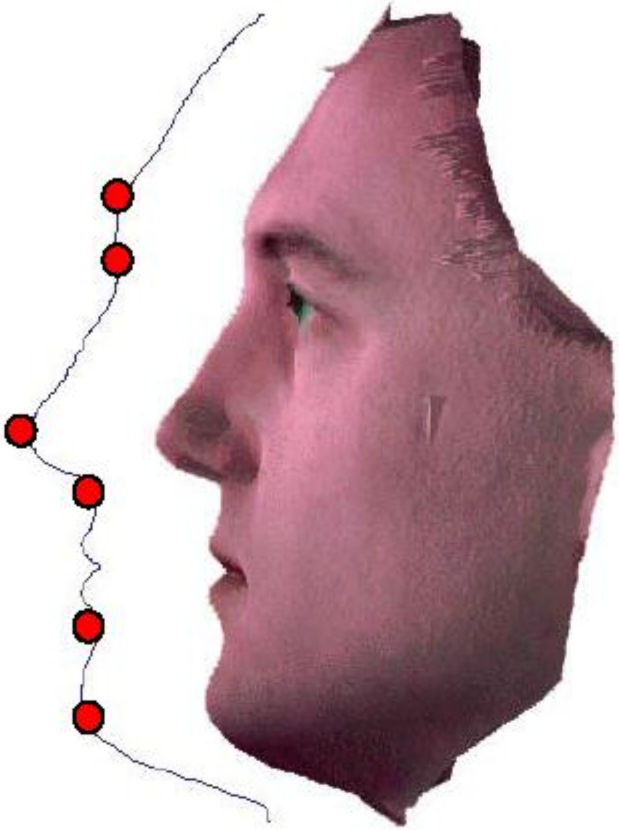


Traditional Approaches

- Feature-based
  - fiducial points
  - distances, angles, areas, etc.
  - Geometrical
- Appearance-based
  - holistic, fiducial regions
  - statistical

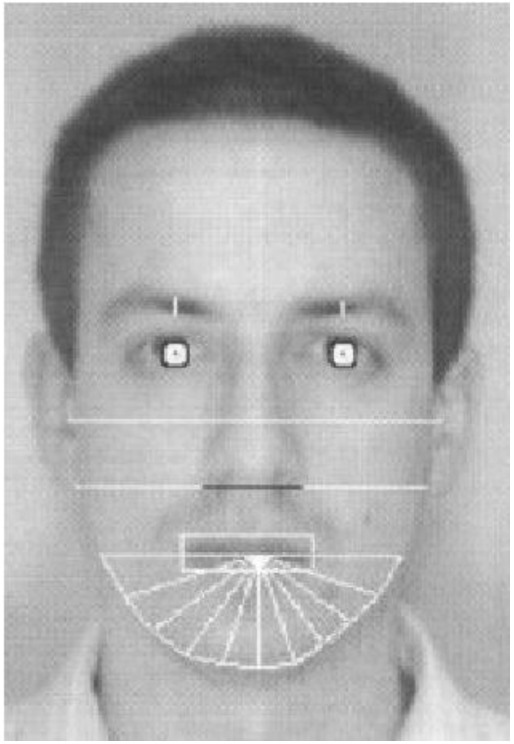
Features: Frontal &Profile





Feature-based Face Recognition

- Eyebrow thickness and vertical position at the eye center position
- A coarse description of the left eyebrow‘s arches
- Nose vertical position and width
- Mouth vertical position, width, height upper and lower lips
- Eleven radii describing the chin shape
- Face width at nose position
- Face width halfway between nose tip and eyes



- R. Brunelli, T. Poggio, “Face Recognition: Features versus Templates”, *IEEE Trans. on PAMI*, Vol. 15, No. 10, pp. 1042-1052, Oct. 1993.

Classification

- Nearest neighbor classifier with Mahalanobis distance as the distance metric.

$$\Delta_j(x) = (x - m_j)^T \Sigma^{-1} (x - m_j)$$

x: input face image

m_j: average vector representing the *j*th person.

Σ : Covariance matrix

T: transpose operator

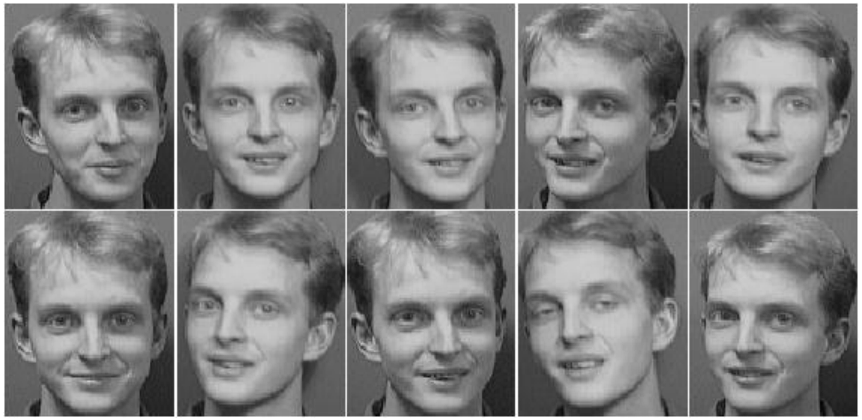
- Different people are characterized only by their average feature vector.
- The distribution is common and estimated by using all the examples in the training set.

21

Appearance Based Approaches

- can be either
  - holistic, i.e. they process the whole face as the input
  - local / fiducial, i.e. they process facial features, such as eyes, mouth, etc. seperately

Whole Faces



Sample sequence from ORL database

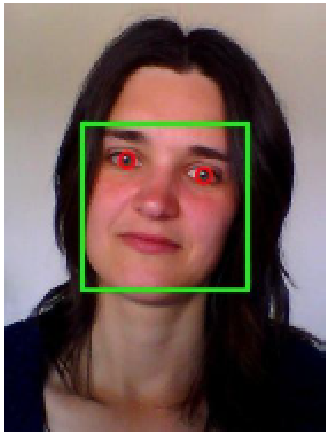
Regions

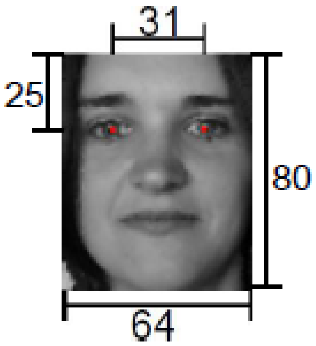
Brunelli & Poggio (1993)

- R. Brunelli, T. Poggio, “Face Recognition: Features versus Templates”, *IEEE Trans. on PAMI*, Vol. 15, No. 10, pp. 1042-1052, Oct. 1993.

Preprocessing Step

- Align faces with facial landmarks
  - e.g. using manually labeled or automatically detected eye centers
  - Normalize face images to a common coordination, remove translation, rotation and scaling factors
  - Crop off unnecessary background





23

Eigenfaces

- A face image defines a point in the high dimensional image space
- Different face images share a number of similarities with each other

⇨They can be described by a relatively low dimensional subspace

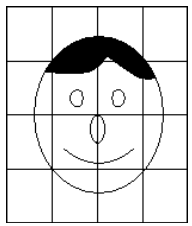
⇨Project the face images into an appropriately chosen subspace and perform classification by similarity computation (distance, angle)

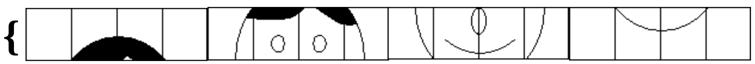
Eigenfaces

- Dimensionality reduction procedure used here is called *Karhunen-Loéve transformation* or *principal component analysis*

- Objective: Find the vectors that best account for the distribution of face images within the entire image space

Principal Component Analysis (PCA)



y = {  **}**

y: face image

$Y = [y_1, y_2, y_3, \ldots, y_K]$ — *Y*: face matrix

$m = (1/K) * \Sigma\, y$ — *m*: Mean face

$C = (Y - m)(Y - m)^T$ — *C*: Covariance matrix

$D = U^T C U$ — *D*: eigenvalues, *U*: eigenvectors

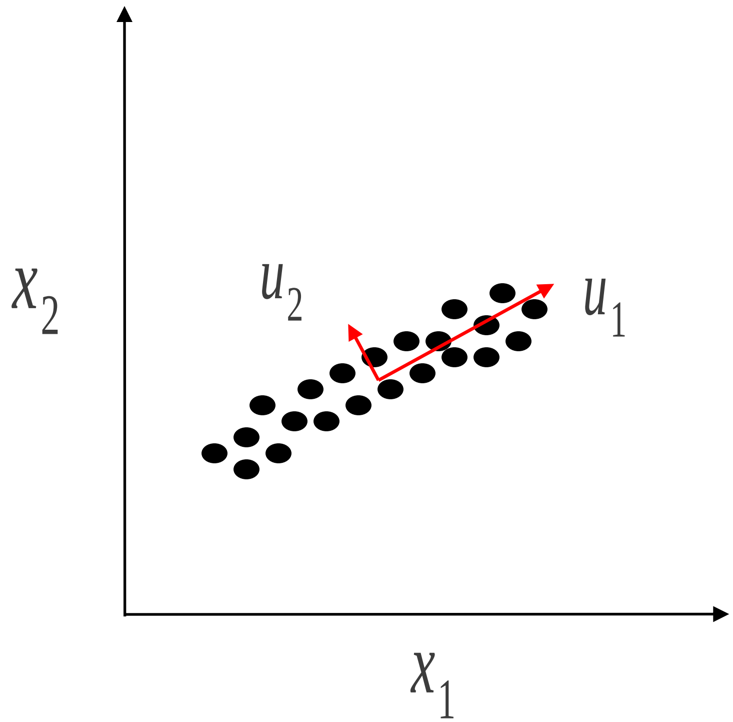
$\Omega = U^T * (y - m)$ — Ω: representation coefficients

Eigenfaces

- Training:
  - Acquire initial set of face images (training set):
    - $Y = [y_1, y_2, y_3, \ldots, y_K]$
  - Calculate the eigenfaces from the training set, keeping only the M images corresponding to the highest eigenvalues
    - $U=(u_1, u_2, .. u_M)$
  - Calculate representation of each known individual *k* in face space
    - $\Omega_k = U^T *(y_k - m)$
- Testing:
  - Project input new image *y* into face space: $\Omega = U^T *(y - m)$
  - Find most likely candidate class *k* by distance computation
    - $\varepsilon_k = || \Omega - \Omega_k ||$, for all Ω_k

Eigenfaces / PCA

Karlsruhe Institute of Technology

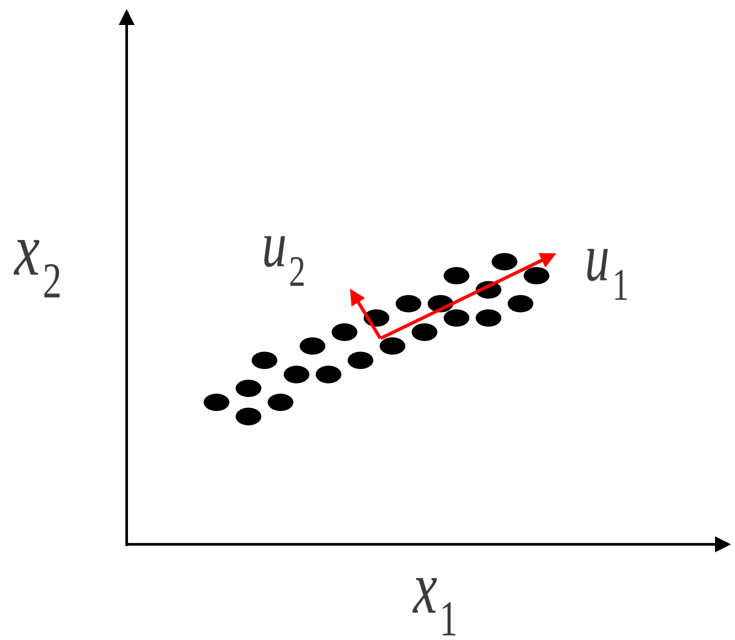


Principal components are the eigenvectors of the covariance matrix of the set of face images

(PCA: see e.g. Duda & Hart, 1973: Pattern Classification, Scene Analyisis)

28

Principal Component Analysis (PCA)



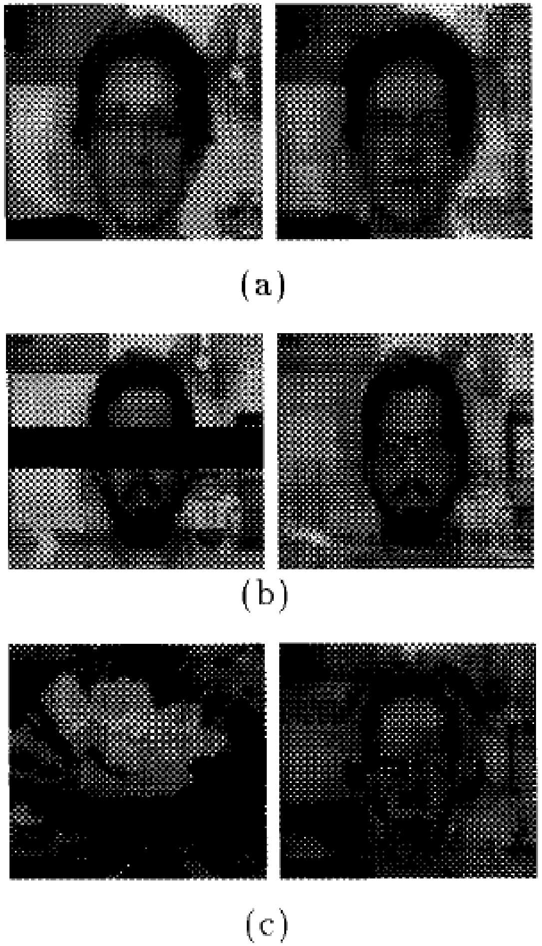
29

Eigenfaces

Principal components are called “eigenfaces” and they span the “face space”

Projections onto the face space

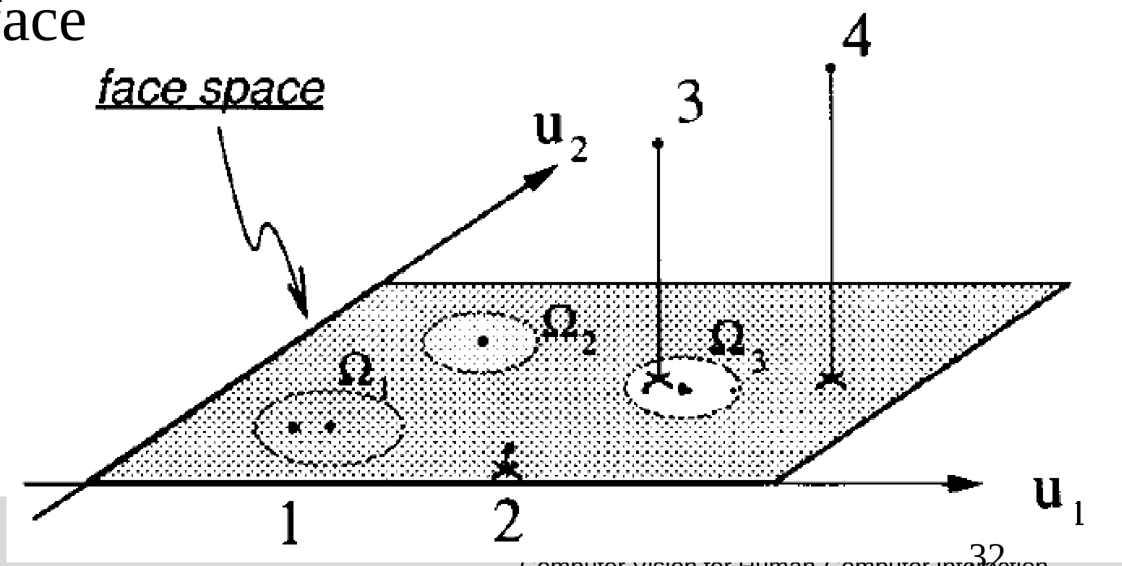
- Images can be reconstructed by their projections in face space
  - $Y_f = \Sigma_{i=1}^{M} \omega_i u_i$
- Appearance of faces in face-space does not change a lot
- Difference of mean-adjusted image (Y-m) and projection Y_f gives a measure of *„faceness“*
  - → *distance from face space (dffs)*
  - Can be used to detect faces



Images (left) and their projections in face space (right)

Projections into face space

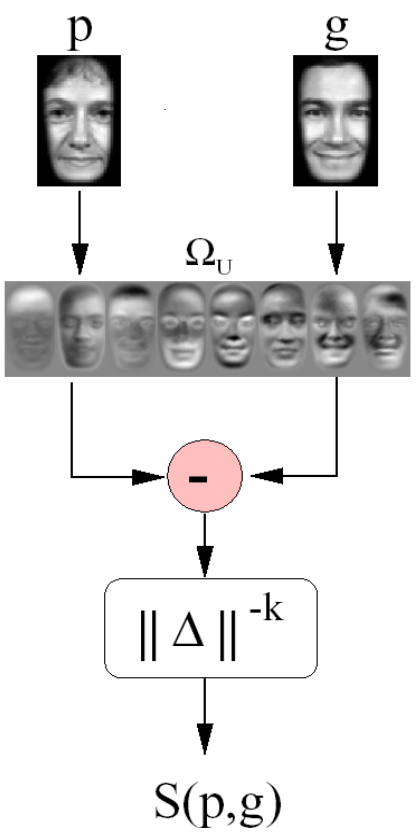
- Case 1: projection of a *known* individual
  - Near face space ($\varepsilon < \theta_\delta$) and near known face Ω_k ($\varepsilon_k < \theta_\varepsilon$)
- Case 2: projection of an *unkown* individual
  - Near face space, far from reference vectors
- Case 3 and 4: not a face
  - Far from face space



Principal Component Analysis (PCA)

Projects all faces
onto a universal
eigenspace to “encode”
via principal components

Uses inverse-distance
as a similarity measure *S(p,g)*
for matching & recognition



- M. Turk and A. Pentland, “Eigenfaces for Recognition”, *Journal of Cognitive Science*, pp. 71-86, 1991.

View-based Eigenspaces

- Extension: View-based eigenspaces for general viewing conditions
- Given: N individuals under M different views
  - Build universal eigenspace from the combination of NM images: “parametric eigenspace”
  - Build “view-based” set of M separate eigenspaces
- Experiments show slight advantage of view-based approach

View-based Eigenspaces

- Build an eigenspace for each view
- Decide input images' direction of view using distance from view space metric
- Do classification in that view-space

- A. Pentland, B. Moghaddam, T. Starner and M. Turk, "View based and Modular Eigenspaces for Face Recognition", *CVPR'94*, pp. 84-91.

Bayesian Face Recognition

- Problem: Simple nearest-neighbour similarity measures do not exploit knowledge of critical appearance varitations

Bayesian similarity measure:

- denotes belief that image differences are caused by typical appearance variations of an individual (caused by expression, etc.)
- compares typical within-class (intrapersonal) variations with between-class (extrapersonal) variations

- B. Moghaddam, T. Jebara, A. Pentland, "Bayesian Face Recognition", *Pattern Recognition*, Vol. 33, No. 11, pp. 1771-1782, Nov. 2000.

Dual PCA (Bayesian)

Intrapersonal Ω_I

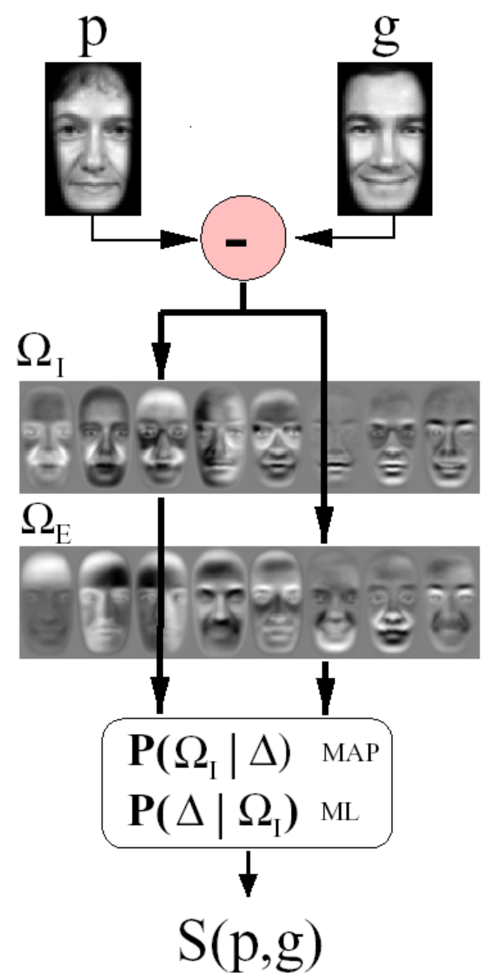
Extrapersonal Ω_E

$$\Omega_I \equiv \{\Delta = x_i - x_j : L(x_i) = L(x_j)\}$$
$$\Omega_E \equiv \{\Delta = x_i - x_j : L(x_i) \neq L(x_j)\}$$

$$S = P(\Omega_I | \Delta) = \frac{P(\Delta|\Omega_I)P(\Omega_I)}{P(\Delta|\Omega_I)P(\Omega_I) + P(\Delta|\Omega_E)P(\Omega_E)}$$

$P(\Delta|\Omega)$ derived from training data

(some tricks are needed …)



S(p,g) defines a probabilistic similarity metric

Face ID: Eigenfaces

- Problems and shortcomings:
  - Eigenfaces do not distinguish between shape and appearance:

    Active Shape Models (ASM)

    Active Appearance Models (AAM)

  - PCA does not use class information:

    PCA projections are optimal for reconstruction from a low dimensional basis, they may not be optimal from a discrimination standpoint:

    "Much of the variation from one image to the next is due to illumination changes." [Moses, Adini, Ullman]

38

Linear Discriminant Analysis (LDA) –Fisherfaces-

- Fischer‘s Linear Discriminant
  - Preserves separability of classes
  - Maximizes ratio of projected between-classes to projected within-class scatter

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

- Between-class scatter

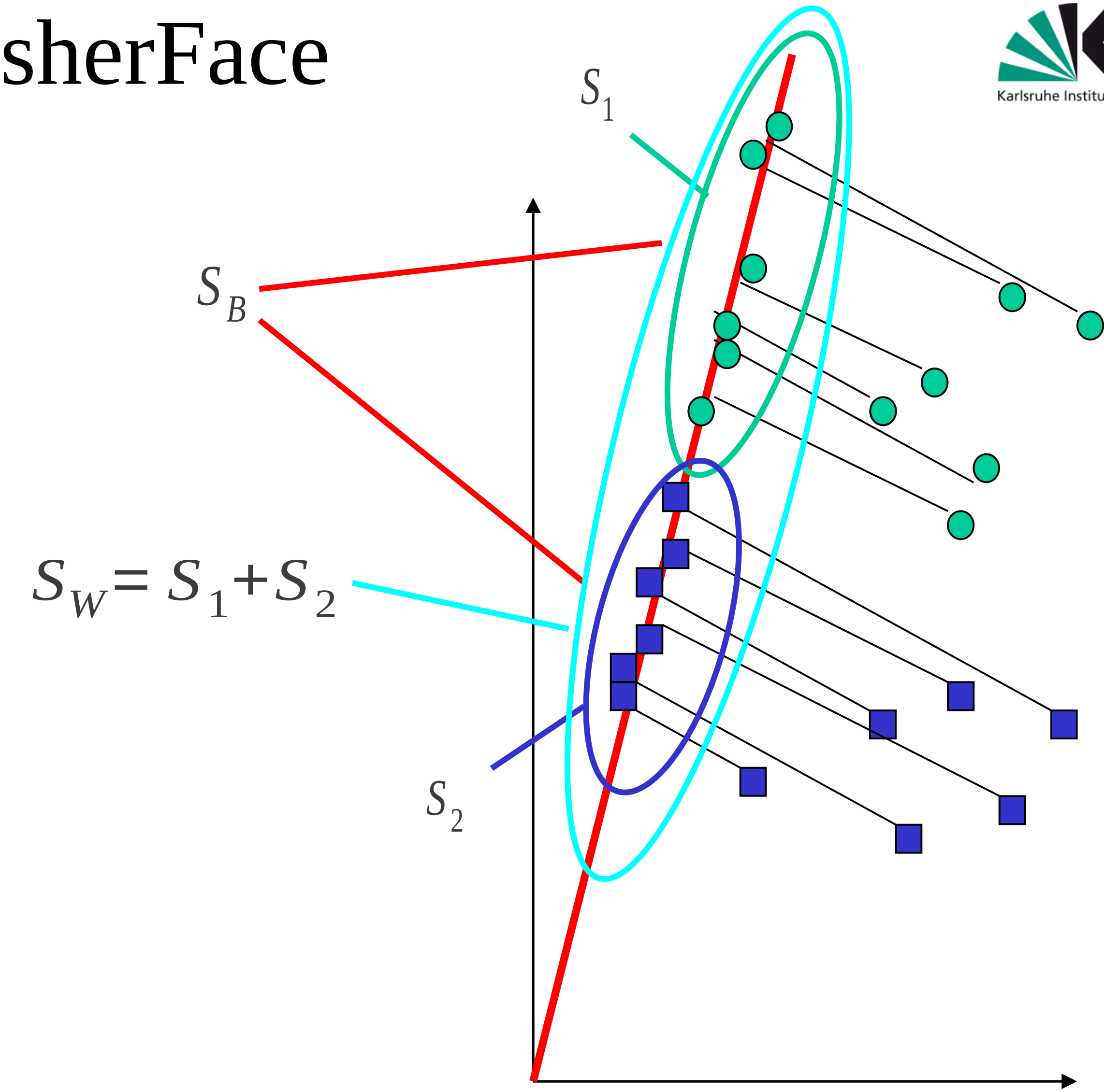
$$S_B = \sum_{i=1}^{c} |x_i| (\mu_i - \mu)(\mu_i - \mu)^T$$

- Within-class scatter

$$S_W = \sum_{i=1}^{c} \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T$$

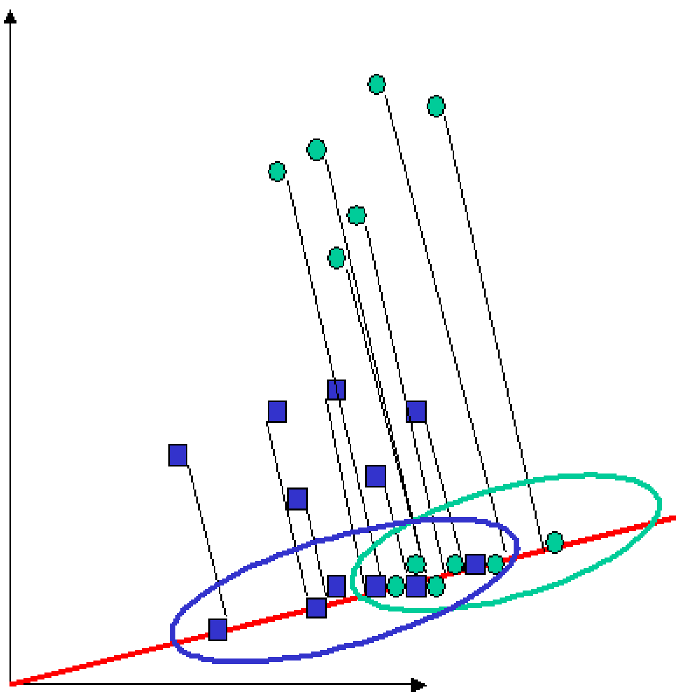
- Where
  - c is the number of classes
  - μ_i is the mean of class X_i
  - $|X_i|$ is number of samples of X_i

FisherFace

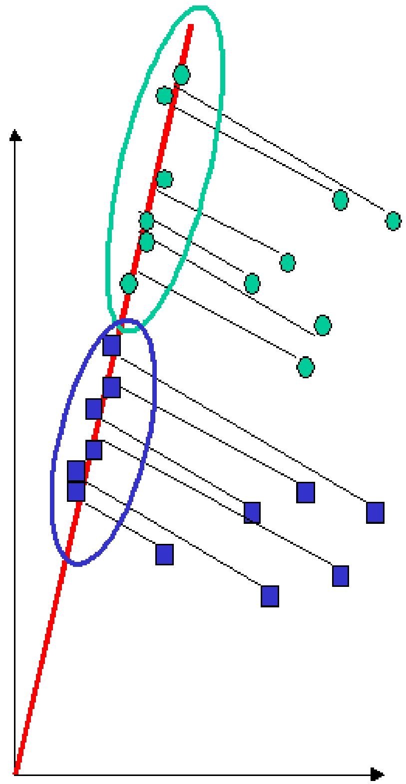


Good separation

Fisherface



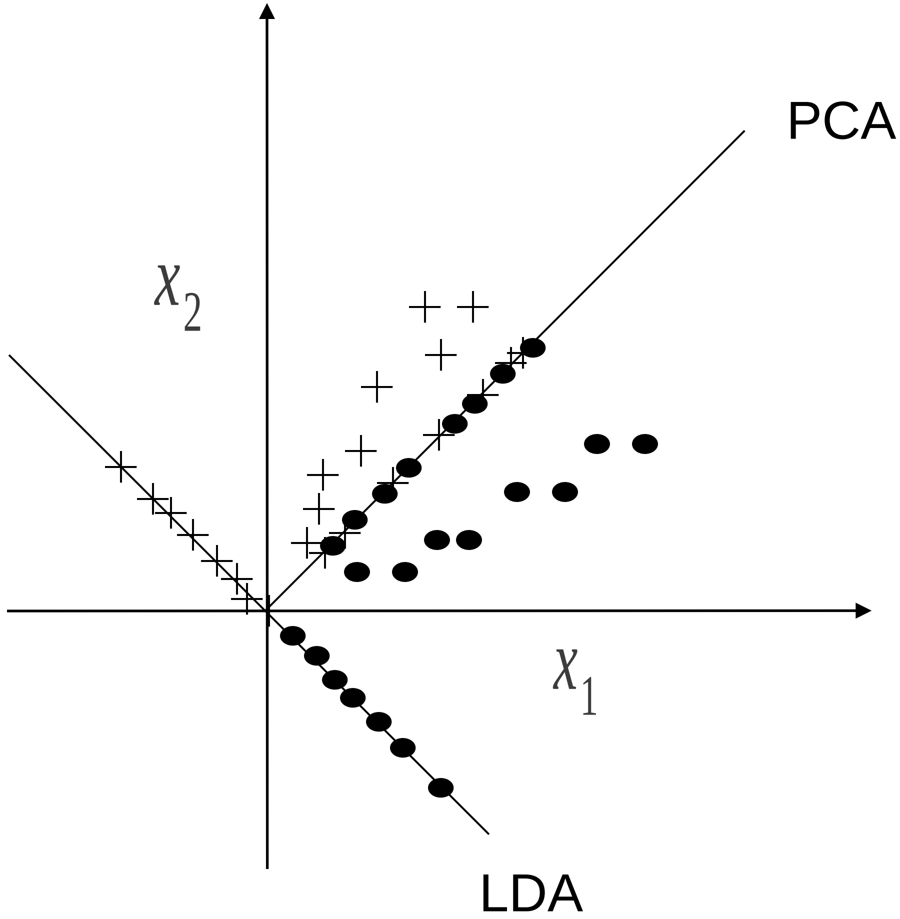
Poor Projection



Good Projection

41

PCA vs. LDA



42

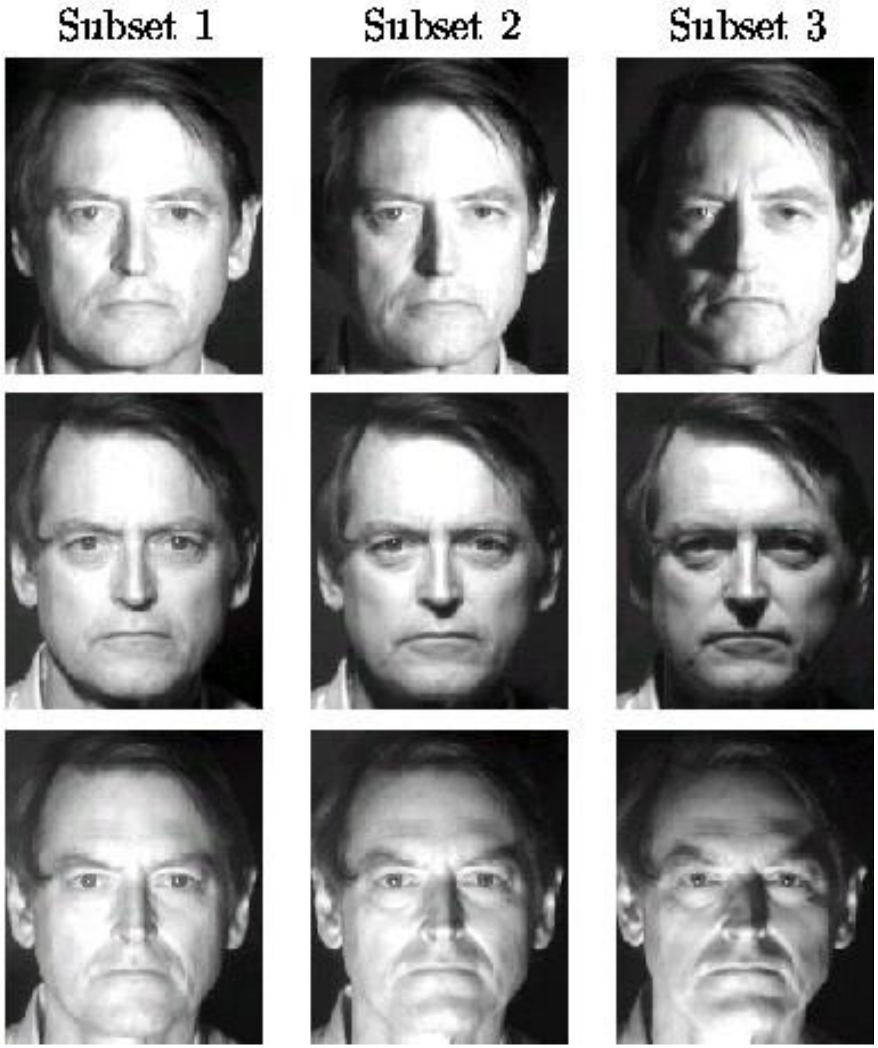
Fisherfaces

- **Fisher's Linear Discriminant projects away the within-class variation (lighting, expressions) found in training set.**

- **Fisher's Linear Discriminant preserves the separability of the classes.**

- P.N. Belhumeur et al., "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", *IEEE Trans. on PAMI*, No. 7, pp. 711-720, 1997.

Fisherfaces: Experiments



Subsets of Harvard Database:

Subset 1: angle of light source is within 15 degr. of camera axis

Subset 2: angle of light source is within 30 degr. of camera axis

Subset 3: angle of light source is within 45 degr. of camera axis

Face ID: Fisherfaces

Method	Subset 1	Subset 2	Subset 3
Eigenface	0.0	4.4	41.5
Eigenface w/o 1^{st} 3	0.0	4.4	27.7
Fisherface	0.0	0.0	4.6

Methods were trained on images from Subset 1

(Belhumeur et al., “Eigenfaces vs. Fisherfaces”, TPAMI, 1997)

Local Appearance based Face Recognition

- Modular Eigen Spaces
- Local PCA
- Local Feature based (e.g. LBP, DCT etc)

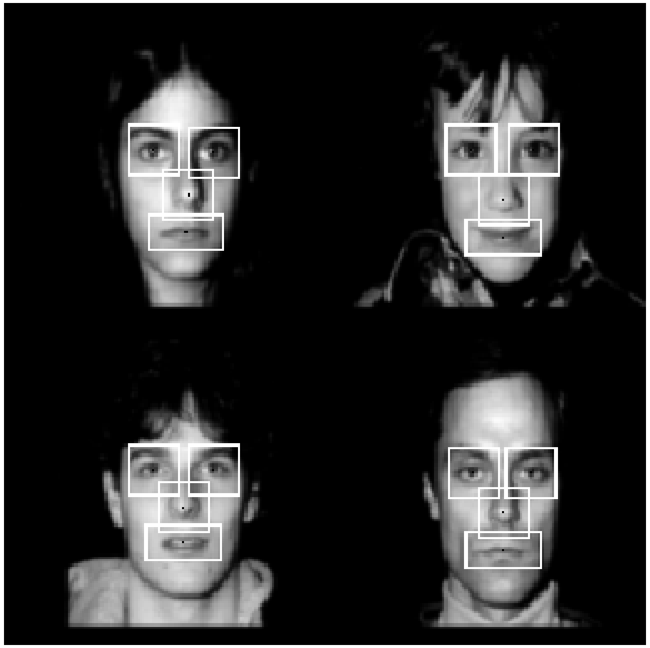
46

Local vs. Holistic Approaches

- Local variations on the facial appearance, i.e. due to different expression, occlusion and lighting, lead to modifications on the entire representation in the holistic approaches, while in local approaches only the corresponding local region is effected.
- Face images contain different statistical illumination –high frequency at the edges, i.e. eyebrows, low frequence at smooth regions, i.e. cheeks. Easier to represent the varying statistics linearly by using local representation.
- Local approaches facilitate the weighting of each local region in terms of their effect on face recognition.

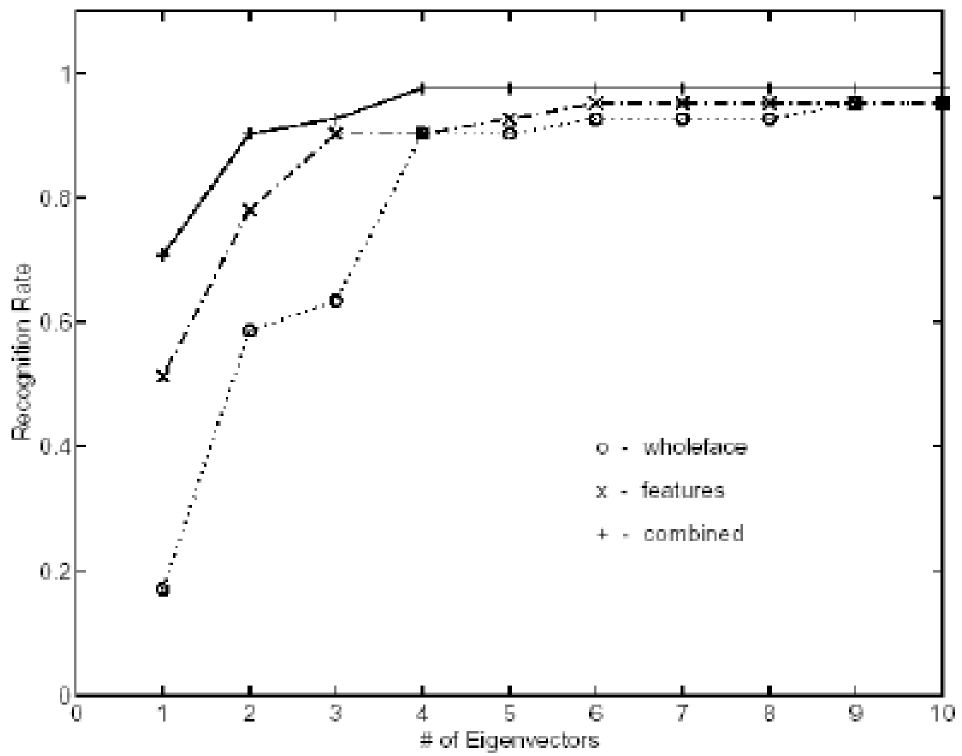
Modular Eigenspaces

Does classification using fiducial regions (eyes, nose, -mouth is excluded in this study-) instead of using entire face.



- A. Pentland, B. Moghaddam, T. Starner and M. Turk, “View based and Modular Eigenspaces for Face Recognition”, *CVPR‘94*, pp. 84-91.

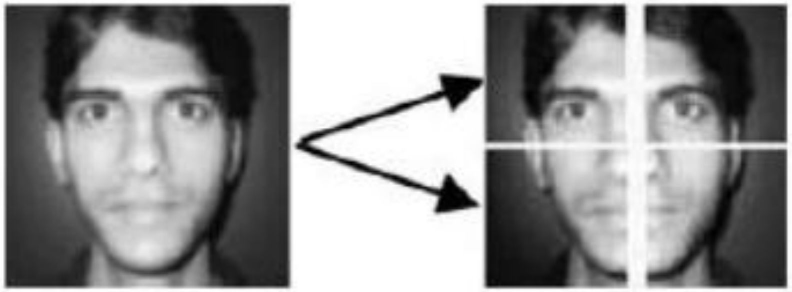
Modular Eigenspaces



A. Pentland et al., “View-based and modular eigenspaces for face recognition“, In *CVPR’94*, 1994.

Local Principal Component Analysis (Modular PCA)

- The face images are divided into N smaller sub-images.
- PCA is applied on each of these sub-images.



- Performed better than global PCA on large variations of illumination and expression
- No improvements under variation of pose

(Gottumukal & Asari, 2003)

Local feature based Face Recognition

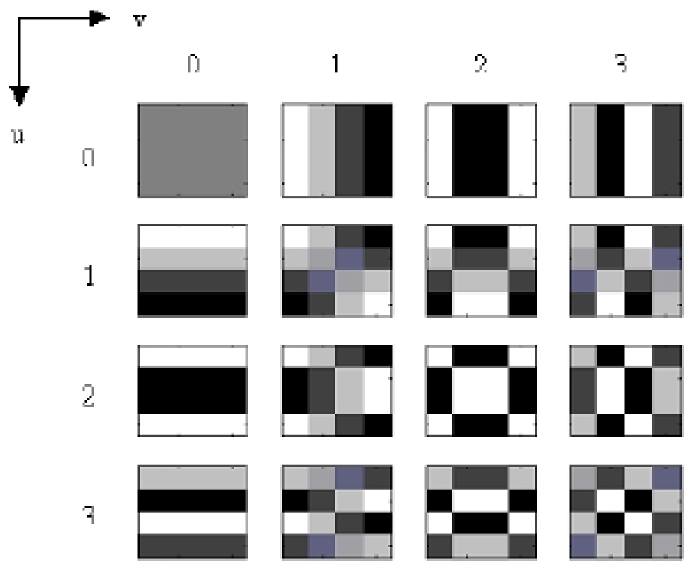
- **Objective:**

  To mitigate the effects of expression, illumination and occlusion variations by performing local analysis and by fusing the outputs of extracted local features at the feature or at the decision level.

Two popular facial descriptions achieving good results
1.Local binary Pattern Histogram LBP, Ahonen et el, 2004
2.Discrete Cosine Transform based Approaches, Ekenel et al 2005.

Discrete Cosine Transform (DCT)

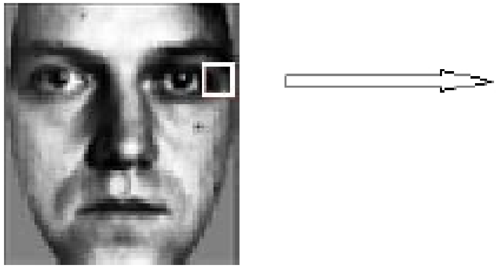
- Image block is represented as weighted sum of cosines at different frequencies
  - -> data independent bases (in contrast to PCA)
  - fast implementation available (O(n log n))



52

Discrete Cosine Transform (DCT)

- Compact representation



415	60	59	26	61	48	10	11
146	9	69	54	8	5	4	5
77	70	28	4	13	1	3	1
33	11	19	9	8	0	2	3
14	21	11	5	4	6	4	2
1	10	0	15	8	1	4	2
1	3	1	4	2	5	4	2
0	2	1	0	2	1	1	1

DCT vs. Principal Component Analysis (PCA)

- In PCA, a data-specific space is built which requires fine alignment to obtain proper bases.
- In DCT, one has data independent bases.

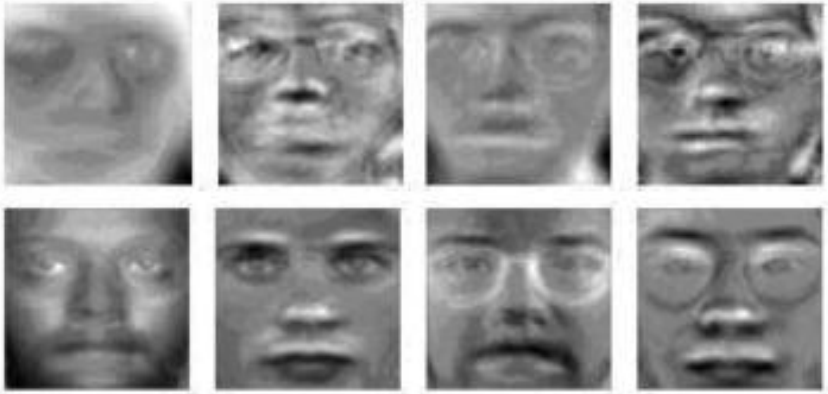


Figure: Eigenfaces bases computed from mis-aligned (top) and well-aligned (bottom) images

DCT

- Divide the input image to blocks of 8x8 pixels size.
- Perform DCT on each block.
- Order the remaining DCT coefficients using zig-zag scan.
- Remove the first DCT coefficients and from the remaining ones select the first M of them.

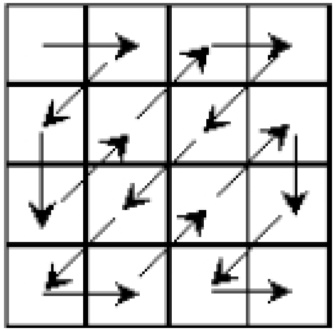


Figure: Zig-zag scan pattern

Local DCT-based Face Reco - Results

(Ekenel & Stiefelhagen, 2005)

Method	Reco. Rate
PCA (20)	75.6%
LDA (14)	80.0%
ICA 1 (40)	77.8%
ICA 2 (40)	72.2%
Global DCT (64)	74.4%
Local DCT (18) + GMM (8) as in [12]	58.9%
Local DCT + Feature Fusion (192)	**86.7%**
Local DCT (10) + Decision Fusion (64)	**98.9%**

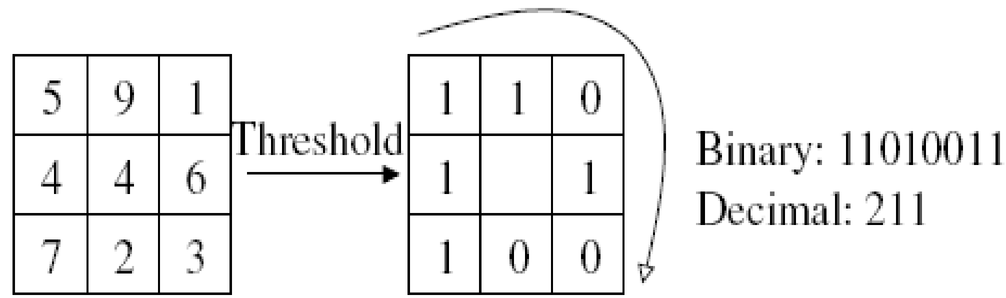
Results on Yale-DB
(15 individuals)

Method	Reco. Rate
PCA (80)	57.1%
LDA (67)	59.5%
ICA 1 (200)	59.1%
ICA 2 (200)	51.8%
Global DCT (256)	44.1%
Local DCT (18) + GMM (8) as in [12]	12.0%
Local DCT + Feature Fusion (640)	**70.9%**
Local DCT (10) + Decision Fusion (64)	**68.5%**

Results on CMUPIE-DB
(68 individuals)

Local Binary Pattern Histogram, Ahonen et al. 2004

$$LBP_{(P,R)}^{U2}$$



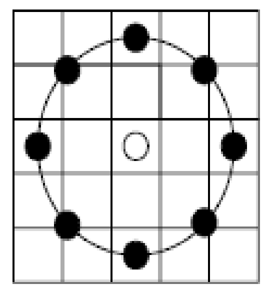


Fig.. (a) the basic LBP operator. (b) The circular (8,2) neighbourhood. The pixel values are bilinearly interpolated whenever the sampling point is not in the centre of a pixel (Ahonen et al, 2004)

- Use any distance measure/Similarity Metric for matching e.g.

- Histogram Intersection

$$I(S,M)=\sum_i \min(S_i, M_i)$$

- Chi Square Statistics χ^2

$$\chi^2(S,M)=\sum_i \frac{(S_i-M_i)^2}{S_i+M_i}$$

Some results

Method	fb	fc	dup I	dup II	lower	mean	upper
LBP, weighted	0.97	0.79	0.66	0.64	0.76	0.81	0.85
LBP, nonweighted	0.93	0.51	0.61	0.50	0.71	0.76	0.81
PCA, MahCosine	0.85	0.65	0.44	0.22	0.66	0.72	0.78
Bayesian, MAP	0.82	0.37	0.52	0.32	0.67	0.72	0.78
EBGM _Optimal	0.90	0.42	0.46	0.24	0.61	0.66	0.71

For a detail and in-depth discussion on different face feature extraction techniques and how they are commonly used, see the relevant book chapter

Sarfraz et al, *Feature Extraction and Representation for face recognition*, Face recognition, INTECH publishers, Austria, 2010, ISBN 978-953-307-060-5

59

State of the art Algorithm's Performance

Face Recognition Grand Challenge (FRGC) database,
Controlled environment, 120 subjects,
Expression variations, time gap

98.5% (98.8%)

FRGC face database,
Uncontrolled environment,
120 subjects,
Expression variations, time gap

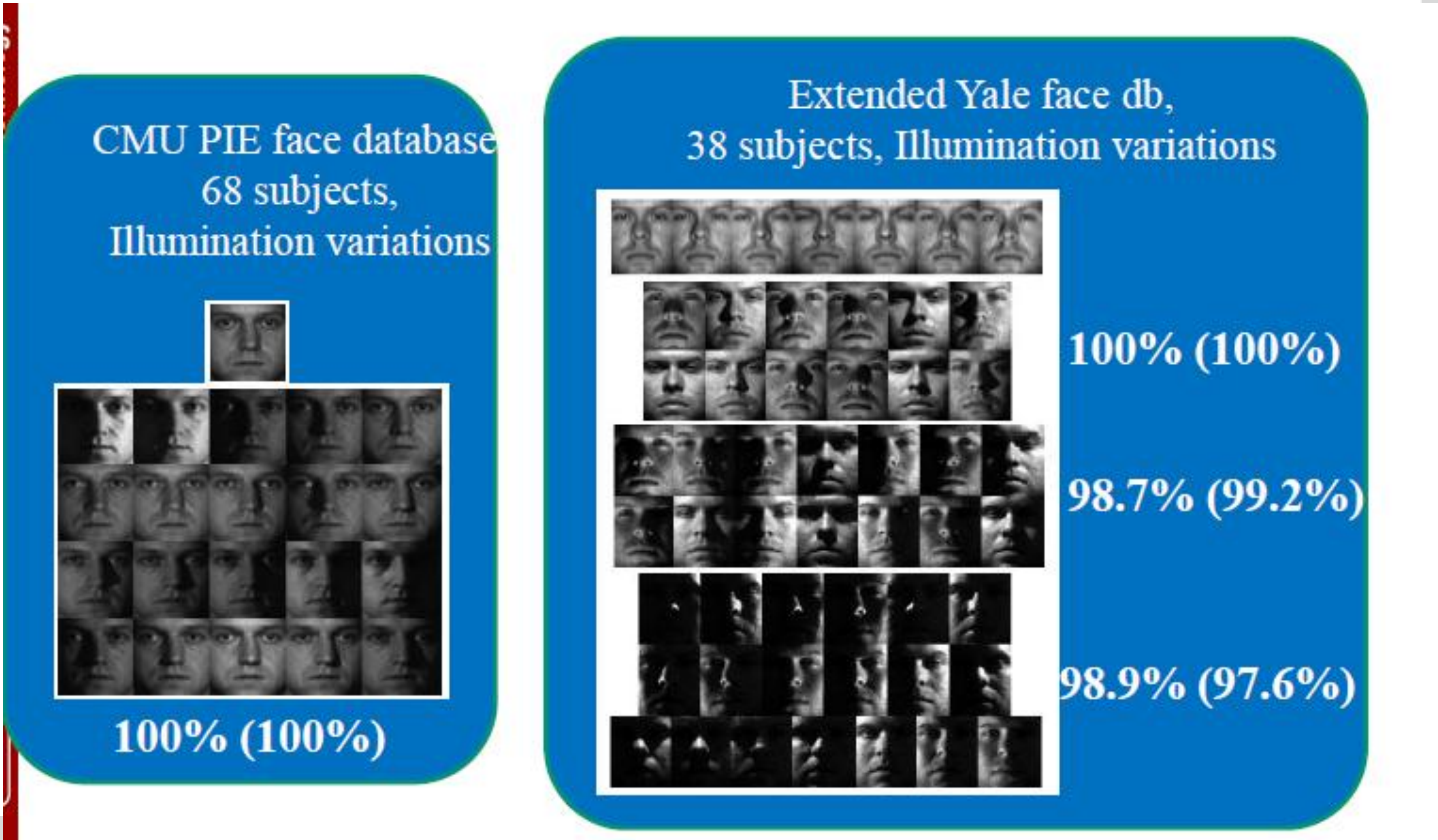
96.2% (80.6%)

AR face database,
110 subjects,
Occlusion

98.2% (97.5%) 97.3% (93.5%)

Results in (): best reported performance by other *task-specific* algorithms

Results- Illumination variation



Relevant Papers

- M. Turk and A. Pentland, “Face Recognition using Eigenfaces”, CVPR 1991.
- M. Turk and A. Pentland, “Eigenfaces for Recognition”, *Journal of Cognitive Science*, pp. 71-86, 1991.
- P.N. Belhumeur, J.P. Hespanha and D.J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection”, *IEEE Trans. on PAMI*, Vol. 19, No. 7, pp. 711-720, 1997.
- W. Zhao, R. Chellappa, P.J.J. Phillips, and A. Rosenfeld, “Face Recognition: A Literature Survey”, *ACM Computing Survey*, Vol. 35, No 4, 399-458, 2003.
- Ahonen, T., Hadid, A. & Pietikainen, M. (2004). Face recognition with local binary patterns, *ECCV*, pp. 469–481.

62

Further Reading

- R. Brunelli, T. Poggio, “Face Recognition: Features versus Templates”, IEEE Trans. on PAMI, Vol. 15, No. 10, pp. 1042-1052, Oct. 1993
- Y. Moses, Y. Adini and S. Ullman, “Face Recognition: The Problem of Compensating for Changes in Illumination Direction”, Proceedings of European Conference on Computer Vision, pp. 286-296, 1994
- A. Pentland, B. Moghaddam, T. Starner and M. Turk, “View based and Modular Eigenspaces for Face Recognition”, *CVPR ‘94*, pp. 84-91.
- B. Moghaddam, T. Jebara, A. Pentland, "Bayesian Face Recognition", *Pattern Recognition*, Vol. 33, No. 11, pp. 1771-1782, Nov. 2000.
- B. Moghaddam, “Principal Manifolds and Probabilistic Subspaces for Visual Recognition”, *IEEE Trans. on PAMI*, Vol. 24, No. 6, pp. 780-788, 2002.
- Sarfraz et al, *Feature Extraction and Representation for face recognition*, Face recognition, INTECH publishers, Austria, 2010, ISBN 978-953-307-060-5