Face Biometric: Algorithms, Performance & Applications

Stan Z. Li

Center for Biometrics and Security Research (CBSR) &
National Lab of Pattern Recognition (NLPR)
Institute of Automation, Chinese Academy of Sciences

ASI-07, Hong Kong, 12 Jan, 2007

Outline

- Introduction
- Subspace Analysis
  - Linear Methods
  - Nonlinear Methods
  - Face Grand Challenges from Subspace Viewpoint
- Face Analysis Methods
  - Face Detection
  - Face Alignment
  - Face Recognition
- Face Recognition Using Near Infrared Images
- Applications

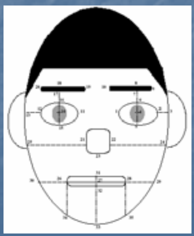
Introduction

Face Recognition Process

1. Face Detection
2. Face Tracking
3. Face Alignment
4. Face Recognition

History (60-70's):
Geometric Feature Based Approach

- Eg. Kanade 1972
- In traditional AI-CV framework
- Image features pre-specified
- Features={type, locations, distances}



Feature	Distance
1	0.5 * ((1,2) + (11,12))
2	0.5 * ((5,6) + (15,16))
3	(3,13)
4	(24,25)
5	(29,30)
6	(34,35)
7	(26,34)
8	(28,35)
9	(26,28)
10	(27,31)
11	(27,32)
12	(32,33)
13	(23,31)
14	(21,22)
15	0.5 * ((13,25) + (3,24))
16	0.5 * ((25,30) + (24,29))
17	0.5 * ((30,34) + (29,35))
18	0.5 * ((1,22) + (11,21))
19	(10,19)
20	0.5 * ((2,9) + (12,20))
21	0.5 * ((9,10) + (19,20))
22	0.5 * ((11,19) + (1,10))
23	0.5 * ((6,7) + (16,17))
24	0.5 * ((7,8) + (17,18))
25	0.5 * ((18,19) + (8,10))
26	0.5 * ((18,20) + (8,9))
27	(11,23)
28	(1,23)
29	0.5 * ((1,28) + (11,26))
30	0.5 * ((12,13)+(2,3))

Table 1: The 30-dimensional feature vector.

History (90's -):
Learning-Based, Subspace Analysis Approach

- Linear Subspace Methods: Eigenface (PCA) and Others
  - Face Representation: Kirby & Sirovich. 1990.
  - Face Recognition: Turk & Pentland. 1991.
- Different from the AI-CV approach
  - Example-based
  - Features Learned
  - Dimension reduction
  - Linear mapping from high-dim to low-dim spaces
- Nonlinear Methods
  - (More contemporary work).

1990s -2004

- Hardware:
  - Visible light imaging
- Algorithms:
  - Appearance Based + Statistical Learning
  - Local Features + Little Learning
- 2002: EyeCU System at MSR TechFest

2005: AuthenMetric System

- Illumination Invariant
- Accurate and Fast
- Real Applications
  - Passport Control at China-Hong Kong/Macau Boarders
  - Access-control in Many Places

Algorithms

Subspace Modeling
Dimension Reduction
Feature Extraction

- Eg: Images of size 64x64
- Dimensionality of image space: 64x64=4096 (pixels)
- Pixel values in {0,…,255}
- 256^4096 > 10^9864 possible configurations in 4096-dim hypercube
- Face pattern living in low dim subspace
- Dimension reduction (features = projected coordinates)

PCA, VQ, NMF, and LNMF

$$X \approx BH$$

Method	Constraints
PCA	b orthonormal vectors
VQ	h unary vectors
ICA	h independent
NMF	b,h non-negative vectors
LNMF	b,h non-negative + h sparse → b ***really*** part-based

PCA Representation

- Basis vectors = Principal eigenfaces

1st 2nd 3rd 4th 5th 6th 7th

- Face as linear combination of eigenfaces

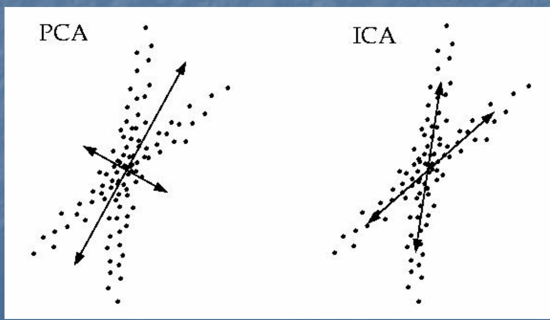
Original Face 1st 2nd 3rd 4th

Y=[0.9569, - 0.1945 , 0.0461, 0.0573, …,]

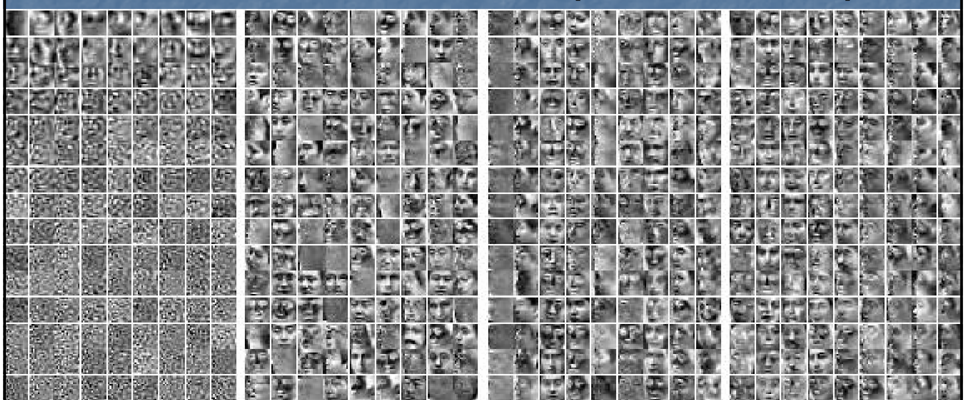
Independent Component Analysis

$$X \approx BH, \quad X = (x_1, \dots, x_N),\, H = (h_1, \dots, h_N)$$

H components as independent as possible



Learning View-Subspaces by Using PCA, ICA, ISA, TICA (Li et al 2001)



	PCA	ICA	ISA	TICA
view-specific	n	Y	Y	Y
View-grouping	n	n	Y	Y
View-ordering	n	n	n	Y

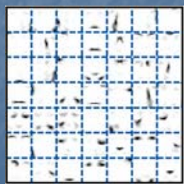
Non-negative Matrix Factorization

- Papers:
  - Lee and Seung, *Nature* , 1999
  - Lee and Seung, *NIPS*, 2001.
- Non-negative Matrix Factorization $X \approx BH$

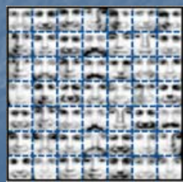
  min D(X||BH), s.t. <u>B,H >=0</u> and $\sum_i b_{ij} = 1$ for all j

Basis Components learned by different methods

Training Example



NMF



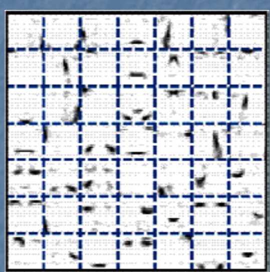
VQ

PCA

Problems with NMF

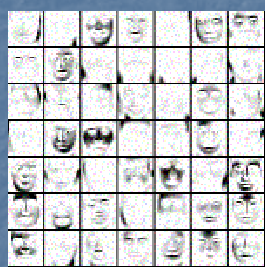
NMF Results Learned From:

Lee-Seung's Data



ORL Data

Our Data

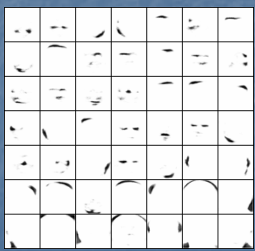


1. Learned components not really localized, part-based
2. Face recognition not very good

Local Non-negative Matrix Factorization

- Additional constraints imposed on NMF
  for spatially localized, part-based representation

Comparative results learned from ORL data:



LNMF

NMF

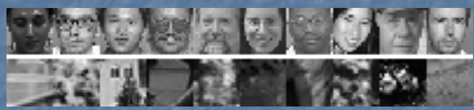
PCA

Nonlinear Subspace Analysis

Face Detection and Recognition - From Manifold Viewpoint

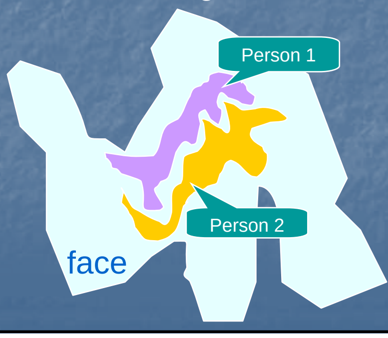
Faces

Nonfaces



Detection

Recognition



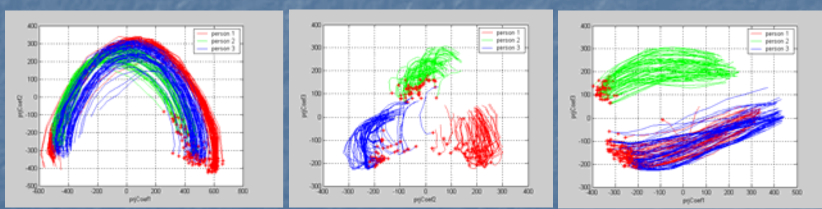
Face Grand Challenges - From Subspace Viewpoint

Challenges in Face Recognition

- Complexity of nonlinear face manifolds
- Problem in Generalizing
  - Limited Training Data
  - When lighting changes
  - When pose changes
  - Daily changes and aging
  - When Camera property change
- Euclidean Geometry Inappropriate in image space

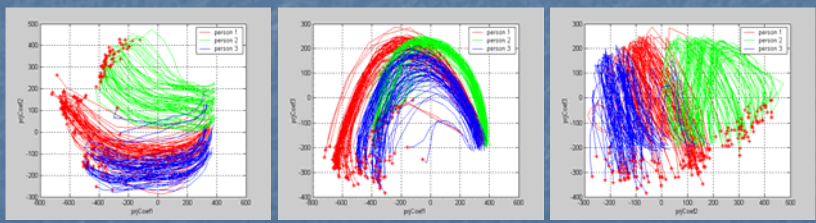
Rotated Faces

in PCA Subspace



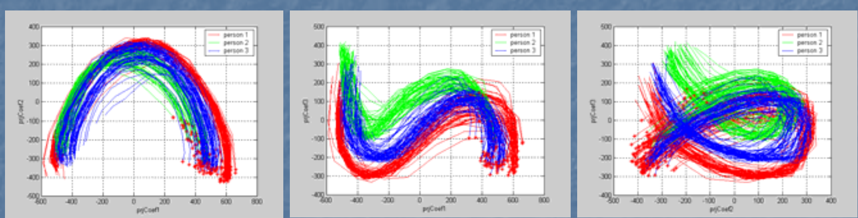
Scaled Faces

in PCA Subspace



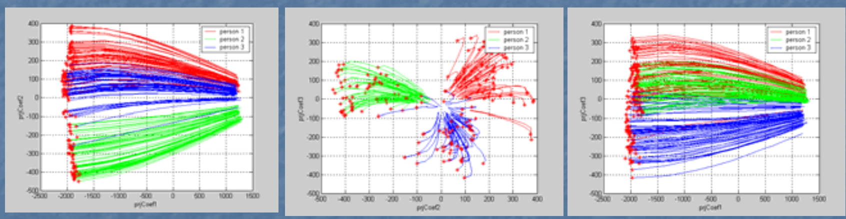
Translated Faces

in PCA Subspace



Manifolds are *Folding* and *Interweaving*

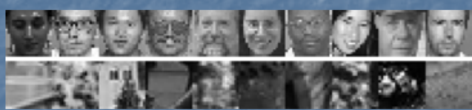
PCA Subspace of “Re-Lighted” Faces



Subspaces in Detection and Recognition

Faces

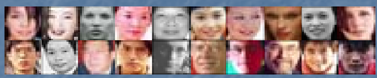
Nonfaces

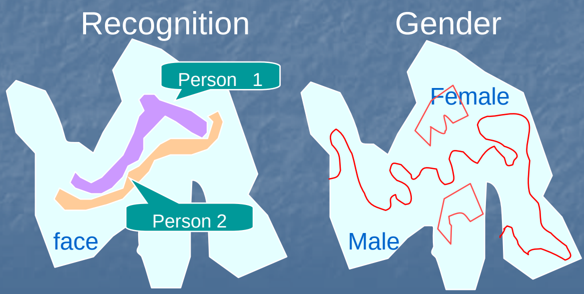


Detection

Recognition

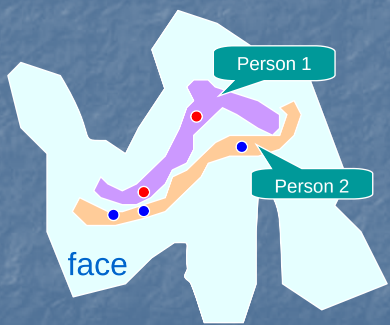
Subspaces in
Face Recognition and Gender Classification





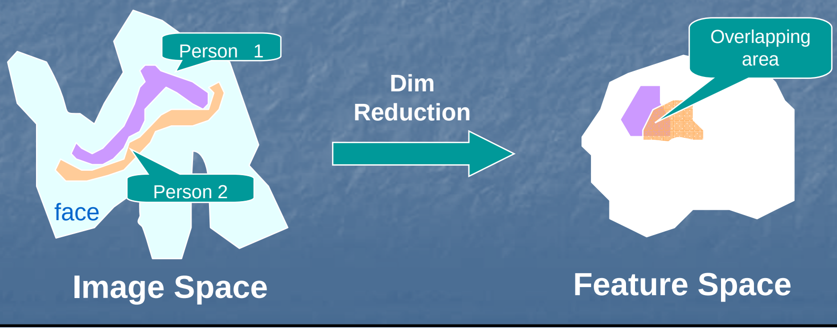
Non-Euclidean Geometry

- Euclidean Geometry Inappropriate
- Need to model manifolds in Non-Euclidean Space
- Geodesic distance



Separability in Image and Feature Spaces

- Individual faces Separable in <u>image space</u>
  - Complex, but separable
- Difficult to separate in <u>feature space</u>
  - Overlapping in feature space due to information loss



Towards Accurate Face Recognition

- Rid of Extrinsic Variations, and Use only Intrinsic Info
  - Option 1: Face Normalization
    - Geometric & Photometric Alignment
  - Option 2: Special Purpose Imaging System
    - Near Infrared Imaging
    - Others
- Make a Powerful Classifier
  - Able to deal with nonlinear variations
  - Framework: Local Features + Boosting Learning

Face Detection

Face Detection: Approach

- Scan the image with subwindows of varying size and location
- Classify a subwindow x into face/nonface
  - Need a “strong classifier” for accurate classification
- Post-processing: Merge multiple detects

Faces

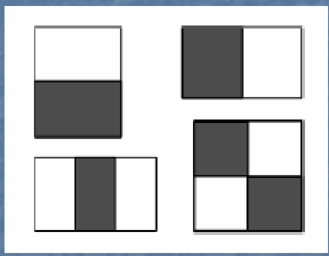
Nonfaces

State-of-the-Art Methods: Local Features + Boosting

- Viola & Jones, 2001
  - Haar Features + AdaBoost + Cascade
- Schneiderman & Kanade, 2000
  - Wavelet Histograms
- Li, et al, 2002
  - Extended Haar Features + FloatBoost + Pyramid
- Haizhou Ai, et al, 2003-2005
  - Omni-view face detection, Haar feature + Boosting + More advanced architecture

AdaBoost Method
(Viola & Jones)

Simple Haar features (Viola & Jones)



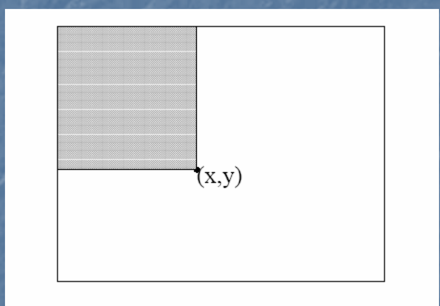
3 rectangular features types:

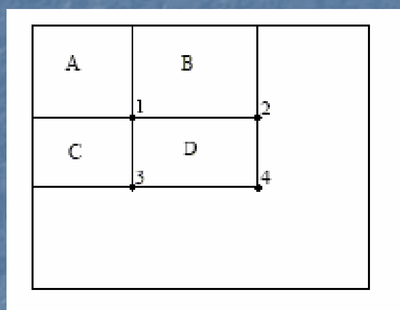
- *two-rectangle feature* type (horizontal/vertical)
- *three-rectangle feature* type
- *four-rectangle feature* type

These rectangular features, as opposed to more expressive steerable filters, can be computed very efficiently using integral images.

Using 24x24 windows → 49,396 features.

Integral Images





AdaBoost Learning

- Proposed by Freund et al 1997, 1998
- Task: Given $\{(x_{i,} y_i)\}$, learns $H_M(x)$ so that $y_i = \text{sign}(H_M(x))$
- Learns and combines a sequence of weak classifiers $h_m(x)$ into a strong classifier

$$H_M(x) = \sum_{m=1}^{M} \alpha_m h_m(x)$$

- h_m (x) are learned in stages to minimize error bound (see later)

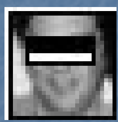
$$J(H_M(x)) = \sum_i e^{-y_i H_M(x_i)}$$

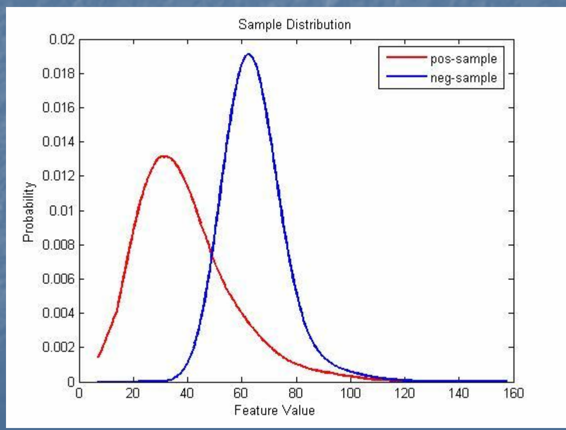
- Associate (x_i , y_i) with weight w_i and reweight after each iteration (see formula later)

Weak Classifiers

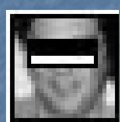
- One WC for a scalar Haar feature
- WC outputs face/nonface by comparing the scalar value with a threshold
- Best threshold obtained by examining the weighted histogram

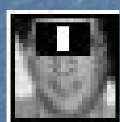
Learning Weak Classifiers Based on Weighted Histogram





Best Features Learned

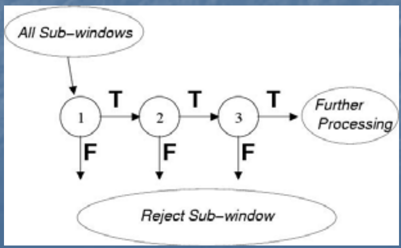




- First features selected by AdaBoost are meaningful and have high discriminative power
- By varying the threshold of the final classifier one can construct a two-feature classifier which has a detection rate of 1 and a false positive rate of 0.4.

Speed-up through Cascade

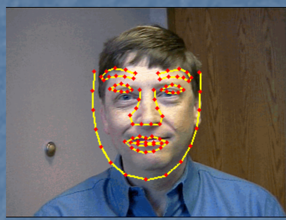
- Simple, boosted classifiers can reject many of negative sub-windows while detecting all positive instances.
- Series of such simple classifiers can achieve good detection performance while eliminating the need for further processing of negative sub-windows.



Face Alignment

Face Alignment

- Input:
  - Face detection/tracking output (location, scale, and pose)
- Output:
  - Accurate localization of facial outline and components
- Purpose:
  - For accurate facial feature extraction



Active Shape Models (ASM)

- Developed by Cootes, Taylor, *et al.*
  - The solution space is constrained by PDM, namely the global shape space.
  - Local appearance models derived at the landmarks converge to the local image evidence.

Formulation of ASM

- Global Shape Model: $S = \bar{S} + Us$
- Local Appearance Models:

$$(x, y) = \underset{(x,y)\in N(x_i^n, y_i^n)}{\arg\min} \| g_i(x, y) - \bar{g}_i \|^2_{\Sigma_i^g}$$

Where $\bar{g}_i$ is the average profile around the i-th landmark, and Σ_i^g is the covariance matrix of the sample profiles for the i-th landmark.

Formulation of ASM

- In each iteration, S_{lm} is obtained from the refinement of the local appearance models, the solution shape s is derived by maximizing the likelihood probability:

$$s = \arg\max_s p(S_{lm} \mid s) = \arg\min_s Eng(S_{lm}; s)$$

where

$$Eng(S_{lm}; s) = \lambda \| S_{lm} - S'_{lm} \|^2 + \| s - s_{lm} \|^2_{\Lambda}$$

Active Appearance Models(AAM)

- Cootes proposed and developed the Active Appearance Model (AAM)
  - Built based on PDM.
  - Shape and texture are combined for the *appearance* modeling.
  - Alignment is guided by minimizing the texture difference between model and ground truth.

Formulation of AAM

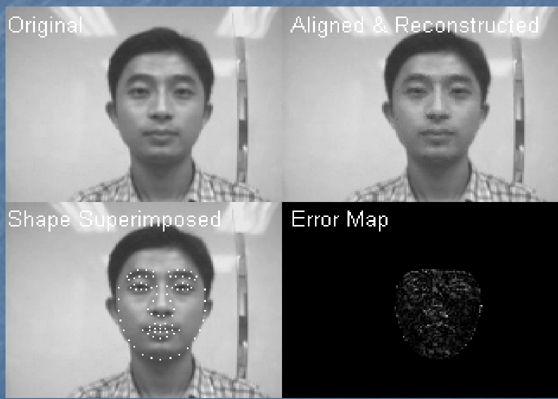
- Shape Model: $s = \bar{s} + Us$
- Texture Model: $T = \bar{T} + Vt$
- Appearance Model:

$$A = \begin{pmatrix} \Lambda s \\ t \end{pmatrix} \qquad A = Wa$$

- The search strategies are based on the linear regression assumptions:

$$\delta a = A_a \delta T \qquad \delta p = A_p \delta T$$

AAM/DAM



Face Recognition

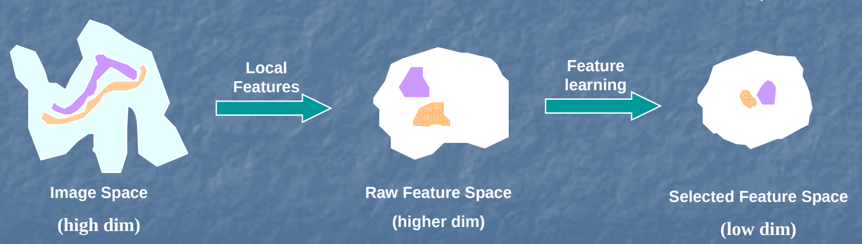
Local Features + AdaBoost Learning

Framework

- Local Features
  - Eg: Haar, Gabor, LBP, Ordinal, etc
  - Having good properties
  - Form a High-Dim Space
- Intra vs Extra Representation for Multi-class Problem
- Statistical Learning
  - 2-Class Classification
  - Training on pos and neg samples
  - Nonlinear classifier: Eg AdaBoos, SVM
  - Learning for
    - Dim reduction (feature selction)
    - Classifier construction

Working in Good Feature Space

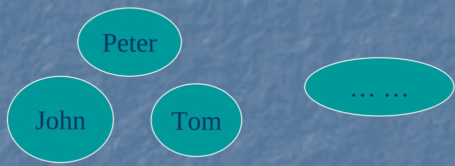
- Map input image to a higher dim local feature space
- Learning to select good features



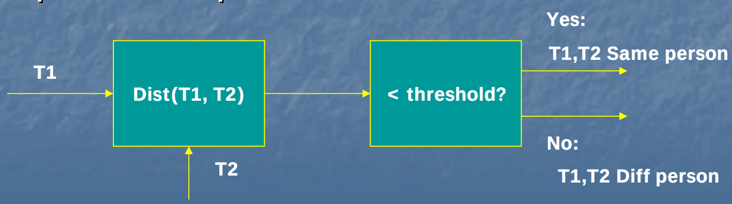
Intra vs Extra Representation: N Class → Two Class

(Baback Moghaddam)

- N persons

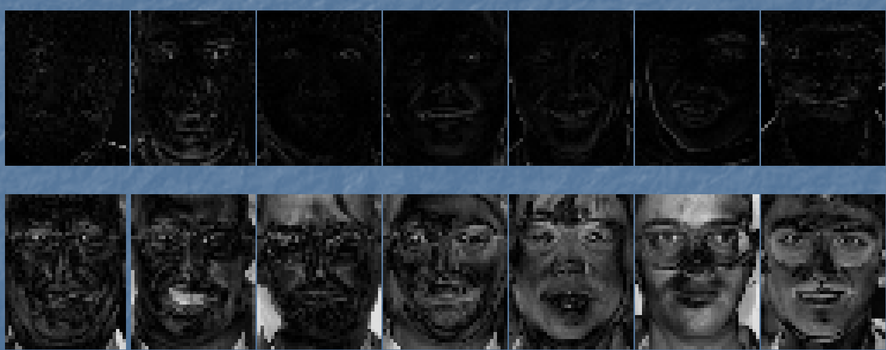


- Compare 2 templates

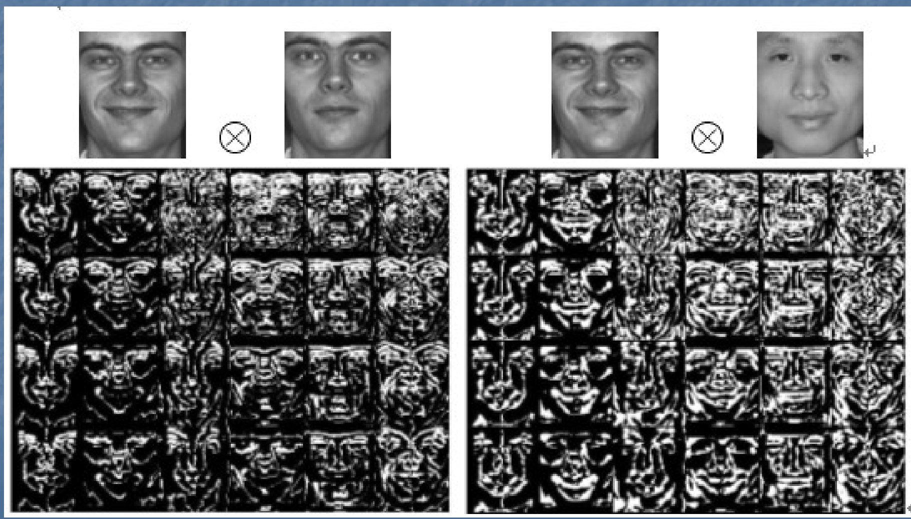


Intra- and Extra- personal Variations in Image Space

(Baback Moghaddam)



Differences of Ordinal Maps



Intra-Difference Extra-Difference

As Result of AdaBoost Learning

- Effective features are selected
- A weak classifier is constructed for each feature
- The weak classifiers are combined into a strong one
- Fusion at both feature and decision levels

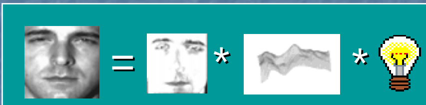
Face Recognition Using NIR Images

Imaging Models

- Face is a 3D
- Physical Imaging Model

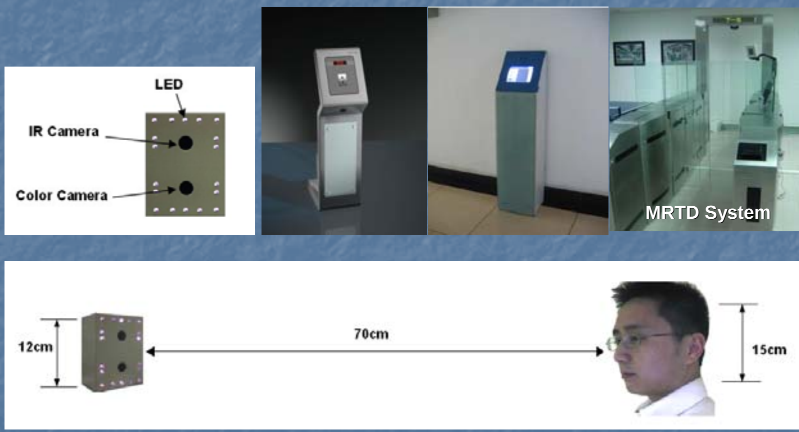
$$I(x,y) = \rho(x,y)\, n^T(x,y)\, s$$

(Lambertian Model)



- Imaging Factors
  - Shape $n(x,y)$ – intrinsic factor
  - Albedo $\rho(x,y)$ – intrinsic factor
  - Illumination $s = (s_1, s_2, s_3)$ – extrinsic factor

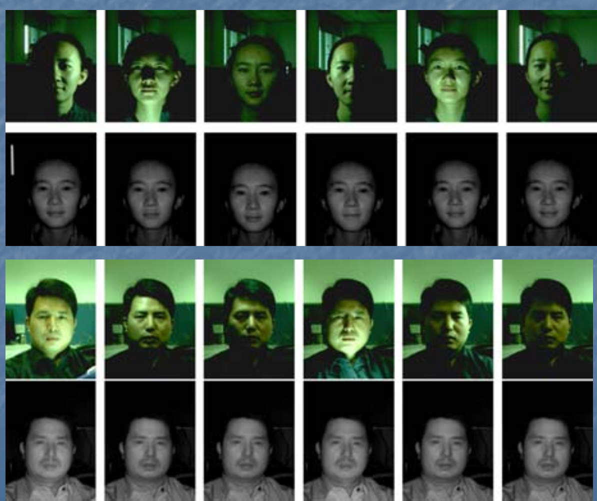
NIR imaging Hardware



$I(x,y)=\rho(x,y)n^T(x,y)s$ with $s = (0,0,1)$

$=\rho(x,y)n_z(x,y)$

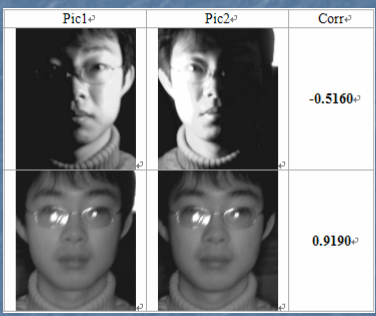
VL vs. NIR Images Under Various Lighting

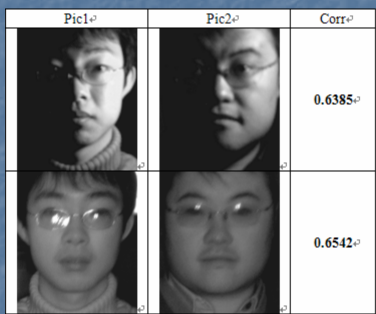


Advantages

- Working in invisible spectrum. VL can be filtered out
- Invisible to human eyes: non-intrusive way of active lighting

Visible Light vs. NIR Images





Correlation Coefficients

NIR+LBP Gives
Illumination Invariant Features

Active NIR Image

$I(x,y) = \rho(x,y) n^T(x,y) s$ with $s = (0,0,1)$

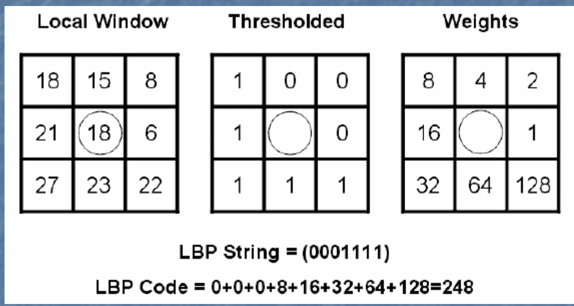
$= \rho(x,y) n_z(x,y)$

$$I(x,y) \propto \kappa \rho(x,y) \cos\theta(x,y)$$

$$I(x,y) = \kappa \rho(x,y) n_z(x,y)$$

- It is subject to an unknown constant κ, or a Monotonic Transform, only
- DOF overcome by use of LBP

Local Binary Pattern (LBP)
(University of Oulu)



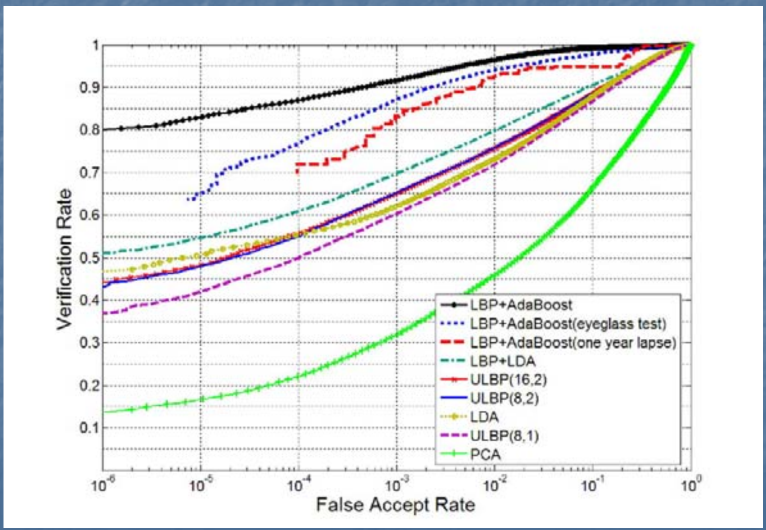
- LBP code of NIR images are invariant to environmental illumination changes

Classifier

- LBP Features+ Boosting Learning
  - LBP Feature Selection
  - Classifier Learning

Performance

Performance Evaluation



AuthenMetric System

- Assumptions
  - For Cooperative Applications
  - Applications: Access control, E-Passport, ATM, etc
- Features
  - Hardware: Active NIR image capture device to minimizes influence of environmental lighting
  - Recognition Engine: Classifier learned using LBP features + AdaBoost
- Live Demo

Applications

Face Biometric Applications

- Consumer products: Eg. Face Logon
- Enterprise: Eg. Time attendance and access control
- Governmental
  - Self-Service Border-crossing (deployed)
    - ShenZhen – Hong Kong Boarder since June 2005
    - Zhuhai – Macau Boarder since April 2006
  - Biometric E-Passport (on-going)

FaceLogon

Biometric Border-Crossing: ShenZhen – HongKong

- 400,000 border-crossings every day
- Two scenarios：Passengers & Vehicle Drivers
- 100+ gates deployed by now
- Two Modalities: Face & Fingerprint
- 1,600,000 people enrolled.
- Verification Speed: 6 sec / crossing
- 35,000,000 crossings since June 2005

CBSR Talents

Thank You

Contact:
Prof. Stan Z. Li
Center for Biometrics and Security Research
Institute of Automation, Chinese Academy of Sciences
szli@cbsr.ia.ac.cn