Empowered by Innovation NEC

IJCB2014

Face Recognition: Beyond the Limit of Accuracy

NEC Corporation
Information and Media Processing Laboratories
Hitoshi Imaoka

h-imaoka@cb.jp.nec.com

What is the hurdle in face recognition?

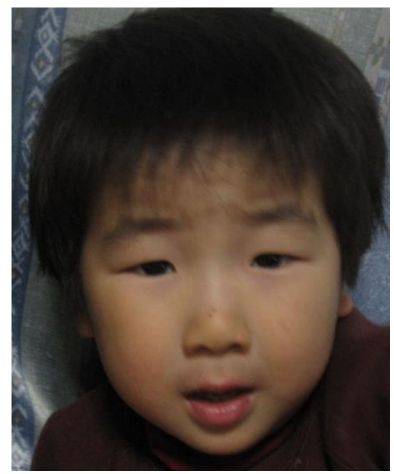
Motivation of my research

"Accuracy" is the most important

Question: Which of these three pictures is me?

Query image

A



B

C

Motivation of my research

Query image

A

B

C

Even in this sample, a lot of problems include

- long term aging change

- facial view, expression, similar face etc.

Why is face recognition so difficult?

Other variations

- view and illumination
- aging change
- facial expression
- makeup
- identical twins
- plastic surgery etc.

Most facial parts can be changed

Outline

- Face recognition algorithm
- Evaluation results by NIST and LFW
- Experimental results
  - Fusion of Human and Automatic Recognition
- Application examples
  - Is face recognition useful tool in our real life?
- Summary

Progress of Face Recognition Algorithm

Processing flow of face recognition algorithm

Query image

Feature Extraction

Target image

Feature Extraction

Feature Matching

- **Distance**
- **Cosine**

etc.

Similarity Score

Mate or Non-mate

Processing flow of face recognition algorithm

Query image

Target image

Feature Extraction

Feature Extraction

Feature Matching

Similarity Score

Mate or Non-mate

How to extract optimal features?

Progress of Face Recognition Algorithm

1st generation (1990～)	2nd generation (2000～)	3rd generation (2010～)
Basic method	**Linear Subspace method**	**Non-Linear method**
Eigenface Fisherface	Sparse Representation Metric learning	Deep Learning

Linear method → **Non Linear method**

Generative model → **Discriminative model**

Simple features → **Complex features**

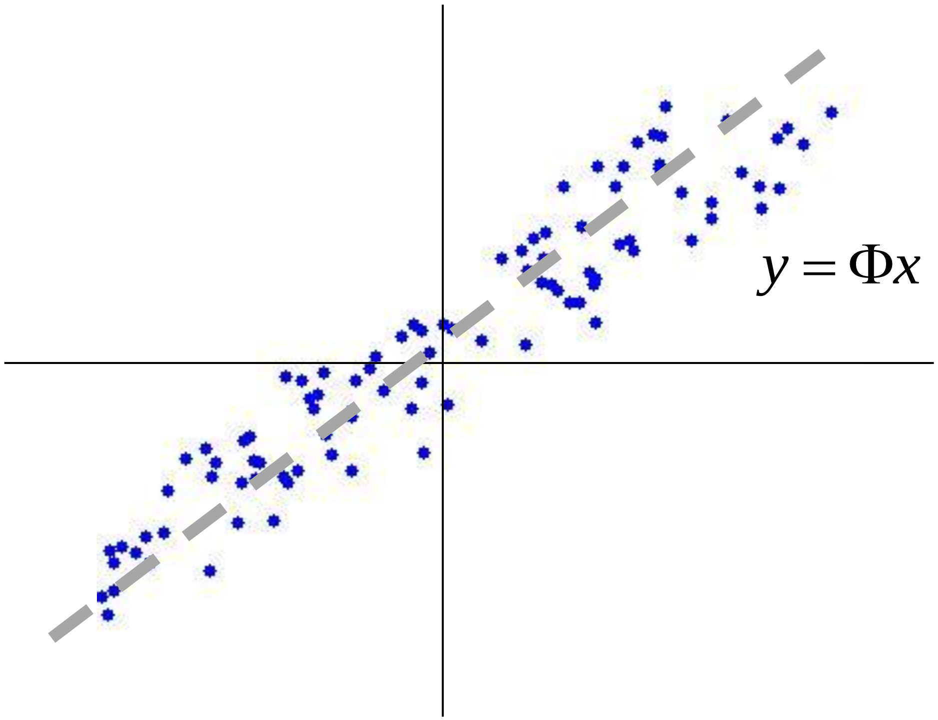
Progress of Face Recognition Algorithm

1st generation (1990～)	2nd generation (2000～)	3rd generation (2010～)
Basic method	**Linear Subspace method**	**Non-Linear method**
Eigenface Fisherface	Sparse Representation Metric learning	Deep Learning

Linear method	→	**Non Linear method**
Generative model		**Discriminative model**
Simple features		**Complex features**

1st generation: Eigenface

Turk, M. A. and Pentland, Alex P. ”*Face recognition using eigenfaces”*. Computer Vision and Pattern Recognition, 1991.



Based on Principal Component Analysis (PCA)

Projection vector is a set of eigenvector of training samples

Top 4 eigenface

Limitation

PCA projection is optimal for reconstruction of face, but may not be optimal for discrimination

1st generation: Fisherface

P. Belhumeur, J. Hespanha, and D. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection”,PAMI, 19(7):711--720, 1997.

- **Based on Linear Discriminant Analysis (LDA)**
  - **Optimal subspace is obtained by maximizing the ratio of between and within class scatter matrix:**

$$r = \left\| \Phi^T S_b \Phi \right\| / \left\| \Phi^T S_w \Phi \right\|$$

S_b : between class scatter matrix

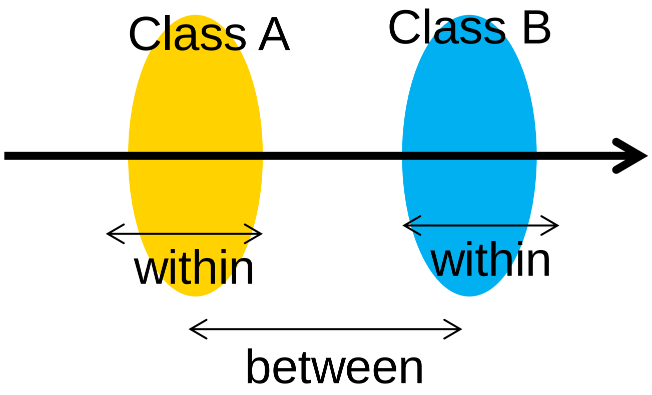
S_w : within class scatter matrix

Four top fisherfaces

http://www.scholarpedia.org/article/Fisherfaces

Limitation

It is difficult to discriminate faces near the individual boundaries



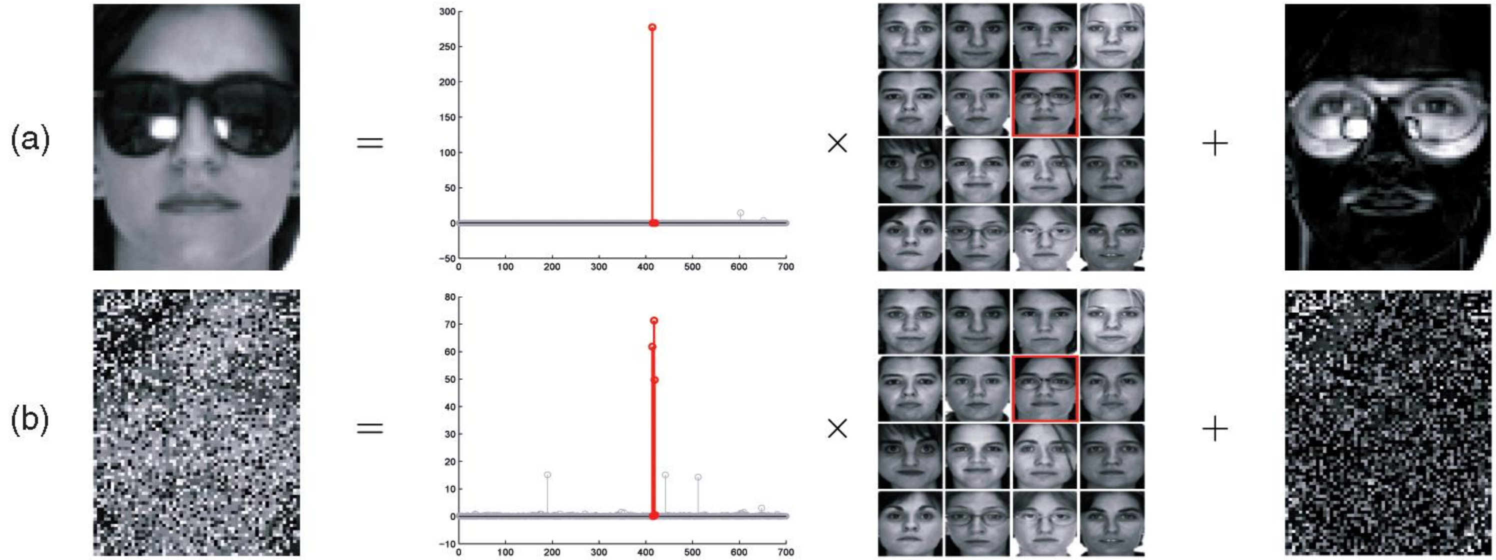
Progress of Face Recognition Algorithm

1st generation (1990～)	2nd generation (2000～)	3rd generation (2010～)
Basic method	**Linear Subspace method**	**Non-Linear method**
Eigenface Fisherface	Sparse Representation Metric learning	Deep Learning

Linear method	**Non Linear method**
Generative model	**Discriminative model**
Simple features	**Complex features**

2nd generation: **Sparse Representation**

- Allen Y. Yang, Arvind Ganesh and Yi Ma, ”The basic idea is to cast **recognition** as a **sparse representation** problem, utilizing new mathematical tools from compressed sensing and L1 minimization”, PAMI 2009.



- **Train sparse matrix under L1 minimization constraint**
- **Decomposed as sparse components and remaining elements**
- **By sparse representation, robust against occluding facial parts**

2nd generation: Metric Learning Approach

- **Metric Learning Approach**

Distance metric between feature xi and xj

$$d(x_i, x_j) = (x_i - x_j)^T M (x_i - x_j)$$

M **is a symmetric positive definite matrix**

design matrix M to discriminate Mate and Non-mate class

Query image

Target image

optimize matrix M

Mate or Non-mate

2nd generation: Metric Learning Approach

- **J. Davis, B. Kulis, P. Jain, S. Sra, and I. Dhillon. "Information theoretic metric learning". In ICML, 2007.**

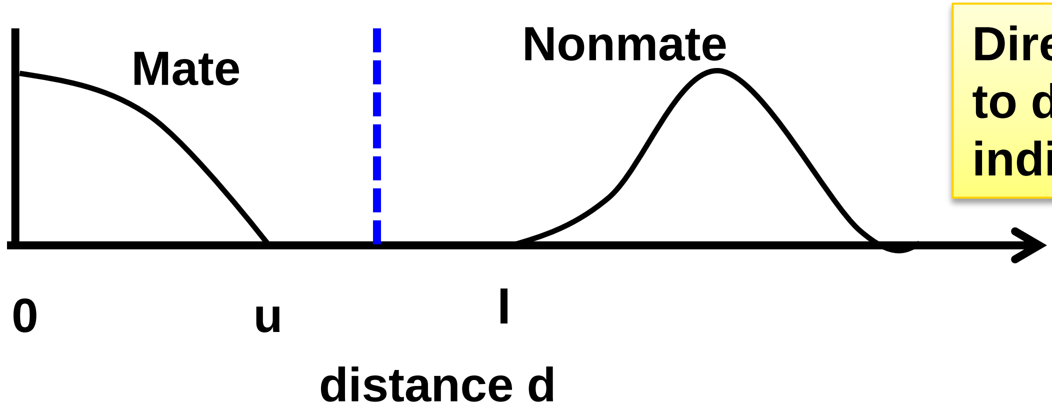
  Objective function : Kullback-Leibler divergence criterion

$$\min_{A} KL(p(x;A_0) \parallel p(x;A))$$

constraints

$$d_A(x_i, x_j) \le u \quad (i,j) \in mate\ pair$$

$$d_A(x_i, x_j) \ge l \quad (i,j) \in nonmate\ pair$$

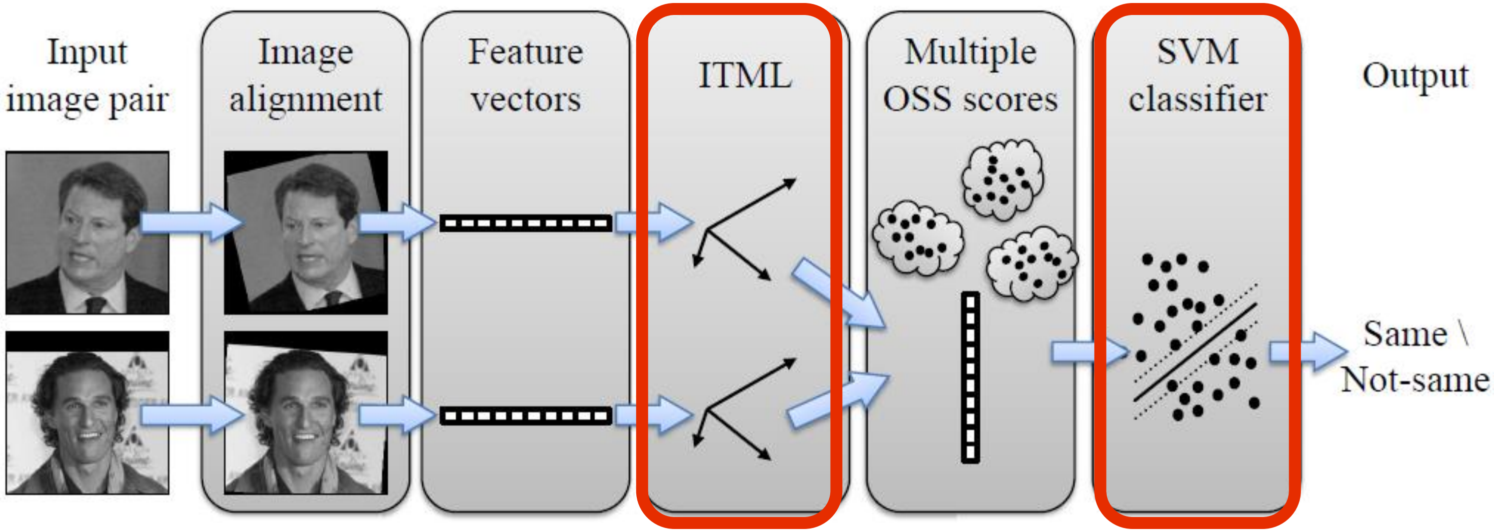


Direct approach to discriminate individual

2nd generation: Metric Learning Approach

YANIV TAIGMAN, LIOR WOLF, AND TAL HASSNER.
MULTIPLE ONE-SHOTS FOR UTILIZING CLASS LABEL INFORMATION.
BRITISH MACHINE VISION CONFERENCE (BMVC), 2009.

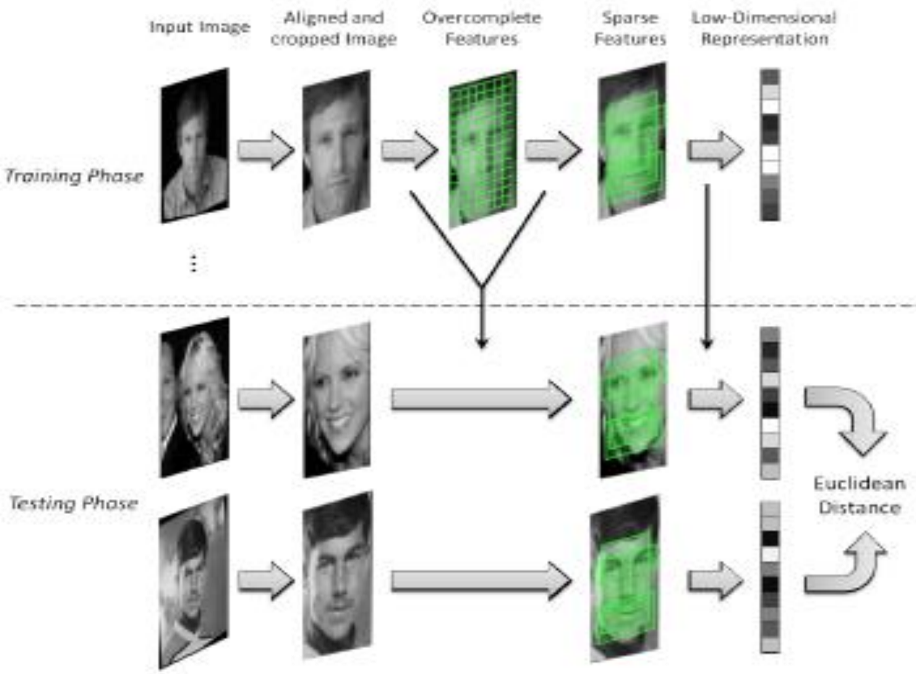
Algorithm using Information theoretic metric learning



LFW DATABASE 1-EER= 89%

2nd generation: Metric Learning Approach

Chang Huang, Shenghuo Zhu, and Kai Yu."**Large Scale Strongly Supervised Ensemble Metric Learning, with Applications to Face Verification and Retrieval.**"*NEC Technical Report TR115*, 2011.



Distance metrics learning is difficult to use in a high dimensional feature space

Joint metric learning : two step approach

- **select effective feature groups from feature pool**
- **train optimal subspace by distance metric learning**

LFW DATABASE 1-EER= 92%

Progress of Face Recognition Algorithm

1st generation (1990～)

Basic method

Eigenface
Fisherface

2nd generation (2000～)

Linear Subspace method

Sparse Representation
Metric learning

3rd generation (2010～)

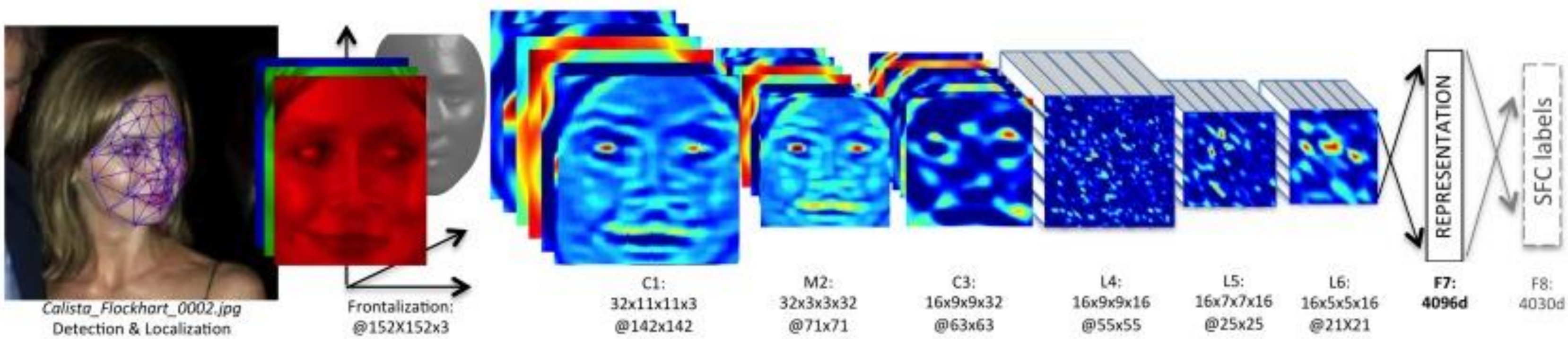
Non-Linear method

Deep Learning

Linear method	→	Non Linear method
Generative model		Discriminative model
Simple features		Complex features

3rd generation: Deep Learning (DeepFace)

- **Align face by 2D and 3D affine transformation**
- **Extract feature vector by deep neural network**
  - **Training data: 4.4million images/ 4030 subjects**
- **Compare features by distance metric**



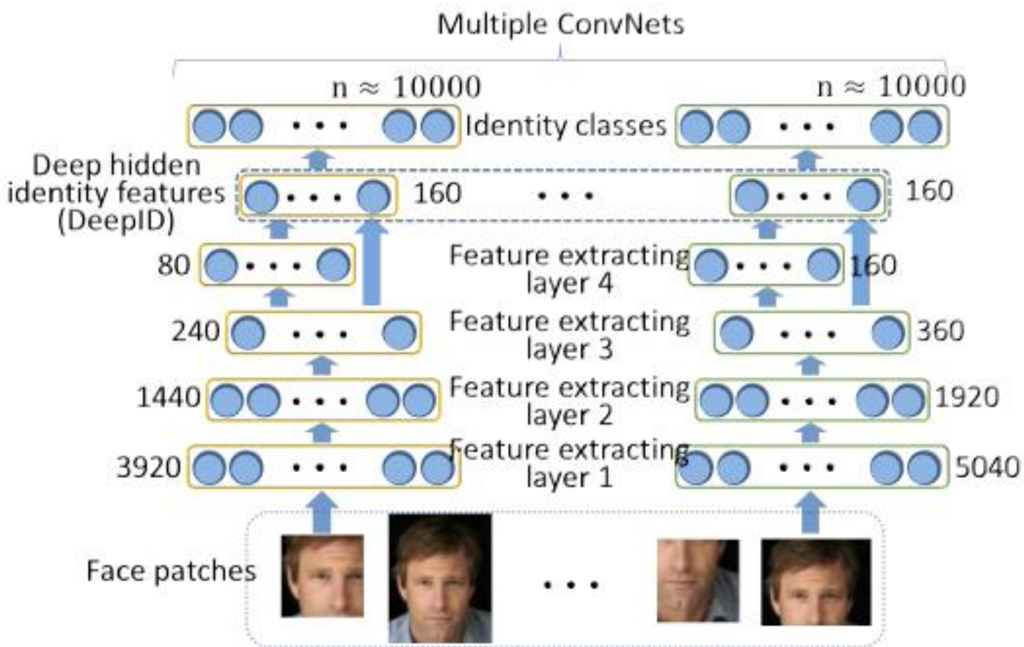
Matthias Hullin, Qionghai Dai; DeepFace: Closing the Gap to Human-Level Performance in Face Verification

LFW DATABASE 1-EER= 97%

3rd generation: Deep Learning (DeepID)

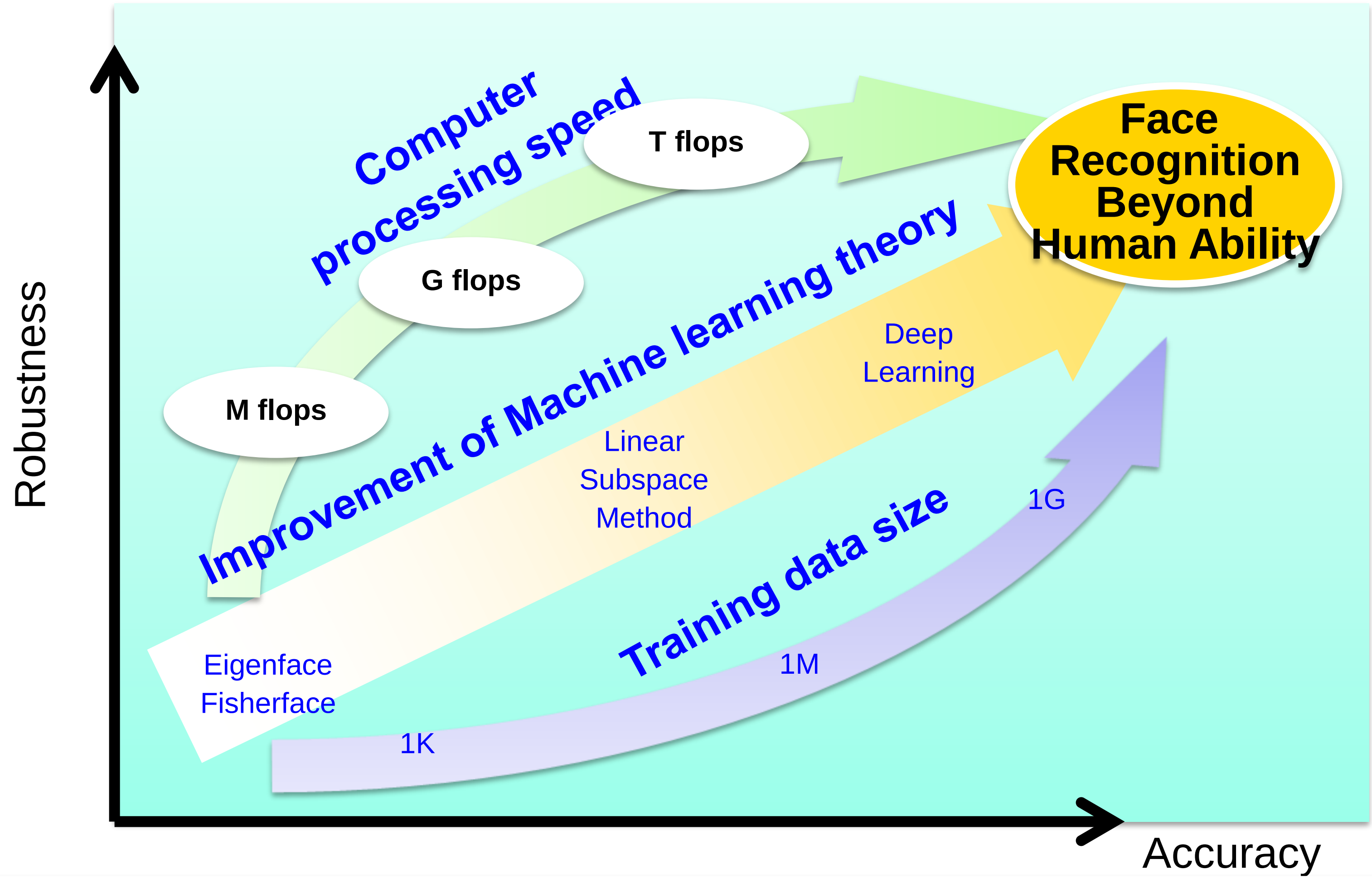
YI SUN, XIAOGANG WANG, AND XIAOOU TANG.
DEEP LEARNING FACE REPRESENTATION BY JOINT IDENTIFICATION-VERIFICATION.

- **Extract facial image dividing several face patches**
- **Fusion of multiple convolutional neural networks**



LFW DATABASE 1-EER= 99%

Direction of face recognition algorithm



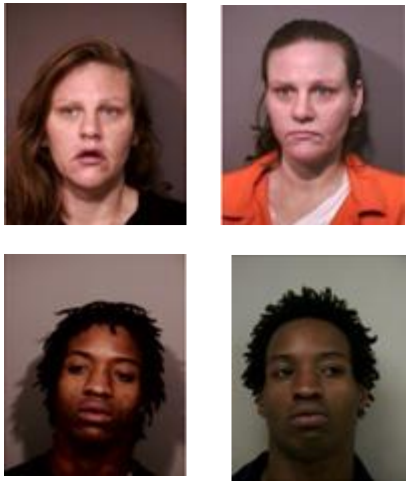
By above 3 elements, computer face recognition accuracy will overtake human recognition ability

Evaluation Result of Face Recognition

NIST benchmark and LFW database evaluation

NIST benchmark

- Controlled images (Criminal operational data)
- Closed data (it is difficult to tune algorithm)
- Algorithm is closed. Only evaluation results is reported.
- Useful to know accuracy in large scale dataset (over 1 million)



Technical Report, 8009, National Institute of Standards and Technology, May 21 2014

LFW (Labeled Faces in the Wild) database

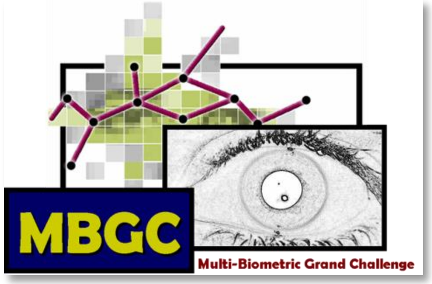
- Uncontrolled images (Web data)
- Open data (it is easy to tune algorithm)
- Most algorithms are open to the public
- Useful to know effectiveness of algorithm in medium size of dataset (16,000 images)

http://vis-www.cs.umass.edu/lfw/

NIST benchmark result

NIST's Face Recognition Evaluation Program

- **NIST benchmark test started in 1993**
- **Purpose**
  - **Independent government evaluations of commercial and academic algorithms**
  - **Identify future research directions for research community**



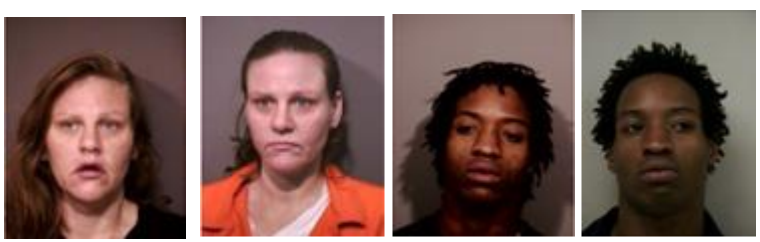
Multiple Biometric Grand Challenge
in 2009

Multiple Biometrics Evaluation
in 2010

Face Recognition Vendor Test
in 2013

Overview of the Face Recognition Vendor Test 2013 (FRVT)

- **Final report published in May 2014**
- **Target applications**
  - **criminal investigations and immigration control**
- **16 participating vendors and universities worldwide**
- **Large scale face database : over 1 million**



High-quality image
Criminal application

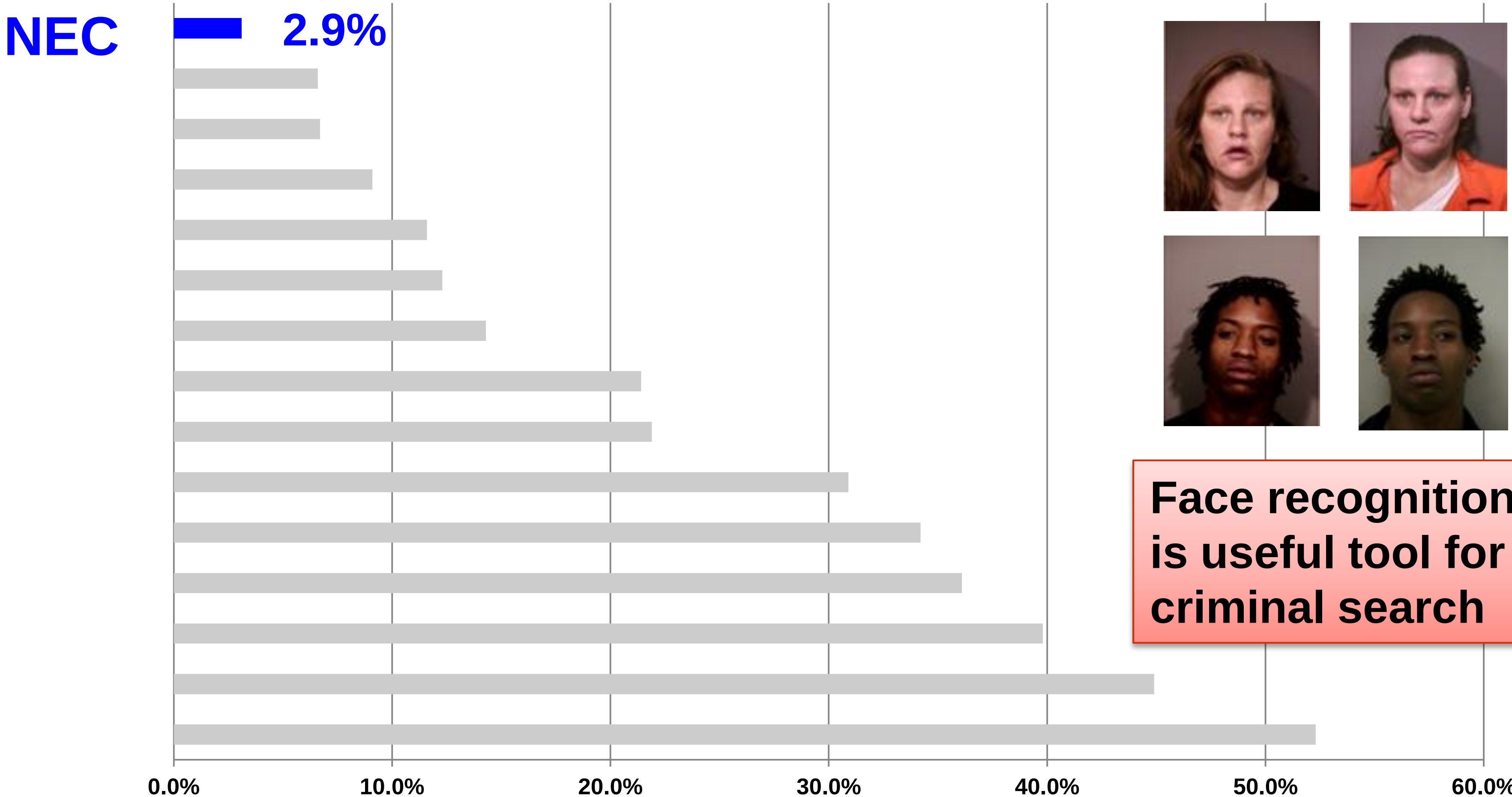
Low-quality image
Surveillance application

Patrick Grother and Mei Ngan, “Face Recognition Vendor Test (FRVT) Performance of Face Identification Algorithms”, Technical Report, National Institute of Standards and Technology, May 21, 2014

Result of Face Recognition Vender Test 2013

Rank-1 miss identification rates in High-quality image

(number of enrolled subject N=160,000) **Mugshot**



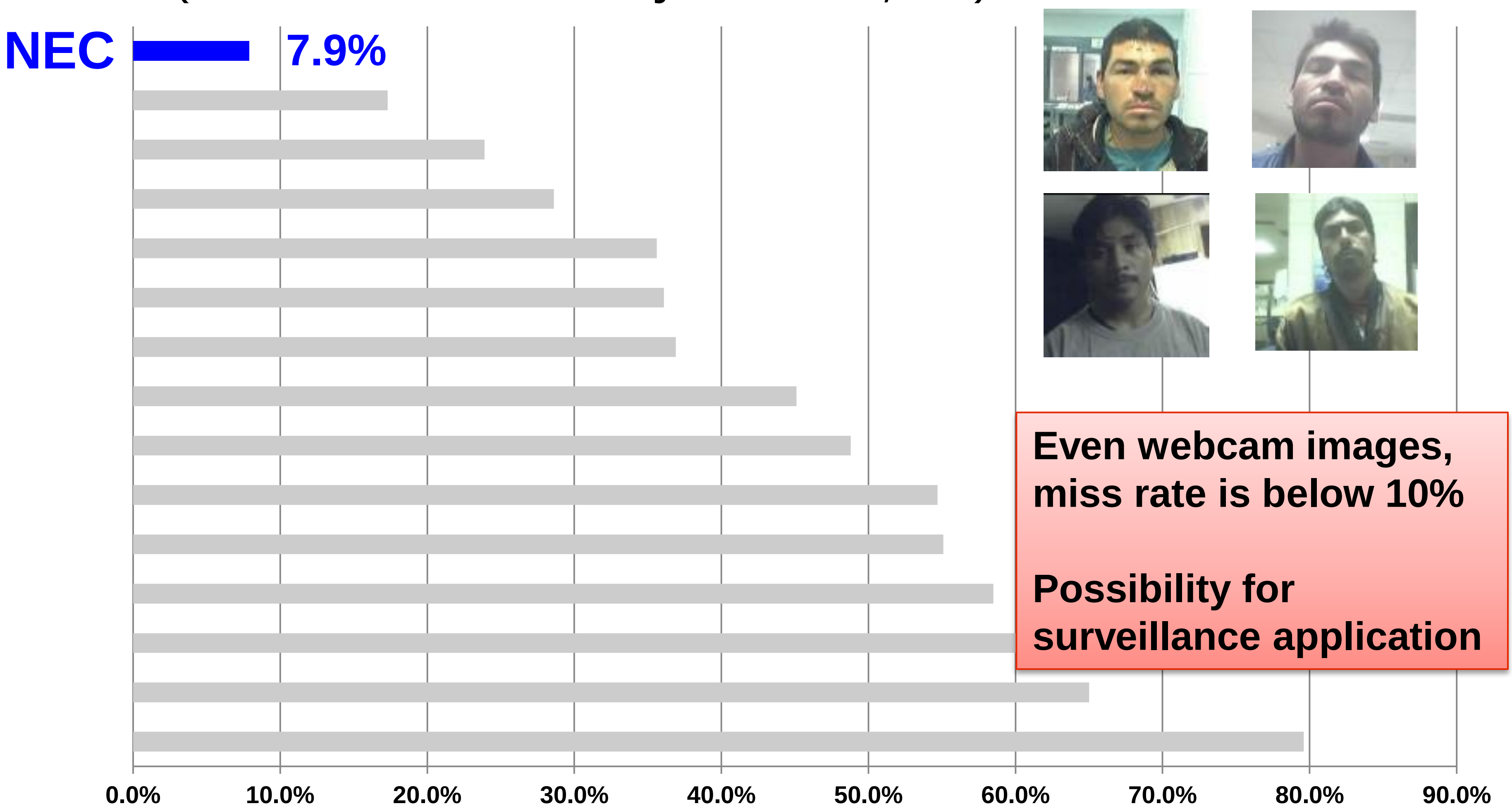
Face recognition is useful tool for criminal search

Technical Report, 8009, National Institute of Standards and Technology, May 21 2014

Result of Face Recognition Vender Test 2013

Rank-1 miss rates in Low-quality image (Webcam)

(number of enrolled subject N=160,000)



Even webcam images, miss rate is below 10%

Possibility for surveillance application

Technical Report, 8009, National Institute of Standards and Technology, May 21 2014

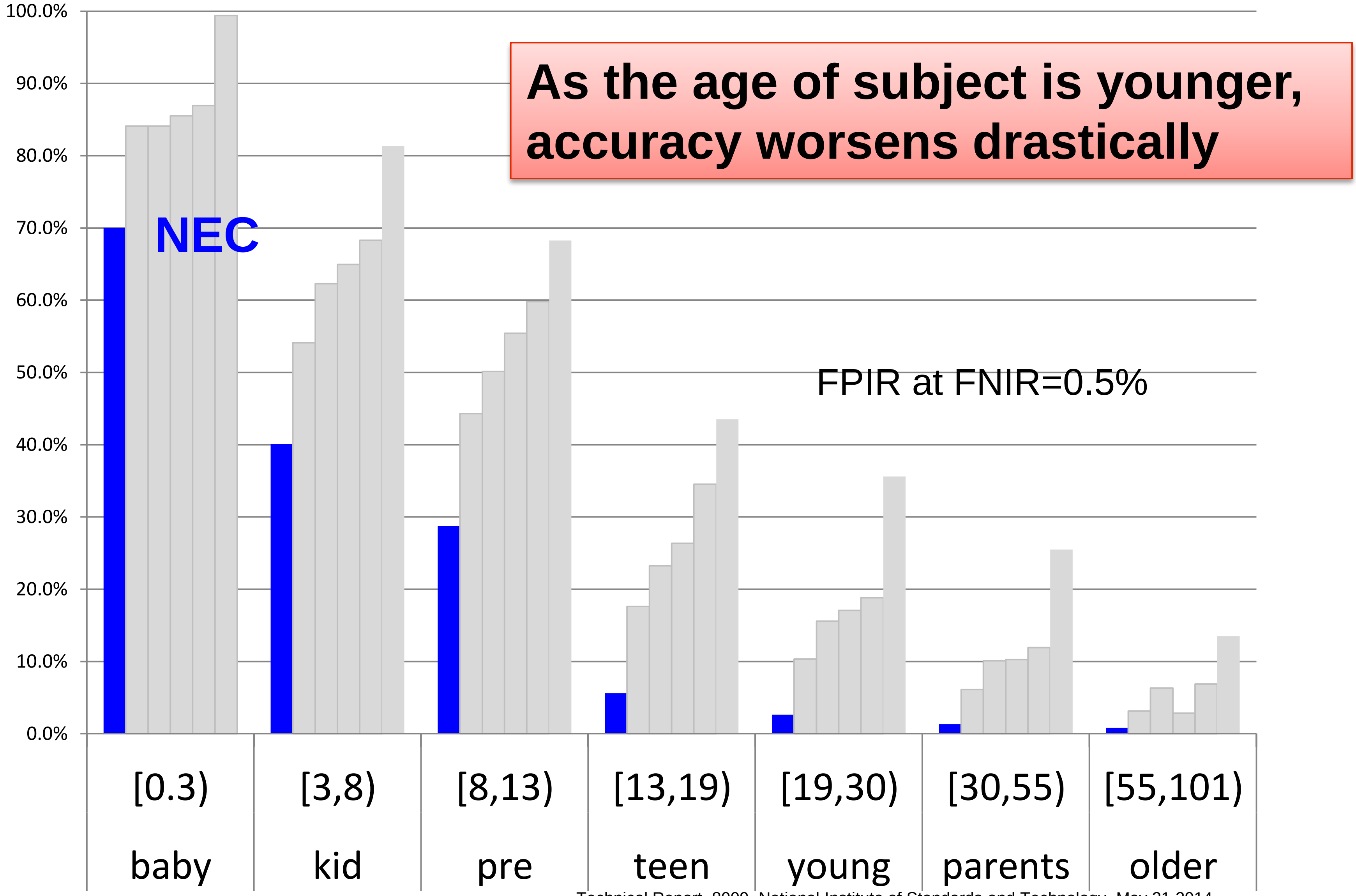
Result of Face Recognition Vender Test 2013

Accuracy dependence on subject age

Group No.	Group Label	Age Range	Search Age Mean	Mated Time Lapse Mean	Mated Count
1	*baby*	$[0, 3)$	2.3	1.6	57
2	*kid*	$[3, 8)$	5.7	2.8	340
3	*pre*	$[8, 13)$	10.7	3.7	533
4	*teen*	$[13, 19)$	17.0	2.5	1447
5	*young*	$[19, 30)$	25.4	2.0	5930
6	*parents*	$[30, 55)$	40.5	2.1	8293
7	*older*	$[55, 101)$	63.6	2.2	2709

Technical Report, 8009, National Institute of Standards and Technology, May 21 2014

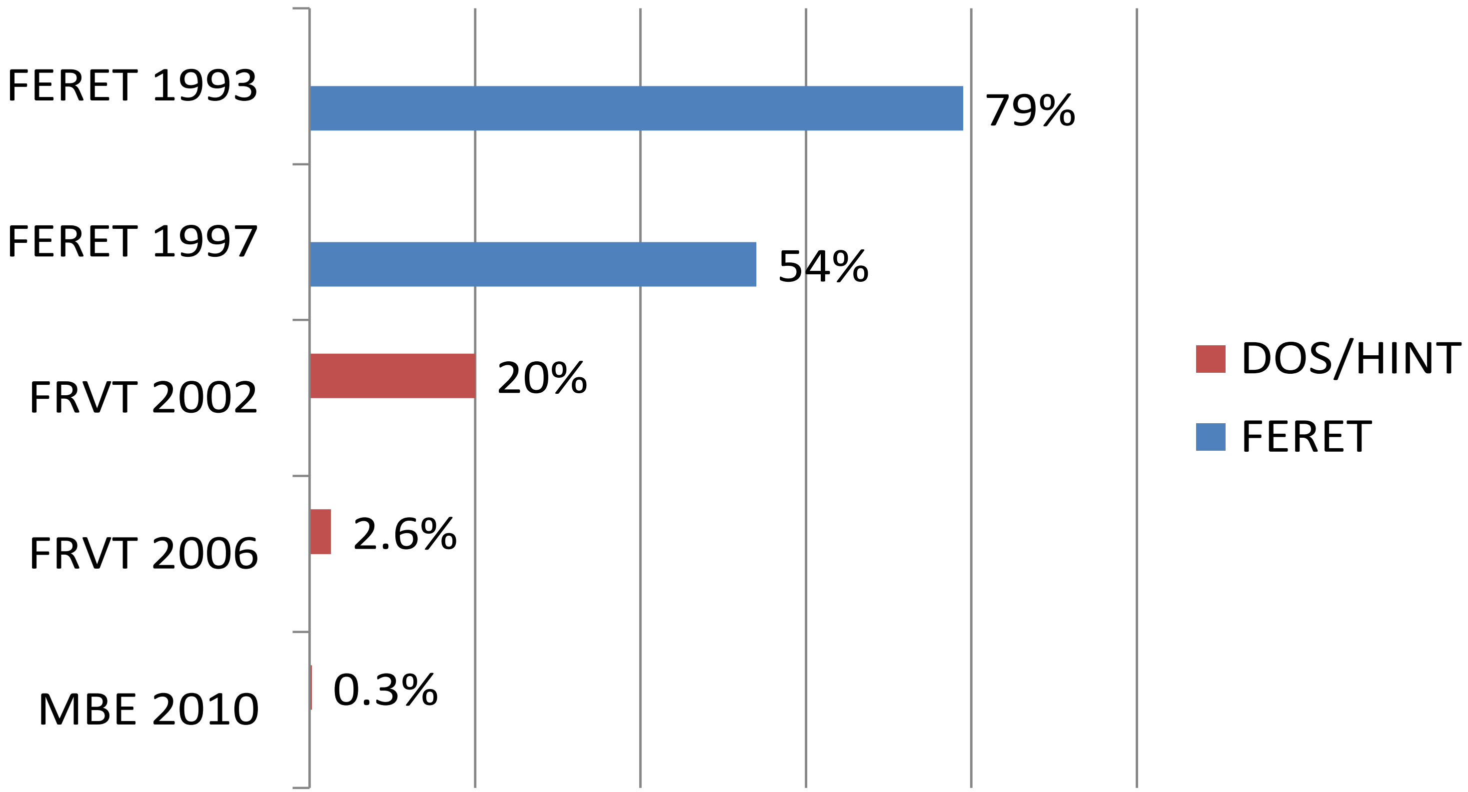
Result of Face Recognition Vender Test 2013



Technical Report, 8009, National Institute of Standards and Technology, May 21 2014

Progress of NIST evaluation result

Remarkable advance in these 20 years



False non-match rate(FNMR) at false match rate(FMR) 0.1%

Report on the Evaluation of 2D Still-Image Face Recognition Algorithms NIST Interagency Report 7709

LFW database Result

LFW database

Uncontrolled dataset

- facial expression
- facial view
- illumination change
- Occlusion (hand etc.)

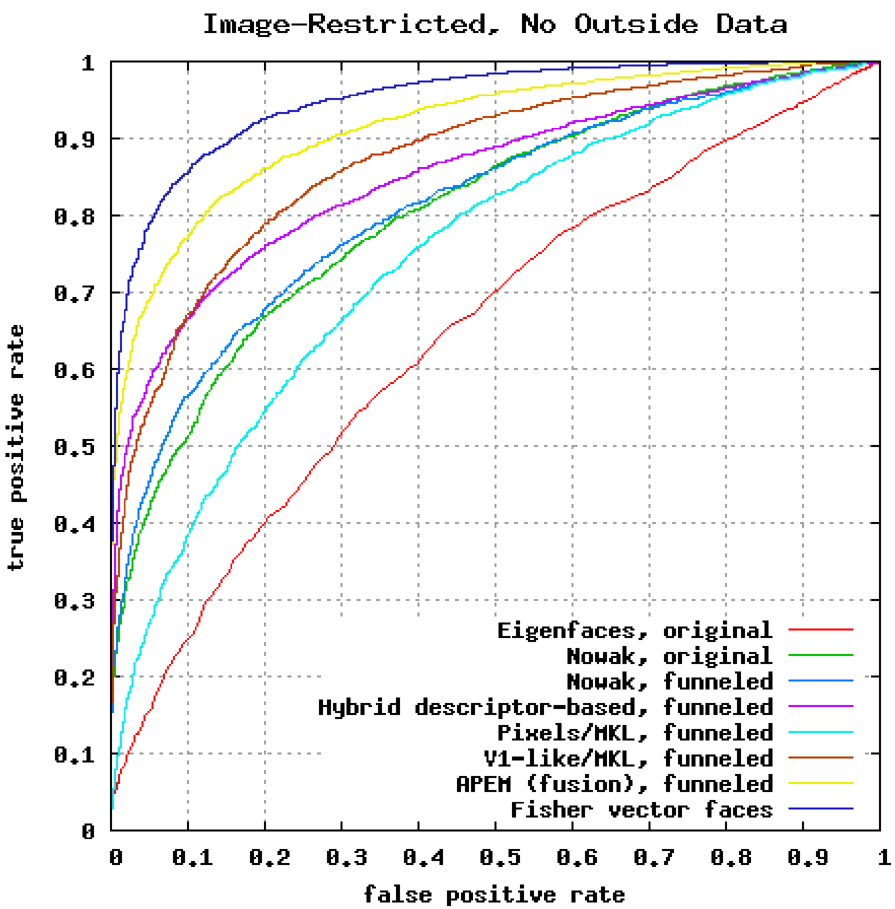
Resolution is not low

- Intra-ocular distance is about 90 pixels.

http://vis-www.cs.umass.edu/lfw/

LFW database result (Image-Restricted, No Outside Data)

Restricted training data : compare accuracy of algorithms



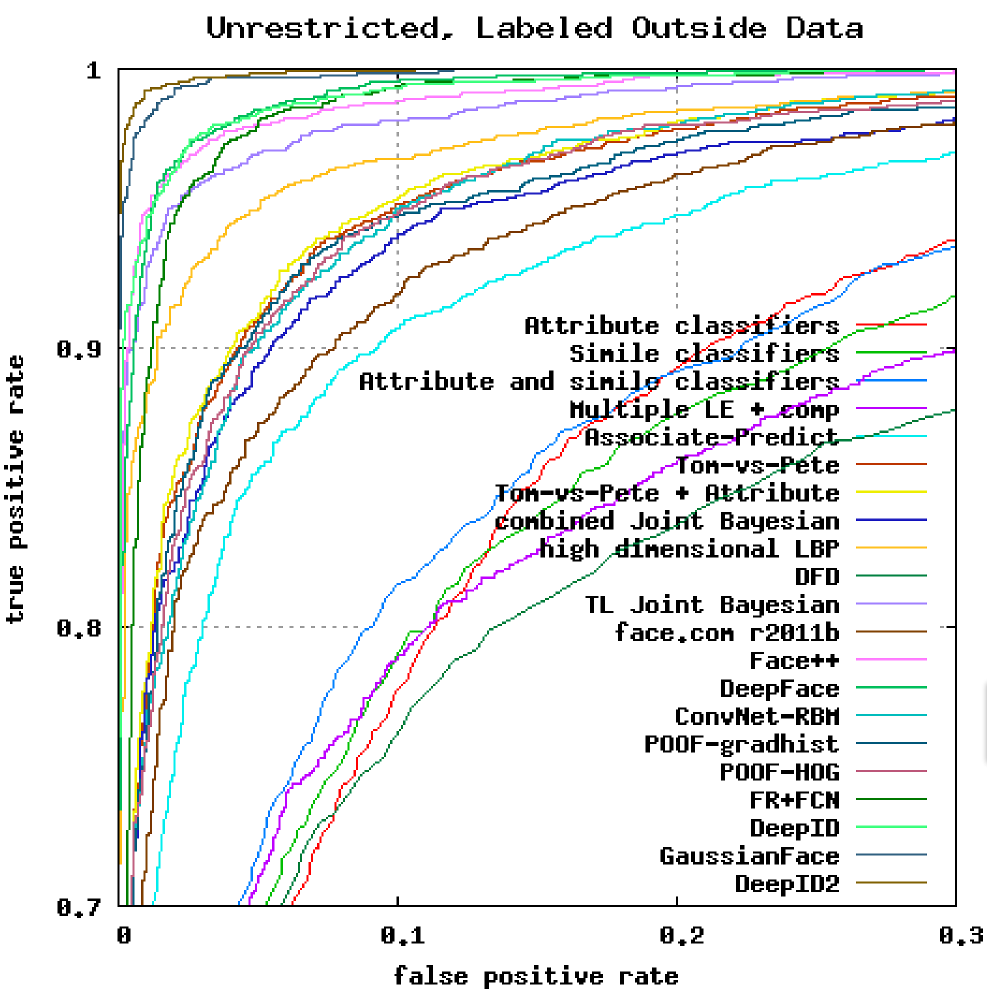
	$\hat{u} \pm S_E$
Eigenfaces[1], original	0.6002 ± 0.0079
Nowak[2], original	0.7245 ± 0.0040
Nowak[2], funneled[3]	0.7393 ± 0.0049
Hybrid descriptor-based[5], funneled	0.7847 ± 0.0051
3x3 Multi-Region Histograms (1024)[6]	0.7295 ± 0.0055
Pixels/MKL, funneled[7]	0.6822 ± 0.0041
V1-like/MKL, funneled[7]	0.7935 ± 0.0055
APEM (fusion), funneled[25]	0.8408 ± 0.0120
MRF-MLBP[30]	0.7908 ± 0.0014
Fisher vector faces[32]	0.8747 ± 0.0149

Best performance : 1-EER=87%

In case that training data size is small, accuracy is not good

LFW database result (Unrestricted, Labeled Outside Data)

Unrestricted training data : limit of accuracy



Recent result

DeepFace-ensemble[41]	0.9735 ± 0.0025
ConvNet-RBM[42]	0.9252 ± 0.0038
POOF-gradhist[44]	0.9313 ± 0.0040
POOF-HOG[44]	0.9280 ± 0.0047
FR+FCN[45]	0.9645 ± 0.0025
DeepID[46]	0.9745 ± 0.0026
GaussianFace[47]	0.9852 ± 0.0066
DeepID2[48]	0.9915 ± 0.0013

Best performance : 1-EER=over 99%

If we can use numerous training data, almost 100% accuracy may be achieved

Summary of evaluation result

- **In the last 20 years, accuracy has improved rapidly**
- **However some obstacles still remain**

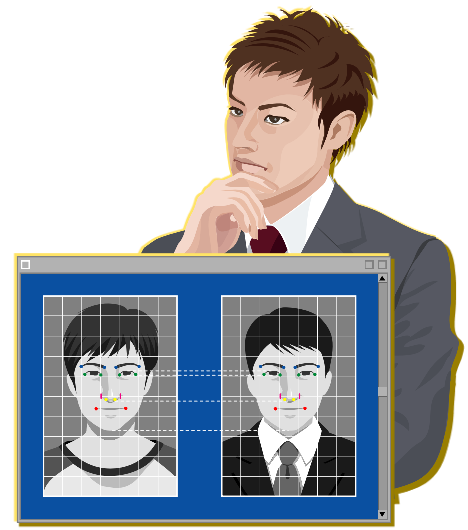
Obstacle factor	Easy	Possible	Difficult
pose (tilt)	frontal	～30 degree	profile
Illumination	normal	severe change	
expression	slight	drastic change	
aging change	within 1 year	～10 years	over decades
subject’s age	over 20 years old	teenager	baby
resolution (intraocular distance)	over 60 pixel	20-30 pixel	under 10 pixel
Occlusion	no	glasses/beard makeup	dark sunglass
other factors	-	ethnicity plastic surgery	Identical twins

Human vision accuracy: Fusion of machine recognition and human recognition

Question

- In the verification task, can the human brain assist the machine generated recognition result?

Machine recognition



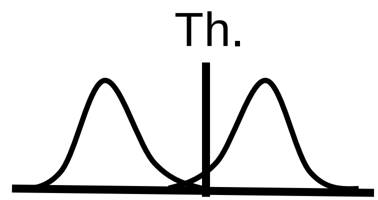
Human recognition

Fusion

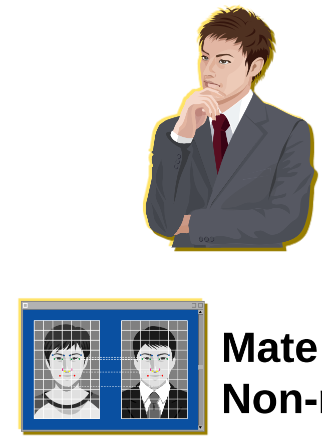
Experimental procedure

Step1

Step2



Step3



Mate or Non-mate?

Calculate matching score by machine recognition

Randomly select same number of mate and nonmate pairs near the threshold

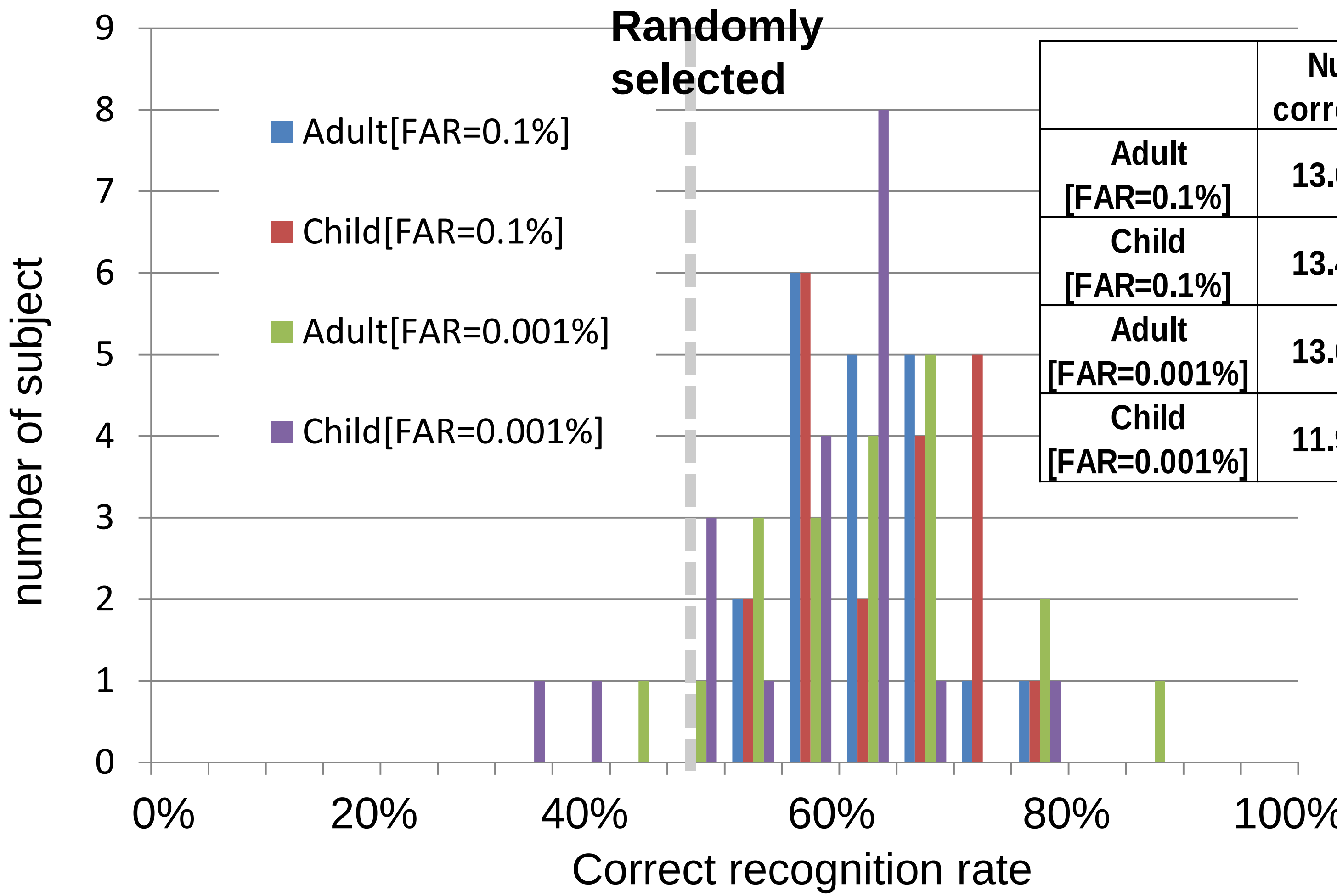
Judge mate or non-mate pair using human brain

- Number of subjects is 20
- Subject knows that mate and non-mate pair is mixed as the same number

Experimental condition

Test Set	Face Database	Threshold	Num. of mate pair	Num. of non-mate pair	EER by machine recognition
1	**Adult** - over 20 years old - aging change over decades	FAR 0.1% similarity is low	10	10	4.0%
2	**Child** - under 10 years old	FAR 0.1% similarity is low	10	10	11.9%
3	**Adult** - over 20 years old - aging change over decades	FAR 0.001% similarity is very high	10	10	4.0%
4	**Child** - under 10 years old	FAR 0.001% similarity is very high	10	10	11.9%

Experimental result by human recognition

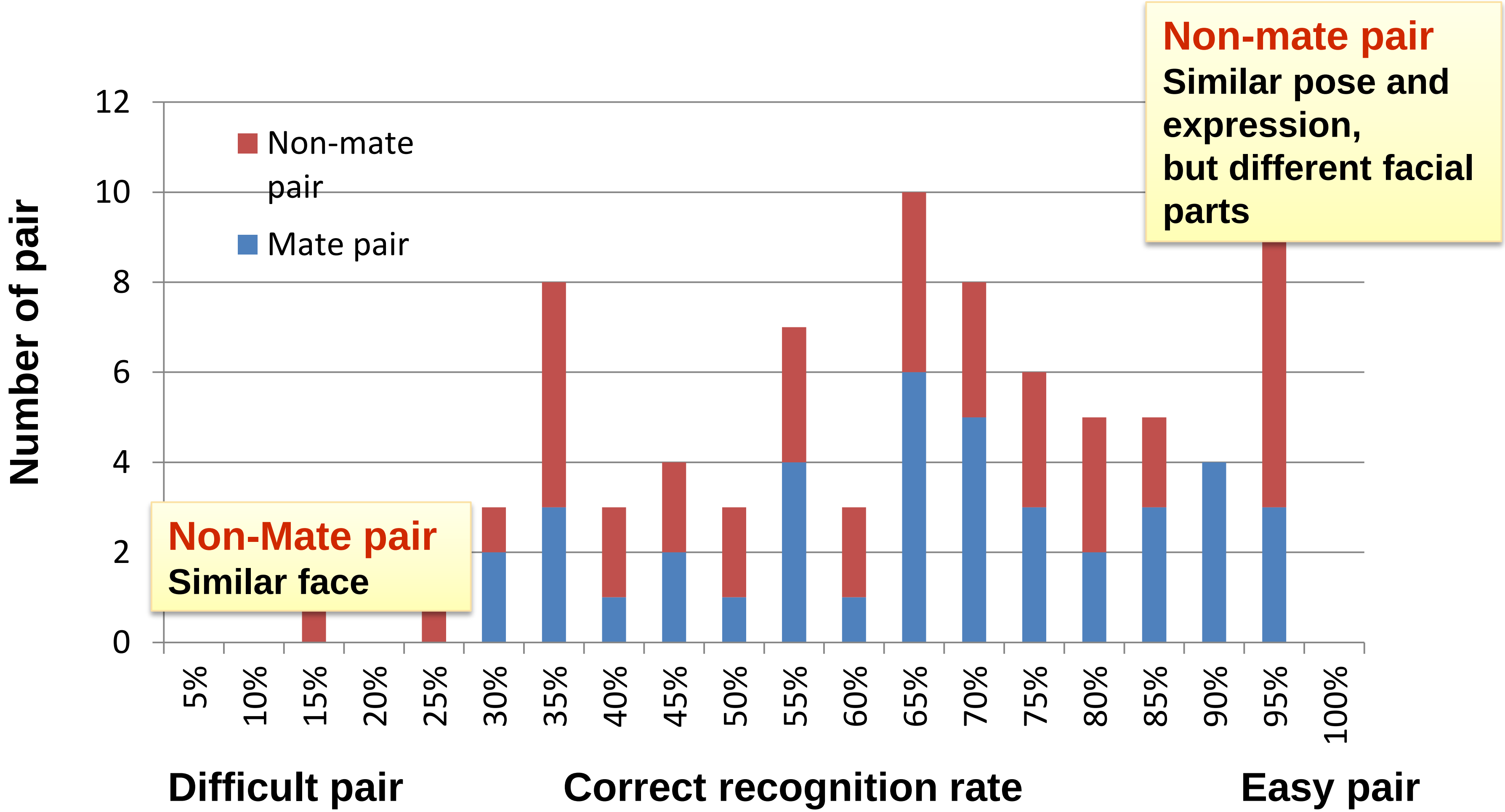


	Num of correct pair	FAR	FRR
Adult [FAR=0.1%]	13.0±1.3	27%	44%
Child [FAR=0.1%]	13.4±1.5	26%	41%
Adult [FAR=0.001%]	13.0±2.2	41%	29%
Child [FAR=0.001%]	11.9±2.1	42%	39%

Human brain may assist to discriminate mate or non-mate pair, but reliability is low

Experimental result by human

Among mate and nonmate pairs, which pair is easy or difficult to discriminate?



Accuracy in the application system of face recognition

Advantage of Face Recognition System

① Face can be recognized **at a distance**

- Hands-free recognition - surveillance application

② **No need for special devices**

- uses tablets, smartphones, and other mobile devices

③ Matched face images can be **confirmed by human**

- human can check the result in case of failure to match

Accuracy is relatively low compared with other types of biometrics

"Improving recognition accuracy" is a key point of face recognition

Introduction of application examples

- Government sector application

1) Hong Kong Immigration System(2004)

2) Boston Marathon Bombings Suspects

- Privatized sector application

3) Terracotta Army

4) Great East Japan Earthquake

Application example (1) Hong Kong Immigration System

'FACE Recognition System (FACES)' to verify the identity of suspects, started operation in 2004

Application Category of the 7th IT Excellence Awards (ImmD) ***Judge's Comment***

- ◆**Over 75% similarity from over 200,000 suspect records in just one second**
- ◆**Over 100 suspects have been successfully detected.**

Identification performance, Aging change, Ethnicity

Application example(1) Hong Kong Immigration System

Automated border control system

- Drive-through face and fingerprint recognition system
- Checkpoints on the Hong Kong - China border, started in 2007

When the driver is recognized, gate opens

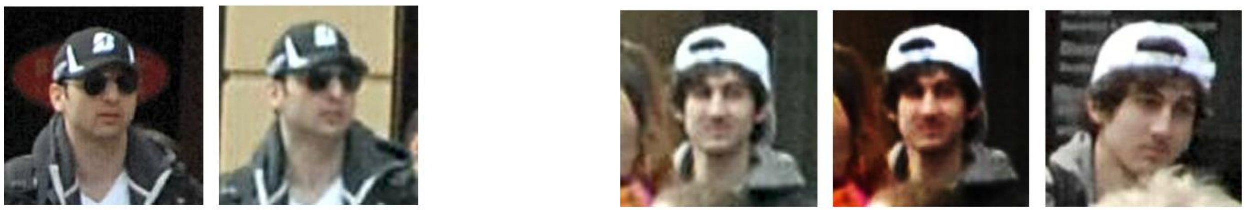
Device moves up and down according to truck seat height

Illumination change

Application example (2)
Facial Recognition Using the Boston Marathon Bombings Suspects

Klontz and Jain, "A Case Study on Unconstrained Facial Recognition Using the Boston Marathon" Technical Report MSU-CSE-13-4 (2013/5/29)

FBI released images of 2 suspects



1a *1b* *2a* *2b* *2c*

Verify identification performance

Suspects arrested in 88 hours

Application example (2)
Facial Recognition Using the Boston Marathon Bombings Suspects

Suspect 1

1a 1b

Suspect 2

2a 2b 2c

Captured image from Video

Query Images

1toN
Matching

Suspect1

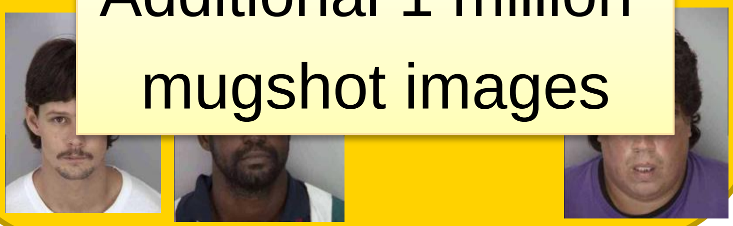
1x 1y 1z

Suspect2

2x 2y 2z

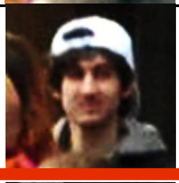
+

Additional 1 million mugshot images



Enrolled Images

Search Result (NEC): ranking (database size =1 million)

Query Image	No filtering	Filterd by age and gender
2a 	213	19
2b 	260	30
2c 	1	1
1a 	12,446	1746
1b 	236,343	42,827

- ◆**Face recognition is useful tool for criminal application**
- ◆**Difficult to identify wearing sunglasses**

Application example (4)
Terra-cotta soldier's face recognition

- **Sculptures of the first emperor of China's army**
- **Buried over 8,000 soldier sculptures**
- **Analyzed sculpture faces using face recognition software**



http://en.wikipedia.org/wiki/Terracotta_Army
http://www.youtube.com/watch?v=LoCr9AEYpCo

Application example (4)

Terra-cotta soldier's face recognition (TV program)

Input feature points manually: eyes, nose, mouth

http://www.youtube.com/watch?v=LoCr9AEYpCo

Application example (4)
Terra-cotta soldier's face recognition (TV program)

Examples of similar pairs

All of them are unique

http://www.youtube.com/watch?v=LoCr9AEYpCo

Save the memory project in the Great East Japan Earthquake

Great East Japan Earthquake

- ***11th March 2011***
- ***Magnitude 9.0***
- **20,000 dead and missing people**

Tsunami and Nuclear accidents

Application example (7)
Save the memory project in the Great East Japan Earthquake

“Save the Memory Project” (collaboration of Ricoh and NEC)

- Earthquake disaster reconstruction project
- Rescue team collect albums and photographs
- Volunteer washed and digitized photographs
- Return photographs to the owner

http://www.ricoh.co.jp/release/2012/0808_1.html

Application example (5)

Save the memory project in the Great East Japan Earthquake

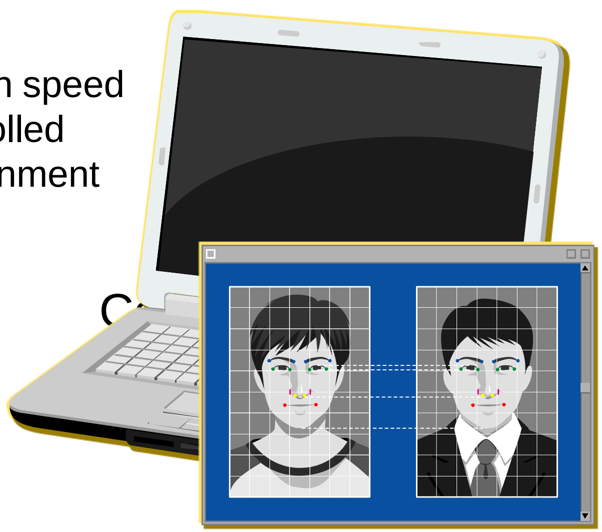
Face recognition is used to search among 150,000 photographs

Face recognition system assisted in returning 12% of the photographs to the owner

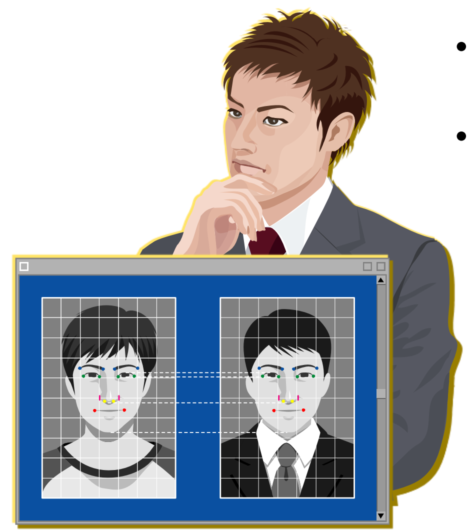
Summary

- **Introduced face recognition technology : 1) algorithms, 2) evaluation results and 3) applications**
- **Face recognition accuracy has improved rapidly in these 20 years**
- **Next 10 years, accuracy will improve more and more beyond limit of human face recognition ability**

- Search speed
- Controlled environment



VS

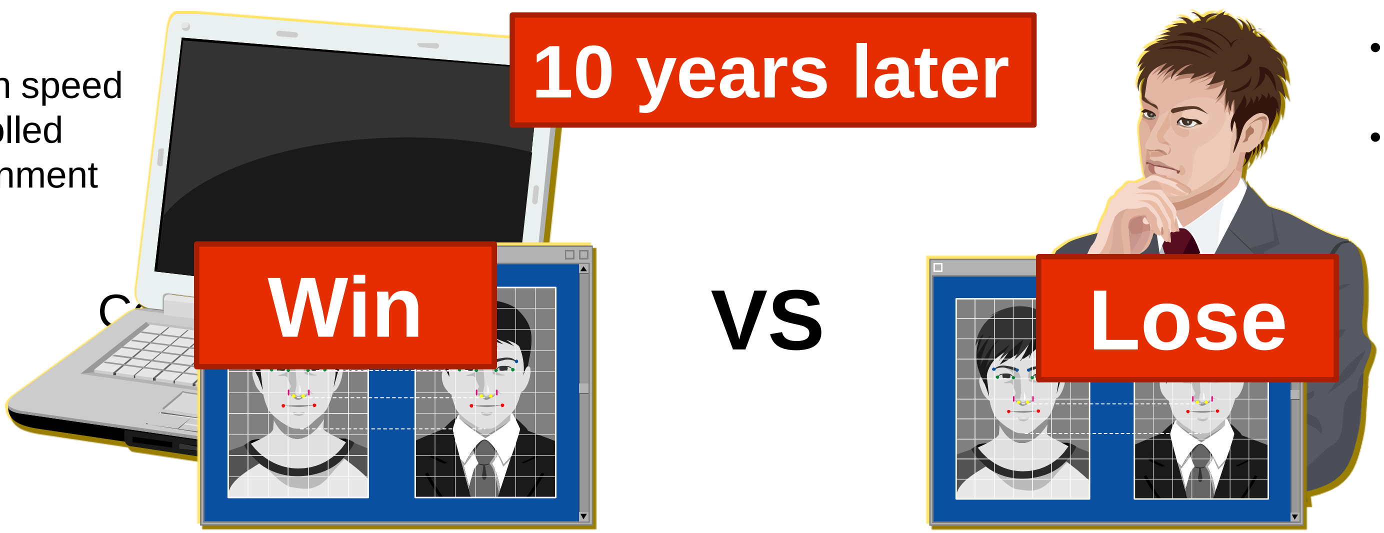


- Uncontrolled environment
- Total judgment using other clue

Summary

- Introduced face recognition technology : 1) algorithms, 2) evaluation results and 3) applications
- Face recognition accuracy has improved rapidly in these 20 years
- Next 10 years, accuracy will be more and more improved beyond limit of human ability

- Search speed
- Controlled environment



- Uncontrolled environment
- Total judgment using other clue

Thank you for your attention