IAPR/IEEE WINTER SCHOOL ON BIOMETRICS

9-13 JANUARY 2017 | HONG KONG BAPTIST UNIVERSITY, HONG KONG

Biometric Data Analysis

Tieniu Tan

Center for Research on Intelligent Perception and Computing (CRIPAC)

National Laboratory of Pattern Recognition (NLPR)

Chinese Academy of Sciences' Institute of Automation (CASIA)

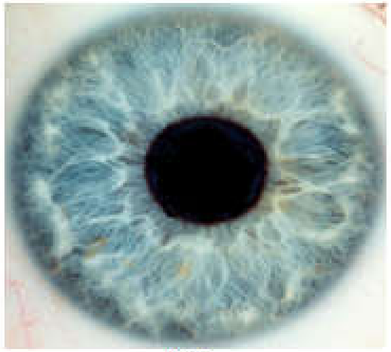
January 9, 2017

Outline

- Preamble
- Identity from Biometric Data
- Gender from Biometric Data
- Ethnicity from Biometric Data
- Age and Affect from Biometric Data
- Conclusions

Biometric Data

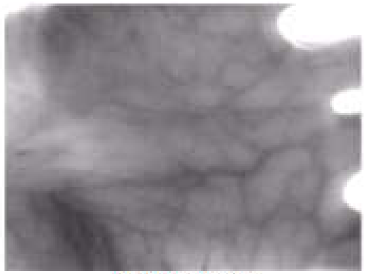
Physiological Modalities

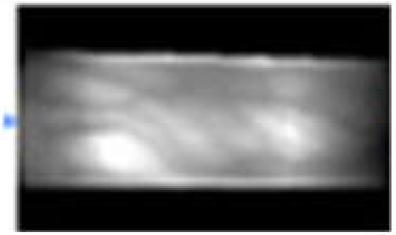

Iris


Face


Fingerprint

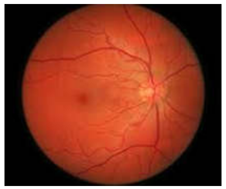

Palmprint

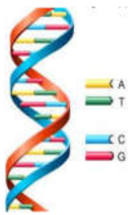

Palm vein


Finger vein


Hand geometry

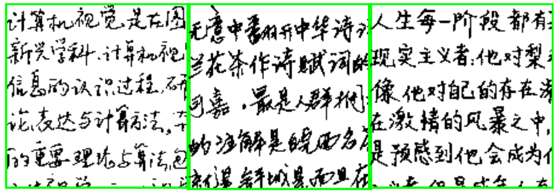

Ear

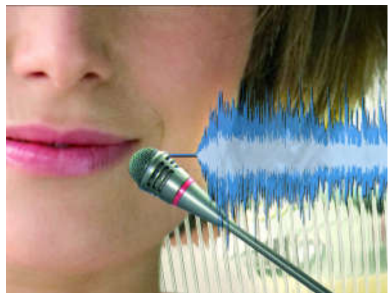

Retina


DNA

Behavioral Modalities


Gait


Handwriting


Voiceprint

Ubiquity of Biometric Data

Mobile phones are widely available sensors for multi-modal biometric data

Internet and social media

TV and video games

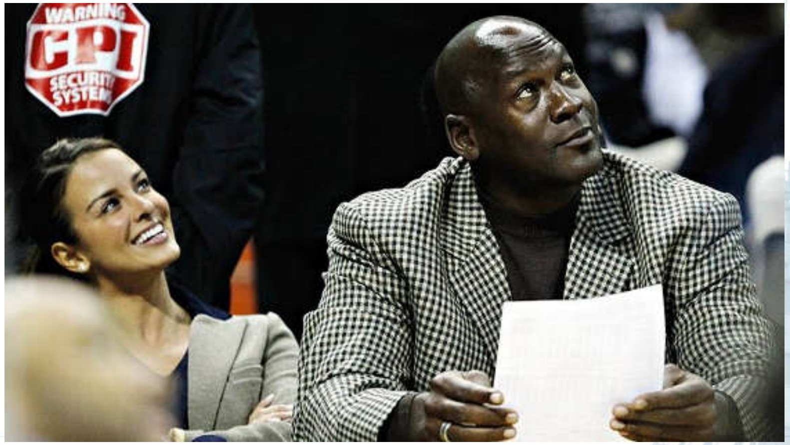
Wearable devices

CCTV cameras

Passport and ID card

Information from Biometric Data

What demographic and affective information can be derived from this face image?



Identity	Rose	Jordan
Gender	Female	Male
Ethnicity	White	Black
Age	27	45
Affect	Happy	Surprised

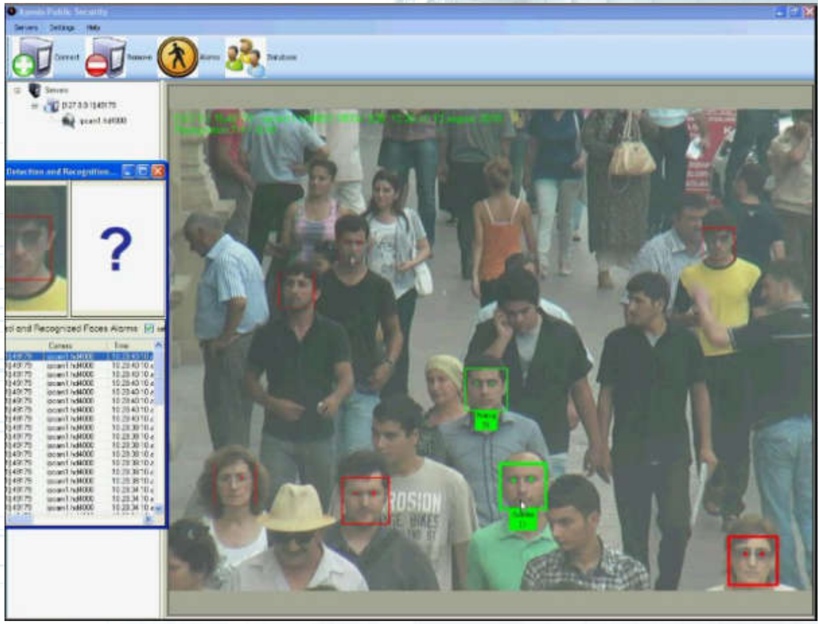
How to determine such information from biometric data?

中国科学院
自动化研究所
INSTITUTE OF AUTOMATION
CHINESE ACADEMY OF SCIENCES

Biometric Data Analysis

—Applications —

Human-Computer (Robot) Interaction



Intelligent visual surveillance

Outline

- Preamble
- Identity from Biometric Data
- Gender from Biometric Data
- Ethnicity from Biometric Data
- Age and Affect from Biometric Data
- Conclusions

Identity from Biometric Data

Fingerprint recognition for mobile authentication

Face recognition for border control

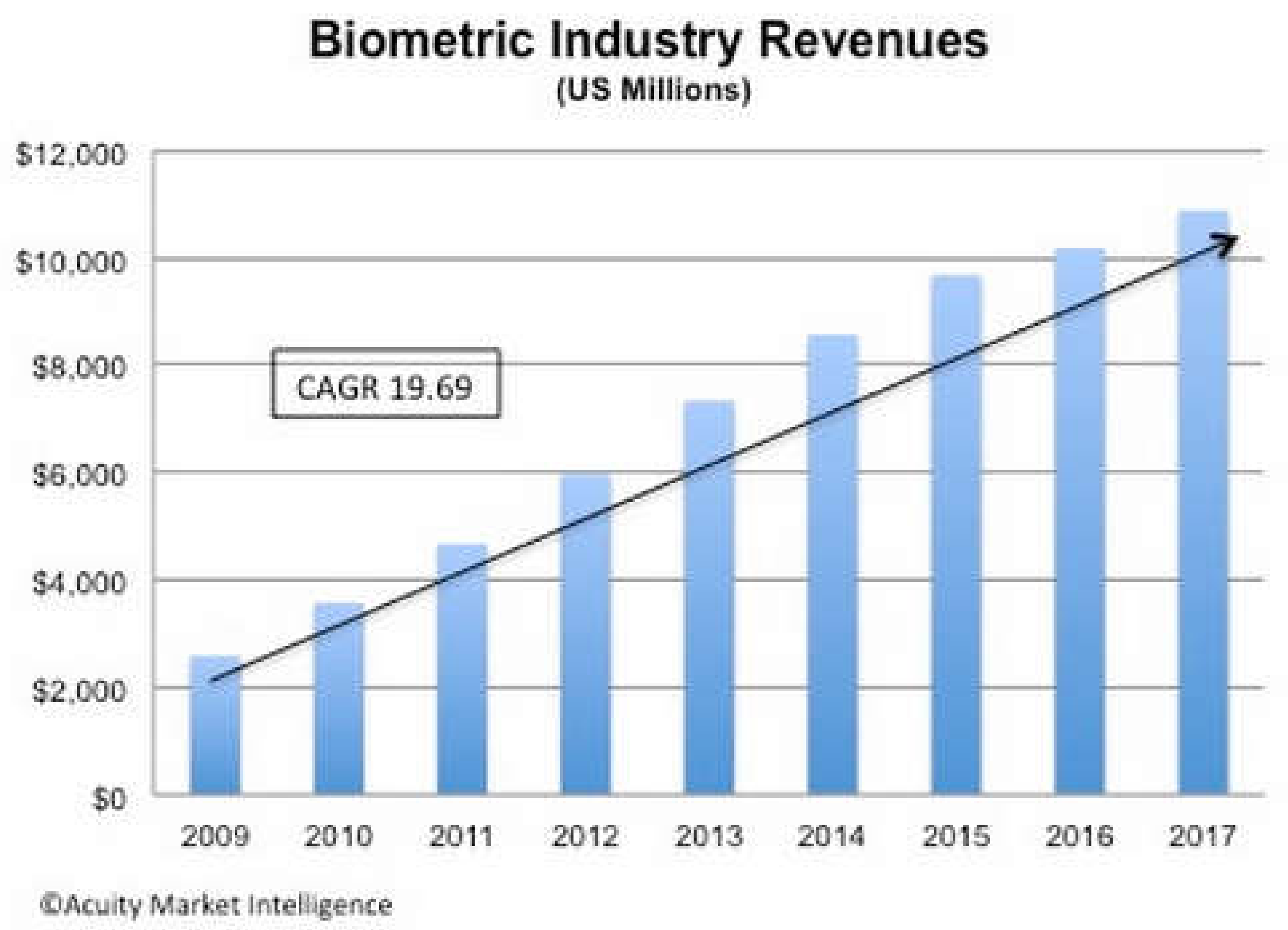
Iris recognition for coal miner identification

Finger vein recognition for ATM authentication

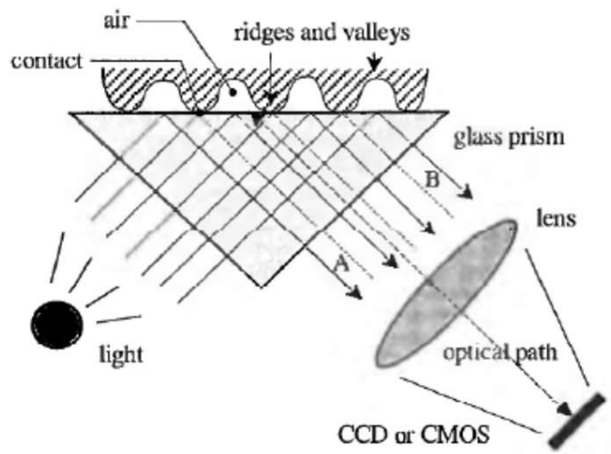
Voiceprint recognition for payment

Signature verification for credit card security

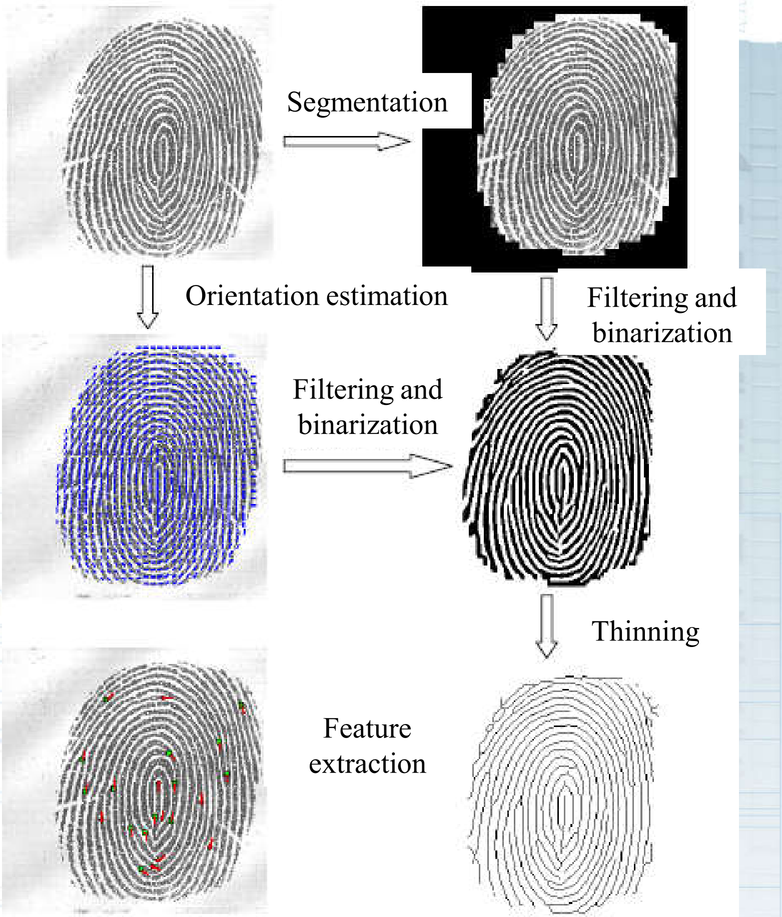
Fast growing market of biometric recognition



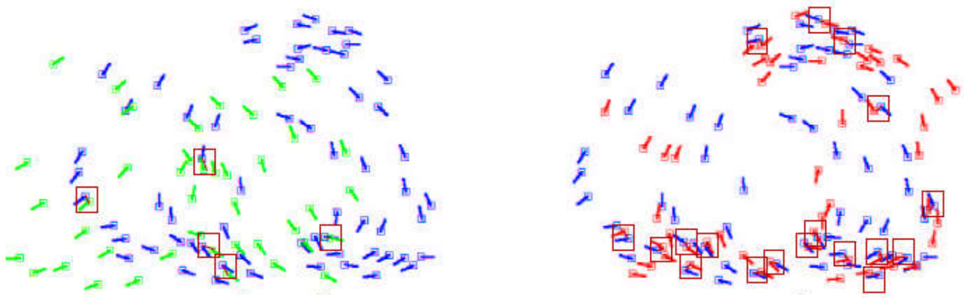
Fingerprint Recognition



Imaging

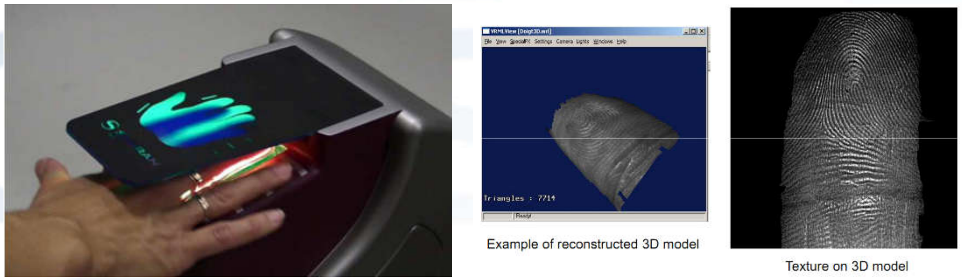


Preprocessing and feature extraction

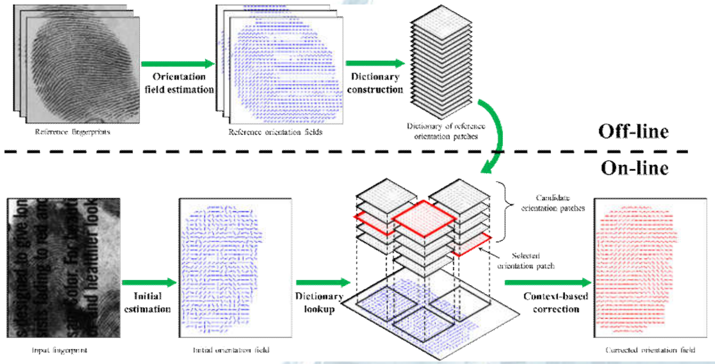


Minutiae matching

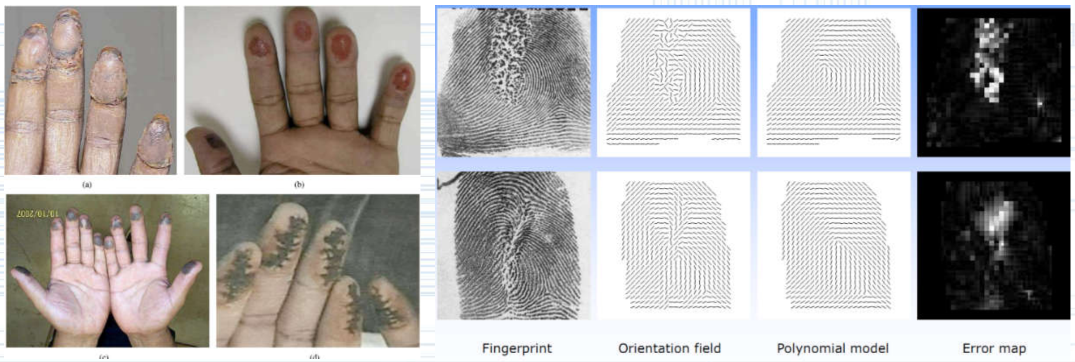
New methods in fingerprint recognition



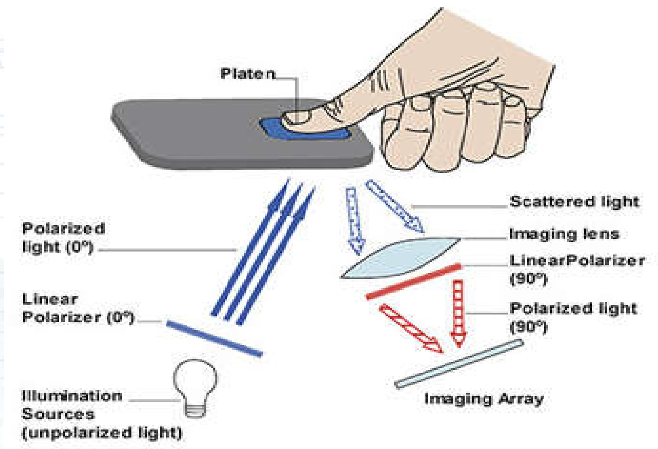
Touchless 3D fingerprint (SAFRAN Morph)



Latent fingerprint recognition (Tsinghua)

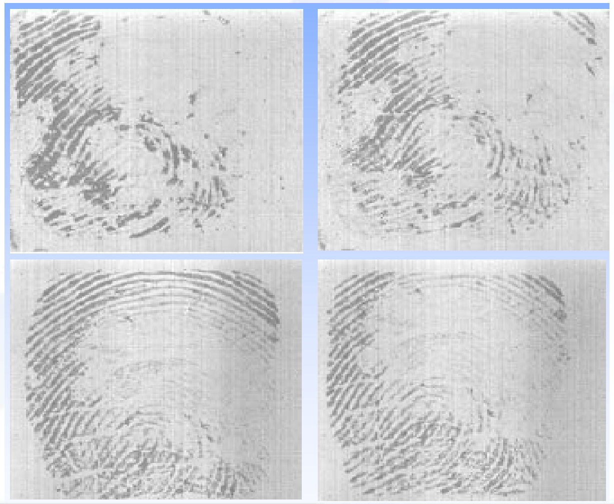


Detection and recognition of altered fingerprint (MSU)

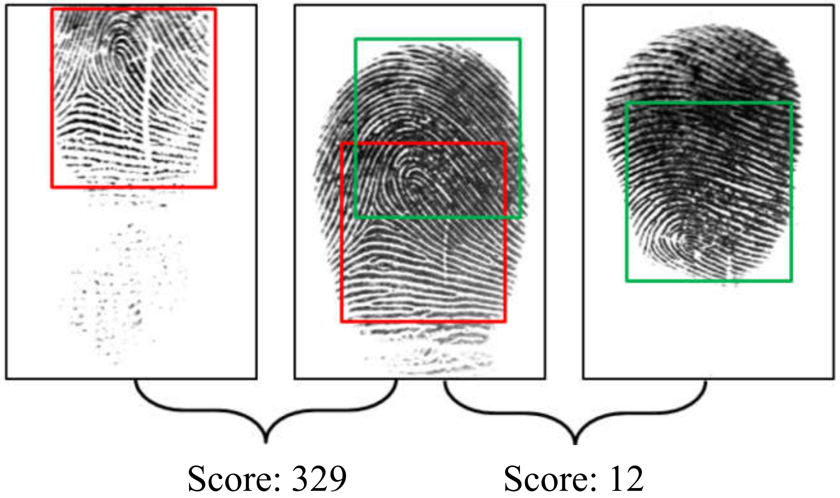


Multispectral Imaging for anti-spoofing (Lumidigm)

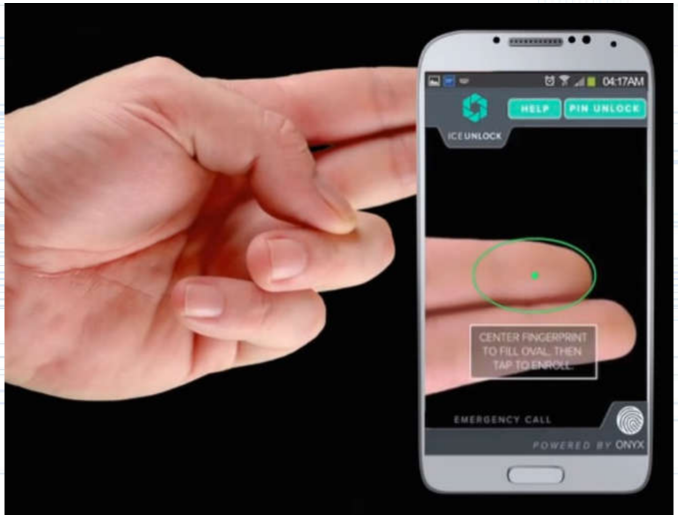
Open Problems of Fingerprint Recognition



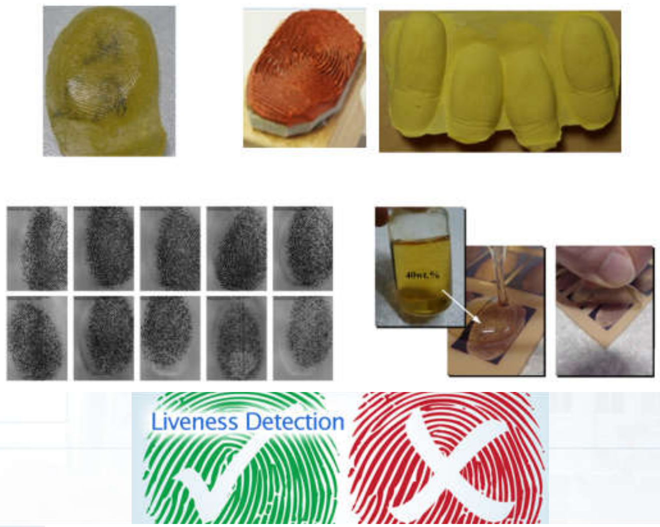
Latent fingerprint images



Distorted fingerprint images

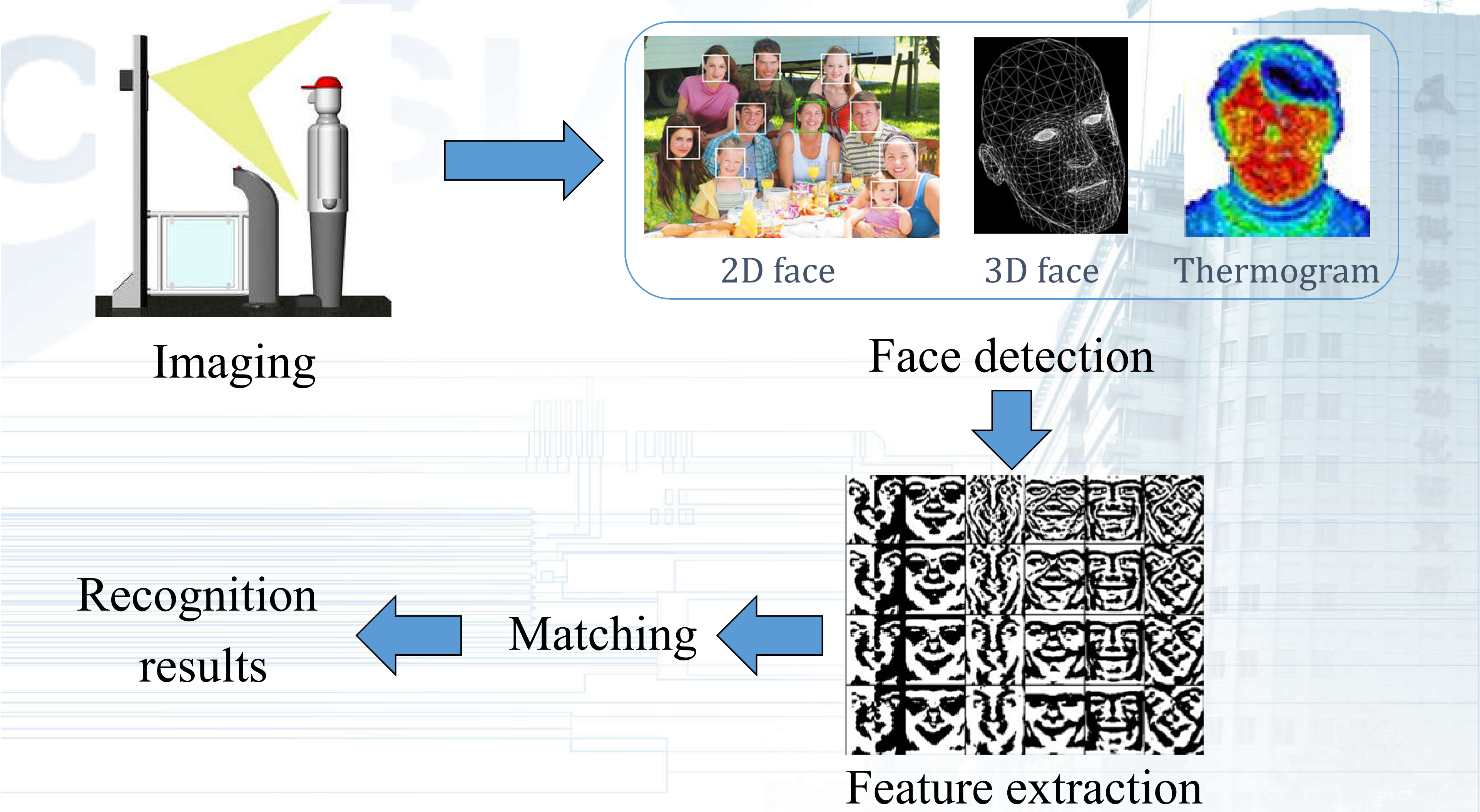


Touchless fingerprint recognition



Fingerprint liveness detection

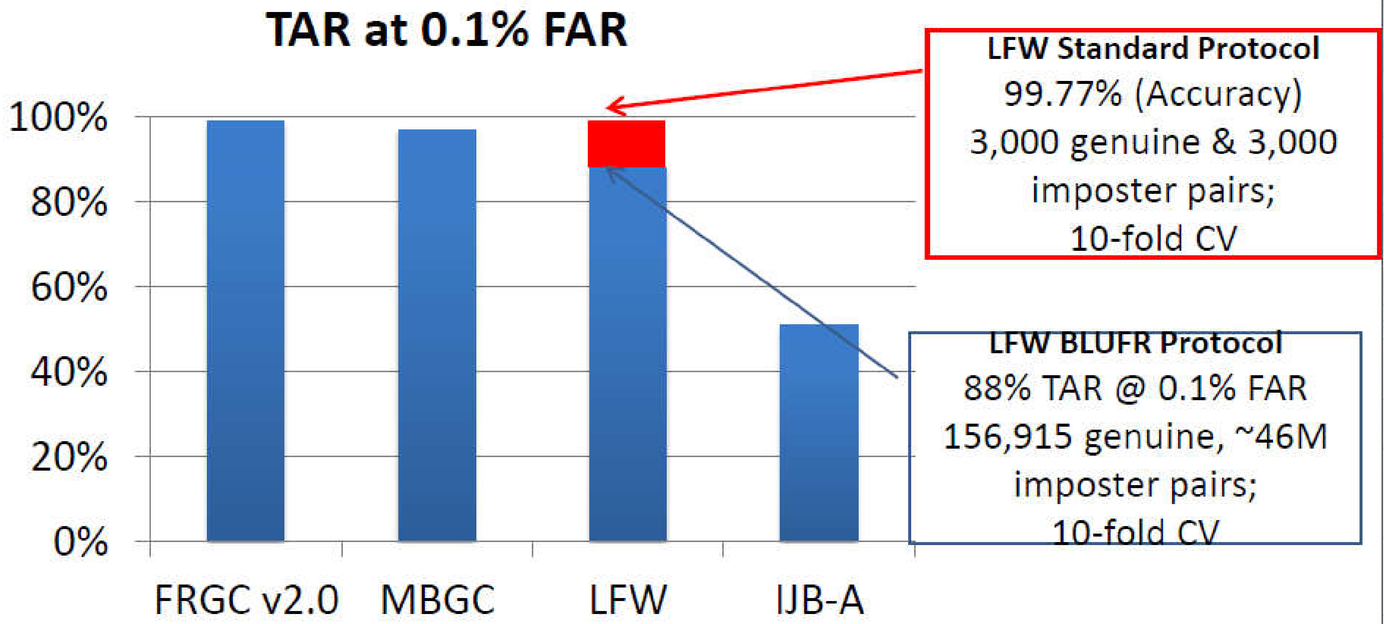
Face Recognition



Popular methods: Gabor/LBP/Ordinal measures/Sparse representations/Deep learning

State-of-the-Art Performance of Face Recognition

FRGC v2.0 (2006) MBGC (2010) LFW (2007) IJB-A (2015)

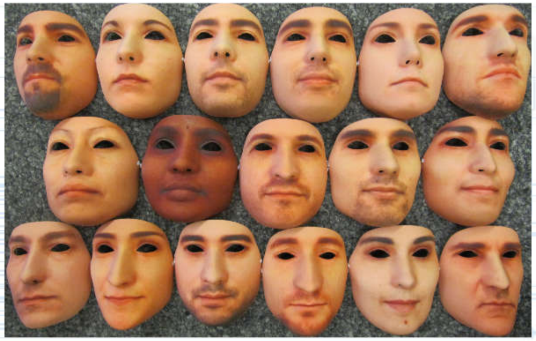


D. Wang, C. Otto and A. K. Jain, "Face Search at Scale: 80 Million Gallery", arXiv, July 28, 2015

Open Problems of Face Recognition

PIE (Pose, Illumination, Expression)

Face recognition in surveillance

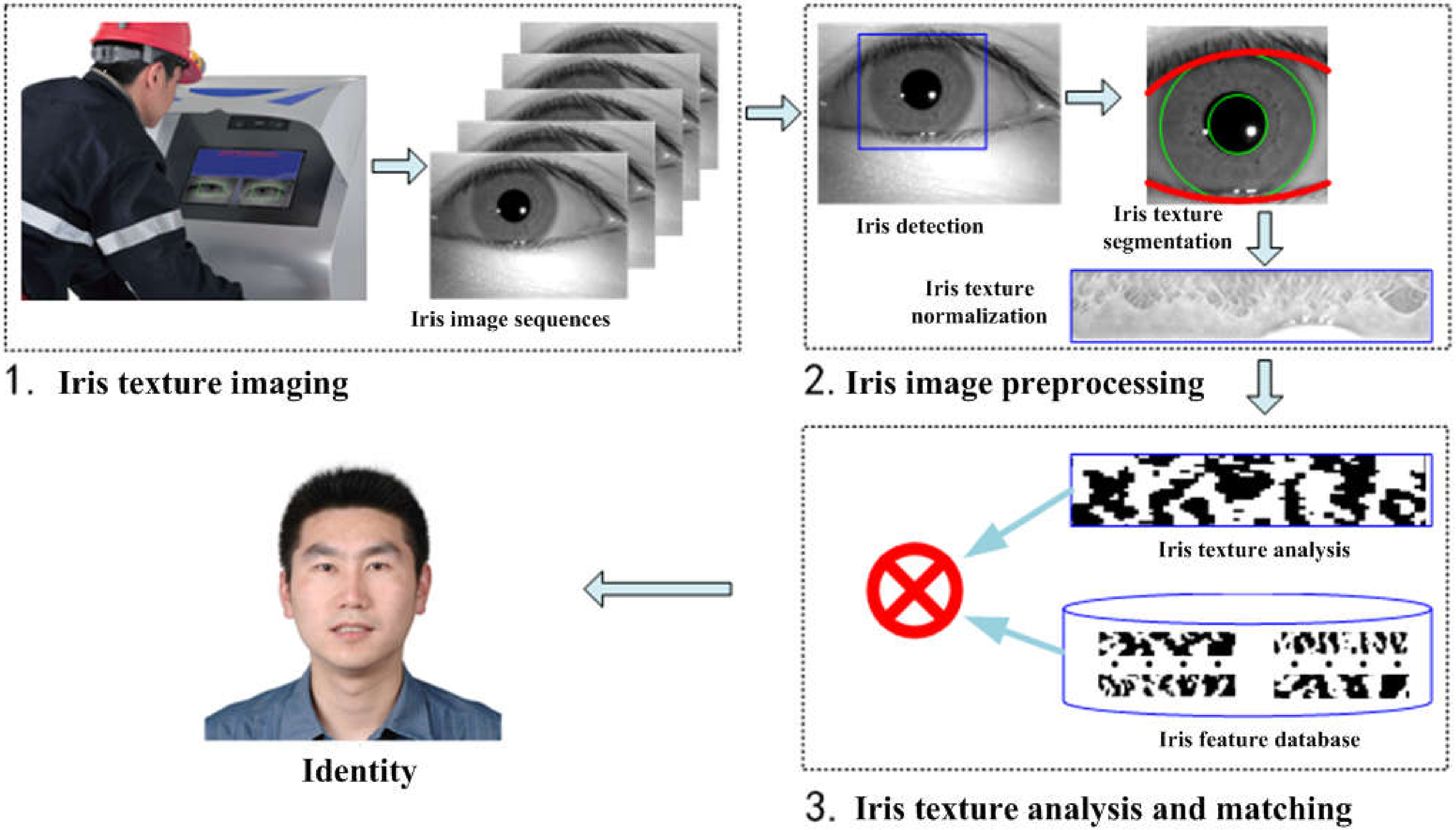


Spoof-attack

Face recognition

Facial aging

Iris Recognition



Iris Recognition at CASIA

1999

2000

2001

2004

2005

2007

2008

2009

2014

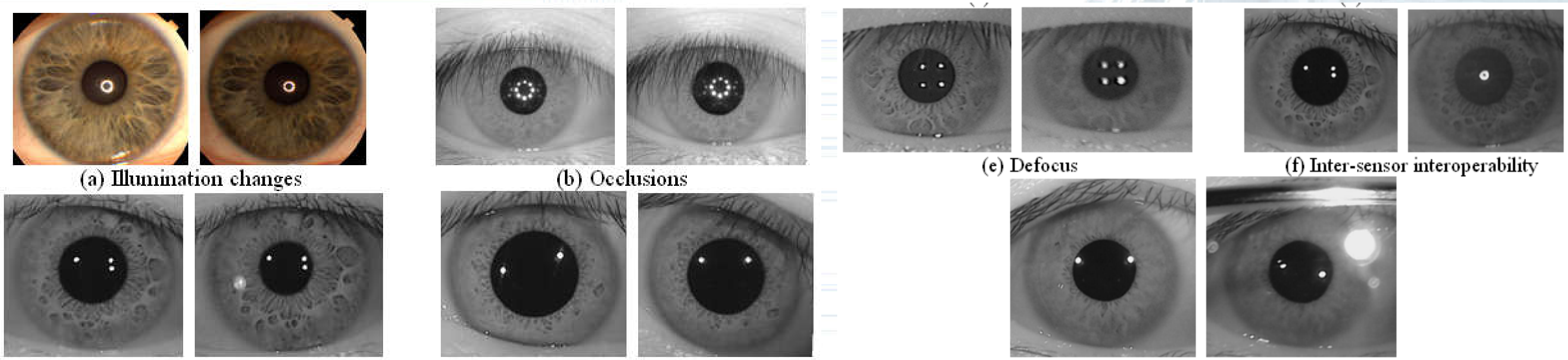
2015

Recent Progress of Iris Recognition

Open Problems of Iris Recognition

Less or unconstrained iris image acquisition

Forensic applications



(a) Illumination changes

(b) Occlusions

(e) Defocus

(f) Inter-sensor interoperability

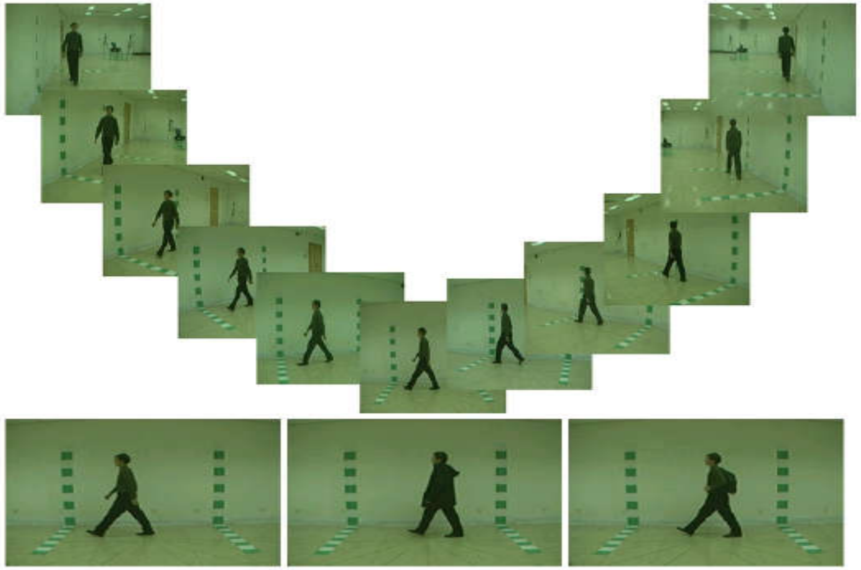
(c) Deformation

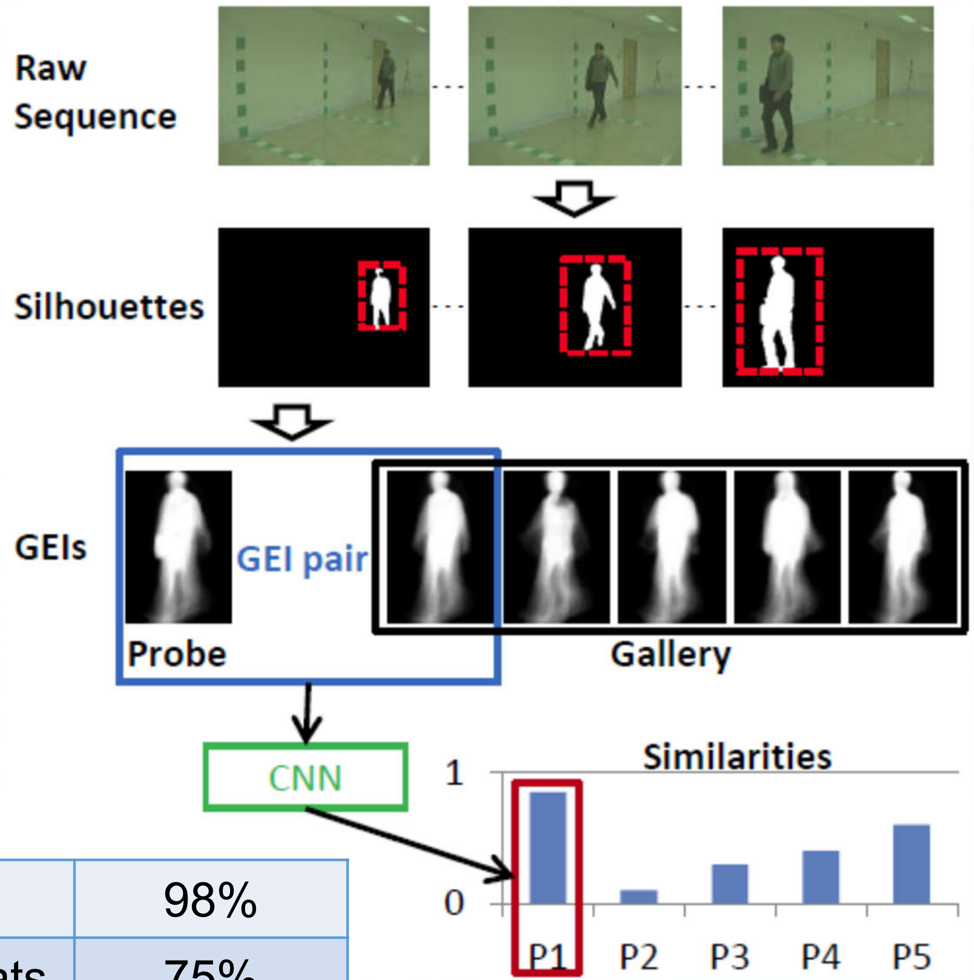
(d) Rotation

(g) Eyeglasses

Poor quality iris images

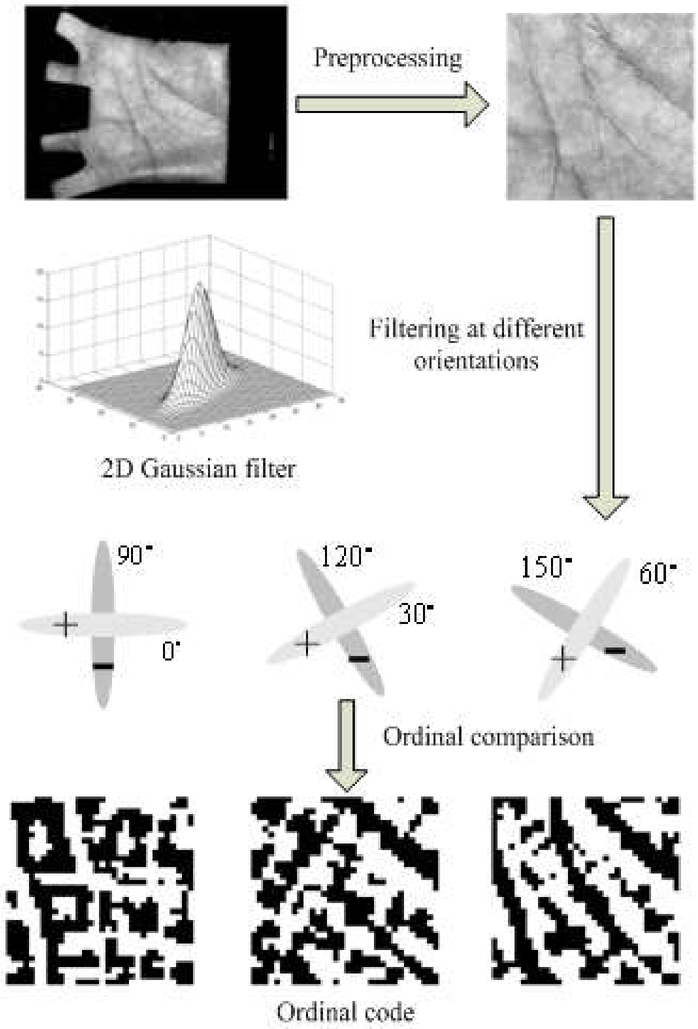
Multi-view Gait Recognition





CASIA-B Gait	cross-view	98%
	cross-view and with coats	75%
	cross-view and with bags	90%
OU-ISIR	cross-view	91%

Ordinal Measures Based Palmprint Recognition



Ear Biometrics

BBC | Menu | Search

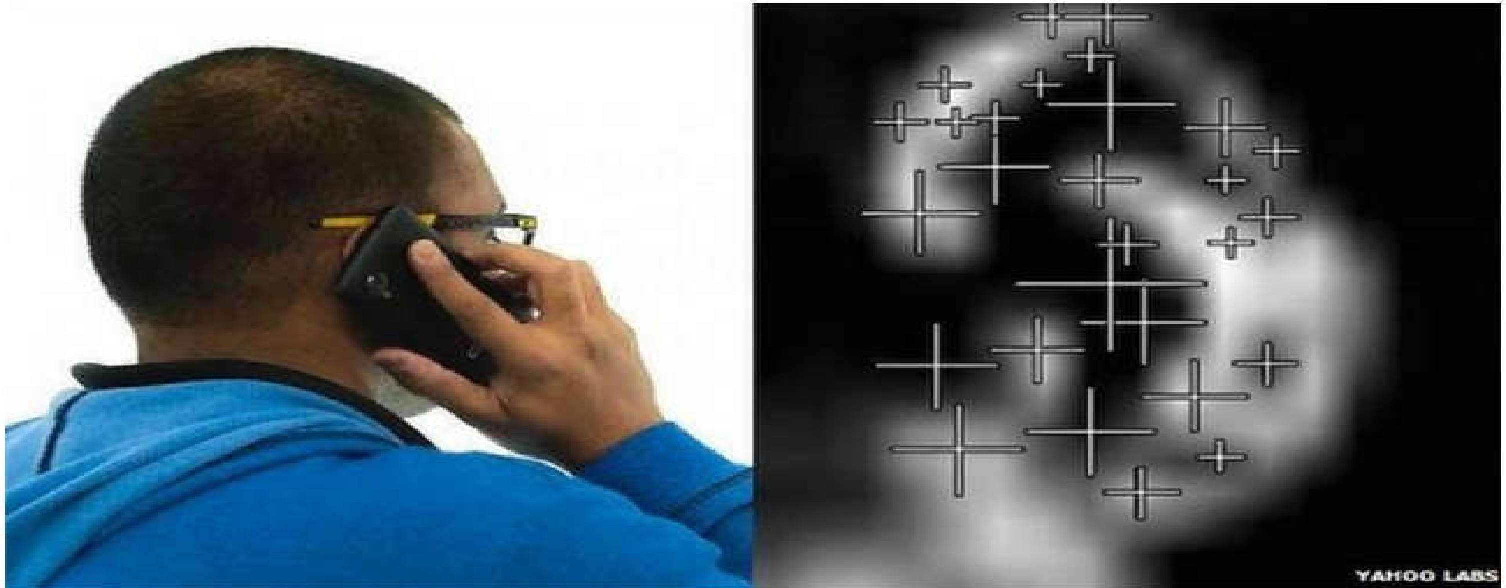
NEWS

Home | Video | World | Asia | UK | Business | Tech | Science | Magazine | Entertainment & Arts | Health | More

Technology

Yahoo tests ear-based smartphone identification system

28 April 2015 | Technology



The system identifies users based on the shape of their ears

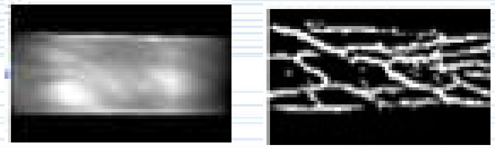
Hand Vein Patterns for Biometric Recognition

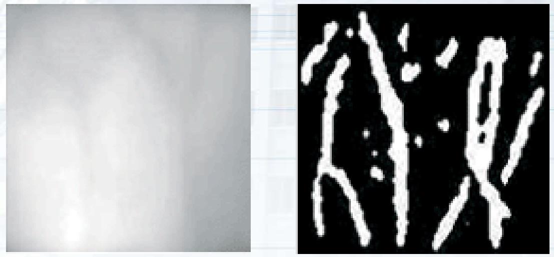
Unique, stable and secure biometric patterns underneath the skin surface

Finger vein

Palm vein

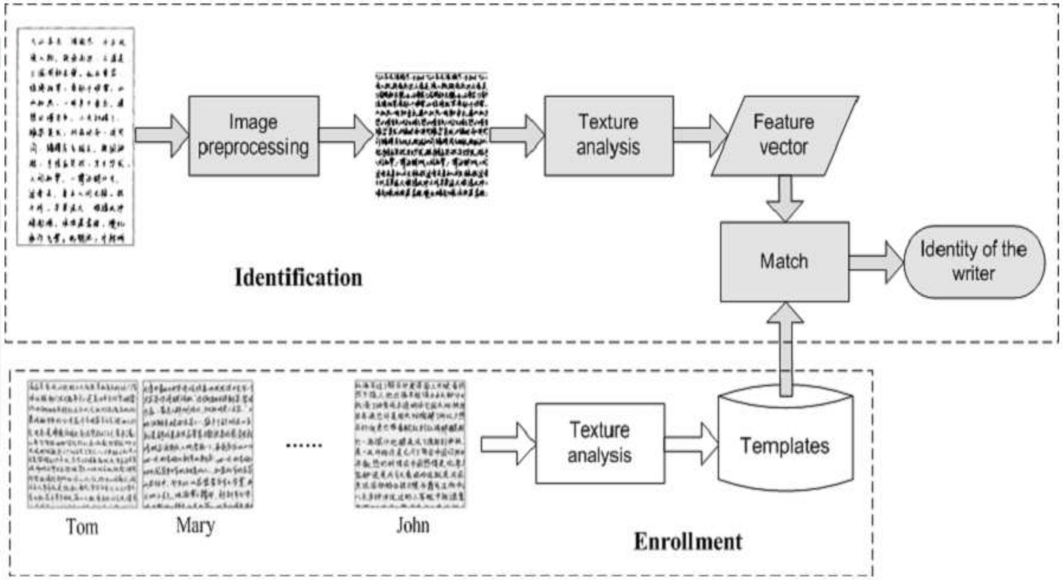
Hand vascular pattern



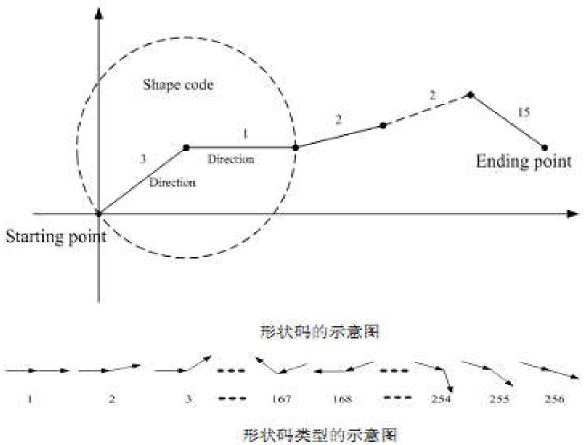


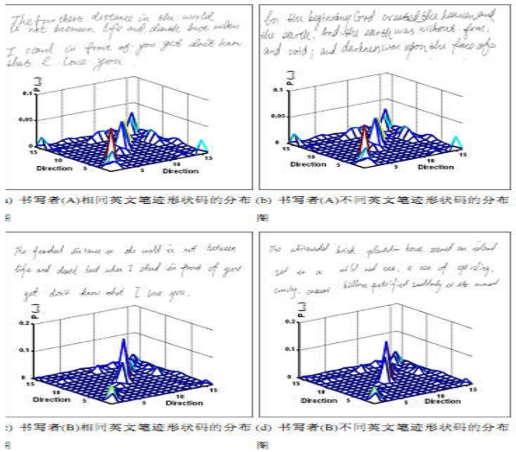
Handwriting Biometrics

www.ia.ac.cn



Handwriting texture analysis for writer identification





Statistical analysis of stroke shape features for writer identification

Challenges of Biometric Identification

50 Years of Biometric Research:

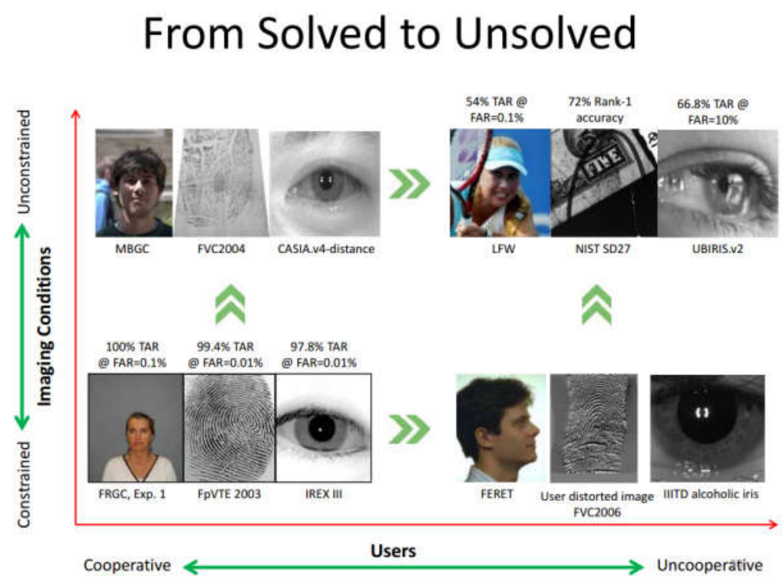
Almost

~~The~~ Solved, The Unsolved, and The Unexplored

Anil Jain
Michigan State University

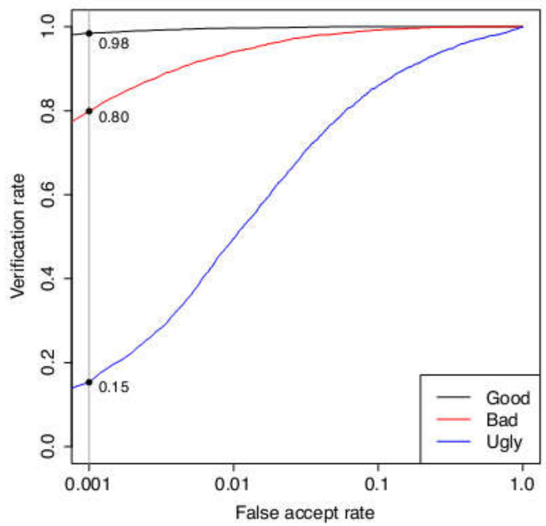
June 5, 2013

Keynote Talk Delivered at the International Conf. on Biometrics, Madrid, Spain, June 5, 2013



Jonathon Phillips
NIST

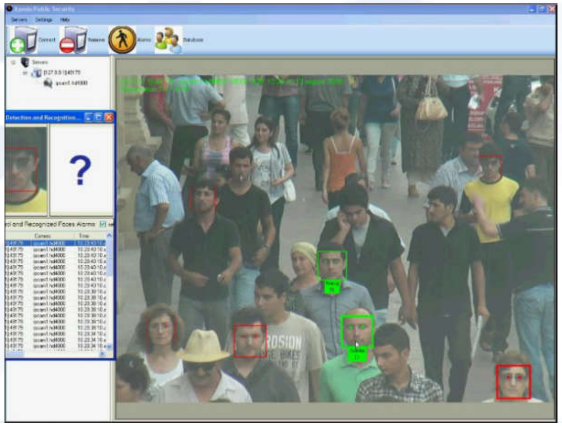
An Introduction to the Good, the Bad, & the Ugly Face Recognition Challenge Problem (FG2011)



Outline

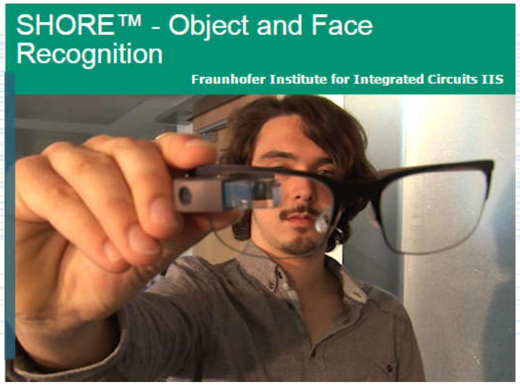
- Preamble
- Identity from Biometric Data
- Gender from Biometric Data
- Ethnicity from Biometric Data
- Age and Affect from Biometric Data
- Conclusions

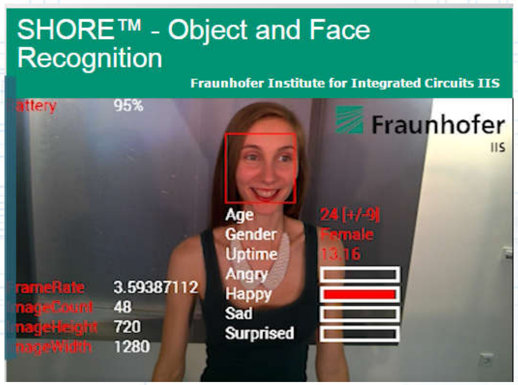
Gender from Biometric Data



Intelligent visual surveillance

Smart vending machine

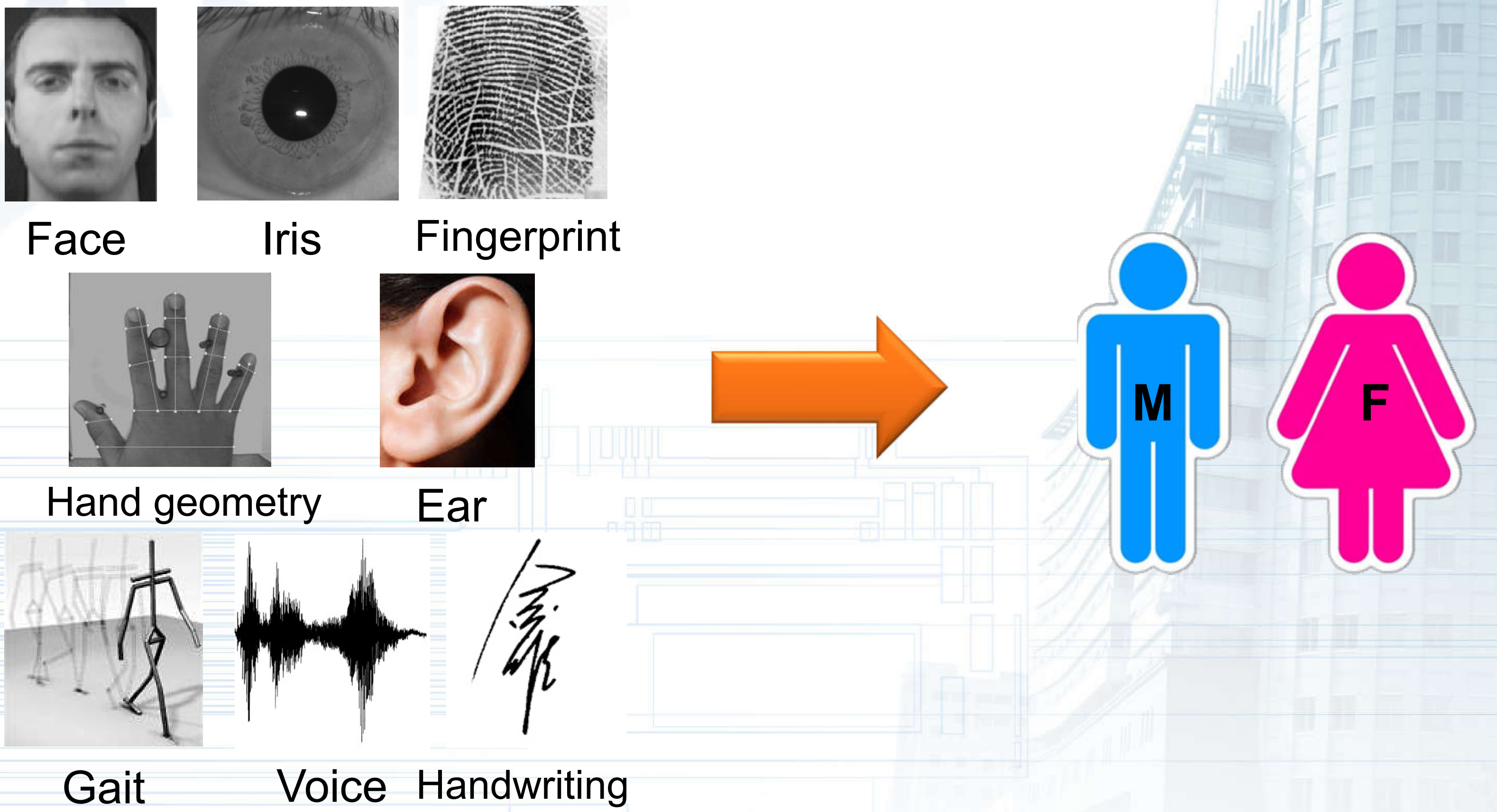




Personal attributes labeling in wearable devices

Gender specific beautification

Main Biometric Modalities for Gender Estimation



Gender from Biometric Data

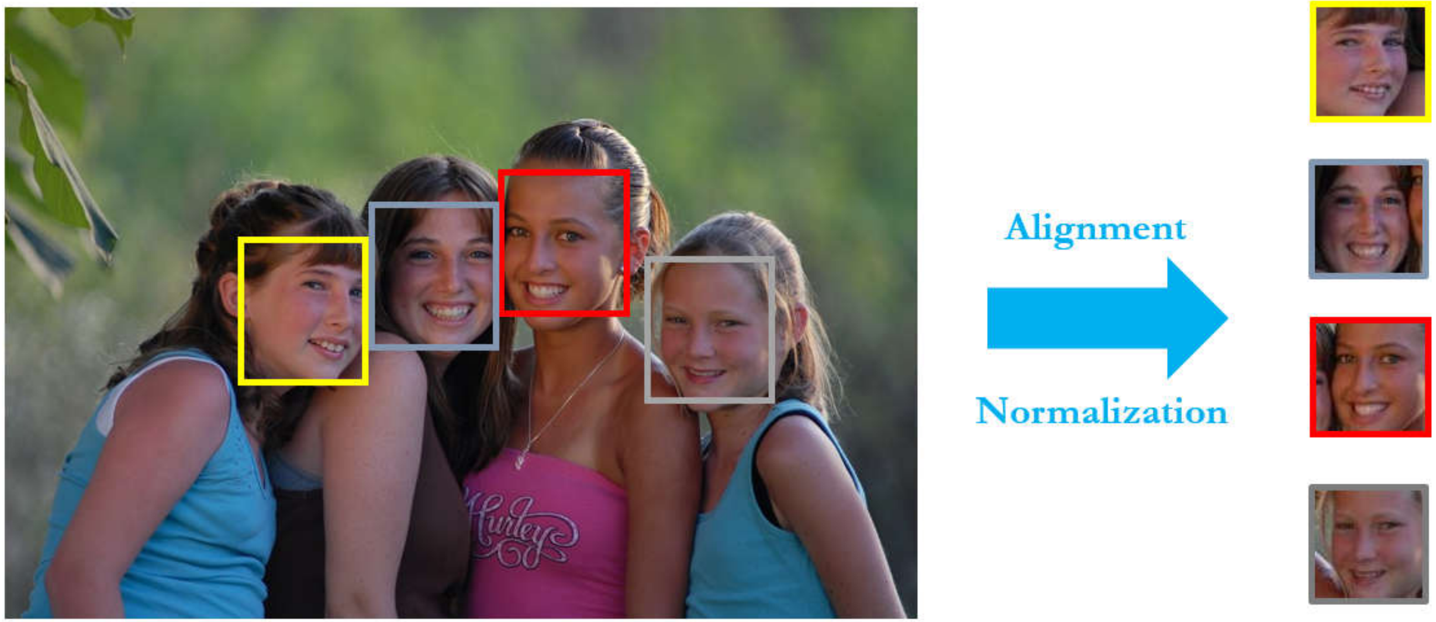
— Face —

- **What are the differences between adult male and female faces? (from human perception)**
  - In general, the nose and nasopharynx are larger in men than in women (Enlow, 1982).
  - Men in general have more prominent brows, more sloping foreheads, and more-deep-set eyes than women (Enlow, 1982).
  - Women generally have less facial hair, not only in the beard region, but also in the eyebrows (Shepherd, 1989).
  - Women appear to have fuller cheeks than men (Shepherd, 1989).

Gender from Biometric Data

— Face —

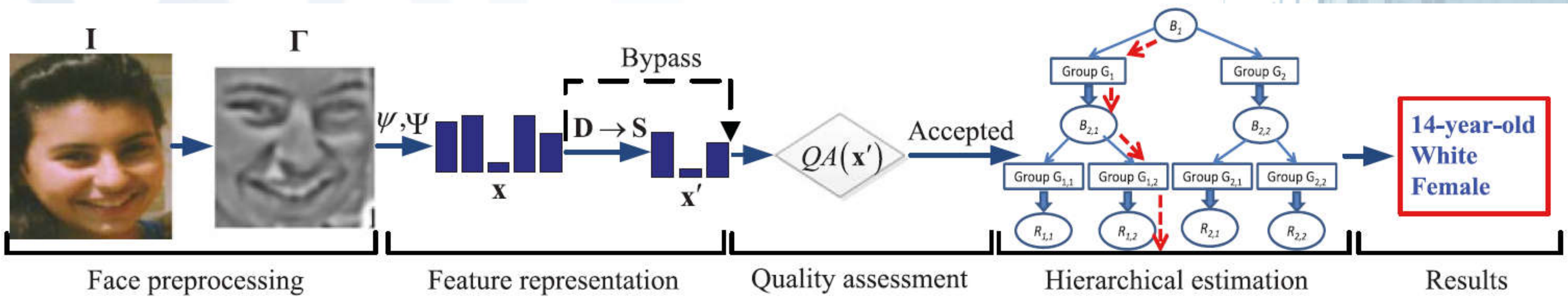
Pipeline



Face Detection → Pre-processing → Feature Extraction → Classifier Training

Gender from Biometric Data

— Face —



Input: single face image

Features: biologically inspired features (BIF)

Classifier: SVM

Output: age, ethnicity, gender

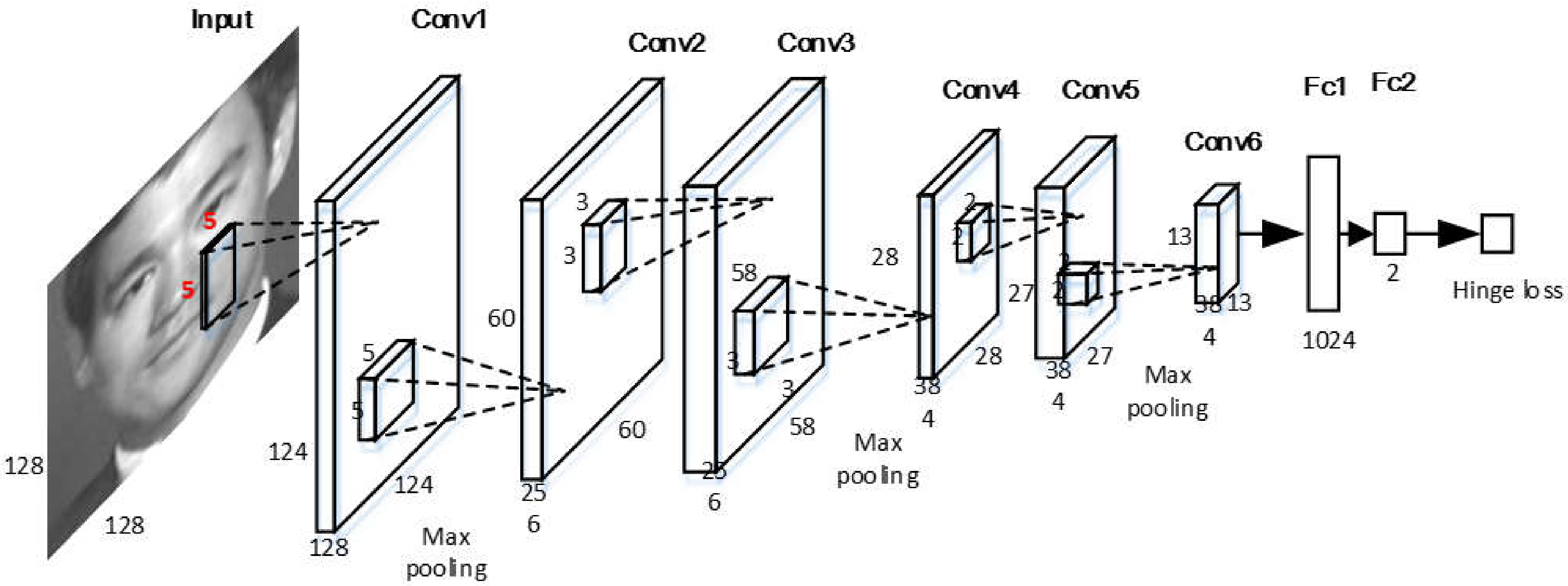
Testing database	Classification task	Proposed - Intra-DB	Proposed - Cross-DB
MORPH II	Gender	97.6	90.5
	Race	99.1	95.3
PCSO	Gender	97.1	88.6
	Race	98.7	93.0

Hu Han, Charles Otto, Xiaoming Liu and Anil K. Jain, “Demographic Estimation from Face Images: Human vs. Machine Performance”, IEEE Trans. PAMI, vol.37, no.6, pp.1148-1161, 2015.

Gender from Biometric Data

— Face —

Our work



Dataset: LFW + images downloaded from the Internet
Training: 11,889 female images + 15,042 male images
Testing: 3,000 female images + 3,000 male images
Correct Classification Rate (CCR): 97.5%

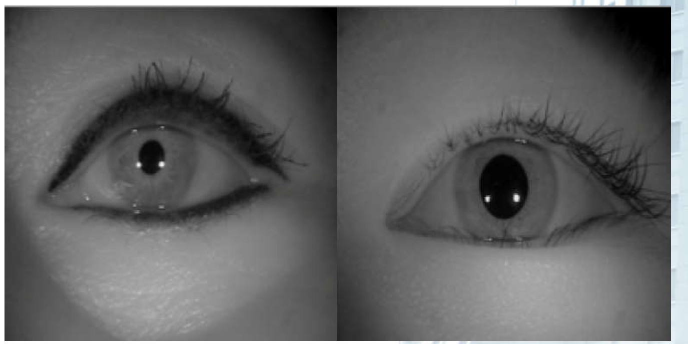
中国科学院
自动化研究所
INSTITUTE OF AUTOMATION
CHINESE ACADEMY OF SCIENCES

Gender from Biometric Data

www.ia.ac.cn

— Iris —

Male



Female

- **Features**
  - **Geometric features (e.g. inter-landmark distance, area, ratio)**
  - **Texture features (e.g. mean and variance of pixel intensity, LBP, wavelet features)**
  - **Statistical features (e.g. statistical distributions of filter response)**
- **Classifiers**
  - **C4.5 tree, SMO, Random Forest, SVM, Naïve Bayes, etc.**

Gender from Biometric Data

— Iris —

Authors	Features	Classifiers	Datasets	Gender(the number of samples)	CCR*
Thomaset *al.* [1]	Geometric and texture features	C4.5 decision tree	unpublished	Female(about 28000)/Male(about 28000)	close to 80%
Lagree *et al.* [2]	Texture features	SMO in WEKA	unpublished	Female(1200)/Male(1200)	90.58%
Bansal*et al.* [3]	Statistical and wavelet features	SVM	unpublished	Female/Male (total 400)	85.68%
Tapia *et al.* [4]	Uniform LBP	SVM	UND	Female/Male(total 3000)	91.33%

* "CCR" means correct classification rate.

[1] Vince Thomas, Nitesh V. Chawla, Kevin W. Bowyer, and Patrick J. Flynn, "Learning to Predict Gender from Iris Images", in Proc. IEEE International Conference on Biometrics: Theory, Applications, and Systems, pp.1–5, 2007.

[2] Stephen Lagree and Kevin W. Bowyer, "Predicting ethnicity and gender from iris texture ", in Proc. IEEE International Conference on Technologies for Homeland Security, pp.440–445, 2011.

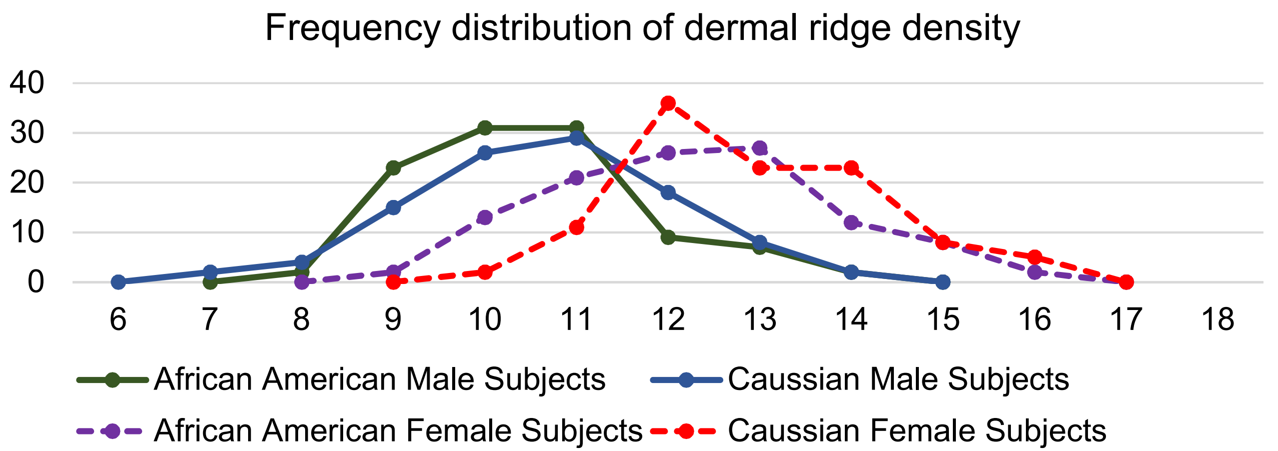
[3] A. Bansal, R. Agarwal, and R.K. Sharma, "Predicting Gender Using Iris Images", Research Journal of Recent Sciences, vol.3, no.4, pp.20–26, 2014.

[4] Juan E. Tapia, Claudio A. Perez and Kevin W. Bowyer, "Gender Classification From Iris Images Using Fusion of Uniform Local Binary Patterns", Lecture Notes in Computer Science. Springer, vol. 8926, pp. 751–763, 2015.

Gender from Biometric Data

— Fingerprint —

- **Observation and statistical analysis**
  - All ridges within the depicted 5mm × 5mm square were summed. This value is referred to as **ridge density** and serves as the basis of comparison.
  - Results show that **women tend to have a significantly higher ridge density than men** and that this trend is upheld in subjects of both Caucasian and African American descent.



Mark A. Acree, "Is there a gender difference in fingerprint ridge density?", Forensic Science International, vol.102, no.1, pp.35-44, 1999.

中国科学院自动化研究所
INSTITUTE OF AUTOMATION CHINESE ACADEMY OF SCIENCES

Gender from Biometric Data

www.ia.ac.cn

— Fingerprint —

- Gender classification for a specific race

Indian	100F+100M
Chinese and Malaysian	100F+100M
Turkish	118F + 88M
Egyptian	372F+380M
Mataco-Mataguay	110F + 99M
Argentinian and Spanish	193F+200M
Spanish Caucasian	100F+100M

Ridge count analysis
in different fingerprint areas

- Females have finer ridges than males.
- Females have more ridges in a given area than males.
- Females have larger ridge density, hence finer ridge details, than males.

[1] N. Kapoor and A. Badiye, "Sex Differences in the Thumbprint Ridge Density in a Central Indian Population", Egyptian Journal of Forensic Sciences, vol.5, no.1, pp:23-29, 2015.
[2] V. C. Nayak, et al., "Sex Differences from Fingerprint Ridge Density in Chinese and Malaysian Population", Forensic Science International, vol.197, no.1-3, pp:67-69, 2010.
[3] E. B. Ceyhan and S. Sagiroglu, "Gender Inference within Turkish Population by Using Only Fingerprint Feature Vectors", IEEE Symposium on Computational Intelligence in Biometrics and Identity Management, 2014.
[4] G. A. Eshak, et al., "Sex Identification from Fingertip Features in Egyptian Population", Journal of Forensic and Legal Medicine, vol.20, no.1, pp: 46-50, 2013.
[5] E. Gutiérrez-Redomero, et al, "Sex Differences in Fingerprint Ridge Density in The Mataco-mataguayo Population", HOMO - Journal of Comparative Human Biology, vol.62, no.6, pp: 487-499, 2011.
[6] E. Gutiérrez-Redomero, et al. "A Comparative Study of Topological and Sex Differences in Fingerprint Ridge Density in Argentinian and Spanish Population Samples", Journal of Forensic and Legal Medicine, vol.20, no.5, pp: 419-429, 2013.
[7] E. Gutiérrez-Redomero, et al., "Variability of Fingerprint Ridge Density in a Sample of Spanish Caucasians and Its Application to Sex Determination", Forensic Science International, vol.180, no.1, pp: 17-22, 2008.

Gender from Biometric Data

— Fingerprint —

Authors	Features	Classifiers	Datasets	Gender(the number of samples)	CCR*
Badawi *et al.* [1]	RTVTR**	Neural Network	unpublished	Female(1100)/Male(1100)	88.5%
Li*et al.* [2]	Bag-of-words features	Discriminative LDA	unpublished	Female(197)/Male(201)	close to 80%
Gupta*et al.* [3]	Discrete Wavelet Transform features	Artificial Neural Network	unpublished	Female(275)/Male(275)	91.45%
Ceyhan*et al.* [4]	Ridge thickness and counts	Naive Bayes	unpublished	Female(300)/Male(300)	95.3%

* "CCR" means correct classification rate.
** "RTVTR" means ridge thickness to valley thickness ratio.
*** "LDA" means Latent Dirichlet Allocation.

[1] Ahmed Badawi, Mohamed Mahfouz, Rimon Tadross and Richard Jantz, "Fingerprint-based Gender Classification", in Proc. International Conference on Image Processing, Computer Vision, pp. 41–46, 2006.

[2] Xiong Li, Xu Zhao, Yun Fu and Yuncai Liu, "Bimodal Gender Recognition from Face and Fingerprint", in Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 2590–2597, 2010.

[3] Samta Gupta and A. Prabhakar Rao, "Fingerprint Based Gender Classification Using Discrete Wavelet Transform & Artificial Neural Network", International Journal of Computer Science and Mobile Computing, pp. 1289–1296, 2014.

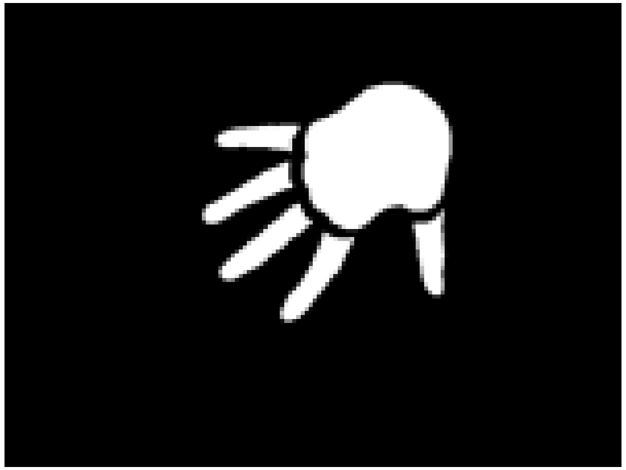
[4] Eyup Burak Ceyhan and Seref Sagiroglu, "Gender Inference within Turkish Population by Using Only Fingerprint Feature Vectors", IEEE Symposium on Computational Intelligence in Biometrics and Identity Management, pp. 146–150, 2014.

Gender from Biometric Data

— Hand geometry —

Table 1. Measurements of hand breadth and length (based on centimeters) [1].

	Population	Hand Breadth (cm)				Hand Length (cm)				Hand Index (%)
Gender	N	min	max	μ	σ	min	max	μ	σ	μ
Male	125	7.3	9.4	8.45	0.40	15.3	21.0	18.89	0.88	44.73
Female	125	6.7	8.8	7.48	0.38	14.8	20.4	17.22	0.92	43.46



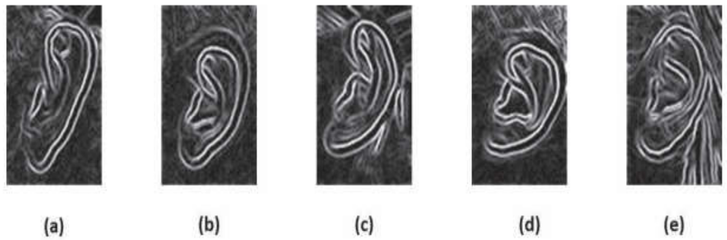
Region and boundary features + LDA

CCR: 98%

Gholamreza Amayeh, George Bebis and Mircea Nicolescu, "Gender Classification from Hand Shape", IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp.1-7, 2008.

Gender from Biometric Data

— Ear —



3D

Histogram of Indexed Shapes (HIS)
+
SVM

CCR: 92.94 ± 1.44%

2D

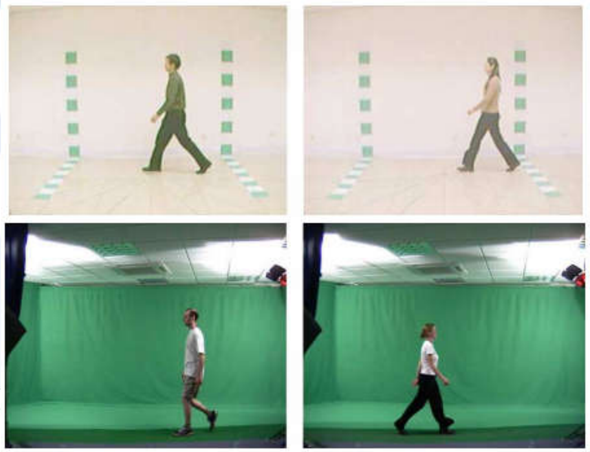
SIFT
+
Support Vector Classification (SVC)

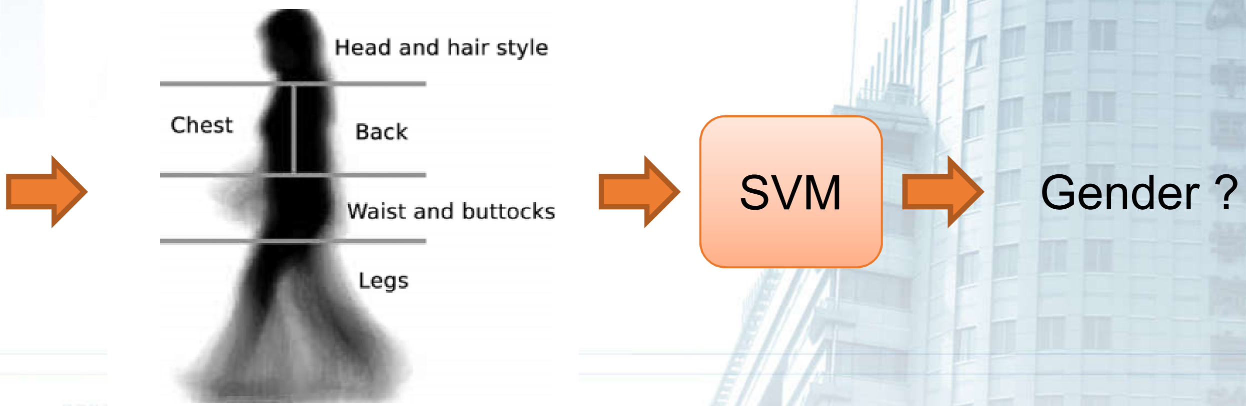
CCR: 97.65% ± 2.06%

[1] Jiajia Lei, Jindan Zhou and Mohamed Abdel-Mottaleb. "Gender Classification Using Automatically Detected and Aligned 3D Ear Range Data", in Proc. International Conference on Biometrics, 2013.
[2] Guangpeng Zhang and Yunhong Wang. "Hierarchical and Discriminative Bag of Features for Face Profile and Ear Based Gender Classification", in Proc. International Joint Conference on Biometrics, 2011.

Gender from Biometric Data

— Gait —





Segmented GEI

TABLE II
CORRECT CLASSIFICATION RATES USING THE CASIA DATABASE (DATASET B)

Method	Dataset	CCR
Lee and Grimson [6] [a]	25 males and 25 females	85.0%
Huang and Wang [14] [b]	25 males and 25 females	85.0%
Li *et al.* [3] [c]	31 males and 31 females	93.28%
44 human observers	31 males and 31 females	95.47%
Proposed method	31 males and 31 females	95.97%

Cross-race experimental results (correct classification rate)

Asian training data, European test data	European training data, Asian test data
87.15%	87.90%

Shiqi Yu, Tieniu Tan, Kaiqi Huang, Kui Jia and Xinyu Wu, "A Study on Gait-Based Gender Classification", IEEE Transactions on Image Processing, vol.18, no.8, pp.1905-1910, 2009.

Gender from Biometric Data

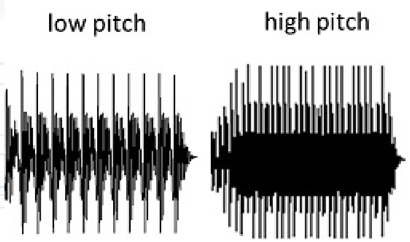
www.ia.ac.cn

— Voice —

Applications of gender from voiceprint:

(1) sort telephone calls by gender for gender sensitive surveys;

(2) enhance speaker adaptation as part of an automatic speech/speaker recognition system.



Pitch features

Mel-frequency cepstral coefficients (MFCC)

Fusion

Table 4. The accuracy of the combining approach

Session	Accuracy
Dialect Digits	98.3%
English Digits	98.7%
Province Phrase	96.7%
Mandarin Digits	99.7%

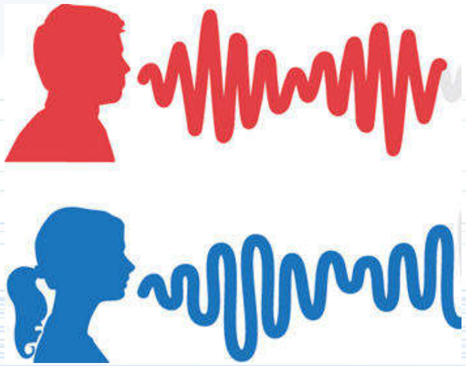
Ting Huang, Yingchun Yang and Zhaohui Wu, "Combining MFCC and Pitch to Enhance the Performance of the Gender Recognition“, in Proc. International Conference on Signal Processing, pp.16-20, 2006.

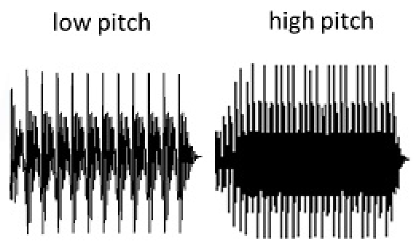
Gender from Biometric Data

www.ia.ac.cn

— Voice —

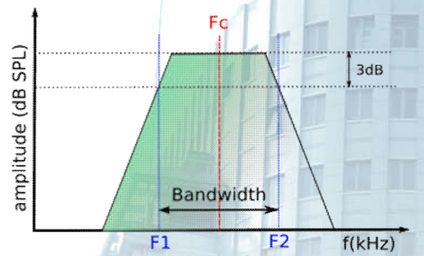
features



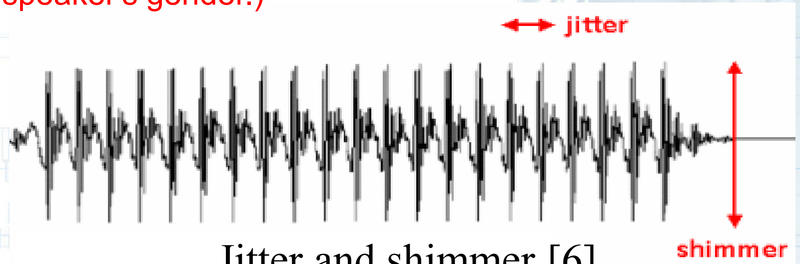


Pitch [1-4]

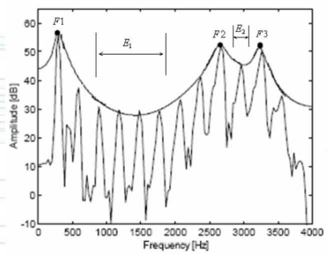
(It is a physiologically distinctive trait of a speaker's gender.)



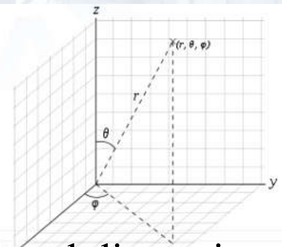
Frequency and bandwidth [2]



Jitter and shimmer [6]



Energy [5]



Fractal dimension and fractal complexity [3]

[1]Yu-Min Zeng, et al., “Robust GMM Based Gender Classification Using Pitch and RASTA-PLP Parameters of Speech”, in Proc. Int. Conf. Mach. Learn. Cybern. pp. 3376-3379, 2006.
[2] Yen-Liang Shue, et al., “The Role of Voice Source Measures on Automatic Gender Classification”, in Proc. IEEE ICASSP, pp. 4493-4496, 2008.
[3] Yingle Fan, et al., “Speaker gender identification based on combining linear and nonlinear features”, in Proc. 7th WCICA. pp. 6745-6749, 2008.
[4] Ting Huang , et al., “Combining MFCC and pitch to enhance the performance of the gender recognition”, in Proc. 8th Int. Conf. Signal Process., 2006.
[5] Deepak S. Deepawale, et al., “Energy estimation between adjacent formant frequencies to identify speaker's gender”, in Proc. 5th Int. Conf. ITNG, pp. 772-776, 2008.
[6] Florian Metze, et al., “Comparison of four approaches to age and gender recognition for telephone applications”, in *Proc. IEEE ICASSP*, pp. IV-1089IV-1092, 2007.

Gender from Biometric Data

— Handwriting —

ICDAR 2013 Competition on Gender Prediction from Handwriting

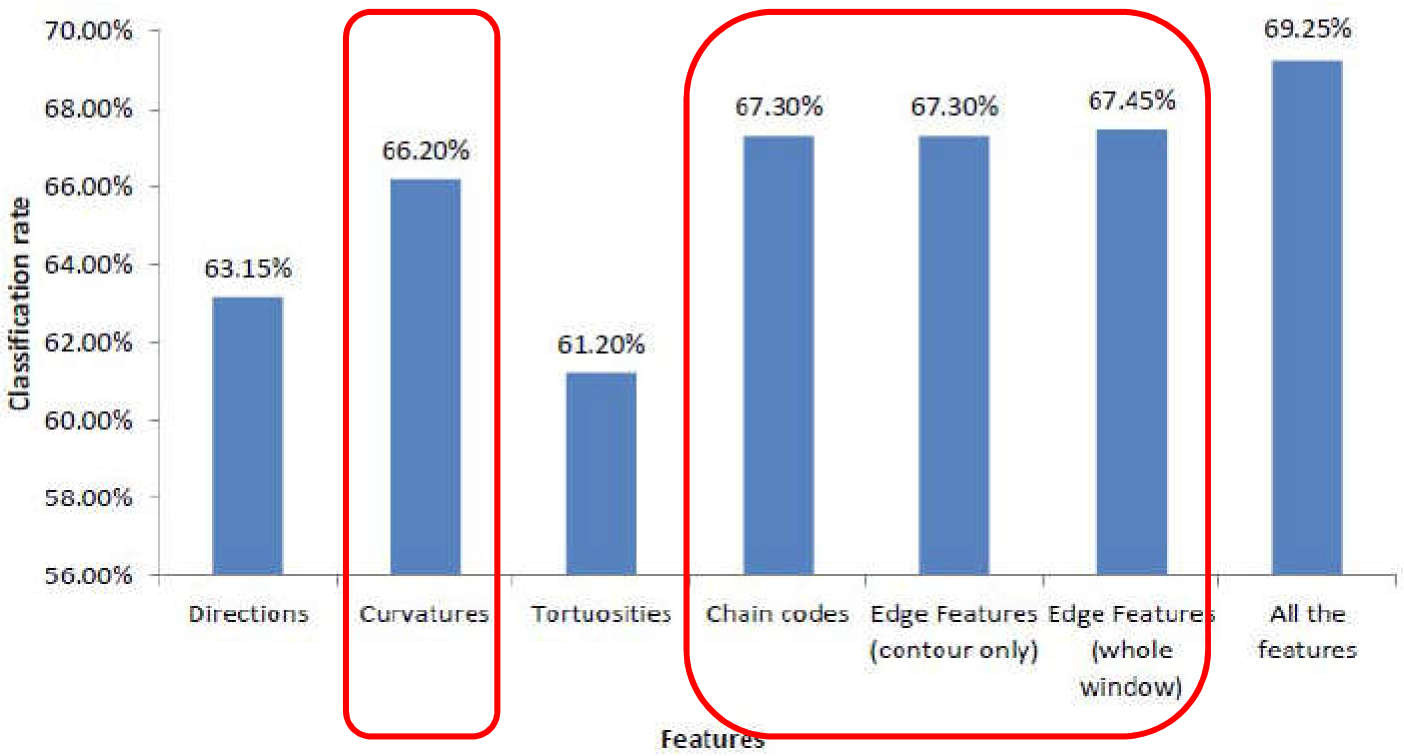


Figure 2: Identification rates of each category of features.

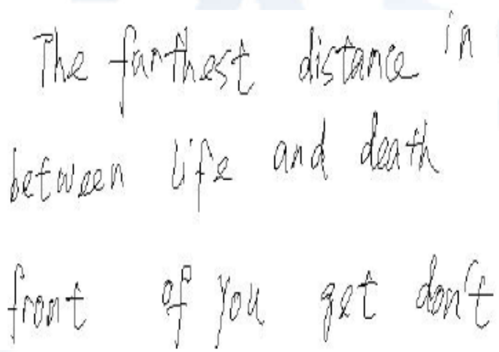
Abdulaali Hassaine, et al., "ICDAR 2013 Competition on Gender Prediction from Handwriting", in Proc. International Conference on Document Analysis and Recognition, pp.1417-1421, 2013.

CASIA 中国科学院自动化研究所 INSTITUTE OF AUTOMATION CHINESE ACADEMY OF SCIENCES

Gender from Biometric Data

www.ia.ac.cn

— Handwriting —



Direction features
Curvature features
Tortuosity features
Chain code features
Edge-based directional features

Random Forest Classifier with kernel discriminant analysis using spectral regression

Gender, age and nationality

QUWI dataset
1,017 writers in both English and Arabic

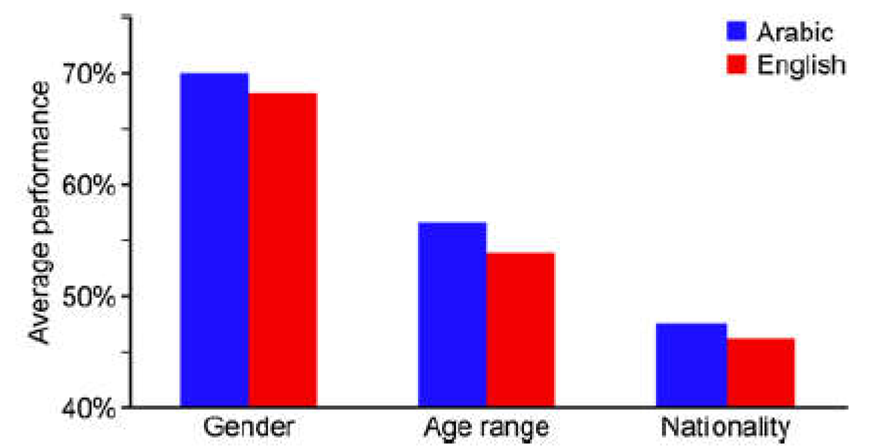
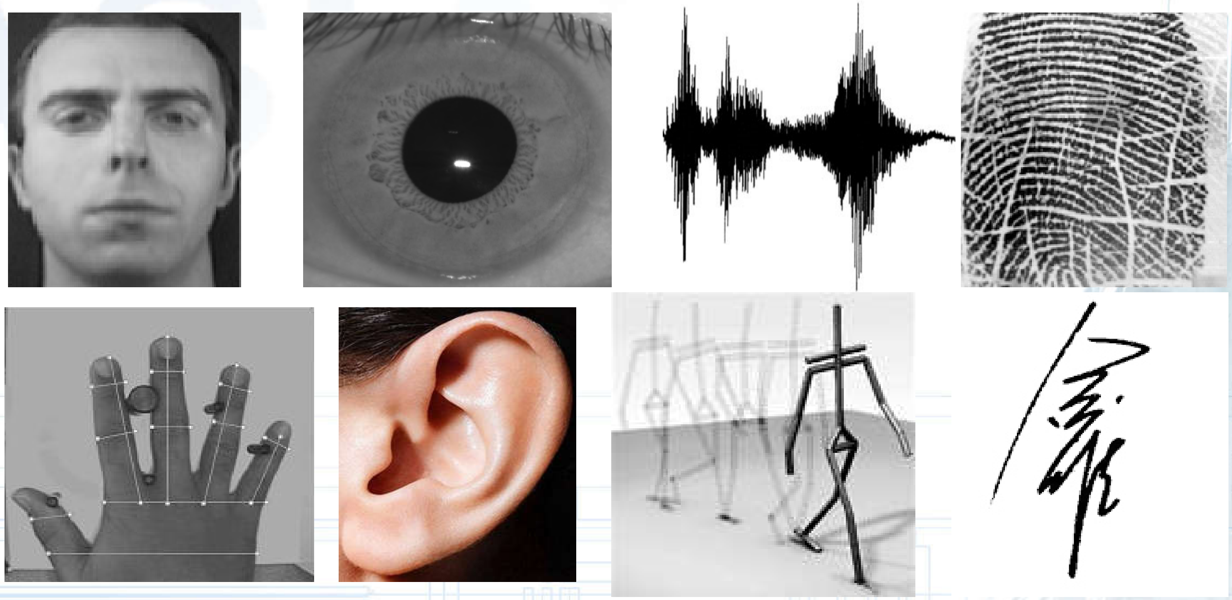


Figure 11 Average performance for Arabic and English texts.

Somaya Al Maadeed and Abdelaali Hassaine, "Automatic Prediction of Age, Gender and Nationality in Offline Handwriting." EURASIP Journal on Image and Video Processing, vol.2004, no.1, pp.1-10, 2014.

Gender from Biometric Data

— Conclusions —



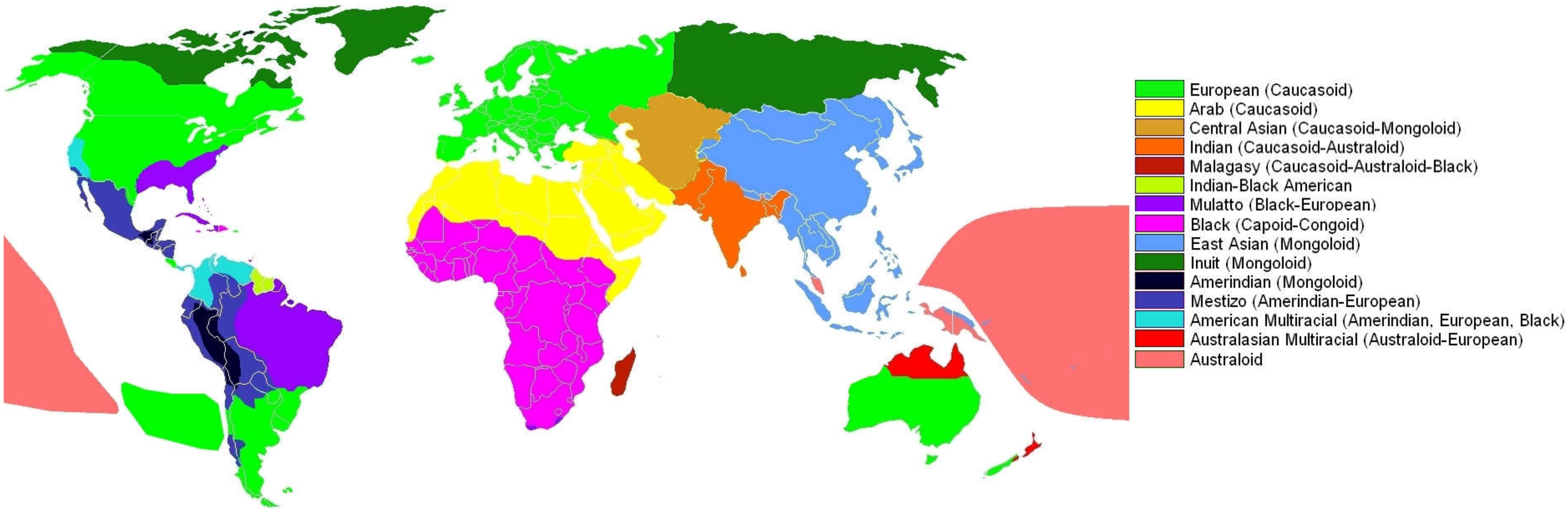
- Common biometric modalities such as face, iris, voice, fingerprint, hand geometry, ear, gait and handwriting have shown promising performance in gender estimation.
- **Future work:** gender from multi-modal biometric data and large-scale databases for algorithm research and evaluation

Outline

- Preamble
- Identity from Biometric Data
- Gender from Biometric Data
- Ethnicity from Biometric Data
- Age and Affect from Biometric Data
- Conclusions

Ethnicity from Biometric Data

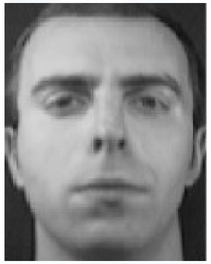
- **Definition from Wiki**
  - **An ethnic group or ethnicity is a socially defined category of people who identify with each other based on common ancestral, social, cultural or national experience. Unlike most other social groups, ethnicity is primarily an inherited status.**
  - **Ethnic groups derived from the same historical founder population often continue to speak related languages and share a similar gene pool.**



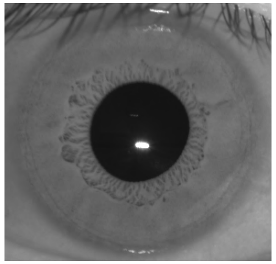
Ethnic Groups in the World by the End of the 20th Century

Ethnicity from Biometric Data

The most popular and informative biometric modality for ethnicity estimation is face.



Face



Iris

Gait

Ethnicity from Biometric Data

Significant facial appearance differences for various ethnicities

Ethnic Group	Facial Characteristics
Asian (Mongoloid)	-Eyes: Narrow, epicanthic fold -Nose: low, average width -Lips: average fullness -Face Shape: short with flat, projected cheekbones, -Hair: thick and straight or slightly wavy, thin facial hair -Skin: yellowish
European (Caucasoid)	-Eyes: double eyelid, exposed tear trough, large -Nose: Prominent, high bridge, narrow -Lips: thin, tight -Face Shape: center of face juts outward, wedge shaped, long face -Hair: wavy or curly, thick body and facial hair (males) -Skin: light or brown
African (Negroid)	-Eyes: large with exposed tear trough -Nose: broad, low -Lips: thick, stretched mouth -Face Shape: long head -Hair: tight curls or heavy waves -Skin: dark (high melanin quantity)

From
Satoshi Hosoi, Erina Takikawa and Masato Kawade. "Ethnicity Estimation with Facial Images", in Proc. IEEE International Conference on Automatic Face and Gesture Recognition, 2004.

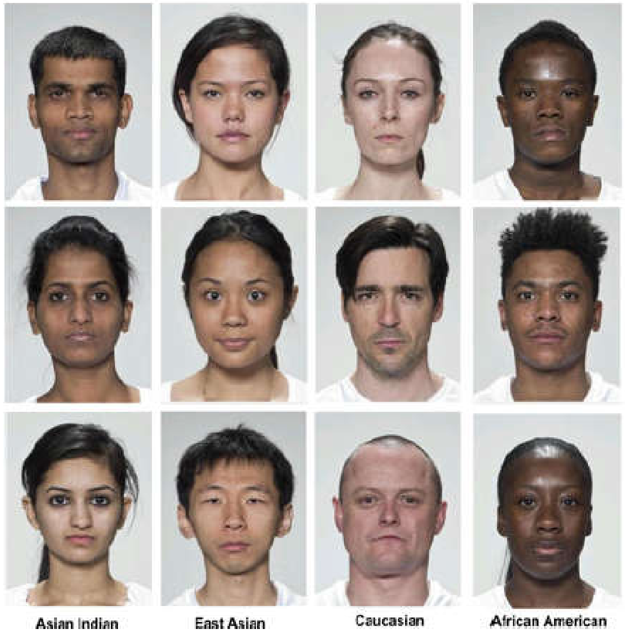
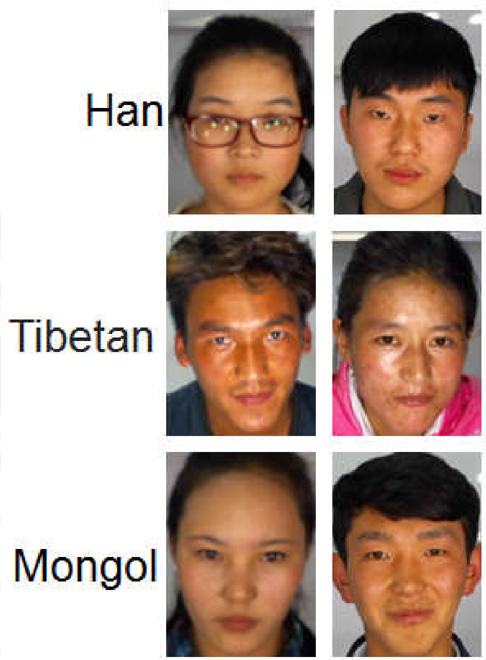


Fig. 3. Illustrative genetic variance distribution of human races, while in practice it is often accepted that 3- to 7-races classification system would be enough for regular applications (Figure source: http://www.faceresearch.org/).

From
Siyao Fu, Haibo He and Zeng-Guang Hou. "Learning Race from Face: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.36, no.12, pp.2483-2509, 2014.



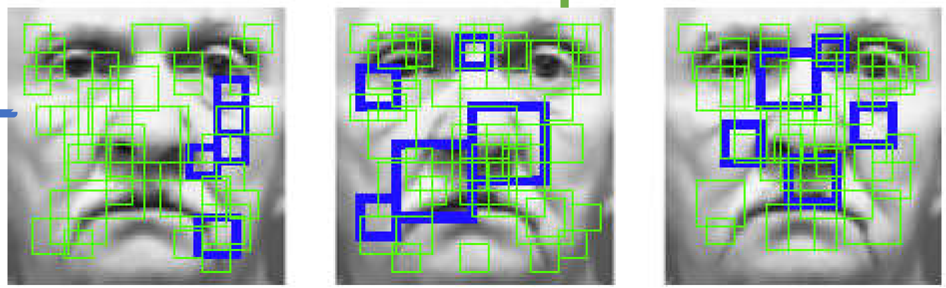
Facial Appearance of three ethnicities in China

Ethnicity from Biometric Data

www.ia.ac.cn

— Face —

For **gender** classification, **besides the features located around the eyes and lip, the jaw** is also found to be salient.



(a) (b) (c)

Fig. 6. The top five (blue) and top 6-50 (green) most informative BIF features selected for estimating (a) age, (b) gender, and (c) race. The rectangle size indicates the scale of the corresponding BIF feature.

Dataset：
a subset of MORPHII
(2000 images）

CCR:
(Black vs White)
99.1%

The most informative features for **age** estimation are located in the regions where wrinkles typically appear, such as the **eye and mouth corners, nasolabial folds, and cheeks**.

For **race (white vs. black)** classification, the most informative features are **around the eyes, nose, and lip**.

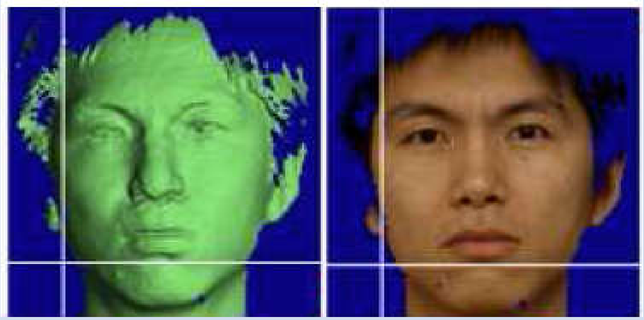
Hu Han, Charles Otto, Xiaoming Liu and Anil K. Jain, "Demographic Estimation from Face Images: Human vs. Machine Performance", IEEE Trans. PAMI, vol.37, no.6, pp.1148-1161, 2015.

Ethnicity from Biometric Data

— 3D Face —

	Features	Results
[1]	Range and intensity images	Error rate (see table below)
[2]	3D mesh	(see table below)

[1]

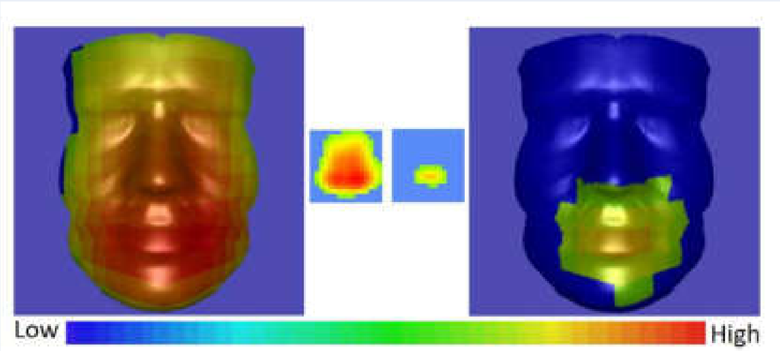


Range and intensity images

Error rate

	Ethnicity	Gender
Range	$3.8\% \pm 0.024$	$14.6\% \pm 0.044$
Intensity	$3.2\% \pm 0.029$	$14.0\% \pm 0.047$
Range + Intensity	$2.0\% \pm 0.016$	$9.0\% \pm 0.030$

[2]



3D mesh

UR3D	White	Asian
White	(99.8 ± 0.5) %	(0.2 ± 0.5) %
Asian	(0.3 ± 0.9) %	(99.7 ± 0.9) %
GMPM	**White**	**Asian**
White	(99.7 ± 0.3) %	(0.3 ± 0.3) %
Asian	(0.4 ± 0.8) %	(99.6 ± 0.8) %

[1] Xiaoguang Lu, Hong Chen and Anil K. Jain, "Multimodal facial gender and ethnicity identification", Advances in Biometrics. Springer Berlin Heidelberg, pp. 554-561, 2005.
[2] Omar Ocegueda, et al., "3D Face Discriminant Analysis Using Gauss-markov Posterior Marginals", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.35, no.3, pp. 728–739, 2013.

中国科学院
自动化研究所
INSTITUTE OF AUTOMATION
CHINESE ACADEMY OF SCIENCES

Ethnicity from Biometric Data

www.ia.ac.cn

— Iris —

Authors	Feature	Classifier	Dataset	Race(the number of samples)	CCR**
Qiu *et al.* [1]	Gabor energy	SVM	CASIA	Asian(1200)/non-Asian(1200)	91.02%
Lagree *et al.* [2]	Statistical features	SMO in WEKA	Self-collection	Caucasian(1200)/Asian(1200)	92.58%
Zhang *et al.* [3]	SIFT+Kmeans+LLC+SPM	SVM	Self-collection	Asian(10000)/non-Asian(1320)	94.28%
Zarei *et al.* [4]	Spot and line features	Neural net	Self-collection	Caucasian(NM*)/non-Asian(NM)	97.50%
Sun *et al.* [5]	SIFT+HVC	SVM	UPOL, ICE2005,UBIRIS, CASIA and Self-collection	Asian(15000)/non-Asian(5549)	97.86%

* "NM" means not mentioned in the paper.
** "CCR" means correct classification rate.

[1] Xianchao Qiu, Zhenan Sun and Tieniu Tan, "Learning appearance primitives of iris images for ethnic classification", in Proc. International Conference on Image Processing, vol. 2, pp. II–405–II–408, 2007.
[2] Stephen Lagree and Kevin W. Bowyer, "Predicting Ethnicity and Gender from Iris Texture", in Proc. IEEE International Conference on Technologies for Homeland Security, pp. 440–445, 2011.
[3] Hui Zhang, Zhenan Sun, Tieniu Tan and Jianyu Wang, "Ethnic Classification Based on Iris Images.", ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, vol. 7098, book section 11, pp. 82–90, 2011.
[4] Anahita Zarei and Mou Duxing, "Artificial Neural Network for Prediction of Ethnicity Based on Iris Texture," in Proc. International Conference on Machine Learning and Applications, pp. 514–519, 2012.
[5] Zhenan Sun, Hui Zhang, Tieniu Tan and Jianyu Wang, "Iris Image Classification Based on Hierarchical Visual Codebook." , IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.36, no.6, pp. 1120–1133, 2014.

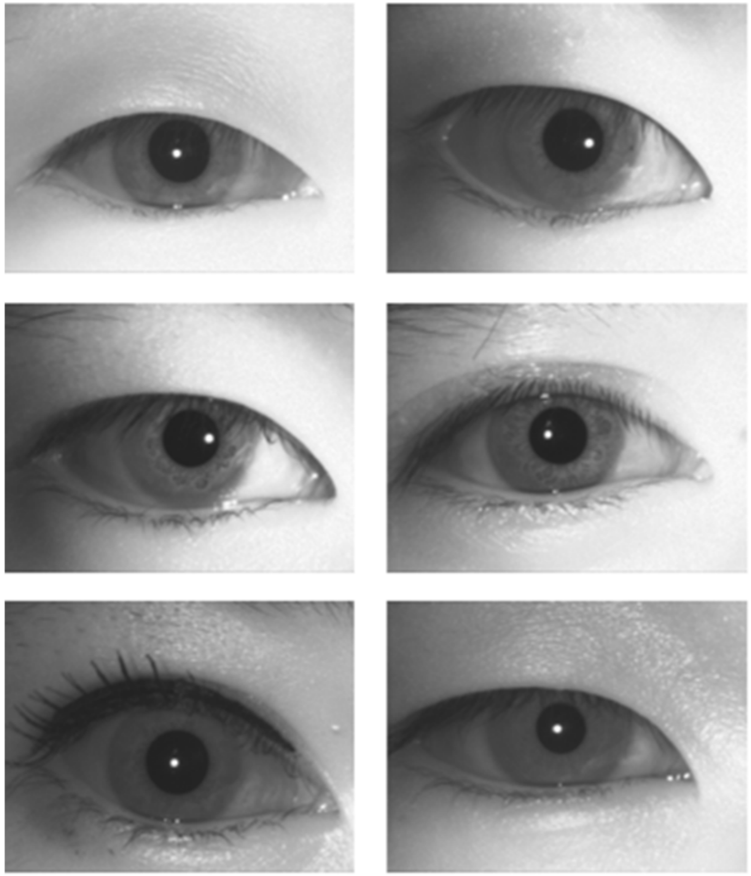
Ethnicity from Biometric Data

— Iris —

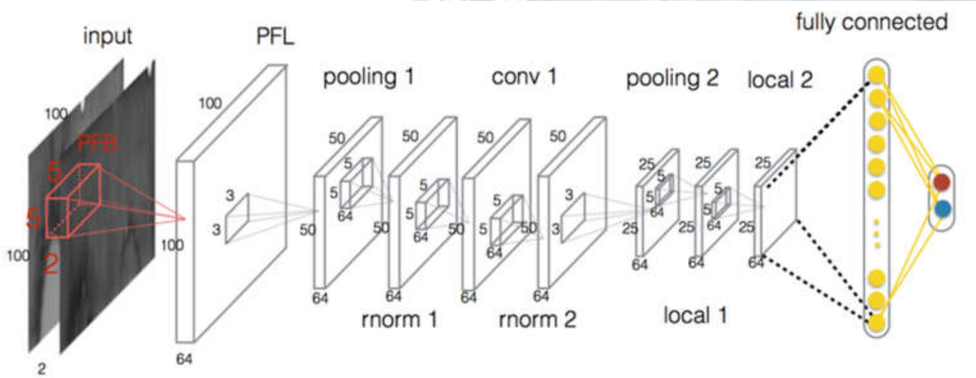
Han

Tibetan

Mongol



Joint gender/ethnicity estimation based on deep learning



Ethnicity from Biometric Data

— Iris —

	Han	Tibetan	Mongol
Male	404 subjects 8,068 images	178 subjects 3,560 images	58 subjects 1,160 images
Female	266 subjects 5,318 images	124 subjects 2,480 images	72 subjects 1,439 images
Total	**670 subjects 13,386 images**	**302 subjects 6,040 images**	**130 subjects 2,599 images**

	CCR
Race prediction	98.09%
Gender prediction	98.46%
Multi-task (race and gender)	Race: 99.05% Gender: 99.23%

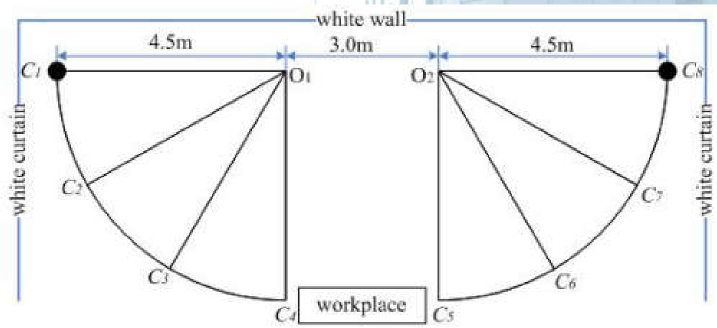
Ethnicity from Biometric Data

www.ia.ac.cn

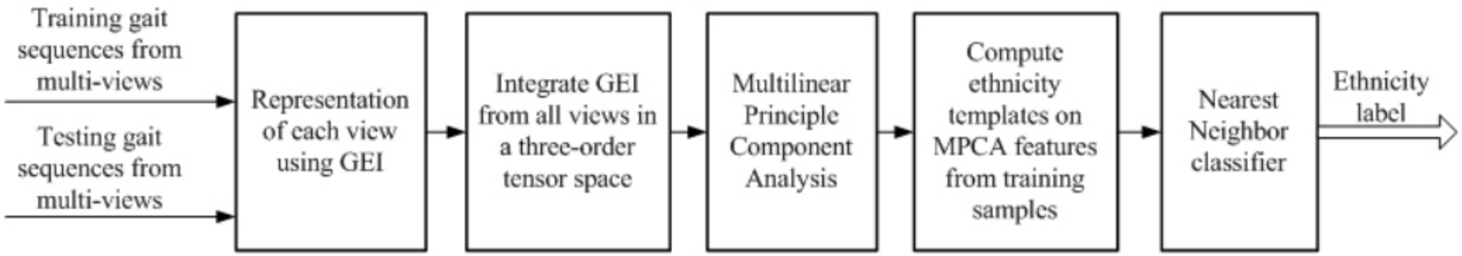
— Gait —

An East Asian subject

An South American subject



Different view angles



Correct classification rate: 84.4%

GEI + multilinear principal component analysis (MPCA) + multi-view gait feature fusion

De Zhang, Yunhong Wang and Bir Bhanu. "Ethnicity Classification Based on Gait Using Multi-view Fusion", in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp.108-115, 2010.

中国科学院自动化研究所
INSTITUTE OF AUTOMATION CHINESE ACADEMY OF SCIENCES

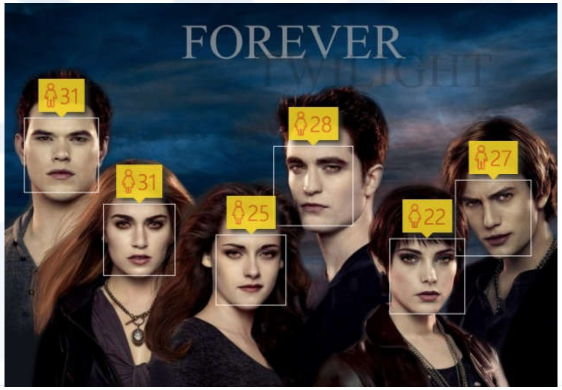
Outline

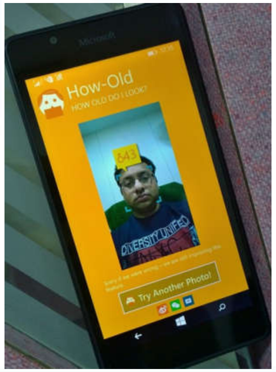
www.ia.ac.cn

- Preamble
- Identity from Biometric Data
- Gender from Biometric Data
- Ethnicity from Biometric Data
- Age and Affect from Biometric Data
- Conclusions

Age and affect from biometric data

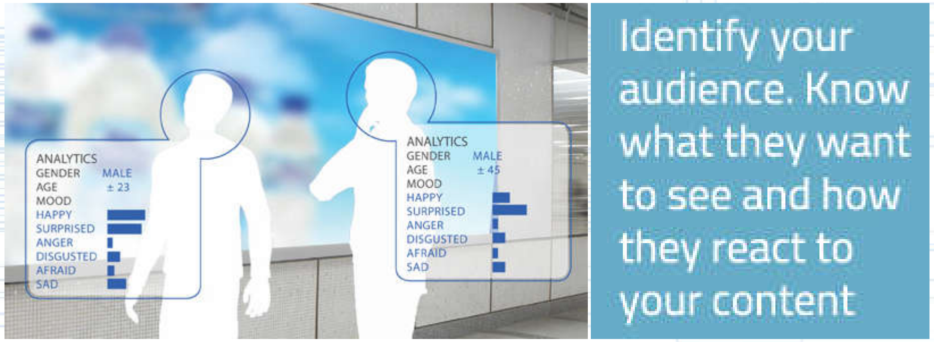
—Applications —





How-Old.net (Microsoft)

Human-Computer (Robot) Interaction

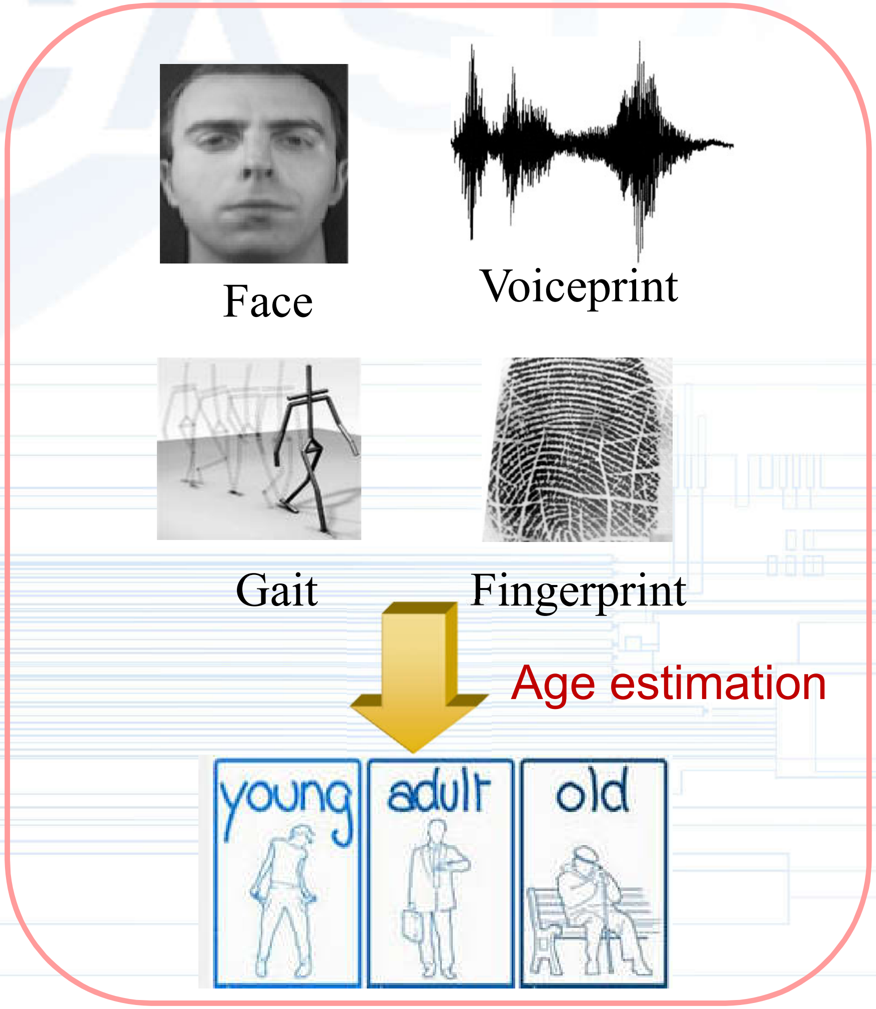


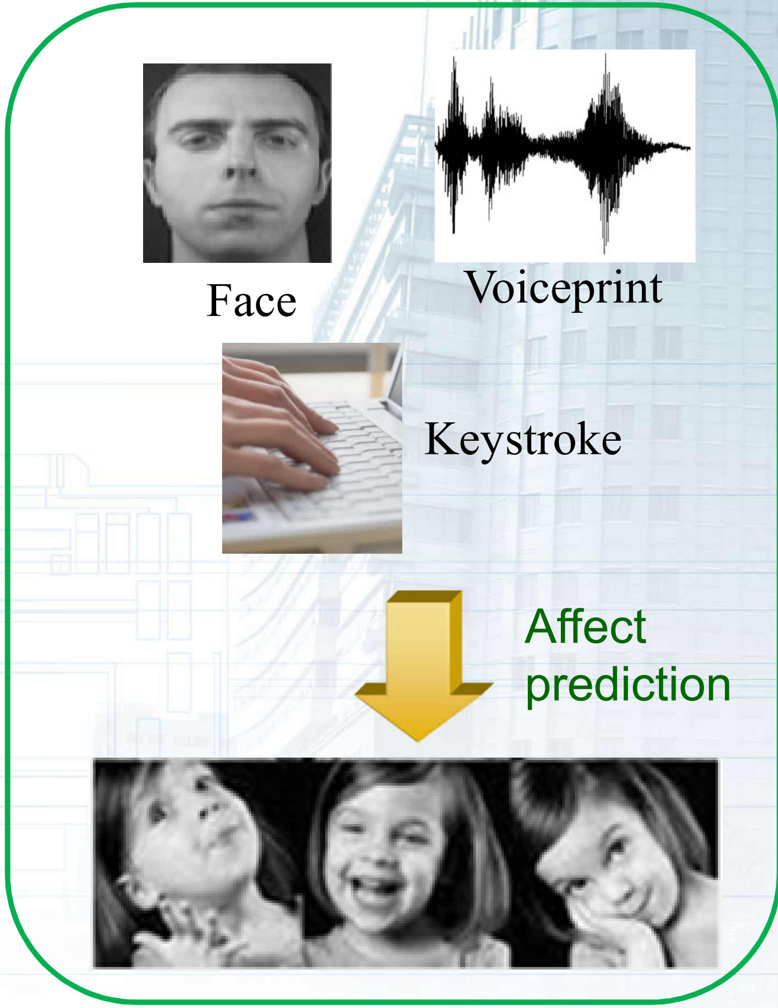
Understand and Predict Your Audience

Driver Monitoring

Age and affect from biometric data

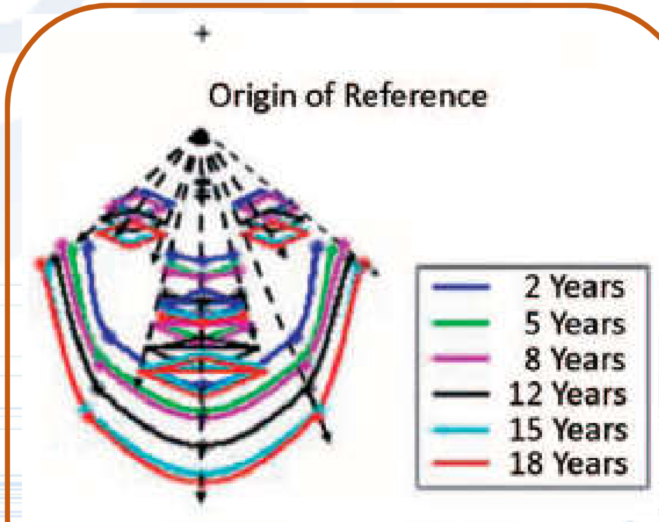
Biometric modalities informative for age and affect prediction





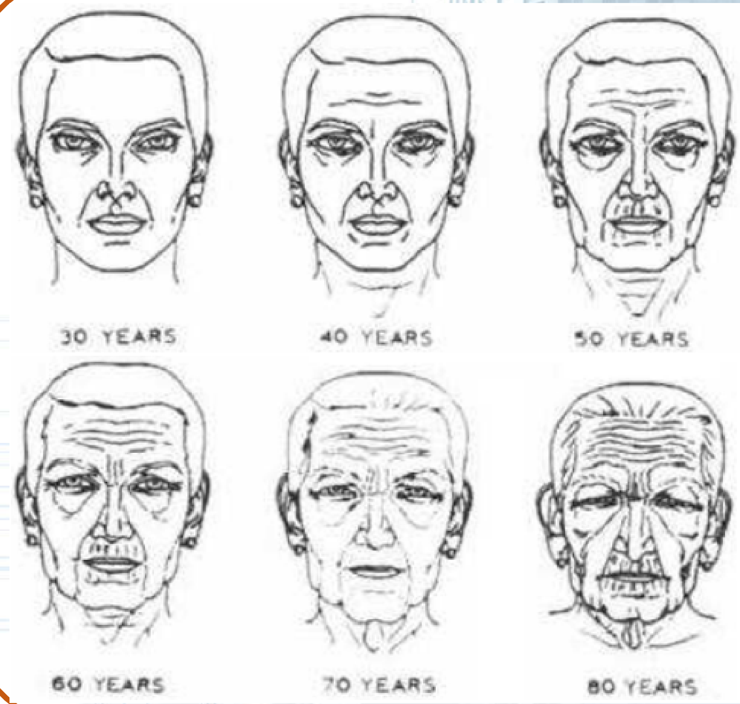
Two Stages of Face Aging

Stage 1: Early Aging



Craniofacial growth model

Stage 2: Adult Aging

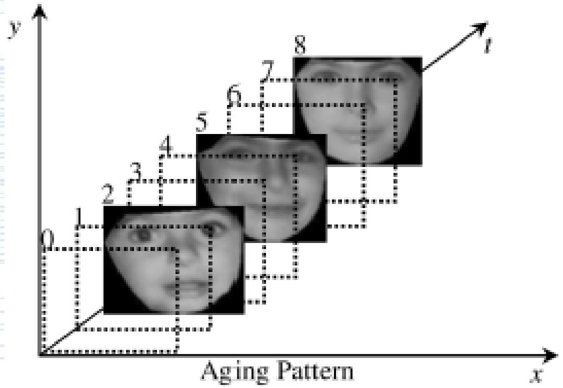


[1] N. Ramanathan and R. Chellappa, "Modeling Age Progression in Young Faces," In Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 387-394, 2006.
[2] M. Gonzalez-Ulloa and E. Flores, "Senility of the Face: Basic Study to Understand Its Causes and Effects," Plastic and Reconstructive Surgery, vol. 36, pp. 239-246, 1965.

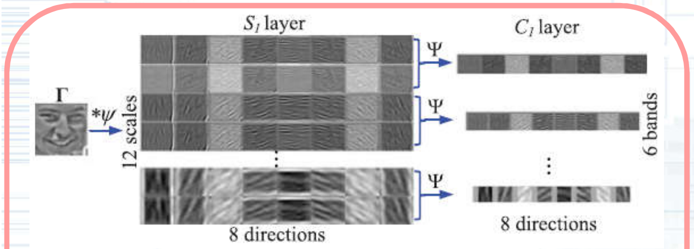
Facial Age Representations

Mean absolute error of age estimation on three public face databases (in Years)

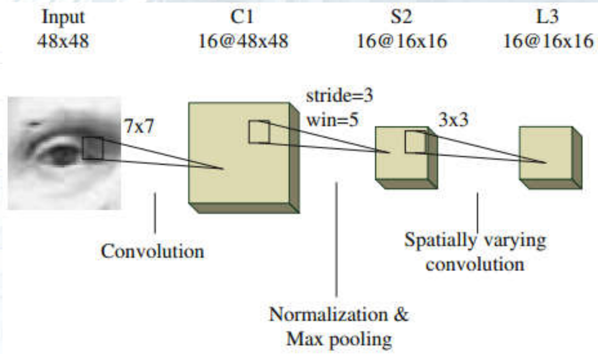
Database	Proposed algorithm		Human workers	
	w/o QA	w/ QA	w/o QA	w/ QA
FG-NET	4.8 ± 6.2	3.8 ± 4.2	4.7 ± 5.0	4.5 ± 4.8
MORPH II	3.8 ± 3.3	3.6 ± 3.0	6.3 ± 4.9	4.3 ± 3.8
PCSO	4.3 ± 3.7	4.1 ± 3.3	7.2 ± 5.7	6.6 ± 4.9



Facial Aging Patterns
(Geng et al., 2007)



Appearance Models
(Han et al. 2015)

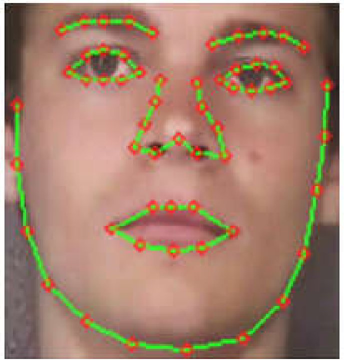


Deep Age Models
(Yi et al., 2014)

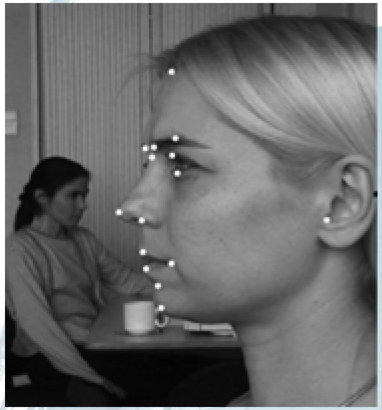
Facial Affect Representations

Experimental results on CK+ database

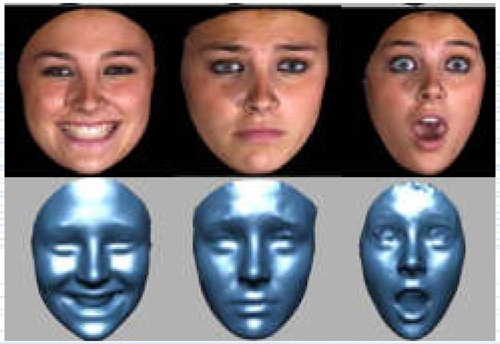
Method	Accuracy(%)
3D SIFT [23]	81.35
HOE [27]	82.26
LBP-TOP [34]	88.99
HOG 3D [14]	91.44
ITBN [30] (15-fold)	86.3
CERT [17]	87.21
MCF [3] (LOSO)	89.4
MSR [21]	91.4
TMS [12] (4-fold)	91.89
Cov3D [22] (5-fold)	92.3
Ours STM	**91.13**
Ours STM-ExpLet	**94.19**



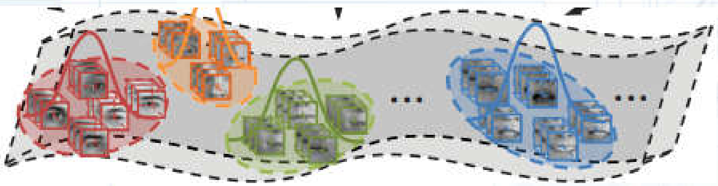
Spatiotemporal Geometric Features (Chang et al., 2006)



Profile View Face (Pantic et al., 2006)



3D Face Models (Yin et al. 2006)



Spatiotemporal Manifold (Liu et al. 2014)

Facial Muscle Action Units (Valstar et al., 2006)

Challenges of Age & Affect from Face

Age Estimation

Individual Differences

DoB: 1973-01-18 1974-10-05

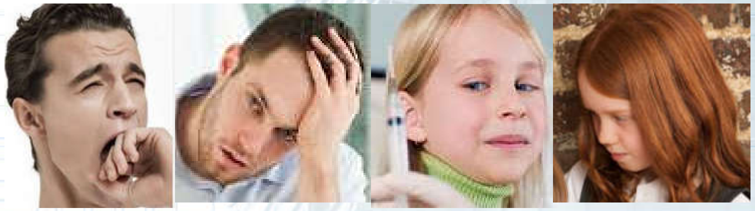
Group Differences (race, gender, etc)

Faces in the Wild

Affect Prediction

Spontaneous Affect

Attitudinal and Non-basic Affect

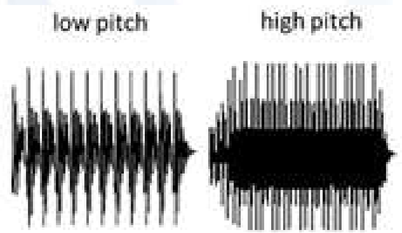


Faces in the Wild

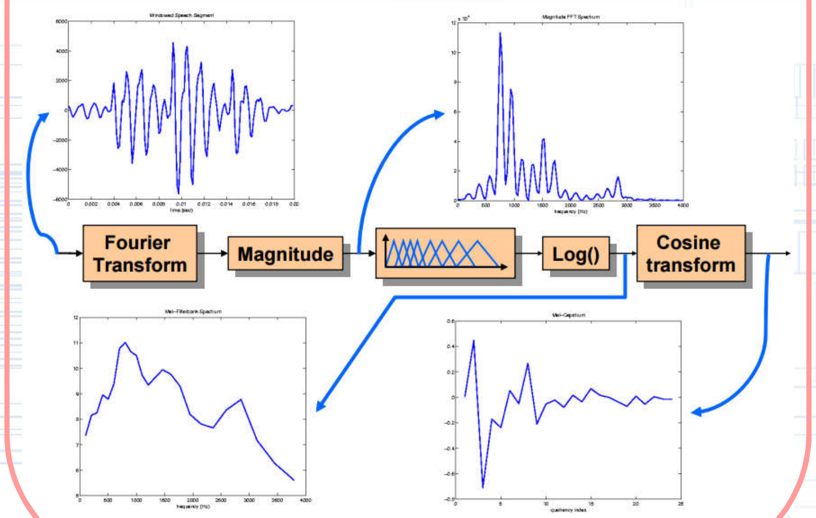
Age & Affect from Voiceprint

Age Estimation

Prosodic Features
(pitch, energy, speech rate, etc)

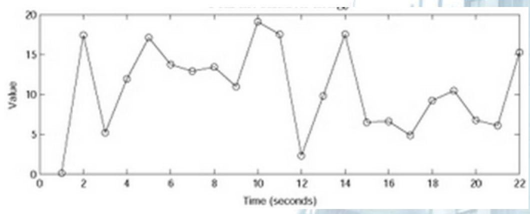


Spectral Features
(MFCC, cepstral features, etc)

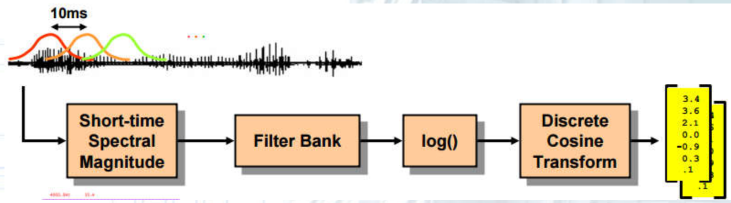


Affect Prediction

Prosodic Features
(pitch, energy, speech rate, etc)

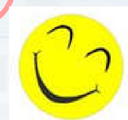


Spectral Features
(MFCC, cepstral features, etc)



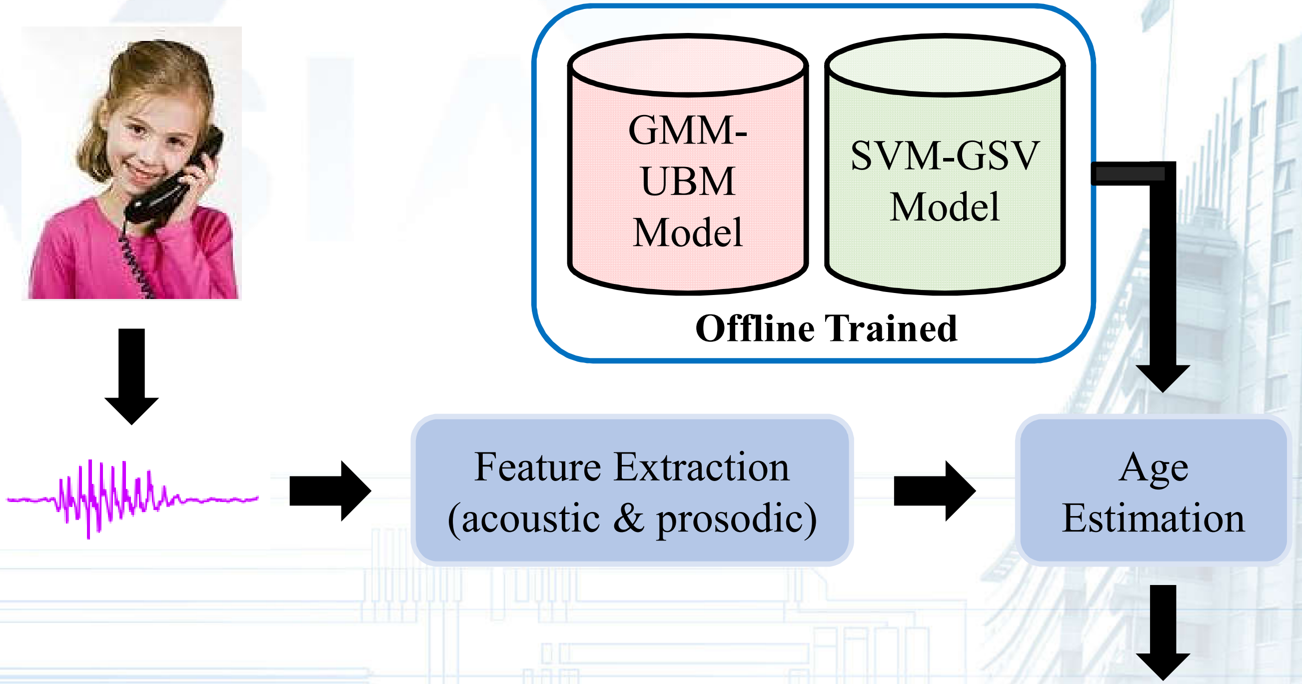
Linguistic Features
(language, discourse, context, etc)

Nice weather, isn't it?



[1] D.A. Reynolds, "Overview of Automatic Speaker Recognition", JHU 2008 Workshop Summer School
[2] Z. Zeng, M. Pantic, G.I. Roisman and T. S. Huang. "A survey of affect recognition methods: Audio, visual, and spontaneous expressions." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 1, pp. 39-58, 2009

Age from Voiceprint



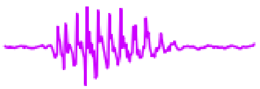
Confusion matrix on the test set of Interspeech 2010 Paralinguistic Challenge (%)

	C	Y	A	S
Children	**71.0**	15.8	5.5	7.8
Youths	7.3	**41.8**	26.2	24.7
Adults	2.2	19.1	**25.3**	53.4
Seniors	4.0	9.8	16.3	**70.0**

[1] M. Li, K. J. Han, and S. Narayanan. "Automatic speaker age and gender recognition using acoustic and prosodic level information fusion."Computer Speech & Language, vol. 27, no.6, pp. 151-167, 2014.

Affect from Voiceprint

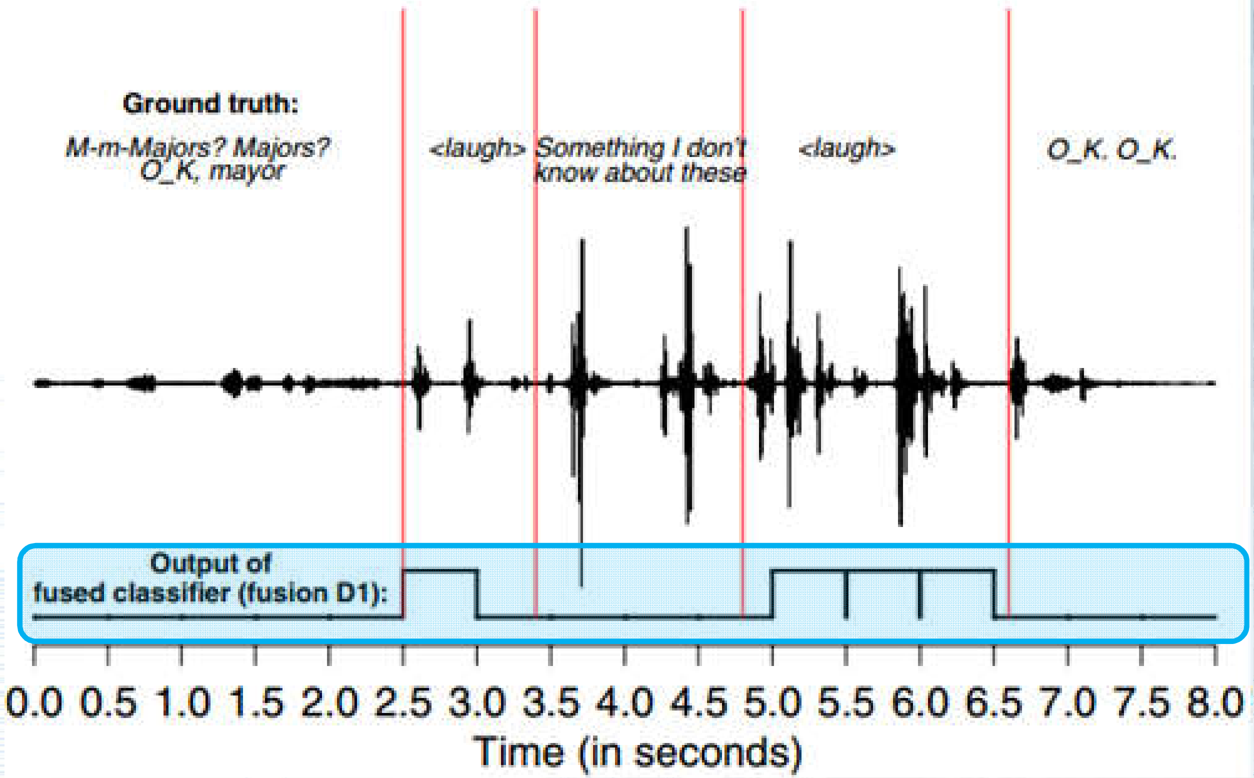
Laughter vs Speech



Feature Extraction (spectral & pitch)

Classifier (GMM + SVM)

Dataset	EER
Meeting recordings	2.9%
Conversation recordings	7.5%



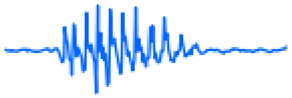
GMM+SVM applied to a fragment of a meeting recording

[1] K.P. Truong, and D. A. Van Leeuwen. "Automatic discrimination between laughter and speech." Speech Communication, vol.49, no.2, pp. 144-158, 2007.

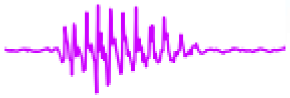
Challenges of Age & Affect from Voiceprint

Age Estimation

Individual Differences

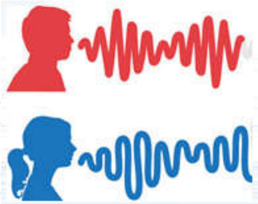


David's voice



John's voice

Group Differences (race, gender, etc)



Naturalistic Audio Recordings

Affect Prediction

Attitudinal and Non-basic Affect

Interest?

Shame?

Fatigue?

Linguistic Feature Extraction

Nice weather, isn't it?

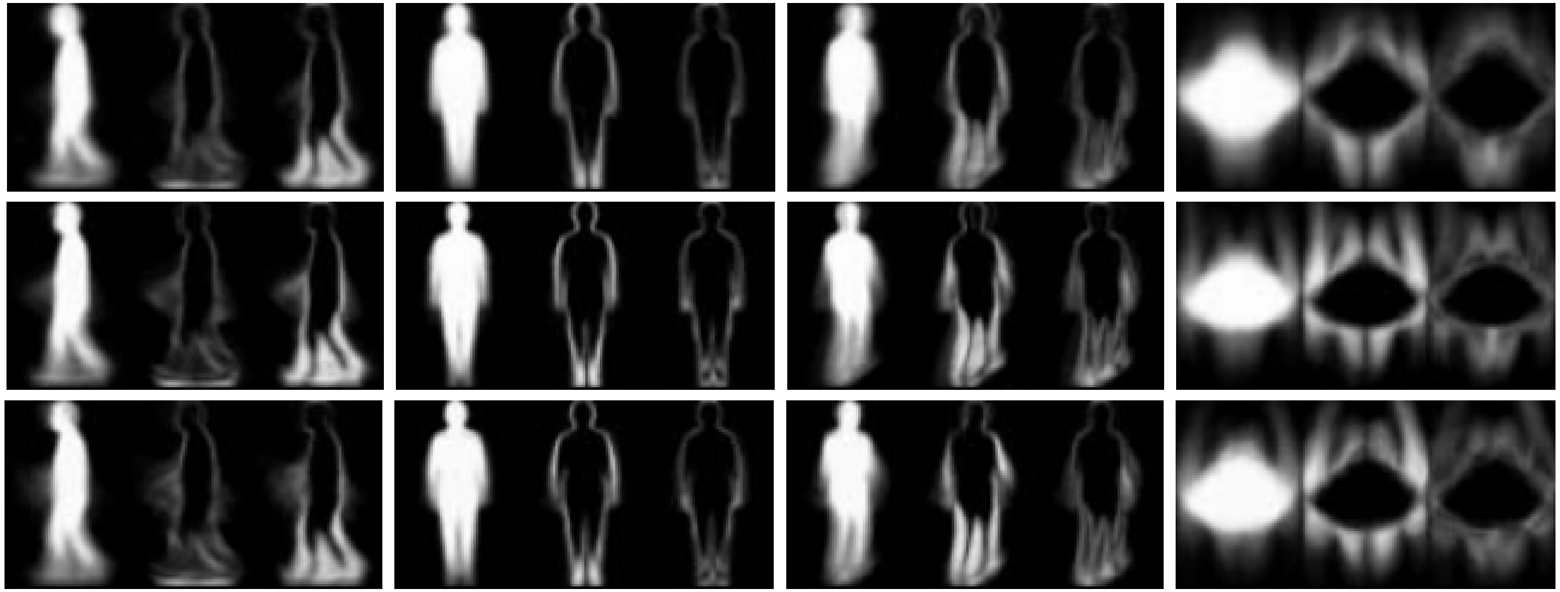
Naturalistic Audio Recordings

Age from Multi-view Gait

Young

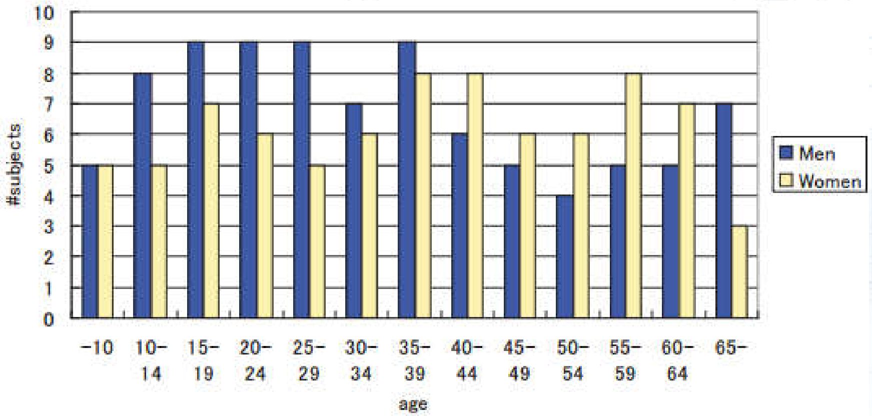
Adult

Elder



left profile　front　right-back　overhead

168 people
(4~75 years old)



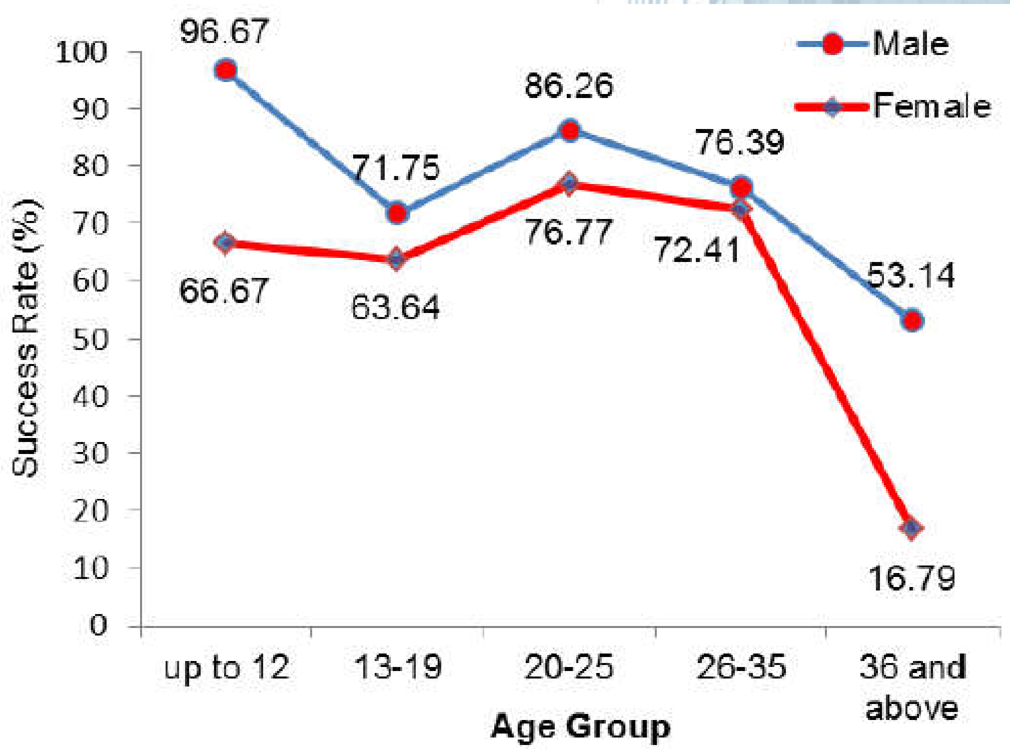
CCR: 94%

[1] Y. Makihara, H. Mannami and Y. Yagi, "Gait Analysis of Gender and Age Using a Large-Scale Multi-view Gait Database", In Proc. Asian Conference on Computer Vision, pp. 440-451, 2011

Age from Fingerprint

Database Statistics

Age Group	Male	Female	Total
Up to 12	70	60	130
13-19	190	320	510
20-25	1050	680	1730
26-35	320	270	590
36 and above	350	260	610
Total Samples	1980	1590	3570

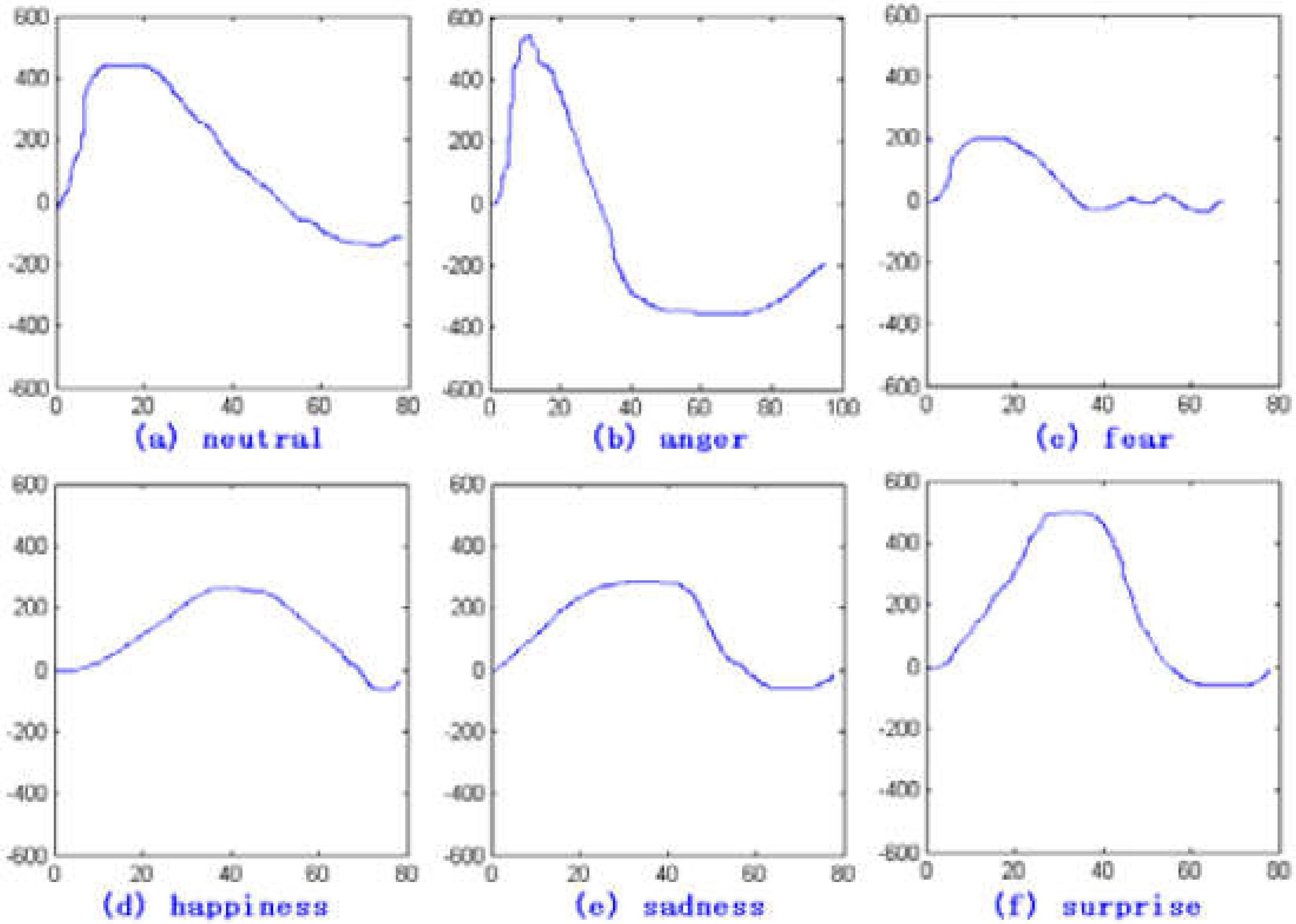


[1] P. Gnanasivam and D.S. Muttan. "Estimation of age through fingerprints using wavelet transform and singular value decomposition." International Journal of Biometrics and Bioinformatics, vol. 6, no. 2, pp. 58-67, 2012.

Affect from Keystroke

www.ia.ac.cn

Keystrokes corresponding to different affect states



[1] H.R. Lv, Z.L. Lin, W.J. Yin and J. Dong, "Emotion recognition based on pressure sensor keyboards.“, in Proc. IEEE Internation Conference on Multimedia and Expo, 2008.

Conclusions

- **Biometric data is becoming ubiquitous with fast development of mobile and wearable devices, social media, surveillance networks and identification systems.**
- **Biometric data can be mined to obtain a wide variety of information including identity, gender, ethnicity, age and affect.**
- **Great opportunities exist in transforming big biometric data to many killer apps.**
- **Many open problems remain to be solved in biometric data analysis. Compared with biometric identification, there is relatively less research on demographic and affective information prediction from biometric data.**

Thank you!

Q & A