# Advanced Techniques for Face-based Biometrics

**Massimo Tistarelli**
*Computer Vision Laboratory*
*University of Sassari – Italy*
*tista@uniss.it*

# The Computer Vision Lab



# The Computer Vision Lab

Vision Lab


# The Computer Vision Lab

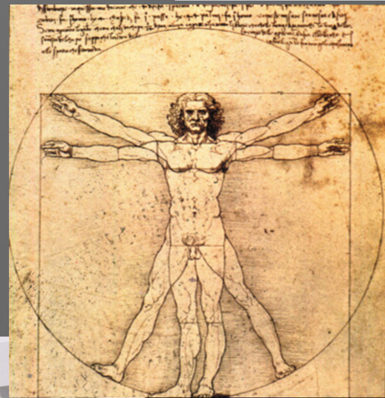
Vision Lab

- Second European Conference on Computer Vision
  May 1992 – *ECCV '92* – S. Margherita Ligure, Italy
- Int.l Workshop on Facial Image Analyisis and Recognition Technology
  May 1998 – Freiburg, Germany
- Int.l Workshop on Biometric Authentication
  June 2002 Copenhagen, Denmark
- Int.l Summer School for Advanced Studies on *Biometrics:*
  - *1st Authentication and Recognition* – June 2-6 2003
  - *2nd Multimodality and System Integration* – June 6-10 2005
  - *3rd New Sensors, Databases and Evaluation* – June 12-16 2006
  - *4th New Technologies and Embedded Systems* – June 11-15 2007
  - *5th New Technologies for Security and Privacy* – June 9-13 2008
  - *6th Multimodality and Identity Manegemnt* – June 8-12 2009
- Fifth IEEE AutoId Workshop – June 7-8 2007, Alghero, Italy
- Third IAPR/IEEE Int.l Conference on Biometrics
  June 2-5 2009, Alghero, Italy

# Credits

- The laboratory staff:
  - Manuele Bicego
  - Gavin Brelstaff
  - Linda Brodo
  - Marinella Cadoni
  - Massimo Gessa
  - Enrico Grosso
  - Andrea Lagorio
  - Ajita Rattani
  - Elif Surer

# Biometrics



- **Biometrics (as in statistics):** "Statistical study of biological observations and phenomena" (*Webster*)

- **Biometrics (as in security industry):** "A measurable, physical characteristic or personal behavioral trait used to recognize the identity, or verify the claimed identity, of an enrollee"

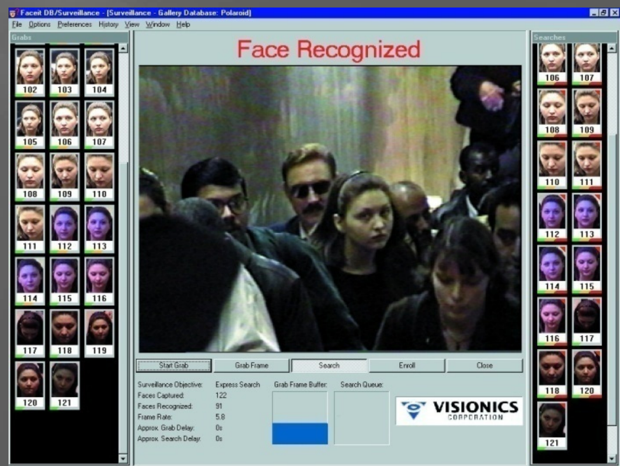
- ***Biometric recognition**: Personal recognition based on "who you are or what you do" as opposed to "what you know" (password) or "what you have" (ID card)*

Vision Lab

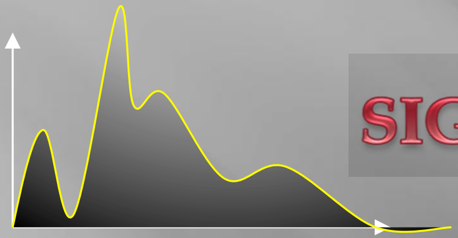
# *Why* face recognition?

- Most natural for humans
- Highly acceptable and non-intrusive
- Highly applicable:
  - Static identity verification
  - Uncontrolled face detection and identification from video


- Medium to High performances
- Not unique (twins)
- Aging and time effects



# How face recognition?



SIGNAL PROCESSING

*Information Theory*

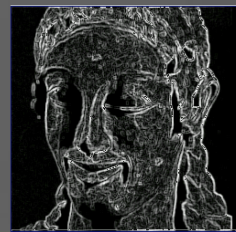
IMAGE PROCESSING

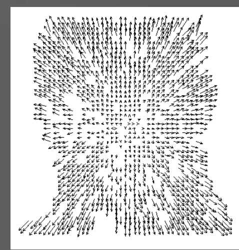
*Pattern Recognition*

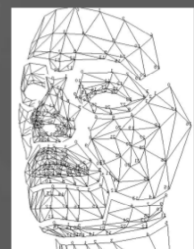
COMPUTER VISION

*Artificial Intelligence*

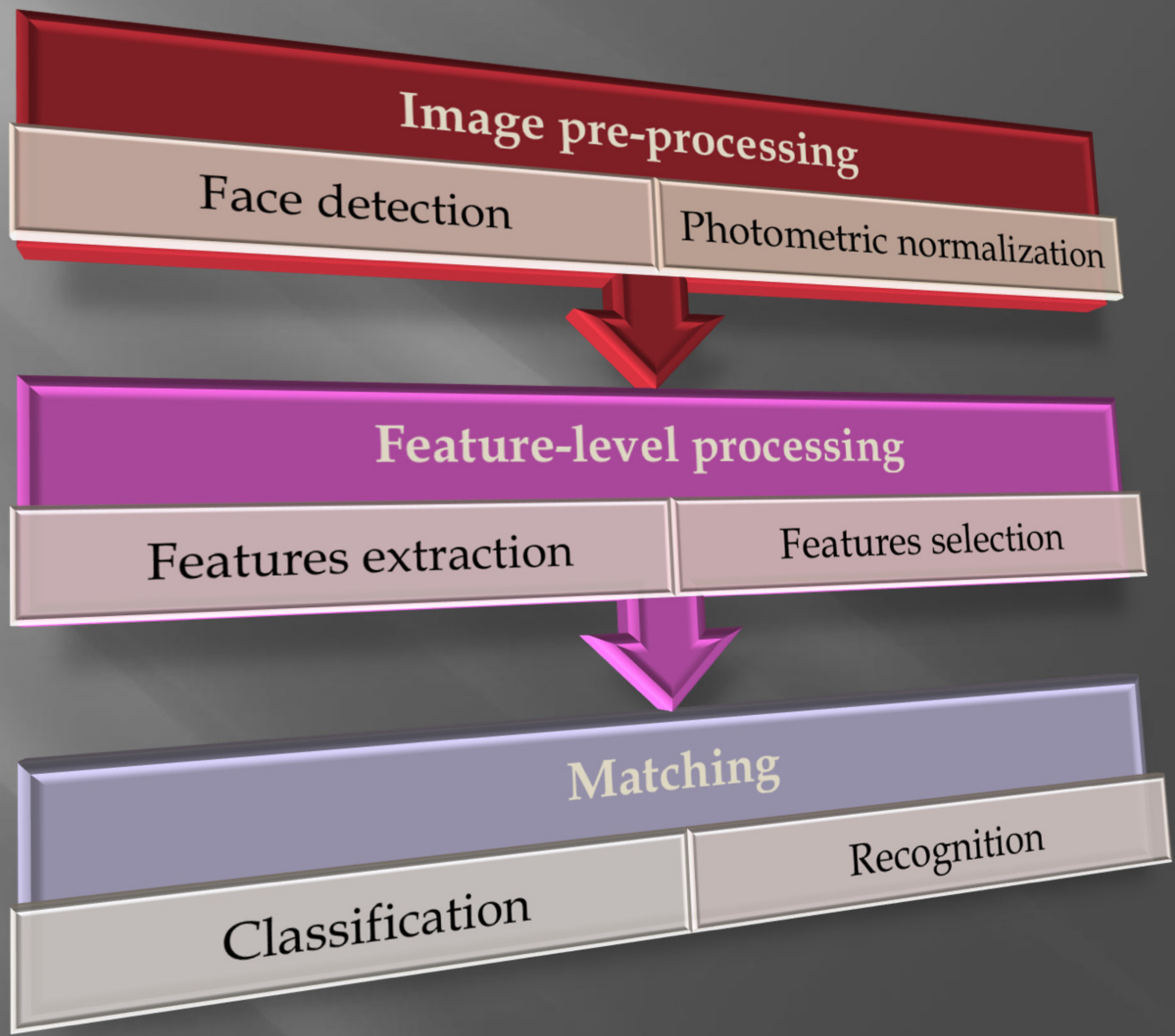
IMAGE UNDERSTANDING







# Fundamentals



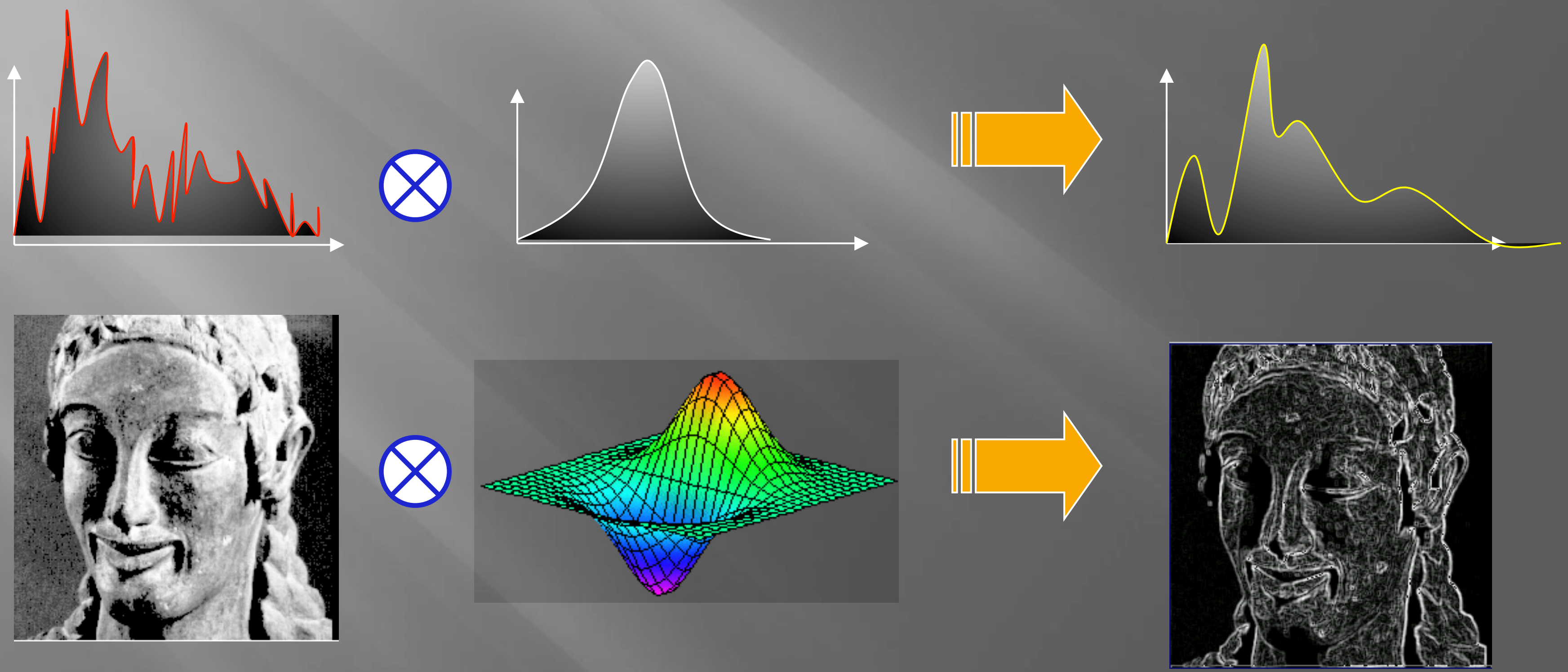
# Basic Image Processing techniques

Vision Lab

- **Convolution and filtering**
  - frequency-tuned feature extraction; edge detection; spatial and temporal derivatives; smoothing
  - Morphological filtering (*base operators; peaks, valleys and contour extraction*)
  - Correlation
- **Statistical image analysis**
  - image moments; gray level structures; Hidden Markov Models; histogram analysis; data classification
  - Hough transform (*grouping of geometrical structures*)
  - Clustering techniques (*data selection; set analysis*)
- **Color analysis**
  - Space conversion
  - Histogram analysis

# Basic Image Processing techniques

- **Convolution:** allows differential analysis, smoothing, frequency-tuned filtering…

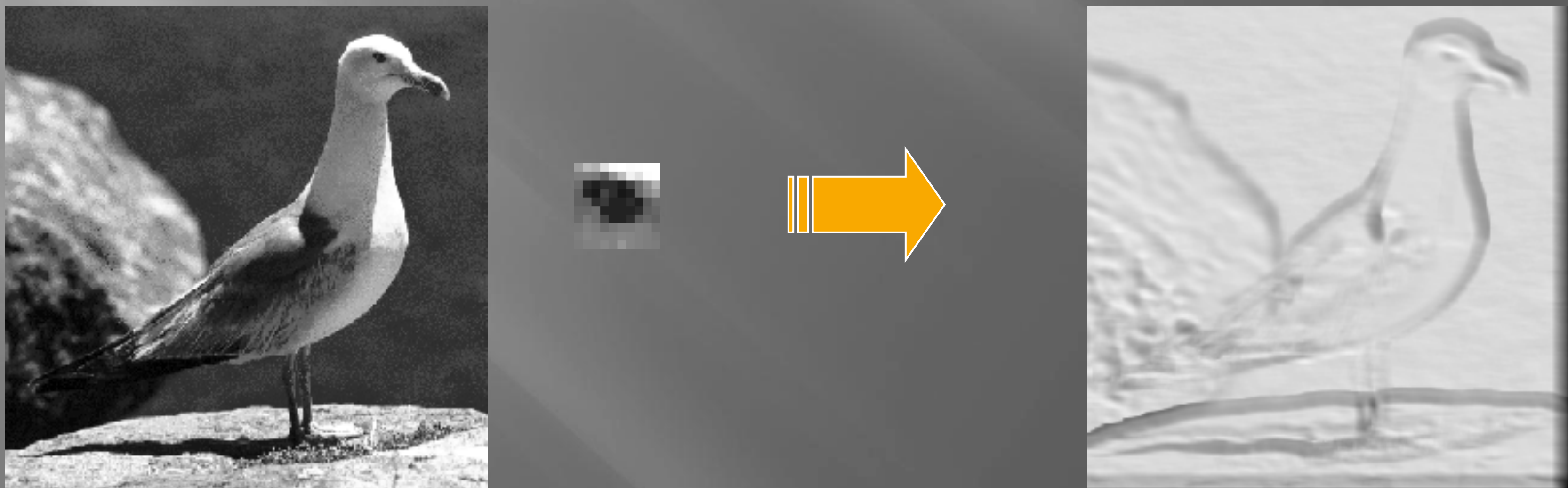


# Basic Image Processing techniques

- **Correlation:** pattern matching, cross-correlation…

$$z(x,y) = f(x,y) \otimes g(x,y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(u,v)g(x+u,y+v)dudv.$$

$$z(x,y) = f(x,y) \otimes g(x,y) = \frac{1}{MN}\sum_{x'=0}^{M-1}\sum_{y'=0}^{N-1} f(x',y')g(x+x',y+y')$$



# Pattern Recognition

Vision Lab

"The real power of human thinking is based on recognizing patterns. The better computers get at **pattern recognition**, the more humanlike they will become".

*Ray Kurzweil, NY Times, Nov 24, 2003*

# Classification and learning

- The regression function $f^*(x) = \mathbb{E}[y|x]$ minimizes:

$$\mathbb{E}\left[|y - f(x)|^2\right]$$

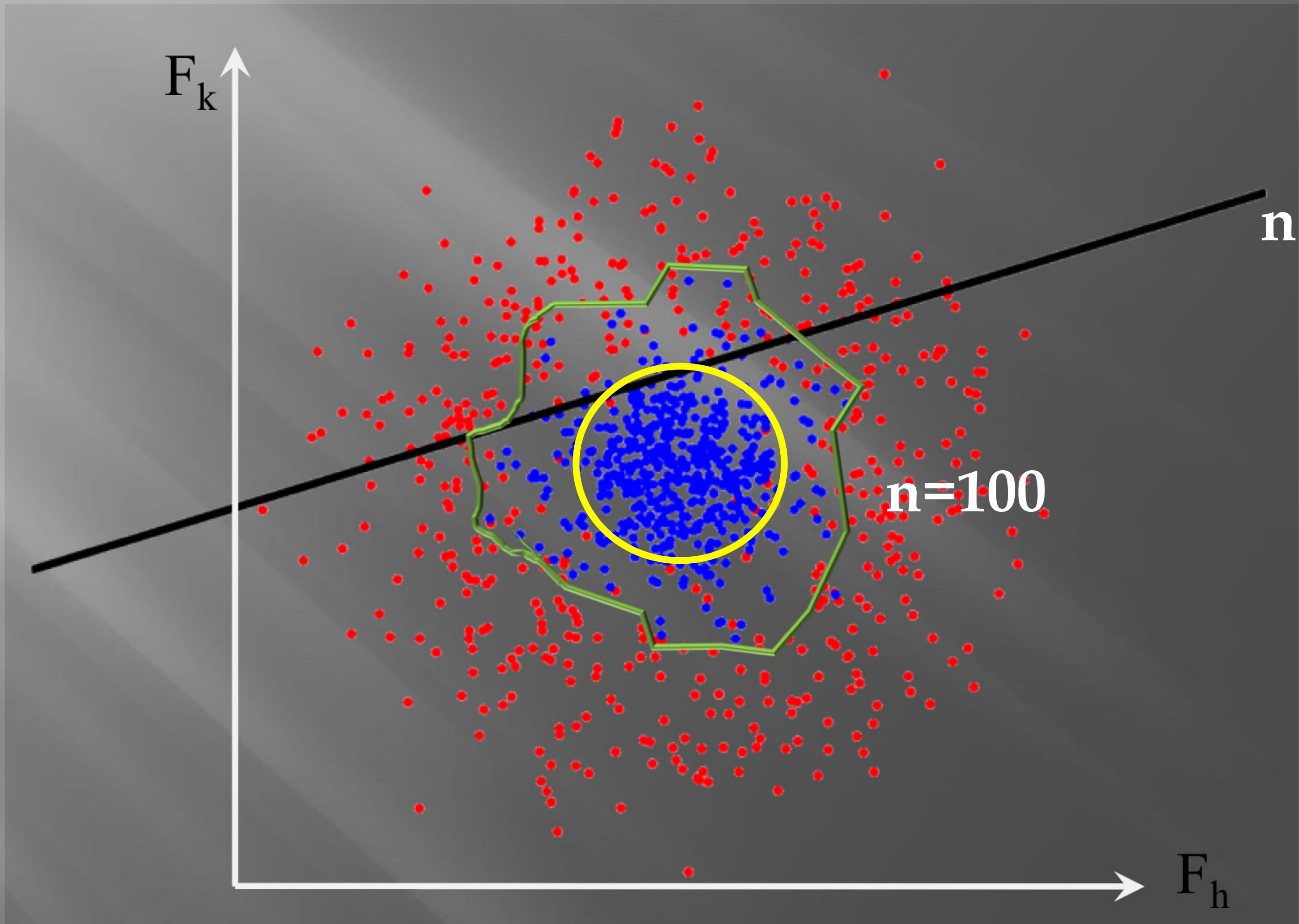
- The Bayes rule $f_b(x) = \mathrm{sign}(f^*(x))$ minimizes:

$$P(yf(x) \leq 0)$$

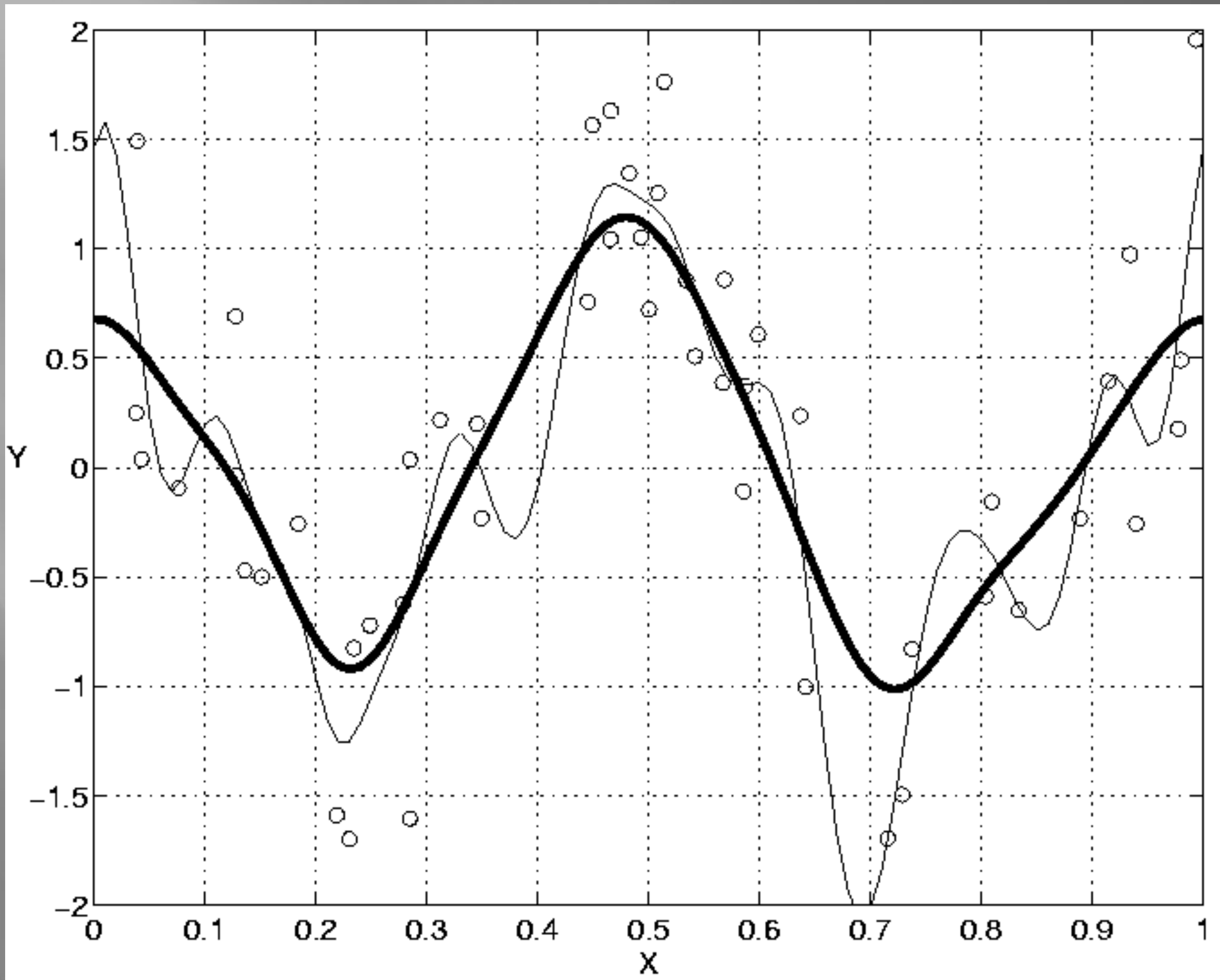
given $(x_1, y_1), \ldots, (x_n, y_n)$ find an estimator $f_n$ such that with high probability

$$\lim_{n \to \infty} \int_X (f_n(x) - f^*(x))^2 p(x) dx = 0$$

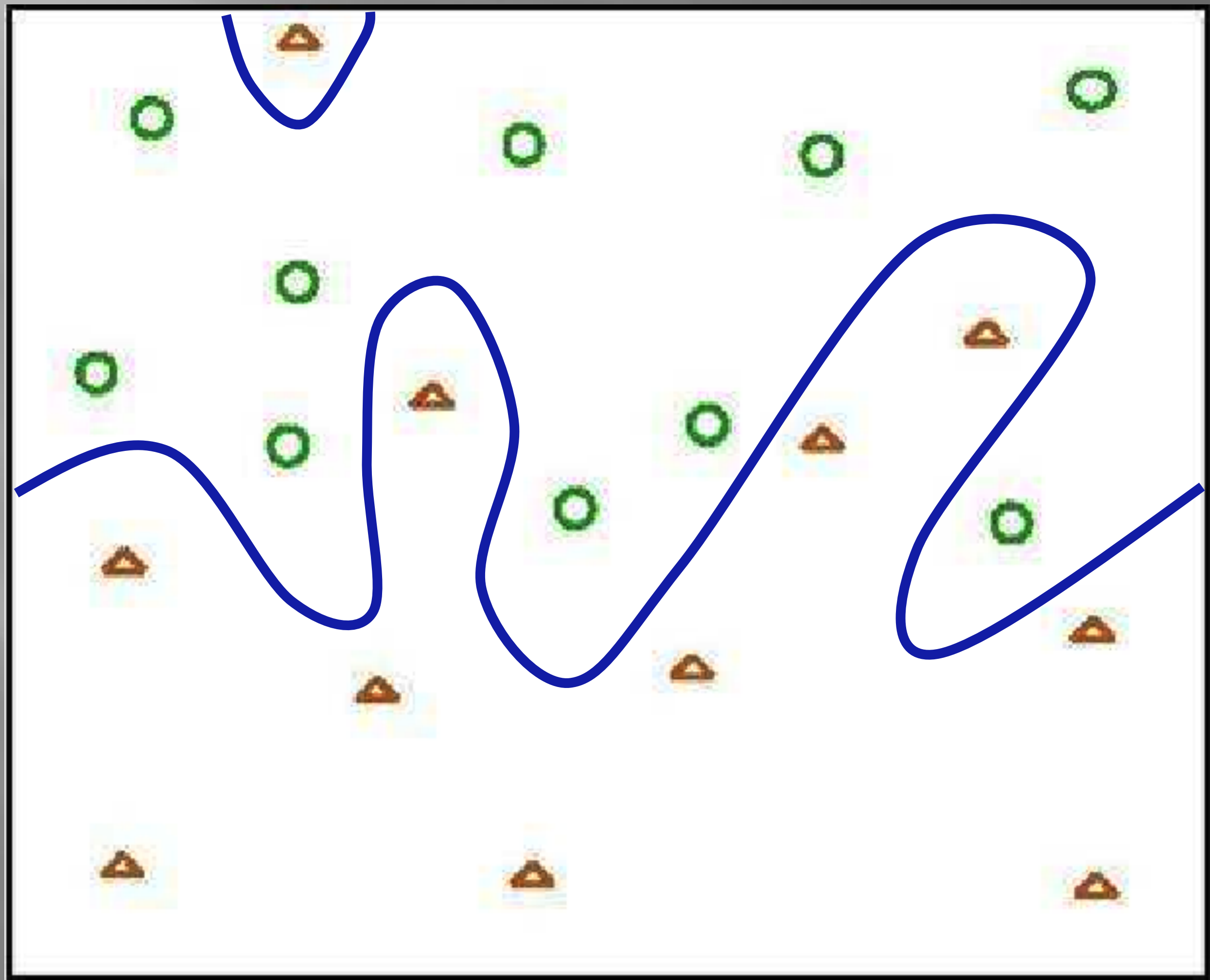
# Classification and learning



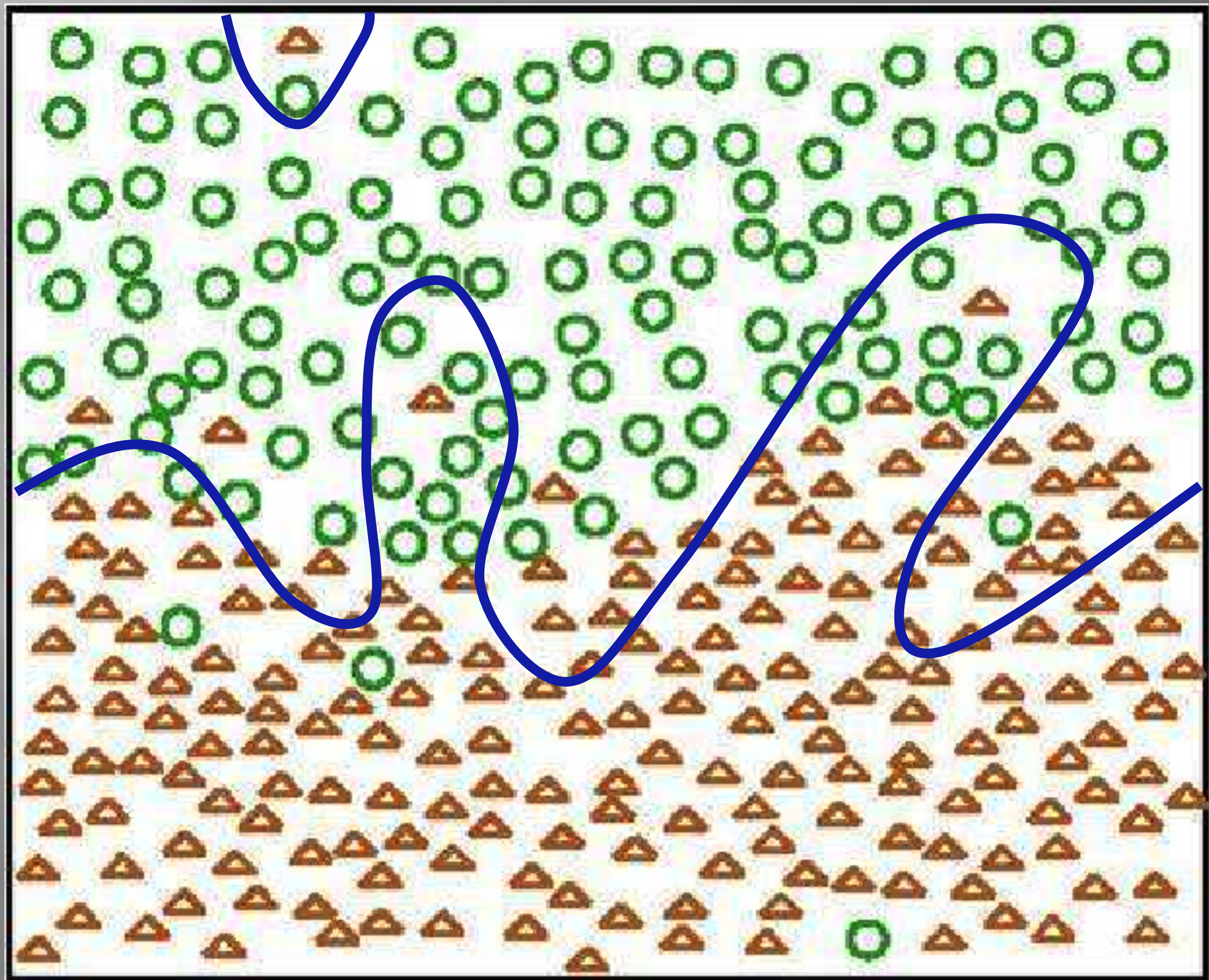
# Classification and learning



# Classification and learning



# Classification and learning



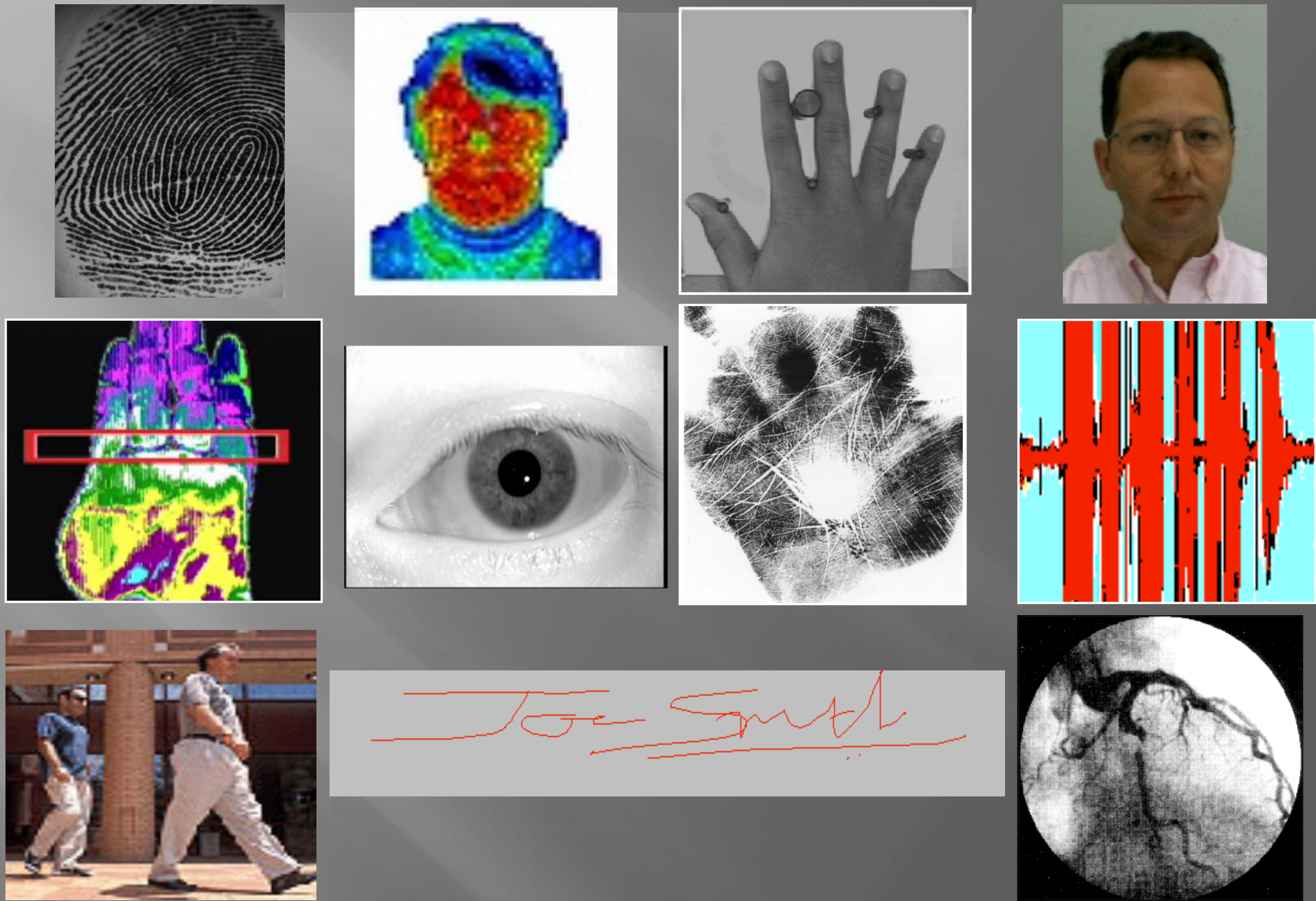
# Must Read!

Vision Lab

- D.H. Ballard and C.M. Brown *Computer Vision (http://homepages.inf.ed.ac.uk/rbf/BOOKS/BANDB/bandb.htm)*
- W.K. Pratt *Digital Image Processing*
- B.K.P. Horn *Robot Vision*
- A.K. Jain *and* S. Lee *Handbook of Face Recognition*
- E. Trucco *and* A. Verri *Introductory Techniques for 3D Computer Vision*
- J. Bigun *Vision with direction*

# FACE AS BIOMETRICS

# Biometric traits



# Biometric traits

- **Universality** (everyone should have this trait)
- **Uniqueness** (no two persons should have the same trait)
- **Permanence** (should be invariant with time)
- **Collectability** (can be measured quantitatively)
- **Performance** (achievable recognition accuracy, resources required, operational/environmental factors)
- **Acceptability** (to what extent people are willing to accept it)
- **Circumvention** (how easy it is to fool the system)

# An *almost* «fair» comparison

Vision Lab

*(from Jain et al. 1998)*

| BIOMETRICS | Universality | Uniqueness | Permanence | Collectability | Performance | Acceptability | Circumvention |
|---|---|---|---|---|---|---|---|
| Face | High | Low | Medium | High | Low | High | Low |
| Fingerprint | Medium | High | High | Medium | High | Medium | Low |
| Hand Geometry | Medium | Medium | Medium | High | Medium | Medium | Medium |
| Iris | High | High | High | Medium | High | Low | High |
| Retinal Scan | High | High | Medium | Low | High | Low | High |
| Signature | Low | Low | Low | High | Low | High | Low |
| Voice | Medium | Low | Low | Medium | Low | High | Low |
| Facial Thermogram | High | High | Low | High | Medium | High | High |

# Genotypic vs phenotipic traits

Vision Lab

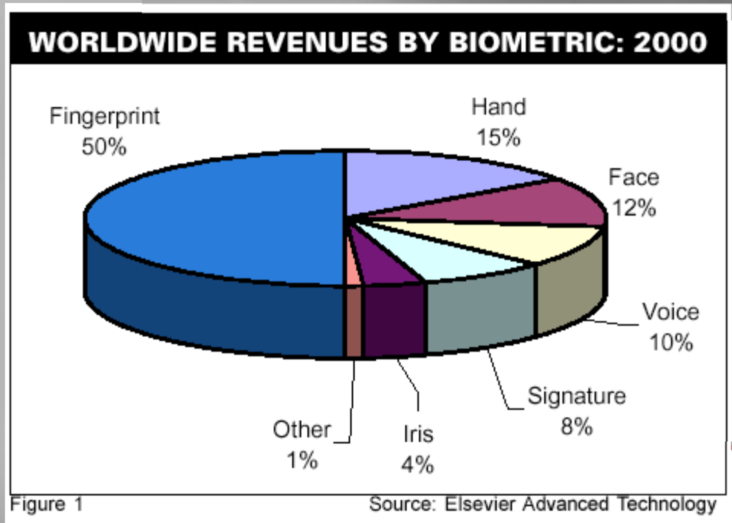
**Biometric traits develop:**

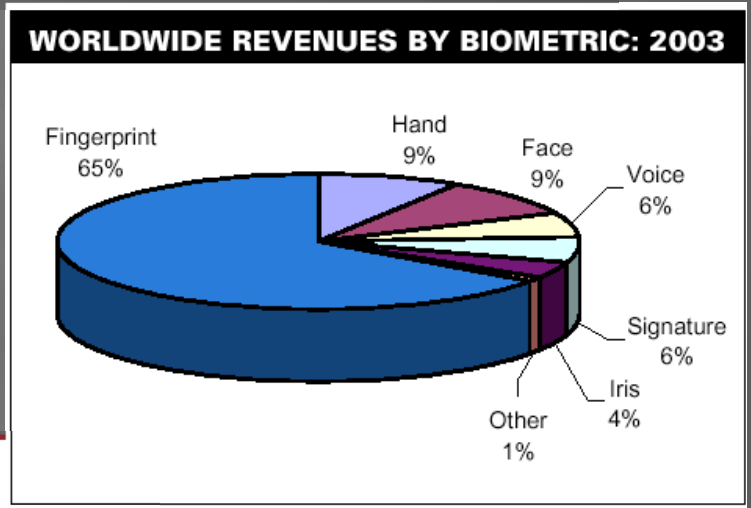
1. through genetics: **Genotypic**
2. through random variations in the early phases of an embryo's development: **Phenotypic**
3. through training: **Behavioral**

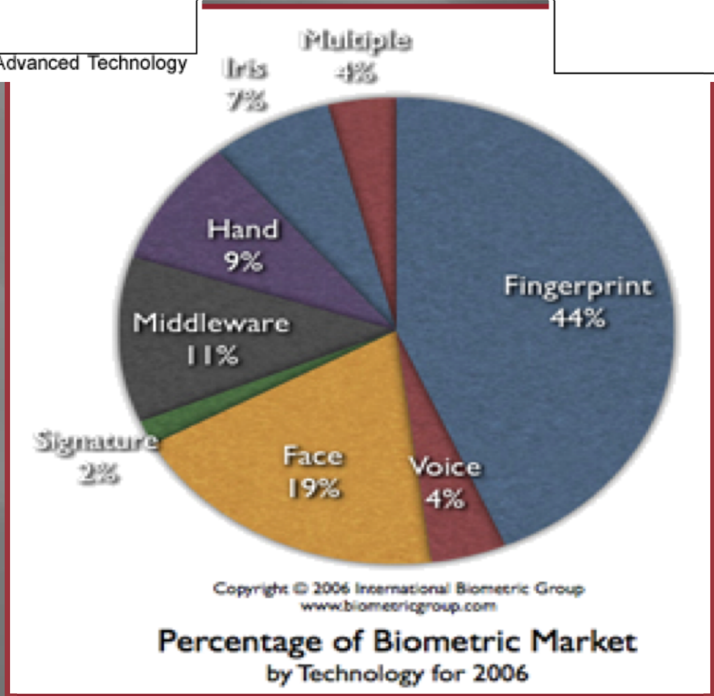
| *Biometric Trait* | *genotypic* | *phenotypic* | *behavioral* |
|---|---|---|---|
| Fingerprint (only minutia) | o | ooo | o |
| Signature (dynamic) | oo | o | ooo |
| **Facial geometry** | ooo | o | o |
| Iris pattern | o | ooo | o |
| Retina (Vein structure) | o | ooo | o |
| Hand geometry | ooo | o | o |
| Finger geometry | ooo | o | o |
| Vein structure of the back of hand | o | ooo | o |
| Ear form | ooo | o | o |
| Voice (Tone) | ooo | o | oo |
| DNA | ooo | o | o |
| Odor | ooo | o | o |
| Keyboard Strokes | o | o | ooo |
| Comparison: Password | | | (ooo) |

**Source:** http://www.bromba.com/faq/biofaqe.htm#entstehen

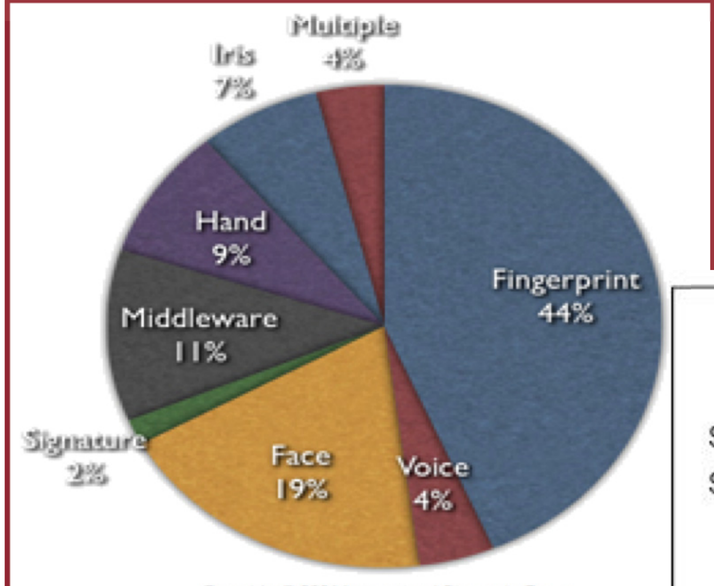
# Market share in biometrics



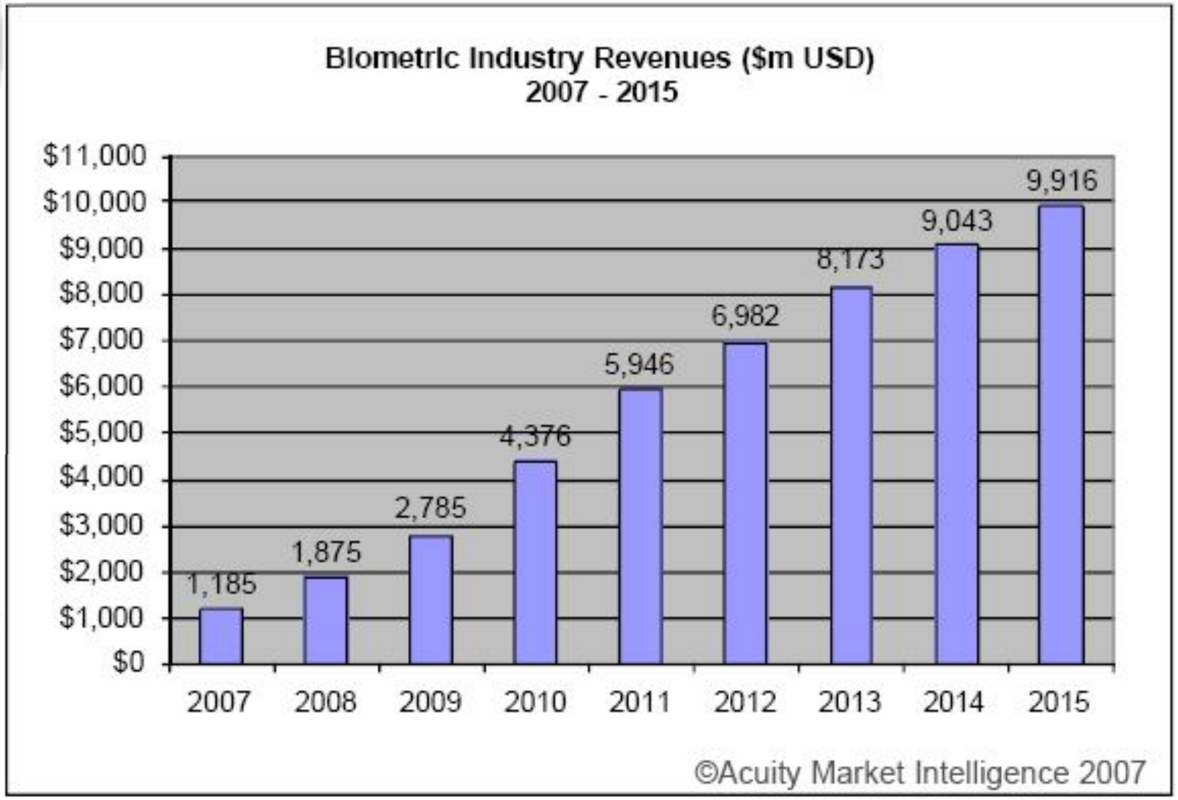



# Market share in biometrics



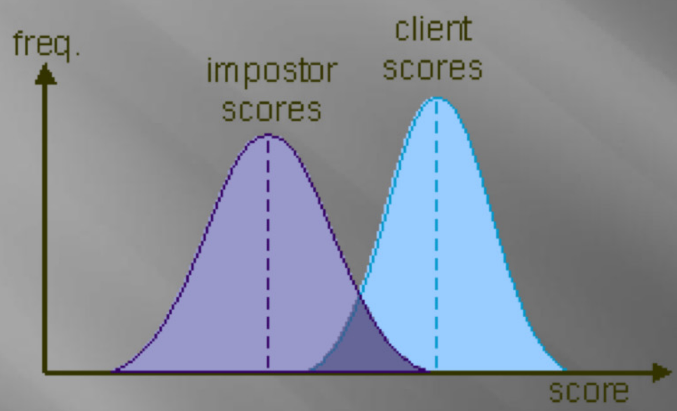

Percentage of Biometric Market
by Technology for 2006

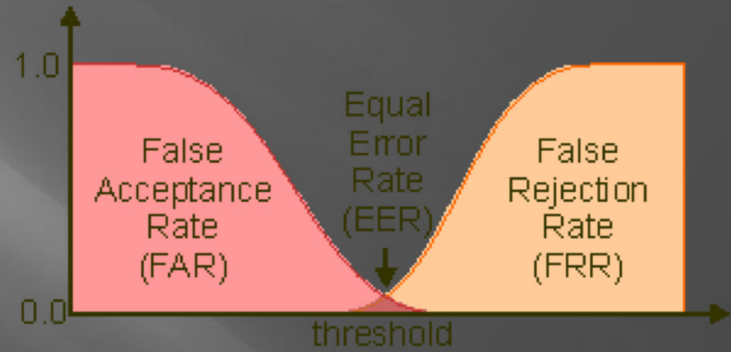


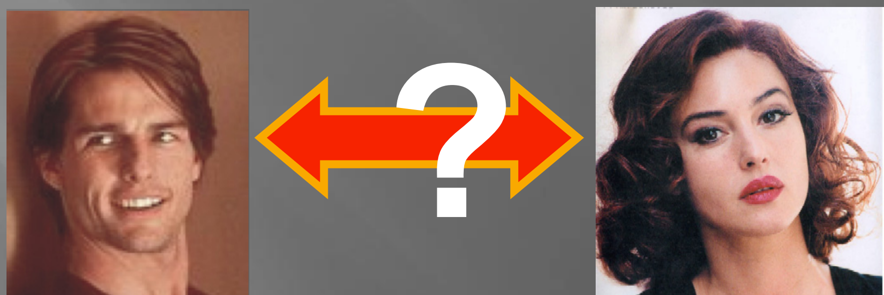
# FROM FACES TO IDENTITIES

# Classification of faces

## A class separation problem:



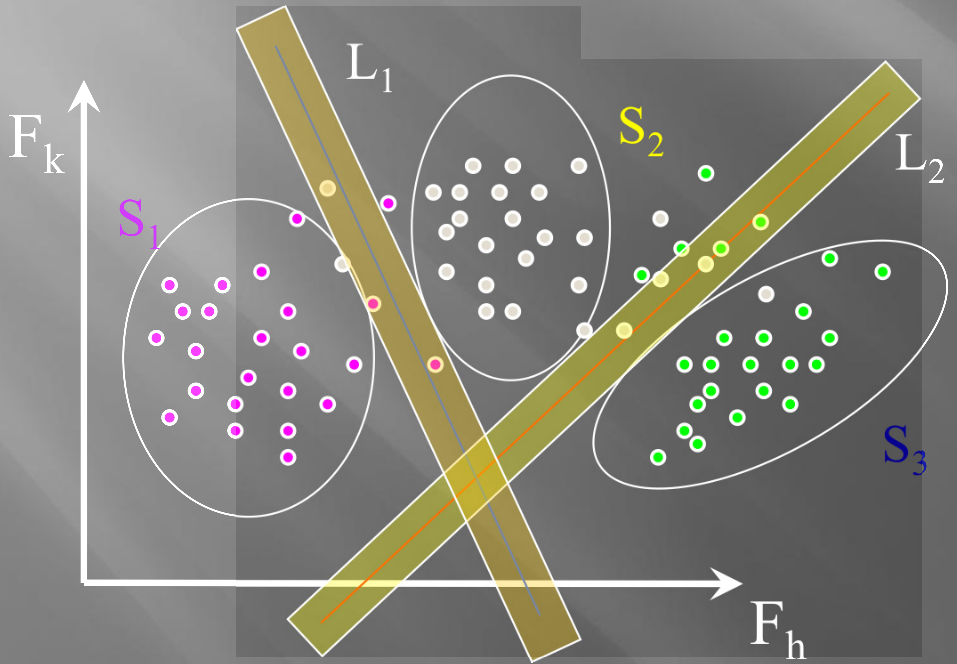



# Identification of faces

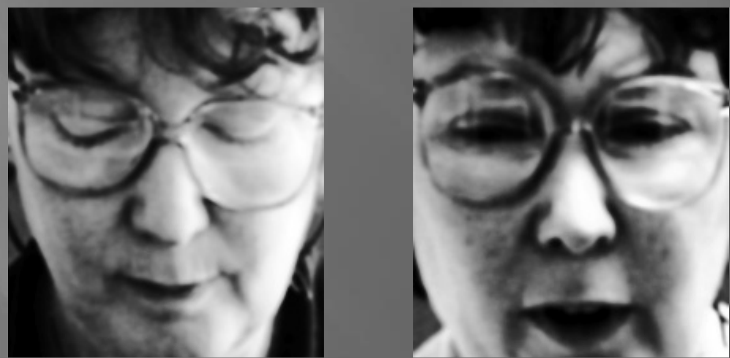
## A class (identity) separation problem:

- Choice of optimal representation
- Inter-class similarity vs intra-class variability




$= \Psi(F_h, F_k)$

# Noise in the raw data



# Inter-class similarity

**Two different people may have very similar appearance**

**FALSE ACCEPTANCE**

www.marykateandashley.com

**Twins**

news.bbc.co.uk/hi/english/in_depth/americas/2000/us_elections

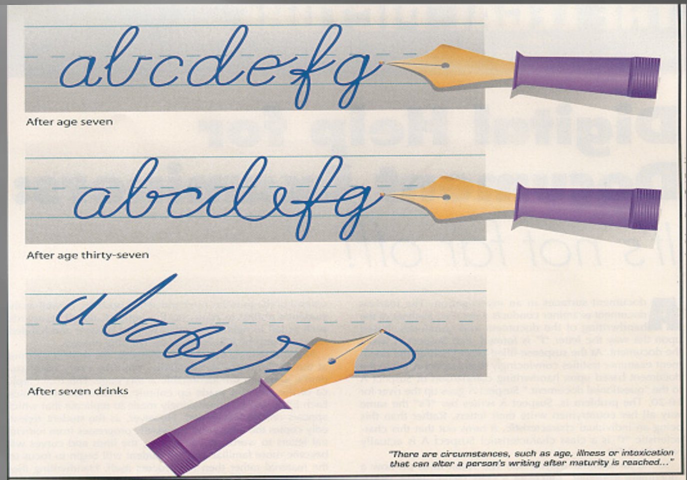
**Father and son**

Vision Lab

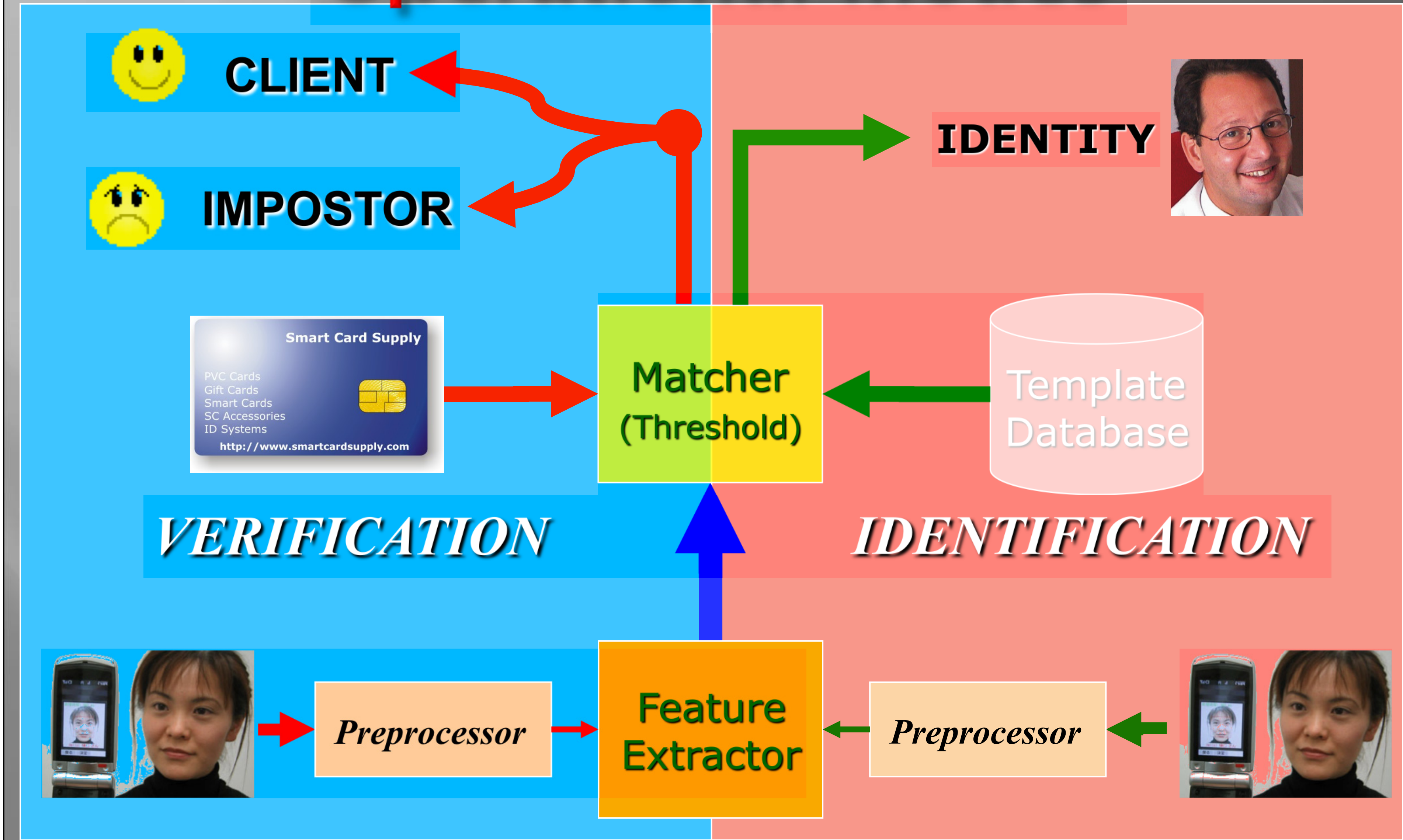
# Intra-class variability

**The same person may present very different biometric samples**

## FALSE REJECTION




# Operational modes



# Template

- A set of features extracted from the raw biometric data of an individual: a prototype of an individual's biometric
- Usually the template is stored along with some biographic details of the individual (e.g., user name, PIN)
- Central database vs. smart cards
- Template selection and update (aging)

# Recognition and Authentication

- Many facts come into play:
  - data features
  - noise in the data
  - data alignment/localizazion
  - acquisition device
  - algorithmic complexity and discrimination power

# FACE RECOGNITION PERFORMANCES

# Performance evaluation

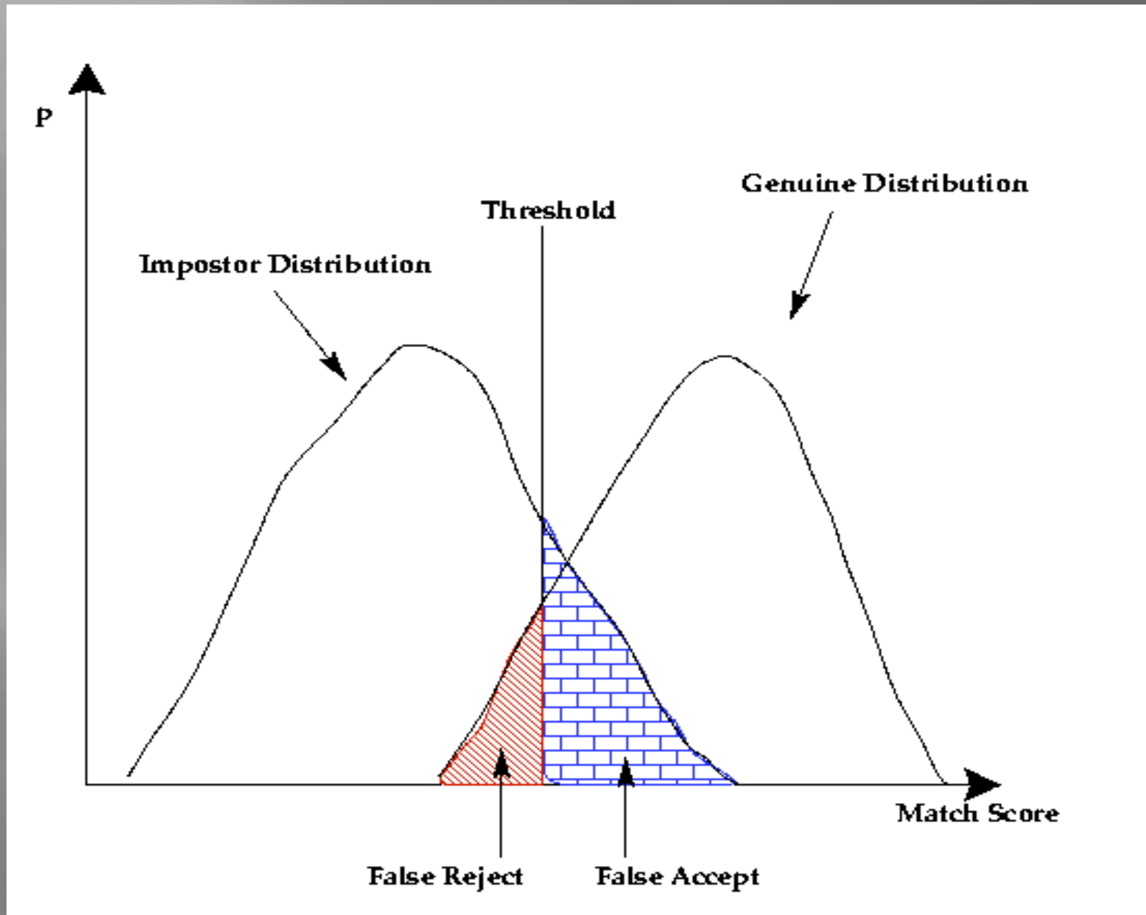
- The overall performance of a biometric system is assessed in terms of its **accuracy**, **speed**, and **storage**
- Factors like cost and ease of use also affect efficacy
- **Biometric systems are not perfect**, and will sometimes mistakenly accept an impostor as a valid individual (a false match) or conversely, reject a valid individual (a false non-match)

**Best Practices**: http://www.cesg.gov.uk/technology/
**FRVT2002**: www.frvt.org/FRVT2002/documents.htm
**FVC 2002**: bias.csr.unibo.it/fvc2002
**NIST SV**: www.nist.gov/speech/tests/spk

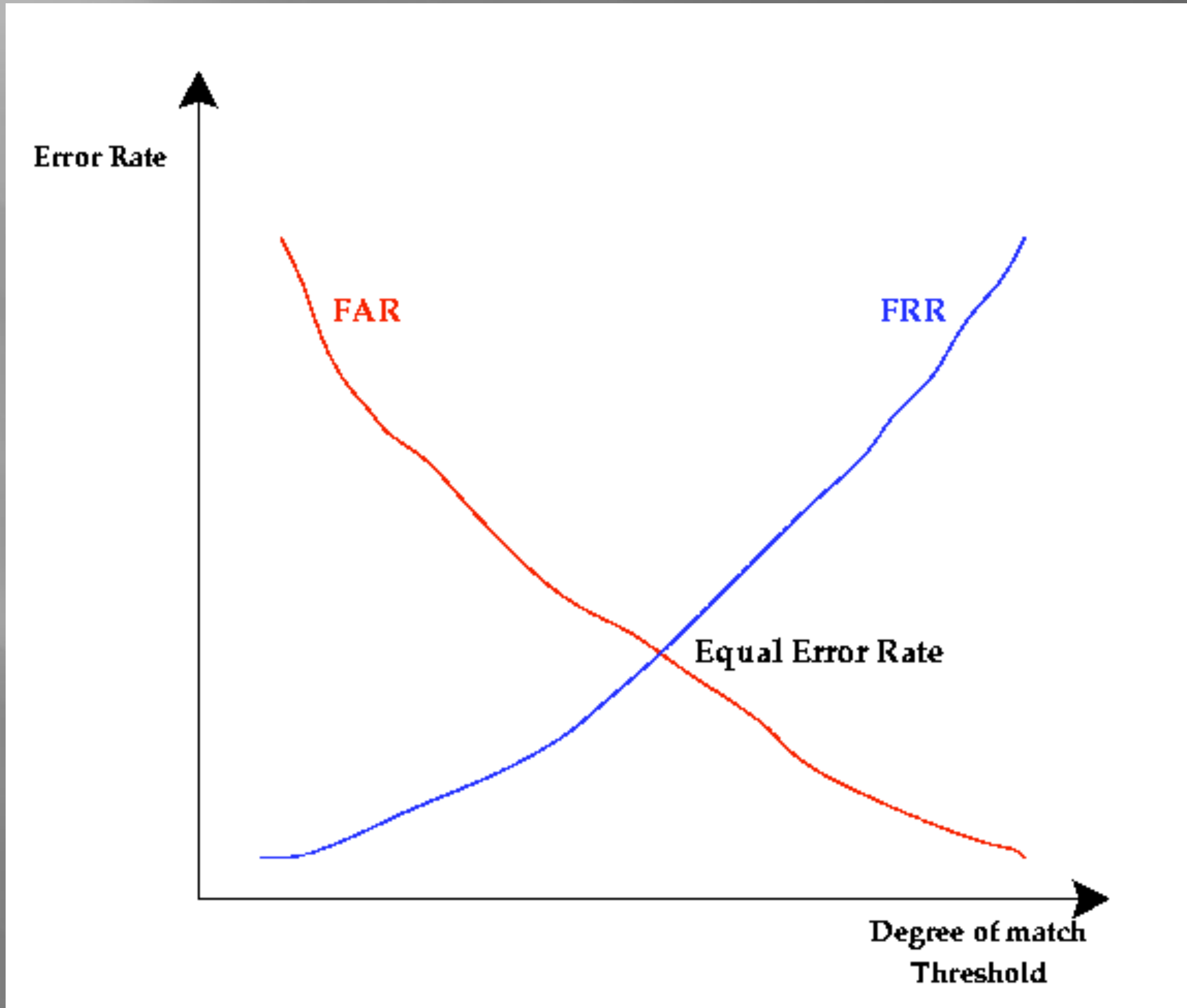
# Performance evaluation

- Failure to **Enroll**
- Failure to **Acquire**
- Impostor and Genuine **Distributions**
- Matching **Threshold**
- False Accept Rate (**FAR**) or False Match Rate
- False Reject Rate (**FRR**) or False Non-match Rate
- Receiver Operating Characteristic (**ROC**) curve
- **Equal Error Rate** or Crossover Rate

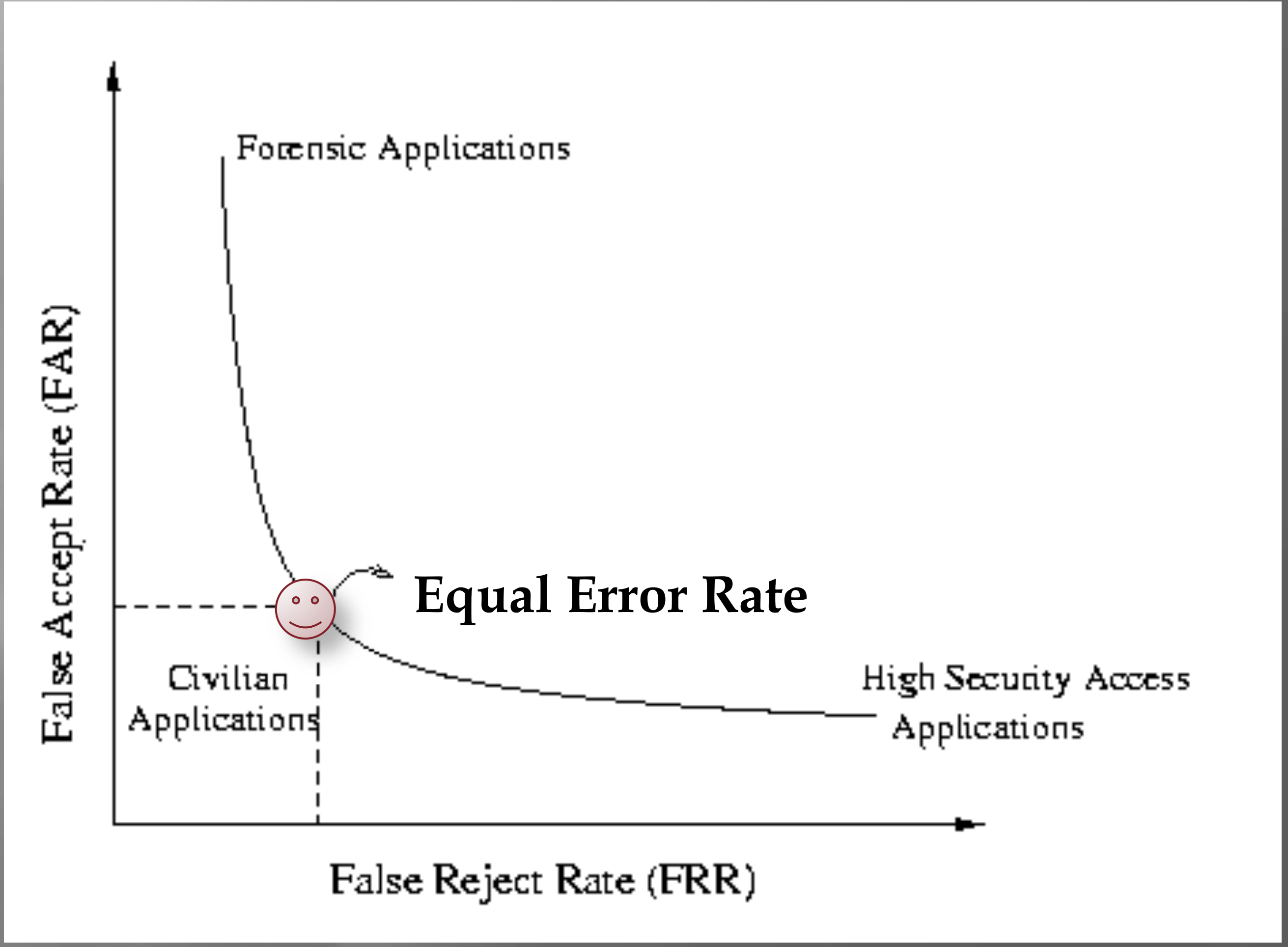
# Probability distributions



# Errors and thresholds



# Receiver Operating Characteristic (ROC) curve



# Evaluation and testing

- How do you test the system, select the database and measure the performance?
- Evaluations should be conducted by independent groups (bias.csr.unibo.it/fvc2006/)
- Size of the database and representative examples of the database should be made available prior to testing
- Test on biometric data previously unseen by the system
- Voice, face and fingerprint systems have undergone the most study and testing

# Face databases

**The appearance of a face is affected by many factors**

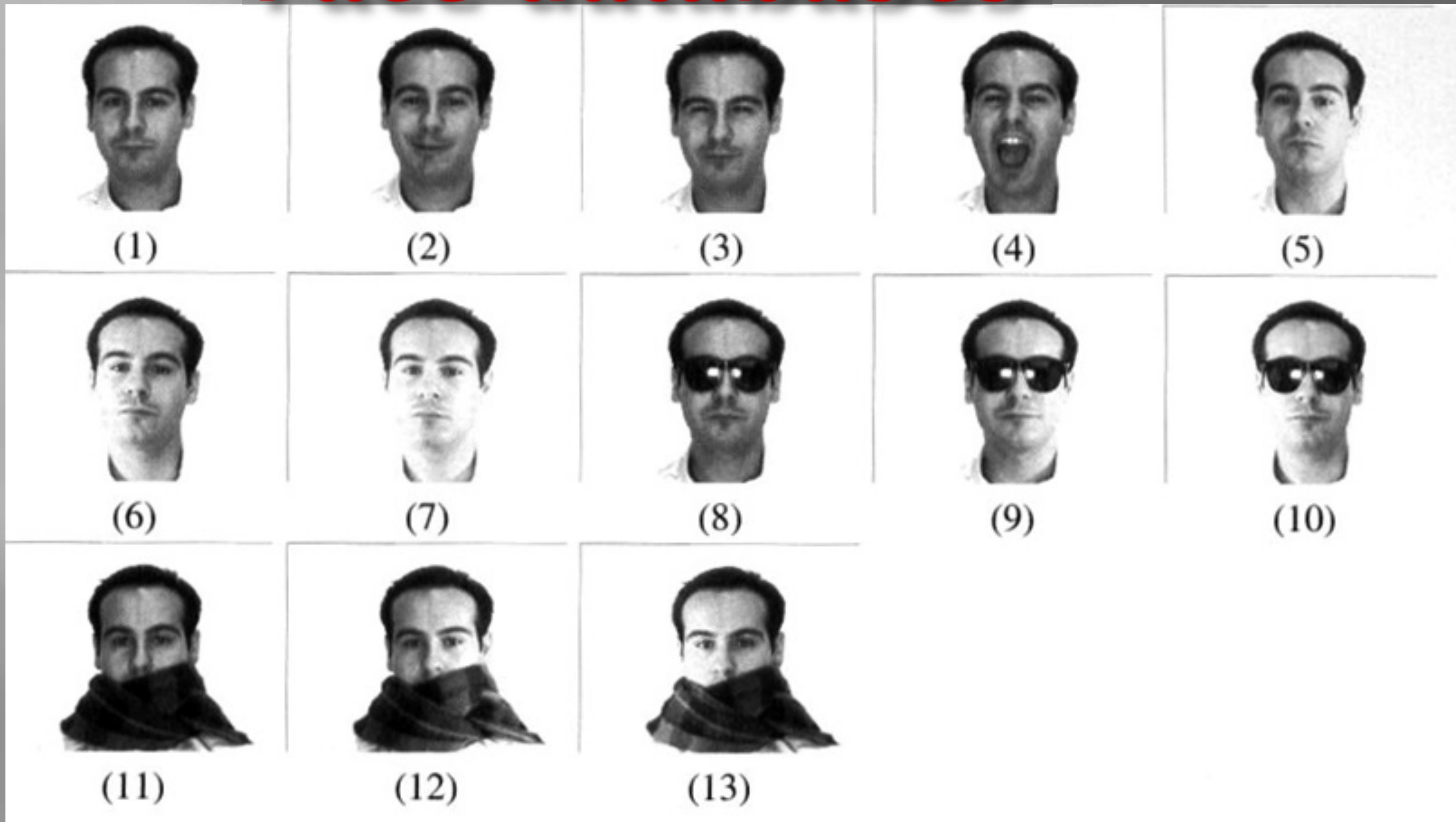
- Identity - Age
- Face pose - Occlusion
- Illumination - Facial hair
- Facial expression

The development of algorithms robust to this variations requires databases of sufficient size that include carefully controlled variations of these factors.

Common databases are necessary to comparatively evaluate algorithms.

Collecting a high quality database is a resource-intensive task.

# Face databases




**AR database.** The conditions are (1) neutral, (2) smile, (3) anger, (4) scream, (5) left light on, (6) right light on, (7) both lights on, (8) sun glasses, (9) sun glasses/left light (10) sun glasses/right light, (11) scarf, (12) scarf/left light, (13) scarf/right light

# Face databases

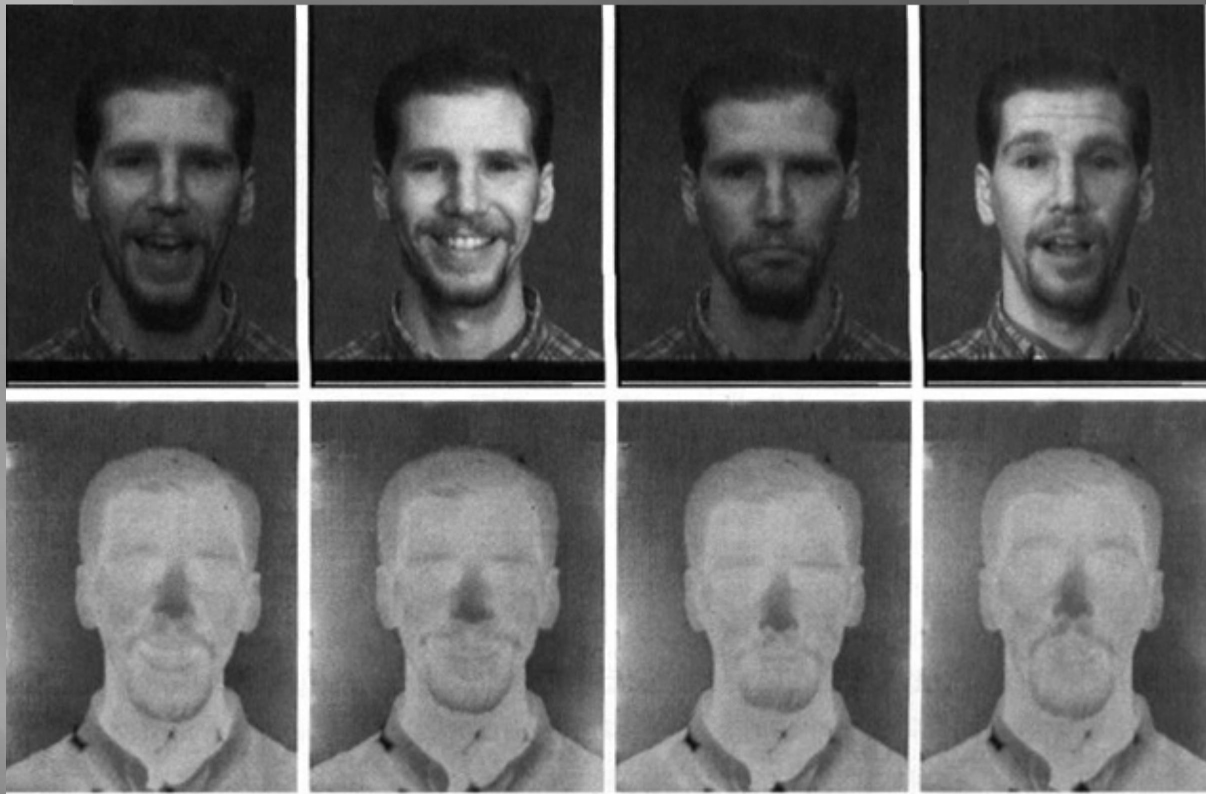
**CAS-PEAL database.** The images were recorded using separate cameras triggered in close succession. The cameras are about $22.5^0$ apart. Subjects were asked to look up, to look straight ahead, and to look down. Shown here are seven of the nine poses currently being distributed.

# Face databases

**Illumination variation in the CAS-PEAL database.** The images were recorded with constant ambient illumination and manually triggered fluorescent lamps.

Also the **CMU PIE database** has been designed to include illumination variations

# Face databases

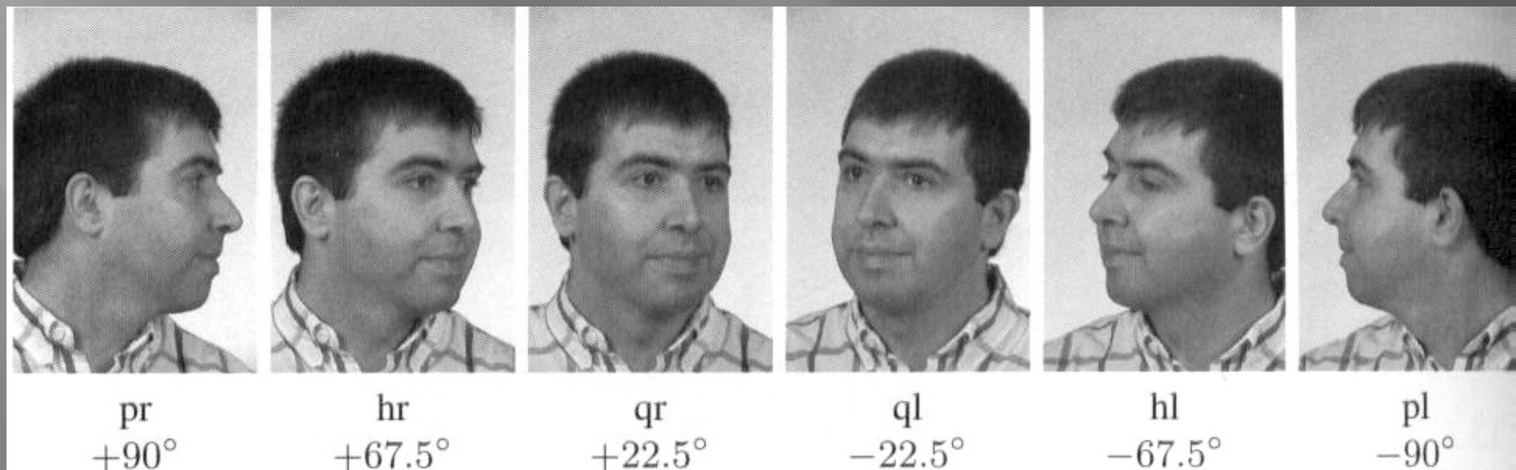


(a) (b) (c) (d)

**Equinox IR database**. The upper row contains visible images and the lower row long-wave infrared images. The categories are (a) vowel (frontal illumination), (b) “smile” (right illumination), (c) “frown” (frontal illumination), (d) “surprise” (left illumination).

# Face databases


**Frontal image categories used in the FERET evaluations**. For images in the fb category, a different facial expression was requested. The fc images were recorded with a different camera and under different lighting conditions. The duplicate images were recorded in a later session, with 0 and 1031 days (duplicate I) or 540 to 1031 days (duplicate II) between recordings.



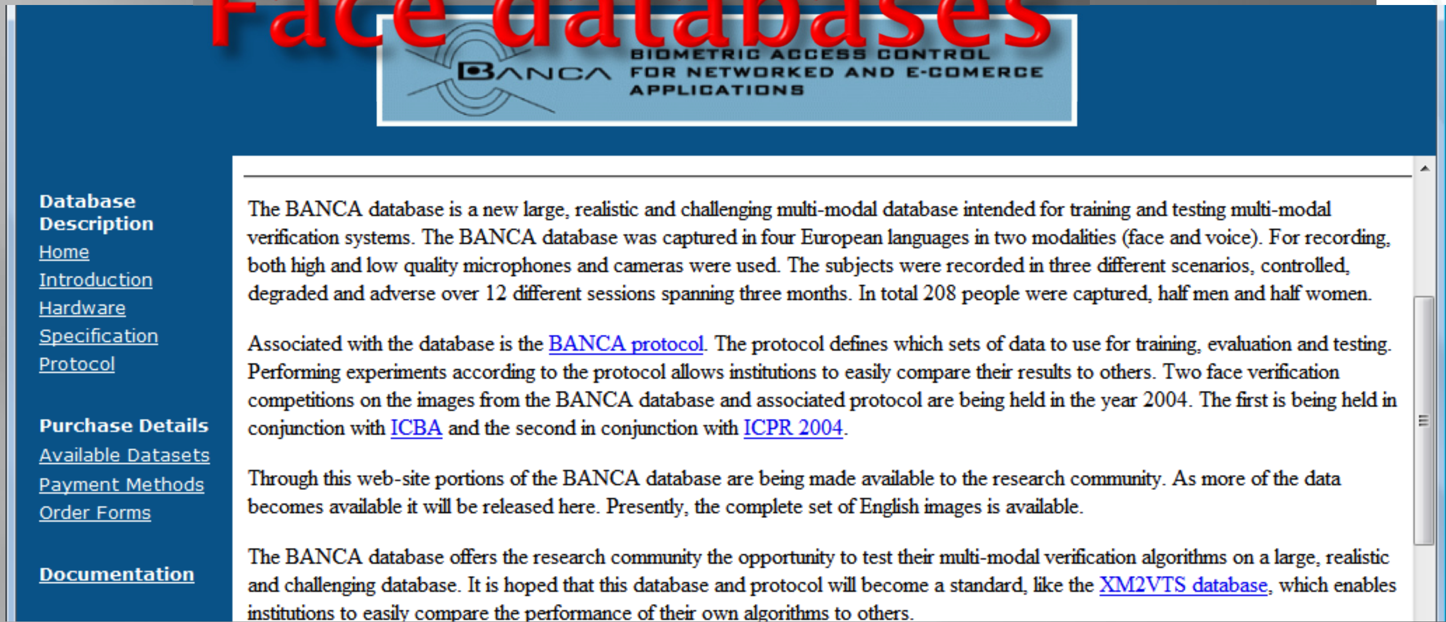

Additional set of pose images from the FERET database: right and left profile (labeled pr and pl), right and left quarter profile (qr, ql), and right and left half profile (hr, hl).

# Face databases

**Notre Dame HumanID database.** Example images of the "unstructured" lighting condition recorded in the hallway outside of the laboratory.

**University of Texas Video Database.** Example images for the different recording conditions of the database. First row: Facial speech. Second row: Laughter. Third row: Disgust.

# Face databases


BIOMETRIC ACCESS CONTROL FOR NETWORKED AND E-COMERCE APPLICATIONS

**Database Description**
Home
Introduction
Hardware
Specification
Protocol

**Purchase Details**
Available Datasets
Payment Methods
Order Forms

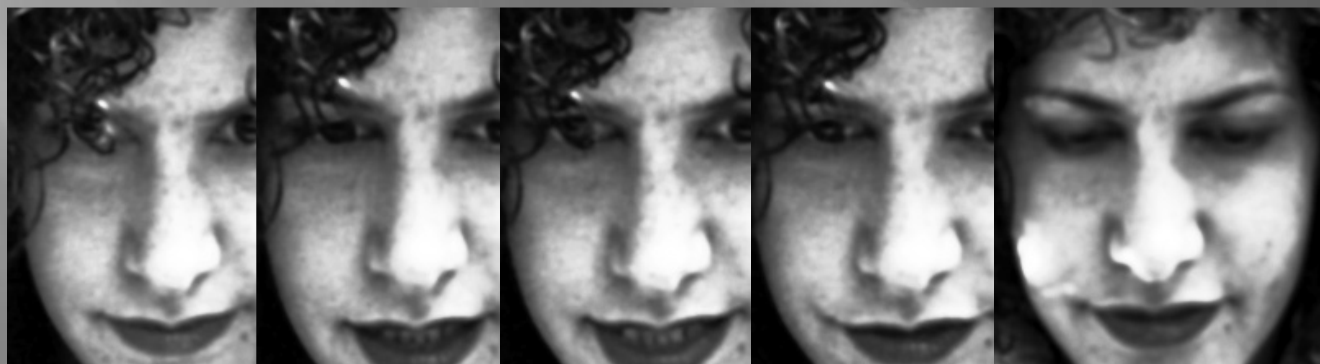
**Documentation**

The BANCA database is a new large, realistic and challenging multi-modal database intended for training and testing multi-modal verification systems. The BANCA database was captured in four European languages in two modalities (face and voice). For recording, both high and low quality microphones and cameras were used. The subjects were recorded in three different scenarios, controlled, degraded and adverse over 12 different sessions spanning three months. In total 208 people were captured, half men and half women.

Associated with the database is the BANCA protocol. The protocol defines which sets of data to use for training, evaluation and testing. Performing experiments according to the protocol allows institutions to easily compare their results to others. Two face verification competitions on the images from the BANCA database and associated protocol are being held in the year 2004. The first is being held in conjunction with ICBA and the second in conjunction with ICPR 2004.

Through this web-site portions of the BANCA database are being made available to the research community. As more of the data becomes available it will be released here. Presently, the complete set of English images is available.

The BANCA database offers the research community the opportunity to test their multi-modal verification algorithms on a large, realistic and challenging database. It is hoped that this database and protocol will become a standard, like the XM2VTS database, which enables institutions to easily compare the performance of their own algorithms to others.



**The BANCA and XM2VTS video databases distributed by the University of Surrey**

# Face databases

BIOMETRIC ACCESS CONTROL FOR NETWORKED AND E-COMERCE APPLICATIONS

## The BANCA Protocol

---

An evaluation protocol defines a set of data, how it should be used by a system to perform a set of experiments and how the system performance should be computed.

In verification, two types of protocols exist; closed-set and open-set. In closed-set verification the population of clients is fixed. This means that the system design can be tuned to the clients in the set. Thus both the adopted representation (features) and the verification algorithm applied in the feature space are based on some training data collected for this set of clients. Anyone who is not in the training set is considered an impostor. The XM2VTS protocol is an example of this type of verification problem formulation.

In open-set verification we wish to add new clients to the list without having to redesign the verification system. In particular, we want to use the same feature space and the same design parameters such as thresholds. In such a scenario the feature space and the verification system parameters must be trained using completely independent data from that used for specifying client models. The BANCA protocol is an example of an open-set verification protocol.

# Face detection databases

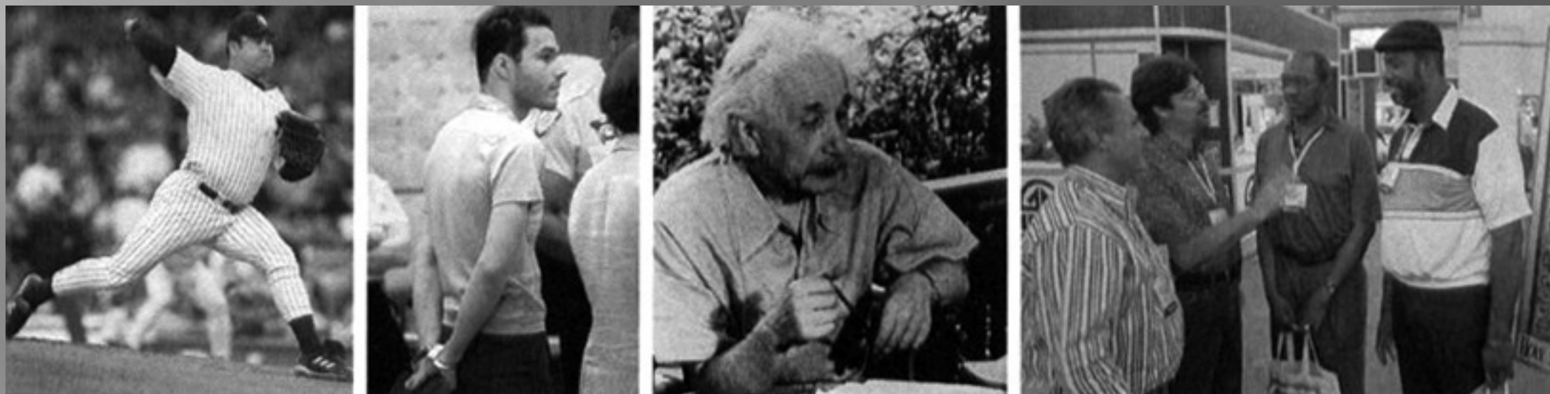
- Face detection algorithms typically have to be trained on face and non-faces images to build up an internal representation of the human face.
- Popular choices are the FERET, MIT, ORL, Harvard, and AR public databases. Nonpublic databases are often also employed.
- These data sets should be representative of real-world data containing faces of various orientations against a complex background.
- In recent years two public data sets emerged as quasi-standard evaluation test sets:
  - The combined MIT/CMU test set for frontal face detection
  - The CMU test set II for frontal and nonfrontal face detection

# Face databases

Example images from the Upright Test Set portion of the MIT/CMU test set.

# Face databases

Example images from the Tilted Test Set portion of the MIT/CMU test set.



**CMU Test Set II**. Most of the faces in this test set are in profile view.

# Face databases

**Cohn-Kanade AU-Coded Facial Expression database.** Examples of emotion-specified expressions from image sequences.

**University of Maryland database**. The images show peak frames taken from an image sequence in which the subjects display a set of facial expressions of their choice.

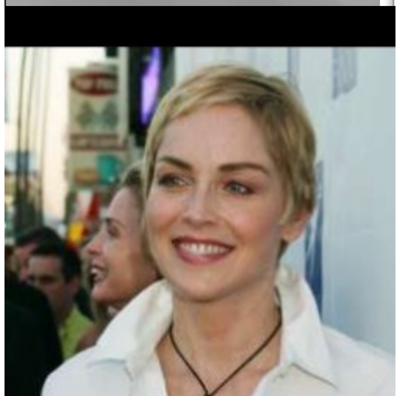
# Face recognition challenges

Vision Lab

**After**

- **FERET**
- **FRVT**
- **FRGC**
- **...**

# Face recognition challenges



## Labeled Faces in the Wild

**Menu**
- LFW Home
  - Mailing
  - Explore
  - Download
  - Train/Test
  - Results
  - Information
  - Errata
  - Reference
  - Contact
  - Support
  - Changes
- UMass Vision

### LFW Home

**New:** Professor Learned-Miller will be running a workshop titled Faces in Real-Life Images at the European Conference on Computer Vision with co-organizers Andras Ferencz and Frederic Jurie.

Welcome to Labeled Faces in the Wild, a database of face photographs designed for studying the problem of unconstrained face recognition. The database contains more than 13,000 images of faces collected from the web. Each face has been labeled with the name of the person pictured. 1680 of the people pictured have two or more distinct photos in the database. The only constraint on these faces is that they were detected by the Viola-Jones face detector. More details can be found in the technical report below.

*last updated: 2007/11/21 1:30 PM EST*
change log

**Mailing list:**
If you wish to receive announcements regarding any changes made to the LFW database, please send email to majordomo@cs.umass.edu with the message body: "subscribe lfw" on a single line.

**Explore the database:**
- Alphabetically by first name:
  [A][Alf][Ang][B][Bin][C][Che][Col][D][Daw][Don][E][Eri][F][G][Goe][H] [I][J] [Jav][Jes][Joh][Jos][K][Kim][L][Lil][M][Mark][Mel][Mik][N][O][P] [Per][Q][R][Ric] [Rog][S][Sha][Ste][T][Tim][U][V][W][X][Y][Z]
- Alphabetically by first name, only people with more than one image:
  [A][B][C][D][E][F][G][H][I][J][K][L][M][N][O][P][Q][R][S][T][U][V][W][X][Y][Z]
- Alphabetically by last name:
  [A][B][C][D][E][F][G][H][I][J][K][L][M][N][O][P][Q][R][S][T][U][V][W][X][Y][Z]
- By number of images per person:

# Face recognition challenges

Vision Lab

Labeled Faces in the Wild

**Menu**
- LFW Home
- UMass Vision

## Results

### Image-Restricted Training Results

| | $\hat{\mu} \pm S_E$ |
|---|---|
| Eigenfaces, original | 0.6002 ± 0.0079 |
| Nowak, original | 0.7245 ± 0.0040 |
| Nowak, funneled | 0.7393 ± 0.0049 |
| MERL | 0.7052 ± 0.0060 |
| MERL+Nowak, funneled | 0.7618 ± 0.0058 |
| Hybrid descriptor-based, funneled | 0.7847 ± 0.0051 |

Table 1: Mean classification accuracy $\hat{\mu}$ and standard error of the mean $S_E$.

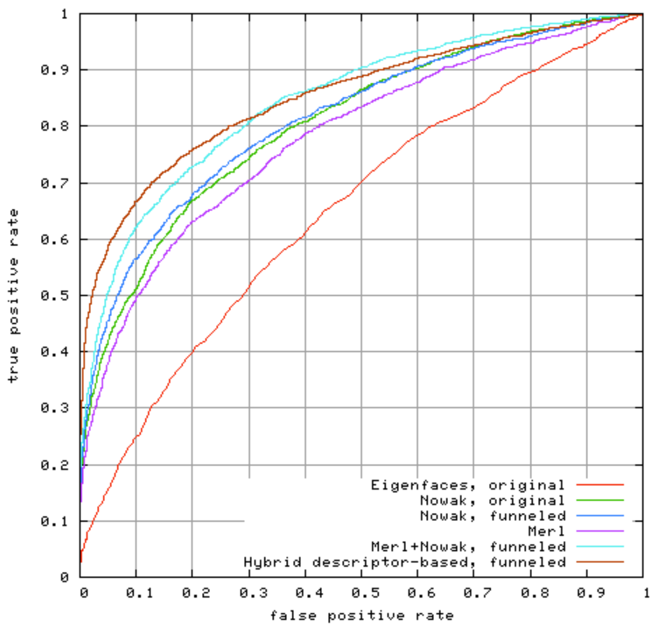



Fig 1: ROC curves averaged over 10 folds of View 2*.

* Each point on the curve represents the average over the 10 folds of (false positive rate, true positive rate) for a fixed threshold.

# Biometrics challenges

2-5 June 2009 University of Sassari, Italy

ICB 2009

Vision Lab

The 3rd IAPR/IEEE International Conference on Biometrics

Home
Call for Papers
Committee
Competitions
Keynote Speakers
Paper Submission
Demo Submission
Awards
Call for Tutorials
Conference Venue
Sponsors
Travel Information
Poster

## ICB2009 Competitions

IAPR

IEEE Biometrics Council

### Competitions Chairs

| | |
|---|---|
| Bernadette Dorizzi | Biosecure Foundation, France |
| Jonathon Phillips | NIST, USA |

### Face recognition from stills and video

This competition is performed under the supervision of Norman Poh from the University of Surrey

### Fingerprint

This competition is performed under the supervision of Raffaele Cappelli from the University of Bologna

### Multimodal Biometric Feature Selection Challenge

This competition is performed under the supervision of Krzysztof Kryszczuk from the Ecole Politechnique Federale de Lausanne

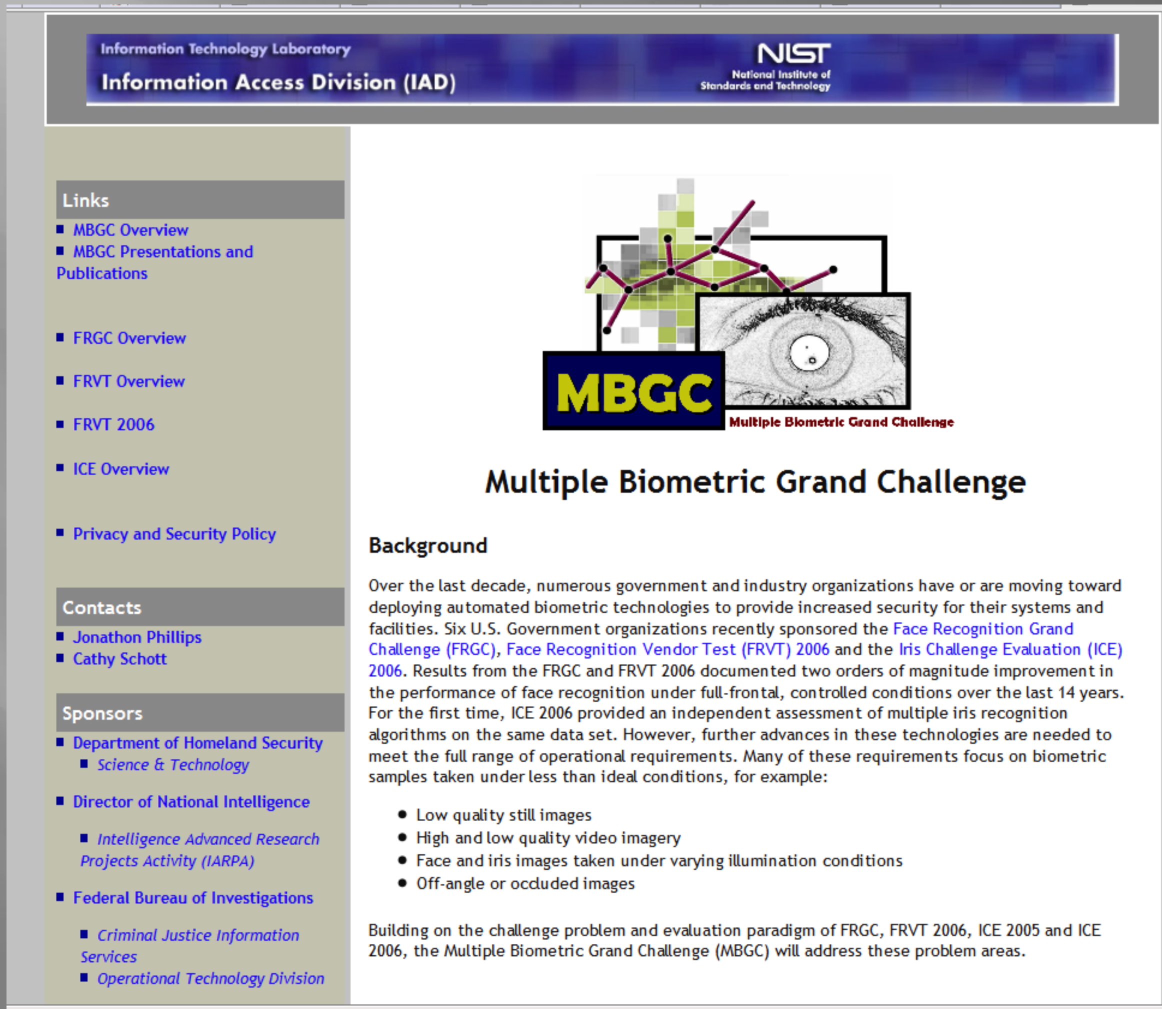
### Multiple Biometrics Grand Challenge

The Multiple Biometrics Grand Challenge is organized and supported by the National Institute of Standards and Technology (NIST). The MBGC is sponsored by multiple U.S. Government Agencies. Dr Jonathon Phillips is the responsible for the NIST MBGC evaluation. Within the framework of ICB program, submissions are encouraged to the MBGC evaluation. The results of the MBGC, together with the other competitions, will be presented at a special conference session.

### Signature verification

This competition is performed under the supervision of Sonia Garcia-Salicetti from the institute TELECOM SudParis

# Biometrics challenges



Information Technology Laboratory
Information Access Division (IAD)
NIST National Institute of Standards and Technology

**Links**

- MBGC Overview
- MBGC Presentations and Publications
- FRGC Overview
- FRVT Overview
- FRVT 2006
- ICE Overview
- Privacy and Security Policy

**Contacts**

- Jonathon Phillips
- Cathy Schott

**Sponsors**

- Department of Homeland Security
  - *Science & Technology*
- Director of National Intelligence
  - *Intelligence Advanced Research Projects Activity (IARPA)*
- Federal Bureau of Investigations
  - *Criminal Justice Information Services*
  - *Operational Technology Division*

MBGC
Multiple Biometric Grand Challenge

## Multiple Biometric Grand Challenge

### Background

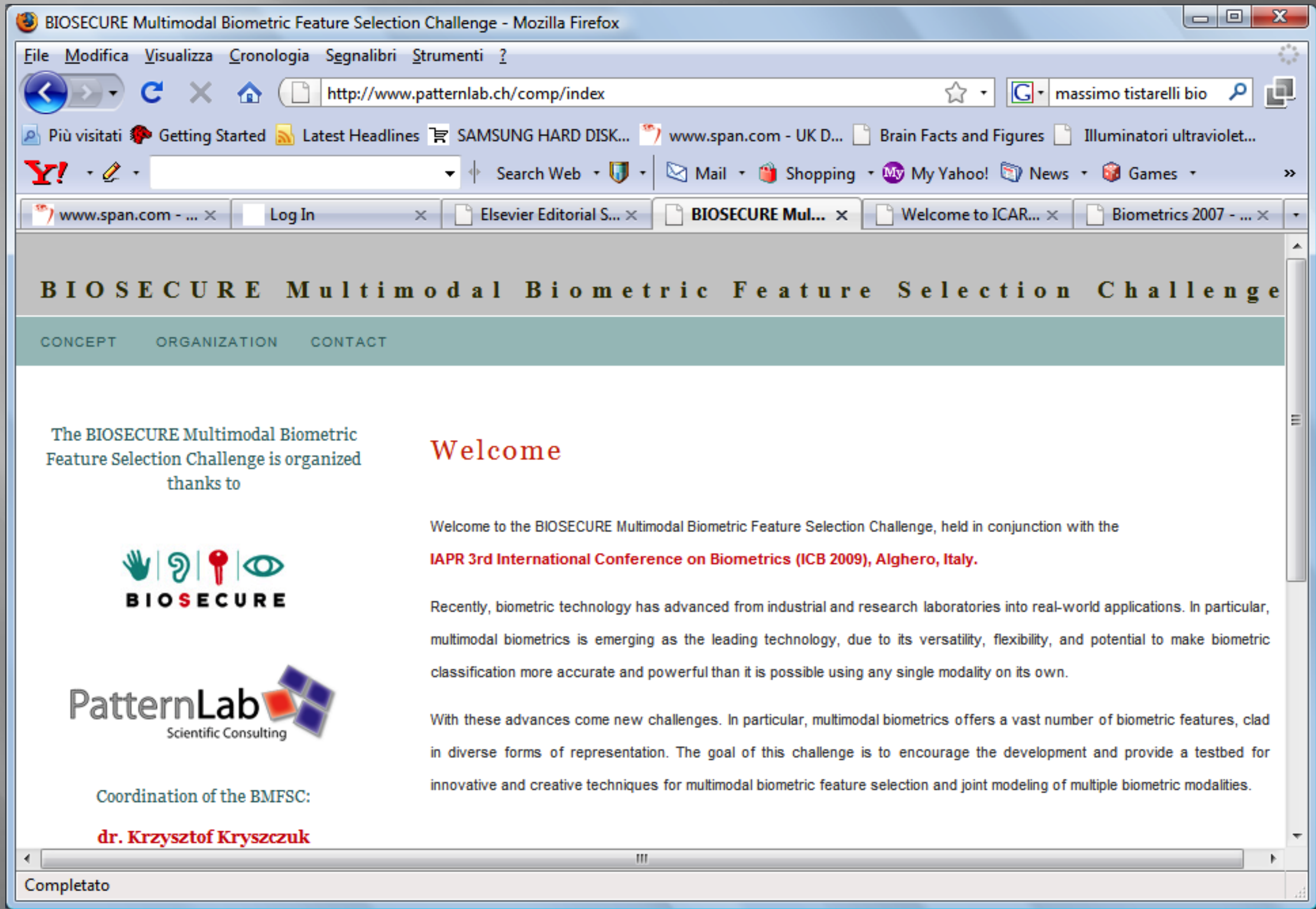
Over the last decade, numerous government and industry organizations have or are moving toward deploying automated biometric technologies to provide increased security for their systems and facilities. Six U.S. Government organizations recently sponsored the Face Recognition Grand Challenge (FRGC), Face Recognition Vendor Test (FRVT) 2006 and the Iris Challenge Evaluation (ICE) 2006. Results from the FRGC and FRVT 2006 documented two orders of magnitude improvement in the performance of face recognition under full-frontal, controlled conditions over the last 14 years. For the first time, ICE 2006 provided an independent assessment of multiple iris recognition algorithms on the same data set. However, further advances in these technologies are needed to meet the full range of operational requirements. Many of these requirements focus on biometric samples taken under less than ideal conditions, for example:

- Low quality still images
- High and low quality video imagery
- Face and iris images taken under varying illumination conditions
- Off-angle or occluded images

Building on the challenge problem and evaluation paradigm of FRGC, FRVT 2006, ICE 2005 and ICE 2006, the Multiple Biometric Grand Challenge (MBGC) will address these problem areas.

# Biometrics challenges

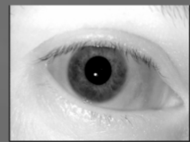
# Biometrics challenges



# Multimodal databases

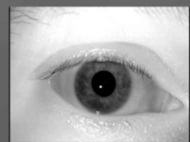
# Multimodal databases

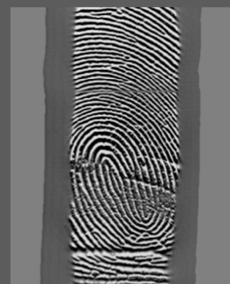










# HUMAN FACE PERCEPTION

Vision Lab

# How do we recognize faces?

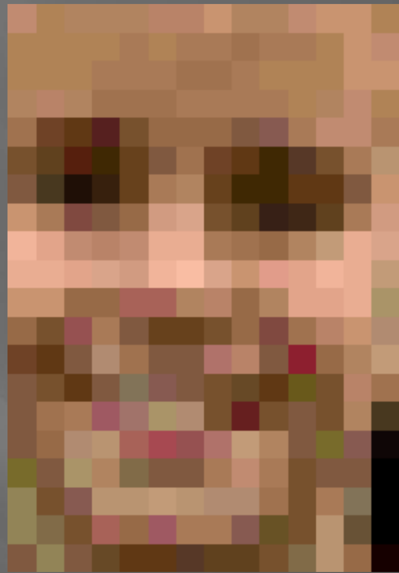
Are both from the same subject?

# How do we recognize faces?

Are both from the same subject?

# Data complexity

**How many pixels are needed to reliably percieve a face?**

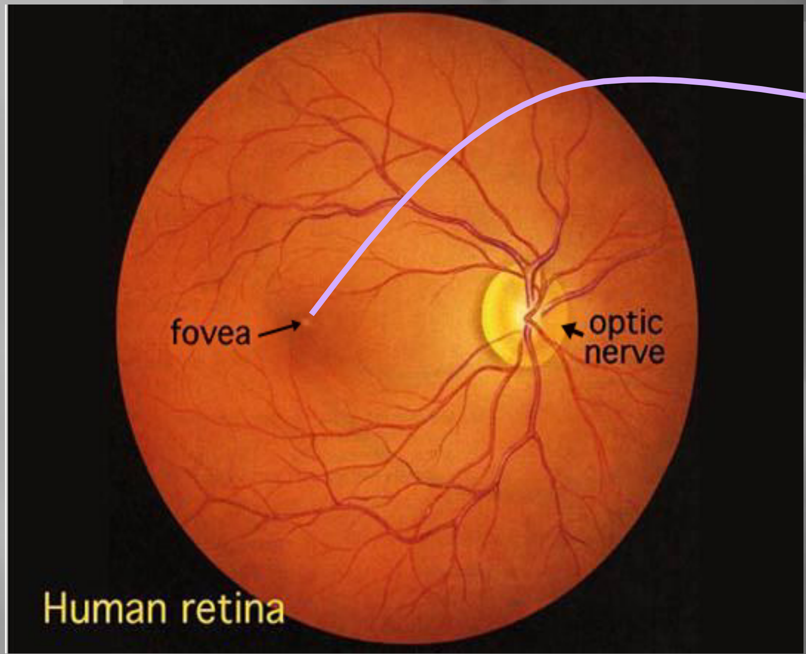


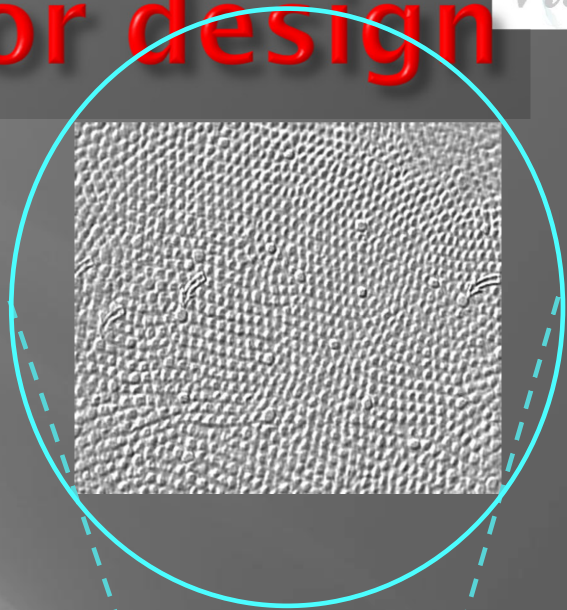
... Not many ... *(20x14)*

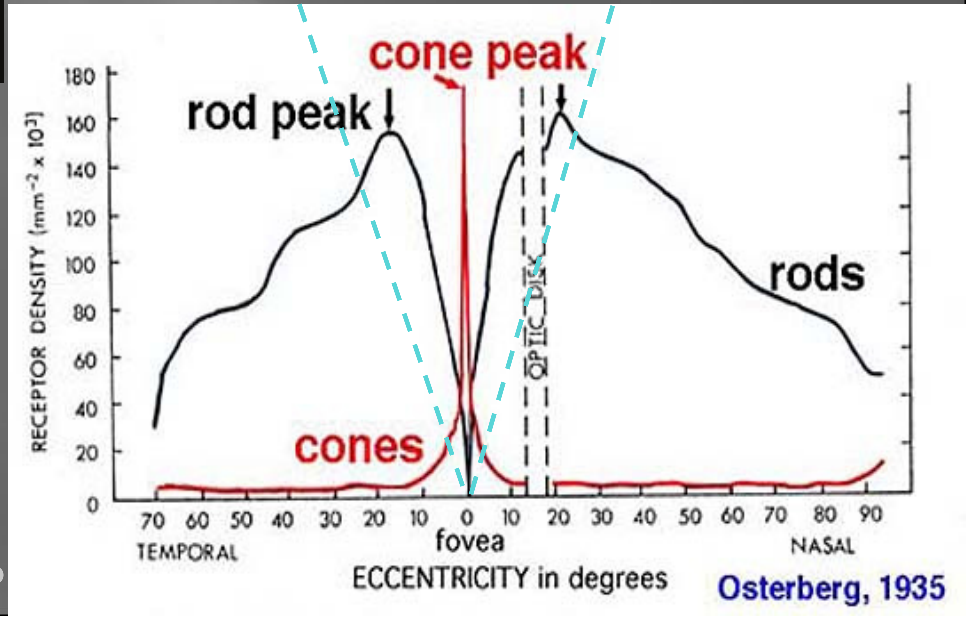
It's more a question of spatial distribution and …

proper frequency tuning

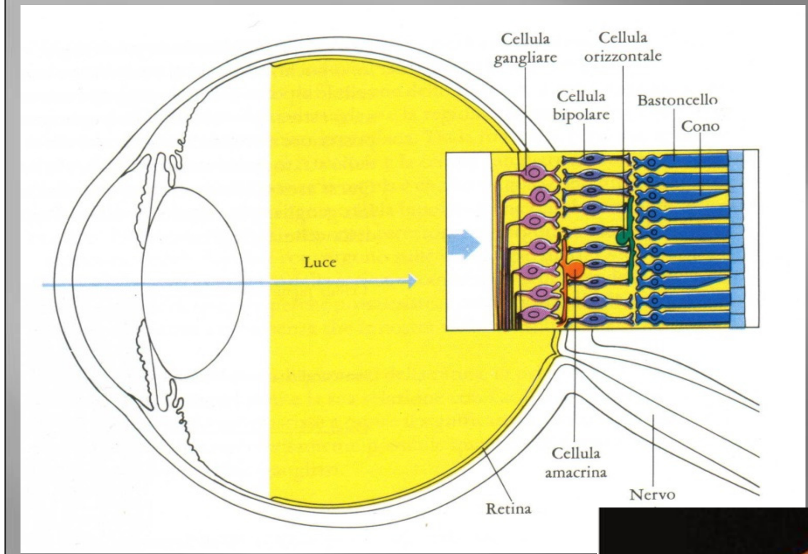
# A unique sensor design





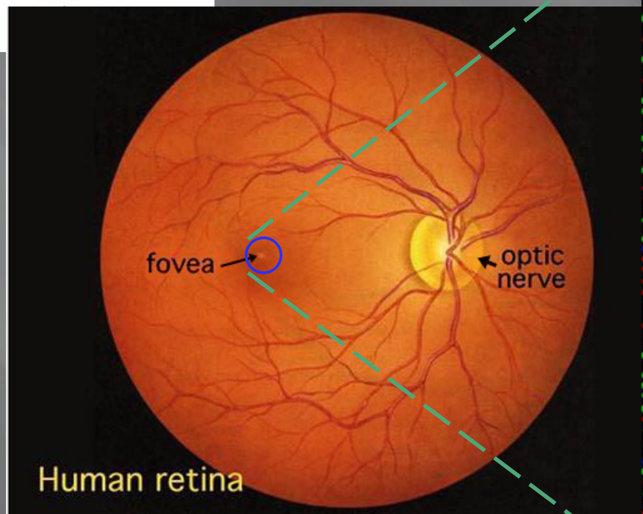
# Data compression

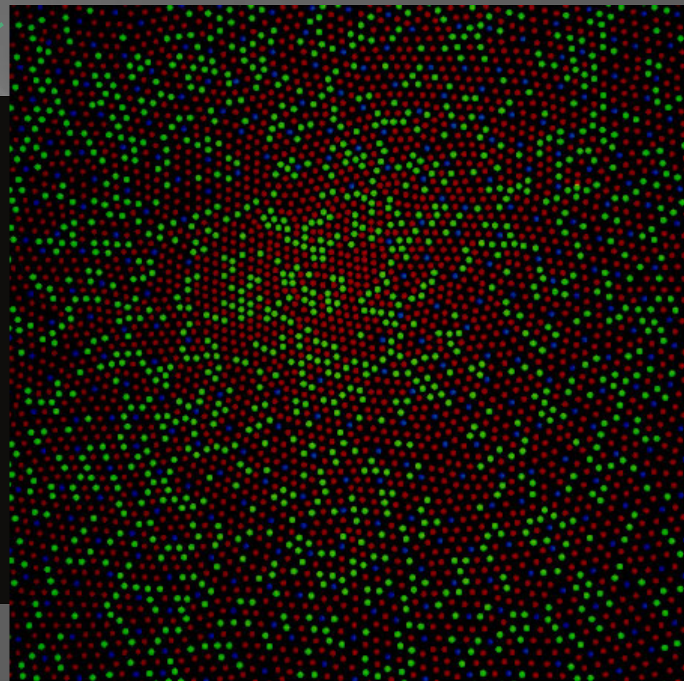



$10^6$ nerve fibers over 200º spherical sensor

6,400,000 cones + 120,000,000 rods

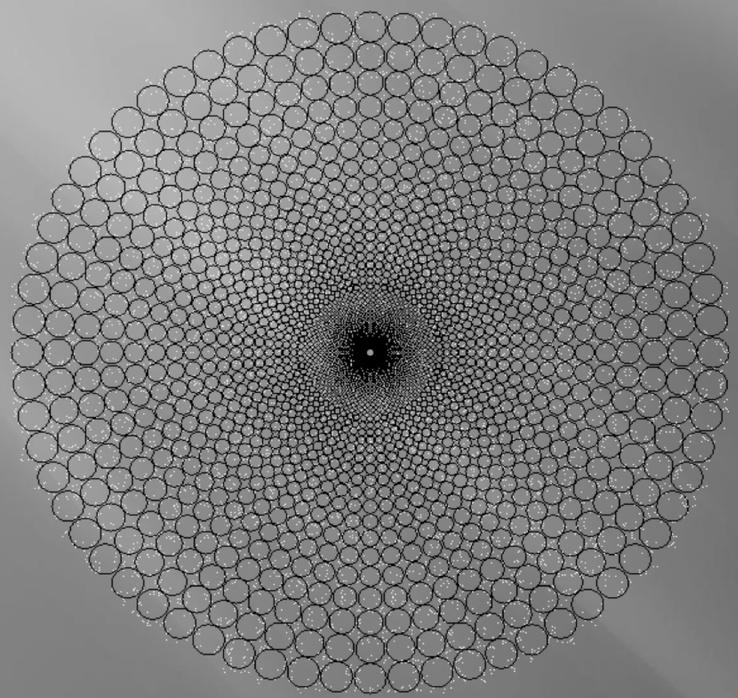
Compression ratio ≅ 1:130

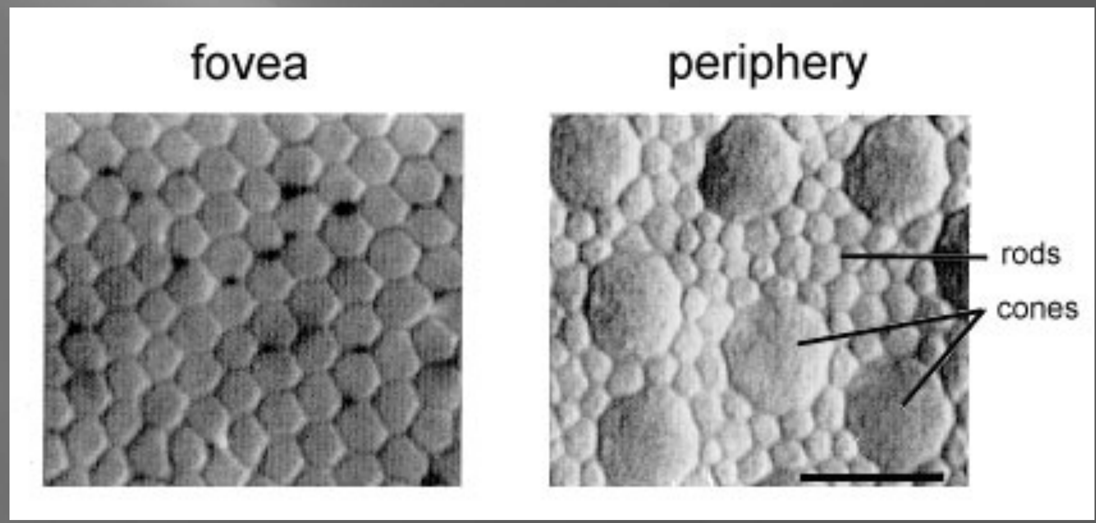


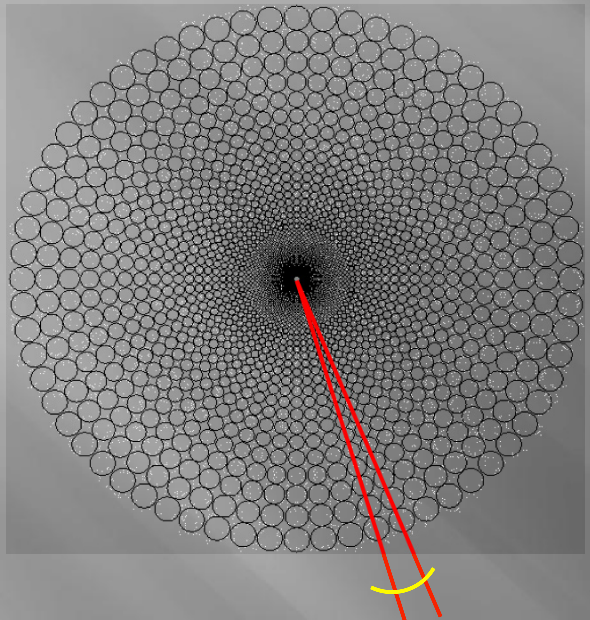
# *Local* analysis: Receptive fields

A good approximation of the cones density over the retina is given by the complex log-polar transform



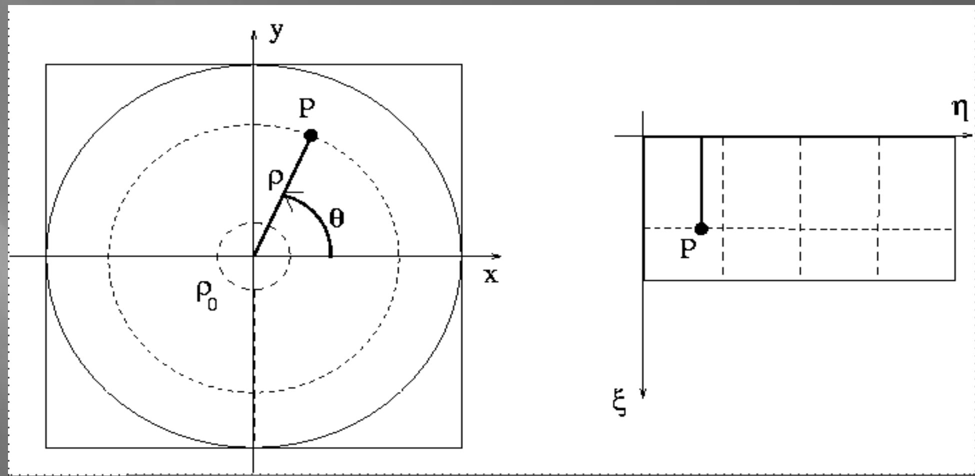


# Space-variant sampling



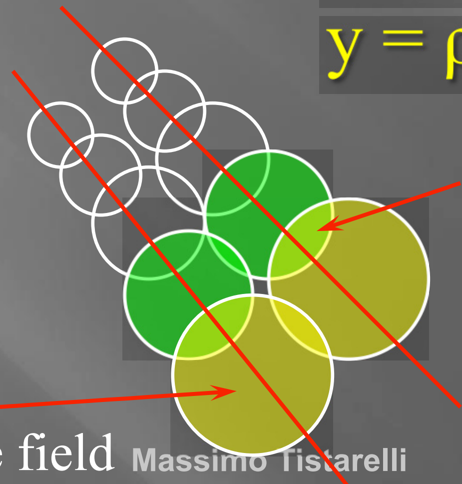
$1/\alpha$

minimal angular resolution




$x = \rho \sin \theta \qquad \xi = \log_a (\rho/\rho_0)$

$y = \rho \cos \theta \qquad \eta = \alpha\, \theta$

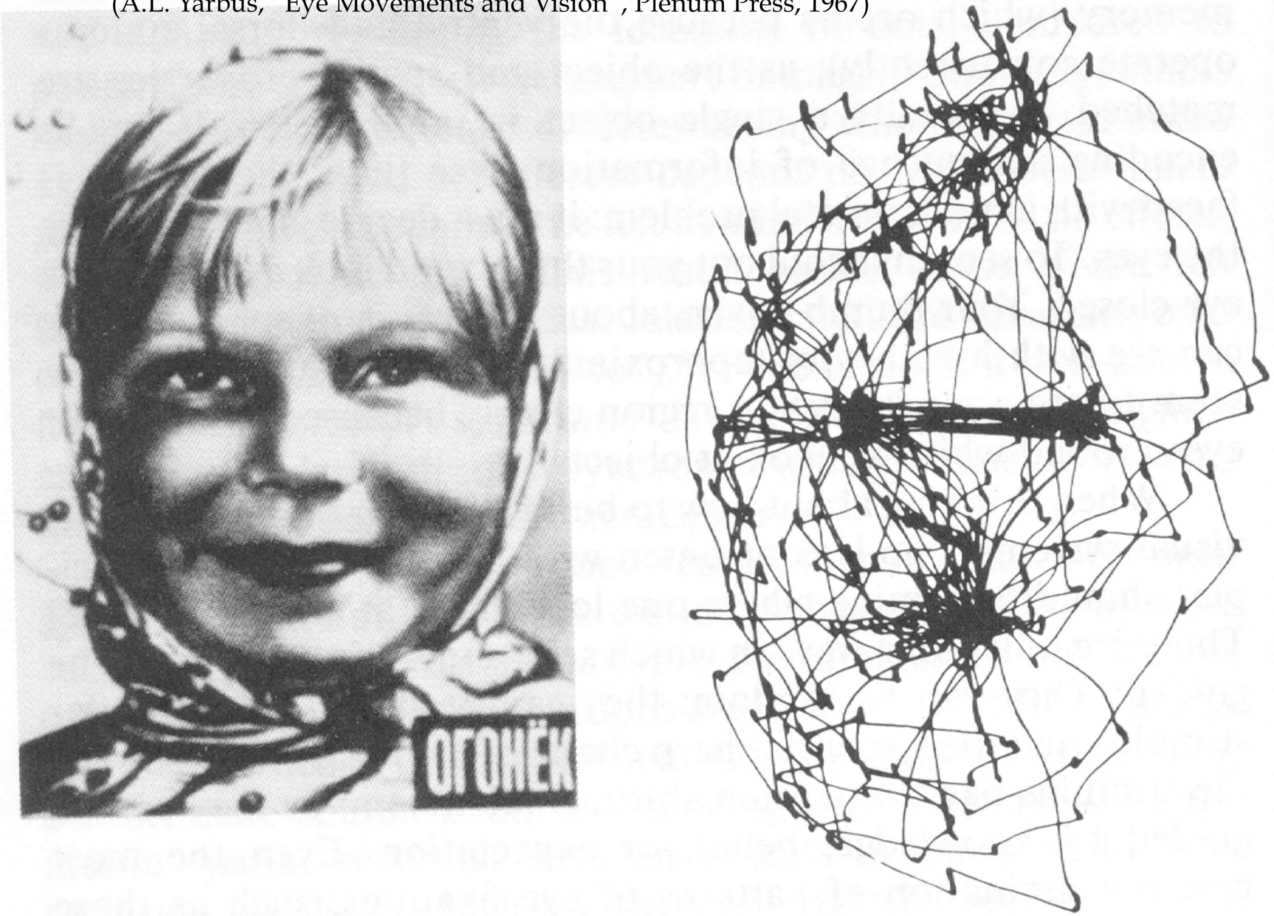


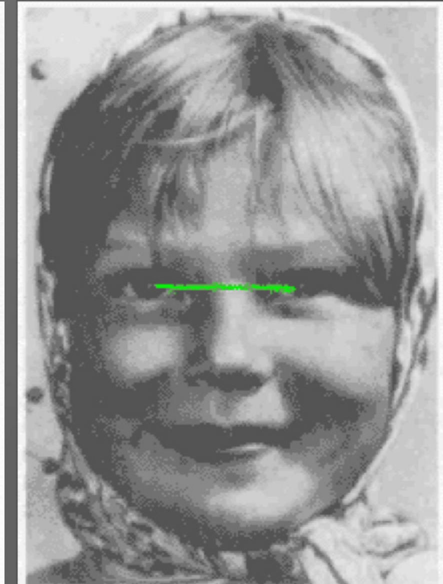
overlap factor (controlled by $\boldsymbol{a}$)

receptive field

# *Context* analysis: Visual attention

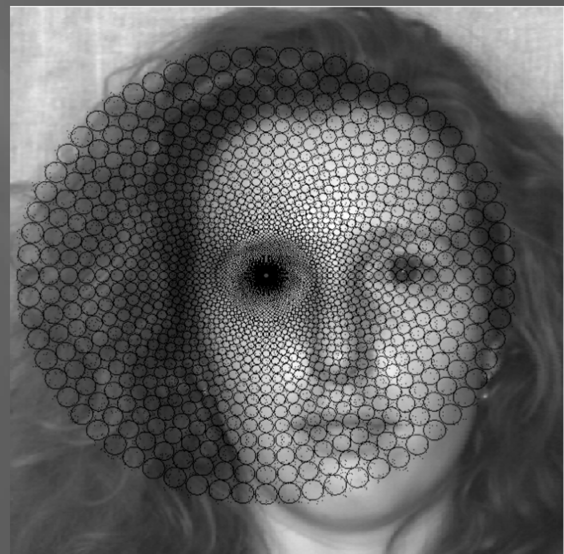
Eye movements while watching a girl's face
(A.L. Yarbus, "Eye Movements and Vision", Plenum Press, 1967)



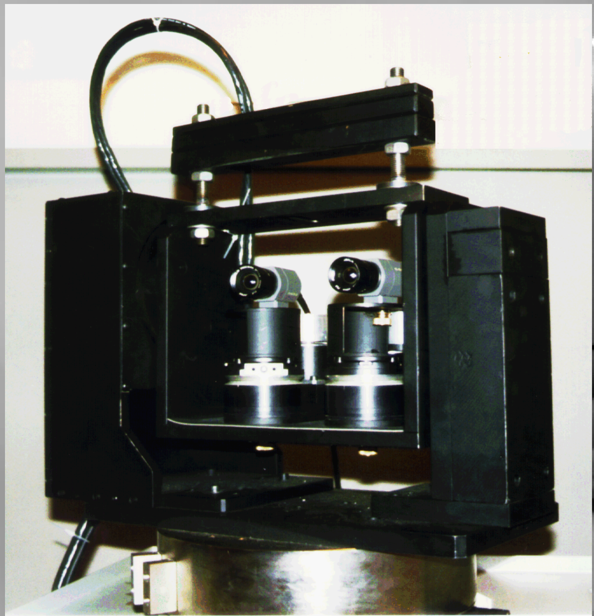



# Visual attention

- **Visual attention requires**
  - a special acquisition sensor/ camera with motor control
  - or a high resolution image for either Hw or Sw sub-sampling



# Space-variant active systems





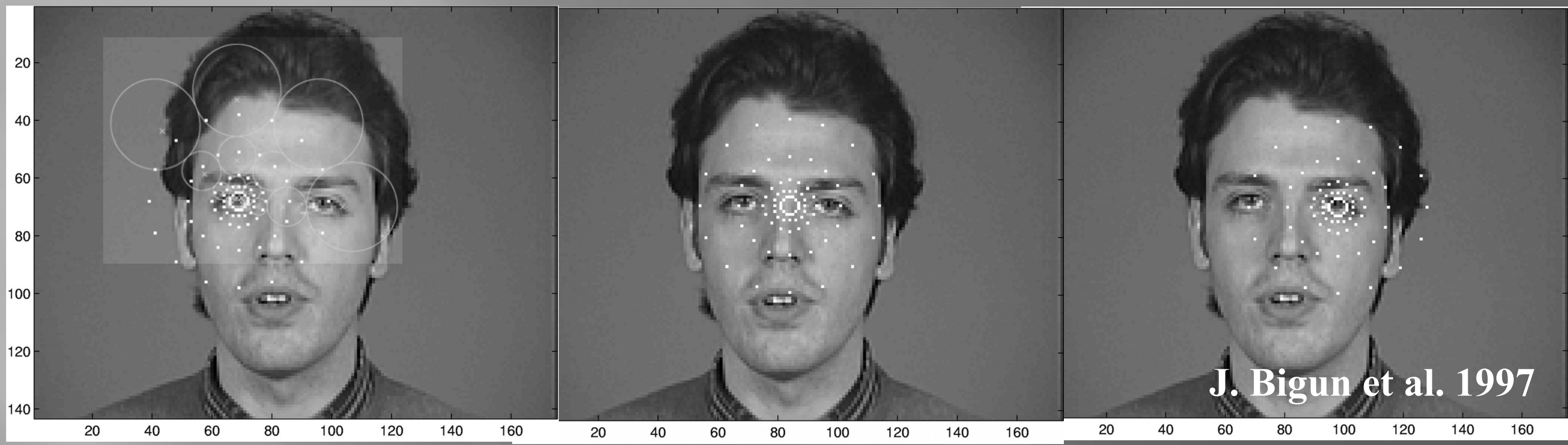
Tistarelli, M., and  Sandini G. (1992) "Dynamic Aspects in Active Vision", *Computer Vision Graphics and Image Processing: Image Understanding*, Special Issue on Purposive and Qualitative Active Vision, Vol. 56, No. 1, pp. 108-129, July 1992

Tistarelli, M. and Grosso, E. (1997) "Active face recognition with an hybrid approach" *Journal of Pattern Recognition Letters*: Special issue on audio-visual person authentication, Vol. 18, pp 933-946, 1997

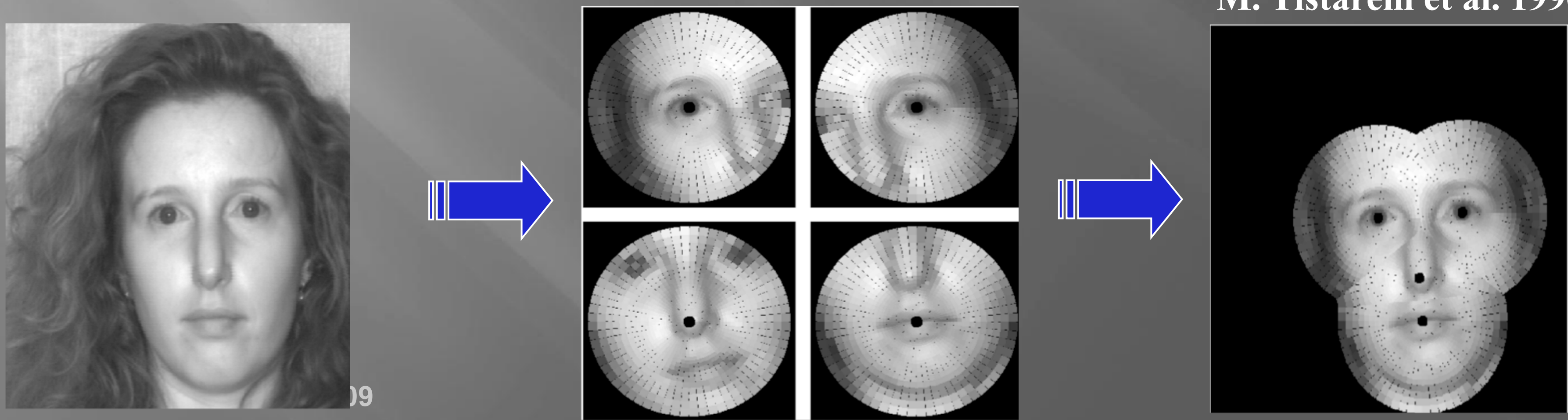
Tistarelli, M. and Grosso, E. (2000) "Active vision-based face authentication" *Image and Vision Computing*: Special issue on Facial Image Analysis, M. Tistarelli ed., Vol. 18, no. 4, pp 299-314, 2000

Vision Lab

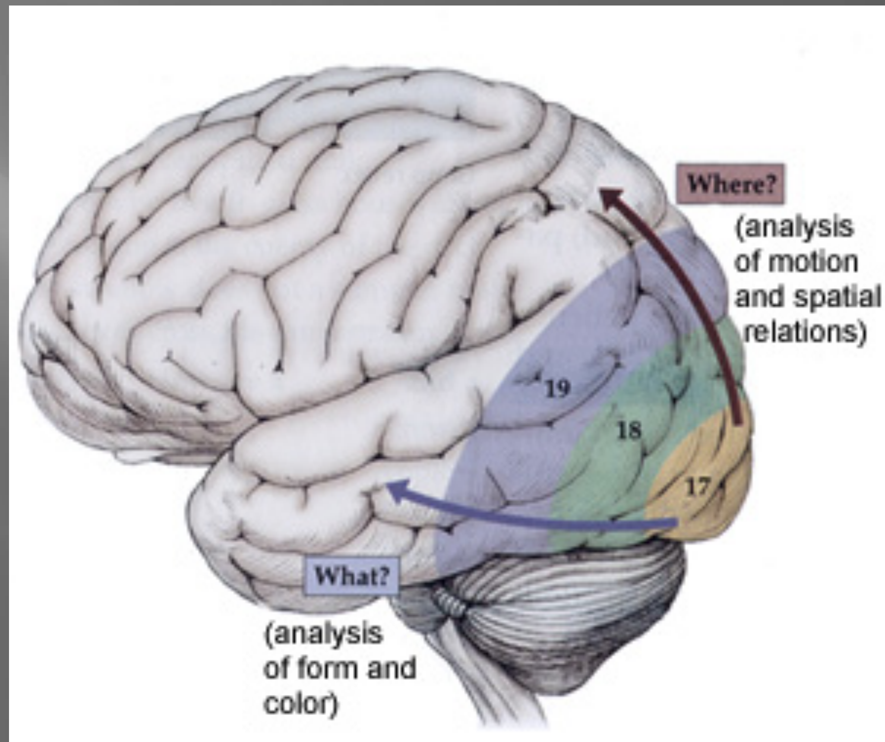
# Space-variant representations




M. Tistarelli et al. 1996



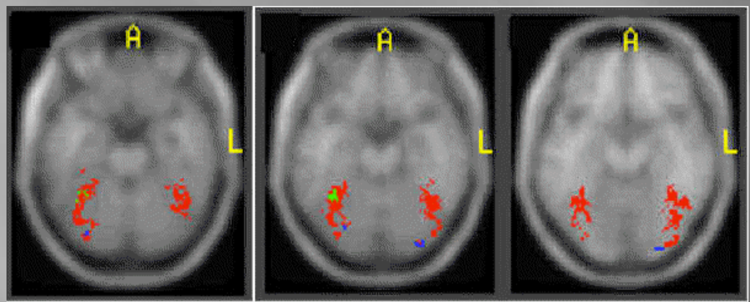
# Where is the *Mom's Neuron?*

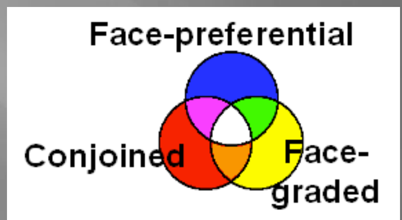



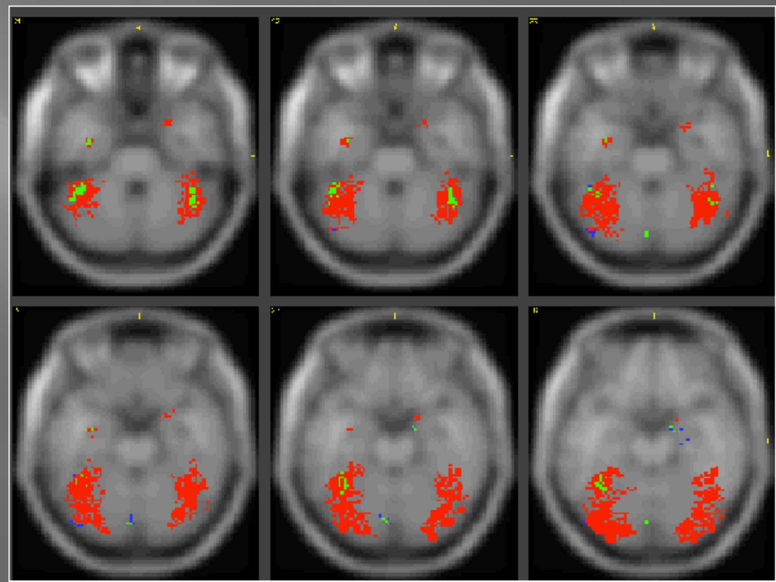
# Functional Magnetic Resonance Imaging


21-06-2009

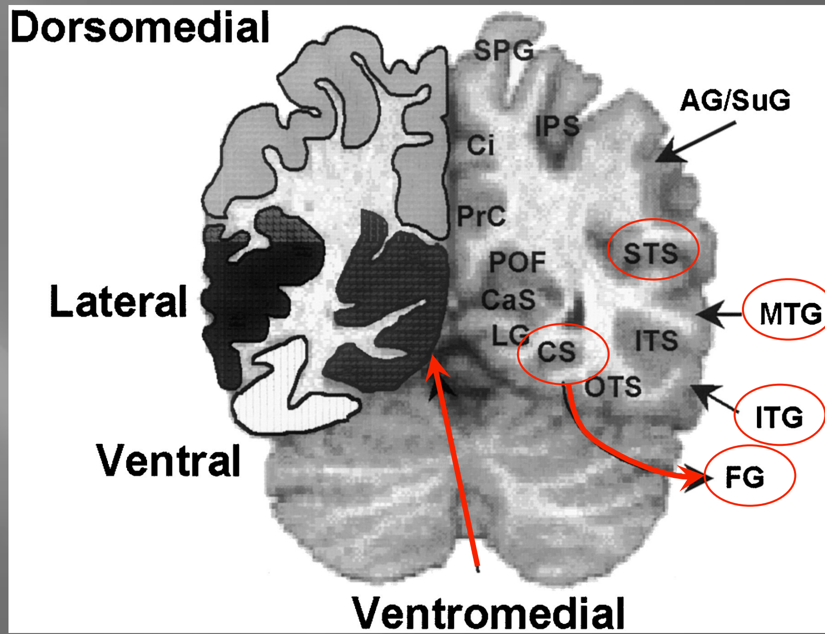
# Brain anatomy and fMRI



# Brain anatomy




**Coronal anatomical image of a brain:**

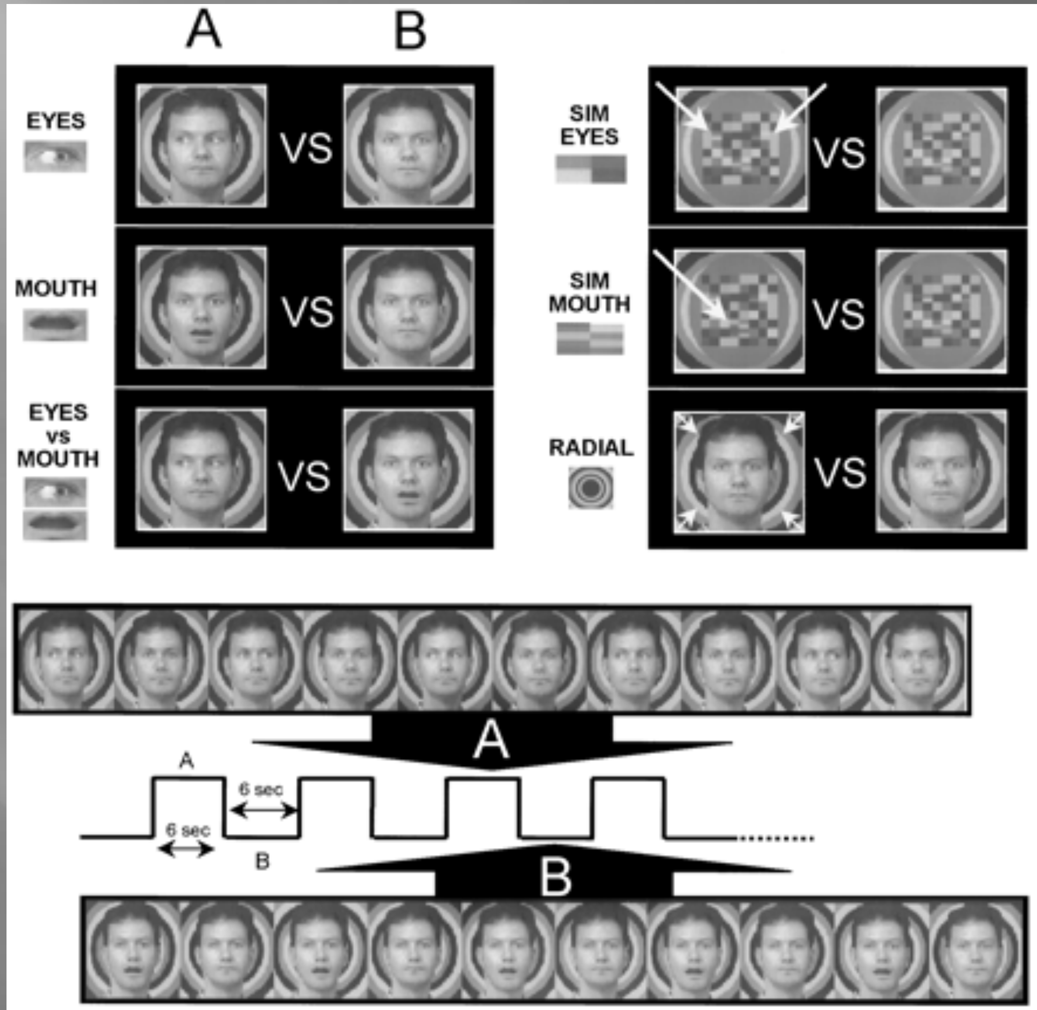
*STS*, Superior temporal sulcus;
*MTG*, middle temporal gyrus;
*ITS*, inferior temporal sulcus;
*ITG*, inferior temporal gyrus;
*FG*, fusiform gyrus;

From **fMRI**, face-selective neurons have been found in the inferior temporal areas, the superior temporal sensory area, the amygdala, the ventral striatum (receiving input from the amygdala) and the inferior convexity.

Specific regions have been reported also in the fusiform gyrus responding significantly to viewing faces.

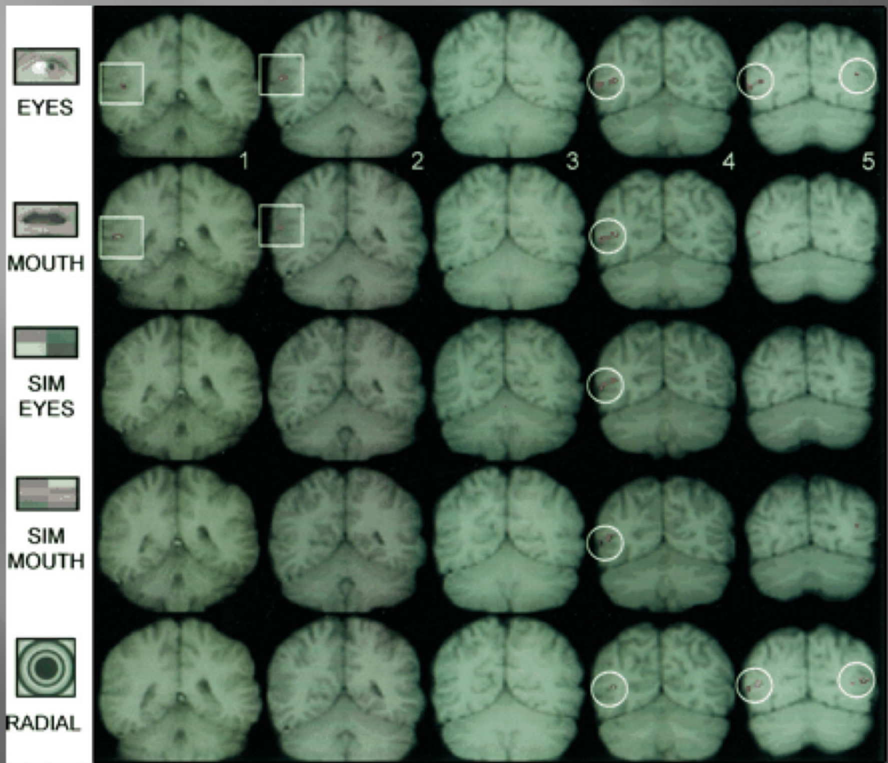
A. R. Damasio, J. Damasio and G. W. Van Hoesen. "Prosopagnosia: Anatomic basis and behavioral mechanisms". Neurology, vol. 32, 331-341, 1982.

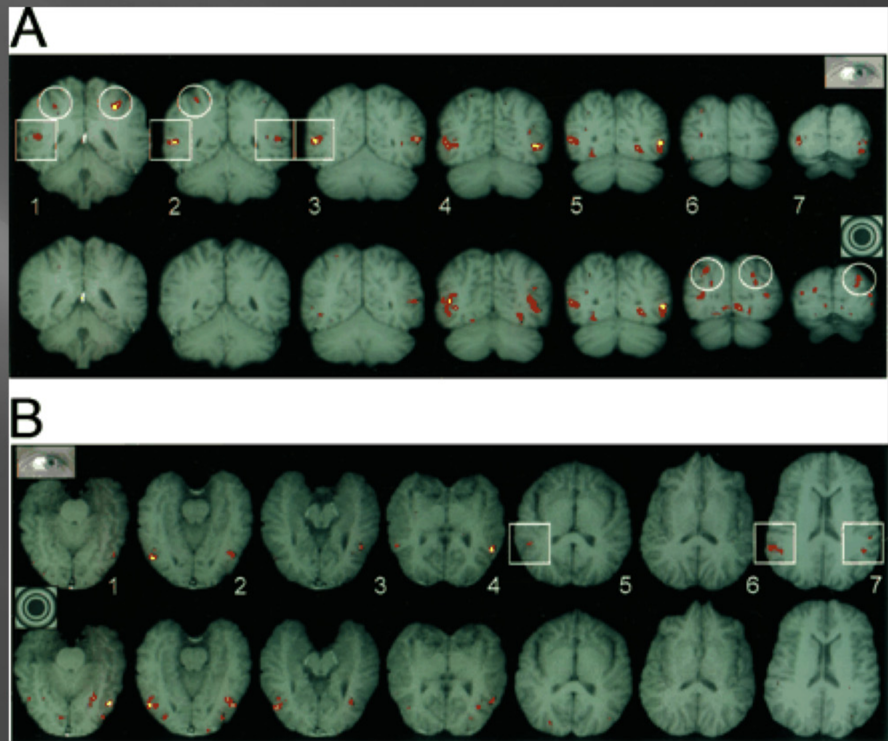
# Brain activation



# Brain activation - fMRI maps

Vision Lab




Aylward et al. *Brain Activation during Face Perception: Evidence of a Developmental Change. J. Cogn. Neurosci..2005; 17: 308-319.*

# Brain activation - fMRI maps

Vision Lab

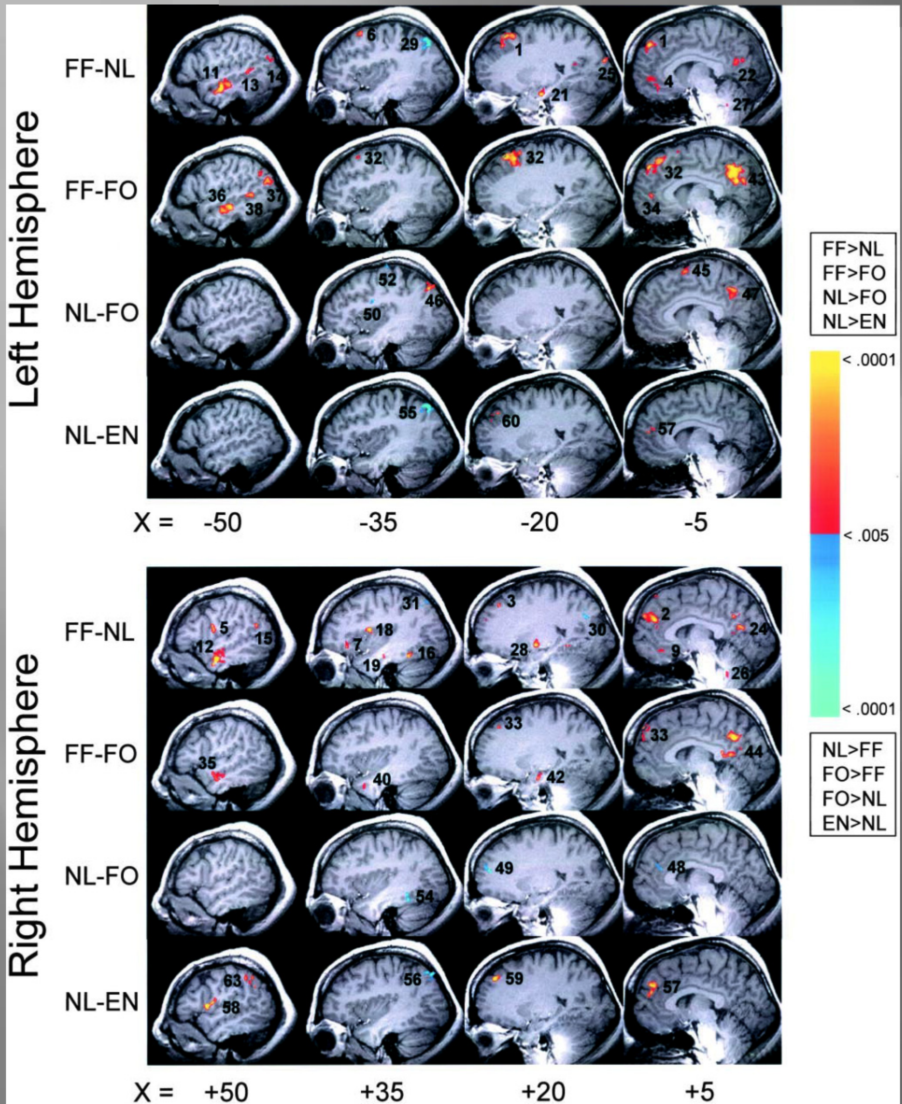
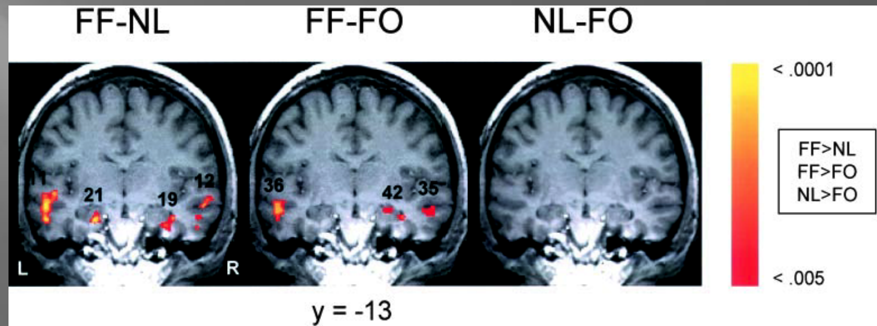



*Figure 2.* Areas of significantly increased (*red–yellow* scale) and decreased (*blue–cyan* scale) MR signal intensity from *t* tests ($p < 0.005$) comparing the three conditions: FF minus NL, FF minus FO, and NL minus FO. Numbers below each image represent millimeters from the interhemispheric fissure (−, left; +, right). Numbers adjacent to activated foci correspond to location numbers (first column) of Tables 1, 2, and 3.

Recognition of 50 Familiar Faces (**FF**) vs 50 Newly Learned Faces (**NL**) and compared to rejection of 50 Foil (**FO** -False Objective) faces. Encoding (**EN**) session for learning new faces.




C. L. Leveroni et al. **"Neural Systems Underlying the Recognition of Familiar and Newly Learned Faces",** The Journal of Neuroscience, January 15, 2000, *20*(2):878–886

# Brain activation - fMRI maps

Vision Lab

C. L. Leveroni et al. **"Neural Systems Underlying the Recognition of Familiar and Newly Learned Faces"**, The Journal of Neuroscience, January 15, 2000, *20*(2):878–886

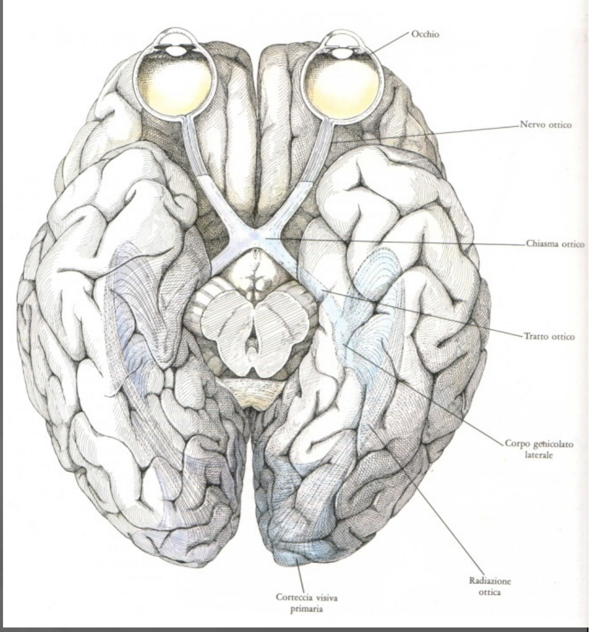
**Table 1. Famous faces (FF) vs newly learned (NL) faces**

| Loc. # | Brain region | BA | vol. (ml) | *x* | *y* | *z* |
|---|---|---|---|---|---|---|
| | FF > NL | | | | | |
| | Frontal Lobe | | | | | |
| 1 | L Superior Frontal | 8 | 2.6 | −15 | 33 | 44 |
| 2 | R Medial Frontal | 9 | 2.4 | 10 | 47 | 25 |
| 3 | R Superior Frontal | 8 | 0.5 | 12 | 40 | 45 |
| 4 | L Medial Frontal | 10 | 0.4 | −6 | 49 | −4 |
| 5 | R Precentral | 6 | 0.4 | 49 | −1 | 13 |
| 6 | L Superior Frontal | 8 | 0.4 | −36 | 15 | 50 |
| 7 | R Inferior Frontal | 47 | 0.3 | 32 | 32 | −7 |
| 8 | R Anterior Cingulate | 32 | 0.3 | 11 | 21 | −7 |
| 9 | R Medial Frontal | 11 | 0.3 | 9 | 35 | −13 |
| 10 | L Medial Frontal | 11 | 0.3 | −6 | 39 | −14 |
| | Temporal Lobe | | | | | |
| 11 | L Middle Temporal | 21 | 2.7 | −51 | −11 | −13 |
| 12 | R Middle Temporal | 21 | 1.9 | 52 | −6 | −18 |
| 13 | L Middle Temporal | 21 | 0.6 | −49 | −42 | 7 |
| 14 | L Middle Temporal | 39 | 0.5 | −46 | −68 | 22 |
| 15 | R Superior Temporal | 22 | 0.5 | 54 | −52 | 15 |
| 16 | R Fusiform | 20/37 | 0.4 | 32 | −46 | −16 |
| 17 | R Middle Temporal | 37 | 0.3 | 43 | −64 | 9 |
| 18 | R Insula | – | 0.3 | 37 | 3 | 11 |
| 19 | R Parahippocampal | 35 | 0.2 | 30 | −14 | −23 |
| 20 | R Parahippocampal | 36 | 0.2 | 24 | −43 | −7 |
| 21 | L Hippocampus | 28 | 0.2 | −19 | −12 | −20 |
| | Parietal/Occipital Lobe | | | | | |
| 22 | L Posterior Cingulate | 23/30 | 1.7 | −4 | −57 | 15 |
| 23 | R Inferior Parietal | 40 | 0.5 | 44 | −30 | 22 |
| 24 | R Posterior Cingulate | 31 | 0.3 | 2 | −57 | 29 |
| 25 | L Extrastriate | 18 | 0.3 | −20 | −89 | 20 |
| | Subcortical | | | | | |
| 26 | R Pons | – | 0.4 | 11 | −43 | −34 |
| 27 | L Pons | – | 0.2 | −10 | −43 | −33 |
| 28 | R Putamen | – | 0.3 | 22 | −7 | −6 |
| | NL > FF | | | | | |
| | Parietal Lobe | | | | | |
| 29 | L Inferior Parietal | 40 | 1.0 | −37 | −64 | 40 |
| 30 | R Superior Parietal | 7 | 0.5 | 23 | −66 | 30 |
| 31 | R Inferior Parietal | 40 | 0.3 | 35 | −67 | 42 |

Region is defined as center of mass. The first column refers to location numbers demarcated in Figures 2 and 3 (italicized numbers indicate locations not shown in figures). Coordinates represent distance in millimeters from anterior commissure: *x* right (+)/left (−); *y* anterior (+)/posterior(−); *z* superior (+)/inferior(−).

# Brain activation

- Parts of the inferior and medial temporal cortex may work together to process faces the amygdala being responsible for assigning significance to faces, and thus **affecting both attention and mnemonic aspects of face processing**.




**C. A. Nelson**. "The development and neural bases of face recognition". Infant and Child Development, vol. 10, 3-18, 2001.

**J. P. Aggleton, M. J. Burton and R. E. Passingham**. "Cortical and subcortical afferents to the amygdala of the rhesus monkey (Macaca mulatta)". Brain Research, vol. 190, 347–368, 1980.

# Face perception and attention

Vision Lab

- A recent fMRI study on individuals with autism and Asperger syndrome showed a failure to activate the fusiform face area during face processing. Damage to fusiform gyrus and to amygdala results in impaired face recognition.

R. T. Schultz, I. Gauthier, A. Klin, R. K. Fulbright, A. W. Anderson, F. R. Volkmar, P. Skudlarski, C. Lacadie, D. J. Cohen and J. C. Gore. "Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome". Archives of General Psychiatry, vol. 57, 331–340, 2000

- Behavioral studies (positive and negative) suggest that the most salient parts for face recognition are, in order of importance: eyes, mouth, and nose.

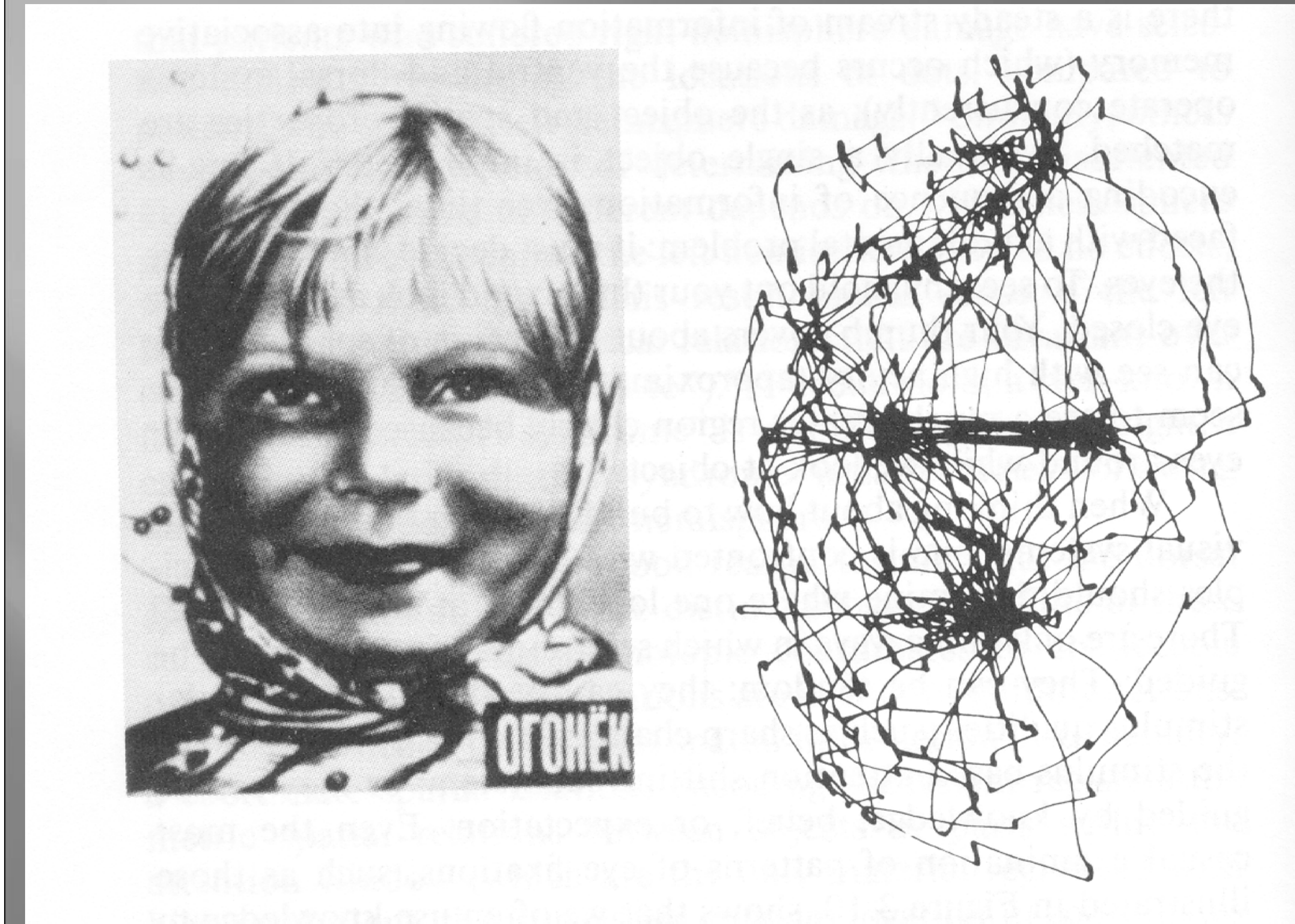
**J. Shepherd**. "Social factors in face recognition". In G. Davies, H. Ellis & J. Shepherd (Eds.), Perceiving and remembering face, London: Academic Press, 55–79, 1981.

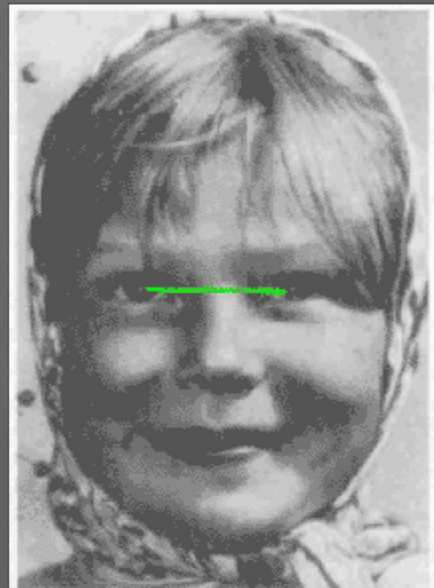
# Visual attention

- Klin reported that adults with autism show abnormal patterns of attention when viewing naturalistic social scenes. These patterns include reduced attention to the eyes and increased attention to mouths, bodies, and objects.

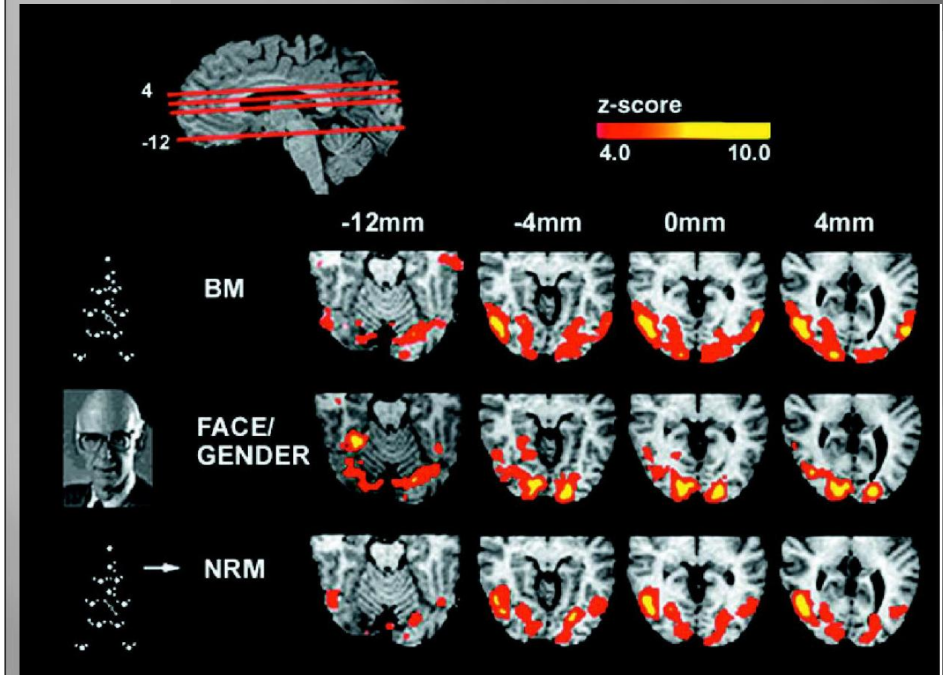

**A. Klin**. "Eye-tracking of social stimuli in adults with autism". *NICHD Collaborative Program of Excellence in Autism*. Yale University, New Haven, CT, May 2001

# Visual attention



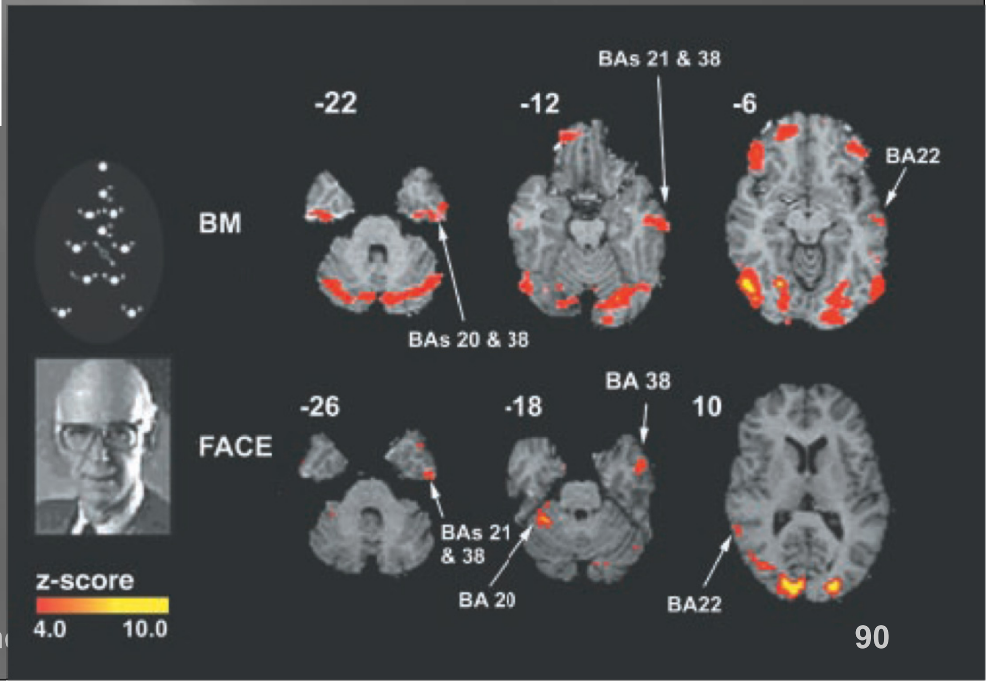
# Face and motion perception



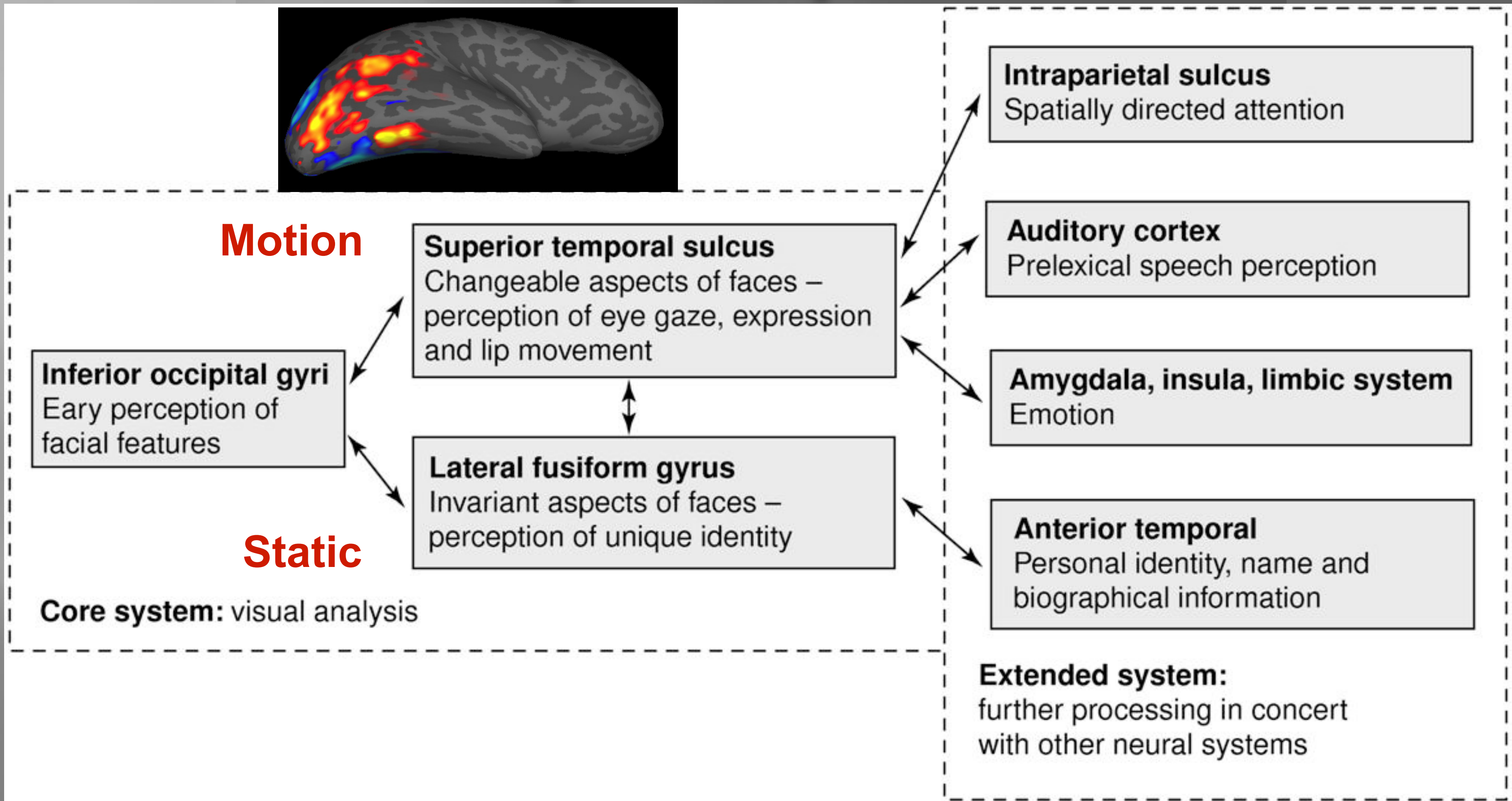

BM = Biological Motion

NR = Non Rigid Motion

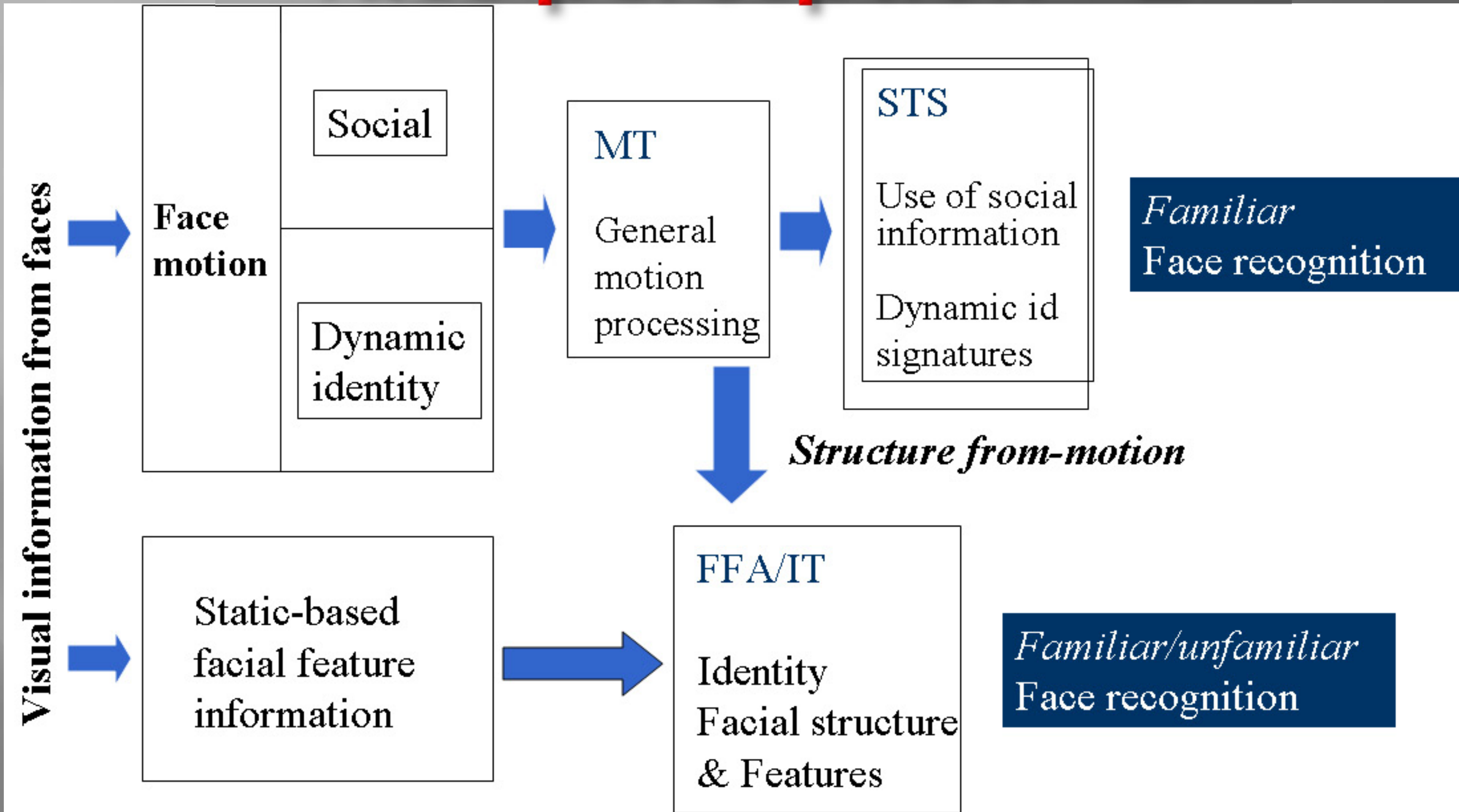
Vaina, L.M., Solomon, J., Chowdhury, S., Sinha, P., Belliveau, J.W., “Functional Neuroanatomy of Biological Motion Perception in Humans”. *Proc. of the National Academy of Sciences of the United States of America*, Vol. 98, No. 20 (Sep. 25, 2001) , pp. 11656-11661



# Neural architecture of face perception - 1



# Neural architecture of face perception - 2




**A. O'Toole**, "Biological Face Recognition", Summer School on Biometrics 2006.

# Face representation

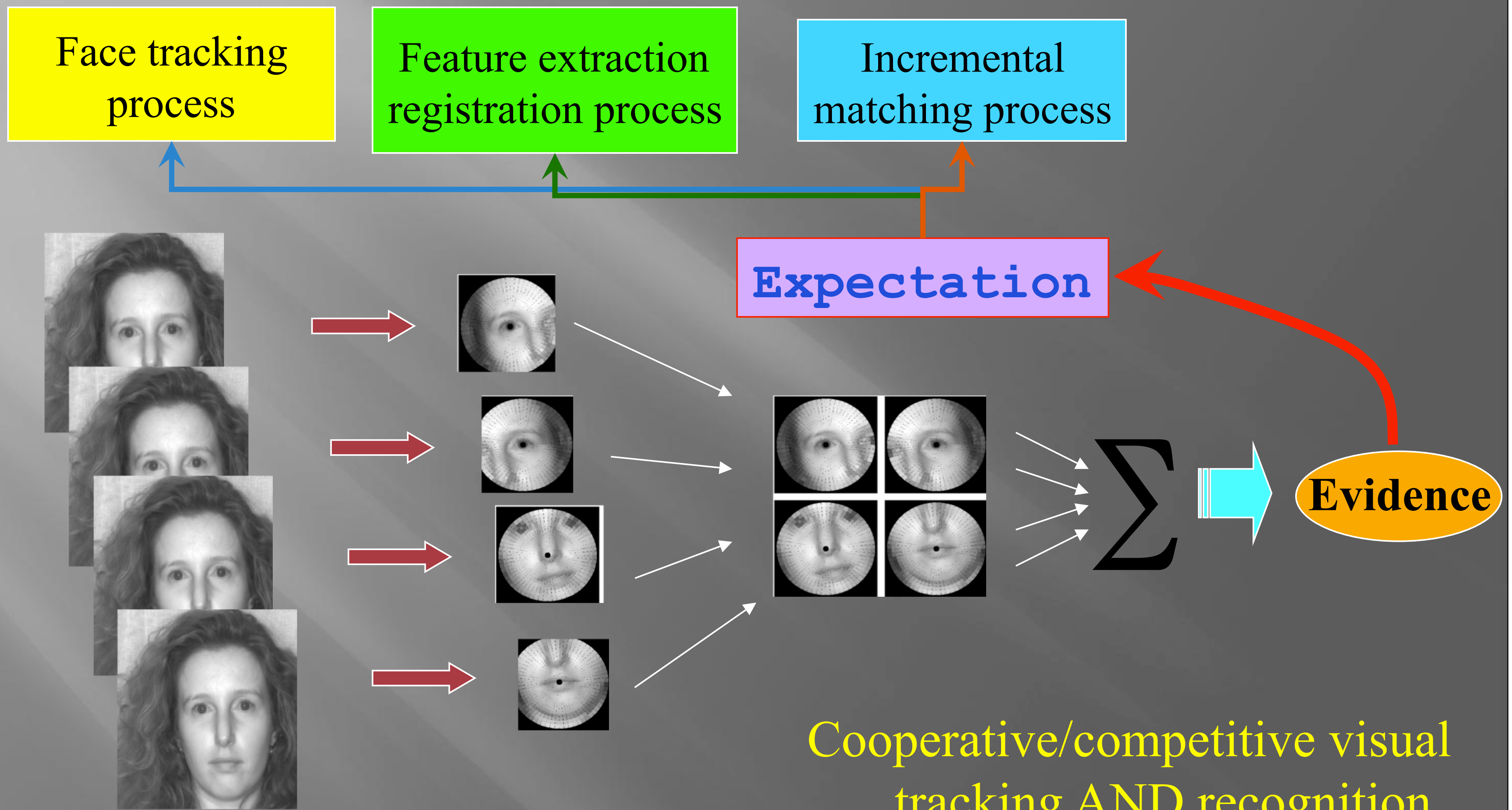
- It is not clear how faces are represented in the HVS, but the representation is formed by a dynamically updated collection of visual fixations.
  - Both foveal and peripheral vision are involved, the former responsible for a more accurate representation.

J.M. Henderson et al. “Eye movements are functional during face learning”, *Memory & Cognition 2005, 33 (1), 98-106.*

D.R. Melmoth et al. “The Effect of Contrast and Size Scaling on Face Perception in Foveal and Extrafoveal Vision”, *Investigative Ophthalmology and Visual Science.* 2000;41:2811-2819.

- This representation includes both iconic data as well as information about the spatial relationship among face elements.

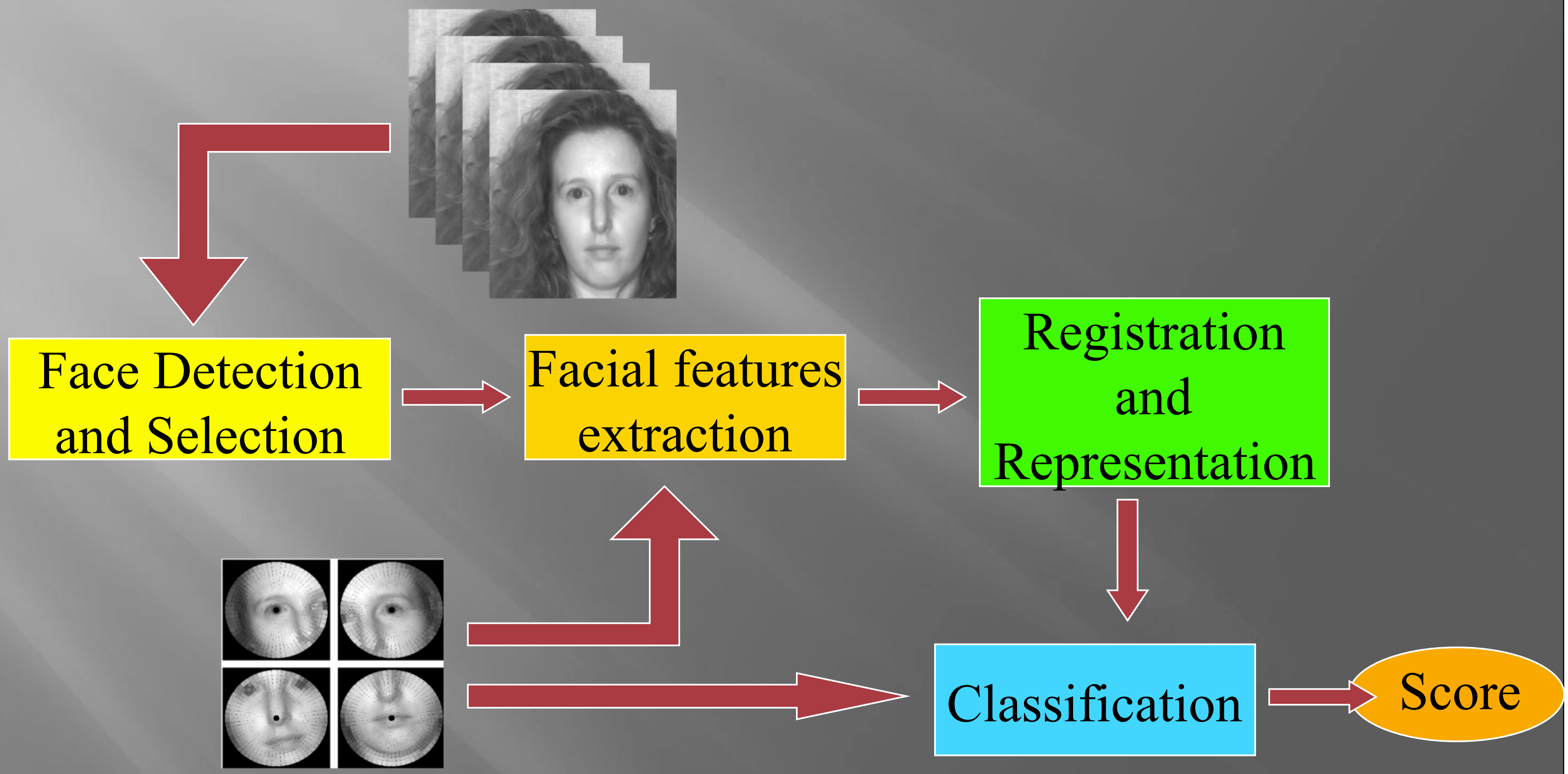
# Biological face perception




Cooperative/competitive visual tracking AND recognition

# FACE RECOGNITION BY MACHINES

# Face matching

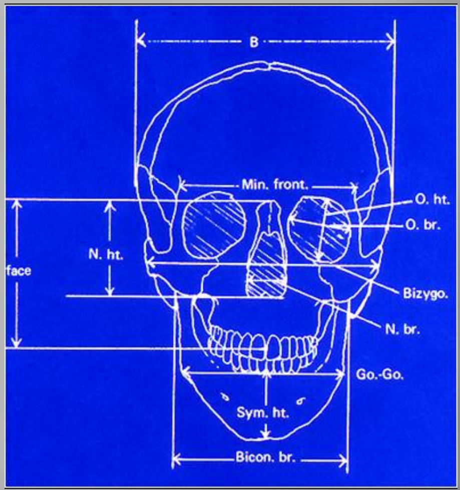


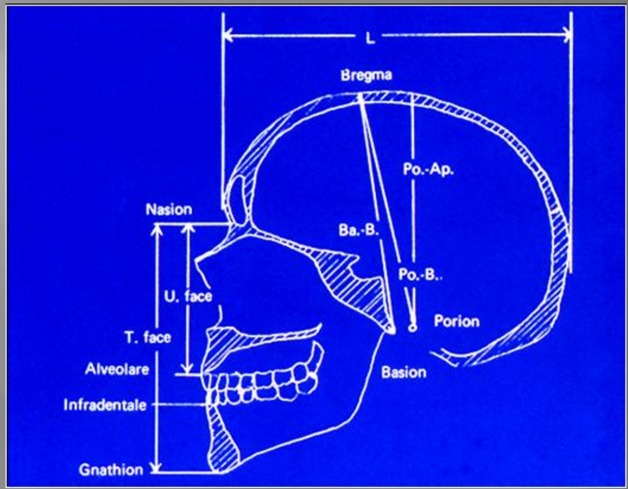
# Automatic face recognition

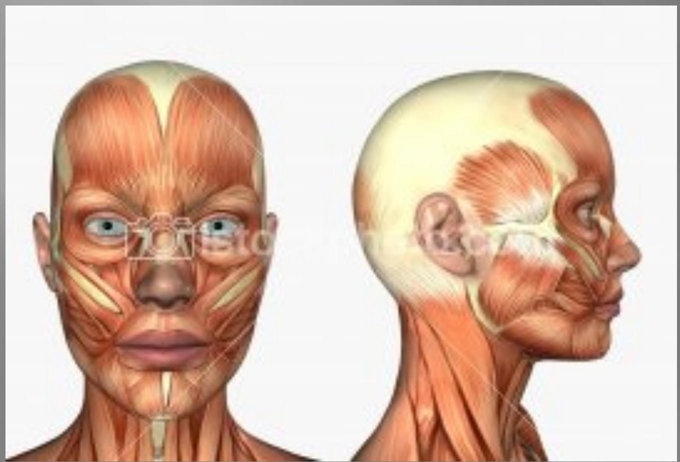
**Face recognition involves several sequential processes:**

1. **Face detection**
2. **Normalization**
3. **Facial features extraction**
4. **Registration/ Representation**
5. **Classification**

# Human face structure



**Craniometric measurements (1)**

- Cranial circumference
- Max. cranial breadth
- Min. frontal breadth
- Bigonial breadth
- Upper facial height
- Basion-Prosthion length
- Nasal breadth (max.)
- Lower nasal breadth
- Orbital breadth
- Biorbital breadth
- Foramen magnum breadth
- Cranial height
- Max. cranial length
- Bizygomatic breadth
- Total facial height
- Basion-Nasion length
- Basal height
- Upper nasal breadth
- Orbital height
- Interorbital breadth
- Palate-external breadth & length
- Palate-internal breadth & length
- …. (more)

# Automatic face recognition

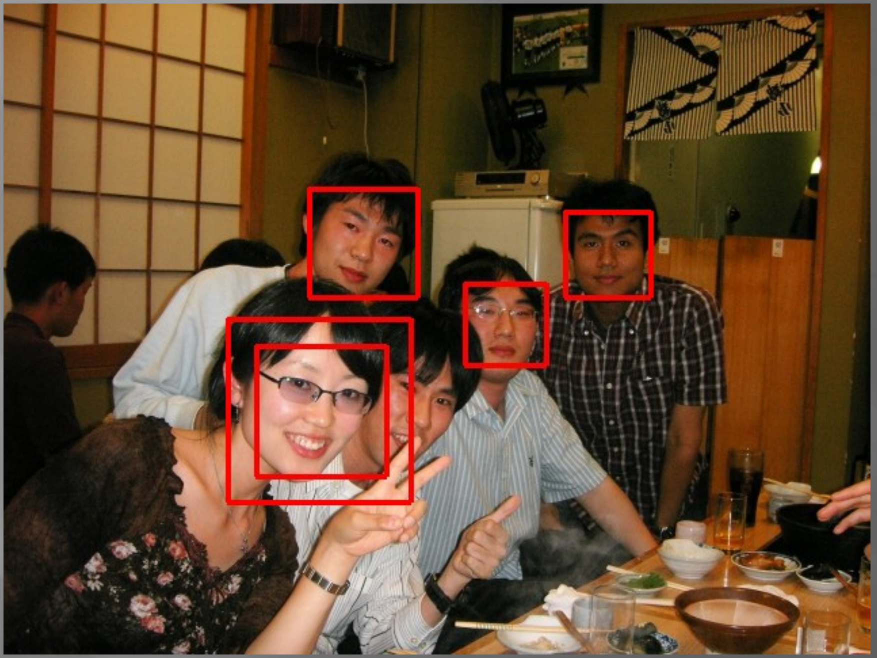
Vision Lab

**Face recognition involves several sequential processes:**

1. Face detection
2. Normalization
3. Facial features extraction
4. Registration/ Representation
5. Classification

# Face detection

Given a set of images find regions which contain instances of a face.



**Sebastian Marcel** - http://www.idiap.ch/~marcel/demos/comparison/opencv/alt/image_66-detect.jpg

# Face detection

**Sebastian Marcel** - http://www.idiap.ch/~marcel/demos/comparison/opencv/alt/image_66-detect.jpg

# Face detection

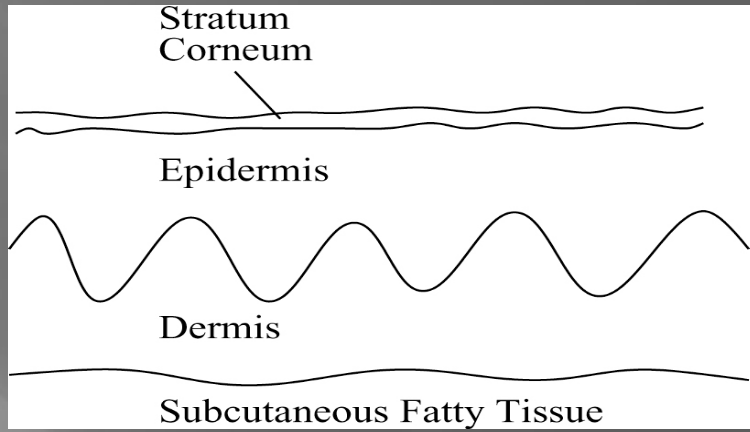
Vision Lab

| Approach | Representative Works |
|---|---|
| Knowledge-based | |
| | Multiresolution rule-based method [170] |
| Feature invariant | |
| – Facial Features | Grouping of edges [87] [178] |
| – Texture | Space Gray-Level Dependence matrix (SGLD) of face pattern [32] |
| – Skin Color | Mixture of Gaussian [172] [98] |
| – Multiple Features | Integration of skin color, size and shape [79] |
| Template matching | |
| – Predefined face templates | Shape template [28] |
| – Deformable Templates | Active Shape Model (ASM) [86] |
| Appearance-based method | |
| – Eigenface | Eigenvector decomposition and clustering [163] |
| – Distribution-based | Gaussian distribution and multilayer perceptron [154] |
| – Neural Network | Ensemble of neural networks and arbitration schemes [128] |
| – Support Vector Machine (SVM) | SVM with polynomial kernel [107] |
| – Naive Bayes Classifier | Joint statistics of local appearance and position [140] |
| – Hidden Markov Model (HMM) | Higher order statistics with HMM [123] |
| – Information-Theoretical Approach | Kullback relative information [89] [24] |

# Face detection
## Skin color

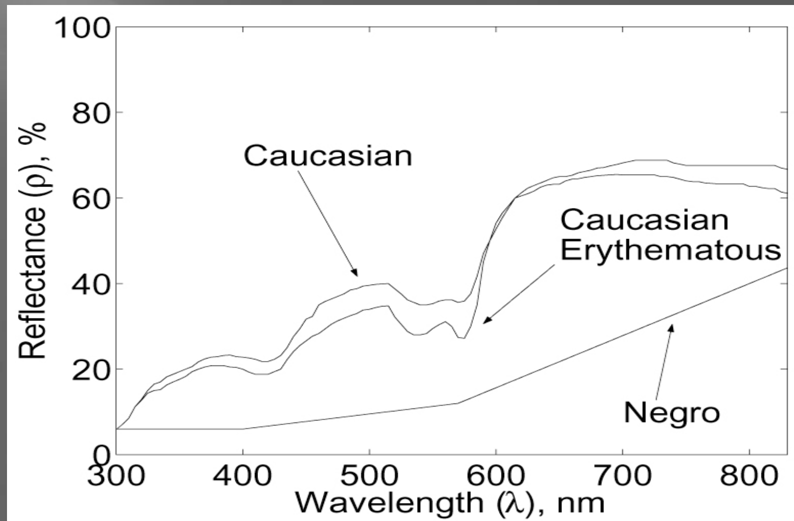
### Algorithms:

- Pixel-based
- Region-based

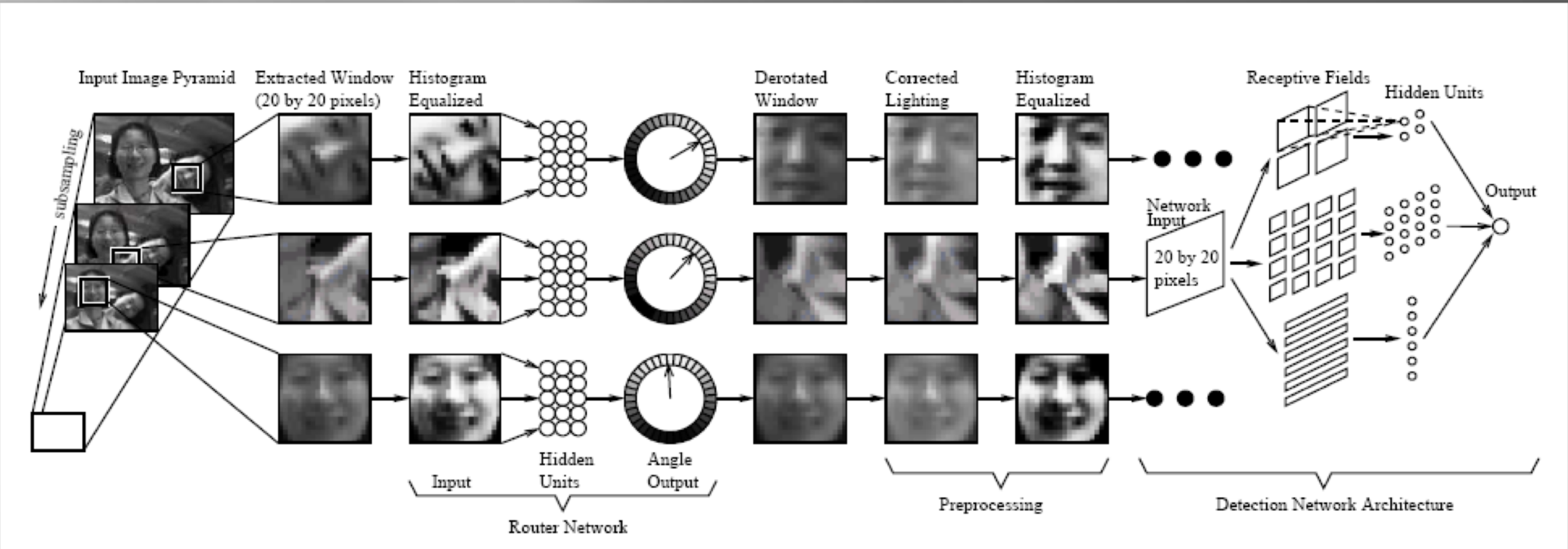



### Approaches:

- Explicitly defined region within a specific color space
- Dynamic skin distribution model



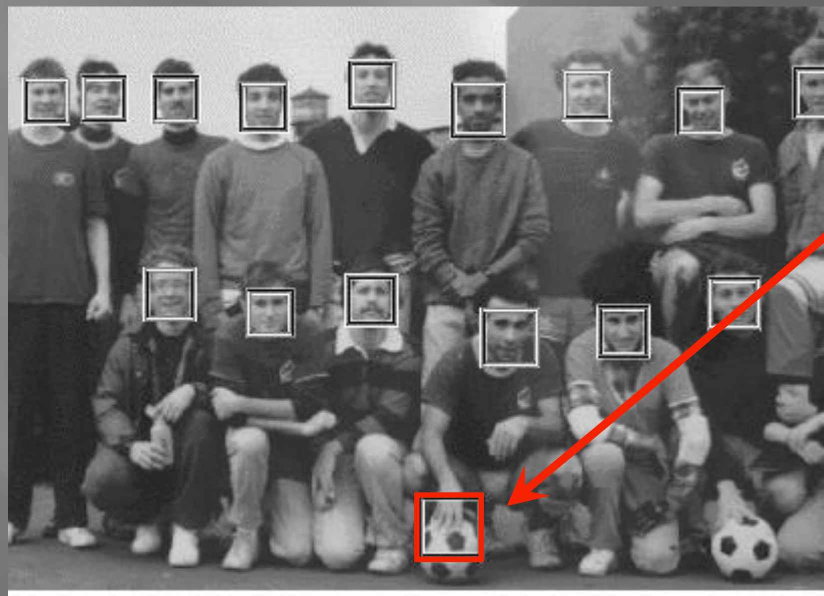
# Face detection
## Neural Networks




**Henry A. Rowley, Shumeet Baluja and Takeo Kanade** (1997).

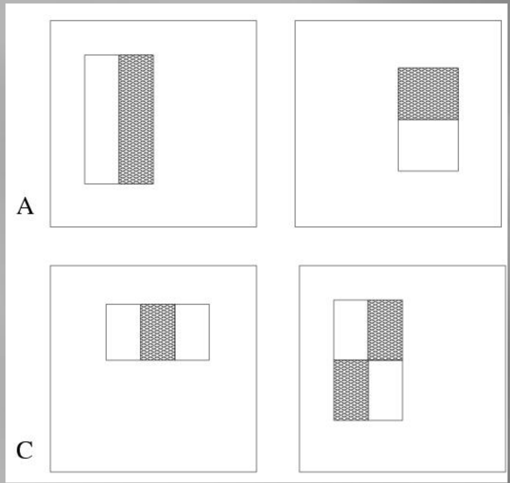
# The Viola-Jones Face detector

- At present the most efficient face detector implemented and largely adopted
- Publicly available within the Open-CV library
- Several improvements are being developed



**P. Viola, M. J. Jones**, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Features for face detection




**3 rectangular features types:**

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

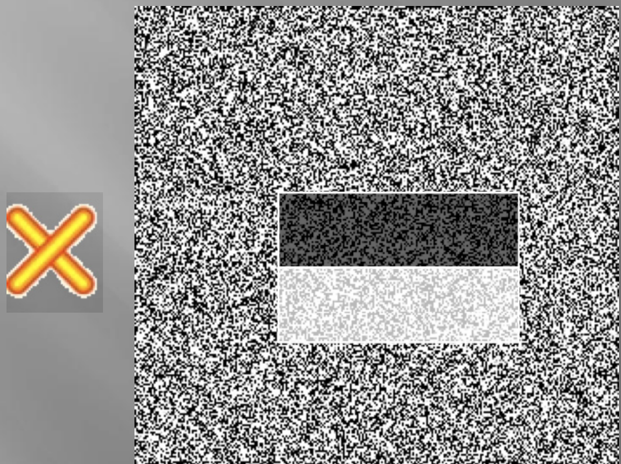
- Using rectangular (Haar) features, as opposed to more complex steerable filters, saves computation time.
- Estimation of the local contrast.

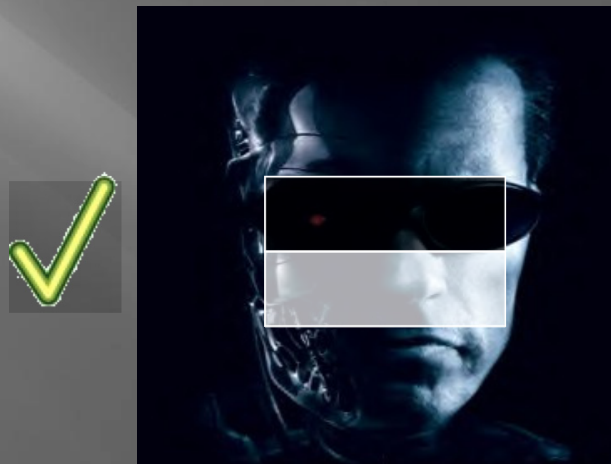


**P. Viola, M. J. Jones**, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Features for face detection

Using a 24x24 pixel base detection window, and all possible combination of location and scale, the full set is 49,396 features!



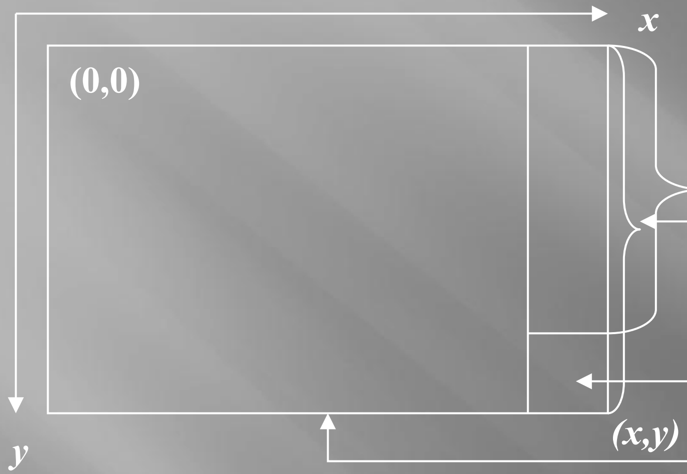




P. Viola, M. J. Jones, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

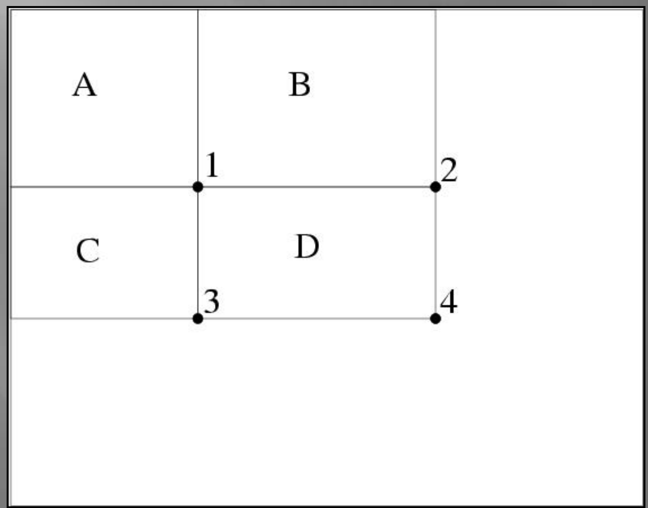
# The *Integral Image*

**The *Integral image* at location (*x*,*y*), is the sum of the pixel values above and to the left of (*x*,*y*), inclusive.**




$$s(x,y) = s(x,y\text{-}1) + I(x,y)$$

$$Ii(x,y) = Ii(x\text{-}1,y) + s(x,y)$$




The value of any rectangular sum is computed in constant time.

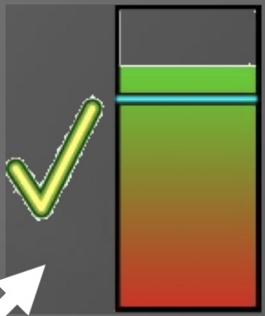
The integral sum inside rectangle ***D*** is computed as:

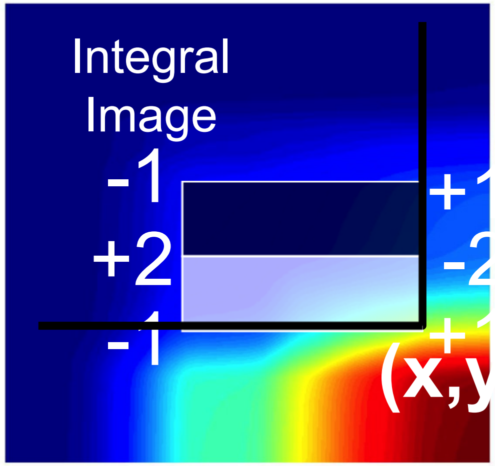
$$Ii(4) + Ii(1) – Ii(2) – Ii(3)$$



**P. Viola, M. J. Jones**, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# The *Integral Image*






P. Viola, M. J. Jones, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Learning classifiers

Vision Lab

- Given a feature set and labeled training set of images several machine learning techniques can be applied.
- 45,396 features are associated with each image sub-window, hence the computation of all features is computationally prohibitive.
- **Hypothesis**: A combination of only a small number of these features can yield an effective classifier.
- **Challenge**: Find these discriminant (weak) features.

**P. Viola, M. J. Jones**, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

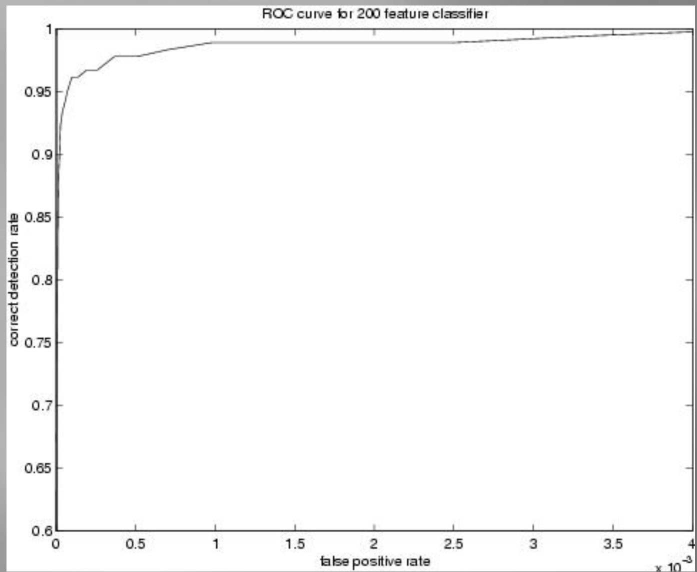
# Learning classifiers

- A weak classifier is a combination of a feature and a threshold, which best separate the examples
  - *K* features
  - *N* thresholds, where *N* is the number of examples
- Total of *KxN* weak classifiers
- For each feature sort the examples based on feature value
  - Select the “best” weak classifiers



P. Viola, M. J. Jones, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Performances: 200 features




The ROC curve of the constructed classifiers indicates that a detection rate of 0.95 can be achieved while maintaining an extremely low false positive rate of approximately $10^{-4}$.

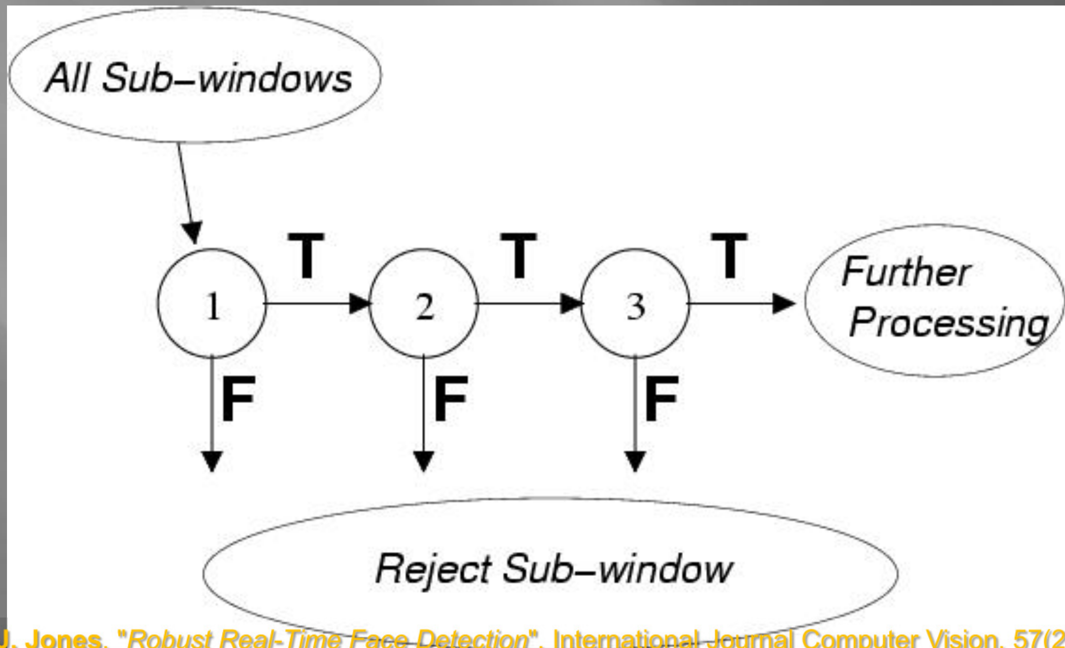
- First features selected by AdaBoost have a high discriminative power
- By varying the threshold of the final classifier a two-feature classifier can be obtained with a detection rate of 1 and a false positive rate of 0.4.



P. Viola, M. J. Jones, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Attentional cascade

- Simple, boosted classifiers can reject many 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.




P. Viola, M. J. Jones, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Attentional cascade

**Processing.** It is essentially identical to the processing performed by a degenerate decision tree:

- only a positive result from a previous classifier triggers the evaluation of the subsequent classifier.

$$IsFace = \sum_{i=1}^{N} F[i].score > thr$$

**Training.** It is also much like the training of a decision tree:

- subsequent classifiers are trained only on examples which passed all the previous classifiers.

  - **The task for classifiers further down the cascade is more difficult.**

**P. Viola, M. J. Jones**, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Performances

- 4,916 positive training examples are hand picked aligned, normalized, and scaled to a base resolution of 24x24 pixels
- 10,000 negative examples are selected by randomly picking sub-windows from 9,500 images not containing faces



P. Viola, M. J. Jones, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Performances: CMU DB

| False positive | 10 | 31 | 50 | 65 | 78 | 95 | 110 | 167 | 422 |
|---|---|---|---|---|---|---|---|---|---|
| Viola-Jones | 78.3% | 85.2% | 88.8% | 89.8% | 90.1% | 90.8% | 91.1% | 91.8% | 93.7% |
| Rowley-Baluja-Kanade | 83.2% | 86.0% | - | - | - | 89.2% | - | 90.1% | 89.9% |
| Schneiderman-Kanade | - | - | - | 94.4% | - | - | - | - | - |
| Roth-Yang-Ajuha | - | - | - | - | 94.8% | - | - | - | - |

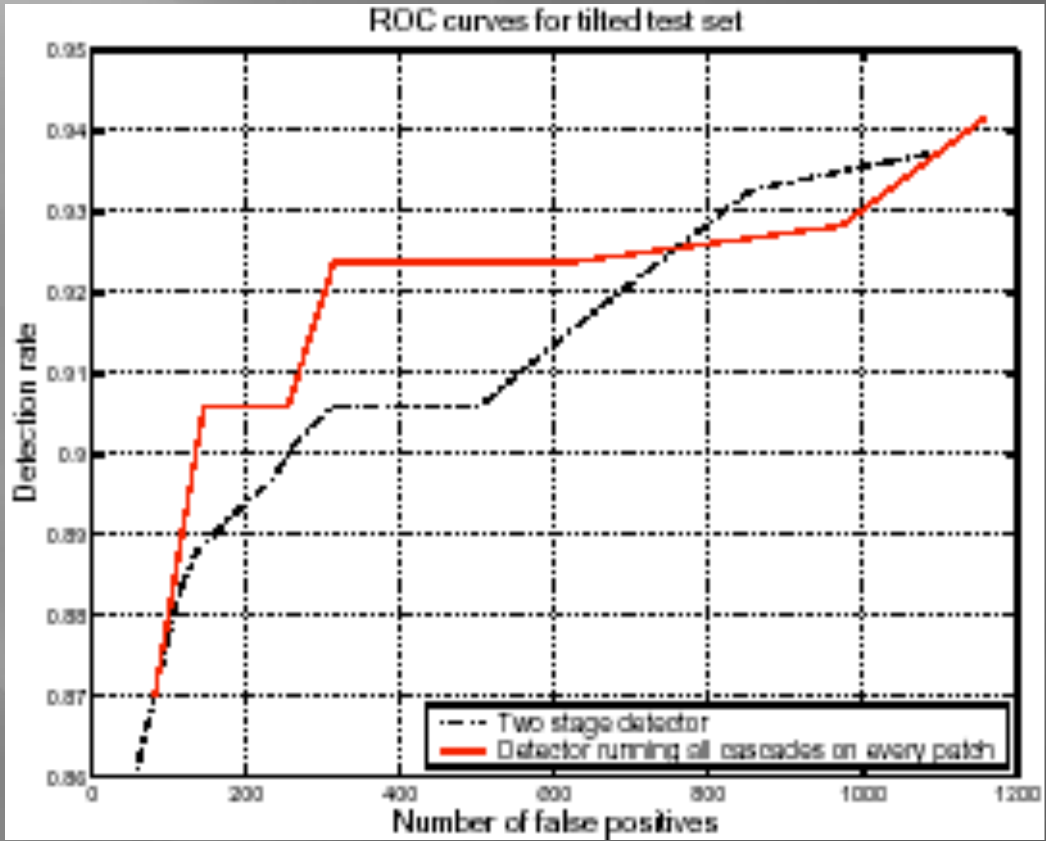
- 130 images
- 505 labeled frontal faces
- The final detector has 32 layers and 4,297 features
- The processing time of a 384 by 288 pixel image on a conventional PC is about .067 seconds.



P. Viola, M. J. Jones, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Performances: CMU DB

Vision Lab




Results over a total of 10,515,781 analyzed image windows

P. Viola, M. J. Jones, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Performances: CMU DBII

P. Viola, M. J. Jones, "*Robust Real-Time Face Detection*", International Journal Computer Vision, 57(2), pp. 137-154, 2004.

# Some improvements…

Vision Lab

Improve localization accuracy by post-filtering the face region

QuickTime™ e un
decompressore
sono necessari per visualizzare quest'immagine.

QuickTime™ e un
decompressore
sono necessari per visualizzare quest'immagine.

QuickTime™ e un
decompressore
sono necessari per visualizzare quest'immagine.

QuickTime™ e un
decompressore
sono necessari per visualizzare quest'immagine.

QuickTime™ e un
decompressore
sono necessari per visualizzare quest'immagine.

QuickTime™ e un
decompressore
sono necessari per visualizzare quest'immagine.

# Automatic face recognition

**Face recognition involves several sequential processes:**

- **Face detection**
- **Normalization**
- **Facial features extraction**
- **Registration/ Representation**
- **Classification**

# The "illumination dilemma"

- **THE Problem**

Histogram equalization

Original Image

# The "*illumination dilemma*"

Vision Lab

- **THE Problem** is related to the *energy* and the ***position*** of the light sources
  - You need:
    - **position and intensity of light sources;**
    - **position and orientation of the face;**
    - **the reflectance map of the face**
- **THE Solution** … simply doesn't exist
- Several ***approximations*** can be guessed

# The "*illumination dilemma*"

## Main techniques:

- Histogram-based adaptive techniques, applied on image patches
- Re-lighting techniques
- Synthesis of illumination-invariant representations (for example the *Hue* component in color space)

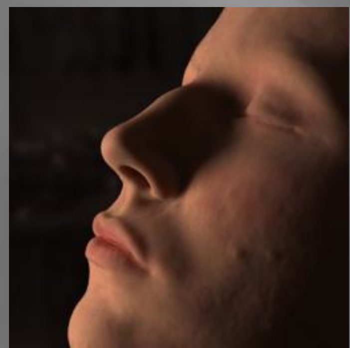
# Skin modelling

Vision Lab

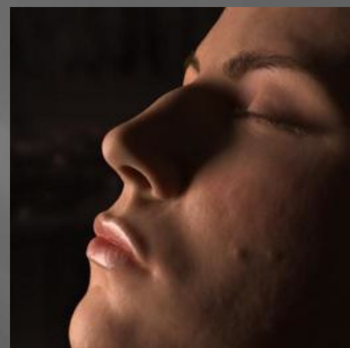
Skin chromaticity map

Diffuse light rendering

Reflectance map of the oily skin layer



Sub-surface reflectance



Final face rendering

***Henrik Wann Jensen,*** *"Digital face cloning", SIGGRAPH'2003 Technical Sketch, San Diego, July 2003. (http://graphics.ucsd.edu/~henrik/papers/face_cloning/)*

# Image re-lighting

$$I(x,y) = R(x,y) \cdot L(x,y) \qquad R(x,y) = \frac{I(x,y)}{L(x,y)} \quad \xrightarrow{?} \quad L(x,y)$$

$$F(L) = \iint_{\Omega} \rho(x,y)\left(L(x,y) - I(x,y)\right)^2 dxdy + \lambda \iint_{\Omega} (L_x^2 + L_y^2)\,dxdy \qquad (1)$$

Anisotropic diffusion
(Lagrange solution of (1))

Isotropic diffusion
(Gaussian filtering)

**R. Gross and V. Brajovic**, "An Image Preprocessing Algorithm for Illumination Invariant Face Recognition", International Conference on Audio- and Video-Based Biometric Person Authentication, 2003.

**D. Jobson, Z. Rahmann and G. Woodell**, "A Multiscale Retinex for Bridging the Gap Between Color Images and the Human Observations of Scenes", IEEE Transanctions on Image Processing, volume 6, Issue 7, 1997.

# Automatic face recognition

**Face recognition involves several sequential processes:**

- Face detection
- Normalization
- Facial features extraction
- Registration/ Representation
- Classification

# Taxonomy of face recognition

## Algorithms

- **Pose-dependent**
  - **Viewer-centered Images**
    - **Appearance-based (Holistic)**: PCA, LDA, ICA, LPP, Kernel . . .
    - **Feature-based (Analytic)**: Gabor sets, LBP . . .
    - **Hybrid**: LFA, EGBM/JETs, AAM . . .
- **Pose-invariant**
  - **Object-centered Models**
    - Gordon et al., 1995 3D from 2D
    - Lengagne et al., 1996 3D from stereo
    - Atick et al., 1996 LFA
    - Yan et al., 1996 Geometric modeling
    - Zhao et al., 2000 Shape from Shading
    - Zhang et al., 2000 3D from video

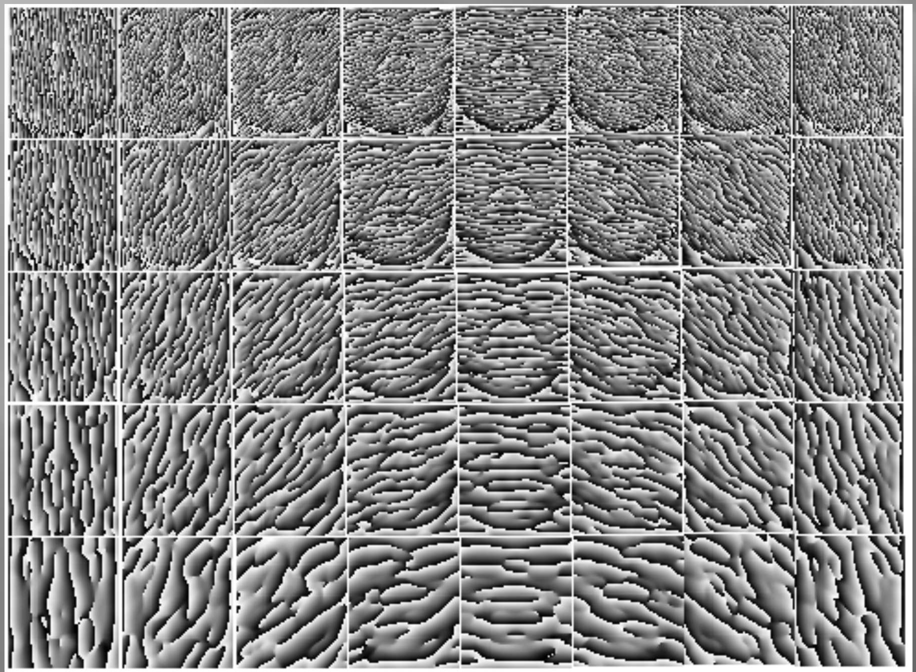
Pose-dependency

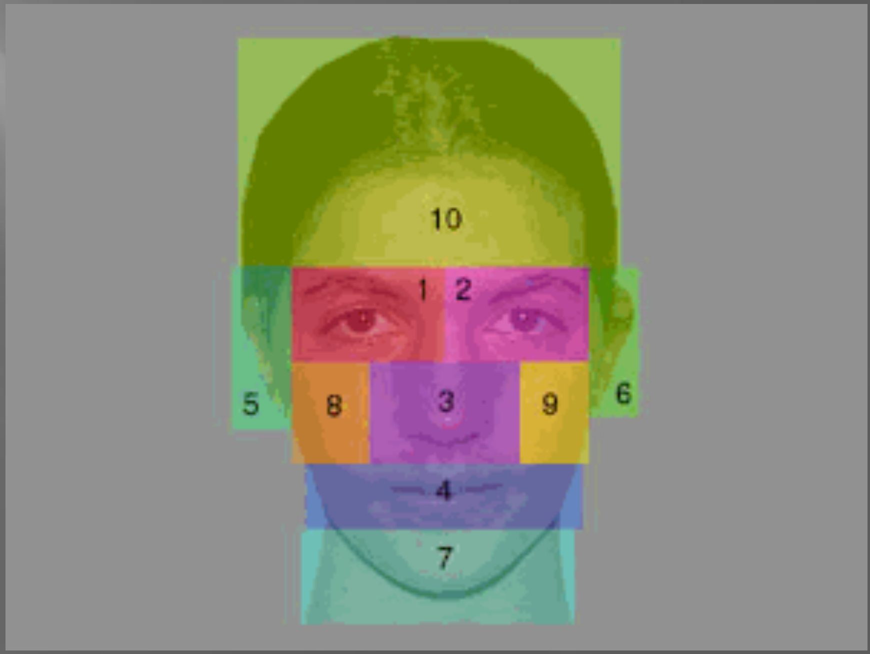
Face representation

Matching features

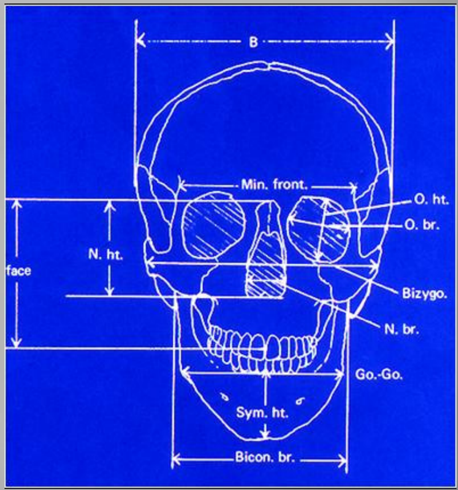
# Understanding *Facial Features*

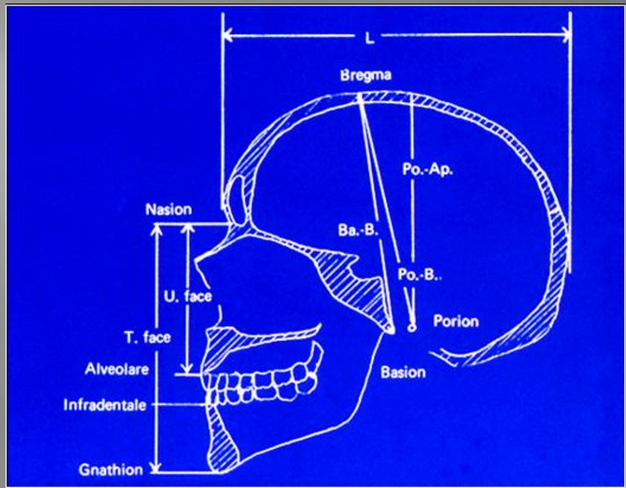
- Gray level oriented patterns/properties
- Physical Landmarks

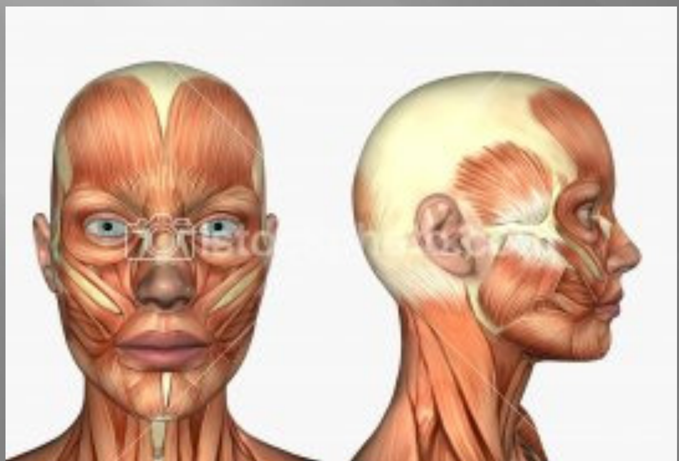




# How to define *facial features*



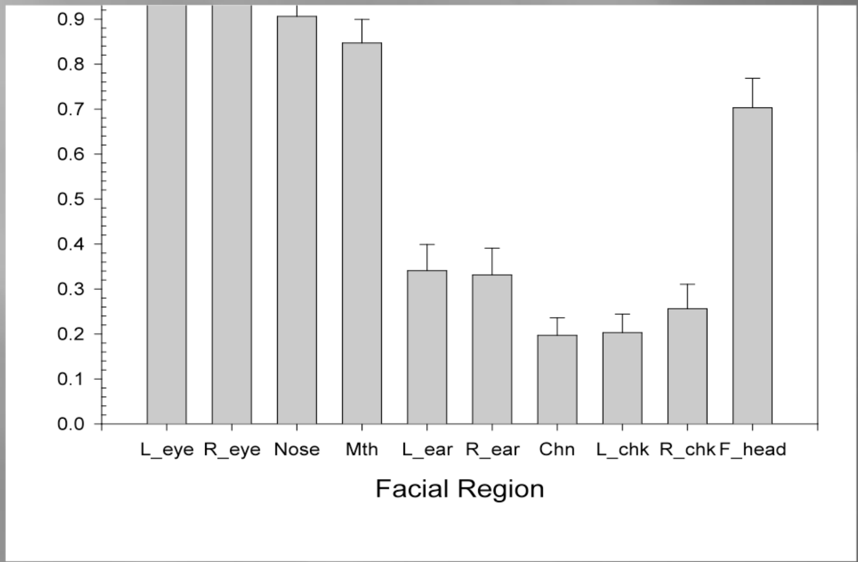



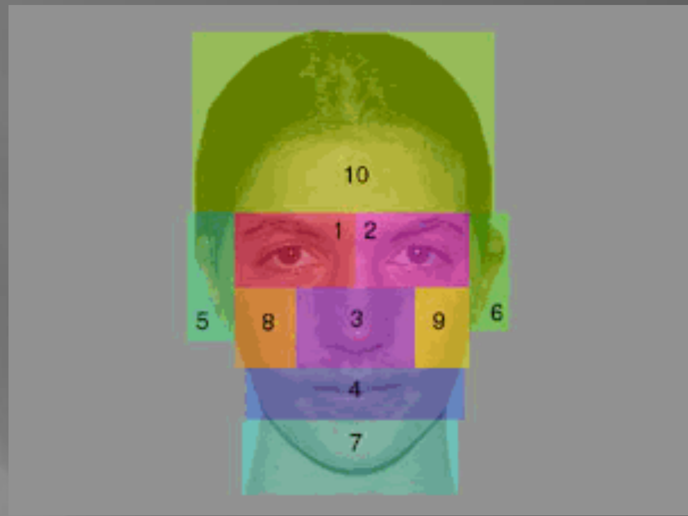



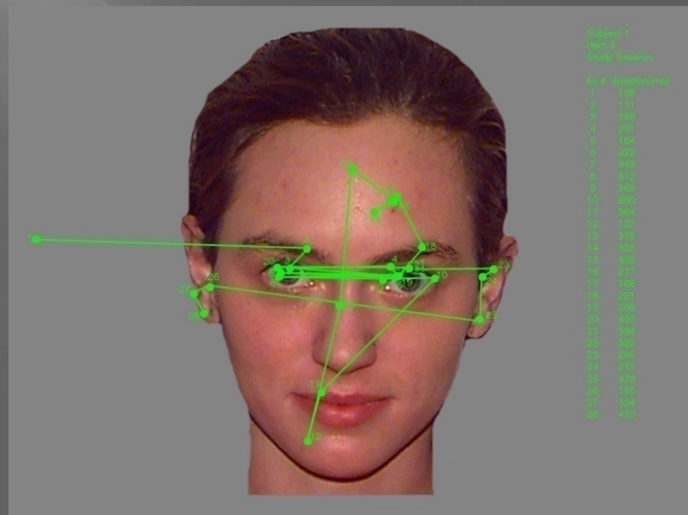
**Craniometric measurements (1)**

Cranial circumference
Max. cranial breadth
Min. frontal breadth
Bigonial breadth
Upper facial height
Basion-Prosthion length
Nasal breadth (max.)
Lower nasal breadth
Orbital breadth
Biorbital breadth
Foramen magnum breadth
Cranial height
Max. cranial length
Bizygomatic breadth
Total facial height
Basion-Nasion length
Basal height
Upper nasal breadth
Orbital height
Interorbital breadth
Palate-external breadth & length
Palate-internal breadth & length
…. (more)

# How to define *facial features*

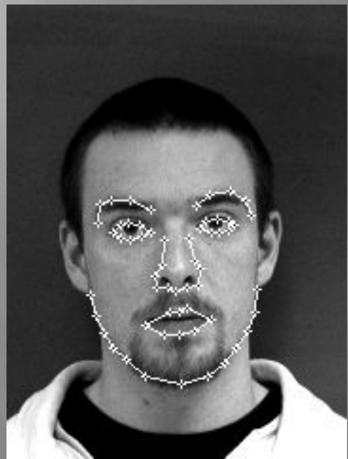


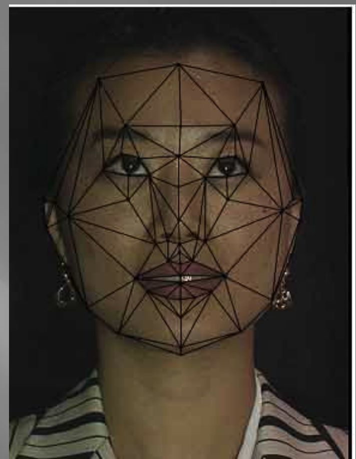



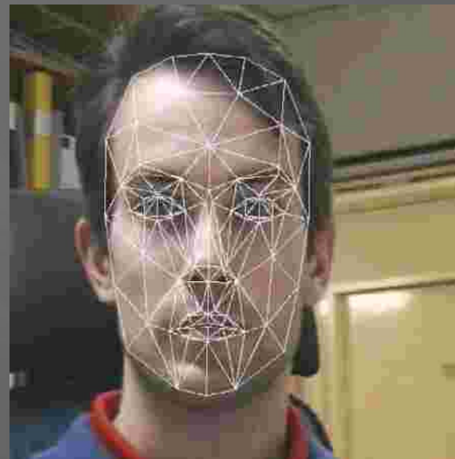



**J.H. Henderson et al.** "Gaze Control for Face Learning and Recognition by Humans and Machine"; in T. Shipley and P. Kellman (Eds.), *From Fragments to Objects: Segmentation and Grouping in Vision*

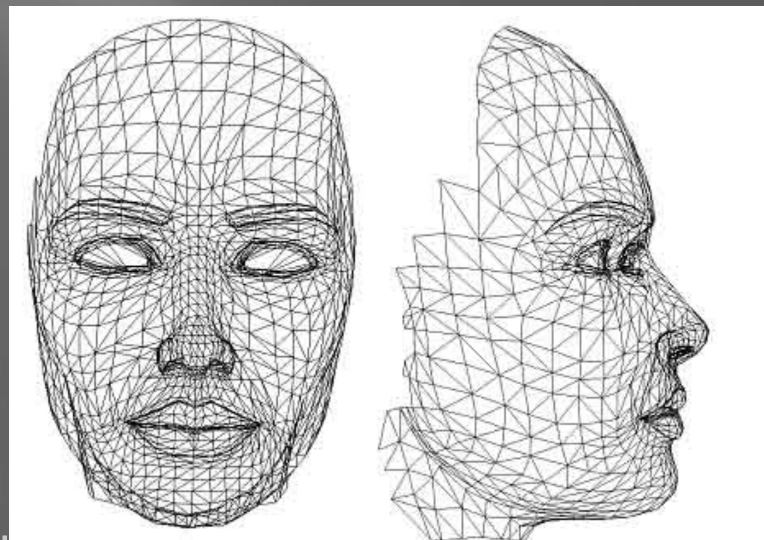
# *Facial features* as 2D/3D Landmarks



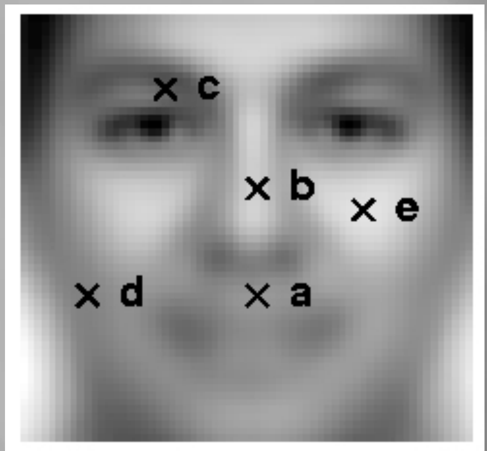




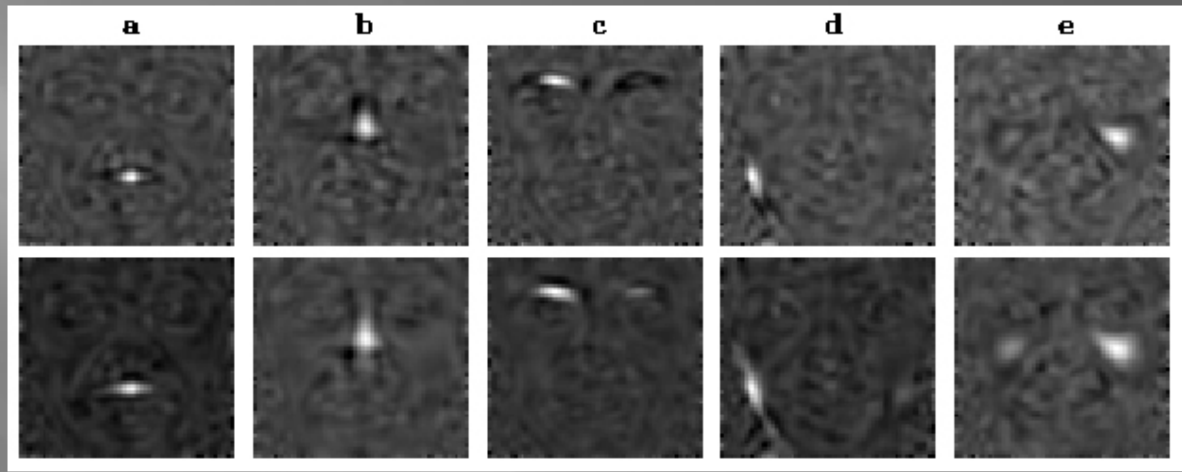
- 2D landmarks can be defined and tracked on face images
- Simple 2D vs complex 3D representations



# How to define *facial features*




Marked average face image



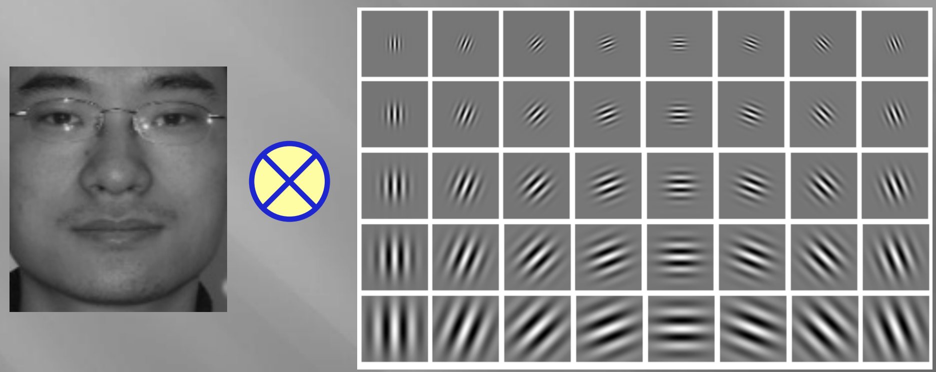

- Five topographic kernels are shown in the top row
- Five corresponding residual correlations (response) in the bottom row.

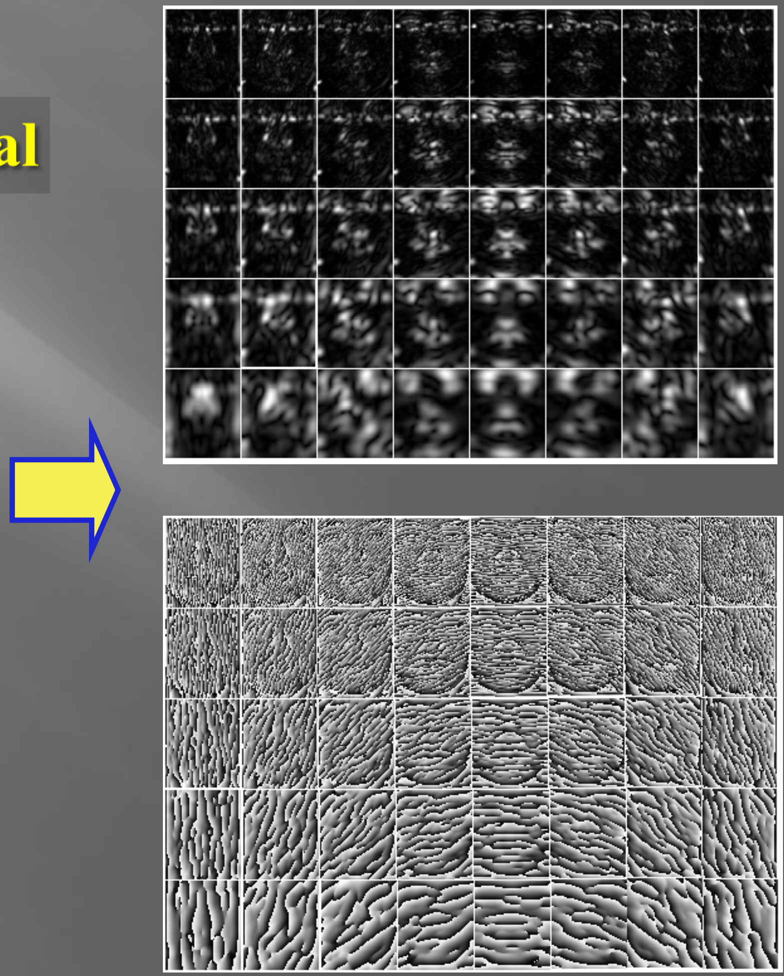
- The **LFA**-based approach (*Local Feature Analyisis*) uses localized kernels, constructed from PCA-based eigenvectors, for extracting topographic facial features (e.g., eyebrows, cheek, mouth, etc.)

# Facial features as 2D patterns

## Gabor wavelets

- Provide a description of the local structure of the facial patterns

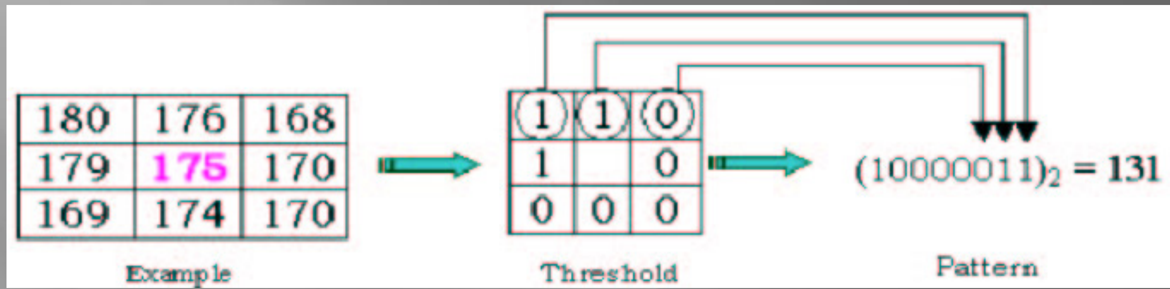




- Convolution with a bank of frequency-tuned filters

# Facial features as 2D patterns

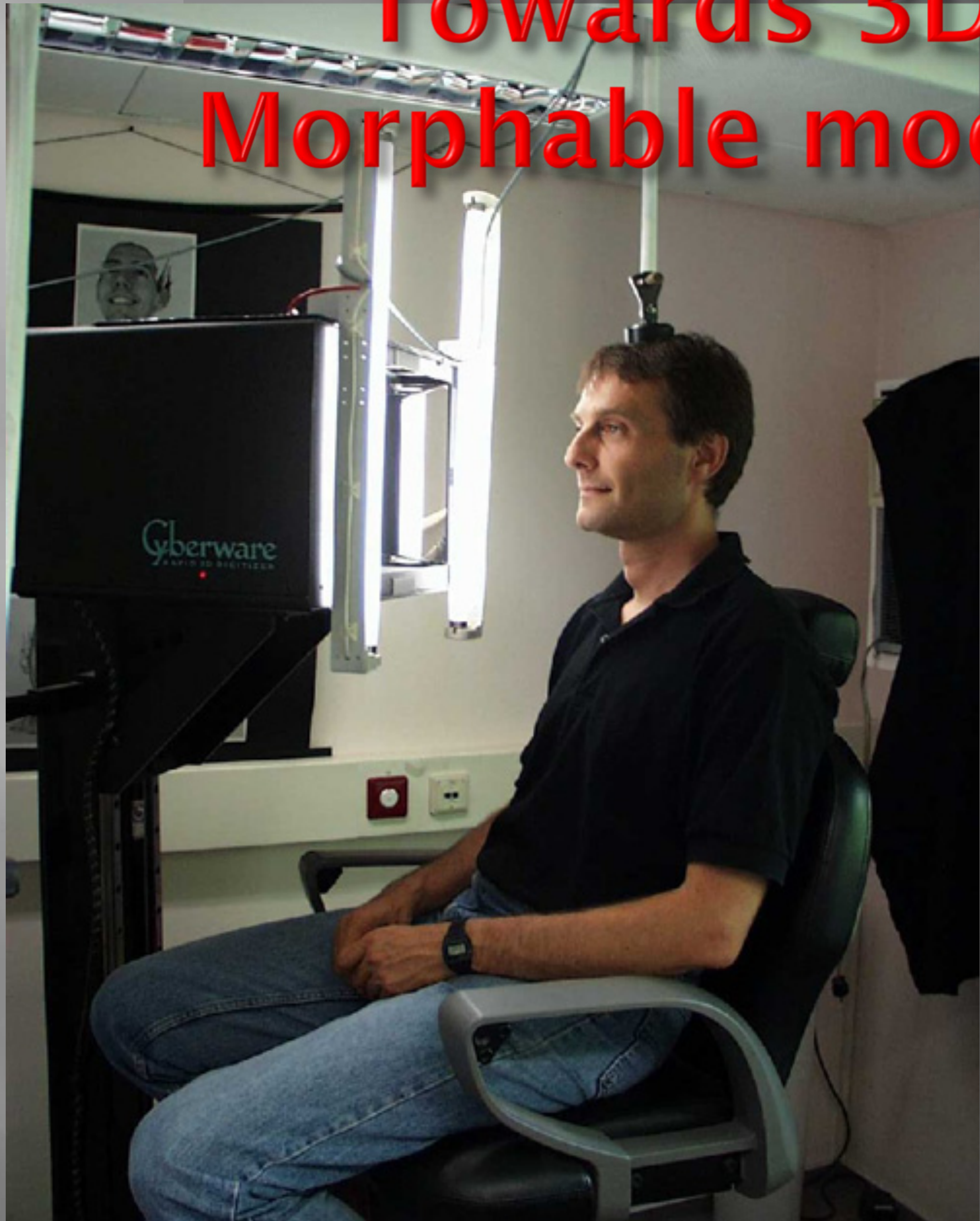
## Local Binary Patterns (LBP)




- Pixels are labeled by thresholding the 3x3neighbourhood with the center value and considering the result as a binary number.
- The histogram of the labels is used as a texture descriptor.

# Towards 3D: Morphable models

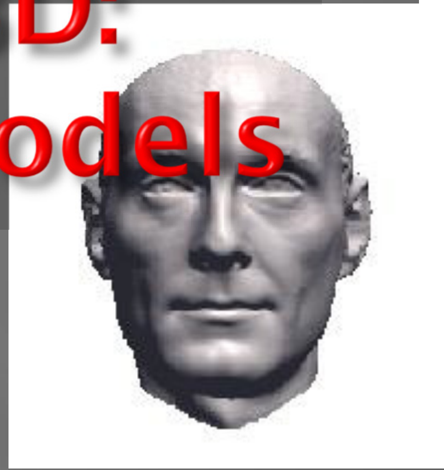
Vision Lab




Laser Scanner

Max Planck Institute
Biologische Kybernetik

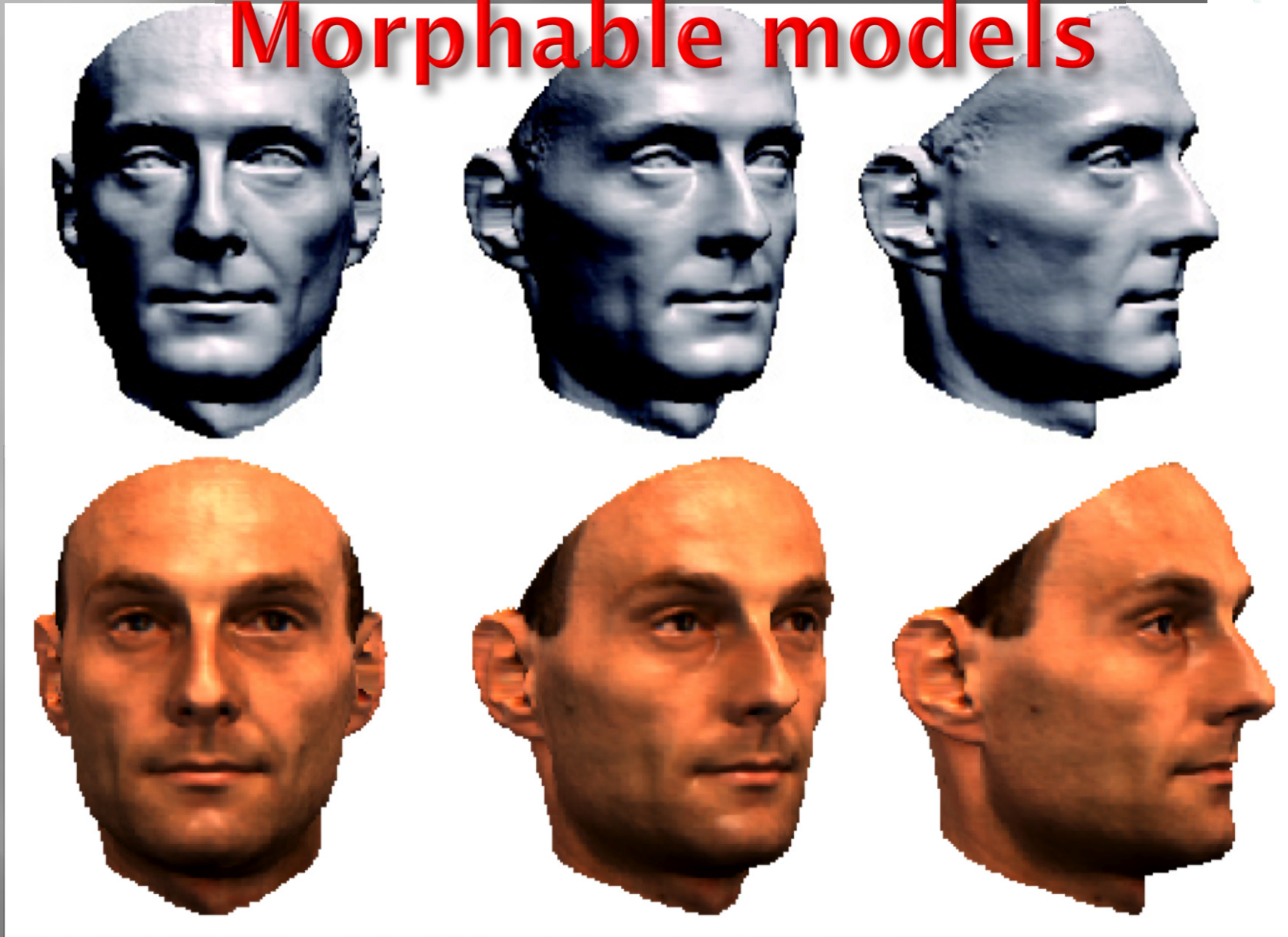
# Towards 3D: Morphable models



3D shape

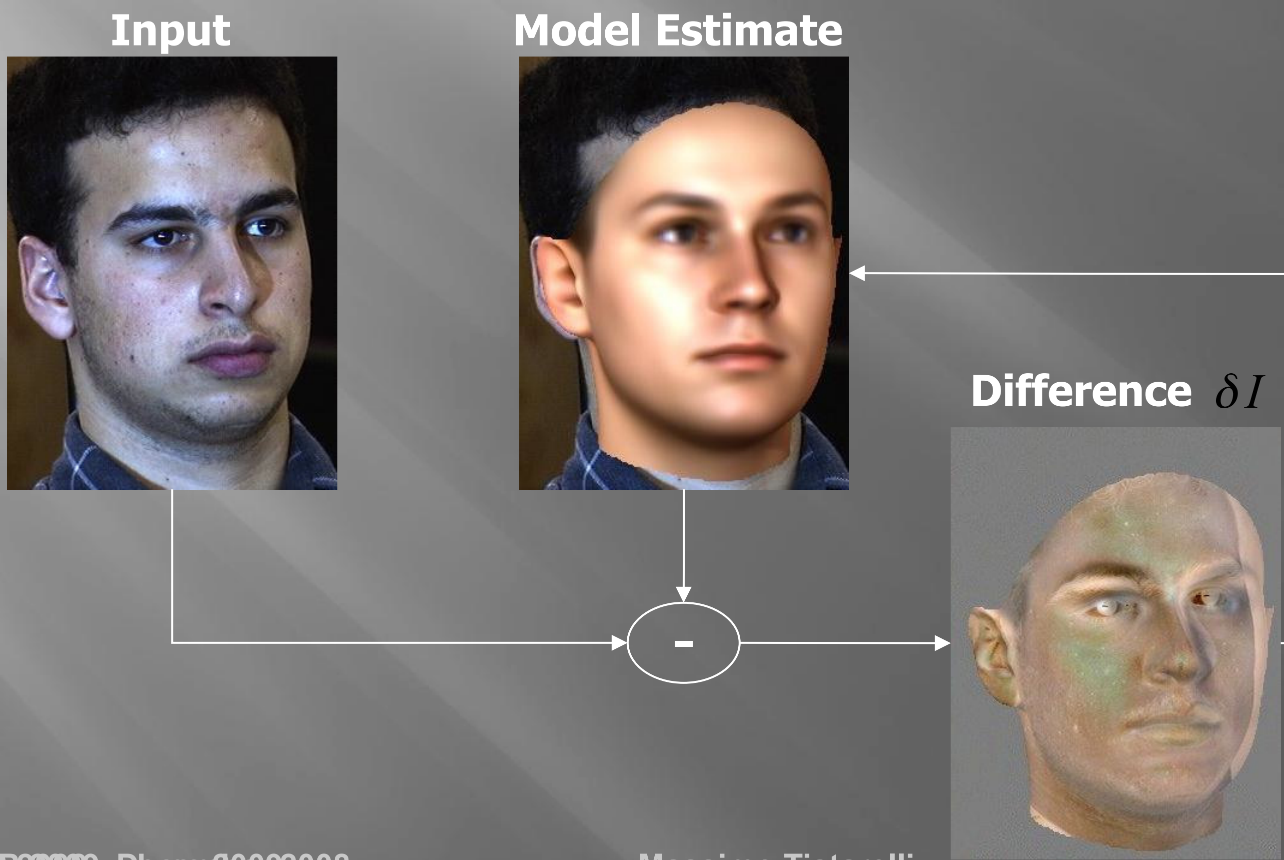
surface reflectance

# Towards 3D: Morphable models



# Towards 3D: Morphable models

- Jones, Poggio 98 : Gradient Descent
- Blanz, Vetter 99 : Stochastic Gradient Descent
- Pighin, Szeliski, Salesin 99 : Levenberg-Marquardt
- Romdhani, Blanz, Vetter 02: Non-linear fitting

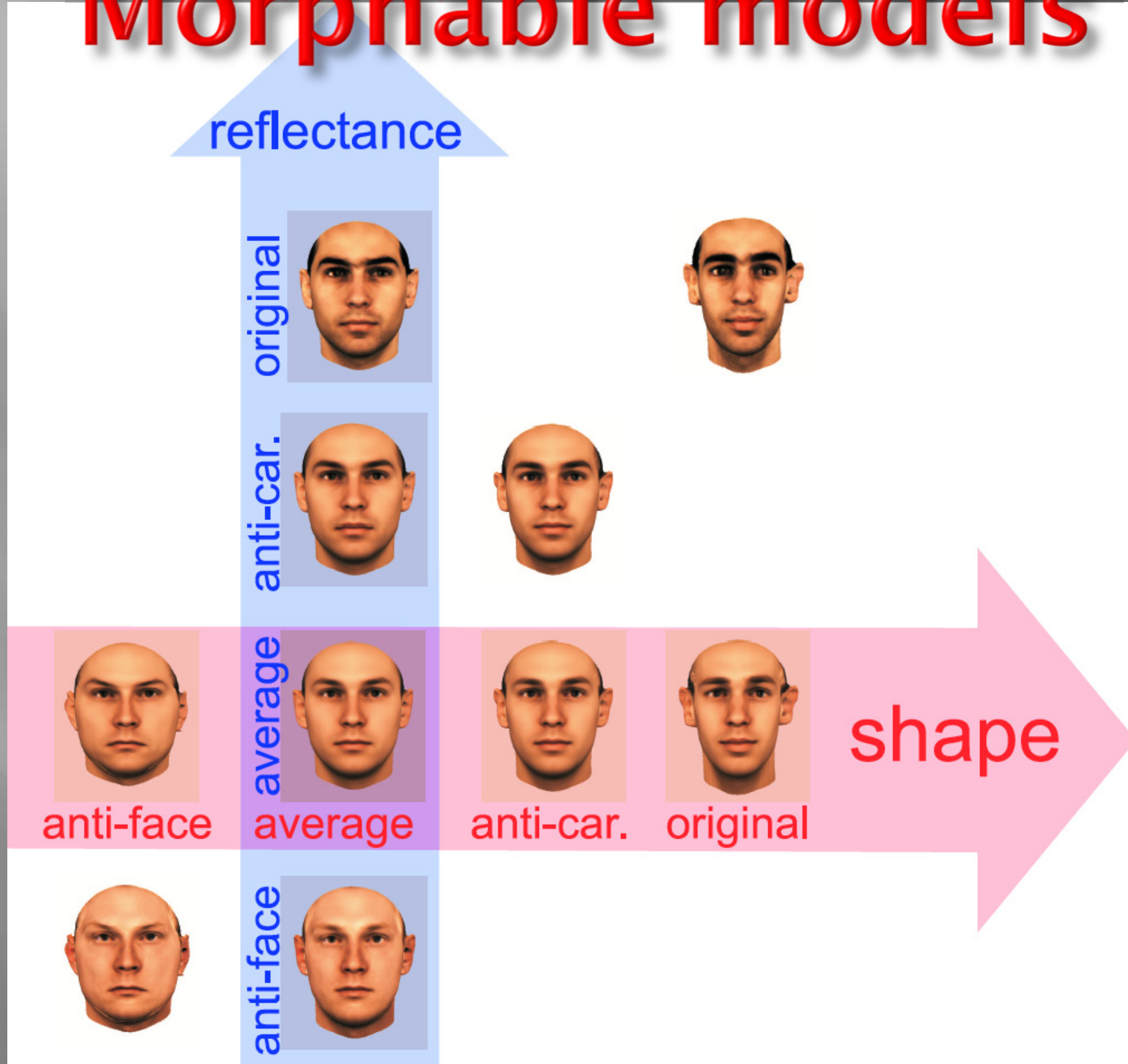



$$\delta\alpha = -\lambda \frac{\partial}{\partial\alpha}\left(\delta I^2\right)$$

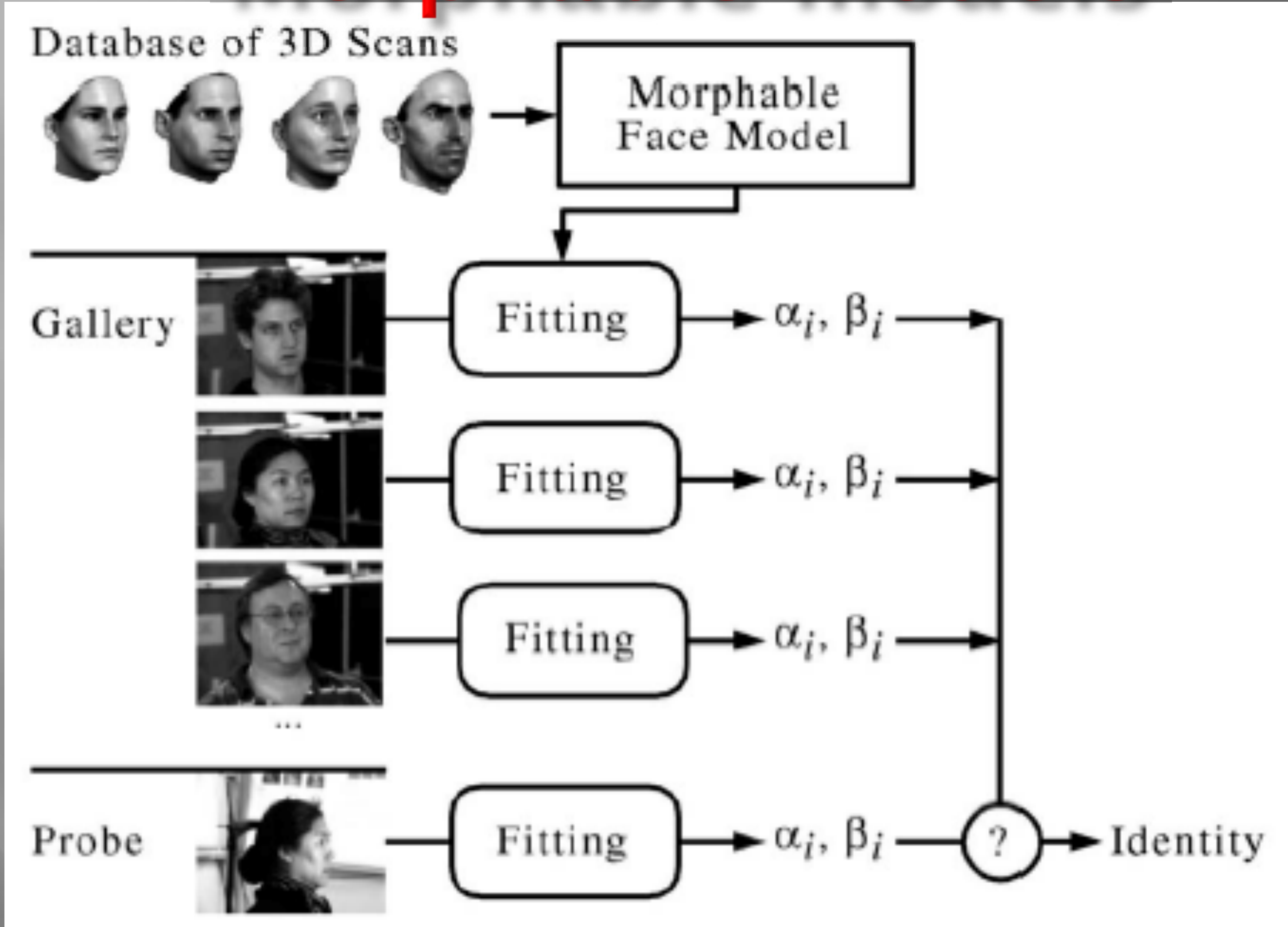
$$\delta\beta = -\lambda \frac{\partial}{\partial\beta}\left(\delta I^2\right)$$

$$\delta\rho = -\lambda \frac{\partial}{\partial\rho}\left(\delta I^2\right)$$

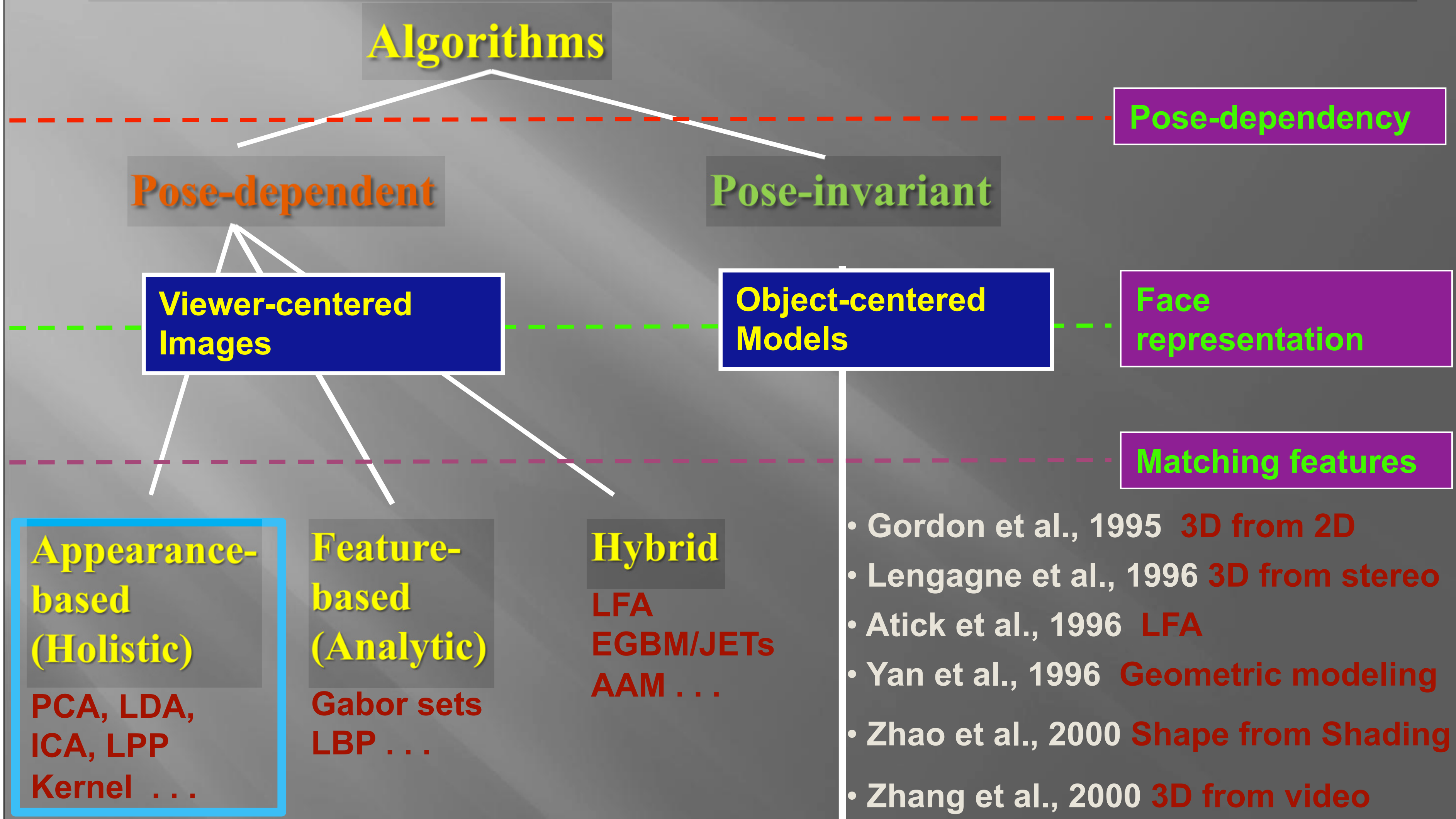
# Towards 3D: Morphable models



# Towards 3D: Morphable models




# Taxonomy of face recognition



# Holistic face recognition

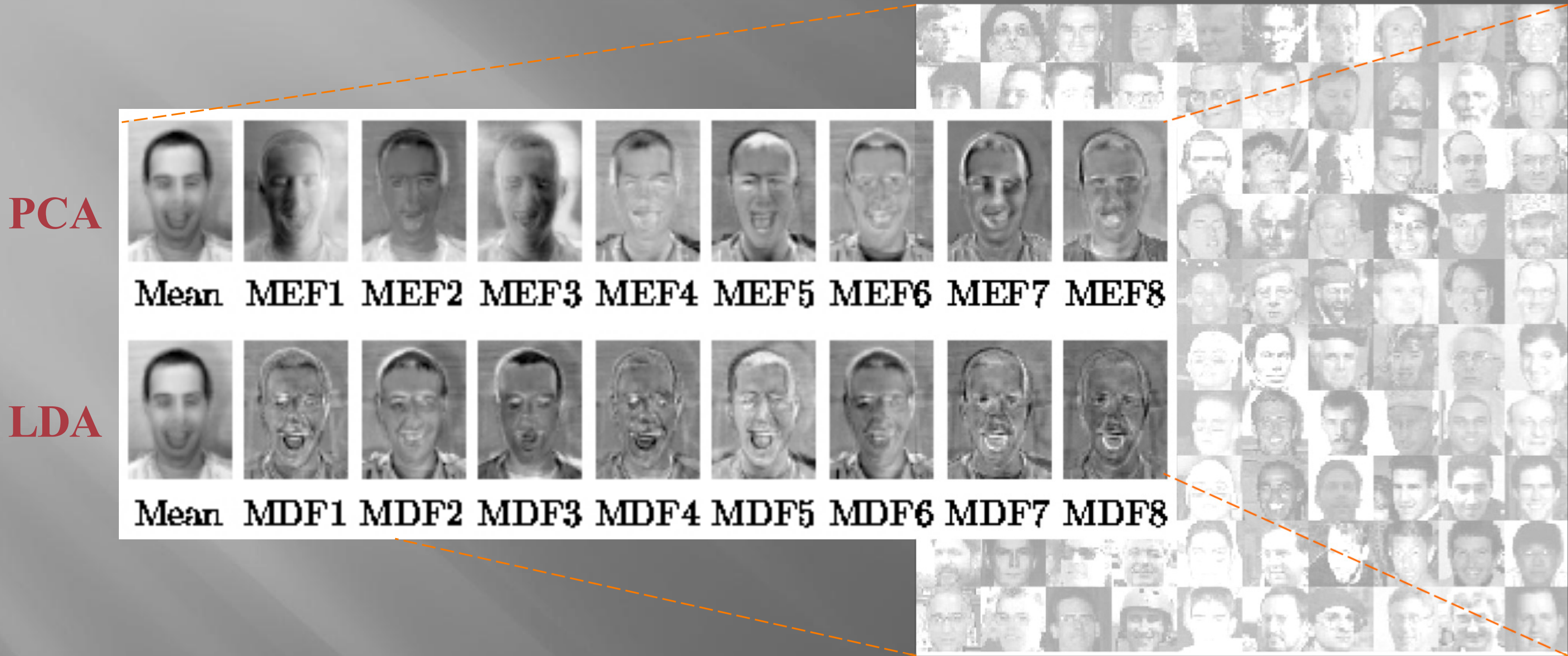
Vision Lab

- The basic idea of many similar approaches is to define a basis of vectors to describe any face in the "*universal space*" of all existing faces…
- The basic tool is the *Singular Values Decomposition*:

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{W}$$

- The **eigenvectors** ($r$ columns of U) of the decomposition define the basis of vectors and the eigenvalues define the "saliency" of each eigenvector (*eigenface)*

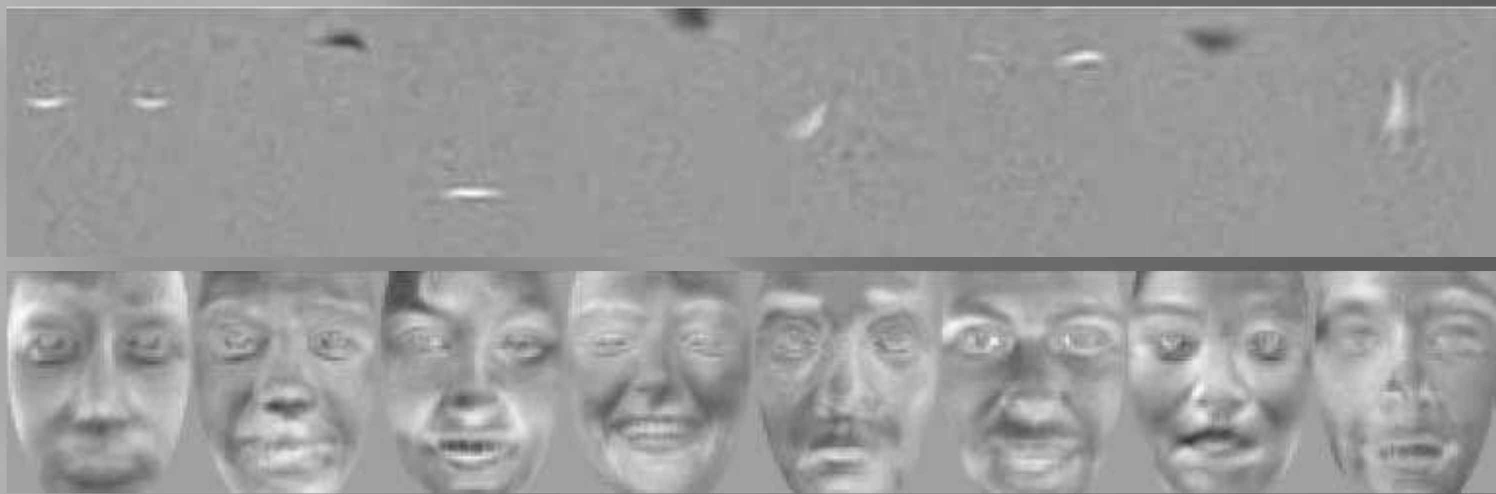
# Holistic face recognition




- The PCA-based and LDA-based approaches make use of a set of orthogonal basis images
- Both provide a compact and global representation of face images

# Holistic face recognition

ICA

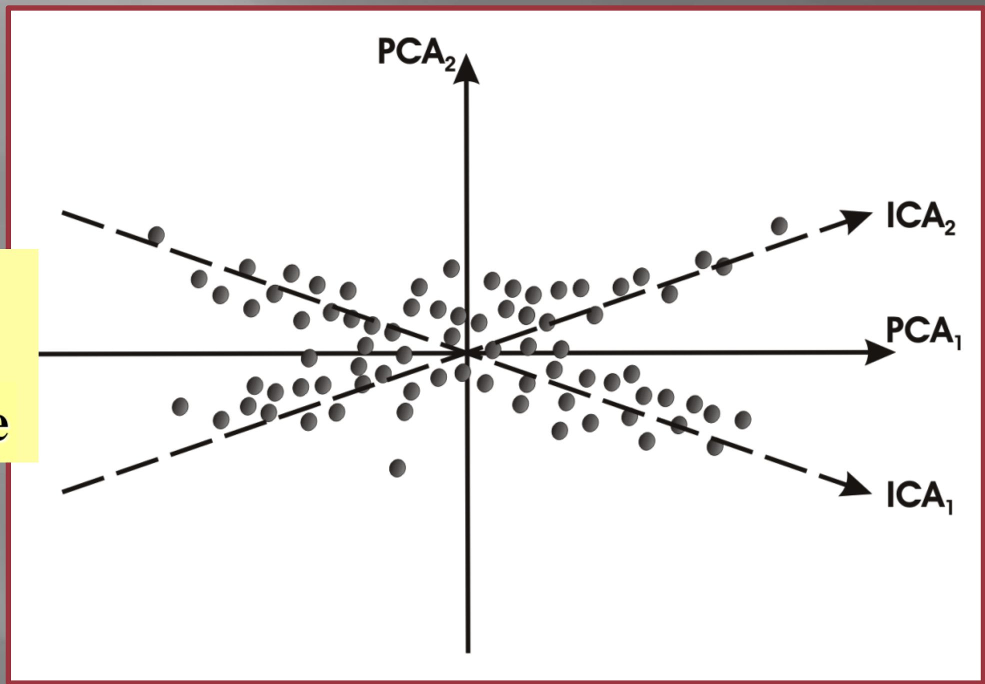


- The Independent Component Analysis (ICA) is based on a higher order optimization to find independent (orthonormal) components for the face sub-spaces
- Better description of the inter-class variability

# Holistic face recognition

Vision Lab

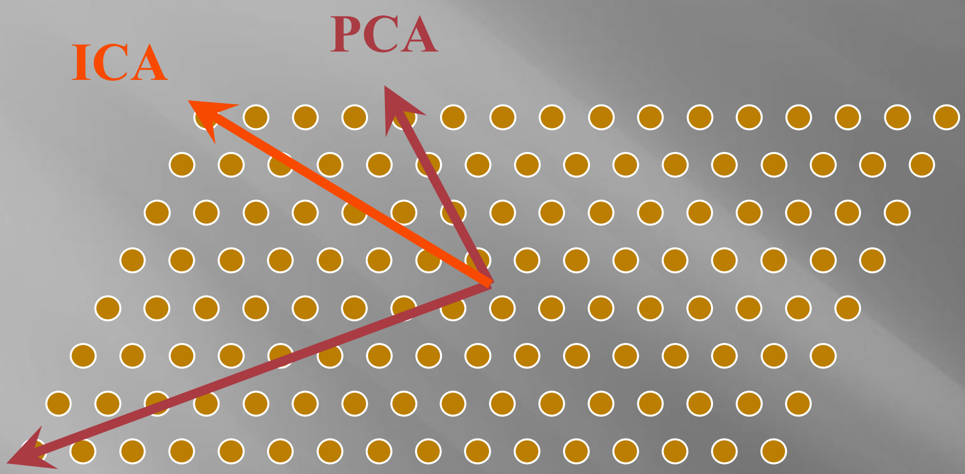
ICA



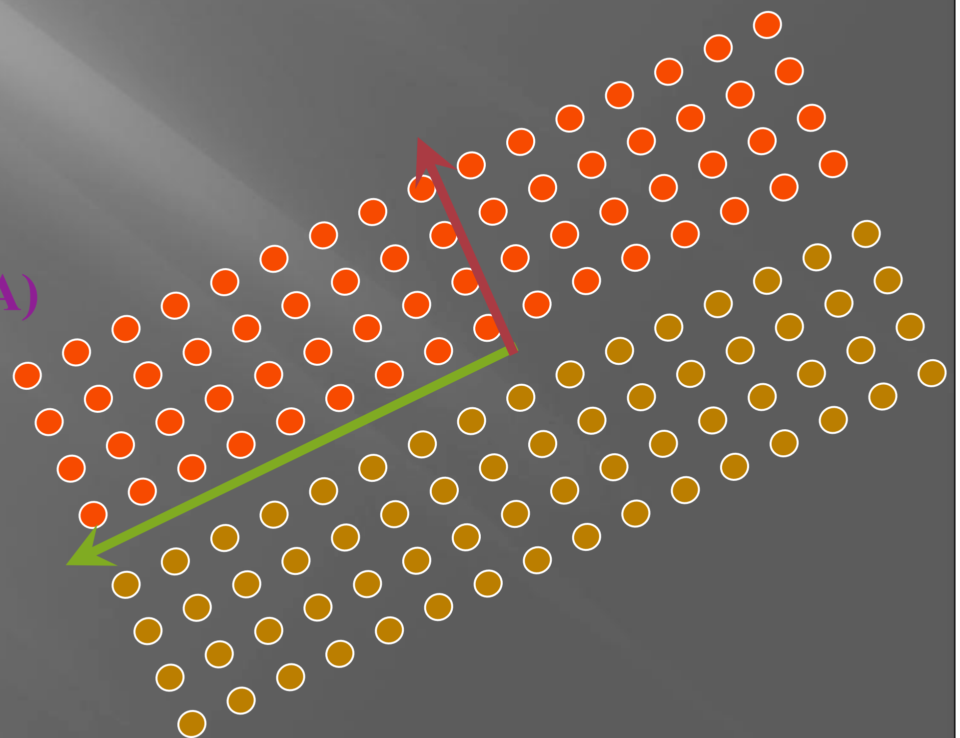

**Orthonormality**
**vs**
**Maximal variance**

PCA

# Holistic face recognition




- Relevant Component Analysis (RCA)
- Fisher Discriminant analysis (FDA)
- Locality Preserving Projections (Laplacianfaces)
- Others . . .

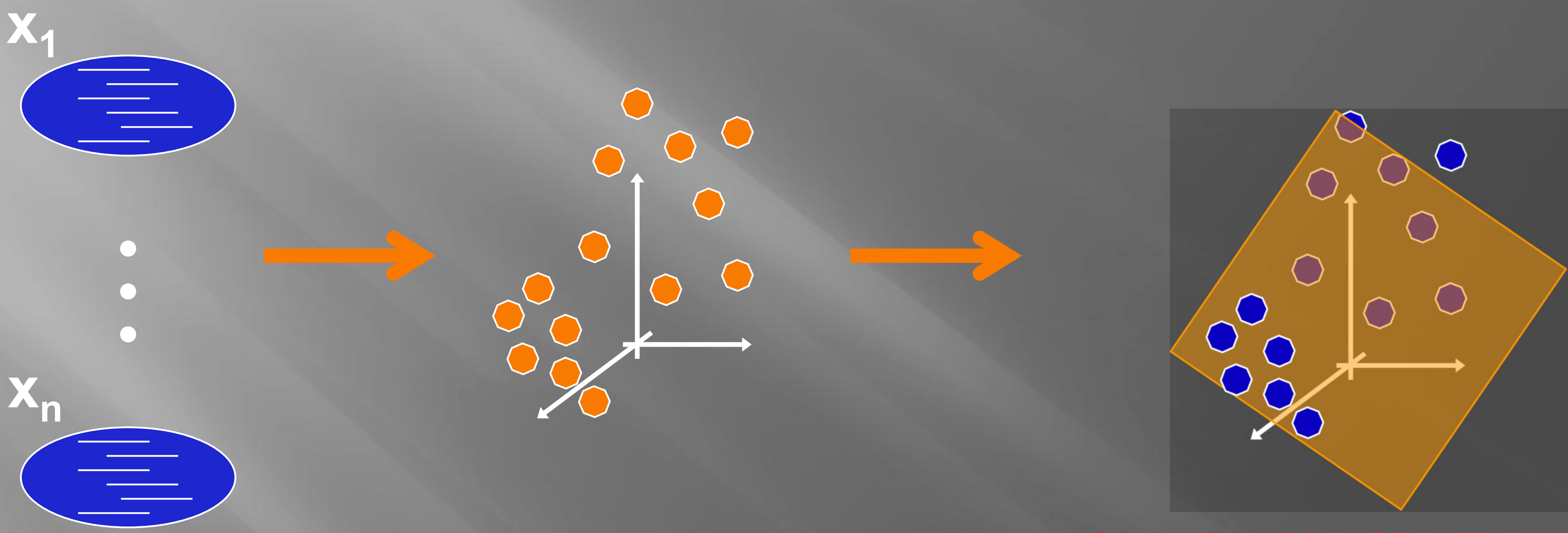


# Kernel methods

- K-PCA; K-ICA; K-LDA… (B. Schölkopf et al. 1998)
- Are all variations of existing face-space representations where the transformation to the lower space is mediated by a kernel function such as Gaussian, polinomial, sigmoid and Radial Basis Functions
- More robust to noise and discretization
- Better separation of classes
- General *Learning Theory*

# Kernel methods

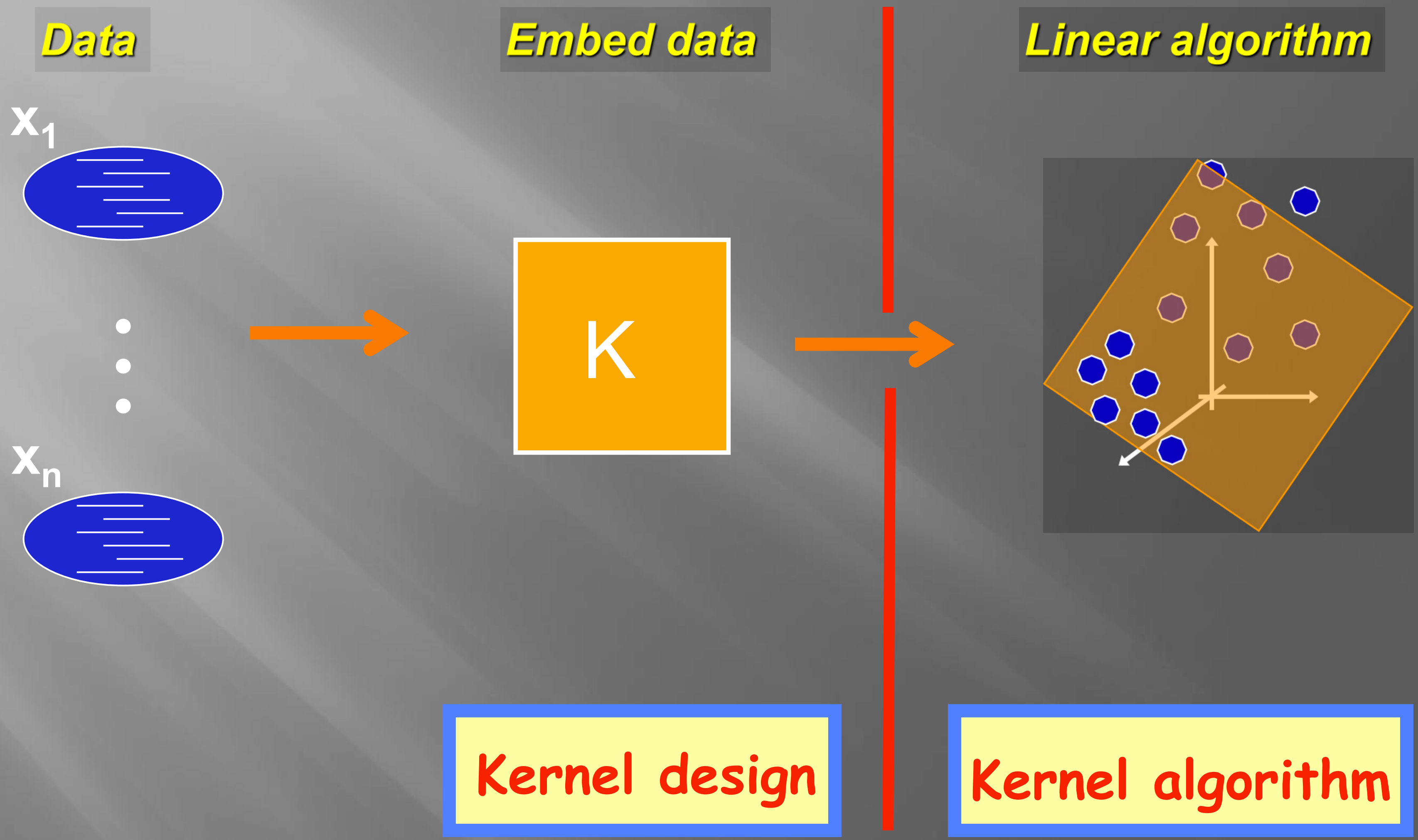
*Data* | *Embed data* | *Linear algorithm*




SVM, MPM, PCA, CCA, FDA…

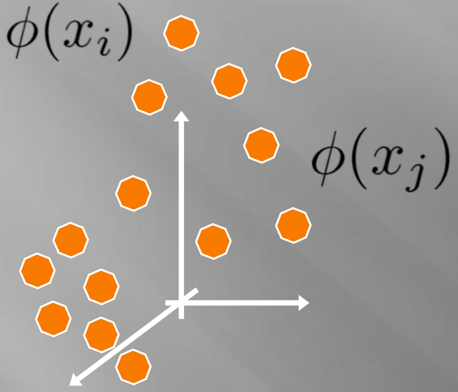
- **if the data is described by numerical vectors: embedding ~ (non-linear) transformation**

# Kernel methods



# Kernel methods

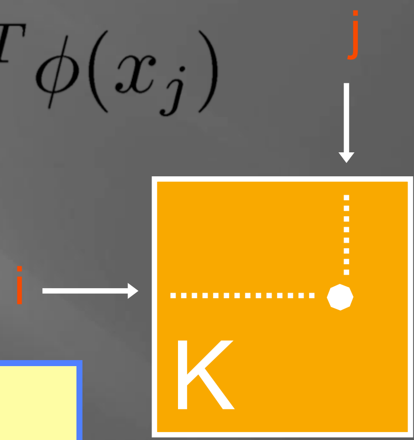
**Embed data**




***IMPLICITLY: Inner product measures similarity***

$$k(x_i, x_j) = \phi(x_i)^T \phi(x_j)$$

$$K_{ij} = \phi(x_i)^T \phi(x_j)$$



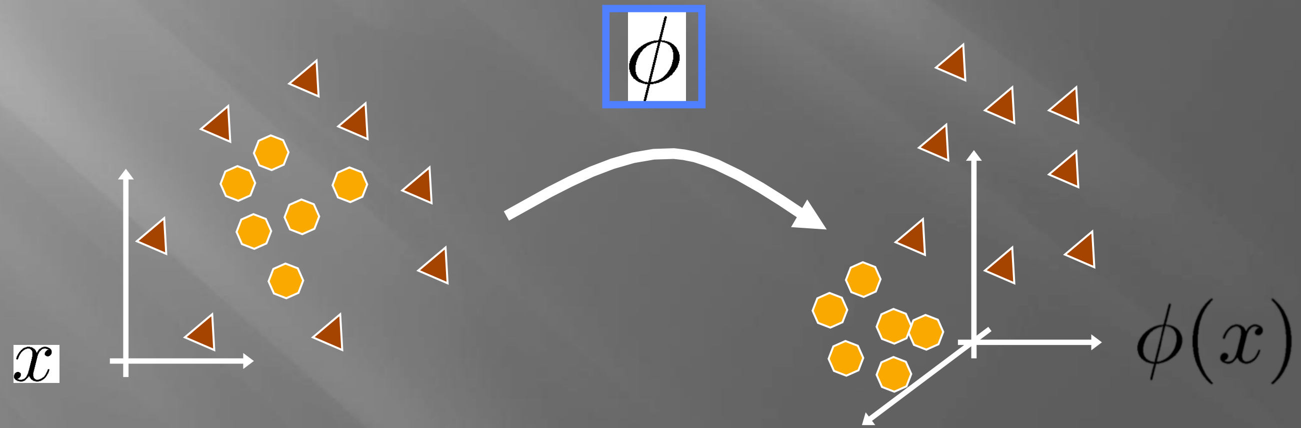

**Property:** Any **symmetric positive definite** matrix specifies a kernel matrix & every kernel matrix is **symmetric positive definite**

# Kernel methods

- Similarity measurement for vector data:
  Gaussian kernel

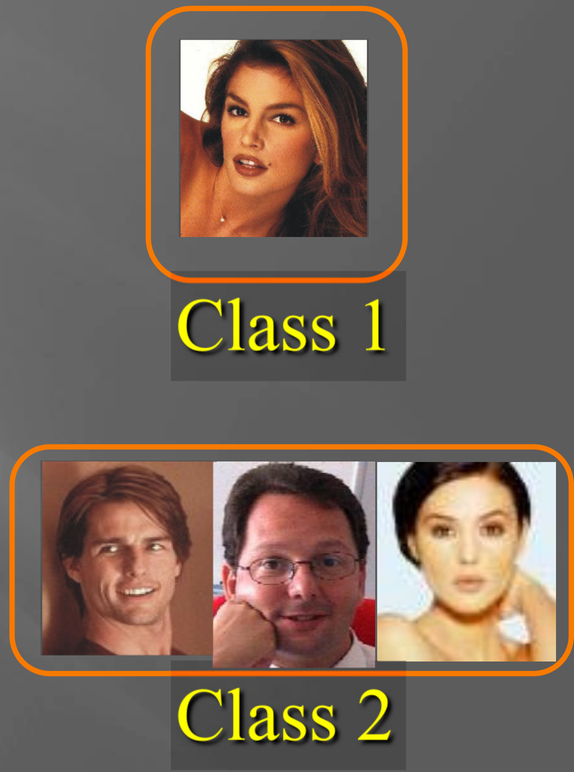
$$k(x,z) = \exp\left(\frac{-\|x-z\|_2^2}{2\sigma}\right)$$

- Corresponds to highly non-linear embedding



# Support Vector Machines

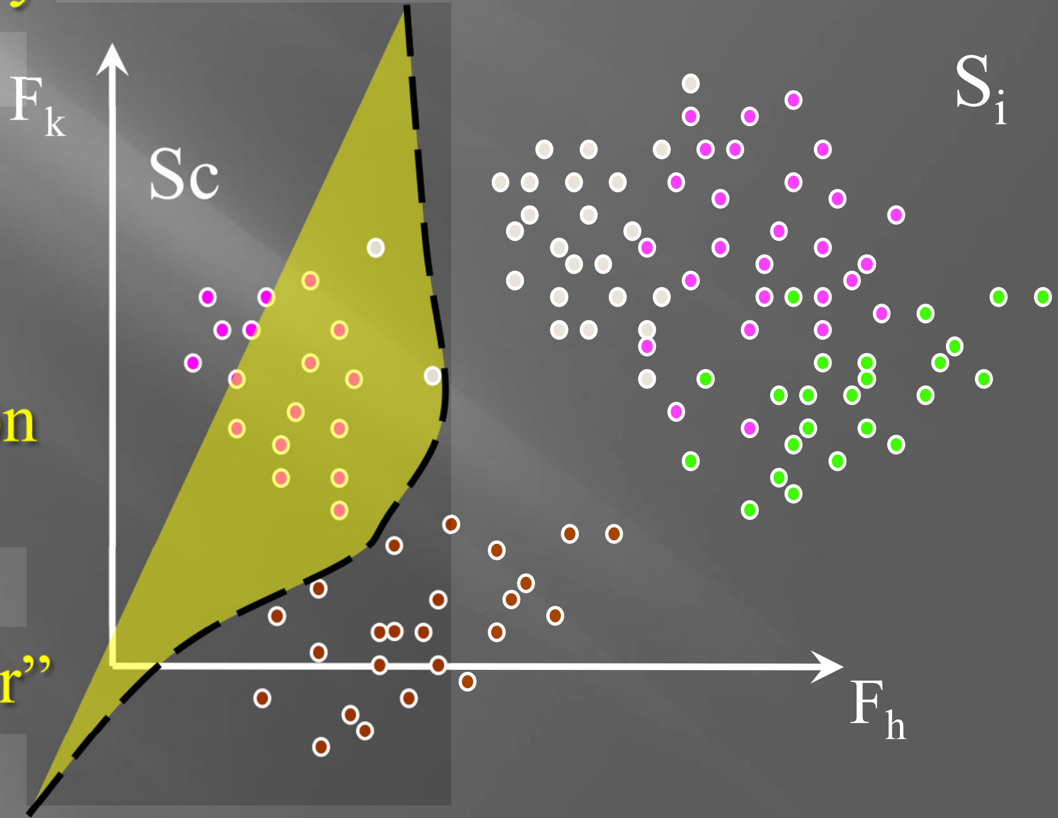
## Support Vector Machines are <u>binary</u> classifiers




**V. Vapnik, S.E. Golowich, A.J. Smola**: Support Vector Method for Function Approximation, Regression Estimation and Signal Processing. Neural Information Processing Systems 1996: 281-287

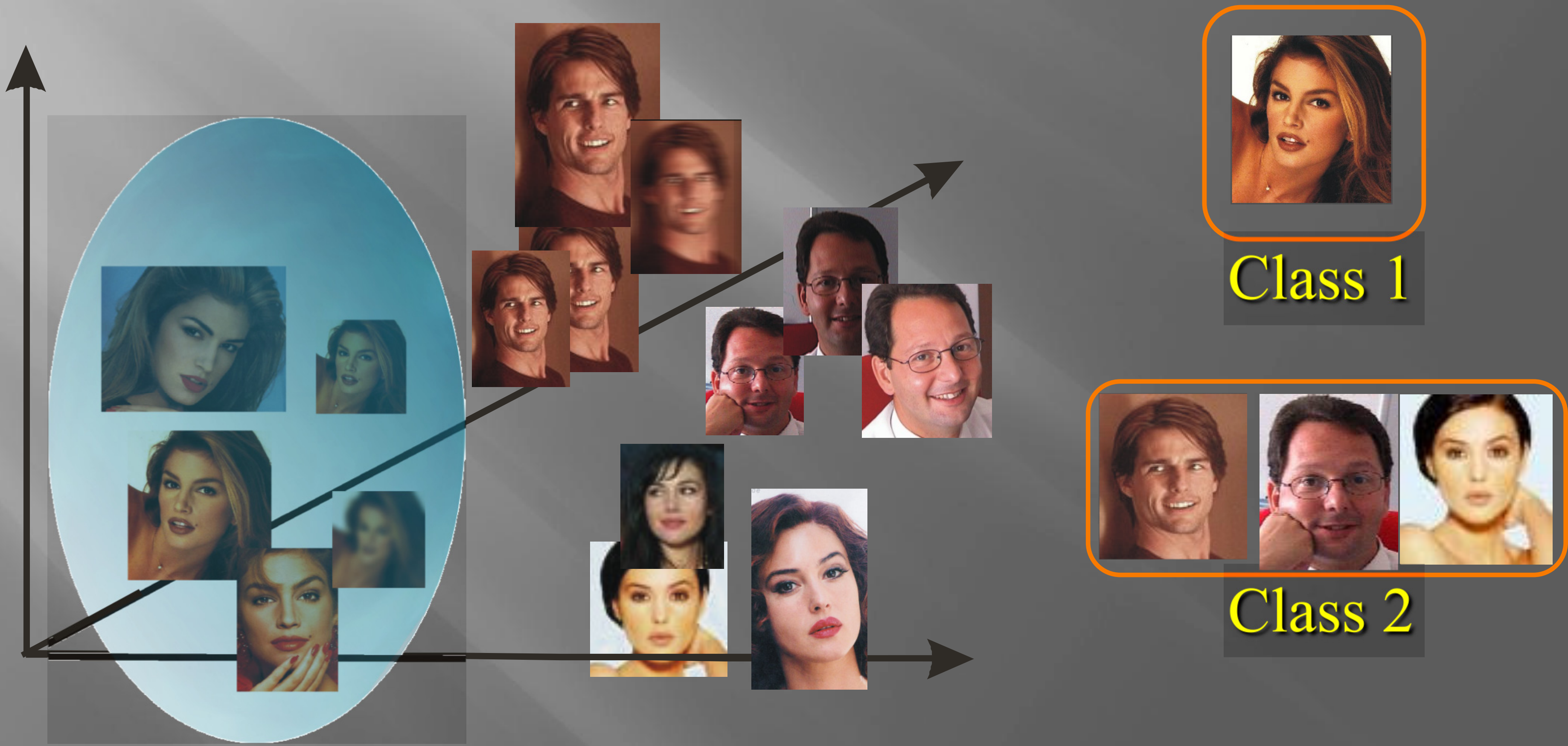
# Support Vector Machines

- The separating hyperplane can be very complex
- Problems may occur with outliers
- The shape of the hyperplane depends on the population of the classes
- Need for an "impostor" class



# One-Class Support Vector Machines

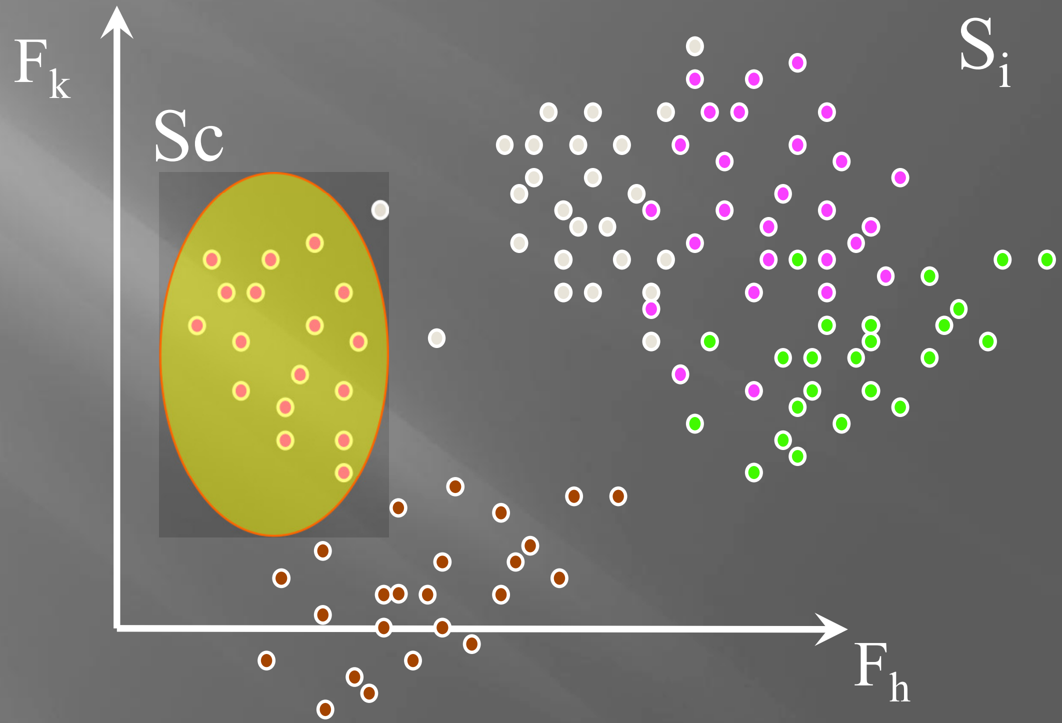
## UNARY classifiers




**Ben-Hur, A., Horn, D., Siegelmann, H., , Vapnik, V.:** « Support vector clustering ». Journal of Machine Learning Research **2** (2001) 125–137

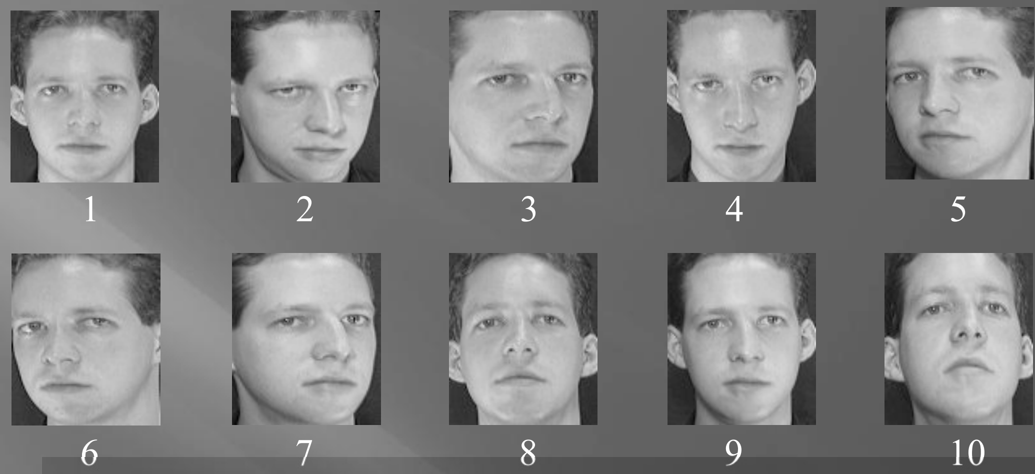
# One-Class Support Vector Machines

- The separating surface is a hyperspehere
- Selectivity can be adjusted by two parameters
- No need for direct “impostor” training




$$\|x_i - a\|^2 \leq R^2 \qquad i = 1 \ldots \ell$$

# One-Class Support Vector Machines




ORL Database 10 frames/sbj

Our video database
95:195 frames/sbj

# One-Class Support Vector Machines

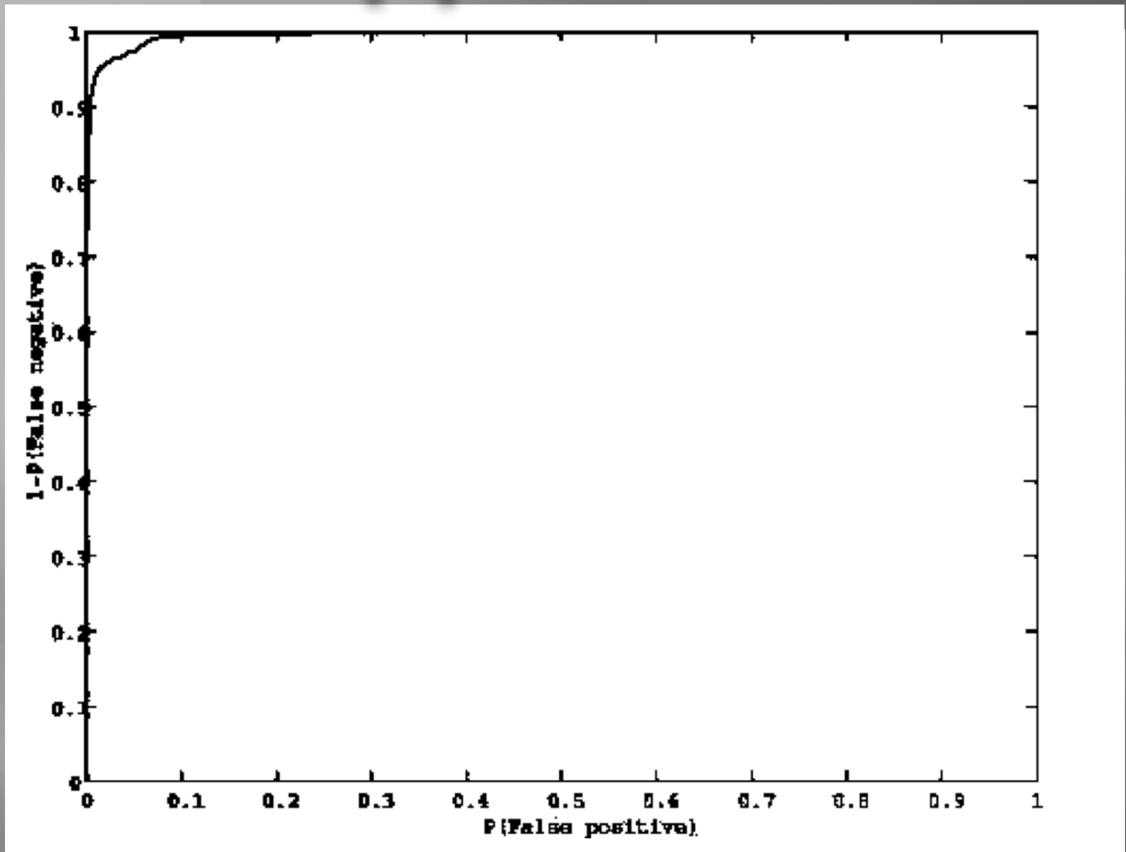
Poor representation (10 frames/sbj)

| Method | Impostor set | Equal Error Rate (EER) |
|---|---|---|
| **Binary SVM** | Bern | 6.00 % |
| | Stirling | 5.50 % |
| | Yale | 6.00 % |
| | *Mixeddata set* | 6.19 % |
| **One-Class SVM** | **NONE** | **5.50 %** |

Rich representation (100 frames/sbj)

| Threshold | Client tests | Impostor tests | Equal Error Rate (EER) |
|---|---|---|---|
| Single threshold | 2275 | 122520 | 3.38 % |
| One threshold per subject | | | 0.56 % |

# One-Class Support Vector Machines




**M. Bicego, E. Grosso, M. Tistarelli**: « Face Authentication Using One-Class Support Vector Machines », S.Z. Li et al. (Eds.): IWBRS 2005, LNCS 3781, pp. 15–22, 2005.

**ROC from our video database (25 sbj.s)**
**EER 0.56 %**

# Hidden Markov Models

Vision Lab

- **Largely applied for classification purposes, in several different contexts:**
  - speech recognition (Rabiner 89)
  - hand-written character recognition (Hu et al. 96)
  - DNA and protein modelling (Hughey et al. 96)
  - gesture recognition (Eickeler et al. 98)
  - complex action classification (Brand et al. 97)
  - behaviour analysis and synthesis (Jebara et al. 99)
  - face recognition (Kohir et al. 98, Eickeler et al 2000)

# Hidden Markov Models

Vision Lab

- **Modelling *sequences of symbols***
  - capture the underlying structure of a set of symbol strings
  - A stochastic generalisation of finite-state automata, governed by probabilistic distributions
- **Markovian Models where states are not directly observable**
  - a probability function is associated to each state: the probability that a symbol is emitted from that state.

This representation can naturally encode time-varying sequences of image regions, objects or … faces

# Hidden Markov Models

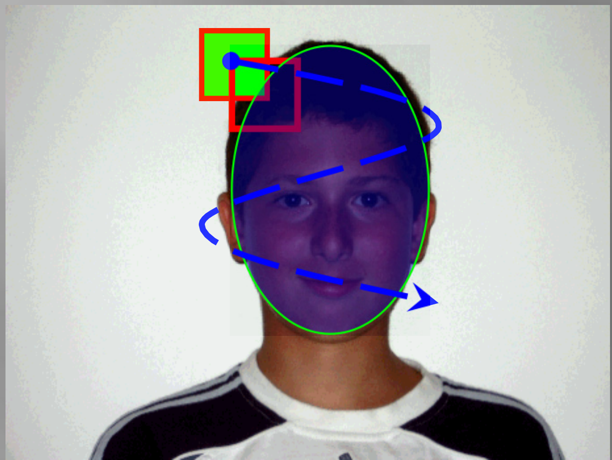
Vision Lab

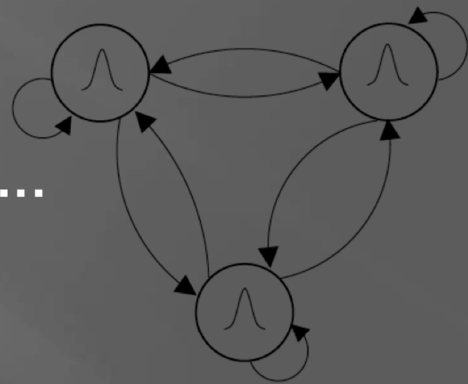
- ***Coding***
  - the image is scanned to obtain a sequence of *T* partially overlapped sub-images
  - for each sub-image the DCT coefficients are computed
  - only the most important *D* coefficients are retained
  - the final sequence is composed by *DxT* symbols
- ***Learning*** **: train one HMM for each subject:**
  - the number of states is fixed *a priori*
  - at the end we have one HMM for each subject
- ***Recognition*** **:**
  - Bayesian scheme: assuming a priori equi probable classes, an object is assigned to the class whose model shows the highest likelihood (Maximum Likelihood scheme)

# Hidden Markov Models

Vision Lab

## Statistical analysis of sequences of patterns



# Hidden Markov Models

## State-of-the-art in face classification (2003)

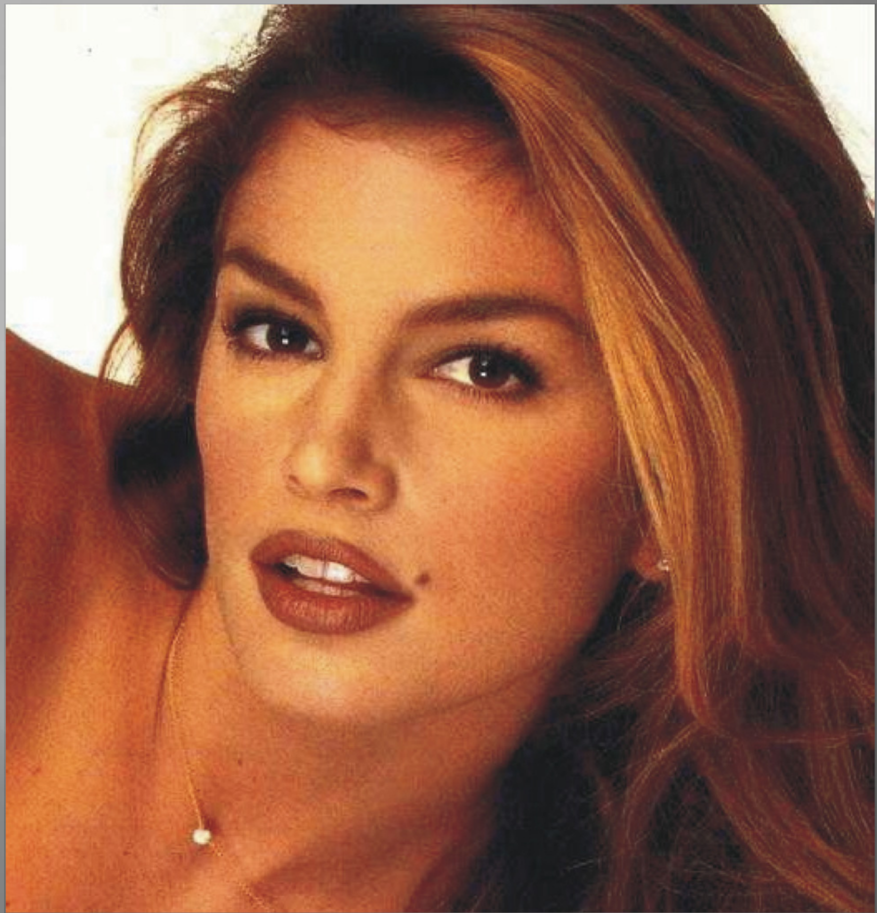
| Classifier | Year | Accuracy |
|---|---|---|
| Top-down HMM + gray tone feat. | 1994 | 87% |
| Eigenface | 1994 | 90.5% |
| Pseudo 2D HMM + gray tone feat. | 1994 | 94.5% |
| Elastic matching | 1997 | 80.0% |
| PDNN | 1997 | 96.0% |
| Continuous n-tuple classifier | 1997 | 97.3% |
| Top-down HMM + DCT coef. | 1998 | 84% |
| Point-matching and correlation | 1998 | 84% |
| Ergodic HMM + DCT coef. | 1998 | 99.5% |
| Pseudo 2D HMM + DCT coef. | 2000 | 100% |
| SVM + PCA coef. | 2001 | 97% |
| Indipendent Component Analysis | 2002 | 85% |
| Gabor filters + rank correlation | 2002 | 91.5% |
| Wavelet + HMM | 2003 | 100% |

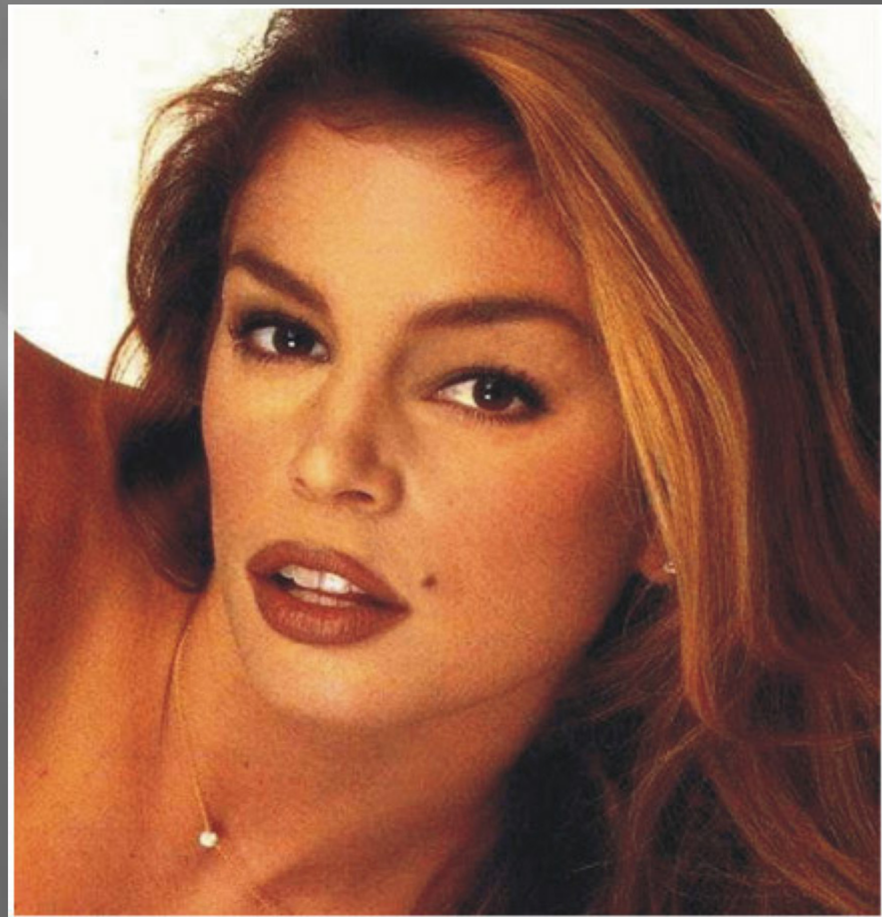
# Automatic face recognition

**Face recognition involves several sequential processes:**

- **Face detection**
- **Normalization**
- **Facial features extraction**
- **Registration/ Representation**
- **Classification**

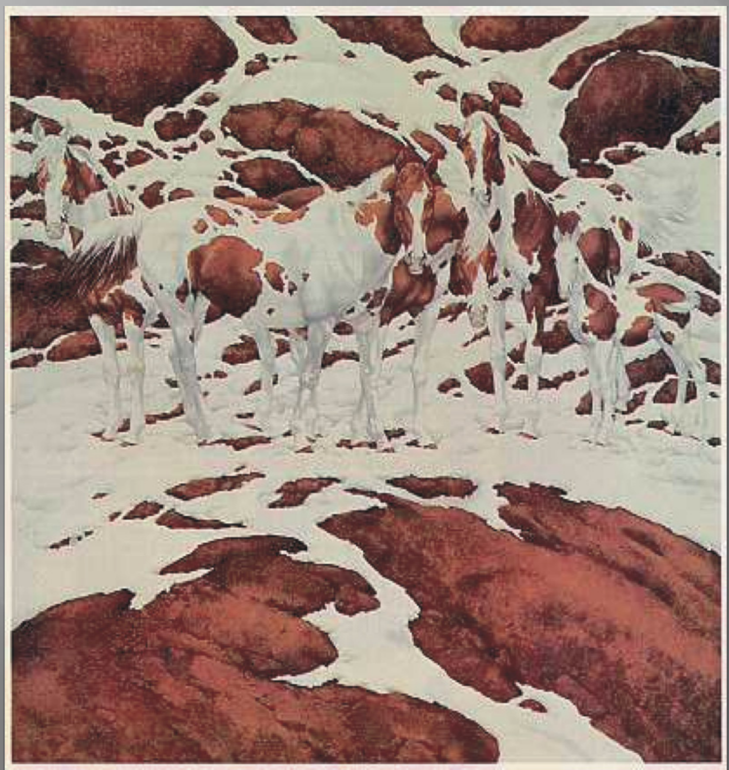
# Face alignment



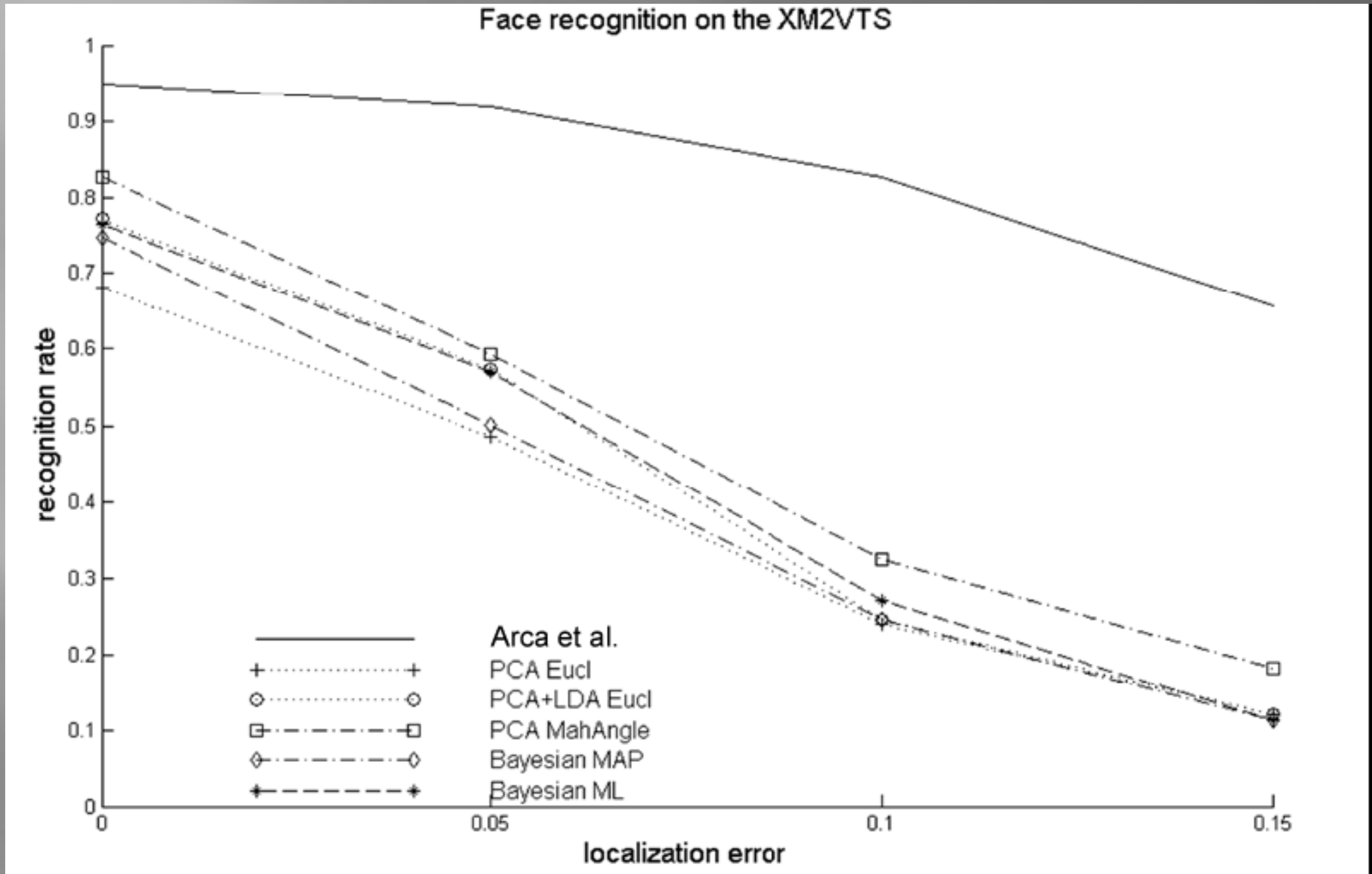


# Image registration

In cognitive psychology it is called "perceptual organization"



# Face alignment



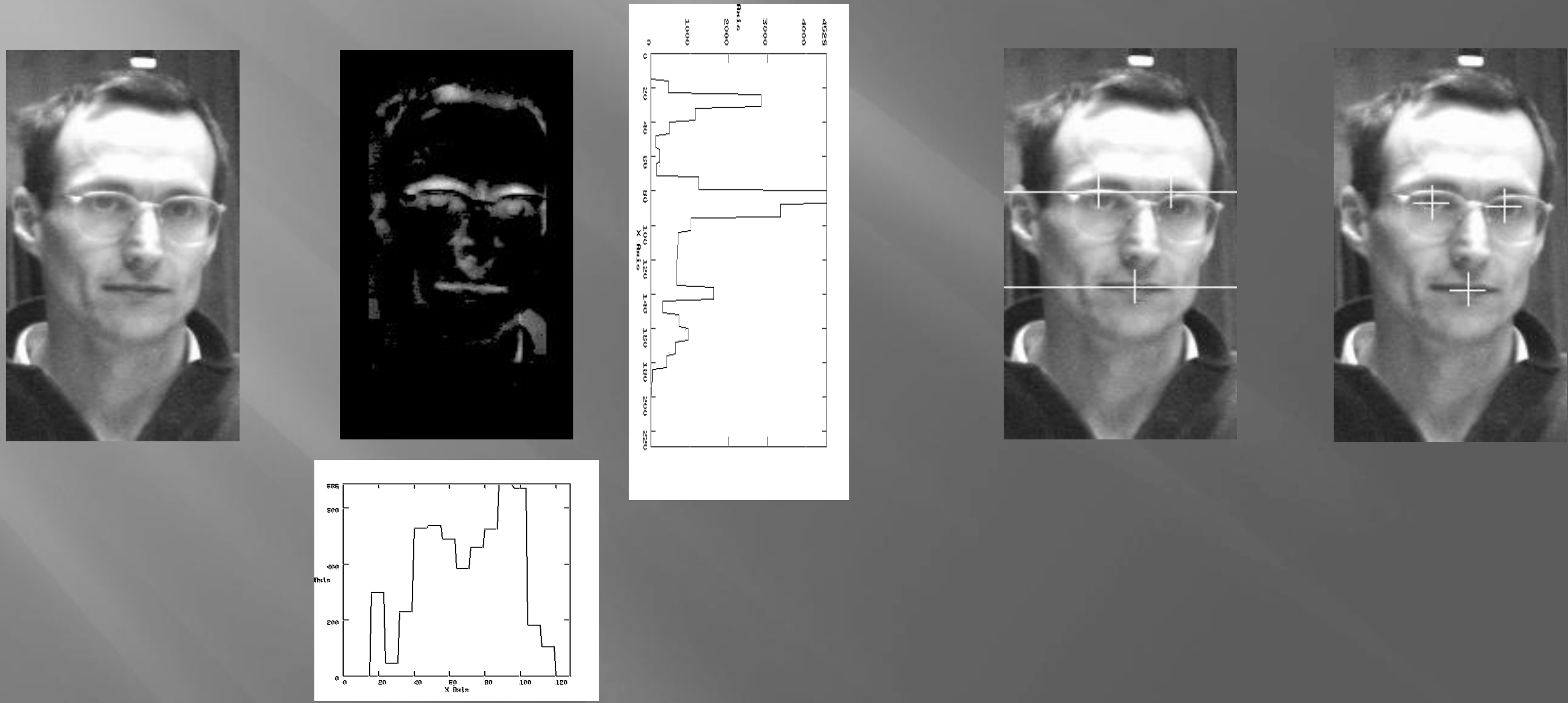

**S. Arca, P. Campadelli, and R. Lanzarotti**. A face recognition system based on automatically determined facial fiducial points. *Pattern Recognition*, 39(3):432–443, 2006.

# Facial landmarks

Vision Lab

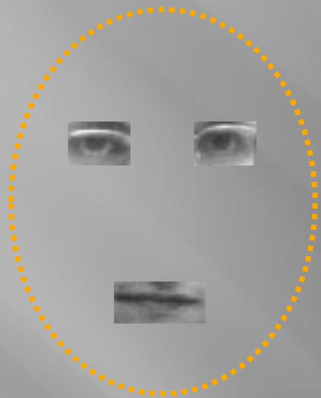
## A two step algorithm :

- Morphological filtering and matching

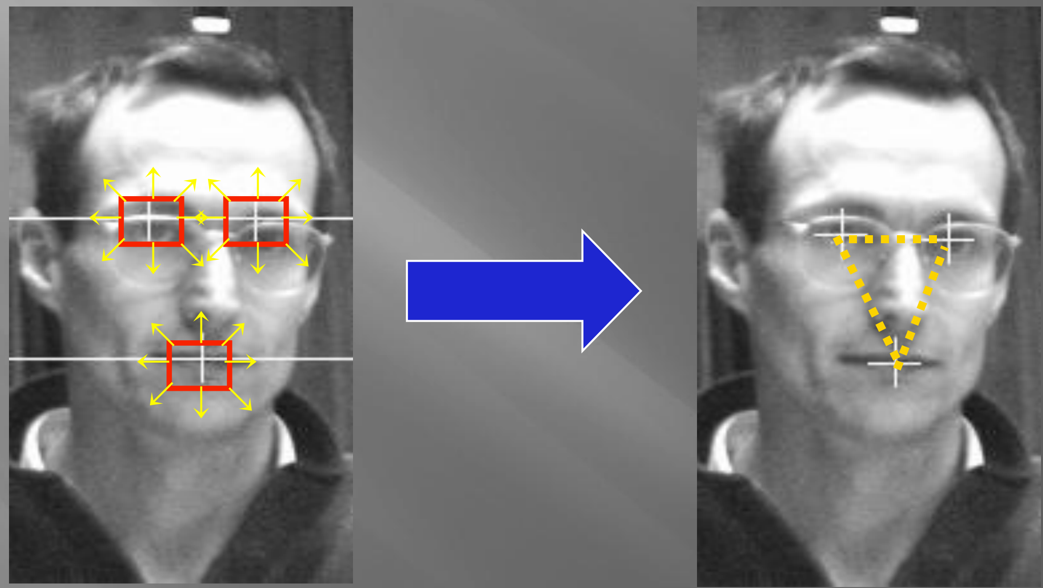


# Facial landmarks





Eyes and Mouth templates

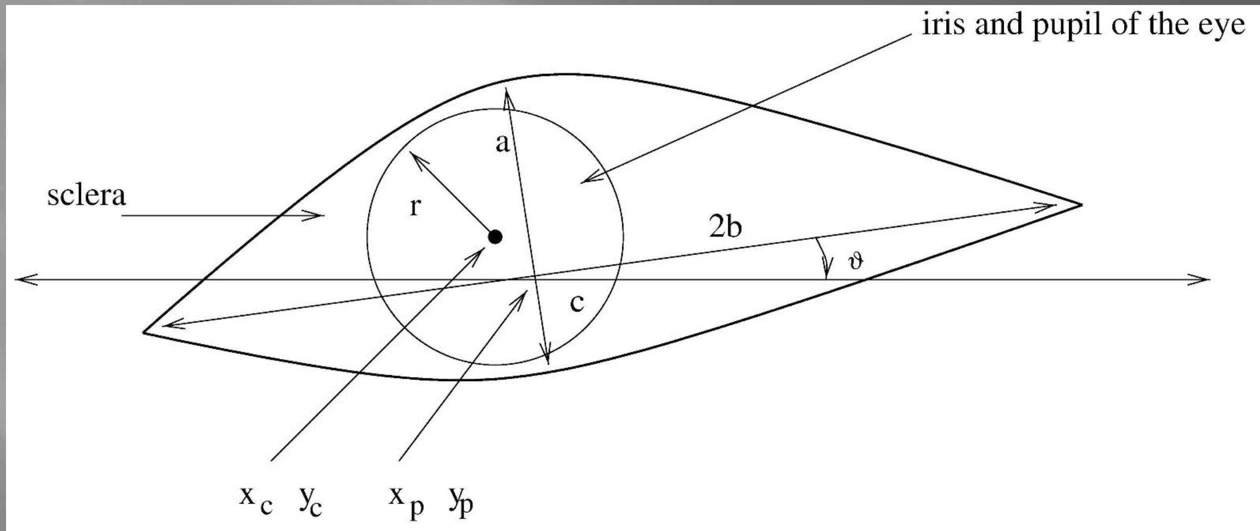


**Validation by correlation score and geometric constraint**

# Facial landmarks

Vision Lab

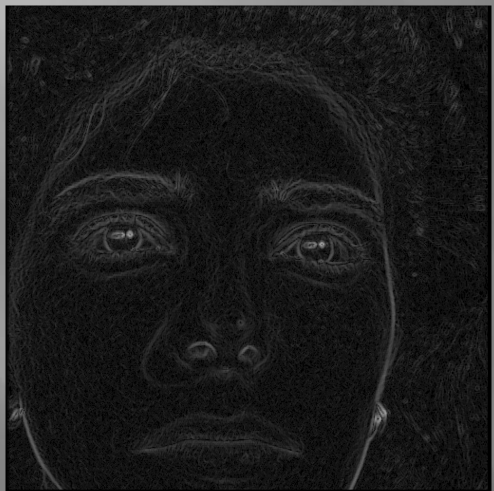
More accurate feature detection requires
***adaptive templates***



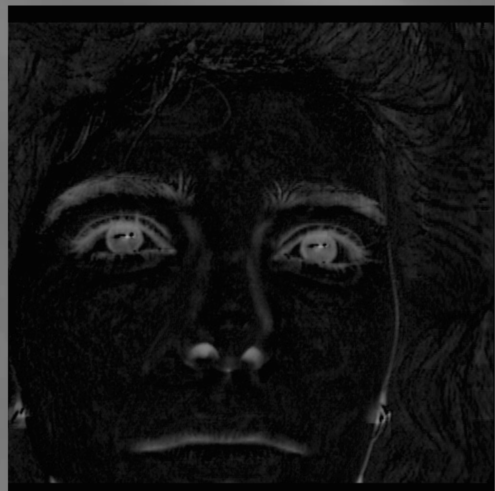

**A.L.Yuille, D. Cohen, and P.Hallinan**, “Feature extraction from faces using deformable templates”, IJCV 92.

# Facial landmarks

**Morphological filtering**



gray level CONTOURS



gray level VALLEYS

gray level PEAKS

# Facial landmarks

# Automatic face recognition

**Face recognition involves several sequential processes:**

1. Face detection
2. Normalization
3. Facial features extraction
4. Registration/ Representation
5. Classification

# Classification and learning

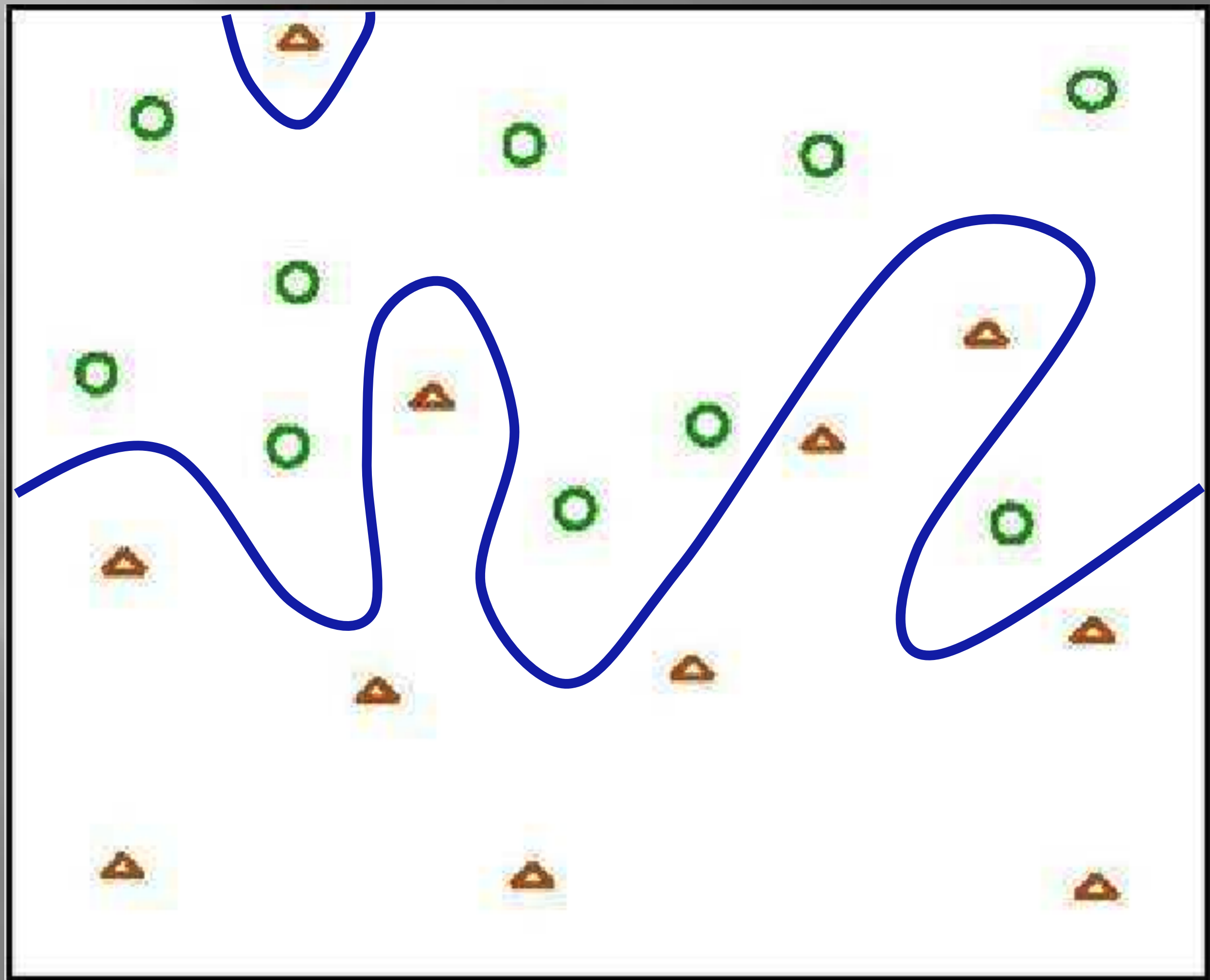
**Classifiers:**

- Bayesian
- SVM / Kernel
- Neural Networks
- HMMs
- GMMs
- …

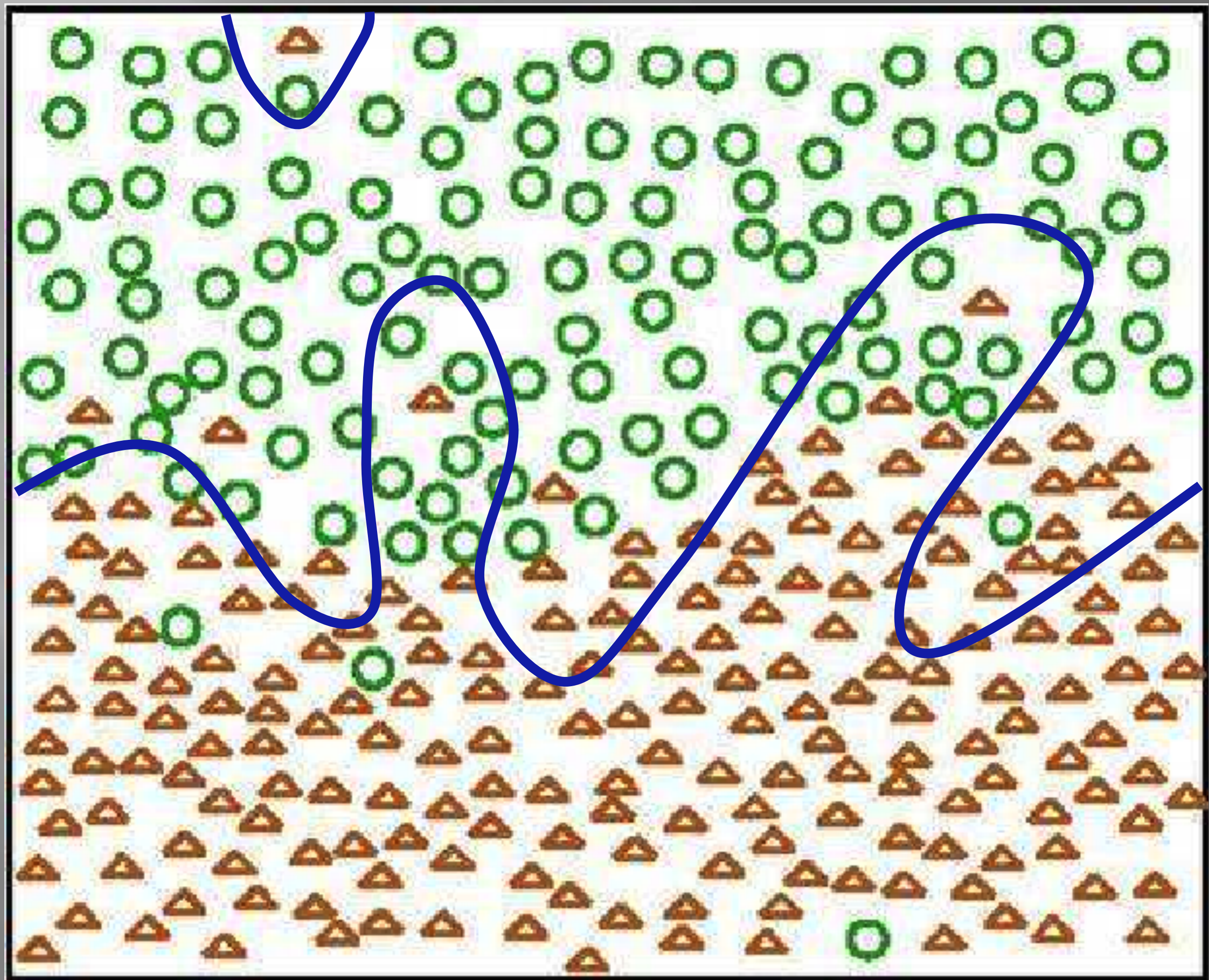
**Features:**

- Gabor
- PCA/LDA/ICA
- LBP
- SIFT
- Haar
- DCT
- Edges/Regions
- Raw data
- …

# Classification and learning



# Classification and learning

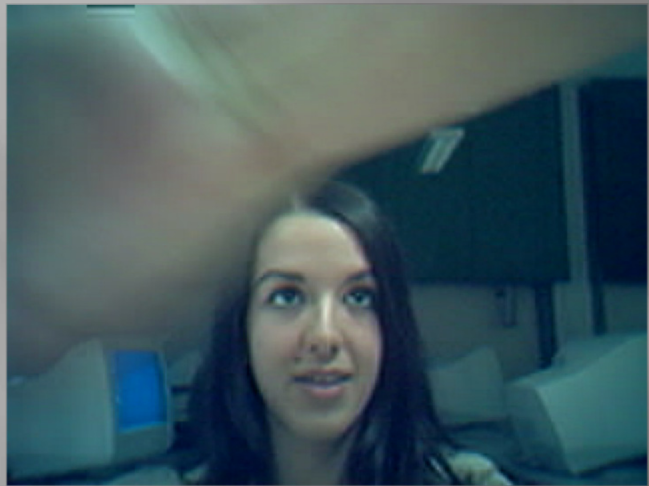


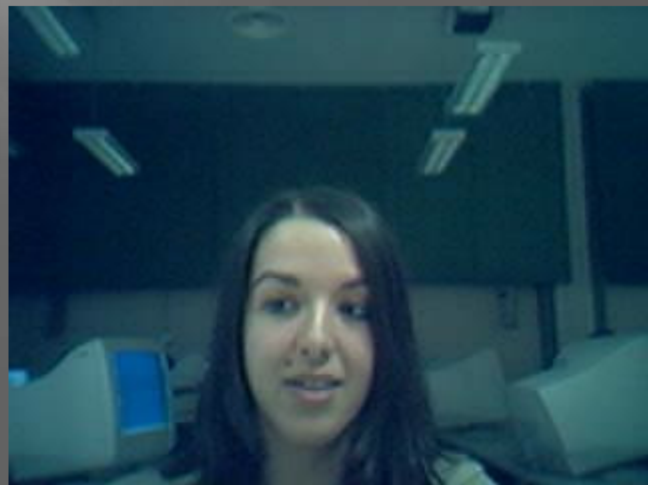
# FACE RECOGNITION FROM VIDEO

# Face recognition from video

**Dynamics in a video stream conveys far more information than a collection of single snapshots**

# Face recognition from video





**Advantages from video:**

- **More data available**
- **Temporal integration**
- **Behavioral cues**
- **Spatial and temporal sampling**

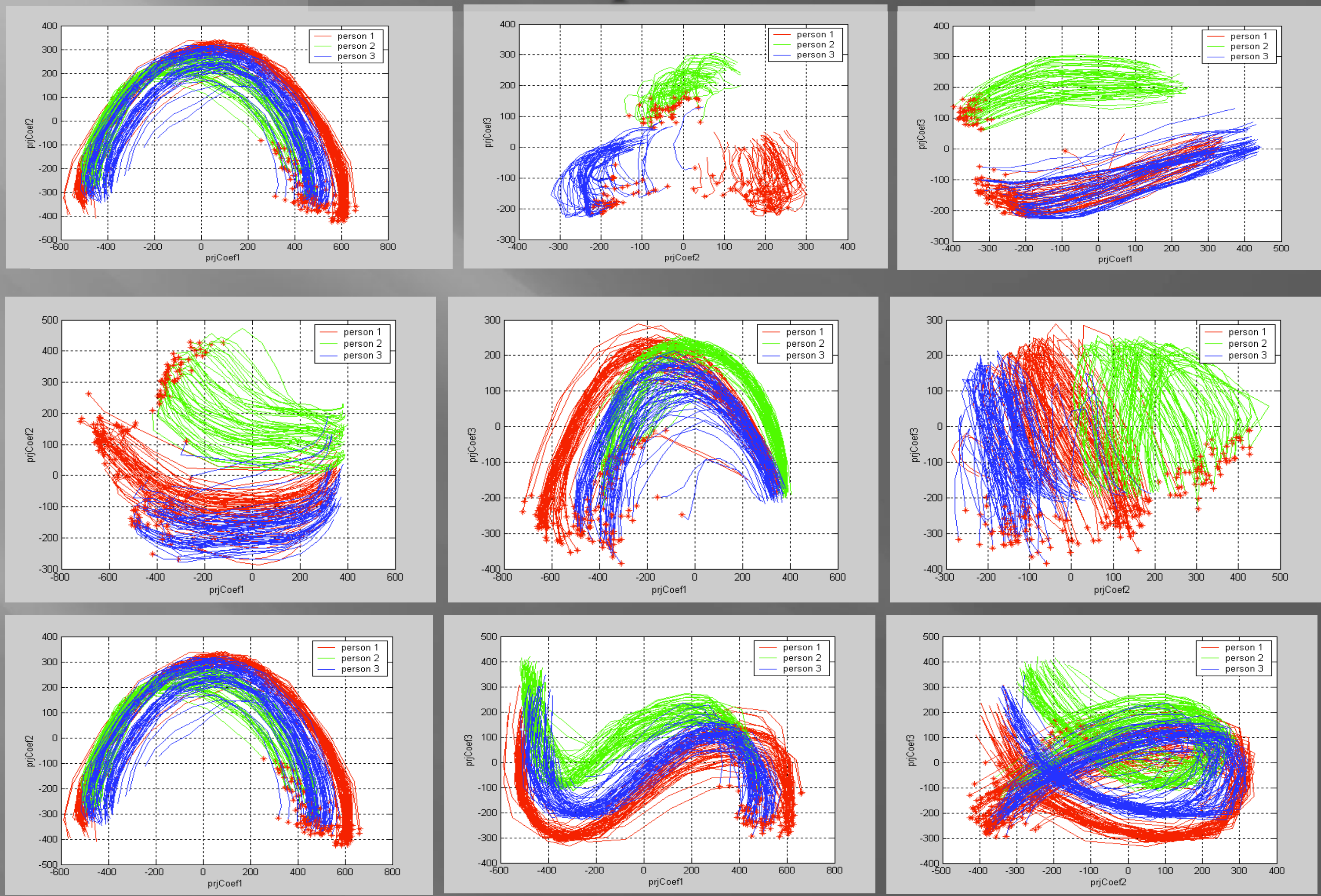
# Effects of motion

## Face sub-spaces manifold

Face rotation

Face scaling

Face translation



# Face recognition from video

- The curse of dimensionality…


Standard VGA (640x480)

1 frame: 300 KByte
30 frames: 1 MByte

Standard video: ~1 MByte/Sec

- …the risk is to have too much data to be processed

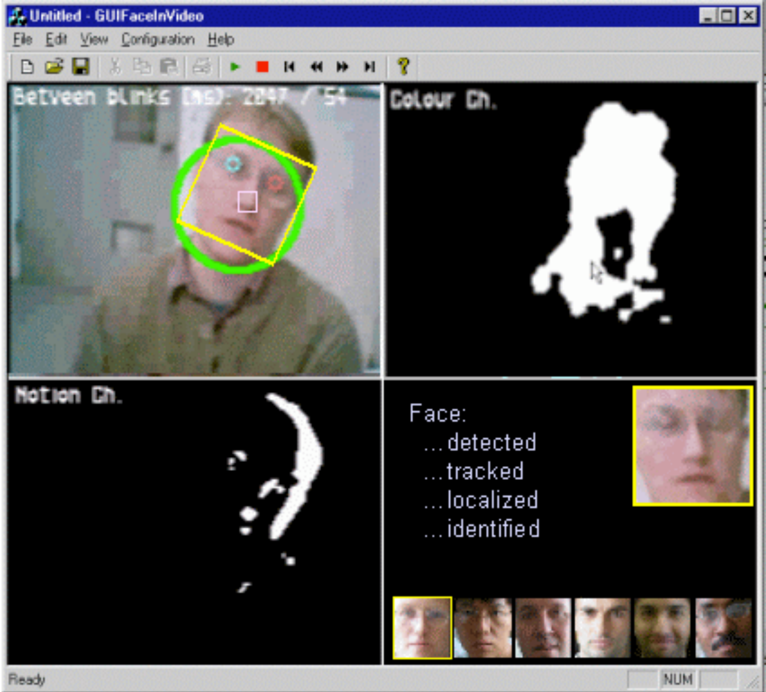
How to exploit the added information in video?

# Face recognition from video

**Not just *more data* to be processed:**

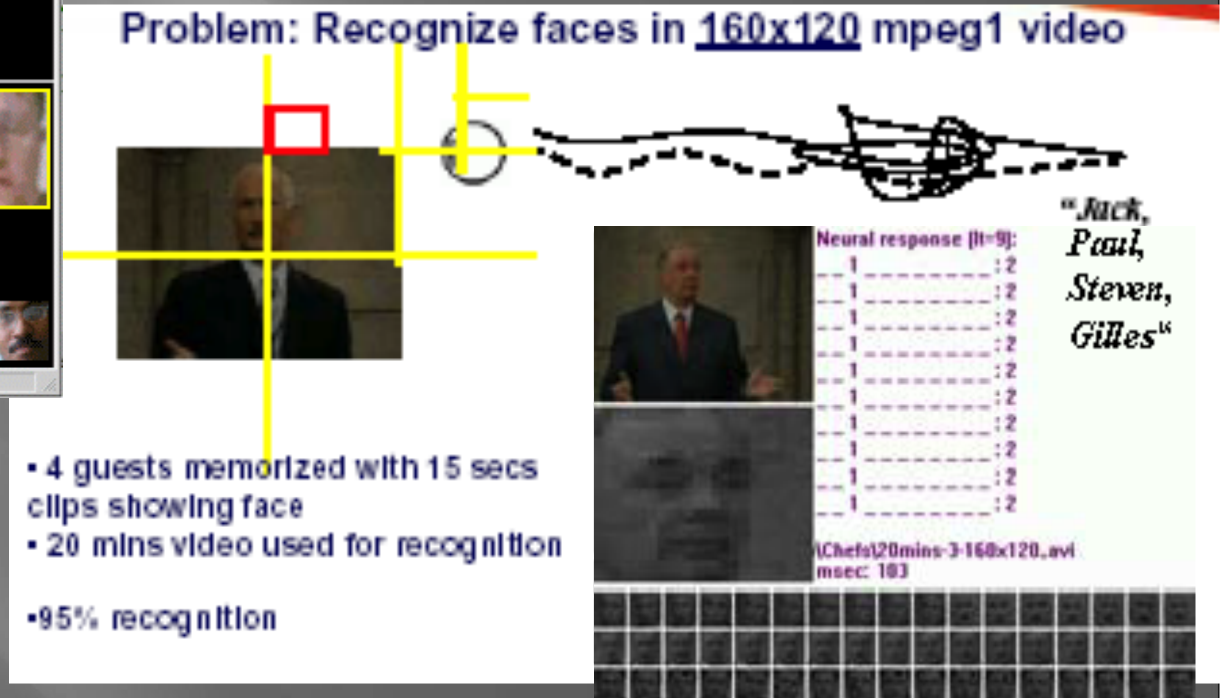
- Data selection (pose, expression, illumination, noise…)
- Multi-data fusion (decision/score/feature level)
- 3D reconstruction/virtual views
- Resolution enhancement
- Expression and emotion analysis
- Behavioral analysis
- *Dynamic* video templates…?

# Face recognition from video



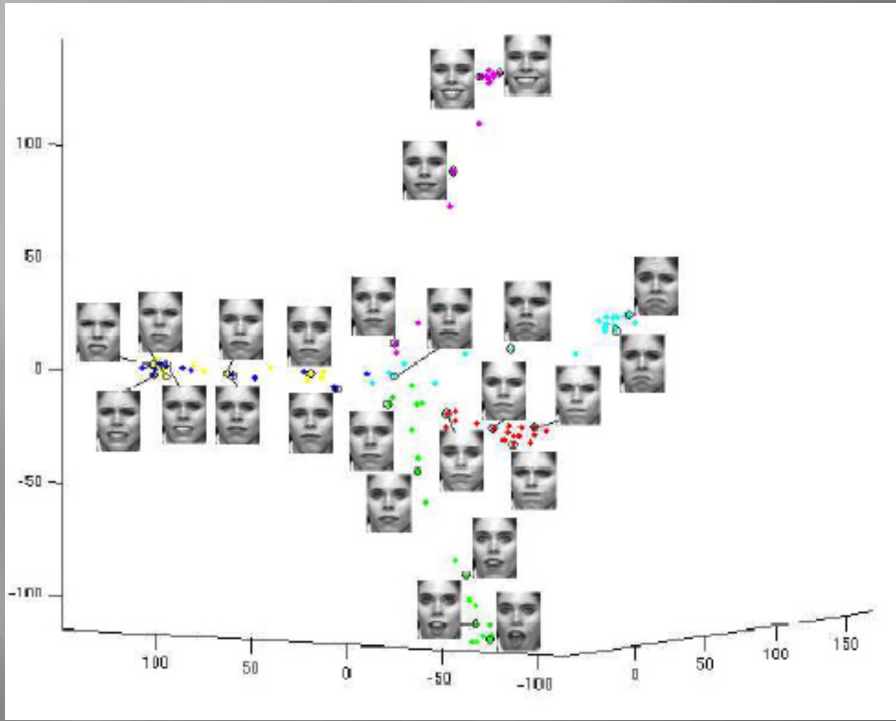

## Neural networks-based model

- **D.O. Gorodnichy**, Institute for Information Technology, National Research Council of Canada



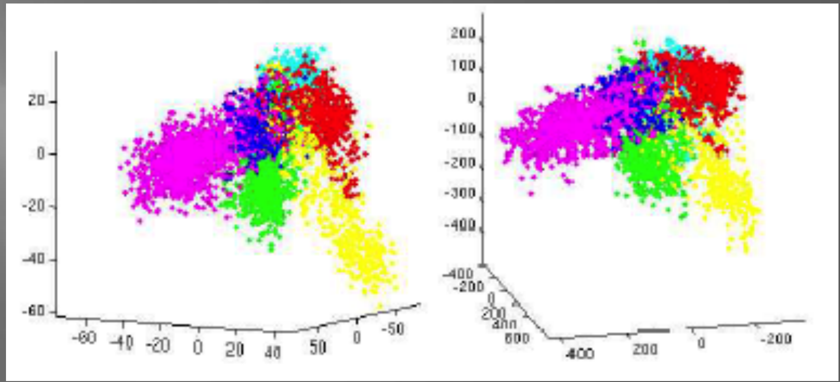

**Goal: Recognition in very low resolution video**

# Face recognition from video



## Multi-dimensional PCA

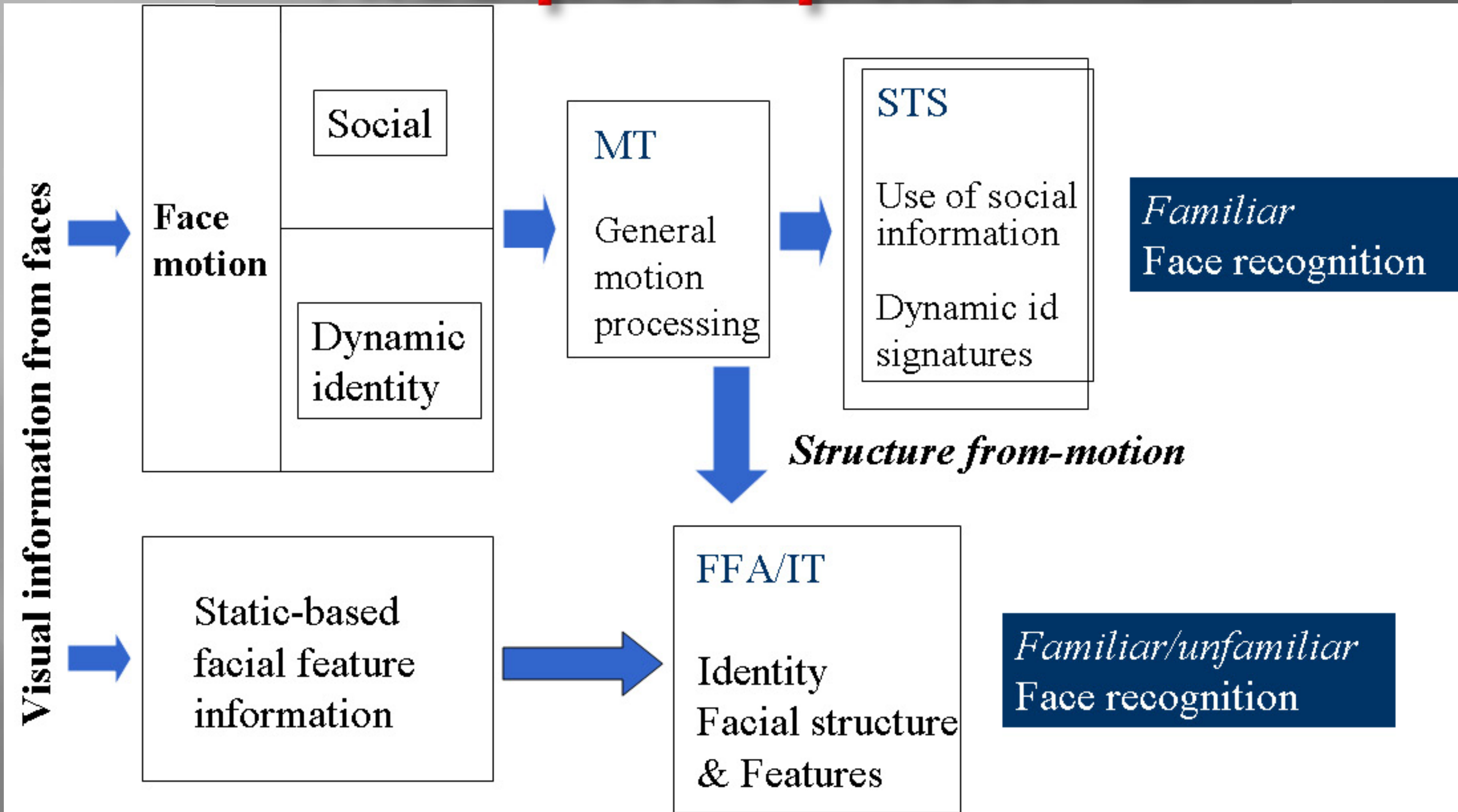
- **Kruger, Chellappa, Kriegman** 2003
- **Shan, Ging, Mc Owan** 2004



**Goal**:

- Recognition from face sequence manifolds
- Facial expression recognition

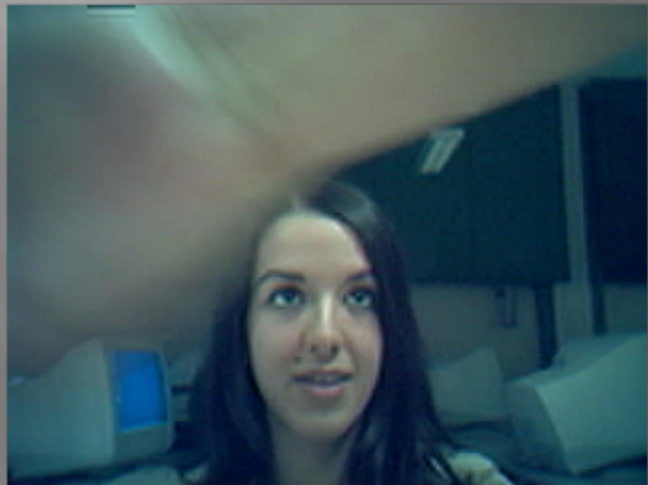
# Neural architecture of face perception - 2

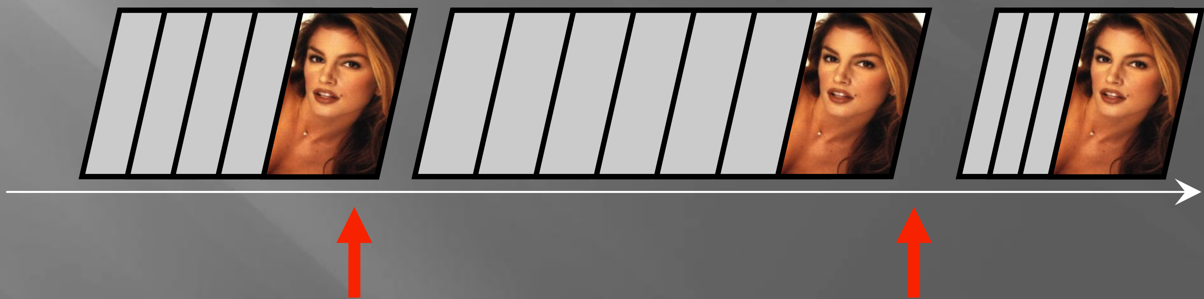



**A. O'Toole**, "Biological Face Recognition", Summer School on Biometrics 2006.

# Face recognition from video

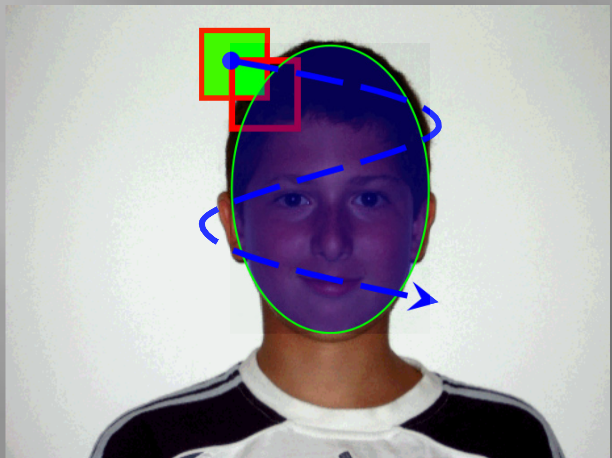
**How to represent photometric and dynamic information at the same time?**

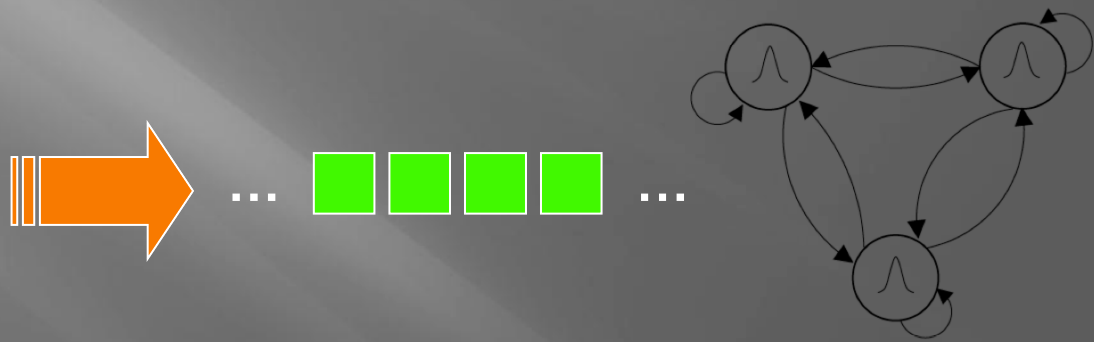




# Hidden Markov Models

Statistical analysis of sequences of patterns





- This idea can be extended to multi-dimensional patterns and sequences … *in several ways*

# Dynamic Hidden Markov Models

Standard HMMs can be extended to multi-dimensional patterns and sequences … *in many ways*

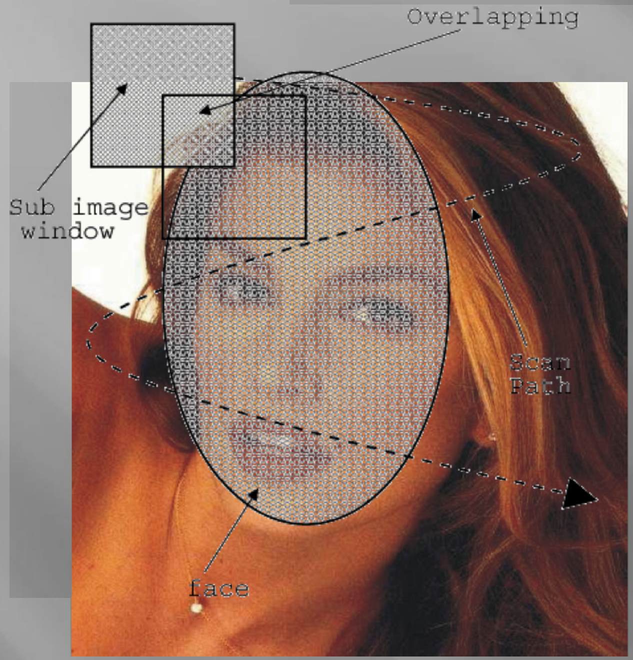
1. Each image is modeled as a single HMM and the sequence of images as a sequence of HMMs
   (A. Hadid and M. Pietikainen. "An experimental investigation about the integration of facial dynamics in video-based face recognition". *Electronic Letters on Computer Vision and Image Analysis*, 5(1):1-13, 2005.)

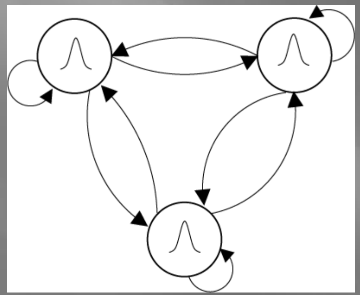
2. The entire video is modeled as a single HMM
   (X. Liu and T. Chen. "Video-based face recognition using adaptive hidden Markov models". In *Proc. Int. Conf. on Computer Vision and Pattern Recognition*, 2003.)

3. The images and the sequence itself are modeled as a complex, hierarchical HMM-based structure
   (M. Bicego, E.Grosso, M. Tistarelli. "Person authentication from video of faces: a behavioral and physiological approach using Pseudo Hierarchical Hidden Markov Models", Int.l Conference on Biometric Authentication 2006, Hong Kong, China, January 2006. )

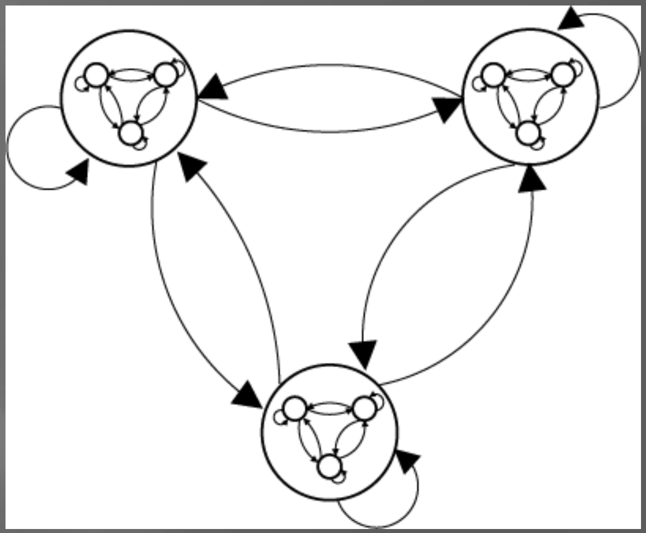
# Pseudo-Hyerarchical HMM

## Physiological and behavioral features



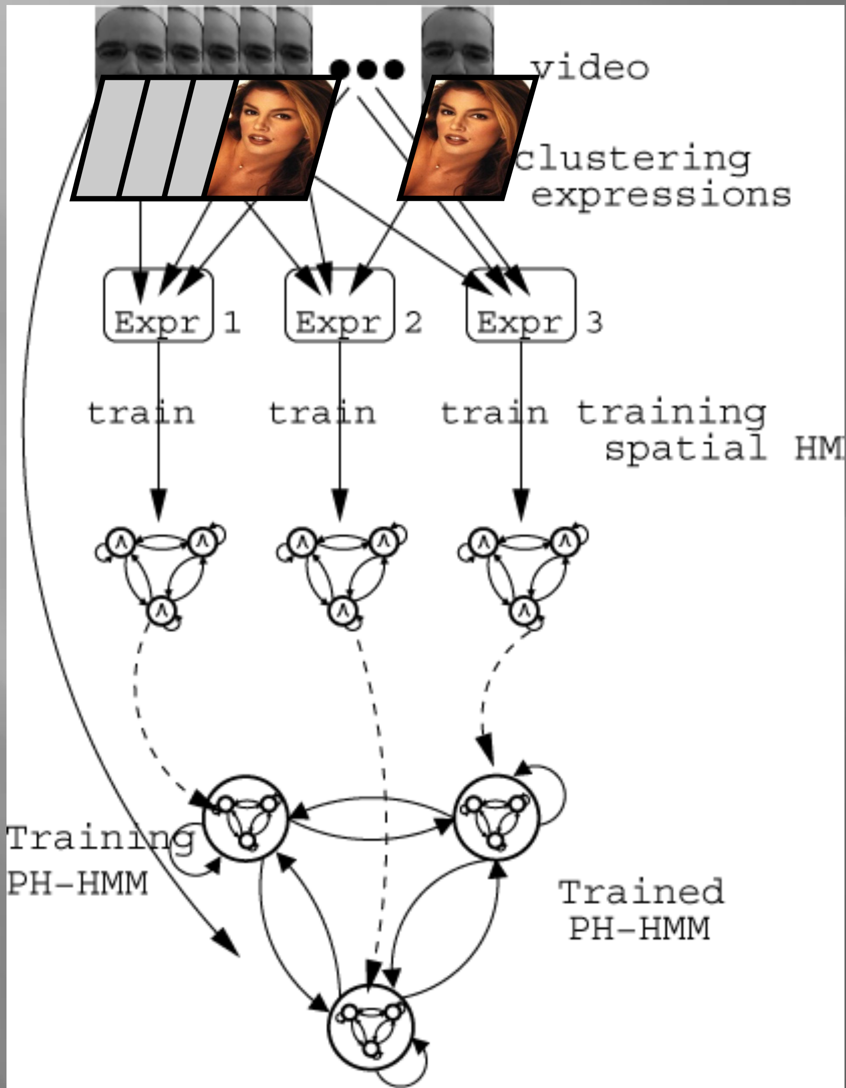



HMM



PH-HMM

# Pseudo-Hyerarchical HMM

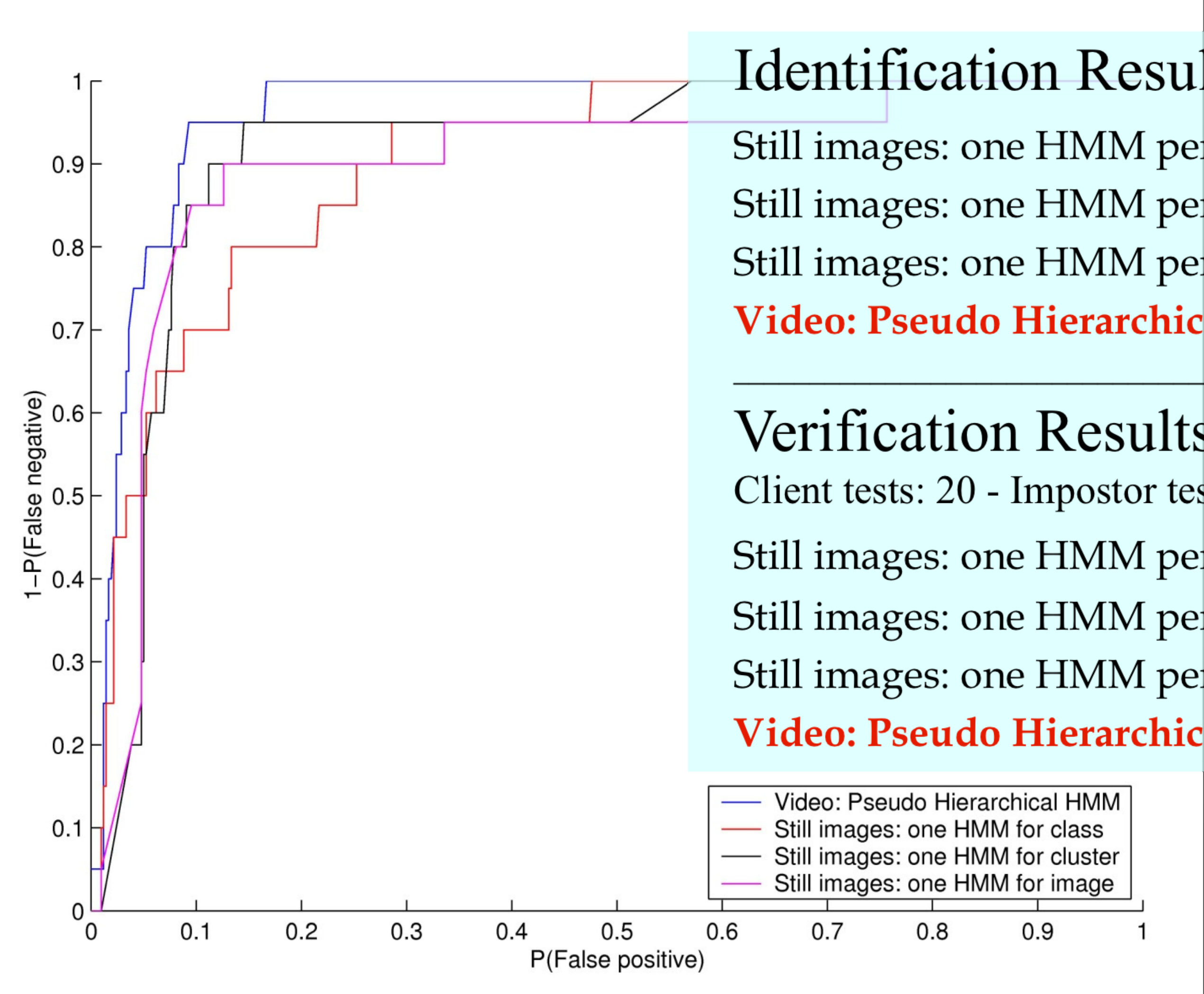



## Pseudo-Hierarchical Hidden Markov Model

The image stream is represented as a collection of sub-streams, each corresponding to a different facial expression or motion.

The change in expression is capturede by the transition matrix of the PH-HMM

# HMM Classification from video




**Identification Results** (recognition rate)

| | |
|---|---|
| Still images: one HMM per class | 11/21 (52.38%) |
| Still images: one HMM per cluster | 14/21 (66.67%) |
| Still images: one HMM per image | 12/21 (57.14%) |
| **Video: Pseudo Hierarchical HMM** | **21/21 (100%)** |

**Verification Results** (EER)

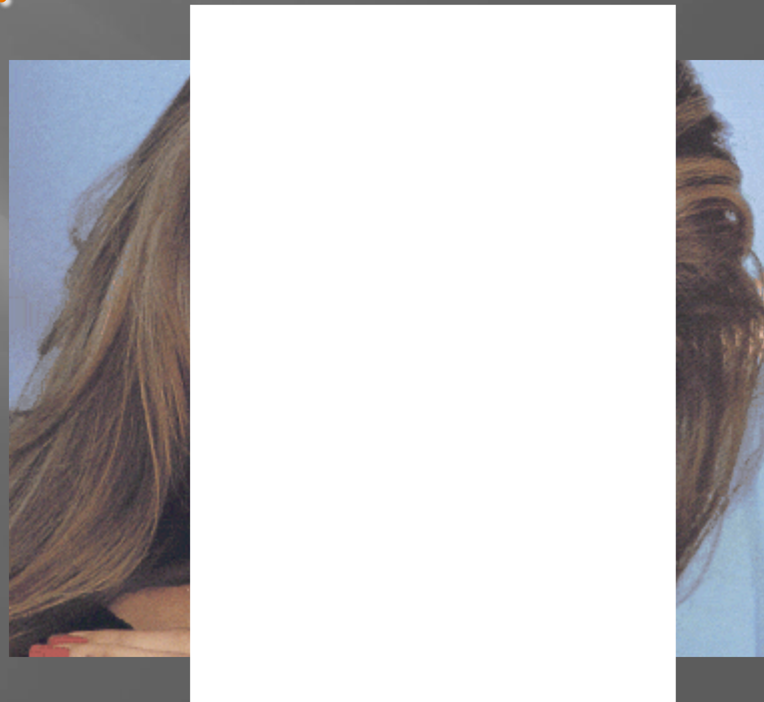
Client tests: 20 - Impostor tests: 420

| | |
|---|---|
| Still images: one HMM per class | 20.24% |
| Still images: one HMM per cluster | 10.60% |
| Still images: one HMM per image | 13.81% |
| **Video: Pseudo Hierarchical HMM** | **6.07 %** |

# SUBJECT-DEPENDENT RECOGNITION

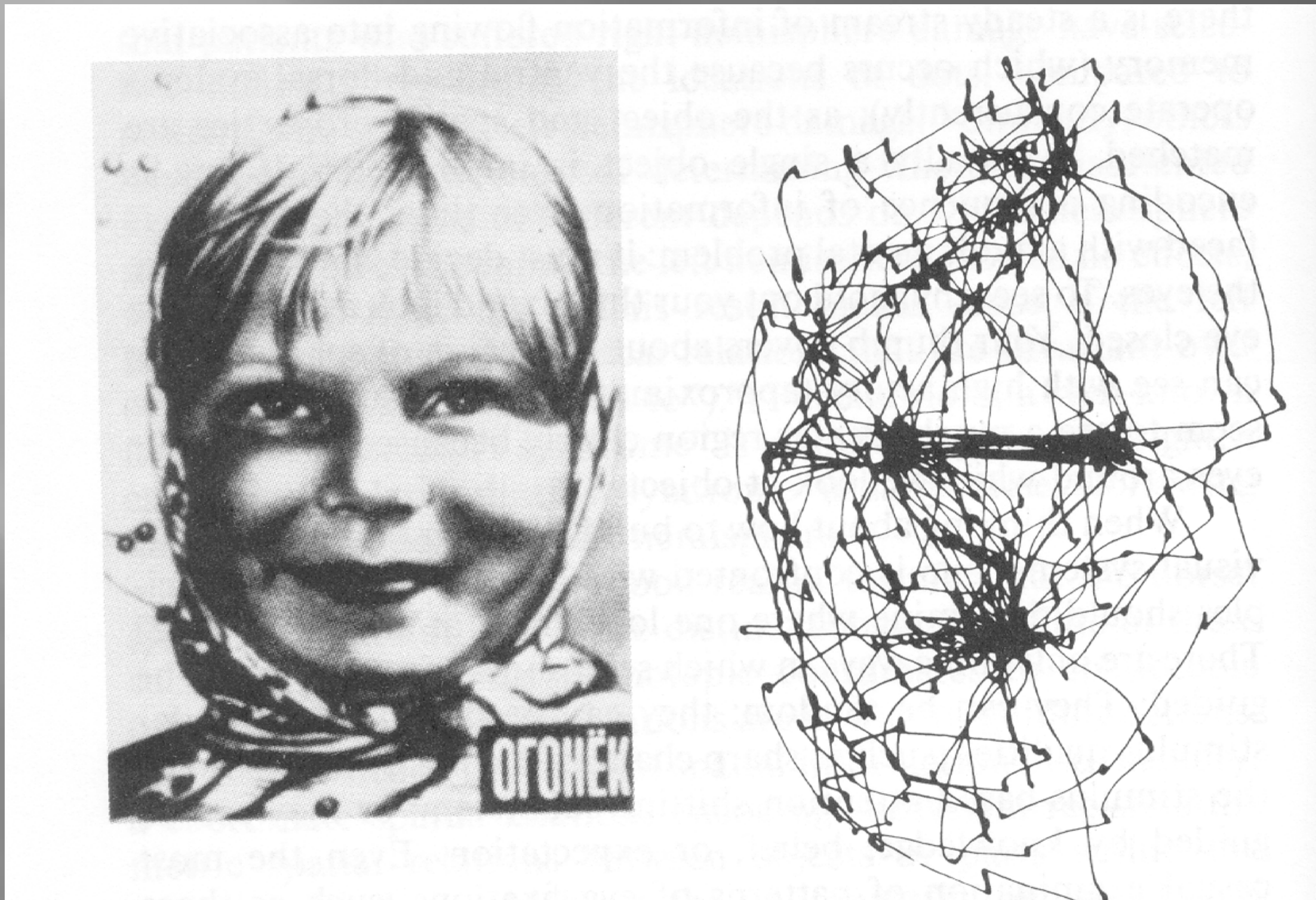
# A simple recognition experiment…

## Are both from the same subject?




Massimo Tistarelli

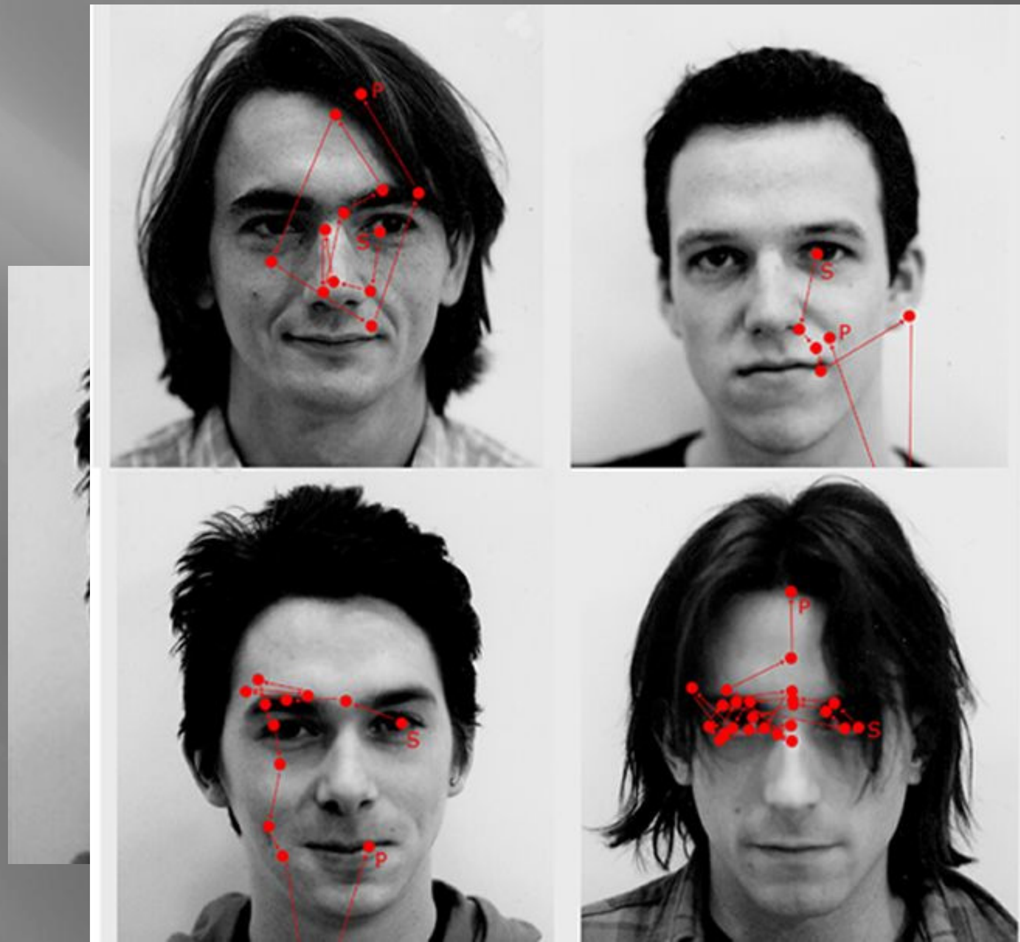
# Fixations



Eye movements while watching a girl's face
(A.L. Yarbus, "Eye Movements and Vision", Plenum Press, 1967)

# Visual tracking

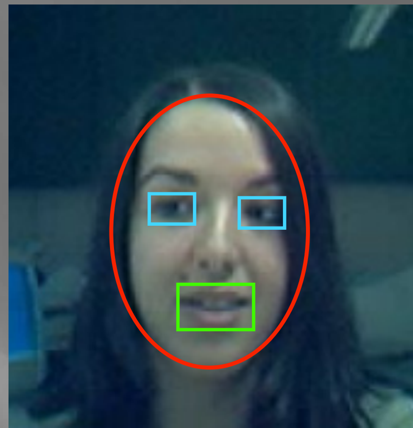
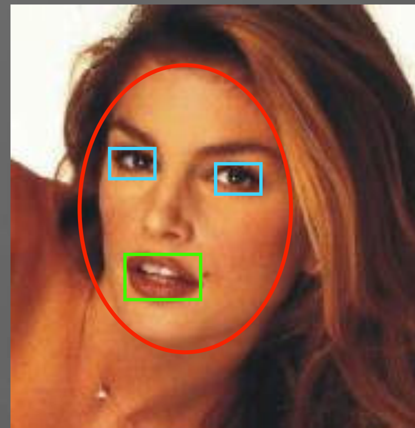
**The analysis of visual tracking patterns is sometimes misleading…**



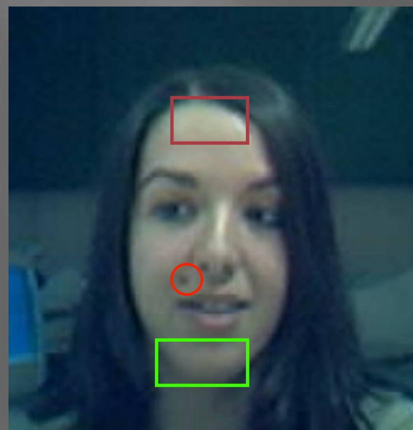
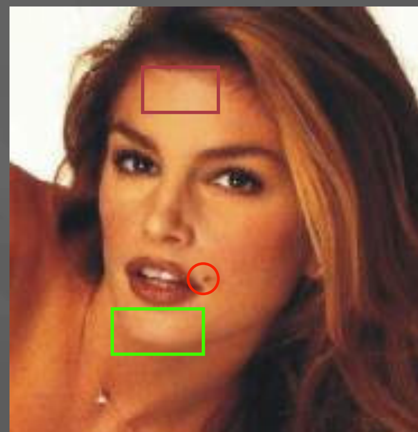
R.C. Miall, University Laboratory of Physiology, Oxford, UK
John Tchalenko, Camberwell College of Arts, London, UK

# Subject-specific representation

For localization and tracking we are interested on what every face has <u>in common</u> (to tell a face from “non-faces”)

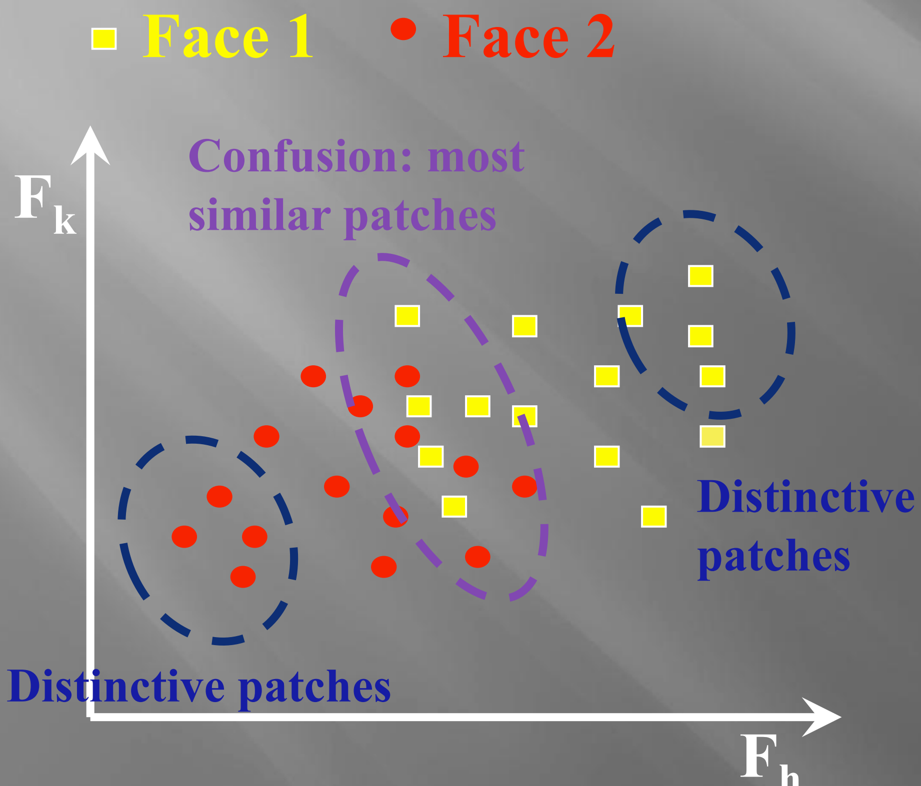
For recognition we are not interested on what faces have in common but rather <u>what differentiate</u> one face from another.

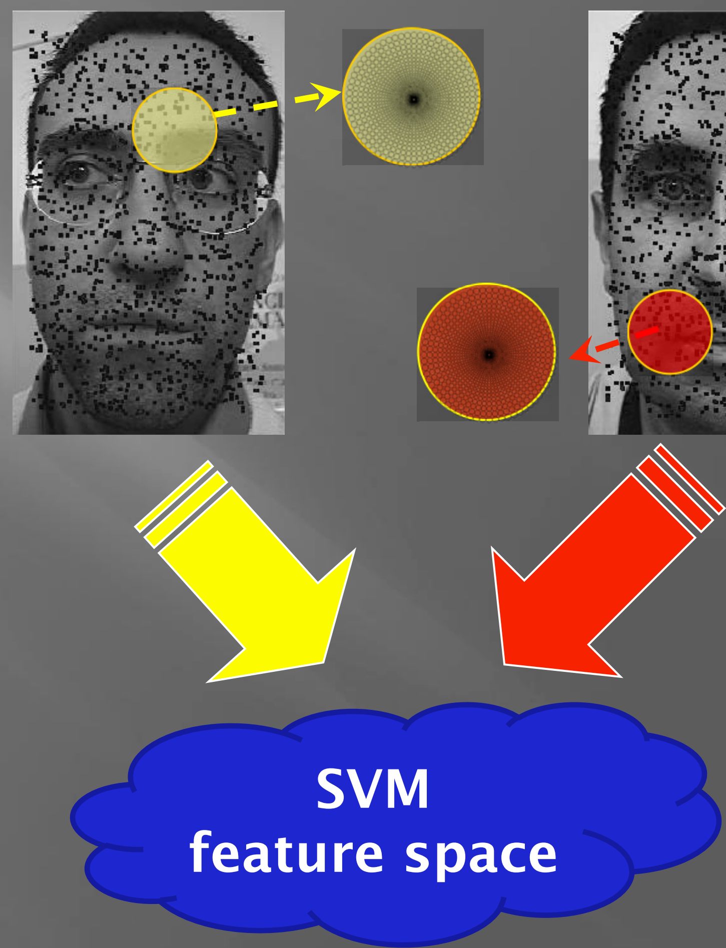



# Subject-specific representation

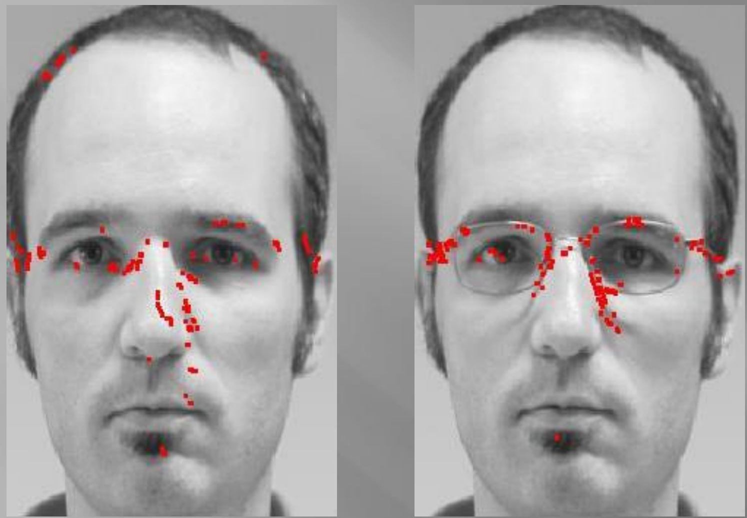
# Subject-specific representation

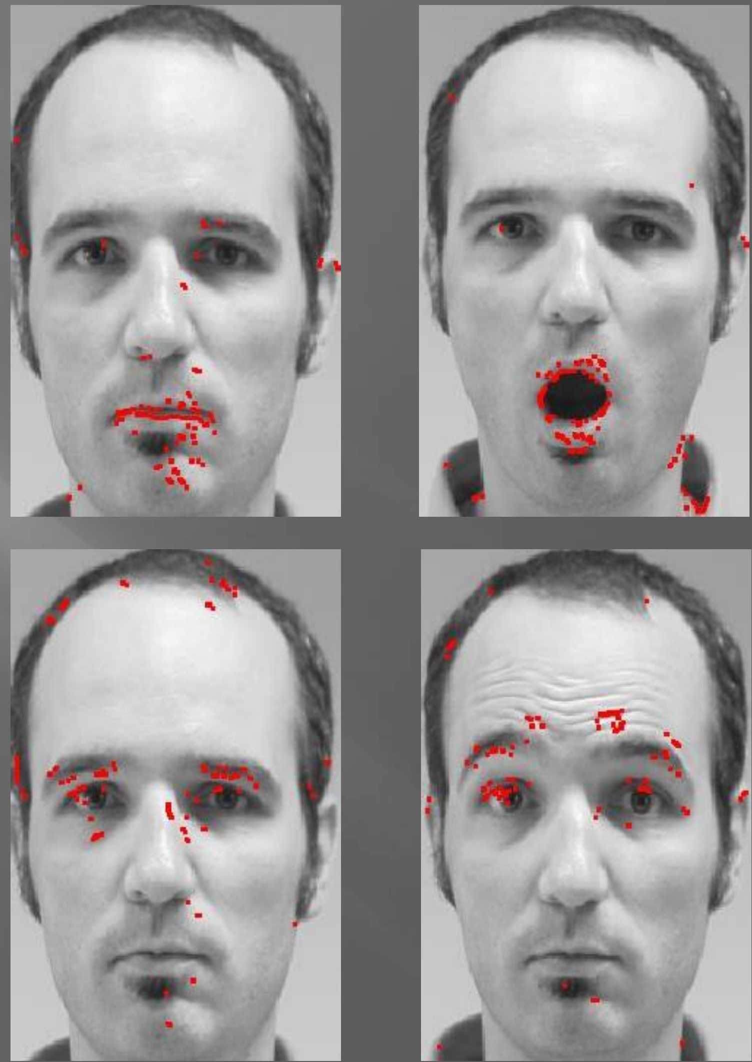
Bicego M., Brelstaff G., Brodo L., Grosso E., Lagorio A. and Tistarelli M. (2007) “Distinctiveness of faces: a computational approach”, ACM Transactions on Applied Perception, Vol. 5, n. 2, 2008.



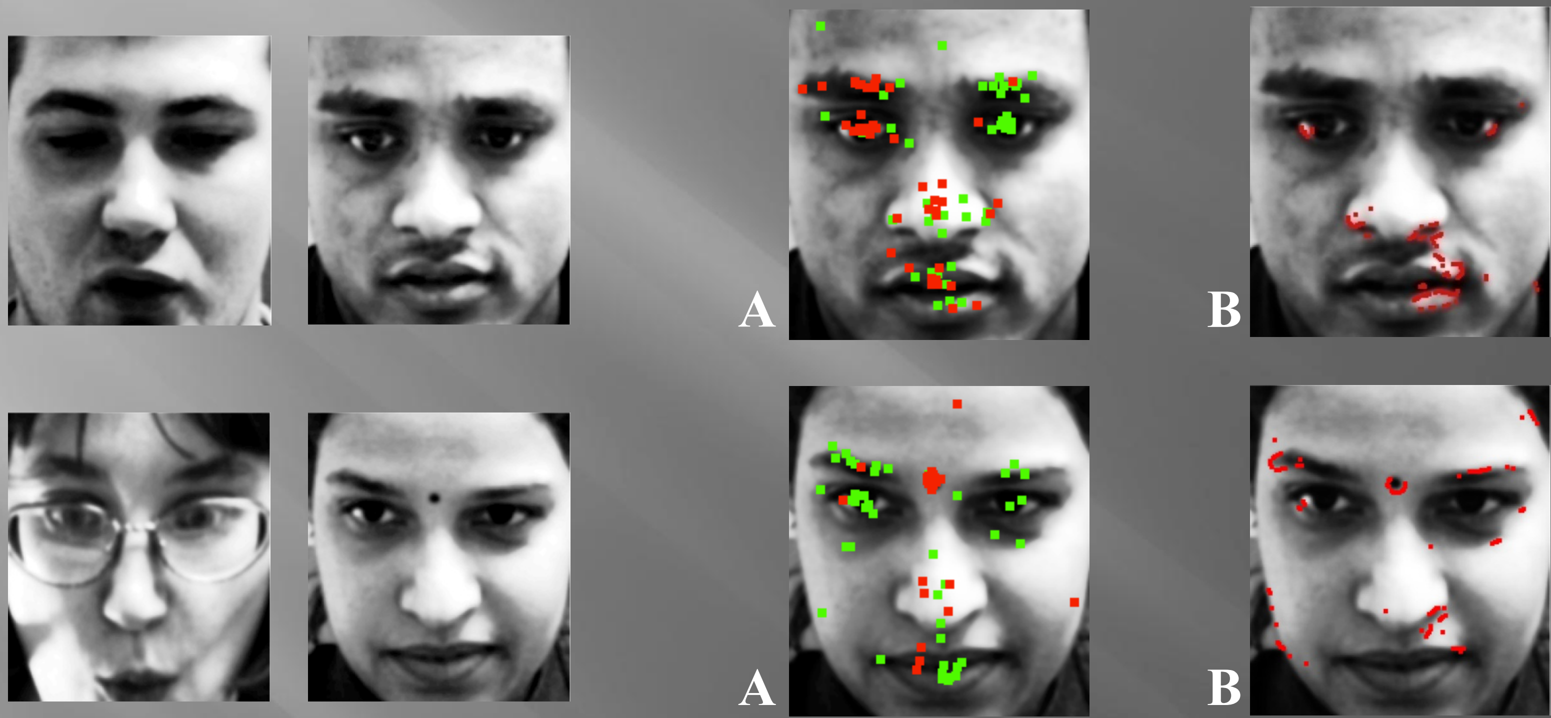
# Subject-specific representation





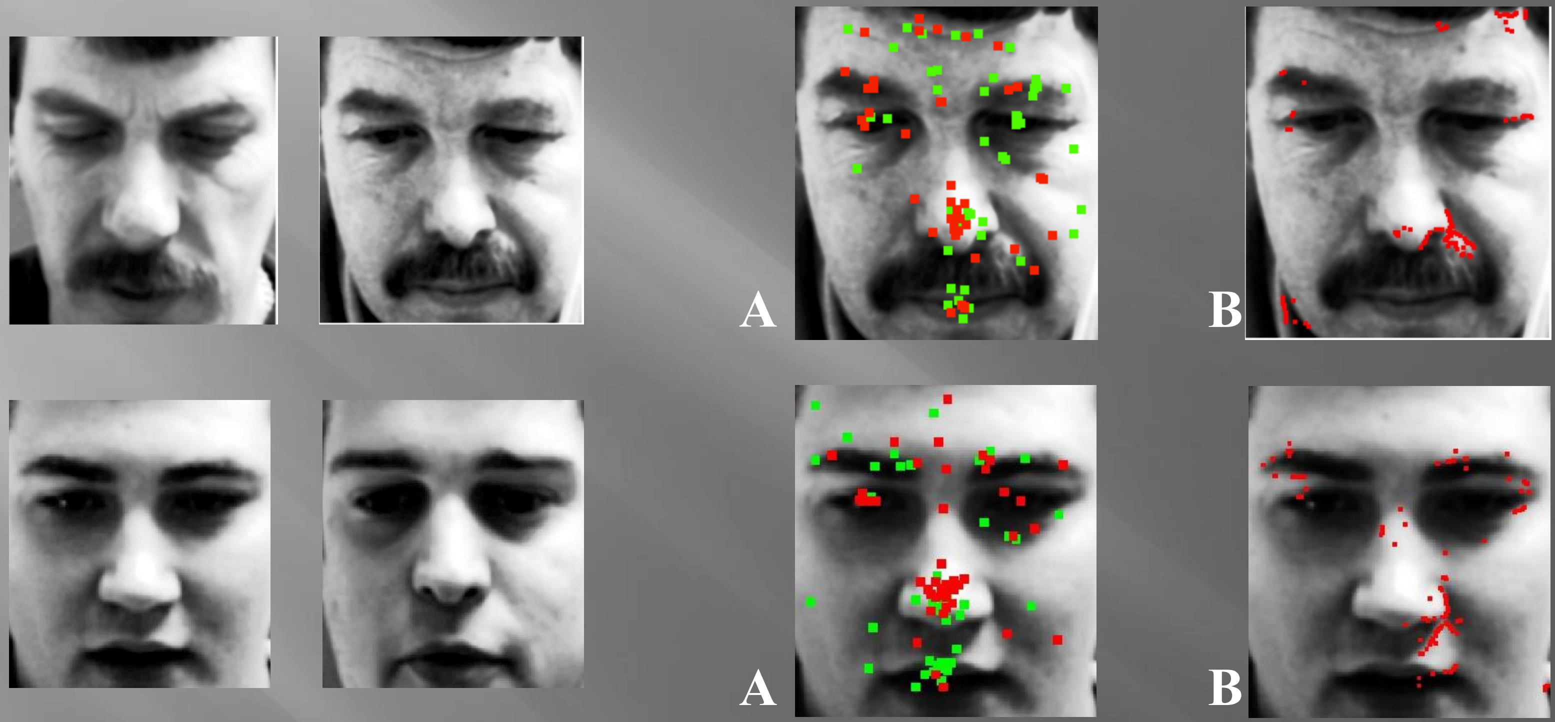
**Three examples of differences extracted from pairs of images of the same person. The 100 most weighted patches are shown**

# Subject-specific representation




(A) perceptual and (B) computational test results of saliency of local facial features. Red and green points in (A) represent the two positions selected by each of the 45 subjects (13 male, 32 female).

# Subject-specific representation




(A) perceptual and (B) computational test results of saliency of local facial features. Red and green points in (A) represent the two positions selected by each of the 45 subjects (13 male, 32 female).

# Identity verification

- 52 subjects from the BANCA database (26 M and 26 F).
- Matched Controlled (MC) BANCA protocol.
- The Weighted Error Rate (WER) is computed from the threshold from G1 on G2 (26 subjects each).
- Image areas are selectively deleted from the face image.

# Subject-specific representation

Quantitative results by selectively masking image areas

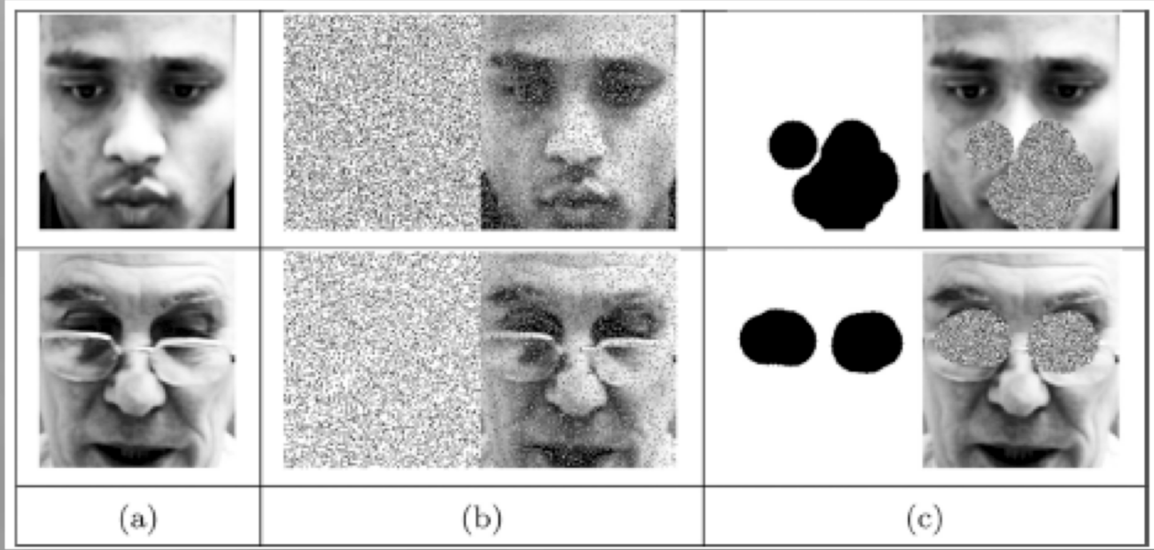


Table II. Different WER for Three Methodologies

| Methodology | WER (R = 0.1) (%) G1 | WER (R = 0.1) (%) G2 | WER (R = 1) (%) G1 | WER (R = 1) (%) G2 | WER (R = 10) (%) G1 | WER (R = 10) (%) G2 | Average |
|---|---|---|---|---|---|---|---|
| All points | 10.76 | 4.85 | 10.58 | 8.21 | 5.10 | 1.97 | 6.91 |
| Del. random pnt | 9.55 | 6.37 | 15.71 | 8.71 | 7.64 | 3.00 | 8.50 |
| Del. most imp. pnt | 11.62 | 7.46 | 19.58 | 13.37 | 9.11 | 4.46 | 10.93 |

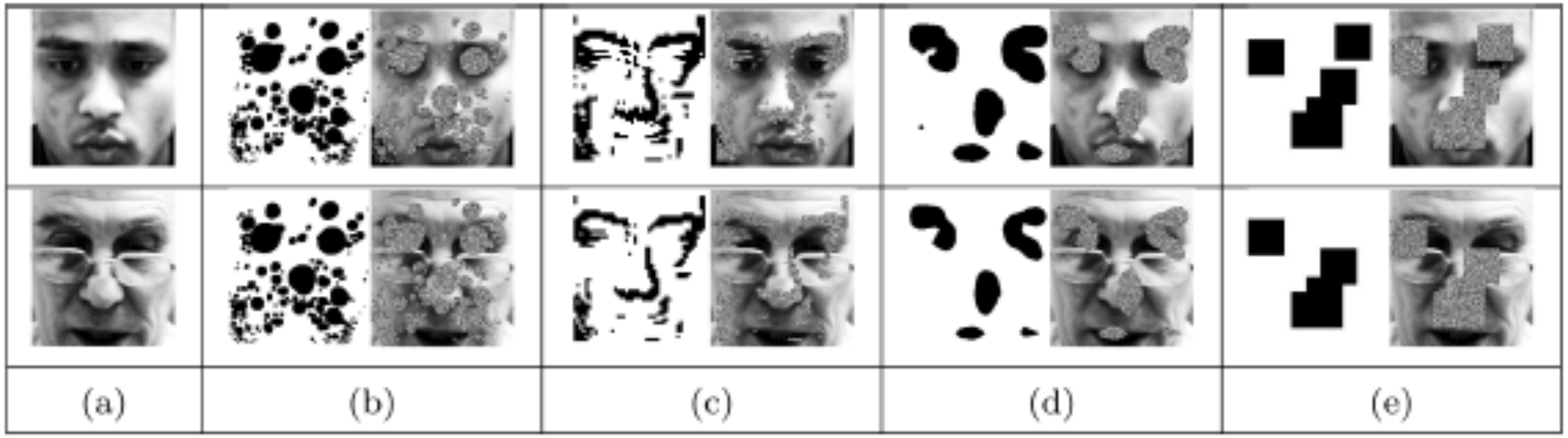
$$\mathrm{WER}(R) = \frac{FRR_{G2}(\theta_{G1}) + R \cdot FAR_{G2}(\theta_{G1})}{1+R}$$

# Subject-specific representation

- Comparative tests
  Lowe [2004] (SIFT),
  Salah et al. [2002],
  Walther [2006],
  Ullman et al. [2002].

Table III. Averaged WER on the Impairment Authentication Test for Each Method

| Method | Averaged WER (%) |
|---|---|
| Lowe [2004] | 10.15 |
| Salah et al. [2002] | 10.19 |
| Walther [2006] | 10.74 |
| Ullman et al. [2002] | 11.87 |
| Our algorithm | 10.93 |



# Selective attention

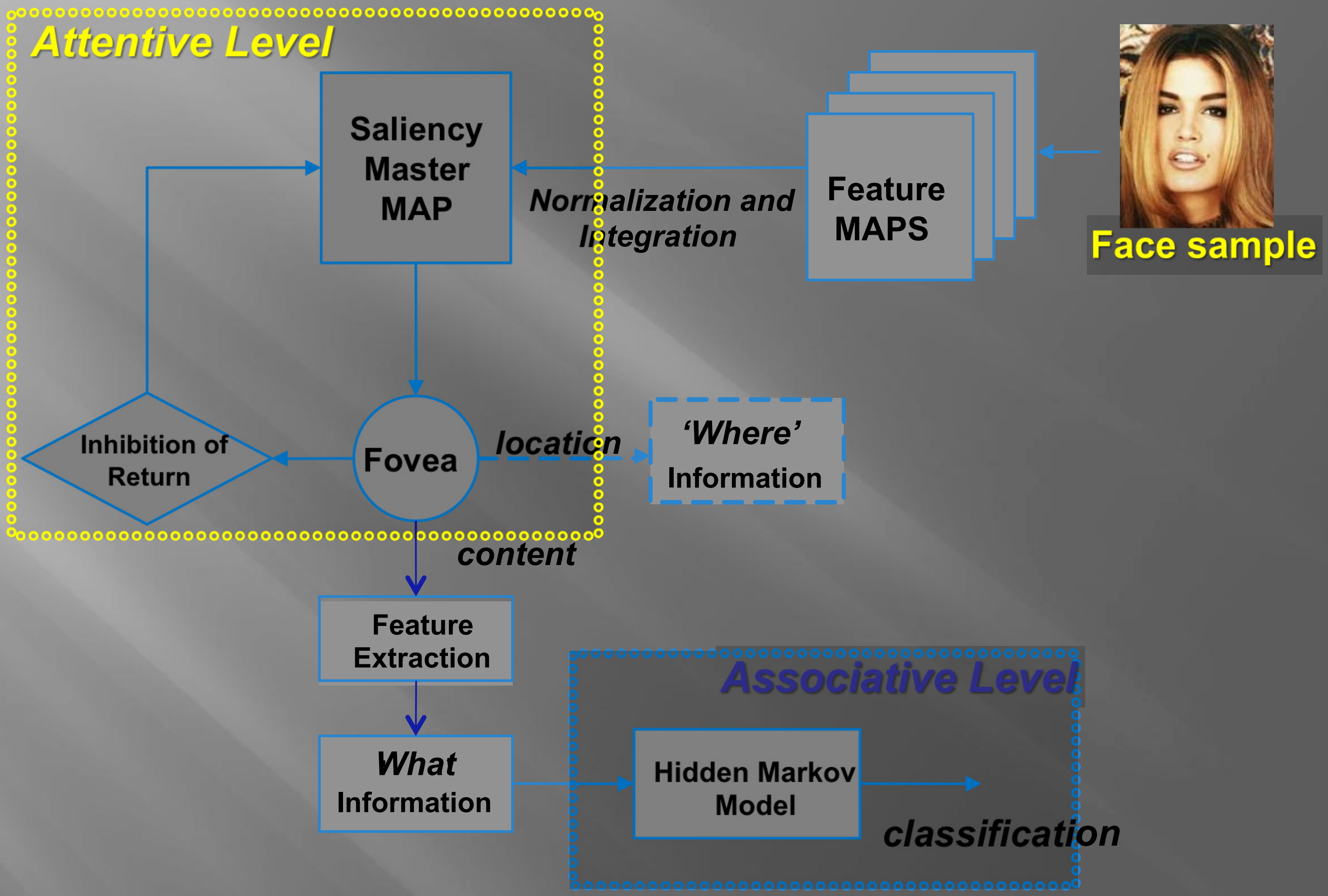
- Starting point: HMM based classification of faces
- "Walking on the face" for obtaining HMM  sequences

Standard raster scan-path

Saliency-based scan-path

**Attention drives face scanning**

# Attention-based classification



# Attention-based classification

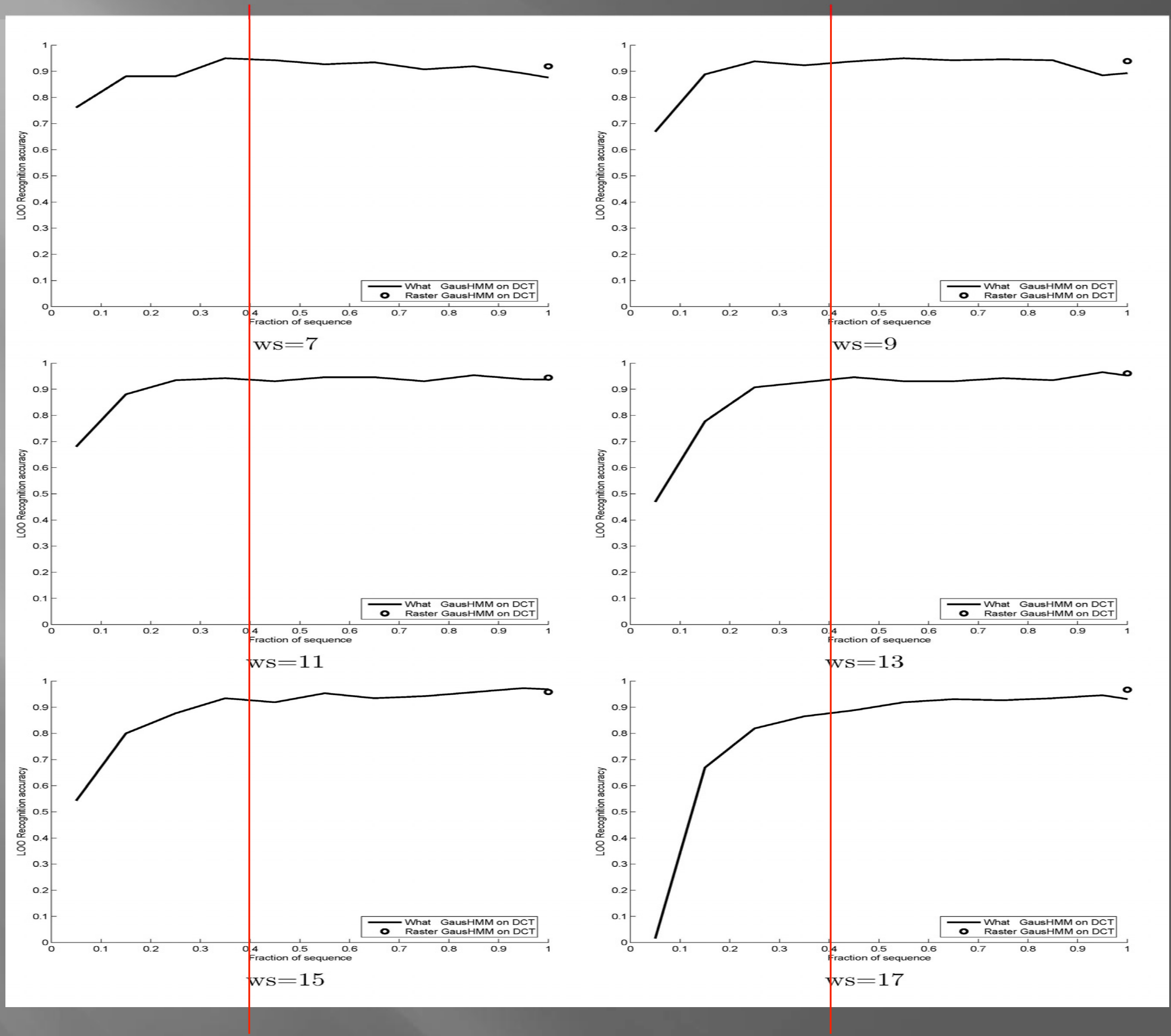
- Experiments on BANCA protocol MC
- Gabor wavelets for saliency map construction
- Employed features: gray levels, DCT coefficients, Haar wavelets

| Window Size | Average Acc. (std) | | Max Acc. | |
|---|---|---|---|---|
| | Biological | Raster | Biological | Raster |
| 7 | 87.62%(2.28%) | 91.92%(1.63%) | 91.15% | 93.08% |
| 9 | 89.31%(1.20%) | 93.92%(0.92%) | 90.38% | 95.00% |
| 11 | 93.69%(1.58%) | 94.46%(1.29%) | 95.77% | 95.77% |
| 13 | 95.23%(0.89%) | 96.08%(0.74%) | 96.15% | 97.31% |
| 15 | 96.85%(1.00%) | 95.85%(1.29%) | 98.08% | 97.31% |
| 17 | 93.15%(1.13%) | 96.69%(0.89%) | 95.00% | 98.08% |

**Table 2.** Comparison between raster and biological scanning

A. A. Salah, M. Bicego, L. Akarun, E. Grosso, M. Tistarelli: "Hidden Markov model-based face recognition using selective attention", *Human Vision and Electronic Imaging XII*, Proc. of SPIE, vol. 6492, (2007)

# Attention-based classification



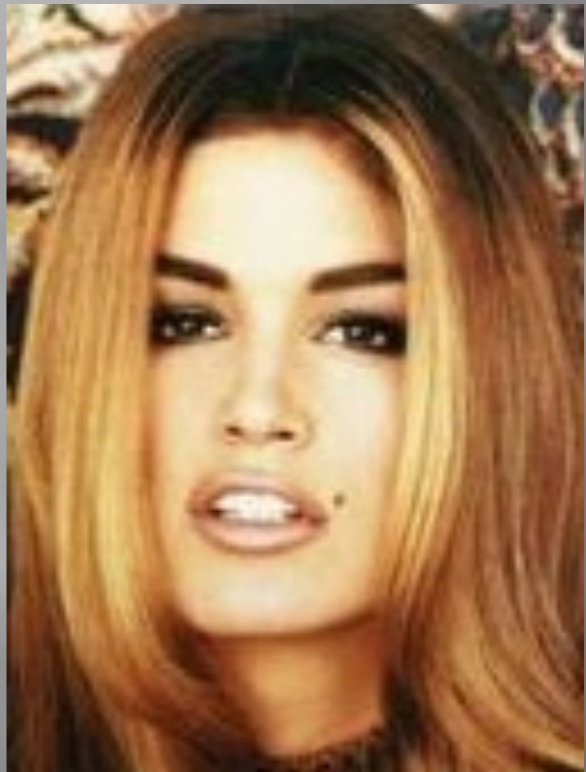
# Attention-based classification

- About 40% of the saccades are sufficient to obtain the same result of raster scan
- In some cases further saccades may decrease the classification performances
- The system can be improved by exploiting also the *where* information

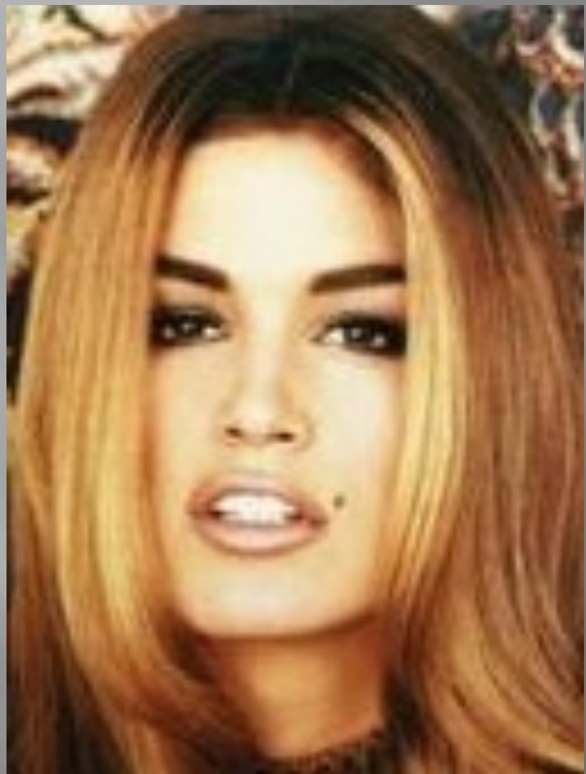
# DEALING WITH AGE PROGRESSION

# Facing aging

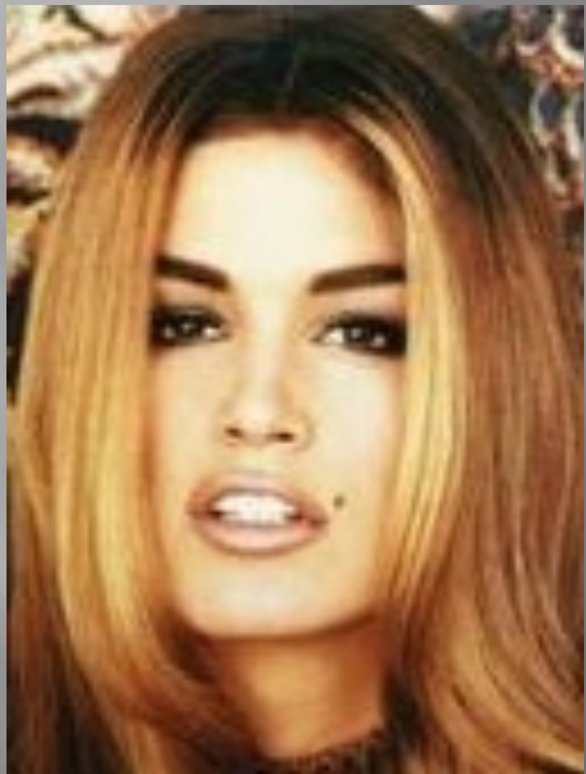
# Facing aging

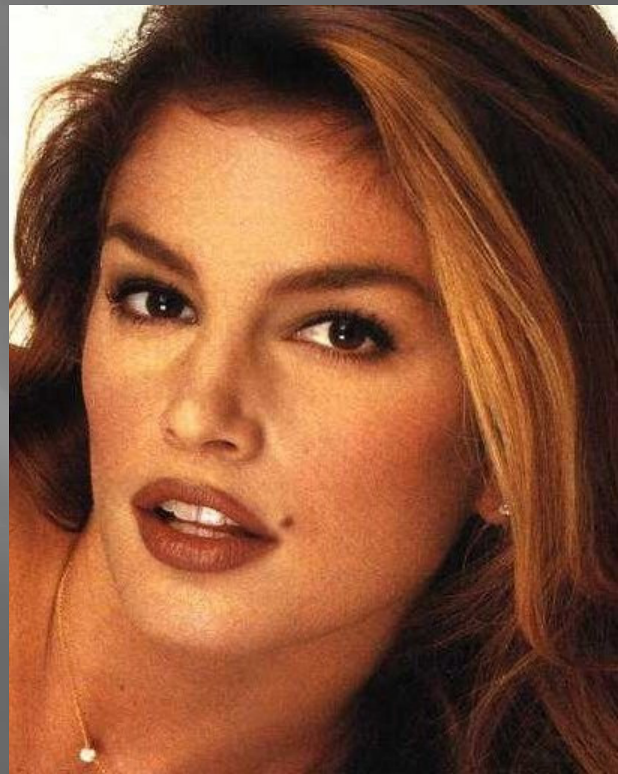


# Facing aging

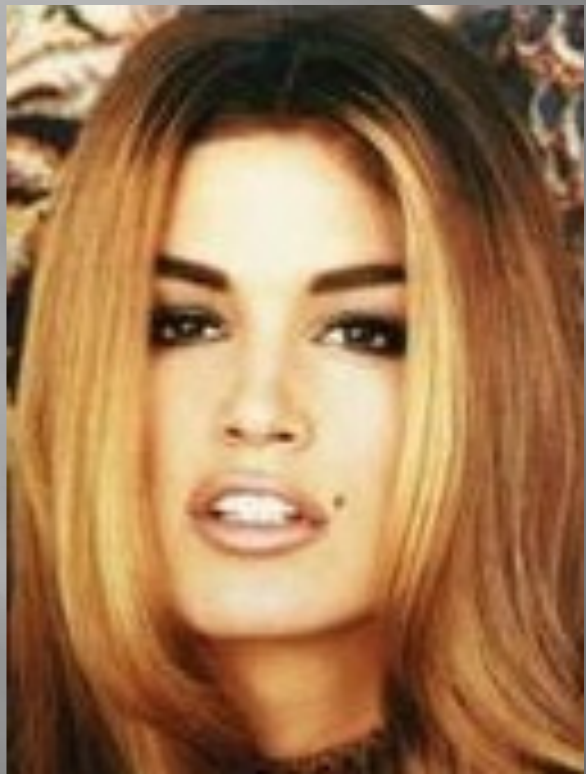


# Facing aging

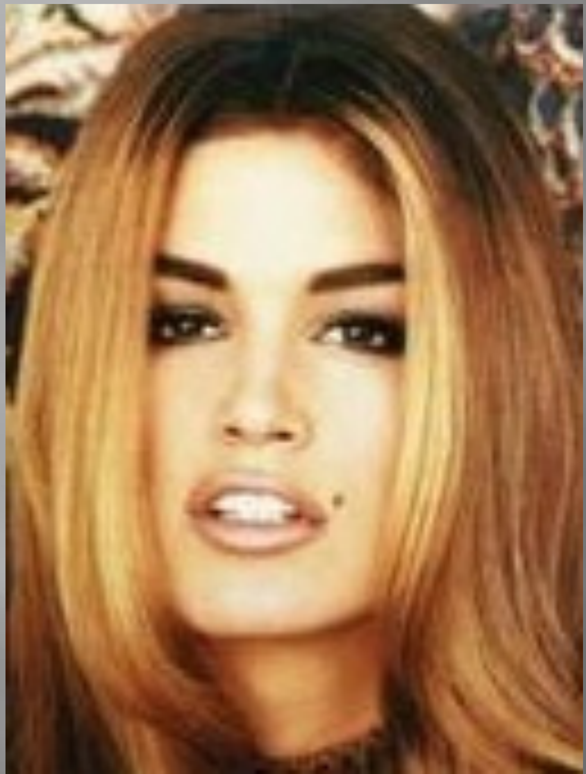




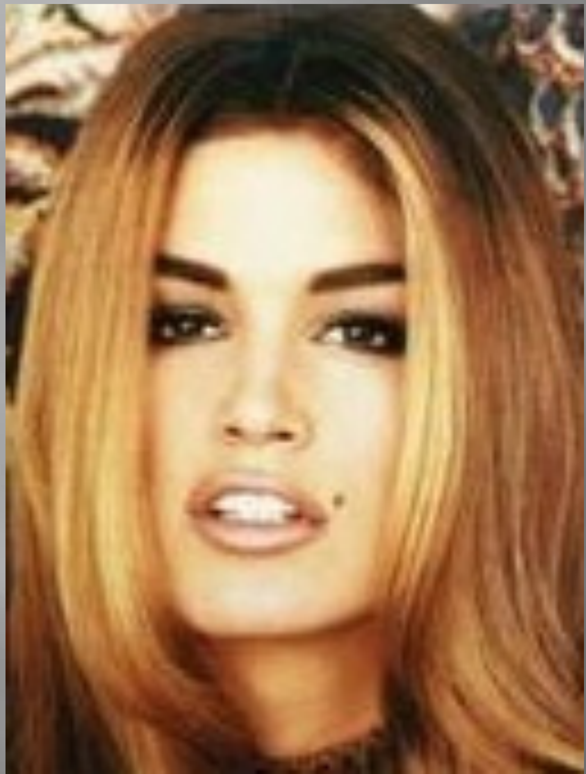
# Facing aging



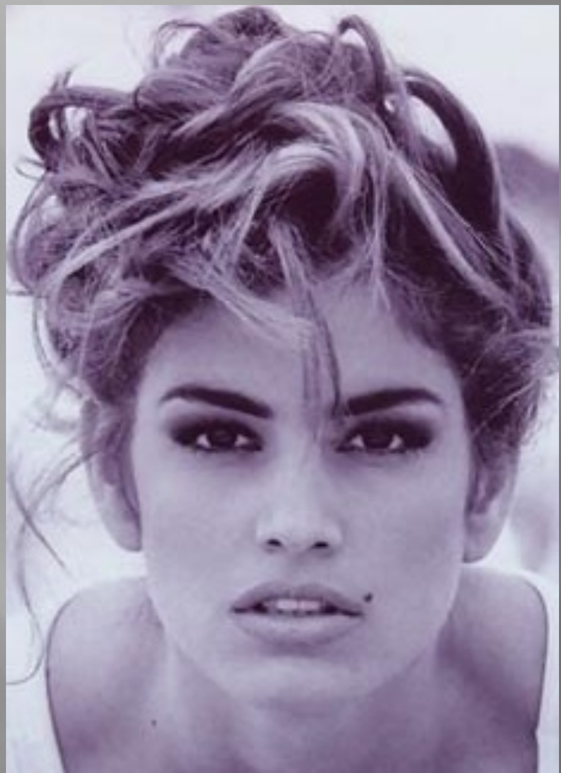
# Facing aging



# Facing aging

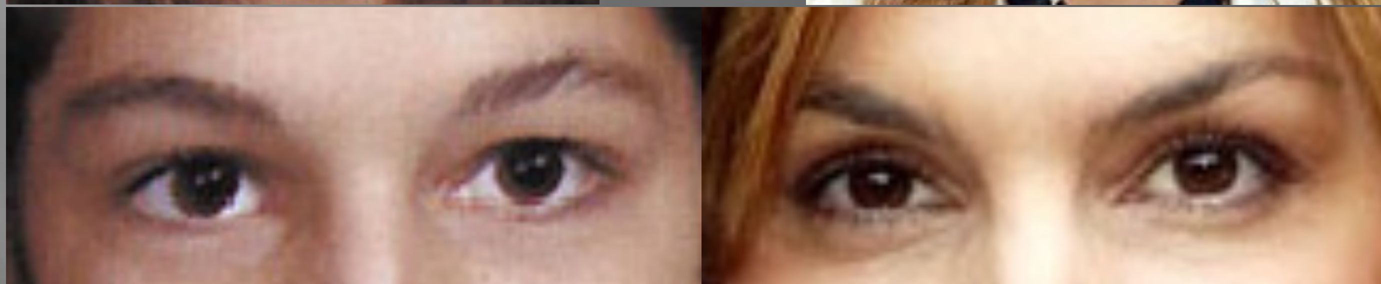


# Facing aging



# Facing aging

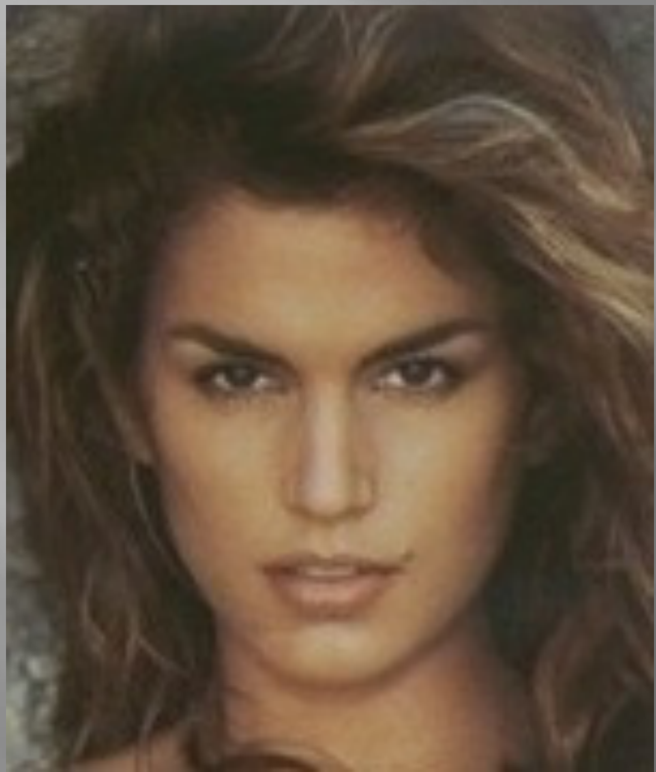
Vision Lab



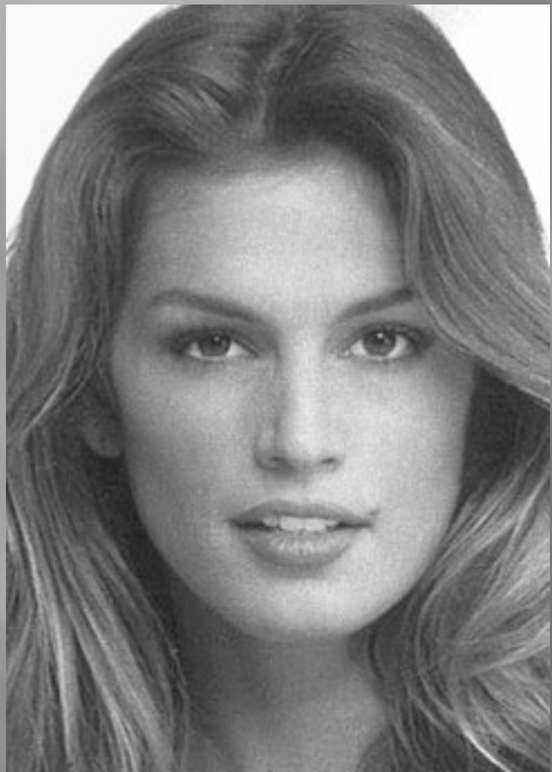
# Facing aging



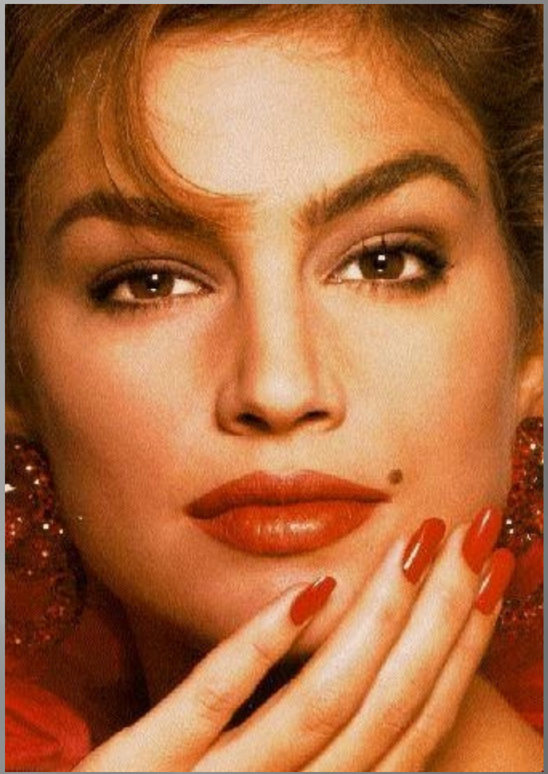

# Facing aging



# Facing aging

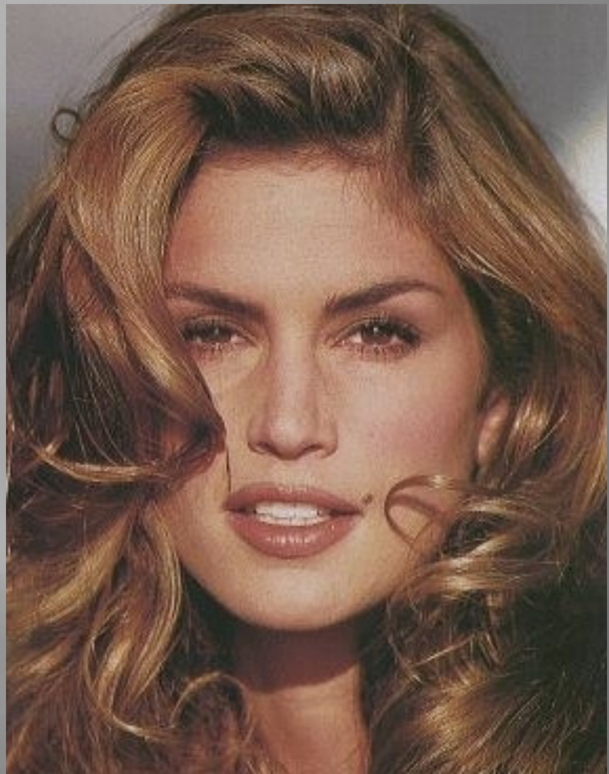


# Facing aging

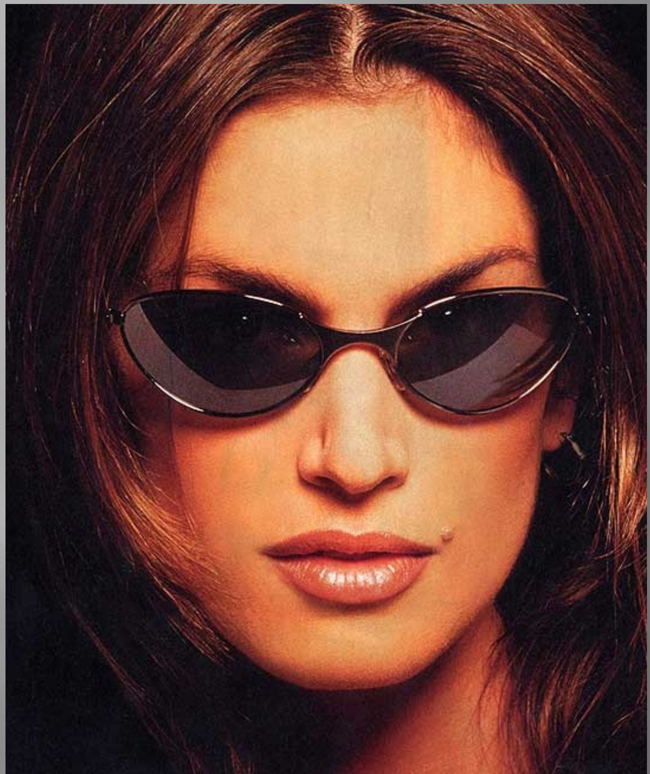


Vision Lab

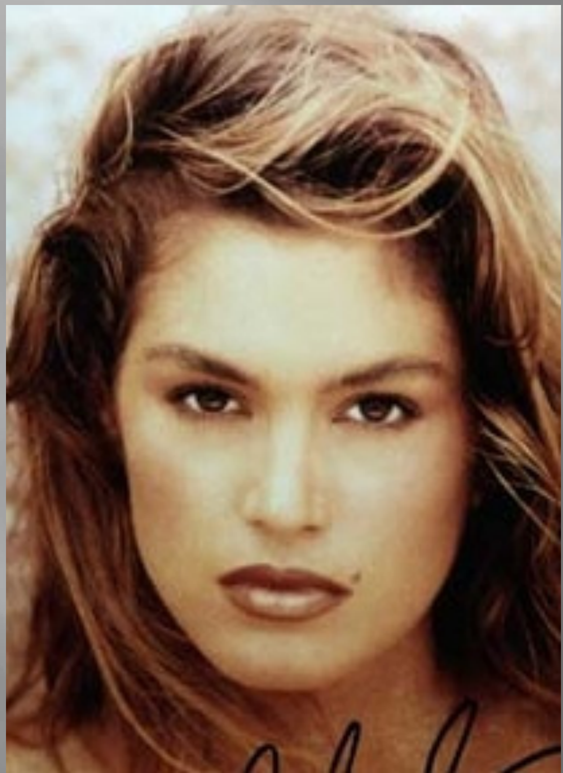
# Facing aging



# Facing aging



# Facing aging



# Facing aging

Vision Lab

# Facing aging

# Effects of age progression

Vision Lab

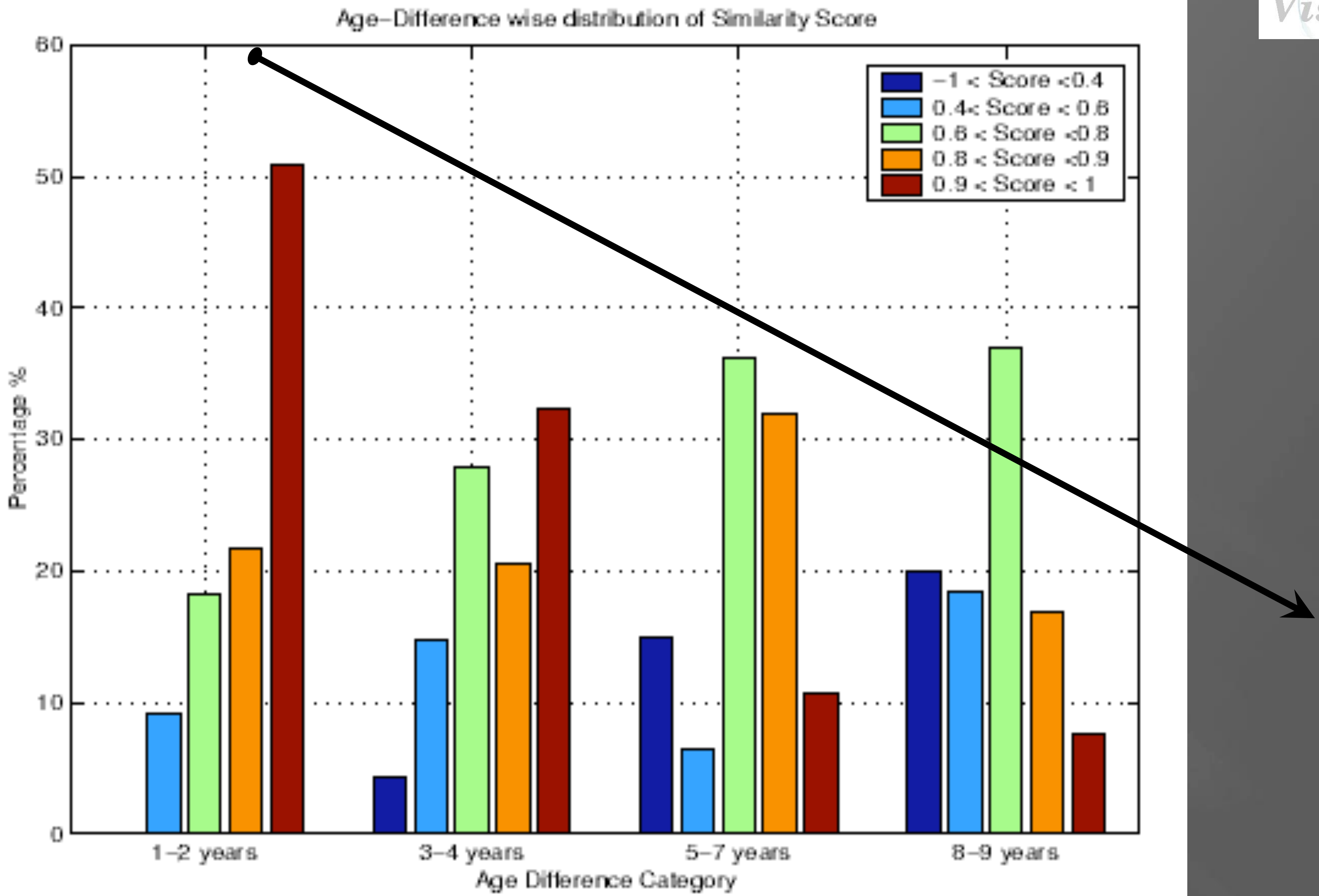
**Passport renewal** →

Variations in pairs of passport images
( 25 years & above )
- Wrinkles
- Facial Hair
- Expression
- Glasses

**Age Based Similarity Measure**

| Age Difference | First Set | | Second Set | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Expression | | Glasses | | Facial Hair | |
| | $\mu$ | $\sigma^2$ | $\mu$ | $\sigma^2$ | $\mu$ | $\sigma^2$ | $\mu$ | $\sigma^2$ |
| *1-2 yrs* | 0.85 | 0.02 | 0.70 | 0.021 | 0.83 | 0.01 | 0.67 | 0.04 |
| *3-4 yrs* | 0.77 | 0.03 | 0.65 | 0.07 | 0.75 | 0.02 | 0.63 | 0.01 |
| *5-7 yrs* | 0.70 | 0.06 | 0.59 | 0.01 | 0.72 | 0.02 | 0.59 | 0.10 |
| *8-9 yrs* | 0.60 | 0.08 | 0.55 | 0.10 | 0.68 | 0.18 | 0.55 | 0.10 |

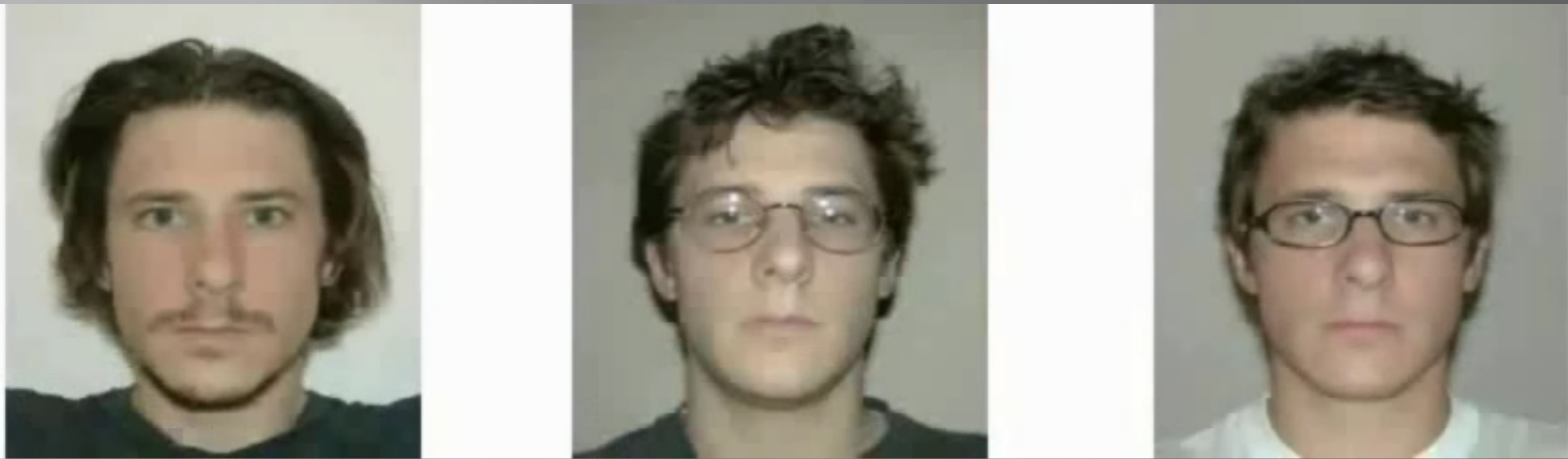
**As age difference increased:**
**mean score decreased, variance increased**




Similarity scores decreased as age difference increases

# Aging … over time

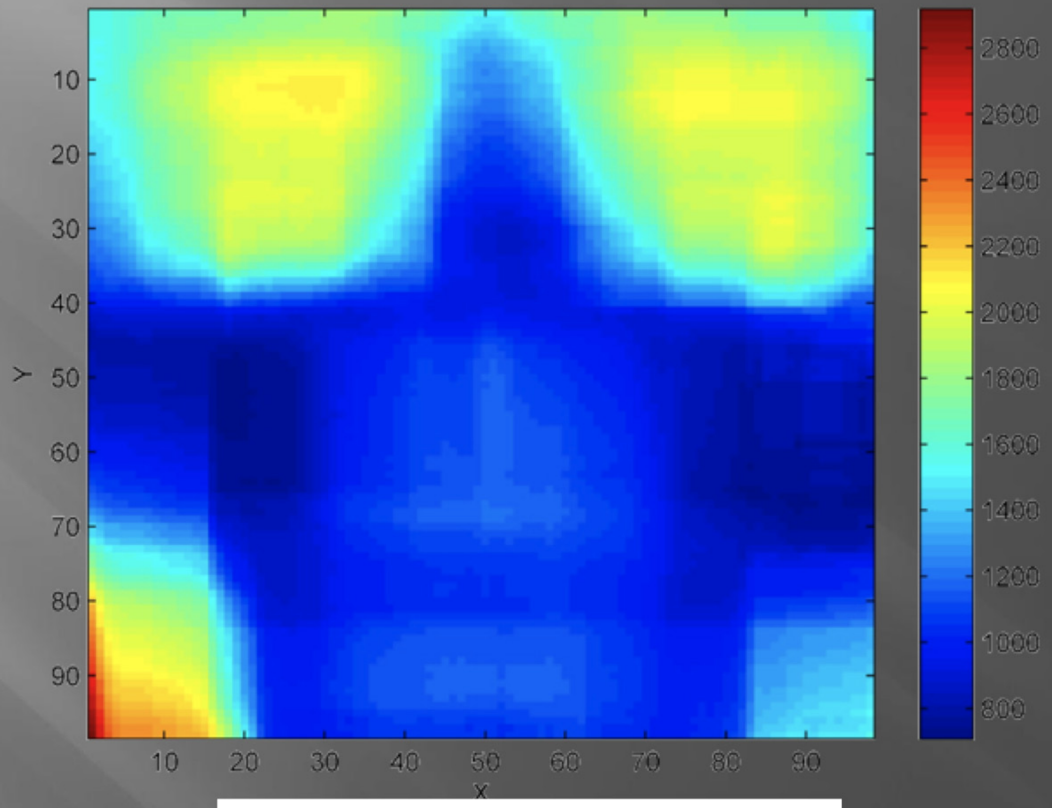

# Aging … over time



$$d_{I_1,I_2}(x,y) = \frac{1}{2}\left(\frac{1}{|P_{I_1}|}\sum_{p\in P_{I_1}}\omega(p) + \frac{1}{|P_{I_2}|}\sum_{q\in P_{I_2}}\omega(q)\right)$$

# Photometric effects

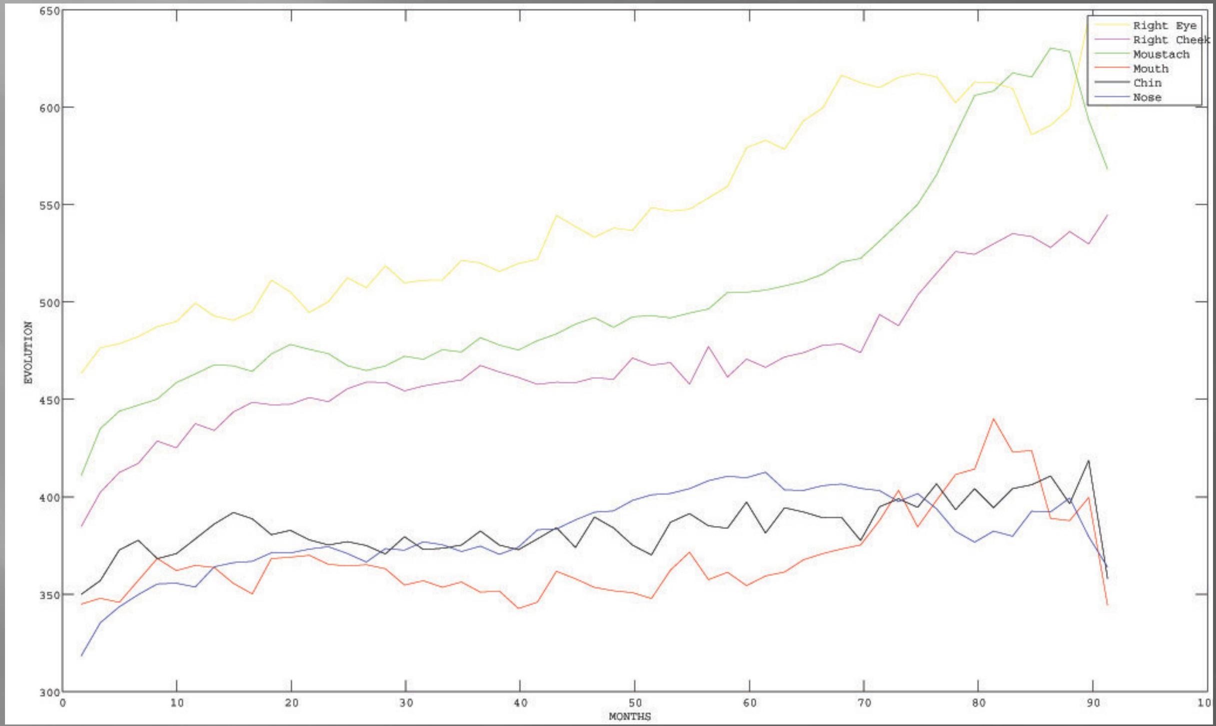
**Time evolution of facial features over 4 years**




$$E(x,y,t) = \frac{1}{T}\sum_{i=1}^{T} d_{I_i,I_{i+t}}(x,y)$$

# Photometric effects

## Time evolution of facial features over 8 years

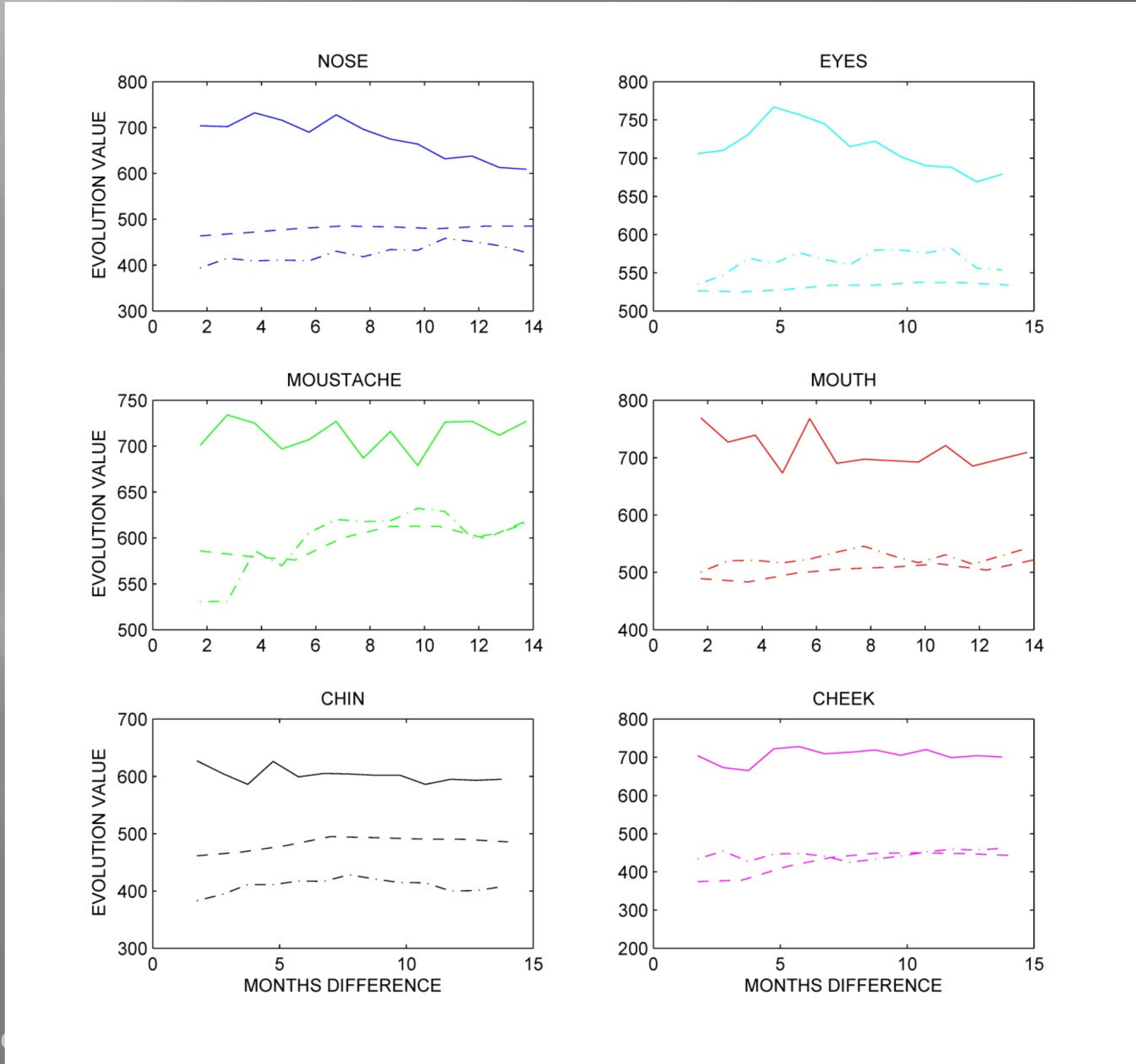



$$E(L,t) = \frac{1}{w_L h_L} \sum_{(x,y)\in L} E(x,y,t)$$

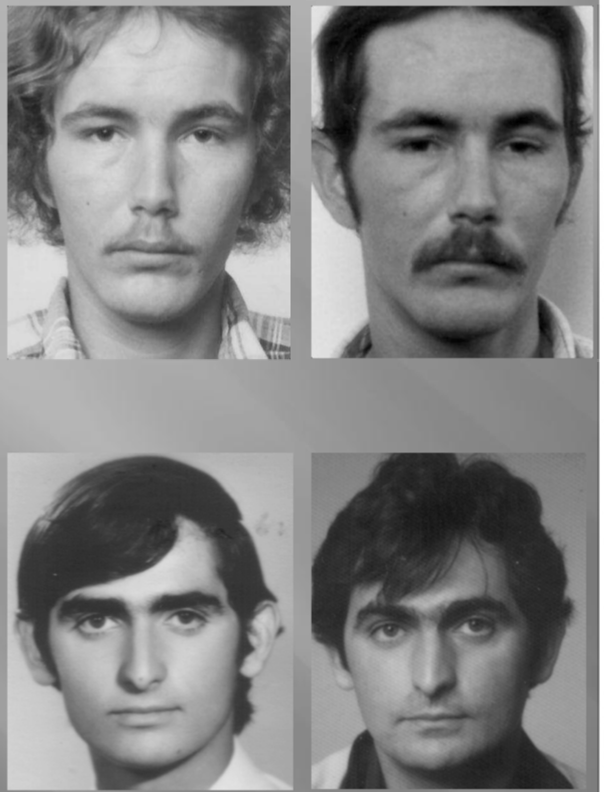
# Photometric effects

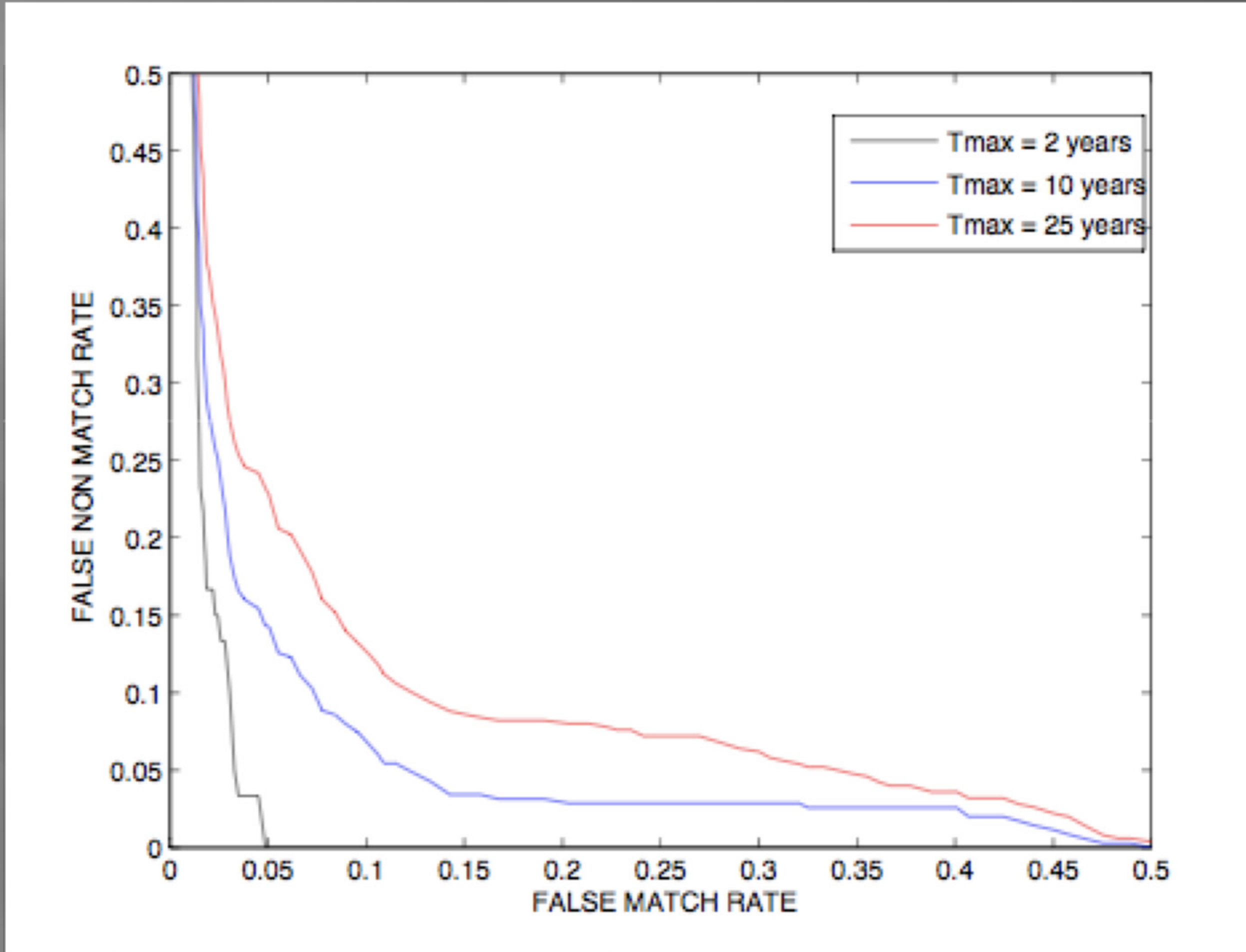
Vision Lab

## Comparative time evolution of features for different subjects



# Face matching across age





# Ageing

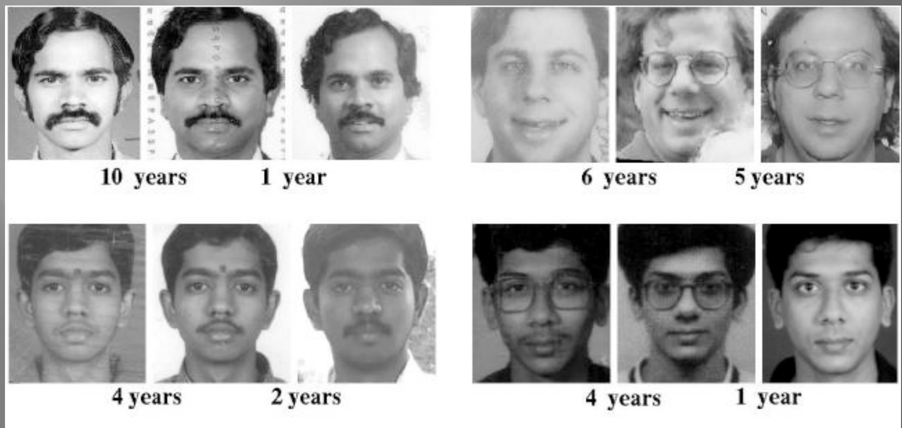
# Face verification across age progression

(R. Chellappa - CVPR 2005, IEEE Trans. IP, 2006)

## Problem Statement :

- Given a pair of age separated face images of an individual, what is the confidence measure in verifying the identity ?
- How does age progression affect facial similarity ?

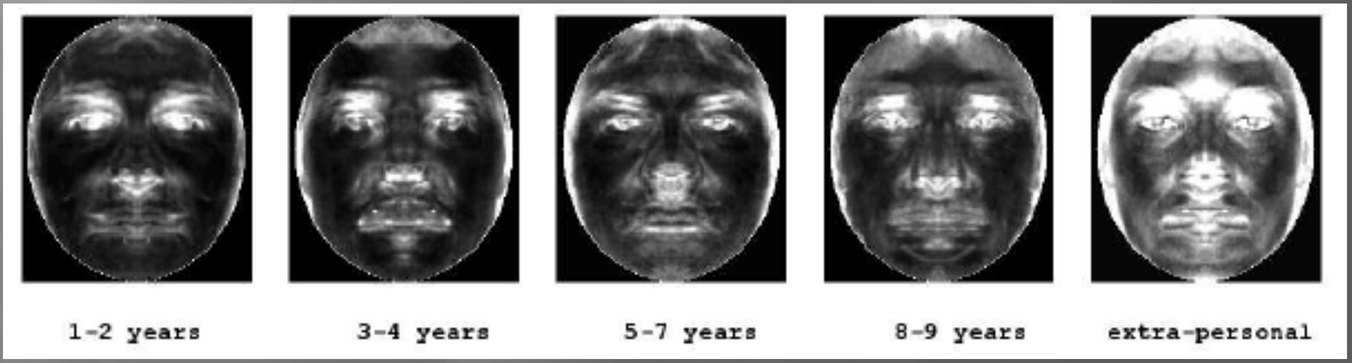
### Passport Image database



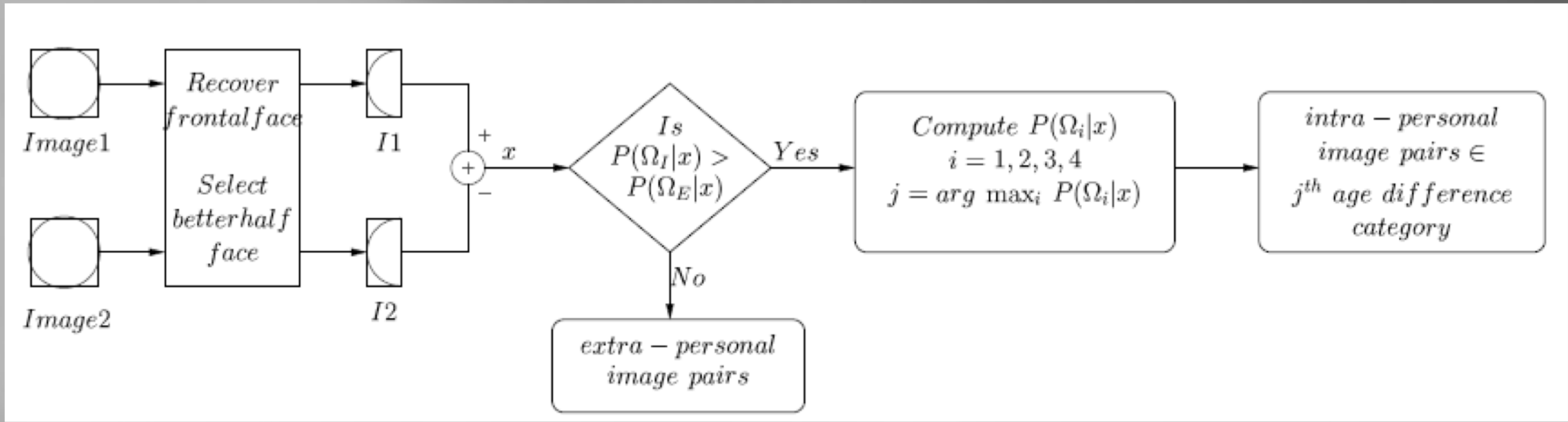

- 465 pairs of passport images
- Age range : 20 yrs to 70 yrs
- Pair-wise age difference
  - ✓ 1 - 2 yrs : 165
  - ✓ 3 - 4 yrs : 104
  - ✓ 5 - 7 yrs : 81
  - ✓ 8 - 9 yrs : 115

# Classifying aging

- In adults, facial aging effects are manifested in the form of wrinkles and other skin artifacts. Loss in elasticity of muscles, loss in facial fat etc. results in the sagging of facial features and hence wrinkles appear on faces.
- The age difference classifier is based on the difference image obtained from age separated face images.



# Classifying aging




$$P(\Omega_I|\mathbf{x}) = \frac{P(\mathbf{x}|\Omega_I)P(\Omega_I)}{P(\mathbf{x}|\Omega_I)P(\Omega_I) + P(\mathbf{x}|\Omega_E)P(\Omega_E)}$$

if intra-personal

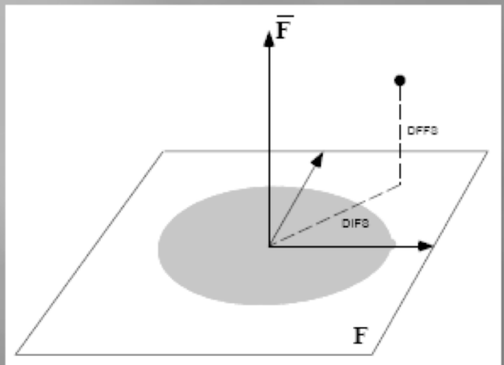
$$P(\Omega_i|\mathbf{x}) = \frac{P(\mathbf{x}|\Omega_i)P(\Omega_i)}{\sum_{j=1}^{4} P(\mathbf{x}|\Omega_j)P(\Omega_j)}$$

Age-difference based intra-personal class

# Classifying aging

**Subspace Density Estimation : Intra Personal class**

**Assume Gaussian distribution of Intra – personal image differences**




(Courtesy : Moghaddam 1997)

$$
\begin{aligned}
P(\mathbf{x}|\Omega_I) &= \frac{\exp(-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T\Sigma^{-1}(\mathbf{x}-\bar{\mathbf{x}}))}{(2\pi)^{N/2}|\Sigma|^{1/2}} \\
&= \frac{\exp(-\frac{1}{2}\sum_{i=1}^{N}\frac{y_i^2}{\lambda_i})}{(2\pi)^{N/2}\prod_{i=1}^{N}\lambda_i^{1/2}} \\
&\simeq \left[\frac{\exp(-\frac{1}{2}\sum_{i=1}^{k}\frac{y_i^2}{\lambda_i})}{(2\pi)^{k/2}\prod_{i=1}^{k}\lambda_i^{1/2}}\right]\cdot\left[\frac{\exp\left(-\frac{\epsilon^2(\mathbf{x})}{2\rho}\right)}{(2\pi\rho)^{(N-M)/2}}\right] \\
&= P_F(\mathbf{x}|\Omega_I)\,.\,\hat{P}_{\bar{F}}(\mathbf{x}|\Omega_I)
\end{aligned}
$$

Principal subspace and its orthogonal complement for Gaussian density

Marginal density in F space

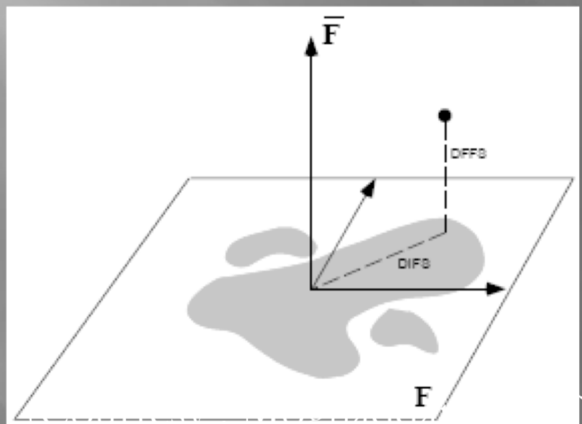
Marginal density in complementary space

# Classifying aging

## Subspace Density Estimation : Extra personal class

Assume feature space F to be estimated by a parametric mixture model (mixture of Gaussian – use EM approach to estimate the parameters)

Assume components of complementary space to be Gaussian

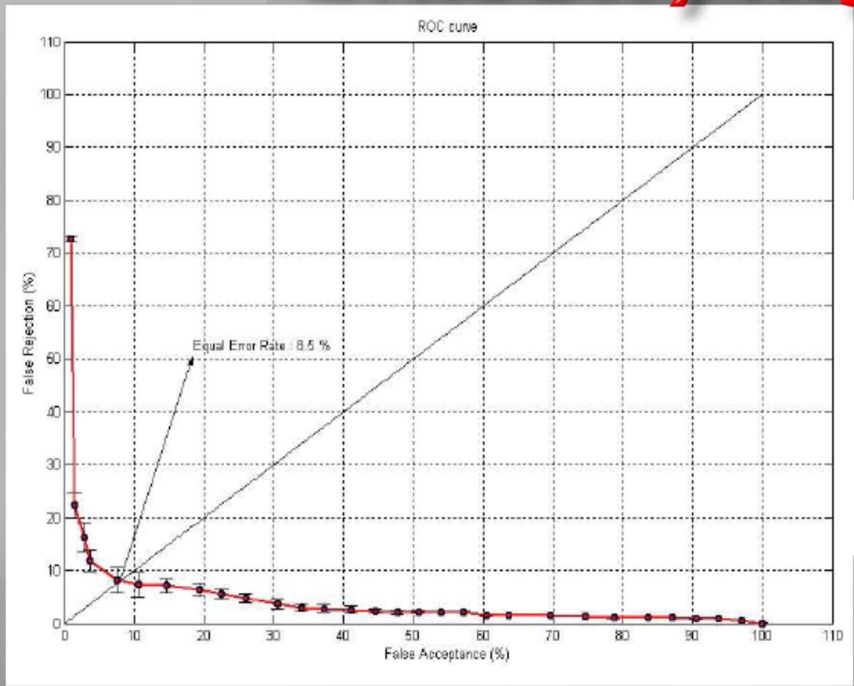



$$\hat{P}(\mathbf{z}|\Omega_E) = P(\mathbf{y}|\Theta^*) \; . \; \hat{P}_{\bar{F}}(\mathbf{z}|\Omega_E)$$

$$P(\mathbf{y}|\Theta) = \sum_{i=1}^{N_c} w_i N(\mathbf{y}; \mu_i, \Sigma_i)$$

$$\Theta^* = argmax \left[ \prod_{i=1}^{M} P(\mathbf{y}_i|\Theta) \right]$$

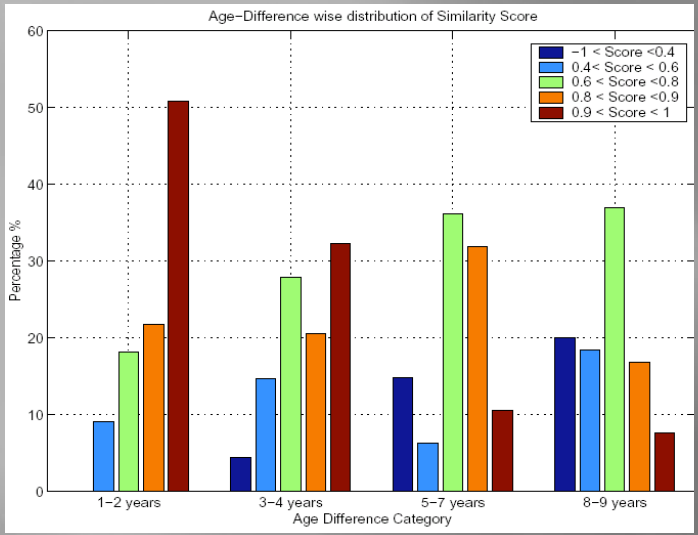
# Classifying aging




| Type | Class | 1-2 yrs | 3-4 yrs | 5-7 yrs | 8-9 yrs |
|---|---|---|---|---|---|
| Original set of images | $\Omega_1$ | **41.0 (1.1)** | 12.0 (6.9) | 9.0 (5.0) | 38.0 (7.2) |
| | $\Omega_2$ | 8.0 (5.0) | **46.0 (5.6)** | 8.0 (4.9) | 37.0 (9.2) |
| | $\Omega_3$ | 10.0 (3.3) | 8.0 (6.3) | **53.0 (4.4)** | 28.0 (6.9) |
| | $\Omega_4$ | 10.0 (2.3) | 12.0 (7.3) | 5.0 (5.4) | **73.0 (8.2)** |

- Intra-personal image pairs with little variations due to facial expressions / glasses / facial hair were more often classified correctly to their age difference category.
- Image pairs with significant variations in the above factors were incorrectly classified under 8-9 yrs category.

# Similarity measures




Similarity measure computed as the correlation of principal components corresponding to 95 % of the variance

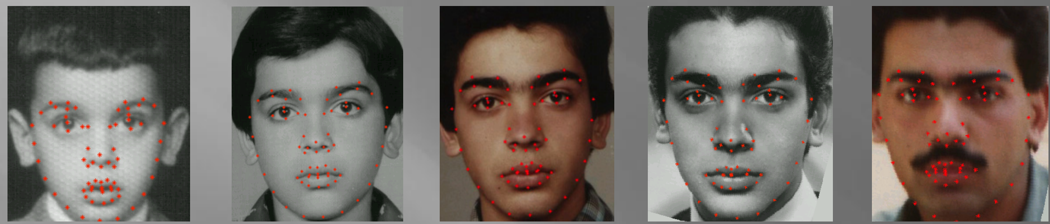
Similarity scores between intra-personal images dropped as age-difference increased

**Age Based Similarity Measure**

| Age Difference | First Set | | Second Set | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Expression | | Glasses | | Facial Hair | |
| | $\mu$ | $\sigma^2$ | $\mu$ | $\sigma^2$ | $\mu$ | $\sigma^2$ | $\mu$ | $\sigma^2$ |
| *1-2 yrs* | 0.85 | 0.02 | 0.70 | 0.021 | 0.83 | 0.01 | 0.67 | 0.04 |
| *3-4 yrs* | 0.77 | 0.03 | 0.65 | 0.07 | 0.75 | 0.02 | 0.63 | 0.01 |
| *5-7 yrs* | 0.70 | 0.06 | 0.59 | 0.01 | 0.72 | 0.02 | 0.59 | 0.10 |
| *8-9 yrs* | 0.60 | 0.08 | 0.55 | 0.10 | 0.68 | 0.18 | 0.55 | 0.10 |

# Craniofacial growth model – young faces

R. Chellappa CVPR 2006



7 yrs 13 yrs 14 yrs 17 yrs 20 yrs

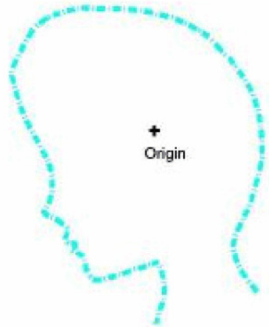
- Craniofacial growth is one of the most prominent manifestations of aging effects in children.
- Skin textural variations are minimal during formative years (but for facial hair during teenage years).
- Modeling age progression in young faces essentially implies modeling the flow of facial features across different ages.

# Craniofacial growth model – young faces

- **Challenges :**
  - Facial growth depends on factors such as gender, ethnicity, age group etc.
  - Facial features grow at different rates during different ages : During infancy and during adolescence, growth spurts are observed over different facial features.
- **Previous work :**
  - Researchers from psychophysics, studied craniofacial growth as a result of internal forces acting on the human cranium.
  - Cardioidal strain, spiral strain, affine shear etc. are some of the transformations that were applied on infant faces (profile views) to study age transformation effects.
- N. Ramanathan and R. Chellappa, CVPR 2006.

# Craniofacial growth model

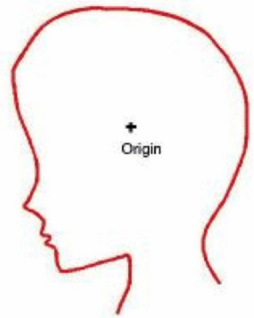
### Affine Shear

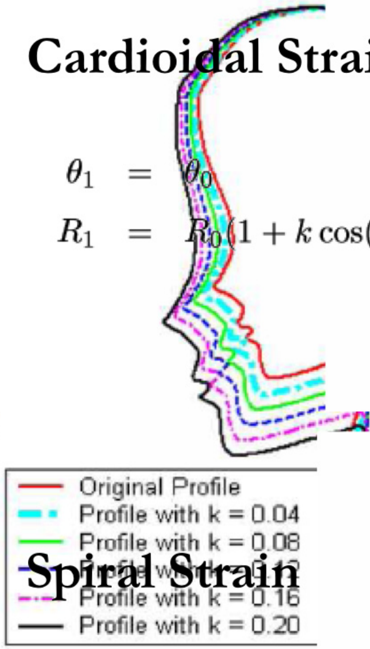



$$
\begin{aligned}
y_1 &= y_0 \\
x_1 &= x_0 + k y_0
\end{aligned}
$$

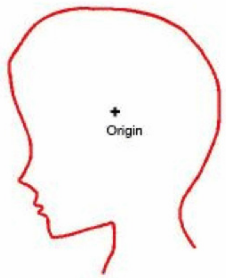
### Cardioidal Strain

$$
\begin{aligned}
\theta_1 &= \theta_0 \\
R_1 &= R_0(1 + k\cos(\theta_0))
\end{aligned}
$$



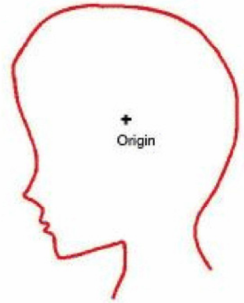

### Revised Cardioidal Strain




$$
\begin{aligned}
R_1 &= R_0 + k(R_0 - R_0\cos(\theta_0)) \\
\theta_1 &= \theta_0
\end{aligned}
$$

### Spiral Strain

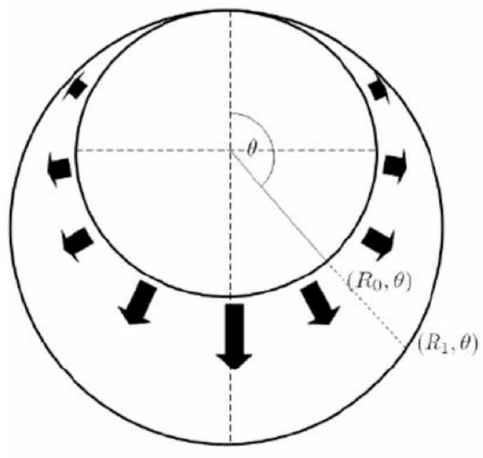
$$
\begin{aligned}
\theta_1 &= \theta_0 \\
R_1 &= R_0(1 + k|(\theta_0)|)
\end{aligned}
$$




Transformations induced by the revised cardioidal strain model reflected growth-related transformations best.

# Craniofacial growth model

Remodeling of Human head with growth is considered analogous to the remodeling a fluid-filled spherical tank with pressure (Mark et al 1980)




$Pressure \propto R_0(1 - cos(\theta))$

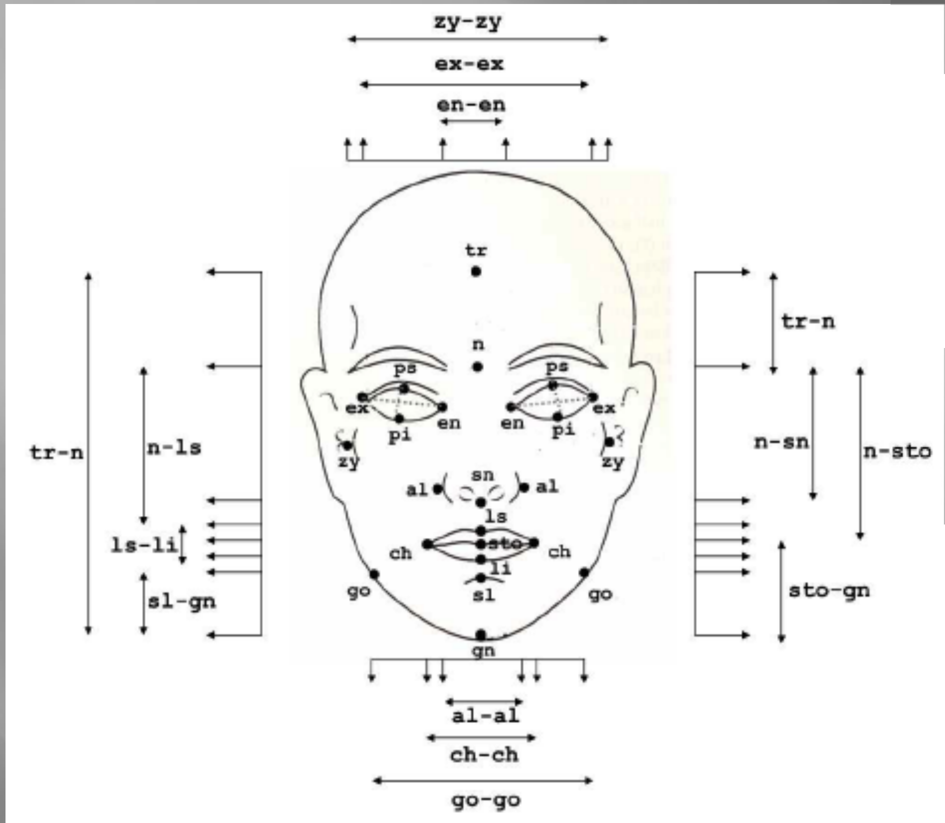
$$
\begin{aligned}
P &\propto R_0(1 - cos(\theta_0)) \\
R_1 &= R_0 + k(R_0 - R_0 \cos(\theta_0)) \\
\theta_1 &= \theta_0
\end{aligned}
$$

k : analogous to the growth parameter

Face anthropometric studies provide measurements extracted across different facial landmarks over different ages (0 to 18 years), thus characterizing facial growth during formative years.

# Craniofacial growth model






Age based facial measurements (compiled in Farkas1994) extracted across 24 facial landmarks were used

Proportion indices (ratios of anthropometric distances) to study facial growth. Eg:

- facial index (n_gn/zy_zy)
- intercanthal index (en_en/ex_ex)

# Craniofacial growth model

- Facial indices (ratios of face anthropometric measurements) can be represented as functions of facial growth parameters.

$$\left[\frac{(n\_gn)_{t_1}}{(zy\_zy)_{t_1}} = c_{t_1}\right] \Rightarrow \left[\frac{(R_{gn})_{t_1} - (R_n)_{t_1}}{2 \times (R_{zy})_{t_1} \times \cos(\theta_{zy})} = c_{t_1}\right] \Rightarrow$$

$$\left[\frac{(R_{gn})_{t_0}(1 + k_{gn}(1 - \cos(\theta_{gn}))) - (R_n)_{t_0}(1 + k_n(1 - \cos(\theta_n)))}{(R_{zy})_{t_0}(1 + k_{zy}(1 - \cos(\theta_{zy})))} = c_{t_1}\right]$$

$$\Rightarrow \left[\alpha_1 k_{gn} + \alpha_2 k_n + \alpha_3 k_{zy} = \alpha_4\right]$$

Growth parameters

# Craniofacial growth model

52 such proportion indices are identified. They result in linear and non-linear equations on facial growth parameters.

$$r_1 : \left[\frac{n-gn}{zy-zy} = c_1\right] \equiv \ \alpha_1^{(1)} k_1 + \alpha_2^{(1)} k_7 + \alpha_3^{(1)} k_{12} = \beta_1$$

$$r_2 : \left[\frac{al-al}{ch-ch} = c_2\right] \equiv \ \alpha_1^{(2)} k_{13} + \alpha_2^{(2)} k_{14} = \beta_2$$

$$r_3 : \left[\frac{li-sl}{sto-sl} = c_3\right] \equiv \ \alpha_1^{(3)} k_4 + \alpha_2^{(3)} k_5 + \alpha_3^{(3)} k_6 = \beta_3$$

$$r_4 : \left[\frac{sto-gn}{gn-zy} = c_4\right] \equiv \alpha_1^{(4)} k_5 + \alpha_2^{(4)} k_7 + \alpha_3^{(4)} k_{12} + \alpha_4^{(4)} k_4^2 + \alpha_5^{(4)} k_7^2 + \alpha_6^{(4)} k_{12}^2 + \alpha_7^{(4)} k_4\, k_7 + \alpha_8^{(4)} k_7\, k_{12} = \beta_4$$

$[k_1,\ k_2,\ \cdots\ k_{15}]$ $\longrightarrow$ Age-based growth parameters

$c_i$ $\longrightarrow$ Ratios of age-based facial measurements

Vision Lab

# Craniofacial growth model

$$f(\mathbf{k}) = \frac{1}{2}\sum_{i=1}^{N}(r_i(\mathbf{k}) - \beta_i)^2$$

$$\mathbf{k}_{i+1} = \mathbf{k}_i - (\mathbf{H} + \lambda diag[\mathbf{H}])^{-1}\nabla f(\mathbf{k}_i)$$

The computation of the growth parameters is formulated as that of solving a non-linear optimization.

We use Levenberg Marquartd optimization to solve for the age-based growth parameters defined over facial landmarks

$$f(\mathbf{x}_i) = k_i \qquad i = 1, \ldots, n$$

$$\mathbf{E} = \int\int_{\Omega} f_{xx}^2(\mathbf{x}) + 2f_{xy}^2(\mathbf{x}) + f_{xx}^2(\mathbf{x})d\mathbf{x}$$

$$f(\mathbf{x}) = p(\mathbf{x}) + \sum_{i=1}^{n}\lambda_i\phi(|\mathbf{x} - \mathbf{x}_i|)$$

$$\phi(\mathbf{x}) = |\mathbf{x}|^2\log(|\mathbf{x}|)$$

Using thin-plate spline formulations, we compute growth parameters over other facial regions

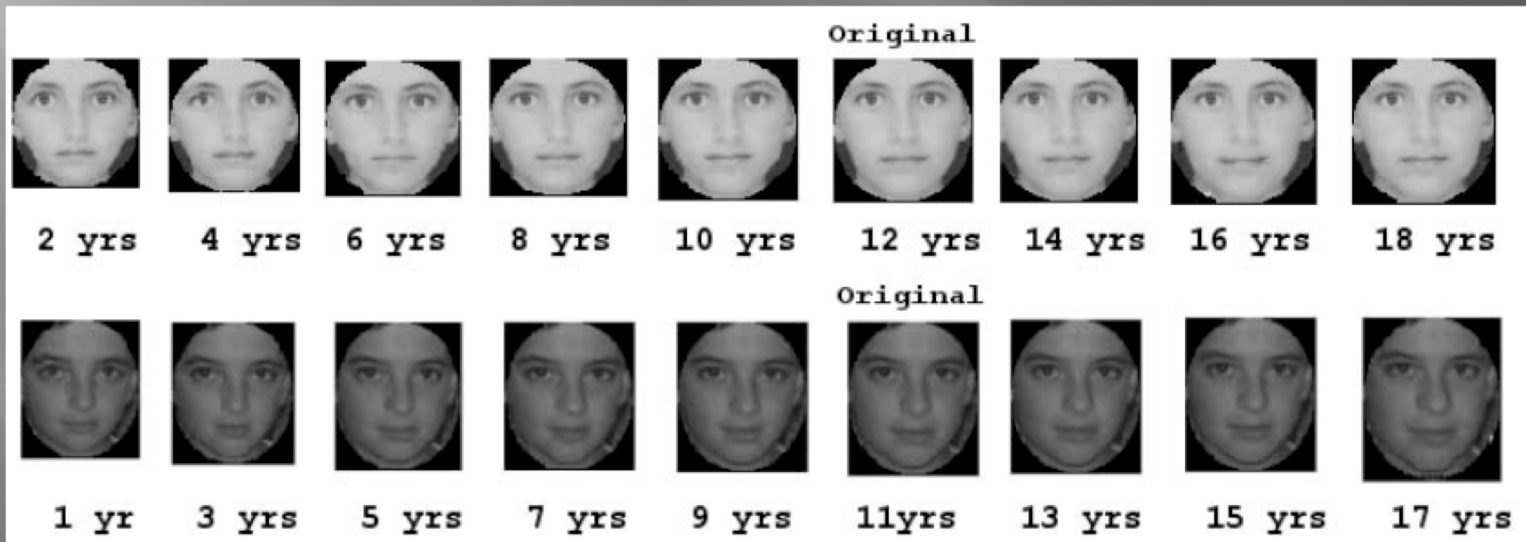
# Age transformation results

$$\begin{aligned} R_q^i &= R_p^i(1 + k_{pq}^i(1 - \cos(\theta_p^i))) \\ \theta_q^i &= \theta_p^i \end{aligned}$$

Transformation from 'p' years to 'q' years : q > p

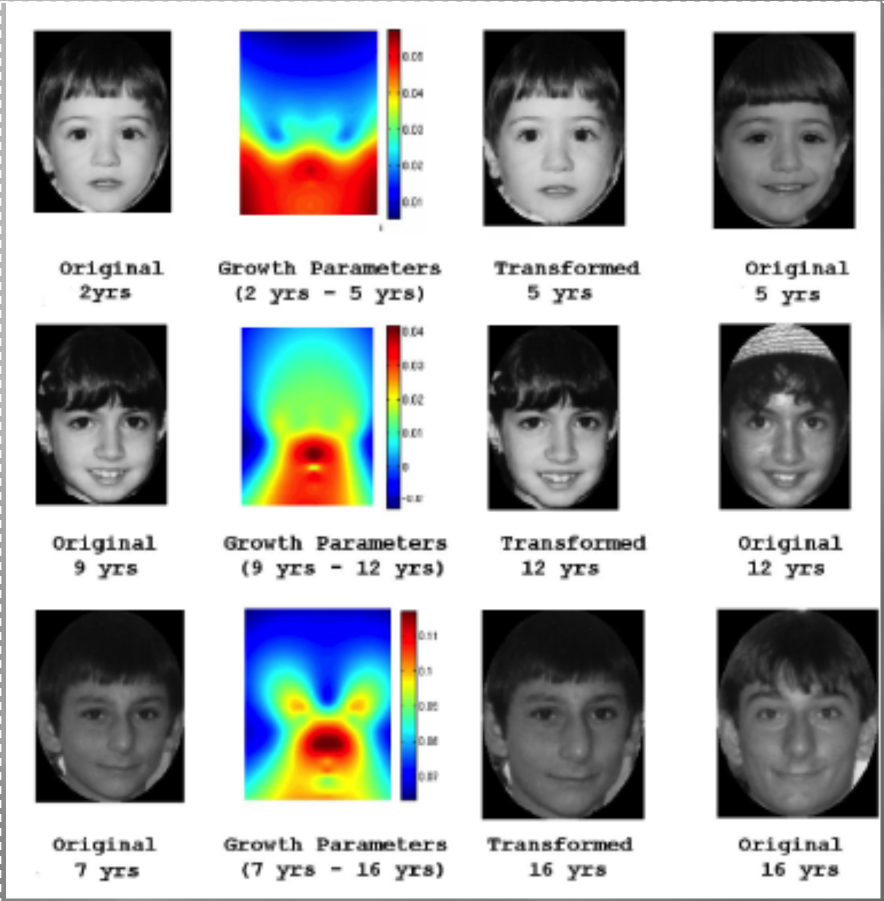
$$\begin{aligned} R_q^i &= \frac{R_p^i}{1 + k_{qp}^i(1 - \cos(\theta_q^i))} \\ \theta_q^i &= \theta_p^i \end{aligned}$$

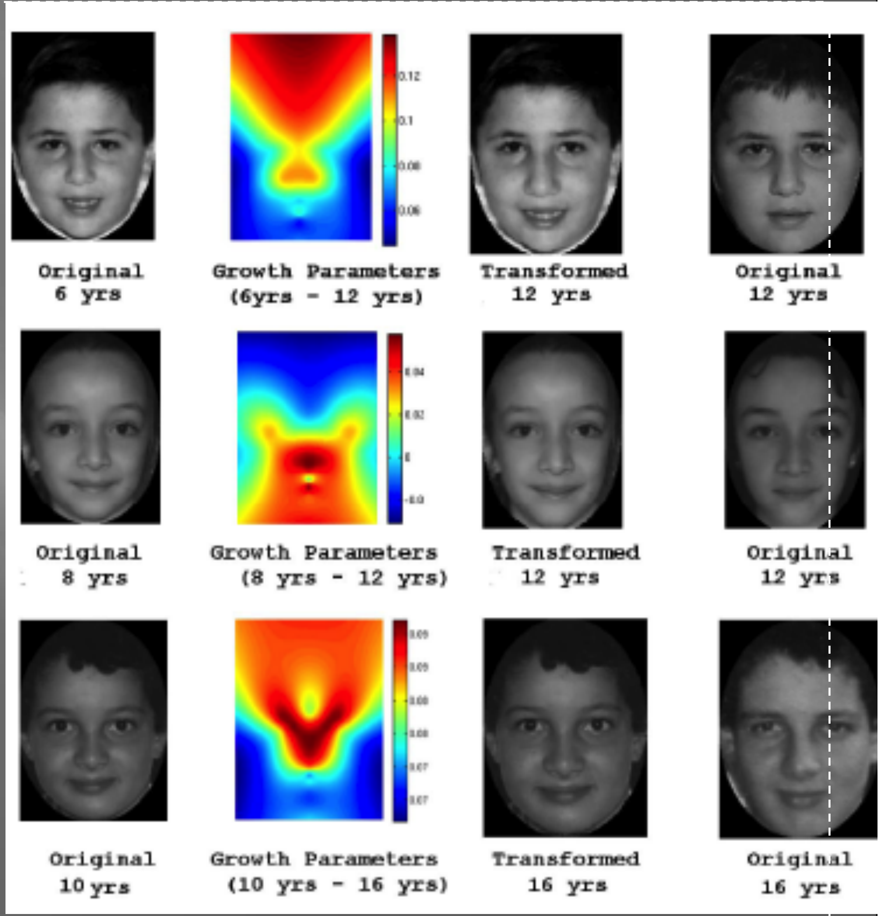
Transformation from 'p' years to 'q' years : q < p



# Age transformation results

Vision Lab



# Face recognition across age progression

- On a database of 233 images of 109 individuals (a few individuals with multiple age separated images), a face recognition experiment (eigenfaces)

- For each probe image (age known apriori), the gallery images are transformed before performing face recognition.

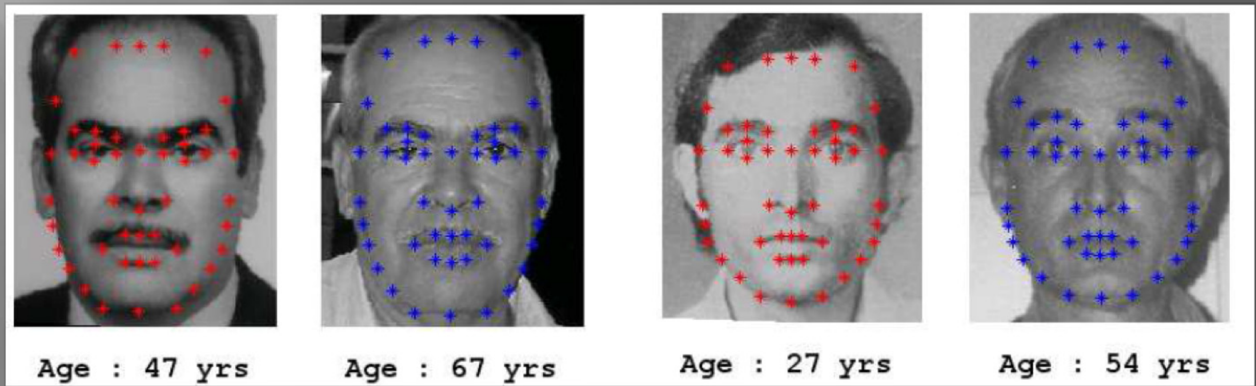
| Approach | Rank 1 | Rank 5 | Rank 10 |
|---|---|---|---|
| No transformation | 8 | 28 | 44 |
| Age transformed | 15 | 37 | 58 |

# Modeling age progression in adult faces

- **Reasons attributed to the appearance of wrinkles and other skin artifacts on faces are**
  - Loss of elasticity of facial muscles & skin : This results in the sagging of facial features
  - Loss of facial fat also causes furrows and wrinkles
  - Repetitive facial expressions
  - Overexposure to sun's rays, smoking etc.

# Modeling age progression in adult faces

- Two fold approach towards modeling facial aging in adults:
  - Shape transformation model: Physically based parametric muscle model that captures the growth of facial features with age.
  - Texture transformation model: An image gradient based texture transformation function that characterizes facial wrinkles.



# Shape transformation Model

- Shape transformation model assuming the form:

$$
\begin{aligned}
x_{t_1}^{(i)} &= x_{t_0}^{(i)} + k^{(i)} \left[P_{t_0}^{(i)}\right]_x \\
y_{t_1}^{(i)} &= y_{t_0}^{(i)} + k^{(i)} \left[P_{t_0}^{(i)}\right]_y
\end{aligned}
$$

Applied pressure

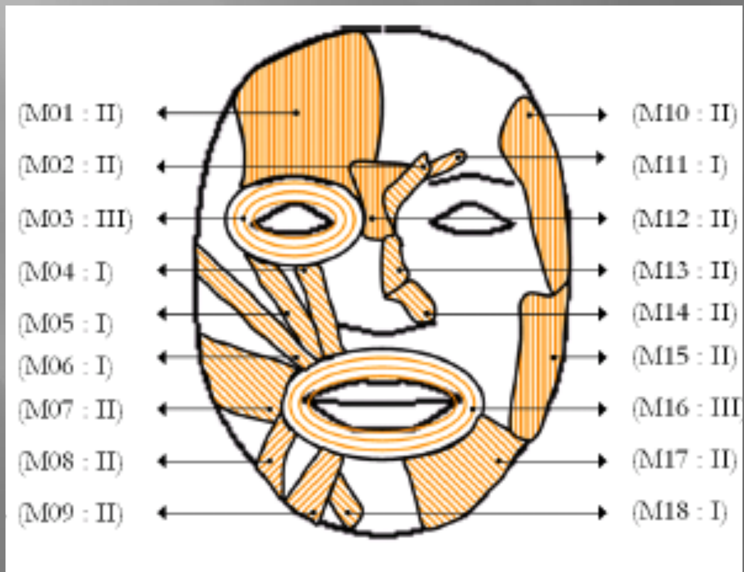
coordinates of the i'th facial feature at ages $t_0$ and $t_1$

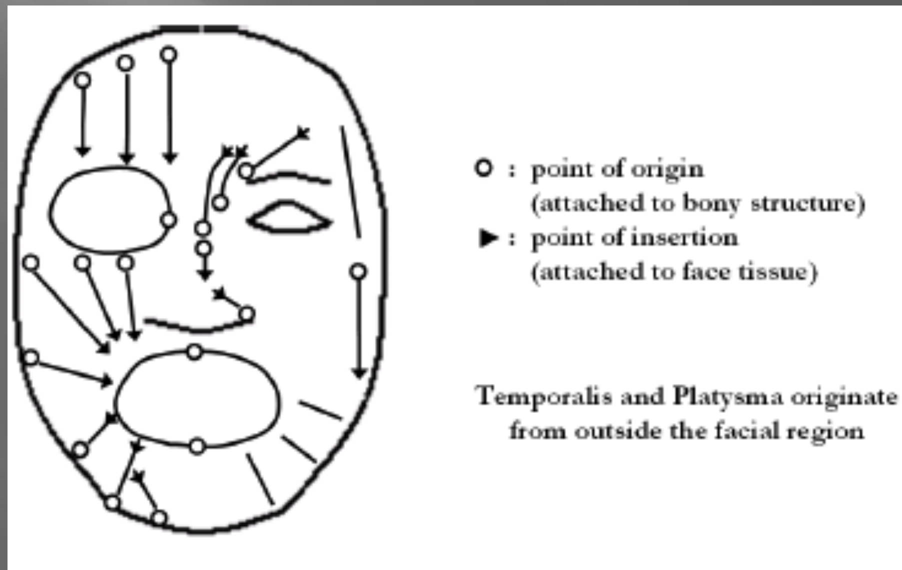
growth parameter

- Facial growth statistics :
  - Age groups in the study : *21 – 30 years, 31 – 40 years, 41 – 50 years, 51 – 60 years, 61 – 70 years.*
  - Identified 48 fiducial features on 100 pairs of age separated images in each age group and collected the facial measurements.

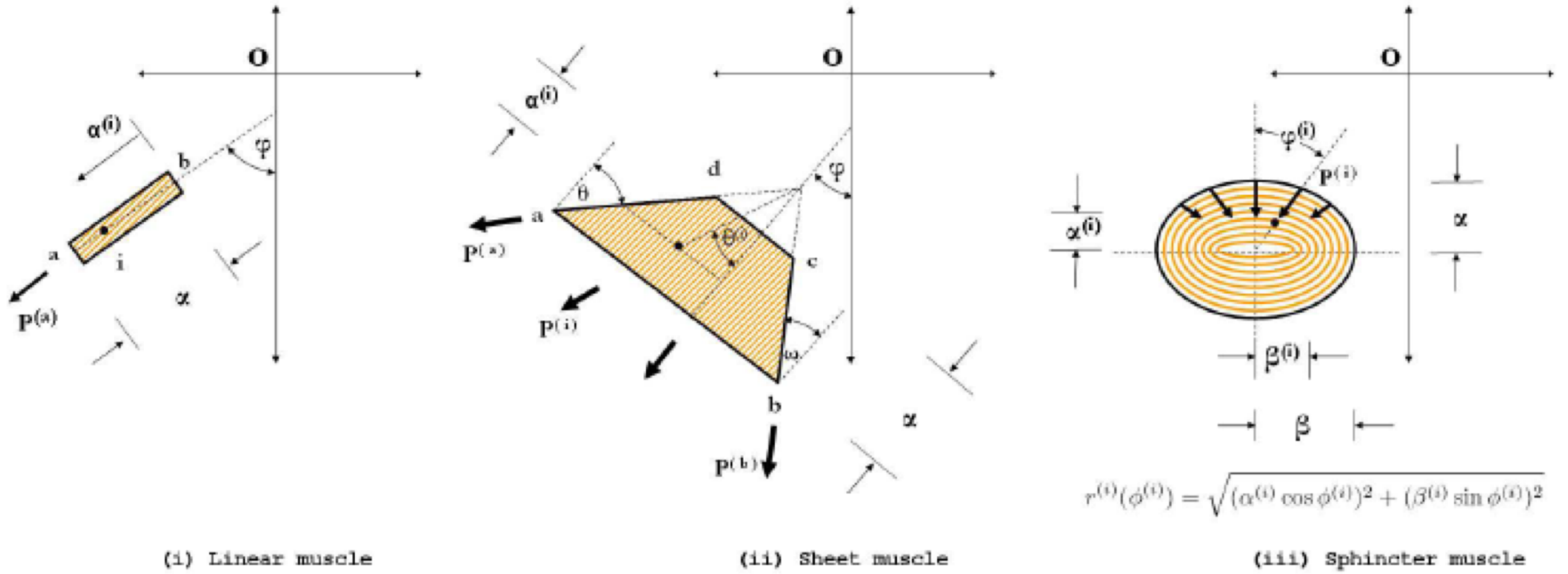
# Parametric muscle model

- Parametric muscle model for faces:
  - Muscles divided into 3 classes : Linear muscles, sheet muscles and sphincter muscles (based on their functionalities and geometric properties)
  - 18 facial muscles chosen



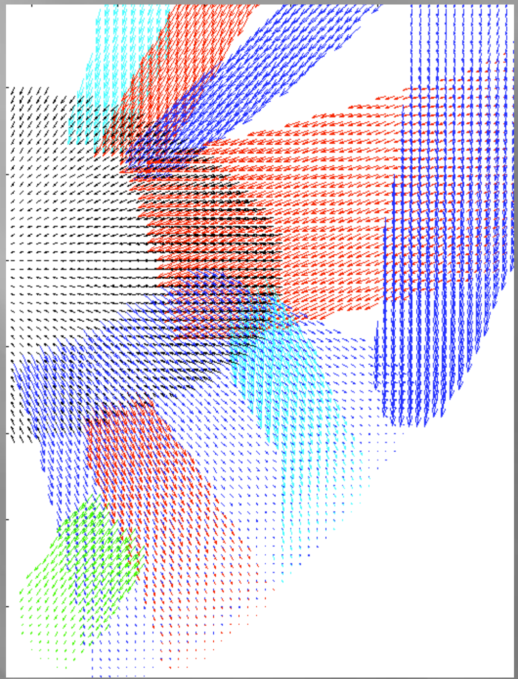
# Parametric muscle model




| Muscle type | Pressure model | Transformation model |
|---|---|---|
| Linear muscle($\{\alpha, \phi\}$) | $P^{(i)} \propto \alpha^{(i)}$ | $x_{t_1}^{(i)} = x_{t_0}^{(i)} + k\,[\alpha^{(i)} \sin\phi]$<br>$y_{t_1}^{(i)} = y_{t_0}^{(i)} + k\,[\alpha^{(i)} \cos\phi]$ |
| Sheet muscle($\{\alpha, \phi, \theta, \omega\}$) | $P^{(i)} \propto \alpha^{(i)} \sec\theta^{(i)}$ | $x_{t_1}^{(i)} = x_{t_0}^{(i)} + k\,[\alpha^{(i)} \sec\theta^{(i)} \sin(\phi + \theta^{(i)})]$<br>$y_{t_1}^{(i)} = y_{t_0}^{(i)} + k\,[\alpha^{(i)} \sec\theta^{(i)} \cos(\phi + \theta^{(i)})]$ |
| Sphincter muscle($\{\alpha, \beta\}$) | $P^{(i)} \propto r^{(i)}(\phi^{(i)}) \cos\phi^{(i)}$ | $x_{t_1}^{(i)} = x_{t_0}^{(i)} + k\,[r^{(i)}(\phi^{(i)}) \cos^2\phi^{(i)}]$<br>$y_{t_1}^{(i)} = y_{t_0}^{(i)} + k\,[r^{(i)}(\phi^{(i)}) \cos\phi^{(i)} \sin\phi^{(i)}]$ |

# Computational aspects

- In order to account for the confluence of different types of facial muscles over different facial features, for instance:



We propose the following shape transformation model

$$\bar{x}_{t_1}^{(i)} = \bar{x}_{t_0}^{(i)} + \sum_{j=1}^{n} \left[\boxed{k_{t_0 t_1}(m_j)}\ \xi_{t_0}^{(i)}(m_j)\right]$$

$$\bar{y}_{t_1}^{(i)} = \bar{y}_{t_0}^{(i)} + \sum_{j=1}^{n} \left[\boxed{k_{t_0 t_1}(m_j)}\ \psi_{t_0}^{(i)}(m_j)\right]$$

Pressure applied on the i'th facial feature at age $t_0$

The objective is to compute the muscle parameters that result in facial shape variations with age

# Computational aspects

- An example :

$$\left[\frac{d_{(15-16)}}{d_{(18-17)}}\right]_{t_1} = \left(\frac{\bar{y}_{t_1}^{(16)} - \bar{y}_{t_1}^{(15)}}{\bar{x}_{t_1}^{(17)} - \bar{x}_{t_1}^{(18)}}\right) = c_{t_1}$$

$$\Rightarrow 2 \times [\bar{y}_{t_0}^{(16)} + k_{t_0 t_1}(5)\, \psi_{t_0}^{(16)}(5) + k_{t_0 t_1}(14)\, \psi_{t_0}^{(16)}(14)] = c_{t_1} \times [\, \bar{x}_{t_0}^{(17)} - \bar{x}_{t_0}^{(18)} + k_{t_0 t_1}(16)\, \xi_{t_0}^{(17)}(16) - k_{t_0 t_1}(11)\, \xi_{t_0}^{(18)}(11) - k_{t_0 t_1}(12)\, \xi_{t_0}^{(18)}(12)\, ]$$

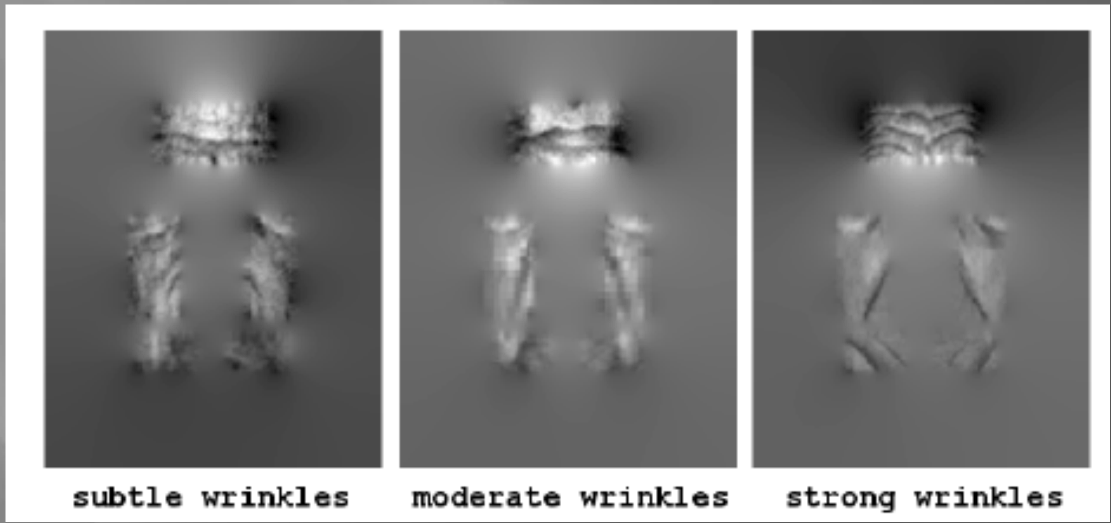
$$\Rightarrow \boxed{\lambda_5\, k_{t_0 t_1}(5) + \lambda_{14}\, k_{t_0 t_1}(14) + \lambda_{16}\, k_{t_0 t_1}(16) + \lambda_{11}\, k_{t_0 t_1}(11) + \lambda_{12}\, k_{t_0 t_1}(12) = \theta}$$

Linear equation in muscle parameters

946 proportion indices are taken into account in our study. The resulting linear equations in muscle parameters are solved using least squares approach.

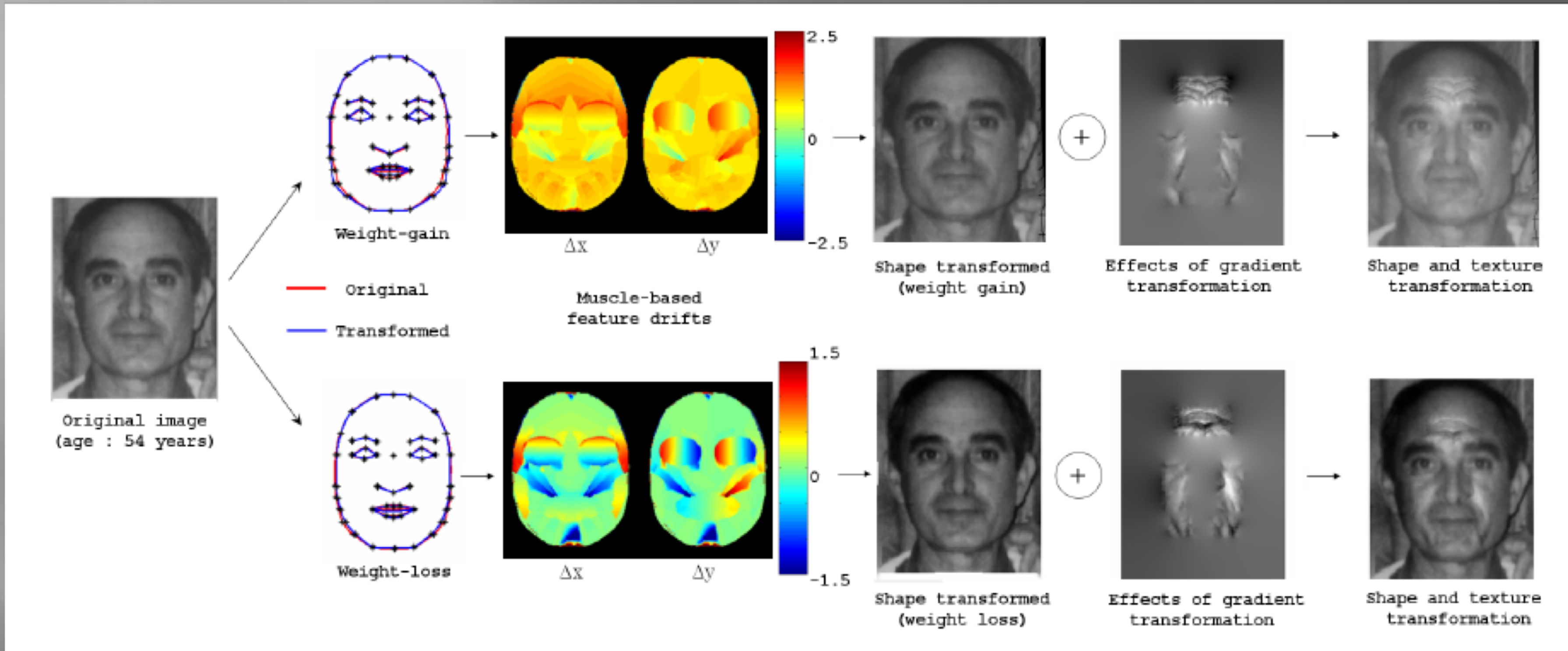
# Texture variation model

- ‘region-based’ facial wrinkle patterns learned from age separated face images across different age groups.
- Wrinkles are characterized using image gradient transformations



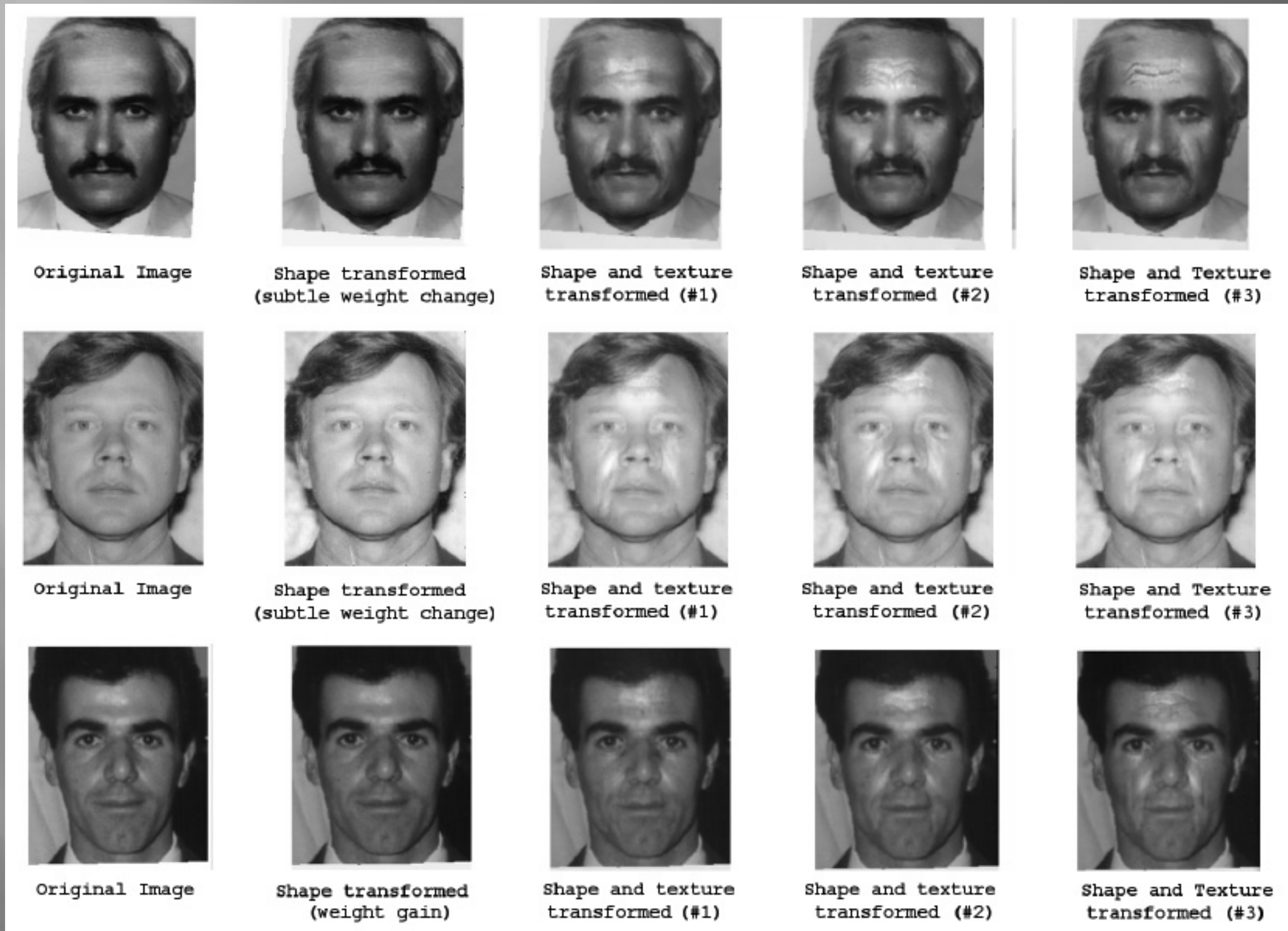

$$\nabla J_{t_2} = \nabla J_{t_1} + \frac{1}{N} \sum_{i=1}^{N} \sum_{n=1}^{4} W_n \cdot ||\nabla I_{t_2}^{(i)} - \nabla I_{t_1}^{(i)}||$$

# Texture variation model

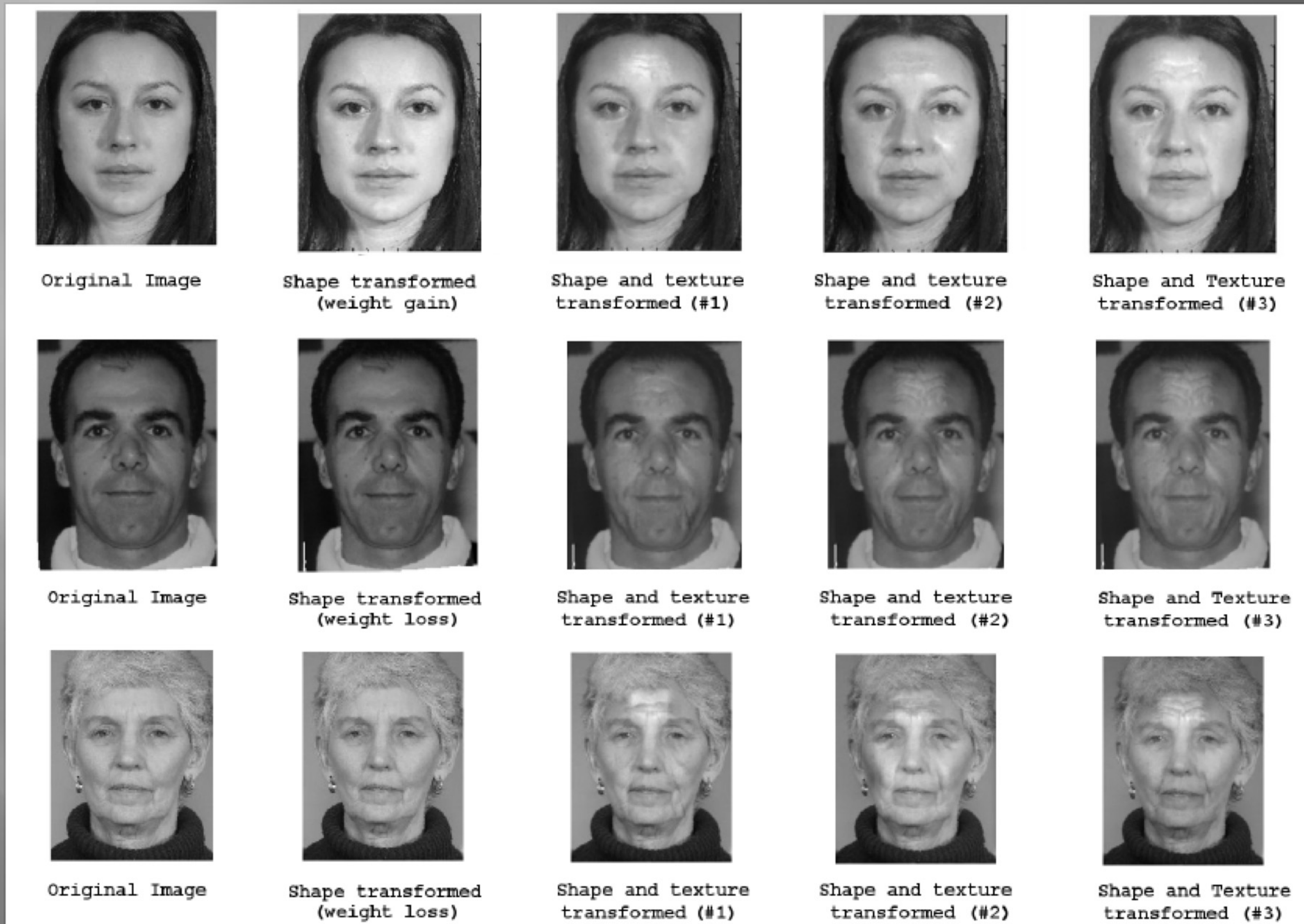


Vision Lab

# Results



# Results


Original Image
Shape transformed (weight gain)
Shape and texture transformed (#1)
Shape and texture transformed (#2)
Shape and Texture transformed (#3)
Original Image
Shape transformed (weight loss)
Shape and texture transformed (#1)
Shape and texture transformed (#2)
Shape and Texture transformed (#3)
Original Image
Shape transformed (weight loss)
Shape and texture transformed (#1)
Shape and texture transformed (#2)
Shape and Texture transformed (#3)

# DATA QUALITY

# Face quality

- The quality of a face biometric sample is strictly related to:
  - The quality of the sampled signal
  - The acquisition procedure
  - The environmental conditions
- These produce two different measures:
  - Related to the 2D photometric structure of the image
  - Related to the acquisition set-up
  - Related to the image pre-processing

# Face quality

Omolara Fatukasi, Josef Kittler, and Norman Poh
"Quality Controlled Multimodal Fusion of Biometric Experts"
Progress in Pattern Recognition, Image Analysis and Applications
pp 881-890 ,LNCS 4756/2008

## QUALITY MEASURES

- Left and Right eye coordinate (x,y) of the original image (extracted from video)
- Reliability of the face detector- this is the output of a classifier that has been trained to give an overall measure of quality given the quality measures 4-16 below
- Brightness
- Contrast
- Focus - this quantifies the sharpness of an image
- Bit per pixel - it measures the colour resolution in terms of bits
- Spatial resolution - the number of pixels between eyes
- Illumination
- Uniform Background – measuring the variance of the background intensity
- Background Brightness – the average intensity of the background
- Reflection - or, specular reflection
- Glasses – face wearing glasses
- Rotation in Plane
- Rotation in Depth
- Frontalness - it measures how much a face image deviates from a typical mug-shot face image

# Face quality

- **Image-related (photometric) quality measures:**
  - Signal to Noise ratio
  - Contrast/Brightness levels
  - Image sharpness
  - Bits per pixel
  - Resolution
  - Specular reflections

**Xiufeng Gao, Stan Z. Li, Rong Liu, and Peiren Zhang**
**"Standardization of Face Image Sample Quality"**
**Advances in Biometrics**
**pp 242-251 ,LNCS 4642/2007**

# Face quality

- **Related to the acquisition set-up**
  - Face/head pose in 3D
  - Background color/brightness
  - Face "frontalness"
  - Illumination
  - Eye glasses
- **Related to image processing**
  - Reliability of the face detector
  - Reliability of the eye/nose/mouth detector
  - Strength of the extracted features

# Exploiting face quality

- **Image level**
  - Contrast enhancement;
  - Histogram equalization;
  - Image re-lighting;
  - Deconvolution;
  - Noise reduction filtering;
  - Color space conversion.

# Exploiting face quality

- **Feature level**
  - Quality-based feature selection;
  - Feature weighting.
- **Matching/Classification level**
  - Quality-driven matching;
  - Data pruning;
  - Expert reliability;
  - Template updating.

# FACE RECOGNITION PERFORMANCES

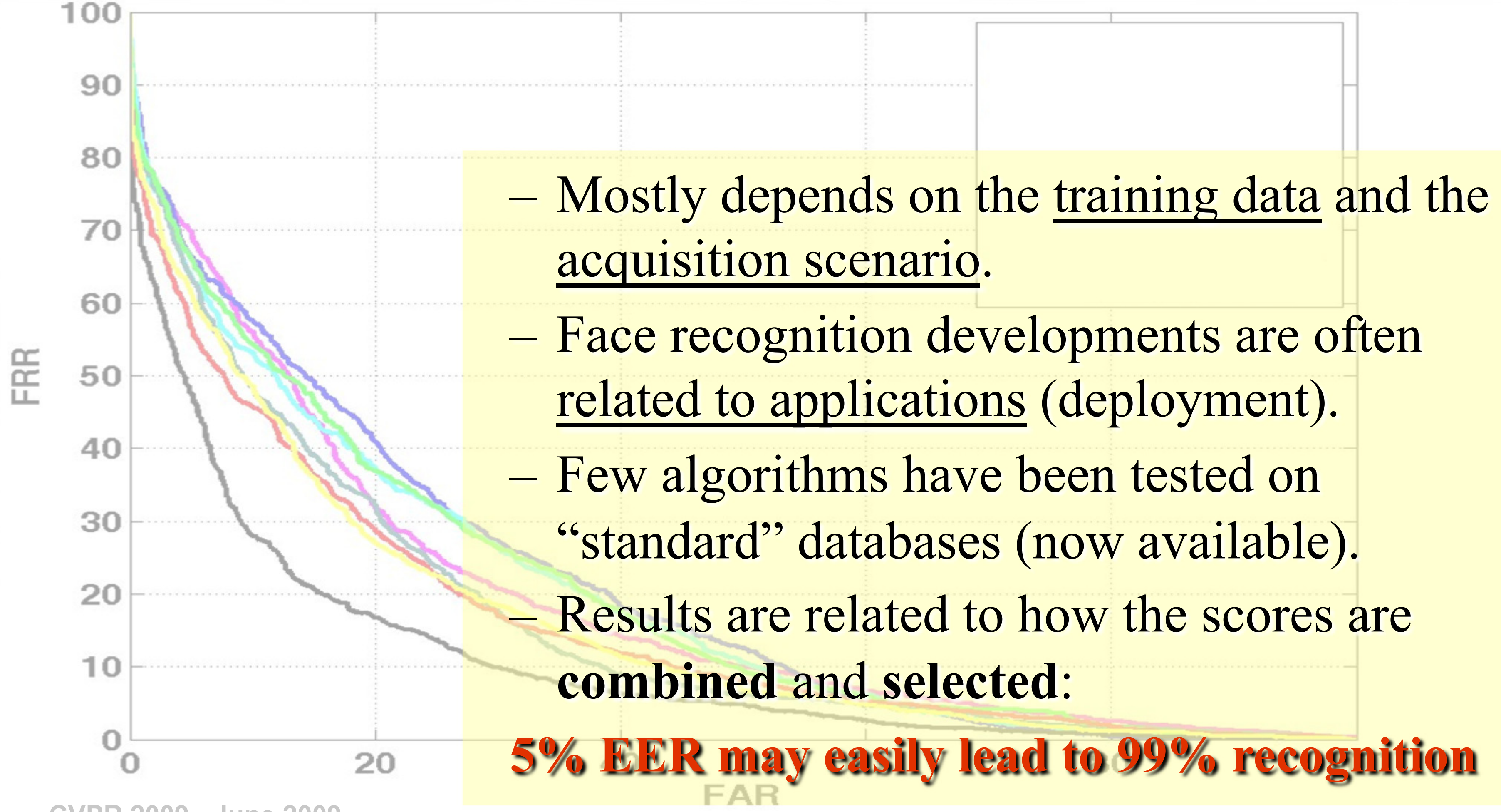
# Face recognition performances

Vision Lab

| Modality | Test Label | Test Parameter | False Reject Rate (FRR) | False Accept Rate (FAR) |
|---|---|---|---|---|
| **Fingerprint** | FpVTE 2003 | US Government operational data | 0.1% | 1% |
| **Fingerprint** | FVC 2006 | Heterogeneous population (young, elderly) | 2.2% | 2.2% |
| **Face** | FRGC 2006 | Controlled Illumination, high-resolution images | 0.8-1.6% | 0.1% |
| **Voice** | NIST 2004 | Text independent, multi-lingual | 5-10% | 2-5% |
| **Iris** | ITIRT 2005 | Indoor environment | 0.99% | 0.94% |
| **Iris** | ICE 2006 | Controlled Illumination, broad quality range | 1.1-1.4% | 0.1% |

# Face recognition performances

## This is a highly controversed subject

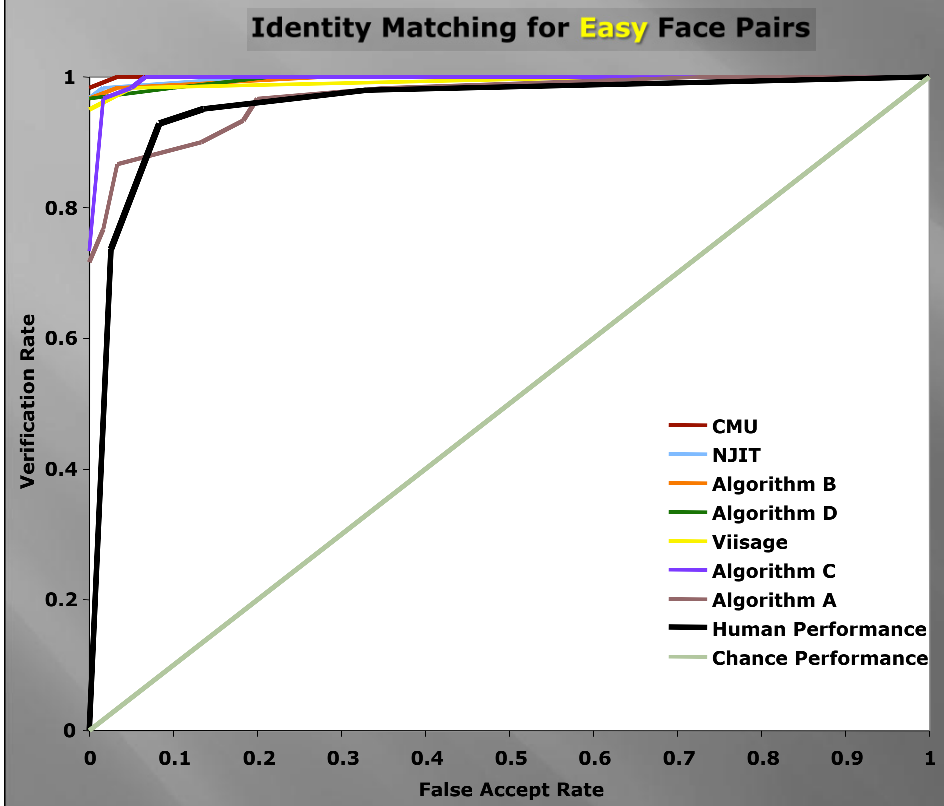


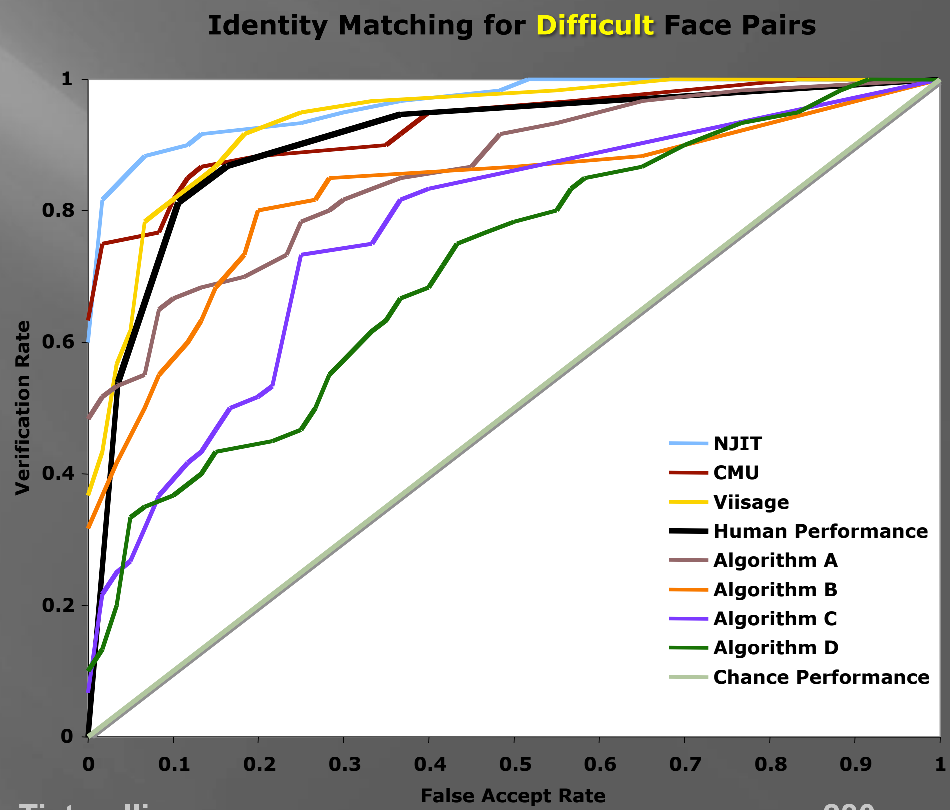

- Mostly depends on the training data and the acquisition scenario.
- Face recognition developments are often related to applications (deployment).
- Few algorithms have been tested on "standard" databases (now available).
- Results are related to how the scores are **combined** and **selected**:

**5% EER may easily lead to 99% recognition**

# Human vs Machine performances

A. O'Toole, J. Phillips, F. Jiang, Ayyad, Pénard & Abdi, "Benchmarking Algorithms Against Humans" *IEEE:T-PAMI, 2007* (in press)



# State-of-the-art

**Good marks:**

- Optimal classifiers (SVM/Bayesian)
- Advanced face-space representation (ICA/LFA)
- Best feature extraction methods (Gabor/LBP)

**Need for improvements:**

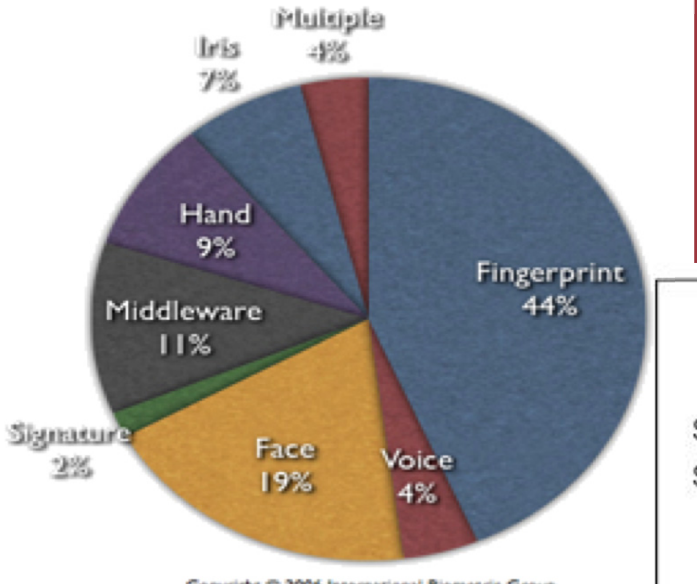
- Video vs mugshots
- Subject-based template
- Inaccurate registration
- Illumination dependent
- Feature selection

# Recent advances

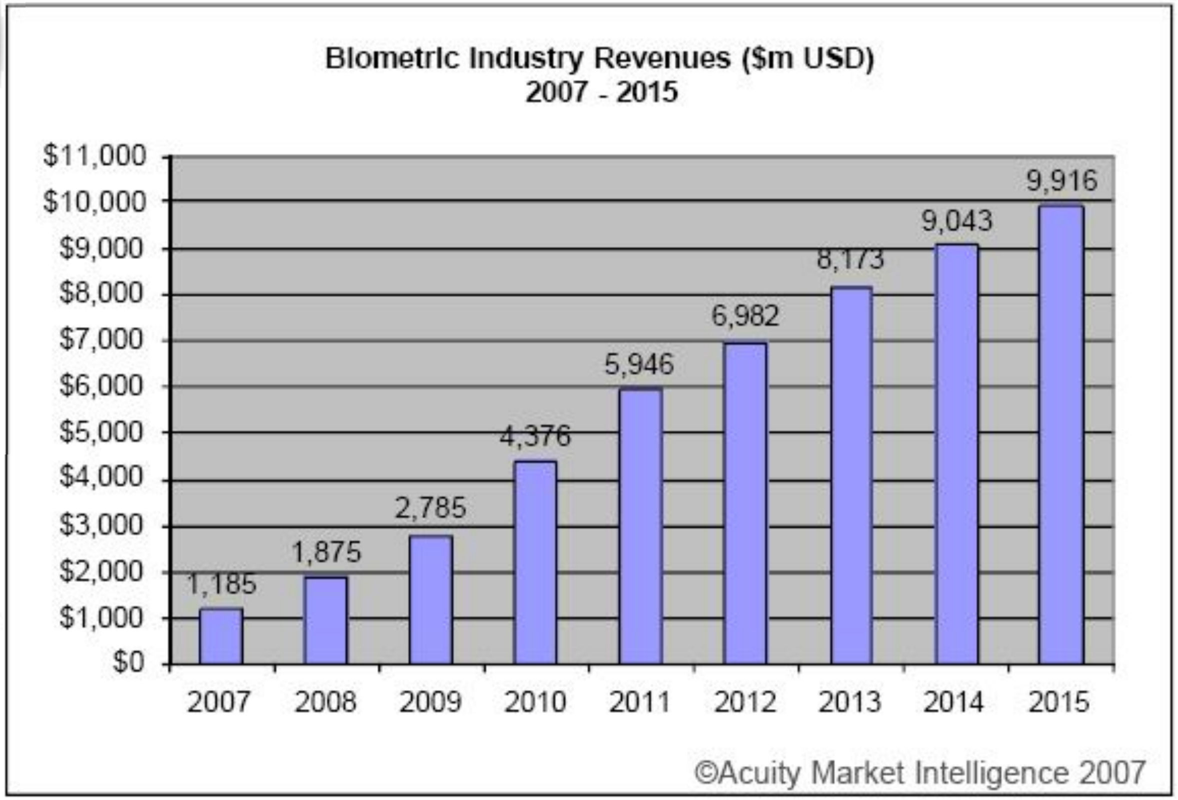
- Spoofing/Camouflage
- Face registration and Facial simmetry
- Face recognition from video sequences
- Subject-based template definition
- Ageing compensation
- Compensation of illumination
  - Multispectral imaging
  - Evaluation of illuminant components
  - Face appearance invariant models

# INDUSTRY STANDARDS AND COMMERCIAL SYSTEMS

# Biometrics market




Percentage of Biometric Market
by Technology for 2006



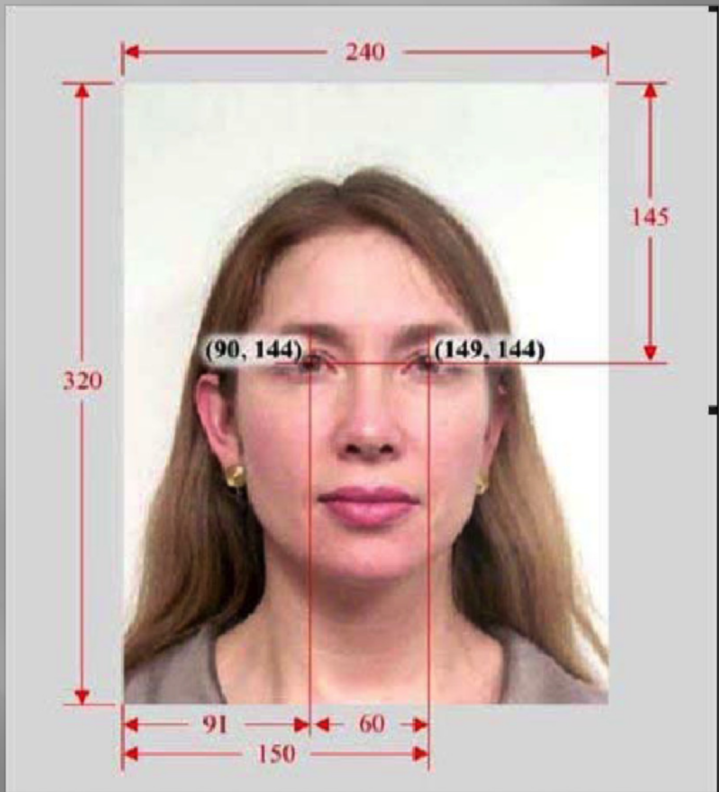
# Commercial FR systems

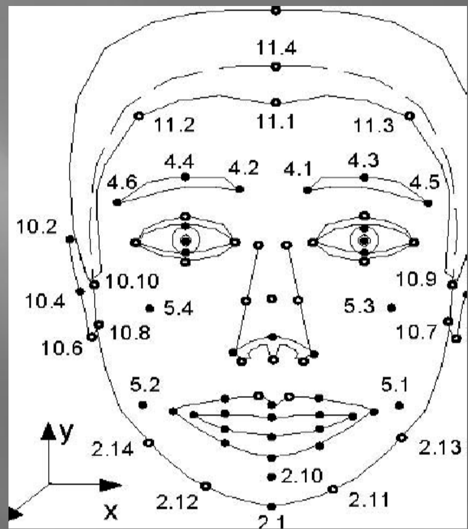
Vision Lab

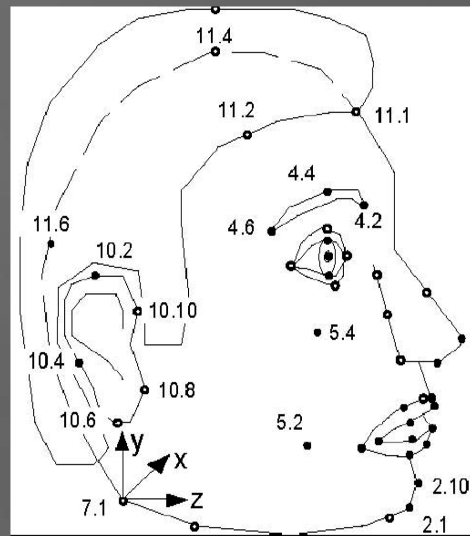
**From**: *www.face-rec.org*

- A4Vision, Inc. *(3D scanner)*
- AcSys Biometrics Corp.
- Animetrics Inc. *(3D shape)*
- C-VIS Computer Vision und Automation GmbH *(Neural Networks)*
- Cognitec Systems GmbH *(LFA)*
- Cybula Ltd. *(3D shape/2D texture)*
- DreamMirh Co., Ltd. *(2D ??? )*
- Geometrix, Inc. *(3D shape)*
- Iconquest *(2D Fractal-based ??? )*
- Identix Inc. *(LFA)*
- Imagis Technologies Inc. *(VISIPHOR Advanced Face Recognition ???)*
- Neven Vision, Inc. *(2D Gabor wavelets; feature-based )*
- Takumi Vision Technologies, Inc. *(Hi-Tech algorithms ??? )*
- Viisage *(Template matching ??? )*
- VisionSphere Technologies Inc. *(2D features with Holistic Feature Code)*

# The *standard* face



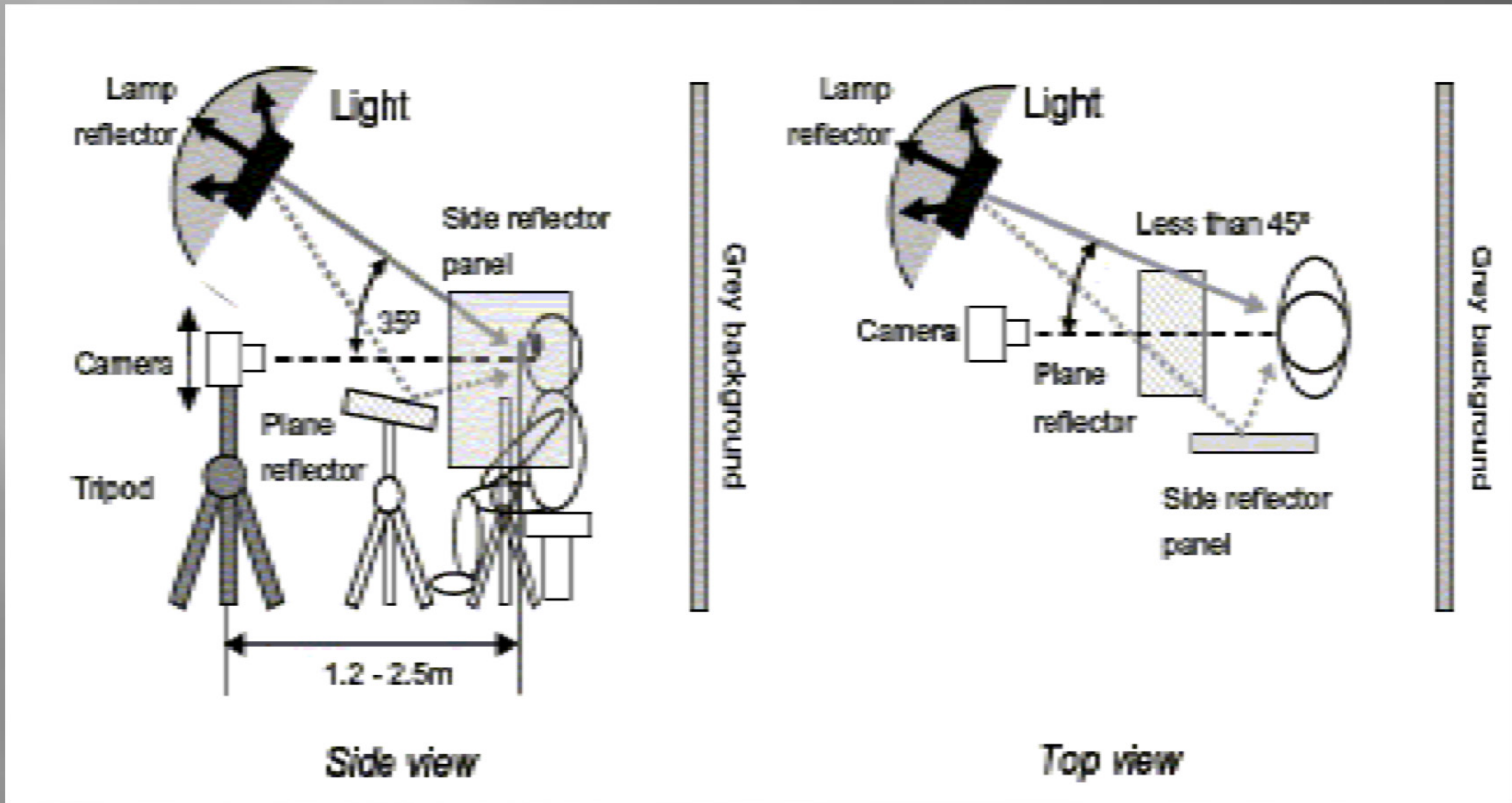




| Width(pixels) | Distance from Eye to Eye (Inclusive) |
|---|---|
| **240(min)** | **60** |
| 480 | 120 |
| 720 | 180 |

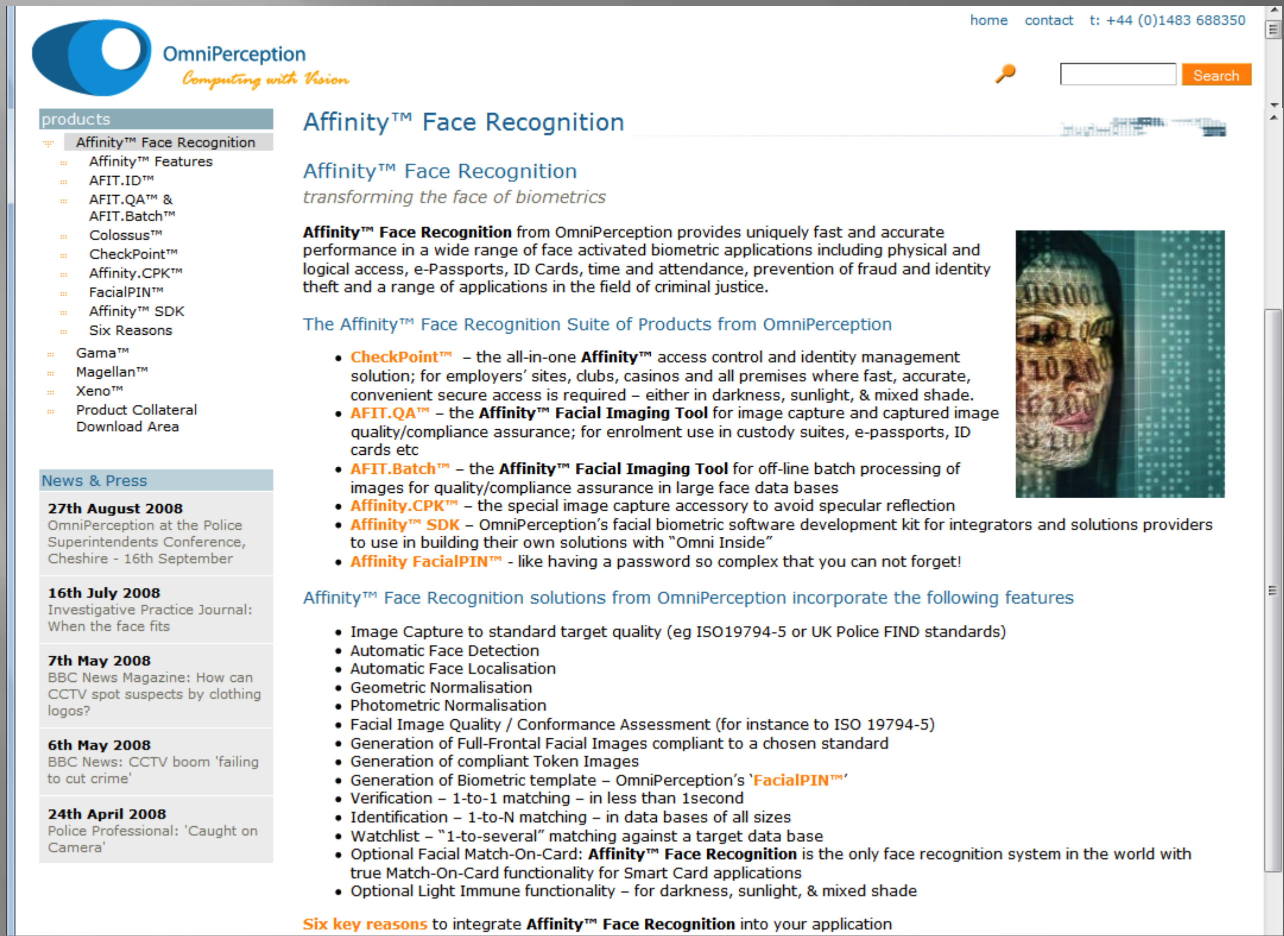
# The *standard* face

# The *standard* face



# Commercial FR systems

# Commercial FR systems



# Commercial FR systems

HOME | COMPANY | INVESTOR RELATIONS | NEWS & EVENTS | CONTACT US

L-1 IDENTITY SOLUTIONS

Protecting and Securing Personal Identities and Assets

SOLUTIONS | PRODUCTS | SERVICES | SUPPORT | PARTNERS | Search

BIOMETRICS
- Search By Biometric
  - Face
  - Fingerprint/Palm
  - Iris
  - Multi-Modal
- Civil Enrollment Systems
- Criminal Booking Systems
- Criminal Investigations
- Multi-Biometric Identification
- Enterprise Data Management
- Verification System Components
- Mobile Identification
- Inmate Management & Identification
- Facial Screening
- Biometric Network Logon
- Software Developer Kits
- Standards
- Associations

SECURE CREDENTIALING
ENTERPRISE ACCESS SOLUTIONS
ENTERPRISE ACCESS DIVISION

Face

L-1 Identity Solutions is a pioneer in the facial recognition arena. We have continually developed state of the art technologies that deliver superior capture and matching capabilities. In fact, L-1's facial recognition technology delivered top tier performance in the recent NIST-sponsored FRVT 2006 evaluation.

The following products use our top ranked technology, providing unique capabilities and unprecedented accuracy. Learn more by clicking on the links below.

| Product | Description | Product Category |
|---|---|---|
| ABIS® System | Scalable multi-biometric search engine that handles fingerprint, face and iris | Multi-Biometric Identification |
| ABIS® System FaceExaminer | A revolutionary, forensic application – unique to L-1 Identity Solutions – that helps analyze, search and identify faces of wanted subjects taken from surveillance video, where image quality is typically poor and subjects may not be directly facing the camera. Uses ABIS® System backend. | Criminal Investigations |
| FaceEXPLORER | Facial recognition enabled mugshot booking solution to help build a valuable library of facial images for search in the criminal identification process. | Criminal Booking Systems Inmate Management & Identification |
| FaceIt® ARGUS | Facial screening system that provides real time identification as subjects pass through or approach security checkpoints. | Facial Screening |
| FaceIt® Quality Assessment SDK | Face image quality evaluation and standards-based formatting tools that comply with industry best practices | Software Developer Kits |

# Commercial FR systems


GRIAULE biometrics

Contact us | English | ( Login )

Home | Products | Downloads | Resources | Partners | Company | Support | Buy

Google Custom Search | search

Home » ICAO Face SDK-Overview

## ICAO Face SDK

**ICAOFace** is an image processing Software Development Kit (SDK) that allows automatic identification systems to handle facial images according to international standards for systems like e-passport.

**ICAO Face SDK 2008** - Griaule launches the new version of the software!

What's new in this version? Click here and find out!

### Overview

Griaule ICAO Face SDK is an image processing Software Development Kit (SDK) that allows automatic identification systems handle facial images following international standards.

Digital facial images are widely used for systems like e-passport. The most common problem when dealing with photos is the heterogeneous quality due to the size, face orientation, color, contrast and other variables.

Griaule ICAO Face SDK processes a digital facial image to improve its quality conforming ICAO (International Civil Aviation Organization) requirements. ICAO is the leading organization on the effort to create standards for the emerging e-passport technology. It includes a standard for digital image quality for facial images, defining requirements for photographic and digital characteristics.

### Key Features

See full feature list

- Automated characteristics detection: Automated detection of eyes and mouth locations and option for manual adjustment;
- Aligned with standards: Several ICAO Full Frontal Image requirements (ISO/IEC FCD 19794-5) based on eyes and mouth detection;
- Biometric data interchange and interoperability available: Compliant with ISO/IEC 19785:1-2004 and others;
- Independence of camera: The software needs only the face image to be used;
- Easy and intuitive: Simple interface available as an ActiveX (COM) and DLL;
- Easy to learn: Lots of samples and the forum will help you integrating in a matter of hours!
- Use your preferred programming language: Wide range of supported programming languages;
- Safe to purchase: 90-days trial available for download.

NEXT STEPS

- Buy on-line
- Download
- Contact

- What's new in ICAO 2008?
- Overview
- Demo On-line
- Features
- Licensing and Editions
- System Requirements
- Manual

Vision Lab

# Near-infra red imaging

## Authentimetric F1 (CASIA-NLPR)

**Main features:**

- Near infra-red imaging
- Accurate eye localization with Gabor features
- LBP features and selection with Ada-boost
- “*smart* enrollment” with 5 snapshots
- Anti-spoofing system with liveness detection

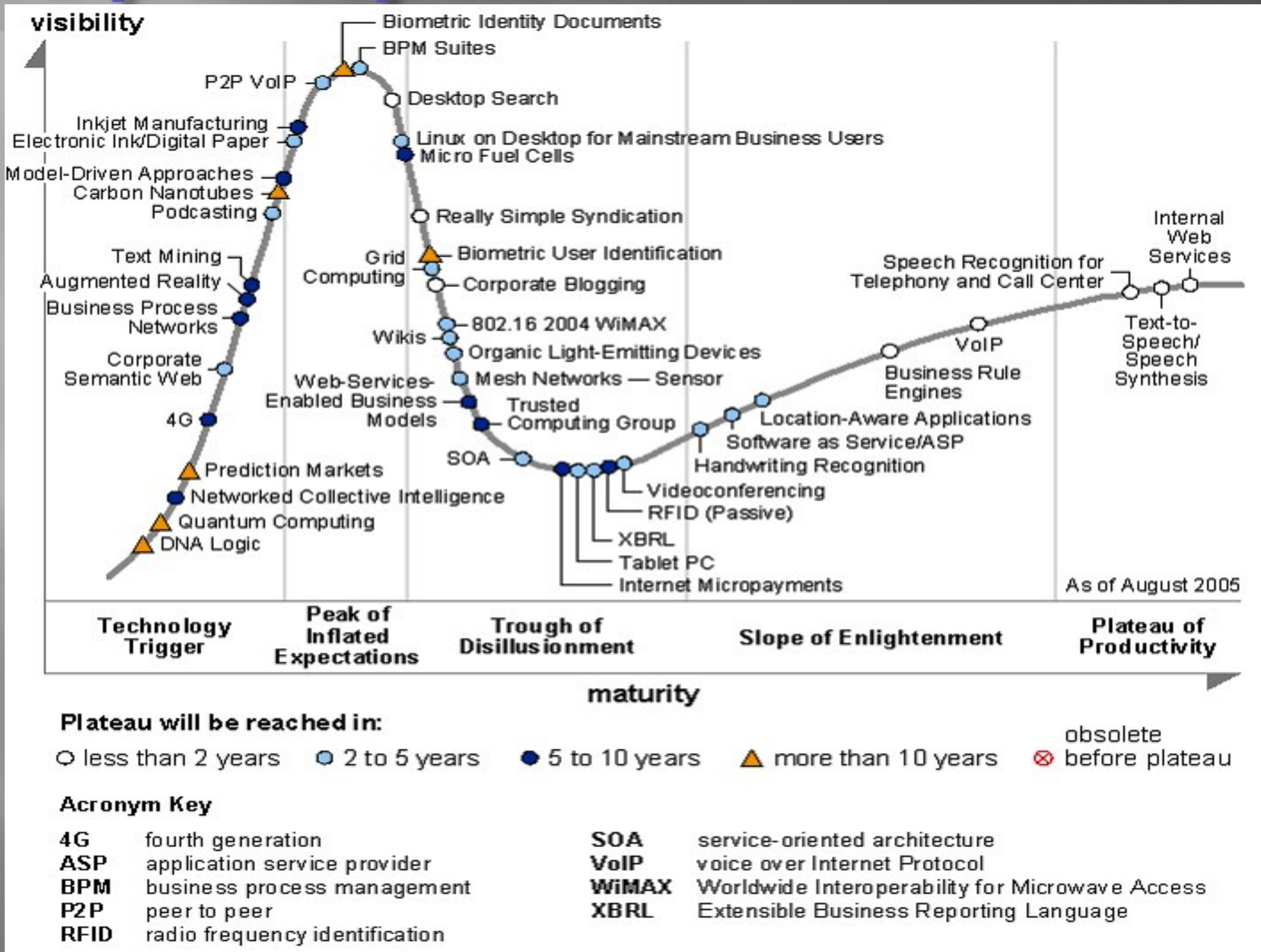
# Near-infra red imaging

Vision Lab

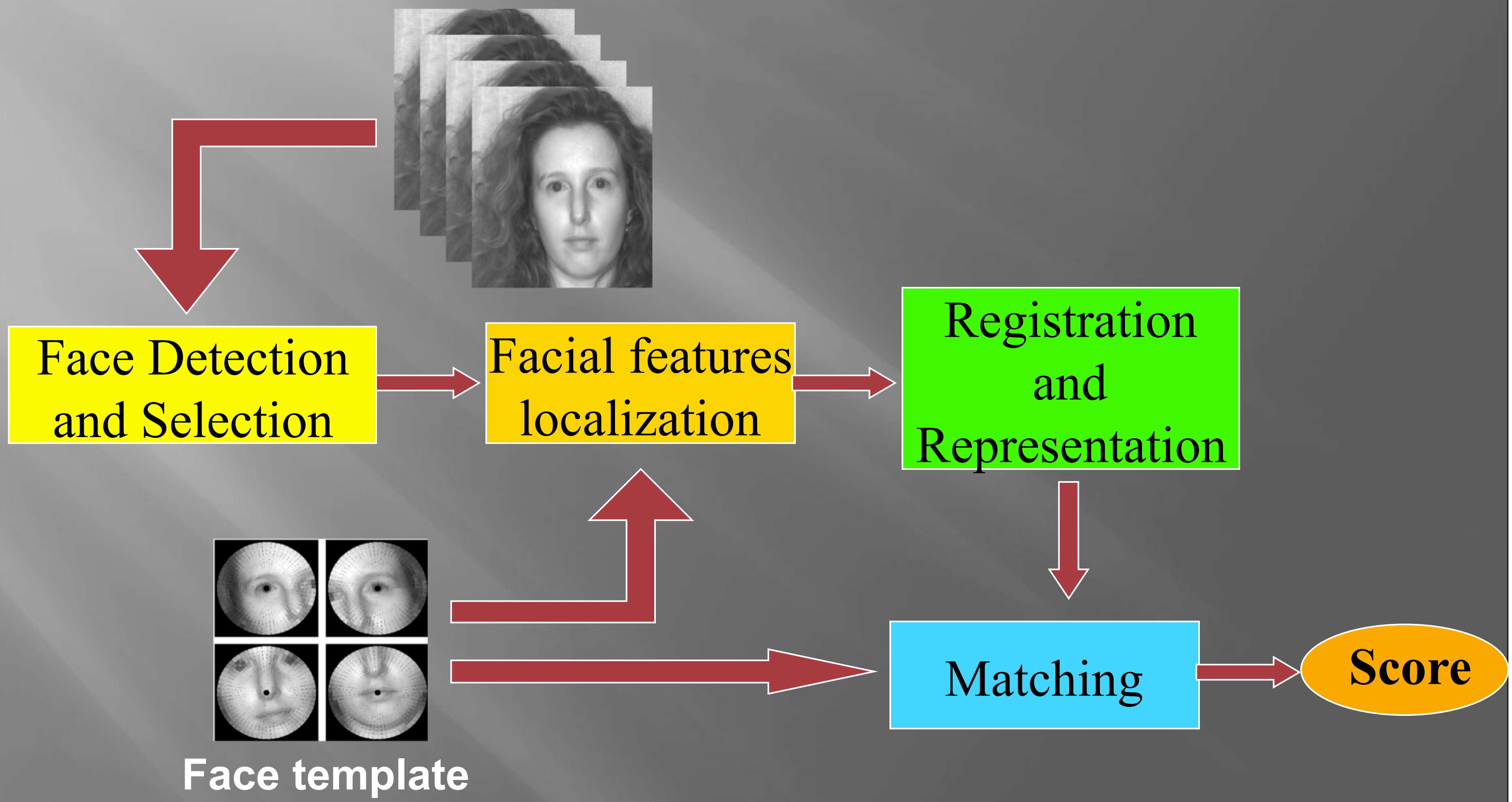
## Authentimetric F1

# MACHINES VS HUMANS

# The *perception* of Biometrics



# Face recognition operation



# Templates and standards

## Eurosmart white paper 2003

**Figure 2: Biometric Template Size**
*Source: Frost & Sullivan*

| Biometric | Bytes Required |
|---|---|
| Finger-scan | 300-1200 |
| Finger geometry | 14 |
| Hand geometry | 9 |
| Iris recognition | 512 |
| Voice verification | 1500 |
| Face recognition | 500-1000 |
| Signature verification | 500-1000 |
| Retina recognition | 96 |

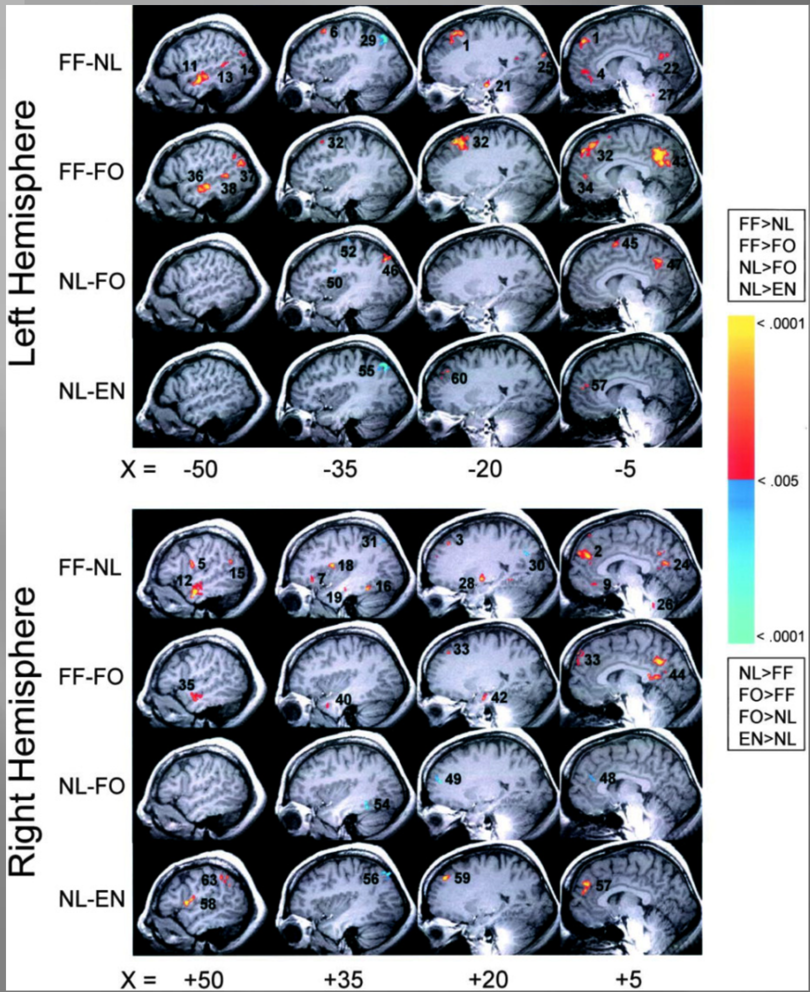
- *It is generally anticipated that any single biometric template would fit within 10 Kbytes of data memory in the storage device (That includes the template itself, the signature or encryption overhead and any additional data required to fulfill the data file structure).*

# Templates and standards…

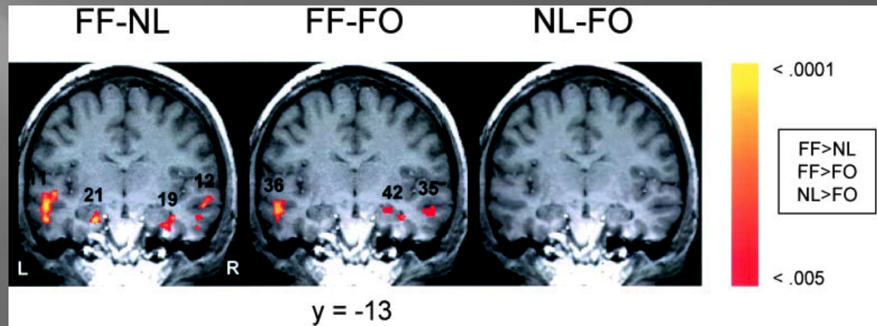
## Federal Computer Week March 2002

- *…And many templates can be stored on one system because each is less than 1K in size, compared to between 150K and 300K for a facial image.*
- ***Face-Id/Face-It** … The small template, at 88 bytes, is used for fast searches, while a larger, 3,518-byte image is used in conjunction with the small template for more thorough searches.*
- ***ID-2000** uses proprietary algorithms to transform the detection information into an encode array, or template. The template size is 532 bytes — small enough to put on a smart card —*
- ***UnMask** recognizes faces using a proprietary method … the resulting template, also called a face code… is less than 1K in size.*

# Templates in the Human Visual System




Recognition of 50 Familiar Faces (**FF**) vs 50 Newly Learned Faces (**NL**) and compared to rejection of 50 Foil (**FO** -False Objective) faces. Encoding (**EN**) session for learning new faces.




C. L. Leveroni et al. **"Neural Systems Underlying the Recognition of Familiar and Newly Learned Faces",** The Journal of Neuroscience, January 15, 2000, *20*(2):878–886

*Figure 2.* Areas of significantly increased (*red–yellow* scale) and decreased (*blue–cyan* scale) MR signal intensity from *t* tests ($p < 0.005$) comparing the three conditions: FF minus NL, FF minus FO, and NL minus FO. Numbers below each image represent millimeters from the interhemispheric fissure (−, left; +, right). Numbers adjacent to activated foci correspond to location numbers (first column) of Tables 1, 2, and 3.

# Face representation in the HVS

Vision Lab

- It is not clear how faces are represented in the HVS, but the representation is formed by a dynamically updated collection of visual fixations.
  - Both foveal and peripheral vision are involved, the former responsible for a "more accurate representation".

J.M. Henderson et al. "Eye movements are functional during face learning", *Memory & Cognition 2005, 33 (1), 98-106.*

D.R. Melmoth et al. "The Effect of Contrast and Size Scaling on Face Perception in Foveal and Extrafoveal Vision", *Investigative Ophthalmology and Visual Science.* 2000;41:2811-2819.

- This representation includes both iconic data as well as information about the spatial relationship among face elements.
  - **What about a "*face template*"?**

# Face representation in the HVS

Vision Lab

**Table 1. Famous faces (FF) vs newly learned (NL) faces**

| Loc. # | Brain region | BA | vol. (m | | | |
|---|---|---|---|---|---|---|
| | FF > NL | | | | | |
| | Frontal Lobe | | | | | |
| 1 | L Superior Frontal | 8 | 2.6 | | | |
| 2 | R Medial Frontal | 9 | 2.4 | | | |
| 3 | R Superior Frontal | 8 | 0.5 | | | |
| 4 | L Medial Frontal | 10 | 0.4 | | | |
| 5 | R Precentral | 6 | 0.4 | | | |
| 6 | L Superior Frontal | 8 | 0.4 | | | |
| 7 | R Inferior Frontal | 47 | 0.3 | | | |
| 8 | R Anterior Cingulate | 32 | 0.3 | | | |
| 9 | R Medial Frontal | 11 | 0.3 | | | |
| 10 | L Medial Frontal | 11 | 0.3 | | | |
| | Temporal Lobe | | | | | |
| 11 | L Middle Temporal | 21 | 2.7 | | | |
| 12 | R Middle Temporal | 21 | 1.9 | | | |
| 13 | L Middle Temporal | 21 | 0.6 | | | |
| 14 | L Middle Temporal | 39 | 0.5 | | | |
| 15 | R Superior Temporal | 22 | 0.5 | | | |
| 16 | R Fusiform | 20/37 | 0.4 | | | |
| 17 | R Middle Temporal | 37 | 0.3 | | | |
| 18 | R Insula | – | 0.3 | | | |
| 19 | R Parahippocampal | 35 | 0.2 | | | |
| 20 | R Parahippocampal | 36 | 0.2 | 24 | −43 | −7 |
| 21 | L Hippocampus | 28 | 0.2 | −19 | −12 | −20 |
| | Parietal/Occipital Lobe | | | | | |
| 22 | L Posterior Cingulate | 23/30 | 1.7 | −4 | −57 | 15 |
| 23 | R Inferior Parietal | 40 | 0.5 | 44 | −30 | 22 |
| 24 | R Posterior Cingulate | 31 | 0.3 | 2 | −57 | 29 |
| 25 | L Extrastriate | 18 | 0.3 | −20 | −89 | 20 |
| | Subcortical | | | | | |
| 26 | R Pons | – | 0.4 | 11 | −43 | −34 |
| 27 | L Pons | – | 0.2 | −10 | −43 | −33 |
| 28 | R Putamen | – | 0.3 | 22 | −7 | −6 |
| | NL > FF | | | | | |
| | Parietal Lobe | | | | | |
| 29 | L Inferior Parietal | 40 | 1.0 | −37 | −64 | 40 |
| 30 | R Superior Parietal | 7 | 0.5 | 23 | −66 | 30 |
| 31 | R Inferior Parietal | 40 | 0.3 | 35 | −67 | 42 |

Region is defined as center of mass. The first column refers to location numbers delineated in Figures 2 and 3 (italicized numbers indicate locations not shown in figures). Coordinates represent distance in millimeters from anterior commissure: $x$ right (+)/left (−); $y$ anterior (+)/posterior(−); $z$ superior (+)/inferior(−).

**BRAIN Total 1400 ml**
**100 billion neurons**
**71.5 Mneurons/ml**

**Maybe we can sketch the network size devoted to process faces….**

C. L. Leveroni et al. **"Neural Systems Underlying the Recognition of Familiar and Newly Learned Faces"**, The Journal of Neuroscience, January 15, 2000, *20*(2):878–886

# Face representation in the HVS

Vision Lab

aces (FF) vs newly learned (NL) faces

| Brain region | BA | vol. (ml) |
|---|---|---|
| FF > NL | | |
| Frontal Lobe | | |
| L Superior Frontal | 8 | 2.6 |
| R Medial Frontal | 9 | 2.4 |
| R Superior Frontal | 8 | 0.5 |
| L Medial Frontal | 10 | 0.4 |
| R Precentral | 6 | 0.4 |
| L Superior Frontal | 8 | 0.4 |
| R Inferior Frontal | 47 | 0.3 |
| R Anterior Cingulate | 32 | 0.3 |
| R Medial Frontal | 11 | 0.3 |
| L Medial Frontal | 11 | 0.3 |
| Temporal Lobe | | |
| L Middle Temporal | 21 | 2.7 |
| R Middle Temporal | 21 | 1.9 |
| L Middle Temporal | 21 | 0.6 |
| L Middle Temporal | 39 | 0.5 |
| R Superior Temporal | 22 | 0.5 |
| R Fusiform | 20/37 | 0.4 |
| R Middle Temporal | 37 | 0.3 |
| R Insula | – | 0.3 |
| R Parahippocampal | 35 | 0.2 |
| R Parahippocampal | 36 | 0.2 |
| L Hippocampus | 28 | 0.2 |
| Parietal/Occipital Lobe | | |
| L Posterior Cingulate | 23/30 | 1.7 |
| R Inferior Parietal | 40 | 0.5 |
| R Posterior Cingulate | 31 | 0.3 |
| L Extrastriate | 18 | 0.3 |
| Subcortical | | |
| R Pons | – | 0.4 |
| L Pons | – | 0.2 |
| R Putamen | – | 0.3 |
| NL > FF | | |
| Parietal Lobe | | |
| L Inferior Parietal | 40 | 1.0 |
| R Superior Parietal | 7 | 0.5 |
| R Inferior Parietal | 40 | 0.3 |

center of mass. The first column refers to location numbers demarcated in Figures 2 and
t distance in millimeters from anterior commissure: $x$ right (+)/left (−); $y$ anterior (+)/p

The BRAIN mass is equal to 1400 ml
Composed of some 100 billion neurons
71.5 Mneurons/ml

Summing up the volumes of all active areas, the total volume is 21,2 ml

**or … 1.5 Bneurons**

… with 12K Synapses/neuron!

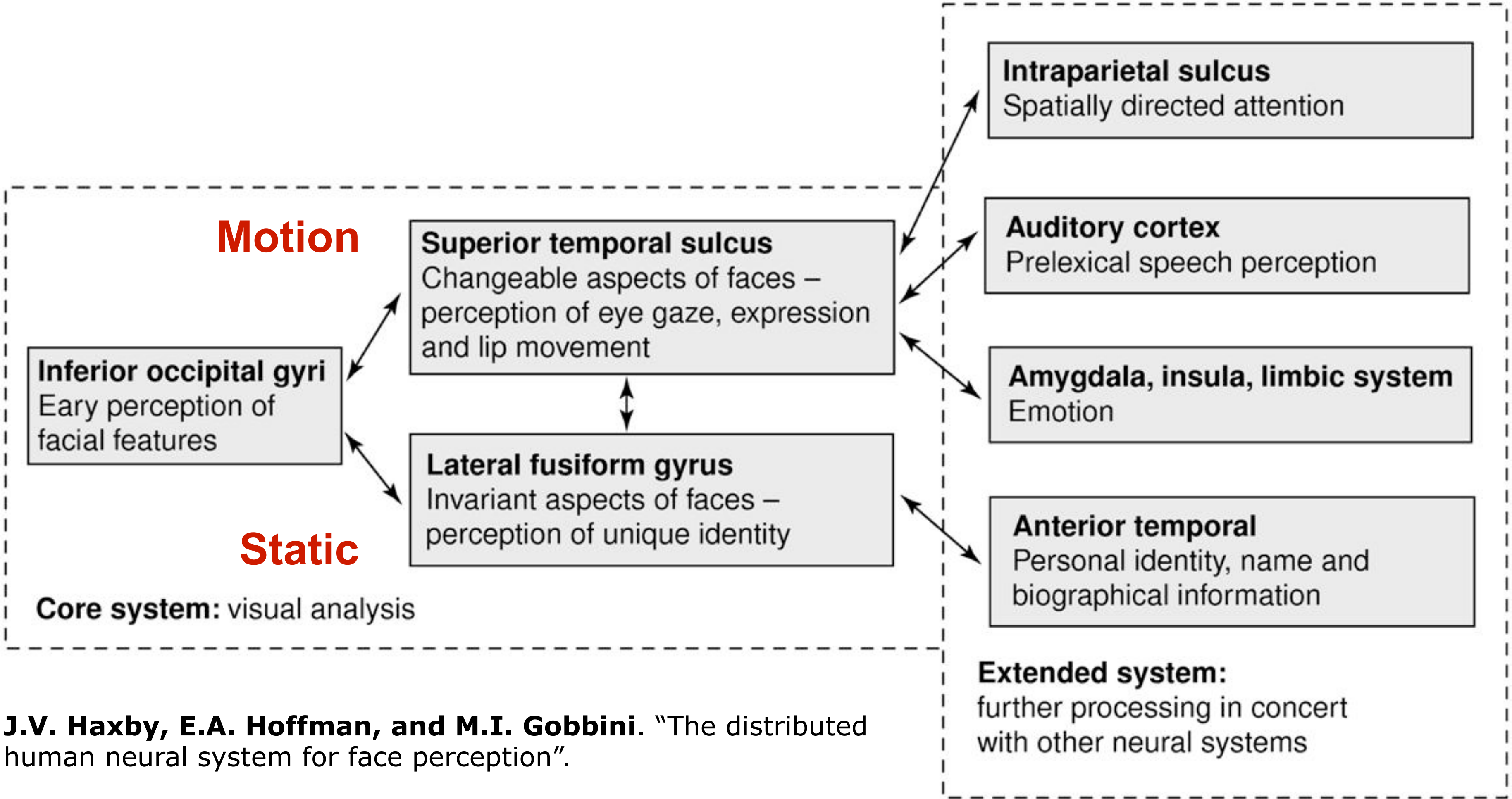
**= 18 trillion synapses!**
**= 2.3 trillion Bytes?**

If we can learn, say 10,000 faces this corresponds to

**220 MB/face**

(or a 7 sec. video stream of 1Kx1K images)

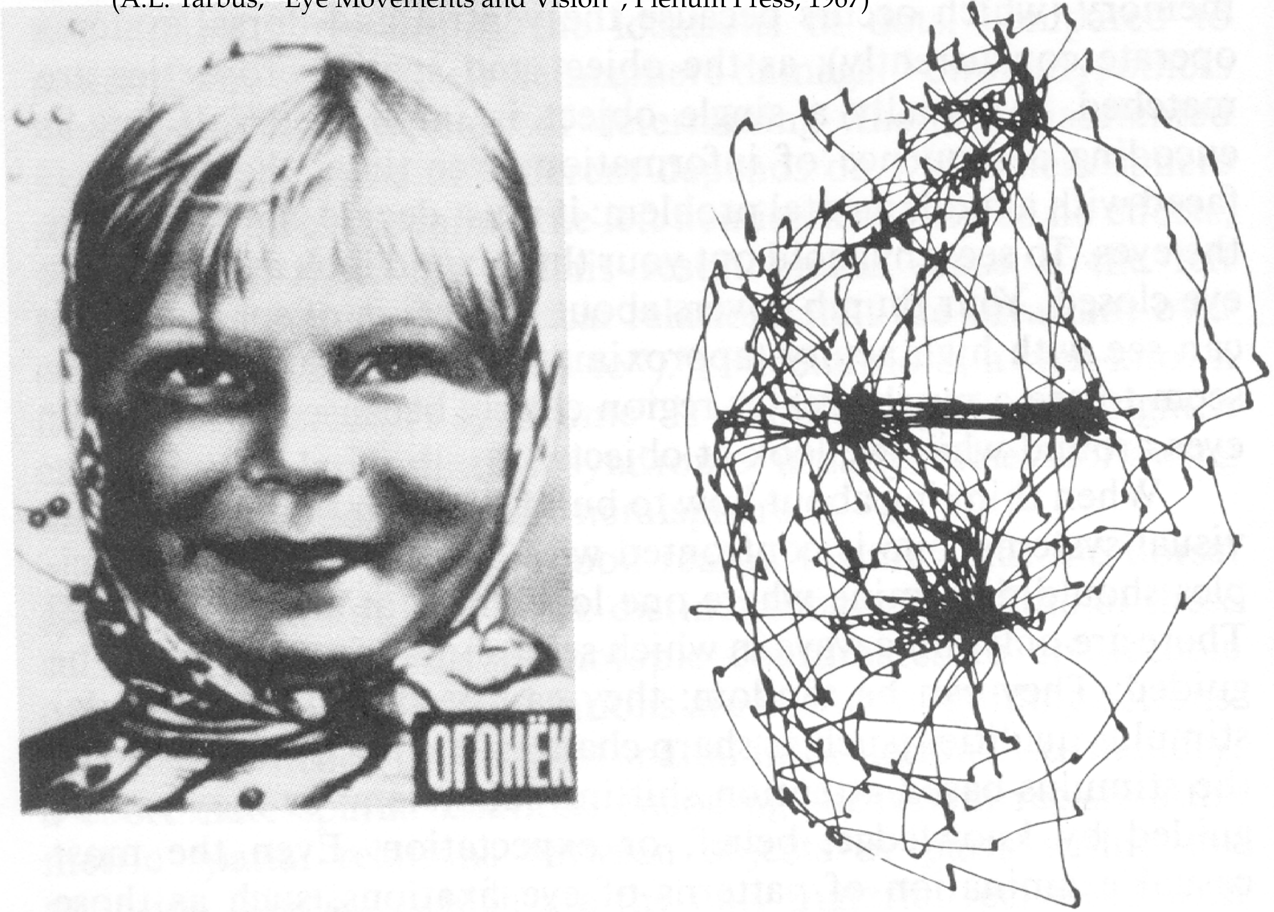
# Neural architecture of face perception - 1

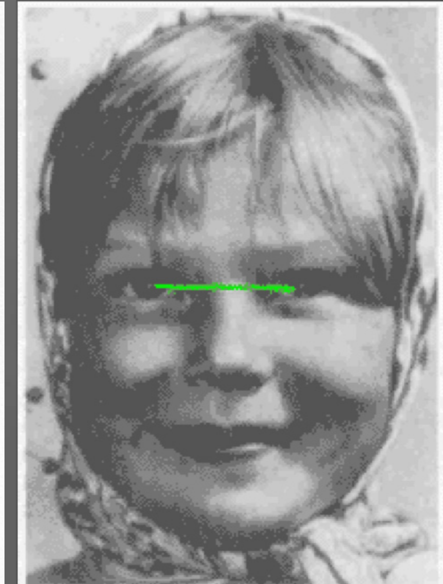



**J.V. Haxby, E.A. Hoffman, and M.I. Gobbini**. "The distributed human neural system for face perception".

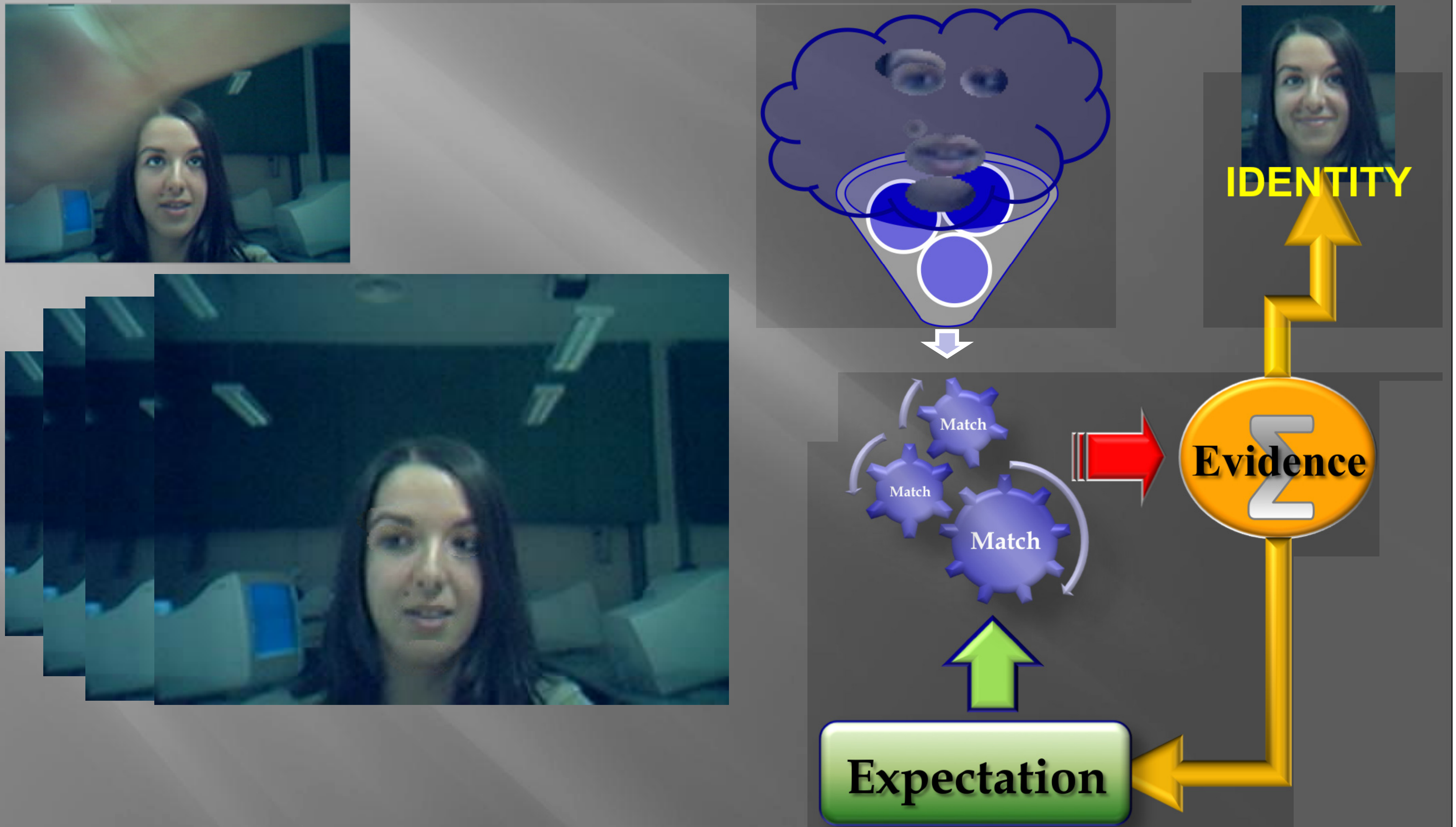
# Parsing the representation space

Eye movements while watching a girl's face
(A.L. Yarbus, "Eye Movements and Vision", Plenum Press, 1967)



# Biological face matching



# Conclusion

- Face recognition IS a mature technology.
  - Offers several advantages over more "traditional" biometric technologies
- Many practical algorithmic and system solutions have been proposed.

➢ Need to investigate more on the role of dynamic information

➢ More basic research on open issues

  ➢ Less *application push* and more *technology pull*

➢ Better understanding of human perception and representation of identity

# Conclusion

- More investments needed.
- More basic research on open issues.
- Need for "universal" databases for evaluation and testing on key scenarios.