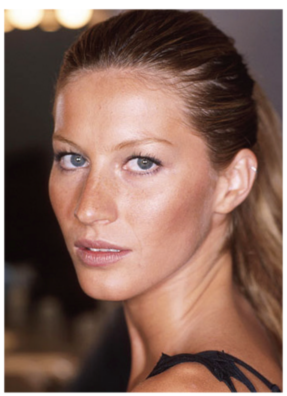
Face Recognition
A tutorial

Filareti Tsalakanidou

Informatics & Telematics Institute

Face recognition problem

- Given a still image or video of a scene, identify or verify one or more persons in this scene using a stored database of facial images

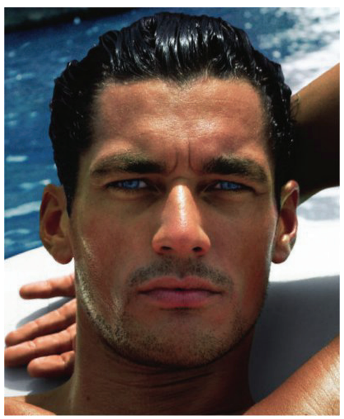


Who is she?

http://picasa.google.com

Face recognition/identification

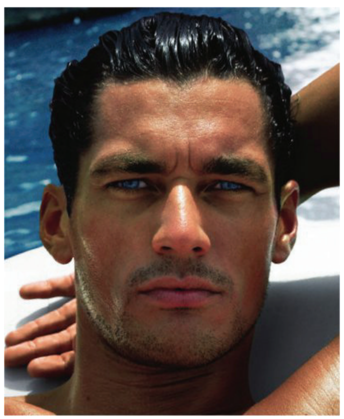
Who is this person?



He is David.

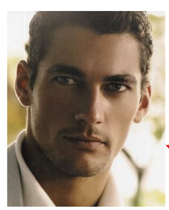
Face authentication/verification

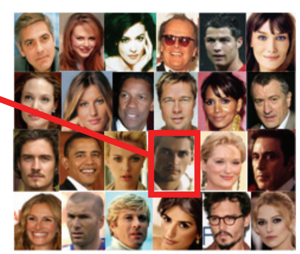
Is he who he claims to be?



Yes, he is.

I am David.





Applications

- Civil applications and law enforcement
  - National ID, passport, driver's license, border control
  - Surveillance of public places (airports, metro stations, etc)
  - Forensic applications
- Security applications for electronic transactions and access control
  - Physical access
  - Secure access to networks and infrastructures
  - e-health, e-commerce, e-banking (and now mobile…)
- Ambient Intelligence
  - Smart homes
  - Natural human-machine interaction
- Wearable systems
  - Memory aids and context-aware systems
- Entertainment
  - Interactive movies, computer games
- Search
  - Picasa 3.5 face recognition application for finding and managing photos

Face recognition in humans

- The human visual system starts with a preference for face-like patterns
- The human visual system devotes special neural mechanisms for face perception
- Facial identity and expression might be processed separately

- Facial features are processed holistically
  - Among facial features eyebrows are most important for recognition!
- Humans can recognize faces in very low dimensional images
  - Tolerance to image degradation increases with familiarity

- Color and texture are as important as shape
- Illumination changes influence generalization
- View-generalization is mediated by temporal association

P. Sinha, B. Balas, Y. Ostrovsky and R. Russell, "Face recognition by humans: 19 results all computer vision researchers should know about", *Proceedings of IEEE,* 2006.

Challenges: Intrapersonal variations

- If people can do it so easily, why can’t computers?
- Intrapersonal (intra-class) variations are variations of the appearance of the same face caused by
  - Illumination variations
  - Pose variations
  - Facial expressions
  - Use of cosmetics and accessories, hairstyle changes
  - Temporal variations (aging, etc)

Challenges: Interclass similarity

- Interclass similarity: different persons may have very similar appearance
  - Twins
  - Relatives
  - Strangers may look alike

http://multiples.about.com

http://www.nypost.com

http:// www.tujefetevigila.com

Challenges: Illumination variations

- Illumination variations may significantly affect the appearance of a face in 2D images
  - Recognition performance may drop more than 40% for images taken outdoors!
  - Humans have difficulties in recognizing familiar faces when light direction changes (e.g. top-lit → bottom-lit)

Yale Face Database B

Yale Face Database B http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html

Challenges: Pose variations

- Difference between two images of the same subject under different view angles is greater than differences between images of two different subjects under the same view

CMU Face In Action (FIA) Database

Rodney Goh, Lihao Liu, Xiaoming Liu and Tsuhan Chen, “The CMU Face In Action (FIA) Database”, *Proc. Int. Workshop on Analysis and Modelling of Faces and Gesture,* 2003.

Challenges: Facial expressions

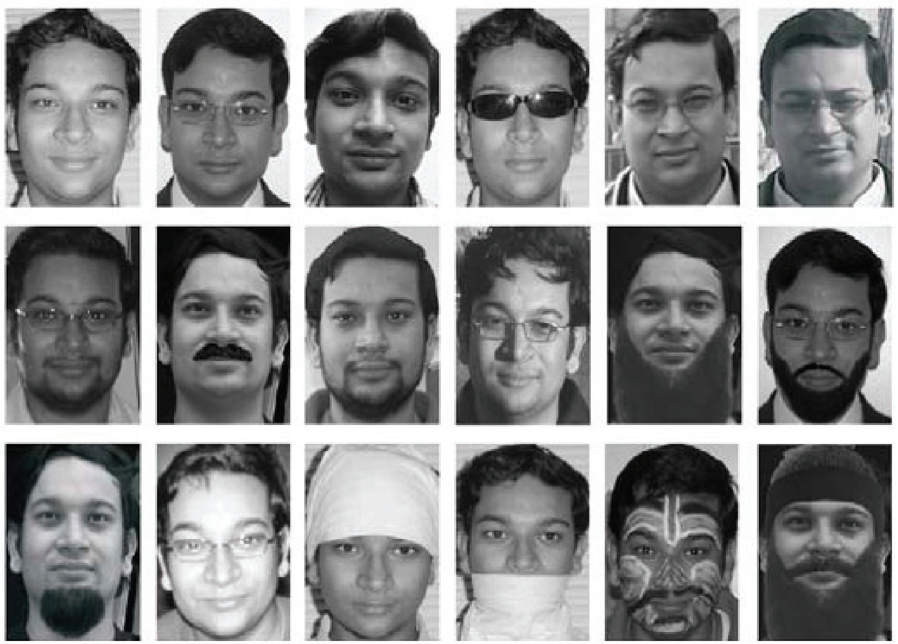
- Facial expressions caused by facial muscle movements may significantly deform the face surface

Binghamton University 3D Facial Expression database

BU-3DFE (Binghamton University 3D Facial Expression) database http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html

Challenges: Disguises

- People may disguise to avoid being recognized…



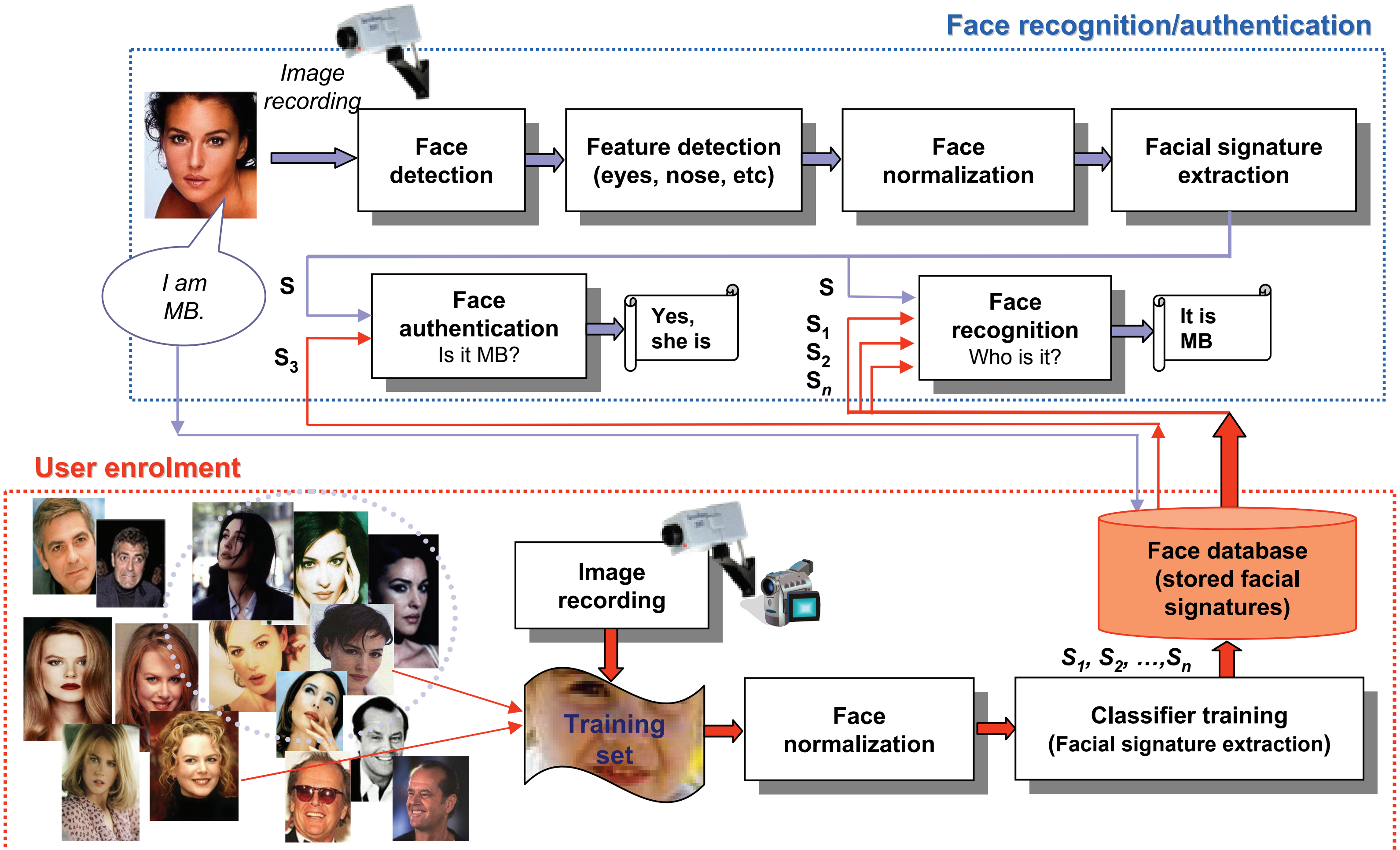
R. Singh, M. Vatsa and A. Noore, "Recognizing Face Images with Disguise Variations", *Recent Advances in Face Recognition,* I-Tech, Vienna, 2008.

Challenges: Information redundancy

- 20x20 facial image
- $256^{400}=2^{3200}$ possible combinations of intensity values
- Total world population as of 8 Oct. 2009 6,789,000,000 ≈ 2^{32}
- That's an extremely high-dimensional space…

Ming-Hsuan Yang, "Recent Advances in Face Detection: A tutorial", *Proc. Int. Conf. on Pattern Recognition,* 2004.

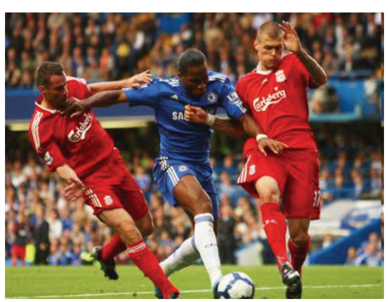
Typical face recognition system architecture



Face detection

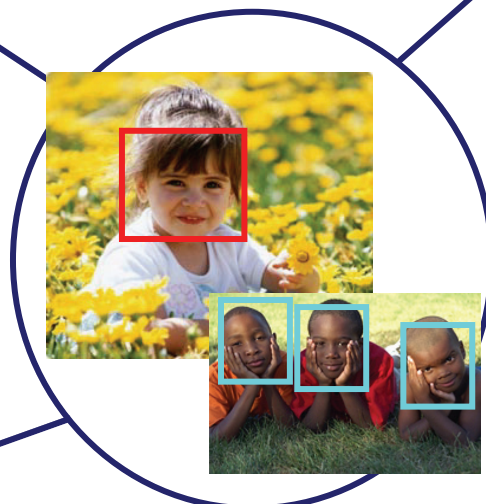
Face detection

- Face detection: find all faces in an image (if any) regardless of their position, scale, in plane rotation, pose, illumination, facial expressions, occlusions
  - First step to every face recognition system
- Face localization: find the exact location of a detected face
  - Detection of salient facial features such as eyes, nose, nostrils, eyebrows, mouth, etc
- Face tracking: detect (“follow”) a face in a video sequence



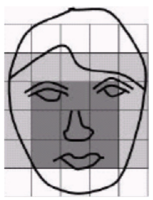
M. M.-H. Yang, D. J. Kriegman, and N. Ahuja, “Detecting Faces in Images: A Survey”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002

Face detection techniques



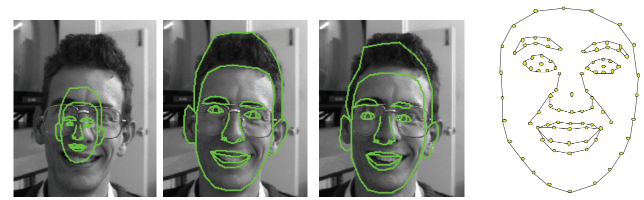
Knowledge-based

Translate knowledge about typical face to a set of rules



Structural matching

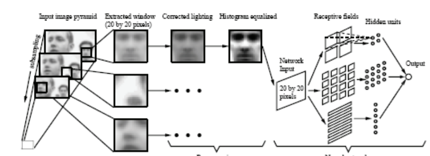
Statistical models of shape appearance based on a set of landmarks



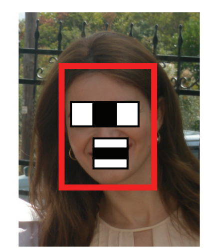
© http://personalpages.manchester.ac.uk/staff/timothy.f.cootes/

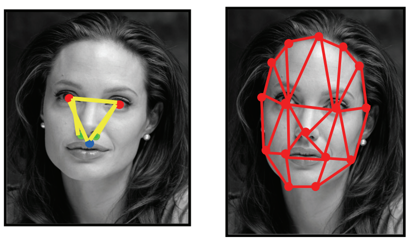
Appearance-based

Learn face characteristics from a representative set of example images using classic machine learning techniques



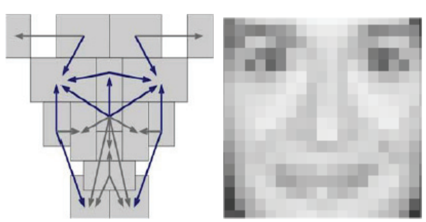
Eigenfaces, Neural Networks, Distribution-based, HMM, Haar features





Feature invariant

Find features of the face invariant to appearance variations (facial features, edges, shape, texture, skin color)



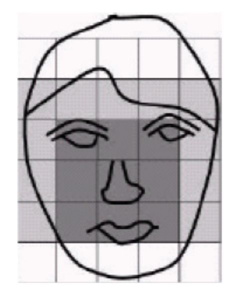
Template matching

Standard patterns stored to describing the face or facial features

M. M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002.
M. H. Yang, "Recent Advances in Face Detection: A tutorial", *Proc. Int. Conf. on Pattern Recognition*, 2004.

Knowledge-based face detection

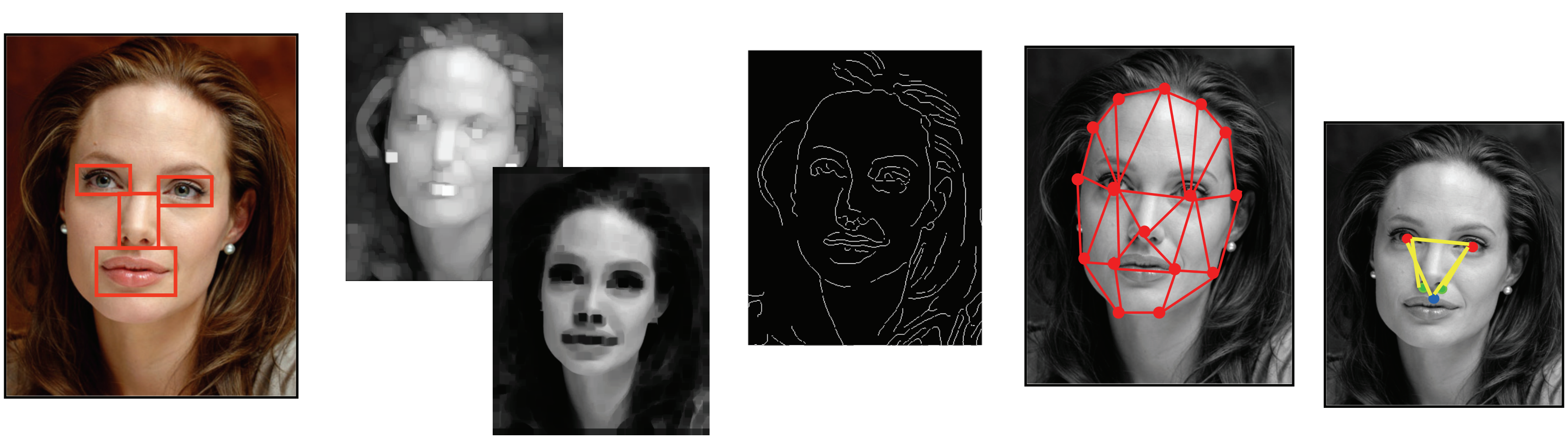
- Define a set of rules to represent the face and use them for detection, e.g.
  - “Center face part has uniform intensity”
  - “Face has two eyes, a nose and a mouth”
- Hierarchical approach examining the face at different resolution levels
  - Lower level: find possible face candidates based on image intensity
  - Mid level: detect edges
  - Higher level: extract facial features (mouth, eyes). Classify the image region as face or non-face
- Not easy to translate knowledge into rules or extend rules to different poses



G. Yang and T. Huang, “Human face detection in a complex background”, *Pattern Recognition*, 1994.
M. H. Yang, “Recent Advances in Face Detection: A tutorial”, *Proc. Int. Conf. on Pattern Recognition*, 2004.

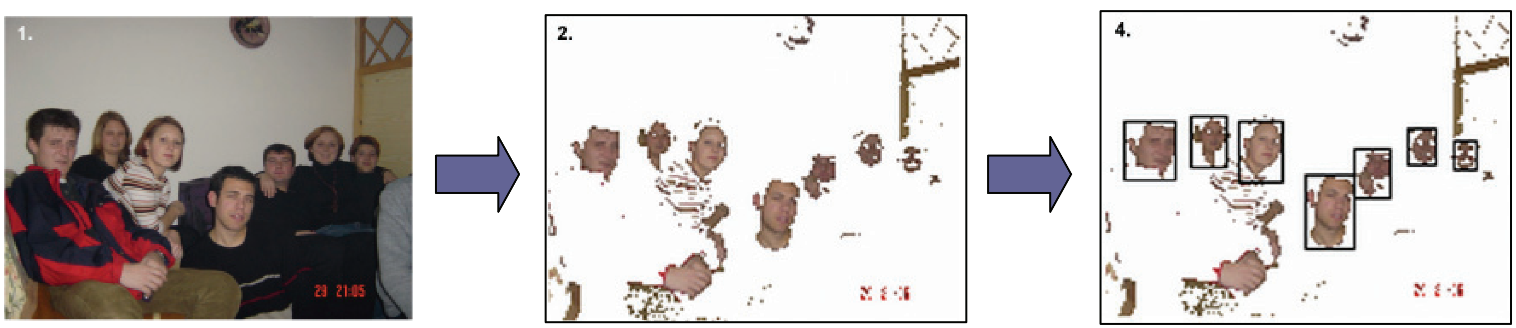
Feature invariant face detection

- Extraction of local facial features (eyes, eyebrows, nose, mouth) using multi-resolution or derivative filters, edge detectors, morphological operations, etc
- Statistical models, neural networks and graph matching used to describe relationships between features
  - Improved invariance under pose or illumination, problems due to occlusions, noise, complex background, etc



Skin color based face detection

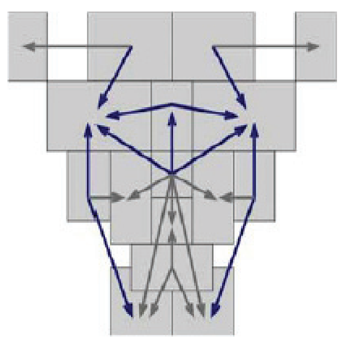
- Each pixel labeled as skin or non-skin
- Connected component analysis and grouping to merge neighbor skin areas
- A candidate region is classified as face if it satisfies some criteria
  - Easy to implement, insensitive to pose and facial expressions, but sensitive to illumination variations and other body-parts or skin-color like regions



J. Kovac, P. Peer and F. Solina, "Human Skin Colour Clustering for Face Detection", *Proc. Int. Conf. on Computer as a Tool,* 2003.

Template based face detection

- Several standard templates stored to describe the face as a whole or the facial features separately
  - Predefined templates based on edges, silhouettes
  - Deformable templates based on facial contours (e.g. Snakes)
- Face detection based on correlation with these templates
- Simple to implement but cannot deal with pose variations



M. H. Yang, "Recent Advances in Face Detection: A tutorial", *Proc. Int. Conf. on Pattern Recognition*, 2004.
B. Scassellati, "Eye finding via face detection for a foevated, active vision system", *Proc. National Conf. on Artificial Intelligence*, 1998.

Appearance based face detection

- Uses statistical analysis and machine learning techniques to learn the “characteristics” of a face from a large set of images
  - PCA, LDA
  - Support Vector Machines
  - Neural Networks
  - Hidden Markov Models
  - Adaboost
- Most successful approach, fast and robust
  - Detection rates 80-90% at a false positive rate of 10%
- Needs to search over scale and space and requires large set of training examples

http://picasa.google.com

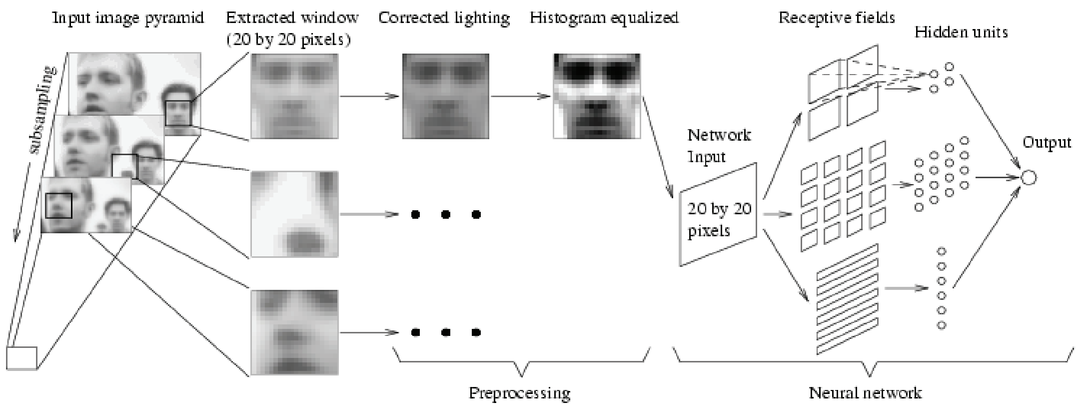
Eigenfaces for detection

- A low-dimensional subspace (face space) is built using a large set of training images and Principal Component Analysis (PCA)
- The distance of an image sub-window from the face space (DFFS) determines its likelihood to represent a face
- Sensitive to pose variations

B. Moghaddam and A. Pentland, "Probabilistic visual learning for object detection", *Int. Conf. on Computer Vision*, 1995

Neural networks

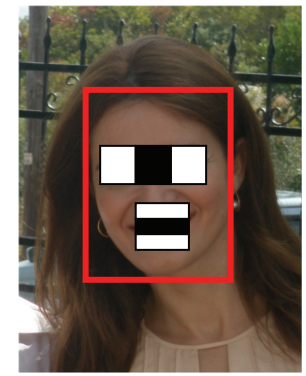
- Two-class pattern recognition problem
  - An image window is classified as face or non-face
  - Requires careful design of network and tuning of parameters and extensive training



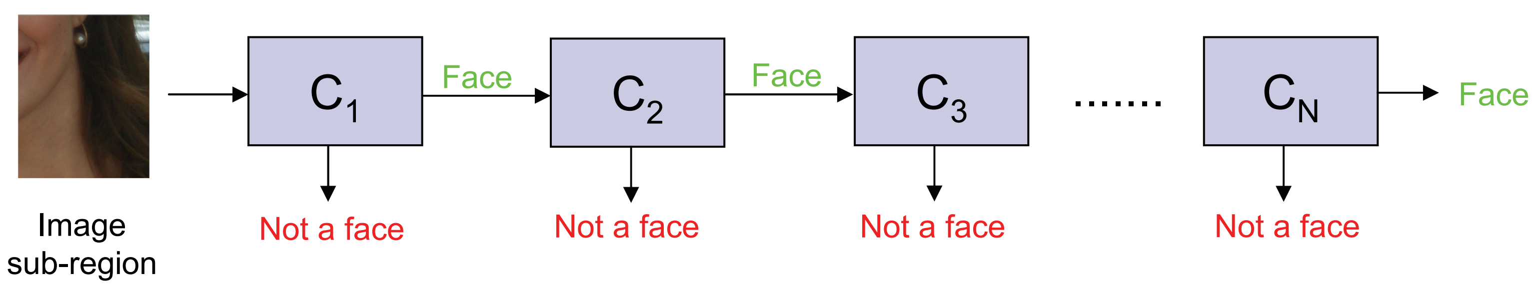
H. Rowley, S. Baluja and T. Kanade, “Neural Network-Based Face Detection”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1998.

Haar features

- Haar features consist of two or more rectangles and encode intensity differences between neighboring areas



- A cascade of boosted classifiers working with Haar features used to classify image regions as face or non-face
  - Classifiers at earlier stages use fewer Haar features
  - Feature selection is based on the Adaboost algorithm → features sorted in order of importance
  - Fast and robust, but time-consuming training (days...)



P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features”, *Proc. Int. Conf. on Computer Vision and Pattern Recognition*, 2001.

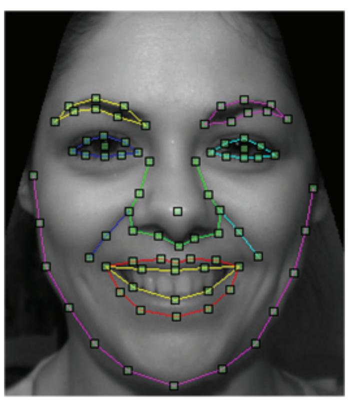
Facial feature detection

- Edge based techniques
- Feature templates for eyes, mouth, nose
  - Problematic when features are occluded or obscured
- Eigenfeatures
  - Eigeneyes, eigenmouth, etc
- Gabor wavelets, multidimensional erotion-dilation
- Structural matching using ASM, AAM, etc
  - More robust under shape or image intensity variation

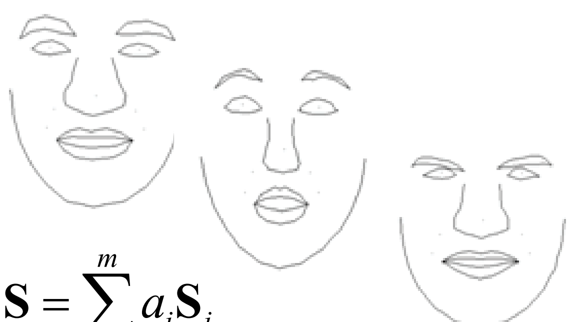
Active Shape Model (ASM)

- Statistical model representing the shape of faces
  - Point distribution model with N points
  - Local appearance model for each point based on image gradient
  - Point and local appearance distributions learned by applying PCA to a set of annotated images
  - The face can be expressed as the sum of a mean shape and a linear combination of basis shapes

- Iterative fitting to find the points that best match the local appearance distributions under constraints imposed by shape



Face model with 81 points



$$\mathbf{S} = \sum_{i=1}^{m} a_i \mathbf{S}_i$$

Modes of shape variation

Fitting results

T.F. Cootes, C.J. Taylor, D.H. Cooper and J. Graham, "Active shape models - their training and application", *Computer Vision and Image Understanding,* 1995.

Active Appearance Model (AAM)

- Single statistical model combining shape and texture
  - Shape model + texture model
  - Correlations between shape and texture are learned to generate the combined appearance model
- Iterative fitting to find model parameters that minimize the difference between the probe image and a synthesized model example (analysis by synthesis)

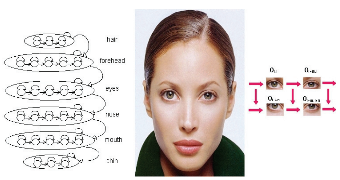
G.J. Edwards, C.J. Taylor and T.F. Cootes, “Face Recognition Using the Active Appearance Model”, *Proc. European Conf. on Computer Vision*, 1998.

2D face recognition techniques

2D face recognition techniques

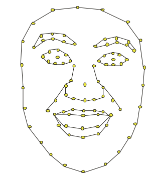
Appearance-based (Holistic)

Eigenfaces, Fisherfaces, ICA, Kernel PCA, Local Feature Analysis (LFA), Hidden Markov Models (HMM)

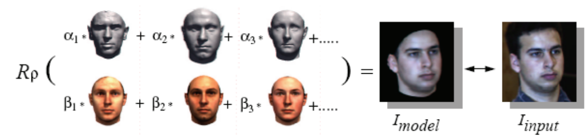


©National Geographic

Model-based

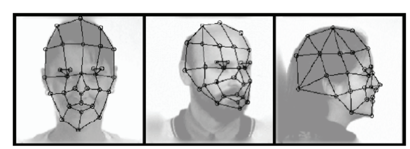


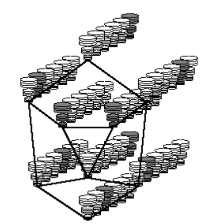
Active Appearance Models
3D morphable models



Feature matching

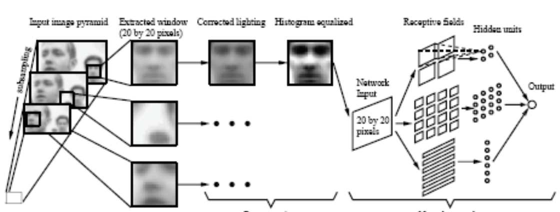
Geometry methods
Elastic Graph Matching





face bunch graph

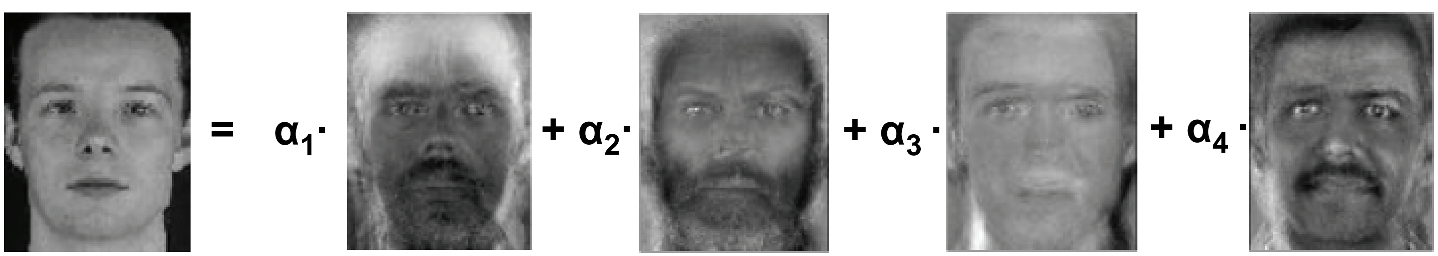
Neural networks



W. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, “Face recognition: a literature survey,” *ACM Computing Surveys*, 2003.

Eigenfaces

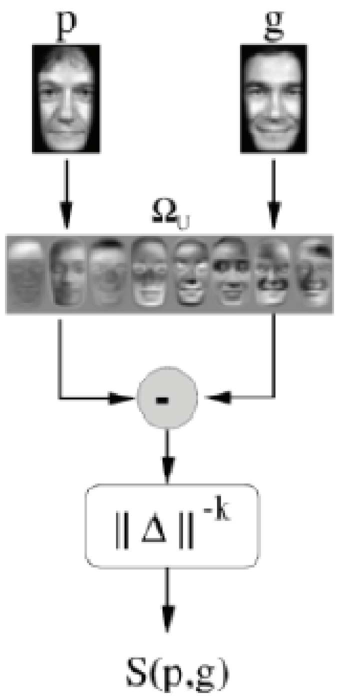
- Principal Component Analysis (PCA) applied on a set of images to extract a lower dimension subspace that best describes the variance of underlying data
  - Dimensionality reduction!
  - The principal components are called eigenfaces due to their face-like appearance
  - A face can be modeled as a linear combination of a small subset of the eigenfaces
  - Face recognition is based on comparing the coefficients of this linear representation
- Easy to implement, but sensitive to pose, needs accurate alignment of probe and gallery



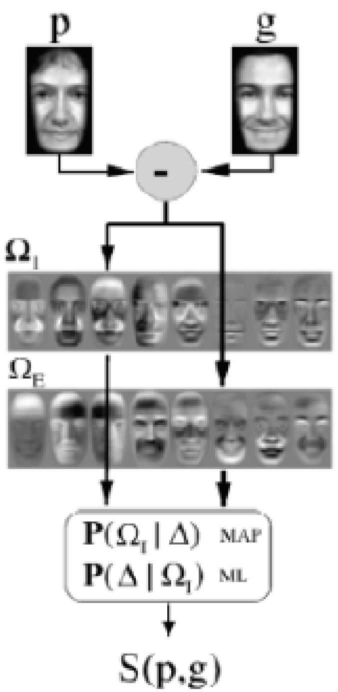
M. Turk, and A. Pentland, “Eigenfaces for recognition”, *Journal of Cognitive Neuroscience*, 1991.

Probabilistic eigenfaces

- A multi-class problem converted into a two-class problem
  - Intrapersonal and extrapersonal classes based on image differences between images of the same person and different persons respectively
  - Use of a probabilistic measure of similarity instead of Euclidean distances
- More robust to illumination variations and facial expressions



Classic eigenfaces

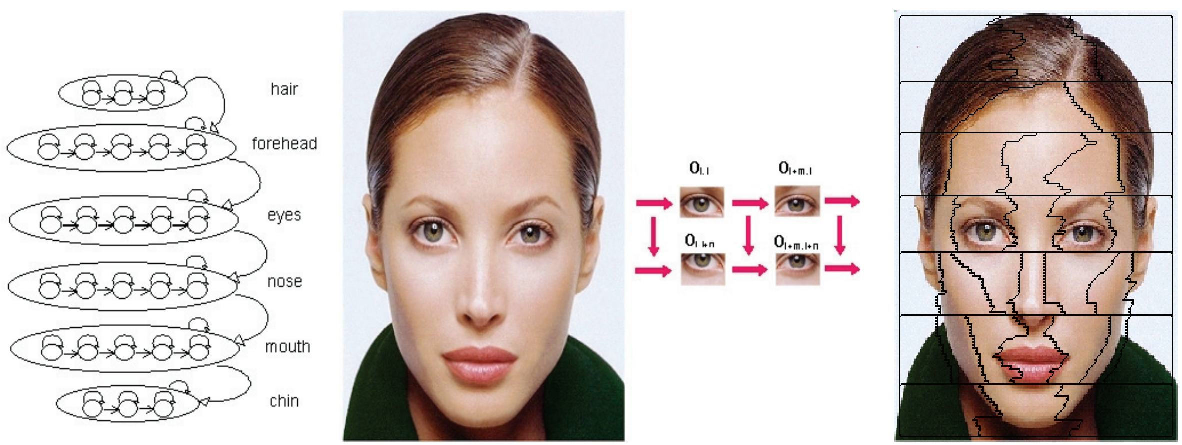


Probabilistic eigenfaces

B. Moghaddam, W. Wahid and A. Pentland, “Beyond Eigenfaces: Probabilistic Matching for Face Recognition”, *Proc. Int. Conf. on Automatic Face and Gesture Recognition*, 1998.

Embedded Hidden Markov Models

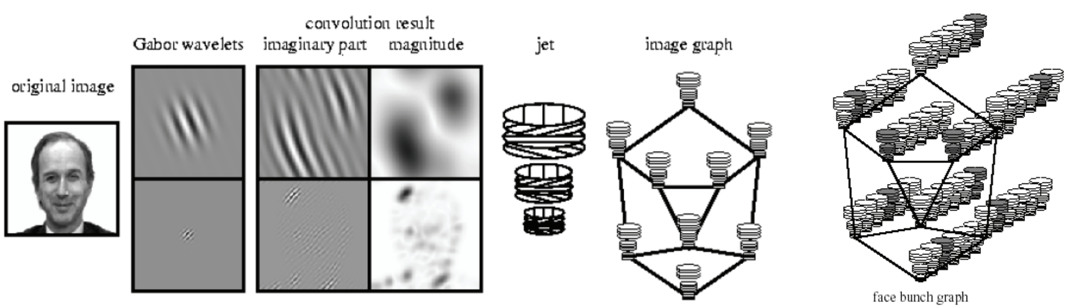
- EHMMs consist of a set of super states along with a set of embedded states
  - Super states model the face from top to bottom
  - Embedded states model the face from left to right
- Model parameters estimated based on observations extracted from training images – one model is trained for each face

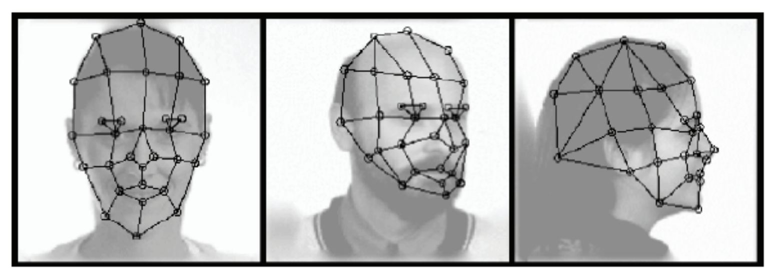


A. V. Nefian and M. H. Hayes III, "Maximum likelihood training of the embedded hmm for face detection and recognition", *Proc. Int. Conf. on Image Processing*, 2000.

Elastic Graph Matching

- The face is represented as a set of local feature vectors (jets) placed over the nodes of a labeled 2D graph (image graph)
  - Features are Gabor wavelet coefficients or morphological operators computed over different scales and orientations
  - Edges are labeled with distance information
- Face classification is based on elastically deforming the grid of a probe image to fit the grid of a gallery image
  - Cost function measuring jet similarity and grid distortion
- Elastic Bunch graph matching (EBGM)
  - Uses a set of jets for every node that correspond to several appearance variations
  - Extract image graph of probe image by fitting the EBGM and then match against gallery graphs

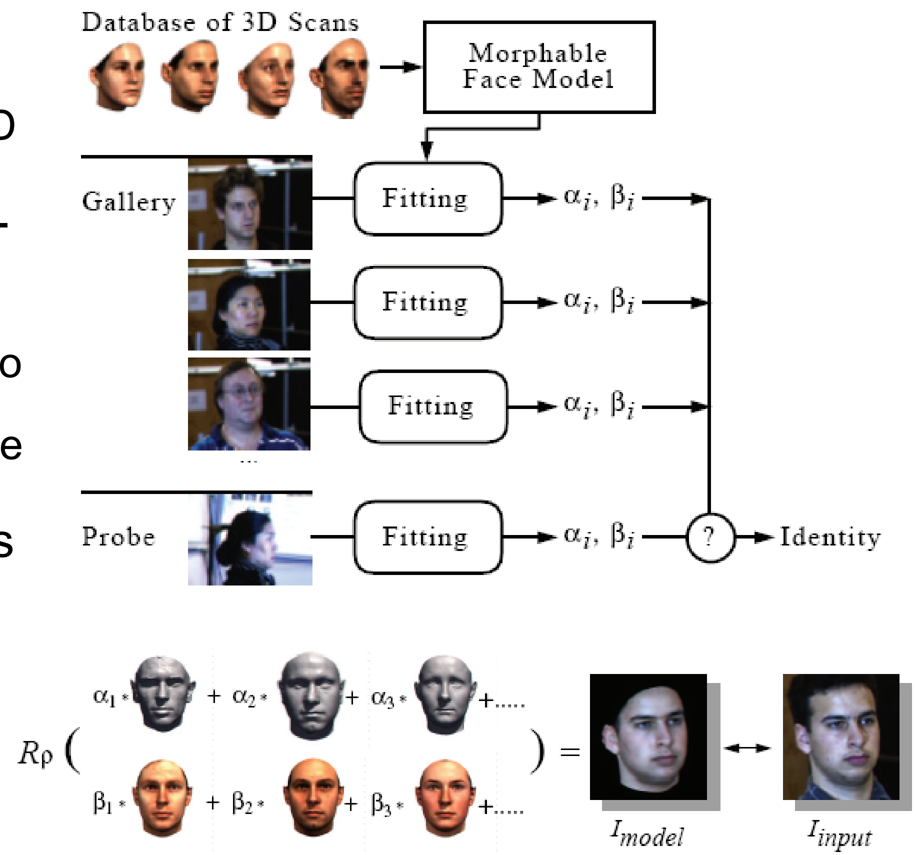




L. Wiskott, J. M. Fellous, and C. von der Malsburg, "Face recognition by elastic bunch graph matching", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, July 1997.
C. Kotropoulos, A. Tefas, and I. Pitas, "Frontal Face Authentication Using Morphological Elastic Graph Matching", *IEEE Trans. on Image Processing*, April 2000.

3D morphable model

- Statistical model built from dense 3D face scans with texture
  - Two distinct models for texture and 3D shape built by applying PCA
- Model fitting based on an analysis-by-synthesis approach
  - Estimate model parameters and face position, orientation and illumination so that the image produced by model rendering is as close as possible to the input image
- Face matching based on Mahalanobis distances of shape and texture parameters



V. Blanz, S. Romdhani, and T. Vetter, "Face Identification across different poses and illuminations with a 3D morphable model", *Proc. Int. Conf. on Automatic Face and Gesture Recognition*, 2002.

Face Recognition Vendor Test 2002

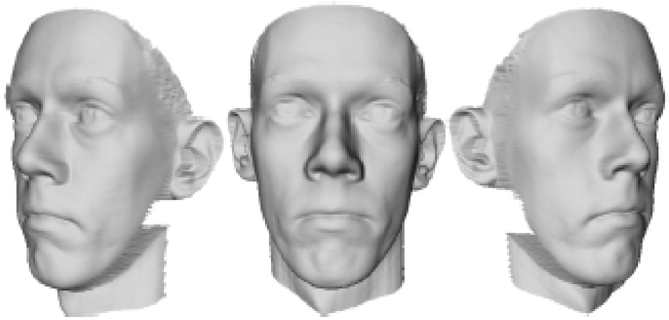
- FRVT2002: Independent evaluation of commercial 2D face recognition systems
  - Performance in large datasets (121589 images, 37437 subjects)
  - Performance under pose and illumination variations, effect of time
- Results
  - 10% FRR for 1% FAR, 18% FRR for 0.1% FAR
  - 73% RR for 37437 subjects, 83% for 1600 subjects, 85% for 800 subjects
  - FR performance decreases approximately linearly with elapsed time between database and new images (5% per year)
  - Recognition rate drops more than 40% for images taken outdoors!
  - Recognition rate drops to 20% for 45° rotations (FAR 1%)
    - Three-dimensional morphable models substantially improve the ability to recognize non-frontal faces (80%)
  - Recognition from video sequences not better than from still images
  - Males are easier to recognize than females (6-9%)
  - Younger people are harder to recognize than older people
- Performance measurements
  - False Acceptance Rate (FAR): Percentage of instances that the system accepts a claimed identity when it shouldn't
  - False Rejection Rate (FRR): Percentage of instances that the system rejects a claimed identity when it shouldn't
  - Recognition Rate (RR): Percentage of instances that a person is correctly recognized by the system

http://www.frvt.org/DLs/FRVT_2002_Evaluation_Report.pdf

3D and 2D+3D face recognition techniques

Why use 3D images for face recognition?

- 3D images represent the 3D structure of the face
  - Rich source of information not captured in 2D images
    - Better at capturing surface geometry
  - Not affected by illumination variations or use of cosmetics
  - Less sensitive to appearance variations
  - Easier to handle pose variations
    - Projective nature of 2D images …
  - Simplifies face & facial feature detection, pose estimation & pose compensation



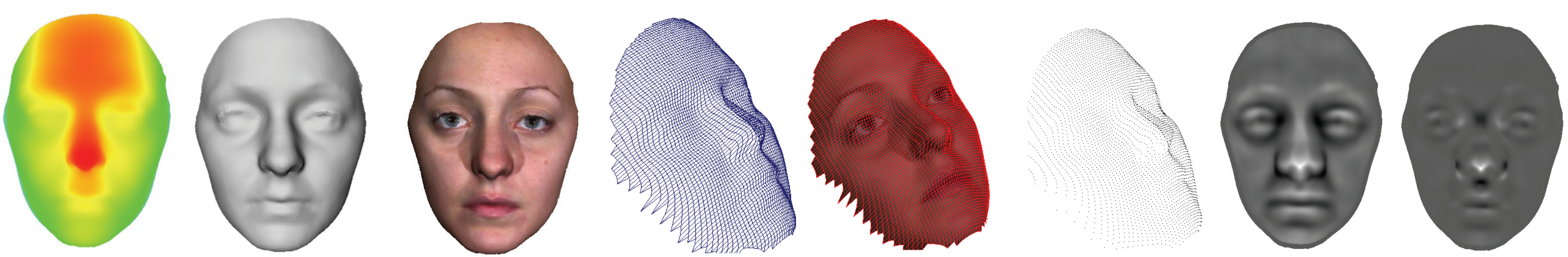
2D+3D face recognition

- If 3D is so much better, should we forget 2D?
  - No! 2D and 3D images provide complementary information about the face
    - Texture is as important as geometry
  - The ideal is a multimodal approach: use both 2D and 3D facial information in all stages of face recognition
    - Experiments have shown that the combination of 2D and 3D offers increased performance compared to 2D or 3D alone
    - Fusion of scores of 2D and 3D classifiers

K. I. Chang, K. W. Bowyer and P.J. Flynn, "An Evaluation of Multimodal 2D+3D Face Biometrics", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2005.

3D facial data representations

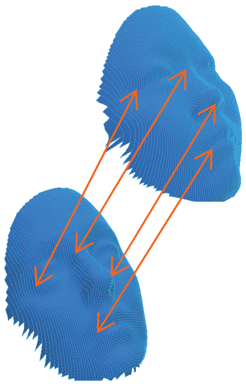
- Range (depth) image
  - the z coordinates of the face points are mapped on a regular *x−y* grid using linear interpolation
  - 2D image where pixel values correspond to distance from camera plane
- Point cloud
  - the set of the 3D coordinates of the points of a face
- 3D mesh
- Curvature
  - each point in the face is described by its curvature (Gaussian, mean, principal)
- Surface normal
  - each point in the face is described by its normal vector



Range image | 3D face surface | 3D face surface overlaid with texture | 3D mesh | 3D mesh + texture | Point cloud | Gaussian curvature | Mean curvature

BU-3DFE (Binghamton University 3D Facial Expression) database http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html

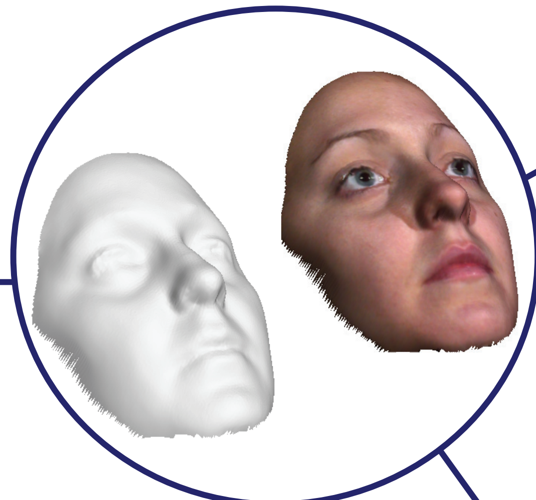
3D face recognition techniques

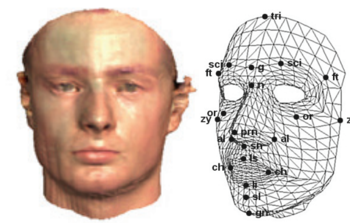


• ICP, Hausdorff distance
• Needs good initial estimation to converge

Point-to-point matching

Surface-based





Model-based

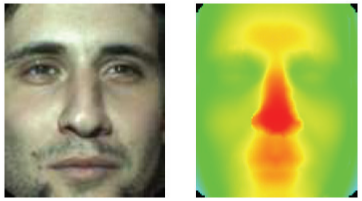
• 3D morphable model, 3D annotated face model
• Analysis-by-synthesis
• Not real-time

Curvature-based

• Point signatures, EGIs
• Sensitive to noise and data quality

Appearance-based

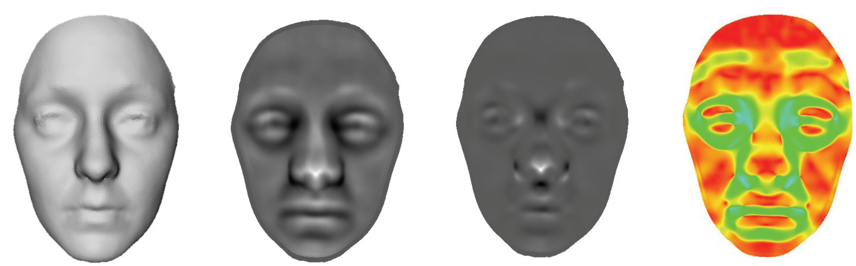
• Eigenfaces, Fisherfaces
• Works on range images only (2D classification…)
• Needs accurate alignment between probe and gallery



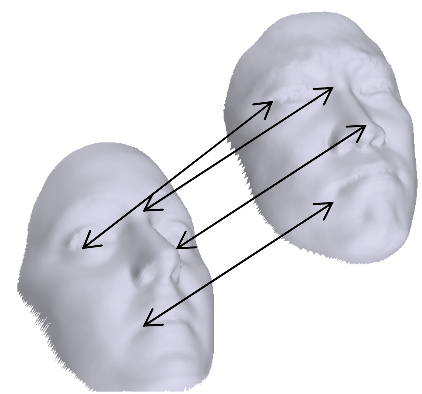
T. Papatheodorou and D. Rueckert, "3D Face Recognition", Face Recognition, I-Tech, Vienna, Austria, 2007.
B. Gokberk, "Principles of 3D Facial Recognition", *1st Biosecure Industrial Committee Meeting*, Feb. 2006.

Surface-based 3D face recognition

- Based on rigidity assumption → use of classic 3D object recognition techniques
  - Use of local curvature features, which are rotation invariant (Point Signatures, EGIs)
  - Use of point-to-point matching (ICP, Hausdorff distance)
- Sensitive to image noise, heavy computation load



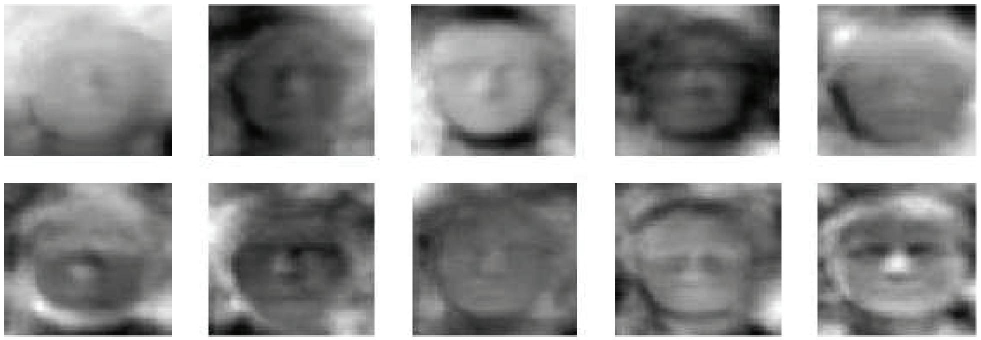
Face surface Gaussian curvature Mean curvature Shape Index



T.K. Kim, S.C. Kee, and S.R. Kim, "Real-time normalization and feature extraction of 3d face data using curvature characteristics", *Proc. Int. Workshop on Robot and Human Interactive Communication*, 2001.
H.T. Tanaka, M. Ikeda, and H. Chiaki, "Curvature-based face surface recognition using spherical correlation: Principal directions for curved object recognition", *Proc. Int. Conf. on Automatic Face and Gesture Recognition*, 1998.
G. Medioni and R. Waupotitsch, "Face recognition and modeling in 3D", *Proc. Int. Workshop on Analysis and Modeling of Faces and Gestures*, 2003.

Appearance-based 3D face recognition

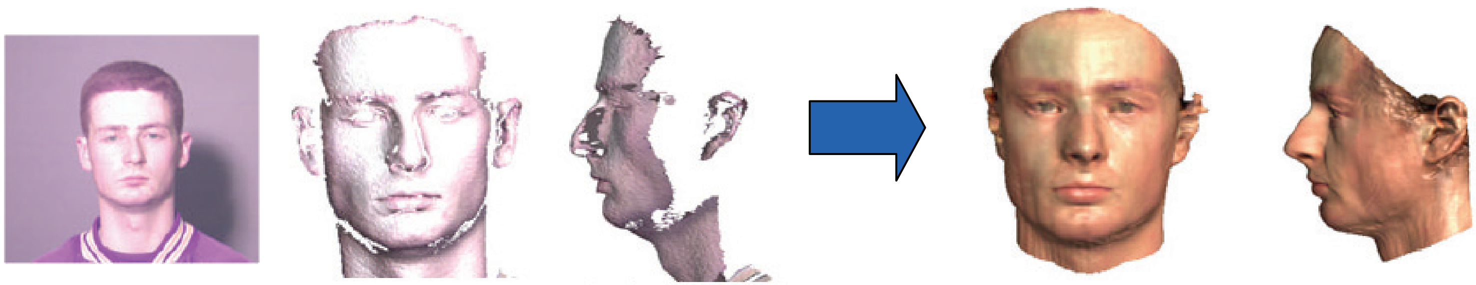
- Eigenfaces or Fisherfaces applied to range images
  - Easy to implement, fast response
  - Requires accurate alignment of probe and gallery images – may be achieved by localizing facial features such as the nose and the eyes
  - Sensitive to pose variations and facial expressions
  - First technique to be used for 2D+3D face recognition [Tsalakanidou et al. 2003]



F. Tsalakanidou, D. Tzovaras,and M.G. Strintzis, "Use of depth and colour eigenfaces for face recognition", *Pattern Recognition Letters*, 2003.
K. Chang, K. Bowyer and P. Flynn, "Face Recognition Using 2D and 3D Facial Data", *Proc. Multimodal User Authentication Workshop*, 2003.

Model-based 3D face recognition

- 3D parametric-morphable model [Blanz2007]



- Non-parametric 3D model [Kakadiaris2007]

V. Blanz, K. Scherbaum, H.P. Seidel, “Fitting a Morphable Model to 3D Scans of Faces”, *Proc. Int. Conf. on Computer Vision,* 2007.

I. Kakadiaris, G. Passalis, G. Toderici, M. Murtuza, Y. Lu, N. Karampatziakis, T.Theoharis, "Three-Dimensional Face Recognition in the Presence of Facial Expressions: An Annotated Deformable Model Approach“, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2007.

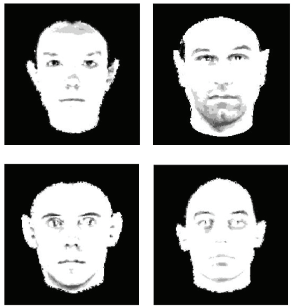
Handling appearance variations

Handling illumination variations

- Illumination invariant face representations
- Subspace methods
- Generative image models
- Illumination estimation and image relighting

Illumination insensitive techniques

- Use of illumination insensitive representations of face images based on egdes, correlation filters, gradient direction, image ratios, face symmetry
  - Quotient image
  - Symmetric shape from shading
- Illumination invariants *do not* exist for Lambertian surfaces



A. Shashua and T.R. Raviv, "The quotient image: Class based re-rendering and recognition with varying illuminations", *IEEE PAMI*, 2001.

Sub-space techniques

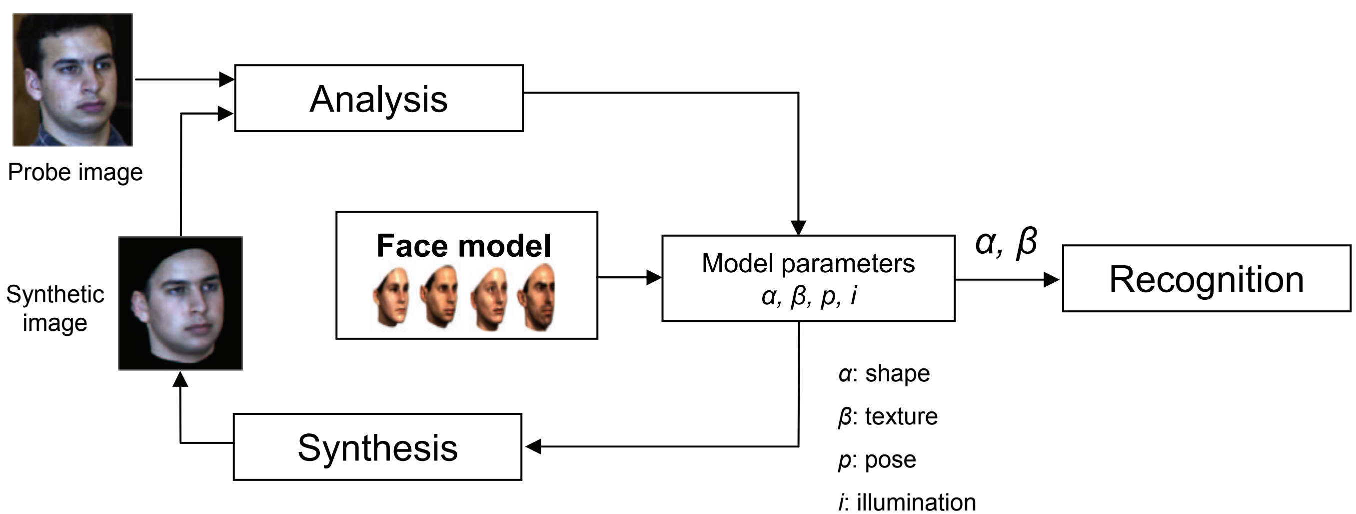
- Modeling of illumination variations using linear subspaces
  - The face is considered a Lambertian surface → the set of images of a face obtained under a wide variety of lighting conditions can be approximated by a low-dimension linear subspace
- Subspace estimation
  - PCA applied to a number of images of the same subject under different illumination
  - Illumination cones
  - Spherical harmonics
- Requirement for large training sets and pixel wise alignment between probe and gallery, reliance to simplified reflectance models

P. Belhumeur and D. Kriegman, "What Is the Set of Images of an Object under All Possible Illumination Conditions?", *International Journal of Computer Vision*, 2001.

L. Zhang and D. Samaras, "Face Recognition Under Variable Lighting Using Harmonic Image Exemplars", *Proc. Int. Conf. on Computer Vision and Pattern Recognition*, 2003.

Generative image models

- Separates intrinsic model parameters of the face (shape, texture) from extrinsic imaging parameters (pose, illumination, camera parameters)
  - Parameterize a new image in terms of the model
    - Model parameters estimated using an analysis-by-synthesis approach
  - Use shape and texture parameters for classification
- Needs large training database, employs time consuming non-linear fitting techniques, requires manual selection of landmarks



V. Blanz, S. Romdhani, and T. Vetter, “Face Identification across different poses and illuminations with a 3D morphable model”, *Proc. Int. Conf. on Automatic Face and Gesture Recognition*. 2002.

Image relighting

- Illumination estimation
  - Much easier when 3D data is available
- Synthesis of novel views resembling the illumination of gallery images
  - Normalization of probe images to diminish the effect of varying illumination conditions
  - Generation of symmetric frontally illuminated images
- Inverse of generative approach

S. Malassiotis, M.G. Strintzis, "Robust face recognition using 2D and 3D data: Pose and illumination compensation", *Pattern Recognition*, 2005.

Handling pose variations

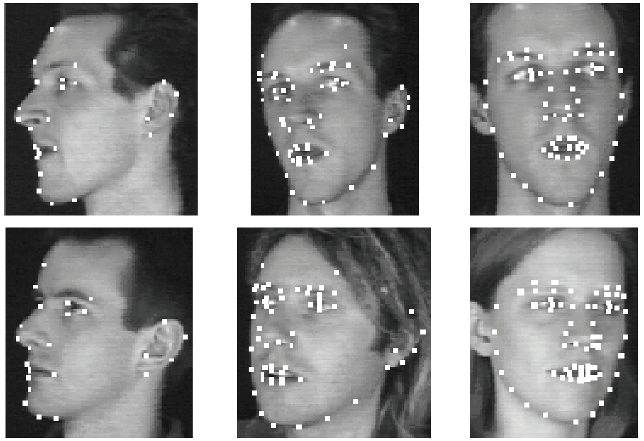
- 3D model-based approaches (generative)
  - Deformable 3D face models or AAMs
- Multi-view approaches
  - Multi-view database images for each subject
- Pose estimation and compensation
  - Creation of normalized frontal views

Model-based approaches

- Deformable 3D face models or AAMs (generative models)
  - Shape, texture, position and pose estimated by fitting the model on a 2D or 3D image
    - Automatic generation of novel views resembling the pose in the probe image – minimization of difference metric (analysis-by-synthesis approach)
  - Classification based on
    - Shape & texture model parameters
    - Similarity between generated view and probe image

Multi-view approaches

- View-based methods
  - Set of separate eigenspaces, each capturing the variations of several individuals under the same pose
  - Set of separate models, e.g. AAMs, to represent appearance under different poses
    - Require extensive enrolment
  - Illumination cones extended for pose (one cone for each pose)



A. Pentland, B. Moghaddam, T. Starner, "View-based and modular eigenspaces for face recognition", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1994.

T.F. Cootes, K. Walker, C.J. Taylor, "View-based active appearance models", *Proc. IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 2000.

Pose estimation & compensation

- Pose estimation based on extraction of salient features
  - Difficult in 2D images due to their projective nature
  - Easier when 3D data is available
- Pose compensation → generation of normalized (frontal) image views
  - Warping procedure between gallery and probe
- Face matching between normalized views

S. Malassiotis, M.G. Strintzis, “Robust face recognition using 2D and 3D data: Pose and illumination compensation”, *Pattern Recognition*, 2005.

Automatic generation of synthetic views under varying pose and illumination

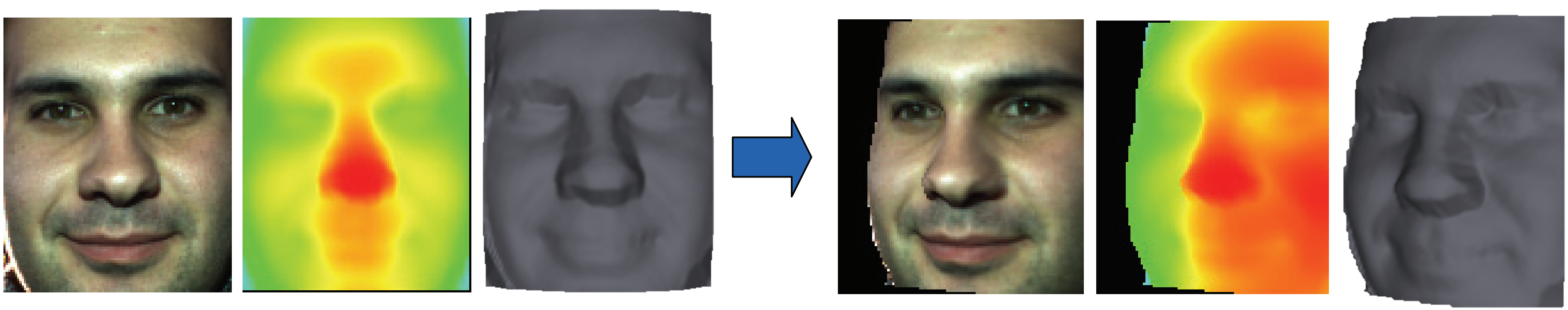
- Subspace techniques rely on the availability of large training sets to achieve high recognition rates → use of domes

- Solution: database enrichment with automatically generated synthetic images with arbitrary pose and illumination using a few 2D+3D frontal views
  - Avoids cumbersome enrolment process
  - Only a few images per person are needed
  - For each original image, a set of synthetic views is generated based on depth data

F. Tsalakanidou, S. Malassiotis and M.G. Strintzis, "Face localization and authentication using color and depth images", *IEEE Trans. on Image Processing,* 2005.

Synthetic images - pose

- Generation of synthetic poses
  - Creation of a 3D face mesh based on depth image
  - Rotation of 3D face mesh
  - Rendering of 3D data using z-buffer algorithm
    - generation of synthetic depth + color image



F. Tsalakanidou, S. Malassiotis and M.G. Strintzis, “Face localization and authentication using color and depth images”, *IEEE Trans. on Image Processing,* 2005.

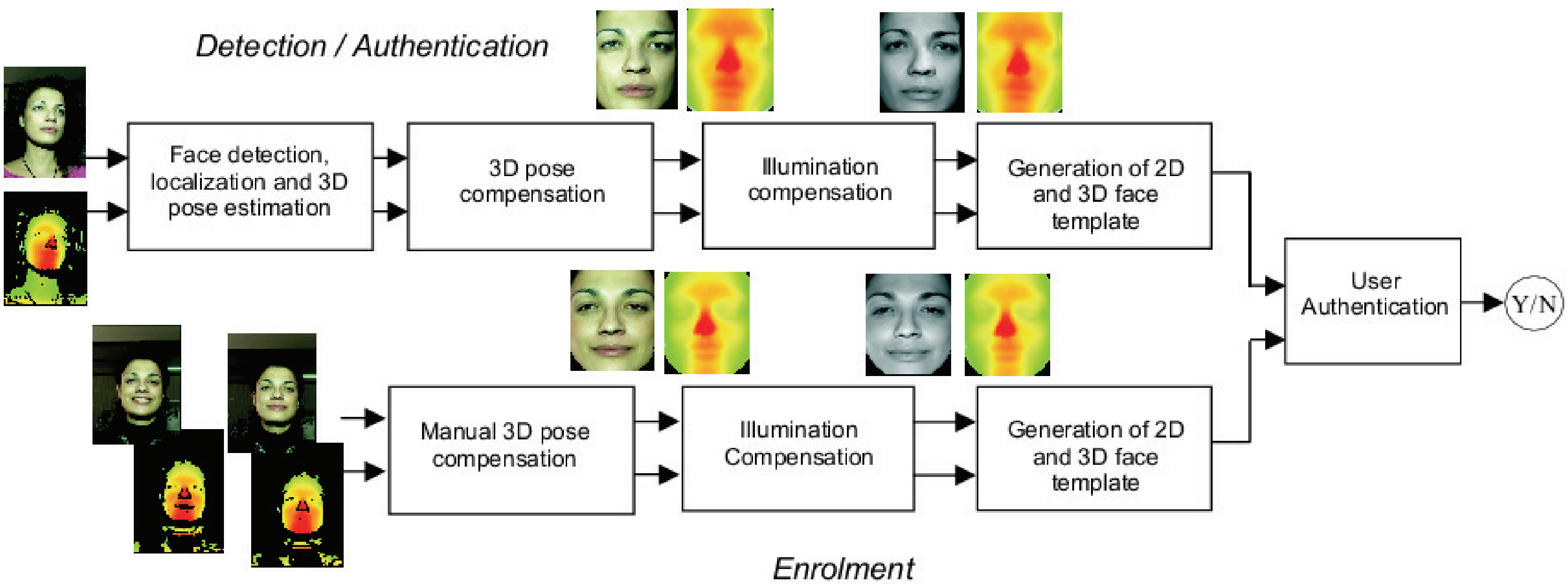
Synthetic images - illumination

- Simulation of heterogeneous shading of the face caused by a directional light
  - Estimate surface normal **n** over each pixel using depth image
  - Define light source **L** based on azimuth angles θ and φ
  - Create a synthetic view illuminated by **L**: $I_s=I_o(k_a+k_d\mathbf{L}\cdot\mathbf{n})$
    - k_a, k_d: weights for ambient light and diffuse reflectance

- Training of classifier (PCA, EHMM) with a large set of synthetic views

F. Tsalakanidou, S. Malassiotis and M.G. Strintzis, “Face localization and authentication using color and depth images”, *IEEE Trans. on Image Processing,* 2005.

Pose & illumination compensation based on 2D+3D images

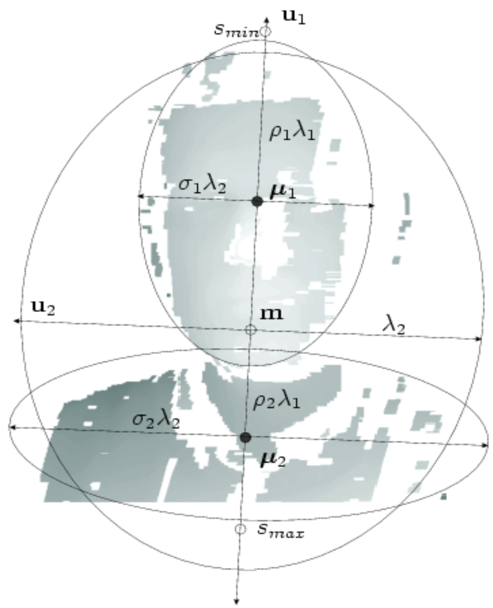


S. Malassiotis, M.G. Strintzis, "Robust face recognition using 2D and 3D data: Pose and illumination compensation", *Pattern Recognition*, 2005.

F. Tsalakanidou, F. Forster, S. Malassiotis, M.G. Strintzis, "Real-time acquisition of depth and color images using structured light and its application to 3D face recognition", *Real Time Imaging*, 2005.

3D pose estimation

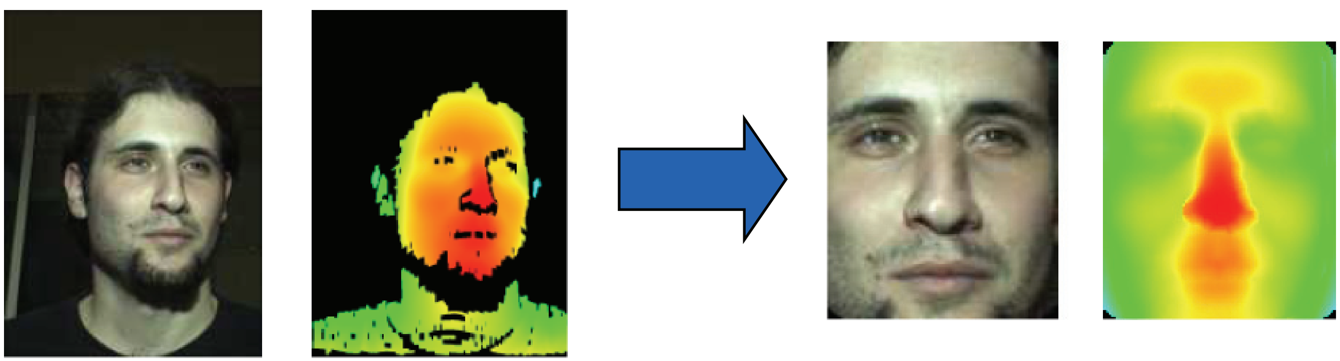
- Face detection based on 3D moments and a-priori knowledge of face geometry
- Nose tip and nose ridge localization based on principal curvatures and 3D face symmetry
- Pose estimation using 3D face symmetry
  - Accuracy: < 2mm RMS (nose tip location), <2.5° RMS (face orientation) up to 30°



S. Malassiotis, M.G. Strintzis, "Robust face recognition using 2D and 3D data: Pose and illumination compensation", *Pattern Recognition*, 2005.
F. Tsalakanidou, F. Forster, S. Malassiotis, M.G. Strintzis, "Real-time acquisition of depth and color images using structured light and its application to 3D face recognition", *Real Time Imaging*, 2005.

3D pose compensation

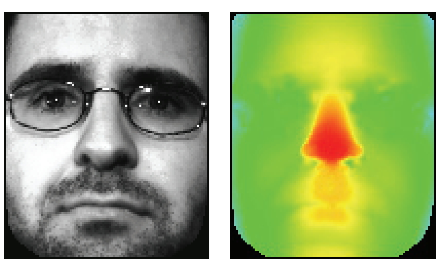
- Estimate the pose in a pair of probe images
- 3D warping to align the probe with the gallery
- Enhance image alignment using ICP
- Generate frontal views

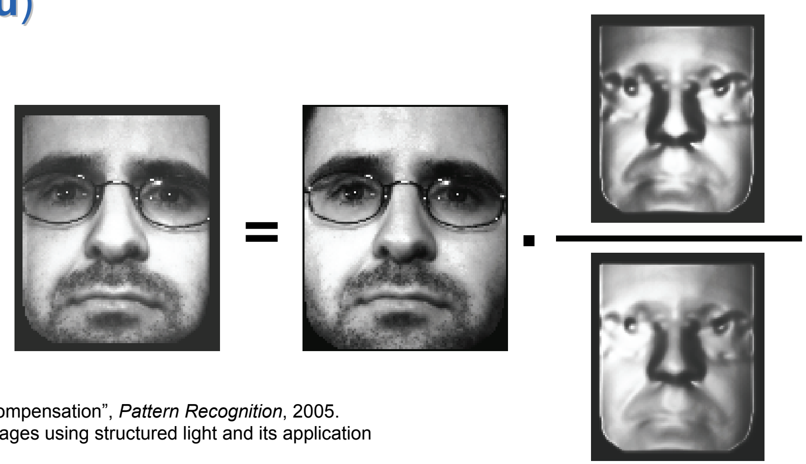


S. Malassiotis, M.G. Strintzis, “Robust face recognition using 2D and 3D data: Pose and illumination compensation”, *Pattern Recognition*, 2005.
F. Tsalakanidou, F. Forster, S. Malassiotis, M.G. Strintzis, “Real-time acquisition of depth and color images using structured light and its application to 3D face recognition”, *Real Time Imaging*, 2005.

Illumination compensation

- Estimation of the light source **L** based on example-based regression
- Relighting with frontal illumination $\mathbf{L}_0$
  - $I_C(\mathbf{u}) = A(\mathbf{u}) \cdot R(I_D, \mathbf{L}, \mathbf{u})$
    - I_C: pose compensated color image
    - I_D: pose compensated depth image
    - A: unknown face albedo
    - R: rendering of the surface with constant albedo
  - $\tilde{I}_c(\mathbf{u}) = A(\mathbf{u}) \cdot R(I_D, \mathbf{L_0}, \mathbf{u})$
    - $\tilde{I}_c$: image relit by $\mathbf{L_0}$
  - $\tilde{I}_c(\mathbf{u}) = I_c(\mathbf{u}) \cdot R(I_D, \mathbf{L_0}, \mathbf{u}) / R(I_D, \mathbf{L}, \mathbf{u})$





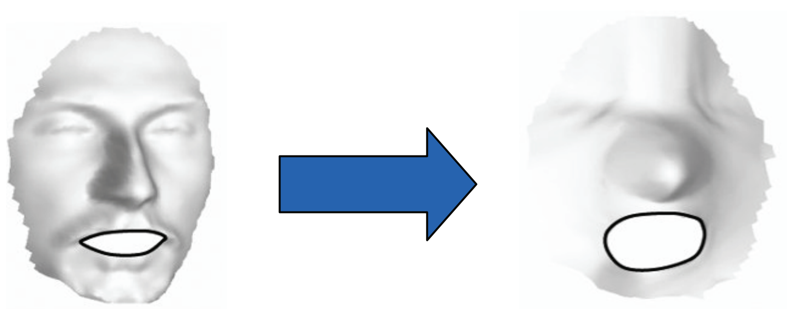
S. Malassiotis, M.G. Strintzis, “Robust face recognition using 2D and 3D data: Pose and illumination compensation”, *Pattern Recognition*, 2005.
F. Tsalakanidou, F. Forster, S. Malassiotis, M.G. Strintzis, “Real-time acquisition of depth and color images using structured light and its application to 3D face recognition”, *Real Time Imaging*, 2005.

Handling facial expressions

- Detecting-excluding deformable regions (e.g. mouth, cheeks)
- Expression invariant representations based on isometry assumption
- Decoupling identity from expression

Canonical forms

- Expression invariant representation based on geodesic distances
  - Geodesic distance: the length of the minimum length curve that connects two points
  - The face is an isometric surface → geodesic distances are preserved
  - Canonical form: a new surface where Euclidean distances between its points correspond to their geodesic distances in the face surface



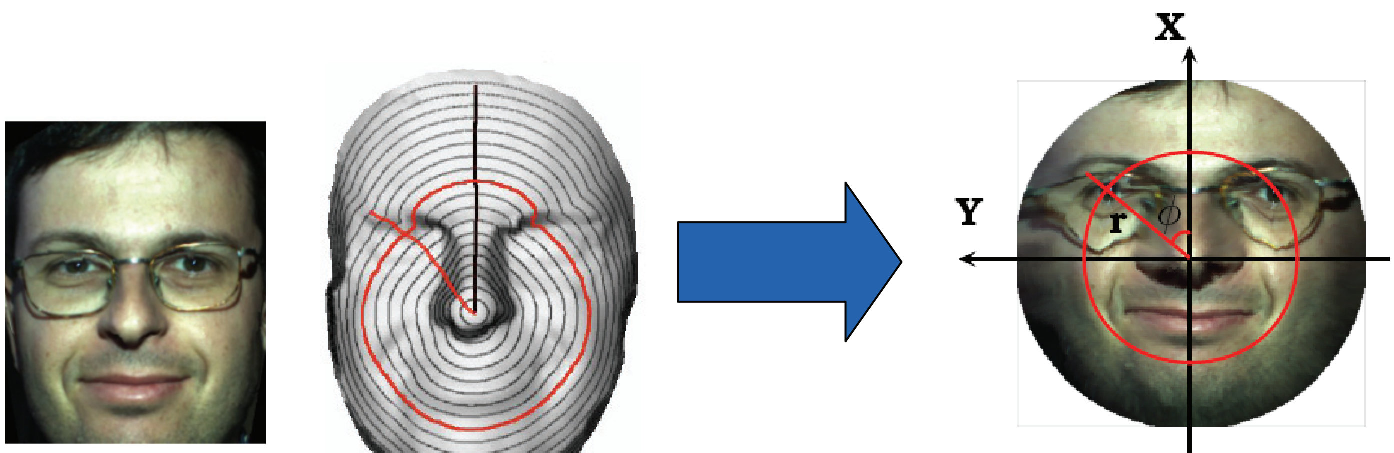
Face surface

Canonical form

A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Expression-invariant representations of faces", *IEEE Trans. on Image Processing,* 2007.

Geodesic polar images

- Expression invariant representation based on polar geodesic coordinates
  - Geodesic circle: the locus of points having the same geodesic distance from the pole (tip of nose)
  - Geodesic circles are mapped to circles on a new plane → 2D expression invariant images representing texture or Gaussian curvature
- Fast warping procedure, becomes a 2D face recognition problem
  - The isometry assumption stands for moderate expressions only…

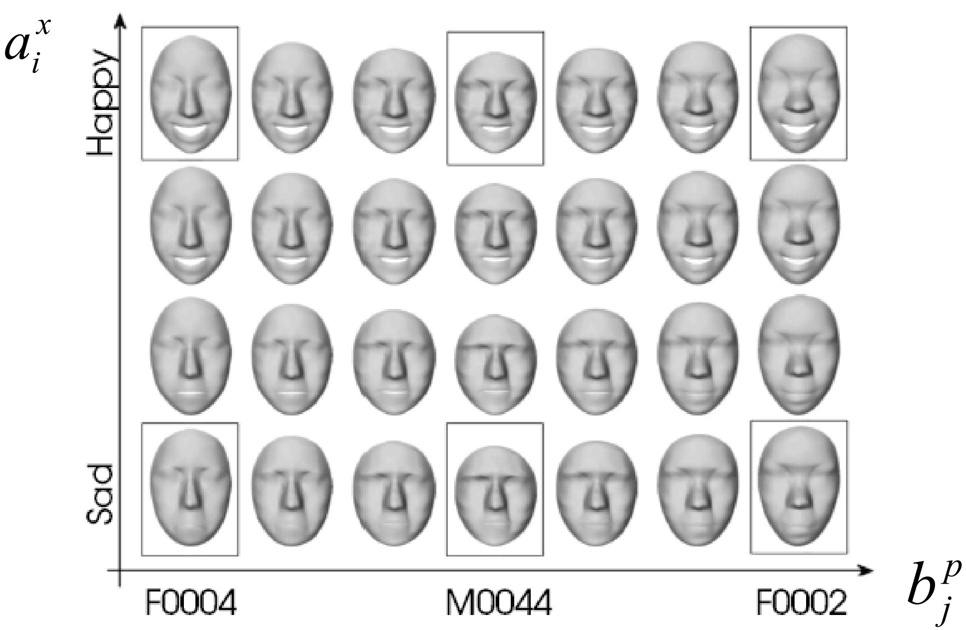


Geodesic polar image

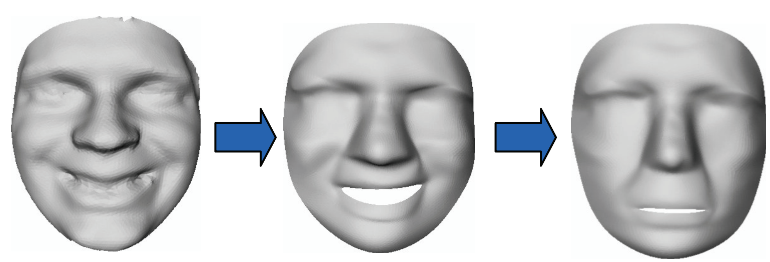
I. Mpiperis, S. Malassiotis and M.G.Strintzis, “3-D Face Recognition With the Geodesic Polar Representation”, *IEEE Trans. on Information Forensics and Security*, 2007

Bilinear models for 3D face recognition

- Basic idea: built a statistical model that effectively decouples identity and expression components of appearance
- Face classification is based on identity parameters



$$\mathbf{v}^{xp} = \sum_{i=1}^{N} \sum_{j=1}^{M} \mathbf{w}_{ij} a_i^x b_j^p$$



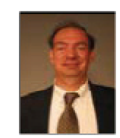
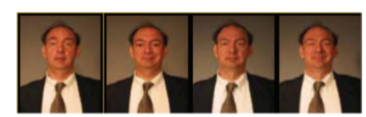
I. Mpiperis, S. Malassiotis and M.G.Strintzis, “Bilinear Models for 3D Face and Facial Expression Recognition”, *IEEE Trans. on Information Forensics and Security*, 2008.
J.B. Tenenbaum and W.T. Freeman, “Separating Style and Content with Bilinear Models”, *Neural Computation*, 2000.

Face Recognition Vendor Test 2006

- FRVT2002: Evaluation of commercial 2D face recognition systems
  - 10% FRR for 1% FAR, 18% FRR for 0.1% FAR
  - Outdoors illumination: Recognition rate drops more than 40% for images taken outdoors!
- FRVT2006: Evaluation of both 2D and 3D commercial face recognition systems
  - 2D: 2-5% FRR for 0.1% FAR
  - 3D: 0.5-3% FRR for 0.1% FAR
  - Uncontrolled illumination: 10-15% FRR for 0.1% FAR
  - Algorithm improvement, sensor improvement

http://www.frvt.org/DLs/FRVT_2002_Evaluation_Report.pdf, http://www.frvt.org/FRVT2006/docs/FRVT2006andICE2006LargeScaleReport.pdf

Face Recognition Grand Challenge (FRGC)

- Aim: Develop still and 3D algorithms to improve performance an order of magnitude over FRVT2002
  - FRVT2002 : verification rate 80% at 0.1% FAR
  - FRVT2006 : verification rate 98%
  - FRGC experiments results
    - 3D vs. 3D : verification rate 97% 
    - High resolution still vs. high resolution still : verification rate 99% 
    - Multi-still vs. multi-still : verification rate 99.99% 

http://www.frvt.org/FRGC/, FRGC presentation in Biometrics Consortium 2005: http://www.frvt.org/FRGC/FRGC_Phillips_BC2005.pdf

Some thoughts…

- Mature technology for biometric authentication especially if used together with other biometrics
- But still no commercial system working robustly under unconstrained conditions
  - Pose, illumination, expressions: open issues
- Unexplored ground: handling disguises and occlusions, exploiting information about hairstyle, aging, etc

Thank you.

Questions?