Biometrics

SBE 4022, Fall 2024

Professor

Ahmed M. Badawi

ambadawi@eng1.cu.edu.eg

http://scholar.google.com.eg/citations?user=r9pLu6EAAAAJ&hl=en

TA:

Mohamed Mutair

Mohamed_mutair@eng.cu.edu.eg

قسم الهندسة الحيوية
الطبية والمنظومات

جامعة القاهرة
كلية الهندسة

Biometrics Modalities:

- √ Introduction to Biometrics, features, and classification**,**
- √ Fingerprint verification,
- √ Face recognition,
- √ Hand geometry, hand veins, finger veins, pam veins
- √ Iris recognition,
- Signature, speaker, keystrokes dynamics verification,
- Gait recognition,
- √ Ear recognition,
- DNA based identification,
- Testing and evaluation of Biometric system,
- Multimodal systems and fusion on sensory, features, and decision levels,
- Biometrics ethics, applications, and current technologies.

What traits qualify to be a biometric?

Universality *Permanence*
Distinctiveness *Collectability*

Some other important requirements:

Performance
Acceptability
Circumvention

Biometric Technologies

1 (worst) ---------- 5 (best)

Technology	Accuracy	Convenience	Cost	Size
Fingerprint	5	5	4	4
Voice	1	5	5	5
Face	2	3	4	3
Hand	3	3	2	2
Iris	5	2	3	3

Face Recognition and Detection

Recognition problems

What is it?

- Object and scene recognition

Who is it?

- Identity recognition

Where is it?

- Object detection

What are they doing?

- Activities

All of these are **classification** problems

Variation
s

Illumination (light)

ORL database

Cairo 2000 database

Pose

Expression

Similarity

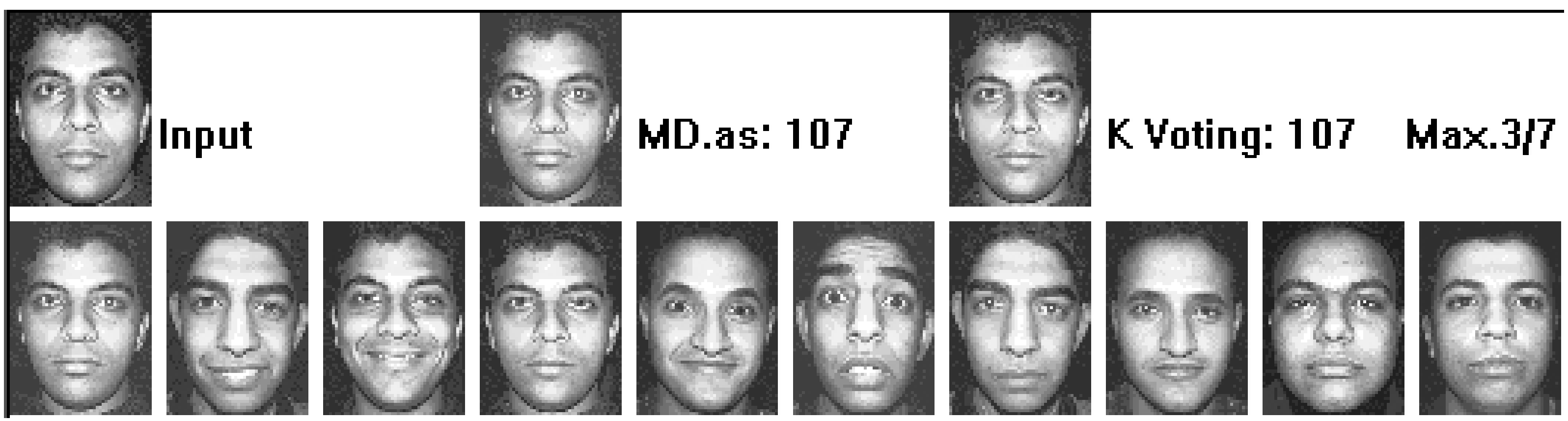
Eg.

Nearest neighbor (min Euclidean distance)

K-voting

NN or SVM

Bayes prob.



Biometrics capturing and recognition environment

Controlled or not

Illumination (light)
Pose
Expression (activity)
Camera sensor (cellphone or High res sensor)
Make up
Occlusions (face parts)

Face recognition is well matured under controlled environment

What is recognition?

A different taxonomy from [Csurka *et al.* 2006]:

Recognition

- Where is *this* particular object?

Categorization

- What *kind* of object(s) is(are) present?

Content-based image retrieval

- Find me something that looks similar

Detection

- Locate *all* instances of a given class

Some of following slides by Steve Seitz

Readings

- C. Bishop, “Neural Networks for Pattern Recognition”, Oxford University Press, 1998, Chapter 1.
- Forsyth and Ponce, Chap 22.3 (through 22.3.2--eigenfaces)
- Turk, M. and Pentland, A. *Eigenfaces for recognition*. Journal of Cognitive Neuroscience, 1991
- Viola, P. A. and Jones, M. J. (2004). Robust real-time face detection. *IJCV*, 57(2), 137–154.

Sources

- Steve Seitz, CSE 455/576, previous quarters
- Fei-Fei, Fergus, Torralba, CVPR'2007 course
- Efros, CMU 16-721 Learning in Vision
- Freeman, MIT 6.869 Computer Vision: Learning
- Linda Shapiro, CSE 576, Spring 2007

Today’s lecture

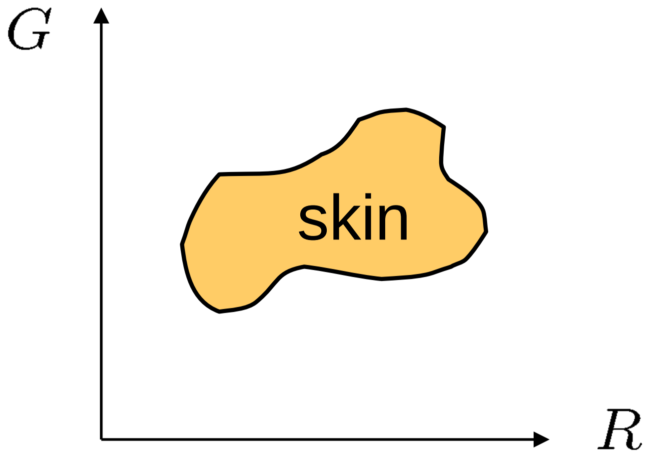
Face recognition and detection

- Color-based skin detection
- Recognition: eigenfaces [Turk & Pentland] and parts [Moghaddan & Pentland]
- Detection: boosting [Viola & Jones]

Face detection

How to tell if a face is present?

Skin detection



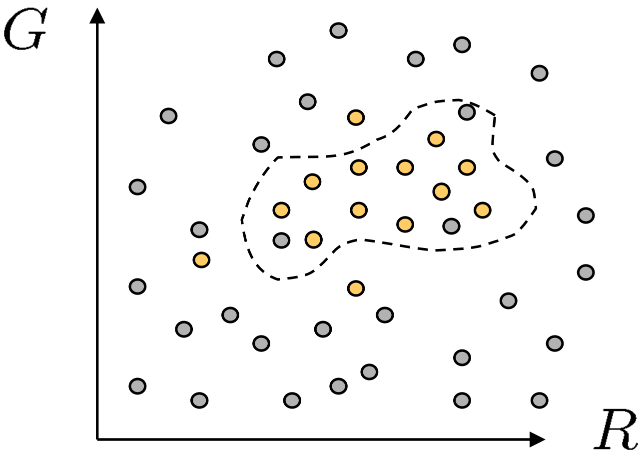
Skin pixels have a distinctive range of colors

- Corresponds to region(s) in RGB color space

Skin classifier

- A pixel X = (R,G,B) is skin if it is in the skin (color) region
- How to find this region?

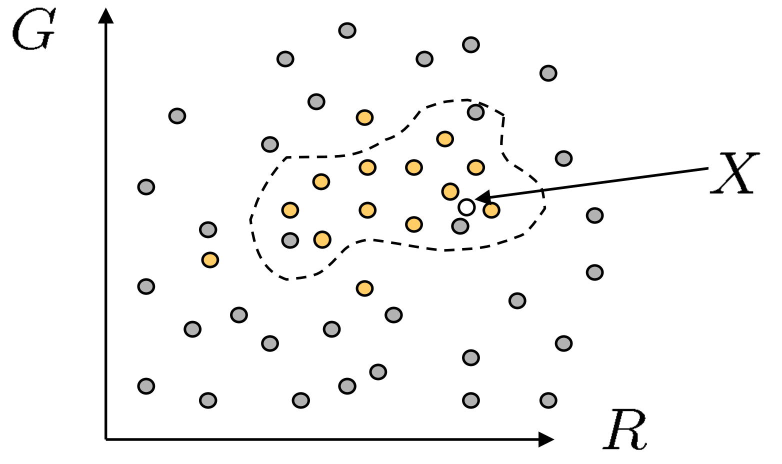
Skin detection



Learn the skin region from examples

- Manually label skin/non pixels in one or more “training images”
- Plot the training data in RGB space
  - skin pixels shown in orange, non-skin pixels shown in gray
  - some skin pixels may be outside the region, non-skin pixels inside.

Skin classifier

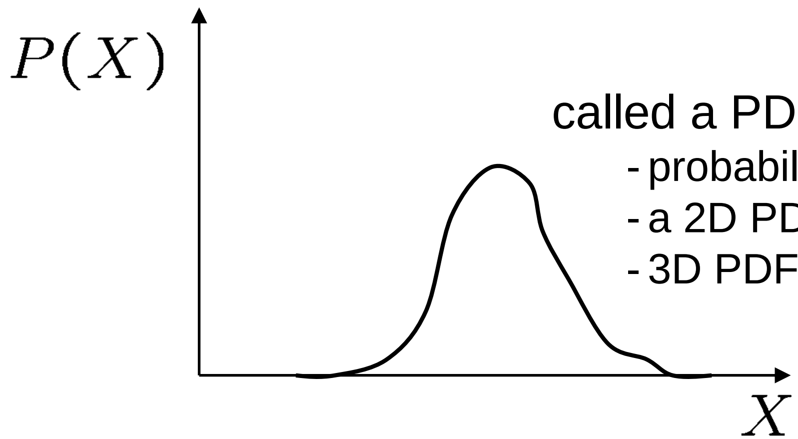


Given X = (R,G,B): how to determine if it is skin or not?

- Nearest neighbor
  - find labeled pixel closest to X
- Find plane/curve that separates the two classes
  - popular approach: Support Vector Machines (SVM)
- Data modeling
  - fit a probability density/distribution model to each class

Probability

- X is a random variable
- P(X) is the probability that X achieves a certain value



called a PDF

- probability distribution/density function
- a 2D PDF is a surface
- 3D PDF is a volume

$$0 \leq P(X) \leq 1$$

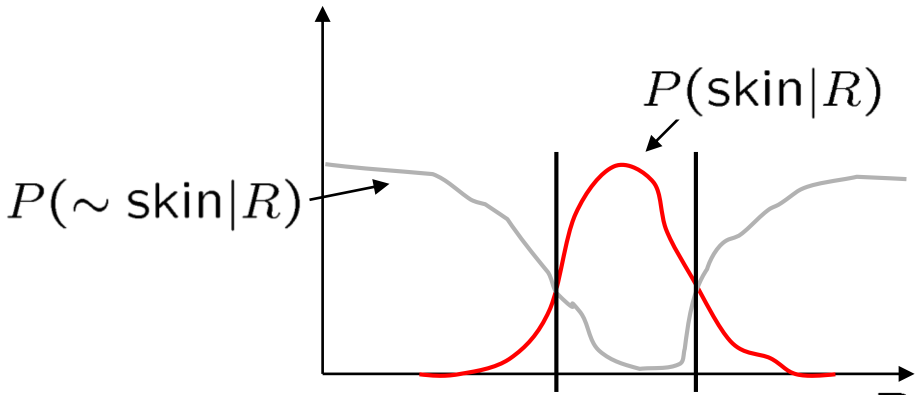
$$\int_{-\infty}^{\infty} P(X)dX = 1$$

continuous X

$$\sum P(X) = 1$$

discrete X

Probabilistic skin classification



Model PDF / uncertainty

- Each pixel has a probability of being skin or not skin

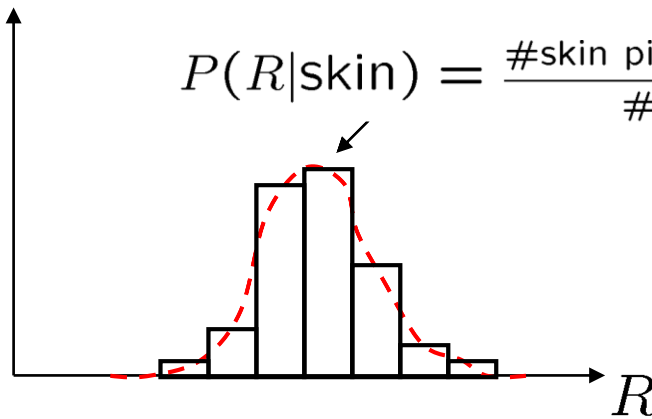
$$P(\sim \text{skin}|R) = 1 - P(\text{skin}|R)$$

Skin classifier

- Given X = (R,G,B): how to determine if it is skin or not?
- Choose interpretation of highest probability

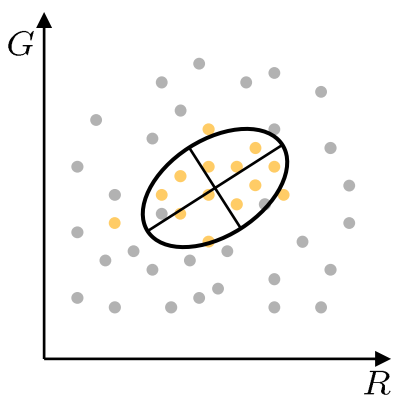
Where do we get $P(\text{skin}|R)$ and $P(\sim \text{skin}|R)$?

Learning conditional PDF's



We can calculate P(R | skin) from a set of training images

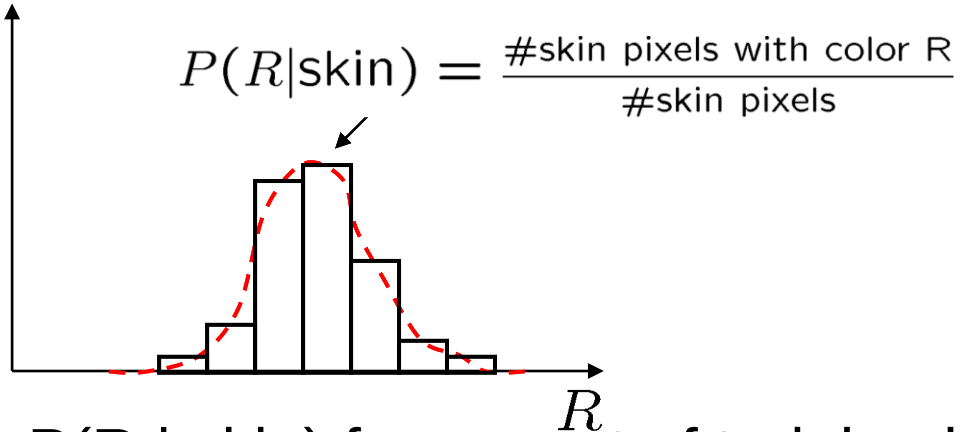
- It is simply a histogram over the pixels in the training images
  - each bin R_i contains the proportion of skin pixels with color R_i
- This doesn't work as well in higher-dimensional spaces. Why not?



Approach: fit parametric PDF functions

- common choice is rotated Gaussian
  - center $\mathbf{c} = \overline{X}$
  - covariance $\sum_X (X - \overline{X})(X - \overline{X})^T$

Learning conditional PDF's



We can calculate P(R | skin) from a set of training images

But this isn't quite what we want

- Why not? How to determine if a pixel is skin?
- We want P(skin | R) not P(R | skin)
- How can we get it?

Bayes rule

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

In terms of our problem:

what we measure (**likelihood**)

domain knowledge (**prior**)

$$P(\text{skin}|R) = \frac{P(R|\text{skin})\ P(\text{skin})}{P(R)}$$

what we want (**posterior**)

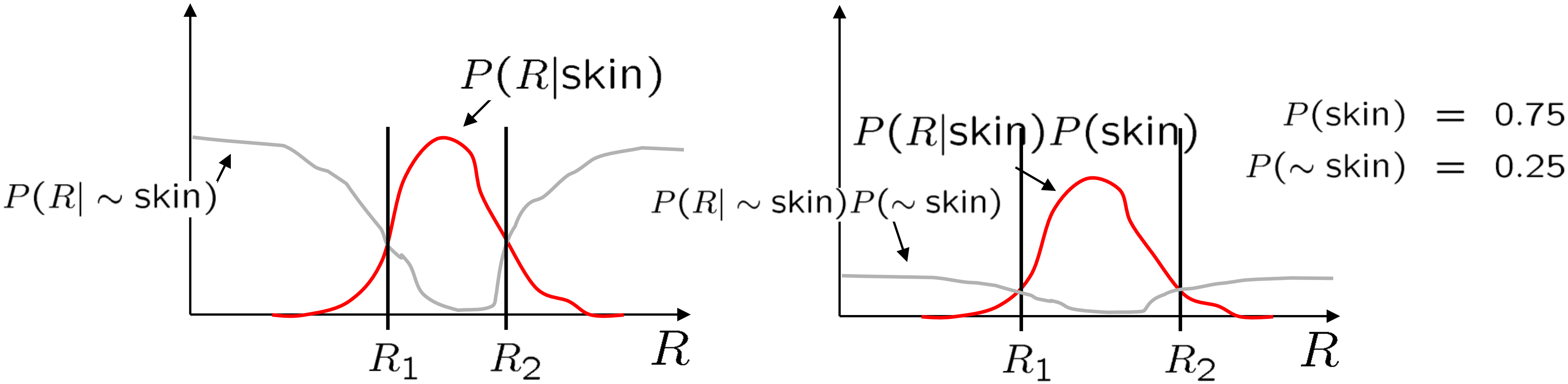
normalization term

$$P(R) = P(R|\text{skin})P(\text{skin}) + P(R|\sim\text{skin})P(\sim\text{skin})$$

What can we use for the prior P(skin)?

- Domain knowledge:
  - P(skin) may be larger if we know the image contains a person
  - For a portrait, P(skin) may be higher for pixels in the center
- Learn the prior from the training set. How?
  - P(skin) is proportion of skin pixels in training set

Bayesian estimation



likelihood

posterior (unnormalized)

Bayesian estimation

- Goal is to choose the label (skin or ~skin) that maximizes the posterior ↔ minimizes probability of misclassification
  – this is called **Maximum A Posteriori (MAP) estimation**

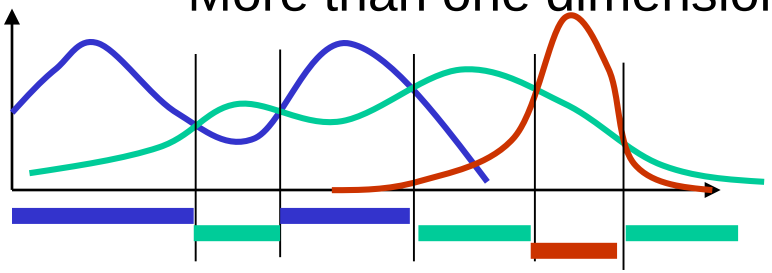
Skin detection results

Figure 25.3. The figure shows a variety of images together with the output of the skin detector of Jones and Rehg applied to the image. Pixels marked black are skin pixels, and white are background. Notice that this process is relatively effective, and could certainly be used to focus attention on, say, faces and hands. *Figure from "Statistical color models with application to skin detection," M.J. Jones and J. Rehg, Proc. Computer Vision and Pattern Recognition, 1999 © 1999, IEEE*

General classification

This same procedure applies in more general circumstances

- More than two classes
- More than one dimension



Example: face detection

- Here, X is an image region
  - dimension = # pixels
  - each face can be thought of as a point in a high dimensional space

H. Schneiderman, T. Kanade. "A Statistical Method for 3D Object Detection Applied to Faces and Cars". CVPR 2000

Today’s lecture

Face recognition and detection

- color-based skin detection
- recognition: eigenfaces [Turk & Pentland] and parts [Moghaddan & Pentland]
- detection: boosting [Viola & Jones]

Eigenfaces for recognition

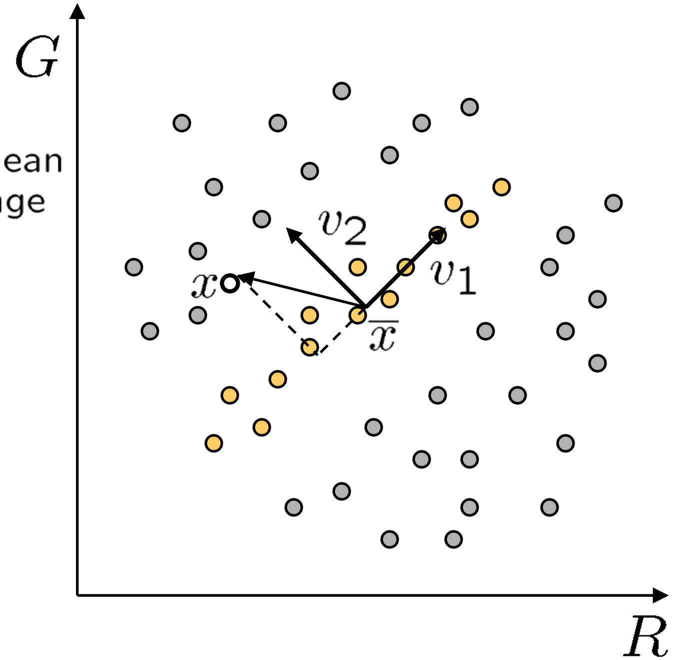
Matthew Turk and Alex Pentland

J. Cognitive Neuroscience

1991

Linear subspaces

$\overline{x}$ is the mean of the orange points



convert **x** into **v₁**, **v₂** coordinates

$$\mathbf{x} \to ((\mathbf{x} - \overline{x}) \cdot \mathbf{v_1}, (\mathbf{x} - \overline{x}) \cdot \mathbf{v_2})$$

What does the $\mathbf{v_2}$ coordinate measure?

- distance to line
- use it for classification—near 0 for orange pts

What does the $\mathbf{v_1}$ coordinate measure?

- position along line
- use it to specify which orange point it is

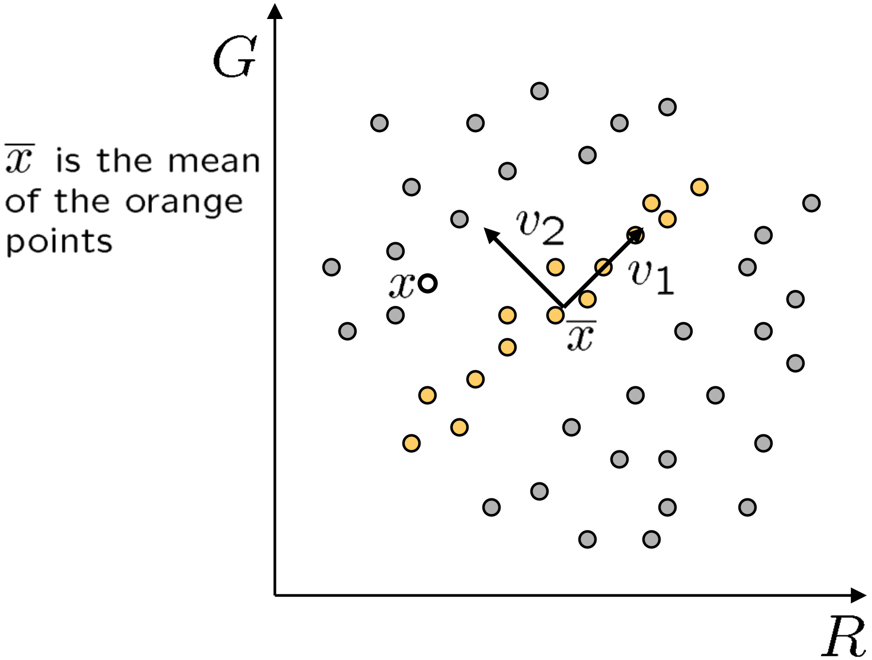
Classification can be expensive:

- Big search prob (e.g., nearest neighbors) or store large PDF's

Suppose the data points are arranged as above

- Idea—fit a line, classifier measures distance to line

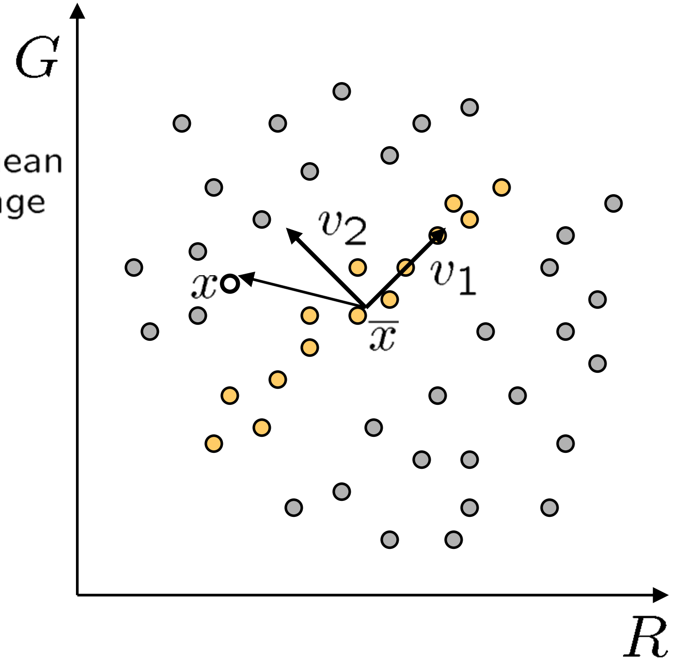
Dimensionality reduction



Dimensionality reduction

- We can represent the orange points with *only* their $\mathbf{v}_1$ coordinates (since $\mathbf{v}_2$ coordinates are all essentially 0)
- This makes it much cheaper to store and compare points
- A bigger deal for higher dimensional problems

Linear subspaces



$\overline{x}$ is the mean of the orange points

Consider the variation along direction **v** among all of the orange points:

$$var(\mathbf{v}) = \sum_{\text{orange point } \mathbf{x}} \|(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}} \cdot \mathbf{v}\|^{2}$$

What unit vector **v** minimizes *var*?

$$\mathbf{v}_2 = min_{\mathbf{v}} \ \{var(\mathbf{v})\}$$

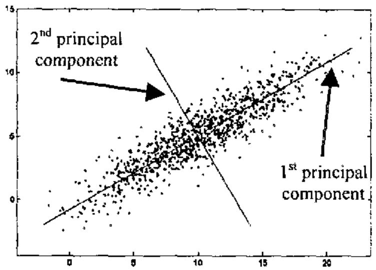
What unit vector **v** maximizes *var*?

$$\mathbf{v}_1 = max_{\mathbf{v}} \ \{var(\mathbf{v})\}$$

$$\begin{aligned} var(\mathbf{v}) &= \sum_{\mathbf{x}} \|(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}} \cdot \mathbf{v}\| \\ &= \sum_{\mathbf{x}} \mathbf{v}^{\mathbf{T}}(\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}}\mathbf{v} \\ &= \mathbf{v}^{\mathbf{T}}\left[\sum_{\mathbf{x}}(\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}}\right]\mathbf{v} \\ &= \mathbf{v}^{\mathbf{T}}\mathbf{A}\mathbf{v} \quad \text{where } \mathbf{A} = \sum_{\mathbf{x}}(\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}} \end{aligned}$$

Solution: $\mathbf{v}_1$ is eigenvector of **A** with *largest* eigenvalue
$\mathbf{v}_2$ is eigenvector of **A** with *smallest* eigenvalue

Principal component analysis



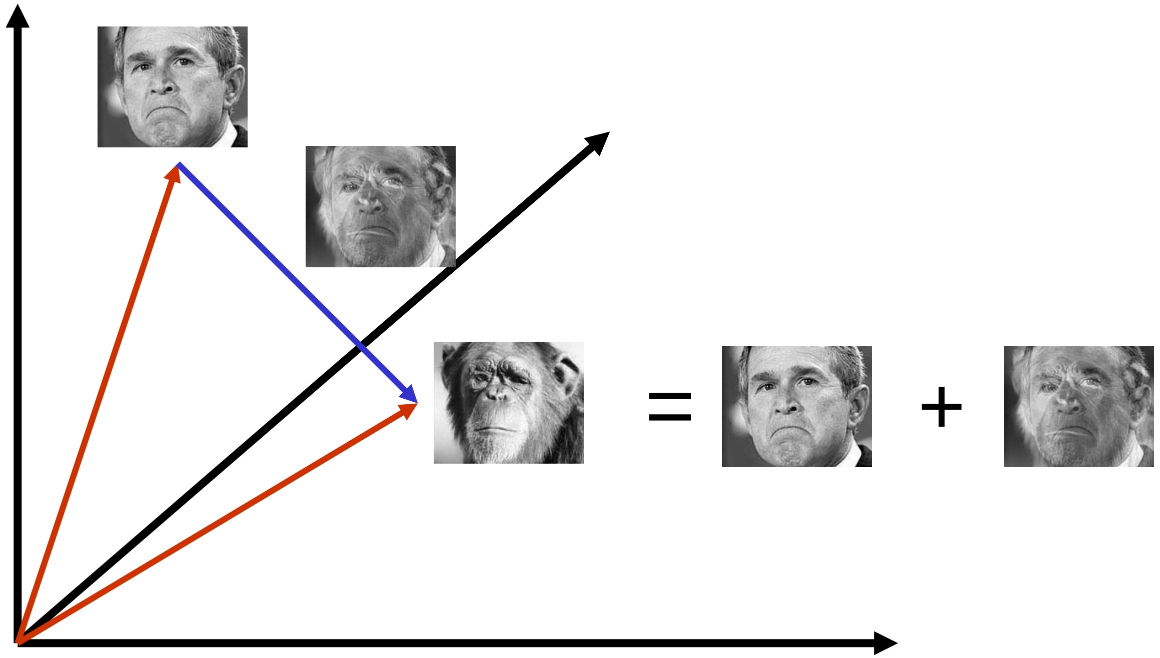
Suppose each data point is N-dimensional

- Same procedure applies:

$$\begin{aligned} var(\mathbf{v}) &= \sum_{\mathbf{x}} \|(\mathbf{x}-\overline{\mathbf{x}})^{\mathrm{T}} \cdot \mathbf{v}\| \\ &= \mathbf{v}^{\mathrm{T}}\mathbf{A}\mathbf{v} \quad \text{where } \mathbf{A} = \sum_{\mathbf{x}} (\mathbf{x}-\overline{\mathbf{x}})(\mathbf{x}-\overline{\mathbf{x}})^{\mathrm{T}} \end{aligned}$$

- The eigenvectors of **A** define a new coordinate system
  - eigenvector with largest eigenvalue captures the most variation among training vectors **x**
  - eigenvector with smallest eigenvalue has least variation
- We can compress the data using the top few eigenvectors
  - corresponds to choosing a “linear subspace”
    - » represent points on a line, plane, or “hyper-plane”
  - these eigenvectors are known as the ***principal components***

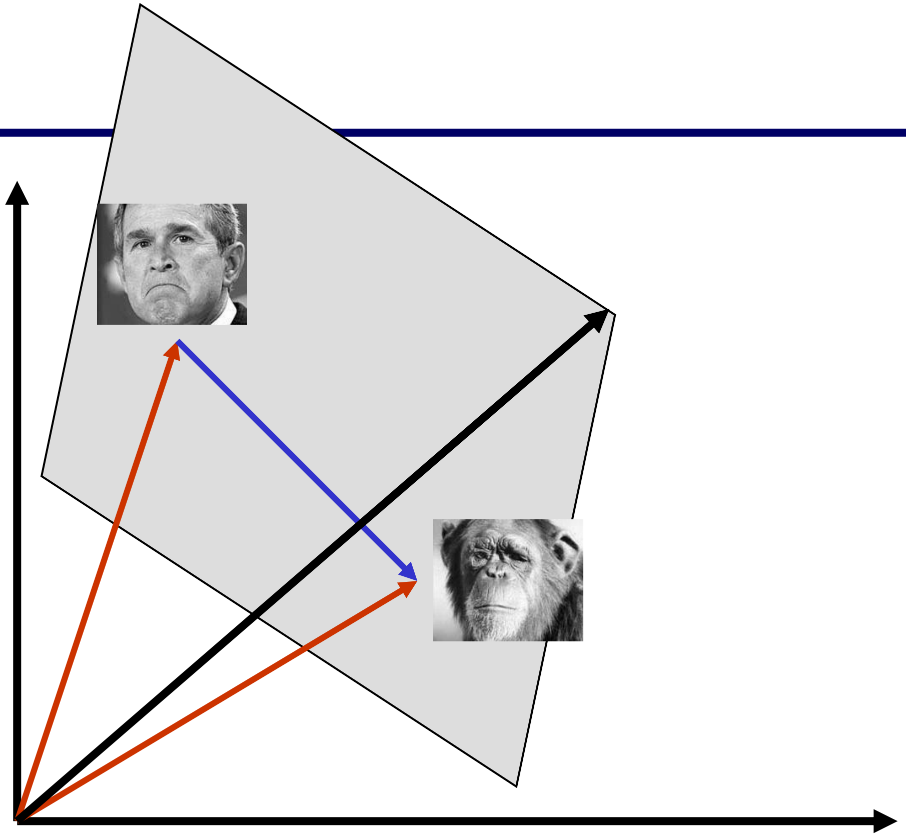
The space of faces



An image is a point in a high dimensional space

- An N x M image is a point in R^{NM}
- We can define vectors in this space as we did in the 2D case

Dimensionality reduction



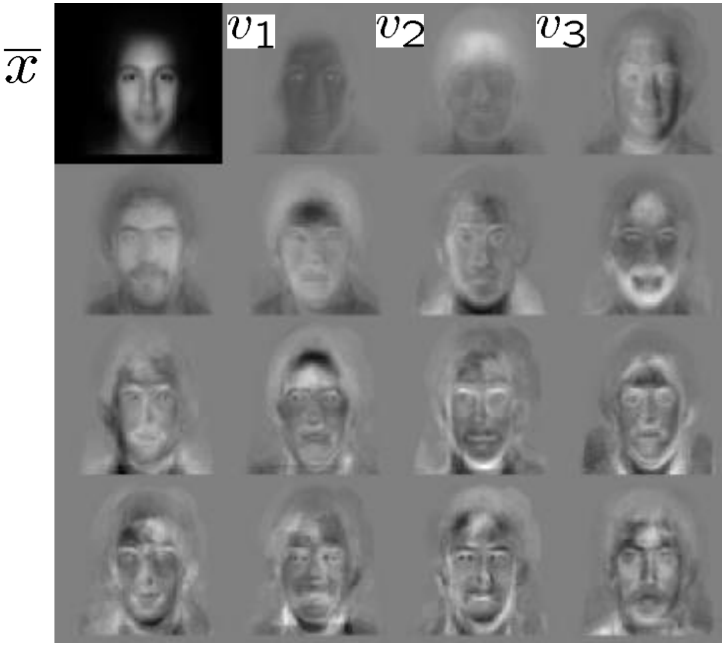
The set of faces is a "subspace" of the set of images

- We can find the best subspace using PCA
- This is like fitting a "hyper-plane" to the set of faces
  - spanned by vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_K}$
  - any face $\mathbf{x} \approx \overline{\mathbf{x}} + a_1\mathbf{v_1} + a_2\mathbf{v_2} + \ldots + a_k\mathbf{v_k}$

Eigenfaces

PCA extracts the eigenvectors of **A**

- Gives a set of vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \ldots$
- Each vector is a direction in face space
  - what do these look like?



Projecting onto the eigenfaces

The eigenfaces $\mathbf{v_1}, \ldots, \mathbf{v_K}$ span the space of faces

- A face is converted to eigenface coordinates by

$$\mathbf{x} \rightarrow (\underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v_1}}_{a_1}, \ \underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v_2}}_{a_2}, \ldots, \ \underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v_K}}_{a_K})$$

$$\mathbf{x} \approx \overline{\mathbf{x}} + a_1\mathbf{v_1} + a_2\mathbf{v_2} + \ldots + a_K\mathbf{v_K}$$

$\mathbf{x}$ $\quad a_1\mathbf{v_1}$ $\quad a_2\mathbf{v_2}$ $\quad a_3\mathbf{v_3}$ $\quad a_4\mathbf{v_4}$ $\quad a_5\mathbf{v_5}$ $\quad a_6\mathbf{v_6}$ $\quad a_7\mathbf{v_7}$ $\quad a_8\mathbf{v_8}$

Recognition with eigenfaces

Algorithm

1. Process the image database (set of images with labels)
   - Run PCA—compute eigenfaces
   - Calculate the K coefficients for each image
2. Given a new image (to be recognized) **x**, calculate K coefficients

$$\mathbf{x} \to (a_1, a_2, \ldots, a_K)$$

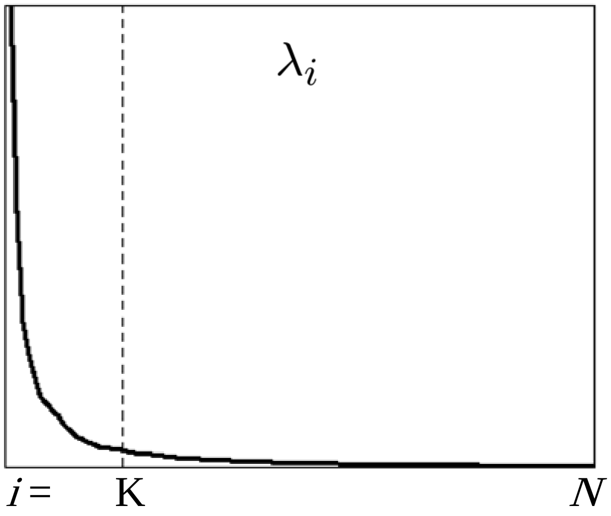
3. Detect if x is a face

$$\|\mathbf{x} - (\overline{\mathbf{x}} + a_1\mathbf{v_1} + a_2\mathbf{v_2} + \ldots + a_K\mathbf{v_K})\| < \text{threshold}$$

1. If it is a face, who is it?
   – Find closest labeled face in database
      » nearest-neighbor in K-dimensional space

Choosing the dimension K

eigenvalues



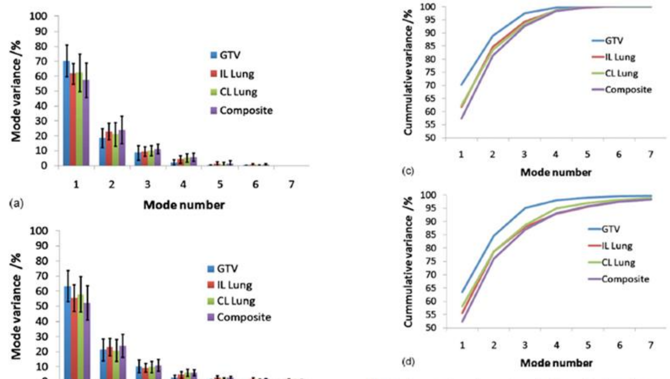


Fig. 5. Average eigenvalue spectrum (variance of each mode as a percentage of the total variance) for the retrospective W2W (a), allscans displaying only first 7/21 modes (b), average cumulative variance W2W (c), and cumulative variance allscans.

How many eigenfaces to use?

Look at the decay of the eigenvalues

- the eigenvalue tells you the amount of variance “in the direction” of that eigenface
- ignore eigenfaces with low variance

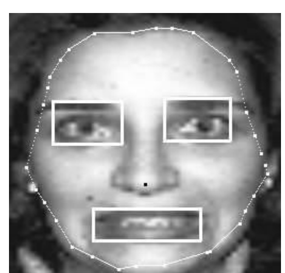
View-Based and Modular Eigenspaces for Face Recognition

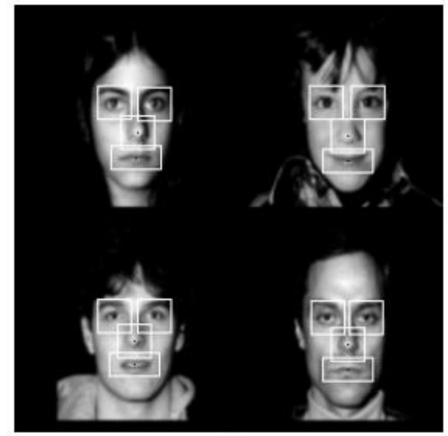
Alex Pentland, Baback Moghaddam and Thad Starner

CVPR'94

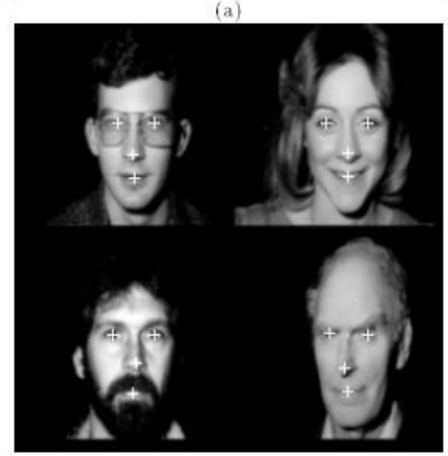
Part-based eigenfeatures

- Learn a separate eigenspace for each face feature
- Boosts performance of regular eigenfaces





(a)



PCA for data modeling and representation

- Intensities/colors I(x,y,z) --------- Appearance
- Geometries or vertex positions (X,Y,Z) ------ Shapes
- Shapes
- General data D(X,Y,Z,…..)
- Inter or intra modeling (eg modeling facial aging, geometries vs time or tumors vs time etc.)

Idea is each sample represents a point in Eigen space (Parameter space)

PCA for 3D Geometry Modeling example

3D ROI models have *homologous* point set correspondence as we warp each point and this point has the same order in all feature vectors (points are corresponding to each other)

[1] Ahmed M Badawi, Elisabeth Weiss, William C Sleeman IV, Chenyu Yan, Geoffrey D Hugo, **"Optimizing principal component models for representing interfraction variation in lung cancer radiotherapy"** Med Phys. 2010 September; 37(9): 5080–5091.

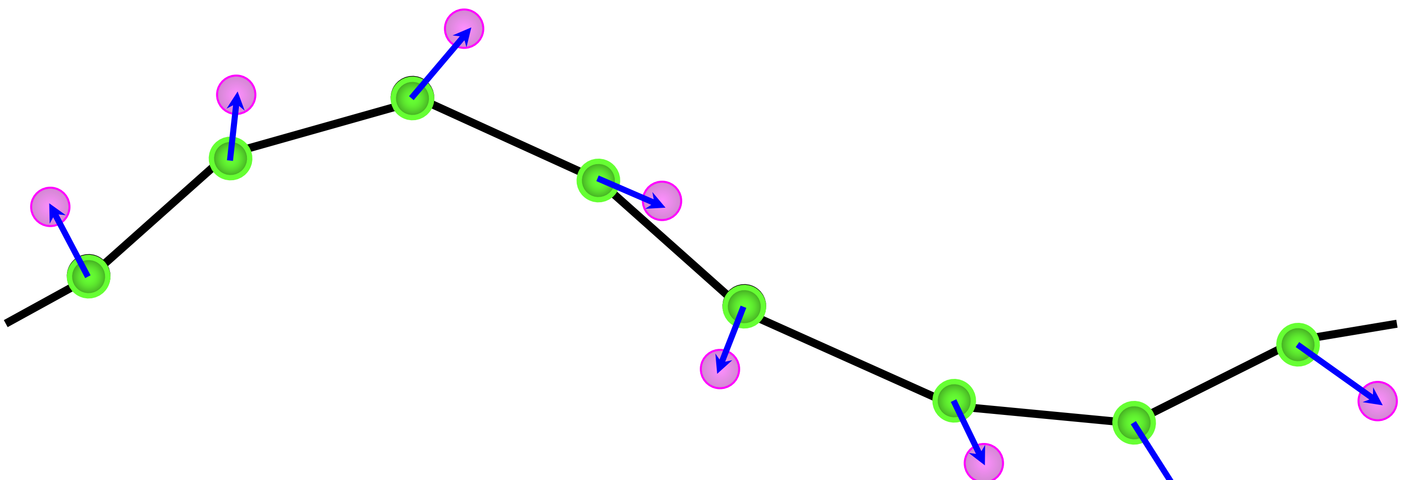
[2] Ahmed M Badawi, Elisabeth Weiss, William C Sleeman IV and Geoffrey D Hugo, "**Classifying geometric variability by dominant eigenmodes of deformation in regressing tumours during active breath-hold lung cancer radiotherapy**", 2012 *Phys. Med. Biol.* **57** 395

[3] M. Sohn, M. Birkner, D. Yan and M. Alber, "**Modelling individual geometric variation based on dominant eigenmodes of organ deformation: implementation and evaluation**," Phys Med Biol **50**, 5893-5908 (2005).

[4] Mohamed Mahfouz, Ahmed Badawi, Brandon Merkl, Emam E. Abdel Fatah, Emily Pritchard, Katherine Kesler, Megan Moore, Richard Jantz, Lee Jantz, **"Patella sex determination by 3D statistical shape models and nonlinear classifiers,"** Forensic Science International vol. 173,2007 , pages,161-170.

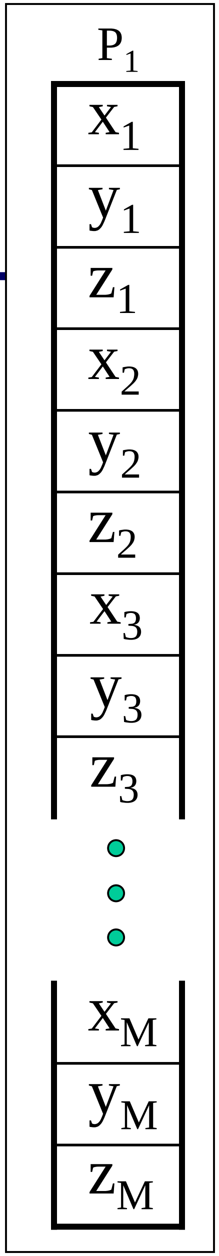
3D feature vector

- **Feature vector formation (3M size)**
- **Positions (X,Y,Z) or dvf (ΔX,ΔY,ΔZ)**



Green: (P_1) 3D Point position at week time, W_t
Purple: (P_2) Deformed (warped) Point position at W_{t+1}

P_i: I = 1 → 7 for W2W samples and I = 1 → 21 for all scans



Principal Components Analysis

Mean Shape subtraction

Samples (Observations)
Vectors

Mea
n

Variations from mean
matrix

$$\begin{bmatrix} p_1 & \vdots & p_2 & \vdots & \ldots & \vdots & p_N \end{bmatrix} - \begin{bmatrix} \bar{p} \end{bmatrix} = \underbrace{\begin{bmatrix} dp_1 & \vdots & dp_2 & \vdots & \ldots & \vdots & dp_N \end{bmatrix}}_{dp}$$

3Mx
N

Principal Components Analysis

In systems Eng. characteristic matrix, we solve for Ax=λx where A is the system characteristic matrix that have the control parameters (poles in system transfer function)

Eigen calculation

$$\underbrace{\frac{1}{N-1} dp\, dp^{\mathrm{T}}}_{COV} q_l = \lambda_l q_l$$

3Mx3M

3Mx1

Diagonalization of COV matrix (3Mx3M) results in eigenvectors q_l

q_l (3Mx1) is the l^{th} eigenvector of $\frac{1}{N-1} dp\, dp^{\mathrm{T}}$

λ_l is the eigenvalue associated with q_l

Eigenvalue =Statistical Variance $\sigma_l^2 = \lambda_l$

Singular Value Decomposition (SVD)

$$\mathrm{p}'(t) = \bar{\mathrm{p}} + \sum_{l=1}^{L} c_l(t)\mathrm{q}_l,$$

$$\mathrm{dp}(t) = \frac{1}{\sqrt{N-1}}(\mathrm{p}(t) - \bar{\mathrm{p}})^T,$$

Faster for large data

$$\mathrm{DP} = \{\mathrm{dp}(t_1)|\ldots\ldots|\mathrm{dp}(t_N)\}.$$

$$\mathrm{DP} = USV^T,$$

DP: 3MxN, S: NxN singular values, U: 3MxN containing eigenvectors, V:NxN

Construction of Organ Geometries Using Eigen Modes

Ranking λ_l (geometric variability)
Principal deformation modes (L) are the dominant eigen modes with largest eigen values (modes which span the space in which the majority of deformation occurs)

$$p = \overline{p} + \sum_{l=1}^{N-1} c_l q_l \qquad \|q_l\| = 1$$

$$p = \overline{p} + \sum_{l=1}^{L} c_l q_l + \varepsilon \qquad \|q_l\| = 1$$

- Deforming the mean shape by a weighted sum of L dominating eigenmodes
- c_l obey Gaussian distribution with corresponding λ_l as variances. Thus the dominating eigenmodes serve as statistical model of individual organ/deformation with only a small number of parameters.

Calculating Optimal Reconstruction Coefficients

PCA model representation:

Weighted sum of L dominant modes

$$p_{i,opt}^{[L]} = \overline{p} + \sum_{l=1}^{L} c_{l,opt}(i) q_l$$

- Optimal Coefficients Calculation:

$$c_{l,opt}(i) = (p_i - \overline{p}) \cdot q_l \qquad with \qquad \|q_l\| = 1; \ \ l = 1, \ldots, L.$$

L: is selected for sum of variance > 90% or 95% of total sum (in some better representations, it can be taken as 98% to minimize errors), selection of L depends on the resolution error in application as trade of with amount of reduction in model (speed of reconstruction)

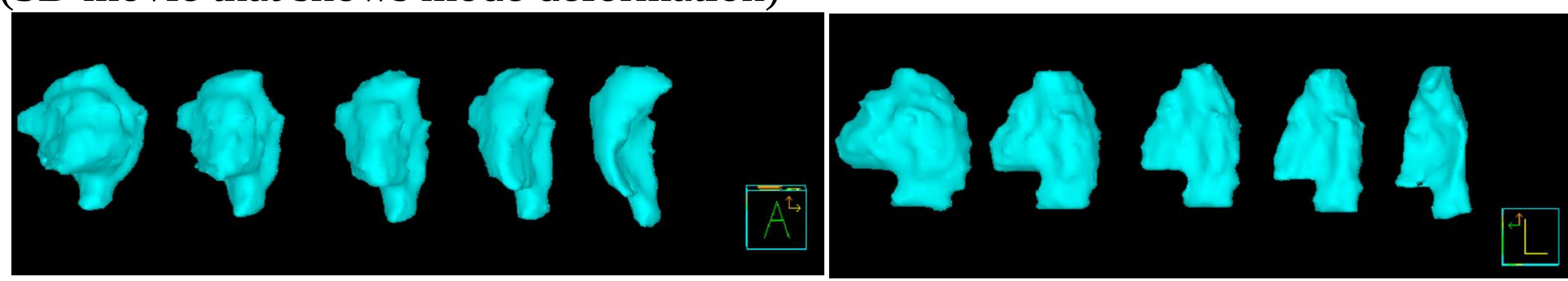
3D Reconstruction of Dominating Eigenmodes

Deformation of the mean geometry ($\bar{p}$) by the respective normed eignevector (q_l) is given by:

$$p_l = \bar{p} \pm \sigma_l . q_l$$

Eigenvalue =Statistical Variance $\qquad \sigma_l^2 = \lambda_l$

One can then construct an animation from $-\sigma$ to $+\sigma$ for small increments (3D movie that shows mode deformation)



2 Views of dominant mode 1 for subject 1 GTV
(Mode1 - 3σ, Mode1 – σ, Mean Mode, Mode1 + σ, Mode1 + 3σ)

Shape Similarity Quantization

Local representation error or local residual:

$$d_{i,j}^{[L]} \qquad i = 1,\ldots,N \quad , j = 1,\ldots,M$$

- Average local residual and stdev (Δd):

$$\bar{d}_{j}^{[L]} = \frac{1}{N}\sum_{i=1}^{N} d_{i,j}^{[L]} \qquad j = 1,\ldots,M$$

Histogram of these values gives an overview of the overall quality of representation.

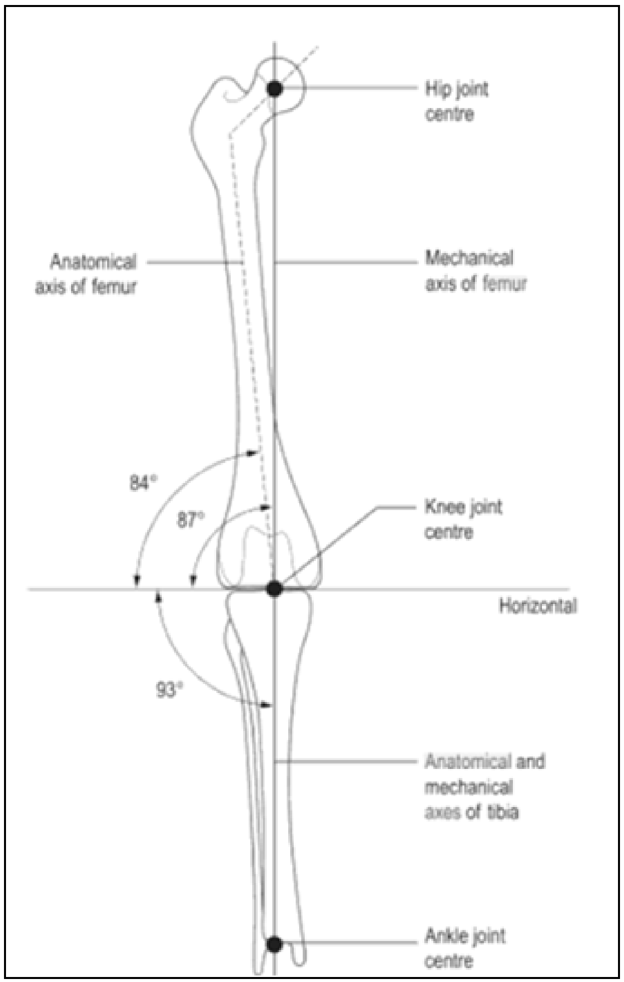
- Overall residual and overall stdev (ΔR):

$$R^{[L]} = \frac{1}{M}\sum_{j=1}^{M} \bar{d}_{j}^{[L]}$$

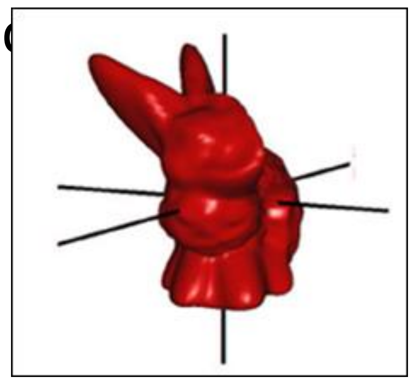
Overall error and stdev gives comprehensive measure for the quality of the PCA model with L eigenmodes

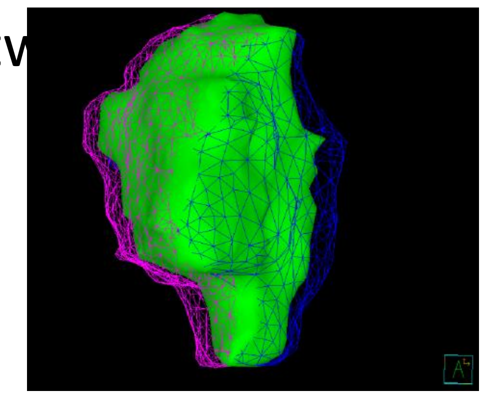
Rotation Calculations for 3D shapes (Rigid alignment)

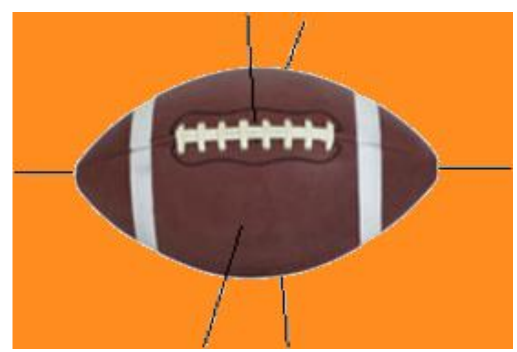
Anatomical and Mechanical axes are defined



- For undefined axes shapes:
- Axes of maximum variance (Principal axes) are used to calculate rotations between 2 ROI's
- Center of geometry is used to c … etw







Principal Eigen Rotation (Principal Axes)

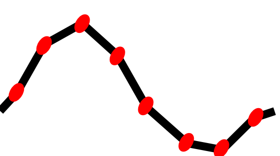
Due to lack of anatomical axes definition to GTV anatomy wrt to its geometry (unlike bones), we calculate rotations between 2 ROI's principal axes.

Principal axes are calculated by PCA using XYZ points as samples (Eigen Vectors are the principal axes)

$$P_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \quad P_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \quad P_3 = \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} \quad P_o = \begin{bmatrix} \bullet \\ \bullet \\ \bullet \end{bmatrix} \quad P_M = \begin{bmatrix} x_M \\ y_M \\ z_M \end{bmatrix}$$

$$\begin{bmatrix} p_1 & \vdots & p_2 & \vdots & \cdots & \vdots & p_M \end{bmatrix} - \begin{bmatrix} \bar{p} \end{bmatrix} = \underbrace{\begin{bmatrix} dp_1 & \vdots & dp_2 & \vdots & \cdots & \vdots & dp_M \end{bmatrix}}_{dp}$$

3x M



Morphable Face Models

Rowland and Perrett ’95

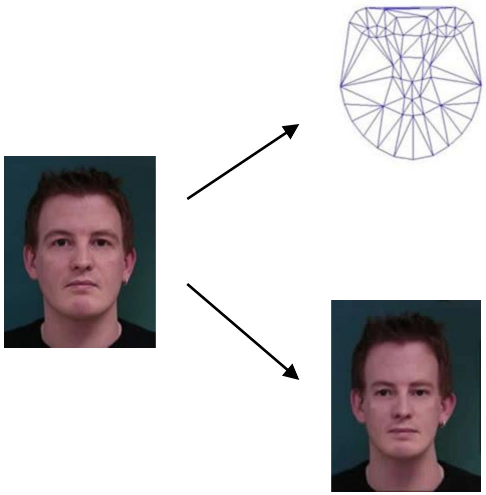
Lanitis, Cootes, and Taylor ’95, ’97

Blanz and Vetter ’99

Matthews and Baker ’04, ‘07

Morphable Face Model

Use subspace to model elastic 2D or 3D ***shape*** variation (vertex positions), in addition to *appearance* variation

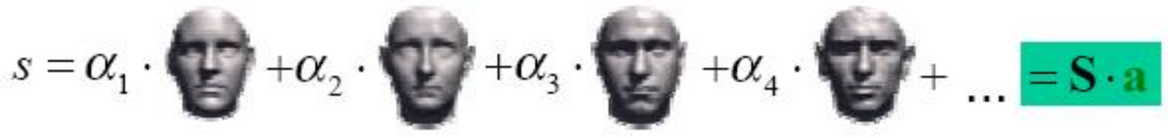


Shape S

Appearance T

Morphable Face Model

$$\mathbf{S}_{model} = \sum_{i=1}^{m} a_i \mathbf{S}_i \qquad \mathbf{T}_{model} = \sum_{i=1}^{m} b_i \mathbf{T}_i$$



3D models from Blanz and Vetter ‘99

Face Recognition Resources

Face Recognition Home Page:

- http://www.cs.rug.nl/~peterkr/FACE/face.html

PAMI Special Issue on Face & Gesture (July '97)

FERET

- http://www.dodcounterdrug.com/facialrecognition/Feret/feret.htm

Face-Recognition Vendor Test (FRVT 2000)

- http://www.dodcounterdrug.com/facialrecognition/FRVT2000/frvt2000.htm

Biometrics Consortium

- http://www.biometrics.org

Today’s lecture

Face recognition and detection

- color-based skin detection
- recognition: eigenfaces [Turk & Pentland] and parts [Moghaddan & Pentland]
- detection: boosting [Viola & Jones]

Example: Face Detection

Scan window over image

Classify window as either:

- Face
- Non-face

Then recognize face by Eigen analysis

Robust real-time face detection

Paul A. Viola and Michael J. Jones

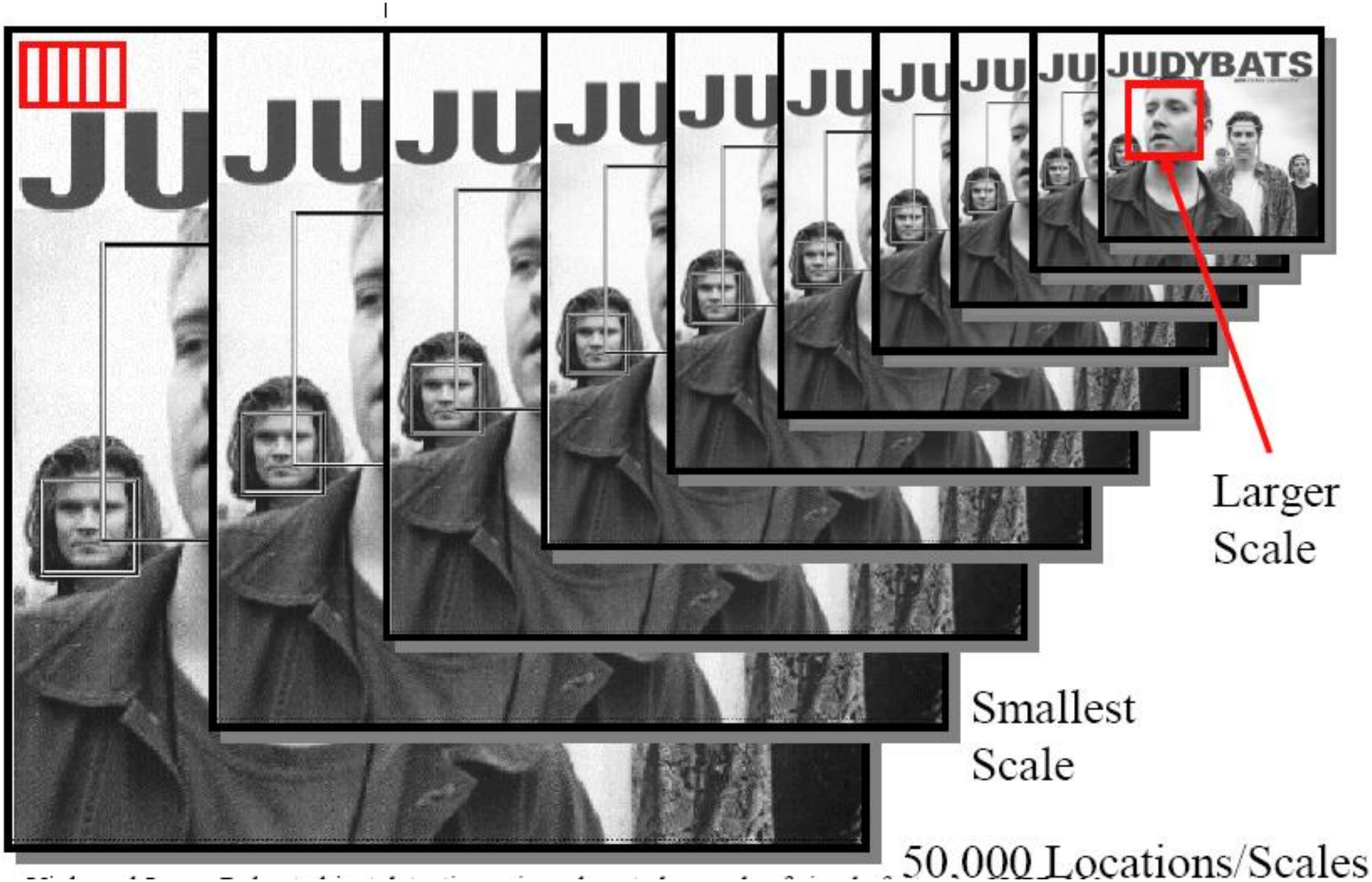
Intl. J. Computer Vision

57(2), 137–154, 2004

(originally in CVPR'2001)

(slides adapted from Bill Freeman, MIT 6.869, April 2005)

Scan classifier over locs. & scales



"Learn" classifier from data

Training Data

- 5000 faces (frontal)
- 10^8 non faces
- Faces are normalized
  - Scale, translation

Many variations

- Across individuals
- Illumination
- Pose (rotation both in plane and out)

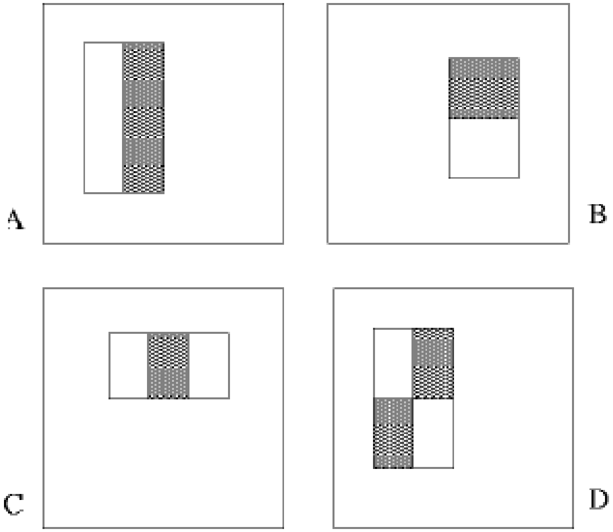
Characteristics of algorithm

- Feature set (…is huge about 16M features)
- Efficient feature selection using AdaBoost
- New image representation: Integral Image
- Cascaded Classifier for rapid detection

- Fastest known face detector for gray scale images

Image features

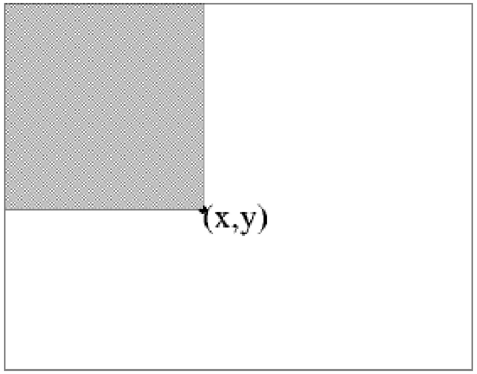
- "Rectangle filters"
  - Similar to Haar wavelet
- Differences between sums of pixels in adjacent rectangles



$$h_t(x) = \begin{cases} +1 & \text{if } f_t(x) > \theta_t \\ -1 & \text{otherwise} \end{cases}$$

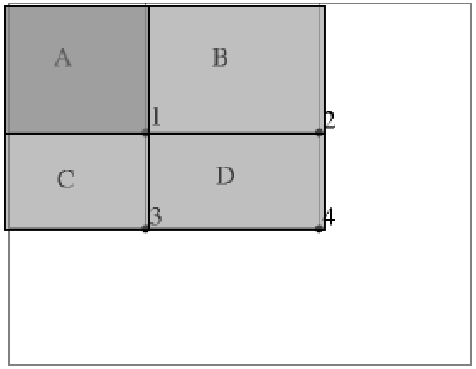
Integral Image

Partial sum $I'(x,y) = \sum_{\substack{x' \le x \\ y' \le y}} I(x', y')$



Any rectangle is

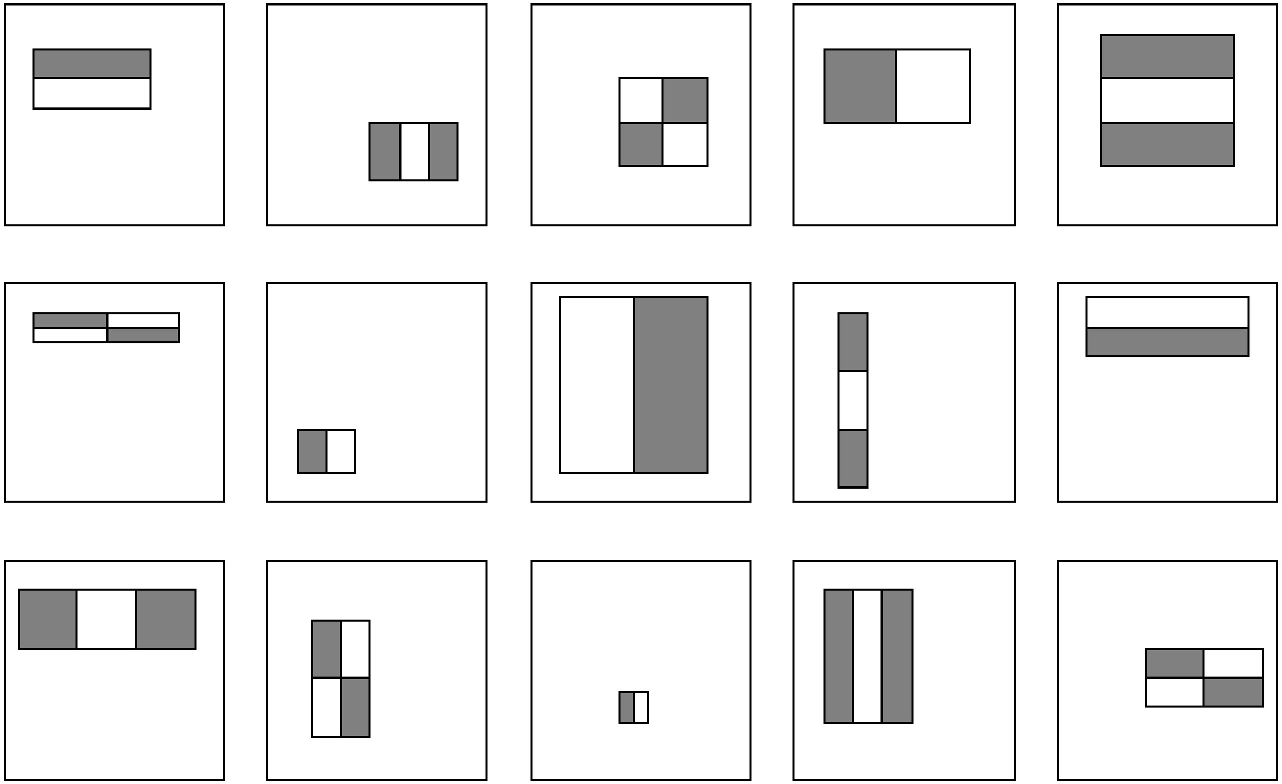
D = 1+4-(2+3)



Also known as:

- *summed area tables* [Crow84]
- *boxlets* [Simard98]

Huge library of filters



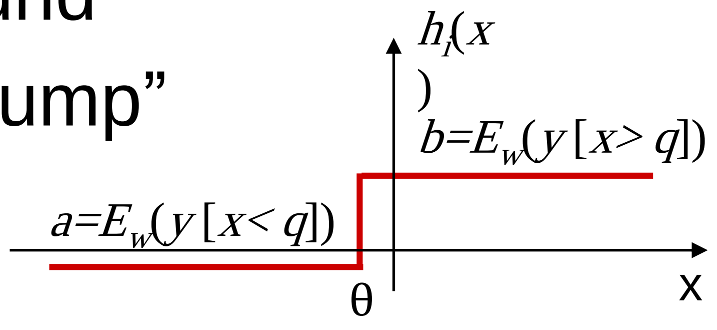
Constructing the classifier

Perceptron yields a sufficiently powerful classifier

$$C(x) = \theta\left(\sum_i \alpha_i h_i(x) + b\right)$$

Use AdaBoost to efficiently choose best features

- add a new $h_i(x)$ at each round
- each $h_i(x_k)$ is a "decision stump"



Constructing the classifier

For each round of boosting:

- Evaluate each rectangle filter on each example
- Sort examples by filter values
- Select best threshold for each filter (min error)

  Use sorting to quickly scan for optimal threshold
- Select best filter/threshold combination
- Weight is a simple function of error rate
- Reweight examples

  (There are many tricks to make this more efficient.)

Good reference on boosting

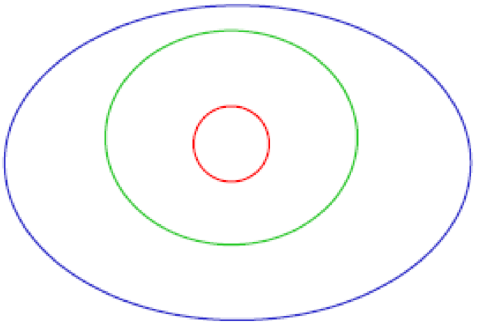
Friedman, J., Hastie, T. and Tibshirani, R. *Additive Logistic Regression: a Statistical View of Boosting*

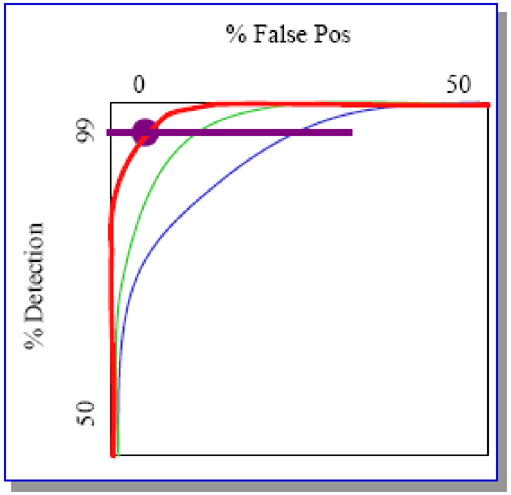
http://www-stat.stanford.edu/~hastie/Papers/boost.ps

"We show that boosting fits an additive logistic regression model by stagewise optimization of a criterion very similar to the log-likelihood, and present likelihood based alternatives. We also propose a multi-logit boosting procedure which appears to have advantages over other methods proposed so far."

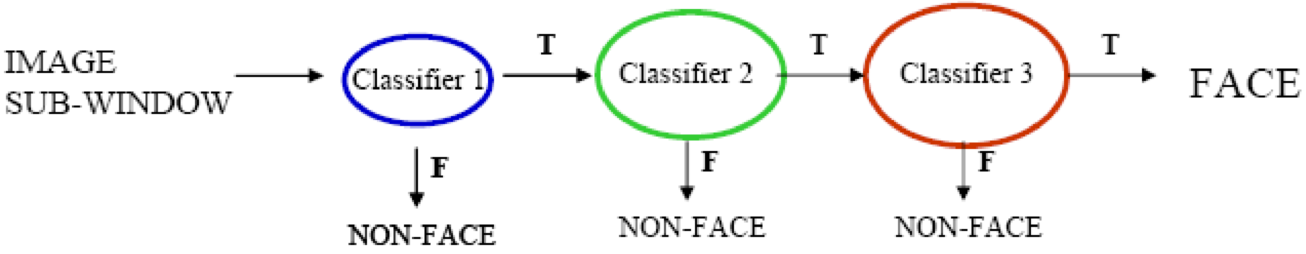
Trading speed for accuracy

Given a nested set of classifier hypothesis classes





Computational Risk Minimization



Speed of face detector (2001)

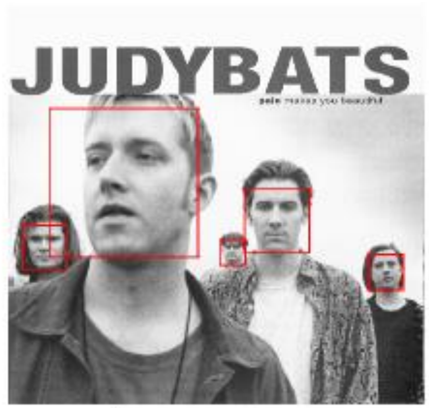
Speed is proportional to the average number of features computed per sub-window.

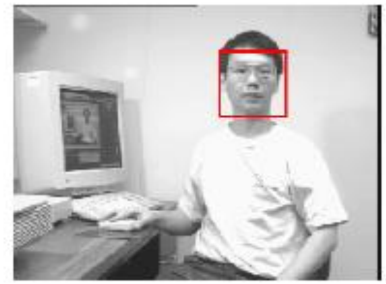
On the MIT+CMU test set, an average of 9 features (/ 6061) are computed per sub-window.

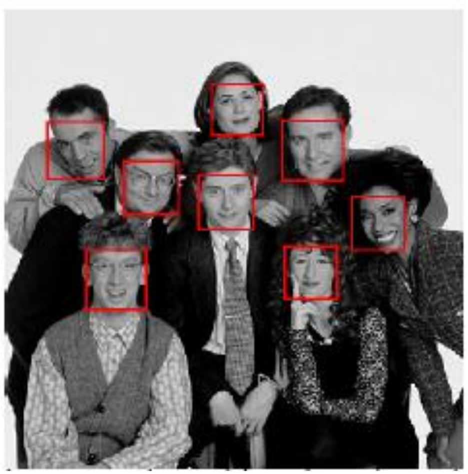
On a 700 Mhz Pentium III, a 384x288 pixel image takes about 0.067 seconds to process (15 fps).

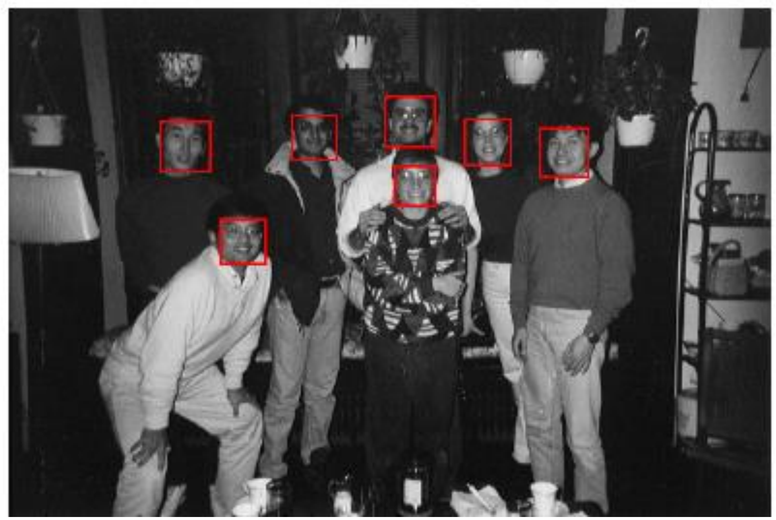
Roughly 15 times faster than Rowley-Baluja-Kanade and 600 times faster than Schneiderman-Kanade.

Sample results









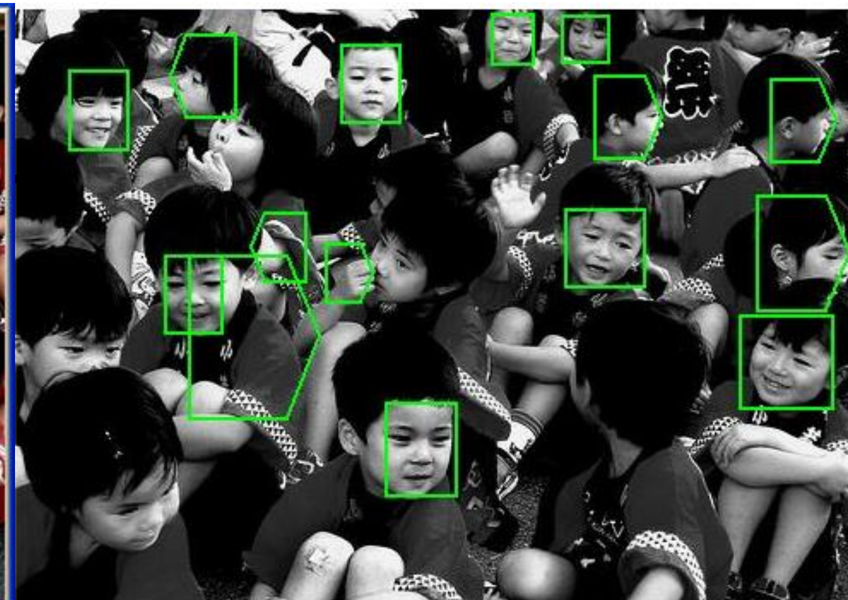
Summary (Viola-Jones)

- Fastest known face detector for gray images
- Three contributions with broad applicability:
  - ❖Cascaded classifier yields rapid classification
  - ❖AdaBoost as an extremely efficient feature selector
  - ❖Rectangle Features + Integral Image can be used for rapid image analysis

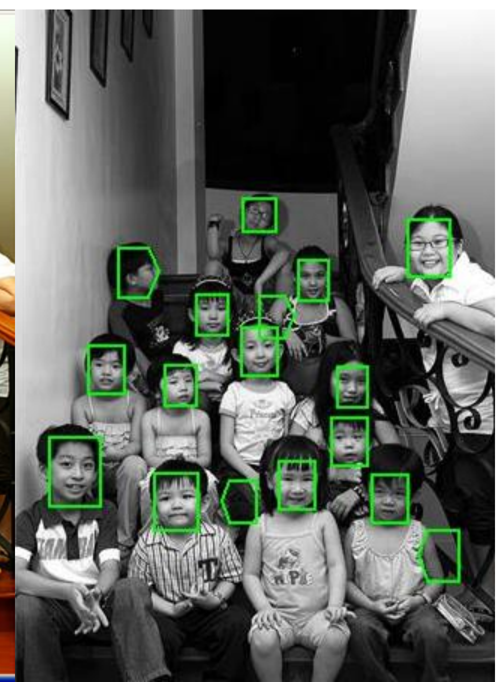
Face detector comparison

Informal study by Andrew Gallagher, CMU, for CMU 16-721 Learning-Based Methods in Vision, Spring 2007

- The Viola Jones algorithm OpenCV implementation was used. (<2 sec per image).
- For Schneiderman and Kanade, Object Detection Using the Statistics of Parts [IJCV'04], the www.pittpatt.com demo was used. (~10-15 seconds per image, including web transmission).

Viola
Jones

Schneiderman
Kanade

Today's lecture

Face recognition and detection

- color-based skin detection
- recognition: eigenfaces [Turk & Pentland] and parts [Moghaddan & Pentland]
- detection: boosting [Viola & Jones]

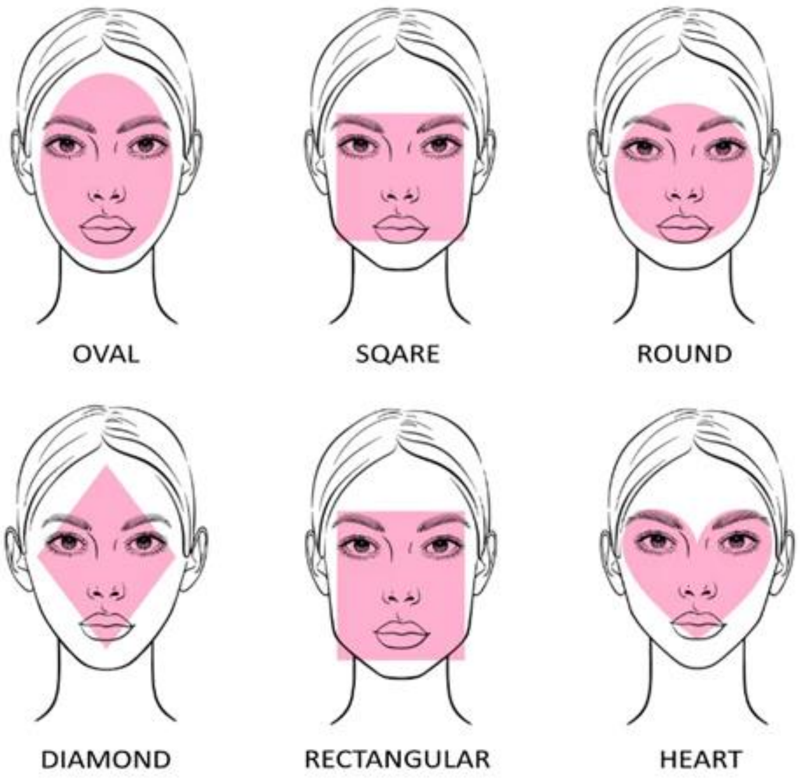
Now YOLOv? network has better performance for detection & for real time apps

Self study as example of detection

https://datascientest.com/en/you-only-look-once-yolo-what-is-it#:~:text=You%20Only%20Look%20Once%20or,the%20mainstays%20of%20computer%20vision.

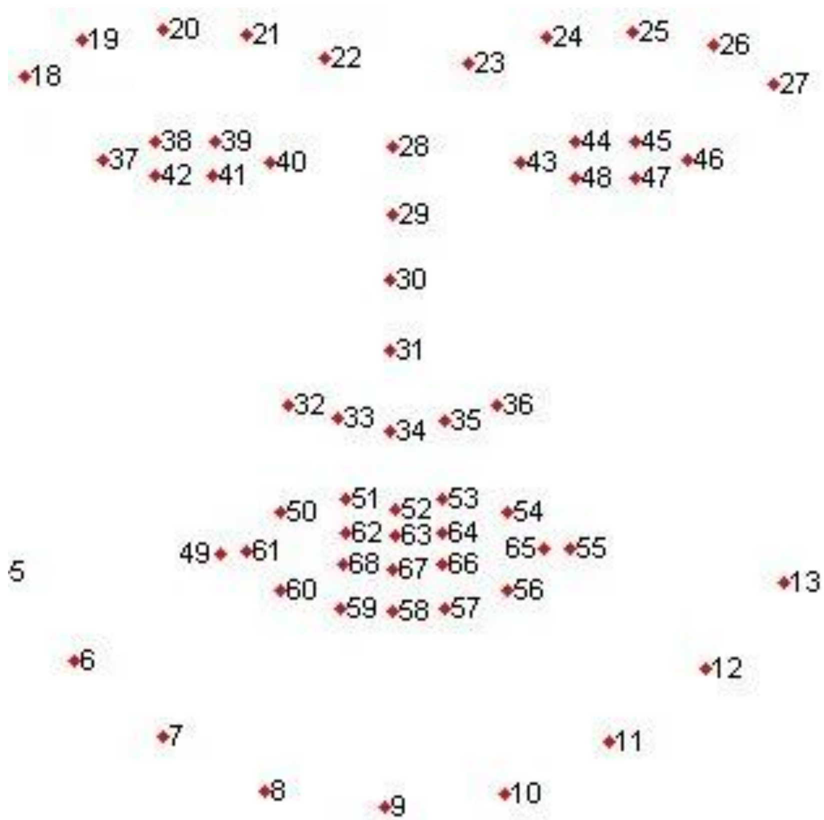
Few previous projects

- Face shape detection



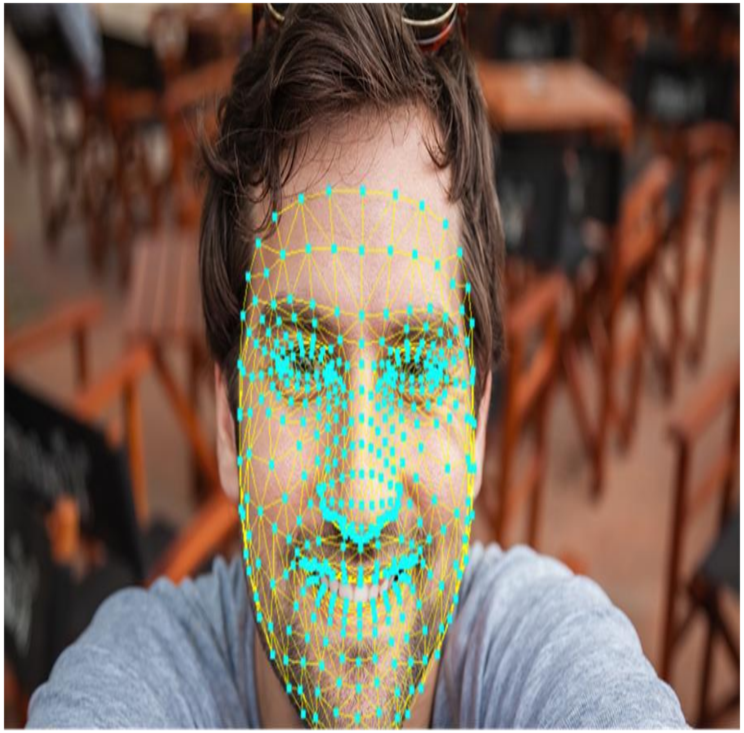
Face shape detection

1. Face detection
2. Facial landmarks detection



Facial landmarks detection

1. Dlib (68, 81 face landmarks)
2. Mediapipe library (face mesh)



Face parameters

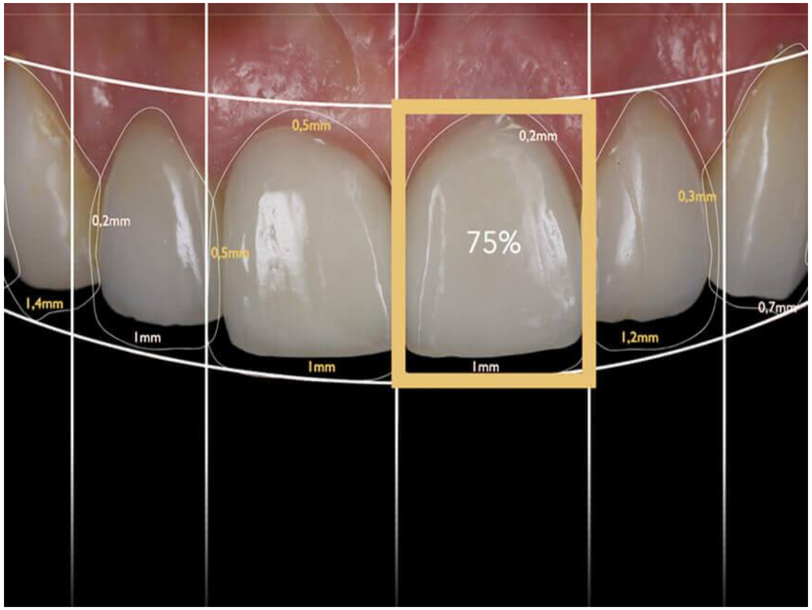
- Face length
- Forehead length
- Jawline length
- Chin width
- Cheekbone width
- Angle a1
- Angle a2
- Angle a3

Face shape detection

Face Shape	Condition
Heart	Forehead Width > Cheekbone Width > Jawline && a1>a2>a3
Oblong	Face Length > (Cheekbone Width ≈ Forehead Width ≈ Jawline) && a1 ≈ 90>a2>a3
Oval	Face Length > Cheekbone Width & Forehead Width > Jawline && a1 ≈a2≈a3
Square	Face Length ≈ Cheekbone Width ≈ Forehead Witch ≈ Jawline && a1 ≈ 90>a2 ≈ a3
Round	(Face Length ≈ Cheekbone Width) > (Forehead Width ≈ Jawline)

Apps and few Previous projects

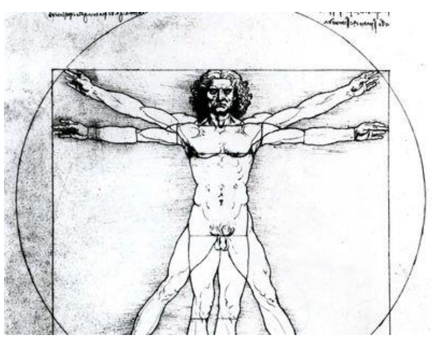
- Digital smile design and Golden Ratio
- Age and gender classification
- Beauty index and the Golden ratio (used in facial and dental surgery and alignment)

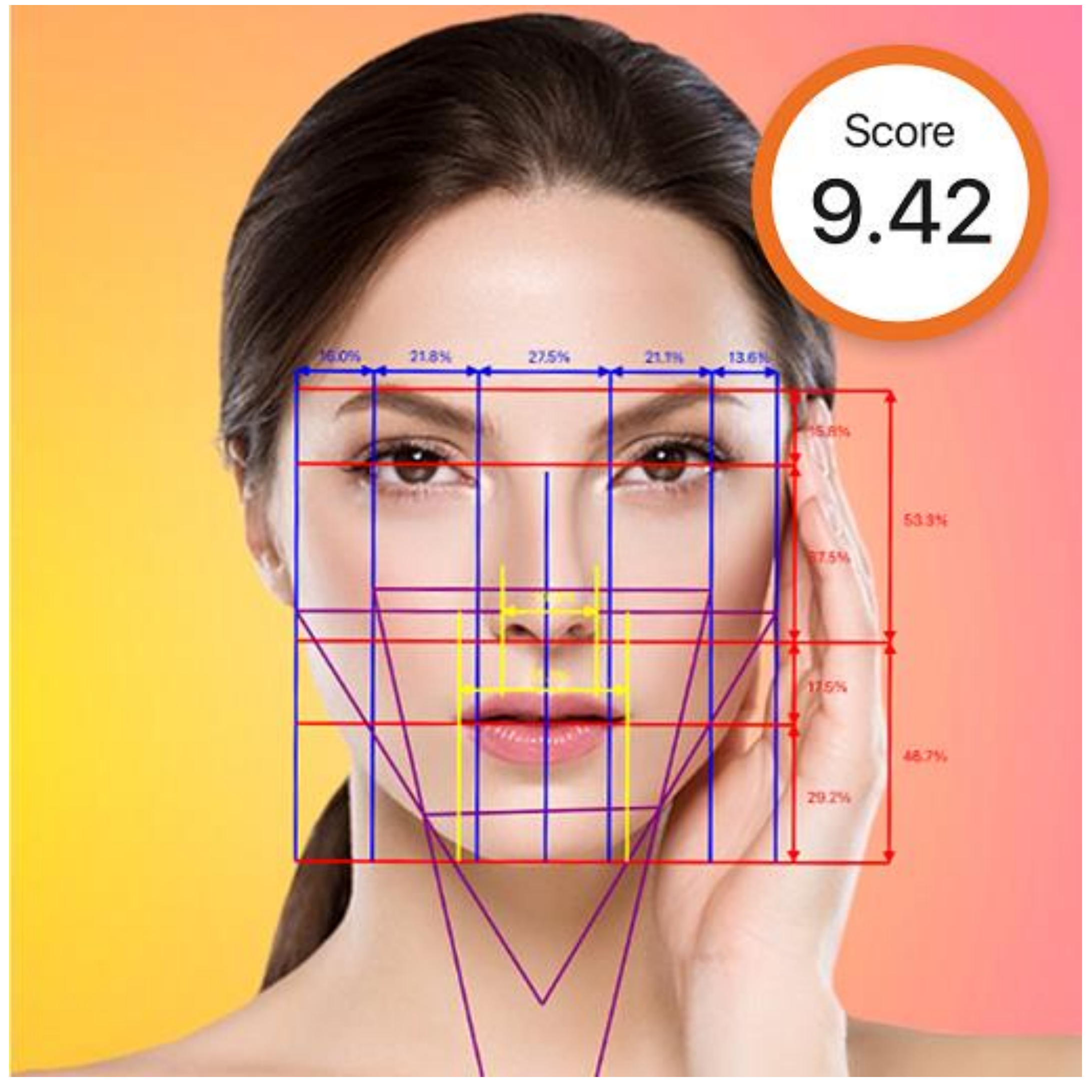


Beauty Score and Golden Ratio

It is suggested that a face is perceived as more aesthetically pleasing when its width is about 1.618 times the width of the mouth. $(1 + \sqrt{5})/2$, often denoted by the Greek letter ϕ.

And so on for eyes, nose, mouse, ..and teeth.





Digital Smile Design
(Dental Midline)

Biometrics

SBE 4022, Fall 2024

Professor

Ahmed M. Badawi

ambadawi@eng1.cu.edu.eg

http://scholar.google.com.eg/citations?user=r9pLu6EAAAAJ&hl=en

TA:

Mohamed Mutair

Mohamed_mutair@eng.cu.edu.eg

قسم الهندسة الحيوية الطبية والمنظومات

جامعة القاهرة
كلية الهندسة

Biometrics Modalities:

- √ Introduction to Biometrics, features, and classification**,**
- √ Fingerprint verification,
- √ Face recognition,
- √ Hand geometry, hand veins, finger veins, pam veins
- √ Iris recognition,
- Signature, speaker, keystrokes dynamics verification,
- Gait recognition,
- √ Ear recognition,
- DNA based identification,
- Testing and evaluation of Biometric system,
- Multimodal systems and fusion on sensory, features, and decision levels,
- Biometrics ethics, applications, and current technologies.

What traits qualify to be a biometric?

Universality *Permanence*

Distinctiveness *Collectability*

Some other important requirements:

Performance

Acceptability

Circumvention

Biometric Technologies

1 (worst) ----------- 5 (best)

Technology	Accuracy	Convenience	Cost	Size
Fingerprint	5	5	4	4
Voice	1	5	5	5
Face	2	3	4	3
Hand	3	3	2	2
Iris	5	2	3	3

Face Recognition and Detection

Recognition problems

What is it?

  Object and scene recognition

Who is it?

  Identity recognition

Where is it?

  Object detection

What are they doing?

  Activities

All of these are **classification** problems

Variations

Illumination (light)

ORL database

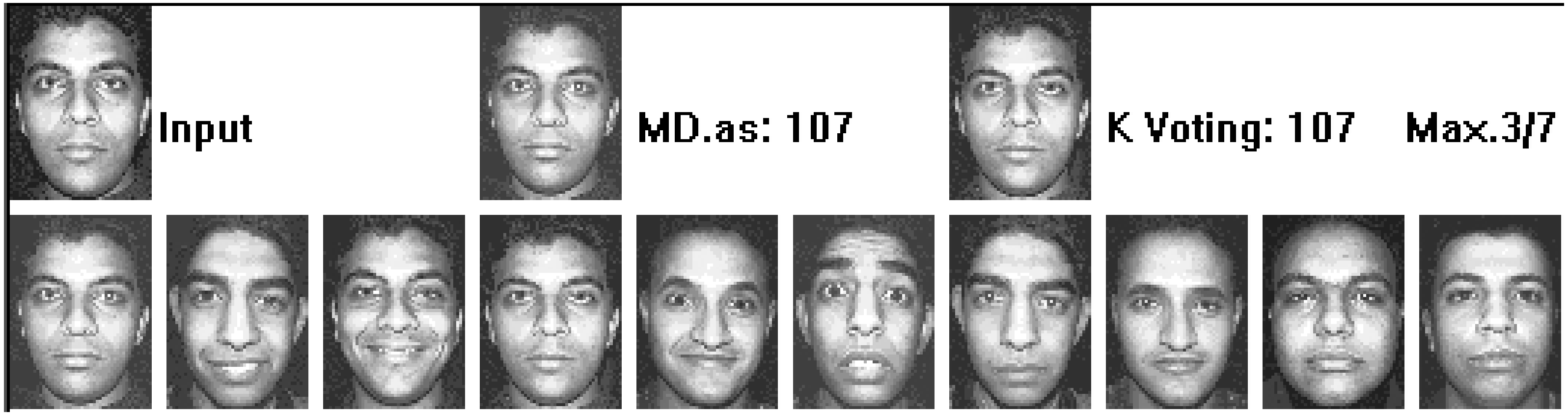
Cairo 2000 database

Pose

Expression

Similarity

Eg.
Nearest neighbor (min Euclidean distance)
K-voting
NN or SVM
Bayes prob.



Biometrics capturing and recognition environment

Controlled or not

Illumination (light)
Pose
Expression (activity)
Camera sensor (cellphone or High res sensor)
Make up
Occlusions (face parts)

Face recognition is well matured under controlled environment

What is recognition?

A different taxonomy from [Csurka *et al.* 2006]:

Recognition

- Where is *this* particular object?

Categorization

- What *kind* of object(s) is(are) present?

Content-based image retrieval

- Find me something that looks similar

Detection

- Locate *all* instances of a given class

Some of following slides by Steve Seitz

Readings

- C. Bishop, “Neural Networks for Pattern Recognition”, Oxford University Press, 1998, Chapter 1.
- Forsyth and Ponce, Chap 22.3 (through 22.3.2--eigenfaces)
- Turk, M. and Pentland, A. *Eigenfaces for recognition*. Journal of Cognitive Neuroscience, 1991
- Viola, P. A. and Jones, M. J. (2004). Robust real-time face detection. *IJCV*, 57(2), 137–154.

Sources

- Steve Seitz, CSE 455/576, previous quarters
- Fei-Fei, Fergus, Torralba, CVPR’2007 course
- Efros, CMU 16-721 Learning in Vision
- Freeman, MIT 6.869 Computer Vision: Learning
- Linda Shapiro, CSE 576, Spring 2007

Today's lecture

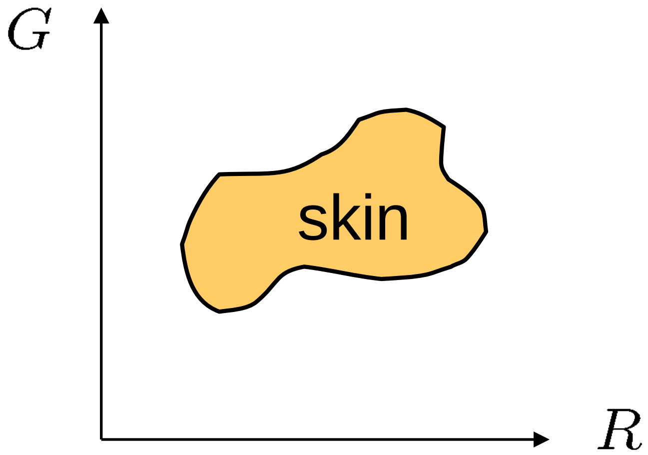
Face recognition and detection

- Color-based skin detection
- Recognition: eigenfaces [Turk & Pentland] and parts [Moghaddan & Pentland]
- Detection: boosting [Viola & Jones]

Face detection

How to tell if a face is present?

Skin detection



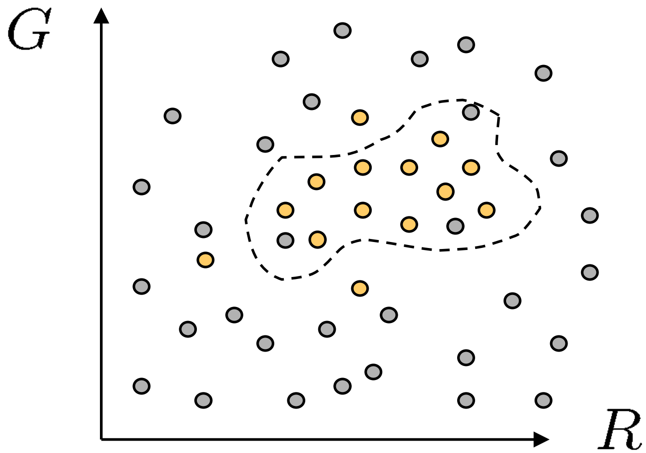
Skin pixels have a distinctive range of colors

- Corresponds to region(s) in RGB color space

Skin classifier

- A pixel X = (R,G,B) is skin if it is in the skin (color) region
- How to find this region?

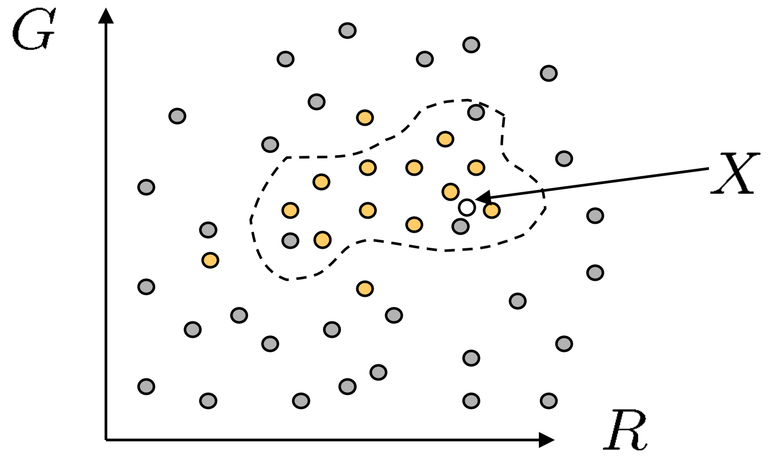
Skin detection



Learn the skin region from examples

- Manually label skin/non pixels in one or more “training images”
- Plot the training data in RGB space
  - skin pixels shown in orange, non-skin pixels shown in gray
  - some skin pixels may be outside the region, non-skin pixels inside.

Skin classifier

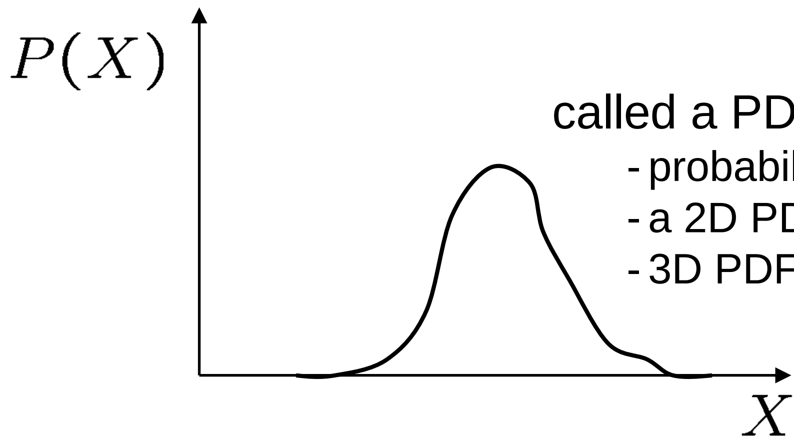


Given X = (R,G,B): how to determine if it is skin or not?

- Nearest neighbor
  - find labeled pixel closest to X
- Find plane/curve that separates the two classes
  - popular approach: Support Vector Machines (SVM)
- Data modeling
  - fit a probability density/distribution model to each class

Probability

- X is a random variable
- P(X) is the probability that X achieves a certain value



called a PDF
- probability distribution/density function
- a 2D PDF is a surface
- 3D PDF is a volume

$$0 \leq P(X) \leq 1$$

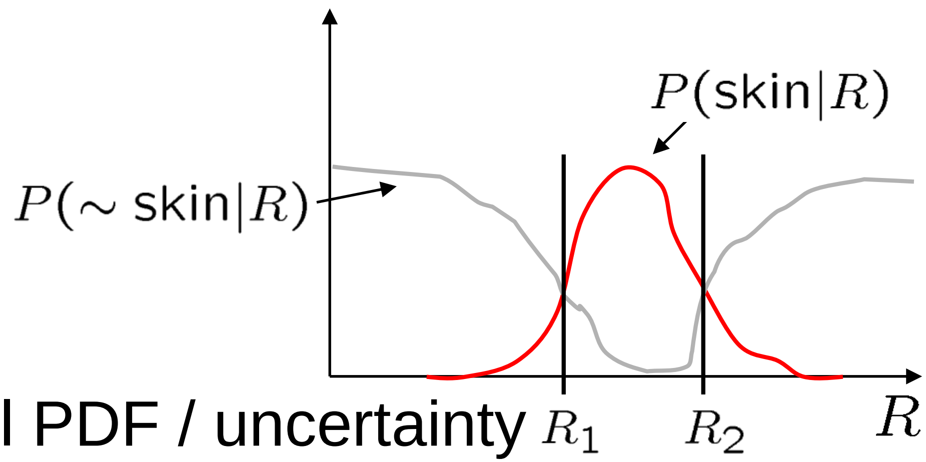
$$\int_{-\infty}^{\infty} P(X)dX = 1$$

continuous X

$$\sum P(X) = 1$$

discrete X

Probabilistic skin classification



Model PDF / uncertainty

- Each pixel has a probability of being skin or not skin

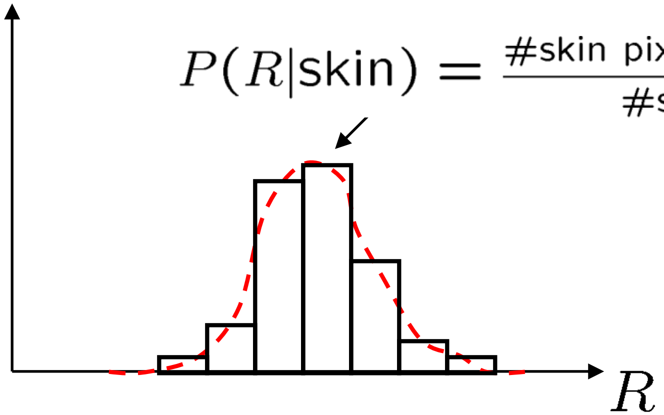
$$P(\sim \text{skin}|R) = 1 - P(\text{skin}|R)$$

Skin classifier

- Given X = (R,G,B): how to determine if it is skin or not?
- Choose interpretation of highest probability

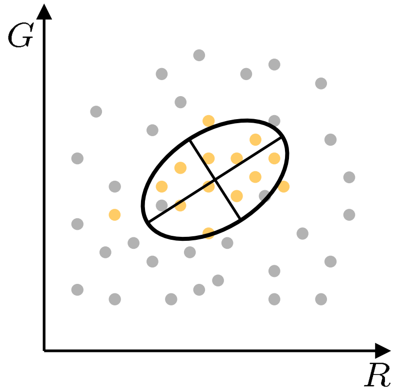
Where do we get $P(\text{skin}|R)$ and $P(\sim \text{skin}|R)$?

Learning conditional PDF's



We can calculate P(R | skin) from a set of training images

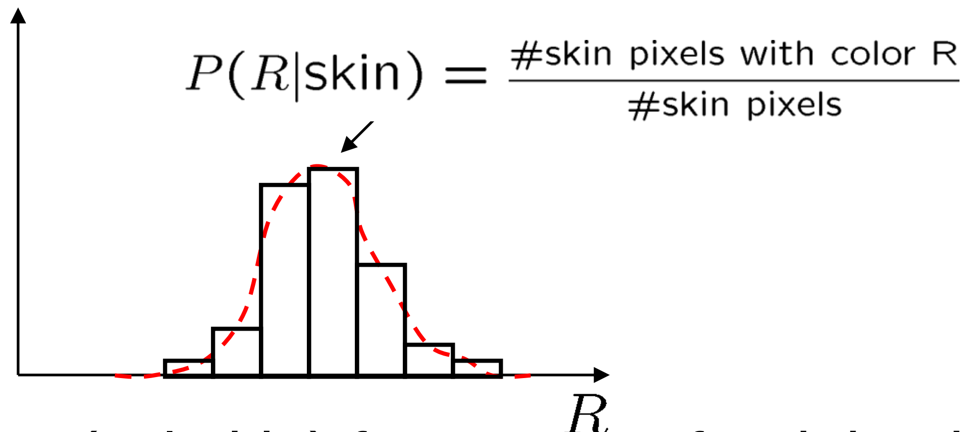
- It is simply a histogram over the pixels in the training images
  - each bin R_i contains the proportion of skin pixels with color R_i
- This doesn't work as well in higher-dimensional spaces. Why not?



Approach: fit parametric PDF functions

- common choice is rotated Gaussian
  - center $\mathbf{c} = \overline{X}$
  - covariance $\sum_X (X - \overline{X})(X - \overline{X})^T$

Learning conditional PDF’s



We can calculate P(R | skin) from a set of training images

But this isn’t quite what we want

- Why not? How to determine if a pixel is skin?
- We want P(skin | R) not P(R | skin)
- How can we get it?

Bayes rule

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$

In terms of our problem:

what we measure (**likelihood**)

domain knowledge (**prior**)

$$P(\text{skin}|R) = \frac{P(R|\text{skin})\ P(\text{skin})}{P(R)}$$

what we want (**posterior**)

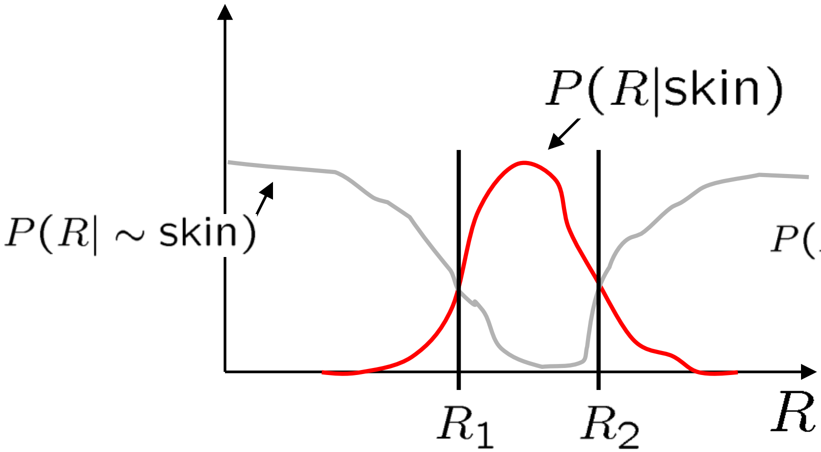
normalization term

$$P(R) = P(R|\text{skin})P(\text{skin}) + P(R|\sim\text{skin})P(\sim\text{skin})$$

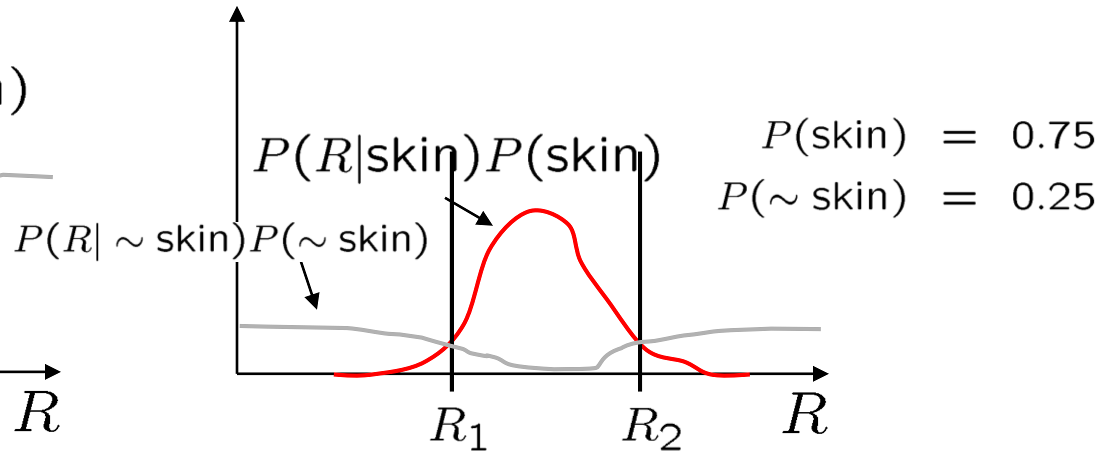
What can we use for the prior P(skin)?

- Domain knowledge:
  - P(skin) may be larger if we know the image contains a person
  - For a portrait, P(skin) may be higher for pixels in the center
- Learn the prior from the training set. How?
  - P(skin) is proportion of skin pixels in training set

Bayesian estimation



likelihood



posterior (unnormalized)

Bayesian estimation

- Goal is to choose the label (skin or ~skin) that maximizes the posterior ↔ minimizes probability of misclassification
  - this is called **Maximum A Posteriori (MAP) estimation**

Skin detection results

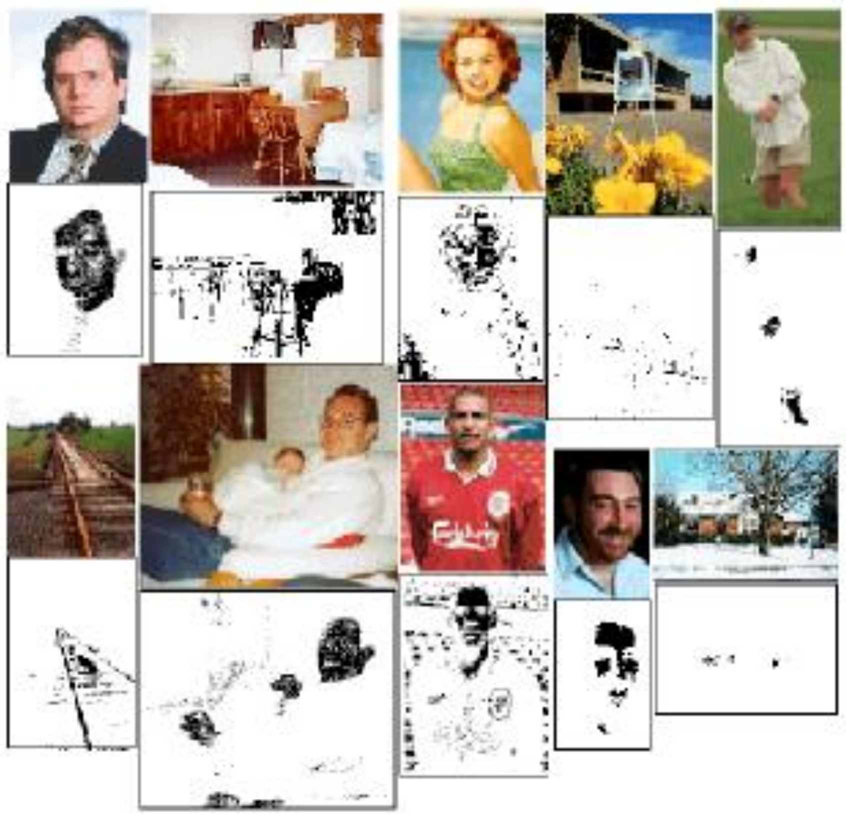
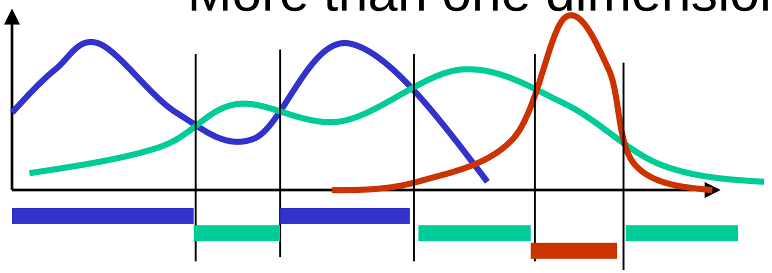


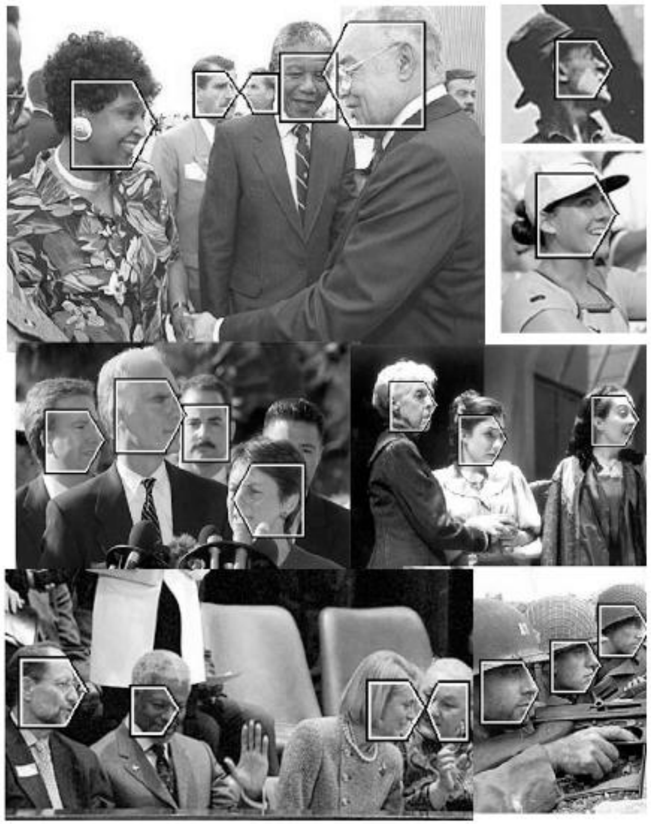
Figure 25.3. The figure shows a variety of images together with the output of the skin detector of Jones and Rehg applied to the image. Pixels marked black are skin pixels, and white are background. Notice that this process is relatively effective, and could certainly be used to focus attention on, say, faces and hands. *Figure from "Statistical color models with application to skin detection," M.J. Jones and J. Rehg, Proc. Computer Vision and Pattern Recognition, 1999 © 1999, IEEE*

General classification

This same procedure applies in more general circumstances

- More than two classes
- More than one dimension





Example: face detection

- Here, X is an image region
  - dimension = # pixels
  - each face can be thought of as a point in a high dimensional space

H. Schneiderman, T. Kanade. "A Statistical Method for 3D Object Detection Applied to Faces and Cars". CVPR 2000

Today's lecture

Face recognition and detection

- color-based skin detection
- recognition: eigenfaces [Turk & Pentland] and parts [Moghaddan & Pentland]
- detection: boosting [Viola & Jones]

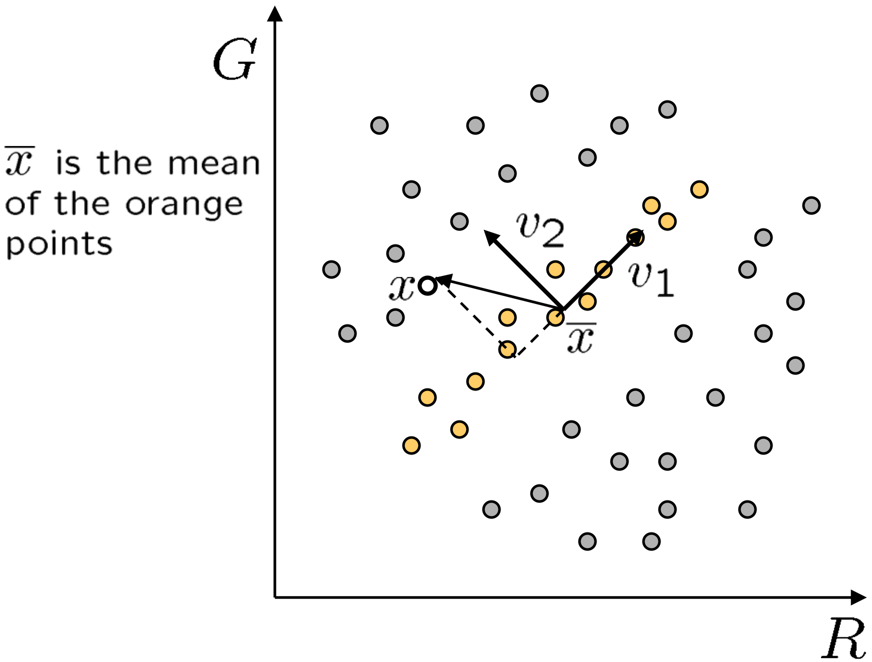
Eigenfaces for recognition

Matthew Turk and Alex Pentland

J. Cognitive Neuroscience

1991

Linear subspaces



convert **x** into **v$_1$**, **v$_2$** coordinates

$$\mathbf{x} \to ((\mathbf{x} - \overline{x}) \cdot \mathbf{v_1}, (\mathbf{x} - \overline{x}) \cdot \mathbf{v_2})$$

What does the **v$_2$** coordinate measure?

- distance to line
- use it for classification—near 0 for orange pts

What does the **v$_1$** coordinate measure?

- position along line
- use it to specify which orange point it is

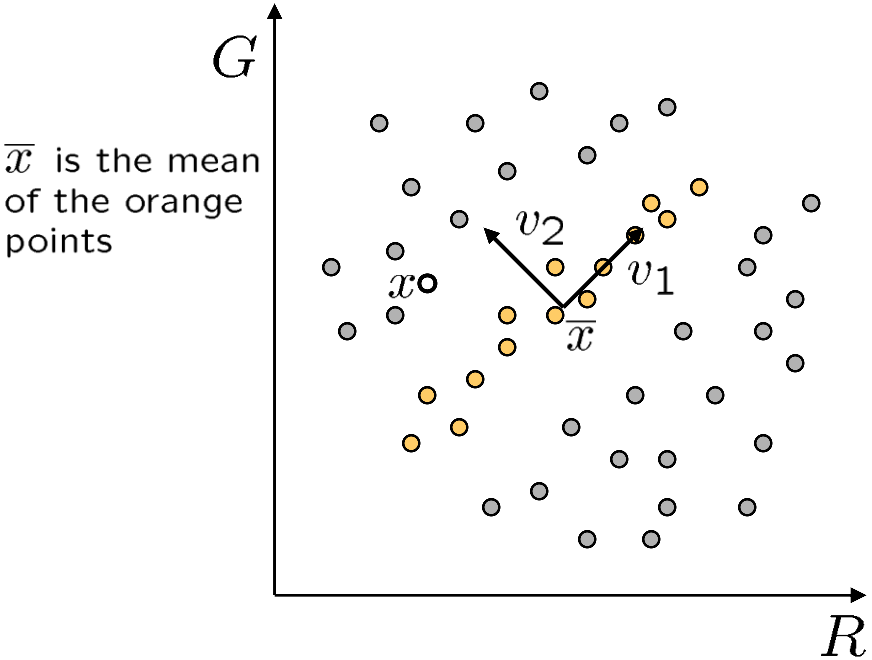
Classification can be expensive:

- Big search prob (e.g., nearest neighbors) or store large PDF's

Suppose the data points are arranged as above

- Idea—fit a line, classifier measures distance to line

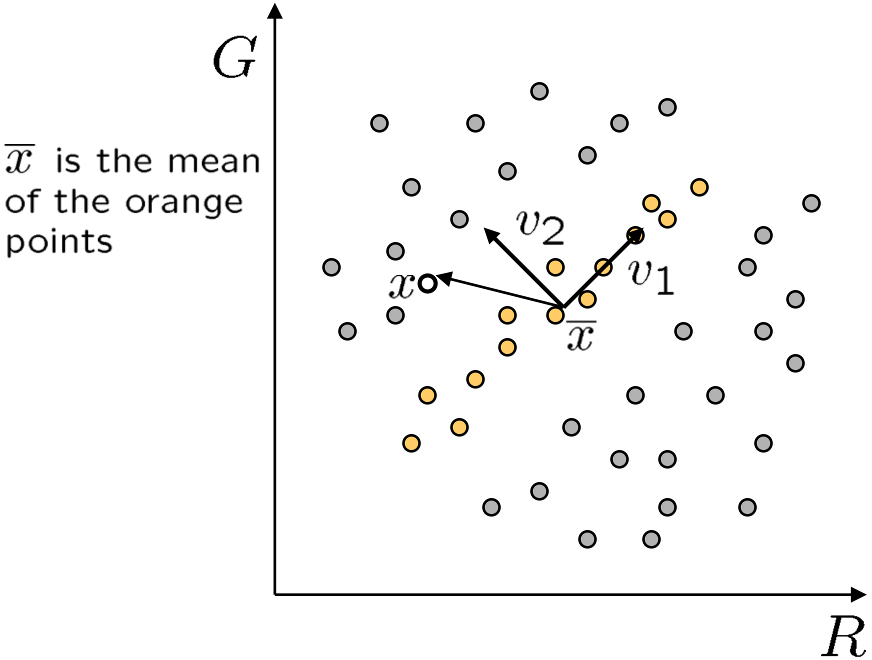
Dimensionality reduction



Dimensionality reduction

- We can represent the orange points with *only* their $\mathbf{v_1}$ coordinates (since $\mathbf{v_2}$ coordinates are all essentially 0)
- This makes it much cheaper to store and compare points
- A bigger deal for higher dimensional problems

Linear subspaces



Consider the variation along direction **v** among all of the orange points:

$$var(\mathbf{v}) = \sum_{\text{orange point } \mathbf{x}} \|(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}} \cdot \mathbf{v}\|^{2}$$

What unit vector **v** minimizes *var*?

$$\mathbf{v}_2 = min_{\mathbf{v}} \; \{var(\mathbf{v})\}$$

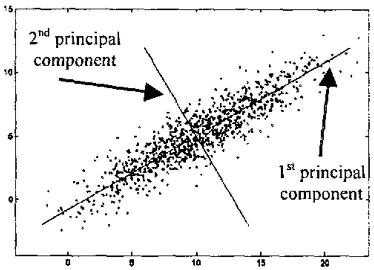
What unit vector **v** maximizes *var*?

$$\mathbf{v}_1 = max_{\mathbf{v}} \; \{var(\mathbf{v})\}$$

$$\begin{aligned} var(\mathbf{v}) &= \sum_{\mathbf{x}} \|(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}} \cdot \mathbf{v}\| \\ &= \sum_{\mathbf{x}} \mathbf{v}^{\mathbf{T}}(\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}}\mathbf{v} \\ &= \mathbf{v}^{\mathbf{T}}\left[\sum_{\mathbf{x}}(\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}}\right]\mathbf{v} \\ &= \mathbf{v}^{\mathbf{T}}\mathbf{A}\mathbf{v} \quad \text{where } \mathbf{A} = \sum_{\mathbf{x}}(\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}} \end{aligned}$$

Solution: $\mathbf{v}_1$ is eigenvector of **A** with *largest* eigenvalue
$\mathbf{v}_2$ is eigenvector of **A** with *smallest* eigenvalue

Principal component analysis



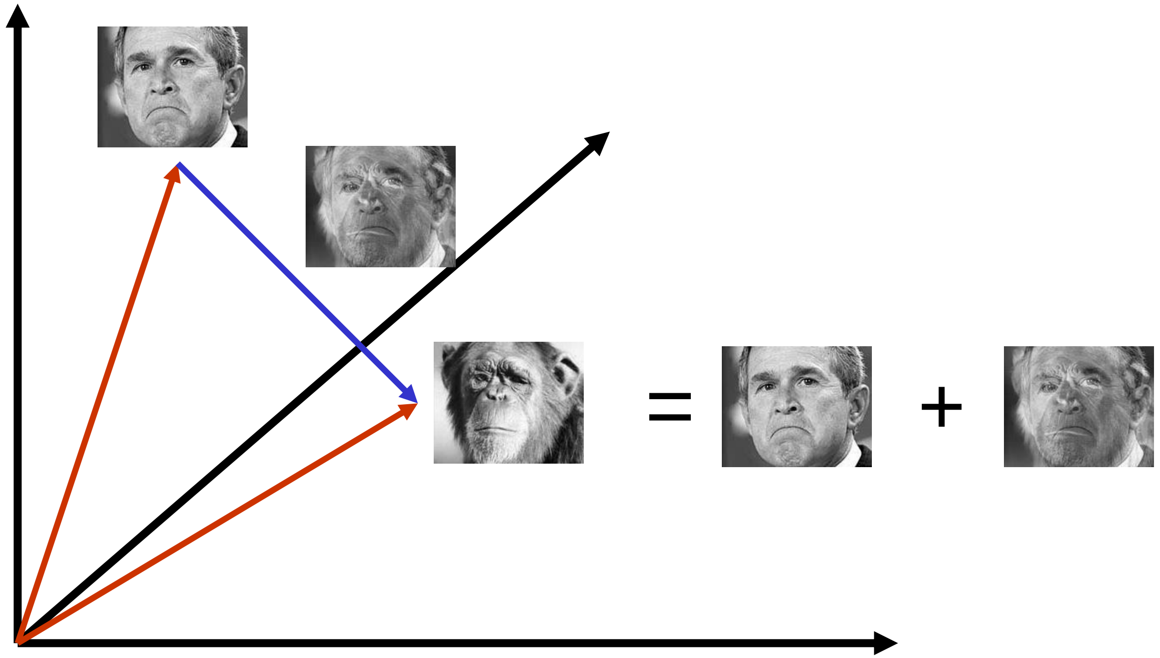
Suppose each data point is N-dimensional

- Same procedure applies:

$$\begin{aligned} var(\mathbf{v}) &= \sum_{\mathbf{x}} \|(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}} \cdot \mathbf{v}\| \\ &= \mathbf{v}^{\mathbf{T}} \mathbf{A} \mathbf{v} \quad \text{where } \mathbf{A} = \sum_{\mathbf{x}} (\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}} \end{aligned}$$

- The eigenvectors of **A** define a new coordinate system
  – eigenvector with largest eigenvalue captures the most variation among training vectors **x**
  – eigenvector with smallest eigenvalue has least variation
- We can compress the data using the top few eigenvectors
  – corresponds to choosing a “linear subspace”
    » represent points on a line, plane, or “hyper-plane”
  – these eigenvectors are known as the ***principal components***

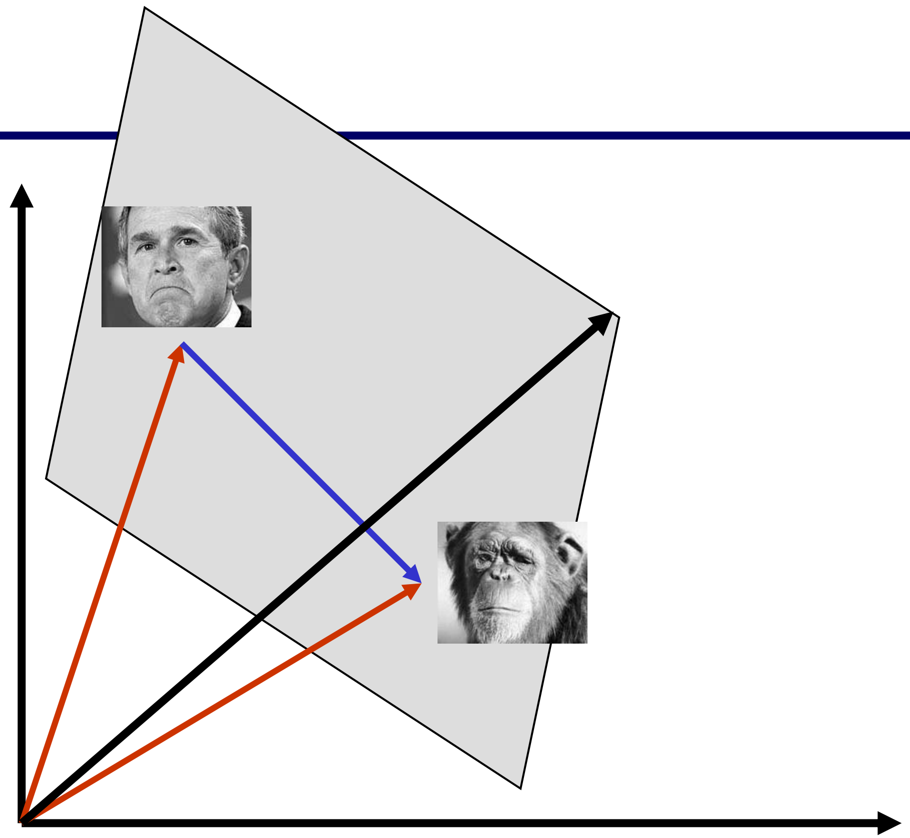
The space of faces



An image is a point in a high dimensional space

- An N x M image is a point in R^{NM}
- We can define vectors in this space as we did in the 2D case

Dimensionality reduction



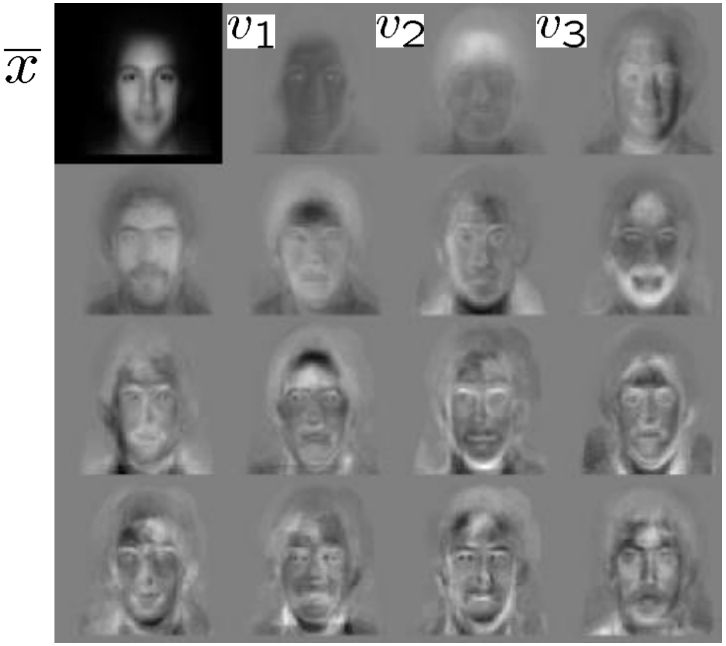
The set of faces is a "subspace" of the set of images

- We can find the best subspace using PCA
- This is like fitting a "hyper-plane" to the set of faces
  - spanned by vectors $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_K}$
  - any face $\mathbf{x} \approx \overline{\mathbf{x}} + a_1\mathbf{v_1} + a_2\mathbf{v_2} + \ldots + a_k\mathbf{v_k}$

Eigenfaces

PCA extracts the eigenvectors of **A**

- Gives a set of vectors $\mathbf{v_1}$, $\mathbf{v_2}$, $\mathbf{v_3}$, ...
- Each vector is a direction in face space
  - what do these look like?



Projecting onto the eigenfaces

The eigenfaces $\mathbf{v_1}, \ldots, \mathbf{v_K}$ span the space of faces

- A face is converted to eigenface coordinates by

$$\mathbf{x} \rightarrow (\underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v_1}}_{a_1}, \underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v_2}}_{a_2}, \ldots, \underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v_K}}_{a_K})$$

$$\mathbf{x} \approx \overline{\mathbf{x}} + a_1\mathbf{v_1} + a_2\mathbf{v_2} + \ldots + a_K\mathbf{v_K}$$

$\mathbf{x}$

$a_1\mathbf{v_1}$ $a_2\mathbf{v_2}$ $a_3\mathbf{v_3}$ $a_4\mathbf{v_4}$ $a_5\mathbf{v_5}$ $a_6\mathbf{v_6}$ $a_7\mathbf{v_7}$ $a_8\mathbf{v_8}$

Recognition with eigenfaces

Algorithm

1. Process the image database (set of images with labels)
   - Run PCA—compute eigenfaces
   - Calculate the K coefficients for each image
2. Given a new image (to be recognized) **x**, calculate K coefficients

$$\mathbf{x} \rightarrow (a_1, a_2, \ldots, a_K)$$

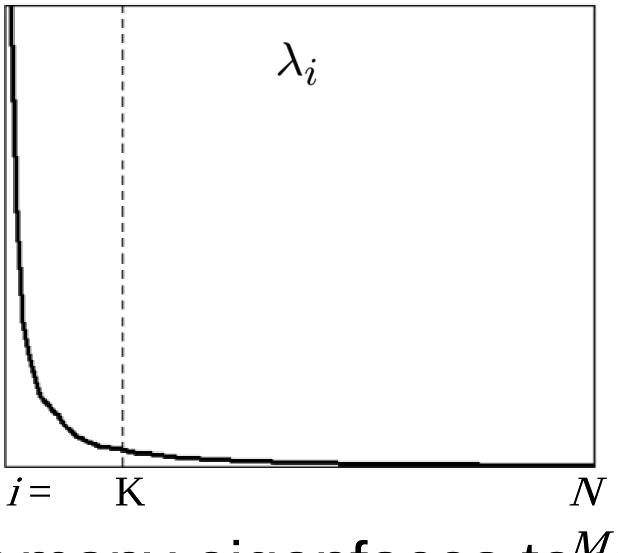
3. Detect if x is a face

$$\|\mathbf{x} - (\overline{\mathbf{x}} + a_1\mathbf{v_1} + a_2\mathbf{v_2} + \ldots + a_K\mathbf{v_K})\| < \text{threshold}$$

1. If it is a face, who is it?
   – Find closest labeled face in database
      » nearest-neighbor in K-dimensional space

Choosing the dimension K

eigenvalues



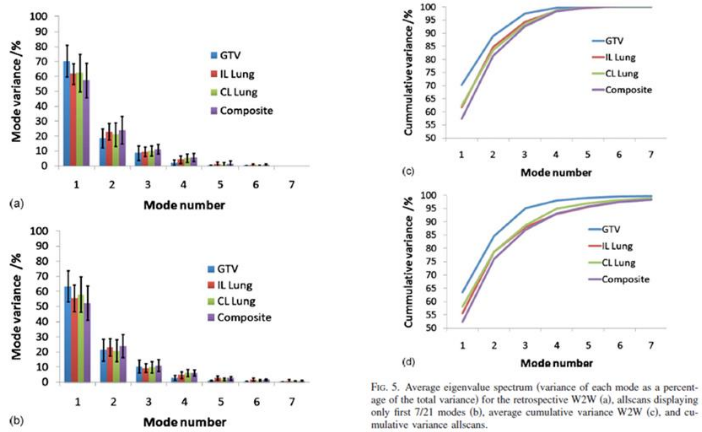


Fig. 5. Average eigenvalue spectrum (variance of each mode as a percentage of the total variance) for the retrospective W2W (a), allscans displaying only first 7/21 modes (b), average cumulative variance W2W (c), and cumulative variance allscans.

How many eigenfaces to use?

Look at the decay of the eigenvalues

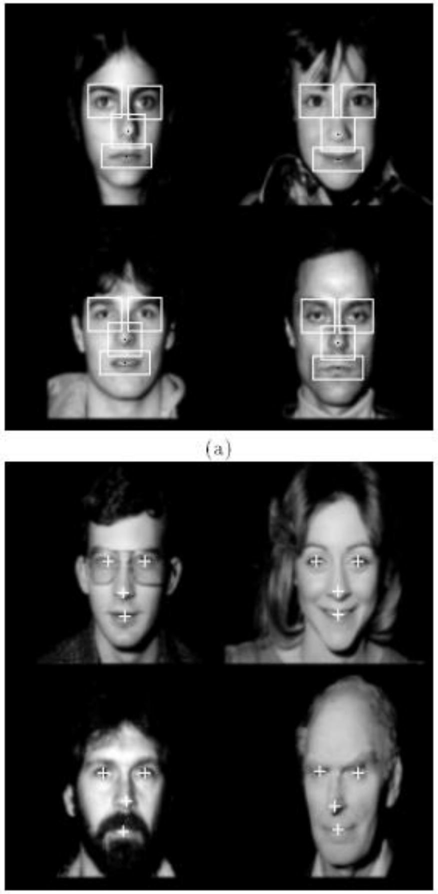
- the eigenvalue tells you the amount of variance “in the direction” of that eigenface
- ignore eigenfaces with low variance

View-Based and Modular Eigenspaces for Face Recognition

Alex Pentland, Baback Moghaddam and
Thad Starner
CVPR'94

Part-based eigenfeatures

- Learn a separate eigenspace for each face feature
- Boosts performance of regular eigenfaces



PCA for data modeling and representation

- Intensities/colors I(x,y,z) ---------□ Appearance
- Geometries or vertex positions (X,Y,Z) ------□ Shapes
- Shapes
- General data D(X,Y,Z,…..)
- Inter or intra modeling (eg modeling facial aging, geometries vs time or tumors vs time etc.)

Idea is each sample represents a point in Eigen space (Parameter space)

PCA for 3D Geometry Modeling example

3D ROI models have *homologous* point set correspondence as we warp each point and this point has the same order in all feature vectors (points are corresponding to each other)

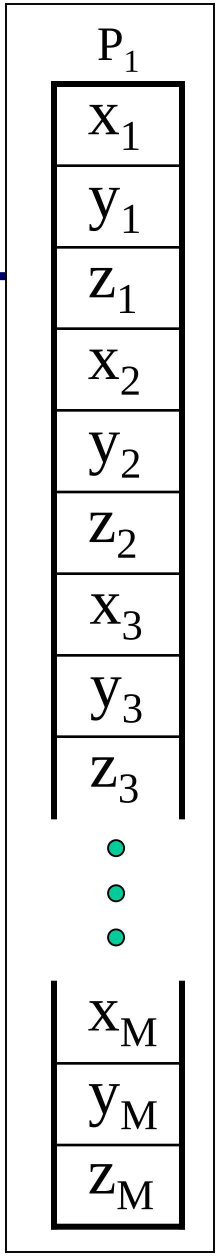
[1] Ahmed M Badawi, Elisabeth Weiss, William C Sleeman IV, Chenyu Yan, Geoffrey D Hugo, **“Optimizing principal component models for representing interfraction variation in lung cancer radiotherapy”** Med Phys. 2010 September; 37(9): 5080–5091.

[2] Ahmed M Badawi, Elisabeth Weiss, William C Sleeman IV and Geoffrey D Hugo, “**Classifying geometric variability by dominant eigenmodes of deformation in regressing tumours during active breath-hold lung cancer radiotherapy**”, 2012 *Phys. Med. Biol.* **57** 395

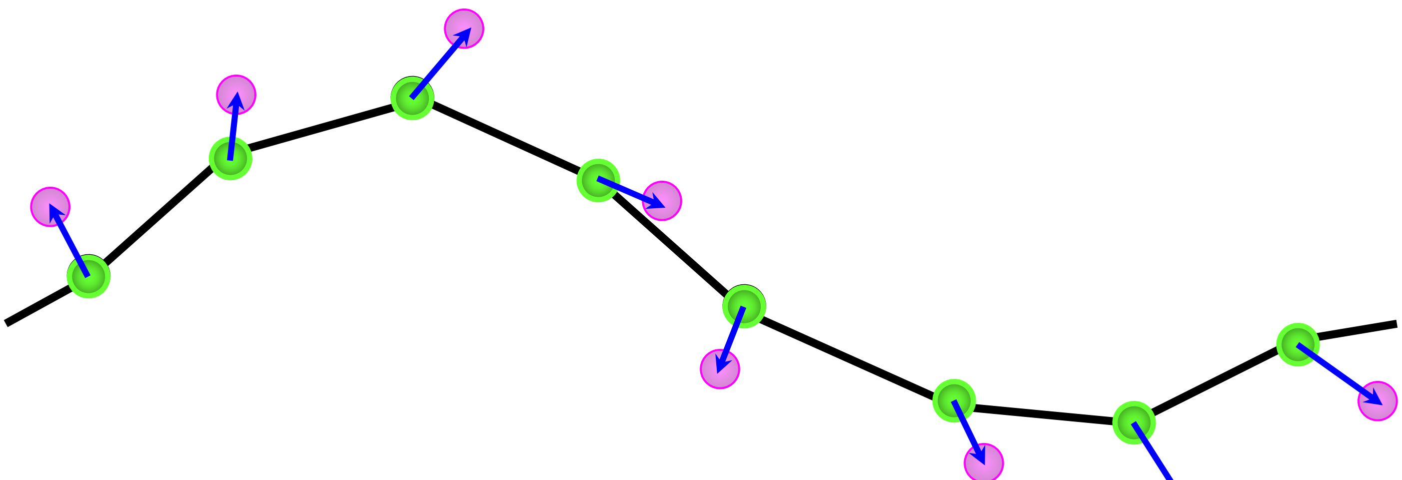
[3] M. Sohn, M. Birkner, D. Yan and M. Alber, "**Modelling individual geometric variation based on dominant eigenmodes of organ deformation: implementation and evaluation**," Phys Med Biol **50**, 5893-5908 (2005).

[4] Mohamed Mahfouz, Ahmed Badawi, Brandon Merkl, Emam E. Abdel Fatah, Emily Pritchard, Katherine Kesler, Megan Moore, Richard Jantz, Lee Jantz, **"Patella sex determination by 3D statistical shape models and nonlinear classifiers,"** Forensic Science International vol. 173,2007 , pages,161-170.

3D feature vector



- **Feature vector formation (3M size)**
- **Positions (X,Y,Z) or dvf (ΔX,ΔY,ΔZ)**



Green: (P_1) 3D Point position at week time, W_t

Purple: (P_2) Deformed (warped) Point position at W_{t+1}

P_i: I = 1 → 7 for W2W samples and I = 1 → 21 for all scans

Principal Components Analysis

Mean Shape subtraction

Samples (Observations) Vectors

Mean

Variations from mean matrix

$$\begin{bmatrix} p_1 & \vdots & p_2 & \vdots & \ldots & \vdots & p_N \end{bmatrix} - \begin{bmatrix} \bar{p} \end{bmatrix} = \underbrace{\begin{bmatrix} dp_1 & \vdots & dp_2 & \vdots & \ldots & \vdots & dp_N \end{bmatrix}}_{dp}$$

3Mx N

Principal Components Analysis

In systems Eng. characteristic matrix, we solve for Ax=λx where A is the system characteristic matrix that have the control parameters (poles in system transfer function)

Eigen calculation

$$\underbrace{\frac{1}{N-1} dp\, dp^{T}}_{COV} q_l = \lambda_l q_l$$

3Mx3M

3Mx1

Diagonalization of COV matrix (3Mx3M) results in eigenvectors q_l

q_l (3Mx1) is the l^{th} eigenvector of $\frac{1}{N-1} dp\, dp^{T}$

λ_l is the eigenvalue associated with q_l

Eigenvalue =Statistical Variance $\sigma_l^2 = \lambda_l$

Singular Value Decomposition (SVD)

$$\mathrm{p}'(t) = \bar{\mathrm{p}} + \sum_{l=1}^{L} c_l(t)\mathrm{q}_l,$$

$$\mathrm{dp}(t) = \frac{1}{\sqrt{N-1}}(\mathrm{p}(t) - \bar{\mathrm{p}})^T,$$

Faster for large data

$$\mathrm{DP} = \{\mathrm{dp}(t_1)|\ldots\ldots|\mathrm{dp}(t_N)\}.$$

$$\mathrm{DP} = USV^T,$$

DP: 3MxN, S: NxN singular values, U: 3MxN containing eigenvectors, V:NxN

Construction of Organ Geometries Using Eigen Modes

Ranking λ_l (geometric variability)

Principal deformation modes (L) are the dominant eigen modes with largest eigen values (modes which span the space in which the majority of deformation occurs)

$$p = \overline{p} + \sum_{l=1}^{N-1} c_l q_l \qquad \|q_l\| = 1$$

$$p = \overline{p} + \sum_{l=1}^{L} c_l q_l + \varepsilon \qquad \|q_l\| = 1$$

- Deforming the mean shape by a weighted sum of L dominating eigenmodes
- c_l obey Gaussian distribution with corresponding λ_l as variances. Thus the dominating eigenmodes serve as statistical model of individual organ/deformation with only a small number of parameters.

Calculating Optimal Reconstruction Coefficients

PCA model representation:

Weighted sum of L dominant modes

$$p_{i,opt}^{[L]} = \overline{p} + \sum_{l=1}^{L} c_{l,opt}(i) q_l$$

- Optimal Coefficients Calculation:

$$c_{l,opt}(i) = (p_i - \overline{p}) \cdot q_l \qquad with \qquad \|q_l\| = 1; \;\; l = 1, \ldots, L.$$

L: is selected for sum of variance > 90% or 95% of total sum (in some better representations, it can be taken as 98% to minimize errors), selection of L depends on the resolution error in application as trade of with amount of reduction in model (speed of reconstruction)

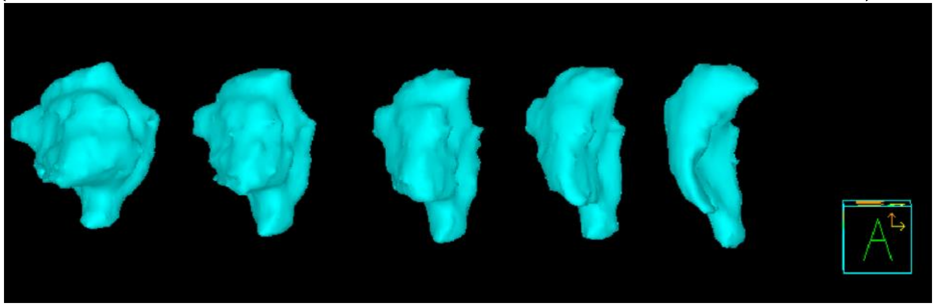
3D Reconstruction of Dominating Eigenmodes

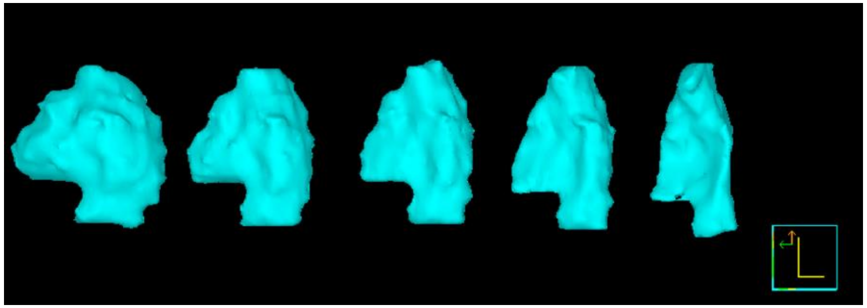
Deformation of the mean geometry ($\bar{p}$) by the respective normed eignevector (q_l) is given by:

$$p_l = \bar{p} \pm \sigma_l . q_l$$

Eigenvalue =Statistical Variance $\sigma_l^2 = \lambda_l$

One can then construct an animation from –σ to +σ for small increments (3D movie that shows mode deformation)





2 Views of dominant mode 1 for subject 1 GTV
(Mode1 - 3σ, Mode1 – σ, Mean Mode, Mode1 + σ, Mode1 + 3σ)

Shape Similarity Quantization

Local representation error or local residual:

$$d_{i,j}^{[L]} \qquad i = 1,\ldots,N \quad , j = 1,\ldots,M$$

- Average local residual and stdev (Δd):

$$\bar{d}_j^{[L]} = \frac{1}{N}\sum_{i=1}^{N} d_{i,j}^{[L]} \qquad j = 1,\ldots,M$$

Histogram of these values gives an overview of the overall quality of representation.

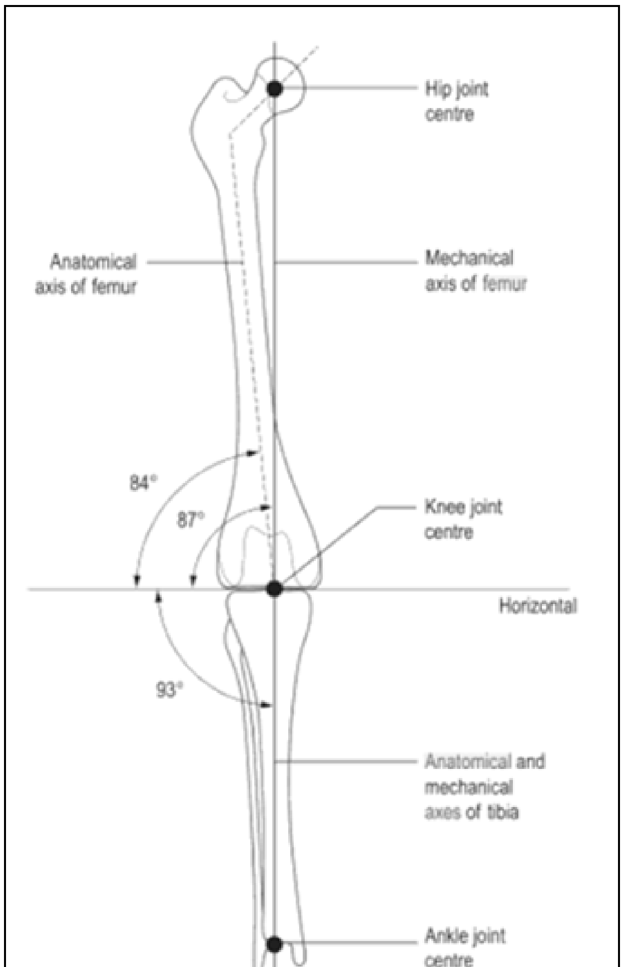
- Overall residual and overall stdev (ΔR):

$$R^{[L]} = \frac{1}{M}\sum_{j=1}^{M} \bar{d}_j^{[L]}$$

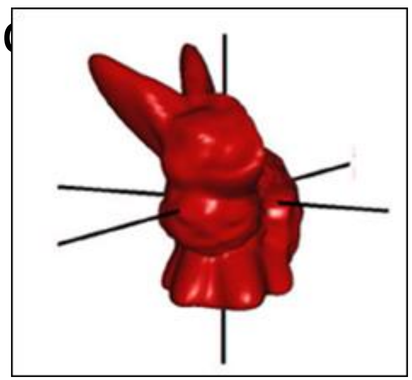
Overall error and stdev gives comprehensive measure for the quality of the PCA model with L eigenmodes

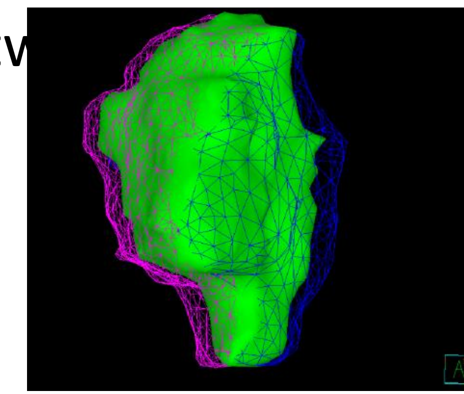
Rotation Calculations for 3D shapes (Rigid alignment)

Anatomical and Mechanical axes are defined



- For undefined axes shapes:
- Axes of maximum variance (Principal axes) are used to calculate rotations between 2 ROI's
- Center of geometry is used to c etw





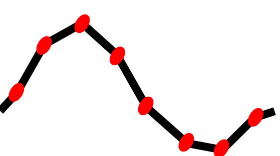
Principal Eigen Rotation (Principal Axes)

Due to lack of anatomical axes definition to GTV anatomy wrt to its geometry (unlike bones), we calculate rotations between 2 ROI's principal axes.

Principal axes are calculated by PCA using XYZ points as samples (Eigen Vectors are the principal axes)

$$P_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \quad P_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \quad P_3 = \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} \quad P_o = \begin{bmatrix} \bullet \\ \bullet \\ \bullet \end{bmatrix} \quad P_M = \begin{bmatrix} x_M \\ y_M \\ z_M \end{bmatrix}$$

$$\underset{3\mathrm{x}\,M}{\left[\begin{array}{c:c:c:c} p_1 & p_2 & \ldots & p_M \end{array}\right]} - \left[\, \bar{p} \,\right] = \underbrace{\left[\begin{array}{c:c:c:c} dp_1 & dp_2 & \ldots & dp_M \end{array}\right]}_{dp}$$



Morphable Face Models

Rowland and Perrett ’95

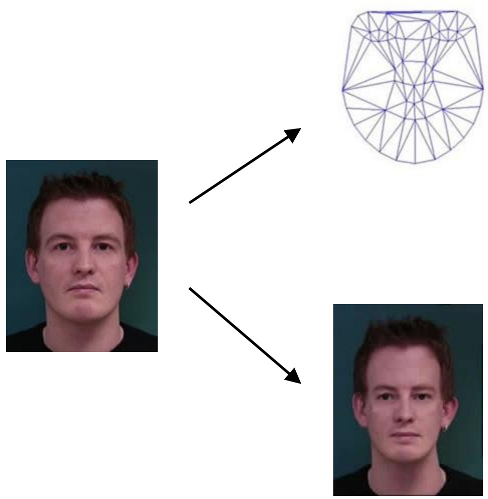
Lanitis, Cootes, and Taylor ’95, ’97

Blanz and Vetter ’99

Matthews and Baker ’04, ‘07

Morphable Face Model

Use subspace to model elastic 2D or 3D ***shape*** variation (vertex positions), in addition to *appearance* variation

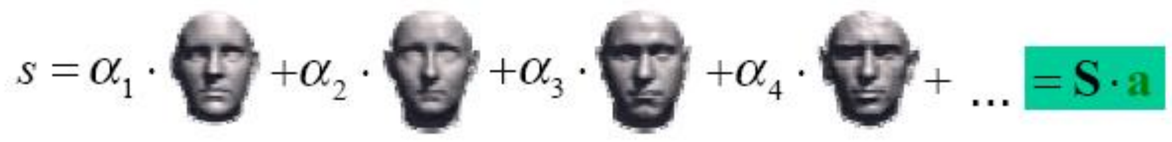


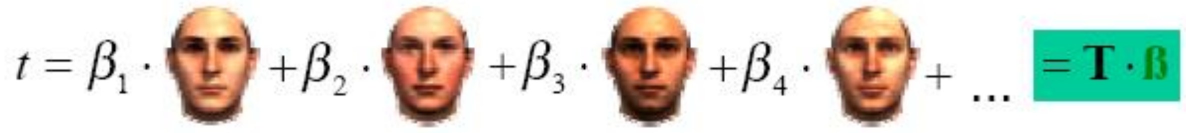
Shape S

Appearance T

Morphable Face Model

$$\mathbf{S}_{model} = \sum_{i=1}^{m} a_i \mathbf{S}_i \qquad \mathbf{T}_{model} = \sum_{i=1}^{m} b_i \mathbf{T}_i$$





3D models from Blanz and Vetter ‘99

Face Recognition Resources

Face Recognition Home Page:

- http://www.cs.rug.nl/~peterkr/FACE/face.html

PAMI Special Issue on Face & Gesture (July '97)

FERET

- http://www.dodcounterdrug.com/facialrecognition/Feret/feret.htm

Face-Recognition Vendor Test (FRVT 2000)

- http://www.dodcounterdrug.com/facialrecognition/FRVT2000/frvt2000.htm

Biometrics Consortium

- http://www.biometrics.org

Today's lecture

Face recognition and detection

- color-based skin detection
- recognition: eigenfaces [Turk & Pentland] and parts [Moghaddan & Pentland]
- detection: boosting [Viola & Jones]

Example: Face Detection

Scan window over image

Classify window as either:

- Face
- Non-face

Then recognize face by Eigen analysis

Robust real-time face detection

Paul A. Viola and Michael J. Jones

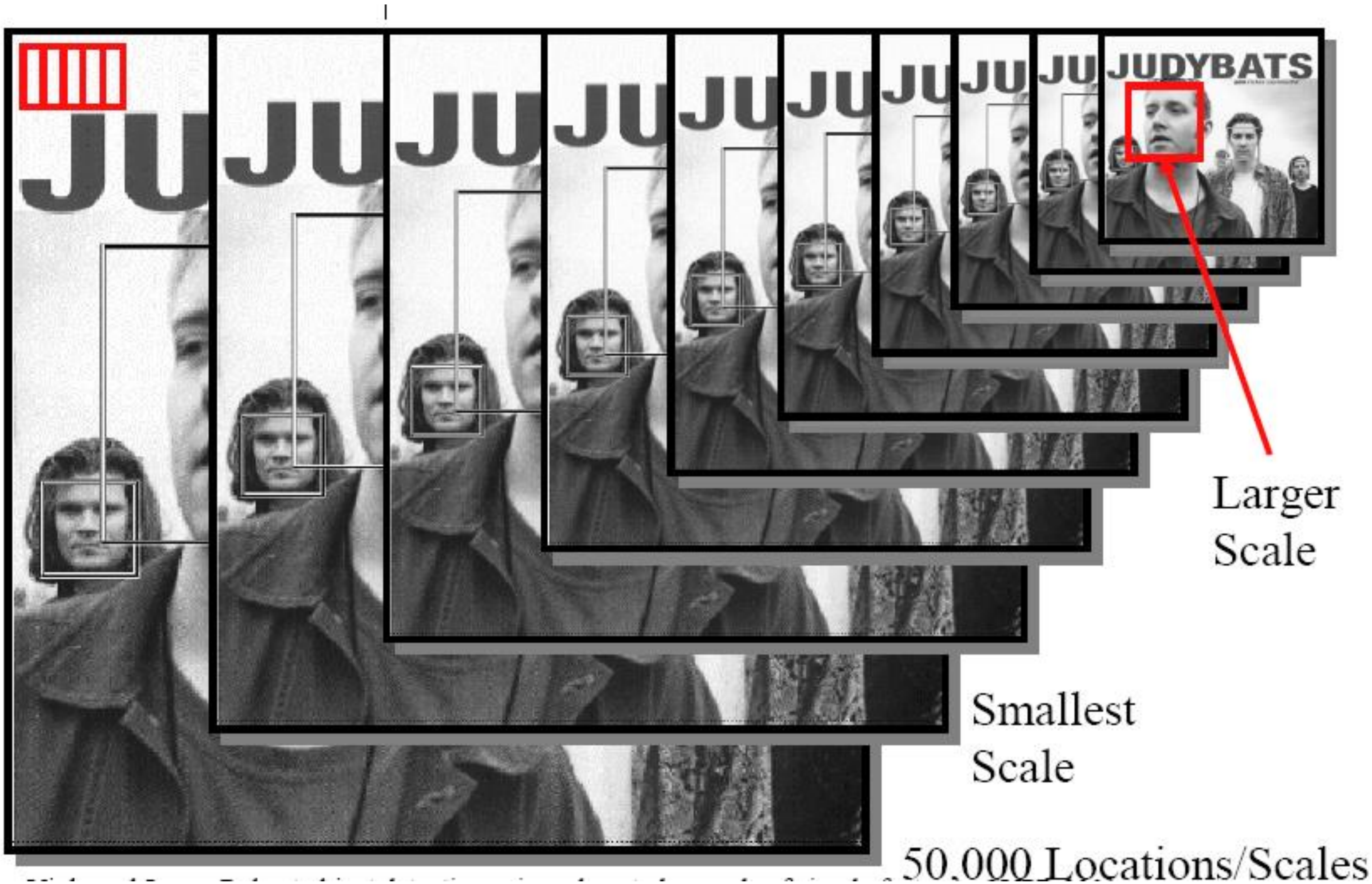
Intl. J. Computer Vision

57(2), 137–154, 2004

(originally in CVPR'2001)

(slides adapted from Bill Freeman, MIT 6.869, April 2005)

Scan classifier over locs. & scales



“Learn” classifier from data

Training Data

- 5000 faces (frontal)
- 10^8 non faces
- Faces are normalized
    - Scale, translation

Many variations

- Across individuals
- Illumination
- Pose (rotation both in plane and out)

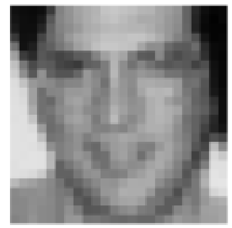
Characteristics of algorithm

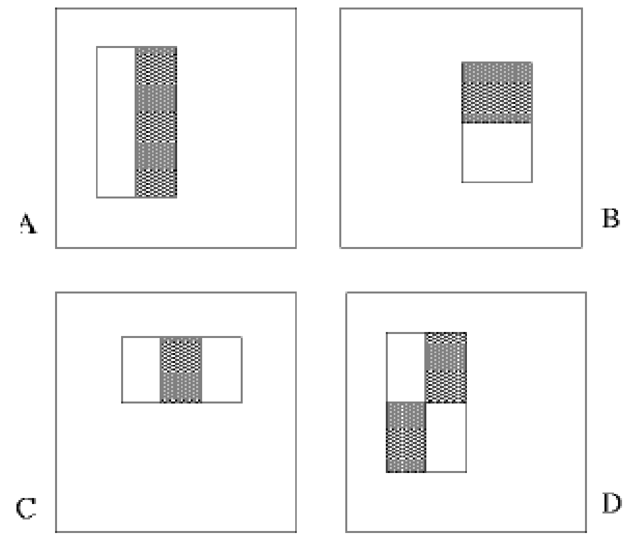
- Feature set (…is huge about 16M features)
- Efficient feature selection using AdaBoost
- New image representation: Integral Image
- Cascaded Classifier for rapid detection

➢ Fastest known face detector for gray scale images

Image features

- "Rectangle filters"
  - Similar to Haar wavelet
- Differences between sums of pixels in adjacent rectangles

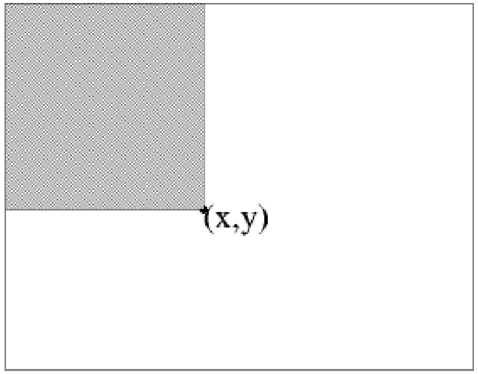




$$h_t(x) = \begin{cases} +1 & \text{if } f_t(x) > \theta_t \\ -1 & \text{otherwise} \end{cases}$$

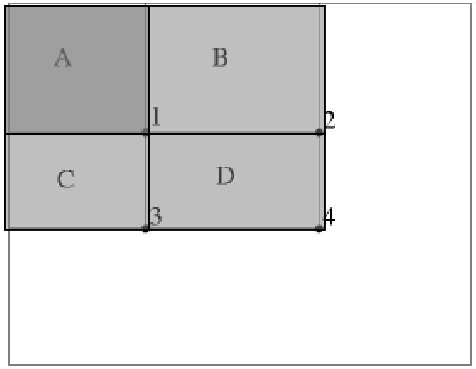
Integral Image

Partial sum $I'(x,y) = \sum_{\substack{x' \le x \\ y' \le y}} I(x',y')$



Any rectangle is

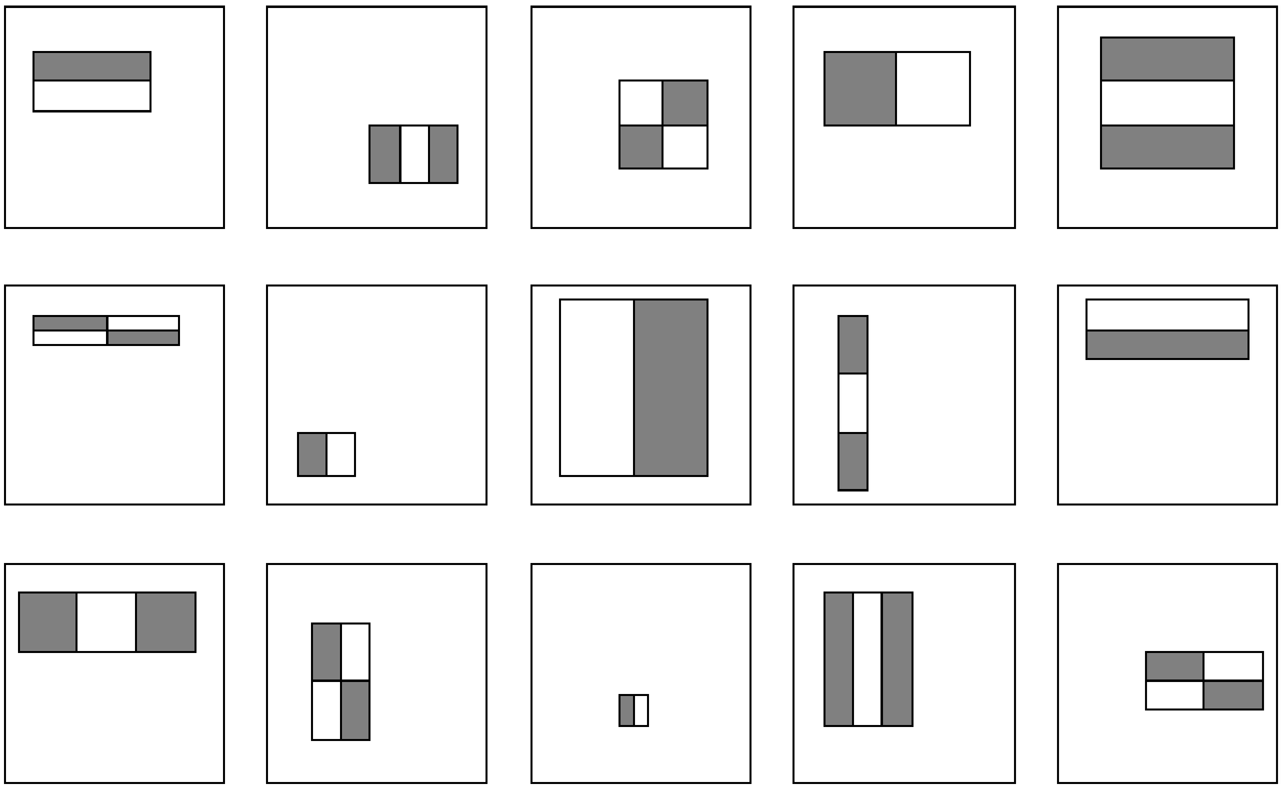
D = 1+4-(2+3)



Also known as:

- *summed area tables* [Crow84]
- *boxlets* [Simard98]

Huge library of filters



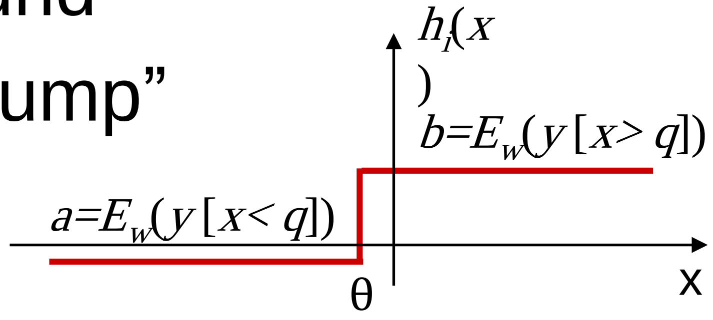
Constructing the classifier

Perceptron yields a sufficiently powerful classifier

$$C(x) = \theta\left(\sum_i \alpha_i h_i(x) + b\right)$$

Use AdaBoost to efficiently choose best features

- add a new $h_i(x)$ at each round
- each $h_i(x_k)$ is a "decision stump"



Constructing the classifier

For each round of boosting:

- Evaluate each rectangle filter on each example
- Sort examples by filter values
- Select best threshold for each filter (min error)
  Use sorting to quickly scan for optimal threshold
- Select best filter/threshold combination
- Weight is a simple function of error rate
- Reweight examples
  (There are many tricks to make this more efficient.)

Good reference on boosting

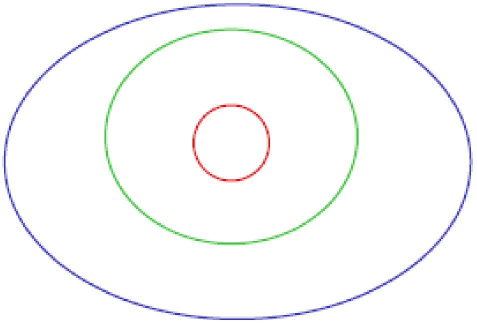
Friedman, J., Hastie, T. and Tibshirani, R.
Additive Logistic Regression: a Statistical View of Boosting

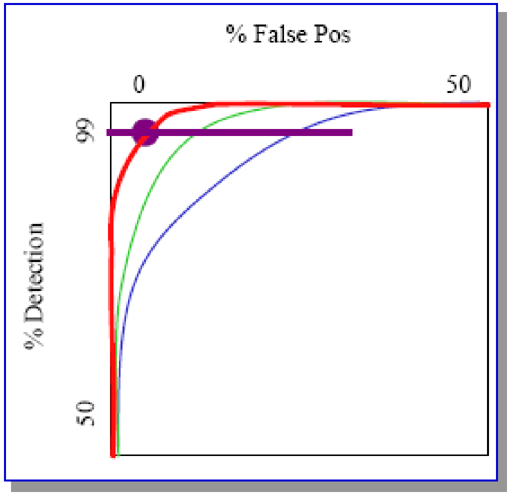
http://www-stat.stanford.edu/~hastie/Papers/boost.ps

"We show that boosting fits an additive logistic regression model by stagewise optimization of a criterion very similar to the log-likelihood, and present likelihood based alternatives. We also propose a multi-logit boosting procedure which appears to have advantages over other methods proposed so far."

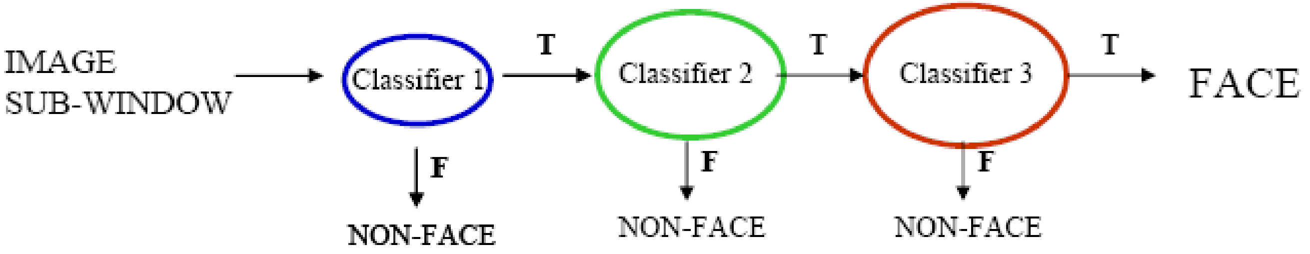
Trading speed for accuracy

Given a nested set of classifier hypothesis classes





Computational Risk Minimization



Speed of face detector (2001)

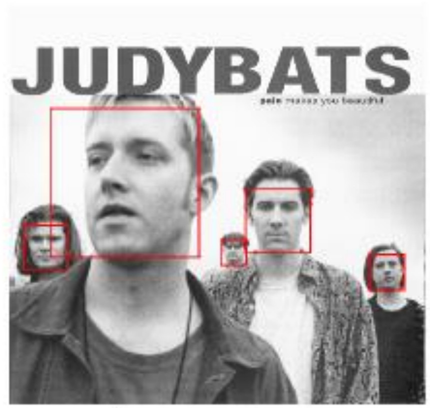
Speed is proportional to the average number of features computed per sub-window.

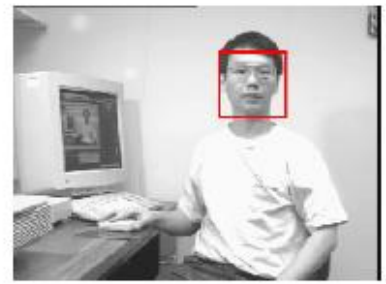
On the MIT+CMU test set, an average of 9 features (/ 6061) are computed per sub-window.

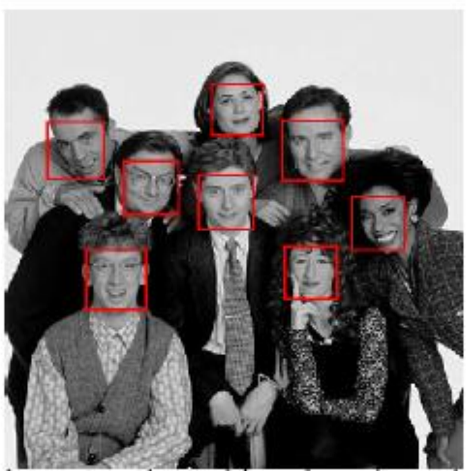
On a 700 Mhz Pentium III, a 384x288 pixel image takes about 0.067 seconds to process (15 fps).

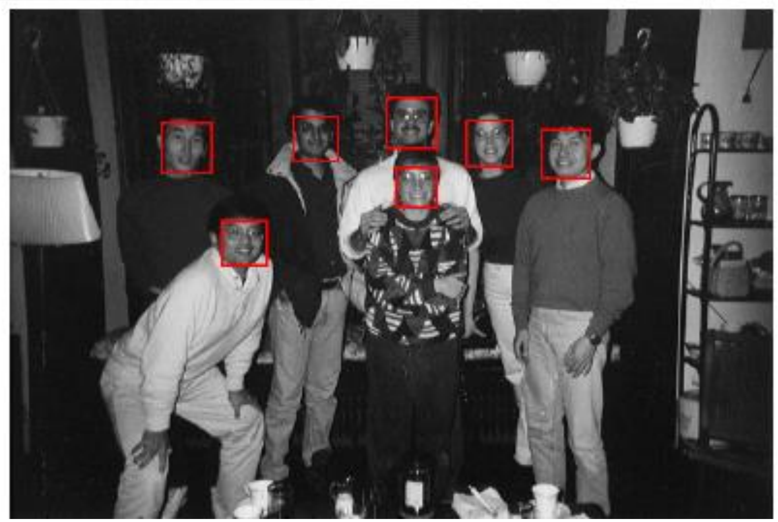
Roughly 15 times faster than Rowley-Baluja-Kanade and 600 times faster than Schneiderman-Kanade.

Sample results









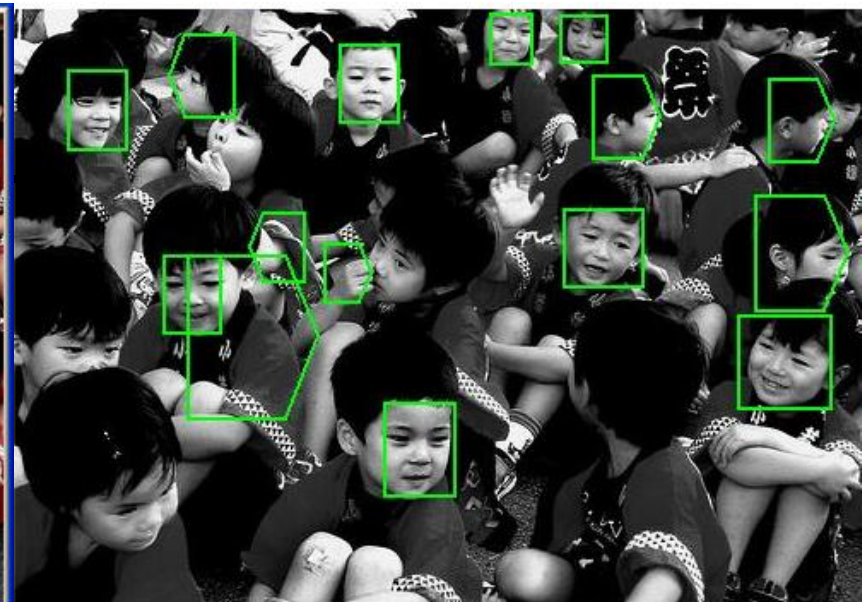
Summary (Viola-Jones)

- Fastest known face detector for gray images
- Three contributions with broad applicability:
  - ❖Cascaded classifier yields rapid classification
  - ❖AdaBoost as an extremely efficient feature selector
  - ❖Rectangle Features + Integral Image can be used for rapid image analysis

Face detector comparison

Informal study by Andrew Gallagher, CMU, for CMU 16-721 Learning-Based Methods in Vision, Spring 2007

- The Viola Jones algorithm OpenCV implementation was used. (<2 sec per image).
- For Schneiderman and Kanade, Object Detection Using the Statistics of Parts [IJCV'04], the www.pittpatt.com demo was used. (~10-15 seconds per image, including web transmission).

Viola
Jones


Schneiderman
Kanade

Today's lecture

Face recognition and detection

- color-based skin detection
- recognition: eigenfaces [Turk & Pentland] and parts [Moghaddan & Pentland]
- detection: boosting [Viola & Jones]

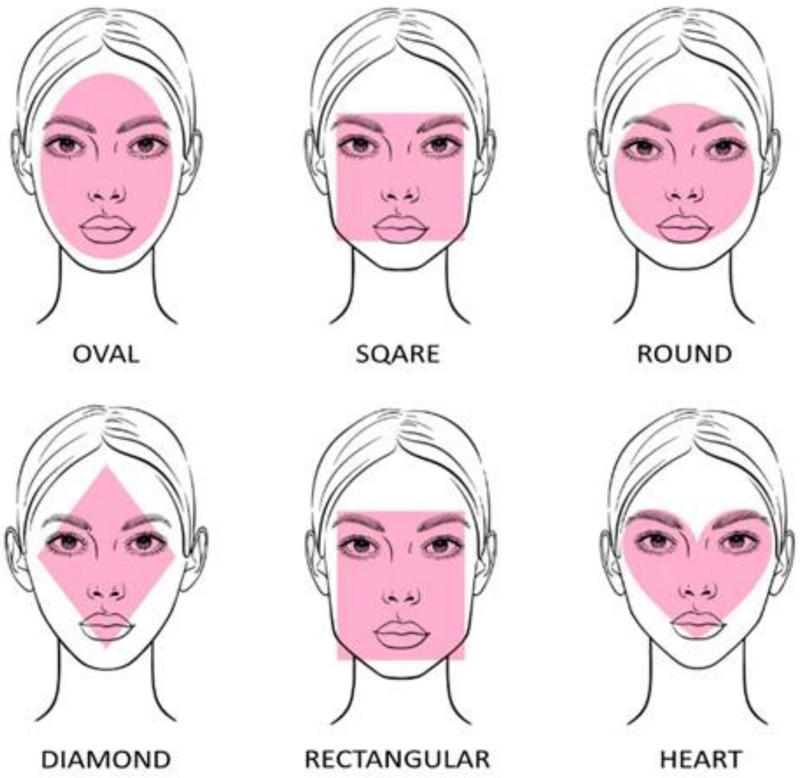
Now YOLOv? network has better performance for detection & for real time apps

Self study as example of detection

https://datascientest.com/en/you-only-look-once-yolo-what-is-it#:~:text=You%20Only%20Look%20Once%20or,the%20mainstays%20of%20computer%20vision.

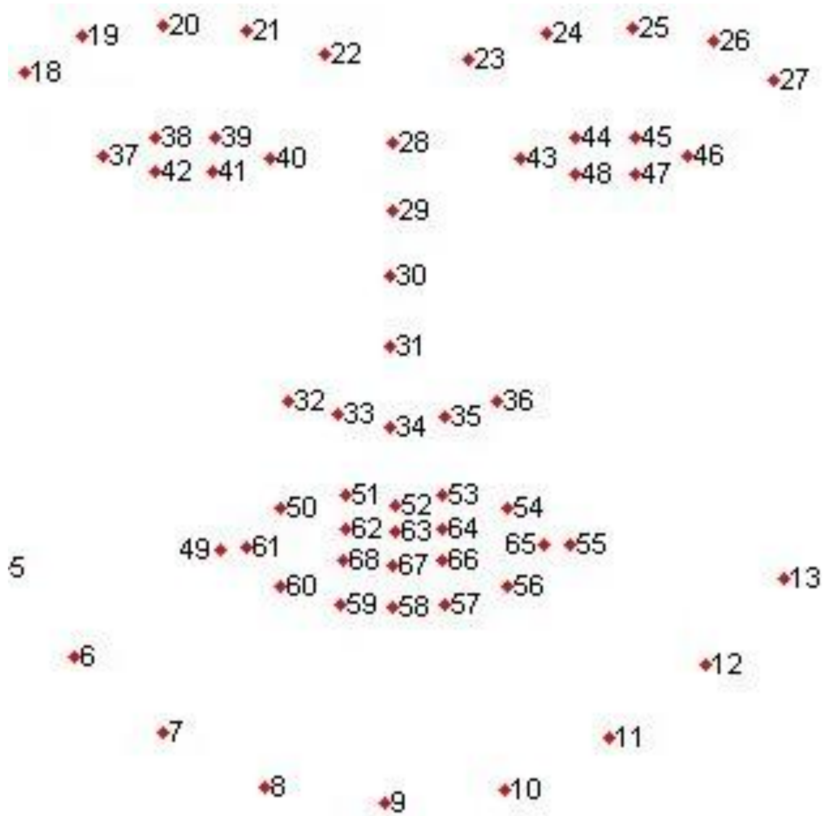
Few previous projects

- Face shape detection



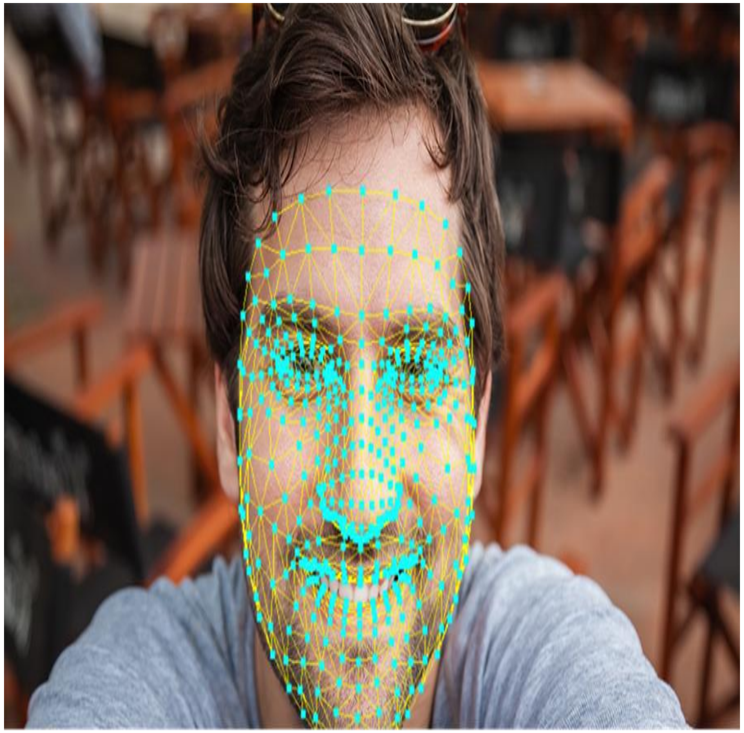
Face shape detection

1. Face detection
2. Facial landmarks detection



Facial landmarks detection

1. Dlib (68, 81 face landmarks)
2. Mediapipe library (face mesh)



Face parameters

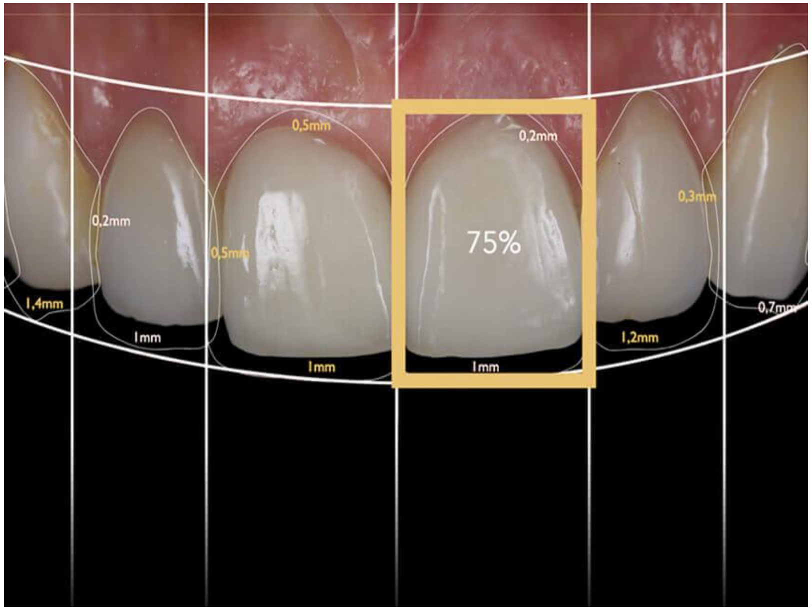
- Face length
- Forehead length
- Jawline length
- Chin width
- Cheekbone width
- Angle a1
- Angle a2
- Angle a3

Face shape detection

Face Shape	Condition
Heart	Forehead Width > Cheekbone Width > Jawline && a1>a2>a3
Oblong	Face Length > (Cheekbone Width ≈ Forehead Width ≈ Jawline) && a1 ≈ 90>a2>a3
Oval	Face Length > Cheekbone Width & Forehead Width > Jawline && a1 ≈a2≈a3
Square	Face Length ≈ Cheekbone Width ≈ Forehead Witch ≈ Jawline && a1 ≈ 90>a2 ≈ a3
Round	(Face Length ≈ Cheekbone Width) > (Forehead Width ≈ Jawline)

Apps and few Previous projects

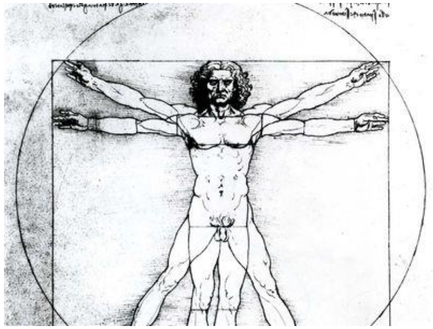
- Digital smile design and Golden Ratio
- Age and gender classification
- Beauty index and the Golden ratio (used in facial and dental surgery and alignment)

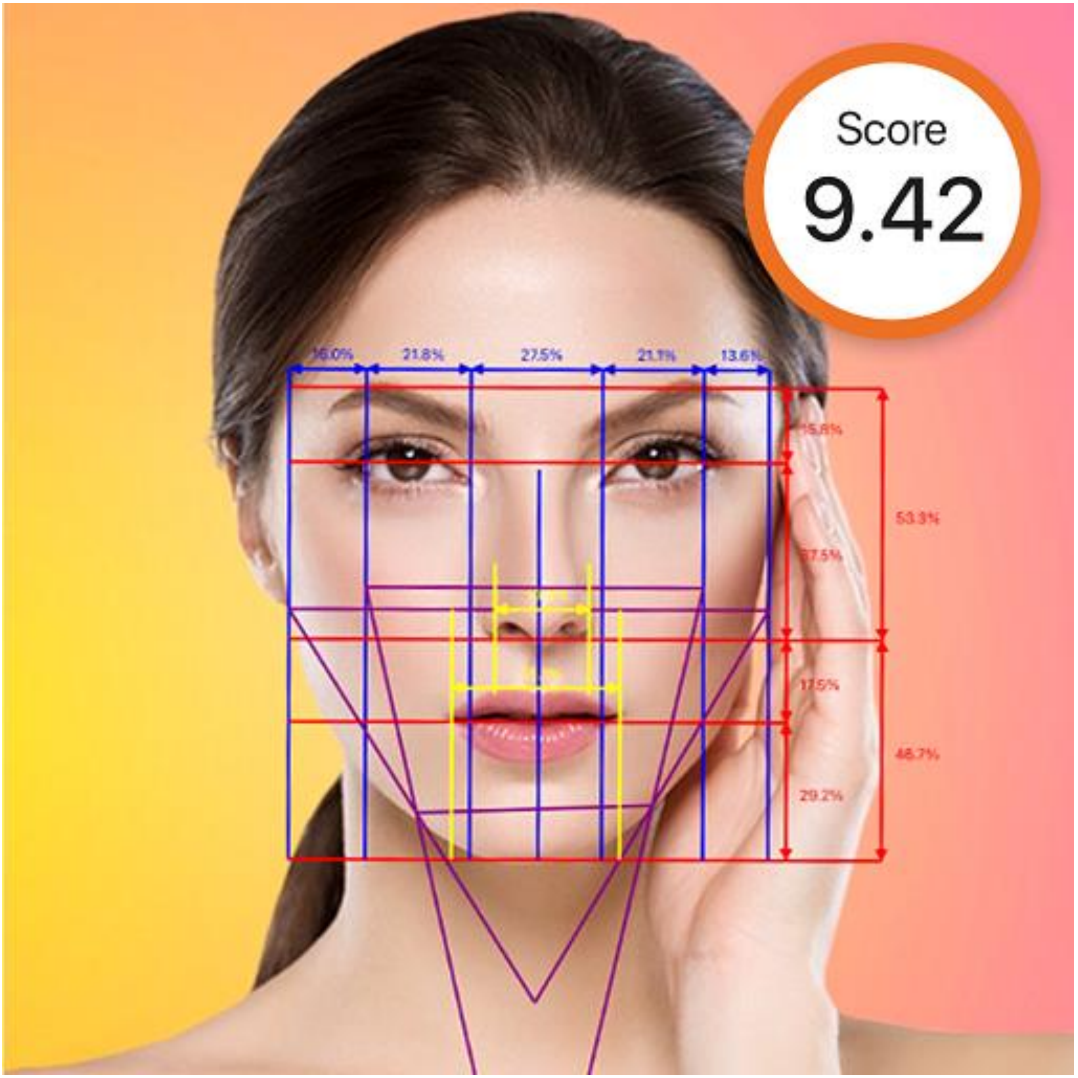


Beauty Score and Golden Ratio

It is suggested that a face is perceived as more aesthetically pleasing when its width is about 1.618 times the width of the mouth. $(1 + \sqrt{5})/2$, often denoted by the Greek letter ϕ.

And so on for eyes, nose, mouse, ..and teeth.





Digital Smile Design
(Dental Midline)