Object Recognition

Object recognition allows computers to detect and identify objects in images or videos.

Learning Objectives

- Understand the concept of object recognition
- Explain the workflow of object recognition systems
- Understand feature extraction
- Identify classification algorithms
- Understand deep learning approaches

What is Object Recognition?

- Object recognition is the process of detecting and identifying objects in images or videos.
- Example: A system analyzes an image and recognizes a 'cat'.

Examples of Object Recognition

- Face recognition in smartphones
- Self-driving cars detecting pedestrians
- Medical systems detecting tumors
- Industrial inspection detecting defects

Object Recognition Pipeline

- Image Acquisition → Preprocessing → Segmentation → Feature Extraction → Classification → Recognition

Image Acquisition

- Images are captured using cameras or sensors.
- Examples: Smartphone cameras, CCTV, satellite cameras, MRI scanners.

Image Preprocessing

- Improves image quality before analysis.
- Examples: Noise reduction, contrast enhancement, normalization.

Image Segmentation

- Separates objects from the background.
- Example: Extracting a car from a road scene.

Feature Extraction

- Important characteristics are extracted from objects.
- Examples: Shape, texture, color, edges.

Example Feature Vector

- Example features extracted from an apple:
- [Radius, Area, Color, Edge]
- [10, 314, Red, Smooth]

Classification

- Classification assigns a label to an object.
- Example result: Object = Apple, Confidence = 92%

Common Classification Algorithms

- K-Nearest Neighbor (KNN)
- Support Vector Machine (SVM)
- Neural Networks

Deep Learning for Object Recognition

- Modern systems use Convolutional Neural Networks (CNN).
- Examples: AlexNet, VGG, ResNet, YOLO.

Example: Face Recognition

Steps:

1. Capture image
2. Detect face
3. Extract features
4. Compare with database
5. Identify person

Challenges in Object Recognition

- Lighting variations
- Object occlusion
- Different viewpoints
- Complex backgrounds

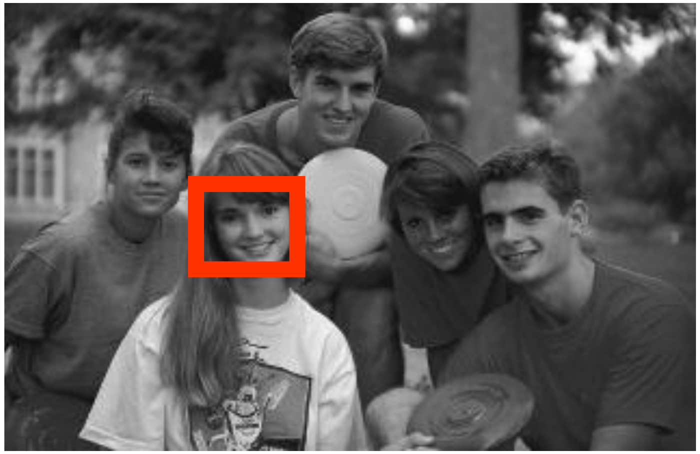
Applications

- Autonomous vehicles
- Healthcare imaging
- Security systems
- Retail product recognition

Object Recognition

Object Detection

Where is Jane?
Where is a Face?
Is there a face in the image?



Object Identification

Who is it?
Is it Jane or Erik?

General Problems of Recognition

Invariance:

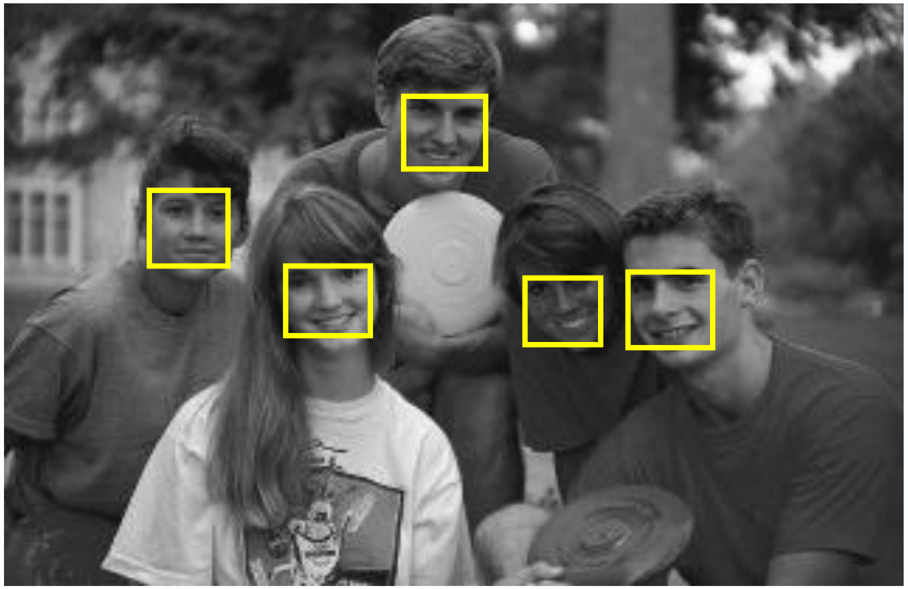
- “External parameters”
  - Pose
  - Illumination
- “Internal parameters”
  - Person identity
  - Facial expression

Applicable to many classes of objects

Object Detection

Task

Given an input image, determine if there are objects of a given class (e.g. faces, people, cars..) in the image and where they are located.



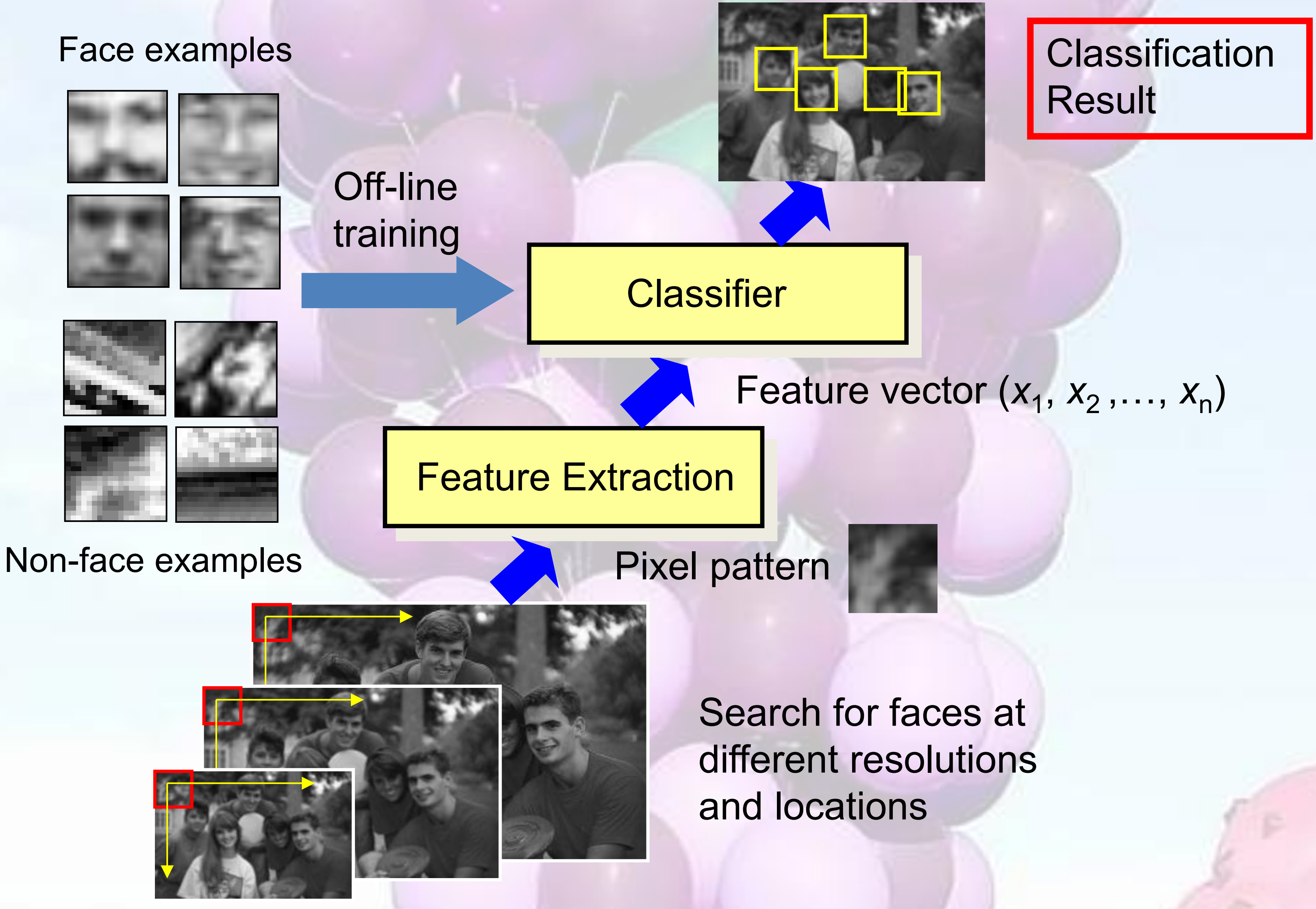
Detection—Problems

1. Classifier must generalize over all exemplars of one class.

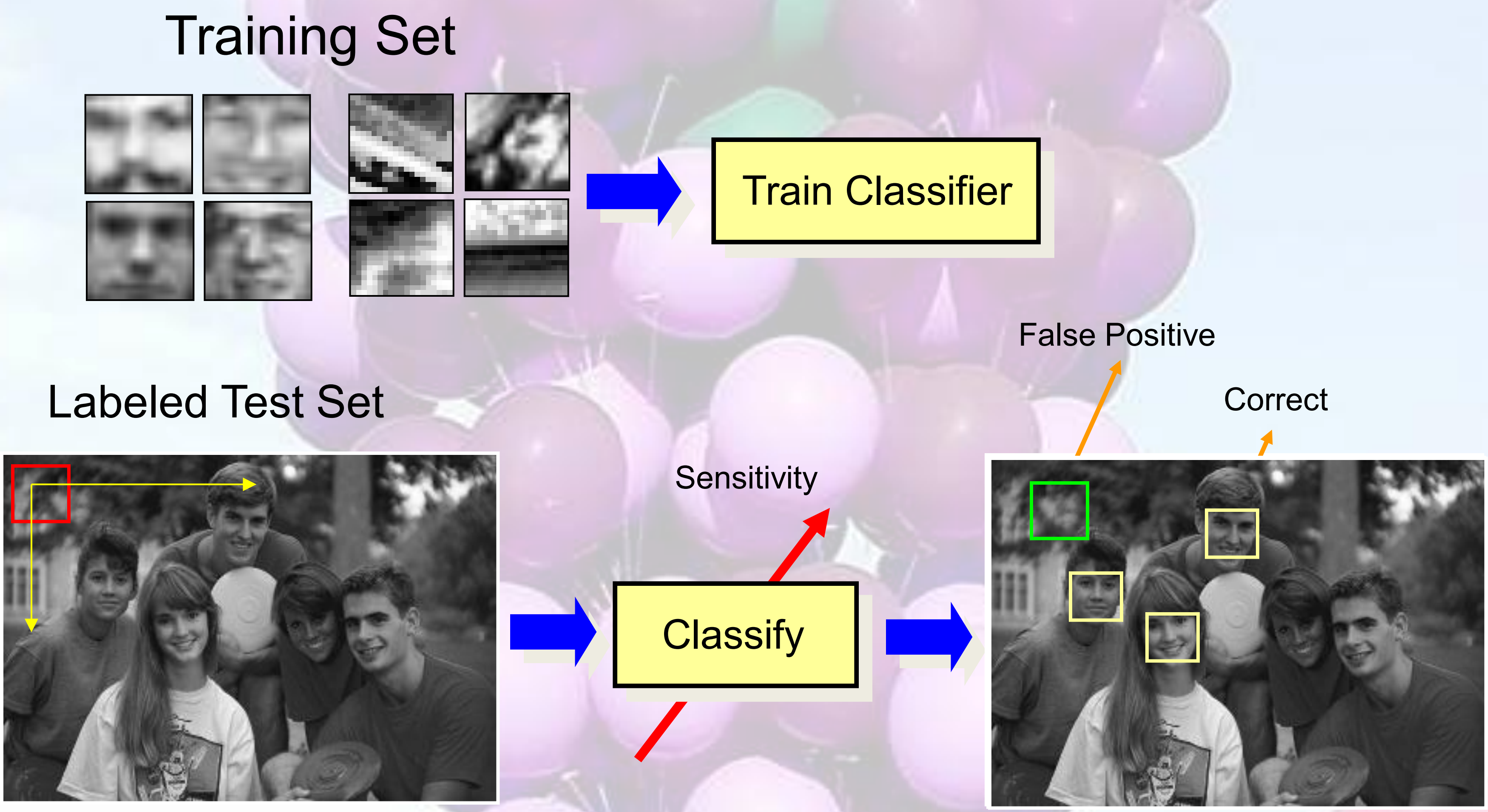
2. Negative class consists of everything else.

3. High accuracy (small FP rate) required for most applications.

Face Detection – basic scheme



Training and Testing



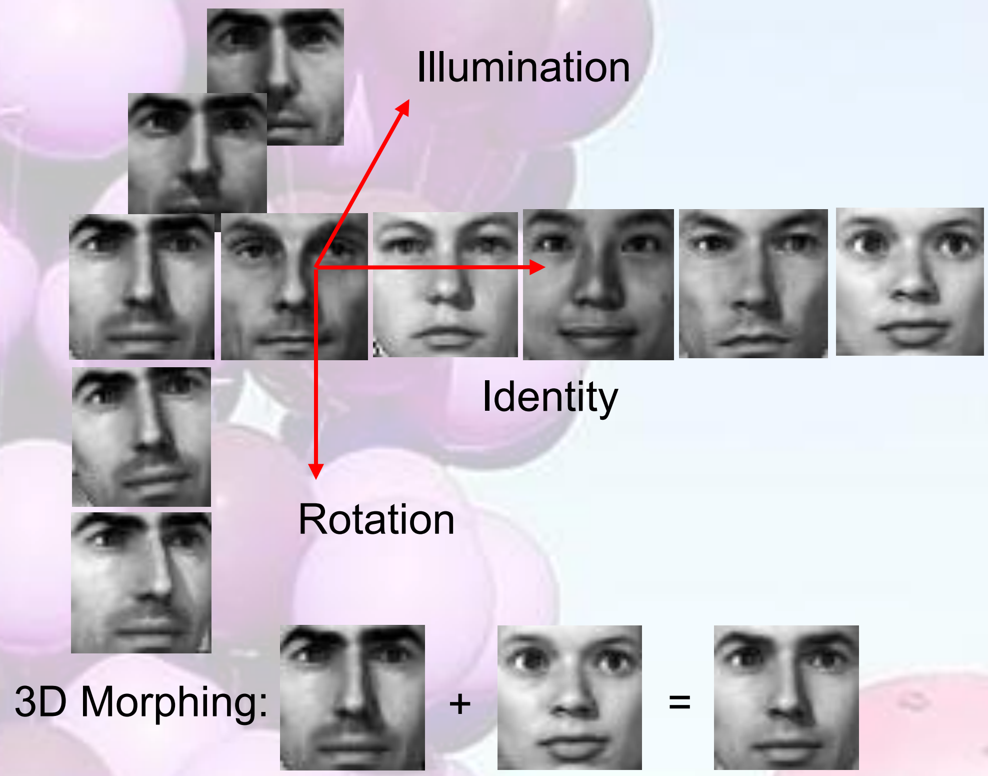
Positive Training Data

Real

2900 faces + 2900 mirrored faces

Synthetic

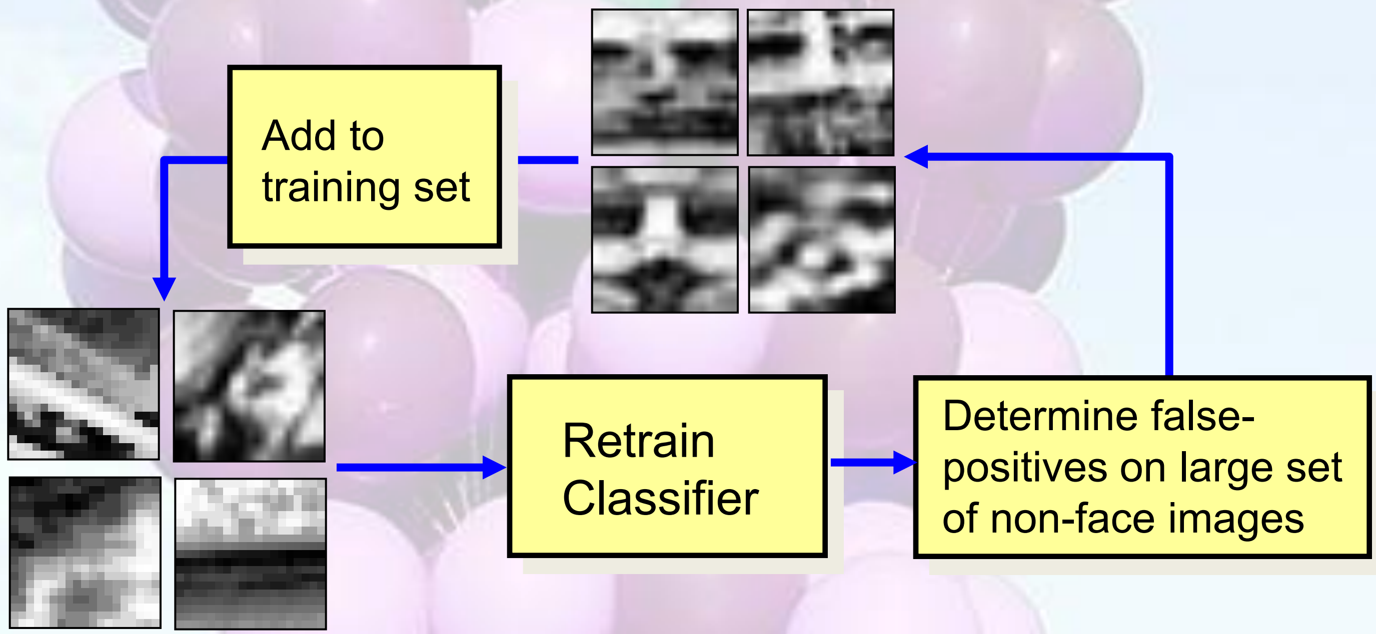
(T. Vetter, Univ. of Freiburg)



Negative Training Data

Problem: 1 face in 116.440 examined windows

Initial set of 25.000 non-faces



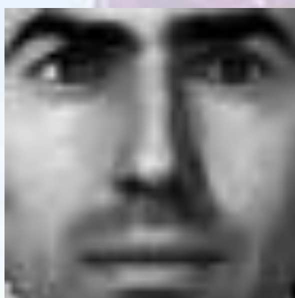
Bootstrapping

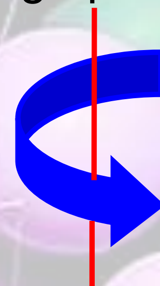
Performance of Global Face Detectors

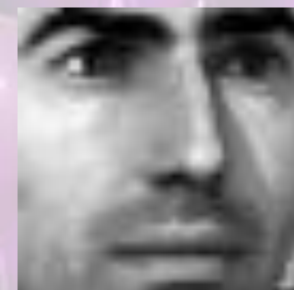
System	Subset of test set 1 23 images, 155 faces		Test set 1 130 images, 507 faces	
	Det. Rate	FPs	Det. Rate	FPs
[Sung 96] Neural Network	84.6%	13	N/A	N/A
[Osuna 98] SVM	74.2%	20	N/A	N/A
[Rowley et al. 98] Single neural network	N/A	N/A	90.9%	738
[Rowley et al. 98] Multiple neural networks	84.5%	8	84.4%	79
[Schneiderman & Kanade 98][3] Naïve Bayes	91.1%	12	90.5%	33
[Yang et al. 99][4] SNoW, multi-scale	94.1%	3	94.8%	78
Our system[5]	84.7% 90.4%	11 26	85.6% 89.9%	9 75

Rotation

Rotation out of
image plane







Rotation in the
image plane

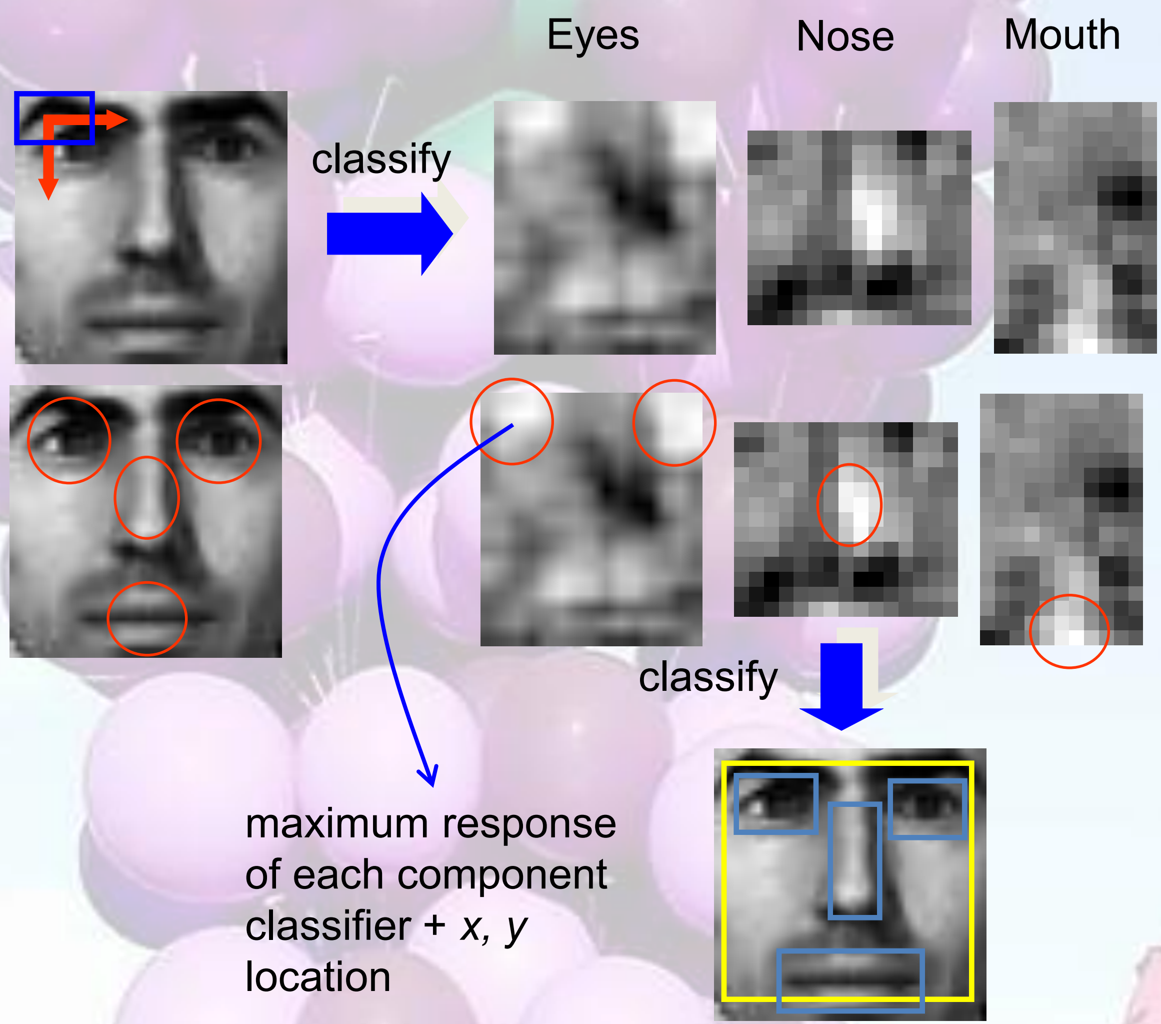
- Component-based classification
- Train on rotated faces

- Rotation invariant features
- Apply 2D rotation to image

Component-based Detection

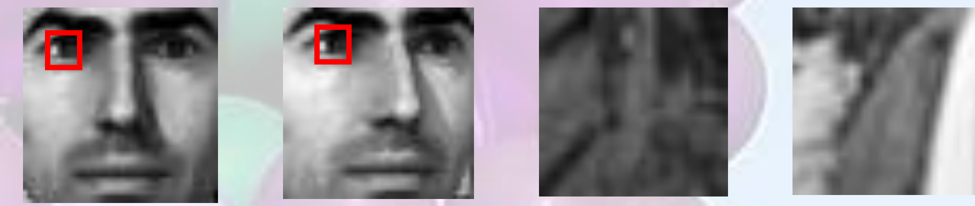
1st Level: Components

2nd Level: Geometrical relation between components

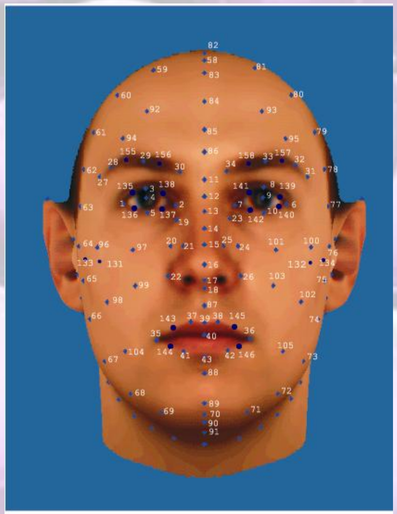


Learning Components

Components:
- discriminatory
- robust against changes in pose and illumination



Synthetic faces:
- 7 different 3-D head models
- 2,500 faces
  Rotation: -30^{o} to $+30^{o}$
- 3-D correspondences for automatic location of components



Learning Components—One Way To Do It

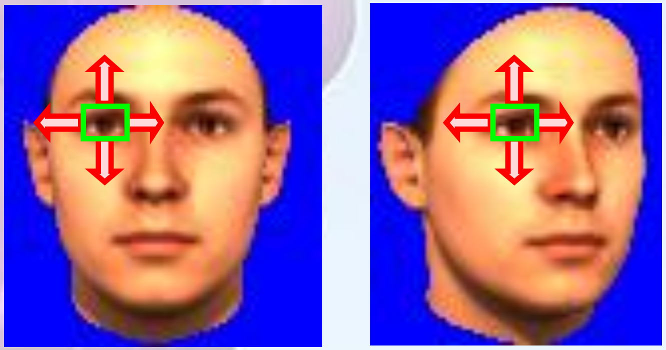
Start with small initial regions

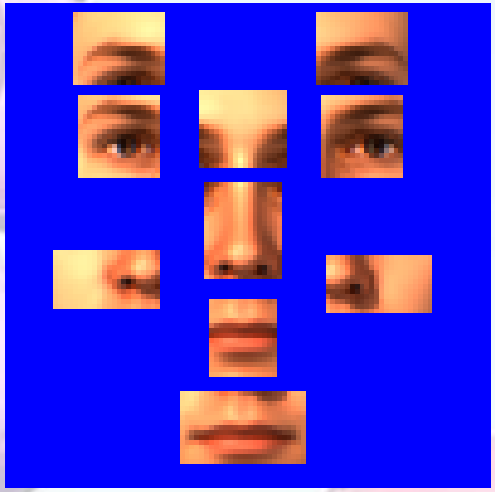
Expand into one of four directions

Extract new components from images

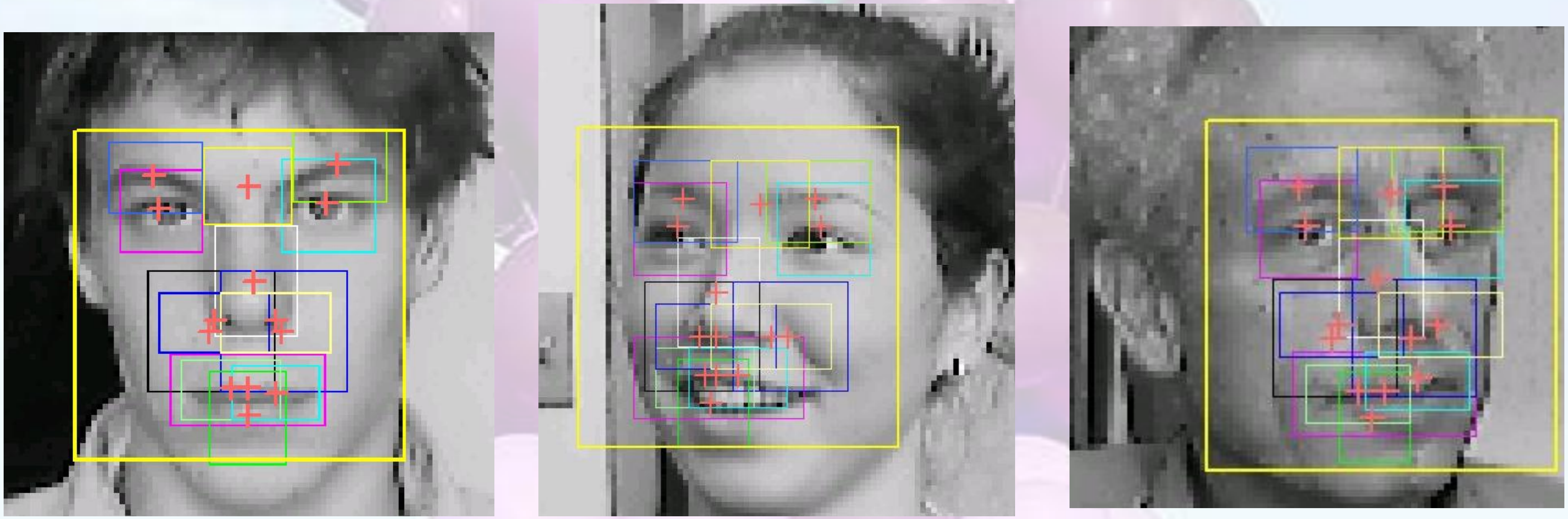
Train SVM classifiers

Choose best expansion according to error bound of SVMs





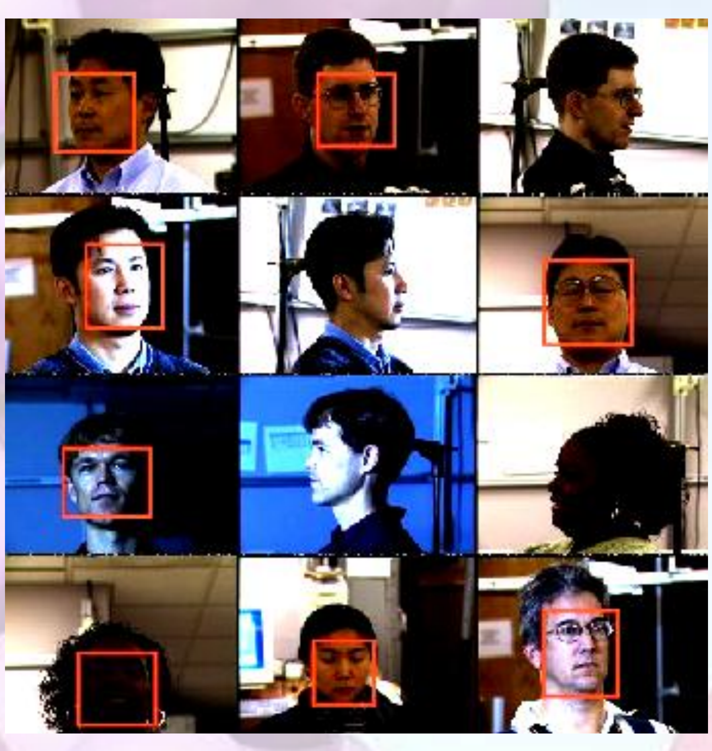
Some Examples



Test on CMU PIE Database

Faces have been manually labeled
(only –45^o to 45^o of rotation)

- About 40,000 faces
- 68 people
- 13 poses
- 43 illumination conditions
- 4 different expressions

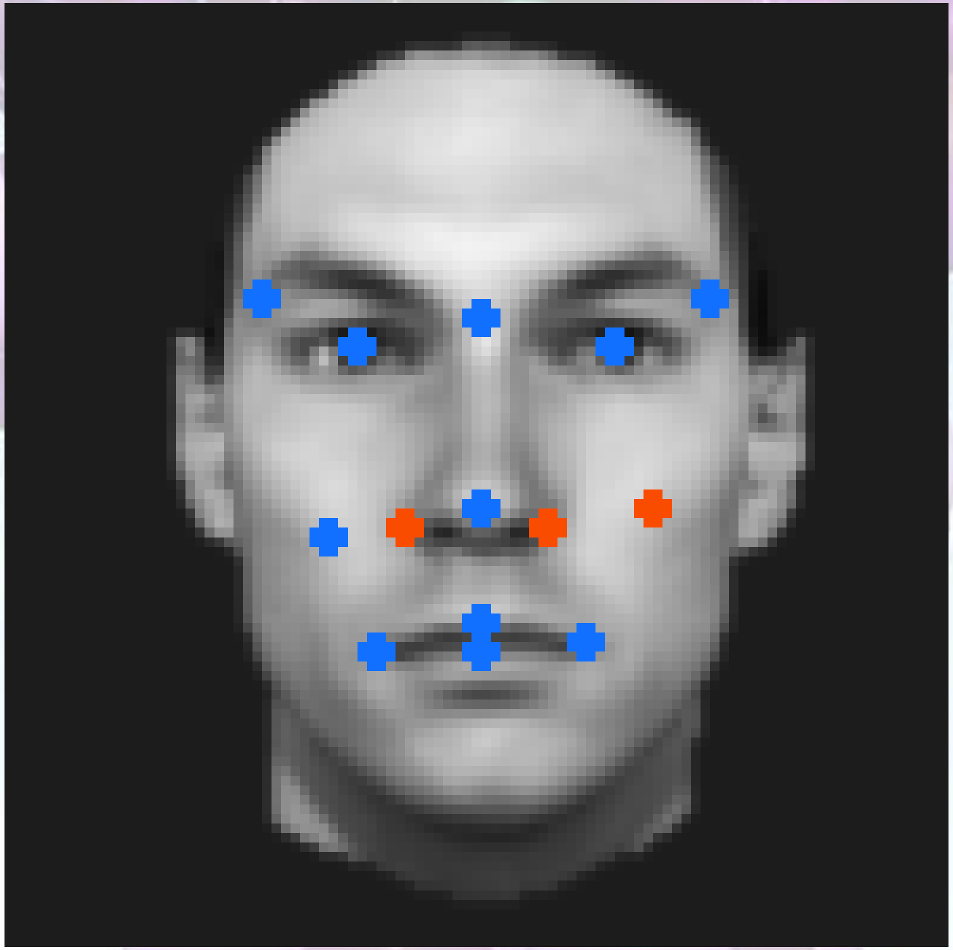


Advances on Component-base Face Detection Stan Bileschi

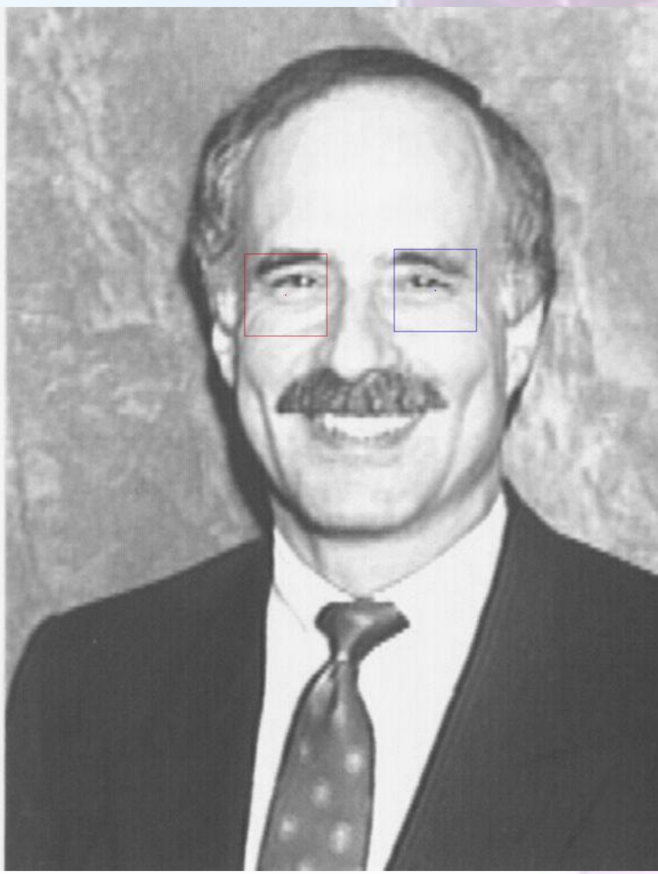
Components are small, and prone to false detection, even within the face.

Stan Bileschi

Often, many components classify correctly, with only a few errors



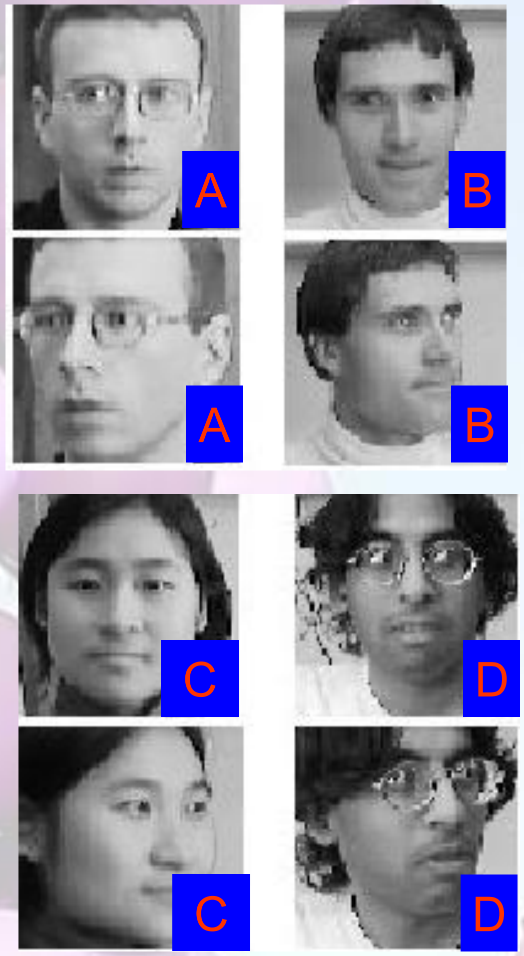
Application: Eye Detection



Identification

Task:

Given an image of an object of a particular class (e.g. face) identify which exemplar it is.



Identification—Problems

1. Multi-class problem

2. Classifier must distinguish between exemplars that might look very similar.

3. Classifier has to reject exemplars that were not in the training database.

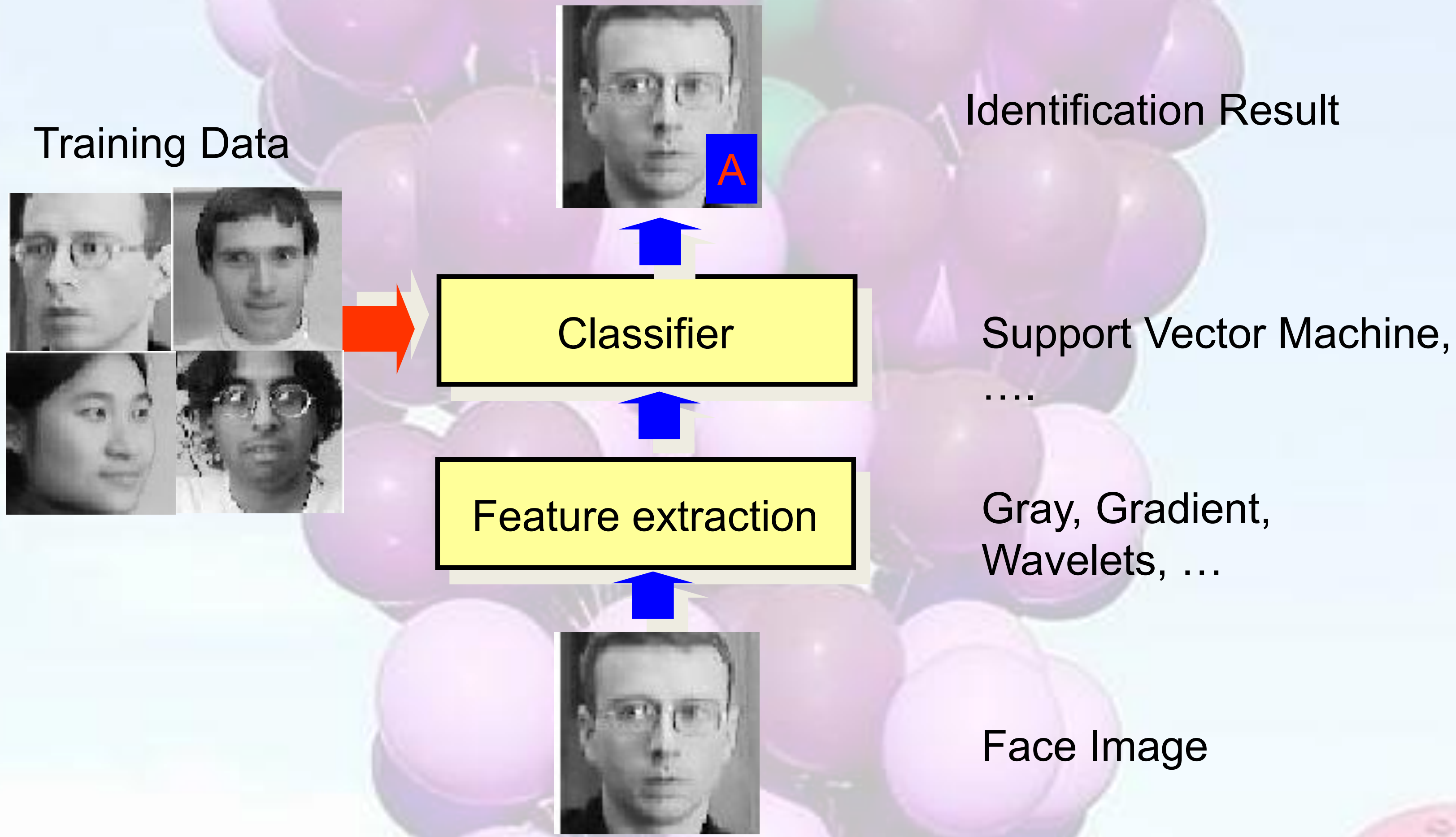
Problems in Face Identification

Limited information in a single face image

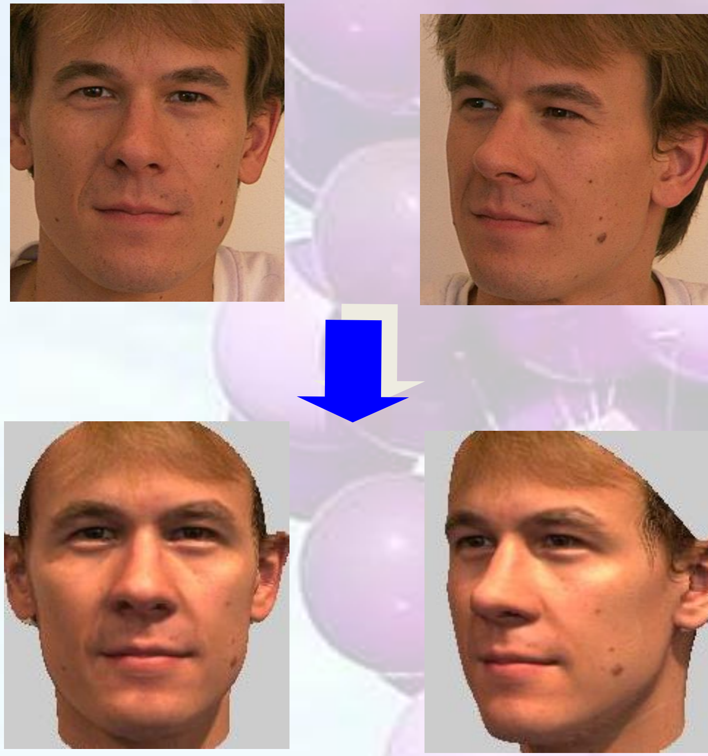
Illumination

Rotation

System Architecture



Morphable Models for Face Identification, Jennifer Huang



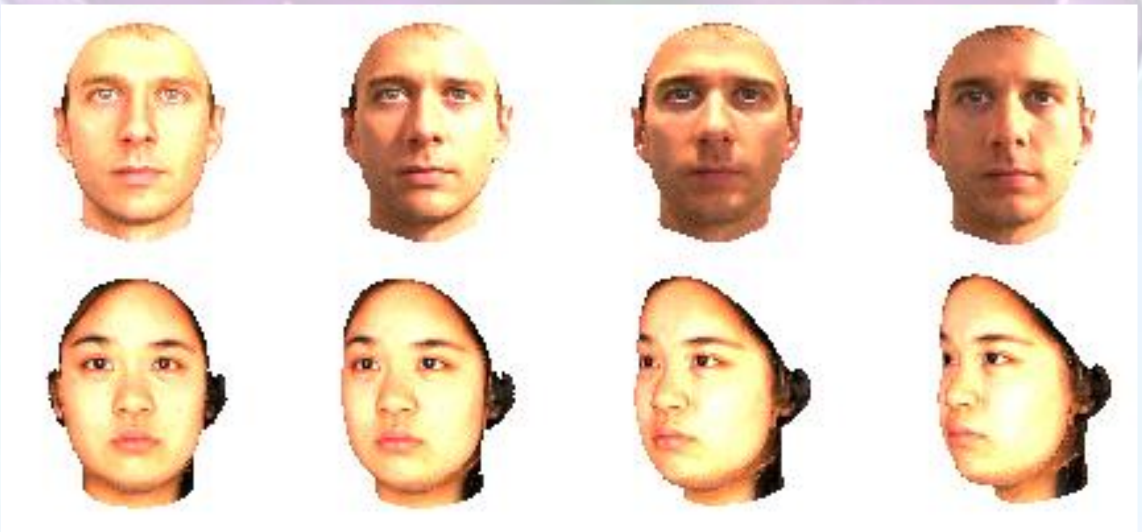
3D morphable models

Training data for component-based face recognition

A Morphable Model for the Synthesis of 3D Faces.
Blanz, V. and Vetter, T. SIGGRAPH'99 Conference Proceedings, pp. 187-194

Jennifer Huang

Some Training Images



Synthetic images are easily rendered from 3D head model under varying illuminations and rotations in depth

Jennifer Huang

Current Work – Testing on Real Images

Problems Encountered:

Detection

Inaccurate Component detection

Recognition

Accuracy of 3D models

Choice of Illumination and Pose

Literature

B. Heisele, A. Verri and T. Poggio: *Learning and Vision Machines*. Proceedings of the IEEE, Visual Perception: Technology and Tools, Vol. 90, No. 7, pp. 1164-1177, 2002.

See also CBCL Web page

Summary

- Object recognition identifies objects using segmentation, feature extraction, and classification.
- Modern systems often use deep learning.