# 9.913 Pattern Recognition for Vision

## Class 9 – Object Detection and Recognition

Bernd Heisele

# Overview

1. Detection
2. Recognition
3. Demo

# General Problems of Detection and Recognition

Rotation/illumination invariance

Images removed due to copyright considerations.

Applicable to many classes of objects

Images removed due to copyright considerations.

# 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.

Photograph by MIT OCW.

# Detection—Problems

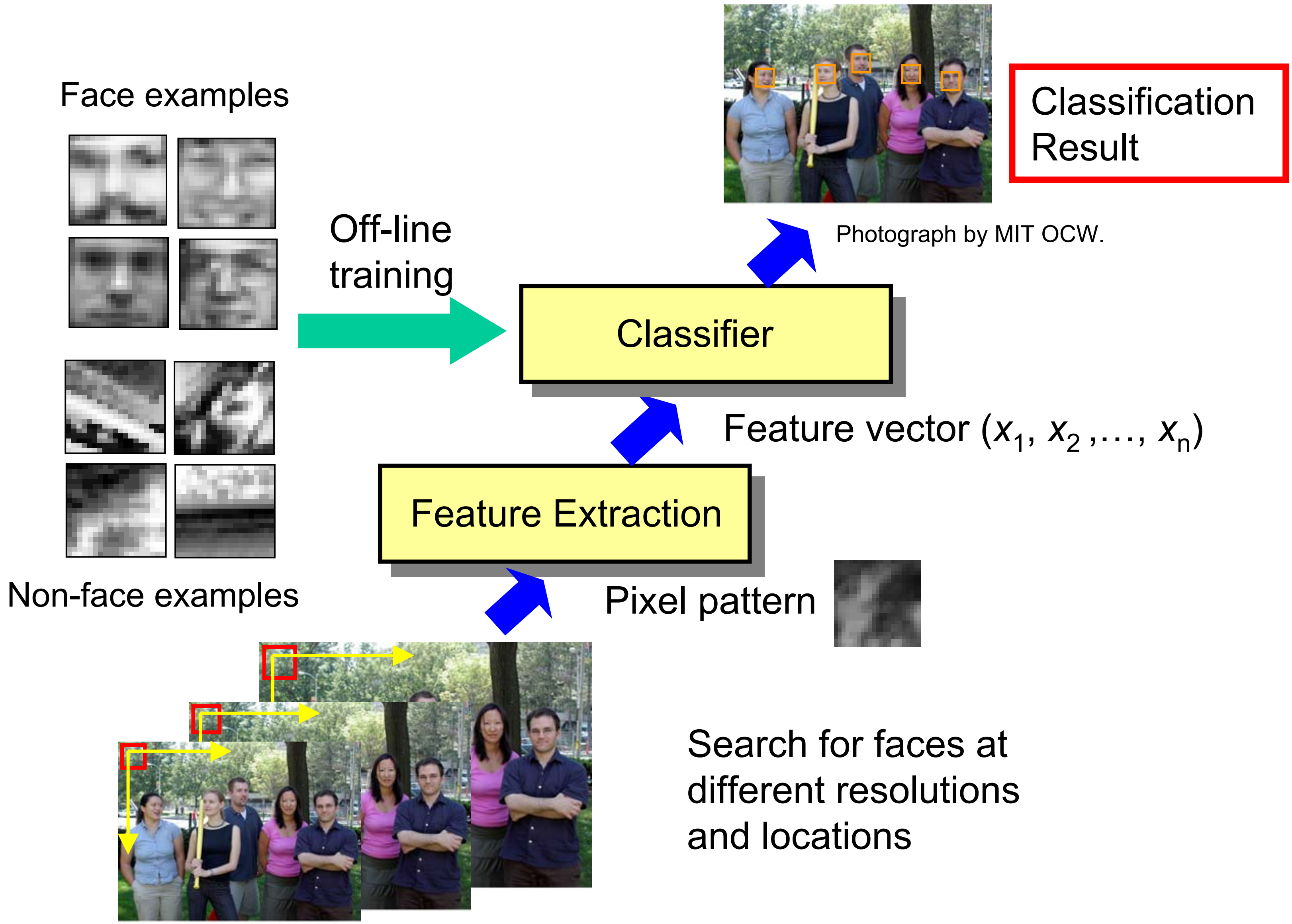
1. Classifier must generalize over all exemplars of one class.

2. Negative class consists of everything else.

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

Images removed due to copyright considerations.

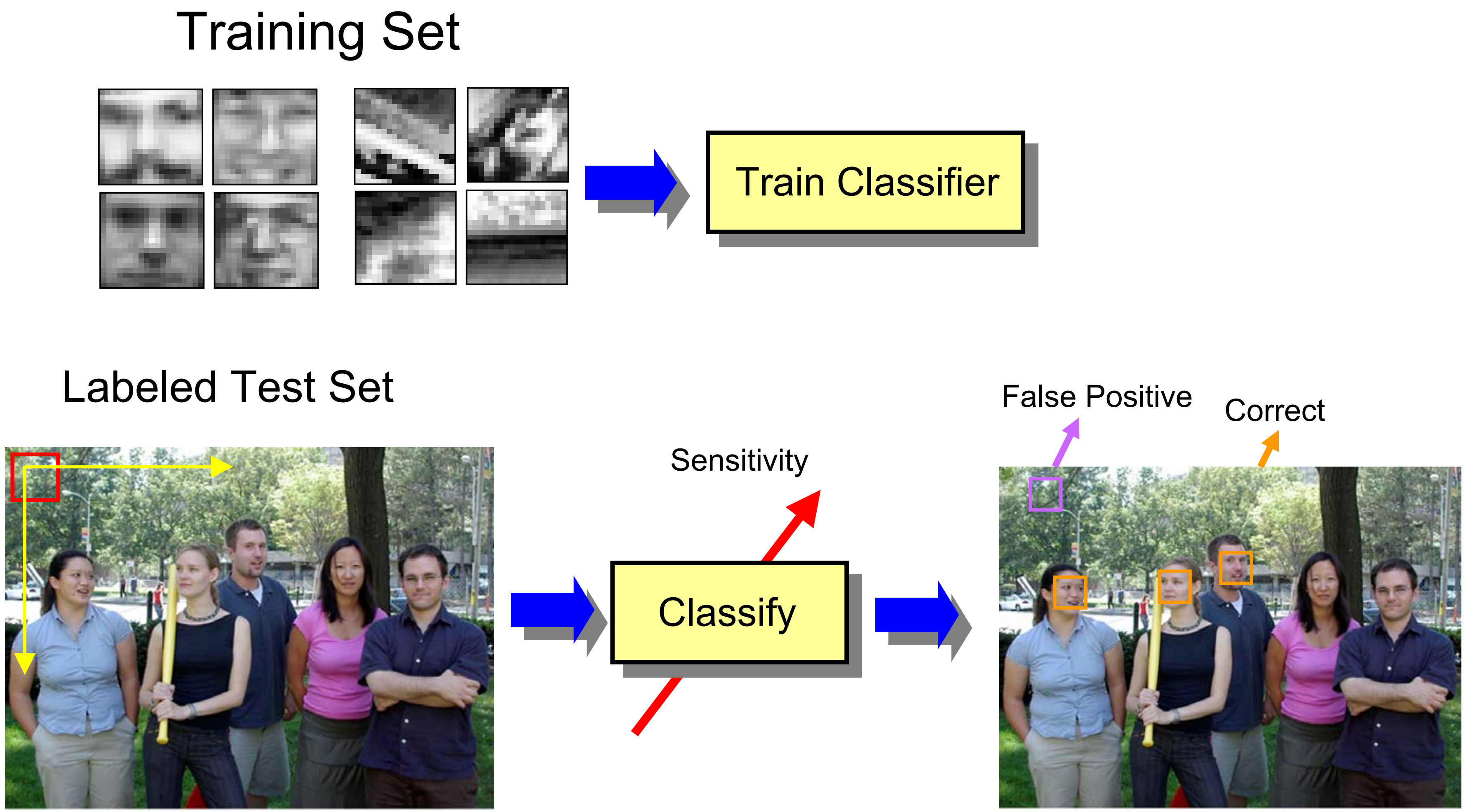
# Face Detection






Photograph by MIT OCW.

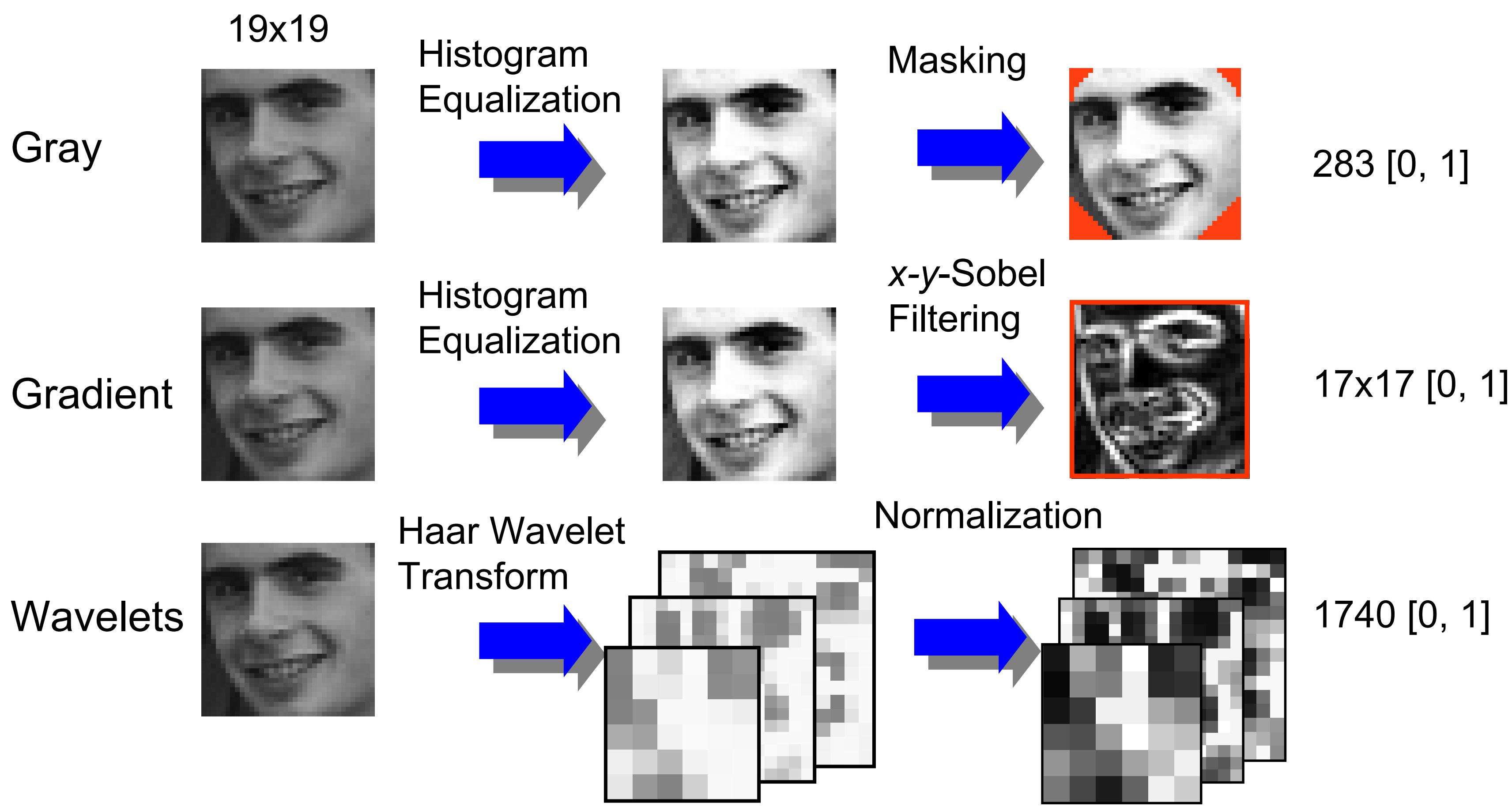
# Training and Testing



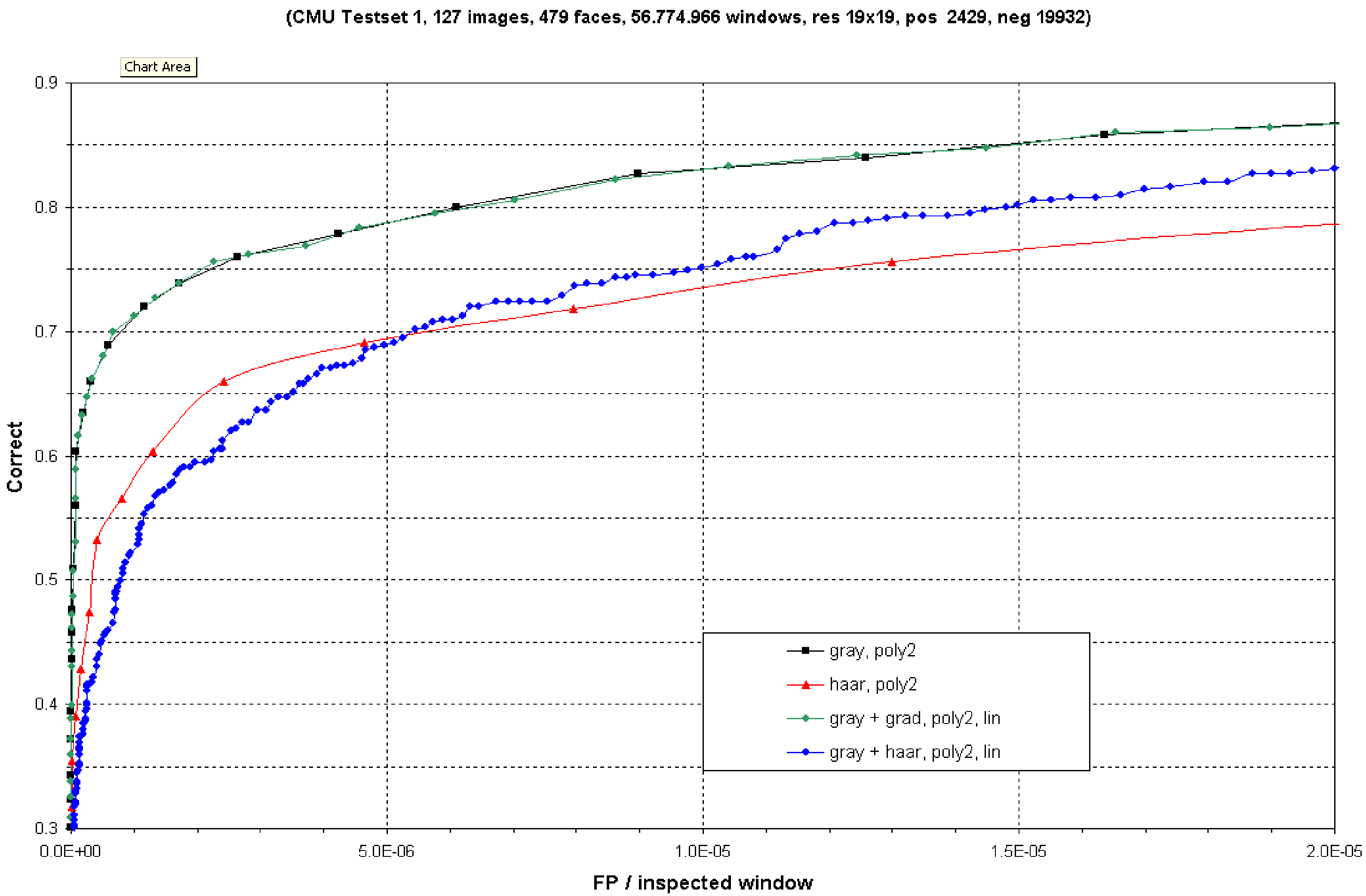
Photograph by MIT OCW.

Photograph by MIT OCW.

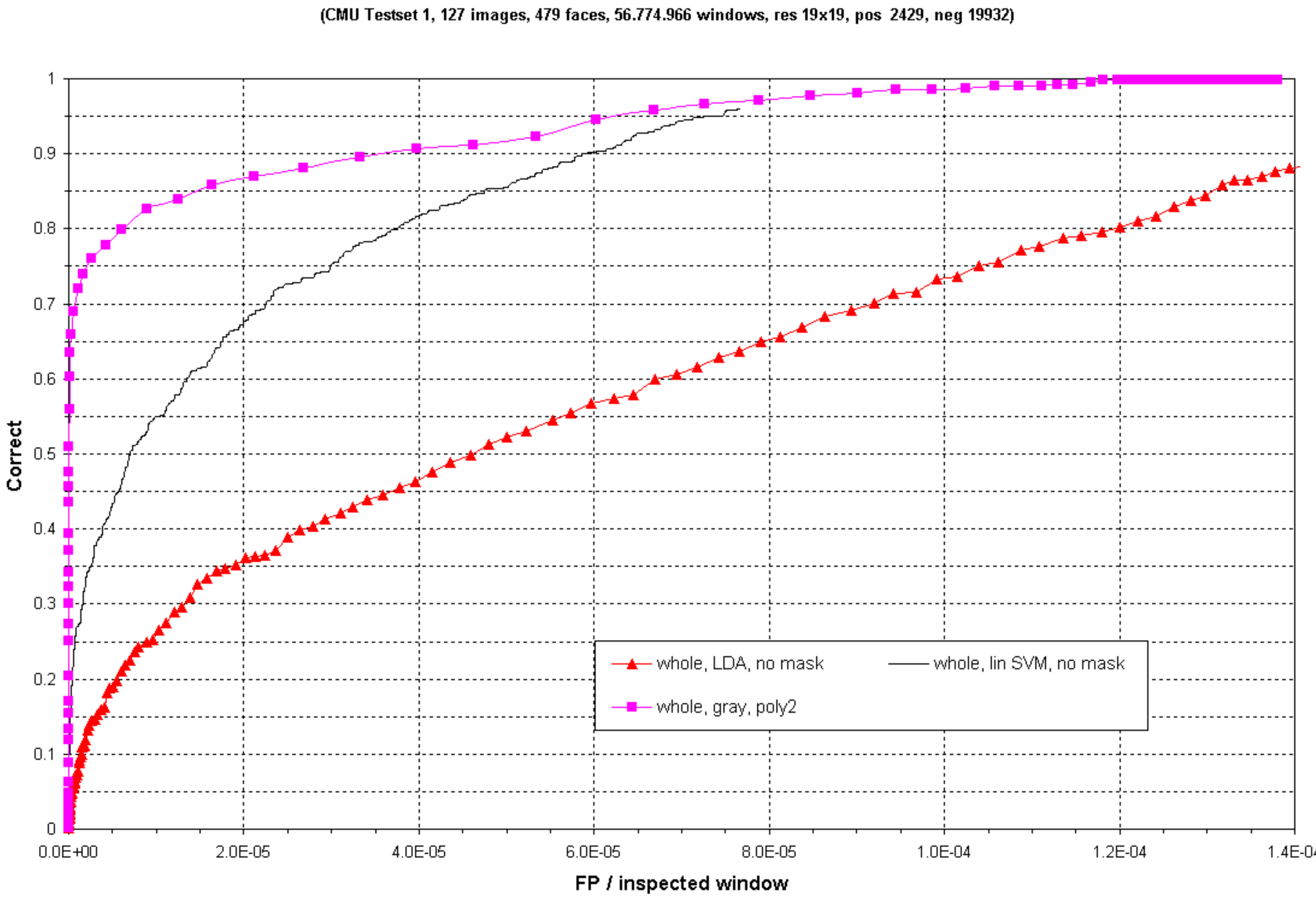
# Image Features



# ROC Image Features



# Classifiers



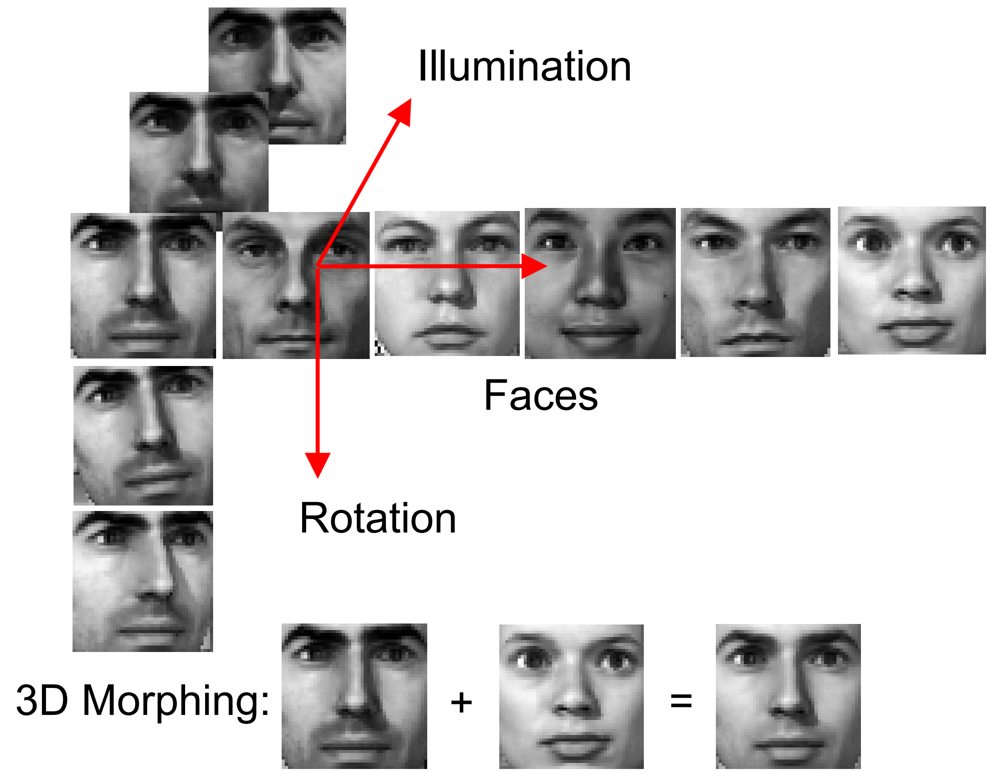
# Positive Training Data

## Real

2900 faces + 2900 mirrored faces

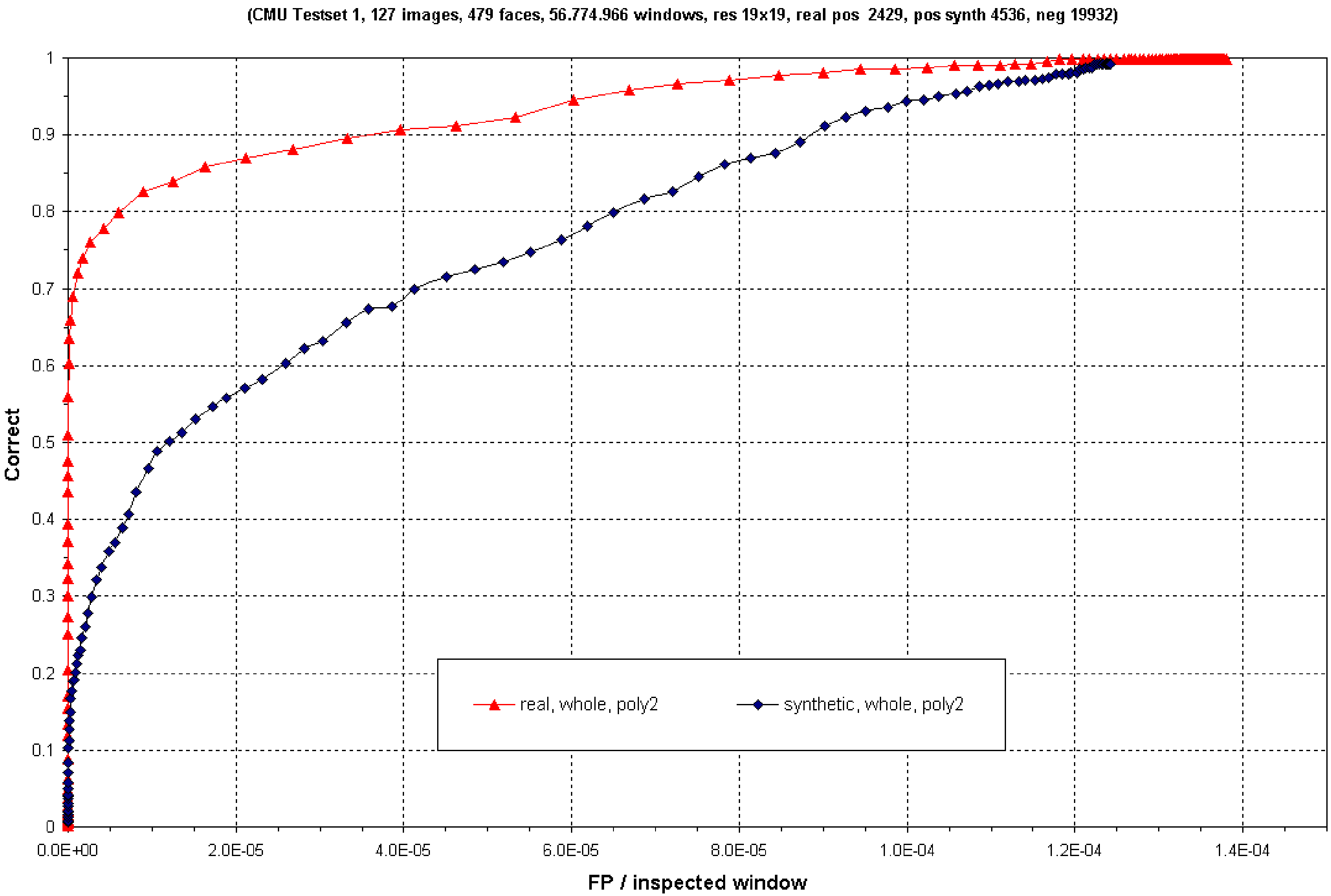
## Synthetic

(T. Vetter, Univ. of Freiburg)



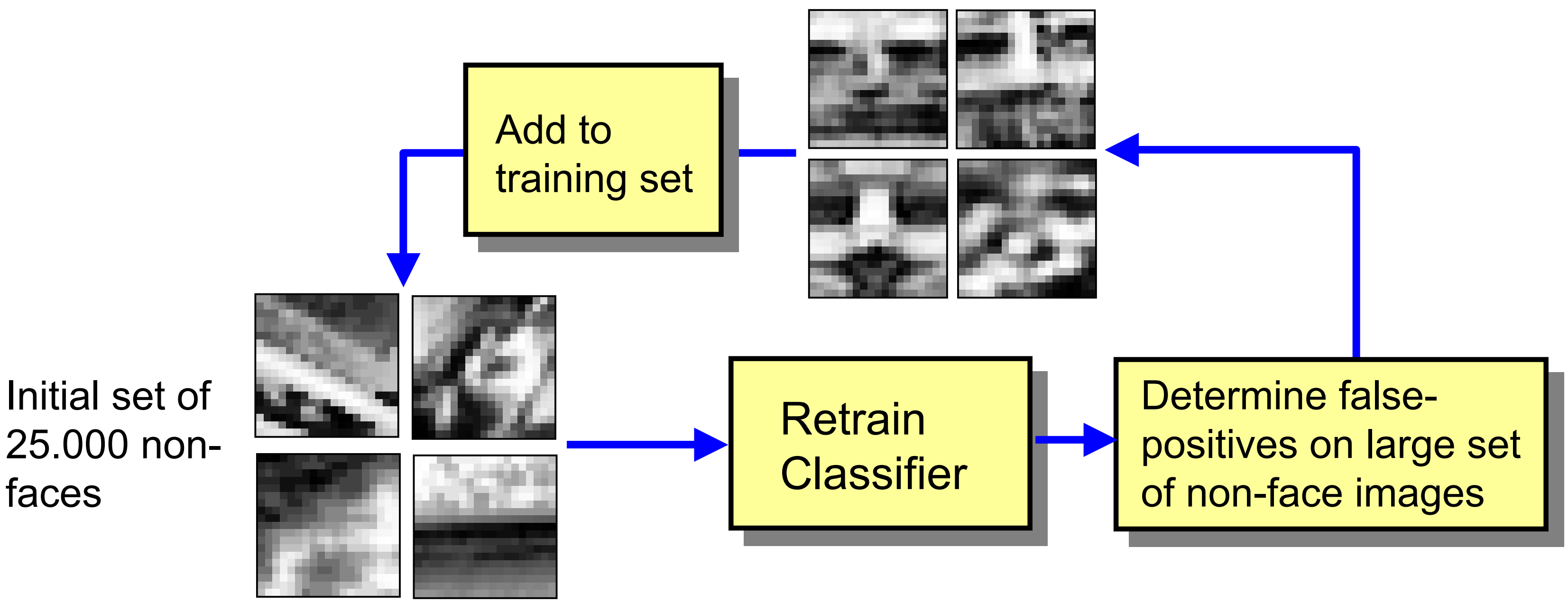

Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Real vs. Synthetic



# Negative Training Data

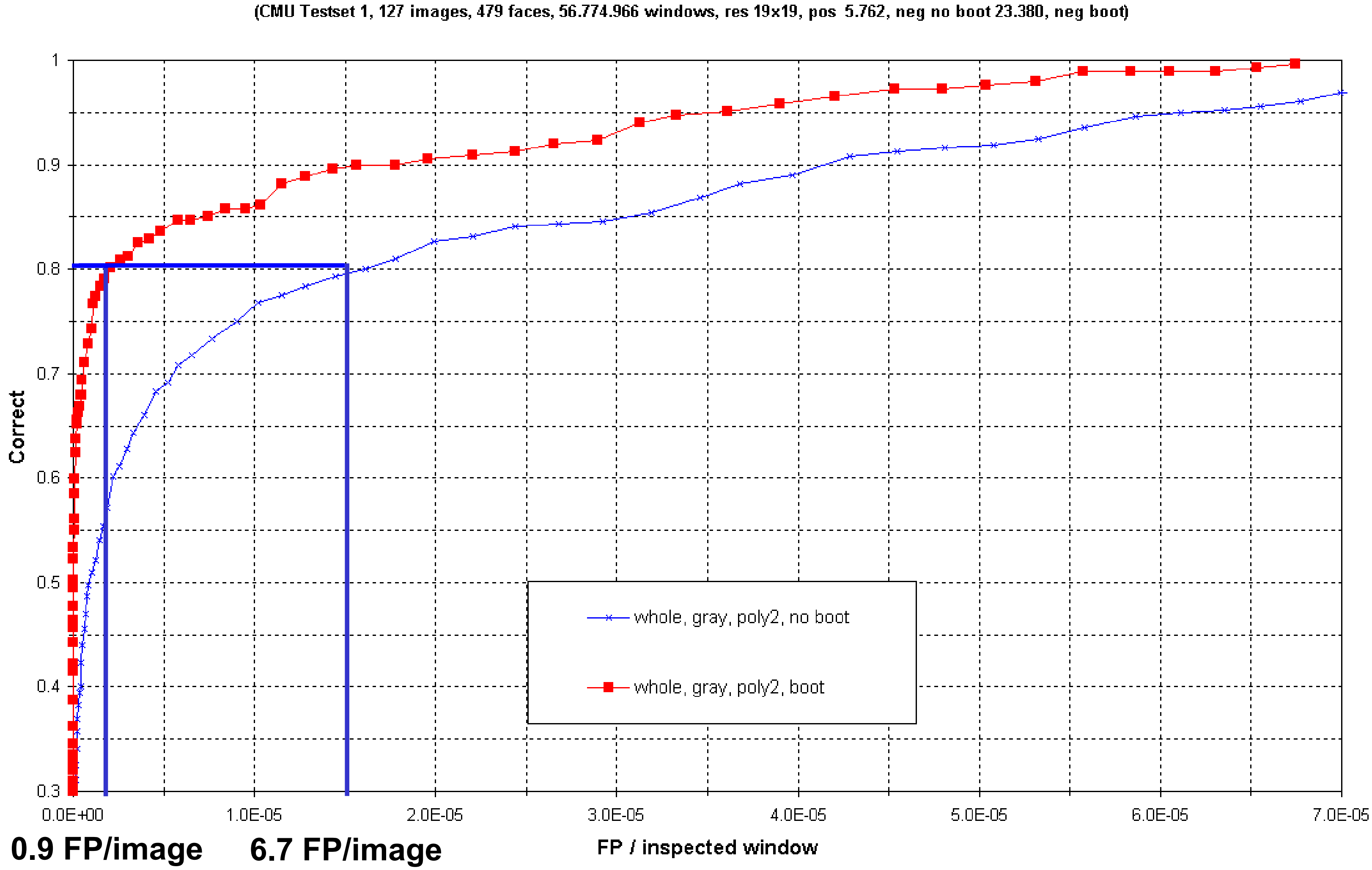
Problem: 1 face in 116.440 examined windows




**Bootstrapping**

Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Bootstrapping

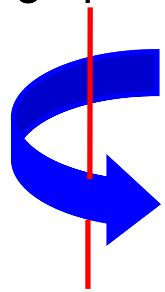


# Performance of Global Face Detectors

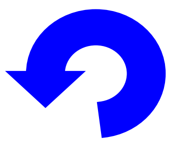
| System | Subset of test set 1<br>23 images, 155 faces | | Test set 1<br>130 images, 507 faces | |
|---|---|---|---|---|
| | Det. Rate | FPs | Det. Rate | FPs |
| [Sung 96]<br>Neural Network | 84.6% | 13 | N/A | N/A |
| [Osuna 98]<br>SVM | 74.2% | 20 | N/A | N/A |
| [Rowley et al. 98]<br>Single neural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

# Rotation

Rotation out of image plane



- Component-based classification
- Train on rotated faces



Rotation in the image plane

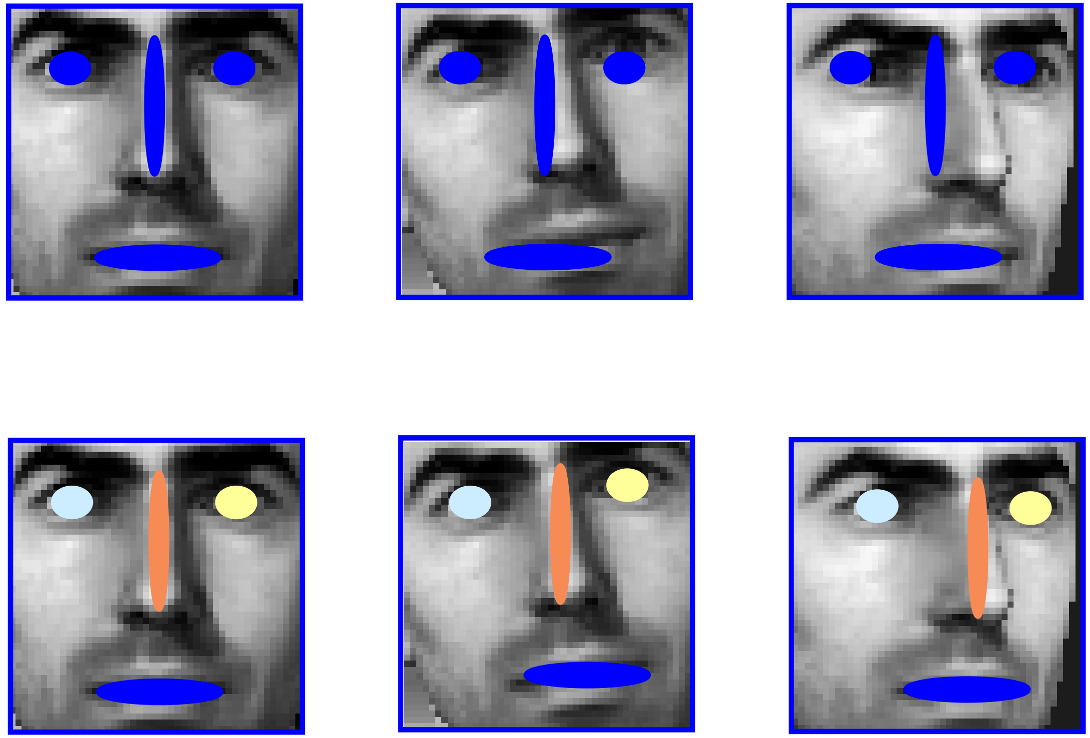
- Rotation invariant features
- Apply 2D rotation to image

Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Global vs. Components

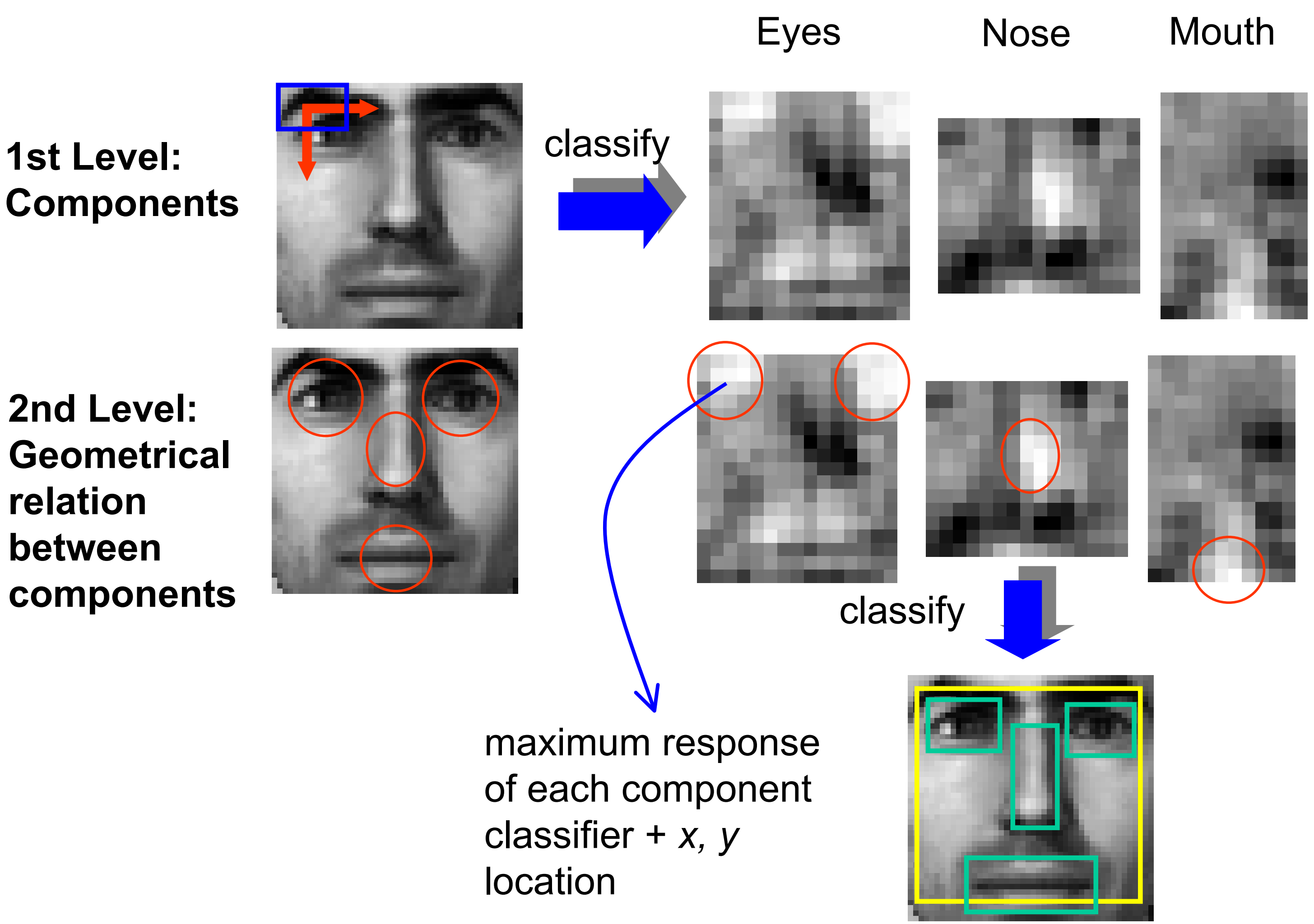
Single template

**Component templates**



Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Component-based Detection

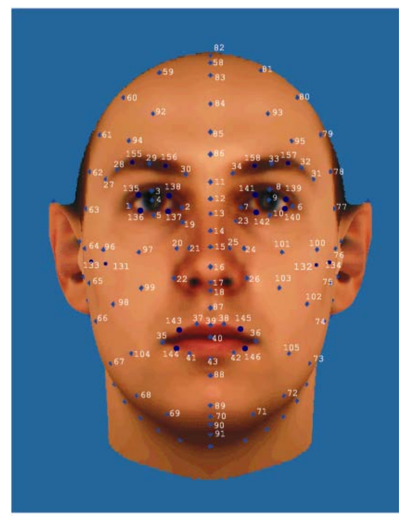


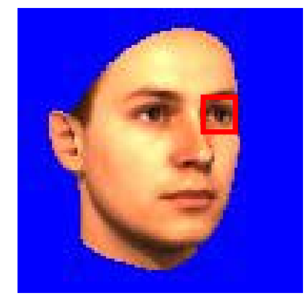

Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

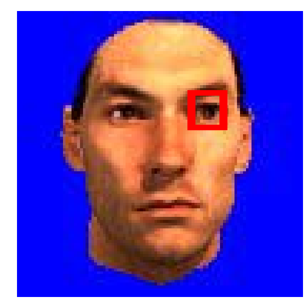
# Learning Components

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

**Synthetic faces:**
- 7 different 3-D head models
- 2,500 faces
  Rotation: -30° to + 30°
- 3-D correspondences for automatic location of components







Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Learning Components—One Way To Do It

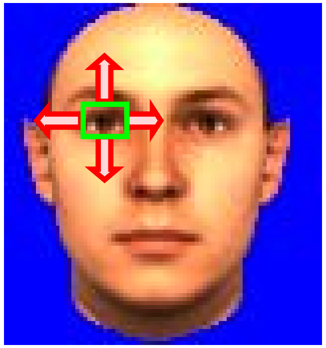
Start with small initial regions

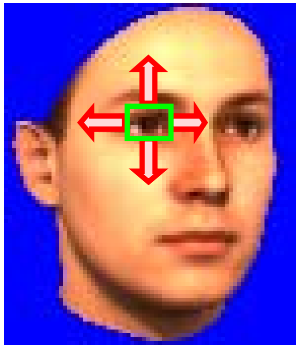
Expand into one of four directions

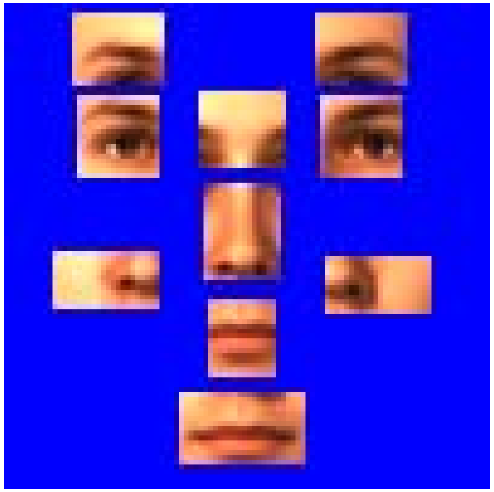
Extract new components from images

Train SVM classifiers

Choose best expansion according to error bound of SVMs

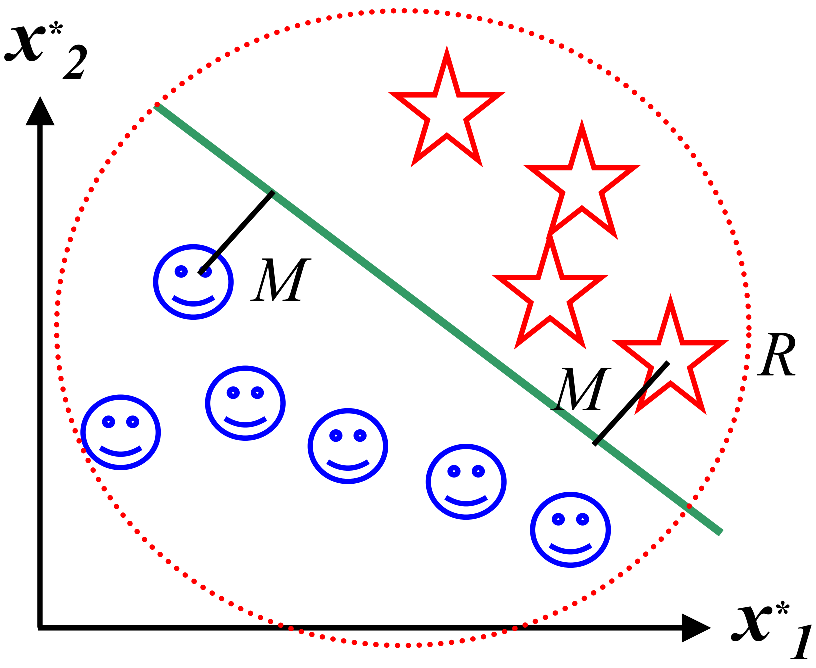






Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Margin, Radius and Expected Error

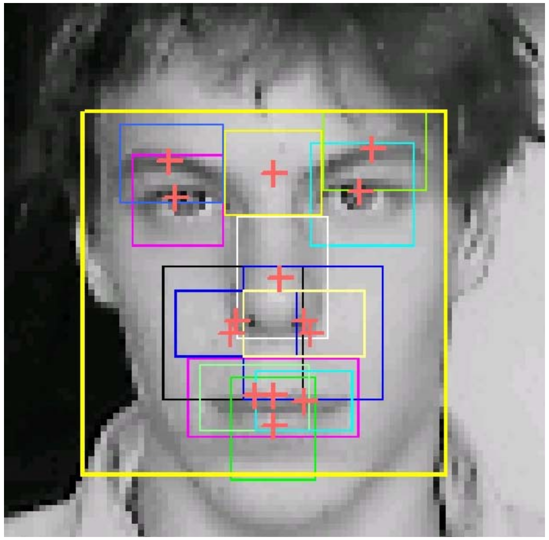
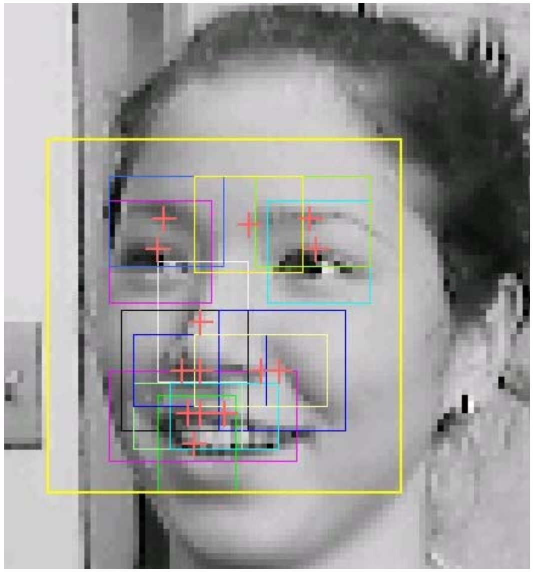
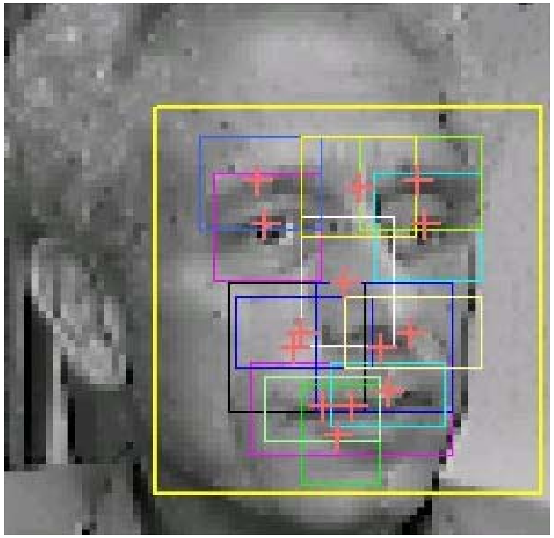



Feature Space

Bound on error
$E < c(R / M)^2$

Cross Validation might be better

# Some Examples

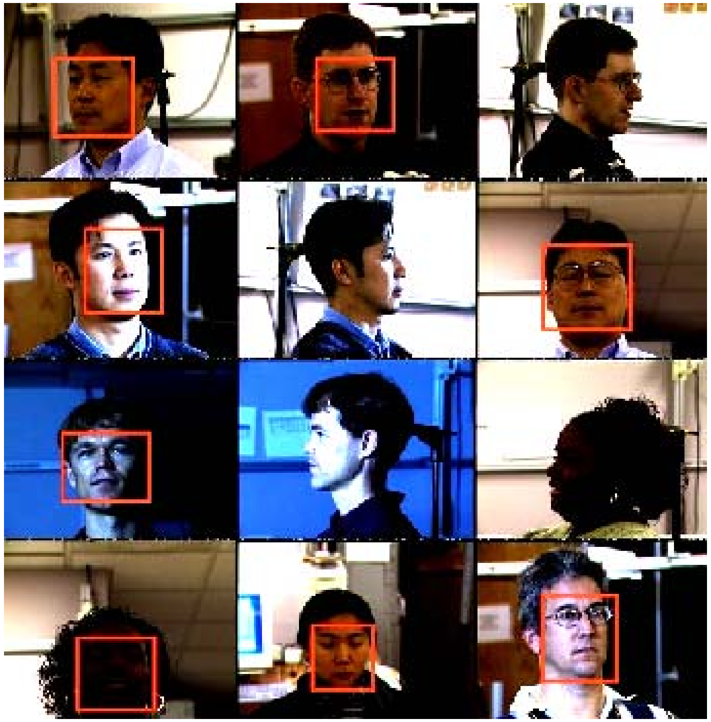




Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Test on CMU PIE Database

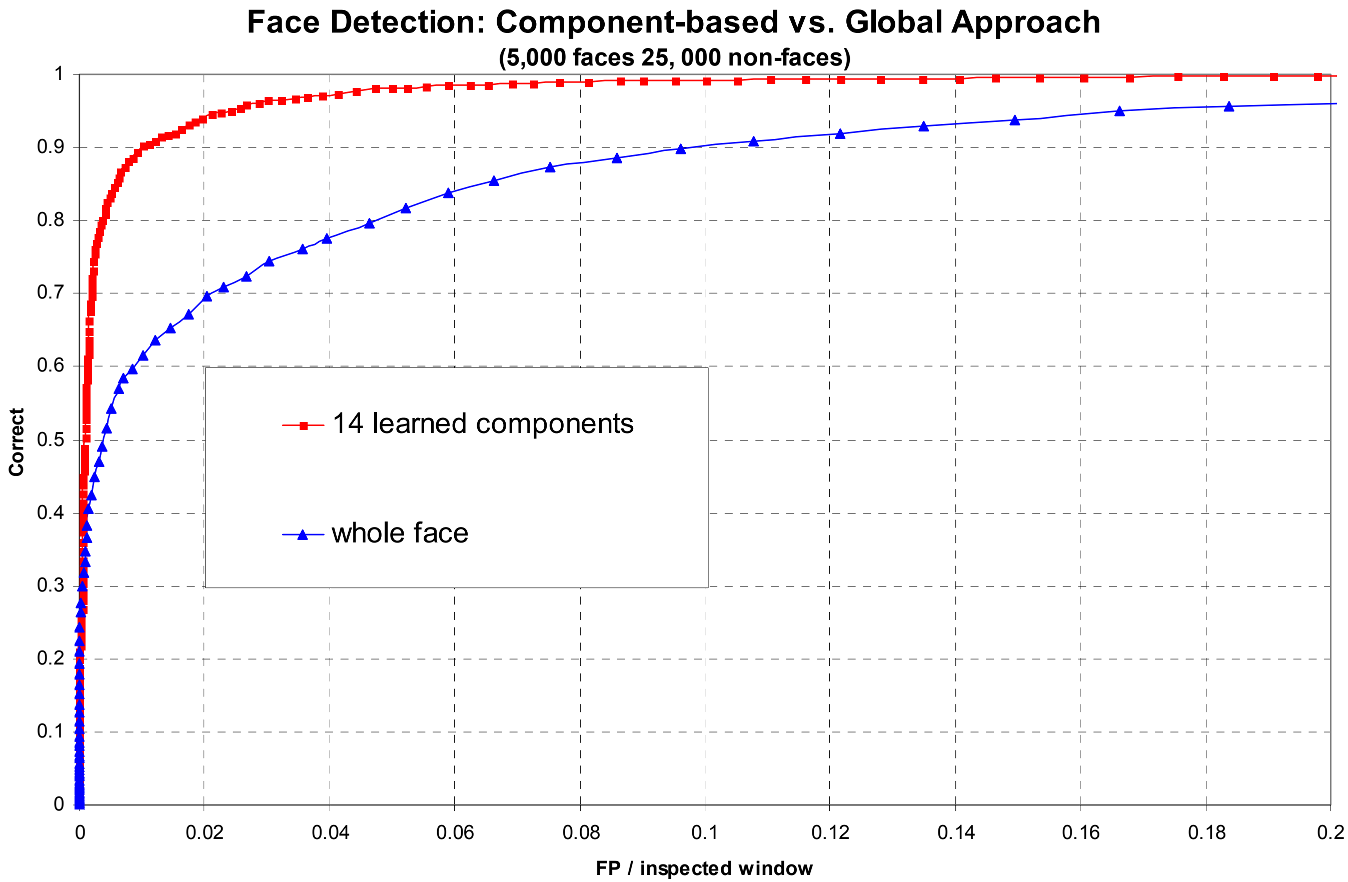
Faces have been manually labeled
(only –45$^o$ to 45$^o$ of rotation)

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



Photographs courtesy of CMU/VASC Image Database at
http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# ROC Component vs. Global




Graph based on work in: Heisele, B., T. Serre, M. Pontil, T. Vetter, and T.Poggio. "Categorization by Learning and Combining Object Parts." In *Advances in Neural Infrmation Processing Systems (NIPS'01)*, Vancouver, Canada.

# Pedestrian Detection

Images removed due to copyright considerations. See Figure 4 in: Papageorgiou, C., and T. Poggio. "A Trainable System for Object Detection." *International Journal of Computer Vision* 38, no. 1 (2000): 15-33.

- Representation: dictionary of Haar wavelets; high dimensional feature space (>1300 features)
- SVM classifier

# Examples

Image removed due to copyright considerations. See Fig. 7P in: Papageorgiou, C., and T. Poggio. "A Trainable System for Object Detection." *International Journal of Computer Vision* 38, no.1 (2000): 15-33.

# Components

- Haar wavelets
- 5 components
- Can deal with partial occlusions

Mohan, A. "Robust Object Detection in Images by Components." Master's Thesis, MIT, 1999.

# Advances on Component-base Face Detection Stan Bileschi

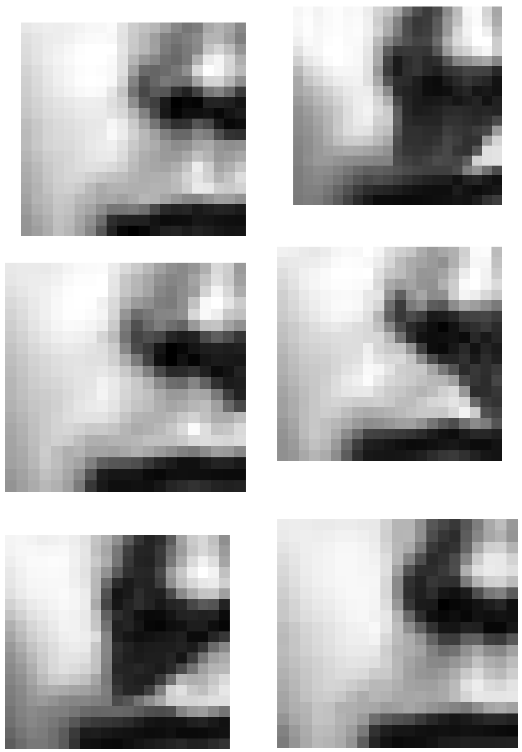
Components are small, and prone to false detection, even within the face.

Photos and figures from: Bileschi, Stan, and Bernd Heisele. "Advances in Component-Based Face Detection." *IEEE International Workshop on Analysis and Modeling of Faces and Gestures* (2003): 149. Courtesy of IEEE, Stan Bileschi, and Bernd Heisele. 

# Training on Faces

## Use the remainder of the face in the negative training set

**Positive**

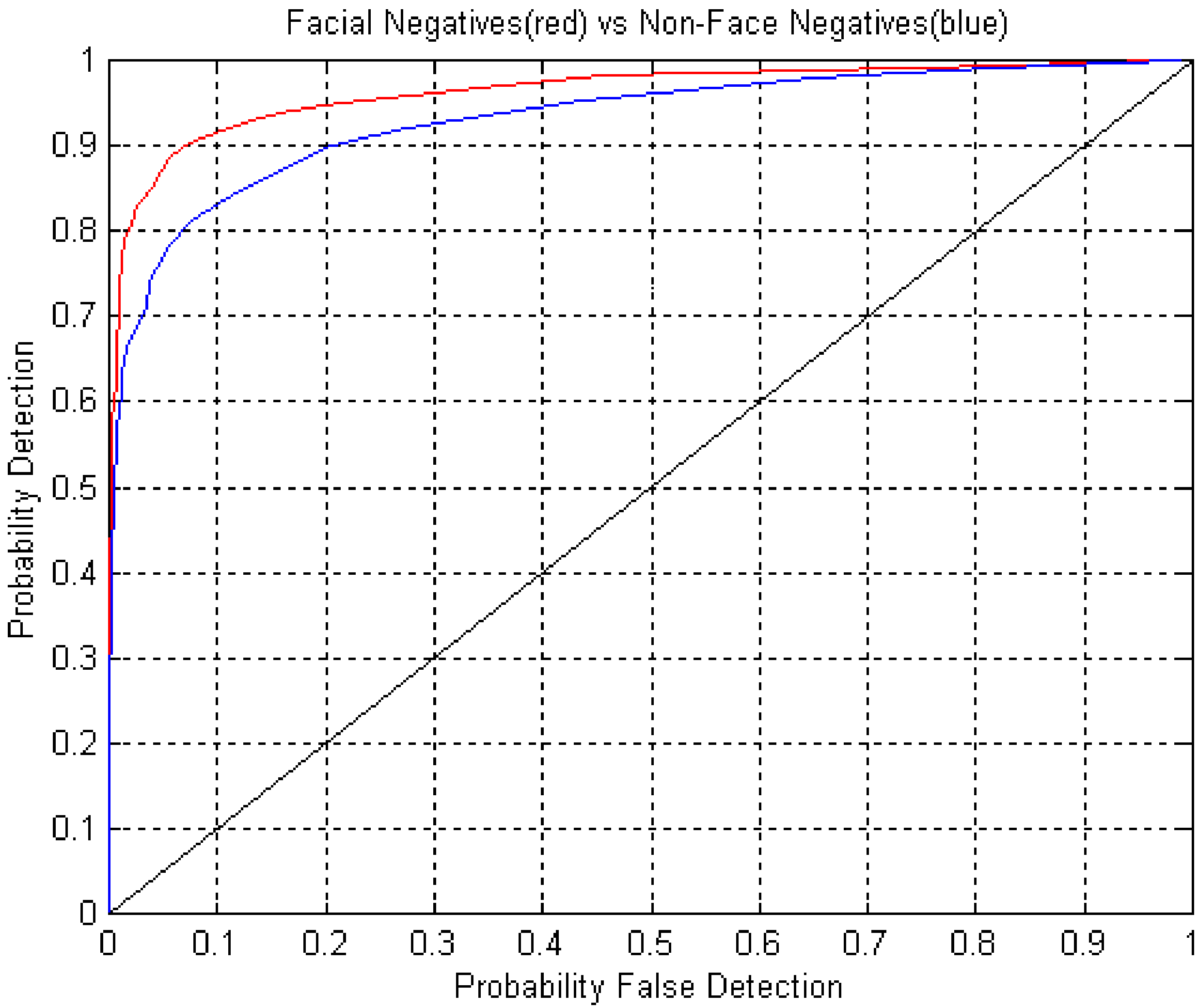


**Negative**

# Training on Faces Only

Red: Trained only with faces.

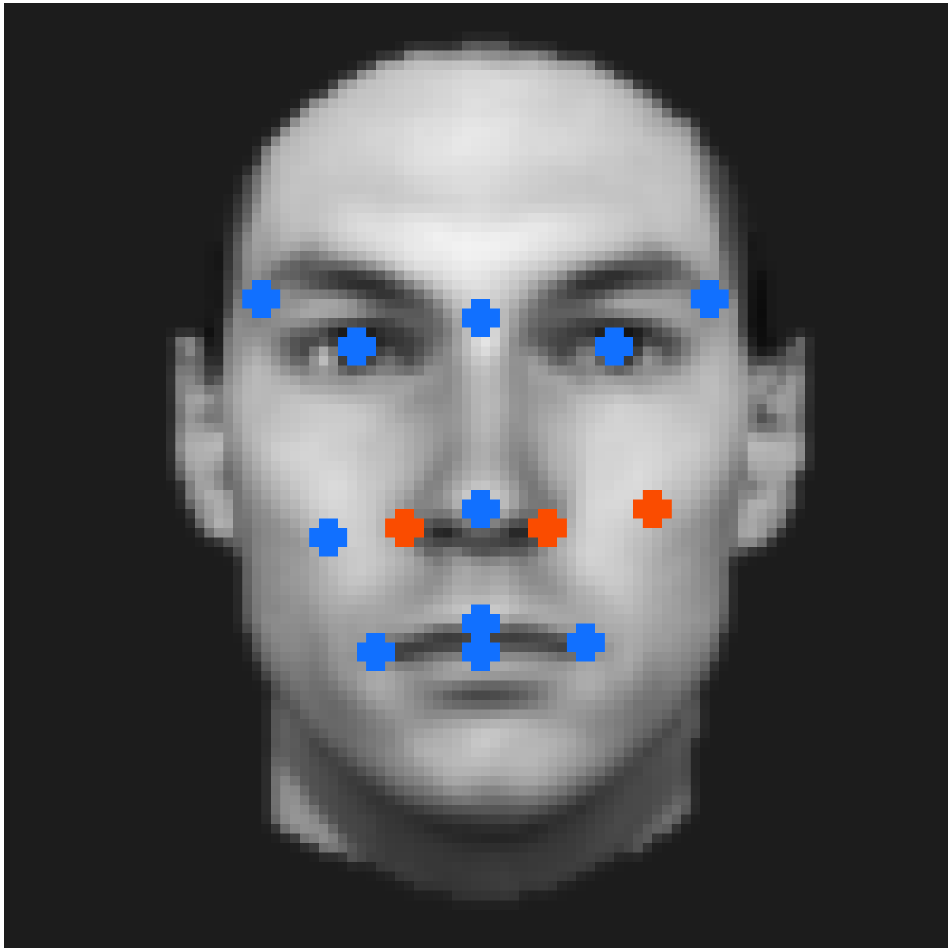
Blue: Trained on faces and non-faces.



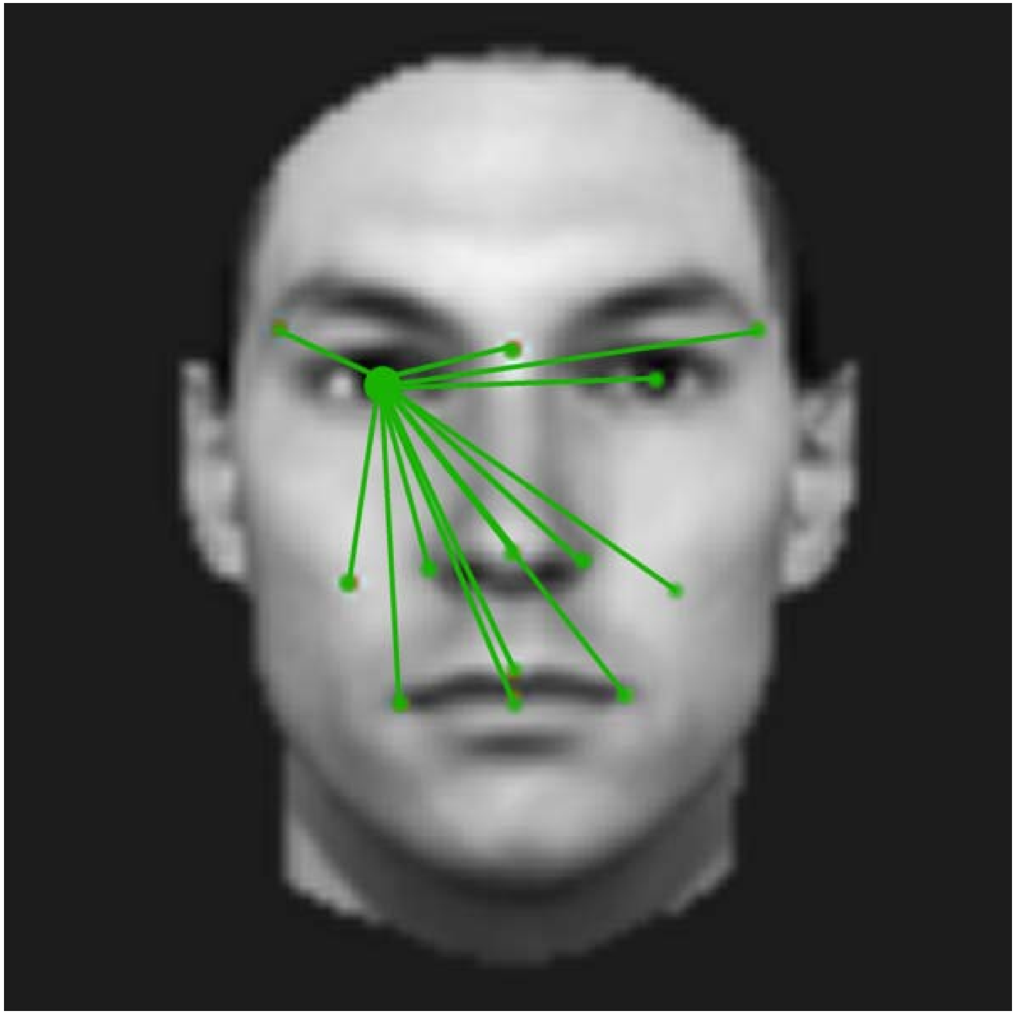

Photos and figures from: Bileschi, Stan, and Bernd Heisele. "Advances in Component-Based Face Detection." *IEEE International Workshop on Analysis and Modeling of Faces and Gestures* (2003): 149. Courtesy of IEEE, Stan Bileschi, and Bernd Heisele. 

# Errors

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

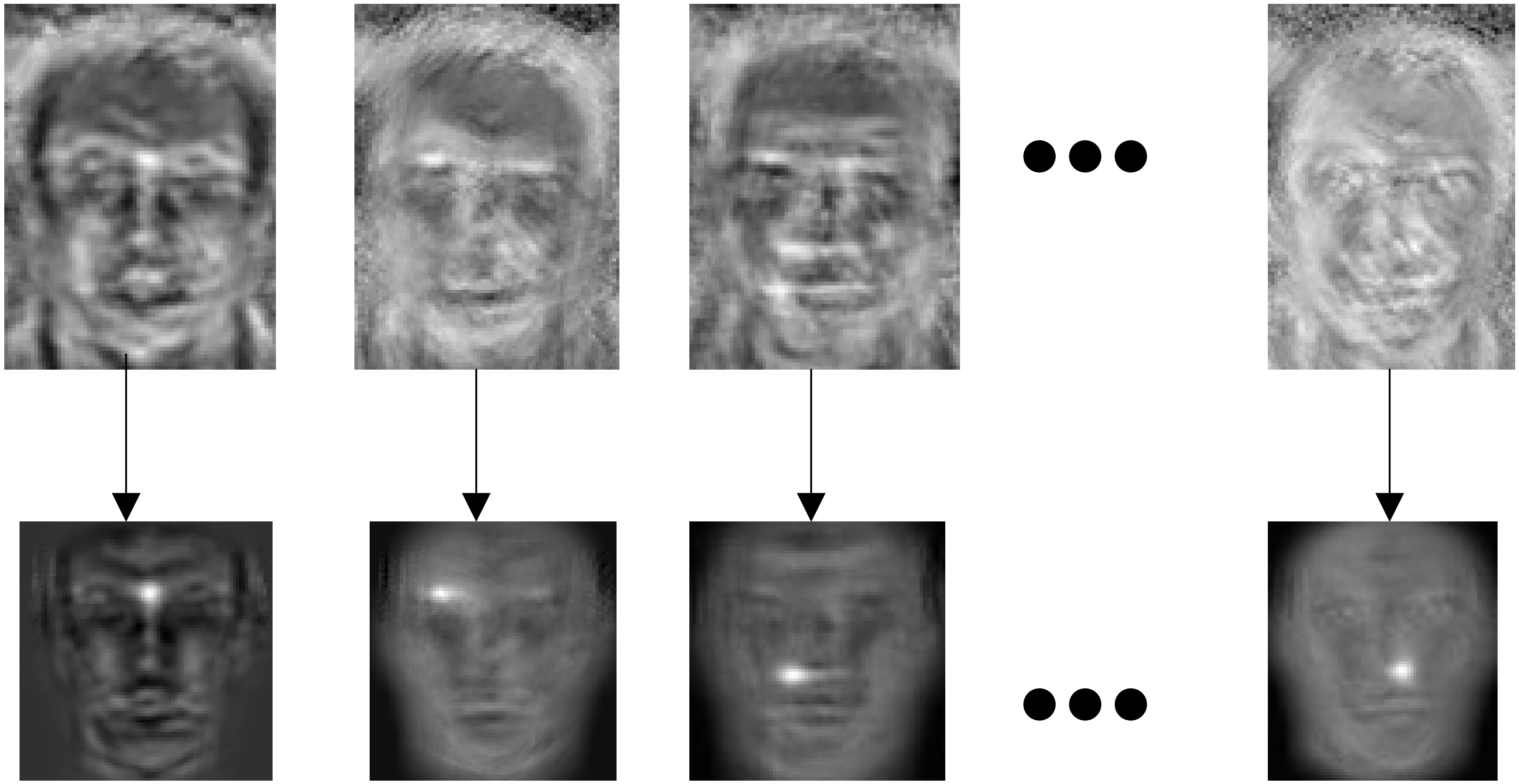


# Using Models of Pair-wise Positions



Photos and figures from: Bileschi, Stan, and Bernd Heisele. "Advances in Component-Based Face Detection." *IEEE International Workshop on Analysis and Modeling of Faces and Gestures* (2003): 149. Courtesy of IEEE, Stan Bileschi, and Bernd Heisele. Copyright 2003 IEEE. Used with Permission.

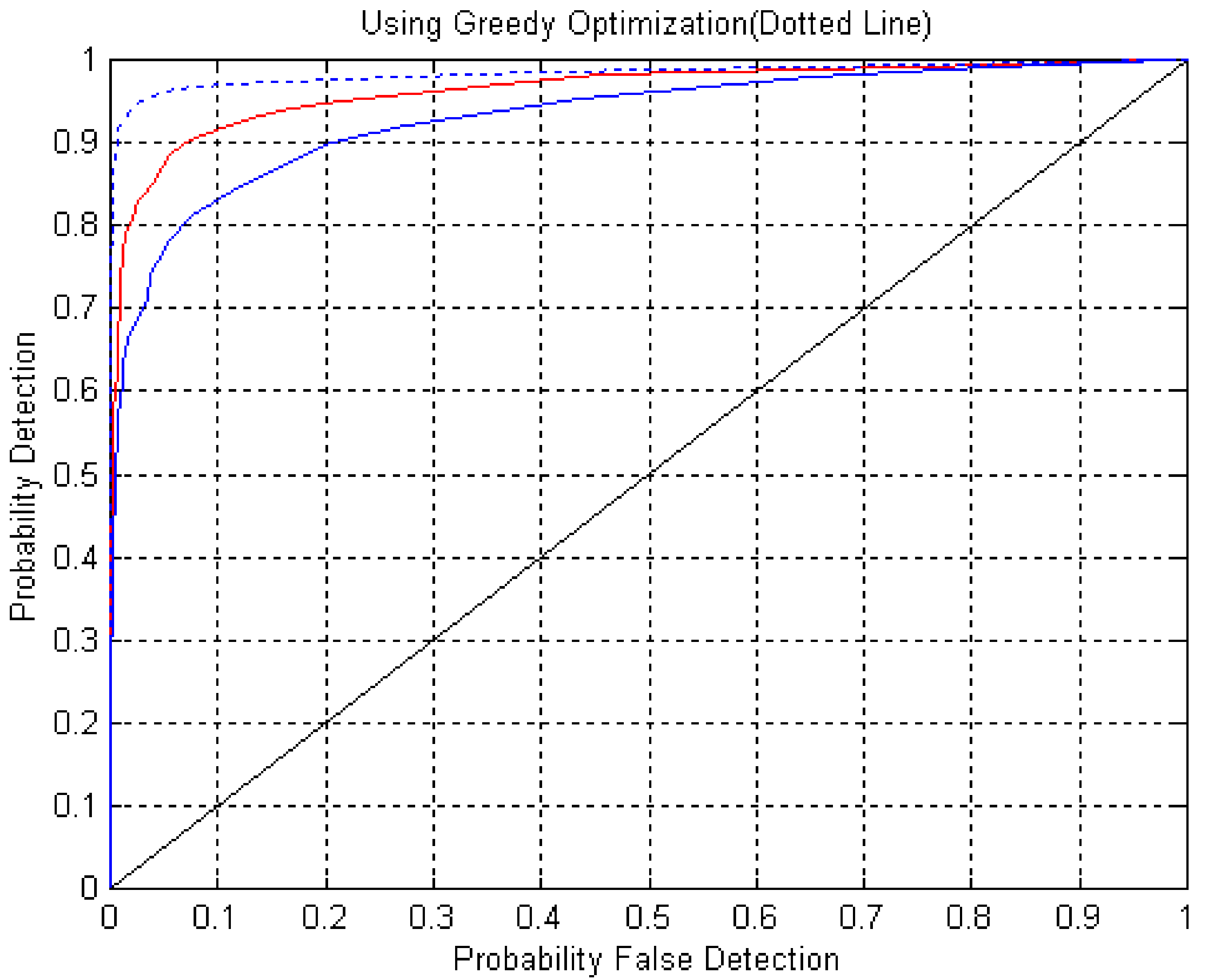
# Pair-wise Biasing Leads to Tightened Result Images



Photos and figures from: Bileschi, Stan, and Bernd Heisele. "Advances in Component-Based Face Detection." *IEEE International Workshop on Analysis and Modeling of Faces and Gestures* (2003): 149. Courtesy of IEEE, Stan Bileschi, and Bernd Heisele. Copyright 2003 IEEE. Used with Permission.

Courtesy of Stan Bileschi. Used with permission.

# Pair-wise Biasing




Photos and figures from: Bileschi, Stan, and Bernd Heisele. "Advances in Component-Based Face Detection." *IEEE International Workshop on Analysis and Modeling of Faces and Gestures* (2003): 149. Courtesy of IEEE, Stan Bileschi, and Bernd Heisele. Copyright 2003 IEEE. Used with Permission.

Courtesy of Stan Bileschi. Used with permission.

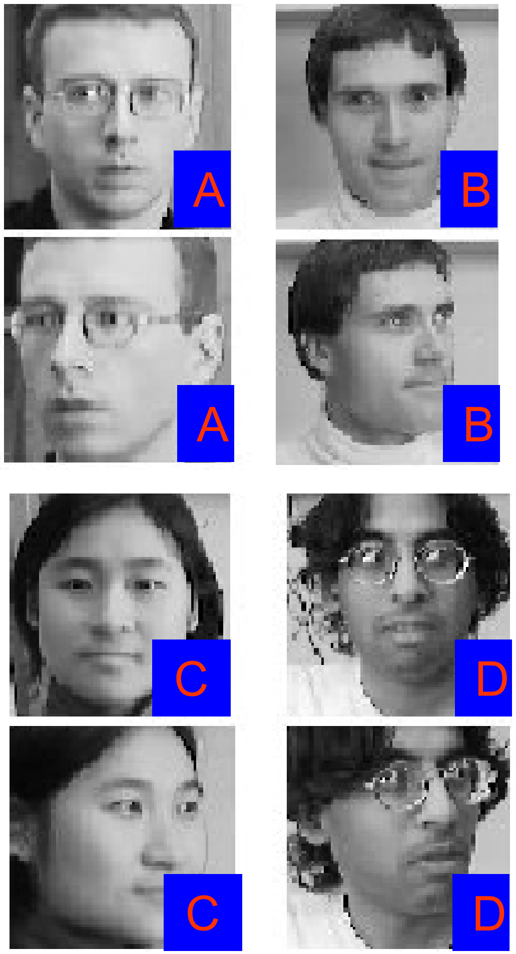
# Application: Eye Detection

Photos and figures from: Bileschi, Stan, and Bernd Heisele. "Advances in Component-Based Face Detection." *IEEE International Workshop on Analysis and Modeling of Faces and Gestures* (2003): 149. Courtesy of IEEE, Stan Bileschi, and Bernd Heisele. 

# Recognition

**Task:**

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




Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Recognition—Problems

1. Multi-class problem

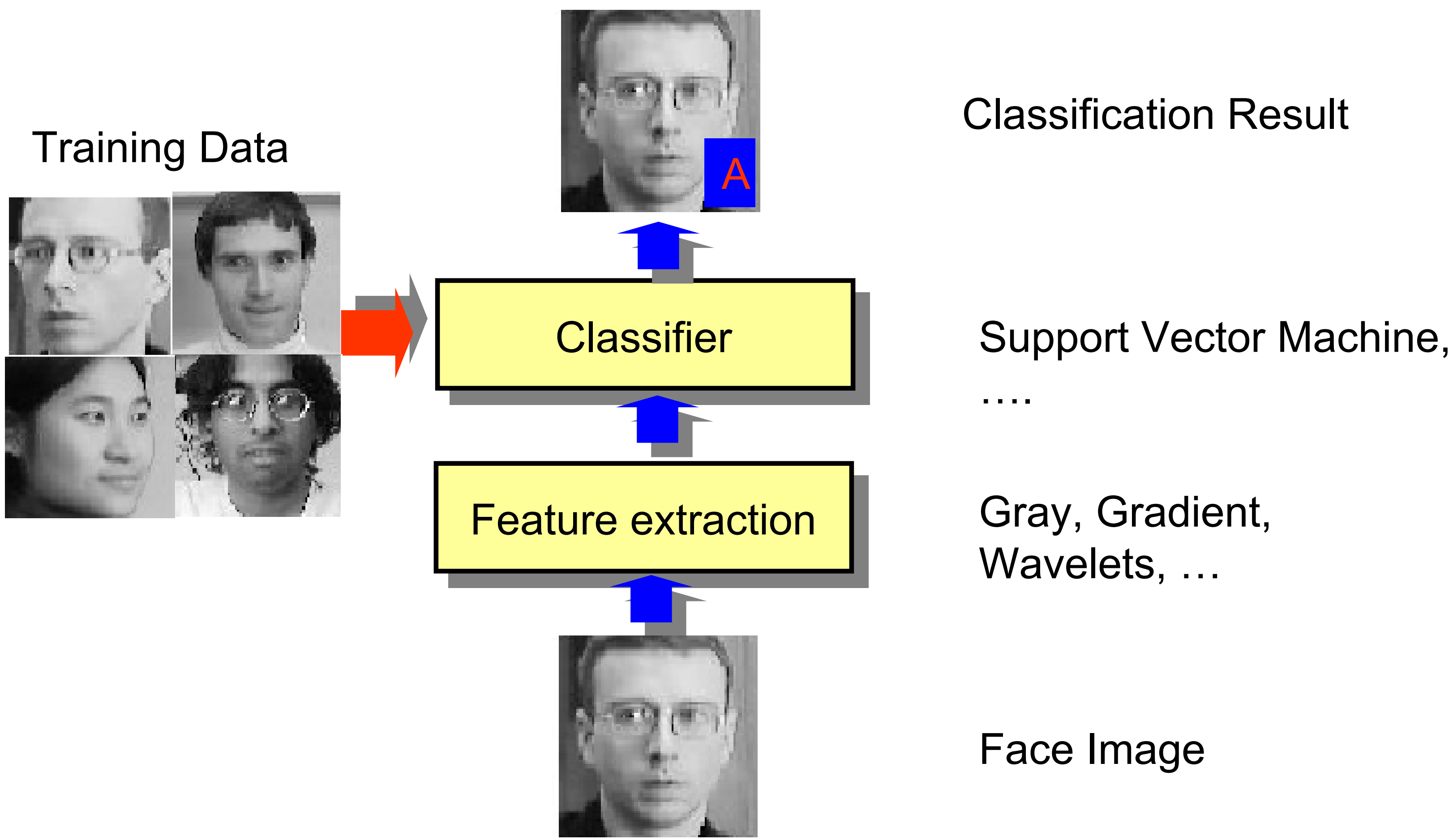
2. Classifier must distinguish between exemplars that might look very similar.

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

# Problems in Face Recognition

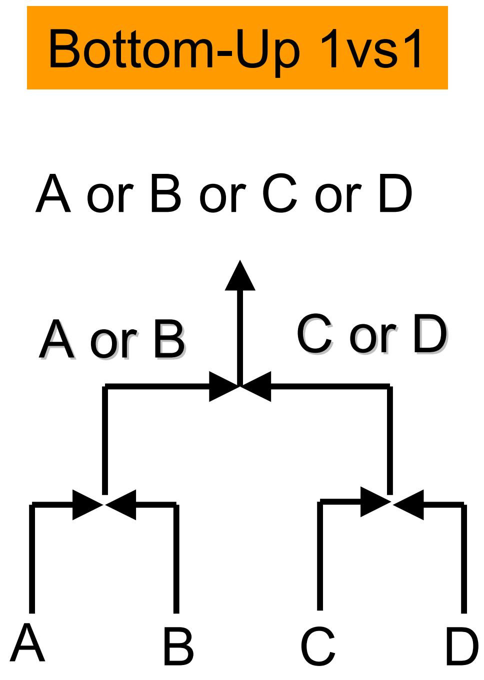
Images removed due to copyright considerations. See: Huang, Jennifer.
"Face Recognition Using Component-Based SVM Classification and Morphable Models."
SVM 2002, LNCS 2388, 2002, pp. 334-341.

# System Architecture

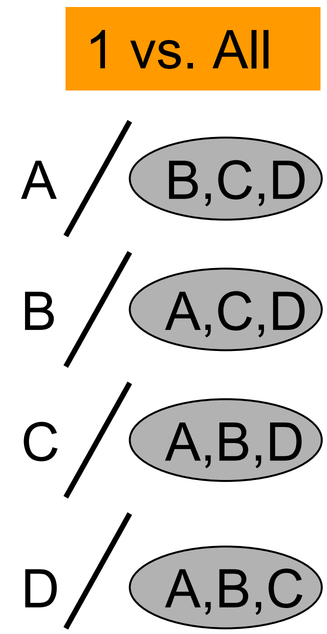



Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/

# Multi-class Classification with SVM



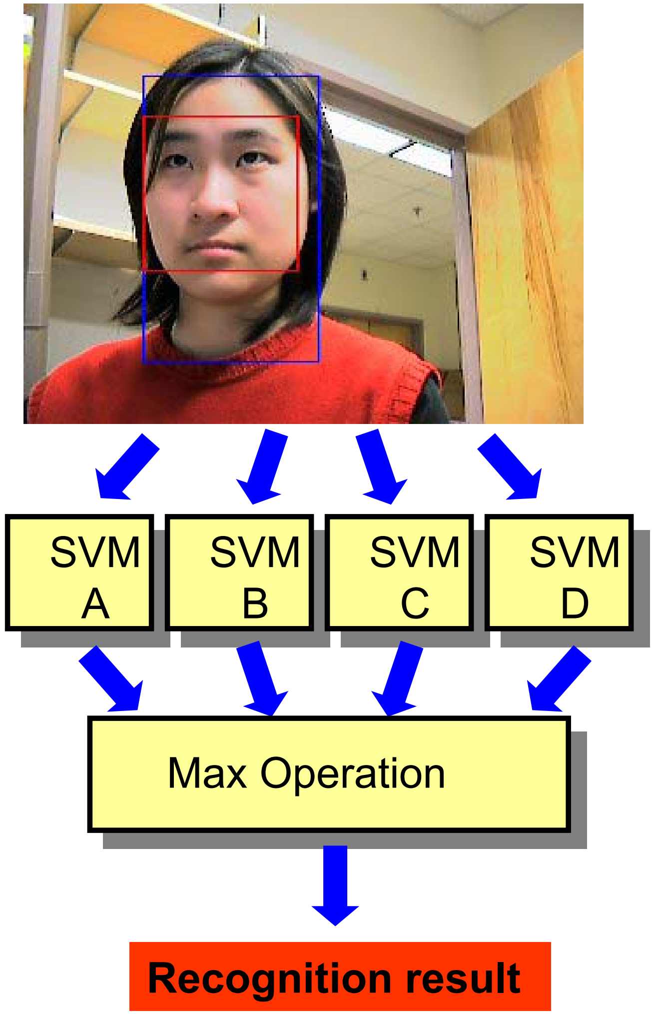

Training: $L$ $(L$-1) / 2
Classification : $L$-1




Training: $L$
Classification : $L$

# Global Approach

Photographs courtesy of CMU/VASC Image Database at
http://vasc.ri.cmu.edu/idb/html/face/frontal_images/



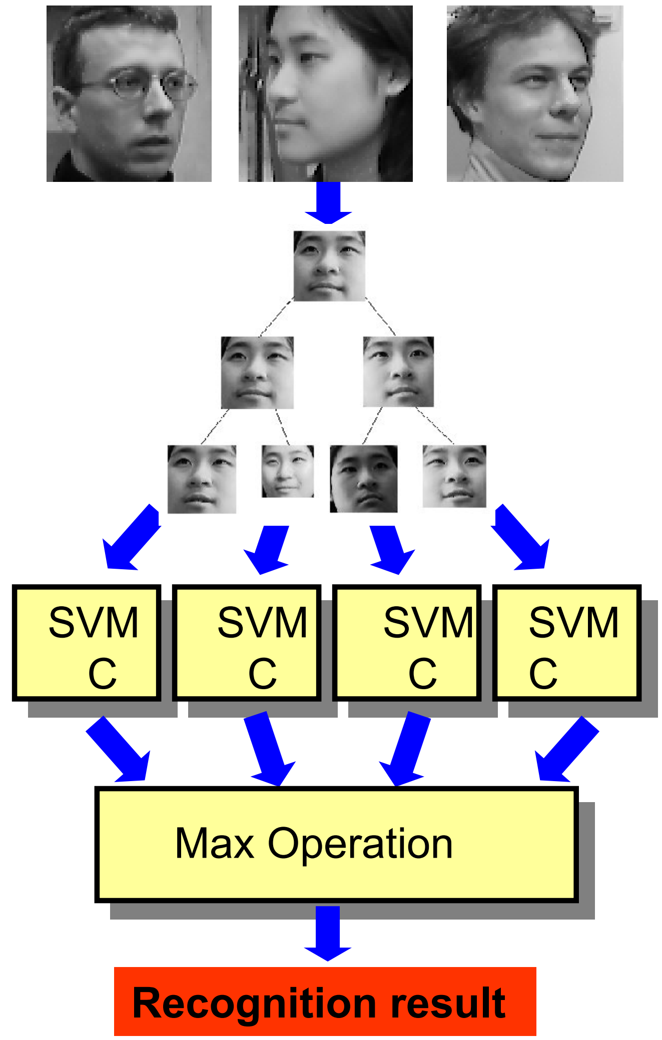

Detect and extract face

Feed gray values into *N* SVMs

Classify based on maximum output

# Global Approach with Clustering

Photographs courtesy of CMU/VASC Image Database at http://vasc.ri.cmu.edu/idb/html/face/frontal_images/




Partition training images of each person into viewpoint-specific
clusters

Train a linear SVM on each cluster

Take maximum over all SVM outputs

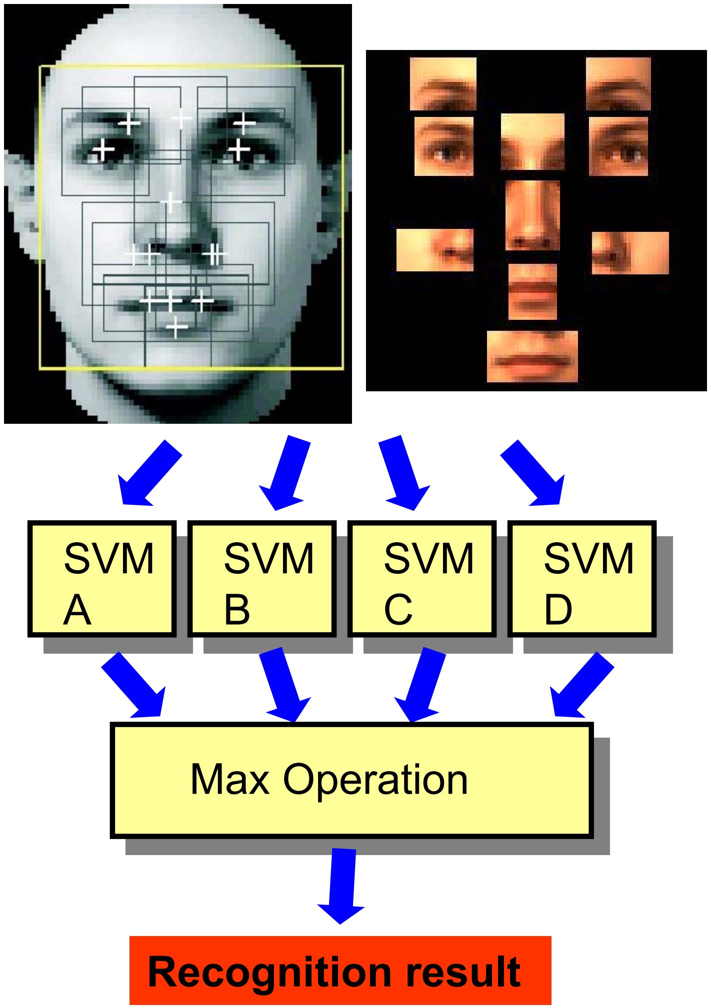
# Component-based Approach

Detect and extract components

Feed gray values of
components to *N* SVMs

Take max. over all SVM outputs

Photographs courtesy of CMU/VASC Image Database at
http://vasc.ri.cmu.edu/idb/html/face/frontal_images/



# Why Components for Recognition?

Images removed due to copyright considerations. See: Huang, Jennifer. "Face Recognition Using Component-Based SVM Classification and Morphable Models." SVM 2002, LNCS 2388, 2002, pp. 334-341.



Component-based Face Recognition with 3D Morphable Models

# More ROC Curves

Image removed due to copyright considerations. See: Heisele, B., P. Ho, and T.Poggio. "Face Recognition with Support Vector Machines: Global Versus Component-based Approach." *International Conference on Computer Vision (ICCV'01).* Vol. 2. Vancouver, Canada, 2001, pp. 688-694.

# Morphable Models for Face Recognition, Jennifer Huang

Images removed due to copyright considerations.
See: Huang, Jennifer. "Face Recognition Using Component-Based SVM Classification and Morphable Models." SVM 2002, LNCS 2388, 2002, pp. 334-341.

Images removed due to copyright considerations.
See: Blanz, V., and Vetter, T. "A Morphable Model for the Synthesis of 3D Faces." SIGGRAPH'99 Conference Proceedings, pp. 187-194.

# Morphable Model

Images removed due to copyright considerations. See: Huang, Jennifer. "Face Recognition Using Component-Based SVM Classification and Morphable Models." SVM 2002, LNCS 2388, 2002, pp. 334-341.

# Some Training Images

Images removed due to copyright considerations. See: Huang, Jennifer.
"Face Recognition Using Component-Based SVM Classification and Morphable Models."
SVM 2002, LNCS 2388, 2002, pp. 334-341.

# Preliminary Results on Synthetic Images

Image removed due to copyright considerations. See: Huang, Jennifer. "Face Recognition Using Component-Based SVM Classification and Morphable Models." SVM 2002, LNCS 2388, 2002, pp. 334-341.

# Current Work – Testing on Real Images

Problems Encountered:

**Detection**

Inaccurate Component detection

**Recognition**

Accuracy of 3D models

Choice of Illumination and Pose

*Jennifer Huang*

# 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