Detecting, Localizing and Classifying Facial Traits from Arbitrary Viewpoints

Matthew Toews and Tal Arbel
Centre for Intelligent Machines
McGill University, Montreal, Canada

McGill Centre for Intelligent Machines

ICCV AMFG
2007

Overview

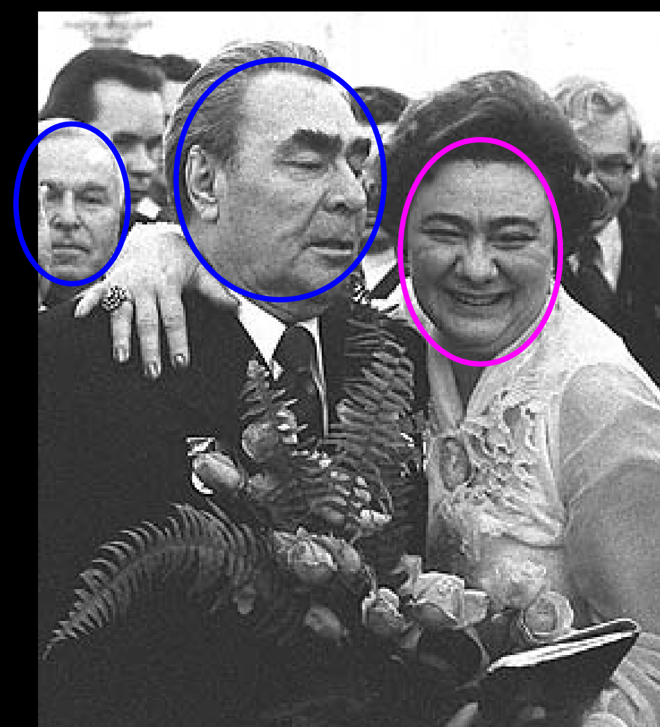
- Problem
- Approach
  - Invariant image features
  - Viewpoint-invariant face modeling
  - Classification
- Experimentation

Facial Traits?

- Gender, age, expression, …

Problem

- In a cluttered scene, detect, localize and classify faces in terms of traits
  - Facial traits: *gender, age, expression, …*

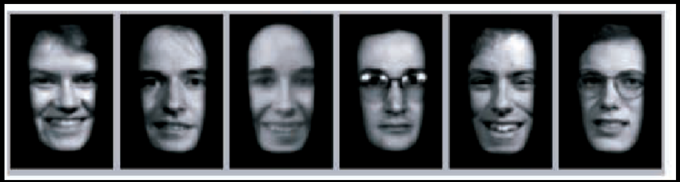


Facial Trait Classification

- Image features
  - Global: principle/independent components
  - Local: Haar wavelets, local regions
- Classification
  - Support vector machines, neural networks, decision trees

Classification: Gender

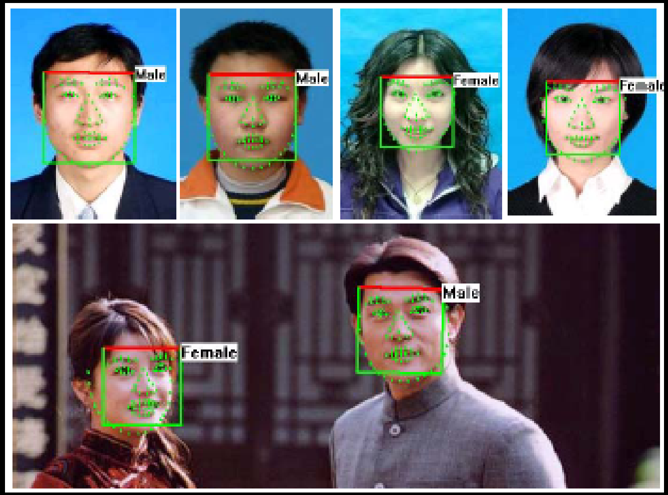
Shakhnarovich et al. 2002



Moghaddam 2002

Classification error: 4-10%

Jain et al. 2004

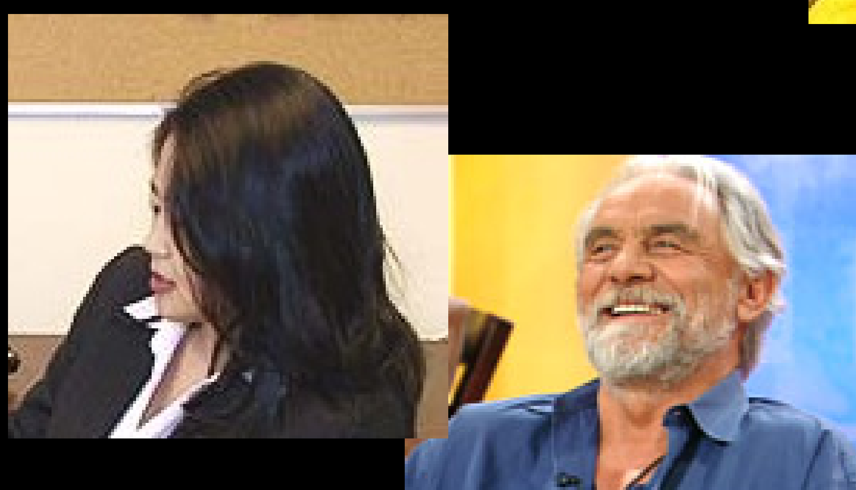


Yang et al. 2006

Baluja et al. 2007

Classification Challenges

- Detecting faces & features required for classification
  - Arbitrary viewpoints?
  - Occlusion?



Detection/Localization Challenges

- Nuisance parameters
  - Illumination changes
  - In-plane geometrical deformations
  - Partial occlusion
  - Intra-class variation
  - In-depth geometrical deformations (viewpoint variation)
- Computational efficiency

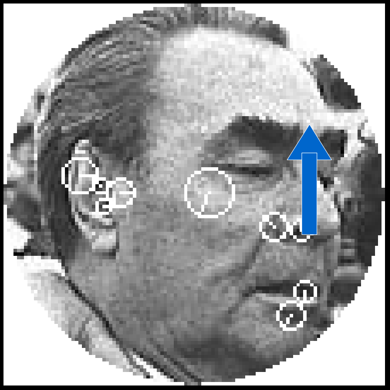
Approach

In an arbitrary cluttered scene,

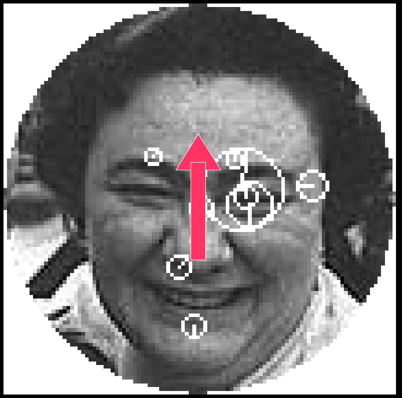
extract features,

detect, localize

and classify faces.



Male



Female

Approach

- Local invariant image features
- Viewpoint-invariant face modeling
- Classification

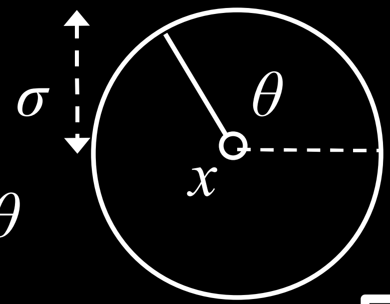
Approach

- Local invariant image features
  - Identify the same image features despite changes in illumination and geometry.
- Viewpoint-invariant face modeling
- Classification

Local Invariant Image Features

Geometry

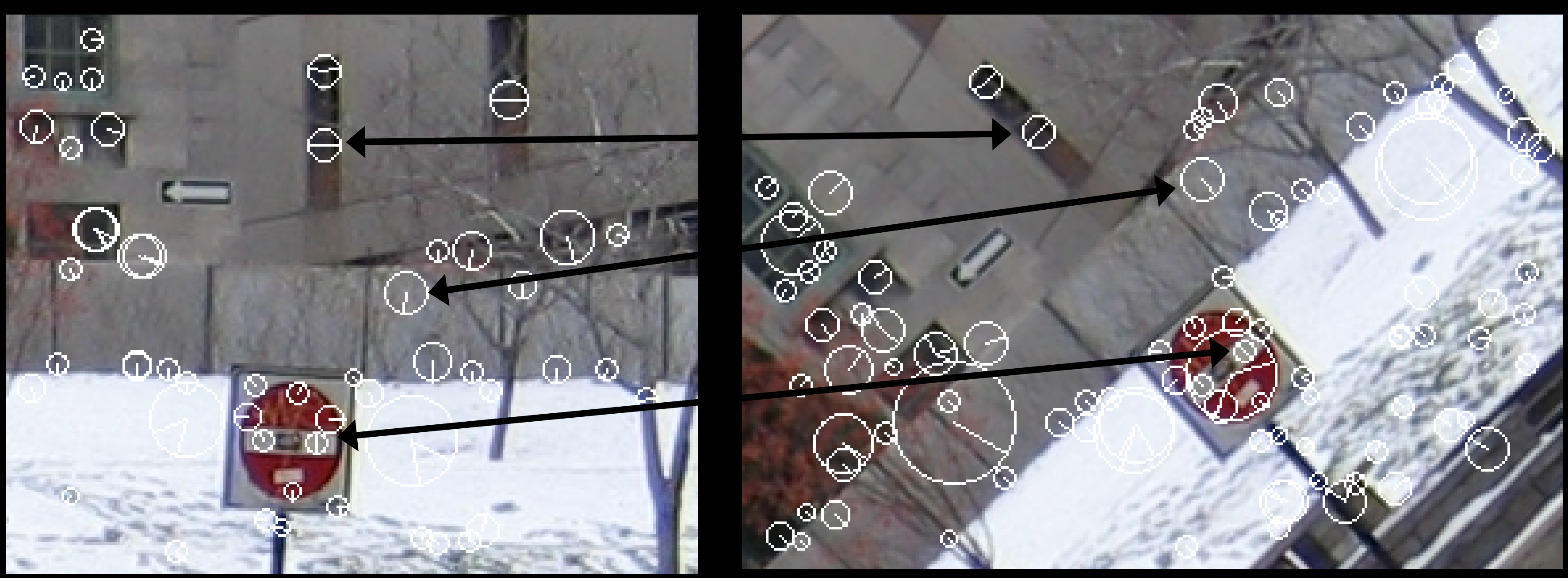
- Location x
- Orientation θ
- Scale σ



Appearance

- Image intensity information
- I.e. Pixels, edges

SIFT: scale-invariant feature transform, Lowe 2004



Local Invariant Image Features

- Address:
  - Nuisance parameters
    - Illumination variation
    - In-plane geometrical variation
    - Occlusion
  - Efficiency
- Do not address:
  - Intra-class variation
  - Viewpoint variation

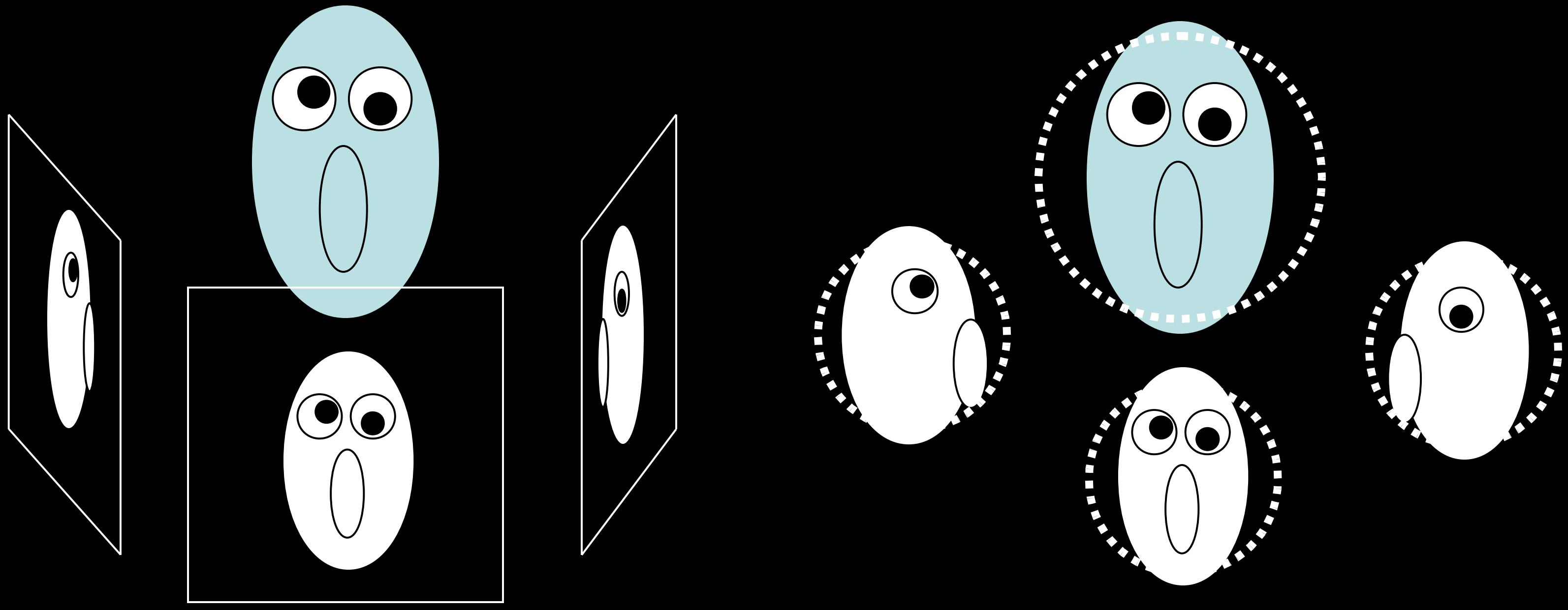
Approach

- Local invariant image features
- Viewpoint-invariant face modeling
  - Identify the same invariant features in arbitrary viewpoints and faces.
- Classification

Viewpoint-Invariant?

Multi-view model

Viewpoint-invariant model



Koenderink et al., 1979

Multi-view Viola-Jones, 2003

Beiderman, 1987

Toews & Arbel, ICPR 2006

Viewpoint-invariance: Advantage

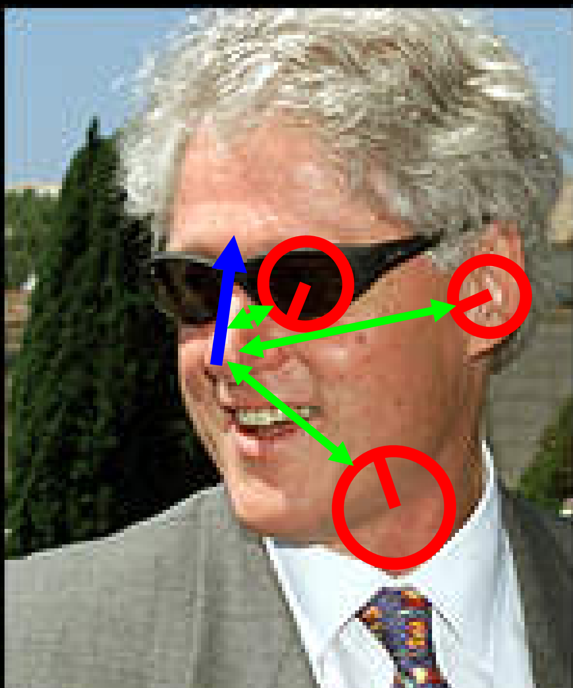
- Simplicity: no viewpoint variable

The Object Class Invariant (OCI)

- A geometrical reference frame that is:
  1) Uniquely defined for each pattern/object class instance.
  2) Invariant to the geometrical transform arising from the imaging process *(perspective projection).*


Toews & Arbel, ICPR 2006

Object Class Invariant Modeling



Object Class Invariant

$o:\{o^b, o^g\}$

Occurrence: binary presence or absence.

Geometry: location, scale, orientation.

Scale-invariant Feature

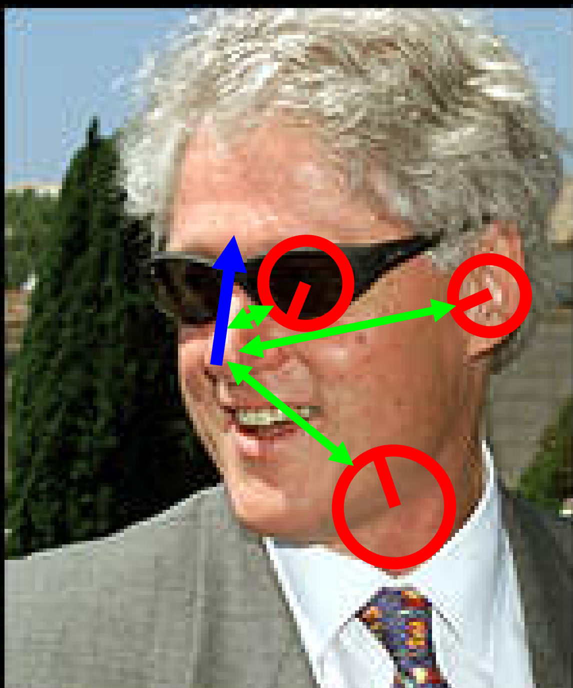
$m_i:\{m_i^{\,b}, m_i^{\,g}, m_i^{\,a}\}$

Transform relating feature and OCI geometries:

$t_i : m_i^{\,g} \rightarrow o^g, o^g = t_i(m_i)$

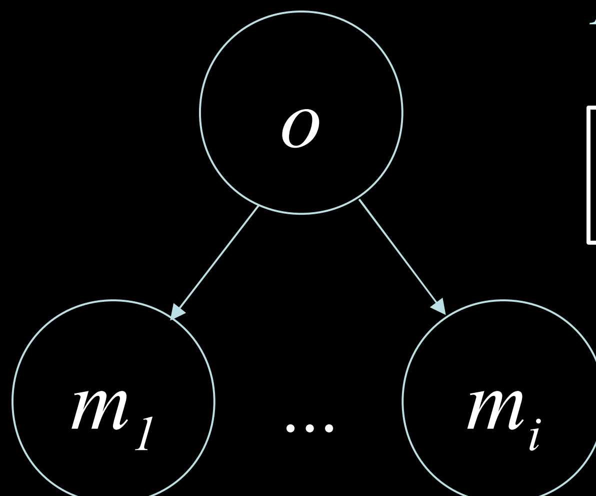
Appearance: derivative histograms.
Note: the OCI is unobservable, and has no appearance!

OCI Model



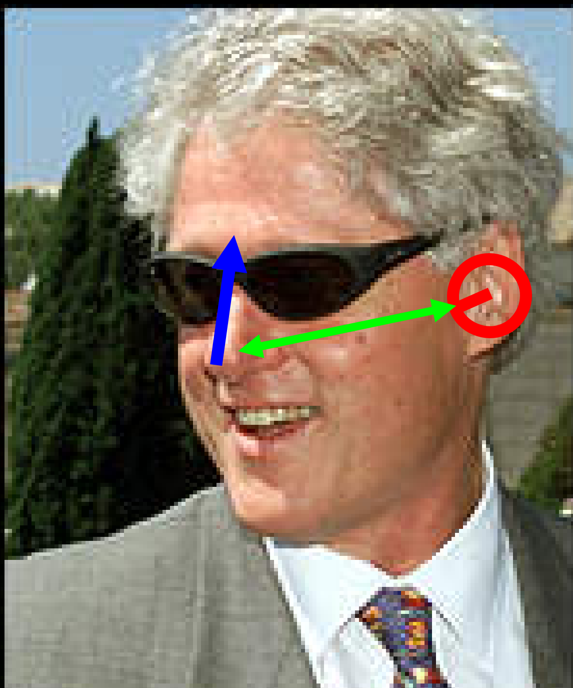
$$p(o \mid \{m_i\}) = \frac{p(o)\,p(\{m_i\} \mid o)}{p(\{m_i\})}$$

$$= \frac{p(o)\prod_i p(m_i \mid o)}{p(\{m_i\})}$$

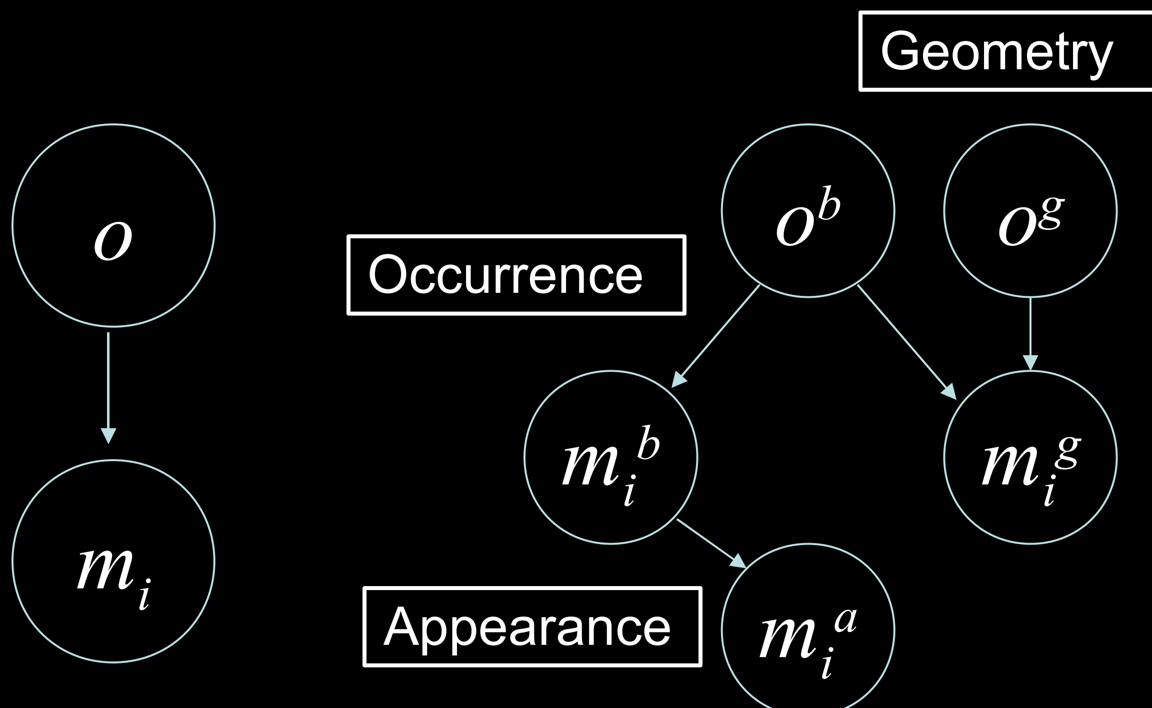


Conditional Feature Independence

OCI Model

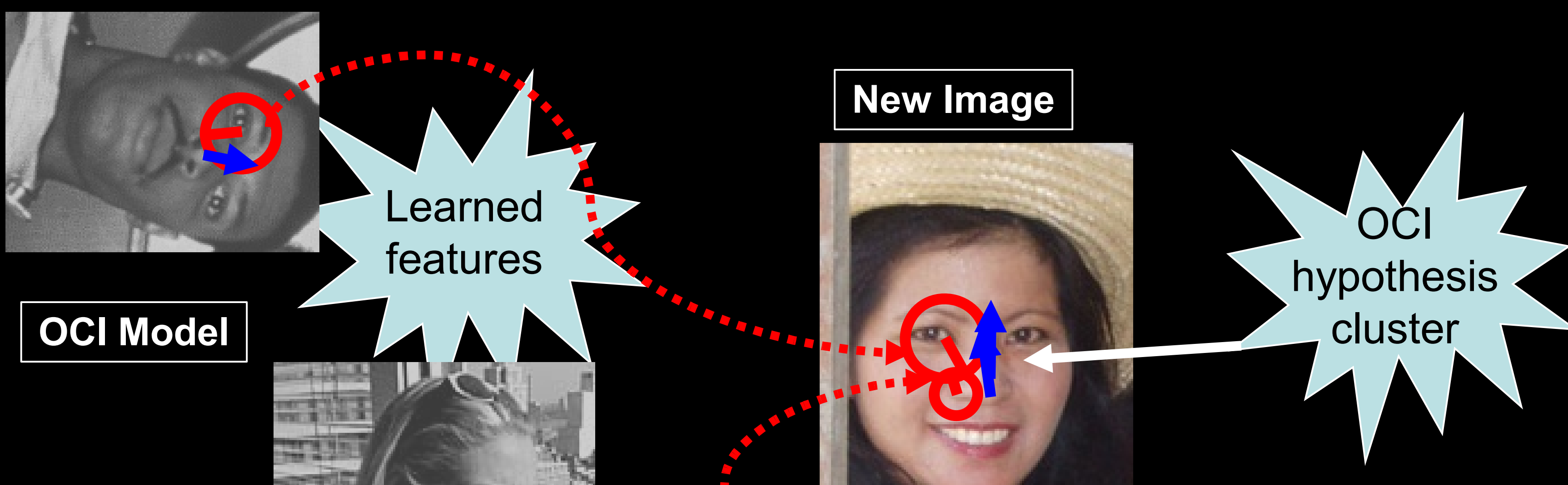


$$p(m_i \mid o) = p(m_i^a \mid m_i^b)\, p(m_i^b \mid o^b)\, p(m_i^g \mid o^b, o^g)$$



OCI Model Fitting

- Identifying OCI instance in new image
  - Probabilistic voting
  - Robust hypothesis clustering



$$o = argmax\left\{ \frac{p(o \mid \{m_i\})}{p(\bar{o} \mid \{m_i\})} \right\}$$

OCI Model Learning

- Labeled OCIs
- Automatic features
- Robust feature clustering
  - Geometric agreement wrt OCI
  - Appearance agreement

Approach

- Local invariant image features
- Viewpoint-invariant face modeling
- Classification
  - Learn a classifier based on invariant image features identified in different viewpoints and faces.

Classification

Female feature

Male feature

Gender neutral feature

Classification

- Bayesian classifier

Visual Trait

$$c : \{c_1, \ldots, c_K\}$$

Feature Occurrence

$$f_i = m_i^{b=1}$$

$$\psi(c) = \frac{p(c \mid \{f_i\})}{p(\bar{c} \mid \{f_i\})} = \frac{p(c)\,p(\{f_i\} \mid c)}{p(\bar{c})\,p(\{f_i\} \mid \bar{c})} = \frac{p(c)}{p(\bar{c})} \prod_i^M \frac{p(f_i \mid c)}{p(f_i \mid \bar{c})}$$

Conditional Feature Independence

Classification

Offset

Likelihood ratio of feature given trait presence vs. absence.

$$\log\psi\ (c) = \log\frac{p(c)}{p(\bar{c})} + \sum_i^M \log\frac{p(f_i \mid c)}{p(f_i \mid \bar{c})}.$$

Weighted feature/trait co-occurrence count

$$\log\frac{p(c)}{p(\bar{c})} = E\left[\sum_i^M \log\frac{p(f_i \mid c)}{p(f_i \mid \bar{c})}\right]$$

Expected log likelihood sum given M detected features.

$$p(f_i \mid c_j) = \frac{k_{ij}}{p(c_j)} + d_{ij}.$$

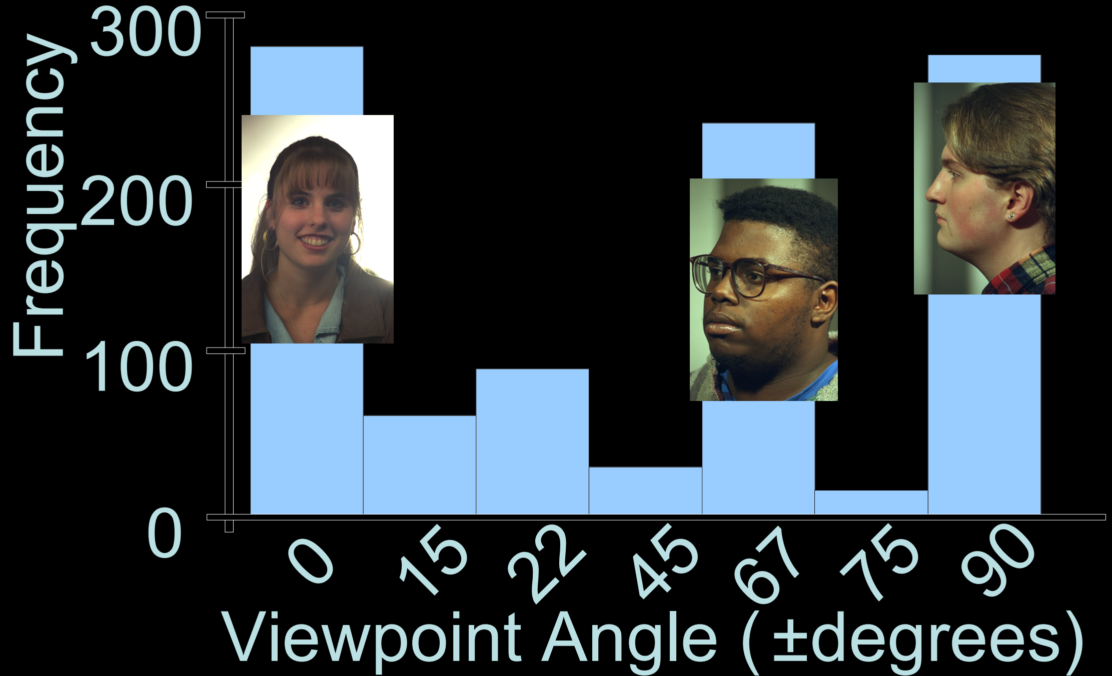
Bayesian estimation:
Binomial likelihood + Dirichlet prior

Experimentation

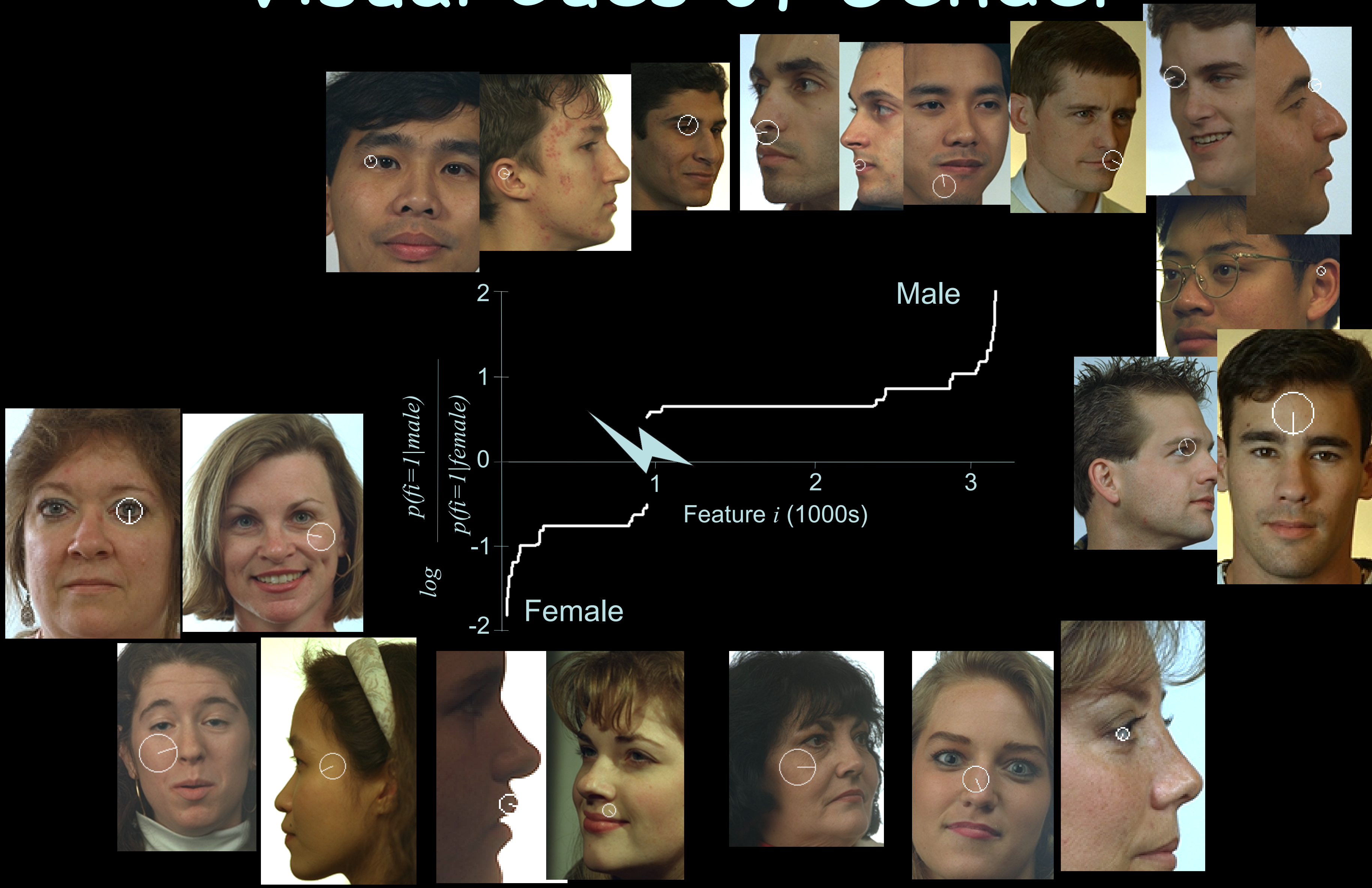
- Task:
  - Detect, localize and classify face gender.
- Training:
  - Learn face model & gender classifier
  - Training data labels: OCI and gender
- Testing:
  - Detect, localize and classify new faces
  - Fully automatic

Experimentation

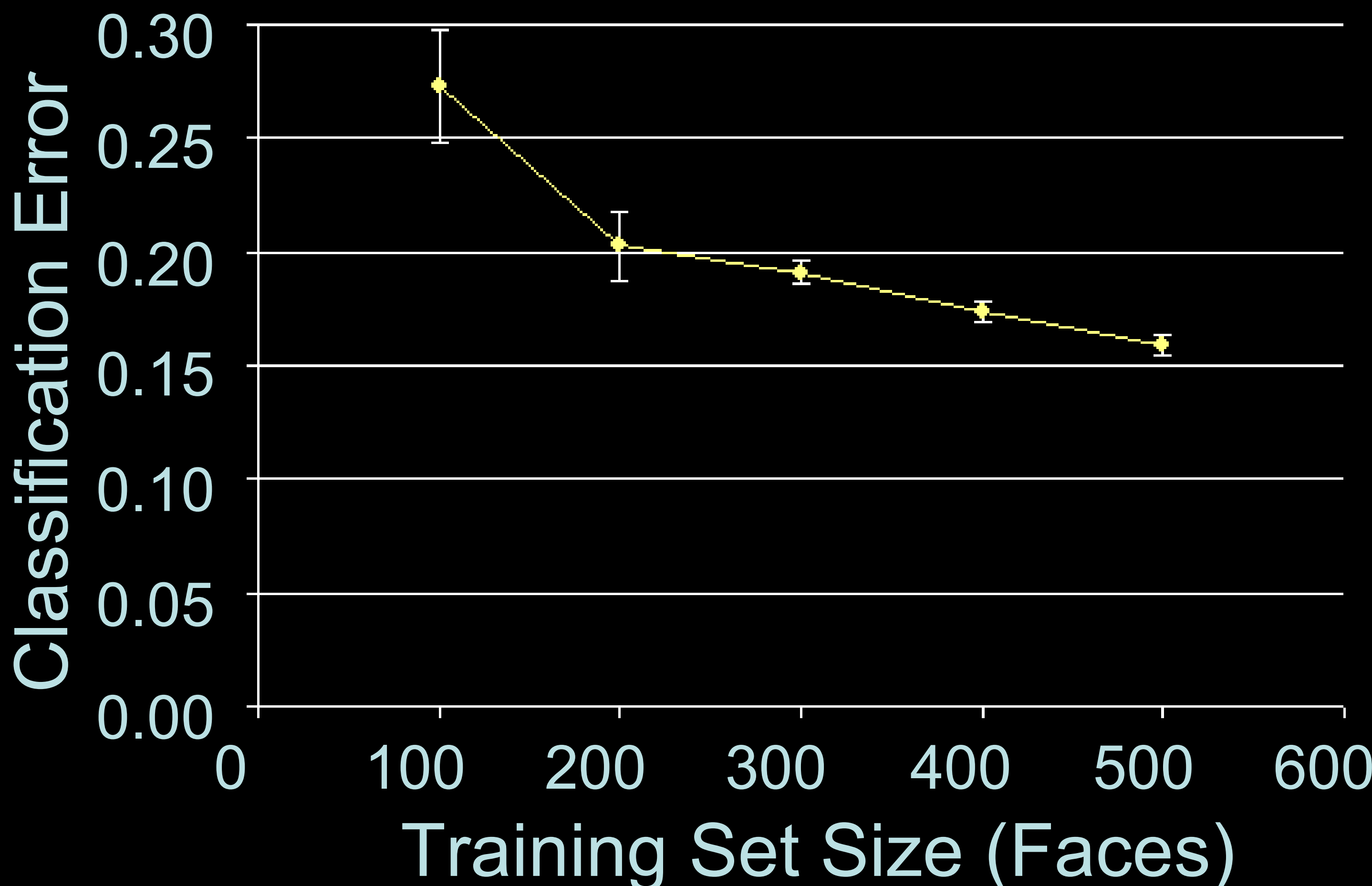
- Database: Color FERET
  - 994 unique subjects, random viewpoint
  - Male:Female = 3:2



Visual Cues of Gender



Classification Error



Classification: Correct

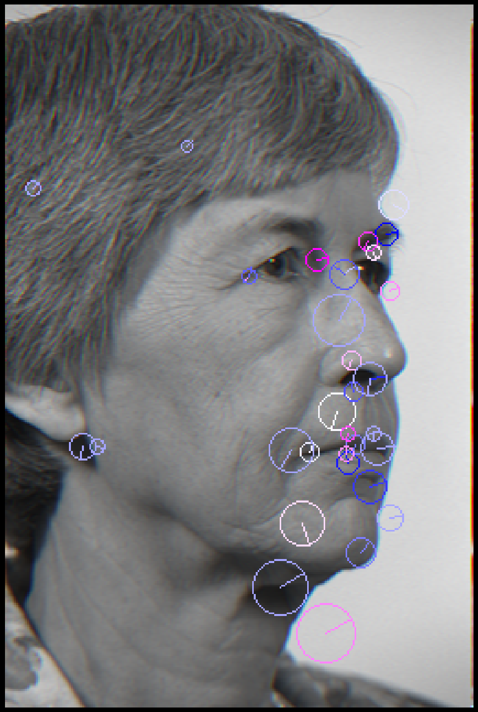
Female

Male

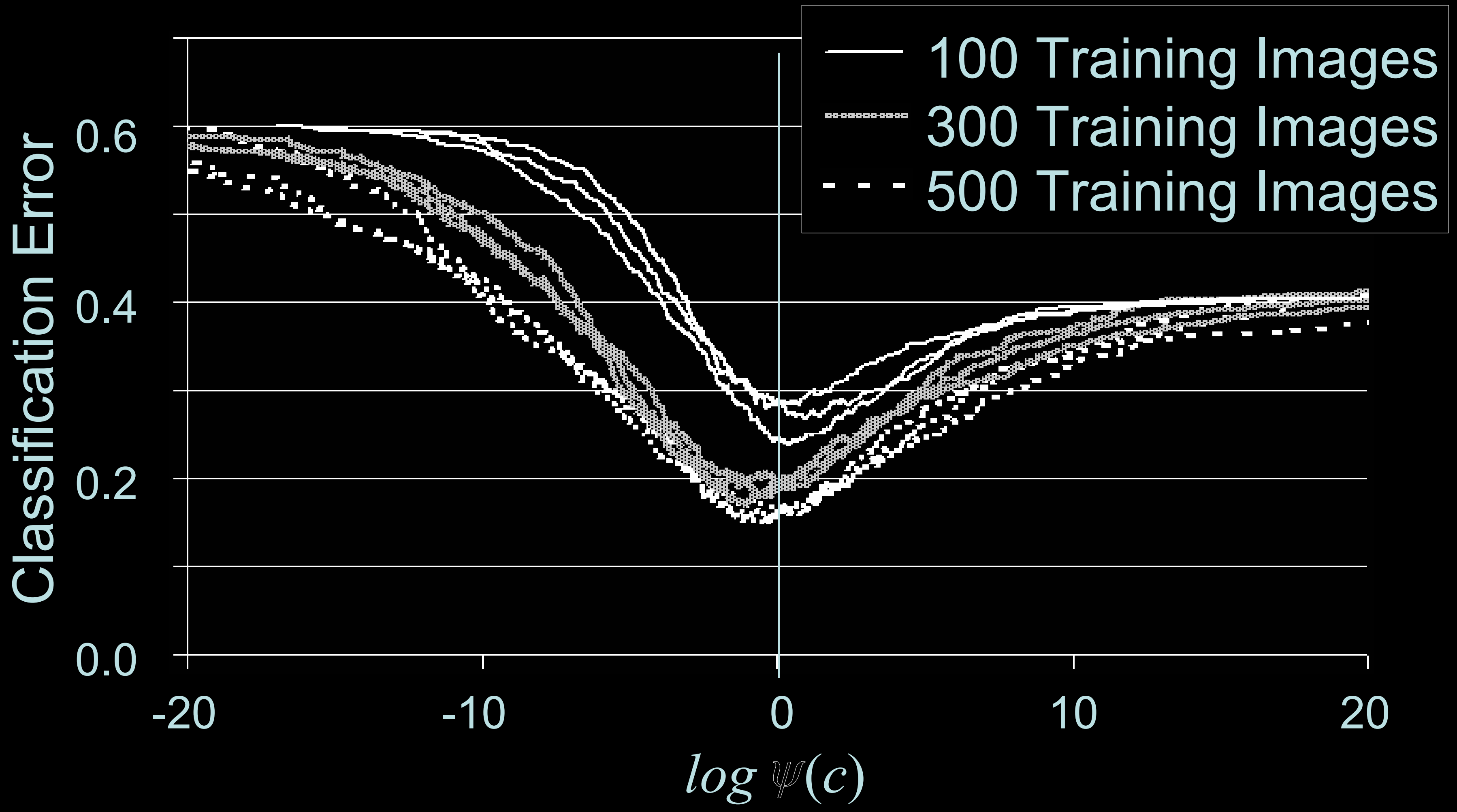
Classification: Incorrect

Localization Error 3.6%

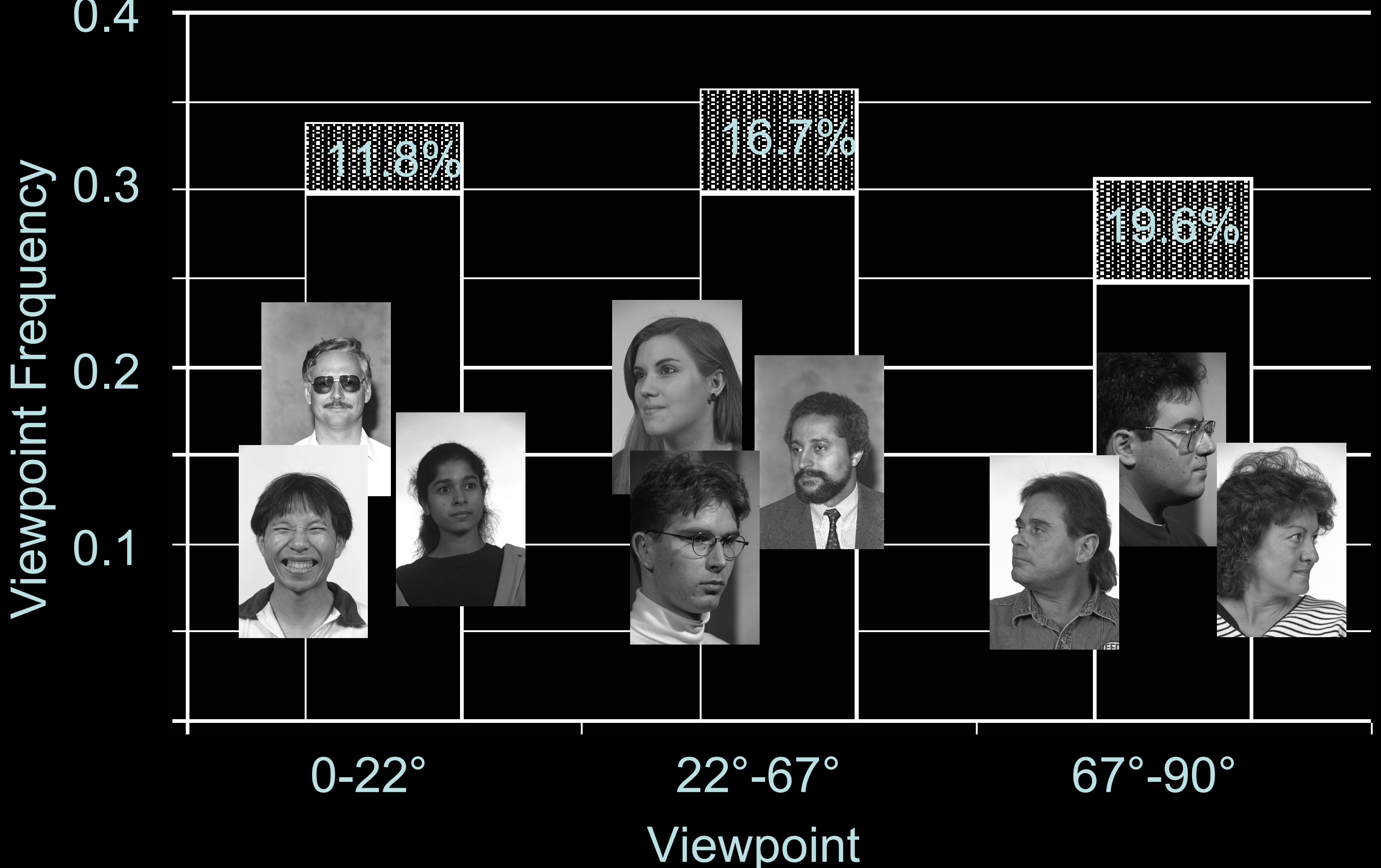
Excessive features of the opposite gender



Thresholds & Error

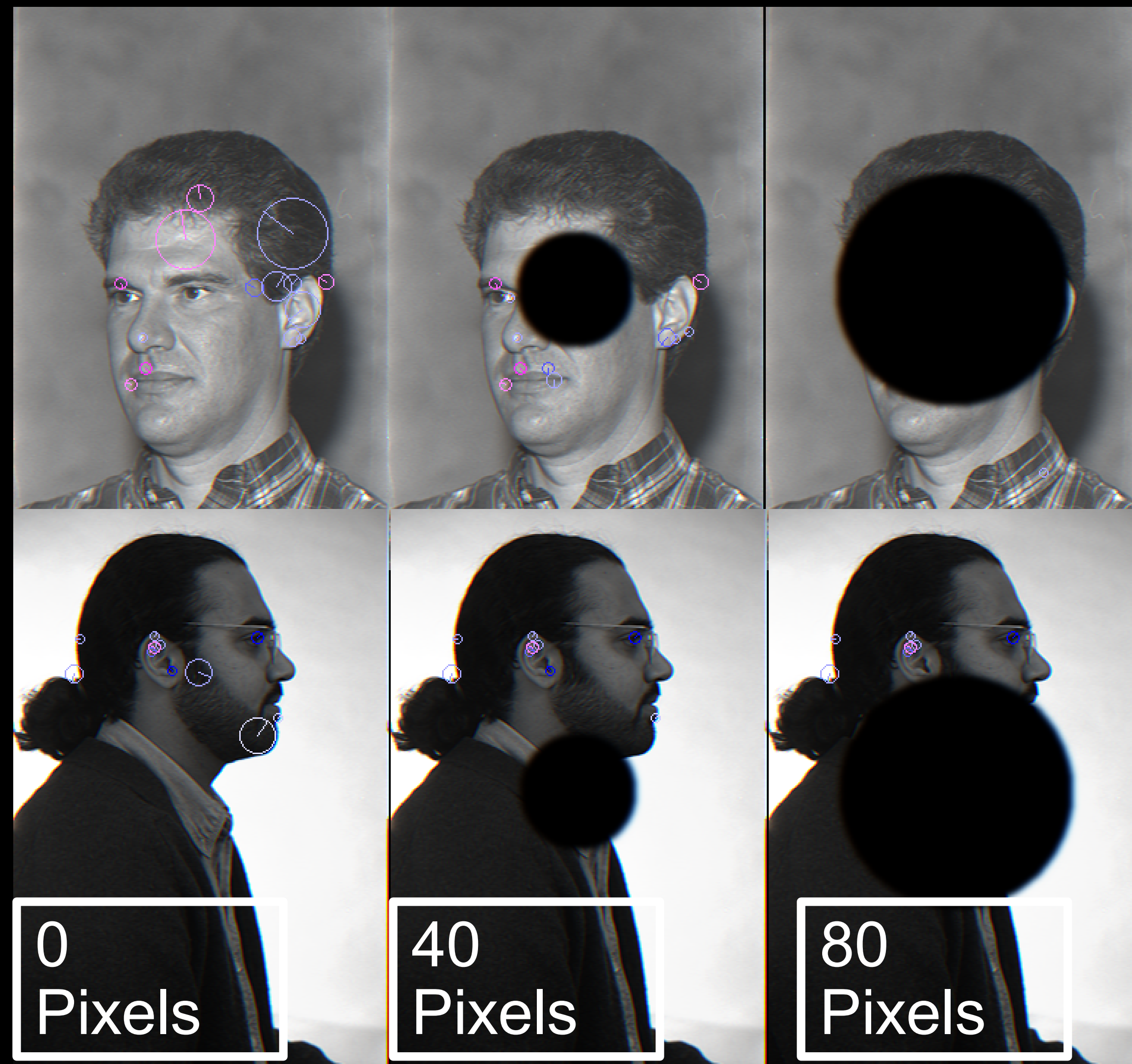


Classification Error over Viewpoint

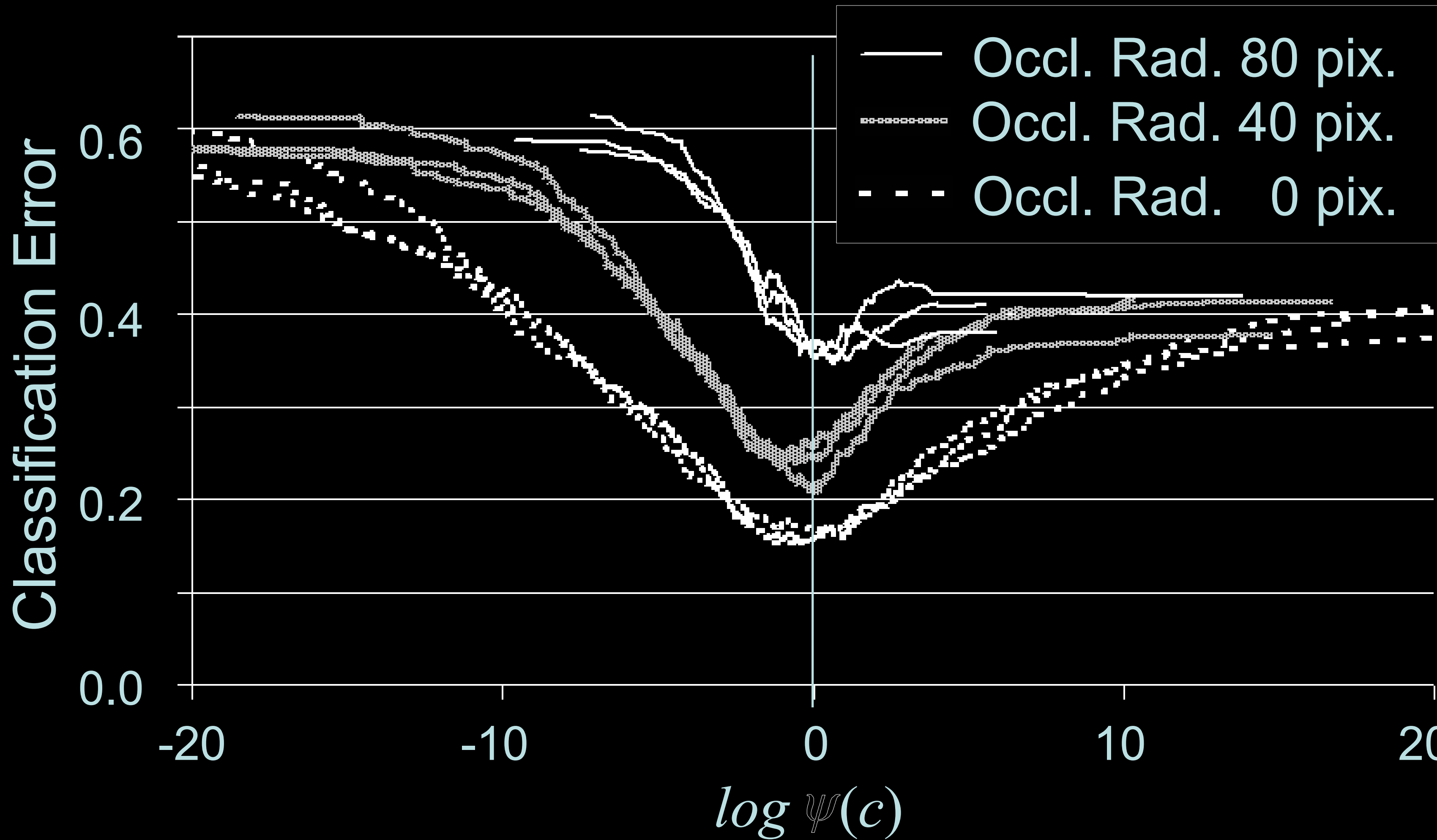


Occlusion

Artificial occluding circle in testing images, vary radius.



Occlusion: Thresholds & Error



Conclusion

- Integrated detection, localization and classification
  - Arbitrary viewpoints
  - Occlusion
- Viewpoint-invariant model
  - No viewpoint information required
- Evaluation
  - Color FERET
  - Large, natural dataset would be ideal