EEE-6561 Fundamentals of Biometric Identification

January 26th, 2018

Lecture #5 Face Detection

Damon L. Woodard, Ph.D.

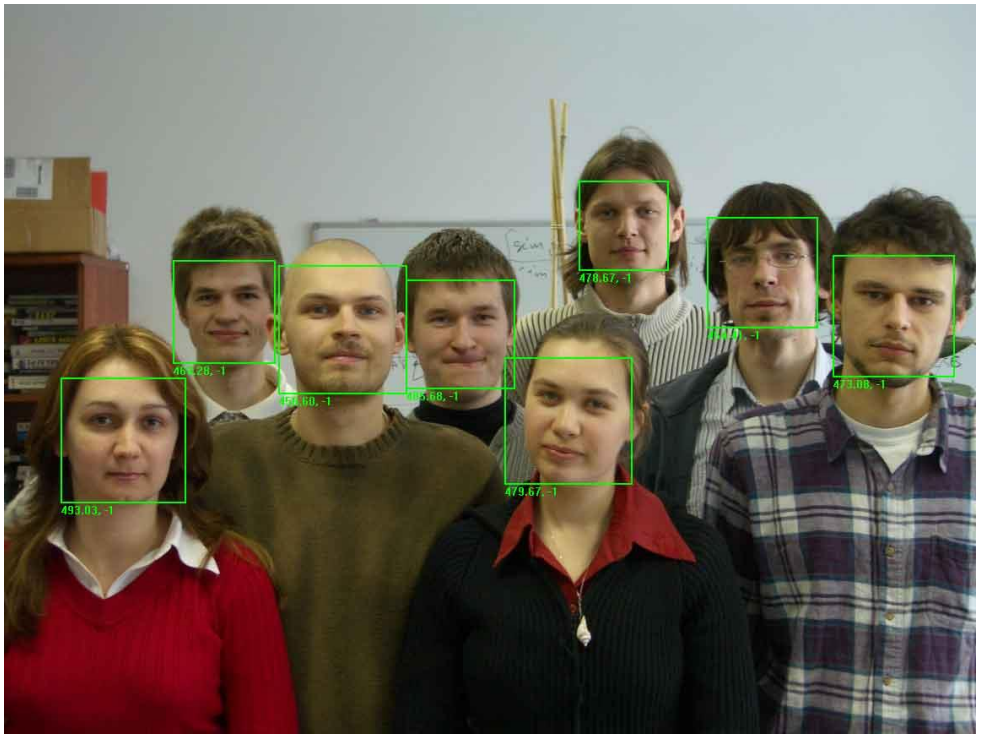
Dept. of Electrical and Computer Engineering

dwoodard@ece.ufl.edu

Face Detection

Definition: Identify and locate human faces in an image regardless of their:

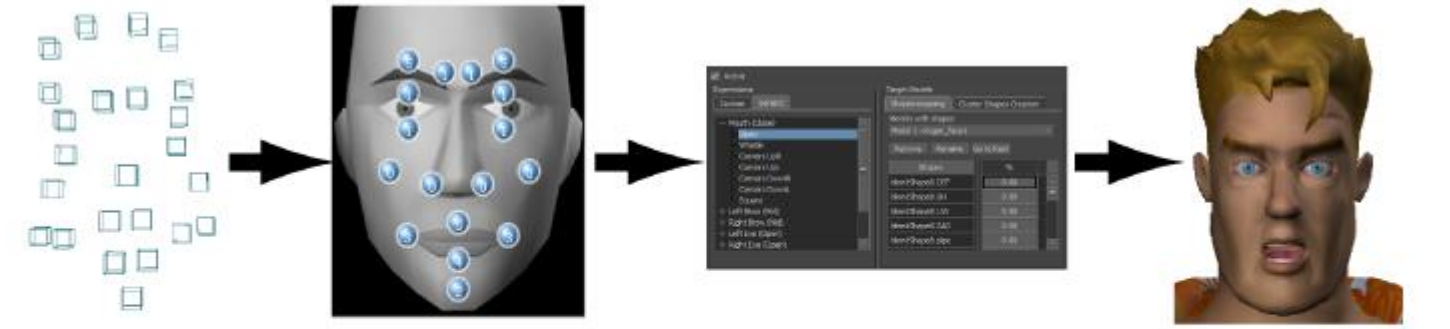
- Position
- Scale
- Orientation (in plane rotation)
- Pose (out-of plane rotation)
- Illumination



Why is Face Detection Important?

- First step for any fully automatic face recognition system
- First step in many surveillance systems
- Lots of applications

Example Application: Facial Motion Capture



facial optical data

Actor Face

Character Face with Generic expressions

Animated head model

Example Application: Facial Recognition

Related Problems

- Face localization:
  - Aim to determine the image position of a single face
  - A simplified detection problem with the assumption that an input image contains only one face
- Facial feature extraction:
  - To detect the presence and location of features such as eyes, nose, nostrils, eyebrow, mouth, lips, ears, etc
  - Usually assume that there is only one face in an image
- **Face recognition (identification)**
- Facial expression recognition
- Human pose estimation and tracking

Challenges of Face Detection

- **Pose (Out-of-Plane Rotation):**
  - frontal, 45 degree, profile, upside down
- **Presence or absence of structural components:**
  - beard, mustache and glasses
- **Facial expression:**
  - face appearance is directly affected by a person's facial expression
- **Occlusion:**
  - faces may be partially occluded by other objects
- **Orientation (In-Plane Rotation):**
  - face appearance directly vary for different rotations about the camera's optical axis
- **Imaging conditions:**
  - lighting (spectra**,** source distribution and intensity) and camera characteristics (sensor response, gain control, lenses), resolution

In One Thumbnail Face Image

- Consider a thumbnail 19×19 face pattern
- 256^{361} possible combination of gray values
  - $256^{361} = 2^{8 \times 361} = 2^{2888}$
- Total world population (as of 2004)
  - $6{,}400{,}000{,}000 \cong 2^{32}$
- 2^{90} times more than the world population!
- Extremely high dimensional space!

Face Detection Research Issues

- **Representation:** How to describe a typical face?
- **Scale:** How to deal with face of different size?
- **Search strategy:** How to spot these faces?
- **Speed:** How to speed up the process?
- **Precision:** How to locate the faces precisely?
- **Post processing**: How to combine detection results?

Face Detector: Ingredients

- **Target application domain:** single image, video
- **Representation:** holistic, feature, etc.
- **Pre processing:** histogram equalization, etc.
- **Cues:** color, motion, depth, voice, etc.
- **Search strategy:** exhaustive, greedy, focus of attention, etc.
- **Classifier design:** ensemble, cascade
- **Post processing:** combing detection results

Detecting Faces in Grayscale Images

- **Knowledge-based methods**
  - Encode human knowledge of what constitutes a typical face (usually, the relationship between facial features)

- **Feature invariant approaches**
  - Aim to find structural features of a face that exist even when the pose, viewpoint, or lighting conditions vary

- **Template matching methods**
  - Several standard patterns stored to describe the face as a whole or the facial features separately

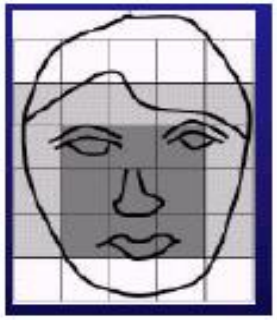
- **Appearance-based methods**
  - The models (or templates) are learned from a set of training images which capture the representative variability of facial appearance

Knowledge-Based Methods

- Top-down approach: Represent a face using a set of human-coded rules
- Example:
  - The center part of face has uniform intensity values
  - The difference between the average intensity values of the center part and the upper part is significant
  - A face often appears with two eyes that are symmetric to each other, a nose and a mouth
- Use these rules to guide the search process

Knowledge-based method [Yang and Huang 94]

- Multi-resolution focus-of-attention approach
- Level 1: (lowest resolution): apply the rule “the center part of the face has 4 cells with a basically uniform intensity” to search for candidates search for candidates
- Level 2: local histogram equalization followed by edge detection
- Level 3: search for eye and mouth features for validation



Knowledge-Based Method: [Kotropoulos & Pitas 94]

- Horizontal/vertical projection to search for candidates

$$HI(x) = \sum_{y=1}^{n} I(x,y) \quad VI(y) = \sum_{x=1}^{m} I(x,y)$$

- Search eyebrow/eyes, nostrils/nose for validation
- Difficult to detect multiple people or in complex background

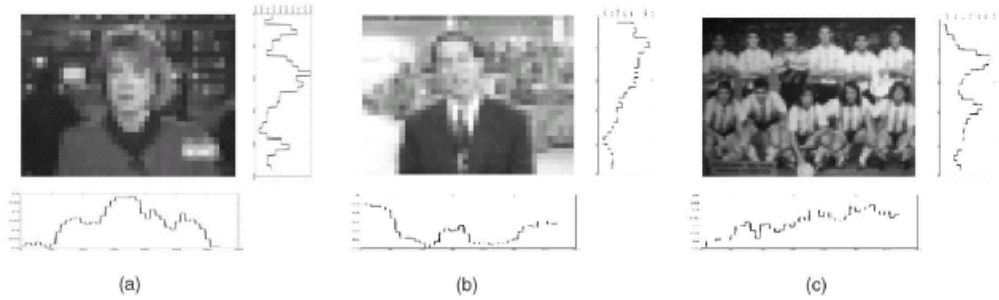


Fig. 3. (a) and (b) n = 8. (c) n = 4. Horizontal and vertical profiles. It is feasible to detect a single face by searching for the peaks in horizontal and vertical profiles. However, the same method has difficulty detecting faces in complex backgrounds or multiple faces as shown in (b) and (c).

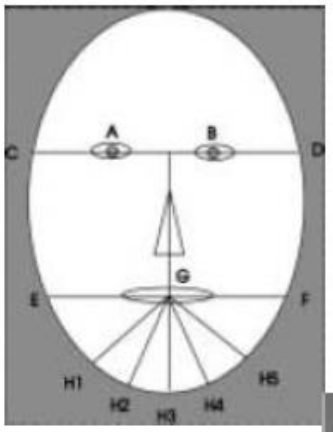
Knowledge-based Methods: Summary

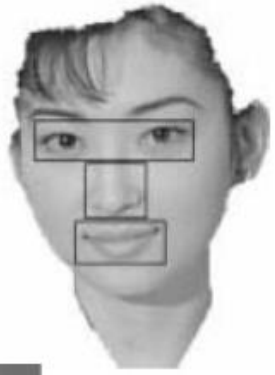
- **Pros:**
  - Easy to come up with simple rules to describe the features of a face and their relationships
  - Based on the coded rules, facial features in an input image are extracted first, and face candidates are identified
  - Work well for face localization in uncluttered background
- **Cons:**
  - Difficult to translate human knowledge into rules precisely: detailed rules fail to detect faces and general rules may find many false positives
  - Difficult to extend this approach to detect faces in different poses: implausible to enumerate all the possible cases

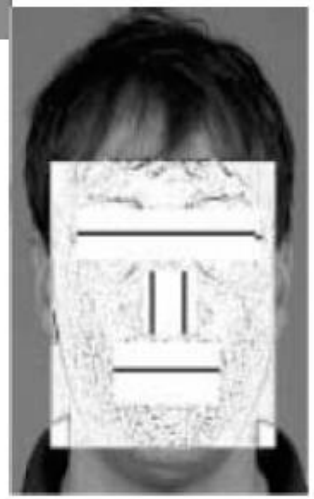
Feature-Based Methods

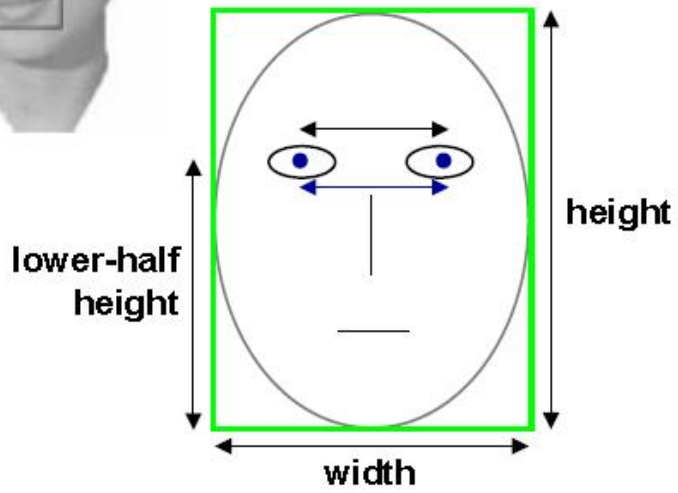
- Bottom-up approach: Detect facial features (eyes, nose, mouth, etc.) first
- Facial features: edge, intensity, shape, texture, color, etc.
- Aim to detect invariant features
- Group features into candidates and verify them

Feature-based Methods









Feature-Based Methods: Summary

- **Pros:**
  - Features are invariant to pose and orientation change

- **Cons:**
  - Difficult to locate facial features due to several corruption (illumination, noise, occlusion)

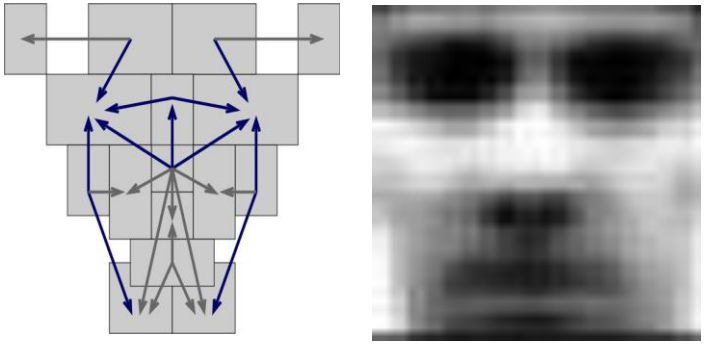
  - Difficult to detect features in complex background

Template matching methods

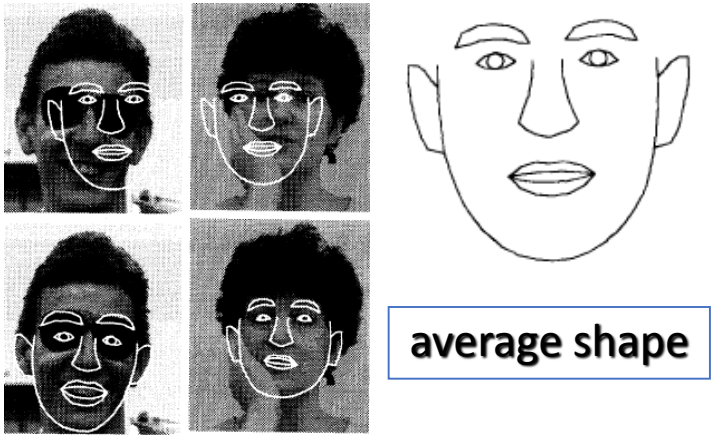
- Several standard patterns stored to describe the face as a whole or the facial features separately
  - Predefined: based on edges or regions
  - Deformable: based on facial contours
- Templates are hand-coded (not learned)
- Use correlation to locate faces

Face Template

- Use relative pair-wise ratios of the brightness of facial regions (14×16 pixels): the eyes are usually darker than the surrounding face [Sinha 94]
- Use average area intensity values rather than absolute pixel values
- See also Point Distribution Model (PDM) [Lanitis et al. 95]



Ration Template [Sinha 94]



[Lanitis et al. 95]