CSE 40537 / 60537: Biometrics

Biometrics Basics 6 / Face Recognition 1
Special pair of guest lectures this week: Vito Štruc

Hailing all the way from Ljubljana!

Don’t miss it!
Cross Validation

Leave one out:

\[ f_n = (x_1, y_1), \ldots, (x_{n-1}, y_{n-1}), (x_n, y_n), (x_{n+1}, y_{n+1}), \ldots (x_N, y_N) \]

\[ e_n = E_{\text{val}}(f) = e(f(x_n), y) \]

\[ E_{\text{cv}} = \frac{1}{N} \sum_{n=1}^{N} e_n \]
Cross Validation

\[ E_{cv} = \frac{1}{3} (e_1 + e_2 + e_3) \]
Cross Validation

Without Validation

With Validation

Image Credit: Y. Abu-Mostafa, *Learning from Data*
$K$-Fold Cross Validation

Generalize: $\frac{N}{K}$ training sessions on $N - K$ points each

Typical procedure is 10-fold cross validation: $K = \frac{N}{10}$
Recall

Recall = \frac{\{\text{Relevant Images}\} \cap \{\text{Retrieved Images}\}}{\{\text{Relevant Images}\}}

Probe

Gallery

Recall = 50\%
Precision

Precision = \frac{|\text{Relevant Images} \cap \text{Retrieved Images}|}{|\text{Retrieved Images}|}

Probe

Precision = 33\%
F-measure

What’s wrong with accuracy?

Sliding window detector

11 faces, and 1,000 negative windows

A “no” detector will be 98.9% accurate!
F-measure

Calculate harmonic mean of precision and recall:

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$
Area Under the Curve

0.9 - 1.0 = excellent
0.8 - 0.9 = good
0.7 - 0.8 = fair
0.6 - 0.7 = poor
0.5 - 0.6 = fail

*use when you can’t pick a threshold

Image adapted from figure appearing in the Journal of Clinical Pathology
Course Roadmap

Basics

Multi-Biometric Fusion

3 Core Modalities

Privacy & Security
A bit of history


- Manual extraction of fiducial points
- Automatic matching on geometry and spatial relationships


- Successfully identified 10 faces from 250 training iterations
- Custom photomultiplier hardware!
A building block for contemporary systems


“I wrote my face recognition program in an assembler language, and ran it on a machine with 10 microsecond cycle time and 20 kB of main memory. It was with pride that I tested the program with 1000 face images, a rare case at the time when testing with 10 images was called a ‘large-scale’ experiment.”

- T. Kanade, Rosenfeld Prize Speech, ICCV 2007
Prof. Kanade’s Computer

Figure 2-3: View of GIRLS subsystem: (1) I/O typewriter; (2) Paper tape reader; (3) CPU of MACC 7/F minicomputer; (4) Graf/pen; (5) Transmission cable to NEAC 2200/200; (6) Storage-type CRT; (7) GIRLS-interface controller; (8) Flying-spot scanner; (9) Character CRT; (10) Mini-disk device

Image Credit: T. Kanade 1973
Prof. Kanade’s Algorithm

Step 1. Detect Edges

Figure 3-2

(a) Original photograph
(b) Printout of the digital gray-level picture
(c) Binary picture

Picture input and line extraction. The dark horizontal line in the upper part is due to the burn in the CRT surface of the FSS used for digitization.

Image Credit: T. Kanade 1973
Prof. Kanade’s Algorithm

Step 2. Detect Parts of Face

Figure 3-3

Typical sequence of the analysis steps.

(a) top of head
(b) cheeks and sides of face
(c) nose, mouth, and chin
(d) chin contour
(e) face-side lines
(f) nose lines
(g) eyes
(h) face axis

Image Credit: T. Kanade 1973
Prof. Kanade’s Algorithm

Step 3. Detect Features
Step 4. Match Facial Parameters
How did he do?

<table>
<thead>
<tr>
<th></th>
<th>Rank of the real target picture</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>B</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>C</td>
<td>2 1 1 1 1</td>
</tr>
<tr>
<td>D</td>
<td>3 3 1 1 1</td>
</tr>
<tr>
<td>E</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>F</td>
<td>4 5 5 5 5</td>
</tr>
<tr>
<td>G</td>
<td>1 2 3 6</td>
</tr>
<tr>
<td>H</td>
<td>2 2 2 1</td>
</tr>
<tr>
<td>I</td>
<td>1 1 1 1 2</td>
</tr>
<tr>
<td>J</td>
<td>2 1 1 1 1</td>
</tr>
<tr>
<td>K</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>L</td>
<td>2 2 2 2 5</td>
</tr>
<tr>
<td>M</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>N</td>
<td>2 1 1 1 1</td>
</tr>
<tr>
<td>O</td>
<td>2 1 1 1 1</td>
</tr>
<tr>
<td>P</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>Q</td>
<td>5 5 2 1</td>
</tr>
<tr>
<td>R</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>S</td>
<td>2 2 1 1 1</td>
</tr>
<tr>
<td>T</td>
<td>3 2 2 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Person</th>
<th>Rank of the real target picture</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>B</td>
<td>2 2 2 2 2</td>
</tr>
<tr>
<td>C</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>D</td>
<td>2 1 1 1 1</td>
</tr>
<tr>
<td>E</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>F</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>G</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>H</td>
<td>6 3 3 5</td>
</tr>
<tr>
<td>I</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>J</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>K</td>
<td>1 2 2 5</td>
</tr>
<tr>
<td>L</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>M</td>
<td>5 5 5 5</td>
</tr>
<tr>
<td>N</td>
<td>1 1 2 2</td>
</tr>
<tr>
<td>O</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>P</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>Q</td>
<td>4 3 2 3</td>
</tr>
<tr>
<td>R</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>S</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>T</td>
<td>1 1 1 1</td>
</tr>
</tbody>
</table>

Average:

- Table 4-3: 1.9 1.8 1.5 1.8
- Table 4-4: 1.8 1.5 1.5 1.8

Number of correct identification:

- Table 4-3: 9 12 14 15
- Table 4-4: 14 15 14 14

Table 4-3
Results of identification test; feature measurement by computer.

Table 4-4
Results of identification test; feature measurement by human.

Image Credit: T. Kanade 1973
Face Processing Pipeline

1973

2015

Acquisition → Normalization → Alignment → Features → Recognition

Image Credit: T. Kanade 1973
Acquisition
Benchmark Data Sets

The appearance of a face is affected by many factors:

- Identity
- Face pose
- Illumination
- Facial expression
- Occlusion
- Facial hair

How do we capture these factors for evaluation?

Controlled vs. Uncontrolled Acquisition?
ORL (1994)

10 images of 40 subjects
Face Databases: FERET (1993-1997)

1196 Gallery Images

fa  fb  duplicate I  fc  duplicate II
frontal  expressions  0-1031 days after  lighting, cameras  540-1031 days after

pr  hr  qr  ql  hl  pl
+90°  +67.5°  +22.5°  -22.5°  -67.5°  -90°

pose
Yale & Yale B (1997)

Yale B:
10 Subjects x 9 Poses x 64 illumination conditions
AR (1998)

(1) neutral, (2) smile, (3) anger, (4) scream, (5) left light on, (6) right light on, (7) both lights on, (8) sun glasses, (9) sun glasses / left light (10) sun glasses / right light, (11) scarf, (12) scarf / left light, (13) scarf / right light
CMU PIE (2000)

68 people x 13 poses x 43 illumination conditions x 4 expressions