The Impact of Image Quality on the Performance of Face Recognition

Abhishek Dutta\textsuperscript{1}, Raymond Veldhuis, Luuk Spreeuwers

University of Twente, Netherlands.

May 25, 2012

33rd WIC Symposium on Information Theory in the Benelux, May 24-25, 2012

\textsuperscript{1}\url{http://abhishekdutta.org}
Introduction

What is the impact of image quality on the performance of a commercial face recognition system?
Introduction

What is the impact of image quality on the performance of a commercial face recognition system?²

² Cognitec FaceVACS SDK 8.4.0
A Quick Question: Fill in the blank

Face recognition is ____________

- difficult

- very difficult
A Quick Question: Fill in the blank

Face recognition is ____________

- difficult
- very difficult

[MultiPIE]
A Quick Question: Fill in the blank

Face recognition is ____________

- difficult
- very difficult

face recognition performance depends on image quality of the pair of images participating in the comparison process.
Measure of Face Recognition Performance

Reference Test \[\rightarrow\] Face Recognition System \[\rightarrow\] Similarity Score
Measure of Face Recognition Performance

Reference → Face Recognition System → Similarity Score

Test

Verification Rate vs. False Accept Rate

random guess
Measure of Face Recognition Performance

Reference → Face Recognition System → Test → Similarity Score

Verification Rate

False Accept Rate

- Good system
- Bad system
- Random guess
Measure of Face Recognition Performance

Reference  $\rightarrow$ Face Recognition System $\rightarrow$ Similarity Score

Verification Rate $\rightarrow$ False Accept Rate

Forbidden Region
Area Under ROC (AUC) = 0.5 (or 50 %)
Pose and Illumination

* : Illumination (flash) placed just above the respective camera
gradual performance degradation as ref. pose moves away from test pose
Pose and Illumination

gradual performance degradation as ref. pose moves away from test pose
Pose and Illumination ...

- gradual reduction in recognition performance as the reference set pose moves away from the pose in the test set
- role of illumination insignificant for same pose (and camera\(^3\)) in the test and reference set

\(^3\)in the MultiPIE data set, if we exactly match the pose, we are also matching all the imaging characteristics
Pose and Illumination ...

- gradual reduction in recognition performance as the reference set pose moves away from the pose in the test set
- role of illumination insignificant for same pose (and camera$^3$) in the test and reference set

Test and reference images having *similar pose and captured by same camera* can greatly improve recognition performance.

$^3$In the MultiPIE data set, if we exactly match the pose, we are also matching all the imaging characteristics
Resolution

Test & ref. pose

Resol

Area Under ROC (AUC)

Reference set resolution

Test set res.
(dist. between eyes in pixels)

- 40x30 (13)
- 60x45 (16)
- 80x60 (18)
- 100x75 (20)
- 120x90 (22)
- 160x120 (25)
- 200x150 (28)
- 640x480 (50)

Pose

Gaussian Noise (mean= 0)
var. = 0.007 var. = 0.3

Illumination

60 x 45 120 x 90

Resolution

60 x 45 120 x 90
Resolution ...

<table>
<thead>
<tr>
<th>Reference set resolution</th>
<th>Area Under ROC (AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40x30</td>
<td>0.5</td>
</tr>
<tr>
<td>60x45</td>
<td>0.5</td>
</tr>
<tr>
<td>80x60</td>
<td>0.5</td>
</tr>
<tr>
<td>100x75</td>
<td>0.5</td>
</tr>
<tr>
<td>120x90</td>
<td>0.6</td>
</tr>
<tr>
<td>160x120</td>
<td>0.7</td>
</tr>
<tr>
<td>200x150</td>
<td>0.8</td>
</tr>
<tr>
<td>640x480</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Test set res. (dist. between eyes in pixels)
- 40x30 (13)
- 60x45 (16)
- 80x60 (18)
- 100x75 (20)
- 120x90 (22)
- 160x120 (25)
- 200x150 (28)
- 640x480 (50)

Analysis:
- Recognition performance improves with the resolution of the test and reference set (as expected).
- Relative difference in pose between the test and reference image determines the extent of influence of resolution.
Noise

Ref. set Gaussian noise variance (mean = 0)

Test set Gaussian noise variance (mean = 0)

Motion Blur (angle = 0)
length = 0
length = 17

Gaussian Noise (mean= 0)

var. = 0.007
var. = 0.3
Analysis:

- recognition performance degrades with the noise in the test and reference set (as expected).
- relative difference in pose between the test and reference image determines the extent of influence of zero mean Gaussian noise.
Blur

Test & ref. pose

Area Under ROC (AUC)

Ref. set motion blur length (angle = 0)

Test set motion blur length (angle = 0)

Gaussian Noise (mean = 0)
var. = 0.007
var. = 0.3

Pose Illumination
60 x 45 120 x 90

Resolution

Motion Blur (angle = 0)

length = 03 length = 17
Analysis:

- recognition performance degrades with the blur in the test and reference set (as expected).
- relative difference in pose between the test and reference image determines the extent of influence of image blur.
Summary

<table>
<thead>
<tr>
<th>Quality</th>
<th>Difference in Area Under ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pose</td>
<td>~ 50%</td>
</tr>
<tr>
<td>Resolution</td>
<td>~ 35%</td>
</tr>
<tr>
<td>Noise (Gaussian)</td>
<td>~ 35%</td>
</tr>
<tr>
<td>Blur (Motion)</td>
<td>~ 20%</td>
</tr>
<tr>
<td>Illumination</td>
<td>~ 20%</td>
</tr>
</tbody>
</table>

- approximately matching non-frontal pose between test and reference images (and using same camera) can greatly improve recognition performance.
Summary

<table>
<thead>
<tr>
<th>Quality</th>
<th>Difference in Area Under ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pose</td>
<td>~ 50%</td>
</tr>
<tr>
<td>Resolution</td>
<td>~ 35%</td>
</tr>
<tr>
<td>Noise (Gaussian)</td>
<td>~ 35%</td>
</tr>
<tr>
<td>Blur (Motion)</td>
<td>~ 20%</td>
</tr>
<tr>
<td>Illumination</td>
<td>~ 20%</td>
</tr>
</tbody>
</table>

- approximately matching non-frontal **pose** between test and reference images (and using same **camera**) can greatly improve recognition performance.

- it is the **relative difference in pose** between the test and reference image that determines the extent of influence that other **quality** parameters like illumination, noise, motion blur, and resolution have on the face recognition performance.
### Summary

<table>
<thead>
<tr>
<th>Quality</th>
<th>Difference in Area Under ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pose</td>
<td>~ 50%</td>
</tr>
<tr>
<td>Resolution</td>
<td>~ 35%</td>
</tr>
<tr>
<td>Noise (Gaussian)</td>
<td>~ 35%</td>
</tr>
<tr>
<td>Blur (Motion)</td>
<td>~ 20%</td>
</tr>
<tr>
<td>Illumination</td>
<td>~ 20%</td>
</tr>
</tbody>
</table>

- Approximately matching non-frontal **pose** between test and reference images (and using same **camera**) can greatly improve recognition performance.
- It is the **relative difference in pose** between the test and reference image that determines the extent of influence that other **quality** parameters like illumination, noise, motion blur, and resolution have on the face recognition performance.
- Face recognition **performance depends** on quality of **image pair**.
Limitations

- We have assumed that the image quality parameters are independent. In reality, all the quality parameters co-exist and presence or absence of one quality parameter (like pose, blur, etc) might affect the behavior of other quality parameters (like resolution, noise, etc).
Limitations

- We have assumed that the image quality parameters are independent. In reality, all the quality parameters co-exist and presence or absence of one quality parameter (like pose, blur, etc) might affect the behavior of other quality parameters (like resolution, noise, etc).

- All the images used in this study were taken from a single image data set. Although test and reference images differed by session, ideally both test and reference images should have been taken from the different data set in order to simulate the conditions present in a real forensic case.
Limitations

- We have assumed that the image quality parameters are independent. In reality, all the quality parameters co-exist and presence or absence of one quality parameter (like pose, blur, etc) might affect the behavior of other quality parameters (like resolution, noise, etc).

- All the images used in this study were taken from a single image data set. Although test and reference images differed by session, ideally both test and reference images should have been taken from the different data set in order to simulate the conditions present in a real forensic case.

- These findings are limited by the inclusion of a specific commercial face recognition system in this study.
Thank You

For more details, refer to:

Notes for Q/A Session

Table: Image quality variations included in this study.

<table>
<thead>
<tr>
<th>Quality</th>
<th>Camera</th>
<th>Flash</th>
<th>Resolution</th>
<th>Motion Blur</th>
<th>Gaus. Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pose and Illumination</td>
<td>$c_i, c_j \in \mathbb{C}$</td>
<td>$f_i, f_j \in \mathbb{F}$</td>
<td>$r_i, r_j = D_0$</td>
<td>0, 0</td>
<td>0, 0</td>
</tr>
<tr>
<td>Resolution</td>
<td>19.1, {∗}</td>
<td>18, {∗∗}</td>
<td>$r_i, r_j \in \mathbb{R}$</td>
<td>0, 0</td>
<td>0, 0</td>
</tr>
<tr>
<td>Gaussian Noise</td>
<td>19.1, {∗}</td>
<td>18, {∗∗}</td>
<td>$r_i, r_j = D_0$</td>
<td>0, 0</td>
<td>$\bar{\sigma}_i, \bar{\sigma}<em>j \in \mathbb{N}</em>{\bar{\sigma}}$</td>
</tr>
<tr>
<td>Motion Blur</td>
<td>19.1, {∗}</td>
<td>18, {∗∗}</td>
<td>$r_i, r_j = D_0$</td>
<td>$l_i, l_j \in \mathbb{B}_l$</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

$\mathbb{C} = \{19.1, 19.0, 04.1, 05.0, 05.1, 14.0, 13.0, 08.0, 08.1\}$, $\mathbb{F} = \{02, 04, 14, 05, 15, 06, 07, 16, 08, 09, 17, 10, 18, 12\}$, $\mathbb{R} = [640 \times 480, \ldots, 60 \times 45]$, $\mathbb{B}_l$ (length in pixels) = [1, 3, 5, 7, 13, 17, 21], Note: angle = 0 $\mathbb{N}_{\bar{\sigma}}$ (variance) = [0.001, 0.007, 0.03, 0.07, 0.1, 0.2]. Note: mean = 0 {∗} = {19.1, 05.1}, {∗∗} = {10, 07}, $D_0 = 640 \times 480$

Table: Test and reference set specifications (source: MultiPIE)

<table>
<thead>
<tr>
<th></th>
<th>Test Set (Probe)</th>
<th>Reference Set (Gallery)</th>
</tr>
</thead>
<tbody>
<tr>
<td>size (image count)</td>
<td>479</td>
<td>442</td>
</tr>
<tr>
<td>person count</td>
<td>319</td>
<td>268</td>
</tr>
<tr>
<td>session</td>
<td>01,03</td>
<td>02,04</td>
</tr>
<tr>
<td>expression</td>
<td>neutral</td>
<td>neutral</td>
</tr>
<tr>
<td>eye annotation</td>
<td>manual</td>
<td>manual</td>
</tr>
</tbody>
</table>
Figure: Camera and flash location for all the images used in this experiment (source: MultiPIE data set)