Occlusions in 3D Face Recognition

Alessandro Colombo

Imaging and Vision Laboratory
DISCo, Dipartimento di Informatica Sistemistica e Comunicazione
Università degli Studi di Milano Bicocca

Lyon, 28 January 2011
Outline

- Preliminaries
- Detection of occluded 3D faces
- Recognizing occluded 3D faces: a restoration approach
- Experimental Results on artificial and real occlusions
- Future Directions
Why occlusions?

Authentication

Identification
Occlusions in 3D
Requirements

- Fully automatic
- Scalable
- *Is an holistic approach possible?*
System design choices

- Occlusion detection and restoration based pipeline.
Occlusion tolerant detection and *pre*-matching normalization.
System design choices

- Holistic approach through occlusion detection and restoration.
Faces are analyzed in order to detect occlusions.
If occlusions are too large the acquisition may be eventually rejected.
System design choices

- Occlusions are cleared and the missing facial surface is reconstructed.
System design choices

- Any state of the art recognition strategy may be applied.
Occlusion Tolerant 3D Face Detector

BEGIN

Curvature analysis

Candidate facial features extraction

Candidate faces generation

For each candidate face

Candidate faces registration

Candidate faces classification

Multiple detection filtering

END

► Hypothesis generation and test structure.
Occlusion Tolerant 3D Face Detector

BEGIN

Curvature analysis

Candidate facial features extraction

Candidate faces generation

For each candidate face

Candidate faces registration

Candidate faces classification

Multiple detection filtering

END

▶ Hypothesis generation.
Occlusion Tolerant 3D Face Detector

For each candidate face

Curvature analysis

Candidate faces generation

Candidate facial features extraction

Candidate faces registration

Candidate faces classification

Multiple detection filtering

END

◄ Hypothesis test.
Curvature Analysis
Curvature Analysis

Mean Curvature

\[ H(x_0, y_0) = \frac{(1+f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1+f_x^2)f_{yy}}{2(1+f_x^2+f_y^2)^{\frac{3}{2}}} \]
Curvature Analysis

Gaussian Curvature

\[ K(x_0, y_0) = \frac{f_{xx} f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^2} \]
Curvature Analysis

HK Classification

<table>
<thead>
<tr>
<th></th>
<th>$K &lt; 0$</th>
<th>$K = 0$</th>
<th>$K &gt; 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H &lt; 0$</td>
<td>Hyperbolic concave</td>
<td>Cylindrical concave</td>
<td>Elliptical concave</td>
</tr>
<tr>
<td>$H = 0$</td>
<td>Hyperbolic symmetric</td>
<td>Planar</td>
<td>Impossible</td>
</tr>
<tr>
<td>$H &gt; 0$</td>
<td>Hyperbolic convex</td>
<td>Cylindrical convex</td>
<td>Elliptical convex</td>
</tr>
</tbody>
</table>
Candidate Features Extraction
Candidate Features Extraction

Candidate Noses

Positive Mean regions;

\[ H(u, v) > T_h \]
Candidate Features Extraction

Elliptical concave regions;

$$|K(u, v)| > T_k$$
Candidate Features Extraction

Region size, shape, average curvature etc.

Final candidate filtering
Candidate faces generation

Candidate Eye and Nose Regions

Eyes Pairs

Eye-Nose Pairs

Face Triangles

Upward face hypothesis

Downward face hypothesis

Left eye hypothesis

Right eye hypothesis

A. Colombo (IVL Laboratory)
Candidate faces rough registration
Candidate faces fine registration

- Global registration through ICP with a mean reference face
Candidate faces fine registration

- Global registration through ICP with a mean reference face
- Projective matcher
Candidate faces fine registration

- Global registration through ICP with a mean reference face
- Projective matcher
- Geometric correspondences rejector
Candidate faces classification

Standard PCA classification
Candidate faces classification

Standard PCA classification

\[ \approx + y_1 \times + y_2 \times + \cdots + y_n \times \]
Candidate faces classification

Standard PCA classification

\[ \sim + y_1 \times + y_2 \times + \cdots + y_n \times = \]
Candidate faces classification

Standard PCA classification

\[
\approx + y_1 \times + y_2 \times + \cdots + y_n \times = D.F.F.S.
\]

\[
D.F.F.S. = \| - \|
\]
Candidate faces classification

Standard PCA classification

\[ \simeq + y_1 \times + y_2 \times + \cdots + y_n \times = \]

\[ D.F.F.S. = \| - \| \| - \| \]
Candidate faces classification

Standard PCA classification

\[
\begin{align*}
\sim & + y_1 \times + y_2 \times + \cdots + y_n \times = \\
D.F.F.S. & = \| - \| \ll \| - \|
\end{align*}
\]

- Missing data?
Candidate faces classification

Standard PCA classification

\[ D.F.F.S. = \| \cdot \| - \| \cdot \| \leq \| \cdot \| - \| \cdot \| \]

- Missing data?
- Occlusions?
Candidate faces classification

Gappy PCA
Candidate faces classification
Gappy PCA
Candidate faces classification

Gappy PCA

\[ \approx + y_1 \times + y_2 \times + \cdots + y_n \times \]
Candidate faces classification

Gappy PCA

\[ \mathbf{\approx} \mathbf{y}_1 \times + \mathbf{y}_2 \times + \cdots + \mathbf{y}_n \times = \]
Candidate faces classification

Gappy PCA

\[ \approx y_1 \times + y_2 \times + \cdots + y_n \times = \]

\[ \min \| \text{mask} \| \]
Candidate faces classification

Gappy PCA

\[
\approx + y_1 \times + y_2 \times + \cdots + y_n \times =
\]

\[
\min || \quad - \quad ||_{mask} = D.F.F.S.
\]
Candidate faces classification
Gappy PCA

\[ \approx \sum y_1 \times \cdots + y_n \times = \min \left\| \frac{\text{mask}}{D.F.F.S.} \right\| \]

- Closed form for GPCA projection; really fast.
Candidate faces classification
Gappy PCA

\[ \approx + y_1 \times + y_2 \times + \cdots + y_n \times = \]

\[ \min \| \begin{array}{c} \text{mask} \end{array} \| = D.F.F.S. \]

- Closed form for GPCA projection; really fast.
- Occlusions?
Candidate faces classification

Rough occlusion detection
Candidate faces classification

Rough occlusion detection
Candidate faces classification
Rough occlusion detection
Candidate faces classification
Rough occlusion detection
Candidate faces classification

Summary

( ) → ( ) → ( ) → ( )
Candidate faces classification

Summary

\[
\begin{pmatrix}
\begin{array}{c}
- \\
\end{array}
\end{pmatrix} \rightarrow \begin{array}{c}
\end{array} \rightarrow \begin{array}{c}
\end{array}
\]

\[
\simeq \begin{array}{c}
\end{array} + y_1 \times \begin{array}{c}
\end{array} + y_2 \times \begin{array}{c}
\end{array} + \cdots + y_n \times \begin{array}{c}
\end{array}
\]

= \begin{array}{c}
\end{array}
Candidate faces classification

Summary

\[ \begin{align*}
(\text{face} - \text{mask}) & \rightarrow \text{mask} \\
\simeq & + y_1 \times + y_2 \times + \cdots + y_n \times \\
D.F.F.S. & = \| \text{face} - \| \text{mask} & &
\end{align*} \]
Candidate faces classification

Summary

\[
\begin{pmatrix}
\end{pmatrix}
\rightarrow
\begin{pmatrix}
\end{pmatrix}
\rightarrow
\begin{pmatrix}
\end{pmatrix}
\]

\[
\simeq + y_1 \times + y_2 \times + \cdots + y_n \times =
\]

\[
D.F.F.S. = \left\| \begin{array}{c}
\end{array} \right\| - \left\| \begin{array}{c}
\end{array} \right\| \text{mask}
\]

\[\text{if } D.F.F.S. < \tau \rightarrow \text{candidate is a face}\]
Candidate faces classification

Summary

\[ \sum y_1 \times \sum y_2 \times \cdots \sum y_n \times \]  

\[ D.F.F.S. = \| \text{mask} \| - \| \]  

- if $D.F.F.S. < \tau \rightarrow \text{candidate is a face}$
- If the occluded area is large the candidate may be rejected
A. Colombo (IVL Laboratory)
Occlusion Detection and Face Restoration

- Need a more precise detection of occluded pixels and reconstruction of the face for maximizing recognition performances
Occlusion Detection and Face Restoration

- a "large" threshold $\tau$ is used to discriminate occluded pixels
a first reconstruction is generated through GPCA projection
Occlusion Detection and Face Restoration

- a smaller threshold $\tau_2$ is used. Morphological operators are also applied
Occlusion Detection and Face Restoration

A new projection with the refined mask is generated
Occlusion Detection and Face Restoration

If the occluded area is large the face may be rejected
Experimental Results

Dataset Requirements:

- large number of subjects
- different occlusion types (object type, location, size of occluded area)

Ground truth needed for an in-depth analysis:

- occluded/non-occluded regions, *pixel precision*
- landmarks, *even if occluded*
- original, *non-occluded surface*
Experimental Results

Dataset Requirements:

- large number of subjects
- different occlusion types (object type, location, size of occluded area)

Ground truth needed for an in-depth analysis:

- occluded/non-occluded regions, *pixel precision*
- landmarks, *even if occluded*
- original, *non-occluded surface*
Artificial Occlusions Generation

- Random or semi-random position and orientation of the object
Artificial Occlusions Generation

Occluding objects
Artificial Occlusions Generation

The Artificially Occluded UND Dataset

UND Dataset:

- 277 subjects (male/female, different races)
- 953 acquisitions (frontal, neutral expression, 2D+3D)
- facial features annotation

A. Colombo (IVL Laboratory)
## Face Detection Results

### Artificially occluded UND Dataset

<table>
<thead>
<tr>
<th>Test Set</th>
<th>Classifier</th>
<th>Total Faces</th>
<th>False Positives</th>
<th>False Negatives</th>
<th>Detected Faces</th>
</tr>
</thead>
<tbody>
<tr>
<td>UND (non-occluded)</td>
<td>PCA</td>
<td>951</td>
<td>3</td>
<td>1</td>
<td>950 (99.9%)</td>
</tr>
<tr>
<td></td>
<td>GPCA</td>
<td>951</td>
<td>18</td>
<td>0</td>
<td>951 (100%)</td>
</tr>
<tr>
<td>UND (artificially occluded)</td>
<td>PCA</td>
<td>951</td>
<td>165</td>
<td>483</td>
<td>468 (49.2%)</td>
</tr>
<tr>
<td></td>
<td>GPCA</td>
<td>951</td>
<td>135</td>
<td>104</td>
<td>847 (89.8%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Object Type</th>
<th>Total Faces</th>
<th>False Positives</th>
<th>False Negatives</th>
<th>Detected Faces</th>
</tr>
</thead>
<tbody>
<tr>
<td>hand with cup</td>
<td>91</td>
<td>5</td>
<td>3</td>
<td>88 (96.7%)</td>
</tr>
<tr>
<td>eyeglasses</td>
<td>184</td>
<td>4</td>
<td>0</td>
<td>184 (100.0%)</td>
</tr>
<tr>
<td>free hands</td>
<td>181</td>
<td>46</td>
<td>41</td>
<td>140 (77.3%)</td>
</tr>
<tr>
<td>newspaper</td>
<td>77</td>
<td>20</td>
<td>18</td>
<td>59 (76.6%)</td>
</tr>
<tr>
<td>hand with phone</td>
<td>97</td>
<td>1</td>
<td>0</td>
<td>97 (100.0%)</td>
</tr>
<tr>
<td>scarf</td>
<td>123</td>
<td>42</td>
<td>29</td>
<td>94 (76.4%)</td>
</tr>
<tr>
<td>scissors</td>
<td>109</td>
<td>9</td>
<td>7</td>
<td>102 (93.6%)</td>
</tr>
<tr>
<td>hat</td>
<td>89</td>
<td>8</td>
<td>6</td>
<td>83 (93.6%)</td>
</tr>
<tr>
<td>all</td>
<td>951</td>
<td>135</td>
<td>104</td>
<td>847 (89.8%)</td>
</tr>
</tbody>
</table>
Face Detection Results

Successes
Face Detection Results

Failures
Face Detection Results
Detector Precision vs Recall

- Precision vs Recall varying the minimum valid area fraction threshold
Face Recognition Results
Artificially occluded UND Dataset

- Fisherfaces recognition approach; Manual normalization
- Training: 474 images; 3 x subject (non-occluded)
- Test: 477 remaining images (occluded)

<table>
<thead>
<tr>
<th>Set</th>
<th>EER</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>non occluded</td>
<td>0.020</td>
<td>99.35%</td>
</tr>
<tr>
<td>occluded</td>
<td>0.488</td>
<td>27.45%</td>
</tr>
<tr>
<td>restored</td>
<td>0.062</td>
<td>96.08%</td>
</tr>
<tr>
<td>exact restoration</td>
<td>0.022</td>
<td>99.35%</td>
</tr>
</tbody>
</table>
Face Recognition Results
Artificially occluded UND Dataset

- Full Automatic Pipeline: face detection, occlusion detection and restoration
- Fisherfaces recognition approach;
- Same training and test sets

- EER drops from 0.062 to 0.115
- IR drops from 96.08% to 81%

<table>
<thead>
<tr>
<th>Occlusion type</th>
<th>Manual normalization</th>
<th>Automatic normalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cup</td>
<td>100%</td>
<td>81%</td>
</tr>
<tr>
<td>Scissors</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Eyeglasses</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Hat</td>
<td>100%</td>
<td>63.16%</td>
</tr>
<tr>
<td>Hand</td>
<td>97.18%</td>
<td>96.49%</td>
</tr>
<tr>
<td>Phone</td>
<td>100%</td>
<td>93.75%</td>
</tr>
<tr>
<td>Newspaper</td>
<td>92%</td>
<td>87.5%</td>
</tr>
<tr>
<td>Scarf</td>
<td>95.92%</td>
<td>46.34%</td>
</tr>
</tbody>
</table>
Face Recognition Results
Artificially occluded UND Dataset

<table>
<thead>
<tr>
<th>subset</th>
<th>from %</th>
<th>to %</th>
</tr>
</thead>
<tbody>
<tr>
<td>small</td>
<td>0%</td>
<td>20%</td>
</tr>
<tr>
<td>medium</td>
<td>20%</td>
<td>40%</td>
</tr>
<tr>
<td>large</td>
<td>40%</td>
<td>82%</td>
</tr>
</tbody>
</table>
Restoration Accuracy
Artificially occluded UND Dataset

- Average absolute difference between depth of corresponding points for each pair of restored and original faces.

- Average absolute z difference between the pixels of pairs of non-occluded acquisitions of the same subject: 0.93 mm.

- GPCA is able to precisely reconstruct faces even if occluded up to 40%.
Occlusion Detection Accuracy
Artificially occluded UND Dataset

- Occlusion detection performances evaluated at the pixel level

\[ \text{precision} = \frac{tp}{tp + fp} \]

\[ \text{recall} = \frac{tp}{tp + fn} \]
Face Recognition Results

Bosphorus Database

- 105 subjects; 4652 acquisitions, 381 occluded.
- Occluding objects:
  - mouth with hand
  - one eye with hand
  - eyeglasses (not sunglasses, normal eyeglasses)
  - hair
# Face Recognition Results

**Bosphorus Database**

<table>
<thead>
<tr>
<th>Method</th>
<th>Occlusion Type</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eye</td>
<td>Mouth</td>
<td>Eyeglasses</td>
<td>Hair</td>
</tr>
<tr>
<td>Alyuz et al.[1] (semi-auto)</td>
<td>93.62%</td>
<td>93.62%</td>
<td>97.87%</td>
<td>89.66%</td>
</tr>
<tr>
<td>Proposed approach (auto)</td>
<td>91.18%</td>
<td>74.75%</td>
<td>94.23%</td>
<td>90.47%</td>
</tr>
</tbody>
</table>

---


A. Colombo (IVL Laboratory)
UMB-DB Database
The University of Milano-Bicocca 3D face Database

- 143 subjects
- 1461 acquisitions; 9 - 12 per subject
- 4 facial expressions
- 578 occluded acquisitions
- 6+ occluding objects
# UMB-DB Database

## Face Detection Preliminary Results

<table>
<thead>
<tr>
<th>Total Faces</th>
<th>False Positives</th>
<th>False Negatives</th>
<th>Detected Faces</th>
</tr>
</thead>
<tbody>
<tr>
<td>1473</td>
<td>56</td>
<td>52</td>
<td>1421 (96.5%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Acquisition type</th>
<th>Number of faces</th>
<th>Detected Faces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutrals</td>
<td>441</td>
<td>437 (99.1%)</td>
</tr>
<tr>
<td>Non-neutral</td>
<td>442</td>
<td>431 (97.5%)</td>
</tr>
<tr>
<td>Occluded</td>
<td>578</td>
<td>553 (95.7%)</td>
</tr>
<tr>
<td>Scarf</td>
<td>151</td>
<td>141 (93.4%)</td>
</tr>
<tr>
<td>Glasses</td>
<td>75</td>
<td>71 (94.7%)</td>
</tr>
<tr>
<td>Hair</td>
<td>33</td>
<td>30 (90.9%)</td>
</tr>
<tr>
<td>Hand</td>
<td>165</td>
<td>150 (90.9%)</td>
</tr>
<tr>
<td>Hat</td>
<td>183</td>
<td>179 (97.8%)</td>
</tr>
<tr>
<td>Misc</td>
<td>28</td>
<td>26 (92.85%)</td>
</tr>
</tbody>
</table>
# UMB-DB Database

## Face Recognition Preliminary Results

<table>
<thead>
<tr>
<th>Normalization</th>
<th>Test set</th>
<th>EER</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual</td>
<td>All cases</td>
<td>0.186</td>
<td>71.0%</td>
</tr>
<tr>
<td>Automatic</td>
<td>All cases</td>
<td>0.195</td>
<td>69.6%</td>
</tr>
<tr>
<td>Automatic</td>
<td>Neutral</td>
<td>0.019</td>
<td>98.0%</td>
</tr>
<tr>
<td>Automatic</td>
<td>Occlusions</td>
<td>0.238</td>
<td>56.5%</td>
</tr>
</tbody>
</table>

![Graph](image-url)
Conclusions

- Main contributions:
  - a fully automatic occlusion tolerant 3D face recognition pipeline
  - occlusion tolerant face detector
  - occlusion detector
  - face restoration module
Conclusions

▶ Main contributions:
  ▶ a fully automatic occlusion tolerant 3D face recognition pipeline
  ▶ occlusion tolerant face detector
  ▶ occlusion detector
  ▶ face restoration module

▶ Considerations:
  ▶ there is enough information to recognize faces even with 40% of the occluded area
  ▶ face detection and normalization is crucial
Conclusions

- Main contributions:
  - a fully automatic occlusion tolerant 3D face recognition pipeline
  - occlusion tolerant face detector
  - occlusion detector
  - face restoration module

- Considerations:
  - there is enough information to recognize faces even with 40% of the occluded area
  - face detection and normalization is crucial

- Work in progress & Future work
  - use of other surface features and 2D data in order to improve occlusion detection, normalization and restoration
  - embed face detection, normalization, occlusion detection and face restoration in an iterative schema
  - explore facial expressions
  - artificial generation of datasets with facial expressions and occlusions
Facile
A system for artificial face datasets generation

Given a face 3D textured model with neutral facial expression, the current version of facile is able to:

- simulate facial expressions through a physical simulator based on a mass-spring system and synthetic muscles.
- render images of the morphed faces from different viewpoints and with different lighting conditions

Features:

- facial expressions are defined by the user on a mean face specifying muscle contractions
- facial expressions are automatically transferred on novel faces
- a dense annotation of facial feature points is automatically generated
- ready to be used with synthetic faces generated through morphable models or morphing techniques
References

Thresholds

- Histogram of the distribution of the pixel by pixel difference between training faces and the mean face