Topology Preserving Graph Matching for Partial Face Recognition

Yueqi Duan, Jiwen Lu, Jianjiang Feng, Jie Zhou
Tsinghua University, China
Face Recognition

- Access control
- Surveillance
- Easy people tagging
Typical Face Recognition System

- Face detection → Face alignment → Face representation → Face classification
Partial Faces Exist in the Wild

- Under crowded scenes:
- Occluded by objects:
Challenges
Challenges

- **Unreliable face alignment**
  - Most face alignment approaches require landmark detection
  - Missing landmarks in partial faces

- **Less discriminative description**
  - Different facial parts of the same person → Large intra-class distance
  - Description of the occluded objects → Small inter-class distance
Challenges

The LFW dataset

- HDLBP: 84.08%
- VGG-16: 97.27%

The partial LFW dataset

- HDLBP: 49.32%
- VGG-16: 71.27%

Partial faces deserve more attention!
Possible Solutions

- Only describe the common facial parts

- Occlusion removal?
  - Difficult to detect occlusions from an unaligned face accurately
  - Description of different facial parts for the same person

- Component-based methods?
Component-Based Methods \[1\]

Keypoint-Based Methods \cite{2,3}


\cite{3} Renliang Weng, Jiwen Lu, and Yap-Peng Tan, Robust Point Set Matching for Partial Face Recogniton, TIP, vol. 25, no. 3, pp. 1163-1176, 2016.
Motivation

- Existing local keypoint-based approaches rely heavily on the descriptors, ignoring the topological structural information.

- The structural information of facial parts is relatively stable, which would enhance the robustness of keypoint matching.
Flowchart

- Feature Extraction
  - SIFT keypoint detector and SiftSurfSILBP descriptor

- Keypoint Filter
  - Lowe’s matching algorithm to remove obvious outliers
  - Lowe’s matching relies on descriptors, which fails to exploit the geometric information

- Topology Preserving Graph Matching
  - Delaunay triangulation to construct the graph
  - Estimate a non-rigid transformation from the probe image to the gallery image

- Face Matching
Estimate a non-rigid transformation to match the graphs

Objective function:
- Textural cost
- Node-wise matching cost
- Edge-wise matching cost

Outlier removal
Face Matching

- We compute the distance between probe and gallery faces as follows:

\[
    d = \frac{\bar{d}}{\sum_{i,j} X_{ij}} = \frac{J_{\text{min}}}{(\sum_{i,j} X_{ij})^2} = \frac{K_t + \lambda_p K_p + \lambda_q K_q}{(\sum_{i,j} X_{ij})^2}
\]

- In proportion to the average loss
- Inverse proportion to the number of matching pairs
Experimental Results

- **LFW**
  - 13233 labeled faces of 5749 subjects
  - Random transformation

- **PubFig**
  - 58797 images of 200 people
  - Random transformation

- **AR**
  - 126 identities with 70 males and 56 females
  - 13 facial images for an identity in a session:
    - 4 with different expressions
    - 3 under various illuminations
    - 3 wearing sunglasses
    - 3 wearing scarves
Evaluation on LFW and PubFig

- The partial LFW dataset
- The partial PubFig dataset

<table>
<thead>
<tr>
<th>Method</th>
<th>VR ± SE</th>
<th>rank = 1</th>
<th>rank = 10</th>
<th>rank = 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDLBP</td>
<td>49.32 ± 1.09</td>
<td>25.00</td>
<td>49.29</td>
<td>57.86</td>
</tr>
<tr>
<td>CNN</td>
<td>71.27 ± 1.38</td>
<td>28.36</td>
<td>51.93</td>
<td>62.29</td>
</tr>
<tr>
<td>CPD-SiftSurfSILBP</td>
<td>61.62 ± 1.19</td>
<td>27.86</td>
<td>52.86</td>
<td>64.29</td>
</tr>
<tr>
<td>MKD-SRC-GTP</td>
<td>68.18 ± 1.77</td>
<td>38.57</td>
<td>62.14</td>
<td>72.14</td>
</tr>
<tr>
<td>MLERPM-SiftSurf</td>
<td>65.55 ± 1.53</td>
<td>37.14</td>
<td>64.29</td>
<td>72.86</td>
</tr>
<tr>
<td>MLERPM-SiftSurfLBP</td>
<td>67.22 ± 1.83</td>
<td>42.86</td>
<td>65.00</td>
<td>74.29</td>
</tr>
<tr>
<td>LAIRPM-SiftSurf</td>
<td>70.40 ± 1.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAIRPM-SiftSurfSILBP</td>
<td>70.73 ± 1.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPSM-SiftSurf</td>
<td>70.81 ± 1.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPSM-SiftSurfSILBP</td>
<td>71.65 ± 1.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPGM</td>
<td>73.48 ± 1.12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Evaluation on AR

- The AR dataset

<table>
<thead>
<tr>
<th>Method</th>
<th>S1-G</th>
<th>S1-S</th>
<th>S2-G</th>
<th>S2-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPD</td>
<td>71.00</td>
<td>75.67</td>
<td>49.33</td>
<td>61.00</td>
</tr>
<tr>
<td>MLERPM</td>
<td>75.00</td>
<td>78.33</td>
<td>53.33</td>
<td>66.67</td>
</tr>
<tr>
<td>LAIRPM</td>
<td>87.33</td>
<td>88.33</td>
<td>56.33</td>
<td>81.33</td>
</tr>
<tr>
<td>MKD-SRC-GTP</td>
<td>82.33</td>
<td>83.33</td>
<td>57.67</td>
<td>76.33</td>
</tr>
<tr>
<td>RPSM</td>
<td>88.67</td>
<td>90.33</td>
<td>63.67</td>
<td>85.67</td>
</tr>
<tr>
<td>TPGM</td>
<td>89.33</td>
<td>91.00</td>
<td>65.00</td>
<td>86.67</td>
</tr>
</tbody>
</table>
Future Works

- **The keypoint-based approach**
  - Exploit higher order structural information for the graph
  - Deep graph matching approaches to learn reliable transformation
  - Usage of facial structure as strong prior knowledge

- **Learning alignment-free local facial descriptor**

- **Partial face alignment**
Thanks!