Robust Face Recognition under Varying Illumination and Occlusion Considering Structured Sparsity

Xingjie Wei, Chang-Tsun Li and Yongjian Hu

Department of Computer Science, University of Warwick

x.wei@warwick.ac.uk

http://warwick.ac.uk/xwei
Face

- People love faces!
  - Biological nature
  - Sensitive to the face pattern

A house with a Hitler face
Face Recognition

• Uncontrolled conditions: large changes in pose, illumination, expression and occlusion, aging... Still challenging
Motivation

• Face recognition in real-world environments often has to confront with uncontrolled and uncooperative conditions
  – illumination changes, occlusion
• Uncontrolled variations are usually coupled
• Less work focuses on simultaneously handling them
Our Method

• Our work deals with the illumination changes and occlusion *simultaneously* considering *structured sparsity*

  represents a test image using the minimal number of *clusters*

Sparse Representation

*flat sparsity*

represents a test image using minimal number of training images from *all classes*
Our Method

• Our work deals with the illumination changes and occlusion \textit{simultaneously} considering \textit{structured sparsity} aided with:
  – \textbf{Structural occlusion dictionary}: better modelling contiguous occlusion

  contiguous occlusion also forms a \textit{cluster} structure
Our Method

Our work deals with the illumination changes and occlusion simultaneously considering structured sparsity aided with:

- **Structural occlusion dictionary**: better modelling contiguous occlusion
- **WLD feature**: robust to illumination changes, remove shadows

Inspired by the psychophysical *Weber’s Law*
Sparse Representation

- Models a test image as a \textit{linear combination} of training images
  - Using minimal number of training images

\[
\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_1 \quad \text{subject to} \quad y = X \alpha
\]
Sparse Representation

- Involves training images from all classes
  - Optimal for reconstruction but not necessary for classification

Using the same number of base vectors
Our Method

• Structured Sparsity
  – Each class form a cluster

\[ X = \begin{bmatrix} x_1, \ldots, x_d, \ldots, x_{n-d+1}, \ldots, x_n \end{bmatrix} \]

\[ X[1] \quad X[s] \]

\[ \alpha = \begin{bmatrix} \alpha_1, \ldots, \alpha_d, \ldots, \alpha_{n-d+1}, \ldots, \alpha_n \end{bmatrix}^T \]

\[ \alpha[1] \quad \alpha[s] \]

cluster structure

\[ y = X \alpha \]
Our Method

- Structured Sparsity
  - Represents a test image using the minimum number of clusters

\[ \hat{\alpha} = \text{arg min}_{\alpha} \| \alpha \|_{2,1} \]

\[ = \text{arg min}_{\alpha} \sum_{i=1}^{s} \sqrt{\sum_{j=1}^{d} \alpha_i^2[j]} \]

subject to \( y = X \alpha \)
Sparse Representation

- Occlusion modelling: identity matrix $I \in \mathbb{R}^{m \times m}$

- Limitation: $I$ is able to represent any image of size $m$

- $\text{size: } m = X \times I \times \alpha$

- $\alpha_e$ sparse
Our method

- Contiguous occlusion: the nonzeros entries are likely to be spatially continuous, are aligned to clusters.

size: $83 \times 60 = 4980$

(index of occlusion base vectors)
Our method

• Structural occlusion dictionary
  – uses the *cluster occlusion dictionary* to replace the *identity matrix* $I$
Our Method

• Extreme illumination + occlusion:
  – coupled occlusion takes up a large ratio of the image
  – not “sparse” error
Our Method

• A different view: extract relevant **features** that reduce the difference

• Using WLD feature
  ✓ Maintain most salient facial features
  ✓ Insensitive to illumination changes
  ✓ Can correct shadow effects

\[
WLD(p) = \arctan\left( \sum_{i=1}^{l} \frac{p_i - p}{p} \right)
\]

Illustrative Example

Reference image  Estimated occlusion

Test image  Reconstruction

Reference image  Estimated occlusion

Test image  Reconstruction

Sparse coefficients

class 1

Residuals

Residuals
Experiments

• Synthetic Occlusion with Extreme Illumination
  – Extended Yale B database
  – Occlusion levels: 0% ~ 50% of the image
Experiments

• Synthetic Occlusion with Extreme Illumination
  – using only the raw pixel intensity as feature

<table>
<thead>
<tr>
<th>Occlusion</th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR-P[15]</td>
<td>100</td>
<td>100</td>
<td>99.8</td>
<td>98.5</td>
<td>90.3</td>
<td>65.3</td>
</tr>
<tr>
<td>CRC-RLS[17]</td>
<td>100</td>
<td>100</td>
<td>95.8</td>
<td>85.7</td>
<td>72.8</td>
<td>59.2</td>
</tr>
<tr>
<td>R-CRC[17]</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>97.1</td>
<td>92.3</td>
<td>82.3</td>
</tr>
<tr>
<td>Proposed SSR-P</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>97.8</td>
<td>85.4</td>
</tr>
</tbody>
</table>

Experiments

• Synthetic Occlusion with Extreme Illumination
  – using WLD feature

<table>
<thead>
<tr>
<th></th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subset 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR-P[15]</td>
<td>86.3</td>
<td>78.5</td>
<td>70.0</td>
<td>53.2</td>
<td>36.7</td>
<td>28.1</td>
</tr>
<tr>
<td>Proposed SSR-P</td>
<td>97.2</td>
<td>93.4</td>
<td>84.8</td>
<td>68.4</td>
<td>53.4</td>
<td>39.9</td>
</tr>
<tr>
<td>SR-G[16]</td>
<td>95.5</td>
<td>88.8</td>
<td>84.2</td>
<td>76.4</td>
<td>66.5</td>
<td>54.7</td>
</tr>
<tr>
<td>SR-W</td>
<td>99.4</td>
<td>99.6</td>
<td>99.4</td>
<td>99.1</td>
<td>99.1</td>
<td>96.6</td>
</tr>
<tr>
<td>Proposed SSR-W</td>
<td>99.6</td>
<td>99.8</td>
<td>99.4</td>
<td>99.4</td>
<td>99.6</td>
<td>98.1</td>
</tr>
<tr>
<td>Subset 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR-P[15]</td>
<td>37.5</td>
<td>26.9</td>
<td>14.3</td>
<td>9.0</td>
<td>7.9</td>
<td>7.3</td>
</tr>
<tr>
<td>Proposed SSR-P</td>
<td>42.6</td>
<td>31.6</td>
<td>23.4</td>
<td>15.3</td>
<td>11.5</td>
<td>10.9</td>
</tr>
<tr>
<td>SR-G[16]</td>
<td>44.2</td>
<td>31.7</td>
<td>32.0</td>
<td>23.8</td>
<td>21.5</td>
<td>17.5</td>
</tr>
<tr>
<td>SR-W</td>
<td>98.0</td>
<td>97.5</td>
<td>96.9</td>
<td>96.9</td>
<td>91.9</td>
<td>83.0</td>
</tr>
<tr>
<td>Proposed SSR-W</td>
<td>98.3</td>
<td>98.0</td>
<td>97.3</td>
<td>95.8</td>
<td>95.4</td>
<td>88.6</td>
</tr>
</tbody>
</table>

Experiments

• Synthetic Occlusion with Extreme Illumination
  – using WLD feature

Experiments

- Disguise with Non-uniform Illumination
  - The AR Database
  - Real occlusion, 2 sessions

Training set  Testing set
Experiments

- Disguise with Non-uniform Illumination

### TABLE III
Recognition rates (%) on the AR database

<table>
<thead>
<tr>
<th></th>
<th>Sunglasses</th>
<th>Scarves</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR-P[15]</td>
<td>42.5</td>
<td>29.8</td>
</tr>
<tr>
<td>Proposed SSR-P</td>
<td>43.5</td>
<td>31.8</td>
</tr>
<tr>
<td>SR-G[16]</td>
<td>74.8</td>
<td>76.0</td>
</tr>
<tr>
<td>SR-W</td>
<td>85.0</td>
<td>89.5</td>
</tr>
<tr>
<td>Proposed SSR-W</td>
<td>87.5</td>
<td>92.0</td>
</tr>
</tbody>
</table>
Thank you

• Questions?

• Xingjie Wei
  • x.wei@warwick.ac.uk
  • http://warwick.ac.uk/xwei
  • Department of Computer Science, University of Warwick