Perception Preserving Projections

Saining Xie ¹, Jiashi Feng ², Shuicheng Yan ², Hongtao Lu ¹

¹ Shanghai Jiao Tong University
² National University of Singapore

Presenter: Chao Zhang
Outline

1. Motivations
2. Problem Formulation
3. Optimization
4. Experimental Results
5. Conclusions
Motivations

**Dimensionality Reduction in Computer Vision**

- An image is a point in a high dimensional space.
  - An $n \times m$ image is a point in $\mathbb{R}^{n \times m}$.
Motivations (cont.)

Eg. 2D Gabor Filter

Eg. Gradient Feature
Problem Formulation: PCA Revisiting

\[ \min_{U^T U = I_r} \mathcal{L}(U) = \| X - U U^T X \|_F^2. \]

**PCA Formulation: Minimize Reconstruction Loss**

PCA does NOT take the feature extraction step into consideration, may incur severe information loss in Specific perception systems.
Problem Formulation: \( P_{\text{erception}} P_{\text{reserving}} P_{\text{rojections}} \)

Many Feature Extractors can be expressed as Linear Operators:

1. Convolution with Linear Filters
   \( f: \mathcal{P}(x) = P_f x = f * x. \)

2. Pixel-wise “masking”
   \( P x, \) where \( P \) is a \( d \times d \) diagonal matrix

3. Sum of Filters
   \( \sum_{k=1}^{K} P_{f_k} x \)

Objective Function of Perception Preserving Projections:

\[
\begin{align*}
\min_{U^T U = I_r} \quad & \mathcal{L}(U) = \| \mathcal{P}'(X) - \mathcal{P}'(U U^T X) \|_F^2, \\
& P' = [(1 - \alpha) P, \alpha I_d]
\end{align*}
\]
Straight-forward solution:
Gradient Descent on Stiefel Manifolds

state-of-the-art off-the-shelf solver:
Cayley Transformation

\[ Q = (I - A)(I + A)^{-1} \]

Produce an orthogonal matrix Q.

(Inspired by Robust PCA work)

We can relax the orthogonal constraints to a rank minimization problem:

$$\min_W \| \mathcal{P}(X) - \mathcal{P}(WX) \|_F^2, \text{ s.t. } \text{rank}(W) \leq r,$$

Since Nuclear Norm is the convex envelope of the rank function.

$$\min_W \|W\|_* + \lambda \|E\|_F^2, \text{ s.t. } \mathcal{P}(X) - \mathcal{P}(WX) = E,$$

Can be solved efficiently by: Alternating Direction Method (ADM)

**Problem:** $\mathcal{P}(WX) + W = C$ ($C$ is a constant matrix)

Encounter a Sylvester equation above in the sub-problem w.r.t variable $W$

Infeasible because $O(n^6)$ time complexity!

Optimization (cont.)

Solution: Linearize the objective function at the point of $W_k$

$$
\mathcal{L}(W, W_k) = \|W\|_* + \langle Y, -\mathcal{P}(W_k X) \rangle + \mu \left\langle \mathcal{P}^* \left( \mathcal{P}(X) - \mathcal{P}(W_k X) - E \right) X^T, W - W_k \right\rangle + \frac{\mu \eta}{2} \|W - W_k\|_F^2.
$$

$$
\mathcal{L}(W, Y, W_k) = \|W\|_* + \frac{\mu \eta}{2} \|W - M_k\|_F^2,
$$

$$
M_k = W_k - \mathcal{P}^* \left( \mathcal{P}(X) - \mathcal{P}(W_k X) - E \right) X^T / \eta + \mathcal{P}^* Y X^T / \mu \eta
$$

**CLOSED FORM** solution (Soft-thresholding):

$$
W_{k+1} = US \frac{1}{\mu \eta} (\Sigma) V^T, \quad S_\varepsilon [x] = \text{sgn}(x) \max(|x| - \varepsilon, 0)
$$

**Speed-up Tricks:**

1. **adaptive penalty strategy.** $\mu_{k+1} = \begin{cases} 
\rho_0 \mu_k, & \text{if } \mu_k \max(\sqrt{\eta} \varepsilon_W, \varepsilon_E) / \|\mathcal{P}(X)\| \leq \varepsilon_2, \\
\mu_k, & \text{otherwise}.
\end{cases}$

2. **PROPACK** for partial SVD

3. **Skinny SVD** to avoid full matrix multiplication

Experimental Results: Synthetic Data

randomly generate 100 gray level patches with Vertical lines and Horizontal outliers

Figure 2: Feature deviation descent curve on the synthetic data along with iterations.

(a) Original images
(b) PCA reconstruction
(c) PPP reconstruction based on gradient descent over Stiefel manifold
(d) PPP reconstruction based on low rank optimization

Figure 4: Examples of the synthetic data and reconstruction results from different methods.
Experimental Results:
Gradient Preserving and Gabor Feature Preserving

Eg. 2D Gabor Filter

Eg. Gradient Feature
Experimental Results: Gradient and Gabor Feature Preserving (cont.)

Figure: Results on FRGC dataset (with multiple classifiers)

Table: Results on Extended Yale-B: Gabor Feature (above) / Gradient Feature (bottom)

<table>
<thead>
<tr>
<th>Dim</th>
<th>PCA+kNN</th>
<th>PPP-S+kNN</th>
<th>PPP-L+kNN</th>
<th>PCA+LDA</th>
<th>PPP-S+LDA</th>
<th>PPP-L+LDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5.82 ± 1.64</td>
<td>43.51 ± 2.63</td>
<td>42.30 ± 2.17</td>
<td>10.40 ± 1.53</td>
<td>44.30 ± 1.17</td>
<td>44.41 ± 2.68</td>
</tr>
<tr>
<td>10</td>
<td>33.56 ± 1.24</td>
<td>68.34 ± 2.81</td>
<td>70.58 ± 2.29</td>
<td>49.89 ± 2.17</td>
<td>69.57 ± 2.60</td>
<td>71.14 ± 2.73</td>
</tr>
<tr>
<td>15</td>
<td>48.66 ± 1.95</td>
<td>77.63 ± 1.25</td>
<td>76.52 ± 1.09</td>
<td>70.13 ± 1.07</td>
<td>78.75 ± 2.73</td>
<td>77.40 ± 1.66</td>
</tr>
<tr>
<td>20</td>
<td>58.39 ± 2.49</td>
<td>81.66 ± 2.83</td>
<td>80.21 ± 2.10</td>
<td>80.54 ± 2.79</td>
<td>83.00 ± 1.69</td>
<td>83.54 ± 2.96</td>
</tr>
<tr>
<td>30</td>
<td>71.36 ± 2.07</td>
<td>84.23 ± 2.26</td>
<td>84.34 ± 3.03</td>
<td>85.12 ± 1.35</td>
<td>85.12 ± 2.85</td>
<td>86.92 ± 1.09</td>
</tr>
<tr>
<td>50</td>
<td>79.08 ± 2.27</td>
<td>82.66 ± 1.20</td>
<td>86.13 ± 0.81</td>
<td>89.15 ± 2.44</td>
<td>86.02 ± 1.75</td>
<td>89.58 ± 1.16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dim</th>
<th>PCA+kNN</th>
<th>PPP-S+kNN</th>
<th>PPP-L+kNN</th>
<th>PCA+LDA</th>
<th>PPP-S+LDA</th>
<th>PPP-L+LDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>4.47 ± 2.92</td>
<td>40.04 ± 3.38</td>
<td>40.16 ± 0.69</td>
<td>17.34 ± 1.93</td>
<td>36.13 ± 3.17</td>
<td>38.59 ± 1.77</td>
</tr>
<tr>
<td>10</td>
<td>41.05 ± 2.93</td>
<td>64.21 ± 1.87</td>
<td>69.46 ± 1.64</td>
<td>62.53 ± 2.04</td>
<td>74.83 ± 3.28</td>
<td>72.71 ± 2.62</td>
</tr>
<tr>
<td>15</td>
<td>57.94 ± 1.32</td>
<td>75.50 ± 4.40</td>
<td>79.08 ± 2.79</td>
<td>82.21 ± 2.22</td>
<td>84.56 ± 1.15</td>
<td>87.81 ± 1.68</td>
</tr>
<tr>
<td>20</td>
<td>70.36 ± 2.94</td>
<td>78.97 ± 0.74</td>
<td>81.77 ± 1.25</td>
<td>87.92 ± 1.04</td>
<td>86.80 ± 2.21</td>
<td>91.05 ± 1.38</td>
</tr>
<tr>
<td>30</td>
<td>79.19 ± 2.91</td>
<td>79.42 ± 2.36</td>
<td>82.89 ± 1.86</td>
<td>90.72 ± 1.28</td>
<td>90.04 ± 2.06</td>
<td>91.39 ± 1.77</td>
</tr>
<tr>
<td>50</td>
<td>83.00 ± 1.97</td>
<td>82.44 ± 2.13</td>
<td>83.11 ± 0.95</td>
<td>91.50 ± 2.26</td>
<td>92.06 ± 2.76</td>
<td>92.17 ± 2.45</td>
</tr>
</tbody>
</table>
Experimental Results:
Gradient and Gabor Feature Preserving (cont.)

when alpha is small, the reconstructed images of different persons look quite similar. However, by extracting the Gabor features, these images can be distinguished correctly.
Conclusion

- Proposed the **perception preserving projection** method, which is able to **preserve** the important information for specific perception system in the image projection process.

- Explicitly **embed** the feature preserving metric provided by a certain type of perception systems into the loss function.

- The results suggest **PPP** can better preserve the **discriminative** and **domain-specific** feature information.

😄 the current framework is quite *naïve*.

**Future work:**
- How can PPP be regarded as an implementation of **unsupervised joint embedding** of different domains?

- Extend the range of its application scenarios.