

CS123a Statistical Machine Learning (Spring 2013): Homework Assignment #5

Chuan Wang

March 19, 2013

(1) In the Fisherface approach, how to deal with the case when the within-class scatter matrix is singular?

Solution: To deal with the singular problem of within-class scatter matrix, the paper first projects the image set(data set) to a lower dimensional space by using PCA to reduce the dimension of the feature space to $N - c$, which is the upper bound of the rank of within-class scatter matrix. In this step, the transformation W_{pca} of PCA is given by:

$$W_{pca} = \arg \max_W |W^\top S_T W| \quad (1)$$

Then apply the standard Fisher's Linear Discriminant to reduce the dimension to $c - 1$. And the transformation W_{fld} of FLD is given by:

$$W_{fld} = \arg \max_W \frac{|W^\top W_{pca}^\top S_B W_{pca} W|}{|W^\top W_{pca}^\top S_W W_{pca} W|} \quad (2)$$

And the optimal transformation W is:

$$W_{opt}^\top = W_{fld}^\top W_{pca}^\top \quad (3)$$

(2) In one of the Eigenface approaches, why the first three principal components were removed? Why discarded the first 3 principal components instead of the first 2 or 4 principal components?

Solution: From Fig.8 in the paper, it is shown that dropping the first three principal components appears to improve the performance. According to the

paper's opinion, the first three principal components are primarily due to lighting variation and removing them can somehow improve the recognition rates. Also, from empirical experiments, it is suggested that the first three principal components explained large percentage of the variation

(3) Under what assumptions, the Eigenface approach is equivalent to the correlation method?

Solution: Assuming that the classification is performed using a nearest neighbor classifier in the reduced feature space and the number of Eigenvalues is chosen to be the number of images in the training set. The Eigenface method is equivalent to the correlation method.

(4) If the training set is representative (i.e., equally and densely covers all conditions) and the computational resource is not a problem, will the correlation approach outperform the Eigenface and Fisherface approaches?

Solution: It has been proved in the paper that the Eigenface method is equivalent to correlation when the number of Eigenfaces equals the size of the training set. So the Eigenfaces method cannot outperform the correlation method. But for the Fisherface method, since the training set is representative, the correlation method will include all the information. And the fisherface method is losing some of the information by using projection. So ideally the correlation method will outperform fisherface.