

1 Names and Faces in the News

In their paper, the authors deal with a clustering problem of facial images, which are compiled from the news. The contribution of the paper is that for each image, to extract names from the captions and to match the names to the faces appeared in the image. The authors note that the problem is not trivial because of the complexity resulted from both ambiguous names extracted from the captions and the fact that facial images tend to vary widely.

2 Approach

The authors first compiled a vast dataset from about a half million news from the Internet. The dataset contains over 45000 images and corresponding captions which contains some names. The images are rectified by using SVM-based facial feature detectors to a canonical form. During the rectification step, some of the images are removed from the dataset according to their rectification score. The remaining ones then are represented in a vectorized form to conduct discriminant analysis. The discriminant analysis is conducted to project the images to a hyper-space in which the dimensionality of the images are reduced. The authors note that this step is an essential part of the proposed algorithm due to the vast size of the dataset. The discriminant analysis is consisted of an approximated form Linear Discriminant Analysis and kernel-based Principal Component Analysis. After projecting the dataset to another space which is smaller in dimension, a small subset of the dataset containing only one common extracted name is used to form a new discriminant space, which is used (to my best understanding) as the initial guess in the clustering step. The dataset is then “cleaned-up” by a variant of k-means algorithm with the initial clusters. The initial guess improves the clustering performance due to less noisy signal the initial guess carry, given the fact that it is a subset of the dataset containing only one common name for each image. Next, two more rejection steps are utilized; the first some clusters are removed from the dataset according to number of the members and the second is a statistical measure of validity of class membership. Lastly, the paper uses a merging step in which visually similar faces corresponding to different names are proposed to be merged. The similarity measure employed is the distance amongst clusters in the discriminant space.

3 Conclusions

The proposed method contains multi-level rejection step in which some subset of the dataset is removed from the dataset. It is expected to see that filtering out the noisy data in these rejection steps should have significant impact on the overall performance. However, the Figure-4 implies that these steps have less significant impact than the cluster structure (The cost of correcting labels by themselves is higher than clustering cost). In other words, the cluster structure is the most significant aspect of the proposed method. Another subject worth to mention is that it is known that kmeans-like clustering algorithms are prone to converge to the local optima. It would be better if the authors would provide more results on clustering process or if they prevent this phenomenon happening by doing some sort of process.