Face Recognition from a Single Training Sample per Person (SSPP)

Pratyaksha , Kamil Mansuri

Abstract

Most of the literature on Face Recognition problem assume that a good number of training samples are available per subject. But in practical scenarios, official records like ID Card, passports etc, usually have only few images of each person. In our project we want to address this problem and perform experiments that demonstrate the failure of conventional PCA and LDA approaches in such cases.We implement the Discrimintive Multi-Manifold Analysis(DMMA) method for learning discriminative features by considering an image as a manifold of patches. We also design a method for increasing the number of data points by considering the elements of manifolds as independent samples. We compare the performance of each.

1. Technical Details

Method 1: Experiments show that using the SSPP training set with conventional supervised learning approaches does not perform well because the number of samples is much lesser than the number of features. With this in mind, we tried to divide each image into **"t"** patches of equal dimensions. So each image is now a manifold with the local patches as points in the manifold. We considered these patches(or d-dimensional vectors) as individual samples and assigned the corresponding subject name to each of these samples. We applied SVM classifier on this dataset.

Method 2: In the next experiment, instead of treating the points of manifolds as individual samples, we implemented a simple manifold-manifold matching method and tried to recognize faces as an unsupervised learning method by calculating manifold-manifold distance of the test subject with each of the train subject and assigning the label corresponding to the subject with minimum distance.

Method 3: Using the concepts from[1], we formulated another manifold-manifold matching method for learning unique projection matrices(W_i) for each subject such that in the projected space, the points (discriminative features) within one manifold are close to each other while those from different manifolds are far from each other. As a result, we have our maximization problem as given in (i) where J_1(across manifold distance) has to be maximized and J_2(within manifold distance) has to be minimized. After simplifying this objective function, the problem reduces to finding the eigen vectors correspoding to the top k eigen values of the matrix $H_1 - H_2$(Given in Appendix). The projection matrix W_i consists of these eigen vectors as columns. Once we have the projection matrices, we can project each manifold point x_{ir}. Similarly for a test image, we use above matrices to project each of it's patches and form manifolds containing the new discriminative feature spaces. We then solve (ii) to find the label for that test image.

2. Results

The results of Method 1 demonstrated that the formulation ignored the geometric information of the patches within a manifold and hence does not perform up to the mark. As a result we tried Method 2, which also had an issue that points(features) corresponding to the same facial parts across two different manifolds were more similar than different parts within the same manifold. Method 3 resolves this issue by solving the objective function as discussed earlier.

For training samples, we used one image each(with normal expression) of 15 different subjects from Yale Faces dataset. We used total of 165 images(having variety of expressions) of these 15 subjects as test samples.Train set accuracy for all the three methods was 100% while test accuracies were < 50 % for Method 1, 71.5% for Method 2 and 77.5% for Method 3. Values of k_1, k_2, σ, t, d and k were selected empirically based on test set performance.

3. Novel Contributions

[1] only talks about Method 3 discussed above. Method 1 and Method 2 were formulated as a part of our comparative study. Method 3 discussed in [1] used different dimensionalities for projected features which in turn increased the computational complexity of our experiment. Applying the projections as well as the reconstruction part seemed complex and expensive especially when the number of subjects and in turn the number of training samples(N) increases. So as a change, we used fixed dimensional projection matrices and simple manifold-manifold distance method for labelling the test set, which gave decent performance.

4. Reference :

Lu, Jiwen, Yap-Peng Tan, and Gang Wang. "Discriminative multimanifold analysis for face recognition from a single training sample per person." IEEE transactions on pattern analysis and machine intelligence 35.1 (2013): 39-51.

Data : http://vision.ucsd.edu/content/yale-face-database

5. Tools Used

Google Colab : For coding the equations
OpenCV : For preprocessing(Cropping) the images

6. Appendix

equation(i):

$max_{W_1,W_2,\cdots,W_N} J(W_1, W_2, W_3, \cdots, W_N)$

= $J_1(W_1, W_2, \cdots, W_N) - J_2(W_1, W_2, \cdots, W_N)$

= $\sum_{i=1}^{N}(\sum_{r=1}^{t}\sum_{p=1}^{k1}||W_i^T x_{ir} - W_i^T x_{irp}||^2 A_{irp}$
$-\sum_{i=1}^{N}(\sum_{r=1}^{t}\sum_{q=1}^{k2}||W_i^T x_{ir} - W_i^T x_{irp}||^2 B_{irp}$

$$A_{irp} = \begin{cases} \exp(-||x_{ir} - x_{irp}||^2/\sigma^2; & if x_{irp} \in N_{inter}^{k_1}(x_{ir}) \\ 0 & \text{Otherwise} \end{cases}$$

$$B_{irq} = \begin{cases} \exp(-||x_{ir} - x_{irq}||^2/\sigma^2; & if x_{irq} \in N_{intra}^{k_2}(x_{ir}) \\ 0 & \text{Otherwise} \end{cases}$$

$$H_1 \triangleq \sum_{r=1}^{t}\sum_{p=1}^{k_1}(x_{ir} - x_{irp})(x_{ir} - x_{irp})^T A_{irp}$$

$$H_2 \triangleq \sum_{r=1}^{t}\sum_{q=1}^{k_2}(x_{ir} - x_{irq})(x_{ir} - x_{irq})^T B_{irq}$$

equation(ii):

$$c = \underset{i}{\operatorname{argmin}}\, d(\mathcal{M}_{\mathcal{T}}, \mathcal{M}_i);$$

where i = 1,2,$\cdots, N$.

here x_{irp} are the pthk$_1$-nearest inter-manifold neighbors and x_{irq} are the qthk$_2$-nearest intramanifold neighbors of x_{ir}.
$N_{inter}^{k_1}(x_{ir})$ and $N_{intra}^{k_2}(x_{ir})$ are the k_1-inter-manifold neighbors and k_2-intra-manifold neighbors of x_{ir}