Final Report for Project 7: Discriminative K-SVD for Dictionary Learning in Face Recognition

Jianfeng Wu, Yu-Hsuan Chuang

***Abstract*—In a sparse-representation-based face recognition scheme, we expected to have a dictionary that has both good representational power and discrimination capability. The Discriminative K-SVD algorithm proposed in [1] achieves the above two goals by incorporating the classification error into the K-SVD algorithm instead of iteratively solving sub-problems to obtain global optimal.**

Keywords - Facial recognition, Sparse representation, K-SVD, Discriminative K-SVD.

I. Introduction

In recent years, many new technical methods have been exploited for face recognition in Computer Vision. Many Sparse representation based techniques, such as image denoising [7], image compression [3], [6] and SRC algorithm, have been performed well in facial recognition. These techniques usually involves sparse coding to approximate an input image by a sparse linear combination of samples from an over-complete dictionary. However, the basic way of forming the dictionary by employing the entire set of training samples may lead to a large size of dictionary, which is not good for sparse solver. Therefore, a lots of methods for dictionary leaning, which aims to learn a small dictionary including few atoms from huge amounts of training data have also been proposed. For example, the K-SVD algorithm efficiently learns an over-complete dictionary with a small size, but the drawback of this method is that it only focuses on representational power, but does not consider discriminative capability. To circumvent this problem, some other methods, such as Supervised Dictionary Learning [5] and LC-KSVD [4], use more sophisticated objective function in dictionary optimization to gain discriminative power for the dictionary. In the paper[1], the proposed Discriminative K-SVD algorithm also unifies the representation power and discriminate ability to train the dictionary and classifier simultaneously. It extends basic K-SVD algorithm by incorporating the classification error into K-SVD algorithm. And this algorithm has been proven effective and efficiency in image classification.

II. Problem Statement

In this project, we are supposed to finish an experiment by using the the proposed dictionary-learning method, Discriminative K-SVD (D-KSVD). D-KSVD method is based on K-SVD algorithm, which can allow the representative power of the dictionary and the discrimination of the linear classifier being considered by the same optimization procedure simultaneously. Besides, the paper also provides a classification method by using sparse coding. We evaluate this proposed method by using two databases, the first one is synthetic data which was randomly generated by ourselves and the second one is Extended YaleB face database.

The rest of this paper is arranged as follows. In Section 3, we first describe the detailed implementation of this algorithm, includes initializing dictionary, initializing classifier parameters, D-KSVD training algorithm and finally the classification algorithm for testing. In the Section 4, we will describe how we evaluate this proposed method using two databases. In the first experiment, we use synthetic data which was randomly generated by ourselves. In the second experiment, we use the Extended YaleB face database as training and testing data. Then we also report the analysis of the experiment results in this Section. Finally, we give a brief summary on D-KSVD methods, and expect some extensions in the future work.

III. Detailed Implementation

A. Initialize Dictionary using K-SVD Algorithm

To initialize the dictionary, we use the K-SVD toolbox which runs the K-SVD dictionary training algorithm on the specified set of signals, returning the trained dictionary D and the signal representation matrix α. For sparsity-based minimization, the optimization problem is given by

$$< D, \alpha >= \underset{D,\alpha}{\operatorname{argmin}} \|Y - D * \alpha\|_2 \ subject\ to\ \|\alpha\|_0 \leqslant T \quad (1)$$

where Y is the set of training signals, α_i is the i-th column of α, and the T is the target sparsity.

After implementing this algorithm, we can calculate the classifier parameters by using the initial dictionary and representation matrix. The detailed calculation steps were described at the following section.

B. Initialize Classifier Parameter

To gain the initial classifier parameter, we use the following equation,

$$W = (\alpha^T \alpha + \beta' * I)^(- 1) * \alpha * H^T \quad (2)$$

where H is a matrix that represents the label of the training images. Each column of H is a vector: h_i = [0,0,...,1...,0,0], where the position of non-zero element indicates the class, β is a scalar, and here we set to 1.

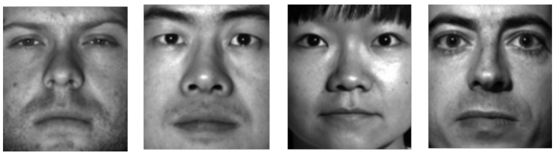

Fig. 1. Extended YaleB database face image

C. D-KSVD Training Algorithm

Based on the K-SVD algorithm, the paper added discrimination ability simultaneously and generated a new algorithm, Discriminative D-KSVD,

$$< D, W, \alpha >= \underset{D,W,\alpha}{\operatorname{argmin}} \| \begin{pmatrix} Y \\ \sqrt{\gamma} * H \end{pmatrix} - \begin{pmatrix} D \\ \sqrt{\gamma} * W \end{pmatrix} * \alpha \|_2 + \beta * \|W\|_2 \ subject\ to\ \|\alpha\|_0 \leqslant T, \tag{3}$$

where W is the parameter for the linear classifier $H = W * \alpha + b$, γ and β are scalars controlling the relative contribution of the corresponding terms and $\|W\|_2$ is the regularization penalty term.

D. Algorithm for Classification

With D-KSVD algorithm, we can get a dictionary D and a classifier W but when we test a new image, the dictionary doesn't support a sparse-coding based representation since D and W are simultaneously normalized. So the paper built a mapping method to normalize the dictionary and corresponding classifier based on the learning result. The desired dictionary D' and corresponding classifier W' can be computed as

$$\begin{aligned} D' &= \{d'_1, d'_2, \cdots, d'_k\} \\ &= \{\frac{d_1}{\|d_1\|_2}, \frac{d_2}{\|d_2\|_2}, \cdots, \frac{d_k}{\|d_k\|_2}\}, \\ W' &= \{w'_1, w'_2, \cdots, w'_k\} \\ &= \{\frac{w_1}{\|d_1\|_2}, \frac{w_2}{\|d_2\|_2}, \cdots, \frac{w_k}{\|d_k\|_2}\}, \end{aligned} \tag{4}$$

where d_i and w_i represent the i-th column of D and W. With this, we can compute the sparse coefficients for a given test image by solving the following sparse-coding problem,

$$< \alpha' >= \underset{\alpha'}{\operatorname{argmin}} \|y - D' * \alpha'\|_2 + \sigma * \|\alpha'\|_0 \tag{5}$$

After calculating the sparse coefficients, we could finally get the label by the equation,

$$l = w' * \alpha'. \tag{6}$$

Here l is a vector that indicated the label of the test image. The label is the index i where l_i is the largest one in all the element in vector l.

IV. Experiments and Analysis

A. Train and test on the Extended YaleB Database

The database we used is the same with paper, Extended YaleB Database [8], which includes 2414 face images belongs to 38 subjects. Each subjects contains roughly 64 images. All these given images were cropped and normalized into 192*168 pixels, Fig. III-C. Before training the models, we processed the face data by projecting the 192*168 pixels images into a R^{504} vector with the RandomFace method [2]. Here, for each image, we firstly normalized the matrix and then we timed the 192*168 matrix with a 504*162*148 randomly generated matrix and got a feature vector$\in R^{504}$.

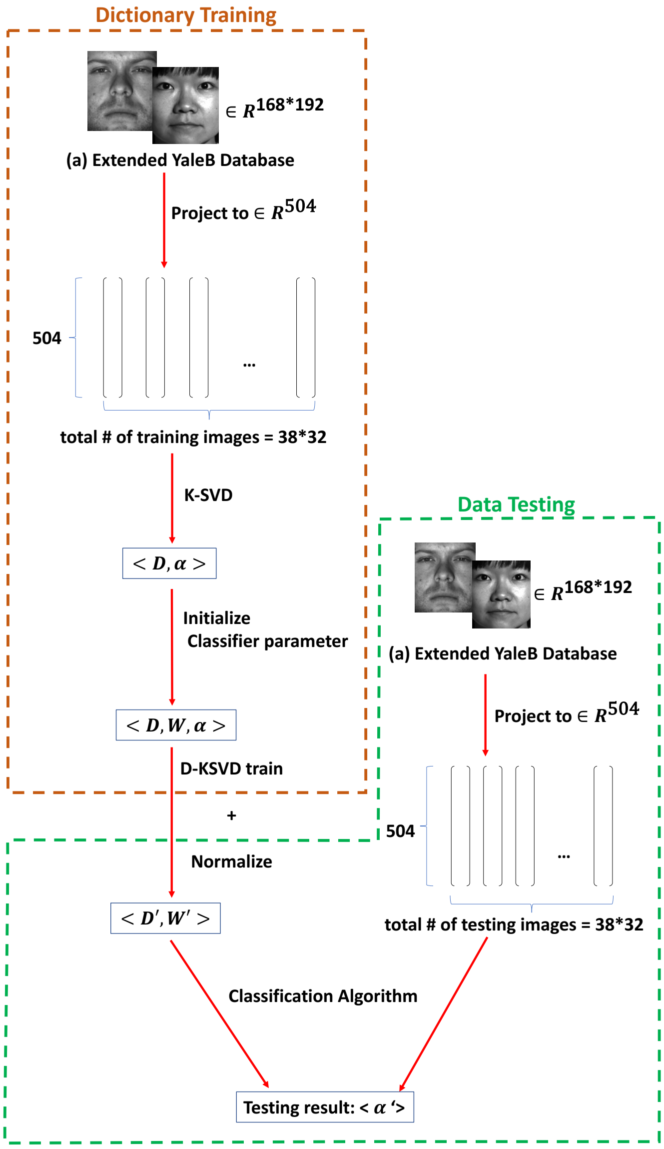

Fig. 2. Implementation flow chart

Then, we started to train the model. For the training dataset, we used half of the images, which means we randomly chose 32 image for each subjects. And the rest images were for testing. Thus the training dataset contained 1216 images, 38*32, and the test data included the rest 1198 images. Initially, we applied K-SVD algorithm, Eqn. 1 to train the dictionary D as

TABLE I
TEST ACCURACY AND TIME COST FOR EXPERIMENT B

Accuracy	Initializing time	Training time	Testing time
0.972454	217.40625(s)	825.796875(s)	170.234375(s)

well as representation matrix. With these, we could initialize the classifier parameter, Eqn. 2. Then, we started to train the model by Discriminative K-SVD, Eqn. 3. We combined the dictionary and classifier to a same matrix and set to 16 and the scalars γ and β, which control the relative contribution of the corresponding coefficient, are both set to 1. By solving the equation, we adopt OMP algorithm [9], which has been proved faster than other L1-optimization method. We set the sparsity prior in the learning model to 16. Finally, we acquired the trained dictionary and classifier. Before testing, we normalized the dictionary and classifier with Eqn. 4.

Finally, we tested the 1198 images with the dictionary and classifier, Eqn. 5. Here we still used the OMP method to solve the equation. Here, the sparsity prior was set to 64. And with the gained corresponding coefficient, we were able to calculate the identify information by Eqn. 6. The accuracy for the experiment is about 97.24%

We run all the experiment on Matlab 2017a under Window 10 system. The processer information is Inter(R) Core(TM) i7-6700 CPU @ 3.40GHz 3.41 GHz.

The time cost for each step is show in Table. I.

B. Train and test on Synthetic Dataset

In this experiment, we generated two clusters for X, the elements inside the same cluster are in the same class, and use this synthetic data as the training data Y. In this experiment, we generated two clusters which followed the Gaussian distribution. The variance is the same but the mean of them is different. For the training dataset, we generated 32 data for each subject. Also, we set 32 data for each subject in test data. Each data is also a vector$\in R^{504}$. The scalars γ and β, which control the relative contribution of the corresponding coefficient, are also both set to 1. And the sparsity prior is still 16 but the atom is only 2, one for each subject.

We think the accuracy for the algorithm depends on the distribution of the data. We assumed that the accuracy would show an increasing tendency with the distance increasing. To prove our idea, we generated different pairs of clusters, of which the distance between the means is different. The difference of the mean is 0.1, 0.2, 0.8, 1, 10 and 100.

We did the experiment for many time. One of the results was shown in Table. II. The result proved our thought that the accuracy will increase with the further distribution of two clusters. When the distance is greater than 1, the accuracy is great than 95%, always 100%. Even when the distance is smaller than the variance, such as 0.1 or 0.2, the test could get a high accuracy which is upto 60%. This prove the robust property of the Discriminate-KSVD algorithm.

For the first experiment, the condition would be more complicated since the difference between each image cannot be very significant. The algorithm is robust enough to distinguish the difference of these features.

TABLE II
ACCURACY OF VARIOUS DISTANCE OF MEAN

	Accuracy
m=0.1	0.546875
m=0.2	0.609375
m=0.8	0.968750
m=1	1.000000
m=10	1.000000
m=100	1.000000

V. CONCLUSION

The proposed D-KSVD algorithm realizes the learned overcomplete dictionary that has both good representational power and discrimination capability, it directly finds all the parameters (the dictionary and the classifier) simultaneously instead of iteratively solving sub-problems to gain global solutions. And the experiment shows the classification results with high accuracy. However, D-KSVD is still a two-step iterative method, and the convergence speed is heavily influenced by the initialization values. Thus, the possible future work includes finding a initialization method to improve speed performance of the algorithms, developing effective ways to incorporate the label term into the objective and to update the dictionary without retraining.

REFERENCES

[1] Aharon, M., Elad, and Bruckstein. "rmK-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation." IEEE Transactions on Signal Processing a Publication of the IEEE Signal Processing Society. 54, no. 11 (2006): 4311-322.

[2] Wright, Yang, Ganesh, Sastry, and Yi Ma. "Robust Face Recognition via Sparse Representation." Pattern Analysis and Machine Intelligence, IEEE Transactions on 31, no. 2 (2009): 210-27.

[3] Bryt, and Elad. "Compression of Facial Images Using the K-SVD Algorithm." Journal of Visual Communication and Image Representation 19, no. 4 (2008): 270-82.

[4] Jiang, Zhuolin, Zhe Lin, and Larry S. Davis. "Learning a discriminative dictionary for sparse coding via label consistent K-SVD." Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011.

[5] Mairal, Julien, Francis Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman. "Supervised Dictionary Learning." 2008.

[6] Bryt, O., and M. Elad. "Improving the K-svd Facial Image Compression Using a Linear Deblocking Method." Electrical and Electronics Engineers in Israel, 2008. IEEEI 2008. IEEE 25th Convention of, 2008, 533-37.

[7] Elad, and Aharon. "Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries." Image Processing, IEEE Transactions on 15, no. 12 (2006): 3736-745.

[8] Kuang-Chih Lee, Ho, and Kriegman. "Acquiring Linear Subspaces for Face Recognition under Variable Lighting." Pattern Analysis and Machine Intelligence, IEEE Transactions on 27, no. 5 (2005): 684-98.

[9] Chen, S., C.F.N. Cowan, and P.M. Grant. "Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks." Neural Networks, IEEE Transactions on 2, no. 2 (1991): 302-09.