Is sparse representation a good and profound explanation for face recognition?

1st Weikang Wang
dept. of electrical engineering
Columbia University
New York, USA
ww2461@columbia.edu

***Abstract*—Sparse representation tools for image recognition was raised up in nearly ten years before and soon became very popular for image classification especially face recognition, for its big success in accuracy and robustness with respect to errors like occlusion and disguise. But there are also some queries that disproved the role of sparse representation used in face recognition problems. So we want to ask whether sparse representation a good and profound explanation for face recognition problems.**

***Index Terms*—Face Recognition, SRC, Sparse Representation**

I. Introduction

Compressing Sensing, or sparse representation, is a very powerful tool for signal processing. About ten years before, in [3], J. Wright et. al first applied sparse representation into face recognition areas and raised up a new algorithm, named SRC algorithm. This algorithm made a huge success dealing with aligned face images and was robust to errors like occlusion and disguise. The SRC algorithm used a hypothesis that the testing image lies in the subspace spanned by its training images of the same identity.

However, though the SRC algorithm has made a big success and was appilled in many areas [2], [4], there are also some people have queries about the sparsity used in face recognition problems and the role of sparsity in the successful SRC algorithm. Shi et. al in [1] states that the low dimensional structure of the SRC algorithm is not true and raises up a new algorithm based on l2-norm minimization. Zhang et. al in [5] studies the working mechanism of the SRC algorithm and states that it is the collaborative representation rather than the sparsity used in SRC algorithm that makes it successful and raises up a new algorithm based on l2-norm minimization, too.

In this paper, we will take a deep look at the analysis of the queries about the role of sparse representation used in face recognition and about the SRC algorithm. We will show that whether this queries are true or not and whether sparse representation a good and profound explanation for face recognition problems, or equivalently, whether SRC algorithm reveals the essence of face recognition problems.

II. Technical Approach

In the paper, we use MATLAB to do experiments on the SRC algorithm and to test the queries raised in the preceding papers.

In chapter III, we first show and test the effectiveness of the SRC algorithm. In chapter IV, we analyze the two main queries of the SRC algorithm and the role of sparsity used in face recognition problems. In chapter V, we derive the conclusion whether sparse representation a good and profound explanation for face recognition problems.

III. Sparse Representation-based Algorithm

The Sparse Representataion-based Algorithm is the generalization of the Nearest Neighbor Algorithm (NN) and the Nearest Subspace Algorithm (NS). The NN algorithm is the compute the distance between the test image and every training image and the identity of the test image is the identity of the training image with the smallest distance. The Nearest distance is the "distance" between the test image and every training sample subspace by calculating the length of projection of the test image onto every subspace spanned by training samples of every identity.

But either the NN algorithm or the NS algorithm is too "local"; they just use one or one kind of training samples to represent the test image. The SRC algorithm is more discriminative. It calculate the most sparse representation of the test image by all training samples. The SRC algorithm is shown below:

$Sparse-representation-based-algorithm$

1 INPUT: a matrix of training samples $A = [A_1, A_2, ..., A_k] \in R^{m \times n}$ for k classes, a test sample $y \in R^m$ and an optimal error tolerance $\epsilon > 0$.
2 Normalize the columns of A to have unit l^2 norm.
3 Solve the l^1 minimization problem:
$\hat{x}_1 = \underset{x}{argmin}\, ||x||_1$ subject to $Ax = y$
Or alternatively, solve
$\hat{x}_1 = \underset{x}{argmin}\, ||x||_1$ subject to $||Ax - y||_2 \leq \epsilon$
4 Compute the residuals $r_i(y) = ||y - A\delta_i(\hat{x}_1)||_2$ for $i = 1, ..., k$.
5 OUTPUT: $identity(y) = \underset{i}{argmin}\, r_i(y)$

We test the SRC algorithm based on the YaleB dataset. We use the first 10 images of every identity as training samples and the 10^{th} to 15^{th} images of every identity as test samples. We use Eigenface as features and use the above SRC algorithm to calculate every identity of every test images. The result is shown as below.

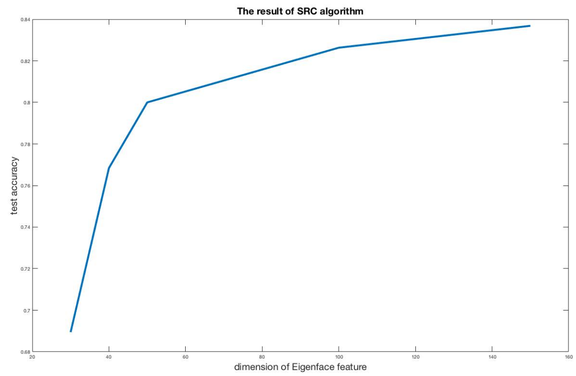

Fig. 1. Result of SRC Algorithm

We calculate the test accuracy for situations of dimension of eigenface feature as 30,40,50,100,150. We can find that the accuracy of the testing is increasing and the highest testing accuracy is around 83%.

IV. Queries

Though the SRC algorithm achieves a big success in face recognition problems. Some people have doubts such that, Is sparse representation really a good representation of face recognition problem? What is the role of sparity in the SRC algorithm? Is the sparity that makes the algorithm successful and can we find a better algorithm? (nearly the same accuracy with less computation). Here we analyze two most prevalent queries about the sparsity used in the SRC algorithm for face recognition problems.

A. Is the training sample space really low-dimensional?

The most important and fundamental hypothesis of the SRC algorithm is that: if we have enough training samples of one indentity of algined face images, then the aligned test image of the same identity lies in the subspace spanned by these training images. But some people have some doubts about this hypothesis. In [1], Shi et. al have done an experiment to disprove this hypothesis. They use the AR dataset. If the hypothesis for SRC algorithm is true, then the dimension of the spaces of all samples of one identity should be less than the number of samples since some samples can be linear represented by others. And then they concluded that the whole sample matrix which is composed of all samples from all identities should also be much less than the number of all samples from all identieis. So they do an experiment that is calculating the singular values of matrix A, where $A = [A_1, A_2, ..., A_k]$, A_k is the sample matrix consisting of all samples from identity k. The result is shown below.

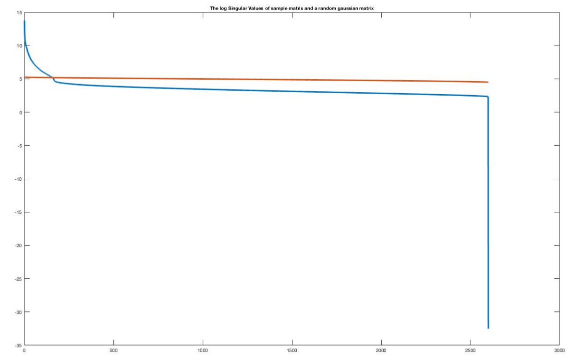

Fig. 2. Singular Values of The Sample matrix and a Gaussian Matrix

From the figure, though the first several singular values are big (Shi et .al states that this due to the common features of face images of all identities), but the figure has long tail, means the singular values of these sample matrix are all non-negligible, thus the sample matrix is full rank, implying that the low rank model is false. The red line in the figure is the singular values for a random gaussian matrix with the same size, which is used for comparison.

But the deduction above could not draw the conclusion that the low dimensional hypothesis is not true for two reasons. Firstly, the low dimensional structure is for only samples for one identity, not for all samples from all identities. Thus the calculation of singular value of the whole sample matrix makes no sense. Secondly, the low dimensional struture needs that all samples with no occlusions. But in the AR dataset, every identity has several samples with sunglasses or scarf, which is a big occlusoin, thus it definitely destroys the low dimensional structure.

Below we calculate the singular values of a sample matrix consisting of samples from one identity with only changes in illumination and facial expressions. We can see that the first three singluar values can count (acutally, only the first one is very big), and the else are very small. Thus this result proves the low dimensional structure is true.

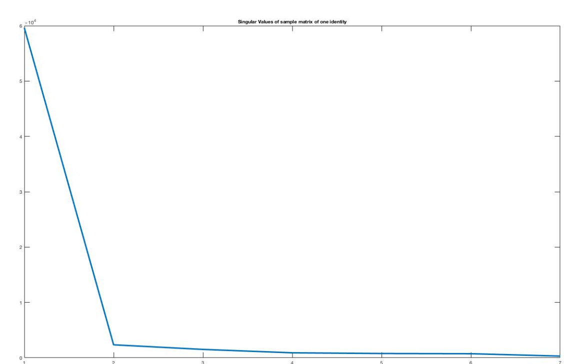

Fig. 3. Singular Values of The Sample matrix of one identity

B. Is collaborative representation rather than l1-norm minimization used in SRC algorithm that makes the algorithm success?

The second query is about the essencial working mechanism of SRC algorithm. Zhang in [5] analyzes it and shows that it is the corroborative representattion used in the SRC algorithm rather than the sparsity (l1-norm minimization) that makes the SRC algorithm successful. But actually, both collaborative representation and sparsity play important role in the algorithm and they do not conflict with each other.

The SRC algorithm arises from Nearest Subspace algorithm as stated before. Actually, if the training samples of one identity is enough, then by our hypothesis, the test image of this identity lies in the subspace spanned by these training samples. So in this situation, both the SRC algorithm and NS algorithm work well. But in many applications in real world, we have very little training samples and we want to identify test samples as many as possible. Thus in this situation, only training samples from one identity is not sufficient to represent the test image. But fortunately, there are many similarities between training samples from different identieis since there are all face images. So we can take advantage of this similarities between identities and use collaborative representation rather than single-identity representation to do identification. In this situation, the NS algorithm won't work, however, the SRC algorithm can still have a good performance. Though collaborative representation plays an important role in the SRC algorithm, this does not mean that the sparity is not essential in the SEC algorithm. In Zhang's paper, [5], they come up with a new algorithm that uses l2-norm minimization rather than l1-norm minimization used in the SRC algorithm to do face recognition. According to the experiment result, this new algorithm is as good as the original SRC algorithm and far more time efficient. Thus the authors conclude that it is the collaborative plays the central role in SRC algorithm rather than l1-norm sparity, so we can use other minimization method such as l2-norm instead of l1-norm in order to implement time efficient algorithm. The new algorithm is shown below.

$CRC - RLS$

1 Normalize the columns of training sample matrix X to have l2-norm

2 Code the test image y over X by
$\rho = Py$
where $P = (X^T X + \lambda I)^{-1} X^T$

3 Compute the regularized residuals
$r_i = ||y - X_i \rho_i||_2 / ||\rho_i||_2$

4 Output the identity of y as
$Identity(y) = argmin_i\{r_i\}$

But actually, from the experiment results of these two algrithms (CRC-RLS and SRC), we can not conclude that the l1-norm minimization used in SRC is not essential and can be replaced by l2-norm minimization. Because the most important role of l1-norm minimization, or equivalently sparsity, is to provide with robustness with respect to errors such as occlusions. Since in the nature environment, errors of an image will always happen at some area of the image, thus the error is pixel-sparsed and we take use of these sparsity of the error to make the identification task work well. But in [5], they only test the face images with different illuminations and facial expressions with these two algorithms and both of them work well. Though they also test the images with sunglasses and scarf, neither of these two algorithms work well. Thus the conclusion that sparsity is not essential for SRC algorithm can not be drawn just from these experiments from [5].

We do experiments on the SRC algorithm to test the robustness with respect to occlusion. We corrupt the sample image by painting some continuous pixel area black (0 value) and use the same SRC algorithm to calculate the test accuracy. The result is shown as below.

We corrupt every image from the 11^{th} pixel in the column and the 11^{th} pixel in the row by $5k \times 5k$ area with whole black error, where $k = 1, 2, ..., 30$. The figure of test accuracy with respect to k.

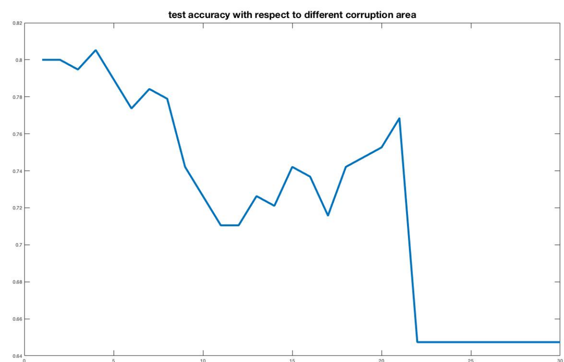

Fig. 4. Test Accuracy with respect to different corruption

We can see from the test accuracy that SRC algorithm is very robust to sparse corruption. It still holds for above 0.7 accuracy until the corruption area has 110×110 pixel area, which is very large since the original image size is 192×168. Thus SRC algorithm holds pretty well to sparse errors. But for the CRC-RLS algorithm, it do not have good robustness to the sparse error as the SRC algorithm. Thus the l1-norm can not be replaced by the l2-norm when we are dealing with images with corruptions.

V. Conclusion

Is sparsity representation a good tool for face recognition problem? Definitely yes since there are lots of algorithms and applications of these tool used in face recognition area that have successful results. However, as for is sparse representation a profound and the intrinsic quality of face recognition problem, in my opinion, my answer is yes. Though the SRC algorithm need all face images to be aligned and is not good for pose changes, but the low dimensional structure of the sample matrix of one identity is true and this gives us very nature thinking that the face recognition is a sparse representation problem. However, as for the alignment requirement for the SRC algorithm, I have some idea that may be helpful. First, since the neural network have big successful today, we can use the neural network to extract features of not aligned face images or face images of different poses and then just use these features to do SRC algorithm. It may have good results since the neural network can extract very deep feature of training samples.

REFERENCES

[1] Qinfeng Shi, Anders Eriksson, Anton Van Den Hengel, and Chunhua Shen. Is face recognition really a compressive sensing problem? In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 553–560. IEEE, 2011.

[2] John Wright, Yi Ma, Julien Mairal, Guillermo Sapiro, Thomas S Huang, and Shuicheng Yan. Sparse representation for computer vision and pattern recognition. *Proceedings of the IEEE*, 98(6):1031–1044, 2010.

[3] John Wright, Allen Y Yang, Arvind Ganesh, S Shankar Sastry, and Yi Ma. Robust face recognition via sparse representation. *IEEE transactions on pattern analysis and machine intelligence*, 31(2):210–227, 2009.

[4] Allen Y Yang, John Wright, Yi Ma, and S Shankar Sastry. Feature selection in face recognition: A sparse representation perspective. *submitted to IEEE Transactions Pattern Analysis and Machine Intelligence*, 2007.

[5] Lei Zhang, Meng Yang, and Xiangchu Feng. Sparse representation or collaborative representation: Which helps face recognition? In *Computer vision (ICCV), 2011 IEEE international conference on*, pages 471–478. IEEE, 2011.