Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017

An Effective Method to Tackle Illumination Problem in Collaborative Representation Based Classification

Mohammed Ahmed Magzoub Albashier
Faculty of Engineering and Technology
Multimedia University
Melaka, Malaysia
magzoub.srcmlc@gmail.com

Alan Tan
Faculty of Engineering and Technology
Multimedia University
Melaka, Malaysia
alan.tan@mmu.edu.my

Azlan Abdaziz
Faculty of Engineering and Technology
Multimedia University
Melaka, Malaysia
azlan.abdaziz@mmu.edu.my

Fawaz Sammani
Faculty of Engineering and Technology
Multimedia University
Melaka, Malaysia
fawaz.sammani@aol.com

Abstract—**many algorithms for face recognition have been used in researches. Sparse representation based classification is an approach that classifies a sample with over complete dictionary. The testing can be recovered via L_1 norm minimization. A newer Approach called Collaborative representation based classification uses the same way as Sparse representative, but it recovers the solution using L_2 norm minimization. Both collaborative representation and sparse representation deal with only a small variation in pose and illumination. In this paper, we propose an approach to tackle the problem of illumination variation in collaborative representation. Our method is a combination between collaborative representation and logarithmic total variation (LTV). In this approach we are using LTV as a pre-processing step to our algorithm. LTV has made a huge impact on the result.**

Keywords—Face Recognition (FR), Sparse representation based classification (SRC), Collaborative representation based classification (CRC), logarithmic total variation (LTV)

I. Introduction

In recent days, Face recognition has occupied the field of many researchers, and it has been used in many applications for the past several years. Because of its overwhelming benefits, such as its high uniqueness, studies have been developing and unstoppable, compared to other applications. Recent researchers showed that in order to achieve a significant result, the face needs to be frontal and normalized [1]. Images naturally are characterized by the high dimension they have, and recently it was found out that the discriminating information of the images can be found in lower dimensional subspaces, any many subspace methods have attracted the interest of numerous studies. [2-5]. However, the subspace learning method only considers the holistic feature of the face images, that are usually sensitive to the variation of illumination. Face recognition classifiers have been further studied, and excellent results were achieved, such as the Nearest Neighbor (NN), Support Vector Machine (SVM), and Hidden Markov [6] [7]. Furthermore, for better achievement, the nearest subspace classifier [8-12] has been developed, which achieved much better results than the nearest subspace (NN). Recently, a new classifier has been developed called SRC, which represents a sample according to an over complete dictionary. The testing image is sparsely coded over the training images, and is then classified by checking class by class, and checking which class gives the least coding error. Recent techniques have been using Sparse Coding for many image restoration applications, with the major help of L0-norm and L1-norm minimization methods. Sparse representation has also occupied the area of pattern classification. Many researchers have managed to utilize Sparse representation to achieve their objectives, like K. Huang and S. Aviyente [13], and J.Wright [14], that managed to achieve robust face recognition using Sparse representation. SRC has truly brought excellent success in face and pattern recognition. Recently, SRC has been further studied to solve the problem of Pose change and illumination variation, as well as the problem of corruption in images. The methods researchers used in [15] and [16] managed to solve this problem. Sparse representation codes a signal Y over a complete dictionary A, such that $Y=A\alpha$, where α is a sparse vector/matrix, where most of its elements are zeros. The sparsity of α can be recognized using L0-norm minimization, which considers the number of non-zeros in α. However, the using L0-norm minimization is considered as NP-hard, so to solve this issue, studies have developed a similar L1- minimization to replace the L0- minimization, which can be represented as: $\|y-Ax\| \le \varepsilon$, where ε is a small constant that represents the error. However, L1-minimization still has a disadvantage, because it takes a considerable amount of time to operate, and many methods were proposed to speed up its process, like Homotophy [17] and Augmented Lagrange Multiplier (ALM) [18], that are mainly used for accuracy and fast speed. Although SRC has achieved successful results, not escaping from its disadvantage, very recently researchers have revealed the operation of SRC and found out that it is the Collaborative representation (CR) that plays the essential role in FR, and makes the FR powerful, and not the L1-norm minimization [19]. The same researchers have also proposed a new approach called CRC that uses L2-norm minimization to regularize the coding coefficient, instead of the time consuming L1-norm minimization. CRC has achieved very close results to SRC, but with much less complexity. However, both SRC and CRC deal only with a small variation in pose and illumination. Researchers have told that the differences caused by variation in pose and illumination is more important than the differences between the individual images [20] [21]. In this paper, we propose a new method to tackle the problem of illumination in CRC based on the logarithmic total variation (LTV) that will be used in face recognition.

II. RELATED WORK

This section mainly discusses SRC and CRC

A. *SRC*

Let's consider that we have n classes, let X=[X1, X2, X3…..Xn]. When a query image is applied, it will be coded as Y=Xα, where α=[α_1, α_2, …α_i,…. α_n], and α_i is the coded matrix associated with class i. After checking class by class, up to n classes, if Y is from the i^{th}class, then the result y=$X_i\alpha_i$ will be the identity of the recognized face, as α_i will best match class X_i because of the significant values it has compared to the zero values in α. These significant values will be able to recognize the identity of the query image Y. SRC is also a generalization of the nearest neighbor classifier (NN) that classifies the test sample based on the best representation in terms of a single training sample, and also the more developed Nearest Subspace classifier (NS), that classifies based on the best representation of all the training samples that are in each class. SRC represents the query image collaboratively by the samples of all classes, using L1-norm minimization, that can effectively overcome the over fitting problem of NN and NS. But what makes SRC important? One fact that is scary in face recognition is that some images are similar. Let's take to dictionaries, Xi and Xj. If we have Xi = Xj + Δ, and Δ is small, then we would receive an error in the identification of the image. To eliminate this problem, we can apply a small sparsity on α_i and α_j, because under a certain condition, one class will yield a smaller error than the other, implying that we have to use not too much, but also not too less training samples. A range between 3-10 would work best. Studies [22] have also developed the L1-minimization, since the L0-minimization is NP-hard, in the following formula: α_i = argminα{||y-X_iα$\|_2^2$ + λ||α$\|_1$}. That formula was achieved after using the identity matrix that is added to the dictionary, in order to code the outlier pixels. We can see that the coding residuals are also characterized by L1-norm to achieve robustness to outlier pixels. This method will effectively reduce the problem of occlusions and corruption, and is referred to as R-SRC. The procedures of standard SRC are summarized in Table 1.

Table 1: The standard SRC Algorithm
1-Normalize the classes, which represent the columns of X via L1-norm minimization. 2-Code y over X: $\hat{\alpha} = \text{argmin}\ \alpha\{\|y - x\alpha\| + \lambda\|\alpha\|1\}$ Where λ represents a positive scalar 3-Calculate the residuals $r_i = \|y - Xi\alpha i\|_2$ Where $\hat{\alpha}_i$ is the coefficient matrix associated with class i. 4- Calculate the result of y, which represents the identity as Identity (y) = arg min$_i$ $\{ri\}$

B. *CRC*

This method was proposed in [19] to show that it's not the sparsity, or the L1-norm minimization that improves the FR accuracy, but the collaborative representation between classes that improves it. The method proposed was to use the L2-norm to regularize the coding coefficient, and the regularized least square equation can help us to find the collaborative representation: $\hat{\alpha}$= argminα{||y-xα$\|_2^2$ + λ||α$\|_2^2$}

The role of l_2-regularization term $\|\alpha\|_2$ is to make the result stable, specifically when x is undetermined, and to give a small amount of sparsity to $\hat{\alpha}$. However, this sparsity is much less than the L1-norm. It can be derived that $\hat{\alpha}$ = PY, where P = $(X^TX+\lambda.I)^{-1}$ X^T is the projection matrix, and it is independent of Y, thus it can be calculated earlier. As soon as a query image Y is applied, Y can be projected onto P easily, and that is what makes CRC very fast and less complex.

III. AN EFFECTIVE METHOD

This section discusses the proposed method to tackle the illumination problem in collaborative representation based classification. We did implement CRC to each subset of the 5 subset in the Yale B data base based on the following :

> 1. Divided the subset images into training data and testing data with different experimental adjustment
> 2. Down sampled the images using imresize into 56 dimensionality.
> 3. Computed the projection matrix of the training images
>
> $$P = (X^TX+\lambda.I)^{-1} X^T$$
>
> 4. Coded y over X by
>
> $$\hat{\alpha} = Py$$
>
> 5. Computed the regularized residuals
>
> $$r_i = \|y - X_i\hat{\alpha}_i\|_2 / \|\hat{\alpha}_i\|_2$$
>
> 6. Get the output identity of y as
>
> $$\text{Identity (y)= arg min}_i\{r_i\}$$

Empirically, we have noticed that the above collaborative representation based classification suffer from illumination variation. Not a lot of work has been done to solve this problem. As a collaborative representation considers apart from sparse representation which have been widely studied in the community of computer vision as well as signal processing. Collaborative representation consider as the fast version of sparse representation. LTV was used as a model as a pre-processing technique for the face recognition algorithm this is the idea behind (CRC_LTV). LTV reduces the notorious halo artifacts and leaves only illumination invariant small scale facial structures with only one simple parameter to set.

The new method that we are purposing can be described as follow:

1 Divided the subset images into training data and testing data with different experimental adjustment
2 Selected the LTV parameter λ as 0.3
3 Input the training and testing images to LTV algorithm to obtain the background hues u, as well as sharp edges v.
4 Down sampled the sharp edges v using imresize into 56 dimensionality.
5 Computed the projection matrix of the training images

$$P = (X^TX+\lambda.I)^{-1} X^T$$

6 Coded y over X by

$$\hat{\alpha} = Py$$

1 Computed the regularized residuals

$$r_i = \|y - X_i\hat{\alpha}_i\|_2 / \|\hat{\alpha}_i\|_2$$

2 Get the output identity of y as

$$\text{Identity (y)= arg min}_i\{r_i\}$$

The above method has proved its effectiveness In Yale B database with all the subsets with different illumination variation as it is going to illustrate in the next part.

IV. EXPERIMENTS AND ANALYSIS

This section, the experiments are carried out on one database known as Extended Yale B [58][20]. Example pictures are shown in Figure. The face images based on the angle between the camera axis and the light direction are divided into 5 subsets. Each subset contains images for 20 different people and the images captured under uncontrolled illumination.

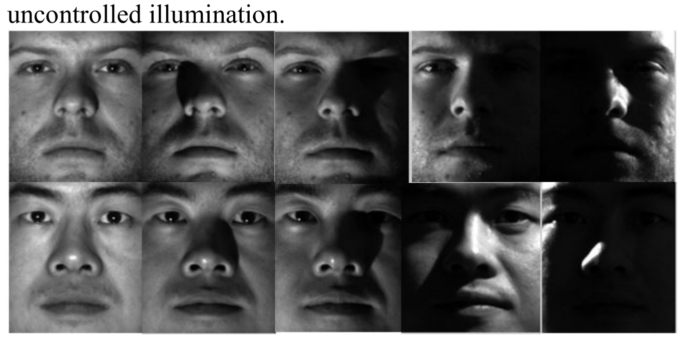


FIGURE EXTENDED YALE B DATABASE

Table 1

The table shows the angle between the camera axis and the light direction the light direction and number of images for each subset.

Subset	Angle	No of images
1	Up to 15°	240
2	Up to 25°	140
3	Up to 50°	220
4	Up to 95°	300
5	Up to 130°	220

Identify applicable funding agency here. If none, delete this text box.

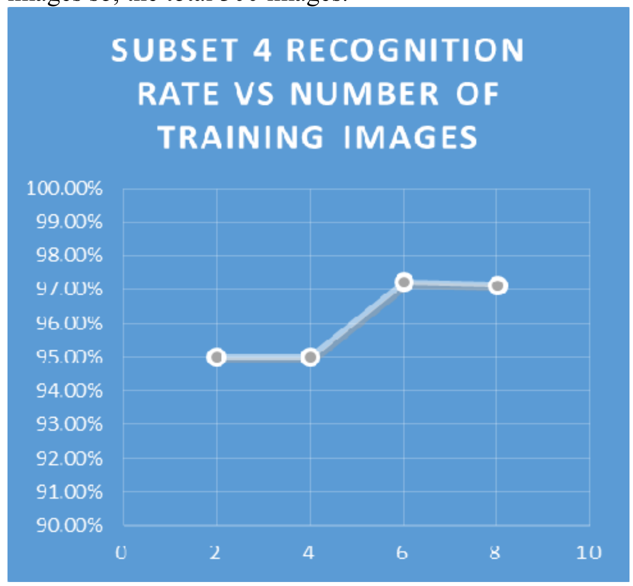
We have implemented our experiment on subset 2,3,4 and 5 as the there is variation in illumination in these subsets. Subset 2 includes 20 individuals each of them have 7 images so, the total 140 images.

Subset 2 Training images	Subset 2 Testing images	Recognition Rate before LTV	Recognition Rate After LTV
3	4	26.25%	100%
4	3	60%	100%
2	5	52%	96%
1	6	37%	90%

Subset 3 includes 20 individuals each of them have 11 images so, the total 220 images.

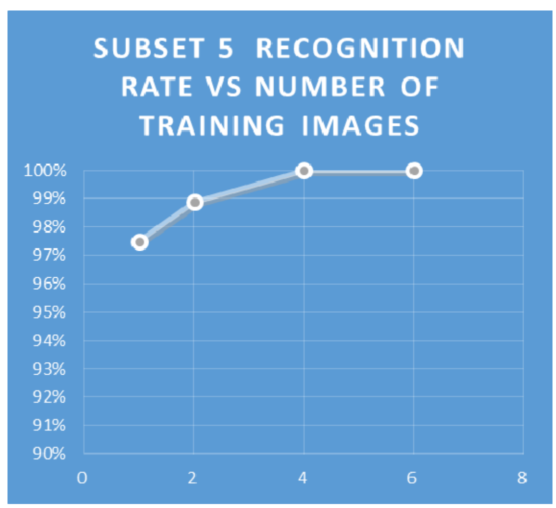
Subset 3 Training images	Subset 3 Testing images	Recognition rate before LTV	Recognition Rate after LTV
6	5	14%	98%
4	7	26.4286%	97.8571%
2	9	23.8889%	95.5556%
1	10	21%	95%

Subset 4 includes 20 individuals each of them have 15 images so, the total 300 images.



Subset 4 Training images	Subset 4 Testing images	Recognition rate before LTV	Recognition Rate after LTV
8	7	4.2857%	97.1429%
6	9	26.6667%	97.2222%
4	11	32.2727%	95%
2	13	23.4615%	95%

Subset 5 includes 20 individuals each of them have 11 images so, the total 220 images.



Subset 5 Training images	Subset 5 Testing images	Recognition rate before LTV	Recognition Rate after LTV
6	5	6%	100%

4	7	22.1429%	100%
2	9	17.2222%	98.8889%
1	10	23.5000%	97.5000%

TABLE 1

BEST CLASSIFICATION RATE OBTAINED BY EACH METHOD

V. CONCLUSION

In conclusion, our main aim was to tackle the problem of illumination variation in collaborative representation based classification. Adding logarithmic total variation as a pre-processing step to our main algorithm (collaborative representation based classification) was our idea to tackle the problem of illumination variation in collaborative representation based classification. LTV made a huge impact on CRC and gave a really excellent result as it has been illustrated in the results and discussion session.

REFERENCES

[1] H. Huang and H. T. He, "Super-resolution method for face recognition using nonlinear mappings on coherent features," IEEE Trans. Neural Netw., vol. 22, no. 1, pp. 121–130, Jan. 2011.

[2] M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cognit. Neurosci., vol. 3, no. 1, pp. 71–86, 1991.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriengman, "Eigenfaces versus Fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711–720, Jul. 1997.

[4] S. Zafeiriou, G. Tzimiropoulos, M. Petrou, and T. Stathaki, "Regularized kernel discriminant analysis with a robust kernel for face recognition and verification," IEEE Trans. Neural Netw. Learn. Syst., vol. 23, no. 3, pp. 526–534, Mar. 2012.

[5] J. B. Tenenbaum, V. deSilva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," Science, vol. 290, no. 5500, pp. 2319–2323, 2000.

[6] A. Timo, H. Abdenour, and P. Matti, "Face recognition with local binary patterns," in Proc. Eur. Conf. Comput. Vis., May 2004, pp. 469–481.

[7] S. F. Xie, S. G. Shan, X. L. Chen, and J. Chen, "Fusing local patterns of gabor magnitude and phase for face recognition," IEEE Trans. Image Process., vol. 19, no. 5, pp. 1349–1361, May 2010.

[8] Y. G. Liu, S. Z. Sam Ge, C. G. Li, and Z. S. You, "K-NS: A classifier by the distance to the nearest subspace," IEEE Trans. Neural Netw., vol. 22, no. 8, pp. 1256–1268, Aug. 2011.

[9] S. Z. Li, "Face recognition based on nearest linear combinations," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 1998, pp. 839–844.

[10] S. Z. Li and J. Lu, "Face recognition using nearest feature line method," IEEE Trans. Neural Netw., vol. 10, no. 2, pp. 439–443, Mar. 1999.

[11] I. Naseem, R. Togneri, and M. Bennamoun, "Linear regression for face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 11, pp. 2106–2112, Nov. 2010.

[12] J. C. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 1794–1801.

[13] K. Huang and S. Aviyente. Sparse representation for signal classification. In NIPS, 2006

[14] J. Wright, A. Y. Yang, A Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. IEEE PAMI, 31(2):210-227, 2009

[15] A. Wagner, J. Wright, A. Ganesh, Z.H. Zhou, and Y. Ma, Towards a practical face recognition system: robust registration and illumination by sparse representation. In CVPR 2009.

[16] Y. Peng, A. Ganesh, J.Wright. W. Xu, and Y. Ma. RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images. Submitted to IEEE PAMI, 2010.

[17] D. Malioutove, M. Cetin, and A. Willsky. Homotophy continuation for sparse signal representation. In ICASSP, 2005

[18] S.J.Kim, K. Koh, M.Lustig, S.Boyd, and D.Gorinevsky. A interior-point method for large-scale l1-regularized least squares. IEEE Journal on Selected Topics in Signal Processing, 1(4):606-617, 2007

[19] Sparse Representation or Collaborative Representation: Which Helps Face Recognition?" In *ICCV 2011*

[20] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction" IEEE Tran. PAMI, 19(7), pp. 721-732, 1997.

[21] W. Zhao, and R. Chellappa, "Robust face recognition using symmetric shape-from-shading" Technical report, Center for Automation Research, University of Maryland, 1999.