International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 18
December 2016

Under the Strong Type Facial Recognition Samples Learn Dictionary Assistant

1. KAMBALAPALLI RAMA DEVI, 2. Mrs.G.SATYA PRABHA

[1]PG Scholar, Department of ECE, SLC's Institute of Engineering and Technology, Piglipur Village, Hayathnagar Mandal, Near Ramoji Film City, Ranga Reddy District, Hyderabad, Telangana

[2]Assistant Professor, Department of ECE, SLC's Institute of Engineering and Technology, Piglipur Village, Hayathnagar Mandal, Near Ramoji Film City, Ranga Reddy District, Hyderabad, Telangana

Abstract In this paper, we address the problem of robust face recognition with undersampled training data. Given only one or few training images available per subject, we present a novel recognition approach, which not only handles test images with large intraclass variations such as illumination and expression. The proposed method is also to handle the corrupted ones due to occlusion or disguise, which is not present during training. This is achieved by the learning of a robust auxiliary dictionary from the subjects not of interest. Together with the undersampled training data, both intra and interclass variations can thus be successfully handled, while the unseen occlusions can be automatically disregarded for improved recognition. Our experiments on four face image datasets confirm the effectiveness and robustness of our approach, which is shown to outperform state-of-the-art sparse representation-based methods.

.I. INTRODUCTION

Face recognition has been an active research topic, since it is challenging to recognize face images with illumination and expression variations as well as corruptions due to occlusion or disguise. A typical solution is to collect a sufficient amount of training data in advance, so that the above intraclass variations can be properly handled. However, in practice, there is no guarantee that such data collection is applicable, nor the collected data would exhibit satisfactory generalization. Moreover, for real-world applications, e.g. e-passport, driving license, or ID card identification, only one or very few face images of the subject of interest might be

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 18
December 2016

captured during the data acquisition stage. As a result, one would encounter the challenging task of undersampled face recognition .Existing solutions to undersampled face recognition can be typically divided into two categories: patch-based methods and generic learning from external data. For patch-based methods, one can either extract discriminative information from patches collected by different images, or utilize/integrate the corresponding classification results for achieving recognition. pattern (LBP) Gabor features or manifold learning while the latter advanced weighted plurality voting or margin distribution optimization Nevertheless, the major concern of patch-based methods comes from the fact that local patches extracted from undersampled training data only contain limited information, especially for the scenario of single-sample face recognition (i.e., one training image per person). As a result, the classification results would degrade significantly when there exists large variations between the query and the gallery ones. Moreover, patch-based methods often assume that the image patches are free from occlusion; this would limit their uses in practical scenarios. In contrast to patch-based approaches for undersampled face recognition, the second type of methods advocate the use of external data which contain the subjects not of interest. These approaches aim at learning the classifiers with improved recognition abilities or modeling the intra-class variations. For example, based on the assumption that the face images of different subjects are independent, adaptive generic learning (AGL) [7] utilized external data for estimating the within-class scatter matrix for each subject to be recognized. Different from AGL which requires the above assumption, Kan et al. further proposed a nonlinear estimation model to calculate the within-class scatter matrix. Different from recent works like employed external data for describing possible intra-class variations when performing recognition. Although promising result have been shownthese approaches require the query image and the external data to exhibit the same type of occlusion, which might not be practical. Since we typically do not have the prior knowledge on the occlusion of concern, how to select external data for learning intra-class variations would become a problem for methods like . Fig. 1. Illustration of our proposed method for undersampled face recognition, in which the gallery set only

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 18
December 2016

contains one or few face images per subject of interest, while the auxiliary dictionary is learned from external data for observing possible image variants. Note that the corrupted image regions of the query input can be automatically disregarded using our proposed method.

Unlike which require the prior knowledge of the occlusion, our approach eliminates such assumptions by introducing a novel classification method based on robust sparse coding. It is worth noting that existing dictionary learning algorithms like KSVD [13] can also be used to learn dictionaries for images from external datasets. However, these learned dictionaries cannot guarantee the recognition performance for the subjects of interest, since KSVD only considers the representation ability of dictionaries. In our work, we jointly solve the tasks of auxiliary dictionary learning and robust sparse coding in a unified optimization framework). This makes our approach able to improve the performance for robust face recognition under the scenario of undersampled training data.

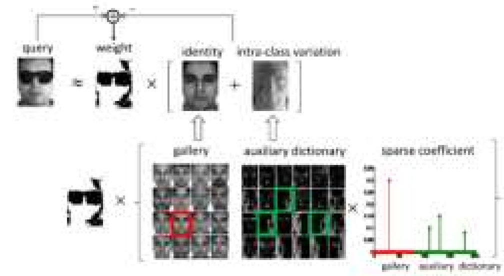

Fig. 1 illustrates our idea of the proposed method. By learning an auxiliary dictionary from an external dataset together with robust sparse coding, the benefits of our approach are threefold.

Firstly, we are able to address undersampled face recognition problem, since only one or few training images of the subjects to be recognized are required for training. Therefore, there is no need to collect a large training dataset for covering image variants for all subjects of interest. Secondly, our approach provides a new tool for recognizing occluded face images by means of robust sparse coding and the auxiliary dictionary, while no assumptions are made about the information on occlusion. Finally, our algorithm for auxiliary dictionary learning allows one to model intra-class variations including illumination and expression changes from external data. By solving both auxiliary dictionary learning and robust face recognition in a unified framework, improved recognition performance can be expected. The remaining of this paper is organized as follows.

II. RELATED WORK

A. SRC and Extended SRC

Recently, Wright et al. proposed sparse representation based classification (SRC) for face recognition. Since our proposed method is extended from SRC, we briefly review this classification technique for the completeness of this paper. Given a test image y, SRC represents y as a sparse linear combination of a codebook D = [D1,D2, · · · ,DL], where Di denotes the training images associated with class i .

Precisely, SRC derives the sparse coefficient x of y by solving the following L1-minimization problem: We note that, compared to (2), the operator δ_(·) in (4) is only applied to xd instead of the entire coefficient vector x. This is because that xa is not associated with any class label information. Although ESRC has shown promising results on undersampled face recognition, there are three concerns with ESRC. Firstly, ESRC directly apply external as A, which might be noisy or contain undesirable artifacts. Secondly, the computation of (3) would be very expensive due to the large size of A. This is due to the fact that ESRC needs the matrix A for covering all intra-class variations of interest. Finally, ESRC regards occlusion as intra-class variations during the collection of A from external data. In other words, ESRC assumes the type of occlusion to be known when collecting external data, which might not be practical.

B. Dictionary Learning for Sparse Coding

Recent research on computer vision and image processing has shown that the learning of data or application-driven dictionaries outperforms approaches using predefined ones . In general, the optimization algorithms for dictionary learning can be designed in an unsupervised or supervised manner. Unsupervised dictionary learning such as MOD [17] or KSVD [13] focuses on data representation, and is suitable for image synthesis tasks like image denoising. Nevertheless, for addressing recognition tasks, one requires supervised dictionary learning strategies which aim at introducing improved discriminative capability for the observed learning model. Several approaches have been proposed by introducing different

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 18
December 2016

classification criteria to the objective function.

.

C. Remarks on SRC-Based Approaches for Face Recognition

We highlight and compare the properties of recent sparse representation based face recognition methods in Table I, in which SRC ESRC ADL SVDL have been discussed in previous two subsections. It is worth mentioning that Yang et al. have proposed an iteratively reweighted sparse coding algorithm to improve SRC for better dealing with outliers such as occlusion or corruption. Another recent work utilized low-rank matrix decomposition with structural incoherence to address the scenario where both training and test data can have occluded images. Both do not require the knowledge of occlusion in test images, but they need a sufficient amount of training data to cover image variants for all subjects of interest. Directly applying the methods of and to undersampled face recognition

can lead to degraded recognition performance. Later in the experiments, we will confirm that our approach outperforms state-of-the-art SRC based methods.

III. OUR PROPOSED METHOD

A. Face Recognition via Robust Auxiliary Dictionary Learning

1) Our Classification Formulation: We now present our classification algorithm for undersampled face recognition via robust auxiliary dictionary learning, as shown in the upper

part of Let y ∈ Rd be the query image and D ∈ Rd×nbe the gallery matrix. The gallery matrix D is composed of data matrices from L classes, i.e. D = [D1,D2, . . . ,DL]. The auxiliary dictionary A ∈ Rd×m is learned from external data, and the detailed algorithms

$$\min_{\mathbf{x}} \rho\left(\mathbf{y} - [\mathbf{D},\ \mathbf{A}]\begin{bmatrix}\mathbf{x}_d \\ \mathbf{x}_a\end{bmatrix}\right) + \lambda\|\mathbf{x}\|_1. \tag{5}$$

where $\mathbf{x} = [\mathbf{x}_d;\ \mathbf{x}_a]$ is the sparse coefficient of $\mathbf{y}$, and the residual function $\rho(\cdot): \mathbb{R}^d \rightarrow \mathbb{R}$ is defined as

$$\rho(\mathbf{e}) = \sum_{k=1}^{d} \rho(e_k),$$

$$\rho(e_k) = -\frac{1}{2\mu}\left(\ln\left(1 + \exp(-\mu e_k^2 + \mu\delta\right) - \ln(1 + \exp \mu\delta)\right), \tag{6}$$

where ek is the kth entry of e = y−[D,A]x, and the parameters μ and δ will be detailed at the end of this subsection. In the theory of robust M-estimators [24], the residual function ρ(·) in (5) is designed to minimize the influence of outliers. Standard residual functions used in robust

M-estimators include Huber, Cauchy, and the Welsch functions. We consider the residual function ρ(·) defined in (6), because

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 18
December 2016

this type of residual functions has shown promising results

$$\begin{aligned}
&\sum_{k=1}^{d} \frac{d\rho(e_k)}{de_k}\frac{de_k}{d\mathbf{x}} + \lambda\partial\|\mathbf{x}\|_1 \\
&= \frac{1}{2}\sum_{k=1}^{d} \frac{d\rho(e_k)}{de_k}\frac{1}{e_k}\frac{de_k^2}{d\mathbf{x}} + \lambda\partial\|\mathbf{x}\|_1 \\
&= \frac{1}{2}\sum_{k=1}^{d} w(e_k)\frac{de_k^2}{d\mathbf{x}} + \lambda\partial\|\mathbf{x}\|_1, \qquad (8)
\end{aligned}$$

where

$$w(e_k) = \frac{d\rho(e_k)}{de_k}\frac{1}{e_k} = \frac{\exp(-\mu e_k^2 + \mu\delta)}{1 + \exp(-\mu e_k^2 + \mu\delta)}. \qquad (9)$$

If $w(e_k)$ in (8) is fixed as a constant, then (8) becomes the derivative of

in recent literatures of robust face recognition and updating W according to (11), where ek is the

kth entry of e. Notice that with W fixed, (12) is in the form of the standard L1-minimization problem, and one can apply existing techniques such as Homotopy, Iterative Shrinkage-Thresholding, or Alternating Direction Method for solving (12).

where $\mathbf{e} = \mathbf{y} - [\mathbf{D}, \mathbf{A}]\mathbf{x}$ and

$$\mathbf{W} = \operatorname{diag}(w(e_1), w(e_2), \ldots, w(e_d))^{1/2}. \qquad (11)$$

From the above derivation, we know that the solution of (5) can be calculated by repeatedly solving

$$\min_{\mathbf{x}} \left\| \mathbf{W}\left(\mathbf{y} - [\mathbf{D},\ \mathbf{A}]\begin{bmatrix}\mathbf{x}_d \\ \mathbf{x}_a\end{bmatrix}\right)\right\|_2^2 + \lambda\|\mathbf{x}\|_1. \qquad (12)$$

IV. EXPERIMENTAL RESULTS

A. Extended Yale B Database

First, we consider the Extended Yale B database [28] for our experiments. This database contains 38 subjects with about 64 frontal face images for each and the face images are taken under various illumination conditions All images are converted into grayscale and are down sampled to 34×30 pixels prior to our experiments. We select 32 subjects from the database to be recognized,

and the remaining 6 subjects are considered as external data(i.e., subjects not of interest) for robust auxiliary dictionary learning .For the 32 subjects of interest, we select 3 images from each of the 32 subjects as the gallery D, and the remaining 61 images for testing. The three gallery images correspond to the three illumination conditions: A+000E+00, A-085E+20, and A+085E+20 (A+085 refers to 85 degrees azimuth, and E+20 refers to 20 degrees elevation [28]). For the training stage of robust auxiliary dictionary learning using external data (i.e., the six subjects not of interest), we choose the same images corresponding to A+000E+00, A-085E+20,

and A+085E+20 as the gallery De, and thus De contains a total of 6×3 images. The probe Ye consists of the randomselection of 29 images from the remaining images of these6 subjects. We

B. AR Database

1) Face Recognition and RADL With the Same Domain: The AR database [30] consists of over 4,000 frontal face images of 126 individuals. The images are taken under different variations, including illumination, expression, and facial occlusion/disguise in two separate sessions. For each session there are thirteen images, in which three images are with sunglasses, another three are with scarfs, and the remaining seven are with illumination and expressions variations. In our experiments, we consider a subset of AR consisting of 50 men and 50 women. All images are converted to grayscale and cropped to 165×120 pixels. We select 80 subjects of interest for training and testing, and the remaining 20 subjects are considered as external data for robust auxiliary dictionary learning. For the scenario of undersampled face recognition, we choose only the neutral image of each of the 80 subjects (40 men and 40 women) in Session 1 as the gallery, and the rest images in Sessions 1 and 2 are for testing,

V. CONCLUSION

We presented a novel learning-based algorithm for undersampled face recognition. We advocated the learning of an auxiliary dictionary from external data for modeling intra-class image variants of interest, and utilized a residual function in a joint optimization formulation for identifying and disregarding corrupted image regions due to occlusion. As a result, the proposed algorithm allows one to recognize occluded face images, or those with illumination and expressions variations, even only one or few gallery images per subject are available during training. Experimental results on four different face image datasets confirmed the effectiveness and robustness of our method, which was shown to outperform state-of-the-art sparse representation and dictionary learning based approaches with or without using external face data.

REFERENCES

[1] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey," Pattern Recognit., vol. 39, no. 9, pp. 1725–1745, Sep. 2006.

[2] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 18
December 2016

[3] J. Zou, Q. Ji, and G. Nagy, "A comparative study of local matching approach for face recognition," IEEE Trans. Image Process., vol. 16,no. 10, pp. 2617–2628, Oct. 2007.

[4] J. Lu, Y.-P. Tan, and G. Wang, "Discriminative multimanifold analysis for face recognition from a single training sample per person," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 1, pp. 39–51, Jan. 2013. 1734 IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 24, NO. 6, JUNE 2015

[5] R. Kumar, A. Banerjee, B. C. Vemuri, and H. Pfister, "Maximizing all margins: Pushing face recognition with kernel plurality," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Nov. 2011, pp. 2375–2382.

[6] P. Zhu, L. Zhang, Q. Hu, and S. C. K. Shiu, "Multi-scale patch based collaborative representation for face recognition with margin distribution optimization," in Proc. 12th Eur. Conf. Comput. Vis., 2012, pp. 822–835.

[7] Y. Su, S. Shan, X. Chen, and W. Gao, "Adaptive generic learning for face recognition from a single sample per person," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2010, pp. 2699–2706.

[8] M. Kan, S. Shan, Y. Su, D. Xu, and X. Chen, "Adaptive discriminant learning for face recognition," Pattern Recognit., vol. 46, no. 9, pp. 2497–2509, Sep. 2013.

[9] W. Deng, J. Hu, and J. Guo, "Extended SRC: Undersampled face recognition via intraclass variant dictionary," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 9, pp. 1864–1870, Sep. 2012.

[10] C.-P. Wei and Y.-C. F. Wang, "Learning auxiliary dictionaries for undersampled face recognition," in Proc. IEEE Int. Conf. Multimedia Expo (ICME), Jul. 2013, pp. 1–6.

[11] M. Yang, L. Van Gool, and L. Zhang, "Sparse variation dictionary learning for face recognition with a single training sample per person," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2013, pp. 689–696.

[12] M. Yang, L. Zhang, J. Yang, and D. Zhang, "Robust sparse coding for face recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2011, pp. 625–632.

[13] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 18
December 2016

sparse representation," IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[14] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 2, pp. 210–227, Feb. 2009.

[15] C.-F. Chen, C.-P. Wei, and Y.-C. F. Wang, "Low-rank matrix recovery with structural incoherence for robust face recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2012, pp. 2618–2625.

[16] I. To˘si´c and P. Frossard, "Dictionary learning: What is the right representation for my signal?" IEEE Signal Process. Mag., vol. 28, no. 2, pp. 27–38, Mar. 2011.

[17] K. Engan, S. O. Aase, and J. Hakon Husoy, "Method of optimal directions for frame design," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., Mar. 1999, pp. 2443–2446.

[18] I. Ramirez, P. Sprechmann, and G. Sapiro, "Classification and clustering via dictionary learning with structured incoherence and shared features," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2010, pp. 3501–3508.

[19] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Fisher discrimination dictionary learning for sparse representation," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Nov. 2011, pp. 543–550.

[20] D.-S. Pham and S. Venkatesh, "Joint learning and dictionary construction for pattern recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2008, pp. 1–8.