Proceedings of the IEEE International Conference on Multimedia and Expo Workshops (ICMEW) 2017 10-14 July 2017

DOMAIN TRANSFER SPARSE REPRESENTATION FOR SINGLE SAMPLE FACE RECOGNITION

Venice Erin Liong[1] *and Haibin Yan*[2*]

Interdisciplinary Graduate School, Nanyang Technological University, Singapore, 639798, Singapore
School of Automation, Beijing University of Posts and Telecommunications, Beijing, 100876, China
veniceer001@e.ntu.edu.sg; eyanhaibin@bupt.edu.cn

ABSTRACT

In this paper, we propose a new single sample face recognition approach under the widely used sparse representation-based classification (SRC) framework. Previous work has shown that SRC only works well when there are sufficient number of training samples per person and not suitable for SSFR. To address this, we propose a domain transfer sparse representation-based classification (DT-SRC) method by using an auxiliary dataset to learn intra-class variations and transferring them into the single-sample training set. Since the auxiliary and training sets are likely captured in different environments, we apply the dictionary learning technique to learn a meta-space to transfer intra-class variations from the auxiliary set to the training set. To achieve this, we minimize the distribution difference of these two datasets in the meta-space so that such information can be effectively transferred. We extend DT-SRC to discriminative DT-SRC (DDT-SRC) by making use of the label information samples in the auxiliary set to exploit more discriminative information in the learned meta-space. Experimental results on three face benchmark datasets demonstrate the effectiveness of the proposed approach.

Index Terms— Face recognition, single-sample face recognition, sparse representation, domain transfer

1. INTRODUCTION

In some real-world applications such as e-passport and ID card verification, there are cases that the training samples only provide one image for each person since it is generally difficult to collect these data multiple times. This problem is called Single Sample Face Recognition (SSFR) in the face recognition community. Here, it is fairly challenging to have a discriminative model which correctly estimates the intra-class and inter-class variations due to the lack of training samples.

*Corresponding author.
This work was supported by the National Natural Science Foundation of China under Grant 61603048, Beijing Natural Science Foundation under Grant 4174101, and the Fundamental Research Funds for the Central Universities.

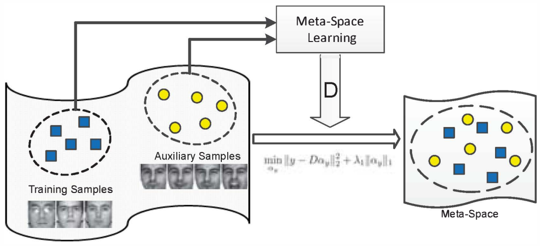

Fig. 1. Basic idea of our proposed DT-SRC method. We use an auxiliary set to model the intra-class variations for SRC-based single-sample face recognition. In this figure, the squares and circles denote face samples from the training set and the auxiliary set, respectively. Since they are captured in different environments, there is large distribution difference in the original feature space. Having learned the meta-space, they are closer to each other and the distribution difference is significantly reduced.

The sparse representation-based classification (SRC) method [1] is one of the most successful recognition approaches in face recognition and many SRC-based face recognition methods have been proposed in the literature. However, it is not suitable for the SSFR problem since it is inconsistent to the basic assumption of the SRC model in which given sufficient number of samples per person in the training set, the testing image can be represented as a sparsely linear combination of the training samples, and it can be recognized as the class which yields the minimal reconstruction error. As shown in previous work, the performance of SRC is not good for SSFR [2], mainly because there are not enough training samples per person to model the intra-class variations.

To make SRC applicable to SSFR, Deng *et al.* [2] presented an extended SRC (ESRC) method by modeling each testing image as a linear space combination of the training set and an additional intra-class variation bases extracted from another auxiliary dataset. More recently, they also proposed a superposed SRC (SSRC) [3] method by using a prototype plus variation model to improve the SRC performance for SSFR. Yang *et al.* [4] proposed a sparse variation dictionary learning

978-1-5386-0560-8/17/$31.00 ©2017 IEEE

(SVDL) method by employing a large generic dataset to learn an adaptive projection and dictionary to model the intra-class variations. While improved performance can be obtained by these SRC-based SSFR methods [2–4], intra-class variations learned from the auxiliary set may be not consistent to those of the training set because they are likely captured in different environments. Hence, these methods suffer from the variations between the training dataset and the auxiliary dataset. To tackle these challenges, we propose in this paper a domain transfer sparse representation-based classification (DT-SRC) to learn a sharable space for auxiliary and training set, in the framework of SRC for SSFR. The basic idea is illustrated in Fig. 1. We apply the dictionary learning technique to learn a meta-space by minimizing their distribution difference so that intra-class variations learned from the auxiliary dataset can be effectively transferred into the training set. We further extend DT-SRC to discriminative DT-SRC (DDT-SRC) by making use of the label information of samples in the auxiliary dataset to extract more discriminative information in the learned meta-space. Experimental results on three benchmark datasets show the effectiveness of our proposed methods.

2. BACKGROUND

In this section, we briefly review two related topics: 1) single-sample face recognition, and 2) transfer learning.

2.1. Single-Sample Face Recognition

Existing SSFR methods can be mainly classified into two categories: robust feature representation and generic discriminative learning. Methods in the first category extract robust local descriptors to enlarge the margin among different persons in the feature space. Representative methods in this category include local binary pattern (LBP), Gabor features, and their combinations. The second category uses an auxiliary generic dataset consisting of multiple samples per person to model the intra-class variations, which can provide more new information to the training set. However, the performance of these methods depends heavily on the selection of the generic set for SSFR. Moreover these methods employ the learned intra-class variations from the auxiliary generic dataset to the testing set directly, which is not optimal because it is still difficult to predict the variations of the training set with single sample per person. In many real world applications, the generic set and the training set are usually captured in different environments so that there is a large distribution gap between these two sets. Hence an adaptive generic set is necessary. For example, Su *et al.* [5] learned a Fisher's linear discriminant (FLD) model a generic training set. While Yang *et al.* [4] performed sparse variation dictionary learning which is adaptively and jointly learned to exploit the correlation between gallery and training set. In this work, we propose a transfer learning approach using sparsity-based learning for SSFR so that the selection of the generic training set is not critical. Deng *et al.* [2] is a sparsity-based face recognition model which extends the SRC by modeling each testing face as a linear space combination of the training set and the intra-class variation bases computed from an auxiliary dataset. While it addressed the SSFR to some extent, intra-class variations of the auxiliary dataset cannot well represent those in the training set because these two sets are usually captured in different environments. In this work, we propose a DT-SRC method to simultaneously learn the intra-class variations and transfer them into the training set so that the SRC classifier is more suitable to SSFR.

2.2. Transfer Learning

Transfer learning aims to transfer knowledge from the source domain to the target domain due to the real nature that two domains are of different distributions. Over the past years, a variety of transfer learning methods have been presented. For example, Pan *et al.* [6] proposed a transfer component analysis which is a subspace learning method which reduces the maximum mean discrepancy (MMD) of labeled and unlabeled data, while also minimizing the reconstruction error of the unlabeled data. Duan *et al.* [7] proposed a domain transfer support vector machine which learns a kernel function and an SVM classifier jointly. Zhang *et al.* [8] proposed a transfer metric learning approach which learn a metric and domain covariances between source and target domain based on a unified convex formulation. Most existing transfer learning methods were developed to learn a transferable feature or model, and none of them has been used in the sparsity-based SSFR applications.

3. PROPOSED APPROACH

3.1. Domain Transfer SRC

Let $X^{tr} = [x_1^{tr}, x_2^{tr}, \cdots, x_{N_0}^{tr}] \in \mathbb{R}^{d \times N_0}$ and $X^{au} = [x_1^{au}, x_2^{au}, \cdots, x_{N_1}^{au}] \in \mathbb{R}^{d \times N_1}$ be the training set and the auxiliary set, where N_0 and N_1 are the number of samples of these two sets, and d is the feature dimension of each sample. Assume there are C classes in X^{au} and K_c samples in the cth class, where $N_1 = \sum_{c=1}^{C} K_c$ and $1 \leq c \leq C$. In X^{tr}, x^p is the single training sample of the pth person, where $1 \leq p \leq N_0$.

We let a given test image y be represented as a linear combination of the training set and the intra-class variation bases as in ESRC [2]:

$$y = X^{tr}\beta + V^{tr}\gamma + z \tag{1}$$

where V^{tr} is the intra-class variation for the training set, β and γ are the representation coefficients in the training set and the variation basis matrix, respectively, and $z \in \mathbb{R}^d$ is a noise term.

Since V^{tr} is unknown and difficult to estimate directly from X^{tr}, ESRC uses V^{au} as an alternative where it assumes that intra-class variations can be shared across different persons. However, face images are usually captured under different conditions, the variations obtained from the auxiliary dataset are likely different from those of the training set. To address this, we propose the following transfer learning approach which aim to learn a new space, where X^{au} can better approximate X^{tr}. Our new space is based on a dictionary D. In the new space, given a sample y, we represent it as coefficient α_y, which can be obtained by the following optimization problem:

$$\min_{\alpha_y} \|y - D\alpha_y\|_2^2 + \lambda_1 \|\alpha_y\|_1 \quad (2)$$

where λ_1 is a parameter to balance these two terms in (2). We aim to learn the dictionary D, so that the training data and the auxiliary data have similar representation. As a result, X^{tr} and X^{au} can be close as much as possible.

Let $X = [X^{tr} X^{au}] \in \mathbb{R}^{d\times N}$ be the augmented data matrix which is a combination of the training set and the auxiliary set, where $N = N_0 + N_1$. We aim to learn a dictionary under which the distribution difference between the training set and the auxiliary set is minimized so that intra-class variations estimated from the auxiliary set can be transferred into the training set. To achieve this, we formulate the following optimization problem:

$$\begin{aligned} \min_{D,A} \quad & \sum_{i=1}^{N} \left(\|x_i - D\alpha_i\|_2^2 + \lambda_1 \|\alpha_i\|_1\right) \\ & + \lambda_2 \left\| \frac{1}{N_0} \sum_{i=1}^{N_0} \alpha_i - \frac{1}{N_1} \sum_{j=N_0+1}^{N} \alpha_j \right\|^2 \end{aligned} \quad (3)$$

where $D \in \mathcal{R}^{d\times K}$ and $A = [\alpha_1, \cdots \alpha_N] \in \mathcal{R}^{K\times N}$ are the learned dictionary and encoding coefficient matrix, respectively, λ_1 and λ_2 are two parameters to balance the effects of different terms. Specifically, $A = [A^{tr} A^{au}]$, where $A^{tr} \in \mathbb{R}^{K\times N_\bullet}$ and $A^{au} \in \mathbb{R}^{K\times N_1}$ are the sparse codes of the samples in the training set and auxiliary set, respectively.

The first term in (3) is to ensure that the learned dictionary D can sparsely reconstruct the augmented dataset and the reconstruction error is minimized. The second term in (3) is to ensure that the distribution difference between the sparse codes of the training set and those of the auxiliary set is minimized so that the intra-class variations in the auxiliary set can be well transferred into the training set. In our study, the maximum mean discrepancy (MMD) criterion [9] which computes the distance between the mean of the samples from the training set and the auxiliary set is applied, which is an effective nonparametric similarity measure to compare two distributions.

The second term of (3) can be simplified as follows:

$$\sum_{i,j=1}^{N} \alpha_i^T \alpha_j S_{ij} = \mathrm{tr}(ASA^T)$$

where S is an affinity matrix defined as:

$$S_{ij} = \begin{cases} \frac{1}{N_0^2}, & x_i, x_j \in X^{tr} \\ \frac{1}{N_1^2}, & x_i, x_j \in X^{au} \\ -\frac{1}{N_0 N_1}, & \text{otherwise} \end{cases} \quad (4)$$

Then, the objective function of (3) can be re-written as:

$$\begin{aligned} \min_{D,A} \quad & (\|X - DA\|_F^2 + \lambda_1 \|A\|_1) + \lambda_2 \mathrm{tr}(ASA^T) \\ \text{subject to} \quad & \|d_k\|^2 \leq 1, 1 \leq k \leq K. \end{aligned} \quad (5)$$

While (5) is not convex for D and A simultaneously, it is convex for D when A is fixed and it is also convex for A when D is fixed. Here, we propose an alternating optimization method to iteratively optimize D and A. For a fixed D, the optimization function can be decomposed into individual ℓ_1-regularized least square problem for α_i as follows:

$$\begin{aligned} \min_{\alpha_i} \quad & \sum_{i=1}^{N} (\|x_i - D\alpha_i\|_2^2 + \lambda_1 \sum_{j=1}^{K} \|\alpha_i^{(j)}\|) \\ & + \lambda_2 (\alpha_i^T \alpha_i S_{ij} + \alpha_i^T \mathbf{t}) \end{aligned} \quad (6)$$

where $\mathbf{t} = 2\sum_{j\neq i} S_{ij}\alpha_j$ and $|\alpha_i^j|$ is the jth element of α_i. We use the feature sign search algorithm [10] to optimize α_i in (6).

For a fixed A, we can update the dictionary D by optimizing the following objective function

$$\begin{aligned} \min_{D} \quad & \|X - DA\|_F^2 \\ \text{subject to} \quad & \|d_k\|^2 \leq 1, 1 \leq k \leq K. \end{aligned} \quad (7)$$

We use the conjugate gradient descent method in [11] to optimize D in (7).

Having learned the dictionary D, the training set X^{tr} and the auxiliary set X^{au} are represented by A^{tr} and A^{au}, respectively. We compute the intra-class variation matrix V_A^{au} similar to [2] in the sparse codes space A^{au}. Given a test image y, we use sparse coding to transform y into α_y by solving (2). Then, the SRC-based SSFR model in (1) can be re-written as

$$\begin{aligned} \alpha_y &= A^{tr}\beta + V_A^{tr}\gamma + z \\ &= A^{tr}\beta + V_A^{au}\gamma + z \end{aligned} \quad (8)$$

Here, V_A^{tr} can be approximated V_A^{au} because A_{tr} and A_{au} are transformed into the same space with the learned dictionary D.

Lastly, the test sample y can be classified as follows:

$$c = \arg\min_i r_i(\alpha_y) \quad (9)$$

where

$$r_i(\alpha_y) = \left\| \alpha_y - \begin{bmatrix} A^{tr} & V_A^{au} \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \end{bmatrix} \right\|_2 \tag{10}$$

3.2. Discriminative Domain Transfer SRC

Since DT-SRC is a unsupervised learning approach, it will be more effective if the label information of the samples in the auxiliary set can be used. In this subsection, we propose a discriminative DT-SRC (DDT-SRC) method that exploits the label information for classification. Our DDT-SRC method is formulated as the following optimization problem:

$$\begin{aligned} \min_{D,A} \quad & (\|X - DA\|_F^2 + \lambda_1 \|A\|_1) + \lambda_2 \text{tr}(ASA^T) \\ & + \lambda_3 \sum_{i,j=1}^{N} \|\alpha_i - \alpha_j\|^2 W_{ij} \end{aligned} \tag{11}$$

where W is an affinity matrix to measure the sample similarity, which is defined as:

$$W_{ij} = \begin{cases} 1, & x_i \text{ and } x_j \text{ are from the the auxiliary set} \\ & \text{and they are from the same class} \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

λ_3 is a parameter to balance the different contributions of different terms in the objective function.

The third term in (11) is to ensure that if two samples are from the same class, their sparse codes in the meta-space should be as similar as possible. The third term in (11) can be simplified as follows:

$$\sum_{i,j=1}^{N} \|\alpha_i - \alpha_j\|^2 W_{ij} = \text{tr}(ALA^T) \tag{13}$$

where $L = \Delta - W$, $\Delta = diag(\delta_1, \cdots, \delta_N)$ and $\delta_i = \sum_{i=1}^{N} W_{ij}$.

Then the objective function in (11) can be re-written as

$$\begin{aligned} \min_{D,A} \quad & (\|X - DA\|_F^2 + \lambda_1 \|A\|_1) + \lambda_2 \text{tr}(ASA^T) \\ & + \lambda_3 \text{tr}(ALA^T) \\ \text{subject to} \quad & \|d_k\|^2 \le 1, 1 \le k \le K. \end{aligned} \tag{14}$$

Let $S' = S + \frac{\lambda_3}{\lambda_2} L$. (14) can be re-written as the same as (5) by replacing S with S'. Similar to DT-SRC, we also apply an alternating optimization method to iteratively optimize D and A by using (6) and (7), and the classification can be performed as the same as that in DT-SRC.

4. EXPERIMENT

4.1. Datasets

The Multi-PIE dataset was captured in four sessions with variations of pose, illumination and expression. There are 249 subjects whose images were collected in Session 1. Among the 249 subjects, the first 100 subjects were selected as the gallery set and the remaining 149 subjects were selected as the auxiliary set used for experiments on the other datasets. For the gallery set, face image of each subject with the frontal pose, illumination 7 and neutral expression was selected as the training set. Among these 100 subjects in Session 1, there are 75, 73 and 79 subjects whose images were also captured in Sessions 2, 3 and 4 (denoted as S2, S3 and S4), respectively. Facial expressions in these three sessions are surprise, smile and neutral, respectively. We selected frontal face images with all the 20 illuminations as the testing set. Hence, there are 1500, 1460 and 1580 images in the testing set for S2, S3, and S4, respectively.

There are 13539 facial images corresponding to 1565 subjects in the FERET dataset. In our experiments, we follow the standard FERET evaluation protocol [12], where five subsets including Fa, Fb, Fc, Dup1, and Dup2 were constructed for SSFR experiments. The Fa set contains 1196 frontal images of 1196 subjects, one image per subject, Fb contains 1195 images of the subjects from Fa but with different facial expressions, Fc contains 194 images captured under different illuminations, Dup1 contains 722 images and Dup2 contains 234 images taken in two different sessions. We took the Fa set as the training set and the other four sets as the testing sets.

The LFW dataset contains 13233 face images of 5749 subjects collected from the web under uncontrolled conditions. Since this dataset was captured in unconstrained environments, there are many natural variations such as pose, illumination, expression, race, occlusion and background. In our experiments, 158 subjects who have more than 9 distinct photos are selected for evaluation. For those subjects who have more than 10 images, we selected the first 10 images for our experiment. For each subject in this dataset, the first image was selected for training and the remaining were used for testing.

4.2. Experimental Settings

We manually cropped and aligned each image into 100×80 according to the eye positions for all images in the above four datasets. In our experiments, we down-sampled each image into 25×20 and used the raw intensity values as the feature representation for classification[1].

There are two parameters λ_1 and λ_2 in DT-SRC and three parameters λ_1, λ_2 and λ_3 in DDT-SRC, where λ_1 controls the sparsity of the representation coefficients, λ_2 and λ_3 regularize the effects of different terms, respectively. In our experiments, we empirically fix $\lambda_1 = 5$, $\lambda_2 = 1000$, and $\lambda_3 = 0.01$ by using the cross-validation strategy on Multi-PIE training dataset. The number of atoms K in DT-SRC and DDT-SRC was also required to pre-specify. In our implementations, K was set as 200 for all datasets. We constructed the auxiliary

[1] Other handcrafted features may be applied.

Table 1. Rank-1 recognition rates (%) of different SSFR methods on the Multi-PIE and FERET datasets.

Method	Multi-PIE			FERET			
	S2	S3	S4	fb	fc	dup1	dup2
NN	5.1	4.1	8.5	82.0	1.6	29.0	12.8
SVM	11.4	9.2	16.9	91.9	65.5	63.3	55.1
SRC [1]	23.3	25.0	30.3	91.3	52.6	58.0	45.7
ESRC [2]	32.1	39.2	46.4	90.1	88.7	69.4	63.7
SSRC [3]	30.1	37.0	44.7	90.0	86.1	69.5	62.8
SVDL [4]	32.6	32.7	45.2	92.5	92.9	73.7	**70.5**
DT-SRC	34.1	38.7	52.0	**92.7**	90.7	72.0	67.5
DDT-SRC	**34.5**	**40.9**	**53.8**	92.2	**93.3**	**75.1**	70.1

set from the Multi-PIE dataset for SSFR experiments on the FERET and LFW datasets. Specifically, we selected 149 subjects (ID index from 101 to 249) from Session 1 of the Multi-PIE dataset. For each subject, we selected face images with 7 different illuminations (illumination 0, 1, 7, 13, 14, 16 and 18) and the neutral expression to construct the auxiliary set for the FERET dataset. Hence, there are 1043 (=149×7) images in the constructed auxiliary set. Since there are large variations in face images from the LFW datasets, we included more face images with different poses from the Multi-PIE dataset to construct the auxiliary set for the LFW dataset. Besides the above mentioned 1043 images, face images with 2 different poses (pose 4-0 and 5-1) and the neutral expression were also added in the auxiliary set. To better model the pose variation, we further flipped these images and combined them to construct an augmented auxiliary set for SSFR experiments on the LFW dataset. For the SSFR experiments on the Multi-PIE dataset, we selected the AR dataset [13] as the auxiliary set, where face images captured in the first session with no occlusion of the whole 100 subjects were selected. Hence, there are 700 (=100×7) images in this auxiliary set.

4.3. Results and Analysis

Comparison with State-of-the-Art SSFR Methods: We compared our proposed methods with six methods which can also address the SSFR problem, including nearest neighbor (NN), support vector machine (SVM), SRC [1], ESRC [2], SSRC [3], and SVDL [4]. Among these six methods, there is no auxiliary set required for NN, SVM, and SRC, and the other three methods such as ESRC, SSRC and SVDL need an auxiliary set to model the intra-class variations. We implemented these methods ourselves except SVM and SVDL because the codes of these two methods are publicly available. For a fair comparison, we tuned the best possible parameters for each method. There is no parameter for NN. For SVM, the RBF kernel was selected to compute the sample similarity and the libsvm toolbox [2] was used in our experiments. For SRC, ESRC and SSRC, we followed the parameter setting as suggested in the original papers [1], [2], [3], where the sparsity parameter is set as 0.05. For SVDL, we followed the same settings in [4] where λ_1, λ_2 and λ_3 were set as 0.001, 0.001 and 0.01[3], and the atom number in the learned dictionary was set as 200, respectively. For all SRC-based methods, the homotopy method [14] was employed to solve the ℓ_1-minimization problem due to its excellent computational efficiency [15]. Table 1 shows the rank-one recognition rates of different SSFR methods on the Multi-PIE and FERET datasets, and Fig. 2 shows the cumulative match curve (CMC) of different SSFR methods on the LFW dataset. We see that our proposed methods consistently outperform the other compared methods in all four datasets. It is important to note that while our performance in these experiments are not state-of-the-art compared to current face recognition methods, we only used the downsampled features to easily show the effectiveness of our approach of exploiting auxiliary datasets for SSFR. Fig. 3 shows two specific recognition examples where both SRC and ESRC fail to work but our DT-SRC succeeds. For the two given test samples in this figure, SRC cannot sparsely reconstruct the test samples because the number of training samples is insufficient in the training set. Since the auxiliary set is constructed from the AR dataset and the variations in the AR dataset are different from those in the Multi-PIE dataset because they are captured from different conditions (different illuminations and expressions), the intra-class variations modeled from the AR dataset are not appropriate to use for the Multi-PIE dataset directly in ESRC. Unlike SRC and ESRC, our proposed DT-SRC successfully recognizes the testing samples because we transferred both the training set and the auxiliary set into a meta-space so that the intra-class variations learned from the AR dataset can be applied to the Multi-PIE dataset.

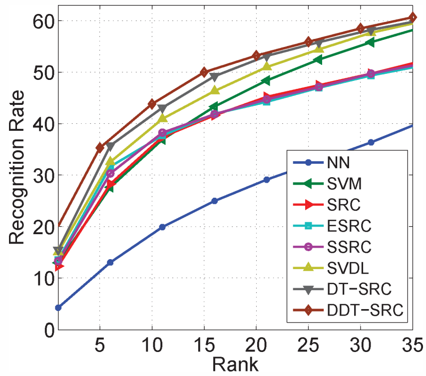

Fig. 2. Cumulative match curve (CMC) of different SSFR methods on the LFW dataset.

Effect Analysis of Different Auxiliary Sets: We investigated the effects of using different auxiliary sets in our proposed DT-SRC and DDT-SRC. Table 2 shows the rank-one recognition rate when different auxiliary sets were used for

[2]http://www.csie.ntu.edu.tw/ cjlin/libsvm/

[3]Parameter sensitivity detail is shown in the Supplementary Material

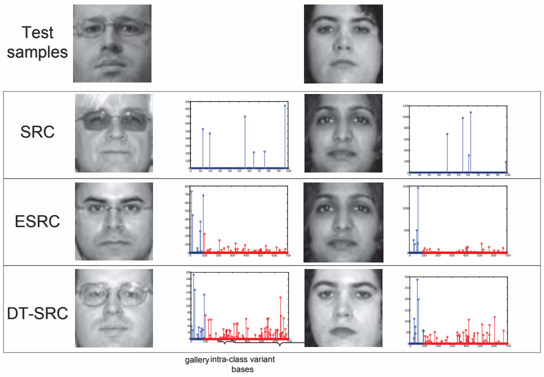

Fig. 3. Two recognition examples where both SRC and ESRC fail to work but our DT-SRC succeeds. The first row shows two testing face images from the S4 set of the Multi-PIE dataset. The second to fourth rows show the recognized subjects, as well as the representation coefficients obtained by SRC, ESRC and DT-SRC, respectively. Best view in color pdf file.

Table 2. Rank-one recognition rates (%) of our proposed methods when different auxiliary sets were used for the SSFR experiments on the CMU Multi-PIE dataset.

Method	Auxiliary set	S2	S3	S4
DT-SRC	AR	34.3	37.3	53.7
DT-SRC	FERET	31.4	28.6	48.7
DT-SRC	AR+FERET	**37.3**	**39.1**	**55.6**
DDT-SRC	AR	35.2	38.7	54.4
DDT-SRC	FERET	35.9	31.0	50.1
DDT-SRC	AR+FERET	**38.5**	**40.4**	**57.1**

SSFR experiments on the CMU Multi-PIE dataset. As shown in this table, using two auxiliary sets achieves higher recognition rates than using a single one. That is because two different auxiliary sets could provide different types of variations which are better to model the intra-class variations. Hence, better recognition performance can be obtained for both DT-SRC and DDT-SRC.

Parameter Analysis: We also investigated the performance of our methods versus different codebook sizes. Fig. 4 shows the rank-one recognition rates of our DT-SRC and DDT-SRC methods on the CMU Multi-PIE and FERET datasets. It can be seen that our method is not sensitive to the codebook size and best recognition rates can be obtained when the codebook size was set as 200 to 300.

5. CONCLUSION

In this paper, we have presented an SRC-based SSR method which uses an auxiliary set to help model the intra-class variation in the SRC-based classification framework. From the experimental results, our proposed DT-SRC and DDT-SRC outperform the other SRC-based SSFR method such as ESRC, SSRC and SVDL when the auxiliary set and the training set are collected from different conditions.

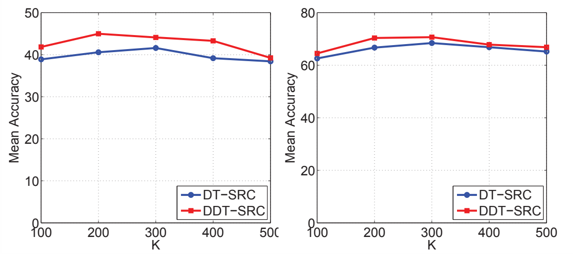

Fig. 4. Rank-one recognition rate (%) of our proposed methods versus different values K on the CMU Multi-PIE (left) and FERET (right) datasets, respectively.

6. REFERENCES

[1] John Wright, Allen Y Yang, Arvind Ganesh, Shankar S Sastry, and Yi Ma, "Robust face recognition via sparse representation," *PAMI*, vol. 31, no. 2, pp. 210–227, 2009.

[2] Weihong Deng, Jiani Hu, and Jun Guo, "Extended src: Undersampled face recognition via intraclass variant dictionary," *PAMI*, vol. 34, no. 9, pp. 1864–1870, 2012.

[3] Weihong Deng, Jiani Hu, and Jun Guo, "In defense of sparsity based face recognition," in *CVPR*, 2013, pp. 399–406.

[4] Meng Yang, Luc Van Gool, and Lei Zhang, "Sparse variation dictionary learning for face recognition with a single training sample per person," in *ICCV*, 2013, pp. 689–696.

[5] Yu Su, Shiguang Shan, Xilin Chen, and Wen Gao, "Adaptive generic learning for face recognition from a single sample per person," in *CVPR*, 2010, pp. 2699–2706.

[6] Sinno Jialin Pan, James T Kwok, and Qiang Yang, "Transfer learning via dimensionality reduction.," in *AAAI*, 2008, vol. 8, pp. 677–682.

[7] Lixin Duan, Ivor W Tsang, Dong Xu, and Stephen J Maybank, "Domain transfer svm for video concept detection," in *CVPR*, 2009, pp. 1375–1381.

[8] Yu Zhang and Dit-Yan Yeung, "Transfer metric learning with semi-supervised extension," *ACM TIS)*, vol. 3, no. 3, pp. 54, 2012.

[9] Sinno Jialin Pan, James T Kwok, and Qiang Yang, "Transfer learning via dimensionality reduction.," in *AAAI*, 2008, pp. 677–682.

[10] Honglak Lee, Alexis Battle, Rajat Raina, and Andrew Ng, "Efficient sparse coding algorithms," in *NIPS*, 2006, pp. 801–808.

[11] Quoc V Le, Alexandre Karpenko, Jiquan Ngiam, and Andrew Y Ng, "Ica with reconstruction cost for efficient overcomplete feature learning.," in *NIPS*, 2011, pp. 1017–1025.

[12] P Jonathon Phillips, Hyeonjoon Moon, Syed A Rizvi, and Patrick J Rauss, "The feret evaluation methodology for face-recognition algorithms," *PAMI*, vol. 22, no. 10, pp. 1090–1104, 2000.

[13] Aleix M Martinez, "The ar face database," *CVC Technical Report*, 1998.

[14] David L Donoho and Yaakov Tsaig, "Fast solution of l1-norm minimization problems when the solution may be sparse," *TIT*, vol. 54, no. 11, pp. 4789–4812, 2008.

[15] A.Y. Yang, Zihan Zhou, A.G. Balasubramanian, S.S. Sastry, and Yi Ma, "Fast ℓ_1-minimization algorithms for robust face recognition," *TIP*, vol. 22, no. 8, pp. 3234–3246, 2013.