Random Sparse Representation for Thermal to Visible Face Recognition

Samira Reihanian‡, Ehsan Arbabi‡ and Behrouz Maham‡
†School of ECE, University of Tehran, Iran
‡School of Engineering, Nazarbayev University, Kazakhstan
Emails: samira.reihanian@nu.edu.kz,earbabi@ut.ac.ir,behrouz.maham@nu.edu.kz

Abstract—Heterogeneous face recognition (HFR) has a prominent importance in sophisticated face recognition systems. Thermal to visible scenario, where the gallery and the probe images are respectively captured in visible and long wavelength infrared (LWIR) band, is one of the most challenging and interesting HFR scenarios. Since the formation of thermal images does not require an external illumination source, the deployment of thermal probe images is practical even in totally darkness conditions such as night security surveillance systems. In this paper, we propose an ensemble classifier which uses the random subspace idea for defining different representations of each image in distinct base learners, and exploits the sparse representation algorithm for the classification of thermal probe images. According to the experimental results, our proposed algorithm leads significant performance improvements in the area of thermal to visible face recognition and achieves the average Rank-1 accuracy of 89.33 percent.

I. INTRODUCTION

Although visible images have traditionally been used in the mechanized face recognition systems, recently researchers are interested in utilizing other parts of electromagnetic spectrum such as long wavelength infrared (LWIR) (7 - 14 μm) because of their special characteristics [1]. Each object, depending on its temperature and emissivity characteristics, emits different ranges of infrared energy. Characteristics of human face and body temperature cause the emission of the face to be in the thermal infrared band, especially in the LWIR band. Thermal infrared cameras produce thermal face images by sensing temperature variations in a face. Since the formation of the thermal image is dependent on intrinsic property of the face, an external illumination source is not required for the formation of this type of images. Thus, acquisition of the thermal images is possible even in total darkness conditions, where the formation of the visible image is absolutely impossible [2]. Thus, in night security surveillance systems, the deployment of thermal probe images is practical, where conventional face recognition systems have failed. However, in majority of the face recognition systems, stored face images in the systems, entitled gallery images, are in the visible modality. Therefore, matching between the visible gallery images and the thermal probe images can solve face recognition problems in variable illumination conditions and even in totally dark circumstances. However, because of the distinct formation mechanisms of these two types of images, there are lots of challenges in the matching process. Hence, thermal to visible face recognition is one of the most challenging heterogeneous face recognition (HFR) scenarios. Since successful algorithms for this scenario are expected to be effective in other HFR scenarios, developing recognition algorithms for this scenario has an enormous importance.

Until now, only few methods (see e.g., [3]-[7]) have been suggested for solving thermal to visible face recognition problem. The first solution for handling the thermal to visible face recognition problem was proposed by Li et al. [4]. This method synthesizes a visible pair of each thermal test image by training a canonical correlation analysis (CCA) model for each test subject. The best reported recognition result for the CCA based method (on a database consists of 47 subjects with 50.06% by performing the testing stage for each subject and considering all the thermal-visible pairs of other subjects as the training set [4]. Choi et al. proposed partial least squares-discriminant analysis (PLS-DA) based approach by correlating the thermal image signatures of a person to the visible image signatures of the same person [5]. For evaluating the PLS-DA based approach, a database consists of 41 subjects with multiple images for each subject was used in [5] and the reported recognition rate was 49.9%. Sarfraz et al. learned a non-linear mapping from thermal to visible spectrum by training a deep neural network [6]. For evaluating the proposed algorithm in [6], Sarfraz et al. considered 41 subjects with multiple images for training the model and 41 subjects for testing purpose. The Rank-1 recognition rate of the algorithm was 55.36%, when one visible image was considered for each test subject in the gallery set. However, when multiple visible images were considered for each test subject in the gallery, the Rank-1 recognition rate was reached to 83.73%. Klare and Jain presented prototype random subspace (P-RS) method for handling any heterogeneous and analogous face recognition scenarios [7]. The main idea of this method was concentrated on definition of the feature vector such that a new feature vector is defined for representation of each image by calculating kernel similarity between the initial feature vector of the image and the initial feature vectors of the training images. Klare and Jain used 333 and 667 test and train subjects, respectively (with one visible and one thermal image for each subject). However, the gallery set used in the experiments was augmented by 10000 different visible images.
from other different subjects. The reported recognition rate for the thermal to visible scenario in [7] was about 46.7%. Although there are some other methods in literatures (see, e.g., [6] and [3]) that yielded better Rank-1 recognition rate compared to the P-RS method, this method has a noticeable importance among other existing methods. The reason of this priority is the database utilized in [7] for the evaluation of this algorithm. It is very clear that by increasing the number of subjects in the gallery, the probability of the correct recognition is reduced. For example, in [6] and [3] only 41 and 55 subjects were considered in the gallery set, respectively. Moreover, Klare and Jain considered one visible image and one thermal image for each subject. Both considerations are more probable in the real world face recognition scenarios. Therefore, in this paper, the performance of our proposed method is compared with the P-RS method on behalf of the other existing methods in this area.

In this paper, we introduce an ensemble classifier [8] for handling the thermal to visible face recognition problem. For the first time, we apply sparse representation classification algorithm (SRC) [9], which is one of the latest successful robust face recognition algorithms and reconstruct the thermal image of a person by means of the visible images of the same person in the gallery set. Our proposed ensemble classifier uses the same learning algorithm in the base learners, while the representations of the images are different in the distinct learners by application of the random subspace idea [10] for definition of the feature vectors. We show that our proposed method can achieve a considerable improvement in terms of the average Rank-1 accuracy of the thermal to visible face recognition scenario.

The remainder of the paper is organized as follows. Section II briefly reviews sparse representation theory and the idea of applying this theory on HFR problems. In Section III, different stages of our proposed algorithm are described in detail. Experimental results are presented in Section IV. Finally, in Section V, we conclude the paper.

II. CLASSIFICATION BASED ON SPARSE REPRESENTATION

In this paper, we reconstruct thermal image of a person by the weighted sum of the visible images of the same person. To achieve this purpose, we exploit sparse representation theory [9]. So in this section, we briefly introduce sparse representation theory [9] and then present how to apply this representation in our heterogeneous face recognition algorithm.

A. Sparse Representation

Sparse representation [9] gives a general solution for the robust face recognition problems especially in harsh circumstances. The goal of the face recognition by sparse representation is to classify a probe sample, by possessing a gallery set that consists of \( N \) different samples from \( n_t \) distinct subjects. The \( i^{th} \) subject has \( n_i \) samples in the gallery set. Thus, the gallery set \( \mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_{n_t}] \), where \( \mathbf{A} \in \mathbb{R}^{m \times N} \), and \( \mathbf{A}_i \in \mathbb{R}^{m \times n_i} \) is associated to the gallery samples of the \( i^{th} \) subject. The sparse representation framework assumes different samples of the specific subject lie on a linear subspace [9]. Thus, by having the sufficient gallery samples of the \( i^{th} \) subject, any probe sample \( \mathbf{y} \) from the \( i^{th} \) class, can be reconstructed as \( \mathbf{y} = \mathbf{A}_i \mathbf{x}_i \), where \( \mathbf{x}_i \) is a \( n_i \)-dimensional vector with scalar elements. Since the class of the probe sample is not known yet, the probe sample can be reconstructed by means of the all gallery samples as \( \mathbf{y} = \mathbf{A} \mathbf{x} \) (where \( \mathbf{x} \in \mathbb{R}^N \), and all elements of \( \mathbf{x} \) are zero, except the elements associated with the \( i^{th} \) class, which are equal to elements of \( \mathbf{x}_i \)). Hence, the class of the probe sample will be determined easily, by calculation of \( \mathbf{x} \). However, for the classification purpose, only the elements of one class must be nonzero in \( \mathbf{x} \). So, in order to guarantee the sparsity of the \( \mathbf{x} \) vector, the desirable \( \mathbf{x} \), which is sparse enough, can be calculated as [9]:

\[
\hat{\mathbf{x}}_0 = \arg\min_{\parallel \mathbf{x} \parallel_0 \text{ subject to } \mathbf{y} = \mathbf{A} \mathbf{x}},
\]

where \( \parallel \cdot \parallel_0 \) denotes the \( l^0 \)-norm. When the system is underdetermined, i.e., \( m < N \), finding the solution of (4) is categorized as NP-hard problems. Fortunately, it is proved that if solution \( \hat{\mathbf{x}}_0 \) is sparse enough, the solution of the \( l^0 \)-minimization problem is equal to the solution of the \( l^1 \)-minimization problem as:

\[
\hat{\mathbf{x}}_1 = \arg\min_{\parallel \mathbf{x} \parallel_1 \text{ subject to } \mathbf{y} = \mathbf{A} \mathbf{x}},
\]

where \( \parallel \cdot \parallel_1 \) denotes the \( l^1 \)-norm and this problem can be solved in a polynomial time [9]. After calculating of the \( \hat{\mathbf{x}}_1 \), the classification of the test sample can be done. Although, if the elements of distinct classes are nonzero in \( \hat{\mathbf{x}}_1 \), classification will not be straightforward. In this situation, the problem can be solved by the definition of the residuals concept [9]. Thus, first approximations of the probe sample are constructed by considering the elements of \( \hat{\mathbf{x}}_1 \) associated with a single class. Assume \( \hat{\mathbf{z}}_i \in \mathbb{R}^N \) denotes the approximation of the \( \hat{\mathbf{x}}_1 \) by means of the class \( i \) elements, where all elements of the \( \hat{\mathbf{z}}_i \) are zero, except the elements associated with the \( i^{th} \) class, these elements are equal to the elements of \( \hat{\mathbf{x}}_1 \). Now, the approximation of the probe sample \( \mathbf{y} \) by means of the class \( i^{th} \) elements, can be calculated as \( \mathbf{A} \hat{\mathbf{z}}_i \). The difference between the probe sample and its approximation by each class determines the residual vector of that class. Finally, the probe sample is assigned to the class with the minimum residual value, i.e. the class with the minimum \( l^2 \)-norm of the residual vector.

B. Sparse Representation for Heterogeneous Face Recognition

In this section, for the first time we introduce applying the sparse representation in the heterogeneous face recognition problems. In the sparse representation approach, the probe image of a person is reconstructed by means of the gallery images of the same person. However, in the heterogeneous face recognition problems, the probe and gallery images are not in the same modality and there is a significant difference between these two types of images. Although this difference is not caused by variable lighting or illumination conditions, the probe image of a person can be considered as the destroyed
III. PROPOSED RANDOM SPARSE REPRESENTATION METHOD

In this section, random sparse representation (RSR) algorithm has been introduced. First, we have preprocessing step. Then, similar to the other classification methods, our algorithm includes two main stages: the training and testing stages.

A. Preprocessing and Feature Extraction Steps

In this paper, the width and the height of the all images have been set to 200 and 240 pixels respectively, the eyes are centralized horizontally at row 115, and the distance between two pupils is set to 75 pixels. Then, by means of the min-max normalization, the minimum and maximum intensity values of each image are set to 0 and 255, respectively.

Existence of a filtering step in the preprocessing stage seems essential in the face recognition algorithms especially heterogeneous cases in order to reduce the modality gap between the probe and gallery images of a subject. Thus, in this study, we have employed two conventional filters as center-surround divisive normalization (CSDN) [11] and difference of Gaussians (DoG) for increasing similarity between the probe and the gallery images of a subject. For applying CSDN filter to each image, the intensity value of each pixel in the image is divided into the mean intensity value of the neighboring pixels within a $s \times s$ neighborhood area [11]. In experiments of this paper $s = 16$. For construction of the DoG filter, a Gaussian filter with smaller width, $\sigma_1 = 2$, is subtracted from another Gaussian filter with larger width, $\sigma_2 = 4$.

Before starting the feature extraction step, all the normalized and filtered images are divided uniformly to some overlapping patches (each patch contains $32 \times 32$ pixels with 50% overlapping with vertical and horizontal neighbors). Then a feature vector extracted from each patch. In this paper we use uniform local binary pattern (LBP) [12] for the feature extraction purpose, which has been used effectively in the heterogeneous face recognition problems [7]. In this paper $n = 8$ neighbors are considered around each pixel at radius $r = 1, 3, 5, 7$. Thus, by considering four different values for radius, and concatenation of the four resulted feature vectors, dimension of the final feature vectors will be 236.

B. Ensemble Classifier

In this paper we use an ensemble classifier [8] consists of a set of individual classifiers, which are entitled base learners. Combination of the decisions of different base learners is used for classification of the test samples. For construction of the ensemble classifier, we use manipulation of the input features, i.e., the base learners utilize the same learning algorithm, but the feature vectors used by distinct learners are different. For the formation of the different feature vectors which are exploited by the base learners, we use random subspace theory [10]. First, all patches of each image are divided to $B$ bags randomly, meanwhile each bag includes a fraction of all the patches, (i.e. each bag is constituted by $\eta$ percent of all the patches) and it is possible for a patch to be in more than one bag or none of the bags. Then, cascading the feature vectors of the patches of each bag represents the final feature vector of the image in the same bag. Since the feature vectors associated with each image are different in distinct bags, the representations of each image will be different in distinct bags. These different representations can be used in an ensemble classifier by the base learners for determining the class of the test sample. In addition, the employment of random subspace theory for the definition of the feature vectors yields reduction of the feature space dimension. As described in section II, in the formulation of the sparse representation algorithm, dimension of the feature space is considered smaller than the number of the images in the gallery set.

C. Training Stage

The training stage of our approach involves two main steps: Principal Component Analysis (PCA) [13] and Linear Discriminant Analysis (LDA) [14]. The execution of these two steps helps to dimension reduction of the feature space, as well as enhancement of the discrimination between the different classes. Note that sparse representation necessitates the reduction of the feature space dimension.

Assume our training set consists of different thermal and visible samples from $n_{train}$ distinct subjects. The $l^{th}$ subject has $c_l$ thermal and visible samples in the training set. Thus, the total training set of the $l^{th}$ base learner, i.e. $S_l$, can be constructed as:

$$S_l = \{F_l(T^1_l), F_l(V^1_l), \ldots, F_l(T^n_{train}), F_l(V^n_{train})\},$$

where $F_l(T^j_l)$ and $F_l(V^j_l) \in \mathbb{R}^m$ indicate the final feature vectors of the $i^{th}$ thermal and visible images of the $j^{th}$ subject extracted by the $l^{th}$ base learner, respectively. By the execution of PCA algorithm, the mean vector (i.e. $\mu_l \in \mathbb{R}^m$) and the PCA mapping matrix, i.e., $W_{lPCA} \in \mathbb{R}^m \times \hat{m}$ are calculated for $l = 1, 2, \ldots, B$, which $\hat{m}$ is the dimension of the resultant low dimensional space. In our experiments, 97 percent of the variance is retained in the PCA step. Although there is a remarkable difference between the thermal and visible images, by considering the feature vectors of these two types of images in the total training set, the variances of both types of images affect the calculation of the mapping matrix, which can help...
more reduction of the differences between the thermal and visible images. After the calculation of $W_l^{\text{PCA}}$ matrix, we use LDA method and calculate LDA mapping matrix, i.e., $W_l^{\text{LDA}} \in \mathbb{R}^{d_l \times n_{\text{train}} - 1}$ in order to further separation of the distinct classes. For performing the LDA algorithm, different thermal and visible images of a training subject are considered in the same class. Thus, by performing the PCA and LDA steps, a total mapping matrix, i.e., $W_l^{\text{Total}}$, is calculated for each of the base learners as:

$$W_l^{\text{Total}} = W_l^{\text{PCA}} \times W_l^{\text{LDA}}. \quad (4)$$

D. Testing Stage

As we know in the thermal to visible face recognition problem, the visible images are available as the gallery set and we want to classify the thermal probe samples. In this study, we assume at least one visible image for each test subject, exists in the gallery set. By applying sparse representation algorithm in the thermal to visible heterogeneous face recognition problem, the thermal probe image is reconstructed by means of the visible gallery images associated with the test person. Suppose $B$ is the number of the bags which is equivalent to the number of the base learners in the ensemble classifier and $Q$ different visible images from $n_{\text{test}}$ distinct subjects are available, and the $i$th subject has $d_i$ different visible images in the gallery set. If $F_l^j(y) \in \mathbb{R}^{d_i}$ indicates the feature vector extracted from $y$ by the $l$th base learner, then gallery set of the $l$th base learner, $G_l$, defines as follows

$$G_l = [F_l^1(V_1^j), \ldots, F_l^1(V_{d_i}^j), \ldots, F_l^Q(V_1^{n_{\text{test}}}), \ldots, F_l^Q(V_{d_i \times n_{\text{test}}})], \quad (5)$$

where $F_l^j(V_j^i) \in \mathbb{R}^{n_{\text{train}} - 1}$ is the final feature vectors of the $i$th visible image of the $j$th test subject by the $l$th base learner that is defined as $F_l^j(V_j^i) = W_l^{\text{Total}}(F_l^j(V_j^i) - \mu_l)$. Moreover if $F_l^j(q) \in \mathbb{R}^{d_i}$ indicates the feature vector extracted from the thermal test image $q$ by the $l$th base learner, $F_l^j(q) \in \mathbb{R}^{n_{\text{train}} - 1}$ is the final feature vectors of the thermal test image used by the $l$th base learner, which is defined as $F_l^j(q) = W_l^{\text{Total}}(F_l^j(q) - \mu_l)$.

By performing sparse representation algorithm for $F_l^j(q)$ in each base learner, the residual values are calculated for different test classes. Since our gallery set contains the visible images of $n_{\text{test}}$ distinct subjects, each base learner calculates $n_{\text{test}}$ distinct residual values for each thermal probe image. Hence, the $l$th base learner, calculates $r_i^l(q)$ for $i = 1, 2, \ldots, n_{\text{test}}$, where $r_i^l(q)$ indicates the residual value of the $i$th class for the thermal test image $q$, which is calculated by the $l$th base learner. Then by combining the residuals of all base learners, class of the test sample can be determined. Thus, first we define total residuals for different classes as:

$$R_l^i(q) = \sum_{l=1}^{B} r_i^l(q), \quad \text{for } i = 1, 2, \ldots, n_{\text{test}} \quad (6)$$

where $R_l^i(q)$ is the total residual of the $i$th class for the thermal test sample $q$, which is defined as the sum of the residuals of the $i$th class by all of the base learners for the thermal test image $q$. The total residuals of different classes can be exploited for final decision about class of the test sample. Fig. 1 demonstrates different steps of the testing stage for calculation of the total residual values in our proposed algorithm. Since we apply two different filters on the thermal and visible images in the preprocessing step, by considering same filter for the probe and gallery images, two distinct residuals can be calculated for the thermal test sample, that we entitle them as $R_{\text{CSDN}}^i(q)$ and $R_{\text{DoG}}^i(q)$. In this paper, fusion strategy is used for the combination of the residuals of two filters. So first, the total residuals of these two filters are added together to form $R_{\text{fusion}}^i$, i.e.

$$R_{\text{fusion}}^i(q) = R_{\text{CSDN}}^i(q) + R_{\text{DoG}}^i(q), \quad \text{for } i = 1, 2, \ldots, n_{\text{test}}. \quad (7)$$

Then, the thermal test sample $q$ is assigned to the class with the minimum $R_{\text{fusion}}^i(q)$ value. Thus, in the fusion strategy, the predicted class of the thermal test sample $q$, which is denoted as $C_{\text{fusion}}(q)$ can be determined as

$$C_{\text{fusion}}(q) = \arg \min_{i} R_{\text{fusion}}^i(q). \quad (8)$$

IV. EXPERIMENTS

A. Database Description

For evaluation accuracy of our proposed algorithm, we select 60 subjects with three pairs of thermal and visible

\footnote{1IEEE OTCBVS WS Series Bench: DOE University Research Program in Robotics under grant DOE-DE-FG02-86NE37968; DOD/TACOM/NAC/ARC Program under grant R01-1344-18; FAA/NSSA grant R01-1344-48/49; Office of Naval Research under grant N000143010022.}
images for each subject from two thermal-visible databases. We use "Dataset 02: IRIS Thermal/Visible Face Database" subset of Object Tracking and Classification Beyond the Visible Spectrum (OTCBVS) database \(^1\) \([15]\), and select 22 subjects with three pairs of thermal and visible images for each subject, however, each pair has different expression (surprised, laughing or angry). Furthermore, we use the Natural Visible and Infrared Facial Expression Database (USTC-NVIE) \([16]\) as well, collected by CCSL of China \([17]\). We select 38 subjects from the posed subset with three pairs of thermal and visible images for each subject, such that each pair has different expression (happiness, anger, sadness, fear, onset, disgust or surprised). Note that all of the thermal and visible images for all subjects are taken in front view.

**B. Results and Discussion**

P-RS method \([7]\) is one of the most successful methods for solving the thermal to visible face recognition problem. Therefore, in this paper the performance of our proposed algorithm is compared with the performance of the P-RS method on behalf of the other existing methods in this area. The P-RS method is implemented precisely in the experiments of this paper, but since we consider three thermal-visible pairs for each subject, the number of the prototypes is three times of the number of the training subjects and each class is represented by six samples in the execution of the LDA step. Moreover, Since we exploit sparse representation in our proposed algorithm, we also compare its performance with the original classification method based on sparse representation (SRC) \([9]\). This method has not been applied yet in thermal to visible face recognition problem. Note that SRC method does not require a separate training set and the visible images of the test subjects are used as the gallery set and thermal probe image is reconstructed by means of the visible gallery images. First, each image is treated as one feature vector simply by concatenating the rows of pixels in the original image. Then, the original feature vectors of the gallery set is projected into a lower dimensional space by means of Eigenfaces \([13]\) or other common algorithm in reduction of the feature space dimension. In this paper by applying Eigenfaces algorithm \([13]\) on the gallery set, the feature vectors of the gallery images and thermal probe image are projected into the lower dimensional space. Then, the resultant low dimensional feature vectors are used by the classification algorithm.

In the experiments of this paper, 30 subjects are used for the training set and the remaining 30 subjects are considered as the testing set. To confirm the validity of the experimental results, the training and testing steps of all algorithms are repeated 10 times, each time the training and testing sets are selected randomly. Thus, the reported recognition rate is average of the obtained results. Note that in the testing stage of the RSR and SRC \([9]\) algorithms, the gallery set is formed by means of the visible images of the 30 test subjects (each subject has three visible images). So our gallery set in this situation consists of 90 feature vectors which belong to the visible images of 30 different test subjects. Since in the testing stage of the P-RS method, just one visible image is needed for each test subject, the gallery set consists of the 30 visible images of the 30 test subjects (one visible image for each test subject). Each time a feature vector associated with the thermal image of a test subject is presented to the algorithm for the purpose of classification.

The reported results for the proposed RSR algorithm are based on the following parameter values: \(B = 30\) and \(\eta = 5\) and as mentioned before, 97 percent of the variance is retained in the PCA step. For validation of the results of the P-RS method, except the number of the images for each training subject, other parameters of the P-RS method are considered as \([7]\) exactly. In the simulation of SRC method \([9]\), the dimension of the feature space is reduced by applying Eigenfaces \([13]\) algorithm. Since the accuracy of the original sparse representation method \([9]\) depends on the size of the feature vectors or equivalently number of the eigenfaces used in the algorithm, in this paper, we tune the number of the eigenfaces in order to maximize the accuracy of the SRC method. Average Rank-1 recognition rates of our proposed algorithm (RSR), the P-RS and the SRC methods are summarized in Table I. According to the results, our proposed method achieves the average Rank-1 accuracy of 89.33 percent. By comparison, P-RS achieves the average Rank-1 accuracy of 69.33 percent on our database. As Klare and Jain mentioned in \([7]\), the performance of the P-RS method is highly dependent on the number of prototypes. By increasing the number of prototypes, first accuracy of the P-RS method is improved, and then, its accuracy is saturated. According to Table I., when number of training subjects is limited (and as a sequence number of the prototypes is limited), our proposed algorithm outperforms P-RS and improves the recognition rate considerably. The sparse representation classification method (SRC) by applying Eigenfaces \([13]\) algorithm for reduction the dimension of the feature space, achieves the average Rank-1 accuracy of 14% at the best manner by considering 42 eigenfaces. Although

---

**TABLE I**

**RECOGNITION RESULTS FOR PROPOSED METHOD (RSR) AND TWO BASELINES: P-RS AND SRC.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank-1 accuracy (%)</td>
<td>89.33±6.34</td>
<td>69.33±7.66</td>
<td>14.00±7.17</td>
</tr>
</tbody>
</table>

---

Fig. 2. CMC plots of the performance of the proposed method (RSR) and two baselines: P-RS and SRC.
for handling the challenging heterogeneous face recognition scenario, thermal to visible face recognition. Decisions of the different base learners are combined together for prediction the class of the thermal test sample. The Rank-1 recognition results demonstrate considerable improvement of around 20 percent in comparison to the most successful algorithm in this area, i.e., (P-RS). The utilization of the sparse representation theory in handling the thermal to visible face recognition problems, and describing the thermal probe image of a person by means of linear combination of the visible gallery images of the same person is one of novelties of this work which yields significant improvement in Rank-1 recognition rate. Note that since the thermal to visible scenario is one of the most challenging HFR scenarios, we expect our proposed algorithm to perform well in other HFR scenarios as well. Therefore, it seems this paper reveals a general framework for HFR problems. The future work will focus on testing our proposed algorithm on other HFR scenarios.

V. CONCLUSION

In this paper, sparse representation classification algorithm is regularized for heterogeneous cases, and is utilized by the different base learners in our proposed ensemble classifier

<table>
<thead>
<tr>
<th>Type of the Filter</th>
<th>CSDN</th>
<th>DoG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank-1 accuracy (%)</td>
<td>84.33±8.61</td>
<td>75.33±9.32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of the Filter</th>
<th>CSDN</th>
<th>DoG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank-1 accuracy (%)</td>
<td>22.98±8.88</td>
<td>21.03±9.02</td>
</tr>
</tbody>
</table>

REFERENCES


