Joint Sparsity-based Robust Multimodal Biometrics Recognition

Sumit Shekhar[1], Vishal M. Patel[1], Nasser M. Nasrabadi[2] and Rama Chellappa[1]

[1]University of Maryland, College Park, USA, [2]Army Research Lab, Adelphi, USA

Abstract. Traditional biometric recognition systems rely on a single biometric signature for authentication. While the advantage of using multiple sources of information for establishing the identity has been widely recognized, computational models for multimodal biometrics recognition have only recently received attention. We propose a novel multimodal multivariate sparse representation method for multimodal biometrics recognition, which represents the test data by a sparse linear combination of training data, while constraining the observations from different modalities of the test subject to share their sparse representations. Thus, we simultaneously take into account correlations as well as coupling information between biometric modalities. Furthermore, the model is modified to make it robust to noise and occlusion. The resulting optimization problem is solved using an efficient alternative direction method. Experiments on a challenging public dataset show that our method compares favorably with competing fusion-based methods.

1 Introduction

Unimodal biometric systems rely on the evidence of a single source of information such as a single iris or fingerprint or face for authentication. Unfortunately these systems often have to deal with some of the following inevitable problems [1]: (a) Noisy data (b) Non-universality: the biometric system based on a single source of evidence may not be able to capture meaningful data from some users. (c) Intra-class variations: in the case of iris recognition a user who wears artificial contact lenses with various patterns can cause these variations. (d) Spoof attack: hand signature forgery is an example of this type of attack. It has been observed that some of the limitations of unimodal biometric systems can be addressed by deploying multimodal biometric systems that essentially integrate the evidence presented by multiple sources of information such as iris, fingerprints and face.

Classification in multibiometric systems is done by fusing information from different modalities. The information fusion can be done at different levels, which can be broadly divided into feature level, score level and rank/decision level fusion. Due to preservation of raw information, feature level fusion can be more discriminative than score or decision level fusion [2]. But, there have been very little effort in exploring feature level fusion in the biometric community. This is because of the different output formats of different sensors, which result in features with different dimensions. Often the features have large dimensions, and fusion becomes difficult at feature level. The prevelant method is feature concatenation, which has been used for different multibiometric settings [3,4]. However, in many scenarios, each modality produces high-dimensional features. In such cases, the method is both impractical and non-robust. It also cannot exploit the constraint that features of different modalities should share the identities.

In recent years, theories of Sparse Representation (SR) and Compressed Sensing (CS) have emerged as powerful tools for efficiently processing data. This has led to a resurgence in interest in the principles of SR and CS for biometrics recognition. See [5,6] and the references therein for a survey of biometrics recognition algorithms using SR and CS. Motivated by the success of SR in unimodal biometric recognition, we propose a joint sparsity-based algorithm for multimodal biometrics recognition. Our method is based on the well known regularized regression method, multi-task multivariate Lasso [7,8]. Figure. 1 presents an overview of our method.

This paper makes the following contributions:

- We present a robust feature level fusion algorithm for multibiometric recognition tasks. Through the proposed joint sparse framework, we can easily handle different dimensions of different modalities by forcing the different features to interact through their sparse coefficients. Further, the proposed algorithm can efficiently handle large dimensional feature vectors.
- We make the classification robust to occlusion and noise by introducing an error term into the optimization framework.
- The algorithm is easily generalizable to handle multiple test inputs from a modality.

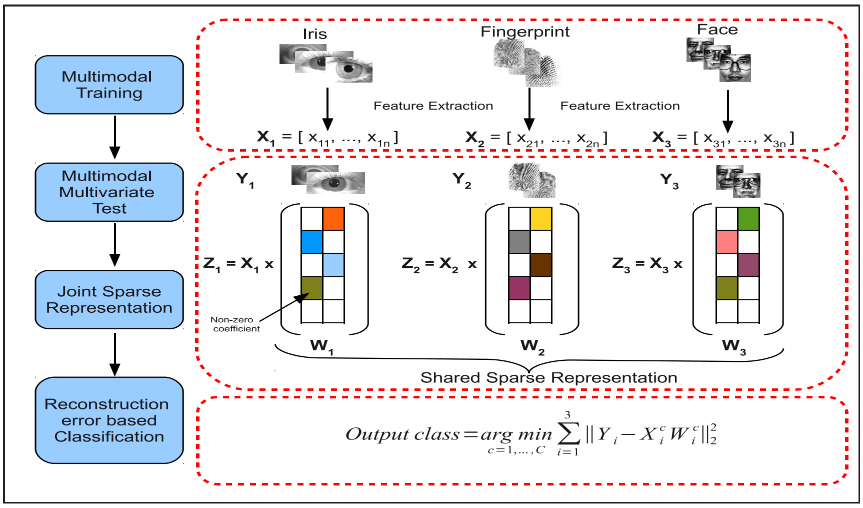

Fig. 1: Overview of our algorithm

2 Joint sparsity-based multimodal biometrics recognition

Consider a multimodal C-class classification problem with D different biometric traits. Suppose there are p_i training samples in each biometric trait. For each biometric trait $i = 1, \ldots, D$, we denote

$$\mathbf{X}^i = [\mathbf{X}^i_1, \mathbf{X}^i_2, \ldots, \mathbf{X}^i_C]$$

as an $n \times p_i$ dictionary of training samples consisting of C sub-dictionaries $\mathbf{X}^i_k$'s corresponding to C different classes. Each sub-dictionary

$$\mathbf{X}^i_j = [\mathbf{x}^i_{j,1}, \mathbf{x}^i_{j,2}, \ldots, \mathbf{x}^i_{j,p_j}] \in \mathbb{R}^{n \times p_j}$$

represents a set of training data from the ith modality labeled with the jth class. Note that n is the feature dimension of each sample and there are p_j number of training samples in class j. Hence, there are a total of $p = \sum_{j=1}^{C} p_j$ many samples in the dictionary $\mathbf{X}^i_C$. In multimodal biometrics recognition problem given a test samples (matrix) $\mathbf{Y}$, which consists of D different modalities $\{\mathbf{Y}^1, \mathbf{Y}^2, \ldots, \mathbf{Y}^D\}$ where each sample $\mathbf{Y}^i$ consists of d_i observations $\mathbf{Y}^i = [\mathbf{y}^i_1, \mathbf{y}^i_2, \ldots, \mathbf{y}^i_d] \in \mathbb{R}^{n \times d_i}$, the objective is to identify the class to which a test sample $\mathbf{Y}$ belongs to. In what follows, we present a multimodal multivariate sparse representation-based algorithm for this problem [7–9].

2.1 Multimodal multivariate sparse representation

We want to exploit the joint sparsity of coefficients from different biometrics modalities to make a joint decision. To simplify this model, let us consider a bimodal classification problem where the test sample $\mathbf{Y} = [\mathbf{Y}^1, \mathbf{Y}^2]$ consists of two different modalities such as iris and face. Suppose that $\mathbf{Y}^1$ belongs to the jth class. Then, it can be reconstructed by a linear combination of the atoms in the sub-dictionary $\mathbf{X}^1_j$. That is, $\mathbf{Y}^1 = \mathbf{X}^1 \boldsymbol{\Gamma}^1 + \mathbf{N}^1$, where $\boldsymbol{\Gamma}^1$ is a sparse matrix with only p_j nonzero rows associated with the jth class and $\mathbf{N}^1$ is the noise matrix. Similarly, since $\mathbf{Y}^2$ represents the same subject, it belongs to the same class and can be represented by training samples in $\mathbf{X}^2_j$ with different set of coefficients $\boldsymbol{\Gamma}^2_j$. Thus, we can write $\mathbf{Y}^2 = \mathbf{X}^2 \boldsymbol{\Gamma}^2 + \mathbf{N}^2$, where $\boldsymbol{\Gamma}^2$ is a sparse matrix that has the same sparsity pattern as $\boldsymbol{\Gamma}^1$. If we let $\boldsymbol{\Gamma} = [\boldsymbol{\Gamma}^1, \boldsymbol{\Gamma}^2]$, then $\boldsymbol{\Gamma}$ is a sparse matrix with only p_j nonzeros rows.

In the more general case where we have D modalities, if we denote $\{\mathbf{Y}^i\}_{i=1}^D$ as a set of D observations each consisting of d_i samples from each modality and let $\boldsymbol{\Gamma} = [\boldsymbol{\Gamma}^1, \boldsymbol{\Gamma}^2, \ldots, \boldsymbol{\Gamma}^D] \in \mathbb{R}^{p \times d}$ be the matrix formed by concatenating the coefficient matrices with $d = \sum_{i=1}^{D} d_i$, then we can seek the row-sparse matrix $\boldsymbol{\Gamma}$ by solving the following ℓ_1/ℓ_q-regularized least square problem

$$\hat{\boldsymbol{\Gamma}} = \arg\min_{\boldsymbol{\Gamma}} \frac{1}{2} \sum_{i=1}^{D} \|\mathbf{Y}^i - \mathbf{X}^i \boldsymbol{\Gamma}^i\|_F^2 + \lambda \|\boldsymbol{\Gamma}\|_{1,q} \tag{1}$$

where λ is a positive parameter and q is set greater than 1 to make the optimization problem convex. Here, $\|\boldsymbol{\Gamma}\|_{1,q}$ is a norm defined as $\|\boldsymbol{\Gamma}\|_{1,q} = \sum_{k=1}^{p} \|\boldsymbol{\gamma}^k\|_q$ where $\boldsymbol{\gamma}^k$'s are the row vectors of $\boldsymbol{\Gamma}$ and $\|\mathbf{Y}\|_F$ is the Frobenius norm of matrix $\mathbf{Y}$ defined as $\|\mathbf{Y}\|_F = \sqrt{\sum_{i,j} Y_{i,j}^2}$. Once $\hat{\boldsymbol{\Gamma}}$ is obtained, the class label associated with an observed vector is then declared as the one that produces the smallest approximation error

$$\hat{j} = \arg\min_j \sum_{i=1}^{D} \|\mathbf{Y}^i - \mathbf{X}^i \boldsymbol{\delta}_j^i(\boldsymbol{\Gamma}^i)\|_F^2, \tag{2}$$

where $\boldsymbol{\delta}_j^i$ is the matrix indicator function defined by keeping rows corresponding to the jth class and setting all other rows equal to zero. Note that the optimization problem (1) reduces to the conventional Lasso [10] when $D = 1$ and $d = 1$. For $D = 1$ (1), it is equivalent to multivariate Lasso [7].

2.2 Robust multimodal multivariate sparse representation

In this section, we consider a more general problem where the data is contaminated by noise. In this case, the observation model can be modeled as

$$\mathbf{Y}^i = \mathbf{X}^i \boldsymbol{\Gamma}^i + \mathbf{Z}^i + \mathbf{N}^i, \quad i = 1, \dots D, \tag{3}$$

where $\mathbf{N}^i$ is a small dense additive noise and $\mathbf{Z}^i \in \mathbb{R}^{n \times d_i}$ is a matrix of background noise (occlusion) with arbitrarily large magnitude. One can assume that each $\mathbf{Z}^i$ is sparsely represented in some basis $\mathbf{B}^i \in \mathbb{R}^{n \times m^i}$. That is, $\mathbf{Z}^i = \mathbf{B}^i \boldsymbol{\Lambda}^i$ for some sparse matrices $\boldsymbol{\Lambda}^i \in \mathbb{R}^{m_i \times d_i}$. Hence, (3) can be rewritten as

$$\mathbf{Y}^i = \mathbf{X}^i \boldsymbol{\Gamma}^i + \mathbf{B}^i \boldsymbol{\Lambda}^i + \mathbf{N}^i, \quad i = 1, \dots D, \tag{4}$$

With this model, one can simultaneously recover the coefficients $\boldsymbol{\Gamma}^i$ and $\boldsymbol{\Lambda}^i$ by taking advantage of that fact that $\boldsymbol{\Lambda}^i$ are sparse

$$\hat{\boldsymbol{\Gamma}}, \hat{\boldsymbol{\Lambda}} = \arg\min_{\boldsymbol{\Gamma}, \boldsymbol{\Lambda}} \frac{1}{2} \sum_{i=1}^{D} \|\mathbf{Y}^i - \mathbf{X}^i \boldsymbol{\Gamma}^i - \mathbf{B}^i \boldsymbol{\Lambda}^i\|_F^2 + \lambda_1 \|\boldsymbol{\Gamma}\|_{1,q} + \lambda_2 \|\boldsymbol{\Lambda}\|_1, \tag{5}$$

where λ_1 and λ_2 are positive parameters and $\boldsymbol{\Lambda} = [\boldsymbol{\Lambda}^1, \boldsymbol{\Lambda}^2, \dots, \boldsymbol{\Lambda}^D]$ is the sparse coefficient matrix corresponding to occlusion. The ℓ_1-norm of matrix $\boldsymbol{\Lambda}$ is defined as $\|\boldsymbol{\Lambda}\|_1 = \sum_{i,j} |\Lambda_{i,j}|$. Note that the idea of exploiting the sparsity of occlusion term has been studied by Wright *et al.* [5].

Once $\boldsymbol{\Gamma}, \boldsymbol{\Lambda}$ are computed, the effect of occlusion can be removed by setting $\tilde{\mathbf{Y}}^i = \mathbf{Y}^i - \mathbf{B}^i \boldsymbol{\Lambda}^i$. One can then declare the class label associated to an observed vector as

$$\hat{j} = \arg\min_j \sum_{i=1}^{D} \|\mathbf{Y}^i - \mathbf{X}^i \boldsymbol{\delta}_j^i(\boldsymbol{\Gamma}^i) - \mathbf{B}^i \boldsymbol{\Lambda}^i\|_F^2. \tag{6}$$

2.3 Optimization algorithm

In this section, we present an algorithm to solve (5) based on the classical alternating direction method of multipliers (ADMM) [11], [12]. Note that the optimization problem (1) can be solved by setting λ_2 equal to zero. Let

$$\mathcal{C}(\boldsymbol{\Gamma}, \boldsymbol{\Lambda}) = \frac{1}{2}\sum_{i=1}^{D} \|\mathbf{Y}^i - \mathbf{X}^i\boldsymbol{\Gamma}^i - \mathbf{B}^i\boldsymbol{\Lambda}^i\|_F^2.$$

In ADMM the idea is to decouple $\mathcal{C}(\boldsymbol{\Gamma}, \boldsymbol{\Lambda})$, $\|\boldsymbol{\Gamma}\|_{1,q}$ and $\|\boldsymbol{\Lambda}\|_1$ by introducing auxiliary variables to reformulate the problem into a constrained optimization problem

$$\min_{\boldsymbol{\Gamma},\boldsymbol{\Lambda},\mathbf{U},\mathbf{V}} \mathcal{C}(\boldsymbol{\Gamma}, \boldsymbol{\Lambda}) + \lambda_1\|\mathbf{V}\|_{1,q} + \lambda_2\|\mathbf{U}\|_1 \quad \text{s. t.} \quad \boldsymbol{\Gamma} = \mathbf{V}, \boldsymbol{\Lambda} = \mathbf{U}. \tag{7}$$

Since, (7) is an equally constrained problem, the Augmented Lagrangian method (ALM) [11] can be used to solve the problem. This can be done by minimizing the augmented lagrangian function $f_{\alpha_\Gamma,\alpha_\Lambda}(\boldsymbol{\Gamma}, \boldsymbol{\Lambda}, \mathbf{V}, \mathbf{U}; \mathbf{A}_\Lambda, \mathbf{A}_\Gamma)$ defined as

$$\mathcal{C}(\boldsymbol{\Gamma}, \boldsymbol{\Lambda}) + \lambda_2\|\mathbf{U}\|_1 + \langle \mathbf{A}_\Lambda, \boldsymbol{\Lambda} - \mathbf{U}\rangle + \frac{\alpha_\Lambda}{2}\|\boldsymbol{\Lambda} - \mathbf{U}\|_F^2 + \lambda_1\|\mathbf{V}\|_{1,q} + \langle \mathbf{A}_\Gamma, \boldsymbol{\Gamma} - \mathbf{V}\rangle + \frac{\alpha_\Gamma}{2}\|\boldsymbol{\Gamma} - \mathbf{V}\|_F^2, \tag{8}$$

where $\mathbf{A}_\Lambda$ and $\mathbf{A}_\Gamma$ are the multipliers of the two linear constraints, and $\alpha_\Lambda, \alpha_\Gamma$ are the positive penalty parameters. The ALM algorithm solves $f_{\alpha_\Gamma,\alpha_\Lambda}(\boldsymbol{\Gamma}, \boldsymbol{\Lambda}, \mathbf{V}, \mathbf{U}; \mathbf{A}_\Lambda, \mathbf{A}_\Gamma)$ with respect to $\boldsymbol{\Gamma}, \boldsymbol{\Lambda}, \mathbf{U}$ and $\mathbf{V}$ jointly, keeping $\mathbf{A}_\Gamma$ and $\mathbf{A}_\Lambda$ fixed and then updating $\mathbf{A}_\Gamma$ and $\mathbf{A}_\Lambda$ keeping the remaining variables fixed. Due to the separable structure of the objective function $f_{\alpha_\Gamma,\alpha_\Lambda}$, one can further simplify the problem by minimizing $f_{\alpha_\Gamma,\alpha_\Lambda}$ with respect to variables $\boldsymbol{\Gamma}, \boldsymbol{\Lambda}, \mathbf{U}$ and $\mathbf{V}$, separately. Different steps of the algorithm are given in Algorithm 1. In what follows, we describe each of the suboptimization problems in detail.

Algorithm 1: Alternating Direction Method of Multipliers (ADMM).

Initialize: $\boldsymbol{\Gamma}_0, \mathbf{U}_0, \mathbf{V}_0, \mathbf{A}_{\Lambda,0}, \mathbf{A}_{\Gamma,0}, \alpha_\Gamma, \alpha_\Lambda$

While not converged do

1. $\boldsymbol{\Gamma}_{t+1} = \arg\min_{\boldsymbol{\Gamma}} f_{\alpha_\Gamma,\alpha_\Lambda}(\boldsymbol{\Gamma}, \boldsymbol{\Lambda}_t, \mathbf{U}_t, \mathbf{V}_t; \mathbf{A}_{\Gamma,t}, \mathbf{A}_{\Lambda,t})$
2. $\boldsymbol{\Lambda}_{t+1} = \arg\min_{\boldsymbol{\Lambda}} f_{\alpha_\Gamma,\alpha_\Lambda}(\boldsymbol{\Gamma}_{t+1}, \boldsymbol{\Lambda}, \mathbf{U}_t, \mathbf{V}_t; \mathbf{A}_{\Gamma,t}, \mathbf{A}_{\Lambda,t})$
3. $\mathbf{U}_{t+1} = \arg\min_{\mathbf{U}} f_{\alpha_\Gamma,\alpha_\Lambda}(\boldsymbol{\Gamma}_{t+1}, \boldsymbol{\Lambda}_{t+1}, \mathbf{U}, \mathbf{V}_t; \mathbf{A}_{\Gamma,t}, \mathbf{A}_{\Lambda,t})$
4. $\mathbf{V}_{t+1} = \arg\min_{\mathbf{V}} f_{\alpha_\Gamma,\alpha_\Lambda}(\boldsymbol{\Gamma}_{t+1}, \boldsymbol{\Lambda}_{t+1}, \mathbf{U}_{t+1}, \mathbf{V}; \mathbf{A}_{\Gamma,t}, \mathbf{A}_{\Lambda,t})$
5. $\mathbf{A}_{\Gamma,t+1} \doteq \mathbf{A}_{\Gamma,t} + \alpha_\Lambda(\boldsymbol{\Gamma}_{t+1} - \mathbf{U}_{t+1})$
6. $\mathbf{A}_{\Lambda,t+1} \doteq \mathbf{A}_{\Lambda,t} + \alpha_\Gamma(\boldsymbol{\Gamma}_{t+1} - \mathbf{V}_{t+1})$

Update step for $\boldsymbol{\Gamma}$: The first suboptimization problem involves the minimization of $f_{\alpha_\Gamma,\alpha_\Lambda}(\boldsymbol{\Gamma},\boldsymbol{\Lambda},\mathbf{V},\mathbf{U};\mathbf{A}_\Lambda,\mathbf{A}_\Gamma)$ with respect to $\boldsymbol{\Gamma}$. It has the quadratic structure, which is easy to solve by setting the first-order derivative equal to zero, and has the following solution

$$\boldsymbol{\Gamma}^i_{t+1} = (\mathbf{X}^{i^T}\mathbf{X}^i + \alpha_\Gamma \mathbf{I})^{-1}(\mathbf{X}^{i^T}(\mathbf{Y}^i - \boldsymbol{\Lambda}^i_t) + \alpha_\Gamma \mathbf{V}^i_t + \mathbf{A}^i_{V,t}),$$

where $\mathbf{I}$ is $p \times p$ identity matrix and $\boldsymbol{\Lambda}^i_t, \mathbf{V}^i_t$ and $\mathbf{A}^i_{V,t}$ are submatrices of $\boldsymbol{\Lambda}_t, \mathbf{V}_t$ and $\mathbf{A}_{V,t}$, respectively.

Update step for $\boldsymbol{\Lambda}$: The second suboptimization problem is similar in nature, whose solution is given below

$$\boldsymbol{\Lambda}^i_{t+1} = (1+\alpha_\Lambda)^{-1}(\mathbf{Y}^i - \mathbf{X}^i\boldsymbol{\Gamma}^i_{t+1} + \alpha_\Lambda \mathbf{U}^i_t - \mathbf{A}^i_{\Lambda,t}),$$

where $\mathbf{U}^i_t$ and $\mathbf{A}^i_{\Lambda,t}$ are submatrices of $\mathbf{U}_t$ and $\mathbf{A}_{\Lambda,t}$, respectively.

Update step for $\mathbf{U}$: The third suboptimization problem is with respect to $\mathbf{U}$, which is the standard ℓ_1 minimization problem which can be recast as

$$\min_{\mathbf{U}} \frac{1}{2}\|\boldsymbol{\Lambda}_{t+1} + \alpha_\Lambda^{-1}\mathbf{A}_{\Lambda,t} - \mathbf{U}\|_F^2 + \frac{\lambda_2}{\alpha_\Lambda}\|\mathbf{U}\|_1. \tag{9}$$

Equation (9) is the well-known shrinkage problem whose solution is given by

$$\mathbf{U}_{t+1} = \mathcal{S}\left(\boldsymbol{\Lambda}_{t+1} + \alpha_\Lambda^{-1}\mathbf{A}_{\Lambda,t}, \frac{\lambda_2}{\alpha_\Lambda}\right),$$

where $\mathcal{S}(a,b) = sgn(a)(|a| - b)$ for $|a| \geq b$ and zero otherwise.

Update step for $\mathbf{V}$: The final suboptimization problem is with respect to $\mathbf{V}$ and can be formulated as

$$\min_{\mathbf{V}} \frac{1}{2}\|\boldsymbol{\Gamma}_{t+1} + \alpha_\Gamma^{-1}\mathbf{A}_{\Gamma,t} - \mathbf{V}\|_F^2 + \frac{\lambda_1}{\alpha_\Gamma}\|\mathbf{V}\|_{1,q}. \tag{10}$$

Due to the separable structure of (10), it can be solved by minimizing with respect to each row of $\mathbf{V}$ separately. Let $\boldsymbol{\gamma}_{i,t+1}, \mathbf{a}_{\Gamma,i,t}$ and $\mathbf{v}_{i,t+1}$ be rows of matrices $\boldsymbol{\Gamma}_{t+1}, \mathbf{A}_{\Gamma,t}$ and $\mathbf{V}_{t+1}$, respectively. Then for each $i = 1, \ldots, p$ we solve the following sub-problem

$$\mathbf{v}_{i,t+1} = \arg\min_{\mathbf{v}} \frac{1}{2}\|\mathbf{z} - \mathbf{v}\|_2^2 + \eta\|\mathbf{v}\|_q, \tag{11}$$

where $\mathbf{z} = \boldsymbol{\gamma}_{i,t+1} - \mathbf{a}_{\Gamma,i,t}\alpha_\Gamma^{-1}$ and $\eta = \frac{\lambda_1}{\lambda_2}$. One can derive the solution for (11) for any q. In this paper, we only focus on the case when $q = 2$. The solution of (11) has the following form

$$\mathbf{v}_{i,t+1} = \left(1 - \frac{\eta}{\|\mathbf{z}\|_2}\right)_+ \mathbf{z},$$

where $(\mathbf{v})_+$ is a vector with entries receiving values $\max(v_i, 0)$.

Our algorithm for multimodal biometrics recognition is summarized in Algorithm 2.

Algorithm 2: Sparse Multimodal Biometrics Recognition (SMBR).

Input: Training samples $\{\mathbf{X_i}\}_{i=1}^D$, test sample $\{\mathbf{Y_i}\}_{i=1}^D$, Occlusion basis $\{\mathbf{B}\}_{i=1}^D$
Procedure: Obtain $\hat{\boldsymbol{\Gamma}}$ and $\hat{\boldsymbol{\Lambda}}$ by solving

$$\hat{\boldsymbol{\Gamma}}, \hat{\boldsymbol{\Lambda}} = \arg\min_{\boldsymbol{\Gamma},\boldsymbol{\Lambda}} \frac{1}{2}\sum_{i=1}^{D} \|\mathbf{Y}^i - \mathbf{X}^i\boldsymbol{\Gamma}^i - \mathbf{B}^i\boldsymbol{\Lambda}^i\|_F^2 + \lambda_1\|\boldsymbol{\Gamma}\|_{1,q} + \lambda_2\|\boldsymbol{\Lambda}\|_1,$$

Output: $\texttt{identity}(\mathbf{Y}) = \arg\min_j \sum_{i=1}^D \|\mathbf{Y}^i - \mathbf{X}^i\boldsymbol{\delta}_j^i(\hat{\boldsymbol{\Gamma}}^i) - \mathbf{B}^i\hat{\boldsymbol{\Lambda}}^i\|_F^2$.

3 Experiments

We evaluated our algorithm for different multi-biometric settings. We tested on a publicly available dataset - the WVU Multimodal dataset [13]. The WVU dataset is one of the few publicly available datasets which allows fusion at image level, hence the proposed feature level fusion technique can be tested. In all the experiments, $\mathbf{B}_i$ was set to be identity for convenience, *i.e.*, we consider error to be sparse in image domain.

3.1 WVU Multimodal Dataset

WVU multimodal dataset is a comprehensive collection of different biometric modalities such as fingerprint, iris, palmprint, hand geometry and voice from subjects of different age, gender and ethnicity. It is a challenging dataset and many of these samples are of poor quality corrupted with blur, occlusion and sensor noise as shown in Figure 2. Out of these, we chose iris and fingerprint modalities for testing the algorithm, giving a total of 6 modalities (4 fingerprint + 2 iris). The evaluation was done on a common subset of 219 subjects having samples in all the modalities.

Fig. 2: Examples of challenging images from the WVU Multimodal dataset, suffering from various artifacts as sensor noise, blur, occlusion and poor acquisition.

Preprocessing Robust pre-processing of images was done before feature extraction. Iris images were segmented following the recent method proposed in [14]. Following the segmentation, 25×240 iris templates were formed by re-sampling using the publicly available code of Masek *et al.* [15]. Fingerprint images were enhanced using the filtering based methods described in [16], and then the core point was detected using the enhanced images [17]. Features were then extracted around the detected core point.

Feature Extraction Gabor features were extracted on the processed images as they have been shown to give good performance on both fingerprints [17] and iris [15]. For fingerprints, the processed images were convolved with Gabor filters at 8 different orientations. Circular tesselations were extracted around the core point for all filtered images. The mean values for each sector were concatenated to form 3600×1 feature vectors. For iris, the templates were convovled with a log-Gabor filter, and vectorized to give 6000×1 dimensional feature.

Experimental Set-up The dataset was randomly divided into 4 training samples per class (1 sample here is 1 data sample each from 6 modalities) and the rest 519 for testing. The recognition result was averaged over 5 runs. The proposed methods were compared with compared with state-of-the-art classification methods such as sparse logistic regression (SLR) [18] and SVM [19]. Although these methods give superior performance on individual modalities, they cannot handle multimodal data. One possible way to handle multimodal data is to use feature concatenation. But, this resulted in feature vectors of size 26400×1 when all 6 modalities are used, and is not useful. Hence, two techniques were explored for fusion. In the first technique, a score-based fusion was followed where the probability outputs for test sample of each modality, $\{\mathbf{y}_i\}_{i=1}^6$ were added together to give the final score vector. Classification was based upon the final score values. For the second technique, the subject chosen by the maximum number of modalities was taken to be from the correct class.

Observations The recognition performances of SMBR-WE (without error) and SMBR-E (with error) were compared with linear SVM and linear SLR classification methods. In the experiments, λ_1 and λ_2 were set to 0.01. Figures 3 and Table 1 demonstrate recognition performance of different methods on three fusion settings - (1) two irises, (2) four fingerprints and (3) all combined.

	SMBR-WE	SMBR-E	SLR-Sum	SLR-Major	SVM-Sum	SVM-Major
4 Fingerprints	**97.9**	97.6	96.3	74.2	90.0	73.0
2 Irises	76.5	**78.2**	72.7	64.2	62.8	49.3
Overall	**98.7**	98.6	97.6	84.4	94.9	81.3

Table 1: Rank one recognition performance for WVU Multimodal dataset.

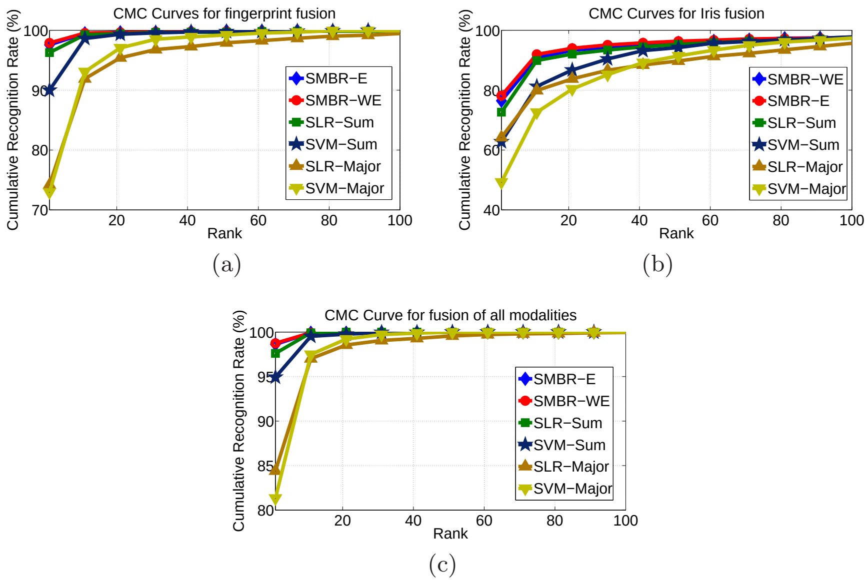

Fig. 3: CMC (Cumulative Match Curve) for multimodal fusion using (a) four fingerprints, (b) two irises and (c) all modalities.

Comparsion of Methods: Clearly, the proposed SMBR approach outperforms existing techniques for all the fusion settings. Both SMBR-E and SMBR-WE have similar performance, though the latter seems to give a slightly better performance. This may be due to the penalty on the sparse error, though the error may not be sparse in image domain. Further, sum-based fusion shows a superior performance over voting-based methods, and SLR performs better than SVM on all the modalities. However, by jointly classifiying all the modalities, SMBR achieves the best performance.

4 Conclusion

We have proposed a novel multimodel multivariate joint sparsity-based algorithm for multimodal biometrics recognition. The algorithm is robust as it explicitly accounts for both noise and occlusion. An efficient algorithm based on alternative direction was proposed for solving the optimization problem. Various experiments show that our method is robust and significantly improves the overall recognition accuracy.

Acknowledgment

The work of SS, VMP and RC was partially supported by a MURI grant from the Army Research Office under the Grant W911NF-09-1-0383.

References

1. Ross, A., Jain, A.K.: Multimodal biometrics: an overview. In: Proc. European Signal Processing Conference. (2004) 1221–1224
2. Klausner, A., Tengg, A., Rinner, B.: Vehicle classification on multi-sensor smart cameras using feature- and decision-fusion. In: IEEE Conf. Dist. Smart Cameras. (2007) 67 –74
3. Rattani, A., Kisku, D., Bicego, M., Tistarelli, M.: Feature level fusion of face and fingerprint biometrics. In: IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems. (2007) 1–6
4. Ross, A.A., Govindarajan, R.: Feature level fusion of hand and face biometrics. In: Proc. of the SPIE. Volume 5779. (2005) 196–204
5. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence **31** (2009) 210–227
6. Patel, V.M., Chellappa, R.: Sparse representations, compressive sensing and dictionaries for pattern recognition. In: Asian Conference on Pattern Recognition. (2010)
7. Yuan, M., Lin, Y.: Model selection and estimation in regression with grouped variables. Journal of the Royal Statistical Society: Series B **68** (2006) 49–67
8. Meier, L., Geer, S.V.D., Bhlmann, P.: The group lasso for logistic regression. Journal of the Royal Statistical Society: Series B **70** (2008) 53–71
9. Nguyen, N.H., Nasrabadi, N.M., Tran, T.D.: Robust multi-sensor classification via joint sparse representation. In: International Conference on Information Fusion. (2011)
10. Tibshirani, R.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B **58** (1996) 267–288
11. Yang, J., Zhang, Y.: Alternating direction algorithms for l1 problems in compressive sensing. SIAM Journal on Scientific Computing **33** (2011) 250278
12. Afonso, M., Bioucas-Dias, J., Figueiredo, M.: An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems. IEEE Transactions on Image Processing **20** (2011) 681–695
13. S. Crihalmeanu, A. Ross, S.S., Hornak, L.: A protocol for multibiometric data acquisition, storage and dissemination. In Technical Report, WVU, Lane Department of Computer Science and Electrical Engineering (2007)
14. Pundlik, S., Woodard, D., Birchfield, S.: Non-ideal iris segmentation using graph cuts. In: IEEE CVPR Workshop. (2008) 1–6
15. Masek, L.: Recognition of human iris patterns for biometric identification. Technical report, The University of Western Australia (2003)
16. S. Chikkerur, C.W., Govindaraju, V.: A systematic approach for feature extraction in fingerprint images. In: Int. Conference on Bioinformatics and its Applications. (2004) 344
17. Jain, A., Prabhakar, S., Hong, L., Pankanti, S.: Filterbank-based fingerprint matching. IEEE Transactions on Image Processing **9** (2000) 846 –859
18. Krishnapuram, B., Carin, L., Figueiredo, M., Hartemink, A.: Sparse multinomial logistic regression: fast algorithms and generalization bounds. IEEE Transactions on Pattern Analysis and Machine Intelligence **27** (2005) 957 –968
19. Burges, C.J.: A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery **2** (1998) 121–167