Joint Sparse Representation for Robust Multimodal Biometrics Recognition

Sumit Shekhar, *Student Member*, *IEEE*, Vishal M. Patel, *Member*, *IEEE*, Nasser M. Nasrabadi, *Fellow*, *IEEE*, and Rama Chellappa, *Fellow*, *IEEE*

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 multimodal sparse representation method, 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 among biometric modalities. A multimodal quality measure is also proposed to weigh each modality as it gets fused. Furthermore, we also kernelize the algorithm to handle nonlinearity in data. The optimization problem is solved using an efficient alternative direction method. Various experiments show that the proposed method compares favorably with competing fusion-based methods.

Index Terms—Multimodal biometrics, feature fusion, sparse representation

1 Introduction

Unimodal biometric systems rely on a single source of information such as a single iris or fingerprint or face for authentication [1]. Unfortunately, these systems have to deal with some of the following inevitable problems [2]:

1. *Noisy data*. Poor lighting on a user's face or occlusion are examples of noisy data.
2. *Nonuniversality*. The biometric system based on a single source of evidence may not be able to capture meaningful data from some users. For instance, an iris biometric system may extract incorrect texture patterns from the iris of certain users due to the presence of contact lenses.
3. *Intraclass variations*. In the case of fingerprint recognition, the presence of wrinkles due to wetness [3] can cause these variations. These types of variations often occur when a user incorrectly interacts with the sensor.
4. *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. Such systems are less vulnerable to spoof attacks, as it would be difficult for an imposter to simultaneously spoof multiple biometric traits of a genuine user. Due to sufficient population coverage, these systems are able to address the problem of nonuniversality.

Classification in multibiometric systems is done by fusing information from different biometric modalities. Information fusion can be done at different levels, 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 [4]. But, feature-level fusion methods have been explored in the biometric community only recently. This is because of the differences in features extracted from different sensors in terms of types and dimensions. Often features have large dimensions, and fusion becomes difficult at the feature level. The prevalent method is feature concatenation, which has been used for different multibiometric settings [5], [6], [7]. However, for high-dimensional feature vectors, simple feature concatenation may be inefficient and nonrobust. A related work in the machine learning literature is multiple kernel learning (MKL), which aims to integrate information from different features by learning a weighted combination of respective kernels. A detailed survey of MKL-based methods can be found in [8]. However, for multimodal systems, weight determination during testing is important, based on the quality of modalities. Also, a corrupted test sample from a modality must be rejected by the algorithm. Such a framework is not yet feasible in the MKL settings. Methods like those given in [9], [10] try to exploit information from data from a different view to improve classifier performance. However, [9] being an unsupervised technique, is not suited for classification tasks, and [10] reduces to the MKL framework in a supervised setting. Similarly, SVM-2k [11] jointly learns SVM for two views, while maximizing the

- *S. Shekhar, V.M. Patel, and R. Chellappa are with the Department of Electrical and Computer Engineering and the Center for Automation Research, UMIACS, University of Maryland, College Park, MD 20742. E-mail: {sshekha, pvishalm, rama}@umiacs.umd.edu.*
- *N.M. Nasrabadi is with the U.S. Army Research Lab, Adelphi, MD 20783. E-mail: nasser.m.nasrabadi@us.army.mil.*

Manuscript received 27 Aug. 2012; revised 11 Mar. 2013; accepted 28 May 2013; published online 10 June 2013.
Recommended for acceptance by S. Sarkar.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-2012-08-0667.
Digital Object Identifier no. 10.1109/TPAMI.2013.109.

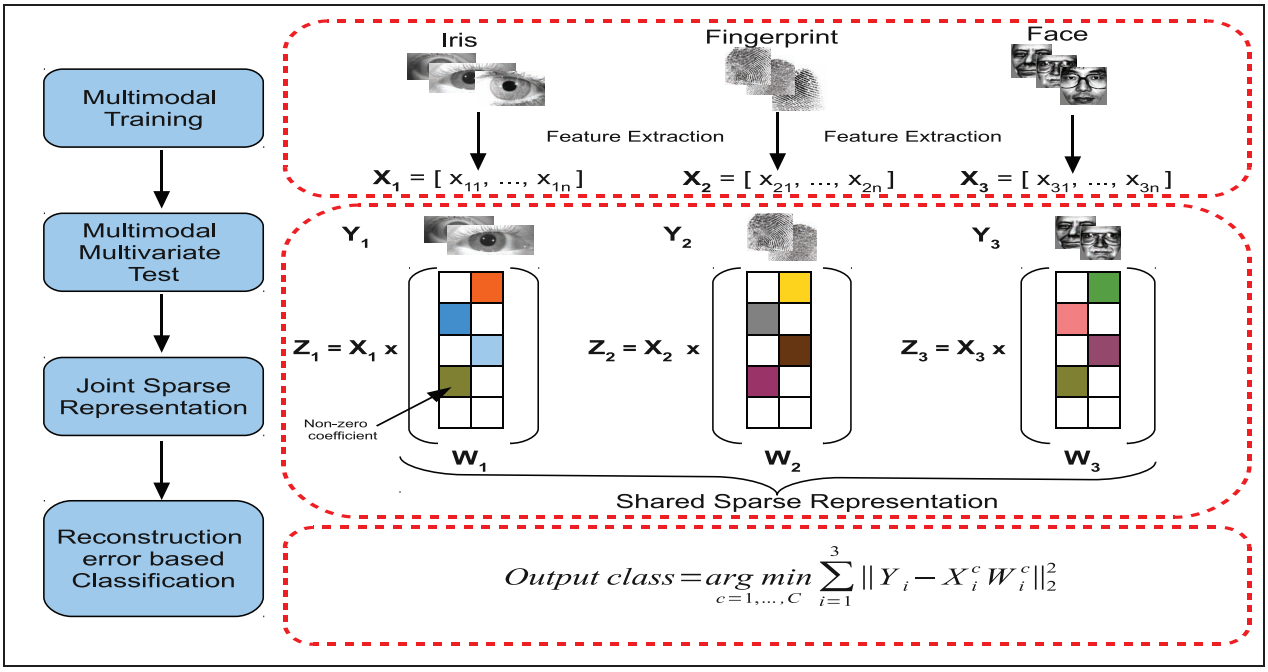


Fig. 1. Overview of our algorithm. The proposed algorithm represents the test data by a sparse linear combination of training data while constraining the observations from different modalities of the test subject to share their sparse representations. Finally, classification is done by assigning the test data to the class with the lowest reconstruction error.

agreement between the projections of data from the two views . It is, however, not clear how this can be extended to multiple views, which is common in multimodal biometrics. A Fisher-discriminant-analysis-based method has also been proposed for integrating multiple views in [12], but it is also similar to MKL with kernel Fisher discriminant analysis as the base learner [13].

In recent years, theories of sparse representation (SR) and compressed sensing (CS) have emerged as powerful tools for efficient processing of data in nontraditional ways [14]. This has led to a resurgence in interest in the principles of SR and CS for biometrics recognition [15]. Wright et al. [16] proposed the seminal sparse representation-based classification (SRC) algorithm for face recognition. It was shown that by exploiting the inherent sparsity of data, one can obtain improved recognition performance over traditional methods especially when data are contaminated by various artifacts such as illumination variations, disguise, occlusion, and random pixel corruption. Pillai et al. [17] extended this work for robust cancelable iris recognition. Nagesh and Li [18] presented an expression-invariant face recognition method using distributed CS and joint sparsity models. Patel et al. [19] proposed a dictionary-based method for face recognition under varying pose and illumination. A discriminative dictionary learning method for face recognition was also proposed by Zhang and Li [20]. For a survey of applications of SR and CS algorithms to biometric recognition, see [14], [15], [21], [22], [23] and the references therein.

Motivated by the success of SR in unimodal biometric recognition, we propose a joint sparsity-based algorithm for multimodal biometrics recognition. Fig. 1 presents an overview of our framework. It is based on the well-known regularized regression method, multitask multivariate Lasso [24], [25]. The proposed method imposes common sparsities both within each biometric modality and across different modalities. The idea of joint sparsity has been explored recently for image classification [26], [27] and segmentation [28]. However, our method is different from these previously proposed algorithms based on joint sparse representation for classification. For example, Yuan and Yan [27] proposed a multitask sparse linear regression model for image classification. This method uses group sparsity to combine different features of an object for classification. Zhang et al. [26] proposed a joint dynamic sparse representation model for object recognition. Their essential goal was to recognize the same object viewed from multiple observations, i.e., different poses. Our method is more general in that it can deal with both multimodal as well as multivariate sparse representations.

This paper makes the following contributions:

- We present a robust feature level fusion algorithm for multibiometric recognition. Through the proposed joint sparse framework, we can easily handle unequal dimensions from different modalities by forcing the different features to interact through their sparse coefficients. Furthermore, the proposed algorithm can efficiently handle large-dimensional feature vectors.
- We make the classification robust to occlusion and noise by introducing an error term in the optimization framework.
- The algorithm is easily generalizable to handle multiple test inputs from a modality.
- We introduce a quality measure for multimodal fusion based on the joint sparse representation.
- Last, we kernelize the algorithm to handle nonlinearity in the data samples.

A preliminary version of this work appeared in [29], which describes just the linear version of the algorithm, robust to noise and occlusion. Furthermore, extensive experimental evaluations are presented here.

1.1 Paper Organization

The paper is organized as follows: In Section 2, we describe the proposed sparsity-based multimodal recognition algorithm, which is kernelized in Section 4. The quality measure is described in Section 3. Experimental evaluations on a comprehensive multimodal data set and a face database are described in Section 5. Finally, in Section 6, we discuss the computational complexity of the method. Concluding remarks are presented in Section 7.

2 Joint Sparsity-Based Multimodal Biometrics Recognition

Consider a multimodal C-class classification problem with D different biometric traits. Suppose there are $p = \sum_{j=1}^{C} p_j$ training samples in each biometric trait, where p_j is the number of training samples in class j. For each biometric trait $i = 1, \ldots, D$, we denote

$$\mathbf{X}^i = \left[\mathbf{X}_1^i, \mathbf{X}_2^i, \ldots, \mathbf{X}_C^i\right]$$

as an $n_i \times p$ dictionary of training samples consisting of C subdictionaries $\mathbf{X}_k^i$ corresponding to C different classes. Each subdictionary

$$\mathbf{X}_j^i = \left[\mathbf{x}_{j,1}^i, \mathbf{x}_{j,2}^i, \ldots, \mathbf{x}_{j,p_j}^i\right] \in \mathbb{R}^{n_i \times p_j}$$

represents a set of training data from the ith modality labeled with the jth class. Note that n_i is the feature dimension of each sample. Elements of the dictionary are often referred to as atoms. In multimodal biometrics recognition problem, given test samples $\mathbf{Y}$, 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 $\mathbf{Y}^i = [\mathbf{y}_1^i, \mathbf{y}_2^i, \ldots, \mathbf{y}_{d_i}^i] \in \mathbb{R}^{n_i \times d_i}$, the objective is to identify the class to which a test sample $\mathbf{Y}$ belongs to. Note that we do not constrain the number of samples per modality to be the same, as assumed in forming the training matrix. In what follows, we present a multimodal multivariate sparse representation-based algorithm for this problem [24], [25], [30].

2.1 Multimodal Multivariate Sparse Representation

We propose to exploit the joint sparsity of coefficients from different biometric 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 subdictionary $\mathbf{X}_j^1$. That is, $\mathbf{Y}^1 = \mathbf{X}^1\mathbf{\Gamma}^1 + \mathbf{N}^1$, where $\mathbf{\Gamma}^1$ 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 $\mathbf{X}_j^2$ with different sets of coefficients $\mathbf{\Gamma}_j^2$. Thus, we can write $\mathbf{Y}^2 = \mathbf{X}^2\mathbf{\Gamma}^2 + \mathbf{N}^2$, where $\mathbf{\Gamma}^2$ is a sparse matrix that has the same sparsity pattern as $\mathbf{\Gamma}^1$. If we let $\mathbf{\Gamma} = [\mathbf{\Gamma}^1, \mathbf{\Gamma}^2]$, then $\mathbf{\Gamma}$ is a sparse matrix with only p_j nonzero rows, as both $\mathbf{Y}^1$ and $\mathbf{Y}^2$ are represented by samples of the jth class.

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 $\mathbf{\Gamma} = [\mathbf{\Gamma}^1, \mathbf{\Gamma}^2, \ldots, \mathbf{\Gamma}^D] \in \mathbb{R}^{p \times d}$ be the matrix formed by concatenating the coefficient matrices with $d = \sum_{i=1}^{D} d_i$, then we can determine the row-sparse matrix $\mathbf{\Gamma}$ by solving the following ℓ_1/ℓ_q-regularized least-squares problem:

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

where λ is a positive parameter and q is set greater than 1 to make the optimization problem convex. Here, $\|\mathbf{\Gamma}\|_{1,q}$ is a norm defined as $\|\mathbf{\Gamma}\|_{1,q} = \sum_{k=1}^{p} \|\boldsymbol{\gamma}^k\|_q$, where $\boldsymbol{\gamma}^k$s are the row vectors of $\mathbf{\Gamma}$ and $\|\mathbf{Y}\|_F$ is the Frobenius norm of the matrix $\mathbf{Y}$ defined as $\|\mathbf{Y}\|_F = \sqrt{\sum_{i,j} Y_{i,j}^2}$. The ℓ_1/ℓ_q regularization seeks a solution with sparse nonzero rows; hence, we get a representation consistent across all the modalities. Once $\hat{\mathbf{\Gamma}}$ is obtained, the class label associated with an observed vector is then declared as the one that produces the smallest approximation error:

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

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 [31] when $D = 1$ and $d = 1$. In the case when $D = 1$, (1) is referred to as multivariate Lasso [24].

2.2 Robust Multimodal Multivariate Sparse Representation

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

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

where $\mathbf{N}^i$ is a small dense additive noise and $\mathbf{Z}^i \in \mathbb{R}^{n_i \times d_i}$ 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 $\mathbf{B}^i \in \mathbb{R}^{n_i \times m_i}$. That is, $\mathbf{Z}^i = \mathbf{B}^i\mathbf{\Lambda}^i$ for some sparse matrices $\mathbf{\Lambda}^i \in \mathbb{R}^{m_i \times d_i}$. For simplicity, we assume $\mathbf{B}^i$ to be orthonormal in this paper. Hence, (3) can be rewritten as

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

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

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

where λ_1 and λ_2 are positive parameters and $\mathbf{\Lambda} = [\mathbf{\Lambda}^1, \mathbf{\Lambda}^2, \ldots, \mathbf{\Lambda}^D]$ is the sparse coefficient matrix corresponding to

occlusion. The ℓ_1-norm of matrix $\mathbf{\Lambda}$ is defined as $\|\mathbf{\Lambda}\|_1 = \sum_{i,j}|\Lambda_{i,j}|$. Note that the idea of exploiting the sparsity of occlusion term has been studied by Wright et al. [16] and Candes et al. [32].

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

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

2.3 Optimization Algorithm

The optimization problem (5) is convex but difficult to solve due to the joint sparsity constraint. In this section, we present an approach based on the classical alternating direction method of multipliers (ADMM) [33], [34] to solve (5). Note that the optimization problem (1) can be solved by setting λ_2 equal to infinity. Let

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

Then, our goal is to solve the following optimization problem:

$$\min_{\mathbf{\Gamma},\mathbf{\Lambda}} \mathcal{C}(\mathbf{\Gamma}, \mathbf{\Lambda}) + \lambda_1\|\mathbf{\Gamma}\|_{1,q} + \lambda_2\|\mathbf{\Lambda}\|_1. \tag{7}$$

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

$$\begin{aligned} &\min_{\mathbf{\Gamma},\mathbf{\Lambda},\mathbf{U},\mathbf{V}} \mathcal{C}(\mathbf{\Gamma}, \mathbf{\Lambda}) + \lambda_1\|\mathbf{V}\|_{1,q} + \lambda_2\|\mathbf{U}\|_1 \quad \text{s.t.} \\ &\mathbf{\Gamma} = \mathbf{V}, \mathbf{\Lambda} = \mathbf{U}. \end{aligned} \tag{8}$$

Since (8) is an equally constrained problem, the augmented Lagrangian method (ALM) [33] can be used to solve the problem. This can be done by minimizing the augmented Lagrangian function $f_{\alpha_\Gamma,\alpha_\Lambda}(\mathbf{\Gamma}, \mathbf{\Lambda}, \mathbf{V}, \mathbf{U}; \mathbf{A}_\Lambda, \mathbf{A}_\Gamma)$ defined as

$$\begin{aligned} &\mathcal{C}(\mathbf{\Gamma}, \mathbf{\Lambda}) + \lambda_2\|\mathbf{U}\|_1 + \langle\mathbf{A}_\Lambda, \mathbf{\Lambda} - \mathbf{U}\rangle + \frac{\alpha_\Lambda}{2}\|\mathbf{\Lambda} - \mathbf{U}\|_F^2 \\ &+ \lambda_1\|\mathbf{V}\|_{1,q} + \langle\mathbf{A}_\Gamma, \mathbf{\Gamma} - \mathbf{V}\rangle + \frac{\alpha_\Gamma}{2}\|\mathbf{\Gamma} - \mathbf{V}\|_F^2, \end{aligned} \tag{9}$$

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 $f_{\alpha_\Gamma,\alpha_\Lambda}(\mathbf{\Gamma}, \mathbf{\Lambda}, \mathbf{V}, \mathbf{U}; \mathbf{A}_\Lambda, \mathbf{A}_\Gamma)$ with respect to $\mathbf{\Gamma}$, $\mathbf{\Lambda}$, $\mathbf{U}$, 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 $\mathbf{\Gamma}$, $\mathbf{\Lambda}$, $\mathbf{U}$, 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: $\mathbf{\Gamma}_0, \mathbf{U}_0, \mathbf{V}_0, \mathbf{A}_{\Lambda,0}, \mathbf{A}_{\Gamma,0}, \alpha_\Gamma, \alpha_\Lambda$

While not converged do

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

2.3.1 Update Step for $\mathbf{\Gamma}$

The first suboptimization problem involves the minimization of $f_{\alpha_\Gamma,\alpha_\Lambda}(\mathbf{\Gamma}, \mathbf{\Lambda}, \mathbf{V}, \mathbf{U}; \mathbf{A}_\Lambda, \mathbf{A}_\Gamma)$ with respect to $\mathbf{\Gamma}$. It has the quadratic structure, which is easy to solve by setting the first-order derivative equal to zero. Furthermore, the loss function $\mathcal{C}(\mathbf{\Gamma}, \mathbf{\Lambda})$ is a sum of convex functions associated with sub-matrices $\mathbf{\Gamma}^i$, one can seek for $\mathbf{\Gamma}_{t+1}^i$, $i = 1, \ldots, D$, which has the following solution:

$$\mathbf{\Gamma}_{t+1}^i = \left(\mathbf{X}^{i^T}\mathbf{X}^i + \alpha_\Gamma\mathbf{I}\right)^{-1}\left(\mathbf{X}^{i^T}\left(\mathbf{Y}^i - \mathbf{B}^i\mathbf{\Gamma}_t^i\right) + \alpha_\Gamma\mathbf{V}_t^i - \mathbf{A}_{\Gamma,t}^i\right),$$

where $\mathbf{I}$ is the $p \times p$ identity matrix and $\mathbf{\Lambda}_t^i$, $\mathbf{\Gamma}_t^i$, and $\mathbf{A}_{\Gamma,t}^i$ are submatrices of $\mathbf{\Lambda}_t$, $\mathbf{\Gamma}_t$, and $\mathbf{A}_{\Gamma,t}$, respectively.

2.3.2 Update Step for $\mathbf{\Lambda}$

The second suboptimization problem is similar in nature; its solution is given below:

$$\mathbf{\Lambda}_{t+1}^i = \left(\mathbf{B}^{i^T}\mathbf{B}^i + \alpha_\Lambda\mathbf{I}\right)^{-1}\left(\mathbf{B}^{i^T}\left(\mathbf{Y}^i - \mathbf{X}^i\mathbf{\Gamma}_{t+1}^i\right) + \alpha_\Lambda\mathbf{U}_t^i - \mathbf{A}_{\Lambda,t}^i\right),$$

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

2.3.3 Update Step for $\mathbf{U}$

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

$$\min_{\mathbf{U}} \frac{1}{2}\left\|\mathbf{\Lambda}_{t+1} + \alpha_\Lambda^{-1}\mathbf{A}_{\Lambda,t} - \mathbf{U}\right\|_F^2 + \frac{\lambda_2}{\alpha_\Lambda}\|\mathbf{U}\|_1. \tag{10}$$

Equation (10) is the well-known shrinkage problem whose solution is given by

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

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

2.3.4 Update Step for $\mathbf{V}$

The final suboptimization problem is with respect to $\mathbf{V}$, which can be reformulated as

$$\min_{\mathbf{V}} \frac{1}{2}\left\|\mathbf{\Gamma}_{t+1} + \alpha_\Gamma^{-1}\mathbf{A}_{\Gamma,t} - \mathbf{V}\right\|_F^2 + \frac{\lambda_1}{\alpha_\Gamma}\|\mathbf{V}\|_{1,q}. \tag{11}$$

Due to the separable structure of (11), 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 $\mathbf{\Gamma}_{t+1}$, $\mathbf{A}_{\Gamma,t}$ and $\mathbf{V}_{t+1}$, respectively. Then for each $i = 1, \ldots, p$, we solve the following subproblem:

$$\mathbf{v}_{i,t+1} = \arg\min_{\mathbf{v}} \frac{1}{2}\|\mathbf{z} - \mathbf{v}\|_2^2 + \eta\|\mathbf{v}\|_q, \tag{12}$$

where $\mathbf{z} = \boldsymbol{\gamma}_{i,t+1} + \mathbf{a}_{\Gamma,i,t}\alpha_\Gamma^{-1}$ and $\eta = \frac{\lambda_1}{\alpha_\Gamma}$. One can derive the solution for (12) for any q. In this paper, we only focus on the case when $q = 2$. The solution of (12) has the following form:

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

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

Our proposed sparse multimodal biometrics recognition (SMBR) method is summarized in Algorithm 2. We refer to the robust method that takes sparse error into account as SMBR-E (SMBR with error), and the initial case where it is not taken into account as SMBR-WE (SMBR without error).

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 $\hat{\mathbf{\Gamma}}$ and $\hat{\mathbf{\Lambda}}$ by solving

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

Output:

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

3 Quality-Based Fusion

Ideally, a fusion mechanism should give more weights to the more reliable modalities. Hence, the concept of quality is important in multimodal fusion. A quality measure based on sparse representation was introduced for faces in [16]. To decide whether a given test sample has good quality or not, its sparsity concentration index (SCI) was calculated. Given a coefficient vector $\boldsymbol{\gamma} \in \mathbb{R}^p$, the SCI is given as

$$SCI(\boldsymbol{\gamma}) = \frac{\frac{C.\max_{j\in\{1,\ldots,C\}}\|\delta_j(\boldsymbol{\gamma})\|_1}{\|\boldsymbol{\gamma}\|_1} - 1}{C-1},$$

where δ_j is the indicator function keeping the coefficients corresponding to the jth class and setting others to zero. SCI values close to 1 correspond to the case where the test sample can be represented well using the samples of a single class, hence are of high quality. On the other hand, samples with SCI close to 0 are not similar to any of the classes, and hence are of poor quality. This can be easily extended to the multimodal case using the joint sparse representation matrix $\hat{\mathbf{\Gamma}}$. In this case, we can define the quality, q_j^i, for sample $\mathbf{y}_j^i$ as

$$q_j^i = SCI(\hat{\mathbf{\Gamma}}_j^i),$$

where $\hat{\mathbf{\Gamma}}_j^i$ is the jth column of $\hat{\mathbf{\Gamma}}^{\mathbf{i}}$. Given this quality measure, the classification rule (2) can be modified to include the quality measure:

$$\hat{j} = \arg\min_j \sum_{i=1}^{D}\sum_{k=1}^{d_i} q_k^i \|\mathbf{y}_k^i - \mathbf{X}^i\boldsymbol{\delta}_j(\mathbf{\Gamma}_k^i)\|_F^2, \tag{13}$$

where $\boldsymbol{\delta}_j$ is the indicator function retaining the coefficients corresponding to jth class.

4 Kernel Space Multimodal Biometrics Recognition

The class identities in the multibiometric data set may not be linearly separable. Hence, we also extend the sparse multimodal fusion framework to kernel space. The kernel function, $\kappa : \mathbb{R}^n \times \mathbb{R}^n$, is defined as the inner product

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \langle\phi(\mathbf{x}_i), \phi(\mathbf{x}_j)\rangle,$$

where ϕ is an implicit mapping projecting the vector $\mathbf{x}$ into a higher dimensional space.

4.1 Multivariate Kernel Sparse Representation

Considering the general case of D modalities with $\{\mathbf{Y}^i\}_{i=1}^{D}$ as a set of d_i observations, the feature space representation can be written as

$$\mathbf{\Phi}(\mathbf{Y}^i) = \left[\phi(\mathbf{y}_1^i), \phi(\mathbf{y}_2^i), \ldots, \phi(\mathbf{y}_d^i)\right].$$

Similarly, the dictionary of training samples for modality $i = 1, \ldots, D$ can be represented in feature space as

$$\mathbf{\Phi}(\mathbf{X}^i) = \left[\phi(\mathbf{X}_1^i), \phi(\mathbf{X}_2^i), \ldots, \phi(\mathbf{X}_C^i)\right].$$

As in joint linear space representation, we have

$$\mathbf{\Phi}(\mathbf{Y}^i) = \mathbf{\Phi}(\mathbf{X}^i)\mathbf{\Gamma}^i,$$

where $\mathbf{\Gamma}^i$ is the coefficient matrix associated with modality i. Incorporating information from all the sensors, we seek to solve the following optimization problem similar to the linear case:

$$\hat{\mathbf{\Gamma}} = \arg\min_{\mathbf{\Gamma}} \frac{1}{2}\sum_{i=1}^{D} \|\mathbf{\Phi}(\mathbf{Y}^i) - \mathbf{\Phi}(\mathbf{X}^i)\mathbf{\Gamma}^i\|_F^2 + \lambda\|\mathbf{\Gamma}\|_{1,q}, \tag{14}$$

where $\mathbf{\Gamma} = [\mathbf{\Gamma}^1, \mathbf{\Gamma}^2, \ldots, \mathbf{\Gamma}^D]$. It is clear that the information from all modalities is integrated via the shared sparsity pattern of the matrices $\{\mathbf{\Gamma}^i\}_{i=1}^{D}$. This can be reformulated in terms of kernel matrices as

$$\hat{\mathbf{\Gamma}} = \arg\min_{\mathbf{\Gamma}} \frac{1}{2}\sum_{i=1}^{D} \left(\text{trace}\left(\mathbf{\Gamma}^{i^T}\mathbf{K}_{\mathbf{X}_i,\mathbf{X}_i}\mathbf{\Gamma}^i\right) - 2\text{trace}\left(\mathbf{K}_{\mathbf{X}_i,\mathbf{Y}_i}\mathbf{\Gamma}^i\right)\right) + \lambda\|\mathbf{\Gamma}\|_{1,q}, \tag{15}$$

where the kernel matrix $\mathbf{K}_{\mathbf{A},\mathbf{B}}$ is defined as

$$\mathbf{K}_{\mathbf{A},\mathbf{B}}(i,j) = \langle\phi(\mathbf{a}_i), \phi(\mathbf{b}_j)\rangle, \tag{16}$$

$\mathbf{a}_i$ and $\mathbf{b}_j$ being ith and jth columns of $\mathbf{A}$ and $\mathbf{B}$, respectively.

4.2 Optimization Algorithm

Similarly to the linear fusion method, we apply the alternating direction method to efficiently solve the problem for kernel fusion. This is done by introducing a new variable $\mathbf{V}$ and reformulating the problem (15) as

$$\arg\min_{\mathbf{\Gamma},\mathbf{V}} \frac{1}{2}\sum_{i=1}^{D} \left(\text{trace}\left(\mathbf{\Gamma}^{i^T}\mathbf{K}_{\mathbf{X}^i,\mathbf{X}^i}\mathbf{\Gamma}^i\right) - 2\text{trace}\left(\mathbf{K}_{\mathbf{X}^i,\mathbf{Y}^i}\mathbf{\Gamma}^i\right)\right) + \lambda\|\mathbf{V}\|_{1,q} \text{ s.t. } \mathbf{\Gamma} = \mathbf{V}. \tag{17}$$

Rewriting the problem using the Lagrangian multiplier $\mathbf{P}_\Gamma$, the optimization problem becomes

$$\arg\min_{\mathbf{\Gamma},\mathbf{V}} \frac{1}{2}\sum_{i=1}^{D} \left(\text{trace}\left(\mathbf{\Gamma}^{i^T}\mathbf{K}_{\mathbf{X}^i,\mathbf{X}^i}\mathbf{\Gamma}^i\right) - 2\text{trace}\left(\mathbf{K}_{\mathbf{X}^i,\mathbf{Y}^i}\mathbf{\Gamma}^i\right)\right) + \lambda\|\mathbf{V}\|_{1,q} + \langle\mathbf{P}_\Gamma, \mathbf{\Gamma} - \mathbf{V}\rangle + \frac{\beta_\Gamma}{2}\left\|\mathbf{\Gamma} - \mathbf{V}\right\|_F^2, \tag{18}$$

where β_Γ is a positive penalty parameter. This upon rearranging reduces to

TABLE 1
WVU Biometric Data

Biometric Modality	# of subjects	# of samples
Iris	244	3099
Fingerprint	272	7219
Palm	263	683
Hand Geometry	217	3062
Voice	274	714

$$\arg\min_{\mathbf{\Gamma},\mathbf{V}} \frac{1}{2}\sum_{i=1}^{D}\left(\text{trace}\left(\mathbf{\Gamma}^{i^T}\mathbf{K}_{\mathbf{X}^i,\mathbf{X}^i}\mathbf{\Gamma}^i\right) - 2\text{trace}\left(\mathbf{K}_{\mathbf{X}^i,\mathbf{Y}^i}\mathbf{\Gamma}^i\right)\right) + \lambda\|\mathbf{V}\|_{1,q} + \frac{\beta_\Gamma}{2}\left\|\mathbf{\Gamma} - \mathbf{V} + \frac{1}{\beta_\Gamma}\mathbf{P}_\Gamma\right\|_F^2. \tag{19}$$

Now, (19) can be solved in a similar way as the linear fusion problem in (5). The optimization method is summarized in Algorithm 3. It should be pointed out that each step has a simple closed-form expression.

Algorithm 3. Alternating Direction Method of Multipliers (ADMM) in kernel space.
Initialize: $\mathbf{\Gamma}_0, \mathbf{V}_0, \mathbf{B}_0, \beta_\Gamma$
While not converged do
1. $\mathbf{\Gamma}_{t+1} = \arg\min_{\mathbf{\Gamma}} \frac{1}{2}\sum_{i=1}^{D}(\text{trace}(\mathbf{\Gamma}^{i^T}\mathbf{K}_{\mathbf{X}^i,\mathbf{X}^i}\mathbf{\Gamma}^i) - 2\text{trace}(\mathbf{K}_{\mathbf{X}^i,\mathbf{Y}^i}\mathbf{\Gamma}^i)) + \lambda\|\mathbf{V}_t\|_{1,q} + \frac{\beta_\Gamma}{2}\left\|\mathbf{\Gamma} - \mathbf{V_t} + \frac{1}{\beta_\Gamma}\mathbf{P}_{\Gamma,t}\right\|_F^2$
2. $\mathbf{V}_{t+1} = \arg\min_{\mathbf{V}} \lambda\|\mathbf{V}\|_{1,q} + \frac{\beta_\Gamma}{2}\left\|\mathbf{\Gamma}_{t+1} - \mathbf{V} + \frac{1}{\beta_\Gamma}\mathbf{P}_{\Gamma,t}\right\|_F^2$
3. $\mathbf{P}_{\Gamma,t+1} = \mathbf{P}_{\Gamma,t} + \beta_\Gamma(\mathbf{\Gamma}_{t+1} - \mathbf{V}_{t+1})$

4.2.1 Update Steps for $\mathbf{\Gamma}_t$

$\mathbf{\Gamma}_{t+1}$ is obtained by updating each submatrix $\mathbf{\Gamma}_t^i$, $i = 1, \ldots, D$, as

$$\mathbf{\Gamma}_t^i = (\mathbf{K}_{\mathbf{X}^i,\mathbf{X}^i} + \beta_\Gamma\mathbf{I})^{-1}\left(\mathbf{K}_{\mathbf{X}^i,\mathbf{Y}^i} + \beta_\Gamma\mathbf{V}_t^i - \mathbf{P}_{\Gamma,t}^i\right), \tag{20}$$

where $\mathbf{I}$ is an identity matrix, and $\mathbf{V}_t^i$ and $\mathbf{P}_{\Gamma,t}^i$ are submatrices of $\mathbf{V}_t$ and $\mathbf{P}_{\Gamma,t}$, respectively.

4.2.2 Update Steps for $\mathbf{V}_t$

The update equation for $\mathbf{V}_t$ is the same as in the linear fusion case using (11) and (12), replacing $\mathbf{A}_{\Gamma,t}$ and α_Γ with $\mathbf{P}_{\Gamma,t}$ and β_Γ, respectively.

4.3 Classification

Once $\mathbf{\Gamma}$ is obtained, classification can be done by assigning the class label as

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

or in terms of kernel matrices as

$$\hat{j} = \arg\min_j \sum_{i=1}^{D}\left(\text{trace}(\mathbf{K}_{\mathbf{Y},\mathbf{Y}}) - 2\text{trace}\left(\hat{\mathbf{\Gamma}}_j^{i^T}\mathbf{K}_{X_j^i,Y}\hat{\mathbf{\Gamma}}_j^i\right) + \text{trace}\left(\hat{\mathbf{\Gamma}}_j^{i^T}\mathbf{K}_{X_j^i,X_j^i}\hat{\mathbf{\Gamma}}_j^i\right)\right). \tag{21}$$

Here, $\mathbf{X}_j^i$ is the subdictionary associated with the jth class and $\hat{\mathbf{\Gamma}}_j^i$ is the coefficient matrix associated with this class.

The classification rule can be further extended to include the quality measure as in (13). But, we skip this step here as we wish to study the effect of kernel representation and quality separately.

Fig. 2. Examples of challenging images from the WVU multimodal data set. The images shown above suffer from various artifacts such as sensor noise, blur, and occlusion.

Multivariate kernel sparse recognition (kerSMBR) algorithm is summarized in Algorithm 4:

Algorithm 4. Kernel Sparse Multimodal Biometrics Recognition (kerSMBR).
Input: Training samples $\{\mathbf{X}_i\}_{i=1}^D$, test sample $\{\mathbf{Y}_i\}_{i=1}^D$
Procedure: Obtain $\hat{\mathbf{\Gamma}}$ by solving
$\hat{\mathbf{\Gamma}} = \arg\min_{\mathbf{\Gamma}} \frac{1}{2}\sum_{i=1}^{D}\left(\text{trace}(\mathbf{\Gamma}^{i^T}\mathbf{K}_{\mathbf{X}_i,\mathbf{X}_i}\mathbf{\Gamma}^i) - 2\text{trace}(\mathbf{K}_{\mathbf{X}_i,\mathbf{Y}_i}\mathbf{\Gamma}^i)\right) + \lambda\|\mathbf{\Gamma}\|_{1,q}$
Output: $\texttt{identity}(\mathbf{Y}) = \arg\min_j \sum_{i=1}^{D}(\text{trace}(\mathbf{K}_{\mathbf{Y},\mathbf{Y}}) - 2\text{trace}(\hat{\mathbf{\Gamma}}_j^{i^T}\mathbf{K}_{X_j^i,Y}\hat{\mathbf{\Gamma}}_j^i) + \text{trace}(\hat{\mathbf{\Gamma}}_j^{i^T}\mathbf{K}_{X_j^i,X_j^i}\hat{\mathbf{\Gamma}}_j^i))$

5 Experiments

We evaluated our algorithm on two publicly available data sets—the WVU multimodal data set [35] and the AR face data set [36]. In the first experiment, we tested on the WVU data set, which is one of the few publicly available data sets that allows fusion at the image level. It is a challenging data set consisting of samples from different biometric modalities for each subject.

In the second experiment, we show the applicability of the proposed approach to fusing information from *weak* biometrics extracted from face images. In particular, the periocular region has been shown to be a useful biometric [37]. Similarly, the nose region has also been explored as a biometric [38]. Sinha et al. [39] have demonstrated that eyebrows are important for face recognition. However, each of these subregions may not be as discriminative as the whole face. The challenge for fusion algorithms is to be able to combine these weak modalities with a strong modality based on the whole face [40]. We demonstrate how our framework can be extended to address this problem. Furthermore, we also show the effects of noise and occlusion on the performance of different algorithms. In all the experiments, $\mathbf{B}_i$ was set to be identity for convenience, i.e., we assume background noise to be sparse in the image domain.

5.1 WVU Multimodal Data Set

The WVU multimodal data set 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, as described in Table 1. It is a challenging data set, as many of these samples are corrupted with blur, occlusion, and sensor noise, as shown in Fig. 2. Out of these, we chose iris and fingerprint modalities for testing the proposed algorithms. In total, there are two iris (right and left iris) and four fingerprint modalities. Also, the evaluation was done on a subset of 219 subjects having samples in both modalities.

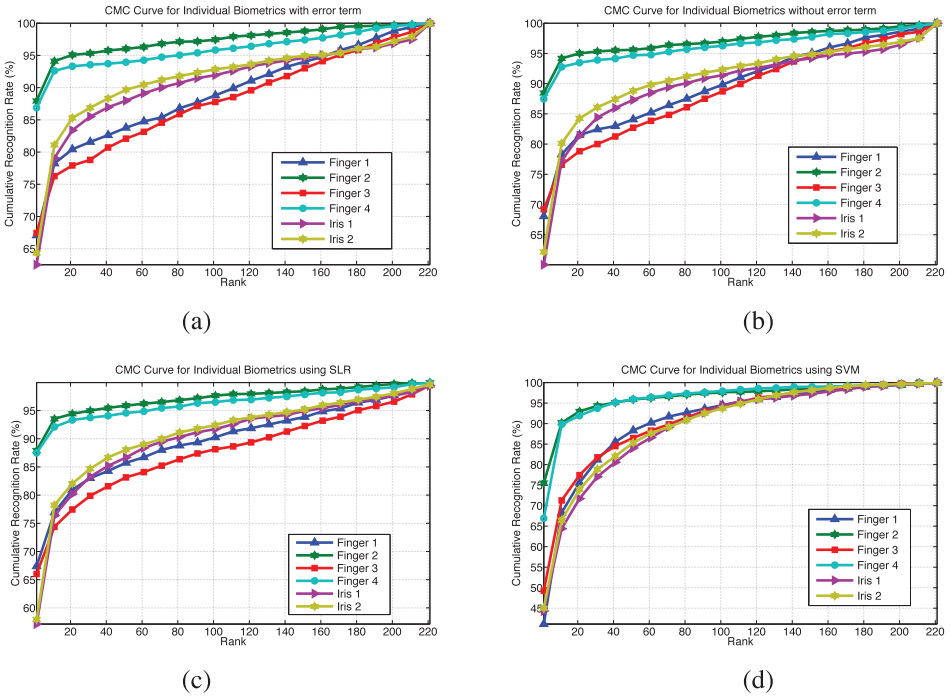


Fig. 3. CMCs for individual modalities using (a) SMBR-E, (b) SMBR-WE, (c) SLR, and (d) SVM methods on WVU data set.

5.1.1 Preprocessing

Robust preprocessing of images was done before feature extraction. Iris images were segmented using the method proposed in [41]. Following the segmentation step, 25×240 iris templates were generated by resampling using the publicly available code of Masek and Kovesi [42]. Fingerprint images were enhanced using the filtering methods described in [43], and then the core point was detected from the enhanced images [44]. Features were then extracted around the detected core point.

5.1.2 Feature Extraction

Gabor features were extracted from the processed images as they have been shown to give good performance on both fingerprints [44] and iris [45]. For fingerprint samples, the processed images were convolved with Gabor filters at eight different orientations. Circular tessellations were extracted around the core point for all the filtered images similar to [44]. The tessellation consisted of 15 concentric bands, each of width 5 pixels and divided into 30 sectors. The mean values for each sector were concatenated to form the feature vector of size $3{,}600 \times 1$. Features for iris images were formed by convolving the templates with a log-Gabor filter at a single scale, and vectorizing the template to give a $6{,}000 \times 1$ dimensional feature.

5.1.3 Experimental Setup

The data set was randomly divided into four training samples per class (one sample here is one data sample each from six modalities) and the remaining 519 samples were used for testing. The recognition result was averaged over five runs. The proposed methods were compared with state-of-the-art classification methods such as sparse logistic regression (SLR) [46] and SVM [47]. As these methods cannot handle multiple modalities, we explored score-level and decision-level fusion methods for combining the results of individual modalities. For score-level fusion, the probability outputs for the 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 decision-level fusion, the subject chosen by the maximum number of modalities was taken to be from the correct class. We further compared with an efficient multiclass implementation of the MKL algorithm [48]. The proposed linear and kernel fusion techniques were tested separately and were compared with linear and kernel versions of SLR, SVM, and MKL algorithms. We denote the score-level fusion of these methods as SLR-Sum and SVM-Sum, and the decision-level fusion as SLR-Major and SVM-Major. The MKL-based method is denoted as MKLFusion. We report the mean and standard deviation of rank-one recognition rates for all the methods. We also show the cumulative match curves (CMCs) for all the classifiers. The CMCs provide the performance measure for biometric recognition systems and has been shown to be equivalent to the ROC of the system [49].

Linear Fusion. The recognition performances of SMBR-WE and SMBR-E were compared with linear SVM and linear SLR classification methods. The parameters λ_1 and λ_2 were set to 0.01:

- *Comparsion of methods*. Fig. 3 and Table 2 show the performance on individual modalities. All the classifiers show a similar trend. The performance for all of them are lower on iris images and fingers 1 and 3. The proposed method show superior performance on all the modalities. Fig. 4 and Table 3 show the recognition performance for different fusion settings. The proposed SMBR approach outperforms existing classification techniques. Furthermore, the CMC curves of the proposed approaches lie above

TABLE 2
Rank-One Recognition Performance on WVU Data Set for Individual Modalities

	Finger 1	Finger 2	Finger 3	Finger 4	Iris 1	Iris 2
SMBR-WE	$\mathbf{68.1 \pm 1.1}$	$\mathbf{88.4 \pm 1.2}$	$\mathbf{69.2 \pm 1.5}$	$\mathbf{87.5 \pm 1.5}$	60.0 ± 1.5	62.1 ± 0.4
SMBR-E	67.1 ± 1.0	87.9 ± 0.8	67.4 ± 1.9	86.9 ± 1.5	$\mathbf{62.5 \pm 1.2}$	$\mathbf{64.3 \pm 1.0}$
SLR	67.4 ± 1.9	87.9 ± 1.3	66.0 ± 2.2	87.5 ± 1.3	57.1 ± 3.0	57.9 ± 2.7
SVM	41.1 ± 5.0	75.5 ± 2.2	49.2 ± 1.6	67.0 ± 8.3	44.3 ± 1.2	45.0 ± 2.9

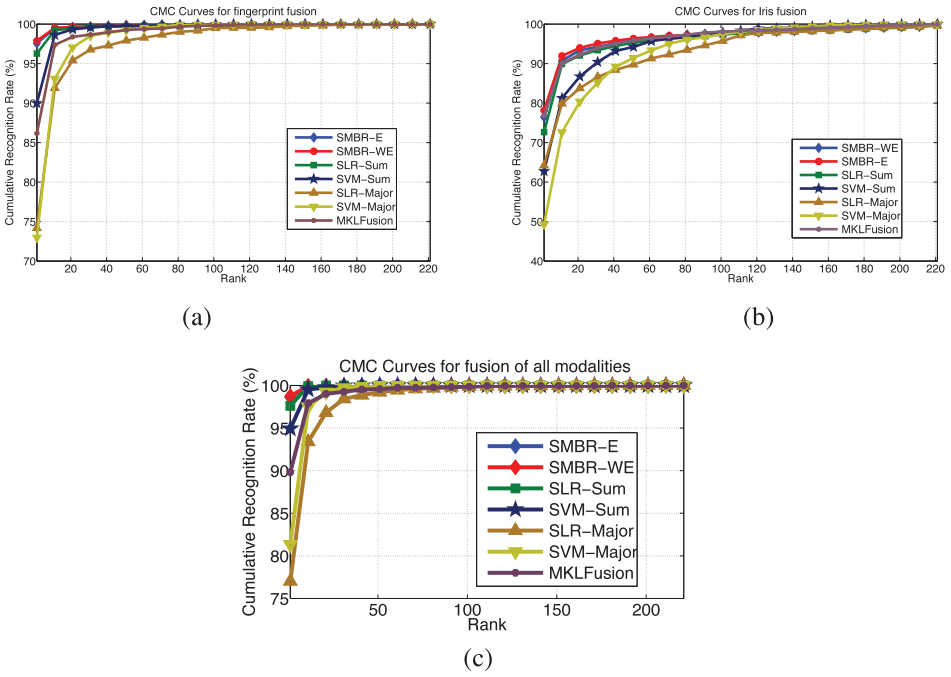


Fig. 4. CMCs (Cumulative Match Curve) for multimodal fusion using (a) four fingerprints, (b) two irises, and (c) all modalities on WVU data set.

TABLE 3
Rank-One Recognition Performance on WVU Data Set for Different Fusion Settings

	SMBR-WE	SMBR-E	SLR-Sum	SLR-Major	SVM-Sum	SVM-Major	MKLFusion
4 Fingerprints	$\mathbf{97.9 \pm 0.4}$	97.6 ± 0.6	96.3 ± 0.8	74.2 ± 0.7	90.0 ± 2.2	73.0 ± 1.5	86.2 ± 1.2
2 Irises	76.5 ± 1.6	$\mathbf{78.2 \pm 1.2}$	72.7 ± 4.0	64.2 ± 2.7	62.8 ± 2.6	49.3 ± 2.0	76.8 ± 2.5
All modalities	$\mathbf{98.7 \pm 0.2}$	98.6 ± 0.5	97.6 ± 0.4	84.4 ± 0.9	94.9 ± 1.5	81.3 ± 1.7	89.8 ± 0.9

the other methods 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 the image domain. Furthermore, sum-based fusion shows a superior performance over voting-based methods. The MKL-based method shows good performance for iris fusion, but the performance drops for other two settings. This may be because by weighing kernels during training, it loses flexibility while testing when number of modalities increase.

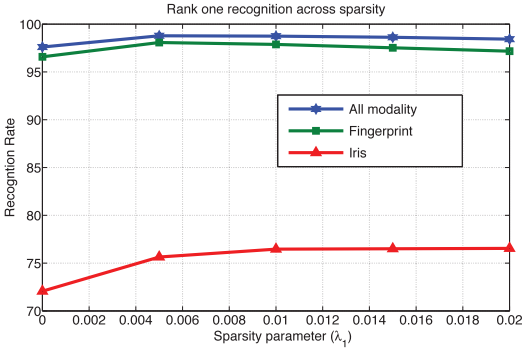


Fig. 5. Variation of recognition performance with different values of sparsity constraint, λ_1.

- *Fusion with quality*. Clearly, different modalities have different levels of performance. Hence, we studied the effect of the proposed quality measure on the performance of different methods. For a consistent comparison, the quality values produced by the SMBR-E method were used for all the algorithms. Table 4 shows the performance for the three fusion settings. The effect of including the quality measure can be studied by comparing with Table 3. Clearly, the recognition rate increases for all the methods across the fusion settings. Again SMBR-E and SMBR-WE give the best performances among all the methods.
- *Effect of joint sparsity*. We also studied the effect of joint sparsity constraint on the recognition performance. For this, the SMBR-WE algorithm was run for different values of λ_1. Fig. 5 shows the rank-one

TABLE 4
Rank-One Recognition Performance on WVU Data Set Using the Proposed Quality Measure

	SMBR-WE	SMBR-E	SLR-Sum	SLR-Major	SVM-Sum	SVM-Major
4 Fingerprints	$\mathbf{98.2 \pm 0.5}$	98.1 ± 0.5	97.5 ± 0.5	86.3 ± 0.6	93.6 ± 1.6	85.5 ± 0.9
2 Irises	76.9 ± 1.2	$\mathbf{78.8 \pm 1.7}$	74.1 ± 1.0	67.2 ± 2.4	64.3 ± 3.3	51.6 ± 2.0
All modalities	$\mathbf{98.8 \pm 0.4}$	98.6 ± 0.3	98.2 ± 0.2	93.8 ± 0.9	95.5 ± 1.5	93.3 ± 1.2

recognition variation across λ_1 values for different fusion settings. All the curves show a sharp increase in performance around $\lambda_1 = 0$. Furthermore, the increase is more for iris fusion, which shows around 5 percent improvement at $\lambda_1 = 0.005$ over $\lambda_1 = 0$. This shows that imposing joint sparsity constraint is important for fusion. Moreover, it helps in regulating fusion performance, when the reconstruction error alone is not sufficient to distinguish between different classes. The performance is then stable across λ_1 values, and starts decreasing slowly after reaching the optimum performance.

- *Variation with number of training samples*. We varied the number of training samples and studied the effect on the proposed method along with SLR-Sum and MKLFusion. Fig. 6 shows the variation for fusion of all the modalities. It can be seen that SMBR-WE and SMBR-E are stable across number of training samples, whereas the performances of SLR-Sum- and MKLFusion-based methods fall sharply. The fall in performance of SLR-Sum and MKLFusion can be attributed to the discriminative approaches of these methods, as well as score-based fusion, as the fusion further reduces the recognition performance when individual classifiers are not good.
- *Comparison with other score-based fusion methods*. Although sum-based fusion is a popular technique for score fusion, some other techniques have also been proposed. We evaluated the performance of the likelihood-based fusion method proposed in [50]. The results are shown in Table 5. The method does not show good performance as it models score distribution as a Gaussian mixture model. However, it is difficult to model score distribution due to large variations in data samples. The method is also affected by the curse of dimensionality.

Kernel Fusion. We further compared the performances of proposed kerSMBR with kernel SVM, kernel SLR, and MKLFusion methods. In the experiments, we used radial basis function (RBF) as the kernel, given as

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{\sigma^2}\right),$$

σ being a parameter to control the width of the RBF. For MKLFusion, we gave linear, polynomial, and RBF kernels as the base kernels for learning:

- *Hyperparameter tuning*. To fix the value of hyperparameter σ, we iterated over different values of σ, $\{2^{-3}, 2^{-2}, \ldots, 2^3\}$, for one set of training and test split of the data. The value of σ giving the maximum performance was fixed for each modality, and the performance was averaged over a few iterations. λ and β_Γ were set to 0.01 and 0.01, respectively.
- *Comparison of methods*. Fig. 7 and Table 6 show the performance of different methods on individual modalities, and Fig. 8 and Table 7 on different fusion settings. Comparison of performance with linear fusion shows that the proposed kerSMBR significantly improves the performance on individual iris modalities as well as iris fusion. The performance on fingerprint modalities is similar; however, the fusion of all six modalities (two iris + four fingerprints) shows an improvement of 0.4 percent. kerSMBR also achieves the best accuracy among all the methods for different fusion settings. kerSLR scores better than kerSVM in all the cases, and it's accuracy is close to kerSMBR. The performance of kerSLR is better than the linear counterpart; however, kerSVM does not show much improvement.

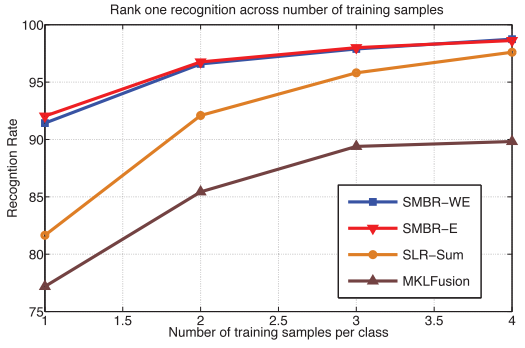


Fig. 6. Variation of recognition performance with number of training samples.

5.2 AR Face Data Set

The AR face data set consists of faces with varying illumination, expression, and occlusion conditions, captured in two sessions. We evaluated our algorithms on a set of 100 users. Images from the first session, seven for each subject, were used as training and the images from the second session, again seven per subject, were used for testing. For testing the fusion algorithms, four weak modalities were extracted from the face images: left and right periocular, mouth, and nose regions. This was done by applying

TABLE 5
Rank One Recognition Performance with the Likelihood-Based Method [50] on WVU Data Set

	2 irises	4 fingerprints	All modalities
SLR-Likelihood	66.6	83.5	75.1
SVM-Likelihood	50.7	31.9	31.0

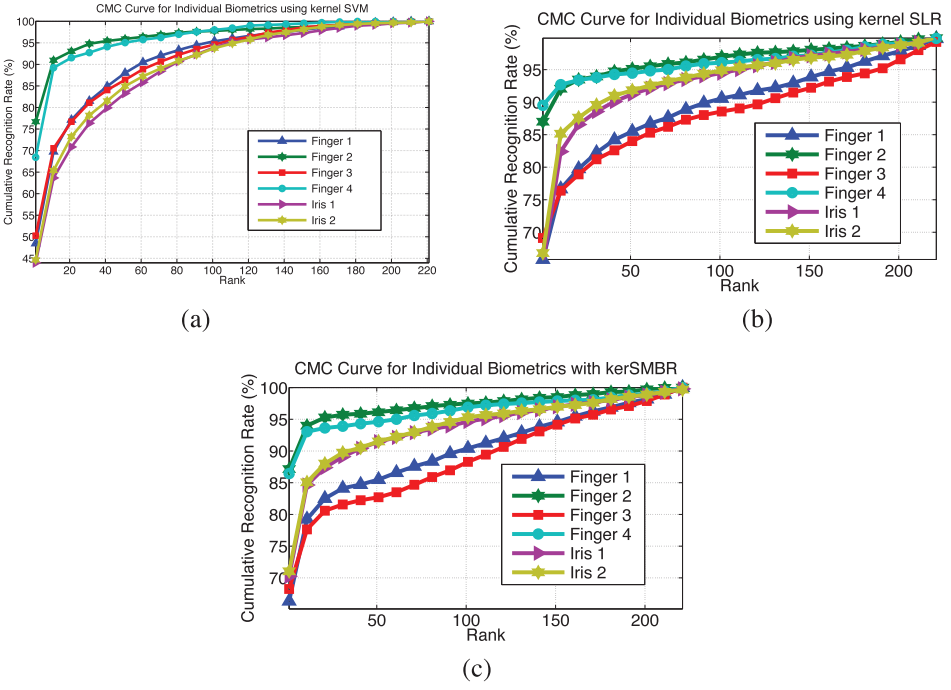


Fig. 7. CMCs for individual modalities using (a) kernel SVM, (b) kernel SLR, and (c) kerSMBR.

TABLE 6
Rank-One Recognition Performance on WVU Data Set for Individual Modalities Using Kernel Methods

	Finger 1	Finger 2	Finger 3	Finger 4	Iris 1	Iris 2
kerSMBR	$\mathbf{66.3 \pm 1.7}$	$\mathbf{87.1 \pm 1.0}$	$\mathbf{69.1 \pm 2.1}$	86.4 ± 1.5	$\mathbf{70.3 \pm 1.8}$	$\mathbf{71.0 \pm 1.6}$
kerSLR	65.8 ± 1.8	86.9 ± 1.7	68.3 ± 2.0	$\mathbf{89.5 \pm 1.6}$	65.1 ± 1.7	66.8 ± 1.1
kerSVM	48.4 ± 5.4	76.7 ± 2.3	50.2 ± 1.9	68.4 ± 7.4	43.9 ± 1.1	44.6 ± 3.0

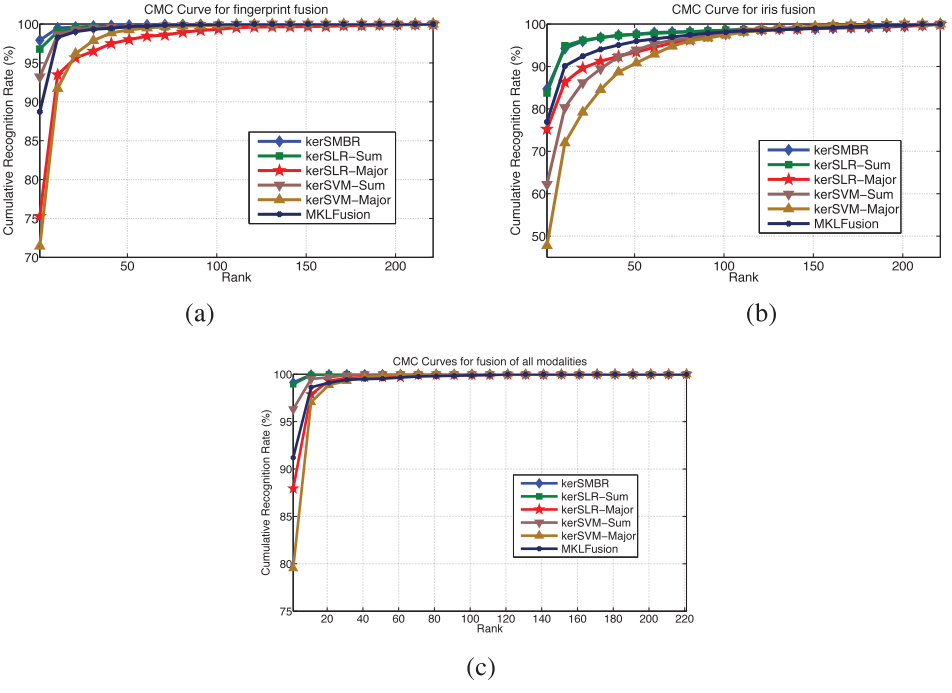


Fig. 8. CMCs for different fusion methods for (a) four fingerprints, (b) two irises, and (c) all modalities.

rectangular masks as shown in Fig. 9, and cropping out the respective regions. These, along with the whole face, were taken for fusion. Simple intensity values were used as features for all of them. The experimental setup was similar to the previous section. The parameter values, λ_1 and λ_2, were set to 0.003 and 0.002, respectively. Furthermore, we also studied the effect of noise and occlusion on recognition performance.

TABLE 7
Rank-One Recognition Performance on WVU Data Set for Different Fusion Settings Using Kernel Methods

	kerSMBR	kerSLR-Sum	kerSLR-Major	kerSVM-Sum	kerSVM-Major	MKLFusion
4 Fingerprints	$\mathbf{97.9 \pm 0.3}$	96.8 ± 0.7	75.2 ± 0.7	93.2 ± 1.2	71.4 ± 1.3	88.7 ± 0.9
2 Irises	$\mathbf{84.7 \pm 1.7}$	83.7 ± 1.8	75.2 ± 1.2	62.2 ± 2.8	47.8 ± 2.4	76.9 ± 2.4
All modalities	$\mathbf{99.1 \pm 0.2}$	98.9 ± 0.1	87.9 ± 0.6	96.3 ± 0.8	79.5 ± 1.6	91.2 ± 1.0

- *Comparison of methods:* Table 8 shows the performance of different algorithms on the face data set. Here, SR shows the classification result using just the whole face. The block sparse method is a recent block sparsity-based face recognition algorithm [23] and FDDL [51] is a state-of-the-art discriminative dictionaries-based technique, but using only a single modality. Clearly, the SMBR approach achieves about 4 percent improvement over other techniques. Thus, robust classification using multiple modalities results in a significant improvement over the current benchmark. Furthermore, a comparison with discriminative methods, such as SLR and SVM, shows that they perform poorly compared to the proposed method. This is because weak modalities are hard to discriminate; hence, score-level fusion with strong modality does not improve performance. On the other hand, by appropriately weighing different modalities, MKLFusion achieves better results. However, by imposing reconstruction and joint sparsity simultaneously, the proposed method is able to achieve the best performance.
- *Effect of noise:* In this experiment, test images were corrupted with white Gaussian noise of increasing variance, σ^2. Comparisons are shown in Fig. 10. It can be seen that both SMBR, SR, and block sparse methods are stable with noise. The performance of other algorithms degrade sharply with the noise level. This also highlights the problem with MKLFusion, as it is not robust to degradation during testing.
- *Effect of occlusion:* In this experiment, a randomly chosen block of the test image was occluded. The recognition performance was studied with increasing block size. Fig. 11 shows the performance of various algorithms with block size. SMBR-E is the most stable among all the methods due to robust handling of error. Recognition rates for other methods fall sharply with increasing block size.
- *Recognition in spite of disguise:* We also performed experiment on the rest of the AR face data set, occluded by sun-glass and scarves. Similar to the above experiment, seven frontal nonoccluded images per subject, from the first session, were used for training, and 12 occluded images per person from both the sessions were used for testing. Again the proposed SMBR-WE and SMBR-E methods outperformed the other methods, as shown in Table 9. SMBR-E method gave the best performance, improving by 17.7 percent over the block sparse method.

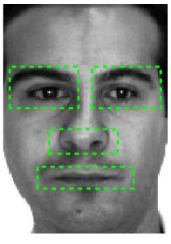


Fig. 9. Face mask used to crop out different modalities.

TABLE 8
Rank-One Performance Comparison of Different Methods on AR Face Data Set

Method	Recognition Rate (%)	Method	Recognition Rate (%)
SMBR-WE	**96.9**	SVM-Sum	86.7
SMBR-E	96	SLR-Sum	77.9
SR	91	FDDL [51]	91.9
Block Sparse [23]	92.2	MKLFusion	89.7

TABLE 9
Rank-One Performance Comparison of Different Methods on Images with Disguise in AR Face Data Set

Method	Scarves	Sun-glass	Overall
SMBR-WE	**86.2**	36.0	61.1
SMBR-E	80.0	**75.0**	**77.5**
SR	45.3	52.3	48.8
Block Sparse [23]	65.8	53.8	59.8
SLR-Sum	72.2	39.6	55.9
SVM-Sum	13.8	42.5	28.1
MKLFusion	47.7	13.0	30.3

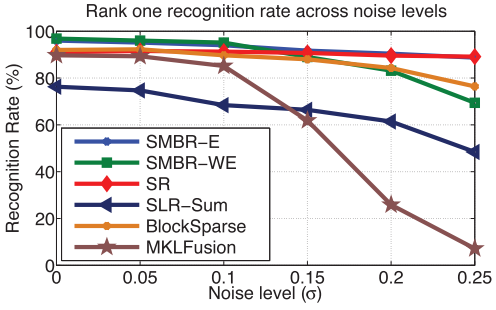


Fig. 10. Effect of noise on rank-one recognition performance for AR face data set.

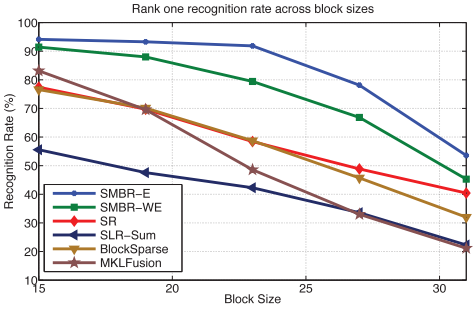


Fig. 11. Effect of occlusion on rank-1 recognition performance for AR face data set.

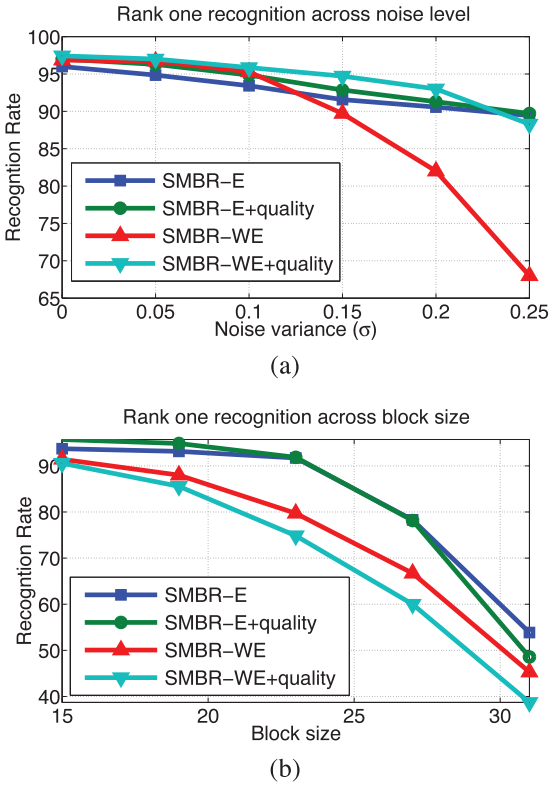


Fig. 12. Effect of quality on recognition performance across (a) noise, (b) random blocks on AR face data set.

- *Quality-based fusion:* Quality determination is an important parameter in fusion here, as strong modality is being combined with weak modalities. We studied the effect of quality measure introduced in Section 3. However, in this case, we fix the quality for strong modality, *viz.* whole face to be 1, while for the weak modalities, the SCI values were taken. The recognition performance for SMBR-E and SMBR-WE across different noise and occlusion levels was studied. Fig. 12 show the performance comparison with the unweighted methods. Using quality, the recognition performance for SMBR-WE goes up to 97.4 percent from 96.9 percent, whereas for SMBR-WE it increases up to 97 percent from 96 percent. Similarly, results improve across different noise levels for both methods. However, SMBR-WE with quality shows worse performance as block size is increased. This may be because it does not handle sparse error; hence, the quality values are not robust.

6 Computational Complexity

The proposed algorithms are computationally efficient. The main steps of the algorithms are the update steps for $\mathbf{\Gamma}$, $\mathbf{\Lambda}$, $\mathbf{U}$, and $\mathbf{V}$. For linear fusion, the update step for $\mathbf{\Gamma}$ involves computing $(\mathbf{X}^{iT}\mathbf{X}^i + \alpha_\Gamma \mathbf{I})^{-1}$, and four matrix multiplications. The first term is constant across iterations and can be precomputed. Matrix multiplication for two matrices of sizes $m \times n$ and $n \times p$ can be done in $\mathcal{O}(mnp)$ time. Hence, for a given training and test data, the computations are linear in feature dimension. Hence, large feature dimensions can be efficiently handled. Similarly, update step for $\mathbf{\Lambda}$ involves matrix multiplication $\mathbf{X}^i\mathbf{\Gamma}^i$. Update steps for $\mathbf{U}$ and $\mathbf{V}$ involve only scalar matrix computations and are very fast. Similarly in the kernel fusion, update for $\mathbf{\Gamma}$ involves calculating $(\mathbf{K}_{\mathbf{X}^i,\mathbf{X}^i} + \beta_\Gamma \mathbf{I})^{-1}$, which can be precomputed. Other steps are similar to linear fusion. Classification step involves calculating the residual error for each class, and is efficient.

7 Conclusion and Discussions

We proposed a novel joint sparsity-based feature level fusion algorithm for multimodal biometrics recognition. The algorithm is robust as it explicitly includes both noise and occlusion terms. An efficient algorithm based on the alternative direction was proposed for solving the optimization problem. We also proposed a multimodal quality measure based on sparse representation. Furthermore, the algorithm was kernelized to handle nonlinear variations. Various experiments have shown that the method is robust and significantly improves the overall recognition accuracy.

An important question is about the theoretical justification of the proposed approaches through ℓ_0/ℓ_1 equivalence. For the special case of $D = 1$, the theory of sparse representation has been reviewed in [52], [53]. However, it has not been addressed yet for the general multimodal setting. This is a challenging problem, and can be investigated as a future direction to this paper.

Acknowledgments

The work of Sumit Shekhar, Vishnal M. Patel, and Rama Chellappa was partially supported by a MURI grant from the US Army Research Office under the Grant W911NF-09-1-0383.

References

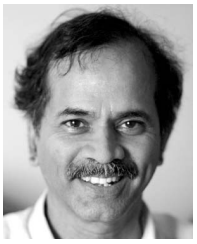
[1] A. Ross, K. Nandakumar, and A.K. Jain, *Handbook of Multibiometrics.* Springer, 2006.
[2] A. Ross and A.K. Jain, "Multimodal Biometrics: An Overview," *Proc. European Signal Processing Conf.*, pp. 1221-1224, Sept. 2004.
[3] P. Krishnasamy, S. Belongie, and D. Kriegman, "Wet Fingerprint Recognition: Challenges and Opportunities," *Proc. Int'l Joint Conf. Biometrics*, pp. 1-7, Oct. 2011.
[4] A. Klausner, A. Tengg, and B. Rinner, "Vehicle Classification on Multi-Sensor Smart Cameras Using Feature- and Decision-Fusion," *Proc. IEEE Conf. Distributed Smart Cameras*, pp. 67-74, Sept. 2007.
[5] A. Rattani, D. Kisku, M. Bicego, and M. Tistarelli, "Feature Level Fusion of Face and Fingerprint Biometrics," *Proc. IEEE Int'l Conf. Biometrics: Theory, Applications, and Systems*, pp. 1-6, Sept. 2007.
[6] X. Zhou and B. Bhanu, "Feature Fusion of Face and Gait for Human Recognition at a Distance in Video," *Proc. Int'l Conf. Pattern Recognition*, vol. 4, pp. 529-532, Aug. 2006.
[7] A.A. Ross and R. Govindarajan, "Feature Level Fusion of Hand and Face Biometrics," *Proc. SPIE*, vol. 5779, pp. 196-204, Mar. 2005.
[8] M. Gönen and E. Alpaydn, "Multiple Kernel Learning Algorithms," *J. Machine Learning Research*, vol. 12, pp. 2211-2268, 2011.
[9] S. Kakade and D. Foster, "Multi-View Regression via Canonical Correlation Analysis," *Pro. 20th Int'l Conf. Learning Theory*, pp. 82-96, 2007.
[10] V. Sindhwani and D. Rosenberg, "An RKHS for Multi-View Learning and Manifold Co-Regularization," *Proc. 25th Int'l Conf. Machine learning*, pp. 976-983, July 2008.
[11] J. Farquhar, H. Meng, S. Szedmak, D. Hardoon, and J. Shawe-taylor, "Two View Learning: SVM-2k, Theory and Practice," *Proc. Advances in Neural Information Processing Systems*, Dec. 2006.

[12] T. Diethe, D. Hardoon, and J. Shawe-Taylor, "Constructing Nonlinear Discriminants from Multiple Data Views," *Machine Learning and Knowledge Discovery in Databases,* pp. 328-343, 2010.
[13] S. Kim, A. Magnani, and S. Boyd, "Optimal Kernel Selection in Kernel Fisher Discriminant Analysis," *Proc. 23rd Int'l Conf. Machine Learning,* pp. 465-472, June 2006.
[14] V.M. Patel and R. Chellappa, "Sparse Representations, Compressive Sensing and Dictionaries for Pattern Recognition," *Proc. Asian Conf. Pattern Recognition,* pp. 325-329, Nov. 2011.
[15] V.M. Patel, R. Chellappa, and M. Tistarelli, "Sparse Representations and Random Projections for Robust and Cancelable Biometrics," *Proc. Int'l Conf. Control, Automation, Robotics, and Vision,* pp. 1-6, Dec. 2010.
[16] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma, "Robust Face Recognition via Sparse Representation," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 31, no. 2, pp. 210-227, Feb. 2009.
[17] J.K. Pillai, V.M. Patel, R. Chellappa, and N.K. Ratha, "Secure and Robust Iris Recognition Using Random Projections and Sparse Representations," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 33, no 9, pp. 1877-1893, Sept. 2011.
[18] P. Nagesh and B. Li, "A Compressive Sensing Approach for Expression-Invariant Face Recognition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 1518-1525, June 2009.
[19] V.M. Patel, T. Wu, S. Biswas, P. Phillips, and R. Chellappa, "Dictionary-Based Face Recognition under Variable Lighting and Pose," *IEEE Trans. Information Forensics and Security,* vol. 7, no. 3, pp. 954-965, June 2012.
[20] Q. Zhang and B. Li, "Discriminative K-SVD for Dictionary Learning in Face Recognition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 2691-2698, June 2010.
[21] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. Huang, and S. Yan, "Sparse Representation for Computer Vision and Pattern Recognition," *Proc. IEEE,* vol. 98, no. 6, pp. 1031-1044, June 2010.
[22] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, "Towards a Practical Face Recognition System: Robust Alignment and Illumination via Sparse Representation," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 34, no. 2, pp. 372-386, Feb. 2012.
[23] E. Elhamifar and R. Vidal, "Robust Classification Using Structured Sparse Representation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 1873-1879, June 2011.
[24] M. Yuan and Y. Lin, "Model Selection and Estimation in Regression with Grouped Variables," *J. Royal Statistical Soc.: Series B,* vol. 68, pp. 49-67, Feb. 2006.
[25] L. Meier, S.V.D. Geer, and P. Bhlmann, "The Group Lasso for Logistic Regression," *J. Royal Statistical Soc.: Series B,* vol. 70, pp. 53-71, Feb. 2008.
[26] H. Zhang, N.M. Nasrabadi, Y. Zhang, and T.S. Huang, "Multi-Observation Visual Recognition via Joint Dynamic Sparse Representation," *Proc. IEEE Int'l Conf. Computer Vision,* pp. 595-602, Nov. 2011.
[27] X.-T. Yuan and S. Yan, "Visual Classification with Multi-Task Joint Sparse Representation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 3493-3500, June 2010.
[28] B. Cheng, G. Liu, J. Wang, Z. Huang, and S. Yan, "Multi-Task Low-Rank Affinity Pursuit for Image Segmentation," *Proc. IEEE Int'l Conf. Computer Vision,* pp. 2439-2446, Nov. 2011.
[29] S. Shekhar, V.M. Patel, N.M. Nasrabadi, and R. Chellappa, "Joint Sparsity-Based Robust Multimodal Biometrics Recognition," *Proc. ECCV Workshop Information Fusion in Computer Vision for Concept Recognition,* Oct. 2012.
[30] N.H. Nguyen, N.M. Nasrabadi, and T.D. Tran, "Robust Multi-Sensor Classification via Joint Sparse Representation," *Proc. Int'l Conf. Information Fusion,* pp. 1-8, July 2011.
[31] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *J. Royal Statistical Soc.: Series B,* vol. 58, pp. 267-288, 1996.
[32] E.J. Candes, X. Li, Y. Ma, and J. Wright, "Robust Principal Component Analysis?" *J. ACM,* vol. 58, pp. 1-37, May 2011.
[33] J. Yang and Y. Zhang, "Alternating Direction Algorithms for l1 Problems in Compressive Sensing," *SIAM J. Scientific Computing,* vol. 33, pp. 250-278, 2011.
[34] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An Augmented Lagrangian Approach to the Constrained Optimization Formulation of Imaging Inverse Problems," *IEEE Trans. Image Processing,* vol. 20, no. 3, pp. 681-695, Mar. 2011.
[35] S.S.S. Crihalmeanu, A. Ross, and L. Hornak, "A Protocol for Multibiometric Data Acquisition, Storage and Dissemination," technical report, Lane Dept. of Computer Science and Electrical Eng., West Virginia Univ., 2007.
[36] A. Martnez and R. Benavente, "The AR Face Database," CVC technical report, June 1998.
[37] U. Park, R. Jillela, A. Ross, and A. Jain, "Periocular Biometrics in the Visible Spectrum," *IEEE Trans. Information Forensics and Security,* vol. 6, no. 1, pp. 96-106, Mar. 2011.
[38] A. Moorhouse, A. Evans, G. Atkinson, J. Sun, and M. Smith, "The Nose on Your Face May Not Be So Plain: Using the Nose as a Biometric," *Proc. Int'l Conf. Crime Detection and Prevention,* pp. 1-6, Dec. 2009.
[39] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell, "Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About," *Proc. IEEE,* vol. 94, no. 11, pp. 1948-1962, Nov. 2006.
[40] H. Li, K.-A. Toh, and L. Li, *Advanced Topics In Biometrics.* World Scientific Publishing Co. Pvt. Ltd., 2012.
[41] S. Pundlik, D. Woodard, and S. Birchfield, "Non-Ideal Iris Segmentation Using Graph Cuts," *Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops,* pp. 1-6, June 2008.
[42] L. Masek and P. Kovesi, "MATLAB Source Code for Biometric Identification System Based on Iris Patterns," technical report, The Univ. of Western Australia, 2003.
[43] C.W.S. Chikkerur and V. Govindaraju, "A Systematic Approach for Feature Extraction in Fingerprint Images," *Proc. Int'l Conf. Bioinformatics and Its Applications,* pp. 344-350, Dec. 2004.
[44] A. Jain, S. Prabhakar, L. Hong, and S. Pankanti, "Filterbank-Based Fingerprint Matching," *IEEE Trans. Image Processing,* vol. 9, no. 5, pp. 846-859, May 2000.
[45] J. Daugman, "How Iris Recognition Works," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 14, no. 1, pp. 21-30, Jan. 2004.
[46] B. Krishnapuram, L. Carin, M. Figueiredo, and A. Hartemink, "Sparse Multinomial Logistic Regression: Fast Algorithms and Generalization Bounds," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 27, no. 6, pp. 957-968, June 2005.
[47] C.J. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery,* vol. 2, pp. 121-167, June 1998.
[48] A. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "SimpleMKL," *J. Machine Learning Research,* vol. 9, pp. 2491-2521, 2008.
[49] R. Bolle, J. Connell, S. Pankanti, N. Ratha, and A. Senior, "The Relation between the ROC Curve and the CMC," *Proc. Fourth IEEE Workshop Automatic Identification Advanced Technologies,* pp. 15-20, 2005.
[50] K. Nandakumar, Y. Chen, S. Dass, and A. Jain, "Likelihood Ratio-Based Biometric Score Fusion," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 30, no. 2, pp. 342-347, Feb. 2008.
[51] X.F.M. Yang, L. Zhang, and D. Zhang, "Fisher Discrimination Dictionary Learning for Sparse Representation," *Proc. IEEE Int'l Conf. Computer Vision,* pp. 543-550, Nov. 2011.
[52] A.M. Bruckstein, D.L. Donoho, and M. Elad, "From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images," *SIAM Rev.,* vol. 51, no. 1, pp. 34-81, Feb. 2009.
[53] M. Elad, *Sparse and Redundant Representations.* Springer, 2010.

Sumit Shekhar received the BTech degree in electrical engineering from the Indian Institute of Technology Bombay in 2009. He is working toward the PhD degree in electrical and computer engineering at the University of Maryland, College Park. He received the University of Maryland James Clarke Graduate Fellowship for 2009-2010. He is a student member of the IEEE.

Vishal M. Patel received the BS degrees in electrical engineering and applied mathematics (with honors) and the MS degree in applied mathematics from North Carolina State University, Raleigh, in 2004 and 2005, respectively. He received the PhD degree from the University of Maryland, College Park, in electrical engineering in 2010. He received an ORAU postdoctoral fellowship in 2010. His research interests are in signal processing, computer vision, and pattern recognition with applications to biometrics and imaging. He is a member of the research faculty at the University of Maryland Institute for Advanced Computer Studies (UMIACS). He is a member of the IEEE, Eta Kappa Nu, Pi Mu Epsilon, and Phi Beta Kappa.

Nasser M. Nasrabadi received the BSc degree in English and the PhD degree in electrical engineering from the Imperial College of Science and Technology (University of London), London, England, in 1980 and 1984, respectively. From October 1984 to December 1984, he was at IBM, United Kingdom, as a senior programmer. During 1985 to 1986, he was with Philips research laboratory in New York as a member of the technical staff. From 1986 to 1991, he was an assistant professor in the Department of Electrical Engineering at Worcester Polytechnic Institute, Worcester, Massachusetts. From 1991 to 1996, he was an associate professor with the Department of Electrical and Computer Engineering at the State University of New York, Buffalo. Since September 1996, he has been a senior research scientist (ST) with the US Army Research Laboratory (ARL) working on image processing and automatic target recognition. He has served as an associate editor for the *IEEE Transactions on Image Processing*, *the IEEE Transactions on Circuits*, *Systems and Video Technology*, and *the IEEE Transactions on Neural Networks*. His interests include hyperspectral imaging, automatic target recognition, statistical machine learning theory, robotics, and neural networks applications to image processing. He is a fellow of the IEEE, ARL, and SPIE.



Rama Chellappa received the BE (Hons.) degree in electronics and communication engineering from the University of Madras, India, in 1975 and the ME (with distinction) degree from the Indian Institute of Science, Bangalore, India, in 1977. He received the MSEE and PhD degrees in electrical engineering from Purdue University, West Lafayette, Indiana, in 1978 and 1981, respectively. During 1981 to 1991, he was a faculty member in the Department of Electrical Engineering-Systems at the University of Southern California (USC). Since 1991, he has been a professor of electrical and computer engineering (ECE) and an affiliate professor of computer science at the University of Maryland (UMD), College Park. He is also affiliated with the Center for Automation Research, the Institute for Advanced Computer Studies (Permanent Member) and is serving as the chair of the Electrical and Computer Engineerintg Department. In 2005, he was named a Minta Martin Professor of Engineering. He received the US National Science Foundation (NSF) Presidential Young Investigator Award, four IBM Faculty Development Awards, an Excellence in Teaching Award from the School of Engineering at USC, and two paper awards from the International Association of Pattern Recognition (IAPR). He received the K.S. Fu Prize from IAPR. He received the Society, Technical Achievement, and Meritorious Service Awards from the IEEE Signal Processing Society. He also received the Technical Achievement and Meritorious Service Awards from the IEEE Computer Society. At UMD, he was elected as a Distinguished Faculty Research fellow, as a Distinguished scholar-teacher, received an Outstanding Innovator Award from the Office of Technology Commercialization, and an Outstanding GEMSTONE Mentor Award from the Honors College. He received the Outstanding Faculty Research Award and the Poole and Kent Teaching Award for Senior Faculty from the College of Engineering. In 2010, he was recognized as an Outstanding ECE by Purdue University. He served as the editor-in-chief of the *IEEE Transactions on Pattern Analysis and Machine Intelligence*. He has served as a General and Technical Program chair for several IEEE international and national conferences and workshops. He is a Golden Core member of the IEEE Computer Society and served as a Distinguished lecturer of the IEEE Signal Processing Society. Recently, he completed a two-year term as the president of the IEEE Biometrics Council. He holds four patents. He is a fellow of the IEEE, IAPR, OSA, and AAAS.