Joint Sparsity-Based Robust Multimodal Biometric Recognition

Sumit Shekhar, Vishal M. Patel, Nasser M. Nasrabadi* and Rama Chellappa

University of Maryland, College Park, USA
* Army Research Laboratory, USA

1. Introduction

• Traditional biometric systems rely on a single biometric signature for identification. Often such systems deal with problems as noisy data, intra-class variations and spoof attack.

• Deploying multimodal biometric systems can help address above problems by integrating information from multiple sources.

• Classification in multimodal systems involve fusion of results from individual modalities.

• Fusion can be done at feature level, score level or decision level. Due to preservation of raw information, feature-level fusion can be more discriminative.

• However, current methods in biometric literature focus on score-level and decision-level fusion methods.

• We present a robust method for feature-level fusion using joint sparsity framework.

2. Challenges and Proposed Framework

• Different sensors have different output formats, i.e., different dimensions and distribution. Often each sensor produces high dimensional features.

• Recently, sparse representation (SR) based techniques have emerged as powerful technique in biometric recognition.

• Motivated by it success, we present a joint sparsity-based algorithm for multimodal classification.

• Given input features from different modalities, we represent them by a sparse linear combination of respective training sets, while constraining them to share their sparse representation.

• An efficient algorithm based on multi-task, multivariate Lasso is presented.

3. Overview of the method

4. Problem Formulation

• For a C-class problem with D biometrics, for each biometric denote the training set as:

\[ X^i = \{X_1^i, X_2^i, \ldots, X_C^i\} \]

where, each sub-dictionary:

\[ X^i = \{X_{1,i}^i, X_{2,i}^i, \ldots, X_{D,i}^i\} \in \mathbb{R}^{D \times d_i} \]

• Given a test matrix \( Y = \{Y_1, Y_2, \ldots, Y_D\} \) with each

\[ Y^i = \{Y_{1,i}, Y_{2,i}, \ldots, Y_{C,i}\} \in \mathbb{R}^{C \times d_i} \]

we seek a matrix \( f \) which minimizes representation error, while enforcing joint sparsity across all test samples (\( q = 2 \)):

\[ F = \arg\min_{F} \frac{1}{2} \sum_{i=1}^{D} \|Y^i - X^iF\|^2 + \lambda \left(\|F\|_{1,q}\right) \]

• To make the algorithm robust to noise and occlusion, we modify the representation of \( Y^i \):

\[ Y^i = X^iF + B^iA^i + N^i \]

where, \( N^i \) is dense additive noise, and \( B^i \) is a basis in which occlusion can be sparsely represented using matrix \( A^i \).

The optimization is modified as:

\[ F, A = \arg\min_{F, A} \frac{1}{2} \sum_{i=1}^{D} \|Y^i - X^iF - B^iA^i\|^2 + \lambda_1 \left(\|F\|_{1,q}\right) + \lambda_2 \left(\|A\|_{1,q}\right) \]

5. Optimization

• We optimize using Alternating Directions of Multiplier Method as summarized:

\[ \begin{align*}
F_{t+1} &= \arg\min_{F} \|Y - X^iF\|^2 + \lambda \left(\|F\|_{1,q}\right) + \rho \langle Y - X^iF, U_t + V_t \rangle \\
U_{t+1} &= \arg\min_{U} \|Y - X^iF - U\|^2 + \lambda_1 \left(\|U\|_{1,q}\right) + \rho \langle Y - X^iF - U, V_t \rangle \\
V_{t+1} &= \arg\min_{V} \|Y - X^iF - U_t\|^2 + \lambda_2 \left(\|V\|_{1,q}\right) + \rho \langle Y - X^iF - U_t, V \rangle \\
A_{t+1} &= \arg\min_{A} \|Y - X^iF - B^iA^i\|^2 + \lambda_1 \left(\|F\|_{1,q}\right) + \lambda_2 \left(\|A\|_{1,q}\right) + \rho \langle Y - X^iF - B^iA^i, A_t \rangle
\end{align*} \]

6. Classification

• After optimization, the test sample \( Y \) can be classified as:

\[ \text{identity}(Y) = \arg\min_{i=1}^{C} \|Y - X^i\|^2 + \lambda \left(\|F\|_{1,q}\right) \]

7. Experimental Results

• We tested on the publicly available WVU Multimodal dataset, using 2 iris + 4 fingerprint modalities for 219 subjects. 4 random samples per class were used for training and rest 519 for testing.