Sparse Discrimination based Multiset Canonical Correlation Analysis for Multi-Feature Fusion and Recognition

Hongkun Ji
jihongkunnust@126.com
Xiaobo Shen
njust.shenxiaobo@gmail.com
Quansen Sun
qssun@126.com
Zexuan Ji
jizexuan840820@hotmail.com

School of Computer Science and Engineering
Nanjing University of Science and Technology
Nanjing 210094, China

Canonical correlation analysis (CCA) [3] is a powerful tool for finding the correlation between two sets of multidimensional variables. It investigates the linear correlations between two sets of random variables and linearly projects two sets of random variables into a lower-dimensional space in which they are maximally correlated. However, most of CCA-based methods are not efficient for multiple (more than two) feature representations classification tasks [5]. As a generalized extension of CCA, multiset canonical correlation analysis (MCCA) [6, 7] was proposed to solve this problem. In recent years, with the development of sparse representation [1, 2, 4, 9], sparse representation based classification (SRC) has been proved to be robust for feature selection and efficient in handling occlusion and corruption. Sparsity preserving projections (SPP) [8] aims to preserve the sparse reconstructive relationship of the data. Extensive experiments on some common datasets show that the projections obtained by SPP are invariant to rotations, rescalings and translations of the data.

However, MCCA usually fails to discover the intrinsic sparse reconstructive relationship and discriminating structure of multiple data spaces in real-world applications. In this paper, by taking discriminative information of within-class and between-class sparse reconstruction into account, we propose a novel algorithm, called sparse discrimination based multiset canonical correlations (SDbMCCs), to explicitly consider both discriminative structure and sparse reconstructive relationship in multiple representation data. In addition to maximizing between-set cumulative correlations, SDbMCC minimizes within-class sparse reconstructive distances and maximizes between-class sparse reconstructive distances, simultaneously. The feasibility and effectiveness of the proposed method is verified on four popular databases (CMU PIE, ETH-80, AR and Extended Yale-B) with promising results.

We characterize the within-class and between-class scatters, i.e., minimizing the average within-class distance and maximizing the average between-class distance, simultaneously. Given m feature sets $\{X^{(i)} = [x_1^{(i)}, x_2^{(i)}, \cdots, x_n^{(i)}] \in \Re^{d_i \times n}\}_{i=1}^m$ extracted from n samples of c classes. Let $\Phi : \Re^n \to \Re^n$ be the characteristic function that selects the coefficients of $s_j^{(i)}$ associated with the same class of the sample $x_j^{(i)}$, where $s_j^{(i)}$ is the sparse weight vector of $x_j^{(i)}$ as in SPP. For $s_j^{(i)} \in \Re^n$, with the assumption that $x_j^{(i)}$ belongs to the kth class, $\Phi(s_j^{(i)})$ is a vector whose only nonzero entries are the entries in $s_j^{(i)}$ that are associated with class k. For each feature set $X^{(i)}$, after the projections of $\{x_1^{(i)}, x_2^{(i)}, \cdots, x_n^{(i)}\}$ onto the projection axis $\alpha^{(i)}$, we can get the within-class scatter defined by

$$\begin{aligned} J_w^{(i)} &= \sum_{i=1}^{n} \| \alpha^{(i)T} x_j^{(i)} - \alpha^{(i)T} X^{(i)} \Phi(s_j^{(i)}) \|^2 \\ &= \| \alpha^{(i)T} X^{(i)} - \alpha^{(i)T} X^{(i)} \Phi(S^{(i)}) \|^2 \\ &= \alpha^{(i)T} X^{(i)} S_\Phi^{(i)} X^{(i)T} \alpha^{(i)} \end{aligned} \quad (1)$$

where $S^{(i)} = [s_1^{(i)}, s_2^{(i)}, \cdots, s_n^{(i)}]$ and $S_\Phi^{(i)} = [I - \Phi(S^{(i)})] \cdot [I - \Phi(S^{(i)})]^T$. Then we can construct the total within-class scatter for all m feature sets as follows:

$$J_w = \sum_{i=1}^{m} J_w^{(i)} = \sum_{i=1}^{m} \alpha^{(i)T} X^{(i)} S_\Phi^{(i)} X^{(i)T} \alpha^{(i)} \quad (2)$$

Obviously, minimizing the total within-class scatter J_w can lead to a better sparse reconstructive relationship of the samples belonging to the same class.

Similarly to the within-class scatter, let $\Psi : \Re^n \to \Re^n$ be the characteristic function that selects the coefficients of $s_j^{(i)}$ associated with the different classes of the sample $x_j^{(i)}$. We can construct the total between-class scatter for all m feature sets as follows:

$$J_b = \sum_{i=1}^{m} J_b^{(i)} = \sum_{i=1}^{m} \alpha^{(i)T} X^{(i)} S_\Psi^{(i)} X^{(i)T} \alpha^{(i)} \quad (3)$$

where $S_\Psi^{(i)} = [I - \Psi(S^{(i)})] \cdot [I - \Psi(S^{(i)})]^T$. Maximizing the total between-class scatter can significantly enhance the discriminative capability.

By considering both discriminative and sparse reconstructive relationship in multiple representation data, we tend to minimize the total within-class sample distance and maximize the total between-class sample distance in the low-dimensional embedding subspace, as well as maximizing the correlations. Combining with the original MCCA objective function, we can construct the model of SDbMCC as follows:

$$\begin{gathered} maxJ(\alpha) = \Sigma_{i=1}^m \Sigma_{j=1}^m \alpha^{(i)T} S_{ij} \alpha^{(j)} - \tau[\eta J_w - (1-\eta) J_b] \\ s.t.\ \Sigma_{i=1}^m \alpha^{(i)T} S_{ii} \alpha^{(i)} = 1. \end{gathered} \quad (4)$$

where $\alpha^T = [\alpha^{(1)T}, \alpha^{(2)T}, \cdots, \alpha^{(m)T}]$, S_{ii} is the within-set covariance matrix of feature set $X^{(i)}$, and S_{ij} is the between-set covariance matrix between feature sets $X^{(i)}$ and $X^{(j)}$. Because the ratio between within-class and between-class sparse representation weights suffers from a variety of noises, we further hope that the tradeoff between between-class scatters and within-class scatters, as well as the tradeoff between correlations and the sparse reconstructive relationship, can be tuned. Therefore, we construct the corresponding regularizations with two tradeoff parameters τ and η. By utilizing Lagrange multiplier, we can solve the optimization problem in Eq. (4).

[1] M. Davenport, M. Duarte, M. Wakin, D. Takhar, K. Kelly, and R. Baraniuk. The smashed filter for compressive classification and target recognition. In *In Proc. IS&T/SPIE Symposiumon Electronic Imaging: Computational Imaging*, 2006.

[2] M. Davenport, M. Wakin, and R. Baraniuk. Detection and estimation with compressive measurements. In *Technical Report*, 2007.

[3] H. Hotelling. Relations between two sets of variates. *Biometrika*, 28: 321–377, December 1936.

[4] K. Huang and S. Aviyente. Sparse representation for signal classification. In *Advances in Neural Information Processing Systems (NIPS)*, 2006.

[5] M. Kan, S.G. Shan, H.H. Zhang, S.H. Lao, and X.L. Chen. Multi-view discriminant analysis. In *European Conference on Computer Vision (ECCV)*, volume 7572, pages 808–821, 2012.

[6] J.R. Kettenring. Canonical analysis of several sets of variables. *Biometrika*, 58:433–451, 1971.

[7] A.A. Nielsen. Multiset canonical correlations analysis and multispectral, truly multitemporal remote sensing data. *IEEE Transactions on image processing*, 11:293–305, 2002.

[8] L.S. Qiao, S.C. Chen, and X.Y. Tan. Sparsity preserving projections with applications to face recognition. *Pattern Recognition*, 43:331–341, 2010.

[9] J. Wright, A. Yang, S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31:210–227, 2009.