Infrared Physics & Technology 78 (2016) 233–240

Contents lists available at ScienceDirect

Infrared Physics & Technology

journal homepage: www.elsevier.com/locate/infrared

Kernel-aligned multi-view canonical correlation analysis for image recognition

Shuzhi Su, Hongwei Ge *, Yun-Hao Yuan

Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi, China

HIGHLIGHTS

- We propose a kernel-aligned multi-view canonical correlation analysis method.
- Each feature vector is transformed into a feature matrix by kernel alignment.
- An optimized combination of multiple basis kernels is automatically learned.
- Extensive experiments have shown the superiority of the method.

ARTICLE INFO

Article history:
Received 21 June 2016
Revised 25 August 2016
Accepted 25 August 2016
Available online 27 August 2016

Keywords:
Canonical correlation analysis
Multi-view feature learning
Kernel alignment
Image recognition

ABSTRACT

Existing kernel-based correlation analysis methods mainly adopt a single kernel in each view. However, only a single kernel is usually insufficient to characterize nonlinear distribution information of a view. To solve the problem, we transform each original feature vector into a 2-dimensional feature matrix by means of kernel alignment, and then propose a novel kernel-aligned multi-view canonical correlation analysis (KAMCCA) method on the basis of the feature matrices. Our proposed method can simultaneously employ multiple kernels to better capture the nonlinear distribution information of each view, so that correlation features learned by KAMCCA can have well discriminating power in real-world image recognition. Extensive experiments are designed on five real-world image datasets, including NIR face images, thermal face images, visible face images, handwritten digit images, and object images. Promising experimental results on the datasets have manifested the effectiveness of our proposed method.

1. Introduction

Canonical correlation analysis (CCA) is an important feature learning technique that was proposed by Hotelling [1] as early as 1936. The technique is a linear method and can learn linear correlation features from two-view raw data [2]. Since raw data from real-world applications usually possess complex nonlinear relationships, CCA is difficult to cater the real-world raw data. To capture nonlinear distribution information of real-world raw data, kernel CCA (KCCA) [3] first projects raw data into higher-dimensional kernel spaces, and then the traditional CCA is implemented in the kernel spaces. Up to now, KCCA has been widely applied in real-world applications, such as facial expression recognition [4], image retrieval [5], blind analysis [6], and radar emitter identification [7]. Based on different motivations, many typical kernel-based CCA methods have been proposed recently, including kernel generalized CCA (KGCCA) [8], gradient descent kernel CCA (GDKCCA) [9], and restricted KCCA [10], and so on. Different from such the kernel-based methods that only employ a single kernel in each view, Zhu et al. [11] proposed a new kernel-based CCA method called mixed KCCA, which projects raw data of each view into a reproducing kernel Hilbert space with the mixture of two kernels, i.e. a linear combination of local and global kernels. Since the mixture of the two kernels can be treated as an independent preprocessing process before KCCA, mixed KCCA still belongs to the domain of the single kernel model [11]. In [12], Zhu et al. proposed a globalized and localized CCA method with multiple empirical kernel mappings, which maps raw single-view data into multiple kernel spaces for obtaining multi-view data. Moreover, the method is a single-view feature learning approach in essence and also employs one single kernel in each view. To reduce the

* Corresponding author.
E-mail addresses: sushuzhi@foxmail.com (S. Su), ghw8601@163.com (H. Ge), yyhzbh@163.com (Y.-H. Yuan).

http://dx.doi.org/10.1016/j.infrared.2016.08.010
1350-4495/

time consumption in real-world applications, scholars developed many parallel approaches [13–18].

As a generalized extension of CCA, multiset CCA (MCCA) [19] aims at seeking a set of linear correlation projection directions for each of multiple (more than two) views, so that the extracted correlation features in the low-dimensional spaces are maximally correlated. Recently, many nonlinear MCCA variants have been designed, but such the variants mainly focus on graph technique. Inspired by graph embedding [20], Shen et al. [21] presented a unified framework of MCCA based on graph embedding, and this framework gives a unified viewpoint to embed different graphs into MCCA. In [22], a new method called graph regularized multiset canonical correlations (GrMCC) has been also proposed for joint feature extraction, which exploits both discriminative and intrinsic geometrical information of raw data by graph regularization technique. In addition, Yuan et al. [23] proposed a Laplacian MCCA (LapMCCA) method, which simultaneously considers local within-view and local between-view correlations by using nearest neighbor graphs. In existing nonlinear MCCA variants, only a few employ kernel technique. For example, Rupnik and Shawe-Taylor [24] described a kernel MCCA (KMCCA) method that could be treated as a multi-view extension of KCCA. To the best of our knowledge, up to now, there still have been no kernel-based MCCA variants that are able to automatically determine an adaptive kernel of each view from multiple kernels.

Existing kernel-based correlation analysis methods mainly employ one single kernel in each view. However, their performance could be limited when an unsuitable kernel is utilized, and it is difficult to choose a beneficial kernel for given data. Meanwhile, they also ignore the fact that only one single kernel is usually insufficient to characterize nonlinear distribution information of a view. To solve the problem, we propose a kernel-aligned multi-view canonical correlation analysis (KAMCCA) method. In our proposed method, a 2-dimensional feature matrix can be constructed by aligning multiple kernel space projections of an original feature vector. Then, based on the feature matrices, we construct a unified optimization problem of correlation feature learning and multiple kernel fusion. By solving the optimization problem in iteration, we can obtain correlation projection directions and multi-kernel fusion vector of each view. Our KAMCCA method has the following three main characteristics. First, KAMCCA can simultaneously employ numerous kernels to better characterize the nonlinear distribution information of each view, so that nonlinear correlation features learned by KAMCCA have better discriminating power in real-world image recognition. Second, KAMCCA also eases the process of determining an adaptive kernel for each view because an optimized combination of multiple basis kernels in each view can be learned automatically. Third, KAMCCA is developed in a general way, and some kernel-based correlation analysis methods such as KCCA and KMCCA are the special cases of KAMCCA. In addition, the multi-kernel extension approach of KAMCCA can provide some inspiration for extending many single-view subspace learning methods (such as linear discriminant analysis (LDA) [25] and locality preserving projection (LPP) [26]) to their multi-kernel variants. To evaluate our proposed method, we design extensive experiments on five real-world image datasets, including NIR face images, thermal face images, visible face images, handwritten digit images, and object images. All experimental results have shown that our proposed method is an effective and relatively robust approach for image recognition.

The rest contents are organized as follows. We briefly review KMCCA in Section 2. Section 3 gives a detailed description of our proposed method. In the next section, extensive experimental results are described and analyzed. Conclusions are discussed in Section 5.

2. Related work

In this section, we briefly review KMCCA. Suppose that M view datasets from the same N objects are given as $\left\{X^{(i)} = \left[x_1^{(i)}, x_2^{(i)}, \ldots, x_N^{(i)}\right] \in R^{d^{(i)} \times N}\right\}_{i=1}^{M}$, and the $r(r = 1, 2, \ldots, N)$ th samples $\left\{x_r^{(i)}\right\}_{i=1}^{M}$ of all views come from the same object. For each view $X^{(i)}(i = 1, 2, \ldots, M)$, we suppose that there exists a nonlinear mapping $\phi^{(i)} : x_r^{(i)} \mapsto \phi^{(i)}\left(x_r^{(i)}\right) \in R^{d_\phi^{(i)} \times 1}$ that can implicitly project original samples into a higher-dimensional kernel space, and $\phi^{(i)}(X^{(i)}) = \left[\phi^{(i)}\left(x_1^{(i)}\right), \phi^{(i)}\left(x_2^{(i)}\right), \ldots, \phi^{(i)}\left(x_N^{(i)}\right)\right] \in R^{d_\phi^{(i)} \times N}$.

KMCCA aims at seeking correlation projection directions $\left\{\alpha_\phi^{(i)} \in R^{d_\phi^{(i)} \times 1}\right\}_{i=1}^{M}$ by the below optimization problem [24]:

$$\begin{aligned} &\max_{\alpha_\phi^{(i)}} \sum_{i=1}^{M} \sum_{j=i+1}^{M} \alpha_\phi^{(i)T} \phi^{(i)}(X^{(i)}) \phi^{(j)}(X^{(j)})^T \alpha_\phi^{(j)} \\ &s.t.\ \alpha_\phi^{(i)T} \phi^{(i)}(X^{(i)}) \phi^{(i)}(X^{(i)})^T \alpha_\phi^{(i)} = 1, i = 1, 2, \ldots, M \end{aligned} \tag{1}$$

Note that we assume that $\phi^{(i)}(X^{(i)})$ has been centered, i.e. $\frac{1}{N}\sum_{r=1}^{N} \phi^{(i)}\left(x_r^{(i)}\right) = 0, i = 1, 2, \ldots, M$. The details about the centering process can be found in [27]. Due to the two equalities $\alpha_\phi^{(i)T} \phi^{(i)}(X^{(i)}) \phi^{(i)}(X^{(i)})^T \alpha_\phi^{(i)} = 1$ and $\alpha_\phi^{(i)T} \phi^{(i)}(X^{(i)}) \phi^{(j)}(X^{(j)})^T \alpha_\phi^{(j)} = \alpha_\phi^{(j)T} \phi^{(j)}(X^{(j)}) \phi^{(i)}(X^{(i)})^T \alpha_\phi^{(i)} (i, j = 1, 2, \ldots, M)$, Eq. (1) can be transformed into the following optimization problem:

$$\begin{aligned} &\max_{\alpha_\phi^{(i)}} \sum_{i=1}^{M} \sum_{j=1}^{M} \alpha_\phi^{(i)T} \phi^{(i)}(X^{(i)}) \phi^{(j)}(X^{(j)})^T \alpha_\phi^{(j)} \\ &s.t.\ \alpha_\phi^{(i)T} \phi^{(i)}(X^{(i)}) \phi^{(i)}(X^{(i)})^T \alpha_\phi^{(i)} = 1,\ i = 1, 2, \ldots, M \end{aligned} \tag{2}$$

According to the kernel trick [27], it is assumed that $\alpha_\phi^{(i)} = \phi^{(i)}(X^{(i)})\beta^{(i)}$, where $\beta^{(i)} \in R^{N \times 1}, i = 1, 2, \ldots, M$. Then, Eq. (2) can be rewritten as follows:

$$\begin{aligned} &\max_{\beta^{(i)}} \sum_{i=1}^{M} \sum_{j=1}^{M} \beta^{(i)T} K^{(i)} K^{(j)T} \beta^{(j)} \\ &s.t.\ \beta^{(i)T} K^{(i)} K^{(i)T} \beta^{(i)} = 1, i = 1, 2, \ldots, M \end{aligned} \tag{3}$$

where $K^{(i)} = \phi^{(i)}(X^{(i)})^T \phi^{(i)}(X^{(i)}) \in R^{N \times N}$ is a kernel matrix, $i = 1, 2, \ldots, M$.

By means of Lagrange multiplier technique [2], Eq. (3) can be converted into a multivariate eigenvalue problem (MEP) [28]:

$$K\beta = \Lambda K_D \tag{4}$$

where $K \in R^{NM \times NM}$ is a black matrix with (i, j) the block element as $K^{(i)}K^{(j)T}$, $i, j = 1, 2, \ldots, M$, $K_D = diag(K^{(1)}K^{(1)T}, K^{(2)}K^{(2)T}, \ldots, K^{(M)}K^{(M)T}) \in R^{NM \times NM}$, $\beta^T = (\beta^{(1)T}, \beta^{(2)T}, \ldots, \beta^{(M)T}) \in R^{1 \times NM}$, and $\Lambda = diag(\lambda_1 I_N, \lambda_2 I_N, \ldots, \lambda_M I_N)$ with multivariate eignenvalues $\{\lambda_i\}_{i=1}^{M}$ and the identity matrix $I_N \in R^{N \times N}$. Since the MEP has no analytical solutions, i.e. exact solutions, some iterative methods [28,24] have been proposed for its solutions.

3. Kernel-aligned multi-view canonical correlation analysis

3.1. Model of KAMCCA

Recent researches [25,11] have shown that it is necessary to consider multiple kernels rather than one single kernel in real-world applications. For each view $X^{(i)}(i = 1, 2, \ldots, M)$, we consider $G_i(G_i \geqslant 2)$ implicit mappings:

$$\phi_g^{(i)} : x_r^{(i)} \mapsto \phi_g^{(i)}\left(x_r^{(i)}\right) \in R^{d_{\phi_g}^{(i)} \times 1}$$