Hindawi
Journal of Electrical and Computer Engineering
Volume 2018, Article ID 8745251, 5 pages
https://doi.org/10.1155/2018/8745251

Research Article

A New Feature Extraction Algorithm Based on Orthogonal Regularized Kernel CCA and Its Application

Xinchen Guo,[1] Xiuling Fan,[2] Xiantian Xi,[3] and Fugeng Zeng[1]

[1]*College of Ocean Information Engineering, Hainan Tropical Ocean University, Sanya, China*
[2]*Henan Xuehang Education and Information Service Co., Zhengzhou, China*
[3]*Shanghai Renyi Technology Co., Ltd., Shanghai, China*

Correspondence should be addressed to Fugeng Zeng; zengfugeng@foxmail.com

Received 3 April 2018; Revised 21 July 2018; Accepted 23 August 2018; Published 29 October 2018

Academic Editor: Jar Ferr Yang

In this paper, an orthogonal regularized kernel canonical correlation analysis algorithm (ORKCCA) is proposed. ORCCA algorithm can deal with the linear relationships between two groups of random variables. But if the linear relationships between two groups of random variables do not exist, the performance of ORCCA algorithm will not work well. Linear orthogonal regularized CCA algorithm is extended to nonlinear space by introducing the kernel method into CCA. Simulation experimental results on both artificial and handwritten numerals databases show that the proposed method outperforms ORCCA for the nonlinear problems.

1. Introduction

Canonical correlation analysis (CCA) is a technique of multivariate statistical analysis, which deals with the mutual relationships of two sets of variables [1–3]. This method extracts the representative variables which are the linear combination of the variables in each group. The relationships between new variables can reflect the overall relationships between two groups of variables [4].

The orthogonal regularization canonical correlation analysis (ORCCA) algorithm [5] is that the original formula of CCA algorithm with orthogonal constraints is substituted for CCA conjugate orthogonalization [6, 7]. When the number of samples is less and the sample distribution patterns of different classifications are different, the ORCCA algorithm has the better ability of classification. A suboptimal solution to eigenvalue decomposition problem can be obtained by introducing two regularization parameters [8]. So, the complexity of time and space for the quadratic optimization problem should be considered at the same time. ORCCA algorithm is the same as CCA algorithm that both their goals look for the linear combinations of the variables in each group. But when the nonlinear relationships between the variables exist, ORCCA algorithm cannot extract effectively the comprehensive variables.

In this paper, the kernel method [9–11] is introduced into ORCCA algorithm, and ORKCCA algorithm is presented. The kernel method maps the linear inseparable data in the low-dimensional space into a higher-dimensional space [12, 13]. In the higher-dimensional space, the characteristics of the data can be extracted and analyzed through the linear method. By introducing kernel function, the computation of the orthogonal regularization canonical correlation analysis extends to a nonlinear feature space. Experimental results show that the accuracies of classification of our method in the nonlinear space are significantly improved. The experimental results show ORKCCA is feasible.

2. Orthogonal Regularized CCA Algorithm

Given n pairs of pairwise samples $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)^T$ and $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)^T$, where $\mathbf{x}_i \in R^p$, $\mathbf{y}_i \in R^q$ $(i = 1, 2, \ldots, n)$. We assume that the samples have been centered. ORCCA algorithm aims at finding a pair of projection directions $\mathbf{a}$ and $\mathbf{b}$ which satisfy the following optimal problem [5].

$$\begin{aligned} &\min_{\mathbf{a},\mathbf{b}} \quad \frac{1}{n}\sum_{i=1}^{n}\left\|\mathbf{a}^T\mathbf{x}_i-\mathbf{b}^T\mathbf{y}_i\right\|^2,\\ &\text{s.t.} \quad \mathbf{a}^T\mathbf{a}=1,\\ &\qquad\quad \mathbf{b}^T\mathbf{b}=1. \end{aligned} \tag{1}$$

The objective function in Equations (1) can be expanded as follows:

$$\frac{1}{n}\sum_{i=1}^{n}\left\|\mathbf{a}^T\mathbf{x}_i-\mathbf{b}^T\mathbf{y}_i\right\|^2=\mathbf{a}^T\mathbf{S}_{xx}\mathbf{a}+\mathbf{b}^T\mathbf{S}_{yy}\mathbf{b}-2\mathbf{a}^T\mathbf{S}_{xy}\mathbf{b}, \tag{2}$$

where $\mathbf{S}_{xx}=(1/n)\mathbf{X}^T\mathbf{X}$, $\mathbf{S}_{yy}=(1/n)\mathbf{Y}^T\mathbf{Y}$, and $\mathbf{S}_{xy}=(1/n)\mathbf{X}^T\mathbf{Y}$.

The optimal model in Equation (1) can be rewritten as

$$\begin{aligned} &\max_{\mathbf{a},\mathbf{b}} \quad 2\mathbf{a}^T\mathbf{S}_{xy}\mathbf{b}-\mathbf{a}^T\mathbf{S}_{xx}\mathbf{a}-\mathbf{b}^T\mathbf{S}_{yy}\mathbf{b},\\ &\text{s.t.} \quad \mathbf{a}^T\mathbf{a}=1,\\ &\qquad\quad \mathbf{b}^T\mathbf{b}=1. \end{aligned} \tag{3}$$

According to the Lagrange multipliers method, Lagrange function is as follows:

$$\begin{aligned} L(\mathbf{a},\mathbf{b}) = {} & 2\mathbf{a}^T\mathbf{S}_{xy}\mathbf{b}-\mathbf{a}^T\mathbf{S}_{xx}\mathbf{a}-\mathbf{b}^T\mathbf{S}_{yy}\mathbf{b}-\lambda_1\left(\mathbf{a}^T\mathbf{a}-1\right)\\ & -\lambda_2\left(\mathbf{b}^T\mathbf{b}-1\right), \end{aligned} \tag{4}$$

where both λ_1 and λ_2 are Lagrange multipliers.

The solutions to Equation (4) are given as follows:

$$\left(\mathbf{S}_{xy}\left(\mathbf{S}_{yy}+\lambda_2\mathbf{I}_q\right)^{-1}\mathbf{S}_{xy}^T-\mathbf{S}_{xx}\right)\mathbf{a}=\lambda_1\mathbf{a}, \tag{5}$$

$$\left(\mathbf{S}_{xy}^T\left(\mathbf{S}_{xx}+\lambda_1\mathbf{I}_p\right)^{-1}\mathbf{S}_{xy}-\mathbf{S}_{yy}\right)\mathbf{b}=\lambda_2\mathbf{b}, \tag{6}$$

where $\mathbf{I}_p$ and $\mathbf{I}_q$ denote identity matrices of size $p*p$ and $q*q$, respectively.

Both λ_1 and λ_2 in Equations (5) and (6) are called regularization parameters. By solving Equation (5), the eigenvalues $\lambda_1^{(1)},\lambda_1^{(2)},\ldots,\lambda_1^{(p)}$ and their corresponding eigenvectors $\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_p$ can be obtained. The eigenvalues $\lambda_2^{(1)},\lambda_2^{(2)},\ldots,\lambda_2^{(q)}$ and their corresponding eigenvectors $\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_q$ can be obtained from Equation (6).

3. Orthogonal Regularized Kernel CCA Algorithm (ORKCCA)

ORCCA algorithm can give the linear relationships between two groups of random variables. But if the linear relationships between two groups of random variables do not exist, the performance of ORCCA will not work well. The kernel method is an effective way to analyze the nonlinear pattern problem. So, the kernel method is introduced into ORCCA algorithm, and ORKCCA algorithm is proposed.

Both Φ_x and Φ_y are nonlinear mappings which map original random variables $\mathbf{x}_i$ and $\mathbf{y}_i$ into $\Phi_x(\mathbf{x}_i)$ and $\Phi_y(\mathbf{y}_i)$ in P-dimensional space $F_x(P>p)$ and Q-dimensional space $F_y(Q>q)$, $i=1,2,\ldots,n$. Let $\mathbf{a}=\Phi_x(\mathbf{X})^T\boldsymbol{\alpha}$, $\mathbf{b}=\Phi_y(\mathbf{Y})^T\boldsymbol{\beta}$, where $\Phi_x(\mathbf{X})\in R^{n\times P}$, $\Phi_y(\mathbf{Y})\in R^{n\times Q}$ $\boldsymbol{\alpha},\boldsymbol{\beta}\in R^n$.

ORCCA is implemented in higher-dimensional spaces F_x and F_y. So, Equation (7) can be obtained by substituting $\mathbf{a}$, $\mathbf{b}$, $\Phi_x(\mathbf{x}_i)$, and $\Phi_y(\mathbf{y}_i)$ into Equation (1) as follows:

$$\begin{aligned} &\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \quad \frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{\alpha}^T\Phi_x(\mathbf{X})\Phi_x(\mathbf{x}_i)-\boldsymbol{\beta}^T\Phi_y(\mathbf{Y})\Phi_y(\mathbf{y}_i)\right\|^2,\\ &\text{s.t.} \quad \boldsymbol{\alpha}^T\Phi_x(\mathbf{X})\Phi_x(\mathbf{X})^T\boldsymbol{\alpha}=1,\\ &\qquad\quad \boldsymbol{\beta}^T\Phi_y(\mathbf{Y})\Phi_y(\mathbf{Y})^T\boldsymbol{\beta}=1. \end{aligned} \tag{7}$$

Expanding the objective function in Equation (7), we get

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n}\left\|\boldsymbol{\alpha}^T\Phi_x(\mathbf{X})\Phi_x(\mathbf{x}_i)-\boldsymbol{\beta}^T\Phi_y(\mathbf{Y})\Phi_y(\mathbf{y}_i)\right\|^2\\ &\quad=\boldsymbol{\alpha}^T\frac{1}{n}\Phi_x(\mathbf{X})\Phi_x(\mathbf{X})^T\Phi_x(\mathbf{X})\Phi_x(\mathbf{X})^T\boldsymbol{\alpha}\\ &\qquad+\boldsymbol{\beta}^T\Phi_y(\mathbf{Y})\Phi_y(\mathbf{Y})^T\Phi_y(\mathbf{Y})\Phi_y(\mathbf{Y})^T\boldsymbol{\beta}\\ &\qquad-2\boldsymbol{\alpha}^T\frac{1}{n}\Phi_x(\mathbf{X})\Phi_x(\mathbf{X})^T\Phi_y(\mathbf{Y})\Phi_y(\mathbf{Y})^T\boldsymbol{\beta}. \end{aligned} \tag{8}$$

Applying the kernel trick to Equation (8), $\mathbf{K}_x$ and $\mathbf{K}_y\in R^{n\times n}$ can be computed, namely, $\mathbf{K}_x=\Phi_x(X)\Phi_x(X)^T=(\Phi_x(x_i)^T\Phi_x(x_j))_{n\times n}=(k(x_i,x_j))_{n\times n}$ $\mathbf{K}_y=\Phi_y(Y)\Phi_y(Y)^T=(\Phi_y(y_i)^T\Phi_y(y_j))_{n\times n}=(k(y_i,y_j))_{n\times n}$, where $k(\cdot,\cdot)$ is kernel function. Centralization is exerted on $\mathbf{K}_x$ and $\mathbf{K}_y$. The optimal model in which the kernel method is introduced can be given by using Equation (9):

$$\begin{aligned} &\max_{\boldsymbol{\alpha},\boldsymbol{\beta}} \quad 2\boldsymbol{\alpha}^T\mathbf{M}_{xy}\boldsymbol{\beta}-\boldsymbol{\alpha}^T\mathbf{M}_{xx}\boldsymbol{\alpha}-\boldsymbol{\beta}^T\mathbf{M}_{yy}\boldsymbol{\beta},\\ &\text{s.t.} \quad \boldsymbol{\alpha}^T\mathbf{K}_x\boldsymbol{\alpha}=1,\\ &\qquad\quad \boldsymbol{\beta}^T\mathbf{K}_y\boldsymbol{\beta}=1, \end{aligned} \tag{9}$$

where $\mathbf{M}_{xy}=(1/n)\mathbf{K}_x^T\mathbf{K}_y$, $\mathbf{M}_{xx}=(1/n)\mathbf{K}_x^T\mathbf{K}_x$, and $\mathbf{M}_{yy}=1/n\mathbf{K}_y^T\mathbf{K}_y$.

According to the Lagrange multiplier method, the Lagrange function is as follows

$$\begin{aligned} L'(\boldsymbol{\alpha},\boldsymbol{\beta}) = {} & 2\boldsymbol{\alpha}^T\mathbf{M}_{xy}\boldsymbol{\beta}-\boldsymbol{\alpha}^T\mathbf{M}_{xx}\boldsymbol{\alpha}-\boldsymbol{\beta}^T\mathbf{M}_{yy}\boldsymbol{\beta}-\zeta_1\left(\boldsymbol{\alpha}^T\mathbf{K}_x\boldsymbol{\alpha}-1\right)\\ & -\zeta_2\left(\boldsymbol{\beta}^T\mathbf{K}_y\boldsymbol{\beta}-1\right), \end{aligned} \tag{10}$$

where ζ_1 and ζ_2 are Lagrange multipliers. Taking the partial derivatives of $L'(\boldsymbol{\alpha},\boldsymbol{\beta})$ with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and letting them zero, we get

$$\begin{cases} \dfrac{\partial L'}{\partial\boldsymbol{\alpha}}=2\left(\mathbf{M}_{xy}\boldsymbol{\beta}-\mathbf{M}_{xx}\boldsymbol{\alpha}-\zeta_1\boldsymbol{\alpha}\right)=0,\\[2ex] \dfrac{\partial L'}{\partial\boldsymbol{\beta}}=2\left(\mathbf{M}_{xy}^T\boldsymbol{\alpha}-\mathbf{M}_{yy}\boldsymbol{\beta}-\zeta_2\boldsymbol{\beta}\right)=0, \end{cases} \tag{11}$$

where $\mathbf{M}_{xx}$ and $\mathbf{M}_{yy}$ are positive semidefinite matrices and ζ_1 and ζ_2 are positive numbers.

So, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be obtained from Equation (11):

$$\alpha = \left(M_{xx} + \zeta_1 I_P\right)^{-1} M_{xy}\beta, \quad (12)$$

$$\beta = \left(M_{yy} + \zeta_2 I_Q\right)^{-1} M_{xy}^T \alpha, \quad (13)$$

where I_P and I_Q are the identity matrices of size $P * P$ and $Q * Q$, respectively.

Equations (14) and (15) can be obtained through replacing $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ with their expressions in Equations (12) and (13), respectively.

$$\left(M_{xy}\left(M_{yy} + \zeta_2 I_Q\right)^{-1} M_{xy}^T - M_{xx}\right)\alpha = \zeta_1 \alpha, \quad (14)$$

$$\left(\mathbf{M}_{xy}^T \left(\mathbf{M}_{xx} + \zeta_1 \mathbf{I}_P\right)^{-1} \mathbf{M}_{xy} - \mathbf{M}_{yy}\right)\boldsymbol{\beta} = \zeta_2 \boldsymbol{\beta}. \quad (15)$$

As like before, both λ_1 and λ_2 in Equations (14) and (15) are called regularization parameters. By solving Equation (14), the eigenvalues $\zeta_1^{(1)}, \zeta_1^{(2)}, \ldots, \zeta_1^{(n)}$ and their corresponding eigenvectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_P$ can be obtained. The eigenvalues $\zeta_2^{(1)}, \zeta_2^{(2)}, \ldots, \zeta_2^{(Q)}$ and their corresponding eigenvectors $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_Q$ can be obtained from Equation (15).

4. Simulation Experiments

In this section, we evaluate our method compared with ORCCA on artificial and handwritten numerals databases.

4.1. Experiment on Artifical Databases. The pairwise samples $\mathbf{X}$ and $\mathbf{Y}$ are generated from the expressions in Equations (16) and (17), respectively.

$$\mathbf{x} = \begin{pmatrix} \theta \\ \cos\dfrac{\theta}{2} \\ \sin 3\theta \\ \sin\theta\cos\theta \end{pmatrix} + \boldsymbol{\varepsilon}_1, \quad (16)$$

$$\mathbf{y} = e^{2\theta}\begin{pmatrix} \cos\theta^2 \\ \sin\dfrac{\theta}{3} \end{pmatrix} + \boldsymbol{\varepsilon}_2, \quad (17)$$

where θ obeys uniform distribution on $[-\pi, \pi]$ and $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are Gaussian noise with standard deviation 0.05. The radial basis function $k(\mathbf{x}, \mathbf{y}) = \exp(-|\mathbf{x} - \mathbf{y}|^2/2\sigma^2)$ is chosen as kernel function, where $\sigma = 1.0$.

4.1.1. Determining Regularization Parameters. For the selection of the regularization parameters, by far there is no reliable method to determine the optimal values. In this paper, in order to simplify the calculation, let $\lambda = \lambda_1 = \lambda_2$ and $\zeta = \zeta_1 = \zeta_2$. The regularization parameters were chosen from 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, and 1. This method is used in the literature [5].

According to Equations (16) and (17), 100 pairs of data are randomly generated as the training samples. Canonical variables are calculated from the ORCCA and ORKCCA algorithms for the different values of regularization parameters. The correlation coefficients of canonical variables are sorted by the descending order. Many pairs of canonical variables can be gained from the two algorithms. For the sake of simplicity, the most representative of the former two groups of canonical variables are examined.

The average value of the correlation coefficients of the former two groups of canonical variables is regarded as criterion that judges the regularization parameters is good or not. The larger the average value is, the better the regularization parameters are.

Table 1 lists the average value of the correlation coefficients of the former two groups of canonical variables for the different values of the regularization parameters.

Table 1 shows that the optimal values of the regularization parameters for the ORCCA and ORKCCA algorithms are 10^{-3} and 10^{-1}, respectively. The optimal regularization parameters are used to perform simulations in the next section.

4.1.2. Simulation Experiment 1. According to Equations (16) and (17), 200 pairs of data are randomly generated as the test samples. For the regularization parameters $\lambda = 10^{-3}$ and $\zeta = 10^{-1}$ in the ORCCA and ORKCCA algorithms, the canonical variables are obtained for test samples, respectively. The correlation coefficients of the canonical variables are sorted in the descending order.

Tables 2 and 3 list the correlation coefficients of the first two groups of canonical variables for ORCCA and ORKCCA algorithms. u_1 and v_1 denote the first group of canonical variables. u_2 and v_2 are the second group of canonical variables.

The experimental results in Tables 2 and 3 show that the correlationships between the same pair of the canonical variables are better than that between the different pairs of canonical variables, especially for nonlinear data.

4.1.3. Simulation Experiment 2. According to Equations (16) and (17), 5 pairs of data are randomly generated as the sample data. Each pair of sample data represents the center data of each class. 100 pairs of data for each class are given by adding Gaussian noise with standard deviation of 0.05 to each class center data. So we have five class data, which contains 100 samples for each class.

100, 175, and 250 pairs of data are chosen from the 500 pairs of the whole data as the training samples, respectively. The rest 400, 325, and 250 pairs of data are the test samples, respectively. The classification experiments based on K-neighbors algorithm are carried out on the test samples data which are preprocessed in the above way. And, the accuracies of classification are given. For the test samples with 400, 325, and 250 pairs of data, the experiments are performed 15 times, respectively. The accuracies of classification for 400, 325, and 250 pairs of data are the averages of the accuracies of classification for the 15 experiments results,

Table 1: The mean values of the correlation coefficients of the former two groups of canonical variables for the different values of the regularization parameters from ORCCA and ORKCCA.

Regularization parameters	Mean values of the correlation coefficients	
	ORCCA	ORKCCA
10^{-5}	0.74	0.80
10^{-4}	0.77	0.84
10^{-3}	**0.82**	0.89
10^{-2}	0.81	0.92
10^{-1}	0.79	**0.93**
1	0.80	0.92

Table 2: The correlation coefficients of the first two groups of canonical variables for ORCCA.

	v_1	v_2
u_1	0.58	0.14
u_2	0.21	0.36

Table 3: The correlation coefficients of the first two groups of canonical variables for ORkCCA.

	v_1	v_2
u_1	0.91	0.08
u_2	0.06	0.83

respectively. Table 4 gives the accuracies of classification for ORCCA and ORKCCA for the test samples with the different number.

In Table 4, the first column is the numbers of the training samples and the second column and the third column are the accuracies of classification for ORCCA and ORKCCA for the training samples with the different number. The experimental results show that the accuracies of classification for ORKCCA are higher than those for ORCCA. So, the performance of ORKCCA outperforms that of ORCCA for the nonlinear problem. The comparison curves of the accuracies of classification for ORCCA and ORKCCA are given in Figure 1.

4.2. Experiments on Handwritten Numerals Databases. The Concordia University CENPARMI database of handwritten Arabic numerals have 10 classes, that is, 10 digits (from 0 to 9), and 600 samples for each. The first 400 samples are used as the training set, and the remaining samples as the test set in each class. Then, the training samples and the test samples are 4000 and 2000, respectively. The handwritten digital images are preprocessed by the method given in [14]. Four kinds of features are extracted as follows: X^{G} (256-dimensional Gabor transformation feature), X^{L} (121-dimensional Legendre moment feature), X^{P} (36-dimensional Pseudo-Zernike moment feature), and X^{Z} (30-dimensional Zernike moment feature).

For the choice of the regularization parameters, let $\lambda = \lambda_1 = \lambda_2$ and $\zeta = \zeta_1 = \zeta_2$. The regularization parameters were chosen from 10^{-5}, 10^{-3}, and 1. The results of our method are

Table 4: Comparisons of the accuracies of classification for ORCCA and ORKCCA.

Numbers of the training samples	ORCCA (%)	ORKCCA (%)
100	65.0	73.1
175	72.4	77.8
250	76.5	83.6

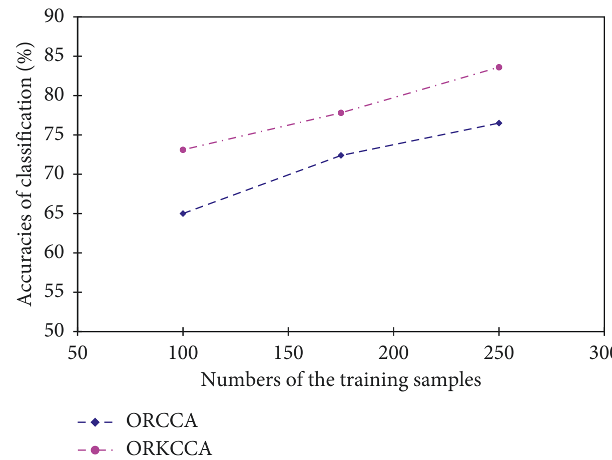

Figure 1: Comparison curves of the accuracies of classification for ORCCA and ORKCCA.

Table 5: Comparisons of the accuracies of classification for ORCCA and ORKCCA in different feature combinations and regularization parameters.

Feature combinations	ORCCA			ORKCCA		
	$\lambda = 10^{-5}$	$\lambda = 10^{-3}$	$\lambda = 1$	$\zeta = 10^{-5}$	$\zeta = 10^{-3}$	$\zeta = 1$
X^{G}-X^{L}	0.9314	0.9300	0.9375	0.9625	0.9681	0.9687
X^{G}-X^{P}	0.9230	0.9228	0.9248	0.9511	0.9525	0.9536
X^{G}-X^{Z}	0.9180	0.9196	0.9196	0.9482	0.9518	0.9520
X^{L}-X^{P}	0.9187	0.9187	0.9190	0.9500	0.9533	0.9545
X^{L}-X^{Z}	0.9200	0.9205	0.9235	0.9574	0.9600	0.9615
X^{P}-X^{Z}	0.7413	0.7413	0.7525	0.8436	0.8450	0.8450

compared with the results of ORCCA in order to verify the effectiveness of ORKCCA. Table 5 lists the accuracies of classification for ORCCA and ORKCCA in different feature combinations and regularization parameters. Experimental results show that (1) the classification effect of the two methods is the best as the regularization parameter is 1; (2) the classification accuracies of ORKCCA are higher than that of ORCCA for different features combinations; (3) the classification accuracies of ORKCCA in the regularization parameters 10^{-5} and 10^{-3} are higher than those of ORCCA in the regularization parameters 1.

5. Conclusions

An orthogonal regularized kernel CCA algorithm for nonlinear problem is presented. By introducing the kernel function, our proposed algorithm is more suitable for solving nonlinear problem. Contrast experiments of ORCCA and ORKCCA are performed on artificial and handwritten numerals databases. Experimental results show that the proposed method outperforms ORCCA for the correlation coefficients of canonical variables and the accuracies of classification on the test data. The experimental results show ORKCCA is feasible.

Data Availability

The experiments in paper were performed by the author Xi 2 years ago. Some troubles happened to his computer. The data can not be gotten from his computer. I'm sorry that the data is unable to be provided.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are grateful to the support of the Hainan Provincial Natural Science Foundation (117150) and the Scientific Research Foundation of Hainan Tropical Ocean University (RHDXB201624).

References

[1] Q. H. Ran, Z. N. Shi, and Y. P. Xu, "Canonical correlation analysis of hydrological response and soil erosion under moving rainfall," *Journal of Zhejiang University-SCIENCE A*, vol. 14, no. 5, pp. 353–361, 2013.

[2] B. K. Sarkar and C. Chakraborty, "DNA pattern recognition using canonical correlation algorithm," *Journal of Biosciences*, vol. 40, no. 4, pp. 709–719, 2015.

[3] R. R. Sarvestani and R. Boostani, "FF-SKPCCA: kernel probabilistic canonical correlation analysis," *Applied Intelligence*, vol. 46, no. 2, pp. 438–454, 2016.

[4] E. Sakar, H. Ünver, S. Keskin, and Z. M. Sakar, "The investigation of relationships between some fruit and kernel traits with canonical correlation analysis in ankara region walnuts," *Erwerbs-Obstbau*, vol. 58, no. 1, pp. 19–23, 2015.

[5] S. D. Hou and Q. S. Sun, "An orthogonal regularized CCA learning algorithm for feature fusion," *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 785–792, 2014.

[6] X. Shen and Q. Sun, "Orthogonal multiset canonical correlation analysis based on fractional-order and its application in multiple feature extraction and recognition," *Neural Process Letters*, vol. 42, no. 2, pp. 301–316, 2015.

[7] Y. H. Yuan, Y. Li, X. B. Shen, Q. S. Sun, and J. L. Yang, "Laplacian multiset canonical correlations for multiview feature extraction and image recognition," *Multimedia Tools and Applications*, vol. 76, no. 1, pp. 731–755, 2015.

[8] X. Xing, K. Wang, T. Yan, and Z. Lv, "Complete canonical correlation analysis with application to multi-view gait recognition," *Pattern Recognition*, vol. 50, pp. 107–117, 2016.

[9] H. Joutsijoki and M. Juhola, "Kernel selection in multi-class support vector machines and its consequence to the number of ties in majority voting method," *Artificial Intelligence Review*, vol. 40, no. 3, pp. 213–230, 2013.

[10] S. Wang, Z. Deng, F. L. Chung, and W. Hu, "From Gaussian kernel density estimation to kernel methods," *International Journal of Machine Learning and Cybernetics*, vol. 4, no. 2, pp. 119–137, 2013.

[11] X. Chen, R. Tharmarasa, T. Kirubarajan, and M. Mcdonald, "Online clutter estimation a Gaussian kernel density estimator for multitarget tracking," *Radar Sonar and Navigation IET*, vol. 9, no. 1, pp. 1–9, 2014.

[12] O. Taouali, I. Jaffel, H. Lahdhiri, M. F. Harkat, and H. Messaoud, "New fault detection method based on reduced kernel principal component analysis (RKPCA)," *International Journal of Advanced Manufacturing Technology*, vol. 85, no. 5, pp. 1547–1552, 2016.

[13] K. Yoshida, J. Yoshimoto, and K. Doya, "Sparse kernel canonical correlation analysis for discovery of nonlinear interactions in high-dimensional data," *BMC Bioinformatics*, vol. 18, no. 1, pp. 108–118, 2017.

[14] Z. Hu, Z. Lou, J. Yang, K. Liu, and C. Suen, "Handwritten digital recognition based on multi-classifier combination," *Chinese Journal Computers*, vol. 22, no. 4, pp. 369–374, 1999.