Application to Three-Dimensional Canonical Correlation Analysis for Feature Fusion in Image Recognition

Xiaogang Gong
School of Electronics and Information Engineering, Sichuan University, Chengdu, China
Email: scugxg@gmail.com

Jiliu Zhou, Huilin Wu, Gang Lei and Xiaohua Li
College of Computer Science, Sichuan University, Chengdu, China
Email: { zhoujl@scu.edu.cn, whl0912@hotmail.com, cdleigang@gmail.com, lxhw@scu.edu.cn }

Abstract—This paper presents a three-dimensional canonical correlation analysis (TCCA) method, and applies it to feature fusion for image recognition. It is an extension of traditional canonical correlation analysis (CCA) and two-dimensional canonical correlation analysis (2DCCA). Considering two views of a three-dimensional data, the TCCA can directly find the relations between them without reshaping the data into matrices or vectors. We stress that TCCA dramatically reduce the computational complexity, compared to the CCA and 2DCCA. To evaluate the algorithm, we are using Gabor wavelet to generate the three-dimensional data, and fusing them at the feature level by TCCA. Some experiments on ORL database and JAFEE database and compared with other methods, the results show that the TCCA not only the computing complexity is lower, the recognition performance is better but also suitable for data fusion.

Index Terms—canonical correlation analysis, feature fusion, three-dimensional

I. INTRODUCTION

Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are very popular approaches to dimensionality reduction in the field of statistics and machine learning. When observations arrive from two sources that share some mutual information a related approach called the Canonical Component Analysis (CCA) was developed by Hotelling (1936). Since then it has been widely used in text data mining[1], face recognition[2], facial expression recognition[3], etc.

CCA is a way of measuring the linear relationship between two multidimensional variables. In this analysis, one finds a linear combination of the first set of variables and a linear combination of the second set of variables such that they both have unit variance and the Pearson correlation coefficient between them is maximum. Thus the obtained pair of linear combinations are called the first canonical variables, and the correlation is called the first canonical correlation. This process is repeated to obtain the second, third, … canonical variables and correlations with the additional restriction that the pair of linear combinations currently being computed are uncorrelated with all the previously obtained pairs. Use of few canonical variables to perform data analysis is in fact a general way of dimension reduction. An important property of CCA is that they are invariant with respect to affine transformations of the variables. The same as PCA and LDA, it has drawbacks: small sample size(SSS) problem, where the dimensionality is higher than the number of vectors in the training set[4], singularity problem of the covariance matrix in the case of the high-dimensional space, and structural problem that samples vector operation loses spatial information. To address these problems, many two-dimensional algorithms, 2DPCA[5], 2DLDA[6], 2DCCA[7, 8] have been proposed.

However, in the real world, most of the source data, such as video sequence, object structure is the three-dimensional. So, it is essential to study the three dimensional version pattern analysis algorithms, which can be used on three dimensional objects directly. Inspire of[7, 8], We propose a method namely three-dimensional Canonical Component Analysis(TCCA) which is an extension of classical CCA and apply it to feature fusion for image recognition.

Data fusion is a multilevel, multifaceted process dealing with the automatic detection, association, correlation, estimation and combination of data from single and multiple sources[9], that achieve improved accuracies and more reliable information. Fusion may be useful for several objectives such as detection, recognition, identification, tracking, change detection, decision making, etc. These objectives may be encountered in many application domains as Defense, Robotics, Medicine, Space, etc., Using an efficient fusion scheme, one may expect significant advantages as:

- Improved confidence in decisions due to the use of complementary information (e.g. silhouette of objects
from visible image, active/non-active status from Infra-Red image, speed and range from radar, etc.).

- Improved performance to countermeasures (it is very hard to camouflage an object in all possible wave-bands).
- Improved performance in adverse environmental conditions. Typically smoke or fog cause bad visible contrast and some weather conditions (rain) cause low thermal contrast (Infra Red imaging), combining both types of sensors should give better overall performance.

Fusion processes are often categorized as low (pixel fusion), intermediate (feature fusion) or high level fusion (decision fusion) depending on the processing stage at which fusion takes place.

Low level fusion combines several sources of raw data to produce new raw data that is expected to be more informative and synthetic than the inputs. Intermediate level fusion combines various features. Those features may come from several raw data sources (several sensors, different moments, etc.) or from the same raw data. In the latter case, the objective is to find relevant features among available features that might come from several feature extraction methods. The objective is to obtain a limited number of relevant features. High level fusion combines decisions coming from several experts.

Feature fusion plays an important role in the process data fusion. The advantage of feature fusion is obvious. As a matter of fact, the different feature vector extracts from the same pattern always reflects the different feature of patterns. By combining these different features, it not only obtains more mutual information of one pattern but also eliminates redundant information.

II. RELATED WORK

Recently, there has been increased interest in the use of CCA for feature fusion in various pattern recognition applications[10-17], in [10], CCA is applied to feature fusion and image recognition for the first time. The method using correlation feature of two groups of feature can find more effective discriminate information. In [11], the same idea is used in a block-based approach where the sample images are divided into two blocks, and canonical features are extracted, which are then combined linearly to obtain better discriminating vectors for recognition.

A multiset integrated canonical correlation analysis(MICCA) is presented in [12], it extracts multiple features from the same patterns by using different feature extraction methods and finally expresses the integral correlation among multi-group features using MICCA to remove the redundancy between features. In [13], a color image CCA (CICCA) is proposed, which can extract three color components and provide the analytical solution. It means the method can directly acquire the typically correlated features of three input datasets.

The above methods combine the features from single modality using CCA while the following methods fuse data from two modalities.

In [14], CCA is used to fuse the two modalities (face and gait) at the feature level that achieves the superior recognition performance in gender recognition. A kernel based CCA[15] is presented to fusion ear and profile face based multimodal biometrics for recognition. The method can extract the nonlinear canonical correlation features as the associated features of ear and profile face. In [16], for better speaker identification, the highly correlated components (speech and lip texture or movement) are fused using CCA to form audiovisual feature synchronization. In [17], the palmprint and finger geometry are segmented from a whole hand image, that gets two feature vectors, then fused them together using CCA improving the average recognition rate.

To address the three-dimensional problem mentioned before, we generate two three-dimensional features using Gabor wavelet filter from one source and using feature level fusion by TCCA method to achieve a higher performance. Fig. 1 shows the proposed feature fusion recognition system.

![Figure 1. Feature fusion recognition system](image)

The rest paper is organized as follows: Section 3 introduces the traditional CCA. Section 4 describes the proposed algorithm TCCA in detail. Section 5 gives two experiments on ORL database and JAFEE database to evaluate algorithm, and Section 6 is the conclusions.

III. CANONICAL COMPONENT ANALYSIS

Given two random variables \( \{ x_t \in \mathbb{R}^m, t = 1,2,\ldots,N \} \) and \( \{ y_t \in \mathbb{R}^n, t = 1,2,\ldots,N \} \), \( \bar{x} \) and \( \bar{y} \) denotes the mean value respectively, centered source data with \( \tilde{x}_t = x_t - \bar{x} \) and \( \tilde{y}_t = y_t - \bar{y} \) can obtain zero mean dataset \( \{ \tilde{x}_t \in \mathbb{R}^m, t = 1,2,\ldots,N \} \) and \( \{ \tilde{y}_t \in \mathbb{R}^n, t = 1,2,\ldots,N \} \). In computer vision, \( \tilde{x} \) and \( \tilde{y} \) can be seen as two views of one observation: Such as the images collected from different sensors, lighting conditions or different transformations (spatial or spectral). So features extracted from two views are correlated. Feature fusion using CCA is that to find a pair of projections \( u_1 \) and \( v_1 \), to make the canonical variables \( \tilde{x}_1^* = u_1^T \tilde{x} \) and \( \tilde{y}_1^* = v_1^T \tilde{y} \) have the maximum correlation coefficient \( \rho_1 \). See (1).
\[
\rho_i = \frac{\text{cov}(u_i^T \tilde{x}, v_i^T \tilde{y})}{\sqrt{\text{var}(u_i^T \tilde{x}) \text{var}(v_i^T \tilde{y})}} = \frac{u_i^T C_{xy} v_i}{\sqrt{(u_i^T C_{xx} u_i)(v_i^T C_{yy} v_i)}}
\]

(1)

Where \(C_{xx}, C_{yy}\) are the autocorrelation matrices of \(\tilde{x}, \tilde{y}\), respectively, while \(C_{xy}\) is the cross-covariance matrix of them.

\[
\begin{align*}
C_{xx} &= \frac{1}{N} \sum_{t=1}^{N} \tilde{x}_t \tilde{x}_t^T \\
C_{yy} &= \frac{1}{N} \sum_{t=1}^{N} \tilde{y}_t \tilde{y}_t^T \\
C_{xy} &= \frac{1}{N} \sum_{t=1}^{N} \tilde{x}_t \tilde{y}_t^T
\end{align*}
\]

The problem about the first pair of matrices with maximum correlation is formulated as following equation:

\[
\begin{align*}
\arg \max_{u_i, v_i} & \quad (u_i^T C_{xy} v_i) \\
\text{s.t.} & \quad (u_i^T C_{xx} u_i) = 1 \\
& \quad (v_i^T C_{yy} v_i) = 1
\end{align*}
\]

The CCA problem can be posed as a constrained optimization problem using Lagrange multipliers and the canonical variables can be calculated by solving a generalized eigenvalue solution.

Then finding a second pair of canonical variables \(\tilde{x}_2^*, \tilde{y}_2^*\), which are uncorrelated with their first canonical variables \(\tilde{x}_1^*, \tilde{y}_1^*\), We only need to analyze a few pairs of canonical variables.

IV. THREE-DIMENSIONAL CANONICAL COMPONENT ANALYSIS

Traditional CCA usually converts a matrix to vector first, such processing changes the spacial structure of source data and also causes dimension disaster. To overcome those problems, two-dimensional canonical correlation analysis(2DCCA)[7] was proposed which processing matrix directly. As much source data is three-dimensional, we expand CCA and 2DCCA to a three-dimensional version as follows:

Let \(A \in R^{I_x \times I_y \times I_z}\) denotes a three-dimensional source data, \(I_x, I_y, I_z\) are the dimensions of \(x, y, z\) directions. \(A_{i,j,k}\) denotes an element of \(A\), where \(1 \leq i \leq I_x\), \(1 \leq j \leq I_y\), \(1 \leq k \leq I_z\). The dot product of two three-dimensional variables is defined as

\[
\langle A, B \rangle = \sum_{i,j,k} A_{i,j,k} B_{i,j,k}
\]

The norm of \(A\) is defined by

\[
\|A\| = \sqrt{\langle A, A \rangle}
\]

As the same as matrix, three-dimensional data in each direction can be flattened into their respective vector space. The \(x\) direction flattened matrix is denoted as \(A_{(x)} \in R^{(I_y \times I_z)}\), \(I_y\) and \((I_y \times I_z)\) are the rows and columns respectively of \(A_{(x)}\). The product of \(A_{(x)}\) and matrix \(U\) is defined as \(U^T A_{(x)}\) or \(A_{(x)} U\), and they are equal. Fig. 2 shows the data is flattened according \(x\) direction.

Consider two three-dimensional source data \(\{A_i \in R^{I_x \times I_y \times I_z}\}_{i=1}^{N}\) and \(\{B_i \in R^{I_x \times I_y \times I_z}\}_{i=1}^{N}\) which are real world of random variables \(\mathcal{A}\) and \(\mathcal{B}\). The mean of \(\mathcal{A}\) and \(\mathcal{B}\) is defined as \(M_A = \frac{1}{N} \sum_{i=1}^{N} A_i\), \(M_B = \frac{1}{N} \sum_{i=1}^{N} B_i\), then centered source data is denoted by \(\tilde{A}_i = A_i - M_A\), \(\tilde{B}_i = B_i - M_B\), as the same as traditional CCA, the purpose of TCCA is to find projections \(u_x, u_y, u_z\) and \(v_x, v_y, v_z\) that makes correlation between \(A_{sx}, u_{sx}, u_{syz} u_z\) and \(B_{sx}, v_{sx}, v_{syz} v_z\) maximum. Inspired of[18], we construct objective function as follows:

\[
\begin{align*}
J_1 &= \arg \max_{\mathbf{u}_x, \mathbf{u}_y, \mathbf{v}_x, \mathbf{v}_y, \mathbf{v}_z} \text{cov}
\langle A_{sx}, u_{sx}, u_{syz} u_z, B_{sx}, v_{sx}, v_{syz} v_z \rangle \\
\text{s.t.} \quad & \text{var}(A_{sx}, u_{sx}, u_{syz} u_z) = 1 \\
& \text{var}(B_{sx}, v_{sx}, v_{syz} v_z) = 1
\end{align*}
\]

The objective function (3) is a non-linear constraints with the high-dimensional non-linear optimization problem. It is difficult to directly find the closed form solution. In this paper, we use alternative numerical iterative method to find the solution. We only discuss the solution of transforms \(\mathbf{u}_x\) and \(\mathbf{v}_x\), which belongs to \(x\)
direction. And the transforms of \( y \), \( z \) directions are the same.

Assuming that \( u_y, v_y \) and \( u_z, v_z \) are fixed. Defining matrices:

\[
\Sigma_{AB}^{x} = \frac{1}{N} \sum_{i=1}^{N} (\tilde{A}_{ixy} u_{xy} z) (\tilde{B}_{ixy} v_{xy} z)^{T}
\]

\[
\Sigma_{AA}^{x} = \frac{1}{N} \sum_{i=1}^{N} (\tilde{A}_{ixy} u_{xy} z) (\tilde{A}_{ixy} v_{xy} z)^{T}
\]

\[
\Sigma_{BB}^{x} = \frac{1}{N} \sum_{i=1}^{N} (\tilde{B}_{ixy} u_{xy} z) (\tilde{B}_{ixy} v_{xy} z)^{T}
\]

Then \( \text{cov}(\tilde{A}_{ixy} u_{xy} z\tilde{B}_{ixy} v_{xy} z) \) can be replaced by \( u_{x}^{T} \Sigma_{AB}^{x} v_{x} \), and (3) can be formulated as the following optimization problem:

\[
\arg\max_{u_{x},v_{x}} v_{x}^{T} \Sigma_{AB}^{x} v_{x}
\]

\[
s.t. \quad u_{x}^{T} \Sigma_{AA}^{x} u_{x} = 1
\]

\[
v_{x}^{T} \Sigma_{BB}^{x} v_{x} = 1
\]

This problem can be reduced to a linear equations by writing the Lagrangian of the optimization problem (let \( \lambda \) denotes the Lagrange’s multiplier) and setting its gradient to zero. This result in:

\[
J = u_{x}^{T} \Sigma_{AB}^{x} v_{x} + \lambda_{u_{x}} (1 - u_{x}^{T} \Sigma_{AA}^{x} u_{x}) + \lambda_{v_{x}} (1 - v_{x}^{T} \Sigma_{BB}^{x} v_{x})
\]

\[
\text{s.t.} \quad u_{x}^{T} \Sigma_{AA}^{x} u_{x} = 1
\]

\[
v_{x}^{T} \Sigma_{BB}^{x} v_{x} = 1
\]

Then we have \( \Sigma_{AB}^{x} v_{x} + \lambda_{v_{x}} (1 - u_{x}^{T} \Sigma_{AA}^{x} u_{x}) + \lambda_{v_{x}} (1 - v_{x}^{T} \Sigma_{BB}^{x} v_{x}) \)

\[
\Sigma_{AB}^{x} U_{x}^{T} = \lambda \Sigma_{AA}^{x} U_{x}^{T} + \lambda \Sigma_{BB}^{x} V_{x}^{T}
\]

With the same method, the transforms \( u_y, v_y \) and \( u_z, v_z \) can be solved when \( u_x, v_x, u_y, v_y \) and \( u_x, v_x, u_y, v_y \) are given respectively. We compute the \( d \) pairs of transforms each has the maximum correlation \( (\rho) \) and consist of projection matrices:

\[
U_{d} = [u_{x1}, u_{x2}, \ldots, u_{xd}] \quad \text{and} \quad V_{d} = [v_{x1}, v_{x2}, \ldots, v_{xd}]
\]

Similarly, we can get the projection matrices \( U_{d}, V_{d} \). In our proposed TCCA, the generalized eigenvalue problem for much smaller size matrices, compared to the traditional CCA, which can reduce the computation cost dramatically. Detailed algorithm process is shown in TABLE I.

### TABLE I. Algorithm of TCCA

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Input: ( A_{x}, B_{y} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1:</td>
<td>centering data: ( \tilde{A}<em>{x} = A</em>{x} - M_{A}, \tilde{B}<em>{y} = B</em>{y} - M_{B} )</td>
</tr>
<tr>
<td>Step 2:</td>
<td>initialization: ( U_{d}, V_{d} )</td>
</tr>
<tr>
<td>Step 3:</td>
<td>for ( t = 1, 2, \ldots, t_{max} )</td>
</tr>
<tr>
<td>(a)</td>
<td>Compute: ( \sum_{AB}^{x}, \sum_{AA}^{x}, \sum_{BB}^{x} ), ( \Sigma_{AB}^{x} = (\Sigma_{AB}^{x})^{T} )</td>
</tr>
<tr>
<td>(b)</td>
<td>Compute: ( \sum_{AB}^{x}, \sum_{AA}^{x}, \sum_{BB}^{x} ), ( \Sigma_{AB}^{x} = (\Sigma_{AB}^{x})^{T} )</td>
</tr>
<tr>
<td>(c)</td>
<td>Compute: ( \sum_{AB}^{x}, \sum_{AA}^{x}, \sum_{BB}^{x} ), ( \Sigma_{AB}^{x} = (\Sigma_{AB}^{x})^{T} )</td>
</tr>
<tr>
<td>(d)</td>
<td>If the norm of projection matrices between two iteration is less than a threshold ( \varepsilon )</td>
</tr>
</tbody>
</table>

End loop.

Step 4: | Output: |
| Projection matrices of \( x \) direction: \( U_{d}, V_{d} \) |
| Projection matrices of \( y \) direction: \( U_{d}, V_{d} \) |
| Projection matrices of \( z \) direction: \( U_{d}, V_{d} \) |

### V. EXPERIMENTS AND RESULTS

Two groups of experiments are designed to testify the effectiveness of the proposed approach. One is face recognition using ORL database. The other is facial expression recognition using JAFFE database.

The three-dimensional data is generated using Gabor filter. The properties of Gabor wavelet are in accordance with the characteristics of human vision for the kernels of Gabor wavelet have the similar structure with the two-dimensional field profiles of the mammalian cortical simple cells. Because of this biological relevance, Gabor wavelet can capture the subtle details of an image, and render salient visual properties such as spatial localization, orientation selectivity and spatial frequency characteristic.

All these features are useful for image understanding and recognition[19]. We encode face...
images using Gabor wavelet from different scale and orientation. Fig. 3 shows a pair of three-dimensional data from Gabor wavelet feature extraction (4 scales, 5 directions and 4 directions, 5 scales for example).

A. Experiment on the ORL database

The ORL database includes 400 images with 119*92 dimensions from 40 individuals, each providing 10 different images. The series of 10 images presents variations in facial expression, in facial position (slightly rotated faces) and in some other details like glass/no glass. The Fig. 4 shows a subset of the ORL database.

To reduce the computation complexity, each image is scaled down to the size of 28*23 pixels. We randomly select 5 to 9 samples from each class for training, and remaining samples are used for testing. This process is repeated for 3 times and 15 different training and test sets are created to evaluate TCCA’s robust. For ease of representation, the experiment is named as Gm/Pn which means that the m random images are selected for training and the n remaining images for testing. We extract 5 scales, 6 directions and 5 directions, 6 scales of total 30 Gabor filters to consist of a pair of 28*23*30 three-dimensional data. The nearest-neighbor classifier is used for recognition. For comparison, gray feature and fusion Gabor feature are chosen. Three methods: PCA, 2DPCA, 2DLDA also are tested. After 3 times of such experiments the average recognition rate is shown in TABLE II.

In ORL database, image size is 28*23, after extracting 5 scales, 6 directions Gabor feature, the data is vector into 28*23*30=19320-dimensional feature in PCA, you encounter a 19320*19320 covariance matrix eigenvalue decomposition, that is high computational complexity and time-consuming. The matrix eigenvalue decomposition using TCCA transforms the data to three small matrices: 28*28, 23*23, 30*30, this process effectively reduces the computational complexity, and as the dimension of the covariance matrix is not high, it reduces the possibility of the singular matrix.

B. Experiment on the JAFFE database

To test the applicability of our algorithm, we chose the JAFFE(Japanese Female Facial Expression) database for facial expression recognition. The database contains 10 females, each of seven different expressions (neutral, sadness, surprise, happiness, disgust, fear, anger). Most of the females have three images, and the others have two or four. A total of 213 images, each image is 200*180 resolution of 256 grayscale. Some examples are shown in Fig. 5.

The results show that using gray feature, 2DPCA as to avoid the image matrix to quantify, can keep the original data structure, and the recognition rate has increased than the one-dimensional vector PCA algorithm, while the 2DLDA as a supervised learning method, which contains the class information has higher recognition rate than the non-supervised learning method 2DPCA. The Gabor features have a stronger description ability of image information, and has a higher recognition rate than the gray feature, although this method as an unsupervised learning method, but because of the role of feature fusion, compared with only a single grayscale or Gabor features, the recognition rate is higher.

In JAFFE database, image size is 28*23, after extracting 5 scales, 6 directions Gabor feature, the data is vector into 28*23*30=19320-dimensional feature in PCA, you encounter a 19320*19320 covariance matrix eigenvalue decomposition, that is high computational complexity and time-consuming. The matrix eigenvalue decomposition using TCCA transforms the data to three small matrices: 28*28, 23*23, 30*30, this process effectively reduces the computational complexity, and as the dimension of the covariance matrix is not high, it reduces the possibility of the singular matrix.

<table>
<thead>
<tr>
<th>(Recognition rate)</th>
<th>G5/P5</th>
<th>G6/P4</th>
<th>G7/P3</th>
<th>G8/P2</th>
<th>G9/P1</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA (gray)</td>
<td>0.9150</td>
<td>0.9313</td>
<td>0.9444</td>
<td>0.9708</td>
<td>0.9667</td>
</tr>
<tr>
<td>2DPCA (gray)</td>
<td>0.9417</td>
<td>0.9437</td>
<td>0.9635</td>
<td>0.9875</td>
<td>0.9750</td>
</tr>
<tr>
<td>2DLDA (Gray)</td>
<td>0.9483</td>
<td>0.9583</td>
<td>0.9639</td>
<td>0.9667</td>
<td>0.9833</td>
</tr>
<tr>
<td>PCA (Gabor)</td>
<td>0.9617</td>
<td>0.9708</td>
<td>0.9722</td>
<td>0.9875</td>
<td>0.9833</td>
</tr>
<tr>
<td>TCCA (Gabor)</td>
<td>0.9733</td>
<td>0.9875</td>
<td>0.9833</td>
<td>0.9917</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

For the convenience, we chose 210 images, which including 3 images each facial expression (total 30 images) of one person for experiment. We extract 8 scales, 5 directions and 8 directions, 5 scales of total 40 Gabor filters to consist of a pair of 25*23*40 three-dimensional data. And flatten source data to two-dimensional matrices: 200*115 and 125*184, according to x and y direction for 2DCCA, while reshape source data to 23000 dimensions for traditional CCA, that leads to a 23000*23000 matrix for eigenvalue decomposition. We randomly select 10, 15, 20 25 images for training and the rest for testing. This process is repeated for 3 times. The best average facial expression recognition rate is shown in TABLE III.
facial expression recognition. Since maintaining the structural information between pixels, 2DPCA, 2DLDA, 2DCCA have a higher recognition rate than PCA, and as considering class information in 2DLD., it does a well performance than 2DPCA. Because of feature fusion, TCCA is better than any other single feature methods. To better illustrate the effectiveness of TCCA, the recognition rates versus the number of dimensions are plotted in Fig. 6.

<table>
<thead>
<tr>
<th>Recognition Rate</th>
<th>G10/P20</th>
<th>G15/P15</th>
<th>G20/P10</th>
<th>G25/P5</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA (gray)</td>
<td>0.5429</td>
<td>0.6254</td>
<td>0.7714</td>
<td>0.8286</td>
</tr>
<tr>
<td>2DPCA (gray)</td>
<td>0.5714</td>
<td>0.6444</td>
<td>0.8143</td>
<td>0.8476</td>
</tr>
<tr>
<td>2DLDA (Gray)</td>
<td>0.5833</td>
<td>0.6762</td>
<td>0.8086</td>
<td>0.8762</td>
</tr>
<tr>
<td>2DCCA (Gabor)</td>
<td>0.5143</td>
<td>0.6381</td>
<td>0.7524</td>
<td>0.8571</td>
</tr>
<tr>
<td>TCCA (Gabor)</td>
<td>0.5881</td>
<td>0.6667</td>
<td>0.8190</td>
<td>0.9143</td>
</tr>
</tbody>
</table>

Table III. Recognition Rate of Different Methods

From the curves in the figure, the conclusions can be given that the TCCA method has the highest recognition rate when feature dimensions are greater than 12. If the feature dimensions are small, 2DPCA, 2DLDA have well performance and 2DCCA is unsteady. It is proving that TCCA has strong robustness and stability.

**VI. CONCLUSIONS**

In this paper, we proposed a three-dimensional Canonical Component Analysis (TCCA) method and apply it to feature fusion and image recognition. Using of correlation features between two groups of features as effective discriminant information is not only suitable for data fusion, but also eliminates the redundant information within the features.

Besides, the method can effectively preserve the spatial structure of the source data and effectively decrease the possibility of the singular problem of high dimensional matrix’s eigenvalue decomposition. Be compared with other methods, experiments show that our proposed TCCA has better recognition rate as well as decreases the computation cost.

**REFERENCES**


Xiaogang Gong, was born in Shaanxi, China, in 1980. He received his Master degree in the area of Condensed Physics from Sichuan University in 2007. He is a PhD candidate at the School of Electronics and Information Engineering, Sichuan University.

His major is Communication and Information System. His main research interests include signal processing, biometrics recognition, especially in human face and gait recognition.