Cross-view Graph Embedding

Zhiwu Huang[1,2], Shiguang Shan[1], Haihong Zhang[3], Shihong Lao[3], Xilin Chen[1]

[1]Key Lab of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]OMRON Social Solutions Co. Ltd, Kyoto, Japan
{zhiwu.huang, shiguang.shan}@vipl.ict.ac.cn, angelazhang@ssb.kusatsu.omron.co.jp, lao@ari.ncl.omron.co.jp, xilin.chen@vipl.ict.ac.cn

Abstract. Recently, more and more approaches are emerging to solve the cross-view matching problem where reference samples and query samples are from different views. In this paper, inspired by Graph Embedding, we propose a unified framework for these cross-view methods called Cross-view Graph Embedding. The proposed framework can not only reformulate most traditional cross-view methods (e.g., CCA, PLS and CDFE), but also extend the typical single-view algorithms (e.g., PCA, LDA and LPP) to cross-view editions. Furthermore, our general framework also facilitates the development of new cross-view methods. In this paper, we present a new algorithm named Cross-view Local Discriminant Analysis (CLODA) under the proposed framework. Different from previous cross-view methods only preserving inter-view discriminant information or the intra-view local structure, CLODA preserves the local structure and the discriminant information of both intra-view and inter-view. Extensive experiments are conducted to evaluate our algorithms on two cross-view face recognition problems: face recognition across poses and face recognition across resolutions. These real-world face recognition experiments demonstrate that our framework achieves impressive performance in the cross-view problems.

1 Introduction

Real world data analytic problems in social media often involve data from multi-view. Typical examples are multi-view face images, which consist of face with different kinds of viewpoints, or contain face with various of resolutions. These different views represent different angles to reveal the fundamental properties of the study subjects. In most cases, the multi-view problem can be solved by converting to multiple two-view problems. Therefore, in this work, we are only focusing on the *cross-view* problem where reference images and query images may come from two different views.

To deal with the cross-view problem, a few methods has been proposed. Two of the most popular unsupervised approaches are Canonical Correlation Analysis (CCA) [1–3] and Partial Least Squares (PLS) [4–6]. Both of them assume that the data consists of feature vectors that arose from two views which are from the same underlying semantic instance. Specifically, they seek for two transforms to project samples from different views into a common space. Another popular method is recently proposed by Li et al.[7]. They developed a CLPM algorithm with an objective function similar as CCA by introducing a penalty weighting matrix to preserve the locality. The common point of the above three methods is that they all try to contract pairwise sample while ignoring the label information of non-corresponding pair of samples.

To make a full use of the supervised information, several supervised methods recently are developed. Sun et al. [8] proposed Discriminative Canonical Correlation Analysis (DCCA) which incorporates the class information into the framework of CCA for recognition with missing samples. Lin et al.[9] proposed a common discriminant feature extraction (CDFE) method to learn a pair of transformations by incorporating both the empirical discriminative power and the local smoothness of the feature transformation. Lei et al. [10] recently presented Coupled Spectral Regression (CSR) by deriving the solutions from the view of spectral regression [11] which uses the label information as its response.

The objective of most cross-view methods is to learn view-specific transformations to project samples from two different views into a common space. Inspired by the idea of graph embedding [12], in this paper, we develop a framework called Cross-view Graph Embedding. Firstly, we show that most previous cross-view methods mentioned above can be reformulated within the proposed framework; Secondly, using the proposed framework, existing single-view algorithms following the original Graph Embedding framework can be easily extended to corresponding cross-view editions; Thirdly, our general framework also facilitates the development of new cross-view data analysis methods. As an example, we propose a new algorithm named Cross-view Local Discriminant Analysis (CLODA). Extensive experiments are conducted to evaluate our proposed algorithms and show that our methods perform impressively on the cross-view face recognition.

The remainder of the paper is organized as follows. Section 2 briefly reviews the original Graph Embedding framework and details our proposed Cross-view Graph Embedding framework for cross-view data analysis methods. Section 3 introduces our new method Cross-view Local Discriminant Analysis (CLODA) developed under our general framework. Section 4 provides the experimental results on two real cross-view face recognition tasks which include face recognition across poses and face recognition across resolutions. Section 5 concludes our work in this paper.

2 Cross-view Graph Embedding: A General Framework for Cross-view Methods

Yan et al. [12] proposed a general framework called Graph Embedding to unify most single-view methods under different constraints. Furthermore, the general framework provides a common perspective in understanding the relationships between all kinds of single-view algorithms and designing new algorithms. Specifically, let an undirected weighted graph $G = \{X, W\}$ be the intrinsic graph with vertex set X and similarity matrix W. And let graph $G^P = \{X, W^P\}$ be the penalty graph whose vertices X are the same as those of G. The vertex set $X = \{x_1, x_2, \ldots, x_N\}$ represents the sample set, and each element of similarity matrix W represents the similarity between a pair of vertices. The graph-preserving criterion is given as follows:

$$\omega^* = \mathop{\arg\min}_{\substack{\omega^T X B X^T \omega = d \\ or\ \omega^T \omega = d}} \sum_{i \neq j} \|\omega^T x_i - \omega^T x_j\|^2 W_{ij}. \tag{1}$$

where B typically is a diagonal matrix for scale normalization and is also defined on the penalty graph. That is, $B = D^P - W^P$, where $D^P_{ij} = \sum_{i \neq j} W^P_{ij}$.

The above Graph Embedding framework elegantly unified algorithms seeking one single-view transform for dimension reduction. However, it is not applicable for cross-view problem. To address this inability, we extend it to Cross-view Graph Embedding. In this section, we first present our theoretical formulation of Cross-view Graph Embedding, followed by the revisiting of previous cross-view methods under the proposed framework. We then show how some representative single-view methods can be correspondingly extended to cross-view editions with the proposed framework.

2.1 Cross-view Graph Embedding

In the cross-view problem, there are two types of sample sets: the query samples and the reference samples captured from two different views. Suppose we have a training set of N_q samples in the query space and N_r samples in the reference space from C classes, denoted by $(X_i^{(q)}, C_i^{(q)})_{i=1}^{N_q}$ and $(X_j^{(r)}, C_j^{(r)})_{j=1}^{N_r}$. To enable the comparison of the query samples and reference samples in practice, it is necessary to learn two different projections that can transform the two different sources to a common space respectively. We denote the projection for the query source by $f_q : X_i^{(q)} \to Y_i^{(q)}$ and the projection for the reference source by $f_r : X_i^{(r)} \to Y_i^{(r)}$.

Let $G = (X^{(q)}, X^{(r)}, W)$ be an undirected weighted graph with two vertex sets $X^{(q)}$, $X^{(r)}$ and similarity matrix $W = \Re^{N_q \times N_r}$. As shown in Figure 1, the original graph could be divided into one inter-graph which is a bipartite graph and two intra-graphs. While the inter-graph describes the relationships between the query samples and the reference samples, the two intra-graphs represent the graph structure of the query samples $X^{(q)}$ and the reference samples $X^{(r)}$ respectively. We denote the inter-graph as $G = ((X^{(q)}, X^{(r)}), W)$ and the two intra-graphs as $G^{(q)} = (X^{(q)}, W^{(q)})$ and $G^{(r)} = (X^{(r)}, W^{(r)})$.

As the original Graph Embedding, we also define an intrinsic graph and a penalty graph for both the inter-graph and the intra-graphs, they are denoted as $G_I = ((X^{(q)}, X^{(r)}), W_I)$, $G_P = ((X^{(q)}, X^{(r)}), W_P)$, $G_I^{(q)} = (X^{(q)}, W_I^{(q)})$, $G_P^{(q)} = (X^{(q)}, W_P^{(q)})$, $G_I^{(r)} = (X^{(r)}, W_I^{(r)})$ and $G_P^{(r)} = (X^{(r)}, W_P^{(r)})$, respectively. In this paper, the goal of the Cross-view Graph Embedding is to seek a common space where the relationships between the vertex pairs in inter-graph and

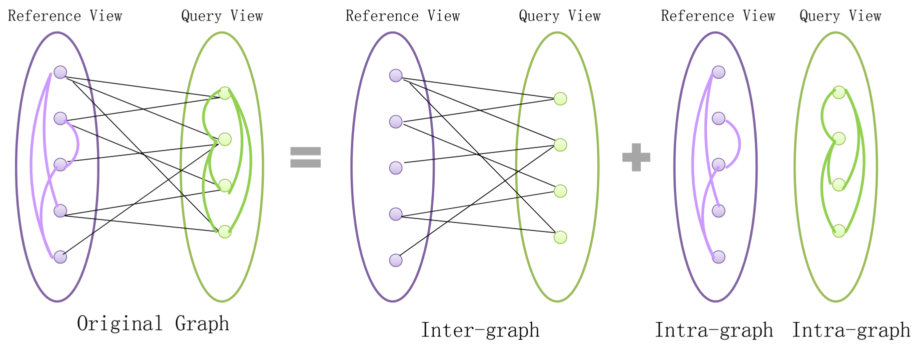

Fig. 1. Illustration of Cross-view Graph Embedding: the original graph could be divided into one inter-graph which is a bipartite graph and two intra-graphs.

intra-graphs are mostly preserved. The formulation of our Cross-view Graph Embedding framework is given as follows:

$$[\omega_q^*, \omega_r^*] = \arg\min \left\{ \sum_{i,j}^{N_q,N_r} Z_{ij}^{(qr)} W_{Iij} + \beta_q \sum_{i \neq j}^{N_q,N_q} Z_{ij}^{(qq)} W_{I}{}_{ij}^{(q)} + \beta_r \sum_{i \neq j}^{N_r,N_r} Z_{ij}^{(rr)} W_{I}{}_{ij}^{(r)} \right\},$$

$$\text{s.t.} \quad \sum_{i,j}^{N_q,N_r} Z_{ij}^{(qr)} W_{Pij} + \beta_q^{'} \sum_{i \neq j}^{N_q,N_q} Z_{ij}^{(qq)} W_{P}{}_{ij}^{(q)} + \beta_r^{'} \sum_{i \neq j}^{N_r,N_r} Z_{ij}^{(rr)} W_{P}{}_{ij}^{(r)} = d, \tag{2}$$

where, *from here on, we formulate the* $Z_{ij}^{(uv)}$ *as follows*:

$$Z_{ij}^{(uv)} = \|\omega_u^T x_i^{(u)} - \omega_v^T x_j^{(v)}\|^2, \tag{3}$$

ω_q^* and ω_r^* are the projection direction for query sample and reference sample. They tend to balance the feature extraction optimization with tuning the parameters β_q, β_r, $\beta_q^{'}$ and $\beta_r^{'}$. Like Eq. (1), the subject of our objective function can also be formulated as:

$$\sum_{i,j}^{N_q,N_r} Z_{ij}^{(qr)} W_{Pij} + \beta_q^{'} \omega_q^T X_q B_q X_q^T \omega_q + \beta_r^{'} \omega_r^T X_r B_r X_r^T \omega_r = d. \tag{4}$$

Following the Appendix in this paper, the above proposed objective function is formulated as a standard generalized eigenvalue problem which can be efficiently solved. The details of the general solution to this framework are presented in the Appendix.

2.2 Revisiting of Existing Cross-view Methods Under Cross-view Graph Embedding Framework

Given the above formulation of Cross-view Graph Embedding, it is necessary to check its capacity of deducting existing cross-view data analysis methods, such as

Table 1. Reformulations of the existing cross-view methods under Cross-view Graph Embedding.

Method	W_I	$W_I^{(q)}/W_I^{(r)}$	W_P	$W_P^{(q)}/B_q/W_P^{(r)}/B_r$
CCA	I	$W_I^{(q)} = W_I^{(r)} = 0$	0	$B_q = B_r = I$
PLS	I	$W_I^{(q)} = W_I^{(r)} = 0$	0	$B_q = (X_q^T X_q)^{-1}$ $B_r = (X_r^T X_r)^{-1}$
CDFE	$\frac{1}{N_1}$, if $C_i^{(q)} = C_j^{(r)}$ $-\frac{\alpha}{N_2}$, if $C_i^{(q)} \neq C_j^{(r)}$	${W_I}_{ij}^{(q)} = \frac{1}{N_q} exp(-\frac{\\|x_i^{(q)} - x_j^{(q)}\\|}{\sigma_q^2})$ ${W_I}_{ij}^{(r)} = \frac{1}{N_r} exp(-\frac{\\|x_i^{(r)} - x_j^{(r)}\\|}{\sigma_r^2})$	0	$W_I^{(q)} = W_I^{(r)} = I$

CCA, PLS and CDFE. Both CCA [1–3] and PLS [4–6] attempt to maximize the correlation between a projection $\omega_q^T X_q$ of X_q and $\omega_r^T X_r$ of X_r, with the difference lying in their constraint conditions. Under our Cross-view Graph Embedding, the detailed reformulations of CCA and PLS are shown in Table 1.

CDFE [9] formulates the learning objective by incorporating both the empirical discriminative power and the local smoothness of the feature transformation. It can be reformulated using our Cross-view Graph Embedding framework as shown in Table 1. Note that N_1 is the number of pairs from the same class, N_2 is the number of pairs from different classes.

Table 1 lists the similarity and constraint matrices for all of the above mentioned methods. As shown in the Table 1, all methods including CCA, PLS and CDFE can be unified into our proposed framework only with different of the similarity matrices and constrain matrices.

2.3 Cross-view Extensions of Single-view Algorithms in Original Graph Embedding Framework

As a framework extended from Graph Embedding, it is also expected that our framework can easily extend existing typical single-view algorithms in Graph Embedding framework to their cross-view editions for cross-view analysis. In this subsection, we will show how PCA, LDA and LPP etc. are reformulated to their cross-view editions respectively under the proposed framework. Table 2 lists the similarity and constraint matrices for all cross-view editions of the above mentioned methods.

Extending PCA to Cross-view PCA (CvPCA): To deal with the cross-view problem, PCA [13, 14] can be extended to the following cross-view edition:

$$[\omega_q^*, \omega_r^*] = \arg\min \sum_{i,j}^{N_q, N_r} Z_{ij}^{(qr)}, \quad \text{s.t.} \quad \omega_q^T \omega_q = 1, \omega_r^T \omega_r = 1. \tag{5}$$

It is clearly that the cross-view PCA (CvPCA) is equal to PLS [4–6]. Therefore, we can conclude that PLS is the cross-view edition of the original PCA.

Table 2. Cross-view extensions of typical single-view algorithms in Graph Embedding framework.

Method	W_I	$W_I^{(q)}/W_I^{(r)}$	W_P	$W_P^{(q)}/B_q/W_P^{(r)}/B_r$
CvPCA	$W_I = I$	0	$W_P = 0$	$W_P^{(q)} = W_P^{(q)} = I$
CvLDA	$W_{Iij} = 1, C_i^{(q)} = C_j^{(r)}$	0	$W_{Pij} = 1, C_i^{(q)} \neq C_j^{(r)}$	$W_P^{(q)} = W_P^{(q)} = 0$
CvLPP	$W_I = exp(-\frac{\Vert y_i^{(q)} - y_j^{(r)} \Vert}{t})$	0	$W_P = 0$	$B_q = D_q, B_r = D_r$

Extending LDA to Cross-view LDA (CvLDA): Similarly, to deal with the cross-view problem, LDA [15, 16] can be extended to the following cross-view edition:

$$[\omega_q^*, \omega_r^*] = \arg\min \sum_{i,j}^{N_q, N_r} Z_{ij}^{(qr)} W_{Iij}, \quad \text{s.t.} \quad \sum_{i,j}^{N_q, N_r} Z_{ij}^{(qr)} W_{Pij} = 1, \tag{6}$$

where

$$W_{Iij} = \begin{cases} 1, & \text{if } C_i^{(q)} = C_j^{(r)}, \\ 0, & \text{if } C_i^{(q)} \neq C_j^{(r)}, \end{cases}, W_{Pij} = \begin{cases} 1, & \text{if } C_i^{(q)} \neq C_j^{(r)}, \\ 0, & \text{if } C_i^{(q)} = C_j^{(r)}. \end{cases} \tag{7}$$

The similarity matrices are detailed in Table 2 under our proposed framework. While those unsupervised approaches like CCA and PLS use part of cross-view information in order to contract pair-wise different view samples only, the Cross-view LDA (CvLDA) contract cross-view samples in the same class and separate samples in different classes.

Extending LPP to Cross-view LPP (CvLPP): Again, to deal with the cross-view sources analysis, LPP [17] could be extended to the following cross-view edition:

$$\begin{aligned} [\omega_q^*, \omega_r^*] = \arg\min \sum_{i,j}^{N_q, N_r} Z_{ij}^{(qr)} W_{Iij}, \\ \text{s.t.} \quad \omega_q^T x_i^{(q)} D_q x_j^{(q)^T} \omega_q = 1, \omega_r^T x_i^{(r)} D_r x_j^{(r)^T} \omega_r = 1. \end{aligned} \tag{8}$$

where $D_{qii} = \sum_{i \neq j} W_{ij}^{(q)}$, $D_{rii} = \sum_{i \neq j} W_{ij}^{(r)}$. Since the dimensions of the query samples and the reference samples in cross-view scenario might be different, the similarity matrix W_I cannot be calculated directly as original LPP. Therefore, we uses CvPCA features of the two cross-view sample sets to calculate the similarity matrix W_I with the definition of original LPP. As shown in Table 2, $y_i^{(q)}$ and $y_j^{(r)}$ represent the CvPCA features of samples $x_i^{(q)}$ and $x_j^{(r)}$.

Table 3. The definitions of similarity matrices in our method CLODA.

W_I/W_P	$W_I^{(q)}/W_I^{(r)}$	$W_P^{(q)}/B_q/W_P^{(r)}/B_r$
$W_{Iij} = exp(-\frac{\Vert y_i^{(q)}-y_j^{(r)}\Vert}{\sigma^2})$	${W_I}_{ij}^{(q)} = exp(-\frac{\Vert x_i^{(q)}-x_j^{(q)}\Vert}{\sigma_q^2})$	${W_P}_{ij}^{(q)} = exp(-\frac{\Vert x_i^{(q)}-x_j^{(q)}\Vert}{\sigma_q^2})$
$W_{Pij} = exp(-\frac{\Vert y_i^{(q)}-y_j^{(r)}\Vert}{\sigma^2})$	${W_I}_{ij}^{(r)} = exp(-\frac{\Vert x_i^{(r)}-x_j^{(r)}\Vert}{\sigma_r^2})$	${W_P}_{ij}^{(r)} = exp(-\frac{\Vert x_i^{(r)}-x_j^{(r)}\Vert}{\sigma_r^2})$

3 Cross-view Local Discriminant Analysis : A New Method Developed from Cross-view Graph Embedding Framework

Besides reformulation and extending, as a general framework, our Cross-view Graph Embedding framework can facilitate the development of new cross-view data analysis methods. In this section, as an example, we present a new approach derived from our framework named Cross-view Local Discriminant Analysis (CLODA).

On one side, CLODA considers the intra-class compactness and inter-class separability in both inter-view and intra-views. On the other side, as CDFE [9], CLODA also introduces the notion *local consistency* into the formulation to regularize the empirical objective and reduce the risk of overfitting. Different from CDFE, we not only use the locality of intra-views but also utilize the locality of inter-view. Formally, we formulate the learning objective to minimize the following objective function:

$$
\begin{aligned}
&[\omega_q^*, \omega_r^*] = \arg\min \left\{ \sum_{i,j}^{N_q,N_r} Z_{ij}^{(qr)} W_{Iij} + \beta_q \sum_{i\neq j}^{N_q,N_q} Z_{ij}^{(qq)} {W_I}_{ij}^{(q)} + \beta_r \sum_{i\neq j}^{N_r,N_r} Z_{ij}^{(rr)} {W_I}_{ij}^{(r)} \right\}, \\
&\text{s.t.} \quad \sum_{i,j}^{N_q,N_r} Z_{ij}^{(qr)} W_{Pij} + \beta_q' \sum_{i\neq j}^{N_q,N_q} Z_{ij}^{(qq)} {W_P}_{ij}^{(q)} + \beta_r' \sum_{i\neq j}^{N_r,N_r} Z_{ij}^{(rr)} {W_P}_{ij}^{(r)} = d.
\end{aligned}
\tag{9}
$$

where the definitions of all similarity matrices are presented in the Table 3. Specifically, the intrinsic graph similarity matrices' entry $W_{I_{ij}}$, ${W_I}_{ij}^{(q)}$ and ${W_I}_{ij}^{(r)}$ are equal to the corresponding value in the table when x_i and x_j belong to the same class and one element of them is among the K nearest neighbors of the other element; 0, otherwise. Likewise, the penalty graph similarity matrices' entry $W_{P_{ij}}$, ${W_P}_{ij}^{(q)}$ and ${W_P}_{ij}^{(r)}$ are equal to the corresponding value in the table when x_i and x_j belong to different classes and one element of them is among the K nearest neighbors of the other element; 0, otherwise. In addition, as CvLPP, we uses CvPCA features of the two cross-view sample sets to calculate the similarity matrices W_I and W_P.

4 Experiments

In this section, we evaluate the proposed framework on cross-view face recognition problems at two large-scale face databases: Multi-PIE [18] and COX-MR multi-resolution face databases. The former is a publicly released dataset, while the latter is collected by ourselves. We will release it with the publication of this work. In the two experiments, we compare our approaches with state-of-the-art cross-view methods and (linear or non-linear) single-view methods for cross-view recognition. Specifically, the parameters of involved methods are setting as follows: *For LPP* [17], we use KNN as the Neighbor Mode, the number of neighborhood $k = 5$, use HeatKernel as the Weight Mode, let $t = 5$; *For SDA* method [19], we manually select two subclasses (samples from the same view construct a subclass) for each class of samples; *For CDFE* [9], we select the best performance by tuning α from 0.1 to 2 and β from 0.1 to 1; *For CSR* [10], we select the best performance by tuning λ from 1 to 21 and η from 1 to 21. *For our Cross-view LPP (CvLPP)* we use 40 dimensions of CvPCA features to calculate the similarity matrix W_I, set the other parameters as original LPP; *For our method CLODA*, we use 40 dimensions of CvPCA features to calculate the similarity matrix W_I and W_P, let all the kernel widths σ,σ_q and σ_r be the mean of Euclidean distances of samples, set $\beta_q = \beta_r = \beta_q^{'} = \beta_r^{'} = \beta$ then select the best performance by tuning β from 0.1 to 1 and number of neighboring points K from 10 to 1000. For all methods, the dimension of the feature space is firstly selected by keeping 95% of the PCA energy.

4.1 Evaluation on the Multi-PIE Database

Multi-PIE [18] collects images of totally 337 subjects in multiple recording sessions and with large variations in pose, illumination, and expression. In this experiment, we select 7 distinct poses (-45^o, -30^o, -15^o, -15^o, 30^o, 45^o, and frontal angel 0^o) with automatic pose estimation (around 5% error). Therefore, these images with different poses form 7 types of views.

For our recognition experiment, we use 137 subjects (Subject ID from 201 to 346) with neutral expression from all 4 sessions at 7 different poses. Among them, 200 subjects (Subject ID from 001 to 200) are used for training and the rest are used for testing. Specifically, images in two different poses are used to learn pairs of projection directions. The frontal face image from the earliest session for each subject is used as the gallery image (137 totally), and all of the remaining images per subject are used as 7 probe sets from -45^o to 45^o, respectively. For each image, we fix the eye centers and crop it to the size of 64×80 pixels.

Table 4 shows the results of all methods for face recognition across poses which is a canonical cross-view problem. The comparisons show that our cross-view edition algorithms such as CvPCA(PLS), CvLDA and CvLPP all work better than corresponding original algorithms (PCA, LDA and LPP) under Graph Embedding. In the cross-view recognition, most existing cross-view methods prevail on the linear or non-linear single-view approaches. It also shows that our new method CLODA consistently outperforms any other methods significantly.

Table 4. Face recognition performance on Multi-PIE Database. For each probe set, the gallery set is the same set which contains the frontal face images.

Method	probe set(rank-1 recognition rate)						
(*gray feature*)	-45^o	-30^o	-15^o	0^o	-15^o	30^o	45^o
PCA	14.18	19.72	39.87	87.67	57.59	19.32	15.85
PCA+PLS(CvPCA)	21.63	72.89	83.72	87.33	81.72	74.24	19.01
PCA+LDA	11.70	53.52	88.70	95.33	85.52	52.54	8.10
PCA+CvLDA	26.95	80.99	93.02	94.33	90.34	80.00	24.30
PCA+LPP	9.22	14.79	30.90	76.33	38.28	16.95	4.26
PCA+CvLPP	11.70	77.25	85.38	92.67	82.41	66.44	11.21
PCA+SDA	13.12	33.80	80.07	93.67	80.69	42.71	9.51
PCA+LE[20]	7.09	7.04	11.63	72.00	23.45	5.42	6.69
PCA+LLE[21]	10.99	12.68	33.55	89.0	60.69	12.32	9.49
PCA+CCA	19.15	78.17	89.04	91.00	89.31	78.31	18.66
PCA+DCCA	21.27	70.42	88.04	93.33	84.48	70.17	18.31
PCA+CDFE	14.89	77.82	91.36	93.33	88.28	72.20	15.49
PCA+CSR	23.40	76.76	89.37	93.33	87.93	76.61	23.24
PCA+CLODA	**29.08**	**83.80**	**94.35**	**95.00**	**92.76**	**82.03**	**29.23**

4.2 Evaluation on the COX-MR Database

The COX-MR database collected by ourselves consists of images from 5 resolutions as shown in Table 5. Since the resolutions are different, we normalize the images with different sizes as detailed in Table 5. On this dataset, we use the images of 300 persons for training, and the images of 500 persons for testing. Furthermore, we subdivide the training set and the test set into 5 categories according to 5 different views (resolutions). Specifically, recognition cross 2 different resolutions is a cross-view problem. Table 5 details the training sets and test sets of our database. Note that COX-MR also contains images under different illumination conditions and facial expressions.

On COX-MR database, we design 10 face verification testings specifically for the cross-resolution scenario. As shown in Table 6, for example, index *SHH* represents the test using *Super High* Resolution testing set as the target set and *High* Resolution testing set as the query set. It is worth mentioning that *the images in the High Resolution set contain kinds of variations in poses while the images in other resolution sets are relatively controlled.*

Since the dimensionalities of images with different resolutions are different, most single-view methods and the cross-view method CSR [10] can not work directly in the cross-resolution recognition. Therefore, Table 6 only shows comparisons of cross-view methods which could deal with the cross-view problem. The performance is measured by the verification rate when false accept rate is 0.001. Seen from the above extensive comparisons, our cross-view editions of PCA, LDA, LPP all work well and our new method CLODA even outperforms the conventional approaches consistently in all evaluations.

Table 5. Details of COX-MR database. The original size of face is measured in pixels between the centers of the eyes; cropped size is meausred in pixels of width and height.

Resolutions	Eye-distance	Cropped-size	Training	Test	Total-num.
Super High *(SH)*	≈350	160×200	2030	3348	5378
High *(H)*	≈100	128×160	2444	3128	5572
Median *(M)*	70∼80	80×100	2719	4287	7006
Low *(L)*	40∼45	64×80	2943	4875	7818
Super Low *(SL)*	18∼20	40×50	2844	4848	7692

Table 6. Face verification performance of COX-MR Database

Method	Protocols(VR (%) at FAR = 0.001)									
(*gray feature*)	*SHH*	*SHM*	*SHL*	*SHSL*	*HM*	*HL*	*HSL*	*ML*	*MSL*	*LSL*
PCA+CvPCA	5.77	8.33	8.89	9.02	57.27	32.08	12.41	51.71	15.92	21.97
PCA+CvLDA	11.52	18.08	19.14	13.60	84.17	70.35	38.06	86.23	47.51	62.49
PCA+CvLPP	6.10	5.93	6.32	4.55	81.86	61.09	25.17	88.16	43.48	63.31
PCA+CCA	10.12	15.93	17.08	13.50	86.23	77.15	45.84	90.88	60.18	71.88
PCA+DCCA	7.80	14.46	14.94	12.10	79.46	67.73	38.77	84.08	54.84	62.52
PCA+CDFE	11.58	16.12	19.23	11.63	86.35	75.99	42.83	90.91	59.95	71.54
PCA+CLODA	**17.56**	**23.30**	**24.17**	**17.58**	**88.78**	**79.52**	**50.02**	**92.34**	**62.74**	**72.87**

4.3 Discussion

In this paper, our proposed Cross-view Graph Embedding framework attempts to deal with the general cross-view problem, where the two views are not necessarily from the same feature space. Although some non-linear manifold learning techniques (e.g., LPP, LE, LLE etc.) can handle the cross-view problem when the two views are from the same feature space, e.g. the cross-pose problem on MultiPIE, they fail to work in the scenarios that the two views are from the different spaces, e.g., cross-resolution problem on COX-MR. Furthermore, in most cases, the two views may not lie in a continuous manifold, e.g., frontal faces and half-profile face images. In these cases, those manifold learning techniques may not work well. Additionally, the existing cross-view methods only attempted to preserve the inter-view discriminant information or the intra-view local structure; besides these two types of information, our method CLODA also preserves the intra-view discriminant and inter-view local structure information, leading to a better generalization.

5 Conclusion

In this paper, we proposed a unified framework called Cross-view Graph Embedding extended from Graph Embedding. We show that, our framework not only can elegantly reformulate most existing cross-view methods but also have the ability to rebuild the typical intra-view algorithms in original Graph Embedding into corresponding cross-view editions. It also facilitates the development of new cross-view feature extraction methods. As an example, we design a novel Cross-view Local Discriminant Analysis (CLODA) method. Extensive experiments show that our methods achieve significant improvement over the traditional methods in the cross-view problems.

Acknowledgements. This work is partially supported by National Basic Research Program of China (973 Program) under contract 2009CB320902; Natural Science Foundation of China (NSFC) under contract Nos. 61025010 and U0835005; and Beijing Natural Science Foundation under contract No. 4111003. Haihong Zhang and Shihong Lao are partially supported by "R&D Program for Implementation of Anti-Crime and Anti-Terrorism Technologies for a Safe and Secure Society", Special Coordination Fund for Promoting Science and Technology of MEXT, the Japanese Government.

6 Appendix

The section presents the general solution of proposed Cross-view Graph Embedding in details. To simplify the further deductions, we introduce the following diagonal matrices notations: $S_I^{(u)}(i,i) = \sum_{j=1}^{N_u} W_{Iij}$, $D_I^{(u)}(i,i) = \sum_{j=1}^{N_u} W_{Iij}^{(u)}$, $S_P^{(u)}(i,i) = \sum_{j=1}^{N_u} W_{Pij}$, $D_P^{(u)}(i,i) = \sum_{j=1}^{N_u} W_{Pij}^{(u)}$, where u denotes the query view(q) or reference view (r). Following Eq.(3) and above definitions, then we rewrite the objective function of Eq.(2) in matrix form as:

$$\begin{aligned}
&\sum_{i,j}^{N_q,N_r} Z_{ij}^{(qr)} W_{Iij} + \beta_q \sum_{i\neq j}^{N_q,N_q} Z_{ij}^{(qq)} W_{Iij}^{(q)} + \beta_r \sum_{i\neq j}^{N_r,N_r} Z_{ij}^{(rr)} W_{Iij}^{(r)} \\
=& tr(\Omega_q^T X_q S_I^{(q)} X_q^T \Omega_q + \Omega_r^T X_r S_I^{(r)} X_r^T \Omega_r - 2\Omega_q^T X_q W_I X_r^T \Omega_r) \\
&+ 2\beta_q tr(\Omega_q^T X_q (D_I^{(q)} - W_I^{(q)}) X_q^T \Omega_q) + 2\beta_r tr(\Omega_r^T X_q (D_I^{(r)} - W_I^{(r)}) X_r^T \Omega_r) \\
=& tr(\Omega_q^T X_q R_I^{(q)} X_q^T \Omega_q + \Omega_r^T X_r R_I^{(r)} X_r^T \Omega_r - 2\Omega_q^T X_q W_I X_r^T \Omega_r)
\end{aligned} \tag{10}$$

where $R_I^{(q)} = S_I^{(q)} + 2\beta_q(D_I^{(q)} - W_I^{(q)})$, $R_I^{(r)} = S_I^{(r)} + 2\beta_r(D_I^{(r)} - W_I^{(r)})$. Likewise, we also simplify the subjective condition of Eq.(2) as following:

$$\begin{aligned}
&\sum_{i,j}^{N_q,N_r} Z_{ij}^{(qr)} W_{Pij} + \beta_q' \sum_{i\neq j}^{N_q,N_q} Z_{ij}^{(qq)} W_{Pij}^{(q)} + \beta_r' \sum_{i\neq j}^{N_r,N_r} Z_{ij}^{(rr)} W_{Pij}^{(r)} \\
&= tr(\Omega_q^T X_q R_P^{(q)} X_q^T \Omega_q + \Omega_r^T X_r R_P^{(r)} X_r^T \Omega_r - 2\Omega_q^T X_q W_P X_r^T \Omega_r)
\end{aligned} \tag{11}$$

where $R_P^{(q)} = S_P^{(q)} + 2\beta_q'(D_P^{(q)} - W_P^{(q)})$, $R_P^{(r)} = S_P^{(r)} + 2\beta_r'(D_P^{(r)} - W_P^{(r)})$. To solve the optimization problem, we introduce the matrices:

$$M_I = \begin{bmatrix} X_q R_I^{(q)} X_q^T & -X_q W_I X_r^T \\ -X_r W_I^T X_q^T & X_r R_I^{(r)} X_r^T \end{bmatrix}, M_P = \begin{bmatrix} X_q R_P^{(q)} X_q^T & -X_q W_P X_r^T \\ -X_r W_P^T X_q^T & X_r R_P^{(r)} X_r^T \end{bmatrix}, \Omega = \begin{bmatrix} \Omega_q \\ \Omega_r \end{bmatrix} \tag{12}$$

Finally, following Eq.(10-12), we transform the optimization problem Eq.(2) to the generalized eigen-decomposition problem $M_I \Omega = \lambda M_P \Omega$.

References

1. Hotelling, H.: Relations between two sets of variates. Biometrika **28** (1936) 321–377
2. Hardoon, D., Szedmak, S., Shawe-Taylor, J.: Canonical correlation analysis: An overview with application to learning methods. Neural Computation **16** (2004) 2639–2664
3. Yi, D., Liu, R., Chu, R., Lei, Z., Li, S.: Face matching between near infrared and visible light images. Advances in Biometrics (2007) 523–530
4. Kotz, S., Read, C., Banks, D.: Encyclopedia of Statistical Sciences. John Wiley (1999)
5. Rosipal, R., Krämer, N.: Overview and recent advances in partial least squares. Subspace, Latent Structure and Feature Selection (2006) 34–51
6. Sharma, A., Jacobs, D.: Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch. In: CVPR. (2011) 593–600
7. Li, B., Chang, H., Shan, S., Chen, X.: Low-resolution face recognition via coupled locality preserving mappings. Signal Processing Letters **17** (2010) 20–23
8. Sun, T., Chen, S., Yang, J., Hu, X., Shi, P.: Discriminative canonical correlation analysis with missing samples. In: CSIE. Volume 6. (2009) 95–99
9. Lin, D., Tang, X.: Inter-modality face recognition. In: ECCV. (2006) 13–26
10. Lei, Z., Li, S.: Coupled spectral regression for matching heterogeneous faces. In: CVPR. (2009) 1123–1128
11. Cai, D., He, X., Han, J.: Spectral regression for efficient regularized subspace learning. In: ICCV. (2007)
12. Yan, S., Xu, D., Zhang, B., Zhang, H., Yang, Q., Lin, S.: Graph embedding and extensions: A general framework for dimensionality reduction. IEEE T-PAMI **29** (2007) 40–51
13. Turk, M., Pentland, A.: Face recognition using eigenfaces. In: CVPR. (1991) 586–591
14. Jolliffe, I., MyiLibrary: Principal component analysis. Volume 2. Wiley Online Library (2002)
15. Fisher, R.: The use of multiple measurements in taxonomic problems. Annals of Human Genetics **7** (1936) 179–188
16. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE T-PAMI **19** (1997) 711–720
17. He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H.: Face recognition using laplacianfaces. IEEE T-PAMI **27** (2005) 328–340
18. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-PIE. Image and Vision Computing **28** (2010) 807–813
19. Zhu, M., Martinez, A.: Subclass discriminant analysis. IEEE T-PAMI **28** (2006) 1274–1286
20. Belkin, M., Niyogi, P.: Laplacian eigenmaps and spectral techniques for embedding and clustering. In: Advances in neural information processing systems 14. (2001) 585–591
21. Roweis, S., Saul, L.: Nonlinear dimensionality reduction by locally linear embedding. Science **290** (2000) 2323–2326