Pairwise Linear Regression Classification for Image Set Retrieval

Qingxiang Feng, Yicong Zhou*, Rushi Lan
Department of Computer and Information Science, University of Macau
Avenida da Universidade, Taipa, Macau, China
fengqx1988@gmail.com, yicongzhou@umac.mo, rslan2012@gmail.com

Abstract

This paper proposes the pairwise linear regression classification (PLRC) for image set retrieval. In PLRC, we first define a new concept of the unrelated subspace and introduce two strategies to constitute the unrelated subspace. In order to increase the information of maximizing the query set and the unrelated image set, we introduce a combination metric for two new classifiers based on two constitution strategies of the unrelated subspace. Extensive experiments on six well-known databases prove that the performance of PLRC is better than that of DLRC and several state-of-the-art classifiers for different vision recognition tasks: cluster-based face recognition, video-based face recognition, object recognition and action recognition.

1. Introduction

In pattern recognition and computer vision fields, images classification tasks (e.g., face recognition, object recognition and action recognition) attract a lot of researchers' attention. Good performance is known to be highly reply on classifiers. A number of classifiers were proposed, such as the nearest neighbor (NN) [4], SVM classifier [15], sparse representation-based classifications (SRC) [20] and linear regression classification (LRC) [13]. These classifiers use a single test sample for classification. Their classification performance, however, is generally dependent on the base or representation of individual test samples. Recently, researchers paid more attention to the image-set-based classification that is a generalization of video-based classification. They both use multiple test samples. In [8], the image-set based face recognition is considered as the same category as video-based face recognition. Several image set-based methods use only benchmarks of video databases for their evaluation [1, 17, 19, 21]. However, the video-based classification is not suitable for some applications. For example, the real-time recognition systems may have difficulty to always obtain videos with a proper length and sufficient detectable images. Therefore, the limited-sample-query-set based classification becomes significative.

For similar classification tasks, dual linear regression classification (DLRC) [3] was proposed as a non-parametric approach. It borrows the idea of LRC [13] and extends LRC from the single-query-sample based method to the image-set-based method. DLRC has a demonstrated better performance than a few well-known methods. However, DLRC considers only the related class-subspace for classification. That is to say, it pays attention only to minimizing the distance between the query set and the related train set.

In the paper, we proposed the pairwise linear regression classification (PLRC) for image set retrieval. As an improved version of DLRC, PLRC introduces the new unrelated subspace to maximize the distance between the query set and the unrelated images set, and utilizes a new combined metric that integrates a related metric and a new unrelated metric for classification. The effectiveness of the proposed PLRC is assessed on six popular databases. They include LFW face database [25], AR face database [12] for cluster-based face recognition, Honda/UCSD database [11], CMU Mobo database [6] for video-based face recognition, the Caltech101 object database [5] for object recognition, UCF50 action database [14] for action recognition.

The main contributions of the paper are as follows:

- We propose the unrelated subspace concept to improve the traditional of DLRC method.
- Using this concept, we introduce two strategies to constitute the unrelated subspace. The optimization problem in (18) of the first strategy is based on the image-set, which is different from the previous optimization problem based on the single test sample in SRC [20] and CRC [24].
- We propose the new unrelated metric, related metric and the combination metric for classification.
- Based on new metrics and two different methods of

*Corresponding author.

constituting the unrelated subspace, we further propose two classifiers called PLRC-I and PLRC-II for image recognition.

2. Dual Linear Recognition Classification

This section briefly reviews the dual linear recognition classification (DLRC) algorithm. Given the training image-set of the c^{th} class as:

$$X_c = [x_1^c \; x_2^c \; \cdots \; x_{N_c}^c] \in R^{q \times N_c}, \tag{1}$$

and the test image-set as:

$$Y = [y_1 \; y_2 \; \cdots \; y_n] \in R^{q \times n}. \tag{2}$$

DLRC tries to find the joint coefficients for restructuring the sample from two image-sets. First, the two image-set will be disposed as:

$$\hat{X}_c = [\hat{x}_1^c \; \hat{x}_2^c \; \cdots \; \hat{x}_{N_c}^c] \in R^{q \times N_c - 1}, \tag{3}$$

and

$$\hat{Z} = [\hat{z}_1 \; \hat{z}_2 \; \cdots \; \hat{z}_n] \in R^{q \times n-1}, \tag{4}$$

where $\hat{x}_i^c = x_i^c - x_{N_c}^c$, $i = 1, 2, \cdots, N_c - 1$, $\hat{z}_i = y_i - y_n$, $i = 1, 2, \cdots, n-1$. In order to obtain the joint coefficient vector of the two image set, the joint image set S^c and test vector s^c can be constituted as:

$$S^c = [\hat{X}_c \; -\hat{Z}] \in R^{q \times (N_c + n - 2)}, \tag{5}$$

and

$$s^c = y_n - x_{N_c}^c. \tag{6}$$

Suppose that $\beta^c \in R^{(N_c+n-2)\times 1}$ is the joint coefficient vector of $\hat{X}_c$ and $\hat{Z}$, and can be calculated by solving the following equation:

$$s^c = S^c \beta^c. \tag{7}$$

Namely, $\beta^c \in R^{(N_c+n-2)\times 1}$ can be solved as:

$$\beta^c = (S^{cT} S^c)^{-1} S^{cT} s^c. \tag{8}$$

According to DLRC, the two reconstructed images r_1 and r_2 from subspaces $\hat{X}_c$ and $\hat{Z}$ can be described as:

$$\begin{aligned} r_1 &= \hat{X}_c [\beta_1^c \; \cdots \; \beta_{N_c-1}^c]^T + x_{N_c}^c, \\ r_2 &= \hat{Z} [\beta_{N_c}^c \; \cdots \; \beta_{N_c+n-2}^c]^T + y_n. \end{aligned} \tag{9}$$

Therefore, the distance measure between the test image set and related training set can be calculated as:

$$d_r^c = \|r_1 - r_2\| = \|s^c - S^c \beta^c\|. \tag{10}$$

3. Pairwise Linear Regression Classification

This section proposes pairwise linear regression classification (PLRC). Its flowchart is shown in Figure 1. The main contents of this section are as follows. First, two strategies of constituting the unrelated subspace are described in Subsection 3.1. Then, the related metric and unrelated metrics are computed in Subsections 3.2 and Subsection 3.3, respectively. Next, the final distance metric for classification, called combination metric, is described in Subsection 3.4. Last, Subsection 3.5 describes the relationship of the DLRC and the proposed PLRC.

3.1. Construction of the Unrelated Subspace

Before constituting the unrelated subspace, the definition of unrelated subspace is described in the Definition 1.

Definition 1: *Suppose that there are M classes. There exists a specific test set with N_c samples belongs to the c^{th} class. If another set U also contains N_c samples from the other $M-1$ classes except the c^{th} class, this set U will be called unrelated subspace.*

According to Definition 1, we need to select c^{th} samples from the other classes except the c^{th} class to constitute the unrelated subspace. In this subsection, two strategies are used to constitute the unrelated subspace, which are described as follows.

3.1.1 Strategy 1

Obtain the training image-set X_c of the c^{th} class by (1) and the test image-set Y by (2). Compute the mean sample of Y as

$$y_{mean} = \frac{1}{n} \sum_{i=1}^{n} y_i. \tag{11}$$

Increase the mean sample as the last element of the original image-set Y and the test image-set will be disposed as

$$\hat{Y} = [\hat{y}_1 \; \hat{y}_2 \; \cdots \; \hat{y_n}] \in R^{q \times n}, \tag{12}$$

where $\hat{y}_i = y_i - y_{mean}$, $i = 1, 2, \cdots, n$. Suppose that we have M classes of subjects, Collect the entire class-specific set X_c defined in (1) to form the complete data model as

$$X = [X_1 \; X_2 \; \cdots \; X_M] \in R^{q \times \sum_{c=1}^{M} N_c}. \tag{13}$$

Compute the mean sample of X as

$$x_{mean} = \frac{1}{MN_c} \sum_{c=1}^{M} \sum_{i=1}^{N_c} x_i^c. \tag{14}$$

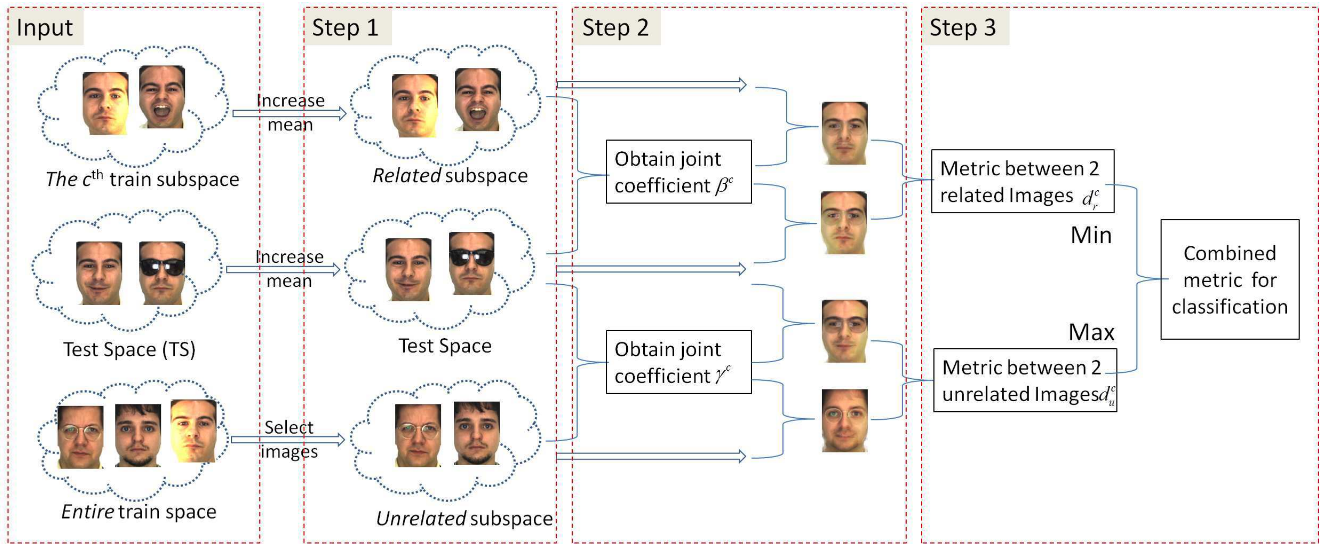

Figure 1. The flowchart of the proposed PLRC. In Step 1, the related subspace and test space of the c^{th} class are reconstituted by increasing the mean sample as the last element of the original sets. The unrelated subspace is chosen from the entire training space with the specific strategy. In Step 2, the joint coefficient of the test space and related subspace is obtained. Next the two spaces will restructure two images with the coefficient. The construction procedure of test space and unrelated subspace is similar. In Step 3, the related metric and the unrelated metric are computed. It is easy to know that the smaller related metric and the larger unrelated metric are conducive to the classification. Next, the combined metric with the related metric and unrelated metric is obtained for classification.

Increase the mean sample as the last element of the original image-set X. The image-set X will be disposed as

$$\hat{X} = [\hat{x}_1^1 \ \hat{x}_2^1 \ \cdots \ \hat{x}_{N_M}^M] \in R^{q \times L}, \tag{15}$$

where $L = \sum_{c=1}^{M} N_c$, $\hat{x}_i^c = x_i^c - x_{mean}^c$, $i = 1, 2, \cdots, N_c$, $c = 1, 2, \cdots, M$. In order to obtain the joint coefficient vector of the two image sets, the joint image set E and test vector e can be constituted as

$$E = [\hat{X} \ \ -\hat{Y}] \in R^{q \times (L+n)} \tag{16}$$

and

$$e = y_{mean} - x_{mean}. \tag{17}$$

Suppose that $\alpha \in R^{(L+n)\times 1}$ is the joint coefficient vector of $\hat{X}$ and $\hat{Y}$, which can be calculated by solving the optimization problem as:

$$\hat{\alpha} = \arg\min_{\alpha} \|e - E\alpha\|^2 + \lambda\|\Gamma\alpha\|^2, \tag{18}$$

where λ is a parameter, Γ is the Tikhonov matrix, which can be computed as:

$$\Gamma = \begin{pmatrix} \|E_1 - e\| & 0 & 0 & 0 \\ 0 & \|E_2 - e\| & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \|E_{L+n} - e\| \end{pmatrix}. \tag{19}$$

The weighted distance between e and E can be computed as:

$$d_i = \|E_i \hat{\alpha}_i - e\|, i = 1, 2, \cdots, (L+n). \tag{20}$$

Because the last n elements of E is the test set, and the N_c elements of the c^{th} class of E are the related set, the $(n + N_c)$ samples will be removed from E as $\hat{E} \in R^{q\times(L-N_c)}$. Similar, suppose $D = [d_1 \ d_2 \ \cdots \ d_{L+n}] \in R^{1\times(L+n)}$, remove $n + N_c$ distances from D as $\hat{D} \in R^{1\times(L-N_c)}$. Then select N_c samples from $\hat{E}$, which is corresponding to the N_c minimum distances of $\hat{D}$, to constitute a subspace. The subspace will be treated as the unrelated class subspace of the c^{th} class and be described as:

$$U_c = [u_1^c \ u_2^c \ \cdots \ u_{N_c}^c] \in R^{q\times N_c}. \tag{21}$$

where $u_i = \hat{E}_j$, $i = 1, 2...N_c$, $j \in \Omega$ and Ω denotes the label set of the N_c distances in $\hat{D}$. The classifier based on this strategy will be called pairwise linear regression-I (PLRC-I). **Noted:** the optimization problem in (18) is based on the image-set, which is different from the previous optimization problem based on the single test sample in SRC [20] and CRC [24].

3.1.2 Strategy 2

The Euclid distance between the y_{mean} and a train sample X_i can be computed as:

$$d_i = \|X_i - y_{mean}\|, i = 1, 2, \cdots, L. \tag{22}$$

where the X_i denotes the i^{th} elements of the entire data model X. Suppose $D = [d_1 \ d_2 \ \cdots \ d_L] \in R^{1\times L}$, remove the elements corresponding to the c^{th} class from D as

$\hat{D} \in R^{1\times(L-N_c)}$. Then select N_c samples from X, which is corresponding to the label of N_c chosen distances from $\hat{D}$, to constitute a subspace. The subspace will be treated as the unrelated class subspace of the c^{th} class and be described as

$$U_c = [u_1^c \; u_2^c \; \cdots \; u_{N_c}^c] \in R^{q\times N_c}. \tag{23}$$

where $u_i = \hat{X}_j,\ i = 1, 2...N_c,\ j \in \Omega$ and Ω denotes the label set of the N_c distances.The classifier based on this strategy will be called pairwise linear regression-II (PLRC-II).

3.2. Related Distance Metric

Obtain the training image-set X_c of the c^{th} class by (1) and the test image-set Y by (2). Compute the mean sample of X_c as:

$$x_{mean}^c = \frac{1}{N_c}\sum_{i=1}^{N_c} x_i^c \tag{24}$$

The image set X will be disposed as:

$$\hat{X}_c = [\hat{x}_1^c \; \hat{x}_2^c \; \cdots \; \hat{x}_{N_c}^c] \in R^{q\times N_c}, \tag{25}$$

where $\hat{x}_i^c = x_i^c - x_{mean}^c, i = 1, 2, \cdots, N_c$. The y_{mean} can be computed by (11). The image set Y can be disposed by (12) as $\hat{Y}$. In order to obtain the joint coefficient vector of the two image set, the joint image set and test vector can be constituted as

$$S_r^c = [\hat{X}_c \quad -\hat{Y}] \in R^{q\times(N_c+n)} \tag{26}$$

and

$$s_r^c = y_{mean} - x_{mean}^c. \tag{27}$$

Suppose that $\beta^c \in R^{(N_c+n)\times 1}$ is the joint coefficient vector of $\hat{X}_c$ and $\hat{Y}$, which can be calculated by solving the following equation:

$$s_r^c = S_r^c \beta_c. \tag{28}$$

and $\beta^c \in R^{(N_c+n)\times 1}$ can be solved as:

$$\beta^c = (S_r^{cT} S_r^c)^{-1} S_r^{cT} s_r^c. \tag{29}$$

Similar to DLRC, the two reconstructed images r_1 and r_2 from subspaces $\hat{X}_c$ and Y can be described as:

$$r_1 = \hat{X}_c[\beta_1^c \; \cdots \; \beta_{N_c}^c]^T + x_{mean}^c \tag{30}$$

and

$$r_2 = \hat{Y}[\beta_{N_c+1}^c \; \cdots \; \beta_{N_c+n}^c]^T + y_{mean}. \tag{31}$$

Use the r_1 and r_2, the distance measure between the test image set and related training set can be calculated as follows:

$$d_r^c = \|r_1 - r_2\| = \|s_r^c - S_r^c \beta^c\|. \tag{32}$$

3.3. Unrelated Distance Metric

The unrelated image set U_c of the c^{th} class can be obtained by (21) or (23). Compute the mean sample of U_c as:

$$u_{mean}^c = \frac{1}{N_c}\sum_{i=1}^{N_c} u_i^c. \tag{33}$$

Increase the mean sample as the last element of the original image-sets U, and it will be disposed as:

$$\hat{U}_c = [\hat{u}_1^c \; \hat{u}_2^c \; \cdots \; \hat{u}_{N_c}^c] \in R^{q\times N_c}, \tag{34}$$

where $\hat{u}_i^c = u_i^c - u_{mean}^c,\ i = 1, 2, \cdots, N_c$. In order to obtain the joint coefficient vector of the two image set U_c and Y, the joint image set and test vector can be constituted as:

$$S_u^c = [\hat{U}_c \quad -\hat{Y}] \in R^{q\times(N_c+n)}, \tag{35}$$

and

$$s_u^c = y_{mean} - u_{mean}^c. \tag{36}$$

Suppose that $\gamma \in R^{(N_c+n)\times 1}$ is the joint coefficient vector of $\hat{U}_c$ and $\hat{Y}$, which can be calculated by solving the following equation:

$$s_u^c = S_u^c \gamma_c, \tag{37}$$

where $\gamma \in R^{(N_c+n)\times 1}$ can be calculated by the least square error as:

$$\gamma = (S_u^{cT} S_u^c)^{-1} S_u^{cT} s_u^c. \tag{38}$$

Similar to DLRC, the two reconstructed images u_1 and u_2 from subspaces $\hat{U}_C$ and $\hat{Y}$ can be described as:

$$u_1 = \hat{U}_c[\gamma_1^c \; \cdots \; \gamma_{N_c}^c]^T + u_{mean}^c \tag{39}$$

and

$$u_2 = \hat{Y}[\gamma_{N_c+1}^c \; \cdots \; \gamma_{N_c+n}^c]^T + y_{mean}. \tag{40}$$

The unrelated distance metric between the test image set and unrelated training set are calculated as:

$$d_u^c = \|u_1 - u_2\| = \|s_u^c - S_u^c \gamma^c\|. \tag{41}$$

3.4. Combined Distance Metric

After obtaining the related metric and the unrelated metric, we need a method to combine these two metrics as a better metric. It is easy to know that the minimum distance of related metric represents the best match while the maximum distance of unrelated metric represents the best match. Therefore, we combine the related metric and the unrelated metric as the combination metric as:

$$d_p^c = d_r^c / d_u^c \tag{42}$$

PLRC selects the class with the minimum distance as:

$$\min_{c^*} d_p^c, c = 1, 2, \cdots, M. \tag{43}$$

Algorithm 1 Pairwise Linear Regression Classification (PLRC)

Inputs The entire training samples x_i^c, $c = 1, 2, \cdots, M$, $i = 1, 2, \cdots, N_c$ and a test image vector $x \in R^{q \times 1}$.

Output Index of x.

1. Two strategies are used to constitute the unrelated subspace. Their main difference is the distance metric between test set and the train set. They are described as:

   Strategy 1: Solve the optimization problem as:

   $$\hat{\alpha} = \arg\min_{\alpha} \ \|e - E\alpha\|^2 + \lambda\|\Gamma\alpha\|^2.$$

   Next, compute the sparse weighted distance as:

   $$d_i = \|E_i\hat{\alpha}_i - e\|, i = 1, 2, \cdots, (L + n).$$

   Strategy 2: Compute the Eculid distance as:

   $$d_i = \|X_i - y_{mean}\|, i = 1, 2, \cdots, L.$$

   Using the distances, we select N_c nearest samples from the other $M - 1$ classes except the c^{th} class to constitute the unrelated subspace as:

   $$U_c = [u_1^c \ u_2^c \ \cdots \ u_{N_c}^c] \in R^{q \times N_c}.$$

2. Use the related subspace and the test set to compute the related distance metric as:

   $$d_r^c = \|r_1 - r_2\| = \|x_r^c - S_r^c\beta^c\|.$$

3. Use the unrelated subspace and the test set to compute the unrelated distance metric as:

   $$d_u^c = \|u_1 - u_2\| = \|s_u^c - S_u^c\gamma^c\|.$$

4. Combine the related metric and the unrelated metric as the combined metric as:

   $$d_p^c = d_r^c / d_u^c$$

5. PLRC selects the class with the minimum distance as:

   $$\min_{c^*} d_p^c, c = 1, 2, \cdots, M.$$

3.5. PLRC vs DLRC

In this subsection, we compare the similarity and difference between PLRC and DLRC as follows.

Similarity: PLRC and DLRC both follow the idea of restructuring the virtual sample of two image set for classification, and use the metric of the test set and the related train set.

Difference: For restructuring the virtual sample, DLRC uses the last sample of the image set plus the variations while PLRC utilizes the mean sample of the image set plus the variations. In addition, PLRC not only includes the information that minimizes the distance between the test set and related train set, but also pays attention to the information that maximizes the distance between the test set and the unrelated train set. However, DLRC considers only the information that minimizes the distance between the test set and related train set.

4. Experimental Results

This section provides extensive experimental results to testify the performance of two proposed classifiers: PLRC-I and PLRC-II. These experiments are carried out using the following vision recognition tasks and databases: image-based face recognition on the LFW face database [25] and AR face database [12], video-based face recognition on Hona/UCSD face database [11] and CMU Mobo face database [6], action recognition on the UCF50 action databases [14], and object recognition on Caltech101 object databases [5].

4.1. Face recognition on image-based databases

This subsection tests the performance of the proposed PLRC for face recognition in wild on LFW face database and face recognition with occlusion on AR database.

4.1.1 Face recognition in wild

LFW face database were captured in unconstrained environments such that there will be large variations in face images including pose,age, race, facial expression, lighting, occlusions, and background, etc. We use the aligned version of the LFW database, LFW-a database, to study the performance of the proposed classifiers. Note that all the images in LFW-a database are a size of 250×250. Following the operations in [3], we manually crop images into a size of 90×78. An subset of LFW contains 62 persons, each people has more than 20 face images, is used for evaluating the algorithms. Our experimental setting is identical to that in [3]. That is, 10 images of each subject are selected to form the training set, while the last 10 image are used as the probe images. The proposed classifiers are compared with following methods: sparse approximated nearest points (SANP) [7, 8], affine hull based image set distance (ASIHD) [2], convex hull based image set distance (CSIHD) [2], manifold discriminant analysis (MDA) [16], SRC+NN [20], LRC+NN [13], and DLRC [3]. All methods use the downscaled images of size of 10×10 and 15×10 as in [3]. The classification results of all methods are illustrated in Table 1. For the images with size of 10×10, the proposed PLRC-II achieves identical performances with the MDA method. Another classifier, PLRC-I, obtains the most satisfactory recognition rate compared with other methods.

Classifier	10×10	15×10
SANP	85.48	92.55
ASIHD	87.10	95.16
CSIHD	90.32	93.55
MDA	93.55	95.16
SRC+NN	85.48	88.71
LRC+NN	77.42	83.87
DLRC	91.94	95.16
PLRC-I	95.16	96.77
PLRC-II	93.55	96.77

Table 1. The recognition rates (RR) of several classifiers on LFW database.

For images with size of 15×10, our proposed PLRC-I and PLRC-II classifiers both obtain the highest recognition rate of 96.77%.

4.1.2 Face recognition with occlusion

In this experiment, we study the performance of the proposed classifiers using the well-known AR database. There are over 4000 face images of 126 subjects (70 men and 56 women) in the database. We use the cropped AR database that includes 2600 face images of 100 individuals. The face images of each individual contain different expressions, lighting conditions, wearing sun glasses and wearing scarf. They were manually cropped into 40×40 pixels.

For this database, the proposed classifiers are compared with following state-of-the-art approaches: SANP [7, 8], ASIHD [2], CSIHD [2], LRC+NN [13], and DLRC [3], respectively. This experiment is run as follows: 5 samples per person are used as Gallery set, and the rest 21 samples are divided into 3 probe sets with the same size. The recognition rates of different classifiers have been presented in Table 2. The experimental results show that the PLRC-II obtains significant improvements compared with SANP, ASIHD, CSIHD, and LRC+NN methods. The result of DLRC is one percentage higher than that of PLRC-II. The recognition rate of PLRC-I surpasses both of DLRC and PLRC-II more than one percentage. From this experiment, we know that the unrelated subspace based on the sparse optimization is better than that based on the Eculid distance for face recognition with occlusion.

4.2. Face Recognition on Video-based Databases

This subsection tests the performance of PLRC for face recognition on video-based databases.

4.2.1 Honda/UCSD face database

There are totally 59 video clips of 20 subjects in the Honda/UCSD data [11]. Each subject has 2 videos at least. 20

Classifier	RR
SANP	77.00
ASIHD	87.67
CSIHD	84.67
LRC+NN	82.33
DLRC	96.00
PLRC-I	97.33
PLRC-II	95.00

Table 2. The recognition rate (RR) of several classifiers on AR database.

Classifier	RR
DCC	70.92
MMD	69.32
MDA	82.05
AHISD	87.18
CHISD	82.05
MSM	74.36
SANP	84.62
RNP	87.18
DLRC	87.18
PLRC-I	89.74
PLRC-II	87.18

Table 3. The recognition rate (RR) of several classifiers on Honda face database.

videos are named training videos and the rest 39 test videos. We conduct the experiment as the identical setting in [8]. That is, the Viola-Jone cascaded face detector is applied to extract the face images from all frame successively in each video.

We carry out the experiment using the first 50 frames in each video for this database. The shared database by [8] is used. For the video clips that contain less than 50 frames, all frames are selected in the experiment. The following methods are chosen for comparison: DCC [10], MMD [18], MDA [16], AHISD [2], CHISD [2], MSM [22], SANP [8], RNP [23], and DLRC [3]. Table 3 lists all recognition rates of these classifiers on this database. We can find that the recognition rates of PLRC-II, AHISD, RNP, and DLRC are all equal 87.18%, which is much better than that of DCC and MMD methods. The PLRC-I classifier obtains the highest accuracy 89.74% for this database.

4.2.2 CMU Mobo face database

The CMU Mobo database [6] consists the video sequences of 25 subjects that were captured on a treadmill. All except for the last sequence contain different videos collected in the following walking patterns, i.e., holding a ball, fast walking, slow walking, and incline walking. Following the common setting, we use the videos of the first 24 subjects to

Classifier	10 random splits
DCC	82.10+2.7
MMD	90.10+2.3
MDA	86.20+2.9
AHISD	91.60+2.8
CHISD	91.20+3.1
MSM	84.30+2.6
SANP	91.80+3.1
RNP	91.90+2.5
DLRC	93.06+3.4
PLRC-I	93.74+4.3
PLRC-II	93.75+4.3

Table 4. The recognition rate (RR) of several classifiers on CMU Mobo face database.

generate the video-based face databases. In [8], the Viola-Jones algorithm is used to obtain the faces from videos. Each image is extracted the local binary pattern (LBP) features. We exploit the public processed LBP histogram features to test different classifiers.

The proposed classifiers are compared with following methods, namely DCC [10], MMD [18], MDA [16], AHISD [2], CHISD [2], MSM [22], SANP [8], RNP [23], and DLRC [3]. To provide fair comparison, this experiment is set identically to [8], namely the "10 random splits" strategy is considered. Table 4 reports the recognition rates of all methods. We can see that the two proposed classifiers achieve similar performance for this database, and they both work better than other comparative algorithms.

4.3. Object Classification

In this experiment, we apply a challenging object database, Caltech101 database, to study the performance of the proposed methods. This database includes over 9000 images for 102 classes (like airplanes, elephant, pyramid, and sunflower, etc). For each image category, there are about 31 to 800 images. In the experiment, we select a subject of 102 subjects, and each contains 30 object images. Each image in this database is about a size of 300×200, which is transformed to spatial pyramid feature [9]. The following methods are selected for comparison, namely CRC+NN [24], RH-ISCRC [26], KCH-ISCRC [26], LRC+NN [13], and DLRC [3]. The experiment here is set as follows:

- **Scheme 1:** 5 samples per class are used Gallery set, the rest 25 samples are divided into 5 probe sets. Each set contains 5 samples.
- **Scheme 2:** 10 samples per class are used Gallery set, the rest 20 samples are divided into 4 probe sets. Each contains 5 samples.

The performances of all testing methods are presented in Table 5. We can observe that PLRC-I and PLRC-II obtain comparable performance on this database using Scheme 1, they lead an improvement of more than two percentages compared with the second best result 70.20 %. For the Scheme 2, the PLRC-II classifier work slightly better than DLRC methods, while PLRC-I outperforms DLRC about 3 percentages.

Classifier	Scheme 1	Scheme 2
KCH-ISCRC	69.81	76.72
RH-ISCRC	65.09	79.17
CRC+NN	60.20	62.50
LRC+NN	41.96	51.47
DLRC	70.20	80.39
PLRC-I	72.75	83.09
PLRC-II	72.35	80.64

Table 5. The recognition rate (RR) of several classifiers on the Caltech101 object database.

Classifier	RR	Run Time
SANP	67.00	18.96874
RH-ISCRC	69.00	5.68937
KCH-ISCRC	65.00	1.28978
DLRC	66.00	1.30427
PLRC-I	72.50	6.26050
PLRC-II	66.00	2.99981

Table 6. The recognition rate (RR) and run time of several classifiers on UCF50 action database.

4.4. Action Recognition

This experiment evaluates the proposed method using the action recognition application. The challenging database, UCF50 database [14], is selected here. There are totally 6,676 videos for 50 action categories in the UCF50 database collected from YouTube. In the experiment, 5000 videos of 50 categories with 100 videos are used. Our proposed method is compared following algorithms: SANP [8], RH-ISCRC [26], KCH-ISCRC [26] and DLRC [3]. To extensively test the performance of all testing methods, our experiments are conducted by following scheme:

- **Scheme:** 20 video samples per class are used Gallery set, the resting 80 video samples are divided into 4 probe set, each contains 20 video samples.

Table 6 reports the recognition rates (RR) of all methods. It can be seen that PLRC-I outperforms all other methods more than 3%.

4.5. Run Time

Run Time is another important aspect to evaluate an algorithm. This experiment compares the run time of the proposed PLRC classifiers with other methods using the

UCF50 action database. All competing classifiers are obtained with a desktop PC with 3.5GHz Intel CPU and 16 GB memory. The comparison results are given in Table 6. The result shows that The computational cost of PLRC-I and PLRC-II are both more than that of DLRC and lower than that of SANP. The computational cost of PLRC-II are about 2 times that of DLRC.

5. Conclusion

In this paper, a novel framework called pairwise linear regression classification (PLRC) is proposed for image recognition. Compared to DLRC, PLRC increases the unrelated subspace for classification. Based on the different methods of constituting the unrelated subspace, two classifiers are proposed in this paper. In order to prove the performance of two classifiers, a lot of experiments are evaluated on six database for three classification tasks: image-based face recognition, video-based face recognition and object recognition. All experimental results confirm the effectiveness of two proposed classification algorithms.

Acknowledgements

This work was supported in part by the Macau Science and Technology Development Fund under Grant FDC-T/106/2013/A3 and by the Research Committee at University of Macau under Grants MYRG2014-00003-FST, and MRG017/ZYC/2014/FST.

References

[1] O. Arandjelović, G. Shakhnarovich, J. Fisher, R. Cipolla, and T. Darrell. Face recognition with image sets using manifold density divergence. In *CVPR*, volume 1, pages 581–588. IEEE, 2005.

[2] H. Cevikalp and B. Triggs. Face recognition based on image sets. In *CVPR*, pages 2567–2573. IEEE, 2010.

[3] L. Chen. Dual linear regression based classification for face cluster recognition. In *CVPR*, pages 2673–2680. IEEE, 2014.

[4] T. M. Cover and P. E. Hart. Nearest neighbor pattern classification. *TIT*, 13(1):21–27, 1967.

[5] L. Fei-Fei, R. Fergus, and P. Perona. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. *Computer Vision and Image Understanding*, 106(1):59–70, 2007.

[6] R. Gross and J. Shi. The cmu motion of body (mobo) database. 2001.

[7] Y. Hu, A. S. Mian, and R. Owens. Sparse approximated nearest points for image set classification. In *CVPR*, pages 121–128. IEEE, 2011.

[8] Y. Hu, A. S. Mian, and R. Owens. Face recognition using sparse approximated nearest points between image sets. *T-PAMI*, 34(10):1992–2004, 2012.

[9] Z. Jiang, Z. Lin, and L. S. Davis. Label consistent k-svd: Learning a discriminative dictionary for recognition. *TPAMI*, 35(11):2651–2664, 2013.

[10] T.-K. Kim, J. Kittler, and R. Cipolla. Discriminative learning and recognition of image set classes using canonical correlations. *T-PAMI*, 29(6):1005–1018, 2007.

[11] K.-C. Lee, J. Ho, M.-H. Yang, and D. Kriegman. Video-based face recognition using probabilistic appearance manifolds. In *CVPR*, volume 1, pages I–313. IEEE, 2003.

[12] A. M. Martínez and A. C. Kak. Pca versus lda. *TPAMI*, 23(2):228–233, 2001.

[13] I. Naseem, R. Togneri, and M. Bennamoun. Linear regression for face recognition. *T-PAMI*, 32(11):2106–2112, 2010.

[14] K. K. Reddy and M. Shah. Recognizing 50 human action categories of web videos. *Machine Vision and Applications*, 24(5):971–981, 2013.

[15] C. Schüldt, I. Laptev, and B. Caputo. Recognizing human actions: a local svm approach. In *ICPR*, volume 3, pages 32–36. IEEE, 2004.

[16] R. Wang and X. Chen. Manifold discriminant analysis. In *CVPR*, pages 429–436. IEEE, 2009.

[17] R. Wang, H. Guo, L. S. Davis, and Q. Dai. Covariance discriminative learning: A natural and efficient approach to image set classification. In *CVPR*, pages 2496–2503. IEEE, 2012.

[18] R. Wang, S. Shan, X. Chen, and W. Gao. Manifold-manifold distance with application to face recognition based on image set. In *CVPR*, pages 1–8. IEEE, 2008.

[19] W. Wang, R. Wang, Z. Huang, S. Shan, and X. Chen. Discriminant analysis on riemannian manifold of gaussian distributions for face recognition with image sets. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2048–2057, 2015.

[20] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *TPAMI*, 31(2):210–227, 2009.

[21] Y. Wu, M. Minoh, M. Mukunoki, and S. Lao. Set based discriminative ranking for recognition. In *ECCV*, pages 497–510. Springer, 2012.

[22] O. Yamaguchi, K. Fukui, and K.-i. Maeda. Face recognition using temporal image sequence. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, pages 318–323. IEEE, 1998.

[23] M. Yang, P. Zhu, L. Van Gool, and L. Zhang. Face recognition based on regularized nearest points between image sets. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on. IEEE,*, pages 1–7. IEEE, 2013.

[24] L. Zhang, M. Yang, and X. Feng. Sparse representation or collaborative representation: Which helps face recognition? In *ICCV*, pages 471–478. IEEE, 2011.

[25] P. Zhu, L. Zhang, Q. Hu, and S. C. Shiu. Multi-scale patch based collaborative representation for face recognition with margin distribution optimization. In *ECCV*, pages 822–835. Springer, 2012.

[26] P. Zhu, W. Zuo, L. Zhang, S. C.-K. Shiu, and D. Zhang. Image set-based collaborative representation for face recognition. *Information Forensics and Security, IEEE Transactions on*, 9(7):1120–1132, 2014.