where

$$J_b = \frac{1}{L}\sum_{i=1}^{L}\sum_{j=1, j\neq l(x_i)}^{M}(x_i - x_{ij}^b)(x_i - x_{ij}^b)^T \tag{18}$$

and

$$J_w = \frac{1}{L}\sum_{i=1}^{L}(x_i - x_i^w)(x_i - x_i^w)^T \tag{19}$$

Afterwards, the objective function can be expressed as

$$\begin{aligned} &\arg\max_{P} \frac{P^T J_b P}{P^T J_w P} \\ &s.t. \quad P^T P = I \end{aligned} \tag{20}$$

In order to address the typical small sample size problem, the term εI is increased without affecting the subspace. Thus, the objective function can be rewritten as

$$\begin{aligned} &\arg\max_{P} \frac{P^T J_b P}{P^T (J_w + \varepsilon I) P} \\ &s.t. \quad P^T P = I \end{aligned} \tag{21}$$

where ε is a small number and I is an identity matrix. By using Lagrange multiplier, the projection matrix $P = [p_1, ..., p_k, ..., p_d]$ that maximizes the objective function, which can be gained by solving the eigen decomposition problem of $\frac{J_b}{J_w + \varepsilon I}$ as

$$J_b p_k = \lambda_k (J_w + \varepsilon I) p_k \quad , k = 1, 2, ..., d \tag{22}$$

where $\lambda_1 \geq ... \geq \lambda_k \geq ... \geq \lambda_d$ is d largest eigenvalues and their corresponding eigenvectors, $p_1, ..., p_k, ..., p_d$ of $\frac{J_b}{J_w}$. It is noted that $P = [p_1, ..., p_k, ..., p_d]$ is a $q \times d$ projection matrix, which can project the original q-element data vector to the new d-element data vector as $w_i = P^T x_i$ for $i = 1, 2...L$.

Classification

In the above Section, DPRC obtains the effective discriminant space $W = \{w_i \in R^{d\times 1}, i = 1, 2, \cdots, L\}$. Using the discriminant space W and Algorithm 1, the approximation projection p is obtained for the c^{th} class subspace. The distance between the test sample and the c^{th} class subspace can be computed as

$$d_c(w) = \|w - p\| \tag{23}$$

DPRC selects the class with the minimum distance

$$\min_{c^*} d_c(w), c = 1, 2, \cdots, M. \tag{24}$$

where $w = P^T x$.

DPRC vs ULDA

This section compares DPRC with a discriminant-based method: Uncertain LDA (ULDA) (Saeidi, Astudillo, and Kolossa 2016). To better explain it, their similarity and difference are given as follows.

- Similarity: They both maximize the following objective function $\max_P J(P) = \max_P \frac{J_b}{J_w}$, where J_b, J_w are the within-class and between-class scatters. This objective function is the same to that in LDA (Haeb-Umbach and Ney 1992).
- Difference: In ULDA, $J_b = S_b + U_b, J_w = S_w + U_w$, where S_b, S_w are the within-class and between-class scatters in LDA. ULDA proposes the uncertain within-class and between-class scatters U_b, U_w. In DPRC, the J_b, J_w can be treated as new projection-based within-class and between-class scatters, which has significantly difference to S_b, S_w and U_b, U_w.

Experimental Results

This section evaluate the proposed PRC and DPRC on several vision recognition databases.

Face recognition

LFW-a database (Zhu et al. 2012) is used in this experiment. Following (Zhang et al. 2015), we apply 158 subjects that have no less than ten samples for evaluation. The experiment set: 5 samples are randomly selected to form the training set, while other 2 samples are exploited for testing. The SRC (Wright et al. 2009), SVM (Schüldt, Laptev, and Caputo 2004), FDDL (Yang et al. 2014), MCT (Zhang et al. 2015), RCR (Yang et al. 2012), ULDA (Saeidi, Astudillo, and Kolossa 2016), ProCRC (Cai et al. 2016) and CRC (Zhang, Yang, and Feng 2011) algorithms are chosen for comparison. Table 1 illustrates the comparison results of all methods. DPRC obtains better performance than PRC. Compared to the exsiting methods, DPRC has more than 3% improvement..

Classifier	Accuracy	Classifier	Accuracy (%)
SRC	44.10	CRC	44.30
SVM	43.30	ULDA	44.30
FDDL	42.00	ProCRC	44.90
MCT	44.90	**PRC**	**46.84**
RCR	36.70	**DPRC**	**47.90**

Table 1: The recognition rate (RR) of several classifiers on LFW face database

Scene classification

The well-known 15 scene database contains 4,485 images of 15 scene categories (Lazebnik, Schmid, and Ponce 2006). Each image is transformed to spatial pyramid feature provided by (Jiang, Lin, and Davis 2013). The following experimental protocol is used (Liu and Liu 2015): 100 images per class are randomly chosen for training and the rest images are used for testing. The D-KSVD (Zhang and Li 2010), LLC (Wang et al. 2010), LC-KSVD (Jiang, Lin, and Davis 2013), ULDA (Saeidi, Astudillo, and Kolossa 2016), LLNMC (Liu and Liu 2015), LLKNNC (Liu and Liu 2015), LRC (Naseem, Togneri, and Bennamoun 2010), CRC (Zhang, Yang, and Feng 2011), SRC (Wright et al. 2009), ProCRC (Cai et al. 2016), DADL (Guo et al. 2016) methods are chosen for comparison. The average classification rate of 10 runs is used to evaluate all methods. From the results