

基于核扩展混合块字典的单样本人脸识别研究

马杲东¹, 吕 非², 童 莹³, 曹雪虹³

(1. 南京邮电大学 通信与信息工程学院, 江苏 南京 210003;

2. 南瑞集团有限公司, 江苏 南京 211106;

3. 南京工程学院 信息与通信工程学院, 江苏 南京 211167)

摘 要:稀疏表示分类(sparse representation-based classification, SRC)在样本数量充足下的人脸识别中具有较好的识别效果。然而由于基本字典缺乏判别性同时过度依赖于字典中每类样本的原子数目,稀疏表示分类在真实情况下的单样本(每类样本只有一张训练样本)人脸识别任务中缺乏鲁棒性。针对以上问题,该文提出了基于核扩展混合块字典的单样本人脸识别方法。首先,对样本进行分块处理,分别对分块图像进行核判别分析(kernel discriminant analysis, KDA)投影降维,提取图像的局部特征信息构成更具判别性的基本块字典;然后,为经过KDA投影之后的分块样本分别构建遮挡字典和类内差异字典来描述样本中的大面积连续遮挡以及光照、表情等类内差异信息,将遮挡字典和类内差异字典共同组合成混合块字典,使混合块字典能够更好地描述测试样本中不同类型的差异信息;最后,将测试样本表示为基本块字典和混合块字典的稀疏线性组合,根据重构残差进行分类识别,从而实现真实情况下的单样本人脸识别。在标准人脸库CAS-PEAL, AR以及真实人脸库LFW和PubFig上的实验结果表明,该方法与其他方法相比有较好的结果。

关键词:稀疏表示分类;核判别分析;人脸识别;混合块字典;单样本

中图分类号: TP273

文献标识码: A

文章编号: 1673-629X(2022)01-0104-07

doi:10.3969/j.issn.1673-629X.2022.01.018

Face Recognition with a Single Training Sample Per Person Based on Kernel Extended Hybrid Block Dictionary Learning

MA Gao-dong¹, LYU Fei², TONG Ying³, CAO Xue-hong³

(1. School of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China;

2. NARI Group Corporation, Nanjing 211106, China;

3. School of Information and Communication Engineering, Nanjing Institute of Technology, Nanjing 211167, China)

Abstract: Sparse representation-based classification (SRC) is effective in face recognition with sufficient samples. However, due to the lack of discriminativeness of its basic dictionary and excessive dependence on the number of atoms of each class of sample in the dictionary, SRC lacks robustness in face recognition tasks of single sample per person (SSPP). Therefore, we propose a single sample face recognition method based on kernel extended hybrid block dictionary. Firstly, the samples are divided into blocks, and the kernel discriminant analysis (KDA) projection dimension reduction is performed on the divided images respectively, and the local feature information of the image is extracted to form a more discriminative basic block dictionary. Then, for the blocked samples after KDA projection, an occlusion dictionary and an intra-class difference dictionary are constructed to describe the large-area continuous occlusion in the sample, as well as the intra-class difference information such as illumination and expression. The occlusion dictionary and the intra-class dictionary are combined to construct the hybrid block dictionary and enable the hybrid block dictionary to better describe the intra-class variation in the test sample. Finally, the test samples are represented as sparse linear combinations of the basic block dictionary and the hybrid block dictionary. The classification is determined by the reconstruction residuals. Experiment on standard face databases CAS-PEAL, AR and real face databases LFW and PubFig shows that the proposed method has better results compared with other methods.

Key words: sparse representation-based classification; kernel discriminant analysis; face recognition; hybrid block dictionary; single sample per person

0 引言

人脸识别技术由于其在图像处理、计算机视觉领域的广泛应用而获得了广泛的关注^[1]。在过去的十几年中,人脸识别技术取得了巨大的成就。但是由于待测试人脸中包含大量光照、阴影、姿势、表情以及遮挡等面部变化,人脸识别仍然是一项极具挑战性的任务。

随着人脸识别的应用场景越来越广泛,在执法、身份识别以及机场监控等实际应用场景中,鉴于有限的存储空间和个人隐私的考虑,每个人可能只有一张样本。当没有足够多的训练样本来预测待测试样本中的差异信息时,人脸识别问题就变得十分困难,这就产生了所谓的单样本人脸识别问题(single sample per person, SSPP)。

如何在这种单样本情况下得到较高的鲁棒性是人脸识别中最重要也是最具挑战性的一个问题。传统的判别子空间方法以及流形学习算法^[2]无法直接用来应用在单样本问题中。Wright等^[3]提出的稀疏表示分类(sparse representation-based classification, SRC)以及Zhang等^[4]提出的协作表示分类(collaborative representation-based classification, CRC)需要使用足够过的样本来描述测试样本,因此无法在单样本人脸识别中取得很好的性能。

近年来,研究者们为了解决单样本的人脸识别问题提出了许多方法。这些方法大致可以分为两类^[5]:全局方法和局部方法。全局方法使用整张人脸图像作为输入来识别测试样本。对于全局方法,有两个主要方向^[6]:一是生成虚拟样本将单样本问题转化为传统多样本问题,二是通用学习方法。生成虚拟样本指的是通过奇异值分解、几何变换等方法将原图像生成多张虚拟样本。例如,彭帆等^[7]提出了基于WSSRC的样本扩充方法,采用一种三层级联的虚拟样本产生方法获取冗余样本,将生成的多种表情和多种姿态的新样本当成训练样本。Zhang等^[8]通过几何变换来创建更多的虚拟训练样本。这些方法的主要缺点是虚拟样本和原始样本总是高度相关,导致虚拟样本所产生的新信息比较有限。

与虚拟样本方法相比,通用学习方法通常借助一个通用训练集来补充原始的SSPP验证集。例如,Wang等^[9]提出一种通用的学习框架来学习通用训练集中每类人的类内差异信息,而通用训练集中的类内差异信息可以被所有人共享。Deng等^[10]对SRC^[3]进行了改进,提出了一种扩展稀疏表示分类器(extended sparse representation-based classification, ESRC),该方法通过通用训练集构建一种辅助的类内差异字典来表示训练样本和测试样本之间的差异,Yang等^[11]通过学习稀疏差异字典来描述人脸中的各种光照、遮挡等

变化。Ji等^[12]提出协作概率标签方法(collaborative probabilistic labels, CPL)通过标签传播来实现人脸识别。尽管这些方法在一定程度上可以提升SSPP的性能,但是这些方法的性能在很大程度上取决于通用训练集的选择。

局部方法利用面部的局部特征来识别待测试样本。一些方法将图像分成几个重叠或者不重叠的图像块,用于生成局部特征,其中每个分块的样本被认为是该人的独立样本。基于这种假设,研究人员对传统的子空间学习方法以及基于表示的分类器进行了扩展,例如模块化PCA^[13]、模块化LDA^[14]、基于补丁的SRC^[3]。这些方法通过整合每个分块的输出结果来实现SSPP人脸识别。但是由于这些方法没有考虑到每个分块之间的相关性,因此存在一定的局限性。PCA^[15]和LDA^[16]等方法不能很好地处理复杂的非线性特征,它们的核化版本KPCA^[17]和KDA^[18]通过非线性核映射将样本映射到一个高维空间中,然后在高维空间中进行PCA和LDA,能够有效地提取人脸的非线性特征。

Huang等^[19]学习了一种核扩展字典(kernel extended dictionary, KED),通过KDA对样本进行判别分析,另外构建了一个遮挡模型来处理遮挡问题。一些方法尝试结合LBP^[20]特征以及Gabor特征等传统的人脸特征提出更具判别性的特征以用于SSPP人脸识别。例如,马振等^[21]将分层LBP特征与金字塔模式HOG特征相融合从而得到更具判别性的特征。Wang等^[22]提出一种三重局部特征的鲁棒联合表示,提取多个方向的局部特征来构建更具判别性的特征。文献[23]中采用旋转主方向梯度直方图特征算子提取非约束人脸图像的多尺度多方向梯度特征。文献[24]同时考虑样本的局部结构信息和全局分布信息,提出了有效的降维算法。

受上述研究工作的启发,该文提出了一种基于核扩展混合块字典(kernel extended hybrid block dictionary, KEHBD)的单样本人脸识别方法,用于解决SSPP人脸识别问题。首先,对样本进行分块处理,分别对分块图像进行核判别分析(kernel discriminant analysis, KDA)投影降维,提取图像的局部特征信息构成更具判别性的基本块字典;然后,为经过KDA投影之后的分块样本分别构建遮挡字典和类内差异字典来描述样本中的大面积连续遮挡以及光照、表情等类内差异信息,将遮挡字典和类内差异字典共同组合成混合块字典,使混合块字典能够更好地描述测试样本中不同类型的差异信息;最后,将测试样本表示为基本块字典和混合块字典的稀疏线性组合,根据重构残差进行分类识别。

1 相关方法

1.1 稀疏表示分类和扩展稀疏表示分类

假设 $X = [x_1, x_2, \dots, x_n] \in R^{d \times n}$ 是验证集中的 n 个样本, 其中 d 是样本的特征维数, x_i 的类别标签是 c_i 。给定一张待测试样本 $y \in R^{d \times 1}$ 。稀疏表示的基本原理就是将测试样本表示为训练样本的稀疏线性组合并通过式(1)来求解稀疏系数 $\beta \in R^{n \times 1}$:

$$\min_{\beta} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \quad (1)$$

令向量 $\delta_c(\beta) \in R^{n \times 1}$ 表示除了与 c 类样本相关的系数之外其余的系数全为0, 通过最小化残差来对 y 进行分类。

$$r_c(y) = \|y - X\delta_c(\beta)\|_2 \quad (2)$$

在很多实际的人脸识别场景中, 由于测试样本中会包含遮挡和污染, SRC 通过式(3)来计算稀疏表示, 其中 $I \in R^{d \times d}$ 是单位矩阵。

$$\min_{\beta, \hat{\beta}} \left\| y - [X, I] \begin{bmatrix} \beta \\ \hat{\beta} \end{bmatrix} \right\|_2^2 + \lambda \left\| \begin{bmatrix} \beta \\ \hat{\beta} \end{bmatrix} \right\|_1 \quad (3)$$

扩展稀疏表示分类(extended sparse representation-based classification, ESRC)使用类内差异字典来代替 SRC 中的单位矩阵作为扩展字典。假设 $X_i = [x_{i,1}, x_{i,2}, \dots, x_{i,n_i}] \in R^{d \times n_i}$ 表示第 i 类样本, μ_i 表示第 i 类样本的标准样本或者该类样本的均值。则第 i 类样本的类内差异表示为:

$$E_i = [x_{i,1} - \mu_i, x_{i,2} - \mu_i, \dots, x_{i,n_i} - \mu_i] \quad (4)$$

将所有类样本的类内差异组合成类内差异字典 E 。然后 ESRC 用 E 来取代 SRC 中的单位矩阵 I , 并通过式(5)计算测试样本的稀疏表示:

$$\min_{\beta, \hat{\beta}} \left\| y - [X, E] \begin{bmatrix} \beta \\ \hat{\beta} \end{bmatrix} \right\|_2^2 + \lambda \left\| \begin{bmatrix} \beta \\ \hat{\beta} \end{bmatrix} \right\|_1 \quad (5)$$

1.2 核判别分析

核判别分析(kernel discriminant analysis, KDA)的基本思想是先通过一个非线性映射 $\varphi: R^d \rightarrow F$, 将样本映射到一个高维特征空间 F , 在此高维特征空间根据 Fisher 准则学习到一个最优投影矩阵 v^φ , 然后用该投影矩阵将样本投影到一个低维空间。在高维空间中学习 v^φ 时, 其计算只需在原始图像中完成, 无需了解非线性映射的具体含义, 只需定义一个与之相对应的核函数。

定义高维特征空间 F 中的内积为:

$$\langle \varphi(x_1), \varphi(x_2) \rangle = k(x_1, x_2) \quad (6)$$

假设 $X = [x_1, x_2, \dots, x_n]$ 是用于训练 KDA 投影矩阵的 \hat{n} 张样本。KDA 的优化目标为:

$$a_{\text{opt}} = \arg\max \left(\frac{a^T K W K a}{a^T K K a} \right) \quad (7)$$

其中, K 为核矩阵, W 为系数矩阵。通过求解特征值

问题 $K W K a = \lambda K K a$ 得到特征向量作为 a_{opt} 。得到一个

投影向量: $(v^\varphi) = \sum_{i=1}^{\hat{n}} a_i \varphi(x_i)$, 取 $C-1$ 个最大特征值对应的特征向量得到 KDA 投影矩阵 $(V^\varphi) = [(v_1^\varphi), (v_2^\varphi), \dots, (v_{C-1}^\varphi)]$ 。使用投影矩阵对一张新的样本 x 进行投影:

$$\langle (V^\varphi), \varphi(x) \rangle = A^T K(:, x) \quad (8)$$

其中, $A = [a_1, a_2, \dots, a_{C-1}]$, $K(:, x) = [k(x_1, x), k(x_2, x), \dots, k(x_n, x)]$ 。

2 核扩展混合块字典

本节介绍核扩展混合块字典的构建。首先对图像进行分块得到多个子块。在 SSPP 的情况下, 验证集中每个人只有一张样本, 因此需要构建判别性强的字典才能用一张样本去表示同类别的待测试样本。由于 KDA 通过非线性映射能够提取复杂的非线性特征, 同时引入类别标签可以使样本的区分度更高。因此对每个子块样本分别进行 KDA 投影, 构建每个子块的基本字典。由于待测试样本中包含大量光照、表情、姿势以及遮挡等类内变化, 通过 KDA 投影并不能消除这些类内变化, 因此为每个子块分别构建混合字典, 分别提取通用训练集中的遮挡信息和非遮挡类内变化信息来构建遮挡字典和类内差异字典, 两者共同组成的混合字典可以很好地描述每个子块中的类内变化。

2.1 图像分块方式

本小节讨论图像的分块方式。相比于整张图像的情况, 对图像进行分块处理一方面可以充分提取图像的局部特征, 另一方面可以有效减少大面积遮挡的干扰。分块处理方法可以分为无重叠的分块和有重叠的分块。由于分块的大小, 形状等因素都会对效果产生不同的影响, 该文首先采用多种无重叠分块方式。如图1所示, 分块1、2、3分别将图像均匀分成 1×2 块, 2×1 块, 2×2 块, 分块4、5、6、7分别将图像分成 4×1 和 1×4 块, 3×3 块, 4×4 块。通过实验在上述分块方式中选出效果最好的分块方式, 然后对该分块方式进行有重叠的分块, 更进一步提取人脸中最为丰富的局部特征, 有重叠的分块将在实验部分阐述。

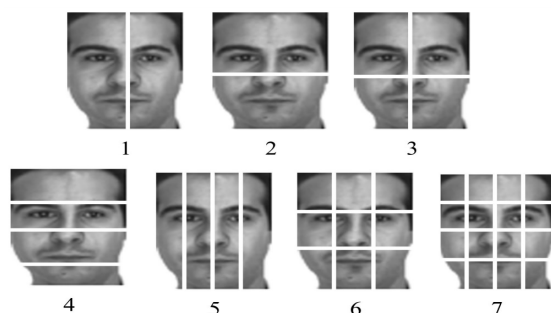


图1 图像分块方式

2.2 构建基本块字典

将样本进行分块处理,每一块的大小相同。 $\mathbf{X}^{\mathfrak{S}} = [x_1^{\mathfrak{S}}, x_2^{\mathfrak{S}}, \dots, x_n^{\mathfrak{S}}] \in R^{d \times \hat{n}}$ 表示第 \mathfrak{S} 个子块用于训练 KDA 投影矩阵的样本, d 表示子块的维度, \hat{n} 表示样本的数量。通过训练得到第 \mathfrak{S} 个子块的投影矩阵 $(\mathbf{V}^{\varphi})^{\mathfrak{S}}$, 将第 \mathfrak{S} 个子块的验证样本 $\mathbf{G}^{\mathfrak{S}} = [g_1^{\mathfrak{S}}, g_2^{\mathfrak{S}}, \dots, g_c^{\mathfrak{S}}] \in R^{d \times c}$ 通过 KDA 投影得到第 \mathfrak{S} 个子块的基本字典 $\mathbf{A}^{\mathfrak{S}}$:

$$\mathbf{A}^{\mathfrak{S}} = \langle (\mathbf{V}^{\varphi})^{\mathfrak{S}}, \varphi(\mathbf{G}^{\mathfrak{S}}) \rangle \quad (9)$$

KDA 的关键在于构建一个合适的核函数。由于使用 LBP 特征^[22]时 χ^2 距离的效果更好,故构建以下基于 LBP 特征的径向基核函数:

$$k(x, x') = \exp\left[-\frac{\chi^2(x, x')}{\sigma^2}\right] \quad (10)$$

其中, $\chi^2(x, x') = \sum_i ((x_i - x'_i)/x_i + x'_i + \varepsilon)$, ε 的取值为 0.01。该文采用多尺度的 LBP 特征作为核函数的输入。首先建立一个比例尺为 $\sqrt{2}$ 的三层图像金字塔,对每个层图像进行 LBP 特征提取,最后将每个比例尺的特征连接起来组合成最终的 LBP 特征。

2.3 构建混合块字典

ESRC 中统一构建类内差异字典,没有区分遮挡与光照,表情等类内差异信息的区别从而导致字典缺乏判别性。为了区分连续的遮挡以及光照表情等类内信息的干扰,为分块后的样本分别构建遮挡字典和类内差异字典,由这两者共同组成混合字典。

假设 $\mathbf{X}_{occ}^{\mathfrak{S}} = [x_1^{\mathfrak{S}}, x_2^{\mathfrak{S}}, \dots, x_n^{\mathfrak{S}}] \in R^{d \times n}$ 是 n 张含有遮挡的样本的第 \mathfrak{S} 个子块, $\mathbf{X}_{nor}^{\mathfrak{S}} = [(x'_1)^{\mathfrak{S}}, (x'_2)^{\mathfrak{S}}, \dots, (x'_n)^{\mathfrak{S}}] \in R^{d \times n}$ 是与 $\mathbf{X}_{occ}^{\mathfrak{S}}$ 相对应的标准样本的第 \mathfrak{S} 个子块,直接将一张遮挡样本 $x^{\mathfrak{S}}$ 与对应的标准样本 $(x')^{\mathfrak{S}}$ 相减可以得到遮挡样本中的遮挡信息 $x^{\mathfrak{S}} - (x')^{\mathfrak{S}}$, 由于 KDA 投影矩阵 $(\mathbf{V}^{\varphi})^{\mathfrak{S}}$ 是根据人脸信息训练得到的,遮挡信息 $x^{\mathfrak{S}} - (x')^{\mathfrak{S}}$ 是一个奇异样本,通过 KDA 投影至低维空间并不能很好地描述低维空间中的遮挡信息。为了构建低维空间中的遮挡字典,先将 $x^{\mathfrak{S}}$ 以及 $(x')^{\mathfrak{S}}$ 进行 KDA 投影至低维空间,再将它们进行相减得到两者之间的差异信息,以构建遮挡字典:

$$\mathbf{D}_1^{\mathfrak{S}} = (\mathbf{X}_{occ}^{\mathfrak{S}})_{\text{KDA}} - (\mathbf{X}_{nor}^{\mathfrak{S}})_{\text{KDA}} \quad (11)$$

其中, $(\mathbf{X}_{occ}^{\mathfrak{S}})_{\text{KDA}} = \langle (\mathbf{V}^{\varphi})^{\mathfrak{S}}, \varphi(\mathbf{X}_{occ}^{\mathfrak{S}}) \rangle$, $(\mathbf{X}_{nor}^{\mathfrak{S}})_{\text{KDA}} = \langle (\mathbf{V}^{\varphi})^{\mathfrak{S}}, \varphi(\mathbf{X}_{nor}^{\mathfrak{S}}) \rangle$ 。

同理,为了构建类内差异字典,将同一类样本的 m 张光照、表情以及姿势变化样本 $\mathbf{X}_{interclass}^{\mathfrak{S}} = [x_1^{\mathfrak{S}}, x_2^{\mathfrak{S}}, \dots, x_m^{\mathfrak{S}}] \in R^{d \times m}$ 投影至低维空间,与这些样本在低维空间中的均值相减得到第 \mathfrak{S} 个子块样本的类内差异信息,从而得到该类样本的类内差异字典:

$$\mathbf{d}^{\mathfrak{S}} = [(x_1^{\mathfrak{S}})_{\text{KDA}} - \mu_{\text{KDA}}^{\mathfrak{S}}, (x_2^{\mathfrak{S}})_{\text{KDA}} - \mu_{\text{KDA}}^{\mathfrak{S}}]$$

$$\mu_{\text{KDA}}^{\mathfrak{S}}, \dots, (x_m^{\mathfrak{S}})_{\text{KDA}} - \mu_{\text{KDA}}^{\mathfrak{S}}] \quad (12)$$

其中, $(x_i^{\mathfrak{S}})_{\text{KDA}} = \langle (\mathbf{V}^{\varphi})^{\mathfrak{S}}, \varphi(x_i^{\mathfrak{S}}) \rangle$, $\mu_{\text{KDA}}^{\mathfrak{S}} = \frac{1}{m} \sum_{i=1}^m (x_i^{\mathfrak{S}})_{\text{KDA}}$ 。假设共有 k 类训练样本,学习每一类样本的类内差异字典,将 k 类的类内差异字典组合起来得到最终的类内差异字典:

$$\mathbf{D}_2^{\mathfrak{S}} = [d_1^{\mathfrak{S}}, d_2^{\mathfrak{S}}, \dots, d_k^{\mathfrak{S}}] \quad (13)$$

将 $\mathbf{D}_1^{\mathfrak{S}}$ 和 $\mathbf{D}_2^{\mathfrak{S}}$ 组合起来,就得到了第 \mathfrak{S} 个子块样本的混合字典 $\mathbf{D}^{\mathfrak{S}} = [\mathbf{D}_1^{\mathfrak{S}}, \mathbf{D}_2^{\mathfrak{S}}]$ 。

3 基于核扩展混合块字典的单样本人脸识别算法

本小节给出文中所提出方法的具体步骤。由于对图像进行分块,使用 SRC 分类器进行分类时,只能计算每一块各自对应的残差,最后将残差进行累和,这样每一个子块之间就失去了关联性,从而损失很多结构信息。文献[25]中提出一种协作表示分类器 (relaxed collaborative representation, RCR), 可用于分块图像的分类。RCR 可以利用不同块样本之间的相似性和判别性来进行编码和分类。因此文中采用 RCR 分类器进行分类,RCR 分类器的参数设置与文献[25]中相同。

算法:基于核扩展混合字典的单样本人脸识别 (KEHBD)。

输入:分成 q 块的训练样本 $\{\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^{\mathfrak{S}}, \dots, \mathbf{X}^q\}$, 其中 $\mathbf{X}^{\mathfrak{S}} = [x_1^{\mathfrak{S}}, x_2^{\mathfrak{S}}, \dots, x_n^{\mathfrak{S}}] \in R^{d \times \hat{n}}$, 验证样本 $\{\mathbf{G}^1, \mathbf{G}^2, \dots, \mathbf{G}^{\mathfrak{S}}, \dots, \mathbf{G}^q\}$, 其中 $\mathbf{G}^{\mathfrak{S}} = [g_1^{\mathfrak{S}}, g_2^{\mathfrak{S}}, \dots, g_c^{\mathfrak{S}}] \in R^{d \times c}$, 测试样本 $\{y^1, y^2, \dots, y^{\mathfrak{S}}, \dots, y^q\}$, 其中 $y^{\mathfrak{S}} \in R^{d \times 1}$ 。

输出:测试样本类别标签。

步骤 1:用训练样本 $\mathbf{X}^{\mathfrak{S}} = [x_1^{\mathfrak{S}}, x_2^{\mathfrak{S}}, \dots, x_n^{\mathfrak{S}}] \in R^{d \times \hat{n}}$ 训练出 KDA 投影矩阵 $(\mathbf{V}^{\varphi})^{\mathfrak{S}}$, 根据公式(9)将验证样本进行 KDA 投影得到基本字典 $\mathbf{A}^{\mathfrak{S}}$ 。

步骤 2:通过公式(11)得到遮挡字典 $\mathbf{D}_1^{\mathfrak{S}}$, 根据公式(13)得到类内差异字典 $\mathbf{D}_2^{\mathfrak{S}}$ 。 $\mathbf{D}_1^{\mathfrak{S}}$ 和 $\mathbf{D}_2^{\mathfrak{S}}$ 共同组成混合字典 $\mathbf{D}^{\mathfrak{S}} = [\mathbf{D}_1^{\mathfrak{S}}, \mathbf{D}_2^{\mathfrak{S}}]$ 。

步骤 3:将测试样本 $y^{\mathfrak{S}} \in R^{d \times 1}$ 进行 KDA 投影, $y_{\text{KDA}}^{\mathfrak{S}} = \langle (\mathbf{V}^{\varphi})^{\mathfrak{S}}, \varphi(y^{\mathfrak{S}}) \rangle$ 。

步骤 4:重复步骤 1~3, 得到每一块的基本字典 $[\mathbf{A}^1, \mathbf{A}^2, \dots, \mathbf{A}^q]$ 和混合字典 $[\mathbf{D}^1, \mathbf{D}^2, \dots, \mathbf{D}^q]$ 。

步骤 5:使用 RCR 分类器^[25]进行分类。

4 实验结果与分析

本节在 CAS-PEAL、AR、LFW 和 PubFig 这四个不同的人脸数据库上进行仿真实验来验证文中方法的有效性。所有的人脸图像裁剪成 120×100 大小。实验

环境为:Win10 64 位操作系统,内存 8 GB, MATLAB R2017a。

4.1 CAS-PEAL 数据库

CAS-PEAL 人脸数据库包含 1 040 类人(595 位男性,445 位女性)共 99 594 张人脸图像。使用其中包含了 1 040 类人的 9 031 张图像的子集进行实验。图 2 展示了 CAS-PEAL 人脸数据库中的部分样本。



图 2 CAS-PEAL 库中的部分图像

在 CAS-PEAL 数据库上目标对象的标准数据集,非目标对象的通用数据集,遮挡数据集以及类内变化数据集的设计如下:

(1)非目标对象的通用数据集包含光照变化的 180 类人和光照变化的 80 类人,每类人有 4 张图像,共 1 040 张变化样本。同时,数据集中还包含每类人 1 张正脸无干扰图像,共 260 张标准样本,他们共同组成非目标对象的通用数据集,用于训练 KDA 投影矩阵。

(2)非目标对象的遮挡数据集包含配饰遮挡的 20 类人,每类人有 4 张图像,共 80 张图像,用于构建遮挡字典。

(3)非目标对象的类内变化数据集包含光照变化 20 类人和表情变化 20 类人,每类人 4 张图像,共 160 张图像,用于构建类内差异字典。

(4)目标对象的验证样本集由 CAS-PEAL 数据库中 767 类人(与上述目标对象不同类的人),每类取 1 张正脸无干扰图像构成,共 767 张样本,用于构建基本字典。

(5)目标对象的测试样本集由 767 类人包含了配饰遮挡、光照、表情、距离、时间以及背景变化的所有样本组成。

4.1.1 非分块实验

为了验证提出的混合字典的有效性,并说明混合字典的有效性是不依赖于图像分块的,首先不对样本进行分块并使用 SRC^[3]分类器进行分类,将该方法记为 KEHD。KEHD 与 KEHBD 的区别在于:KEHD 省略了图像分块的步骤;KEHD 采用 SCR 分类器进行分类。

将 KEHD 与 SRC^[3]、ESRC^[10]、LDA^[16]、KDA^[18]以及 KED^[19]进行比较,上述所有的方法都是基于相同的 LBP 特征。对于 SRC 和 ESRC,鉴于 LBP 特征的高维性,使用 PCA 将特征维数降至 600。表 1 是不同方法在 CAS-PEAL 库上的识别率。

表 1 CAS-PEAL 数据库中的识别率 %

方法	识别率
SRC	82.27
ESRC	90.58
LDA	85.17
KDA	86.84
KED	91.16
KEHD	92.17

从表中可以看出,KEHD 识别率为 92.17%,在所有方法中最高。可见即使在没有分块的情况下,文中提出的核扩展混合字典由于融合了遮挡信息以及类内差异信息,对不同的干扰信息都具有较高的鲁棒性。

4.1.2 分块实验

本节实验是为了对不同的分块方式进行对比。将 KEHBD 按照图 1 的 7 种分块方式进行分块,分别记为 KEHBD-1 ~ KEHBD-7。实验结果如表 2 所示。

表 2 CAS-PEAL 数据库中 KEHBD 的识别率 %

方法	识别率
KEHBD-1	94.96
KEHBD-2	94.50
KEHBD-3	96.15
KEHBD-4	95.92
KEHBD-5	95.12
KEHBD-6	96.00
KEHBD-7	95.43

由表 2 可知,无论是哪种分块方式都比不分块时的识别率高,这是由于进行分块之后可以提取图像的局部特征,通过 RCR^[25]分类器可以将每一块的特征进行融合从而提升识别率。在 KEHBD-1 ~ KEHBD-7 中,第三种分块方式(将图像分成 4 * 4 块)的识别率最高,达到了 96.15%。选择分块 3 的分块方式对图像进行重叠分块,分别设置重叠率为 0.5(相邻两块的重叠面积为 50%)和 0.7。重叠率为 0.5 时会将图像分成 9 块。重叠率为 0.7 时分成 16 块。表 3 是不同重叠率下 KEHBD 的识别率。

表 3 CAS-PEAL 数据库中 KEHBD 不同重叠率下的识别率 %

重叠率	识别率
0	96.15
0.5	98.13
0.7	97.52

可以看出,设置了重叠率之后识别率进一步提高。这是由于设置了重叠率之后,能够在整张人脸范围内提取出最具有判别性的特征从而提高了识别率。但是

如果重叠率太高会增加计算的负担,因此综合考虑选择分块3的分块方式,以及0.5的重叠率为最佳的分块方式,如图3所示。



图3 最佳分块方式

4.2 AR 数据库

AR人脸数据库中包含126类人的4000多张正面对齐人脸。每类有26张图像,分为两个阶段,每个阶段13张图像,其中标准图像1张,光照变化图像3张,表情变化3张,眼镜遮挡3张,围脖遮挡3张。实验选取了100类样本。图4为AR人脸库中某类人的样本图像。



图4 AR库中某一类人的人脸图像

在AR数据库上目标对象的标准数据集,非目标对象的通用数据集,遮挡数据集以及类内变化数据集的设计如下:

- (1)非目标对象的通用数据集包含70类人的每类前7张无遮挡样本,共490张样本,用于训练KDA。
- (2)非目标对象的遮挡数据集包含70类人的每类第二阶段6张遮挡样本,共420张样本,用于构建遮挡字典。
- (3)非目标对象的类内变化数据集包含70类人的每类第二阶段6张无遮挡样本,共420张样本,用于构建类内差异字典。
- (4)目标对象的验证样本集包含30类人的每类第一张标准样本(与上述目标对象不同类的人),共30张样本,用于构建基本字典。
- (5)目标对象的测试样本包含30类人的每类剩余25张样本,共750张样本。

表4是所有方法在AR人脸库中的识别率,KEHBD是按图3最佳分块方式的情况。由表4可以看出,KEHBD取得了最高的识别率,证明了KEHBD的有效性。

表4 AR数据库中单样本情况的识别率 %

方法	识别率
SRC	84.67
ESRC	94.00
LDA	80.65

续表4

方法	识别率
KDA	82.27
KED	86.83
KEHBD	96.93

4.3 LFW 和 PubFig 数据库

LFW人脸数据库和PubFig人脸数据库中的人脸面部都是在不受环境约束和不准确对齐的情况下获得的,这对于人脸识别具有很大的挑战性。这两个数据库中的部分图像如图5和图6所示。



图5 LFW库中某一类人的人脸图像



图6 PubFig库中某一类人的人脸图像

由于LFW和PubFig数据集中的人脸不对齐,选择遮挡样本以及相应的标准样本较困难,因此只构建类内差异字典来描述样本中的差异信息。在LFW数据库上目标对象的标准数据集,非目标对象的通用数据集,类内变化数据集的设计如下:

- (1)非目标对象的通用数据集包含100类人的每类5张样本,共500张样本,用于训练KDA。
- (2)非目标对象的类内变化数据集包含100类人的每类5张样本(与通用数据集不重叠),共500张样本,用于构建类内差异字典。
- (3)目标对象的验证样本集包含58类人的每类第一张标准样本(与上述目标对象不同类的人),共58张样本,用于构建基本字典。
- (4)目标对象的测试样本由58类人的每类剩余全部样本组成。

在PubFig数据库上的设计如下:

- (1)非目标对象的通用数据集包含70类人的每类10张样本,共700张样本,用于训练KDA。
- (2)非目标对象的类内变化数据集包含70类人的每类剩余10张样本,共700张样本,用于构建类内差异字典。
- (3)目标对象的验证样本集包含30类人的每类第一张标准样本(与上述目标对象不同类的人),共30张样本,用于构建基本字典。
- (4)目标对象的测试样本由30类人的每类剩余19张样本组成,共570张样本。

表 5 展示了不同方法在 LFW 和 PubFig 中的识别率, 由于数据库的难度较高, 因此选择分块 3 中 0.7 重叠率的分块方式, 一张图像会被分成 16 块 (参考图 3)。从表中可以看出, 在 LFW 和 PubFig 中 KEHBD 在性能上均优于其他算法, LFW 中 KEHBD 的识别率达到了 65.94%, PubFig 数据库中 KEHBD 的识别率达到了 34.56%, 与其他方法相比有很大程度上提升, 可见 KEHBD 在非控环境下仍然有较强的鲁棒性。由于人脸图像不对齐, 传统的方法在这种 SSPP 识别场景中缺乏鲁棒性。文中方法通过构建核扩展混合块字典, 通过分块提取图像中的局部特征, 并为每个分块图像构建混合字典去描述可能的误差, 从而明显提高了识别的准确率。

表 5 LFW 和 PubFig 数据库中的识别率 %

数据库	SRC	ESRC	LDA	KDA	KED	KEHBD
LFW	32.72	33.89	35.80	37.96	38.13	65.94
PubFig	23.16	30.81	29.30	30.00	30.23	34.56

5 结束语

针对单样本情况下人脸识别中存在的问题, 提出一种基于核扩展混合块字典 (KEHBD) 的单样本人脸识别方法。首先, 对样本进行分块处理, 分别对分块图像进行核判别分析投影降维, 提取图像的局部特征信息构成更具判别性的基本块字典; 然后, 为经过 KDA 投影之后的分块样本分别构建遮挡字典和类内差异字典来描述样本中的大面积连续遮挡以及光照、表情等类内差异信息, 将遮挡字典和类内差异字典共同组合成混合块字典, 使混合块字典能够更好地描述测试样本中不同类型的差异信息。从而实现真实情况下的单样本人脸识别。

参考文献:

- [1] SNEHA S S. Face recognition techniques; a survey[J]. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 2018, 4(1): 221-226.
- [2] BELHUMEUR P N, HESPANHA J P, KRIEGMAN D J. Eigenfaces vs. fisherfaces; recognition using class specific linear projection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720.
- [3] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 31(2): 210-227.
- [4] ZHANG L, YANG M, FENG X. Sparse representation or collaborative representation; which helps face recognition? [C]//2011 international conference on computer vision. Barcelona, Spain: IEEE, 2011: 471-478.
- [5] GAO S, JIA K, ZHUANG L, et al. Neither global nor local: regularized patch-based representation for single sample per person face recognition[J]. International Journal of Computer Vision, 2015, 111(3): 365-383.
- [6] LU J, TAN Y P, WANG G. Discriminative multim manifold analysis for face recognition from a single training sample per person[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 39-51.
- [7] 彭帆, 徐志京, 叶丽. 基于 WSSRC 单样本人脸识别及样本扩充方法研究[J]. 计算机测量与控制, 2016, 24(10): 154-157.
- [8] ZHANG D, CHEN S, ZHOU Z H. A new face recognition method based on SVD perturbation for single example image per person [J]. Applied Mathematics and Computation, 2005, 163(2): 895-907.
- [9] WANG J, PLATANOTIS K N, LU J, et al. On solving the face recognition problem with one training sample per subject [J]. Pattern Recognition, 2006, 39(9): 1746-1762.
- [10] DENG W, HU J, GUO J. Extended SRC: undersampled face recognition via intraclass variant dictionary[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(9): 1864-1870.
- [11] YANG M, VAN GOOL L, ZHANG L. Sparse variation dictionary learning for face recognition with a single training sample per person [C]//Proceedings of the IEEE international conference on computer vision. Sydney, NSW, Australia: IEEE, 2013: 689-696.
- [12] JI H K, SUN Q S, JI Z X, et al. Collaborative probabilistic labels for face recognition from single sample per person[J]. Pattern Recognition, 2017, 62: 125-134.
- [13] GOTTUMUKKAL R, ASARI V K. An improved face recognition technique based on modular PCA approach[J]. Pattern Recognition Letters, 2004, 25(4): 429-436.
- [14] CHEN S, LIU J, ZHOU Z H. Making FLDA applicable to face recognition with one sample per person[J]. Pattern Recognition, 2004, 37(7): 1553-1555.
- [15] ABDI H, WILLIAMS L J. Principal component analysis[J]. Wiley Interdisciplinary Reviews: Computational Statistics, 2010, 2(4): 433-459.
- [16] IZENMAN A J. Linear discriminant analysis[M]//Modern multivariate statistical techniques. New York, NY: Springer, 2013: 237-280.
- [17] SCHÖLKOPF B, SMOLA A, MÜLLER K R. Kernel principal component analysis[C]//International conference on artificial neural networks. Berlin, Heidelberg: Springer, 1997: 583-588.
- [18] BAUDAT G, ANOUAR F. Generalized discriminant analysis using a kernel approach[J]. Neural Computation, 2000, 12(10): 2385-2404.