

多层融合深度局部 PCA 子空间 稀疏优化特征提取模型

胡正平, 陈俊岭

(燕山大学信息科学与工程学院, 河北秦皇岛 066004)

摘 要: 子空间方法是主要利用全局信息的经典模式识别方法, 随着深度学习思想的引入, 局部自学习结构特征模型得到大家的关注. 利用深度学习原理, 本文提出一种多层融合的深度局部子空间稀疏优化特征自学习抽取模型解决目标识别问题. 首先, 对训练样本集通过最小化重构误差得到第一层的主成分 (Principal Component Analysis, PCA) 特征映射矩阵; 然后, 通过 L_1 范数约束对特征映射结果进行稀疏优化, 提高算法鲁棒性. 接着, 在第二层映射层以第一层的特征输出为输入, 进行同样的特征矩阵学习操作, 最终将图像映射至深层 PCA 子空间; 然后, 对各个映射层的特征提取结果进行加权融合, 进行二值化哈希编码和直方图分块编码, 提取图像的深度子空间稀疏特征. 在 FERET、AR、Yale 等经典人脸数据库以及 MNIST、CIFAR-10 等目标数据库上的实验结果表明, 该算法可以取得较高的识别率以及较好的光照、表情、人脸朝向鲁棒性, 并且相对于卷积神经网络等深度学习框架具有结构简洁、收敛速度快等优点.

关键词: 深度学习; 多层融合; 子空间; 稀疏优化

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112 (2017)10-2383-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2017.10.011

Feature Extraction Model Based on Multi-layered Deep Local Subspace Sparse Optimization

HU Zheng-ping, CHEN Jun-ling

(School of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066004, China)

Abstract: Subspace method is classical pattern recognition method, that uses global information mainly to denote an image. Recently, with the introduction of deep learning, the feature extraction model based on local self-learning has attracted more and more attention. By using the theory of deep learning, this paper presents a new feature extraction model based on multi-layered deep local subspace sparse optimization to solve the problem of object recognition. Firstly, we calculate the PCA mapping matrix on the first layer by minimizing the reconstruction error on the training sample set, then we optimize the feature mapping results through L_1 norm to enhance the robustness of algorithm. Secondly, we use the output of the first layer as the input of second layer, then we implement same actions of feature learning. In this way we can map the image to deep PCA subspace. Finally we merge these feature extraction results from different layers with weighting and encode the merged feature with binary hash code and histogram segment code. After that, we obtain the multi-layered deep local subspace sparse feature. The experimental results on face database of FERET, AR, Yale and target database of MNIST, CIFAR-10 show that this feature extraction model can achieve high recognition rate and robustness for illumination, expression and pose. At the same time, compared with the convolutional neural networks, our algorithm owns the advantages of simple structure and fast convergent rate.

Key words: deep learning; multi-layered fusion; subspace; sparse optimization

1 引言

人脸识别技术是模式识别中最具挑战性的研究方向之一,在身份认证、安保监控方面有着广泛应用.人脸识别的发展主要分为三个阶段,从最开始的基于子空间的识别算法,到浅层机器学习算法,再到深度学习人脸识别方法.成熟的子空间人脸识别方法主要包括 Turk 等^[1]提出的 Principal Component Analysis (PCA) 特征脸 (Eigenface) 方法和 Belhumeur 等^[2]提出的 Fisher 脸 (FisherFace) 方法,奠定了代数特征人脸识别的理论基础.之后兴起的以 Back Propagation (BP) 神经网络^[3]、Support Vector Machine (SVM) 支持向量机和稀疏表示^[4]为代表的浅层机器学习人脸识别算法,使得人们在像素数据层面对图像有了更深层次的理解.2006 年 Hinton 等^[5]提出多隐层神经网络具有更为优异的特征学习能力,并且其在网络训练上的难度可以通过“逐层初始化”的方式来降低,开启了深度学习的大门.近年来随着深度学习研究的深入,越来越多的学者试图将深度学习与子空间方法相结合,其中 Chan 等^[6]提出的 PCA Network (PCANet) 方法,为子空间深度化指明了方向.

卷积神经网络 (Convolutional Neural Network, CNN) 作为目前应用最为广泛的深度学习模型,所提取的图像深层特征要优于 Local Binary Patterns (LBP)、Histogram of Oriented Gradients (HOG)^[7]、Gabor^[8] 等人构建的图像特征,其在人脸识别方面的表现也相当出色.

在 2012 年的 ImageNet 大赛上基于 CNN 的深度学习模型在两个比赛中均取得第一名^[9],并超出第二名至少 10%. 2014 年 Taigman 等^[10]提出的 DeepFace 算法,在 LFW 人脸数据库上达到 97% 以上的正确率.之后由 Sun 等^[11-13]提出的 DeepID 算法更是将 CNN 人脸认证的准确率提高到了 99.75%. 卷积神经网络虽然准确率,但其参数众多,其参数调整过程不仅复杂而且需要依赖大量的先验知识,收敛速度缓慢,为此研究人员试图寻找更为简单有效的深度结构框架. Lu 等^[14]将卷积神经网络中的映射核改为加权 PCA (WPCA) 映射矩阵,采用生成码本的方式融合得到最终的特征向量,提高了网络收敛性能. Liong 等^[15]提出的深度 PCA (Deep PCA, DPCA) 人脸识别算法,通过构建双层 PCA 映射网络,结合 ZCA 白化,提取深度子空间特征. Chan 等^[16]在 DPCA 基础上进行扩展,在 2DPCA^[17] 框架内,加入哈希编码和直方图分块提取,构建了 PCANet 算法,在人脸识别、手写字体识别、物体识别方面的识别率都达到 95% 以上.在此基础上, Huang 等^[18]将 PCANet 与线性回归分类器相结合,应用于人脸识别,在算法鲁棒性方面有所提高.

本文提出一种多层融合深度 PCA 子空间稀疏优化特征提取算法,在经典 PCA 特征脸方法中融入深度结构思想,将图像映射至深层 PCA 子空间,并将不同层次子空间特征映射结果进行加权融合,通过稀疏优化算法进行迭代以增加特征映射鲁棒性.在 FERET、AR、Yale 等人脸数据库上的实验表明,该方法在单训练样本条件下,其正确率与传统子空间识别方法相比有很大提高,在算法鲁棒性方面相比 PCANet 等新兴方法也有所改善.在 MNIST、CIFAR-10 等目标数据库上的实验结果表明该方法和卷积神经网络深度学习框架的识别率相当,但其结构简洁、收敛速度快.

2 深度 PCA 原理

2.1 DPCA 原理

深度学习结构模型在一定程度上与人脑分层信息处理机制^[19]相似, Sun Y 等^[13]通过对卷积神经网络的研究,从数据层面阐述深度分层特征模型在特征提取过程中所具有的稀疏性、特征选择性和特征鲁棒性,提出特征多层映射能够更好的对图像语义进行描述.因此越来越多的研究人员致力于寻找新的有效的多层特征模型, Liong 等^[15]提出的深度 PCA (DPCA) 算法属于子空间深度化改造的典型算法.

DPCA 算法流程图如图 1 所示.在第一层对输入图像首先进行 ZCA 白化处理,去除特征之间相关性,降低冗余,然后进行 PCA 映射,提取输入图像的主成分信息.在第二层中,以上一层的特征输出为输入,进行同样操作,得到深层映射后的主成分特征.最终,将两层特征映射结果进行加权融合,送入分类器进行训练,在单训练样本条件下使用欧氏距离分类器进行预测.

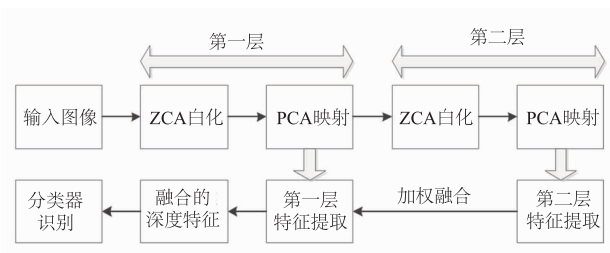


图1 DPCA原理框图

2.2 PCANet 模型

2015 年 Chan 等^[16]提出的 PCANet 识别算法是深度子空间模型的又一典型算法.它在 DPCA 算法的基础上进行了大量扩充,以单向的 2DPCA^[17] 为核心,在完成双层特征映射之后,对输入特征进行直方图分块特征提取,然后经过哈希编码,得到鲁棒性更强的深度特征,具体流程如图 2 所示.

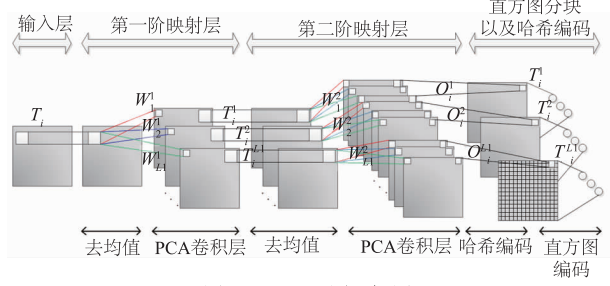


图2 PCANet原理框图

(1) 输入层

假设有 N 个训练样本 $\{T_i\}_{i=1}^N$, 样本尺寸为 $m \times n$, PCA 滤波器尺寸为 $k_1 \times k_2$, 首先, 将第 i 个训练样本分块, 减去均值之后转换成向量形式:

$$\begin{aligned} \bar{X}_i &= [\bar{x}_{i,1}, \bar{x}_{i,2}, \dots, \bar{x}_{i,mn}] \\ \text{s. t. } \bar{x}_{i,j} &= x_{i,j} - \frac{1}{k_1 k_2} \mathbf{1}^T \mathbf{x}_{i,j} \mathbf{1} \end{aligned} \quad (1)$$

此时输入训练样本可以表示为向量集形式:

$$\mathbf{X} = [\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N] \in \mathbf{R}^{k_2 \times k_1 \times Nmn} \quad (2)$$

(2) 第一映射层

首先, 对输入的数据矩阵 \mathbf{X} 通过最小化重构误差来得到 PCA 滤波器核:

$$\begin{aligned} \min_{V \in \mathbf{R}^{k_1 \times k_2}} \|\mathbf{X} - \mathbf{V}\mathbf{V}^T \mathbf{X}\|_F^2 \\ \text{s. t. } \mathbf{V}^T \mathbf{V} = \mathbf{I}_{L_1} \end{aligned} \quad (3)$$

训练得到的第一层滤波器核为:

$$\begin{aligned} \mathbf{W}_l^1 &= \text{mat}_{k_1, k_2}(\mathbf{q}_l(\mathbf{X}\mathbf{X}^T)) \in \mathbf{R}^{k_1 \times k_2}; \\ l &= 1, 2, \dots, L_1 \end{aligned} \quad (4)$$

其中 $\mathbf{q}_l(\mathbf{X}\mathbf{X}^T)$ 代表 $\mathbf{X}\mathbf{X}^T$ 的第 l 个特征向量。

(3) 第二映射层

第二层, 输入数据为前一层映射输出结果:

$$T_i^l = T_i * W_l^1; \quad i = 1, 2, \dots, N \quad (5)$$

同样, 在第二层中对输入数据去均值后进行 PCA 分析, 得到该层的 PCA 卷积核:

$$\begin{aligned} \mathbf{W}_l^2 &= \text{mat}_{k_1, k_2}(\mathbf{q}_l(\mathbf{Y}\mathbf{Y}^T)) \in \mathbf{R}^{k_1 \times k_2}; \\ l &= 1, 2, \dots, L_2 \end{aligned} \quad (6)$$

通过 PCA 卷积核 \mathbf{W}_l^2 将输入数据映射至深层子空间, 得到深层子空间主成分特征:

$$O_i^l = \{T_i^l * W_l^2\}_{l=1}^{L_2} \quad (7)$$

(4) 输出层

在输出层, 对提取到的特征进行二值化哈希编码:

$$T_i^l = \sum_{l=1}^{L_2} 2^{l-1} \mathbf{H}(T_i^l * W_l^2) \quad (8)$$

其中 $\mathbf{H}(\cdot)$ 为类 Heaviside 阶跃函数, 去除特征值中的负值部分。编码完成后, 对得到的特征图进行直方图分块处理, 得到最终的特征输出:

$$f_i = [\text{Bhist}(T_i^1), \dots, \text{Bhist}(T_i^{L_1})]^T \in \mathbf{R}^{(2^{L_1})^{L_2}} \quad (9)$$

3 多层融合深度局部子空间稀疏优化特征提取模型

3.1 层间融合深度结构扩展

深度结构模型本质上就是将信号在多层次模型中进行逐层映射, 在每一层都进行相应的映射特征提取。实验表明^[13,16,19], 在深层模型中, 底层特征多为目标结构信息, 顶层特征则多为目标抽象语义信息, 因此将各个层面的特征进行融合来进行特征表达, 将比单纯使用顶层特征更为全面而准确。Wang^[20]等在通过卷积神经网络解决年龄分类问题过程中, 同样采取了层间特征融合手段, 对深度结构中每层映射结果通过 PCA 降维进行下采样, 融合得到多层深度特征。

如图3所示, 在本文所提出的算法中, 同样借鉴了这种特征级联思想, 对第一层和第二层子空间信息进行加权融合, 得到更为全面的人脸主成分信息。在特征提取过程中, 输入图像为 \mathbf{X} , 尺寸为 $a \times b$, 第一层的 PCA 滤波器个数为 L_1 , 第二层的滤波器个数为 L_2 , 第一层的 PCA 映射矩阵可表示为 $\mathbf{W}_l^1 (l = 1, 2, \dots, L_1)$, 第二层映射矩阵可表示为 $\mathbf{W}_l^2 (l = 1, 2, \dots, L_2)$, 则第一层的映射子空间特征可表示为:

$$\mathbf{F}_m^1 = \{\mathbf{X} * \mathbf{W}_l^1\}_{l=1}^{L_1} \in \mathbf{R}^{a \times b}; m = 1, 2, \dots, L_1 \quad (10)$$

第二层映射子空间特征可表示为:

$$\begin{aligned} \mathbf{F}_n^2 &= \{\{\mathbf{F}_m^1 * \mathbf{W}_l^2\}_{m=1}^{L_1}\}_{l=1}^{L_2} \in \mathbf{R}^{a \times b}; \\ n &= 1, 2, \dots, L_1 \times L_2 \end{aligned} \quad (11)$$

将两层特征进行加权融合, 得到融合的深度特征 \mathbf{F} :

$$\mathbf{F} = [\mathbf{F}_1^1 \times \lambda, \dots, \mathbf{F}_{L_1}^1 \times \lambda; \mathbf{F}_1^2, \dots, \mathbf{F}_{L_1 \times L_2}^2] \quad (12)$$

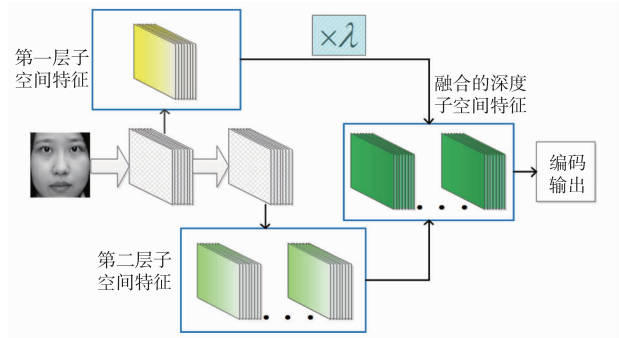


图3 深度子空间特征融合

3.2 子空间 2DPCA L_1 范数稀疏优化

人脑视觉皮层在处理图像信号过程中具有结构化、层次化稀疏性^[19], 同样深度学习结构在信号响应的过程中也表现出这种特性。实验证明^[13], 深度神经网络的神经元响应稀疏性主要表现在两个方面: 一是对于单张图片样本, 只有半数神经元进行激活响应; 二是对于每个神经元, 只会响应半数图片样本。因此人们在构

建深度神经网络时会着重对其稀疏性进行约束,其中具有代表性的有稀疏滤波(Sparse Filtering)算法^[21]以及线性修正单元^[22](Rectified Linear Units, ReLU)激活函数算法.在此基础上,人们对子空间方法也进行了相应的稀疏化改造,包括 Lu 等^[14]提出的稀疏加权 PCA(WPCA)算法和 Wang 等^[23]提出的 L1 范数约束的 2DPCA 算法.

在本文所提出算法中,同样借助 L1 范数对子空间单向 2DPCA 映射核进行稀疏优化,使得每层特征输出呈现出视觉稀疏性.假设有 K 个训练样本图像 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K]$,单向 2DPCA 目的在于找出最优投影矩阵 \mathbf{U} 以及投影系数 $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_K]$,通过优化以下目标函数:

$$\min_{\mathbf{U}, \mathbf{Z}} \frac{1}{K} \sum_{i=1}^K \|\mathbf{X}_i - \mathbf{U}\mathbf{Z}_i\|_F^2 \quad (13)$$

由于在人脸识别中,不可避免的要引入光照、表情等干扰因素,导致重构系数 \mathbf{Z} 丧失原有的稀疏性,因此需要引入一个额外的稀疏误差矩阵 \mathbf{E} 参与重构:

$$\mathbf{X} \approx \mathbf{U}\mathbf{Z} + \mathbf{E} \quad (14)$$

重新构造目标函数 $L(\mathbf{Z}, \mathbf{E})$:

$$L(\mathbf{Z}, \mathbf{E}) = (1/2) \|\mathbf{X} - \mathbf{U}\mathbf{Z} - \mathbf{E}\|_F^2 + \lambda \|\mathbf{E}\|_1 \quad (15)$$

优化求解最优的 \mathbf{Z} 和 \mathbf{E} :

$$\begin{aligned} \min_{\mathbf{Z}, \mathbf{E}} L(\mathbf{Z}, \mathbf{E}) \\ \text{s. t. } \mathbf{U}^T \mathbf{U} = \mathbf{I} \end{aligned} \quad (16)$$

这里 \mathbf{X} 代表训练样本数据矩阵, \mathbf{Z} 代表子空间投影系数, \mathbf{E} 为误差矩阵, λ 为规则化参数,通过 l_1 范数约束,可以求出具有最佳稀疏性的误差矩阵 \mathbf{E}_{opt} 和 \mathbf{Z}_{opt} ,这里通过迭代来完成优化.

第一步:固定 \mathbf{E}_{opt} ,此时目标函数可表示为 \mathbf{Z} 的函数 $J(\mathbf{Z})$:

$$J(\mathbf{Z}) = (1/2) \|\mathbf{X} - \mathbf{E}_{\text{opt}} - \mathbf{U}\mathbf{Z}\|_F^2 \quad (17)$$

其导数为:

$$\begin{aligned} \frac{\partial J(\mathbf{Z})}{\partial \mathbf{Z}} &= \frac{\partial \left[\frac{1}{2} \text{tr}(\mathbf{Z}^T \mathbf{Z}) - \text{tr}[\mathbf{U}\mathbf{Z}(\mathbf{X} - \mathbf{E}_{\text{opt}})^T] \right]}{\partial \mathbf{Z}} \\ &= \mathbf{Z} - \mathbf{U}^T (\mathbf{X} - \mathbf{E}_{\text{opt}}) \end{aligned} \quad (18)$$

令导数为零,求得全局最优值 \mathbf{Z}_{opt} :

$$\mathbf{Z}_{\text{opt}} = \mathbf{U}^T (\mathbf{X} - \mathbf{E}_{\text{opt}}) \quad (19)$$

第二步:固定 \mathbf{Z}_{opt} ,则目标函数可以表示为:

$$J(\mathbf{E}) = (1/2) \|\mathbf{E} - (\mathbf{X} - \mathbf{U}\mathbf{Z}_{\text{opt}})\|_2^2 + \lambda \|\mathbf{E}\|_1 \quad (20)$$

解这个凸优化问题,可得全局最优值 \mathbf{E}_{opt} :

$$\begin{aligned} \mathbf{E}_{\text{opt}} &= S_\lambda (\mathbf{X} - \mathbf{U}\mathbf{Z}_{\text{opt}}) \\ \text{s. t. } S_\tau(x) &= \text{sgn}(x) \cdot (|x| - \tau) \end{aligned} \quad (21)$$

循环迭代第一步和第二步,得到优化的重构系数 \mathbf{Z}_{opt} 和误差矩阵 \mathbf{E}_{opt} ,在本文实验中设计迭代次数为 10 次,得到的重构系数 \mathbf{Z}_{opt} 即为子空间特征.

3.3 多层融合深度局部子空间稀疏优化算法流程

多层融合深度子空间稀疏优化深度特征提取流程如图 4 所示.对于一个双层深度子空间模型,假设第一层的映射核个数为 L_1 ,第二层的映射核个数为 L_2 ,输入图像为 \mathbf{X} .在第一层映射过程中,对训练样本通过最小化重构误差得到第一层的 PCA 映射矩阵 $\{\mathbf{U}_1^l\}_{l=1}^{L_1}$,然后针对子空间重构系数进行迭代优化,迭代 10 次之后得到第一层子空间稀疏特征 $\{\mathbf{Z}_1^l\}_{l=1}^{L_1}$.在第二层中以第一层的特征输出 $\{\mathbf{Z}_1^l\}_{l=1}^{L_1}$ 为输入,再次进行最小化重构和稀疏映射,得到第二层子空间稀疏特征 $\{\mathbf{Z}_2^l\}_{l=1}^{L_2}$,然后再将前后两层得到的映射结果进行加权融合,得到深度稀疏子空间特征 $\mathbf{Z} = \{\mathbf{Z}_1^1 \times \lambda, \mathbf{Z}_1^2 \times \lambda, \dots, \mathbf{Z}_1^{L_1} \times \lambda; \mathbf{Z}_2^1, \mathbf{Z}_2^2, \dots, \mathbf{Z}_2^{L_2}\}$,再经过哈希编码和直方图分块编码,得到最终的鲁棒深度 PCA 子空间特征.

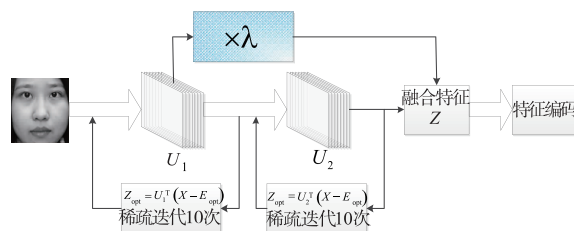


图4 多层融合的深度子空间稀疏优化算法

4 实验仿真

本文在 FERET、Yale、AR 等人脸数据库上评估算法在人脸识别方面的准确性和鲁棒性,在 CIFAR-10、Minist 等数据库上验证算法在目标识别方面的性能.在进行人脸识别实验时采用欧氏距离分类器,在单训练样本的条件下进行识别;在目标识别方面测试时采用支持向量机作为模式识别分类器.采用双层网络结构,第一层映射核个数 $L_1 = 8$,第二层映射核个数 $L_2 = 8$,映射核的尺寸均为 7×7 ,采用半覆盖式分块采样,迭代次数为 10 次,迭代收敛误差 $\tau = 0.05$.本文采用 CMU_PIE 人脸数据库作为训练样本对特征提取模型进行训练,CMU_PIE 人脸数据库包含来自 68 人的 41 368 张人脸图片,包含不同的姿态、光照、表情变化.这里从中选取 1 万张相对规范的人脸样本进行训练.实验证明多层融合深度子空间稀疏优化算法由于在特征训练和提取过程中针对噪声干扰进行了稀疏优化,因此其在光照、表情方面的鲁棒性相对于 PCANet 算法有一定提升,同时在正确率方面也远远超过传统的单层子空间识别方法.

4.1 光照、表情人脸识别

为验证算法在面对光照、表情、人脸朝向等干扰因素时的鲁棒性,选用 FERET 和 AR 两个人脸库进行仿真实验.本文用到 FERET 数据库子集共 200 人,每人 7

张图片,包括 1 张正面人脸及 6 张不同朝向人脸,其中 1 张具有表情因素干扰,1 张具有光照因素干扰,图片大小为 64×64 . AR 数据库子集共 100 人,每人 7 张图片,包括 1 张正面人脸图像及 6 张具有表情、光照变化的人脸图像,图片大小 64×64 . 采用单训练样本模式进行实验,将人脸数据库中每个人正面中性表情均匀光照人脸图片作为训练样本,其余图片作为测试样本,依据测试样本和训练样本之间深层子空间欧式距离进行分类. 图 5 给出 FERET 数据库中的部分人脸图像,图 6 给出 AR 数据库中的部分人脸图像,表 1 给出在这两个数据库子集上本文算法与最新的 PCANet 算法在识别率方面的对比情况.



图5 FERET人脸数据库示例图像



图6 AR人脸数据库示例图像

表 1 存在表情、朝向、光照干扰情况下两种方法的识别率比较

数据库	干扰类型	样本总数	PCANet 算法识别率	本文算法识别率
AR	光照、表情	600	95.83%	99.17%
FERET	光照、人脸朝向	1200	70.83%	75.00%

在原文中^[16],PCANet-2 算法在 FERET 数据库上达到 97.26% 的准确率,这主要是由于 Chan 等所构建的 PCANet-2 算法是在大型人脸数据库 MultiPIE 上进行训

练的,MultiPIE 人脸数据库包含数十万张人脸训练样本,而本文实验由于条件限制,选择在 CMU_PIE 人脸数据库中的 1 万张训练样本进行试验,使得学习得到的特征映射矩阵性能相对于原文中的算法模型会有所降低. 不过实验证明在同等训练水平下,本文提出的多层融合深度子空间稀疏优化人脸识别算法由于在特征训练和提取过程中针对存在的干扰噪声进行了稀疏优化,因此其相对于之前的 PCANet 深度学习框架,在存在光照、表情、人脸朝向干扰的情况下,识别率提高了 4% ~ 5% 左右,充分说明本文提出的算法具有较好的鲁棒性,能够在 PCANet 算法的基础上更好的应对干扰问题.

4.2 深度化模型-单层局部模型对比实验

深度学习理论指出多隐层神经网络相对于单一结构网络系统具有更好的特征提取能力. 很多学者都尝试对传统子空间方法进行深度化改造,将其扩展为两层甚至多层网络,具有代表性的包括 Chan 等^[16]提出的 PCANet 算法以及 Lu 等^[14]提出的联合特征学习 (Joint Feature Learning, JFL) 人脸识别算法. 与经典的卷积神经网络等深度学习框架不同,本文所提出算法也属于单层特征提取算法多层堆叠的深度化改造,为了验证分层扩展后的特征提取性能,我们在 FERET 数据库、AR 数据库、Yale B 数据库上进行仿真实验,对比经过深度化改造的特征提取模型与传统特征提取模型的性能,为了着重比较模型特征提取能力,需要简化人脸识别过程中用到的分类器,因此采用最简单的欧氏距离分类器,使得最终识别效果主要取决于模型所提取特征对图像语义的表达能. 表 2 给出各个方法之间的实验对比结果.

实验证明,在同等硬件条件、均使用欧氏距离分类器的情况下,PCANet 算法以及本文提出的深层 PCA 子空间方法都要优于以往的单层全局以及局部特征提取方法. 尤其是在处理 FERET 这种存在较为严重的光照、人脸朝向干扰的人脸数据库时,经典的单层特征提取手段在欧氏距离分类器条件下只能达到 50% 左右的识别率. 可见深层特征提取模型提取的层次化子空间稀疏特征能够更好的对原始图像语义进行描述.

表 2 本文方法与经典特征模型性能对比分析

特征类型	FERET 数据库(1200 张)		AR 数据库(600 张)		Yale B 数据库(2376 张)		平均识别率(%)
	样本数	识别率(%)	样本数	正确率(%)	样本数	正确率(%)	
Gabor	623	51.92%	465	77.50%	1890	80.29%	69.90%
LBP	701	58.41%	532	88.67%	1972	83.00%	76.69%
HOG	487	40.58%	424	70.67%	1683	70.83%	60.78%
Fisher	802	66.83%	560	93.33%	2016	84.85%	81.67%
PCA	779	64.92%	543	90.50%	2071	87.16%	80.86%
PCANet-2	850	70.83%	575	95.83%	2150	90.49%	85.72%
本文算法	900	75.00%	595	99.17%	2249	94.65%	89.61%

4.3 深度化模型-卷积神经网络对比实验

本文提出的算法不仅在处理单训练样本人脸识别问题上表现良好,其在多训练样本目标识别方面也具有优异特征学习能力. 在手写数字数据库 MNIST 和目标识别数据库 CIFAR-10 上进行仿真实验,将本文算法与卷积神经网络算法进行对比,评估算法在准确率、时间消耗等方面的性能.

MNIST 数据库是深度学习模型测试时常用数据库,主要采集美国中学生手写的数字样本,如图 7 所示. 这里选用 60000 张图片作为训练样本,10000 张样本作为测试样本. 类似的 CIFAR-10 是一个广泛应用于深度学习训练的目标识别库,包含有 10 种目标,70000 张图片,如图 8 所示,这里同样选择灰度化之后的 60000 张图片作为训练样本,10000 张图片作为测试样本. 在 MNIST 数据库和 CIFAR-10 数据库上进行仿真实验时,先在训练样本的基础上学习得到深度特征提取模型,然后再对训练样本进行特征提取,将提取到的特征送入支持向量机中进行训练,测试过程中也是先提取测试样本的深度子空间稀疏特征,然后送入支持向量机中进行分类预测. 表 3 给出了卷积神经网络与本文算法在以上两个数据库上的仿真结果.

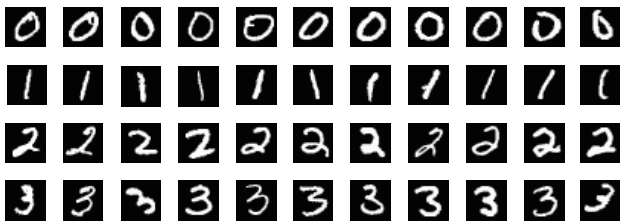


图7 MNIST数据库样本示例

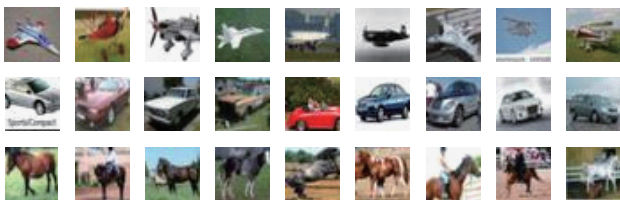


图8 CIFAR-10数据库样本示例

表3 CNN 与本文算法性能对比分析

数据库	卷积神经网络		本文算法	
	识别率 (%)	时间消耗 (min)	识别率 (%)	时间消耗 (min)
MNIST	99.15%	63	99.00%	10
CIFAR-10	81.93%	306	78.81%	60

实验证明,在同等硬件条件下(CPU 主频 3.10GHz,内存 64G),本文所提出算法在 CIFAR-10 数据库上的识别率相对于卷积神经网络降低了大约 3%,但

训练和识别的时间消耗只为卷积神经网络的五分之一;类似的在 MNIST 数据库上本文算法识别率相对于卷积神经网络只降低了 0.15%,但模型训练的时间损耗为原来的六分之一. 不仅如此,在可调参数数量、模型收敛速度等方面本文算法也要优于卷积神经网络,充分证明了多层融合深度子空间稀疏优化算法的特征提取性能.

5 结论

受深度学习理论启发,本文重点研究经典子空间识别方法经过深度化改造之后的特征提取能力,对 PCA 算法进行深度分层扩展之后将其用于图像深度特征提取. 于此同时为了解决人脸图像识别中存在的光照、表情、朝向等干扰问题,增加算法的稳定性,向其中加入稀疏优化算法,通过对特征映射矩阵进行迭代优化来增强其映射的鲁棒性. 在 FERET、Yale、AR 等人脸数据库上的实验表明,该算法在单训练样本、只使用欧氏距离分类器进行分类的情况下,对于存在光照、表情、人脸朝向干扰的数据库仍能够达到较高的识别率,并且相对于单层局部特征提取模型具有更高程度的抽象语义特征提取能力;在 MNIST 和 CIFAR-10 等目标识别数据库上的实验表明在多训练样本、借助支持向量机分类的情况下,该方法与卷积神经网络的识别率相当,并且在时间消耗、模型复杂度、模型收敛速度方面要优于卷积神经网络.

参考文献

- [1] TURK M, PENTLAND A. Eigenfaces for recognition [J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86.
- [2] BELHUMEUR P N, HESPANHA J P, KRIEGMAN D. Eigenfaces vs Fisherfaces: recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720.
- [3] LI Yong-qiang, PAN Jin. Face recognition algorithm based on improved BP neural network [J]. International Journal of Security and Its Applications, 2015, 9(5): 175-184.
- [4] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210-227.
- [5] HINTON G, SALAKHUTDINOV R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504-507.
- [6] CHAN Tsung Han, JIA Kui, GAO Shenghua, LU Jiwen, ZENG Zinan, MA Yi. PCANet: A simple deep learning baseline for image classification [J]. IEEE Transactions on Image Processing, 2015, 24(12): 5017-5032.

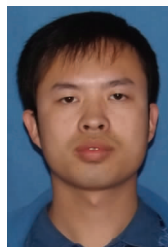
- [7] TAN Hengliang, YANG Bing, MA Zhengming. Face recognition based on the fusion of global and local HOG features of face images [J]. *Computer Vision*, 2014, 8(3): 224 – 234.
- [8] LIU Chengjun, WECHSLER H. Independent component analysis of Gabor features for face recognition [J]. *IEEE Transactions on Neural Networks*, 2003, 14(4): 919 – 928.
- [9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [A]. *Advances in Neural Information Processing Systems* [C]. Lake Tahoe, NV, United states: NIPS, 2012. 1097 – 1105.
- [10] TAIGMAN Y, YANG Ming, RANZATO M, WOLF L. DeepFace: Closing the gap to human-level performance in face verification [A]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [C]. Columbus, OH, United states: CVPR, 2014. 1701 – 1708.
- [11] SUN Yi, WANG Xiaogang, TANG Xiaoou. Deep learning face representation from predicting 10,000 classes [A]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [C]. Columbus, OH, United states: CVPR, 2014. 1891 – 1898.
- [12] SUN Y, CHEN Y, WANG X, et al. Deep learning face representation by joint identification-verification [A]. *Advances in Neural Information Processing Systems* [C]. Montreal, Canada: NIPS, 2014. 1988 – 1996.
- [13] SUN Yi, WANG Xiaogang, TANG Xiaoou. Deeply learned face representations are sparse, selective, and robust [A]. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [C]. Boston, MA, USA: CVPR, 2015. 2892 – 2900.
- [14] LU Jiwen, LIONG V E, WANG Gang, MOULIN P. Joint feature learning for face recognition [J]. *IEEE Transactions on Information Forensics and Security*, 2015, 10(7): 1371 – 1383.
- [15] LIONG V E, LU J, WANG G. Face recognition using Deep PCA [A]. *IEEE Transactions on Information, Communications and Signal Processing (ICICS)* [C]. Beijing, China: ICICS, 2013. 1 – 5.
- [16] CHAN T H, JIA K, GAO S, et al. PCANet: A simple deep learning baseline for image classification? [J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 5017 – 5032.
- [17] YANG J, ZHANG D, FRANGI A F, et al. Two-dimensional PCA: a new approach to appearance-based face representation and recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(1): 131 – 137.
- [18] HUANG J, YUAN C. Weighted-PCANet for Face Recognition [A]. *Neural Information Processing* [C]. Montréal, Quebec, Canada: Springer International Publishing, 2015. 246 – 254.
- [19] 焦李成, 赵进, 杨淑媛, 刘芳, 谢雯. 稀疏认知学习、计算与识别的研究进展 [J]. *计算机学报*, 2015, 38(10): 1 – 18.
JIAO Li-cheng, ZHAO Jin, YANG Shu-yuan, LIU Fang, XIE Wen. Research advances on sparse cognitive learning, computing and recognition [J]. *Chinese Journal of Computers*, 2015, 38(10): 1 – 18. (in Chinese)
- [20] WANG Xiaolong, GUO Rui, KAMBHAMETTU C. Deeply-learned feature for age estimation [A]. *IEEE Winter Conference on Applications of Computer Vision (WACV)* [C]. Waikoloa, HI, United states: IEEE, 2015. 534 – 541.
- [21] NGIAM Jiquan, KOH Pangwei, Chen Zhenghao, BHASKAR Sonia, NG ANDREW Y. Sparse filtering [A]. *Advances in Neural Information Processing Systems* [C]. Granada, Spain: NIPS, 2011. 1125 – 1133.
- [22] DAHL G E, SAINATH T N, HINTON G E. Improving deep neural networks for LVCSR using rectified linear units and dropout [A]. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* [C]. Vancouver, BC, Canada: IEEE, 2013. 8609 – 8613.
- [23] WANG D, LU H. Object tracking via 2DPCA and L1-regularization [J]. *Signal Processing Letters*, 2012, 19(11): 711 – 714.

作者简介



胡正平 男, 1970 年生于四川仪陇县, 博士、教授、博士生导师, 中国电子学会高级会员, 中国图像图形学会高级会员, 目前研究方向为稀疏表示、模式识别。

E-mail: hzp@ysu.edu.cn



陈俊岭 男, 1991 年生于河北唐山, 燕山大学信息科学与工程学院信息与通信工程专业硕士研究生, 主要研究方向为深度学习分类。

E-mail: 1095988845@qq.com