

一种基于 Curv-SAE 特征融合的人脸降维和识别方法

张志禹 刘思媛

(西安理工大学自动化与信息工程学院 西安 710048)

摘要 相比于传统的降维算法,深度学习中的栈式自编码器(Stacked Autoencoder,SAE)能够有效地学习特征并实现高效降维,然而对输入特征极其敏感。第二代离散曲波变换(Discrete Curvelet Transform,DCT)能够提取出人脸的各向信息(包含边缘和概貌特征),确保 SAE 的输入特征充分,从而弥补了其不足。因此,提出了一种基于 Curv-SAE 特征融合的人脸识别降维算法,即对人脸图像进行 DCT 得到特征脸并将其作为 SAE 的输入特征进行训练,特征融合后将其输入到分类器中进行识别。在 ORL 和 FERET 人脸数据库上的实验表明,与小波变换相比,曲波的特征信息更丰富;与传统的降维算法相比,SAE 的特征表达更充分且识别精度更高。

关键词 深度学习,人脸识别,第二代离散曲波变换,栈式自编码器,降维

中图分类号 TP391.41 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.10.049

Method of Face Recognition and Dimension Reduction Based on Curv-SAE Feature Fusion

ZHANG Zhi-yu LIU Si-yuan

(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract Compared with the traditional dimension reduction algorithm, stacked autoencoders (SAE) in deep learning can effectively learn the features and achieve efficient dimension reduction, but its performance depends on the input characteristics. The second generation discrete curvelet transform can extract the information of human faces, including edge and overview features, and ensure that the input features of SAE are sufficient, thus making up for the shortages of SAE. Therefore, a new recognition and dimension reduction algorithm based on Curv-SAE feature fusion was proposed. Firstly, the face images are processed by DCT to generate the Curv-faces, which are trained as input characteristics of SAE. And then different layers of features are used for the final classification of identification. Experimental results on ORL and FERET face databases show that the feature information of curvelet transform is more abundant than the wavelet transform. Compared with the traditional dimension reduction algorithms, the feature expression of SAE is more complete and the recognition accuracy is higher.

Keywords Deep learning, Face recognition, The Second generation discrete curvelet transform, Stacked autoencoders, Dimension reduction

1 概述

人脸识别是目前常用的特征识别技术中更为安全可靠的身份鉴定技术。它具有特征唯一、操作隐蔽性强、准确率高等优点,在金融、司法、网络信息安全、安保等应用领域中发挥着重要作用。

人脸识别实质上是特征提取和分类的过程,其中基于多尺度的分析算法是图像特征提取的常用算法。鉴于小波变换具有时频局限性和多分辨分析的良好特性,Cheng 等^[1]提出利用二维 Gabor 小波变换提取局部特征;Chien 等^[2]应用多尺度小波变换提取特征脸,实现了较高的识别率。然而小波变换存在图像边缘奇异性的问题,并不能有效地提取图像边缘信息,人脸图像的轮廓信息容易缺失,进而影响人脸识别的准确度。与小波变换相比,曲波变换在处理曲线上具有显著

的优势,它能够提取出曲线的细节特征,并且噪声含量较小,能够将人脸的重要信息更精确地稀疏表达。对曲线的提取如图 1 所示。

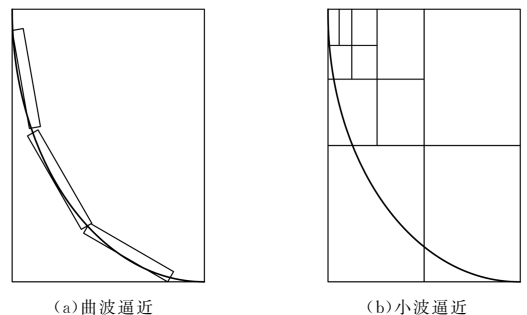


图 1 曲线逼近对比

Fig. 1 Contrast of curve approximation

到稿日期:2017-09-17 返修日期:2017-12-27 本文受国家自然科学基金资助重大项目(41390454)资助。

张志禹(1966—),男,博士,教授,主要研究方向为图像处理、阵列信号处理、模式识别,E-mail:songwz464237@126.com(通信作者);刘思媛(1993—),女,硕士生,主要研究方向为图像处理、深度学习、模式识别,E-mail:siyuanliu822@126.com。

高维数据制约着机器学习,因此数据降维方法成为模式识别领域的热门话题之一^[3]。人脸识别中,人脸原始图像具有高维特征向量,为了减少计算复杂度和运算时间,必须将高维数据映射到低维空间再进行分类和识别。经典的降维算法有主成分分析(PCA)^[4]、线性判别分析(LDA)^[5]和独立成分分析(ICA)^[6]等。当数据呈线性形式时,这些方法具有良好的降维效果,然而,当数据呈非线性的高维复杂形式时,这些方法往往会丢失原图像的重要信息,从而不能有效地降维^[7]。自2000年以来,流形学习被认为是非线性降维方法的一个分支,其中 Isomap^[8]和 LLE^[9]得到了迅速发展。虽然流形学习能将非线性结构的高维数据很好地映射到低维空间,然而大多数流形学习方法^[10]都存在不收敛的问题,当原始数据集增加或减少几个数据量时,得到的结果往往存在明显差异,鲁棒性较弱,且该方法对噪声敏感,因此在实际工程应用中很难大施拳脚。栈式自编码器(SAE)^[11]是深度学习的一个分支结构,它由多层稀疏自编码器组成,通过逐层训练的方式自我学习图像特征。自编码器的本质是神经网络,这种“类脑”的形式能够模拟大脑并捕捉到训练样本中最显著的特征。与此同时,SAE是欠完备自编码器,后一层的输入节点数比前一层的输出节点数少,这种特殊结构能够使高维数据在不丢失重要信息的前提下实现降维。2006年,Hinton等^[12]训练了一个深层自编码器并逐渐缩小隐层,在30个单元时达到极值。相比30维的PCA,其生成的编码所产生的重构误差更少,所学习到的特征更容易定性解释。因此,SAE具有优于其他方法的降维和特征表达效果。然而,它的性能优劣依赖于输入层的特征充分与否,如果特征表达不充分、噪声含量大,那么训练好的SAE模型并不理想。

综上所述,本文提出了一种基于特征脸下(Curv脸)SAE特征融合的人脸识别降维算法,将提取的特征作为SAE的输入,通过无监督逐层训练和有监督调优得到降维后的人脸图像特征;同时将多层特征融合,用softmax分类器进行分类识别。在ORL和FERET数据集上的实验结果表明,Curv脸能够充分地表达各向人脸特征,同时SAE有良好的学习特征和降维能力,多层特征融合比单层的识别精度更高。

2 Curvelet 变换

2.1 Curvelet 变换概述

Curvelet 变换最先由 Candes 等^[13]提出,它在很大程度上弥补了小波变换在处理图像边缘时存在的奇异性缺陷,能够准确地提取各个方向上的曲线信息,去噪性能也优于小波变换。

Curvelet 变换有两代演变过程。第一代曲波变换以脊波变换理论为基础,其过程主要是先对图像进行子带分解,然后将不同尺度的子带图像平滑分块并规范化,最后实现脊波变换。第一代曲波变换虽能很好地处理边缘信息缺失问题,但会产生较高的时间和空间复杂度,而且曲波金字塔的分解会带来很大的数据冗余量,参数多且复杂,不利于调试工作的进行。因此,Candes 等在2005年提出了第二代曲波变换^[14]。第二代曲波变换的方法结构完全有别于第一代,其在理论上与脊波并不存在直接联系,参数少且简单,同时细节处理的效果优于第一代,因此更适用于人脸识别算法。本文采用的即为第二代曲波变换。

2.2 离散 Curvelet 变换

在笛卡尔坐标系下,Curvelet 的半径窗为方形同心环,局部窗为“楔形”,如图2所示。

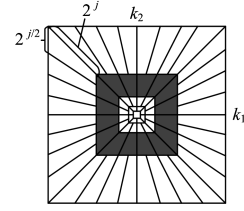


图2 离散 Curvelet 空间频域图

Fig. 2 Diagram of discrete Curvelet space frequency domain

采用一个带通半径窗函数 $\tilde{W}_j(\omega)$ 和一个角度窗函数 $V_j(\omega)$ 来表达此频域窗 $\tilde{U}_j(\omega)$,其表达式为:

$$\tilde{U}_j(\omega) = \tilde{W}_j(\omega)V_j(\omega) \quad (1)$$

其中,

$$\tilde{W}_j(\omega) = \sqrt{\Phi_{j+1}^2(\omega) - \Phi_j^2(\omega)}, j \geq 0 \quad (2)$$

$$V_j(\omega) = V(2^{-j/2}\omega_2/\omega_1) \quad (3)$$

Φ 是低通一维窗口的内积,满足:

$$\Phi_j(\omega_1, \omega_2) = \Phi(2^{-j}\omega_1)\Phi(2^{-j}\omega_2), 0 \leq \Phi \leq 1 \quad (4)$$

因此得到“楔形”附近的频率为:

$$2^j \leq \omega_1 \leq 2^{j+1}, -2^{[-j/2]} \leq \frac{\omega_1}{\omega_2} \leq 2^{[-j/2]} \quad (5)$$

引入斜率:

$$\tan \theta_l = l \cdot 2^{[-j/2]} \quad (6)$$

那么

$$\tilde{U}_{j,l}(\omega) = W_j(\omega)V_j(S_{\theta_l}\omega) \quad (7)$$

其中,剪切矩阵 $S_{\theta_l} = \begin{bmatrix} 1 & 0 \\ -\tan \theta_l & 1 \end{bmatrix}$,表示斜率是均匀分布的。

因此 Curvelet 变换的最终离散形式为:

$$c(j, l, k) = \int \hat{f}(S_{\theta_l}\omega)\tilde{U}_j(\omega)e^{i(b,\omega)}d\omega \quad (8)$$

本文采用第二代快速离散 Curvelet 变换,下面给出该算法在频域中的一种实现方法——USFFT 算法。

1)对笛卡尔坐标系下给定的二维函数做二维快速傅里叶变换(FFT),得到二维频域表示: $\hat{f}[n_1, n_2], n_1 \geq -n/2, n_2 < n/2$ 。

2)在频域中,对于每一对尺度和方向参量 (j, l) ,重采样 $\hat{f}[n_1, n_2]$ 得到采样值 $\hat{f}[n_1, n_2 - n_1 \tan \theta_l], (n_1, n_2) \in P_j$ 。其中 $P_j = \{(n_1, n_2) : n_{1,0} \leq n_1 < n_{1,0} + L_{1,j}, n_{2,0} \leq n_2 < n_{2,0} + L_{2,j}\}$ 且 $L_{1,j}$ 是关于 2^j 的参量, $L_{2,j}$ 是关于 $2^{j/2}$ 的参量,其分别表示窗函数 $\tilde{U}_j[n_1, n_2]$ 的支撑区间的长宽分量。

3)将内插后的 \hat{f} 与频域窗 \tilde{U}_j 相乘得到: $\hat{f}[n_1, n_2] = \hat{f}[n_1, n_2 - n_1 \tan \theta_l]\tilde{U}_j[n_1, n_2]$ 。

4)最后对 $\tilde{f}_{j,l}$ 经过二维FFT逆变换得到离散 Curvelet 系数 $c^D\{j\}$ 。

3 栈式自编码器(SAE)

自编码器是一种神经网络,包含输入层、隐藏层和输出重

构层,其主要含义是输出层重构出输入信息,使输出尽可能等于输入。其包括编码和解码两个过程。

编码是指把输入 $x \in R^m$ 映射到隐藏层表示 $h(x) \in R^n$ 的过程,计算过程如下:

$$h(x) = \sigma_h(Wx + b) \quad (8)$$

其中, $W \in R^{n \times m}$ 为编码权值矩阵; $b \in R^n$ 为编码偏置; $\sigma_h(x)$ 为映射函数,通常情况下取 sigmoid 函数或者 tanh 函数,最常用的是前者 $\sigma_h(x) = \frac{1}{1 + e^{-x}}$ 。

解码是指将隐层 $h(x)$ 再映射到输出层,以实现重构输入 x 的功能,其计算过程如下:

$$\phi = \sigma_\phi(W'h(x) + b') \quad (9)$$

其中, $W' \in R^{m \times n}$ 为解码权值矩阵; $b' \in R^m$ 为解码偏置; $\sigma_\phi(x)$ 为映射函数,同 $\sigma_h(x)$ 。

对于 M 个输入样本 (x^k, y^k) ($1 \leq k \leq M$), 输入为 x^k , 期望输出为 y^k , 实际输出为 ϕ^k , 其目标函数即重构误差损失函数为:

$$J_E(\omega, b) = \frac{1}{2} \sum_{k=1}^M \sum_{i=1}^N (\phi_i^k - y_i^k)^2 \quad (10)$$

稀疏自编码器增加了对隐层神经元激活的稀疏性约束,以便大多数隐层神经元处于非激活状态,则优化后的目标函数为:

$$\begin{aligned} J_{\text{Sparse}} &= J_E + \beta \sum_i KL(\rho \parallel \tilde{\rho}_i) \\ &= \frac{1}{2} \sum_{k=1}^M \sum_{i=1}^N (\phi_i^k - y_i^k)^2 + \beta \sum_i KL(\rho \parallel \tilde{\rho}_i) \end{aligned} \quad (11)$$

其中, ρ 是稀疏性参数,通常取为一个接近于 0 的正实数; $\tilde{\rho}_i$ 是第 i 个隐层神经元的平均激活值; β 为惩罚因子; $KL(\rho \parallel \tilde{\rho}_i)$ 为惩罚项,其表示概率分布 $\{\rho, 1 - \rho\}$ 和 $\{\tilde{\rho}_i, 1 - \tilde{\rho}_i\}$ 之间的 KL 散度,定义为:

$$KL(\rho \parallel \tilde{\rho}_i) = \rho \log \frac{\rho}{\tilde{\rho}_i} + (1 - \rho) \log \frac{1 - \rho}{1 - \tilde{\rho}_i} \quad (12)$$

栈式自编码器由多层稀疏自编码器组成,其结构如图 3 所示。其前一层稀疏自编码器的输出作为后一层稀疏自编码器的输入。根据栈式自编码器的输入层、隐藏层到输出层,把相邻的两层看作一个受限玻尔兹曼机(RBM),从最底层开始采用无监督学习算法(CD-K)逐层对全部 RBM 进行训练,训练当前层时其他层的参数保持不变,将这一过程称为“贪婪逐层无监督预训练”。通过这种有效的方式能够得到合适的初始网络参数值,再用有监督学习算法(梯度下降法)对网络参数进行整体调优,其参数更新法则为:

$$W^l \leftarrow W^l - \eta \frac{\partial J_{\text{Sparse}}}{\partial W^l} \quad (13)$$

$$b^l \leftarrow b^l - \eta \frac{\partial J_{\text{Sparse}}}{\partial b^l} \quad (14)$$

其中, η 为学习率。设层数为 L , 计算误差 δ_i^k 的表达式为:

$$\delta_i^k = (\phi^k - y^k) \circ \sigma_\phi(u_i^k) \quad (15)$$

其中, $u_i^k = W^l h_{i-1}^k + b^l$ 。计算误差 δ_i^k 的表达式为:

$$\delta_i^k = [(W^{l+1})^T \delta_{i+1}^k] \circ \sigma_\phi(u_i^k), 1 \leq l \leq L-1 \quad (16)$$

最后的偏导式为:

$$\begin{aligned} \frac{\partial J_{\text{Sparse}}}{\partial W^l} &= \sum_{k=1}^M \delta_i^k (h_{i-1}^k)^T \\ &, 1 \leq l \leq L \\ \frac{\partial J_{\text{Sparse}}}{\partial b^l} &= \sum_{k=1}^M \delta_i^k \end{aligned} \quad (17)$$

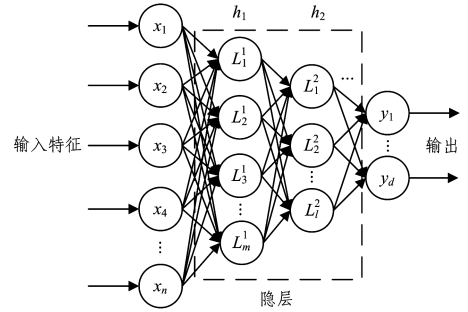


图 3 栈式自编码器的结构

Fig. 3 Structure of stacked autoencoder

经训练最终得到一个高效的数据降维结构模型。最后将降维后得到的有效特征输入到 softmax 分类器中进行分类和识别,从而得到较好的识别效果。

4 基于 Curv-SAE 特征融合的人脸降维识别

本文将提取的特征脸作为 SAE 的输入进行训练,逐层学习人脸特征,通过充分地学习特征来实现降维,从而进行人脸识别。相对于传统的特征提取方法和数据降维方法,该算法提取到的特征具有更丰富的曲线信息,如边缘轮廓、唇缘、眼睛周边细节等,同时 SAE 自主学习特征的方式能够捕捉更显著的特征。本文算法的具体步骤如下:

1) 对人脸库图像进行预处理,分别将 ORL 库的图像尺寸 112×92 和 FERET 库的图像尺寸 80×80 转换成 128×128 的统一尺寸,然后进行直方图均衡化,尽量去除光照等变化带来的影响。

2) 对预处理后的人脸库图像进行第二代快速离散 Curvelet 变换, 128×128 的图像被分为 3 个层次(coarse 层、中间 detail 层和 fine 层)。选取 $C^D\{1\}$ 作为曲波脸。

3) 将曲波脸作为 SAE 的输入,训练多层 SAE 网络模型。

4) 将带有标签的数据输入到 SAE 中得到期望输出,用梯度下降法微调网络。

5) 测试训练好的模型,将测试集输入到 SAE 中。

6) 将 SAE 的不同层特征融合,作为 softmax 分类器的输入,最后得到准确率 acc ,其计算公式为:

$$acc = \frac{\text{测试集标签} = \text{分类器输出类别}}{\text{总测试集数}} \times 100\%$$

5 实验结果与分析

本文仿真平台为 matlab R2014a。为了验证算法的有效性,分别在 ORL 人脸数据库和 FERET 数据库上进行实验。ORL 人脸库共有 40 个对象,有男性和女性且年龄差异较大,每个对象有 10 张图像,共 400 张图像,图像尺寸为 92×112 ,该库是目前使用得最广泛的标准数据库。FERET 人脸库包括了一个通用人脸库以及通用测试标准,共 1400 张图像,其中包含 200 个人每人 7 张的 80×80 的灰度图像。

本实验随机选取训练测试样本,各实验 10 次,取最大值作

为最终结果。部分人脸图像和预处理后的效果如图4所示。

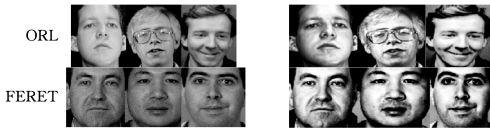


图4 人脸库部分图像

Fig. 4 Part of images in face database

5.1 Curv 脸的结果分析

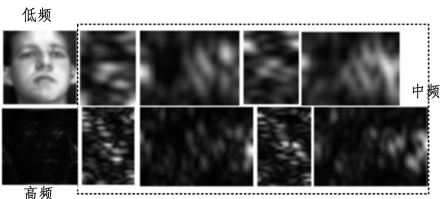
曲波分解结果分为四层三频。coarse 层为低频第一层, 包含一个方向参量, 其大小为 32×32 ; detail 层为中频, 包含第二、三层, 每层有 32 个方向参量, 即 4 个大方向中包括 8 个小方向; fine 层为最后一层, 是只有一个方向参量的 128×128 的高频图像。由于能量集中于低频且噪声含量小, 其包含了图像的全貌, 因此选取 coarse 层作为特征脸。曲波分解层次结构如表 1 所列。

表 1 128×128 曲波分解后的层次结构

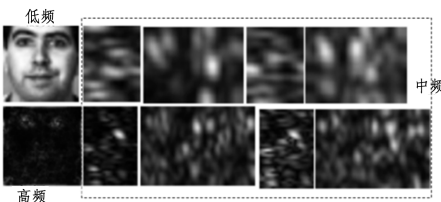
Table 1 128×128 hierarchical structure after curvelet decomposition

层次	尺度	方向个数	分解后形式
coarse 层	$C^D\{1\}$	1	32×32
detail 层	$C^D\{2\}$	$32(4 \times 8)$	2 个 16×12 , 2 个 12×16
	$C^D\{3\}$	$32(4 \times 8)$	2 个 22×32 , 2 个 32×22
fine 层	$C^D\{4\}$	1	128×128

本文对 ORL 和 FERET 人脸库的全部图像提取 Curv 脸。其中中频只显示 8 个小方向中的一幅。为验证算法的有效性, 采用 2D 小波变换提取特征进行对比, 结果如图 5 所示。



(a) ORL 中一个对象的 Curvelet 变换



(b) FERET 中一个对象的 Curvelet 变换



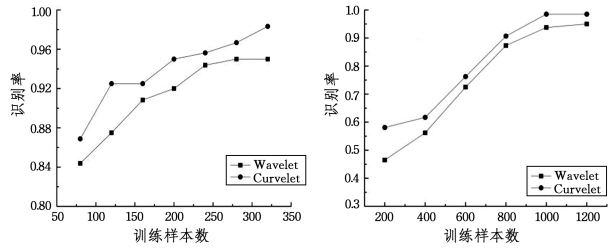
(c) ORL 中一个对象的 Wavelet 变换 (d) FERET 中一个对象的 Wavelet 变换

图 5 2D 小波变换和 Curvelet 变换提取特征脸的对比

Fig. 5 Comparison of feature faces extracted by 2D wavelet transformation and Curvelet transformation

Curvelet 提取特征的过程实质上是初步降维的过程, 对全局人脸特征提取出有效部分, 在特征充分表达的情况下实现第一次粗提取和粗降维。由图 5 可知, 与小波变换相比, 曲波变换后的低频信息的去噪性能更好, 其低频图像特征较充分, 在边缘部分更清晰地体现了人脸轮廓细节。后文在 ORL

和 FERET 人脸库上的实验都是基于此特征脸进行的。不改变 SAE 结构(详见 5.2 节), 只使用不同的特征脸, 则不同训练样本数的 Curvelet 和 Wavelet 得到的识别率如图 6 所示。



(a) ORL 人脸库

(b) FERET 人脸库

图 6 Curvelet 和 Wavelet 的识别率对比

Fig. 6 Contrast of Curvelet and Wavelet accuracy

由图 6 可知, 训练样本数越多, 识别率越高。图 6(a) 中, Wavelet 的训练样本数在 250 以上时识别率几乎趋于 0.96, 而 Curvelet 仍处上升趋势, 识别率也比 Wavelet 高出 0.01~0.05。在图 6(b) 中, 训练样本数多于 800 后变化较小, 同样 Curv 脸的识别率较高。该结果验证了 Curv 脸的有效性。

5.2 ORL 人脸库上 SAE 结果分析

将 SAE 层数设置为 1024-786-441-(40-240)。SAE 与不同降维方法在 ORL 人脸库上的识别精度如图 7 所示, 参数设置如表 2 所列。

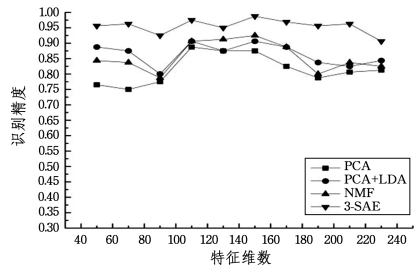


图 7 不同降维算法的识别精度对比

Fig. 7 Comparison of recognition accuracy of different dimension reduction algorithms

表 2 ORL 人脸库参数设置

Table 2 Parameters setting of ORL faces database

ρ	学习率	动量项	maxiter	隐层数	训练样本
0.05	0.1	0.5	200	2	任意 6 张/人

3-SAE 代表三层 SAE(包含输出层)。图 7 结果表明, 与传统降维算法 PCA、PCA+LDA、NMF 相比, 3-SAE 的识别精度明显提高。对于降维后的不同维数变化, 传统算法的鲁棒性较差, 识别精度波动幅度较大且精度不高, 特别是特征维度低于 40 维 PCA 的识别精度仅为 0.75, 而 SAE 的识别精度为 0.95。由此得出在降维后特征维数较低的情况下, SAE 仍能保持稳定性, 鲁棒性更好。这源于 SAE 具有自我学习特征的优势, 使其能应对不同维数下的特征表达。为了验证 SAE 的层数对精度的影响, 测试不同降维后维数的识别精度和训练时间, 结果如表 3 所列。表 3 表明降维后的特征维数越高, SAE 的训练时间就越长, 而精度在 150 维左右达到最大。对于不同隐藏层数的 SAE, 150 维时 3-SAE 和 4-SAE 的训练时间相差不大, 而此时两层隐层的 SAE 的精度已经达到饱和, 无需再增加层数来提高识别精度。综上所述, 选取 3-SAE 且

降维数为 150 时识别效果最佳,识别率为 98.75%,训练时间为 531.34 ms。

表 3 在 ORL 人脸库上的精度和训练时间

Table 3 Accuracy and training time on ORL faces

特征维数	Curv-(2-SAE)		Curv-(3-SAE)		Curv-(4-SAE)	
	精度/%	时间/ms	精度/%	时间/ms	精度/%	时间/ms
50	80.25	509.47	89	517.58	90.75	605.01
100	85.5	802.53	95.25	810.05	96.5	834.12
150	90.5	1695.42	97.75	1809.43	97.5	1926.44
200	92	2080.08	93.75	2336.11	94.25	2459.02

5.3 FERET 人脸库上 SAE 结果分析

ORL 人脸库是小样本量数据库,受光照、姿态等变化影响较小。FERET 是大样本数据库,含噪量较大,包含 1400 张光照、姿态、肤色各异的人脸图像。为了进一步验证算法的鲁棒性和有效性,在 FERET 库上进行相同对比实验,结果如图 8 所示。参数设置同 5.2 节,最大迭代次数为 500,训练样本为每人任意 4 张,其余不变。不同降维算法的识别精度如图 9 所示。



图 8 FERET 中部分变化明显的图像

Fig. 8 Part of obvious changes in images on FERET

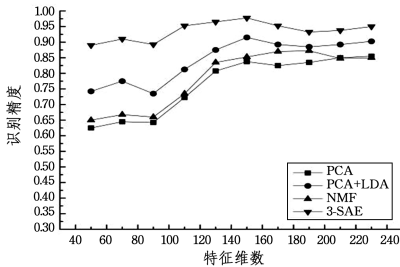


图 9 不同降维算法的识别精度对比

Fig. 9 Comparison of recognition accuracy of different dimension reduction algorithms

由图 9 可知,对于大样本数据,3-SAE 的优势相对 ORL 库更为明显。特征维数在 50~90 之间时,传统降维算法的识别精度在 0.625~0.775 之间,而 SAE 在 0.9 左右;SAE 从低维到高维的识别精度仍能保持稳定,这进一步说明了 SAE 在大样本数据上仍具有较高的鲁棒性和高效性。

同样测试不同降维后维数的识别精度和训练时间,结果如表 4 所列。

表 4 FERET 人脸库上的精度和训练时间

Table 4 Accuracy and training time on FERET faces

特征维数	Curv-(2-SAE)		Curv-(3-SAE)		Curv-(4-SAE)	
	精度/%	时间/ms	精度/%	时间/ms	精度/%	时间/ms
50	85	175.22	95.63	199.74	96.25	205.43
100	88.75	290.94	97.5	272.13	97.5	270.56
150	95.63	522.58	98.75	531.34	98.75	535.22
200	92.5	679.15	96.25	634.20	97.5	665.63

表 4 结果表明,在 FERET 库中特征维度降到 150 左右也能实现较高的识别精度,3-SAE 和 4-SAE 的训练时间相差不大,因此亦可选取 3-SAE 为最佳优化结构。

综合两个人脸库的实验结果可知,因深度学习的自我无

监督学习机能,栈式自编码器能够自动适应不同的数据量并作出相应的调整,以达到较充分的特征表达,实现同等情况下相比其他算法更优越的降维效果;针对光照、姿态、肤色变化的影响,SAE 的鲁棒性更强。

5.4 特征融合结果对比

5.4.1 特征融合对比

将不同隐层与输出层的特征融合后可以进一步提高识别精度,降低误识率。实验结果如图 10 所示。

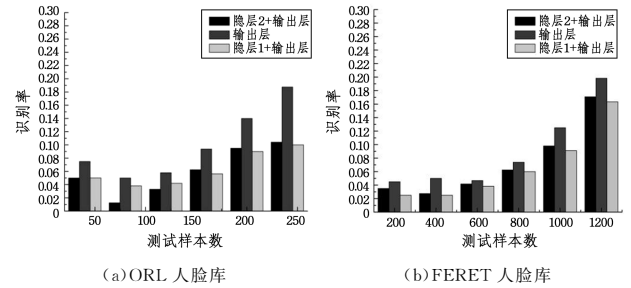


图 10 不同测试样本数的特征融合后的误识率

Fig. 10 Error rates of different test samples after feature fusion

由图 10 的结果可知,特征融合后的误识率更低,这是由于隐层的特征弥补了最后输出层的缺失信息,使特征表达更完整。对于不同隐层的融合,由于隐层 1 的神经节点更多,含有更丰富的细节,因此误识率更低。然而融合后的特征维数约为 900,失去了数据降维本身的意义,加深了分类器的复杂度,不利于快速识别。因此,选取隐层 2 与输出层特征融合,在 ORL 和 FERET 的测试样本分别为 80 和 400 时,其达到最低误识率,分别为 1.25% 和 2.75%。

5.4.2 不同算法的对比

不同算法的识别率如表 5 所列。

表 5 不同算法的识别率

Table 5 Accuracy of different algorithms

不同降维算法	识别率 (%)	
	ORL 上的识别率	FERET 上的识别率
PCA	87.5	82.5
LDA	80.5 ^[16]	78.6 ^[19]
PCA+LDA	90.5	85
NMF	92.5	92.5
KDA ^[17]	95.5	—
LPP	85.94 ^[18]	84.72 ^[20]
LLE	86.5 ^[16]	88.14 ^[20]
ISOMAP	89.5 ^[16]	74.73 ^[20]
SparseAE ^[16]	94.9	—
本文算法	98.75	97.75

由表 5 可知,PCA、LDA、PCA+LDA 在两个人脸库中的识别率并不高,这是因为这些算法是线性分析形式,而图像是非线性结构,易受到光照、姿态等变化的影响;而 NMF 算法虽然得到了 92.5% 的较高识别率,但收敛速度慢,全局最小点不容易找到;KDA 虽然在 ORL 上的识别率较高,为 95%,但主要原因在于核加速,其本质的线性特征并未改变;流形学习 ISOMAP、LLE、LPP 等虽然是非线性结构,能较好地将高维特征映射到低维空间上,但对噪声敏感,鲁棒性较差;SparseAE 是稀疏编码器,并未达到多层栈式结构,其识别率为 94.9%。而本文算法的特征学习能更充分地表达,在实现降维后也能实现较高的识别率,在 ORL 和 FERET 上的识别率

(下转第 305 页)

- cluster[J]. Journal of Systems Architecture, 2014, 60(2): 234-244.
- [12] MILANI B A, NAVIMIPOUR N J. A comprehensive review of the data replication techniques in the cloud environments[J]. Journal of Network & Computer Applications, 2016, 64(C): 229-238.
- [13] SONG J, LI T T, YAN Z X, et al. Energy-Efficiency Model and Measuring Approach for Cloud Computing[J]. Journal of Software, 2012, 23(2): 200-214. (in Chinese)
宋杰, 李甜甜, 闫振兴, 等. 一种云计算环境下的能效模型和度量方法[J]. 软件学报, 2012, 23(2): 200-214.
- [14] ZOU L. Research of Replica Selection Strategy based in Ant Algorithm in Data Grid[D]. Nanjing: Nanjing University of Information Science & Technology, 2014. (in Chinese)
邹露. 基于蚂蚁算法的数据网格副本选择策略研究[D]. 南京: 南京信息工程大学, 2014.
- [15] CAMERON D G, MILLAR A P, CARVAJAL-SCHIAFFINO R, et al. OptorSim: A Simulation Tool for Scheduling and Replica Optimization in Data Grids[OL]. <http://cds.cern.ch/record/865684/files/p707.pdf>.

(上接第 271 页)

分别为 98.75% 和 97.75%, 有较强的鲁棒性。

结束语 SAE 的自我学习特征能够实现数据的有效降维且不丢失重要信息, 而依赖于输入特征的缺陷由 Curv 特征脸来弥补, 鉴于此, 本文提出了一种基于 Curv 脸的 SAE 特征融合自降维的人脸识别算法。本文通过实验验证了算法的有效性, 在不同人脸库上验证了算法具有较高的识别精度和较强的鲁棒性。本文在两个数据集上对该算法进行验证, 下一步工作将在更多的人脸库上进行验证, 并优化结构和算法。

参 考 文 献

- [1] CHENG J L, WECHSLER H. A Gabor Feature Classifier for Face Recognition[C]//IEEE International Conference on Computer Vision, 2001: 270-275.
- [2] CHIEN J T, WU C C. Discriminant Wavelet-faves and Nearest Feature classifiers for Face Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002(24): 1644-1649.
- [3] HU Z H, SONG Y L. Image reduction and reconstruction based on autoencoder network [J]. Journal of Electronic and Information Technology, 2009, 31(5): 1189-1193. (in Chinese)
胡昭华, 宋耀良. 基于 Autoencoder 网络的数据降维和重构[J]. 电子与信息学报, 2009, 31(5): 1189-1193.
- [4] MEE D A, RAT A C. Enhanced Face Recognition Through Variation of Principle Component Analysis(PCA)[C]//International Conference on Industrial and Information. 2007: 347-351.
- [5] BEL P N, HES J P, KRR D J, et al. Fisherfaces: Recognition Using Class Specific Linear Projection[J]. IEEE Trans. on PAMI, 1997, 19(7): 711-720.
- [6] MART T, LEO M, D'ORAZIO. Facial Feature Extraction by Kernel Independent Component Analysis[C]//IEEE Conference on Advanced Video and Signal Based Surveillance, 2005: 210-275.
- [7] HU Z H, SONG Y L. Image reduction and reconstruction based on a continuous self-coding network [J]. Data Collection and Processing, 2010, 25(3): 318-324. (in Chinese)
胡昭华, 宋耀良. 基于一种连续自编码网络的图像降维和重构[J]. 数据采集与处理, 2010, 25(3): 318-324.
- [8] TEN J B, SILVA V D, LANGFO J C. A Global Geometric Framework for Nonlinear Dimensionality Reduction[J]. Science, 2000, 290(5500): 2319-2323.
- [9] ROUEIS S T, SAUL L K. Nonlinear Dimensionality Reduction by Locally Linear Embedding [J]. Science, 2000, 290(5500): 2323-2326.
- [10] LIANG W, DAVID S. Learning and Matching of Dynamic Shape Manifolds for Human Action Recognition[J]. IEEE Transactions on Image Processing, 2007, 16(6): 1646-1661.
- [11] HINTON G E, OSIN S. A Fast Learning Algorithm for Deep Belief Nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [12] HINTON G E, SALAK R R. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(9): 504-507.
- [13] CANDES E J, DONOHO D L. Curvelet-A Surprisingly Effective Nonadaptive Representation for Objects with Edges[M]//Curve and Surface Fitting; SaintMalo. TN: Vanderbilt Univ Press, 1999.
- [14] CANDES E J, DEMANET L, DONOHO D L. Fast Discrete Curvelet Transforms[R]. Applied and Computational Mathematics. California: California Institute of Technology, 2005.
- [15] VINCENT P, LARO H, BENGIO Y, et al. Extracting and Composing Robust Features with Denoising Autoencoders[C]//Proceedings of the 25th International Conference on Machine Learning. 2008: 1096-1103.
- [16] ZHU X T, ZHANG Y Z, WANG F D. Research on data dimension reduction method based on sparse self-coding network[J]. Journal of Shenyang University of Technology, 2016, 35(5): 39-44. (in Chinese)
朱啸天, 张艳珠, 王凡迪. 一种基于稀疏自编码网络的数据降维方法研究[J]. 沈阳理工大学学报, 2016, 35(5): 39-44.
- [17] CHEN D Y. Research on dimension reduction algorithm of manifold learning and its application in face recognition[D]. Jiangsu: Jiangnan University, 2014. (in Chinese)
陈达遥. 流形学习降维算法研究及其在人脸识别中的应用[D]. 江苏: 江南大学, 2014.
- [18] WANG H Y. Research on manifold learning feature extraction method and face recognition in subspace [D]. Harbin: Harbin Polytechnic University, 2017. (in Chinese)
王海燕. 子空间的流形学习特征提取方法及人脸识别研究[D]. 哈尔滨: 哈尔滨理工大学, 2017.
- [19] GUO J X, CHEN W. Feature recognition based on HOG multi-feature fusion and random forest [J]. Computer Science, 2013, 40(10): 279-284. (in Chinese)
郭金鑫, 陈玮. 基于 HOG 多特征融合与随机森林的人脸识别[J]. 计算机科学, 2013, 40(10): 279-284.
- [20] FADHLAN K Z, AMIR A S, YASIR M M. Robust Face Recognition Against Expressions and Partial Occlusions[J]. International Journal of Automation and Computing, 2016, 13(4): 319-337.