

融合局部特征与深度置信网络的人脸表情识别

王琳琳¹, 刘敬浩¹, 付晓梅²

¹天津大学电气自动化与信息工程学院, 天津 300072;

²天津大学海洋科学与技术学院, 天津 300072

摘要 针对传统人脸表情识别(FER)方法所提取的表情特征较为单一,同时对于表情分类器的选择存在局限性的问题,提出一种融合局部特征与深度置信网络(DBN)的FER方法。该方法首先从人脸表情图像中切割出眉毛眼睛部位与嘴巴部位这2种包含丰富表情信息的局部表情图像,对其分别提取包含纹理信息的Log-Gabor特征与包含形状信息的二阶梯度方向直方图特征,并将这2种特征相融合,获得更有效的表情特征,然后利用融合后的特征训练DBN模型,并用训练后的DBN模型进行表情识别。利用本文方法在三种表情库上进行实验,识别率可分别达到96.30%、97.39%以及95.73%,表明本文方法可有效提高人脸表情识别率。

关键词 图像处理; 表情识别; 特征融合; Log-Gabor特征; 二阶梯度方向直方图特征; 深度置信网络

中图分类号 TP391

文献标识码 A

doi: 10.3788/LOP55.011002

Facial Expression Recognition Based on Fusion of Local Features and Deep Belief Network

Wang Linlin¹, Liu Jinghao¹, Fu Xiaomei²

¹School of Electrical and Information Engineering, Tianjin University,

Tianjin 300072, China;

²School of Marine Science and Technology, Tianjin University, Tianjin 300072, China

Abstract The traditional facial expression recognition (FER) methods only extract single expression feature. Meanwhile, the choice of expression classifiers has limitations. To solve these problems, we propose a FER method based on the fusion of local features and deep belief network (DBN). Firstly, the eyebrows and eyes part and mouth part with rich expression information are extracted as local expression images. In order to attain more effective expression features, the Log-Gabor features with texture information and second-order histogram of gradient direction features with shape information are extracted and fused from local expression images. DBN model is trained with fusion features. The trained DBN model is used to recognize the facial expression. The experimental results show that the recognition rates of the proposed method on three databases are 96.30%, 97.39% and 95.73%. The proposed method effectively improves the recognition rate of facial expression.

Key words image processing; expression recognition; feature fusion; Log-Gabor features; second-order histogram of gradient direction features; deep belief network

OCIS codes 100.3008; 100.2960; 100.5010; 150.1135

1 引言

人脸表情识别(FER)技术是一种典型的多分类

情感计算技术^[1],广泛应用于图像理解、虚拟现实、医疗、人机交互等领域,是近年来机器视觉研究的热点^[2]。对人脸表情识别的研究,主要集中在人脸表

收稿日期: 2017-05-27; 收到修改稿日期: 2017-08-07

基金项目: 国家自然科学基金(61571323)

作者简介: 王琳琳(1992—),女,硕士研究生,主要从事模式识别与机器学习方面的研究。

E-mail: wanglinlin@tju.edu.cn

导师简介: 刘敬浩(1963—),男,硕士,副教授,主要从事计算机通信与模式识别方面的研究。

E-mail: liujinghao@tju.edu.cn

情的特征提取以及表情分类器的选择方面。

表情识别的关键在于如何提取出完整、稳健而富有区分性的表情特征^[3]。Gabor^[4]小波变换常用于提取图像纹理特征。Log-Gabor^[5]滤波器相较于Gabor滤波器,具有不受光照影响的相位一致性,以及符合人类视觉系统认知的特性^[6]。梯度方向直方图(HOG)特征^[7]常用于描述图像边缘形状。相比传统的HOG特征,二阶HOG^[8]特征包含了更多信息,因此,可以进一步提高HOG特征对于形状和表情的描述能力。特征融合可以获得丰富的特征信息^[9]。文献[10]将面部显著区域的局部二值模式(LBP)直方图与HOG特征相融合进行表情识别。文献[11]提出了一种融合时空运动LBP特征与多方向融合Gabor直方图特征的表情识别方法,实验结果显示,相较于单独使用其中一种特征进行表情识别,融合特征能得到更好的识别效果。人脸表情特征融合可分为基于整体脸部特征^[12]和基于局部脸部特征^[13]。文献[13]将眉毛眼睛和嘴巴部位分割出来进行特征提取与表情分类,得到了很好的分类效果。

在分类方法选择方面,目前的表情识别研究多采用支持向量机(SVM)^[14]、K最邻近规则(KNN)^[15]、隐马尔可夫模型(HMM)^[16]等机器学习方法。传统的机器学习方法多属于浅层学习方法,当样本数量增多、多样性增强时,浅层学习在复杂情况下的运算能力与适应能力受到限制^[17]。2006年,Hinton等^[18]提出了深度置信网络(DBN)模型。DBN是一种模拟人类视觉的认知过程,通过逐层训练网络,实现高维度非线性数据特征抽取与分类的深度学习方法。DBN模型作为一种有效的分类识别方法,适用于表情识别之中。文献[19]先提取了人脸的HOG特征,通过DBN模型进行人脸解析,再提取局部器官的Gabor特征,通过堆叠自动编码器进行表情识别;但是,该研究所采用的特征为传统的HOG与Gabor特征,没有采用改进的HOG与Gabor特征或融合特征以进一步提高识别率。文献[20]根据AdaBoost算法的思想,将若干个DBN作为弱分类器进行级联,组成强DBN分类器进行表情识别;但这种方法需要训练的DBN模型数量较多,计算复杂度高,在模型训练阶段耗时过长。

本文提出一种基于人脸局部器官图像,融合了Log-Gabor特征和二阶HOG特征与DBN模型的表情识别算法。该方法首先分割出眉毛眼睛部位和嘴巴部位2种局部表情图像,对局部器官图像提取Log-Gabor特征与二阶HOG特征并进行融合,然

后将融合后的特征输入到DBN模型中进行训练,最后利用训练后的DBN模型进行表情识别。实验结果表明,本文方法具有很高的表情识别率。

2 Log-Gabor 与二阶 HOG 特征

2.1 Log-Gabor 特征

Gabor特征主要包含空间尺度、空间位置、方向选择等图像局部纹理信息^[21]。二维Gabor滤波器定义为

$$G_{\alpha,\beta}(z) = \frac{\|\mathbf{k}_{\alpha,\beta}\|^2}{\sigma^2} \exp\left(-\frac{\|\mathbf{k}_{\alpha,\beta}\|^2 \|z\|^2}{2\sigma^2}\right) \times \left[\exp(i\mathbf{k}_{\alpha,\beta} \cdot z) - \exp\left(-\frac{\sigma^2}{2}\right)\right], \quad (1)$$

式中 $z=(x,y)$ 为图像上定点位置, α 和 β 分别表示Gabor滤波器的方向和尺度(频率), $\mathbf{k}_{\alpha,\beta}$ 为滤波器的中心频率, $\mathbf{k}_{\alpha,\beta}=k_{\beta}(\cos\phi,\sin\phi)^T$, $k_{\beta}=2^{(-\beta+2)/2}\pi$, $\phi=\alpha\cdot\pi/K$, K 为方向个数, i 为复数算子, σ 为滤波器的带宽, $\|\cdot\|$ 表示模。

由Gabor滤波器的函数可知,对于偶对称Gabor滤波器:如果高斯函数的标准偏差 σ 大于中心频率的1/3,2个高斯函数在 origin 重合产生直流分量,则导致无法处理亮度变化较大的图像,同时无法构造互成正交的滤波器;如果标准偏差 σ 等于中心频率的1/3,带宽大约为1倍频,则导致带宽不能任意扩展。

Log-Gabor滤波器是Gabor滤波器的一种改进。与Gabor滤波器相比:1) Log-Gabor滤波器没有直流分量,带宽可以扩展,能够处理亮度变化较大的图像;2) Log-Gabor滤波器的传递函数在高频端有一个延长的尾巴,更适合对自然图像进行编码,克服了传统Gabor函数在低频表示过度而在高频表示不足的缺点。二维Log-Gabor滤波器定义为

$$L(f,g) = \exp\left\{-\frac{[\lg(f_1/f_0)]^2}{2[\lg(k/f_0)]^2}\right\} \exp\left[-\frac{g_1^2}{(2\delta)^2}\right], \quad (2)$$

式中 $f_1=f\cos\gamma+g\sin\gamma$, $g_1=-f\sin\gamma+g\cos\gamma$, f_0 为滤波器的中心频率, γ 为滤波器的方向角度, k 控制 f_1 方向的带宽, δ 控制 g_1 方向的带宽。

为了获得人脸表情图片的Log-Gabor特征,需要将人脸表情图片与Log-Gabor滤波器进行卷积。若一幅图像的灰度值为 $I(x,y)$,那么经过Log-Gabor滤波器后的结果为 $F(x,y)=I(x,y)*L_{\mu,\nu}(x,y)$,其中 $L_{\mu,\nu}(x,y)$ 表示 μ 尺度、 ν 方向的

Log-Gabor 滤波器, $F(x, y)$ 为 Log-Gabor 滤波器滤波后的结果, $*$ 为卷积运算。图 1 所示为对一幅

人脸表情图像提取眉毛眼睛部位图像后, 进行 5 个尺度 8 个方向的 Log-Gabor 幅值特征提取的结果。

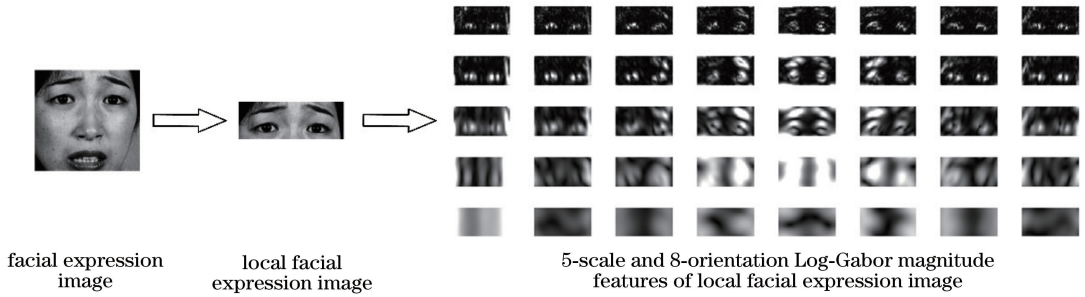


图 1 人脸局部表情图像的 Log-Gabor 幅值特征

Fig. 1 Log-Gabor magnitude features of local facial expression image

2.2 二阶 HOG 特征

HOG 特征通过计算图像单元的梯度方向直方图, 能够有效地描述图像的边缘和形状。二阶 HOG 特征在 HOG 特征的基础之上, 通过计算图像内多个单元直方图元素之间的相互关系, 对一阶直

方图特征提取二阶统计特征, 从而提高了特征对于人脸表情的描述的区别能力。二阶 HOG 的具体步骤如下。

1) 采用一阶梯度算子 $[-1, 0, 1]$ 计算图像中每个点的梯度幅值和方向, 计算公式为

$$g(x, y) = \sqrt{[I(x+1, y) - I(x-1, y)]^2 + [I(x, y+1) - I(x, y-1)]^2}, \quad (3)$$

$$\varphi(x, y) = \arctan \frac{I(x, y+1) - I(x, y-1)}{I(x+1, y) - I(x-1, y)}, \quad (4)$$

式中 $I(x, y)$ 为图像在点 (x, y) 处的灰度值, $g(x, y)$ 为梯度幅值, $\varphi(x, y)$ 为梯度方向。

2) 将图像分为 $m \times n$ 个大小相同的单元, 并统计每个单元的梯度方向直方图。

3) 取相邻若干个图像单元组成 $p \times q$ 个图像块, 将图像块内的所有直方图元素进行相互之间的成对组合, 并采用调和平均值计算一种成对关系

值^[22], 获得图像块的成对关系向量。

4) 对得到的图像块成对关系向量进行 L2 范数归一化, 减少光照影响。将归一化后的向量按顺序串联起来, 从而得到整个图像的二阶 HOG 特征。

对一幅人脸表情图像提取眉毛眼睛部位得到 $128 \text{ pixel} \times 48 \text{ pixel}$ 的局部图像, 将局部图像划分为 $16 \text{ pixel} \times 16 \text{ pixel}$ 的单元, 取相邻 2×2 个单元组成一个图像块, 图像块之间重叠 2 个单元, 对其中一个图像块内的梯度直方图元素进行成对组合, 得到该图像块的二阶 HOG 特征, 如图 2 所示。

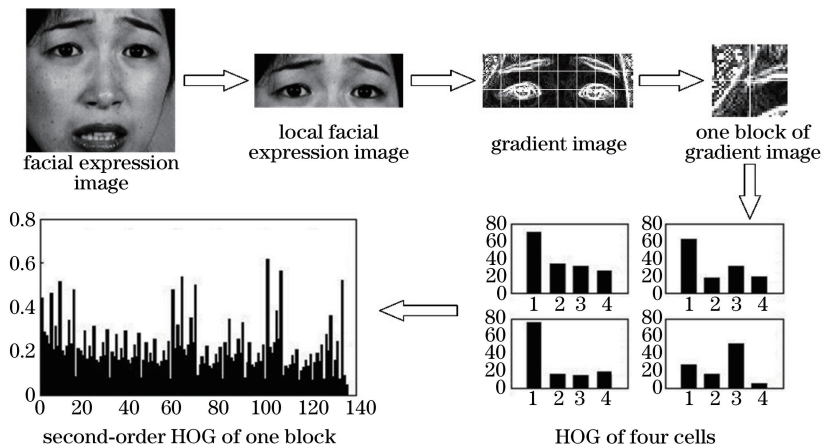


图 2 人脸局部表情图像的二阶 HOG 特征

Fig. 2 Second-order HOG features of local facial expression image

3 融合局部特征与 DBN 的人脸表情识别

3.1 DBN 模型

DBN 是一种深度学习网络,由多层受限玻尔兹曼机(RBM)和一层反向传播(BP)神经网络堆叠而成。RBM 是 DBN 的关键组成部分。RBM 是一种具有二分结构的无向图模型,由可视层 v 和隐含层 h 组成,每层内部无连接,可视层与隐含层之间全连接,连接权重为 w 。可视层与隐含层单元都为二元值,即 0 或 1。

对于可视层与隐含层的所有结点,RBM 的能量函数定义为

$$E(v, h | \theta) = - \sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m w_{ij} v_i h_j, \quad (5)$$

式中 $\theta = (w_{ij}, a_i, b_j)$ 为 RBM 的参数, a_i 为可视层单元的偏置值, b_j 为隐含层单元的偏置值, w_{ij} 为可视层与隐含层之间的连接权重, n 和 m 分别为可视层与隐含层的神经元数目。由能量函数可以得到可视层与隐含层的联合概率分布为

$$P(v, h | \theta) = \frac{\exp[-E(v, h | \theta)]}{Z(\theta)}, \quad (6)$$

式中 $Z(\theta) = \sum_v \sum_h \exp[-E(v, h | \theta)]$ 是归一化函数。可视层与隐含层节点之间无连接,所以各个节点之间相互独立。当可视层节点状态已知时,隐含层节点的激活概率为

$$P(h_j = 1 | v, \theta) = \sigma\left(b_j + \sum_{i=1}^n v_i w_{ij}\right), \quad (7)$$

式中 $\sigma(x) = 1/(1 + e^{-x})$ 。当隐含层节点状态已知时,可视层节点的激活概率为

$$P(v_i = 1 | h, \theta) = \sigma\left(a_i + \sum_{j=1}^m h_j w_{ij}\right). \quad (8)$$

RBM 采用迭代训练的方法,训练出可以拟合给定训练数据的参数 θ 。通过计算在训练集上的极大对数似然函数,可以得到参数 θ 。通过对比散度(CD)算法^[23],可以得到参数 θ 的更新规则:

$$\Delta w_{ij} = \epsilon (\langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{recon}}), \quad (9)$$

$$\Delta a_{ij} = \epsilon (\langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{recon}}), \quad (10)$$

$$\Delta b_{ij} = \epsilon (\langle h_j \rangle_{\text{data}} - \langle h_j \rangle_{\text{recon}}), \quad (11)$$

式中 ϵ 为学习率, $\langle \cdot \rangle_{\text{data}}$ 表示数据的分布期望, $\langle \cdot \rangle_{\text{recon}}$ 表示通过 CD 算法得到的分布期望。

DBN 模型的训练过程主要分为预训练和微调。第一步通过无监督贪婪的方式自底向上逐层训练

RBM,下层 RBM 隐含层的输出作为向上一层 RBM 的可视层输入。逐层训练后的 RBM 可以从高维数据中提取出更有区别度的低维数特征。第二步通过有监督的方式采用 BP 神经网络对数据进行分类,同时将误差自顶向下传播,对 RBM 网络进行微调,使 RBM 整个网络的参数达到最优。典型的 DBN 结构如图 3 所示。

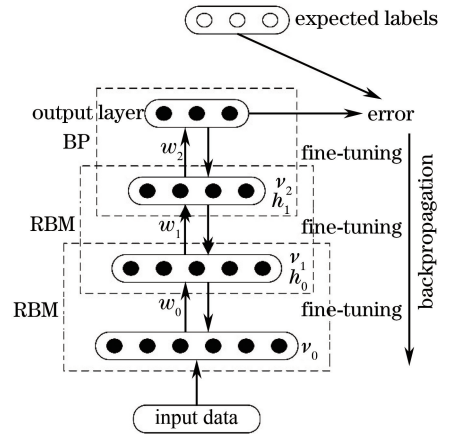


图 3 DBN 结构图

Fig. 3 Structure of DBN

3.2 融合局部特征与 DBN 的人脸表情识别方法

图 4 为融合局部特征与 DBN 的人脸表情识别方法的流程图,具体步骤如下:

- 1) 从人脸表情图像中切割出眉毛眼睛部位和嘴巴部位的表情图像。
- 2) 对局部人脸表情图像分别提取描述纹理的 Log-Gabor 特征与描述边缘和形状的二阶 HOG 特征。
- 3) 采用串联方式将这 2 种特征进行融合,形成融合局部表情特征向量。
- 4) 采用融合后的表情特征训练 DBN 模型。
- 5) 利用训练后的 DBN 模型进行表情分类。

4 实验与结果分析

4.1 数据库预处理与特征提取

选用日本女性人脸表情(JAFFE)库、Cohn-Kanade(CK)表情库和 Extended Cohn-Kanade(CK+)表情库进行实验,实验样本图像如图 5 所示。

JAFFE 表情库中包含 10 人的 213 张表情图像,实验随机选取每人每种表情 1~2 张共 105 张作为训练图像,余下的 108 张作为测试图像。CK 表情库中包含 97 人的 486 个表情序列共 8795 张表情图像,从 CK 表情库中选取 20 个人的表情图像,选取每人每种表情图像 2~3 张,共 303 张表情图像,实验随机选取每人每种表情图像 1~2 张,共 150 张

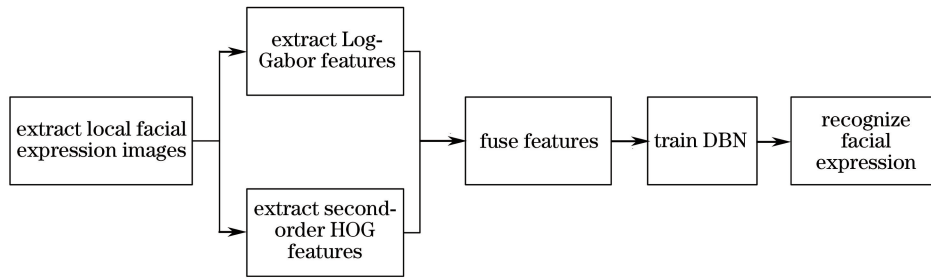


图4 融合局部特征与DBN的人脸表情识别流程

Fig. 4 Flowchart of facial expression recognition based on fusion of local features and DBN

作为训练表情,余下的153张图片作为测试图像。CK+表情库包含123人的593个表情序列共10708张表情图像,选取带标签的327个表情序列中每个序列的最后一张峰值表情图像作为实验图像,包含愤怒

45张、厌恶59张、恐惧25张、高兴69张、悲伤28张、惊讶83张,以及轻视18张。随机选取每种表情图像的1/2共计163张作为训练图像,余下的164张作为测试图像。每次实验重复3次取平均值。

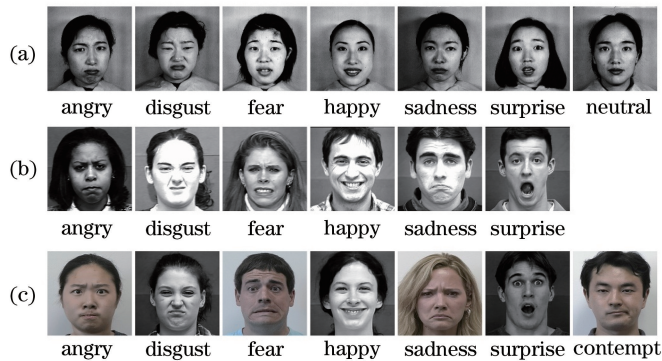


图5 样本图像。(a) JAFFE表情库;(b) CK表情库;(c) CK+表情库

Fig. 5 Sample images. (a) JAFFE database; (b) CK database; (c) CK+ database

对3种表情库进行预处理,首先提取出大小为 $128 \text{ pixel} \times 128 \text{ pixel}$ 的纯脸图像,其中CK+表情库中的彩色图像需要转化为灰度图像。然后在纯脸图像的基础上根据人脸“三庭五等”的结构,切割出大小为 $128 \text{ pixel} \times 48 \text{ pixel}$ 的眉毛眼睛部位局部图像,以及大小为 $80 \text{ pixel} \times 32 \text{ pixel}$ 的嘴巴部位局部图像,如图6所示。

对预处理提取到的2种局部表情图像,分别提取5尺度8方向的Log-Gabor幅值特征,并将2种局部图像的Log-Gabor特征按顺序进行串联。将预处理得到的2种局部表情图像,分别划分为 $16 \text{ pixel} \times 16 \text{ pixel}$ 的单元,每个图像块由相邻 2×2 个单元组成,图像块之间重叠2个单元,分别提取出两种局部图像的二阶HOG特征,并按顺序进行串联。将Log-Gabor特征与二阶HOG特征进行串联,得到融合之后的局部特征。

4.2 DBN参数设置

RBM预训练的学习率设置为0.05,每层RBM

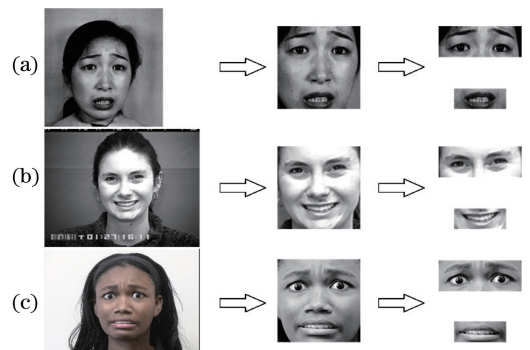


图6 表情库图像预处理示例。(a) JAFFE表情库;(b) CK表情库;(c) CK+表情库

Fig. 6 Examples of facial expression database image preprocessing. (a) JAFFE database; (b) CK database; (c) CK+ database

的迭代次数为10,BP神经网络的学习率设置为0.05,迭代次数设置为1000。采用批训练的方式提高训练速度,批训练样本数设置为100。DBN网络结构的深度对表情检测效果有着重要的影响,本文

通过实验的方式确定 DBN 模型网络中 RBM 的层数,由于实验所用的表情数据库样本数较小,所以 DBN 模型所需的 RBM 层数较少,分别设置 DBN 模型中 RBM 层数为 1~4,RBM 隐含层节点数设置为 50。实验的硬件环境为 3.7 GHz Intel Core i7 CPU,16 GB RAM 计算机,软件环境为 Matlab R2014a。不同 RBM 网络层数的 DBN 模型表情识别率如图 7 所示,训练与识别时间如表 1 所示。

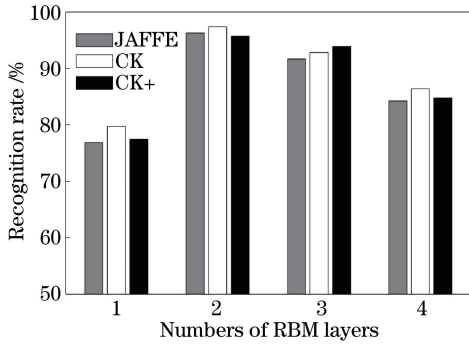


图 7 不同 RBM 层数的 DBN 模型的表情识别率

Fig. 7 Expression recognition rate of DBN with different RBM layers

表 1 不同层数 RBM 的 DBN 模型训练与识别时间

Table 1 Training and recognition time of DBN with different RBM layers

| Database | 1 RBM layer | 2 RBM layers | 3 RBM layers | 4 RBM layers |
|----------|-------------|--------------|--------------|--------------|
| JAFFE | 344.86 | 228.90 | 265.38 | 711.62 |
| CK | 337.76 | 402.75 | 537.88 | 669.64 |
| CK+ | 369.32 | 461.98 | 542.23 | 743.40 |

由图 7 与表 1 结果可知,当 DBN 模型中 RBM 的层数为 1 时,模型的训练与识别时间较长,且识别率不高,这是由于 RBM 层数过少导致重构误差较大,模型收敛较慢且识别率不佳。当 RBM 的层数为 2 时,在 3 种数据库中进行实验的表情识别率均最高,且模型训练与识别时间相对较短。当层数继续增加时,DBN 模型的识别率有所下降,同时所需要的训练与识别时间逐渐增加。因此,本文 DBN 模型最终采用 2 层 RBM 结构。

4.3 与其他方法对比

为了验证本文方法对于表情分类识别的有效性,在局部表情图像中分别提取 Gabor 特征、Log-Gabor 特征、HOG 特征、二阶 HOG 特征,然后再采用 DBN 模型在 JAFFE 表情库、CK 表情库和 CK+ 表情库分别进行表情识别,不同特征的识别率对比结果如表 2 所示。

表 2 不同特征的识别率

Table 2 Recognition rate based on different features %

| Feature | JAFFE database | CK database | CK+ database |
|------------------|----------------|-------------|--------------|
| Gabor | 87.96 | 90.20 | 88.42 |
| Log-Gabor | 93.52 | 94.77 | 93.29 |
| HOG | 85.19 | 88.24 | 86.83 |
| Second order HOG | 92.59 | 94.12 | 92.68 |
| Log-Gabor+ | 96.30 | 97.39 | 95.73 |
| Second order HOG | | | |

由表 2 可知,与单独采用 Gabor 特征、Log-Gabor 特征、HOG 特征、二阶 HOG 特征相比,所提出的融合 Log-Gabor 与二阶 HOG 特征的方法,在 JAFFE 表情库、CK 表情库以及 CK+ 表情库上的识别率更高,分别达到 96.30%、97.39%、95.73%,验证了融合特征的有效性。

为了验证 DBN 模型对于表情分类识别的有效性,将本文方法同 KNN 与 SVM 算法进行对比。对局部图像提取 Log-Gabor 特征与二阶 HOG 特征并进行特征融合,之后采用不同的分类方法在 JAFFE 表情库、CK 表情库与 CK+ 表情库分别进行表情识别。SVM 采用广泛使用的 LIBSVM 工具包,选用 C-SVC 类型,核函数采用径向基核函数(RBF),Gamma 参数为 0.5,惩罚因子 $C=10$ 。对比实验结果如表 3 所示。

表 3 不同算法的识别率

Table 3 Recognition rate of different algorithms %

| Algorithm | JAFFE database | CK database | CK+ database |
|-----------|----------------|-------------|--------------|
| KNN | 75.00 | 78.43 | 77.44 |
| SVM | 82.41 | 83.01 | 81.10 |
| DBN | 96.30 | 97.39 | 95.73 |

由表 3 可知,在相同融合特征条件下,DBN 模型对于表情的整体识别率在不同数据库中均高于传统的浅层学习方法(SVM 与 KNN)。具有深度学习结构的 DBN 模型通过无监督的预训练与有监督的微调,对底层特征逐层进行特征提取,最终学习到适合表情识别的特征,提高了表情识别的效果。

将本文方法同近年来学者们所提出的其他方法,比如局部定向数字(LDN)特征+SVM^[26]、HOG+bagging 极限学习机(ELM)^[27]、LBP+多任务稀疏学习算法(MTSL)^[28]、复合局部二值模式(CLBP)+SVM^[29]、保持全局和局部多样性的特征提取算法(GLDPE)^[30]以及 HOG+DBN+Gabor+

堆自动编码器(SAE)^[19]等方法,在 JAFFE 表情库、CK 表情库和 CK+表情库的识别结果进行对比,对比结果如表 4~6 所示。

表 4 JAFFE 表情库上不同方法识别率对比

Table 4 Comparison of recognition rate of different methods on JAFFE database

| Method | Recognition rate /% |
|--------------------------------------|---------------------|
| PHOG+LBP+SVM ^[24] | 87.43 |
| Local Gabor+RFLD+KNN ^[25] | 89.67 |
| LDN+SVM ^[26] | 90.60 |
| HOG+bagging ELM ^[27] | 94.37 |
| Proposed method | 96.30 |

表 5 CK 表情库上不同方法识别率对比

Table 5 Comparison of recognition rate of different methods on CK database

| Method | Recognition rate /% |
|--------------------------------------|---------------------|
| Local Gabor+RFLD+KNN ^[25] | 91.51 |
| LBP+MTSL ^[28] | 91.53 |
| CLBP+SVM ^[29] | 94.20 |
| GLDPE ^[30] | 97.08 |
| Proposed method | 97.39 |

表 6 CK+表情库上不同方法识别率对比

Table 6 Comparison of recognition rate of different methods on CK+ database

| Method | Recognition rate /% |
|--|---------------------|
| Geometric features+LBP+SVM ^[31] | 90.08 |
| HOG+DBN+Gabor+SAE ^[19] | 91.11 |
| PHOG+LBP+SVM ^[24] | 94.63 |
| Boosted DBN ^[20] | 96.70 |
| Proposed method | 95.73 |

由表 4~6 的对比结果可知,在 3 种数据库中,同其他方法相比,本文方法均有较好的识别效果。在 JAFFE 库中,本文方法较其他方法的识别率至少提高了 1.93%,尤其是与文献[24]所采用的融合金字塔梯度分布直方图(PHOG)与 LBP 特征,并基于 SVM 进行分类的方法相比,识别率提高了 8.87%。在 CK 表情库中,本文方法的识别率较其他方法至少提高了 0.31%,与文献[25]所采用的基于局部 Gabor 特征、回归 fisher 线性判别分析(RFLD)降维,以及 KNN 分类的识别方法相比,识别率提高了 5.88%。在 CK+表情库中,虽然本文方法的识别率略低于文献[20]所采用的 Boosted DBN 方法 0.97%,但是文献[20]方法需要构建 80 个 DBN 模型,而本文方法只需要构建 1 个 DBN 模型,大大减少了计算的复杂度与计算时间。同时,本

文方法与文献[31]所采用的基于几何特征与 LBP 特征融合、SVM 分类识别的方法相比,识别率提高了 5.65%。

5 结 论

采用融合局部特征与 DBN 模型进行表情识别。从表情图像中提取眉毛眼睛部位与嘴巴部位作为局部表情图像,有效地减少了冗余信息。提取局部表情图像的 Log-Gabor 特征与二阶 HOG 特征,并相融合,融合后的特征同时包含了纹理特征与形状、边缘特征,具有更加丰富的表情信息。DBN 模型通过深度学习网络,能够更好地提取到更深层次的特征,进一步提高了识别率。将本文方法应用在 JAFFE 表情库、CK 表情库以及 CK+表情库上,识别率分别达到 96.30%、97.39%与 95.73%,证明了本文方法对于表情识别的有效性。在今后的研究中,应进一步探索如何优化模型,使其得以应用于存在部分遮挡的人脸表情识别中。

参 考 文 献

- [1] Chen J, Takiguchi T, Ariki Y. Facial expression recognition with multithreaded cascade of rotation-invariant HOG [C]. International Conference on Affective Computing and Intelligent Interaction, IEEE, 2015: 636-642.
- [2] Negri P, Acevedo D, Mejail M, *et al.* Facial expression recognition: A comparison between static and dynamic approaches [C]. International Conference on Pattern Recognition Systems, 2016: 1-6.
- [3] Liu S S, Tian Y T, Wan C, *et al.* Facial expression recognition method based on Gabor multi-orientation features fusion and block histogram [J]. Acta Automatica Sinica, 2011, 37(12): 1455-1463.
刘帅师, 田彦涛, 万川. 基于 Gabor 多方向特征融合与分块直方图的人脸表情识别方法[J]. 自动化学报, 2011, 37(12): 1455-1463.
- [4] Ye Z, Bai L, Nian Y J. Hyperspectral image classification algorithm based on Gabor feature and locality-preserving dimensionality reduction[J]. Acta Optica Sinica, 2016, 36(10): 1028003.
叶珍, 白麟, 粘永健. 基于 Gabor 特征与局部保护降维的高光谱图像分类算法[J]. 光学学报, 2016, 36(10): 1028003.
- [5] Liu Y, Wu X J. Image recognition algorithm based on Log-Gabor wavelet and Riemannian manifold

- learning [J]. Pattern Recognition and Artificial Intelligence, 2015, 28(10): 946-952.
- 刘元, 吴小俊. 基于 Log-Gabor 滤波与黎曼流形学习的图像识别算法[J]. 模式识别与人工智能, 2015, 28(10): 946-952.
- [6] Yuan W Q, Fan Y G, Ke L. Palmprints recognition method based on the phase consistency combined with Log-Gabor filter[J]. Acta Optica Sinica, 2010, 30(1): 147-152.
- 苑玮琦, 范永刚, 柯丽. 相位一致性和对数 Gabor 滤波器相结合的掌纹识别方法[J]. 光学学报, 2010, 30(1): 147-152.
- [7] Liu B, Zhang X, Hu C H, *et al.* Gesture recognition method of hand over face occlusion in color and depth images[J]. Laser & Optoelectronic Progress, 2016, 53(6): 061001.
- 刘斌, 赵兴, 胡春海, 等. 面向颜色深度图像手脸近距离遮挡的手势识别[J]. 激光与光电子学进展, 2016, 53(6): 061001.
- [8] Cao H, Yamaguchi K, Naito, *et al.* Pedestrian recognition using second-order HOG feature[C]. 9th Asian Conference on Computer Vision (ACCV 2009), 2009: 628-634.
- [9] Zhang S H, He H, Kong L F. Fusing multi-feature for video occlusion region detection based on graph cut[J]. Acta Optica Sinica, 2015, 35(4): 0415001.
- 张世辉, 何欢, 孔令富. 融合多特征基于图割实现视频遮挡区域检测[J]. 光学学报, 2015, 35(4): 0415001.
- [10] Liu Y, Li Y, Ma X, *et al.* Facial expression recognition with fusion features extracted from salient facial areas[J]. Sensors, 2017, 17(4): 712.
- [11] Zhao L, Wang Z, Zhang G. Facial expression recognition from video sequences based on spatial-temporal motion local binary pattern and Gabor multiorientation fusion histogram[J]. Mathematical Problems in Engineering, 2017, 2017: 1-12.
- [12] Jia Q, Gao X, Guo H, *et al.* Multi-layer sparse representation for weighted LBP-patches based facial expression recognition[J]. Sensors, 2015, 15(3): 6719-6739.
- [13] Hu M, Jiang H, Wang X H, *et al.* Precise local feature description for facial expression recognition [J]. Journal of Image and Graphics, 2014, 19(11): 1613-1622.
- 胡敏, 江河, 王晓华, 等. 精确局部特征描述的表情识别[J]. 中国图象图形学报, 2014, 19(11): 1613-1622.
- [14] Li Y, Mavadati S M, Mahoor M H, *et al.* A unified probabilistic framework for measuring the intensity of spontaneous facial action units [C]. IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, 2013: 1-7.
- [15] Wang Q, Jia K, Liu P. Design and implementation of remote facial expression recognition surveillance system based on PCA and KNN algorithms [C]. International Conference on Intelligent Information Hiding and Multimedia Signal Processing. 2016: 314-317.
- [16] Siddiqi M H, Ali R, Idris M, *et al.* Human facial expression recognition using curvelet feature extraction and normalized mutual information feature selection [J]. Multimedia Tools and Applications, 2016, 75(2): 935-959.
- [17] Liu D W, Han L, Han X Y. High spatial resolution remote sensing image classification based on deep learning [J]. Acta Optica Sinica, 2016, 36(4): 0428001.
- 刘大伟, 韩玲, 韩晓勇. 基于深度学习的高分辨率遥感影像分类研究[J]. 光学学报, 2016, 36(4): 0428001.
- [18] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets [J]. Neural Computation, 2006, 18(7): 1527-1554.
- [19] Lv Y, Feng Z, Xu C. Facial expression recognition via deep learning [C]. International Conference on Smart Computing, IEEE, 2015: 303-308.
- [20] Liu P, Han S, Meng Z, *et al.* Facial expression recognition via a boosted deep belief network [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1805-1812.
- [21] Xia J, Pei D, Wang Q Z, *et al.* Face recognition based on local adaptive ternary derivative pattern coupled with Gabor feature [J]. Laser & Optoelectronics Progress, 2016, 53(11): 111004.
- 夏军, 裴东, 王全州, 等. 融合 Gabor 特征的局部自适应三值微分模式的人脸识别[J]. 激光与光电子学进展, 2016, 53(11): 111004.
- [22] Zhang Y, Liao Q Z, Luo Y. Head pose estimation fusing the second order HOG and CS-LBP[J]. CAAI Transactions on Intelligent Systems, 2015, 10(5): 741-746.
- 张毅, 廖巧珍, 罗元. 融合二阶 HOG 与 CS-LBP 的头部姿态估计[J]. 智能系统学报, 2015, 10(5): 741-746.
- [23] Hinton G E. Training products of experts by

- minimizing contrastive divergence [J]. *Neural Computation*, 2002, 14(8): 1771-1800.
- [24] Happy S L, Routray A. Robust facial expression classification using shape and appearance features [C]. *Eighth International Conference on Advances in Pattern Recognition*, 2015: 1-5.
- [25] Gu W, Xiang C, Venkatesh Y V, *et al.* Facial expression recognition using radial encoding of local Gabor features and classifier synthesis [J]. *Pattern Recognition*, 2012, 45(1): 80-91.
- [26] Rivera A R, Castillo J R, Chae O. Local directional number pattern for face analysis: face and expression recognition [J]. *IEEE Transactions on Image Processing*, 2013, 22(5): 1740-1752.
- [27] Ghimire D, Lee J. Extremelearning machine ensemble using bagging for facial expression recognition [J]. *Journal of Information Processing Systems*, 2014, 10(3): 443-458.
- [28] Zhong L, Liu Q, Yang P, *et al.* Learning multiscale active facial patches for expression analysis [J]. *IEEE Transactions on Cybernetics*, 2015, 45(8): 1499-1510.
- [29] Ahmed F, Bari H, Hossain E. Person-independent facial expression recognition based on compound local binary pattern (CLBP) [J]. *International Arab Journal of Information Technology*, 2014, 11(2): 195-203.
- [30] Li Y Q, Li Y J, Li H B, *et al.* Fusion of global and local various feature for facial expression recognition [J]. *Atca Optica Sinica*, 2014, 34(5): 0515001. 李雅倩, 李颖杰, 李海滨, 等. 融合全局与局部多样性特征的人脸表情识别 [J]. *光学学报*, 2014, 34(5): 0515001.
- [31] Ghimire D, Jeong S, Yoon S, *et al.* Facial expression recognition based on region specific appearance and geometric features [C]. *Tenth International Conference on Digital Information Management, IEEE*, 2016: 142-147.