

面向非特定表情的加权和稀疏分类方法^①

蒋行国, 冯 彬, 李志丰

(桂林电子科技大学 信息与通信学院, 桂林 541004)

摘 要: 针对在非特定人脸表情识别中, 表情纹理特征的利用率不高问题, 提出了一种改进的加权局部二值模式(LBP)和稀疏表示相结合的人脸表情识别方法. 为了有效利用面部器官的局部纹理信息, 采用改进的加权 LBP 算子提取人脸局部纹理特征, 然后用获取的特征值组成训练样本, 最后根据稀疏表示理论进行表情分类. 在 JAFFE 和 CK 人脸库上的实验结果表明, 该方法对非特定人脸表情的识别效果有了明显提高.

关键词: 人脸表情识别; 稀疏表示; 纹理特征提取; 局部二值模式(LBP)

Nonspecific Facial Expression Recognition via the Sparse Representation and Weighted

JIANG Xing-Guo, FENG Bin, LI Zhi-Feng

(School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract: On the person-independent facial expression recognition, the utilization rate of the facial expression texture is not high. Facing with the problem of the person-independent face, this paper proposes a method about facial expression recognition based on the improved weighted local binary pattern (LBP) and sparse representation. In order to use the local texture information of the facial organs effectively, first it uses the improved weighted LBP operator to extracting the local texture feature, the extracted features to construct the training samples, and classified via the sparse representation last. Experimental results show a better performance on the JAFFE and CK database.

Key words: facial expression recognition; sparse representation; texture feature exaction; local binary pattern(LBP)

人脸表情传递着丰富的个人情感信息, 并且在人与人之间的交流中扮演着重要的角色, 因此为了推动更加智能和自然的人机交互, 人脸表情识别在过去几十年得到广泛研究, 并且成为一个研究热点. 非特定人脸的特点决定了对它的研究有一定的难度和更深的意义. 表情识别主要分为两部分: 特征提取和分类器.

由于人脸表情的本质特征大多藏在图像的高维空间中, 把人脸信息映射到最有效的低维空间相当重要. 一般的, 特征提取方法用于表示人脸的整体特征或局部特征. 文献[1,2]证明局部特征表示在表情识别中有很好效果. 因此, 本文提出一种改进的加权 LBP 局部特征提取算法.

常用的识别分类器有最近邻分类(NNC)、线性支持向量机(SVM)^[3]、最近子空间分类(NSC)^[4]、线性判别分类(LDA)^[3]等. 近几年, 压缩感知中的稀疏表示

(SR)被不断地应用于信号处理领域. 随着 Ma^[5,6]等人提出了基于稀疏表示分类(SRC)的人脸识别, SRC 在人脸识别领域表现出强大的优势. Huang Ming-wei^[7]对 LBP+稀疏表示和稀疏分类的表情识别结果进行融合, 比传统的识别方法有更好的识别率、鲁棒性, 但未能有效的利用表情特征. 对于非特定人脸, 表情特征相对于人脸特征复杂, 在同个图像中, 人脸特征会干扰表情特征的处理, SRC 识别算法的性能也不理想. 如何有效地利用表情特征和 SRC 分类是提高非特定表情识别效果的关键.

针对上述的问题和非特定人脸问题, 本文提出一种改进的加权 LBP 和稀疏表示相结合的表情识别方法. 通过改进的加权 LBP 提取人脸主要器官的纹理特征, 并有效地给表情特征分配权重, 最终结合稀疏表示分类实现表情识别, 提高运算时间. 在 JAFFE 和

^① 收稿时间:2013-12-02;收到修改稿时间:2014-01-03

CK 人脸表情库的实验结果表明, 该算法具有较好的识别效果.

1 基于改进的加权 LBP 的特征提取

1.1 LBP 直方图

LBP 是一种灰度尺度不变的纹理描述算子, 可对图像局部邻近区域的纹理信息进行度量 and 提取. 最初 LBP 算子由 Ojala^[8]等人于 1996 年提出, LBP 能够对人脸图像细小模块信息的纹理特征进行有效编码, 已经被引入到人脸图像分析领域中描述人脸. LBP 方法首先计算图像中某点像素与该点领域在灰度上的二值关系, 然后对二者关系按一定规则加权形成像素的 LBP 值, 如式(1)所示:

$$LBP(P, R) = \sum_{n=0}^{n=P-1} s(g_n - g_c) 2^n \quad (1)$$

其中 P 表示领域里的像素点数, R 表示领域半径, g_c 是中心像素点, $g_n (n=0, 1, \dots, P-1)$ 是 g_c 的近邻点, 函数 $s(x)$ 的定义为:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

下面以一个 3×3 邻域为例简要介绍原始 LBP 算子的计算原理, 如图 1 所示. 其中原始 LBP 算子的参数为 $P=8, R=1$. 如果邻域像素大于中心像素值, 那么对应的二进制编码给定 1; 反之, 为 0. 然后将得出的二进制数按顺时针连起来, 将二进制数转换成十进制数, 从而得到了中心像素的 LBP 值.

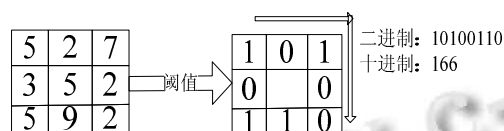


图 1 原始 LBP 算子的计算原理

把图像 $f(x, y)$ 的 LBP 值统计, 组成 LBP 特征直方图, 公式 $g(A)$ 表示为:

$$H_i = \sum_{x,y} g\{f(x, y) = i\}, i = 0, 1, \dots, m-1 \quad (3)$$

其中 m 表示直方图的 bins 数目 (即在原始算子中 $m=2^P=256$), 函数定义如下:

$$g(A) = \begin{cases} 1, & A \text{ 为真} \\ 0, & A \text{ 为假} \end{cases} \quad (4)$$

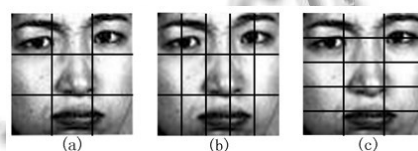
经过上式(3)统计出的 LBP 直方图, 包含了原图像大量的局部特征信息, 但不能区分哪些是表情特征, 仅仅依靠这些局部特征, 不能达到好的识别效果. 人

脸不同面部器官对表情的贡献不同, 眼睛、眉毛和嘴巴对表情分类有着关键作用, LBP 并不能直接提取这些器官的纹理特征.

1.2 改进的加权 LBP

针对以上所描述的 LBP 存在的不足, 对人脸图像进行改进的加权 LBP 特征提取. 本文把所要处理的图像均匀地划分为几个非重叠的小方格 (子图像), 分别对每个子图像进行 LBP 直方图统计. 文献^[9]指出对原始的人脸图像进行均匀划分, 不仅能取得很好地效果, 而且在不增加算法复杂性的同时提高表情特征的提取率.

常用的三种不同划分方法, 分别是 3×3 、 3×5 、 5×3 等多种方格划分法. 图 2 为 JAFFE 表情库^[10]单张人脸的三种划分情况, 从左至右, 依次为 3×3 划分, 3×5 划分, 5×3 划分. 由于人脸的特殊结构, 不同的方格划分方式所得到的 LBP 直方图统计信息存在很大的差异. 很多划分方法存在不够局部化, 不能有效地描绘出方格里包含的眼睛、鼻子、嘴巴等主要特征等问题. 本文使用图 2(b) 的划分方式 3×5 , 它符合面部 “三庭五眼” 的结构特征, 能够得到较好的实验结果. 这样, 每一个子图像块都包含了块内的结构信息, 诸如眼睛、鼻子、嘴巴之类的; 这些子图像结合又体现了它们之间的拓扑关系信息, 这些特征信息均有利于后续的分类识别.



(a) 3×3 划分 (b) 3×5 划分 (c) 5×3 划分

图 2 JAFFE 单张人脸划分示意图:

如何有效地利用子图像的信息? 通常地, 人为地给人脸中的每个子图像分配权重, 但权值的选择并不能有效地放映表情特征的局部信息和子图像间的结构关系. 本文通过度量多种表情种类间的子图像的纹理信息差异, 找出判别表情种类的主要的子图像, 即分析每个子图像的分类性能, 自适应地给每个子图像分配权重. 文中, 借助卡方分布 (χ^2) 来测量两个直方图的差异来分析权重. 通过下式(5)测量两个直方图的差异, 又称为卡方距离.

$$\chi^2(S, M) = \sum_i \frac{(S_i - M_i)^2}{S_i + M_i} \quad (5)$$

其中 S 和 M 分别表示不同图像的直方图. 用 S 表示一张人脸图像, 经 $n_1 \times n_2$ 划分后的子图像矩阵用 S_i 表示, 其中 $i=1, 2, \dots, n_1 \times n_2$ 分别按从左到右从上到下的顺序表示每一个子图像. M_i 为该个人脸的 7 种表情图像中的第 i 个子图像的平均值. 对于每一个子图像, 通过卡方距离来分析这个子图像的分类性能 w_i , 如下式(6)所示.

$$w_i = \frac{1}{N} \sum_{j=1}^N \chi^2(S_{i,j}, M_i) \quad (6)$$

其中, N 为样本的种类数, $\chi^2(S_{i,j}, M_i)$ 表示两个图像对应的第 i 个子图像直方图之间的卡方距离. 类间的差异程度影响分类效果, 经常被用来分类. 卡方距离越大, 则此子图像类间差别极大, 即它的信息对表情的分类有很大影响.

通过上式(6), 计算出每个子图像对应的权值, 能自适应地根据训练人脸的信息给图像中的子图像分配权值. 通过子图像的权重找出人脸的一些区域比其他区域包含着更多表情分类的有用信息. 使用自适应加权算法, 筛选出对表情分类有利的表情特征, 降低与表情无关纹理的影响. 有效地结合分类器利用这些表情特征的权值, 能提高识别性能.

2 稀疏表示分类器

在 1.2 节中, 在每张人脸表情中, 分别求得对应子图像的 LBP 值后, 乘以其对应的 w_i 得到特征值 a , 把 a 按列排放, 构成 $A=[a_{1,1} \ a_{1,2} \ \dots \ a_{i,1} \ a_{i,2} \ \dots \ a_{i,n} \ \dots \ a_{7,n}]$, 其中 i 表示类别, n 为该类的图像个数. 根据稀疏表示分类理论, 测试样本 y 可以表示为: $y = Ax$; 其中由训练样本构成. 给出一个属于第 i 类的测试样本 y , 那么 y 可以被训练样本张开的空间中线性表达出来.

$$y = a_{i,1}v_{i,1} + a_{i,2}v_{i,2} + \dots + a_{i,n}v_{i,n} \quad (7)$$

测试样本 y 表示为来自同一类别训练样本的线性组合, 则解 $x_1 = [0, \dots, 0, a_{i,1}, a_{i,2}, \dots, a_{i,n}, 0, \dots, 0]$, 只有第 i 类的值是非 0 元素, 其他类别的系数为零或绝对值接近零. 系数向量 a 中包含了大量有利于分类的信息. 因此 SRC 的关键在于求解稀疏向量 x_1 . 由压缩感知理论^[11]可知, x_1 中非零稀疏越少, 越容易判断 y 的种类, 这是一个 NP 难问题. 相关文献^[12]指出, 最小 l_0 范数解和最小 l_1 范数解在某种情况下等效. 因此, 这个问题可以通过求解最小 l_1 范数问题来解决, 即:

$$x_1 = \arg \min \|x\|_1 \quad \text{s.t.} \quad Ax = y \quad (8)$$

测试样本 y 所属类别的公式为:

$$\min r_i(y) = \|y - A\delta_i(x_i)\|_2 \quad (9)$$

其中 $\delta_i(x_i) \in R^n$, 是第 i 组系数中非零的数为系数 x_1 中与 i 对应的那些数. $r_i(y) = \|y - A\delta_i(x_i)\|_2$ 表示的是测试样本 y 与线性表示 $A\delta_i(x_i)$ 的残差值, 我们认为残差值最小的对应类别 i 为测试样本 y 的类别.

综上所述, 稀疏表示分类的步骤总结如下:

①对表情库的每副人脸图像进行改进的加权 LBP 特征提取.

②用特征提取后的训练样本组成矩阵 A .

③利用 L1_LS 算法求解 l_1 范数问题:

$$x_1 = \arg \min \|x\|_1 \quad \text{s.t.} \quad Ax = y \quad (8)$$

④计算残差值: $r_i(y) = \|y - A\delta_i(x_i)\|_2$ 其中 $i=1, 2, \dots, k$ (10)

⑤判定测试样本的类别:

$$\text{identify}(y) = \arg \min_i r_i(y) \quad (11)$$

上述步骤中, 第 3 步就是求解稀疏解 x_1 , 稀疏表示理论指出在一定条件下可通过求解 l_1 范数问题来求解稀疏解. 其中求解 l_1 范数问题的解法有很多, 例如 L1-Magic, L1-LS, CVX 等. 本文使用的是 L1_LS 解法. 第 4 步中, k 为类别总数, 根据实际情况我们可以在式子(10)中再减去一个误差值, 使得计算残差值更精确. 最后找出最小的残差值 $r_i(y)$, 确定测试样本的类别 i .

3 实验结果与分析

实验环境: 实验主要是在 PC 机用 Matlab 进行仿真. PC 的配置为处理器 Intel Core i5-2450M(双核 2.50GHz), 内存 4GB, 硬盘 500GB, 显卡 NVIDIA GeForce GT 525M. 软件版本: 微软的 WIN7 的 32 位操作系统, Matlab R2011b.

3.1 在 JAFFE 人脸数据库上的实验

对人脸库的图像进行几何归一化、灰度归一化、滤波等预处理. 反复实验证明, 图像大小为 64×64 时, 识别效果最好. JAFFE 人脸图像经过预处理后大小为 64×64 , 取 KA 的 7 种表情预处理后的图显示如下图 3, 从左到右依次为厌恶、恐惧、开心、自然、伤心、惊奇、愤怒 7 种表情:

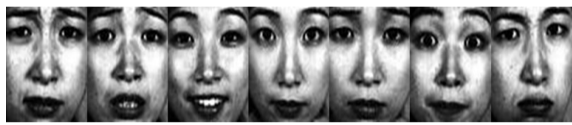


图 3 KA 的 7 种表情预处理后的图像

JAFFE 人脸库是日本的女性人脸表情库, 它含有 7 种表情类别、10 个人, 每个人每种表情有 3 张图片. 本实验针对的是非特定人的表情识别. 每次从 10 个人中取出其中 9 个人的图像进行训练, 余下 1 人的表情为测试样本, 取得实验结果并重复进行 10 次, 确保每人都能单独成为测试样本, 将 10 次实验结果取平均作为最终识别结果. 通过多次实验比较了 3 种识别方法, SRC、SRC+LBP 和本文的算法, 结果如表 1 中所示. 其中, PCA+SVM 和 LDA+SVM 算法在文献^[13]中提及, 是两种最经典的表情识别算法.

表 1 几种不同算法在 JAFFE 库的比较

算法	表情识别率(%)	时间(s)
PCA+SVM	56.3	
LDA+SVM	58.2	
SRC	60.05	1.37
SRC+LBP	60.48	0.46
本文算法	67.14	0.51

由表 1 可看出, SRC、SRC+LBP 比经典表情识别算法 PCA+SVM 和 LDA+SVM 在识别率有提高, 而本文的算法比 SVM 算法有很大的提高. 在 JAFFE 库里的表情识别方面, SRC 和 SRC+LBP 算法的识别率差不多, 但 SRC 识别时间比 SRC+LBP 要长, 本文的算法在这几种算法中, 识别率最高, 识别时间接近 SRC+LBP. 故本文的算法识别性能比较好.

3.2 在 CK 人脸数据库上的实验

本文所使用的 Cohn-Kanade 人脸库含有 18 个人、7 种表情, 每个人每种表情有 5 张图片. 本实验同样针对的是非特定人的表情识别. 每次从 18 个人中取出其中 17 个人的表情图像作为训练样本, 剩下 1 个人的表情作为测试样本, 取得实验结果并重复进行 18 次, 且确保每个人都能单独成为测试样本, 将 18 次实验结果取平均作为最终识别结果. 比较了 3 种识别方法: SRC、SRC+LBP 和本文的算法, 结果如表 2 中所示.

表 2 几种不同算法在 CK 库的比较

算法	表情识别率(%)	时间(s)
SRC	52.86	8.22
SRC+LBP	53.06	3.02
本文算法	56.23	3.46

从表 1 和表 2 可以看出 SRC+LBP 与 SRC 的识别率差不多, SRC+LBP 的识别时间要短很多, 但 SRC+LBP 算法没能充分利用局部纹理特征. 本文的算法的识别率比 SRC 和 SRC+LBP 有了明显的提高, 它能提高纹理特征的有效性, 通过图像分块加权的方式强化表情种类的分类信息, 而且本文的算法识别速度接近 SRC+LBP. 因此, 本文算法在非特定人脸表情识别上能有较好的识别效果. 使用 SRC 分类器中, 训练样本的个数越多, 运算的时间越长, 所以在 CK 库的实验时间比 JAFFE 库的要长.

综上所述, 本文提出的分块加权 LBP 和稀疏表示相结合的表情识别算法, 能有效地利用人脸表情的局部特征, 提高局部特征的有效率, 结合稀疏表示分类器提高了表情识别性能. 本文的算法比 SRC 和 SRC+LBP 算法的识别效果好.

4 结语

本文针对表情识别领域的非特定人脸问题, 提出了一种改进的加权 LBP 和稀疏表示相结合的算法. 该算法提高局部纹理特征的有效利用率, 通过改进加权的 LBP 自适应地加强重要器官的表情分类信息, 最后结合稀疏表示分类器实现非特定的人脸表情识别. 通过实验结果分析, 本文的算法相比 LBP+SRC 和 SRC 算法的识别性能有了极大的提升.

参考文献

- 1 Buciu I, Pitas I. Subspace image representation for facial expression analysis and face recognition and its relation to the human visual system. Organic Computing of Springer Verlag, 2008: 303–320.
- 2 Zhi RC, Ruan QQ, Wang ZF. Facial expression recognition via sparse representation. IEICE Trans. on Information and Systems, 2012, 9(9): 2347–2350.
- 3 Tan XY, Chen SC, Zhou ZH. Face recognition from a single image per person: a survey. Pattern Recognition, 2006, 39(9): 1725–1745.

- 4 Ho J, Yang M, Lim J, et al. Clustering appearances of objects under varying illumination conditions. IEEE International Conference On Computer Vision And Pattern Recognition. 2001. 11–18.
- 5 Wright J, Ma Y, et al. Robust Face Recognition via Sparse Representation. Pattern Analysis And Machine Intelligence, 2009, 31(2): 210–217.
- 6 Jia K, Chan TH, Yi M. Robust and practical face recognition via structured sparsity. European Conference on Computer Vision (ECCV), 2012.
- 7 Huang MW, Wang ZW, Ying ZL. A new method for facial expression recognition based on sparse representation plus LBP. The 3rd International Congress on Image and Signal Processing. Yantai. 2010. 1750–1754.
- 8 Ojala T, Pietikainen M, Harwood D. A comparative study of texture measures with classification based on feature distributions. Pattern Recognition, 1996, 29(1): 51–59.
- 9 Ahonen T, Hadid A, Pietikainen M. Face recognition with local binary patterns. Proc. of Computer Vision Conference(ECCV 2004). Berlin, Springer. 2004. 469–481.
- 10 Lyons MJ, Miyuki K, Jiro G. Japanese Female Facial Expressions (JAFFE). Database of digital images(1997). <http://www.kasrl.org/jaffe.html>. 1997.
- 11 焦李成, 杨淑媛, 刘芳, 等. 压缩感知回顾与展望. 电子学报, 2011, 39(7): 1651–1662.
- 12 Candès EJ, Wakin M. An Introduction to Compressive Sampling. Signal Processing Magazine, 2008, 25(2): 21–30.
- 13 Ying ZL, Li JW, Zhang YW. Facial expression recognition based on 2 dimensional feature extraction. ICSP2008. Beijing. 2008. 1440–1444.