

文章编号: 2095-2163(2020)05-0052-04

中图分类号: TP391

文献标志码: A

融合 CLBP 与几何显著特征的人脸情感识别

王从澳, 黄润才, 孙延标, 杨 彬, 孙刘成

(上海工程技术大学 电子电气工程学院, 上海 201600)

摘要: 针对局部二值模式(LBP)对非一致性光照敏感和随机噪声不稳健等问题, 提出融合完整局部二值模式(CLBP)与几何显著特征的情感识别算法。利用 Dlib 库提取面部 68 个特征点, 通过选取显著特征点计算中立及巅峰时刻欧几里德距离, 构建特征比例向量。采用 CLBP 提取面部细粒度纹理特征, 串联融合其特征直方图, 利用随机森林进行表情分类。实验表明该算法在 CK+数据集上具有 92.8% 的识别准确率, 优于传统表情识别算法。

关键词: CLBP; 纹理特征; 随机森林; 表情分类

Facial emotion recognition based on CLBP and geometric salient features

WANG Congao, HUANG Runcai, SUN Yanbiao, YANG Bin, SUN Liucheng

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201600, China)

[Abstract] Aiming at the problems of local binary pattern (LBP) being sensitive to non-uniform illumination and random noise instability, this paper proposes an emotion recognition algorithm that combines complete local binary pattern (CLBP) with geometrically significant features. Firstly, 68 feature points of the face were extracted using the Dlib library, and the feature vector was calculated by selecting salient feature points to calculate the Euclidean distance at the neutral and peak times. Then use CLBP to extract facial fine-grained texture features, and then fuse the feature histograms in series to use random forest for facial expression classification. Experiments show that the algorithm has a recognition accuracy of 92.8% on the CK + dataset, which is better than the traditional expression recognition algorithm.

[Key words] CLBP; texture features; random forest; expression classification

0 引言

面部情感识别是计算机视觉领域的一个重要研究方向。近年来, 由于其在医疗保健、视频监控、辅助驾驶以及人机交互等方面的广泛应用, 引起了越来越多的关注。在识别过程中, 根据不同的特征表示向量, 人脸情感识别可以分为两大类: 静态图像的人脸情感识别 (facial expression recognition, FER) 和动态序列的人脸情感识别。在静态 FER 中, 特征表示使用单个图像的空间信息进行编码, 而动态的方法则是考虑输入面部表情序列中连续帧之间的时间关系^[1]。除了这两种传统的基于视觉的方法, 还有其他基于音频、生理信号等多模态信息的方法进行面部表情的辅助识别。

多数传统的表情识别使用手工提取特征或浅层机器学习的方法, 如: 局部二进制模式 (local binary pattern, LBP), 三个正交平面的 LBP (LBP-TOP), 非负矩阵分解和稀疏学习等。有人使用主成分分析 (principal components analysis, PCA) 和独立成分

分析 (independent component analysis, ICA) 识别不对称面部单元的情感表达; M. Lyons 等人使用 gabor 滤波器实现多分辨率和多方位的面部表情识别; 还有人通过级联 4 个支持向量机, 使用主动外观模型 (active shape model, AAM) 对 4 种基本表情进行分类。

在人脸表情识别中, 显著区域的几何特征变化和局部的细粒度纹理特征都是面部表情的重要标识, 基于 LBP (局部二值模式) 特征提取存在的光照敏感和噪声不稳健, 几何特征选取复杂, 信息冗余等问题。本文提出融合 CLBP 与几何显著特征的人脸表情特征提取方式, 通过串联融合其特征向量直方图, 构建完整的表情特征向量, 利用随机森林分类器对人脸面部表情进行分类。

1 人脸情感识别框架

本文提出的融合 CLBP 与几何显著特征的人脸情感识别算法实现过程如图 1 所示。首先, 对输入人脸面部表情图像进行数据预处理, 包括图像旋转、

作者简介: 王从澳 (1995-), 男, 硕士研究生, 主要研究方向: 情感识别、图像处理; 黄润才 (1966-), 男, 博士, 副教授, 主要研究方向: 智能计算、计算机网络与应用; 孙延标 (1995-), 男, 硕士研究生, 主要研究方向: 多机器人系统、网络通信; 杨 彬 (1996-), 男, 硕士研究生, 主要研究方向: 多信息融合、图像处理; 孙刘成 (1993-), 男, 硕士研究生, 主要研究方向: 文本挖掘、深度学习。

收稿日期: 2020-03-13

裁剪以及尺寸归一化与灰度归一化等操作, 根据面部稳定点间距作为缩放比例系数, 进行尺寸归一化

标定, 同时利用直方图均衡化增强面部区域的对比度, 进一步消除光照强度的不利影响。

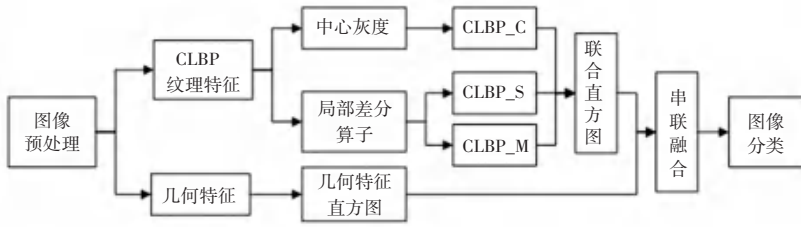


图 1 识别算法过程图

Fig. 1 Recognition algorithm flowchart

预处理完成后, 对表情图像进行特征向量的提取。(1)几何特征:使用 Dlib 库对人脸的 68 个特征点进行标记, 根据其几何显著特征点“中立”与“巅峰”时刻的欧几里得距离, 提取几何特征的特征直方图。(2)纹理特征:采用 CLBP 对面部纹理特征进行提取, 分别得出 CLBP_C, CLBP_S, CLBP_M 三部分特征直方图, 通过并联的形式构建面部完整纹理特征向量, 即联合直方图。(3)使用串联融合的方式对面部几何特征与纹理特征提取的直方图向量进行融合, 生成完整面部表情特征向量。(4)使用随机森林分类器对表情图像分类, 得出最终分类结果。

1.1 几何特征提取

表情图像经预处理完成后, 利用 Dlib 库标记出人脸眉毛, 眼睛, 鼻子, 嘴唇和脸部轮廓上的 68 个特征点坐标, 为提取脸部情感特征做准备。人脸面部 68 个坐标点标记如图 2 所示。

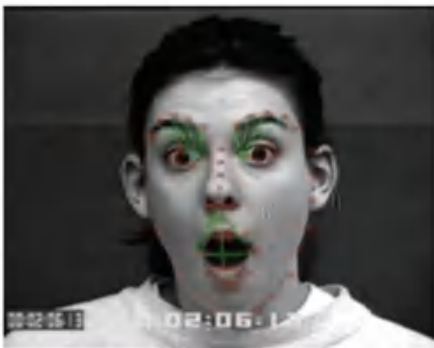


图 2 人脸 68 个特征点

Fig. 2 68 feature points on the face

相关实验研究表明, 人在经历不同的情绪变化时, 眉毛的变化幅度、嘴的大小宽高比例以及眼睛的变化幅度能够较准确的体现情感特征。因此本文利用左右两边的眉毛、眼睛和嘴部轮廓变化作为主要特征提取部位, 选取左右内眼角到眉毛 4 个坐标点的距离之和、左右眼部轮廓开合距离、鼻子到左右嘴角轮廓、嘴部高度、宽度等作为 8 个主要特征向量,

分别计算其“中立”与“巅峰”时刻特征值的欧几里得距离(记为 $d_1, d_2, d_3 \dots d_8$)。

本文几何特征构造是通过选取几何不变特征点来构造尺寸, 旋转和位移不变性的比例特征向量, 因此选取左眼角到眉头的距离 L1, 右眼角到眉头的距离 L2, 鼻尖到上嘴唇的距离 L3, 作为特征基准距离, 分别计算 8 个主要特征向量的比例特征向量, 将其作为模型训练输入, 构造的 8 个特征比例向量如表 1 所示。

表 1 构造的 8 个特征向量

Tab. 1 Constructed 8 feature vectors

特征向量	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8
人脸部位	左眉	右眉	左眼	右眼	左嘴角	右嘴角	嘴高	嘴宽
特征比例	d_1/L_1	d_2/L_2	d_3/L_1	d_4/L_2	d_5/L_3	d_6/L_3	d_7/L_3	d_8/L_3

最终用 8 个特征向量 $T = (T_1, T_2, T_3 \dots T_8)$ 表示不同人脸部位显著特征区域的特征比例向量, 对特征向量进行向量归一化处理, 消除数量级差异, 用以构建几何特征直方图。

1.2 CLBP 纹理特征提取

对于人脸的纹理特征提取, 主要有主成分分析法 (PCA)、线性鉴别分析法 (LDA) 和局部二值模式 (LBP) 等。LBP 作为一种经典的面部纹理特征提取算法, 通过比较中心像素点与领域像素点的差值来描述图像的局部结构, 但通常在提取特征时会出现非一致性光照和随机噪声不稳健等问题。因此, 本文提出使用完整局部二值模式 (CLBP) 进行面部纹理特征提取, 该算子由两部分组成: (1) 中心像素 LBP 算子 (CLBP_C); (2) 局部差分符号数值变换算子 (LDSMT): ①符号 LBP 算子 (CLBP_S); ②梯度 LBP 算子 (CLBP_M)。

完整局部二值模式 (CLBP) 算子的各模块如图 3 所示, (a) 图表示 3×3 采样块图像, (b) 图是局部差分符号数值变换算子 (LDSMT), (c) 图表示差分算子的符号向量 (CLBP_S), (d) 图表示差分算子的梯度向量 (CLBP_M)。

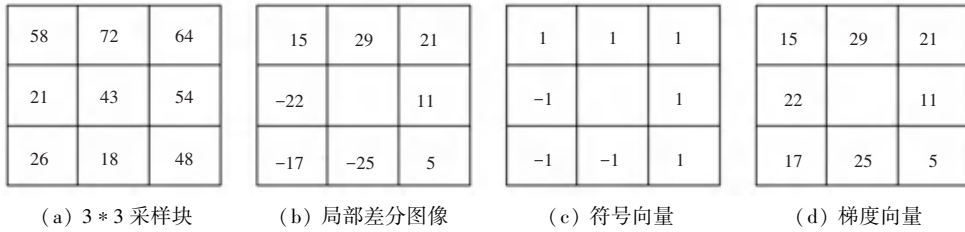


图3 CLBP算子

Fig. 3 CLBP operator

(1) 中心像素的 LBP 算子 (CLBP_C)

$$CLBP_C_{p,R} = t(g_c, c_i). \quad (1)$$

其中, g_c 表示中心像素灰度值, c_i 表示采样块图像的中心领域灰度值。

$$t(x, c) = \begin{cases} 1, & x \geq c; \\ 0, & x < c. \end{cases} \quad (2)$$

其中, x 表示各邻域像素值, c 表示采样块中心像素, 比较中心像素与周围邻域像素的数值, 大于中心像素的灰度置 1, 小于中心像素的置 0, 然后进行二进制编码。

(2) 局部差分符号数值变换算子 (LDSMT)

$$LDSMT_p = S_p * M_p, \quad (3)$$

$$S_p = \text{sign}(I_p - I_c), M_p = |I_p - I_c|, \quad (4)$$

其中, $I_p - I_c$ 表示差分变换算子, S_p 表示符号向量, M_p 表示梯度向量, 因差分变换算子不能直接作为特征描述子, 故使用 S_p 和 M_p 两部分表示。

$$S_p = \begin{cases} 1, & I_p \geq I_c; \\ -1, & I_p < I_c. \end{cases} \quad (5)$$

CLBP_S 与 CLBP_C 计算类似, 比较中心像素与周围邻域像素的大小, 大于中心像素的灰度值置 1, 小于中心像素的置 -1。

$$CLBP_M_{p,R} = \sum_{p=0}^{P-1} t(M_p, C) 2^p. \quad (6)$$

$$t(x, c) = \begin{cases} 1, & x \geq c; \\ 0, & x < c. \end{cases} \quad (7)$$

其中, c 表示设定阈值, 该算法中阈值使用 3*3 采样块图像的平均灰度值。

最终得到三个特征描述符 CLBP_C, CLBP_S, CLBP_M, 通过级联的形式将其融合, 得到纹理特征直方图。

1.3 特征融合

在几何特征提取中, 首先计算几何显著特征点间的距离, 然后计算其不同部位的比例特征向量, 最后对特征向量进行转换和编码。为了最大程度让提取的特征值能够体现某一特定情绪的特点, 算法通过计算人脸“中立”与“巅峰”时刻两组特征点的欧几里得距离作为变量, 表示采样点的特征信息。变量转换如(8)式:

$$\beta = T_{\max} - T_{\min}. \quad (8)$$

对转换变量按照 CLBP 模式进行编码:

$$t(x, e) = \begin{cases} 1, & x \geq e; \\ 0, & x < e. \end{cases} \quad (9)$$

其中, e 表示几何显著特征点转换变量 β 的平均值, 最后计算所有特征比例向量的特征直方图。

特征直方图的融合通常包含串联和并联两种形式。经实验验证, 在该算法中, 串联直方图形式具有更高的准确率和鲁棒特性。因此, 本文采用串联直方图的形式将几何特征直方图与 CLBP 提取的纹理特征直方图进行串联融合, 将融合后的特征描述符作为分类模型进行训练^[2]。

1.4 随机森林分类

本文使用随机森林分类器对融合后的特征描述符进行分类。随机森林作为一种典型的集成学习方法, 通过构建多棵决策树投票以获得更准确和稳定的预测或分类。使用引导重采样, 从原始样本中抽取决策树, 并使用特征的随机子集对决策树进行训练。随机森林中回归树生成过程如下:

假设有 $D = \{(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)\}$ 数据集, 目的是找到对应函数 $f: X \rightarrow Y$, 其中 X 表示输入, Y 表示输出。

(1) 随机森林从 D 中随机选择 n 个观测值进行替换以形成重采样引导样本。

(2) 每棵树根据 M 个特征中 m 个特征子集进行生长, 这里将特征子集设置为 $M = 3$, 然后在每个节点上随机选择 m 个特征, 并根据基尼纯度选择 m 个特征中性能最好的进行分割。

(3) 决策树不需要剪枝即可生长到最大深度。

通过加权各决策树的投票结果进行总体预测, 与其他基于决策树的集成学习方法不同, 随着决策树的数目递增, 随机森林不易发生过拟合, 且特征的随机选择能够最大化降低决策树之间的相关性, 提升预测结果准确率。

2 实验与结果分析

2.1 实验环境

本实验程序的电脑运行配置为 Windows10 专业

版操作系统,采用 2.7GHz. Inter 酷睿 CPU 处理器, 16GB 内存。

实验将 CK+数据库中 70%的表情图像作为训练集,其余 30%作为测试集,将预处理后的表情图像作为特征提取的输入图像,分别提取其 CLBP 纹理特征和几何显著特征,融合其特征向量直方图后送入随机森林分类器进行训练。

2.2 实验结果分析

实验对比了几何特征,CLBP 纹理特征和融合特征提取等方法在 CK+数据库上的识别准确率,分类结果分别如表 2,表 3,表 4 所示,其中粗体字表示正确的分类结果。

表 2 几何显著特征提取的混淆矩阵

Tab. 2 Confusion matrix for geometric salient feature extraction

表情	愤怒	失望	高兴	悲伤	惊讶	恐惧	中立
愤怒	78.5	4.3	0	9.2	0	4.3	3.8
失望	4.0	83.6	0	2.3	3.7	2.7	3.6
高兴	1.7	0	95.2	0	0	3.0	0
悲伤	2.1	4.3	0	88.9	0	2.3	2.3
惊讶	0	0	2.3	0	96.2	1.5	0
恐惧	6.4	1.5	0	7.4	0	76.6	8.0
中立	12.3	9.8	0	3.4	0	1.8	72.8

表 3 CLBP 纹理特征提取的混淆矩阵

Tab. 3 CLBP texture feature extraction confusion matrix

表情	愤怒	失望	高兴	悲伤	惊讶	恐惧	中立
愤怒	83.3	1.0	1.2	6.0	0	3.0	5.4
失望	2.5	82.5	2.6	2.4	0	6.7	3.4
高兴	2.1	0	94.0	0	2.3	1.6	0
悲伤	3.0	4.2	1.6	90.2	0	0	0.9
惊讶	0	0.6	1.6	0	97.1	0.8	0
恐惧	1.6	4.6	3.0	5.4	3.0	82.0	0.4
中立	2.8	3.0	0	7.5	0	6.9	79.9

表 4 融合特征提取的混淆矩阵

Tab. 4 Confusion matrix for feature extraction

表情	愤怒	失望	高兴	悲伤	惊讶	恐惧	中立
愤怒	88.4	0.5	0	4.6	0	2.4	4.0
失望	0.2	92.5	0	2.1	0.6	1.0	3.6
高兴	0	0	100	0	0	0	0
悲伤	3.5	0	1.2	93.0	0	0	2.1
惊讶	0.6	0	0.6	0	96.7	2.2	0
恐惧	0	0.5	1.2	2.0	0.6	92.0	3.6
中立	4.5	3.2	0	3.5	0	2.0	86.7

为了将本文提出的人脸表情识别算法与最新方法进行比较,采取了留一法交叉验证的设计方案,同样也可以在随机森林内部对误差的无偏估计进行计算,从而达到相同的实验对比效果。

由实验数据可以看出,几何显著特征与 CLBP 纹理特征融合后的表情图像分类比单一几何特征或纹理特征进行识别的准确率高。除了惊讶表情外,融合特征对失望,恐惧,中立的识别准确率均明显高于几何显著特征和纹理特征,且各表情平均识别准

确率高出几何特征 8.3%,高出纹理特征 5.8%。本文算法的综合识别准确率高达 92.8%,结果高于其他类似表情识别算法。表 5 是本文算法与其他算法的性能对比。

表 5 不同算法在 CK+上的性能对比

Tab. 5 Performance comparison of different algorithms on CK +

论文来源	算法	准确率/%
Li et al ^[3]	LBP+BP	91.35
Fa et al ^[4]	Hog+光流	83.7
Li and Lam ^[5]	DBM+SVM	86.8
Shan C ^[6]	LBP+SVM	88.9
Ghi et al ^[7]	GF+LBP+SVM	90.08
本文算法	GF+CLBP+RF	92.8

3 结束语

本文提出了融合 CLBP 与几何显著特征的特征提取方法。利用 Dlib 库对人脸 68 个关键点进行标注,根据人脸表情变化显著区域构建特征比例向量,融合几何显著特征和 CLBP 提取的细粒度纹理特征作为输入特征向量进行表情分类。实验证明,该算法在 CK+数据库上具有较高的识别准确率和鲁棒特性,优于传统的单一特征提取人脸表情识别算法。

下一步的研究计划将在本文算法基础上,进一步完善特征提取方式,融入深度特征描述符,将深度卷积神经网络提取的面部深度特征与人工提取特征相融合,更完整的提取出人脸表情变化的显著特征信息,从而提升表情识别的准确率。再利用不同场景下多个标准人脸表情数据库进行试验验证和测试,使模型具有更好的泛化能力和鲁棒特性。

参考文献

- [1] ZHAO G, PIETIKAINEN M. Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2007, 29:915-928.
- [2] 王晓华,李瑞静,胡敏,等. 融合局部特征的面部遮挡表情识别 [J]. 中国图象图形学报, 2016, 21(11):1473-1482.
- [3] 李敦宇,薛雷,张旭. 融合 LBP 卷积神经网络的面部表情识别 [J]. 工业控制计算机, 2018, 31(05):69-70.
- [4] FAN X, T. A spatial-temporal framework based on histogram of gradients and optical flow for facial expression recognition in video sequences [J]. Pattern Recognition, 2015, 48(11):3407-3416.
- [5] LI J, LAM E Y. Facial expression recognition using deep neural networks [C]//2015 IEEE International Conference on Imaging Systems and Techniques (IST). IEEE, 2015: 1-6.
- [6] SHAN C, GONG S, MCONWAN P W. Robust facial expression recognition using local binary patterns [C]//IEEE International Conference on Image Processing, Genova, Italy, 2005:2-370-3.
- [7] 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.