

基于分块 CBP 特征和稀疏表示的三维人脸表情识别^①



马 杰, 蔡轶珩, 盛 楠

(北京工业大学 信息学部, 北京 100124)

通讯作者: 蔡轶珩, E-mail: caiyiheng@bjut.com

摘 要: 在三维人脸表情识别中, 基于局部二值模式 (LBP) 算子算法与传统的特征提取算法相比具有特征提取准确、精细、光照不变性等优点, 但也有直方图维数高、判别能力差、冗余信息大的缺点. 本文提出一种通过对整幅图像进行多尺度分块提取 CBP 特征的 CBP 算法, 能够更有效的提取分类特征. 再结合使用稀疏表达分类器实现对特征进行分类和识别. 经实验结果表明, 与传统 LBP 算法和 SVM 分类识别算法对比, 文中算法用于人脸表情的识别的识别率得到大幅度提高.

关键词: LBP; CBP; 特征提取; 稀疏表; 特征分块提取

引用格式: 马杰, 蔡轶珩, 盛楠. 基于分块 CBP 特征和稀疏表示的三维人脸表情识别. 计算机系统应用, 2019, 28(2): 196-200. <http://www.c-s-a.org.cn/1003-3254/6692.html>

3D Facial Expression Recognition Using Block CBP Features and Sparse Representation

MA Jie, CAI Yi-Heng, SHENG Nan

(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: In 3D facial expression recognition, operator algorithm based on LBP has the characteristics of accurate, precise, and invariable illumination, while it also has the disadvantages of high dimensions of histogram, poor discriminating ability, and large redundant information, comparing with traditional feature extraction method. In this study, a CPB algorithm is proposed based on multi-scaled block CBP feature extraction system for the classification of facial expressions that are represented in 3D, designating to extract features more efficiently. Then, sparse expression classifier is used to classify and identify the features. The experimental results demonstrate that the facial expression recognition rate has been greatly improved, comparing with traditional LBP algorithm and SVM algorithm.

Key words: LBP; CBP; feature extraction; sparse expression; feature block extraction

三维人脸表情识别是现在计算机视觉领域研究的热点问题, 在人机交互, 医疗保健, 公共安全等领域都有着重要的应用. 现有的三维人脸表情识别算法主要包括特征提取和特征分类两个主要部分. 常用的特征提取算法有 PCA^[1], LDA, 活动外观模型 (Active Appearance Model, AAM)^[2], Gabor 小波变换^[3], 局部二

元模式 (Local Binary Pattern, LBP)^[4]等. 其中, 基于 AAM 的方法虽然能可靠地提取人脸表情特征, 但具有计算复杂、初始参数难以确定等缺点; 基于 Gabor 小波变换的方法为了提取多尺度和方向信息, 使用了多个滤波通道对人脸图像进行卷积运算, 非常耗时间和耗内存, 实时性不高, 而且提取的 Gabor 特征向量的维

① 基金项目: 国家重点研发计划课题 (SQ2017YFC170323); 北京市教委科技项目 (2017YFC1703302)

Foundation item: National Key Research and Development Program of China (SQ2017YFC170323); Science and Technology Program of Education Committee of Beijing Municipality (2017YFC1703302)

收稿时间: 2018-05-22; 修改时间: 2018-06-15; 采用时间: 2018-06-28; csa 在线出版时间: 2019-01-28

数很大,存在大量的冗余,不利于后续的分类识别;基于LBP的特征提取方法以其计算简单、提取特征速度快、所需存储空间小、对光照变化和单调灰度变化具有很好的鲁棒性、能有效地描述图像的局部纹理信息等优点,被广泛使用.常用的分类识别方法有支持向量机(Support Vector Machine, SVM)^[5,6]、神经网络^[7]以及基于稀疏表示的分类(Sparse Representation based Classification, SRC)^[8]等.

本文基于中心化二值模式算子(CBP特征)和稀疏表示的三维人脸表情识别算法框架,提出对原始图像采用先分块后分别提取特征的方法,以克服全局CBP特征对表情分类表达不够的缺点,最后利用稀疏表示实现三维人脸表情识别.为了验证本文方法的识别性能,实验分析了分块提取CBP特征和稀疏表示的分类效果,分别比较了本文特征与LBP算子、CBP算子,以及稀疏表示与SVM的分类性能.

1 方法框架

本文方法分为面部多尺度分块,局部CBP特征提取和稀疏表示分类三部分,如图1所示.

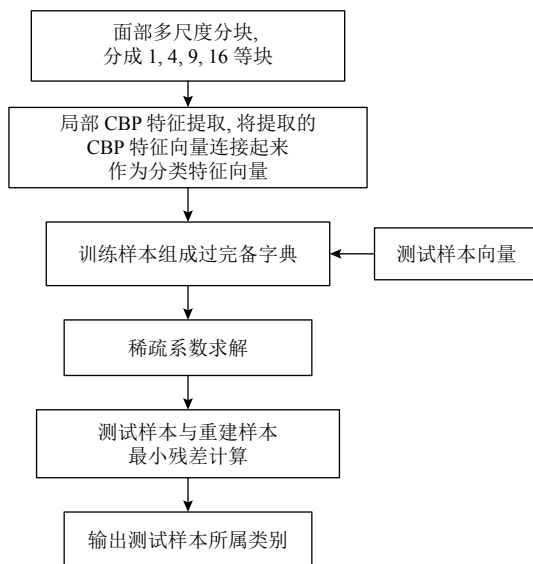


图1 本文方法的算法流程图

1) 面部多尺度分块.为了更好地表达出原图像人脸表情的细节特征,提高分类准确率,本文首先对原始图像进行了不同尺度分块.分块数量分别为1, 4, 6, 9, 16块,如图2所示.

2) 局部CBP特征提取.由于CBP特征对于图像

的纹理特征具有较强的描述能力,而这些特征是表情分类的重要依据,本文基于前步分块结果,提取各分块的局部CBP特征,并计算CBP特征直方图作为分类特征向量.

3) 稀疏表示分类.同表情的人脸样本可看作一类,其特征向量处于同一个线性空间,任何一个测试样本都可以用来自于该类的训练样本进行线性表示;将所有的训练样本特征向量构成字典,则测试样本特征向量在该字典上的表示是稀疏的,同时该稀疏系数包含了样本的类别信息.基于此,本文利用稀疏表示进行人脸识别.

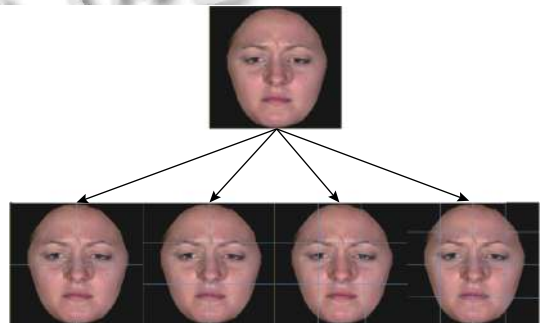


图2 人脸区域划分图

1.1 分块CBP特征提取

在人脸表情的识别中,多数的人脸表情特征都集中在上半脸眼睛、眉毛周围和下半脸的嘴巴附近等局部地区,这些局部特征信息对表情识别的正确率具有重要作用,但是常用的整体图片的CBP直方图序列分类特征向量,不能很好的对这些局部信息进行有效的表示.

基于CBP算子表情识别的基本原理是将中心点像素加入到传统的LBP算子当中,并且赋予它最高权值,通过比较中心像素点与周围环状邻域的像素灰度差值得出图像的CBP编码,即CBP特征图.然后把图像的CBP编码直方图组成的直方图序列作为特征向量用于表情识别.本文采用邻域为8,半径为1的CBP算子,公式表示为:

$$CBP(8, 1) = \sum_{l=0}^3 s(g_n - g_{n+4})2^n + s\left(g_c - \frac{1}{9}\left(\sum_{n=0}^7 g_n + g_c\right)\right)2^4$$

$$= s(g_0 - g_4)2^0 + s(g_1 - g_5)2^1 + s(g_2 - g_6)2^2$$

$$+ s(g_3 - g_7)2^3 + s(g_c - g_T)2^4 \quad (1)$$

式(1)中, $g_T = (g_0 + g_1 + g_2 + g_3 + g_4 + g_5 + g_6 + g_7 + g_c)/9$

式(1)中 g_C 为中心像素点像素值, $g_i(i=1, 2, 3, \dots, 7)$ 为邻域像素点像素值, g_T 为平均像素值.

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

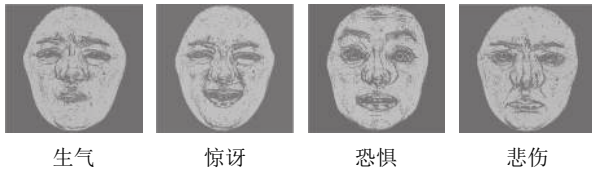


图3 不同表情的CBP特征图

1.2 稀疏表达分类器

近些年来, 稀疏表示(SRC)^[8]已经成为信号处理与相关应用领域的重要概念之一, 其基本思想是用含有 n 个不同类别的数据集构造过完备字典, 通过求解测试数据在过完备字典下的稀疏表示系数来判断其类别.

稀疏表示的原理为: 把同一类别的样本归于一个特征空间, 当样本属于某一类时, 可由该类特征空间的其它样本进行线性描述, 线性表示能够找到稀疏解. 因此, 测试样本的特征向量可以被训练样本的同一类型特征稀疏表示, 此时构造的过完备字典完全由训练样本本身构成.

所有的训练样本可以构成一个稀疏表示的过完备字典 $D = [D_1, D_2, \dots, D_n]$ 那么字典中的每列对应一类人脸表情. 若样本 y 属于第 k 类测试样本, 那么可以由过完备字典表示为:

$$y = DA \quad (3)$$

式(3)中, A 表示 y 在 D 上的投影系数矩阵. 具体公式如下:

$$A = [A_1^1, \dots, A_m^1, A_1^2, \dots, A_m^2, A_1^n, \dots, A_m^n] \quad (4)$$

式(4)中, 为测试样本 y 在 D 中的第 i 类第 j 个训练样本上的投影系数. 根据稀疏表示的分类原理可知, 当样本 y 在字典 D 中属于第 k 类训练样本的投影系数不全为0时, 则样本与 y 属于第 k 类人脸表情, 且在其它类别的投影系数全部为0. 因此, 找到式(3)的最稀疏解成为判断样本 y 所属类别的关键. 通过式(5)求解可以得到特征向量近似解.

$$\hat{A} = \arg \min \|A\|_1 \quad \text{s.t.} \quad \|DA - y\| < \varepsilon \quad (5)$$

在稀疏表示分类的过程中, 采取的计算稀疏投影矩阵方法为正交匹配追踪算法(Orthogonal Matching

Pursuit, OMP)^[10]. 与匹配追踪算法^[11](Matching Pursuit, MP)相比. OMP算法在分解的每一步对所选择的全部原子进行正交化处理, 这使得在精度要求相同的情况下, OMP算法的收敛速度更快. 最终求得稀疏投影系数 A , 并使用同类别的训练样本重建测试样本, 计算测试样本与重建样本之间的残差值, 然后通过判断最小残差值实现人脸表情识别分类.

2 相关实验

本文实验采取的硬件平台为英特尔 i7-7400 处理器, 软件包括 Opencv3.0, Matlab2014b, Visual studio 2015 等开发工具.

本文实验采取 BU-3DFE^[12]数据库进行了大量的实验. BU-3DFE 数据集共有来自不同地区, 不同年龄的 100 个人的共计 2500 幅不同表情图像. 其中女性 56 人, 男性 44 人. 每个人的人脸表情信息包括中性(NE)、生气(AN)、厌恶(DI)、恐惧(FE)、高兴(HA)、悲伤(SA)和惊讶(SU), 其中除中性表情外, 其余六种表情均含有四种不同的强度等级. 图4展示了同一人的不同强度的4种表情图.

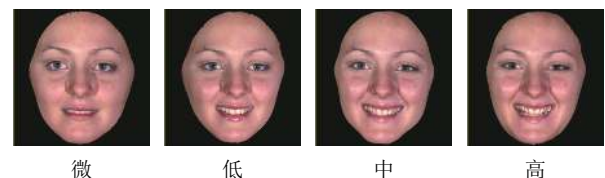


图4 不同强度等级人脸表情图

为了有效地进行实验, 对从 BU-3DFE 数据库中选取实验材料提出如下规则: 不只采用最高强度的三维人脸表情模型进行实验, 而是四个强度均被采用, 用以验证本算法对于细微表情的变化是否拥有良好的鲁棒性; 对于 100 个随机选取的三维人脸表情模型, 50 个用作稀疏表示训练, 50 个用作测试样本, 且单一实验重复 20 次, 然后计算平均识别准确度作为实验结果.

本文主要进行两个实验, 用于验证基于稀疏表示的三维人脸表情识别方法的有效性.

实验 1. 首先, 提取人脸表情深度图像的 CBP 特征和 LBP 特征, 然后根据提取的特征进行训练, 分别使用稀疏表示与 SVM 分类器进行训练, 最后, 利用得到的模型进行测试, 得到三维人脸表情识别结果. 具体结果如表 1 所示.

表1 三维人脸识别结果对比

| 相关工作 | 通用特征分类器 | 平均识别准确度 (%) |
|-----------------------------|----------|-------------|
| 本算法 | CBP+稀疏表示 | 87.21 |
| | CBP+SVM | 85.78 |
| | LBP+SVM | 70.27 |
| | LBP+稀疏表示 | 75.90 |
| Wang 等人 ^[13] | 距离特征+LDA | 83.6 |
| Berretti 等人 ^[14] | SIFT+SVM | 78.4 |

根据表1实验数据,我们可以得出如下结论:

(1) 基于稀疏表示分类器的三维人脸表情识别准确率高于基于SVM分类器的识别效果.从表中可以看出,基于稀疏表示的三维人脸表情识别针对CBP特征与LBP特征的识别准确率分别为87.21%与75.90%,均高于基于SVM的三维人脸表情识别针对CBP特征与LBP特征的85.78%与70.27%.这说明稀疏表示分

类方法对于三维人脸表情识别来说具有更好的分类识别能力.提出的基于CBP特征与稀疏表示方法的三维人脸表情识别效果最高,识别准确率达到87.21%;

(2) 基于CBP特征的三维人脸表情识别效果远优于基于LBP特征的识别效果.表2中,基于CBP特征的三维人脸表情识别准确率分别为85.78%和87.21%,比基于LBP特征的识别准确率分别提高了15.51%和11.31%;

(3) 本文所采用的基于CBP特征与稀疏表示分类的算法框架,同Wang等人^[13]和Berretti等人^[14]的算法相比具有明显的优势,分类识别的正确率分别高出3.6%和8.81.实验1的特征提取对象是整个人脸表情图像,人脸表情不同区域对于人脸表情的描述能力存在差异,因此为了进一步验证本文提出的分块CBP特征的性能,本文基于实验1中性能最好的CBP+稀疏表示算法框架,进行了实验2的对比分析.

表2 三维人脸识别结果 (%)

| 分块 | 高兴 | 生气 | 恐惧 | 惊讶 | 悲伤 | 厌恶 | 平均值 |
|------|------|------|------|------|------|------|------|
| 1分块 | 92.1 | 80.2 | 87.5 | 93.8 | 82.7 | 87.0 | 87.2 |
| 4分块 | 93.3 | 82.4 | 89.2 | 94.1 | 84.1 | 87.9 | 88.5 |
| 6分块 | 95.7 | 83.0 | 88.4 | 94.9 | 84.6 | 89.2 | 89.3 |
| 9分块 | 98.8 | 85.7 | 90.6 | 95.9 | 87.4 | 91.3 | 91.6 |
| 16分块 | 96.4 | 83.5 | 89.2 | 95.1 | 85.2 | 89.4 | 89.8 |

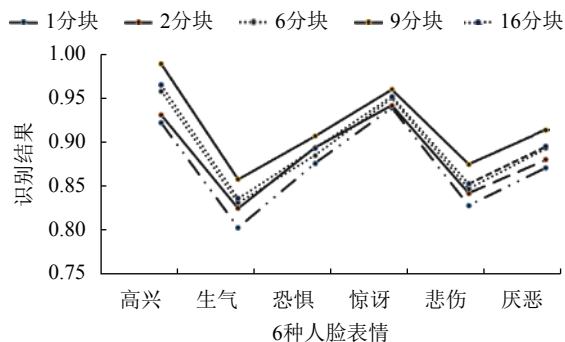


图5 不同分块数识别结果曲线

从实验结果,可以得出以下结论:

(1) 随着人脸表情区域分块增加,基于CBP特征与稀疏表示的三维人脸表情识别平均准确率呈增长趋势.从表中可以看出,基于1分块、4分块、6分块、9分块与16分块的三维人脸表情识别的平均准确率分别为87.2%、88.5%、89.3%、91.6%和89.8%,三维人脸表情平均识别准确率先增加后减少,其中9分块分类表现最佳.

(2) 随着区域分块数目的增加,三维人脸不同表情

的识别准确率大多呈上升趋势.从图中可以看出,对于中性、高兴、愤怒、惊讶、悲伤与厌恶这6种人脸表情,随着人脸分块数的增加,识别准确率随着增加,仅有恐惧表情例外,这说明了人脸不同区域对于不同的表情影响程度不同,合理地进行人脸区域分块能够有效地提升人脸表情识别准确率:

3 总结与展望

本文针对三维人脸表情识别,提出了基于分块CBP特征与稀疏表示的三维人脸表情识别方法,阐述了基于该算法的三维人脸表情识别具体过程.最后实验部分,针对CBP特征与LBP特征、稀疏表示与SVM分别进行了比较,验证了基于CBP特征与稀疏表示进行三维人脸表情识别算法框架的优越性.在此基础上,对人脸表情图像进行均匀分块,并对不同分块图像进行CBP特征提取,以稀疏表示为分类器,实现三维人脸表情识别,最后得出了合理的区域分块能有效地提升三维人脸表情识别准确率的结论.

在本文算法基础上, 可以将七种表情分类拓展到更多表情, 以满足对现实生活中的人类表情的分类需求. 此外, 可以在本文工作基础上进一步进行不均匀分块, 或者特定分块, 自适应权值等思路. 可以更加充分的利用人脸面部不同区域对于表情识别的不同影响, 从而有效提高识别正确率.

参考文献

- 1 Chen JJ, Zhao Z, Sun H, *et al.* Facial expression recognition based on PCA reconstruction. 2010 5th International Conference on Computer Science & Education. Hefei, China. 2010. 195–198.
- 2 Edwards GJ, Taylor CJ, Cootes TF. Interpreting face images using active appearance model. Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. Nara, Japan. 1998. 300–305.
- 3 Zhang HH, Zhang BL, Huang WM, *et al.* Gabor wavelet associative memory for face recognition. IEEE Transactions on Neural Networks, 2005, 16(1): 275–278. [doi: [10.1109/TNN.2004.841811](https://doi.org/10.1109/TNN.2004.841811)]
- 4 Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions. Pattern Recognition, 1996, 29(1): 51–59. [doi: [10.1016/0031-3203\(95\)00067-4](https://doi.org/10.1016/0031-3203(95)00067-4)]
- 5 Heisele B, Serre T, Poggio T. A component-based framework for face detection and identification. International Journal of Computer Vision, 2007, 74(2): 167–181. [doi: [10.1007/s11263-006-0006-z](https://doi.org/10.1007/s11263-006-0006-z)]
- 6 Zaidi NA, Squire DM. Local adaptive SVM for object recognition. 2010 International Conference on Digital Image Computing: Techniques and Applications. Sydney, NSW, Australia. 2010. 196–201.
- 7 Rázuri JG, Sundgren D, Rahmani R, *et al.* Automatic emotion recognition through facial expression analysis in merged images based on an artificial neural network. 2013 12th Mexican International Conference on Artificial Intelligence. Mexico City, Mexico. 2013. 85–96.
- 8 Wright J, Yang AY, Ganesh A, *et al.* Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210–227. [doi: [10.1109/TPAMI.2008.79](https://doi.org/10.1109/TPAMI.2008.79)]
- 9 Mallat SG, Zhang ZF. Matching pursuits with time-frequency dictionaries. IEEE Transactions on Signal Processing, 1993, 41(12): 3397–3415. [doi: [10.1109/78.258082](https://doi.org/10.1109/78.258082)]
- 10 Pati YC, Rezaiifar R, Krishnaprasad PS. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. Proceedings of 27th Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA. 1993. 40–44.
- 11 Cai TT, Wang L. Orthogonal matching pursuit for sparse signal recovery with noise. IEEE Transactions on Information Theory, 2011, 57(7): 4680–4688. [doi: [10.1109/TIT.2011.2146090](https://doi.org/10.1109/TIT.2011.2146090)]
- 12 Yin LJ, Wei XZ, Sun Y, *et al.* A 3D facial expression database for facial behavior research. 7th International Conference on Automatic Face and Gesture Recognition. Southampton, UK. 2006. 211–216.
- 13 Wang J, Yin LJ, Wei XZ, *et al.* 3D facial Expression recognition based on primitive surface feature distribution. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, NY, USA. 2006. 1399–1406.
- 14 Berretti S, Ben Amor B, Daoudi M, *et al.* 3D facial expression recognition using SIFT descriptors of automatically detected keypoints. The Visual Computer, 2011, 27: 1021–1036. [doi: [10.1007/s00371-011-0611-x](https://doi.org/10.1007/s00371-011-0611-x)]