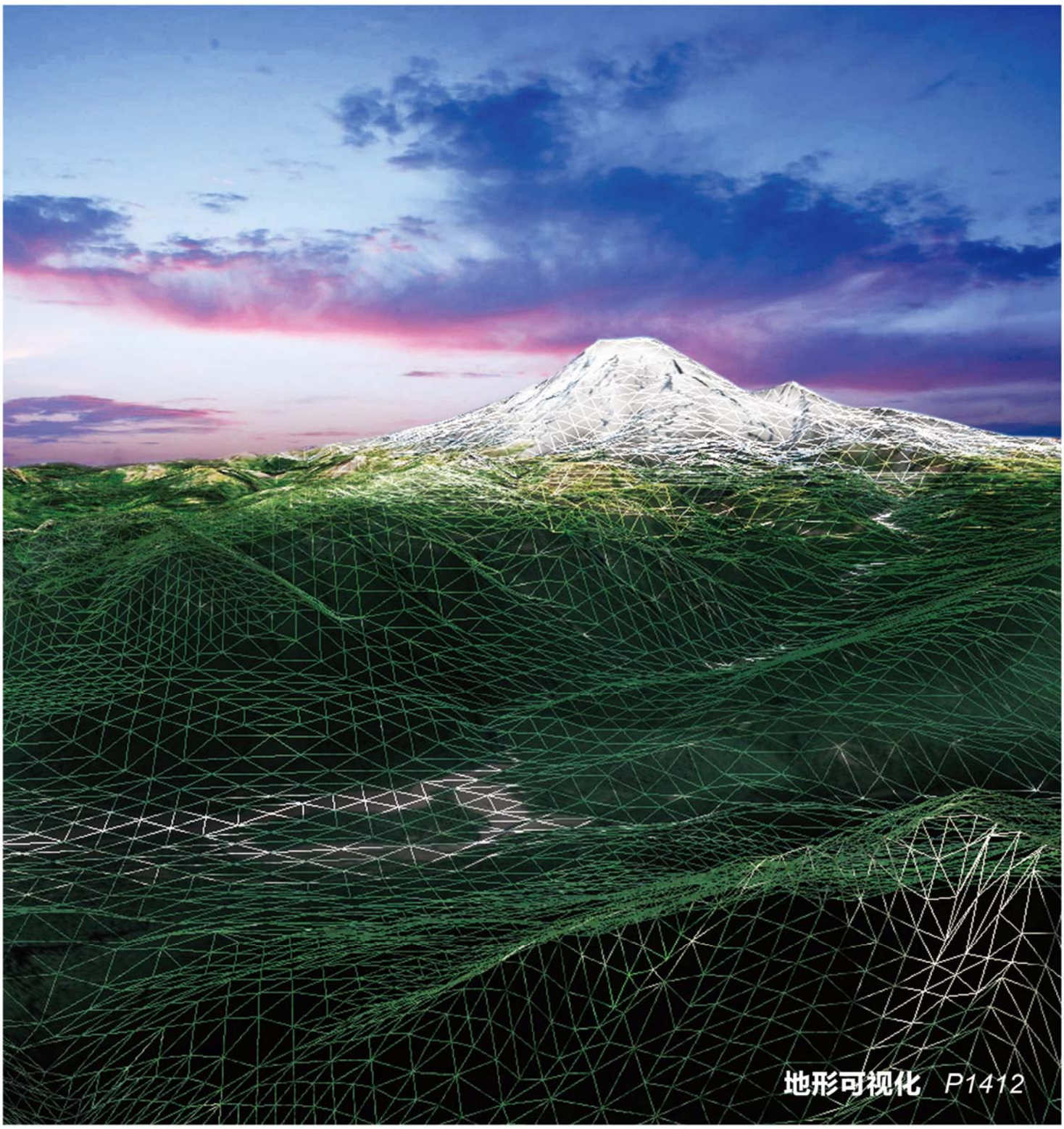
JIG
JOURNAL OF IMAGE AND GRAPHICS

中国图象图形学报

主办：中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

2015
10
VOL.20

ISSN1006-8961
CN11-3758/TB



地形可视化 P1412

中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）

第20卷第10期（总第234期）

2015年10月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主　　编 李小文
编辑出版 《中国图象图形学报》编辑出版委员会
邮政信箱 北京9718信箱
邮　　编 100101
电子信箱 jig@radi.ac.cn
电　　话 010-64807995
网　　址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号
总 发 行 北京报刊发行局
订　　购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
（邮政信箱：北京399信箱　邮编：100048）
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
P.O.Box 9718, Beijing, P.R.China
Zip code 100101
E-mail jig@radi.ac.cn
Telephone 010-64807995
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048,P.R.China))
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 50.00元

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

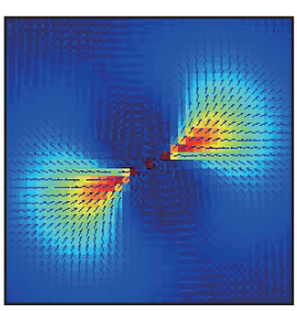
All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

2015年10月 第20卷 第10期（总第233期）
中国图象图形学报 Zhongguo Tuxiang Tuxing Xuebao

目次

JPEG图像双重压缩偏移量估计的篡改区域自动检测定位(第1304页)

基于表情传输的交互式照片编辑(第1366页)



结合张量投票和Snakes模型的SAR图像道路提取(第1403页)

图像处理和编码

基于模式特征的H.264/AVC可逆视频水印
李淑芝，张翔，邓小鸿，吴晓燕 1285

二阶总广义变分小波修复模型
许建楼，郝岩，张冀 1297

JPEG图像双重压缩偏移量估计的篡改区域自动检测定位
赵洁，郭继昌，张艳，张众维 1304

图像分析和识别

基于非对称局部梯度编码的人脸表情识别
胡敏，程轶红，王晓华，任福继，许良凤，黄晓音 1313

采用积分图块间距离检测图像边缘
贾迪，孟琭，孙劲光，李思慧，赵明远 1322

自适应属性加权2维FCM分割算法
侯晓凡，吴成茂 1331

利用层间相关性的岩心CT图像半自动分割方法
徐永进，滕奇志，吴晓红，卿粼波 1340

图像理解和计算机视觉

在线特征选取的多示例学习目标跟踪
周志宇，彭小龙，吴迪冲，朱泽飞 1346

融合外观和运动特征的在线目标分割
张雷，李成龙，汤进，高思晗 1358

基于RGB-D深度相机的室内场景重建
梅峰，刘京，李淳芃，王兆其 1366

计算机图形学

带平台伸缩函数的参数曲线变形
张莉，余慧芳，檀结庆 1374

视景建模中树木纹理图像的随机变形网格方法
施冠羽，欧阳清 1384

基于表情传输的交互式照片编辑
刘锦云，彭宏京 1390

遥感图像处理

结合张量投票和Snakes模型的SAR图像道路提取
符喜优，张风丽，王国军，邵芸 1403

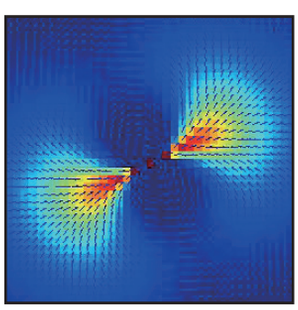
GPU Tessellation全球地形可视化方法
李尚林，郑利平，张迎凯，李琳 1412

Volume 20, Number 10
Published October 16, 2015

CONTENTS
JOURNAL OF IMAGE AND GRAPHICS

Automatic detection and localization of image forgery regions based on offset estimation of double JPEG compression (P1304)

Interactive photo editing based on expression transmission (P1366)



Road extraction from SAR images using tensor voting and Snakes model(P1403)

Image Processing and Coding

Reversible video watermarking algorithm for H.264/AVC based on mode feature
Li Shuzhi, Zhang Xiang, Deng Xiaohong, Wu Xiaoyan 1285

Second order total generalized variational model for wavelet inpainting
Xu Jianlou, Hao Yan, Zhang Ji 1297

Automatic detection and localization of image forgery regions based on offset estimation of double JPEG compression
Zhao Jie, Guo Jichang, Zhang Yan, Zhang Zhongwei 1304

Image Analysis and Recognition

Facial expression recognition based on asymmetric region local gradient coding
Hu Min, Cheng Yihong, Wang Xiaohua, Ren Fuji, Xu Liangfeng, Huang Xiaoyin 1313

Edge detection method of block distance combining with summed area table
Jia Di, Meng Lu, Sun Jinguang, Li Sihui, Zhao Mingyuan 1322

Adaptive weighted two-dimensional histogram FCM segmentation algorithm
Hou Xiaofan, Wu Chengmao 1331

Semi-segmentation of rock CT images using the correlation of adjacent frames
Xu Yongjin, Teng Qizhi, Wu Xiaohong, Qing Linbo 1340

Image Understanding and Computer Vision

Object tracking based on multiple instance learning with online feature selection
Zhou Zhiyu, Peng Xiaolong, Wu Dichong, Zhu Zefei 1346

Online object segmentation via fusing appearance and motion features
Zhang Lei, Li Chenglong, Tang Jin, Gao Sihan 1358

Improved RGB-D camera based indoor scene reconstruction
Mei Feng, Liu Jing, Li Chunpeng, Wang Zhaoqi 1366

Computer Graphics

Deformation of parametric curves based on platform extension function
Zhang Li, Yu Huifang, Tan Jieqing 1374

Method of random transmogrification mesh for tree texture images in scene modeling
Shi Guanyu,Ouyang Qing 1384

Interactive photo editing based on expression transmission
Liu Jinyun, Peng Hongjin 1390

Remote Sensing Image Processing

Road extraction from SAR images using tensor voting and Snakes model
Fu Xiyou, Zhang Fengli, Wang Guojun, Shao Yun 1403

GPU Tessellation method of global terrain visualization
Li Shanglin, Zheng Liping, Zhang Yingkai, Li Lin 1412

© 中国图象图形学报版权所有

中图法分类号：TP391　　**文献标识码**：A　　**文章编号**：1006-8961(2015)10-1313-09

论文引用格式：Hu M, Cheng Y H, Wang X H, Ren F J, Xu L F, Huang X Y. Facial expression recognition based on asymmetric region local gradient coding[J]. Journal of Image and Graphics, 2015, 20(10): 1313-1321. [胡敏, 程轶红, 王晓华, 任福继, 许良凤, 黄晓音. 基于非对称局部梯度编码的人脸表情识别[J]. 中国图象图形学报, 2015, 20(10): 1313-1321.] [DOI: 10.11834/jig.20151004]

基于非对称局部梯度编码的人脸表情识别

胡敏[1,2]，程轶红[1,2]，王晓华[1,2]，任福继[1,2]，许良凤[1]，黄晓音[1,2]

1. 合肥工业大学计算机与信息学院，合肥　230009；2. 情感计算与先进智能机器安徽省重点实验室，合肥　230009

摘　要：目的　针对局部梯度编码算子(LGC)只能在固定大小邻域内提取图像纹理特征的不足，提出了一种非对称邻域LGC算子(AR-LGC)多尺度融合的表情特征提取方法。**方法**　首先，对归一化的表情图像进行Gauss滤波处理；然后，对图像进行分块，对每个子块图像中每一像素点，采用不同邻域大小的AR-LGC算子得到两个二进制序列，将两个序列作按位逻辑异或得到一个新的序列，对此序列进行编码，计算每个子块的直方图分布，级联各子块直方图构成人脸表情的特征；最后用SVM分类器进行表情分类识别。**结果**　该算法在JAFFE库和CK库上进行实验，分别取得了95.24%和96.83%的平均识别率，并与CBP(中心化二值模式)、LBP(局部二值模式)、LGC和AR-LBP(非对称局部二值模式)算法进行了比较，在JAFFE库的平均识别率分别比CBP、LBP、LGC、AR-LBP高5.6%、4.85%、3.71%、2.40%，在CK库的平均识别率分别比CBP、LBP、LGC、AR-LBP高3.66%、2.50%、2.17%、1.66%，实验结果表明，该算法可以较准确地进行人脸表情识别。**结论**　本文所提的表情特征提取方法通过融合不同梯度不同尺度子邻域间的强度关系，可以很好地表达图像的局部特征和全局特征，与典型的特征提取算法的对比实验也表明了本文算法的有效性，表明本文算法适用于静态人脸表情图像的识别。

关键词：表情识别；非对称局部梯度编码；特征提取；多尺度融合；支持向量机(SVM)

Facial expression recognition based on asymmetric region local gradient coding

Hu Min[1,2], Cheng Yihong[1,2], Wang Xiaohua[1,2], Ren Fuji[1,2], Xu Liangfeng[1], Huang Xiaoyin[1,2]

1. *School of Computer and Information of Hefei University of Technology, Hefei*　230009, *China*;
2. *Anhui Province Key Laboratory of Affective Computing and Advanced Intelligent Machine, Hefei*　230009, *China*

Abstract: Objective To overcome the deficiency of local gradient coding, which only extracts texture feature in neighborhoods of a fixed size, we propose a novel multi-scale local gradient coding fusion method based on asymmetric regions for feature extraction of facial expressions. **Method** A normalized face image is preprocessed by using a Gaussian filter to reduce the impact of noise. Then, the preprocessed expression image is divided into several blocks. For each pixel of each sub-block image, multiple and differently sized operators of local gradient coding based on asymmetric regions are used to obtain two binary sequences. These binary sequences are fused into a new binary sequence according to the logical XOR.

收稿日期：2015-04-29；**修回日期**：2015-06-12

基金项目：国家自然科学基金项目(61432004)；安徽省自然科学基金项目(1408085MKL16)；国家自然科学青年基金项目(61300119)

第一作者简介：胡敏(1967—)，女，教授，硕士生导师，2004年于合肥工业大学获计算机应用技术专业博士学位，主要研究方向为图形图像处理、计算机视觉等。E-mail: uhnim@163.com

Supported by: National Natural Science Foundation of China (61432004); Young Scientists Fund of National Natural Science Foundation of China (61300119)

© 中国图象图形学报版权所有

The new binary sequence is then encoded, each sub-block histogram distribution is statistically analyzed, and all the sub-block histograms are cascaded into the texture features of a facial expression. Finally, the process of expression classification is completed with the SVM method. **Result** Experiments using the proposed method are performed using the JAFFE database and CK database. The average recognition rate for JAFFE is 95.24%, whereas that for CK is 96.83%. The proposed method is compared with LBP, CBP, LGC, and AR-LBP. Experimental results demonstrate that the proposed approach for the JAFFE database achieves recognition rates that are 5.6%, 4.85%, 3.71%, and 2.40% higher than those achieved with CBP, LBP, LGC, and AR-LBP, respectively. As for the CK database, the proposed approach achieves recognition rates that are 3.66%, 2.50%, 2.17%, and 1.66% higher than those achieved with CBP, LBP, LGC, and AR-LBP, respectively. Cross validation results show that the proposed method for facial expression recognition has excellent accuracy. **Conclusion** Through the fusion of the intensities between neighborhoods of different gradients and scales, the multi-scale local gradient coding fusion method based on asymmetric regions combined with block histograms can perform well in local and global feature description. Experiment results show that the proposed method is better than typical feature extraction algorithms and is suitable for static facial expression recognition.

Key words: expression recognition; asymmetric region local gradient coding; feature extraction; multi-scale fusion; SVM

0 引 言

人脸表情包含着丰富的信息,在人际交往中,它是一种非常重要的信息传递方式。心理学家 Mehrabian 的研究结果表明,在人们交流中,7% 的信息通过语言来传递的,38% 的信息通过声调来传递的,而 55% 的信息是通过面部表情来传递的[1]。因此一直以来,人脸表情识别得到了广泛的研究,并成为一个研究热点[2-6]。

人脸表情识别主要分为 3 个步骤:图像预处理与人脸检测、特征提取、分类识别。其中,特征提取是人脸表情识别系统的重要部分,也是提高识别精度的重要一步。目前,用于特征提取的几大流行方法有:Gabor 小波[7]、活动外观模型(AAM)[8]、主成分分析(PCA)、局部二值模式(LBP)等。Gabor 小波从多尺度多方向来提取图像的特征,特征维数较高,计算比较复杂。文献[9]对预处理后的表情图像先提取 Gabor 特征,然后采用 PCA(主成分分析) + LDA(线性判别分析)算法进行特征选择,最后采用 K 近邻方法分类,该算法虽降低特征维数,但计算比较复杂,实时性较低。与此相比,文献[10]提出的局部二值模式纹理描述法简单有效,具有良好的灰度不变性和旋转不变性,是一个很好的纹理描述子,已经广泛应用于目标检测、人脸识别、纹理分类、图像匹配、人脸表情识别等领域[11]。但在实际应用中,传统 LBP 算子难以适应不同尺度下提取纹理特征的需要,许多学者在具体应用中对 LBP 进行了不少相应的改进。Wang 等人[12]提出 LBP 金字塔特征,通过多尺度分析构建人脸图像金字塔,采用 LBP 算子提取各层图像的 LBP 特征建立图像的 LBP 金字塔,分块统计各层 LBP 特征谱并级联构成图像的特征。Singh 等人[13]提出一种完全局部二值模式,融合图像差值符号特征、差值幅度特征和中心像素点灰度特征,提取了比 LBP 更充分的纹理特征的同时增加了时间复杂度。Han 等人[14]针对传统 LBP 算子缺少对图元的相位分析,提出一种融合图元旋转不变性和相位统计信息的纹理分析算法,解决了欠分类问题,但不能解决由噪声等原因引起的错误分类。文献[15]提出了一种局部梯度编码(LGC)算法,通过从水平、垂直、对角线 3 个梯度方向上比较邻域间强度大小进行编码,提取图像纹理特征,用支持向量机(SVM)来进行分类识别,该算法考虑了眉毛、眼睛、嘴巴等区域特征点的变化趋势,鉴别能力比传统 LBP 算子强,但文献中该算子的邻域大小固定,不能在大的尺度下很好地提取图像纹理特征,同时易受噪声影响。

在文献[16-17]的启发下,针对文献[15]中算法的缺点,提出一种非对称邻域局部梯度编码描述算子(AR-LGC)。该算子具有可扩展性,具有恒定的特征维数,能反映局部区域的纹理细微变化。实验选用日本女性人脸表情库(JAFFE)和 Cohn-Kanade(CK)库。首先用 Gauss 滤波对图像进行预处理,然后将图像分为不同子块,在各子块内利用 AR-LGC 算子提取特征并用直方图统计信息,最后采用 SVM 分类器对表情特征进行分类。

© 中国图象图形学报版权所有

1 图像预处理

图像常常被强度随机信号所污染，为了降低噪声的影响，提高识别精度，先对图像进行 Gauss 滤波处理。Gauss 滤波器是一类根据 Gauss 函数的形状来选择权值的线性平滑滤波器。图像上的噪声一般服从 Gauss 正态分布，Gauss 滤波可以消除噪声干扰。2 维 Gauss 函数表达式为

$$G(x,y) = \exp\left\{-\frac{1}{2}\left(\left(\frac{x}{\sigma_x}\right)^2+\left(\frac{y}{\sigma_y}\right)^2\right)\right\} \quad (1)$$

Gauss 滤波器宽度(决定着平滑程度)是由方差参数 σ_x，σ_y 表征的，这些参数表征了对图像进行平滑的程度，其值越大，Gauss 滤波器的频带就越宽，平滑程度就越好。令 $I(x,y)$ 表示输入图像的灰度分布，那么图像的 Gauss 滤波可以定义为

$$F(x,y) = I(x,y) \otimes G(x,y) \quad (2)$$

式中，“⊗”表示卷积运算，$F(x,y)$ 表示滤波后输出图像。

2 改进的局部梯度编码算法

2.1 局部梯度编码算法

传统 LBP 仅考虑中心像素点和周围邻域的灰度关系，未考虑周围邻域间的灰度关系，图 1(a)(b)为 3×3 大小的图像灰度值示例，显然，两幅图像中周围 8 个邻域的像素点灰度值存在差别，但若用 LBP 算子进行编码则结果相同。因此文献[15]提出了 LGC 算子。LGC 算子的邻域大小为 3×3，如图 2 所示，该算子从水平、垂直、对角线 3 个梯度上比较邻域间的灰度大小，该算法为

$$\begin{aligned} LGC(x,y) = {} & s(g_0-g_2)\times 2^7 + s(g_7-g_3)\times 2^6 + \\ & s(g_6-g_4)\times 2^5 + s(g_0-g_6)\times 2^4 + \\ & s(g_1-g_5)\times 2^3 + s(g_2-g_4)\times 2^2 + \\ & s(g_0-g_4)\times 2^1 + s(g_2-g_6)\times 2^0 \end{aligned} \quad (3)$$

式中 $s(x)$ 函数为

$$s(g_f-g_c) = \begin{cases} 1 & g_f > g_c \\ 0 & g_f \leqslant g_c \end{cases} \quad (4)$$

图 1(a)(b)周围 8 个邻域的像素点灰度值均小于中心像素点的灰度值，LBP 编码为$(00000000)_2$，但 8 个邻域之间的灰度关系存在差别，图 1(a)(b)的

1	5	5
1	6	5
1	1	5

(a) LGC: $(00001001)_2$

1	1	1
1	6	1
1	1	1

(b) LGC: $(00000000)_2$

图 1 不同灰度分布示意图(LBP:$(00000000)_2$)

Fig. 1 Examples of different gray distribution (LBP:$(00000000)_2$)

g_0	g_1	g_2
g_7	g_c	g_3
g_6	g_5	g_4

图 2 LGC 的 3×3 大小邻域的灰度示意图

Fig. 2 The 3×3 template of LGC operator

LGC 编码为$(00001001)_2$、$(00000000)_2$，这样就将图 1(a)和图 1(b)区分开。与传统 LBP 算子相比，LGC 算子具有更强的鉴别能力。

2.2 非对称 LGC 编码算法

LGC 算子考虑了中心像素周围邻域之间的灰度关系，但是算子邻域大小是固定的，不可扩展，不能在更大的尺度下很好地提取图像的纹理特征，因此在文献[16]提出的 AR-LBP 算法的启发下，结合文献[15]，提出了一种非对称邻域的局部梯度编码描述子，即 AR-LGC 描述子。该描述子的邻域大小为$(2m+1)\times(2n+1)$，其中 $1\leqslant m\leqslant\left\lfloor\frac{h-1}{2}\right\rfloor$，$1\leqslant n\leqslant\left\lfloor\frac{w-1}{2}\right\rfloor$，$h$ 和 w 分别为图像的高度和宽度，符号$\lfloor\ \rfloor$表示向下取整。将该邻域分成 9 个子邻域，如图 3 所示。每个子邻域记为 R_i，每个子邻域的像素

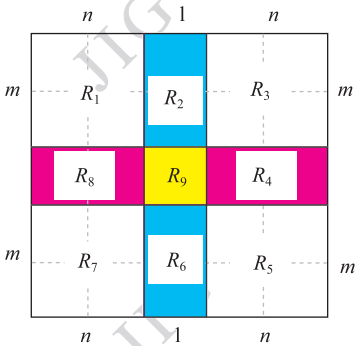


图 3 AR-LGC 的 9 个子邻域示意图

Fig. 3 Nine sub region of AR-LGC operator

© 中国图象图形学报版权所有

值记为 $g_{R_i}, i=1,2,\cdots,9$。

图 3 中，黄色子邻域 R_9 大小为 1×1，为中心像素点，未着色的 4 个子邻域 R_1、R_3、R_5、R_7 的大小为 $m\times n$，蓝色的 2 个子邻域 R_2、R_6 的大小为 $m\times1$，紫色的 2 个子邻域 R_4、R_8 的大小为 $1\times n$。

$$g_{R_i}=\frac{1}{N_i}\sum_{j=1}^{N_i}p_{ij} \tag{5}$$

式中，N_i 表示第 i 个子邻域内像素点的个数。p_{ij} 表示第 i 个子邻域第 j 个像素点的像素值。

AR-LGC 描述子从水平、垂直、对角线 3 个梯度上来比较子邻域之间的强度关系，依次考察了局部邻域内 3 对水平梯度、3 对垂直梯度及 2 对对角线梯度子邻域像素值差值的大小关系，顺序为

g_{R_1}	g_{R_2}	g_{R_3}
g_{R_8}	g_{R_9}	g_{R_4}
g_{R_7}	g_{R_6}	g_{R_5}

根据比较的结果生成一串 8 位二进制码，再将该串二进制码转换成十进制后即为此局部邻域的纹理描述值。

二进制序列：

$P_0:s(g_{R_1}-g_{R_3})$，$P_1:s(g_{R_8}-g_{R_4})$，

$P_2:s(g_{R_7}-g_{R_5})$，$P_3:s(g_{R_1}-g_{R_7})$，

$P_4:s(g_{R_2}-g_{R_6})$，$P_5:s(g_{R_3}-g_{R_5})$，

$P_6:s(g_{R_1}-g_{R_5})$，$P_7:s(g_{R_3}-g_{R_7})$

AR-LGC 的编码具体为

$$AR-LGC(x,y)=\sum_{i=0}^{7}P_i\times2^i \tag{6}$$

式中，P_i 表示二进制序列第 i 位的值。

AR-LGC 描述子具有扩展性，可以在较大的尺度下提取图像的丰富纹理特征并且特征维数是恒定的。

2.3 异或 AR-LGC 编码算法

AR-LGC 描述子邻域大小是可扩展的，特征维数是固定的，能在大尺度下描述图像的主要纹理特征，但同时存在一些不足：1）邻域大小的可扩展性可能导致该描述子忽略局部邻域内强度不断变化的趋势；2）每个子邻域的值取的是子邻域内所有像素点的均值，纹理平滑化会降低该描述子的鉴别能力。

在数字电路中，当二进制 a 和 b 相同时输出为 0，不同时输出为 1，这就是异或运算。在本文的特征提取中当考虑两个尺度 AR-LGC 算子时，若在梯度上邻域的强度关系相同，则结果为 1；若在梯度上邻域的强度关系不同，则结果为 0。这正好与异或运算的逻辑规则相符。同与运算、或运算相比，异或运算更符合本文需求，并且从其结果中可以看出不同尺度上梯度方向强度关系是否是一致的。

因此对 AR-LGC 做了相应的改进，提出了异或 AR-LGC（XOR-AR-LGC）编码算法。对同一像素点，将通过两个不同大小的邻域窗口得到的两个序列 PA 和 PB 进行按位异或得到一个新的序列 P，对新的序列 P 进行编码，该值作为该像素点的纹理特征值。设序列 PA 和 PB 的第 i 位为 PA_i 和 PB_i，XOR-AR-LGC 编码具体为

$$XOR\text{-}AR\text{-}LGC(x,y)=\sum_{i=0}^{L-1}(PA_i\oplus PB_i)\times2^i \tag{7}$$

式中，“⊕”表示逻辑异或操作，L 为序列的长度，$L=8$。

XOR -AR-LGC 的编码过程的示意图如图 4 所示。从图 4 可以看到，以两个邻域大小 5×5 和 3×3 为例，同一像素点不同尺度上邻域间存在强度变

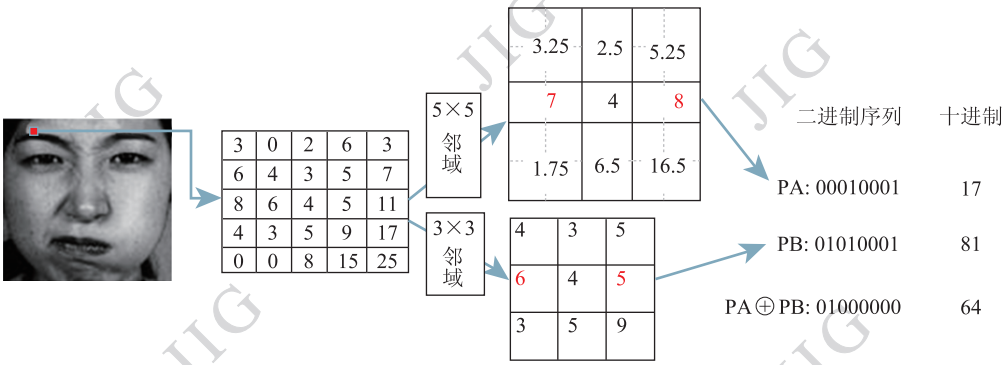


图 4 异或 AR-LGC 编码示意图

Fig. 4 Schematic diagram of the XOR AR-LGC coding

© 中国图象图形学报版权所有

化，如中心像素点 R_9（$g_{R_9}=4$）的 5×5 邻域红色部分子邻域 R_8 像素值小于 R_4 像素值（$g_{R_8}=7<g_{R_4}=8$），其 AR-LGC 算子的编码为 PA：$(00010001)_2$，而 3×3 邻域红色部分子邻域 R_8 像素值大于 R_4 像素值（$g_{R_8}=6>g_{R_4}=5$），其 AR-LGC 算子的编码为 PB：$(01010001)_2$，PA 与 PB 按位异或得到异或 AR-LGC 的编码 PA ⊕ PB：$(01000000)_2$。

异或 AR-LGC 通过融合水平、垂直、对角线 3 个梯度方向上不同尺度的子邻域间强度变化关系来进行编码，它继承 AR-LGC 的优点，能更好地提取图像的纹理特征，同时具有更高的鉴别能力。

2.4 异或 AR-LGC 特征提取

直方图能够有效地描述纹理图像的全局特征，然而直接对整幅表情图像进行直方图统计会丢失很多结构上的细节信息，因此先将表情图像划分为若干个大小相等且互不重叠的子块，然后分别对每个子块区域内的特征统计其直方图，将每个子块的直方图级联起来，级联后的直方图表示该表情图像的特征。

第 t 个子块的分布直方图定义为

$$H_t(i)=\sum_{y=0}^{h-1}\sum_{x=0}^{w-1}I(XOR-AR-LGC(x,y)=i),\quad i=0,1,\cdots,255 \tag{8}$$

式中，h 和 w 分别表示子块的高度和宽度，

$$I(A)=\begin{cases}1 & A\text{ 为真}\\ 0 & A\text{ 为假}\end{cases} \tag{9}$$

整幅表情图像的直方图定义为

$$\boldsymbol{H}=\{H_1,H_2,\cdots,H_t,\cdots,H_N\} \tag{10}$$

式中，H_t 表示第 t 个子块的直方图，N 为子块的个数，$t\in\{1,2,\cdots,N\}$。

本文特征提取过程具体步骤为：

1）将经过预处理后的表情图像分成 $n\times n$ 个子块；

2）对每个子块采用异或 AR-LGC 算法进行编码，得到异或 AR-LGC 的纹理图像；

3）统计每个子块的异或 AR-LGC 直方图；

4）级联 n^2 个子块的直方图，级联后的直方图表示该幅表情图像的特征，具体如图 5 所示。

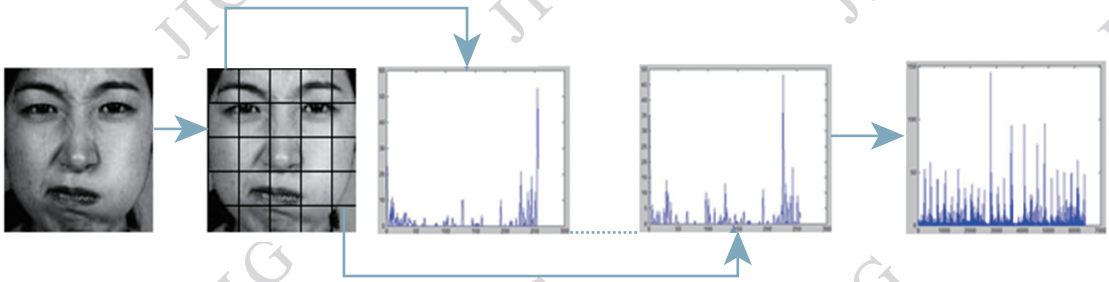


图 5 异或 AR-LGC 特征提取过程

Fig. 5 The feature extraction process using XOR AR-LGC operator

AR-LGC 的邻域大小是可变的，该描述子是可扩展的，它可以通过实验选择一个合适大小的邻域组合来提取图像的纹理特征。异或 AR-LGC 融合不同梯度的不同尺度上邻域间的强度变化，继承 AR-LGC 的优点，比 AR-LGC 的鉴别能力强，能提取出更具鉴别能力的图像纹理特征。分块直方图包含了更多邻域内的信息，兼顾局部的细微变化和整体变化，有助于提高识别精度。

3 实验描述及结果分析

为了测试异或 AR-LGC 特征的性能，在日本女性人脸表情库（JAFFE）和 Cohn-Kanade（CK）库上进行人脸表情识别实验。

3.1 实验描述

JAFFE 库包括 10 个人的表情图像，每人有 7 种表情：生气、厌恶、恐惧、高兴、中性、悲伤、惊讶。每种表情有 3 或 4 幅样本，共 213 幅表情图像。

CK 库包含了 200 个由 18～30 岁的 210 个成年人近 2 000 幅不同表情灰度图像序列，每一种表情均由一系列的动态图像构成。其中 69% 是女性，31% 是男性，81% 是欧美人，13% 是黑人，6% 的其他人种。每个人有 6 种表情：生气、厌恶、恐惧、高兴、悲伤、惊讶。选取 CK 库中部分人的所有表情进行实验。

实验步骤如下：

1）对人脸表情图像进行预处理，再将预处理后的图像进行尺寸归一化，统一大小为 96×96 像素，

© 中国图象图形学报版权所有

如图 6 所示。

(a) JAFFE 库人脸图像预处理示例

(b) CK 库人脸图像预处理示例

图 6　预处理示例

Fig. 6　Preprocessing image samples

2)用本文算法提取图像的异或 AR-LGC 特征作为人脸表情特征。

3)采用 SVM 分类器来进行人脸表情识别。

3.2　实验结果分析

在 JAFFE 库中随机选取每个人每种表情的 2 幅图像作为训练样本,剩余的每个人每种表情的 1 幅图像作为测试样本,因此,每次实验选取 140 幅图像作训练样本和 70 幅图像作测试样本; 在 CK 库中选取其中 18 人的每种表情 1 ～3 幅图像作为训练样本,在剩余图像中选取每种表情 1 ～3 幅图像作为测试样本,因此实验中 CK 库训练样本库和测试样本库中 6 种表情(悲伤,高兴,惊讶,恐惧,生气,厌恶)样本数分别为(37,38,38,30,30,28)和(36,38,38,30,30,28);每个库上的实验循环 3 次,得到平均识别率。

3.2.1　对 AR-LGC 邻域大小选择的研究

由于 AR-LGC 子邻域取的是均值,邻域的大小不能过大或过小,过大的邻域不能表示该像素点的周围强度变化,过小的邻域对噪声点很敏感,易受其影响,所以在 JAFFE 库和 CK 库上对选取不同邻域大小的 AR-LGC 做了研究,实验结果如表 1 和表 2 所示。表 1、表 2 中 H 和 W 分别表示 AR-LGC 描述子邻域的高度和宽度。

从表 1、表 2 中可以看出,随邻域的宽度和高度的增大,识别率提高,但达到一定大小时再增大,识别率又下降,两个表中都是邻域大小为 5×5 时识别率达到最大值 93.35% 和 95.67%。实验中异或 AR-LGC 选用的两个邻域大小为 3×3 和 5×5。

表 1　JAFFE 库不同邻域大小的 AR-LGC 的识别率比较

Table 1　The recognition rate of different size of AR-LGC on JAFFE database　/%

邻域高度 H	邻域宽度 W			
	3	5	7	9
3	91.53	91.60	90.58	88.21
5	91.96	93.35	92.46	87.76
7	92.47	89.64	89.18	84.01
9	90.11	91.06	88.22	83.04

表 2　CK 库不同邻域大小的 AR-LGC 的识别率比较

Table 2　The recognition rate of different sizes of AR-LGC on CK database　/%

邻域高度 H	邻域宽度 W			
	3	5	7	9
3	94.66	94.83	93.5	90
5	95	95.67	94.5	86.5
7	94	93.5	89.66	85.5
9	90	90.5	86.5	83.5

3.2.2　本文算法的实验结果及分析

利用 SVM 分类器进行分类时,常用的核函数有高斯核函数和多项式核函数。高斯核函数具有很强的局部性,学习能力强,但泛化能力较弱;多项式函数具有良好的全局特性以及很强的泛化能力。故本实验选用多项式核函数作为 SVM 的核函数,并采用“one-vs-one”的方式实现多分类。多项式核函数下的 SVM 分类器的决策函数为

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{l} a_i y_i (x_i \cdot x + 1)^d + b\right) \quad (11)$$

式中,l 为训练样本数,a_i、x_i 和 y_i 分别为第 i 个训练样本的拉格朗日系数、特征向量和类别,b 为最优超平面的偏移量。d 表示多项式次数,次数越低,泛化能力越强但学习能力越弱。选取 $d=4$,因为在该次数下即可以使算法达到良好的识别率又可以使算法有很好的泛化能力。

实验其他相关参数设置为:图像分块大小为 5×5,异或 AR-LGC 的两个邻域大小为 3×3 和 5×5。

在 JAFFE 库上的实验结果如表 3 所示。从表 3 中可以看出,本文算法在 JAFFE 库上的平均识别率为 95.24%,表情出现误识别。其中,识别错误的表情大多数都存在相似性,如图 7(a)恐惧误识别为生气,图 7(b)生气误识别为悲伤,因

© 中国图象图形学报版权所有

为这些表情在眉毛、眼睛、鼻子上的细节变化比较相似。

在 CK 库上的实验结果如表 4 所示。从表 4 中可以看出，本文算法在 CK 库上的平均识别率为 96.83%，该算法对生气、厌恶、悲伤 3 种表情出现误识别情况，原因是这些表情之间有很多的雷同处，属于微表情误识别。

(a) 误识别生气　(b) 误识别为悲伤

图 7　误识别表情示例

Fig. 7　Mistaken image samples

表 3　基于 JAFFE 人脸表情库的实验结果

Table 3　Three groups of experiment results on JAFFE database

表情	测试样本数	正确识别数			平均识别率/%
		第 1 次	第 2 次	第 3 次	
生气	10	10	10	9	96.67
厌恶	10	8	10	10	93.33
恐惧	10	10	9	8	90
高兴	10	10	10	9	96.67
中性	10	10	10	9	96.67
悲伤	10	9	10	10	96.67
惊讶	10	10	10	9	96.67
总计	70	67	69	64	95.24

表 4　基于 CK 人脸表情库的实验结果

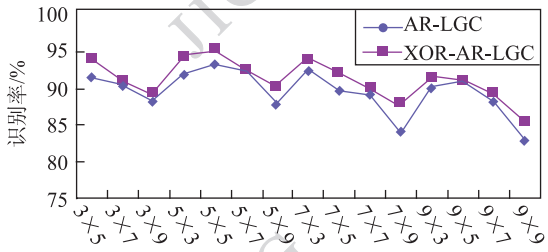
Table 4　Three groups of experiment results on CK database

表情	测试样本数	正确识别数			平均识别率/%
		第 1 次	第 2 次	第 3 次	
生气	30	28	28	28	93.33
厌恶	28	26	25	26	91.66
恐惧	30	30	29	29	97.78
高兴	38	38	38	38	100
悲伤	36	34	35	34	95.37
惊讶	38	38	38	38	100
总计	200	194	193	194	96.83

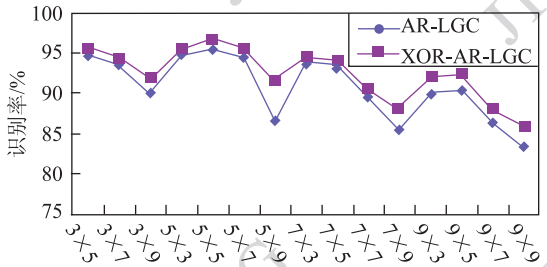
3.2.3　不同算法的识别率分析

为表明本文算法的有效性，还进行了以下实验：1）在不同大小邻域下比较异或 AR-LGC 与 AR-LGC，实验结果如图 8 所示；2）将本文方法与 CBP（中心化二值模式）、LBP、LGC、AR-LBP（非对称局部二值模式）作比较，比较结果如表 5、表 6 所示。这些算法都是在同一实验环境下实现的。

从图 8（a）（b）的折线图中可以看出，在不同大小邻域下，异或 AR-LGC 的识别率大部分均高于 AR-LGC 的识别率，表明了异或 AR-LGC 能较好地提高识别率。



(a) JAFFE 库上的识别率比较



(b) CK库上的识别率比较

图 8　异或 AR-LGC 与 AR-LGC 在不同大小邻域下的识别率比较

Fig. 8　Rate comparison of XOR-AR-LGC and AR-LGC in different sizes of region

表 5　各算法识别性能的比较

Table 5　Average recognition rate comparison of different algorithms

算法	识别率/%	
	JAFFE 库	CK 库
CBP + SVM	89.64	93.17
LBP + SVM	90.39	94.33
LGC + SVM	91.53	94.66
AR-LBP + SVM	92.84	95.17
本文	95.24	96.83

© 中国图象图形学报版权所有

表 6　各算法其他性能的比较结果

Table 6　Comparison in other features of different algorithms

算法	特征维数	特征提取时间/(ms/幅)	识别时间/(ms/幅)
CBP + SVM	800	10	10
LBP + SVM	6 400	12	13.55
LGC + SVM	6 400	12	13.55
AR-LBP + SVM	6 400	12.35	13.8
本文	6 400	12.5	14

从表 5、表 6 中可以看出,本文算法在 JAFFE 库上的识别率达到 95.24%,在 CK 库上识别率达到 96.83%,识别率较高,同时一定程度上满足了实时性的要求。但是与 CBP 相比,该算法特征维数大,特征提取时间和识别时间较长,在识别率上有所提高。在适当的范围内增加空间和时间的开销来提高识别率是可行的,因此,实验表明了该算法的有效性。

4　结　论

针对传统 LBP 算法和已有的 LGC 算法的不足,提出了 AR-LGC 算子,并对该算子做了进一步的改进,得到异或非对称邻域 LGC。该算法从中心像素的边缘响应强度出发,考虑中心像素的子邻域间的强度关系,并且通过融合不同尺度不同梯度上邻域间的强度进行编码,提取图像特征。与表情识别主流算法相比,本文算法无论在时间复杂度上还是在识别效果上都表现出了一定的优越性:

1)AR-LGC 算子比较了中心像素点的周围邻域间的强度关系,考虑了眼睛、眉毛、嘴巴等在梯度方向上的渐变趋势,该算子具有更好的鉴别能力;

2)本文算法的邻域大小是可扩展的,它可以在更大的尺度下较好的提取图像的纹理特征,并且特征维数是恒定的,不随邻域的增大而增大;

3)本文算法融合了不同尺度不同方向的邻域间的关系来进行编码的,形成融合特征,融合后的特征比单尺度下的特征更具鉴别能力,提高识别率。实验结果表明该算法的识别能力强,而且过程简单,易于实现。

本文算法是基于正面人脸和静态图像的表情识别,主要考虑图像的空间域强度信息,计算量不大,具有较高的识别率,适用于静态人脸表情识别。然而,表情是一个动态过程,如何有效地利用动态信息提高姿态变化大、视频序列中表情的识别率是下一步所要研究的工作。

参考文献(References)

[1] Mehrabian A. Communication without words[J]. Psychological today, 1968, 2:53-55.

[2] Xue Y L, Mao X, Guo Y. The research advance of facial expression recognition in human computer interaction[J]. Journal of Images and Graphics, 2009, 14(5): 764-772. [薛雨丽,毛峡,郭叶. 人机交互中的人脸表情识别研究进展[J]. 中国图象图形学报, 2009, 14(5): 764-772.] [DOI: 10.11834/jig.20090503]

[3] Li Y Q, Li Y J, Li H B. Fusion of global and local various feature for facial expression recognition[J]. Acta Optica Sinica, 2014, 34(5): 0515001(1-7). [李雅倩, 李颖杰, 李海滨. 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(1-7).]

[4] Eleftheriadis S, Rudovic O, Pantic M. Discriminative shared gaussian processes for multiview and view-invariant facial expression recognition[J]. IEEE Transactions on Image Processing, 2015, 24(1): 189-204.

[5] Yu J, Wang Z F, Li R. A simultaneous motion tracking and facial expression recognition algorithm[J]. Acta Electronic Sinica, 2015, 43(2): 371-376. [於俊, 汪增福, 李睿. 一种同步人脸运动跟踪与表情识别算法[J]. 电子学报, 2015, 43(2): 371-376.]

[6] Calvo M G, Fernández-Martín A, Nummenmaa L. Facial expression recognition in peripheral versus central vision: role of the eyes and the mouth[J]. Psychological research, 2014, 78(2): 180-195.

[7] Yao W, Sun Z X, Zhang Y. Optimal Gabor feature for facial expression recognition[J]. Journal of Computer-aided Design & Computer Graphics, 2008, 20(1): 79-84. [姚伟, 孙正兴, 张岩. 面向脸部表情识别的 Gabor 特征选择方法[J]. 计算机辅助设计与图形学学报, 2008, 20(1): 79-84.]

[8] Wang L T, Ding X Q, Fang C. A novel method for robust and automatic facial features localization[J]. Acta Automatic Sinica, 2009, 35(1): 9-16. [王丽婷, 丁晓青, 方驰. 一种鲁棒的全自动人脸特征点定位方法[J]. 自动化学报, 2009, 35(1): 9-16.]

[9] Deng H B, Jin L W. Facial expression recognition based on local Gabor filter bank and PCA + LDA[J]. Journal of Images and Graphics, 2007, 12(2): 322-329. [邓洪波, 金连文. 一种基于局部 Gabor 滤波器组及 PCA + LDA 的人脸表情识别方法[J]. 中国图象图形学报, 2007, 12(2): 322-329.] [DOI:

© 中国图象图形学报版权所有

10.11834/jig.20070224]

[10] Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions [J]. Pattern Recognition, 1996, 29(1): 51-59.

[11] Zhang J Y, Zhao H P, Chen S. Face recognition based on weighted local binary pattern with adaptive threshold[J]. Journal of Electronics & Information Technology, 2014,36(6):1327-1333. [张洁玉, 赵鸿萍, 陈曙. 自适应阈值及加权局部二值模式的人脸识别[J]. 电子与信息学报, 2014, 36(6): 1327-1333.]

[12] Wang W, Huang F F, Li J W, et al. Face description and recognition by LBP pyramid[J]. Journal of Computer-Aided Design & Computer Graphics, 2009, 21(1): 94-100.

[13] Singh S, Maurya R, Mittal A. Application of complete local binary pattern method for facial expression recognition[C]//Proceedings of the 4th International Conference on Intelligent Human Computer Interaction. Nanchang, China: IEEE, 2012: 1-4.

[14] Han Y B, Yi J Q, Li J P. Research of an improved LBP algorithm in texture classification based on rotation invariance and statistical phase distribution[J]. Chinese Journal of Computers, 2011, 34(3):583-592. [韩延彬, 尹建芹, 李金屏. 基于图元旋转不变性和相位统计信息的 LBP 算法在纹理分类中的研究[J]. 计算机学报, 2011, 34(3):583-592.]

[15] Tong Y, Chen R, Cheng Y. Facial expression recognition algorithm using LGC based on horizontal and diagonal prior principle [J]. Optik-International Journal for Light and Electron Optics, 2014, 125(16): 4186-4189.

[16] Naika C L S, Das P K, Nair S B. Asymmetric region local binary pattern operator for person-dependent facial expression recognition [C]//Proceedings of International Conference on Computing, Communication and Applications. Dindigul, India: IEEE, 2012: 1-5.

[17] Ahsan T, Jabid T, Chong U P. Facial expression recognition using local transitional pattern on Gabor filtered facial images[J]. IETE Technical Review, 2013, 30(1): 47-52.

© 中国图象图形学报版权所有