

JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象学报 图形学报

2021
12
VOL.26

ISSN1006-8961
CN11-3758/TB





第26卷第12期（总第308期）
2021年12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

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.

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
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 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@aircas.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218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 Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

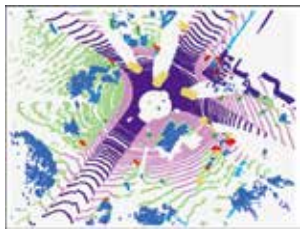
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

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 60.00元



环视鱼眼图像处理深度学习
研究进展(第2778页)



结合空间深度卷积和残差
的大尺度点云场景分割
(第2848页)



曼哈顿世界环境下消失点非线
性优化估计算法(第2931页)

综述

深度对抗视觉生成综述

谭明奎, 许守恺, 张书海, 陈奇 2751

人机交互中的智能体情感计算研究

刘婷婷, 刘箴, 柴艳杰, 王瑾, 王媛怡 2767

环视鱼眼图像处理深度学习研究进展

包俊, 刘宏哲, 褚文博 2778

图像处理和编码

图像分解与色彩先验下的多曝光图像融合

李嫛源, 王琴, 朱智勤, 齐观秋 2800

双尺度分解和显著性分析相结合的红外与可见光图像融合

霍星, 邹韵, 陈影, 檀结庆 2813

轻量级注意力特征选择循环网络的超分辨率重建

徐雯捷, 宋慧慧, 袁晓彤, 刘青山 2826

区域级通道注意力融合高频损失的图像超分辨率重建

周波, 李成华, 陈伟 2836

图像分析和识别

结合空间深度卷积和残差的大尺度点云场景分割

刘盛, 黄圣跃, 程豪豪, 沈家瑜, 陈胜勇 2848

细粒度图像分类的互补注意力方法

赵勋, 王家宝, 李阳, 王亚鹏, 苗壮 2860

联合分块谱带优选和深度特征的高光谱人脸识别

谢志华, 李毅, 牛杰一 2870

时空双仿射微分不变量及骨架动作识别

李琪, 墨瀚林, 赵婧涵, 郝宏翔, 李华 2879

多尺度分割先验迭代重加权低秩恢复显著性检测

张荣国, 郑维佳, 赵建, 胡静, 刘小君 2892

融合环境特征与改进YOLOv4的安全帽佩戴检测

葛青青, 张智杰, 袁琰, 李秀梅, 孙军梅 2904

图像理解和计算机视觉

关节轴角先验对3维人体重建结果的影响

姚砾, 张幼安, 张梦雪, 万燕 2918

曼哈顿世界环境下消失点非线性优化估计算法

韩冰心, 卢惠民, 于清华, 张礼廉 2931

双目机器视觉及RetinaNet模型的路侧行人感知定位

连丽容, 罗文婷, 秦勇, 李林 2941

遥感图像处理

遥感图像中油罐目标精确定位与参数提取

江哈, 张月婷, 郭嘉逸, 赵鑫, 李芳芳, 黄丽佳, 胡玉新, 雷斌, 丁赤飏 2953

参数合成空间变换网络的遥感图像一致性配准

陈颖, 张祺, 李文举, 石艳娇, 陈磊 2964

改进Mask R-CNN模型的海洋锋检测

徐慧芳, 黄冬梅, 贺琪, 杜艳玲, 覃学标 2981

噪声鲁棒的轻量级深度遥感场景图像分类检索

王亚鹏, 李阳, 王家宝, 赵勋, 苗壮 2991

改进U-Net型网络的遥感图像道路提取

杨佳林, 郭学俊, 陈泽华 3005

弱监督深度语义分割网络的多源遥感影像水体检测

李鑫伟, 李彦胜, 张永军 3015

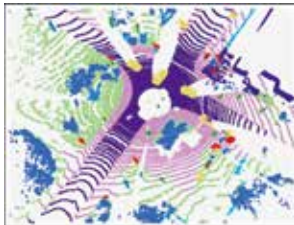
总目次 I

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Research progress of fisheye image processing based on deep learning(P2778)



A deep residual network with spatial depthwise convolution for large-scale point cloud semantic segmentation(P2848)



Vanishing point estimation based on non-linear optimization in manhattan world environments (P2931)

Review

A review on deep adversarial visual generation	
Tan Mingkui, Xu Shoukai, Zhang Shuhai, Chen Qi	2751
Agent affective computing in human-computer interaction	
Liu Tingting, Liu Zhen, Chai Yanjie, Wang Jin, Wang Yuanyi	2767
Research progress of fisheye image processing based on deep learning	
Bao Jun, Liu Hongzhe, Chu Wenbo	2778

Image Processing and Coding

Multi-exposure image fusion based on image decomposition and color prior	
Li Yuanyuan, Wang Qin, Zhu Zhiqin, Qi Guanqiu	2800
Dual-scale decomposition and saliency analysis based infrared and visible image fusion	
Huo Xing, Zou Yun, Chen Ying, Tan Jieqing	2813
Lightweight attention feature selection recursive network for super-resolution	
Xu Wenjie, Song Huihui, Yuan Xiaotong, Liu Qingshan	2826
Region-level channel attention for single image super-resolution combining high frequency loss	
Zhou Bo, Li Chenghua, Chen Wei	2836

Image Analysis and Recognition

A deep residual network with spatial depthwise convolution for large-scale point cloud semantic segmentation	
Liu Sheng, Huang Shengyue, Cheng Haohao, Shen Jiayu, Chen Shengyong	2848
Complemented attention method for fine-grained image classification	
Zhao Xun, Wang Jiabao, Li Yang, Wang Yapeng, Miao Zhuang	2860
Hyperspectral face recognition based on partition bands optimal selection and deep features	
Xie Zhihua, Li Yi, Niu Jieyi	2870
Spatio-temporal dual affine differential invariants for skeleton-based action recognition	
Li Qi, Mo Hanlin, Zhao Jinghan, Hao Hongxiang, Li Hua	2879
Iterative reweighted low-rank recovery salient object detection with multiscale segmentation prior	
Zhang Rongguo, Zheng Weijia, Zhao Jian, Hu Jing, Liu Xiaojun	2892
Safety helmet wearing detection method of fusing environmental features and improved YOLOv4	
Ge Qingqing, Zhang Zhijie, Yuan Long, Li Xiumei, Sun Junmei	2904

Image Understanding and Computer Vision

The impact of joint axis angle prior on the results of 3D human body reconstruction	
Yao Li, Zhang Youan, Zhang Mengxue, Wan Yan	2918
Vanishing point estimation based on non-linear optimization in manhattan world environments	
Han Bingxin, Lu Huimin, Yu Qinghua, Zhang Lilian	2931
Roadside pedestrian detection and location based on binocular machine vision and RetinaNet	
Lian Lirong, Luo Wenting, Qin Yong, Li Lin	2941

Remote Sensing Image Processing

Accurate localization and parameter extraction of oil tank in remote sensing images	
Jiang Han, Zhang Yueting, Guo Jiayi, Zhao Xin, Li Fangfang, Huang Lijia, Hu Yuxin, Lei Bin, Ding Chibiao	2953
Consistent registration of remote sensing images in parametric synthesized spatial transformation network	
Chen Ying, Zhang Qi, Li Wenju, Shi Yanjiao, Chen Lei	2964
Ocean front detection method based on improved Mask R-CNN	
Xu Huifang, Huang Dongmei, He Qi, Du Yanling, Qin Xuebiao	2981
A robust lightweight deep neural network for remote sensing scene image classification and retrieval under label noise	
Wang Yapeng, Li Yang, Wang Jiabao, Zhao Xun, Miao Zhuang	2991
Road extraction method from remote sensing images based on improved U-Net network	
Yang Jialin, Guo Xuejun, Chen Zehua	3005
Weakly supervised deep semantic segmentation network for water body extraction based on multi-source remote sensing imagery	
Li Xinwei, Li Yansheng, Zhang Yongjun	3015

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2021)12-2870-09

论文引用格式: Xie Z H, Li Y and Niu J Y. 2021. Hyperspectral face recognition based on partition bands optimal selection and deep features. Journal of Image and Graphics, 26(12): 2870-2878 [谢志华, 李毅, 牛杰一. 2021. 联合分块谱带优选和深度特征的高光谱人脸识别. 中国图象图形学报, 26(12): 2870-2878] [DOI:10.11834/jig.200158]

联合分块谱带优选和深度特征的高光谱人脸识别

谢志华*, 李毅, 牛杰一

江西科技师范大学光电子与通信重点实验室, 南昌 330031

摘要: **目的** 高光谱人脸数据具有丰富的鉴别信息。最优谱带选择和谱内特征表示是高光谱人脸识别的关键。基于高光谱波段范围为 400~1 090 nm 和采样间隔为 10 nm 的高光谱成像人脸数据, 本文提出一种分块谱带选择和 VGG (Visual Geometry Group) 网络的高光谱人脸识别方法。**方法** 为了优化适合人脸识别的谱带组合, 基于人脸关键点, 提出分块局部二值模式 (local binary pattern, LBP) 特征的 AdaBoost 支持向量机 (support vector machine, SVM) 谱带选择方法。基于卷积神经网络结构建立一个面向高光谱人脸特点的深度网络 (VGG12), 提取谱带内特征。融合不同谱带的深度特征, 利用三层堆栈自编码器 (stack auto-encoder, SAE) 抽取谱间特征。对提取的谱间和谱内特征, 采用最近邻分类器完成最后的识别。**结果** 为了验证提出方法的有效性, 在公开的高光谱人脸数据集 UWA-HSFD (University of Western Australia hyperspectral face database) 和 PolyU-HSFD (Hong Kong Polytechnic University hyperspectral face database) 上进行对比试验。结果显示, 基于分块 LBP 特征的谱带选择算法优于传统基于整幅图像像素的方法, 提出的 VGG12 网络相比已有深度学习网络, 仅保留少量 (6~8 个) 谱带, 在两个数据集上都取得了最高的识别率 (96.8% 和 97.2%), 表明传统可见光人脸深度网络并不适合高光谱人脸识别。**结论** 实验结果表明, 高光谱数据用于人脸识别中, 谱带选择与深度学习结合是有效的, 本文方法联合有监督深度网络 (VGG12) 和无监督学习网络 (SAE) 挖掘谱内和谱间鉴别特征, 在降低深度网络训练复杂度的同时取得了较其他深度网络更好的识别性能。

关键词: 高光谱人脸识别; 局部二值模式 (LBP); VGG 网络; 谱带选择; AdaBoost SVM

Hyperspectral face recognition based on partition bands optimal selection and deep features

Xie Zhihua*, Li Yi, Niu Jieyi

Key Laboratory of Optic-Electronic and Communication, Jiangxi Sciences and Technology Normal University, Nanchang 330031, China

Abstract: **Objective** The applications of hyperspectral face imaging have been getting more and more concerned due to the low cost of hyperspectral camera. Hyperspectral imaging has been providing sufficient distinctive information along the spectral dimension for unconstrained face recognition system. The hyperspectral face images have been taking challenges such as low signal to noise ratio, high computational complexity and cross-band misalignment. High data dimensionality can result in high computational costs. Not all spectral bands are equally informative and discriminative. The useless spectral band usage may even produce noises and weaken the performance. Spatial feature extraction and optimal bands selection

收稿日期: 2020-05-13; 修回日期: 2020-08-26; 预印本日期: 2020-09-02

* 通信作者: 谢志华 xie_zhihua68@aliyun.com

基金项目: 国家自然科学基金项目 (61861020); 江西省教育厅科技项目 (GJJ190578)

Supported by: National Natural Science Foundation of China (61861020)

have been two vital issues for robust hyperspectral face recognition. The feature representation based on deep learning has obtained excellent performance in image classification tasks. The deep neural networks to extract spatial features for hyperspectral face data have been focused on. A novel hyperspectral face recognition method based on automatic bands selection and visual geometry group network (VGG net) with the success of VGG net in traditional face recognition. **Method** First, an improved AdaBoost algorithm has been built up to select the optimal bands for different facial regions. One face has been divided into 3×5 partition mode based on the robust 68 landmarks from the cascade regression method to get the adaptive spectral features for different facial organs. Next, support vector machine (SVM) has been opted as a weak classifier for AdaBoost ensemble learning for bands selection. SVM is more suitable for small sample and high dimension sets, and lower sensitivity to imbalanced datasets, which are the characteristics of hyperspectral face data. Thus, the AdaBoost SVM algorithm has been used to select optimal bands (ranging from 450 nm to 1 090 nm) for hyperspectral face recognition. The AdaBoost SVM is based on partition local binary pattern (LBP) features. Second, a convolutional neural network (CNN) structure has been built for deep feature extraction in each hyperspectral face image based on VGG12 deep network. Therefore, VGG12 network can fulfill the deep feature extraction in spatial domain. Third, a three-layer stacked auto-encoder (SAE) has been trained to mine the discriminative and compact features amongst different bands. Finally, the nearest neighbor classifier has been applied to get the final results. The automatic bands selection is based on partition LBP features instead of original images. The optimal bands selection can choose the most suitable bands for different regions in faces. The second novelty has combined the supervised deep network (VGG12) with unsupervised learning network (SAE) to fully mine within and between spectral features in hyperspectral faces. **Result** To verify the effectiveness of the proposed hyperspectral face recognition method, the comparative experiments have been conducted on public UWA hyperspectral face database (UWA-HSFD) and Hong Kong Polytechnic University hyperspectral face database (PolyU-HSFD) databases where the spectral bands vary from 400 to 1 090 nm with the spectral interval of 10 nm. The two datasets have been preprocessed and normalized to 220×180 hyperspectral face images. The ratio of the test set of the training set is 1 : 1. The first half samples have been used as the training set and the rest have been used as the test set. The VGG12 network has been trained based on the multi-spectral bands cycle. The objective function of the network is Adam. The super parameters of VGG12 are set as follows: $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, epoch = 10. The learning rate of SAE network is 0.000 1 and epoch is 10. The all experiments have been implemented on the computer platform whose configuration is of 2.6 GHz CPU, 8 GB GPU. In contrast experiment 1, the bands selection algorithm based on partition LBP features has been superior to the traditional method based on the whole faces. The main reason is that partition LBP features can improve the influence of non-alignment factors of hyperspectral facial data and select optimal bands combination for the different face regions. In contrast experiments 2, the method with SAE model can get a higher recognition rate than that without SAE model by introducing little computational complexity (about 15% consuming time). In contrast experiment 3, compared with the existing deep learning network, the proposed VGG12 network, retaining only 6 ~ 8 spectral bands, achieves the highest recognition rates (96.8% and 97.2%) on both databases. **Conclusion** The experiment results have shown that the combination of bands selection and VGG12 is feasible and effective for the performance improvement of hyperspectral face recognition.

Key words: hyperspectral face recognition; local binary pattern (LBP); VGG net; bands selection; AdaBoost SVM

0 引言

随着传感器和计算能力的飞速发展,高光谱成像已经成为可靠的、准确的新型成像技术,其设备价格也有了大幅度下降,目前广泛应用于农业、生物医学和计算机视觉等领域。其中,高光谱人脸成像由于具有丰富的鉴别信息,成为提高人脸识别性能的

有效途径之一(Osia 和 Bourlai,2014)。已有研究显示,人脸的不同部位包含明显的光谱差异性,而高光谱成像能捕获这些变异性特征,使高光谱图像数据含有较多的鉴别信息,提高了非约束人脸识别的鲁棒性(Pan 等,2003)。选择高光谱人脸数据的最优谱带和自动挖掘高光谱人脸数据鉴别特征是实现鲁棒高光谱人脸识别的两个关键步骤(Chang 等,2010)。

根据高光谱成像包含丰富鉴别信息的特点, Di 等人(2010)建立了公开的高光谱人脸数据库,将高光谱人脸投影到 2 维图像空间,通过经典的欧氏距离进行识别。为提高高光谱人脸识别中特征提取的有效性,Shen 和 Zheng(2012)提出了基于 Gabor 小波分解的高光谱人脸识别方法。Uzair 等人(2015)提出了基于空间谱带信息融合的高光谱人脸识别方法,并比较了常用可见光识别算法在高光谱人脸识别系统上的效果。为降低高光谱人脸的谱间冗余度,魏冬梅等人(2017)将谱带优选与 Gabor 变换相结合,通过测试分析不同谱带的变异与相关性,去除了相关性大的谱带对识别性能的影响。Ghasemzadeh 和 Demirel(2018)将 3 维小波变换应用高光谱人脸数据,在空间域和时间域同时进行小波变换,提取空谱联合特征用于高光谱人脸识别。谢志华等人(2019)提出了基于多谱带循环训练的深度学习高光谱人脸识别方法,表明了卷积神经网络对高光谱人脸识别特征提取的有效性。

目前,高光谱人脸识别研究主要集中在设计有效的手工特征和统计特征(Shen 和 Zheng, 2012)。此外,谱带选择主要是基于整体图像的经验谱带提取,不能实现面向人脸图像的自动最优谱带提取(魏冬梅等,2017)。鉴于高光谱成像的低信噪比、数据复杂性和人脸不同波段光谱性能的差异,研究适合人脸识别的高光谱谱带组合和面向高光谱人脸识别的深度特征是非常有必要的(Ghasemzadeh 和 Demirel, 2018)。本文在借鉴深度卷积神经网络可见光人脸识别研究的经验上,考虑高光谱人脸成像分布的特殊性,提出了基于分块局部二值模式(local binary pattern, LBP)特征 AdaBoost SVM (support vector machine) 谱带选择和小型 VGG net (visual geometry group network) 网络的高光谱人脸识别方法,并利用堆栈自编码器(stack auto-encoder, SAE)实现谱带间深度特征的有效融合与分类。提出的高光谱人脸识别方法框图如图 1 所示, η 表示初始波段数。该高光谱人脸识别的主要贡献有:1) 为了改善高光谱人脸数据“非对齐”因素对高光谱人脸谱带选择的影响,提出采用 LBP 直方图特征代替原有基于像素的特征,并采用分块方式选取人脸不同部位的最优谱带组合;2) 联合有监督深度网络(VGG12)和非监督学习网络(SAE)充分挖掘高光谱人脸的谱内和谱间特征。

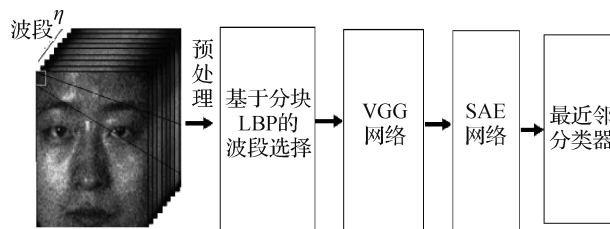


图 1 高光谱人脸识别框图

Fig. 1 The flow chart of hyperspectral face recognition

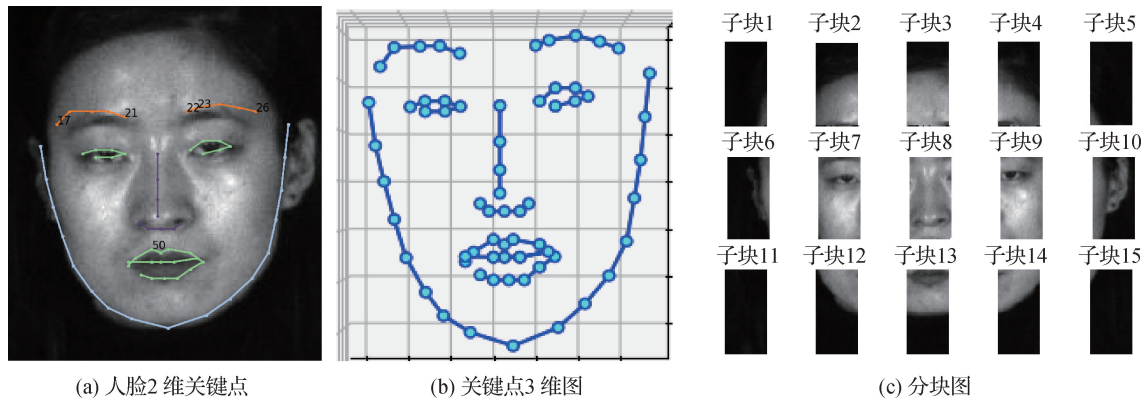
1 谱带优选

已有高光谱图像分类研究显示,随着参与运算谱带数目的增加,会出现分类精度先增后降的现象,称为高光谱图像分类中的休斯现象(魏冬梅等, 2017)。所以,有效的谱带选择可以减少冗余谱带对识别性能的影响。特别地,对基于深度网络的高光谱人脸特征提取,可以大幅提升深度学习网络训练的收敛速度和计算复杂度(杨楠等,2018)。

传统的高光谱人脸识别的谱带选择都是基于整幅图像,假设谱间人脸严格对齐且不同部位的最优谱带组合都相同,谱带选择性能还有待进一步提高(于洋等,2017)。为了改善高光谱人脸数据的非对齐因素对高光谱人脸谱带选择的影响,本文提出采用 LBP 直方图特征代替原有基于像素的特征,并采用分块方式为人脸不同部分提取对应的最优谱带组合。本文的 LBP 分块方式依据人脸三庭五眼的格局(魏冬梅等,2017),将高光谱人脸按 3×5 格式分块。为了保证不同的人脸器官划分到不同的子块,采用级联回归的方法,自动标注人脸 68 (0 ~ 67 号) 个关键点,并构建对应的 3D (three dimensions) 关键点(Bulat 和 Tzimiropoulos, 2020)。子块划分的横坐标采用序号为 17、21、22、26 的关键点横坐标;纵坐标采用序号为 23 和 50 的关键点纵坐标。人脸 3×5 分块图如图 2 所示。可以看出,人脸的主要特征:眼睛、鼻子、嘴巴和额头分在不同的小块中,这些小块表示不同器官具有相似的频谱属性。基于每个小块的 LBP 直方图特征,采用 AdaBoost SVM 选择谱带,自动抽取不同生理器官的最优谱带。

1.1 高光谱人脸的 LBP 特征表示

局部二值模式具有简单、有效地提取图像中微结构信息的优点,广泛应用于提取图像局部特征

图2 人脸 3×5 分块图Fig. 2 Face partition mode (3×5) ((a) 2D landmarks in a face; (b) landmarks in 3D coordinate; (c) partition maps)

(Ryu 等, 2017)。LBP 特征提取后, 通常采用模式编码的直方图统计特征来进行图像纹理表示, 对于图像间的非对齐影响, 具有一定的鲁棒性 (Bulat 和 Tzimiropoulos, 2020)。LBP 模式编码原理为

$$LBP(c) = \sum_{p=0}^{P-1} s(\tau_p - \tau_c) 2^p$$

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

式中, $LBP(c)$ 表示中心点 c 的编码模式, s 为符号函数, τ_c 是中心像素值, τ_p 是周围像素, $p \in [0, P-1]$, P 是周围像素个数, 编码后的 LBP 特征有 2^P 种。为了保留图像中的空间位置信息, LBP 模式编码后, 对其进行分块模式直方图统计, 具体公式为

$$H(r) = \sum_{x_c=2}^{N-1} \sum_{y_c=2}^{M-1} f(LBP(c), r) \quad (2)$$

式中, $H(r)$ 表示直方图函数, c 为编码的中心元素, $LBP(c)$ 由式 (1) 定义, r 表示编码值; 当 x 等于 y 时, $f(x, y)$ 取值为 1, 否则 $f(x, y)$ 取值为 0; N 和 M 表示分块 LBP 区域的长和宽。

1.2 基于 AdaBoost SVM 的谱带选择

本文对高光谱人脸图像采用 3×5 的方式进行分块, 每一分块采用 LBP 直方图作为该块的谱带选择特征, 降低非对齐因素对谱带选择的影响。AdaBoost 算法作为一种代表性的同质集成学习方法, 通过改变训练数据的分布和弱分类器的权重, 将多个弱分类器组合成一个强分类器 (朱健翔 等, 2006)。本文基于每个谱带特征建立弱分类器, 按照 AdaBoost 算法思路, 集成选择弱分类就完成了高光谱的谱带选择 (曹莹 等, 2013)。AdaBoost 通过改变数据间分布差异, 能够充分融合不同谱带数据的互补特

性, 从而为不同人脸区域自动提取最优的谱带组合。

由于支持向量机 (support vector machine, SVM) 在非均匀数据分布下的良好性能, 本文方法采用 SVM 作为 AdaBoost 的基学习器模型。假设谱带选择任务是从 η 个多光谱谱带中提取 K 个最优组合谱带来完成人脸识别。在 AdaBoost 的初始化阶段, 每个训练样本 $\mathbf{x}$ 都设置为 $1/M$, M 为训练样本的个数。在 AdaBoost 的第一轮循环, 确定性能最优的谱带为所提取谱带。训练集中每个样本的权重根据预测结果是否正确来更新。如果样本 $\mathbf{x}$ 的预测正确, 则减少其相应的权重, 相反如果预测错误, 则增加权重。在后续轮循环中, 基于对上轮所选谱带的互补性, 每次都选择最优的谱带。最后, 选择最优的 K 个谱带, AdaBoost 集成所选 K 个已经学习的 SVM 模型。基于最优的 K 个谱带和 K 个 SVM 分类器, 通过加权投票的规则得到集成学习结果。假设谱带选择训练数据为 $\{(\mathbf{x}_{([\lambda_1, \dots, \lambda_\eta], i)}, y_i)\}$, $i = 1, \dots, M$, $y_i \in \{1, \dots, C\}$, $\mathbf{B}$ 表示初始谱带集合, $\mathbf{T}$ 为构建的基分类器集合, $\mathbf{F}$ 为输出的波段集合, η 表示初始波段数, 基于分块 LBP 特征的 AdaBoost SVM 谱带优选算法流程如下:

- 1) 初始化权重: $w_i = 1/M$, 定义谱带 $\mathbf{B} = 1:\eta$;
- 2) $\mathbf{T} = \emptyset$, $\mathbf{F} = \emptyset$,
For $k = 1$ to K do,
For $b = \mathbf{B}$ do,
 $\mathbf{D}_b = \{(\mathbf{x}_{(\lambda_b, 1)}, y_1), \dots, (\mathbf{x}_{(\lambda_b, M)}, y_M)\}$,
基于 SVM 方法, 用权重 w_i 从样本集 $\mathbf{D}_d$,
中学到一个 $\mathbf{h}_b$ 基分类器,

Acc_b = 分类模型 h_b 的识别率,

$T = T \cup \{Acc_b\}$,

End,

计算 $\arg \max_{b' \in B} T_{b'}$, 将识别率最高的谱带 b' 作为提取的谱带 $b' \in B$,

$$\varepsilon_{b'} = \sum_{h_{b'}(x(\lambda_{b'}, i)) \neq y_i} w_i,$$

$$\beta_{b'} = \varepsilon_{b'} / (1 - \varepsilon_{b'}),$$

For 1 to M :

If $h_{b'}(x(\lambda_{b'}, i)) = y_i$,

Then $w_i = w_i \times \beta_{b'}$,

$$s = \sum_i w_i$$

For all $w_i, w_i = w_i / s$,

$$W_k = \log(1/\beta_{b'}),$$

$$B = B \setminus \{b'\},$$

$$F = F \cup \{b'\},$$

End;

3) 输出谱带选择结果: F 。

2 VGG12 网络模型

目前, 高光谱人脸数据库相对于可见光人脸数据库, 规模较小。如果沿用已有可见光人脸深度网络(VGG net 和 Deep ID)对高光谱人脸数据进行表示, 容易导致过拟合问题(Szegedy 等, 2015; Simon-yan 和 Zisserman, 2015)。基于 VGG net 网络基本结构(于洋 等, 2017), 构建了面向高光谱人脸数据的小型卷积神经网络模型(张国山 等, 2018), 共 12 层, 每个卷积核的大小都是 3×3 , 结构如图 3 所示, 本文将该深度模型简称为 VGG12 网络。

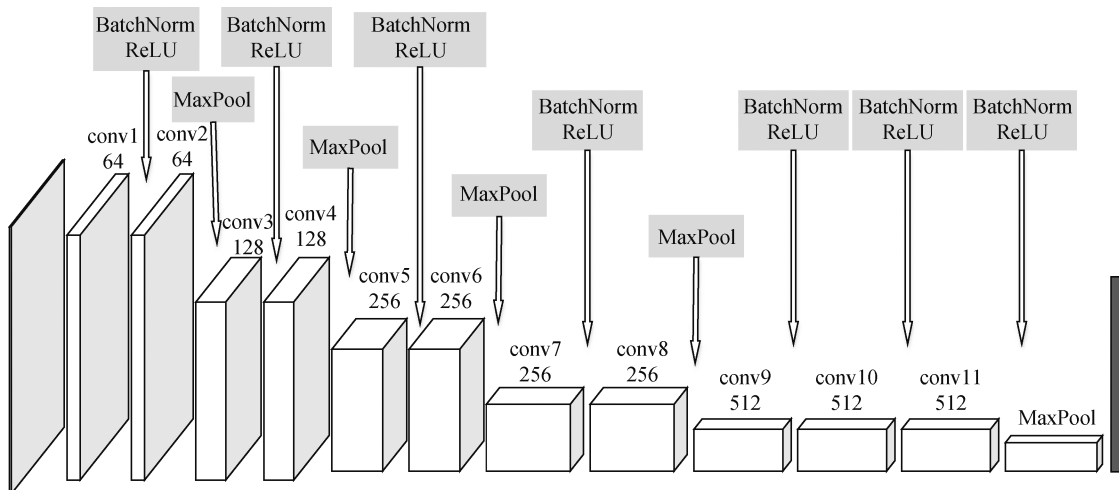


图 3 高光谱人脸网络模型(VGG12)

Fig. 3 Hyperspectral face deep neural network model (VGG12)

在 VGG12 中, 第 1 层和第 2 层(即 conv1 和 conv2)为 64 通道的卷积层, 步长(stride)取值为 1, 周围填充(padding)设置为 1, 达到 conv1 和 conv2 不降维, 提取更多的特征。conv2 卷积后接一个最大池化层(pool)。conv3 卷积层的卷积核数量在池化后增加一倍, 因此 conv3 的图像特征大小是 conv2 的两倍。经过 conv3、conv4 以及池化后, 特征维数降低到原始特征的 1/4。从 conv1 层一直到 conv8 层, 都是两个连续卷积层后接一个最大池化层的结构。conv9、conv10、conv11 这 3 层都是具有相同数量和大小卷积核的卷积运算。这 3 层后面接一个 MaxPool 操作, 达到减少特征个数的目的。最后一

层(第 12 层)即全连接层, 完成特征输出。

用于训练 VGG12 网络的损失函数为交叉熵与 SoftMax 相结合(Krizhevsky 等, 2012)。VGG12 网络的训练基于传统 VGG net 三通道数据处理的特点, 采用多谱带循环方式对高光谱人脸数据进行处理。在训练阶段, 第 1、2、3 谱带组合为一个样本, 第 2、3、4 构成一个样本, 依次循环到倒数第 3 个谱带。VGG12 网络训练采用适应性矩估计(Adam)优化算法, 学习率(α)为 0.001, 一阶矩估计的指数衰减率 β_1 为 0.9, β_2 为 0.999, batch 大小为 10。

VGG12 是一个典型的有监督深度学习网络。对于训练好的 VGG12 网络, 利用全连接层的特征作

为每个谱带的输出特征,同时将不同谱带的特征并联成深度网络的输出特征。由于每个谱带都对应一个深度特征矢量,VGG12输出层的简单谱带并联后,输出的特征维有一定增加,且简单融合没有考虑深度特征对识别的重要程度(即不同谱带深度特征的表达能力不一样),融合效率还有提升空间。

3 基于SAE的特征降维和融合

波段优选可以为高光谱人脸识别选择最优波段子集,VGG12网络提取了波段中的空间深度特征。为了对简单连接后的谱带深度特征进行有效融合和降维,同时整体考虑谱带深度特征间的不同表示能力,本文利用具有“特征转化”和“数据表示”优越性能的堆栈自编码器(Parchami等,2017),对不同波段的VGG12网络输出谱带特征进行深度融合。堆栈自编码器是一种无监督的特征学习网络,利用反向传播算法,使目标输出值等于输入值,具有较好的特征转换和数据降维能力。自编码器优化问题为

$$\min \| \text{sigm}(\mathbf{W}^T \text{sigm}(\mathbf{W}x^*)) - x \|_2 \quad (3)$$

式中, x 表示原始数据, x^* 表示污染后的数据, sigm 表示sigmoid函数, $\mathbf{W}$ 表示一个 $m \times n$ 的矩阵。为了有效进行谱带深度特征融合和降维,基于堆栈自编码器良好的信号表达性能,本文采用简单的3层SAE网络对连接后的深度网络特征降维。3层SAE网络结构如图4所示,第1、2、3层神经元个数分别为 $K^*VL/2$ 、 $K^*VL/4$ 、 $K^*VL/8$,其中 K 为波段数, VL 表示VGG12网络的输出维数。自编码器输入特征维数为VGG12网络最后全连接层的输出特征在波段上的联合。

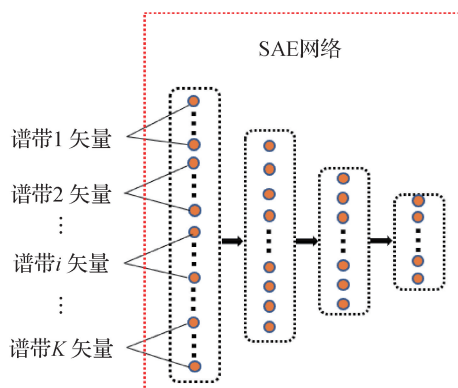


图4 三层SAE结构图

Fig. 4 The three-layers structure of SAE

4 实验和分析

目前,在光谱人脸识别研究中,已经公开了UWA-HSFD (University of Western Australia hyperspectral face database) 和 PolyU-HSFD (Hong Kong Polytechnic University hyperspectral face database) 两个数据集(Uzair等,2015)。UWA-HSFD数据库的光谱范围为450~1090 nm,谱带间隔为10 nm,共33个谱带和70个人,每人120幅图像。PolyU-HSFD是香港理工大学公开的高光谱人脸数据库,谱带范围为400~720 nm,谱带取样间隔为10 nm,共33个谱带和48个人,每人至少正脸、左侧脸、右侧脸3个样本。试验中,两个数据集预处理归一化为 220×180 像素的高光谱人脸图像,对两个数据库的训练集和测试集的划分比例为1:1,即不同光谱人脸图像的前一半样本作为训练集,剩下的做测试集。高光谱数据采用多谱带循环训练方式训练VGG12网络。该网络训练的优化准则为Adam方法,超参数设置如下: α 为0.001, β_1 为0.9, β_2 为0.999,epoch为10。SAE网络的学习率为0.0001,epoch为10。

为了确定谱带选择算法中谱带的个数 K ,在 K 取值不同的情况下,本文提出的方法(不包含SAE模块和最近邻居分类器)在两个数据库的实验识别率如图5所示。可以看出,在谱带选择量达到一定数值的时候,高光谱人脸的识别率不会一直随着谱带数的增加而提高;PolyU-HSFD在谱带个数(K)增加到8的时候,得到了最好的识别率95.2%, K 在8~10之间识别率保持不变,当 K 大于10以后识别率缓慢下降;UWA-HSFD在 K 从1增加到6时,

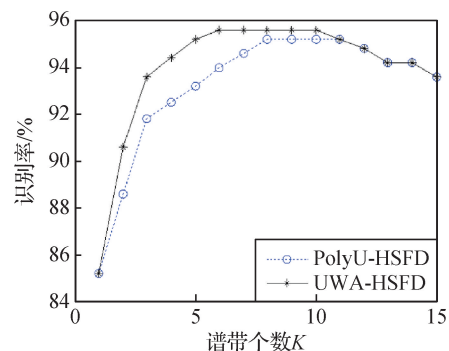


图5 不同 K 值识别结果

Fig. 5 The recognition rates of different K values

识别率是单调增加的,当 K 等于 6~10 时取得了最高的识别率 95.8%。基于上述结果,为了避免高光谱人脸识别中的休斯现象,同时缓解冗余谱带对后续深度网络训练的影响,本文波段选择对于 PolyU-HSFD 数据 K 取值为 8,UWA-HSFD 数据库 K 取值为 6。

为了验证本文提出的谱带选择算法的性能,试

验分别用像素特征和分块 LBP 直方图特征作为 AdaBoost SVM 谱带选择算法的输入特征,进行高光谱人脸识别,对比结果如图 6 所示。可以看出,基于分块 LBP 特征的谱带选择算法优于传统基于整幅图像像素的方法,主要原因是分块 LBP 特征能有效缓解多光谱图像中的非对齐影响,而且能为不同的人脸区域选择不同的最优谱带组合。

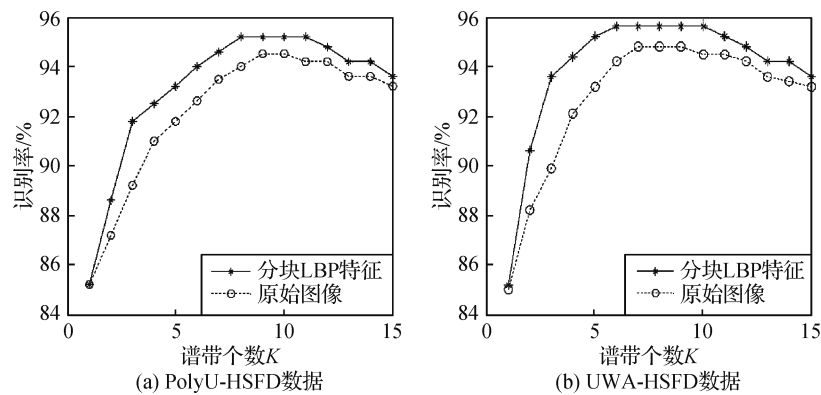


图 6 不同特征谱带选择识别结果

Fig. 6 The recognition rates of different bands selection algorithm((a) PolyU-HSFD; (b) UWA-HSFD)

为了验证本文提出的高光谱人脸识别方法的有效性,在 UWA 和 PolyU 两个高光谱数据集上,与 AlexNet(Krizhevsk 等,2012)、VGG net(Simonyan 和 Zisserman, 2015)、GoogLeNet (Mollahosseini 等, 2016)、Deep ID(Sun 等,2014) 和 DeepFace(Taigman 等,2014) 等深度网络及其他高光谱人脸识别方法进行对比实验,结果如表 1 所示。可以看出,1) 提出的高光谱人脸识别方法相比已有深度学习网络和方法,在两个数据集上都取得了最高的识别率(96.8% 和 97.2%),说明传统可见光人脸深度网络不适合高光谱人脸深度特征的提取。2) 多谱带循环 + VGG12 方法(谢志华 等,2019) 的识别结果(PolyU 为 94.8% ,UWA 为 95.2) 是手动选择中质量最好的谱带测试结果,但比本文提出方法的识别率要低,说明本文提出的分块自动谱带选择方法比手动根据图像质量选择谱带更适合高光谱人脸识别任务,能实现自动地最优谱带选择。3) 相比于所有谱带融合,SAE 网络对于谱带选择后的识别率提升更明显,说明提出的 SAE 谱带特征融合模块和谱带优选的结合能有效改善高光谱人脸识别的性能。

基于谱带选择特征和原始谱带特征下,VGG12 网络和 SAE 网络训练在单个 GPU(NVIDIA GTX 8 GB) 配置下测试,不同结构网络训练消耗时间对

表 1 不同深度网络结果

Table 1 The recognition rates of different methods

网络	识别率/%	
	PolyU-HSFD	UWA-HSFD
AlexNet	65.3	60.5
VGG net	83.6	86.7
GoogLeNet	81.4	74.6
Deep ID	83.6	71.8
DeepFace	74.2	75.2
VGG12	87.6	90.2
VGG + SAE	87.1	90.7
多谱带循环 + VGG12	94.8	95.2
谱带优选 + VGG12	95.2	95.8
谱带优选 + VGG12 + SAE	96.8	97.2

注:加粗字体表示各列最优结果。

比结果如表 2 所示。可以看出,1) 谱带选择由于挑选了鉴别和互补能力强的波段组合,改善了深度网络训练的效率。2) 在原始谱带下,SAE 在两个数据集上的复杂度都有一定的增加;但是经过谱带优选后,由于选择后的谱带具有较好的互补性,可以有效降低深度特征融合的难度,SAE 训练复杂度有较大的改善。在 PolyU-HSFD 数据集上 SAE 的训练时间

从 124 min 减少到 28 min; 在 UWA-HSFD 数据集上训练耗时从 133 min 减少到 35 min。在最后的整体框架中, SAE 训练耗时在原有基础上增加了约 15%。

综合表 1 和表 2, 谱带选择和深度网络的结合是有效的, 提出的谱带选择算法能够在降低深度网络训练复杂度的同时改善识别性能; SAE 以较小的时间复杂度进一步提升了最终的识别精度。

最后, 将本文方法与已有手工特征方法进行对比, 结果如表 3 所示。可以看出: 1) 提出的 VGG12 深度网络比传统手工特征取得了更高的识别率, 说明深度特征对高光谱人脸特征表示的有效性。2) 谱带选择对深度网络识别性能的提升效果比传统方法性能的提升更明显, 说明提出的谱带选择改善了深度网络模型的泛化能力, 缓解了深度网络训练的过拟合问题。

表 2 网络训练时间对比表

Table 2 Consuming time of different network structures

输入特征	时间/min	
	PolyU-HSFD	UWA-HSFD
谱带选择特征 + VGG	243	251
原始谱带特征 + VGG	327	369
谱带选择特征 + VGG + SAE	271	286
原始谱带特征 + VGG + SAE	451	502

注: 加粗字体表示各列最优结果。

表 3 深度特征与手工特征对比结果

Table 3 Recognition results of deep features and crafted features

特征表示	识别率/%	
	PolyU-HSFD	UWA-HSFD
2 维 Gabor 变换	85.9	81.6
3 维小波变换	81.6	83.6
LBP	79.5	78.7
VGG12	87.6	90.2
谱带选择 + Gabor 变换	87.1	85.7
谱带选择 + 3 维小波变换	81.2	84.2
谱带选择 + LBP	83.3	84.1
谱带选择 + VGG12 + SAE	96.8	97.2

注: 加粗字体表示各列最优结果。

5 结 论

针对高光谱人脸数据模式复杂、冗余度高的特点, 本文提出了波段选择结合深度特征的高光谱人脸识别方法。为抽取适合高光谱人脸识别的谱带, 采用基于分块 LBP 直方图的 AdaBoost SVM 谱带选择算法, 提取人脸不同区域的最优波段。基于高光谱人脸成像的特点和数据规模限制, 构建了一个有效的 VGG12 深度学习网络。为充分发挥监督和非监督特征学习方法的特征表示能力, 将 VGG12 全连接层特征组合后, 送入 3 层 SAE 进行特征降维和融合。最后, 基于最近邻分类器输出识别结果。

在高光谱人脸数据库的实验结果表明, 有效的谱带选择能改善深度网络对高光谱人脸判别特征的代表能力, 提出的深度学习高光谱人脸识别方法较手工特征提取方法和 Deep ID、DeepFace、VGG net 等深度网络方法取得了更高的识别率。

本文联合波段优先和深度特征表示来提高高光谱人脸识别的性能。波段选择和深度网络优化互相独立, 无法达到端到端的联合优化。后续工作考虑将波段选择模块集成到整个深度网络中, 研究联合优化“波段选择和空间特征表示”这两个目标的网络模型, 进一步探索新的深度网络结构以实现鲁棒的高光谱人脸识别。

参考文献 (References)

- Bulat A and Tzimiropoulos G. 2020. Hierarchical binary CNNs for landmark localization with limited resources. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42 (2): 343-356 [DOI: 10.1109/TPAMI.2018.2866051]
- Gao Y, Miao Q G, Liu J C and Gao L. 2013. Advance and prospects of AdaBoost algorithm. *Acta Automatica Sinica*, 39(6): 745-758 (曹莹, 苗启广, 刘家辰, 高琳. 2013. AdaBoost 算法研究进展与展望. *自动化学报*, 39(6): 745-758) [DOI: 10.3724/SP.J.1004.2013.00745]
- Chang H, Koschan A, Abidi B and Abidi M. 2010. Fusing continuous spectral images for face recognition under indoor and outdoor illuminants. *Machine Vision and Applications*, 21 (2): 201-215 [DOI: 10.1007/s00138-008-0151-1]
- Di W, Zhang L, Zhang D and Pan Q. 2010. Studies on hyperspectral face recognition in visible spectrum with feature band selection. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40 (6): 1354-1361 [DOI: 10.1109/TSMCA.2010.2052603]

- Ghasemzadeh A and Demirel H. 2018. 3D discrete wavelet transform-based feature extraction for hyperspectral face recognition. *IET Biometrics*, 7(1): 49-55 [DOI: 10.1049/iet-bmt.2017.0082]
- Krizhevsky A, Sutskever I and Hinton G E. 2012. ImageNet classification with deep convolutional neural networks//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA; ACM Press: 1097-1105 [DOI: 10.1145/3065386]
- Mollahosseini A, Chan D and Mahoor M H. 2016. Going deeper in facial expression recognition using deep neural networks//Proceedings of 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Placid, USA; IEEE: 1-10 [DOI: 10.1109/WACV.2016.7477450]
- Osia N and Bourlai T. 2014. A spectral independent approach for physiological and geometric based face recognition in the visible, middle-wave and long-wave infrared bands. *Image and Vision Computing*, 32(11): 847-859 [DOI: 10.1016/j.imavis.2014.06.010]
- Pan Z H, Healey G E, Prasad M and Tromberg B J. 2003. Face recognition in hyperspectral images//Proceedings of 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Madison, USA; IEEE: 334-339 [DOI: 10.1109/CVPR.2003.1211372].
- Parchami M, Bashbaghi B, Granger E and Sayed S. 2017. Using deep autoencoders to learn robust domain-invariant representations for still-to-video face recognition//Proceedings of the 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). Lecce, Italy; IEEE: 1-6 [DOI: 10.1109/AVSS.2017.8078553]
- Ryu B, Rivera A R, Kim J and Chae O. 2017. Local directional ternary pattern for facial expression recognition. *IEEE Transactions on Image Processing*, 26(12): 6006-6018 [DOI: 10.1109/TIP.2017.2726010]
- Shen L L and Zheng S H. 2012. Hyperspectral face recognition using 3D Gabor wavelets//Proceedings of the 21st International Conference on Pattern Recognition. Tsukuba, Japan; IEEE: 1574-1577
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2020-05-13]. <https://arxiv.org/pdf/1409.1556.pdf>
- Sun Y, Wang X G and Tang X O. 2014. Deep learning face representation from predicting 10 000 classes//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA; IEEE: 1891-1898 [DOI: 10.1109/CVPR.2014.244]
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2015. Going deeper with convolutions//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA; IEEE: 1-9 [DOI: 10.1109/CVPR.2015.7298594]
- Taigman Y, Yang M, Ranzato M A and Wolf L. 2014. DeepFace: closing the gap to human-level performance in face verification//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, USA; IEEE: 1701-1708 [DOI: 10.1109/CVPR.2014.220]
- Uzair M, Mahmood A and Mian A. 2015. Hyperspectral face recognition with spatio-spectral information fusion and PLS regression. *IEEE Transactions on Image Processing*, 24(3): 1127-1137 [DOI: 10.1109/TIP.2015.2393057]
- Wei D M, Zhang L R, Hu N N, Liu L, Ma N and Zhao Y F. 2017. Hyperspectral face recognition with spatial-spectral fusion information and Gabor feature. *Transactions of Beijing Institute of Technology*, 37(10): 1077-1083 (魏冬梅, 张立人, 胡楠楠, 刘璐, 马娜, 赵曰峰. 2017. 联合空谱信息和 Gabor 特征的高光谱人脸识别算法. 北京理工大学学报, 37(10): 1077-1083) [DOI: 10.15918/j. tbit1001-0645.2017.10.016]
- Xie Z H, Jiang P, Yu X H and Zhang S. 2019. Hyperspectral face recognition system based on VGGNet and multi-band recurrent network. *Journal of Computer Applications*, 39(2): 388-391 (谢志华, 江鹏, 余新河, 张帅. 2019. 基于 VGGNet 和多谱带循环网络的高光谱人脸识别系统. 计算机应用, 39(2): 388-391) [DOI: 10.11772/j. issn. 1001-9081.2018081788]
- Yang N, Nan L, Zhang D Y and Ku T. 2018. Research on image interpretation based on deep learning. *Infrared and Laser Engineering*, 47(2): #0203002 (杨楠, 南琳, 张丁一, 库涛. 2018. 基于深度学习的图像描述研究. 红外与激光工程, 47(2): #0203002) [DOI: 10.3788/IRLA201847.0203002]
- Yu Y, Zhang X, Liao Y and Zhang Z Y. 2017. Visible-near infrared hyperspectral imaging combined with chemometric methods to distinguish human facial information. *Spectroscopy and Spectral Analysis*, 37(8): 2339-2345 (于洋, 张欣, 廖奕, 张卓勇. 2017. 可见—近红外高光谱成像结合化学计量学分辨人面部信息. 光谱学与光谱分析, 37(8): 2339-2345) [DOI: 10.3964/j. issn. 1000-0593(2017)08-2339-07]
- Zhang G S, Zhang P C and Wang X B. 2018. Visual place recognition based on multi-level feature difference map. *Infrared and Laser Engineering*, 47(2): #0203004 (张国山, 张培崇, 王欣博. 2018. 基于多层次特征差异图的视觉场景识别. 红外与激光工程, 47(2): #0203004) [DOI: 10.3788/IRLA201847.0203004]
- Zhu J X, Su G D and Li Y C. 2006. Facial expression recognition based on Gabor feature and Adaboost. *Journal of Optoelectronics · Laser*, 17(8): 993-998 (朱健翔, 苏光大, 李迎春. 2006. 结合 Gabor 特征与 Adaboost 的人脸表情识别. 光电子 · 激光, 17(8): 993-998) [DOI: 10.3321/j. issn:1005-0086.2006.08.022]

作者简介



谢志华, 1977 年生, 男, 教授, 主要研究方向为计算机视觉。

E-mail: xie_zhihua68@aliyun.com

李毅, 男, 硕士研究生, 主要研究方向为模式识别。

E-mail: linfenliyi@gmail.com

牛杰一, 男, 硕士研究生, 主要研究方向为图像处理。

E-mail: niu_jieyi@163.com