



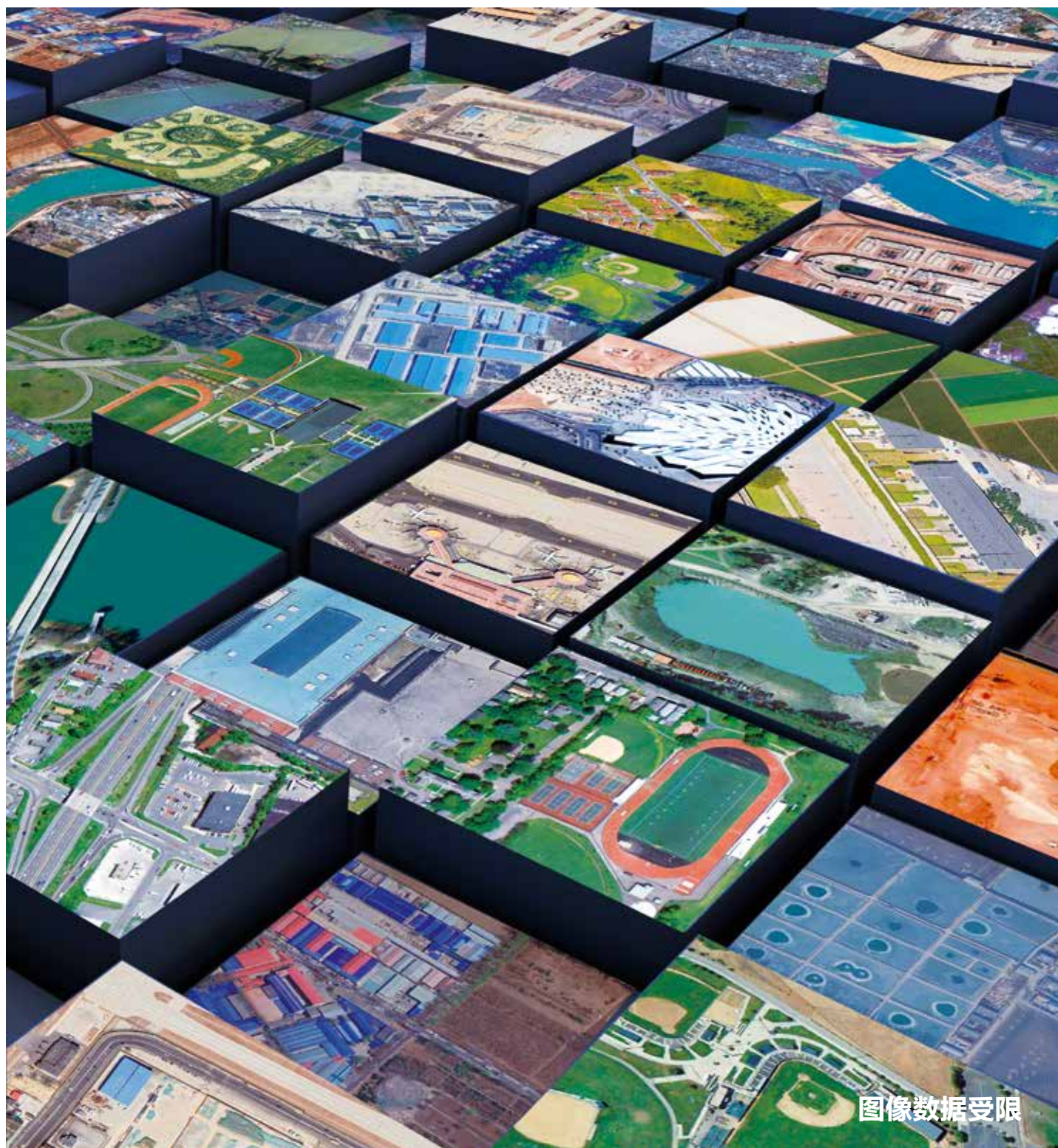
JOURNAL OF IMAGE AND GRAPHICS

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

2022
10
VOL.27

ISSN1006-8961
CN11-3758/TB



图像数据受限



第27卷第10期（总第318期）
2022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

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

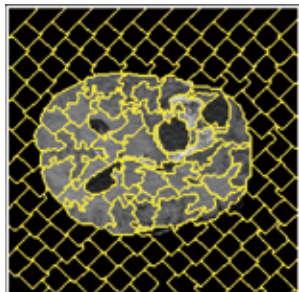
Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

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

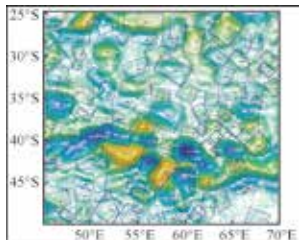
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



MRI脑肿瘤图像的超像素/体素分割及发展现状(第2897页)



融合多尺度特征与全局上下文信息的X光违禁物品检测(第3043页)



融合多尺度旋转锚机制的海洋中尺度涡自动检测(第3092页)

图像数据受限

《中国图象图形学报》图像数据受限专栏简介

刘怡光, 孙显, 赵启军, 魏秀参, 王琦, 陈秀妍 2801

数据受限条件下的多模态处理技术综述

王佩瑾, 闫志远, 容雪娥, 李俊希, 路晓男, 胡会扬, 严启炜, 孙显 2803

图像数据受限下的处理与分析

刘怡光 2835

面向跨模态行人重识别的单模态自监督信息挖掘

吴岸聪, 林城栋, 郑伟诗 2843

小样本条件下的RGB-D显著性物体检测

何静, 傅可人 2860

综述

面向目标检测的对抗样本综述

袁珑, 李秀梅, 潘振雄, 孙军梅, 肖蕾 2873

MRI脑肿瘤图像的超像素/体素分割及发展现状

方玲玲, 王欣 2897

图网络层级信息挖掘分类算法综述

魏文超, 蔺广逢, 廖开阳, 康晓兵, 赵凡 2916

个性化图像美学评价的研究进展与趋势

祝汉城, 周勇, 李雷达, 赵佳琦, 杜文亮 2937

视盘和视杯分割在计算机辅助青光眼诊断中的应用综述

方玲玲, 张丽榕 2952

图像处理和编码

多监督损失函数光滑化图像超分辨率重建

孟志青, 张晶, 邱健数 2972

轻量级注意力约束对齐网络的视频超分重建

靳雨桐, 宋慧慧, 刘青山 2984

面向图像修复的增强语义双解码器生成模型

王倩娜, 陈焱 2994

图像分析和识别

动态模态交互和特征自适应融合的RGBT跟踪

王福田, 张淑云, 李成龙, 罗斌 3010

采用Transformer网络的视频序列表情识别

陈港, 张石清, 赵小明 3022

对抗型半监督光伏面板故障检测

卢芳芳, 牛然, 杜海舟, 杨振辰, 陈菁菁 3031

融合多尺度特征与全局上下文信息的X光违禁物品检测

李晨, 张辉, 张邹铨, 车爱博, 王耀南 3043

高速公路场景的车路视觉协同行车安全预警算法

汪长春, 高尚兵, 蔡创新, 陈浩霖 3058

结合卷积神经网络与曲线拟合的人体尺寸测量

马燕, 殷志昂, 黄慧, 张玉萍 3068

医学图像处理

U-Net支气管超声弹性图像纵膈淋巴结分割

刘羽, 吴蓉蓉, 唐璐, 宋宁宁 3082

遥感图像处理

融合多尺度旋转锚机制的海洋中尺度涡自动检测

杜艳玲, 刘倩倩, 王丽丽, 徐鑫, 魏泉苗, 宋巍 3092

集成注意力机制和扩张卷积的道路提取模型

王勇, 曾祥强 3102

空域协同自编码器的高光谱异常检测

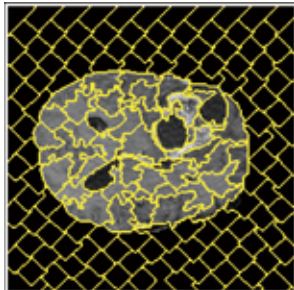
樊港辉, 马泳, 梅晓光, 黄珺, 樊凡, 李峰 3116

自适应权重金字塔和分支强相关的SAR图像舰船检测

郭伟, 申磊, 曲海成, 王雅萱, 林畅 3127

CONTENTS

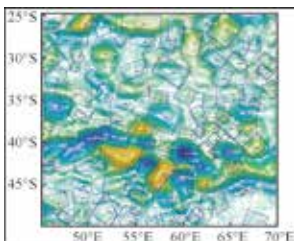
JOURNAL OF IMAGE AND GRAPHICS



The review of superpixel/voxel segmentation of MRI brain tumor images(P2897)



Integrated multi-scale features and global context in x-ray detection for prohibited items(P3043)



Multi-scale rotating anchor mechanism based automatic detection of ocean mesoscale eddy(P3092)

Limited Image Data

Review of multimodal data processing techniques with limited data

Wang Peijin, Yan Zhiyuan, Rong Xuee, Li Junxi, Lu Xiaonan, Hu Huiyang, Yan Qiwei, Sun Xian ... 2803

The processing and analyzing derived of limited image data

Liu Yiguang 2835

Single-modality self-supervised information mining for cross-modality person re-identification

Wu Ancong, Lin Chengzhi, Zheng Weishi 2843

RGB-D salient object detection of using few-shot learning

He Jing, Fu Keren 2860

Review

Review of adversarial examples for object detection

Yuan Long, Li Xiumei, Pan Zhenxiong, Sun Junmei, Xiao Lei 2873

The review of superpixel/voxel segmentation on MRI brain tumor images

Fang Lingling, Wang Xin 2897

Survey of graph network hierarchical information mining for classification

Wei Wenchao, Lin Guangfeng, Liao Kaiyang, Kang Xiaobing, Zhao Fan 2916

The review of personalized image aesthetics assessment

Zhu Hancheng, Zhou Yong, Li Leida, Zhao Jiaqi, Du Wenliang 2937

The review of optic disc and optic cup segmentation applications in computer-aided glaucoma diagnosis

Fang Lingling, Zhang Lirong 2952

Image Processing and Coding

Multi-supervision loss function based smoothed super-resolution image reconstruction

Meng Zhiqing, Zhang Jing, Qiu Jianshu 2972

Super-resolution Video frame reconstruction through lightweight attention constraint alignment network

Jin Yutong, Song Huihui, Liu Qingshan 2984

Enhanced semantic dual decoder generation model for image inpainting

Wang Qianna, Chen Yi 2994

Image Analysis and Recognition

RGBT tracking based on dynamic modal interaction and adaptive feature fusion

Wang Futian, Zhang Shuyun, Li Chenglong, Luo Bin 3010

Video sequence-based human facial expression recognition using Transformer networks

Chen Gang, Zhang Shiqing, Zhao Xiaoming 3022

Generative adversarial networks based semi-supervised fault detection for photovoltaic panel

Lu Fangfang, Niu Ran, Du Haizhou, Yang Zhenchen, Chen Jingjing 3031

Integrated multi-scale features and global context in x-ray detection for prohibited items

Li Chen, Zhang Hui, Zhang Zouquan, Che Aibo, Wang Yaonan 3043

Vehicle-road visual cooperative driving safety early warning algorithm for expressway scenes

Wang Changchun, Gao Shangbing, Cai Chuangxin, Chen Haolin 3058

The convolution neural network and curve fitting based human body size measurement

Ma Yan, Yin Zhiang, Huang Hui, Zhang Yuping 3068

Medical Image Processing

U-Net-based mediastinal lymph node segmentation method in bronchial ultrasound elastic images

Liu Yu, Wu Rongrong, Tang Lu, Song Ningning 3082

Remote Sensing Image Processing

Multi-scale rotating anchor mechanism based automatic detection of ocean mesoscale eddy

Du Yanling, Liu Qianqian, Wang Lili, Xu Xin, Wei Quanmiao, Song Wei 3092

Road extraction model derived from integrated attention mechanism and dilated convolution

Wang Yong, Zeng Xiangqiang 3102

Spatial-coordinated autoencoder for hyperspectral anomaly detection

Fan Ganghui, Ma Yong, Mei Xiaoguang, Huang Jun, Fan Fan, Li Hao 3116

Ship detection in SAR images based on adaptive weight pyramid and branch strong correlation

Guo Wei, Shen Lei, Qu Haicheng, Wang Yaxuan, Lin Chang 3127

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2022)10-2937-15

论文引用格式: Zhu H C, Zhou Y, Li L D, Zhao J Q and Du W L. 2022. The review of personalized image aesthetics assessment. Journal of Image and Graphics, 27(10): 2937-2951 (祝汉城, 周勇, 李雷达, 赵佳琦, 杜文亮. 2022. 个性化图像美学评价的研究进展与趋势. 中国图象图形学报, 27(10): 2837-2951) [DOI:10.11834/jig.210211]

个性化图像美学评价的研究进展与趋势

祝汉城^{1,2}, 周勇^{1,2*}, 李雷达³, 赵佳琦^{1,2}, 杜文亮^{1,2}

1. 中国矿业大学计算机科学与技术学院, 徐州 221116; 2. 矿山数字化教育部工程研究中心, 徐州 221116;
3. 西安电子科技大学人工智能学院, 西安 710071

摘要: 图像美学评价方法是当前研究的热点问题。图像美学评价分为大众化和个性化两种。大众化图像美学评价主要研究大多数人对图像共同的审美感知评估, 而个性化图像美学评价可以针对用户的个性化审美感知进行评估。现有的研究工作主要集中在大众化图像美学评价上, 但是由于人们对图像的审美体验具有高度主观性, 研究针对特定用户的个性化图像美学评价方法更加符合现实意义。目前研究人员针对个性化图像美学评价展开了相关研究, 并取得了一定的研究进展。但是现有的文献中缺少对个性化图像美学评价方法的综述, 本文针对个性化图像美学评价的研究进展与趋势进行概述。首先分析图像美学评价的研究现状与发展趋势; 然后针对现阶段的个性化图像美学评价模型进行概述, 将现有的个性化图像美学评价模型总结为基于协同过滤的模型、基于用户交互的模型和基于审美差异的模型, 并分析这3类模型主要的设计思路以及优缺点; 最后介绍个性化图像美学评价在精准营销、个性化推荐系统、个性化视觉增强和个性化艺术设计上的应用前景, 并指出未来研究工作在主观特性分析和知识驱动建模等方面的发展方向。

关键词: 个性化图像美学; 评价方法; 审美体验; 主观特性; 知识驱动

The review of personalized image aesthetics assessment

Zhu Hancheng^{1,2}, Zhou Yong^{1,2*}, Li Leida³, Zhao Jiaqi^{1,2}, Du Wenliang^{1,2}

1. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China;
2. Engineering Research Center of Mine Digitization of Ministry of Education, Xuzhou 221116, China;
3. School of Artificial Intelligence, Xidian University, Xi'an 710071, China

Abstract: The multimedia imaging technology can meet people's visual demands to a certain extent. People can easily obtain high-quality images through mobile devices, so people begin to pay more attention to their aesthetic experience of images, which makes the image aesthetics assessment (IAA) method become a hotspot issue and frontier technology in the current image processing and computer vision fields. Intelligent IAA can be developed to imitate people's aesthetic perception of images and predict the results of aesthetic evaluation automatically. Aesthetic-preference images can be screened

收稿日期: 2021-03-19; 修回日期: 2021-05-17; 预印本日期: 2021-05-25

* 通信作者: 周勇 yzhou@cumt.edu.cn

基金项目: 国家自然科学基金项目(62101555, 62171340, 62002360); 江苏省自然科学基金项目(BK20210488, BK20201346, BK20181354); 中国博士后科学基金面上项目(2022M713379); 中央高校基本科研业务费专项基金资助项目(2021QN1071); 江苏省六大高峰人才基金项目(2015-DZXX-010)

Supported by: National Natural Science Foundation of China (62101555, 62171340, 62002360); Natural Science Foundation of Jiangsu Province, China (BK20210488, BK20201346, BK20181354); Project Funded by China Postdoctoral Science Foundation (2022M713379); Fundamental Research Funds for the Central Universities (2021QN1071); Six Talent Peaks High-level Talents in Jiangsu Province (2015-DZXX-010)

out. Consequently, IAA is critical to be applied in photography, beauty, photo album management, interface design, and marketing. Generally, IAA can be classified into two categories, including generic image aesthetics assessment (GIAA) and personalized image aesthetics assessment (PIAA). Early researches believe that people have a consensus on the aesthetic experience of images, and leverage the general photography rules to describe most people's visual aesthetics on images, which are usually affected by many factors, such as light intensity, color richness, and composition. Most of the current GIAA model can predict most people's aesthetic evaluation results of images. GIAA models can be divided into three aesthetic-related tasks like classification, score regression and distribution prediction. The aesthetic classification task is focused on dividing the image into two classes of "high" and "low" according to the aesthetic experience of most people. The research goal of the aesthetic score regression task can predict the aesthetic score of an image. This task leverages the average aesthetic ratings of most people as the image aesthetic score for regression modeling. However, the two tasks shown above need to convert different people's aesthetic ratings of images into a unified result ("high" or "low" and score). Label uncertainty is derived from people's aesthetic experience of images, which makes it difficult for the consensus result. Therefore, the predictable aesthetic distribution is more concerned to reflect people's subjectivity. The goal of the aesthetic distribution prediction task can predict the aesthetic distribution results of multiple people's ratings of an image. This task predicts the aesthetic distribution straightforward and converts the aesthetic distribution result into aesthetic scores and aesthetic classes. Consequently, recent GIAA models researches are mainly focused on the task of aesthetic distribution prediction. Although the aesthetic distribution prediction task of the GIAA model can reflect people's subjectivity of image aesthetics to a certain extent, the task can realize people's visual aesthetic preferences from the image level only. Besides, it is more realistic to develop the PIAA model for specification beneficial from the growth of customized services. Therefore, the PIAA model has received great attention recently, which can predict the accurate aesthetic results for customized users. We introduce the existing PIAA models published from 2014 to 2020 due to the lack of reviews on PIAA models. Generally, the PIAA model faces two challenges for specific users as mentioned below: First, PIAA is a typical small sample learning task. This is because the PIAA model is a real-time system for specific users, which cannot force users to annotate a large number of images aesthetically. In addition, a small number of image samples can just be obtained for model training. Second, the user's subjective characteristics become important factors to affect their aesthetic perception of images since the user's aesthetic experience of images is highly subjective. Meanwhile, users' aesthetic experience is influenced by many subjective factors like emotion and personality traits. Therefore, the existing framework of the PIAA model is mainly divided into two stages based on machine learning or deep learning. In the first stage, the GIAA dataset rated by a large number of users is used to obtain the prior knowledge of the PIAA model through supervision training for the small sized sample learning issue of the PIAA task. In the second stage, a user's PIAA dataset is used for fine-tuning on the prior knowledge model for the high subjectivity of users' image aesthetic experience, and the subjective knowledge of users is integrated to obtain the PIAA model that conforms to the user's aesthetic perception. The existing PIAA models can be divided into three categories like collaborative filtering based PIAA models, user interaction based PIAA models, and aesthetic difference based PIAA models. To demonstrate the differences between these three PIAA models, we first introduce each of the three PIAA models separately. Then, we summarize the design clues, pros and cons of existing PIAA models. Meanwhile, we introduce the commonly used datasets and evaluation criteria of PIAA models, as well as the application prospect of PIAA models in precision marketing, personalized recommendation systems, personalized visual enhancement, and personalized art design. Finally, the future direction of the PIAA model is predicted in subjective characteristic analysis and knowledge-driven modeling.

Key words: personalized image aesthetics; evaluation metrics; aesthetic experience; subjective characteristics; knowledge-driven

0 引言

随着移动互联网的迅速发展,人们在日常生活中可以轻易通过移动设备获取大量图像数据,极大丰富了人们的视觉体验。图像的美与美感成为人们追求精神生活的重要部分(王伟凝等,2012)。图像美学评价(image aesthetics assessment, IAA)成为当前图像处理与计算机视觉领域的热点研究问题(Deng等,2017b;金鑫等,2018;Hosu等,2020)。图像美学评价可以应用于图像增强(Chaudhary等,2018;王浩等,2017)、图像检索(Vo等,2019;柯圣财等,2017)、相册管理(Kuzovkin,2019)和推荐系统(Zhang等,2017;黄立威等,2018)等多个实用场景。例如,当用户在图像搜索系统中输入山景这个词时,希望看到的图像内容是色彩丰富、令人愉悦的山景或构图较好的山峰,而不是灰色或模糊的山景。目前,图像美学评价已经出现在一些实际应用中,并取得一定的成功。例如,自动设计广告海报的鲁班系统、自动美化图像的美图秀秀软件和自动选取视频封面的腾讯视频软件等。图像美学评价具有重要的理论研究意义和实际应用价值。

早期的图像受设备成像技术和图像处理技术的限制会引入不同程度的噪声,研究人员主要通过图像质量评价(image quality assessment, IQA)方法来量化图像的失真程度(王志明,2015)。随着图像成像和处理等技术的逐渐成熟,人们不仅可以相对容易地获取到高质量图像,并且要求图像既具有高质量内容,又具有美感。例如,手机用户拍照时,对拍摄的图像不仅要求内容清晰,而且要求符合用户的审美体验。由于人类对图像的审美体验受到图像客观内容和人类自身心理因素的双重影响,因此图像美学评价需要心理学、美学和计算机视觉等多个交叉学科的共同支持,是极具挑战性的研究课题。

根据研究对象的不同,图像美学评价方法可分为两类:大众化图像美学评价(generic IAA, GIAA)和个性化图像美学评价(personalized IAA, PIAA)(Ren等,2017)。早期的IAA方法主要是将人类的视觉审美看成具有一致性,从图像层面上提取美学特征,研究面向大多数人的GIAA模型(王伟凝等,2014)。GIAA模型按照不同的研究目标主要可以分为3类:美学二分类(王伟凝等,2016)、美学分数

回归(Kong等,2016)和美学分布预测(Jin等,2018)。由于人们在日常生活中的审美评判标准存在较大差异,一千个人眼中有一千个哈姆雷特等谚语也充分说明人与人之间存在不同的审美偏好与兴趣,精准服务和私人定制等消费理念逐渐深入人心。因此,研究针对特定用户的PIAA模型(王伟凝等,2020)可以满足人们对个性化定制服务的需求,在现实生活中具有更加重要的研究意义与应用价值。图1展示了大众化与个性化图像美学评价方法的区别,其中美学分数范围为1~5分,分数越高表示图像的美学质量越高。可以看出,PIAA方法相比于GIAA方法可以更精确地推断特定用户对图像的审美评分。

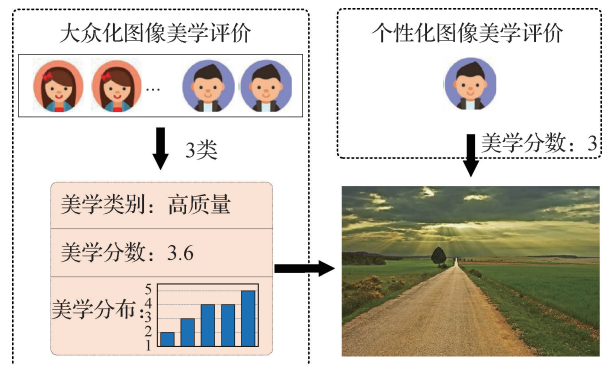


图1 大众化与个性化图像美学评价的区别

Fig. 1 Difference between generic and personalized image aesthetic assessment

目前,国内外关于图像美学评价的相关综述文献(王伟凝等,2012;Deng等,2017b;金鑫等,2018;白茹意等,2019;鲁越等,2020)主要集中在针对GIAA模型的研究上。王伟凝等人(2012)主要介绍早期利用手工特征和传统的机器学习来构建图像美学评价模型;Deng等人(2017b)从实验角度总结基于不同视觉特征(手工特征和深度特征)的方法在多个图像美学评价数据库上的对比结果和性能分析;金鑫等人(2018)回顾了图像美学评价的发展历史,从方法、新挑战和数据库构建的角度综述了该领域关键技术的发展情况;白茹意等人(2019)和鲁越等人(2020)从艺术绘画图像分类的角度详细分析了绘画图像美学评价方法的研究现状及发展。这些综述文献主要是从大众审美的角度探讨图像美学评价方法,缺少对PIAA模型相关研究进展的分析与总结。为进一步拓展和补充现有针对图像美学评价的综述文献,本文详细概括了PIAA模型的研究进

展与趋势。

1 图像美学评价

1.1 研究现状

从古至今,对美好事物的不断追求一直是人类精神生活中的重要部分(罗利建,2014),例如古代的艺术绘画和现代的摄影技术等。随着物质生活水平不断提高,人们不再满足于物品在功能上的实用性,还要求它们在视觉上具有美感(金鑫等,2018)。而图像作为人们日常生活中最广泛使用的视觉内容,使得图像美学成为当前美学与计算机视觉领域的重要研究课题(罗利建,2014)。图像美学的研究目标是利用科学的计算方法确定人类对图像在视觉上表现出的审美感知,也就是判断人类对给定图像的审美体验(Palmer等,2013)。对人类的视觉审美体验进行有效建模可以了解人们所需的审美情感诉求,在人们进行艺术创作时(如绘画和电影制作等)形成一套指导方针,以诱导感知者的神经系统产生情感上的审美共鸣(Cavanagh,2005)。

长久以来,研究人员对人类的审美体验是否可以通过科学方法进行计算存在一定争议,一些研究人员认为美学不可能通过科学方法进行研究,因为科学方法是客观的和合理的,而美学具有主观性和不确定性(Nanay等,2019)。但是随着科学计算方法的不断进步,大多数研究人员认为虽然人类的视觉审美体验是主观的,但是这并不阻碍利用客观的计算方法进行研究(Jacobsen,2006)。例如,人们对颜色的审美是主观的,并且不同人之间也存在很大的审美差异,然而这并没有阻碍色觉成为一门完善以及技术成熟的科学体系(Koenderink,2010)。虽然人类对图像的审美感知存在主观性和不确定性,但是科学的计算方法就是从人类的视觉审美感知中寻找规律,进而对图像的美感进行评估。因此,图像美学评价是一种主观性较强但可以通过客观计算方法进行研究的课题。

目前,图像美学评价吸引了国内外科研机构和院校的关注与研究。具有代表性的有清华大学、中国科学院大学、西安电子科技大学、北京电子科技学院、华南理工大学、山东财经大学、香港中文大学、美国谷歌研究院、美国宾夕法尼亚州立大学、美国加州大学欧文分校、美国罗格斯大学、新加坡南洋理工大

学和德国蒂宾根大学等。这些机构院校在计算机视觉与图像处理等相关领域顶级期刊和会议发表了图像美学评价相关研究成果,如《中国图象图形学报》、《计算机辅助设计与图形学学报》、TIP (IEEE Transactions on Image Processing) 和 CVPR (IEEE Conference on Computer Vision and Pattern Recognition) 等。图2展示了近10年在国内外期刊和会议发表的图像美学评价论文数量(Web of Science 和中国知网的统计结果)。可以看出,关于图像美学评价的研究成果呈逐年增长趋势,这些研究成果推动了图像美学评价的快速发展。此外,多媒体领域顶级国际会议(Association for Computing Machinery International Conference on Multimedia, ACM MM)于2020年举办了第1届面向多媒体美学与技术质量评价研讨会(The 1st Workshop on Aesthetic and Technical Quality Assessment of Multimedia),专门就当前图像美学评价的最新研究趋势、存在问题和未来方向进行意见交流和讨论,以期加快该领域的研究进展(Hosu等,2020)。

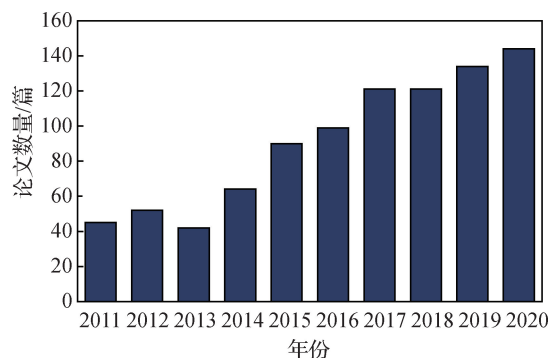


图2 近10年发表的图像美学评价论文数量

Fig. 2 The number of papers on image aesthetics assessment published in the past ten years

1.2 研究进展

图像美学评价方法的目的是利用可计算的智能系统有效地模拟人类对图像的审美体验,并自动给出对图像美学质量的感知评估分数(金鑫等,2018)。神经科学的研究表明人类的视觉审美体验是一种信息处理过程,包括5个阶段:感知、内隐记忆整合、内容与风格的明确分类和认知掌握和评价,最终产生审美判断(Leder等,2004)。从早期的心理美学(Martindale,2007)到现代的神经理美学,研究人员一致认为,无论来源、文化和经验如何,人类的审美体验与视觉刺激引起的感觉之间存在一定的关

联性(Zeki,2013)。因此,早期的研究人员认为人类在对图像的审美体验上通常具有共识性,可以通过摄影专家总结出的通用美学规则来描述图像的美感,这些规则会受到多种因素的影响,如光照强度、色彩丰富度和构图方式(Wells,2015)等;为了方便计算,早期研究方法通常利用机器学习或深度学习构建评价模型对图像进行美学二分类(王伟凝等,2016)或美学分数回归(Kong等,2016)。

图3展示了美学二分类和美学分数回归任务,其中美学分数范围为1~5分,分数越高表示图像美学质量越高。美学二分类任务的研究目标是将图像按照大多数人的审美评估分成“高”和“低”两类,这种任务相对简单且相应研究较为成熟。美学分数回归任务的研究目标是预测得到图像的美学评分,这种任务通常将多数人对图像美学的平均评分作为图像美学分数进行回归建模。可以看出这两种研究任务主要是通过确定的结果(“高”或“低”和分数)来研究符合大多数人审美体验的大众化图像美学评价(GIAA)模型。

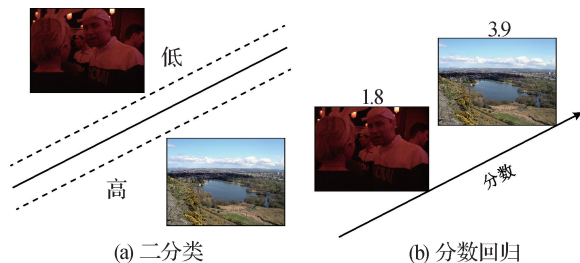


图3 图像美学二分类和分数回归的两种任务

Fig. 3 Two tasks of image aesthetics binary classification and score regression ((a) binary classification; (b) score regression)

但是,不论是美学二分类任务还是美学分数回归任务,都需要将不同人对图像的审美评估处理成统一结果(“高”/“低”以及分数),这种操作会在一定程度上引入标签不确定性,主要是因为人类对图像的审美具有较强的主观性,会导致这种统一的结果无法有效地描述不同人对图像审美感知结果。因此,可以反映人们对图像审美主观性的美学分布预测任务受到了研究者的关注(Hou等,2020;Xu等,2020)。如图4所示,美学分布预测任务的研究目标是预测得到不同人对图像评分的美学分布结果,其中美学分数范围为1~5分。该研究任务直接对图像获取的美学分布进行建模预测,并且可以通过计

算方法将美学分布转化成美学分数和美学二分类结果(Zhang等,2019b,2020;Zeng等,2020)。因此,针对GIAA模型的研究逐渐转向对美学分布预测任务的研究(Talebi和Milanfar,2018;Cui等,2019;Fang等,2018;Zhang等,2019b,2020;Zeng等,2020)。

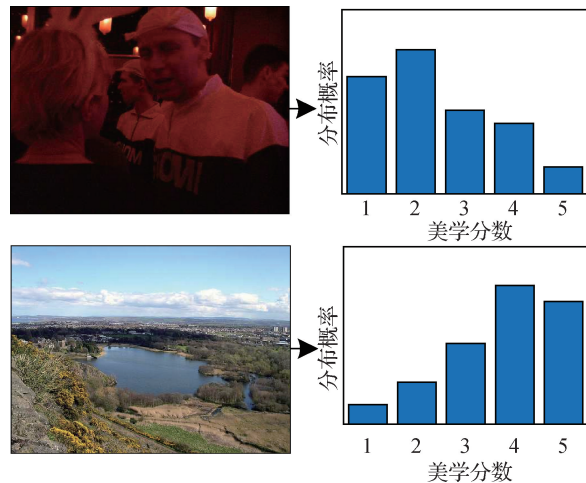


图4 图像美学分布预测任务

Fig. 4 Image aesthetic distribution prediction task

虽然图像美学分布可以在一定程度反映出人们对图像审美的主观性,但是该研究任务还只是从图像层面上来考虑人类的视觉审美特性,所以美学分布预测难以有效地推断特定用户对图像的审美评估。图5展示了5位用户对同一幅图像的美学评分。可以看出,不同用户对图像的审美差异较大,所以仅从面向多数人的审美体验设计的GIAA模型难以有效地适用于现实生活中以用户为中心的应用场景,如个人相册管理(Ceroni,2018)、个性化图像增强(Bianco等,2020)和个性化推荐系统(Zhang等,2019a)等。考虑到人类对图像的审美体验存在较强的个性化和主观性,因而构建符合特定用户审美体验的PIAA模型是当前图像美学评价研究的重要发展方向。

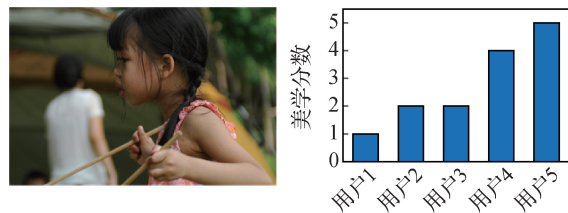


图5 图像和5位用户对应的美学评分

Fig. 5 Image and aesthetic rating corresponding to five users

2 个性化图像美学评价

2.1 面临挑战

个性化图像美学评价任务需要对特定用户的视觉审美进行研究。因为用户的审美体验受地域、年龄、性格、情感和行为习惯等多重因素的影响,因此针对特定用户的PIAA模型的研究相对复杂和困难,面临巨大挑战。首先,PIAA是一种典型小样本学习任务(Ren等,2017)。PIAA模型是面向特定用户的即时系统,无法要求用户对大量图像进行美学标注,通常情况下仅能获取用户对少量图像样本的标注数据用于模型构建。目前基于数据驱动的深度卷积神经网络的有效训练需要大量样本数据(Krizhevsky等,2012),因此,仅通过少量训练样本构建符合特定用户审美体验的PIAA模型是一个较大挑战。其次,由于用户对图像的审美体验具有高度主观性(Kim等,2020b),所以用户的主观特性是影响其对图像审美感知的重要因素。但是由于用户的主观

特性较为复杂,导致用户的审美体验会受到情感情绪(Joshi等,2011)和性格特征(Swami和Furnham,2014)等多方面影响,因此如何有效地提取用户主观特性并融入到PIAA模型中也是一个较大挑战。虽然面临上述挑战,但是由于PIAA模型可以较为精确地推断出特定用户对图像的个性化美学评价结果,在面向个体用户的智能终端上(例如,智能手机、智能手表和平板电脑等)有更加广阔的应用前景。

2.2 整体框架

图6是现有的PIAA模型的整体框架图,该框架分为两个阶段。第1阶段,主要针对PIAA任务的小样本学习问题,利用大量用户进行美学评价的GIAA数据集通过监督训练获取美学先验知识模型;第2阶段,主要针对用户视觉审美体验中的主观性问题,利用特定用户的PIAA数据集对美学先验知识模型进行微调训练,得到符合该用户个性化审美体验的PIAA模型。现有的PIAA模型分为3类:基于协同过滤的PIAA模型、基于用户交互的PIAA模型和基于审美差异的PIAA模型。

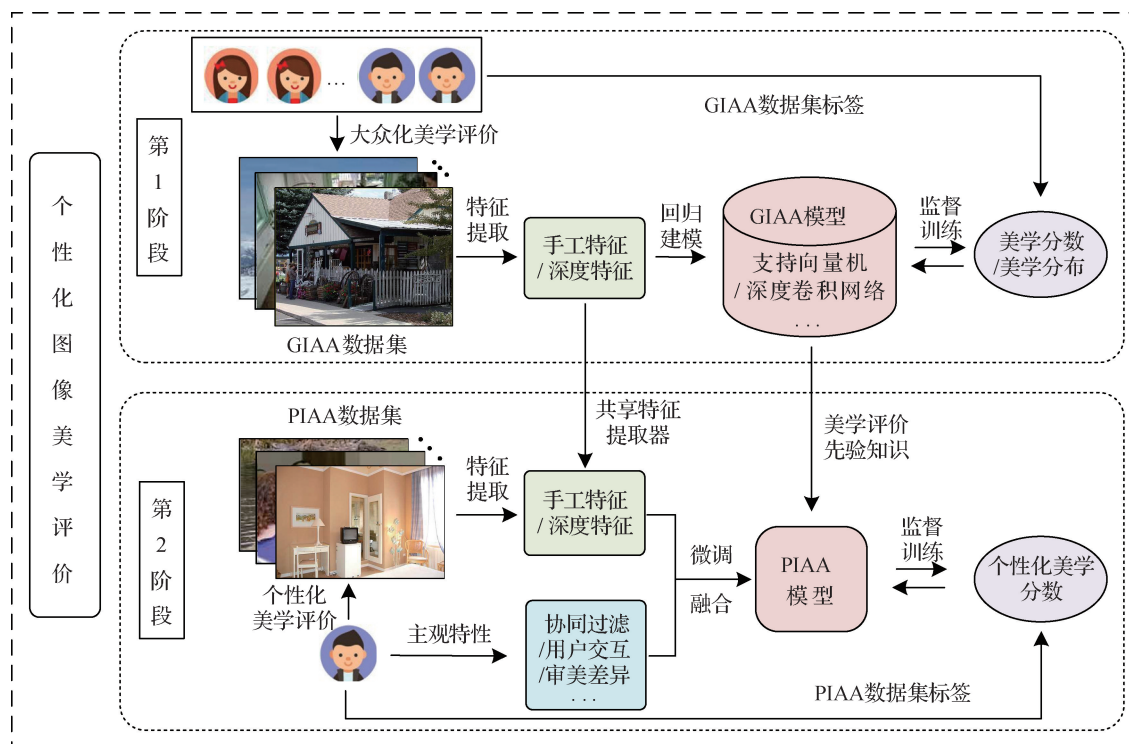


图6 个性化图像美学评价整体框架图

Fig. 6 The overall framework of personalized image aesthetics assessment

2.2.1 基于协同过滤的PIAA模型

由于针对特定用户的PIAA是小样本学习任

务,直接通过少量的标注样本训练PIAA模型存在极大的不确定性。因此需要借助于其他相似用户的

美学先验知识构建 PIAA 模型。如图 7 所示, 协同过滤是一种可以有效利用不同用户之间相似的审美偏好来推断特定用户的个性化图像美学评价方法 (O'Donovan 等, 2014)。

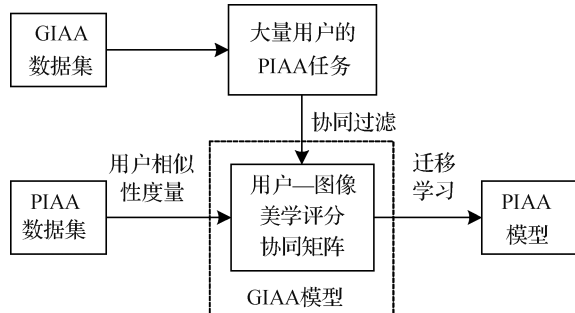


图 7 基于协同过滤的 PIAA 模型典型框架

Fig. 7 The typical framework of the PIAA model based on collaborative filtering

Deng 等人 (2017a) 提出使用社交网络上用户标记为“偏好”的样本图像表征用户的审美感知。首先通过协同过滤的方法提取相应偏好图像的视觉特征 (He 等, 2017) 来构建基于大众化审美感知的 GIAA 模型; 然后通过对已训练的 GIAA 模型进行个性化审美的迁移学习来实现针对特定用户的 PIAA 模型。Wang 等人 (2018) 在图像美学评价数据集 AVA (aesthetic visual analysis) (Murray 等, 2012) 的基础上收集用户对图像的文本评论数据。首先利用协同过滤方法设计了一个带有用户与图像关系编码输入的卷积神经网络 (convolutional neural network, CNN) 模型来训练 GIAA 模型, 然后提出一种注意力机制, 通过融合图像和用户评论信息来捕获用户对图像语义标签和兴趣区域, 最后得到符合用户个性化审美偏好的 PIAA 模型。Cui 等人 (2020) 首先利用大规模的图像美学评价数据集训练出 GIAA 模型, 然后在社交媒体平台上收集用户的偏好图像, 并结合用户的个性化审美偏好和大众化审美标准来解决用户偏好行为的不确定性, 最后通过协同过滤的方法对每个用户与图像之间的成对排序进行优化的方式来解决数据稀疏问题, 得到基于 CNN 的 PIAA 模型。从上述分析可以看出, 基于协同过滤的 PIAA 模型借助于用户对图像的审美相似性构成协同矩阵, 并通过目标用户与其他用户之间的相似性度量获取该用户对图像的美学评估。

2.2.2 基于用户交互的 PIAA 模型

基于用户交互的 PIAA 模型与基于协同过滤的 PIAA 模型类似, 主要目的也是解决小样本学习的问题。不同之处在于基于交互的 PIAA 模型需要目标用户自身的交互行为, 而基于协同过滤的 PIAA 模型则依赖于目标用户与其他用户相似的审美偏好。基于用户交互的 PIAA 模型如图 8 所示。

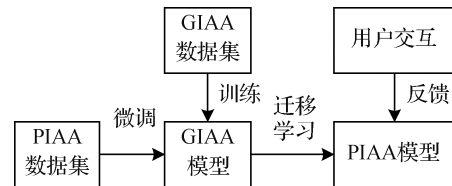


图 8 基于用户交互的 PIAA 模型典型框架

Fig. 8 The typical framework of the PIAA model based on user interaction

Yeh 等人 (2014) 提出一种针对用户审美偏好的交互式反馈系统, 通过用户对图像的颜色、纹理和组成等通用手工特征的权重进行调整来反馈其正向和负向的审美偏好行为, 实现用户对图像的个性化审美排序。Park 等人 (2017) 针对用户的个性化审美偏好提出一种基于支持向量机 (support vector machine, SVM) 的组合学习框架, 包括支持向量机回归 (support vector regression, SVR) 和排序支持向量机 (ranking SVM, R-SVM)。其中, SVR 用于学习一个基于大众化审美的 GIAA 模型, R-SVM 用于从交互系统中学习用户的个性化审美偏好来调整模型, 并通过最近邻搜索算法在 GIAA 数据集中识别与用户的 PIAA 数据集最相似的部分图像数据对该组合学习框架进行优化, 最终得到用户的 PIAA 模型。Lyu 等人 (2018) 提出一种基于用户交互的图像美学排序模型, 将特定用户偏好的一系列图像作为输入, 并输出符合用户审美偏好的图像美学排序结果。考虑到特定用户的主观性和不确定性, 通过检索算法从 GIAA 数据集中提取与用户指定图像相似的数据集来描述其视觉审美偏好, 并基于该用户特定数据集中的图像美学属性分布, 通过用户多次的交互排序得到符合此用户的 PIAA 模型。从上述分析可以看出, 基于用户交互的 PIAA 模型可以通过用户在 PIAA 模型的训练过程中对个性化美学排序结果不断地进行调整, 从而获取符合用户个性化审美体验的 PIAA 模型。

2.2.3 基于审美差异的 PIAA 模型

除了上述两类 PIAA 模型,一些研究人员认为用户的个性化审美与图像大众化审美之间存在的差异会受到用户主观因素和图像客观因素影响,基于此动机,现有的研究工作提出了基于审美差异的 PIAA 模型,其典型框架如图 9 所示。

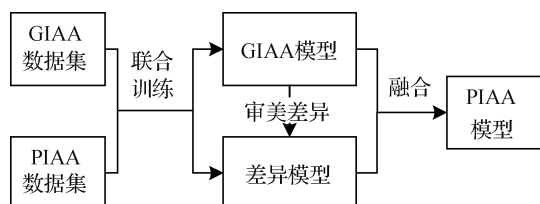


图 9 基于审美差异的 PIAA 模型典型框架

Fig.9 The typical framework of the PIAA model
based on aesthetic difference

Ren 等人(2017)针对用户的个性化审美偏好问题提出了两个用于个性化图像美学评价的数据集,并通过研究发现用户的个性化审美与图像大众化审美之间的差异与图像内容和美学属性之间存在着密切的关联性。利用用户的个性化美学评分与图像大众化美学分数的差异值作为预测目标,通过对基于 CNN 的 GIAA 模型进行微调训练来获取 PIAA 模型。Wang 等人(2019)提出一种包含大众化网络和辅助网络的元学习框架,首先利用大众化网络学习 GIAA 模型,并利用辅助网络学习图像的内容的特征,然后利用元学习的快速适应能力通过辅助特征学习特定用户的个性化美学评分与图像大众化美学分数之间的差异分数,最终通过对差异分数与大众化美学评分进行求和得到 PIAA 模型。Li 等人(2020)将用户的性格特征作为该用户个性化审美与大众化审美之间差异的主观因素,提出一种基于多任务学习的个性化图像美学评价框架。该框架包括两个阶段,第 1 阶段,提出一个具有共享权重的多任务学习网络,用于同时预测图像的美学分布和偏好该图像的用户性格特征。为了获取 GIAA 模型和用户性格特征的共同表征,构建一个孪生网络,将美学数据和性格数据用于联合训练多任务学习网络模块。第 2 阶段,基于多任务模块中预测得到用户的性格特征和大众化审美分数,进一步引入任务间融合学习模块,最终生成针对特定用户的 PIAA 模型。Zhu 等人(2022)发现基于平均美学的 GIAA 先验模型难以体现不同用户对图像审美差异的多样性,为

了学习不同用户对图像进行美学评价时更可靠的美学评价先验知识,提出一种基于双层梯度优化元学习的 PIAA 模型,直接通过对大量用户的 PIAA 任务进行训练构建美学评价先验知识模型,然后通过目标用户少量的 PIAA 数据集对先验知识模型进行微调,最终获得符合该用户个性化审美体验的 PIAA 模型。

从上述分析可以看出,基于审美差异的 PIAA 模型可以通过分析造成不同用户之间审美差异的影响因素,并利用目标用户的偏好图像提取相应特征,从已训练的 GIAA 模型进行迁移学习,得到该用户的 PIAA 模型。

2.3 3 类 PIAA 模型的优缺点

表 1 总结了现有 3 类 PIAA 模型的主要特性以及优缺点。可以看出,现有基于协同过滤的 PIAA 模型可以利用不同用户对图像的审美偏好建立协同矩阵,在对目标用户进行 PIAA 建模时,通过协同矩阵中相似用户的审美偏好得到该用户对图像的个性化审美。但是这在实际情况通常难以实现,因为用户之间的审美是相对稀疏的,无法确保目标用户与其他用户对同一幅图像都进行过美学评价。现有基于用户交互的 PIAA 模型可以根据用户的实时反馈获取较为鲁棒的性能,但是该方法需要用户进行在线的交互行为,这通常难以适用于在现实生活中的离线场景应用。现有基于审美差异的 PIAA 模型通过先验模型中学习不同用户之间的审美差异,可以利用目标用户的少量标注样本进行微调得到,但是该方法难以有效地量化用户个性化审美与大众化审美之间差异的影响因素,这是因为用户的个性化审美偏好会受到其自身的多重主观特性影响,导致获取的 PIAA 模型存在一定的不确定性。

综上所述,基于协同过滤和用户交互的 PIAA 模型局限于需要用户进行在线交互的应用场景;基于审美差异的 PIAA 模型依赖于可以有效地衡量用户个性化审美相对于图像大众化审美的差异的主观性因素。因而,上述不同类型的 PIAA 模型需要在不同的前提条件下获取用户的主观特性,无法在统一的实验设置下对现有 PIAA 模型的性能进行定量评估。面对上述问题,如何通过对用户在日常生活中外化出的审美行为进行分析,有效捕获影响用户对图像审美体验的主观性因素,并借助基于小样本学习的深度框架将相应的主观性因素融入到 PIAA

表1 现有的3类PIAA模型的主要特性和优缺点

Table 1 The characteristics, advantages, and shortcomings of three kinds of the existing PIAA models

PIAA 模型	分类	主要特性	优点	缺点
Deng 等人(2017a)	协同过滤	用户偏好图像之间相似性度量	协同矩阵有效反映不同用户之间的审美偏好	依赖相似用户之间共同的审美偏好
Wang 等人(2018)	协同过滤	用户与图像文本之间相关性编码	可以有效获取用户对图像的注意力特性	依赖相似用户之间共同的审美偏好和评论信息
Cui 等人(2020)	协同过滤	用户偏好图像之间相似性度量	协同矩阵有效反映不同用户之间的审美偏好	依赖相似用户之间共同的审美偏好
Yeh 等人(2014)	用户交互	基于用户反馈的美学排序与选取	实时反馈,鲁棒性较强	需要用户进行在线反馈
Park 等人(2017)	用户交互	基于最近邻搜索的美学排序	排序模型的鲁棒性较强	需要用户进行交互行为
Lyu 等人(2018)	用户交互	基于用户交互的美学再排序	排序模型的鲁棒性较强	需要用户实时地进行美学再排序
Ren 等人(2017)	审美差异	基于美学属性的差异分数建模	美学差异模型参数量少,较容易训练	美学属性难以较好地表征用户的个性化审美偏好
Wang 等人(2019)	审美差异	基于元学习的差异分数建模	基于元学习的差异模型的迁移能力较强	内容特征难以较好地表征用户个性化审美偏好
Li 等人(2020)	审美差异	融入用户的性格特征	可以较好地描述不同用户之间的审美差异	大众化先验知识模型难以有效捕获用户个性化审美
Zhu 等人(2022)	审美差异	基于元学习的共享先验知识建模	可以捕获到不同用户之间的审美差异	对用户审美的主观性因素描述不足

模型的学习中,是提高 PIAA 模型有效性和鲁棒性的关键。

3 评价指标和数据库

3.1 评价指标

对于特定用户的 PIAA 任务来说,图像美学分数的相对大小相比于绝对大小更具研究意义,所以图像美学评价模型的预测结果与真实结果之间的排序一致性是非常重要的评价标准(Ren 等,2017;Lyu 等,2018;Li 等,2020)。PIAA 模型的性能通常采用斯皮尔曼相关系数(Spearman rank order correlation coefficient, SROCC)进行衡量。假设 s_i 和 s'_i 分别表示第 i 幅测试图像在真实美学分数和预测美学分数中的排序等级,则真实美学分数和预测美学分数之间的等级差为

$$d_i = s_i - s'_i \quad (1)$$

SROCC 计算为

$$f_{\text{SROCC}} = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \quad (2)$$

式中, N 表示测试图像的数量, SROCC 的取值范围在 $-1 \sim 1$ 之间。SROCC 的值越高表示 PIAA 模型的性能越好。

3.2 PIAA 数据库

现有不少针对 IAA 的研究建立的数据库,但大多是针对 GIAA 任务构建的,如 AVA(aesthetic visual analysis)(Murray 等,2012)、AADB(aesthetics and attributes database)(Kong 等,2016)、CUHK-PQ(Chinese University of Hong Kong photo quality)(Tang 等,2013)、DPChallenge(digital photography challenge)(Datta 等,2008)和 Photo.net(Joshi 等,2011)等,这使得 GIAA 模型的研究相对成熟且发展较快。随着社交网络和在线众包技术的快速发展,一些研究人员针对特定用户的个性化审美体验建立了公开的 PIAA 数据库,如 FLICKR-AES(Flickr images with aesthetics annotation dataset)和 REAL-CUR(real album curation dataset)(Ren 等,2017)。AADB 是针对 GIAA 任务建立的数据库,但是由于 AADB 数据库包含每个用户对图像进行美学评分的身份信息,因此该数据库也可以用于 PIAA 任务。本文主要是对 PIAA 模型的研究综述,因此主要对 AADB、

FLICKR-AES 和 REAL-CUR 这 3 个公开数据库进行详细介绍,主要信息如表 2 所示。

表 2 3 个公开数据库的主要信息
Table 2 The main information of three public databases

数据库	特点	图像数量/幅	用户数量	分数范围
AADB	美学属性	10 000	190	[1, 5]
FLICKR-AES	在线评分	40 000	210	[1, 5]
REAL-CUR	个人相册	2 872	14	[1, 5]

2016 年美国加州大学欧文分校的 Kong 等人 (2016) 建立了一个基于美学属性的 AADB 数据库。该数据库包含 10 000 幅图像,其中,每一幅图像都收到至少 5 位来自亚马逊土耳其机器人 (Amazon mechanical Turk) (<http://www.mturk.com>) 在线用户的美学评分和 11 种美学属性,共 190 位用户对图像进行美学标注,并且数据库提供了每个用户的身份 (identity, ID) 信息。美学分数范围为 1~5 之间,用户根据 11 种美学属性对图像给出正负两类标签。正标签表示图像具有某种美学属性,负标签表示图像不具有某种美学属性。11 种美学属性分别为有趣内容、强调对象、光照良好、色彩和谐、色彩生动、景深程度、运动模糊、三分构图、平衡元素、重复程度和对称程度。该数据库的训练集包含 8 500 幅图像,验证集包含 500 幅图像,测试集包含 1 000 幅图像。AADB 数据库提供了每位用户的身份信息,可用于 PIAA 模型的性能验证。

美国罗格斯大学建立了两个 PIAA 数据库:

FLICKR-AES 和 REAL-CUR (Ren 等,2017)。其中,FLICKR-AES 数据库收集了来自图像分享网站 Flickr 上的 40 000 幅图像,并利用 Amazon mechanical Turk 在线收集了 210 位用户对图像的美学评分,每一幅图像收到 5 位不同用户的美学评分。图 10 展示了 FLICKR-AES 数据库中一幅示例图像和进行打分的 5 位用户的身份 (ID) 信息和美学评分。其中美学分数范围为 1~5 分,分数越高表示图像美学质量越高。可以看出,不同用户之间的美学评分差异较大。该数据库提供了用户的身份信息,可以根据用户 ID 获取同一位用户对不同图像的美学评分结果,研究针对该用户的 PIAA 模型。在 FLICKR-AES 数据库中,将 37 位用户进行美学评分的 4 737 幅图像组成测试集,每位用户进行美学评价的图像数量范围在 105~171 之间 (平均值为 137)。其余由 173 位用户标注的 35 263 幅图像作为训练集。这样划分可以保证训练集和测试集中的用户进行标注的图像之间互不重叠,符合在实际场景应用中用户只会对自己的照片进行美学评分情况。因此,FLICKR-AES 数据库成为近期个性化图像美学评价方法 (Ren 等,2017;Lyu 等,2018;Li 等,2020;Zhu 等,2022) 进行性能验证的重要数据库。

为了验证在真实场景下用户对相册中照片的美学排序和评分,Ren 等人 (2017) 提出了 REAL-CUR 数据库。该数据库由 14 位不同用户的个人相册以及相应用户提供的美学评分,每个相册包含的图像数量从 197~222 幅。由于个人相册中的每幅图像只会收到相册所拥有用户的美学评分,该数据库



(a) 示例图像

用户ID: A2B6WQG0A9CKXQ	美学评分: 2
用户ID: A2I1 8CP0G6WMYB	美学评分: 3
用户ID: A1RILUP7EBYOQ1	美学评分: 3
用户ID: A24GJXAH1I09ZU	美学评分: 5
用户ID: AI39WUF9N0XJ9	美学评分: 2

(b) 用户身份信息

图 10 FLICKR-AES 数据库中的示例图像以及对应用户的身份 (ID) 信息和美学评分

Fig. 10 Example image with the corresponding users' identity (ID) information and aesthetic score in FLICKR-AES database ((a) example image; (b) users' identity)

收集的图像美学评分是每位用户对个人相册进行审美评估的结果。因此,REAL-CUR数据库是使用真实用户对自己相册中的图像进行美学评价的数据集。

4 PIAA 模型的应用

PIAA 的研究旨在利用可计算的智能系统构建模型来模拟特定用户的审美思维对图像的美感进行评估。由于人们对图像的审美体验具有较强的主观性,往往涉及心理学、艺术美学和计算机图形学等多个交叉学科的融合。因此,PIAA 模型的研究面临全新的问题与挑战。同样,PIAA 模型的研究在人机交互、个性化搜索引擎与推荐系统(冀振燕等,2017)、个性化视觉增强(Kim等,2020a)和个性化广告海报设计(Song等,2019)等经济、艺术与科技的融合领域有着重要的应用价值。

随着计算机技术和硬件制造技术的不断进步,可计算的智能设备逐渐开始面向个体用户并朝着小型化和个性化发展,如可穿戴设备、移动手机和平板电脑等。用户利用这些移动智能设备可以随时随地进行购物和娱乐等消费行为,而用户之间的消费行为和审美体验往往存在较大差异,因此,需要根据不同的用户偏好设计个性化的私人定制服务。例如,阿里巴巴的鲁班系统可以根据用户个性化的审美体验自动设计出千人千面的商品广告,提升了投放广告的点击率;OPPO R17 手机的立体自定义美颜功能可以根据用户不同的审美偏好打造个性化的成像效果,深受用户欢迎。这些案例说明当今社会人们希望智能系统能够理解他们独特的审美偏好,并尽可能满足他们的审美需求。此外,百度公司于2020年9月—2021年1月举办的AI(artificial intelligence)沉浸式互动艺术体验展体现了人们对艺术作品的追求。但是由于人们对艺术作品的感受具有高度主观性,因而在利用AI技术创造艺术作品时需要考虑不同人的个性化审美体验。并且,通过研究用户的个性化审美可以分析用户潜在的行为倾向和心理状态,对用户的情感分析(Zhao等,2018)、性格分析(Zhu等,2018)以及心理疾病辅助诊断(Guntuku等,2017)具有重要的应用价值。总而言之,PIAA 模型的研究在当今科技、经济和医疗和艺术等领域的融合与发展上具有广阔的应用场景。

但是,现有PIAA模型在实际应用中还存在较大问题,主要原因有:1)用户的审美体验在实际场景存在较大的不确定性,导致难以有效地量化用户在审美感知中的主观特性;2)现有技术对PIAA模型的研究尚处于起步阶段,相关的数据库还较为匮乏,难以直接应用于实际场景。因此,在实际应用中,需要借助于更有效的模型来提取用户更为鲁棒的审美主观特性,并创建与应用场景相关的数据库来学习更加有效的PIAA模型。

5 结 语

移动互联网的快速发展使用户对个性化产品和服务产生了极大需求。针对图像美学的个性化评价符合时代的发展要求。本文首先对图像美学评价的研究现状和发展趋势进行分析,然后总结当前个性化图像美学评价方法的整体研究框架和面临的问题与挑战。PIAA模型的研究主要面临训练样本有限和美学评估的高度主观性两个主要挑战。PIAA模型主要可分为基于协同过滤的PIAA模型、基于用户交互的PIAA模型和基于审美差异的PIAA模型,虽然取得了一定进展,但是还存在不少问题。基于协同过滤的PIAA模型和基于用户交互的PIAA模型需要借助于其他相似用户或用户的交互行为,这不利于现实情况下的离线系统;而基于审美差异的PIAA模型依赖于对用户的主观特性的有效客观量化,导致PIAA模型存在不确定性。目前对PIAA模型的研究还处于起步阶段,如何利用有限的样本进行模型训练和提取描述用户主观特性的有效特征是问题的关键。

针对PIAA模型的研究,小样本学习和主观性较强仍然是两个主要问题。对此,总结未来的研究重点和发展方向如下:

1)研究知识驱动的PIAA模型。现有的PIAA方法大都需要数据驱动的机器学习或深度学习方法进行模型构建(Ren等,2017;Li等,2020),这些方法的良好表现很大程度上依赖于高质量的标注数据。由于用户对图像的审美标注具有高度主观特性,所以高质量的标注数据需要极为精准的主观性实验。在现实应用中,由于人工标注成本较高,耗费大量人力物力,针对PIAA模型的小样本学习问题,如何在少量训练样本情况下获取有效的先验知识模

型具有重要研究价值。并且,随着近些年自监督学习(Jing 和 Tian,2021)、强化学习(Li 等,2018)和元学习(Wortsman 等,2019)等知识驱动的深度学习技术在计算机视觉领域广泛应用,基于知识驱动的PIAA模型是下一步研究的重要方向。

2)研究融入用户主观特性的PIAA模型。在用户对图像的审美体验中,不同用户之间存在审美差异的关键因素是用户之间具有不同的主观特性。因此,在研究用户的个性化审美中,不仅要借助于用户对图像美学评价的标注数据,而且还需要借助于用户的自身主观特性。针对上述问题,可以通过对用户外化出的行为进行分析,获取可以全面描述影响其对图像审美偏好的多重主观特性,例如情绪情感(Zhao 等,2018)和性格特征(Zhu 等,2018)等。同时,利用基于多模态融合理论的深度学习技术融入用户的多重主观特性,构建出鲁棒性较好的PIAA模型。

3)构建具有统一的用户主观特性和个性化图像美学评价的数据库。从本文的分析中可以看出,研究用户对图像的个性化美学感知需要借助用户的主观特性,但是现有工作很少将这两项研究内容进行结合分析。究其原因,是由于目前尚无相关的数据库可以同时标注用户的主观特性和其对图像的个性化美学评分(Ren 等,2017)。因此,在未来的研究工作中,亟需开发出具有统一的用户主观特性和个性化图像美学评价数据库,通过该数据库可以更加深入地研究用户对图像审美感知的个性化评价问题。

参考文献 (References)

- Bai R Y, Guo X Y, Jia C H and Geng H J. 2019. Overview of research methods of painting aesthetics. *Journal of Image and Graphics*, 24(11): 1860-1881 (白茹意, 郭小英, 贾春花, 耿海军. 2019. 绘画图像美学研究方法综述. *中国图象图形学报*, 24(11): 1860-1881) [DOI: 10.11834/jig.190118]
- Bianco S, Cusano C, Piccoli F and Schettini R. 2020. Personalized image enhancement using neural spline color transforms. *IEEE Transactions on Image Processing*, 29: 6223-6236 [DOI: 10.1109/TIP.2020.2989584]
- Cavanagh P. 2005. The artist as neuroscientist. *Nature*, 434(7031): 301-307 [DOI: 10.1038/434301a]
- Ceroni A. 2018. Personal photo management and preservation//Mezaris V, Niederée C and Logie R H, eds. *Personal Multimedia Preservation: Remembering or Forgetting Images and Video*. Cham: Springer: 279-314 [DOI: 10.1007/978-3-319-73465-1_8]
- Chaudhary P, Shaw K and Mallick P K. 2018. A survey on image enhancement techniques using aesthetic community//Proceedings of International Conference on Intelligent Computing and Applications. Singapore, Singapore: Springer: 585-596 [DOI: 10.1007/978-981-10-5520-1_53]
- Cui C R, Liu H H, Lian T, Nie L Q, Zhu L and Yin Y L. 2019. Distribution-oriented aesthetics assessment with semantic-aware hybrid network. *IEEE Transactions on Multimedia*, 21(5): 1209-1220 [DOI: 10.1109/TMM.2018.2875357]
- Cui C R, Yang W Y, Shi C, Wang M, Nie X S and Yin Y L. 2020. Personalized image quality assessment with social-sensed aesthetic preference. *Information Sciences*, 512: 780-794 [DOI: 10.1016/j.ins.2019.10.011]
- Datta R, Li J and Wang J Z. 2008. Algorithmic inferencing of aesthetics and emotion in natural images: an exposition//Proceedings of the 15th IEEE International Conference on Image Processing. San Diego, USA: IEEE: 105-108 [DOI: 10.1109/ICIP.2008.4711702]
- Deng X, Cui C R, Fang H D, Nie X S and Yin Y L. 2017a. Personalized image aesthetics assessment//Proceedings of 2017 ACM on Conference on Information and Knowledge Management. Singapore, Singapore: ACM: 2043-2046 [DOI: 10.1145/3132847.3133052]
- Deng Y B, Loy C C and Tang X O. 2017b. Image aesthetic assessment: an experimental survey. *IEEE Signal Processing Magazine*, 34(4): 80-106 [DOI: 10.1109/MSP.2017.2696576]
- Fang H D, Cui C R, Deng X, Nie X S, Jian M W and Yin Y L. 2018. Image aesthetic distribution prediction with fully convolutional network//Proceedings of the 24th International Conference on Multimedia Modeling. Bangkok, Thailand: Springer: 267-278 [DOI: 10.1007/978-3-319-73603-7_22]
- Guntuku S C, Yaden D B, Kern M L, Ungar L H and Eichstaedt J C. 2017. Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*, 18: 43-49 [DOI: 10.1016/j.cobeha.2017.07.005]
- He X N, Liao L Z, Zhang H W, Nie L Q, Hu X and Chua T S. 2017. Neural collaborative filtering//Proceedings of the 26th International Conference on World Wide Web. Perth, Australia: ACM: 173-182 [DOI: 10.1145/3038912.3052569]
- Hosu V, Saupe D, Goldluecke B, Lin W S, Cheng W H, See J and Wong L K. 2020. From technical to aesthetics quality assessment and beyond: challenges and potential//Proceedings of 2020 Joint Workshop on Aesthetic and Technical Quality Assessment of Multimedia and Media Analytics for Societal Trends. Seattle, USA: ACM: 19-20 [DOI: 10.1145/3423268.3423589]
- Hou J W, Yang S and Lin W S. 2020. Object-level attention for aesthetic rating distribution prediction//Proceedings of the 28th ACM International Conference on Multimedia. Seattle, USA: ACM: 816-824 [DOI: 10.1145/3394171.3413695]

- Huang L W, Jiang B T, Lyu S Y, Liu Y B and Li D Y. 2018. Survey on deep learning based recommender systems. *Chinese Journal of Computers*, 41(7): 1619-1647 (黄立威, 江碧涛, 吕守业, 刘艳博, 李德毅. 2018. 基于深度学习的推荐系统研究综述. *计算机学报*, 41(7): 1619-1647) [DOI: 10.11897/SP. J. 1016. 2018. 01619]
- Jacobsen T. 2006. Bridging the arts and sciences; a framework for the psychology of aesthetics. *Leonardo*, 39(2): 155-162 [DOI: 10.1162/leon.2006.39.2.155]
- Ji Z Y, Yao W N and Pi H Y. 2017. Research on personalized image retrieval and recommendation. *Journal of Beijing University of Posts and Telecommunications*, 40(3): 19-30 (冀振燕, 姚伟娜, 皮怀雨. 2017. 个性化图像检索和推荐的研究. *北京邮电大学学报*, 40(3): 19-30) [DOI: 10.13190/j.jbupt.2017.03.002]
- Jin X, Wu L, Li X D, Chen S Y, Peng S W, Chi J Y, Ge S M, Song C G and Zhao G. 2018. Predicting aesthetic score distribution through cumulative jensen-shannon divergence//Proceedings of the 32nd AAAI Conference on Artificial Intelligence and the 30th Innovative Applications of Artificial Intelligence Conference and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence. New Orleans, USA: AAAI Press: 10
- Jin X, Zhou B, Zou D Q, Li X D, Sun H B and Wu L. 2018. Image aesthetic quality assessment; a survey. *Science and Technology Review*, 36(9): 36-45 (金鑫, 周彬, 邹冬青, 李晓东, 孙红波, 吴乐. 2018. 图像美学质量评价技术发展趋势. *科技导报*, 36(9): 36-45) [DOI: 10.3981/j.issn.1000-7857.2018.09.005]
- Jing L L and Tian Y L. 2021. Self-supervised visual feature learning with deep neural networks; a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(11): 4037-4058 [DOI: 10.1109/TPAMI.2020.2992393]
- Joshi D, Datta R, Fedorovskaya E, Luong Q T, Wang J Z, Li J and Luo J B. 2011. Aesthetics and emotions in images. *IEEE Signal Processing Magazine*, 28(5): 94-115 [DOI: 10.1109/MSP.2011.941851]
- Ke S C, Zhao Y W, Li B C and Peng T Q. 2017. Image retrieval based on convolutional neural network and kernel-based supervised hashing. *Acta Electronica Sinica*, 45(1): 157-163 (柯圣财, 赵永威, 李弼程, 彭天强. 2017. 基于卷积神经网络和监督核哈希的图像检索方法. *电子学报*, 45(1): 157-163) [DOI: 10.3969/j.issn.0372-2112.2017.01.022]
- Kim H U, Koh Y J and Kim C S. 2020a. PieNet; personalized image enhancement network//Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK: Springer: 374-390 [DOI: 10.1007/978-3-030-58577-8_23]
- Kim W H, Choi J H and Lee J S. 2020b. Objectivity and subjectivity in aesthetic quality assessment of digital photographs. *IEEE Transactions on Affective Computing*, 11(3): 493-506 [DOI: 10.1109/TAFFC.2018.2809752]
- Koenderink J J. 2010. *Color for the Sciences*. Cambridge: MIT Press
- Kong S, Shen X H, Lin Z, Mech R and Fowlkes C. 2016. Photo aesthetics ranking network with attributes and content adaptation//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands: Springer: 662-679 [DOI: 10.1007/978-3-319-46448-0_40]
- 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: Curran Associates Inc: 1097-1105
- Kuzovkin D. 2019. *Assessment of Photos in Albums Based on Aesthetics and Context*. Rennes: University of Rennes
- Leder H, Belke B, Oeberst A and Augustin D. 2004. A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95(4): 489-508 [DOI: 10.1348/0007126042369811]
- Li D B, Wu H K, Zhang J G and Huang K Q. 2018. A2-RL: aesthetics aware reinforcement learning for image cropping//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 8193-8201 [DOI: 10.1109/CVPR.2018.00855]
- Li L D, Zhu H C, Zhao S C, Ding G G and Lin W S. 2020. Personality-assisted multi-task learning for generic and personalized image aesthetics assessment. *IEEE Transactions on Image Processing*, 29: 3898-3910 [DOI: 10.1109/TIP.2020.2968285]
- Lu Y, Guo C, Lin Y L, Zhuo F and Wang F Y. 2020. Computational aesthetics of fine art paintings; the state of the art and outlook. *Acta Automatica Sinica*, 46(11): 2239-2259 (鲁越, 郭超, 林懿伦, 卓凡, 王飞跃. 2020. 绘画艺术图像的计算美学: 研究前沿与展望. *自动化学报*, 46(11): 2239-2259) [DOI: 10.16383/j.aas.c200358]
- Luo L J. 2014. *Emotional Aesthetics — On the Beauty and Aesthetics*. Beijing: China Economic Publishing House (罗利建. 2014. 情感美学: 论美和美感. 北京: 中国经济出版社)
- Lyu P, Wang M, Xu Y B, Peng Z, Sun J Y, Su S M, Zhou B and Xu M L. 2018. USAR; an interactive user-specific aesthetic ranking framework for images//Proceedings of the 26th ACM International Conference on Multimedia. Seoul, Korea (South): ACM: 1328-1336 [DOI: 10.1145/3240508.3240635]
- Martindale C. 2007. Recent trends in the psychological study of aesthetics, creativity, and the arts. *Empirical Studies of the Arts*, 25(2): 121-141 [DOI: 10.2190/B637-1041-2635-16 NN]
- Murray N, Marchesotti L and Perronnin F. 2012. AVA: a large-scale database for aesthetic visual analysis//Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE: 2408-2415 [DOI: 10.1109/CVPR.2012.6247954]
- Nanay B, Smith M, Irvin S and Schellekens E. 2019. Is psychology relevant to aesthetics? A symposium. *Estetika: The Central European Journal of Aesthetics*, 56(1): 87-138 [DOI: 10.33134/eeja.185]

- O'Donovan P, Agarwala A and Hertzmann A. 2014. Collaborative filtering of color aesthetics//Proceedings of the Workshop on Computational Aesthetics. Vancouver, Canada: ACM: 33-40 [DOI: 10.1145/2630099.2630100]
- Palmer S E, Schloss K B and Sammartino J. 2013. Visual aesthetics and human preference. *Annual Review of Psychology*, 64: 77-107 [DOI: 10.1146/annurev-psych-120710-100504]
- Park K, Hong S, Baek M and Han B. 2017. Personalized image aesthetic quality assessment by joint regression and ranking//Proceedings of 2017 IEEE Winter Conference on Applications of Computer Vision. Santa Rosa, USA: IEEE: 1206-1214 [DOI: 10.1109/WACV.2017.139]
- Ren J, Shen X H, Lin Z, Mech R and Foran D J. 2017. Personalized image aesthetics//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 638-647 [DOI: 10.1109/ICCV.2017.76]
- Song X M, Han X J, Li Y K, Chen J Y, Xu X S and Nie L Q. 2019. GP-BPR: personalized compatibility modeling for clothing matching//Proceedings of the 27th ACM International Conference on Multimedia. Nice, France: ACM: 320-328 [DOI: 10.1145/3343031.3350956]
- Swami V and Furnham A. 2014. *Personality and Aesthetics Preferences*. Cambridge: Cambridge University Press
- Talebi H and Milanfar P. 2018. NIMA: neural image assessment. *IEEE Transactions on Image Processing*, 27(8): 3998-4011 [DOI: 10.1109/TIP.2018.2831899]
- Tang X O, Luo W and Wang X G. 2013. Content-based photo quality assessment. *IEEE Transactions on Multimedia*, 15(8): 1930-1943 [DOI: 10.1109/TMM.2013.2269899]
- Vo N, Jiang L, Sun C, Murphy K, Li L J, Li F F and Hays J. 2019. Composing text and image for image retrieval-an empirical odyssey//Proceedings of 2019 IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 6439-6448 [DOI: 10.1109/CVPR.2019.00660]
- Wang G L, Yan J C and Qin Z. 2018. Collaborative and attentive learning for personalized image aesthetic assessment//Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm, Sweden: ACM: 957-963
- Wang H, Zhang Y, Shen H H and Zhang J Z. 2017. Review of image enhancement algorithms. *Chinese Optics*, 10(4): 438-448 (王浩, 张叶, 沈宏海, 张景忠. 2017. 图像增强算法综述. *中国光学*, 10(4): 438-448) [DOI: 10.3788/CO.20171004.0438]
- Wang W N, Ma X D, Su J J, Luo J B and Xu X M. 2020. User image aesthetic representation based on personalized generation template aggregation. *Journal of South China University of Technology (Natural Science Edition)*, 48(3): 1-9 (王伟凝, 马学东, 苏俊杰, 罗杰波, 徐向民. 2020. 基于个性化生成模板集聚的用户图像美感表征. *华南理工大学学报(自然科学版)*, 48(3): 1-9) [DOI: 10.12141/j.issn.1000-565X.190442]
- Wang W N, Su J J, Li L M, Xu X M and Luo J B. 2019. Meta-learning perspective for personalized image aesthetics assessment//Proceedings of 2019 IEEE International Conference on Image Processing. Taipei, China: IEEE: 1875-1879 [DOI: 10.1109/ICIP.2019.8803119]
- Wang W N, Wang L, Zhao M Q, Cai C J, Shi T T and Xu X M. 2016. Image aesthetic classification using parallel deep convolutional neural networks. *Acta Automatica Sinica*, 42(6): 904-914 (王伟凝, 王励, 赵明权, 蔡成加, 师婷婷, 徐向民. 2016. 基于并行深度卷积神经网络的图像美感分类. *自动化学报*, 42(6): 904-914) [DOI: 10.16383/j.aas.2016.c150718]
- Wang W N, Yi J J and He Q H. 2012. Review for computational image aesthetics. *Journal of Image and Graphics*, 17(8): 893-901 (王伟凝, 蚁静绒, 贺前华. 2012. 可计算图像美学研究进展. *中国图象图形学报*, 17(8): 893-901) [DOI: 10.11834/jig.20120801]
- Wang W N, Yi J J, Xu X M and Wang L. 2014. Computational aesthetics of image classification and evaluation. *Journal of Computer-Aided Design and Computer Graphics*, 26(7): 1075-1083 (王伟凝, 蚁静绒, 徐向民, 王励. 2014. 可计算的图像美学分类与评估. *计算机辅助设计与图形学学报*, 26(7): 1075-1083)
- Wang Z M. 2015. Review of no-reference image quality assessment. *Acta Automatica Sinica*, 41(6): 1062-1079 (王志明. 2015. 无参考图像质量评价综述. *自动化学报*, 41(6): 1062-1079) [DOI: 10.16383/j.aas.2015.c140404]
- Wells L. 2015. *Photography: A Critical Introduction*. London: Routledge
- Wortsman M, Ehsani K, Rastegari M, Farhadi A and Mottaghi R. 2019. Learning to learn how to learn: self-adaptive visual navigation using meta-learning//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 6743-6752 [DOI: 10.1109/CVPR.2019.00691]
- Xu M N, Zhong J X, Ren Y R, Liu S and Li G. 2020. Context-aware attention network for predicting image aesthetic subjectivity//Proceedings of the 28th ACM International Conference on Multimedia. Seattle, USA: ACM: 798-806 [DOI: 10.1145/3394171.3413834]
- Yeh C H, Barsky B A and Ouhyoung M. 2014. Personalized photograph ranking and selection system considering positive and negative user feedback. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 10(4): #36 [DOI: 10.1145/2584105]
- Zeki S. 2013. Clive Bell's "Significant Form" and the neurobiology of aesthetics. *Frontiers in Human Neuroscience*, 7: #730 [DOI: 10.3389/fnhum.2013.00730]
- Zeng H, Cao Z S, Zhang L and Bovik A C. 2020. A unified probabilistic formulation of image aesthetic assessment. *IEEE Transactions on Image Processing*, 29: 1548-1561 [DOI: 10.1109/TIP.2019.2941778]
- Zhang J, Yang Y, Tian Q, Zhuo L and Liu X. 2017. Personalized social image recommendation method based on user-image-tag model.

- IEEE Transactions on Multimedia, 19(11): 2439-2449 [DOI: 10.1109/TMM.2017.2701641]
- Zhang J, Yang Y, Zhuo L, Tian Q and Liang X. 2019a. Personalized recommendation of social images by constructing a user interest tree with deep features and tag trees. IEEE Transactions on Multimedia, 21(11): 2762-2775 [DOI: 10.1109/TMM.2019.2912124]
- Zhang X D, Gao X B, Lu W and He L H. 2019b. A gated peripheral-foveal convolutional neural network for unified image aesthetic prediction. IEEE Transactions on Multimedia, 21(11): 2815-2826 [DOI: 10.1109/TMM.2019.2911428]
- Zhang X D, Gao X B, Lu W, He L H and Li J. 2020. Beyond vision: a multimodal recurrent attention convolutional neural network for unified image aesthetic prediction tasks. IEEE Transactions on Multimedia, 23: 611-623 [DOI: 10.1109/TMM.2020.2985526]
- Zhao S C, Yao H X, Gao Y, Ding G G and Chua T S. 2018. Predicting personalized image emotion perceptions in social networks. IEEE Transactions on Affective Computing, 9(4): 526-540 [DOI: 10.1109/TAFFC.2016.2628787]
- Zhu H C, Li L D, Wu J J, Zhao S C, Ding G G and Shi G M. 2022. Personalized image aesthetics assessment via meta-learning with bilevel gradient optimization. IEEE Transactions on Cybernetics, 52(3): 1798-1811 [DOI: 10.1109/TCYB.2020.2984670]
- Zhu H C, Li L D, Zhao S C and Jiang H Y. 2018. Evaluating attributed personality traits from scene perception probability. Pattern Recognition Letters, 116: 121-126 [DOI: 10.1016/j.patrec.2018.09.027]

作者简介

祝汉城,男,博士后,主要研究方向为图像美学评价。

E-mail: zhuhanheng@cumt.edu.cn

周勇,通信作者,男,教授,主要研究方向为计算机视觉和大数据挖掘。E-mail: yzhou@cumt.edu.cn

李雷达,男,教授,主要研究方向为多媒体情感计算、视觉质量评价和多媒体信息安全。E-mail: ldli@xidian.edu.cn

赵佳琦,男,副教授,主要研究方向为计算机视觉和行人重识别。E-mail: jiaqizhao@cumt.edu.cn

杜文亮,男,讲师,主要研究方向为计算机视觉和遥感图像处理。E-mail: wldu@cumt.edu.cn