

# 图像美学质量评价技术发展趋势

金鑫<sup>1</sup>, 周彬<sup>2</sup>, 邹冬青<sup>3</sup>, 李晓东<sup>1</sup>, 孙红波<sup>1</sup>, 吴乐<sup>1</sup>

1. 北京电子科技学院计算机科学与技术系, 北京 100070
2. 北京航空航天大学计算机学院, 北京 100191
3. 北京市商汤科技开发有限公司, 北京 100084

**摘要** 图像质量评价是利用计算机模拟人类视觉系统自动评价图像的失真程度, 而图像美学质量评价是利用计算机模拟人类对美的感知与认知, 自动评价图像的“美感”, 分析图像在构图、颜色、光影、景深、虚实等美学因素影响下所形成的美感刺激, 这是计算美学与计算机视觉、心理学、虚拟现实等领域的学科交叉新方向, 其在摄影摄像、电影电视、电子商务、服装设计、平面设计、文稿设计、工业设计、汽车行业、建筑行业、美容美妆等多个行业具有良好的应用前景。本文回顾了图像美学质量评价的发展历史, 从基于人工设计美学特征的方法、基于美学特征深度学习的方法、图像美学质量评价的新任务、图像美学质量评价数据集构建等方面综述了该领域关键技术的发展情况, 并展望了图像美学质量评价及其关键技术的发展趋势。

**关键词** 图像美学质量评价; 计算美学; 虚拟现实; 计算艺术; 视觉艺术

随着照相机、摄像机、深度摄像机、智能手机快速普及, 图像、视频、图形等可视内容数据与日俱增, 可视内容感知理解已经成为可视计算、计算机视觉、计算摄像学等科学研究领域及其交叉方向国际前沿的研究方向。其中图像美学质量评价(image aesthetic quality assessment, IAQA)是近期可视内容感知理解方向中的研究热点。

传统的图像质量评价(image quality assessment, IQA)旨在利用计算机模拟人类视觉系统自动评价图像的失真程度, 主要针对图像在采集、压缩、处理、传输及显示等过程产生图像质量下降情况, 通常包括成像条件差而引起的失真, 有损压缩引起的失真、噪声, 图像传输过程中受信道衰减影响引起的失真等。虽然目的都是获得与主观评价结果相一致的客观评价价值, 但图

像美学质量评价旨在利用计算机模拟人类对美的感知与认知, 自动评价图像的“美感”, 即图像美学质量的计算机评价, 主要针对拍摄或绘画的图像在构图、颜色、光影、景深、虚实等美学因素影响下形成的美感刺激。审美是人类与生俱来的智能活动和需求, 相对于“教会”计算机识别物体、场景、事件等确定性识别类基本技能, 如何让计算机也能感知“美”, 进而能够自动发现“美”和生成“美”是计算美学与计算机视觉、心理学、虚拟现实等交叉研究方向的新挑战。近期, 人工智能技术在物体识别、人脸识别、行为识别等领域取得了较大的进步, 使得计算机或机器人的识物、辨人、认事等方面的能力与人类水平的差距逐渐缩小, 甚至在某些特定应用场景下超过了人类的智能水平。然而, 计算机或机器人在发现美、加工美、创造美等方面的能力还未

收稿日期: 2018-04-23; 修回日期: 2018-05-03

基金项目: 国家自然科学基金项目(61402021)

作者简介: 金鑫, 讲师, 研究方向为计算美学、虚拟现实、计算机视觉, 电子信箱: jinxin@besti.edu.cn; 周彬(通信作者), 讲师, 研究方向为虚拟现实、计算机图形学、机器人, 电子信箱: zhoubin@buaa.edu.cn

引用格式: 金鑫, 周彬, 邹冬青, 等. 图像美学质量评价技术发展趋势[J]. 科技导报, 2018, 36(9): 36-45; doi: 10.3981/j.issn.1000-7857.2018.09.005

取得长足的进步,与人类的美学智能还有较大差距。

如今,人们对美好生活的向往日益增强,所谓的“美丽经济”已经渗透到人们生产生活的各个方面:摄影摄像、电影电视、电子商务、服装设计、平面设计、文稿设计、工业设计、汽车行业、建筑行业、美容美妆等多个行业中均需要在“美感”方面下功夫才能获得更好的经济效益,人们已经不满足仅仅在功能上的齐全,更需要在视觉上的美观。

目前图像美学质量评价的应用才刚刚走出实验室,还处于起步阶段,已经有一些较为成功的案例。例如阿里巴巴的“鲁班”人工智能设计师能够自动设计商品海报,如图1所示,根据雷锋网报道,该系统于2016年“双十一购物狂欢节”当日上线,当年自动设计了1.7亿幅商品海报,商品的点击率提升了100%。如果全靠设计师人工完成,假设每张图耗时20 min,需要100个设计师连续做300年。2017年该系统已经能够每秒设计8000幅海报,共设计了4亿幅海报。“鲁班”系统具有如此强大性能的奥秘就在于其图像美学质量评价引擎,能够从众多自动设计的海报中自动挑选出美学质量较高的一批。



图1 阿里巴巴“鲁班”人工智能设计师设计的商品海报  
(图片来源:雷锋网)

Fig. 1 A product poster which is automatically designed by the Luban system from Alibaba corporation

其他案例还包括:在线视频网站爱奇艺利用图像美学质量评价自动挑选每一段视频的封面,以提高视频的点击率;美团点评公司利用图像美学质量评价自动选择商家首图,以提高商家的用户点击率;时光相册

公司利用图像美学质量评价自动推荐最佳照片,为用户提供更好的照片管理体验;美图公司利用图像美学质量评价自动对用户的人像照片进行美化等;研究人员利用图像美学质量评价对输入图像进行自动裁减,输出美学质量高的图像,以及利用图像美学质量评价对图像搜索引擎返回的图像进行排序,以给用户提高美学质量的检索结果等。

感知和创作具有美感的艺术作品的经验知识需要通过较长时间的训练,艺术具有抽象、模糊、难传授等特点,而计算具有具体、精确、易重用等特点,如图2所示。在物质生活日益丰富的今天,需要为人们提供创造美的便捷工具。如果使计算机或机器人完成艺术创作中的大量非创造性体力劳动,就能够使得人类更好地集中精力发挥创作才能,从而使得更多的人能够创作以前只有专业艺术家才能创作的艺术作品,这首先需要计算机能够识别“美”。



图2 艺术与计算

Fig. 2 Art and computation

## 1 计算美学与图像美学质量评价

1) 计算美学。艺术中的情感即审美的情感,是一种无功利的具有人类普遍性的情感。计算美学(computational aesthetics)是人工智能、美学、心理学等学科发展到一定程度时出现的新兴研究方向,其目的在于使得计算机具有类似人类发现美、改造美、创造美的美学智能,并且研究人类感知美的内在机理,探索人类美学智能与大自然和谐美丽的奥秘。

2) 计算视觉美学。人类约80%的认知信息来自于视觉感知通道,视觉艺术是人类对可视内容信息的一种艺术性加工处理,计算视觉美学是虚拟现实、计算机视觉、计算摄影学、情感计算等领域的研究热点。其在工业设计、服装设计、产品设计、影像视觉、城市规划、环境保护等多个方面有重要的应用前景。

3) 图像美学质量评价。在计算视觉美学的研究中,使得计算机或机器人能够“发现图像美”并且“理解

图像美”,是当前最活跃的研究主题,也是计算视觉美学研究的第一步。如图3所示,拍摄同样一个场景,通常人们会认为左边图的美感高于右边的图。计算机或机器人是否也能像人一样分辨“美”与“丑”?计算机只有在理解美的基础上,才能智能地美化图像,进而能够创造出美的视觉艺术作品。



图3 图像美感的人类评价

Fig. 3 Human evaluation of image aesthetics

图像美学质量评价的主要难点在于以下3点。

(1) 与物体识别的语义特征相比,人类审美的奥秘还未出现可量化的科学解释,图像美学特征点选取难度较大。

(2) 图像美学质量评价除了具备一定的“客观性”(共识性)之外,还具有很强的“主观性”。

(3) 人类对于图像美感的评价存在多种形式,例如分“美”与“丑”,给出数值评分、语言评价等。

## 2 图像美学质量评价研究现状

由于图像美学评价引起了国内外研究机构及学者的广泛关注,目前关注或者与该研究方向相关的主要国内外高校与研究机构有:微软亚洲研究院、清华大学、北京航空航天大学、北京电子科技学院、中国科学院自动化研究所、北京理工大学、复旦大学、山东财经大学、香港中文大学、中国台湾信息科学研究所、美国加州大学欧文分校、西班牙巴塞罗那自治大学、新加坡南洋理工大学、英国萨里大学、美国罗格斯大学、美国奥多比公司(Adobe)、德国蒂宾根大学等。这些高校和研究机构近期在计算机视觉、多媒体、计算机图形学等领域国际权威期刊《Transactions on Multimedia》《Transactions on Image Processing》,以及国际权威会议美国人工智能促进会国际人工智能会议(AAAI Conference on Artificial Intelligence)、国际计算机视觉会议(International Conference on Computer Vision)、国际计算机视觉

与模式识别会议(Computer Vision and Pattern Recognition)、美国计算机协会国际多媒体会议(ACM International Conference on Multimedia)、美国计算机协会信息检索特别兴趣小组信息检索研究与发展会议(ACM SIGIR Conference on Research and Development in Information Retrieval)、欧洲计算机视觉会议(European Conference on Computer Vision)等发表多篇学术论文,说明该方向是瞄准国际前沿符合科学发展趋势的研究方向。

除此之外,国际计算美学会议(Computational Aesthetics, CAe)从2005年开始,至今已经连续举办了12届;计算机视觉三大顶级国际会议之一——欧洲计算机视觉会议(European Conference on Computer Vision, ECCV)从2012年开始举办面向艺术分析的计算机视觉研讨会(Workshop on Computer Vision for Art Analysis),专门研究用计算机视觉技术进行艺术分析,已经连续举办了3届,并将在2018年的ECCV会议上继续举办第4届;北京电子科技学院可视计算与安全实验室将在2018年度国际人工智能与机器人研讨会(International Symposium on Artificial Intelligence and Robotics, ISAIR 2018)上举办第1届计算美学与视觉艺术分析特别分会(Special Session on Special Session on Computational Aesthetics and Visual Arts Analysis, CAVA),该会议将在2018年11月在南京举办,将讨论计算美学领域的最新进展,同时进一步推进计算美学的发展。

以下将从基于人工设计美学特征的方法、基于美学特征深度学习的方法、图像美学质量评价的新任务、图像美学质量评价数据集构建4方面介绍该领域关键技术的发展情况。

### 2.1 基于人工设计美学特征的图像美感分类与评分

2004年,微软亚洲研究院与清华大学自动化系联合提出了一种方法,能够自动区分专业摄影师拍摄的照片与普通用户拍的照片<sup>[1]</sup>,该项工作被认为是图像美学质量评价最早的一篇论文,研究者收集了16643幅专业摄影师拍摄的照片,以及12897幅普通用户拍摄的照片,使用了21类共846维图像低层特征学习分类模型,从而对测试图像进行美感分类。此外,研究者还收集了379幅图像,每一幅图像请16名测试者(8名男性与8名女性)根据图像美感进行评分(1~5档:差、不好、中等、好、卓越),将每幅图像的平均得分作为其真实的人类评分数据,由此学习回归模型,能够对测试图像进行美感评分,如图4所示。



图4 图像美感计算机评分与人类评分

Fig. 4 Image aesthetics assessed by computers and human beings

随后的大部分图像美学质量评价研究工作,通过设计多种视觉特征,拟合人类对图像美学质量的评价结果,依据视觉美学质量对图像进行分类与数值评价,能够给出大众认可的自动评价结果,通常分为以下3个基本步骤。

1) 依据美学标准,收集图像数据集,根据专家知识将该数据集分为美学质量“高”和“低”2个子集(专业/非专业、艺术/非艺术等,或者给出每幅图像的美学质量评分)。

2) 设计多种图像低层与高层特征,并在图像数据集上提取这些特征。

3) 利用机器学习(K算法近邻分类、支撑向量机、随机森林、回归分析、Boosting、朴素贝叶斯等)根据提取的特征与人工分类结果训练分类器,依据视觉美学质量对输入图像进行分类,或者对提取的特征与人类美学质量评分进行回归分析,实现对输入图像视觉美学质量的数值评价。

基于人工设计美学特征的图像美感分类与评分技术的发展共经历了2个主要的发展阶段:内容无关的图像美感分类与评分、基于内容的图像美感分类与评分。

#### 1) 内容无关的图像美感分类与评分。

研究人员设计多种可计算视觉特征,并且利用这些视觉特征,依据视觉美学质量对照片和绘画进行自动分类与数值评价,在早期的研究工作中,研究人员仅仅提取图像的全局特征<sup>[1-2]</sup>,接着,Datta等<sup>[3]</sup>引入了局部区域特征,随后的一些工作又加入了局部区域之间的对比特征<sup>[4-6]</sup>。

随后研究人员又提出了基于单一美学因素(颜色和谐性)的特征<sup>[7]</sup>、高层描述属性特征(构图属性、场景属性、天气属性)<sup>[8]</sup>、低层通用图像描述符(generic image descriptors)<sup>[9]</sup>等特征。然而上述研究均没有考虑不同内

容的图像具有不同的特征(例如为人像与风景照片设计不同的特征),这样对分类或评分的效果有很大影响,因为不同类型的图像内容具有不同的美学特征<sup>[10-13]</sup>。

#### 2) 基于内容的图像美感分类与评分。

2010年欧洲计算机视觉会议(European Conference on Computer Vision, ECCV 2010)上,金鑫等<sup>[14]</sup>提出了“基于内容的图像美学质量评价”的新研究路线。会议论文匿名评委认为该论文是“利用计算模型学习与表示艺术光影的有意义的一步”。2015年,北京航空航天大学将这项工作进行了扩展和总结,发表于图像处理领域顶级期刊《Transactions on Image Processing》<sup>[15]</sup>。

2011年国际计算机视觉会议(International Conference on Computer Vision, ICCV 2011)上,香港中文大学汤晓鸥<sup>[16]</sup>等发表的论文引用了文献[14],对比2篇论文的内容发现:文献[16]是该研究路线的延续和进一步扩展,将17613幅图像按内容分为7类,并且根据图像内容设计了一系列图像显著区域及其特征的提取方法。

2013年12月,文献[16]研究进一步深入后,成果发表于多媒体领域著名期刊《IEEE Transactions on Multimedia》<sup>[17]</sup>。文献[17]根据7种内容的图像(风景、植物、动物、夜景、人像、静物、建筑),分析并提取了一系列基于内容的图像特征,利用基于内容的图像特征与支撑向量机,进行图像视觉美学质量的分类,能够显著提高图像视觉质量评价的准确度(图5<sup>[16]</sup>)。近期基于内容的图像美学质量评价受到了越来越多的研究者的关注,大多集中于人脸图像和人像图像方面<sup>[18-20]</sup>。

## 2.2 基于美学特征深度学习的图像美感分类与评分

2014年以来,图像美感分类与评分的研究工作全面进入了深度学习时代:自动提取图像的美学特征<sup>[21-34]</sup>,研究人员改造了多种用于图像识别的卷积神经网络,并且将图像风格、图像内容等额外信息显示或隐式地建模

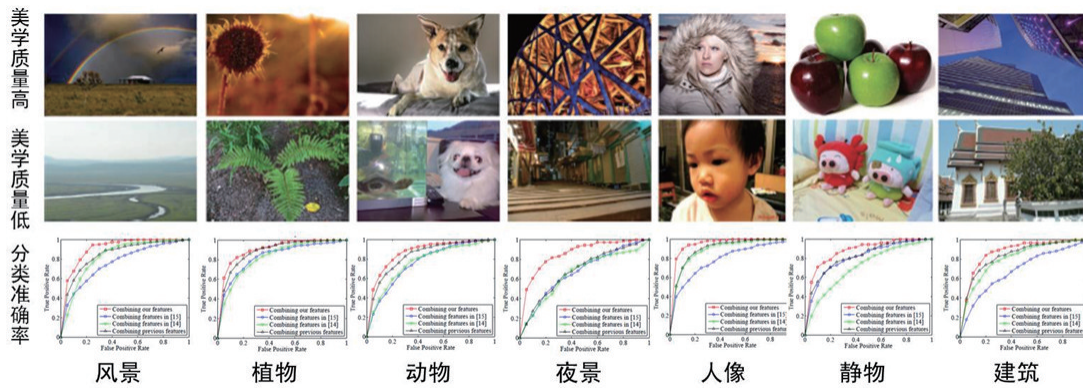


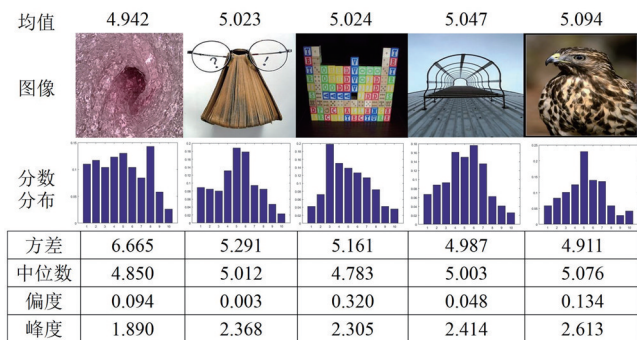
图5 7类不同内容的图像按其美学质量“高”“低”的分类准确率

Fig. 5 Classification accuracies of image aesthetic quality (high or low) on 7 categories

于改造后的卷积神经网络中,用于美感分类与评分,其分类准确率比手工设计的特征相比有较大提高。北京电子科技学院提出的 ILGNet-Inc.V4 目前在公开的 AVA 数据集(a Large-scale database for aesthetic visual analysis)<sup>[35]</sup>上的分类准确率排名国际第一<sup>[28][30]</sup>。文献[34]给出了图像美感分类与评分方面的综述。

### 2.3 图像美学质量评价的新任务:美感分布、美学因素、美学描述等

图像美感分类与评分,本质上是利用1个标量来表示图像的美感,这在很大程度上忽略了人类审美在一定共识上的多样性、主观性、个性化,这也是美学评价与一般图像识别最大的差别,一般的图像识别几乎都有标准答案,共识率很高,产生争议的图像较少。图像美学评分的概率分布能在一定程度上描述审美的主观性,例如方差能够在一定程度上描述人类对一幅图像的共识度,峰度能够在一定程度上描述一幅图像的受欢迎程度等,如图6所示。



以直方图表示图像美感评分的分布,5幅评分接近5分的图像,其评分的分布各异,并且各种统计量有较大差异(AVA数据集)

图6 图像美感分布

Fig. 6 Image aesthetic score distribution

早在2011年,中国科学院自动化研究所的Wu等发表的文献[36]就在图像美感分布的预测上进行了初步探索,随后文献[37]-[39]给出了统计意义的图像美感主观性分析,然而文献[36]的研究工作淹没在大量(200余篇论文)图像美感单标量值客观评价的研究工作中。北京电子科技学院金鑫等<sup>[40]</sup>在2017年9月提出了图像美感分布预测方法,设计了一种新的卷积神经网络:RS-CJS(cumulative Jensen-Shannon divergence based deep convolutional neural networks,基于JS散度的深度卷积神经网络),在深度学习时代重新启动了这个方向,在AVA数据集上取得了目前排名国际第一的美感分布预测性能。

除了美感分布之外,在美学因素方面,2016年美国加州大学欧文分校Kong等<sup>[27]</sup>设计了一个新的图像美学数据集美学因素数据集(aesthetics and attributes database, AADB),包含了8种美学因素(平衡、颜色和谐型、趣味、景深、光照、主体、三分、颜色丰富性)的二值分类评价(在单个美学因素上的“好”与“不好”),然而其美学因素的评价过于简单,不适合用于分析美学评价的主观性与多样性(图7<sup>[27]</sup>)。

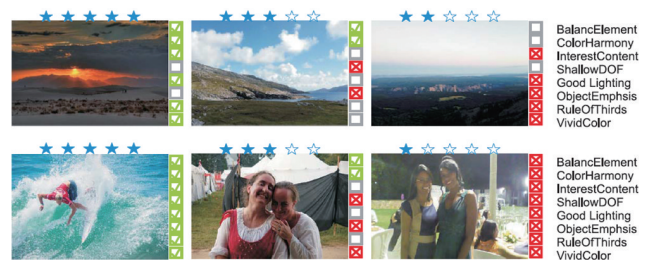


图7 AADB数据集示例,提供了8种美学因素的二值分类评价  
Fig. 7 Samples from AADB dataset, which provide binary classification assessment of 8 aesthetic attributes

2017年,中国台湾信息科学研究所的Chang等<sup>[41]</sup>发布了一个新的图像美学数据集照片语言评论数据集(photo critique captioning dataset, PCCD),首次在图像美学数据集中加入了多美学因素的语言评论信息,并结合卷积神经网络(convolutional neural networks, CNN)与长短记忆单元(long-short term memory, LSTM)进行了图像美学语言评论的预测(图8<sup>[41]</sup>)。该数据集的标注较为全面,包含了1个总体和6个美学因素的评价分数、分布以及多人语言评论,然而其数据量太少(4307幅),难以满足大型深度神经网络对于训练样本的规模需求,并且其数据来源Gurushots.com已经停止发布更多的评价数据。



**Photo Critique Captioning:** racing makes for interesting pictures because of the speed the movement the bright colors

**Image Captioning:** a man riding a motorcycle down a street

第1行为图像美学语言评论(高速运动和明亮的色彩使得整幅图像显得很有趣),第2行为图像内容简述(在街上骑着摩托车的人)

图8 PCCD数据集示例

Fig. 8 A sample from PCCD

2017年,英国萨里大学的Collomosse<sup>[42]</sup>将图像美学质量评价用于图形检索系统。北京理工大学的Wang等<sup>[44]</sup>将图像美学质量评价用于图像自动裁减系统;山东财经大学的Cui等<sup>[43]</sup>将图像美学质量评价用于基于美学的图像检索系统,能够自动裁减输入图像,输出美学质量较高的裁减图像,美国Ren等<sup>[45]</sup>提出了个性化的图像美学评价方法与系统。

2018年,德国蒂宾根大学的Schwarz等<sup>[46]</sup>通过在线图像分享网站Flickr,构建了一个大型图像美学数据集多用户共识大型图像数据集(multi-user agreements and assemble a large dataset, AROD),包含了38万幅图像,标注了图像美学评分的标注数据,其美学评分标注数据是由图像查看次数与点赞次数计算得出,可以用

于图像美学评分估计。2018年,复旦大学的Wang等<sup>[47]</sup>利用AVA数据集构建了AVA-Reviews数据集,包含了AVA数据集中的4万幅图像,每幅图像跟随了6条语言评论,该文献利用CNN与循环神经网络(recurrent neural network, RNN)相结合的神经网络结构同时预测图像的美学分类与语言评论,然而该AVA-Reviews数据集的规模仍然不大,并且语言评论的标注没有考虑多美学因素。

#### 2.4 图像美学质量评价基准数据集构建方法

图像美学质量评价是一个近十几年才开始引起研究人员关注的一个方向,从一开始就没有走基于规则的路线,而是直接遵循了据驱动的路线,因此图像美学质量评价基准数据集构建就成为了该方向研究的关键前提条件。在图像美学质量人工主观评价得分获取方面,目前主要分为实验室内的人工打分实验<sup>[1,6,14,48]</sup>(例如香港中文大学的CUHKPO数据集<sup>[16]</sup>)、在线图像分享打分网站下载收集<sup>[2-3,35]</sup>(例如美国宾夕法尼亚州立大学的photo.net数据集<sup>[3]</sup>)、众包(crowd sourcing)评价方法<sup>[19,49]</sup>。

2012年,由于图像视觉美学质量分析与度量受到越来越多的研究人员的关注,西班牙巴塞罗那自治大学计算机视觉中心的Murray等<sup>[35]</sup>构建了一个面向图像视觉美学质量分析与度量的大规模图像数据库(a large-scale database for aesthetic visual analysis, AVA),采用了从在线图像分享打分网站(dpchallenge.com)下载收集的打分方法,数据集包含255530幅有效图像,每一幅图像都由多个网站注册的人类艺术家会员评分(1~10分),每幅图被评分的人数范围为78~539,平均210人参与评分。该数据集是图像美学评价领域公认的基准测试集,目前有200余篇论文引用了该数据集。其标注数据质量较高,能够支持美感分类、美感评分、美感分布的学习,并且超过60种内容的摄影样式语义标记。然而该数据库并没有考虑图像的拍摄场景、相机参数以及后期处理带来的影响。

由于实验室内人工主观评价打分实验的参与人数有限,难以代表各个种族、年龄、国别等对美的认知差异,而在线图像分享打分网站下载收集的方法虽然参与者众多,然而难以控制图像来源,并且难以控制认知实验设计的各项参数,容易引入多种实验误差,因此一些研究人员采用众包的方案<sup>[19,49]</sup>,图像美学质量评价打分实验的图像来源可以严格控制,而利用互联网众包模式可以获得在一定范围内可控的众多参与者。

2013年,新加坡南洋理工大学的Agrawal等<sup>[50]</sup>分析了在线图像美学质量自动评价引擎(aesthetic quality inference engine, ACQUINE)与人工打分的差异,利用众包的方法获取图像美学质量的人工打分<sup>[50]</sup>,其参与打分的照片来源于33个摄影师采用相同的相机(全都调至自动拍摄模式)拍摄的7种限定的场景内容(景物、建筑、喷泉、人像、斑马线、户外、其他),然后在一个大型在线众包标注平台——亚马逊土耳其机器人 Amazon Mechanical Turk (AMT)上进行人工打分实验。该文认为,让测试者直接给出图像质量美学得分较难,而让测试者仅仅给出“喜欢”还是“不喜欢”的二值答案较容易,根据168个AMT测试者的打分结果统计图像美学质量得分。

随后出现的新型图像美学评价数据集包括 AADB

(2016年)<sup>[27]</sup>、PCCD(2017年)<sup>[41]</sup>、AROD(2018年)<sup>[46]</sup>、AVA-Reviews(2018年)<sup>[47]</sup>等,均已在2.3节中作了介绍。

### 3 结论与建议

目前在图像、图形、视频3种主要的可视媒体中,图像美学评价的研究较多,可以将上述研究现状总结为5层任务:美感分类是指给定一幅图像,输出“好”和“不好”或者美学质量“高”或“低”2个类别;美感评分是给出图像的美学质量评分,表现为一个连续数值;美感分布是给出图像的美学质量分数分布直方图;美学因素是给出图像的光影、配色、构图、模糊、运动、趣味等多个方面的评价;美学描述给出图像美学方面的语言评论,如图9所示。



图9 图像美学质量评价的5层任务

Fig. 9 The 5 layers of topics of image aesthetic assessment

对图像美学质量评价技术的分析总结与发展建议如下。

1) 美学评价5层模型与其对应的数据量和研究论文数量总体成倒三角形分布:层级越高,目前的数据量和论文数量越少,标注质量越低。

2) 目前该方向的研究主要集中于客观评价,即第1层美感分类(200+篇学术论文)与第2层美感评分(30+篇美学论文)。第3层美感分布的研究,目前还不到10篇论文,第4层美学因素的研究有2篇论文,第5层美学描述的研究只有2篇论文涉及。

3) 美学质量评价5层任务中,可用有标注数据的数量逐层递减。数据主要集中于前3层:美感分类、评分、分布(30万+),第4层美学因素可用数据小于1万

(AADB),第5层美学描述数据小于0.5万(PCCD),建议加强图像美学质量评价高层任务的研究力度,增加高层任务所需要的标注数据量,以及增加更丰富的图像美学质量自动评价新任务,并进一步拓展到视频美学质量与图形美学质量评价的研究。

4) 与图像识别(代表性数据集 ImageNet: 1400万+的有标注数据)等多项计算机视觉任务相比,图像美学的数据获取难度较大,整体数据规模较小(代表性数据集AVA:25万+有标注数据),主流图像美学数据集来自于西方国家,代表了西方的审美观,这对于研究东方人、中国人美学智能极为不利。建议继续扩大图像美学数据集的规模,尤其是需要建立中国人自己的图像美学数据集。

5) 美学质量评价 5 层任务中, 数据的标注质量逐层递减。前 3 层的数据标注, 典型代表 AVA 数据集中每一幅图像至少有 78 名艺术家标注, 平均每幅图像的标注人数为 210。而第 4 层和第 5 层每一幅图像的标注人数不到 10 人, 难以支撑美学评价多样性的分析, 建议增强图像美学数据集的标注质量与标注多样性。

6) 目前图像美学评价所运用的主流技术为深度神经网络, 其性能大大超越了之前的手工设计美学特征, 然而深度神经网络学习的特征可解释性较差, 难以支撑对人类美学智能的深入研究, 需要打开图像美学深度学习特征的黑匣子, 建议研究美学特征与人类美学智能的对应关系, 进一步探索人类美学智能的奥秘。

### 参考文献 (References)

- [1] Tong H, Li M, Zhang H, et al. Classification of digital photos taken by photographers or home users[C]//Proceedings of the 5th Pacific Rim Conference on Advances in Multimedia Information Processing. Heidelberg: Springer-Verlag, 2004: 198-205
- [2] Ke Y, Tang X, Jing F. The design of high-level features for photo quality assessment[C]//Proceedings of the 19th IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway NJ: IEEE, 2006: 419-426.
- [3] Datta R, Joshi D, Li J, et al. Studying aesthetics in photographic images using a computational approach[C]//Proceedings of the 9th European Conference on Computer Vision. Heidelberg: Springer-Verlag, 2006: 288-301.
- [4] Luo Y, Tang X. Photo and video quality evaluation: Focusing on the subject[C]//Proceedings of the 10th European Conference on Computer Vision (ECCV). Heidelberg: Springer-Verlag, 2008: 386-399.
- [5] Wong L K, Low K L. Saliency-enhanced image aesthetics class prediction[C]//Proceedings of the 16th IEEE International Conference on Image Processing. Piscataway NJ: IEEE, 2009: 997-1000.
- [6] Li C, Chen T. Aesthetic visual quality assessment of paintings [J]. IEEE Journal of Selected Topics in Signal Processing, 2009, 3(2): 236-252.
- [7] Nishiyama M, Okabe T, Sato I, et al. Aesthetic quality classification of photographs based on color harmony[C]//Proceedings of the 24th IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway NJ: IEEE, 2011: 33-40.
- [8] Dhar S, Ordonez V, Berg T L. High level describable attributes for predicting aesthetics and interestingness[C]//Proceedings of the 24th IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway NJ: IEEE, 2011: 1657-1664.
- [9] Marchesotti L, Perronnin F, Larlus D, et al. Assessing the aesthetic quality of photographs using generic image descriptors [C]//Proceedings of the 13th IEEE International Conference on Computer Vision. Piscataway NJ: IEEE, 2011: 1784-1791.
- [10] Hurter B. The best of photographic lighting techniques and images for digital photographers[M]. 2nd ed. New York: Amherst Media, 2007.
- [11] Hunter F, Biver S, Fuqua P. Light: Science and magic: An introduction to photographic lighting[M]. 3rd ed. New York: Focal Press, 2007.
- [12] Grey C. Master lighting guide for portrait photographers[M]. New York: Amherst Media, 2004.
- [13] Praker D. Basics photography: Lighting[M]. Switzerland: AVA Publishing, 2007.
- [14] Jin X, Zhao M, Chen X, et al. Learning artistic lighting template from portrait photographs[C]//Proceedings of the 11th European Conference on Computer Vision. Heidelberg: Springer-Verlag, 2010: 101-114.
- [15] Chen X, Jin X, Wu H, et al. Learning templates for artistic portrait lighting analysis[J]. IEEE Transactions on Image Processing, 2015, 24(2): 608-618.
- [16] Luo W, Wang X, Tang X. Content-based photo quality assessment[C]//Proceedings of the 13th IEEE International Conference on Computer Vision. Piscataway NJ: IEEE, 2011: 2206-2213.
- [17] Tang X, Luo W, Wang X. Content-based photo quality assessment[J]. IEEE Transactions on Multimedia, 2013, 15(8): 1930-1943.
- [18] Gray D, Yu K, Xu W, et al. Predicting facial beauty without landmarks[C]//Proceedings of the 11th European Conference on Computer Vision. Heidelberg: Springer-Verlag, 2010: 434-447.
- [19] Li C, Gallagher A, Loui A C, et al. Aesthetic quality assessment of consumer photos with faces[C]//Proceedings of the IEEE International Conference on Image Processing. Piscataway NJ: IEEE, 2010: 3221-3224.
- [20] Khan S S, Vogel D. Evaluating visual aesthetics in photographic portraiture[C]//Proceedings of the 8th Annual Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging. New York: ACM, 2012: 55-62.
- [21] Lu X, Lin Z, Jin H, et al. RAPID: Rating pictorial aesthetics using deep learning[C]//Proceedings of the ACM International Conference on Multimedia. New York: ACM, 2014: 457-466.
- [22] Kao Y, Wang C, Huang K. Visual aesthetic quality assessment with a regression model[C]//Proceedings of 2015 IEEE

- International Conference on Image Processing. Piscataway NJ: IEEE, 2015. 1583–1587.
- [23] Lu X, Lin Z, Shen X, et al. Deep multi-patch aggregation network for image style, aesthetics, and quality estimation[C]// Proceedings of 2015 IEEE International Conference on Computer Vision. Piscataway NJ: IEEE, 2015: 990–998.
- [24] Dong Z, Tian X. Multi-level photo quality assessment with multi-view features[J]. *Neurocomputing*, 2015, 168: 308–319.
- [25] Wang W, Zhao M, Wang L, et al. A multi-scene deep learning model for image aesthetic evaluation[J]. *Signal Processing Image Communication*, 2016, 47(C): 511–518.
- [26] Kao Y, Huang K, Maybank S. Hierarchical aesthetic quality assessment using deep convolutional neural networks[J]. *Signal Processing Image Communication*, 2016, 47: 500–510.
- [27] Kong S, Shen X, Lin Z, et al. Photo aesthetics ranking network with attributes and content adaptation[C]// Proceedings of 14th European Conference on Computer Vision. Heidelberg: Springer-Verlag, 2016: 662–679.
- [28] Jin X, Chi J, Peng S, et al. Deep image aesthetics classification using inception modules and fine-tuning connected layer [C]// Proceedings of the 8th International Conference on Wireless Communications and Signal Processing. Piscataway NJ: IEEE, 2016, doi: 10.1109/WCSP.2016.7752571.
- [29] Ma S, Liu J, Chen C W. A-Lamp: Adaptive layout-aware multi-patch deep convolutional neural network for photo aesthetic assessment[C]// Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway NJ: IEEE, 2017: 722–731.
- [30] Jin X, Wu L, Li X D, et al. ILGNet: Inception modules with connected local and global features for efficient image aesthetic quality classification using domain adaptation[J]. *arXiv.org*, 2018, arXiv:1610.02256v3.
- [31] Jin B, Segovia M V O, Süsstrunk S. 2016. Image aesthetic predictors based on weighted CNNs[C]// Proceedings of 2016 IEEE International Conference on Image Processing. Piscataway NJ: IEEE, 2016: 2291–2295.
- [32] Wang Z, Liu D, Chang S, et al. Image aesthetics assessment using deep chatterjee's machine[C]// Proceedings of 2017 International Joint Conference on Neural Networks. Piscataway NJ: IEEE, 2017: 941–948.
- [33] Hou L, Yu C P, Samaras D. Squared earth mover's distance-based loss for training deep neural networks[J]. *arXiv.org*, 2017, arXiv: 1611.05916
- [34] Deng Y, Chen C L, Tang X. Image aesthetic assessment: An experimental survey[J]. *IEEE Signal Processing Magazine*, 2016, 34(4): 80–106.
- [35] Murray N, Marchesotti L, Perronnin F. AVA: A large-scale database for aesthetic visual analysis[C]// Proceedings of the 25th IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway NJ: IEEE, 2012: 2408–2415.
- [36] Wu O, Hu W M, Gao J. Learning to predict the perceived visual quality of photos[C]// Proceedings of 2011 IEEE International Conference on Computer Vision. Piscataway NJ: IEEE, 2011: 225–232.
- [37] Park T S, Zhang B T. Consensus analysis and modeling of visual aesthetic perception[J]. *IEEE Transactions on Affective Computing*, 2015, 6(3): 272–285.
- [38] Kim W H, Choi J H, Lee J S. Subjectivity in aesthetic quality assessment of digital photographs: Analysis of user comments [C]// Proceedings of the 23rd ACM international conference on Multimedia. New York: ACM, 2015: 983–986.
- [39] Kim W H, Choi J H, Lee J S. Objectivity and subjectivity in aesthetic quality assessment of digital photographs[J]. *IEEE Transactions on Affective Computing*, 2018, doi: 10.1109/TAFFC.2018.2809752.
- [40] Jin X, Wu L, Li X, et al. Predicting aesthetic score distribution through cumulative Jensen-Shannon divergence[C]// Proceedings of AAAI Conference on Artificial Intelligence (AAAI). New York: AAAI, 2018.
- [41] Chang K Y, Lu K H, Chen C S. Aesthetic critiques generation for photos[C]// Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway NJ: IEEE, 2017: 3534–3543.
- [42] Collomosse J, Tu B, Wilber M, et al. Sketching with Style: Visual search with sketches and aesthetic context[C]// Proceedings of IEEE International Conference on Computer Vision. Piscataway NJ: IEEE, 2017: 2679–2687.
- [43] Cui C, Fang H, Deng X, et al. Distribution-oriented aesthetics assessment for image search[C]// Proceedings of International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2017: 1013–1016.
- [44] Wang W, Shen J. Deep cropping via attention box prediction and aesthetics assessment[C]// 2017 IEEE International Conference on Computer Vision. Piscataway NJ: IEEE, 2017: 2205–2213.
- [45] Ren J, Shen X, Lin Z, et al. Personalized image aesthetics[C]// Proceedings of 2017 IEEE International Conference on Computer Vision. Piscataway NJ: IEEE, 2017: 638–647.
- [46] Schwarz K, Wieschollek P, Lensch H P A. Will people like your image?[J]. *arXiv.org*, 2018, arXiv:1611.05203.
- [47] Wang W S, Yang S, Zhang W S, et al. Neural aesthetic image reviewer[J]. *arXiv.org*, 2018, arXiv:1802.10240.
- [48] Yanulevskaya V, Uijlings J, Bruni E, et al. In the eye of the beholder: Employing statistical analysis and eye tracking for analyzing abstract paintings[C]// Proceedings of the 20th ACM

- International Conference on Multimedia. New York: ACM, 2012: 349–358.
- [49] Agrawal A, Premachandran V, Kakarala R. Rating image aesthetics using a crowd sourcing approach[C]//Workshops of the 6th Pacific–Rim Symposium on Image and Video Technology. New York: Springer–Verlag New York Inc., 2013, 8334: 24–32.
- [50] Datta R, Wang J Z. ACQUINE: Aesthetic quality inference engine—real time automatic ratings of photo aesthetics[C]//Proceedings of the ACM International Conference on Multimedia Information Retrieval. New York: ACM, 2010: 421–424.

## Image aesthetic quality assessment: A survey

JIN Xin<sup>1</sup>, ZHOU Bin<sup>2</sup>, ZOU Dongqing<sup>3</sup>, LI Xiaodong<sup>1</sup>, SUN Hongbo<sup>1</sup>, WU Le<sup>1</sup>

1. Department of Computer Science and Technology, Beijing Electronics Science and Technology Institute, Beijing 100070, China
2. School of Computer Science and Engineering, Beihang University, Beijing 100191, China
3. Sense Time, Beijing 100084, China

**Abstract** Image quality assessment (IQA) is to measure the perceived image degradation while image aesthetic quality assessment (IAQA) is to automatically give a review of aesthetics of an image. The composition, color, illumination, depth of field, blurriness, sharpness, etc. are analyzed by means of computer to analog the aesthetic perception and cognition of human beings. IAQA is a novel interdisciplinary field of computer aesthetics and computer vision, psychology, and virtual reality. IAQA can be used for photography, movie and television, e-commerce, fashion design, graphic design, document typesetting, industry design, automobile industry, architectural design, beauty makeup, etc. In this article, we give a brief history of IAQA. Then we review the methods based on hand-crafted aesthetic feature design, the methods based on deep aesthetic feature learning, new topics in IAQA, and dataset collection for IAQA. Finally, we give the future directions of IAQA.

**Keywords** image aesthetic quality assessment; computational aesthetics; virtual reality; computational art; visual art ●



(责任编辑 刘志远)