



中国图形学报

ISSN1006-8961
CN11-3758/TB



第27卷第8期（总第316期）
2022年8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

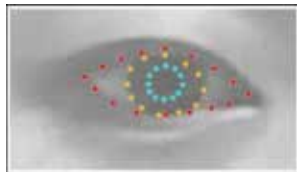
Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

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

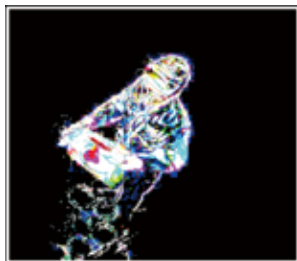
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

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

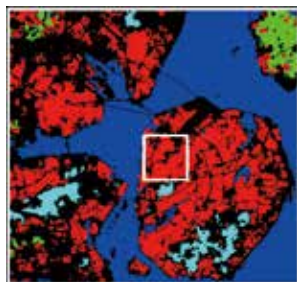
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



自然光普通摄像头的眼部分割及特征点定位数据集ESLD (第2329页)



双视图三维卷积网络的工业装箱行为识别(第2368页)



多源特征自适应融合网络的高分遥感影像语义分割(第2516页)

综述

虚拟现实图像客观质量评价研究进展

周玉, 汪一, 李雷达, 高陈强, 卢兆林 2313

数据集论文

自然光普通摄像头的眼部分割及特征点定位数据集ESLD

张俊杰, 孙光民, 郑鲲, 李煜, 付晓辉, 慈康怡, 申俊杰, 孟凡超, 孔江萍, 张玥 2329

图像处理和编码

双尺度顺序填充的深度图像修复

陈东岳, 朱晓明, 马腾, 宋园园, 贾同 2344

二维码位流长度最小化算法

袁泰凌, 徐昆 2356

图像分析和识别

双视图三维卷积网络的工业装箱行为识别

胡海洋, 潘健, 李忠金 2368

面向航拍图像中工程车辆检测与识别的改进胶囊网络

钟映春, 郑海阳, 张文祥, 王波, 罗志勇 2380

用于骨架行为识别的多维特征融合注意力机制

姜权晏, 吴小俊, 徐天阳 2391

高性能整数倍稀疏网络行为识别研究

臧影, 刘天娇, 赵曙光, 杨东升 2404

面向工业零件分拣系统的低纹理目标检测

闫明, 陶大鹏, 普园媛 2418

融合策略优选和双注意力的单阶段目标检测

戴坤, 许立波, 黄世旸, 李鉴铃 2430

融合视觉词与自注意力机制的视频目标分割

季传俊, 陈亚当, 车洵 2444

基于Gaussian-Hermite矩的旋转运动模糊不变量

郭锐, 贾丽, 郝宏翔, 墨瀚林, 李华 2458

图像理解和计算机视觉

RGB-D语义分割: 深度信息的选择使用

赵经阳, 余昌黔, 桑农 2473

自纠正噪声标签的人脸美丽预测

甘俊英, 吴必诚, 翟懿奎, 何国辉, 麦超云, 白振峰 2487

面向非对称特征注意力和特征融合的太赫兹图像检测

曾文健, 朱艳, 沈韬, 曾凯, 刘英莉 2496

面向纹理平滑的方向性滤波尺度预测模型

林俊彦, 刘春晓, 章金凯, 李泓易 2506

遥感图像处理

多源特征自适应融合网络的高分遥感影像语义分割

张文凯, 刘文杰, 孙显, 许光奎, 付琨 2516

结合上下文编码与特征融合的SAR图像分割

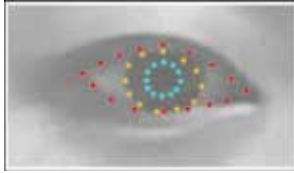
范艺华, 董张玉, 杨学志 2527

一阶全卷积遥感影像倾斜目标检测

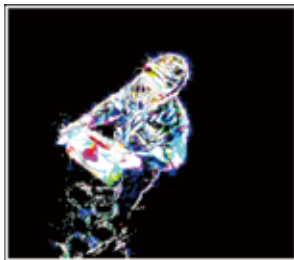
周院, 杨庆庆, 马强, 薛博维, 孔祥楠 2537

CONTENTS

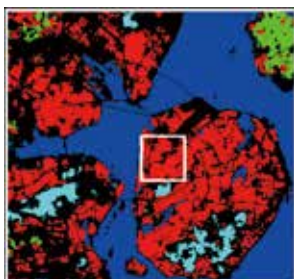
JOURNAL OF IMAGE AND GRAPHICS



ESLD: eyes segment and landmark detection in the wild (P2329)



Dual-view 3D ConvNets based industrial packing action recognition(P2368)



Multi-source features adaptation fusion network for semantic segmentation in high-resolution remote sensing images(P2516)

Review

Research progress in objective quality evaluation of virtual reality images

Zhou Yu, Wang Yi, Li Leida, Gao Chenqiang, Lu Zhaolin 2313

Dataset

ESLD: eyes segment and landmark detection in the wild

Zhang Junjie, Sun Guangmin, Zheng Kun, Li Yu, Fu Xiaohui, Ci Kangyi, Shen Junjie, Meng Fanchao, Kong Jiangping, Zhang Yue 2329

Image Processing and Coding

Depth image recovery based on dual-scale sequential optimized filling

Chen Dongyue, Zhu Xiaoming, Ma Teng, Song Yuanyuan, Jia Tong 2344

Minimization of bit stream length of QR codes

Yuan Tailing, Xu Kun 2356

Image Analysis and Recognition

Dual-view 3D ConvNets based industrial packing action recognition

Hu Haiyang, Pan Jian, Li Zhongjin 2368

Improved capsule network method for engineering vehicles detection and recognition in aerial images

Zhong Yingchun, Zheng Haiyang, Zhang Wenxiang, Wang Bo, Luo Zhiyong 2380

M2FA: multi-dimensional feature fusion attention mechanism for skeleton-based action recognition

Jiang Quanyan, Wu Xiaojun, Xu Tianyang 2391

Action recognition analysis derived of integer sparse network

Zang Ying, Liu Tianjiao, Zhao Shuguang, Yang Dongsheng 2404

Texture-less object detection method for industrial components picking system

Yan Ming, Tao Dapeng, Pu Yuanyuan 2418

Single stage object detection algorithm based on fusing strategy optimization selection and dual attention mechanism

Dai Kun, Xu Libo, Huang Shiyang, Li Yunling 2430

Visual words and self-attention mechanism fusion based video object segmentation method

Ji Chuanjun, Chen Yadang, Che Xun 2444

Rotational motion blur invariants based on Gaussian-Hermite moments

Guo Rui, Jia Li, Hao Hongxiang, Mo Hanlin, Li Hua 2458

Image Understanding and Computer Vision

RGB-D semantic segmentation: depth information selection

Zhao Jingyang, Yu Changqian, Sang Nong 2473

Self-correcting noise labels for facial beauty prediction

Gan Junying, Wu Bicheng, Zhai Yikui, He Guohui, Mai Chaoyun, Bai Zhenfeng 2487

Terahertz image detection combining asymmetric feature attention and feature fusion

Zeng Wenjian, Zhu Yan, Shen Tao, Zeng Kai, Liu Yingli 2496

Texture-smoothing-oriented directional filtering scales-predicting model

Lin Junyan, Liu Chunxiao, Zhang Jinkai, Li Hongyi 2506

Remote Sensing Image Processing

Multi-source features adaptation fusion network for semantic segmentation in high-resolution remote sensing images

Zhang Wenkai, Liu Wenjie, Sun Xian, Xu Guangluan, Fu Kun 2516

The integrated contextual encoding and feature fusion SAR images segmentation method

Fan Yihua, Dong Zhangyu, Yang Xuezhi 2527

Improved one-stage fully convolutional network for oblique object detection in remote sensing imagery

Zhou Yuan, Yang Qingqing, Ma Qiang, Xue Bowei, Kong Xiangnan 2537

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2022)08-2313-16

论文引用格式: Zhou Y, Wang Y, Li L D, Gao C Q and Lu Z L. 2022. Research progress in objective quality evaluation of virtual reality images. Journal of Image and Graphics, 27(08): 2313-2328 (周玉, 汪一, 李雷达, 高陈强, 卢兆林. 2022. 虚拟现实图像客观质量评价研究进展. 中国图象图形学报, 27(08): 2313-2328) [DOI:10.11834/jig.210949]

虚拟现实图像客观质量评价研究进展

周玉^{1,2}, 汪一^{3,1}, 李雷达^{4*}, 高陈强⁵, 卢兆林^{1,2}

1. 中国矿业大学信息与控制工程学院, 徐州 221116; 2. 徐州市第一人民医院, 徐州 221116; 3. 江苏师范大学科文学院, 徐州 221132; 4. 西安电子科技大学人工智能学院, 西安 710071; 5. 重庆邮电大学通信与信息工程学院, 重庆 400065

摘要: 随着科学技术的发展, 虚拟现实(virtual reality, VR)技术逐渐渗透到医疗、教育、军事和娱乐等众多领域, 并凭借在各个领域广阔的应用前景而受到广泛关注。鉴于视觉质量是决定VR技术能否成功应用的关键, 且图像是VR应用最基础和最重要的视觉信息载体, VR图像质量评价已经成为质量评价领域的重要前沿性研究方向。与传统图像质量评价类似, VR图像质量评价可以分为主观质量评价和客观质量评价。由于客观质量评价相比主观质量评价具有成本低、稳定性高和应用范围广等优点, 对VR图像客观质量评价的研究受到了国内外学者的高度重视。目前, 关于VR图像客观质量评价的研究已经取得了一定进展, 但是文献中缺少对该方向相关研究方法的综述。基于此, 本文针对VR图像客观质量评价的研究进行概述。首先, 对VR图像质量评价的研究现状进行分析。然后, 重点对现有的VR图像客观质量评价模型进行综述。具体地, 根据模型是否需要使用原始无失真图像信息作为参考, 将现有的VR图像客观质量评价模型划分为全参考型和无参考型两大类。其中, 全参考型方法进一步划分为基于峰值信噪比/结构相似度的方法和基于传统机器学习的方法。根据特征表达空间的不同, 无参考型VR图像质量评价模型划分为3类: 基于等矩形投影表达空间的方法、基于其他投影表达空间的方法和基于实际观看空间的方法。介绍完各类模型后, 分别对其相应的优缺点进行分析。此外, 本文对VR图像客观质量评价模型的性能评价指标和现有VR图像质量评价数据库进行了归纳。最后对VR图像客观质量评价模型的应用进行了介绍, 并指出了未来的研究可能的发展方向。

关键词: 图像质量评价; 客观评价; 虚拟现实(VR); 球面图像; 等矩形投影(ERP)

Research progress in objective quality evaluation of virtual reality images

Zhou Yu^{1,2}, Wang Yi^{3,1}, Li Leida^{4*}, Gao Chenqiang⁵, Lu Zhaolin^{1,2}

1. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China; 2. Xuzhou First Peoples Hospital, Xuzhou 221116, China; 3. Jiangsu Normal University Kewen College, Xuzhou 221132, China; 4. School of Artificial Intelligence, Xidian University, Xi'an 710071, China; 5. School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract: Virtual reality (VR) technology has gradually penetrated into many fields such as medical education, military and entertainment. Given visual quality is the key to the successful application of VR technology, and the image is visual information benched carrier of VR applications, VR image quality evaluation has become an important frontier research direction for quality evaluation. Just like traditional image quality evaluation, VR image quality evaluation can be divided

收稿日期: 2021-10-13; 修回日期: 2021-12-15; 预印本日期: 2021-12-22

* 通信作者: 李雷达 ldli@xidian.edu.cn

基金项目: 国家自然科学基金项目(62001475); 江苏省自然科学基金项目(BK20200649)

Supported by: National Natural Science Foundation of China(62001475); Natural Science Foundation of Jiangsu Province, China(BK20200649)

into subjective quality evaluation and objective quality evaluation. Among them, the subjective quality evaluation methods refer to images scoring through human eyes followed by some data processing steps to obtain the subjective scores, while the objective quality evaluation methods focus on the methods of images scoring based on designing the mathematical model using computers to simulate the subjective scoring results as closely as possible. Compared to the subjective quality evaluation, the objective quality evaluation has its priorities of lower cost, stronger stability and wider application scope. Many researchers in the scientific research institutions and colleges have dedicated to studying the objective quality evaluation of VR images. Our executive summary is focused on the research of objective quality evaluation of VR images. First, the current situation of VR image quality evaluation is summarized. Then, the existing objective quality evaluation models of VR images are mainly introduced and analyzed. According to whether the models need to use the original undistorted image information as modeling reference, current objective quality evaluation models for VR images are divided into two types, including the full-reference (FR) and no-reference (NR) types. The ground truth based FR models need the completed original image information as reference, while the NR models can achieve quality evaluation of a distorted VR image without any reference information. Specifically, the FR models can be divided into two categories for VR images, including the peak-signal-to-noise ratio/structural similarity (PSNR/SSIM) based methods and the machine learning based methods. The latter first extract features from VR images and then train the quality evaluation model via the support vector regression method or the random forest method. The NR models are further divided into three categories; the equirectangular projection (ERP) expression space based methods, the other projection expression spaces based methods and the actual viewing space based methods. These models are classified according to the space in which features are extracted. For the first kind of models, the raw spherical VR image is first transferred to the ERP space for feature expression, while it is converted to other projection spaces or the actual viewing space for the other two kinds of methods. The reason why the ERP expression space based methods are listed as a separate category is that the ERP space is the default and the mostly used projection space. Followed the space transformation, there are also two options for the sequential quality evaluation consisting of the traditional quality evaluation methods and the deep-learning based methods. For the traditional ones, the manual features are first extracted and fused to generate the final quality score or fed into the model trainer to obtain the quality assessment model. For the deep-learning based methods, both the feature extraction and quality prediction steps are conducted based on deep neural networks. The pros and cons of them are further analyzed. In addition, the performance evaluation indexes of the objective quality evaluation models of VR images are introduced, which are consistent to other image quality evaluations. The detailed existing VR image databases are summarized subsequently. Finally, our review focuses on the applications of VR image objective quality evaluation models and predicts the future potentials further.

Key words: image quality evaluation; objective evaluation; virtual reality (VR); spherical image; equirectangular projection (ERP)

0 引言

随着社会生产力和信息技术的持续发展,虚拟现实(virtual reality, VR)技术作为一种能够模拟360°真实世界场景,为人们提供仿佛置身现实生活中的真实感、沉浸感和互动感的技术,其发展速度相当迅猛(曹凡,2019;周忠等,2015)。尤其是随着5G时代的到来,VR相关产业迎来了全新的发展机遇。VR应用已经成为新一代信息技术中的重要前沿性研究方向,具有广阔的应用前景和不可估量的市场价值(郑菲,2020)。

VR应用的最终目标是为用户提供逼真的、身临其境的体验。用户在头部佩戴一定硬件设备后,只需要旋转头部即可看到360°内的场景,从而获得沉浸式体验,如图1所示(Lim等,2018)。显然,VR应用开发过程中最关注的问题是用户体验质量的高低,其主要受VR内容视觉质量的影响,原因在于视觉是人类感觉中最重要的一种。VR图像是VR内容在视觉方面最基础也是最重要的信息载体。在实际应用中,VR图像的获取往往需要经过图像采集、拼接、投影映射、编码、传输、存储和显示等多个过程。其中的每个过程均可能在VR图像中引入失真,如采集过程的模糊和噪声、投影失真、编码过程

的压缩失真和拼接过程的重影失真等等。这些失真均会影响用户对VR应用的视觉体验质量。因此,研究VR图像的质量评价(VR image quality assessment, VR IQA)具有重要意义。其研究成果能够用于指导VR相关技术的设计和优化,进而提高VR应用的用户视觉体验质量。

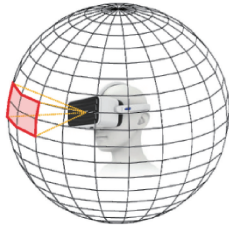


图1 用户进行VR体验时的情景展示(Lim等,2018)

Fig.1 The scene presentation of users' VR experience(Lim et al.,2018)

目前,图像质量评价的研究众多。从被评价图像的内容角度,现有研究可以分为针对自然图像的质量评价和非自然图像的质量评价,如屏幕图像和色调映射图像的质量评价(Wang等,2016;Fang等,2017,2020)。从研究方法来划分,图像质量评价分为主观质量评价和客观质量评价两大类(顾颀,2015;吴金建,2014;周玉,2019)。其中,主观质量评价是指通过人眼观看对图像质量进行打分,而客观质量评价是指通过设计数学模型来模拟人眼对图像质量的感知,以尽可能实现和主观评价一致的评价结果。相比主观质量评价,客观质量评价具有明显的成本低、稳定性高和实用性强等优点(Zhou等,2018,2019)。因此,客观图像质量评价受到更广泛的关注。尽管目前的研究已经在自然图像和部分非自然图像的客观质量评价方面取得重要进展,并已经在一些领域得到初步应用,但是这些方法在评价VR图像质量时却有较大的性能局限性。造成这种结果的原因是多方面的:其一,实际VR系统中的图像为3维球面形式,而传统质量评价方法主要针对2维平面图像进行设计。虽然为了方便图像传输和存储,3维球面VR图像通常被非线性投影到2维空间,但是投影过程往往会在生成的2维全景图像中引入失真(Jiang等,2021a),而这种失真并不真实存在于原始VR图像中。例如,最常用的等矩形投影(equirectangular projection,ERP)方法会在图像两极区域带来严重的几何形变,如图2(b)所示。此

外,投影后图像和原始VR图像的像素分布呈非线性关系,这较大幅度地改变了图像中原本的失真分布特性和结构,造成了难以对投影后全景图像和原始VR图像进行质量关系建模的问题。其二,传统质量评价模型主要针对模糊、压缩失真和噪声等传统类型失真进行评价,而VR图像中特有的失真(如拼接失真)与这些失真的特性大不相同(Chen等,2020;Madhusudana和Soundararajan,2019;Li等,2020a)。基于以上分析,亟需有针对性地设计VR图像的客观质量评价模型。



(a) 球面VR图像



(b) ERP形式图像

图2 一幅球面VR图像及相应的ERP形式图像

Fig.2 A spherical VR image and the corresponding ERP form image((a)spherical VR image;(b)ERP form image)

越来越多的科研机构 and 科研院校加入到该项研究工作中,例如上海交通大学、江西财经大学、宁波大学、中国科学技术大学、浙江大学、北京航空航天大学、西安电子科技大学、武汉大学、辽宁大学、美国斯坦福大学、美国德克萨斯大学、韩国科学技术院、印度科技学院等。目前的研究工作主要包括VR图像主观质量评价和客观质量评价两方面的研究。如前所述,后者比前者具有更广泛的实际应用价值,因此本文主要对VR图像客观质量评价的相关研究进行综述,内容框架如图3所示。图中ERP格式图像中的红色边框图像块及位于右上角的放大图是为了突出失真和无失真图像之间的差异。

根据模型中是否需要无失真的高质量VR图像信息作为参考,本文将现有客观质量评价模型划分为全参考(full-reference,FR)型和无参考(no-reference,NR)型,并对各模型的具体实现方法进行梳理,对各类模型的优缺点进行分析。其中,FR模型是指需要整幅无失真VR图像作为参照,而NR模型则无需任何参考信息即可实现对失真VR图像的质量预测(Fang等,2018;Zhang等,2015;Min等,2020)。更具体地,将FR型VR图像质量评价模型进一步划分为基于峰值信噪比/结构相似度(peak-

signal-to-noise ratio/structural similarity, PSNR/SSIM) 的方法和基于传统机器学习的方法;根据特征表达空间的不同,将 NR 型 VR 图像质量评价模型划分为 3 类:基于 ERP 表达空间的方法、基于其他投影表达空间的方法和基于实际观看空间的方法。除此之外,本文还对 VR 图像质量评价数据库、性能评价指标和模型应用进行总结,同时指出该方向研究可能的发展趋势。

尽管国内外已经有多篇关于图像质量评价的综述文献,但是并没有专门针对 VR 图像客观质量评价的综述。具体地,在现有的图像质量评价综述文献中,王志明(2015)、Niu 等人(2019)、Yang

等人(2019)、Athar 和 Wang(2019) 仅针对传统图像的质量评价模型或相关数据库进行介绍,并未提及 VR 图像质量评价的相关信息。Zhai 和 Min(2020)、方玉明等人(2021) 虽然在对图像质量评价方向研究进行综述的同时,简单介绍了 VR 图像质量评价,但是由于相应工作并不是专门针对 VR 图像质量评价进行研究,相关介绍较为笼统,不够详细全面和有针对性,其中缺少针对 VR 图像客观质量评价模型的详细介绍、深入分析和总结。因此,本文是对现有图像质量评价综述的扩展与补充,更是对 VR 图像客观质量评价综述的补缺。

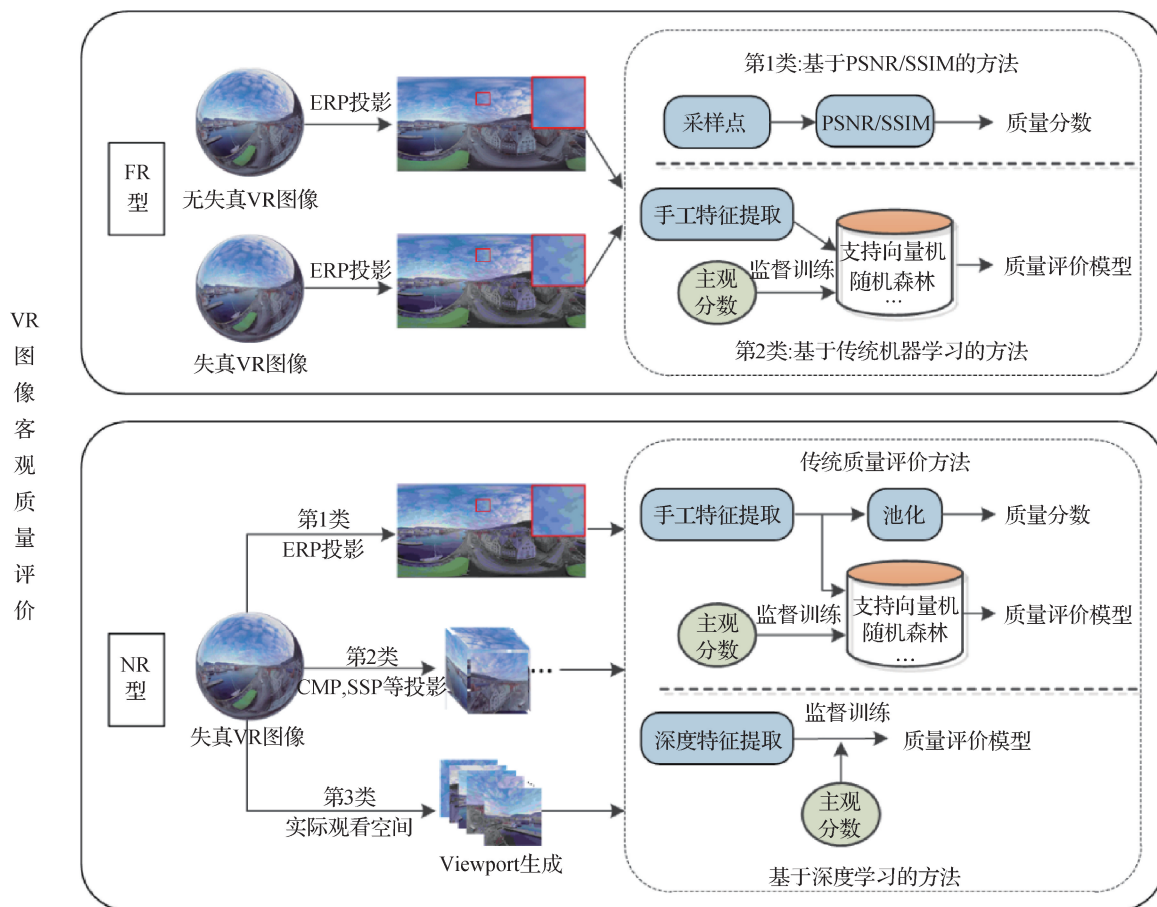


图3 VR图像客观质量评价整体框架图

Fig. 3 The overall framework of objective quality assessment of VR images

1 VR图像客观质量评价

类似于传统质量评价,根据模型中是否需要使用无失真VR图像的信息作为参考,现有的VR图

像客观质量评价模型可以分为FR和NR两种类型。下面分别对两种类型中的各个模型进行详细介绍,并对各类型方法的优缺点进行分析与总结。

1.1 FR型VR图像质量评价方法

该类方法的设计思路主要分为两类。第1类是

直接借助目前发展较为成熟的传统2维图像质量评价方法,即PSNR/SSIM(Wang等,2004)进行VR图像质量评价。更具体地,该类方法首先将原始球面VR图像投影到ERP表达空间,然后在表达空间中通过PSNR/SSIM方法实现VR图像质量评价。例如,Yu等人(2015)首先在球面VR图像上进行均匀采样,然后将失真和无失真球面图像投影到ERP表达空间并确定每个采样点分别在原始无失真ERP图像和失真ERP图像中的对应点,然后利用PSNR方法计算各对应点之间的距离。最后,将所有采样点对应的PSNR均值作为VR图像的质量分数。该方法命名为球面PSNR(spherical PSNR, S-PSNR)方法。Zakharchenko等人(2016)提出一种基于卡斯特抛物线投影(craster parabolic projection)空间的PSNR方法(CPP-PSNR)。作者将图像映射到该空间进行表达的原因是他们通过研究发现该空间的像素分布和球面域分布更相近,因此使用该空间能够避免位置变动带来的计算误差。将图像映射到CPP空间后,通过在该空间计算PSNR值来预测VR图像的质量分数。Sun等人(2017b)以ERP空间和立方体面片投影(cubemap projection, CMP)空间为例,提出首先借助PSNR方法在表达空间中计算像素级误差,然后利用各个位置像素在真实球面空间中的映射面积计算各像素权重,并将加权平均结果作为VR图像的质量分数。类似地,Chen等人(2018)提出了一种利用球面空间映射面积计算各位置像素结构相似性权重的VR图像质量评价方法。

第2类是采用传统机器学习的方法。该类方法首先进行手工特征提取,然后训练质量评价模型。例如,Madhusudana和Soundararajan(2019)提出了一种针对拼接失真的VR图像质量评价模型。该模型采用特征提取—质量评价模型训练的思路,首先设计了两组特征来描述拼接失真(模糊、重影和几何失真)在VR图像中引起的结构和空间一致性的变化,并分别从失真VR图像和原始无失真VR图像中进行提取。然后,计算失真和无失真图像之间的特征差值,并将它们作为支持向量回归器的输入训练质量评价模型。郑珂珂(2017)重点介绍了VR全景视频拼接技术,研究了影响全景视频质量的关键因素,并提出了一种VR全景视频拼接质量评估模型。该模型从拼接后图像与原始子图像在亮度/

色彩上的不一致性、模糊和错位等方面分别进行特征提取,然后采用反向传播算法进行质量评价模型的训练。

以上模型设计思路清晰,易于理解,但是在实际应用中具有较大局限性,主要原因在于:1)与传统FR型质量评价模型类似,它们均需要原始无失真VR图像作为参考,而实际应用中的无失真VR图像难以获得(Yue等,2019;Wu等,2020)。2)FR型方法要求失真图像与原始无失真图像之间具有相同的分辨率,且像素之间具有相同的位置关系。而360°VR图像一般是由多幅具有重叠区域的子图像拼接而成,因此VR图像和原始子图像分辨率不同,这极大地限制了该类型方法在现实中的应用。

1.2 NR型VR图像质量评价方法

根据特征表达空间的不同,NR型VR图像质量评价模型可以划分为3类:第1类是基于ERP表达空间的模型;第2类是基于其他投影表达空间的模型;第3类是基于实际观看空间的模型。其中,将ERP表达空间的模型单独划分为一类的原因是ERP是目前默认使用的、最常用的VR图像投影映射方法(Yang等,2018)。接下来,对3类NR型VR图像质量评价模型分别进行详细介绍和分析。

1.2.1 基于ERP表达空间的方法

该类VR图像质量评价方法的主要思想是直接通过对ERP表达形式图像进行特征提取和特征融合得到质量分数。如Kim等人(2020)提出了一种基于分块的VR图像质量预测网络模型。该模型首先将ERP形式图像划分为大小均匀且不相互重叠的图像块,然后采用深度学习的方法对每个块进行特征编码和空间位置编码,并将位置特征和空间特征融合作为空间特征权重,最后通过计算所有图像块的加权和得到质量分数。在模型训练阶段,该方法采用敌对学习策略通过原始无失真ERP图像对分数预测网络的学习进行监督,以获得更好的分数预测模型。Liu等人(2021)提出了一种基于结构和自然性特征的VR图像质量评价模型。对于ERP形式全景图像,首先在梯度域提取空间和局部特征进行结构失真描述,同时直接从图像中提取亮度、颜色和熵特征进行自然性描述。最后将所有特征输入随机森林回归器训练质量评价模型。

由于ERP空间图像为2维平面形式,所以该类方法也较为简单和易于理解。然而,ERP表达空间

中的图像存在明显的拉伸形变,尤其是越靠近图像的两极区域,拉伸形变越明显,如图2(b)所示。这种拉伸效应不仅造成图像结构上的改变,也使图像中的失真特性发生变化。图像在表达空间和实际观看空间中的差异会造成客观评价结果与主观评价结果的不一致,从而降低该类方法的评价准确性。

1.2.2 基于其他投影表达空间的方法

该类模型主要是为了克服 ERP 形式图像在两极处会产生明显拉伸形变的问题,它们首先将 VR 图像投影映射到其他表达空间,例如分段球面投影空间(segmented spherical projection, SSP)或 CMP 空间等,然后在新的表达空间进行质量评价。Zheng 等人(2020)提出了一种基于 SSP 的 VR 图像质量评价模型。投影到 SSP 空间的图像分成了3部分,包括两个两极区域和一个赤道区域,然后分别对各区域进行特征提取。在特征提取过程中,分别利用两类区域的热力图作为相应区域的权重进行特征表达,以模拟人类视觉特性。最后将所有特征输入随机森林回归器训练最终的质量评价模型。Jiang 等人(2021a)提出了3种基于 CMP 空间的 VR 图像质量评价模型。首先,将待评价图像投影到 CMP 空间得到6个面图像。第1种模型是直接对6个面图像进行特征提取和特征融合得到质量分数;第2种模型是利用注意力机制作为特征权重计算质量分数;第3种模型是将热点图同样进行 CMP 投影得到热点图的6个面图像,然后根据注意力强弱,从 VR 图像6个面中提取4种不同注意力级别的特征,并通过特征融合获得质量分数。

相比 ERP 表达空间,这些空间解决了 ERP 映射在两极区域引入拉伸形变失真的问题。然而,它们同样存在一些其他的映射失真,仍然无法等同于实际中用户体验的球面 VR 图像。综上,第1类和第2类基于投影表达空间的模型均依赖于表达空间的优劣。对于该问题,如何寻找表达空间图像和实际观看空间图像之间的对应关系并进行合理建模是解决问题的关键。

1.2.3 基于实际观看空间的方法

该类模型主要是通过模拟人类在现实中观看 VR 内容时的真实过程或特性,以获得与人类主观评价更接近的客观质量评价结果。目前,该类方法均是基于视窗(viewport)的方法,其目的是模拟人眼在某一时刻视角范围有限这一特性,即用户在观看

360° VR 内容时的某一时刻,只能一次性看到 viewport 范围内的内容,只有通过旋转头部才能观看整个 360° 的内容。此外,用户视觉感知到的内容实际为球面内容的切面图,如图4所示。当位于中心 O 处的用户朝着 OO' 方向观看时(O' 为视线方向与球面交点),用户视觉捕捉到的图像内容为以 O' 为中心的视窗范围内的球体切面图 $ABCD$ 。该图称为用户实际观看到的 viewport 图。基于此,该类方法首先从 ERP 形式图像中恢复出实际观看空间中的 viewport 图,然后通过度量 viewport 图的质量来计算整个 VR 图像的质量分数。如 Xu 等人(2021)提出一种基于局部和全局评价的 VR 图像质量评价模型。其中,局部评价分支首先借助关键点检测以及热力图获取若干关键 viewport 图,再借助 ResNet18 进行特征提取,最后利用图卷积网络进行局部质量分数预测。全局分支采用 Zhang 等人(2020)提出的深度双线性卷积神经网络进行。最后融合局部和全局特征得到最终质量分数。Sun 等人(2020)提出了一种多通道卷积神经网络的 VR 图像质量评价模型。该模型通过从 ERP 图像中获得前、后、左、右、上、下6个方向的 viewport 图来模拟用户的真实观看过程。获取 viewport 图像后,构建卷积神经网络分别对每个图像进行特征提取,最后将多方向特征融合进行质量分数预测。类似地,Zhou 等人(2021)同样从 ERP 形式图像中获取6个方向 viewport 图,然后借助多任务学习思想,设计了使用失真鉴别网络辅助质量预测网络的多流网络模型。Sui 等人(2021)提出了一种新颖的思路来进行 VR 图像客观质量评价。作者将用户旋转头部来观看 360° VR 内容的真实观看过程想象为用户不动,VR 内容自行旋转的过程。另外,结合 viewport 的概念以及观看

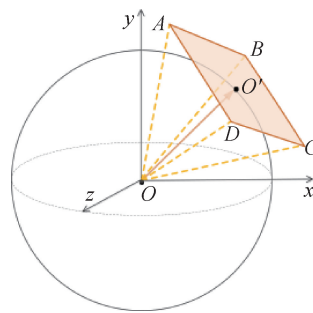


图4 在中心点 O 处的用户沿 OO' 方向观看内容的展示

Fig. 4 Illustration of the viewport image received by the viewers at center O in the OO' direction

的连续性,作者进一步将该过程想象为一段 viewport 大小的 2D 视频在用户面前进行播放。因此该方法首先将 ERP 图像转化为由大量 viewport 图像构成的 2D 视频,最后采用现有 2D 视频质量评估方法计算质量分数。

该类模型相比前两类更加符合用户实际观看特性,但是仍然存在以下问题,具体为:1)模型中选取的 viewport 图像与真实观看情况下用户选取的 viewport 图像不一致问题。现有文献中,并未针对用户在真实观看过程中如何选取 viewport 以及不同用户对于 viewport 的选取是否具有统计规律等问题进行研究。尽管 Xu 等人(2021)借助关键点检测和热力图来确定 viewport 图像,但是 ERP 空间中检测的关键点和热力图未必与原始球面空间中的关键点和热力图一致。如果不一致,将会造成 viewport 提取的误差和最终评价结果的不准确。2)以上模型并没有考虑各个 viewport 之间的关联性。实际观看时,用户不断移动视线获取不同 viewport 内容之后,会将各个 viewport 内容整合进行质量评价。综上,研究人类选取 viewport 以及大脑整合处理 viewport 信息的机理是该类模型取得突破的关键。

1.3 各类 VR 图像质量评价方法的优缺点

表 1 中对各类主流 VR 图像质量评价方法进行

了总结概括和分析,主要包括每种类型方法的主要特性、优点和缺点等信息。从表中可以看到,FR 型 VR 图像质量评价方法的优点是设计思路同传统 FR 方法类似,所以易于理解,且计算速度快。但其存在的主要问题是:1)依赖 PSNR/SSIM 方法和手工特征性能的优劣;2)依赖原始参考信息,但实际应用中参考信息难以获得,所以这极大限制了这类方法应用。相比之下,NR 类型方法则克服了 FR 方法依赖原始参考信息这一缺点,在实际场景中具有更加广泛的应用,因此具有更大的研究价值和更重要的研究意义。由于 ERP 表达空间图像与传统图像类似,均是 2 维平面图像,所以基于 ERP 表达空间的方法同样具有易于理解的优点。但是该表达空间的图像在两极区域具有明显的拉伸形变,这与实际观看空间中图像的特性相差甚远,因此这限制了该类方法的性能。鉴于 ERP 表达空间的这一问题,多种其他的表达空间陆续提出。这些表达空间与实际观看空间特性更加接近,但是仍然不可避免地存在一些特有的投影映射失真。为了解决上述问题,研究者们致力于研究基于实际观看空间的 VR 图像质量评价方法。如前文所述,目前这类方法主要是基于 viewport 的方法。该类方法更好地模拟了人类在单一时刻视觉范围有限的特点,即用户在观看 360°VR 内

表 1 各类 VR 图像质量评价方法的主要特性和优缺点总结

Table 1 Summarization of the characteristics, advantages, and shortcomings of each kind of VR image quality metrics

类型	子类型	主要特性	优点	缺点	代表性文献
FR 型	基于传统 2 维质量评价模型的方法	首先将球面 VR 图投影到 ERP 空间,然后通过 PSNR/SSIM 方法进行质量评价	设计思路清晰,易于理解,计算速度快	依赖 PSNR/SSIM 法,需要参考信息,有较大应用局限	Yu 等人(2015) Zakharchenko 等人(2016) Sun 等人(2017a) Chen 等人(2018)
	基于传统机器学习的方法	在 ERP 空间提取手工特征,然后训练质量评价模型	方法简单,易于理解	依赖手工特征,需要参考信息,有较大应用局限	Madhusudana 和 Soundararajan(2019) 郑珂珂(2017)
NR 型	基于 ERP 表达空间的方法	直接通过对 ERP 表达形式图像进行特征提取和融合得到质量分数	无需参考信息,易于理解	两极区域存在拉伸形变,与实际观看空间差异较大	Kim 等人(2020) Liu 等人(2021)
	基于其他投影表达空间的方法	首先将 VR 图像投影到其他表达空间,然后在新空间进行质量评价	无需参考信息,更接近实际观看空间特性	存在其他映射失真;与实际观看空间差异较大	Zheng 等人(2020) Jiang 等人(2021a)
	基于实际观看空间的方法	通过模拟用户观看 VR 内容时的真实特性,设计评价模型	无需参考信息,更符合用户实际观看特性	模型中选取的 viewport 与真实观看时用户选取的 viewport 不一致;未考虑各 viewport 的关联性	Xu 等人(2021) Sun 等人(2020) Zhou 等人(2021) Sui 等人(2021)

容时的某一时刻,只能观看到视窗范围内的图像内容。然而,如何通过建立数学模型来更科学地模拟人类选取 viewport 的原则以及各 viewport 之间的关联性却是影响该类方法性能的关键因素。

综上所述,各类方法均有一定的优缺点。相比而言,NR 类型方法比 FR 方法具有更大的发展前景和空间。对于前两类 NR 类型方法,如何探索表达空间与真实空间的关系并合理建模是关键问题。对于第 3 类方法,探索实际观看空间特性或用户观看 VR 内容时的视觉特性是性能进一步提升的重要突破口。

2 评价指标

VR IQA 的评价指标采用最常用的皮尔森线性相关系数(Pearson linear correlation coefficient, PLCC)、斯皮尔曼秩相关系数(Spearman rank order correlation coefficient, SRCC)和均方根误差(root mean square error, RMSE)。其中,PLCC 与 RMSE 指标用来评价客观质量评价模型预测分数的准确性,SRCC 指标用来评价客观质量评价算法预测的分数与主观评价分数在单调性方面的一致性(Jiang 等,2021b; Jiang,2018)。PLCC 和 SRCC 的取值范围均是 $[0, 1]$ 。PLCC 和 SRCC 数值越大, RMSE 数值越小,代表相应模型的性能越好。下面给出 3 个性能评价指标的详细计算方法,即

$$P_{\text{PLCC}} = \frac{\sum_{i=1}^R (p_i - \bar{p})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^R (p_i - \bar{p})^2} \sqrt{\sum_{i=1}^R (q_i - \bar{q})^2}} \quad (1)$$

式中, R 表示数据库中包含的测试图像数量, p_i 和 q_i 分别表示第 i 幅图像的主观分数和映射后的客观质量评价分数, $\bar{p}$ 和 $\bar{q}$ 分别表示它们的均值。

$$P_{\text{SRCC}} = 1 - \frac{6 \sum_{i=1}^R d_i^2}{R(R^2 - 1)} \quad (2)$$

式中, d_i 表示主观分数和客观分数的排序差异

$$d_i = X_i - Y_i \quad (3)$$

X_i 和 Y_i 分别表示第 i 幅图像的主观分数和客观分数在数据库所有图像中的排序序号。

$$P_{\text{RMSE}} = \sqrt{\frac{1}{R} \sum_{i=1}^R (p_i - q_i)^2} \quad (4)$$

需要说明的是,在计算 PLCC 和 RMSE 两个指

标之前,需要完成五参数的非线性逻辑映射,目的是将所有的客观质量评价方法的质量评价分数统一到同一范围内(Zhan 等,2017; Li 等,2021)。五参数逻辑映射的计算方式为(Shao 等,2021; Wu 等,2018)

$$f(x) = \tau_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\tau_2(x - \tau_3)}} \right) + \tau_4 x + \tau_5 \quad (5)$$

式中, x 表示客观质量评价算法预测的客观分数, $f(x)$ 表示逻辑映射后的客观分数, τ_i ($i = 1, 2, 3, 4, 5$) 是拟合参数。其中,拟合参数的计算是利用 MATLAB 中提供的非线性最小二乘法函数 nlinfit 实现(Li 等,2020b)。最终的拟合参数实质上是能够使映射后的客观分数和主观分数之间的误差平方和最小的参数组合。对于函数 nlinfit,其输入包括主观分数、相应的客观分数、非线性函数和初始参数,而输出包括拟合参数、残差和雅克比(Jacobian)矩阵。关于逻辑映射的更多信息可以从 Video Quality Experts Group (VQEG) (2000) 获得。

除以上 3 个评价指标外,还有一些学者设计了 D-test、L-test 和 P-test 实验分别用来验证模型区分失真图像和无失真图像的能力、判别失真程度的能力以及从图像对中判别质量好坏的能力。关于这 3 种测试的更多信息可从方玉明等人(2021)的论文中获得。

3 数据库

目前已经有多个针对 VR 图像质量评价的数据库,包括压缩 VR 图像质量数据库(compressed VR image quality database, CVIQD) (Sun 等,2017a)、全方位图像质量评价数据库(omnidirectional image quality assessment database, OIQA) (Duan 等,2018)、压缩 VR 图像质量数据扩展版(compressed VR image quality database, CVIQ) (Sun 等,2020)、图像和视频工程实验室 3 维 VR 图像质量评价数据库(Laboratory for Image Video Engineering Three-Dimensional VR Image Quality Assessment database, LIVE 3D VR IQA) (Chen 等,2020)、多失真视觉显著性质量库(multi-distortions visual attention quality dataset, MVAQD) (Jiang 等,2021a)、印度科学研究院拼接图像质量评价数据库(Indian Institute of Science Stitched Image Quality Assessment database, ISIQA) (Madhusudana 和 Soundararajan,2019)、VR IQA (Sui

等,2021)和宁波大学高动态范围全方位图像库(Ningbo University high dynamic range omnidirectional image database, NBU-HOID)(Cao等,2021)。下面分别对各个数据库进行详细介绍。

1) CVIQD 数据库是由上海交通大学图像通信与信息处理研究所于2017年建立,共提供了5幅原始无失真VR图像和由3种编码方法压缩生成的165幅压缩图像,3种编码方法为JPEG,H.264/AVC和H.265/HEVC。所有原始图像均是采用Insta360 4 K球形VR视频摄像机获得,分辨率均为 $4\,096 \times 2\,048$ 像素,场景包含教学楼、操场、湖泊、雕塑和广场。主观实验采用单激励(single-stimulus, SS)法,主观分数以平均意见分数(mean opinion scores, MOS)形式给出,MOS值越高代表图像质量越好。

2) OIQA 数据库是由上海交通大学图像通信与网络工程研究所于2018年构建。该数据集共包含16幅原始无失真VR图像和320幅对应的失真图像,失真类型包括JPEG压缩失真、JPEG2000压缩失真、高斯模糊和高斯噪声4种类型,每种类型又包含5种失真等级。库中所有图像均采用ERP形式,分辨率范围为 $11\,332 \times 5\,666$ 像素到 $13\,320 \times 6\,660$ 像素不等。主观实验同样采用SS法。测试人员观看图像后,采用10分离散制打分法,分数越高代表图像质量越好。此外,主观实验过程中还记录了测试人员的头动和眼动数据,并在数据库中提供了相关数据。该数据库已经公开,下载链接为:https://mega.nz/#!FqxxRQRR!4Ju2qcmmo6Ced_7nRBXXqAaDejqxjH2uUFnXIeyE2ts。

3) CVIQ 数据库是CVIQD数据库的扩展版本,由同一研究所于2019年建立,共提供了16幅原始图像和3种编码方法处理生成的528幅压缩图像。压缩编码方法和建立CVIQD库时使用的相同。16幅原始图像中有12幅是采用Insta360 4 K球形VR视频摄像机获得,剩余4幅来自于JVET的测试视频。库中图像的分辨率均为 $4\,096 \times 2\,048$ 像素,相比CVIQD数据库,该库包含的图像内容更多样化,包含城镇、风景、人物和物体等。主观实验同样采用SS方法,主观分数采用MOS值形式。该数据库获取链接为:<https://github.com/sunwei925/CVIQDatabase>。

4) LIVE 3D VR IQA 数据库是由美国德克萨斯大学电气与计算机工程系于2019年建立。与其他数据库不同的是,该数据库中提供了双目视图,即对

于每一组待评价图像,都包含左、右两张视图。数据库中共包含15组原始图像。每一组原始无失真图像中分别加入了6种类型失真,包括高斯噪声、高斯模糊、拼接失真、下采样失真、VP9压缩失真和H.265压缩失真。对于每种类型的失真,作者又设置了不同的失真等级,最终产生了450组失真图像。库中所有图像均采用Insta360 Pro摄像机拍摄获得,场景涉及白天/夜间、晴天/阴天以及室内/室外等。主观实验过程中,测试图像的分辨率由原始的 $7\,680 \times 3\,840$ 像素下采样到 $4\,096 \times 2\,048$ 像素,采用0~100的连续打分制,主观分数以差分平均意见分数(difference mean opinion score, DMOS)形式给出,其数值越小代表图像质量越好。该数据库也提供了眼动数据。数据库获取链接为:<http://live.ece.utexas.edu/research/VR3D/index.html>。

5) MVAQD 数据库是由宁波大学信息科学与工程学院于2019年建立。数据库共包含315幅 360° 图像,其中,15幅为原始无失真图像,300幅为失真图像。失真图像是通过将5种类型失真分别加入各原始无失真图像获得,包括JPEG压缩失真、JP2K压缩失真、HEVC帧内编码失真、白噪声和高斯模糊。每种类型又包含4种失真等级,形成20种组合。主观实验采用5分制法。需要注意的是,尽管该数据库称为多失真数据库,但是每一幅图像中仅仅包含一类失真,命名为多失真的原因是数据库中一共包含5类失真。数据库获取链接:<https://github.com/Jianghao2019/MVAQD>。

6) ISIIQA 数据库是由印度科学研究院建立,共提供了26组拼接前子图像和拼接后的264幅拼接图像。每一幅拼接图像都是由4~5幅具有重叠区域的子图像拼接而成。该数据库中的失真只包含拼接失真。拼接过程一般包括图5中展示的几个过程,由于其中的曝光补偿和图像融合阶段通常会默认进行色彩校正,所以在建立该数据库时没有单独设置色彩校正过程。图像内容包含有建筑物、花园、室内及公共场所。主观实验时采用连续百分制打分法,主观分数采用MOS值形式。该数据库获取链接为:<https://pavanm.github.io/stitched-qa/>。

7) VR IQA 数据库是由江西财经大学信息管理学院于2020年建立。数据库中一共提供了72幅VR图像,包括36幅原始无失真图像和相应的36幅失真图像。原始图像中有24幅是利用Insta360 Pro 2

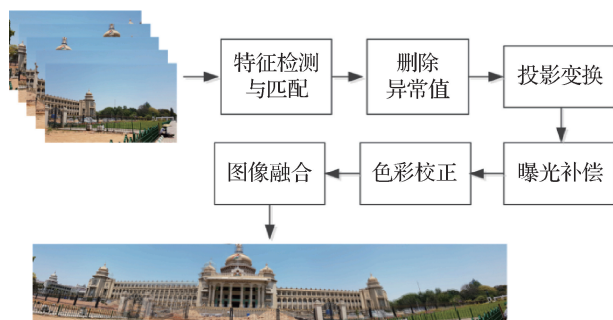


图5 拼接技术流程

Fig. 5 The process of the stitched technology

摄像机拍摄获得,这些图像中加入了拼接失真。其余的12幅图像是从网络上下载而来,其中加入了H.265压缩失真。库中所有图像的分辨率均为 $7\,860 \times 3\,840$ 。主观实验采用1~5分的连续打分制,分数越高代表图像质量越好。实验过程中同时采集了测试人员的头动数据和眼动数据。此外,为研究用户观看起始点和观看时间对评分的影响,作者在实验中设置了两个起始点和两种观看时间。该数据库的获取链接为: <https://github.com/xiangjieSui/img2video>。

8) NBU-HOID 数据库是由宁波大学信息科学与工程学院于2021年建立。该数据库是第1个研究高动态范围VR图像质量评价的数据库。相比其他数据库里的图像,高动态范围图像具有更宽的动态范围、更高的对比度和更多的图像细节。库中一共提供了16组原始无失真VR图像和320幅失真的多曝光VR图像。其中的失真由JPEG XT编码和色调映射引入。主观实验过程中采用9分制评分法,主观分数采用MOS值形式。该数据集可以从以下链接获得: <https://github.com/caoliuyan/NBU-HOID>。

9) 除以上VR图像质量评价数据库外,一些科研机构和院校也相继建立了一些VR视频质量评价数据库,例如上海交通大学图像通信与网络工程研究所建立的沉浸式视频质量评价(immersive video quality assessment database, IVQAD)数据库(Duan等,2017)、北京航空航天大学电子信息工程学院建立的全方位视频视觉质量评价数据库(visual quality assessment dataset of omnidirectional video, VQA-OV)(Li等,2018)和全景视频数据库(Xu等,2017)等。本文仅针对静态VR图像的客观质量评价进行综

述,所以不对动态VR视频质量评价数据库进行介绍和总结。

表2中对8个公开的VR图像质量评价数据库的主要信息进行了详细总结。

4 算法性能对比

本文测试了现有主流的FR型和NR型图像质量评价方法在各个VR数据集上的性能。其中,FR型方法包括PSNR、SSIM、S-PSNR和CPP-PSNR,NR型方法包括盲图像质量评价方法(blind image quality index, BIQI)(Moorthy和Bovik,2010)、基于失真类型鉴别的图像真实性和完整性评价方法(distortion identification-based image verity and integrity evaluation, DIIVINE)(Moorthy和Bovik,2011)、综合局部自然图像质量评价器的方法(integrated local natural image quality evaluator, IL-NIQE)(Zhang等,2015)、基于多幅伪参考图像的盲评价(blind multiple pseudo reference images, BMPRI)方法(Min等,2018)、基于质量鉴别图像对推断的质量评价方法(quality-discriminable image pairs inferred quality index, dipIQ)(Ma等,2017)和多任务端到端优化的深度神经网络(multi-task end-to-end optimized deep neural network, MEON)(Ma等,2018)。

表3展示了相应的实验结果,并对各个库上的最优性能进行了加粗显示。其中,PLCC和SRCC数值越高,代表相应算法的预测准确性越高,单调性与主观结果更符合。需要注意的是,由于ISIQA数据库中未提供与失真图像相同分辨率的原始无失真图像,这使得FR型方法的性能无法测试,因此表中相应位置处用“-”表示。从表中的实验结果可以得出以下结论:1)各个方法在不同数据库上的性能表现有所差异,没有一种方法能够在所有数据集上都能获取比较理想的性能,这说明算法的鲁棒性有待提高;2)FR型方法普遍比NR型方法性能好,主要原因是该类型方法使用了原始信息作为参考。但是该类型方法在实际应用中相比NR方法具有更大局限性。3)现有主流质量评价方法在VR IQA数据集上的性能普遍较差,最高的PLCC和SRCC数值仅为0.3左右。其主要原因是数据集中的拼接失真与传统类型失真特性相差甚远,现有方法无法很好地进行评价。

表 2 8 个公开的 VR 图像质量评价数据库的主要信息总结
Table 2 Summary of the main information of eight public databases for VR IQA

数据库	建立院校	年份	原始图/失真图数量/幅	失真类型	MOS/DMOS	数据库链接
CVIQD	上海交通大学	2017	5/165	JPEG 压缩 H. 264/AVC 压缩 H. 265/HEVC 压缩	MOS	包含在 CVIQ 数据库中
OIQA	上海交通大学	2018	16/320	JPEG 压缩 JPEG2000 压缩 高斯模糊 高斯噪声	MOS	https://mega.nz/#!FqxxRQRR!4Ju2qcmmo6Ced_7nRBXXqAaDejqxjH2uUfnXleyE2ts
CVIQ	上海交通大学	2019	16/528	JPEG 压缩 H. 264/AVC 压缩 H. 265/HEVC 压缩 高斯噪声 高斯模糊	MOS	https://github.com/sunwei925/CVIQDatabase
LIVE 3D VR IQA	美国德克萨斯大学	2019	15/450 (双目视图)	拼接失真 下采样 VP9 压缩 H. 265 压缩	DMOS	http://live.ece.utexas.edu/research/VR3D/index.html
MVAQD	宁波大学	2019	15/300	JPEG 压缩 JP2K 压缩 HEVC 帧内编码 白噪声 高斯模糊	MOS	https://github.com/Jianghao2019/MVAQD
ISIQA	印度科学研究院	2019	26 组/264	拼接失真	MOS	https://pavanm.github.io/stitched-qa/
VR IQA	江西财经大学	2020	36/36	拼接失真 H. 265 压缩	MOS	https://github.com/xiangjieSui/img2video
HOID	宁波大学	2021	16/320	JPEG XT 编码和色调映射	MOS	https://github.com/caoliuyan/NBU-HOID

表 3 现有主流的全参考和无参考质量评价方法在各 VR 数据集上的性能

Table 3 Performance of existing state-of-the-art FR and NR quality assessment metrics on seven VR databases

数据库	性能指标	FR				NR					
		PSNR	SSIM	S-PSNR	CPP-PSNR	BIQI	DIIVINE	IL-NIQE	BMPRI	dipIQ	MEON
CVIQD	PLCC	0.813	0.745	0.782	0.703	0.565	0.597	0.646	0.619	0.625	0.655
	SRCC	0.786	0.717	0.754	0.701	0.545	0.572	0.632	0.604	0.621	0.649
OIQA	PLCC	0.492	0.856	0.716	0.707	0.473	0.616	0.597	0.431	0.701	0.749
	SRCC	0.497	0.880	0.712	0.703	0.414	0.581	0.573	0.338	0.691	0.717
CVIQ	PLCC	0.786	0.897	0.785	0.779	0.548	0.438	0.626	0.627	0.706	0.665
	SRCC	0.757	0.885	0.761	0.754	0.442	0.413	0.534	0.621	0.623	0.567
LIVE 3D VR IQA	PLCC	0.651	0.718	0.662	0.604	0.527	0.502	0.553	0.742	0.722	0.702
	SRCC	0.560	0.694	0.608	0.527	0.468	0.443	0.484	0.740	0.608	0.629
MVAQD	PLCC	0.683	0.758	0.703	0.721	0.432	0.583	0.321	0.601	0.664	0.570
	SRCC	0.671	0.743	0.714	0.682	0.495	0.596	0.304	0.602	0.652	0.585
ISIQA	PLCC	—	—	—	—	0.278	0.303	0.338	0.392	0.814	0.805
	SRCC	—	—	—	—	0.293	0.501	0.285	0.404	0.799	0.770
VR IQA	PLCC	0.239	0.051	0.224	0.206	0.104	0.108	0.304	0.292	0.312	0.342
	SRCC	0.103	0.049	0.107	0.108	0.103	0.104	0.210	0.208	0.227	0.353

注:加粗字体表示每行最优结果,“—”表示 FR 类型方法无法在没有提供原始参考图的 ISIQA 数据库进行测试。

5 VR IQA 模型的应用

VR IQA 模型的应用主要包括 VR 相关技术的参数优化、评价与选择。其中,VR 相关技术是指从图像采集、处理、传输和存储到用户端的投影和显示整个环节中的各项技术。参数优化是指寻找某一技术中参数的最佳值,具体方法为:将参数进行迭代遍历生成不同参数设置下的图像,然后将 VR IQA 模型作为“裁判”,裁决出最高质量的图像,则其对应的参数被认为是最优参数。例如,图像拼接技术是生成 VR 图像的关键技术,该技术中往往包含多个参数,且不同参数下拼接的图像效果大不相同。在这种情况下,如何来评价哪种参数设置下拼接效果好则是关键问题。显然,通过人眼来判断最可靠。但是,参数的可选择范围通常较大,所以这种方法成本太高,费时费力,实用性较差。而一个优秀的 VR IQA 模型则能够充当“人眼”,对不同参数设置下拼接图像的质量进行评价,从而判断出哪组参数下拼接效果最好,则该参数被认为是最优参数。另外,对于同一种需求,不同的学者会研究出不同的技术方法。同样以拼接技术为例,目前的方法多种多样(Xu 和 Mulligan,2010;Chang 等,2014;Nie 等,2021;Yuan 等,2021),且每个方法可能都有各自的适用场景。在实际应用中,面对待处理的图像,如何选择最适合的方法是另一个关键问题。类似地,对于一组待拼接图像,可以采用现有拼接算法依次对图像进行拼接,然后采用 VR IQA 模型来裁断出哪种算法拼接效果最好,从而确定最终选择的拼接技术。

6 结 语

6.1 当前研究总结

VR 技术凭借在医疗、军事、娱乐和教育等众多领域的广阔应用前景引起了各界密切关注。其中对于 VR 图像客观质量评价的研究已经成为图像处理领域热门的前沿性研究方向。本文主要对目前的 VR 图像客观质量评价相关研究进行综述。现有模型包含 FR 型和 NR 型两类。对于 FR 型的方法,它们依赖于原始无失真的高清 VR 图像,这极大限制了在实际场景中的应用。对于 NR 类型方法,本文根据特征表达空间的不同,将其划分为 3 类,包括基

于 ERP 表达空间的模型、基于其他投影表达空间的模型和基于实际观看空间的模型。前两类模型在将原始球面 VR 图像投影到其他空间进行表达时,会在图像中引入一定的投影失真,从而改变图像结构和图像中失真的分布特性,造成评价的不准确性。对于第 3 类模型,其相对于前面两类模型是个更好的思路,具有更大的发展空间,因为客观质量评价模型的最终目标是代替人眼来进行质量评估,所以研究用户在实际观看空间的特性是解决问题的突破口。目前该类型的方法均是基于 viewport 的方法,它们对于 viewport 的选取与融合等处理并没有真正地模拟人类的视觉信息处理特性,这造成了质量预测性能方面的“天花板”。综上,对于 VR 图像客观质量评价的研究仍然具有相当大的发展空间。

6.2 未来发展方向

虽然 VR IQA 的研究已经逐渐成为 IQA 领域的研究热点,但是相比该领域其他较为成熟的研究方向,它仍处于起步阶段,未来仍然有许多亟需发展的方向,主要包括以下几点:

1) 研究基于小样本学习的 VR IQA 模型。数据量不足一直是 IQA 领域中各个研究方向普遍存在的问题。尤其对于 VR 图像质量评价,其主观评价相比传统 IQA 的主观评价难度更大、成本更高,因而数据量问题在 VR 图像的客观图像质量评价中显得更为突出。目前,最大的 VR 图像数据库(CVIQ 数据库)也仅仅包含 528 幅图像,这对于深度学习而言远远不够。面对该问题,除了尽可能扩大数据库之外,研究如何在小样本情况下进行有效的质量评估更是关键问题,这也是未来研究中的重中之重。

2) 研究用户实际观看特性并建模。佩戴了头戴设备的用户在实际观看 360° VR 内容时的观看特性与观看 2 维平面图的特性具有巨大差异。目前的研究中对观看特性的研究与建模严重不足。尽管 Sui 等人(2021)研究了不同观看起始点、观看路径和观看时间对质量评价的影响,但是缺少对用户观看特性的统计性研究和更深入分析。例如面对一个新的 VR 内容,用户选择起始点和观看路径的依据、不同用户的观看特性是否具有一定的统计规律以及如何对相关特性进行合理建模等问题都极大影响 VR 图像客观质量评价的研究,因此这也是重要研究方向之一。

3) 研究面向视觉体验舒适度的 VR IQA 模型。

失真是造成视觉质量下降的重要因素,因此量化图像失真程度对于传统2维图像质量评价具有重要意义。然而,用户在体验VR应用时更注重的是体验舒适度而不仅仅是失真的多少。图像中的失真是造成用户在生理和心理方面不舒适的重要原因之一,从而降低用户对视觉质量的感知,这与传统2维图像大不相同。因此,研究失真与体验舒适度的关系并进行合理建模是另一关键问题。目前,关于图像舒适度的研究主要集中于普通立体3D图像,而对于360°VR图像舒适度评价的相对较少。

4)构建多维复合失真的VR图像数据集。在现实环境中,VR图像在复杂的采集、拼接、编码和传输等过程中引入的失真往往呈现多维复合特性。现有的VR图像客观质量评价模型并没有考虑多种失真对质量的共同作用,而阻碍该方向发展的一个重要原因是缺少相应的模型测试数据集。现有数据集存在以下问题:(1)单幅图像中只含有一类失真,而现实中的图像往往同时包含多类失真;(2)大多数数据集仅包含模糊、压缩失真、噪声等传统类型失真。仅有少数数据库中包含VR图像中特有的拼接失真和投影失真等,而它们的特性与传统失真特性大不相同,所以需要在数据集中加入这些特有失真。(3)现有数据集中加入的传统类型失真均是通过计算机模拟生成的失真,而真实失真往往比模拟失真更加复杂,且两者特性差异较大,这会间接限制相应质量评价模型的可扩展性和实际应用。

参考文献(References)

- Athar S and Wang Z. 2019. A comprehensive performance evaluation of image quality assessment algorithms. *IEEE Access*, 7: 140030-140070 [DOI: 10.1109/ACCESS.2019.2943319]
- Cao F. 2019. Overview of VR technology research and development at home and abroad. *China Science and Technology Information*, (5): 36-37 (曹凡. 2019. 国内外VR技术研发现状综述. *中国科技信息*, (5): 36-37 [DOI: 10.3969/j.issn.1001-8972.2019.05.010])
- Cao L Y, Jiang G Y, Jiang Z D, Yu M, Qi Y B and Ho Y S. 2021. Quality measurement for high dynamic range omnidirectional image systems. *IEEE Transactions on Instrumentation and Measurement*, 70: #5012415 [DOI: 10.1109/TIM.2021.3093940]
- Chang C H, Sato Y and Chuang Y Y. 2014. Shape-preserving half-projective warps for image stitching//*Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Columbus, USA; IEEE: 3254-3261 [DOI: 10.1109/CVPR.2014.422]
- Chen M X, Jin Y Z, Goodall T, Yu X X and Bovik A C. 2020. Study of 3D virtual reality picture quality. *IEEE Journal of Selected Topics in Signal Processing*, 14(1): 89-102 [DOI: 10.1109/JSTSP.2019.2956408]
- Chen S J, Zhang Y X, Li Y M, Chen Z Z and Wang Z. 2018. Spherical structural similarity index for objective omnidirectional video quality assessment//*Proceedings of 2018 IEEE International Conference on Multimedia and Expo (ICME)*. San Diego, USA; IEEE: 1-6 [DOI: 10.1109/ICME.2018.8486584]
- Duan H Y, Zhai G T, Min X K, Zhu Y C, Fang Y and Yang X K. 2018. Perceptual quality assessment of omnidirectional images//*2018 IEEE International Symposium on Circuits and Systems (ISCAS)*. Florence, Italy; IEEE: 1-5 [DOI: 10.1109/ISCAS.2018.8351786]
- Duan H Y, Zhai G T, Yang X K, Li D and Zhu W H. 2017. IVQAD 2017: an immersive video quality assessment database//*Proceedings of 2017 International Conference on Systems, Signals and Image Processing (IWSSIP)*. Poznan, Poland; IEEE: 1-5 [DOI: 10.1109/IWSSIP.2017.7965610]
- Fang Y M, Sui X J, Yan J B, Liu X L and Huang L P. 2021. Progress in no-reference image quality assessment. *Journal of Image and Graphics*, 26(2): 265-286 (方玉明, 眭相杰, 鄢杰斌, 刘学林, 黄丽萍. 2021. 无参考图像质量评价研究进展. *中国图象图形学报*, 26(2): 265-286) [DOI: 10.11834/jig.200274]
- Fang Y M, Yan J B, Du R G, Zuo Y F, Wen W Y, Zeng Y and Li L D. 2020. Blind quality assessment for tone-mapped images by analysis of gradient and chromatic statistics. *IEEE Transactions on Multimedia*, 23: 955-966 [DOI: 10.1109/TMM.2020.2991528]
- Fang Y M, Yan J B, Li L D, Wu J J and Lin W S. 2018. No reference quality assessment for screen content images with both local and global feature representation. *IEEE Transactions on Image Processing*, 27(4): 1600-1610 [DOI: 10.1109/TIP.2017.2781307]
- Fang Y M, Yan J B, Liu J Y, Wang S Q, Li Q H and Guo Z M. 2017. Objective quality assessment of screen content images by uncertainty weighting. *IEEE Transactions on Image Processing*, 26(4): 2016-2027 [DOI: 10.1109/TIP.2017.2669840]
- Gu K. 2015. Research on Perceptual and Statistical Image Quality Assessment and Its Application. Shanghai: Shanghai Jiao Tong University (顾颀. 2015. 基于感知和统计模型的图像质量评价技术及应用研究. 上海: 上海交通大学)
- Jiang G Y, Song H, Yu M, Song Y and Peng Z J. 2018. Blind tone-mapped image quality assessment based on brightest/darkest regions, naturalness and aesthetics. *IEEE Access*, 6: 2231-2240 [DOI: 10.1109/ACCESS.2017.2782320]
- Jiang H, Jiang G Y, Yu M, Zhang Y, Yang Y, Peng Z J, Chen F and Zhang Q B. 2021a. Cubemap-based perception-driven blind quality assessment for 360-degree images. *IEEE Transactions on Image Processing*, 30: 2364-2377 [DOI: 10.1109/TIP.2021.3052073]

- Jiang Q P, Peng Z Y, Yue G H, Li H and Shao F. 2021b. No-reference image contrast evaluation by generating bidirectional pseudo references. *IEEE Transactions on Industrial Informatics*, 17 (9): 6062-6072 [DOI: 10.1109/TII.2020.3035448]
- Kim H G, Lim H T and Ro Y M. 2020. Deep virtual reality image quality assessment with human perception guider for omnidirectional image. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(4): 917-928 [DOI: 10.1109/TCSVT.2019.2898732]
- Li C, Xu M, Du X Z and Wang Z L. 2018. Bridge the gap between VQA and human behavior on omnidirectional video: a large-scale dataset and a deep learning model//*Proceedings of the 26th ACM Multimedia Conference*. Seoul, Korea (South): ACM: 932-940 [DOI: 10.1145/3240508.3240581]
- Li J, Zhao Y F, Ye W H, Yu K W and Ge S M. 2020a. Attentive deep stitching and quality assessment for 360° omnidirectional images. *IEEE Journal of Selected Topics in Signal Processing*, 14(1): 209-221 [DOI: 10.1109/JSTSP.2019.2953950]
- Li L D, Zhou Y, Gu K, Yang Y Z and Fang Y M. 2020b. Blind realistic blur assessment based on discrepancy learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(11): 3859-3869 [DOI: 10.1109/TCSVT.2019.2947450]
- Li L D, Zhou Y, Wu J J, Li F and Shi G M. 2021. Quality index for view synthesis by measuring instance degradation and global appearance. *IEEE Transactions on Multimedia*, 23: 320-332 [DOI: 10.1109/TMM.2020.2980185]
- Lim H T, Kim H G and Ra Y M. 2018. VR IQA NET: deep virtual reality image quality assessment using adversarial learning//*Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing*. Calgary, Canada: IEEE: 6737-6741 [DOI: 10.1109/ICASSP.2018.8461317]
- Liu Y, Yu H W, Huang B Q, Yue G H and Song B Y. 2021. Blind omnidirectional image quality assessment based on structure and natural features. *IEEE Transactions on Instrumentation and Measurement*, 70: #5014011 [DOI: 10.1109/TIM.2021.3102691]
- Ma K D, Liu W T, Liu T L, Wang Z and Tao D C. 2017. DipIQ: blind image quality assessment by learning-to-rank discriminable image pairs. *IEEE Transactions on Image Processing*, 26(8): 3951-3964 [DOI: 10.1109/TIP.2017.2708503]
- Ma K D, Liu W T, Zhang K, Duanmu Z F, Wang Z and Zuo W M. 2018. End-to-end blind image quality assessment using deep neural networks. *IEEE Transactions on Image Processing*, 27(3): 1202-1213 [DOI: 10.1109/TIP.2017.2774045]
- Madhusudana P C and Soundararajan R. 2019. Subjective and objective quality assessment of stitched images for virtual reality. *IEEE Transactions on Image Processing*, 28(11): 5620-5635 [DOI: 10.1109/TIP.2019.2921858]
- Min X K, Zhai G T, Gu K, Liu Y T and Yang X K. 2018. Blind image quality estimation via distortion aggravation. *IEEE Transactions on Broadcasting*, 64(2): 508-517 [DOI: 10.1109/TBC.2018.2816783]
- Min X K, Zhai G T, Zhou J T, Farias M C Q and Bovik A C. 2020. Study of subjective and objective quality assessment of audio-visual signals. *IEEE Transactions on Image Processing*, 29: 6054-6068 [DOI: 10.1109/TIP.2020.2988148]
- Moorthy A K and Bovik A C. 2010. A two-step framework for constructing blind image quality indices. *IEEE Signal Processing Letters*, 17(5): 513-516 [DOI: 10.1109/LSP.2010.2043888]
- Moorthy A K and Bovik A C. 2011. Blind image quality assessment: from natural scene statistics to perceptual quality. *IEEE Transactions on Image Processing*, 20(12): 3350-3364 [DOI: 10.1109/TIP.2011.2147325]
- Nie L, Lin C Y, Liao K, Liu S C and Zhao Y. 2021. Unsupervised deep image stitching: reconstructing stitched features to images. *IEEE Transactions on Image Processing*, 30: 6184-6197 [DOI: 10.1109/TIP.2021.3092828]
- Niu Y Z, Zhong Y N, Guo W Z, Shi Y Q and Chen P K. 2019. 2D and 3D image quality assessment: a survey of metrics and challenges. *IEEE Access*, 7: 782-801 [DOI: 10.1109/ACCESS.2018.2885818]
- Shao F, Fu Z Q, Jiang Q P, Jiang G Y and Ho Y S. 2021. Transformation-aware similarity measurement for image retargeting quality assessment via bidirectional rewarping. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(5): 3053-3067 [DOI: 10.1109/TSMC.2019.2917496]
- Sui X J, Ma K D, Yao Y R and Fang Y M. 2021. Perceptual quality assessment of omnidirectional images as moving camera videos. *IEEE Transactions on Visualization and Computer Graphics*: #3050888 [DOI: 10.1109/TVCG.2021.3050888]
- Sun W, Gu K, Zhai G T, Ma S W, Lin W S and Calle P L. 2017a. CVIQD: subjective quality evaluation of compressed virtual reality images//*Proceedings of 2017 IEEE International Conference on Image Processing (ICIP)*. Beijing, China: IEEE: 3450-3454 [DOI: 10.1109/ICIP.2017.8296923]
- Sun W, Min X K, Zhai G T, Gu K, Duan H Y and Ma S W. 2020. MC360IQA: a multi-channel CNN for blind 360-degree image quality assessment. *IEEE Journal of Selected Topics in Signal Processing*, 14(1): 64-77 [DOI: 10.1109/JSTSP.2019.2955024]
- Sun Y L, Lu A and Yu L. 2017b. Weighted-to-spherically-uniform quality evaluation for omnidirectional video. *IEEE Signal Processing Letters*, 24(9): 1408-1412 [DOI: 10.1109/LSP.2017.2720693]
- Video Quality Experts Group (VQEG). 2000. Final report from the video quality experts group on the validation of objective models of video quality assessment [EB/OL]. [2020-07-08]. http://videoclarity.com/PDF/COM-80E_final_report.pdf
- Wang S Q, Gu K, Ma S W and Gao W. 2016. Joint chroma downsampling and upsampling for screen content image. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(9): 1595-1609 [DOI: 10.1109/TCSVT.2015.2461891]

- Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600-612 [DOI: 10.1109/TIP.2003.819861]
- 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]
- Wu J J. 2014. Image Information Perception and Quality Assessment based on the Human Visual System. Xi'an: Xidian University (吴金建. 2014. 基于人类视觉系统的图像信息感知和图像质量评价. 西安: 西安电子科技大学)
- Wu J J, Ma J P, Liang F H, Dong W S, Shi G M and Lin W S. 2020. End-to-end blind image quality prediction with cascaded deep neural network. *IEEE Transactions on Image Processing*, 29: 7414-7426 [DOI: 10.1109/TIP.2020.3002478]
- Wu Q B, Li H L, Meng F M and Ngan K N. 2018. Toward a blind quality metric for temporally distorted streaming video. *IEEE Transactions on Broadcasting*, 64(2): 367-378 [DOI: 10.1109/TBC.2017.2786023]
- Xu W and Mulligan J. 2010. Performance evaluation of color correction approaches for automatic multi-view image and video stitching//*Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, USA: IEEE: 263-270 [DOI: 10.1109/CVPR.2010.5540202]
- Xu J H, Zhou W and Chen Z B. 2021. Blind omnidirectional image quality assessment with viewport oriented graph convolutional networks. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(5): 1724-1737 [DOI: 10.1109/TCSVT.2020.3015186]
- Xu M, Li C, Liu Y F, Deng X and Lu J X. 2017. A subjective visual quality assessment method of panoramic videos//*Proceedings of 2017 IEEE International Conference on Multimedia and Expo (ICME)*. Hong Kong, China: IEEE: 517-522 [DOI: 10.1109/ICME.2017.8019351]
- Yang J C, Liu T L, Jiang B, Song H B and Lu W. 2018. 3D panoramic virtual reality video quality assessment based on 3D convolutional neural networks. *IEEE Access*, 6: 38669-38682 [DOI: 10.1109/ACCESS.2018.2854922]
- Yang X H, Li F and Liu H T. 2019. A survey of DNN methods for blind image quality assessment. *IEEE Access*, 7: 123788-123806 [DOI: 10.1109/ACCESS.2019.2938900]
- Yuan Y T, Fang F M and Zhang G X. 2021. Superpixel-based seamless image stitching for UAV images. *IEEE Transactions on Geoscience and Remote Sensing*, 59(2): 1565-1576 [DOI: 10.1109/TGRS.2020.2999404]
- Yu M, Lakshman H and Girod B. 2015. A framework to evaluate omnidirectional video coding schemes//*Proceedings of 2015 IEEE International Symposium on Mixed and Augmented Reality*. Fukuoka, Japan: IEEE: 31-36 [DOI: 10.1109/ISMAR.2015.12]
- Yue G H, Hou C P and Zhou T W. 2019. Blind quality assessment of tone-mapped images considering colorfulness, naturalness, and structure. *IEEE Transactions on Industrial Electronics*, 66(5): 3784-3793 [DOI: 10.1109/TIE.2018.2851984]
- Zakharchenko V, Choi K P and Park J H. 2016. Quality metric for spherical panoramic video//*Proceedings Volume 9970, Optics and Photonics for Information Processing X*. San Diego, USA: SPIE: 1-9 [DOI: 10.1117/12.2235885]
- Zhai G T and Min X K. 2020. Perceptual image quality assessment: a survey. *Science China Information Sciences*, 63(11): #211301 [DOI: 10.1007/s11432-019-2757-1]
- Zhan Y B, Zhang R and Wu Q. 2017. A structural variation classification model for image quality assessment. *IEEE Transactions on Multimedia*, 19(8): 1837-1847 [DOI: 10.1109/TMM.2017.2689923]
- Zhang L, Zhang L and Bovik A C. 2015. A feature-enriched completely blind image quality evaluator. *IEEE Transactions on Image Processing*, 24(8): 2579-2591 [DOI: 10.1109/TIP.2015.2426416]
- Zhang W X, Ma K D, Yan J, Deng D X and Wang Z. 2020. Blind image quality assessment using a deep bilinear convolutional neural network. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(1): 36-47 [DOI: 10.1109/TCSVT.2018.2886771]
- Zheng F. 2020. Design and Research on Real-Time AI Reporting of VR E-Sports News. Harbin: Harbin Institute of Technology (郑菲. 2020. 虚拟现实电竞新闻实时智能报道的设计研究. 哈尔滨: 哈尔滨工业大学)
- Zheng K K. 2017. Research on Stitching Quality Assessment Method of Virtual Reality Panoramic Video. Xiamen: Xiamen University (郑珂珂. 2017. VR全景视频的拼接质量评价方法研究. 厦门: 厦门大学)
- Zheng X L, Jiang G Y, Yu M and Jiang H. 2020. Segmented spherical projection-based blind omnidirectional image quality assessment. *IEEE Access*, 8: 31647-31659 [DOI: 10.1109/ACCESS.2020.2972158]
- Zhou Y. 2019. Study on the Objective Quality Assessment Metrics of Virtual View Synthesis. Xuzhou: China University of Mining and Technology (周玉. 2019. 面向虚拟视角合成的客观质量评价方法研究. 徐州: 中国矿业大学)
- Zhou Y, Li L D, Wang S Q, Wu J J, Fang Y M and Gao X B. 2019. No-reference quality assessment for view synthesis using DoG-based edge statistics and texture naturalness. *IEEE Transactions on Image Processing*, 28(9): 4566-4579 [DOI: 10.1109/TIP.2019.2912463]
- Zhou Y, Li L D, Wu J J, Gu K, Dong W S and Shi G M. 2018. Blind quality index for multiply distorted images using biorder structure degradation and nonlocal statistics. *IEEE Transactions on Multimedia*, 20(11): 3019-3032 [DOI: 10.1109/TMM.2018.2829607]
- Zhou Y, Sun Y J, Li L D, Gu K and Fang Y M. 2021. Omnidirectional

image quality assessment by distortion discrimination assisted multi-stream network. IEEE Transactions on Circuits and Systems for Video Technology, 32: 1767-1777 [DOI: 10. 1109/TCSVT. 2021. 3081162]

Zhou Z, Zhou Y and Xiao J J. 2015. Survey on augmented virtual environment and augmented reality. Scientia Sinica Informationis, 45(2): 157-180 (周忠, 周颐, 肖江剑. 2015. 虚拟现实增强技术综述. 中国科学: 信息科学, 45(2): 157-180) [DOI: 10. 1360/N112014-00076]

作者简介



周玉, 1990 年生, 女, 讲师, 主要研究方向为图像质量评价。

E-mail: zhouy@cumt.edu.cn



李雷达, 通信作者, 男, 教授, 主要研究方向为人工智能、计算机视觉、视觉质量评估。

E-mail: ldli@xidian.edu.cn

汪一, 男, 助教, 主要研究方向为图像质量评价。

E-mail: wangyi1567@126.com

高陈强, 男, 教授, 主要研究方向为计算机视觉与机器学习、图像处理。E-mail: gaocq@cqupt.edu.cn

卢兆林, 男, 副教授, 主要研究方向为计算机视觉、图像处理。

E-mail: zhaolinlu@cumt.edu.cn