



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

2022
05
VOL.27

ISSN1006-8961
CN11-3758/TB



低质图像增强



第27卷第5期（总第313期）

2022年5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

《中国图象图形学报》低质图像增强专刊简介

李波, 朱策, 操晓春, 山世光, 石争浩, 潘金山, 刘家瑛, 胡海苗, 任文琦, 陈秀妍 1335

综述

水下光学图像重建方法研究进展

王柯俨, 黄诗芮, 李云松 1337

单幅图像去雨数据集和深度学习算法的联合评估与展望

胡明娣, 吴怡, 宋尧, 杨静冰, 张瑞芳, 王红, 孟德宇 1359

低光照图像增强算法综述

马龙, 马腾宇, 刘日升 1392

图像与视频质量评价综述

程茹秋, 余烨, 石岱宗, 蔡文 1410

图像质量评价研究综述——从失真的角度

鄢杰斌, 方玉明, 刘学林 1430

数据集论文

面向真实水下图像增强的质量评价数据集

顾约瑟, 姜求平, 邵枫, 高伟 1467

图像去雾去雨

适合跨域目标检测的雾霾图像增强

郭强, 浦世亮, 张世峰, 李波 1481

沙尘图像色彩恢复及增强卷积神经网络

石争浩, 刘春月, 任文琦, 都双丽, 赵明华 1493

融合3D注意力和Transformer的图像去雨网络

王美华, 柯凡晖, 梁云, 范衡, 廖磊 1509

联合形态学滤波和卷积稀疏编码的图像去雨

黄淑英, 许亚婷, 杨勇, 管巨伟 1522

多尺度渐进式残差网络的图像去雨

卢贝, 盖杉 1537

低照图像增强

改进融合策略下透明度引导的逆光图像增强

赵明华, 程丹妮, 都双丽, 胡静, 石程, 石争浩 1554

融合注意力机制和上下文信息的微光图像增强

赵兴运, 孙帮勇 1565

多残差联合学习的水下图像增强

陈龙, 丁丹丹 1577

零参考样本下的逆光图像深度学习增强方法

王知音, 张二虎, 石争浩, 段敬红 1589

图像超分辨

跨尺度耦合的连续比例因子图像超分辨率

吴瀚霖, 李宛谕, 张立保 1604

用于单幅模糊图像超分辨的Transformer融合网络

刘花成, 任文琦, 王蕊, 操晓春 1616

图像修复

全细节增强多曝光图像融合

陈彬, 谭新成, 伍世虔 1632

基于卷积自编码生成式对抗网络的高分辨率破损图像修复

侯向丹, 刘昊然, 刘洪普 1645

融合参考先验与生成先验的老照片修复

刘继鑫, 陈瑞, 安仕鹏 1657

区域注意力机制引导的双路虹膜补全

张志礼, 张慧, 王甲, 夏玉峰, 刘亮, 李佩佩, 何召锋 1669

融合注意力机制的模糊图像多尺度复原

陈紫柠, 张宏怡, 曾念寅, 李寒 1682

低光照图像增强算法综述
(第1392页)融合3D注意力和Transformer
的图像去雨网络(第1509页)融合注意力机制的模糊图像
多尺度复原(第1682页)

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



The review of low-light image enhancement(P1392)



3D attention and Transformer based single image deraining network(P1509)



Attention mechanism embedded multi-scale restoration method for blurred image(P1682)

Review

- An optical reconstruction based underwater image analysis
Wang Keyan, Huang Shirui, Li Yunsong 1337
- The integrated evaluation and review of single image rain removal based datasets and deep learning methods
Hu Mingdi, Wu Yi, Song Yao, Yang Jingbing, Zhang Ruifang, Wang Hong, Meng Deyu 1359
- The review of low-light image enhancement
Ma Long, Ma Tengyu, Liu Risheng 1392
- The critical review of image and video quality assessment methods
Cheng Ruqiu, Yu Ye, Shi Daizong, Cai Wen 1410
- The review of distortion-related image quality assessment
Yan Jiebin, Fang Yuming, Liu Xuelin 1430

Dataset

- A real-world quality evaluation dataset for enhanced underwater images
Gu Yuese, Jiang Qiuping, Shao Feng, Gao Wei 1467

Image defogging/deraining

- Cross-domain object detection based foggy image enhancement
Guo Qiang, Pu Shiliang, Zhang Shifeng, Li Bo 1481
- Convolutional neural networks for sand dust image color restoration and visibility enhancement
Shi Zhenghao, Liu Chunyue, Ren Wenqi, Du Shuangli, Zhao Minghua 1493
- 3D attention and Transformer based single image deraining network
Wang Meihua, Ke Fanhui, Liang Yun, Fan Zhun, Liao Lei 1509
- Image deraining algorithm based on morphological filtering and convolution sparse coding
Huang Shuying, Xu Yating, Yang Yong, Guan Juwei 1522
- Single image rain removal based on multi scale progressive residual network
Lu Bei, Gai Shan 1537

Low illumination image enhancement

- An improved fusion strategy based on transparency-guided backlit image enhancement
Zhao Minghua, Cheng Danni, Du Shuangli, Hu Jing, Shi Cheng, Shi Zhenghao 1554
- An attention mechanism and contextual information based low-light image enhancement method
Zhao Xingyun, Sun Bangyong 1565
- Joint multi-residual learning for underwater image enhancement
Chen Long, Ding Dandan 1577
- Deep learning based backlight image enhancement method derived of zero-reference samples
Wang Zhiyin, Zhang Erhu, Shi Zhenghao, Duan Jinghong 1589

Image super resolution

- Cross-scale coupling network for continuous-scale image super-resolution
Wu Hanlin, Li Wanyu, Zhang Libao 1604
- A super-resolution Transformer fusion network for single blurred image
Liu Huacheng, Ren Wenqi, Wang Rui, Cao Xiaochun 1616

Image inpainting

- Overall detail-enhanced multi-exposure images fusion
Chen Bin, Tan Xincheng, Wu Shiqian 1632
- High-resolution damaged images restoration based on convolutional auto-encoder generative adversarial network
Hou Xiangdan, Liu Haoran, Liu Hongpu 1645
- Reference prior and generative prior linked distorted old photos restoration
Liu Jixin, Chen Rui, An Shipeng 1657
- Region attention mechanism based dual human iris completion technology
Zhang Zhili, Zhang Hui, Wang Jia, Xia Yufeng, Liu Liang, Li Peipei, He Zhaofeng 1669
- Attention mechanism embedded multi-scale restoration method for blurred image
Chen Zining, Zhang Hongyi, Zeng Nianyin, Li Han 1682

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2022)05-1410-20

论文引用格式: Cheng R Q, Yu Y, Shi D Z and Cai W. 2022. The critical review of image and video quality assessment methods. Journal of Image and Graphics, 27(05): 1410-1429 (程茹秋, 余烨, 石岱宗, 蔡文. 2022. 图像与视频质量评价综述. 中国图象图形学报, 27(05): 1410-1429) [DOI: 10.11834/jig.210314]

图像与视频质量评价综述

程茹秋¹, 余烨^{1,2*}, 石岱宗¹, 蔡文¹

1. 合肥工业大学计算机与信息学院, 合肥 230009; 2. 工业安全与应急技术安徽省重点实验室, 合肥 230009

摘要: 图像/视频的获取及传输过程中, 由于物理环境及算法性能的限制, 其质量难免会出现无法预估的衰减, 导致其在实际场景中的应用受到限制, 并对人的视觉体验造成显著影响。因此, 作为计算机视觉领域的一项重要任务, 图像/视频质量评价应运而生。其目的在于通过构建计算机数学模型来衡量图像/视频中的失真信息以判断其质量的好坏, 达到自动预测质量的效果。在城市生活、交通监控以及多媒体直播等多个场景中具有广泛的应用前景。图像/视频质量评价研究取得了长足的发展, 为计算机视觉领域中其他任务提供了一定的便利。本文在广泛调研前人研究的基础上, 回顾了整个图像/视频质量评价领域的发展历程, 分别列举了传统方法和深度学习方法中一些具有里程碑意义的算法和影响力较大的算法, 然后从全参考、半参考和无参考 3 个方面分别对图像/视频质量评价领域的一些文献进行了综述, 具体涉及的方法包含基于结构信息、基于人类视觉系统和基于自然图像统计的方法等; 在 LIVE (laboratory for image & video engineering)、CSIQ (categorical subjective image quality database)、TID2013 等公开数据集的基础上, 基于 SROCC (Spearman rank order correlation coefficient)、PLCC (Pearson linear correlation coefficient) 等评价指标, 对一些具有代表性算法的性能进行了分析; 最后总结当前质量评价领域仍存在的一些挑战与问题, 并对其进行了展望。本文旨在为质量评价领域的研究人员提供一个较全面的参考。

关键词: 图像/视频质量评价 (I/VQA); 结构信息; 人类视觉系统 (HVS); 自然图像统计 (NSS); 深度学习

The critical review of image and video quality assessment methods

Cheng Ruqiu¹, Yu Ye^{1,2*}, Shi Daizong¹, Cai Wen¹

1. School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230009, China;

2. Anhui Province Key Laboratory of Industry Safety and Emergency Technology, Hefei 230009, China

Abstract: Images and videos quality has a great impact on the information acquisition of human behavior visual system. Image/video quality assessment (I/VQA) can be regarded as a key factor for multimedia and network video services nowadays. The quality assessment methods are mainly segmented into qualitative and quantitative models. Qualitative methods conduct quality assessment based on the human's eyes, which cost lots of manpower and time consuming resources. Quantitative quality assessment simulates human observation and it can automatically forecast input quality for those of I/VQA researchers. Our review is critical reviewed quality assessment (QA) methods like full reference (FR), reduced reference (RR) and no reference (NR) methods, which are categorized based on the "clean" data capability (the reference image with no distortion). The full reference methods assess the quality of the distorted images with comparison of the "clean"

收稿日期: 2021-05-12; 修回日期: 2021-08-05; 预印本日期: 2021-08-12

* 通信作者: 余烨 yuye@hfut.edu.cn

基金项目: 安徽省重点研发计划资助 (201904d07020010); 国家自然科学基金项目 (61906061)

Supported by: Provincial Key R&D Program of Anhui (201904d07020010); National Natural Science Foundation of China (61906061)

data. The pros of these methods are better performance, low complexity and good robustness. The cons are that the reference images are challenged to obtain in the in-situ scenarios. Compared with the full reference methods, the reduced reference methods reduce the number of the reference images based on featured reference data to predict the quality. All reference data are not required for the no reference methods. Traditional IQA methods are mainly based on structural similarity, human visual system (HVS) and natural scene statistical theory (NSS). The structure similarity based methods assesses the measurement quality derived of the structural information changes; the HVS based methods are based on some human eyes features; Natural scene statistics based methods fit the transformed coefficients distribution of images or videos and compare the gap of reference coefficients and test coefficients. The emerging deep learning methods based on convolutional neural network (CNN) extract image features through convolutional operations and implement logistic regression to update the models. The learning capability of the IQA-oriented CNN has its priority. VQA models are mainly divided into two categories. One category is that the temporal video quality can be obtained based on the IQA methods for single frame, and then integrate the quality of all frames. The integration methods are mainly divided into general average and weighted average methods in which the weights can be obtained in terms of manual setting or learning. The other one is for the three dimensions (3D) video as mentioned below: First, extracting the coefficient distribution by 3D transformation or using 3D CNN to extract features, and then fitting the coefficient distribution or mapping the features to obtain the final quality. Compared to traditional methods, learning-based methods have higher complexity, but better performance. Most of the current VQA methods also use CNN as the backbone structure, which assists in the overall model construction. Our critical review analyzes the growth of I/VQA and lists some representative algorithms. Then, we review some I/VQA based literatures from two aspects, including traditional methods and deep learning-based methods, respectively. The capability of representative algorithms is analyzed derived of Spearman rank order correlation coefficient (SROCC) and Pearson linear correlation coefficient (PLCC) evaluation indexes in terms of laboratory for image & video engineering (LIVE), categorical subjective image quality database (CSIQ), TID2013 and other datasets. Finally, the challenging issues of quality assessment are summarized and predicted.

Key words: image/video quality assessment (I/VQA); structural information; human vision system (HVS); natural scene statistics (NSS); deep learning

0 引言

随着社会与科技的发展,图像和视频已成为人们生活中使用最广泛的数据媒介。据不完全统计,目前国内安装的摄像头约1.76亿个,一个200 W像素的摄像头工作一天的容量约为42.1 GB,按此计算,这些摄像头每天产生的视频约为6.9 EB,再加上抓拍的图像,数据量十分庞大。这些数据在智慧城市、交通安全和生活服务等多个领域发挥着重要作用。海量图像和视频数据,其质量难以保证,使得对其应用带来影响。如何对这些数据的质量进行评价逐渐成为计算机视觉领域的一项重要任务,具有十分重要的应用前景。

在视频监控中,通过图像/视频质量评价(image/video quality assessment, I/VQA)可以预测设备状态,以及对存在问题的设备进行维修或更换;在网络直播中,通过I/VQA可以分析视频质量,以改善终端用户体验。同时,I/VQA也逐渐成为众

多图像和视频处理任务的前提。在图像超分辨率重建中,可以基于生成图像的质量来对模型性能进行判断;在图像/视频压缩中,基于I/VQA可以保证压缩后的码率;在图像增强领域,可以基于增强图像质量的评价来引导整个模型更新。

I/VQA方法分为主观和客观两类。主观方法通过人为打分的方式获得平均主观得分(mean opinion score, MOS)或平均主观得分差异(differential mean opinion score, DMOS),目前仅用于需要构造数据集的情况下,基于主观方法获取图像的MOS作为数据集的标签。由于该方法需要耗费大量的人力,成本高且效率低下,因此,I/VQA主要以客观评价方法为主。本文也集中在对客观评价方法的综述上。

客观评价方法主要分为全参考(full-reference, FR)、半参考(reduced-reference, RR)和无参考(no-reference, NR)3类。其中,NR方法又称为盲图像/视频质量评价(blind image/video quality assessment, BIQA/BVQA)方法。FR方法在对测试图像或视频

进行质量评价时,需要原始的参考图像或视频作为对比;RR方法仅需从参考数据中提取部分信息作为对比;NR方法不需要参考数据,能够直接对测试图像进行质量评价,是目前质量评价领域的研究重点,在3种方法中实用性最强。

随着图像采集、监控技术的发展,I/VQA在受到众多研究者关注的同时,也面临着以下挑战:1)客观I/VQA方法如何能更好地模拟人眼主观感知;2)如何能对变化多端的失真均进行有效评估,使其具有实用性;3)如何平衡I/VQA模型的性能与速度,提高模型效率。面对上述挑战,研究者们提出了很多I/VQA方法,尤其是随着深度学习技术的发展,许多新的设计思路和评价方法相继提出。

本文将对I/VQA领域的研究进展进行全面综述,分别从传统I/VQA方法和深度学习时代下的一些新兴方法两个方面展开论述,同时对一些经典

I/VQA算法的性能进行评估和分析,最后对I/VQA领域存在的一些问题进行总结和展望。

与Chandler(2013)、方玉明等人(2021)以及Zhai和Min(2020)近期的综述不同,本篇综述中贯通了IQA与VQA领域的关联,贯通了传统方法与基于深度学习方法的区别与关联,综述更加全面,对整个QA行业的发展起到了借鉴和推动作用。

1 I/VQA方法的发展历程

质量评价领域自20世纪70年代逐渐发展至今,目前虽没有统一的衡量标准,但依然诞生了很多具有影响力的算法。本文对质量评价领域自2000年以后的发展历程进行总结,将一些有影响力的工作按照时间线的顺序绘制出来,如图1所示,其中RRED为reduced reference entropic differencing。

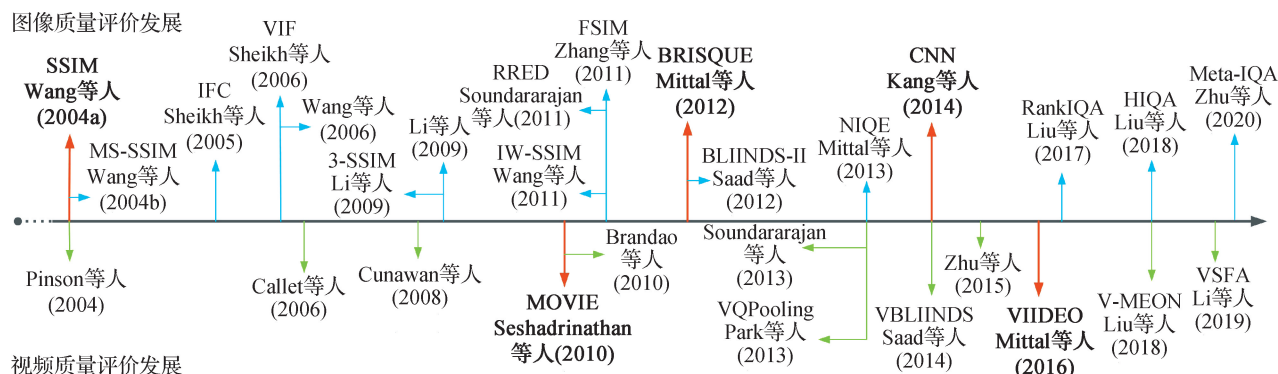


图1 图像与视频质量评价的发展历程(加粗的算法被广泛认为较为经典且性能较好)

Fig. 1 Chronology of classical methods for I/VQA task (the bold algorithms are widely considered to be classic and good performance)

图1中,里程碑式的事件被认为有3个:

1) 基于结构相似性(structural similarity, SSIM) IQA方法(Wang等,2004a)的提出。其改变了以往主流质量评价方法PSNR(peak signal to noise ratio)和MSE(mean square error)仅关注图像中像素信息改变的现状,从而使基于结构信息的主观感知逐渐成为评价任务的重心。

2) 基于自然图像统计(natural scene statistics, NSS)的BRISQUE(blind/referenceless image spatial quality evaluator)算法(Mittal等,2012)。这是较早使用自然图像统计处理IQA的工作,它解决了此前大多方法计算开销大、速度不够快、鲁棒性和性能不佳的问题。

3) 2014年基于深度学习的质量评价方法的提

出(Kang等,2014)。该方法证明将深度学习用于质量评价任务中,可以大幅提升算法的性能和鲁棒性,同时也开启了基于深度学习的IQA研究的先河,是IQA领域的一个重要转折点。

传统IQA中具有影响力的一些算法还包括:在SSIM基础上扩展的MS-SSIM(multi-scale structural similarity)(Wang等,2004b)、3-SSIM(three component SSIM)(Li和Bovik,2009)和IW-SSIM(information content weighting SSIM)(Wang和Li,2011);基于信息保真度的IFC(information fidelity criterion)(Sheikh等,2005a)和VIF(visual information fidelity)(Sheikh等,2006);人类视觉系统(human visual system, HVS)相关的FSIM(feature similarity)(Zhang等,2011);基于变换域系数的方法(Wang等,2006;

Li 和 Wang, 2009); 基于 NSS 的 BLIINDS-II (blind image integrity notator using DCT statistics) (Saad 等, 2012) 和 NIQE (natural image quality evaluator) (Mittal 等, 2013) 等。

进入深度学习时代后, IQA 中具有影响力的算法包括: 基于质量排序思想的 RankIQA (Liu 等, 2017)、基于生成对抗网络的 HIQA (hallucinated-IQA) (Lin 和 Wang, 2018) 以及基于元学习的 Meta-IQA (Zhu 等, 2020) 等。这些深度学习方法在一定程度上减少了对参考图像和 MOS 的需求, 促进了 NR-IQA 方法的应用和发展。

在 VQA 中, 虽然没有诸如 SSIM 这类里程碑式的算法, 但有如下两个被认为较经典且贡献较大的算法:

1) 2010 年提出的 MOVIE (motion-based video integrity evaluation index) 算法 (Seshadrinathan 和 Bovik, 2010)。该方法中创新性的时间和空间质量评价模块使其在性能上实现了较大的提升, 引起了很多研究人员的关注, VQA 任务逐渐得到重视。

2) 2016 年提出的 VIIDEO (video intrinsic integrity and distortion evaluation oracle) 算法作为首个完全无参考的视频质量评价方法, 极大地提升了 VQA 算法的实用性。

其他具有一定影响力的方法包括: 减少带宽的方法 (Pinson 和 Wolf, 2004; Gunawan 和 Ghanbari, 2008); 基于时间与运动信息设计的方法 (Le Callet 等, 2006) 和 VQPooling (video quality pooling) (Park 等, 2013); 基于 NSS 的 VBLIINDS (video BLIINDS) (Saad 等, 2014; Zhu 等, 2015) 等。这些传统方法与传统 IQA 方法一样着重于性能的提升。深度学习方法出现后, 诞生了基于 3 维卷积神经网络 (three dimension convolutional neural network, 3D CNN) 的 V-MEON (video multi-task end-to-end optimized neural network) (Liu 等, 2018b) 和基于 2D CNN 的 VSFA (Li 等, 2019) 等方法。这些方法仍以提升性能为主, 这是因为当前大多数 VQA 算法的性能仍达不到要求, 也无法减少对 MOS 等外部信息的依赖。

整体上看, IQA 任务目前聚焦于提升性能的同时提升算法实用性和泛化性, 减少对外部条件的依赖; VQA 则着重于性能方面的提升, 同时也关注着算法的泛化性问题。

2 传统的 I/VQA 方法

传统的 I/VQA 方法主要包含特征提取、特征表示和特征比对/映射 3 个步骤。首先基于手工设计的特征来进行图像/视频中一些能够表示失真信息特征的提取, 然后利用合适的模型对这些特征进行整合和表示, 最后将其与原始特征之间的差距映射为质量 (全/半参考) 或者直接将其映射为质量 (无参考)。该流程在质量评价方法的设计中被广泛采用。下面分别从 FR、RR 和 NR 等 3 个方面对 I/VQA 方法进行综述。

2.1 传统的 IQA 方法

2.1.1 全参考 IQA 方法 (FR-IQA)

全参考方法作为 IQA 领域中最早使用的方法, 已经发展了很长一段时间, 它将测试图像与原始图像进行对比来评估测试图像的质量。由于具备原始数据作为参考, 该方法能够较为精确地衡量图像质量, 如早期的 MSE 和 PSNR 方法已经在 H. 264 等编码标准中用于评价压缩后的图像质量。然而, 由于这两种方法被证明与人眼主观感知不符, 研究人员又设计了一系列基于结构信息和 HVS 的方法。

1) 基于结构信息的方法。基于人眼能够从场景中自适应提取结构信息的假设, Wang 等人 (2004a) 提出一种新的 IQA 方法——结构相似性算法 SSIM, 该方法通过比较失真图像和参考图像之间结构信息的变化来衡量图像质量。

由于 SSIM 能够有效衡量亮度和对比度的退化, 在工业上用来衡量广播电视和有线电视的图像质量, 在超分辨率和图像去模糊任务中也有广泛应用。SSIM 的出现让质量评价任务从以客观感知质量为主转变为以人眼主观感知为中心。

随后, 很多基于 SSIM 的改进方法相继提出。受图像观看距离和显示分辨率的影响, Wang 等人 (2004b) 提出多尺度结构相似性方法 MS-SSIM, 以对 SSIM 方法的单尺度进行扩展, 通过保留多次滤波后的结果, 在多个尺度上衡量结构信息的变化。为解决 SSIM 对模糊图像质量评价效果不佳的问题, Chen 等人 (2006) 提出 GSSIM (gradient-based structural similarity) 方法, 使用梯度替换结构相似度中的对比度和结构度量, 梯度的变化能够衡量由模糊导致的质量下降。此外, 基于模糊会减弱图像的

语义信息这一特性,Zhang 等人(2013)提出 ESSIM (edge strength similarity-based image quality metric) 方法,通过计算边缘强度相似性来衡量由语义感知导致的质量退化。

上述方法均在其第 1 个步骤(特征提取)上进行改进,然而,第 2 个步骤(特征表示)中使用的池化方法也十分重要,部分研究者在此步骤中设计了不同池化方法形成最后的特征表示。Li 和 Bovik (2009)提出 3-SSIM 方法,通过将图像分解为边缘、纹理和平滑 3 个不同区域,在池化过程中对其分配不同权重,以解决 SSIM 对模糊图像评价效果较差的问题。Wang 和 Li(2011)则对多种池化方法进行分析,提出基于信息容量的加权池化方法 IW-SSIM,该方法能够应用在大多数全参考方法的特征表示步骤中,并有效提升算法的主观相关性。

2) 基于人类视觉系统的方法。HVS 作为一个十分复杂且精细的光学“仪器”,具有亮度适应性、视觉掩蔽、多尺度和边缘敏感等多种特性,这些特性广泛应用于质量评价。

IFC (Sheikh 等,2005a) 和 VIF (Sheikh 等,2006) 作为较早考虑 HVS 特性的方法,通过衡量图像的信息保真度,量化测试图像和参考图像之间共享信息的退化来评价图像质量。Ninassi 等人(2007)研究人眼对图像的视觉注意力机制,通过加强显著区域信息、减少非显著区域失真的影响来改进质量评价模型。Larson 和 Chandler(2010)认为 HVS 采用双重策略来确定图像质量,即对于失真严重的图像,HVS 会选择忽略失真,寻找图像内容来获得感知;而在轻微失真图像上,HVS 着重于寻找失真信息来获得感知,基于此提出了 MAD (most apparent distortion) 方法。根据生理学上的研究“视觉上可识别的特征与相位一致性有关”,Zhang 等人(2011)使用相位一致性和梯度特征结合提出了 FSIM 方法。高敏娟等人(2020)则利用 Grünwald-Letnikov 分数阶导数同时衡量图像的全局变化与局部变化,能够有效模拟 HVS 的多尺度特性。

2.1.2 半参考 IQA 方法(RR-IQA)

半参考方法常用于图像/视频的传输系统,受传送通路限制,一般基于图像的部分特征而不是完整的参考图像来完成测试图像的质量评价。常见方法有基于 NSS 和自由能(free energy)的方法。

1) 自然图像统计。自然图像统计方法是通过

量化图像变换域系数分布的退化程度来衡量图像质量。如图 2 所示,对原始和模糊图像分别进行 DCT (discrete cosine transform) 和系数建模,能够观察到,模糊图像 DCT 系数的分布与原始图像的相比,出现了偏差,通过对这种偏差的评估可以衡量模糊图像的质量。

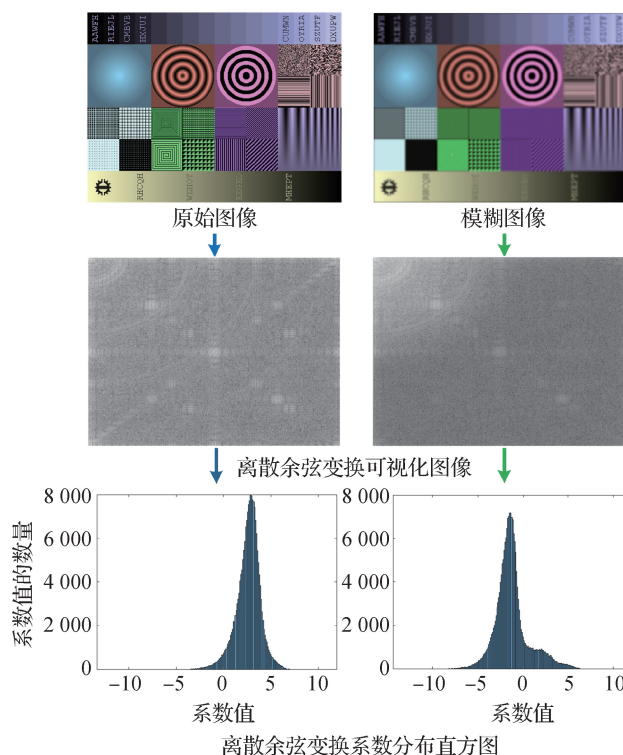


图 2 原始图像(左侧)和模糊图像(右侧)的离散余弦变换可视化图像和系数分布直方图

Fig. 2 DCT images and coefficient distribution histograms of reference image(left) and blur image(right)

Wang 等人(2006)提出一种构造质量感知图像的方法,即将原始图像的小波系数分布编码进原始图像,形成质量感知图像;测试时对此系数进行解码,通过解码后系数分布与原系数分布的偏差来量化图像的质量。Li 和 Wang(2009)使用相对熵量化原始和失真图像变换系数分布间的非对称性差异,以此来预测图像质量。上述方法直接对变换域系数进行建模,而 Ma 等人(2011)则提出将 DCT 系数重新组织成 3 级系数树,然后使用广义高斯分布对系数进行建模,并采用城市块距离(city block distance, CBD)来衡量两个分布之间的差异以预测质量。除使用广义高斯分布外,Xue 和 Mou(2010)还提出基于韦伯分布建模的半参考方法。

2) 自由能原理。脑理论和神经科学研究中的

自由能原理将外界场景的感知和理解建模为一个主动推理过程,在此过程中,大脑试图用一个内部生成模型来描述人眼所看到的场景。即利用内部生成模型,大脑会对遇到的视觉场景做出相应预测。因此,可以基于视觉输入和大脑预测之间的差异来衡量图像质量。

Zhai 等人(2012)提出计算参考图像和输入图像之间自由能水平的差值来衡量图像质量的方法,该方法对平移、旋转等不会导致视觉质量变化的因素不敏感。Liu 等人(2018a)提出 FSI (free energy principle and sparse representation based index for IQA)方法,利用稀疏表示分别预测参考图像和受损图像,然后基于预测误差熵之间的差值衡量图像质量。Zhu 等人(2019)则提出使用多通道自由能的质量评价方法,在图像小波变换域使用稀疏表示和自由能理论提取特征,最后使用支持向量回归(support vector regression, SVR)得到预测质量。

2.1.3 无参考 IQA 方法(NR-IQA)

无参考方法直接建立图像与 MOS 之间的映射关系,从根本上解决参考图像的需求问题,在实际生活中具有非常大的应用价值。然而,在缺少参考图像的情况下,无参考方法的设计难度也大大增加。

1) 面向压缩失真的评价方法。早期的无参考方法主要用于压缩失真图像的质量评价。图像压缩过程中的变换操作会丢失图像原有的一些信息,造成图像失真。Wang 等人(2002)针对 JPEG 压缩过程中块边界不连续产生的块效应和模糊,通过计算图像块边界差异的平均值和图像信号强度来评价图像质量。Babu 等人(2007)提取边缘幅值、边缘长度、背景活动和背景亮度等信息作为特征映射为最终的质量,以衡量 JPEG 失真。Marziliano 等人(2004)主要针对模糊失真和压缩编码所导致的振铃失真进行衡量。Sheikh 等人(2005b)则使用小波变换域的自然场景双状态模型来衡量 JPEG2000 压缩失真。

除 JPEG 压缩方法外,其他压缩算法还有很多,但影响主观感知的失真不仅只在压缩过程中产生。因此,通用型的 NR-IQA 方法逐渐成为研究的重点。

2) 自然图像统计(NSS)。无参考方法中的 NSS 方法通过对变换域系数分布构建统计模型,并从模型中提取相关特征信息,使用机器学习算法(如:支持向量机),将 MOS 作为标签进行监督训练。测试

时首先提取测试图像系数分布的特征,再输入进训练模型中映射为质量。Saad 等人(2012)提出 BLI-INDS-II 方法,使用 DCT 系数拟合的函数均值和方差作为频率变化系数,将系数分为若干子带提取特征,最后整合所有特征并使用简单贝叶斯分布来预测质量。金波等人(2012)则通过失真图像对数域的小波系数分布来预测原始图像的小波系数分布,最后量化两种分布之间的差距来衡量图像质量。Fang 等人(2015)关注对比度失真所产生的影响,从构建的 NSS 模型中提取能够反映对比度失真的均值、标准差和熵等 5 个特征,最后使用 SVR 将这些特征映射为质量。

上述方法为获得系数分布使用的线性变换极大地增加了模型计算复杂度。为减少计算量,提升计算速度,Mittal 等人(2012)提出直接在图像像素域进行处理的 BRISQUE 方法,通过提取平均去除对比度归一化系数,拟合 4 个方向邻近系数乘积分布,得到系数特征,使用支持向量机将这些特征映射为质量分数。Mittal 等人(2013)提出了完全无参考的自然图像质量评估器 NIQE,使用多元高斯模型拟合 125 幅自然图像的变换系数,并将最终质量表示为失真图像拟合模型和多元高斯模型之间的距离差,整个方法无需训练,速度快且没有 MOS 的需求。

尽管基于 NSS 的方法对某些失真有不错的效果,但由于噪声多样性和图像内容复杂性,这些方法还不能有效地对各种失真进行正确表示,从而限制了模型的性能。

2.2 传统的 VQA 方法

视频作为图像在时间维上排列形成的一种数据,不仅包含图像的空间信息,还包含时间维上连贯的运动信息。因此,其设计主要从以下两部分考虑:1) 空间维上借鉴 IQA 方法或思路;2) 时间维上运用加权平均或 3 维处理等方法进行建模。与 IQA 相同,VQA 也可以分为 FR、RR 和 NR 共 3 类。

2.2.1 全参考 VQA 方法(FR-VQA)

由于 FR-IQA 中 SSIM 方法的成功,很多 FR-VQA 方法直接将基于结构信息的方法扩展到视频域。Wang 和 Li(2007)计算参考视频与测试视频对应帧的 SSIM 值,并以运动向量作为权重信息对得到的帧质量分数进行加权平均,获得最终质量。Moorthy 和 Bovik(2010)提出基于运动补偿方法对时间维质量进行评估,视频帧的空间质量则使用基

于结构相似性的算法获得。Seshadrinathan 等人(2010)提出将人眼延迟效应作为池化方法去整合帧级的 SSIM 值。随后又提出了 MOVIE 方法(Seshadrinathan 和 Bovik, 2010), 分别使用时间和空间维度上的特征向量来评估视频质量。与以上分别对时间和空间提取特征、然后再整合的思想不同, Wang 等人(2012)直接在时空域提取结构信息用于质量评价。

基于结构信息的 VQA 方法着重于空间信息的提取, 时间维仅采用池化方法来处理, 导致了大量时间信息的丢失。因此, 部分研究者基于 HVS 特性来进行 VQA 的研究, 该类方法基于某种 3 维变换同时捕捉时间和空间维上的信息作为特征, 避免了池化方法简单将质量分数平均而丢失时间信息的问题。Aydin 等人(2010)从众多生理学研究分析了灵长类动物眼球的亮度适应性和视觉通道掩盖机制, 提出亮度适应模块、3 维时空对比敏感度函数和视觉掩盖预测, 并基于此来解决 VQA 问题。He 等人(2017)则基于 HVS 具有能量压缩表示的特性, 提出使用 3 维 DCT 系数提取的信息表示能量压缩特征, 系数的均值和方差则用来衡量视频频率的变化, 最后计算特征间差异来衡量视频质量。运动信息在视频帧中频繁出现且不断变化, 基于此特性可以衡量时间维上的失真。Vu 等人(2011)将 MAD 方法扩展到视频域, 提出 ST-MAD (spatial temporal MAD) 方法, 该方法将基于光流得到的运动向量作为权重, 对空间特征进行加权映射得到最终的质量。Manasa 和 Channappayya(2016)则直接量化局部光流统计的退化程度来评估视频的质量。Wang 等人(2012)提出基于结构信息扩展的方法, 使用 3 维结构张量同时衡量空间边缘特征和时间运动信息, 能够解决图像内容突变所带来的影响。Yan 和 Mou(2018)则将运动信息分为简单与复杂两部分, 使用梯度标准差估计不同复杂程度的运动信息, 以此评价视频质量。

2.2.2 半参考 VQA 方法(RR-VQA)

图像质量评价中半参考方法应用于传输系统, 对带宽以及内存要求较高。因此, 视频质量评价中 RR 方法更需要考虑减少高带宽占用问题。Pinson 和 Wolf(2004)提出一种半参考方法, 该方法利用系数选择机制对变换后的小波系数进行处理, 最终能够使用较低的带宽保持不错的主观相关性。Gu-

nawan 和 Ghanbari(2008)对原始图像与参考图像进行处理, 得到谐波增益和损耗信息, 然后从这些信息中分析视频质量。Zeng 和 Wang(2010)直接考察复杂小波变换域中的局部相位结构来评价视频质量。Wang 等人(2015)研究了 HVS 的对比度敏感特性后, 提出一种时空信息选择机制, 以降低所需参考信息率。

一般地, IQA 中的自然图像统计方法同样能够扩展至视频域使用。Soundararajan 和 Bovik(2013)使用高斯纹理尺度模型对帧和帧差的小波系数进行建模, 基于系数分布的差异信息计算得到熵差, 以此来量化时空失真对视频质量产生的影响。Bampis 等人(2017)提出了空间效率熵差质量评价方法 (spatial efficient entropic differencing for quality assessment, SpEED-QA), 将帧差块上条件熵差的分布和帧差的差集平均值系数作为特征。Ma 等人(2013)对相邻视频帧使用基于块的运动估计, 并对处理后的视频帧直方图统计量进行建模, 最后基于 CBD 计算原始视频与参考视频序列直方图之间的距离, 获得视频质量。

RR-VQA 的另一研究重点为在时间维上整合已提取的空间特征。Le Callet 等人(2006)通过时间延迟结构整合频率内容、时间内容和块效应 3 个感知特征来预测质量。Park 等人(2013)发现不同视频帧质量会有很大变化, 影响视频的全局质量, 从而提出一种自适应池化方法 VQPooling 去整合视频中每一帧的质量。而 Zhu 等人(2014)则使用一个由全局线性函数、局部分配模块和质量校准构成的映射模型来预测视频质量。

2.2.3 无参考 VQA 方法(NR-VQA)

由于缺少原始参考视频, 早期无参考方法的设计比较困难, 一般用于评估经过压缩后失真的视频的质量。Brandao 和 Queluz(2010)提出一种衡量 H. 264/AVC 压缩编码失真的方法, 分为编码误差估计和编码误差感知加权两个步骤, 通过极大似然估计和线性预测的结合来进行参数估计, 并将这些参数作为模型的特征来预测质量。李晨昊等人(2019)从 H. 264 编码流中提取参数进行统计分析, 以衡量视频内容与运动信息, 然后使用神经网络将特征映射为视频质量。宋巍等人(2020)提出针对水下视频质量评价的无参考方法, 将帧间特征和视频编码参数输入线性回归模型, 进行训练得到视频质量。

除了面向压缩失真方面的研究, NSS 方法在无参考视频质量评价中也得到应用。Saad 等人(2014)提出一种盲视频质量评价方法 VBLIINDS, 该方法是 IQA 方法 BLIINDS-II 的扩展, 它将帧差变换系数拟合模型提取的特征及 NIQE 方法提取的特征, 送入 SVR 并映射为视频质量。同样是对帧进行处理, Zhu 等人(2015)计算图像交流系数和系数块子带的特征, 将所有特征整合并输入神经网络, 映射得到质量分数。Mittal 等人(2016)基于帧差的带通滤波系数进行处理, 从该系数中提取特征来预测视频质量。考虑到当邻近两帧所包含的空间失真完全相同时, 帧差会抵消这两帧的失真信息, Li 等人(2016a)使用 3 维 DCT 进行处理, 拟合系数的分布来提取形状参数等特征, 利用主成分分析对最终的特征信息去除冗余后, 使用 SVR 映射为质量分数。

3 基于深度学习的 I/VQA

2012 年 AlexNet (Alex 等, 2012) 的提出促进了深度学习技术的发展, 使其在图像分割、视频理解等众多计算机视觉任务中逐渐占据主流地位。在质量评价任务中, 深度学习技术主要应用于 NR-I/VQA 领域, 使无参考方法在实用性和性能上均得到了提升。

3.1 基于深度学习的 NR-IQA 方法

传统 NR-IQA 方法通过手工设计特征衡量图像质量。在应用深度学习之后, 通过借助 MOS 进行监督训练、映射特征并得到最终的图像质量。

图像分块作为传统 IQA 方法广泛使用的一大策略, 仍然应用于基于深度学习的方法之中。以图

像块为单位进行训练, 可以增加特征的丰富性和细节信息。Kang 等人(2014)较早采用 CNN 模型解决 IQA 问题, 通过对图像进行分块处理(如图 3 所示), 将每幅图像划分为 N 个 32×32 像素大小且无重叠的图像块, 给每个块分配一个与原图相同的标签。CNN 将图像块视为单独的一幅图像, 提取其特征, 并将得到的特征映射为质量分数, 最后对质量分数进行平均以获得整个图像的质量评价分数。考虑到上述方法的网络结构简单, Kang 等人(2015)在此基础上加深了网络结构, 将模型从单任务改为多任务, 分别预测图像质量和可能存在的失真类别。Fu 等人(2016)则在网络中添加非线性转换和归一化, 以处理包含多种失真的图像。Bosse 等人(2016)根据人眼观察图像时注意力的不同, 提出用加权平均来替代简单平均, 使用两个全连接层作为网络第 2 个分支以输出权重信息, 这些权重是通过网络学习得到的, 能够自适应地调节对不同图像块的注意力。同样考虑视觉感知的不同, Chetouani (2018) 直接在网络输入之前进行处理, 基于 Harel 等人(2006)使用的图显著性方法提取显著区域, 然后将这些区域划分为图像块, 输入进网络进行训练以获得图像质量。He 等人(2019)提出 VRPON (visual residual perception optimized network) 方法, 对图像进行多次重叠采样得到大量图像块, 提取内容特征和失真特征, 使用基于显著图的加权模块对特征进行整合以获得最终质量。朱惠娟等人(2021)利用眼动仪的视点分布图生成基于视觉重要性的权重池模块, 对输入图像进行多次滤波, 从而得到多尺度图像用于模型训练。

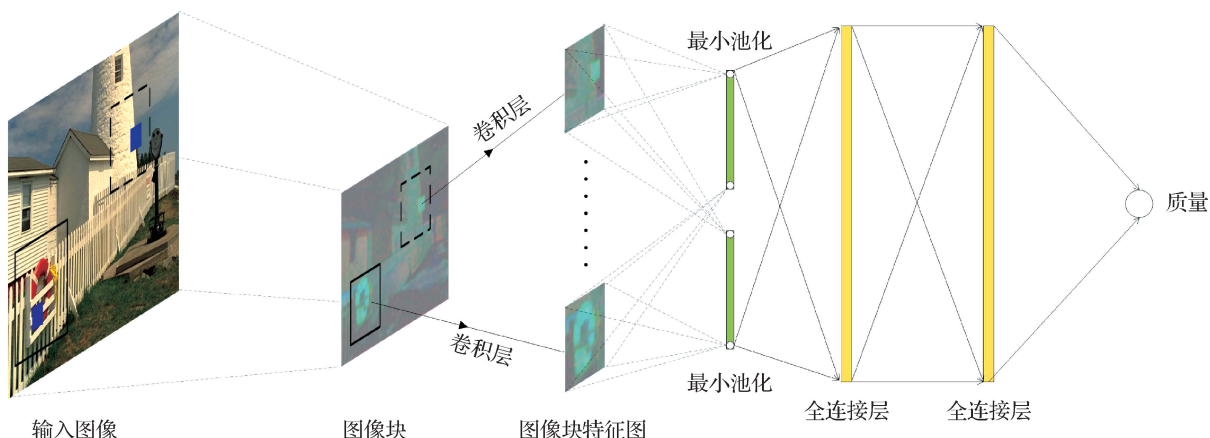


图3 第1个使用深度学习来解决图像质量评价任务的模型结构

Fig.3 Model structure of the first model that used deep learning for IQA task

尽管图像分块能够让模型提取更丰富的特征,但后续的监督训练仍需要 MOS 作为标签,而标定海量数据的 MOS 需要耗费大量的人力和物力,这给模型的监督训练带来了极大阻碍。为缓解此问题,Liu 等人(2017)提出基于图像质量排序进行图像质量评估的 RankIQA 方法,该方法在不含 MOS 的图像上添加不同等级失真,利用暹罗网络对上述图像质量进行排序,最后使用具有 MOS 的数据集对网络进行微调,能够在 MOS 数量有限的情况下扩增训练数据。Niu 等人(2019)同样采用学习质量排序的思想,首先生成图像对,并基于暹罗网络学习排序模型,最后通过将当前测试图像块与其他图像的图像块进行比较,计算该图像块优于其他图像块的次数,获得测试图像的相对质量分数。Chen 等人(2021a)对自然图像(源域)和屏幕内容图像(目标域)运用无监督的领域自适应方法,模型中的编码模块同时提取源域特征和排序信息,回归模块则使用基于源域排序信息生成的伪标签来对目标域进行训练。

基于质量排序的方法减少了 MOS 的使用并对训练数据进行了扩增,这对深度学习方法来说十分关键,使得模型在鲁棒性和性能上均有提升,让研究者们不再约束于常规方法的设计思路,开始尝试一些新思路来设计算法。

自 2014 年 Ian 等人(2014)提出生成对抗网络(generative adversarial nets, GAN)以来,由于其在图像生成、图像去噪、图像转换和图像超分辨率等领域的优良表现,不断有研究者开始基于 GAN 来完成质量评价任务。Ren 等人(2017)提出恢复性对抗网络方法 RAN4IQA(restorative adversarial net),其中,网络恢复器用于恢复原始参考图像;鉴别器用于判别是恢复后的图像还是失真图像。Lin 和 Wang(2018)在失真图像的基础上使用 GAN 生成一个幻觉图像,将其作为原始参考图像,然后融合两幅图像的特征进行训练来获得图像的质量分数,该方法一定程度上解决了参考图像难以获取的问题。此外,考虑到恢复图像的质量会影响最后评价结果,Ma 等人(2020)提出不直接恢复原始图像,而是恢复图像内容的方法,即使用 GAN 来推导失真图像的内容信息,然后基于多路 CNN 质量评估模块对失真图像进行评估。

在深度学习模型方面,一些研究者们尝试用更深更复杂的网络模型来解决 IQA 问题。Li 等人

(2016b)提出深层网络 DeepCNN,基于分类任务对网络进行预训练,然后对参数进行微调以适应质量评价任务。Gao 等人(2018)提出 BLINDER(blind image quality prediction via multi-level deep representations)方法,使用预训练的 VGG16(Visual Geometry Group)网络提取特征来预测质量。Li 等人(2018)则使用残差网络 ResNet-50 来提取内容特征和高层语义特征,将这两种特征映射为质量。Su 等人(2020)以自适应超网络(hyper network, HN)为导向,基于 HN 模块理解图像内容,对质量预测网络进行调整,以适应图像内各种失真。Zhu 等人(2020)则使用基于元学习的无参考方法 MetaIQA 对质量评价模型的学习过程进行学习。

3.2 基于深度学习的 NR-VQA 方法

深度学习时代的 VQA 方法注重帧与帧之间的关联及其可能对视觉感知造成的影响,主要分为基于帧处理并融合时序模块的 2D CNN 与基于视频级处理的 3D CNN 两类。

与传统 VQA 方法类似,深度学习下的 2D CNN 模型也将处理时间和空间信息的模块分别考虑,不同的是,后者额外加入了时序模块对提取的特征进行建模,以加深视频帧之间的关联,最后模型所达到的效果也超越了传统方法。Wang 等人(2018)提出 COME(convolutional neural network and multi-regression based evaluation)方法使用 2D CNN 提取帧级特征,同时从视频帧差中提取运动信息作为时间维特征,将这些特征送入多任务回归模型,映射为最终的视频质量。通过将人眼时空视觉感知纳入研究范围,Kim 等人(2018)使用 CNN 和卷积神经聚合网络量化时空视觉感知,并使用基于注意力机制的时间维池化方法捕捉失真在时间维上的变化。Chen 等人(2021b)提出对 2D CNN 提取的空间特征进行深度监督学习,在时间维上利用多层感知器处理不同时间分辨率,后续的通道级自注意和跳跃连接模块将上层的所有特征聚合并输入进最后的映射模块得到质量。

除时间维池化方法作为时序模块外,Li 等人(2019)提出了基于循环神经网络(recurrent neural network, RNN)的质量评价方法。该方法考虑了 HVS 的内容感知特性,将提取的特征送入门控循环单元(RNN 的一种)来捕捉失真在时间维上的变化,最后提出时间延迟效应启发的时间维池化方法,对

帧级特征进行加权并映射得到整个视频的质量。进一步地,基于上述模型,Li 等人(2021)提出了多数数据集联合训练的方法。Varga 和 Szirányi(2019)也提出基于 RNN 的模型,使用长短期记忆单元来捕捉视频的时间维信息。Shahreza 等人(2019)则使用双向长短期记忆单元来构建更强的时间依赖性,充分捕捉视频时间维的特征信息。Agarla 等人(2020)分别对视频每一帧的质量属性和语义内容进行编码,最后由 RNN 来编码特征并进行回归。Chen 等人(2020)则使用 RNN 构建循环时间分辨率模块,每一层 RNN 单元的输出信息都会被送入前面所有层和本层对应的 RNN 单元中,能够提取更强的时间依赖性。

基于帧处理的 2D CNN 模型将视频看做图像的组合,在输入阶段丢失了帧间相关联的时间信息,从而影响了时序模块的有效性。因此,部分研究者提出基于 3D CNN 的模型对多帧同时进行处理,获得时空特征,同时,时间维上的卷积操作,可以减少时间长度,降低带宽的占用量。Liu 等人(2018b)使用 3D CNN 提取压缩视频的时空特征,利用编码分类器子任务对 CNN 进行预训练,再结合回归子任务对整个网络进行优化。Yang 等人(2018)也使用 3D CNN 作为基础,不同的是输入阶段将视频在时空上进行切片,使用运动强度作为权重进行质量整合。Feng 等人(2021)仅仅将 3D CNN 作为时间模块以提取时间维的特征信息,空间维则使用 2D CNN 提取特征来预测质量。除了对普通视频的质量评价外,Wu 等人(2019)还使用 3D CNN 设计算法对虚拟现实视频进行了质量评价。

4 数据集和各算法模型性能比较

4.1 数据集

在质量评价任务中,具有 MOS 的数据集是客观 I/VQA 算法实验的基础。用于 IQA 的人为失真数据集主要有 LIVE(laboratory for image & video engineering)(IQA)(Sheikh 和 Bovik,2006)、CSIQ(categorical subjective image quality database)(Larson 和 Chandler,2010)、TID2008/2013(Ponomarenko 等,2015)和 KADID/KADIS(konstanz artificially distorted image quality database/set)(Lin 等,2019)等;自然失真数据集有 BID(blurred image database)(Ciancio 等,2011)、CID2013(camera image database)(Virtanen 等,2015)、LIVE Challenge(Ghadiyaram 和 Bovik,2015)、KonIQ(konstanz image quality)(Hosu 等,2020)和 SPAQ(smartphone photography attribute and quality)(Fang 等,2020),其中后 3 者的数据规模较大。除了具有基本的 MOS 信息和图像属性外,SPAQ 还包含图像类别和 EXIF(exchangeable image file format)信息,是其中最为全面的一个数据集。在 VQA 任务中,人为失真数据集主要有 LIVE(VQA)(Seshadrinathan 等,2010)和 CSIQ(Vu 和 Chandler,2014)等;自然失真数据集有 LIVE-Qualcomm(Ghadiyaram 等,2018)、CVD2014(camera video database)(Nuutinen 等,2016)和 KoNViD-1k(konstanz natural video database)(Hosu 等,2017)等。上述数据集的详细信息如表 1 和表 2 所示。

表 1 图像质量评价数据集

Table 1 Image quality assessment datasets

数据集	参考图像/幅	分辨率/像素	失真数/种类	失真图像/幅	失真类型
LIVE IQA	29	512 × 768	5	779	人为失真
CSIQ	30	512 × 512	6	866	人为失真
TID2008	25	384 × 512	17	1 700	人为失真
TID2013	25	384 × 512	24	3 000	人为失真
KADID	81	—	25	10 215	人为失真
KADIS	140 000	—	25	700 000	人为失真
KonIQ	—	1 024 × 768	—	10 073	自然失真
BID	—	—	—	585	自然失真
CID2013	—	1 600 × 1 200	—	480	自然失真
SPAQ	—	—	—	11 125	自然失真
LIVE C	—	500 × 500	—	1 162	自然失真

注:最后一行是 LIVE Challenge 数据集的简写;“—”表示无相应数据。

表 2 视频质量评价数据集
Table 2 Video quality assessment datasets

数据集	参考视频/条	分辨率/像素	失真数/种类	失真视频/条	失真类型
LIVE VQA	10	—	4	150	人为失真
CSIQ Video	12	832 × 480	6	216	人为失真
LIVE Q	—	1 920 × 1 080	6	208	自然失真
CVD2014	—	1 280 × 720	—	234	自然失真
KoNViD-1k	—	960 × 540	—	1 200	自然失真

注:第 3 个数据集是 LIVE Qualcomm 数据集的简写;“—”表示无相应数据。

4.2 各算法性能比较

主要对上述一些具有代表性的算法模型性能进行评估和分析,为更好地反映模型与主观感知之间的相关性和单调性,采用了斯皮尔曼秩相关系数 (Spearman rank order correlation coefficient, SROCC)、皮尔逊线性相关系数 (Pearson linear correlation coefficient, PLCC) 和均方根误差 (root mean squared error, RMSE) 的评价指标。SROCC 判断预测值与真实值所处等级之间的相关性,即判断二者是否接近,以检验预测是否准确。PLCC 通过对两组数据的分布和分布之间的线性距离进行计算,判断预测的准确概率。RMSE 计算预测值与真实值之间的偏差,通过偏差反映预测的准确性。实验中传统方法的实验性能从调研的论文中获取,部分具有开源代码的深度学习方法 (如 CNN 方法) 是通过重新训练并测试得出性能。

SROCC 和 PLCC 的取值范围是 0~1,数值越大代表模型的主观相关性越高,即模型的精度越高。RMSE 取值范围没有限制,其值越小代表模型误差越小,模型越精准。

4.2.1 传统 IQA 模型性能分析

选取 11 种经典的传统 IQA 算法,对其模型性能进行比较和分析。这些算法包括 PSNR、SSIM、MS-SSIM、FSIM、ESSIM、IW-SSIM、IFC 和 VIF 等。PSNR 为早期最为广泛使用的方法;SSIM 为公认的极大促进了质量评价领域发展的算法;MS-SSIM、ESSIM 和 IW-SSIM 作为结构信息的典型改进算法,可以说明结构信息的有效性;IFC 和 VIF 作为信息保真度的两种典型方法被引用,以与结构信息的方法进行对比;FSIM 说明 HVS 的有效性,并同时与前几种方法进行对比;基于自然图像统计的方法选取了 BLINDS-II、BRISQUE 和 NIQE 等 3 种经典的无参考方法,以便与全参考方法进行

对比。

实验主要在早期被广泛认同的 LIVE 图像数据集 (LIVE-IQA)、近几年提出的 CISQ 数据集和具有丰富失真类型的 TID2013 数据集上开展,实验结果如表 3 所示。

从表 3 中可以看出,作为最早使用的全参考方法 PSNR,其在 3 个数据集上的 SROCC 和 PLCC 值都低于其他全参考方法,这是因为 PSNR 仅计算像素值之间的差值,不符合主观视觉感知。PSNR 的 RMSE 值也大于其他方法,这说明模型的误差较大。

为改善此问题而提出的 SSIM 方法假设人类视觉感知受结构信息的影响,与 PSNR 相比,其在 3 个数据集上的相关性都有明显的提升,表明结构信息能够有效地表达主观感知。基于 SSIM 的一系列改进算法 MS-SSIM、ESSIM 和 IW-SSIM 在 SSIM 的基础上进一步提升精度和泛化性,从表 3 中可以看出,其性能均有不同程度的提升。而基于 HVS 设计的 FSIM 算法参考了 SSIM 设计思想使用的是梯度和相位一致性结合的特征相似性,在 3 个数据集上的 SROCC、PLCC 和 RMSE 指标都达到了传统算法的最优效果。

基于信息保真度的 IFC 方法模拟了视觉系统对失真信息的处理,虽然其性能指标要略差于 SSIM,但其改进算法 VIF 在 3 个数据集上的 SROCC、PLCC 和 RMSE 与 FSIM 算法不相上下。

表 3 的最后 3 行是基于 NSS 的经典无参考方法的实验结果,从中可以看出 NR 方法不使用参考图像,与 FR 方法相比,其在 3 个数据集上的实验效果都有明显下降。完全无参考方法 NIQE 与使用 MOS 做监督训练的 BRISQUE 和 BLINDS-II 方法相比,在性能上不相上下,这证明了自然统计模型的产生不一定需要使用大量的数据做监督训练,无监督的方式 also 具有很大的前景。

表 3 传统 IQA 方法在 LIVE、CSIQ 和 TID2013 数据集上的性能
Table 3 Performance of traditional IQA methods on LIVE, CSIQ and TID2013 datasets

类型	算法	LIVE			CSIQ			TID2013		
		SROCC	PLCC	RMSE	SROCC	PLCC	RMSE	SROCC	PLCC	RMSE
	PSNR	0.863	0.859	13.367	0.806	0.800	0.158	0.687	0.675	1.137
	SSIM(Wang 等,2004a)	0.948	0.945	8.126	0.875	0.861	0.133	0.742	0.790	0.851
	MSSIM(Wang 等,2004b)	0.944	0.943	9.095	0.914	0.900	0.114	0.786	0.820	0.711
FR-IQA	IFC(Sheikh 等,2005a)	0.923	0.925	6.656	0.748	0.838	0.143	0.569	0.736	0.909
	VIF(Sheikh 等,2006)	0.963	0.960	7.673	0.919	0.928	0.098	0.750	0.809	0.789
	FSIM(Zhang 等,2011)	0.963	0.960	7.678	0.924	0.912	0.108	0.802	0.859	0.635
	ESSIM(Zhang 等,2013)	0.961	0.963	—	0.933	0.940	—	—	—	—
	IW-SSIM(Wang 和 Li,2011)	0.957	0.952	8.347	0.921	0.914	0.106	—	—	—
	BLINDS-II(Saad 等,2012)	0.912	0.916	8.80	0.780	0.832	—	0.536	0.628	—
NR-IQA	BRISQUE(Mittal 等,2012)	0.939	0.942	9.54	0.775	0.817	—	0.573	0.651	—
	NIQE(Mittal 等,2013)	0.913	0.915	—	0.821	0.865	—	—	—	—

注:加粗字体表示各列最优结果;“—”表示对应文献未给出数据。

从表 3 中还能明显地看出,所有方法在 TID2013 上的实验效果比在其他两个数据集上的效果差了很多。这是由于早期的传统方法针对特定数据集设计手工特征,其泛化能力受到限制,虽然它们在 LIVE 上已经达到了性能瓶颈,却仍无法衡量 TID2013 中的多种失真情况。

4.2.2 基于深度学习的 IQA 模型性能分析

选取 8 种基于深度学习的经典算法进行性能分析,这些算法包括 CNN、DeepCNN、BLINDER 和 Hyper-IQA 等。选择的原因说明如下:CNN 方法为较早使用深度学习的方法;DeepCNN 作为随后提出的深度学习方法,与 CNN 方法进行对比;BLINDER 和 Hyper-IQA 作为近几年较新颖的算法代表;RankIQA 则是使用了图像对的质量排序新思想设计的算法;VRPON 和 HIQA 为基于 GAN 的方法中性能较好的代表。

实验主要在 LIVE 图像数据集和 TID2013 数据集上开展,实验结果如表 4 所示。较早使用深度学习的 NR-IQA 方法是 CNN,其在 LIVE 数据集上的 SROCC 和 PLCC 均达到了 0.95,在 TID2013 上则均高于 0.71,远远超过了传统 NR 方法 NIQE(NIQE 的性能如表 3 所示)的性能,与 SSIM 这类 FR 方法不

相上下,展现出了深度学习方法在质量评价领域的巨大潜力。此后基于卷积神经网络设计的 VRPON 在复杂失真数据集 TID2013 上的效果较 CNN 方法有 10% 的提升。

表 4 基于深度学习的无参考 IQA 方法在 LIVE-IQA 和 TID2013 数据集上的性能
Table 4 Performance of some deep learning-based NR-IQA methods on LIVE-IQA and TID2013 datasets

算法	LIVE-IQA		TID2013	
	SROCC	PLCC	SROCC	PLCC
CNN(Kang 等,2014)	0.956	0.953	0.718	0.778
DeepCNN(Li 等,2016b)	0.935	0.956	—	—
Rank-IQA(Liu 等,2017)	0.980	0.982	0.780	0.792
HIQA(Lin 和 Wang,2018)	0.983	0.989	0.879	0.838
BLINDER(Gao 等,2018)	0.966	0.959	0.819	0.838
VRPON(He 等,2019)	0.955	0.961	0.837	0.840
Hyper-IQA(Su 等,2020)	0.983	0.966	0.903	0.910
RAN4IQA(Ren 等,2017)	0.960	0.970	0.871	0.842

注:加粗字体表示各列最优结果;“—”表示对应文献未给出数据。

DeepCNN 和 BLINDER 作为对骨干网络进行改进的方法,其性能与 CNN 相比,除 DeepCNN 的 SROCC 指标比 CNN 的低之外,其他的均有所提升,但提升幅度较小,这说明一般的骨干网络已经足够胜任 IQA 任务,在骨干网络方面的改进效果不明显。

基于近些年广泛研究的 GAN 设计的 HIQA 和 RAN4IQA 方法,在两个数据集上都具有非常优秀的性能,尤其是在 TID2013 上,SROCC 达到了 0.87。之所以能够达到如此高的相关性,是因为利用 GAN 对失真图像恢复并产生假参考图像,两幅图像之间包含了更加丰富的语义信息以及差异信息。

作为近期提出的算法 Hyper-IQA,在 LIVE-IQA 上 SROCC 和 PLCC 指标分别达到了 0.98 和 0.96,而在 TID2013 数据集上,其 SROCC 和 PLCC 都达到了 0.9 以上。这是因为其采用了内容理解模块及自适应网络模块来学习人眼的感知规则,在表 3 和表 4 共 20 个算法中具有更优秀的实验效果,已经超越作为全参考方法中优秀算法之一 FSIM 的性能,使无

参考方法能够在性能上与全参考方法媲美。

4.2.3 VQA 模型性能分析

1) 基于 LIVE-VQA 数据集的分析。选取了 10 种经典的 VQA 模型,进行其性能比较和分析。这些模型包括:基于运动信息的 FR 方法 ST-MAD 和 MOVIE;时空维池化方法 VQPooling 基于 NSS 的经典方法;VBLIINDS 和 Li 等人(2016a);完全无参考方法 VIDEO;近些年新兴的深度学习方法 COME 和 Varga 和 Szirányi(2019)。

由于早期的传统方法大多以 LIVE-VQA 数据集上的实验为评价标准,因此选用 LIVE-VQA 作为实验数据集,进行实验结果的评估。

如表 5 所示,Wireless 代表在传输 H.264 压缩比特流出错失真视频上的实验效果;IP 代表在无线网络传输错误的失真视频上的实验效果;MPEG-2 代表 MPEG-2 压缩失真视频上的实验效果;H.264 代表 H.264 压缩失真视频上的实验效果;ALL 则代表在所有失真类型即整个 LIVE-VQA 数据集上进行的实验效果。

表 5 LIVE-VQA 数据集上的各种 VQA 方法的性能
Table 5 Performance of various methods on LIVE-VQA dataset

算法	不同失真类型的 SROCC					不同失真类型的 PLCC				
	Wireless	IP	MPEG2	H.264	ALL	Wireless	IP	MPEG2	H.264	ALL
ST-MAD(Vu 等,2011)	0.806	0.768	0.904	0.847	0.824	0.812	0.790	0.909	0.842	0.830
MOVIE(Seshadrinathan 等,2010)	0.802	0.715	0.766	0.773	0.789	0.838	0.762	0.790	0.759	0.811
MOVIE(VQPooling)(Park 等,2013)	0.802	0.806	0.831	0.850	0.843	0.850	0.801	0.844	0.845	0.861
VBLIINDS(Saad 等,2014)	0.815	0.779	0.839	0.869	0.759	0.951	0.946	0.924	0.893	0.881
Zhu 等人(2015)	-	-	-	-	0.880	-	-	-	-	0.835
VIDEO(Mittal 等,2016)	0.532	0.612	0.674	0.556	0.624	0.625	0.737	0.692	0.631	0.651
Li 等人(2016)	0.815	0.826	0.912	0.961	0.782	0.939	0.964	0.951	0.976	0.891
Speed-QA(Bampis 等,2017)	-	-	-	-	0.784	-	-	-	-	0.774
COME(Wang 等,2018)	0.833	0.771	0.881	0.857	0.821	0.917	0.913	0.933	0.929	0.852
Varga 和 Szirányi(2019)	-	-	-	-	0.703	-	-	-	-	0.691

注:加粗字体表示最优结果;“-”表示对应文献未给出数据。

MOVIE 和 ST-MAD 作为基于运动信息的全参考方法,从表 5 上看其 SROCC 和 PLCC 都达到 0.78 以上。这表明运动信息是有效的,但仍有较大的提升空间。从理论上讲,研究者们期望基于相应 IQA

算法延展的 VQA 算法也能达到同样优秀的性能,但事实并非如此。如表 5 所示,基于 IQA 方法 BLIINDS-II 设计的 VBLIINDS 方法,其在 LIVE-VQA 数据集上的 SROCC 和 PLCC 分别只有 0.75 和 0.88,

而 BLIINDS-II 在 LIVE-IQA 数据集上的 SROCC 和 PLCC 都能够达到 0.91, 由此可以看出, 基于 IQA 延展的 VQA 方法其性能出现了很大的下降。这也说明, 视频的时间维信息丢失是影响质量评价的重要因素之一。基于此, 出现了一些重点考虑时间维整合信息的方法, 如 VQPooling 方法, 将其应用在 MOVIE, 在 LIVE-VQA 数据集上的 SROCC 上达到了 0.84, 与不使用 VQPooling 的 MOVIE 方法相比较, 性能有了 5% 的提升。同时 Zhu 等人 (2015) 提出的时间维整合信息的算法也有不错的性能, 其 SROCC 和 PLCC 分别为 0.88 与 0.83, 高于使用 VQPooling 模块的 MOVIE 方法。

完全无参考方法 VIIDEO 在 LIVE-VQA 上的 SROCC 只有 0.62, 这说明完全无参考方法在视频上的效果尚有待提高, 这是由于时间维的变化信息给完全无参考的结果带来了无法预估的误差, 也是目前 VQA 算法的难点所在。

从表 5 中可以看出, Li 等人 (2016b) 的方法在针对 4 种失真分别评价时具有非常好的效果, 其 SROCC 和 PLCC 指标均高于其他方法; 在对 4 种失真同时进行评价时, PLCC 指标也是最高的, 然而, 其 SROCC 指标却有明显下降, 仅高于 VIIDEO 及 Varga 和 Szirányi (2019) 的方法。这可能是由于算法中使用的拟合函数在受到某种失真影响时其分布产生了变化, 但在遭受另一种失真时, 上述失真所产生的影响被当前失真产生的影响所抵消, 此时算法对于这两种失真评价效果就会减弱, 导致整体性能有明显下降。

随着深度学习方法在 IQA 中的成功应用, VQA

任务中也不断有基于深度学习的方法出现, COME 方法的 SROCC 和 PLCC 指标都达到了 0.8 以上, 说明深度学习方法在 VQA 中是有效的。但 Varga 和 Szirányi (2019) 的方法却只达到了 0.7, 这是由于该方法本是针对自然失真数据集设计, 并使用自然失真数据集对骨干网络进行了预训练。这也显现出深度学习方法的一大问题: 算法的最终实验结果受训练集数据的影响。

2) 基于 KoNViD-1k 和 LIVE-Qualcomm 的分析。随着 VQA 任务的发展, 基于 LIVE-VQA 数据集设计的算法已经无法满足 VQA 的需求, 因此研究者们提出了更具挑战性的自然失真数据集。本节则是基于 LIVE-Qualcomm 和 KoNViD-1k 两个应用较多的数据集进行算法性能的评价。评价的模型包括: 传统的 VQA 方法 VBLIINDS 和 VIIDEO; 创新使用 RNN 相关模块设计的算法 (Li 等, 2019; Varga 和 Szirányi, 2019; Agarla 等, 2020)。

如表 6 所示, 在自然失真数据集上, 传统方法 VBLIINDS 和 VIIDEO 的效果都非常差。VBLIINDS 在两个数据集上的 SROCC 和 PLCC 均为 0.5 ~ 0.7, 而 RMSE 的差别很大, 说明模型的精度不稳定; VIIDEO 方法则基本无法对视频的质量做出有效评价。上述结果表明, 算法中手工设计的特征难以捕捉视频中的自然失真。基于深度学习的 VSFA、Varga 和 Szirányi (2019) 以及 Agarla 等人 (2020) 的方法都达到不错的效果, 在两个数据集上的 SROCC 和 PLCC 都在 0.7 之上。

上述实验结果表明, 无论是针对人为失真还是自然失真, 网络自适应提取的特征能够有效地表示

表 6 KoNViD-1k 和 LIVE-Qualcomm 数据集上的一些经典 VQA 方法性能

Table 6 Performance of some classical VQA methods on KoNViD-1k and LIVE-Qualcomm datasets

算法	数据集					
	KoNViD-1k			LIVE-Qualcomm		
	SROCC	PLCC	RMSE	SROCC	PLCC	RMSE
VBLIINDS (Saad 等, 2014)	0.600	0.591	0.493	0.560	0.686	9.005
VIIDEO (Mittal 等, 2016)	0.298	0.303	0.610	0.127	0.101	12.308
VSFA (Li 等, 2019)	0.755	0.744	0.469	0.737	0.732	8.863
Varga 和 Szirányi (2019)	0.849	0.867	—	—	—	—
Agarla 等人 (2020)	0.810	0.810	0.391	0.740	0.772	7.931

注: 加粗字体表示各列最优结果; “—”表示对应文献未给出数据。

失真信息,但最终的实验效果仍有很大提升空间。

4.3 小结

从前面的实验分析中,可以得出如下结论:

1)从方法上看,早期传统方法手工设计的特征在人为失真数据集上相关性较好,但在复杂失真数据集或自然失真数据集上则效果较差。

2)从实验性能上看,深度学习方法更有希望解决 NR 方法在自然失真数据集上性能差的问题。

3)从整个质量评价任务看,VQA 领域的研究尚不够成熟,仍需注重性能上的提升,而 IQA 领域则更注重解决一些算法实用性上的问题。

5 思考与展望

传统 I/VQA 方法从结构信息、模拟人类视觉原理和自然图像统计等方面开展研究,提出了很多非常优秀的算法。然而,传统 I/VQA 方法主要针对某类失真进行特征设计,其泛化性受到限制。随着深度学习的发展,学习特征能力的增强,更多优秀的无参考方法被陆续提出,性能上也逐渐超越了传统的全参考方法,在缓解原始参考数据缺乏及带来带宽增大等难题的同时,进一步推动了 NR-I/VQA 方法的发展。

目前,I/VQA 领域仍然面临如下问题和挑战:

1)I/VQA 模型的实际应用问题。无论是传统方法还是深度学习方法,无论是 FR 还是 NR 方法,目前大多数的研究都集中在主观相关性的问题上。然而,在实际应用中,如:H. 264/AVC 压缩编码、图像增强和图像超分辨率重建,仍然使用较早的 PSNR 方法来衡量图像/视频的质量。现有的一些 I/VQA 算法仍然难以得到应用,这说明当前质量评价任务的相关研究仍局限于理论,其实际应用仍面临挑战。

2)模型的泛化性问题。虽然目前很多 I/VQA 算法已经达到了很好的性能,然而,它们都局限于评估此算法的数据集,很多算法仅仅在单一数据集上性能较高,在其他数据集上却难以达到同样的效果。虽然随着深度学习技术的发展,数据集规模和失真类型的增加,模型的泛化性有了进一步的增强,然而,现实生活中存在的失真类型种类繁多,要想使算法能同时适合对不同失真的评价,仍然面临很大的困难。

3)模型的效率问题。实际场景中只有模型的性能高且计算量足够小时才能得到应用。FR 和 RR 方法虽然均达到较好的性能,但由于参考数据不易获取,限制了这些方法的使用。NR 方法虽然不需要原始参考数据,但传统 NR 方法的性能较差,而基于深度学习的 NR 方法,由于其包含一些复杂的网络结构,参数量较大,尤其是使用 3D CNN 和 VGG 等参数较多的网络时。因此,如何提升 I/VQA 模型的效率,仍然值得思考。

未来若干年,基于深度学习的 NR 方法仍然是质量评价领域的主流。如何利用深度学习技术为 NR 设计出更鲁棒、性能更优越的算法,以适应不同失真类型及不同场景内容,仍然有待进一步探索。

6 结语

本文从 FR、RR 和 NR 出发,对传统和基于深度学习的 I/VQA 进行了综述,通过对各种方法的分析,介绍了 I/VQA 领域的发展历程。同时,在 LIVE-I/VQA、TID2013 和 KoNViD-1k 等经典数据集上对各种经典算法的性能进行了比较和分析,并针对当前 I/VQA 领域尚存在的挑战进行了思考和展望。

无论是传统的方法还是基于深度学习的 I/VQA 方法,其在速度和性能方面都有各自的特点,适合不同的应用场景。传统方法由于参数量少、硬件要求低,主要应用于对速度有较高要求的场景。而基于深度学习的方法评价精度高,但计算量大,主要适用于追求精度的场景。各种 I/VQA 方法在模拟人类感知时尚存在较大差距,实际情况中可以根据不同需求进行选择,以达到最优的效果。

参考文献 (References)

- Agarla M, Celona L and Schettini R. 2020. No-reference quality assessment of in-capture distorted videos. *Journal of Imaging*, 6(8): #74 [DOI: 10.3390/jimaging6080074]
- Alex K, Ilya S and Geoffrey E. H. 2012. ImageNet classification with deep convolutional neural networks. *Communication of the ACM*, 60(6):84-90 [DOI: 10.1145/3065386]
- Aydin T O, Čadík M, Myszkowski K and Seidel H P. 2010. Video quality assessment for computer graphics applications. *ACM Transactions on Graphics*, 29(6): #161 [DOI: 10.1145/1882261.1866187]
- Babu R V, Suresh S and Perkis A. 2007. No-reference JPEG-image

- quality assessment using GAP-RBF. *Signal Processing*, 87(6): 1493-1503 [DOI: 10.1016/j.sigpro.2006.12.014]
- Bampis C G, Gupta P, Soundararajan R and Bovik A C. 2017. SpEED-QA: spatial efficient entropic differencing for image and video quality. *IEEE Signal Processing Letters*, 24(9): 1333-1337 [DOI: 10.1109/LSP.2017.2726542]
- Bosse S, Maniry D, Wiegand T and Samek W. 2016. A deep neural network for image quality assessment//*Proceedings of 2016 IEEE International Conference on Image Processing (ICIP)*. Phoenix, USA; IEEE: 3773-3777 [DOI: 10.1109/ICIP.2016.7533065]
- Brandao T and Queluz M P. 2010. No-reference quality assessment of H.264/AVC encoded video. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(11): 1437-1447 [DOI: 10.1109/TCSVT.2010.2077474]
- Chandler D M. 2013. Seven challenges in image quality assessment: past, present, and future research. *International Scholarly Research Notices*, 2013: #905685 [DOI: 10.1155/2013/905685]
- Chen B L, Li H L, Fan H F and Wang S Q. 2021a. No-reference screen content image quality assessment with unsupervised domain adaptation. *IEEE Transactions on Image Processing*, 30: 5463-5476 [DOI: 10.1109/TIP.2021.3084750]
- Chen G H, Yang C L and Xie S L. 2006. Gradient-based structural similarity for image quality assessment//*Proceedings of 2006 International Conference on Image Processing*. Atlanta, USA; IEEE: 2929-2932 [DOI: 10.1109/ICIP.2006.313132]
- Chen P F, Li L D, Ma L, Wu J J and Shi G M. 2020. RIRNet: recurrent-in-recurrent network for video quality assessment//*Proceedings of the 28th ACM International Conference on Multimedia*. New York, USA; ACM: 834-842 [DOI: 10.1145/3394171.3413717]
- Chen P F, Li L D, Wu J J, Zhang Y B and Lin W S. 2021b. Temporal reasoning guided QoE evaluation for mobile live video broadcasting. *IEEE Transactions on Image Processing*, 30: 3279-3292 [DOI: 10.1109/tip.2021.3060255]
- Chetouani A. 2018. Convolutional neural network and saliency selection for blind image quality assessment//*Proceedings of the 25th IEEE International Conference on Image Processing (ICIP)*. Athens, Greece; IEEE: 2835-2839 [DOI: 10.1109/icip.2018.8451654]
- Ciancio A, Da Costa A L N T T, Da Silva E A B, Said A, Samadani R and Obrador P. 2011. No-reference blur assessment of digital pictures based on multifeature classifiers. *IEEE Transactions on Image Processing*, 20(1): 64-75 [DOI: 10.1109/TIP.2010.2053549]
- Fang Y M, Ma K D, Wang Z, Lin W S, Fang Z J and Zhai G T. 2015. No-reference quality assessment of contrast-distorted images based on natural scene statistics. *IEEE Signal Processing Letters*, 22(7): 838-842 [DOI: 10.1109/LSP.2014.2372333]
- 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, 睦相杰, 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, Zhu H W, Zeng Y, Ma K D and Wang Z. 2020. Perceptual quality assessment of smartphone photography//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, USA; IEEE: 3674-3683 [DOI: 10.1109/CVPR42600.2020.00373]
- Feng Y J, Li S M and Chang Y L. 2021. Multi-scale feature-guided stereoscopic video quality assessment based on 3d convolutional neural network//*Proceedings of 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Toronto, Canada; IEEE: 2095-2099 [DOI: 10.1109/ICASSP39728.2021.9414231]
- Fu J, Wang H L and Zuo L X. 2016. Blind image quality assessment for multiply distorted images via convolutional neural networks//*Proceedings of 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Shanghai, China; IEEE: 1075-1079 [DOI: 10.1109/icassp.2016.7471841]
- Gao F, Yu J, Zhu S G, Huang Q M and Tian Q. 2018. Blind image quality prediction by exploiting multi-level deep representations. *Pattern Recognition*, 81: 432-442 [DOI: 10.1016/j.patcog.2018.04.016]
- Gao M J, Dang H S, Wei L L, Wang H L and Zhang X D. 2020. Combining global and local variation for image quality assessment. *Acta Automatica Sinica*, 46(12): 2662-2671 (高敏娟, 党宏社, 魏立力, 王海龙, 张选德. 2020. 结合全局与局部变化的图像质量评价. *自动化学报*, 46(12): 2662-2671) [DOI: 10.16383/j.aas.c190697]
- Ghadiyaram D and Bovik A C. 2015. LIVE in the wild image quality challenge database [DB/OL]. [2021-05-27]. <http://live.ece.utexas.edu/research/ChallengeDB/index.html>.
- Ghadiyaram D, Pan J, Bovik A C, Moorthy A K, Panda P, and Yang K C. 2018. In-capture mobile video distortions: a study of subjective behavior and objective algorithms. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(9): 2061-2077 [DOI: 10.1109/TCSVT.2017.2707479]
- Gunawan I P and Ghanbari M. 2008. Reduced-reference video quality assessment using discriminative local harmonic strength with motion consideration. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(1): 71-83 [DOI: 10.1109/TCSVT.2007.913755]
- Harel J, Koch C and Perona P. 2006. Graph-based visual saliency//*Proceedings of the 19th International Conference on Neural Information Processing Systems (NIPS)*. Vancouver, Canada; NIPS: 545-552 [DOI: 10.5555/2976456.2976525]
- He L H, Lu W, Jia C C and Hao L. 2017. Video quality assessment by compact representation of energy in 3D-DCT domain. *Neurocomputing*, 269: 108-116 [DOI: 10.1016/j.neucom.2016.08.143]
- He L H, Zhong Y Z, Lu W and Gao X B. 2019. A visual residual perception optimized network for blind image quality assessment. *IEEE*

- Access, 7: 176087-176098 [DOI: 10.1109/access.2019.2957292]
- Hosu V, Hahn F, Jenadeleh M, Lin H, Men H, Szirányi T, Li S J and Saupe D. 2017. The konstanz natural video database (KoNViD-1k)//Proceedings of the 9th International Conference on Quality of Multimedia Experience (QoMEX). Erfurt, Germany: IEEE: 1-6 [DOI: 10.1109/QoMEX.2017.7965673]
- Hosu V, Lin H H, Sziranyi T and Saupe D. 2020. KonIQ-10k: an ecologically valid database for deep learning of blind image quality assessment. *IEEE Transactions on Image Processing*, 29: 4041-4056 [DOI: 10.1109/TIP.2020.2967829]
- Ian J G, Jean P A, Mehdi M, Bing X, David W F, Sherjil O, Aaron C and Yoshua B. 2014. Generative Adversarial Nets//Proceedings of Advances in Neural Information Processing Systems. Montreal, Canada: NIPS 2672-2680 [DOI: 10.1109/NIPS.1406.2661]
- Jin B, Li C F and Wu X J. 2012. No-reference image quality assessment based on natural scene statistics and wavelet. *Journal of Image and Graphics*, 17(1): 33-39 (金波, 李朝锋, 吴小俊. 2012. 结合 NSS 和小波变换的无参考图像质量评价. *中国图象图形学报*, 17(1): 33-39 [DOI: 10.11834/jig.20120105]
- Kang L, Ye P, Li Y and Doermann D. 2014. Convolutional neural networks for no-reference image quality assessment//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE: 1733-1740 [DOI: 10.1109/CVPR.2014.224]
- Kang L, Ye P, Li Y and Doermann D. 2015. Simultaneous estimation of image quality and distortion via multi-task convolutional neural networks//Proceedings of 2015 IEEE International Conference on Image Processing (ICIP). Quebec City, Canada: IEEE: 2791-2795 [DOI: 10.1109/ICIP.2015.7351311]
- Kim W, Kim J, Ahn S, Kim J and Lee S. 2018. Deep video quality assessor: from spatio-temporal visual sensitivity to a convolutional neural aggregation network//Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer: 224-241 [DOI: 10.1007/978-3-030-01246-5_14]
- Larson E C and Chandler D M. 2010. Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of Electronic Imaging*, 19(1): # 011006 [DOI: 10.1117/1.3267105]
- Le Callet P, Viard-Gaudin C and Barba D. 2006. A convolutional neural network approach for objective video quality assessment. *IEEE Transactions on Neural Networks*, 17(5): 1316-1327 [DOI: 10.1109/TNN.2006.879766]
- Li C F and Bovik A C. 2009. Three-component weighted structural similarity index//Proceedings of the SPIE 7242, Image Quality and System Performance VI. San Jose, USA: SPIE: 72420Q [DOI: 10.1117/12.811821]
- Li C H, Zhuo L and Li J F. 2019. Content-based H.264 non-reference video quality evaluation model. *Measurement and Control Technology*, 38(1): 106-110, 135 (李晨昊, 卓力, 李嘉峰. 2019. 基于内容的 H.264 无参考视频质量评价模型. *测控技术*, 38(1): 106-110, 135) [DOI: 10.19708/j.ckjs.2019.01.023]
- Li D Q, Jiang T T and Jiang M. 2018. Exploiting high-level semantics for no-reference image quality assessment of realistic blur images//Proceedings of the 25th ACM International Conference on Multimedia. Mountain View, USA: ACM: 378-386 [DOI: 10.1145/3123266.3123322]
- Li D Q, Jiang T T and Jiang M. 2019. Quality assessment of in-the-wild videos//Proceedings of the 27th ACM International Conference on Multimedia. Nice, France: ACM: 2351-2359 [DOI: 10.1145/3343031.3351028]
- Li D Q, Jiang T T and Jiang M. 2021. Unified quality assessment of in-the-wild videos with mixed datasets training. *International Journal of Computer Vision*, 129(4): 1238-1257 [DOI: 10.1007/s11263-020-01408-w]
- Li Q and Wang Z. 2009. Reduced-reference image quality assessment using divisive normalization-based image representation. *IEEE Journal of Selected Topics in Signal Processing*, 3(2): 202-211 [DOI: 10.1109/JSTSP.2009.2014497]
- Li X L, Guo Q and Lu X Q. 2016a. Spatiotemporal statistics for video quality assessment. *IEEE Transactions on Image Processing*, 25(7): 3329-3342 [DOI: 10.1109/TIP.2016.2568752]
- Li Y M, Po L M, Feng L T and Yuan F. 2016b. No-reference image quality assessment with deep convolutional neural networks//Proceedings of 2016 IEEE International Conference on Digital Signal Processing (DSP). Beijing, China: IEEE: 685-689 [DOI: 10.1109/ICDSP.2016.7868646]
- Lin H H, Hosu V and Saupe D. 2019. KADID-10k: a large-scale artificially distorted IQA database//Proceedings of the 11th International Conference on Quality of Multimedia Experience (QoMEX). Berlin, Germany: IEEE: 1-3 [DOI: 10.1109/QoMEX.2019.8743252]
- Lin K Y and Wang G X. 2018. Hallucinated-IQA: no-reference image quality assessment via adversarial learning//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, America: IEEE: 732-741 [DOI: 10.1109/CVPR.2018.00083]
- Liu W T, Duanmu Z F and Wang Z. 2018b. End-to-end blind quality assessment of compressed videos using deep neural networks//Proceedings of the 26th ACM international conference on Multimedia. Seoul, Korea(South): ACM: 546-554 [DOI: 10.1145/3240508.3240643]
- Liu X L, Van De Weijer J and Bagdanov A D. 2017. RankIQA: learning from rankings for no-reference image quality assessment//Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE: 1040-1049 [DOI: 10.1109/ICCV.2017.118]
- Liu Y T, Zhai G T, Gu K, Liu X M, Zhao D B and Gao W. 2018a.

- Reduced-reference image quality assessment in free-energy principle and sparse representation. *IEEE Transactions on Multimedia*, 20(2): 379-391 [DOI: 10.1109/tmm.2017.2729020]
- Ma J P, Wu J J, Li L D, Dong W S and Xie X M. 2020. Active inference of GAN for no-reference image quality assessment//Proceedings of 2020 IEEE International Conference on Multimedia and Expo (ICME). London, UK: IEEE: 1-6 [DOI: 10.1109/ICME46284.2020.9102895]
- Ma L, Li S N, Zhang F and Ngan K N. 2011. Reduced-reference image quality assessment using reorganized DCT-based image representation. *IEEE Transactions on Multimedia*, 13(4): 824-829 [DOI: 10.1109/TMM.2011.2109701]
- Ma L, Ngan K N and Xu L. 2013. Reduced reference video quality assessment based on spatial HVS mutual masking and temporal motion estimation//Proceedings of 2013 IEEE International Conference on Multimedia and Expo (ICME). San Jose, America: IEEE: 1-6 [DOI: 10.1109/ICME.2013.6607611]
- Manasa K and Channappayya S S. 2016. An optical flow-based full reference video quality assessment algorithm. *IEEE Transactions on Image Processing*, 25(6): 2480-2492 [DOI: 10.1109/TIP.2016.2548247]
- Marziliano P, Dufaux F, Winkler S and Ebrahimi T. 2004. Perceptual blur and ringing metrics; application to JPEG2000. *Signal Processing; Image Communication*, 19(2): 163-172 [DOI: 10.1016/j.image.2003.08.003]
- Mittal A, Moorthy A K and Bovik A C. 2012. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 21(12): 4695-4708 [DOI: 10.1109/TIP.2012.2214050]
- Mittal A, Saad M A and Bovik A C. 2016. A completely blind video integrity oracle. *IEEE Transactions on Image Processing*, 25(1): 289-300 [DOI: 10.1109/TIP.2015.2502725]
- Mittal A, Soundararajan R and Bovik A C. 2013. Making a “completely blind” image quality analyzer. *IEEE Signal Processing Letters*, 20(3): 209-212 [DOI: 10.1109/LSP.2012.2227726]
- Moorthy A K and Bovik A C. 2010. Efficient video quality assessment along temporal trajectories. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(11): 1653-1658 [DOI: 10.1109/TCSVT.2010.2087470]
- Ninassi A, Le Meur O, Le Callet P and Barba D. 2007. Does where you gaze on an image affect your perception of quality? Applying visual attention to image quality metric//Proceedings of 2007 IEEE International Conference on Image Processing. San Antonio, America: IEEE: II-169-II-172 [DOI: 10.1109/ICIP.2007.4379119]
- Niu Y Z, Huang D, Shi Y Q and Ke X. 2019. Siamese-network-based learning to rank for no-reference 2D and 3D image quality assessment. *IEEE Access*, 7: 101583-101595 [DOI: 10.1109/access.2019.2930707]
- Nuutinen M, Virtanen T, Vaahteranoksa M, Vuori T, Oittinen P and Häkkinen J. 2016. CVD2014—a database for evaluating no-reference video quality assessment algorithms. *IEEE Transactions on Image Processing*, 25(7): 3073-3086 [DOI: 10.1109/TIP.2016.2562513]
- Park J, Seshadrinathan K, Lee S and Bovik A C. 2013. Video quality pooling adaptive to perceptual distortion severity. *IEEE Transactions on Image Processing*, 22(2): 610-620 [DOI: 10.1109/TIP.2012.2219551]
- Pinson M H and Wolf S. 2004. A new standardized method for objectively measuring video quality. *IEEE Transactions on Broadcasting*, 50(2): 312-322 [DOI: 10.1109/TBC.2004.834028]
- Ponomarenko N, Jin L N, Ieremeiev O, Lukin V, Egiazarian K, Astola J, Vozel B, Chehdi K, Carli M, Battisti F and Kuo C C J. 2015. Image database TID2013: peculiarities, results and perspectives. *Signal Processing; Image Communication*, 30: 57-77 [DOI: 10.1016/j.image.2014.10.009]
- Ren H Y, Chen D Q and Wang Y Z. 2017. RAN4IQA: restorative adversarial nets for no-reference image quality assessment//Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, America: AAAI: 7308-7314
- Saad M A, Bovik A C and Charrier C. 2012. Blind image quality assessment; a natural scene statistics approach in the DCT domain. *IEEE Transactions on Image Processing*, 21(8): 3339-3352 [DOI: 10.1109/TIP.2012.2191563]
- Saad M A, Bovik A C and Charrier C. 2014. Blind prediction of natural video quality. *IEEE Transactions on Image Processing*, 23(3): 1352-1365 [DOI: 10.1109/TIP.2014.2299154]
- Seshadrinathan K and Bovik A C. 2010. Motion tuned spatio-temporal quality assessment of natural videos. *IEEE Transactions on Image Processing*, 19(2): 335-350 [DOI: 10.1109/TIP.2009.2034992]
- Seshadrinathan K, Soundararajan R, Bovik A C and Cormack L K. 2010. A subjective study to evaluate video quality assessment algorithms//Proceedings of SPIE 7527, Human Vision and Electronic Imaging XV. San Jose, America: SPIE: 75270H [DOI: 10.1117/12.845382]
- Shahreza H O, Amini A and Behroozi H. 2019. No-reference video quality assessment using recurrent neural networks//Proceedings of the 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS). Shahrood, Iran: IEEE: 1-5 [DOI: 10.1109/icspis48872.2019.9066015]
- Sheikh H R and Bovik A C. 2006. Image information and visual quality. *IEEE Transactions on Image Processing*, 15(2): 430-444 [DOI: 10.1109/TIP.2005.859378]
- Sheikh H R, Bovik A C and Cormack L. 2005b. No-reference quality assessment using natural scene statistics; JPEG2000. *IEEE Transactions on Image Processing*, 14(11): 1918-1927 [DOI: 10.1109/TIP.2005.854492]
- Sheikh H R, Bovik A C and De Veciana G. 2005a. An information fidel-

- ity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing*, 14(12): 2117-2128 [DOI: 10.1109/TIP.2005.859389]
- Sheikh H R, Wang Z, Cormack L and Bovik A C. 2006. LIVE image quality assessment database release 2 [DB/OL]. [2021-04-08]. <http://live.ece.utexas.edu/research/Quality/subjective.htm>.
- Song W, Liu S M, Huang D M, Wang W J and Wang J. 2020. Non-reference underwater video quality assessment method for small size samples. *Journal of Image and Graphics*, 25(9): 1787-1799 (宋巍, 刘诗梦, 黄冬梅, 王文娟, 王建. 2020. 适用小样本的无参考水下视频质量评价方法. *中国图象图形学报*, 25(9): 1787-1799) [DOI: 10.11834/jig.200025]
- Soundararajan R and Bovik A C. 2013. Video quality assessment by reduced reference spatio-temporal entropic differencing. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(4): 684-694 [DOI: 10.1109/TCSVT.2012.2214933]
- Su S L, Yan Q S, Zhu Y, Zhang C, Ge X, Sun J Q and Zhang Y N. 2020. Blindly assess image quality in the wild guided by a self-adaptive hyper network//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, America; IEEE: 3664-3673 [DOI: 10.1109/CVPR42600.2020.00372]
- Varga D and Szirányi T. 2019. No-reference video quality assessment via pretrained CNN and LSTM networks. *Signal, Image and Video Processing*, 13(8): 1569-1576 [DOI: 10.1007/s11760-019-01510-8]
- Virtanen T, Nuutinen M, Vaahteranoksa M, Oittinen P and Häkkinen J. 2015. CID2013: a database for evaluating no-reference image quality assessment algorithms. *IEEE Transactions on Image Processing*, 24(1): 390-402 [DOI: 10.1109/TIP.2014.2378061]
- Vu P V, Vu C T and Chandler D M. 2011. A spatiotemporal most-apparent-distortion model for video quality assessment//*Proceedings of the 18th IEEE International Conference on Image Processing*. Brussels; IEEE: 2505-2508 [DOI: 10.1109/ICIP.2011.6116171]
- Vu V P and Chandler D M. 2014. ViS3: an algorithm for video quality assessment via analysis of spatial and spatiotemporal slices. *Journal of Electronic Imaging*, 23(1): #013016 [DOI: 10.1117/1.JEI.23.1.013016]
- Wang C F, Su L and Zhang W G. 2018. COME for no-reference video quality assessment//*Proceedings of 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. Miami, USA; IEEE: 232-237 [DOI: 10.1109/MIPR.2018.00056]
- Wang M M, Zhang F and Agraftotis D. 2015. A very low complexity reduced reference video quality metric based on spatio-temporal information selection//*Proceedings of 2015 IEEE International Conference on Image Processing (ICIP)*. Quebec, Canada; IEEE: 571-575 [DOI: 10.1109/ICIP.2015.7350863]
- Wang Y, Jiang T T, Ma S W and Gao W. 2012. Novel spatio-temporal structural information based video quality metric. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(7): 989-998 [DOI: 10.1109/TCSVT.2012.2186745]
- Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004a. 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 and Li Q. 2007. Video quality assessment using a statistical model of human visual speed perception. *Journal of the Optical Society of America A*, 24(12): B61-B69 [DOI: 10.1364/JOSAA.24.000B61]
- Wang Z and Li Q. 2011. Information content weighting for perceptual image quality assessment. *IEEE Transactions on Image Processing*, 20(5): 1185-1198 [DOI: 10.1109/TIP.2010.2092435]
- Wang Z, Sheikh H R and Bovik A C. 2002. No-reference perceptual quality assessment of JPEG compressed images//*Proceedings of International Conference on Image Processing*. Rochester, America; IEEE: 1 [DOI: 10.1109/ICIP.2002.1038064]
- Wang Z, Simoncelli E P and Bovik A C. 2004b. Multiscale structural similarity for image quality assessment//*Proceedings of the 37th Asilomar Conference on Signals, Systems and Computers*. Pacific Grove, USA; IEEE: 1398-1402 [DOI: 10.1109/ACSSC.2003.1292216]
- Wang Z, Wu G X, Sheikh H R, Simoncelli E P, Yang E H and Bovik A C. 2006. Quality-aware images. *IEEE Transactions on Image Processing*, 15(6): 1680-1689 [DOI: 10.1109/TIP.2005.864165]
- Wu P, Ding W X, You Z X and An P. 2019. Virtual reality video quality assessment based on 3d convolutional neural networks//*Proceedings of the 26th IEEE International Conference on Image Processing (ICIP)*. Taipei, China; IEEE: 3187-3191 [DOI: 10.1109/icip.2019.8803023]
- Xue W F and Mou X Q. 2010. Reduced reference image quality assessment based on Weibull statistics//*Proceedings of the 2nd International Workshop on Quality of Multimedia Experience (QoMEX)*. Trondheim, Norway; IEEE: 1-6 [DOI: 10.1109/QoMEX.2010.5518131]
- Yan P and Mou X Q. 2018. Video quality assessment based on motion structure partition similarity of spatiotemporal slice images. *Journal of Electronic Imaging*, 27(3): #033019 [DOI: 10.1117/1.JEI.27.3.033019]
- 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]
- Zeng K and Wang Z. 2010. Temporal motion smoothness measurement for reduced-reference video quality assessment//*Proceedings of 2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. Dallas, America; IEEE: 1010-1013 [DOI: 10.1109/ICASSP.2010.5495316]
- 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]
- Zhai G T, Wu X L, Yang X K, Lin W S and Zhang W J. 2012. A psychovisual quality metric in free-energy principle. IEEE Transactions on Image Processing, 21 (1): 41-52 [DOI: 10.1109/TIP.2011.2161092]
- Zhang L, Zhang L, Mou X Q and Zhang D. 2011. FSIM: a feature similarity index for image quality assessment. IEEE Transactions on Image Processing, 20(8): 2378-2386 [DOI: 10.1109/TIP.2011.2109730]
- Zhang X D, Feng X C, Wang W W and Xue W F. 2013. Edge strength similarity for image quality Assessment. IEEE Signal Processing Letters, 20(4): 319-322 [DOI: 10.1109/LSP.2013.2244081]
- Zhu H C, Li L D, Wu J J, Dong W S and Shi G M. 2020. MetaIQA: deep meta-learning for no-reference image quality assessment//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 14131-14140 [DOI: 10.1109/CVPR42600.2020.01415]
- Zhu H J, Zong P and Cong Y H. 2021. Multi-scale image quality assessment method based on weight pool. Computer Engineering and Applications, 57(3): 215-221 (朱惠娟, 宗平, 丛玉华. 2021. 基于权重池的多尺度图像质量评估方法. 计算机工程与应用, 57(3): 215-221) [DOI: 10.3778/j.issn.1002-8331.2008-0365]
- Zhu K F, Barkowsky M, Shen M M, Le Callet P and Saupe D. 2014. Optimizing feature pooling and prediction models of VQA algorithms//Proceedings of 2014 IEEE International Conference on Image Processing (ICIP). Paris, France: IEEE: 541-545 [DOI: 10.1109/ICIP.2014.7025108]
- Zhu K F, Li C Q, Asari V and Saupe D. 2015. No-reference video quality assessment based on artifact measurement and statistical analysis. IEEE Transactions on Circuits and Systems for Video Technology, 25(4): 533-546 [DOI: 10.1109/TCSVT.2014.2363737]

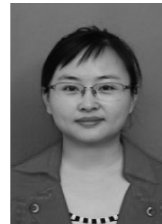
- Zhu W H, Zhai G T, Min X K, Hu M H, Liu J, Guo G D and Yang X K. 2019. Multi-channel decomposition in tandem with free-energy principle for reduced-reference image quality assessment. IEEE Transactions on Multimedia, 21 (9): 2334-2346 [DOI: 10.1109/TMM.2019.2902484]

作者简介



程茹秋, 1997 年生, 男, 硕士研究生, 主要研究方向为图像/视频质量评价。

E-mail: 1269130224@qq.com



余烨, 通信作者, 女, 副教授, 硕士生导师, 主要研究方向为计算机视觉、虚拟现实与可视化。

E-mail: yuye@hfut.edu.cn

石岱宗, 男, 硕士研究生, 主要研究方向为目标检测。

E-mail: 1359926433@qq.com

蔡文, 女, 硕士研究生, 主要研究方向为计算机视觉。

E-mail: 15226555110@163.com