



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 图形学报

2021
02
VOL.26

ISSN1006-8961
CN11-3758/TB



深度神经网络结构搜索 P245



第26卷第2期（总第298期）

2021年2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@radi.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@radi.ac.cn
Telephone 010-58887035
Website www.cjig.cn

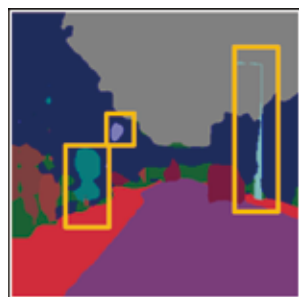
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

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



单幅图像刚体目标姿态估计方法综述(第0334页)



DeepLabv3plus-IRCNet: 小目标特征提取的图像语义分割(第0391页)



融合视觉风格和标签约束的少数民族服装图像解析(第0402页)

前沿进展

NeurIPS 2020观察与分析

林宙辰, 王奕森 0229

学者观点

深度神经网络结构搜索综述

唐浪, 李慧霞, 颜晨倩, 郑侠武, 纪荣嵘 0245

无参考图像质量评价研究进展

方玉明, 睦相杰, 鄢杰斌, 刘学林, 黄丽萍 0265

轨道病害视觉检测: 背景、方法与趋势

王建柱, 李清勇, 张靖, 甘津瑞 0287

RMFS-CNN: 遥感图像分类深度学习新框架

赵雪梅, 吴军, 陈睿星 0297

综述

深度学习在医学影像智能处理中的应用与挑战

左艳, 黄钢, 聂生东 0305

视频车辆黑烟检测算法研究进展

张天琪, 杨伟东, 张姣姣, 彭凯 0316

单幅图像刚体目标姿态估计方法综述

杨步一, 杜小平, 方宇强, 李佩阳, 王阳 0334

流体运动估计光流算法研究综述

邵绪强, 杨艳, 刘艺林 0355

图像处理和编码

帧内块复制中的位移矢量参数编码算法

赵利平, 林涛, 杨玉芬, 胡珂立, 彭华 0368

图像分析和识别

遮挡判定下多层次重定位跟踪算法

姜文涛, 金岩, 刘万军 0378

DeepLabv3plus-IRCNet: 小目标特征提取的图像语义分割

刘文, 王海荣, 周北京 0391

融合视觉风格和标签约束的少数民族服装图像解析

张茜, 刘骊, 甘霖, 付晓东, 刘利军, 黄青松 0402

孪生导向锚框RPN网络实时目标跟踪

尚欣茹, 温尧乐, 奚雪峰, 胡伏原 0415

融合密度和精细分数的行人检测

甄烨, 王子磊, 吴枫 0425

图像理解和计算机视觉

融合边缘保持与改进代价聚合的立体匹配算法

程德强, 李海翔, 寇旗旗, 于泽宽, 庄焕东, 吕晨 0438

医学图像处理

Seg-CapNet: 心脏MRI图像分割神经网络模型

刘畅, 林楠, 曹仰杰, 杨聪 0452

遥感图像处理

自适应感受野机制遥感图像分割模型

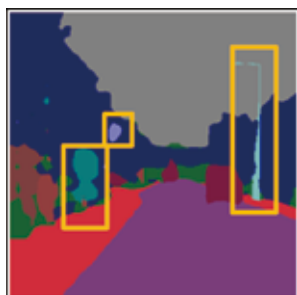
刘航, 汪西莉 0464

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Review of rigid object pose estimation from a single image (P0334)



DeepLabv3plus-IRCNet: an image semantic segmentation method for small target feature extraction (P0391)



Clothing parsing of Chinese minorities via the fusion of visual style and label constraints (P0402)

Frontier

Report of NeurIPS 2020

Lin Zhouchen, Wang Yisen 0229

Scholar View

Survey on neural architecture search

Tang Lang, Li Huixia, Yan Chenqian, Zheng Xiawu, Ji Rongrong 0245

Progress in no-reference image quality assessment

Fang Yuming, Sui Xiangjie, Yan Jiebin, Liu Xuelin, Huang Liping 0265

Visual inspection of rail defects: background, methodologies, and trends

Wang Jianzhu, Li Qingyong, Zhang Jing, Gan Jinrui 0287

RMFS-CNN: new deep learning framework for remote sensing image classification

Zhao Xuemei, Wu Jun, Chen Ruixing 0297

Review

Application and challenges of deep learning in the intelligent processing of medical images

Zuo Yan, Huang Gang, Nie Shengdong 0305

Video black smoke detection methods for vehicles: a survey

Zhang Tianqi, Yang Weidong, Zhang Jiaojiao, Peng Kai 0316

Review of rigid object pose estimation from a single image

Yang Buyi, Du Xiaoping, Fang Yuqiang, Li Peiyang, Wang Yang 0334

Review of optical flow algorithms in fluid motion estimation

Shao Xuqiang, Yang Yan, Liu Yilin 0355

Image Processing and Coding

Displacement vector coding algorithm for intra block copy

Zhao Liping, Lin Tao, Yang Yufen, Hu Keli, Peng Hua 0368

Image Analysis and Recognition

Multilevel relocation tracking algorithm under occlusion decision

Jiang Wentao, Jin Yan, Liu Wanjun 0378

DeepLabv3plus-IRCNet: an image semantic segmentation method for small target feature extraction

Liu Wen, Wang Hairong, Zhou Beijing 0391

Clothing parsing of Chinese minorities via the fusion of visual style and label constraints

Zhang Qian, Liu Li, Gan Lin, Fu Xiaodong, Liu Lijun, Huang Qingsong 0402

Target tracking system based on the Siamese guided anchor region proposal network

Shang Xinru, Wen Yaole, Xi Xuefeng, Hu Fuyuan 0415

Pedestrian detection method based on density and score refinement

Zhen Ye, Wang Zilei, Wu Feng 0425

Image Understanding and Computer Vision

Stereo matching algorithm based on edge preservation and improved cost aggregation

Cheng Deqiang, Li Haixiang, Kou Qiqi, Yu Zekuan, Zhuang Huandong, Lyu Chen 0438

Medical Image Processing

Seg-CapNet: neural network model for the cardiac MRI segmentation

Liu Chang, Lin Nan, Cao Yangjie, Yang Cong 0452

Remote Sensing Image Processing

Remote sensing image segmentation model based on an adaptive receptive field mechanism

Liu Hang, Wang Xili 0464

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2021)02-0265-22

论文引用格式: 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(02): 0265-0286(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(02): 0265-0286) [DOI:10.11834/jig.200274]

无参考图像质量评价研究进展

方玉明, 眭相杰, 鄢杰斌, 刘学林, 黄丽萍

江西财经大学信息管理学院, 南昌 330032

摘要: 图像质量评价一直是图像处理和计算机视觉领域的一个基础问题, 图像质量评价模型也广泛应用于图像/视频编码、超分辨率重建和图像/视频视觉质量增强等相关领域。图像质量评价主要包括全参考图像质量评价、半参考图像质量评价和无参考图像质量评价。全参考图像质量评价和半参考图像质量评价分别指预测图像质量时参考信息完全可用和部分可用, 而无参考图像质量评价是指预测图像质量时参考信息不可用。虽然全参考和半参考图像质量评价模型较为可靠, 但在计算过程中必须依赖参考信息, 使得应用场景极为受限。无参考图像质量评价模型因不需要依赖参考信息而具有较强的适用性, 一直都是图像质量评价领域研究的热点。本文主要概述2012—2020年国内外公开发表的无参考图像质量评价模型, 根据模型训练过程中是否需要用到主观分数, 将无参考图像质量评价模型分为有监督学习和无监督学习的无参考图像质量评价模型。同时, 每类模型分成基于传统机器学习算法的模型和基于深度学习算法的模型。对基于传统机器学习算法的模型, 重点介绍相应的特征提取策略及思想; 对基于深度学习算法的模型, 重点介绍设计思路。此外, 本文介绍了图像质量评价在新媒体数据中的研究工作及图像质量评价的应用。最后对介绍的无参考图像质量评价模型进行总结, 并指出未来可能的发展方向。

关键词: 图像质量评价; 人类视觉系统; 视觉感知; 自然统计特征; 机器学习; 深度学习

Progress in no-reference image quality assessment

Fang Yuming, Sui Xiangjie, Yan Jiebin, Liu Xuelin, Huang Liping

School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China

Abstract: Image quality assessment (IQA) has been a fundamental issue in the fields of image processing and computer vision. It has also been extensively applied to other relevant research areas, such as image/video coding, super-resolution and visual enhancement. In general, IQA consists of subjective and objective evaluations. Subjective evaluation always refers to estimating the visual quality of images by subject, with the goal of building test benchmarks. Objective evaluation typically resorts to computational algorithms (i. e., IQA models) to make visual quality predictions, and its ultimate objective is to provide consistent judgment with subjects. The effectiveness of objective IQA models must be verified on test benchmarks built via subjective evaluation. Undoubtedly, subjective evaluation cannot be fully embedded into multimedia processing applications because such process is time-consuming and labor-intensive. By contrast, an objective IQA model can work efficiently as an important module in multimedia processing applications, playing roles in visual image quality monitoring, image filtering, and visual quality enhancement. Given their availability, research on objective IQA models has

收稿日期: 2020-06-09; 修回日期: 2020-07-20; 预印本日期: 2020-07-27

基金项目: 国家自然科学基金项目(61822109); 霍英东教育基金会高等院校青年教师基金项目(161061)

Supported by: National Natural Science Foundation of China(61822109)

elicited considerable attention from industries and academia. Objective IQA models can be classified into three categories: full-reference (FR), reduced-reference (RR), and no-reference/blind (NR) models. FR and RR models denote that reference information for estimating the visual quality of images is completely and partially available, respectively. Meanwhile, an NR model indicates that reference information is unavailable for visual quality prediction. Although reference-based IQA models (i.e., FR and RR models) are relatively reliable, their applications are limited to specific scenarios due to their dependence on reference information. By contrast, NR-IQA models are more flexible than reference-based models because they are free from the constraint of reference information. Consequently, NR-IQA models have consistently been a popular research topic over the past decades. In this study, we introduce NR-IQA models published from 2012 to 2020 to provide a comprehensive survey on feature engineering and end-to-end learning techniques in NR-IQA. In accordance with whether subjective quality scores are involved in training procedures, NR-IQA models are classified into two categories: opinion-aware/supervised and opinion-unaware/unsupervised NR-IQA models. To present a clear and integrated description, each category is further divided into two subclasses: traditional machine learning-based models (MLMs) and deep learning-based models (DLMs). For the former subclass, we mostly investigate their individual feature extraction schemes and the principle behind these schemes. In particular, a widely adopted feature extraction approach in MLMs, namely, natural scene statistics (NSS), is introduced in this study. The principle of NSS is as follows: some visual features of quality perfect images follow certain associated distributions; meanwhile, different types of distortions will break this rule in corresponding methods. On the basis of this observation/fact, researchers have proposed many NSS-based NR-IQA methods, in which the estimated parameters of the established distributions are used as quality-aware features. Thereafter, a machine learning algorithm is selected to train the IQA models. Another well-known feature extraction approach described in this study relies on dictionary learning, which is frequently accompanied by sparse coding. The core component of this type of feature extraction approach is to learn a dictionary by searching for a group of over-complete bases. Then, these over-complete bases are used to build a reference system for image representation. A test image can be concretely represented directly or indirectly by the constructed dictionary by using sparse indexes or cluster centroids. Image representations are further used as quality-aware features to capture variations in image quality. For the latter subclass (i.e., DLMs), the design principles described in detail in this paper mostly correspond to different architectures of deep neural networks. In particular, we introduce three different schemes for designing opinion-aware DLMs and commonly used strategies in opinion-unaware DLMs. To guarantee length balance among various contents and clearly exhibit the differences between NR-IQA models designed for natural images and other types of images, we introduce them separately in subsections. In addition, we provide a brief introduction into IQA research on new media, including virtual reality, light field, and underwater sonar images, along with the applications of IQA models. Finally, an in-depth conclusion about NR-IQA models is drawn in the last section. We summarize the current achievements and limitations of MLMs and DLMs. Furthermore, we highlight the potential development trends and directions of NR-IQA models for further improvements from the perspectives of image contents and NR-IQA models.

Key words: image quality assessment (IQA); human visual system (HVS); visual perception; natural scene statistics (NSS); machine learning; deep learning

0 引言

图像数据爆发式增长,广泛出现在人们的生产生活中。图像的视觉质量与用户的体验密切相关,是决定计算机视觉应用的重要因素之一(Fang等, 2017c, 2020b, c)。不同的内在因素如成像设备老化、损伤等(Fang等, 2020c)和外在因素如拍摄环境、数据存储、压缩、传输和处理等(Zhang等,

2013)都与图像质量紧密相关,每个环节都可能引入噪声。同时,高质量图像也是众多计算机视觉应用追求的目标,如超分辨率重建(Yang等, 2019a)、图像修复(Dong等, 2019)和图像去噪(Xu等, 2018)等。

图像质量评价(image quality assessment, IQA)是指通过对图像信号进行相关特性分析,量化图像的视觉失真程度(吴金建, 2014; 顾镭, 2015; 鄢杰斌, 2018)。评价方法主要包括主观质量评价和客

观质量评价。主观质量评价是指由主观判断获取图像视觉质量;客观质量评价是指通过客观算法自动计算图像的视觉质量。根据计算图像视觉质量时是否需要参考图像的信息,客观质量评价可分成3类:全参考质量评价(full reference IQA, FR-IQA)(Wang等,2004)、半参考质量评价(reduced reference IQA, RR-IQA)(Rehman和Wang,2012)和无参考质量评价(no reference IQA, NR-IQA)(Mittal等,2012)。FR-IQA模型和RR-IQA模型主要通过通过对图像的视觉特征进行分析,量化参考图像和失真图像之间的差异,计算失真图像的视觉质量。相对于FR-IQA模型和RR-IQA模型,NR-IQA模型在计算失真图像

的视觉质量时不需要任何参考图像的信息,在实际应用系统中具有更广泛的应用前景。

本文重点介绍 NR-IQA 模型的相关研究进展和未来发展趋势,内容框架如图1所示。根据 NR-IQA 模型训练时是否需要图像的主观分数,将 NR-IQA 模型分为有监督学习的无参考模型(需要主观分数)和无监督学习的无参考模型(不需要主观分数),并对每类模型分别对基于传统机器学习的模型和基于深度学习的模型进行梳理。同时,考虑到不同内容图像的性质差异和内容篇幅的平衡,对自然图像和其他内容图像(如图2所示)的质量评价模型分别介绍。

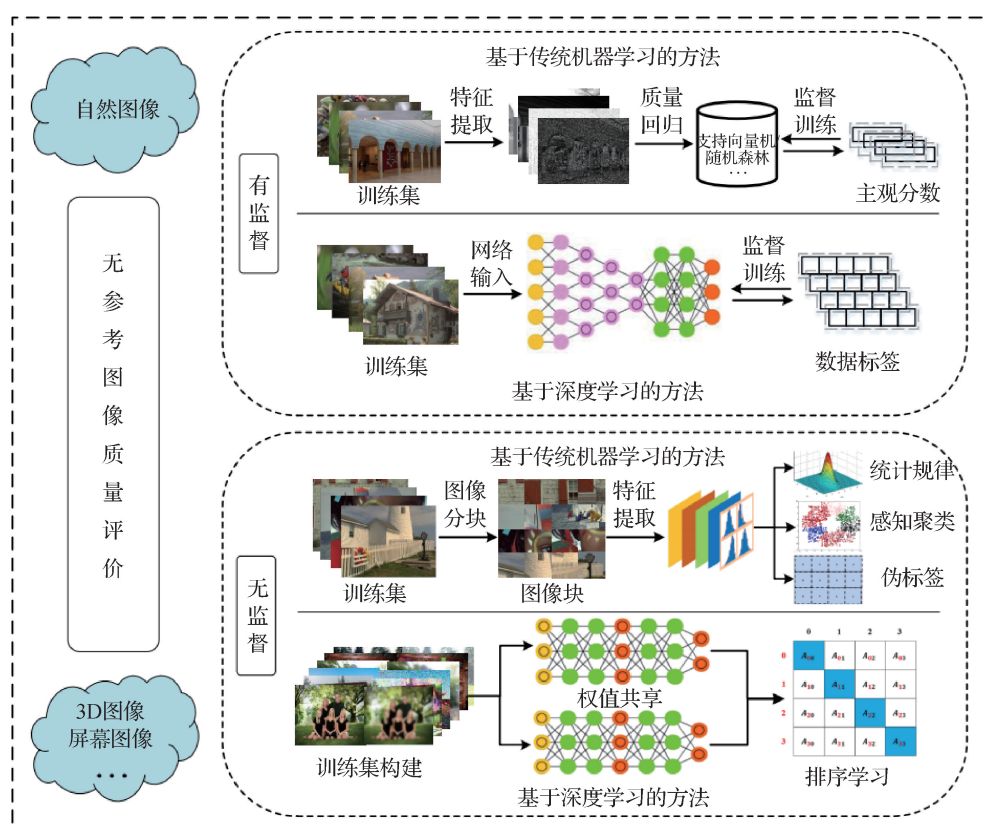


图1 本文内容框架

Fig. 1 The framework of this paper



图2 不同内容图像示例

Fig. 2 Visual examples of different types of images

((a) natural images; (b) screen content images; (c) 3-D images; (d) tone-mapped images)

目前,国内外关于 IQA 的综述文献主要包括基于传统机器学习的针对自然图像的 NR-IQA 方法(王志明,2015)、自然图像和 3 维图像质量评价算法(Niu 等,2019)、基于深度学习的自然图像质量评价方法(Yang 等,2019c)、针对自然图像的 IQA 算法(Athar 和 Wang,2019)、介绍各种内容图像的质量评价算法(Zhai 和 Min,2020)。本文详细介绍 NR-IQA 模型,涵盖了基于传统机器学习算法的模型和基于深度学习算法的模型。同时,分别介绍自然图像和其他内容图像的模型及它们之间的区别。因此,本文是上述综述文献很好的补充。

1 有监督学习的 NR-IQA 模型

有监督学习的 NR-IQA 模型指的是模型训练过程中需要用到主观数据,模型的输入为图像或者图像特征,输出为主观分数,通过迭代训练,获得最终的 NR-IQA 模型。有监督学习的 NR-IQA 模型包括基于传统机器学习算法的有监督学习模型和基于深度学习算法的有监督学习模型。前者依赖先验知识手动提取视觉特征,后者通过端到端的方式学习特征表达。

1.1 基于传统机器学习的模型

基于传统机器学习的有监督学习 NR-IQA 模型旨在设计有效的视觉特征表达方法,能较好地反映图像视觉质量变化,并通过传统机器学习算法如支持向量机(support vector machine, SVM)或随机森林(random forest, RF)等学习视觉特征到图像视觉质量的映射模型,基本框架如图 3 所示。此类模型的研究重点在于特征提取,考虑到不同内容篇幅的平衡以及自然图像和其他内容图像如屏幕图像和合成图像等的内容差异,对自然图像和其他内容图像的模型分别介绍。

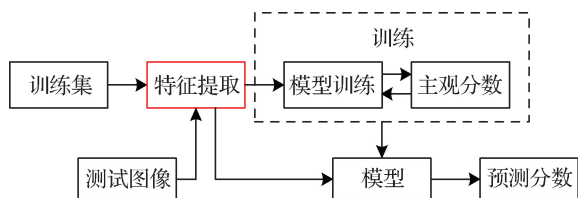


图 3 基于传统机器学习的有监督无参考图像质量评价模型基本框架

Fig. 3 The general framework of the opinion-aware NR-IQA models based on traditional machine learning algorithms

1.1.1 自然图像

在基于传统机器学习算法的 NR-IQA 模型中,广泛使用的是自然场景统计(natural scene statistics, NSS)特征。NSS 是指“质量完好”图像的视觉特征(如亮度、梯度等)服从一定的分布规律,而不同类型或不同程度的失真会对这种分布规律产生相应的影响。研究者利用这种规律,提出了一系列 NR-IQA 模型。

Mittal 等人(2012)较早提出使用空域下的 NSS 特征来构建图像质量评价模型,称为 BRISQUE (blind/referenceless image spatial quality evaluator)。经研究发现,视觉质量完好的图像的亮度值呈现广义高斯分布 (generalized Gaussian distribution, GDD)。同时,像素对的亮度值乘积呈现非对称广义高斯分布 (asymmetric generalized Gaussian distribution, AGGD)。亮度计算方式为

$$\hat{I} = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + c} \quad (1)$$

$$\mu(i,j) = \sum_{h=-H}^H \sum_{l=-L}^L \omega_{h,l} I_{h,l}(i,j) \quad (2)$$

$$\sigma(i,j) = \sqrt{\sum_{h=-H}^H \sum_{l=-L}^L \omega_{h,l} (I_{h,l}(i,j) - \mu(i,j))^2} \quad (3)$$

式中, $I(i,j)$ 表示图像 I 在位置 (i,j) 处的像素值, μ 和 σ 分别表示以 (i,j) 为中心的长宽为 $2H$ 和 $2L$ 的局部区域的均值和方差, ω 为高斯滤波器。均值为 0 的 GDD 模型具体表示为

$$f(x; \alpha, \sigma^2) = \frac{\alpha}{2\zeta \Phi(1/\alpha)} \exp\left(-\left(\frac{|x|}{\zeta}\right)^\alpha\right) \quad (4)$$

$$\zeta = \sigma \sqrt{\frac{\Phi(1/\alpha)}{\Phi(3/\alpha)}} \quad (5)$$

$$\Phi(a) = \int_0^\infty t^{a-1} e^{-t} dt \quad (6)$$

式中, x 表示像素值, α 和 σ^2 分别表示 GGD 模型的形状参数和方差参数, Φ 表示伽马函数 (gamma), a 的取值大于 0。AGGD 模型表达可参考 Mittal 等人(2012)介绍的方法。该方法认为,噪声的存在将破坏亮度值的分布,因此使用拟合得到的 GDD 和 AGGD 模型的参数作为失真图像的特征,并使用 SVM 训练特征与图像视觉质量之间的映射模型。

与 Mittal 等人(2012)方法的设计思路类似, Fang 等人(2015)研究发现图像的矩特征(均值、方差、峰度和偏态)和信息熵呈现某种分布(如高斯分

布), 并提出使用拟合得到的分布函数的输出值作为失真图像的特征, 以构建 NR-IQA 模型。董宏平和刘利雄(2014)提出利用互信息对图像邻近像素相关性进行建模, 在原始图像及其对应的局部标准差图和亮度图上提取多尺度特征, 最后使用 SVM 训练 NR-IQA 模型。贾惠珍等人(2014)融合 NSS 特征和感知特征构建 NR-IQA 模型, 其中, NSS 特征为归一化系数拟合的高斯分布模型的方差和形状参数; 感知特征是指低级视觉特征, 包括相位、梯度和熵。

上述方法(Mittal 等, 2012; Fang 等, 2015; 贾惠珍等, 2014)特征提取过程都需要拟合函数分布。为此, 研究者研究了其他简便的特征提取方式, 即特征提取过程不需要拟合函数分布。Xue 等人(2014)使用局部对比度统计特征感知图像变化, 统计特征包括梯度幅值和高斯差分响应。Gu 等人(2015b)提出一种基于自回归模型的图像清晰度评价方法, 首先计算每个像素点的自回归系数, 然后根据自回归系数的能量及对比差异计算图像的清晰度。Gu 等人(2015a)将自由能机制引入 NR-IQA 研究中, 提出基于自由能机制和自然性的质量相关特征提取方法。王俊平等(2017)在 NR-IQA 算法研究中引入了平坦区域和非平坦区域两个概念, 从图像的非平坦区域提取视觉特征来感知图像质量的变化, 提取的视觉特征包括空间特征和离散余弦变换特征, 最后通过均衡广义回归神经网络训练 NR-IQA 模型。Wu 等人(2019)提出一种基于方向选择视觉感知机理的 NR-IQA 方法, 根据方向选择机制, 计算图像方向直方图作为感知图像失真的特征。Li 等人(2016b)提出使用直方图表示亮度统计特征和结构统计特征, 以构建 NR-IQA 模型。Li 等人(2016a)提出一种多失真图像质量评价模型, 先将图像转换到梯度域, 然后通过局部二值化模式在梯度域下提取多阶导数特征。

基于传统机器学习算法的 NR-IQA 模型的另一种常用的视觉特征提取方法是字典学习。字典学习是指从图像中抽象出有代表性的原子, 组成一个字典, 再通过字典来表征图像。字典的表示形式为

$$\mathbf{D} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_K] \quad (7)$$

式中, $\mathbf{D}$ 表示字典, $\mathbf{p}_k$ 表示图像块或者特征向量, K 表示字典中元素的数量。该类特征提取方法的一般框架如图 4 所示。

Ye 等人(2012)提出一种基于字典学习的 NR-

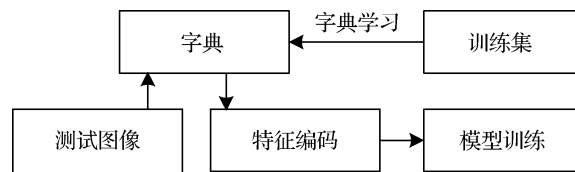


图 4 基于字典学习的特征提取方法的一般框架

Fig. 4 The general framework of dictionary learning based feature extraction methods

IQA 算法, 称为 CORNIA (codebook representation for no-reference image assessment), 首先通过 K-Means 聚类方法, 构建一个码本, 然后通过软分配编码和最大池化获取图像视觉感知特征。高飞和高新波(2014)提出一种主动特征学习方法并应用在 NR-IQA 中, 该学习方法综合考虑样本的表征性和样本间的差异性, 筛选出包含最大信息量的图像块作为字典原子, 并利用学习到的字典提取图像的特征向量, 用于表征图像质量的变化。Xu 等人(2016)提出一种融合高阶统计特征的 NR-IQA 方法, 类似于 CORNIA (Ye 等, 2012), 首先使用 K-Means 构建码本, 并使用归一化图像块均值、方差和协偏度表示码本中的每个图像块。之后, 计算测试图像包含的图像块与码本中图像块之间的高阶统计特征差异, 并作为测试图像的特征, 用于后续的模式训练。Jiang 等人(2017)使用稀疏编码 (sparse coding) 和 K-SVD (K-singular value decomposition) 学习字典, 提出一种基于字典的多阶段特征编码策略, 提取图像特征。

1.1.2 其他内容图像

其他内容图像主要指屏幕图像 (Fang 等, 2017b, 2018b)、超分辨率重建图像 (Ma 等, 2017c)、3 维图像 (Wang 等, 2015)、色调映射图像 (Yeganeh 和 Wang, 2013) 等其他内容形式的图像。因为获取方式的差异或处理方式的不同, 导致针对自然图像 (Sheikh 等, 2006) 设计的模型在其他内容图像上的性能较差, 因而较多研究者针对其他内容图像的特性, 设计出有针对性的 NR-IQA 模型。

Fang 等人(2018b)基于人类视觉系统 (human visual system, HVS) 对亮度和纹理的敏感特性, 提出使用直方图提取亮度统计特征以及高阶导数纹理特征, 感知屏幕图像质量的变化。Chen 等人(2013a)提出一种基于 NSS 特征的 3 维图像质量评价方法, 其中, NSS 特征包括 2 维亮度统计特征和 3 维视差统计特征。考虑到双目视觉特性 (双目竞争和双目

融合), Fang 等人(2019b)提出融合单目视觉特征和双目视觉特征的 3 维图像质量评价模型。其中,单目视觉特征为左右视图的亮度统计特征;双目视觉特征为 3 维图像的结构特征和深度特征。Ma 等人(2017c)较早开展了超分辨率图像质量评价研究,同时使用空域和频域特征量化超分辨率重建过程中引入的噪声,并提出一个两阶段的回归模型训练超分辨率图像质量评价模型。陈曦等人(2019)将 NSS 引入深度图像质量评价研究中,首先使用 Canny 算子检测图像边缘失真区域,然后在该区域上提取一阶特征和二阶特征,共同作为深度 RF 模型的输入。Fang 等人(2020a)考虑色彩和梯度统计特征,提出一种有效的色调映射图像质量评价模型。

1.1.3 小结

基于传统机器学习的 NR-IQA 模型研究主要聚焦在特征提取上,结合主流的机器学习算法训练模型。虽然该类 NR-IQA 模型在小数据集上能够表现出较好的性能,但存在较大的过拟合风险。同时,手工设计特征的表达能力有限,当数据集较大时该类模型无法表现出优势。

此外,针对性的设计也可能使该类算法的通用性受到极大限制。如针对自然图像设计的 NR-IQA 算法(Mittal 等,2012)在屏幕图像质量评价任务中的性能有较大下降,Fang 等人(2018b)提出的方法在屏幕图像质量评价任务中的性能超过了其他相关算法,而在自然图像评价任务中表现的竞争力相对在屏幕图像质量评价任务中的竞争力有所下降。

1.2 基于深度学习的模型

不同于传统手工提取特征的 NR-IQA 模型,基于深度学习的模型可以端到端学习图像与图像质量之间的映射关系,且性能远超过基于传统机器学习的模型,该类模型的重点在于卷积神经网络(convolutional neural network, CNN)的设计,框架如图 5 所示。

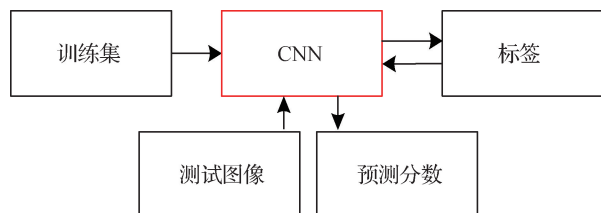


图 5 基于深度学习的有监督 NR-IQA 模型的一般框架

Fig. 5 The general framework of the opinion-aware NR-IQA models based on deep learning algorithms

1.2.1 自然图像

Kang 等人(2014)将 CNN 引入 NR-IQA 模型的设计中,该模型的输入为 32×32 像素的图像块,模型包括 1 个卷积层和 1 个池化层,两路池化操作将每幅特征图都映射为一个最大值和一个最小值,最后加入 2 个全连接层,并通过简单的线性回归预测质量分数。Kim 等人(2019b)提出了一个基于 CNN 的深度 NR-IQA 模型,称为 DIQA(deep image quality assessor)。DIQA 的训练分为两个阶段:第 1 阶段训练可学习客观误差图的 CNN;第 2 阶段使用人类主观分数微调 CNN 模型,使模型进一步学习人类视觉感知行为,预测人类主观意见分数。Yan 等人(2019c)提出了一个基于双流 CNN 结构的 NR-IQA 模型,使用结构相同的两个子网络分别提取失真图及其对应梯度图的特征,为解决图像局部失真分布不均匀的问题,在网络中加入了一个基于区域的全卷积层(Dai 等,2016),用于提取图像块中心周围的信息,生成 $9 \times C$ 个通道的特征图,分别对应 9 个区域和 C 种失真,最后两个子网络提取的特征图经池化、特征连接等操作,映射为图像块质量分数。Zhang 等人(2020)提出了一个适用于合成失真及真实失真的深度双线性 NR-IQA 模型。对合成失真图像基于已有数据库 Waterloo Exploration Database(Ma 等,2017a)和 PASCAL VOC 2012(pattern analysis, statistical modelling and computational learning visual object classes 2012)(Everingham 等,2014)构建预训练数据集,使用已知的失真类型及失真水平作为图像的真实标签,通过多分类任务预训练 CNN 网络。采用预训练的 VGG16 (visual geometry group 16-layer net)模型(Simonyan 和 Zisserman,2015)提取失真图像特征。最后使用双线性汇合(bilinear pooling)技术融合两部分特征图,再通过全连接层进行质量预测。

上述模型(Kang 等,2014; Kim 等,2019b; Yan 等,2019c)都是直接学习图像至主观分数之间的映射,除此之外,研究者提出将图像映射为质量感知的特征向量解决 NR-IQA 问题。Talebi 和 Milanfar(2018)提出了一个用于图像美感评估及质量评价的新模型,称为 NIMA(neural image assessment)。NIMA 模型并不是将图像划分为高质量或低质量或者直接回归得到质量分数,而是使用 CNN 提取图像的高级特征,为每幅图像生成分数范围为 1 ~ 10 的

评分直方图,并计算图像属于不同评分的可能性,让模型更加贴近人类评分。Li 等人(2019a)分析了图像语义信息对质量评价的重要作用,提出了基于语义特征聚合的 NR-IQA 方法。模型以图像块为输入,先使用 ResNet(residual net)(He 等,2016)提取高层语义特征图,然后对特征图进行全局池化得到一组特征向量,计算同一维度特征的均值及方差、分位数、矩等统计学特征,并对图像块特征进行融合,最后利用局部最小二乘回归(partial least squares regression,PLSR)对3个特征进行降维和学习,得到相应的3个质量分数后进行平均,得到最终图像的质量分数。

除了直接端到端学习图像至质量分数之间映射的模型(Kang 等,2014; Kim 等,2019b; Yan 等,2019c; Li 等,2019a),另一类模型包含多个模块,如相似性图预测模块和质量分数预测模块等。Pan 等人(2018)提出了一个基于深度 CNN 的 NR-IQA 模型。该模型主要分为生成网络和池化网络两部分。生成网络为 U-Net 结构(Ronneberger 等,2015),网络以失真图像块为输入,然后经过4个上采样和4个下采样操作,对应大小的特征图使用短连接(skip connection)相关联,采用多种 FR-IQA 方法生成不同的相似度图,并分别作为标签引导网络生成不同的质量图。同时,将生成网络生成的不同类型的质量图融合,然后输入池化网络中进行回归,使用人类主观评分作为标签指导训练。Lin 和 Wang(2018)提出基于生成对抗网络(generative adversarial network,GAN)的 NR-IQA 模型,该模型由质量感知生成网络、判别网络和质量回归网络3部分组成。生成网络以失真图作为输入,目标是生成一幅伪参考图,判别网络的目标是区分生成的伪参考图和真实参考图。质量预测网络以失真图及伪参考图的差异图作为输入,预测质量分数。与上述模型(Lin 和 Wang,2018)类似,Li 等人(2019b)提出的模型包含差异图生成和特征提取两个模块,该模型先使用深度神经网络(deep neural network,DNN)生成差异图,然后再使用传统方法从差异图中提取空间熵和光谱等锐度感知特征,基于感知特征训练 NR-IQA 模型。

上述基于深度学习的 NR-IQA 模型均为单一任务,有研究将 NR-IQA 任务分成多个子任务,即利用其他任务辅助 IQA 任务,提高 NR-IQA 模型的性能。Kang 等人(2015)基于 Kang 等人(2014)模型提出

一个紧凑的多任务 CNN 模型,在原有网络模型(Kang 等,2014)的基础上,增加了1个卷积层并减小卷积核感受野。该模型的全连接层后接1个线性回归层和1个逻辑回归层,分别用于估计图像质量分数和识别失真。其中,质量分数估计为主要任务,使用 L1 范数进行参数调整;失真识别为次要任务,使用负对数似然进行损失函数的训练。Ma 等人(2018b)提出一种基于多任务端到端学习的 NR-IQA 模型,称为 MEON(multi-task end-to-end optimized network)。MEON 包括失真识别和质量预测两个子网络,与传统多任务网络训练方法不同,MEON 的训练过程分为两阶段,第1阶段训练失真类型识别子网络,第2阶段基于预训练的共享卷积层参数和第1个子网络的参数训练质量预测任务的参数。卷积层后两路全连接网络分别用于计算输入图像属于某种失真类型的概率及其对应的感知质量分数。最后,融合层将两路特征向量合并,得到图像整体的质量分数。Yan 等人(2019b)提出了一个基于 DNN 多任务学习的 NR-IQA 模型,该模型包含两个任务:自然性预测任务和质量分数预测任务。自然性预测任务用于辅助质量分数预测任务,帮助输入图像更好地映射为最终的质量分数。Yang 等人(2019b)提出了一个基于 DNN 和显著性辅助的 NR-IQA 模型,该模型主要包含两个子任务:视觉显著性预测和图像质量预测,两个子任务在前期共享特征提取器。从框架上看,Yang 等人(2019b)提出的模型与 MEON(Ma 等,2018b)类似,它们的不同之处在于,Yang 等人(2019b)提出的模型使用显著性预测任务代替了原来的失真预测任务。

1.2.2 其他内容图像

针对自然图像的模型直接使用图像和主观分数端到端训练即能获得较好的性能(Kang 等,2014),无需考虑图像的复杂性和其他因素。然而,其他内容图像展现出与自然图像不同的性质,如观看3维图像时产生的双目融合和双目竞争现象(Chen 等,2013a)、HVS 对屏幕内容图像中的文本部分和图像部分敏感性不一致的问题(Yang 等,2015)等,这些特性使得针对自然图像设计的深度学习模型无法直接应用在其他内容图像上。针对其他内容图像的特性,研究人员设计了特定的 NR-IQA 模型。

针对3维图像,Lyu 等人(2016)受双目视觉理论的启发,提出一个基于双目自相似度(binocular

self-similarity, BS) 和双目融合 (binocular integration, BI) 的无参考 3 维图像质量评价方法。其中, BS 指数通过计算左视图和由视差估计图与右视图合成的左视图之间的相似性得到; BI 指数计算过程相对复杂, 首先构建一个预测图像局部区域质量分数的 DNN 模型, 该模型输入为图像块, 用 FR 方法计算的分数作为标签进行训练。然后通过增益控制模型融合左右视图 DNN 预测分数得到 BI 指数。最后, 通过池化 BS 指数和 BI 指数, 获得 3 维图像最终的质量分数。Oh 等人(2017) 建立了一个基于局部到全局特征聚合的深度无参考 3 维图像质量评价模型, 该模型的训练包括两个阶段: 第 1 阶段为局部质量分数预测, 使用 3 维图像的 FR-IQA 模型获取 3 维图像的标签数据, 再利用计算的标签指导基于块的 CNN 模型进行预训练。第 2 阶段为全局质量预测, 在基于块的 CNN 模型上添加局部特征融合层, 然后将局部特征聚合为全局特征, 并通过多层全连接网络转化为预测分数。Zhou 等人(2019) 受 HVS 多层结构及层次间交互现象的启发, 提出了一个 3 维图像质量评价网络模型, 该模型包含 2 个结构相同的子网络, 均包含 5 个卷积层和 2 个全连接层。这 2 个子网络分别提取左右视图的特征, 并在第 2 层和第 5 层分别进行交互操作, 以获取低级双目特征和高级双目特征。最后, 连接左右视图单目特征和高级双目特征向量, 并使用两个全连接层进一步融合特征, 输出 3 维图像的质量分数。Kim 等人(2019a) 提出了一个双目融合深度网络 (binocular fusion deep network, BFN) 用于预测立体图像视觉舒适度。BFN 主要包含空间特征编码、双目特征编码和视觉舒适度预测 3 部分。空间特征编码部分由多个不同尺度的卷积层组成, 用于提取左右视图的不同感受野的空间特征; 双目特征编码部分通过逐层融合左右视图不同尺度的空间特征来学习 3 维图像潜在的双目特征; 视觉舒适度预测部分将学习到的潜在双目特征进行非线性映射, 从而得到最后的视觉舒适度分数。同时, 还定义了一个视差正则化网络, 用于辅助 BFN 的训练。Fang 等人(2019a) 考虑到大脑内在推导机制 (Wu 等, 2013) 和 HVS 中的双目竞争机制 (Chen 等, 2013a), 提出一种基于双通道融合 CNN 的 3 维图像质量评价模型, 该模型的双通道为权值共享网络, 设计思想来源于 VGG16 (Simonyan 和 Zisserman, 2015)。左右视图分别经过单一

通道, 获得相应的特征图, 然后连接左右视图的特征图, 并使用 2 个卷积层进一步融合左右视图信息, 最后通过 2 个全连接层将融合后的特征图转化成质量分数。Yan 等人(2020) 进一步提出了一种基于多尺度特征融合网络的 3 维图像质量评价模型。不同于仅融合高层抽象信息的模型 (Fang 等, 2019a), 该模型 (Yan 等, 2020) 分别融合 3 维图像的低层、中层和高层抽象信息, 再通过 2 个卷积层进一步融合多层信息, 最后通过全连接网络将多层融合信息映射为主观分数。

针对超分辨率图像, Fang 等人(2018c) 提出基于 CNN 的超分辨率图像质量评价模型, 该模型仅包含 2 个卷积层和 3 个全连接层。Zhou 等人(2020a) 提出一种双通道 CNN 的超分辨率图像质量评价模型。双通道包括结构分支和纹理分支, 均由 5 个卷积层和 2 个全连接层组成, 分别输入超分辨率图像的结构图和纹理图, 然后将两个分支的输出连接, 作为后续 2 层全连接网络的输入, 以输出超分辨率图像的主观分数。Yan 等人(2019a) 将超分辨率图像质量评价模型用于辅助图像超分辨率重建, 首先提出一个 FR 超分辨率图像质量评价深度模型, 并将该 FR 模型用于辅助超分辨率重建网络的训练。

针对屏幕图像, Yue 等人(2018) 提出一种双通道的屏幕内容图像质量评价网络, 两个通道分别输入归一化图和内在推导机制 (internal generative mechanism, IGM) (Wu 等, 2013) 分解非预测部分, 输入大小均设置为 227×227 像素。经过两个通道的处理, 可分别得到 1 个特征向量, 然后将 2 个特征向量连接, 再通过 2 个全连接层网络处理, 输出屏幕图像的分数。Jiang 等人(2020) 提出一种基于孪生网络 (siamese network) 的屏幕图像质量评价模型, 整体框架通过排序学习指导子网络判别能力的学习。子网络为一个多任务学习网络, 包括失真分类子模块和分数预测子模块。训练时, 排序任务、分类任务和预测任务共同学习, 测试时, 只需要输出预测任务的值。

1.2.3 小结

总体而言, 基于深度学习的 NR-IQA 模型在大多数数据库上的表现都超过了基于传统机器学习的模型, 性能接近饱和。比如, Yan 等人(2019b) 提出的模型在 LIVE 2D 数据库 (Sheikh 等, 2006) 上测试性能, 皮尔逊线性相关系数达到了 0.984; Zhou 等

人(2019)提出的模型在 LIVE 3D Phase I 数据库 (Moorthy 等,2013)上测试性能,皮尔逊线性相关系数达到了 0.973。

由于大部分公开主观数据库的数据量较小,如 LIVE 2D 数据库 (Sheikh 等,2006)仅含 29 幅参考图像和 779 幅失真图像,数据量较小的训练集很可能造成过拟合问题,这对模型的可扩展性将造成较大的负面影响。

2 无监督学习的 NR-IQA 模型

目前,针对无监督学习的 NR-IQA 模型的研究相对较少。无监督学习的 NR-IQA 模型指的是模型训练过程中不需要主观数据。主要包括针对自然图像内容的模型和针对其他内容图像的模型。

2.1 基于传统机器学习的无监督 NR-IQA 模型

基于传统机器学习的无监督 NR-IQA 模型致力于构建有效的图像质量感知参考系模型。根据感知参考系构建方式的不同,可将模型分为 3 类:基于 NSS 特征的模型、基于伪标签学习的模型和基于失真分类度量的模型,如图 6 所示。

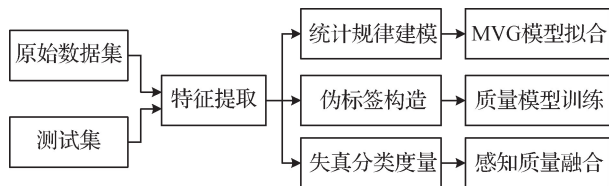


图6 基于传统机器学习的无监督 NR-IQA 模型一般框架

Fig.6 The general framework of the opinion-unaware NR-IQA models based on traditional machine learning algorithms

1) 基于 NSS 特征的 NR-IQA 模型。NIQE (natural image quality evaluator) 是 Mittal 等人 (2013) 提出的一种基于 NSS 特征的无监督 IQA 模型,首先使用质量完好的图像块的 NSS 特征进行拟合,得到多元高斯 (multivariate Gaussian, MVG) 模型,然后利用统计规律进行失真偏差的度量,以实现对图像的质量预测。基于 NIQE (Mittal 等, 2013), Zhang 等人 (2015) 通过整合结构统计特征、多尺度方向和频率统计特征以及颜色统计特征,提出 IL-NIQE (integrated local NIQE) 算法,将拟合的 MVG 模型作为参考标准,计算预测图像的马氏距离,从而得到图像的质量预测分数。类似于 NIQE (Mittal 等, 2013) 和

IL-NIQE (Zhang 等, 2015), Liu 等人 (2020) 提出了一种基于自由能法则的无监督 NR-IQA 模型,对原始图像降质过程中产生的结构、自然性和感知质量变化进行测量,从而进一步量化图像的质量。Wu 等人 (2015) 提出基于局部模式统计的 NR-IQA 方法。为了降低计算复杂度,该方法仅采用图像像素的 4 个邻域来计算局部二值编码,共包含 6 种不同的二值模式。同时,通过对自然图像中的失真进一步分析,发现局部二值模式具有较好的失真识别能力,进而选择与 6 种二值模式之一相关的局部加权统计数据作为图像质量的度量。该类方法通过巧妙构建质量相关统计特征,利用自然图像中的统计规律进行图像质量计算,具有较好的质量预测性能。

2) 基于伪标签学习的 NR-IQA 模型。Xue 等人 (2013) 提出一种基于图像块的质量感知聚类的 NR-IQA 方法,首先通过 FR-IQA 算法 (Zhang 等, 2011) 计算每一个图像块的质量分数,并根据所有图像块质量分数范围划分为 L 大类,然后基于提取的结构特征,通过 K-Means 算法将各大类图像块分成 K 小类。其中,每一小类对应一个聚类中心,即一个图像块 (结构特征和视觉质量分数)。给定失真图像,首先提取图像块,计算图像块与 L 大类中各小类聚类中心的距离。然后融合每大类中距离最小聚类中心的视觉分数,计算得到图像块的分数。最后,使用平均权重策略,得到失真图像的视觉质量分数。Ye 等人 (2014) 提出了一种无监督 IQA 模型 BLISS (blind learning of image quality using synthetic scores), 作者通过将多种有效的 FR-IQA 模型结合,获得单一合成质量分数,用具有高相关性的合成质量分数作为主观分数近似值,以进一步训练相应的 IQA 模型。其中,不同的 FR-IQA 方法计算的分数通过无监督的排序聚集进行融合,且合成得到的质量分数的相应性能超过单一的 FR-IQA 方法的性能。针对屏幕图像,Gu 等人 (2017) 提出一种无参考屏幕图像质量评价方法,通过提取描述图像复杂性、屏幕内容统计、全局亮度和细节清晰度等 4 个方面相关的图像特征,利用现有的高性能全参考屏幕图像质量评价方法计算质量分数,并将该质量分数作为训练数据集标签,得到最终的屏幕图像质量评价模型。

3) 基于失真分类度量的 NR-IQA 模型。传统 NR-IQA 算法通常通过测量失真图像与完美质量图

像(参考图像)的偏差来预测失真图像的质量。与传统 NR-IQA 方法不同,Min 等人(2018)提出“伪”参考图像(pseudo reference image, PRI)的概念,并设计了基于 PRI 的 NR-IQA 模型。其中,作为参考基准的 PRI 由失真图像生成,并假定其失真最严重。针对特定失真,如块效应和锐化等,设计不同的度量方法计算失真图像与 PRI 之间的结构相似性,并将针对特定失真的多种度量方法集成为通用的 NR-IQA 模型。

现实中的图像往往包含多种失真,针对单一特定失真的 IQA 模型难以有效度量多种失真。Zhang 和 Chandler(2018)首先将图像的失真类型进行分类,然后针对不同的失真类型提取 NSS 特征,使用 SVM 训练得到相应的回归模型,估计失真参数并保存,最后计算失真图像的视觉质量分数。Gu 等人(2014)进一步提出了针对单一失真和多重失真的 NR-IQA 模型,称为 SISBLIM(six-step blind metric)。首先通过使用现有的方法估计失真图像中的噪声程度,然后利用 BM3D 方法(Dabov 等,2007)对失真图像进行去噪。考虑到不同的失真和由图像内容产生的掩码效应之间的相互作用,采用基于自由能理论的模型对联合效应进行量化,并通过噪声、模糊、JPEG 压缩以及联合效应的加权质量估计的线性组合得到图像的质量分数。

2.2 基于深度学习的无监督 NR-IQA 模型

基于传统机器学习算法的无监督 NR-IQA 模型通过构建统计模型度量图像质量的差异,基于深度学习的无监督 NR-IQA 模型通过构造有效训练数据,构建神经网络模拟排序信息的学习过程,实现图像与质量分数之间的映射,基本框架如图 7 所示。该类方法一定程度上解决了数据驱动模型中训练数据匮乏的问题,且性能远高于基于传统机器学习的 NR-IQA 模型。

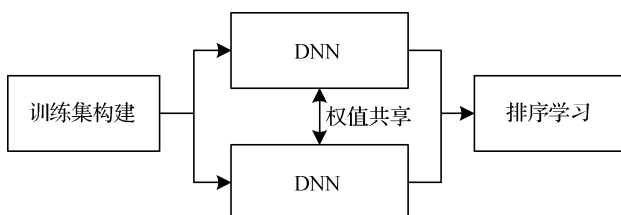


图 7 基于深度学习的无监督 NR-IQA 模型一般框架

Fig. 7 The general framework of the opinion-unaware NR-IQA models based on deep learning algorithms

2017 年, Liu 等人(2017)提出一种基于排序学习的 NR-IQA 方法,先利用构建的包含人工合成失真的排序图像集(Gao 等,2015)进行模型预训练,通过设计权值共享网络实现根据图像质量对图像进行排序,然后利用现有 IQA 数据库中的图像和相应主观质量分数对整个网络进行微调,得到最终的 IQA 模型。同时,提出了在权重共享网络中进行有效反向传播的方法。与传统标准采样和硬负采样策略相比,该优化方法具有收敛速度快、损失小的优点。Ma 等人(2017b)设计了基于权重共享的全连接神经网络 IQA 模型,该模型输入为图像对,采用 CORNIA(Ye 等,2012)提取图像的视觉特征,通过融合图像视觉差异的不确定性及图像视觉质量优劣训练 NR-IQA 模型。Ma 等人(2019)提出一个基于多标签学习的 NR-IQA 模型,通过对收集的高清图像添加若干种类的人工合成失真得到失真图像数据集,同时利用若干种现有 IQA 方法对失真图像数据集进行质量分数预测,即每幅失真图像都包含对应的若干种 IQA 方法计算得到的质量分数;再通过图像对生成机制生成大量图像对,图像对对应的标签通过两两比较若干 IQA 方法计算的质量分数值得到,若图像对中图像 1 的 IQA 质量分数高于图像 2 的 IQA 质量分数(默认 IQA 质量分数越高,图像质量越好),则图像对标签为 1,反之为 0;通过成对排序学习,利用生成的大量图像对训练集对网络模型进行训练以得到最终的 NR-IQA 模型。该方法提出的训练图像数据扩增方法,在一定程度上解决了现有 IQA 模型因图像数据不足而可能导致模型过拟合的问题,且所提算法可高效结合现有 IQA 方法,模型简单有效,可扩展能力强。

2.3 小结

与基于传统机器学习的有监督 NR-IQA 模型类似,基于传统机器学习的无监督 NR-IQA 模型同样依赖手工特征提取,其性能也受到特征表达能力的影响。同时,常用的使用 FR-IQA 模型获取图像伪标签的方式存在明显缺陷。另外,有限的训练数据内容(Mittal 等,2013)也是影响这类模型可扩展能力的关键因素。

基于深度学习的无监督 NR-IQA 模型(Liu 等,2017)可以充分利用海量的无标签数据,通过对比的方式或者借用 FR-IQA 模型产生伪标签(Ma 等,2019),并用于模型的训练。该方式可有效提升算

法性能和算法扩展性,值得进一步研究。

3 新型媒体中的图像质量评价研究

上述基于有监督学习和基于无监督学习的 NR-IQA 模型均针对较为常见的图像内容设计。针对虚拟现实(virtual reality, VR)图像、光场(light field)图像和水下声呐(underwater sonar)图像等新型媒体数据的 NR-IQA 研究逐渐引起关注。

全景图像作为 VR 技术中重要的信息载体之一,与传统 2D 图像存在较大差异,其存储与传输更加困难,容易引入不同的失真从而影响全景图像的质量。在沉浸式的视觉环境下,人们对全景图像的视觉观看体验受到众多因素的影响,如起始点、观看时间以及观看路径等,并且目前针对全景图像质量评价的相关工作较少(Xu 等,2020a;Sui 等,2020)。最近,深度学习技术也应用于全景图像质量评价。Sun 等人(2020)提出一种基于多通道卷积神经网络的无参考全景图像质量评价方法,通过将全景图像分别投影到 6 个视点图以保持与 VR 设备中场景内容观看一致。该方法模型由多通道卷积神经网络和图像质量回归两部分内容组成,其中多通道卷积神经网络部分用于有效提取图像中的相应特征,而质量回归部分则用于融合所提特征并进行特征到质量分数的映射学习。Kim 等人(2020)提出基于对抗神经网络的全景图像质量评价方法,将全景图像分割成若干图像块并估计其局部质量和相应权重,通过融合局部权重与局部质量得到最终的全景图像质量。Xu 等人(2020b)提出基于视点导向的图神经网络的无参考全景图像质量评价方法,通过选择具有较高可见概率的选定视图作为图节点,建立全景图像各视点间的相互依赖关系,并经过图卷积网络对所提图像进行推理。另外,在不进行视口采样的情况下,可同时利用整体全景图像直接获得全局质量,从而根据观看体验提高性能。

光场记录了光线的分布,包含更为丰富的场景信息。近年来提出了许多光场处理技术,但这些操作一定程度上会带来视觉质量的损失,为了对光场图像的视觉质量进行有效评估,一些方法相继提出(Fang 等,2018a)。Shi 等人(2019a,b)提出了光场图像无参考质量评估方法,利用双目视觉特征度量空间质量和角度一致性。并在该方法的基础上,进

一步提出了基于张量理论的无参考光场图像质量评价方法(Zhou 等,2020b),通过考虑光场图像的全局自然性、局部频率特性和各视图间的结构相似性分布,设计度量光场图像空间质量的主成分空间特征和角度一致性质量的张量角变化指数。

在水下探测任务中,由于可见光的穿透深度有限,传统光学成像和分析方法不能有效适用,声呐成像成为首选替代方法。而复杂动态的水下环境的捕获和传输条件不可避免地导致声呐图像的视觉质量下降,并阻碍进一步的识别、分析与理解。为了有效测量声呐图像质量,研究者提出一些可靠的质量指标。Kalwa 和 Madsen(2004)提出了一种前视声呐图像质量评估方法,通过提取各种低层特征来确定声呐图像质量是否足以实现可靠的障碍物识别。Chen 等人(2019)提出一种新的用于感知声呐图像质量的无参考轮廓退化测量方法,通过从频率域和空间域提取描述轮廓信息的变换系数矩阵的稀疏性作为特征,计算滤波前后提取的特征的比值来测量声呐图像的轮廓退化程度。最终,通过学习基于特征聚合的支持向量回归模块来获取轮廓退化程度与声呐图像质量的关系。

4 图像质量评价模型的应用

IQA 模型的应用主要包括参数选择和模型优化。参数选择是指借助 IQA 模型在参数空间自动寻找最优参数序列;模型优化是指利用 IQA 模型去优化其他模型,如超分辨率重建、去噪和图像复原等,目的是让其他模型的输出质量更好。模型优化应用中需要 IQA 模型必须可导,而 NR-IQA 往往不符合该要求,在模型优化中主要采用 FR-IQA 模型。

1) 参数选择。Fang 等人(2017a)通过感知相似性方法(Fang 等,2014)迭代优化,选择最优的重定向算子序列。该研究中,重定向算子包括线裁剪、裁剪、变形和缩放,感知相似性方法是一种针对重定向图像设计的 IQA 算法。通过迭代寻优,最终可获得一个最优的重定向算子序列 $\{O_1, O_2, \dots, O_n\}$, O 表示重定向算子, n 表示迭代次数。与 Fang 等人(2017a)寻找最优操作组合的方法不同,Fang 等人(2020b)尝试使用 IQA 模型选择多曝光图像融合算法最优参数。

2) 模型优化。Ma 等人(2015)提出一种色调映射算法,通过交替迭代优化目标—结构保真度和统计自然性,将高动态范围图像转化为低动态范围图像。类似地, Ma 等人(2018a)使用一种多曝光融合图像质量评价模型优化多曝光图像融合算法。这两个模型优化算法均使用梯度下降法,一次完整的迭代过程针对同一幅图像,比较耗时。针对算法耗时问题, Ma 等人(2020b)提出一种基于 CNN 的快速多曝光图像融合算法,核心部分是一个语义聚合网络(Chen 等, 2017a),通过多曝光融合图像质量评价模型作为目标函数优化网络。IQA 方法也广泛应用于其他计算机视觉研究,如图像去噪(Guo 等, 2019)、超分辨率重建(Yan 等, 2019a)和图像修复(Zhao 等, 2017; Ding 等, 2020)等。

5 评价指标

IQA 模型的性能通常从预测准确性、单调性和一致性 3 个方面进行评估,对应于皮尔逊线性相关系数(Pearson linear correlation coefficient, PLCC)、斯皮尔曼等级相关系数(Spearman rank-order correlation coefficient, SRCC)和均方根误差(root mean square error, RMSE) 3 个评价指标(Video Quality Experts Group(VQEG), 2000)。

PLCC 用于评估 IQA 模型预测的准确性。在计算 PLCC 前,需对客观分数和主观分数进行非线性回归操作。用于非线性回归的 logistic 函数计算为

$$p(Q) = \beta_1 \left[\frac{1}{2} - \frac{1}{1 + e^{(\beta_2(Q - \beta_3))}} \right] + \beta_4 Q + \beta_5 \quad (8)$$

式中, Q 表示原始的客观质量分数; p 为经过回归操作的客观质量分数; $\beta_1, \beta_2, \beta_3, \beta_4$ 和 β_5 为模型参数。PLCC 计算为

$$PLCC = \frac{\sum_{i=1}^N (s_i - \bar{s})(p_i - \bar{p})}{\sqrt{\sum_{i=1}^N (s_i - \bar{s})^2 \sum_{i=1}^N (p_i - \bar{p})^2}} \quad (9)$$

式中, s_i 和 p_i 分别表示第 i 幅图像的主观质量分数及客观质量分数, $\bar{s}$ 和 $\bar{p}$ 分别表示主观质量分数平均值和客观质量分数平均值。

SRCC 用于衡量 IQA 模型预测结果的单调性,其计算式为

$$SRCC = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \quad (10)$$

式中, N 表示样本数量; d_i 表示第 i 幅图像主观质量分数排名与客观质量分数排名的差值。

RMSE 用于评估 IQA 模型预测的一致性,其计算式为

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^N (s_i - p_i)^2} \quad (11)$$

Ma 等人(2017a)提出了 3 种新的评价指标,分别为原始图/失真图可辨别性测试(D-test)、基于序列排序一致性测试(L-test)和配对偏好一致性测试(P-test)。

D-test 将原始图和失真图视为两类,量化 IQA 模型区分原始图和失真图的能力。给定一个数据集,令 q_i 表示 IQA 模型预测的第 i 幅图像的分数,原始图像和失真图像的索引分别为 S_p 和 S_d 。基于模型预测结果,可以找到最优阈值 T ,使正确分类率最大化,分类率的计算式为

$$D = R(T) = \frac{1}{2} \left(\frac{|S_p \cap S_{p*}|}{|S_p|} + \frac{|S_d \cap S_{d*}|}{|S_d|} \right) \quad (12)$$

式中, $S_{p*} = \{i | q_i > T\}$, $S_{d*} = \{i | q_i \leq T\}$, D 取值范围为 $[0, 1]$, 值越大,表示 IQA 模型的可分性越好。

L-test 的目的是评估 IQA 模型在对具有相同内容、相同失真类型但不同失真程度的图像进行评级时的鲁棒性。假设图像质量在任意失真类型下随失真水平的增加而单调下降。若一个数据库包含 M 幅图像、 N 种失真类型和 L 种失真水平,使用平均 SRCC 系数来衡量排序的一致性,具体计算式为

$$L_s = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N SRCC(l_{ij}, p_{ij}) \quad (13)$$

式中, l_{ij} 和 p_{ij} 分别表示第 i 幅失真类型为 j 的失真图像的失真水平和模型预测分数。

P-test 比较了 IQA 模型在质量可分的图像对(quality-discriminable image pairs, DIPs)上的偏好预测。在图像对的偏好预测上,一个好的 IQA 模型应该与 DIPs 具有一致性。假设 DIPs 的数量为 Y ,模型预测正确的 DIPs 数量为 Y_c ,配对偏好一致性比率的计算式为

$$P = \frac{Y_c}{Y} \quad (14)$$

式中,指标 P 越接近 1,表示 IQA 模型的性能越好。

6 常用的数据库

6.1 自然图像数据库

自然图像数据库主要包括 5 个公开的合成失真数据库和 4 个真实失真数据库。

6.1.1 合成失真数据库

1) IVC (images and video-communications) 数据库 (le Callet 和 Autrusseau, 2005) 由法国南特大学的 IRCCyN (Institut de Recherche en Communications et Cybernétique de Nantes) 实验室在 2005 年建立,包含 10 幅参考图像和 235 幅失真图像,失真类型包括 JPEG 压缩、JPEG2000 压缩、局部自适应分辨率编码压缩和高斯模糊失真等 4 种,每种失真类型包括 5 个失真等级。数据库中 185 幅失真图像的平均主观得分 (mean opinion score, MOS) 由 15 位实验者主观评分统计获得。其中, MOS 的取值范围为 $[1, 5]$, 值越大说明图像质量越好。

2) LIVE (Laboratory for Image and Video Engineering) 数据库 (Sheikh 等, 2006) 由美国奥斯汀大学德州分校 LIVE 实验室在 2006 年建立,包含 29 幅参考图像和 779 幅失真图像,失真类型包括 JPEG 压缩、JPEG2000 压缩、白噪声、高斯模糊和 JPEG2000 快速尺度衰落失真等 5 类,每种失真类型包括 5~6 个失真等级。数据库提供每个失真图像对应的平均主观得分差 (differential mean opinion score, DMOS), 大约 25 000 个统计数据,由 161 个实验者测试得到。DMOS 的取值范围为 $[0, 100]$, 值越小说明图像质量越好。

3) TID2008 (Tampere image database 2008) 数据库 (Ponomarenko 等, 2009) 由芬兰坦佩雷理工大学在 2008 年建立,包含 25 幅参考图像和 1 700 幅失真图像,失真类型包括 JPGE 压缩、JPGE2000 压缩、加性高斯噪声、加性噪声、高斯模糊和对比度变换等 17 类,每种失真类型包括 4 个失真等级。数据库提供图像的 MOS 值及对应的标准差, MOS 的取值范围为 $[0, 9]$ 。

4) CSIQ (categorical subjective image quality) 数据库 (Larson 和 Chandler, 2009) 由美国俄克拉荷马州大学于 2009 年建立,包含 30 幅参考图像和 866 幅失真图像,失真类型包括对比度退化、加性高斯彩噪

声、高斯模糊、高斯白噪声、JPEG 和 JPEG2000 等 6 类,每种失真类型包括 3~5 个失真等级。数据库提供测试图像的 DMOS 值,其中,约 5 000 个数据由 25 位实验者测试统计获得,数据库 DMOS 的取值范围为 $[0, 1]$ 。

5) TID2013 (Tampere image database 2013) 数据库 (Ponomarenko 等, 2015) 由芬兰坦佩雷理工大学于 2013 年建立,包含 25 幅参考图像和 3 000 幅失真图像,失真类型包括 JPEG 压缩、JPEG2000 压缩、对比度变换、颜色饱和度变换、加性高斯噪声、加性噪声、高频噪声和高斯模糊等 24 类,每种失真类型包括 5 个失真等级。数据库提供测试图像的 MOS 值,其中,524 340 个数据由 971 位实验者测试统计得到,数据库 MOS 的取值范围为 $[0, 9]$ 。

6.1.2 真实失真数据库

1) LIVE Challenge 数据库 (Ghadiyaram 和 Bovik, 2016) 由美国奥斯汀大学德州分校的 LIVE 实验室于 2016 年建立,包含 1 162 幅真实失真图像,主观分数通过众包的方式收集,受试者超过 8 100 人,收集的主观分数超过 350 000, MOS 的取值范围为 $[3.42, 92.43]$ 。

2) KonIQ-10k (Konstanz authentic image quality database) 数据库 (Hosu 等, 2020) 由德国康斯坦茨大学于 2020 年建立,包含 10 073 幅真实失真图像,主观分数通过众包的方式获得,标注者为 1 459 人,收集的主观数据包括 120 万。同时,该数据库还提供了图像的属性和 EXIF (exchangeable image file format) 信息。

3) SPAQ (smartphone photography attribute and quality) 数据库 (Fang 等, 2020c) 由江西财经大学于 2020 年建立,主要针对手机成像质量评价,包含来自 66 个不同成像设备收集的 11 125 幅图像。主观实验在标准实验室环境下进行, MOS 的取值范围在 $[0, 100]$ 。同时,该数据库提供图像属性信息、类别信息和 EXIF 信息。

4) 美国奥斯汀大学德州分校 LIVE 实验室在 2020 年建立的数据库 (Ying 等, 2020), 包含 39 810 幅图像,将近 12 万个图像块。主观分数通过众包的方式获得,受试者共 7 865 人。

6.2 其他内容图像数据库

6.2.1 3 维图像数据库

1) LIVE 3D Image Quality Database Phase I 数

据库 (Moorthy 等, 2013) 和 Phase II 数据库 (Chen 等, 2013a, b) 由美国奥斯汀大学德州分校的 LIVE 实验室在 2013 年建立。LIVE 3D Phase I 数据库由 20 幅参考图像和 365 幅对称失真图像构成, 失真类型包括 JPEG2000 压缩、JPEG 压缩、加性高斯白噪声、快衰落和高斯模糊等 5 类。其中, JPEG2000 压缩、JPEG 压缩、高斯白噪声和快衰落各对应 80 幅失真图像, 高斯模糊对应 45 幅失真图像。数据库以 DMOS 的形式提供每幅失真图像的人类主观分数, 值越低说明图像质量越好。除此之外, 数据库还提供每幅参考图像对应的深度图及视差图。LIVE 3D Phase II 数据库由 8 幅参考图像和 360 幅失真图像构成, 图像的失真类型与 Phase I 数据库一致, 每种失真类型对应 72 幅图像。与 Phase I 数据库不同, 该数据库同时提供了对称和非对称失真的 3 维图像。其中, 对称失真图像 240 幅, 非对称失真图像 120 幅。

2) Waterloo IVC (image and vision computing) 3D Phase I 和 Phase II 图像质量数据库 (Wang 等, 2015) 由加拿大滑铁卢大学于 2015 年建立。Waterloo IVC 3D Phase I 数据库由 78 幅失真单视图和 330 幅 3 维图像构成, 失真图像由 6 幅原始 3 维图像生成, 失真类型包括加性白色高斯噪声、高斯模糊和 JPEG 压缩等 3 类, 每种失真类型对应 4 个失真等级。Waterloo IVC 3D Phase II 数据库由 130 幅失真单视图和 460 幅失真 3 维图像构成, 失真图像由 10 幅原始 3 维图像生成, 图像的失真类型及失真程度与 Phase I 数据库一致。

6.2.2 超分辨率图像数据库

1) 上海交通大学在 2017 年建立的超分辨率图像数据库 (Ma 等, 2017c) 包括 1 620 幅由低分辨率图像通过超分辨率算法获得的超分辨率图像、30 幅真实高分辨率图像和 60 幅低分辨率图像, MOS 的范围为 $[0, 10]$ 。

2) SRID (super-resolution reconstructed image database) 数据库由西安电子科技大学在 2017 年建立 (Wang 等, 2017), 包含 480 幅失真超分辨率图像, 这些图像由低分辨率图像使用 2 种插值方法和 6 种超分辨率增强算法获得, 数据库没有提供未失真的超分辨率图像, 图像的主观质量分数以 MOS 的形式提供, MOS 的取值范围为 $[0, 10]$ 。

6.2.3 色调映射图像数据库

1) ESPL-LIVE HDR (embedded signal processing

laboratory-laboratory for image and video engineering, high dynamic range) 数据库 (Kundu 等, 2017) 由美国奥斯汀大学德州分校的 LIVE 实验室在 2017 年建立, 包含来自 605 个高质量 HDR 场景使用 3 种 HDR 处理方法获得的 1 811 幅色调映射图像。图像使用 4 种色调映射算子、5 种多曝光融合方法和 2 个后处理效应生成。数据库图像的主观分数通过众包平台大量的主观实验获得。

2) TMID (tone mapped image database) 数据库 (Yeganeh 和 Wang, 2013) 由加拿大滑铁卢大学在 2013 年建立, 包含 15 幅 HDR 图像及 120 幅由 8 个不同的色调映射算子生成的色调映射图像, 每幅色调映射图像对应一个 MOS 分数, MOS 分数的范围为 $[1, 8]$ 。

6.2.4 多曝光图像融合数据库

Deghosting IQA 数据库 (Fang 等, 2020b) 由江西财经大学在 2018 年建立, 包含 20 个动态自然场景的多曝光图像序列及由 9 种多曝光融合算法生成的 180 幅融合图像。数据库采用 2AFC (two-alternative forced choice) 方法进行主观分数收集。共 60 位受试者参与主观实验, 融合图像的主观分数以 MOS 形式提供。

6.2.5 合成图像数据库

ESPL 数据库 (Kundu 和 Evans, 2014; Kundu 等, 2018) 由美国奥斯汀大学德州分校的 LIVE 实验室在 2014 年和 2017 年建立。ESPL 数据库 1.0 版本 (Kundu 和 Evans, 2014) 包括从视频游戏和动画电影中选取的 222 幅高质量合成图像, 失真类型共 9 种, 每种失真对应 5 个失真程度, 整个数据库共提供 10 212 幅合成图像。ESPL 数据库 2.0 版本 (Kundu 等, 2018) 包括 25 幅从 1.0 版本数据库中选取的高质量彩色合成图像, 每幅图像分别添加 5 种失真, 每种失真对应 4 个失真等级, 生成 20 幅失真图像, 进而获得 500 幅失真图像。数据库采用单刺激连续评价方法进行主观实验, 64 名受试者参与主观实验, 最终采用 52 名受试者的评分作为最终主观分数, 主观分数以 DMOS 形式提供。

6.2.6 屏幕内容图像数据库

1) SIQAD (screen image quality assessment database) 数据库 (Yang 等, 2015) 由新加坡南洋理工大学于 2015 年建立, 包含 20 幅参考图像和 980 幅失真图像, 失真类型包括高斯噪声、高斯模糊、运动模

糊、对比度模糊、JPEG 压缩、JPEG2000 压缩和基于层分解的压缩等 7 类, 每种失真对应 7 个失真等级。

2) SCID (screen content image database) 数据库 (Ni 等, 2017) 由华侨大学于 2017 年建立, 包含 40 幅参考图像和 1 800 幅失真图像, 失真类型包括高斯噪声、高斯模糊、运动模糊、对比度模糊、颜色对比度变化、颜色量化失真、JPEG 压缩、JPEG2000 压缩和基于层分解的压缩等 9 类, 每种失真对应 5 个失真等级。

6.3 新型媒体数据内容图像数据库

6.3.1 虚拟现实数据库

1) OIQA (omnidirectional IQA) 数据库 (Duan 等, 2018) 由上海交通大学于 2018 年建立, 包含 16 幅全景参考图像和 320 幅全景失真图像, 等距投影格式, 失真类型包括 JPEG 压缩、JPEG2000 压缩、高斯噪声和高斯模糊等 4 类, 每种失真对应 5 种失真程度。所有参考图像由专业全景相机获得, 共 20 位受试者参与主观实验, 实验过程中记录受试者的头动和眼动数据, 且采用 10 分等级制进行图像质量评分, 分数越高代表图像质量越好, 数据库的主观分数为 MOS 形式, 取值范围为 [1, 10]。

2) CVIQD (compressed 360-degree image database) 数据库 (Sun 等, 2018) 由上海交通大学于 2018 年建立, 包含 16 幅全景参考图像和 528 幅全景失真图像, 失真类型包括 JPEG 压缩、H. 264 压缩和 H. 265 压缩等 3 类。所有全景图像由全景相机拍摄得到, 共 20 位受试者进行主观实验, 实验采用单刺激法进行主观评分。数据库主观分数以 MOS 形式提供, 分数范围为 [0, 100]。

3) 南京大学在 2018 年建立的数据库 (Huang 等, 2018) 旨在研究空间分辨率和 JPEG 压缩对全景图像质量的共同影响关系, 包含 12 幅全景参考图像和 144 幅全景失真图像, 图像的分辨率包含 720 p、1 080 p、2 K 和 4 K 等 4 种, 通过对参考图像进行不同压缩率的 JPEG 压缩得到一定数量的失真图像。共 98 位受试者参与主观实验, 实验过程中采用单刺激法进行主观评分。数据库主观分数以 MOS 形式提供, 取值范围为 [0, 100]。

6.3.2 光场图像数据库

1) Win5-LID (windowed 5 degree of freedom light field image database) 数据库 (Shi 等, 2018) 由中国科学技术大学于 2018 年建立, 包含 10 幅参考图像和

220 幅失真图像, 由 6 个真实场景和 4 个合成场景组成, 失真类型包括 HEVC 编码、JPEG 压缩、线性插值、近邻插值和两种神经网络模型处理得到的类型等 6 类。数据库采用双刺激连续评价方法进行主观实验, 共 20 名受试者参与实验, 受试者的评分作为最终主观分数, 主观分数以 MOS 形式提供。

2) MPI-LFA (Max Planck Institut Informatik-Light Field Archive) 数据库 (Adhikarla 等, 2017) 由德国 MPI 在 2017 年建立, 包含 14 幅参考图像和 336 幅失真图像, 失真类型包括 HEVC 编码、深度图量化、基于光流估计的图像变形、线性插值、近邻插值和高斯模糊等 6 类。数据库采用双选择两两比较的方法进行主观实验, 40 名受试者参与实验, 主观分数以 DMOS 的形式提供。

3) 意大利罗马大学在 2017 年建立的光场图像数据库 (Paudyal 等, 2017) 包含 16 幅参考图像和 256 幅失真图像, 失真类型包括 HEVC 编码、JPEG 压缩、JPEG2000 压缩和稀疏视差编码等 4 类, 每种失真对应 4 种失真程度。数据库采用成对比较的方法进行主观分数收集, 19 名受试者参与主观实验, 主观分数以 DMOS 的形式提供, 取值范围为 [1, 12]。

4) VALID (visual quality assessment for light field images dataset) 数据库 (Viola 和 Ebrahimi, 2018) 由瑞士洛桑理工大学于 2018 年建立, 包含 5 幅参考图像和 40 幅失真图像, 失真类型包括 HEVC 编码、VP9 编码和 3 种现有提供光场压缩失真的方法得到的类型等 5 类。数据库采用双刺激可视化评价方法进行主观实验, 图像主观分数以 MOS 的形式提供。

6.3.3 水下声呐图像数据库

SIQD (sonar image quality database) 数据库 (Chen 等, 2017b) 由厦门大学于 2017 年建立, 包含 40 幅参考声呐图像和 800 幅失真图像, 失真类型包括 ComGBR 压缩、分层压缩编码下的无偏失真、ComGBR 编码下比特流失真和分层压缩编码下比特流失真等 4 类, 每种失真类型对应 5 种失真程度。数据库提供的主观分数形式为 MOS。

7 结 语

7.1 当前发展总结

NR-IQA 模型分为基于传统机器学习算法的

NR-IQA 模型和基于深度学习的 NR-IQA 模型。

基于传统机器学习算法的 NR-IQA 模型在小数据集上的表现较好,但性能已趋于饱和。总体而言,该类模型主要存在两个局限:1) 该类模型依赖手工设计的特征,而优异的手工特征设计需要足够的先验知识和实践技巧;2) 传统机器学习算法数据表示能力不足,虽然在小数据集上的表现较好,但当数据集增大时,该类型算法表现较差。在 IQA 模型 (Ma 等, 2020a) 中,基于传统机器学习算法的无监督模型 NIQE 和 IL-NIQE 表现较好,超过了大部分有监督 NR-IQA 模型,说明基于传统机器学习算法的有监督 NR-IQA 模型可能存在过拟合问题。但 NIQE 和 IL-NIQE 在真实失真数据库上表现较差,反映出该类模型较弱的可扩展能力。

基于深度学习的 NR-IQA 模型在各类 IQA 数据库上表现都不错,但在合成失真自然图像和 3 维图像数据库上的表现已趋于饱和,在屏幕图像、色调映射图像及合成图像等其他内容数据库和自然失真图像数据库上的表现还存在一定的上升空间。该类模型的可扩展性主要受限于训练数据不足,主要的解决思路是增加训练数据,可通过成对比较法生成二值标签或者通过可信赖的 FR-IQA 模型生成伪标签。得益于训练数据的大幅增加和 CNN 强大的数据表达能力,这类模型 (Zhang 等, 2020; Ma 等, 2019) 展现出较强的性能及可扩展能力。

7.2 未来发展方向

无参考图像质量评价在评价内容和客观 NR-IQA 模型方面均存在潜在的发展方向。就评价内容而言,未来仍需要投入大量精力对真实失真图像、特定应用场景下的失真图像以及新媒体图像数据进行研究。就 NR-IQA 模型而言,基于无监督学习的 NR-IQA 模型是未来 IQA 领域的一个研究重点。与计算机视觉等领域遇到的问题类似,数据仍然是必须面对的一个问题。

1) 评价内容。真实失真图像 (Fang 等, 2020c; Ghadiyaram 和 Bovik, 2016) 主要是指自然图像,其失真复杂且失真类型未知。现有的 IQA 数据库大多通过人工添加常见噪声生成,且获取的失真图像内容形式较为单一,这些失真图像仅仅来源于数量有限的高质量图像。在真实应用系统中,需要处理的图像/视频数据内容形式复杂多样。并且,这些数据在生成、传输、处理等过程中可能经历了多重失真,

引入的失真无法完全使用计算机进行模拟生成。这不仅限制了现有合成失真数据库的应用价值,还限制了针对人工添加噪声数据设计的客观评价模型的可扩展性。虽然陆续提出了自然失真图像质量评价的解决方案 (Fang 等, 2020c; Zhang 等, 2020), 该类研究仍然具有较大的挑战性,提升空间较大。此外,众所周知,设计通用的 NR-IQA 模型是困难的。同时,在实际应用系统中,由于应用场景存在差异以及需求不一致等情况,使用同一种算法无法较好地预测不同场景下、不同需求下的失真图像质量,如高动态图像的色调映射造成的图像降质 (Fang 等, 2020a)、虚拟视觉合成图像 (Kundu 和 Evans, 2014; Kundu 等, 2018) 的非局部噪声。虽然对特定应用场景下的 NR-IQA 国内外均有相关研究 (Fang 等, 2018b; Chen 等, 2013a; Fang 等, 2020a), 但依然存在失真复杂和主观数据缺乏两个主要问题。总体而言,特定应用场景下的 NR-IQA 模型目前还比较缺乏,针对特定应用场景设计有效的 NR-IQA 模型具有重要意义。同时,针对新型媒体数据的 NR-IQA 研究相对较少,值得进一步探索。

2) NR-IQA 模型。直接端到端学习的 NR-IQA 模型受限于训练数据,测试性能较高,但可扩展能力不强。主要策略包括通过多任务学习辅助 NR-IQA (Ma 等, 2018b) 和通过排序学习增加训练样本 (Zhang 等, 2020; Ma 等, 2019)。然而,现有的多任务学习 NR-IQA 模型主要通过单一任务 (如失真分类、图像显著性等) 改进模型性能,或者基于数据扩增思想 (Zhang 等, 2020; Ma 等, 2019), 如何自适应地融合多任务并设计鲁棒的 NR-IQA 模型,值得进一步研究。

除此之外,基于元学习的 NR-IQA 模型 (Zhu 等, 2020) 设计值得关注,引入迁移学习 (Muandet 等, 2013; Li 等, 2018) 解决 NR-IQA 模型跨域的问题值得探索。

参考文献 (References)

- Adhikarla V K, Vinkler M, Sumin D, Mantiuk R K, Myszkowski K, Seidel H P and Didyk P. 2017. Towards a quality metric for dense light fields//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 3720-3729 [DOI: 10.1109/CVPR.2017.396]

- 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]
- Chen M J, Cormack L K and Bovik A C. 2013a. No-reference quality assessment of natural stereopairs. *IEEE Transactions on Image Processing*, 22(9): 3379-3391 [DOI: 10.1109/TIP.2013.2267393]
- Chen M J, Su C C, Kwon D K, Cormack L K and Bovik A C. 2013b. Full-reference quality assessment of stereopairs accounting for rivalry. *Signal Processing: Image Communication*, 28(9): 1143-1155 [DOI: 10.1016/j.image.2013.05.006]
- Chen Q F, Xu J and Koltun V. 2017a. Fast image processing with fully-convolutional networks//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice, Italy: IEEE: 2516-2525 [DOI: 10.1109/ICCV.2017.273]
- Chen W L, Gu K, Lin W S, Xia Z F, le Callet P and Cheng E. 2019. Reference-free quality assessment of sonar images via contour degradation measurement. *IEEE Transactions on Image Processing*, 28(11): 5336-5351 [DOI: 10.1109/TIP.2019.2910666]
- Chen W L, Yuan F, Cheng E and Lin W S. 2017b. Subjective and objective quality evaluation of sonar images for underwater acoustic transmission//*Proceedings of 2017 IEEE International Conference on Image Processing*. Beijing, China: IEEE: 176-180 [DOI: 10.1109/ICIP.2017.8296266]
- Chen X, Li L D, Li Q Y, Han X X and Zhu H C. 2019. No-reference quality assessment of depth images based on natural scenes statistics. *Computer Science*, 46(6): 256-262 (陈曦, 李雷达, 李巧月, 韩习习, 祝汉城. 2019. 基于自然场景统计的深度图像质量无参考评价方法. *计算机科学*, 46(6): 256-262) [DOI: 10.11896/j.issn.1002-137X.2019.06.038]
- Dabov K, Foi A, Katkovnik V and Egiazarian K. 2007. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8): 2080-2095 [DOI: 10.1109/TIP.2007.901238]
- Dai J F, Li Y, He K M and Sun J. 2016. R-FCN: object detection via region-based fully convolutional networks//*Proceedings of the 30th International Conference on Neural Information Processing Systems*. Barcelona, Spain: ACM: 379-387
- Ding K Y, Ma K D, Wang S Q and Simoncelli E P. 2020. A comparative study of image quality assessment models through perceptual optimization [EB/OL]. [2020-05-08]. <https://arxiv.org/pdf/2005.01338v1.pdf>
- Dong H P and Liu L X. 2014. No-reference image quality assessment in mutual information domain. *Journal of Image and Graphics*, 19(3): 484-492 (董宏平, 刘利雄. 2014. 互信息域中的无参考图像质量评价. *中国图象图形学报*, 19(3): 484-492) [DOI: 10.11834/jig.20140220]
- Dong W S, Wang P Y, Yin W T, Shi G M, Wu F F and Lu X T. 2019. Denoising prior driven deep neural network for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(10): 2305-2318 [DOI: 10.1109/TPAMI.2018.2873610]
- 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//*IEEE International Symposium on Circuits and Systems*. Florence, Italy: IEEE: 1-5 [DOI: 10.1109/ISCAS.2018.8351786]
- Everingham M, Eslami S M A, Van Gool L, Williams C K I, Winn J and Zisserman A. 2014. The pascal visual object classes challenge: a retrospective. *International Journal of Computer Vision*, 111(1): 98-136 [DOI: 10.1007/s11263-014-0733-5]
- Fang Y M, Fang Z J, Yuan F N, Yang Y, Yang S Y and Xiong N N. 2017a. Optimized multioperator image retargeting based on perceptual similarity measure. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 47(11): 2956-2966 [DOI: 10.1109/TSMC.2016.2557225]
- 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, Wei K K, Hou J H, Wen W Y and Imamoglu N. 2018a. Light filed image quality assessment by local and global features of epipolar plane image//*Proceedings of the 4th IEEE International Conference on Multimedia Big Data*. Xi'an, China: IEEE: 1-6 [DOI: 10.1109/BigMM.2018.8499086]
- Fang Y M, Yan J B, Du R G, Zuo Y F, Wen W Y, Zeng Y and Li L D. 2020a. Blind quality assessment for tone-mapped images by analysis of gradient and chromatic statistics [EB/OL]. [2020-05-08]. <https://ieeexplore.ieee.org/document/9082863>
- Fang Y M, Yan J B, Li L D, Wu J J and Lin W S. 2018b. 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. 2017b. 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]
- Fang Y M, Yan J B, Liu X L and Wang J H. 2019a. Stereoscopic image quality assessment by deep convolutional neural network. *Journal of Visual Communication and Image Representation*, 58: 400-406 [DOI: 10.1016/j.jvcir.2018.12.006]
- Fang Y M, Yan J B, Wang J H, Liu X L, Zhai G T and Le Callet P. 2019b. Learning a no-reference quality predictor of stereoscopic images by visual binocular properties. *IEEE Access*, 7: 132649-132661 [DOI: 10.1109/ACCESS.2019.2941112]
- Fang Y M, Yuan Y, Li L D, Wu J J, Lin W S and Li Z Q. 2017c. Performance evaluation of visual tracking algorithms on video sequences with quality degradation. *IEEE Access*, 5: 2430-2441 [DOI: 10.1109/ACCESS.2017.2666218]
- Fang Y M, Zeng K, Wang Z, Lin W S, Fang Z J and Lin C W. 2014. Objective quality assessment for image retargeting based on structural

- similarity. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 4(1): 95-105 [DOI: 10.1109/JETCAS.2014.2298919]
- Fang Y M, Zhang C, Yang W H, Liu J Y and Guo Z M. 2018c. Blind visual quality assessment for image super-resolution by convolutional neural network. *Multimedia Tools and Applications*, 77(22): 29829-29846 [DOI: 10.1007/s11042-018-5805-z]
- Fang Y M, Zhu H W, Ma K D, Wang Z and Li S T. 2020b. Perceptual evaluation for multi-exposure image fusion of dynamic scenes. *IEEE Transactions on Image Processing*, 29: 1127-1138 [DOI: 10.1109/TIP.2019.2940678]
- Fang Y M, Zhu H W, Zeng Y, Ma K D and Wang Z. 2020c. Perceptual quality assessment of smartphone photography//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 3677-3686 [DOI: 10.1109/CVPR42600.2020.00373]
- Gao F and Gao X B. 2014. Active feature learning and its application in blind image quality assessment. *Chinese Journal of Computers*, 37(10): 2227-2234 (高飞, 高新波. 2014. 主动特征学习及其在盲图像质量评价中的应用. *计算机学报*, 37(10): 2227-2234) [DOI: 10.3724/SP.J.1016.2014.02227]
- Gao F, Tao D C, Gao X B and Li X L. 2015. Learning to rank for blind image quality assessment. *IEEE Transactions on Neural Networks and Learning System*, 26(10): 2275-2290 [DOI: 10.1109/TNNLS.2014.2377181]
- Ghadiyaram D and Bovik A C. 2016. Massive online crowdsourced study of subjective and objective picture quality. *IEEE Transactions on Image Processing*, 25(1): 372-387 [DOI: 10.1109/TIP.2015.2500021]
- 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)
- Gu K, Zhai G T, Lin W S, Yang X K and Zhang W J. 2015b. No-reference image sharpness assessment in autoregressive parameter space. *IEEE Transactions on Image Processing*, 24(10): 3218-3231 [DOI: 10.1109/TIP.2015.2439035]
- Gu K, Zhai G T, Yang X K and Zhang W J. 2014. Hybrid no-reference quality metric for singly and multiply distorted images. *IEEE Transactions on Broadcasting*, 60(3): 555-567 [DOI: 10.1109/TBC.2014.2344471]
- Gu K, Zhai G T, Yang X K and Zhang W J. 2015a. Using free energy principle for blind image quality assessment. *IEEE Transactions on Multimedia*, 17(1): 50-63 [DOI: 10.1109/TMM.2014.2373812]
- Gu K, Zhou J, Qiao J F, Zhai G T, Lin W S and Bovik A C. 2017. No-reference quality assessment of screen content pictures. *IEEE Transactions on Image Processing*, 26(8): 4005-4018 [DOI: 10.1109/TIP.2017.2711279]
- Guo S, Yan Z F, Zhang K, Zuo W M and Zhang L. 2019. Toward convolutional blind denoising of real photographs//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, USA: IEEE: 1712-1722 [DOI: 10.1109/CVPR.2019.00181]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- 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]
- Huang M K, Shen Q, Ma Z, Bovik A C, Gupta P, Zhou R B and Cao X. 2018. Modeling the perceptual quality of immersive images rendered on head mounted displays: resolution and compression. *IEEE Transactions on Image Processing*, 27(12): 6039-6050 [DOI: 10.1109/TIP.2018.2865089]
- Jia H Z, Sun Q S and Wang T H. 2014. Blind image quality assessment based on perceptual features and natural scene statistics. *Journal of Image and Graphics*, 19(6): 859-867 (贾惠珍, 孙权森, 王同罕. 2014. 结合感知特征和自然场景统计的无参考图像质量评价. *中国图象图形学报*, 19(6): 859-867) [DOI: 10.11834/jig.20140606]
- Jiang Q P, Shao F and Jiang G Y. 2017. MSFE: blind image quality assessment based on multi-stage feature encoding//*Proceedings of 2017 IEEE International Conference on Image Processing*. Beijing, China: IEEE: 3160-3164 [DOI: 10.1109/ICIP.2017.8296865]
- Jiang X H, Shen L Q, Yu L W, Jiang M X and Feng G R. 2020. Noreference screen content image quality assessment based on multi-region features. *Neurocomputing*, 386: 30-41 [DOI: 10.1016/j.neucom.2019.12.027]
- Kalwa J and Madsen A L. 2004. Sonar image quality assessment for an autonomous underwater vehicle//*Proceedings of 2004 World Automation Congress*. Seville, Spain: IEEE, 2004: 33-38
- 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*. Quebec City, Canada: IEEE: 2791-2795 [DOI: 10.1109/ICIP.2015.7351311]
- Kim H G, Jeong H, Lim H T and Ro Y M. 2019a. Binocular fusion net: deep learning visual comfort assessment for stereoscopic 3D. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(4): 956-967 [DOI: 10.1109/TCSVT.2018.2817250]

- 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]
- Kim J, Nguyen A D and Lee S. 2019b. Deep CNN-based blind image quality predictor. *IEEE Transactions on Neural Networks and Learning Systems*, 30(1): 11-24 [DOI: 10.1109/TNNLS.2018.2829819]
- Kundu D and Evans B L. 2014. Spatial domain synthetic scene statistics//*Proceedings of the 48th Asilomar Conference on Signals, Systems and Computers*. Pacific Grove, USA; IEEE: 948-954 [DOI: 10.1109/ACSSC.2014.7094593]
- Kundu D, Choi L K, Bovik A C and Evans B L. 2018. Perceptual quality evaluation of synthetic pictures distorted by compression and transmission. *Signal Processing: Image Communication*, 61: 54-72 [DOI: 10.1016/j.image.2017.11.004]
- Kundu D, Ghadiyaram D, Bovik A C and Evans B L. 2017. Large-scale crowdsourced study for tone-mapped HDR pictures. *IEEE Transactions on Image Processing*, 26(10): 4725-4740 [DOI: 10.1109/TIP.2017.2713945]
- Larson E C and Chandler D M. 2009. Categorical subjective image quality CSIQ database [DB/OL]. [2020-07-08]. <http://vision.okstate.edu/csiq/>
- le Callet P and Autrusseau F. 2005. Subjective quality assessment IRCyN/IVC database [DB/OL]. [2020-07-08]. <http://www2.irc-cyn.ec-nantes.fr/ivcdb/>
- Li D, Yang Y X, Song Y Z and Hospedales T M. 2018. Learning to generalize: meta-learning for domain generalization//*Proceedings of the 32nd AAAI Conference on Artificial Intelligence*. New Orleans, USA: AAAI Press: 3490-3497
- Li D Q, Jiang T T, Lin W S and Jiang M. 2019a. Which has better visual quality: the clear blue sky or a blurry animal? *IEEE Transactions on Multimedia*, 21(5): 1221-1234 [DOI: 10.1109/TMM.2018.2875354]
- Li L D, Zhou Y, Gu K, Yang Y Z and Fang Y M. 2019b. 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 Q H, Lin W S and Fang Y M. 2016a. No-reference quality assessment for multiply-distorted images in gradient domain. *IEEE Signal Processing Letters*, 23(4): 541-545 [DOI: 10.1109/LSP.2016.2537321]
- Li Q H, Lin W S, Xu J T and Fang Y M. 2016b. Blind image quality assessment using statistical structural and luminance features. *IEEE Transactions on Multimedia*, 18(12): 2457-2469 [DOI: 10.1109/TMM.2016.2601028]
- 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, USA; IEEE: 732-741 [DOI: 10.1109/CVPR.2018.00083]
- 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*. Venice, Italy; IEEE: 1040-1049 [DOI: 10.1109/ICCV.2017.118]
- Liu Y T, Gu K, Zhang Y B, Li X, Zhai G T, Zhao D B and Gao W. 2020. Unsupervised blind image quality evaluation via statistical measurements of structure, naturalness, and perception. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(4): 929-943 [DOI: 10.1109/TCSVT.2019.2900472]
- Lyu Y Q, Yu M, Jiang G Y, Shao F, Peng Z J and Chen F. 2016. Noreference stereoscopic image quality assessment using binocular self-similarity and deep neural network. *Signal Processing: Image Communication*, 47: 346-357 [DOI: 10.1016/j.image.2016.07.003]
- Ma C, Yang C Y, Yang X K and Yang M H. 2017c. Learning a no-reference quality metric for single-image super-resolution. *Computer Vision and Image Understanding*, 158: 1-16 [DOI: 10.1016/j.cviu.2016.12.009]
- Ma K D, Duanmu Z F, Wang Z, Wu Q B, Liu W T, Yong H W, Li H L and Zhang L. 2020a. Group maximum differentiation competition; model comparison with few samples. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4): 851-864 [DOI: 10.1109/TPAMI.2018.2889948]
- Ma K D, Duanmu Z F, Wu Q B, Wang Z, Yong H W, Li H L and Zhang L. 2017a. Waterloo exploration database; new challenges for image quality assessment models. *IEEE Transactions on Image Processing*, 26(2): 1004-1016 [DOI: 10.1109/TIP.2016.2631888]
- Ma K D, Duanmu Z F, Yeganeh H and Wang Z. 2018a. Multi-exposure image fusion by optimizing a structural similarity index. *IEEE Transactions on Computational Imaging*, 4(1): 60-72 [DOI: 10.1109/TCI.2017.2786138]
- Ma K D, Duanmu Z F, Zhu H W, Fang Y M and Wang Z. 2020b. Deep guided learning for fast multi-exposure image fusion. *IEEE Transactions on Image Processing*, 29: 2808-2819 [DOI: 10.1109/TIP.2019.2952716]
- Ma K D, Liu W T, Liu T L, Wang Z and Tao D C. 2017b. 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. 2018b. 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]
- Ma K D, Liu X L, Fang Y M and Simoncelli E P. 2019. Blind image quality assessment by learning from multiple annotators//*Proceedings of 2019 IEEE International Conference on Image Processing*.

- Taipei, China; IEEE: 2344-2348 [DOI: 10.1109/ICIP.2019.8803390]
- Ma K D, Yeganeh H, Zeng K and Wang Z. 2015. High dynamic range image compression by optimizing tone mapped image quality index. *IEEE Transactions on Image Processing*, 24 (10): 3086-3097 [DOI: 10.1109/TIP.2015.2436340]
- Min X K, Gu K, Zhai G T, Liu J, Yang X K and Chen C W. 2018. Blind quality assessment based on pseudo-reference image. *IEEE Transactions on Multimedia*, 20(8): 2049-2062 [DOI: 10.1109/TMM.2017.2788206]
- 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, 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, Su C C, Mittal A and Bovik A C. 2013. Subjective evaluation of stereoscopic image quality. *Signal Processing: Image Communication*, 28(8): 870-883 [DOI: 10.1016/j.image.2012.08.004]
- Muandet K, Balduzzi D and Schölkopf B. 2013. Domain generalization via invariant feature representation//*Proceedings of the 30th International Conference on Machine Learning*. Atlanta, USA; ACM: 10-18
- Ni Z K, Ma L, Zeng H Q, Chen J, Cai C H and Ma K K. 2017. ESIM: edge similarity for screen content image quality assessment. *IEEE Transactions on Image Processing*, 26(10): 4818-4831 [DOI: 10.1109/TIP.2017.2718185]
- 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]
- Oh H, Ahn S, Kim J and Lee S. 2017. Blind deep S3D image quality evaluation via local to global feature aggregation. *IEEE Transactions on Image Processing*, 26(10): 4923-4936 [DOI: 10.1109/TIP.2017.2725584]
- Pan D, Shi P, Hou M, Ying Z F, Fu S Z and Zhang Y. 2018. Blind predicting similar quality map for image quality assessment//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA; IEEE: 6373-6382 [DOI: 10.1109/CVPR.2018.00667]
- Paudyal P, Battisti F, Sjöström M, Olsson R and Carli M. 2017. Towards the perceptual quality evaluation of compressed light field images. *IEEE Transactions on Broadcasting*, 63 (3): 507-522 [DOI: 10.1109/TBC.2017.2704430]
- Ponomarenko N, Jin L N, Ieremeiev O, Lukin V, Egiazarian K, Astola J, Vozel B, Chehdi K, Carli M, Battisti F and Jay Kuo C C. 2015. Image database TID2013: peculiarities, results and perspectives. *Signal Processing: Image Communication*, 30: 57-77 [DOI: 10.1016/j.image.2014.10.009]
- Ponomarenko N, Lukin V, Zelensky A, Egiazarian K, Carli M and Battisti F. 2009. TID2008-a database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics*, 10(4): 30-45
- Rehman A and Wang Z. 2012. Reduced-reference image quality assessment by structural similarity estimation. *IEEE Transactions on Image Processing*, 21 (8): 3378-3389 [DOI: 10.1109/TIP.2012.2197011]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany; Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Sheikh H R, Wang Z, Cormack L and Bovik A C. 2006. LIVE image quality assessment database release 2 [DB/OL]. [2020-07-08]. <https://live.ece.utexas.edu/research/quality/>
- Shi L K, Zhao S Y and Chen Z B. 2019a. Belif: blind quality evaluator of light field image with tensor structure variation index//*Proceedings of 2019 IEEE International Conference on Image Processing*. Taipei, China; IEEE: 3781-3785 [DOI: 10.1109/ICIP.2019.8803559]
- Shi L K, Zhao S Y, Zhou W and Chen Z B. 2018. Perceptual evaluation of light field image//*Proceedings of the 25th IEEE International Conference on Image Processing*. Athens, Greece; IEEE: 41-45 [DOI: 10.1109/ICIP.2018.8451077]
- Shi L K, Zhou W, Chen Z B and Zhang J L. 2019b. No-reference light field image quality assessment based on spatial-angular measurement. *IEEE Transactions on Circuits and Systems for Video Technology*, 30 (11): 4114-4128 [DOI: 10.1109/TCSVT.2019.2955011]
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition//*Proceedings of the 3rd International Conference on Learning Representations*. San Diego, USA: [s. n.]: 1-14
- Sui X J, Ma K D, Yao Y R and Fang Y M. 2020. Omnidirectional images as moving camera videos [EB/OL]. [2020-07-08]. <https://arxiv.org/pdf/2005.10547.pdf>
- Sun W, Gu K, Ma S W, Zhu W H, Liu N and Zhai G T. 2018. A large-scale compressed 360-degree spherical image database: from subjective quality evaluation to objective model comparison//*Proceedings of the 20th IEEE International Workshop on Multimedia Signal Processing*. Vancouver, Canada; IEEE: 1-6 [DOI: 10.1109/MMSP.2018.8547102]
- 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]

- Talebi H and Milanfar P. 2018. NIMA: neural image assessment. *IEEE Transactions on Image Processing*, 27(8): 3998-4011 [DOI: 10.1109/TIP.2018.2831899]
- 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
- Viola I and Ebrahimi T. 2018. VALID: visual quality assessment for light field images dataset//Proceedings of the 10th International Conference on Quality of Multimedia Experience. Cagliari, Italy: IEEE: 1-3 [DOI: 10.1109/QoMEX.2018.8463388]
- Wang G C, Li L D, Li Q H, Gu K, Lu Z L and Qian J S. 2017. Perceptual evaluation of single-image super-resolution reconstruction//Proceedings of 2017 IEEE International Conference on Image Processing. Beijing, China: IEEE: 3145-3149 [DOI: 10.1109/ICIP.2017.8296862]
- Wang J H, Rehman A, Zeng K, Wang S Q and Wang Z. 2015. Quality prediction of asymmetrically distorted stereoscopic 3D images. *IEEE Transactions on Image Processing*, 24(11): 3400-3414 [DOI: 10.1109/TIP.2015.2446942]
- Wang J P, Li C and Chen W H. 2017. Blind image quality assessment based on DCT features in the non-flat region and EGRNN. *Chinese Journal of Computers*, 40(11): 2492-2505 (王俊平, 李超, 陈伟华. 2017. 基于图像非平坦区域 DCT 特性和 EGRNN 的盲图像质量评价. *计算机学报*, 40(11): 2492-2505) [DOI: 10.11897/SP.J.1016.2017.02492]
- 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, Lin W S, Shi G M and Liu A M. 2013. Perceptual quality metric with internal generative mechanism. *IEEE Transactions on Image Processing*, 22(1): 43-54 [DOI: 10.1109/TIP.2012.2214048]
- Wu J J, Zhang M, Li L D, Dong W S, Shi G M and Lin W S. 2019. No-reference image quality assessment with visual pattern degradation. *Information Sciences*, 504: 487-500 [DOI: 10.1016/j.ins.2019.07.061]
- Wu Q B, Wang Z and Li H L. 2015. A highly efficient method for blind image quality assessment//Proceedings of 2015 IEEE International Conference on Image Processing. Quebec City, Canada: IEEE: 339-343 [DOI: 10.1109/ICIP.2015.7350816]
- Xu J, Zhang L and Zhang D. 2018. External prior guided internal prior learning for real-world noisy image denoising. *IEEE Transactions on Image Processing*, 27(6): 2996-3010 [DOI: 10.1109/TIP.2018.2811546]
- Xu J H, Zhou W and Chen Z B. 2020b. Blind omnidirectional image quality assessment with viewport oriented graph convolutional networks [EB/OL]. [2020-07-08]. <https://arxiv.org/pdf/2002.09140v1.pdf>
- Xu J T, Ye P, Li Q H, Du H Q, Liu Y and Doermann D. 2016. Blind image quality assessment based on high order statistics aggregation. *IEEE Transactions on Image Processing*, 25(9): 4444-4457 [DOI: 10.1109/TIP.2016.2585880]
- Xu M, Li C, Zhang S Y and le Callet P. 2020a. State-of-the-art in 360° video/image processing: perception, assessment and compression. *IEEE Journal of Selected Topics in Signal Processing*, 14(1): 5-26 [DOI: 10.1109/JSTSP.2020.2966864]
- Xue W F, Mou X Q, Zhang L, Bovik A C and Feng X C. 2014. Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features. *IEEE Transactions on Image Processing*, 23(11): 4850-4862 [DOI: 10.1109/TIP.2014.2355716]
- Xue W F, Zhang L and Mou Z Q. 2013. Learning without human scores for blind image quality assessment//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE: 995-1002 [DOI: 10.1109/CVPR.2013.133]
- Yan B, Bare B, Ma C X, Li K and Tan W M. 2019a. Deep objective quality assessment driven single image super-resolution. *IEEE Transactions on Multimedia*, 21(11): 2957-2971 [DOI: 10.1109/TMM.2019.2914883]
- Yan B, Bare B and Tan W M. 2019b. Naturalness-aware deep no-reference image quality assessment. *IEEE Transactions on Multimedia*, 21(10): 2603-2615 [DOI: 10.1109/TMM.2019.2904879]
- Yan J B. 2018. No-Reference Stereoscopic Image Quality Assessment Based on Binocular Visual Perception. Nanchang: Jiangxi University of Finance and Economics (鄢杰斌. 2018. 基于双目视觉感知的无参考三维图像视觉质量评价. 南昌: 江西财经大学)
- Yan J B, Fang Y M, Huang L P, Min X K, Yao Y R and Zhai G T. 2020. Blind stereoscopic image quality assessment by deep neural network of multi-level feature fusion//Proceedings of 2020 IEEE International Conference on Multimedia and Expo. London, UK: IEEE: 1-6 [DOI: 10.1109/ICME46284.2020.9102888]
- Yan Q S, Gong D and Zhang Y N. 2019c. Two-stream convolutional networks for blind image quality assessment. *IEEE Transactions on Image Processing*, 28(5): 2200-2211 [DOI: 10.1109/TIP.2018.2883741]
- Yang H, Fang Y M and Lin W S. 2015. Perceptual quality assessment of screen content images. *IEEE Transactions on Image Processing*, 24(11): 4408-4421 [DOI: 10.1109/TIP.2015.2465145]
- Yang S, Jiang Q P, Lin W S and Wang Y T. 2019b. SGDNet: an end-to-end saliency-guided deep neural network for no-reference image

- quality assessment//Proceedings of the 27th ACM International Conference on Multimedia. Nice, France: ACM: 1383-1391 [DOI: 10.1145/3343031.3350990]
- Yang W M, Zhang X C, Tian Y P, Wang W, Xue J H and Liao Q M. 2019a. Deep learning for single image super-resolution: a brief review. *IEEE Transactions on Multimedia*, 21(12): 3106-3121 [DOI: 10.1109/TMM.2019.2919431]
- Yang X H, Li F and Liu H T. 2019c. A survey of DNN methods for blind image quality assessment. *IEEE Access*, 7: 123788-123806 [DOI: 10.1109/ACCESS.2019.2938900]
- Ye P, Kumar J and Doermann D. 2014. Beyond human opinion scores: blind image quality assessment based on synthetic scores//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE: 4241-4248 [DOI: 10.1109/CVPR.2014.540]
- Ye P, Kumar J, Kang L and Doermann D. 2012. Unsupervised feature learning framework for no-reference image quality assessment//Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE: 1098-1105 [DOI: 10.1109/CVPR.2012.6247789]
- Yeganeh H and Wang Z. 2013. Objective quality assessment of tone-mapped images. *IEEE Transactions on Image Processing*, 22(2): 657-667 [DOI: 10.1109/TIP.2012.2221725]
- Ying Z Q, Niu H R, Gupta P, Mahajan D, Ghadiyaram D and Bovik A C. 2020. From patches to pictures (PaQ-2-PiQ): mapping the perceptual space of picture quality//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 3575-3585 [DOI: 10.1109/CVPR42600.2020.00363]
- Yue G H, Hou C P, Yan W Q, Choi L K, Zhou T W and Hou Y H. 2018. Blind quality assessment for screen content images via convolutional neural network. *Digital Signal Processing*, 91: 21-30 [DOI: 10.1016/j.dsp.2018.12.007]
- 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]
- 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 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 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]
- Zhang X F, Xiong R Q, Fan X P, Ma S W and Gao W. 2013. Compression artifact reduction by overlapped-block transform coefficient estimation with block similarity. *IEEE Transactions on Image Processing*, 22(12): 4613-4626 [DOI: 10.1109/TIP.2013.2274386]
- Zhang Y and Chandler D M. 2018. Opinion-unaware blind quality assessment of multiply and singly distorted images via distortion parameter estimation. *IEEE Transactions on Image Processing*, 27(11): 5433-5448 [DOI: 10.1109/TIP.2018.2857413]
- Zhao H, Gallo O, Frosio I and Kautz J. 2017. Loss functions for image restoration with neural networks. *IEEE Transactions on Computational Imaging*, 3(1): 47-57 [DOI: 10.1109/TCI.2016.2644865]
- Zhou W, Chen Z B and Li W P. 2019. Dual-stream interactive networks for no-reference stereoscopic image quality assessment. *IEEE Transactions on Image Processing*, 28(8): 3946-3958 [DOI: 10.1109/TIP.2019.2902831]
- Zhou W, Jiang Q P, Wang Y W, Chen Z B and Li W P. 2020a. Blind quality assessment for image superresolution using deep two-stream convolutional networks. *Information Sciences*, 528: 205-218 [DOI: 10.1016/j.ins.2020.04.030]
- Zhou W, Shi L K, Chen Z B and Zhang J L. 2020b. Tensor oriented noreference light field image quality assessment. *IEEE Transactions on Image Processing*, 29: 4070-4084 [DOI: 10.1109/TIP.2020.2969777]
- 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. Seattle, USA: IEEE: 14143-14152 [DOI: 10.1109/CVPR42600.2020.01415]

作者简介



方玉明, 1984年生, 男, 教授, 主要研究方向为计算机视觉、多媒体信号处理、视觉质量评估。

E-mail: leo.fangyuming@foxmail.com

眭相杰, 男, 硕士研究生, 主要研究方向为计算机视觉。

E-mail: suixiangjie2017@163.com

鄢杰斌, 男, 博士研究生, 主要研究方向为计算机视觉。

E-mail: jiebinan@foxmail.com

刘学林, 男, 硕士研究生, 主要研究方向为图像处理。

E-mail: 409544864@qq.com

黄丽萍, 女, 硕士研究生, 主要研究方向为图像处理。

E-mail: 2227994134@qq.com