



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

2020

09

VOL.25

ISSN1006-8961
CN11-3758/TB



遥感边缘智能技术 P1719



第25卷第9期 (总第293期)

2020年9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 顾行发

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 Gu Xingfa

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

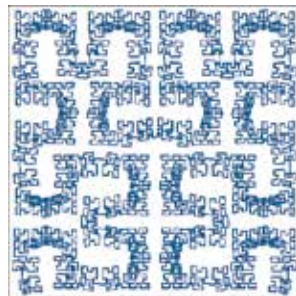
国外发行代号 M1406

国内邮发代号 82-831

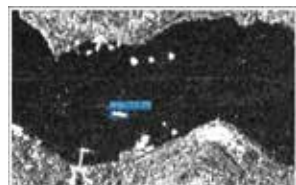
国内定价 60.00元



第一人称视角下的社会力优化多行人跟踪(第1869页)



分形曲线生成的频域方法(第1904页)



双向特征融合的数据自适应SAR图像舰船目标检测模型(第1943页)

学者观点

遥感边缘智能技术研究进展及挑战

孙显, 梁伟, 刁文辉, 曹志颖, 冯瑛超, 王冰, 付琨 1719

综述

红外弱小目标检测算法综述

李俊宏, 张萍, 王晓玮, 黄世泽 1739

目标检测中的尺度变换应用综述

申奉臻, 张萍, 罗金, 刘松阳, 冯世杰 1754

图像处理和编码

分层特征融合注意力网络图像超分辨率重建

雷鹏程, 刘丛, 唐坚刚, 彭敦陆 1773

适用小样本的无参考水下视频质量评价方法

宋巍, 刘诗梦, 黄冬梅, 王文娟, 王建 1787

图像分析和识别

结合面料属性和触觉感测的织物识别

邢寅初, 刘骊, 付晓东, 刘利军, 黄青松 1800

结合GAN的轻量级模糊车牌识别算法

段宾, 符祥, 江毅, 曾接贤 1813

非贪婪的鲁棒性度量学习算法

曾凡霞, 张文生 1825

不规则像素簇显著性检测算法

李明旭, 翟东海 1837

全局区域相异度阈值驱动构建的稀疏尺度集模型

王浩宇, 沈占锋, 柯映明, 许泽宇, 李硕, 焦淑慧 1848

图像理解和计算机视觉

分裂合并运动分割的多运动视觉里程计方法

王晨捷, 张云, 赵青, 王伟, 尹露, 罗斌, 张良培 1859

第一人称视角下的社会力优化多行人跟踪

杨廷召, 刘骊, 付晓东, 刘利军, 黄青松 1869

引入辅助损失的多场景车道线检测

陈立潮, 徐秀芝, 曹建芳, 潘理虎 1882

计算机图形学

旋转骨架法在二值图像分形维数计算中的应用

纪佑军, 何杰, 韩海水, 程忠钊, 曾保全 1894

分形曲线生成的频域方法

陈伟, 乔洁雯, 周晨 1904

医学图像处理

融合残差注意力机制的UNet视盘分割

侯向丹, 赵一浩, 刘洪普, 郭鸿湧, 于习欣, 丁梦园 1915

深层聚合残差密集网络的超声图像左心室分割

吴宣言, 缙新科, 朱子重, 魏域林, 王凯 1930

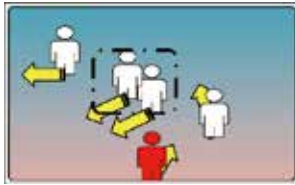
遥感图像处理

双向特征融合的数据自适应SAR图像舰船目标检测模型

张筱晗, 姚力波, 吕亚飞, 简涛, 赵志伟, 藏洁 1943

CONTENTS

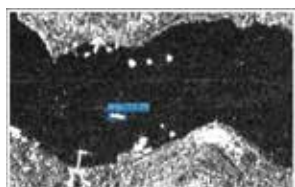
JOURNAL OF IMAGE AND GRAPHICS



Multi-pedestrian tracking optimized by social force model under first-person perspective(P1869)



Fractal curve generation method based on the frequency domain (P1904)



Data-adaptive single-shot ship detector with a bidirectional feature fusion module for SAR images (P1943)

Scholar View

Progress and challenges of remote sensing edge intelligence technology

Sun Xian, Liang Wei, Diao Wenhui, Cao Zhiying, Feng Yingchao, Wang Bing, Fu Kun 1719

Review

Infrared small-target detection algorithms: a survey

Li Junhong, Zhang Ping, Wang Xiaowei, Huang Shize 1739

Scale changing in general object detection: a survey

Shen Fengcan, Zhang Ping, Luo Jin, Liu Songyang, Feng Shijie 1754

Image Processing and Coding

Hierarchical feature fusion attention network for image super-resolution reconstruction

Lei Pengcheng, Liu Cong, Tang Jiangang, Peng Dunlu 1773

Non-reference underwater video quality assessment method for small size samples

Song Wei, Liu Shimeng, Huang Dongmei, Wang Wenjuan, Wang Jian 1787

Image Analysis and Recognition

Cloth recognition based on fabric properties and tactile sensing

Xing Yinchu, Liu Li, Fu Xiaodong, Liu Lijun, Huang Qingsong 1800

Lightweight blurred car plate recognition method combined with generated images

Duan Bin, Fu Xiang, Jiang Yi, Zeng Jiexian 1813

Robust distance metric learning algorithm based on nongreedy solution

Zeng Fanxia, Zhang Wensheng 1825

Significance detection method with irregular pixel clusters

Li Mingxu, Zhai Donghai 1837

Sparse scale set model based on global regional dissimilarity threshold

Wang Haoyu, Shen Zhanfeng, Ke Yingming, Xu Zeyu, Li Shuo, Jiao Shuhui 1848

Image Understanding and Computer Vision

Multi-motion visual odometry based on split-merged motion segmentation

Wang Chenjie, Zhang Yun, Zhao Qing, Wang Wei, Yin Lu, Luo Bin, Zhang Liangpei 1859

Multi-pedestrian tracking optimized by social force model under first-person perspective

Yang Tingzhao, Liu Li, Fu Xiaodong, Liu Lijun, Huang Qingsong 1869

Multi-scenario lane line detection with auxiliary loss

Chen Lichao, Xu Xiuzhi, Cao Jianfang, Pan Lihu 1882

Computer Graphics

Application of skeleton rotating method in fractal dimension calculate of binary image

Ji Youjun, He Jie, Han Haishui, Cheng Zhongzhao, Zeng Baoquan 1894

Fractal curve generation method based on the frequency domain

Chen Wei, Qiao Jiewen, Zhou Chen 1904

Medical Image Processing

Optic disk segmentation by combining UNet and residual attention mechanism

Hou Xiangdan, Zhao Yihao, Liu Hongpu, Guo Hongyong, Yu Xixin, Ding Mengyuan 1915

Left ventricular segmentation on ultrasound images using deep layer aggregation for residual dense networks

Wu Xuanyan, Gou Xinke, Zhu Zizhong, Wei Yulin, Wang Kai 1930

Remote Sensing Image Processing

Data-adaptive single-shot ship detector with a bidirectional feature fusion module for SAR images

Zhang Xiaohan, Yao Libo, Lyu Yafei, Jian Tao, Zhao Zhiwei, Zang Jie 1943

中图分类号: TN911.73 文献标识码: A 文章编号: 1006-8961(2020)09-1787-13

论文引用格式: 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(09): 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(09): 1787-1799) [DOI: 10.11834/jig.200025]

适用小样本的无参考水下视频质量评价方法

宋巍¹, 刘诗梦¹, 黄冬梅^{1,2}, 王文娟¹, 王建¹

1. 上海海洋大学信息学院, 上海 201306; 2. 上海电力大学, 上海 201306

摘要: **目的** 视频质量评价是视频技术研究的关键之一。水下环境比其他自然环境更加复杂, 自然光在深水中被完全吸收, 拍摄所用的人工光源在水中传播时会发生光吸收、色散和散射等情况, 同时受水体浑浊度和拍摄设备等因素影响, 导致水下视频具有高度的空间弱可视性和时间不稳定性, 常规视频质量评价方法无法对水下视频进行准确、有效的评价。本文考虑水下视频特性, 提出一种适用小样本的结合空域统计特性与编码的水下视频质量评价方法。**方法** 基于水下视频成像特性, 建立新的水下视频数据库, 设计主观质量评价方法对所有视频进行1~5分质量标注。从水下视频中提取视频帧图像, 针对空间域计算图像失真统计特性, 然后结合视频编码参数, 通过训练线性模型权重系数完成水下视频的质量评价。**结果** 实验表明, 与几种主流的质量评价方法相比, 本文水下视频质量评价方法与人类视觉感知的相关性最高, 模型评价结果与主观质量评价结果的皮尔森线性相关系数 PCC (Pearson's correlation coefficient) 为 0.840 8, 斯皮尔曼等级秩序相关系数 SROCC (Spearman's rank order correlation coefficient) 为 0.832 2。通过比较各方法评价结果与真实值的均方误差 (mean square error, MSE), 本文方法 MSE 值最小, 为 0.113 1, 说明本文的质量评价结果更加稳定。**结论** 本文通过空间域单帧图像自然场景统计特性和视频编码参数融合的方式, 提出的无参考水下视频质量评价方法, 能够很好地运用小样本水下视频数据集建立与人类视觉感知高度相关的评价模型, 为水下视频做出更准确的质量评价。

关键词: 视频质量评价; 客观质量评价模型; 水下视频; 自然场景统计; 编码参数

Non-reference underwater video quality assessment method for small size samples

Song Wei¹, Liu Shimeng¹, Huang Dongmei^{1,2}, Wang Wenjuan¹, Wang Jian¹

1. College of Information, Shanghai Ocean University, Shanghai 201306, China;

2. Shanghai University of Electric and Power, Shanghai 201306, China

Abstract: Objective The application of underwater video technology has a history of more than 60 years. This technology plays an important role in promoting research on marine bioecology, fish species, and underwater object detection and tracking. Video quality assessment is one of the key areas being studied in video technology research. Such assessment is especially vital for underwater videos because underwater environments are more complex than atmospheric ones. On the one hand, natural sunlight is seriously absorbed in deep water, and the artificial light used in video shooting suffers from light absorption, dispersion, and scattering due to water turbidity and submarine topography. As a result, underwater videos have blurred picture, low contrast, color cast, and uneven lighting. On the other hand, underwater video quality is affected

收稿日期: 2020-01-17; 修回日期: 2020-03-09; 预印本日期: 2020-03-16

基金项目: 国家自然科学基金项目 (61702323)

Supported by: National Natural Science Foundation of China (61702323)

by the limitation of photography equipment and the influence of water flow. When shooting a moving object, the lens hardly stabilizes and turns unsmooth. Compared with videos shot in natural scenes, underwater videos are characterized by large lens movement, shaking, and serious out of focus. These characteristics make it difficult for conventional video quality assessment (VQA) methods to evaluate underwater video accurately and effectively. Thus, the “quality” of underwater videos must be redefined, and an effective quality assessment method must be established. In this study, we establish an underwater video dataset by considering underwater video imaging characteristics, annotate its video quality via subjective quality assessment, and propose an objective underwater video quality assessment model on the basis of spatial naturalness and video compression index.

Method First, a new underwater video dataset is established to 1) collect several underwater videos captured in real deep sea environments for processing as source data; 2) filter these videos preliminarily to include different underwater scenes; 3) cut the preliminary screened videos at intervals of 10 seconds; 4) refilter the short video sequences to cover different shoot characteristics and color diversity, thus generating 25 video sequences with rich color information, different video contents, and different underwater video features; and 5) expand the dataset using different frame rates and bit rates as compression parameters. A total of 250 ($25 + 25 \times 3 \times 3$) video sequences are obtained. Then, subjective quality assessment is conducted. Absolute category rating is used by 20 participants to annotate all the 250 videos with scores ranging from 1 to 5. Then, we consider influences on the underwater video quality from the aspects of spatial, temporal, and compression features. The spatial features are expressed by natural scene statistics distortion characteristics in the spatial domain and are calculated using the blind/referenceless image spatial quality evaluator (BRISQUE) algorithm. The temporal features are expressed by optical flow motion features. We first compute the dense optical flow matrix between adjacent frames and then extract the mean and variation of overall optical flows and the mean and variation of the main objects in the video. Compression features use resolution, frame rate, and bit rate, which are easy-to-access video coding parameters. Considering the redundancy and relevancy of these potential features, we analyze the correlations among the features and between the features and the subjective quality scores. Then, we select 21 features as influence factors, which only contain 18 spatial natural characteristics and three compression indexes. Lastly, we establish a linear model with the selected features to evaluate underwater video quality objectively through linear regression with cross validation.

Result Experimental results show that the proposed underwater video quality assessment model based on spatial naturalness and compression index can obtain the highest correlation with subjective scores in comparison with several mainstream quality assessment models, including two underwater image quality indices (underwater image quality measure (UIQM) and underwater color image quality evaluation (UCIQE)), a natural image quality distortion index (BRISQUE), and a video quality assessment model (video intrinsic integrity and distortion evaluation oracle (VIIDEO)). Performance evaluation is based on Pearson’s correlation coefficient (PCC), Spearman’s rank order correlation coefficient (SROCC) and the mean squared errors (MSE) between the predicted video quality scores of each model and the subjective scores. On the test video dataset, our method achieves the highest correlation ($PCC = 0.8408$, $SROCC = 0.8322$) and a minimum MSE value of 0.1131. This result indicates that our proposed method is more stable and can predict video quality more accurately than other methods. By contrast, the video quality assessment model VIIDEO can hardly provide correct results, whereas UIQM and UCIQE demonstrate poor performance with a PCC and SROCC of 0.3 ~ 0.4. In addition, BRISQUE performs relatively better than the other methods although still poorer than our method.

Conclusion Underwater videos are characterized by blurred picture, low contrast, color distortion, uneven lighting, large lens movement, and out of focus. To achieve an accurate assessment of underwater video quality, we fully consider the characteristics and shooting conditions of underwater videos and establish a labeled underwater video dataset with subjective video quality assessment. By fitting a linear regression model for subjective quality scores with natural statistical characteristics of video frames and video compression parameters, we propose an objective underwater video quality assessment model. The proposed nonreference underwater video quality assessment method is suitable to establish a prediction model that is highly related to human visual perception, with a small sample size of underwater video dataset.

Key words: video quality assessment; objective quality assessment model; underwater video; natural scene statistics; compression parameters

0 引言

随着深海探测技术的不断发展,水下视频在海底探索和研究中的作用日益凸显。而视频质量评价在视频技术研究中处于重要地位。对水下视频准确、有效的质量评价不仅是直接衡量水下视频质量的重要指标,也是获取高质量水下视频的重要前提,还是评价水下图像/视频修复或增强技术的重要手段,有助于水下视频质量的持续提高和有用信息的提取与识别。

水下视频质量一方面受复杂的水下成像条件影响,自然光在深海中完全吸收,拍摄时的辅助人工光源在海水中传播时发生光吸收、散射和色散,同时受海水浑浊度和海底地形等影响,在空间维度上呈现画面模糊、对比度低、色彩失真和光照不均等特点(图1(a)(b));另一方面受拍摄设备限制和水流等影响,移动拍摄时镜头不易稳定,固定镜头拍摄移动物体时,镜头转动生涩,相比于自然场景下拍摄的视频,水下视频具有镜头移动晃动大(图1(c)),以及失焦严重(图1(d))等问题。因此,大气环境下建立的自然场景视频质量评价指标或方法可能无法对水下视频质量做出准确合理评价,需要重新定义水下视频的“质量”,并建立有效的质量评价方法和模型。

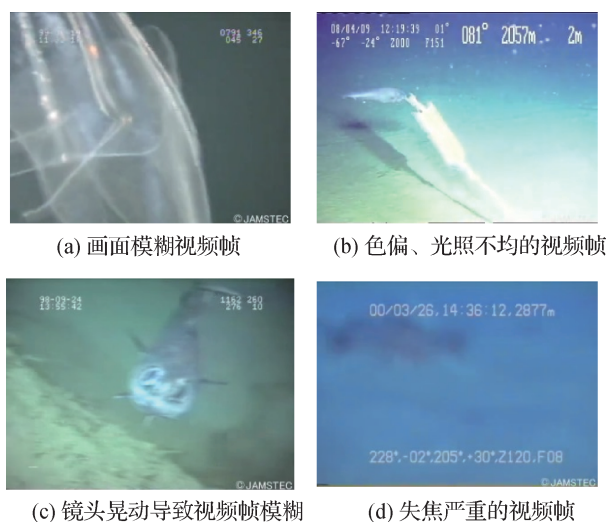


图1 水下视频帧示例

Fig.1 Examples of underwater video frames((a) hazy frame; (b) color cast and uneven lighting frame; (c) blur frame due to wobbly camera shot; (d) out-of-focus video frame)

国内外在水下视频质量评价(video quality assessment, VQA)方面的研究成果较少,研究主要集中于水下图像质量评价方法。少数几个水下视频质量评价模型,如Moreno-Roldán等人(2017, 2018)提出的模型,存在样本不真实和模型过拟合等问题。水下视频质量评价模型在理论和实用性方面都面临挑战。相对而言,自然场景下的视频质量评价的研究已经比较成熟,有主观质量评价和客观质量评价两种,前者是通过人眼对视频的主观感受来评估视频质量,后者通过模型模拟人类视觉感知量化指标衡量视频质量。典型的视频主观VQA方法主要由国际视频质量专家组VEQG(video quality expert group)提出,并被国际电信联盟(International Telecommunication Union, ITU)标准化(ITU-T, 1999; ITU-R, 2002, 2007)。通过主观评价得到的结果,通常认为最符合人类感知的质量,但是主观评价方法普遍存在耗时长、用户评价负担重和容易产生个体偏见等问题,更重要的是,主观评价不能自动化,因此,建立可靠的客观VQA是目前水下视频质量评价的一个有效的途径。客观VQA面临的挑战是要准确模拟人类视觉感知,从而替代复杂耗时的主观质量评价。ITU根据是否需要参考原始视频,把客观VQA分为全参考(full-reference, FR)、部分参考(reduced-reference, RR)和无参考(no-reference, NR)方法。考虑到水下视频的特性,通常无法获得理想的参考视频,FR和RR方法在水下视频质量评价中的实用性有限,因此,对水下视频质量评价一般采用无参考/盲预测的方法。

本文构建了水下视频数据集,然后从时空域和视频编码参数分别讨论了相关特征对水下视频质量的影响,并利用最终选定的特征建立了基于空域自然特性与编码参数的无参考水下视频质量评价模型(no-reference underwater video quality assessment, NR-UVQA),该模型能够以较小的计算复杂度,获得与人类视觉感知较高的相关性。本文贡献概述如下:1)充分考虑水下视频成像特性,建立水下视频质量评价数据集;2)通过分析视频空间全帧统计特性、光流运动特性和编码参数对视频质量的影响,获取与感知质量相关的少量特征,建立NR-UVQA模型;3)与多种优秀的质量评价模型对比,分析认为本文方法在评价结果和观察者评级之间能够获得更高的相关性。

1 相关工作

视频由在时间上连续播放的单帧图像组成,因此单帧图像的质量对于整体视频的质量具有绝对的意义。许多学者提出的视频质量评价方法都是在图像质量评价方法的基础上做出的创新。其中,自然图像统计(natural scene statistics, NSS)(Field, 1999)模式是一种常用的图像特征提取方法,例如,Moorthy 和 Bovik(2011)基于 NSS 以图像真实性完整性为指标提出无参考图像质量评价方法(distortion identification-based image verity and integrity evaluation, DIIVINE),在 LIVE IQA 数据库进行全面评估,取得与较好的 FR 质量评价等同的效果。Mittal 等人(2012)提出一种基于 NSS 的失真通用盲/无参考质量评估算法(blind/referenceless image spatial quality evaluator, BRISQUE),该算法不计算特定图像失真特性,而是使用局部归一化亮度系数的场景统计来量化可能的损失,能够以非常低的计算复杂度准确表示图像质量。Mittal 等人(2013)基于简单的空间域 NSS 模型提出自然图像质量评估器(natural image quality evaluator, NIQE),在不知道图像失真特性的情况下,能够得到较准确的质量评价结果。

针对水下图像特性,也有许多专门用于水下图像质量评价的成果。例如, Yang 和 Sowmya(2015)提出水下彩色图像质量评价指标(underwater color image quality evaluation, UCIQE),该指标提取 CIELab 空间统计特征中与观察者感知相关度最高的 3 个质量度量:色度、饱和度和对比度,将这些参数线性组合用来预测图像质量,最终能够在预测结果和观察者评级之间获得很好的相关性。Panetta 等人(2016)受到人类视觉系统属性的启发,提出了一种无参考的水下图像质量评价方法(underwater image quality measure, UIQM),针对水下图像的模糊、低对比度和色偏的特点,采用 3 种水下图像属性测量:水下图像色彩测量(underwater image colorfulness measure, UICM)、水下图像清晰度测量(underwater image sharpness measure, UISM)和水下图像对比度测量(underwater image contrast measure, UIConM)来表征水下图像质量。Wang 等人(2018)基于水下吸收和散射特征的成像分析,提出新的水下彩色图像质量评估方法 CCF(colorfulness, con-

trast, fog density),该方法将色度、对比度和雾密度作为评价指标。郭继昌等人(2017)将深度学习网络框架与随机森林回归模型相结合,无需参考图像就能得到与观察者感知质量相关性很高的预测结果,但模型的参数调节需要原始数据,且深度学习网络对小样本数据容易过拟合。Wang 等人(2019)对比信息熵(ENTROPY), BRISQUE0, NIQE, UIQM 和 UCIQE 等 5 种方法评估水下图像增强效果,结果显示 BRISQUE, UIQM 和 UCIQE 各有优劣,而 ENTROPY, NIQE 不能准确反映水下图像质量。

无论是水下图像还是自然场景图像,对图像进行质量评价都能取得较好的评价结果,但将图像质量评价方法直接运用于水下视频的质量评价,存在明显的局限性。图像质量评价方法能够用于评估每一帧图像的质量,但视频具有运动特性,仅评估视频帧的质量不能很好地体现视频运动特性对视频质量的影响。Saad 等人(2014)提出一种 Video BLINEDS 的盲参考视频质量评价模型,以离散余弦变换系数的统计特性表示视频的时空属性,结合运动一致性对视频质量分数进行预测。其使用 LIVE VQA 数据库(Seshadrinathan 等, 2010)进行训练,然后在该数据库和 EPFL-Polimi 数据库(De Simone 等, 2010)上进行评估性能,取得与对比 FR 方法同等效果。Mittal 等人(2016)提出 VIIDEO(video intrinsic integrity and distortion evaluation oracle),通过评估视频内在完整性和失真程度,克服对主观得分的依赖性和其他失真,在视频质量评价中取得比 Video BLINEDS 更加接近人类视觉感知的结果。尽管这些方法都能取得优秀的结果,但是在水下视频质量评价的实际应用中,应避免复杂的处理任务、涉及大量图像或特征处理的方法,容易导致过拟合。因此,需要一种低计算复杂度的方法去训练水下视频质量评价客观模型,从而得到更加准确的水下视频质量评价结果。

Moreno-Roldán 等人(2017)提出两种用于水下视频质量评估的机器学习模型:基于曲面拟合回归的 NR 参数模型和基于序数逻辑回归的 RR 混合模型,两种模型均以比特率和帧率作为参数进行训练,主要考虑了水下传输对图像质量的影响。2018 年,他们进一步提出一种无参考水下视频质量评价方法(Moreno-Roldán 等, 2018),该方法使用视频单帧和相邻帧差的 NSS 特征建立线性模型来表征水下视

频质量,在他们建立的主观评价数据集上取得一个不错的结果,但该方法所采用的特征值数量多达41个,而且全部是基于NSS的统计特征,对于小样本的数据集易过拟合。因此,针对水下视频特点,设计专门应用于水下视频质量评价的方法仍是需要解决的问题。

本文从水下视频的特点出发,建立水下视频数据集,并设计一种适用于小样本的无参考水下视频质量评价模型。该方法不针对特定失真类型,最终取得与人类视觉感知相关性最高的评价结果。

2 主观质量评价数据集

本文提出的无参考视频质量评价算法目标是取得接近主观水下视频质量评价的结果。由于不存在公开的水下视频质量评价数据集,在构建本文的NR-UVQA算法进行水下视频质量评价实验之前,需要建立一个水下视频数据集。数据集构建过程如图2所示。

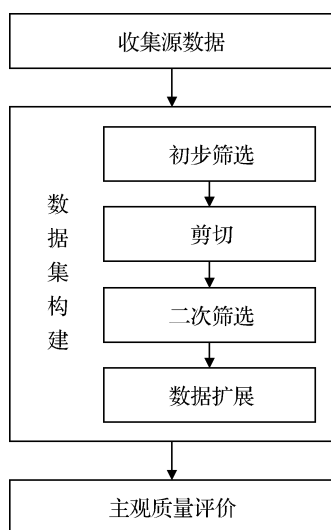


图2 数据集构建流程图

Fig. 2 Flowchart of dataset construction

2.1 源数据

用来测试的水下视频主要来源于日本海洋地球科学技术局的深海图像 E-library (J-EDI)。考虑到样本的多样性,共收集60个不同时间、不同场景、不同分辨率以及不同拍摄环境的水下视频,并根据视频拍摄内容、色彩信息和镜头特点等信息对源数据建立了人为描述。在图3中举例说明了4个水下视频的描述。图3(a)视频描述为:拍摄时间:

2004. 10. 08;拍摄地点:Lau Basin VaiLiliSite;水深:1 717.2~1 716.7 m;特点:色彩丰富、静止主体、镜头少许缩放和转动。图3(b)视频描述为:拍摄时间:2007. 10. 03;拍摄地点:Kushiro Canyon;水深:3 896.5~3 896.3 m;特点:高亮度、低色彩饱和度,海底静止主体,镜头拉近、转动。图3(c)视频描述为:拍摄时间:2016. 02. 21;水深:3 419.0~3 387.0 m;特点:绿色偏暗,缓慢匀速移动拍摄海底岩石。图3(d)视频描述为:拍摄时间:2015. 05. 05;拍摄地点:Sanriku;水深:799.7~799.7 m;特点:绿色偏色、不清晰,主体在海底轻微摆动,镜头拉近,有轻微晃动。

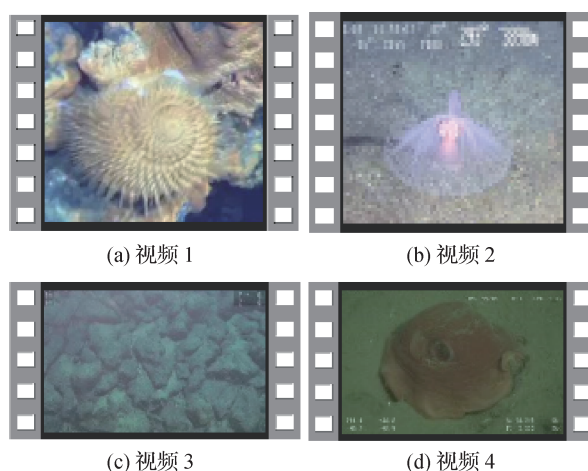


图3 水下视频源数据示例

Fig. 3 Examples of underwater videos

((a) video 1; (b) video 2; (c) video 3; (d) video 4)

2.2 数据集构建

通过大量观察,本文总结出水下视频具有以下特点。从视频画面来看:

1) 视频色彩多呈现蓝、蓝紫、蓝绿、绿、黄绿和暗黑等;

2) 视频多模糊、浑浊;

3) 补光充足多出现曝光现象,导致拍摄主体反光或画面泛白等情况;

4) 补光不足导致视频昏暗,细节内容较少。

从视频拍摄角度来看,视频在水下拍摄过程中,受设备条件限制,具有以下几种情况:

1) 镜头转动不灵活,晃动大,易失焦;

2) 镜头移动拍摄,主体易出画面框;

3) 镜头焦距缩放容易导致失焦,甚至长时间无法对焦;

4) 静止拍摄时,非固定相机会随水流轻微摆动。

为了在筛选的视频中体现上述水下视频特点,本文根据视频内容和画面进行初步筛选。视频内容包含简单的海底几乎静止的动植物、复杂的不规则岩石布局 and 移动的动植物(植物随水流)。初选出 21 个视频作为代表性样本和潜在的测试序列。再将这些时长不同的视频以 10 s 为间隔进行剪切。根据视频拍摄出现的几种情况进行二次筛选并保证视频色彩多样性,最终选出 25 个视频序列作为水下视频质量评价数据集的原始数据。

在图像/视频的质量评价中色彩信息通常是重要的评价方面之一。本文在原始视频的选取中,充分考虑视频画面色彩多样性,视频序列包含的色彩跨度大:R 通道像素均值范围在(9,162.4),G 通道像素均值范围在(12.6,184.2),B 通道像素均值范围在(18.3,198.2)。在图 4 所示的 R、G、B 3 维坐

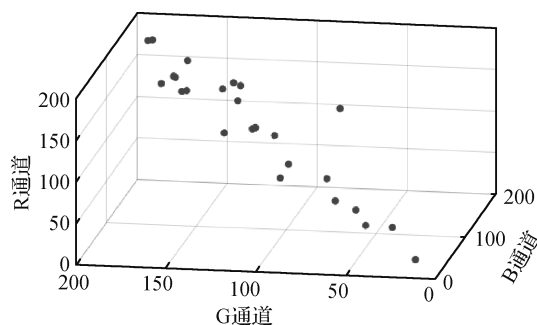


图 4 25 个视频序列的 RGB 像素平均值的 3 维分布

Fig. 4 3D distribution of average R, G and B intensities of 25 video sequences

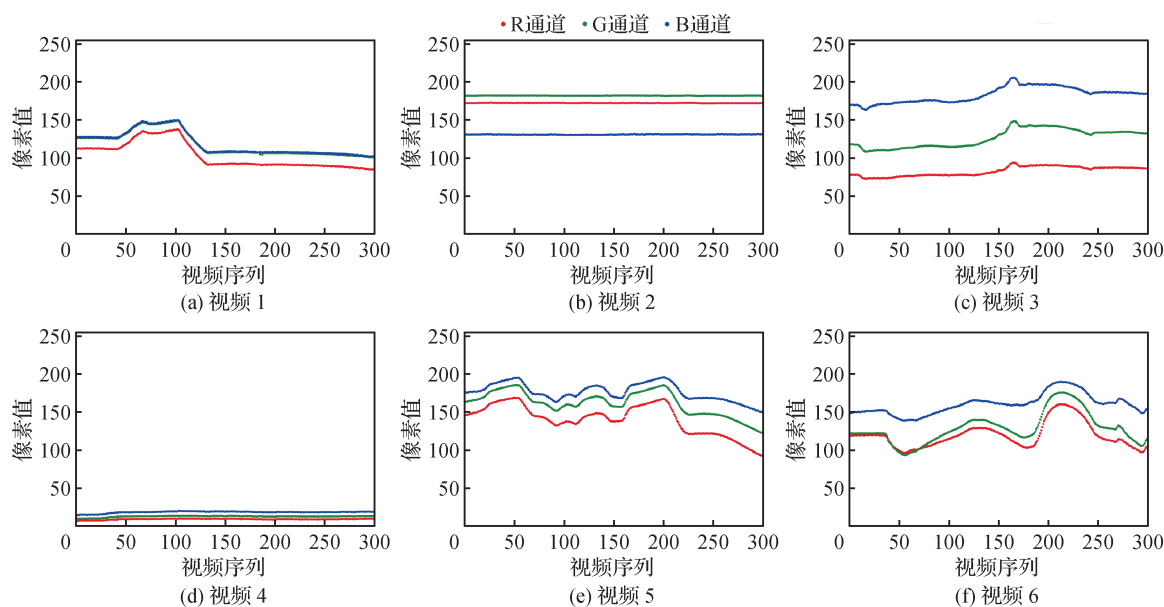


图 5 不同视频的 R、G、B 像素值的逐帧变化图

Fig. 5 The change of R, G, and B values with frames for different videos

((a) video 1; (b) video 2; (c) video 3; (d) video 4; (e) video 5; (f) video 6)

标图显示这些视频序列的色彩分布比较分散,没有集中在某一区域。

图 5 展示了 6 个代表性视频序列的 R、G、B 通道像素值逐帧变化的散点图,图下方为 3 个通道的均值。可以看出,这些视频序列色彩分布情况不同,且视频色彩既有平缓变化(如图 5(b)(d)),也有剧烈变化(如图 5(e))。

此外,比特率、帧率和内容变化是对质量分数影响较大的变量。视频运动量和空间细节会影响视觉的可感知性(Legge 和 Foley,1980;Zhong 等,2004),因此,要确保进行主观视频质量评估的视频序列具有不同的运动信息和空间信息。通常由空间感知信息(spatial information, SI)和时间感知信息(temporal information, TI)(ITU-T,1999)表示这种场景内容变化。文中通过式(1)和式(2)逐帧计算 SI 和 TI,然后取视频帧序列中的最大 SI 值和 TI 值作为视频的 S 和 T,即

$$S = \max_{\text{time}} \{ \text{std}_{\text{space}} [\text{Sobel}(P_n(i,j))] \} \quad (1)$$

$$T = \max_{\text{time}} \{ \text{std}_{\text{space}} [Q_n(i,j)] \} \quad (2)$$

式中, $Q_n(i,j) = P_n(i,j) - P_{n-1}(i,j)$, $P_n(i,j)$ 表示在时刻 n 时,视频帧像素坐标 (i,j) 点的亮度值。 $\text{Sobel}(\cdot)$ 是对各视频帧进行 Sobel 算子边缘检测(Duda 和 Hart,1973)。 $\text{std}_{\text{space}}$ 运算符表示像素矩阵内亮度值的标准偏差, $\max_{\text{time}}$ 运算符表示取参数最大值。

将所有视频的 SI 和 TI 值作散点图展示(如

图6),可以看出,除一个视频的SI和TI值比较高之外,本文最终选择的视频序列在空间和时间复杂度上的分布比较均匀,能够全面反映不同空间和时间复杂度视频的质量结果。

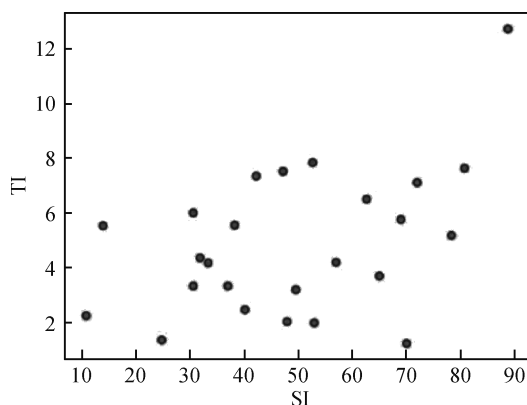


图6 25个原始视频的SI和TI值

Fig. 6 SI and TI values of 25 original videos

考虑到比特率和帧率对质量分数的影响,将最终选定的25个视频序列进行扩展。其中24个视频的比特率在780~820 kbps之间,帧率为29.97 帧/s,1个视频比特率为1 820 kbps,帧率为25 帧/s。对视频的压缩编码,使用超过原视频本身的参数没有意义,因此,本文中以表1所示参数值进一步产生质量受损视频序列,共生成250 ($25 \times 3 \times 3 + 25$)个样本。在图7中以帧率为5 帧/s的视频为例,展示了不同比特率的压缩视频的第1帧图像,可以看出,经过压缩的视频与原始视频在细节上的有明显的差别,不同比特率的压缩视频之间也能看出细微的差别。

2.3 主观质量评价

对数据集的主观质量评价,本文采用P. 910 (ITU-T, 1999)中描述的绝对分类评分ACR (absolute category rating)方法。这种经典方法时间短,并且一次评价一个视频,与典型的视频观看方式最接近。这里采用1~5分的打分制表示5级评级。

由于海底环境涉及生物、地质和海底导航等多方面范畴,很难对质量表达做出直接量化。为此,通过定性分析,结合实际环境,在深入理解视频质量的内涵和属性的基础上,设计深海视频质量的多维语义度量,从画面感受质量、视频流畅程度和主体是否可辨等几个角度进行评判。

表1 视频数据集参数的范围

Table 1 Range of video dataset parameters

参数	不同的视频编码参数		
比特率/kbps	96	200	500
帧率/(帧/s)	5	10	25

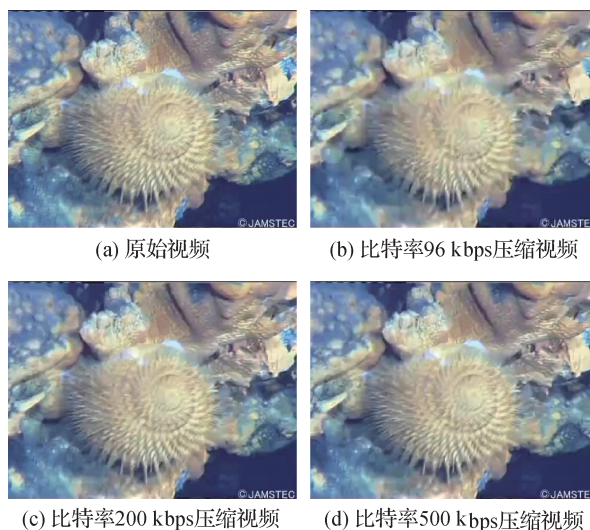


图7 帧率为5 帧/s的视频帧示例

Fig. 7 Example of a video frame with a frame rate of 5 frame/s ((a) original video; (b) compressed video with the bitrate 96 kbps; (c) compressed video with the bitrate 200 kbps; (d) compressed video with the bitrate 500 kbps)

参加这项研究的受试者由10名普通受试者和10名从事海洋图像和视频处理方面研究的受试者组成,均无视频质量评估经验。在开始实验之前为受试者做一个简短的培训,包括观看相对较好的5个水下视频序列和相对较差的5个水下视频,并熟悉测试程序。

研究中所有视频全部观看需要1 h,为尽量减少观众疲劳的影响,每隔20 min休息3 min。视频播放系统是专门为此开发的HTML5应用程序,播放界面如图8所示。为避免质量判断中的上下文和记忆效应,每次观看时,视频序列随机播放。视频播放窗口固定在屏幕中央1 400像素宽的区域,高度自动适应,右侧为打分框,播放完毕后进行打分,点击“下一个”进入下一条视频序列。由20名受试者依次进行打分,最终,得到20个评价结果。去除结果中的异常值,计算平均意见得分(mean opinion score, MOS),即得到视频主观质量结果。



图8 HTML5 质量评价应用程序界面

Fig. 8 Quality assessment application interface with HTML5

3 水下视频客观质量评价模型

本文提出的 NR-UVQA 模型,从视频图像特征、视频运动特征和编码参数 3 个方面出发,分析不同特征对视频质量的影响,将对视频质量影响较大的特征作为质量评价指标输入线性模型,最终得到水下视频质量评价模型。

3.1 全帧统计特征及分析

视频单帧的空间失真是感知视频质量的重要因素。根据自然场景统计和视频统计理论 (NSS/NVS) (Ruderman, 1994),未失真的图像表现出某种统计特性,当出现失真时,这些统计特性就会丢失。BRISQUE 算法 (Mittal 等, 2012) 使用局部亮度归一化的场景统计量化局部块损失。现将 BRISQUE 算法得到的每一视频帧的质量分数取均值作为视频质量分数,即

$$Q_s = \text{mean}\{B(P_n)\} \quad (3)$$

式中, $B(P_n)$ 表示 BRISQUE 算法计算时刻 n 时,视频帧的质量分数。

BRISQUE 算法首先使用式(4)对输入视频帧的亮度值 $I(i, j)$ 进行去均值对比度归一化处理,变换后的亮度 $\hat{I}(i, j)$ 称为 MSCN (mean subtracted contrast normalized) 系数,即

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + 1} \quad (4)$$

$$\mu(i, j) = \sum_{l=1}^M \sum_{k=1}^N w_{k,l} I_{k,l}(i, j) \quad (5)$$

$$\sigma(i, j) = \sqrt{\sum_{l=1}^M \sum_{k=1}^N w_{k,l} (I_{k,l}(i, j) - \mu(i, j))^2} \quad (6)$$

式中, $i \in \{1, 2, \dots, M\}$, $j \in \{1, 2, \dots, N\}$, M 和 N 为图像尺寸, $w = \{w_{k,l} | k = -K, \dots, K; l = -L, \dots, L\}$ 是 2 维循环对称复高斯加权重 (2D circularly-symmetric Gaussian weighting), 这里 $K = L = 3$ 。

BRISQUE 算法接下来通过拟合 MSCN 系数广义高斯分布 (generalized Gaussian distribution, GGD) (Sharifi 和 Leon-Garcia, 1995) 的形状和方差参数得到特征 b_1, b_2 。失真情况下,相邻像素之间会产生干扰, BRISQUE 将原始图像与相邻像素 (水平方向, 垂直方向, 主对角方向和次对角方向) 的 MSCN 系数两两相乘,再拟合其非对称广义高斯分布 (asymmetric generalized Gaussian distribution, AGGD) (Lasmar 等, 2009) 的形状、均值、左方差和右方差参数得到特征 $b_3, \dots, b_{18}$ (具体描述如表 2)。降低一半分辨率进行相同操作得到特征 $b_{19}, \dots, b_{36}$, 最终获得 36 个统计特征。

表 2 图像特征摘要

Table 2 Feature summary

特征 ID	特征描述	计算方式
b_1, b_2	形状, 方差	MSCN 系数 GGD 拟合
$b_3, \dots, b_6$	形状, 均值, 左方差, 右方差	水平方向 AGGD 拟合
$b_7, \dots, b_{10}$	形状, 均值, 左方差, 右方差	垂直方向 AGGD 拟合
$b_{11}, \dots, b_{14}$	形状, 均值, 左方差, 右方差	主对角方向 AGGD 拟合
$b_{15}, \dots, b_{18}$	形状, 均值, 左方差, 右方差	次对角方向 AGGD 拟合

本文中建立的水下视频数据集,样本数量小,过多的冗余参数会导致模型泛化性能差。通过分析,希望减少特征数量也能取得很好的性能。表 3 为求取原始图像和降低分辨率图像对应特征之间的皮尔森线性相关系数 (Pearson's correlation coefficient, PCC)。可以看出,除了 $b_8 - b_{26}$ 特征对,其他的特征对之间都有很高的相关性 (PCC 相关系数均 > 0.7)。为了验证只使用其中一组特征也能取得良好的效果,分别使用原始图像的 18 个特征和降低分辨率图像的 18 个特征对水下视频数据集进行建模,并计算预测值与主观质量之间的皮尔森线性相关系数 PCC 和斯皮尔曼等级秩序相关系数 (Spearman's rank order correlation coefficient, SROCC) (表 4)。

从表4的相关性分析结果可以看出,降低分辨率图像的18个特征参数能得到与使用全部特征很接近的结果,因此,本文将特征 $b_{19}, \dots, b_{36}$ 用做衡量水下视频质量的指标,记为 $F_1(f_1, \dots, f_{18})$ 。

表3 两组特征之间的皮尔森线性相关系数

Table 3 PCC between two sets of features

特征 ID	PCC	特征 ID	PCC
$b_1 - b_{19}$	0.888 5	$b_{10} - b_{28}$	0.989 0
$b_2 - b_{20}$	0.983 9	$b_{11} - b_{29}$	0.911 0
$b_3 - b_{21}$	0.860 4	$b_{12} - b_{30}$	0.806 4
$b_4 - b_{22}$	0.706 4	$b_{13} - b_{31}$	0.946 3
$b_5 - b_{23}$	0.932 0	$b_{14} - b_{32}$	0.977 3
$b_6 - b_{24}$	0.831 1	$b_{15} - b_{33}$	0.912 6
$b_7 - b_{25}$	0.878 5	$b_{16} - b_{34}$	0.789 0
$b_8 - b_{26}$	0.126 4	$b_{17} - b_{35}$	0.943 4
$b_9 - b_{27}$	0.877 5	$b_{18} - b_{36}$	0.974 3

表4 视频特征和主观质量评价结果的

PCC和SROCC相关系数

Table 4 PCC and SROCC coefficients between video features and subjective quality evaluation results

特征 ID	PCC	SROCC
$b_1, \dots, b_{36}$	0.755 0	0.739 7
$b_1, \dots, b_{18}$	0.586 7	0.576 1
$b_{19}, \dots, b_{36}$	0.720 9	0.706 6

3.2 光流运动特征及分析

与图像质量评价不同,视频需要考虑运动特性对质量的影响。光流法不仅携带物体的运动信息,还包含景物的3维结构信息,能够很好地刻画目标运动过程,因此,本文采用光流法表示视频帧序列的运动特性。

光流法是利用像素强度的时域变化和相关性来确定各自像素位置的“运动”。水下视频画面模糊、雾化严重,所以采用传统的稠密光流法,计算帧间光流矩阵。两帧之间的稠密光流矩阵采用OpenCV的calcOpticalFlowFarneback函数获得。对得到的光流矩阵提取光流运动特征,包括整体运动特性和视频中主体的运动特征。

视频整体的运动特征用帧间运动距离的均值

μ_{opt} 和方差 σ_{opt}^2 来表示,即

$$\mu_{\text{opt}} = \frac{1}{K-1} \sum_{k=1}^{K-1} d_{\text{opt}}^k \quad (7)$$

$$\sigma_{\text{opt}}^2 = \frac{1}{K-1} \sum_{k=1}^{K-1} (d_{\text{opt}}^k - \mu_{\text{opt}})^2 \quad (8)$$

$$d_{\text{opt}} = \{d_{\text{opt}}^1, d_{\text{opt}}^2, \dots, d_{\text{opt}}^{N-1}\} \quad (9)$$

$$d_{\text{opt}}^k = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \sqrt{d_{x,ij}^2 + d_{y,ij}^2} \quad (10)$$

式中, K 为视频帧数, $K-1$ 为帧间光流矩阵数, M, N 为视频尺寸。 d_{opt}^k 是第 k 帧光流矩阵的运动距离, $d_{x,ij}$ 、 $d_{y,ij}$ 分别为光流矩阵 (i, j) 点水平方向和垂直方向的运动距离。视频光流矩阵的均值和方差体现了景物运动的程度和波动(例如水下视频拍摄时,拍摄设备不灵活的转向会导致视频运动的突然增加)。

除了视频整体的运动体现,视频中主体的运动状态能体现水下视频的科学有用性,也是本文所考虑的。设置阈值 ε ,统计第 k 帧光流矩阵中超过这一阈值的像素点个数 $\text{num}_{k,\text{pix}}$ (式(11)),表示运动明显的像素点个数,并计算这些像素点的均值(式(12)) and 方差(式(13))。本文中阈值取光流矩阵的均值

$$\text{num}_{k,\text{pix}} = \sum_{i=1}^M \sum_{j=1}^N \delta_{ij} \quad (11)$$

$$\text{式中}, \delta_{ij} = \begin{cases} 0 & \sqrt{d_{x,ij}^2 + d_{y,ij}^2} < \varepsilon \\ 1 & \sqrt{d_{x,ij}^2 + d_{y,ij}^2} \geq \varepsilon \end{cases}$$

$$\mu_{\text{opt_num}} = \frac{1}{K-1} \sum_{k=1}^{K-1} \text{num}_{k,\text{pix}} \quad (12)$$

$$\sigma_{\text{opt_num}}^2 = \frac{1}{K-1} \sum_{k=1}^{K-1} (\text{num}_{k,\text{pix}} - \mu_{\text{opt_num}})^2 \quad (13)$$

将 μ_{opt} 、 σ_{opt}^2 和 $\mu_{\text{opt_num}}$ 、 $\sigma_{\text{opt_num}}^2$ 作为视频的运动特征 $F_2(f_{19}, f_{20}, f_{21}, f_{22})$ 进行讨论。

3.3 编码参数及分析

本文从视频的空间域和时间域讨论,得到视频的空间统计特征和时间运动特征,但是视频的编码特性对视频质量的影响也是不容忽视的。在数据处理阶段,很容易得到视频的分辨率(像素)、帧率(帧/s)、码率(kbps)信息,这些编码参数是视频处理非常常见也容易获取的参数,因此,本文将视频的分辨率、帧率、码率信息作为编码特征 $F_3(f_{23}, f_{24}, f_{25})$ 进行讨论。

3.4 水下视频质量评价模型

将上述视频特征:空间特征 $F_1(f_1, \dots, f_{18})$,光

流运动特征 $F_2(f_{19}, f_{20}, f_{21}, f_{22})$ 以及编码参数 $F_3(f_{23}, f_{24}, f_{25})$ 作为 25 个输入变量 f_i 建立多变量线性回归模型 (multivariate linear regression model, MLR), 每个水下视频的主观质量得分为

$$y = \sum_{i=1}^{25} \omega_i f_i \tag{14}$$

式中, ω_i 是每个特征 f_i 的加权系数。在获得每个视频的 25 个特征之后, 将 250 个测试水下视频数据集分成两个不重叠的部分: 80% 训练集, 20% 测试集建模。进行交叉验证, 重复该过程 200 次, 计算每次预测结果和主观 MOS 值之间的皮尔森线性相关系数 (PLCC) 和斯皮尔曼等级次序相关系数 (SROCC), 将相关系数的中值作为性能度量结果。为了说明每组特征对主观分数的相关程度进行如表 5 所示的相关性对比实验。

由表 5 可以看出, 视频运动特征 F_2 单独预测视频质量的结果与主观结果的相关性很低, 说明运动特性并不能很好地体现视频质量的好坏。而且在分别与 F_1 、 F_3 组合预测的结果, 提升并不明显, 甚至 F_1 、 F_2 、 F_3 预测结果与主观结果的相关性低于 F_1 、 F_3 组合的结果。因此, 最终本文采用 F_1 和 F_3 共 21 个特征值作为模型的输入变量。

表 5 单个特征和组合特征与主观质量评价结果的皮尔森线性相关系数和斯皮尔曼等级次序相关系数

Table 5 PCC and SROCC between single feature and feature combinations and subjective quality evaluation results		
特征 ID	PCC	SROCC
F_1	0.720 9	0.706 6
F_2	0.331 1	0.348 4
F_3	0.599 4	0.532 9
F_1 、 F_2	0.746 4	0.717 7
F_1 、 F_3	0.841 6	0.832 8
F_2 、 F_3	0.643 5	0.615 5
F_1 、 F_2 、 F_3	0.838 5	0.826 7

注: 加粗字体为每列最优值。

4 性能评估

使用本文建立的水下视频数据集开展实验, 与

目前最优的两种水下图像质量评价方法 UIQM (Panetta 等, 2016) 和 UCIQE (Yang 和 Sowmya, 2015), 基于自然统计特性的图像质量评价方法 BRISQUE (Mittal 等, 2012), 以及一种视频质量评价方法 VIIDEO (Mittal 等, 2016) 进行对比实验, 将与主观评价得分 (MOS) 的相关性作为指标, 评估所提出算法的性能。

使用 UIQM、UCIQE 和 BRISQUE 图像质量评价方法对水下视频进行评价时, 对每一帧视频的质量进行评价, 然后取全部视频帧的均值作为水下视频的质量评价结果。为了达到 3 种对比方法的最佳性能, 本文用自己构建的水下数据集分别对 UIQM、UCIQE 和 BRISQUE 模型进行了训练, 获得适合本数据集的权重系数线性模型, 然后用于对测试视频集进行预测。

由于这 3 个对比模型均为线性模型, 所以可以采用 3.4 节相同的方式训练。以 UIQM 方法为例, 首先对训练数据集中的每个视频序列, 计算其所有视频帧的水下图像色彩测量 UICM、清晰度测量 UISM、对比度测量 UIConM; 然后求取视频序列的 UICM 均值、UISM 均值、UIConM 均值作为线性模型的输入变量, 通过线性回归拟合主观评价结果, 得到训练好的模型和权重; 最后在测试集上进行预测。其他两个方法类似。VIIDEO 是视频质量评价模型, 能够直接用于水下视频的质量评价。

图 9 展示了几种方法在测试阶段的视频质量预测结果与主观质量评价结果的散点图, 其中横坐标 x 轴为主观质量评分, 纵坐标 y 轴为模型预测的质量评分。以主观质量评分为真值, 理想情况下预测值和真值完全匹配, 散点落在主对角线 ($y = x$) 上。因此, 图 9 中散点分布越靠近主对角线, 表示模型预测越准确。可以看出, BRISQUE 和本文方法的预测值分布接近主对角线, 说明这两种方法在水下视频质量的评价中均有良好的性能, 但本文方法使用了更少的参数, 模型的泛化性能更好; VIIDEO 方法的表现可以说是预测失败。

除了各方法与主观质量评价分数的散点图和拟合线性关系, 相关性性能评估结果如表 6 所示。从主观评分分别与 UIQM、UCIQE、VIIDEO、BRISQUE 和本文方法预测评分之间的相关系数 PLCC 和 SROCC 可以看出, VIIDEO 几乎不能得出正确的结果 (相关性系数 < 0.1), UIQM、UCIQE 的预测结

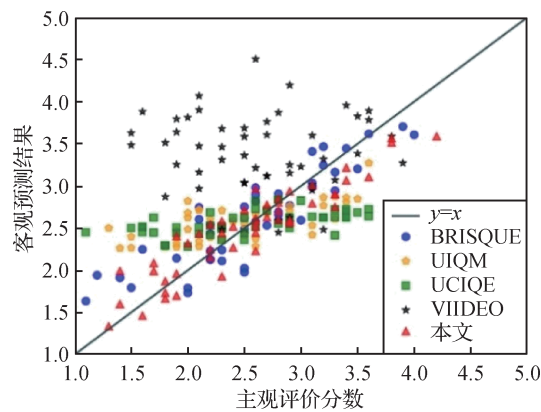


图 9 UIQM、UCIQE、BRISQUE、VIIDEO 和本文方法
主观质量分数与预测结果分数的散点图

Fig. 9 Scatter plots between the predicted scores by
different methods and subjective quality scores

表 6 PLCC 和 SROCC 性能指标结果比较

Table 6 Performance comparison of different methods

方法	PCC	SROCC
UIQM	0.410 3	0.314 1
UCIQE	0.312 4	0.338 2
BRISQUE	0.763 1	0.747 5
VIIDEO	-0.009 0	0.005 0
本文	0.841 6	0.832 8

注:加粗字体为每列最优值。

果不佳(相关系数 < 0.6),BRISQUE 可以得到一个不错的结果,本文算法则得到与主观质量分数最高的相关性($PCC = 0.840\ 8$, $SROCC = 0.832\ 2$),说明本文模型评价的结果更符合人类视觉感知。

为了更加直观地说明算法之间的质量评价结果,随机选择 6 个水下视频(图 10)进行测试,得到的评分情况如表 7。对比观察具体视频的模型评价结果与主观评价分值,可以看出本文方法和 BRISQUE 都与主观得分比较接近,其他方法对某些视频(例如图 7(e) 视频)也有更接近主观分数的预测结果。但是,单一的视频质量评价结果的准确度并不能说明该方法的好坏。本文通过计算各方法评价结果与主观分数的均方误差(mean square error, MSE)(表 8),来判断不同的方法对实际视频质量预测的稳定性。通过比较 MSE,可以认为本文方法比其他方法更稳定。



图 10 随机参数的 6 个水下视频

Fig. 10 Six underwater videos with random parameters

((a) video 1; (b) video 2; (c) video 3;
(d) video 4; (e) video 5; (f) video 6)

表 7 各类方法的预测结果比较

Table 7 Comparison of predicted scores
by different methods

方法	图 10					
	视频 1	视频 2	视频 3	视频 4	视频 5	视频 6
UIQM	2.8	2.4	2.8	2.6	2.4	2.5
UCIQE	2.6	2.4	2.7	2.7	2.2	2.5
BRISQUE	3.6	1.7	3.2	3.3	2.7	1.8
VIIDEO	2.9	3.7	4.2	3.3	4.5	3.5
本文	3.6	1.4	3.1	3.0	2.6	1.5
主观分数	3.8	1.4	3.6	2.4	2.3	1.8

表 8 各类方法预测结果的 MSE

Table 8 MSE of prediction results by different methods

	方法				
	UIQM	UCIQE	BRISQUE	VIIDEO	本文
MSE	0.312 7	0.332 0	0.165 3	1.392 2	0.113 1

注:加粗字体为最优值。

5 结 论

由于水下拍摄环境复杂,且受拍摄设备限制,拍摄的水下视频多存在画面模糊、对比度低、色彩失

真、光照不均、镜头移动晃动大和失焦等问题。针对这些问题,本文考虑水下视频特点和水下视频拍摄情况,建立水下视频数据库,并通过主观评价打分的方式对视频质量进行标注;然后从空间维度、时间维度和视频参数几个方面进行讨论,提出了一种基于空域自然特性与编码参数的水下视频质量评价方法。该方法无需参考视频,能够很好地适用于小样本数据,准确性高,结果符合人类视觉感知。

由于实验条件所限,本文还存在一些不足。首先,色彩信息通常对图像/视频的质量评价有重要影响,但在主观质量评价时,没有要求评价主体从图像色彩上对视频质量进行评价,而是要求从画面质量、视频流畅度和主体可辨的总体感受性进行评价。主要是考虑:一方面本文在制作深海视频数据集时尽量考虑了色彩的多样性,评价者对画面质量的评价本身已隐含了对颜色的视觉感受性;另一方面该方法视频数据大部分来自水深 500 ~ 5 000 m 的深海,在无光环境下借助人工辅助光源拍摄到的图像或视频广泛存在色偏、色彩不够丰富等问题,人们(在大气环境中建立起来)的普遍色彩经验在这种情况下不适用,对色彩方面的评价难度较大。

下一步工作将更多地考虑水下视频各方面的特殊性,研究更合理的主观质量评价方法。本文数据集在数据量上也有局限,仍无法很好地涵盖水下视频的特点,在接下来的工作中将从数据的多样化出发,补充数据集,增加模型鲁棒性。

参考文献 (References)

- De Simone F, Tagliasacchi M, Naccari M, Tubaro S and Ebrahimi T. 2010. A H. 264/AVC video database for the evaluation of quality metrics//Proceedings of 2010 IEEE International Conference on Acoustics, Speech and Signal Processing. Dallas: IEEE: 2430-2433 [DOI: 10.1109/ICASSP.2010.5496296]
- Duda R O and Hart P E. 1973. Pattern Classification and Scene Analysis. New York: Wiley
- Field D J. 1999. Wavelets, vision and the statistics of natural scenes. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 357(1760): 2527-2542 [DOI: 10.1098/rsta.1999.0446]
- Guo J C, Li C Y, Zhang Y and Gu X Y. 2017. Quality assessment method for underwater images. Journal of Image and Graphics, 22(1): 1-8 (郭继昌, 李重仪, 张艳, 顾翔元. 2017. 面向水下图像的质量评价方法. 中国图象图形学报, 22(1): 1-8) [DOI: 10.11834/jig.20170101]
- ITU-R. 2002. Methodology for the subjective assessment of the quality of television pictures. Recommendation BT. 500-11 [EB/OL]. [2020-01-01]. <https://www.itu.int/rec/R-REC-BT.500>
- ITU-R. 2007. Methodology for the subjective assessment of video quality in multimedia applications. Recommendation BT. 1788 [EB/OL]. [2020-01-01]. <https://www.itu.int/rec/R-REC-BT.1788/en>
- ITU-T. 1999. Subjective video quality assessment methods for multimedia applications. Recommendation P. 910 [EB/OL]. [2020-01-01]. <https://www.itu.int/rec/T-REC-P.910/en>
- Lasmar N E, Stitou Y and Berthoumieu Y. 2009. Multiscale skewed heavy tailed model for texture analysis//Proceedings of the 16th IEEE International Conference on Image Processing. Cairo: IEEE: 2281-2284 [DOI: 10.1109/ICIP.2009.5414404]
- Legge G E and Foley J M. 1980. Contrast masking in human vision. Journal of the Optical Society of America, 70(12): 1458-1471 [DOI: 10.1364/JOSA.70.001458]
- 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. 2011. Blind image quality assessment: from natural scene statistics to perceptual quality. IEEE Transactions on Image Processing, 20(12): 3350-3364 [DOI: 10.1109/TIP.2011.2147325]
- Moreno-Roldán J M, Luque-Nieto M Á, Poncela J and Otero P. 2017. Objective video quality assessment based on machine learning for underwater scientific applications. Sensors, 17(4): 664 [DOI: 10.3390/s17040664]
- Moreno-Roldán J M, Poncela J, Otero P and Bovik A C. 2018. A no-reference video quality assessment model for underwater networks. IEEE Journal of Oceanic Engineering, 45(1): 342-346 [DOI: 10.1109/JOE.2018.2869441]
- Panetta K, Gao C and Agaian S. 2016. Human-visual-system-inspired underwater image quality measures. IEEE Journal of Oceanic Engineering, 41(3): 541-551 [DOI: 10.1109/JOE.2015.2469915]
- Ruderman D L. 1994. The statistics of natural images. Network: Computation in Neural Systems, 5(4): 517-548 [DOI: 10.1088/0954-898X/5_4_006]
- 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, Soundararajan R, Bovik A C and Cormack L K.

2010. Study of subjective and objective quality assessment of video. IEEE Transactions on Image Processing, 19(6): 1427-1441 [DOI: 10.1109/TIP.2010.2042111]
- Sharifi K and Leon-Garcia A. 1995. Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video. IEEE Transactions on Circuits and Systems for Video Technology, 5(1): 52-56 [DOI: 10.1109/76.350779]
- Wang Y, Li N, Li Z Y, Gu Z R, Zheng H Y, Zheng B and Sun M M. 2018. An imaging-inspired no-reference underwater color image quality assessment metric. Computers & Electrical Engineering, 70: 904-913 [DOI: 10.1016/j.compeleceng.2017.12.006]
- Wang Y, Song W, Fortino G, Qi L Z, Zhang W Q, and Liotta A. 2019. An experimental-based review of image enhancement and image restoration methods for underwater imaging. IEEE Access, 7: 140233-140251 [DOI: 10.1109/ACCESS.2019.2932130]
- Yang M and Sowmya A. 2015. An underwater color image quality evaluation metric. IEEE Transactions on Image Processing, 24(12): 6062-6071 [DOI: 10.1109/TIP.2015.2491020]
- Zhong Y, Richardson I, Sahraie A and McGeorge P. 2004. Influence of task and scene content on subjective video quality//Campilho A, Kamel M, eds. International Conference on Image Analysis and Recognition. Porto: Springer: 295-301 [DOI: 10.1007/978-3-540-30125-7_37]

作者简介



宋巍, 1977年生, 女, 教授, 主要研究方向为计算机视觉、图像/视频处理、海洋大数据分析。

E-mail: wsong@shou.edu.cn



王建, 通信作者, 女, 讲师, 主要研究方向为海洋大数据分析、遥感技术、空间信息技术。

E-mail: wangjian@shou.edu.cn

刘诗梦, 女, 硕士研究生, 主要研究方向为图像/视频处理、视频质量评价。E-mail: ml70500823@st.shou.edu.cn

黄冬梅, 女, 教授, 博士生导师, 主要研究方向为海洋遥感技术、空间地理信息系统。E-mail: dmhuang@shou.edu.cn

王文娟, 女, 讲师, 主要研究方向为软件工程, 水产品冷链物流与质量溯源, workflow 技术。E-mail: wangwj@shou.edu.cn