

基于深度网络和视觉特性的无参考图像质量评价方法

高方远¹, 何立火²

(1.北京航空航天大学数学与系统科学学院, 北京 102206)

(2.西安电子科技大学电子工程学院, 陕西 西安 710071)

[摘要] 图像质量评价是图像处理和计算机视觉领域的基础性问题, 对于视觉信息的采集、处理和分析系统性能的评判具有重要的意义. 现有的无参考型图像质量评价方法都是基于自然统计规律的, 或者构建单一网络模型, 并未考虑视觉感知特性, 使得最终的评价结果与主观感受间存在较大差异. 为此, 本文提出一种结合多种网络特性和最差视觉感知特性的无参考型图像质量评价方法. 首先, 提取图像的去均值对比度归一化特征, 将特征图进行重叠分块; 然后, 构建 VGG 与 Inception 相结合的深度网络, 对图像块提取深度感知特征; 最后, 将分块图像的质量分数集合进行排序, 利用视觉感知最差情况加权策略对序列进行加权求和, 得到最终的图像质量分数. 在国际公开的质量评价数据库 CSIQ、LIVE 和 TID2013 上的实验结果表明, 本文方法取得了优于现有方法的主客观一致性性能.

[关键词] 图像质量评价, 深度网络, 视觉特性, 最差情况加权策略

[中图分类号] TP391 **[文献标志码]** A **[文章编号]** 1001-4616(2019)03-0020-07

No-reference Image Quality Assessment Based on Deep Network and Visual Characteristics

Gao Fangyuan¹, He Lihuo²

(1.School of Mathematics and Systems Science, Beihang University, Beijing 102206, China)

(2.School of Electronic Engineering, Xidian University, Xi'an 710071, China)

Abstract: Image quality assessment (IQA) is a fundamental problem in image processing and computer vision. IQA is very important for the performance evaluation of acquisition, processing and analysis systems of visual information. Most of the existing no-reference (NR) image quality assessments are based on the natural statistical characteristics, or design deep network to predict the image quality. These methods have not considered the visual characteristics, which produces the large difference in image quality score between the objective and subjective methods. To overcome this problem, this paper proposes a NR-IQA method based on the deep network and visual characteristics. Firstly, mean subtracted contrast normalized map of the image is calculated, and the map is resampled into overlapped patches randomly. Then the deep network is designed by combining the VGG net and Inception net to predict the quality scores of patches. Finally, the scores are ranked and the worst-case pooling is utilized to weight the scores to obtain the image quality score. Experimental results on public CSIQ, LIVE and TID2013 databases show that the proposed method performs consistently with the subjective perception and has a better performance than state-of-the-art methods.

Key words: image quality assessment, deep network, visual characteristic, worst-case pooling mechanism

图像作为传递信息的主要载体之一, 不管在海、陆、空的未知探索, 还是在文、体、商的日常生活中都发挥着十分重要的作用. 然而, 由于成像和采集设备的不完善, 处理和分发技术的缺陷, 以及重建和显示系统的限制, 图像在获取、传输、处理、存储和重现的过程中会产生各种各样的失真, 导致最终得到的图像会存在一定程度的降质, 从而会对人们理解和认识客观世界带来很大的障碍. 因此, 设计一种与人类对实际场景感知质量相一致的图像质量评价方法, 并用来优化和提升图像质量评价采集、处理和显示系统, 对于

收稿日期: 2019-07-05.

基金项目: 国家自然科学基金 (61876146).

通讯联系人: 何立火, 博士, 副教授, 研究方向: 图像质量评价. E-mail: lhhe@mail.xidian.edu.cn

认识客观世界的自然规律具有非常重要的意义. 对于图像质量来说,最为直接的评价方法是主观评价方法,依靠多人观看一幅图像并进行质量的综合评判,费时费力、成本高^[1]. 因此,迫切需要设计一种能自动、精确地预测和评价图像感知质量的客观评测方法.

传统的图像质量评价方法主要是通过对比原始图像和失真图像像素之间或者特征之间的差异性来预测失真图像的质量. 基于像素域的方法主要有均方误差(mean squared error, MSE)和峰值信噪比(peak signal-to-noise ratio, PSNR),都未考虑像素之间的结构关系. 为此, Wang Z 等人提出了结构相似度方法(structural similarity, SSIM)^[2],主要基于人眼对于图像结构信息的敏感性,构建和比较原始图像与失真图像之间的结构信息相似度,衡量和预测图像质量. 后续, Zhang L 等人又提出了基于特征的结构相似度方法(feature similarity index, FSSIM)^[3]. 而 Sheikh H R 等人则从信息论角度,将质量评价过程建模为信息提取的过程,构建了视觉信息保真度评价方法(visual information fidelity, VIF)^[4]. 此外,利用广义高斯分布(generalized Gaussian distribution, GGD)、KL 散度(Kullback-Leibler divergence, KLD)、梯度强度(gradient strength)、视觉稀疏性(visual sparsity)^[5]、自由能量最小化准则(free energy minimization)等视觉特性还设计了一系列图像质量评价方法.

上述方法都需要参考原始图像或者原始图像的特征信息,这通常不能满足实际应用场景的需求,主要原因是实际场景中很难或者不能获取到原始图像. 因此,需要设计和构建一种不依赖于原始图像信息的图像质量评价方法,即无参考型图像质量评价方法. 该方法主要是构建图像失真模型和质量预测方法,可分为两类方法. 第一类主要基于传统提取特征构建质量评价方法,如基于自然场景统计特性(natural scene statistics, NSS)^[6]、去均值对比度归一化特征(mean subtracted contrast normalized, MSCN)^[7]、局部二值模式(local binary pattern, LBP)等特征构建的图像质量评价方法,预测图像的“非自然性”和失真程度. 但是这些方法都未能达到与视觉感知较一致的评价结果. 为此,学者们提出了第二类方法,主要基于深度网络提取图像特征并构建无参考型质量评价方法. Doermann D 等人提出了基于卷积神经网络模型的无参考型图像质量评价方法^[9],构建了一个 5 层的简单网络结构,获得较好效果. 同年, Tang H X 等人提出了基于半监督的校正网络评价图像质量^[8]. 随后, Doermann D 等人改建了原来的网络结构,构建了新的多任务卷积神经网络^[10],可同时评价失真程度和失真类型. 此后, Lee S 等人构建了一个更深层的网络模型^[11],并且设计了多个卷积网络同时提取不同类型的特征. Ren H Y 等人则利用增强型对抗生成网络构建无参考型图像质量评价方法^[12],取得了较好的效果. 最近, Wiegand T 等人提出了基于卷积神经网络的无参考型和全参考型质量评价方法^[13]. Lee S 等人改进了之前的工作,使其更能捕获图像的失真特性^[14].

但是现有的这些基于深度网络的无参考型图像质量评价方法存在一些局限性. 它们在设计网络结构模型的时候并未考虑视觉感知特性,部分方法训练过程相对复杂,模型训练的数据量不足. 这些问题导致模型扩展性较差,与主观感知的结果一致性不好. 为此,本文提出了一种基于深度特征和感知特性的无参考型图像质量评价方法. 主要结合了两种网络提取图像的感知特征,并利用视觉特性中的最差情况融合机制预测图像质量,获得了较好的主客观一致性. 同时,模型结构相对简单,易于扩展和实际应用.

1 基于深度网络和视觉特性的无参考型图像质量评价方法

本文提出了一种新的基于深度网络和视觉特性的无参考型图像质量评价方法. 首先,提取去均值对比度归一化特征,利用随机采样与重采样的方法对图像进行分块;然后,构建 VGG 深度特征提取网络与 Inception 感知网络,预测每一块图像的感知质量;最后,基于视觉感知的最差感知效应,构建最差情况加权融合机制,获得最终的图像质量. 算法总体框图如图 1 所示.

1.1 去均值对比度归一化特征提取

给定一幅图像 I ,提取其去均值对比度归一化特征:

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

式中, (i, j) 代表图像的第 i 行、第 j 列的像素点. C 为非零常数,防止分母为零. μ 和 σ 分别为图像的均值和方差. $k \in \{1, 2, 3\}$, 表示彩色图像的 R、G 和 B 3 通道. 对于灰度图像,单通道提取去均值归一化特征;对于彩色图像,分别针对 3 个彩色通道单独提取去均值归一化特征.

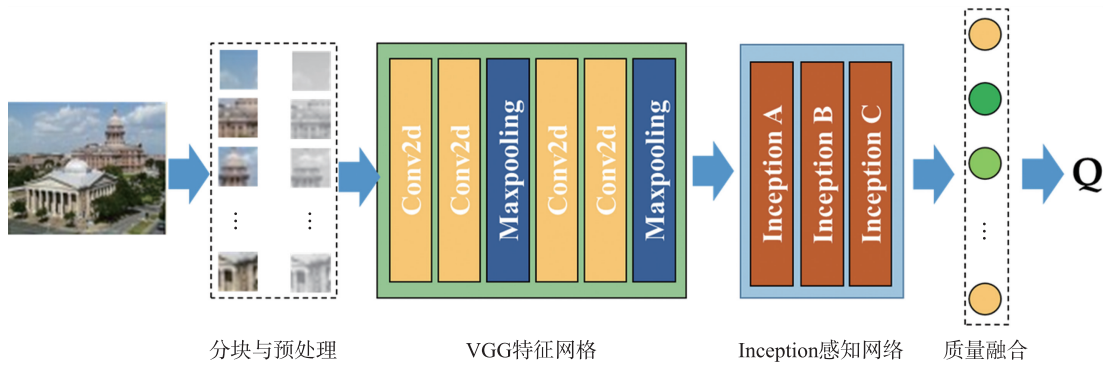


图 1 基于深度网络与感知特性的无参考型图像质量评价框图

Fig. 1 The proposed NR-IQA framework based on deep network and perceptual characteristics

针对提取特征后的图像进行重采样和分块,给定图像 \hat{I} ,重采样图像块

$$I_l = \text{resample}(\hat{I}), l \in \{1, 2, \dots, N\}, \quad (2)$$

式中, N 为块的数量,块大小为 $64 * 64$.

1.2 基于 VGG 和 Inception 的深度网络构建与学习

(1) 网络构建

算法框图 1 所示,本文构建的深度网络分为两个模块:VGG 的全卷积网络和 Inception 的感知特征网络. VGG 前面 4 层的全卷积网络,与视觉的感受野的原理相一致,因此本文利用了 VGG 网络的前 4 层卷积作为提取视觉的初级感知特征. 其结构参数为:

- C1-1 卷积层,输入 $224 * 224 * 3$,滤波器大小: $3 * 3 * 3$,滤波器个数:64,输出: $224 * 224 * 64$,
- C1-2 卷积层,输入 $224 * 224 * 3$,滤波器大小: $3 * 3 * 3$,滤波器个数:64,输出: $224 * 224 * 64$,
- P1 池化层,输入: $224 * 224 * 64$,滤波器大小: $2 * 2$,滤波器个数:64,输出: $112 * 112 * 64$,
- C2-1 卷积层,输入: $112 * 112 * 64$,滤波器大小: $3 * 3 * 64$,滤波器个数:128,输出: $112 * 112 * 128$,
- C2-2 卷积层,输入: $112 * 112 * 64$,滤波器大小: $3 * 3 * 64$,滤波器个数:128,输出: $112 * 112 * 128$,
- P2 池化层,输入: $112 * 112 * 128$,滤波器大小: $2 * 2$,滤波器个数:128,输出: $56 * 56 * 128$.

在质量感知特征提取网络中,基于 Inception 的结构,设计了基于多种 Inception 类型堆叠的深度网络结构. 在框图 1 中,Inception A 堆叠了 3 个,Inception B 堆叠了 2 个,Inception C 堆叠了 4 次,它们的结构如图 2 所示.

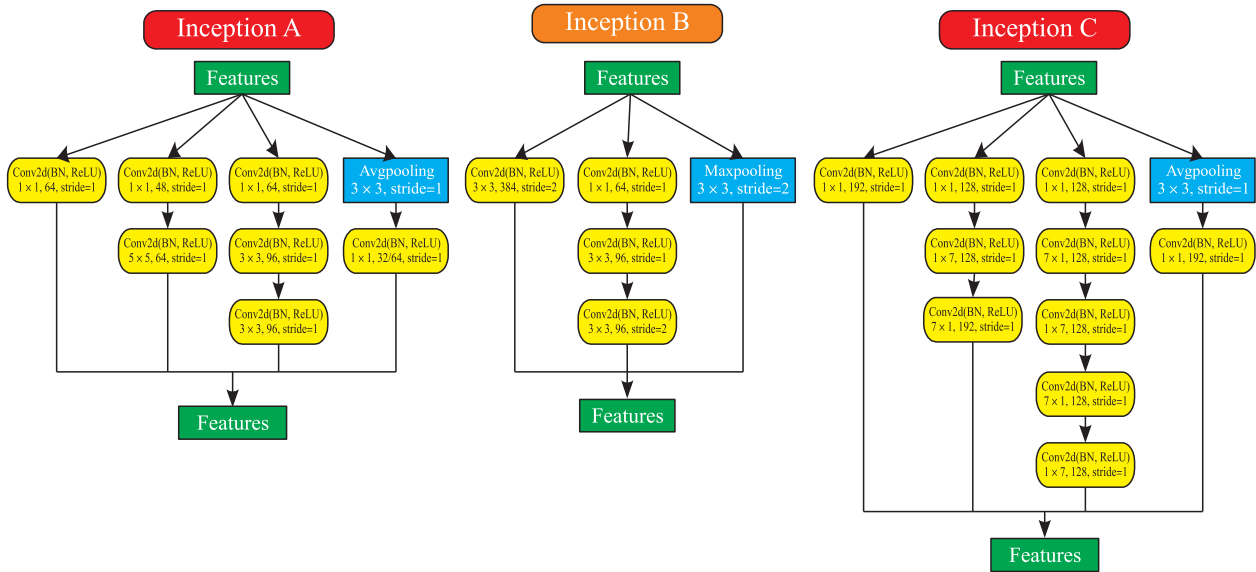


图 2 3 种 Inception 网络结构

Fig. 2 Three types of Inception networks

(2) 模型训练和测试

在模型训练过程中,利用 VGG 和 Inception 网络在 ImageNet 上训练好的参数,对本文构建的模型进行参数初始化. ImageNet 包含了较多的图像内容,训练好的网络能够对不同内容进行识别,对质量评价特征的提取具有一定的内容感知的基础,在此基础上进行本文质量评价模型的训练,既能较快又能较好地达到图像质量预测的目的. 训练过程中,还是利用了质量评价数据库的 80% 随机选取的图像,分块之后每块的质量标签为该图像的主观评价分数.

在测试过程中,针对测试图像提取与训练时候一致的分块方式和去均值归一化特征提取方法,得到分块图像特征. 利用训练好的感知质量预测网络对图形块进行质量预测,获得图像块的感知质量分数

$$Q_l = \text{VGG_Inception}(I_l). \quad (3)$$

1.3 基于视觉感知特性的质量预测

视觉感知的过程通常是不均匀的,这除了体现在视觉显著性上,还体现在最差视觉感知特性,即人类视觉系统对于图像中局部出现较大失真的情况,人类就会降低对整个图像的质量感知,判定为较低质量分数的图像. 因此,本文提出了利用分段加权的方法,融合块图像质量分数,得到整体图像的质量分数.

首先,对块图像质量分数进行排序大到小排序

$$\{\hat{Q}_1, \hat{Q}_2, \dots, \hat{Q}_N\} = \text{rank}(\{Q_1, Q_2, \dots, Q_N\}). \quad (4)$$

然后,对排序后的块质量分数进行加权求和,其中权值为

$$w_l = 1 + \frac{1}{(M-1)} \cdot \left\lfloor \frac{l-1}{\lfloor \frac{N}{M} \rfloor} \right\rfloor, \quad (5)$$

式中, M 为将排序后的序列按照序号分成的几个层级. 本文中 $M=2$, 即将前一半评分好一些的保持原始值,将后面一半评分差一些的加大其权重,得到最终质量分数

$$Q = \frac{1}{\sum_l w_l} \sum_l w_l \cdot \hat{Q}_l. \quad (6)$$

2 实验结果与分析

为验证本文方法的有效性和先进性,设计了主客观一致性实验.

数据库:本文方法在国际公认的图像质量评价数据库上进行实验,包括 LIVE 数据库^[15]、CSIQ 数据库^[16]和 TID2013 数据库^[17]. LIVE 数据库包含 982 张失真图像,5 种失真类型;CSIQ 数据库包含 900 张失真图像,6 种失真类型;TID2013 数据库 3000 张失真图像,24 种失真类型.

对比方法:对比典型的全参考型方法,有峰值信噪比方法 (PSNR)、结构相似度评价方法 (SSIM, Wang et al. 2004)、特征结构相似度方法 (FSIM, Zhang et al. 2011),主要是为图像质量评价做最基准的参考. 对比的基于深度网络的无参考型图像质量评价方法有 CNN++ (2015)^[10]、BIECON (2017)^[11]、deepIQA (2018)^[13]、DIQA (2019)^[14],它们是目前具有较好主客观一致性的方法.

评价准则:通过质量评价算法得出的预测质量分数,使用准确性、单调性和一致性评价标准来评估客观评价方法的分数与主观感知分数之间的一致程度. 主要包括皮尔逊相关系数 (Pearson correlation coefficients, PLCC) 和 Spearman 等级次序相关系数 (Spearman's rank correction coefficient, SROCC)^[18],计算公式如下:

$$PLCC = \frac{\sum_i^n (q_i - \bar{q})(q_i^* - \bar{q}^*)}{\sqrt{\sum_i^n (q_i - \bar{q})^2} \sqrt{\sum_i^n (q_i^* - \bar{q}^*)^2}}, \quad (7)$$

式中, q_i 表示第 i 幅图像的客观评价分数经过非线性拟合曲线映射后的分数, q_i^* 表示主观评价分数; \bar{q} 和 \bar{q}^* 分别表示所有预测分数与主观评测分数的平均值, n 则为用于测试的全部图像的总数.

$$SROCC = 1 - \frac{6 \sum_i^n (rq_i - rq_i^*)^2}{n(n^2 - 1)}, \quad (8)$$

式中, rqi 和 rqi^* 分别表示主观和客观分数分别按相同顺序(由小到大,或由小到大)排序后,对于第 i 幅图像的成绩在各自序列中的序号.

实验结果如表 1 所示,本文方法相对于现有的大多数典型的基于深度的无参考型方法, CNN++和 BIECON, 具有更好的主客观一致性, 与 deepIQA 和 DIQA 具有差不多的性能. 在 CSIQ 和 LIVE 数据库上, 甚至比典型的全参考型方法(PSNR、SSIM 和 FSIM)都具有更好的性能, 能很好地反映图像的视觉感知质量.

表 1 本文方法与对比方法的实验结果

Table 1 Comparison results between the proposed method and reference methods				
方法	指标	TID2013	CSIQ	LIVE
PSNR	SROCC	0.889	0.806	0.866
	PLCC	0.847	0.800	0.856
SSIM	SROCC	0.856	0.876	0.913
	PLCC	0.867	0.861	0.906
FSIM	SROCC	0.963	0.924	0.964
	PLCC	0.932	0.912	0.960
CNN++	SROCC	0.843	—	0.928
	PLCC	0.804	—	0.897
BIECON	SROCC	0.717	0.815	0.958
	PLCC	0.762	0.823	0.960
deepIQA	SROCC	0.885	0.871	0.960
	PLCC	0.913	0.891	0.971
DIQA	SROCC	0.825	0.884	0.975
	PLCC	0.850	0.915	0.977
本文方法	SROCC	0.811	0.948	0.964
	PLCC	0.840	0.960	0.965

表 2~4 给出了本文方法在 LIVE、CSIQ 和 TID2013 数据库上的单独失真的评价性能. 从表中可以看出, 本文方法对于不同失真类型都能获得较好的主客观一致性. 但是对于 TID2013 数据库中的第 15 种类型失真效果有些欠缺. 这主要是该失真类型为局部块的强度变化, 由于这种局部块的突变, 视觉很容易感知这种变化和失真. 算法中的最差感知策略其实已经考虑了这种情况, 但是块质量的融合过程并未通过数据进行训练, 没有完全符合数据的特性, 因此主客观一致性有所欠缺. 总体来说, 本文方法对于压缩失真、模糊失真、颜色失真、亮度失真等大多数失真都获得了较好的性能.

表 2 本文方法在 LIVE 数据库上的结果

Table 2 Results on LIVE database				
LIVE	SROCC	PLCC	KROCC	RMSE
ff	0.943 8	0.911 0	0.807 0	11.024 7
gblur	0.965 2	0.960 6	0.848 9	8.779 2
jpeg	0.978 2	0.984 8	0.888 6	5.672 8
jp2k	0.948 9	0.960 3	0.794 9	10.441 9
wn	0.982 6	0.898 0	0.900 1	6.821 7
All	0.963 7	0.965 0	0.846 8	6.461 6

表 3 本文方法在 CSIQ 数据库上的结果

Table 3 Results on CSIQ database				
CSIQ	SROCC	PLCC	KROCC	RMSE
awgn	0.952 6	0.954 0	0.771 9	0.058 1
blur	0.947 8	0.962 7	0.897 6	0.072 2
contrast	0.792 2	0.797 6	0.804 3	0.125 7
fnoise	0.843 0	0.824 4	0.764 1	0.096 5
jpeg	0.967 5	0.989 1	0.802 3	0.059 6
jp2k	0.964 0	0.957 2	0.845 6	0.078 6
All	0.948 3	0.959 9	0.813 8	0.084 9

表 4 本文方法在 TID2013 数据库上的结果

Table 4 Results on TID2013 database

TID2013	SROCC	PLCC	KROCC	RMSE
#1	0.978 5	0.978 6	0.653 3	0.463 6
#2	0.948 5	0.944 3	0.799 2	0.239 7
#3	0.968 5	0.976 4	0.813 3	0.458 1
#4	0.914 6	0.926 5	0.720 0	0.238 1
#5	0.968 5	0.990 1	0.764 6	0.479 9
#6	0.933 9	0.937 3	0.846 7	0.218 1
#7	0.923 1	0.935 6	0.733 3	0.414 6
#8	0.917 7	0.944 7	0.553 3	0.242 8
#9	0.900 8	0.962 0	0.626 6	0.692 7
#10	0.941 5	0.986 0	0.673 3	0.915 6
#11	0.922 3	0.930 4	0.680 0	0.983 2
#12	0.853 8	0.934 7	0.406 7	1.240 8
#13	0.943 6	0.916 0	0.673 3	0.560 2
#14	0.874 7	0.902 8	0.589 3	0.601 6
#15	0.167 7	0.188 6	0.226 7	0.759 3
#16	0.681 0	0.744 1	0.320 4	0.700 2
#17	0.803 1	0.813 9	0.440 0	0.692 0
#18	0.843 9	0.820 1	0.653 3	0.293 2
#19	0.948 1	0.931 7	0.769 0	0.369 5
#20	0.917 7	0.933 3	0.655 3	0.688 5
#21	0.955 4	0.962 2	0.839 8	0.216 7
#22	0.932 5	0.928 0	0.771 2	0.310 0
#23	0.908 5	0.970 9	0.613 3	0.793 5
#24	0.932 3	0.961 6	0.879 9	0.219 8
All	0.810 9	0.839 9	0.756 0	0.302 4

3 总结与展望

本文提出了一种基于视觉感知特性与深度特征相融合 of 无参考型图像质量评价方法. 该方法首先对图像进行分块和去均值对比度归一化特征提取, 然后构建结合 VGG 网络和 Inception 网络特性的深度模型, 预测图像块的感知质量, 最后基于视觉感知的特性, 设计图像块质量的融合方法, 获得最终的图像质量. 实验验证了该方法与主观评价之间具有较好的一致性, 也比现有的方法具有更好的性能. 但是, 基于深度网络的无参考型图像质量评价方法具有不可解释性, 并且也较少地考虑视觉感知的特性, 与多媒体大数据场景质量评价的应用还有较大距离. 因此, 在后续的研究中, 设计结合更多视觉特性以及能在大规模场景中应用的图像质量评价方法将具有更广阔的前景.

[参考文献]

- [1] WANG Z, BOVIK A C. Modern image quality assessment[M]. San Rafael, CA, USA: Morgan & Claypool, 2006.
- [2] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE transactions on image processing, 2004, 13(4): 600–612.
- [3] ZHANG D, ZHANG L, MOU X Q, et al. FSIM: a feature similarity index for image quality assessment[J]. IEEE transactions on image processing, 2011, 20(8): 2378–2386.
- [4] SHEIKH H R, BOVIK A C. Image information and visual quality[J]. IEEE transactions on image processing, 2006, 15(2): 430–444.
- [5] HE L H, TAO D G, LI X L, et al. Sparse representation for blind image quality assessment[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Providence, Rhode Island, USA, 2012: 1146–1153.
- [6] SHEIKH H R, BOVIK A C, CORMACK L. No-reference quality assessment using natural scene statistics: JPEG2000[J]. IEEE transactions on image processing, 2005, 14(11): 1918–1927.

- [7] KUNDU D, EVANS B L. Spatial domain synthetic scene statistics[C]//48th Asilomar Conference on Signals, Systems and Computers(Asilomar). Pacific Grove, California, USA, 2014:948-954.
- [8] TANG H X, NEEL J, ASHISH K. Blind image quality assessment using semi-supervised rectifier networks[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Columbus, Ohio, USA, 2014:2877-2884.
- [9] KANG L, YE P, LI Y, et al. Convolutional neural networks for no-reference image quality assessment[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Columbus, Ohio, USA, 2014:1733-1740.
- [10] KANG L, YE P, LI Y, et al. Simultaneous estimation of image quality and distortion via multi-task convolutional neural networks[C]//IEEE International Conference on Image Processing(ICIP). Quebec City, Canada, 2015:2791-2795.
- [11] KIM J, LEE S. Fully deep blind image quality predictor[J]. IEEE journal of selected topics in signal processing, 2017, 11(1): 206-230.
- [12] REN H Y, CHEN D Q, WANG Y Z. RAN4IQA: restorative adversarial nets for no-reference image quality assessment[C]//The Thirty-Second AAAI Conference on Artificial Intelligence(AAAI). New Orleans, Louisiana, USA, 2018:7308-7314.
- [13] BOSSE S, MANIRY D, MÜLLER K, et al. Deep neural networks for no-reference and full-reference image quality assessment[J]. IEEE transactions on image processing, 2018, 27(1): 206-219.
- [14] KIM J, NGUYEN A, LEE S. Deep CNN-based blind image quality predictor[J]. IEEE transactions on neural networks and learning systems, 2019, 30(1): 11-24.
- [15] SHEIKH H R, WANG Z, CORMACK L, et al. Live image quality assessment database release 2(LIVE II) [DB/OL]. [2019-01-18]. <http://live.ece.utexas.edu/research/quality>.
- [16] LARSON E C, CHANDLER D M. Categorical image quality(CSIQ) database [DB/OL]. [2019-01-18]. <http://vision.okstate.edu/csiq>.
- [17] PONOMARENKO N, IEREMEIEV O, LUKIN V, et al. A new color image database TID2013: Innovations and Results[M]//Blanc-Talon J, Kasinski A, Philips W, et al. Advanced Concepts for Intelligent Vision Systems. Berlin, Heidelberg: Springer-Verlag, 2013.
- [18] VQEG. Validation of reduced-reference and no-reference objective models for standard definition television, Phase I [DB/OL]. [2019-01-18]. <http://www.vqeg.org/>.

[责任编辑:顾晓天]