

文章编号 1004-924X(2025)18-2944-18

基于图像画面信息及其视觉感知的 无参考图像质量评价

姚军财^{1,2*}, 申 静¹

(1. 南京工程学院 计算机工程学院, 江苏 南京 211167;
2. 西安交通大学 信息与通信工程学院, 陕西 西安 710049)

摘要:为了提出一种符合视觉感知且精度、泛化性和复杂性综合效益较高的图像质量评价(IQA)方法,以实现更好地控制图像处理并满足其实际应用需求,结合人眼视觉系统(HVS)的对比敏感、亮度感知非线性、视觉感知舒适度和纹理复杂度感知等特性,基于图像亮度、色度、纹理、清晰度和局部对比度等画面信息和特征,提出了一种无参考 IQA 方法及其模型 BCTCSP。在该方法中,首先分析图像亮度、灰度分布、颜色深度及其饱和度、亮度非线性感知、视觉感知舒适度等与图像质量之间的关系,提出图像亮度色度及其视觉感知对 IQA 贡献和影响的量化和计算方法;再结合灰度梯度共生矩阵计算和统计图像纹理特征,并采用纹理加权平均和 HVS 复杂目标感知模型,提出量化和计算图像纹理信息熵及其视觉感知对 IQA 的贡献和影响的方法;接着,计算图像每一点的对比度值和觉察阈值,结合 HVS 的对比敏感度特性及其模型和掩蔽特性,量化和计算图像局部对比度及其视觉感知对 IQA 的贡献和影响;之后,采用锐度、信噪比、高频成分占比和分辨率 4 个因子描述图像清晰度,提出其量化和计算方法,得到图像画面的清晰度指标;最后综合 4 个方面的因素,构建 IQA 模型,量化其度量标准。同时,采用 6 个国际开源数据库(TID2013,CSIQ,LIVE,IVC,SPAQ 和 Koniq 10k)中 6430 幅失真图像进行测试和验证,并就精度、复杂性、泛化性、以及其模型效果方面,与 28 个现有典型 IQA 模型进行对比。实验结果表明,所提 BCTCSP 模型的精度 PLCC 值在 6 个数据库中最低可达 0.892 1,最高达到 0.966 4,6 个数据库的加权 PLCC 达到 0.917 4,其综合效益高于该 28 个现有 IQA 模型。综合理论和实验结果表明,BCTCSP 是一种有效的、性能优越的无参考 IQA 模型。

关键词:图像质量评价;图像画面;视觉感知;纹理信息熵

中图分类号:TP391;TN919.8 **文献标识码:**A

doi:10.37188/OPE.20253318.2944 **CSTR:**32169.14.OPE.20253318.2944

No reference image quality assessment based on image tableau information and their visual perception

YAO Juncai^{1,2*}, SHEN Jing¹

(1. School of Computer Engineering, Nanjing Institute of Technology, Nanjing 211167, China;
2. School of information and Communications Engineering, Xi'an Jiaotong University,
Xi'an 710049, China)

* Corresponding author, E-mail: yaojc@njit.edu.cn

收稿日期:2025-04-03;修订日期:2025-06-18.

基金项目:国家自然科学基金(No. 61301237);江苏省自然科学基金面上项目(No. BK20201468);江苏省高校“青蓝工程”中青年学术带头人资助项目(苏教师函[2022]29号)和南京工程学院创新基金重大项目(No. CK-JA202304)

Abstract: It aimed to propose an image quality assessment (IQA) method that conformed to visual perception and had high comprehensive benefits of accuracy, generalization, and complexity, which would better control image processing and meet its practical application needs. Based on the image tableau information and their features such as brightness, chromaticity, texture, clarity, and local contrast, and considering human perception effects such as contrast sensitivity, non-linear perception of brightness, visual comfort, and texture complexity perception, a no-reference IQA method, namely BCTCSP, was proposed. In BCTCSP, firstly, by analyzing the relationship between image quality and image brightness, grayscale distribution, color depth and saturation, non-linear perception of brightness, and visual perception comfort, a quantitative and computational method was proposed to obtain the contribution and impact of image brightness, chromaticity, and visual perception on IQA. Then, combining the gray-gradient co-occurrence matrix to calculate and statistically analyze image texture features, and using texture weighted averaging and the HVS complex object perception model, a method was proposed to quantify and calculate the contribution and impact of image texture information entropy and its visual perception on IQA. Next, the contrast value and detection threshold of each point in the image were calculated, subsequently combining the contrast sensitivity characteristics and their models, and masking properties of HVS, the contribution and impact of local contrast and visual perception of the image on IQA were quantified and calculated. Afterwards, four factors including sharpness, signal-to-noise ratio, proportion of high-frequency components, and resolution, were used to describe the clarity of the image, and their quantification and calculation methods were proposed to obtain the clarity index of the image. Finally, synthesizing four factors, an IQA model was constructed, and its measurement standards were quantified. Meanwhile, 6 430 distorted images from 6 open databases (TID2013, CSIQ, LIVE, IVC, SPAQ, and Koniq-10k) were tested and verified, and in terms of accuracy, complexity, generalization, and their comprehensive benefits, BCTCSP was compared with 28 existing and typical IQA models. The experimental results show that the accuracy PLCC of the proposed model reaches a minimum of 0.892 1, a maximum of 0.966 4 among the 6 databases, and the weighted PLCC of 6 databases reaches 0.917 4. Its comprehensive benefits are higher than those of the 28 existing IQA models. The comprehensive results indicate that the proposed model is an effective and high-performance NR-IQA model.

Key words: image quality assessment; image tableau; visual perception; texture information entropy

1 引 言

随着人工智能、信息技术和光电技术的迅猛发展,图像技术已广泛应用于各行各业。在实际应用中,往往为了实现特定目标和预期效果,需要对图像进行各种各样的处理。所以,在图像的获取、显示、量化、压缩、存储、传输等过程中,由于成像系统、算法模型、处理方法、存储设备和传输条件等的不完善,图像极大可能遭受不同程度的失真和降质,直接影响了用户的感知效果、使用价值和服务质量,迫切需要通过获取图像质量的结果来调节和控制相关环节和过程^[1]。于是,如何评估图像质量成为了关键问题。一种有效而又实用

的图像质量评价(Image Quality Assessment, IQA)准则,不仅能自动评估图像质量,而且还能实时准确有效的跟踪其质量、优化其算法、指导其参数等^[2]。所以,至今对 IQA 进行了大量研究,提出了较多方法和指标,但仍不能满足当前复杂多变且日益飞速增长的应用需求。因此,其工作具有重要研究意义和实际应用价值^[1-3]。

至今,国内外在 IQA 方面做了大量工作,报道了较多的全参考(Full-Reference, FR) IQA 模型,代表性的有 MSE^[1], PSNR^[1], VSNR^[4], SSIM^[5], FSIMc^[6], GMSD^[7]和 VSI^[8]等。但在图像成像、显示和处理等过程中,无参考(No-Reference, NR) IQA 因不需要源参考图像,其应用更

具有实用价值,以致得到了更多关注。目前比较典型的有:BRISQUE^[9],DIIVINE^[10],COR-NIA^[11],BLIINDS-II^[12],MEON^[13]等。但综合考虑 NR-IQA 模型的精度、计算复杂性、泛化性能,以及构建模型的人工智能方法、网络模型、训练样本等,现有的 NR-IQA 模型真正应用于实际的非常少,究其原因,目前主要有以下困境。

(1) 图像质量描述以及其特征提取问题。对于 IQA,关键是图像的质量如何描述,往往采用一定数量的质量特征量来综合描述。对于 NR-IQA 模型,常常采用机器学习方法或数学建模方法设计合适的特征量模型来提取质量特征并量化。但目前的问题是:其特征量越多,越有代表性,则 IQA 模型的精度会越高;然而其泛化性能则越低,模型也越复杂。质量描述特征量与模型的精度、复杂性和泛化性能需要相互制约协调,即为了实现所提 IQA 模型具有较高的性能,需要选择和提取少量而又有效的图像质量特征量^[1,9,11];但如何实现,是目前的难点。

(2) 模型构建方法和性能问题。目前,在构建 NR-IQA 模型时,往往热衷于采用人工智能方法。但主要因为:构建的网络模型参数较多,其泛化性能较差;需要提取大量的图像特征;需要采用大量的样本进行训练,并且训练的样本需要多元化;设计的网络模型非常复杂等原因致使所构建的模型要么比较复杂,要么泛化性低,要么精度不高,其样本的数量、特征、设计网络的深度等等,对精度的影响均较大^[2,9,14],从而导致其真正能实际应用的非常少。如何设计合理的网络实现所提模型的精度、复杂性和泛化性能达到最大化效益,是目前研究的热点。而相对于采用数学建模方法来构建 NR-IQA 模型,其泛化性能和复杂性一般好于采用人工智能方法构建的 IQA 模型,但往往需要提取较少而有效的图像特征,才能保证所提模型的精度、泛化性能和复杂性的优势,其也是目前研究的困境^[2,9,14]。

(3) 模型泛化性能问题。在 IQA 的研究中,泛化性能一直是一个技术难点,即当所提 IQA 模型在交叉数据库中测试和应用时,其精度往往均不高,特别是当前报道较多的采用人工智能的方法构建 NR-IQA 模型,如 BLIINDS-II 模型^[12],其在 LIVE 数据库中的精度 Pearson 线性相关系数

和 Spearman 等级相关系数值为 0.930,0.931^[11],而在 TID2013 和 CSIQ 数据库中的精度其值只有 0.628,0.536^[12,15]和 0.534,0.574^[12,16]。并且,为了提升其泛化性能,往往采用更多、更广泛、更多元化的样本来训练,如此虽然提升了其泛化性能,但同时也增加了其模型的复杂性。目前因其泛化性能很大程度上决定着 IQA 模型的应用效果,所以得到了较多的关注。

(4) HVS 特性及其模型在 IQA 中的应用问题。当前,IQA 建模时,融入人眼视觉系统(Human Visual System, HVS)特性是一种非常有效提升模型精度的途径,特别是在采用人工智能方法构建 IQA 模型时,往往增加 HVS 特性模块,但该特性的融入却导致了 IQA 模型复杂性的增加^[14,17]。因此,在保障模型的复杂性不会有太大的增加而其精度又有较高提升的情况下,如何有效的融入是目前研究的关注点。

综上所述,在 IQA 建模中,选择合适的方法,提取少量而有效的图像特征,融入适当而有效的 HVS 特性及其模型,能够较好的提升 NR-IQA 模型的性能。为此,本文综合考虑 IQA 模型精度、复杂性和泛化性能,结合 HVS 的对比敏感、亮度感知非线性、纹理复杂度感知以及亮度感知舒适度等特性,基于图像亮度、色度、清晰度、纹理和局部对比度等特征,提出了一种 NR-IQA 方法;并采用 6 个开源数据库进行了测试,其结果表明,所提模型具有较好的性能。

2 IQA 贡献和影响因素及其特征提取

对于任意一幅图像,一般其画面呈现的信息主要有亮度、色度、纹理、清晰度和局部对比度等。结合图像画面信息特征和视觉感知效果,影响图像质量和对 IQA 有贡献的因素主要有:图像亮度和色度及其视觉感知(包括视觉感知舒适度)、图像纹理信息及其感知、局部对比及其视觉感知、图像清晰度,各因素及其特征提取、量化和计算方法具体说明如下。

2.1 图像亮度色度及其视觉感知对 IQA 的贡献和影响

人类感知图像时,其画面亮度和色度占主要

信息。结合 HVS 特性和图像特征分析,图像亮度色度及其感知对 IQA 的影响和贡献主要包括:(1) 图像亮度和色度的强度(intensity)大小对图像呈现质量的贡献,(2) 灰度分布(均衡)和颜色深度及饱和度,(3) 亮度和色度的强度的非线性感知,(4) 图像整体亮度(亮度和色度的辐射强度)视觉感知舒适度对 IQA 的贡献和影响。对于该 4 个因子(Factors),其描述、量化和提取计算方法如下。

2.1.1 图像亮度和色度的强度大小对 IQA 的影响和贡献

单从图像的整体亮度上分析,亮度对图像呈现的质量影响较大^[18]。一般当亮度低于 60 坎德拉/平方米(cd/m²)时,无论是图像的亮度分量还是色度分量,其整体呈现强度(即亮度和色度的强度之和,一般习惯称之为图像亮度)越高,往往图像的表现力越好,其质量呈现越好^[19-20];当亮度达到该值时,其对图像呈现效果的贡献达到最大;随着亮度的增加,亮度对图像呈现的贡献趋于稳定(此处不考虑过曝的情况)。同时考虑局部亮度和不同 R,G,B 比例和颜色表现效果的影响,图像亮度和色度的强度大小对图像质量的影响和贡献的计算方法如式(1):

$$F_{\text{Intensity}}(block_i) = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} \cdot \begin{bmatrix} 100 \times [1 - \exp(-0.015 \times \bar{I}_R(block_i)^2)] \\ 100 \times [1 - \exp(-0.015 \times \bar{I}_G(block_i)^2)] \\ 100 \times [1 - \exp(-0.015 \times \bar{I}_B(block_i)^2)] \end{bmatrix}_{\text{Norm}}, \quad (1)$$

其中: $\lambda_1, \lambda_2, \lambda_3$ 为 R,G,B 分量的比例系数,与颜色混合系数一致,取值为 0.299, 0.587, 0.114; $\bar{I}_R(block_i), \bar{I}_G(block_i)$ 和 $\bar{I}_B(block_i)$ 分别为三分量图中子块的平均亮度值。式(1)为计算图像中每一子块的结果,将其归一化,再按子块的数目计算整个图像的平均值,其平均值作为 IQA 影响和贡献的第 1 个质量特征量的分量。为了方便,记 $F_{\text{Intensity}}$ 为 F_1 。

2.1.2 图像灰度分布(均衡)和颜色深度及饱和度对 IQA 的影响和贡献

图像灰度分布和颜色深度及饱和度对图像的呈现效果作用较大,宽泛的图像灰度范围、颜

色深度及饱和度能更好的表现图像细节。颜色深度及饱和度的体现主要依赖其三刺激值,从图像显示的角度上,即依赖于图像 R,G,B 三分量图的灰度值范围,即灰度级在 0~255 之间的数目 N_R, N_G, N_B 。为此,图像灰度分布(均衡)和颜色深度及饱和度对 IQA 的影响和贡献($F_{\text{gray_distribution}}$)采用灰度均衡程度来描述,其计算如式(2):

$$F_{\text{gray_distribution}} = \lambda_1 \frac{N_R}{255} + \lambda_2 \frac{N_G}{255} + \lambda_3 \frac{N_B}{255}, \quad (2)$$

其中: $\lambda_1, \lambda_2, \lambda_3$ 同样为 R,G,B 三分量的比例系数,与式(1)中的参数取值一致。为了方便,记 $F_{\text{gray_distribution}}$ 为 F_{gd} ,其作为 IQA 贡献和影响的第 1 个质量特征量的另外一分量。

2.1.3 亮度和色度的强度的非线性感知对 IQA 的影响和贡献

人眼感知到的图像亮度与显示器的图像实际显示亮度并非线性关系,在低亮度环境中,人眼对亮度的变化更敏感,高亮度环境中则反之^[18-19]。则在 IQA 中,为了使客观 IQA 分数更符合人眼真实感知图像而得出的主观 IQA 结果,在进行 IQA 之前,采用人眼亮度感知非线性模型^[20](式(3))对 RGB 图像进行处理,得到亮度感知三分量图 R_p, G_p, B_p 。

$$\begin{aligned} R_p &= K_{11} \lg I_R + K_{12} \\ G_p &= K_{21} \lg I_G + K_{22}, \\ B_p &= K_{31} \lg I_B + K_{32} \end{aligned} \quad (3)$$

其中: $K_{11}, K_{12}, K_{21}, K_{22}, K_{31}, K_{32}$ 为常数, I_R, I_G, I_B 为 RGB 三分量图的灰度值。

本因子主要是考虑人眼亮度感知的非线性对 IQA 的影响和贡献,在所提 IQA 模型中涉及到亮度感知时,所对应的图像亮度均采用式(3)处理后的亮度,以模拟人眼对亮度的感知。在所提 IQA 方法中,为了考虑图像局部特征对 IQA 的影响,将图像分为 8 pixels×8 pixels 大小的子块进行处理,并按照式(3)计算每一个子块的亮度感知三分量图,分别记为 R_{pb}, G_{pb}, B_{pb} 。下面特征量计算中的图像即为亮度感知处理后的 RGB 图及其子块图像。

2.1.4 视觉感知舒适度对 IQA 的贡献和影响

图像的整体亮度对其画面高质量的呈现有较大的贡献,一般在亮度低于 60 cd/m²时,亮度越大,呈现效果也越好,图像质量越高;但人眼观察不同亮度的图像时,亮度对人眼的刺激造成的

人眼视觉感知舒适度对 IQA 结果也有较大影响。对于常见的显示器,一般人眼对图像的亮度在 $50 \sim 70 \text{ cd/m}^2$ 时,视觉感知舒适度最高(视觉感知舒适度与亮度的关系如图 1 中的红色虚线);亮度较低和较高时,视觉感知舒适度均不高。因而上所述的图像亮度和色度的强度的大小对 IQA 的贡献,即第一个质量特征量的第一个分量 F_1 ,需要考虑视觉感知舒适度的影响,即需要综合考虑图

$$F_{1p}(block_i) = \beta_1 F_1(block_i) + \beta_2 \text{Effect}_{\text{HVS_comfort_block}}(block_i)$$

$$\text{Effect}_{\text{HVS_comfort_block}} = \begin{cases} \exp[-\alpha(x - \delta)^2], & x \leq 60 \\ \sum_{i=0}^6 \omega_i x^i, & x > 60 \end{cases}$$

$$\begin{aligned} \alpha &= 0.0001, \delta = 60, x = \bar{I}(block_i), \omega_6 = -1.32 \times 10^{-10}, \\ \omega_5 &= 2.71 \times 10^{-8}, \omega_4 = -1.81 \times 10^{-6}, \omega_3 = 2.896 \times 10^{-5}, \\ \omega_2 &= 0.000843, \omega_1 = 0.00251, \omega_0 = 0.0057, \end{aligned} \quad (4)$$

其中:参数 $\alpha, \beta, \delta, \omega$ 为基于实验数据拟合和分析的经验值; $\bar{I}(block_i)$ 为任一子块的平均亮度。

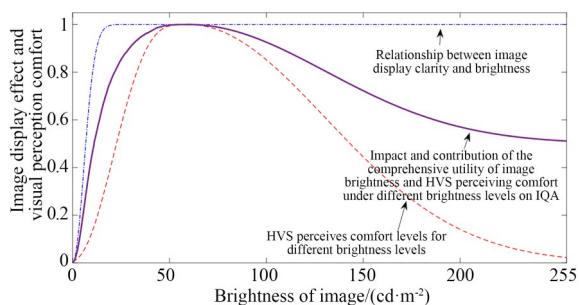


图 1 图像亮度及其人眼感知舒适度对 IQA 的贡献和影响
Fig. 1 TContribution and impact of image brightness and HVS perceived comfort on IQA

实际操作中,首先依据式(4)得到每一子块的结果,再按照子块的数目求平均,得到图像亮度对图像质量呈现及其感知效果(视觉感知舒适度)的综合结果对 IQA 的贡献,记为 F_{1p} ,其作为最终的对 IQA 贡献和影响的第 1 个质量特征量的第一个分量。

综合式(2)和式(4),则对于每一子块,图像亮度色度及其感知对 IQA 的贡献和影响 $F_{1cp}(block_i)$ 的计算如式(5),其结果作为 IQA 贡献和影响的第 1 个质量特征量。

$$F_{1cp}(block_i) = \sigma_1 F_{1p}(block_i) + \sigma_2 F_{gd}(block_i), \quad (5)$$

其中: σ_1 和 σ_2 是综合比例参数,通过分析数据的重要性和数据拟合得出其经验值。

像整体亮度对图像呈现质量的影响和对人眼感知效果的影响;则综合式(1)(图 1 中的蓝色点划线)和图 1 中人眼感知舒适度与亮度的关系(彩图见期刊电子版),可得“图像整体亮度对图像质量呈现及其人眼感知效果(视觉感知舒适度)”对 IQA 的综合贡献和影响可以按照图 1 中的紫色曲线变化,记对 IQA 的综合贡献和影响后的结果值为 F_{1p} ,则对于每一子块,其计算如式(4):

2.2 图像纹理信息熵及其感知对 IQA 的贡献和影响

图像纹理信息熵及其感知结果(Texture Information Entropy and its Perception, TIEP)主要用来描述图像画面内容信息、复杂程度及其视觉感知效果和结果。图像画面内容及其复杂程度采用纹理信息熵来描述,并采用视觉模型处理,其结果即为感知效果和结果。图像纹理一般采用图像的灰度和梯度的统计来描述,文中采用灰度梯度共生矩阵来计算^[14,19]。为了更好的表征图像纹理特征,所提 IQA 方法中,基于现有的纹理特征统计方法,提出了一种图像纹理信息熵的概念和计算方法,其定义为:采用图像纹理的加权平均信息量来描述图像的纹理信息的量的大小;其计算方法为:首先,对于任意一幅图像 $I(i, j)$,统计其灰度分布,计算其梯度,得到其灰度图、梯度图、以及其灰度梯度共生矩阵,分别记为: $gray(i, j)$, $grad(i, j)$, $H(gray(i, j), grad(i, j))$, 并对 H 实施归一化处理;然后,将所有各等级的梯度对应的乘以归一化共生矩阵中的值;最后,采用类似于 Huffman 编码中求加权平均码长的思想和基于信息熵的概念^[1,11],对所有相乘后的值求和,其求和的结果可以描述图像纹理的加权平均信息量(其类似于 Huffman 编码中的加权平均码长),记其为图像纹理信息熵($E_{\text{texture_information}}$)。其量化计算如式(6):

$$E_{\text{texture_information}} = \sum_{x=0, y=0}^{x=L-1, y=L-1} \|y\| \cdot \frac{H(x, y)}{m \times n}, \quad (6)$$

其中: m, n 为图像行和列数; L 为灰度等级(取 256); L_g 为梯度等级, 算法中采用 32 个级别; $H(x, y)$ 定义为灰度和梯度的集合, $x = \text{gray}(i, j)$, $y = \text{grad}(i, j)$, H 的定义如式(7)^[14]:

$$\{H(x, y) \mid x=0, 1, \dots, L-1; y=0, 1, \dots, L_g-1\}. \quad (7)$$

IQA 方法中, 结合 HVS 特征^[19-20]; HVS 对不同图像纹理复杂程度的感知效果不同, 感知纹理结果有差异。通过主观实验测试、数据分析拟合以及前人经验, 人眼感知图像复杂纹理的结果可按式(8)中的规律进行变化和量化计算:

$$F_{\text{Tiep}} = \begin{cases} a_1 \cdot E_{\text{ti}} + b_1, & E_{\text{ti}} < 0.15 \\ 0.95a_2, & 0.15 \leq E_{\text{ti}} \leq 0.25 \\ a_3 \exp(-b_3 \cdot E_{\text{ti}}) + c_3, & E_{\text{ti}} > 0.25 \end{cases}, \quad (8)$$

$$E_{\text{ti}} = E_{\text{texture_information}}$$

其中: a, b, c 是常数, 通过实验可获得。则其计算结果可描述图像纹理特征及其复杂性感知结果。

在所提 NR-IQA 方法实际操作过程中, 首先计算每一子块的 $E_{\text{texture_information}}(block_i)$ 和 $F_{\text{Tiep}}(block_i)$, 再求出图中的所有子块的 F_{Tiep} 的平均值; 则从 IQA 上, F_{Tiep} 平均值能够很好的描述图像画面中目标特征及其感知效果和结果这一方面对 IQA 的贡献, 其作为 IQA 贡献和影响的第 2 个质量特征量。

2.3 局部对比度及其视觉感知对 IQA 的贡献和影响

HVS 对图像的感知及其质量的评价, 对于每一点来说均受到其周围环境的影响^[18, 21]; 即, 对于图像上任意一点, 当其周围的任意一相邻点与之的对比差异小于人眼对比度敏感阈值 C_T 时, HVS 不能感知其差异, 其对图像质量的主观评价没有贡献; 而当相邻两点的对比差异大于 C_T 且越大时, 人眼分辨其效果越好, 表现出图像的轮廓线条越分明, 其图像质量越好。则在 IQA 中, 需要考虑图像中局部对比度及其视觉感知效果和结果, 其结果能够很好的说明图像局部质量的优劣程度; 所提算法中, 图像局部采用子块大小区域。则图像中局部对比度(C)及其视觉感知效果和结果的量化计算方法如下。

(1) 对于每一点, 计算其周围 8 pixels × 8 pixels 大小的子块区域中各像素点对其的对比影响, 即以该点为中心, 计算该点与其子块中其余 63 个点的亮度 I 的对比度 C , 记为 $C_{\text{block}}(i+s, j+t)$, 其计算如式(9):

$t)$, 其计算如式(9):

$$C_{\text{block}}(i+s, j+t) = \left| \frac{I(i, j) - I(i+s, j+t)}{I(i, j) + I(i+s, j+t)} \right|. \quad (9)$$

(2) 因为子块上每一点的空间位置和亮度色度不同, 其与该点的对比差异和感知效果就不同。则以该点为中心, 计算子块上每一点与该点的空间频率 f ; 再以观察距离为 50 cm, 子块大小为观察对象, 计算视角 θ , 得到子块上每点的角频率 f_θ ; 结合 f_θ , 利用 HVS 的对比敏感度函数(CSF)的亮度模型(Barten 等人^[19]提出), 计算子块上每点与该点之间的对比度敏感阈值 C_T , 记为 $C_{T_block}(i+s, j+t)$, 其计算如式(10):

$$C_T = k_1 \cdot f_\theta \cdot \exp(-k_2 \cdot f_\theta) [1 + k_3 \cdot \exp(k_2 \cdot f_\theta)]^{0.5}, \quad (10)$$

其中: f_θ 为像素点 $(i+s, j+t)$ 与中心点构成的周期性条纹的角频率; $\bar{L}$ 为目标的平均亮度; w 为空间视角^[16], $k_1 = 540 \cdot (1 + 0.7/\bar{L})^{-0.2} \cdot \{1 + 12[w \cdot (1 + f_\theta/3)^2]^{-1}\}$, $k_2 = 0.3 \cdot (1 + 100/L)^{0.15}$, $k_3 = 0.06$ 。

(3) 对于每一点, 对应计算其 C/C_T , 其值越大, 表明该两相邻点差异越大, 其显示效果越好, 越易分辨, 其主观感知质量越好; 最后将子块上所有点的 C/C_T 结果求平均, 以其平均值作为该子块大小区域中各点相互间的对比差异及其感知效果和结果的平均。

(4) 对于每个子块, 其上 C 的大小范围 ($C_{\text{block_max}} \sim C_{\text{block_min}}$) 对整个子块的感知效果有较大影响; C 的分布范围越宽, 其分辨效果越好; 为了更好的考虑图像局部(即子块)感知对 IQA 的影响, 加入 C 的分布范围的影响。

以上过程可用式(11)进行量化计算, 其既体现了局部对比度差异, 又结合了 HVS 的对比敏感度特性, 较好的描述了图像局部对比及其视觉感知方面对 IQA 的影响。

$$F_{\text{cp}}(block_i) = \xi_1 \frac{C_{\text{block_max}} - C_{\text{block_min}}}{C_{\text{max}}} + \xi_2 \left(\frac{1}{63} \cdot \sum \frac{C_{\text{block}}(i+s, j+t)}{C_{T_block}(i+s, j+t)} \right), \quad (11)$$

其中: ξ_1 和 ξ_2 为常数和经验值, C_{max} 为图像的 C 的最大值。

实际操作中, 首先依据式(9)和(10)得到每一子块的结果, 再按照子块的数目求平均, 得到此因素对其整幅图像的 IQA 的贡献, 其作为 IQA

贡献和影响的第 3 个质量特征量。

2.4 图像清晰度对 IQA 的贡献和影响

图像清晰度能够直接且较好的反映图像的质量,清晰度是图像细节边缘变化的敏锐程度,当观察图像时,人眼对图像的模糊程度非常敏感。因此,在 IQA 中,图像画面的清晰度起到非

$$F_{SSHR} = \chi_1 \text{Effect}_{\text{sharpness}} + \chi_2 \text{Effect}_{\text{SNR}} + \chi_3 \text{Effect}_{\text{HFCP}} + \chi_4 \text{Effect}_{\text{resolution}} =$$

$$\chi_1 \frac{\sum_{i=1}^{m \times n} \sum_{\text{area}=1}^8 \left| \frac{dI}{dx} \right|}{m \times n} + \chi_2 10 \lg \frac{I_{\text{signal}}}{I_{\text{noise}}} + \chi_3 \frac{f_{s, \text{high}}}{f_{s, \text{max}} - f_{s, \text{min}}} + K_r, \quad (12)$$

其中:锐度采用点锐度算法函数来计算^[22], $m \times n$ 为图像大小, dx 表示距离增量(即像素间隔), dI 表示灰度变化幅值;信噪比中采用图像的均值、方差、峰度和偏度等来判断图像中噪声的类型和大小;高频成分占比是基于高频成分越多图像越清晰、高频成分越少图像越模糊的原理,统计高频成分占比来度量, f_s 为图空间频率;分辨率往往为一定值,取为定值 K_r ;参数 χ_1, χ_2, χ_3 和 χ_4 为经验值。实际操作过程中,每一因子归一化后代入式(10)中计算,再就整个图像按子块的各个值进行归一化处理,并求平均。其计算结果 F_{SSHR} 作为 IQA 贡献和影响的第 4 个质量特征量。

3 IQA 方法和模型

3.1 NR-IQA 方法设计

分析 HVS 特性和图像特征,人眼在感知图像

常重要的作用。影响图像清晰程度的因素有很多,文中主要采用锐度(Sharpness)、信噪比(Signal to Noise Ratio, SNR)、高频成分占比(High-Frequency Components Proportion, HFCP)和分辨率(Resolution)4 个因子来描述,记为 F_{SSHR} ,其量化和计算如式(12):

时获取的信息主要是图像的画面信息;单从图像信息呈现上,画面信息主要包括图像的亮度色度、画面内容、局部对比度效果、清晰度 4 个方面的信息^[20,23]。为此,文中以此 4 个方面的信息作为图像画面信息。则从 IQA 的效果上, HVS 接受到的画面信息是图像画面信息及其相应信息的视觉感知效果的综合体^[24-25]。基于此种图像理解方式,对于任何一幅图像,可以将图像的亮度色度及其感知、图像画面纹理及其复杂度感知、清晰程度、局部对比度及其视觉感知, 4 个方面的结果作为图像信息及其视觉感知结果,即图像的视觉信息,并从这几个方面评价图像质量的优劣。基于此,所提 IQA 方法如下,其设计方案流程如图 2。

(1)首先从亮度对图像质量的贡献以及对视觉舒适度的影响两个方面对 IQA 的贡献和影响进行描述和量化,并考虑图像的灰度分布和颜色深

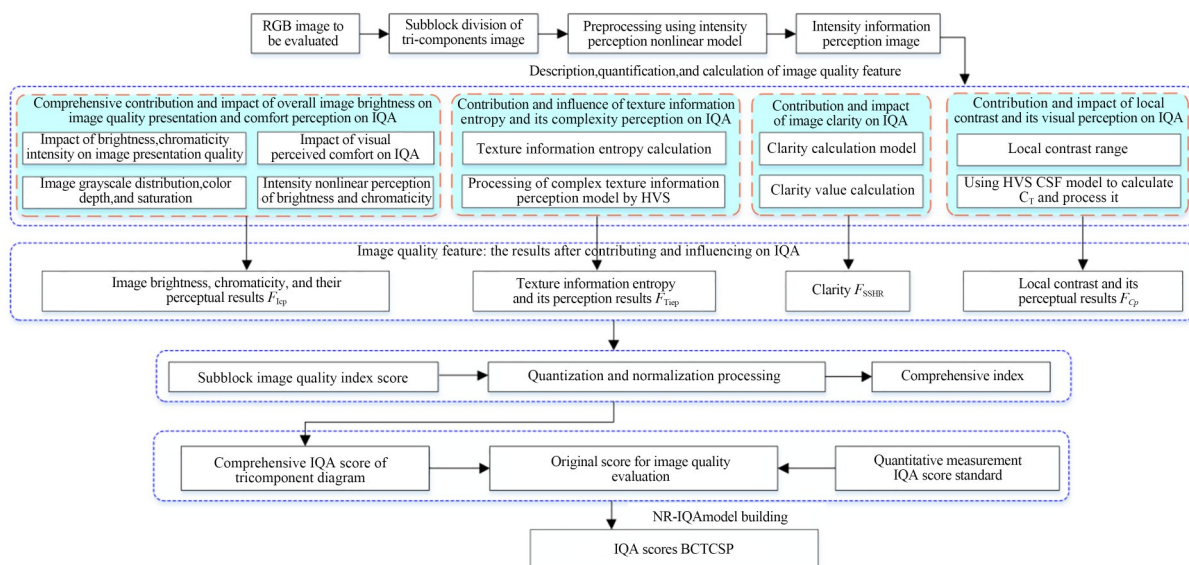


图 2 所提 NR-IQA 方法流程图

Fig. 2 Flow chart of the proposed NR-IQA method

度及饱和度对 IQA 的影响,且采用 HVS 的亮度强度非线性模型对图像的亮度和色度分量进行预处理,得到模拟人眼亮度色度感知图,并提取图像质量特征量 F_{Icp} ;

(2)再基于所提图像画面纹理信息熵的定义、描述和计算方法,结合 HVS 复杂纹理感知特性及其数学模型,量化和计算图像纹理信息熵及其感知结果对 IQA 的贡献和影响,提取其图像质量特征量 F_{Tiep} ;

(3)同时,计算图像每一点的对比度值和觉察阈值,结合 HVS 的对比敏感度特性及其模型和掩蔽特性,量化和计算图像局部对比度及其视觉感知结果对 IQA 的贡献和影响,提取其图像质量特征量 F_{Cp} ;

(4)并且分析清晰度与图像画面特征之间的联系,采用锐度、信噪比、高频成分占比和分辨率 4 个因子描述图像清晰度,结合所提其量化和计算方法,提取其图像质量特征量 F_{SSHR} ;

(5)最终,对上述 4 个图像质量特征量进行归一化处理,以其为质量指标,并对该 4 个指标进行综合,且量化和度量 IQA 评价标准,构建 IQA 模型。

3.2 NR-IQA 模型

结合 HVS 特性,依据上述所提方法和 4 个方面特征量的提取和计算方法,对感知的 RGB 图像的三分量图 R_p, G_p, B_p 的每一子块进行处理和计算,得到感知三分量图的每个子块的“图像亮度色度及其感知 F_{Icp} 、纹理信息熵及其感知 F_{Tiep} 、局部对比度及其感知 F_{Cp} 、清晰度 F_{SSHR} ”的图像质量特征量值,对其进行归一化处理,再按照图像的子块数目对归一化结果求平均,得到整个分量图的各个特征值,并将每个三分量图的 4 个质量特征量按照式(13)进行综合(综合后其最大值为 4),再按照 RGB 颜色混合方法中三分量的比例系数综合得到整个图像的质量描述值,并对其进行归一化处理,其结果即为图像的 IQA 分数。

$$\text{BCTCSP} = \sum \lambda_i \cdot [\mu_1 F_{\text{Icp}}(x) + \mu_2 F_{\text{Tiep}}(x) + \mu_3 F_{\text{Cp}}(x) + \mu_4 F_{\text{SSHR}}(x)]$$

$$(\lambda_i = \lambda_1, \lambda_2, \lambda_3; x = R, G, B), \quad (13)$$

其中: $\lambda_1, \lambda_2, \lambda_3$ 同样为 R, G, B 三分量的比例系数,与式(1)中的参数取值一致。 μ 为权重因子,基于实验数据分析和理论分析可得其经验值,其获得

方法为:进行一定的主观实验,专门研究图像亮度色度、纹理、对比度、清晰度等单独对其 IQA 的影响,分析其结果;同时采用一定量的数据进行拟合,并参考文献[1,7,9,11,25]中 IQA 模型的各个指标的重要性分析和参数值,综合这些因素,以数据拟合参数为主,以实验分析和文献中的参数为参照,对 μ 值进行经验性设定。其中,清晰度对 IQA 的影响敏感度最高,其 μ 值相对较大;纹理指标变化影响敏感度次之,对比度再次之;亮度色度在亮度低于 60 cd/cm^2 时,影响较大,当亮度高于该值后,影响较小。文中式(4),式(5),式(8),式(11),式(12)中均采用经验值确定其参数,其方法均按照本方法思路获取其参数值。

记上述方法为基于图像画面信息(亮度色度、纹理信息熵、局部对比度和清晰度)特征及其视觉感知的 NR-IQA 方法,其整个计算表达式记为 BCTCSP(NR IQA method based on image information (brightness and chromaticity, texture information entropy, local contrast and sharpness and their HVS perception)模型(为了方便,也可称之为 4F 模型)。其物理意义和 IQA 分数量化尺度说明为:采用该模型计算得出的 IQA 分数在 $[0, 1]$ 之间,其取值为 1 时,表示该图像的质量达到最佳;为 0 时,其质量最差。

4 实验仿真和测试结果

采用 6 个开源数据库,即 LIVE^[11], TID2013^[15], CSIQ^[16], IVC^[24], SPAQ^[25] 和 Koniq10k^[26] 中的图像对所提方法和模型进行实验测试。6 个数据库中分别选用 779, 3 000, 866, 185, 800 和 800 幅图像(共 6 430 幅)作为测试对象,测试得出 6 个数据库中每幅图像的 IQA 客观分数。为更好的说明所提模型的性能,基于开源数据库中的主观 IQA 分数 MOS 或 DMOS (Differential MOS)^[11,15-16,24-26],计算主客观 IQA 分数之间的一致性,其参数一般包括 PLCC (Pearson Linear Correlation Coefficient), SROCC (Spearman Rank Order Correlation Coefficient), RMSE (Root Mean Square Error) 和 OR (Outlier ratio)^[11];并得出主客观 IQA 结果间之散点图;其结

果如图 3, 图上的曲线均是采用 Logistic 函数(5 参数)拟合的结果。

从实验测试结果图 3 中可得, (1) 对于 6 个数据库, 每个库中的所有图像的整体评价效果较好, 精度较高, 其 PLCC 值最低达到了 0.892 1, 最高达到 0.966 4, 6 个数据库的加权 PLCC 精度达到 0.917 4; 则表明, 该模型不仅表现出了较高

的精度, 而且从泛化性能上看, 在 6 个数据库中的精度均较高, 表现为对不同数据库中图像具有较好的泛化性能。(2) 从散点图上点的离散程度上分析, 其均较低, 表明其主客观 IQA 结果一致性较好, 模型评价效果较好。综合 6 个数据库的整体评价效果表明, 所提 NR-IQA 模型不仅具有较好的精度, 而且其泛化性能亦较好。

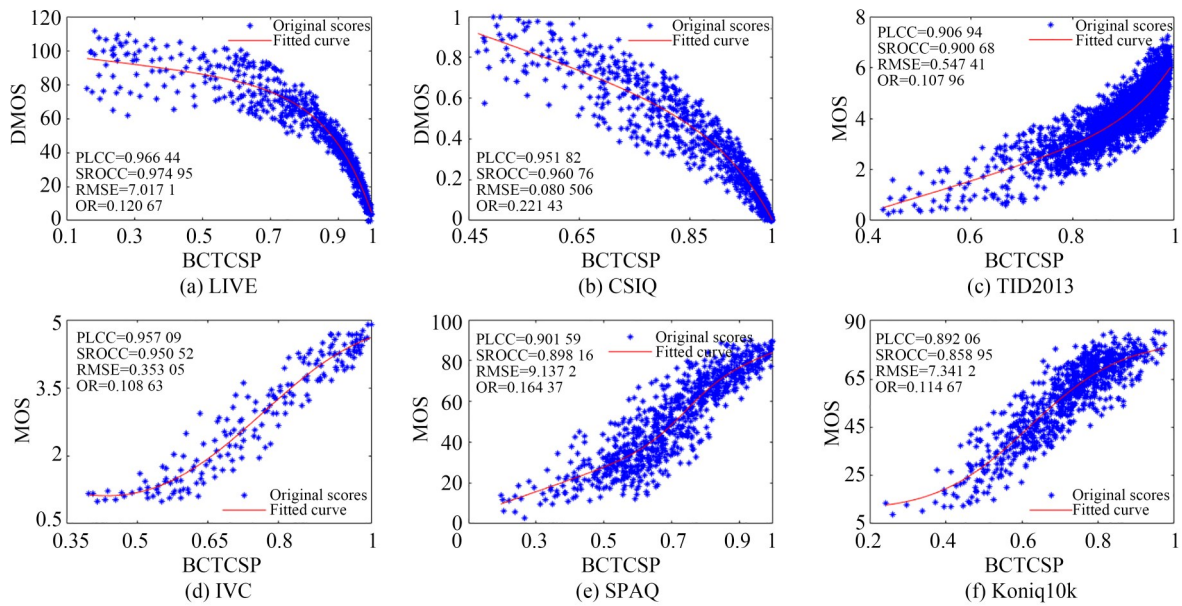


图 3 所提 NR-IQA 模型在 6 个数据库中的测试结果

Fig. 3 Tested results of the proposed NR-IQA model in 6 databases

5 对比分析和讨论

IQA 模型构建中, 需要重点分析和讨论 IQA 模型的 3 个重要性能, 即精度、泛化性和复杂性^[3,14]。为更好说明, 对所提模型进行以下分析探讨: (1) 与现有典型的或最近提出的 FR 和 NR 的 IQA 模型的实验结果进行对比分析, 探讨所提模型的精度; (2) 针对不同失真类型图像, 对模型失真类型的泛化性能进行对比分析, 研究其泛化性能; (3) 结合 IQA 模型算法运行的时间, 对比分析和探讨其模型复杂性。

5.1 精度对比分析和讨论

为更好说明所提模型的精度优劣, 基于 3 个数据库 (LIVE, CSIQ 和 TID2013) 的 IQA 结果, 结合其库中提供的 IQA 主观分数 MOS 或 DMOS^[11,15-16], 计算 4 个相关性参数, 就模型精度

PLCC 和 SROCC 值, 将所提 BCTCSP 模型与现有典型的或最近提出的 14 种 FR-IQA 模型和 14 种 NR-IQA 模型进行对比分析, 其结果如表 1。

结合图 3, 对比分析表 1 中各 IQA 模型的实验结果, 可得: (1) 相对于 14 种现有典型的 FR-IQA 模型, 所提模型的 IQA 效果在 3 个数据库中, 均表现出更高的精度, 其高出幅度在 3 个数据库中加权平均 PLCC 值, 相对于 14 种模型平均为 8.45%; (2) 对于 14 种现有的 NR-IQA 模型, 其精度在 LIVE 数据库中一般均较高, 其 PLCC 和 SROCC 值达到 0.9 以上, 而在 CSIQ 或 TID2013 中, 部分 IQA 模型精度并不高, 甚至较低, 不到 0.6; 所提 BCTCSP 模型在 3 个数据库中的精度均较高, 均达到 0.9 以上; 除 Mfca 模型外 (Mfca 与所提模型在 3 个数据库中互有优势^[35]), 均超过了其余 13 种现有模型的精度 PLCC 和 SROCC

表 1 基于 3 个数据库中 IQA 结果的 PLCC 和 SROCC 对比 BCTCSP 模型与 28 种现有模型的性能

Tab. 1 Performance comparison between BCTCSP and 28 existing models based on PLCC and SROCC of IQA results from 3 databases

Databases		LIVE(779)		TID2013(3000)		CSIQ(866)		Weighting	
IQA models		PLCC	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC
NR	BRISQUE[9]	0.942 4	0.939 5	0.900 0	0.883 0	0.937 0	0.909 0	0.914 0	0.897 3
	DIIVINE[10]	0.917 0	0.916 0	0.794 0	0.795 0	0.855 0	0.835 0	0.826 0	0.822 8
	CORNIA[11]	0.935 0	0.942 0	0.904 0	0.893 0	0.934 0	0.914 0	0.914 8	0.905 1
	BLIINDS-II[12]	0.930 2	0.930 6	0.628 0	0.536 0	0.534 0	0.574 0	0.661 2	0.609 3
	IL-NIQE[12]	0.906 0	0.902 0	0.648 0	0.521 0	0.865 0	0.822 0	0.731 7	0.641 0
	MEON[13]	0.955 0	0.951 0	0.912 0	0.912 0	0.944 0	0.932 0	0.925 2	0.922 3
	CONTRIQUE[21]	0.961 0	0.960 0	0.857 0	0.843 0	0.955 0	0.942 0	0.892 7	0.881 1
	CaHDC[27]	0.964 0	0.965 0	0.862 0	0.878 0	0.903 0	0.914 0	0.886 8	0.899 3
	HyperIQA[28]	0.966 0	0.962 0	0.858 0	0.840 0	0.942 0	0.923 0	0.891 8	0.875 9
	TRIQ[29]	0.965 0	0.949 0	0.858 0	0.846 0	0.838 0	0.825 0	0.872 2	0.859 4
	GraphIQA[31]	0.980 0	0.979 0	0.862 1	0.853 7	0.959 0	0.947 0	0.899 9	0.892 1
	LPIPS[33]	0.934 0	0.932 0	0.749 0	0.670 0	0.896 0	0.876 0	0.807 4	0.752 3
	MFca[35]	0.970 0	0.965 0	0.965 0	0.959 0	0.976 0	0.969 0	0.967 9	0.961 9
	AKD-IQA[38]	0.956 7	0.948 3	0.859 2	0.847 6	0.923 4	0.910 8	0.887 5	0.876 3
	BCTCSP[39]	0.966 4	0.975 0	0.906 9	0.900 7	0.951 8	0.960 8	0.925 2	0.924 4
FR	PSNR[1,4]	0.872 3	0.875 6	0.677 3	0.687 6	0.751 2	0.805 8	0.723 8	0.741 2
	PSNRHVS[4,15]	0.918 2	0.923 4	0.720 9	0.706 1	0.823 8	0.830 3	0.773 2	0.765 7
	VSNR[4]	0.923 1	0.927 4	0.740 6	0.719 7	0.812 9	0.826 8	0.784 7	0.774 5
	SSIM[5]	0.944 9	0.947 9	0.789 5	0.741 7	0.855 9	0.869 9	0.827 9	0.800 2
	MSSIM[5]	0.949 0	0.951 3	0.837 0	0.787 2	0.900 0	0.913 2	0.867 5	0.838 2
	FSIM[6]	0.961 3	0.964 5	0.867 9	0.851 0	0.903 2	0.915 5	0.890 1	0.882 1
	GMSD[7]	0.960 3	0.960 3	0.859 0	0.804 4	0.947 1	0.949 8	0.892 4	0.857 7
	VSI[8]	0.948 2	0.952 4	0.900 0	0.896 5	0.927 9	0.942 2	0.913 3	0.914 4
	GRIDS[14]	0.956 0	0.954 0	0.868 0	0.861 0	0.943 0	0.942 0	0.896 7	0.891 7
	GP-SSIM[32]	0.899 0	0.913 0	0.814 0	0.794 0	0.918 0	0.953 0	0.847 6	0.843 6
	DISTS[33]	0.954 0	0.954 2	0.855 0	0.830 0	0.928 4	0.928 9	0.885 3	0.869 3
	DeepSKLD[34]	0.972 0	0.965 9	0.901 2	0.878 3	0.965 2	0.968 3	0.925 0	0.909 8
	DMM[36]	0.915 0	0.918 0	0.868 0	0.863 0	0.732 0	0.711 0	0.850 5	0.843 9
	PEL[37]	0.966 5	0.965 7	0.902 0	0.879 6	0.967 2	0.962 7	0.925 0	0.909 5

值;且其超出幅度在 3 个数据库中加权平均 PLCC 值,相对于 13 种模型平均为 9.25%。综合 FR 和 NR 的 IQA 模型的对比分析,所提模型既表现出了较高的精度性能,又表现出了较好的泛化性能。

分析其原因,主要有:(1)对于 FR-IQA 而言,模型的精度和泛化性能主要由提取的特征和结合的 HVS 特性等方面决定;14 种现有的 FR-IQA 模型,往往针对性的结合了部分特征,且其特征对图像质量的贡献并不一定高,甚至没有结

合 HVS 特性及其模型,所以,其模型精度一般不高,其 PLCC 或 SROCC 值超过 0.9 的不多,适应不同数据库的泛化性能也较低;而所提模型结合的图像特征是亮度色度、局部对比度、清晰度和纹理信息,其与图像质量有更高的相关性,其相关度均在 0.8 以上,清晰度和纹理信息的相关度甚至超过 0.95^[25,29,35];另外,所提模型结合了 HVS 的多个特性及其模型,所以其精度能进一步的提高。(2)对于 NR-IQA 模型而言,主要采用机器学习或深度学习等人工智能的方法构建 IQA

模型,其模型的精度在报道中往往均较高,但限于不同数据库及其图像特征、训练样本数量和类型、网络模型和结构等,其泛化性能往往受到较大影响^[31-32],常常在其它不同数据库中测试时,其精度不一定达到较高的效果,所以其实际应用中并不多见,上述所提 14 种现有的 NR-IQA 模型,除 BRISQUE 方法外,其余均是如此,特别是 BLIINDS-II 和 IL-NIQE 模型,其影响更为明显。(3)BRISQUE 和所提 BCTCSP 模型主要采用数学建模的方法构建 IQA 模型,同时在其参数的拟合中也采用了大量的样本训练获得,其较好的保证了构建模型的泛化性能。另外,采用提取具体的图像特征构建 IQA 模型,具有较好的实际特征量的物理意义,而利用人工智能方法和提取特征

以构建 IQA 模型,较多的图像特征量物理意义不明确,在改变数据库测试和实际应用时,其精度会明显降低,表现出其泛化性能较弱。则在 IQA 模型的构建中,需要重点考虑选择有效的图像特征,同时结合 HVS 特性及其模型。

5.2 模型失真类型泛化性对比分析和讨论

采用所提模型对 3 个数据库(即 LIVE,CSIQ 和 TID2013)中所有失真类型的每一类失真图像单独进行仿真,预测得到每一类失真图像的客观 IQA 分数,并结合其开源数据库中提供的 MOS/DMOS 主观质量分数^[11,15-16],计算每一类失真图像主客观评价分数之间的一致性参数。基于相关性参数 PLCC 值,将其结果与 7 种常见 IQA 模型的实验结果进行对比,其结果如图 4~图 6;其

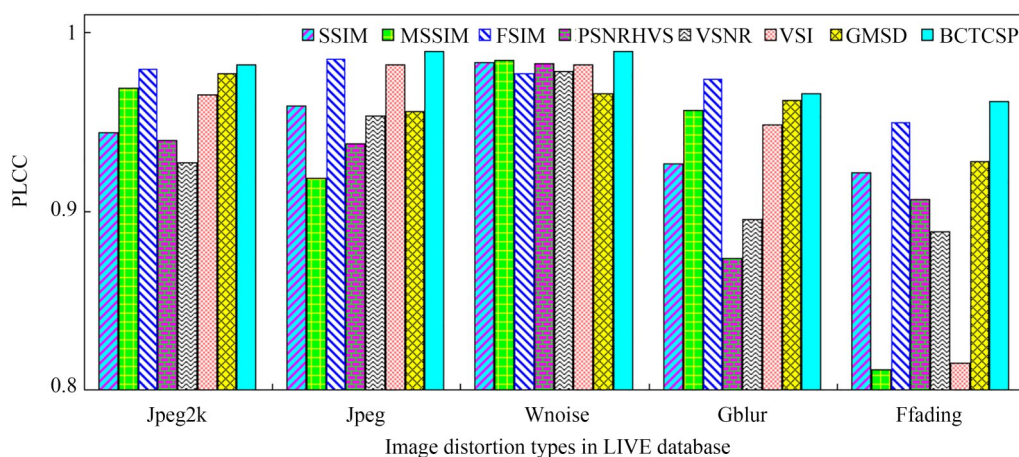


图 4 基于 LIVE 库中 5 种失真类型图像的 IQA 结果的 PLCC 值对比 BCTCSP 模型与 7 种常见模型

Fig. 4 Comparing BCTCSP with 7 common models based on PLCC of IQA results of 5 distortion types in LIVE

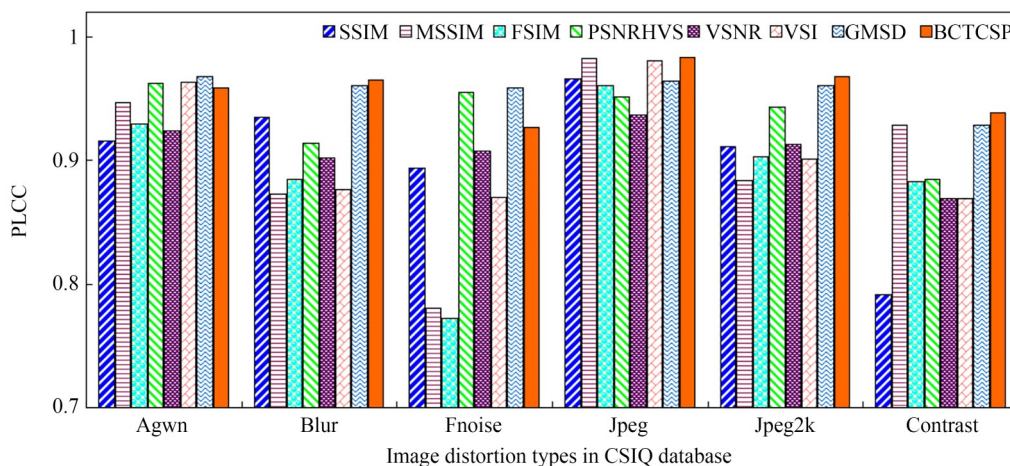


图 5 基于 CSIQ 库中 6 种失真类型图像的 IQA 结果的 PLCC 值对比 BCTCSP 模型与 7 种常见模型

Fig. 5 Comparing BCTCSP with 7 common models based on PLCC of IQA results of 6 distortion types in CSIQ

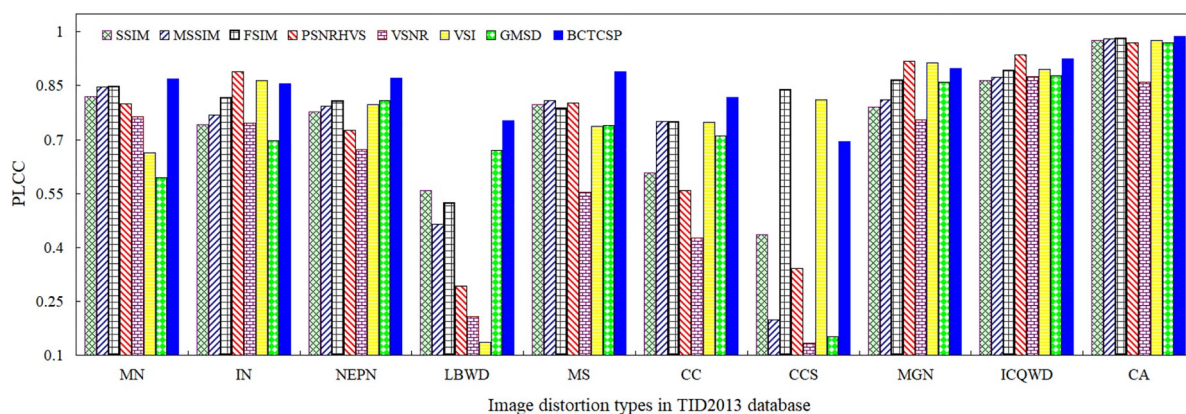


图 6 基于 TID2013 库中 10 种失真类型图像的 IQA 结果的 PLCC 值对比 BCTCSP 模型与 7 种常见模型

Fig. 6 Comparing BCTCSP with 7 common models based on PLCC of IQA results of 10 distortion types in TID2013

中,图 6 是选择 TID2013 数据库中效果有明显优

势的 10 种失真类型图像评价的结果。对比分析图 4~图 6 的结果,直观可得:(1)对于 3 个数据库中所有不同失真类型的图像,所提 BCTCSP 模型评价结果的 PLCC 值波动性明显小于 7 种现有模型的,表明 BCTCSP 模型具有更好的稳定性;(2)对于 3 个数据库中每种失真类型,BCTCSP 模型评价结果的 PLCC 值均能达到 7 种模型中的最高或次高,表明所提模型具有较高的精度和较好的泛化性能;(3)对于每种失真类型,BCTCSP 模型的 PLCC 值基本上均高于现有广泛使用的 SSIM 模型的结果。

5.3 模型复杂性对比分析和讨论

为了更好的说明所提 BCTCSP 模型的复杂性以及为其应用提供参考,对所提模型的复杂性进行分析,并与 20 种现有 IQA 模型(即 PSNR,

GMSD, SSIM, MSIM, VSI, FSIMc, VSNR,

PSNRHVS, DISTIS, DeepSKLD, PEL, BRISQUE, DIIVINE, BLINDS-II, CORNIA, IL-NIQE, HyperIQA, GraphIQA, LPIPS, AKD-IQA 模型)的复杂性进行对比。

在 IQA 研究中,其模型的复杂性往往采用模型算法的平均运算时间来定量描述^[3,14]。为此,对所提模型采用 Matlab 编程,分别对 3 个数据库中任选的 200 幅图像进行评价,以其评价过程中运算平均每 10 幅图像的时间来定量描述模型的复杂性,结合该模型在 3 个数据库中的加权精度 PLCC 值,将其结果(运算时间和 PLCC 值)与 20 种现有 IQA 模型的结果进行对比,得到精度和复杂性共同分析的结果,其如图 7。

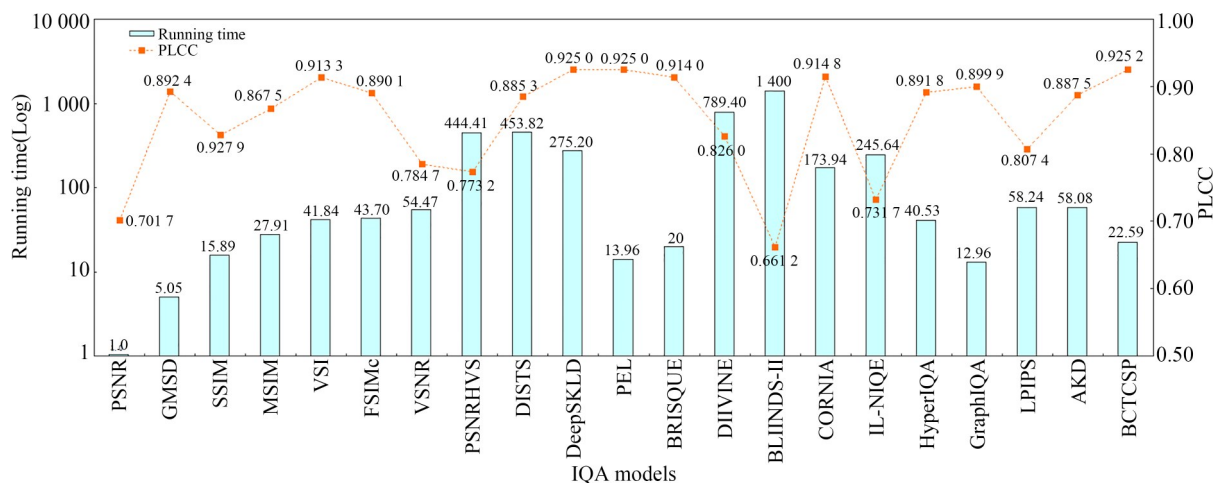


图 7 基于模型评价图像平均运算时间对比所提模型和 20 种现有模型的复杂性

Fig. 7 Comparing complexity between BCTCSP and 20 existing models based on the average running time of evaluating images

图 7 中,因为不同条件下计算的结果不同,采用相对效果来说明更为合适,即 BCTCSP 和 20 种现有模型的算法运算时间均取 PSNR 的相对值。对于大小不同的图像的模型运行时间,按照加权相同像素大小后的时间为其模型运行时间结果^[9,14]。本文实验中,其环境为:64 位操作系统的 Notebook PC, x64 的处理器,其为 Intel(R) Core (TM) i7-8550U CPU @ 1.8 GHz 1.99 GHz。

对比分析图 7 中 21 种模型的精度及其 IQA 时的算法运算耗时,可得:(1) BCTCSP 的复杂性低于其中 14 种模型,即 MSSIM, VSI, FSIMc, VSNR, PSNRHVS, DISTs, DeepSKLD, DIIVINE, BLINDS-II, CORNIA, IL-NIQE, HyperIQA, LPIPS 和 AKD-IQA,且其模型在 3 个数据库中的加权 IQA 精度 PLCC 和 SROCC 值均明显高于该 14 种模型。(2) 对于 PSNR, SSIM, GMSD, PEL, BRISQUE 和 GraphIQA 6 种 IQA 模型, BCTCSP 的复杂性明显高于 PSNR 和 GMSD,与 SSIM, PEL, BRISQUE 和 GraphIQA 的较接近,但在精度上,无论是在 3 个数据库中每一个库的结果,还是其加权结果,均明显好于 PSNR, SSIM 和 BRISQUE,稍高于 GMSD, PEL 和 GraphIQA;对于 GMSD 和 GraphIQA 来说,其在 3 个数据库中的精度均较高,但其泛化性稍差于 BCTCSP,在 TID2013 库中,其 PLCC 和 SROCC 分别只有 0.859, 0.804 和 0.862, 0.854,而对 BCTCSP 来说,其精度无论是在单个数据库中的评价,还是 3 个数据库的加权结果,均表现出了较好的稳定性,其精度 PLCC 和 SROCC 值均超过了 0.9。综合精度、复杂性和泛化性能,所提模型的综合效益好于该 20 种现有 IQA 模型。分析其原因,主要如下。

(1) IQA 模型构建方式不同影响其算法运算速度。

对于 FR-IQA 模型:① PSNRHVS, VSNR, FSIMc, VSI, MSSIM 结合了较多的图像特征和较复杂的 HVS 特性及其数学模型,特别是 HVS 模型,且需要进行图像时空变换以及色度空间变换,导致其计算量较大;② DISTs 和 DeepSKLD 均是通过深度网络学习方法构建的 IQA 模型,均

需要较多的图像特征。

对于 NR-IQA 模型:① DIIVINE, BLINDS-II, CORNIA, IL-NIQE, HyperIQA, LPIPS, AKD-IQA 模型主要是基于机器学习或深度学习构建的 IQA 方法,其设计的网络中参数较多,计算比较复杂;② 对于 BRISQUE 和所提模型,主要采用数学建模的方法,结合的图像特征较有限,所以其计算量相对较小;③ HyperIQA 和 PEL 虽也是基于深度学习方法构建的 IQA 方法,但 GraphIQA 通过单次前向传播完成预测,无循环和迭代过程,所有节点共享同一组学习权重,避免为每个节点定制化计算,所以大幅减少计算;PEL 主要针对光场图像特性,通过光场特有降维策略和选择性计算,采用浅层 CNN 特征而非端到端深度网络,从而计算量大大减少。

从所提模型的算法计算上看,其方法主要是提取图像的数据特征,主要针对的是空域上数值的计算,相对于复杂的神经网络模型和大量参数的回归模型来说,所提模型消耗的计算时间会少很多,所以,其复杂性较低。

(2) 构建 IQA 模型的质量指标影响算法运算速度。

① 对于采用深度学习或机器学习方法构建的 IQA 模型,如所对比的 DIIVINE, BLINDS-II, DISTs 和 DeepSKLD 等模型,其一般均需大量的图像特征指标,构建一定深度的网络,反复迭代和学习,同时融入一定的注意力机制,才能保证 IQA 的精度和泛化性能,但正因如此处理后,所构建 IQA 模型的参数较多,其计算量往往较大。② 所提模型主要结合了图像亮度色度、纹理信息熵、局部对比度和清晰度 4 大指标进行数学建模,其中,图像亮度色度、纹理信息熵和清晰度 3 个质量指标的计算主要针对的是图像空域像素数值上的统计和计算,其计算量相对较小,通过对该 3 个指标的单独实验,测试其耗时,其相对于 PSNR 计算耗时依次为 4.12, 2.14, 3.32;对于局部对比度及其视觉感知对 IQA 的影响指标,由于需要结合 HVS 特性的 Barten 模型计算,其计算量偏大,通过对该指标的单独实验,其耗时相对于 PSNR 为 20.26。该

测试实验中,其条件与本文前面一致,重复 10 次的平均值作为最终数据。从单独分析构建所提模型的关键质量指标的计算复杂度结果上看,其质量指标的计算量均不大,从而保证了所提模型的计算复杂性较低。

5.4 模型效果对比分析和讨论

5.4.1 基于 IQA 效果较差的部分失真类型图像的 IQA 结果对比分析 SSIM 与所提模型

综合精度、泛化性和复杂性,目前应用最为广泛的 IQA 模型是 SSIM;且在现有 IQA 模型中,其对 Local block-wise distortions of different intensity (LBWD), Contrast change (CC) 和 Change of color saturation (CCS)^[14-15]等失真类型的图像的 IQA 效果往往较差。为了说明所提模型的效果,单独对此 3 种类型的失真图像进行实验仿真和测试,比较 SSIM 与所提模型的精度。TID2013 和 CSIQ 数据库中,CC 失真图像共 241

幅^[14-15],LBWD 和 CCS 失真图像各 125 幅^[15];采用所提模型对其进行评价,并结合开源数据库中的主观 IQA 分数 MOS/DMOS^[15-16],计算 4 个相关性参数值。同时,采用 SSIM 对该 491 幅失真图像进行评价,其结果与所提模型的结果进行对比,其结果如图 8 和图 9。

对比分析图 8~图 9 的结果可得:(1)对于两个数据库中的 CC 失真图像,BCTCSP 的 IQA 精度高于 SSIM,其 PLCC 值平均高 15.53%;对于 TID2013 中的 CCS 和 LBWD 失真图像,BCTCSP 的评价精度比 SSIM 的分别提高了 55.58% 和 19.84%,表明所提模型具有较好的精度和泛化性能;(2)两个数据库中 3 组数据加权后,相对于 SSIM,所提模型的 PLCC 值提高了 23.65%。综合表明,所提模型具有较好的 IQA 效果,对 IQA 效果较差的部分失真类型的图像的 IQA 也有较好的效果。

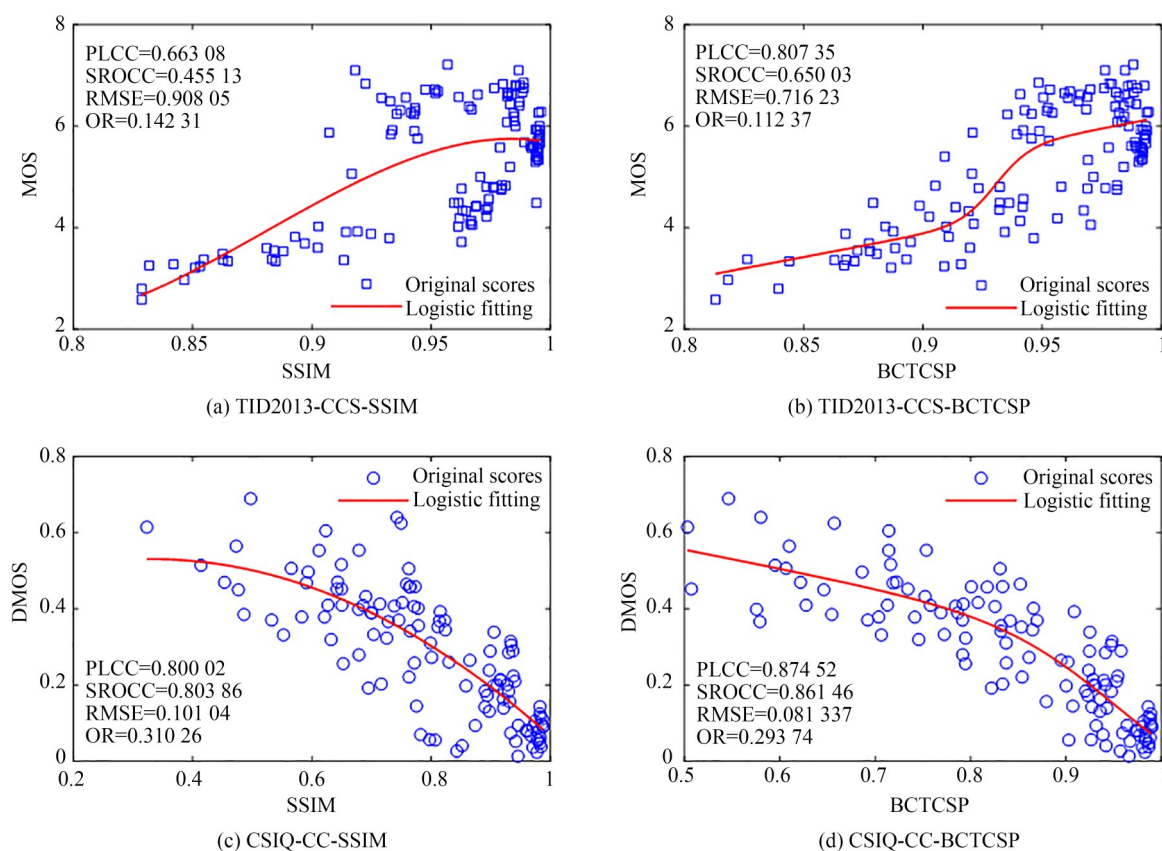


图 8 基于 TID2013 和 CSIQ 数据库中 CC 失真图像的 IQA 结果对比 SSIM 和 BCTCSP 模型的性能

Fig. 8 Performance comparison between SSIM and BCTCSP based on IQA results of CC distorted images in TID2013 and CSIQ

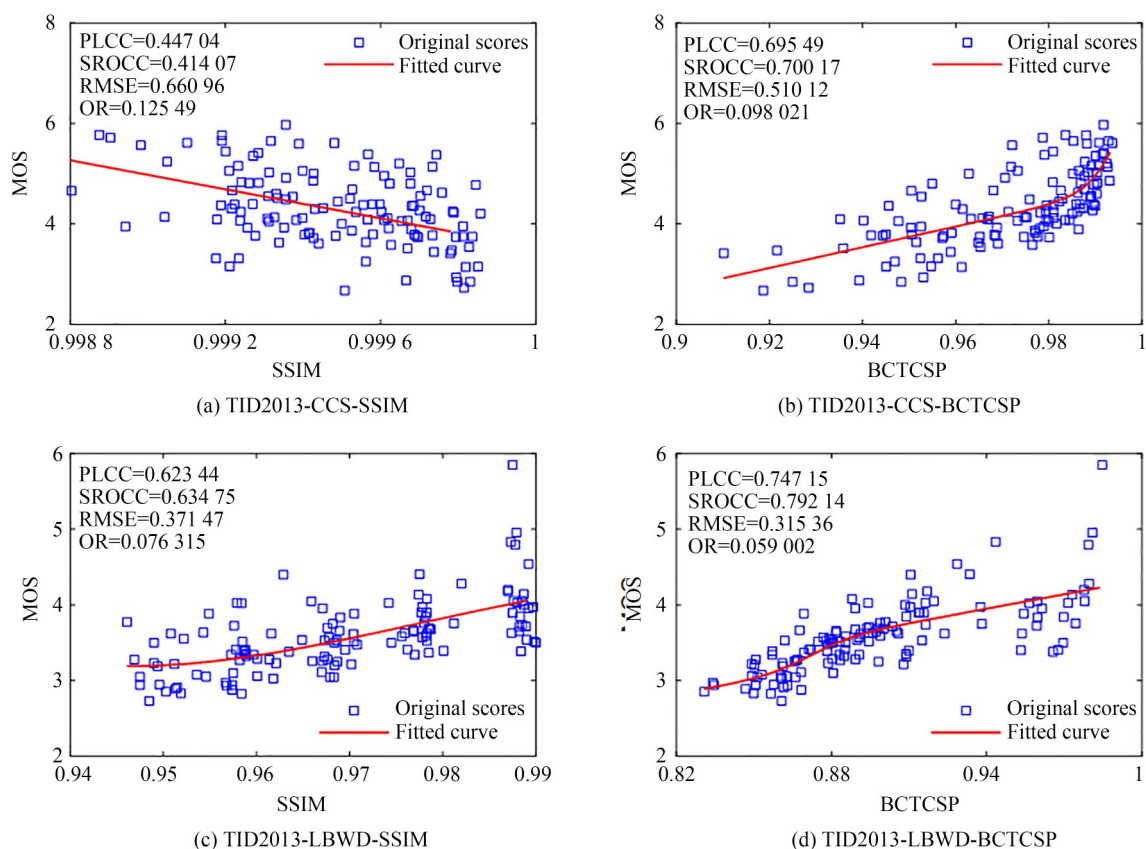


图9 基于TID2013数据库中CCS和LBWD失真图像的IQA结果对比SSIM和BCTCSP模型的性能

Fig. 9 Performance comparison between SSIM and BCTCSP based on IQA results of CC and LBWD distorted images in TID2013

5. 4. 2 与4个常见模型对比分析IQA效果

同时,为了更好地说明所提模型的优势,结合精度和泛化性能,针对两个数据库中的CC图像,采用另外4个常见模型分别对其进行仿真实验,得到其IQA分数,并计算4个相关性参数。基于该4个参数比较所提IQA模型与该4种IQA模型的性能,其对比结果如图10。

对比分析图10中的结果,可得:(1)针对TID2013中的CC失真图像,所提模型的评价精度均好于其4种现有IQA模型的精度;(2)在评价CSIQ中的CC失真图像时,所提模型的评价精度与GMSD和FSIMc的接近,但GMSD和FSIMc在评价TID2013中的CC失真图像时,其评价精度比较低,表明所提模型的泛化性能优于GMSD和FSIMc模型。基于精度和泛化性能的综合效益,表明所提模型性能好于该4种现有IQA模型。

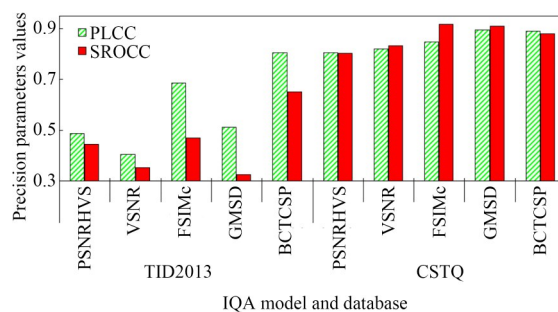


图10 基于TID2013和CSIQ库中CC失真图像的IQA结果的PLCC和SROCC值比较所提模型与4种常见模型的性能

Fig. 10 Performance comparison between 4 common models and BCTCSP based on the PLCC and SROCC of IQA results of CC distorted images in TID2013 and CSIQ

6 结 论

结合HVS的对比敏感、亮度感知非线性、视

觉感知舒适度和纹理复杂度感知等特性,基于图像亮度色度、清晰度、纹理和局部对比度特征,提出了一种 NR-IQA 方法和模型 BCTCSP。在 BCTCSP 中,首先采用 HVS 强度感知非线性模型对图像进行预处理,得到模拟人眼亮度色度感知图;再分别计算感知图像的亮度色度及其感知结果、纹理信息熵及其感知结果、局部对比度及其视觉感知结果、以及清晰度等 4 个质量指标量,以此 4 个指标为基础描述图像质量,并对其指标进行量化和归一化处理;最终综合 4 个指标,并量化度量指标,从而构建 BCTCSP 模型。并采用了 6 个国际开源数据库中的 6 430 幅失真图像对

其进行了仿真,测试和验证了其性能,且就精度、复杂性、泛化性、以及其综合效益方面,与现有的 28 个 IQA 模型进行了对比分析,其结果表明,在 6 个数据库中,BCTCSP 的精度 PLCC 值最低为 0.892 1,最高可达 0.966 4;6 个数据库的加权 PLCC 精度达到 0.917 4,其综合效益高于该 28 个模型。综合理论和实验结果表明,BCTCSP 是一种有效的、性能较好的 NR-IQA 模型。

作者贡献声明:

姚军财:方法和模型提出,论文构思和撰写;
申静:代码编写、实验设计和数据分析。

参考文献:

- [1] ZHAI G T, MIN X K. Perceptual image quality assessment: a survey[J]. *Science China Information Sciences*, 2020, 63(11): 211301.
- [2] ZHOU M L, LANG S J, ZHANG T P, *et al.* Attentional feature fusion for end-to-end blind image quality assessment [J]. *IEEE Transactions on Broadcasting*, 2023, 69(1): 144-152.
- [3] 李佳欣,段发阶,傅晓,等. 基于纹理奇异值分解的全参考图像质量评价[J]. *光学精密工程*, 2025, 33(1): 107-122.
LI J X, DUAN F J, FU X, *et al.* Full-reference image quality assessment based on texture singular value decomposition[J]. *Opt. Precision Eng.*, 2025, 33(1): 107-122. (in Chinese)
- [4] CHANDLER D M, HEMAMI S S. VSNR: a wavelet-based visual signal-to-noise ratio for natural images[J]. *IEEE Transactions on Image Processing*, 2007, 16(9): 2284-2298.
- [5] 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.
- [6] ZHANG L, 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.
- [7] XUE W F, ZHANG L, MOU X Q, *et al.* Gradient magnitude similarity deviation: a highly efficient perceptual image quality index[J]. *IEEE Transactions on Image Processing*, 2014, 23(2): 684-695.
- [8] ZHANG L, SHEN Y, LI H Y. VSI: a visual saliency-induced index for perceptual image quality assessment[J]. *IEEE Transactions on Image Processing*, 2014, 23(10): 4270-4281.
- [9] GAO Y X, MIN X K, CAO Y Q, *et al.* No-reference image quality assessment: obtain MOS from image quality score distribution[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2025, 35(2): 1840-1854.
- [10] LIU Y, YIN X H, WANG Y, *et al.* HVS-based perception-driven No-reference omnidirectional image quality assessment[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 72: 5003111.
- [11] SHEIKH H R, WANG Z, BOVIK A C. LIVE Image and Video Quality Assessment Database [EB/OL]. (2021-6-20). <http://live.ece.utexas.edu/research/quality>.
- [12] ZHANG L, ZHANG L, BOVIK A C. A feature-enriched completely blind image quality evaluator [J]. *IEEE Transactions on Image Processing*, 2015, 24(8): 2579-2591.
- [13] MA K D, LIU W T, ZHANG K, *et al.* End-to-end blind image quality assessment using deep neural networks[J]. *IEEE Transactions on Image Processing*, 2018, 27(3): 1202-1213.
- [14] SHEN W H, ZHOU M L, LUO J, *et al.* Graph-represented distribution similarity index for full-reference image quality assessment[J]. *IEEE Transactions on Image Processing*, 2024, 33: 3075-3089.
- [15] PONOMARENKO N, JIN L, IEREMEIEV O, *et al.* Tampere Image Database 2013 TID2013,

- version 1.0 [OL]. (2015-03-23). <http://www.ponomarenko.info/tid2013.htm>.
- [16] LARSON E C, CHANDLER D M. The CSIQ Image Database [EB/OL]. (2021-6-20). <http://vision.okstate.edu/?loc=csiq>.
- [17] YANG J C, BIAN Z L, LIU J C, *et al.* No-reference quality assessment for screen content images using visual edge model and AdaBoosting neural network[J]. *IEEE Transactions on Image Processing*, 2021, 30: 6801-6814.
- [18] YAO J C. Measurements of human vision contrast sensitivity to opposite colors using a cathode ray tube display [J]. *Chinese Science Bulletin*, 2011, 56(23): 2425-2432.
- [19] BARTEN P G J. Evaluation of subjective image quality with the square-root integral method [J]. *Journal of the Optical Society of America A*, 1990, 7(10): 2024-2031.
- [20] NADENAU M. *Integration of Human Color Vision Models into High Quality Image Compression* [D]. PhD thesis, École Polytechnique Fédérale de Lausanne, Switzerland, 2000.
- [21] MADHUSUDANA P C, BIRKBECK N, WANG Y L, *et al.* Image quality assessment using contrastive learning[J]. *IEEE Transactions on Image Processing*, 2022, 31: 4149-4161.
- [22] 曾海飞, 韩昌佩, 李凯, 等. 改进的梯度阈值图像清晰度评价算法[J]. *激光与光电子学进展*, 2021, 58(22): 285-293.
- ZENG H F, HAN C P, LI K, *et al.* Improved gradient threshold image sharpness evaluation algorithm [J]. *Laser & Optoelectronics Progress*, 2021, 58(22): 285-293. (in Chinese)
- [23] 江本赤, 卞仕磊, 史晨阳, 等. 基于色貌尺度相位一致性的全参考图像质量评价[J]. *光学精密工程*, 2023, 31(10): 1509-1521.
- JIANG B C, BIAN S L, SHI C Y, *et al.* Full reference image quality assessment based on color appearance-based phase consistency [J]. *Opt. Precision Eng.*, 2023, 31(10): 1509-1521. (in Chinese)
- [24] CALLET P L, Autrusseau F. Subjective Quality Assessment IRCCyN/IVC Database [EB/OL]. (2021-6-20). <http://www2.irccyn.ec-nantes.fr/ivcdb/>.
- [25] FANG Y M, ZHU H W, ZENG Y, *et al.* Perceptual quality assessment of smartphone photography [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020. Seattle, WA, USA. IEEE, 2020.
- [26] HOSU V, LIN H H, SZIRANYI T, *et al.* KonIQ-10k: an ecologically valid database for deep learning of blind image quality assessment [J]. *IEEE Transactions on Image Processing*, 2020, 29: 4041-4056.
- [27] WU J J, MA J P, LIANG F H, *et al.* End-to-end blind image quality prediction with cascaded deep neural network [J]. *IEEE Transactions on Image Processing*, 2020, 29: 7414-7426.
- [28] SU S L, YAN Q S, ZHU Y, *et al.* Blindly assess image quality in the wild guided by a self-adaptive hyper network [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020. Seattle, WA, USA. IEEE, 2020.
- [29] YOU J Y, KORHONEN J. Transformer for image quality assessment [C]. 2021 IEEE International Conference on Image Processing (ICIP). September 19-22, 2021, Anchorage, AK, USA. IEEE, 2021: 1389-1393.
- [30] YANG Y, DING Y Q, CHENG M, *et al.* No-reference quality assessment for contrast-distorted images based on gray and color-gray-difference space [J]. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2023, 19(2): 1-20.
- [31] SUN S M, YU T, XU J H, *et al.* GraphIQA: learning distortion graph representations for blind image quality assessment [J]. *IEEE Transactions on Multimedia*, 2022, 25: 2912-2925.
- [32] BAKUROV I, BUZZELLI M, SCHETTINI R, *et al.* Full-reference image quality expression via genetic programming [J]. *IEEE Transactions on Image Processing*, 2023, 32: 1458-1473.
- [33] DING K Y, ZHONG R J, WANG Z H, *et al.* Adaptive structure and texture similarity metric for image quality assessment and optimization [J]. *IEEE Transactions on Multimedia*, 2023, 26: 5398-5409.
- [34] LIAO X R, WEI X K, ZHOU M L, *et al.* Image quality assessment: measuring perceptual degradation via distribution measures in deep feature spaces [J]. *IEEE Transactions on Image Processing*,

- 2024, 33: 4044-4059.
- [35] WANG Z S, YUAN L, ZHAI G T. Channel attention for No-reference image quality assessment in DCT domain[J]. *IEEE Signal Processing Letters*, 2024, 31: 1274-1278.
- [36] CHEN B L, ZHU H W, ZHU L Y, *et al.* Debiased mapping for full-reference image quality assessment[J]. *IEEE Transactions on Multimedia*, 2025, 27: 2638-2649.
- [37] BEZERRA S A C, BEZERRA JÚNIOR S A C, DE S PIO J L, *et al.* Perceptual error logarithm: an efficient and effective analytical method for full-reference image quality assessment[J]. *IEEE Access*, 2025, 13: 68587-68606.
- [38] HU B, CHEN W Z, ZHENG J, *et al.* No-reference image quality assessment *via* inter-level adaptive knowledge distillation[J]. *IEEE Transactions on Broadcasting*, 2025, 71(2): 581-592.
- [39] MA P C, LIU L X, XIAO C Z, *et al.* Omnidirectional image quality assessment with mutual distillation [J]. *IEEE Transactions on Broadcasting*, 2025, 71(1): 264-276.

作者简介:



姚军财(1979—),男,博士,教授。毕业于西安交通大学,获得博士学位,主要从事图像和视频技术,光信息处理技术,机器学习,计算机视觉与模式识别等方面的研究工作。E-mail: yaojc-nj@njit.edu.cn

通讯作者:



申 静(1981—),女,硕士,副教授。主要从事图像技术,信息安全,光信息处理技术,计算机视觉与模式识别等方面的研究工作。E-mail: shenjnj@njit.edu.cn