

文章编号 1004-924X(2026)07-1170-19

双尺度画面细节信息引导的全参考图像质量 客观评价

史晨阳^{1,2,3*}, 吴俊杰¹, 袁瀚成¹, 吴路路¹

- 安徽工程大学 人工智能学院, 安徽 芜湖 241000;
- 安徽工程大学 智能汽车线控底盘系统安徽省重点实验室, 安徽 芜湖 241000;
- 安徽工程大学 安徽省先进检测与智能感知重点实验室, 安徽 芜湖 241000

摘要:在工业视觉系统中,全参考客观评价模型是图像质量评价领域的研究热点。本文提出一种双尺度画面细节信息引导的全参考图像质量评价模型。将测试图像转换至特定颜色空间,分离亮度与色度通道。基于参考图像、失真图像及其融合图像之间的梯度交互,构建亮度通道梯度相似度分量,并结合色度通道方向一致性分析生成颜色相似度分量。通过融合亮度通道的频谱残差、边缘信息与自适应对比度特征,构建双尺度画面细节累积分量,以描述细节信息的累积效应。最后,通过特征系数对上述各分量的标准差特征进行加权融合,得到最终评价结果。为验证本模型的可靠性,依据 Spearman 秩序相关系数(SROCC)、Pearson 线性相关系数(PLCC)、Kendall 秩序相关系数(KROCC)和均方根误差(Root Mean Square Error, RMSE)4项评价标准,在 LIVE, CSIQ, TID2008, TID2013 及 KADID-10K 数据库上进行测试。实验结果表明,本模型在上述数据库中测得的 PLCC 值最低为 0.876 8(KADID-10K),最高达到 0.967 8(LIVE),SROCC 值最低为 0.864 8(TID2013),最高达到 0.961 0(CSIQ)。与多种典型及深度学习全参考图像质量评价模型相比,本模型在计算效率方面优势明显,同时在预测精度与泛化性方面展现出良好的综合性能。

关键词:图像质量评价;全参考;双尺度;画面细节信息

中图分类号:TP391.41 文献标识码:A

doi:10.37188/OPE.20263407.1170 CSTR:32169.14.OPE.20263407.1170

Dual-scale tableau detail information-guided full-reference objective assessment on image quality

SHI Chenyang^{1,2,3*}, WU Junjie¹, YUAN Hancheng¹, WU Lulu¹

- School of Artificial Intelligence, Anhui Polytechnic University, Wuhu 241000, China;
- Anhui Province Key Laboratory of Intelligent Car Wire-Controlled Chassis System, Anhui Polytechnic University, Wuhu 241000, China;
- Anhui Provincial Key Laboratory of Advanced Detection and Intelligent Perception, Anhui Polytechnic University, Wuhu 241000, China)

* Corresponding author, E-mail: shichenyang@ahpu.edu.cn

收稿日期:2026-01-09;修订日期:2026-02-03.

基金项目:国家自然科学基金资助项目(No. 52005003);安徽省教育厅重点项目(No. 2022AH050983);安徽省车载显示集成系统工程研究中心开放基金项目(No. VDIS2023C01);智能汽车线控底盘系统安徽省重点实验室开放课题基金项目(No. QCKJJ202508);安徽工程大学引进人才科研启动基金项目(No. 2021YQQ027);安徽省先进检测与智能感知重点实验室开放课题基金项目(No. JCKJ2025B09)

Abstract: In industrial vision systems, full-reference image quality assessment (FR-IQA) has emerged as a major research focus. To address this problem, a novel FR-IQA model, termed a dual-scale tableau detail-guided model, is proposed. First, the test image is transformed into a specific color space to decouple luminance and chrominance channels. Subsequently, a luminance gradient similarity component is constructed based on gradient interactions among the reference image, the distorted image, and their fused representation. This component is further integrated with a chrominance orientation consistency analysis to derive a color similarity measure. Meanwhile, a dual-scale tableau detail accumulation component is developed by fusing spectral residual, edge, and adaptive contrast features of the luminance channel, enabling effective characterization of cumulative detail information. Finally, the standard deviation features of these components are weighted and aggregated through feature coefficients to produce the overall quality score. The reliability of the proposed model is validated on the LIVE, CSIQ, TID2008, TID2013, and KADID-10K databases using four evaluation metrics: Spearman rank-order correlation coefficient (SROCC), Pearson linear correlation coefficient (PLCC), Kendall rank-order correlation coefficient (KROCC), and root mean square error (RMSE). Experimental results indicate that the PLCC ranges from 0.876 8 on KADID-10K to 0.967 8 on LIVE, while the SROCC varies from 0.864 8 on TID2013 to 0.961 0 on CSIQ. The proposed model demonstrates superior computational efficiency compared with state-of-the-art and deep learning-based FR-IQA methods, while maintaining robust and generalizable predictive performance.

Key words: image quality assessment; full reference; dual-scale; tableau detail information

1 引言

随着视觉传感器与通信技术的迅猛发展,图像作为关键信息载体,其质量直接影响后续处理的准确性与可靠性。然而,在图像采集、传输及处理过程中,视觉系统可能由于各种因素的影响,不可避免地在图像中会引入白噪声、模糊及压缩伪影等多种失真现象,导致其质量显著下降^[1]。特别是在工业视觉、水下成像^[2]、医学影像^[3]等专业场景中,图像质量的退化会直接影响后续分析与决策的准确性。为实现对图像失真程度的快速、精准评估,图像质量评价(Image Quality Assessment, IQA)研究应运而生^[4],其核心目标是构建与人类视觉系统(Human Visual System, HVS)感知高度一致的评价体系,为工业视觉等实际场景提供高效可靠的质量判断依据。

IQA模型主要分为主观评价与客观评价两大类。主观评价以人类观察者的主观感受为核心,其结果最契合HVS特性^[5]。其中,双刺激损伤尺度(Double-Stimulus Impairment Scale, DSIS)和双刺激连续质量尺度(Double-Stimulus Continuous Quality-Scale, DSCQS)是两类经典

的主观评价范式^[6]。这两种方法的核心共性在于:通过强化“无损伤-有损伤”的直接关联,凸显二者的感知差异,为质量评价提供精准依据。这一核心逻辑也为客观IQA模型的优化提供了关键认知启发,即质量评估应建立在参考与失真图像的动态关联及差异累积基础上。然而,主观评价方法存在处理周期长、获取成本高、易受个体差异影响等局限性,难以满足工业视觉等实际场景的高效需求。相比之下,客观评价方法通过构建与主观评价结果高度一致的算法模型,实现图像质量的快速准确评估,已成为IQA领域的研究热点^[7-8]。根据对参考图像的依赖程度,客观IQA方法可进一步细分为全参考(Full-Reference, FR)、部分参考(Reduce-Reference, RR)及无参考(No-Reference, NR)三类^[9-11]。其中,FR-IQA模型无需前期耗时及复杂的训练过程,直接利用参考图像的完整信息进行评价,在工业生产领域发挥着不可或缺的作用。

MSE (Mean Square Error) 和 PSNR (Peak Signal to Noise Ratio) 模型是最基础的FR-IQA模型,其物理意义简单、计算过程简便。然而,由于其未能充分考量图像结构特征及HVS感知特

性,在相关数据库的验证测试中与主观评价结果的一致性较差^[12]。为提升评价相关性,研究人员开始模拟 HVS 感知机制进行建模^[13]。Wang 等提出的结构相似性 (Structural Similarity, SSIM) 模型从亮度、对比度和结构三个维度对图像质量进行统计分析^[14],显著提升了客观 IQA 模型与 HVS 感知的契合度。Sheikh 等基于信息保真度的准则提出 IFC (Information Fidelity Criterion)^[15]与视觉信息保真度 VIF (Visual Information Fidelity)^[16]。上述模型虽然进一步提高了预测精度,但仍未达到产业应用需求,且计算复杂度较高。近年来, Yang 等提出 RVSIM (Riesz Transform and Visual Contrast Sensitivity-Based Feature Similarity)^[17]模型,该模型基于 Riesz 变换和视觉对比敏感度的特征相似性指数,在解决特定失真时,具有较好结果,但泛化性能偏差。Man-souri^[18]等提出 SSVD (Structural Singular Value Decomposition) 模型,通过分析失真图像在原始图像奇异向量矩阵上的反射来评估结构信息,但在处理对比度变化等特定失真时仍有提升空间。Li 等^[19]提出 TSVD (Texture-Based Singular Value Decomposition),该模型在 SSVD 基础上引入梯度特征增强纹理提取效果,但在泛化性和细节失真量化方面仍存在局限。

除上述 IQA 模型的研究路径外,众多研究人员通过模拟 HVS 的画面细节 (Tableau Detail, TD) 感知,设计了一系列代表性 FR-IQA 模型,但仍存在明显局限。利用图像边缘等结构信息表征质量的核心特征^[20-22],如 FSIM (Feature Similarity)^[20]和 CAGS (Color Appearance and Gradient Similarity)^[22]等,上述模型多采用静态梯度算子提取边缘特征,缺乏参考图像与失真图像间的动态关联,导致对边缘与结构失真的捕捉不够精准。利用 HVS 接收的彩色视觉信息相关机制,模型强化对色度信息的考量,如 CPCs (Color Appearance-Based Phase Consistency and Contrast Similarity)^[23]和 VCGS (Visual Saliency with Color Appearance and Gradient Similarity)^[24]等模型采用不同的颜色空间获取色度信息,但在颜色失真的精细量化方面仍有提升空间,未能充分体现 HVS 颜色感知与图像失真的协同影响。TD 是反映图像感知质量的关键维度,现有 IQA 模型

对 TD 的表征多依赖单一特征 (如对比度^[25]、边缘密度^[26]、频谱残差^[27]等),未能考虑多种特征复合处理所带来的影响。另外,客观 IQA 模型相关特征的提取过程仅考虑参考图像与失真图像间差异计算所得结果,未考虑模拟人类观察者的主观感受,导致细微细节损伤难以被精准捕捉,限制了模型对复杂失真场景的适配性,如 CVSS (Contrast and Visual Saliency Similarity)^[25]。

综上所述,现有模型在边缘模糊、结构粘连等失真中,因使用静态梯度算子而缺乏动态关联;在颜色偏移、饱和度下降等失真中,对色度信息的精细量化与协同影响建模不足;在纹理丢失、对比度退化等复杂失真中,因依赖单一特征而难以捕捉画面细节的累积损伤。为此,本文借鉴 DSIS/DSCQS 的双刺激对比思想,通过参考图像、失真图像及其融合图像间的梯度交互,增强对结构动态差异的表征。在 LHM 颜色空间中采用通道乘法融合,提升对颜色协同失真的量化能力。最后,融合频谱残差、方向梯度与自适应对比度构建双尺度画面细节累积分量,以模拟细节损失的累积感知。所提模型将上述三类特征有机协同,构建一个可完成三维感知的 FR-IQA 模型,从而实现对结构、颜色及画面细节失真的全面量化。实验结果表明,该模型在多个公开数据库上均表现出与主观评价高度一致的良好性能,与其他主流模型相比,在预测精度、泛化性与计算复杂度上方面展现出显著优势。

2 原理

本节提出双尺度画面细节 (Dual-scale Tableau Detail, DTD) 信息引导的客观 FR-IQA 模型,该模型通过 TD 信息累积效应、梯度相似性和色度相似性的三维感知体系量化图像失真。其中,画面细节信息累积效应经双尺度融合频谱残差、梯度方向权重和局部对比度特征等画面细节处理得到。再通过参考图像与失真图像的比例融合,构建二者的融合图像,并通过融合相似计算得到梯度融合度特征。基于 LHM 颜色空间,通过色度相似性图度量捕捉颜色偏移。最后,利用标准差归一化与动态权重进行融合处理上述得到的结果,实现对图像质量的综合评估。

2.1 画面细节信息累积处理

TD是构成图像感知质量的基础要素,其丰富度、清晰度与保真度直接影响人眼对图像质量的整体判断^[28]。为全面表征TD信息,介绍一种TD信息累积处理方法,旨在模拟HVS处理参考图像与失真图像时,对整体TD的感知、比较与差异累积过程。本方法的设计借鉴了人类视觉系统的多尺度感知与细节累积机制,通过融合频谱残差、方向梯度与自适应对比度三类互补特征,以全面刻画细节的结构、显著性及层次信息。其整体流程如图1所示。

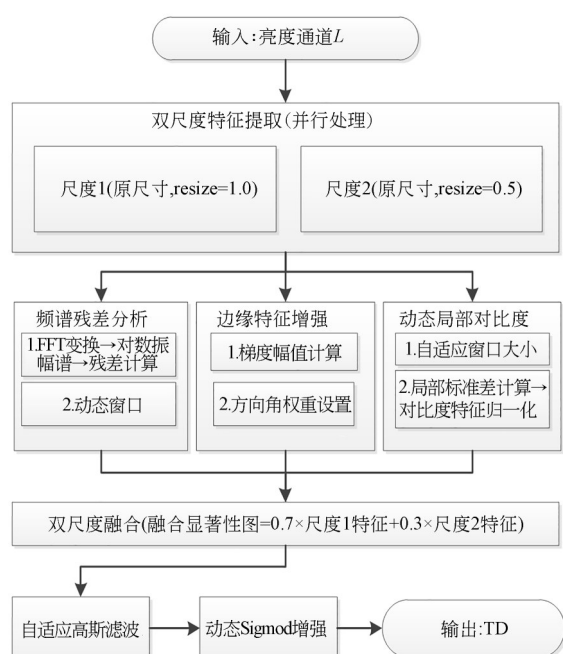


图1 画面细节信息提取过程

Fig. 1 Process of tableau detail information extraction

以输入图像为失真图像(如图2(a))为例,首先,对失真图像的亮度通道(L_D)按不同尺度 s 缩放后,通过对数幅度谱与局部平均谱的残差,并经逆傅里叶变换重构,得到尺度 s 下的频谱残差特征:

$$F_{\text{spec}, D, s} = |\text{ifft2}(\exp((A_{D, s} - \bar{A}_{D, s}) + j \cdot \text{angle}(F_{n, s}^{\text{fit}})))|^2, \quad (1)$$

其中: $F_{n, s}^{\text{fit}}$ 为傅里叶变换结果, $A_{D, s}$ 为对数幅度谱, $\bar{A}_{D, s}$ 为局部平均谱, j 为虚数单位, \exp 用于重构频域矩阵的复数形式, ifft2 表示二维逆傅里叶变换,平方操作用于增强显著区域与背景的差异。利

用上述方法处理图2(a)可得到其频谱残差结果,如图2(b)所示。

为更贴合HVS的方向敏感性特征,并依据自然场景中主导边缘的统计规律,进一步提取方向敏感梯度特征。HVS对水平与垂直方向的结构信息的敏感度比对角方向更高^[29]。为模拟此感知特性,并结合相关参数的性能测试,依据方向角 $\omega_{\theta, D, s}$ 对梯度幅值的重要程度进行加权,赋予水平方向权重1.0,垂直方向0.8,斜向0.5。

$$F_{\text{grad}, D, s} = G_{\text{mag}, D, s} \cdot \omega_{\theta, D, s}, \quad (2)$$

利用上述方法处理图2(a)可得到其处理结果,如图2(c)所示。同时,引入动态窗口局部对比度特征,通过局部标准差反映图像区域内的明暗差异,窗口大小随尺度增大而增大,以捕捉更全局的对比度差异:

$$F_{\text{cont}, D, s} = \text{stdfilt}(L_{D, s}, \mathbf{I}_{k_s \times k_s}), \quad (3)$$

其中, k_s 为动态窗口大小,尺度1为 7×7 ,尺度2为 9×9 , \mathbf{I} 为全1矩阵,确保窗口内所有像素参与标准差计算。 stdfilt 为全1窗口的局部标准差计算函数,利用上述方法处理图2(a)可得到其处理结果,如图2(d)所示。然后,对梯度方向与对比度特征进行全局最大值归一化,与频谱残差特征进行加权融合,得到尺度 s 下的画面细节信息:

$$S_{D, s} = 0.7 \cdot F_{\text{spec}, D, s} + 0.6 \cdot F_{\text{grad}, D, s} + 0.4 \cdot F_{\text{cont}, D, s}, \quad (4)$$

公式(4)中的融合权重通过控制变量法实验优化确定。具体测试过程详见图7与图8。相关权重的相对取值($0.7 > 0.6 > 0.4$)表明,编码视觉显著性的频谱残差特征贡献最为关键,刻画主体结构的方向梯度特征次之,而描述微观层次的局部对比度特征同样具有显著且稳定的正向影响,此结论亦与表3中的数据相互印证。

为克服单一尺度对多尺寸细节捕捉的不足,采用双尺度并行处理策略(原始尺度与半尺度),获取更加全面的TD信息。单尺度TD图如图2(f)所示,虽能定位主体,但对小尺寸物体(如面部结构)及复杂背景(如纹理)的区分力有限。双尺度TD图如图2(e)所示,通过固定权重($q_1 = 0.7, q_2 = 0.3$)融合,可显著提升细节覆盖的完整性:

$$S_{D, \text{multi}} = q_1 \cdot S_{D, 1} + q_2 \cdot S_{D, 2}. \quad (5)$$

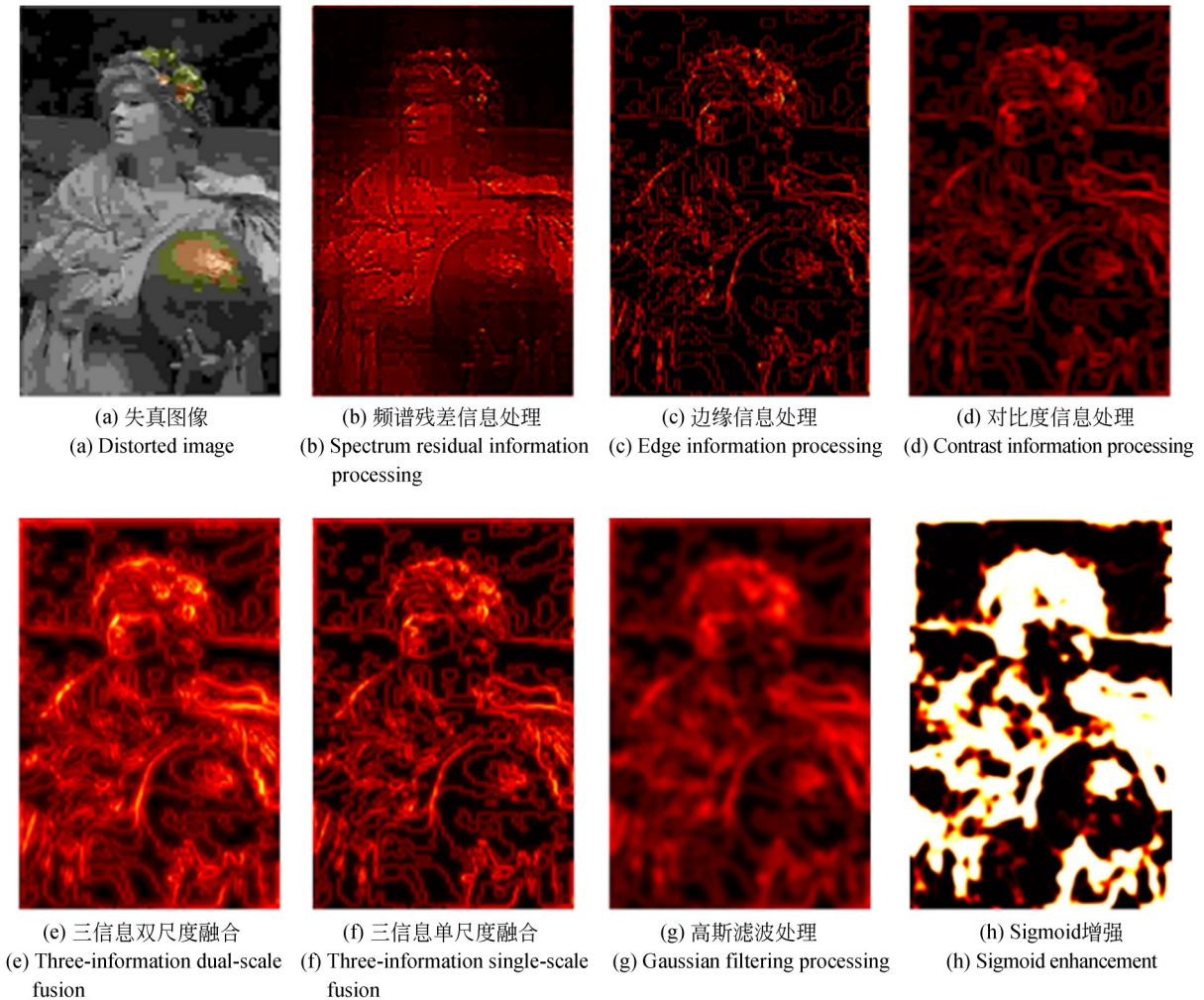


图2 画面细节信息提取过程及对比图

Fig. 2 Process of extracting tableau details and comparison images

双尺度融合权重同样基于控制变量法实验优化确定,具体测试过程详见图9。该权重的分配旨在实现原始尺度与半尺度信息的最优平衡:赋予原始尺度更高的权重是为了确保其丰富的高频细节信息在质量评估中占据主导地位,而适度的半尺度权重则能够有效引入经过下采样平滑后的上下文信息,从而增强模型对噪声及尺度变化的稳定性。

通过对比图2(c)~2(e)的处理结果,频谱残差、梯度与对比度三者的双尺度融合结果,在关键细节的强度与区分度上均具有较好结果,验证了TD融合策略的有效性。

为进一步优化细节图的质量,后续引入自适应高斯滤波(图2(g))有效平滑了孤立噪声点(如墙面斑点),同时保留了关键边缘(如面部,头发)的锐利度。并采用动态Sigmoid函数进行对比度

增强,获得失真图像 TD_D :

$$\mu = \text{mean}(S_{D, \text{gauss}}(:)), \sigma = \text{std}(S_{D, \text{gauss}}(:)), \quad (6)$$

$$TD_D = \frac{1}{1 + \exp\left(-\frac{3}{\sigma + \epsilon} (S_{D, d_{\max}} - \mu)\right)}, \quad (7)$$

其中: μ, σ 为局部窗口的均值和标准差,计算时对图像边界采用镜像填充方式,窗口大小随着尺度的增大而增大(尺度1为 7×7 ,尺度2为 9×9)。 ϵ 为极小常数防止除零,系数8用于控制Sigmoid曲线的陡峭程度:当 σ 较小时(图像对比度高),曲线更陡峭,显著区域与背景差异更突出;当 σ 较大时(图像对比度低),曲线更平缓,避免过度增强导致的噪声放大。经动态Sigmoid增强后,如图2(h)所示,显著区域(如抱着的球、墙的边缘等)对比度强化效果突出,形成视觉聚焦中心。

利用前述处理过程,同样可得到参考图像

TD_R , 借鉴 DSIS 和 DSCQS 的双刺激对比思想, 利用通过公式(8)得到 TD 信息累积量 TDI :

$$TDI = TD_R + TD_D. \quad (8)$$

与现有 IQA 相似性度量不同, 此累积处理并非计算两幅图的“趋同度”, 而是直接累加失真引起的所有细节特征偏移量。这种“镜像计算、差异累积”的机制确保了 TD_D 能够精准反映失真对 TD_R 所表征的细节结构的破坏, 并将细微的细节失真在空间上进行累加与放大。在度量所有细节区域失真程度的总和基础上, 更符合人类对图像整体细节劣化的感知逻辑。

2.2 梯度相似性

梯度幅度的计算通常依赖于特定的算子, 梯度算子卷积运算来近似图像 $f(x)$ 的垂直梯度 $G_y(x)$ 和水平梯度 $G_x(x)$: $G_x(x) = h_x * f(x)$ 和 $G_y(x) = h_y * f(x)$, 其中 h_x 和 h_y 分别代表水平梯度算子和垂直梯度算子, $*$ 表示卷积操作。一阶梯度幅度定义为 $G(x) = \sqrt{G_x^2(x) + G_y^2(x)}$ 。Sobel 算子^[30]、Prewitt 算子^[30] 和 Schar 算子^[31] 是常用的近似一阶导数的梯度算子。在所提出的 IQA 模型中需将梯度信息与其他特征(色度、显著性)融合。Prewitt 算子提取的梯度图与后续的相似度计算兼容性更好, 其均衡的边缘响应能减少因过度敏感(如 Schar)或欠敏感(如简单差分)导致的特征偏差, 确保结构相似性评估的稳定性。

使用 Prewitt 算子计算参考图像(R)和失真图像(D)的亮度 L 通道的梯度大小并通过以下

SSIM 诱导方程计算梯度相似性(GS):

$$GS(x) = \frac{2G_R(x)G_D(x) + C_1}{G_R^2(x) + G_D^2(x) + C_1}, \quad (9)$$

其中: 参数 C_1 是一个用于控制数值稳定性的常数。梯度相似性(GS)在文献[21]、[24]、[32-33]中被广泛使用, 其在测量图像失真的有效性也在文献[32]中进行了深入研究。

上述计算过程无法有效体现参考图像与测试图像之间的关联。特别是 D 与 R 的边缘信息区域出现颜色差异时, 该边缘差异很难被捕捉。因此, 可以通过 D 与 R 融合获取新的融合图像(F), 增强相近颜色处的边缘, 并将 F 与 D, R 分别进行相似计算, 获得最终的边缘相似图^[34]。 F 通过融合 R 和 D 的亮度 L 通道按比例计算得到: $F = \alpha \times R + (1 - \alpha)D$ 。并利用式(10)与(11)得到补充梯度相似性结果:

$$GS_{RF}(x) = \frac{2G_R(x)G_F(x) + C_1}{G_R^2(x) + G_F^2(x) + C_1}, \quad (10)$$

$$GS_{DF}(x) = \frac{2G_D(x)G_F(x) + C_1}{G_D^2(x) + G_F^2(x) + C_1}, \quad (11)$$

其中: G_F 是融合图像 F 的梯度幅度。最后, 将式(9)~式(11)利用式(12)进行融合:

$$\overline{GS}(x) = GS(x) + [GS_{DF}(x) - GS_{RF}(x)]. \quad (12)$$

为了验证融合梯度的性能, 从 LIVE 数据库中选择了具有相同规格但不同失真的参考图像和失真图像, 并对图像进行了分析, 如图 3 所示。

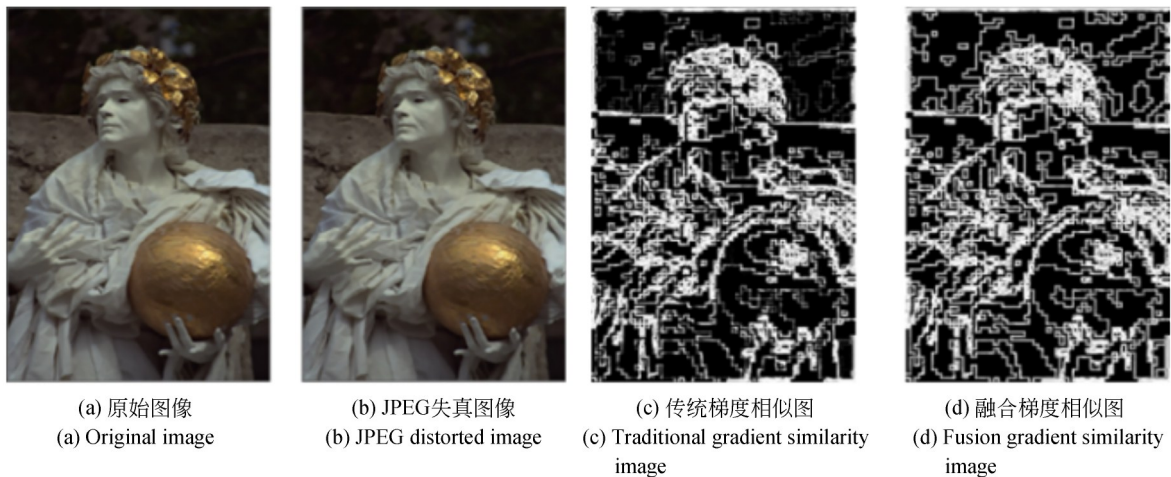


图3 融合梯度与传统梯度的相似比较

Fig. 3 Comparison of similarity of fused and traditional gradients

从图 3 中可以看出,图 3(d)中的亮点及线条等边缘信息相对于图 3(c)较多,说明融合梯度算法对参考图像与失真图像的边缘差异具有较强的表征能力。这些结果表明 \overline{GS} 在预测性能上优于 GS 梯度,更符合 HVS 观察结果,可以更好地表达失真程度。

2.3 色度相似性

根据引言部分的分析,为了有效度量彩色图像质量,色度相似性计算必不可少。同时,合适的图像颜色空间的选择可有效增强参考图像与失真图像的差异度量。通过常用图像颜色空间的测试(具体测试过程见 3.2 节),LHM 被选为本文所提模型的颜色空间,其计算过程如下:

$$\begin{bmatrix} L \\ H \\ M \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ 0.300 & 0.040 & -0.350 \\ 0.340 & -0.600 & 0.170 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (13)$$

进一步对 H, M 通道分别计算颜色相似度。在 LHM 颜色空间中, H 通道(色调)与 M 通道(色度幅度)分别表征了颜色的类别属性与强度属性,而人类视觉系统对颜色的整体感知依赖于二

者的协同作用。为有效建模这种感知协同性,本文通过下面的乘法策略融合各通道的相似度,以实现图像颜色差异更为全面、准确的评估。

$$CS(x) = \frac{2H_R(x)H_D(x) + C_2}{H_R^2(x) + H_D^2(x) + C_2} \times \frac{2M_R(x)M_D(x) + C_2}{M_R^2(x) + M_D^2(x) + C_2}, \quad (14)$$

其中: C_2 是一个用于控制数值稳定性的常数, H_R, H_D 分别代表参考图像与失真图像的 H 通道图, M_R, M_D 分别代表参考图像与失真图像的 M 通道图。

该融合策略的核心机理在于,当任一颜色属性发生显著失真(对应通道相似度降低)时,其乘积结果会非线性地大幅下降,从而对局部严重的颜色失真产生强响应,避免其被另一属性保持良好所掩盖。为直观验证该设计,图 4 将乘法融合结果与加法融合结果进行了对比。由图 4(c)和 4(d)可见, H 通道在蓝色区域(眼睛周边)因压缩伪影产生显著差异;由图 4(e)和 4(f)看出, M 通道对绿色区域(衣服)的色度偏移敏感;通过对比图 4(g)与 4(h)可见,在关键失真区域(如鼻子、嘴巴)乘法融合响应强度与对比度均明显高于加

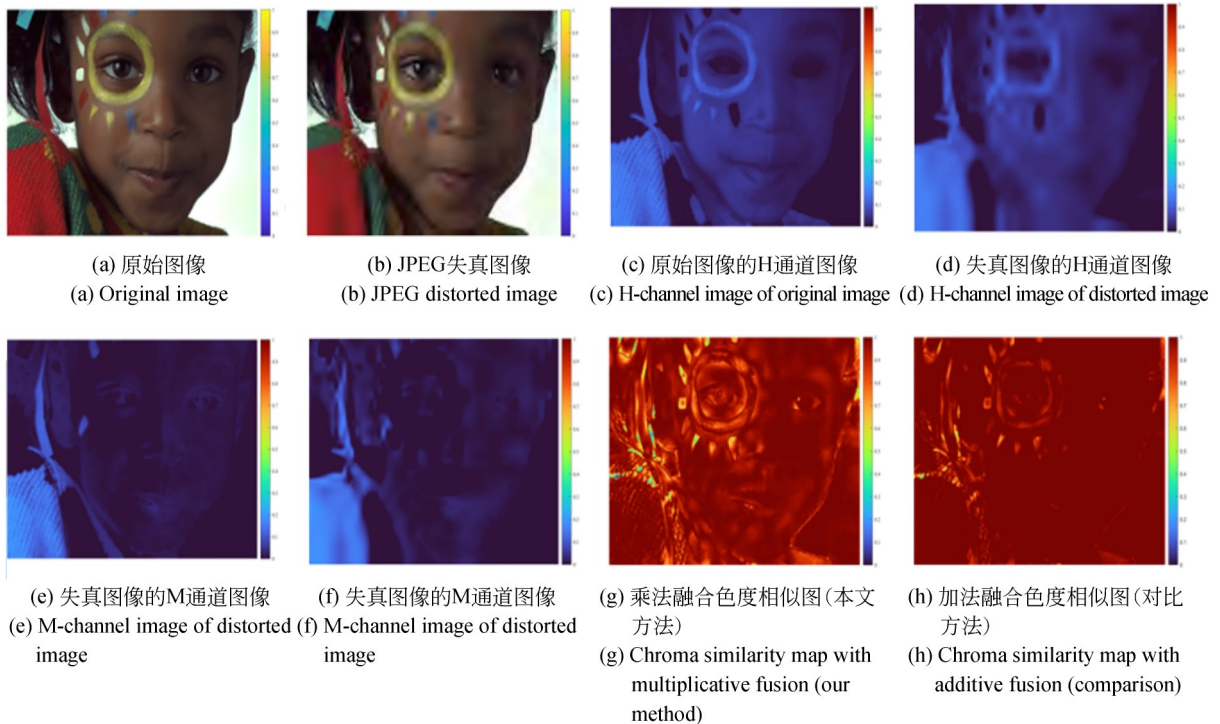


图 4 HM 通道色度分离与失真响应

Fig. 4 HM channel chromatic separation and distortion response

法融合,能更突出人眼敏感的严重失真;而加法融合的结果则相对平缓,部分失真特征(如眼睛周边的蓝色伪影)被弱化。这一视觉对比直观体现了乘法融合在捕捉颜色协同失真方面的优势。

2.4 DTD 模型

利用前述画面细节信息累积处理,梯度相似图及色度相似图,通过在空间上池化可得本文所

提模型,即 DTD。计算流程图如图 5 所示,其源代码链接为 : <https://github.com/AlAlien/DTD>,其池化公式为:

$$DTD = w_1 \cdot \sigma(\overline{GS}) + w_2 \cdot \sigma(CS) + w_3 \cdot \sigma(TDI), \quad (15)$$

其中: w_1, w_2 和 w_3 是表示特征系数,具体取值参数优化过程见 3.2 节。标准差归一化(σ)可以

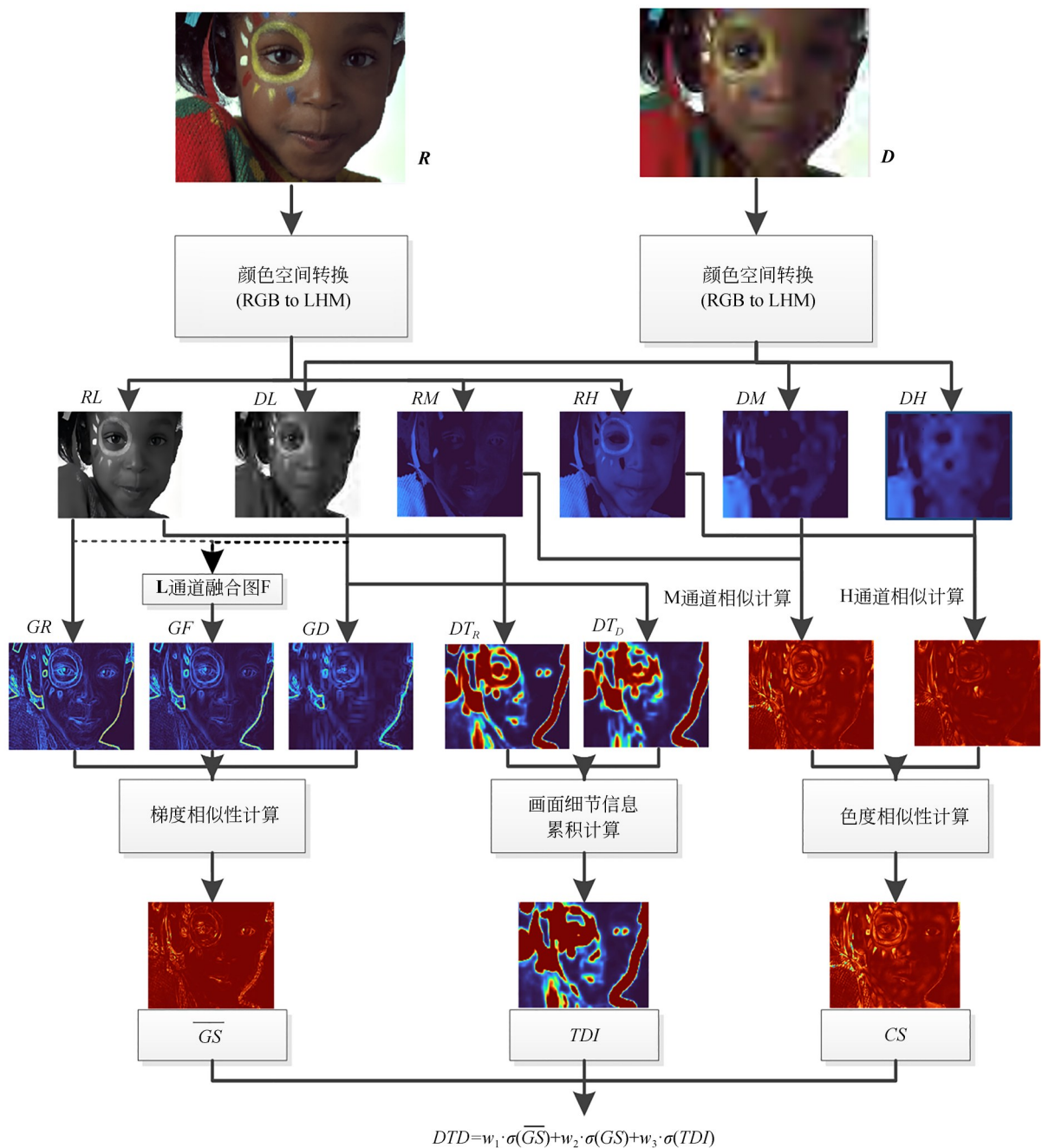


图 5 DTD 的计算过程

Fig. 5 Computational process of DTD

有效地突出图像的特征差异,消除各特征图的数量纲差异。对于输入图像 A 来说, σ 计算公式为:

$$\sigma_A = \sqrt{\frac{1}{M \times N - 1} \sum_{i=1}^M \sum_{j=1}^N (A_{i,j} - \mu_A)^2}, \quad (16)$$

其中: $M \times N$ 为矩阵的维度, μ_A 为矩阵元素的均值。

3 实验结果与分析

3.1 数据库与评价标准

为了评估所提模型的有效性,实验选择了5个公开的图像数据库进行模型验证和比较,包括 TID2008, TID2013, CSIQ, LIVE 和 KADID-10K。上述数据集涵盖多样化的失真类型、差异化的失真程度以及复杂的图案纹理,为开展全面的性能检测提供了理想的测试平台。上述数据库中,每组图像均附有对应的主观评价分数 (Mean Opinion Scores, MOS) 与平均主观分数差异 (Differential Mean Opinion Scores, DMOS)。4个数据库的详细信息汇总于表1。

表1 IQA数据集
Tab.1 IQA databases

数据库	失真类型数	失真图像数	主观分数类型
TID2013	24	3 000	MOS
TID2008	17	1 700	MOS
CSIQ	6	866	DMOS
LIVE	5	779	DMOS
KADID-10K	25	10 125	MOS

为科学评估模型对人类视觉感知的预测能力,将 IQA 客观模型输出的图像质量分数与人类观察者主观评定值进行拟合分析。Spearman 秩序相关系数 (Spearman Rank-Order Correlation Coefficient, SROCC)、Pearson 线性相关系数 (Pearson Linear Correlation Coefficient, PLCC)、Kendall 秩序相关系数 (Kendall Rank-Order Correlation Coefficient, KROCC) 和均方根误差 (Root Mean Squared Error, RMSE) 是广泛采用的四大核心评价指标。

为了计算 PLCC 和 RMSE 指标,采用 logistic 回归^[35]得到与主观判断相同的量表值:

$$p(s) = \beta_1 \left[\frac{1}{2} - \frac{1}{1 + \exp(\beta_2(s - \beta_3))} \right] + \beta_4 s + \beta_5, \quad (17)$$

其中: $\beta_i = 1, 2, \dots, 5$ 是通过最小二乘拟合确定的参数, s 为原始的客观评价分数, $p(s)$ 为回归后的客观评价分数。

3.2 模型的参数优化

如前所述,为了更好地表述彩色图像的质量变化,需要进行图像特定颜色空间的转换,如 YCbCr^[36], LHM^[37], YIQ^[38] 和 YUV^[39]。为了获取与本文所提模型最适配的颜色空间,首先在4个常用公共数据库中进行颜色空间筛选对比测试,其结果如表2所示,排名最高的评估标准值以粗体显示。从表2的数据可以看出, LHM 在4种均匀颜色空间中具有最佳的整体性能,因此在本文提出的模型中,将测试 RGB 图像被转换到 LHM 均匀颜色空间中。

DTD 的参数变量还包括 $\alpha, C_1, C_2, \omega_1, \omega_2$ 和 ω_3 。其中,经过初步测试, α 设置为 0.5 较符合

表2 颜色空间筛选对比测试结果

Tab.2 Color space screening and comparison test results

Color space	TID2013		TID2008		CSIQ		LIVE		Direct average	
	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC
RGB	0.819 1	0.844 9	0.844 2	0.846 4	0.936 5	0.934 4	0.949 9	0.948 7	0.887 4	0.893 6
YUV	0.859 9	0.871 0	0.889 4	0.878 6	0.950 6	0.946 9	0.952 8	0.950 9	0.913 2	0.911 9
LHM	0.868 3	0.878 2	0.895 9	0.884 4	0.955 4	0.951 4	0.954 2	0.951 1	0.918 5	0.916 3
YIQ	0.859 4	0.870 3	0.888 2	0.876 5	0.951 4	0.948 0	0.952 7	0.950 7	0.912 9	0.911 4
YCbCr	0.854 9	0.873 8	0.869 4	0.864 1	0.933 7	0.925 7	0.952 6	0.950 5	0.902 7	0.903 5

入双尺度下频谱残差(SR)、方向敏感梯度(G)及局部对比度(C)3种TD信息,取得最优性能($PLCC=0.8848$)。

本文模型中TD内部特征融合权重及双尺度融合权重的设定均非主观选择,而是基于系统化的实验优化。具体确定过程遵循以下准则:以TID2013数据库为优化平台,以最大化PLCC为统一目标函数,采用控制变量法进行测试。首先,确定TD内部频谱残差、方向梯度与局部对比度三者的融合权重。优化过程以表3第5组($GS+CS$)为性能基线,首先对频谱残差权重 W_{SR} 在 $[0,1]$ 范围内以步长0.1进行一维网格搜索,获得最优值 $W_{SR}=0.7$ 。然后,固定频谱残差权重,对梯度权重 W_G 和对比度权重 W_C 在 $[0,1] \times [0,1]$ 的二维空间中以步长0.1进行联合网格搜索,通过性能曲面分析获得最优组合 $(W_G, W_C)=(0.6, 0.4)$ 。结果对应为表3第12组。上述优化过程如图7和图8所示。

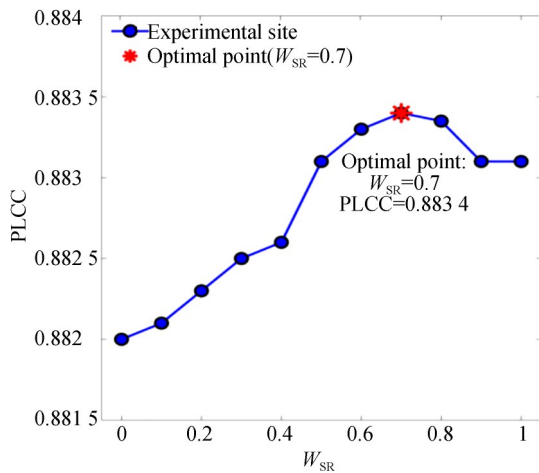


图7 频谱残差权重的一维优化曲线

Fig. 7 One-dimensional optimization curve of spectrum residual weight

对于双尺度融合权重公式(6),同样采用网格搜索策略,对原始尺度权重 q_1 在0.1至0.9范围内以步长0.1进行精细扫描,确定在 $q_1=0.7$ (对应 $q_2=0.3$)时模型性能达到峰值,且在 $q_1 \in [0.6, 0.8]$ 内性能保持稳定高值,表明双尺度权重具有较好的稳定性。上述优化过程如图9所示。

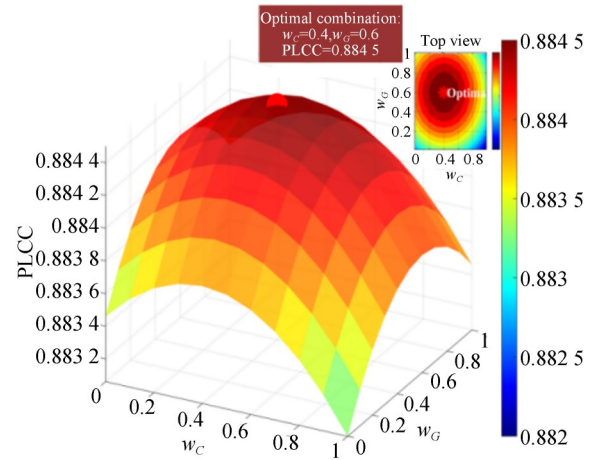


图8 梯度与对比度权重的二维优化曲面

Fig. 8 Two-dimensional optimization surface of gradient and contrast weight

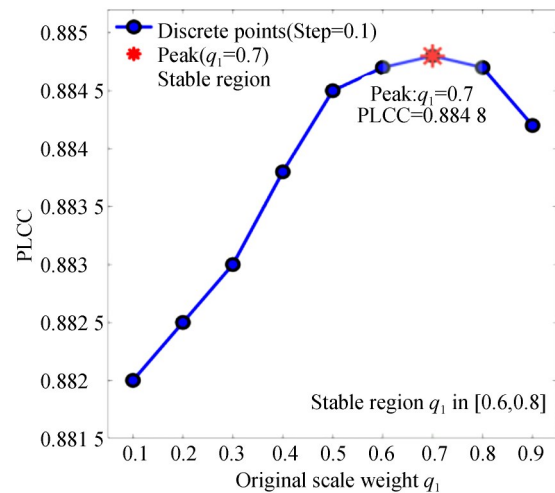


图9 双尺度权重优化分析曲线

Fig. 9 Double scale weight optimization analysis curve

3.4 整体性能对比

为了证明所提模型的性能优势,将所提出的DTD模型与其他典型的模型进行了比较,包括SSIM^[14], VIF^[16], RVSIM^[17], SSVD^[18], FSIM^[20], CAGS^[22], VCGS^[24]和CVSS^[25],以及近年来提出的模型CPCCs(2023)^[23]和TSVD(2025)^[19]。表4中,4个数据库中每个评价指标对应最好的3个结果使用粗体突出显示。此外,对4个数据库中的评价结果进行加权平均和直接平均计算,其中每个数据库的权值是依据其失真图像数量确定的。

从表 4 可以看出,DTD 模型对所有数据库的性能都较好。尤其在 CSIQ, TID2008, 和 TID2013 数据库中,该模型始终位列前三。对于 LIVE 数据库,DTD 模型仍具有竞争性,与前三的差距很小。虽然 VCGS 和 TSVD 可以在 TID2008 数据库中得到前三的结果,但是其整体性能与 DTD 还有一定差距。此外,在 PLCC 指标的加权平均值和直接平均值中,所提模型的性能位列第一。DTD 模型在所有图像质量评估模型中前三排名数量最高(20 次),其次是 TSVD (13 次),VCGS(13 次)。从统计意义更强的加权平均与直接平均指标来看,DTD 模型在多数评价指标上也均表现领先,例如其 PLCC 加权平均为 0.909 7,SROCC 加权平均为 0.900 6。此外,DTD 模型在所有数据库测试得到的

SROCC 和 PLCC 值均大于 0.88,所以可以认为 DTD 模型既具有较高的性能,又具有较好的泛化性。特别地,由于 TSVD 模型代码未开源,相关数据有所缺失,所以表 4 仅提供可查阅的 TSVD 论文中的相关数据。但与 TSVD 模型性能的比较中可以发现,DTD 模型与 TSVD 模型的性能差距较小,在某些数据库上表现甚至更优,与当前最新的研究结果相比具有竞争性。综上所述,本文模型具有很好的精度值,显示出良好的竞争力。为了进一步验证 DTD 模型在不同类型的失真情况下具有更好的性能优势,通过 TID2013 数据库测得的拟合离散点图进行展示,结果如图 10 所示。从图中可以看出,与最近提出模型相比,DTD 模型的客观评分与主观评分之间存在着高度的相关性。

表 4 对比不同典型 IQA 模型的 4 个数据库性能

Tab. 4 Performance comparison of state-of-the-art IQA models on four databases

Model	TID2008				TID2013				CSIQ			
	SROCC	KROCC	PLCC	RMSE	SROCC	KROCC	PLCC	RMSE	SROCC	KROCC	PLCC	RMSE
SSIM	0.774 9	0.576 8	0.773 2	0.851 1	0.741 7	0.558 8	0.789 5	0.760 8	0.875 6	0.690 7	0.861 3	0.133 4
VIF	0.749 1	0.586 0	0.808 4	0.789 9	0.676 9	0.514 7	0.772 0	0.788 0	0.919 5	0.753 7	0.927 7	0.098 0
FSIM	0.884 0	0.699 1	0.876 2	0.646 8	0.851 0	0.666 5	0.876 9	0.595 9	0.931 0	0.769 0	0.919 2	0.103 4
RVSIM	0.737 5	0.562 8	0.795 4	0.813 3	0.675 7	0.514 6	0.782 5	0.771 9	0.897 9	0.723 4	0.923 6	0.100 7
CVSS	0.900 1	0.721 5	0.896 1	0.595 6	0.806 9	0.633 1	0.840 6	0.671 5	0.958 0	0.817 3	0.958 9	0.074 5
SSVD	0.884 7	0.710 7	0.894 1	0.601 0	0.811 2	0.646 8	0.871 5	0.608 0	0.897 5	0.725 6	0.887 8	0.120 8
VCGS	0.897 5	0.717 1	0.877 6	0.643 3	0.892 6	0.716 6	0.900 0	0.540 4	0.944 3	0.790 6	0.930 1	0.096 4
CAGS	0.823 1	0.628 9	0.809 1	0.788 6	0.831 6	0.646 9	0.844 5	0.663 9	0.919 8	0.748 7	0.901 4	0.113 7
CPCCs	0.891 6	0.704 0	0.886 8	0.620 1	0.859 2	0.673 8	0.878 1	0.593 2	0.950 5	0.800 9	0.953 7	0.079 0
TSVD	0.891 0	0.739 0	0.894 1	—	0.894 9	0.706 7	0.897 7	—	0.966 5	0.806 0	0.937 4	—
DTD	0.906 7	0.733 3	0.902 4	0.578 0	0.864 8	0.684 9	0.884 8	0.578 3	0.961 0	0.82 10	0.957 8	0.075 5

Model	LIVE				权重平均			直接平均		
	SROCC	KROCC	PLCC	RMSE	SROCC	KROCC	PLCC	SROCC	KROCC	PLCC
SSIM	0.947 9	0.796 3	0.944 9	8.945 5	0.794 2	0.610 8	0.814 0	0.835 0	0.655 7	0.842 2
VIF	0.963 6	0.828 2	0.960 4	7.613 7	0.764 6	0.604 9	0.826 1	0.827 3	0.670 7	0.867 1
FSIM	0.964 5	0.836 3	0.961 3	7.529 6	0.884 7	0.710 1	0.892 8	0.907 6	0.742 7	0.908 4
RVSIM	0.960 0	0.820 3	0.957 0	7.927 4	0.763 8	0.600 6	0.830 7	0.817 8	0.655 3	0.864 6
CVSS	0.967 2	0.840 6	0.965 1	7.157 3	0.875 1	0.711 5	0.889 3	0.908 0	0.753 1	0.915 2
SSVD	0.959 3	0.819 2	0.956 0	8.060 5	0.860 9	0.695 8	0.890 0	0.888 2	0.725 6	0.902 3
VCGS	0.976 8	0.875 2	0.967 6	7.903 5	0.913 3	0.750 3	0.908 3	0.927 8	0.774 9	0.918 8
CAGS	0.973 4	0.865 8	0.964 0	8.325 1	0.858 7	0.682 6	0.857 2	0.887 0	0.722 6	0.879 7
CPCCs	0.965 3	0.835 6	0.961 6	7.498 3	0.896 4	0.730 0	0.901 0	0.916 7	0.753 6	0.920 1
TSVD	0.960 7	0.820 1	0.967 0	—	0.897 2	0.732 1	0.905 9	0.928 3	0.768 0	0.924 1
DTD	0.957 1	0.827 0	0.967 8	5.822 0	0.900 6	0.733 4	0.909 7	0.922 4	0.766 5	0.928 2

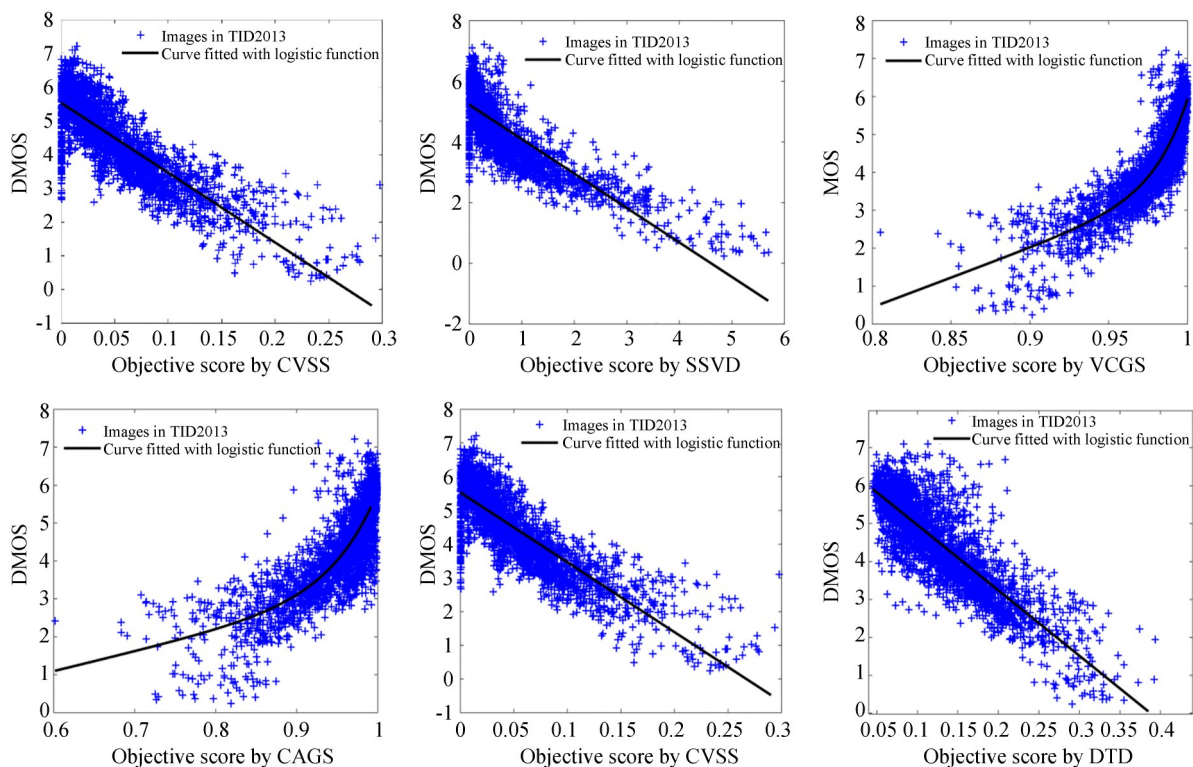


图 10 基于 TID2013 数据库的主观 MOS 与 IQA 模型计算结果拟合对比

Fig. 10 Scatter plots of subjective MOS against scores calculated by IQA models prediction for TID2013 databases

3.5 基于深度学习模型的跨数据库性能对比

为全面验证模型泛化能力并与前沿方法对标,本节进一步将 DTD 模型与近年来代表性 10 种基于深度学习的 IQA 模型进行了对比,同时引入大规模图像质量评价数据集 KADID-10K 进行跨库验证,结果如表 5 所示。由于代码开源的原

因,相关模型的测试结果均来与已发表的文献^[40-41]。KADID-10K 数据库含 10 125 张失真图像,涵盖 25 种失真类型,能够更客观、全面地检验模型在复杂失真场景下的泛化性与稳定性。表 5 中,各项指标前三名已加粗,DTD 模型在包括 KADID-10K 在内的 5 个数据库上均表现稳定,指

表 5 与基于深度学习的 IQA 模型的 5 个数据库性能对比

Tab. 5 Performance comparison with deep learning based IQA models of five databases

Model	LIVE		CSIQ		TID2008		TID2013		KADAD-10K	
	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC
SSIM-CNN	0.965 0	0.963 0	0.952 0	0.946 0	0.770 0	0.737 0	0.759 0	0.752 0	—	—
DeepSIM	0.968 0	0.974 0	0.919 0	0.919 0	0.876 0	0.887 0	0.872 0	0.846 0	—	—
DeepIQA	0.940 0	0.947 0	0.901 0	0.909 0	0.917 0	0.908 0	0.834 0	0.831 0	—	—
PieAPP	0.908 0	0.919 0	0.877 0	0.892 0	0.610 0	0.788 0	0.859 0	0.876 0	—	—
LPIPS	0.954 0	0.954 0	0.928 0	0.929 0	0.705 0	0.668 0	0.855 0	0.830 0	0.809 0	0.814 0
WaDIQaM	0.959 0	0.960 0	0.908 0	0.899 0	—	—	0.868 0	0.856 0	0.775 0	0.762 0
MetaIQA	0.971 0	0.968 0	0.959 0	0.946 0	—	—	0.865 0	0.816 0	0.856 0	0.851 0
DISTS	0.954 0	0.954 0	0.928 0	0.929 0	0.705 0	0.668 0	0.855 0	0.830 0	0.809 0	0.814 0
HyperIQA	0.965 0	0.949 0	0.838 0	0.825 0	—	—	0.858 0	0.846 0	0.855 0	0.850 0
TIQA	0.955 0	0.960 0	0.844 0	0.852 0	—	—	0.855 0	0.835 0	0.752 0	0.739 0
DTD(our)	0.968 0	0.957 0	0.958 0	0.961 0	0.902 0	0.907 0	0.885 0	0.865 0	0.877 0	0.878 0

标与多数深度学习模型相当。而对比的深度学习模型在不同库间性能波动较大,且依赖大量训练数据,在实用性与泛化性上与DTD存在差距。结果进一步表明,DTD不仅预测精度高,还具有出色的跨数据库适应能力与稳定性。

3.6 单个失真类型的性能对比

一个优秀的IQA模型需在各种失真类型上均表现出优异的预测能力。本节基于表6汇总的结果,评估了各模型在多种失真类型上的性能表现。鉴于TID2013数据库涵盖了TID2008

的所有失真类型,故未单独列出TID2008的结果。本研究选用SROCC作为核心评价指标,因其PLCC、RMSE和KROCC等指标呈现相似趋势。为此,本节从3个主要数据库中选取了35组涵盖不同失真类型的图像进行测试。对于每个数据库中的每种失真类型,表6中SROCC值排名前三的结果均已加粗标示。进一步,对于最新TSVD模型,与DTD对比TID2013数据库24种失真类型的SROCC结果,结果如图11所示。

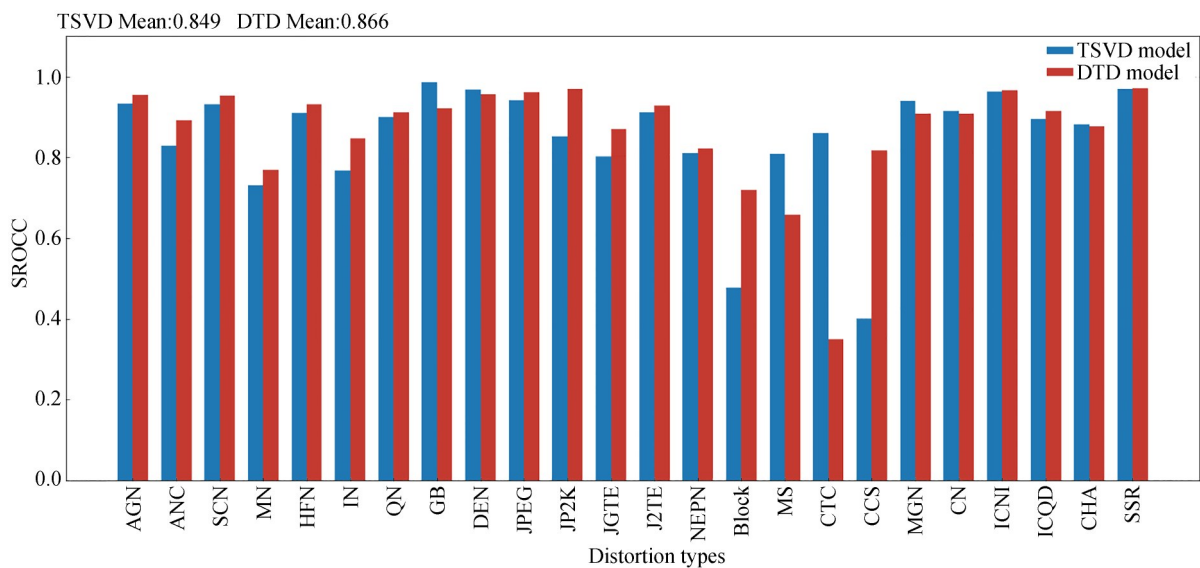


图11 DTD与TSVD在TID2013数据库中不同失真类型SROCC值对比

Fig. 11 PLCC comparison of different distortion types between DTD and TSVD on TID2013

由表6和图10可知,DTD模型在绝大多数失真类型上均表现出显著的性能优势。尤其是在TID2013和LIVE数据库所涵盖的失真类型上,其SROCC值普遍高于其他典型模型。在CSIQ数据库上,DTD模型同样保持竞争力,其表现与当前先进模型相当。对于35种失真类型,每个模型的评价结果各有优势,其中SROCC值大于0.95的数量最多模型是DTD(17次),其次是CAGS(14次)。对于CAGS模型,其最小SROCC值为0.3711,失真类型为CCS,该类型DTD模型SROCC值为0.8172。对比DTD与最新模型TSVD,DTD在所有测试失真类型中,有24个单一失真类型的测试结果优于TSVD,占所有测试失真类型的69%。此外,在局部块状扭曲(Block)和颜色饱和度变化(CCS)失真上,TS-

VD的SROCC值仅为0.4768和0.4015;而DTD模型则取得了0.7189和0.8172的结果,相对提升了51%和104%,且在Block测试中表现最好,显著优于所有对比模型(其余模型均未超过0.67),在CCS测试中结果排名第二。值得注意的是,模型在不同失真类型上的性能表现存在差异:在CCS(颜色饱和度变化)失真上DTD表现优异(SROCC=0.8172),而在CTC(对比度变化)失真上表现相对有限(SROCC=0.3498)。这一差异源于两种失真的本质机制与模型特征提取方式的匹配程度不同。CCS失真主要表现为色度通道的全局强度变化,可直接被色度相似性(CS)分量敏感捕获;而CTC失真本质是亮度通道的全局非线性映射,当前模型的结构特征(梯度相似度(GS)与画面细节累

表 6 IQA 模型在不同失真类型下的 SROCC 值

Tab. 6 SROCC values of IQA model under different distortion types

数据库	失真类型	SSIM	VIF	FSIM	RVSIM	CVSS	SSVD	VCGS	CAGS	CPCCs	TSVD	DTD	
TID2013	AGN	0.867 1	0.899 4	0.910 1	0.864 5	0.940 1	0.922 1	0.936 8	0.935 9	0.932 5	0.932 6	0.954 5	
	ANC	0.772 6	0.829 9	0.853 7	0.812 4	0.863 9	0.828 9	0.855 9	0.865 3	0.843 4	0.828 7	0.891 5	
	SCN	0.851 5	0.883 5	0.890 0	0.841 2	0.907 7	0.936 9	0.931 5	0.927 6	0.911 9	0.931 5	0.953 8	
	MN	0.776 7	0.845 0	0.809 4	0.822 8	0.771 5	0.732 5	0.807 0	0.752 6	0.781 5	0.731 4	0.769 3	
	HFN	0.863 4	0.897 2	0.909 4	0.883 7	0.909 7	0.899 5	0.916 2	0.915 9	0.908 0	0.910 2	0.9307	
	IN	0.750 3	0.853 7	0.825 1	0.882 3	0.745 7	0.767 1	0.868 2	0.836 1	0.781 4	0.767 1	0.847 0	
	QN	0.865 7	0.785 4	0.880 7	0.751 9	0.886 9	0.857 2	0.883 1	0.871 8	0.899 3	0.899 6	0.911 9	
	GB	0.966 8	0.965 0	0.955 1	0.974 2	0.934 8	0.949 9	0.954 9	0.961 4	0.926 8	0.986 2	0.921 0	
	DEN	0.925 4	0.891 1	0.933 0	0.894 7	0.942 7	0.948 6	0.945 4	0.946 6	0.946 6	0.968 2	0.957 1	
	JPEG	0.920 0	0.919 2	0.933 9	0.930 6	0.952 1	0.931 8	0.959 7	0.958 5	0.946 3	0.941 1	0.960 8	
	JP2K	0.946 8	0.951 6	0.958 9	0.952 1	0.958 7	0.968 8	0.968 6	0.962 0	0.954 7	0.852 7	0.969 7	
	JGTE	0.849 3	0.840 9	0.861 0	0.843 4	0.861 3	0.844 1	0.895 5	0.864 4	0.883 2	0.802 5	0.870 8	
	J2TE	0.882 8	0.876 1	0.891 9	0.885 4	0.885 1	0.933 2	0.920 4	0.925 0	0.880 5	0.911 5	0.928 6	
	NEPN	0.7821	0.772 0	0.793 7	0.750 2	0.820 1	0.808 5	0.788 7	0.783 3	0.831 8	0.810 2	0.821 8	
	Block	0.572 0	0.530 6	0.553 2	0.603 7	0.515 2	0.476 8	0.432 6	0.601 5	0.662 0	0.476 8	0.718 9	
	MS	0.775 2	0.627 6	0.748 7	0.607 9	0.715 0	0.750 7	0.764 6	0.744 1	0.710 7	0.808 2	0.657 5	
	CTC	0.377 5	0.838 6	0.467 9	0.152 6	0.294 0	0.449 4	0.468 7	0.451 4	0.314 7	0.859 5	0.349 8	
	CCS	0.414 1	0.309 9	0.835 9	0.399 4	0.261 4	0.356 2	0.781 6	0.371 1	0.584 0	0.401 5	0.817 2	
	MGN	0.780 3	0.846 8	0.856 9	0.823 3	0.879 9	0.860 0	0.894 8	0.870 0	0.868 8	0.939 2	0.909 1	
	CN	0.856 6	0.894 6	0.913 5	0.898 4	0.935 1	0.926 3	0.926 8	0.916 8	0.934 5	0.915 2	0.909 1	
	LCNI	0.905 7	0.920 4	0.948 5	0.915 6	0.962 9	0.964 9	0.955 6	0.957 4	0.965 6	0.963 4	0.967 0	
	ICQD	0.854 2	0.841 4	0.881 5	0.803 6	0.910 8	0.894 8	0.905 0	0.906 0	0.906 1	0.894 8	0.914 4	
	CHA	0.877 5	0.884 8	0.892 5	0.915 4	0.852 3	0.881 6	0.886 1	0.876 8	0.857 7	0.882 1	0.876 6	
	SSR	0.946 1	0.935 3	0.957 6	0.943 9	0.960 5	0.970 0	0.963 5	0.958 0	0.960 3	0.970 1	0.971 6	
	LIVE	JP2K	0.961 4	0.969 6	0.964 9	0.962 2	0.971 9	0.968 1	0.984 1	0.982 2	0.970 7	0.947 8	0.981 6
		JPEG	0.976 4	0.984 6	0.980 8	0.982 2	0.983 6	0.980 5	0.984 9	0.983 6	0.980 4	0.980 5	0.959 6
AWGN		0.969 4	0.985 8	0.966 7	0.969 2	0.980 9	0.981 6	0.989 6	0.983 7	0.966 9	0.966 9	0.985 4	
GB		0.951 7	0.972 8	0.972 0	0.964 3	0.966 2	0.938 7	0.976 3	0.964 1	0.968 9	0.968 9	0.976 0	
FF	0.955 6	0.965 0	0.944 2	0.958 1	0.959 2	0.939 2	0.968 3	0.963 3	0.968 0	0.939 5	0.970 5		
CSIQ	AGWN	0.897 4	0.957 5	0.938 0	0.937 7	0.967 0	0.947 4	0.967 1	0.965 2	0.949 0	0.952 4	0.967 7	
	JPEG	0.954 6	0.970 5	0.966 2	0.959 9	0.968 9	0.959 3	0.964 3	0.957 3	0.968 2	0.956 3	0.960 6	
	JP2K	0.960 6	0.967 2	0.968 3	0.969 3	0.977 7	0.969 2	0.974 7	0.954 5	0.973 0	0.979 5	0.975 9	
	AGPN	0.892 2	0.951 1	0.905 9	0.931 9	0.951 6	0.939 2	0.962 4	0.949 2	0.932 6	0.959 5	0.959 6	
	GB	0.960 9	0.974 5	0.978 2	0.971 5	0.978 9	0.971 3	0.973 4	0.957 4	0.975 4	0.986 5	0.971 3	
GCD	0.792 2	0.934 5	0.953 9	0.879 2	0.932 4	0.869 0	0.955 3	0.927 3	0.927 4	0.965 2	0.947 8		

积(TDI))均基于亮度通道,GS对边缘相对强弱模式的全局变化不敏感;TDI虽包含局部对比度特征,但其基于局部窗口的计算方式难以有效表征全局性的亮度关系重组,从而导致模型对CTC失真的评估能力未能达到预期。这一分析也揭示了未来模型的改进方向。在对各类失真类型的评估中,DTD模型也表现出优异的性能一致性:其SROCC在71%的失真类型中达到0.9以上,在89%的失真类型中达到0.8以上。这一结果显著优于其他IQA模型。

综合上述针对不同失真类型的评估结果,DTD模型在多个数据库及广泛的失真类型上均表现出更高的预测精度和更强的泛化能力。

3.7 模型计算效率比较

除预测精度外,计算效率是设计高质量IQA模型需考量的关键因素。为量化分析计算复杂度,本研究在MATLAB R2022a平台(Intel Core i5处理器,16GB内存)上,选用TID2008数据库的1700幅失真图像进行测试,并参考本领域的其他研究^[22-25,42],以单对彩色图像的平均运行时间作为DTD与多种开源FR-IQA模型计算效率的依据,结果如表7所示。DTD模型平均耗时仅0.065 s,显著低于多数对比模型,且与PSNR,SSIM的运行时间差距不大,并在预测精度方面,显著优于上述模型。较高的计算效率得益于以下几个方面的设计:采用动态下采样自适应约束处理尺度,显著降低了数据量;全程使用向量化矩阵运算,并调用imfilter,fft2等高度优化的内置函数;梯度、颜色与显著性特征的提取过程具备天然并行性;融合阶段仅涉及标准差计算与加权求和,计算开销极低。在实际应用中,本文所提模型可显著提升大规模图像质量评估任务的执行效率,满足实时处理需求。

表7 不同FR-IQA模型的计算效率对比

Tab.7 Computing efficiency comparison of different IQA models

IQA模型	运行时间/s	IQA模型	运行时间/s
PSNR	0.011	SSIM	0.023
FSIM	0.108	SPSIM	0.123
RVSIM	0.457	SSVD	0.068
VCGS	0.224	CAGS	0.127
CPCCs	0.214	DTD	0.065

4 结论

本文提出了一种FR-IQA模型,即DTD模型。该模型通过双尺度TD信息累积图,梯度相似图以及色度相似图,分别表征了TD信息累积效应,单一结构失真和单一颜色失真。该模型将图像转为LHM颜色空间,动态下采样保留结构,通过双尺度处理策略,提取频谱残差信息、方向加权边缘信息特征和自适应局部对比度信息,并采用7:3的主次权重融合双尺度特征。随后,进行感知优化处理,包括自适应高斯混合滤波以及动态Sigmoid增强,生成符合人眼视觉特性的累积效应图。然后,通过Prewitt算子计算梯度,构建梯度相似性矩阵,增强色度方向一致性获得颜色相似性。最后,将TD累积效应、梯度相似性和颜色相似性权重的标准差按权重融合,并通过非线性映射得到本文的计算模型。模型中的主要参数通过控制变量法确定。为了证明所提出的模型的优异性能,本文在5个常用的IQA数据库(LIVE,CSIQ,TID2008,TID2013及KADID-10K)上进行了系统测试,共计包含190幅参考图像与16470幅失真图像。实验除与10个具有代表性的传统FR-IQA模型进行对比外,还与10个近期发表的基于深度学习的IQA模型进行了跨数据库性能比较。实验结果显示,本文模型在5个数据库中的PLCC最低值为0.8768,最高值可达0.9678,其稳定性和精度均优于大多数模型;在单一失真类型测试中,本文模型测得的SROCC达到0.9以上,占71%的失真类型(TID2013),表明其具有良好的泛化能力;通过运行时间对比,验证了本文模型具有较高的计算效率。综上所述,本文提出的客观FR-IQA模型具有较为稳定的算法性能。未来,将进一步优化模型,提升预测精度和泛化性能,同时争取通过自建工业视觉系统IQA数据库,完成本领域测试对象从单一数据库到实际工业场景的转换。

作者贡献说明:

史晨阳:论文审核与编辑;

吴俊杰:方法提出,论文构思与撰写;

袁瀚成:数据整理;

吴路路:实验数据分析。

参考文献:

- [1] 高敏娟, 党宏社, 魏立力, 等. 全参考图像质量评价回顾与展望[J]. 电子学报, 2021, 49(11): 2261-2272.
GAO M J, DANG H SH, WEI L L, *et al.* Review and prospect of full reference image quality assessment[J]. *Acta Electronica Sinica*, 2021, 49(11): 2261-2272. (in Chinese)
- [2] 常志文, 王立忠, 梁晋, 等. 基于图像块分解融合的水下标定图像增强[J]. 中国光学(中英文), 2024, 17(4): 810-822.
CHANG ZH W, WANG L ZH, LIANG J, *et al.* Underwater calibration image enhancement based on image block decomposition and fusion[J]. *Chinese Optics*, 2024, 17(4): 810-822. (in Chinese)
- [3] 程丽君, 孙正, 孙美晨, 等. 一种基于前向成像模型的光声层析图像重建方法[J]. 中国光学(中英文), 2024, 17(2): 444-455.
CHENG L J, SUN ZH, SUN M CH, *et al.* A photoacoustic tomography image reconstruction method based on forward imaging model[J]. *Chinese Optics*, 2024, 17(2): 444-455. (in Chinese)
- [4] 韩昊男, 钱锋, 吕建成, 等. 图像去雾方法质量评价[J]. 光学精密工程, 2022, 30(6): 721-733.
HAN H N, QIAN F, LÜ J W, *et al.* Image dehazing method quality assessment[J]. *Optics and Precision Engineering*, 2022, 30(6): 721-733. (in Chinese)
- [5] 惠子薇, 何坤, 冯霖, 等. 基于视觉特性的图像质量评价[J]. 计算机工程, 2023, 49(7): 189-195.
HUI Z W, HE K, FENG B, *et al.* Image quality assessment based on visual characteristics[J]. *Computer Engineering*, 2023, 49(7): 189-195. (in Chinese)
- [6] SINGLA A. *Comparison of Different Subjective Test Methods* [M]. Assessment of Visual Quality and Simulator Sickness for Omnidirectional Videos. Cham: Springer Nature Switzerland, 2023: 51-64.
- [7] VENKATARAMANAN A K, STEJEREAN C, KATSAVOUNIDIS I, *et al.* One transform to compute them all: efficient fusion-based full-reference video quality assessment[J]. *IEEE Transactions on Image Processing*, 2024, 33: 509-524.
- [8] ZHOU W, YUE G H, ZHANG R Z, *et al.* Reduced-reference quality assessment of point clouds via content-oriented saliency projection[J]. *IEEE Signal Processing Letters*, 2023, 30: 354-358.
- [9] 陈楠, 蔡柏亲, 杨恭宁, 等. 考虑结构一致性和内容完整性的重定向图像质量评价[J]. 福州大学学报(自然科学版), 2023, 51(4): 517-524.
CHEN N, CAI B Q, YANG G N, *et al.* Image re-targeting quality assessment using content integrity and structural consistency[J]. *Journal of Fuzhou University (Natural Science Edition)*, 2023, 51(4): 517-524. (in Chinese)
- [10] 张威, 赵世灵, 刘银豪, 等. 多尺度时空特征聚合的全参考视频质量评价[J]. 计算机工程与应用, 2023, 59(18): 154-162.
ZHANG W, ZHAO SH L, LIU Y H, *et al.* Full reference video quality assessment based on multi-scale spatiotemporal feature aggregation[J]. *Computer Engineering and Applications*, 2023, 59(18): 154-162. (in Chinese)
- [11] JIANG B C, BIAN S L, SHI CH Y, *et al.* Full-reference image quality assessment based on multi-channel visual information fusion[J]. *Applied Sciences*, 2023, 13(15): 8760.
- [12] NARWARIA M, LIN W S. SVD-based quality metric for image and video using machine learning[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2012, 42(2): 347-364.
- [13] 于天河, 柳梦瑶. 基于人眼视觉系统的图像质量评价方法[J]. 北京邮电大学学报, 2023, 46(2): 129-136.
YU T H, LIU M Y. Image quality evaluation method based on human visual system[J]. *Journal of Beijing University of Posts and Telecommunications*, 2023, 46(2): 129-136. (in Chinese)
- [14] 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.
- [15] SHEIKH H R, BOVIK A C, DE VECIANA G. An information fidelity criterion for image quality assessment using natural scene statistics[J]. *IEEE Transactions on Image Processing*, 2005, 14(12): 2117-2128.
- [16] SHEIKH H R, BOVIK A C. Image information and visual quality[J]. *IEEE Transactions on Image Processing*, 2006, 15(2): 430-444.
- [17] YANG G Y, LI D S, LU F, *et al.* RVSIM: a feature similarity method for full-reference image quality assessment[J]. *EURASIP Journal on Image and Video Processing*, 2018, 2018(1): 6.

- [18] MANSOURI A, MAHMOUDI-AZNAVEH A. SSVD: Structural SVD-based image quality assessment[J]. *Signal Processing: Image Communication*, 2019, 74: 54-63.
- [19] 李佳欣, 段发阶, 傅骁, 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]. *Optics and Precision Engineering*, 2025, 33(1): 107-122. (in Chinese)
- [20] 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.
- [21] SUN W, LIAO Q M, XUE J H, *et al.* SPSIM: a superpixel-based similarity index for full-reference image quality assessment [J]. *IEEE Transactions on Image Processing*, 2018, 27(9): 4232-4244.
- [22] 史晨阳, 林燕丹. 基于图像色貌和梯度特征的图像质量客观评价[J]. *物理学报*, 2020, 69(22): 401-412.
SHI CH Y, LIN Y D. Objective image quality assessment based on image color appearance and gradient features[J]. *Acta Physica Sinica*, 2020, 69(22): 401-412. (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 SH L, SHI CH Y, *et al.* Full reference image quality assessment based on color appearance-based phase consistency [J]. *Optics and Precision Engineering*, 2023, 31(10): 1509-1521. (in Chinese)
- [24] SHI CH Y, LIN Y D. Full reference image quality assessment based on visual salience with color appearance and gradient similarity [J]. *IEEE Access*, 2020, 8: 97310-97320.
- [25] JIA H Z, ZHANG L, WANG T H. Contrast and visual saliency similarity-induced index for assessing image quality [J]. *IEEE Access*, 2018, 6: 65885-65893.
- [26] 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.
- [27] HOU X D, ZHANG L Q. Saliency detection: a spectral residual approach[C]. *2007 IEEE Conference on Computer Vision and Pattern Recognition. June 17-22, 2007, Minneapolis, MN, USA.* IEEE, 2007: 1-8.
- [28] 姚军财, 申静. 基于图像画面信息及其视觉感知的无参考图像质量评价[J]. *光学精密工程*, 2025, 33(18): 2944-2961.
YAO J C, SHEN J. No reference image quality assessment based on image tableau information and their visual perception[J]. *Optics and Precision Engineering*, 2025, 33(18): 2944-2961. (in Chinese)
- [29] CAMPBELL F W, ROBSON J G. Application of Fourier analysis to the visibility of gratings [J]. *The Journal of Physiology*, 1968, 197(3): 551-566.
- [30] JAIN R, KASTURI R, SCHUNCK B G. *Machine Vision*[M]. New York: McGraw-Hill, 1995.
- [31] SONKA M, HLAVAC V, BOYLE R. *Image Processing, Analysis, and Machine Vision*[M]. 3rd ed. Toronto: Thomson Learning, 2008.
- [32] 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.
- [33] KIM D O, HAN H S, PARK R H. Gradient information-based image quality metric [J]. *IEEE Transactions on Consumer Electronics*, 2010, 56(2): 930-936.
- [34] ZIAEI NAFCHI H, SHAHKOLAEI A, HEDJAM R, *et al.* Mean deviation similarity index: efficient and reliable full-reference image quality evaluator[J]. *IEEE Access*, 2016, 4: 5579-5590.
- [35] XIANG Z F, HU D F, XIONG H C, *et al.* Benefit of chemotherapy in stage III nasopharyngeal carcinoma: Analysis of the surveillance, epidemiology, and end results database[J]. *Oral Oncology*, 2021, 117: 105284.
- [36] NODA H, NIIMI M, KOREKUNI J. Simple and efficient colorization in YCbCr color space [C]. *18th International Conference on Pattern Recognition (ICPR'06). August 20-24, 2006, Hong Kong, China.* IEEE, 2006: 685-688.
- [37] GEUSEBROEK J M, VAN DEN BOOMGAARD R, SMEULDERS A W M, *et al.* Color invariance [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(12): 1338-1350.

- [38] KWOK S H. Efficient gamut clipping for color image processing using LHS and YIQ [J]. *Optical Engineering*, 2003, 42(3): 701.
- [39] WEN X, PAN Z X, HU Y X, *et al.* Generative adversarial learning in YUV color space for thin cloud removal on satellite imagery [J]. *Remote Sensing*, 2021, 13(6): 1079.
- [40] YANG S D, WU T H, SHI S W, *et al.* MANIQA: multi-dimension attention network for No-reference image quality assessment [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. June 19-20, 2022, *New Orleans, LA, USA*. IEEE, 2022: 1190-1199.
- [41] ASLAM M A, XU W, AHMED N, *et al.* VRL-IQA: visual representation learning for image quality assessment[J]. *IEEE Access*, 2024, 12: 2458-2473.
- [42] 万丙辰, 张选德. 基于元学习和失真感知的图像质量评价[J]. *液晶与显示*, 2024, 39(11): 1519-1531.
- WAN B CH, ZHANG X D. Image quality assessment based on meta-learning and distortion perception [J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(11): 1519-1531. (in Chinese)

作者简介:



史晨阳(1991—),男,安徽寿县人,博士,讲师,2021年于复旦大学获得博士学位,主要从事图像质量评价、机器视觉及光学系统设计方面的研究。E-mail: shichenyang@ahpu.edu.cn



吴俊杰(2002—),男,安徽宣城人,硕士研究生,主要从事图像质量评价方面的研究。E-mail:2241611103@stu.ahpu.edu.cn