



基于图像色貌和梯度特征的图像质量客观评价

史晨阳 林燕丹

Objective image quality assessment based on image color appearance and gradient features

Shi Chen-Yang Lin Yan-Dan

引用信息 Citation: *Acta Physica Sinica*, 69, 228701 (2020) DOI: 10.7498/aps.69.20200753

在线阅读 View online: <https://doi.org/10.7498/aps.69.20200753>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

基于图像内容对比感知的图像质量客观评价

Objective assessment of image quality based on image content contrast perception

物理学报. 2020, 69(14): 148702 <https://doi.org/10.7498/aps.69.20200335>

基于图像内容视觉感知的图像质量客观评价方法

Objective assessment method of image quality based on visual perception of image content

物理学报. 2018, 67(10): 108702 <https://doi.org/10.7498/aps.67.20180168>

基于计算全息和 θ 调制的彩色图像加密方法

Color image encryption method based on computer generated hologram and θ modulation

物理学报. 2019, 68(11): 110502 <https://doi.org/10.7498/aps.68.20182264>

基于光照-反射成像模型和形态学操作的多谱段图像增强算法

Multispectral image enhancement based on illuminance-reflection imaging model and morphology operation

物理学报. 2018, 67(21): 210701 <https://doi.org/10.7498/aps.67.20181288>

天光背景下混浊大气中成像质量的分析方法

Image quality analysis method under background radiation in turbid atmosphere

物理学报. 2018, 67(8): 088701 <https://doi.org/10.7498/aps.67.20172625>

基于多引导滤波的图像增强算法

Image enhancement based on multi-guided filtering

物理学报. 2018, 67(23): 238701 <https://doi.org/10.7498/aps.67.20181425>

基于图像色貌和梯度特征的图像质量客观评价*

史晨阳¹⁾ 林燕丹^{1)2)†}

1) (复旦大学信息科学与工程学院, 光源与照明工程系, 上海 200433)

2) (复旦大学工程与应用技术研究院, 超越照明研究所, 上海 200433)

(2020年5月19日收到; 2020年7月12日收到修改稿)

图像质量评价 (IQA) 方法需要考虑如何从主观视觉度量结果出发, 设计出符合该结果的客观图像质量评价方法, 应用到相关实际问题中. 本文从视觉感知特性出发, 量化色度和结构特征信息, 提出了基于色貌和梯度两个图像特征的图像质量客观评价模型. 两个色貌新指标 (vividness 和 depth) 是色度特征信息提取算子; 梯度算子用来提取结构特征信息. 其中, vividness 相似图一方面作为特征提取算子计算失真图像局部质量分数, 另一方面作为图像全局权重系数反应每个像素的重要程度. 为了量化所提模型的主要参数, 根据通用模型性能评价指标, 使用 Taguchi 实验设计方法进行优化. 为了验证该模型的性能, 使用 4 个常用图像质量数据库中的 94 幅参考图像和 4830 幅失真图像进行对比测试, 从预测精度、计算复杂度和泛化性进行分析. 结果表明, 所提模型的精度 PLCC 值在 4 个数据库中最低实现 0.8455, 最高可以达到 0.9640, 综合性能优于 10 个典型和近期发表的图像质量评估 (IQA) 模型. 研究结果表明, 所提模型是有效的、可行的, 是一个性能优异的 IQA 模型.

关键词: 图像质量评价, 人类视觉特性, 色貌, 梯度**PACS:** 87.57.C-, 07.05.Pj, 42.30.-d**DOI:** 10.7498/aps.69.20200753

1 引言

随着各种多媒体通信系统中彩色图像内容和成像设备的迅速发展, 传统的灰度传输方式被彩色传输所取代. 在这种转变下, 感知图像质量评估 (image quality assessment, IQA) 在众多的视觉数据处理应用中发挥了重要的作用, 特别是针对由于压缩和传输而产生的失真评价, 如模糊、块效应和白噪声等^[1]. 在过去的几十年里, 人们设计了多种客观的 IQA 模型来评估图像质量^[2]. 人眼视觉系统 (human visual system, HVS) 是一种理想的视觉信息接收系统, 也是用主观判断^[3]评价图像质量最可靠的方法. 通常, 主观评价的结果由平均意见分数 (mean opinion scores, MOS) 和差异意见分

数 (differential mean opinion scores, DMOS) 表示, 且这两种主观评分是图像质量客观评价性能的评估依据.

目前, 典型的图像客观质量评价方法主要有: 1) 基于数学统计的方法, 如 MSE (mean squared error) 和 PSNR (peak signal to noise ratio) 等; 2) 基于 HVS 感知特性的方法, 如 SSIM (structural similarity), IFC (information fidelity criterion), FSIM (feature similarity) 等; 3) 基于机器学习的方法^[4], 该方法计算精度较高, 但需要长时间样本训练, 且计算结果过度依赖训练样本, 难以应用到实际问题中. 上述基于数学统计的方法拥有较好的运行效率, 但准确性较基于 HVS 感知特性的方法存在不足. 在基于 HVS 感知特性的方法中, 具有代表性的 SSIM^[5] 方法假设人类对于一个场景的视

* 国家重点研发计划 (批准号: 2017YFB0403700) 资助的课题.

† 通信作者. E-mail: ydlin@fudan.edu.cn

觉感知可以用提取结构信息进行表征,其主要提取信息包括亮度、对比度和结构.该方法因其设计思路的新颖性和特征信息的计算方式,为后续相关研究提供很多启发和思路.这种基于 SSIM 计算方式的图像质量评价模型也被叫做全参考图像质量评价模型. Sheikh 和 Bovik^[6] 提出 IFC^[7] 及其衍生方法 VIF (visual information fidelity), 这两个方法通过原始图像和失真图像的交互信息保真度进行比较和计算,虽然其性能较 SSIM 有所提升,但计算效率较差,难以广泛应用.基于 SSIM 模型,其他研究者又提出很多扩展模型.其中具有代表性的模型有:加入 HVS 多尺度特性的 MS-SSIM(multi-scale SSIM) 方法^[8];根据图像全局信息分布,通过信息加权的方式得到的 IW-SSIM(information content weighted SSIM) 方法^[9]. Larson 和 Chandler^[10] 基于最明显失真统计提出 MAD(most apparent distortion algorithm) 方法,该方法的不足是计算复杂性较高. Zhang 等^[11,12] 针对人类视觉系统对图像的理解主要依赖于图像的底层特征的特点,提出了两种新的基于特征相似度的全参考图像质量评价指标 RFSIM (riesz transforms based feature similarity index) 和 FSIM (feature similarity). 在 FSIM 中,作者针对灰度图像将相位一致性 (phase congruency) 和梯度幅值 (gradient magnitude) 结合,并使用权重系数获得单位质量分数;此外,为解决彩色图像评价问题,加入色度算子,引申出 FSIMc 模型.基于梯度结果, Liu 等^[13] 提出 GSM (gradient similarity metric) 方法.上述提到的 RFSIM, FSIMc 和 GSM 方法的精度较好,但计算泛化性能欠佳.近年,基于对比度和视觉显著性, Jia 等^[14] 提出 CVSS (contrast and visual saliency similarity) 方法,该方法评价灰度图像的性能较好,但是针对日常生活用的彩色图像,性能仍有不足. Yao 等^[15] 基于图像内容对比度感知特性,提出 MPCC (IQA model based on the perception of the contrast of image contents) 方法.

一般来说,上述较好的模型在空间和频率域描述了结构信息、亮度信息、对比度信息和颜色信息.通过上述分析,可以发现目前的 IQA 模型在设计方法、复杂性和泛化性能仍有不足.所以,本文选择基于 HVS 感知特性构建图像质量客观评价模型,以基本物理理论为基础,构建具有较高的精

度、泛化性和较低计算复杂度模型.目前常见的全参考模型在进行特征相似计算时(如前文中提到的 IW-SSIM, RFSIM, FSIMc, GSM, CVSS 等模型),多参照 SSIM 的计算方式,本文也是参考 SSIM 进行相似计算全参考图像质量评价模型.

通常情况下,彩色图像的质量可以通过在单个 RGB 通道上使用灰度模型,然后连接通道分数来评估.由于这种方法的精度有限,因此需要提出更精细的利用颜色感知特性的模型.此外,彩色图像的质量可以用类似于比色法的方法来评估.均匀的颜色空间(如 CIELAB)可以描述成对颜色之间感知到的颜色差异^[16,17]. CIELAB 中的圆柱极坐标与刺激的颜色属性相关.因此,利用该公式可以更好地在像素级表征彩色图像的感知失真.虽然这些公式对比较均匀色块很有用,但需要通过优化,与复杂的真实图像数据保持一致.随着人眼对色彩表象的感知能力的不断提高,一些新的模型被提出,并利用相似计算公式对这些模型进行修正. Lee 和 Plataniotis^[18,19] 提出了基于正常三通道色貌指标的 IQA 方法(即亮度、色调和色度),从而获得较好的彩色图像评价性能.因为色貌指标可以代表彩色图像在 HVS 感知方式下的颜色偏差,所以为了获得与主观评价更高的一致性,可以考虑选择色貌指标评价图像的色彩信息.

前文已提到多种图像特征信息,为了获得图像的质量分数,需要选择合适的特征.本文提出了一种新的基于相似性的 IQA 模型,该模型主要提取图像的色彩信息和结构信息两个部分.结合 HVS 对于色彩感知和图像轮廓特征的机理,选择色貌和梯度两个指标提取上述两个特征信息,并量化得到对应的图像质量分数.目前,彩色图像已应用于很多领域,所以色彩信息需要在图像质量评价时重点考虑.色貌新指标 (vividness 和 depth)^[20] 考虑色度和亮度的同步变化,更符合人眼视觉感知的机理,可以体现图像质量的色彩特征变化.除此之外,选择 vividness 指标作为所提模型的权重系数,进一步突出所提模型对于色彩信息的重视.梯度是目前常用的图像结构特征的提取算子,作为补充特征,可以完善所提模型对于图像质量的评价.通过仿真测试,与其他典型的模型相比,所提模型具有更低的复杂度和更好的预测效果.

2 彩色图像质量评价模型

本节提出基于色貌和梯度相似 (color appearance and gradient similarity, CAGS) 方法来量化彩色图像的感知视觉质量. 该模型是具有通用性, 对于常见的彩色和灰度图像始终表现良好的评价效果. 该模型的输入图像是两个具有相同空间分辨率的 RGB 图像 (X 和 Y), 分别称为原始图像和失真图像. 同时, 它们在尺寸上匹配并且正确对齐. 输出质量评分记为 $CAGS(X, Y)$, 其范围在 0 到 1 之间 (当两幅图像完全相同时, 1 表示最佳质量评分).

该模型包含两个相似图, 一个是利用 CIELAB 颜色空间中的色貌相似图来测量颜色失真, 色貌分布图可以在像素级上表示两幅图像之间的颜色差异, 与 HVS 感知更加兼容^[20]. 对于结构失真, 由于梯度具有的优越性^[12,13], 所以可以利用梯度推导出另一个相似图. 最后, 基于文献^[12,21]提到的权重系数方法, 将上述两个相似图进行合并和池化.

2.1 色貌相似图

为了更好地表征 HVS 感知到的图像色貌, 可以将原始的 RGB 图像转化为更符合人类视觉的颜色空间, 使用 CIELAB 色空间. 由于 CIELAB 色空间和孟塞尔系统之间紧密联系, 同时分离成亮度 (L^*) 和色度 (a^* , b^* 或 C_{ab}^* , H_{ab}^*), 限制了 CIELAB 在实际应用中描述颜色在明度和色度共变时的有效性, 例如在艺术、设计、比较着色剂效果和彩色图像评估等领域. 因此, 两个新的 CIELAB 色貌指标

被提出, 即 vividness 和 depth, 来扩展 CIELAB 作为色貌变量的有效性, 如图 1 所示^[20]. 从图 1 中可以发现, 为了区分两个颜色, 需要对两个颜色的对应参数进行比较, 上述两个色貌指标除了可以表征亮度和色度特性以外, 还可以作为色貌坐标, 计算颜色差别. 经过色空间从 RGB 到 CIELAB 变换后, X 的每个像素包含三个颜色分量: 亮度通道 L^* 、红绿通道 a^* 和蓝黄通道 b^* , 并通过 (1) 式和 (2) 式计算得到 vividness (V_{ab}^*) 图和 depth (D_{ab}^*) 图.

$$V_{ab}^* = \sqrt{(L^*)^2 + (a^*)^2 + (b^*)^2} = \sqrt{(L^*)^2 + (C_{ab}^*)^2}, \quad (1)$$

$$D_{ab}^* = \sqrt{(100 - L^*)^2 + (a^*)^2 + (b^*)^2} = \sqrt{(100 - L^*)^2 + (C_{ab}^*)^2}. \quad (2)$$

为了量化颜色分量中的视觉失真, 使用了两个色度描述符: vividness 相似图 S_V 和 depth 相似图 S_D . 根据 (1) 式和 (2) 式, 原始图像 X 和失真图像 Y 的局部色貌分布图由 V_1 和 D_1 , V_2 和 D_2 表示. 图 2 是一个选自 LIVE 数据库的例子, 图 2(a) 为参考图像, 图 2(b) 为高斯模糊失真后的图像. 图 2(c)、图 2(d)、图 2(e) 和图 2(f) 分别是对应参考图像和失真图像的 V_1 , V_2 , D_1 和 D_2 图像. 然后, 根据 SSIM 相似计算的方式将被比较的两幅图像的色貌相似图定义为

$$S_V = \frac{2V_1 \cdot V_2 + K_V}{V_1^2 + V_2^2 + K_V}, \quad (3)$$

$$S_D = \frac{2D_1 \cdot D_2 + K_D}{D_1^2 + D_2^2 + K_D}, \quad (4)$$

其中 K_V 和 K_D 都是常量, 增加模型的稳定性, 保证分母不为 0^[5]. 将 vividness 相似图 S_V 和 depth 相似图 S_D 通过简单幂函数相乘的方式结合, 可以得到原始图像 X 和失真图像 Y 的色貌相似图 S_{CA} , 如下式所示:

$$S_{CA} = (S_V)^\alpha \cdot S_D, \quad (5)$$

其中 α 表征两个色貌指标在色貌相似计算中的相对权重, 区间是 $[0, 1]$. 较大的 α 值, 表示 vividness 在色貌相似计算中的相对权重越高. 为了说明色貌指标的有效性, 图 2(g) 是连接 vividness 相似图和 depth 相似图的色貌相似图. 图 2(g) 中颜色较深的点, 表征该区域的图像色彩差别越大. 从图 2(g) 中可以清楚地看到, 色貌相似图可以有效表征失真图像与参考图像之间的差别.

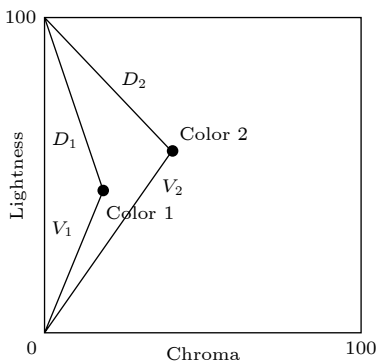


图 1 颜色 1 和 2 的 vividness (V_{ab}^*) 和 depth (D_{ab}^*) 维度表征, 线段长度定义对应属性^[20]

Fig. 1. Dimensions of vividness, V_{ab}^* , and depth, D_{ab}^* for colors 1 and 2. Line lengths define each attribute^[20].

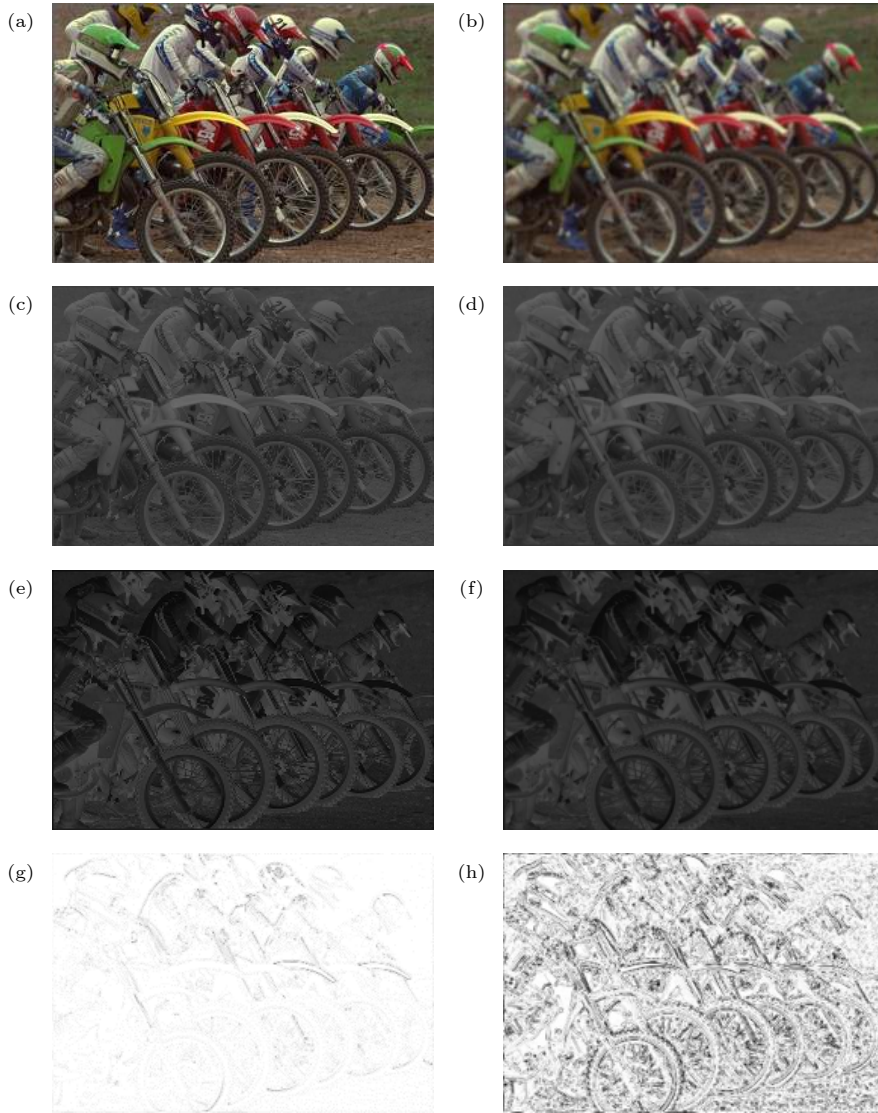


图 2 从 LIVE 数据库中提取的典型图像 (a) 为参考图像; (b) 为高斯模糊畸变类型的失真图像; (c) 和 (e) 分别是参考图像的 Vividness 和 Depth 图; (d) 和 (f) 分别是失真图像的 Vividness 和 Depth 图; (g) 是色貌相似图; (h) 为梯度相似图

Fig. 2. Typical images extracted from LIVE: (a) The reference image; (b) the distorted vision of it by Gaussian blur distortion type; (c) and (e) are the vividness and depth map of the reference image, respectively; (d) and (f) are the vividness and depth map of the distorted image; (g) the color appearance similarity map by connecting the vividness and depth similarity map; (h) the gradient similarity map.

2.2 梯度相似图

目前, 有几种不同的算子可以计算图像梯度, 如 Prewitt 算子^[22]、Sobel 算子^[22]、Roberts 算子^[23]和 Scharr 算子^[23]. 离散域梯度幅值的计算通常是基于上述算子, 并利用差值表示图像函数. 图像 X 的垂直梯度 G_y 和水平梯度 G_x 计算方式是利用卷积得到: $G_x = h_x * X$ 和 $G_y = h_y * X$ (参见 (6) 式和 (7) 式), 其中 h_x 和 h_y 分别为水平和垂直梯度算子, $*$ 表示卷积. 则梯度的定义是 $G(x) = \sqrt{G_x^2 + G_y^2}$. 在提出的 IQA 模型中, 这些计算方式

的执行效果是相同的. 梯度表征结构信息的原理是通过梯度算子的卷积处理图像, 保留包含显著边缘区域的信息, 去除平滑区域的信息.

$$G_x = \frac{1}{16} \begin{bmatrix} 3 & 0 & -3 \\ 10 & 0 & -10 \\ 3 & 0 & -3 \end{bmatrix} * X, \quad (6)$$

$$G_y = \frac{1}{16} \begin{bmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix} * X. \quad (7)$$

在本文中, 经过前期梯度算子的比较, 发现 Scharr 在本文模型中可以更大程度地提取结构信

息, 获得较好的结果. 因此, Scharr 算子被选择用来计算参考图像和失真图像 CIELAB 色空间亮度 (L^*) 通道的梯度 (G_1 和 G_2). 通过这两个梯度, 可以得到两个图像的梯度相似图 S_G , 其计算公式如下所示:

$$S_G = \frac{2G_1 \cdot G_2 + K_G}{G_1^2 + G_2^2 + K_G}, \quad (8)$$

其中参数 K_G 为控制数值稳定性的常数. 梯度相似图 (S_G) 在文献 [24–26] 中被广泛应用, 并研究了它在图像失真测量中的作用. 如图 2(h) 所示, 图中颜色较深的点, 表征该区域的图像结构差别越大. 因此, 梯度相似图是对 HVS 结构信息的一种有效表征.

2.3 CAGS 方法

根据前文提取的色貌相似和梯度相似图, 在 IQA 任务中定义一个新的模型——色貌和梯度相似指标 (color appearance and gradient Similarity index, CAGS), 其程序代码请参见 <https://github.com/AlAlien/CAGS>, 计算公式如下所示:

$$\text{CAGS} = \frac{\sum_{\Omega} S_G \cdot S_{CA} \cdot V_m}{\sum_{\Omega} V_m}, \quad (9)$$

其中 Ω 表示空间域; V_m ($V_m = \max(V_1, V_2)$) 用于表征两个图像的局部重要性权重.

人们普遍认为, 不同的位置对 HVS 图像质量

的视觉感知有不同的贡献, 因此, 在得到最终质量分数时可以考虑视觉系统的注视点. 由于人类视觉皮层对亮度和色度很敏感, 所以某个位置的 vividness 值可以反映出它是一个可感知的重要点的权重大小. 由于上述原因, 在提出的模型中选择色貌分布图 (V_m) 来描述局部区域的视觉重要性. CAGS 的计算过程如图 3 所示. 本文所提模型, 以色貌特征为基础, 结合梯度特征, 表征图像质量. 色貌特征处理图像颜色信息差别, 梯度特征计算图像结构信息不同, 通过颜色信息和结构信息的综合计算, 得到图像的客观评分.

在本文中, K_V , K_D , K_G 均为常数, 可以方便地应用于所有数据库. 此外, α 为了应用于所有数据库也需要确定. 在以往的研究中, 试错法是处理这类问题的一种常用方法, 但在时间和消耗上都不现实. 近年来, Taguchi 等 [27] 实验设计方法被广泛应用于解决多参数寻优问题. 后续, 将使用 Taguchi 方法优化模型中的相关参数.

3 实验结果

3.1 性能评价标准和参考数据库

在我们的研究中, 选择 4 个公共数据库进行模型验证和比较, 即 TID2013[28], CSIQ[29], LIVE[30] 和 IVC[31], 表 1 列出了上述每个数据库的基本信息. 这 4 个数据库是 IQA 研究中最常用的集合, 涵

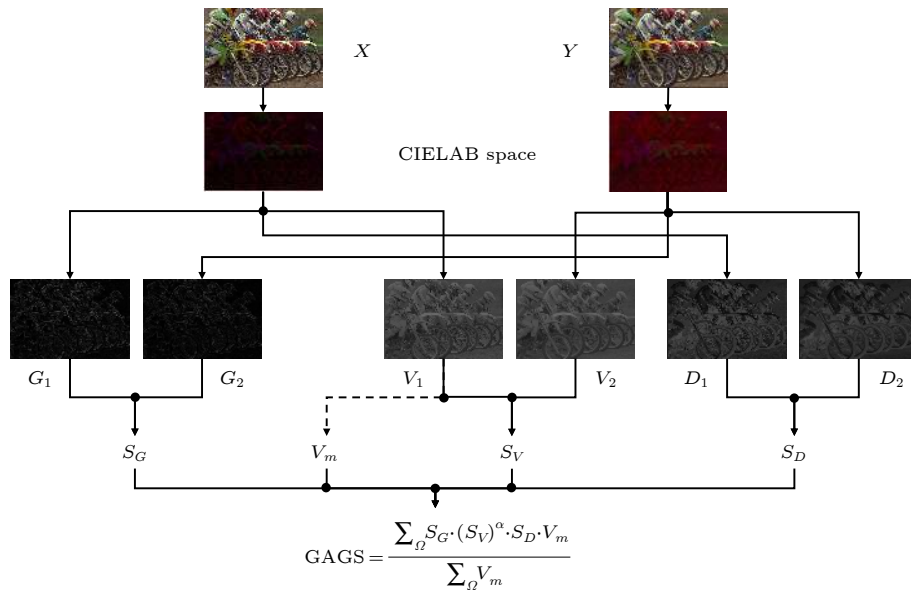


图 3 本文提出的 IQA 模型 CAGS 的计算过程

Fig. 3. Illustration for the computational process of the proposed IQA model CAGS.

表 1 IQA 数据库基本信息
Table 1. Benchmark test databases for IQA.

数据库	原始图像数量	失真图像数量	失真类型	观察者
TID2013	25	3000	24	971
CSIQ	30	866	6	35
LIVE	29	779	5	161
IVC	10	185	4	15

盖了现实应用中常见的各种失真. 它们用主观的评价来表征图像质量 (MOS 或 DMOS), 可以作为设计的模型进行合理的基准测试. 上述数据库中的失真图像是从一组原始图像处理得到的, 这些原始图像反映了足够的颜色复杂性和边缘/纹理细节的多样性, 包括人类、自然场景和人造物体的图片.

为了评估一个模型是否能够预测人类观察者的感知, 将使用客观评价模型计算出的图像质量分数与观察者所评定的值进行拟合. IQA 模型常用的四个评价标准是: 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)^[2,32].

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

$$p(s) = \beta_1 \left[\frac{1}{2} - \frac{1}{1 + \exp(\beta_2(s - \beta_3))} \right] + \beta_4 s + \beta_5, \quad (10)$$

其中 β_1, \dots, β_5 都是拟合参数, s 代表原始 IQA 分数, $p(s)$ 是回归后的 IQA 分数^[30].

3.2 参数优化

Taguchi 方法是 Taguchi 和 Konishi 共同开发

表 3 采用 $L_9(3^4)$ 直交表的实验设计及 IVC 数据库测试结果
Table 3. Design with a $L_9(3^4)$ orthogonal array for IVC database.

实验序号	A	B	C	D	SROCC	SROCC的S/N	RMSE	RMSE的S/N
1	1	1	1	1	0.9300	-0.6303	0.4113	7.7168
2	1	2	2	2	0.9192	-0.7318	0.4533	6.8723
3	1	3	3	3	0.9096	-0.8230	0.4825	6.3301
4	2	1	2	3	0.9171	-0.7517	0.4596	6.7524
5	2	2	3	1	0.9173	-0.7498	0.4672	6.6099
6	2	3	1	2	0.9291	-0.6388	0.4142	7.6558
7	3	1	3	2	0.9114	-0.8058	0.4735	6.4936
8	3	2	1	3	0.9279	-0.6500	0.4174	7.5890
9	3	3	2	1	0.9195	-0.7290	0.4481	6.9725

的一种有效的实验设计工具^[33]. 根据前文的分析, 为了适用于所有数据库, K_V, K_D, K_G 和 α 是 CAGS 模型优化的主要变量参数. 根据相关研究, 将各影响因素的水准数及各水准的值选取在适当的范围内, 如表 2 所列. 对于每个参数, 都选择了 3 个值. 两个色貌相似度对应参数应该确定相同的水准值, 这些值的选择基于文献^[34]. 为了得到梯度相似图的最佳相关系数, 可以在 $[0, 100]$ 中设置 K_G ^[26], 因此定义水准值为 10, 50 和 100. 相对权重参数 α , 应该限制在 $[0, 1]$, 所以选择 0.1, 0.5, 1 作为水准值.

选取参数和水准后, 在直交表中进行组合实验, 如表 3 所列. 采用试错法进行实验需要 3^4 次, 而利用 Taguchi 方法设计实验只需进行 9 次即可, 大大减少了实验时间. 直交表配置为 $L_9(3^4)$ 型. 选择 IVC 数据库作为优化参数的参考集, 表征量化质量特征的 SROCC, RMSE 值及其 S/N (signal to noise) 值均包含在表 3 中. 较大的 S/N 值表征其对应的水准具有较好评价效果. 对于 SROCC 来说, 性能具有望大性, 即数值越大表征效果越好. 因此, SROCC 的 S/N 值 (ε_1) 可以根据 (11) 式计算. 相反, RMSE 应该越少越好. 因此, RMSE 的 S/N 值 (ε_2) 可以根据 (12) 式计算.

表 2 变量参数及其控制水准
Table 2. Influence factors and level setting for CAGS.

代号	参数表述	水准数	水准一	水准二	水准三
A	K_V	3	0.002	0.02	0.2
B	K_D	3	0.002	0.02	0.2
C	K_G	3	10	50	100
D	α	3	0.1	0.5	1

$$\varepsilon_1 = -10 \log \frac{\sum_{i=1}^t (1/q_i^2)}{t}, \quad (11)$$

$$\varepsilon_2 = -10 \log \frac{\sum_{i=1}^t (q_i^2)}{t}, \quad (12)$$

其中 q_i 表示与第 i 次试验对应的 SROCC 和 RMSE 的值; t 为每次试验的总数.

特定的参数组合对 SROCC 和 RMSE 两个评价标准的影响可以通过计算该特定参数水准三次实验相应的平均 S/N 值来评估. 不同水准值对应的 S/N 值可以在图 4 中清晰地识别出来, 其值越高表明该水准对应的评价效果越好.

从图 4 可以得到下述直观合理的结果. 例如, 较小的参数 C 值, 可以获得较大的 S/N 值; 同时因其对应的 S/N 值变化幅度最大, 表明其影响程度也是最大的. 对于其他 3 个参数, 最佳组合为 A2-B2-D1. 对于参数 C , 由于其影响程度最大, 需

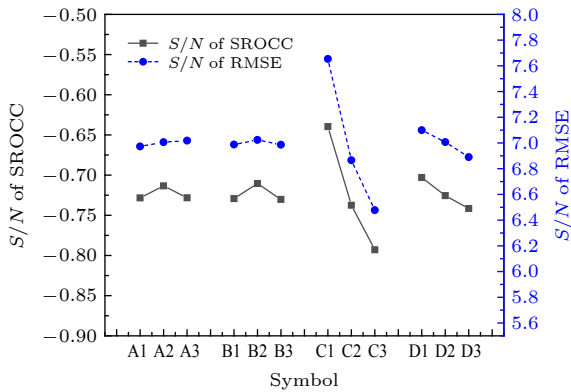


图 4 SROCC 和 RMSE 对应的不同水准的 S/N 值
Fig. 4. S/N ratio of different levels for SROCC and RMSE.

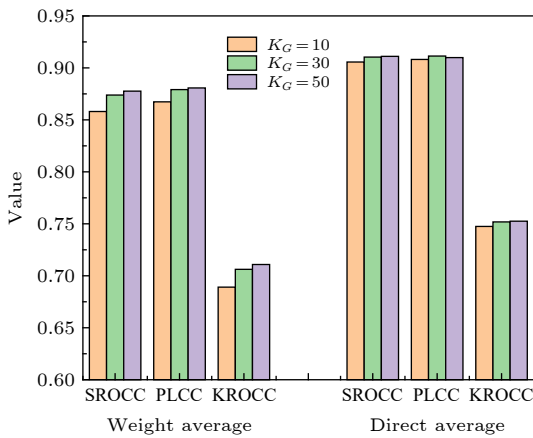


图 5 不同 K_G 值对模型性能的影响
Fig. 5. Performance of different K_G values.

要将其进一步调整为更合适的值, 以达到更好的拟合程度.

使用本文提出的 CAGS 方法, 通过计算四种数据库在不同 K_G 下 SROCC, PLCC 和 KROCC 的加权平均值 (根据每个数据库中失真图像的数量分配不同的权重) 和直接平均值, 确定 K_G 值, 如图 5 所示. 从图 5 中可获得最佳的 K_G 值, 所以最终 CAGS 模型的最佳参数是 $K_V = K_D = 0.02$, $K_G = 50$, $\alpha = 0.1$.

4 讨论

4.1 IQA 模型的性能对比

一个理想的 IQA 模型应该具有良好的性能, 并且在不同类型的失真情况下具有良好的一致性. 本节将所提出的模型与其他典型的模型进行了比较, 包括 SSIM^[5], IW-SSIM^[9], IFC^[6], VIF^[7], MAD^[10], RFSIM^[11], FSIMc^[12] 和 GSM^[13], 以及近年提出的方法 CVSS(2018)^[14] 和 MPCC(2020)^[15]. 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, 根据 Wang 和 Li^[9], 4 个数据库的 SROCC, PLCC 和 KROCC 结果的加权和直接平均值也如表 4 所列, 以评估整体性能. 每个数据库的权值是根据数据库中包含的失真图像的数量确定的.

从表 4 可以看出, 我们提出的模型对所有数据库的性能都较好. 特别是, CAGS 模型在 LIVE 和 IVC 数据库的排名前三名. 对于 TID2013, CAGS 模型具有竞争性, 与前三的差距很小. 对于 CSIQ, CAGS 模型仅比前三模型表现稍差. 虽然 MAD, CVSS 和 MPCC 可以从 CSIQ 中得到最好的结果, 但是它们的性能对其他三个数据库不如我们提出的模型的性能. 此外, CAGS 模型还是三个指标加权和直接平均值的前三名. 综上, 获得前三数量最多的模型是 FSIM (16 次), CAGS (15 次) 和 CVSS (14 次). 而且, 对于所有数据库的 SROCC 和 PLCC 值均大于 0.8316, 所以可以认为 CAGS 模型既具有较高的性能, 又具有较好的泛化性. 特别的, 由于 MPCC 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 MPCC 相关数据. 但与 MPCC 模型性能的比较中可以发现, CAGS 模型与 MPCC 模型的性能差距较小, 与当前最新的研究结果相比具有竞争性.

表 4 对比不同 IQA 模型的 4 个数据库性能
Table 4. Performance comparison of IQA models on four databases.

数据库		SSIM	IW-SSIM	IFC	VIF	MAD	RFSIM	FSIM _C	GSM	CVSS	MPCC	Proposed
TID2013	SROCC	0.7417	0.7779	0.5389	0.6769	0.7807	0.7744	0.8510	0.7946	0.8069	0.8452	0.8316
	PLCC	0.7895	0.8319	0.5538	0.7720	0.8267	0.8333	0.8769	0.8464	0.8406	0.8616	0.8445
	RMSE	0.7608	0.6880	1.0322	0.7880	0.6975	0.6852	0.5959	0.6603	0.6715	0.6293	0.6639
	KROCC	0.5588	0.5977	0.3939	0.5147	0.6035	0.5951	0.6665	0.6255	0.6331	—	0.6469
CSIQ	SROCC	0.8756	0.9213	0.7671	0.9195	0.9466	0.9295	0.9310	0.9108	0.9580	0.9569	0.9198
	PLCC	0.8613	0.9144	0.8384	0.9277	0.9502	0.9179	0.9192	0.8964	0.9589	0.9586	0.9014
	RMSE	0.1334	0.1063	0.1431	0.0980	0.0818	0.1042	0.1034	0.1164	0.0745	0.0747	0.1137
	KROCC	0.6907	0.7529	0.5897	0.7537	0.7970	0.7645	0.7690	0.7374	0.8171	—	0.7487
LIVE	SROCC	0.9479	0.9567	0.9259	0.9636	0.9669	0.9401	0.9599	0.9561	0.9672	0.9660	0.9734
	PLCC	0.9449	0.9522	0.9268	0.9604	0.9675	0.9354	0.9503	0.9512	0.9651	0.9622	0.9640
	RMSE	8.9455	8.3473	10.2643	7.6137	6.9073	9.6642	7.1997	8.4327	7.1573	7.4397	8.3251
	KROCC	0.7963	0.8175	0.7579	0.8282	0.8421	0.7816	0.8366	0.8150	0.8406	—	0.8658
IVC	SROCC	0.9018	0.9125	0.8993	0.8964	0.9146	0.8192	0.9293	0.8560	0.8836	—	0.9195
	PLCC	0.9119	0.9231	0.9093	0.9028	0.9210	0.8361	0.9392	0.8662	0.8438	—	0.9298
	RMSE	0.4999	0.4686	0.5069	0.5239	0.4746	0.6684	0.4183	0.6088	0.6538	—	0.4483
	KROCC	0.7223	0.7339	0.7202	0.7158	0.7406	0.6452	0.7636	0.6609	0.6957	—	0.7488
权重平均	SROCC	0.8051	0.8376	0.6560	0.7750	0.8456	0.8306	0.8859	0.8438	0.8628	—	0.8737
	PLCC	0.8321	0.8696	0.6786	0.8353	0.8752	0.8650	0.8987	0.8730	0.8820	—	0.8772
	KROCC	0.6270	0.6662	0.5002	0.6158	0.6819	0.6575	0.7160	0.6775	0.7020	—	0.7044
直接平均	SROCC	0.8668	0.8921	0.7828	0.8641	0.9022	0.8658	0.9178	0.8794	0.9039	—	0.9111
	PLCC	0.8769	0.9054	0.8071	0.8907	0.9164	0.8807	0.9214	0.8901	0.9021	—	0.9099
	KROCC	0.6920	0.7255	0.6154	0.7031	0.7458	0.6966	0.7589	0.7097	0.7466	—	0.7526

为了更好地说明所提出的 IQA 模型在不同类型的失真情况下具有更好的性能, 使用 TID2013 数据库的散点图进行比较, 如图 6 所示. 从图 6 中可以看出, 与其他模型相比, CAGS 模型的客观评分与主观评分之间存在着高度的相关性.

出现上述结果的原因主要有: 1) CAGS 模型结合 HVS 特性, 同时提取色度信息和结构信息, 并利用权重系数表征局部重要性, 所以其较其他模型具有较好的性能和泛化性; 2) 在颜色信息提取方面, 考虑亮度和色度的协同变化, 提高了颜色信息提取的准确性, 更好的表征图像失真程度.

4.2 IQA 模型对不同失真类型的性能对比

一个稳定 IQA 模型需要产生良好的性能, 才能对每种失真类型进行一致的预测. 在本节中, 对每种失真类型的模型性能进行了对比, 结果如表 5 所列. 选择 SROCC 作为评价指标, 因为其他指标 (PLCC, RMSE 和 KROCC) 具有类似的效果. 由于 IVC 数据库中每种失真类型的失真图像的数量

太少, 无法用统计的方式表达结果. 因此, 选取了三个主要数据库中的 35 组失真图像. 对于每个数据库和每种失真类型, 排名前三 SROCC 值的 IQA 模型的结果用粗体突出显示. 进一步, 对于最新 MPCC 模型, 与 CAGS 对比 TID2013 数据库 PLCC 的结果, 如图 7 所示.

对比分析表 5 和图 7 中不同模型对于不同失真类型的失真图像评价结果, 可得: 1) CAGS 模型与其他典型 IQA 模型相比具有明显的统计优势, 总体效果最好. 特别是, CAGS 模型在 TID2013 和 LIVE 数据库的排名前三的数量较多. 而对于 CSIQ, CAGS 模型具有竞争性, 与前三的差距很小. 获得前三数量最多的模型是 CAGS (21 次), CVSS (19 次) 和 FSIM (14 次). 2) 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 CAGS (14 次) 和 CVSS (14 次). 对于上述两个模型, 其最小 SROCC 值分别是 0.3711 和 0.2614, 且 SROCC 值小于 0.5 都是 2 个类型 (CTC 和 CCS). 所以, CAGS 模型

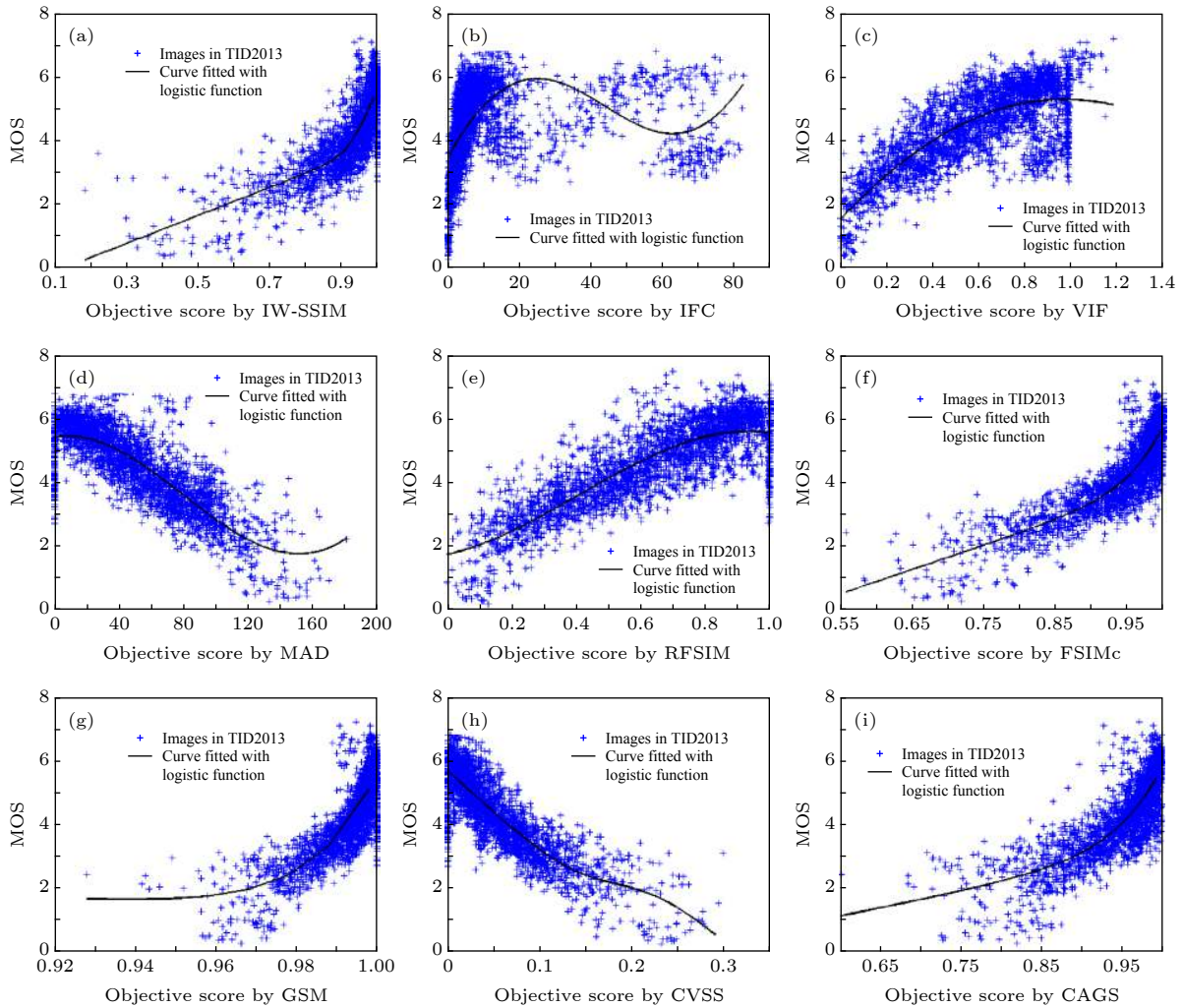


图 6 基于 TID2013 数据库的主观 MOS 与 IQA 模型计算结果拟合对比 (a) IW-SSIM; (b) IFC; (c) VIF; (d) MAD; (e) RFSIM; (f) FSIMc; (g) GSM; (h) CVSS; (i) CAGS

Fig. 6. Scatter plots of subjective MOS against scores calculated by IQA models' prediction for TID2013 databases: (a) IW-SSIM; (b) IFC; (c) VIF; (d) MAD; (e) RFSIM; (f) FSIMc; (g) GSM; (h) CVSS; (i) CAGS.

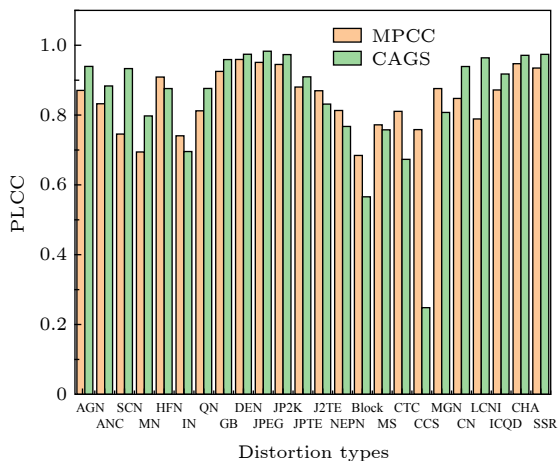


图 7 CAGS 与 MPCC 在 TID2013 数据库种不同失真类型 PLCC 值对比

Fig. 7. PLCC comparison of different distortion types between CAGS and MPCC on TID2013.

的波动程度更小, 性能更加稳定, 具有较好的泛化性. (3) 对比 CAGS 和 MPCC 模型, SROCC 值排名前三次数分别是 21 次和 6 次; SROCC 值大于 0.95 次数分别是 14 次和 9 次; 在 TID2013 数据库不同失真类型的 PLCC 值, 在 24 种失真类型中, CAGS 在 15 种失真类型中具有较高的精度; CAGS 和 MPCC 的最小 SROCC 值分别是 0.3711 和 0.5396; 对比结果表明 CAGS 比最新的 MPCC 模型具有明显的精度优势, 而 MPCC 具有更好的泛化性. 综上可以得出结论, 与其他 IQA 模型相比, CAGS 模型预测的客观评分与主观评价高度相关.

通过以上对于不同失真类型的图像质量评价精度的对比分析, 表明 CAGS 模型在不同数据库中和在不同失真类型中都具有较好的精度和泛化性.

表 5 IQA 模型的不同失真类型 SROCC 值对比
Table 5. SROCC values of IQA models for different types of distortions.

数据库	失真类型	SSIM	IW-SSIM	IFC	VIF	MAD	RFSIM	FSIM _C	GSM	CVSS	MPCC	Proposed
TID2013	AGN	0.8671	0.8438	0.6612	0.8994	0.8843	0.8878	0.9101	0.9064	0.9401	0.8666	0.9359
	ANC	0.7726	0.7515	0.5352	0.8299	0.8019	0.8476	0.8537	0.8175	0.8639	0.8187	0.8653
	SCN	0.8515	0.8167	0.6601	0.8835	0.8911	0.8825	0.8900	0.9158	0.9077	0.7396	0.9276
	MN	0.7767	0.8020	0.6932	0.8450	0.7380	0.8368	0.8094	0.7293	0.7715	0.7032	0.7526
	HFN	0.8634	0.8553	0.7406	0.8972	0.8876	0.9145	0.9094	0.8869	0.9097	0.8957	0.9159
	IN	0.7503	0.7281	0.6208	0.8537	0.2769	0.9062	0.8251	0.7965	0.7457	0.6747	0.8361
	QN	0.8657	0.8468	0.6282	0.7854	0.8514	0.8968	0.8807	0.8841	0.8869	0.7931	0.8718
	GB	0.9668	0.9701	0.8907	0.9650	0.9319	0.9698	0.9551	0.9689	0.9348	0.9218	0.9614
	DEN	0.9254	0.9152	0.7779	0.8911	0.9252	0.9359	0.9330	0.9432	0.9427	0.9510	0.9466
	JPEG	0.9200	0.9187	0.8357	0.9192	0.9217	0.9398	0.9339	0.9284	0.9521	0.8964	0.9585
	JP2 K	0.9468	0.9506	0.9078	0.9516	0.9511	0.9518	0.9589	0.9602	0.9587	0.9160	0.9620
	JPTE	0.8493	0.8388	0.7425	0.8409	0.8283	0.8312	0.8610	0.8512	0.8613	0.8571	0.8644
	J2 TE	0.8828	0.8656	0.7769	0.8761	0.8788	0.9061	0.8919	0.9182	0.8851	0.8409	0.9250
	NEPN	0.7821	0.8011	0.5737	0.7720	0.8315	0.7705	0.7937	0.8130	0.8201	0.7753	0.7833
	Block	0.5720	0.3717	0.2414	0.5306	0.2812	0.0339	0.5532	0.6418	0.5152	0.5396	0.6015
	MS	0.7752	0.7833	0.5522	0.6276	0.6450	0.5547	0.7487	0.7875	0.7150	0.7520	0.7441
	CTC	0.3775	0.4593	0.1798	0.8386	0.1972	0.3989	0.4679	0.4857	0.2940	0.7814	0.4514
	CCS	0.4141	0.4196	0.4029	0.3009	0.0575	0.0204	0.8359	0.3578	0.2614	0.7054	0.3711
	MGN	0.7803	0.7728	0.6143	0.8486	0.8409	0.8464	0.8569	0.8348	0.8799	0.8766	0.8700
	CN	0.8566	0.8762	0.8160	0.8946	0.9064	0.8917	0.9135	0.9124	0.9351	0.8174	0.9168
LCNI	0.9057	0.9037	0.8160	0.9204	0.9443	0.9010	0.9485	0.9563	0.9629	0.8095	0.9574	
ICQD	0.8542	0.8401	0.6006	0.8414	0.8745	0.8959	0.8815	0.8973	0.9108	0.8596	0.9060	
CHA	0.8775	0.8682	0.8210	0.8848	0.8310	0.8990	0.8925	0.8823	0.8523	0.8094	0.8768	
SSR	0.9461	0.9474	0.8885	0.9353	0.9567	0.9326	0.9576	0.9668	0.9605	0.9178	0.9580	
CSIQ	AWGN	0.8974	0.9380	0.8431	0.9575	0.9541	0.9441	0.9359	0.9440	0.9670	0.9329	0.9652
	JPEG	0.9543	0.9662	0.9412	0.9705	0.9615	0.9502	0.9664	0.9632	0.9689	0.9564	0.9573
	JP2 K	0.9605	0.9683	0.9252	0.9672	0.9752	0.9643	0.9704	0.9648	0.9777	0.9630	0.9545
	AGPN	0.8924	0.9059	0.8261	0.9511	0.9570	0.9357	0.9370	0.9387	0.9516	0.9517	0.9492
	GB	0.9608	0.9782	0.9527	0.9745	0.9602	0.9643	0.9729	0.9589	0.9789	0.9664	0.9574
	CTC	0.7925	0.9539	0.4873	0.9345	0.9207	0.9527	0.9438	0.9354	0.9324	0.9399	0.9273
LIVE	JP2 K	0.9614	0.9649	0.9113	0.9696	0.9676	0.9323	0.9724	0.9700	0.9719	0.9608	0.9822
	JPEG	0.9764	0.9808	0.9468	0.9846	0.9764	0.9584	0.9840	0.9778	0.9836	0.9674	0.9836
	AWGN	0.9694	0.9667	0.9382	0.9858	0.9844	0.9799	0.9716	0.9774	0.9809	0.9457	0.9837
	GB	0.9517	0.9720	0.9584	0.9728	0.9465	0.9066	0.9708	0.9518	0.9662	0.9561	0.9641
	FF	0.9556	0.9442	0.9629	0.9650	0.9569	0.9237	0.9519	0.9402	0.9592	0.9627	0.9633

4.3 不同模型间计算复杂性对比

IQA 模型的运行效率也是一个重要的考虑因素. 使用 2.5 GHz Intel Core i5 CPU 和 8 G RAM 的型号电脑比较不同模型的运行时间, 软件平台为 MATLAB R2013b. 表 6 列出了每个模型用于比较一对彩色图像的时间, 分辨率为 512×512 (选

自 IVC 数据库). 由此表 6 结果可知, CGGS 模型具有适中的计算复杂度. 具体来说, 它比一些具有良好预测性能的现代 IQA 模型运行得更快, 如 IW-SSIM, IFC, VIF 和 MAD. 由于 CAGS 模型仅包含色度信息和结构信息两个方面的内容, 即可有效评价图像质量, 同时无需循环计算, 所以 CAGS 模型计算复杂度较低.

表 6 计算复杂度对比
Table 6. Time cost comparisons.

IQA模型	运行时间/s	IQA模型	运行时间/s
PSNR	0.0186	RFSIM	0.1043
SSIM	0.0892	FSIMc	0.3505
IW-SSIM	0.6424	GSM	0.1018
IFC	1.1554	CVSS	0.0558
VIF	1.1825	MPCC	—
MAD	2.7711	CAGS	0.4814

5 结 论

我们提出了一种新颖且性能良好的 IQA 模型, 即 CAGS 模型. 该模型由色貌相似图和梯度相似图组成, 分别表征了颜色失真和结构失真. 其中, 色貌相似图由 vividness 和 depth 两部分组成, 这两部分是表征颜色特征的两个新指标. 同时, 选取 vividness 图作为权重系数, 定义局部图像的重要性. 最后, 根据权重关系将上述两个相似图合并得到所提模型. 模型中的主要参数通过 Taguchi 方法确定, 使用该方法可大大减少实验时间. 为了证明所提出的模型的优异性能, 使用 4 个常用的 IQA 数据库共 94 幅参考图像和 4830 幅失真图像进行测试, 与另外 10 个最典型的或被近期发表的 IQA 模型的结果进行了比较. 结果表明, CAGS 模型的精度 PLCC 在 4 给数据库中最低实现 0.8455, 最高可以达到 0.9640, 综合性能优于其他 IQA 模型. 与其他计算复杂度适中的模型相比, CAGS 模型具有更好的预测精度和泛化性. 未来, 我们将进一步优化模型, 提升预测精度和泛化性能.

参考文献

[1] Yao J C, Liu G Z 2018 *Acta Phys. Sin.* **67** 108702 (in Chinese) [姚军财, 刘贵忠 2018 物理学报 **67** 108702]
 [2] Athar S, Wang Z 2019 *IEEE Access* **7** 140030
 [3] Lin W S, Kuo C C J 2011 *J. Vis. Commun. Image R.* **22** 297
 [4] Chang H W, Zhang Q W, Wu Q G, Gan Y 2015 *Neurocomputing* **151** 1142
 [5] Wang Z, Bovik A C, Sheikh H R, Simoncelli E P 2004 *IEEE Trans. Image Process.* **13** 600
 [6] Sheikh H R, Bovik A C 2006 *IEEE Trans. Image Process.* **15** 430
 [7] Sheikh H R, Bovik A C, de Veciana G 2005 *IEEE Trans.*

Image Process. **14** 2117
 [8] Wang Z, Simoncelli E P, Bovik A C 2003 *37th Asilomar Conference on Signals, Systems and Computer Pacific Grove, CA, November 9–12, 2003* pp1398–1402
 [9] Wang Z, Li Q 2011 *IEEE Trans. Image Process.* **20** 1185
 [10] Larson E C, Chandler D M 2010 *J. Electron. Imaging* **19** 011006
 [11] Zhang L, Zhang L, Mou X Q 2010 *IEEE International Conference on Image Processing Hong Kong, Peoples of China, September 26–29, 2010* pp321–324
 [12] Zhang L, Zhang L, Mou X Q, Zhang D 2011 *IEEE Trans. Image Process.* **20** 2378
 [13] Liu A M, Lin W S, Narwaria M 2012 *IEEE Trans. Image Process.* **21** 1500
 [14] Jia H Z, Zhang L, Wang T H 2018 *IEEE Access* **6** 65885
 [15] Yao J C, Shen J 2020 *Acta Phys. Sin.* **69** 148702 (in Chinese) [姚军财, 申静 2020 物理学报 **69** 148702]
 [16] Robertson A R 1990 *Color Res. Appl.* **15** 167
 [17] Mahny M, Vaneycken L, Oosterlinck A 1994 *Color Res. Appl.* **19** 105
 [18] Lee D, Plataniotis K N 2015 *IEEE Trans. Image Process.* **24** 3950
 [19] Lee D, Plataniotis K N 2014 *2014 International Conference on Acoustics, Speech and Signal Processing (ICASSP) Florence, Italy, May 4–9, 2014* pp166–170
 [20] Berns R S 2014 *Color Res. Appl.* **39** 322
 [21] Zhang L, Shen Y, Li H Y 2014 *IEEE Trans. Image Process.* **23** 4270
 [22] Jain R C, Kasturi R, Schunck B G 1995 *Machine Vision (New York: McGraw-Hill)* pp140–185
 [23] Sonka M, Hlavac V, Boyle R 2008 *Image Processing, Analysis and Machine Vision (3rd Ed.) (Stanford: Cengage Learning)* p77
 [24] Xue W F, Zhang L, Mou X Q, Bovik A C 2014 *IEEE Trans. Image Process.* **23** 684
 [25] Kim D O, Han H S, Park R H 2010 *IEEE Trans. Consum. Electr.* **56** 930
 [26] Nafchi H Z, Shahkolaei A, Hedjam R, Cheriet M 2016 *IEEE Access* **4** 5579
 [27] Taguchi G, Yokoyama Y, Wu Y 1993 *Taguchi Methods, Design of experiments (Dearborn, MI: ASI Press)* pp59–63
 [28] Ponomarenko N, Jin L, Ieremeiev O, Lukin V, Egiazarian K, Astola J, Voze B, Chehdi K, Carli M, Battisti F, Kuo C C J 2015 *Signal Process. Image Commun.* **30** 57
 [29] Larson E C, Chandler D M <http://vision.eng.shizuoka.ac.jp/mod/page/view.php?id=23>[2020-7-13]
 [30] Sheikh H R, Sabir M F, Bovik A C 2006 *IEEE Trans. Image Process.* **15** 3440
 [31] Ninassi A, Le Callet P, Autrusseau F 2006 *Conference on Human Vision and Electronic Imaging XI San Jose, CA, USA, January 16–18, 2006* p1
 [32] Wang S Q, Gu K, Zeng K, Wang Z, Lin W S 2016 *Proc. SPIE 6057 Comput. Graph. Appl.* **38** 47
 [33] Lin C H, Wu C C, Yang P H, Kuo T Y 2009 *J. Disp. Technol.* **5** 323
 [34] Preiss J, Fernandes F, Urban P 2014 *IEEE Trans. Image Process.* **23** 1366

Objective image quality assessment based on image color appearance and gradient features*

Shi Chen-Yang¹⁾ Lin Yan-Dan^{1)2)†}

1) (*Department of Light Sources and Illuminating Engineering, School of Information Science and Technology, Fudan University, Shanghai 200433, China*)

2) (*Institute of Future Lighting, Academy for Engineering & Technology, Fudan university, Shanghai 200433, China*)

(Received 19 May 2020; revised manuscript received 12 July 2020)

Abstract

With the rapid development of color image contents and imaging devices in various kinds of multimedia communication systems, conventional grayscale counterparts are replaced by chromatic ones. Under such a transition, the image quality assessment (IQA) model needs to be built by subjective visual measurement, designed in accordance with the results, and applied to the related practical problems. Based on the visual perception characteristics, chromaticity and the structure feature information are quantified, and an objective IQA model combining the color appearance and the gradient image features is proposed in this paper, namely color appearance and gradient similarity(CAGS) model. Two new color appearance indices, vividness and depth, are selected to build the chromatic similarity map. The structure information is characterized by gradient similarity map. Vividness map plays two roles in the proposed model. One is utilized as feature extractor to compute the local quality of distorted image, and the other is as a weight part to reflect the importance of local domain. To quantify the specific parameters of CAGS, Taguchi method is used and four main parameters, i.e., K_V , K_D , K_G and α , of this model are determined based on the statistical correlation indices. The optimal parameters of CAGS are $K_V = K_D = 0.02$, $K_G = 50$, and $\alpha = 0.1$. Furthermore, the CAGS is tested by utilizing 94 reference images and 4830 distorted images from the four open image databases (LIVE, CSIQ, TID2013 and IVC). Additionally, the influences of the 35 distortion types on IQA are analyzed. Massive experiments are performed on four publicly available benchmark databases between CAGS and other 10 state-of-the-art and recently published IQA models, for the accuracy, complexity and generalization performance of IQA. The experimental results show that the accuracy PLCC of the CAGS model can achieve 0.8455 at lowest and 0.9640 at most in the four databases, and the results about commonly evaluation criteria prove that the CAGS performs higher consistency with the subjective evaluations. Among the 35 distortion types, the two distortion types, namely contrast change and change of color saturation, CAGS and mostly IQA models have the worst influence on IQA, and the CAGS yields the highest top three rank number. Moreover, the SROCC values of CAGS for other distortion types are all larger than 0.6 and the number of SROCC value larger than 0.95 is 14 times. Besides, the CAGS maintains a moderate computational complexity. These results of test and comparison above show that the CAGS model is effective and feasible, and the corresponding model has an excellent performance.

Keywords: image quality assessment, human visual perception, color appearance, gradient

PACS: 87.57.C-, 07.05.Pj, 42.30.-d

DOI: 10.7498/aps.69.20200753

* Project supported by the National Key R&D Program of China (Grant No. 2017YFB0403700).

† Corresponding author. E-mail: ydlin@fudan.edu.cn