

基于显著性目标分类的无参考模糊图像质量评价

沈飞鹏^{1*}, 朱彤¹, 张赫男^{1,2}, 陈正豪¹

¹中国矿业大学信息与控制工程学院, 江苏 徐州 221116;

²中国矿业大学地下空间智能控制教育部工程研究中心, 江苏 徐州 221116

摘要 近年来有大量关于无参考模糊图像质量评价的研究,但是目前很多方法都忽略了图像内容对评价结果的影响。针对纯背景的无显著性目标图像和含背景的显著性目标图像的模糊评价方式是不同的,基于人眼注意力机制,前者侧重于图像的整体模糊,而后者更侧重于图像的局部细节模糊。整体模糊指的是图像整体内容的锐度信息,局部细节模糊指的是图像不同位置的局部锐度信息,二者可以将视觉显著性和图像内容更好地结合起来。针对上述问题,提出了一种基于显著性目标分类的无参考模糊图像质量评价方法。首先提出了一种基于显著性检测的目标分类算法,对待评价图像进行显著性目标分类,然后根据分类结果提取其局部模糊特征和全局模糊特征,最后对这两个特征进行融合得到最终的质量评估分数。实验结果表明,该算法不仅在 BLUR 数据库上取得最优的评价效果,同时在 LIVE、CSIQ 和 TID2013 数据库上也有较好的结果,具有很好的鲁棒性。此外,本文算法在各数据库中也表现出了优异的统计性能。

关键词 图像处理; 模糊图像; 质量评价; 显著性; 目标分类

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.2210015

Non-Reference Blur Image Quality Evaluation Based on Saliency Object Classification

Shen Feipeng^{1*}, Zhu Tong¹, Zhang Henan^{1,2}, Chen Zhenghao¹

¹School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China;

²Engineering Research Center of Intelligent Control for Underground Space, Ministry of Education, China University of Mining and Technology, Xuzhou, Jiangsu 221116

Abstract In recent years, there have been a large number of studies on the quality evaluation of non-reference blur images, but many methods ignore the influence of image content on the evaluation results. The blur evaluation methods of the no-saliency object image with pure background and the saliency object image with background are different. Based on the human attention mechanism, the former focuses on the overall blur of the image, while the latter focuses more on the local detail blur of the image. Overall blur refers to the sharpness information of the overall content of the image, while local detail blur refers to the local sharpness information of different locations of the image. The two can better combine visual salience and image content. To solve the above problems, this paper proposes a non-reference blur image quality evaluation method based on saliency object classification. Firstly, this paper proposes an object classification algorithm based on saliency detection, which classifies the saliency objects of the evaluation image, and extracts the local and global blur features according to the classification results. Finally, the two features are fused to obtain the final quality evaluation score. The experimental results show that the algorithm not only achieves the optimal evaluation effect on the BLUR database, but also has good results on the

收稿日期: 2020-12-03; 修回日期: 2021-01-05; 录用日期: 2021-02-12

基金项目: 国家自然科学基金(61771473, 61379143)

通信作者: *641542849@qq.com

LIVE, CSIQ, and TID2013 databases, with good robustness. In addition, the algorithm in this paper also shows excellent statistical performance in various databases.

Key words image processing; blur image; quality evaluation; saliency; object classification

OCIS codes 100.2000; 100.2960; 110.3000

1 引言

随着电子信息技术的高速发展,人们对图像质量的品质要求不断提升,这增加了实际应用对图像质量评价的迫切需求。图像质量评价在图像处理、视频监控、信息采集等领域有着重要的应用。近年来,图像质量评价已经成为国内外学者的研究热点^[1-4],图像质量评价分为主观评价方法和客观评价方法。主观评价方法依靠人眼主观评价,准确率高且符合人眼视觉特性,但这种方法费时费力,面对海量待评价图像数据的及时性要求,无法实现大规模应用。客观评价方法是指设计计算模型来模拟视觉系统的图像感知过程,这种方法不仅易于实现,而且高效快捷。图像的客观质量评价方法主要有三种类型:全参考型、半参考型和无参考型。全参考型的方法要求完全提供原始质量版本的参考图像,半参考型的方法需要提取参考图像的某些特征,这两种方法存在一定的局限性,通常情况下无法访问无失真的参考图像,因此无参考图像质量评价方法^[5]更具有研究意义。

获取、处理和传输图像等过程会产生各种失真,其中模糊是图像最常见的失真类型,直接决定了图像的清晰程度,而清晰度是影响自然图像感知质量的主要因素之一^[6]。无参考模糊图像的质量评估是一个被广泛研究的领域,国内外学者已经提出了许多无参考模糊图像质量评价模型^[7]。Marziliano 等^[8]提出通过计算边缘宽度来测量图像的清晰度,将整幅图像的平均边缘宽度作为评价分数。Ferzli 等^[9]提出了最小可见模糊(JNB)模型,并结合概率求和模型对图像的模糊度进行评价。随后 Narvekar 等^[10]提出了一种基于 JNB 的累积模糊检测概率(CPBD)方法。这一类方法将模糊与边缘宽度相结合,但是忽略了人眼的视觉特性,模糊背景对图像质量评价结果存在干扰。Hassen 等^[11]提出了一种基于局部相位一致性(LPC)的清晰度评估方法,该方法虽然取得了不错的效果,但是只考虑图像局部细节模糊而忽略了图像的整体模糊,算法的准确性和单调性还有待提升。王红玉等^[12]提出了一种基于有效再模糊的无参考图像质量评价(NRIQAVR)算法,该算法基于全局特征处理,忽略

了图像局部细节的模糊。Zhan 等^[13]通过结合最大梯度和梯度变化性提出了一种简单的清晰度评估方法(GMVG),该算法没有考虑人眼视觉注意力机制,忽略了视觉显著性对图像质量的影响^[14]。周圆等^[15]提出了一种基于模糊检测概率变换的模糊图像质量评价算法,该算法直接将视觉显著性引入评价算法中,但忽略了纯背景图像边缘块的模糊度,这一类图像是没有显著性目标的,显著性特征会干扰最终的评价结果,此外算法的评价结果过于依赖显著性检测算法的性能。

针对现有研究的不足之处,本文提出了基于显著性目标分类的无参考模糊图像质量评价方法。充分考虑到人眼视觉注意力机制^[16],有显著性目标的图像和无显著性目标的图像对于人眼主观评价的影响是不同的,因此首先对图像内容进行有无显著性目标分类,根据分类的结果选择不同的特征计算方式。最后计算图像的局部模糊特征和全局模糊特征,进行特征融合得到最终的质量分数,局部特征用来衡量图像的局部细节模糊,全局模糊特征用来衡量图像的整体模糊。

2 算法

显著性目标图像(SOI)是含有前景和背景的图像,前景指的是显著性区域,无显著性目标图像(NOI)指纯背景图像,没有前后景。如图 1 所示,图 1(a)中的野狼为显著性目标,属于 SOI,而图 1(b)是一幅纯背景图像,无显著性目标,属于 NOI。根据人类视觉系统(HVS)的心理视觉特性^[17],对于 SOI 来说,人眼更加关注显著性目标区域,其显著性区域对图像的清晰度影响更大,同时显

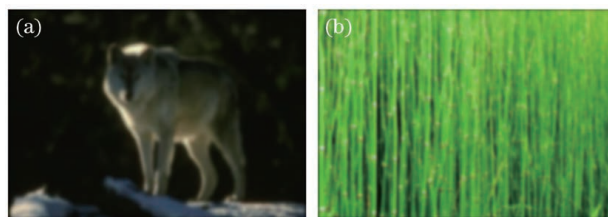


图 1 两幅不同内容的图像。(a) 显著性目标图像;
(b) 无显著性目标图像

Fig. 1 Two images with different content. (a) Saliency object image; (b) no-saliency object image

显著性目标区域外的背景模糊程度有可能造成背景虚化,进而影响评价判断;而对于无显著性目标的图像来说,人眼关注的是图像的整体,边缘区域和中心区域的模糊程度对图像整体清晰度的影响区别不大。

根据以上研究结论,本文提出了基于显著性目标分类的图像质量评价(SOA)算法,算法框架如图 2 所示,本文模型主要由图像分类、特征提取以及

特征融合三部分组成。为了判别输入图像的内容有无显著性目标,在现有显著性检测算法^[18]的基础上,本文提出了一种基于显著性检测的目标分类(SDOC)算法。首先对输入图像进行预处理,根据显著性分类算法判断图像属于 SOI 还是 NOI,然后根据分类结果提取图像的局部模糊特征和全局模糊特征,进行特征融合,得到最终的评价模型,下文将展开介绍图像分类、特征提取以及特征融合。

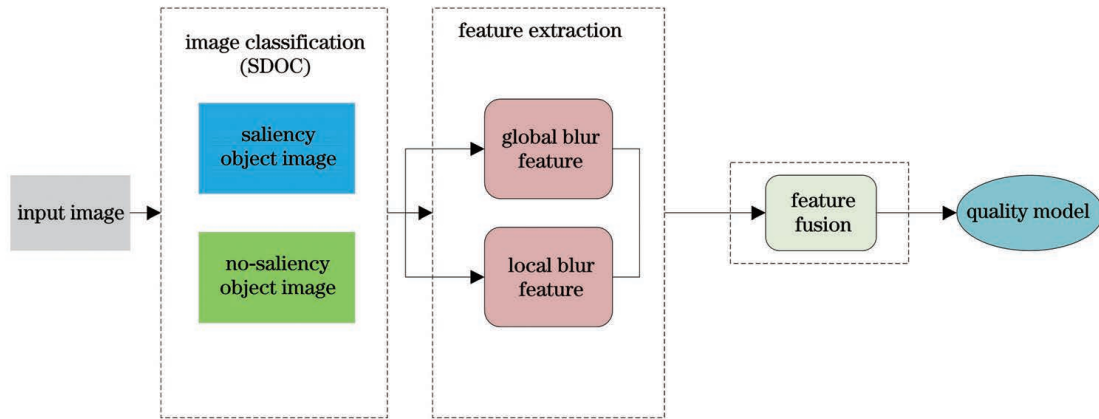


图 2 本文提出的质量评价方法框架图

Fig. 2 Framework of quality evaluation method proposed in this paper

2.1 基于显著性检测的目标分类

SDOC 算法框架如图 3 所示,该算法使用了两次级联的显著性检测,一级显著性检测是对输入图像进行显著性检测,二级显著性检测是在一级显著性检测的基础上对其显著性区域进行显著性检测,这样一方面可以避免显著性目标图像的目标过大对分类结果产生干扰,另一方面可以更准确地提取图

像的显著性区域信息。根据两次级联的显著性检测和区域分割结果来判断输入图像内容是否含有显著性目标,首先对输入图像进行显著性检测,得到一级显著性图,然后利用显著性区域分割算法得到一级显著性区域分割图,计算一级显著性区域分割图和输入图像的面积比值 ω_1 。接着对一级显著性区域分割图进行二级显著性检测,得到二级显著性区域

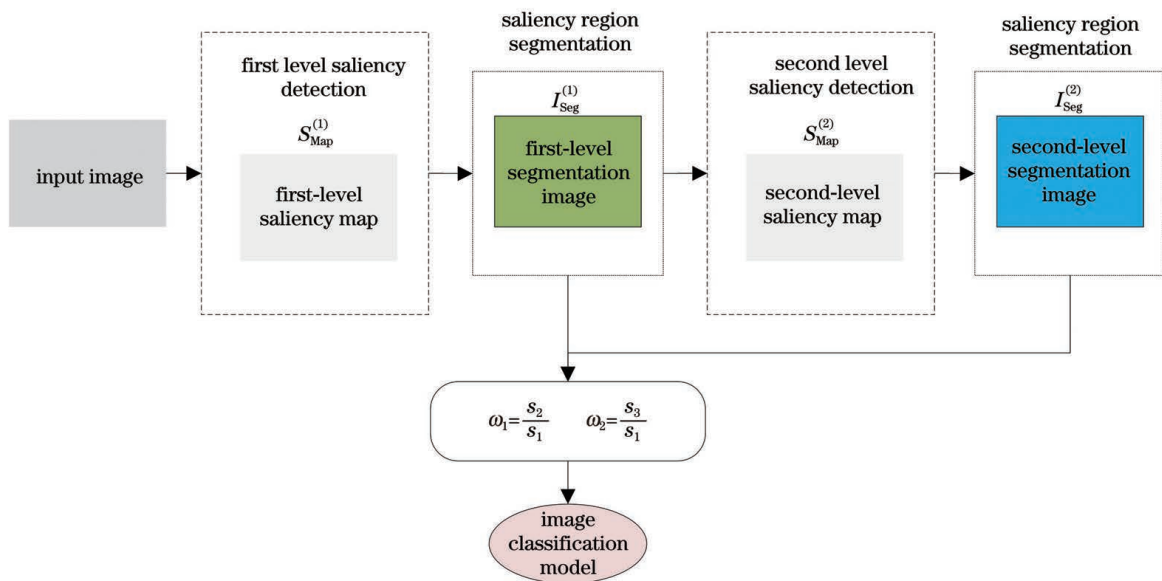


图 3 SDOC 算法框架图

Fig. 3 Framework diagram of SDOC algorithm

分割图并计算二级显著性区域分割图和输入图像的面积比值 ω_2 。最后设置两个阈值 t_1 和 t_2 , 若 $\omega_1 \leq t_1$ 且 $\omega_2 \leq t_2$, 则表明原图像含有显著性目标, 否则表明原图像无显著性目标。

图 4 是显著性区域分割的示例图, 其中: 图 4(a) 为显著性目标图像, 图 4(b) 和图 4(c) 分别

为对应的一级显著性区域分割图和二级显著性区域分割图, 图 4(d) 为无显著性目标图像, 图 4(e) 和图 4(f) 分别为对应的一级显著性区域分割图和二级显著性区域分割图。通过对比发现, 两类图像的显著性区域分割图的 ω_1 和 ω_2 是不同的, 如图 4(b)、图 4(e)、图 4(c) 和图 4(f) 所示。



图 4 显著性区域分割示例图。(a) 有显著性目标图像的示例图; (b) 图 4(a) 的一级显著性区域分割图; (c) 图 4(a) 的二级显著性区域分割图; (d) 无显著性目标图像的示例图; (e) 图 4(d) 的一级显著性区域分割图; (f) 图 4(d) 的二级显著性区域分割图

Fig. 4 Example images of saliency region segmentation. (a) Example image with saliency object; (b) first-level saliency region segmentation map of Fig. 4(a); (c) second-level saliency region segmentation map of Fig. 4(a); (d) example image without object; (e) first-level saliency region segmentation map of Fig. 4(d); (f) second-level saliency region segmentation map of Fig. 4(d)

为了验证上述发现, 将 BLUR 数据库 144 幅图像依次代入算法中进行目标分类, 得到每幅图像的一级显著性区域的面积比 ω_1 和二级显著性区域的面积比 ω_2 , 然后将所有结果值以散点图形式展示, 如图 5 所示, 其中: 横轴表示 ω_1 值, 纵轴表示 ω_2 值。SOI 的 ω_1 值集中在 0.25 和 0.75 之间, ω_2 值集中在 0 和 0.5 之间。NOI 的 ω_1 值集中在 0.75 和 1 之间, ω_2 值集中在 0.5 和 1 之间。两类图像的 ω_1 和

ω_2 值表现出了聚类特性, 表明实现了图像有无显著性目标分类的目的, 基于此, 阈值 t_1 和 t_2 设置为 0.5。图 5 中的数据点为 144 个, 然而图中出现了很多数据点重叠现象, 其原因是数据库中相同内容但不同模糊程度的图像的数据值基本一致, 这一方面造成了图中数据点重叠现象, 另一方面也说明了模糊程度对分类结果的影响较小, 分类算法具有很高的鲁棒性。

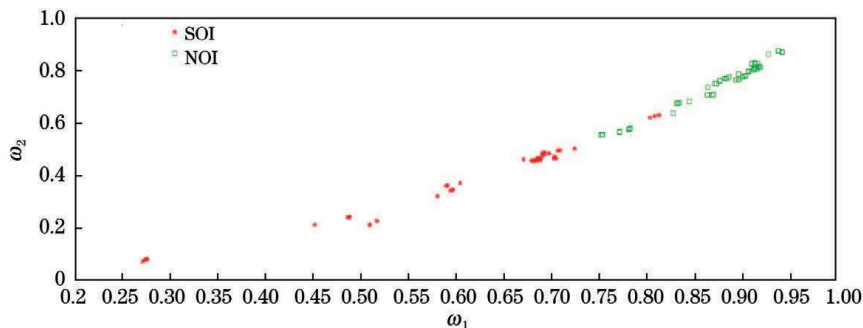


图 5 BLUR 数据库分类结果散点图

Fig. 5 Scatter plot of BLUR database classification results

2.2 全局模糊特征

图像的全局模糊特征和局部模糊特征根据

SDOC 算法的分类结果以及获取的显著性区域信息进行提取。本文利用图像的梯度特征来计算图像的

模糊程度,梯度特征能有效描述图像的结构信息^[19],

可以更好地体现图像的内容,梯度特征可表示为

$$G(i, j) = \sqrt{[I^{(1)}(i, j) - I^{(1)}(i, j + 1)]^2 + [I^{(1)}(i, j) - I^{(1)}(i + 1, j)]^2}, \quad (1)$$

式中: G 为图像 $I^{(1)}$ 对应的梯度图矩阵, 即 $G(i, j)$ 表示像素 $I^{(1)}(i, j)$ 所对应的梯度值矩阵。若输入图像属于 SOI, 为了降低背景虚化图像边缘锐度对整体质量的评估干扰, 要对梯度图 G 进行池化处理。然后计算图像梯度的最大值 g_{\max} 、标准差值 g_{var} 和均值 g_{mean} , 梯度最大值可以有效描述图像整体的锐度信息^[13]。根据 g_{var} 和 g_{mean} 可以计算出梯度图的梯度变化值 g_{cv} , g_{cv} 可以根据梯度值的变化来描述图像内容的变化情况, 其中

$$g_{\text{cv}} = \frac{g_{\text{var}}}{g_{\text{mean}}}. \quad (2)$$

最后根据 g_{\max} 和 g_{cv} 计算出输入图像 $I^{(1)}$ 的全局模糊特征, 用 q_1 表示:

$$q_1 = g_{\max}^{\alpha_1} \cdot g_{\text{cv}}^{1-\alpha_1}, \quad (3)$$

式中: α_1 为设置的可调节参数。

2.3 局部模糊特征

为了获取图像不同区域的模糊信息, 首先对图像 $I^{(1)}$ 进行分块操作, 图像 $I^{(1)}$ 局部模糊特征是基于显著性目标分类算法得到的, 与图像所属类别有关。两类图像的局部特征提取算法是不同的, SOI 的局部模糊信息需要将前景和背景分开考虑, 而 NOT 是纯背景图像, 需要综合考虑图像的不同区域的局部模糊信息。将图像内容无显著性目标的局部特征提取算法标记为 L_{fe_1} , 将图像内容含有显著性目标的局部特征提取算法标记为 L_{fe_2} 。

2.3.1 L_{fe_1} 算法

由于此类图像内容无显著性目标, 图像不存在前景和背景之分, 即使边缘块也会影响图像的整体清晰度。 L_{fe_1} 算法首先计算块集合中每一个图像块的梯度图矩阵, 每个图像块的梯度信息可以有效描述图像不同区域的结构特征, 然后计算每个图像块的梯度最大值, 用来表示每个图像的局部锐度信息^[13], 同时根据(2)式计算出图像块的梯度变化值, 其中: 第 m 个图像块 $I_p^{(m)}$ 的最大梯度值标记为 $g_{\max}^{(m)}$, 梯度变化值标记为 $g_{\text{cv}}^{(m)}$ 。接着分别对 N 个图像块的最大梯度值和梯度变化值进行几何平均, 得到 g_{\max_1} 和 g_{cv_1} :

$$g_{\max_1} = \sqrt[N]{\prod_{m=1}^N g_{\max}^{(m)}}, \quad (4)$$

$$g_{\text{cv}_1} = \sqrt[N]{\prod_{m=1}^N g_{\text{cv}}^{(m)}}. \quad (5)$$

最后根据(6)式得到图像 $I^{(1)}$ 的局部模糊特征值 q_2 :

$$q_2 = g_{\max_1}^{\alpha_2} \cdot g_{\text{cv}_1}^{1-\alpha_2}, \quad (6)$$

式中: α_2 为设置的可调节参数。

2.3.2 L_{fe_2} 算法

此类图像的局部模糊特征获取需要将前景和背景分开考虑, 设置一个阈值 t_3 , 由图 5 可知, t_3 取值 0.4 较为合理, 可以有效实现前后景分离。 L_{fe_2} 算法首先根据 t_3 对原图像 $I^{(1)}$ 进行前后景分离, 背景即为图像的背景。若 $\omega_2 \geq t_3$, 表明显著性目标过大, 则一级显著性区域分割图 $I_{\text{Seg}}^{(1)}$ 在 $I^{(1)}$ 中对应位置即为图像 $I^{(1)}$ 的前景区域, 否则前景区域为二级显著性区域分割图 $I_{\text{Seg}}^{(2)}$ 所对应位置, 此时图像的显著性目标较小, 一级显著性区域无法准确描述图像的显著性区域信息。然后分别计算出前景区域和背景区域的局部模糊特征 q_f 和 q_b , 计算步骤与 L_{fe_1} 算法相同。最后将前景和背景的模糊特征进行融合得到的图像局部模糊特征 q_2 :

$$q_2 = q_f^s \cdot q_b^{1-s}, \quad (7)$$

式中: 上标 s 代表面积, 由 ω_2 大小决定, 若 $\omega_2 \leq t_2$, 表明图像的显著性目标区域较小, 此时 s 等于 s_1 , 否则 s 等于 s_2 。

2.4 特征融合

本文从局部和全局分别提取图像 $I^{(1)}$ 对应的模糊特征, 全局特征是从整体上考虑图像内容对图像清晰度的影响, 局部特征综合考虑了图像不同区域块的模糊特征信息。将基于显著性目标分类的局部和全局特征进行融合的过程不仅引入了人眼注意力机制, 同时也兼顾了图像的内容信息, 融合公式为

$$Q_{\text{blur}} = q_1^{1-\beta} + q_2^\beta, \quad (8)$$

式中: β 为可调节参数。由(8)式可得到图像 $I^{(1)}$ 的质量评价指标。

3 实 验

本文对比算法为 JNB^[9]、CPBD^[10]、LPC^[11]、NRIQAVR^[12] 和 GMVG^[13], 其中: JNB、CPBD、

LPC 和 GMVG 来自论文的公开代码, NRIQAVR 则是根据论文进行复现仿真的。本文的评价指标通过客观质量评价分数与数据库主观分数拟合得出, 评价指标包括 PLCC (其值可用 A_{PLCC} 表示)、SROCC (其值可用 A_{SROCC} 表示) 和 KROCC。其中: PLCC 指的是皮尔逊线性相关系数, 描述客观评价算法与主观评价之间的相关性, 衡量算法的准确性, PLCC 值越大表示算法的准确性越高; SROCC 指的是斯皮尔曼秩相关系数, 衡量客观评价算法预测的单调性, SROCC 值越大表示算法的单调性越好; KROCC 指的是肯德尔秩次相关系数, 衡量算法预测的单调性, KROCC 值越大表示算法的单调性越好。

表 1 各数据库的参数介绍

Table 1 Introduction of parameters of each database

Database	Distortion type	Number of samples	Resolution	Subjective evaluation	Range
LIVE	Gaussian blur	174	768×512	DMOS	[0, 100]
CSIQ		150	512×512	DMOS	[0, 1]
TID2013		125	512×384	MOS	[0, 9]
BLUR		144	481×321	MOS	[0, 10]

3.2 相关参数验证

根据(3)式, α_1 取不同值时, 将全局特征 q_1 作为单独的评估指标在各数据库中进行测试, 测试结果如图 6 所示, LIVE 和 CSIQ 数据库主观评价方式为 DMOS, 因此 PLCC 和 SROCC 值越小, 评价结果就越优, 而 TID2013 和 BLUR 数据库采用的是

3.1 数据库

关于模糊图像质量评价的公共数据库, 数据库主观得分一般采用平均意见得分 (MOS) 或者差分平均意见得分 (DMOS), 其作用是量化图像的主观感知质量。DMOS 是在 MOS 的基础上计算得到的, 与 MOS 相反, DMOS 值越大, 表明对应图像的主观感知质量越差。目前数据库主要有三个, 分别为: LIVE 图像质量评价数据库^[20]、TID2013 图像质量评价数据库^[21] 和 CSIQ 质量评价数据库^[22]。此外, 本文还在 BLUR^[23] 数据库进行了测试, 该数据库由 144 幅不同模糊程度的图像构成, 包括 72 幅有显著性目标的图像和 72 幅无显著性目标的图像。各数据库的参数介绍如表 1 所示。

MOS 评价方式, PLCC 和 SROCC 值越大, 评价结果就越优, 下文算法中均采用绝对值进行分析。其中图 6(a) 是在 LIVE 数据库中的测试结果: 当 α_1 取值为 0.64 时, PLCC 达到最小值 -0.9594; 当 α_1 取值为 0.67 时, SROCC 达到最小值 -0.9660, 因此 α_1 取 0.64 和 0.67 的均值 (0.655) 时, 在 LIVE 数据

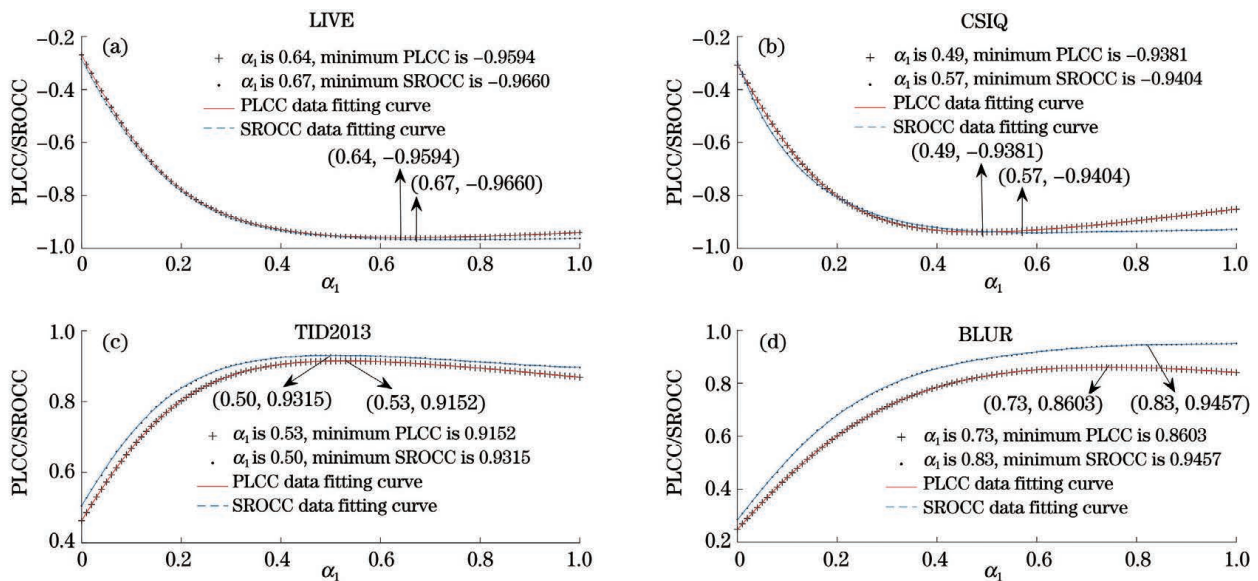


图 6 全局模糊特征在各数据库的测试结果。(a) LIVE 数据库; (b) CSIQ 数据库; (c) TID2013 数据库; (d) BLUR 数据库
Fig. 6 Test results of global blur features in each database. (a) LIVE database; (b) CSIQ database; (c) TID2013 database; (d) BLUR database

中测试结果达到最佳效果。同理,CSIQ、TID2013 和 BLUE 数据库中最佳 α_1 值分别为:0.53、0.515 和 0.78,最终 α_1 取这 4 组数的均值时,全局模糊特征在 4 个数据库中综合性能达到最佳,即 α_1 取值为 0.62。局部模糊特征是对图像块的梯度进行相关计算,其本质上与全局模糊特征一样,因此默认 α_2 和 α_1 取值一致,即 0.62。

当 β 取不同值时,将 SOA 算法在各数据库中进行验证分析,测试结果如图 7 所示:图 7(a)为 PLCC

结果图,图 7(b)为 SROCC 结果图。由图 7(a)可知,当 β 分别取 0.26,0.28,0.18,0.42 时,SOA 算法分别在各数据库中的算法准确度达到最大值,因此 β 取均值 0.285,此时算法在各数据库中的综合准确度最高。由图 7(b)可知,当 β 分别取 0.22,0.22,0.12,0.40 时,SOA 分别在各数据库中的算法单调性最好,因此 β 取均值 0.24,此时算法在各数据库中的综合单调性最优。最终, β 取 0.285 和 0.24 的均值,即 0.2625,此时 SOA 算法在各数据库中的综合性能最优。

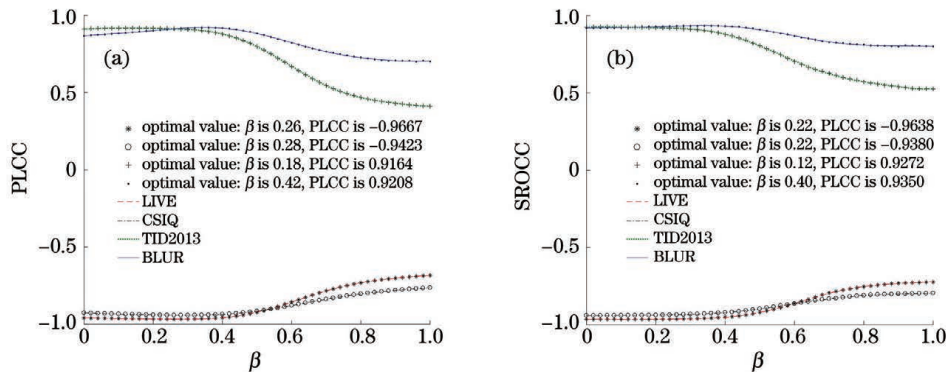


图 7 β 取不同值时,SOA 算法在各数据库中的测试结果。(a)PLCC;(b)SROCC

Fig. 7 Test results of SOA algorithm in each database for different β . (a) PLCC; (b) SROCC

3.3 算法性能分析

本文首先在 LIVE 数据库中进行客观评价实验,6 种无参考模糊图像质量评价算法的测试结果如表 2 所示,表中每种算法的 PLCC、SROCC 和 KROCC 的最大的两个值均用粗体标出。由表 2 可以看出:SOA 算法的 PLCC 值为 0.9667,排名第一,远优于其余 5 种方法;SOA 算法的 SROCC 值和 KROCC 值分别为 0.9622 和 0.8369,排名第二,远优于 JNB、CPBD、LPC 和 NRIQAVR 方法,与 GMVG 算法的测试结果相差不大。以上结果表明本文方法在 LIVE 数据库测试中具有最高的准确性和良好的单调性。

为了验证算法的鲁棒性,在数据库 CSIQ 和

表 2 LIVE 数据库测试结果

Table 2 LIVE database test results

Algorithm	PLCC	SROCC	KROCC
JNB	0.8221	0.8419	0.6651
CPBD	0.9123	0.9429	0.8043
LPC	0.8172	0.9594	0.8241
NRIQAVR	0.8023	0.8441	0.7060
GMVG	0.9335	0.9633	0.8377
SOA	0.9667	0.9622	0.8369

TID2013 中分别进行测试。测试结果如表 3、4 所示:SOA 算法的 PLCC 值分别是 0.9421 和 0.9130,

表 3 CSIQ 数据库测试结果

Table 3 CSIQ database test results

Algorithm	PLCC	SROCC	KROCC
JNB	0.2508	0.7624	0.5971
CPBD	0.8292	0.8846	0.7081
LPC	0.9096	0.9068	0.7197
NRIQAVR	0.9152	0.8874	0.7688
GMVG	0.9264	0.9500	0.8153
SOA	0.9421	0.9358	0.7892

表 4 TID2013 数据库测试结果

Table 4 TID2013 database test results

Algorithm	PLCC	SROCC	KROCC
JNB	0.6876	0.6902	0.5137
CPBD	0.8491	0.8520	0.6470
LPC	0.8382	0.8888	0.6839
NRIQAVR	0.7819	0.7967	0.6533
GMVG	0.9125	0.9274	0.7600
SOA	0.9130	0.9185	0.7430

排名第一,都是最优值,这表明本文方法在 CSIQ 和 TID2013 数据库测试中准确性最高;SROCC 值和 KROCC 值最大的是 GMVG 算法和 SOA 算法,SOA 算法排名第二,表明本文方法在 CSIQ 和 TID2013 数据库测试中单调性良好。

为了进一步说明显著性目标分类对客观质量评价的影响程度,本文在 BLUR 数据库进行算法对比测试,该数据库中 SOI 和 NOI 样本数均衡,如表 5 所示。SOA 算法的 PLCC 和 SROCC 在 6 种客观模糊图像质量评价算法中都是最大的,性能排名第一,说明显著性目标分类对于算法的性能提升作用很明显。综合表 2~5 数据可得:JNB、CPBD 和 NRIQAVR 算法性能最差;LPC 算法在 BLUR 数据库测试中性能略低于本文方法,而在 LIVE、CSIQ 和 TID2013 数据库测试中性能较差,表明 LPC 算法的鲁棒性欠佳;GMVG 算法在 LIVE、CSIQ、TID2013 数据库中的准确性指标(PLCC)远不如本文方法,单调性指标(SROCC 和 KROCC)略高于本文方法,而在 BLUR 数据库测试中性能指标均较差,表明

GMVG 算法的鲁棒性也欠佳。综上所述,本文所提的 SOA 算法在 LIVE、CSIQ 和 TID2013 数据库测试中表现出了最好的鲁棒性,同时在 BLUR 数据库测试中 PLCC 和 SROCC 评价指标都是最优的,即本文方法综合性能最优。SOA 算法的客观评价分数与各数据库上主观评价得分的评估结果散点图如图 8 所示,拟合曲线呈单调递增趋势,再次表明算法的良好鲁棒性,本文所提方法更符合人眼视觉原理。

表 5 BLUR 数据库测试结果
Table 5 BLUR database test results

Algorithm	PLCC	SROCC	KROCC
JNB	0.6295	0.6691	0.7621
CPBD	0.8944	0.9077	0.8318
LPC	0.9108	0.9049	0.7058
NRIQAVR	0.8261	0.8448	0.6876
GMVG	0.7918	0.9000	0.7441
SOA	0.9173	0.9348	0.7786

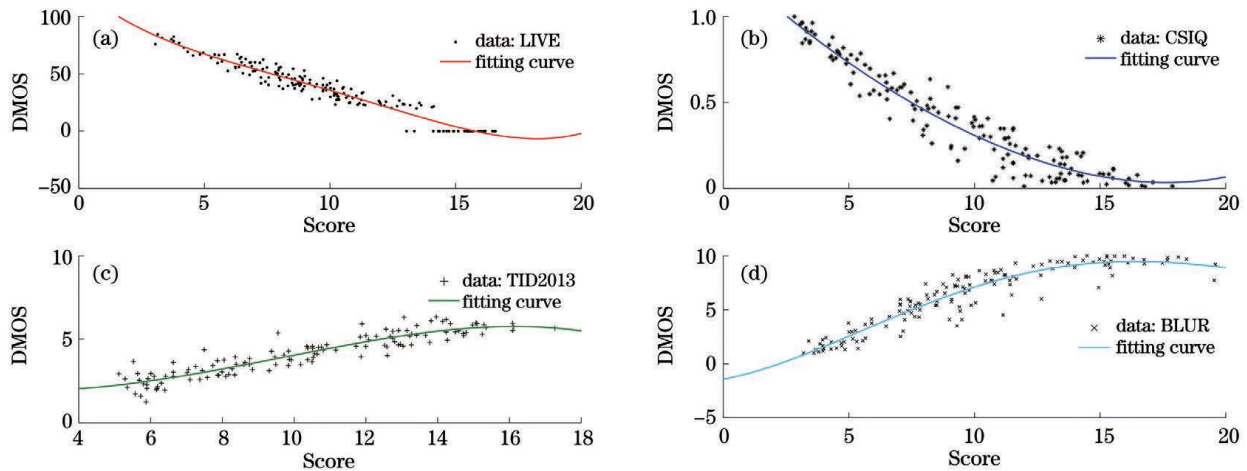


图 8 SOA 算法在各数据库的结果散点图以及对应的拟合曲线。(a)LIVE;(b)CSIQ;(c)TID2013;(d)BLUR

Fig. 8 Scatter plots of results of SOA algorithm for each database and corresponding fitting curves. (a) LIVE; (b) CSIQ; (c) TID2013; (d) BLUR

为进一步证明本文算法的性能特点,对本文算法进行预测曲线对比实验以及算法性能统计实验。在 LIVE 数据库中,对性能较优的算法的评估结果进行预测曲线对比,结果如图 9 所示,其中图 9(a)~(d)分别是 CPBD、LPC、GMVG 以及 SOA 算法的预测曲线结果图,对于非线性拟合,数据散点越接近预测曲线,说明算法性能越好。由图 9 可知:CPBD 和 LPC 算法的曲线有局部波动现象,数据点距离预测曲线也较分散;GMVG 和 SOA 算法的曲线较平滑,呈单调下降趋势,曲线无明显波

动,而 SOA 算法曲线波动更小,数据点更均匀地分散在预测曲线周围。因此,通过对比可得,本文所提 SOA 算法具有更好的主观感知一致性,所得结果从不同角度再次验证了 SOA 更好的算法性能。

基于上述实验结果,本文使用 F 检验来评估每种算法相对于 SOA 算法的统计显著性。具体来说,使用 F 检验评估 SOA 算法的客观得分和主观得分之间的预测性能误差,即算法的准确性误差 σ_1 和单调性误差 σ_2 ,F 检验是一种度量不同算法之间的误差性能的有效工具^[24]。其中 σ_1 和 σ_2 计算方法为

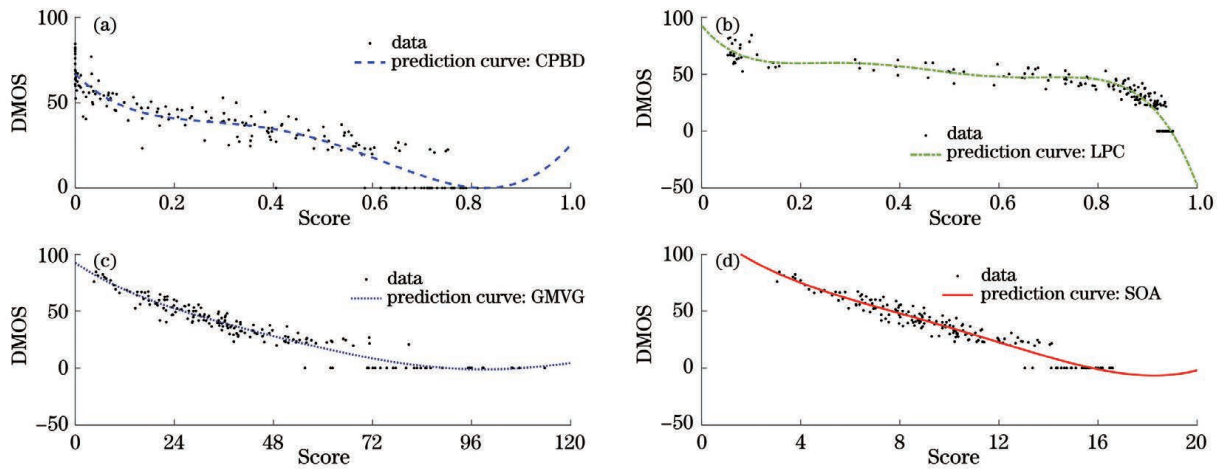


图 9 不同算法的预测曲线对比图。(a)CPBD;(b)LPC;(c)GMVG;(d)SOA

Fig. 9 Comparison of prediction curves of various algorithms. (a) CPBD; (b) LPC; (c) GMVG; (d) SOA

$$\sigma_1 = (1 - |A_{PLCC}|) \times 100, \quad (9)$$

$$\sigma_2 = (1 - |A_{SROCC}|) \times 100, \quad (10)$$

PLCC 和 SROCC 值越接近于 1,说明误差越小,本文取 σ_1 和 σ_2 的均值作为算法的综合性性能误差值 σ , 6 种算法在各数据库中的 σ 值如表 6 所示。将待对比算法统称为算法 A。为了对比算法 A 与算法 SOA 的统计性能,首先需要计算 F 检验分数(其值用 F 表示):

$$F = \sigma_A^2 / \sigma_{SOA}^2, \quad (11)$$

式中: σ_A 和 σ_{SOA} 分别表示待对比算法和 SOA 算法的综合性性能误差值。此外,还使用了一个阈值 $F_{critical}^{[24]}$ 来比较算法之间的统计性能,本文 F 检验的自由度为各数据的样本数量,置信度设置为 0.9,利用 MATLAB 的 finv 函数即可计算各数据的阈值 $F_{critical}$ 。如表 7 所示,LIVE、CSIQ、TID2013 和 BLUR 数据库的阈值 $F_{critical}$ 分别为 1.2150,1.2335,1.2586,1.2388。然后根据 F 检验分数和阈值 $F_{critical}$ 综合比较算法的统计性能:若 $F > F_{critical}$,说明待对比算法 A 的预测误差更大,即 SOA 算法统

表 6 不同算法在各数据库中的综合性性能误差值

Table 6 Comprehensive performance error values of different algorithms for each database

Algorithm	LIVE	CSIQ	TID2013	BLUR
JNB	16.800	49.340	31.110	35.070
CPBD	7.240	14.310	14.945	9.895
LPC	11.170	9.180	13.650	9.215
NRIQAVR	17.680	9.870	21.070	16.455
GMVG	5.160	6.180	8.005	15.410
SOA	3.555	6.105	8.425	7.395

表 7 各数据库的 F 检验相关参数

Table 7 F test related parameters of each database

Database	Confidence level	Freedom	$F_{critical}$
LIVE	0.9	174	1.2150
CSIQ		150	1.2335
TID2013		125	1.2586
BLUR		144	1.2388

计性能明显优于算法 A;若 $1/F_{critical} < F < F_{critical}$,说明算法 A 和 SOA 的预测误差相差不大;否则,说明 SOA 算法的预测误差明显比算法 A 差,即 SOA 算法的统计性能较差。SOA 算法在各数据库的性能统计结果如表 8 所示,0 表示与 SOA 算法统计性能基本一致,1 表示 SOA 算法统计性能更优,-1 表示对比算法统计性能更优。由表 8 可知,SOA 算法在 LIVE 和 BLUR 数据库中的统计性能明显优于其他对比算法,而在 CSIQ 和 TID2013 数据库中,除与 GMVG 算法统计性能相差不大外,SOA 算法明显优于其余算法。

综上所述,本文方法能够有效模拟人眼主观感

表 8 SOA 算法在各数据库中的性能统计结果

Table 8 Performance statistical results of SOA algorithm for each database

Algorithm	LIVE	CSIQ	TID2013	BLUR
JNB	1	1	1	1
CPBD	1	1	1	1
LPC	1	1	1	1
NRIQAVR	1	1	1	1
GMVG	1	0	0	1

知特性,具有很高的准确性和很好的单调性,各数据库中的测试结果表明了 SOA 算法良好的鲁棒性,此外 SOA 算法也表现出了优异的统计性能。

4 结 论

提出了基于显著性目标分类的无参考模糊图像质量评价方法。该算法首先对图像进行有无显著性目标判定,然后根据图像内容的分类结果针对性地提取其全局模糊特征和局部模糊特征,再进行特征融合,得到最终的质量分数。从实验结果可以看出,本文方法能够有效避免图像内容对图像质量评价的影响,更符合人眼视觉原理,同时算法综合考虑图像局部细节模糊和整体模糊,具有良好的鲁棒性。但是本文算法只考虑有无显著性目标这两种情况,未来可以加入更多的分类信息,同时也可以将本文方法应用于其他失真类型的图像。

参 考 文 献

- [1] Ye Y Q, Hu W J. No-reference quality assessment method for inpainting Thangka image based on multiple features[J]. *Laser & Optoelectronics Progress*, 2020, 57(8): 081105.
叶雨琪, 胡文瑾. 基于多特征的彩色唐卡修复图像无参考质量评价方法[J]. *激光与光电子学进展*, 2020, 57(8): 081105.
- [2] Zhou K, Wu C M, Li C X. Quality assessment of blind color images using quaternion Fourier transform[J]. *Laser & Optoelectronics Progress*, 2020, 57(18): 181021.
周珂, 吴成茂, 李昌兴. 基于四元数傅里叶变换的盲彩色图像质量评价[J]. *激光与光电子学进展*, 2020, 57(18): 181021.
- [3] Yu W, Xu J J, Liu Y Y, et al. No-reference quality evaluation for gamut mapping images based on natural scene statistics[J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141006.
余伟, 徐晶晶, 刘玉英, 等. 基于自然场景统计的色域映射图像无参考质量评价[J]. *激光与光电子学进展*, 2020, 57(14): 141006.
- [4] Li Y M, Shu S B, Li Z W, et al. Atmospheric correction and image quality assessment of WFV camera in GaoFen-1 satellite[J]. *Acta Optica Sinica*, 2020, 40(20): 2001004.
李妍敏, 舒双宝, 李志文, 等. 高分一号 WFV 相机的大气校正与图像质量评估[J]. *光学学报*, 2020, 40(20): 2001004.
- [5] Huang S Y, Sang Q B. No-reference stereo image quality assessment based on image fusion[J]. *Laser & Optoelectronics Progress*, 2019, 56(7): 071004.
黄姝钰, 桑庆兵. 基于图像融合的无参考立体图像质量评价[J]. *激光与光电子学进展*, 2019, 56(7): 071004.
- [6] Kim H, Kim J, Oh T, et al. Blind sharpness prediction for ultrahigh-definition video based on human visual resolution[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(5): 951-964.
- [7] Athar S, Wang Z. A comprehensive performance evaluation of image quality assessment algorithms[J]. *IEEE Access*, 2019, 7: 140030-140070.
- [8] Marziliano P, Dufaux F, Winkler S, et al. Perceptual blur and ringing metrics: application to JPEG2000[J]. *Signal Processing: Image Communication*, 2004, 19(2): 163-172.
- [9] Ferzli R, Karam L J. A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)[J]. *IEEE Transactions on Image Processing*, 2009, 18(4): 717-728.
- [10] Narvekar N D, Karam L J. A No-reference image blur metric based on the cumulative probability of blur detection (CPBD)[J]. *IEEE Transactions on Image Processing*, 2011, 20(9): 2678-2683.
- [11] Hassen R, Wang Z, Salama M M A. Image sharpness assessment based on local phase coherence[J]. *IEEE Transactions on Image Processing*, 2013, 22(7): 2798-2810.
- [12] Wang H Y, Feng J, Niu W, et al. No-reference image quality assessment based on re-blur theory[J]. *Chinese Journal of Scientific Instrument*, 2016, 37(7): 1647-1655.
王红玉, 冯筠, 牛维, 等. 基于再模糊理论的无参考图像质量评价[J]. *仪器仪表学报*, 2016, 37(7): 1647-1655.
- [13] Zhan Y B, Zhang R. No-reference image sharpness assessment based on maximum gradient and variability of gradients[J]. *IEEE Transactions on Multimedia*, 2018, 20(7): 1796-1808.
- [14] Cheng D Q, Shao L R, Liu W L, et al. Non-reference blur image quality evaluation based on saliency[J]. *Video Engineering*, 2018, 42(7): 1-4.
程德强, 邵丽蓉, 刘威龙, 等. 基于显著性的无参考模糊图像质量评价[J]. *电视技术*, 2018, 42(7): 1-4.
- [15] Zhou Y, Wang K, Zhang H X, et al. Blur image quality assessment method based on blur detection probability variation[J]. *Laser & Optoelectronics Progress*, 2020, 57(10): 101004.
周圆, 王凯, 张皓翔, 等. 基于模糊检测概率变化的模糊图像质量评价算法[J]. *激光与光电子学进展*,

- 2020, 57(10): 101004.
- [16] Liu H T, Heynderickx I. Visual attention in objective image quality assessment: based on eye-tracking data[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(7): 971-982.
- [17] Wu J J. Image information perception and quality assessment based on the human visual system[D]. Xi'an: Xidian University, 2014: 9-15.
吴金建. 基于人类视觉系统的图像信息感知和图像质量评价[D]. 西安: 西安电子科技大学, 2014: 9-15.
- [18] Borji A, Cheng M M, Hou Q B, et al. Salient object detection: a survey[J]. Computational Visual Media, 2019, 5(2): 117-150.
- [19] Zhan Y B. Image quality assessment based on visual perception [D]. Hefei: University of Science and Technology of China, 2018: 89-95.
詹忆冰. 基于视觉感知特性的图像质量评价方法研究[D]. 合肥: 中国科学技术大学, 2018: 89-95.
- [20] Sheikh H R, Sabir M F, Bovik A C. A statistical evaluation of recent full reference image quality assessment algorithms [J]. IEEE Transactions on Image Processing, 2006, 15(11): 3440-3451.
- [21] Ponomarenko N, Jin L N, Ieremeiev O, et al. Image database TID2013: peculiarities, results and perspectives[J]. Signal Processing: Image Communication, 2015, 30: 57-77.
- [22] Larson E C, Chandler D M. Most apparent distortion: full-reference image quality assessment and the role of strategy [J]. Journal of Electronic Imaging, 2010, 19(1): 011006.
- [23] Liu H T, Wang J L, Redi J, et al. An efficient no-reference metric for perceived blur[C]//3rd European Workshop on Visual Information Processing, July 4-6, 2011, Paris, France. New York: IEEE Press, 2011: 174-179.
- [24] Li L, Wu D, Wu J, et al. Image sharpness assessment by sparse representation [J]. IEEE Transactions on Multimedia, 2016, 18(6): 1085-1097.