

DOI:10.11918/j.issn.0367-6234.201709138

图像质量评价:融合视觉特性与结构相似性指标

朱新山^{1,2}, 姚思如¹, 孙彪¹, 钱永军¹

(1. 天津大学 电气自动化与信息工程学院, 天津 300072;

2. 信息安全国家重点实验室(中国科学院信息工程研究所), 北京 100093)

摘 要: 图像质量评价是多媒体领域的一个基本问题。在现有结构相似性方案的基础上,增加最小可辨失真和视觉显著性评价用于构造一个新的图像质量评价指标。首先,考虑到人眼视觉系统的掩蔽效应,设计了失真图像的修正模型,以 Sigmoid 函数为基础,利用最小可辨失真阈值和图像的像素变化绝对值构建失真图像修正模型,对失真图像进行修正,使其符合人眼的感知效果;然后,考虑到人眼视觉系统的视觉注意特性,设计了图像区域权重模型,利用视觉显著性图像表征图像感兴趣内容,作为图像不同区域的权重;最后,计算修正后的失真图像与原图像的局部结构相似性,并利用区域权重对局部区域质量进行加权平均获得全局图像的质量评价。实验结果表明,与同类度量指标相比,本文提出的指标在评价图像局部质量方面,更符合人的主观感知效果;在主观数据拟合方面,其均方根误差、相关系数等指标都得到提升;在运行效率方面,具有一个适中的计算复杂度,远低于性能优越的度量指标的运行时间。

关键词: 图像质量评价;人眼视觉系统;结构相似性指标;最小可辨失真;显著性图像

中图分类号: TN911.73

文献标志码: A

文章编号: 0367-6234(2018)05-0121-08

Image quality assessment: Combining the characteristics of HVS and structural similarity index

ZHU Xinshan^{1,2}, YAO Siru¹, SUN Biao¹, QIAN Yongjun¹

(1. School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China; 2. State Key Laboratory of Information Security (Institute of Information Engineering, Chinese Academy of Sciences), Beijing 100093, China)

Abstract: Image quality assessment is a basic problem in the multimedia field. A new image quality metric is constructed based on Structural Similarity (SSIM) index, by exploiting Human Visual System (HVS) characteristics. First, considering the masking effects of HVS, the distorted image is preprocessed by a distortion model with the input of the error between it and the original one and just noticeable distortions (JNDs) derived from a human visual model, where Sigmoid function is explored. The process makes the visible errors in the modified distorted image to be more notable. Second, considering the visual attention characteristics of HVS, the image area weight model is designed to quantify the importance of image local regions for the visual quality. The interesting content of an image can be represented by the saliency image, from which the weights of different regions are obtained. Finally, the local SSIM between the modified distorted image and original image is calculated, and the global image quality metric can be expressed by weighting all local quality with the normalized regional weights. Compared with the state-of-the-art image metrics, the proposed metric fits subjective visual quality better in evaluating the local image quality, has better performance in terms of the mean square error, correlation coefficient and other indicators for predicting the subjective image quality, and has a moderate computational complexity, well below the run time of the superior performance metrics.

Keywords: image quality assessment; human visual system; structural similarity index; just noticeable distortion; saliency image

图像质量评价(Image Quality Assessment, IQA)在很多实际应用中起着至关重要的作用,例如图像增强^[1]、编码^[2]、采集^[3]、水印^[4]和压缩^[5]等。IQA 不仅用于比较和评价图像处理系统和算法,而且它还可嵌

入到图像处理算法中以优化其性能。因此,IQA 已经成为一个非常重要的研究课题,受到广泛关注。

目前,一类具有代表性的方法是文献[6]提出的基于结构相似性(Structural Similarity, SSIM)的 IQA 度量指标。根据人眼能从图像中提取结构信息的理论,SSIM 指标利用均值、方差和协方差衡量图像局部亮度、对比度和结构相似性,进而度量图像整体的相似性,其评价结果与人主观感知较为一致。然而,SSIM 仍然有一定的局限性。为增强客观图像

收稿日期: 2017-09-26

基金项目: 国家自然科学基金(61401303;51578189);国家留学基金(201506255067);信息安全国家重点实验室开放课题(2017-MS-11);天津大学自主创新基金(1705)

作者简介: 朱新山(1977—),男,副教授

通信作者: 孙彪, sunbiao@tju.edu.cn

质量评价的准确性,研究者在 SSIM 指标基础上进行改进,涌现出大量改进 SSIM 的指标。

由于人眼对于图像边缘信息比较敏感,且图像边缘对于理解图像信息至关重要。研究者提出了很多基于梯度的 SSIM 改进指标^[7-12]。文献[7]提出基于梯度幅值相似性偏差的图像质量评价指标,对梯度幅值相似性使用标准差的方式进行池化。文献[10]利用图像的相位一致特征和梯度幅值特征,提出基于特征相似指数(FSIM)的图像质量评价指标。文献[11]提出基于梯度相似性(GSM)指标,将亮度和梯度相似性相结合。文献[12]将加权的梯度信息与 SSIM 相结合,可有效提高预测的准确性。文献[13]提出基于内在生成机制的方法,将图像分成有序区域和无序区域进行质量预测,采用多尺度和自适应权重结合两区域质量得到最终的质量预测值。文献[14]利用结构张量、对比度掩蔽和邻域掩蔽,对 SSIM 指标的结构比较函数和对比度比较函数进行改进。为有效地对几何变换图像的质量进行评价,文献[15]利用均值分段的标准差相关函数对 SSIM 的结构比较函数进行改进,并且考虑图像锐度的影响,引入锐度比较函数。文献[16]证明了权重应与局部信息内容成比例,提出基于信息内容加权的结构相似性度量指标(IWSSIM)。

在现有的 SSIM 改进方案中,没有充分考虑人眼掩蔽效应、视觉选择注意机制对图像质量评价的影响。为此,本文结合最小可辨失真(just noticeable distortion, JND)和视觉显著性评价设计了改进的 SSIM。首先,利用 JND 模型进行原始图像的视觉阈值分析,并结合绝对误差图和 Sigmoid 函数进行失真图像的修正;然后,利用视觉显著性模型提取子区域的权重;最后,结合 SSIM 指标对各子区域加权得到最终的图像质量。在 LIVE2 和 CSIQ 数据库上对本文指标进行测试,并与当前多种具有代表性的指标进行广泛比较。

1 基本的 SSIM

SSIM 指标使用图像块结构信息的变化量衡量图像质量,其基本思想是结合参考图像和失真图像之间的局部亮度、对比度和结构相似性得到图像质量。给定 $\mathbf{x}$ 和 $\mathbf{y}$ 是参考图像和失真图像相同位置处对应的图像块,其大小为 $M \times N$ (高 $\times$ 宽)。

局部亮度相似性的度量使用图像局部平均值信息,对图像块 $\mathbf{x}$ 平均值定义如下:

$$\mu_x = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N x(i, j). \quad (1)$$

式中 $x(i, j)$ 为图像 $\mathbf{x}$ 在点 (i, j) 处的像素值。

局部对比度相似性的估计使用图像局部标准差

信息,对图像块 $\mathbf{x}$ 标准差定义如下:

$$\sigma_x = \left[\frac{1}{MN-1} \sum_{i=1}^M \sum_{j=1}^N (x(i, j) - \mu_x)^2 \right]^{1/2}. \quad (2)$$

结构相似性使用图像块 $\mathbf{x}$ 和 $\mathbf{y}$ 的相关性进行度量,其定义为

$$\sigma_{xy} = \frac{1}{MN-1} \sum_{i=1}^M \sum_{j=1}^N (x(i, j) - \mu_x)(y(i, j) - \mu_y). \quad (3)$$

利用上述定义,局部亮度相似性 $l(\mathbf{x}, \mathbf{y})$ 、对比度相似性 $c(\mathbf{x}, \mathbf{y})$ 和结构相似性 $s(\mathbf{x}, \mathbf{y})$ 为

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (4)$$

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (5)$$

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}. \quad (6)$$

式中: μ_y 、 σ_y 分别为图像块 $\mathbf{y}$ 的平均值和标准差,其定义与 μ_x 、 σ_x 类同;常数 $C_i \geq 0 (i = 1, 2, 3)$, 其数值一般较小,用以保持上述各式的稳定性。

参考图像和失真图像之间的 SSIM 是图像局部质量的平均值,表达为

$$\text{SSIM} = \frac{1}{K} \sum_{i=1}^K [l(\mathbf{x}_i, \mathbf{y}_i)]^\alpha [c(\mathbf{x}_i, \mathbf{y}_i)]^\beta [s(\mathbf{x}_i, \mathbf{y}_i)]^\gamma. \quad (7)$$

式中: $\mathbf{x}_i$ 和 $\mathbf{y}_i$ 分别是参考图像和失真图像的第 i 个图像块; K 是图像块的个数; $\alpha > 0$ 、 $\beta > 0$ 和 $\gamma > 0$ 是各个分量的权重,用以调整相对重要性。

虽然 SSIM 在度量图像失真时考虑了像素点与其邻域像素之间的相关性,描述了图像结构信息,但 SSIM 指标仍然存在缺陷。首先,SSIM 没有考虑 HVS 掩蔽效应对视觉感知的影响。其次,SSIM 指标平等对待图像中视觉感兴趣程度不同的区域,忽略了视觉选择注意机制对图像质量评价的影响。因而,SSIM 的性能受到极大的限制。

2 基于 JND 模型和视觉显著性评价的图像相似性度量

为改进基本的 SSIM,本文引进 JND 和视觉显著性评价,设计方案见图 1。

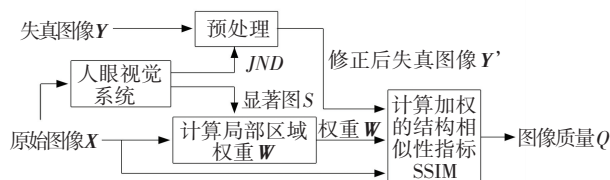


图 1 改进的 IQA 系统流程图

Fig.1 The improved IQA system

从图1可看出,本文的设计方案与原始SSIM的区别在于:失真图像需先经过一个预处理得到修正后的失真图像,且对图像的局部质量进行加权得到最终的图像质量评价指标。

本文利用JND模型和Sigmoid函数对失真图像进行预处理,以保留人眼可见失真,使图像的失真更符合人的主观感受。利用视觉显著性模型提取原始图像的显著图,以模拟图像不同区域的受关注程度。相同的失真发生在图像显著区域与非显著区域的影响不同,因此在进行图像局部质量融合时应对不同区域施加不同的权重。

3 基于JND模型的失真图像修正

JND模型在图像视觉质量评价的研究中具有举足轻重的作用。使用JND模型可得出每个像素点的视觉阈值,并从而修正图像,使得失真图像更能反映图像的可见失真量。

3.1 JND模型

文献[17]中给出一种空间域的JND模型。该模型采用亮度掩蔽和对比度掩蔽的非线性加权来计算JND值。

亮度掩蔽阈值取决于图像 $\mathbf{x}$ 各像素邻域的加权亮度值,像素邻域加权亮度值为

$$\bar{\mathbf{x}} = \frac{1}{\|\mathbf{B}\|_1} (\mathbf{x} * \mathbf{B}). \quad (8)$$

式中: $\|\cdot\|_1$ 为1范数; $\mathbf{B}$ 是一个 5×5 的低通滤波器;符号 $*$ 为卷积。

对比度掩蔽阈值的度量利用图像的梯度信息,文献[17]中采用像素邻域4个方向中最大的平均加权梯度

$$\mathbf{G} = \max_{k=1,2,3,4} \left\{ \frac{1}{2 \cdot \|\mathbf{g}_k\|_1} (\mathbf{x} * \mathbf{g}_k) \right\}. \quad (9)$$

式中 $\mathbf{g}_k (k=1,2,3,4)$ 是 5×5 的高通滤波器。

不同区域的视觉阈值应取不同值。因此,在对比度掩蔽阈值的度量中加入图像的边缘权重系数 $\mathbf{W}_e$,以调整图像边缘区域与非边缘区域的对比度阈值。即

$$\mathbf{W}_e = \mathbf{L} * \mathbf{h}. \quad (10)$$

式中: $\mathbf{L}$ 是图像边缘图; $\mathbf{h}$ 是高斯低通滤波器。

利用上述定义,图像亮度掩蔽阈值 T^l 和对比度掩蔽阈值 T^c 定义如下:

$$T^l(i,j) = \begin{cases} 17(1 - \sqrt{\frac{\bar{x}(i,j)}{127}}) + 3, & \bar{x}(i,j) \leq 127; \\ \frac{3}{128}(\bar{x}(i,j) - 127) + 3, & \bar{x}(i,j) > 127. \end{cases} \quad (11)$$

$$T^c(i,j) = \beta \cdot G(i,j) \cdot W_e(i,j); \quad (12)$$

式中: β 是控制参数; $\bar{x}(i,j)$ 、 $G(i,j)$ 和 $W_e(i,j)$ 分别为 $\bar{x}$ 、 G 和 W_e 在坐标 (i,j) 处的值。

结合亮度掩蔽阈值和对比度掩蔽阈值得到图像像素JND阈值

$$T(i,j) = T^l(i,j) + T^c(i,j) - C \cdot \min(T^l(i,j), T^c(i,j)). \quad (13)$$

式中 C 为亮度通道所占的重叠效应。

式(8)~(10)中的 $\mathbf{B}$ 、 $\mathbf{g}_k$ 、 $\mathbf{L}$ 和 $\mathbf{h}$ 的设计详见文献[17]。

3.2 失真图像的修正

由于亮度掩蔽和对比度掩蔽效应,并不是图像中所有的变化都能引起人眼的感知。图像失真程度在JND阈值以内,人眼无法感知,并且对图像相似性评价的影响很小甚至没有。相反,图像失真量超过JND阈值,对图像相似性的评价产生较大的影响。

利用JND模型建立预处理滤波器,对失真图像进行滤波预处理,移除人眼无法感知的失真,只保留可见失真更符合主观感知。该过程如下:

Step 1: 计算原始图像 $\mathbf{X}$ 中每个像素的视觉阈值 T 。利用式(13)进行视觉阈值的计算。

Step 2: 失真图像修正。为使图像质量评价更符合人眼的主观感知,对失真图像 $\mathbf{Y}$ 的每个像素进行修正

$$Y'(i,j) = X(i,j) \cdot \mathbf{1}\{|D(i,j)| \leq T(i,j)\} + [Y(i,j) - \text{sign}(D(i,j)) \cdot \lambda(i,j) \cdot T(i,j)] \cdot \mathbf{1}\{|D(i,j)| > T(i,j)\}. \quad (14)$$

式中: $\mathbf{1}(\cdot)$ 是示性函数; $|\cdot|$ 为绝对值; $\mathbf{Y}'$ 是修正之后的失真图像;误差 $D(i,j) = X(i,j) - Y(i,j)$; $\text{sign}(\cdot)$ 是符号函数; $\lambda(i,j) \geq 0$ 是惩罚因子。

从式(14)中可看出,当绝对误差小于JND值时,修正后的失真图像与原始图像的像素值相同。由于失真区域易受到人眼关注,相应的图像主观评价质量会降低。所以,绝对误差大于JND值时,失真图像取决于误差和JND值,增大误差,减小图像质量评价。

惩罚因子 $\lambda(i,j)$ 对于图像质量评价具有重要作用, $\lambda(i,j)$ 与图像每个像素的感知误差 $R(i,j) = |D(i,j)|/T(i,j)$ 有关。本文设计 $\lambda(i,j)$ 为Sigmoid函数,表达为

$$\lambda(i,j) = \frac{1}{1 + \exp(-R(i,j))}. \quad (15)$$

$\lambda(i,j)$ 与 $R(i,j)$ 之间的关系如图2所示,可看出,惩罚因子随感知误差的增大而逐渐趋近于1,这表示可见性误差的影响增加了。

使用修正的失真图像计算SSIM,可改善IQA的性能。具体实例如图3和图4所示。

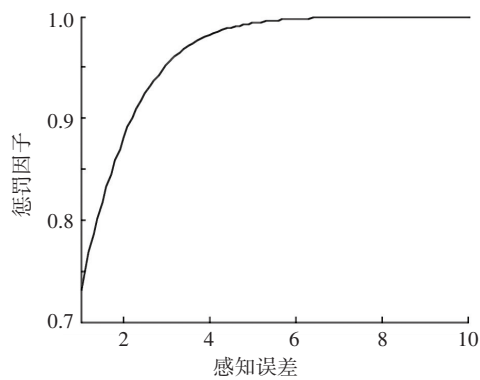


图 2 惩罚因子 $\lambda(i, j)$ 与感知误差 $R(i, j)$ 的关系
Fig.2 The relationship between penalty factor $\lambda(i, j)$ and perceptual error $R(i, j)$

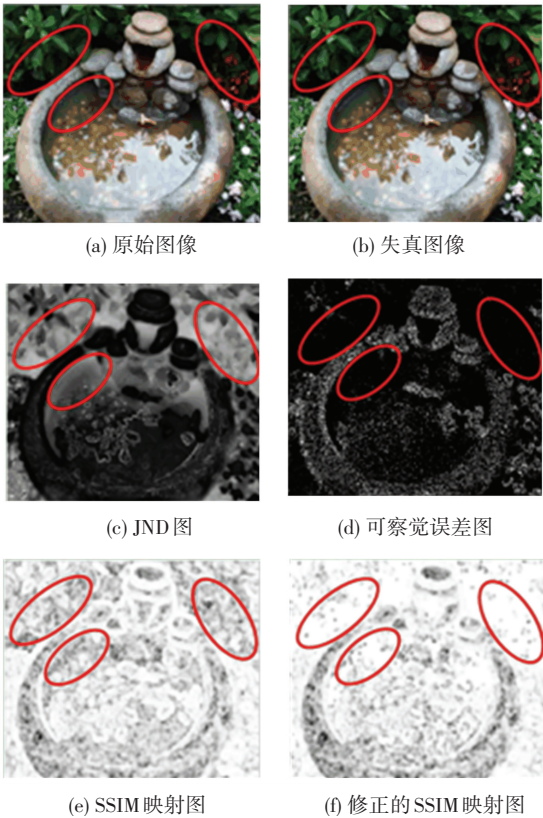


图 3 基于 JND 的失真图像修正“coinsinfountain”在图像上的有效性
Fig.3 The effectiveness of JND - based distortion image modification on “coinsinfountain” image

图 3(a) 为原始图像;图 3(b) 为失真图像;图 3(c) 为基于 JND 模型计算的原始图像的 JND 图;图 3(d) 为图 3(a) 和图 3(b) 之间的可察觉误差,图像中的黑色部分不会影响视觉;图 3(e) 是图 3(a) 和图 3(b) 之间的 SSIM 映射图;图 3(f) 是图 3(a) 与修正失真图之间的 SSIM 映射图. 映射图中亮度越高则表明原始图像与失真图像在此处的相似性越高,失真图像的质量损失越少,图像质量越好.

图 3 标记区域(椭圆)为纹理区域,由于对比度掩蔽效应,人眼对于纹理区域的敏感度较低. 从图中标记区域可以看出,其可察觉误差很低(图 3(d)),图像质量映射图的相应区域应比较明亮. 然而图 3(e) 对应区域亮度暗淡,与人的视觉判断相反. 图 3(f) 可看出对应区域很明亮,与人眼判断一致.

图 4 给出在另外一幅图像上显示的例子,其各子图与图 3 中相应子图的表示含义相同. 图 4 标记区域背景亮度较低,人眼对此处的失真敏感度较低,且从图 4(b) 可看出此区域的失真不可见. 同理,从图 4 可做出与图 3 类似的结论.

以上实例说明,利用 JND 和 Sigmoid 函数对失真图像进行修正可消除不可察觉误差,使可见失真更符合人眼主观感知,提高度量性能.

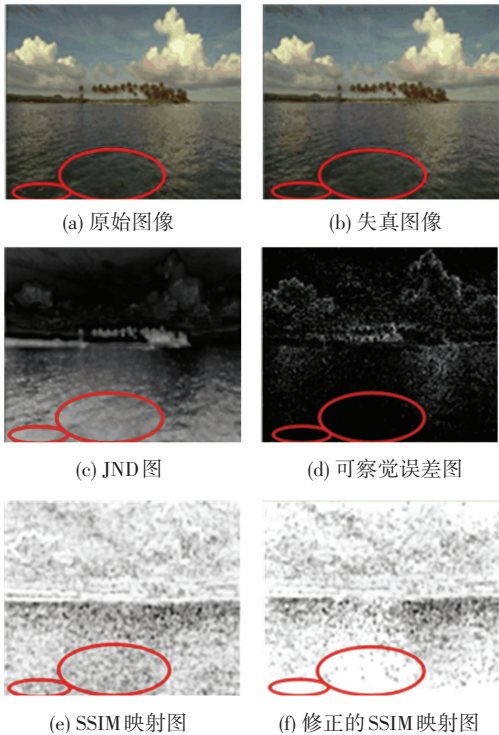


图 4 基于 JND 的失真图像修正在“ocean”图像上的有效性
Fig.4 The effectiveness of JND - based distortion image modification on “ocean” image

4 基于视觉显著性评价的局部质量融合

人类视觉注意机制通常会关注图像中重要的刺激而抑制不重要的信息,所以还需考虑局部显著性对于图像质量的影响,根据受关注的程度而区别对待图像局部区域. 图 5 展示了不同图像的显著性效果,图 5(a) 和 5(b) 为原图,其相应的显著性图如图 5(c) 和 5(d) 所示. 可看出,显著性图像越亮的部分,其显著性越高,即关注度越高,对图像质量的影响更大. 本文利用图像的局部显著性来改善图像质

量评价指标.

4.1 确定局部区域权重

本文采用基于相位谱的四元数傅里叶变换(PQFT)模型^[18]对图像的显著性进行度量. PQFT模型将图像每个像素用颜色、亮度和运动向量进行表示,通过计算其傅里叶变换的相位谱提取图像的显著性区域. 提取图像显著性时,运动向量为零.

利用 PQFT 模型处理原始图像 X 得到显著图 S . 按照图像质量评价指标中图像块的划分方法,设计尺寸相同的均值滤波器,对显著图 S 进行均值滤波,得到滤波后的显著图 W . 将 W 中每个坐标点的数值 w_i 作为区域权重,即利用显著图 S 中各子区域的平均显著值表征图像区域权重.

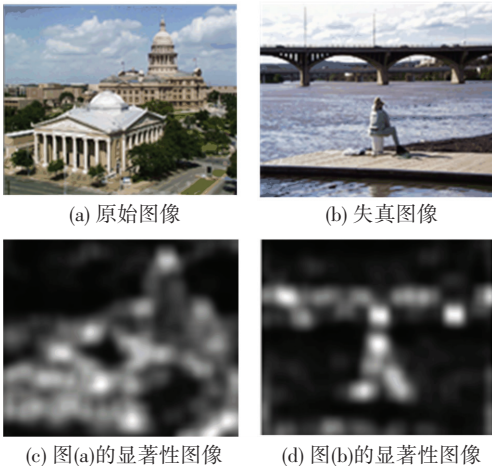


图5 图像显著性检测结果

Fig.5 Image saliency detection results

4.2 融合

将图像各子区域的权重进一步进行归一化处理,得到子区域的相对权重为

$$w'_i = \frac{w_i}{\sum_{i=1}^K w_i}. \quad (16)$$

式中 K 是显著图的分块数.

选取 SSIM 指标度量原始图像 X 与修正失真图像 Y' 之间相对应子块的质量 q_i

$$q_i = l(x_i, y'_i) \cdot c(x_i, y'_i) \cdot s(x_i, y'_i). \quad (17)$$

利用图像的显著性信息对图像各局部质量进行线性加权,得到整幅图像的客观评价价值 Q

$$Q = \sum_{i=1}^K w'_i \cdot q_i. \quad (18)$$

式中 x_i 和 y'_i 分别是原始图像和修正失真图像的第 i 个图像块.

从式(18)中可看出,权重 w_i 越大的区域对图像质量的影响越显著,这比基本的 SSIM 把所有区域视为对图像质量具有相同的影响更合理. 另外,

通过实行式(16)的权重归一化处理,可确保客观评价价值 Q 取值在 0 和 1 之间.

5 实验结果与分析

5.1 测试数据库与评价性能指标

为验证图像质量评价指标的性能,本文采用 LIVE2^[19]和 CSIQ^[20]图像数据库进行仿真实验,并和已有的指标进行比较. LIVE2 图像库包含 29 幅原始图像和 779 幅对应的失真图像,共有 5 种失真类型,其中 JPEG2000 失真(jpeg2k)175 幅, JPEG 失真(jpeg)169 幅,白噪声失真(wn)145 幅,高斯模糊失真(gblur)145 幅以及快速瑞利衰减失真(fastfading)145 幅. CSIQ 图像库包含 30 幅原始图像和 866 幅对应的失真图像,共有 6 种失真类型,分别是加性高斯白噪声、加性高斯彩色噪声、JPEG 失真、JPEG2000 失真、高斯模糊和全局对比度退化. 同时, LIVE2 和 CSIQ 图像库都给出每幅失真图像相应的差分主观分值(DMOS),DMOS 值越小表示失真图像的主观质量越好.

本文指标与 7 种具有代表性的全参考 IQA 指标进行性能比较,包括 PSNR、SSIM^[6]、IFC^[21]、VIF^[22]、FSIM^[10]、IWSSIM^[16]和 GSM^[11]. 其中,PSNR 是峰值信噪比;IFC 和 VIF 是基于信息保真度原则的评价指标;IWSSIM 是采用信息内容加权的 SSIM 算法;FSIM 和 GSM 是基于 SSIM 结构化方法构建的算法,其思想分别为结合相位一致相似性和梯度相似性,利用图像的梯度相似性和亮度相似性.

本文采用以下 4 种参数对 IQA 指标性能进行评价:均方根误差(RMSE)、Pearson 相关系数(PLCC)、Spearman 秩相关系数(SROCC)和 Kendall 秩相关系数(KROCC). RMSE 和 PLCC 用于评估 IQA 指标的预测准确性,SROCC 和 KROCC 用于评估 IQA 指标的预测单调性. RMSE 值越小,SROCC、KROCC 和 PLCC 值越接近于 1,则表明客观评价结果越与主观感知一致.

为客观反映图像质量评价指标的性能,需对客观评价价值 Q 和 DMOS 值进行非线性回归处理. 本文采用 5 参数 Logistic 函数

$$f(Q) = \beta_1 \cdot \left(\frac{1}{2} - \frac{1}{1 + \exp(\beta_2 \cdot (Q - \beta_3))} \right) + \beta_4 \cdot Q + \beta_5, \quad (19)$$

式中 $\beta_i (i = 1, 2, \dots, 5)$ 是拟合参数.

5.2 感知性能评价

为验证本文指标的感知性能,图 6 给出原始图像及其对应的不同类型的失真图像. 图 6(b)~(f)的 DMOS 值分别为 25.743 2、28.869 1、48.072 7、52.573 9、58.507 7;SSIM 值分别为 0.958 5、0.963 7、

0.815 1、0.541 4、0.679 2; FSIM 值分别为 0.976 9、0.981 5、0.926 7、0.867 4、0.836 4; 本文指标评价价值分别为 0.961 5、0.945 0、0.792 3、0.680 3、0.655 6. 不同指标的客观评价价值应随 DMOS 的增大而减小. 但只有本文指标的评价价值的变化趋势与主观评价价值一致. 这说明本文指标可有效度量图像的失真量, 且与视觉效果更一致.

5.3 单一失真图像上的比较

为全面地评估 IQA 指标预测失真引起的图像质量变化的能力, 计算本文指标在不同失真类型下的性能评价指标, 并与其他 7 个最具代表性的 IQA 指标进行比较. 表 1 列出了不同 IQA 指标在 LIVE2 数据库上不同失真类型的性能比较结果. 对于每一个不同失真类型的性能评价指标, 用粗体标出三个表现最好的指标. 从表 1 可看出, 本文指标与 VIF、FSIM 指标性能相当, 性能指标位于前三的次数分别为 13 次、10 次和 13 次, 优于其他 IQA 指标. 与 FSIM、VIF 相比, 本文指标更适合处理 wn 和 fastfading 失真, jpeg 和 gblur 失真次之, 而其他类型失真的性能指标只是略微低于表现最好的指标. 此外, 在不同类型失真情况下, 本文指标具有更严格的

预测单调性, 其 SROCC 和 KROCC 值共有 8 次位于前三, 多于其他 IQA 指标.

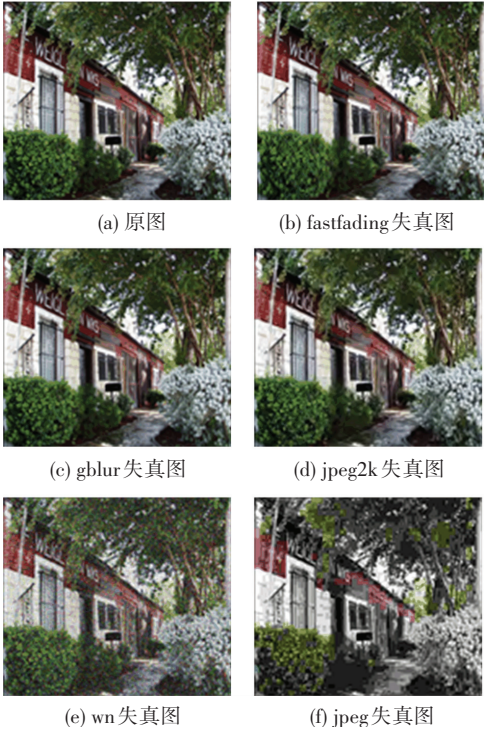


图 6 测试图像及其失真图像

Fig.6 The test image and its distorted images

表 1 不同 IQA 指标对 LIVE2 数据库中不同失真类型的评价结果比较

Tab.1 Comparison of IQA indices for each distortion type on LIVE2 database

评价指标	失真类型	PSNR	SSIM	IFC	VIF	FSIM	IWSSIM	GSM	Proposed
RMSE	jpeg2k	7.475	4.714	6.971	4.350	4.276	4.658	4.278	4.380
	jpeg	8.169	5.296	6.812	5.357	5.232	5.402	5.072	5.366
	wn	2.640	2.940	4.459	2.458	3.770	3.045	3.185	2.638
	gblur	9.761	4.990	4.356	3.582	3.571	3.380	5.131	3.732
	fastfading	7.511	4.867	4.123	7.928	5.310	5.419	5.940	4.521
PLCC	jpeg2k	0.887	0.957	0.903	0.963	0.965	0.958	0.965	0.963
	jpeg	0.860	0.944	0.905	0.942	0.945	0.941	0.948	0.942
	wn	0.986	0.983	0.960	0.988	0.972	0.982	0.980	0.986
	gblur	0.784	0.948	0.961	0.974	0.974	0.977	0.945	0.971
	fastfading	0.890	0.955	0.968	0.876	0.946	0.944	0.933	0.961
SROCC	jpeg2k	0.890	0.953	0.892	0.956	0.957	0.950	0.959	0.956
	jpeg	0.841	0.912	0.866	0.885	0.923	0.882	0.903	0.913
	wn	0.985	0.969	0.938	0.985	0.965	0.967	0.977	0.976
	gblur	0.782	0.952	0.959	0.968	0.971	0.972	0.952	0.969
	fastfading	0.890	0.955	0.963	0.867	0.950	0.944	0.940	0.960
KROCC	jpeg2k	0.704	0.805	0.697	0.816	0.816	0.8	0.821	0.815
	jpeg	0.637	0.741	0.676	0.689	0.773	0.690	0.730	0.750
	wn	0.894	0.852	0.778	0.898	0.839	0.848	0.869	0.871
	gblur	0.585	0.801	0.830	0.84	0.853	0.854	0.806	0.846
	fastfading	0.706	0.820	0.839	0.704	0.817	0.809	0.792	0.834

5.4 整体性能的比较

本文指标与不同 IQA 指标在 LIVE2 数据库上交失真实验的散点图如图 7 所示,横轴是 IQA 指标预测的客观评价分,纵轴是 DMOS 值,图中曲线由式 (19) 非线性拟合得到. 从图 7 中可看出,与其他 IQA 指标相比,本文指标散点图中的点相对均匀地分布在拟合曲线两侧,且拟合结果更接近于线性. 因此,本文指标预测的客观评价值与主观评价值更一致.

不同 IQA 指标在 LIVE2 数据库上交失真性能比较结果见表 2. 对于每一个性能评价指标,用粗体标出三个表现最好的指标. 本文指标 RMSE 为 6.030 6, PLCC 为 0.9272, SROCC 为 0.924 5, KROCC 为 0.757 6. 与相比较的 IQA 指标相比,其 RMSE 值最小,PLCC、SROCC 和 KROCC 值最大. 所以,本文指标在整体上性能优于其他 IQA 指标,具有更高的预测准确性和更严格的预测一致性,与人的主观感知更为一致.

不同指标在 CSIQ 数据库上的性能测试结果如表 3 所示. 对于每一个性能评价指标,仍然用粗体指标出三个表现最好的指标. 从表中可看出,本文指标在 CSIQ 数据库中依然具有很好的评价性能,其 RMSE 值最小,PLCC、SROCC 和 KROCC 值最大. 这些测试足以说明本文指标是有效的.

5.5 运行效率

为评价 IQA 指标的运行效率,在 LIVE2 数据库上运行每个被测试的 IQA 指标,通过多次运行,获得每幅图像客观评价指标计算所需时间的平均值. 实验硬件平台是 i3-3240CPU@3.40GHz 和 8G 内存的台式机,所用软件为 MATLAB R2014a,每幅图像的大小为 632×505 像素. 从表 4 可看出,本文指标的运行时间是 0.126 秒,处于所比较指标的中游. 虽然运行时间长于 SSIM,但远低于 FSIM、IWSSIM 等性能较好指标的运行时间. 因此,本文提出的 IQA 指标获得了令人满意的运行效率.

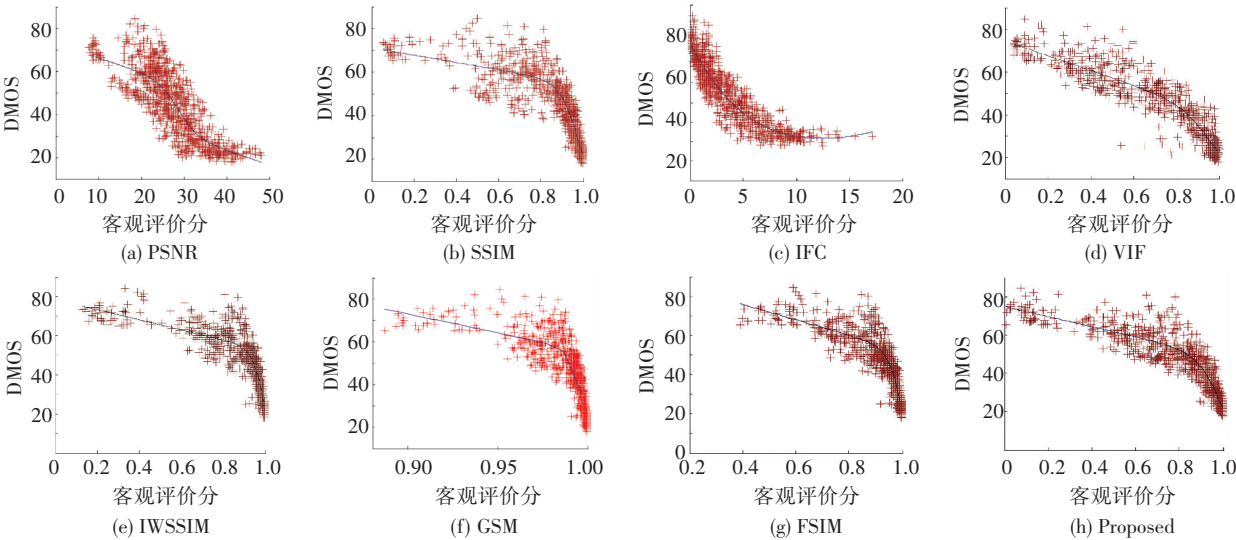


图 7 不同 IQA 指标在 LIVE2 数据库中的散点图

Fig.7 Scatter plot for different IQA indices on LIVE2 database

表 2 不同 IQA 指标在 LIVE2 数据库中整体性能比较

Tab.2 Performance comparison of different IQA indices on LIVE2 database

评价指标	PSNR	SSIM	IFC	VIF	FSIM	IWSSIM	GSM	Proposed
RMSE	9.087 1	6.724 0	6.641 4	6.257 5	6.181 0	6.135 7	6.509 6	6.030 6
PLCC	0.825 6	0.908 7	0.911 0	0.921 4	0.923 4	0.924 6	0.914 7	0.927 2
SROCC	0.819 7	0.899 9	0.912 8	0.922 0	0.922 6	0.921 4	0.912 0	0.924 5
KROCC	0.617 1	0.718 3	0.733 6	0.752 6	0.753 4	0.753 2	0.735 2	0.757 6

表 3 不同 IQA 指标在 CSIQ 数据库中整体性能比较

Tab.3 Performance comparison of different IQA indices on CSIQ database

评价指标	PSNR	SSIM	IFC	VIF	FSIM	IWSSIM	GSM	Proposed
RMSE	0.157 6	0.133 4	0.143 8	0.098	0.107 7	0.106 3	0.116 4	0.095 1
PLCC	0.799 8	0.861 3	0.836 6	0.927 7	0.912 0	0.914 4	0.896 4	0.934 4
SROCC	0.800 5	0.875 6	0.767 1	0.919 5	0.924 2	0.921 2	0.910 7	0.938 0
KROCC	0.598 4	0.690 7	0.589 2	0.753 7	0.756 1	0.752 9	0.736 8	0.777 7

表 4 不同 IQA 指标执行时间比较
Tab.4 Comparisons of computation time

IQA 指标	PSNR	SSIM	IFC	VIF	FSIM	IWSSIM	GSM	Proposed
时间/s	0.005	0.024	1.114	0.24	0.514	0.523	0.03	0.126

6 结 论

本文提出了一种基于最小可辨失真和视觉显著性评价的图像质量评价指标. 在该方案中,利用空间域 JND 模型与 Sigmoid 函数对失真图像进行修正. 通过 JND 修正失真图像使其保留人眼可见失真,通过 Sigmoid 函数设计感知误差的惩罚系数,用于改进局部图像质量的度量. 最后,利用视觉显著性模型提取图像中各子区域的权重,根据权重融合局部图像质量得到最终的图像质量评价结果. 在 LIVE2 和 CSIQ 图像数据库上与其他最具代表性的 IQA 指标进行了比较. 结果表明,本文所提指标与人的主观感知质量具有更好的一致性,运行效率较高.

参考文献

[1] GU Ke, ZHAI Guangtao, YANG Xiaokang, et al. Automatic contrast enhancement technology with saliency preservation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25(9): 1480–1494. DOI: 10.1109/TCSVT.2014.2372392.

[2] TIAN Xin, LI Tao, TIAN Jinwen, et al. Prediction method for image coding quality based on differential information entropy [J]. Entropy, 2014, 16(2): 990–1001. DOI: 10.3390/e16020990.

[3] KEDZIERSKI M, WIERZBICHI D. Radiometric quality assessment of images acquired by UAV's in various lighting and weather conditions [J]. Measurement, 2015, 76: 156–169. DOI: 10.1016/j.measurement.2015.08.003.

[4] PAPAKOSTAS G A, TSOUGENIS E D, KOULOURIOTIS D E. Moment-based local image watermarking via genetic optimization [J]. Applied Mathematics and Computation, 2014, 227: 222–236. DOI: 10.1016/j.amc.2013.11.036.

[5] ZIMBICO A, SCHNEIDER F, MAIA J. Comparative study of the performance of the JPEG algorithm using optimized quantization matrices for ultrasound image compression [C]// ISSNIP Biosignals and Biorobotics Conference. Piscataway NJ: IEEE, 2014: 89–94. DOI: 10.1109/BRC.2014.6880973.

[6] WANG Zhou, 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. DOI: 10.1109/TIP.2003.819861.

[7] XUE Wufeng, ZHANG Lei, MOU Xuanqin, et al. Gradient magnitude similarity deviation: a highly efficient perceptual image quality index [J]. IEEE Transactions on Image Processing, 2014, 23(2): 684–695. DOI: 10.1109/TIP.2013.2293423.

[8] CHEBBI E, BENZARTI F, AMIRI H. An improvement of structural similarity index for image quality assessment [J]. Journal of Computer Science, 2014, 10(2): 353–360. DOI: 10.3844/jcssp.2014.353.360.

[9] WANG Zhengyou, LI Liying, WU Shuang, et al. A new image qual-

ity assessment algorithm based on ssim and multiple regressions [J]. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2015, 8(11): 221–230. DOI: 10.14257/ij-sip.2015.8.11.20.

[10] ZHANG Lin, ZHANG Lei, MOU Xuanqin, et al. FSIM: A feature similarity index for image quality assessment [J]. IEEE Transactions on Image Processing, 2011, 20(8): 2378–2386. DOI: 10.1109/TIP.2011.2109730.

[11] LIU Anmin, LIN Weisi, NARWARIA M. Image quality assessment based on gradient similarity [J]. IEEE Transactions on Image Processing, 2012, 21(4): 1500–1512. DOI: 10.1109/TIP.2011.2175935.

[12] LI Qiaohong, FANG Yuming, LIN Weisi, et al. Gradient-weighted structural similarity for image quality assessments [C]// IEEE International Symposium on Circuits and Systems. Piscataway NJ: IEEE, 2015: 2165–2168. DOI: 10.1109/ISCAS.2015.7169109.

[13] WU Jinjian, LIN Weisi, SHI Guangming, et al. Perceptual quality metric with internal generative mechanism [J]. IEEE Transactions on Image Processing, 2013, 22(1): 43–54. DOI: 10.1109/TIP.2012.2214048.

[14] FEI Xuan, XIAO Liang, SUN Yubao, et al. Perceptual image quality assessment based on structural similarity and visual masking [J]. Signal Processing-Image Communication, 2012, 27(7): 772–783. DOI: 10.1016/j.image.2012.04.005.

[15] LEE D, LIM S. Improved structural similarity metric for the visible quality measurement of images [J]. Journal of Electronic Imaging, 2016, 25(6): 063015. DOI: 10.1117/1.JEI.25.6.063015.

[16] WANG Zhou, LI Qiang. Information content weighting for perceptual image quality assessment [J]. IEEE Transactions on Image Processing, 2011, 20(5): 1185–1198. DOI: 10.1109/TIP.2010.2092435.

[17] YANG Xiaokang, LIN Weisi, LU Z, et al. Motion-compensated residue preprocessing in video coding based on just-noticeable-distortion profile [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(6): 742–752. DOI: 10.1109/TCSVT.2005.848315.

[18] GUO Chenlei, MA Qi, ZHANG Liming. Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform [C]// Proceedings-IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway NJ: IEEE, 2008: 2908–2915. DOI: 10.1109/CVPR.2008.4587715.

[19] SHEIKH H R, WANG Zhou, CORMACK L, et al. LIVE image quality assessment database release 2 [EB/OL]. [2006–11] <http://live.ece.utexas.edu/research/quality/subjective.htm>.

[20] LARSON E C, CHANDLER D M. Categorical Image Quality (CSIQ) database 2009 [EB/OL]. [2009–12–14] <http://vision.eng.shizuoka.ac.jp/course/view.php?id=6>.

[21] SHEIKH H R, BOVIK A C, VECIANA G D. An information fidelity criterion for image quality assessment using natural scene statistics [J]. IEEE Transactions on Image Processing, 2005, 14(12): 2117–2128. DOI: 10.1109/TIP.2005.859389.

[22] SHEIKH H R, BOVIK A C. Image information and visual quality [J]. IEEE Transactions on Image Processing, 2006, 15(2): 430–444. DOI: 10.1109/TIP.2005.859378. (编辑 苗秀芝)