

基于非局部梯度的图像质量评价算法

高敏娟^① 党宏社^① 魏立力^② 张选德^{*①}

^①(陕西科技大学电气与信息工程学院 西安 710021)

^②(宁夏大学数学统计学院 银川 750021)

摘 要: 图像质量评价研究的目的在于模拟人类视觉系统对图像质量的感知过程, 构建与主观评价结果尽可能一致的客观评价算法。现有的很多算法都是基于局部结构相似设计的, 但人对图像的主观感知是高级的、语义的过程, 而语义信息本质上是非局部的, 因此图像质量评价应该考虑图像的非局部信息。该文突破了经典的基于局部信息的算法框架, 提出一种基于非局部信息的框架, 并在此框架内构建了一种基于非局部梯度的图像质量评价算法, 该算法通过度量参考图像与失真图像的非局部梯度之间的相似性来预测图像质量。在公开测试数据库TID2008, LIVE, CSIQ上的数值实验结果表明, 该算法能获得较好的评价效果。

关键词: 图像质量评价; 人类视觉系统; 非局部梯度

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2019)05-1122-08

DOI: 10.11999/JEIT180597

Image Quality Assessment Algorithm Based on Non-local Gradient

GAO Minjuan^① DANG Hongshe^① WEI Lili^② ZHANG Xuande^①

^①(College of Electrical and Information Engineering, Shaanxi University of
Science & Technology, Xi'an 710021, China)

^②(School of Mathematics and Statistics, Ningxia University, Yinchuan 750021, China)

Abstract: The goal of Image Quality Assessment (IQA) research is to simulate the Human Visual System's (HVS) perception process of assessing image quality and construct an objective evaluation algorithm that is as consistent as the subjective evaluation result. Many existing algorithms are designed based on local structural similarity, but human subjective perception of images is a high-level, semantic process, and semantic information is essentially non-local, so image quality assessment should take the non-local information of the image into consideration. This paper breaks through the classical framework based on local information, and proposes a framework based on non-local information. Under the proposed framework, an image quality assessment algorithm based on non-local gradient is also presented. This algorithm predicts image quality by measuring the similarity between the non-local gradients of reference image and the distorted image. The experimental results on the public test database TID2008, LIVE, and CSIQ show that the proposed algorithm can obtain better evaluation results.

Key words: Image Quality Assessment (IQA); Human Visual System (HVS); Non-local gradient

1 引言

图像是人类获取外界信息的主要媒介, 基于图像的信息交流通常涉及图像的获取、压缩、存储和传输等过程^[1]。图像的获取受硬件条件和环境因素的限制可能会引入高斯噪声, 图像压缩也可能会带

来图像的块效应、振铃和模糊等现象。这些不可避免的图像质量损失影响着图像信息表达的充分性和准确性, 影响人们的视觉感知效果^[2,3]。为了降低这些视觉可感知的图像质量损失, 需量化图像质量的退化程度, 以便于在处理过程对图像质量进行保持、控制或增强。因此, 设计有效的图像质量评价算法具有重要的研究意义。

通常人是图像的最终感知者, 因此图像质量应当以人的主观评价为准。但主观评价容易受观测者的年龄、情绪、经验等个体因素影响, 消除个体因素影响的办法是邀请大量的观测者对同一幅图像进

收稿日期: 2018-06-19; 改回日期: 2018-12-18; 网络出版: 2018-12-26

*通信作者: 张选德 zhangxuande@sust.edu.cn

基金项目: 国家自然科学基金(61871260, 61603234, 61362029, 61461043)

Foundation Items: The National Natural Science Foundation of China (61871260, 61603234, 61362029, 61461043)

行评分, 然后对评分值进行统计处理。这种方法较为可靠, 但其成本高, 耗时长, 且不能嵌入实时系统中^[4], 因此有必要研究客观图像质量评价(Image Quality Assessment, IQA)。客观图像质量评价研究的目的在于模拟人类视觉系统(Human Visual System, HVS)对图像质量的感知过程, 构建与主观评价尽可能一致的图像质量算法。根据评价过程中有无真实图像作为参考, IQA可分为全参考(Full Reference, FR)、部分参考(Reduced Reference, RR)和无参考(No Reference, NR)图像质量评价^[5]。本文研究全参考图像质量评价。

图像处理领域通常假定图像属于 $L_2(R^2)$ 空间, 因此 L_2 范数引导的均方误差(Mean Squared Error, MSE)^[6]和峰值信噪比(Peak Signal Noise Ratio, PSNR)^[7]作为最基本的图像质量指标沿用至今, 但这些指标既不考虑图像像素之间的相关性, 也不考虑HVS的感知机制, 其评价结果与主观评价一致性较差。针对MSE和PSNR的不足, 文献^[8]中假定HVS适用于从视觉场景中提取结构信息, 提出了基于结构相似(Structure SIMilarity, SSIM)的图像质量评价算法。该算法较之PSNR和MSE, 对图像质量的度量与主观评价的一致性明显提高, 被认为是图像质量评价领域具有里程碑意义的算法。随后, 研究者基于图像局部结构相似的算法框架, 提出一系列的改进算法^[9-11], 以求获得与主观评价更好的一致性。

目前大多数算法都是基于图像的局部信息设计的, 基于SSIM模型构建的MS-SSIM^[9], 3-SSIM^[10], IW-SSIM^[11]等IQA算法, 均采用度量局部结构失真的方法获得局部质量图。基于梯度特征设计的IQA算法^[12-18], 性能与SSIM相比有了提高, 但梯度特征的提取通常只考虑以当前像素点为中心的邻域内的其他像素, 仅反映了局部变化。图像质量评价的关键在于模拟人对图像质量的主观感知过程。当人观测一幅图像时, 并不逐一地考察图像的局部结构或局部变化, 而是能够感知到图像中包含的语义信息。例如, 当人对一幅人脸图像进行观测时, 总是能够感知到眼睛、鼻子、头发等语义信息, 而不同语义信息的失真对整幅图像视觉质量的影响也往往不同。如给人脸图像中不同语义信息区域加入相同量的噪声, 如果噪声加在头发的区域, 整幅图像依然具有较好的视觉质量; 如果噪声加在眼睛、鼻子区域, 主观感知到的整幅图像的视觉质量就会变差。以上分析说明人对图像的主观感知是语义的过程, 而语义信息本质上是非局部的, 因此图像质量算法的设计应该考虑非局部信息。基于这一研究动机, 本文的工作主要有两个方面: (1)突破经典

的基于局部信息的算法框架, 提出一种基于非局部信息的框架; (2)在非局部信息的框架内构建了一种基于非局部梯度(Non-local Gradient SIMilarity, NGSIM)IQA算法。在测试数据库上完成的数值实验结果表明, NGSIM能够获得与主观评价更好的一致性。

2 基于梯度IQA算法

梯度常被应用于IQA算法设计中, 基于特征相似性(Feature SIMilarity, FSIM)^[12]、梯度相似性(Gradient SIMilarity, GSIM)^[13]、梯度模值相似度离差(Gradient Magnitude Similarity Deviation, GMSD)^[14]、边界强度相似性(Edge Strength SIMilarity, ESSIM)^[15]以及文献^[16-18]等FRIQA算法设计都用到了图像的梯度特征。

记参考图像为 $\mathbf{r} = [r_1, r_2, \dots, r_N]^T$, 失真图像为 $\mathbf{d} = [d_1, d_2, \dots, d_N]^T$, 其中, N 为像素总数。

2.1 FSIM算法

文献^[12]基于HVS对图像理解的底层视觉特征提出FSIM算法, 以相位一致性和梯度模为特征, 计算参考图像和失真图像的局部结构失真。分别记 PC_r, PC_d 为 $\mathbf{r}$ 和 $\mathbf{d}$ 的相位一致性, GM_r, GM_d 为 $\mathbf{r}$ 和 $\mathbf{d}$ 的梯度模, 则 $\mathbf{r}$ 和 $\mathbf{d}$ 在像素点 i 处PC相似性和GM相似性分别定义为式(1)、式(2), i 处的局部质量定义为式(3)

$$S_{PC}(i) = \frac{2PC_r(i) \cdot PC_d(i) + T_1}{PC_r^2(i) + PC_d^2(i) + T_1} \quad (1)$$

$$S_{GM}(i) = \frac{2GM_r(i) \cdot GM_d(i) + T_2}{GM_r^2(i) + GM_d^2(i) + T_2} \quad (2)$$

$$S_L(i) = [S_{PC}(i)]^\alpha [S_{GM}(i)]^\beta \quad (3)$$

其中, T_1 和 T_2 均为常量, 参数 α, β 用于调节 S_{PC}, S_{GM} 的权重, 文中取 $\alpha = \beta = 1$ 。

以相位一致性的较大值 $PC_m(i) = \max(PC_r(i), PC_d(i))$ 对 $S_L(i)$ 进行视觉显著性加权, FSIM算法定义为

$$FSIM = \frac{\sum_{i \in \Omega} S_L(i) \cdot PC_m(i)}{\sum_{i \in \Omega} PC_m(i)} \quad (4)$$

FSIM算法采用两种初级特征并考虑HVS的视觉显著特性, 取得了较好的评价效果, 现已成为IQA研究领域地方基准(benchmark)算法, 但算法需进行多次滤波运算使得执行效率偏低, 且相位一致性特征也有其局限性, 不能对人眼视觉中敏感的边缘信息进行很好的表达^[19]。

2.2 GMSD算法

文献[14]基于FRIQA的两步框架,从结构特征的保留程度和池化策略两方面考虑设计算法。以梯度作为描述图像对比度的有效初级视觉特征,计算 $\mathbf{r}$ 和 $\mathbf{d}$ 每一点处梯度模的相似性度量图像局部质量,具体定义为

$$\text{GMS}(i) = \frac{2\text{GM}_r(i) \cdot \text{GM}_d(i) + T_3}{\text{GM}_r^2(i) + \text{GM}_d^2(i) + T_3} \quad (5)$$

GMSD算法定义为

$$\text{GMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{GMS}(i) - \text{GMSM})^2} \quad (6)$$

其中, T_3 为一常量。 $\text{GMSM} = \frac{1}{N} \sum_{i=1}^N \text{GMS}(i)$ 。该算法对图像的整体质量预测采用离差的池化策略,即GMSD值越大表明图像的失真越严重,主观感知的质量越差。

2.3 ESSIM算法

文献[15]假定图像中被感知到的边界是HVS语义认知的基础,提出基于边界强度相似性的ESSIM算法。定义参考图像 $\mathbf{r}$ 在第 i 个像素的边界强度为

$$\mathbf{E}(r, i) = \max(\mathbf{E}_i^{1,3}(r), \mathbf{E}_i^{2,4}(r)) \quad (7)$$

其中, $\mathbf{E}_i^{2,4}(r) = |\partial r_i^2 - \partial r_i^4|^p$, $\mathbf{E}_i^{1,3}(r) = |\partial r_i^1 - \partial r_i^3|^p$, $\partial r_i^1, \partial r_i^2, \partial r_i^3, \partial r_i^4$ 为4个方向上的方向导数。

为保证 $\mathbf{d}$ 和 $\mathbf{r}$ 在相同位置的边界强度能够反映相同方向的变化, $\mathbf{d}$ 在第 i 个像素的边界强度定义为

$$\mathbf{E}(d, i) = \begin{cases} \mathbf{E}_i^{1,3}(d), & \mathbf{E}(r, i) = \mathbf{E}_i^{1,3}(r) \\ \mathbf{E}_i^{2,4}(d), & \mathbf{E}(r, i) = \mathbf{E}_i^{2,4}(r) \end{cases} \quad (8)$$

ESSIM算法定义为

$$\text{ESSIM}(r, d) = \frac{1}{N} \sum_{i=1}^N \frac{2\mathbf{E}(r, i)\mathbf{E}(d, i) + T_4}{(\mathbf{E}(r, i))^2 + (\mathbf{E}(d, i))^2 + T_4} \quad (9)$$

ESSIM算法取得了比FSIM略好的评价效果,且具有较低的运算复杂度。事实上,ESSIM算法设计已考虑到HVS对图像中语义信息的敏感性,且其提取的边界特征也通常分布在图像的非局部区域。

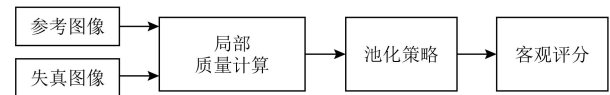
以上考虑梯度特征设计的IQA算法,在一定程度上提高了算法性能,但也有其局限性:(1)图像某一像素点的梯度,通常由以该像素点为中心,与梯度算子大小相同的图像块和梯度算子卷积计算得到。与该像素点相关的视觉关注区域仅考虑了像素点周围局部邻域,是一种局部处理的方法,而主观感知不仅仅关注于图像局部区域;(2)基于梯度的

IQA算法中图像梯度特征的提取通常仅反映了图像特定方向上亮度值的空间分布特征,未考虑像素点周围其它方向上的对比变化,人类视觉观察图像时不仅局限于关注到特定几个方向;(3)FSIM, GMSD算法中用到了梯度模特征。不同的水平和垂直方向的梯度值,其梯度模也可能相同,因此梯度模相似性比较不能完全刻画图像结构信息差异。

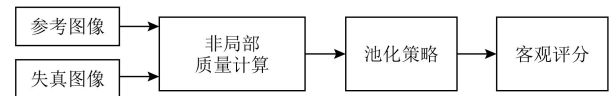
3 基于局部信息和非局部信息的全参考图像质量评价(FRIQA)框架

许多FRIQA设计的算法框架通常具有两个步骤,首先比较参考图像和失真图像在局部的相似性,得到每一处局部区域的局部质量图(Local Quality Map, LQM),再通过一定的池化方法以LQM预测整幅图像质量,基于局部信息的FRIQA框架如图1(a)。

但是,HVS对图像主观质量的感知并不逐一地考察像素点的局部结构,而是一个高级的、语义的过程,这些语义信息往往分布在图像非局部区域。现有IQA算法框架以局部失真度量图像质量的退化,不能够很好地刻画HVS的特性。本文突破现有的基于局部信息的算法框架,提出一种基于非局部的算法框架,首先通过比较参考图像和失真图像的非局部信息相似性,计算非局部质量图(Non-Local Quality Map, NLQM),再将NLQM以一定的方式聚合,获得整幅图像的客观质量评分,具体的算法框架如图1(b)。



(a) 基于局部信息的FRIQA框架



(b) 基于非局部信息的FRIQA框架

图1 基于局部和非局部信息的FRIQA模型两步框架

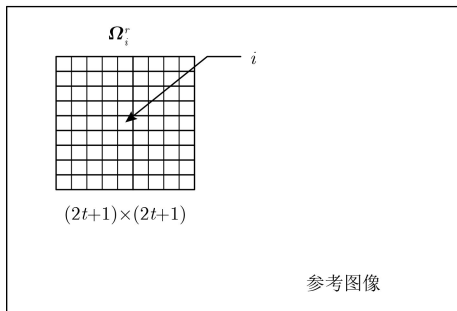
4 基于非局部梯度IQA算法

记参考图像为 $\mathbf{r} = [r_1, r_2, \dots, r_N]^T$,以像素点 i 为圆心、 t 为半径的方邻域记作 Ω_i^r ,该邻域的大小为 $(2t+1) \times (2t+1)$,如图2所示。

参考图像 $\mathbf{r}$ 在像素点 i 处的非局部梯度(Non-local Gradient, NG)记作 $\text{NG}_r(i)$,定义为

$$\text{NG}_r(i) = |r_j - r_i|^{\frac{1}{2}}, \quad j \in \Omega_i^r \quad (10)$$

其中, r_i 为参考图像在 i 处的灰度值, $\text{NG}_r(i)$ 表示以 i 为心、 t 为半径的方邻域 Ω_i^r 内所有像素点灰度值

图2 参考图像中以*i*为中心、*t*为边长的方邻域

r_j 相对于该像素点 r_i 的变化, $\text{NG}_r(i)$ 的计算结果可用 $(2t+1) \times (2t+1)$ 的矩阵表示。与局部梯度计算的意义不同, 局部梯度仅度量图像在水平和垂直两个方向上的变化, 而非局部梯度可以度量图像在多个方向上的变化。

同样地, 记失真图像 $\mathbf{d} = [d_1, d_2, \dots, d_N]^T$ 在*i*点处的非局部梯度为 $\text{NG}_d(i)$, 记以*i*为心的*t*为半径的方邻域为 Ω_i^d , 则 $\text{NG}_d(i)$ 定义为

$$\text{NG}_d(i) = |d_j - d_i|^{\frac{1}{2}}, \quad j \in \Omega_i^d \quad (11)$$

其中, $\text{NG}_d(i)$ 可用 $(2t+1) \times (2t+1)$ 的矩阵表示。

根据式(10)、式(11)计算出 $\mathbf{r}$ 与 $\mathbf{d}$ 在每一点处的非局部梯度 $\text{NG}_r(i)$ 和 $\text{NG}_d(i)$, 通过度量二者的相似性预测*i*点处的视觉感知质量。记非局部梯度 $\text{NG}_r(i)$ 中的元素为 $\text{NG}_r(i)(j)$, 其中 $j \in \Omega_i^r$ 。 $\text{NG}_r(i)$ 反映参考图像 $\mathbf{r}$ 在*i*点处非局部区域内的变化, 其平均变化程度可由均值计算估计 $\mu_r(i)$, 定义为

$$\mu_r(i) = \frac{1}{m} \sum_{j=1}^m \text{NG}_r(i)(j) \quad (12)$$

其中, $m = (2t+1)^2$ 。

类似地, 记 $\text{NG}_d(i)$ 中的元素为 $\text{NG}_d(i)(j)$, $j \in \Omega_i^d$ 。均值 $\mu_d(i)$ 反映失真图像 $\mathbf{d}$ 在*i*处的平均变化, 定义为

$$\mu_d(i) = \frac{1}{m} \sum_{j=1}^m \text{NG}_d(i)(j) \quad (13)$$

非局部梯度 $\text{NG}_r(i)$ 和 $\text{NG}_d(i)$ 平均变化程度的相似性度量定义为

$$S_1(i) = \frac{2\mu_r(i) \cdot \mu_d(i) + C_1}{\mu_r^2(i) + \mu_d^2(i) + C_1} \quad (14)$$

在统计学中, 方差衡量随机变量分布的离散程度, 协方差度量两个随机变量之间的相关性。非局部梯度 $\text{NG}_r(i)$ 与 $\text{NG}_d(i)$ 的相似性可用二者的协方差来刻画。

式(15)计算 $\text{NG}_r(i)$ 的方差, 描述非局部梯度 $\text{NG}_r(i)$ 中各元素值 $\text{NG}_r(i)(j)$ 相对于均值 $\mu_r(i)$ 的离

散程度, 方差越大, 离散程度越大, 表明变化差异性越大。

$$\sigma_r^2(i) = \frac{1}{m-1} \sum_{j=1}^m (\text{NG}_r(i)(j) - \mu_r(i))^2 \quad (15)$$

类似地, $\text{NG}_d(i)$ 的方差计算定义为

$$\sigma_d^2(i) = \frac{1}{m-1} \sum_{j=1}^m (\text{NG}_d(i)(j) - \mu_d(i))^2 \quad (16)$$

式(17)计算 $\text{NG}_r(i)$ 和 $\text{NG}_d(i)$ 的协方差, 度量二者之间的相关性, 反映 $\text{NG}_r(i)$ 和 $\text{NG}_d(i)$ 变化趋势的一致性。

$$\sigma_{r,d}(i) = \frac{1}{m-1} \sum_{j=1}^m (\text{NG}_r(i)(j) - \mu_r(i)) \cdot (\text{NG}_d(i)(j) - \mu_d(i)) \quad (17)$$

非局部梯度 $\text{NG}_r(i)$ 和 $\text{NG}_d(i)$ 结构相似性度量定义为

$$S_2(i) = \frac{2\sigma_{r,d}(i) + C_2}{\sigma_r^2(i) + \sigma_d^2(i) + C_2} \quad (18)$$

式(14)、式(18)中, 参数 C_1, C_2 均为正常量以增加计算结果的稳定性。

结合式(14)、式(18), $\mathbf{r}$ 与 $\mathbf{d}$ 在每一像素点*i*处的非局部梯度 $\text{NG}_r(i)$ 和 $\text{NG}_d(i)$ 的相似性度量定义为

$$\begin{aligned} \text{SIM}(\text{NG}_r(i), \text{NG}_d(i)) &= [(2\mu_r(i) \cdot \mu_d(i) + C_1)(2\sigma_{r,d}(i) + C_2)] / [(\mu_r^2(i) + \mu_d^2(i) + C_1)(\sigma_r^2(i) + \sigma_d^2(i) + C_2)] \end{aligned} \quad (19)$$

最后, 以逐点像素的非局部梯度相似性预测图像整体质量, 采用均值池化的方法计算图像的客观评分

$$\text{NGSIM}(r, d) = \frac{1}{N} \sum_{i=1}^N \text{SIM}(\text{NG}_r(i), \text{NG}_d(i)) \quad (20)$$

5 数值实验

5.1 测试数据库及算法评价指标

本文实验数据为3大公开图像测试数据库TID2008^[20], CSIQ^[21], LIVE^[22]。TID2008包括25幅参考图像、17种失真类型、4种失真等级, 共3000幅失真图像; CSIQ包括30幅参考图像、5种失真类型, 4~5种失真等级, 共866幅失真图像; LIVE包括29幅参考图像、5种失真类型, 4~5种失真等级, 共779幅失真图像。

采用4种常用的算法性能评价指标, 斯皮尔曼等级相关系数(SROCC)、坎德尔等级相关系数(KROCC)、皮尔逊线性相关系数(PLCC)和均方根

误差(RMSE)。SROCC和KROCC指标反映主客观评分的单调性, 绝对值越接近1, 则单调性越好; PLCC和RMSE指标评价算法的预测准确性, 计算时需对客观评分和主观评分进行非线性映射, 映射函数采用式(21)

$$Q_P = \beta_1 \left[\frac{1}{2} - \frac{1}{1 + \exp(\beta_2(Q - \beta_3))} \right] + \beta_4 Q + \beta_5 \quad (21)$$

其中, Q 和 Q_P 分别为IQA算法客观评分及非线性映射后的分值, $\beta_1, \beta_2, \beta_3, \beta_4$ 和 β_5 是可选择的参数。好的IQA算法应具有较高的SROCC, KROCC, PLCC值和较低的RMSE值。

5.2 实验参数

本文算法需设定参数有 C_1, C_2 和 t 。取 $C_1 = (K_1 L)^2$, $C_2 = (K_2 L)^2$, 其中 L 为图像取值动态范围, K_1, K_2 为很小的常量。 t 的大小决定非局部梯度搜索区域范围, 非局部窗口大小与HVS观测感知的尺度相关。一般情况下, HVS感知到的非局部变化区域通常在以当前像素点为中心的一定范围内, 离像素点较远的区域对其影响较小。同时为了降低算法的计算复杂度, 非局部区域选取像素点周围半径为 t 的区域。实验参数通过在TID2008子集上实验来确定, 子集包含数据库中前8个参考图像及相应的544幅失真图像, 以在该子集上获得最高的SROCC值确定实验参数, 根据实验数据, 分别取 K_1, K_2 和 t 的值为0.04, 0.10和21。

5.3 实验结果比较

本文实验环境为MATLAB R2011a, 所有的实验均在DELL optiplex7040, Intel Core i7-6700, CPU@3.4 GHz, 8 G RAM的计算机上进行。实验比较本文所提NGSIM算法和其他IQA算法性能,

比较算法包括PSNR, VSNR^[4], SSIM, MS-SSIM, IW-SSIM, FSIM, ESSIM, GSIM和GMSD。实验中对比算法直接采用作者公布的代码。另外, 实验中如输入为彩色图像, 则仅其亮度成分进行评价。NGSIM与对独立噪声评价效果较好的PSNR算法比较, 可反映本文算法对图像噪声的敏感性。对比算法SSIM与其改进算法MS-SSIM, IW-SSIM以图像局部失真度量图像质量, 选取该类算法作为比较可考查非局部信息对图像质量评价的有效性。FSIM, ESSIM算法用到了梯度特征, GSIM, GMSD为基于梯度特征的算法, 其对比结果可反映非局部梯度特征与局部梯度特征相比对IQA算法评价准确率的影响。

首先验证算法在整体数据库上的评价效果, 表1列出了10种不同IQA算法在3大数据库上的评价结果, 排名首位的实验结果已加粗显示。从表中可看出, NGSIM出现8次, ESSIM出现2次, GMSD和FSIM均为1次。在CSIQ数据库上, 4个指标均显示NGSIM优于其他算法, 具有较好的图像质量预测性能。在TID2008数据库上, SROCC, KROCC指标结果显示, NGSIM算法在图像主客观评价分值单调性能比较上略高于GMSD, 均好于其他算法; 在预测准确性方面, ESSIM和NGSIM效果相当, 优于其他算法。在LIVE数据库上, FSIM, ESSIM和GMSD分别在SROCC, KROCC, PLCC指标上表现出好的评价性能, NGSIM的均方根误差RMSE指标差于ESSIM, 优于FSIM和GMSD。表中实验结果显示, 本文基于非局部梯度的算法NGSIM在3个数据库中不同指标上均明显优于基于梯度的算法GSIM。

为进一步验证本文算法的评价效果, 针对图像单一失真类型进行实验, 选取SROCC作为评价指

表 1 10种不同IQA算法在TID2008, CSIQ, LIVE数据库的实验结果比较

数据库	性能指标	PSNR	VSNR	SSIM	MS-SSIM	IW-SSIM	FSIM	ESSIM	GMSD	GSIM	NGSIM
TID2008	SROCC	0.524	0.704	0.774	0.852	0.855	0.880	0.884	0.891	0.855	0.892
	KROCC	0.369	0.534	0.576	0.654	0.663	0.694	0.704	0.708	0.665	0.713
	PLCC	0.530	0.682	0.773	0.842	0.857	0.873	0.885	0.879	0.846	0.886
	RMSE	1.137	0.981	0.851	0.729	0.689	0.652	0.624	0.640	0.715	0.622
CSIQ	SROCC	0.805	0.810	0.875	0.913	0.921	0.924	0.932	0.957	0.912	0.962
	KROCC	0.608	0.624	0.690	0.739	0.752	0.756	0.768	0.813	0.740	0.825
	PLCC	0.800	0.800	0.861	0.899	0.914	0.912	0.922	0.954	0.897	0.961
	RMSE	0.157	0.157	0.133	0.114	0.106	0.100	0.101	0.079	0.115	0.073
LIVE	SROCC	0.875	0.927	0.947	0.944	0.956	0.963	0.962	0.960	0.955	0.950
	KROCC	0.686	0.761	0.796	0.792	0.817	0.833	0.839	0.823	0.813	0.815
	PLCC	0.872	0.923	0.944	0.943	0.952	0.959	0.953	0.960	0.943	0.946
	RMSE	13.36	10.50	8.944	9.095	8.347	7.678	7.003	7.62	9.037	7.455

标。表2中列出了10种评价算法在3大数据库上对每一种失真类型的评价效果(TID2008中列出6种), 加粗显示了排名第1的实验结果, 从表中可看出PSNR对噪声类型的失真有较好的评价效果。NGSIM对空间位置相关噪声SCN失真评价效果明显, 较好地刻画HVS对固有空间频率的敏感性; NGSIM在

TID2008, CSIQ对JPEG有损压缩的图像失真质量预测较好, 对有损图像压缩技术进行失真评估的任务非常有意义; 其次, 算法对JPEG2000压缩图像在数据传输通道中错误解码, FNIOSE, FAST-FA取得较好的评价效果。

图3给出了TID2008上“平均意见分值”MOS

表2 10种不同IQA算法在TID2008, CSIQ, LIVE数据库单一失真性能(SROCC)的比较

数据库	失真类型	PSNR	VSNR	SSIM	MS-SSIM	IW-SSIM	FSIM	ESSIM	GMSD	GSIM	NGSIM
TID2008	AWN	0.907	0.772	0.811	0.809	0.786	0.857	0.885	0.918	0.857	0.902
	ANMC	0.899	0.779	0.803	0.805	0.792	0.851	0.813	0.898	0.809	0.873
	SCN	0.917	0.766	0.815	0.819	0.771	0.848	0.913	0.913	0.890	0.929
	JPEG	0.872	0.917	0.925	0.934	0.918	0.928	0.943	0.952	0.939	0.956
	JP2K	0.813	0.951	0.962	0.973	0.973	0.977	0.975	0.980	0.975	0.958
	J2TE	0.831	0.790	0.858	0.852	0.820	0.854	0.879	0.883	0.892	0.926
CSIQ	AWGN	0.936	0.924	0.897	0.947	0.938	0.926	0.949	0.968	0.944	0.966
	JPEG	0.888	0.903	0.954	0.963	0.966	0.965	0.964	0.965	0.963	0.966
	JP2K	0.936	0.948	0.960	0.968	0.968	0.968	0.967	0.972	0.964	0.974
	FNIOSE	0.933	0.908	0.892	0.933	0.905	0.923	0.943	0.950	0.938	0.962
	BLUR	0.929	0.944	0.960	0.971	0.978	0.972	0.962	0.971	0.958	0.967
	CONTRST	0.862	0.870	0.792	0.952	0.953	0.942	0.939	0.904	0.950	0.946
LIVE	JPEG2	0.895	0.955	0.961	0.962	0.964	0.971	0.980	0.971	0.958	0.972
	JPEG	0.880	0.965	0.976	0.981	0.980	0.983	0.981	0.978	0.909	0.960
	AWGN	0.985	0.978	0.969	0.973	0.966	0.965	0.976	0.974	0.977	0.993
	BLUR	0.782	0.941	0.951	0.954	0.972	0.970	0.991	0.957	0.951	0.939
	FASTFA	0.890	0.902	0.955	0.947	0.944	0.949	0.947	0.942	0.939	0.956

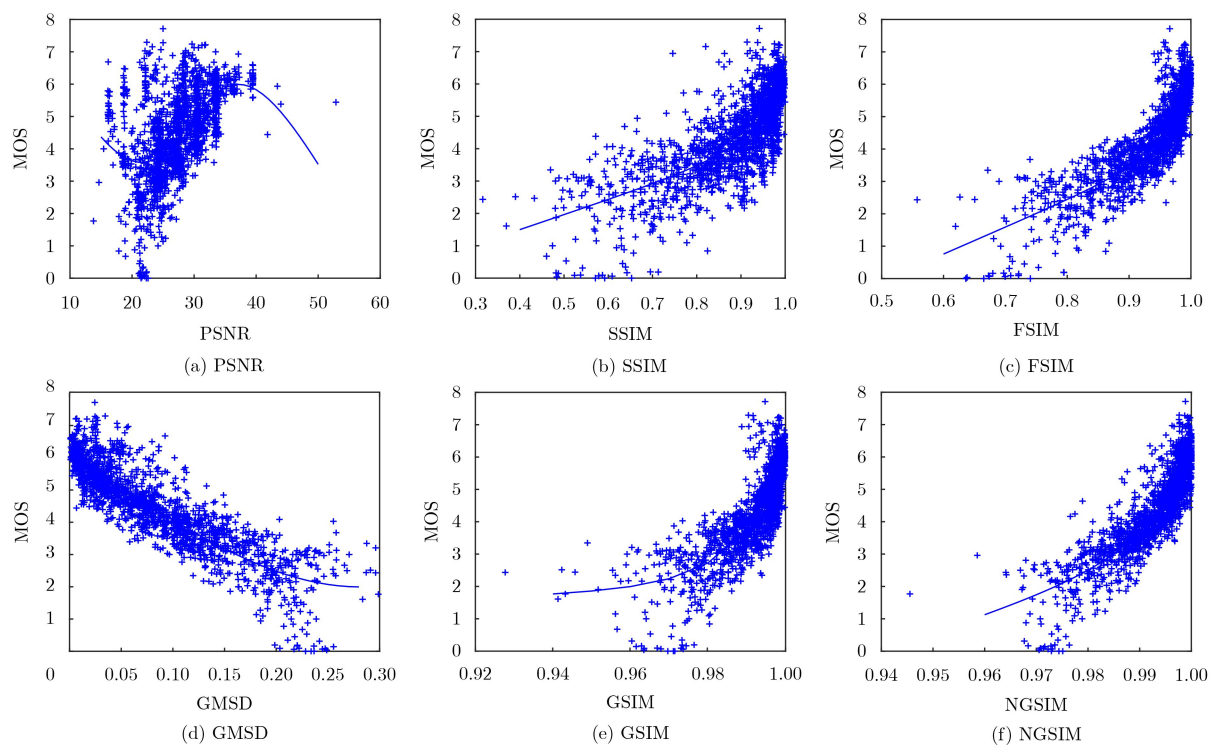


图3 6种算法在TID2008数据库中的散点图

值与6种具有代表性算法客观评分的散点图。算法包括PSNR, SSIM, FSIM, GMSD, GSIM, NGSIM, 从图中可见NGSIM与主观评价MOS有更好的一致性。

6 结束语

本文工作的贡献主要有两点: (1) 突破经典的基于局部信息的算法框架, 提出一种基于非局部信息的框架。本文讨论了现有大多数IQA算法是以局部信息处理的方法设计的, 并分析了人类视觉系统的主观感知是高级的、语义的过程, 而语义信息是非局部的, IQA算法设计应考虑非局部信息; (2) 在基于非局部信息的框架内构建了一种基于非局部梯度的图像质量评价算法, 本算法通过度量参考图像与失真图像的非局部梯度之间的相似性来预测图像质量, 在公开测试数据库中与目前主流算法进行实验比较, 本文算法取得较好的结果。

参 考 文 献

- [1] BAE S H and KIM M. A novel image quality assessment with globally and locally consilient visual quality perception[J]. *IEEE Transactions on Image Processing*, 2016, 25(5): 2392–2406. doi: [10.1109/TIP.2016.2545863](https://doi.org/10.1109/TIP.2016.2545863).
- [2] WANG Hanli, FU Jie, LIN Weisi, *et al.* Image quality assessment based on local linear information and distortion-specific compensation[J]. *IEEE Transactions on Image Processing*, 2017, 26(2): 915–926. doi: [10.1109/TIP.2016.2639451](https://doi.org/10.1109/TIP.2016.2639451).
- [3] DI E C and JACOVITTI G. A detail based method for linear full reference image quality prediction[J]. *IEEE Transactions on Image Processing*, 2017, 27(1): 179–192. doi: [10.1109/TIP.2017.2757139](https://doi.org/10.1109/TIP.2017.2757139).
- [4] CHANDLER D M and HEMAMI S S. VSNR: A wavelet-based visual signal-to-noise ratio for natural images[J]. *IEEE Transactions on Image Processing*, 2007, 16(9): 2284–2298. doi: [10.1109/TIP.2007.901820](https://doi.org/10.1109/TIP.2007.901820).
- [5] 褚江, 陈强, 杨曦晨. 全参考图像质量评价综述[J]. *计算机应用研究*, 2014, 31(1): 13–22. doi: [10.3969/j.issn.1001-3695.2014.01.003](https://doi.org/10.3969/j.issn.1001-3695.2014.01.003).
CHU Jiang, CHEN Qiang, and YANG Xichen. Review on full reference image quality assessment algorithms[J]. *Application Research of Computers*, 2014, 31(1): 13–22. doi: [10.3969/j.issn.1001-3695.2014.01.003](https://doi.org/10.3969/j.issn.1001-3695.2014.01.003).
- [6] WANG Zhou and BOVIK A C. Mean squared error: love it or leave it? A new look at signal fidelity measures[J]. *IEEE Signal Processing Magazine*, 2009, 26(1): 98–117. doi: [10.1109/MSP.2008.930649](https://doi.org/10.1109/MSP.2008.930649).
- [7] HUYNH-THU Q and GHANBARI M. Scope of validity of PSNR in image/video quality assessment[J]. *Electronics Letters*, 2008, 44(13): 800–801. doi: [10.1049/el:20080522](https://doi.org/10.1049/el:20080522).
- [8] 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](https://doi.org/10.1109/TIP.2003.819861).
- [9] WANG Zhou, SIMONCELLI E P, and BOVIK A C. Multiscale structural similarity for image quality assessment[C]. *Proceedings of 37th IEEE Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, USA, 2003: 1398–1402.
- [10] LI Chaofeng and BOVIK A C. Three-component weighted structural similarity index[C]. *SPIE Conference on Image Quality and System Performance*, San Jose, USA, 2009, 7242: 72420Q–72420Q-9.
- [11] WANG Zhou and 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](https://doi.org/10.1109/TIP.2010.2092435).
- [12] 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](https://doi.org/10.1109/TIP.2011.2109730).
- [13] LIU Anmin, LIN Weisi, and 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](https://doi.org/10.1109/TIP.2011.2175935).
- [14] 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](https://doi.org/10.1109/TIP.2013.2293423).
- [15] ZHANG Xuande, FENG Xiangchu, WANG Weiwei, *et al.* Edge strength similarity for image quality assessment[J]. *IEEE Signal Processing Letters*, 2013, 20(4): 319–322. doi: [10.1109/LSP.2013.2244081](https://doi.org/10.1109/LSP.2013.2244081).
- [16] WANG Tonghan, JIA Huizhen, and SHU Huazhong. Full-reference image quality assessment algorithm based on gradient magnitude and histogram of oriented gradient[J]. *Journal of Southeast University*, 2018, 48(2): 276–281. doi: [10.3969/j.issn.1001-0505.2018.02.014](https://doi.org/10.3969/j.issn.1001-0505.2018.02.014).
- [17] NI Zhangkai, MA Lin, ZENG Huanqiang, *et al.* Gradient direction for screen content image quality assessment[J]. *IEEE Signal Processing Letters*, 2016, 23(10): 1394–1398. doi: [10.1109/LSP.2016.2599294](https://doi.org/10.1109/LSP.2016.2599294).
- [18] DING Li, HUANG Hua, and ZANG Yu. Image quality

- assessment using directional anisotropy structure measurement[J]. *IEEE Transactions on Image Processing*, 2017, 26(4): 1799–1809. doi: [10.1109/TIP.2017.2665972](https://doi.org/10.1109/TIP.2017.2665972).
- [19] 张帆, 张偲雅, 李珍珍. 基于对称相位一致性的图像质量评价方法[J]. *激光与光电子学进展*, 2017, 54(10): 194–202. doi: [10.3788/LOP54.101003](https://doi.org/10.3788/LOP54.101003).
- ZHANG Fan, ZHANG Ruoya, and LI Zhenzhen. Image quality assessment based on symmetry phase congruency[J]. *Laser & Optoelectronics Progress*, 2017, 54(10): 194–202. doi: [10.3788/LOP54.101003](https://doi.org/10.3788/LOP54.101003).
- [20] PONOMARENKO N, LUKIN V, ZELENSKY A, *et al*. TID2008: A database for evaluation of full-reference visual quality assessment metrics[OL]. <http://www.ponomarenko.info/papers/mre2009tid.pdf>. 2016.10.
- [21] LARSON EC and CHANDLER D. Categorical subjective image quality (CSIQ) database[OL]. <http://vision.okstate.edu/csiq>, 2016.10.
- [22] SHEIKH H R, WANG Zhou, BOVIK A C, *et al*. Image and video quality assessment research at LIVE[OL]. <http://live.ece.utexas.edu/rese-arch/quality/>. 2016.10.
- 高敏娟：女，1984年生，博士生，研究方向为图像处理、图像质量评价。
- 党宏社：男，1962年生，教授，博士生导师，研究方向为工业过程与优化、计算机控制、图像处理。
- 魏立力：男，1965年生，教授，研究方向为应用统计与数据分析。
- 张选德：男，1979年生，教授，博士生导师，研究方向为图像恢复、图像质量评价、稀疏表示和低秩逼近理论。