

局部和全局特征融合的色调映射图像质量评价

侯春萍, 李浩, 岳广辉[†]

(天津大学 电气自动化与信息工程学院, 天津 300072)

摘要:人类视觉系统首先粗略地感知全局区域,然后精细地感知局部区域的图像质量.针对色调映射图像的质量评价问题,考虑人眼视觉机制的特性,提出一种融合局部和全局特征的无参考图像质量评价算法.首先从全局特征出发,考虑了颜色矩、全局熵和欠曝光/过曝光条件下的明暗分布特性,得到相应的全局特征;然后结合局部对比度、局部熵和分块小波能量,得到相应的局部特征;最后,融合全局特征和局部特征,使用支持向量回归进行特征训练,建立图像特征空间与主观质量分数的关系,得到图像质量评价模型.在公开的 ESPL-LIVE HDR 数据库上验证,实验结果表明,提出的方法与主观评分有较高的一致性,并且性能优于目前较优秀的无参考图像质量评价算法.

关键词:图像质量评价;人类视觉系统;色调映射;无参考

中图分类号:TP391

文献标志码:A

Quality Assessment of Tone-mapped Images Using Local and Global Features

HOU Chunping, LI Hao, YUE Guanghui[†]

(School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China)

Abstract: Human Visual System (HVS) first roughly perceives global areas, then centers on the detailed local areas for the perception of image quality. In this paper, a novel blind Image Quality Assessment (IQA) algorithm was proposed for tone-mapped images by combining local and global features. First, the global features were extracted based on color moments, global entropy and bright/dark pixels' distribution under overexposure/underexposure conditions. Then, local contrast, local entropy and wavelet energy based on blocks were utilized to extract local features. Finally, global features were combined with local features to constitute a final feature vector. And all these feature vectors mentioned above were trained using Support Vector Regression (SVR) to generate a model, which bridges the feature space with quality space. Extensive experiments on a public ESPL-LIVE HDR database have demonstrated that the proposed method has a high consistency with subjective evaluation and outperforms state-of-the-art no-reference IQA metrics.

Key words: Image Quality Assessment (IQA); Human Visual System (HVS); tone mapping; No-reference (NR)

* 收稿日期:2018-07-02

基金项目:国家自然科学基金资助项目(61520106002, 61471262, 61731003), National Natural Science Foundation of China(61520106002, 61471262, 61731003)

作者简介:侯春萍(1957—),女,天津人,天津大学教授,博士生导师

[†] 通讯联系人, E-mail: yueguanghui@tju.edu.cn

图像质量评价是图像处理领域的研究热点^[1],可以用于监控图像质量、评估图像处理算法以及优化图像传输系统^[2].近年来,图像质量评价发展迅速,国内外学者提出了众多性能优异的评价算法^[3-6],现有的图像评价算法主要分为全参考、半参考和无参考质量评价算法.全参考和半参考评价算法依赖于参考图像的信息,具有有限的应用场景,而无参考质量评价算法仅需要失真图像便可完成图像的质量评估工作,具备更宽泛的应用范围.开发有效的无参考图像质量评价算法已逐渐成为目前研究的重点.无参考算法主要分为具体失真^[6,7-11]和通用失真^[4,12-15]类型的算法,这些算法主要针对常见的失真类型(如模糊、压缩、噪声等).而色调映射图像主要包含异常曝光、颜色失真.因此,开发适用于色调映射的图像质量评价算法越来越引起研究者的注意.色调映射是通过色调映射算子将高动态范围(HDR, High Dynamic Range)的图像转化为低动态范围(LDR, Low Dynamic Range)图像以适应显示设备的技术^[16],在这个转化中不可避免会造成图像质量降低,因此评价色调映射图像的质量具有重要意义. Gu^[17]等认为高质量色调映射图像包含更多细节,通过亮化或暗化原始色调映射图像的亮度,提取了9个全局熵特征,应用支持向量机来建立特征与主观分数之间的关系. Yue^[18]等通过局部二值模式和灰度共生矩阵提取色调映射图像的局部纹理和结构特征,通过支持向量回归将特征向量映射为质量分数. Jiang^[19]等定义了色调映射图像的最亮区和最暗区,并结合自然图像统计特征和颜色信息进行特征提取,最后使用随机森林回归得到图像的质量分数. Jiang^[20]等考虑了视觉信息、局部结构和自然性,即通过局部稀疏表示和全局统计分析提取特征,最后用极限学习机实现从特征空间到质量分数的映射.综上所述,色调映射图像质量评价仍处于逐步发展完善的阶段.然而,现有算法仍旧存在两方面的不足:一方面算法性能仍然存在很大的提升空间,另一方面很少结合全局特征和局部特征进行分析, Jiang^[20]等虽然综合了全局和局部特征,但是没有充分考虑色调映射图像自身的特性,比如在欠曝光/过曝光条件下的明暗分布特性.因此,提出一种结合局部和全局特征并且符合色调映射图像自身特性的质量评价方案尤为必要.

针对色调映射图像质量评价问题,本文提出了一种局部和全局特征融合的无参考图像质量评价方法.本方法不需要参考图像信息,在公开数据集上性能优于现有的无参考评价算法.

1 无参考图像质量评价算法

本文提出的无参考色调映射图像质量评价算法主要分为三个步骤:

- 1)全局和局部特征提取;
- 2)模型建立;
- 3)质量预测.

算法流程如图1所示.

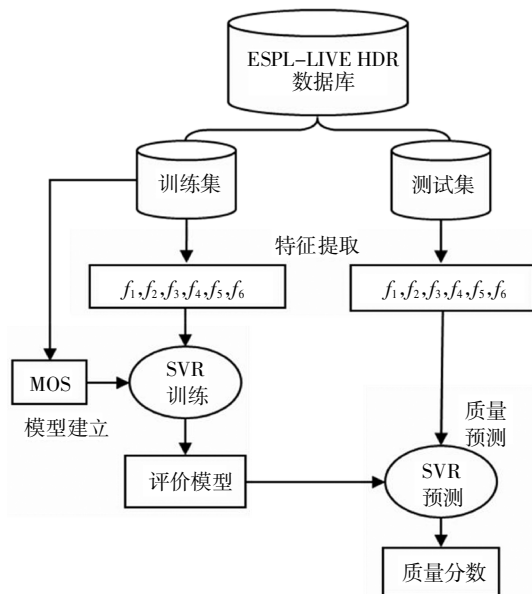


图1 算法流程

Fig.1 Flowchart of the proposed method

选取 ESPL-LIVE HDR 数据库作为数据集,将图片分为训练集和测试集.首先在训练集上进行特征提取,将提取的特征和训练集上的主观质量分数(Mean Opinion Value, MOS)通过机器学习方法进行训练,得到一个质量评价模型;然后在划分的测试集上同样进行特征提取,并将其输入训练好的模型中进行预测,即可得到相应的质量分数.

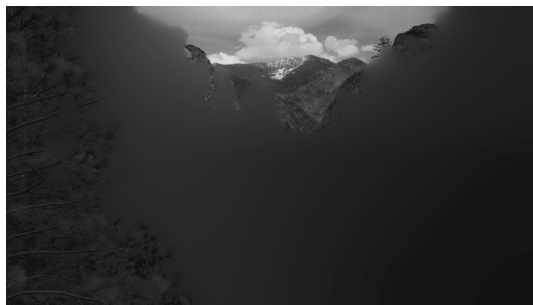
其中,算法的核心是特征提取,本文从色调映射图像的失真特点出发,考虑了颜色、明暗分布和信息量等特征,结合人眼观察物体从全局到局部的特性,分别提取图像局部特征、全局特征,用于评估图像质量.

2 特征提取

视觉生理学研究^[21]表明,人眼视觉系统(Human Visual System, HVS)在观察物体时,首先从整体感知

图像的亮度、颜色等全局信息,继而深入局部区域观察图片的细节信息.由于此特性反映了人眼的普遍规律,因此广泛运用于质量评价过程并且取得了较好的效果^[7,8,20,22].本文模拟 HVS 的机理,从全局和局部两个层次提取图片特征.

色调映射算子在将 HDR 图像转化为 LDR 图像过程中,不可避免地会造成颜色、亮度和信息量等信息的丢失.图 2 是从 ESPL-LIVE HDR 数据库中选出的三张具有代表性的色调映射图片.三张图片的 MOS 值都比较小,分别为 16.94,18.31,31.94.其中,MOS 值越小,表明人眼主观判断该图像的质量越差.



(a)欠曝光



(b)过曝光



(c)颜色失真

图 2 ESPL-LIVE HDR 数据库中低主观质量分数的色调映射图片

Fig.2 Tone-mapped HDR images with poor subjective quality from the ESPL-LIVE HDR database

通过对这三幅图片观察,读者能够发现图 2(a)和(b)分别处于一种类欠曝光和过曝光的状态,图像的细节丢失严重;图片(c)尽管保留了基本图像内

容,然而呈现一种颜色失调的外观.这三幅图像反映了由色调映射算子处理后生成图像的基本失真特性.

针对上述色调映射图像的特性,本算法首先提取全局特征,采用颜色矩来表征颜色特征,定义欠曝光/过曝光的明暗分布来表征非正常的曝光,并用全局熵来量化色调映射图像的信息.然后,进行局部特征的分析,定义局部对比度来度量人眼对色调映射图像的感观变化,计算基于块的局部熵来反映图像的细节信息,并结合分块小波能量来表示图像的轮廓和细节信息.最后,融合全局特征和局部特征,得到最终的特征向量,以反映色调映射图像整体和细节的特性.

2.1 全局特征分析

2.1.1 颜色矩

考虑到色调映射算子非线性映射易造成色彩信息丢失的特点,本文首先提取图像的颜色信息用于质量评价.颜色矩是一种有效的颜色量化手段,本文采用 Stricker 和 Orengo^[23]提出的颜色矩来表示图像的颜色特征.具体而言,在 RGB 颜色空间,首先分解得到每个颜色通道,然后在每个颜色通道分别计算一阶矩均值、二阶矩标准差、三阶矩偏度,计算公式如下:

$$E_i = \frac{1}{N} \sum_{j=1}^N p_{ij} \quad (1)$$

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - E_i)^2 \right)^{1/2} \quad (2)$$

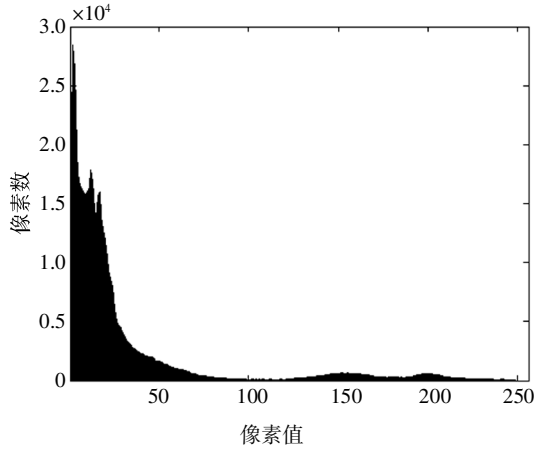
$$s_i = \left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - E_i)^3 \right)^{1/3} \quad (3)$$

式中: p_{ij} 表示第 i 个颜色通道的第 j 个像素; N 表示颜色通道内总像素数,式(1)(2)(3)分别用于求均值、标准差和偏度.通过对三个颜色通道处理,总计得到 9 个颜色特征,记作 f_i .

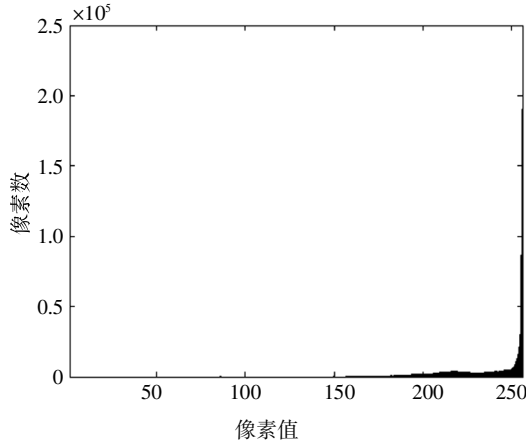
2.1.2 明暗分布

从图 2(a)和(b)可以看出,色调映射图像的一些区域呈现全黑或发白的现象,类似于图像欠曝光或者过曝光,大面积的全黑或者发白意味着大多数像素比较集中分布于低像素值或者高像素值区域,这是因为 HDR 图像在通过色调映射算子转化为 LDR 图像时,图像动态范围减少造成亮度分布不均匀,从而引起质量退化^[24].图 3(a)是图 2(a)欠曝光图像的亮度分布,像素数大多集中在像素值较小的区域,使得整张图片的亮度分布失衡,表现为大部分暗区.相反,图 3(b)是图 2(b)过曝光图像的亮度分

布,像素数多集中分布于像素值较大的区域,整张图片呈现大面积亮区。



(a) 欠曝光图像亮度分布



(b) 过曝光图像亮度分布

图3 非正常曝光图像亮度分布

Fig.3 Image luminance distribution under abnormal exposure

针对色调映射图像这一特征,本文根据经验将像素范围[0 255]分为适度曝光和非正常曝光区域,非正常曝光又可细分为过曝光和欠曝光。三个区域的分割简单遵从平均分布规律,欠曝光对应暗区,过曝光对应亮区,分别定义如下:

1)暗区:像素值分布于[0 85]的图像区域;

2)适度曝光区:像素值分布于[86 169]的图像区域;

3)亮区:像素值分布于[170 255]的图像区域。

定义亮区(暗区)像素数占总像素数的比例为亮区比(暗区比)用于反应图像亮(暗)程度,分别记作 Ω 和 Φ 。计算公式如下:

$$\Omega = \frac{1}{N} \sum_{j=1}^m p_{ij} \quad (4)$$

$$\Phi = \frac{1}{N} \sum_{j=1}^n p_{ij} \quad (5)$$

在灰度图上分别计算 Ω 和 Φ ,将得到的明暗分布特征记作 f_2 。

2.1.3 全局熵

如图3所示,色调映射图像由于非正常曝光,像素值会趋近暗区或者亮区,而暗区和亮区之间的区域则分布较少,使得物体辨识度降低,因此图像的细节丢失严重。信息熵反映信号的复杂程度,本文采用信息熵来量化色调映射图像的信息量:

$$H(D) = - \sum_{l=0}^{255} P_l(D) \log_2 P_l(D) \quad (6)$$

式中: $H(D)$ 表示全局熵; $P_l(D)$ 表示第 l 个灰度级上的概率密度,对8-bits图而言, l 的最大值为255。为方便读者,将全局熵特征记作 f_3 。

综上所述,记全局特征为 f_G ,则 $f_G = [f_1, f_2, f_3]$ 。

2.2 局部特征分析

经过全局特征分析,图像整体反映失真的特征向量通过 f_G 表示。进一步进行局部特征分析,从图像细节中提取质量敏感的特征。

2.2.1 局部对比度

色调映射图像由于动态范围减少,在颜色或者亮度上呈现明显的差异。对比度反映了图像颜色或者亮度在视觉感知中的差异,本文采用对比度来量化人眼对色调映射图像的观感变化。不同图像区域有不同的对比度,定义基于块的局部对比度如下:

$$LC = \frac{1}{P \cdot Q} \sum_{B=1}^P \sum_{B=1}^Q \left(\frac{I_{\max,B} + I_{\min,B} + C_1}{I_{\max,B} - I_{\min,B} + C_2} \right) \quad (7)$$

式中:设置块尺寸为16像素×16像素; P, Q 分别为图像分割后纵向块数和横向块数; $I_{\max,B}$ 和 $I_{\min,B}$ 分别表示图像块中最大像素值和最小像素值;设置 $C_1 = C_2 = 1$ 。

在 R, G, B 三个颜色通道内分别计算 LC 的值,得到3维局部对比度特征,记为 f_4 。

2.2.2 局部熵

全局熵在一定程度上量化了色调映射图像的信息,但是不能反映图像中的细节。因此,提出基于块的局部熵,分块方法与局部对比度相同,分割后一个大小为 $M \times N$ 的图像块中每个灰度值的概率为 ρ_{ij} :

$$\rho_{ij} = \frac{f(i,j)}{\sum_{i=1}^M \sum_{j=1}^N f(i,j)} \quad (8)$$

式中: $f(i,j)$ 表示图像 (i,j) 处的灰度,将概率 ρ_{ij} 代入公式(6)即可得到一个图像块的局部熵 H_L .对一张图所有块的 H_L 计算平均值和标准差,得到局部熵特征,记作 f_5 .

2.2.3 分块小波能量

色调映射图像的明暗分布和色彩失真会使得图像轮廓和细节信息发生变化.如图2(a)(b)所示,类似欠曝光和过曝光的条件下,图像的整体轮廓难以分辨,图2(c)的颜色失真则使得图像的局部细节发生改变.小波变换是一种时频分析方法,广泛应用于图像处理领域^[25-26],相比于傅里叶变换,小波可以进行时频域分析、多分辨率分析,可以从信号中提取有效信息.通过小波变换可以将图像分解为低频和高频成分,其中,低频成分包含了图像的主要信息,而高频成分反映了图像的细节信息,可以在不同的频带上计算小波能量^[27]从而表示图像轮廓和细节信息的变化.图4表示二维离散小波变换(单层分解)的例子,(a)表示原始图像,(b)表示分解图像.



(a)原始图像

(b)分解图像

图4 二维离散小波变换例子

Fig.4 Example of two-dimensional DWT

对比图4(a)(b)可以发现,(b)图的左上角是小波分解的低频区域,因而保留了(a)图的主要轮廓信息,而(b)的其他位置则是小波分解的高频区域,从图中可以看出一些边缘点、线等图像细节.由于色调映射图像的明暗分布和颜色失调会使得图像的轮廓和细节发生变化,因此可以通过小波变换提取图像的低频信息以表征轮廓,提取图像的高频信息以反映细节.具体的二维离散小波变换见框图5.

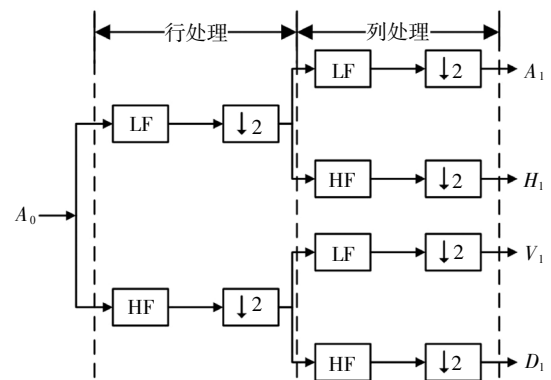


图5 二维离散小波变换

Fig.5 Two-dimensional DWT

图5表示二维离散小波变换. A_0 代表输入的色调映射图像,对应图4(a). A_1, H_1, V_1 和 D_1 分别表示 A_0 的近似分解系数、水平分解系数、垂直分解系数以及对角分解系数,对应图4(b)的左上、右上、左下和右下4个位置.LF、HF分别表示低通滤波器和高通滤波器.

具体地, A_0 经过行低通滤波和列低通滤波处理得到近似分解系数 A_1 , A_1 很大程度上保留了原始图像的低频信息.如图4(b)左上角所示,能够清晰看出图像的轮廓.而 H_1, V_1, D_1 这三个方向上的分解系数,在一定程度上包含了图像的高频分量.如图4(b)右上、左下和右下所示,能够看出图像的一些细节信息.因此, A_1, H_1, V_1 和 D_1 包含了图像的低频和高频成分,能够反映色调映射图像的轮廓和细节信息.

为了进一步表示图像的局部信息,在DWT域将图像分成大小为 $M \times N$ 块,用 C 统一表示近似、水平、垂直和对角方向上的小波分解系数,计算图像块的分解系数,并在系数上进行能量运算.

小波能量可以表示为:

$$E = \sum_{x=1}^M \sum_{y=1}^N (C(x,y))^2 \quad (9)$$

式中: E 为一个图像块的小波能量,包含了近似、水平、垂直和对角方向的能量,对一张图所有图像块的小波能量取均值和标准差,记该特征为 f_6 .

综上所述,将局部特征表示为 f_L ,则 $f_L = [f_4, f_5, f_6]$.经过全局特征和局部特征提取,记总的特征向量为 f ,则 $f = [f_L, f_G]$.

3 模型建立和质量预测

获取多维图像特征后,如何利用这些特征得到

图像的质量是无参考图像质量评价算法下一步的任务.现阶段,机器学习技术被众多学者所采纳,其中较为广泛应用的是 SVR^[4,18].为保证与对比算法的公平性,本文亦采用 SVR 进行评价模型的建立和质量分数的预测.具体而言,首先利用 SVR 对高维的训练集特征 f 进行训练,从而得到一个质量评价模型.标准形式的 SVR 算法可以表示如下^[28]:

$$\begin{aligned} \min_{\omega, b, v, \hat{v}} & \frac{1}{2} \|\omega\|_2^2 + t \sum_{i=1}^z (v_i + \hat{v}_i) \\ \text{s.t.} & \omega^T \phi(x_i) + b - y_i \leq \varepsilon + v_i, \\ & y_i - \omega^T \phi(x_i) - b \leq \varepsilon + \hat{v}_i, \\ & v, \hat{v} \geq 0, i = 1, 2, \dots, z \end{aligned} \quad (10)$$

式中: v 和 $\hat{v}$ 是松弛变量; b 为偏置; x_i 是第 i 张图像的 25 维特征向量; y_i 是与之相关的主观质量分数; z 是训练集样本数. $\phi(x_i)^T \phi(x_i)$ 表示核函数.为方便读者观察,本文进一步给出 25 维特征及描述,见表 1.

表 1 特征提取总结

Tab.1 Summary of extracted features

特征标号	特征描述
f_1	9 维颜色矩特征,公式(1)(2)(3)
f_2	2 维明暗分布特征,公式(4)(5)
f_3	1 维全局熵特征,公式(6)
f_4	3 维局部对比度特征,公式(7)
f_5	2 维局部熵特征,公式(8)(6)
f_6	8 维分块小波能量特征,公式(9)

结合图 1 和表 1, 本文提取了 25 维训练集特征,并将相应的 MOS 值一起通过 SVR 训练,得到一个评价模型.给定一张测试集图片,将提取的测试集特征向量送入模型,即可预测其质量分数.

4 实验结果和分析

4.1 数据库描述

本文选取 ESPL-LIVE HDR Database^[29]作为数据集来验证提出算法的有效性.ESPL-LIVE HDR Database 是目前国际上公开的最大的色调映射图像数据集,因此可以作为检验色调映射无参考图像质量评价算法的理想数据集.该数据库是由德克萨斯大学奥斯汀分校的图像视频工程实验室建立的,总计 1 811 张 LDR 图像,处理的方法主要包括色调映射算子,多曝光融合算法以及软件后期处理.每张图

片的 MOS 值是通过 5 000 多名被试在众包平台打分得到的.数据库不仅提供了每张图片的 MOS 值,还提供了相关的标准差分数.值得注意的是,ESPL-LIVE HDR 数据库没有公开 HDR 图像,因此,将本文算法与一些无参考的评价算法比较,并不涉及全参考方法.

4.2 性能评估准则

为了客观检验图像质量评价算法的性能,采用视频质量专家组建议的四种性能评估准则,分别是皮尔森线性相关系数(Pearson Linear Correlation Coefficient, PLCC)、斯皮尔曼秩相关系数(Spearman Rank-order Correlation Coefficient, SRCC)、肯德尔秩相关系数(Kendall's Rank Correlation Coefficient, KRCC)和均方根误差(Root Mean-Squared Error, RMSE).

PLCC 反映了预测的精度,可以通过下式计算:

$$\text{PLCC} = \frac{\sum_{i=1}^n (s_i - \bar{s}) \cdot (p_i - \bar{p})}{\sqrt{\sum_{i=1}^n (s_i - \bar{s})^2} \cdot \sqrt{\sum_{i=1}^n (p_i - \bar{p})^2}} \quad (11)$$

式中: s_i, p_i 分别表示第 i 张测试集图片的主观质量分数和预测质量分数; $\bar{s}, \bar{p}$ 分别表示测试集图片主观分数的均值和预测分数的均值.

SRCC 反映了预测的单调性,定义为:

$$\text{SRCC} = 1 - \frac{1}{n \cdot (n-1)} \sum_{i=1}^n (s_i - p_i)^2 \quad (12)$$

式中: s_i, p_i 分别表示第 i 张图片的主观分数和预测分数在主观分数序列和预测分数序列中的排序位置.

KRCC 同样用于衡量预测的单调性,可表示为:

$$\text{KRCC} = \frac{2 \cdot (N_c - N_d)}{n \cdot (n-1)} \quad (13)$$

式中: N_c 和 N_d 分别代表数据库中一致对和非一致对的数量.

RMSE 反映了预测的一致性,通过下式计算:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (s_i - \bar{p})^2} \quad (14)$$

上述四个评估准则中,PLCC、SRCC 和 KRCC 越大, RMSE 越小,则说明该客观质量评价算法有较好的性能.

为了减少预测分数非线性的影响,在计算 PLCC 和 RMSE 之前使用一个五参数逻辑回归函数:

$$Q_p = \beta_1 \cdot \left(\frac{1}{2} - \frac{1}{\exp(\beta_2 \cdot (q - \beta_3))} \right) + \beta_4 \cdot q + \beta_5 \quad (15)$$

式中: q 是原始 IQA 评估指标的值; Q_p 是回归后得到的质量分数; $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ 是该逻辑回归函数的参数。

4.3 实验细节

在正式实验开始之前,需要对本算法的参数设置以及实验设置进行说明。

1) 局部特征提取中块尺寸选取

特征提取是算法的核心步骤,好的特征能够在很大程度上反映图像的属性,局部特征计算时涉及到块尺寸选取,体现在局部对比度、局部熵和分块小波能量的提取过程中。本文通过实验讨论,设置块分割尺寸为 16 像素×16 像素(详细讨论见 4.4)。

2) 训练集/测试集划分

基于学习的方法中,测试集和数据集的比例直接影响最终的实验结果。为保证比较的公平性,本文

遵循前人工作,将整个数据库 80% 图片作为训练集,剩余 20% 作为测试集。

4.4 性能比较

为了验证算法的性能,将提出的算法与目前性能优异的无参考质量评价算法进行比较,这些算法可以分为两类:一是针对自然图像质量评价的算法,例如 BRISQUE^[4], ILNIQE^[12], SSEQ^[13], NFERM^[14], NRSL^[15], NIQE^[30] 和 BSD^[31];二是针对色调映射图像的评价算法,包括 Yue's method^[18], BLIQUE-TMI^[20] 和 BTMQI^[32]。为了与提出的算法公平对比,上述算法中基于学习的算法均将数据库划分为 80% 训练集和 20% 测试集。为避免实验结果的随机性、增强结果的可信度,将训练集、测试集随机划分 1 000 次,并将 1 000 次测试结果的中值作为本文最终的实验结果。不同 IQA 算法性能比较如表 2 所示。粗体表示每种评价准则下性能最优的算法。由于 BLIQUE-TMI 没有公布源码,因此用“—”表示表 2 中缺失的值。

表 2 不同 IQA 算法性能比较

Tab.2 Comparison results of different IQA metrics

Criteria	NIQE ^[30]	ILNIQE ^[12]	SSEQ ^[13]	BRISQUE ^[4]	BSD ^[32]	NFERM ^[14]	NRSL ^[15]	BTMQI ^[31]	Yue ^[18]	BLIQUE-TMI ^[20]	Proposed
PLCC	0.076 8	0.267 8	0.538 7	0.390 1	0.382 7	0.487 0	0.551 6	0.617 5	0.701 9	0.712	0.736 7
SRCC	0.090 4	0.225 9	0.532 7	0.349 3	0.374 7	0.478 9	0.541 5	0.615 2	0.691 5	0.704	0.735 8
KRCC	0.060 3	0.150 9	0.369 3	0.236 7	0.253 2	0.329 3	0.380 5	0.483 4	0.498 9	—	0.538 3
RMSE	10.001	9.446 2	8.415 2	9.215 9	9.238 1	8.729 2	8.329 1	7.867 4	7.132 7	—	6.745 6

从表 2 可以得出以下结论:

1) 本文提出的算法性能胜过各种无参考图像质量评价算法,其预测分数与人眼的主观评估有较高的一致性,具有较高的精确度和单调性;

2) BTMQI, Yue's method, BLIQUE-TMI 以及本文提出的算法普遍优于其他自然图像的评价算法性能。究其原因, NIQE, ILNIQE 等算法采用自然图像特征,适用于分析常见的失真类型,比如高斯白噪声、模糊失真和压缩失真等,这些失真与色调映射图像自身的特性有很大不同,因此性能并不好。而 BTMQI, Yue's method 和 BLIQUE-TMI 则是考虑了色调映射图像在结构、纹理和色彩等方面的信息,更符合色调映射图像自身的特性,因而性能较之自然图像评价算法有很大提升;

3) 本文提出的算法性能优于 BTMQI, Yue's method 和 BLIQUE-TMI 这些同样针对色调映射的

IQA 算法。究其原因,提出的算法结合了图像的全局特征和局部特征,并且考虑了色调映射图像非正常曝光条件下明暗分布的差异。正如图 1(a)(b) 所示的欠曝光和过曝光图像,色调映射算子将 16 比特或者 32 比特的 HDR 图像转化为适合在设备上显示的 8 比特 LDR 图像,这一过程会造成信息损失,其中一个突出的表现就是图像明暗分布不均匀,严重影响了图像质量。因此,提出的算法不仅符合人眼感知图像的过程,而且进一步分析了色调映射图像的特性,从而在性能上优于其他算法。

基于训练的算法在很大程度上会依赖训练集的大小,本文进一步讨论不同训练集大小对提出算法性能的影响。具体而言,训练集划分从 20% 开始,每次 10% 的比例增加直到 80%,相应测试集比例互补的减少,迭代 1 000 次,取相应评估准则的中值。结果如表 3 所示。

表 3 不同训练集大小的性能结果
Tab.3 Performance results with different training set sizes

Train – Test	PLCC	SRCC	KRCC	RMSE
20%~80%	0.664 3	0.666 0	0.473 9	7.500 5
30%~70%	0.687 4	0.688 9	0.493 8	7.291 0
40%~60%	0.703 9	0.704 9	0.508 0	7.129 1
50%~50%	0.714 1	0.714 9	0.517 6	7.025 3
60%~40%	0.722 5	0.723 1	0.525 6	6.920 4
70%~30%	0.730 3	0.729 5	0.531 9	6.843 7
80%~20%	0.736 7	0.735 8	0.538 3	6.745 6

从表 3 可以看出,随着训练图片数据量的增加,相应的算法性能评估指标都有所提升(RMSE 减小),与现有的基于学习的无参考评价算法^[18,20]结论一致.结合表 2 和表 3 发现,即使只采用了 20%的比例作训练集,性能依旧领先其他采用 80%比例作训练集的自然图像评价算法.

块尺寸的大小决定了局部信息的多少,在局部特征提取过程中,局部对比度、局部熵和分块小波能量都与块尺寸的选取有关.本文设置块尺寸为 8 像素 *8 像素,16 像素 *16 像素,32 像素 *32 像素以及 64 像素 *64 像素,以分析块尺寸对算法性能的影响.训练集占数据集的 80%,改变块尺寸,并迭代 1000 次,对 PLCC,SRCC,KRCC 以及 RMSE 的结果取中值.实验结果如图 6 所示.

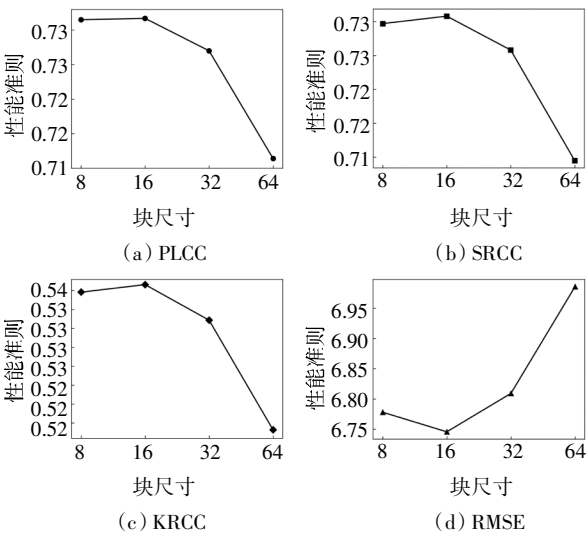


图 6 不同块尺寸的性能结果
Fig.6 Performance results with different block sizes

从图 6 可以看出,不同块尺寸对四个评价准则

的影响并不大,但是块尺寸为 16 像素 *16 像素的效果略好于其它块尺寸.因此,本文在局部特征提取阶段设置块尺寸为 16 像素 *16 像素.

5 结 论

考虑人眼视觉特性,针对色调映射图像,提出了一种结合全局和局部特征的无参考图像质量评价算法.全局特征方面,考虑了颜色矩、明暗分布和信息熵,从图像整体提取特征;局部特征方面,考虑了局部对比度、局部熵和分块小波能量,从图像区域细节入手提取特征.利用 SVR 将融合了全局和局部特征的向量映射到质量分数上.相比于前人研究,更多结合了局部和全局特征分析色调映射图像的特性,提取的特征对质量感知有很强的敏感性.实验结果表明,提出的算法性能优于目前优秀的无参考自然图像评价算法和无参考色调映射图像评价算法,与人眼的主观感知有着很高的一致性,可以应用于监控图像质量等领域.未来的研究中,重点在于色调映射视频的质量研究,使得算法能够评估视频质量从而优化视频传输系统.同时,对于明暗分布特征,本文采用简单的动态范围分割以评估图像质量,初步验证了该想法的有效性.今后会尝试建立极端曝光情形下的图像数据库,根据图像区域分割和曝光特点以期进一步提高算法性能.

参考文献

[1] 侯春萍,刘月,岳广辉,等. 沃尔什哈达玛变换域的无参考图像质量评价[J]. 湖南大学学报(自然科学版),2017,44(8):137—144.
HOU C P,LIU Y,YUE G H,*et al.* No-reference image quality assessment on Walsh Hadamard transform domain [J]. Journal of Hunan University (Natural Sciences),2017,44(8):137—144. (In Chinese)
[2] YUE G,HOU C,ZHOU T. Blind quality assessment of tone-mapped images considering colorfulness,naturalness,and structure [J]. IEEE Transactions on Industrial Electronics,2019,66(5):3784—3793.
[3] WANG Z,BOVIK A C,SHEIKH H R,*et al.* Image quality assessment:from error visibility to structural similarity [J]. IEEE Transactions on Image Processing,2004,13(4):600—612.
[4] MITTAL A,MOORTHY A K,BOVIK A C. No-reference image quality assessment in the spatial domain [J]. IEEE Transactions on Image Processing,2012,21(12):4695—4708.

- [5] YUE G, HOU C, GU K, *et al.* Analysis of structural characteristics for quality assessment of multiply distorted images [J]. IEEE Transactions on Multimedia, 2018, DOI: 10.1109/TMM.2018.2807589.
- [6] YUE G, HOU C, GU K, *et al.* No reference image blurriness assessment with local binary patterns [J]. Journal of Visual Communication and Image Representation, 2017, 49: 382—391.
- [7] YUE G, HOU C, ZHOU T. Effective and efficient blind quality evaluator for contrast distorted images [J]. IEEE Transactions on Instrumentation and Measurement, 2018, DOI: 10.1109/TIM.2018.2868555.
- [8] LU Q, ZHOU W, LI H. A no-reference image sharpness metric based on structural information using sparse representation [J]. Information Sciences, 2016, 369: 334—346.
- [9] LI L, WU D, WU J, *et al.* Image sharpness assessment by sparse representation [J]. IEEE Transactions on Multimedia, 2016, 18(6): 1085—1097.
- [10] LI L, XIA W, LIN W, *et al.* No-reference and robust image sharpness evaluation based on multiscale spatial and spectral features [J]. IEEE Transactions on Multimedia, 2017, 19(5): 1030—1040.
- [11] LI L, ZHOU Y, LIN W, *et al.* No-reference quality assessment of deblocked images [J]. Neurocomputing, 2016, 177(C): 572—584.
- [12] ZHANG L, ZHANG L, BOVIK A C. A feature-enriched completely blind image quality evaluator [J]. IEEE Transactions on Image Processing, 2015, 24(8): 2579—2591.
- [13] LIU L, LIU B, HUANG H, *et al.* No-reference image quality assessment based on spatial and spectral entropies [J]. Signal Processing Image Communication, 2014, 29(8): 856—863.
- [14] GU K, ZHAI G, YANG X, *et al.* Using free energy principle for blind image quality assessment [J]. IEEE Transactions on Multimedia, 2014, 17(1): 50—63.
- [15] LI Q, LIN W, XU J, *et al.* Blind image quality assessment using statistical structural and luminance features [J]. IEEE Transactions on Multimedia, 2016, 18(12): 2457—2469.
- [16] REINHARD E, WARD G, PATTANAIK S, *et al.* High dynamic range imaging: acquisition, display, and image-based lighting [M]. USA: Princeton University Press, 2005.
- [17] GU K, ZHAI G, LIU M, *et al.* Details preservation inspired blind quality metric of tone mapping methods [C]// IEEE International Symposium on Circuits and Systems. IEEE, 2014: 518—521.
- [18] YUE G, HOU C, GU K, *et al.* Biologically inspired blind quality assessment of tone-mapped images [J]. IEEE Transactions on Industrial Electronics, 2018, 65(3): 2525—2536.
- [19] JIANG G, SONG H, YU M, *et al.* Blind tone-mapped image quality assessment based on brightest/darkest regions, naturalness and aesthetics [J]. IEEE Access, 2018, 6: 2231—2240.
- [20] JIANG Q, SHAO F, LIN W, *et al.* BLIQUE-TMI: Blind quality evaluator for tone-mapped images based on local and global Feature Analyses [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2017, 29(2): 323—335.
- [21] GRAHAM N V S. Visual pattern analyzers [M]. New York: Oxford University Press, 1989.
- [22] BAE S H, KIM M. A novel image quality assessment with globally and locally consistent visual quality perception [J]. IEEE Transactions on Image Processing, 2016, 25(5): 2392—2406.
- [23] STRICKER A S M A, ORENGO M. Similarity of color images [J]. Proc Spie Storage & Retrieval for Image & Video Databases, 1995, 2420: 381—392.
- [24] WANG Y F, HUANG Q, HU J. Image enhancement based on adaptive demarcation between underexposure and overexposure [C]// 2017 International Conference on Progress in Informatics and Computing (PIC). Nanjing, 2017: 196—199.
- [25] 李树涛, 孙凤梅. 基于 Curvelet 和小波变换的纹理图像识别 [J]. 湖南大学学报(自然科学版), 2008, 35(4): 51—54.
LI S T, SUN F M. Texture recognition based on Curvelet and wavelet transformation [J]. Journal of Hunan University (Natural Sciences), 2008, 35(4): 51—54. (In Chinese)
- [26] 周四望, 刘龙康. 基于小波变换的图像零树压缩感知方法 [J]. 湖南大学学报(自然科学版), 2017, 44(2): 129—136.
ZHOU S W, LIU L K. Image zerotree compressed sensing based on wavelet transform [J]. Journal of Hunan University (Natural Sciences), 2017, 44(2): 129—136. (In Chinese)
- [27] WU X Q, WANG K Q, ZHANG D. Wavelet energy feature extraction and matching for palmprint recognition [J]. Journal of Computer Science and Technology, 2005, 20(3): 411—418.
- [28] SCHOLKOPF B, SMOLA A J. Learning with kernels: Support vector machines, regularization, optimization, and beyond [M]. London: MIT Press, 2001.
- [29] KUNDU D, GHADIYARAM D, BOVIK A C, *et al.* Large-scale crowd sourced study for tone-mapped HDR pictures [J]. IEEE Transactions on Image Processing, 2017, 26(10): 4725—4740.
- [30] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a “completely blind” image quality analyzer [J]. IEEE Signal Processing Letters, 2013, 20(3): 209—212.
- [31] LI Q, LIN W, FANG Y. BSD: Blind image quality assessment based on structural degradation [J]. Neurocomputing, 2017, 236: 93—103.
- [32] GU K, WANG S, ZHAI G, *et al.* Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure [J]. IEEE Transactions on Multimedia, 2016, 18(3): 432—443.