

基于人眼感知的无参考色调映射图像质量评价

马华林^{1,2} 张立燕¹

¹(浙江工商职业技术学院 浙江 宁波 315012)

²(宁波大学信息科学与工程学院 浙江 宁波 315012)

摘要 为了在标准显示器上显示高动态范围(High Dynamic Range, HDR)图像,需要用色调映射算子(Tone-mapped Operator, TMO)把 HDR 图像压缩成低动态范围的色调映射(Tone-Mapped, TM)图像。为了评价不同的色调映射算子的性能,提出一种盲参考 TM 图像质量评价方法。用聚类算法把图像分成暗区、中间区、亮区,在三个区分别提取信息熵和面积比率特征;在显著性区域提取信息熵和面积比率特征,在亮度通道和 RGB 颜色通道提取自然度特征;用支持向量回归模型训练得到质量预测分数。在 TMID 库上的实验结果中,PLCC 达到 0.883 0, SROCC 达到 0.831 3,表明该算法十分有效;与色调映射图像质量评价算法相比具有更好的性能。

关键词 色调映射图像 图像质量评价 聚类 自然度 显著性区域

中图分类号 TP391 文献标志码 A DOI:10.3969/j.issn.1000-386x.2020.07.026

NO-REFERENCE TONE-MAPPED IMAGE QUALITY ASSESSMENT BASED ON HUMAN EYE PERCEPTION

Ma Hualin^{1,2} Zhang Liyan¹

¹(Zhejiang Business Technology Institute, Ningbo 315012, Zhejiang, China)

²(Faculty of Information Science and Engineering, Ningbo University, Ningbo 315012, Zhejiang, China)

Abstract In order to display a high dynamic range(HDR) image on a standard monitor, it is necessary to use Tone-mapped Operator(TMO) to compress HDR image into Tone-mapped(TM) image with low dynamic range. In order to evaluate the performance of different TMOs, this paper proposes a no-reference TM image quality evaluation method. The image was divided into dark area, middle area and bright area by using the clustering algorithm, and the information entropy and area ratio features were extracted respectively in the three areas; the information entropy and area ratio features were extracted in the significant region, and the naturalness feature was extracted from the brightness channel and the RGB color channel; we used support vector regression model to obtain quality prediction scores. The experimental results on the TMID library show that PLCC reaches 0.883 0 and SROCC reaches 0.831 3. It shows that the proposed algorithm is very effective. Compared with the tone mapping image quality evaluation algorithm, it has better performance.

Keywords Tone-mapped image Image quality assessment Clustering Naturalness Salient region

0 引言

HDR 图像相比低动态范围(Low Dynamic Range, LDR)图像能够表示更大的亮度范围,其亮度范围大约是 10^{-4} cd/m² ~ 10^5 cd/m²[1-2]。LDR 图像能表示的动

态范围不超过 3 个数量级,但人类视觉系统在真实场景中可以接受的动态范围可以达到 6 个数量级。因此,HDR 图像对用户来说感受更真实更有吸引力。随着成像和计算机图形学技术的发展,HDR 图像的获得越来越容易。然而,HDR 显示设备比较昂贵^[3],超出了普通消费者的承受范围。为了解决这个问题,产业

界和学术界开发了许多色调映射算子^[4],TMO 能够把 HDR 图像转换成 LDR 图像,转换后的 LDR 图像被称为色调映射图像 (Tone-mapped Image, TMI)。由于 TMI 相对 HDR 图像缩小了动态范围,不可避免地产生失真,例如亮度信息丢失、结构信息丢失、不自然的颜色等。针对一个 HDR 图像,不同的 TMO 产生的效果不一样。因此,TMI 质量评价对于选择合适的 TMO 和改进 TMO 本身具有重要的研究意义。

TMI 质量评价分为主观质量评价和客观质量评价,而早期的 TMO 性能评价主要采用主观质量评价。因为人眼是 TMI 的最终接受者,主观质量评分具有准确性高的特点。然而,主观质量评价^[5]具有三方面的缺点:主观质量评价比较耗时费力,需要昂贵的 HDR 显示器和多个测试人员多次实验;主观质量不能嵌入到图像处理系统来改进 TMO;主观评价由于人为的一些不确定因素导致评价的误差,例如对于主观评分相近的两个图像人眼也很难判断哪个图像更好一点。传统的客观质量评价方法假定参考图像和测试图像具有相同的动态范围,因此传统的客观质量评价方法不能直接评价 TMI。

近年来,全参考 TMI 质量评价算法取得了丰硕的成果。Yeganeh 等^[5]是最先关注 TMI 质量评价的学者之一。他们首次建立了色调映射图像数据库 (Tone-mapped Image Database, TMID),可以下载并用于评价 TMI 质量评价算法的性能。同时,他们提出了一种全参考 TMI 质量评价方法 (Tone Mapped Image Quality Index, TMQI)。这个方法的基本思想是高质量的 TMI 不但要保护 HDR 图像的结构信息,还要保留图像的自然场景统计特性 (Natural Scene Statistics, NSS)。Nafchi 等^[6]基于图像的局部相位信息提出了 FSITM 方法,该方法考虑了图像的颜色信息,但没有考虑图像的自然度。Kundu 等^[7]针对 TMQI 均匀池化的缺点,在 TMQI 基础上加入了视觉注意力模型,运用感知池化策略提升质量评价算法的性能。Xie 等^[8]使用字典学习技术在稀疏域提取局部结构相似度和全局自然度,合并这两个特征提出了一种全参考质量评价算法 SMTI。鉴于色调映射图像失真类型多的特点,Hadizadeh 等^[9]从结构保真度、自然度、亮度、颜色等方面提取了八类特征来评价 TMI 质量。

无参考 TMI 质量评价相对全参考 TMI 质量评价具有更强的实用性。由于 TMI 失真通常不会出现模糊、块效应等类型的失真,传统的无参考质量评价算法不适合评价 TMI。Gu 等^[10]结合信息量、结构、自然度三方面提出了 BTMQI 算法,该算法取得了很好的效

果,但没有考虑 TMI 的颜色失真。Yue 等^[11]模拟人脑对颜色信息的处理过程来评价 TMI 质量,从简单细胞的响应中提取纹理信息特征,从复杂细胞的响应中提取结构和纹理信息。本文提出一种基于人眼感知的无参考色调映射图像质量评价方法,主要贡献如下:

(1) 首次采用聚类算法针对 TMI 高亮区、低暗区曝光异常的特点把色调映射图像分成暗区、中间区、亮区,根据物理亮度距离与聚类中心亮度值的特性,优化了使用 K-means 聚类对 TMI 分区的方法。相对不分区的色调映射图像质量评价方法显著提高算法性能。

(2) 在色调映射图像质量评价研究中首次使用非负矩阵分解方法,根据 TMI 灰度图像非负矩阵分解获得的系数特性,提出了一种新的图像显著性区域识别方法。

(3) 结合亮度和颜色两种信息来提取图像自然度,发现了亮度自然度和颜色自然度的互补性,相对单独使用颜色自然度或者亮度自然度显著提高了算法的 SROCC 指标。

1 评价方法设计

根据 Weber 定律,人眼能感受到的最小亮度变化应满足 $\Delta S = KS$,其中 K 为常量,当亮度值 S 相近时,对应的能被人眼感知的最小亮度变化范围 ΔS 也相近,按照祖母神经元理论人眼感受到亮度变化的原因是由一组感受 $S + \Delta S$ 亮度的神经细胞兴奋所致。因此,聚集亮度感知变化范围相似的细胞为同一个类别,分别分析每一类细胞感知到的图像质量。本文首先把 TMI 转换成灰度图像,然后使用 K-means 聚类算法把灰度图像按照亮度值进行聚类,每一类对应一个区域,在多个区域分别提取特征来评价图像质量。

根据 TVI 曲线人眼视觉系统具有 5 个亮度适应等级^[12],因此分别把色调映射图像分成 3 个区域、4 个区域、5 个区域在 TMID^[5] 图像数据库上进行实验。例如对于三个区域的划分,首先把色调映射图像用 K-means 聚类算法分成 3 个区域,分别提取每个区域的信息熵特征,然后用 SVM 聚合三个信息熵特征预测图像质量评分。选取 Pearson 线性相关系数 (Pearson Linear Correlation Coefficient, PLCC)、Spearman 秩相关系数 (Spearman Rank-order Correlation Coefficient, SROCC) 作为评价指标分别评价图像预测值与主观评分的准确性和一致性。

如表 1 所示,把色调映射图像划分成三个区域时,信息熵特征的性能指标是最好的,这三个区域分别是亮区、中间区、暗区。色调映射过程会缩小 TMI 的动

态范围,导致 TMI 图像的高亮区和低暗区细节丢失,从而影响 TMI 的质量。如图 1 所示,亮区包含了曝光过度的区域(高亮区域),同理,对于曝光不足的图像,暗区包含低暗区,因此三个区域的划分体现了 TMI 图像的特性。

表 1 不同区域划分方法信息熵特征的性能指标

| 区域类别 | PLCC | SROCC |
|------|----------------|----------------|
| 三个区域 | 0.836 8 | 0.730 6 |
| 四个区域 | 0.802 2 | 0.678 4 |
| 五个区域 | 0.831 0 | 0.730 2 |



(a) 灰度图 (b) 分成三个区域后

图 1 TMI 数据库中的一幅图

K-means 聚类时只考虑像素点到聚类中心的距离来确定像素点的类别,没有考虑人眼对亮度感知的非线性特性。设聚类中心亮度值为 S_1 ,感知亮度为 P_1 ,某个像素点像素亮度值为 S_2 ,感知亮度为 P_2 。根据费希纳定理:

$$\begin{cases} P_1 = \lg(S_1) \\ P_2 = \lg(S_2) \end{cases} \quad (1)$$

则亮度感知差 $P_2 - P_1$ 和 S_2 与 S_1 的关系为:

$$\begin{cases} P_2 - P_1 = \lg(S_1/S_2) \\ S_2 = S_1 \times 10^{P_2 - P_1} \end{cases} \quad (2)$$

在 K-means 算法中 S_1 和 S_2 的物理亮度距离为:

$$D = (S_2 - S_1)^2 = S_1^2 \times (10^{P_2 - P_1} - 1)^2 \quad (3)$$

对于相近的人眼感知差 $P_2 - P_1$,物理亮度距离与聚类中心的亮度值成正比,即人眼对暗区聚类时与聚类中心的距离要短一些,对中间区域聚类时与聚类中心的距离要长一些。因此,采用如下算法对色调映射图像的三个区域进行调整。

设 L 为图像像素亮度值, C_1, C_2, C_3 分别为暗区、中间区、亮区的聚类中心,则三个区域的分类如下:

$$\begin{cases} \text{暗区}: 0 \leq L < C_1 \\ \text{中间区}: C_1 \leq L < C_3 \\ \text{亮区}: C_3 \leq L \end{cases} \quad (4)$$

如表 2 所示,采用式(4)对三个区域的分区进行调整后,信息熵特征的 PLCC 和 SROCC 都明显提升,

说明分区的调整方法符合人眼的感知特性。

表 2 区域调整前后信息熵特征的性能指标

| 情况 | PLCC | SROCC |
|-------|----------------|----------------|
| 分区调整前 | 0.836 8 | 0.730 6 |
| 分区调整后 | 0.845 0 | 0.754 2 |

TMI 失真主要表现为信息细节的丢失、不自然的颜色、全局亮度和对比度的失真。因此,本文算法综合考虑了 TMI 的失真特点。如图 2 所示,本文主要提取了聚类感知特征、显著性区域特征、自然度特征。聚类感知特征在亮度域上提取,先把 TMI 转换成灰度图,根据亮度信息进行聚类,把图像分成亮区、暗区、中间区三个区域,在每个区域分别提取面积比率和信息熵两个特征。生理学和心理学的一些证据^[13]表明人眼观察一幅图像时先看图像的整体,人脑会抑制图像中高频出现的特征,视觉注意力容易关注偏差比较大的区域,即显著性区域。假设图像中一个 $M \times M$ 区域具有亮区、暗区、中间区这三个区域中两个或两个以上区域的像素,这样的区域称为混合区域。混合区域中像素间的亮度值差异比较大,具有显著性区域的特点。本文通过对 TMI 的非负矩阵分解得到测试图像对应的系数,对系数进行分析识别出 TMI 的混合区域,在混合区域提出块比例、信息熵特征。由于 TMO 处理过程缩小了 HDR 图像的动态范围,影响了 TMI 的自然度,因此提取了亮度通道和颜色通道的自然统计特征。最后使用机器学习方法对所有特征进行回归,提出了无参考 TMI 质量评价方法。

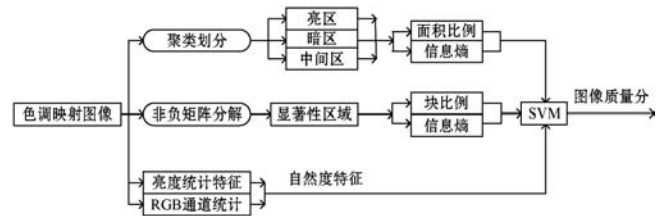


图 2 无参考 TMI 质量评价框架

1.1 聚类感知特征

在视觉注意处理中,人类视觉系统对视觉细胞产生的刺激进行融合,通过不同的信息特征进行聚类,形成人类视觉系统的注意力分配图,因此聚类是人类视觉系统的固有功能。设 $m \times n$ 的灰度图像 $I, I(x, y)$ 表示图像 I 中 (x, y) 像素点的亮度值,设暗区、中间区、亮区的像素亮度值集合分别为 R_L, R_M, R_H , 则:

$$\begin{cases} R_L = \{0 \leq I(x, y) < C_1 \mid 0 \leq x \leq m, 0 \leq y \leq n\} \\ R_M = \{C_1 \leq I(x, y) < C_3 \mid 0 \leq x \leq m, 0 \leq y \leq n\} \\ R_H = \{I(x, y) \geq C_3 \mid 0 \leq x \leq m, 0 \leq y \leq n\} \end{cases} \quad (5)$$

式中: C_3 、 C_1 为亮区和暗区的聚类中心。信息熵是衡量信息量的有效方法,设 p 为概率密度,对 R_L 、 R_M 、 R_H 、 I 分别求信息熵为:

$$E_{\Omega} = - \int p(x) \log p(x) dx \quad \Omega \in \{L, M, H, G\} \quad (6)$$

E_L 、 E_M 、 E_H 分别表示暗区、中间区、亮区的熵。考虑到人眼观察图像时先整体后局部的特点,同时提取全局信息熵 E_G 。

由于TMI图像容易出现过曝光或欠曝光的区域,这些区域的大小会影响图像的质量,因此提取了三个区域的面积比例作为特征。设 $N(\cdot)$ 函数表示计算图像或者图像块的像素个数,则每个区域的面积比例可以表示为:

$$Ratio_{\Omega} = \frac{N(R_{\Omega})}{N(I)} \quad \Omega \in \{L, M, H\} \quad (7)$$

$Ratio_L$ 、 $Ratio_H$ 、 $Ratio_M$ 分别表示暗区、亮区、中间区的面积比率。聚类感知特征向量 $F_{cluster}$ 为:

$$F_{cluster} = \{E_L, E_M, E_H, E_G, Ratio_L, Ratio_M, Ratio_H\}$$

1.2 显著性区域特征

当人眼看一幅图像时注意力会被吸引到图像中的一部分区域,这部分区域被称为显著性区域,是大部分人认为图像中重要的或者显著的部分。显著性区域的图像质量显然影响人眼对图像整体质量的评价。Goferman等^[13]认为人眼观察一幅图像时先看图像的整体,人脑会抑制图像中高频出现的特征,视觉注意力容易关注偏差比较大的区域,即显著性区域。本文对TMI灰度图像的非负矩阵分解获得对应图像的系数,通过对系数的直方图的分析,提出了一种混合区域的识别方法,然后对混合区域提取信息熵,块比率等特征。

许多研究表明稀疏表示符合人脑对图像信号的认知,非负矩阵分解(NMF)与稀疏表示的字典学习类似。非负矩阵分解是把一个数据矩阵 M 分解为两个非负矩阵 W 和 S 的乘积, W 为特征矩阵对应字典学习的原子矩阵, S 为编码矩阵。数据矩阵 M 可以看作特征矩阵 W 中每一列和 S 中对应系数的线性组合,由于 S 的非负性, M 是由 W 中的每一列按照 S 决定的权重系数累加而成,由于 W 中的每一列就是一个图像块,因此非负矩阵分解与部分组成整体的直观认知相符。

为了得到特征矩阵 W ,需要建立一个训练库。本文从TMID数据库中选出了不同场景和纹理的10幅色调映射图像,如图3所示,先把每一幅TMI转换成灰度图像,然后把灰度图像分成 $\sqrt{m} \times \sqrt{m}$ 的图像块,把每幅图像块转换成 m 维的列向量 M_i 。组合所有训练图像块的列向量生成 $m \times n$ 大小的训练样本库 $M =$

$[M_1, M_2, \dots, M_n]$ 。本文中 $m = 4900, n > 20000$ 。



图3 从TMID数据库中选出的原始训练图像

对于训练样本库 M ,NMF的目标是寻找特征矩阵 $W = [W_1, W_2, \dots, W_r] \in \mathbf{R}^{m \times r}$ 和 $S = [S_1, S_2, \dots, S_n] \in \mathbf{R}^{r \times n}$ 来近似训练样本矩阵 M :

$$M = WS \quad (8)$$

式中: r 是大于零的整数,表示特征矩阵中列向量的个数。 W 和 S 的寻找过程可以转化为如下优化问题:

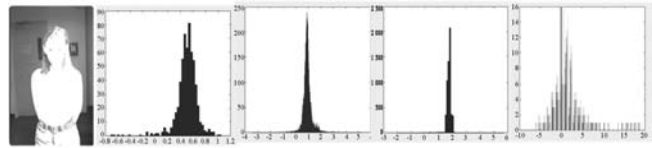
$$\langle W, S \rangle = \arg \min_{W \geq 0, S \geq 0} \|X - WS\|_F^2 \quad (9)$$

本文用Lin^[14]的方法计算出了 W 和 S , W 就是通过训练库训练得到的特征矩阵。对于一个测试图像块转化而成的列向量 $T_i \in \mathbf{R}^{m \times 1}$,则得到 T_i 非负矩阵分解后的编码矩阵 $F_i \in \mathbf{R}^{r \times 1}$ 为:

$$T_i = WF_i \Rightarrow F_i = (W^T W)^{-1} W^T T_i \quad (10)$$

式中: $(W^T W)^{-1} W^T$ 是 W 的伪逆矩阵。对于测试图像 $T = [T_1, T_2, \dots, T_n]$ 获得编码矩阵 $F = [F_1, F_2, \dots, F_n]$, n 为一幅测试图像包含的图像块的个数。

TMI的混合性区域是区域中像素来自两种或两种以上聚类感知区域,例如一个图像块有三分之一的亮区像素和三分之二的中间区像素。把亮区、暗区、中间区、混合区图像块分别进行非负矩阵分解,对每个区域的编码矩阵进行直方图分析,如图4所示。



(a) 原图 (b) 暗区系数 (c) 中间区系数 (d) 亮区系数 (e) 混合区系数

图4 亮区、暗区、中间区、混合区编码矩阵直方图

可以看出,图像暗区、中间区、亮区的编码矩阵的系数值比较小,混合区系数有较大值。因此可以通过分析编码矩阵的最大值来判断显著性区域,设阈值 TH ,定义显著性区域 S 如下:

$$S = \{T_i \mid \max(F_i) > TH, F_i = (W^T W)^{-1} W^T T_i\} \quad (11)$$

$$i = 1, 2, \dots, n$$

信息熵能有效衡量图像的信息量,根据式(6)对 S 的所有像素求信息熵获得显著性区域的信息量 E_S 。显著性区域面积越大,对图像质量的影响越大,因此提取显著性区域的面积比例 $Ratio_S$:

$$Ratio_S = \frac{N(S)}{N(I)} \quad (12)$$

式中: $N(\cdot)$ 函数表示计算图像或者图像块的像素个数; I 表示图像。显著性特征 F_s 为:

$$F_s = \{E_s, Ratio_s\}$$

1.3 自然度特征

HDR 图像经过色调映射后可能会出现曝光过度或者曝光不足的现象,造成 TMI 看起来不自然。然而,高质量的 TMI 不应当破坏其自然特性。TMI 的自然度失真主要体现在图像过亮、过暗、不自然的颜色,因此本文考虑基于亮度和颜色提取自然度特征。

(1) 亮度自然度特征。本文采用了 TMQI^[5]的自然度统计模型,该模型使用 3 000 幅包括动物、夜景、建筑、草地等 14 个不同类别场景的自然图像。首先把每一幅图像转化成灰度图像,然后把灰度图像分割成 11×11 的图像块,分别求出每个图像块的均值与标准差,最后统计一幅图像中所有图像块的均值与标准差的均值,获得一幅图像的均值和标准差。图像的均值、标准差分别与高斯概率密度函数和 Beta 概率密度函数能够很好地拟合。两个概率密度函数计算如下:

$$P_m(m) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{m - \mu_m}{2\alpha_m}\right] \quad (13)$$

$$P_d(d) = \frac{(1-d)^{\beta_d-1} d^{\alpha_d-1}}{B(\alpha_d, \beta_d)} \quad (14)$$

式中: $B(\cdot)$ 是 Beta 函数,模型的参数设置为 $\mu_m = 115.94$, $\alpha_m = 27.99$, $\alpha_d = 4.4$, $\beta_d = 10.1$ 。TMQI 的亮度和对比度的联合概率如下:

$$N = \frac{1}{K} P_m P_d \quad (15)$$

式中: K 是随着 P_m 和 P_d 改变的标准化因子, $K = \max\{P_m, P_d\}$ 使得统计自然度 N 标准化。

(2) 颜色自然度。本文采用 Wang 等^[15]提出的基于颜色空间的自然度统计方法,该方法表明局部标准化颜色系数(local normalized color coefficients, LNCC)服从高斯分布, LNCC 可以表示为:

$$LNCC(i, j) = \frac{I_c(i, j) - u(i, j)}{\sigma(i, j) + 1} \quad (16)$$

式中: $I_c(i, j)$ 是给定图像的 C 颜色通道某个像素的值, (i, j) 是图像的空域坐标。

$$u(i, j) = \sum_{k=-7}^7 \sum_{l=-7}^7 w_{k,l} I_c(i-k, j-l) \quad (17)$$

$$\sigma(i, j) = \sqrt{\sum_{k=-7}^7 \sum_{l=-7}^7 w_{k,l} [I_c(i-k, j-l) - u(i-k, j-l)]^2} \quad (18)$$

采用广义高斯函数(generalized Gaussian distribution, GGD)来拟合 LNCC 参数。GGD 概率密度函数如下:

$$f(x, \alpha, \sigma^2) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} \exp\left(-\left(\frac{|x|}{\beta}\right)^\alpha\right) \quad (19)$$

式中: $\Gamma(\cdot)$ 是伽马函数。 β 为:

$$\beta = \sigma \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}} \quad (20)$$

式中: α 是形状参数, σ^2 是尺度参数,针对每一个颜色通道 C ,用 GGD 进行拟合获得高斯分布参数(α_c , σ_c^2)。由于 RGB 颜色空间没有分离亮度、色度、饱和度, RGB 颜色空间也不是完全独立的颜色,因此文献[15]认为 RGB 颜色通道不适合做图像质量评价,但其实验对象为低动态图像,本文在 TMID 数据库上对 RGB、LAB、YCbCr 颜色空间进行实验。

如表 3 所示, RGB 颜色空间自然度特征获得了最好的性能,因此选择 RGB 颜色空间来提取颜色自然度特征 F_c :

$$F_c = \{\alpha_c, \sigma_c^2 \mid C \in \{R, G, B\}\}$$

表 3 RGB、LAB、YCbCr 颜色空间自然度特征性能指标

| 颜色空间 | PLCC | SROCC |
|-------|----------------|----------------|
| RGB | 0.787 5 | 0.669 8 |
| LAB | 0.728 0 | 0.542 3 |
| YCbCr | 0.758 3 | 0.573 8 |

1.4 聚合策略

本文提取了聚类感知、显著性区域、自然度三类特征共 16 个,设 V 是 TMI 的特征向量,可以表示为:

$$V = \{F_c, F_s, N, F_c\} \quad (21)$$

特征提取之后需要一个回归模型来建立一个函数,这个函数能够映射特征向量到主观图像质量评价分。由于支持向量机(Support Vector Machine, SVM)在图像处理领域得到广泛的应用,因此把训练图像的特征向量和对应的 MOS 值输入支持向量机训练出一个预测模型 $f(\cdot)$,把测试图像的特征向量输入预测模型获得图像的客观质量预测值 Q 。

2 实验结果与分析

2.1 实验数据库和性能指标

本文在 TMID 数据库上对提出的算法进行了相关性性能指标验证。TMID 数据库由加拿大滑铁卢大学开发,数据库中包括 15 幅不同场景的 HDR 图像,在每一幅 HDR 图像上使用 8 种 TMO 总共生成 120 幅 TMI。主观实验邀请了 20 位测试人员,分别将每幅 HDR 图像对应的 8 幅 TMI 按主观质量的好坏进行排序,序号

1~8 作为分数,然后将每幅 TMI 的平均分作为主观质量分数。

根据视频质量专家组 (video quality expert group, VQEG) 的建议,PLCC 和 SROCC 两个参数经常被用来检测各种图像质量评价方法的性能。SROCC 用来评价客观质量评价算法和主观评价的一致性。PLCC 用来评价客观质量评价算法和主观评价的准确性。在计算 PLCC 之前,为了消除主观质量评价引入的非线性因素,需对客观质量评价结果和主观质量评价结果进行非线性拟合,本文采用了五参数 Logistic 拟合方法。PLCC 和 SROCC 值越大表示客观质量评价算法的性能越好。

2.2 训练集大小的影响

采用交叉验证的方法来评价本文提出的无参考色调映射质量评价方法。首先将 TMID 图像库中的 TMI 分为两个不重叠的图像子集。第一子集为训练图像,包含了数据库中 $120 \times n\%$ 的测试色调映射图像,并用这部分图像建立本文所需的预测模型。第二个子集为余下的 $120 \times (1 - n\%)$ 色调映射图像并作为测试图像。为了消除性能偏差,本文随机重复 1 000 次之前的训练-测试过程,并将这 1 000 次测试所得到的性能指标的中值作为最终的性能指标。如表 4 所示,随着训练集尺寸的增大,预测模型的性能逐渐提高,但当训练集达到 80% 时,SROCC 性能指标提升不再明显。

表 4 不同大小训练集的性能指标

| 训练-测试 | PLCC | SROCC |
|----------|---------|---------|
| 50% -50% | 0.846 7 | 0.775 9 |
| 60% -40% | 0.852 5 | 0.793 7 |
| 70% -30% | 0.872 4 | 0.822 3 |
| 80% -20% | 0.883 0 | 0.831 3 |
| 90% -10% | 0.895 0 | 0.831 5 |

2.3 特征分析

本文算法提取了三类特征,具体包括聚类感知特征、自然度特征、显著性区域特征,如表 5 所示。聚类感知特征包括全局信息熵、三个区域的信息熵、三个区域的面积比率,全局信息熵可以看作是整个图像为一类的特殊情况。信息熵越大,说明图像的信息量越大,对图像分亮区、中间区、暗区提取的信息熵相对全局信息熵具有更好的性能指标,说明对图像聚类分区域感知图像质量的方法是有效的。三个区域信息熵特征和全局信息熵的组合显著提升了性能指标,说明三个区域信息熵特征可以弥补全局信息熵特征的不足。由于

TMI 图像容易出现过曝光或欠曝光的区域,这些区域的大小会影响图像的质量,三个区域的面积比例 $Ratio_{\Omega}$ 和 E_L 、 E_M 、 E_H 、 E_C 的组合提升了性能指标。自然度特征包括亮度通道特征 N 和 RGB 颜色通道特征 F_C ,自然度显然与聚类感知特征具有很好的互补性,自然度和聚类感知特征结合后性能指标大大提高。显著性区域提取的混合区域特征 E_S 和块比例 $Ratio_S$ 结合使用比单独使用效果好很多,说明这两个特征具有互补作用。聚类感知特征、自然度、显著性特征联合使用后算法的性能指标进一步提高,说明显著性特征对 TMI 质量有一定的影响。

表 5 本文算法中各特征的性能指标

| 特征 | PLCC | SROCC |
|-----------------------------|---------|---------|
| E_C | 0.680 6 | 0.652 3 |
| E_L, E_M, E_H | 0.793 3 | 0.694 9 |
| E_L, E_M, E_H, E_C | 0.845 0 | 0.754 2 |
| $Ratio_L, Ratio_M, Ratio_H$ | 0.741 6 | 0.635 3 |
| $F_{cluster}$ | 0.871 9 | 0.770 5 |
| N | 0.715 9 | 0.664 9 |
| F_C | 0.787 5 | 0.669 8 |
| $F_{cluster}, F_C, N$ | 0.873 3 | 0.826 4 |
| E_S | 0.612 0 | 0.528 3 |
| $Ratio_S$ | 0.735 6 | 0.593 3 |
| $E_S, Ratio_S$ | 0.793 5 | 0.698 1 |
| $F_{cluster}, F_C, N, F_S$ | 0.883 0 | 0.831 3 |

2.4 算法比较

为了验证本文算法的准确性和有效性,在 TMID 数据库上与现有的代表性算法进行了比较,如表 6 所示。选择了有代表性的全参考和无参考图像质量评价方法作为对比。全参考图像质量评价方法包括 TMQI^[5]、FSITM^[6]、NITI^[16]。无参考图像质量评价方法包括 BRISQUE^[17]、BLINDS-II^[18]、Yue^[11]、BT-MQI^[10]。TMQI、FSITM、NITI 等专门为 TMI 设计的全参考图像质量评价方法要优于 BRISQUE、BLINDS-II 等针对 LDR 图像的无参考图像质量评价方法,说明针对 LDR 的无参考图像质量评价方法不适合 TMI 的质量评价。BTMQI 的性能超越了 TMQI、FSITM 等全参考方法,说明对 TMI 质量评价,无参考质量评价方法也能取得很好的效果。实验结果表明在 TMID 数据库上,对于 PLCC 和 SROCC 两个性能指标,本文算法优于现有代表性的算法。

表 6 本文算法与代表性算法的比较

| 类型 | 算法 | PLCC | SROCC |
|-----|------------|---------|---------|
| 全参考 | TMQI | 0.771 5 | 0.740 7 |
| | FSITM | 0.752 5 | 0.702 8 |
| | NITI | 0.878 9 | 0.830 1 |
| 盲参考 | BLIINDS-II | 0.501 9 | 0.442 9 |
| | BRISQUE | 0.548 1 | 0.481 0 |
| | Yue | 0.824 3 | 0.732 1 |
| | BTMQI | 0.854 1 | 0.828 2 |
| | 本文算法 | 0.883 0 | 0.831 3 |

3 结 语

本文受 Weber 定律启发,采用聚类感知 TMI 图像的质量。用 K-means 聚类算法把 TMI 分成暗区、中间区、亮区,提取三个区域的信息熵、面积比率和全局信息熵作为特征。通过分析图像编码矩阵的系数提取显著性区域,并在显著性区域提取信息熵、面积比率特征。在亮度和颜色通道提取了自然度特征。最后用 SVM 融合全部提取的特征获得 TMI 的质量分。实验表明,本文算法在 TMID 数据库上具有很好的性能。

参 考 文 献

- [1] Chen H W, Zhu R D, Li M C, et al. Pixel-by-pixel local dimming for high-dynamic-range liquid crystal displays[J]. Optics Express, 2017, 25(3):1973-1984.
- [2] Artusi A, Mantiuk R K, Richter T, et al. JPEG XT: a compression standard for HDR and WCG images [standards in a nutshell] [J]. IEEE Signal Processing Magazine, 2016, 33(2):118-124.
- [3] Shambulinga M, Thanuja T C, Ramanathan N. Tone mapping of HDR image for LDR display using guided image filter [J]. International Journal of Scientific & Engineering Research, 2013, 4(6):3034-3037.
- [4] Ma K, Yeganeh H, Zeng K, et al. High dynamic range image compression by optimizing tone mapped image quality index[J]. IEEE Transactions on Image Processing, 2015, 24(10):3086-3097.
- [5] Yeganeh H, Wang Z. Objective quality assessment of tone-mapped images[J]. IEEE Transactions on Image Processing, 2013, 22(2):657-667.
- [6] Nafchi H Z, Shahkolaei A, Moghaddam R F, et al. FSITM: a feature similarity index for tone-mapped images[J]. IEEE Signal Processing Letters, 2015, 22(8):1026-1029.
- [7] Kundu D, Evans B L. Visual attention guided quality assessment of tone-mapped images using scene statistics[C]// IEEE International Conference on Image Processing. IEEE,

2016:96-100.

- [8] Xie L J, Zhang X, Wang S Q, et al. Quality assessment of tone-mapped images based on sparse representation [C]// IEEE International Symposium on Circuits and Systems. IEEE, 2016:2218-2221.
 - [9] Hadizadeh H, Bajić I V. Full-reference object quality assessment of tone-mapped images[J]. IEEE Transactions on Multimedia, 2018, 20(2):392-404.
 - [10] Gu K, Wang S Q, Zhai G T, 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.
 - [11] Yue G H, Hou C P, Gu K, et al. Biologically inspired blind quality assessment of tone-mapped images[J]. IEEE Transactions on Industrial Electronics, 2018, 65(3):2525-2536.
 - [12] Khan I R, Rahardja S, Khan M M, et al. A tone-mapping technique based on histogram using a sensitivity model of the human visual system[J]. IEEE Transactions on Industrial Electronics, 2018, 65(4):3469-3479.
 - [13] Goferman S, Zelnik-Manor L. Context-aware saliency detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(10):1915-1926.
 - [14] Lin C J. Projected gradient methods for nonnegative matrix factorization[J]. Neural Computation, 2007, 19(10):2756-2779.
 - [15] Wang Q, Chu J, Xu L, et al. A new blind image quality framework based on natural color statistic[J]. Neurocomputing, 2016, 173(3):1798-1810.
 - [16] Song Y, Jiang G Y, Yu M, et al. Naturalness index for a tone-mapped high dynamic range image [J]. Applied Optics, 2016, 55(35):10084-10091.
 - [17] 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.
 - [18] Saad M A, Bovik A C, Charrier C. Blind image quality assessment: a natural scene statistics approach in the DCT domain[J]. IEEE Transactions on Image Processing, 2012, 21(8):3339-3352.
- ~~~~~
- (上接第 113 页)
- [17] 王要峰,崔艳. 基于方向链码去除骨架图像毛刺算法[J]. 计算机应用, 2013, 33(S1):193-194, 198.
 - [18] 瞿中,蒋玉萍,文倩云. 一种方向链码扫描与跟踪的图像细化后期处理算法[J]. 计算机科学, 2015, 42(2):292-295.
 - [19] 秦筱斌,蔡超,周成平. 一种有效的骨架毛刺去除算法[J]. 华中科技大学学报(自然科学版), 2004(12):28-31.
 - [20] 孙孝坤,黄继风. 国际音标图像字符细化方法[J]. 图学学报, 2018, 39(2):214-220.