



# 超高清内容清晰度 用户体验质量评价

## Quality of Experience Estimation of Ultra-High Definition Content

朱文瀚/ZHU Wenhan, 翟广涛/ZHAI Guangtao, 陶梅霞/TAO Meixia,  
杨小康/YANG Xiaokang, 张文军/ZHANG Wenjun  
(上海交通大学, 中国 上海 200240)  
(Shanghai Jiao Tong University, Shanghai 200240, China)

**摘要:**针对多媒体行业对超高清内容清晰度用户体验评价的迫切需求,提出了一种有效的无参考质量评价算法,以预测目标内容的用户感知体验,并区分原始4K和伪4K内容。通过对目标内容进行分割,利用局部方差选择了3个代表性子块代替全局来提高计算效率。针对超高清内容的特性,提取了复杂度特征、频域特征和像素统计特征。采用支持向量回归的方法将这些提取的特征融合为一个质量指标,以预测目标内容的质量分数。实验结果表明,本模型可以有效地评估用户感知体验,并具有良好的辨别真假4K内容的能力。

**关键词:**用户体验质量;无参考质量评价;超高清;自由能原理;频域分析;自然图像统计

**Abstract:** In response to the urgent demand for assessing the quality of experience of ultra-high definition content in multimedia industries, a non-reference quality assessment model is proposed to predict the perceptual quality of the target content and distinguish pristine 4K and pseudo 4K contents. Our model segments the image and chooses three representative patches by local variances to improve computing efficiency. According to the characteristics of ultra-high definition content, complexity features, frequency domain features and pixel statistics features are extracted from the representative patches. The support vector regressor is employed to aggregate these extracted features as an overall quality metric to predict the quality score of the target image. The experimental results demonstrate that the proposed method can effectively evaluate quality of user experience and is capable of distinguishing true and pseudo 4K contents.

**Keywords:** quality of experience; non-reference quality assessment; ultra-high definition; free-energy principle; frequency domain analysis; natural scene statistics

DOI: 10.12142/ZTETJ.202101009

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20210114.0948.002.html>

网络出版日期: 2021-01-14

收稿日期: 2020-12-16

随着数字电视与多媒体行业的高速发展,超高清内容已经成为新一代电视、电脑显示器甚至手机屏幕的流行配置。由于超高清图像和视频在改善用户体验方面有着很重要的作用,在适当的观看距离下,4K视

频有生动的细节呈现,清晰度高,能显著增强视觉体验,因此,超高清内容成了时下最火热的话题之一。2012年,国际电信联盟(ITU)发布超高清电视的国际标准——ITU-R BT.2020建议书<sup>[1]</sup>。该标准正式规范了4K分辨率为3 840 × 2 160像素,宽高比为16:9。此后,各国制定了相应的超高清图像和视频标准,以更加规

范该行业<sup>[2]</sup>。例如,中国国家广播电视影视总局发布了用于生产和节目交换的超高清电视系统的参数值(GY/T 307—2017)和针对超高清电视图像质量的主观评估方法(T/CSMPTE 3—2018)。在消费市场上,各种电子设备制造商以4K为卖点,宣称其数字设备支持超高清内容。许多网络视频运营商还推出了超高

基金项目:国家自然科学基金(61831015);国家重点研发计划(2019YFB1405902)

清节目源,例如,Netflix、YouTube、乐视网、优酷和百视通都有4K视频直播服务。此外,智能手机行业将其注意力转向4K,越来越多的高端智能手机可以拍摄和生成4K图像和视频为卖点。

然而,超高清行业的发展同样会带来一些问题。根据Akamai最近的统计数据,只有21%的美国家庭网速在15 Mbit/s以上,这一传输速率被认为是有效播放4K视频的最低门槛。一些调查显示,虽然中国消费了全球约80%的4K电视,但是大部分视频信号仍是高清水平。此外,为了推广4K这一新兴卖点,一些内容提供商或个人在网络上传播大量虚假4K视频。尽管这些“高端”的“4K”视频具有与自然4K内容相同的分辨率,但其往往模糊且缺乏细节,无法满足消费者的需求。这些虚假的4K视频在存储和传输过程中占用了大量的内存和带宽资源,但却无法为用户提供相应的高质量体验。因此,如何将这些伪超高清内容从真实的超高清内容中辨识出来显得尤为重要。

图像质量评价作为一种预测图像的感知质量的方法,在过去的20年中得到了广泛研究<sup>[3]</sup>。一般而言,图像质量评价可以分为主观图像质量评价和客观图像质量评价<sup>[4]</sup>。其中,主观质量评价被认为是判断图像感知质量的最准确方法。研究者们通过建立许多主观的图像质量数据库来提供各种质量和相应的真实质量分数的图像,以促进客观模型的发展。与主观评价相比,客观评价可以自动、高效地预测失真图像的感知质量,具有可重复性高、速度快的特点,是质量评价领域的研究重点。根据参考图像的可用信息,客观的质量评价算法通常可以分为全参考、半参考和无参考算法。其中,全参考质量评

价模型可以利用参考图像的全部信息。均方误差(MSE)、峰值信噪比(PSNR)和结构相似性算法(SSIM)<sup>[5]</sup>是全参考领域的3种最经典的算法。半参考质量评价模型则只能使用一部分参考图像的信息,例如参考图像的几个特征值,但仍可以大大减少传输参考图像时的信息量<sup>[6]</sup>。此外,在大多数的现实场景中,由于参考图像并不存在,无参考图像质量评价则可以发挥出作用,这是因为它不需要参考图像就可以准确地评估失真图像的感知质量。根据方法论的不同,无参考质量评价模型大致可以分为3大类:基于自然图像统计的模型<sup>[7]</sup>、基于机器学习的模型<sup>[8]</sup>和基于人眼视觉系统的模型<sup>[9]</sup>。

目前,大多数图像质量评价方法都针对普通的低分辨率图像或人为制作的失真图像。与这类图像不同,超高清图像具有非常高的分辨率,而人眼很难区分真实的超高清图像和通过插值算法得到的伪4K图像。据我们所知,目前还没有专门针对这项任务而设计的算法。因此,预测超高清图像的质量、区分真伪超高清图像是一个全新的挑战。这值得我们去研究现有的无参考质量评价模型是否可以胜任此任务,同时值得我们去研究针对超高清图像质量的新算法。

## 1 算法设计

### 1.1 图像分解预处理

超高清图像的分辨率比一般的图像大很多,这会显著增加算法的计算量,造成算法运算时间过长,不利于算法的实际应用。因此,我们首先尝试将一个输入图像切成多个子图像,以获得最具代表性的一个或几个子图像来代表整个输入图像,然后在这些选定的子图像上执行后续的特

征提取,以减少算法的计算量。

在给定一个4K图像 $I$ 的条件下,我们首先将 $I$ 划分为 $16 \times 9$ 个子图像 $I_{i,j}$ ,其中 $i \in \{1, 2, \dots, 16\}$ ,  $j \in \{1, 2, \dots, 9\}$ 。这使得子图像 $I_{i,j}$ 的宽度像素和高度像素均为240,在随后的计算过程中具有良好的属性。由于人类的拍摄习惯和节目拍摄技巧,最重要和最具吸引力的内容往往集中在图像的中心而不是边缘。因此,为了避免代表性的子图像出现在图像的边缘,例如带有电视台徽标、电视节目名称、字幕和人们不太关注的图像内容的子图像,我们缩小了选择范围:从左侧的第三列到右侧的第三列,以及从顶部的第二行到底部的第二行。

然后,我们依据图像复杂度的特性来选择代表性的子图像。由于具有高复杂度的子图像具有更多样化的内容,这些内容可能更具吸引力并且更加重要,因此,我们采用了局部方差作为依据。局部方差是一个可以有效反映图像结构信息、对高频信息高度敏感而又比较简单的特征。对于子图像 $I_{i,j}$ 中的像素点 $(m, n)$ ,局部方差 $\sigma^2(m, n)$ 可以表示为:

$$\sigma^2(m, n) = \sum_{k=-K}^K \sum_{l=-L}^L \omega_{k,l} (I_{k,l}(m, n) - \mu(m, n))^2, \quad (1)$$

其中,  $\omega = \{\omega_{k,l} | k = -K, \dots, K, l = -L, \dots, L\}$  是一个二维的圆对称高斯加权函数, $K$ 和 $L$ 分别表示该函数窗宽和窗高的尺寸, $\mu$ 表示图像的局部均值。接下来,我们计算子图 $I_{i,j}$ 的平均局部方差,并将其作为最终的选择依据,如公式(2)所示:

$$\sigma^2(I_{i,j}) = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N \sigma^2(m, n), \quad (2)$$

其中, $M$ 和 $N$ 分别表示图像子块的宽度和高度。将所有子图的局部方差

进行排序后,我们选择了3个拥有最大局部方差的子图作为代表性子图  $I_{1,2,3}$ ,以用于后续算法的实现。

## 1.2 复杂度特征提取

在基于人类视觉系统建模的无参考图像质量评价研究中,很多学者研究自由能原理,并取得了良好的研究成果。自由能原理是在脑神经科学领域里被提出的,用于量化人脑的感知、行为和学习的过程<sup>[10]</sup>。在图像处理领域中,自由能被证明可以很好地表征图像复杂度特征,并且和图像质量高度相关<sup>[9]</sup>。因此,本文中,我们尝试使用自由能原理模型来模拟人脑预测图像的过程,并提取图像复杂度特征。

基于自由能的大脑原理的一个基本前提是,认知过程受人脑内部生成模型的控制。当人的大脑收到一个“惊喜”时,大脑会在其内部生成模型,主动预测有意义的信息并消除残留的不确定性,以生成一个预测结果,来解释大脑的感知。

我们假设用于视觉感知的内部生成模型  $G$  是参数化的,它通过调整参数向量  $\theta$  来解释视觉场景。具体来说,在给定一个输入图像  $I$  的条件下,它对大脑产生的“惊喜”,可以通过在模型参数向量  $\theta$  空间上的联合分布  $P(I, \theta|G)$  来表示:

$$-\log P(I|G) = -\log \int P(I, \theta|G) d\theta. \quad (3)$$

为了使公式(3)便于理解,我们引入了一个辅助项  $Q(\theta|I, G)$  分别加入分子和分母中,可以得到公式(4):

$$-\log P(I|G) = -\log \int Q(\theta|I, G) \frac{P(I, \theta|G)}{Q(\theta|I, G)} d\theta, \quad (4)$$

其中,  $Q(\theta|I, G)$  是这个图像的参数模型的辅助后验分布。该分布可以认

为是人脑计算的模型参数  $P(\theta|I, G)$  的真实后验的近似。大脑将通过感知图像,调整  $Q(\theta|I, G)$  的参数  $\theta$ , 来最小化近似后验  $Q(\theta|I, G)$  与真实后验  $P(\theta|I, G)$  之间的差距,用于更好地解释这个输入信号。为了简化表达,我们在后续分析中删除了生成模型符号  $G$ 。利用詹森不等式,我们可以得到公式(5):

$$-\log P(I) \leq -\int Q(\theta|I) \frac{P(I, \theta)}{Q(\theta|I)} d\theta. \quad (5)$$

然后,基于统计物理和热力学,我们定义公式(5)的右边部分为自由能:

$$F(\theta) = -\int Q(\theta|I) \frac{P(I, \theta)}{Q(\theta|I)} d\theta. \quad (6)$$

可以看出,由于  $-\log P(I) \leq F(\theta)$ , 因此,  $F(\theta)$  定义了一个图像“惊喜”的上确界。注意到  $P(I, \theta) = P(\theta|I)P(I)$ , 我们可以继续将公式(6)转化为公式(7):

$$\begin{aligned} F(\theta) &= \int Q(\theta|I) \frac{Q(\theta|I)}{P(\theta|I)P(I)} d\theta = \\ &-\log P(I) + \int Q(\theta|I) \frac{Q(\theta|I)}{P(\theta|I)} d\theta = \\ &-\log P(I) + \text{KL}(Q(\theta|I)||P(\theta|I)), \quad (7) \end{aligned}$$

其中,  $\text{KL}(\cdot)$  表示近似后验分布和真实后验分布之间的 Kullback-Leibler 散度。

本文中,我们利用稀疏表示的方法来近似人脑的内部生成模型<sup>[11]</sup>。具体来说,我们利用一个提取算子  $O_s(\cdot)$  选取输入图像的一个图像块  $x_s \in \mathbb{R}^B$ , 其中  $B$  是图像块的尺寸。那么基于一个完备字典  $Y \in \mathbb{R}^{B \times U}$ , 图像块  $x_s$  的稀疏表示是通过计算一个向量  $\alpha_s \in \mathbb{R}^U$  来表示的,如公式(8)所示:

$$\alpha_s^* = \arg \min_{\alpha_s} \frac{1}{2} \|x_s - Y\alpha_s\|_2 + \lambda \|\alpha_s\|_p, \quad (8)$$

其中,  $\alpha_s$  是提取的图像块的表示系数向量,  $U$  表示稀疏表示模型里的原子的个数。  $\lambda$  是一个正常数,用于平衡重建保真度的权重约束条件和稀疏惩罚条件。通过计算每一个图像块的表示系数向量,整个输入图像  $I$  的稀疏表示可以表示为:

$$\hat{I} = \sum_{s=1}^{n_p} O_s^T(Y\alpha_s^*) \cdot O_s^T(\mathbf{1}_B), \quad (9)$$

其中,  $\hat{I}$  表示输入图像的稀疏表示,可以被认作人脑对于输入图像的理解。“ $\cdot$ ”表示两个矩阵按元素对相除,  $n_p$  为图像块的数量,  $O_s^T(\cdot)$  表示  $O_s(\cdot)$  的转置,  $\mathbf{1}_B$  表示尺寸为  $B$  的全1矩阵。

根据自由能理论,自由能表示输入图像与人脑内部生成模型得到的最佳预测图像之间的差异。因此,我们可以将自由能定义为输出图像的预测残差的熵。预测残差可以表示为:

$$RE = |I - \hat{I}|. \quad (10)$$

而其熵可以表示为:

$$H = -\sum_{i=0}^{255} p_i \log_2 p_i, \quad (11)$$

其中,  $p_i$  表示预测残差的第  $i$  个灰度级的概率密度。最终我们可以得到输出图像的自由能的值。本文中,我们采用自由能作为图像复杂度特征。

## 1.3 频域特征提取

通常,插值方法会平滑目标图像,造成图像中低频信息增加,高频信息减少。因此,频域特征在这项任务中很有效果。本文中,我们采用离散余弦变换(DCT)获得超高清图像的频域特性。在给定一个输入图像  $I$  的情况下,我们可以获得其DCT系数

$D$ 。由于数据在直角坐标系是二维的,因此,我们将直角坐标系转换到极坐标系中,降低维度,以便于后续的处理。随后,我们计算图像在对数尺度下沿极半径方向的频谱能量 $E_\rho$ ,如公式(12)所示:

$$E_\rho = \log(1 + \sum_{\rho} |D|^2 / f^2), \quad (12)$$

其中, $\rho$ 表示极半径方向, $D$ 表示图像 $I$ 的DCT系数。 $f = f_s / N$ ,其中 $f_s$ 表示采样率。

通过大量的实验,我们发现了真伪4K图像能量谱和累积能量谱上的特征。图1给出了一对真伪4K图像标准化后的能量谱和累积能量谱的示意图。在图1(a)中,黑色曲线 $P_1$ 表示真4K图像,红色曲线 $P_2$ 表示伪4K图像,它们都是从低分辨(例如2K、1080p、720p等)的图像上采样得到的。蓝色实线 $P_3$ 是一条辅助线,经过点 $P_1$ 与 $P_2$ 的交点 $P$ 。 $p_x$ 和 $p_y$ 分别为交点 $P$ 的横坐标和纵坐标。蓝色虚线 $P_4$ 表示一个辅助图像,在整个频率上具有相同的能量,且能量高于或低于 $p_y$ 。图1(b)中, $E_i$ 为 $P_i$ 的累积能量谱( $i = 1, 2, 3, 4$ )。由于是标准化后的累积能量谱, $P_3$ 和 $P_4$ 为相同斜率的一条过原点的线段。

由频域能量谱与累积能量谱的关系可知:

$$E_i(\omega) = \int_0^\omega P_i(x) dx, \quad (13)$$

$$P_i(\omega) = \frac{\partial E_i(\omega)}{\partial \omega}. \quad (14)$$

通过大量的实验统计,我们对原始分辨率为4K的图像,以及从2K、1080p、720p 3种分辨率插值得到的伪4K图像的累积能量谱进行了拟合,发现在这4种情况下,它们的特性均近似满足: $E_i(\omega) \approx a_i \omega^{b_i}$ 。

由于 $E_i$ 的二阶导数小于零,所以它们都是凹函数。因此,我们可以发现曲线上的单点具有和 $E_4$ 相同的斜率,如黑色曲线上的点 $b$ ,红色曲线上的点 $e$ 。绿色的虚线是与蓝线平行的辅助线。点 $b$ 和点 $e$ 分别为累积能量谱曲线 $E_1$ 和 $E_2$ 与绿色虚线的交点,如图1(b)所示。这些单点在 $E_i$  ( $i = 1, 2$ )和 $E_4$ 之间的最大距离记为 $L_1$ 和 $L_2$ 。我们利用这些距离作为算法的频域特征。

实际上, $E_i$  ( $i = 1, 2, 3, 4$ )是图像标准化后的累积能量谱,它们都从原点开始,到点 $(\omega_m, 1)$ 结束,其中 $\omega_m$ 是最大频率。由于不同真假4K图像的特征各不相同,因此其频域能谱和辅助图像的频域能量谱 $P_4$ 位置关系可以分为3种情况:(1) $P_3$ 通过 $P_1$ 和 $P_2$ 的交点;(2) $P_4$ 的值大于 $p_y$ ,如图1(a)

所示;(3) $P_4$ 的值小于 $p_y$ 。下面我们将分3种情况讨论所提出特征的有效性。

(1)考虑 $P_4 = p_y$ 的情况,我们有: $P_1 < P_2, (\omega \leq p_x)$ 。同时, $E_i$  ( $i = 1, 2, 3$ )在 $\omega = p_x$ 处有着相同的斜率,这意味着 $E_1$ 和 $E_2$ 在同一水平坐标中分别拥有单个点,如图1(b)中所示的点 $b$ 和点 $a$ 。因此我们可以推导出公式(15):

$$\int_0^{p_x} P_1 d\omega < \int_0^{p_x} P_2 d\omega, (\omega \leq p_x). \quad (15)$$

在图1(b)中,公式(15)表明线段 $ad$ 的长度大于 $bd$ ,等价于 $ac > bc$ 。根据相似三角形原理,我们有 $L_1 < L_3$ 。即真4K图像的所提特征小于伪4K图像。

(2)考虑 $P_4 > p_y$ 的情况,当 $P_1 = P_2 = P_4$ 时,我们有 $\omega_1 < \omega_2$ ,其中 $\omega_1$ 和 $\omega_2$ 分别是 $P_1$ 和 $P_4$ 、 $P_2$ 和 $P_4$ 的两个交点,则有:

$$\omega_a = \omega_b = \omega_d = \omega_1 < \omega_2 = \omega_e = \omega_g. \quad (16)$$

同时,由于 $P_2 \geq P_4, \omega \in [\omega_1, \omega_2]$ ,我们可以得到: $y_f - y_e < y_e - y_a$ ,其中 $y_i$ 表示点 $i$ 的纵坐标。因此,线段 $ac$ 的长度大于 $ef$ 。根据相似三角形原则,我们有 $L_3 < L_2$ 。结合情况(1),我们可以得到 $L_1 < L_3 < L_2$ ,即真4K图像的所提特征小于伪4K图像。

(3)考虑 $P_4 < p_y$ 的情况,由于 $P_1 > P_2$ ,当 $\omega > p_x$ 时,我们有:

$$\int_{p_x}^{\omega_1} P_1(x) dx > \int_{p_x}^{\omega_2} P_2(x) dx, \quad (17)$$

可以进一步得到:

$$\int_0^{\omega_1} P_1(x) dx > \int_0^{\omega_2} P_2(x) dx. \quad (18)$$

公式(18)说明 $L_1 > L_2$ ,即真4K图像的所提特征小于伪4K图像。

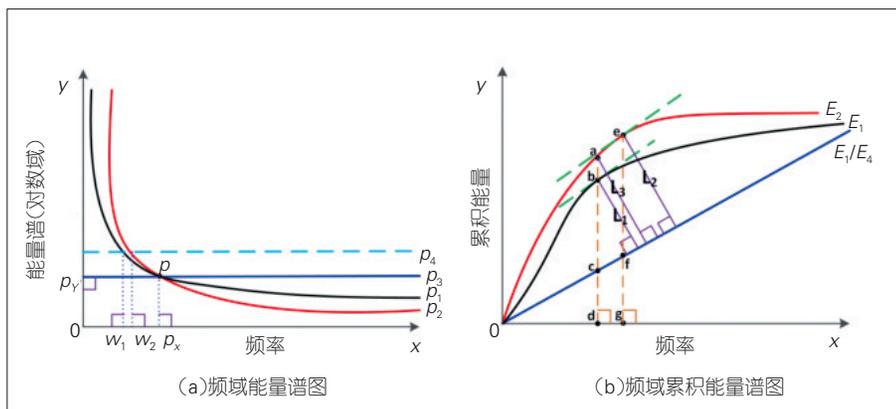


图1 一对真伪超高清图像标准化的频域能量谱和累积能量谱示意图

综上所述,我们所提取的频域成分的特征,可以有效地描述4K图像的真假,敏感于超高清图像的质量。因此,在本文中,我们将其定义为本算法的频域特征。

#### 1.4 像素统计特征提取

作为一种对图像质量很敏感的信息,自然图像统计特征在图像质量评价领域被广泛应用。因此,本算法在像素层面上,也考虑了统计信息特征来提升算法的性能。我们使用了局部的均值去除对比度归一化方法来表征超高清图像的质量变化。

具体来说,对于一张给定的输入图像 $I$ ,我们将其转化为灰度图 $Z$ ,利用公式(19)获得其均值去除对比度归一化系数:

$$\hat{Z}(x, y) = \frac{Z(x, y) - \mu(x, y)}{\sigma(x, y) - 1}, \quad (19)$$

其中, $\hat{Z}(x, y)$ 为坐标点 $(x, y)$ 处的均值去除对比度归一化系数。 $\mu(x, y)$ 和 $\sigma(x, y)$ 分别表示在坐标点 $(x, y)$ 处的局部均值和局部标准差,可以通过公式(20)和公式(21)来计算:

$$\mu(x, y) = \frac{1}{K} \sum_{k=-K}^K \sum_{l=-L}^L \omega_{k,l} Z(x+k, y+l), \quad (20)$$

$$\sigma(x, y) = \sqrt{\sum_{k=-K}^K \sum_{l=-L}^L \omega_{k,l} (Z(x+k, y+l) - \mu(x, y))^2}, \quad (21)$$

其中,与1.1节相似, $\omega_{k,l}$ 是一个二维的圆对称高斯加权窗函数。在以前的研究中我们观察到,对于自然图像,均值去除对比度归一化系数与单位正态高斯特征具有很强的相关性<sup>[12]</sup>。我们尝试利用该系数的属性,来判断真实4K图像和从不同原始分辨率插值以及不同插值算法得到的伪4K图像。为了证明在此任务中使用均值去除对比度归一化系数的效果,我们选择了一个真实4K图像,如图2(a)所示。接着,我们分别计算了该图像与其对应的具有不同的原始分辨率和不同插值算法的伪4K版本的均值去除对比度归一化系数值,如图2(b)所示。

可以看出,真实4K图像及其不同4K版本的分布是可区分的。真实4K图像的分布显示出类似高斯的外观,而其他不同的伪4K版本则以自己的方式偏离了这种特性。这表明该系数的属性在区分真实4K图像和伪4K图像中起着积极的作用。为了数学化描述均值去除对比度归一化系数的分布,我们采用广义高斯分布来有效描述真实4K图像和伪4K图

像统计数据谱。零均值的广义高斯分布可定义为:

$$G(x; \alpha, \sigma^2) = \frac{\alpha}{2\beta(1/\alpha)} \exp(-(\frac{|x|}{\beta})^\alpha), \quad (22)$$

其中, $\beta = \sigma \sqrt{\Gamma(1/\alpha)/\Gamma(3/\alpha)}$ 。伽马方程可以表示为公式(23):

$$\Gamma(\varphi) = \int_0^\infty \varphi^{\varphi-1} e^{-\varphi} d\varphi, \quad \varphi > 0, \quad (23)$$

其中, $\alpha$ 和 $\sigma$ 是两个参数,分别改变此高斯分布的幅度和方差。我们采用该广义高斯分布模型,近似超高清图像的均值去除对比度归一化系数分布。 $\alpha$ 和 $\sigma$ 这两个参数被提取作为本算法像素统计特征。由于多尺度处理有助于改善质量评价模型的预测分数与人类感知之间的相关性,我们从两个尺度上提取特征,包括原始比例和降低两倍分辨率的图像。

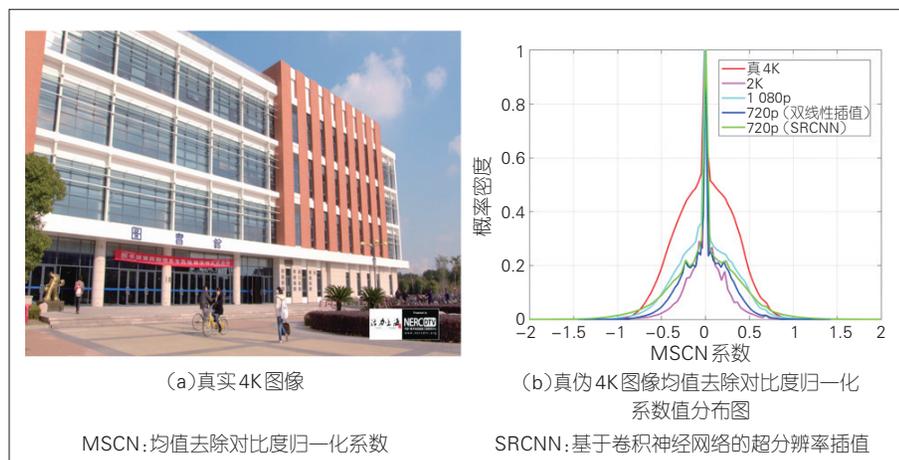
#### 1.5 特征融合和模型表示

为了聚合上述提取的与超高清图像质量相关的特征,并生成质量评价模型以预测目标图像的质量分数,在综合考虑了回归器的有效性和模型的计算速度后,我们利用支持向量回归(SVR)方法聚合提出的特征,并采用LIBSVM软件包来学习有径向基函数(RBF)内核的模型<sup>[13]</sup>。

## 2 实验过程和分析

### 2.1 实验数据构成

为了测试算法的有效性,我们首先构建了真伪超高清图像的数据库,并从几个现有的超高清视频序列库中获得了50多段视频序列。然后,我们从这些具有不同图像内容的视频序列中提取总共350张真实4K图像,得到了真实4K内容数据集。这些素材内容非常广泛,包括室外场景、室内场景、建筑物、角色、动物、静



▲图2 均值去除对比度归一化系数的效果测试

物、夜景、运动场景、电影和电视剧片段。接着,我们将真实的4K图像下采样为具有2K、1080p和720p 3种分辨率的图像。接着,我们通过14种不同的插值方法将它们都上采样到4K分辨率。总共有2802个伪4K图像构成了伪4K内容数据集。

## 2.2 实验方案

基于上述所构建的数据库,我们将提出的算法与现有的客观质量评价模型进行了比较。由于本项任务没有用于确定4K图像真实性的参考图像,因此我们仅选择无参考质量评价算法来衡量性能。我们采用了11种最新的无参考质量评价模型与我们提出的模型进行比较以预测准确性。这11种模型分别是NIQE<sup>[7]</sup>、QAC<sup>[14]</sup>、IL-NIQE<sup>[15]</sup>、LPSI<sup>[16]</sup>、HOSA<sup>[8]</sup>、BRISQUE<sup>[17]</sup>、BPRI<sup>[18]</sup>、BMPRI<sup>[19]</sup>、NFEDM<sup>[9]</sup>、CPBD<sup>[20]</sup>和GMLF<sup>[21]</sup>(NIQE、QAC……GMLF是指不同的无参考质量评价模型)。

根据质量评价领域的传统评估方法,我们使用4个通用评估标准来衡量所有比较的无参考质量评价模型的性能,它们分别是斯皮尔曼等级相关系数(SRCC)、肯德尔等级相关系数(KRCC)、皮尔逊线性相关系数(PLCC)和均方根误差(RMSE)。此外,我们还计算了3个准确性指标:精确率(Precision)、召回率(Recall)和准确率(Accuracy),以比较算法的性能和判断4K图像的真实性。

为了对所提出的模型进行训练,我们将测试材料随机分为两组:训练集和测试集,它们分别包含80%和20%的图像。我们使用训练集训练提出的模型,并使用测试集测试其性能。为了保证模型的鲁棒性,我们将此过程重复了1000次。这1000次重复的中值结果被认为是最终性能。

## 2.3 实验结果和分析

表1给出了所有算法的性能结果。其中,Precision\_T和Precision\_F分别表示真4K图像和伪4K图像素材组的精确率,而Recall\_T和Recall\_F分别表示真4K图像和伪4K图像素材组的召回率。由表1可知,在传统指标中,与传统图像质量评价数据库中的性能结果相比,所有算法的性能均不算出色。例如,这些指标中SRCC和PLCC值均不超过0.9,而通常这些指标在传统的图像质量评价数据库上会超过0.9。造成这种现象的主要原因是真实的4K图像与其对应的伪4K图像之间的差距很小,肉眼难以分辨。对于传统的人为失真来说,这项任务中的差异微乎其微,甚至很多伪4K图像的质量都要优于传统质量评价数据库里的参考图像。从结果上看,我们算法的性能明显优于其他主流的无参考质量评价模型。我们提出的方法的SRCC值超过0.8,

PLCC值接近0.85,而其他算法的SRCC值大都低于0.7,PLCC值低于0.8。

通过分析分类算法中常用的指标精确率、召回率和准确率的结果,我们还可以得出这样的结论:每个模型都具有较强的判断能力,而伪4K图像的判断准确度要优于真4K图像。此外,我们提出的算法具有最佳的性能,综合判断精度超过97%。因此,我们的算法具有优秀的区分真实和伪4K图像的能力,并且这种能力与主观感知分数呈正相关关系。

## 3 结束语

本文中,我们设计了一种新的无参考质量评价模型来评价超高清内容清晰度的用户体验质量。基于超高清内容的特性,我们在目标内容上分别提取复杂度特征、频率特征和像素统计特征,采用具有最高局部方差的3个子图代替完整的目标图像以改

▼表1 提出算法和对比算法的性能结果

指标	BPRI <sup>[18]</sup>	BMPRI <sup>[19]</sup>	BRISQUE <sup>[17]</sup>	CPBD <sup>[20]</sup>	NFERM <sup>[9]</sup>	GMLF <sup>[21]</sup>
SRCC	0.3506	0.3594	0.6651	0.5963	0.6708	0.2387
PLCC	0.5614	0.6534	0.6696	0.6194	0.6662	0.2376
KRCC	0.2956	0.2308	0.5061	0.4315	0.4990	0.1594
RMSE	13.5928	12.4345	12.2003	12.8950	12.2497	14.2550
Precision_T	0.9097	0.5890	0.8795	0.5522	0.9653	0.3433
Precision_F	0.9727	0.9441	0.9477	0.9349	0.9299	0.9196
Recall_T	0.7771	0.5486	0.5629	0.4686	0.3971	0.3600
Recall_F	0.9904	0.9522	0.9904	0.9525	0.9982	0.9140
准确率	0.9667	0.9074	0.9429	0.8988	0.9315	0.8525

指标	HOSA <sup>[8]</sup>	NIQE <sup>[7]</sup>	IL-NIQE <sup>[15]</sup>	LPSI <sup>[16]</sup>	QAC <sup>[14]</sup>	(pro.)
SRCC	0.7153	0.5223	0.3819	0.5782	0.6866	0.8136
PLCC	0.7296	0.5691	0.3437	0.7629	0.6427	0.8447
KRCC	0.5299	0.3797	0.2593	0.5051	0.5204	0.6472
RMSE	11.4445	13.5061	15.4249	10.6193	12.5836	7.9403
Precision_T	0.6613	0.7550	0.1469	0.7188	0.4868	0.9748
Precision_F	0.9496	0.9442	0.9081	0.9942	0.8920	0.9807
Recall_T	0.5914	0.5371	0.4600	0.7886	0.2114	0.9138
Recall_F	0.9622	0.9782	0.6663	0.9936	0.9722	0.9914
准确率	0.9210	0.9293	0.6434	0.9708	0.8877	0.9781

KRCC:肯德尔等级相关系数  
PLCC:皮尔逊线性相关系数

RMSE:均方根误差  
SRCC:斯皮尔曼等级相关系数

(pro.):本文提出的算法

善计算效率。支持向量回归的方法被用于回归这些特征到一个整体质量指标上。实验表明,在预测超高清内容清晰度的用户体验质量方面,本方法优于其他最新的无参考质量评价模型,并且具有良好的区分原始和伪超高清图像的能力。本算法的研究将会对超高清内容清晰度用户体验评估领域的发展起到积极的促进作用。

#### 参考文献

- [1] ITU. Parameter values for ultra-high definition television systems for production and international programme exchange: ITU-R BT.2020 [S]. 2012
- [2] SUGAWARA M, CHOI S Y, WOOD D. Ultra-high-definition television (rec. ITU-R BT.2020): a generational leap in the evolution of television standards in a nutshell [J]. IEEE signal processing magazine, 2014, 31(3): 170-174. DOI: 10.1109/msp.2014.2302331
- [3] ZHAI G T, MIN X K. Perceptual image quality assessment: a survey [J]. Science China information sciences, 2020, 63(11): 211301. DOI: 10.1007/s11432-019-2757-1
- [4] ZHU W H, ZHAI G T, MIN X K, et al. Multi-channel decomposition in tandem with free-energy principle for reduced-reference image quality assessment [J]. IEEE transactions on multimedia, 2019, 21(9): 2334-2346. DOI: 10.1109/tmm.2019.2902484
- [5] 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. DOI: 10.1109/tip.2003.819861
- [6] SOUNDARARAJAN R, BOVIK A C. RRED indices: reduced reference entropic differencing for image quality assessment [J]. IEEE transactions on image processing, 2012, 21(2): 517-526. DOI: 10.1109/tip.2011.2166082
- [7] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a "completely blind" image quality analyzer [J]. IEEE signal processing letters, 2013, 20(3): 209-212. DOI: 10.1109/lsp.2012.2227726
- [8] XU J T, YE P, LI Q H, et al. Blind image quality assessment based on high order statistics aggregation [J]. IEEE transactions on image processing, 2016, 25(9): 4444-4457. DOI: 10.1109/tip.2016.2585880
- [9] ZHAI G, WU X, YANG X, et al. A psychovisual quality metric in free-energy principle [J]. IEEE transactions on image processing, 2012, 21(1): 41-52. DOI: 10.1109/tip.2011.2161092
- [10] KARL F. The free-energy principle: a unified brain theory? [J]. Nature reviews neuroscience, 2010, 11(2): 127-138. DOI: 10.1038/nrn2787
- [11] LIU Y T, ZHAI G T, GU K, et al. Reduced-reference image quality assessment in free-energy principle and sparse representation [J]. IEEE transactions on multimedia, 2018, 20(2): 379-391. DOI: 10.1109/tmm.2017.2729020
- [12] RUDERMAN D L. The statistics of natural images [J]. Network: computation in neural systems, 1994, 5(4): 517-548. DOI: 10.1088/0954-898X\_5\_4\_006
- [13] SCHÖLKOPF B, SMOLA A J, WILLIAMSON R C, et al. New support vector algorithms [J]. Neural computation, 2000, 12(5): 1207-1245. DOI: 10.1162/089976600300015565
- [14] XUE W F, ZHANG L, MOU X Q. Learning without human scores for blind image quality assessment[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA: IEEE, 2013: 995-1002. DOI: 10.1109/cvpr.2013.133
- [15] LIN ZHANG, LEI ZHANG, BOVIK A C. A feature-enriched completely blind image quality evaluator [J]. IEEE transactions on image processing, 2015, 24(8): 2579-2591. DOI: 10.1109/tip.2015.2426416
- [16] WU Q B, WANG Z, LI H L. A highly efficient method for blind image quality assessment [C]//2015 IEEE International Conference on Image Processing (ICIP). Quebec City, QC, Canada: IEEE, 2015: 339-343. DOI: 10.1109/icip.2015.7350816
- [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. DOI: 10.1109/tip.2012.2214050
- [18] MIN X K, GU K, ZHAI G T, et al. Blind quality assessment based on pseudo-reference image [J]. IEEE transactions on multimedia, 2018, 20(8): 2049-2062. DOI: 10.1109/tmm.2017.2788206
- [19] MIN X K, ZHAI G T, GU K, et al. Blind image quality estimation via distortion aggravation [J]. IEEE transactions on broadcasting, 2018, 64(2): 508-517. DOI: 10.1109/tbc.2018.2816783
- [20] NARVEKAR N D, KARAM L J. A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection [C]//2009 International Workshop on Quality of Multimedia Experience. San Diego, CA, USA: IEEE, 2009: 87-91. DOI: 10.1109/qomex.2009.5246972
- [21] XUE W, MOU X, ZHANG L, et al. Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features [J]. IEEE transactions on image processing, 2014, 23(11): 4850-4862. DOI: 10.1109/tip.2014.2355716

#### 作者简介



朱文瀚, 上海交通大学电子信息与电气工程学院、人工智能研究院、教育部人工智能重点实验室在读博士研究生; 研究领域包含图像感知质量评价、图像视频信号处理; 发表国际学术论文17篇。



翟广涛, 上海交通大学电子信息与电气工程学院院长助理、教授、博士生导师,《Displays》主编,《中国科学: 信息科学》编委, IEEE电路与系统分会视觉信号处理与通信技术委员会(CAS VSPC)成员、多媒体系统及应用技术委员会(MSA)成员, 中国电子学会青年科学家俱乐部副主席, 上海市图象图形学会副理事长; 研究方向为多媒体信号处理等; 发表国际期刊论文100余篇。



陶梅霞, 上海交通大学电子信息与电气工程学院教授、博士生导师, IEEE Fellow, 中国电子学会信息论分会副主任委员, 曾任《IEEE Transactions on Wireless Communications》《IEEE Transactions on Communications》

《IEEE Journal of Selected Areas in Communications》等期刊的编委或客座编委; 获2019年IEEE通信学会马可尼论文奖、2013年IEEE通信学会海因里希兹论文奖; 主要从事无线通信与网络基础研究, 包括无线缓存、边缘计算及5G关键技术等; 发表国际期刊论文80余篇、国际会议论文100余篇。



杨小康, 上海交通大学人工智能研究院常务副院长、人工智能教育部重点实验室主任、教育部“长江学者”特聘教授、国家杰出青年科学基金获得者、国家“万人计划”创新领军人才、IEEE Fellow,《IEEE Transactions on Multimedia》《IEEE Signal Processing Letters》编委; 研究领域为图像处理与机器学习; 主持国家重点研发专项、“973”项目、国家自然科学基金项目等10余项, 获国家科技进步二等奖、中国电子学会自然科学一等奖、上海市科技进步一等奖等多个奖项; 发表国际学术论文200余篇, 申请发明专利50余项。



张文军, 上海交通大学教授、教育部“长江学者”特聘教授、国家杰出青年科学基金获得者、“973”项目首席科学家、国家自然科学基金委创新群体学术带头人、IEEE Fellow, 曾任国家高清晰度电视功能样机系统研发项目总体组组长、数字电视国家工程研究中心首席科学家、教育部未来媒体网络协同创新中心主任, 国际未来广播电视合作研究计划技术委员会主席; 主要从事图像通信与数字电视、宽带无线传输、系统芯片设计等工作, 获国家科技进步二等奖(2项)、何梁何利基金科学与技术进步奖、上海市科技进步一等奖(4项)、上海市科技功臣奖。