使用高斯过程回归模型对模糊和块效应图像进行频域盲质量评估

Maryam Viqar[1], Athar A. Moinuddin[1], Ekram Khan[1] 和 M. Ghanbari[2]

[1] 阿里格尔穆斯林大学阿里格尔 Z. H. 工程与技术学院电子工程系
[2]英国科尔切斯特埃塞克斯大学电子与计算学院

数字对象标识符： https://doi.org/10.1016/j.image.2022.116651

摘要

大多数标准图像和视频编解码器均基于块进行编码，压缩图像/视频会因压缩比不同而产生不同的失真。在较低压缩比下，图像会出现模糊；随着压缩程度提高，则会出现块效应。通常，为减轻块效应，会对图像进行低通滤波，但这会导致模糊加重。此外，散景模式图像中也常见此类现象：有意模糊背景会造成模糊，而压缩则会引起块效应和整体模糊。因此，这类视觉媒体会同时受到块效应和模糊失真的影响。此外，噪声也是一种常见失真。现有的大多数质量评估研究都是分别量化这些失真。本文提出一种方法，可对存在上述失真的图像进行无参考整体质量测量，既能分别测量，也能联合测量。该方法通过计算离散频率变换（DFT）低频和高频系数绝对值之和（定义为幅值和）来实现。将幅值和处于特定范围内的块数（包括交流系数为零的块数），以及这些幅值和中最大100个值和最小100个值的均值用作特征向量。随后，将这些特征输入基于机器学习（ML）的高斯过程回归（GPR）模型，以量化图像质量。仿真结果表明，所提方法能够评估受到块效应、模糊、噪声及其组合失真的图像质量。与许多先进方法相比，该方法速度较快，因此适用于实时质量监测应用。

关键词：图像质　　量评估、多重失真、块效应、模糊、离散傅里叶变换

1. 引言

许多受生物系统启发的技术被应用于人工智能系统开发、数据处理、计算方法、优化算法等任务中。由于人类是众多应用的最终用户，因此很多时候需要采用能够高度复现人类生物系统的技术。其中一个应用领域是图像质量评估（IQA），其方法旨在根据人类视觉系统（HVS）感知视觉媒体的质量。这类质量评估方法复现了视觉皮层中神经元的主要功能，例如对视觉信号进行频率分解和除法归一化变换 [1]。

1.1 动机：

如今，图像和视频已成为通过 Instagram、Facebook、YouTube 等社交媒体平台分享信息的常用媒介。此外，全球移动/互联网数据流量呈指数级增长，据估计，到2022年，视觉内容将占互联网总流量的约80% [2]。尽管图像和视频是现代数字时代信息分享的主要来源，但最常见的图像和视频在采集、传输、压缩、处理等不同阶段中，可能会受到多重失真的影响。这些操作会在采集和传输过程中引入噪声，在 JPEG/MPEG 压缩过程中产生块效应和模糊，在滤波或处理过程中造成模糊，等等。这些失真的共同作用

© <2021>。此稿件版本依据 CC-BY-NC-ND 4.0 许可提供

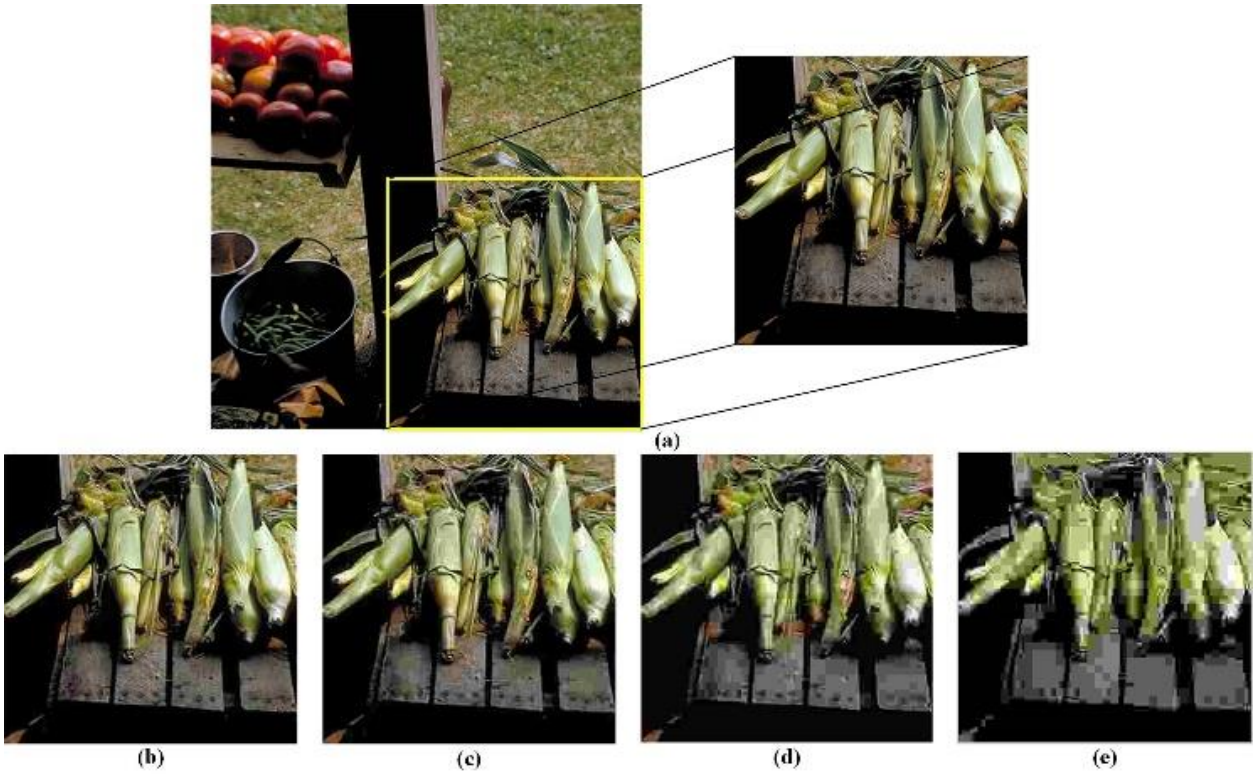

图1 （a）CSIQ 数据库中的原始图像 *[5],* 不同压缩比（CR）下的 JPEG压缩图像：-（b）CR=*20*，（c）CR=*40*，（d）CR =*65* 和（e）CR=*120*

对图像质量造成的影响，在感知上不同于图像中单一失真所造成的影响。因此，开发针对多重失真的图像质量评价方法至关重要，且这一领域近来已受到广泛关注。

图像质量可以通过主观或客观方式进行评估。尽管主观质量评价具有很高的一致性，但它需要大量人类观察者，因此是一个昂贵且耗时的过程，不适用于实时应用。相比之下，客观图像质量评估技术无需任何人工干预，且速度更快。客观质量评价技术可分为三类：全参考（FR），其需要原始图像来估计目标图像的质量；减参考（RR），其不需要完整的参考图像，而是使用其部分特征；以及无参考（NR）方法，即在没有参考（或原始）图像的情况下估计质量。对于大多数实时应用，无参考（NR，也称为盲）图像质量评估技术是最合适的方法。

为应对流量前所未有的增长并促进视觉信息的快速交换，采用高效的图像/视频压缩技术已不可避免。大多数高效压缩技术都属于有损压缩，并会引入一定程度的失真，如图 1 所示。压缩程度越高，图像/视频越可能出现严重失真，因此需要妥善权衡压缩效率与图像/视频质量。为此，需要一种快速、准确的盲质量评估技术，以便实时权衡质量与压缩程度。在编码图像/视频中，最常见的两种失真是块效应和模糊。在基于块的 JPEG（联合图像专家组）和 H.26x/MPEG（动态图像专家组）编码图像/视频中，当相邻块的量化直流系数存在显著差异时，就会出现块效应。另一方面，当图像以低压缩比编码导致高频分量被移除，或采用低通滤波（或标准视频编码器中的环路滤波）等处理时，编码图像会显得模糊

© <2021>。本稿版本依据 CC-BY-NC-ND 4.0 许可发布

以去除块效应（尤其是在视频编码中）。这两种与压缩相关的失真，即块效应和模糊，在图像/视频中往往相伴出现。据观察，在极低压缩率下，最初会出现模糊；随着压缩率进一步提高，块效应便会显现 [3]。一些研究者利用块效应与模糊之间的这种关系，设计了更高效的图像/视频质量评价（IQA）工具。例如，如果可以获取参考图像或其部分信息，那么通过测量块边界处的增益和损失，就能设计出一种高效的 IQA 方法 [4]。为验证这一观察结果，我们将 CSIQ 数据库 [5] 中的一幅图像以不同压缩比（CR）进行压缩，所得解码图像如图 1 所示。为便于观察，我们裁剪并放大了黄色方框中的图像区域。图 1(a) 中的无失真图像以 CR= 20、40、65 和 120 进行压缩，结果分别如图 1(b)–(e) 所示。首先，在较低压缩比（CR=20）下，与原图相比，图 1(b) 中的图像开始变得不够清晰，玉米表皮的细线也逐渐消失。随着 CR 增至 65，模糊进一步加重。当压缩比变得很高时（CR=120），块效应便十分明显，如图 1(e) 所示。从这个例子可以直观地看出，基于块的压缩技术会同时引入模糊和块效应。因此，任何压缩图像/视频都包含这两种失真的混合。此外，环路滤波也会导致这两种失真同时出现，尤其是在视频编码器环路中。此外，加性白高斯噪声（AWGN）也是一种常见失真，它通常会引入高频随机分量，从而影响图像的感知质量。本研究旨在开发一种图像质量评价方法，以准确评估受到块效应、模糊和/或噪声（单独或共同）失真的图像质量。

1.2 相关工作：

现有的盲评或无参考图像质量评价方法可分为两大类，即特定失真方法和通用方法。有时，失真的性质可能是已知的，例如量化、相机运动造成的失真、相机内部固有的噪声源、低通滤波等。然而，在许多情况下，失真的来源和性质并不明确。顾名思义，特定失真方法旨在衡量单一特定类型失真下的图像质量。当图像受到一种已知类型的失真时，可以采用这类方法。相比之下，通用方法旨在评估不同类型的单一失真。

在特定失真的图像质量评价方法中，我们在此考虑专门针对模糊、块效应或噪声设计的方法。可通过测量边缘扩散 [6]- [8], 、测量恰可察觉模糊（JNB）[9]- [11], , 或计算离散余弦变换（DCT）块上的峭度 [12] 来评估模糊图像的质量。在 [13], 中，作者使用每个像素的最大局部变化（MLV）来估计图像的清晰度。MLV 被赋予权重，并以其标准差衡量清晰度程度，再据此量化模糊程度。Zhang 等人在 [14] 中利用结构相似性指数（SSIM）的概念，通过对图像重新模糊来检测其中的模糊。这是一个两步框架：首先使用高斯低通滤波器（GLPF）对图像进行滤波，然后分析直方图中的相应变化，以确定模糊程度。在 [15], 中，作者使用精确泽尼克矩与梯度幅值（EZMGM）建立模糊度量。H. Liu 等人在 [16] 中提出了一种利用边缘模糊和神经网络的模糊度量，该方法对于背景经过有意模糊处理的图像效果良好。

同样，已有大量技术被开发用于测量块效应导致的失真 [17]-[24]。这些技术大多利用频域特性，主要是离散傅里叶变换 (DFT)[17], [21], 和离散余弦变换 (DCT) [18], [20]。DFT 和 DCT 的组合也被用于评估 JPEG压缩图像的质量 [22], , 其中作者使用 DCT 变换块中的零值系数数量，并以 DFT 生成的质量图加权，将其作为特征来评估块效应的程度。尽管这一概念简单有效，但在大多数标准数据库上的准确度较差。[23], 提出了一种基于梯度的技术，该技术针对每个 8×8 像素的非重叠块，确定两个参数，即亮度变化（相邻块之间）和块效应程度（通过将水平和垂直导数相加进行评估），

© <2021>。本文稿版本依据 CC-BY-NC-ND 4.0 许可发布

以测量图像质量。总体而言，该方法准确度较高，但如果块被平坦区域包围，则无法准确估计质量。Zhu 等人 [24] 提出了一种基于质量因子 (QF) 和对已压缩 JPEG 图像进行重新压缩的技术。[25]- [28]提出的方法是为含噪图像的质量评估而开发的。

尽管上述大多数方法在处理单一失真方面表现良好，但对于受到多种失真影响的图像，它们的效率会大幅下降。如今，图像很可能同时受到多种失真的影响，因此需要多种通用图像质量评价方法，以便在存在多种已知或未知失真的情况下，准确评估图像质量。接下来将回顾一些相关的通用方法。

在 [3] 中提出了一种基于频域方法的全参考方法，能够评估同时受到模糊和块效应失真的图像质量。现有许多通用的无参考图像质量评估方法，例如 [29] -[39]。其中大多数方法基于图像的自然场景统计（NSS）[29] -[34], [37], [39], 等，其依据是：当图像受到失真影响时，其统计参数会发生变化，而这些变化可用于估计图像中的失真程度。NSS 特征既可以在空间域或变换域中提取，也可以结合二者进行提取。Moorthy 和 Bovik 提出了一种名为 DIIVINE（基于失真识别的图像真实性与完整性评估）[31] 的技术，用于分两个步骤评估图像质量。第一步识别失真类型，第二步量化失真程度。该方法利用小波变换获取统计特征，并使用高斯尺度混合模型对这些特征进行建模。该方法采用了较大的特征规模（共 88 个特征），显然使其过于复杂。BLIINDS2（使用 DCT 统计的无参考图像完整性评估器）[32] 是一种使用较少基于 DCT 的 NSS 特征的方法。在该方法中，利用广义高斯密度（GGD）模型对每个块的 DCT 系数进行建模，并将 GGD 模型的参数输入贝叶斯回归模型，以量化图像质量。尽管 BLINDS2 [32] 方法相比 DIIVINE [31] 方法具有更高的效率，但由于 BLIINDS2 [32] 中采用了基于块特征的非线性排序，导致其处理时间较长，从而运行缓慢，并限制了其在许多应用中的使用 [29]。BRISQUE（无参考/无基准图像空间质量评估器）[29] 是一种基于空间域 NSS 的方法，它利用图像的均值减除对比度归一化（MSCN）系数构建质量评估指标。尽管 BRISQUE 是运行速度最快的图像质量评价方法之一（就计算时间而言），但对于 TID2013 数据库中的某些特定类型图像，其性能较差 [40]。Xue 等人 [33] 联合利用梯度幅值（GM）和高斯拉普拉斯（LoG），分别研究局部亮度变化和强度变化对图像质量的影响，因为二者与人类视觉系统（HVS）密切相关。该方法能够准确测量受到块效应影响的图像质量，但对于模糊图像的准确性相对较低。FRIQUEE（基于特征图的无参考图像质量评估引擎）方法 [37] 综合利用了空间域和变换域特征。其涉及不同颜色空间中的大量特征，因此复杂且耗时。Li 等人在 [39] 中提出了一种基于对数收缩分布的无参考方法，用于利用图像中的空间相关性和方向相关性。该方法存在数据依赖问题，其跨数据库验证的准确性并不令人满意。上述方法均基于机器学习，而 NIQE [41] 和 IL-NIQE [42] 等无需学习的通用方法也已被提出。M H Pinson [43] 的工作从消费类应用的角度引起了人们对视觉媒体评估的关注。BRISQUE [29], 、NIQE [41], 、IL-NIQE [42], 等方法已针对多重失真进行了测试，但发现它们对于这些失真以及未见过的数据同样表现不佳。此外，还提出了一些多重失真方法 [44]-[47]，用于评估存在两种或更多混合类型失真时的图像质量。Li 等人提出了 G WH- GLBP [44] ，用于评估图像中的多重失真。该方法基于结构信息，利用测试图像梯度图中的局部二值模式（LBP）构建加权直方图。Miao 等人 [45] 提出了另一种基于相位一致性域中 LBP 的方法，以考虑低级特征。该方法还采用梯度幅值作为加权因子。六步无参考度量（SISBLIM）[46] 是一种基于联合效应预测和基于 HVS 的融合模型的无需训练度量。VQEG（视频质量专家组）[48] 已针对项目开展了大量视觉媒体质量评估工作，例如

© <2021>。本文稿版本依据 CC-BY-NC-ND 4.0 许可公开

MOAVI（基于关键指标的视听质量监测）[49], [50] 和 NORM（无参考指标）。MOAVI 项目提出了一组用于定义视听信号服务质量的关键指标。NORM 项目致力于无参考指标的研究、开发和资源共享。

大多数用于评估存在多重失真时图像质量的通用图像质量评估方法，要么时间复杂度较高，要么准确率较低。对于多重失真图像，由于各种失真的联合影响，许多最先进的方法在测试同时受到多种失真影响的图像时表现不佳。此外，多重失真图像质量评估领域可能因缺少基准数据库且问题本身颇具挑战性而未受到足够关注。因此，我们需要一种能够以更高准确率和更低复杂度评估多重失真图像质量的图像质量评价方法。

为克服在多重失真条件下评估图像质量的图像质量评价方法存在的不足，我们提出一种基于离散傅里叶变换的技术，用于估计单独或两种及以上失真组合所造成的块效应、模糊和噪声对图像质量的影响。尽管离散傅里叶变换和 MSCN 系数各自已广泛应用于许多图像质量评价方法中，但将二者优势结合起来的研究却不多。我们认为，所提方法的创新之处在于利用 MSCN 系数的离散傅里叶变换来构建一种图像质量评价方法。现有的大多数 MSCN 方法使用单变量概率模型拟合分布，难以充分捕捉相关性。另一个局限是只考虑相邻像素，而失真通常会同时破坏多个像素对，因此这种做法不适用于复杂类型的失真。为克服这些局限，我们提出一种基于 MSCN 和离散傅里叶变换的独特方法。具体而言，本研究的创新之处在于：（i）结合 MSCN 与离散傅里叶变换，从而获得高度去相关的低维空间；（ii）考虑失真对适当大小系数块中多个像素（而非仅相邻像素）的影响。所提技术首先计算输入图像的 MSCN 系数。随后，利用8×8 分块离散傅里叶变换，将原始图像及其 MSCN 系数转换至频域。再将每个块的离散傅里叶变换系数划分为低、中、高频带。将高频带和低频带系数相加，得到求和参数；这些参数由原始图像和 MSCN 系数的离散傅里叶变换获得，并用于提取特征。对于归一化后的求和参数，统计落在特定范围内的块数（包括零值系数），并计算这些求和参数中 100 个最大值和 100 个最小值的均值，以构成特征向量。然后将所有这些特征输入采用指数核的高斯过程回归模型[51]，以估计图像质量。与为单独测量这些失真而开发的方法相比，所提方法的准确率显著更高。

本文的组织结构如下：第 2 节简要回顾了所提方法中使用的 DFT 和 MSCN。第 3 节详细介绍所提方法。第 4 节给出仿真结果并讨论这些结果。最后，第 5 节对本文进行总结。

2. 背景

由于所提图像质量评价方法利用图像的 DFT 域特征，并结合输入图像的 MSCN 系数的 DFT，因此本节将简要回顾 DFT 和 MSCN 系数的计算及其特性。

2.1 离散傅里叶变换（DFT）：

众所周知，图像中的失真会影响 其频率分布特性。因此，图像高、低频分量强度的相对变化可用于判断图像是否存在块效应、噪声和模糊。这些变化可通过 DFT 轻松捕捉，因此 DFT 显然是本研究的理想选择。图 2 展示了从 LIVE [52] 数据库中选取的八幅图像，其中图 2(a) 和

© <2021>。本手稿版本依据 CC-BY-NC-ND 4.0 许可公开。

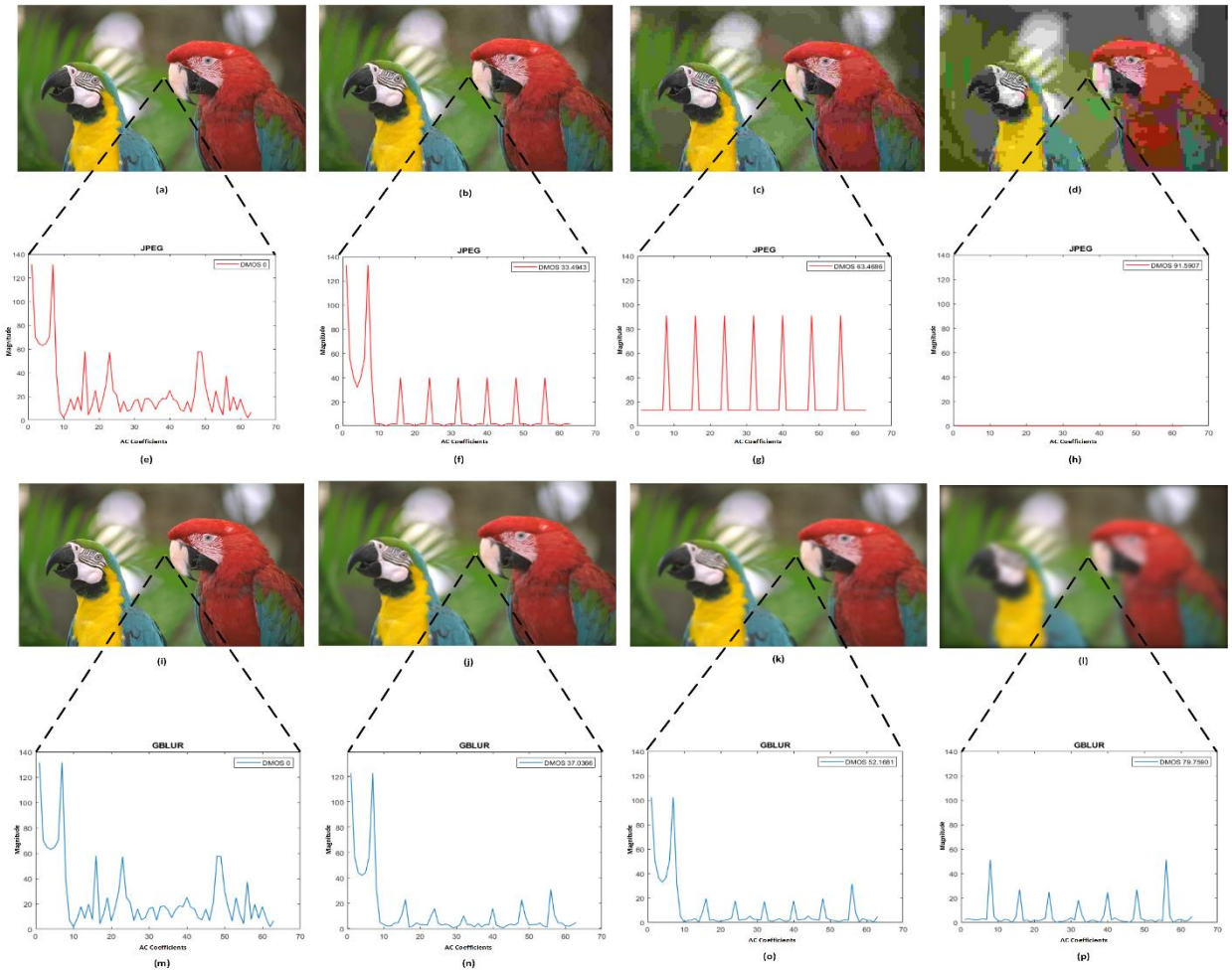

图 2 来自 LIVE *[52]* 数据库的 JPEG 图像（a）–（d）和高斯模糊（GBLUR）类型图像（i）–（l），以及相应的离散傅里叶变换系数图（e）–（h）和（m）–（p）。

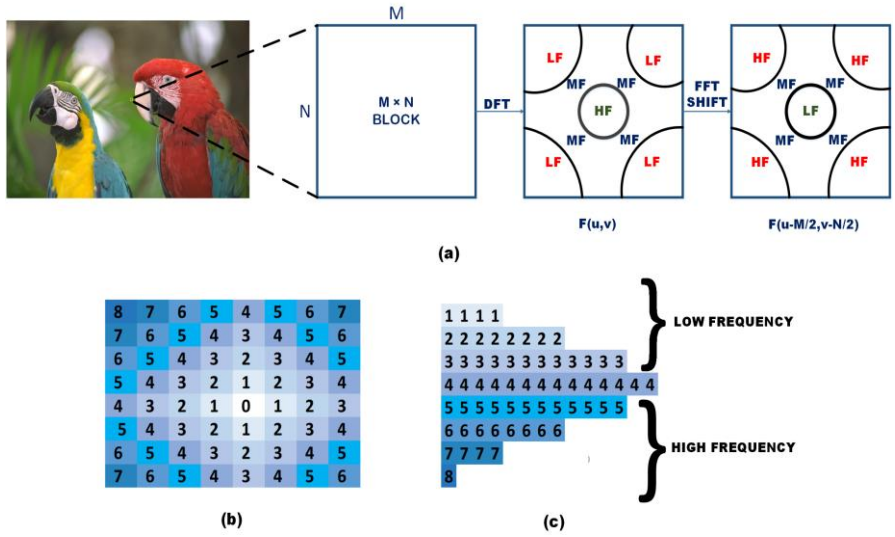

图 3（a）经频移后的离散傅里叶变换系数块级表示；（b）使用曼哈顿距离按频率递增顺序索引的离散傅里叶变换分量块级表示；（c）使用曼哈顿距离将频带划分为索引行，从而分离高频带和低频带

© <2021>。此手稿版本依据 CC-BY-NC-ND 4.0 许可公开提供

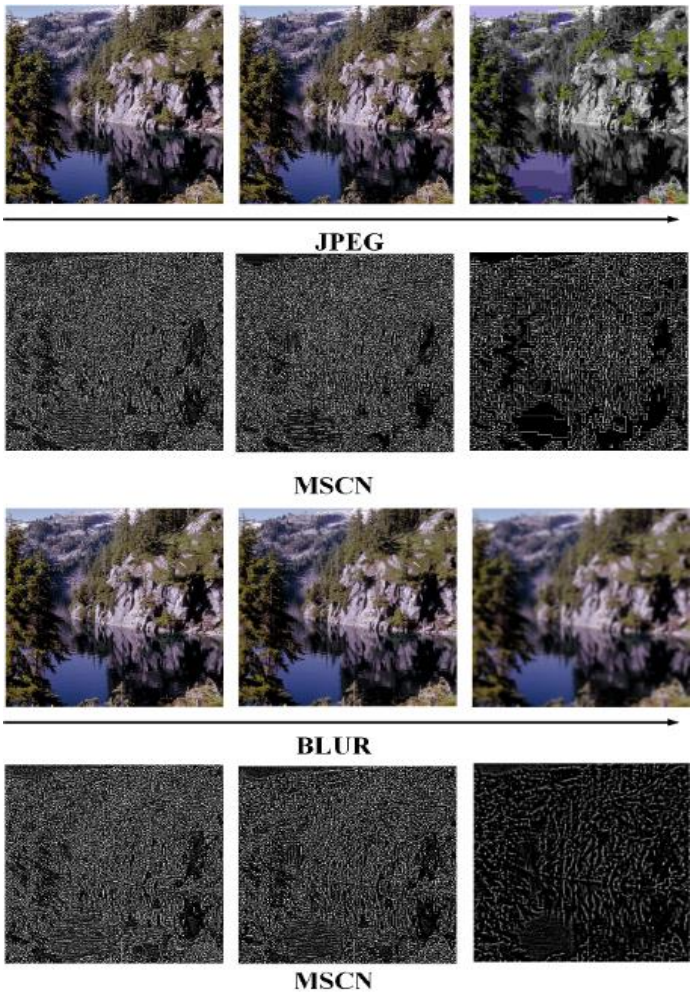

图4 (a) 第一行包含 JPEG压缩图像，第三行包含来自 CSIQ 数据库 *[5]* 的模糊图像。第二行展示了第一行图像对应的均值减除对比度归一化图像，第四行展示了第三行模糊图像对应的均值减除对比度归一化图像。

2(i) 是原始图像，而图 2(b)-(d) 是 JPEG 压缩图像，按质量递减（差分平均意见分数递增）的顺序排列。类似地，图 2(j)-(l) 是经过高斯模糊处理的图像，按差分平均意见分数递增的顺序排列。对每个区块应用基于区块的 8×8 离散傅里叶变换，并在图 2(e)-(h) 和图 2(m)-(p) 中分别展示了与图 2(a)-(d) 和图 2(j)-(l) 所示图像中位置相对应（黄色方框标示）的区块的离散傅里叶变换系数幅值。从这些图中可以看出，离散傅里叶变换系数的分布会随图像中失真（伪影）的类型和程度而变化。离散傅里叶变换系数幅值的这些波动可用于图像质量评估。二维离散傅里叶变换 F(u,v) 可按公式 (1) 计算，其中 I(m,n) 是灰度图像（或图像区块），包含 N×M 个像素。

$$F(u,v) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I(m,n)e^{-j2\pi(\frac{u}{M}m+\frac{v}{N}n)} \quad (1)$$

为了直观呈现离散傅里叶变换的频谱对称性，通常会对频谱进行循环移位，将零频率 (u=0, v=0) 移至中心，该点也被视为频谱平面的原点，如图 3(a) 所示。频移可用公式 (2) 表示：

$$F_1(u,v) = F\left(u-\frac{M}{2}, v-\frac{N}{2}\right) \quad (2)$$

© <2021>。此稿件版本依据 CC-BY-NC-ND 4.0 许可提供

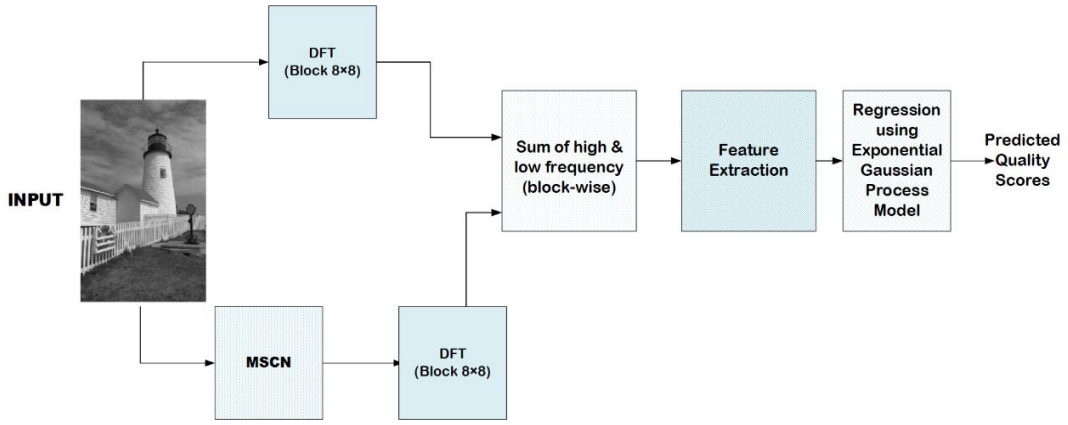

图 5 所提图像质量评价方法的框图

其中，$F_1(u, v)$ 是 $F(u, v)$ 经频移（水平移位 M/2、垂直移位 N/2）后的版本，$F(u, v)$ 定义见公式 (1)。频移后，使用直流系数与块内其他系数之间的曼哈顿距离为各块中的系数编索引。因此，位于 (u, v) 处的频率系数索引 'i' 计算如下:

$$i = |u_0 - u| + |v_0 - v| \tag{3}$$

其中，(u_0, v_0) 表示中心位置 (0,0) 处直流系数的坐标。根据公式 (3)，无论频率系数的方向如何，均按曼哈顿距离以递增频率进行数值索引。然后，可根据频率索引"i"，将 8×8 块的离散傅里叶变换系数划分为低频、中频和高频，如图 3(b) 所示。这里，8×8 块的离散傅里叶变换系数根据其曼哈顿距离从 0 到 8 编制索引。位于中心的直流分量索引为"0"。离中心越远，频率索引"i"越大。在所提方法中，我们将使用此索引来分离低频和高频离散傅里叶变换系数。

2.2 均值减除对比度归一化系数（MSCN）:

人类视觉系统（HVS）已经适应了其自然环境。这一特征体现在自然场景统计（NSS）模型 [53] 中。在过去几十年里，研究人员一直利用 NSS 建立多种图像质量评估指标[29] - [34], [37], [39], [41], [42], 等，这表明 NSS 与人类感知相关。由 NSS 模型导出的一个重要参数是 MSCN [29]。在数学上，它可以表示为:

$$\hat{I}(m,n) = \frac{I(m,n) - \mu(m,n)}{\sigma(m,n) + 1} \tag{4}$$

在公式（4）中，灰度图像表示为 $I(m,n)$，$m \in \{1,2,.....M\}$（M 表示图像高度），$n \in \{1,2,.....N\}$（N 表示图像宽度）为空间索引。

$$\mu(m,n) = \sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l} I_{k,l}(m,n) \tag{5}$$

$$\sigma(m,n) = \sqrt{\sum_{k=-K}^{K} \sum_{l=-L}^{L} w_{k,l} (I_{k,l}(m,n) - \mu(m,n))^2} \tag{6}$$

公式5和公式6分别计算局部均值和对比度。这里，$w=\{w_{k,l} | k = -K, \ldots, K, l = -L, \ldots L\}$ 是高斯加权函数。其采样范围延伸至3个标准差，其中 K=L=3，并重新缩放至单位体积。就对称性而言，它具有二维圆对称性。这种归一化执行的操作与以下操作几乎相同

© <2021>。本文稿版本依据 CC-BY-NC-ND 4.0 许可发布，许可条款见

视觉脑中的处理方式：每个神经元响应都由其相邻神经元响应的能量进行归一化 [1]。视觉皮层细胞会产生非线性响应，尤其是在高对比度情况下，此时基于线性的模型会失效 [54]。从图像中减去均值有助于消除不均匀照明的影响。方差场本身能够显示局部边缘和高对比度细节。组合归一化过程将强度值调整至归一化范围内，留下高度去相关的均匀图像，其中低能量边缘包含重要的结构信息 [29]。这种归一化有助于消除空间冗余，使图像呈现均匀外观。其作用模式与视觉神经元的非线性响应相似。图4展示了从CSIQ 数据库 [5] 中选取的两组图像，其中第1行展示一组JPEG压缩图像，第3行则由按失真程度递增排列的模糊图像组成。第2行和第4行分别展示第1行JPEG图像和第3行模糊图像经过MSCN 操作后的对应图像。可以在第2行和第4行中观察到低能量的残余边界。

3. 所提图像质量评估方法

图 5 展示了所提图像质量评估算法的框图，该算法用于分别或联合估计受到块效应、模糊和噪声失真的图像质量。其工作流程如下。首先，对灰度输入图像进行归一化，以获得相应的均值减除对比度归一化图像。随后，利用 8×8 离散傅里叶变换，将无失真图像和均值减除对比度归一化系数转换到频域。每幅图像中每个块的频率系数被划分为低频和高频频带。完成频域变换后，通过对由灰度图像和均值减除对比度归一化图像得到的离散傅里叶变换块中的高频和低频频率系数分别求和，获得四个求和参数。对于这些求和参数，利用块计数方差确定归一化因子（NF）。使用 NF 对求和参数进行归一化后，统计这些参数值处于特定范围内的块数量。这些块计数被用作特征。通常，无失真自然图像的频域分布呈尖峰状且具有厚尾，而含噪图像则相反。由于噪声会影响高频分量，因此还计算高频求和参数中一百个最大值和一百个最小值的均值，并将其作为特征。为了将特征映射到质量尺度上，使用带[51] 指数核的高斯过程回归模型。下面将介绍所提方法的各个组成部分。

3.1 将离散傅里叶变换与均值减除对比度归一化相结合的动机：

在所提工作中同时使用离散傅里叶变换和均值减除对比度归一化，其动机可从离散傅里叶变换块中零值系数的数量来解释。Golestaneh 等人 [22]首次提出利用离散傅里叶变换零值系数来估计图像中像素块的失真；本研究也采用这一方法评估图像质量。为此，我们选取同一图像的三个版本（一个原始版本和两个失真版本），如图 6(a)、6(d) 和 6(g) 所示。这些图像取自 LIVE 数据库 [52]，，对应的差异平均意见分数分别为 0、47.28 和 83.55。也就是说，图 6(g) 中的图像比图 6(d) 中的图像失真更严重（质量更差）。随后，对每幅图像进行 8×8 离散傅里叶变换（直接变换，或在 MSCN 归一化后变换）。接着，将零值系数映射为黑色像素，将非零值系数映射为白色像素。直接进行离散傅里叶变换并映射后的图像见图 6(b)、6(e) 和 6(h)，分别对应图 6(a)、6(d) 和 6(g) 中的图像。同样，图 6(c)、6(f) 和 6(i) 展示了图 6(a)、6(d) 和 6(g) 中图像经 MSCN 归一化后，其离散傅里叶变换系数映射成黑白图像的结果。从这些图中可以看出，随着失真加重，离散傅里叶变换零值系数（黑色像素）的数量也随之增加。此外，对比图 6(e) 与 6(f) 以及图 6(h) 与 6(i) 可以发现，与仅使用离散傅里叶变换系数相比，结合 MSCN 的离散傅里叶变换能够更好地区分局部失真。

为进一步论证为何所提方法同时考虑经过和未经均值减除对比度归一化处理的离散傅里叶变换系数，我们从 LIVE 数据库中选取了六组 JPEG 图像（原始图像及其失真版本）[52]，其中六幅原始图像（图像1、图像2、……、图像6）如图7所示。每组图像（包括原始图像和失真图像）均通过8×8 离散傅里叶变换进行转换，并计算变换后图像中的零值系数数量。为了了解零值离散傅里叶变换系数与相应图像质量（以差异平均意见分数值表示）之间的相关性，我们将每组图像的零值离散傅里叶变换系数数量与差异平均意见分数值进行绘图比较

© <2021>。本文稿版本依据 CC-BY-NC-ND 4.0 许可提供

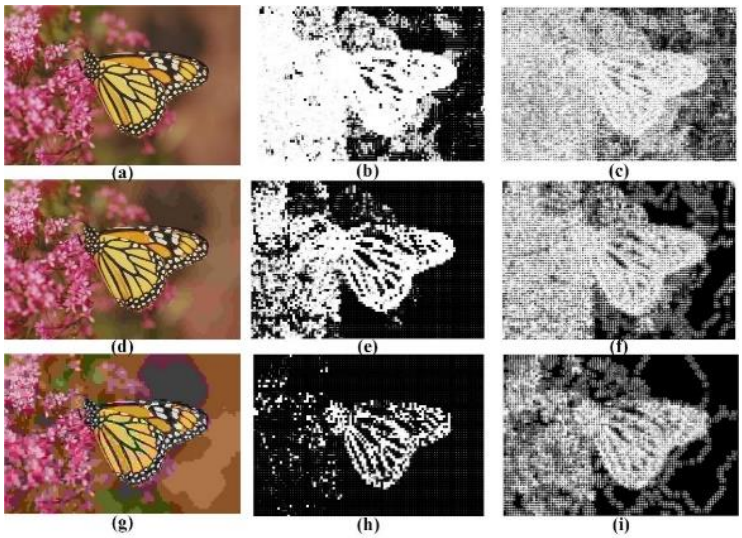

图6 LIVE *[52]* 数据库中的三幅 JPEG 图像，其 DMOS 分别为 (a) 0、(d) 47.28 和 (g) 83.55；灰度图像的离散傅里叶变换结果分别见 (b)、(e) 和 (h)，对 MSCN 系数应用离散傅里叶变换的结果见 (c)、(f) 和 (i)。

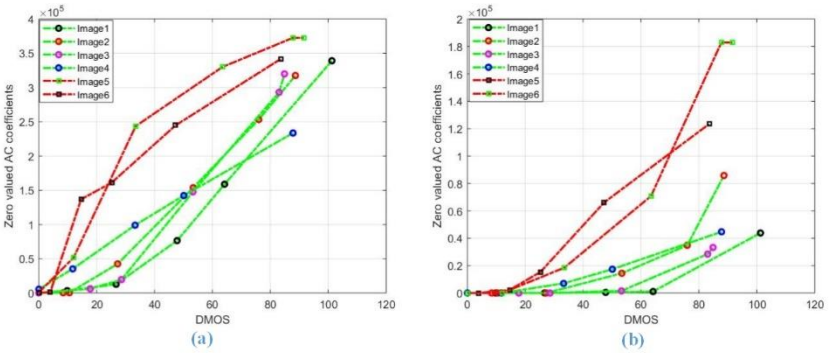



图7 图像1、2、3、4、5和6：来自 LIVE *[52]* 数据库的原始图像。(a) 灰度图像离散傅里叶变换的零值系数图；(b) MSCN 图像离散傅里叶变换的零值系数图（图像1、2、3、4、5和6组成的集合，包含原始图像以及 JPEG LIVE *[52]* 数据库中的所有失真版本）。

© <2021>。本文稿版本依据 CC-BY-NC-ND 4.0 许可提供

（取自数据库），如图7(a)所示。从图中可以看出，对于前4幅图像（图像1-4），零值系数的数量与差分平均意见分数高度相关，但对于图像5和图像6，这些参数之间的相关性较差。值得注意的是，这两幅图像（图像5和图像6）存在局部模糊区域。因此，经过离散傅里叶变换的图像中的零值交流系数数量，无法与局部失真图像的图像质量相关联。然而，对同一组图像进行 MSCN 归一化后再进行离散傅里叶变换，局部失真图像中的零值系数数量与相应图像的差分平均意见分数具有良好的相关性，如图7(b)中的红色曲线所示；但对于全局失真图像（图像1-4），相关性较差。需要指出的是，对于这些图像（图像1-4），灰度图像的离散傅里叶变换系数中零值系数的数量具有良好的相关性，这一点从图7(a)中可以看出。因此，为了更好地评估每幅失真图像的质量，应计算经过和未经过 MSCN 处理的离散傅里叶变换图像的特征。随后将这些特征输入机器学习算法进行进一步处理。

3.2 频率分离：

如前所述，对输入图像逐块应用离散傅里叶变换（块大小为 8×8 像素）。与此同时，从图像中提取 MSCN 系数，以考虑受人类视觉系统启发的自然场景统计模型的特性。随后，使用分块离散傅里叶变换将计算得到的 MSCN 系数也变换到频域。由此，获得图像的两种频域表示（一种经过 MSCN 处理，另一种未经过处理）。设灰度输入图像 $(I(m,n))$ 及其经 MSCN 归一化后的版本 $(\hat{I}(m,n))$ 的离散傅里叶变换分别为 F_g (u, v) 和 F_m(u, v)。可以观察到，对于图像中的不同类型失真，其低频分量和高频分量的能量分布可能不同。此外，由于离散傅里叶变换系数所含的能量等于其绝对值的平方，因此我们仅考虑离散傅里叶变换系数的幅值。分别将 F_g(u, v) 和 F_m(u, v) 的绝对值记为 $\left|F_g(u,v)\right|$ 和 $|F_m(u,v)|$ 。

此外，如第 2.1 节前文所述，为了区分低频和高频，会根据块的直流系数的曼哈顿距离为块的离散傅里叶变换系数编制索引，如式（3）所定义。频率索引相同的系数归为一组，如图 3(c) 所示；所有频率索引为 1–3 的交流系数标记为低频（LF）系数，索引为 4 的系数标记为中频（MF）系数，索引为 5–8 的系数标记为高频（HF）系数。在所提方法中，仅使用属于 LF 和 HF 频带的系数进行特征提取。如果 $F_{k,g} = \{F_g(u,v)\}_k$ 和 $F_{k,m} = \{F_m(u,v)\}_k$分别表示离散傅里叶变换块中所有具有 k^{th} 频率索引的离散傅里叶变换系数集合（分别为不经过和经过均值减除对比度归一化处理），那么，图像在离散傅里叶变换域中的 LF 和 HF 频带（分别为不经过和经过均值减除对比度归一化处理）是频率索引为 1–3 和 5–8 的所有离散傅里叶变换系数的集合，对于每个块，可根据式（7）–（10）表示为：

$$LF_g = \cup_{k=1}^{3}\{F_{k,g}\} \tag{7}$$

$$HF_g = \cup_{k=5}^{8}\{F_{k,g}\} \tag{8}$$

$$LF_m = \cup_{k=1}^{3}\{F_{k,m}\} \tag{9}$$

$$HF_m = \cup_{k=5}^{8}\{F_{k,m}\} \tag{10}$$

3.3 四个求和参数：

为了确定特征向量，我们提出将分别属于集合LF_g、 HF_g、 LF_m、 $and\ HF_m$的所有离散傅里叶变换系数的幅值求和，这些集合分别如公式 (11)–(14) 所定义，并将其作为特征来量化失真图像的质量。

© <2021>。本文稿版本依据 CC-BY-NC-ND 4.0 许可公开

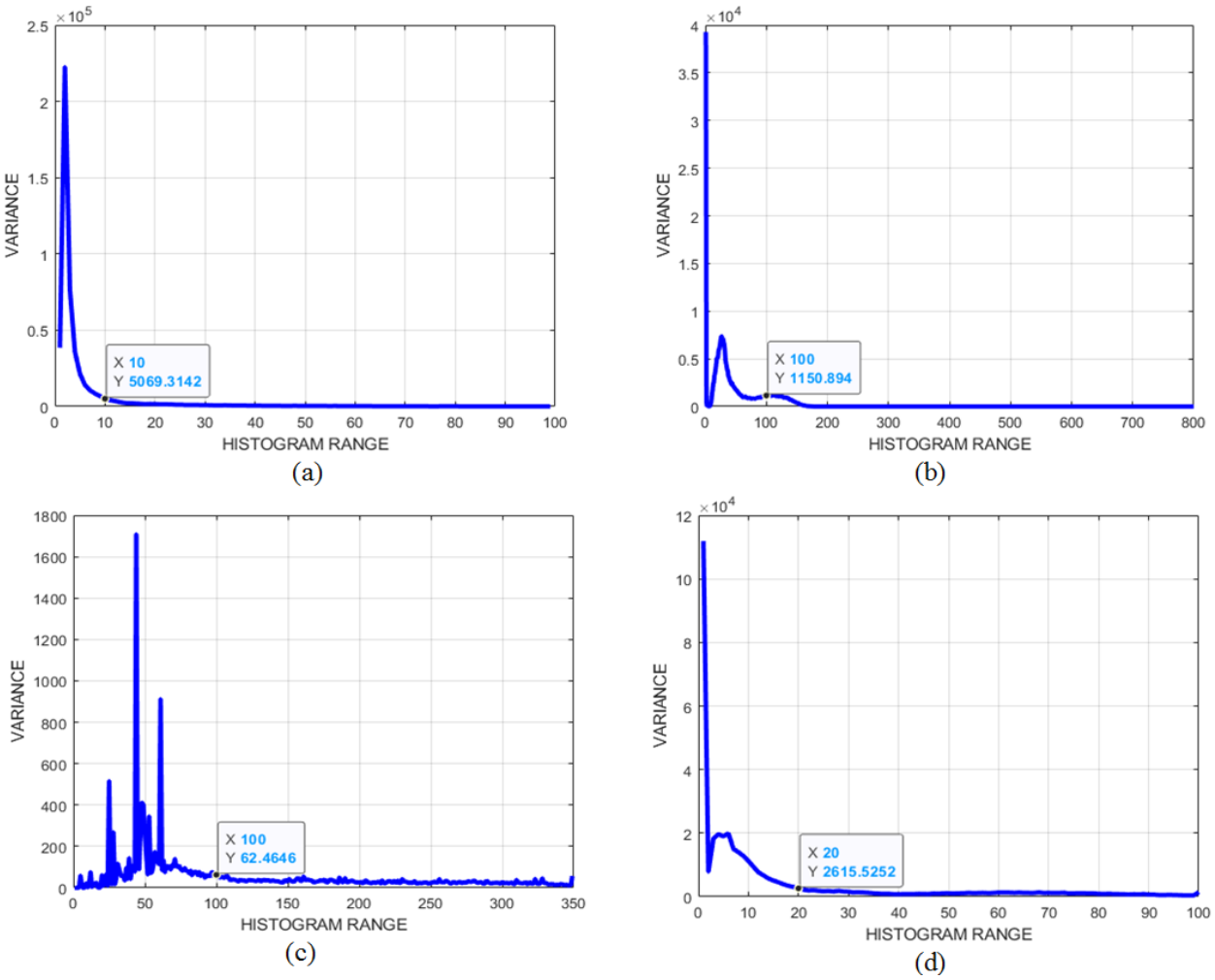

图 8 求和参数的方差分布 (a) S_g^{LF} (b)S_m^{LF} (c)S_g^{HF}(d) S_m^{HF}

$$S_g^{LF} = \cup_{all\ blocks} \sum_{k=1}^{3} |F_{k,g}| \tag{11}$$

$$S_g^{HF} = \cup_{all\ blocks} \sum_{k=5}^{8} |F_{k,g}| \tag{12}$$

$$S_m^{LF} = \cup_{all\ blocks} \sum_{k=1}^{3} |F_{k,m}| \tag{13}$$

$$S_m^{HF} = \cup_{all\ blocks} \sum |F_{k,k=5}^{8}{}_{m}| \tag{14}$$

由于求和参数 S_g^{LF}、 S_g^{HF}、 S_m^{LF} 和 S_m^{HF} 的取值范围各不相同，并且还取决于失真的性质，因此在进一步处理时，应将每个求和参数归一化到 0 到 1 的范围内，具体讨论如下。

3.4 归一化因子:

上文公式（11）至（14）中定义的与离散傅里叶变换（DFT）低频和高频分量（包括经过和未经过均值减除对比度归一化（MSCN）的分量）相关的四个参数，其取值可能因图像内容以及失真的性质和程度而异。为了了解这四个求和参数的取值范围，从 GBLUR（高斯模糊）和 JPEG 压缩的 LIVE 数据库 [52] 子集中随机选取了 100 幅图像。对于每幅图像，

© <2021>。本手稿版本依据 CC-BY-NC-ND 4.0 许可提供

均计算了公式（11）至（14）中定义的四个求和参数。随后，计算这些求和参数在不同取值范围内所包含的块数。针对每个求和参数，分别统计 100 幅图像中某一特定箱值（对应于块数）出现的次数，并将这些次数的方差绘于图 8 中。利用与每个箱对应的数值的方差来确定各求和参数的最大范围。图 8(a) 的箱大小为 100，而其余参数［图 8(b)、图 8(c) 和图 8(d)］的箱大小为 1。从图 8 可以清楚看出，对于大多数图像，四组求和参数 S_g^{LF}、 S_g^{HF}、 S_m^{LF} 和 S_m^{HF} 中各参数的取值分别位于 0–10（箱大小为 100，图 8(a)）、0–100（图 8(c)）、0–100（图 8(b)）和 0–20（图 8(d)）范围内。为了将它们统一到 0 至 1 的范围内，需要进行归一化。将每个求和参数的最大范围视为归一化因子（NF），从而使每个求和参数都能表示在 0 至 1 的尺度上。四组求和参数 S_g^{LF}、 S_g^{HF}、S_m^{LF} 和 S_m^{HF} 的归一化因子（NF）分别为 1000、100、100 和 20。确定求和参数后，将其分别除以对应的 NF，以确保这些参数的归一化值保持在 0 至 1 的尺度上。需要注意的是，归一化因子确保两类失真均在统一尺度内定义。

3.5 特征提取：

最后，按如下方式提取特征。由于所有求和参数均归一化到 0 至 1 的范围，其中，每个求和参数取值为零对应于图像中的交流系数为零。因此，将每个求和参数取值为零的块数用作四个特征（每个求和参数对应一个特征）。整个取值范围（零值除外）被划分为四个相等区间：0-0.25（不包括 0）、0.25-0.50、0.50-0.75 和 0.75-1.0。每个求和参数在这四个区间内的块数将构成 4×4=16 个特征。由于图像尺寸可能不同，因此，对于落在不同区间内的每个求和参数，采用块所占的百分比而非块的绝对数量作为特征向量。由于共有四个求和参数，且每个参数分为 5 类，因此以此方式得到 20 个特征。这 20 个特征向量定义于表 1 中，其数值为属于各类别的块数除以图像中的块总数。特征 f_1, f_2, f_3, f_4 和 f_5 分别表示求和参数集合 S_g^{LF} 中的归一化值为 0、处于 0-0.25、0.25-0.50、0.50-0.75 和 0.75-1.0 范围内的块所占百分比。类似地，f_6-f_{10} 对应求和参数集合 S_m^{LF}，f_{11}-f_{15} 对应 S_g^{HF}，f_{16}-f_{20} 对应 S_m^{LF}，f_{11}-f_{15} 对应 S_g^{HF}，f_{16}-f_{20} 对应 S_m^{HF}。此外，通过分别计算求和参数集合 S_g^{HF} 和 S_m^{HF} 中最大的一百个值和最小的一百个值的均值，额外提取四个特征（f21- f24）。因此，从四组求和参数中共得到 24 个特征。

3.6 回归阶段：

基于机器学习的回归模块用于将特征向量映射到相应的质量评分。高斯过程回归（GPR）是一种基于核函数的监督方法，具有非参数性质。根据所使用的核函数，它有许多变体，例如线性核、指数核、平方指数核以及多个核函数的组合等。本研究使用指数高斯过程回归模块，将图像的频域自然场景统计特征映射为其质量评分。

高斯过程（GP）是一种随机过程，可表示为函数上的分布 ξ，它将输入特征空间 F（在本研究中为特征向量）映射到输出空间 D（此处为由图像的差分平均意见分数组成的空间）。它可以用均值 $m(f)$ 和协方差 $C(f, f')$（也作为高斯过程核）建模，分别如公式 15 和公式 16 所定义。

$$m(f) = E[\xi(f)] \tag{15}$$

$$C(f, f') = E[\big(\xi(f) - m(f)\big)\big(\xi(f') - m(f')\big)] \tag{16}$$

在公式16中，f 和 f′ 是集合 F 中的元素，E 是期望算子。在所提方法中，输入空间 F 是表 I 中列出的从 f1 到 f20 的特征集合。随后，GP 是 m(f) 和 C(f, f') 的非线性函数，并在公式17中定义如下：

© <2021>。本手稿版本依据 CC-BY-NC-ND 4.0 许可公开发布

$\xi(f) \sim GP(m(f),\ C(f,f'))$ 它实现了 i 的映射 (17)

通过施加如下限制，将输入 $f_i \in$ F 映射到输出空间 $d_i \in$D：事先已知一组输入-输出训练点 $\{(f_i, d_i)|i = 1, . , . , . , n\}$，以便获得后验分布。GP 模型的输出 d 是一个带噪声的观测值，表示为：

$$d = \xi(f) + \acute{\varepsilon} \tag{18}$$

其中，$\acute{\varepsilon}$ 是加性高斯噪声$N(0, \sigma^2)$。为从测试数据集中获得预测的差异平均意见分数，使用该后验分布。协方差（即高斯过程模型的核函数）也是已知的，其超参数在训练阶段进行优化。训练输出 d_{tr} 与测试输出 d_p 的联合分布定义如下：

$$\begin{bmatrix} d_{tr} \\ d_p \end{bmatrix} \sim N(0, \begin{bmatrix} C(F,F) + \sigma_n^2 I & C(F,F_*) \\ C(F,F) & C(F,F_*) \end{bmatrix}) \tag{19}$$

其中，F 和 F_* 分别是训练数据和测试数据的特征向量。噪声的标准差表示为$\sigma_n^2 I$，其中 I 是单位矩阵（N 维）。为获得预测函数，以d_{tr} 为条件对观测值d_p 进行处理，使得

$$d_p|F, d_{tr}, F_* \sim N(\overline{d_p}, V(d_p)) \tag{20}$$

其中，$\overline{d_p} = C(F_*, F)[C(F,F) + \sigma_n^2 I]^{-1} d_{tr},$ (21)

$$V(d_p) = C(F_*, F_*) - C(F_*,)\,[C(F,F) + \sigma_n^2 I]^{-1}\, C(F, F_*) \tag{22}$$

GP 使用该预测分布来预测输出 d_p。测试输出（即预测的差分平均意见分数）为 d_p，如公式 20 所定义。关于 GP 的更多详情，请参阅 [51]。本研究使用 MATLAB 18a 机器学习工具箱中的 Regression Learner App 对 GP 模型进行训练和验证。

4. 结果与讨论

本节将分析所提模型，以考察其在相关知名单失真数据库 LIVE [52], CSIQ [5], TID2013 [40] 和 KADID-10k [55] 上的性能是否符合人类感知。为测试最先进方法和所提方法在多重失真情形下的准确性，采用 LIVEMD [56],I VL MD [57], [58], MDID2013 [46] 和 MDID [59] 数据集；对于散景模式图像，则采用 IBBI 数据库 [16], [60]。用于比较的方法包括特定失真方法，如 Ferzli [9], EZMGM [15],Golestaneh [22], Zhan [23], Zhu [24], H. Liu [16]，以及通用方法，如 DIIVINE [31],BLIINDS2 [32], CORNIA [38], BRISQUE [29], GMLoG [33], NFERM [34], FRIQUEE [37], BJLC [39], NIQE [41]和 IL-NIQE [42]。此外，为了公平比较，还纳入了一些专为多重失真图像设计的方法，具体包括 GWH-GLBP [44], Miao [45], Zhou[47] 和 SISBLIM [46]。这些方法可分为两类：基于机器学习的方法和非机器学习方法。图像质量评价方法 DIIVINE [31], BLIINDS2 [32], CORNIA [38], BRISQUE [29], GMLoG [33], NFERM [34], FRIQUEE [37], BJLC [39], H. Liu [16], GWH-GLBP [44], Miao [45] 和 Zhou[47] 属于基于机器学习的方法，而 NIQE [41], IL-NIQE [42], Golestaneh [22], Zhan [23], Zhu [24], Ferzli [9], Bahrami [13], EZMGM [15] 和 SISBLIM [46] 属于非机器学习方法。

4.1 数据库和评估标准：

现有多个数据库，其中包含遭受不同类型和不同程度失真的图像。部分所使用数据库的详细信息如下：

- LIVE [52] 由29幅参考图像组成，其中233幅受到JPEG压缩影响，174幅受到高斯模糊影响，另有145幅图像遭受加性白高斯噪声（AWGN）类型的失真，该数据库在此称为LIVE。
- CSIQ [5] 数据库包含 30 幅参考图像，以及 JPEG、模糊和加性白高斯噪声子数据集中各 150 幅失真图像，以下分别简称为 CSIQ。

© <2021>。此稿件版本依据 CC-BY-NC-ND 4.0 许可公开

- TID2013 [40] 数据库包含共 24 种失真类型和 25 幅参考图像，其中 120 幅 JPEG压缩，120幅经受GBLUR，另有120幅受到加性白高斯噪声影响，在本研究中统称为TID2013。每个子集中有5幅图像因并非自然图像而未被纳入。
- KADID-10k [55] 是最大的基于单一失真的数据集，由从81幅参考图像获得的10,125幅失真图像组成。本研究包括GBLUR、JPEG压缩和加性白高斯噪声子集，每个子集分别包含405幅失真图像。
- LIVEMD数据库 [56] 共包含405幅失真图像，其中270幅为多重失真图像。它由15幅参考图像组成，这些图像受到三种失真类型的影响（高斯模糊、JPEG压缩和白高斯噪声）。除受单一失真类型影响的各个独立子集外，该数据库还包含两个受多重失真影响的子集，类型为：(i) 高斯模糊后接JPEG压缩；以及(ii) 高斯模糊后接白噪声（WN）。这两个子集用于评估多重失真图像的质量，本文中称为LIVEMD。
- IVL MD数据库[57], [58] 包含10幅原始图像，并被划分为两个子集，共包含750幅多重失真图像。该数据库由两个子集组成：1）模糊-JPEG，其中每幅原始图像经过7个等级的高斯模糊处理，随后每幅所得的模糊图像再经过5个等级的JPEG压缩；2）噪声-JPEG，其中每幅参考图像受到10个等级的高斯噪声影响，随后每幅含噪图像再经过4个等级的JPEG压缩。
- 多重失真图像数据库2013（MDID2013）数据库[46] 包含324幅多重失真图像，这些图像由12幅原始图像依次经过三种失真（高斯模糊、JPEG和白噪声）生成。
- 多重失真图像数据库（MDID）[59] （不同于MDID2013）包含20幅参考图像和1600幅多重失真图像，这些图像通过引入五种失真生成：高斯噪声、高斯模糊、对比度变化、JPEG和JPEG2000压缩。在本研究中，使用了受高斯噪声、高斯模糊和JPEG影响的图像进行分析，并将其称为MDID。
- IBBI [16], [60] 数据库包含12张参考图像和60张具有有意模糊背景。

在性能评估中，主要采用了不同的准则：斯皮尔曼等级相关系数（SROCC）、皮尔逊线性相关系数（PLCC）、肯德尔等级相关系数（KROCC）和均方根误差（RMSE）。SROCC、KROCC 和 PLCC 的取值范围为 [-1, 1], , 而 RMSE 为非负实数。这些参数反映了预测评分与主观评分之间的相关性。较低的 RMSE 和较高的 SROCC/PLCC/KROCC 表明与主观判断具有良好的相关性。由于人类响应系统是非线性的，主观评分和预测评分中可能存在非线性。因此，采用逻辑

表2 LIVE、CSIQ 和 TID2013 数据库中 JPEG、高斯模糊和白噪声类型失真的 SROCC、皮尔逊线性相关系数、KROCC 和均方根误差值。最佳的两个结果以粗体标示。

方法	LIVE			CSIQ			TID2013	
	SROCC	皮尔逊线性相关系数	KROCC	均方根误差	SROCC	皮尔逊线性相关系数	KROCC	均方根误差	SROCC	皮尔逊线性相关系数	KROCC	均方根误差
DIIVINE [31]	0.831	0.826	0.685	15.24	0.832	0.839	0.642		0.146	0.675	0.635	0.507	1.016
BLIINDS2 [32]	0.907	0.914	0.737	11.08	0.849	0.873	0.653		0.131	0.751	0.781	0.564	0.820
NIQE [41]	0.924	0.916	0.755	10.98	0.856	0.870	0.663		0.132	0.736	0.741	0.529	0.898
IL-NIQE [42]	0.902	0.908	0.722	11.62	0.871	0.874	0.682		0.155	0.863	0.868	0.616	0.655
CORNIA [38]	0.945	0.951	0.812	8.75	0.916	0.881	0.763		0.153	0.899	0.938	0.686	**0.563**
BRISQUE [29]	0.957	0.960	0.819	7.77	0.914	0.936	0.752		0.095	0.832	0.852	0.663	0.688
GMLoG [33]	0.962	0.957	0.823	7.79	0.926	0.914	0.788		0.113	0.931	0.926	0.763	0.603
NFERM [34]	0.942	0.955	0.810	8.48	0.921	0.933	0.776		0.101	0.908	0.911	0.691	0.615
FRIQUEE [37]	0.963	0.953	0.823	8.55	0.939	0.895	0.795		0.149	0.926	0.845	0.755	0.629
BJLC [39]	0.968	0.963	0.834	**7.61**	0.923	0.928	0.785		0.102	0.941	0.885	0.783	0.618
所提方法	0.979	0.981	0.861	**5.67**	0.932	0.948	0.794	0.096	0.922	0.928	0.748	0.599	

© <2021>。本手稿版本依据 CC-BY-NC-ND 4.0 许可提供

表 3 LIVE、CSIQ 和 TID2013 数据库中 JPEG、高斯模糊和白噪声类型单一失真的 SROCC。最佳的两项结果以粗体标出。

方法	JPEG			BLUR			WN		
	LIVE	CSIQ	TID2013	LIVE	CSIQ	TID2013	LIVE	CSIQ	TID2013
Golestaneh *[22]*	0.954	0.923	0.891	-	-	-	-	-	-
Zhan *[23]*	0.961	0.940	0.936	-	-	-	-	-	-
Zhu *[24]*	0.961	**0.952**	**0.952**	-	-	-	-	-	-
Ferzli *[9]*	-	-	-	0.943	0.886	0.857	-	-	-
EZMGM *[15]*	-	-	-	0.958	0.932	0.944	-	-	-
Bahrami *[13]*	-	-	-	0.954	0.925	0.936	-	-	-
DIIVINE [31]	0.914	0.888	0.826	0.938	0.856	0.943	0.979	0.940	0.906
BLIINDS2 [32]	0.951	0.898	0.775	0.926	0.924	0.858	0.941	0.922	0.741
NIQE [41]	0.943	0.882	0.880	0.939	0.906	0.835	0.961	0.835	0.852
IL-NIQE[42]	0.944	0.904	0.883	0.924	0.867	0.864	0.977	0.866	0.904
CORNIA [38]	0.945	0.906	0.891	**0.960**	0.914	0.922	0.979	0.941	0.935
BRISQUE [29]	0.963	0.904	0.869	0.947	0.919	0.893	0.981	0.957	0.900
GMLoG [33]	0.965	0.916	0.923	0.938	0.915	0.929	0.978	0.943	0.946
NFERM [34]	0.967	0.922	0.911	0.948	0.897	0.926	0.980	0.938	0.931
FRIQUEE [37]	0.955	0.934	0.912	0.956	**0.942**	0.945	0.975	0.941	**0.950**
BJLC [39]	**0.968**	**0.950**	0.946	0.954	0.930	**0.968**	**0.986**	**0.962**	**0.958**
所提方法	**0.979**	0.948	**0.958**	**0.972**	**0.944**	**0.951**	**0.984**	**0.960**	0.947

表 4 LIVE、CSIQ 和 TID2013 数据库中 JPEG、高斯模糊和白噪声类型单一失真的皮尔逊线性相关系数。最佳的两个结果以粗体标出。

方法	JPEG			BLUR			WN		
	LIVE	CSIQ	TID2013	LIVE	CSIQ	TID2013	LIVE	CSIQ	TID2013
Golestaneh *[22]*	0.970	0.954	0.938	-	-	-	-	-	-
Zhan *[23]*	0.978	**0.963**	0.962	-	-	-	-	-	-
Zhu *[24]*	0.971	**0.984**	0.962	-	-	-	-	-	-
Ferzli *[9]*	-	-	-	0.911	-	0.852	-	-	-
EZMGM *[15]*	-	-	-	0.956	**0.928**	**0.942**	-	-	-
Bahrami *[13]*	-	-	-	**0.959**	0.907	0.875	-	-	-
DIIVINE [31]	0.934	0.897	0.899	0.937	0.898	0.844	0.971	0.899	0.882
BLIINDS2 [32]	0.943	0.912	0.889	0.899	0.901	0.825	0.932	0.897	0.714
NIQE [41]	0.935	0.893	0.882	0.936	0.897	0.819	0.946	0.855	0.851
IL-NIQE [42]	0.942	0.908	0.876	0.918	0.885	0.868	0.974	0.858	0.917
CORNIA [38]	0.965	0.877	**0.963**	0.955	0.882	0.941	0.965	0.942	0.937
BRISQUE [29]	0.973	0.946	0.951	0.951	**0.928**	0.863	0.974	0.938	0.810
GMLoG [33]	0.953	0.907	0.925	0.942	0.920	0.931	0.966	0.944	0.925
NFERM [34]	**0.981**	0.944	0.907	0.937	0.897	0.928	**0.979**	0.936	0.939
FRIQUEE [37]	0.947	0.885	0.813	0.949	0.905	0.881	**0.982**	**0.947**	**0.955**
BJLC [39]	0.961	0.942	0.873	0.952	0.926	0.853	-	-	-
所提方法	**0.986**	0.949	**0.966**	**0.971**	**0.941**	**0.953**	0.977	**0.965**	**0.946**

函数在计算 PLCC 和 RMSE 之前对评分进行逻辑映射 [61]。该函数可定义如式（23）所示，其中 x 和 f(x) 分别为回归前后的预测评分，$\beta_1, \beta_2, \beta_3, \beta_4$，以及 β_5 为回归模型参数。

$$f(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{\exp(\beta_2 (x - \beta_3))} \right) + \beta_4 x + \beta_5 \tag{23}$$

© <2021>。本文稿版本依据 CC-BY-NC-ND 4.0 许可提供

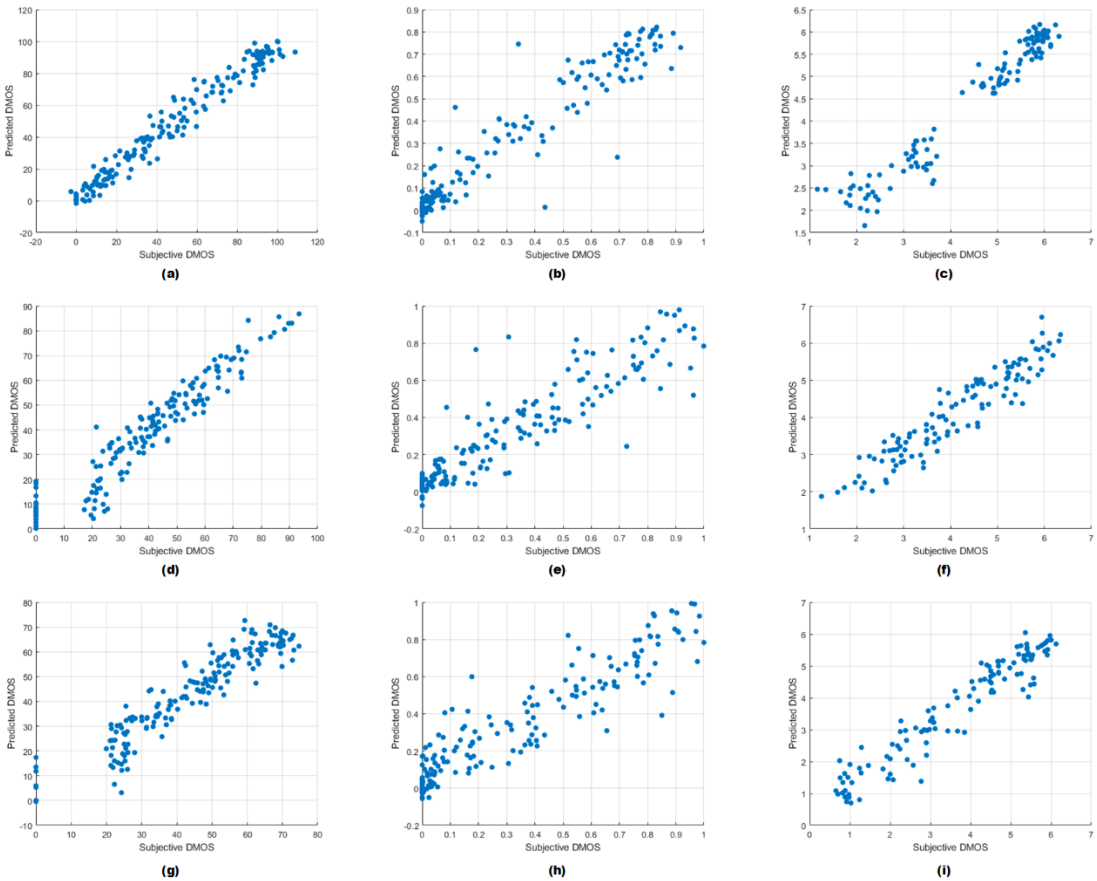

图9 不同数据集的标准分数与预测分数之间的关系图。(a) 用于 JPEG 压缩图像的 LIVE *[52]* 数据集 (b) 用于 JPEG 压缩图像的 CSIQ *[5]* 数据集 (c) 用于 JPEG 压缩图像的 TID*2013 [40]* 数据集 (d) 用于模糊失真的 LIVE *[52]*数据集 (e) 用于模糊失真的 CSIQ *[5]* 数据集 (f) 用于模糊失真的 TID*2013 [40]* 数据集

表 5 KADID-10k 上 JPEG、高斯模糊和白噪声图像数据集以及 IBBI 数据库上的散景模式图像的 SROCC 和皮尔逊线性相关系数。表现最佳的两种方法以粗体标出。KADID-10k IBBI

	SROCC			PLCC			SROCC	PLCC
	JPEG	高斯模糊	WN	JPEG	高斯模糊	WN	BLUR	BLUR
DIIVINE [31]	0.766	0.787	0.625	0.805	0.794	0.648	0.978	0.955
BLIINDS2 [32]	0.789	0.759	0.718	0.812	0.806	0.682	0.971	0.974
ILNIQE [42]	0.871	0.865	0.737	0.908	0.843	0.694	0.964	0.972
NIQE [41]	0.848	0.872	0.829	0.928	0.893	0.826	0.978	0.965
BRISQUE [29]	0.789	0.814	0.583	0.810	0.828	0.588	**0.980**	**0.984**
GMLoG [33]	0.811	0.809	0.847	0.856	0.845	0.816	0.972	**0.981**
NFERM [34]	0.865	0.884	0.864	0.916	**0.938**	0.857	0.984	0.972
FRIQUEE [37]	**0.896**	**0.910**	**0.871**	**0.948**	0.936	**0.884**	0.960	0.962
H. Liu [16]	-	-	-	-	-	-	-	0.973
所提方法	**0.917**	**0.934**	**0.904**	**0.966**	**0.957**	**0.902**	**0.985**	0.977

4.2 各数据库上的性能：本文分析了所提方法在五个数据库的单一失真情况下的性能，即 LIVE[52], CSIQ [5],

© <2021>。此手稿版本依据 CC-BY-NC-ND 4.0 许可公开发布

表6 JPEG 和模糊失真类型的 SROCC 对比交叉验证。最佳性能以粗体标示。

LIVE:CSIQ LIVE:TID

	JPEG	高斯模糊	JPEG	高斯模糊
Golestaneh *[22]*	0.882	--	0.871	--
Zhan *[23]*	0.937	--	0.927	--
Zhu *[24]*	**0.941**	--	**0.951**	--
Ferzli *[9]*	--	0.874	--	0.851
EZMGM *[15]*	--	0.922	--	0.925
Bahrami *[13]*	--	**0.924**	--	**0.938**

表7 JPEG和模糊类型失真的SROCC与皮尔逊线性相关系数交叉验证比较。最优的两项结果以粗体标示

	LIVE:CSIQ (SROCC)		LIVE:CSIQ (皮尔逊线性相关系数)				LIVE:TID2013 (SROCC)		LIVE:TID2013 (皮尔逊线性相关系数)		
	JPEG	高斯模糊	白噪声	JPEG	高斯模糊	白噪声	JPEG	高斯模糊	白噪声	JPEG	高斯模糊	WN
DIIVINE [31]	0.845	0.888	0.876	0.831	0.862	0.851	0.845	0.888	0.868	0.837	0.835	0.857
BLIINDS2 [32]	0.857	0.783	0.893	0.858	0.803	0.898	0.857	0.783	0.736	0.876	0.774	0.712
CORNIA [38]	0.895	**0.913**	0.750	0.887	0.854	0.763	0.895	**0.913**	0.736	0.882	0.905	0.738
BRISQUE [29]	0.889	0.872	0.899	0.905	0.871	0.882	0.889	0.872	0.820	0.871	0.856	0.844
ILNIQE [42]	0.899	0.857	0.848	0.891	0.874	0.836	0.867	0.834	0.885	0.868	0.840	0.863
NIQE [41]	0.882	0.895	0.809	0.863	0.881	0.828	0.862	0.815	0.816	0.884	0.811	0.823
GMLoG [33]	0.895	0.911	0.896	0897	0.894	0.905	0.895	0.911	0.899	0.892	**0.918**	0.903
NFERM [34]	0.913	0.883	0.916	0.918	0.882	0.917	0.913	0.883	0.905	0.853	0.891	**0.922**
FRIQUEE [37]	0.835	0.877	0.849	0.817	**0.896**	0.844	0.835	0.877	0.810	0.838	0.862	0.846
BJLC [39]	**0.941**	0.899	**0.932**	-	-	-	**0.916**	0.891	0.903	-	-	-
所提方法	**0.937**	**0.946**	0.956	0.945	0.932	0.961	0.948	**0.939**	0.938	0.946	**0.937**	**0.948**

KADID-10k [55], IBBI [16], [60] 和 TID2013 [40]，并与其他最先进方法进行比较。所考察方法的源代码，即 DIIVINE [31], BLIINDS2 [32], CORNIA [38], BRISQUE [29],GMLoG [33], NFERM [34], FRIQUEE [37], GWH-GLBP [44], NIQE [41], IL-NIQE [42], Miao [45] 和 SISBLIM[46]，均从相应作者的网站获取。其中，基于机器学习的方法按照与所提方法相同的方式，在所考察的数据集上进行训练。对于包括所提方法在内的每种方法，每个数据集均随机划分为训练集和测试集，比例为 80%:20%；第 4.6 节中则采用 50%:50% 的比例。为确保公平分析，报告结果为 1000 次独立训练集—测试集划分迭代所得的中位数。对于 NIQE [41] 和 IL-NIQE [42], 1000等免训练方法，为保持一致性，在测试集上进行试验，并将相应的中位数作为结果。对于 BJLC [39], Zhou [47], Golestaneh [22], Zhan [23], Zhu [24], Ferzli[9], EZMGM [15], Bahrami [13], 和 H. Liu [16] 等其他方法的性能评估结果，则取自相应的原始论文或最相关的研究文章。由于 Golestaneh [22], Zhan [23], Zhu [24] 方法针对 JPEG失真而设计，而 Ferzli [9], EZMGM [15], Bahrami [13] 则针对高斯模糊类型的失真，因此这些方法的分数仅报告对应特定失真类型的结果。对于 Zhou [47], 提出的方法，仅报告其在 MDID2013 [46] 和 LIVEMD [56] 上的分数，因为原论文只提及了这两个数据库上的结果。

包含所有经受高斯模糊、JPEG 和白噪声类型失真的图像的合并数据库 LIVE [52], 、CSIQ [5] 和 TID 2013 [40] 的实验结果如表2所示。采用粗体突出显示的表现最佳的两种方法的结果表明，与其他现有先进方法相比，所提方法始终取得良好结果。表3和表4分别表示三个数据库（即 LIVE [52], 、CSIQ [5] 和 TID 2013 [40]）中各自包含经受高斯模糊、JPEG 和白噪声类型失真的图像的单独失真数据库的预测单调性和预测准确度。从这些结果可以清楚地观察到

© <2021>。本手稿版本依据 CC-BY-NC-ND 4.0 许可公开发布

对于所有数据库，所提方法始终位列表现最佳的两种方法之列，但 CSIQ [5] 数据库中的 JPEG压缩图像和 TID2013 数据库中的白噪声图像除外。与其他方法相比，所提方法始终表现良好，并且与 FRIQUEE [37] 和 BJLC [39] 方法相当。具体而言，对于 TID2013 [40] ，BRISQUE [29], 、GMLoG [33] 和 NFERM [34] 等广受认可的方法在比较皮尔逊线性相关系数值时表现显著较差，而所提方法在 SROCC 和皮尔逊线性相关系数值方面均取得了最佳表现。

此外，为了检验所提方法更广泛的有效性，我们将其在 KADID-10k [55] 和 IBBI [16], [60] 数据库上的 SROCC 和皮尔逊线性相关系数与其他方法进行比较，结果如表 5 所示。值得注意的是，本研究用于分析的 KADID-10k 数据库包含受到单一类型失真的图像，即高斯模糊、JPEG压缩和白噪声；而 IBBI [16], [60] 数据库则包含有意模糊的背景图像。从结果可以看出，对于 KADID-10k 数据库中的三种失真，所提方法在两项评价指标（SROCC 和皮尔逊线性相关系数）上均稳居表现最佳的两种方法之列。然而，对于 IBBI [16], [60] 数据库中的图像，所提方法仍优于大多数当前最先进的方法，但性能略逊于 BRISQUE [29] 和 GMLoG [33] 方法。与 H. Liu [16] 方法（该方法主要针对 IBBI [16], [60] 数据集开发，是一种质量评估方法）相比，所提方法表现更优。值得注意的是，[16] 的结果引自其原论文。

图9(a)-(f)展示了在两种失真和三个所考虑的数据库上，预测的差异平均意见分数与主观差异平均意见分数的六个散点图，进一步证明了所提方法的预测准确度。需要指出的是，在理想条件下，性能最佳的模型应具有高度线性且紧密聚集的散点分布。从图9的散点图中可以看出，在三个数据库中，两种失真均呈现出线性性能，但CSIQ [5] 数据库中的高斯模糊失真除外。这可能是由于失真对颜色分量的影响，因为我们仅考虑了图像的灰度版本。此外，如图9所示，LIVE [52] 和TID2013 [40] 数据库中JPEG和高斯模糊类型图像的散点高度聚集，这与表3和表4中分别列出的SROCC（0.979，0.972）、（0.958，0.951）以及皮尔逊线性相关系数（0.986，0.971）、（0.966，0.953）一致。所提方法在所有性能指标上均表现出优越性能，因此，对于独立或联合受到JPEG和高斯模糊伪影影响的图像，它是实现准确质量评估的优选算法。

4.3 跨数据库性能：

在上一节中，仅展示了特定数据库和特定失真的结果。为了评估所提方法在不受特定数据库影响时的性能，并考察其泛化能力，本文通过在具有相似失真的不同数据库上训练和测试各方法，进行了跨数据库验证。这确保了从一个数据库中学到的内容并非仅适用于该数据库，也可用于其他数据库。

在本研究中，使用 LIVE [52] 数据库中的图像进行训练，并使用 CSIQ [5] 和 TID2013 [40] 进行测试。由此得到两个跨数据库组合（训练-测试），分别记为 LIVE:CSIQ 和 LIVE:TID2013；各跨数据集及各失真对应的结果（SROCC 和皮尔逊线性相关系数）列于表7。对于失真特定方法，表6仅列出了 SROCC。可以看出，所提方法在两个数据库的两种失真上均表现最佳。在 LIVE:CSIQ 的交叉验证中，BJLC [39] 针对 JPEG 类型失真的 SROCC 值略优于所提方法；而对于 LIVE:TID2013 中的 JPEG压缩图像，所提方法的 SROCC 结果更好。比较 LIVE:TID2013 的 SROCC 值时，Zhu[24] 在 JPEG 类型失真上的表现也更优，但在基于 LIVE:CSIQ 的交叉验证中并非如此。对于白噪声类型的失真，所提方法在大多数情况下也优于其他方法，表现出良好的相关性得分。在单个数据库上的表现方面，BJLC [39] 始终表现良好；但在跨数据库验证中，其相关性得分显著较低，因此是一个高度依赖数据库和训练的模型。显然，所提模型表现突出，是一种鲁棒性强且数据库无关的模型。

© <2021>。此稿件版本依据 CC-BY-NC-ND 4.0 许可向公众开放

4.4 多失真数据库上的性能分析：

表 8 JPEG、高斯模糊和白噪声组合下多种失真图像的 SROCC 和 KROCC 对比。最佳的两项性能以粗体标出。

	LIVEMD		IVL MD		MDID2013		MDID	
	SROCC	KROCC	SROCC	KROCC	SROCC	KROCC	SROCC	KROCC
GWH-GLBP [44]	0.944	0.805	0.889	0.717	0.908	0.719	**0.892**	**0.708**
Miao [45]	**0.955**	**0.810**	**0.941**	**0.798**	**0.921**	**0.744**	**0.845**	**0.661**
NIQE [41]	0.858	0.726	0.837	0.644	0.614	0.456	0.657	0.467
IL-NIQE [42]	0.900	0.763	0.872	0.691	0.707	0.538	0.693	0.493
BLIINDS2 [32]	0.887	0.753	0.803	0.617	0.808	0.623	0.773	0.568
BRISQUE [29]	0.912	0.782	0.815	0.621	0.819	0.642	0.766	0.559
GMLoG [33]	0.833	0.704	0.901	0.731	0.838	0.657	0.771	0.568
NFERM [34]	0.898	0.759	0.871	0.688	0.851	0.674	0.803	0.603
SISBLM [46]	0.907	0.771	0.848	0.672	0.894	0.686	0.744	0.538
ZHOU [47]	0.943	-	-	-	0.907	-	-	-
所提方法	**0.957**	**0.813**	**0.950**	**0.816**	**0.912**	**0.727**	0.842	0.648

表9 JPEG、高斯模糊和白噪声组合下多种失真图像的皮尔逊线性相关系数和均方根误差比较。最佳的两项结果以粗体标示

	LIVEMD		IVL MD		MDID2013		MDID	
	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE	PLCC	RMSE
GWH-GLBP [44]	0.948	5.873	0.916	6.736	0.895	0.025	**0.897**	**1.048**
Miao [45]	**0.959**	**5.484**	**0.941**	**6.378**	0.910	0.024	0.828	1.618
NIQE [41]	0.887	8.933	0.837	7.256	0.648	0.037	0.678	2.766
IL-NIQE [42]	0.914	7.536	0.871	7.018	0.709	0.034	0.733	2.217
BLIINDS2 [32]	0.904	7.864	0.803	7.884	0.842	0.028	0.785	2.015
BRISQUE [29]	0.928	7.141	0.815	7.697	0.833	0.029	0.781	2.059
GMLoG [33]	0.872	9.164	0.901	6.841	0.830	0.028	0.791	1.956
NFERM [34]	0.917	7.459	0.872	6.943	0.875	0.027	0.811	1.873
SISBLM [46]	0.925	7.194	0.822	7.587	0.885	0.025	0.809	1.985
ZHOU [47]	0.951	5.747	-	-	**0.919**	**0.018**	-	-
所提方法	**0.954**	**5.739**	**0.976**	**6.150**	**0.916**	**0.021**	**0.839**	**1.611**

为了证明所提方法也能处理具有多重失真的图像，我们还在四个多重失真数据集 LIVEMD [56], IVL MD [57], MDID2013 [46] 和 MDID [59]中的图像上进行了测试；这些数据集中的原始图像均受到 JPEG压缩、模糊和白噪声的组合失真。机器学习和非机器学习方法的训练与测试流程与第4.1节所述相同。我们将其性能与先进的单失真（通用）方法 BLIINDS2 [32], BRISQUE [29], GMLoG [33],NFERM [34], NIQE [41] 和 IL-NIQE [42]，以及多失真方法 GWH-GLBP [44], Miao [45], 和 SISBLIM [46] 进行比较，并使用这些方法的源代码（在 LIVEMD、IVL MD、MDID2013 和 MDID 上）；结果列于表8和表9。对于 Zhou [47], 提出的方法，LIVEMD 和 MDID2013 的得分直接取自原论文。值得注意的是，著名的通用方法在评估图像中的多重失真时表现相对较差且能力有限。从表8所列的相关性得分，即 SROCC 和 KROCC，可以分析得出，除在 MDID 上的表现略逊于 Miao [45] 和 GWH-GLBP [44] 外，所提技术在所有数据库上处理多重失真组合时均表现良好。但与此同时，在比较皮尔逊线性相关系数得分时，它优于 Miao [45] 。该方法还使用

© <2021>。本手稿版本依据 CC-BY-NC-ND 4.0 许可公开

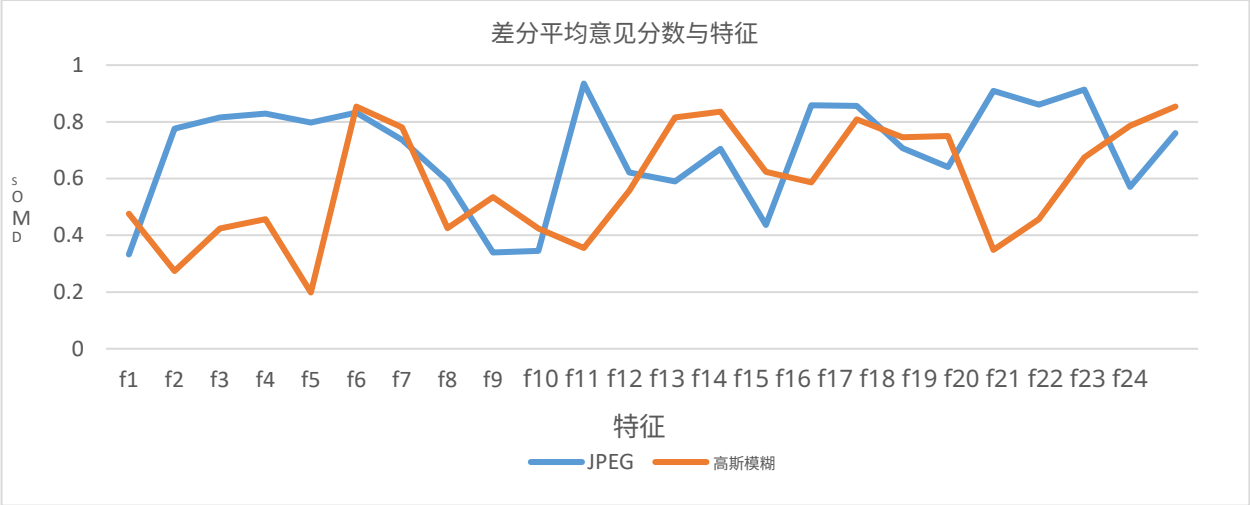

图 10 两种失真类型下，特征与人类感知的 SROCC 分数之间的相关性

表 10 LIVEMD数据库中不同方法在 PLCC、SROCC 和 RMSE 指标上的统计分析。1 2 3 4 5 6 7

1	___	___	1__	111	111	000	000
2	___	___	___	111	111	000	000
3	0__	___	___	111	111	000	000
4	000	000	000	___	111	000	000
5	000	000	000	000	___	000	000
6	111	111	111	111	111	___	__1
7	111	111	111	111	111	__0	___

四个多失真数据库上的皮尔逊线性相关系数和均方根误差得分列于表9中。对于几乎所有子集，所提方法的表现均优于 GWH-GLBP [44] 和 Zhou [47] 。为进一步研究，在 LIVEMD [56] 数据库上对这些方法进行了统计分析（t检验）。由于 NIQE [41] 和 IL-NIQE[42] 是不依赖主观意见的无学习方法，因此将其排除。采用 1 至 7 的数字分别表示方法 NFERM [34], 、GWH-GLBP [44], 、BRISQUE [29], 、BLIINDS2 [32], 、GMLoG [33], 、Miao [45] 和所提方法。符号“1”、“_”和“0”表示与对应列中的指标相比，所在行的方法显著更优、相近或更差（置信水平为 95%）。表10展示了这些统计分析结果，其中每个条目中的符号分别对应基于皮尔逊线性相关系数、SROCC 和均方根误差标准得出的结果。可以看出，除均方根误差标准下的 Miao [45] 外，所提方法在统计上优于所有基准技术。

4.5 不同特征的贡献:

为了清楚地了解特征与主观评分之间的关系，我们在 LIVE [52] 数据库上进行了实验，使用的数据集包含 GBLUR 和 JPEG 子集中的所有图像。每个特征都分别使用高斯过程回归学习器进行训练和测试，训练测试比例为 80-20 %，并重复迭代 1000 次，以获得所有可能的组合。图10分别展示了 JPEG 和 GBLUR 图像的中位数评分。从图中可以看出，基于零值的特征（即 f6、f11 和 f16）分别与人类感知 GBLUR 和 JPEG 图像的质量高度相关。同样，由于 f21 和 f22 是基于高频的特征，因此在 JPEG压缩图像上表现良好。本研究采用了两种并行方法：一种仅基于离散傅里叶变换，另一种则是在均值减除对比度归一化后进行离散傅里叶变换。为了说明这种组合在质量评估方面具有更好的适用性，我们列出了三种可能情形（即离散傅里叶变换、均值减除对比度归一化+离散傅里叶变换，以及

© <2021>。此手稿版本依据 CC-BY-NC-ND 4.0 许可公开提供

离散傅里叶变换与均值减除对比度归一化+离散傅里叶变换的组合）在表11中的 SROCC 和 PLCC 评分。很明显，在 LIVE [52], CSIQ [5] 和 TID2013 [40]这三个数据库中，离散傅里叶变换与均值减除对比度归一化+离散傅里叶变换的组合均取得了最高评分。

4.6 回归模型的选择：

为了证明采用高斯过程回归模型作为所提取特征的映射工具是合理的，我们对支持向量机（SVM）和高斯过程回归模型在两种不同验证比例（80-20 和 50-50）下获得的 SROCC 和皮尔逊线性相关系数得分进行了比较，如表12所示。显然，高斯过程回归模型取得了最佳性能，验证了其在该模型中的适用性。

表11　三种方法的 SROCC 和皮尔逊线性相关系数数值。最佳性能以粗体标出　.

数据库		LIVE		TID2013 C		*SIQ*	
失真类型		JPEG	高斯模糊	JPEG	高斯模糊	JPEG	高斯模糊
I. 均值减除对比度归一化+离散傅里叶变换	SROCC	0.964	0.958	0.904	0.911	0.903	0.918
II. 离散傅里叶变换	SROCC	0.962	0.943	0.924	0.914	0.890	0.912
I 与 II 的组合	SROCC	**0.979**	**0.972**	**0.958**	**0.951**	**0.948**	**0.944**
I. 均值减除对比度归一化+离散傅里叶变换	PLCC	0.971	0.965	0.918	0.922	0.926	0.906
II. 离散傅里叶变换	PLCC	0.965	0.938	0.960	0.934	0.918	0.924
I 与 II 的组合	PLCC	**0.986**	**0.971**	**0.966**	**0.953**	**0.949**	**0.941**

表12 SROCC 和 PLCC 值对比 基于所用技术的图像质量评价指标。最佳表现　ce
使用 b 高亮显示 旧版人脸。

数据库		*LIVE*		TID2013		*CSIQ*		交叉验证比例
失真类型		JPEG	高斯模糊	JPEG	高斯模糊	JPEG	高斯模糊	
GPR	SROCC	**0.979**	**0.972**	**0.958**	**0.951**	**0.948**	**0.944**	80-20
SVM	SROCC	0.961	0.965	0.937	0.924	0.917	0.910	80-20
GPR	SROCC	0.975	0.962	0.953	0.933	0.926	0.928	50-50
SVM	SROCC	0.941	0.953	0.927	0.913	0.910	0.904	50-50
GPR	PLCC	**0.986**	**0.971**	**0.966**	**0.953**	**0.949**	**0.941**	80-20
SVM	PLCC	0.958	0.947	0.939	0.918	0.903	0.926	80-20
GPR	PLCC	0.974	0.959	0.953	0.942	0.933	0.927	50-50
SVM	PLCC	0.932	0.941	0.929	0.899	0.893	0.920	50-50

表13 不同方法的运行时间对比（秒）方法 DIIVINE

	[31]	BRISQUE [29]	NFERM [34]	GMLoG [33]	FRIQUEE [37]	Miao [45]	所提方法
Time	38.6	0.41	82.3	0.19	40.2	5.2	4.81

4.7 运行时间比较：

为了测量并比较运行时间复杂度，所提方法以及一些现有方法，例如 DIVINE [31], BRISQUE [29], NFERM [34], GMLoG [33], FRIQUEE [37] 和 Miao [45]，均在一台配备英特尔（R）至强（R）处理器的计算机上运行，该处理器主频为 2.13 GHz，配备 20 GB 可用内存、64 位操作系统和 1 TB 硬盘。每种算法的运行时间如表13所示。这些结果针对从 LIVE 数据库 [52] 中选取的 100 幅不同尺寸图像组成的图像组，其包含高斯模糊和 JPEG 类型失真（50%-50%）。

由表13可见，所提算法在运行时间复杂度方面表现中等。与 BRISQUE [29] 和 GMLoG [33] 相比，其时间复杂度表现略逊一筹，但优于 DIIVINE [31], NFERM [34], FRIQUEE [37] 和 Miao [45]。不过，值得注意的是，尽管所提方法的运行时间复杂度不及 GMLoG [33] 和 BRISQUE [29]，但其准确性高于这两种模型。因此，该模型综合考虑性能和准确性（SROCC、皮尔逊线性相关系数、KROCC 和均方根误差）以及时间复杂度后，成为评估单独或联合受到块效应、模糊和噪声失真的图像的最适宜算法。

© <2021>。此手稿版本依据 CC-BY-NC-ND 4.0 许可提供

5. 结论

本研究提出了一种新方法，用于评估单独或联合受到块效应、模糊或噪声破坏的图像质量。该模型以人类感知为启发，采用变换域离散傅里叶变换和 MSCN 系数，提取最适合人类视觉系统的自然场景属性。所提模型能够有效、准确地评估存在模糊、块效应和噪声组合失真的图像质量。所提方法速度较快且准确度高。这是一种全盲图像质量评估方法，能够估计同时受到块效应、模糊和噪声失真的图像的整体质量，并适用于评估通过带限低通信道接收的高度 JPEG/MPEG/H.26x 压缩图像/视频的质量。在现代电子设备之间交换图像时，经常会遇到这种情况。未来，我们计划将该方法扩展到视频质量评估（视频质量评估），用于对高效视频编码和多功能视频编码编码的视频进行实时质量监测。

参考文献

[1] Z. Wang, 和 A. C. Bovik.《降参考和无参考图像质量评估》。IEEE信号处理杂志 28，第 6 期（2011）： 29-40。

https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_2022_Forecast_Highlights.pdf[3] M. T. Qadri、K. T. Tan 和 M. Ghanbari。《用于图像质量

Gh 。《分类图像质量（CSIQ）数据库》。（2010）http://vision.okstate.edu/csiq（2010）。[6] P. Ma 》。载于《国际图像处理会议论文集》 第 3 卷，第 III-III 页。IEEE，2002。[7] P. Marziliano、F. Dufaux、S. Winkler 和 T. Ebrahimi。《感知模糊和振铃度量：在 JPEG2000 中的应用》。《信号处理：图像通信》19，第 2 期（2004）： 163-172。[8] EePing On 2003 N. D. Narvekar 和 L. J 图像模糊度量》。《Optik》127，第 14 期（2016

mi 和 A. C. Kot。《一种基于最大局部变化的快 Bo Dong 和 Z. Jiang。《基于结构相似性指数的无参考模糊图像质量评估》。《应用科学》8，第 10 期（2018 nesran、W. A. Jassim、Y.-Poh Yu 和 K. N. Ngan。《使用精确定位矩形模糊核对模糊图像进行无参考模糊质量评估》。《高等表征研究所期刊》258，第 17 期（2016）

Ingrid。《一种高效的无参考感知模糊度量》。载于

00 年国际图像处理会议论文集（目录号 00CH37101）》，第 3 卷，第

的无参考测量》。载于《2001 年 IEEE 国际声学、语音和信号处理会议论文集（目录号 01CH37221）》，第 3 卷，第 1725 A. C. Bovik。《JPEG 压缩图像的无参考感知质量评估》。载于《图像处理国际会议论文集》，第 1 卷。IEEE，2002

© <2021>。本手稿版本依据 CC-BY-NC-ND 4.0 许可提供

[20] C.-Su Park、J.-Hyung Kim 和 S.-Jea Ko。《在像素域和 DCT 域中快速无参考测量块效应》。《数学成像与视觉 [illegible] 种无参考块效应度量》。载于《泛太平洋多媒体会议》，第 112-123 页。Springer，柏林，海德堡，2010。[22] S. A. Golestaneh 和 D. M. Chandler。《基于质量相关性图的 IPEG 图像无参考质量评估》。IEEE 信号处理快报 21，第 2 期（2013）：155-158。[23] Y. Zhan 和 [illegible] IPEG 图像质量评估》。IEEE 信号处理快报 24，第 6 期（2017）：[illegible] 页。IEEE，2017。[25] D. Zoran 和 Y. Weiss。《自然图像中的尺度不 [illegible] 像处理会议》，第 1857-1860 页。IEEE，2011。[27] G. Zhai、A. Kaup、J. Wang 和 X. Yang。《一种含噪图像无参考质量评估的双模型方法》。《亚太信号与信息处理学会信号与信息处理汇刊》4（2015）。[28] C. Deng、S. Wang、A. C. Bovik、G.-Bin Huang 和 B. Zhao。《利用子带峰度进行含噪图像无参考质量评估》。IEEE 控制论汇刊 50，第 3 期（2019）：1146-1156。[29] A. Mittal、A. K. Moorthy 和 A. C. Bovik。《空间域无参考图像质量评估》。IEEE 图像处理汇刊 21，第 12 期（2012）：4695-4708。[30] M. A. Saad、A. C. Bovik 和 Christophe Charrier。《基于 DCT 统计的无参考图像质量指数》。IEEE 信号处理快报 17，第 6 期（2010）：583-586。[31] A. K. Moorthy 和 A. C. Bovik。《无参考图像质量评估：从自然场景统计到感知质量》。IEEE 图像处理汇刊 20，第 12 期（2011）：3350-3364。[32] [illegible] W. Xue、X. Mou、L. Zhang、A. C. Bovik 和 X. Feng。《利用梯度幅值与拉普拉斯特征的联合统计进行无参考图像质量评估》。IEEE 图像处理汇刊 23，第 11 期（2014）：4850-4862。[34] K. Gu、G. Zhai、X. Yang 和 W. Zhang。《利用自由能原理进行无参考图像质量评估》。IEEE 多媒体汇刊 17，第 1 期（2014）：50 [illegible] 和 A. C. Bovik。《基于空间熵和频谱熵的无参考图像质量 [illegible] （2020）：115680。[37] D. Ghadiyaram 和 [illegible] 框架》。载于《2012 年 IEEE 计算机视觉与模式识别会议》，第 1098-1105 页。IEEE，2012。[39] Q. Li、W. Lin、K. Gu、Y. Zhang 和 Y. Fang。《基于联合对数对比度统计的无参考图像质量评估》。《神经计算》331（2019）：189-198。[40] N. Ponomarenko、L. Lin、A. Jeremeiev、V. Lukin 和 [illegible] 据库 TID2013：特点、结果与展望》。《信 [illegible] 209-212。[42] L. Zhang、L. Zhang 和 A. C. [illegible] 质量无参考度量的分析》。美国国家电信和信息管理局技术备忘录 20-547。[44] Q. Li、W. Lin 和 Y. Fang。《梯度域中多重失真图像的无参考质量评估》。IEEE 信号处理快报 23，第 4 期（2016）：541-545。[45] X. Miao、H. Chu、H. Liu、Y. Yang 和 X. Li。《基于相位一致性与梯度幅值的多重失真图像质量评估》。《信号处理：图像通信》79（2019）：54-62。[46] Ke Gu、Guangtao Zhai、Xiaokang Yang 和 Wenjun Zhang。《针对单一及多重失真图像的混合型无参考质量指标》。IEEE 广播汇刊 60，第 3 期（ 2014 ）：第 555-567 页。

© <2021>。本手稿版本依据 CC-BY-NC-ND 4.0 许可提供

[47] Wujie Zhou、Lu Yu、Yaguan Qian、Weiwei Qiu、Yang Zhou 和 Ting Luo。“基于单基二进制编码的多重失真图像深度盲质量评估器。”《视觉通信与图像表示》60（2019）：305-311。[48] 视频质量专家组（VQEG），2012年7月，http://www.vqeg. org/。[49] E. Wyckens、M. Leszczuk 和 S. Borer。“通过关键指标（MOAVI）监测视听质量。”载于第四届新兴网络智能国际会议 EMERGING 2012。2012。[50] M. Leszczuk、M. Hanusiak、I. Blanco、A. Dziech、J. Derkacz、E. Wyckens 和 S. Borer。“用于视听质量监测的关键指标。”载于2014年第22届信号处理与通信应用会议（SIU），第2301-2305页。IEEE，2014。[51] C. E. Rasmussen。“机器学习中的高斯过程。”载于机器学习暑期学校，第63-71页。Springer，柏林，海德堡，2003。[52] H. R. Sheikh，“LIVE图像质量评估数据库第二版。”

）：517-548。[54] M. Carandini、D. J. Heeger 和 J. A. Movshon。“猕猴初级视觉皮层简单

pe。

IEEE，2019。[56] D. Jayaraman、A. Mittal、A. K. Moorthy 和 A. C. Bovik。“多重失真图像的客观质量评估。”载于2012年第四十六届阿西洛马信号、系统与计算机会议论文集（ASILOMAR），第1693-1697页。IEEE，2012。[57] S. Corchs、F. Gasparini。“用于图像质量评估的多失真数据库。”载于计算色彩成像国际研讨会，第95-104页。Springer，Cham，2017。[58] S. Corchs、F. Gasparini、R. Schettini。“含噪图像-JPEG压缩：主观与客观图像质量评估。”IS&T/SPIE电子成像，国际光学与光子学学会。2014年1月7日；90。[59] W. Sun、F. Zhou 和 Q. Liao，“MDID：用于图像质量评估的多重失真图像数据库”，《模式识别》，第61卷，第153–168页，2017年1月。[60] http://ivc.univ-nantes.fr/en/databases/Intentionally_Blurred_Background_Images/[61] 视频质量专家组。《视频质量专家组关于视频质量评估客观模型验证的最终报告，第二阶段（FR_TV2）》，2003。ftp://ftp.its.bldrdoc.gov/dist/ituvidq/Boulder _ VQEG _ jan _ 04/ VQEG _ PhaseII _ FRTV _ Final _ Report _ SG9060E.doc

© <2021>。本手稿版本依据 CC-BY-NC-ND 4.0 许可发布：