

基于三维与二维特征融合的 无参考点云质量评价

刘太伟, 郁梅*, 屠仁伟

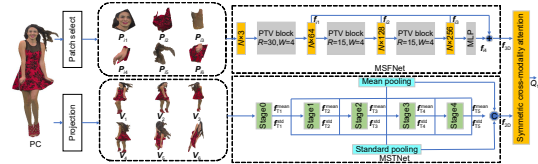
宁波大学信息科学与工程学院, 浙江 宁波 315211

摘要: 随着点云数据在虚拟现实、计算机视觉、机器人等领域中的广泛应用, 点云获取与处理中的失真评价正成为一个重要的研究问题。考虑到点云三维信息对几何失真敏感、点云二维投影图包含丰富的纹理和语义信息, 提出一种基于三维与二维特征融合的无参考点云质量评价方法, 以有效结合点云的三维与二维特征信息, 提高点云质量评价的准确性。对于三维特征提取, 先对点云进行最远点采样, 以选取的点为中心生成互不重叠的点云子模型, 尽可能地覆盖整个点云模型, 利用多尺度三维特征提取网络提取体素和点的特征。对于二维特征提取, 先对点云进行正交6面投影, 再通过多尺度二维特征提取网络提取纹理和语义信息。最后, 考虑到人类视觉系统处理不同类型信息时会存在分割处理和交织融合的过程, 设计一个对称跨模态注意模块融合三维和二维特征。在5个公开点云质量评价数据库上的实验结果显示, 所提方法的皮尔逊线性相关系数 (Pearson's linear correlation coefficient, PLCC) 分别达到 0.9203、0.9463、0.9125、0.916 和 0.921, 表明与现有的代表性点云质量评价方法相比, 所提方法更优。

关键词: 点云质量评价; 三维特征; 二维特征; 对称跨模态注意模块

中图分类号: TP394.1

文献标志码: A



刘太伟, 郁梅, 屠仁伟. 基于三维与二维特征融合的无参考点云质量评价 [J]. 光电工程, 2025, 52(4): 250001

Liu T W, Yu M, Tu R W. No-reference point cloud quality assessment based on fusion of 3D and 2D features[J]. *Opto-Electron Eng*, 2025, 52(4): 250001

No-reference point cloud quality assessment based on fusion of 3D and 2D features

Liu Taiwei, Yu Mei*, Tu Renwei

Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China

Abstract: With the wide application of point clouds in virtual reality, computer vision, robotics and other fields, the assessment of distortions resulted from point cloud acquisition and processing is becoming an important research topic. Considering that the three-dimensional information of point clouds is sensitive to geometric distortion and the two-dimensional projection of point clouds contains rich texture and semantic information, a no-reference point cloud quality assessment method based on the fusion of three-dimensional and two-dimensional features is proposed to effectively combine the three-dimensional and two-dimensional feature information of point cloud and improve the accuracy of point cloud quality assessment. For 3D feature extraction, the farthest point sampling is firstly implemented on the point cloud, and then the non-overlapping point cloud sub-models centered on the selected points are generated, to cover the whole point cloud model as much as possible and use a multi-scale 3D

收稿日期: 2025-01-01; 修回日期: 2025-02-26; 录用日期: 2025-02-28

基金项目: 国家自然科学基金 (62071266)

*通信作者: 郁梅, yumei@nbu.edu.cn。

版权所有©2025 中国科学院光电技术研究所

feature extraction network to extract the features of voxels and points. For 2D feature extraction, the point cloud is first projected with orthogonal hexahedron projection, and then the texture and semantic information are extracted by a multi-scale 2D feature extraction network. Finally, considering the process of segmentation and interweaving fusion that occurs when the human visual system processes different types of information, a symmetric cross-modal attention module is designed to integrate 3D and 2D features. The experimental results on five public point cloud quality assessment datasets show that the Pearson's linear correlation coefficient (PLCC) of the proposed method reaches 0.9203, 0.9463, 0.9125, 0.9164 and 0.9209 respectively, indicating that the proposed method has advanced performance compared with the existing representative point cloud quality assessment methods.

Keywords: point cloud quality assessment; 3D feature; 2D feature; symmetric cross-modal attention module

1 引言

点云 (Point cloud, PC) 是空间中大量三维离散点的集合^[1]。点云数据从获取到应用的过程包括采集、压缩、传输、渲染等步骤, 其间都会不可避免地引入失真, 从而降低点云的视觉质量^[2]。如何准确评价彩色点云的视觉质量是优化点云处理算法、提升用户视觉体验质量的关键。因此, 需要开发可靠且符合人眼视觉感知的客观点云质量评价 (point cloud quality assessment, PCQA) 方法来优化点云处理过程, 提高用户的视觉体验质量。

根据参考点云信息的可用性可以将客观点云质量评价方法分为全参考 (full-reference, FR)、半参考 (reduced-reference, RR) 和无参考 (no-reference, NR) 方式, 这三类方法又可分为基于传统特征和基于深度学习的方法。基于传统特征的点云质量评价方法有点对点^[3]、点对面^[4]、点对网格^[5]、面对面^[6]和基于改进的峰值信噪比^[7]、广义 Hausdorff 距离^[8]和 Mahalanobis 距离^[9]等。Meynet 等^[10]考虑到曲率、曲面近似等特征也可以用于衡量点云失真, 提出了一种基于局部邻域高斯加权曲率统计的点云质量评价方法 (PC-MSDM)。这些方法仅侧重于对参考点云几何失真进行度量, 因此其对于彩色点云质量的评价性能有待提高。除了对点云进行几何特征提取, 颜色等更加符合人眼视觉感知的特征被引入点云客观质量评价, 主要方法有 PCQM^[11]、GraphSIM^[12]、PointSSIM^[13]、BQE-CVP^[14]、IW-SSIM^[15]、MFPCQA^[16]、PRPCQA^[17]等。PCQM 通过将点云几何和颜色特征进行最优线性加权组合评价点云质量。GraphSIM 通过图相似度和颜色梯度评价点云质量。PointSSIM 通过比较局部拓扑结构和颜色分布评价点云质量。BQE-CVP 结合视觉掩蔽效应对点云进行区域分割, 实现了基于视觉感

知的点云无参考质量评价。IW-SSIM 考虑了注意力机制和信息内容加权。MFPCQA 提取了点云多层次特征, 并对其排序和组合整合为高级特征, 来增强模型的泛化能力。为了解决点云评价过程中缺少参考点云的问题, PRPCQA 首先对失真点云进行泊松曲面重构, 构造了伪参考点云, 再结合二维和三维特征进行质量评价。

随着计算智能 (人工智能) 的发展, 现有主流的点云质量评价方法以深度学习方法为主。由于点云数据量巨大, 无法直接对点云处理, 早期基于深度学习的方法通常先用不同的投影方式将 3D 点云或其相关特征投影到 2D 规则网格中, 主要方法有 PQA-Net^[18]、IT-PCQA^[19]、PM-BVQA^[20]等。PQA-Net 通过一个多视图联合特征提取与融合模块提取投影特征, 利用点云失真类型标签预训练模型初始参数来评价点云质量。IT-PCQA 为了扩大训练集, 将自然图像作为源域, 点云投影作为目标域, 并通过无监督对抗域自适应推断点云质量。为了降低点云正交投影的信息损耗, PM-BVQA 利用点云的 VPCC (video based point cloud compression) 投影, 设计了一个多尺度特征融合网络, 联合颜色、几何特征评价点云质量。Tu 等^[21]在利用 VPCC 投影的基础上, 进一步考虑了点云的显著性等特征; 后续为了能够直接对点云进行处理, 对点云分块输入网络。Tliba 等^[22]先将点云划分为 3D Patch, 利用双流图卷积神经网络分别提取几何信息和颜色信息, 并用交叉注意机制模块融合两路信息来度量点云质量。

上述点云质量评价方法仅使用点云三维特征或二维特征, 没有考虑点云不同模态特征间的互补。为了充分利用点云不同模态特征, 使点云客观质量评价模型与主观评价相一致, 本文提出了一种基于三维与二维特征融合的无参考点云质量评价方法, 其主要贡献

可以总结为以下三点:

1) 为了更好地感知点云几何失真, 设计了一个点云多尺度三维特征提取网络, 提取基于体素和基于点的三维特征, 基于点的分支可以提取几何失真特征, 基于体素的分支可以提取语义特征, 两者互补, 提高了点云质量评价的准确性。

2) 点云二维投影图与人眼视觉之间存在一定的感知关系。为了提取投影图中丰富的内容感知特征, 设计了一个多尺度二维特征提取网络, 提取纹理和语义信息, 以更好地度量点云颜色失真, 实现与三维特征的互补。

3) 为了解决人类视觉系统处理不同类型信息时会存在分割处理和交织融合的过程, 设计了一个对称跨模态注意模块 (symmetric cross-modality attention, SCMA), 探究了二维特征和三维特征之间的相互作用, 并在特征融合过程中加入了多头注意力模块, 使其更符合人眼感知特性。

2 所提方法

所提方法共有两个分支, 其总体框图如图 1 所示, 其中 f_s^s 表示点云子模型 P_i 第 s 尺度基于点和基于体素的特征 ($s=1,2,3,4$), R 、 W 分别表示体素分辨率大小和 3D 窗口大小, Q_1 表示网络预测的点云模型质量分数。在三维特征提取分支中, 首先对点云模型进行预处理, 即对点云进行最远点采样, 以采样点为中心将其分割成固定数量、固定点数并且不重叠的点云子模型; 然后将点云子模型输入多尺度三维特征提取网络提取每个点云子模型的基于体素和基于点的多尺度三维特征。在二维特征提取分支中, 首先对点云模型进行正交投影, 得到 6 个视点的点云投影图, 然后将

其输入改进的二维多尺度特征提取网络中, 提取投影图丰富的二维内容感知特征。最后, 通过一个对称跨模态注意模块进行特征融合得到点云质量感知特征。

2.1 Patch 选取

点云数据量巨大, 直接将其输入到神经网络提取三维特征存在内存不足的问题。因此, 可将点云划分为更小的 3D 点云子模型进行处理。在给定同一个点云模型的情况下, 不同失真程度、失真类型的点云子模型之间的比较应对应点云的相同区域。然而, 不同失真程度、失真类型的点云可能有不同数量的点, 例如点云下采样、G-PCC Octree 编码压缩等失真可能会减少点的数量。而通过固定点云子模型的中心点, 可以保证点云子模型之间的对应关系。为了使生成的点云子模型尽可能覆盖整个点云模型, 采用最远点采样 (farthest point sampling, FPS) 策略^[23], 从输入点云模型中采样固定数量的点, 这些采样点被用作生成点云子模型的中心点。假定对每一个点云模型使用 FPS 从点云中选取 N 个采样点, 将每个点云子模型 $P_i (1 \leq i \leq m)$ 的点数固定为 n' , 可以得到 $m (m \leq N)$ 个不重叠点云子模型, 表示为

$$m = \frac{N \times r}{n'} \quad (1)$$

式中: r 为控制点云子模型间重叠比例的超参数。接下来, 对每个中心点应用 K 最近邻 (K-nearest neighbor, KNN) 算法来选取其 $n'-1$ 个最近邻点。

2.2 三维特征提取

点云的三维特征能更好地反映点云的几何失真, 为了提取点云子模型的三维特征, 设计了一个改进的多尺度三维特征提取网络 (3D multi-scale feature extraction network, MSFNet), 该网络包含三个点-体

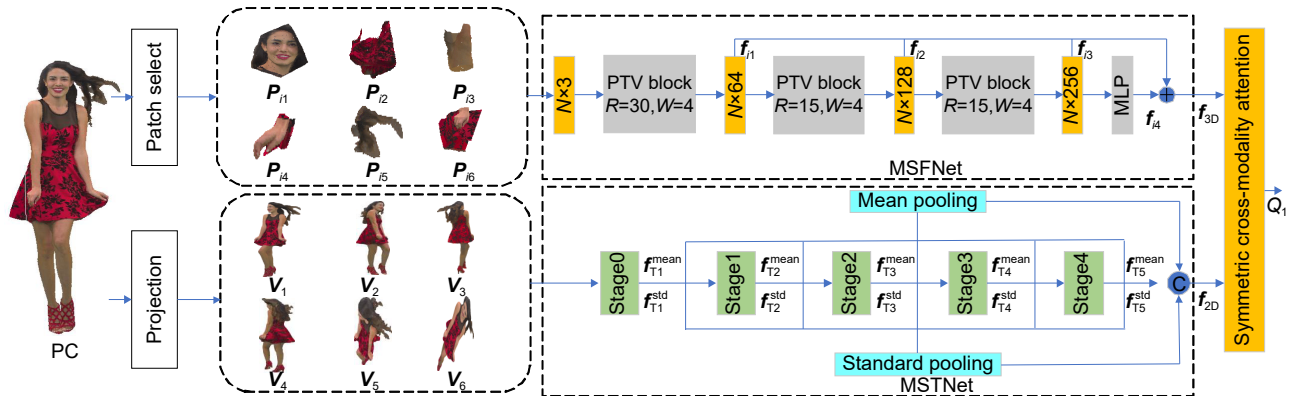


图 1 所提方法框图

Fig. 1 Framework of the proposed method

素转换器 (point-voxel transformer, PVT) 模块, 并通过多层感知器 (multi-layer perceptron, MLP) 生成输出特征。每个 PVT 模块有两个分支, 体素分支能够从空间体素中提取丰富的语义特征; 基于点的分支能够尽可能保留点云模型位置信息完整度, 避免位置信息的损失, 结构框图如图 2 所示。其中, 体素分支主要有三个步骤, 体素化、特征聚合和逆体素化。体素化和逆体素化操作主要是将输入的点云子模型 P_i 映射到一组新的体素特征 V , 并将体素特征转换回基于点的特征 F_{local} 。为了降低在体素化操作过程中的计算复杂度, 引入了稀疏窗口自注意模块 (sparse window self-attention module, S-SWA)。在基于点的分支中, 为了更好地聚合位置信息, 引入了相对注意 (relative attention, RA) 操作, 得到基于点的特征 F_{global} 。

对任意一个点云子模型 P_i , 将其输入多尺度三维特征提取网络 (MSFNet), 经过 PVT 模块后可得到基于点和基于体素的聚合特征, 然后通过多层感知器 (MLP) 提取每个点云子模型不同尺度的特征, 再通过最大值池化和 Concat 操作聚合每个点云子模型的三

维特征, $\text{Max}(\cdot)$ 表示最大值池化。上述过程可表示为

$$f_i^1 = \text{Max}(w_1(P_i)), \quad (2)$$

$$f_i^2 = \text{Max}(w_2(w_1(P_i))), \quad (3)$$

$$f_i^3 = \text{Max}(w_3(w_2(w_1(P_i)))), \quad (4)$$

$$f_i^4 = \text{Max}(w_4(w_3(w_2(w_1(P_i))))) , \quad (5)$$

式中: $f_i^1 \in \mathbf{R}^{64}$, $f_i^2 \in \mathbf{R}^{128}$, $f_i^3 \in \mathbf{R}^{256}$, $f_i^4 \in \mathbf{R}^{512}$; w_1 、 w_2 、 w_3 、 w_4 分别为 MLP 四个全连接层的共享权值。最后, 将得到的 4 个特征向量进行 Concat 拼接, 得到第 i 个点云子模型的三维多尺度特征向量 f_{3D} , 表示为

$$f_{2D} = [f_i^1, f_i^2, f_i^3, f_i^4] \in \mathbf{R}^{1408} . \quad (6)$$

2.3 二维特征提取

为了提取点云二维投影图丰富的纹理和语义信息, 设计了一个多尺度二维特征提取网络 (2D multi-scale feature extraction network, MSTNet) 从点云二维投影图中提取二维内容感知特征。该网络共有 5 级特征提取阶段, 每级特征提取网络结构如图 3 所示。为了扩充训练集, 首先在 ImageNet^[24] 上对网络模型预训练,

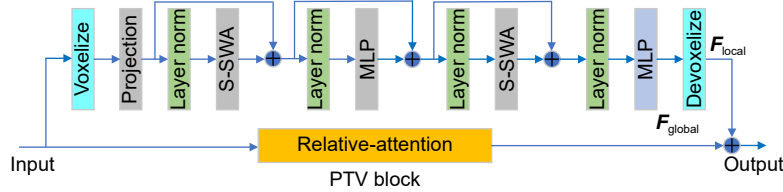


图 2 PVT 模块框图

Fig. 2 Framework of the PVT block

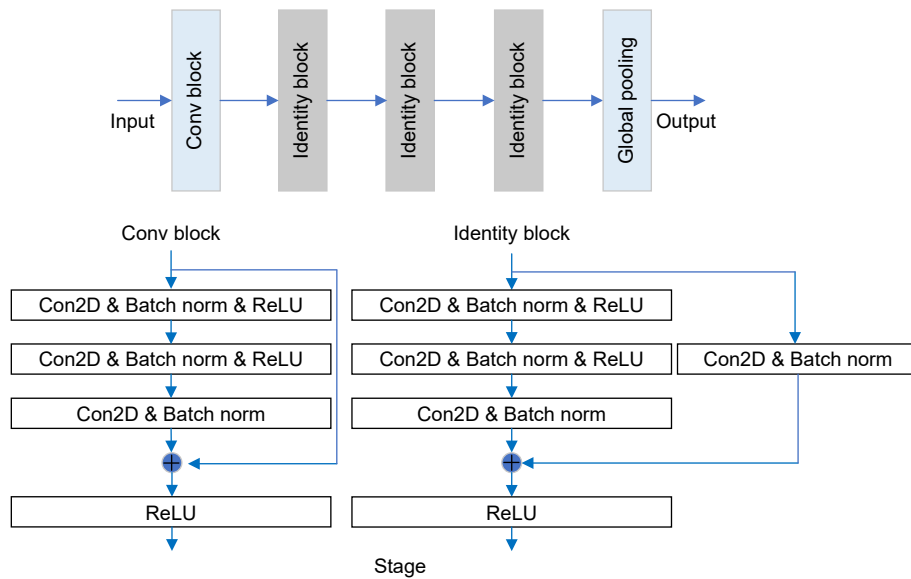


图 3 特征提取模块框图

Fig. 3 Framework of the feature extraction block

再在点云数据集微调网络参数。对点云模型进行 6 面投影, 可以获得 6 个视点的投影图, 令 V_n 表示投影图, 其中 $n=1,2,3,\dots,6$ 。随后, 利用所构建的预训练 MSTNet 模型提取每一视点投影图的内容感知特征图 F_n , 表示为

$$F_n = \text{MSTNet}(V_n). \quad (7)$$

采用空间全局平均池化操作去除不同视点投影图间的大量冗余信息。考虑到不同视点投影图在观察时也会存在一定的依赖关系, 采用空间全局标准差池化操作来保留不同视点投影图间的依赖信息。第 n 个视点投影图的全局平均池化特征 f_{Tn}^{mean} 和全局标准差池化特征 f_{Tn}^{std} 表示为

$$f_{Tn}^{\text{mean}} = \text{GP}_{\text{mean}}(F_n), \quad (8)$$

$$f_{Tn}^{\text{std}} = \text{GP}_{\text{std}}(F_n), \quad (9)$$

式中: $\text{GP}_{\text{mean}}(\cdot)$ 、 $\text{GP}_{\text{std}}(\cdot)$ 分别表示全局平均池化和全局标准池化。最后, 将特征向量进行 Concat 操作得到每个点云视点投影图的内容感知特征向量 f_{2D} , 表示为

$$f_{2D} = \text{Concat}(f_{Tn}^{\text{mean}}, f_{Tn}^{\text{std}}). \quad (10)$$

2.4 对称跨模态注意模块

为了更好地实现点云二维特征和三维特征的互补, 使用一个对称跨模态注意 (symmetric cross-modality attention, SCMA) 模块进行特征融合, SCMA 模块的结构框图如图 4 所示, 其中 Q 、 K 、 V 分别表示多头注意力机制中的查询 (query)、键 (key) 和值 (value)。

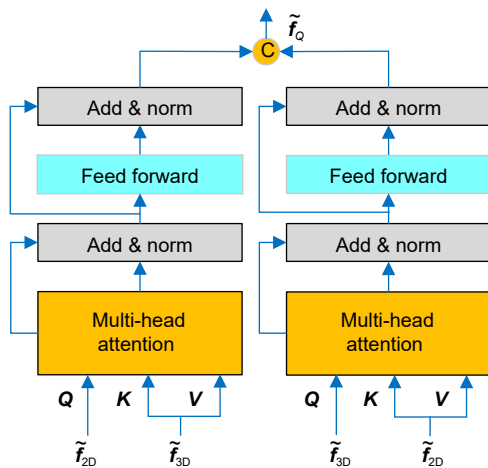


图 4 SCMA 模块框图

Fig. 4 Framework of the SCMA

给定点云模型的二维特征 f_{2D} 和三维特征 f_{3D} , 为了进行特征融合, 用线性映射调整到同一维度, 表

示为

$$\tilde{f}_{2D} = w_{2D} f_{2D}, \quad (11)$$

$$\tilde{f}_{3D} = w_{3D} f_{3D}, \quad (12)$$

式中: $\tilde{f}_{2D}$ 、 $\tilde{f}_{3D}$ 分别表示调整后的点云二维特征和三维特征; w_{2D} 、 w_{3D} 分别表示可学习的线性投影映射参数。令 $\Gamma(\cdot)$ 、 $\beta(\cdot)$ 、 h_u 分别表示多头注意力操作、注意力函数和第 u 个头, 特征融合的过程可表示为

$$\Gamma(Q, K, V) = (h_1 \oplus h_2 \oplus \dots \oplus h_n), \quad (13)$$

$$h_u = \beta(Q W_u^Q, K W_u^K, V W_u^V) |_{u=1}^n, \quad (14)$$

$$\beta(Q, K, V) = \text{softmax}(Q K^T / \sqrt{d}) V, \quad (15)$$

式中: $\oplus$ 表示逐元素相加操作; W_u^Q 、 W_u^K 、 W_u^V 表示可学习的线性映射参数; d 表示缩放因子。

令 $\psi(\cdot)$ 、 $\tilde{f}_Q$ 分别表示对称跨模态注意操作和最终的质量回归特征。对通过对称跨模态注意模块获得的多模态特征、二维特征和三维特征进行 Concat 操作, 得到最终的质量回归特征, 可表示为

$$\tilde{f}_Q = \tilde{f}_{2D} \oplus \tilde{f}_{3D} \oplus \psi(\tilde{f}_{2D}, \tilde{f}_{3D}). \quad (16)$$

最后, 通过两个全连接层可以将质量特征回归到预测的质量分数中, 损失函数 (Loss) 采用均方误差 (mean squared error, MSE), 表示为

$$\text{Loss} = \|Q_p - Q_i\|^2, \quad (17)$$

式中: Q_i 为失真点云平均主观质量评分 (mean opinion score, MOS); Q_p 为网络预测分数。

3 实验结果与分析

本节将首先介绍实验设置与环境, 包括数据库、实验环境和参数、质量评价标准。其次, 与现有的代表性点云质量评价方法进行比较, 验证所提方法的有效性和先进性。最后, 进行了消融实验, 验证所提方法各模块的有效性。

3.1 实验设置与环境

为了验证所提方法的有效性, 在 CPCD2.0^[25]、SJTU-PCQA^[26]、IRPC^[27]、ICIP2020^[28] 和 M-PCCD^[29] 彩色点云数据集上进行了性能评估。CPCD2.0 数据集包括 10 个参考点云与 360 个失真点云, 失真类型包含编码失真与高斯噪声下的几何失真、纹理失真以及双重失真。SJTU-PCQA 数据集包含 9 个参考点云与 378 个失真点云, 失真类型包括八叉树压缩、降尺度、高斯噪声失真, 以及降尺度与噪声的双重失真。

IRPC、ICIP2020 和 M-PCCD 数据集分别包含 6 个参考点云与 54 个失真点云、6 个参考点云与 90 个失真点云和 8 个参考点云与 232 个失真点云, 三个数据集失真类型都包含三种不同的编码失真。

实验环境如下: CPU 为 intel Xeon Silver 4210 2.20 GHz, RAM 为 32 GB, GPU 为 GeForce RTX 3090 24 G、Pycharm、PyTorch1.1.0、python3.6。对于批量优化的学习率使用基于梯度方差的 ADAM 方法自适应地控制每个参数, 参数设置为 ϵ 为 10^{-6} , learning rate 为 10^{-4} , epochs 为 300。采用 K-Fold 的训练方式来划分训练集和测试集, 对于 CPCD2.0、SJTU-PCQA、IRPC、CPCD2.0 子集、ICIP2020 和 M-PCCD 数据集分别采用 10-Fold、9-Fold、6-Fold、10-Fold、6-Fold 和 8-Fold, 遍历所有场景再取中值作为最终结果。

斯皮尔曼等级相关系数 (Spearman's rank-order correlation coefficient, SROCC)、肯德尔等级相关系数 (Kendall's rank-order correlation coefficient, KROCC)、PLCC 和均方根误差 (root mean square error, RMSE) 是评价图像/视频质量客观评价方法性能优劣的四个常用指标。SROCC 和 KROCC 表示预测的单调性, PLCC 和 RMSE 表示预测的准确性。更好的图像/视频质量客观评价方法应该具有更大的 SROCC/KROCC/PLCC 和更小的 RMSE。

3.2 总体性能对比

为了进一步验证所提方法的有效性, 分别在不同

数据集上与现有有代表性的点云质量评价方法进行了对比。在 CPCD2.0、SJTU-PCQA、IRPC 三个数据集上的实验结果如表 1 所示。由于文献 [15]、[16] 没有公开算法代码, 表 1 中的数据参考文中提供的指标。在 SJTU-PCQA、IRPC 数据集上, 基于传统特征的全参考质量评价方法 GraphSIM、PCQM 同时考虑几何和颜色特征, 性能明显更优于仅侧重于对点云几何失真度的全参考质量评价方法 P2point、P2plane。由于 CPCD2.0 点云数据集包含编码下的纹理失真、几何失真、纹理和几何双重失真, 导致 PCQM 算法在该数据库上的性能降低。考虑人眼感知特性的无参考点云质量评价方法 (如 BEQ-CVP、IW-SSIM) 缩小了与全参考方法的性能差距, 在 SJTU-PCQA 数据集上的性能甚至优于大部分全参考方法。为了解决点云数据量巨大的问题, 基于深度学习的方法主要从点云不同方法的投影图中提取特征来评价点云质量。PQA-Net 通过点云 6 面正交投影图提取特征, 并用失真类型标签预训练网络, 其性能优于大多数基于传统特征提取的方法。由于 VPPCQA 和 PRPCQA 主要对压缩编码失真点云进行质量评价, 为公平起见, 在包含不同编码失真的 CPCD2.0 子集、ICIP2020、M-PCCD 三个数据集进行对比实验, 其结果如表 2 所示。VPCC 投影可以得到纹理投影图和几何投影图, 减少了点云正交投影的信息损耗, VPPCQA 设计了双流网络分别提取 VPCC 投影图的纹理信息和几何信息, 并考虑了点云显著性特征, 说明人眼感知特性对点云质量评价的重要性。PRPCQA 通过泊松曲面重构构造

表 1 不同方法在不同数据库上的总体性能对比

Table 1 Comparison of overall performance of different methods on different datasets

Type	Metric	CPCD2.0				SJTU-PCQA				IRPC			
		PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE
Full-reference	P2point_MSE ^[3]	0.6784	0.5491	0.4142	0.8617	0.4721	0.4096	0.2857	2.1394	0.3357	0.3281	0.2146	0.9313
	P2plane_Hausdorff ^[4]	0.4061	0.3786	0.2663	1.0718	0.3752	0.4609	0.3354	2.4467	0.3925	0.2541	0.1975	0.9089
	P2plane_MSE ^[4]	0.6914	0.5692	0.4385	0.8474	0.5651	0.4956	0.3514	2.0022	0.4296	0.2564	0.1957	0.9089
	PC-MSDM ^[10]	0.6254	0.5321	0.3842	0.9152	0.4123	0.3241	0.2189	2.2110	0.2729	0.1519	0.1063	0.9515
	PointSSIM-G ^[13]	0.5343	0.5533	0.4238	0.9915	0.3860	0.3649	0.2792	2.2410	0.6183	0.5951	0.4693	0.7760
	PointSSIM-C ^[13]	0.7457	0.6891	0.4863	0.7814	0.4561	0.4185	0.3172	2.1598	0.6648	0.5638	0.4211	0.7376
	PCQM ^[11]	0.4813	0.3408	0.2615	1.0281	0.7771	0.7420	0.5624	1.5274	0.5611	0.5611	0.3033	0.8184
	GraphSIM ^[12]	0.8553	0.8296	0.6234	0.6077	0.8900	0.8800	-	1.1300	0.9400	0.7600	-	0.2100
No-reference	BEQ-CVP ^[14]	0.7950	0.7890	0.5983	0.7218	0.9192	0.8972	0.7343	0.9717	0.7265	0.7298	0.5427	0.6586
	IW-SSIM ^[15]	-	-	-	-	0.7949	0.7833	-	-	0.0911	0.1339	-	-
	MFPCQA ^[16]	-	-	-	-	0.8972	0.8894	-	0.6488	-	-	-	-
	Proposed	0.9203	0.8996	0.7494	0.4195	0.9463	0.9248	0.8231	0.3854	0.9125	0.8566	0.7018	0.4529

伪参考点云评价点云质量,说明点云参考信息对评价算法的重要性。但上述方法仅考虑点云的二维特征,没有考虑三维特征,而所提方法结合了点云三维特征和二维特征,进行了更全面的特征提取,同时也考虑点云的颜色、纹理、人眼感知等特征,所以性能优于大部分现有有代表性的点云质量评价方法。在ICIP2020点云数据集对比的两种方法是基于传统特征的,本文采用深度学习方法,由于数据集样本较少,算法性能低于PRPCQA算法。

3.3 模块设计与消融实验

为了验证所提方法各模块的有效性,在CPCD2.0彩色点云数据集上进行了消融实验,表3为不同特征的实验结果,包括仅提取点云二维特征,仅提取点云三维特征,以及提取联合特征三种不同的方式。从实验结果来看,由于点云投影图可以提供丰富的纹理和语义信息,从投影图提取特征的评价方法也有着不错的性能。虽然仅提取点云三维特征的方法性能欠佳,但其可以反映点云的几何失真,因而将其与二维特征相结合,起到相互补充、相互增强的作用,取得更好的性能指标。

表4给出了CPCD2.0点云数据集上所提方法采用不同特征融合方法得到的实验结果,验证其所提

方法性能的影响,其中“Add”为不同模态特征相加,“Concat”代表特征级联的方法,“SCMA”为采用对称跨模态注意模块进行特征融合。实验结果表明,使用对称跨模态注意模块进行特征融合的方法性能最好,主要原因是该模块可以更好地融合点云的二维和三维特征,使两种特征相互增强,并且在特征融合的过程中加入了多头注意力机制。

选取过多的点云子模型会增加计算复杂度,降低系统效率,选取过少则点云三维特征不足。表5给出了CPCD2.0点云数据集上点云子模型数量和每个子模型点数的不同组合策略对所提方法性能的影响。在训练策略相同和总采样点数固定的情况下,验证了几种不同的组合方法。实验结果显示,点云子模型个数与每个子模型包含点数最佳组合为6×2048。其原因在于分析数量过多的点云子模型可能会导致原点云中的信息冗余,影响有效点提供的信息,而数量过少的点云子模型又不能提供足够的信息。文中点云子模型的数量为6。

为了验证点云子模型中心点选取对所提方法性能的影响,表6给出了几种不同的采样策略在CPCD2.0点云数据集上的性能对比,包括随机采样、格点采样、几何采样和最远点采样。实验结果表明,用最远点采

表2 不同方法在不同数据库上的总体性能对比

Table 2 Comparison of overall performance of different methods on different datasets

Type	Metric	CPCD2.0 subset				ICIP2020				M-PCCD			
		PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE
No-reference	PRPCQA ^[17]	0.8591	0.8351	0.7014	0.5581	0.9310	0.9307	0.8814	0.3011	0.9144	0.9241	0.6857	0.5107
	VPPCQA ^[21]	0.8343	0.8460	0.6578	0.6046	0.9114	0.9264	0.8311	0.3965	0.9147	0.9322	0.6923	0.4903
	Proposed	0.9100	0.8842	0.7348	0.4371	0.9160	0.8827	0.8564	0.3672	0.9215	0.9400	0.7638	0.4720

表3 CPCD2.0数据集上不同特征的性能对比

Table 3 Performance comparison of different features on CPCD2.0 dataset

f_{2D}	f_{3D}	PLCC	SROCC	KROCC	RMSE
√		0.9096	0.8879	0.6792	0.4391
	√	0.5148	0.4754	0.5148	0.9596
√	√	0.9203	0.8996	0.7494	0.4195

表4 采用不同特征融合方法的性能对比

Table 4 Performance comparison of different feature fusion methods

Method	PLCC	SROCC	KROCC	RMSE
Add	0.8985	0.8757	0.7163	0.4626
Concat	0.9141	0.8914	0.7370	0.4236
SCMA	0.9203	0.8996	0.7494	0.4195

表 5 点云子模型数量及其包含点数对所提方法性能的影响

Table 5 The impact of the number of point cloud sub-models and their inclusion points on the performance of the proposed method

Parameter	PLCC	SROCC	KROCC	RMSE
48×256	0.9060	0.8847	0.7327	0.4495
24×512	0.9049	0.8902	0.7249	0.4557
12×1024	0.8821	0.8607	0.6949	0.4993
6×2048	0.9203	0.8996	0.7494	0.4195

表 6 不同点云子模型中心点生成方法的性能对比

Table 6 Performance comparison of different center point generation methods for point cloud sub-models

Sampling method	PLCC	SROCC	KROCC	RMSE
Random sampling	0.8427	0.8236	0.7129	0.4572
Cell sampling	0.8668	0.7707	0.7018	0.4327
Geometric sampling	0.8643	0.8695	0.6821	0.4311
Farthest point sampling	0.9203	0.8996	0.7494	0.4195

样生成点云子模型中心点的方法性能最好, 主要原因是采用最远点采样可以尽可能地覆盖整个点云模型, 生成的点云子模型可以更好地表征点云特征。

为了验证点云 patch 数量对所提方法性能的影响, 表 7 给出了采用不同 patch 数量在 CPCD2.0 点云数据集上的性能对比。实验结果表明, 当 patch 数量为 6 时, 所提方法的性能最好。其主要原因是较少的 patch 不能较好地覆盖整个点云模型和表征点云特征; 而 patch 数量过多会造成信息冗余, 使系统复杂度增大, 降低系统性能。

表 7 不同点云 patch 数量的性能对比

Table 7 Performance comparison of the number of different cloud points' patch

Patch number	PLCC	SROCC	KROCC	RMSE
2	0.8321	0.8014	0.6929	0.4723
4	0.8718	0.8207	0.7381	0.4532
6	0.9203	0.8996	0.7494	0.4195
8	0.9000	0.8712	0.7301	0.4330

4 结 论

本文提出了一种基于三维特征与二维特征融合的无参考点云质量评价方法, 能在点云二维投影图提取丰富的纹理和语义信息, 在三维空间提取基于点和基于体素的特征, 并通过一个对称跨模态注意模块进行二维特征和三维特征的融合和互补。在 CPCD2.0 点云数据集上的消融实验表明, 使用特征互补后的方法

性能比仅使用二维或三维特征的方法其 PLCC 分别提升了 0.0107、0.4055。与先进的点云质量评价方法相比, 所提方法具有更好的性能。

利益冲突: 所有作者声明无利益冲突

参考文献

- [1] Tao Z Y, Li H, Dou M S, et al. Multi-resolution feature fusion for point cloud classification and segmentation network[J]. *Opto-Electron Eng*, 2023, **50**(10): 230166.
陶志勇, 李衡, 豆森森, 等. 融合多分辨率特征的点云分类与分割网络[J]. *光电工程*, 2023, **50**(10): 230166.
- [2] Gao W, Sun S, Zheng H, et al. OpenDMC: an open-source library and performance evaluation for deep-learning-based multi-frame compression[C]. *Proceedings of the 31st ACM International Conference on Multimedia*, 2023: 9685-9688
- [3] Mekuria R N, Li Z, Tulvan C, et al. Evaluation criteria for PCC (point cloud compression)[J]. *ISO/IEC JTC*, 2016, 1: N16332.
- [4] Tian D, Ochimizu H, Feng C, et al. Geometric distortion metrics for point cloud compression[C]//2017 *IEEE International Conference on Image Processing (ICIP)*, 2017: 3460-3464.
<https://doi.org/10.1109/ICIP.2017.8296925>.
- [5] Cignoni P, Rocchini C, Scopigno R. Metro: measuring error on simplified surfaces[J]. *Comput Graph Forum*, 1998, **17**(2): 167-174.
- [6] Alexiou E, Ebrahimi T. Point cloud quality assessment metric based on angular similarity[C]//2018 *IEEE International Conference on Multimedia and Expo (ICME)*, 2018: 1-6.
<https://doi.org/10.1109/ICME.2018.8486512>.
- [7] Javaheri A, Brites C, Pereira F, et al. Improving psnr-based quality metrics performance for point cloud geometry[C]//2020 *IEEE International Conference on Image Processing*, 2020: 3438-3442. <https://doi.org/10.1109/ICIP40778.2020.9191233>.
- [8] Javaheri A, Brites C, Pereira F, et al. A generalized Hausdorff distance based quality metric for point cloud geometry[C]//2020 *Twelfth International Conference on Quality of Multimedia Experience*, 2020: 1-6.

- <https://doi.org/10.1109/QoMEX48832.2020.9123087>.
- [9] Javaheri A, Brites C, Pereira F, et al. Mahalanobis based point to distribution metric for point cloud geometry quality evaluation[J]. *IEEE Signal Process Lett*, 2020, **27**: 1350–1354.
- [10] Meynet G, Digne J, Lavoué G. PC-MSDM: a quality metric for 3D point clouds[C]//2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), 2019: 1–3. <https://doi.org/10.1109/QoMEX.2019.8743313>.
- [11] Meynet G, Nehmé Y, Digne J, et al. PCQM: a full-reference quality metric for colored 3D point clouds[C]//2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX), 2020: 1–6. <https://doi.org/10.1109/QoMEX48832.2020.9123147>.
- [12] Yang Q, Ma Z, Xu Y L, et al. Inferring point cloud quality via graph similarity[J]. *IEEE Trans Pattern Anal Mach Intell*, 2020, **44**(6): 3015–3029.
- [13] Alexiou E, Ebrahimi T. Towards a point cloud structural similarity metric[C]//2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2020: 1–6. <https://doi.org/10.1109/ICMEW46912.2020.9106005>.
- [14] Hua L, Jiang G Y, Yu M, et al. BQE-CVP: blind quality evaluator for colored point cloud based on visual perception[C]//2021 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), 2021: 1–6. <https://doi.org/10.1109/BMSB53066.2021.9547070>.
- [15] Liu Q, Su H L, Duanmu Z F, et al. Perceptual quality assessment of colored 3D point clouds[J]. *IEEE Trans Vis Comput Graph*, 2023, **29**(8): 3642–3655.
- [16] Lv J Y, Su H L, Long J C, et al. Point cloud quality assessment using multi-level features[J]. *IEEE Access*, 2024, **12**: 47755–47767.
- [17] Tu R W, Jiang G Y, Yu M, et al. Pseudo-reference point cloud quality measurement based on joint 2-D and 3-D distortion description[J]. *IEEE Trans Instrum Meas*, 2023, **72**: 5019314.
- [18] Liu Q, Yuan H, Su H L, et al. PQA-Net: deep no reference point cloud quality assessment via multi-view projection[J]. *IEEE Trans Circuits Syst Video Technol*, 2021, **31**(12): 4645–4660.
- [19] Yang Q, Liu Y P, Chen S H, et al. No-reference point cloud quality assessment via domain adaptation[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 21147–21156. <https://doi.org/10.1109/CVPR52688.2022.02050>.
- [20] Tao W X, Jiang G Y, Jiang Z D, et al. Point cloud projection and multi-scale feature fusion network based blind quality assessment for colored point clouds[C]//Proceedings of the 29th ACM International Conference on Multimedia, 2021: 5266–5272. <https://doi.org/10.1145/3474085.3475645>.
- [21] Tu R W, Jiang G Y, Yu M, et al. V-PCC projection based blind point cloud quality assessment for compression distortion[J]. *IEEE Trans Emerg Top Comput Intell*, 2023, **7**(2): 462–473.
- [22] Tliba M, Chetouani A, Valenzise G, et al. PCQA-graphpoint: efficient deep-based graph metric for point cloud quality assessment[C]//ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023: 1–5. <https://doi.org/10.1109/ICASSP49357.2023.10096610>.
- [23] Qi C R, Yi L, Su H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space[C]. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017: 5105–5114.
- [24] Dodge S, Karam L. Understanding how image quality affects deep neural networks[C]//2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX), 2016: 1–6. <https://doi.org/10.1109/QoMEX.2016.7498955>.
- [25] He Z Y, Jiang G Y, Yu M, et al. TGP-PCQA: texture and geometry projection based quality assessment for colored point clouds[J]. *J Vis Commun Image Represent*, 2022, **83**: 103449.
- [26] Yang Q, Chen H, Ma Z, et al. Predicting the perceptual quality of point cloud: a 3D-to-2D projection-based exploration[J]. *IEEE Trans Multimedia*, 2021, **23**: 3877–3891.
- [27] Javaheri A, Brites C, Pereira F, et al. Point cloud rendering after coding: impacts on subjective and objective quality[J]. *IEEE Trans Multimedia*, 2021, **23**: 4049–4064. <https://doi.org/10.1109/TMM.2020.3037481>.
- [28] Perry S, Cong H P, da Silva Cruz L A, et al. Quality evaluation of static point clouds encoded using MPEG codecs[C]//2020 IEEE International Conference on Image Processing (ICIP), 2020: 3428–3432. <https://doi.org/10.1109/ICIP40778.2020.9191308>.
- [29] Alexiou E, Viola I, Borges T M, et al. A comprehensive study of the rate-distortion performance in MPEG point cloud compression[J]. *APSIPA Trans Signal Inf Process*, 2019, **8**(1): e27.

作者简介



刘太伟(1990-), 男, 山东淄博人, 博士研究生, 2017 年于大连工业大学获得硕士学位, 现为宁波大学信息科学与工程学院博士研究生, 主要从事点云质量评价等方面的研究。

E-mail: 2001100033@nbu.edu.cn



【通信作者】郁梅(1968-), 女, 江苏无锡人, 博士, 教授, 博士生导师, 2000 年于韩国 Ajou 大学获得博士学位, 主要从事多媒体信号处理与通信、计算成像、视觉感知与编码、图像与视频质量评价等方面的研究。

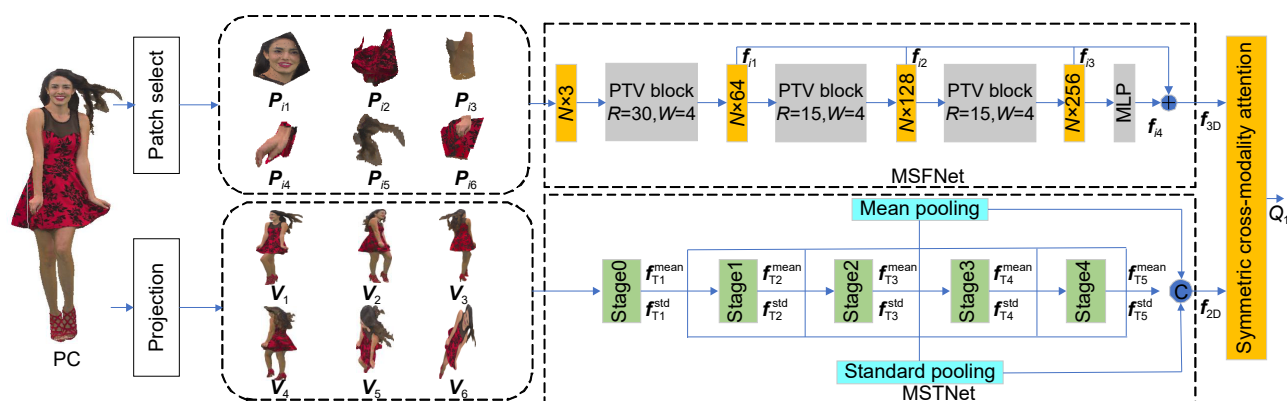
E-mail: yumei@nbu.edu.cn



扫描二维码, 获取PDF全文

No-reference point cloud quality assessment based on fusion of 3D and 2D features

Liu Taiwei, Yu Mei*, Tu Renwei



Framework of the proposed method. R and W denote the voxel resolution and the size of 3D window, respectively

Overview: Point clouds are widely used in virtual reality, computer vision, robotics and other fields, and distortion assessment in point cloud acquisition and processing is becoming an important research topic. Considering that the three-dimensional (3D) information of point cloud is sensitive to geometric distortion and the two-dimensional (2D) projection of point cloud contains rich texture and semantic information, this paper proposes a no-reference point cloud quality assessment method to effectively combine the 3D and 2D feature information of point cloud and improve the accuracy of quality assessment. The farthest point sampling is firstly implemented on the point cloud, and then the non-overlapping point cloud sub-models centered on the selected points are generated, to cover the whole point cloud model as much as possible. For each point cloud sub-model, an improved 3D multi-scale feature extraction network (MSFNet) is designed to extract the features of voxels and points. MSFNet contains three point-voxel transformer (PVT) modules and generates output features through a multilayer perceptron. Each PVT module has two branches. The voxel branch can extract rich semantic features from spatial voxels; the point-based branch can retain the integrity of the point cloud sub-model position information as much as possible and avoid the loss of position information. For 2D feature extraction, the point cloud is first projected with orthogonal hexahedron projection to obtain the corresponding projection maps. To extract the rich texture and semantic information from the 2D projection maps, a 2D multi-scale feature extraction network (MSTNet) is designed to extract 2D content-aware features. Considering that there may be a large amount of redundant information and certain dependency relationships between different viewpoint projection maps, MSTNet uses spatial global average pooling operation to remove redundant information and spatial global standard deviation pooling operation to preserve the dependency information between different viewpoint projection maps. Finally, considering the process of segmentation and interweaving fusion that occurs when the human visual system processes different modality information, to better fuse the 2D and 3D features of the point cloud, so that the two modality features can enhance each other, a symmetric cross-modality attention module is designed to integrate the 3D and 2D features, and a multi-head attention mechanism is added in the feature fusion process. The experimental results on five public point cloud quality assessment datasets show that the Pearson's linear correlation coefficient (PLCC) of the proposed method reaches 0.9203, 0.9463, 0.9125, 0.9164, and 0.9209, respectively, indicating that the proposed method has advanced performance compared with the existing representative point cloud quality assessment methods.

Liu T W, Yu M, Tu R W. No-reference point cloud quality assessment based on fusion of 3D and 2D features[J]. *Opto-Electron Eng*, 2025, 52(4): 250001; DOI: 10.12086/oee.2025.250001

Foundation item: National Natural Science Foundation of China(62071266)

Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China

* E-mail: yumei@nbu.edu.cn