基于内容导向显著性投影的点云简化参考质量评估

周伟、岳光辉、张瑞增、秦一鹏和刘汉涛

arXiv:2301.07681v1 [cs.MM] 18 Jan 2023

摘要——**许多稠密三维点云已被用于表示视觉对象，以替代传统图像或视频。为了评估各类点云的感知质量，本文提出一种新颖且高效的点云简化参考质量指标，该指标基于内容导向显著性投影（RR-CAP）。具体而言，我们首次尝试通过下采样操作，将参考点云和失真点云简化为投影显著性图。通过这一过程，我们解决了将大体量原始点云传输至用户端进行质量评估的问题。随后，受人类视觉系统（HVS）特性的启发，我们结合内容导向相似性与统计相关性测量，得出失真点云的客观质量分数。最后，我们在SJTU-PCQA和WPC数据库上进行了大量实验。实验结果表明，所提出的算法优于现有简化参考和无参考质量指标，并显著缩小了与最先进全参考质量评估方法之间的性能差距。此外，我们通过消融测试展示了各项所提技术组件的性能变化。**

索引词—**点云、简化参考质量指标、视觉内容、显著性投影、人类视觉系统。**

I. 引言

MASSIVE 视觉数据正日益融入我们的日常生活。尤其是近年来，随着三维采集和渲染技术的快速发展，点云已成为最受欢迎的沉浸式媒体格式之一。由此催生了许多现实应用，例如自动潜水、混合现实、遥感等 [1], [2]。通常，稠密点云是一种 3D模型，由空间中一组离散点构成，用于表示物体。每个点都具有几何坐标（即 x、 y、z）和光度属性（例如 RGB颜色信息）。与传统图像/视频类似，点云在多媒体信号处理链中也会经历各种伪影，涉及采集、压缩、传输、重建和显示等环节 [3]–[5]。换言之，这些过程不可避免地会导致用户端的质量下降。因此，点云的质量评估对于监测并确保用户获得满意的体验质量至关重要。

由于人类是点云的最终接收者，主观质量评估是最可靠的方法。通过开展相关主观测试，研究人员已建立了多个点云质量数据库 [6]–[9]。然而，此类主观测试通常耗时、成本高且劳动强度大。因此，如何基于人类视觉系统（HVS）有效设计客观质量评估方法，是一个具有挑战性且前景广阔的研究方向。

一般而言，点云客观质量指标分为三类：全参考（FR）、无参考（NR）和降参考（RR）。当原始信息完全可用时，全参考方法通过比较失真点云与相应原始参考点云之间的差异或相似度来进行评估。最早的全参考指标用于 MPEG 点云压缩标准化工作，包括点到点 [10], 、点到面 [11], 、点到网格 [12], 和面到面 [13], 等方法。这些方法计算参考点云与失真点云之间的距离偏差。随后，Yang 等人 [14] 提出了基于图信号处理的图相似性指数（GraphSIM）方法，并将其扩展为多尺度变体 [15]。受结构相似性指数（SSIM） [16],的启发，Alexiou 等人 [17] 利用几何、法向量、曲率值和颜色信息构建了 PointSSIM。此外，在 [18], 中，Meynet 等人使用局部曲率统计构建了 PC-MSDM，并进一步根据曲率与颜色特征的最优加权线性组合提出了点云质量指标（PCQM） [19]。Lu 等人 [20]通过比较三维边缘相似性来量化点云质量。除了直接比较三维点云模型外，另一类全参考方法是从不同视角将点云投影到多个二维图像中。随后可以采用主流的二维质量评估算法，例如峰值信噪比（PSNR）和结构相似性指数（SSIM）。

在实际场景中，可能无法获取原始参考点云的完整信息。为预测失真点云的视觉质量，研究人员从手工设计和基于学习两个方面开发了多种无参考方法。例如，Zhang等人 [21] 提出了NR-3DQA，该方法基于三维自然场景统计（NSS）评估点云质量。在 [22],中，研究人员通过多任务学习设计了一种点云质量评估网络（PQA-Net）。此外，通过视图投影，还可以采用无参考方式的二维质量度量

本研究部分受国家自然科学基金委员会项目（62001302）和广东省基础与应用基础研究基金（2019A1515111205）资助。W. Zhou 任职于加拿大安大略省滑铁卢市滑铁卢大学电气与计算机工程系，邮编 N2L 3G1（电子邮箱：wei.zhou@uwaterloo.ca）。G. Yue 任职于中国深圳大学生物医学工程学院，深圳 518060（电子邮箱：yueguanghui@szu.edu.cn）。R. Zhang 任职于中国北京理工大学，北京 100081（电子邮箱：hireason@163.com）。Y. Qin 和 H. Liu 任职于英国卡迪夫卡迪夫大学计算机科学与信息学学院，邮编 CF24 4AX（电子邮箱：qiny16@cardiff.ac.uk；liuh35@cardiff.ac.uk）。

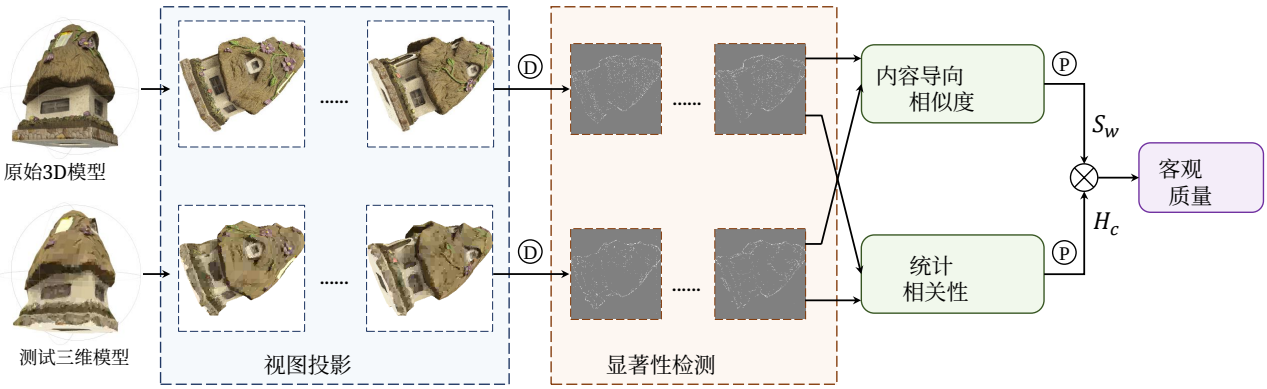


图1：所提出的 RR-CAP 整体框架，其中D◯ 和P◯ 分别表示下采样和池化操作。

例如，具有代表性的自然场景统计方法包括无参考图像空间质量评价器（BRISQUE）[23] 和自然图像质量评价器（NIQE） [24]。当能够获取部分参考信息时，全参考指标可提供一种折中方案。然而，文献中针对点云的全参考质量评估方法寥寥无几。Liu等人 [25]利用量化参数估计V-PCC压缩点云的感知质量。据我们所知，目前只有一种名为PCMRR[26] 的通用全参考指标，它依赖于从原始点云中提取的一小组统计特征，涉及几何、法向量和颜色。PCMRR方法的设计理念未考虑人类视觉系统的特性。

为了解决上述问题，我们提出了基于内容导向显著性投影的 RR 指标（RR-CAP）。本文的主要贡献总结如下：

1）受人类视觉系统（HVS）特性的启发，我们提出了首个通过显著性投影进行图像式全参考点云质量评估的方法。2）在所提出的框架中，我们设计了内容导向相似性和统计相关性测量。3）我们提出的质量度量能够减轻大量点数据带来的传输负担，同时在主观评分数据库上测试时仍表现出竞争力。

本文其余部分安排如下。第二节首先详细介绍所提出的 RR-CAP 方法。第三节展示我们在公开可用的点云质量数据库上进行的实验结果与分析。最后，第四节给出结论。

II. 所提出的 RR-CAP 指标

我们提出的 RR-CAP 整体框架如图 1 所示。为减少参考端的大规模点云数据，我们首先通过视图投影提取下采样后的显著性图。随后结合内容导向相似性与统计相关性，估计测试点云的感知质量。

A. 投影平面中的显著性提取

当受试者浏览目标点云时，会综合所有视图形成质量感知，从而产生最终体验。因此，我们将原始 3D 模型和测试三维模型投影到多个二维图像平面上。这里，我们采用六个垂直投影 [6]，以均匀覆盖大部分可视内容。

由于引入伪影可能会改变显著性行为，且人类视觉系统对显著区域中的失真更加敏感，因此视觉显著性在质量评估中起着至关重要的作用[27], [28]。视图投影后，我们通过图像签名提取显著性图 [29]。具体而言，对于每幅投影图像 I，我们首先将其下采样为分辨率更低的图像：

$$\tilde{I}(i,j) = A * I(si, sj), \tag{1}$$

其中，s 表示下采样尺度。$i = 1、2、\ldots\ldots、{}_{s}^{I}$ 和 $j = 1、2、\ldots\ldots、{}_{s}^{J}$ 分别是下采样图像的行索引和列索引。A 是一个低通滤波器，且 $*$ 表示卷积运算。

然后，下采样投影的图像签名定义为离散余弦变换（DCT）系数 [30] 的符号函数，可按如下方式计算：

$$\hat{I} = \mathrm{sign}(\mathrm{DCT}(\tilde{I})). \tag{2}$$

得到变换格式的图像签名后，我们可以通过逆离散余弦变换（IDCT）将其转换回空间域，并计算显著性图：

$$m = \mathrm{IDCT}(\hat{I}) \circ \mathrm{IDCT}(\hat{I}), \tag{3}$$

其中，$\circ$ 表示哈达玛积。如图 1 所示，与大多数现有研究将视觉显著性用作加权图不同，在我们提出的框架中，下采样后的显著性图用作简化参考信息，从而可以减少原始点云的大量传输数据。

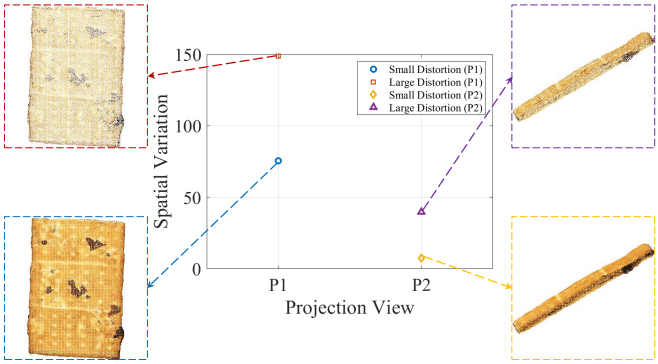


图 2：用于说明视觉内容与质量退化之间关系的示例，其中 P1 和 P2 是两个投影视图。

B. 内容导向的相似性测量

直观而言，我们可以衡量原始投影显著性图与测试投影显著性图之间的图像结构相似性。给定失真显著性图 m_d 和相应的参考显著性图 m_r，我们按如下方式计算 m_d 与 m_r 之间的结构相似性：

$$S = \frac{(2\mu_r\mu_d + C_1)(2\sigma_{rd} + C_2)}{(\mu_r^2 + \mu_d^2 + C_1)(\sigma_r^2 + \sigma_d^2 + C_2)}, \tag{4}$$

其中，μ_r、μ_d、σ_r^2、σ_d^2 和 σ_{rd} 分别表示原始参考显著性图和失真显著性图对应的局部均值、方差和协方差。C_1 和 C_2 是稳定常数。

在对所有视图的相似度进行池化的同时，我们采用了一种与人类视觉系统感知更加一致的内容导向加权策略。考虑到通常使用 Sobel滤波器 [31], 通过空间信息来量化内容，因此这里将空间变化作为权重，并按以下方式计算：

$$w = \left|\mathrm{std}\left[\mathrm{Sobel}\left(I_d\right)\right] - \mathrm{std}\left[\mathrm{Sobel}\left(I_r\right)\right]\right|, \tag{5}$$

其中，std[·] 表示在图像空间的像素上计算的标准差。I_r 和 I_d 分别为参考投影图像和失真投影图像。设 n 为投影视图的数量。内容导向相似性测量的计算如下：

$$S_w = \frac{1}{n}\sum S^w. \tag{6}$$

为展示所提出的内容导向加权策略如何发挥作用，我们提供了两组投影视图，分别具有较小和较大的失真程度，如图 2 所示。可以观察到，不同视图中的视觉内容通常具有不同的空间变化。此外，失真程度较大的投影视图会产生更大的空间变化。

C. 统计相关性测量

除了内容导向的相似性测量之外，显著性图中的统计信息对于点云的感知质量也很重要。因此，对于每个测试点云，我们通过以下公式构建统计相关性测量：

$$H_c = \frac{1}{n}\sum \frac{\mathrm{E}\left[h_r h_d\right] - \mathrm{E}\left[h_r\right]\mathrm{E}\left[h_d\right]}{\sqrt{\mathrm{E}\left[h_r^2\right] - \left(\mathrm{E}\left[h_r\right]\right)^2}\sqrt{\mathrm{E}\left[h_d{}^2\right] - \left(\mathrm{E}\left[h_d\right]\right)^2}}, \tag{7}$$

其中，E[·] 表示期望算子。h_r 和 h_d 分别是原始参考显著性图和失真显著性图的统计直方图。

通过上述步骤，我们得到两种质量测量，即 S_w 和 H_c，分别用于衡量内容导向相似性和统计相关性。最后，失真点云的客观质量分数可计算如下：

$$Q = S_w \cdot H_c. \tag{8}$$

III. 结果与分析

本节将我们提出的 RR-CAP 与当前先进的质量评估方法在公开的点云主观评分质量数据库上进行比较。此外，我们还开展了消融研究，以进一步验证各个所提出技术组件的性能。

A. 实验方案

我们在 SJTU-PCQA [6]和 WPC [7] 点云质量数据库上进行了大量实验。两个数据库均为每个失真点云样本提供平均意见分（MOS），作为真实主观质量。

- SJTU-PCQA数据库：由 9 个参考点云样本组成。通过几何高斯噪声、颜色噪声、降尺度、基于八叉树的压缩及其混合失真等多种失真类型，生成了 378 个失真点云。每种失真类型均包含六个级别。

- WPC数据库：包含 20 个原始点云和四种失真类型，共生成 740 个失真点云。具体而言，其中有 60 个经下采样失真的点云、180 个含高斯噪声的失真点云、320 个经 G-PCC (T) 压缩的点云，以及 180 个经 V-PCC 压缩的点云。

采用四种广泛使用的评价准则进行性能比较，包括斯皮尔曼秩相关系数（SROCC）、肯德尔秩相关系数（KROCC）、皮尔逊线性相关系数（PLCC）和均方根误差（RMSE）。其中，SROCC 和 PLCC/RMSE 分别用于衡量预测单调性和准确性。KROCC 用于衡量两个测量量之间的序数关联。优秀的质量度量应使 SROCC、KROCC 和 PLCC 的性能值接近 1，而 RMSE 值接近 0。请注意，在计算不同质量度量的 PLCC 和 RMSE 之前，会执行非线性拟合过程 [32]，将预测质量分数映射到主观分数的公共尺度空间。

表 I：客观质量指标在 SJTU-PCQA 和 WPC 数据库上的性能比较。FR、NR 和 RR 的最佳性能值以粗体标出。

Ref	Type	方法	SJTU-PCQA数据库				WPC数据库			
			SROCC	KROCC	PLCC	RMSE	SROCC	KROCC	PLCC	RMSE
FR	基于模型的	GraphSIM [14]	0.8483	0.6448	0.8449	1.5721	0.5831	0.4194	0.6163	17.1939
		PointSSIM [17]	0.6867	0.4964	0.7136	1.7001	0.4542	0.3278	0.4667	20.2733
		PCQM [19]	**0.8544**	**0.6586**	**0.8653**	**1.2162**	**0.7434**	**0.5601**	**0.7499**	**15.1639**
	基于图像	PSNR	0.2422	0.1077	0.2317	2.3124	0.4235	0.3080	0.4872	15.8133
		SSIM [16]	0.2987	0.1919	0.3476	2.1770	0.3878	0.3234	0.4944	15.7749
NR	基于模型	NR-3DQA [21]	**0.7144**	**0.5174**	**0.7382**	**1.7686**	**0.6479**	**0.4417**	**0.6514**	**16.5716**
	基于图像	BRISQUE [23]	0.2051	0.1121	0.2241	2.2428	0.3781	0.2444	0.4176	22.5414
		NIQE [24]	0.2214	0.1512	0.3764	2.2671	0.3887	0.2551	0.3957	22.5502
RR	基于模型	PCMRR [26]	0.4816	0.3362	0.6191	1.9342	0.3097	0.2082	0.3433	21.5302
	基于图像	**所提出的 RR-CAP**	**0.7577**	**0.5508**	**0.7691**	**1.5512**	**0.7162**	**0.5260**	**0.7307**	**15.6485**

B. 性能比较

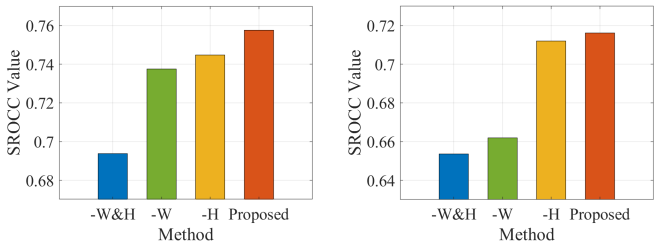
为验证所提出的 RR-CAP 方法的性能，我们将其与现有的客观质量评估方法进行比较，包括全参考、无参考和 RR 指标。结果如表 I 所示。总体而言，参与比较的指标可分为两类：基于模型的方法和基于图像的方法。基于模型的指标直接作用于 3D 模型，包括 GraphSIM [14],PointSSIM [17], PCQM [19],NR-3DQA [21], 和 PCMRR [26]。基于图像的指标将 3D 模型投影到二维图像平面上，然后对相应的投影执行质量评估。参与比较的基于图像的指标包括 PNSR、SSIM [16], BRISQUE [23], 和 NIQE [24]。请注意，PCMRR 是唯一的 RR 方法，且属于基于模型的指标。也就是说，目前尚无专为点云设计的基于图像的 RR 指标研究。因此，我们尝试填补这一空白，并基于人类视觉系统（HVS）的特性提出一种基于图像的 RR 点云度量。

根据定量结果表，我们进行如下分析。首先，传统的基于图像的指标在估计感知质量时效果较差，因为它们没有考虑点云的特性。其次，我们提出的 RR-CAP 优于当前先进的无参考和 RR 指标，包括基于模型和基于图像的指标。
这表明，尽管所提出的方法是一种基于图像的指标，但其性能仍然优越。值得注意的是，我们提出的指标显著优于 PCMRR。这主要是因为我们的框架考虑了人类视觉系统的感知特性和观看过程。第三，我们发现，与基于模型的全参考质量评估方法相比，所提出的 RR 指标能够大幅缩小性能差距。

C. 消融测试

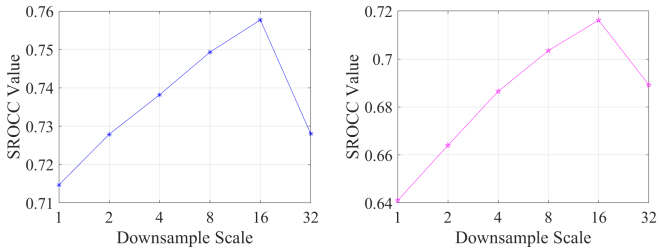
对所提出方法的各项技术组件和参数进行性能变化测试颇有意义。在图3和图4中，我们展示了针对内容导向加权策略、统计直方图特征和下采样尺度的测试结果。

从图中可以看出，随着加权策略和统计信息的加入，我们的方法逐步得到改进。这验证了所提出技术组件的有效性。此外，较大的下采样尺度能够节省更多传输资源。为了获得最佳性



(a) (b)

图3：各项技术组件对性能的影响，其中“-”表示移除相应项目。W 和 H 分别表示所提出的加权策略和统计直方图特征。(a) 在 SJTU-PCQA数据库上运行；(b) 在 WPC数据库上运行。



(a) (b)

图4：性能随下采样尺度的变化。（a）在 SJTU-PCQA 数据库上运行；（b）在 WPC数据库上运行。

能并减轻传输负担，我们在所提出的框架中选择了16作为下采样尺度。

IV. 结论

本文提出了一种通用的新型RR指标，用于点云的客观质量评估。受人类视觉系统特性的启发，我们提出的方法以显著性投影为基础，随后进行内容导向相似性计算和统计相关性测量。实验表明，与最先进的质量指标相比，所提出的RR-CAP与主观评分具有高度一致性。该指标的实现可在以下网址获取：
https://github.com/weizhou-geek/RR-CAP。

参考文献

[1] Y. Guo、H. Wang、Q. Hu、H. Liu、L. Liu 和 M. Bennamoun，“三维点云深度学习：综述”，*IEEE*模式分析与机器智能汇刊，第43卷，第12期，第4338–4364页，2020年。 [2] S. Schwarz、M. Preda、V. Baroncini、M. Budagavi、P. Cesar、P. A. Chou、R. A. Cohen、M. Krivoku´ca、S. Lasserre、Z. Li 等，“点云压缩的新兴MPEG标准”，*IEEE*新兴与精选电路和系统专题期刊，第9卷，148，2018 1，133期，第 – 页。[3] S. Sun、T. Yu、J. Xu、W. Zhou 和 Z. Chen，“GraphIQA：学习用于盲图像质量评估的失真图表示”，*IEEE*多媒体汇刊，2022年。 [4] S. Chikkerur、V. Sundaram、M. Reisslein 和 L. J. Karam，“ 客观视频质量评估方法：分类、综述与性能比较 ”，*IEEE*广播汇刊，第57卷，第2期，第165–182页，2011年。 [5] W. Zhou、Q. Yang、Q. Jiang、G. Zhai 和 W. Lin，“基于结构引导重采样的三维稠密点云盲质量评估”，*arXiv*预印本 *arXiv:2208.14603*，2022年。[6] Q. Yang 、 H. Chen 、 Z. Ma 、 Y. Xu 、 R. Tang 和 J. Sun，“ 预测点云的感知质量：基于三维到二维投影的探索”，*IEEE*多媒体汇刊，第23卷，3891，2020 3877页。[7] Q. Liu、H. Su、Z. Duanmu、W. Liu 和 Z. Wang，“彩色三维点云的感知质量评估”，*IEEE*可视化与计算机图形学汇刊，2022年。 [8] E. Alexiou、I. Viola、T. M. Borges、T. A. Fonseca、R. L. De Queiroz 和 T. Ebrahimi，“MPEG点云压缩率失真性能的全面研究”，*APSIPA*信号与信息处理汇刊，第8卷，2019年。 [9] A. Javaheri、C. Brites、F. Pereira 和 J. Ascenso，“编码后的点云渲染：对主观和客观质量的影响”，*IEEE*多媒体汇刊，第23卷，第4049–4064页，2020年。 [10] R. Mekuria、Z. Li、C. Tulvan 和 P. Chou，“点云压缩的评估准则”，*ISO/IEC MPEG*，第16332号，2016年。 [11] D. Tian、H. Ochimizu、C. Feng、R. Cohen 和 A. Vetro，“点云压缩的几何失真度量”，载于*IEEE*国际图像处理会议，2017年，第3460–3464页。 [12] P. Cignoni、C. Rocchini 和 R. Scopigno，“Metro：测量简化曲面上的误差”，载于计算机图形论坛，第17卷，第2期。Wiley在线图书馆，1998年，第167–174页。 [13] E. Alexiou 和 T. Ebrahimi，“基于角相似度的点云质量评估指标”，载于*IEEE*国际多媒体与博览会会议，2018年，第1–6页。 [14] Q. Yang、Z. Ma、Y. Xu、Z. Li 和 J. Sun，“通过图相似度推断点云质量”，*IEEE*模式分析与机器智能汇刊，2020年。 [15] Y. Zhang、Q. Yang 和 Y. Xu，“MS-GraphSIM：通过多尺度图相似度推断点云质量”，载于第*29*届*ACM*国际多媒体会议论文集，2021年，第1230–1238页。 [22] Q. Liu、H. Yuan、H. Su、H. Liu、Y. Wang、H. Yang 和 J. Hou，“PQA-Net：通过多视图投影进行无参考点云质量深度评估”，*IEEE*视频技术电路与系统汇刊，第31卷，第12期，第4645–4660页，2021年。

[16] Z. Wang、A. C. Bovik、H. R. Sheikh 和 E. P. Simoncelli，“图像质量评估：从误差可见性到结构相似性”，*IEEE*图像处理汇刊，第 13 卷，第 4 期，第 600–612 页，2004 年。[17] E. Alexiou 和 T. Ebrahimi，“迈向点云结构相似性度量”，载于*IEEE*国际多媒体与博览会研讨会，2020 年，第 1–6 页。[18] G. Meynet、J. Digne 和 G. Lavou´e，“PC-MSDM：一种三维点云质量度量”，载于*IEEE*国际多媒体体验质量会议，2019 年，第 1–3 页。[19] G. Meynet、Y. Nehm´e、J. Digne 和 G. Lavou´e，“PCQM：一种彩色三维点云全参考质量度量”，载于*IEEE*国际多媒体体验质量会议，2020 年，第 1–6 页。[20] Z. Lu、H. Huang、H. Zeng、J. Hou 和 K.-K. Ma，“基于三维边缘相似性测量的点云质量评估”，*IEEE*信号处理快报，第 29 卷，第 1804–1808 页，2022 年。[21] Z. Zhang、W. Sun、X. Min、T. Wang、W. Lu 和 G. Zhai，“三维彩色点云和网格模型的无参考质量评估”，*IEEE*视频技术电路与系统汇刊，第 32 卷，第 11 期，第 7618–7631 页，2022 年。[23] A. Mittal、A. K. Moorthy 和 A. C. Bovik，“空间域中的无参考图像质量评估”，*IEEE*图像处理汇刊，第 21 卷，第 12 期，第 4695–4708 页，2012 年。[24] A. Mittal、R. Soundararajan 和 A. C. Bovik，“打造一个‘完全盲’的图像质量分析器”，*IEEE*信号处理快报，第 20 卷，第 3 期，第 209–212 页，2012 年。[25] Q. Liu、H. Yuan、R. Hamzaoui、H. Su、J. Hou 和 H. Yang，“应用于基于视频的点云压缩码率控制的简化参考感知质量模型”，*IEEE*图像处理汇刊，第 30 卷，第 6623–6636 页，2021 年。[26] I. Viola 和 P. Cesar，“用于点云内容视觉质量评估的简化参考度量”，*IEEE*信号处理快报，第 27 卷，第 1660–1664 页，2020 年。[27] W. Lin、S. Lee 等，“三维点云和网格的视觉显著性与质量评估：综述”，*APSIPA*信号与信息处理汇刊，第 11 卷，第 1 期，2022 年。[28] X. Min、K. Gu、G. Zhai、M. Hu 和 X. Yang，“面向自然场景和屏幕内容图像的显著性驱动简化参考质量指数”，信号处理，第 145 卷，第 127–136 页，2018 年。[29] X. Hou、J. Harel 和 C. Koch，“图像签名：突出稀疏显著区域”，*IEEE*模式分析与机器智能汇刊，第 34 卷，第 1 期，第 194–201 页，2011 年。[30] M. A. Saad、A. C. Bovik 和 C. Charrier，“盲图像质量评估：一种基于DCT域自然场景统计的方法”，*IEEE*图像处理汇刊，第 21 卷，3352，2012 8，3339，页码顺序错误。[31] H. Yu 和 S. Winkler，“图像复杂度与空间信息”，载于*IEEE*国际多媒体体验质量研讨会，2013 年，第 12–17 页。[32] A. M. Rohaly、P. J. Corriveau、J. M. Libert、A. A. Webster、V. Baroncini、J. Beerends、J.-L. Blin、L. Contin、T. Hamada、D. Harrison 等，“视频质量专家组：当前成果与未来方向”，载于视觉通信与图像处理，第 4067 卷。SPIE，2000 年，第 742–753 页。