

[PaperTranslation]Point Cloud Saliency Detection by Local and Global Feature Fusion

相关信息: 2019 TIP <https://ieeexplore.ieee.org/abstract/document/8726371>

标题: 局部和全局特征融合的点云显著性检测

作者: Xiaoying Ding , Wei si Lin , Fellow, IEEE, Zhenzhong Chen , Senior Member , IEEE, and Xinfeng Zhang , Member , IEEE

出处: 2019 IEEE TRANSACTIONS ON IMAGE PROCESSING

摘要

受人类视觉系统特点的启发, 提出了一种新颖的3D点云视觉显著区域检测方法。首先, 根据与局部环境的差异来评估每个点的局部独特性 (local distinctness) ; 然后, 将点云分解成小簇, 计算每个簇的初始全局稀有度值 (global rarity value) ; 然后使用随机游走排序方法 (random walk ranking method) 对所有簇中的每个点引入簇级别的全局稀有性细化; 最后, 提出了一种优化框架, 将局部区别和全局稀有度值结合起来, 以获得点云的最终显著性检测结果。我们将所提出的方法与几种相关算法进行了比较, 并将其应用于一些计算机图形应用程序, 例如兴趣点检测、视点选择和网格简化; 实验结果证明了所提方法的优越性能。

索引词: 3D点云, 视觉感知, 显著性。

I .介绍

由于随着技术的飞速发展, 3D 点云数据的获取越来越方便, 这使得 3D 点云处理越来越受到关注 [1]; 近年来, 一些 3D 点云处理方案也被提出用于以人为中心的视觉计算应用, 例如基于 3D 点云的物体检测 [2]、识别 [3]、定位 [4] 和配准 [5]; 然而, 这些方法主要利用基于信号的数学测量, 例如曲率 [6], 而没有考虑 3D 模型上的人类视觉系统特征。

视觉显著性是人类视觉系统的一个重要特征, 它描述了给定场景下的人类注意力分布或眼球运动[7]; 检测这种视觉上的显著区域是计算机视觉和计算机图形学领域[6]中的重要组成部分; 然而, 大多数显著性检测工作都集中在 2D 图像 [8] 和视频 [9] 上, 只有少数作品计算 3D 网格 [10] 和点云的显著性; 由于 3D 模型通常包含**大量数据并且更容易被噪声影响而失真**, 因此用传统方法分析此类数据的可能在计算上很昂贵, 使得 3D 模型的显著性检测比 2D 数据更具挑战性, 并且现有的在 2D 数据上设计的显著性检测方法也很难直接应用于 3D 模型; 最近, 已经提出了一些在 3D 网格上进行显著性检测的算法, 这些方法主要依赖于利用拓扑信息, 例如光谱特性 (spectral properties) [11] 或测地距离 (geodesic distances) , 然而, 由于点云中缺乏拓扑信息及其不同的采样密度, 将这些算法简单地扩展到无组织的点云可能是不切实际的; 为了解决这些限制, 我们专注于**无组织**的 3D 点云的显著性检测。

由于人类的视觉注意力被吸引到与其他区域不同或变化的区域 [1], 3D 点云的显著性检测可以被视为**寻找感知上重要的区域**, 这些区域相对于其周围区域是唯一的 [6]; 寻找这些感知上重要的区域已成为许多应用程序的有用工具, 例如视点选择 [12]、点云配准 [13]、**简化** [14]、分割 [15] 和缩放 [16]; 例如, 当我们对 3D 点云模型进行简化时, 我们希望在显著区域中保留更多的细节, 通过应用这种人类感知启发的显著性度量来指导网格简化应用, 可以保留更多的细节, 从而提供更好的可视化结果; 与使用诸如曲率 [17] 之类的纯数学测量相比, 这种受人类感知启发的显著性测量可以在 3D 点云 [6] 的显示和处理中提供更具视觉吸引力的结果; 为了考虑人类视觉系统[18], 在所提出的方法中考虑了局部独特性和全局稀有性线索; 局部独特性解释了该点与其局部环境的差异, 而全局稀有性作用在于检测整个独特区域并

抑制频繁出现的区域；由于局部独特区域如果在模型周围多次出现，则可能具有较低的全局稀有性，因此神经科学和心理物理学的许多研究表明，应同时考虑局部特征和全局特征，这两个特征都得到心理学证据的支持[19]并已成功应用于二维图像中的显著性检测，例如[20]中提出的显著性检测算法；除了[20]中提出的算法外，许多提出的算法也将这两个线索结合在一起，以获得令人满意的图像显著性检测结果；例如，Wang等人[21]提供了一种利用多尺度局部对比度和全局空间分布来检测显著性的算法；在[22]中，Borji和Itti提供了一种新颖的显著性检测算法，可以测量不同颜色空间中的补丁稀有度

(patch rarities)，并将它们组合成最终的显著性检测图；根据这些算法，局部和全局线索的结合比使用单一线索可以提高显著性检测的准确性，并且更符合人类视觉系统的特征。

本文提出了一种新颖的点云显著性检测算法，该算法同时利用了局部独特性和全局稀有性线索。为了测量局部独特性，我们将其定义为描述符空间中一个点与其 k 最近邻点之间的距离，并利用快速点特征直方图 (Fast Point Feature Histogram FPFH) 描述符 [23] 来提取每个点的局部几何特征；为了测量全局稀有度，使用[24]中提出的超体素分割方法 (supervoxel segmentation method) 将点云分解为几个簇，并通过计算一个簇与描述符空间中剩余簇之间的距离来获得初始全局稀有性；然后采用随机游走排序方法，通过考虑每个点的局部几何特征，对所有簇中的每个点引入簇级的全局稀有性细化；由于局部几何特征表示相邻点之间的关系，使用它来细化初始全局稀有度将提供更精细的全局稀有度结果；最后，提出了一个有效的优化框架，将每个点的局部显著性和全局稀有度值结合起来，得到最终的显著性检测结果；Max Planck模型的显著性检测结果示例如图1所示，根据图1，与[1]中报道的结果相比，我们的检测结果更集中在眼睛等面部特征上，与[25]相比，提供了更好的显著性检测结果；请注意，所提出算法的支持材料可在线获得 (<http://iip.whu.edu.cn/PCSSaliency/>)。

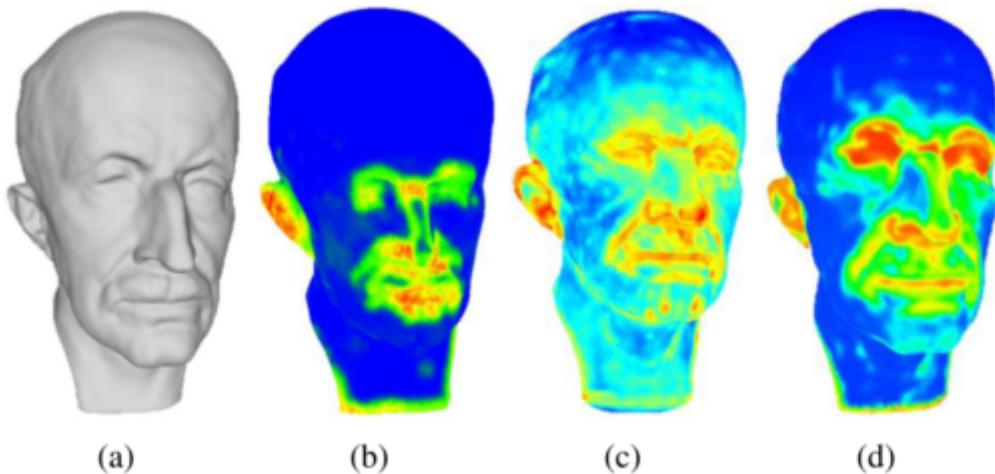


图1 Max Planck 模型的显著性检测结果；(a) 原始模型、(b) 提出的显著性检测结果、(c) 和 (d) 分别在 [1] 和 [25] 中报告的显著性检测结果；与 [1] 和 [25] 相比，我们在该模型上的显著性检测更集中于面部特征，并且具有更好的显著性检测结果

本文的主要贡献有两个；第一个是在对所有簇中的每个点引入簇级全局稀有性细化时考虑每个点的局部几何特征，与计算[25]中不考虑局部几何特征影响的点簇概率矩阵相比，他们的方法不可避免地会忽略点云的一些详细信息，为了解决这个限制，我们在随机游走过程中使用每个点的局部几何特征作为指导，并生成更详细的点级全局稀有结果；第二个贡献是提出的优化框架，现有的显著性检测算法通常对不同的显著性线索进行简单的求和或线性组合来获得最终的显著性检测结果，没有采取不同的显著性线索的特征，考虑到显著特征，所提出的自适应优化框架可以利用不同显著性线索之间的内在关系，从而获得更令人满意的结果。

本文的其余部分组织如下：第二节介绍了显著性检测的相关工作；在第三节中，详细说明了所提出的显著性检测方法；第四节描述了实验结果，第五节提供了结论。

II. 相关工作

在本节中，我们首先总结了图像和视频中视觉显著性检测的相关工作，然后我们讨论一些计算 3D 网格和点云的视觉显著性的工作。

A. 在2D数据上的视觉显著性

视觉显著性检测是 HVS [26] 中注意力机制的模拟，它描述了人类注意力分布 [27] 或给定场景的眼球运动；显著性检测大致可以分为两类，**自下而上的显著性检测**和**自上而下的显著性检测**；前者利用低级特征[28]，例如方向和纹理，例如[29]中提出的方法；而后者主要利用高级知识[30]，例如[31]中使用的深度聚焦 (depth-from-focus) 线索。

显著性检测问题已经研究了几十年，并且已经提出了许多显著性检测方法[32]，特别是对于二维图像中的显著性检测；在[33]中，提出了一种基于扩散的图像显著性检测算法，该算法首先构建一个两层稀疏图，然后使用流形排序扩散方法；在[34]中，Ishikura 等人通过利用两个基于全局极值的特征来计算显著区域，以细化通过计算局部感知色差的多尺度极值获得的潜在显著区域。

除了将显著性检测应用于二维图像外，它还应用于视频等其他数据；例如，在[35]中，提出了一种视频显著性检测算法，该算法融合了**基于运动线索的颜色显著性以获得低秩连贯性线索**，然后用它们来指导**时空显著性扩散**；在[36]中，通过深度结合帧间和帧内信息提出了一种新算法，然后应用能量优化来改善显著性检测结果。

除了视频显著性检测之外，最近的工作还专注于对包含 3D 信息的图像进行显著性检测，例如 RGB-D 图像和全景图像；前者代表一张 RGB 图像及其对应的深度图像，而后者则记录了 3D 世界中的整个场景；对于 RGB-D 图像显著性检测，Cong 等人 [37] 提出了一种方法，该方法首先测量深度信息的置信度，然后通过整合颜色和深度信息来计算紧凑性显著性。随后将不同的显著性线索整合起来以获得最终的显著性检测结果；在 [38] 中，Niu 等人提出了一种利用全局视差对比度和领域知识来计算 RGB-D 显著性的算法，并且在 [39] 中，提出了一种基于学习的算法用于 RGB-D 显著性检测，在他们的论文中，单目特征首先被用来捕捉个体显著性的内在特征，然后他们利用补偿特征来保持立体匹配区域之间的一致性；对于全景图像中的显著性检测，Sitzmann 等人[40]捕获并分析了凝视和头部方向数据，以帮助更好地了解人们如何探索虚拟环境，此外，他们还探索了几个与虚拟现实中显著性预测相关的应用，在 [41] 中，Zhu 等人提出了一种算法来帮助预测观众在观看全景图像时的头部运动；在[42]中，Hu 等人提出了一种在线代理，可以通过选择聚焦于最显著对象的视角来导航 360 度视频。

B. 在3D数据上的视觉显著性

尽管显著性检测已应用于各种 2D 数据，但很少有人提出处理 3D 数据的方法；点云是许多传感器的原始输出数据，例如激光雷达或 Microsoft Kinect 传感器，并且可以通过评估已知的数学描述或通过插入一组给定的无组织 3D 点云来重建 3D 网格；例如，在 [43] 中，Hoppe 等人提出了一种新的算法来从无组织的点云重建表面，而不假设数据点上的任何结构，Mitra 和 Nguyen [44] 提出了一种使用局部最小二乘拟合方法估计所有样本点的表面法线的方法，直接检测 3D 点云而不是 3D 网格上的显著区域将大大提高计算效率，因为它避免了使用点云生成耗时的网格；此外，3D 网格通常由于存储了表面信息而具有更大的数据量，检测点云上的显著区域将大大降低存储容量并消除进行表面重建所引入的错误。

网格 (mesh) 显著性检测已经研究了很长时间，并且已经提出了许多算法；早期网格显著性检测算法通常通过在其二维投影中应用显著性检测来评估模型的显著性；例如，Yee 等人 [45] 将显著性检测应用于投影 3D 动态场景的参考图像，并使用显著性图来减少计算全局照明解决方案的计算量；Frintrop 等人 [46] 提供了一种方法，该方法可以通过将 3D 场景的渲染图像输入注意力系统来检测潜在感兴趣的区域，从而加速 3D 对象检测和分类；Howlett 等人[47]捕获了呈现 3D 对象的 2D 图像的人类注视数据，并使用人类注视数据来帮助确定 3D 对象的显著特征。

最近，受到二维图像显著性检测工作的启发，提出了许多网格显著性检测算法[48]；例如，Lee 等人 [6] 以依赖于尺度的方式计算网格显著性，该方式使用中心环绕 (center-surround) 机制，因为它具有识别与其周围环境不同的区域的直观吸引力，这种中心环绕操作首先由 Itti 等人[48]提出，这是一种有用的图像显著性检测技术；此外，Castellani 等人 [49] 提出了一种使用联合多尺度机制的新型网格显著性检测方法，该方法也是由 Itti 等人 [48] 提出的；Wu 等人[50]提供了一种基于全局稀有度的网格显著性检测算法，该算法也被 Cheng [20] 等人用作显著性检测原理；Leifman 等人 [12] 计算了局部和全局的网格显著性，并考虑了周围的环境，这是由 [51] 中提出的上下文感知算法推动的。

尽管针对 3D 网格中的显著性检测提出了各种算法，但很少有显著性模型可以处理 3D 点云；参考文献 [1] 提出了一种鲁棒算法，可以帮助点云中的显著性检测，他们的方法同时考虑了低水平和高水平的独特性，并已通过各种应用程序进行了评估；Tasse 等人 [25] 提供了一种基于聚类的 3D 点云显著性检测算法，该算法利用聚类唯一性和空间分布来制定最终的显著性检测结果。

III. 提出的方法

在本文中，我们专注于 3D 点云的显著检测；为了解决这个问题，我们考虑了局部独特性和全局稀有性，这是心理学证据支持的人类视觉系统的两个基本特征[19]；由于这两个特征在二维显著性检测中取得了令人满意的结果，我们利用它们来改进点云的显著性检测；图2展示了提出的框架，首先评估每个点的局部特征，捕捉耳朵和鼻子的轮廓等局部显著特征，然后全局稀有度找到更大的显著区域，例如整个耳朵，最后，提出了一种优化框架，将局部独特性和全局稀有度值结合起来，以获得点云的最终显著性检测结果。

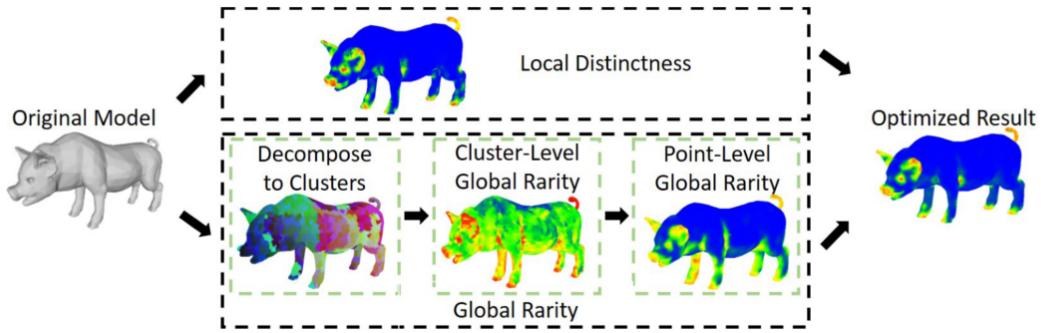


图 2. 算法框架；首先评估每个点的局部独特性，然后将点云分解为若干个簇，得到每个点的全局稀有度值，最后，提出了一种优化框架，将局部独特性和全局稀有性值结合起来，以获得点云中每个点的最终显著性值

A. 局部独特性

为了计算每个点的局部独特性，需要一个描述符来表征局部几何特征；描述符应该对噪声具有鲁棒性和计算效率，并且具有局部几何特征的良好表达能力[1]；我们已经尝试了几个描述符，例如 Zernike 系数 [52]、SHOT [53] 描述符，发现 FPFH [23] 描述符更有效地与我们的点云的局部几何特征相协调。

为了计算查询点 $p_i \in P$ 的 FPFH 描述符，其中 P 表示点云，选择其在半径为 r 的球体中的 k 个相邻点，并使用 Darboux uvn 框架 ($u = n_i, v = (p_j - p_i) \times u, w = u \times v$) 定义如下 [23]：

$$\begin{aligned} \alpha &= v \cdot n_j \\ \phi &= (u \cdot (p_j - p_i)) / \|p_j - p_i\| \\ \theta &= \arctan(\omega \cdot n_j, u \cdot n_j) \end{aligned} \quad (1)$$

其中 p_j 是属于点 p_i 的 k 邻域的点。

使用上述量化的角度，可以计算查询点与其相邻之间的关系，称为简化点特征直方图 (SPFH) [23]，稍后对于每个查询点，SPFH 值用于获取该点的最终 FPFH 描述符，指定为：

$$FPFH(p_i) = SPFH(p_i) + \frac{1}{k} \sum_{j=1}^k \frac{SPFH(p_j)}{\|p_i - p_j\|} \quad (2)$$

其中 k 是点 p_i 的最近邻点数， $\|\cdot\|$ 表示给定度量空间中点 p_i 和 k 最近邻点 p_j 之间的 L2 距离，如 [23] 中所述。

由于 3D 点云中噪声和不同采样密度的影响，欧几里得距离不适用于测量两点之间的相异性。按照[1]中的方法，使用卡方距离来衡量点之间的相异性。给定 P 中两个点 p_i 和 p_j ，卡方距离 $\chi^2(p_i, p_j)$ 计算为：

$$\chi^2(p_i, p_j) = \sum_{n=1}^N \frac{(FPFH_n(p_i) - FPFH_n(p_j))^2}{FPFH_n(p_i) + FPFH_n(p_j)} \quad (3)$$

其中 N 表示 FPFH 描述符中的 bin 数量。

由于点 p_i 与其周围点不同，则它是不同的，因此局部不同性可以定义为与其周围点的不同点的加权和，公式定义为：

$$D(p_i) = 1 - \exp\left(-\frac{1}{R} \sum_{j=1}^R \frac{\chi^2(p_i, p_j)}{1 + \|p_i - p_j\|}\right) \quad (4)$$

其中 $D(p_i)$ 表示点 p_i 的局部区别值，R 是点 p_i 考虑的周围点的数量；注意 R 是根据点云的采样密度设置的。

为了防止标记过多的不同点，等式 (5) 中设计了一个简单但有效的分段抑制函数来减少不同点的数量；

$$F(p_i) = \begin{cases} D(p_i) & D(p_i) \geq h \\ D(p_i) * D(p_i) & \text{otherwise} \end{cases} \quad (5)$$

其中 $F(p_i)$ 表示点 p_i 的放大局部区别值，h 表示控制要抑制的不同点范围的阈值。在我们的实验中，我们使用了一个自适应阈值 h，它等于点云中前 20% 的局部区别值；我们使用这个阈值是因为与其他阈值相比，它可以获得更好的局部区别计算结果，并且图像显著性检测中的许多其他算法也使用了这个阈值，例如 Borji [54] 和 Kanan 等人提出的算法 [30]。

B. 全局稀有度

为了计算点云的全局稀有度，我们首先将点云分解为小簇，以减少需要考虑的区域数量；使用[24]中提出的体素云连通性分割算法 (The voxel cloud connectivity segmentation algorithm) 用于分解点云，Papon 的方法利用体素关系产生与场景在三个维度上的空间几何形状完全一致的过分割，从而获得了最小的信息损失，提高了计算效率。

为了表征每个簇的几何特征，为簇 c_i 定义了平均 FPFH 描述符 $FPFH(c_i)$ 。请注意， $FPFH(c_i)$ 是簇 c_i 中所有点的 FPFH 描述符的平均值；由于全局稀有度被认为是一个区域在整个点云上发生的概率的倒数 [22]，因此通过考虑每个集群的不同权重，以下计算集群 c_i 的初始全局稀有度：

$$G(c_i) = 1 - \exp\left(-\frac{1}{N} \sum_{j=1}^N \frac{\chi^2(c_i, c_j)}{1 + \|c_i - c_j\|}\right) \quad (6)$$

其中 $G(c_i)$ 是簇 c_i 的初始全局稀有度值，N 是点云中的簇数， $\|c_i - c_j\|$ 表示集群 c_i 和 c_j 的中心之间的 L2 距离。请注意，卡方距离 $\chi^2(c_i, c_j)$ 是使用 Eqn 计算的 (3)。

使用方程式中的公式(6)，得到每个簇的初始全局稀有度值，然后我们需要对所有集群中的每个点引入集群级别的全局稀有度；Tasse 等人[25] 定义了一个点簇概率矩阵来表示点 p_i 属于簇 c_j 的概率，尽管在构建概率矩阵时考虑了每个簇的空间分布和描述符空间的差异，但他们的方法没有考虑每个点的局部几何特征的影响，这在对象分类等应用中显示出令人印象深刻的成果[55]；不使用每个点的局部几何特征计

算全局稀有度将不可避免地忽略来自原始模型的一些详细信息，并降低全局稀有度计算的准确性。

随机游走是一种算法，它计算从每个未播种点（unseeded point）开始的随机游走者首先到达每个播种点（seeded points）的概率，并将预定义的值分配给具有最大概率的未播种点；它广泛应用于图像分割、图像显著性检测等图像处理领域；受 [56] 的启发，该方法使用随机游走排序方法将先验显著性估计引入图像中的每个像素，我们将随机游走排序算法应用于所提出的方法，以帮助将集群级全局稀有度引入点级全局稀有度通过考虑每个点的局部几何特征，对所有簇中的每个点进行计算；请注意，此引入过程与某些图像处理算法（例如 [57]、[58]）中使用的二值化方法完全不同，二值化方法的工作原理类似于优化框架，有助于在突出显著区域的同时消除非显著区域，相反，此过程通过计算其首先到达每个播种点的概率来帮助将全局稀有度值分配给每个未播种点；图3显示了使用随机游走排名算法对所有集群中的每个点进行集群级全局稀有性细化的示例；模型（a）表示初始集群级全局使用方程式（6）中的公式计算的稀有度结果；模型（b）显示了使用随机游走排名方法获得的点级全局稀有度细化；与模型（a）相比，模型（b）的视觉效果更好，牛的嘴角等模型细节得到了更好的保留，同时有效地消除了对牛身体的不准确检测结果。

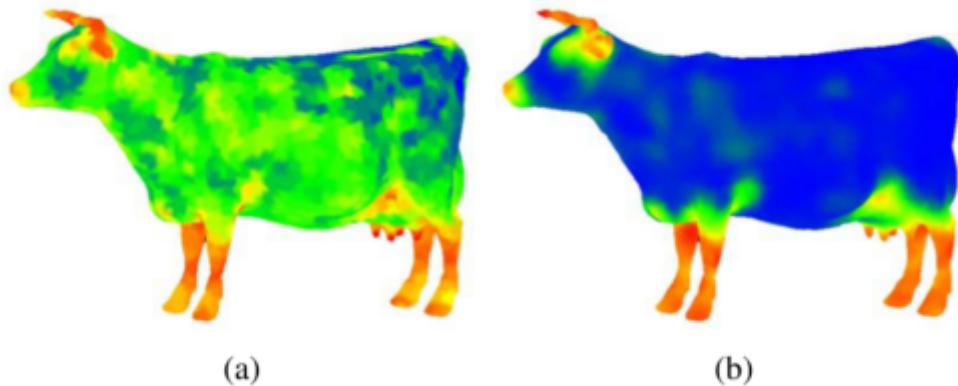


图 3. 初始 (a) 集群级全局稀有度结果与 (b) 使用随机游走排序方法获得的点级全局稀有度细化结果的比较

使用为方程式（6）中的每个集群计算的初始全局稀有度值，我们可以将点云中的一些簇根据它们的簇级全局稀有度值标记为显著簇或非显著簇；请注意，由于一个簇包含许多点，因此最接近显著或非显著簇中心的点将被标记为显著或非显著播种点，以便在所提出的算法中进行后续计算；例如，如果一个簇的全局稀有度值高于阈值 th_1 或低于阈值 th_2 ，则最靠近该簇中心的点将被标记为显著播种点或非显著播种点，而属于该簇的其他点将被标记为未播种点；对于全局稀有度值既不高也不低的簇，簇中的所有点将被分类为非播种点；所提出算法中使用的 th_1 和 th_2 的阈值描述为：

$$th_1 = \text{mean}(G) + \frac{1}{2}\Delta$$

$$th_2 = \text{mean}(G)$$
(7)

其中 $\Delta = \text{最大值}(G) - \text{平均值}(G)$ 。

当选择播种点时，我们从每个非播种点开始给出一个随机游走者，并计算这个随机漫步者首先到达每个播种点的概率；例如，如果播种点的数量设置为 K ，那么将为每个非播种点构建 K 元组向量，以记录从该位置开始的随机步行者首先到达每个 K 个播种点的概率；然后，对于每个非播种点，我们根据其 K 元组向量选择最可能的目的地，以获得点级别的全局稀有性细化；后面将详细描述如何应用随机游走排序算法来帮助对所有集群中的每个点引入集群级别的全局稀度细化。

根据簇级全局稀有性结果选择播种点后，用节点 V 和边 E 构建图 $G = (V, E)$ ，其中 V 对应点云中的所有点， E 代表两点之间的连接；不同于二维图像可以根据任意两个像素的特征和距离建立连接[56]，在 3D 模型中，由于数据结构的不同，任意两点之间的连接要复杂得多[59]；例如，给定一个 3D 点云 P ， P 中始终存在一个点 p_i 和一个点 p_j ，它们不能从同一视点看到，这表明点 p_i 的全局稀有度值不太可能对点 p_j 产

生影响，所以我们需要偏置随机游走器以避免沿着点 p_i 和点 p_j 之间的连接移动；按照[60]中提出的方法，我们将此连接的权重设置为零，以抑制随机游走者沿边缘移动；考虑到判断两个不同点之间的连接是否应该设置为零的复杂性，我们简化了过程，只计算点 p_i 与其 k -最近邻点之间的连接，并为点 p_i 和它之间的连接分配权重为零；点云 P 中的其他点。通过选择较小的 k ，我们可以假设点 p_i 对其 k 最近邻点有很大的影响。

点 p_i 和 k 最近邻点 p_j 之间的连接通过使用每个点的局部几何特征计算的权重矩阵来量化，如公式(8)所示：

$$w_{ij} = \exp\left(-\frac{\|FPFH(p_i) - FPFH(p_j)\|^2}{\sigma_1^2}\right) \quad (8)$$

其中 w_{ij} 表示点 p_i 和 p_j 之间的连接； σ_1 是一个控制常数，在我们的实验中设置为 3.7；请注意，在所提出的方法中，由 FPFH 描述符表示的每个点的局部几何特征用于计算权重矩阵，这与在 2D 图像中使用 CIELab 颜色描述符 [56] 不同；在构建权重矩阵之后，可以进一步构建拉普拉斯矩阵 L ，如 [56] 中所述。 p^s

令 p^s 表示属于类 s 的点的概率向量，它由两个子集 p_u^s 和 p_{seeds}^s 组成； p_u^s 表示未播种点的概率向量，而 p_{seeds} 显示点云中的播种点，固定值为 0 或 1；优化的 p^s 可以通过最小化 Dirichlet 积分来实现 [56]；按照 [61] 中提出的方法，添加了方程 (9) 中的拟合约束，以帮助将 Dirichlet 积分限制为接近初始集群级全局稀有度值。

$$Dir[p^s] = \frac{1}{2}(p^s)^T L(p^s) + \frac{\eta}{2}(p^s - Y)^T(p^s - Y) \quad (9)$$

其中 η 是控制参数， Y 表示继承初始集群级全局稀有度值的逐点指示向量；

通过将 $Dir[p^s]$ 相对于 p_u^s 的微分设置为零，除播种点之外的所有簇中的点的 p_u^s 优化解可以计算为 [61]：

$$p_u^s = (L_u + \eta I)^{-1}(-B^T p_{seeds}^s + \eta Y_u^s) \quad (10)$$

使用上面提到的随机游走排序方法，我们通过考虑每个点的局部几何特征，对所有簇中的每个点引入初始集群级全局稀有性细化。

C. 优化框架

在得到每个点的局部显著性值和全局稀有度值之后，我们需要将这两个线索整合起来，得到最终的显著性检测结果；现有的显著性检测算法通常应用不同显著性线索的简单求和或线性组合来获得最终的显著性检测结果[65]；然而，由于权重系数通常是固定的，线性组合不能为所有数据提供适当的显著性检测结果，权重系数的组合可能适用于数据 A，但无法为数据 B 提供良好的结果，因此需要更自适应的集成机制；为了解决这个问题，我们利用不同显著性线索之间的内在关系，提出了一种新的优化框架。

受 [66] 中提出的方法的启发，我们将集成不同显著性线索的问题建模为点云中所有点的显著性值的优化；方程 (11) 中的损失函数将显著点值 1 和非显著点值 0 分配给损失函数，因为人类视觉系统倾向于将项目组合在一起并获得统一的显著性检测结果[67]，而不是孤立的突出点；

$$L = \sum_{k=1}^2 \sum_{i=1}^N \left(\frac{s_i^2}{Z_k(p_i)} + Z_k(p_i)(1 - s_i)^2 \right) + \sum_{j \in R} W_{ij} (s_i - s_j)^2 \quad (11)$$

其中 $Z \in \{F, G\}$ 与 $Z1(p_i)$ 表示放大的局部区别值 $F(p_i)$, $Z2(p_i)$ 表示点级全局稀有度细化 $G(p_i)$ 。 N 等于点云中点的个数, R 表示点 p_i 的相邻点; 注意, 在实际操作中分母可能等于0, 所以如果 $F(p_i)$ 和 $G(p_i)$ 的值接近于0, 则在 $F(p_i)$ 和 $G(p_i)$ 上加上一个很小的常数0.1。

由于局部独特性和全局稀有性对人类视觉系统都很重要, 因此损失函数考虑了这两个特征的对称性, 可以证明损失函数是一个带有 s_i 的凸函数, 因此可以通过对 s_i 取导数并使其为零来最小化损失函数; 通过最小化损失函数, 损失函数中的第一项鼓励具有较高局部显著性值和较高全局稀有性值的点获得较高的显著性值, 并抑制具有较低局部显著性值和较低全局稀有性值的点; 第二项鼓励最终显著性检测结果的平滑性; 对于局部显著性值较高而全局稀有度值较低的点 p_i 或局部显著性值较低而全局稀有度值较高的点 p_j , 将考虑其他点的影响来计算它们的显著性值。

损失函数中的权重系数 W_{ij} 表示 k 最近邻点 p_j 对点 p_i 的加权影响, 定义为:

$$W_{ij} = \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_2^2}\right) \quad (12)$$

其中 σ_2 是控制常数, 设置为 0.02 和 $\|\cdot\|$ 是点 p_i 和 p_j 之间的 L2 距离。

IV. 实验结果

在本节中, 我们将我们的方法与几种最先进的基于网格和基于点云的显著性检测算法进行比较, 以展示我们的性能; 此外, 还提供了消融研究 (ablation study), 以证明局部区别计算、全局稀有度计算和所提出的优化框架的有效性; 最后, 我们讨论了所提出的算法如何应用于图形应用, 例如兴趣点检测、视点选择和网格简化。

A. 显著性检测结果评估

由于技术的快速发展, 已经提出了许多显著性检测算法; 例如, Yun 和 Sim [68] 提出了一种用于彩色 3D 点云数据的新算法, 以及 Nouri 等人[69] 提出了一种用于 3D 彩色网格数据的方法; 此外, Jeong 和 Sim [10] 提出了一种算法来处理半规则网格的显著性检测, Lau 等人[70]设计了一种算法来预测网格的触觉显著性; 虽然这些算法都取得了相当满意的显著性检测结果, 但输入数据类型与提出的算法有很大不同, 因此我们不将提出的算法与这些方法进行比较; 在本节中, 将所提出的算法与几种最先进的基于网格或基于点云的显著性检测算法进行比较, 以展示其性能。

所提出的方法与[25]、[62]中提出的基于点云的算法、[11]、[63]中提出的基于网格的算法以及[64]中收集的地面实况数据的比较显示在图 4, 从图中可以看出, [25]中提出的方法在检测章鱼末端等显著区域时更有用, 但在章鱼身体和手掌处失败, 这可能是由于使用欧几里得空间进行聚类, 导致平面上的聚类描述符发生了显著变化; [62]中提出的方法产生了相当正确的显著性结果, 但也突出了一些非显著区域, 例如泰迪模型的轮廓和手掌; [63]中提出的方法在花瓶模型上捕获了良好的显著性检测结果, 但未能检测到人体模型上的显著性区域, 这可能是由背景查询的选择引起的, 因为胸部的补丁彼此非常相似, 如果选择属于这些区域的一个补丁作为背景, 则与所选补丁相关性较高的其他补丁更有可能被视为背景; 对于[11]中提出的方法, 它可以很好地检测显著区域, 但更有可能提供具有不同显著值的相似区域, 这可能是由于模型缺乏空间对称性; 与上述方法相比, 该方法的显著性检测结果更加一致, 并且在大多数情况下可以检测到模型的更好的显著区域, 例如 Igea 模型在人体模型的面部以及四肢和胸部, 然

而，所提出的算法还需要进一步改进，因为它经常无法检测到圆柱形状的正确显著区域，例如椅子模型的腿和章鱼模型的脚。



图 4 不同显著性检测结果对比从上到下：使用所提出的方法计算的显著性检测结果，[25]和[62]中报道的基于点云的显著性检测结果，[11]，[63]中报道的基于网格的显著性检测结果和 groundtruth[64]中报告的模型；请注意，蓝色表示低显著值，红色表示高显著值

B. 消融研究

为了更好地说明局部独特性和全局稀有性特征的有效性，我们进行了消融研究，我们报告了仅使用局部独特性特征或全局稀有性特征来帮助点云显著性检测的结果；此外，我们还提供了使用线性组合和提出的优化框架之间的比较，以显示使用优化框架更好的集成结果；图5显示了仅使用局部显著性特征和全局稀有性特征的显著性检测结果，以及使用线性组合和提出的优化框架的集成结果。

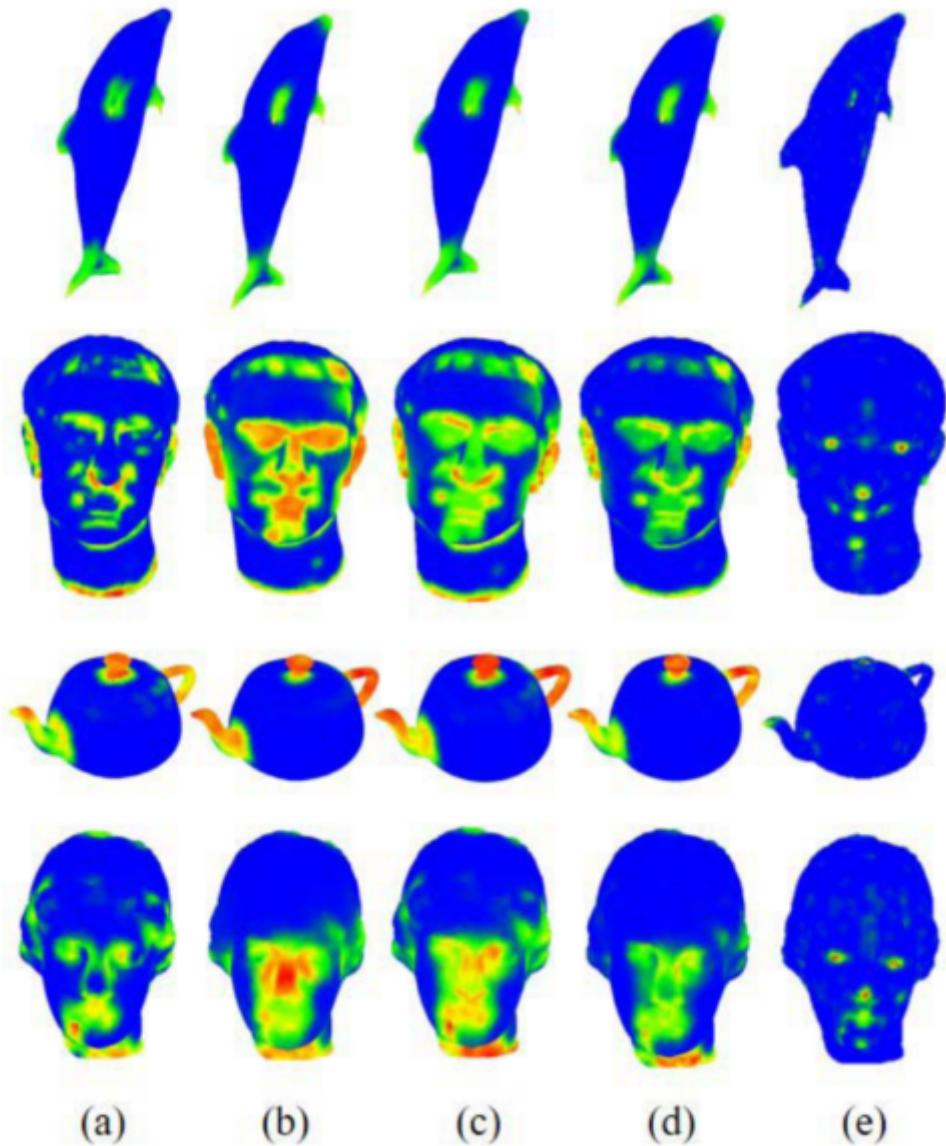


图5 使用不同特征和不同集成算法的显著性检测结果；(a)列显示使用局部显著性特征的显著性检测结果，而(b)列表示使用全局稀有性特征的显著性检测结果；(c)列显示了使用线性组合的显著性整合结果；(d)列显示了使用建议的优化框架的显著性整合结果；(e)列展示了[64]中报告的模型的groundtruth

根据图5，与全局稀有特征相比，使用局部显著性特征的显著性检测结果突出了较小的区域；例如图中第四行，局部特征主要突出嘴巴、眼睛和鼻子区域的轮廓，未能突出整个鼻子，而全局稀有特征则突出整个面部区域。但在第二行，全局稀有度特征也突出了包括脸颊在内的面部区域，这在地面实况数据中并不显著。这表明单独使用局部显著性特征或全局稀有性特征很难获得令人满意的显著性检测结果，这两个特征是互补的。根据实验结果，局部区别特征无法检测到海豚的头部，而全局稀有特征可以正确检测到海豚的头部。如果使用线性组合来整合局部独特性和全局稀有性特征，它将无偏差地组合两个特征。例如，在第三列的第一行和第三行中，可以观察到线性组合不能消除由局部显著性特征检测到的不正确的显著区域，例如海豚的头部和茶壶身体的斑点。它只是直接添加相应区域的显著值，因此这些不正确的显著区域将保留在积分结果中并影响视觉性能。当我们使用优化框架得到的积分结果与线性组合进行比较时，例如第四列的第一行和第三行，可以观察到优化框架保留了正确的显著区域，例如海豚的头部并消除不正确的显著区域，提供更好的视觉性能。上述实验结果证明了结合局部独特性和全局稀有性特征的优越性，并证明了利用所提出的优化框架的有效性。

除了定性分析外，我们还提供定量评估以使说明更加清晰。我们首先使用[71]中提出的显著性误差(SE)来帮助评估。这个指标非常直观，较低的SE代表更好的显著性检测结果。该指标定义为：

$$SE = \sum_{i=1}^N |S(i) - S_{gt}(i)|$$

其中 N 表示模型中点的总数，而 $S(i)$ 表示使用算法计算的点 i 的归一化显著值， $S_{gt}(i)$ 是 [64] 中报告的归一化 Schelling 分布值，用作 groundtruth。

我们还采用了在 2D 图像显著性评估中广泛使用的评估指标 CC 分数。CC 分数越高，表示算法的检测精度越高；定量评价结果见表一。

根据表1所提出的优化框架在所有四个模型中取得了最低的 SE 分数。在海豚模型和茶壶模型中，优化框架也获得了最高的CC分数。对于Man模型和Igea模型，CC分数排名第三位，但分别非常接近 0.22 和 0.30。这可能是因为这两个模型的局部区分度结果很不理想，使得优化框架很难获得更好的CC分数。

	Column (a)		Column (b)		Column (c)		Column (d)	
	SE	CC	SE	CC	SE	CC	SE	CC
Dolphin	17.1	0.44	31.4	0.40	23.9	0.43	11.0	0.45
Man	36.6	0.15	73.0	0.22	54.3	0.22	32.0	0.19
Teapot	31.0	0.34	28.1	0.35	29.4	0.35	28.0	0.36
Igea	23.7	0.25	27.3	0.30	25.2	0.30	16.9	0.29

Column (a) to (d) represent the column (a) to (d) in Fig. 5.

表1：消融研究中的定量评估

C. 兴趣点检测

3D兴趣点是指在其位置上具有独特性的点[76]；它可以提供语义上重要的局部特征，并且对于许多图形应用程序来说都是必不可少的，例如网格分割 [77] 和配准 [78]；为了更好地说明所提出算法的性能，我们将其应用于帮助 3D 兴趣点检测，并根据 [76] 中提出的 3D 兴趣点检测benchmark评估兴趣点检测结果；benchmark由两个数据集组成，数据集 A 包含 24 个三角形网格，使用 23 个参与者标记的兴趣点数据，而数据集 B 包含 43 个三角形网格（包括数据集 A 中的所有模型）并使用 16 个参与者标记的兴趣点数据，Dutagaci 等人开发了一个应用程序，要求人类参与者在三角形网格上标记兴趣点； Dutagaci 等人使用人工标记的点构建了groundtruth数据，并将其与基于假阴性错误 (FNE)、假阳性错误 (FPE) 和加权未命中错误 (WME) 的六种不同的兴趣点检测技术进行了比较；六种兴趣点检测技术包括：显著点 [49]，显著网格[6]、SD-corners [75]、3D-Harris [72]、3D-SIFT [74] 和基于 HKS 的兴趣点检测方法 [73]；评估矩阵的计算简述如下：

1) 假负误差：令 G 表示模型 M 的真实点；对于兴趣点 $g \in G$ ，点 g 的测地线邻域计算为：

$$C_r(g) = \{p \in M | d(g, p) \leq r\}, \quad (13)$$

其中 $d(g, p)$ 表示点 g 和点 p 之间的测地距离，而 r 表示测地线邻域的半径。

我们将 D 表示为通过评估方法检测到的兴趣点集，如果存在不靠近 G 中任何其他点的点 $a \in D \cap C_r(g)$ ，则正确检测到点 g [76]。令 N_c 为 G 中正确检测到的点数， N_g 为 G 中点的总数。FNE 计算为：
 $FNE = 1 - N_c/N_g$ 。

2) 误报错误：对于每个正确检测到的点 $g \in G$ ，都有一个对应点 $a \in D$ 。因此 D 中没有 G 中对应点的所有点都被视为误报。令 N_d 表示通过评估方法检测到的兴趣点的数量，FPE 定义为：
 $FPE = 1 - N_c/N_d$ 。

3) 误码错误：对于半径为 r 的测地线邻域，地面实况点可能由 n_i 个人类参与者标记，因此 WME 定义为：

4)

$$WME = 1 - \frac{1}{\sum_{i=1}^{N_g} n_i} \sum_{i=1}^{N_g} n_i \delta_i \quad (14)$$

和：

$$\delta_i = \begin{cases} 1 & \text{if } g_i \text{ is detected} \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

其中 g_i 属于 ground-truth 点。

FNE、FPE 和 WME 都是介于 0 和 1 之间的度量；算法执行的 FNE、FPE 和 WME 越低，该算法在兴趣点检测方面的性能越好。

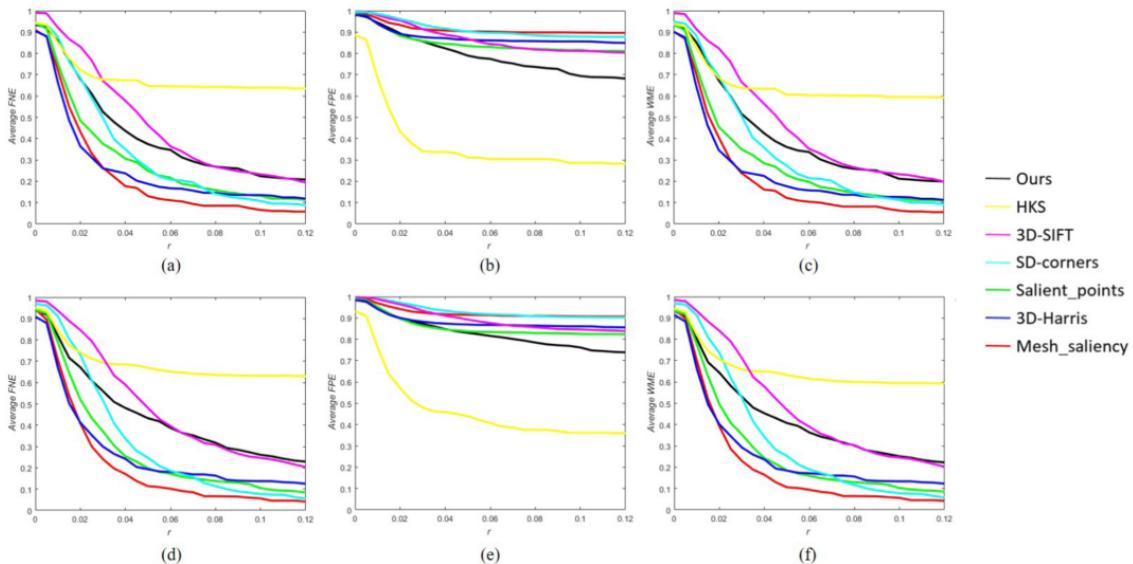


图 6. 使用 FNE、FPE 和 WME 在兴趣点检测基准上评估我们的算法与其他兴趣点检测方法的性能的比较结果。前三图为数据集 A 的评估结果，后三图为数据集 B 的评估结果。不同颜色的线表示每种兴趣点检测算法对 groundtruth 兴趣点的评估分数， $\sigma = 0.03$ 和 $n = 8$ 。

在 [25] 和 [1] 之后，我们的兴趣点被定义为阈值上的局部最大值，该阈值被计算为所有局部最大值的平均值；图 6 展示了我们的方法对使用 FNE、FPE 和 WME 随着半径 r 增加的兴趣点检测基准的评估结果，并与 [76] 中提到的其他六种兴趣点检测算法进行了比较；根据图 6，与其他方法相比，基于 HKS [73] 的方法的 FPE 显著降低，因为基于 HKS [73] 的兴趣点检测方法发现的兴趣点非常少，但精度更高；图 6 中的结果还表明，我们的方法比除 HKS [73] 之外的其他方法实现了更低的 FPE，而 FNE 略高于某些方法，但在这两种方法中仍然与 3D-SIFT [74] 算法相当两个数据集，由于一种方法很难同时实现较低的 FPE 和较低的 FNE，因此假阳性和假阴性值之间的平衡需要量化；根据[25]，我们使用以下公式量化余额：

$$F_{\beta} = (1+\beta^2) \frac{(1-FPE)(1-FNE)}{\beta^2(1-FPE) + (1-FNE)} \quad (16)$$

其中 $(1-FPE)$ 表示精度， $(1-FNE)$ 是召回率； F_{β} 表示 FPE 和 FNE 之间的量化平衡， β 用作准确率和召回率之间的权重系数。

使用等式 (16) 中的公式，我们计算了前面提到的所有兴趣点检测方法的量化平衡值；量化的余额值如表2所示，根据表2、还可以观察到，当召回率与精度同等重要时，HKS算法的性能最好，而当召回率是精度和容差半径的两倍时，所提算法的实验值最好；数据集 A 设置为 $r = 0.06$ 和 $r = 0.09$ ，数据集 B 设置为 $r = 0.09$ ；虽然 HKS 算法取得了相当令人满意的结果，但其代价是较高的假阴性错误和加权未命中错误，这可能不是适用于一些需要较高召回率的应用；除了进行兴趣点检测评估外，图 7 提供了该方法和其他算法获得的兴趣点检测结果示例，与 [25] 中报告的基于聚类显著性的兴趣点检测结果相比，提出的方法检测膝盖处更好的兴趣点和Armadillo模型的故事。

	Dataset A						Dataset B					
	$r=0.03$		$r=0.06$		$r=0.09$		$r=0.03$		$r=0.06$		$r=0.09$	
	$F_{1.0}$	$F_{2.0}$										
Ours	0.25	0.36	0.33	0.47	0.40	0.54	0.20	0.30	0.28	0.42	0.35	0.50
Salient Points [49]	0.23	0.37	0.28	0.46	0.30	0.49	0.22	0.36	0.27	0.46	0.28	0.48
3D-Harris [72]	0.21	0.37	0.24	0.42	0.25	0.43	0.20	0.35	0.22	0.40	0.23	0.42
HKS [73]	0.43	0.35	0.47	0.39	0.48	0.40	0.38	0.33	0.44	0.38	0.46	0.40
Mesh Saliency [6]	0.15	0.29	0.17	0.34	0.18	0.35	0.13	0.27	0.15	0.31	0.16	0.32
3D-SIFT [74]	0.13	0.21	0.25	0.39	0.30	0.47	0.10	0.17	0.20	0.34	0.25	0.41
SD-corners [75]	0.11	0.21	0.18	0.34	0.21	0.39	0.08	0.17	0.15	0.29	0.17	0.33

$\sigma = 0.03$ is used to compute the F_{β} value.

表2.兴趣点检测的 F_{β} 值

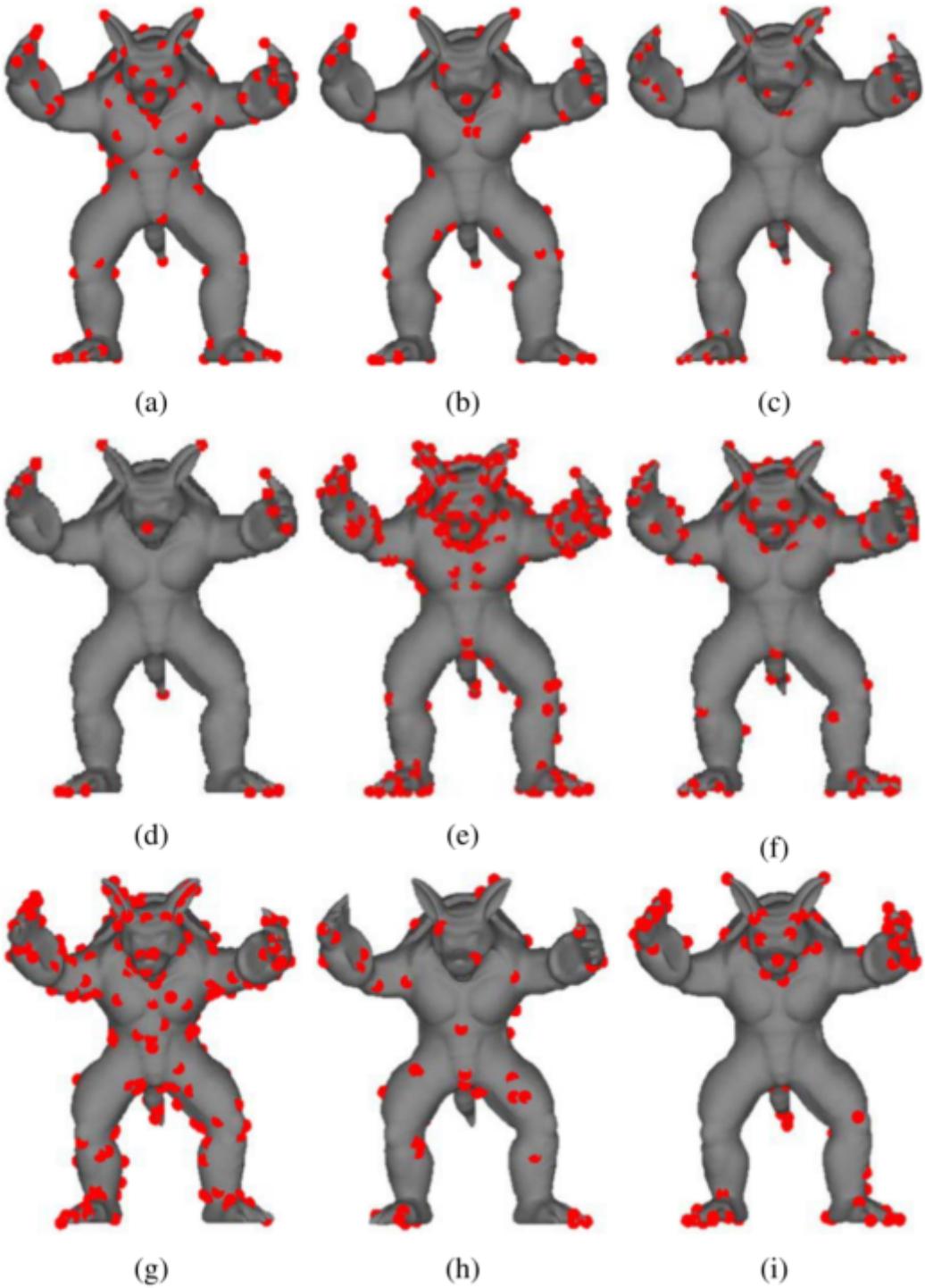


图7：通过所提出的方法和其他算法计算的兴趣点检测结果示例；请注意，(a) 基于 $\sigma = 0.03$ 和 $n = 2$ 呈现；(a) Ground truth；(b) 我们的结果；(c) 聚类显著性[25]；(d) HKS [73]；(e) 网格显著性 [6]；(f) Salient points [49]；(g) Scale dependent [75]；(h) 3D-SIFT [74]；(i) 3D Harris [72]

D. 视点选择

随着遥感技术的发展，3D模型的获取更加方便；越来越多的3D模型越来越重视自动化选择可以最大程度地阐明模型最重要特征的最佳视点；最近，许多论文都解决了这个问题，但这些算法中的大多数都无法处理点云；因此，我们应用所提出的方法来帮助点云的自动视点选择。

对于给定的点云，视点选择是自动确定最能描述点云的信息量最大的视点[12]；因此，视点选择的目的是找到一个可以最大化点云上可见点的总显著值的视点，我们将在后面描述视点选择的整个过程。

最初，候选视点是通过对一个限定点云模型的球体进行均匀采样来生成的[79]，在[12]之后，半径被设置为紧密边界球体长度的两倍；点云模型的候选视点数量约为 2600 个。对于点云模型 P ， S 表示模型 P 中每个点的显著性检测结果；给定视点 v ，令 $F(v)$ 表示从视点 v 可见的点集。根据 [6]，我们定义视点 v 的显著性值总和为：

$$U(v) = \sum_{p_i \in F(v)} S(p_i) \quad (17)$$

其中 $U(v)$ 表示视点 v 的总显著值， p_i 表示从视点 v 可见的点。

对于每个候选视点，计算 $U(v)$ 。然后我们找到最佳视点 $v_p = \text{argmax } U(v)$ 。请注意，我们没有执行梯度下降法来找到 [6] 中描述的局部最大值，因为候选视点的数量已经足够了。图8给出了基于所提出方法的视点选择结果、[80] 中报告的语义驱动视点选择结果和 [12] 中报告的基于感兴趣表面区域的视点选择结果之间的比较。

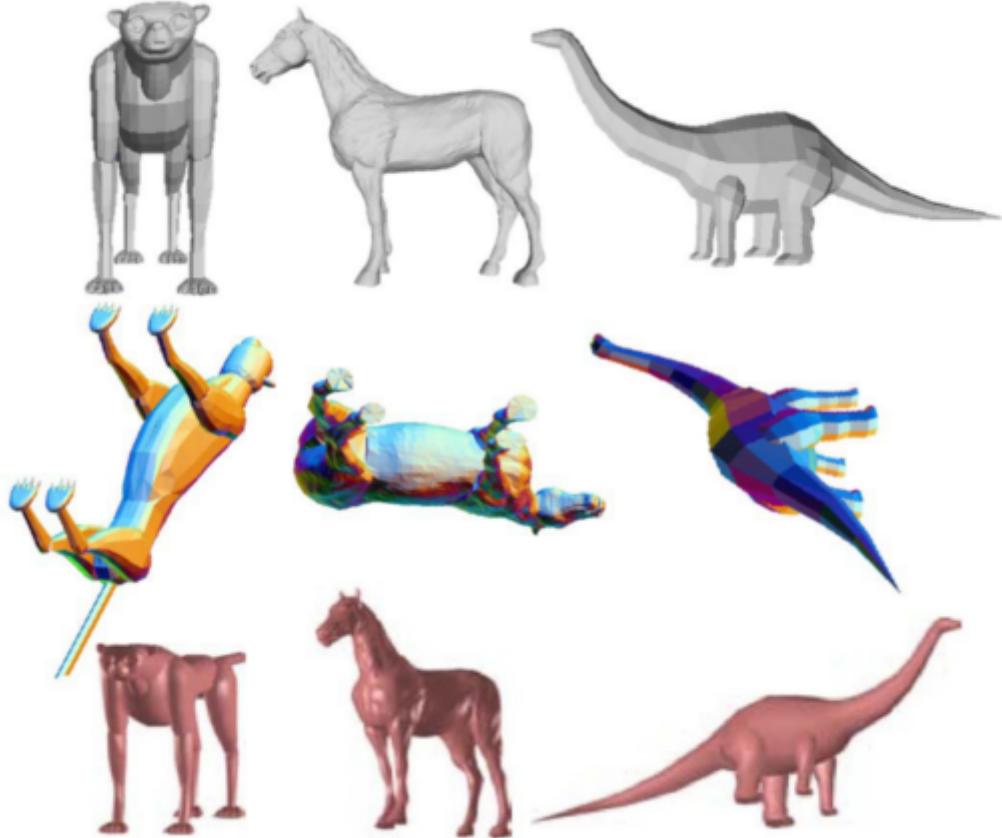


图 8：视点选择结果；对比从上到下，所提出方法的结果、[80]中报告的结果和[12]中报告的结果

根据图8，所提出的方法将狗的面部检测为模型上最显著的区域，因此选择的视点主要集中在面部区域；尽管与 [80] 和 [12] 中报告的结果相比，我们选择的视点未能展示模型的故事，它很可能被人类视觉系统采用，因为人类很可能会被动物的脸所吸引；对于马模型，所提出的算法和[12]选择的视点更有助于识别模型，因为这两种算法都可以更好地描述模型的轮廓；对于恐龙模型，三种算法选择的视点具有相当的可比性。

E. 网格简化

高细节模型呈现更逼真的场景，通常出现在游戏和视频中。然而，它们通常需要更大的存储空间并增加计算的复杂性[81]；由于简化高度详细的 3D 模型对计算机图形学领域至关重要，我们应用所提出的方法来帮助简化网格。

显著性检测结果识别出 3D 模型上视觉上的显著区域，因此它可以指导网格简化过程，以帮助保留重要的细节；许多网格显著性检测算法已被应用于帮助网格简化，例如，Limper 等人 [83] 提出了一种使用局部曲率熵的网格显著性检测算法，并应用他们的算法来帮助网格简化；在[11]中，提出了一种新算法，该算法使用网格数据的光谱属性，然后使用检测到的显著区域来指导简化过程；Zhao等人 [14]还提出了一种基于显著性检测的简化算法，该算法将更多的点分配给显著区域；虽然上面提到的论文都应用了网格显著性估计来帮助网格简化并取得了很好的效果，但是它们的代码很少可用，并且它们使用的模型彼此之间存在很大差异，因此很难对我们的算法及其方法在网格简化结果进行比较；为了评估我们方法的有效性，我们修改了 Garland 和 Heckbert [82] 提出的基于二次曲线的简化方法 (QEM)，方法是使用检测到的显著性值对二次曲线进行加权，QEM 方法通过重复收缩具有最小二次误差的顶点对来简化网格 [6]；下面描述网格简化的整个过程。

给定网格模型 M 上的顶点 $v = [v_x, v_y, v_z, 1]^T$ ，令 F 表示入射到顶点 v 的三角形平面集合， S 为网格模型 M 的显著性检测结果。对于平面 $f \in F$ ，由方程 $ax + by + cz + d = 0$ 和 $a_2 + b_2 + c_2 = 1$ 定义，平面 f 可以描述为 $f = [a, b, c, d]^T$ 。按照[6]中的方法，网格模型M的显著性结果S用于指导简化收缩过程；顶点 v 的误差定义为：

$$\Delta(v) = \Delta([v_x, v_y, v_z, 1]^T) = W(v) \sum_{f \in F} (f^T v)^2, \quad (18)$$

其中 $W(v)$ 是在顶点 v 处放大的显著性值 $S(v)$ ，放大算子指定为：

$$W(v) = \begin{cases} \lambda S(v) & \text{if } S(v) \geq \alpha \\ S(v) & \text{otherwise} \end{cases} \quad (19)$$

其中 λ 是放大参数， α 是控制要放大的显著点范围的阈值。请注意，在本文中， λ 设置为 100， α 设置为论文 [6] 中的前 30% 显著性值。

对于模型中的每个顶点，我们使用公式 (18) 中的公式计算其误差，然后，如果两个点 (v_1, v_2) 是边，我们在初始化阶段选择点对进行收缩；在每对收缩期间，都需要考虑收缩成本。对于给定的收缩 $(v_1, v_2) \rightarrow \bar{v}$ ，我们将收缩误差定义为顶点 v_1 和顶点 v_2 的误差之和。为了执行这个收缩，需要找到一个最小化收缩误差的位置。通过求解 $\partial\Delta/\partial x = \partial\Delta/\partial y = \partial\Delta/\partial z = 0$ ，可以得到 \bar{v} 的最优位置。

在计算出每对的最优收缩目标 \bar{v} 以及收缩成本后，所有对都使用收缩值的成本按升序排序，成本最小的对 (v_1, v_2) 被收缩。然后更新所有涉及顶点 v_1 的对的成本，并重复整个收缩过程，直到剩余顶点的数量满足需要；图9展示了使用不同方法的简化结果的比较。从第一行可以看出，与 QEM [82] 方法相比，由显著性引导的简化结果（第三、第四和第五列）在龙的面部和脚部保留了更好的细节；更具体地说，我们的方法保留了更好的面部和脚部细节，并且分布在龙身上的顶点数量最少；根据第二行，与[6]中提出的网格显著性算法相比，所提出的方法在龙的眼睛、嘴巴和角上保留了更多的顶点，并且在不使用模型的拓扑结构的情况下达到了与[50]相当的结果。

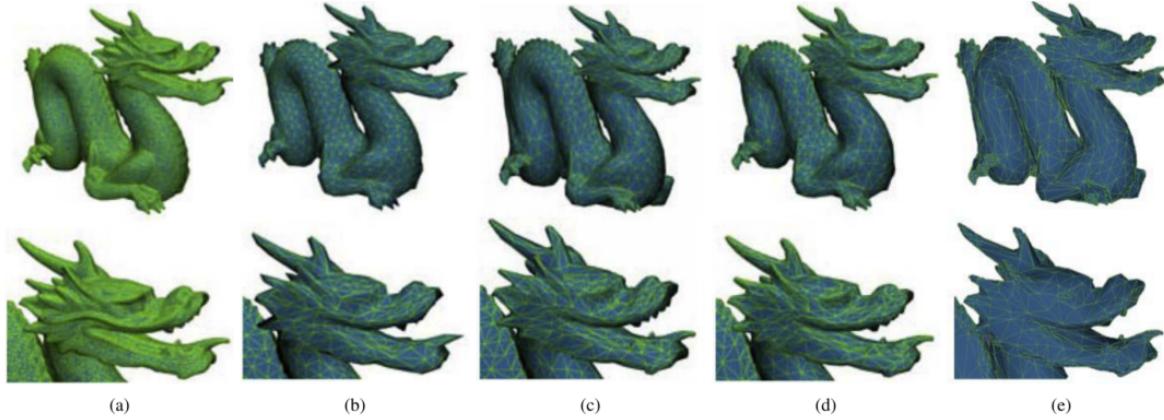


图9：简化结果对比[50]中报告了左侧四列；从左到右：龙模型（30k个顶点），使用QEM[82]算法生成的简化结果（2500个顶点），由[6]（2500个顶点）和[50]（2500个顶点）引导的简化结果，使用所提出方法的简化结果（2500个顶点）；(a) 原始模型；(b) Qslim [82]；(c) 网格显著性[6]；(d) 全局稀有性[50]；(e) 我们的结果

V. 结论

本文提出了一种利用局部独特性和全局稀有性线索的点云显著性检测方法，该线索得到了心理学证据的支持；与其他点云显著性检测算法相比，主要贡献在于在采用随机游走排序方法时，考虑到每个点的局部几何特征，对所有聚类中的每个点引入簇级的全局稀有性细化，由于点云中的噪声，超体素分割可能不准确；随机游走排序方法不依赖于超体素分割，可以减少分割不精确对全局稀度计算的影响，由于局部几何特征代表了相邻点之间的关系，因此可以在随机游走过程中作为指导来生成更精细的全局稀有结果；此外，我们还提出了一种自适应优化框架，该框架可以考虑不同显著性线索之间的内在关系，有效地结合不同显著性线索，实现最先进的显著性检测结果；为了证明该方法的有效性，我们将其应用于三个计算机图形学应用：兴趣点检测、视点选择和网格简化，这些实验证明了我们的性能。