

采用动态分组和投票机制的三维模型兴趣点提取方法

舒振宇^{1,2)}, 但文字^{1)*}, 辛士庆³⁾

¹⁾ (浙江大学软件学院 宁波 315100)

²⁾ (浙大宁波理工学院计算机与数据工程学院 宁波 315100)

³⁾ (山东大学计算机科学与技术学院 青岛 266237)

(22151059@zju.edu.cn)

摘要: 三维模型的兴趣点提取是计算机图形学中的一个基本问题, 针对该问题提出一种采用投票机制和动态分组兴趣点的提取方法, 主要包括 3 个模块. 其中, 顶点编码器模块融合三维模型上的顶点的邻域和全局信息, 生成顶点语义序列; 概率分布预测网络模块将顶点语义序列映射成投票概率序列和置信度序列, 并生成兴趣点概率分布; 动态分组模块通过设定不同的概率阈值, 从概率分布上进行兴趣点分组提取. 所提方法在 SHREC2011 和 KeyPointNet 数据集上进行了实验, 采用 FNE, FPE, BHD 和 CD 评价指标进行比较, 结果表明, 与传统方法相比, FPE 至少减少 0.2, FPE 至少减少 0.18; 与现有机器学习的方法相比, BHD 平均减少 0.011, CD 平均减少 0.002, 在兴趣点提取效果方面有了显著的提升.

关键词: 兴趣点; 语义序列; 分组; 三维模型

中图分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.202*.2023-00594

Extraction of Points of Interest on 3D Models Based on Voting and Dynamic Grouping

Shu Zhenyu^{1,2)}, Dan Wenyu^{1)*}, and Xin Shiqing³⁾

¹⁾ (School of Software, Zhejiang University, Ningbo 315100)

²⁾ (School of Computer and Data Engineering, NingboTech University, Ningbo 315100)

³⁾ (School of Computer Science and Technology, Shandong University, Qingdao 266237)

Abstract: Extracting points of interest on 3D model is a basic problem in computer graphics, Aiming at this problem, a method using voting mechanism and dynamic grouping is proposed, which mainly includes 3 modules. Among them, the vertex encoder module fuses neighborhood and global information of vertices to generate vertex semantic sequences; The probability distribution prediction network module maps the vertex semantic sequence into a probability voting sequence and a confidence sequence to generate a probability distribution of points of interest; The dynamic grouping module groups and extracts points of interest from the probability distribution by setting different probability thresholds. The proposed method is tested on SHREC2011 and KeyPointNet data sets and use FNE, FPE, BHD and CD as evaluation metrics. Compared with traditional methods, FNE is reduced by at least 0.2 and FPE is reduced by at least 0.18; compared with existing machine learning methods, BHD is reduced by 0.011 on average and CD is reduced by 0.002 on average, the proposed method has a significant improvement for extracting points of interest.

Key words: points of interest; semantic sequence; grouping; 3D model

三维模型兴趣点(points of interest, POI)检测是计算机图形学中的一个基本问题,旨在自动检测和抽取三维模型上人眼感兴趣的顶点.在三维模型上,任意顶点是否是 POI 是一个主观性的问题,并没有明确的数学定义来约束,而是由人的主观认知来决定.不同的人对 POI 的主观认知存在细微的差别,但大体上是一致的.许多图形学领域的工作中都需要进行 POI 检测,如三维模型分类分割^[1-3]、三维形状检索^[4-6]、三维动画建模^[7]、人脸表情识别^[8]等.随着图形学技术的实际应用与相关领域研究的进步,人们对三维 POI 检测方法的准确性与效率提出了越来越高的要求.

1 相关工作

三维模型 POI 检测的方法可以分为基于特征的方法和基于学习的方法 2 类.基于特征的方法依赖于模型表面的曲率、局部形状变化、几何信息等特征来进行 POI 检测,包括计算模型表面上的点云密度、模型网格上的高斯曲率、平均曲率、局部坐标等,然后通过各种筛选和聚类方法识别 POI;这类方法虽然具有较高的准确性和可解释性,但需要手动选择特征参数和分类器参数,并且在不同模型之间的通用性和稳健性难以保证.基于学习的方法使用深度学习模型,从大量的三维模型数据中学习 POI 的特征来实现自动化的 POI 检测,这种方法具有强大的泛化能力和鲁棒性,并且可以快速检测新三维模型上的 POI.

1.1 基于特征的方法

在三维模型上,面片和顶点的几何特征描述符是基于特征的方法进行 POI 检测的重要手段.几何特征描述符用数学语言的方式描述三维模型的几何特征,已被广泛使用,表现出色的几何特征描述符有形状直径函数(shape diameter function, SDF)^[9]、平均测地线距离(average geodesic distance, AGD)^[10]、高斯曲率(Gaussian curvature, GC)^[11]、全局特征尺度不变热核特征(scale-invariant heat kernel signatures, SIHKS)^[12]、波核特征(wavelet kernel signature, WKS)^[13]等.

Sipiran 等^[14]提出一种基于 Harris 算子的三维模型 POI 检测器,用自适应技术确定顶点的邻域,然后在邻域内计算顶点的 Harris 响应;该方法对多种转换具有鲁棒性,然而遗漏 POI 和错误 POI 较多.

Dutagaci 等^[15]为三维模型构建手工标记的 POI,并提出假阳性误差(false positive error, FPE),假阴性误差(false negative error, FNE)和加权遗漏误差(weighted miss error, WME)这 3 种评价指标来评估方法预测 POI 的表现,这些评价指标在后来的方法中被广泛使用.Castellani 等^[16]提出一种利用检测和匹配三维模型对象的多个视图的显著点检测方法,首先对顶点进行显著性检测,检测出少数的稀疏显著点;然后对每个显著点用隐马尔可夫模型建模;最后通过评估隐马尔可夫之间的成对相似性来权衡不同的视图点之间的匹配关系.Lin 等^[17]提出为三维模型上每个点计算响应函数提取 POI 的方法,使用邻接顶点之间到彼此所在切平面的欧几里得距离和切平面法向量之间的夹角作为几何特征来计算响应函数,响应函数定义在多尺度空间中能够很好地区分 POI,这些有局部最大的响应函数值的顶点作为 POI 候选点,最后使用优化方法评估其质量和分数得到 POI.

1.2 基于学习的方法

基于学习的方法从大量的三维模型数据样本中建立机器学习模型,以数据驱动的方式自动地学习顶点的几何特征与 POI 之间的复杂映射关系,弥补了基于特征的方法难以处理复杂映射关系的缺陷.

Creusot 等^[18]较早提出基于学习的 POI 检测方法,通过聚焦于三维模型的面片形状联合使用多种几何特征描述符,建立的机器学习模型能够产生与手工标记 POI 有关的 POI 概率图,然后从概率图中进一步提取 POI.Shu 等^[19]构建一个 4 层全连接神经网络来学习顶点的几何特征到 POI 概率值之间的映射关系,然后使用密度峰值聚类方法^[20]从概率分布中提取 POI.Alessio 等^[21]提出一种使用机器学习来检测 POI 的方法,将 POI 的检测视作三维模型表面上顶点的二分类任务,通过训练一个随机森林分类器提取未标记模型上的 POI,由于标记的 POI 只占所有顶点中的很小一部分,该方法使用重采样处理正负样本的不平衡问题.Shu 等^[22]提出一种通过多视角投影来检测 POI 的方法,从不同视角为三维模型生成多幅投影二维图像,然后在这些二维图像上用神经网络检测 POI,最后再逆映射到三维模型上得到最终的 POI;该方法取得较为不错的效果,但不可避免地存在遮挡问题.Zhu 等^[23]提出一种学习隐式表示的 POI 提取方法,通

过神经网络学习 POI 所在球面的有向距离场的隐式表示, 然后使用霍夫变换从有向距离场表示中进一步提取 POI. Li 等^[24]提出一种在点云上的无监督 POI 检测方法, 对输入的点云做随机生成的空间位置变换生成点云对, 然后通过网络对点云对预测 POI, 将 2 个点云对的 POI 在同一坐标系下的距离作为损失函数来优化网络, 该方法保证了姿势不变性, 在各种姿势下的点云都能给出稳健的结果. 舒振宇等^[25]提出一种基于交替优化的 POI 提取方法, 通过对比 POI 提取结果与顶点标签的差异优化顶点的标签, 并重新训练来优化网络实现更好的效果, 但在处理三维模型精细区域时仍有待提高, 且需要反复训练, 因此需要较大的时间开销.

现有方法难以处理三维模型上较为精细的局部区域, 原因是这些局部区域中的顶点之间具有极其相似的几何特征, 因此难以感知这些顶点之间的细微差异, 在这些局部区域提取 POI 时就会出现较大的误差. 针对这种不足, 本文提出一种采用投票机制和动态分组的 POI 提取方法, 能够较

好地处理三维模型上的精细区域, 感知这些精细区域中顶点之间的细微差异, 在判定 POI 时给出更准确的判断.

2 本文方法

本文方法由顶点语义编码、概率分布预测网络和动态分组提取 POI 这 3 个模块组成, 其结构如图 1 所示. 在顶点语义编码模块中, 顶点语义编码器提取三维模型上顶点的邻域和全局信息, 并使用注意力机制进一步编码输出顶点语义序列; 在概率分布预测网络模块中, 投票网络和置信度网络分别学习从顶点语义序列到投票概率序列、置信度序列之间的映射关系, 将顶点语义序列解码成投票概率序列和置信度序列, 将两者结合输出概率分布; 在动态分组提取 POI 模块中, 动态分组通过设定从低到高的概率阈值参数, 从概率分布中采用分组策略来提取 POI.

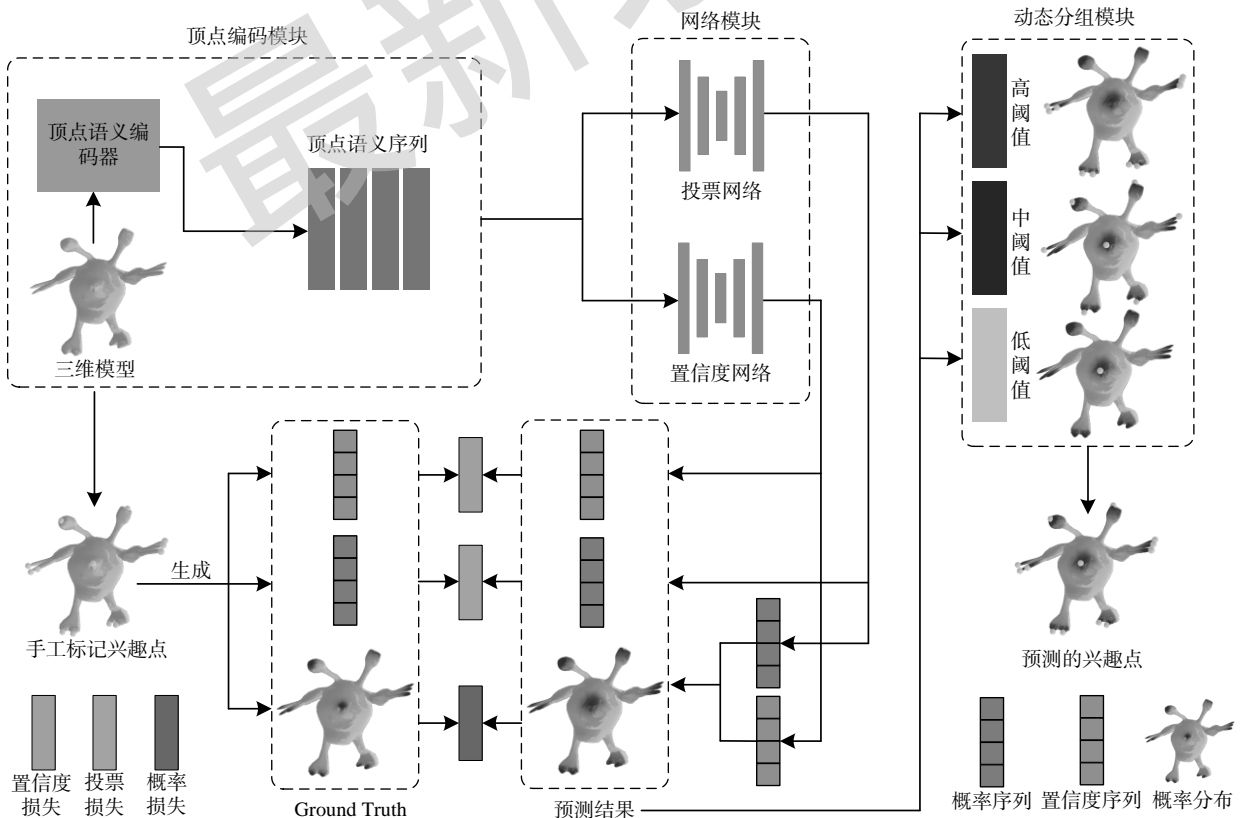


图 1 本文方法结构

2.1 顶点语义编码

顶点语义编码模块的编码过程如图 2 所示. 设输入三维模型的顶点集为 $V = \{v_1, \dots, v_t\}$, 其中, t

表示顶点的数目; 对 V 使用在点云处理任务中广泛使用的最远点采样 (farthest point sampling, FPS), 获得全局顶点序列 $P_g = \{v_{g,1}, \dots, v_{g,N_g}\}$. 对 V 中的采样

顶点 v_i 用测地线距离, 使用 K 近邻算法提取 v_i 的邻域顶点序列 $v_{i,1}, \dots, v_{i,K} \in E(v_i)$, 对于 $E(v_i)$ 中的每个顶点提取 AGD, WKS, SDF, GC 和 SIHKS 这 5 种几何特征描述符, 得到 v_i 的邻域顶点特征 $\mathbf{n}_{i,1}, \dots, \mathbf{n}_{i,K} \in \mathbf{N}_i^{K \times D}$, 其中, D 表示每个邻域顶点提取 5 种几何特征描述符后的特征维度; 然后在采样顶点 v_i 的邻域顶点特征 $\mathbf{N}_i^{K \times D}$ 上对自身使用注意力机制提取采样顶点 v_i 的局部特征 $\mathbf{L}_i^{K \times D}$, 利用 $\mathbf{N}_i^{K \times D}$ 在全局顶点序列 \mathbf{P}_g 上使用注意力机制得到采样顶点 v_i 的全局特征 $\mathbf{G}_i^{K \times D}$; 最后将两者结合, 得到采样顶点 v_i 的顶点语义序列 $\mathbf{F}_i^{K \times D}$.

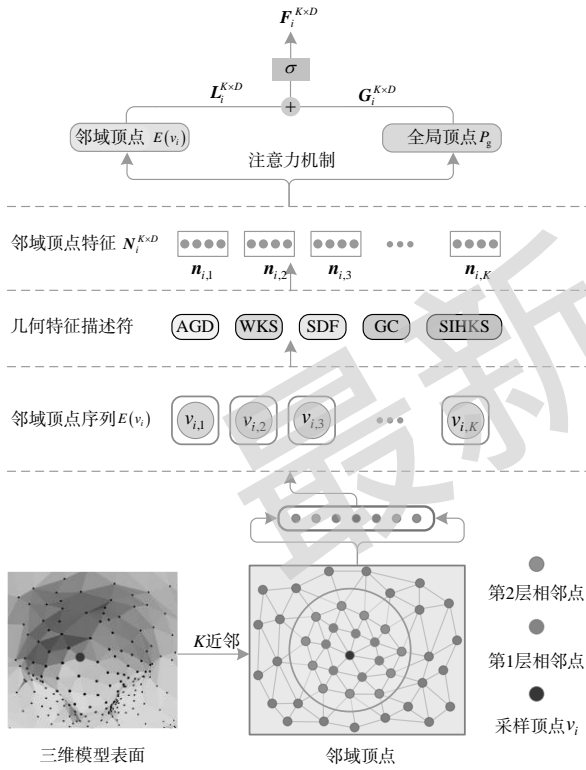


图 2 顶点语义序列的生成过程

图 3 所示为在采样顶点 v_i 的邻域顶点特征 $\mathbf{N}_i^{K \times D}$ 上, 利用注意力机制融合顶点 v_i 局部信息和全局信息的过程, 分为 2 部分: 对自身使用注意力机制获得局部特征 $\mathbf{L}_i^{K \times D}$, 对全局顶点序列 \mathbf{P}_g 使用注意力机制获得全局特征 $\mathbf{G}_i^{K \times D}$.

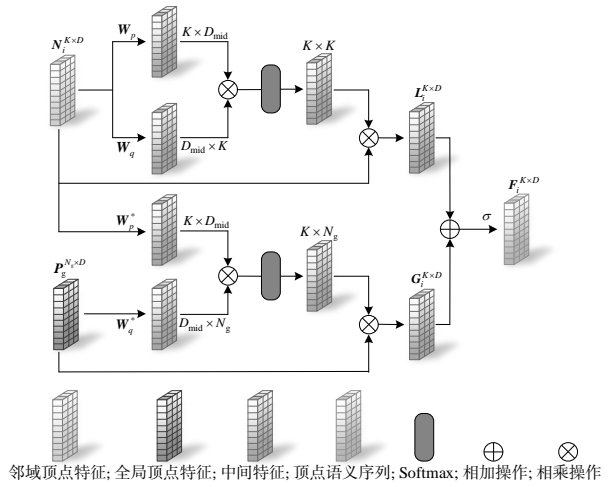


图 3 融合采样顶点的局部信息和全局信息的过程

2.1.1 编码采样顶点的局部特征

获得采样顶点 v_i 的局部信息 $\mathbf{L}_i^{K \times D}$ 的过程为

$$\begin{cases} \mathbf{L}_i^{K \times D} = \{l_{i,1}, \dots, l_{i,K}\} \\ l_{i,k} = A(S(\mathbf{n}_{i,k}, \mathbf{n}_{i,j}) \mathbf{n}_{i,j}, \forall \mathbf{n}_{i,j} \in \mathbf{N}_i^{K \times D}) \\ S(\mathbf{n}_{i,k}, \mathbf{n}_{i,j}) = \text{Softmax}(\mathbf{W}_p(\mathbf{n}_{i,k})^T \mathbf{W}_q(\mathbf{n}_{i,j}) / \sqrt{D_{\text{mid}}}) \end{cases}$$

其中, A 表示聚合函数; S 表示计算采样顶点 v_i 的邻域顶点特征 $\mathbf{N}_i^{K \times D}$ 中 $\mathbf{n}_{i,k}$ 与 $\mathbf{n}_{i,j}$ 之间的权重系数的映射函数; \mathbf{W}_p 和 \mathbf{W}_q 表示 2 个独立的线性变换, 实现从特征维度从 $R^D \mapsto R^{D_{\text{mid}}}$ 的变换, D_{mid} 表示变换的中间维度.

2.1.2 编码采样顶点的全局特征

为了获取采样顶点 v_i 的全局特征, 本文使最远点采样, 从原始三维模型中采样 N_g 个全局顶点, 得到顶点集 $\mathbf{P}_g = \{v_{g,1}, \dots, v_{g,N_g}\}$. 同样地, 对 \mathbf{P}_g 中的每个全局顶点提取 AGD, WKS, SDF, GC 和 SIHKS 这 5 种几何特征描述符, 得到三维模型的全局顶点特征 $\mathbf{P}_g^{N_g \times D} = \{\mathbf{n}_{g,1}, \dots, \mathbf{n}_{g,N_g}\}$. 利用采样顶点 v_i 的邻域顶点特征 $\mathbf{N}_i^{K \times D}$, 在全局顶点特征 $\mathbf{P}_g^{N_g \times D}$ 上用注意力机制得到采样顶点 v_i 的全局特征 $\mathbf{G}_i^{K \times D}$ 的过程为

$$\begin{cases} \mathbf{G}_i^{K \times D} = \{g_{i,1}, \dots, g_{i,K}\} \\ g_{i,k} = A^*(S^*(\mathbf{n}_{i,k}, \mathbf{n}_{g,j}) \mathbf{n}_{g,j}, \forall \mathbf{n}_{g,j} \in \mathbf{P}_g^{N_g \times D}) \\ S^*(\mathbf{n}_{i,k}, \mathbf{n}_{g,j}) = \text{Softmax}(\mathbf{W}_p^*(\mathbf{n}_{i,k})^T \mathbf{W}_q^*(\mathbf{n}_{g,j}) / \sqrt{D_{\text{mid}}}) \end{cases}$$

其中, A^* 表示聚合函数; S^* 表示计算采样顶点 v_i 的邻域顶点特征中的 $\mathbf{n}_{i,k}$ 与全局顶点特征 $\mathbf{P}_g^{N_g \times D}$ 中

的 $n_{g,j}$ 之间的权重系数的映射函数; \mathbf{W}_p^* 和 \mathbf{W}_q^* 表示 2 个独立的线性变换, 实现特征维度从 $R^D \mapsto R^{D_{\text{mid}}}$ 的变换. 然后将顶点 v_i 的局部特征与全局特征相加, 并用 Sigmoid 激活函数 σ 得到采样顶点 v_i 的顶点语义序列 $\mathbf{F}_i^{K \times D}$, 公式为

$$\mathbf{F}_i^{K \times D} = \sigma(\mathbf{L}_i^{K \times D} + \mathbf{G}_i^{K \times D}).$$

2.2 生成 Ground Truth

本文利用三维模型上手工标记的 POI, 为三维模型生成 Ground Truth 对应的概率分布、置信度序列和投票概率序列.

设有 t 个顶点的三维模型的顶点集为 $V = \{v_1, \dots, v_t\}$, 在 V 中手工标记的 POI 顶点集为 $P = \{v_{p,1}, \dots, v_{p,m} | v_{p,j} \in V, 1 \leq j \leq m\}$, 其中, m 表示手工标记的 POI 的数量. 本文为 V 生成一个概率分布 $Y = \{y_1, \dots, y_t\}$, $\forall v_i \in V$, 为 v_i 生成其邻域顶点的置信度序列 $\mathbf{W}_i = \{w_{i,1}, \dots, w_{i,K}\}$ 和投票概率序列 $\mathbf{C}_i = \{y_{i,1}^*, \dots, y_{i,K}^*\}$, 其中, K 表示邻域顶点序列的大小, 其值与第 2.1 节使用 K 近邻算法提取邻域顶点序列使用的参数一致.

设 $P_{i,j}$ 为 P 中 $v_{p,j}$ 为 V 中顶点 v_i 计算的投票概率的参考值, 表示为

$$P_{i,j} = 1 - \frac{d(v_i, v_{p,j})}{D_{\max}(v_{p,j})};$$

其中, d 表示三维模型上 2 个顶点之间的测地线距离, $D_{\max}(v_{p,j})$ 表示 V 中所有顶点到 $v_{p,j}$ 最大的测地线距离. 设 P_i 为 P 中所有手工标记 POI 给出 v_i 的投票概率的参考值中的最大值, 即

$$P_i = \max_{1 \leq j \leq m} (P_{i,j}).$$

本文利用高斯分布函数对 P_i 作进一步处理, 将线性映射变成非线性映射, 得到 v_i 对其邻域顶点的投票概率值 y_i^* 为

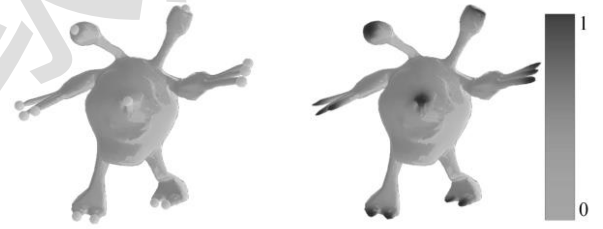
$$y_i^* = \frac{1}{\varepsilon \sqrt{2\pi}} e^{-\frac{(P_i - \mu)^2}{2\varepsilon^2}},$$

本文实验中, ε 为 0.12, μ 为 1.5. 在以上定义的基础上, 为顶点 v_i 生成概率 y_i 、邻域权重序列 \mathbf{W}_i 和邻域投票概率值序列 \mathbf{C}_i 的过程可以表示为

$$\begin{cases} v_{m_i} = \arg \max_{v_{p,j}} (P_{i,j}) \\ \mathbf{R}_i = \left\{ \sigma(\mathbf{n}_{i,1}^T \mathbf{n}_{m_i}), \dots, \sigma(\mathbf{n}_{i,K}^T \mathbf{n}_{m_i}) \right\} \\ \mathbf{W}_i = \text{Softmax}(\mathbf{R}_i) \\ \mathbf{C}_i = \{y_{i,1}^*, \dots, y_{i,K}^*\} \\ y_i = \mathbf{W}_i^T \mathbf{C}_i \end{cases},$$

其中, m_i 表示计算 P_i 使用的 $v_{p,j}$ 在 V 中表示的顶点索引; \mathbf{R}_i 表示 v_i 的邻域顶点特征 $\mathbf{n}_{i,1}, \dots, \mathbf{n}_{i,K} \in N_i^{K \times D}$ 中, 每个顶点特征与 v_{m_i} 的特征 \mathbf{n}_{m_i} 与计算相似度的序列, 然后归一化得到 \mathbf{W}_i ; \mathbf{C}_i 表示 v_i 的邻域顶点投票概率序列, 通过 \mathbf{W}_i 与 \mathbf{C}_i 得到 v_i 最终的概率值 y_i .

如图 4 为根据手工标记的 POI 为三维模型生成 Ground Truth 对应的概率分布. 可以看出, 这样生成的先验概率分布更准确, 有利于提高 POI 的预测效果.



a. 手工标记 POI

b. 概率分布结果

图 4 根据手工标记 POI 生成的概率分布

2.3 概率分布预测网络

如图 5 所示, 概率分布预测网络分为 3 部分: (1) 顶点语义编码器, 其结构如图 3 所示, 提取顶点语义序列的过程见第 2.1.1 节和第 2.1.2 节. 对顶点提取几何特征描述符时, 使用的 AGD, WKS, SDF, GC 和 SIHKS 这 5 种几何特征描述符的维度分别为 1, 19, 1, 1 和 100, 因此, 顶点语义编码器输出的顶点语义序列为 $\mathbf{F}_i^{K \times 122}$. (2) 置信度预测网络. 输入顶点语义序列 $\mathbf{F}_i^{K \times 122}$, 为顶点的邻域顶点预测一个置信度序列 \mathbf{W}_i^* . (3) 概率投票网络. 输入顶点语义序列 $\mathbf{F}_i^{K \times 122}$, 为顶点的邻域顶点预测对该顶点的投票概率序列 \mathbf{C}_i^* . 利用置信度序列和投票概率序列得到预测的概率值 p_i^* , 公式为

$$\begin{cases} \mathbf{C}_i^* = \{p_{i,1}, \dots, p_{i,K}\} \\ \mathbf{W}_i^* = \{w_{i,1}^*, \dots, w_{i,K}^*\} \\ p_i^* = (\mathbf{W}_i^*)^T \mathbf{C}_i^* \end{cases}.$$

设 $F_{in,out}$ 表示输入维度为 in , 输出维度为 out 的全连接层. 在置信度预测网络中, 本文使用 4 个全连接层 $F_{122,64}$, $F_{64,32}$, $F_{32,16}$ 和 $F_{16,1}$ 来预测置信度

序列 W_i^* .

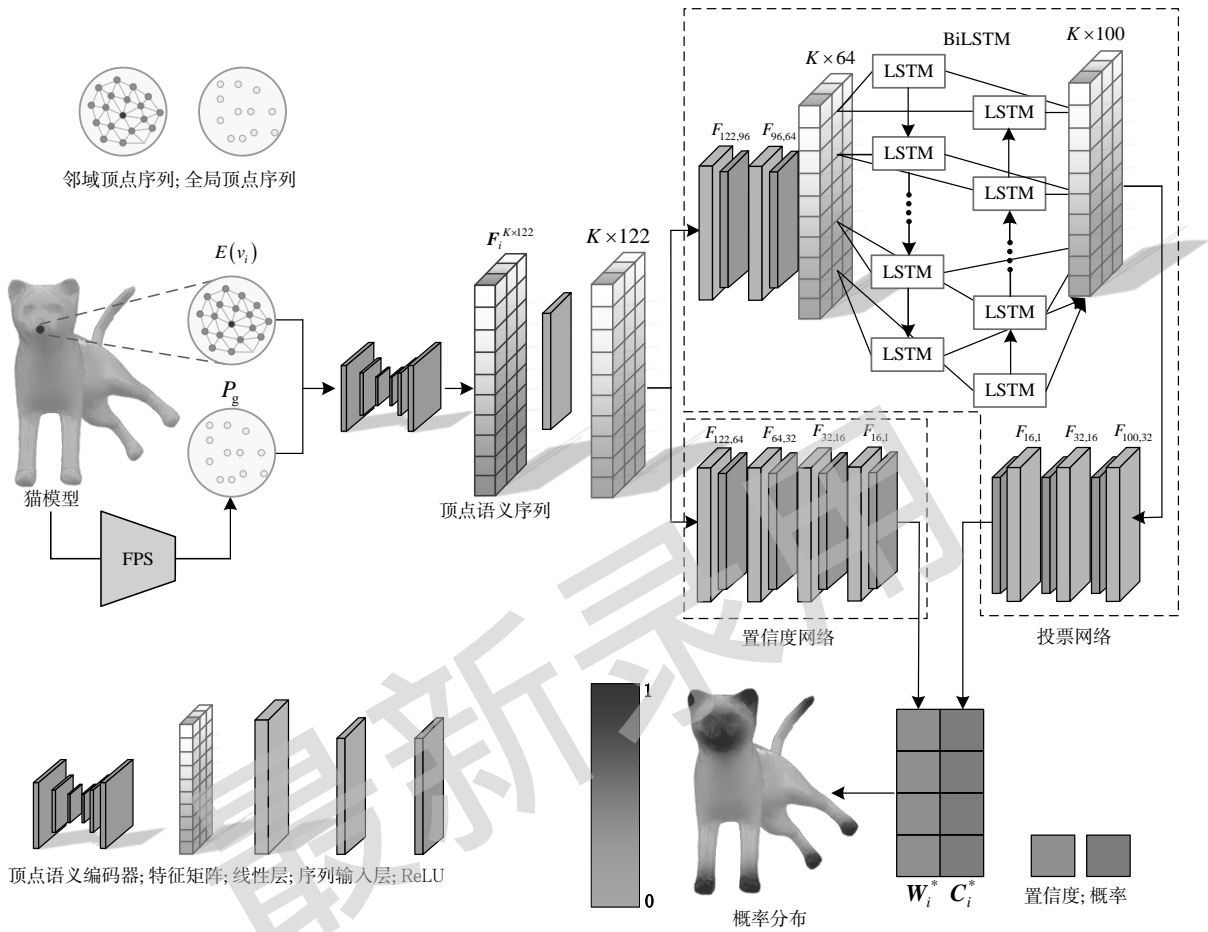


图 5 概率分布预测网络

在概率预测网络中, 本文使用处理序列任务中广泛使用的双向长短期记忆网络 (bi-directional long short-term memory, BiLSTM)^[26-27] 来处理输入的顶点语义序列 $F_i^{K \times 122}$, 得到预测的邻域顶点的投票概率值序列 C_i^* . 具体过程如下: 顶点语义序列 $F_i^{K \times 122}$ 先经过序列输入层, 再经过 $F_{122,96}$ 和 $F_{96,64}$ 这 2 个全连接层之后维度变为 $K \times 64$; 然后由 BiLSTM 层从正反 2 个方向进一步提取序列特征, BiLSTM 中, 每个 LSTM 的输入之一均为 $K \times 64$ 中由 $F_{96,64}$ 生成的 64 维特征向量. 本文设置 BiLSTM 内部隐藏层为 100, 序列经过 BiLSTM 后的输出维度为 $K \times 100$; 然后经过 $F_{100,32}$, $F_{32,16}$ 和 $F_{16,1}$ 这 3 个全连接层输出预测的投票概率值序列 C_i^* .

2.4 损失函数

概率分布预测网络的损失分为邻域置信度损失 L_w 、概率损失 L_p 和邻域投票概率损失 L_c 这 3

部分, 分别表示预测的置信度序列 W_i^* 、概率 p_i^* 、投票概率值序列 C_i^* 与 Ground Truth 生成的置信度序列 W_i 、概率 y_i 、投票概率值序列 C_i 之间的损失, 均使用均方误差 (mean square error, MSE) 损失表示为

$$\begin{cases} L_p = \frac{1}{t} \sum_{i=1}^t (p_i^* - y_i)^2 \\ L_w = \frac{1}{t} \sum_{i=1}^t \|W_i^* - W_i\|_2 \\ L_c = \frac{1}{t} \sum_{i=1}^t \|C_i^* - C_i\|_2 \end{cases}$$

将这三部分损失结合起来, 得到网络模块的损失函数 $L = L_p + \alpha L_w + \beta L_c$, 其中, α , β 表示控制 L_w 和 L_c 的超参数.

2.5 动态分组提取 POI

本文提出一种在三维模型概率分布上采用动态分组策略提取 POI 的方法. 该方法基于顶点分组策略, 先从概率分布上分割出多个概率值较大

的分组, 其通常对应于三维模型上具有特殊语义的部分; 然后对这些具有特殊语义的分组进一步提取预测的 POI. 分组策略的基本思路如图 6 所示.

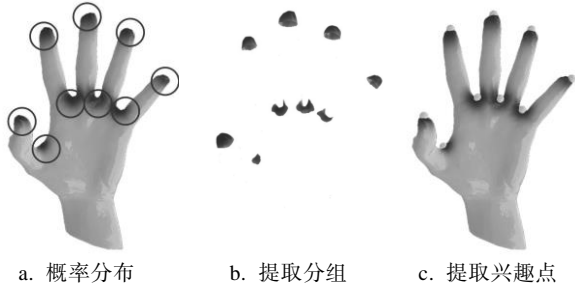


图 6 分组提取策略

设三维模型的顶点集为 $V = \{v_1, v_2, \dots, v_n\}$, 预测的概率分布为 $Y = \{y_1, y_2, \dots, y_n\}$, $d(v_i, v_j)$ 表示顶点 v_i 和 v_j 之间的测地线距离, $e(v_i, v_j)$ 表示顶点 v_i 和 v_j 之间是否相邻的分段函数, 相邻为 1 否则为 0, h 表示所有相邻顶点之间的平均测地线距离. $e(v_i, v_j)$ 和 h 可表示为

$$e(v_i, v_j) = \begin{cases} 1, & \text{if } v_i, v_j \text{ 相邻} \\ 0, & \text{otherwise} \end{cases},$$

$$h = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n d(v_i, v_j) \cdot e(v_i, v_j)}{\sum_{i=1}^{n-1} \sum_{j=i+1}^n e(v_i, v_j)}.$$

$$\begin{cases} k = j \times 0.01 \\ P_k = \{v_{p,1}, \dots, v_{p,A_k}\} \\ I = \bigcup_{j=0}^{100} P_k^{\text{POI}} \end{cases}.$$

分组提取 POI 的方法步骤如下:

(1) 设定一个概率阈值 k , 从预测的概率分布中筛选出所有概率值不低于 k 的顶点构成顶点集 $P_k = \{v_{p,1}, \dots, v_{p,A_k}\}$, 其中, A_k 表示顶点集 P_k 中的顶点数目. 初始时刻, 将每个顶点看作为一个局部的分组, 得到分组集合 $R = \{g_1, \dots, g_{A_k}\}$. 对 $\forall v_{p,i} \in P_k$, $H(v_{p,i})$ 表示在分组集合 R 中 $v_{p,i}$ 所在的分组, 即 $v_{p,i} \in H(v_{p,i})$, $H(v_{p,i}) \in R$. 对 $\forall v_{p,i} \in P_k, \forall v_{p,j} \in P_k$, 若满足 $d(v_{p,i}, v_{p,j}) \leq h$, 则将顶点 $v_{p,i}$ 和顶点 $v_{p,j}$ 所在的局部分组合并为一个新的分组, 即对 R 做如下更新

$$R = R - H(v_{p,i}) - H(v_{p,j}) + (v_{p,i}) \cup H(v_{p,j}).$$

(2) 对 R 更新完后得到最终的分组集 $R = \{g_{f,1}, \dots, g_{f,m}\}$, 对 R 中每个分组求解一个概率值最大的顶点作为 POI, 得到最终的预测 POI 顶点集

$$P_k^{\text{POI}} = \{\max(g_{f,1}), \dots, \max(g_{f,m})\}.$$

在分组提取中需要设定概率阈值参数 k , 其大小影响最终的分组结果. 若 k 过大, 将导致分组只筛选出概率值较大的顶点, 只有那些概率值较大的分组才会被提取 POI, 造成一些较小概率值顶点中的 POI 被遗漏; 而 k 过小时, 则会筛选出过多的概率值较小的顶点, 多个分组之间会浮现出一些新的顶点作为桥梁将这些分组合并, 从而造成分组的减少. 针对在分组提取时无论如何设置概率阈值 k 都无法全面提取到所有 POI 的问题, 进一步改进动态分组, 从较低的概率阈值 k 开始, 并逐渐增大 k 进行分组, 从而兼顾不同 k 的分组结果, 实现全面地从概率分布上提取 POI. 设 P_k^{POI} 表示在三维模型概率分布上, 对由概率阈值 k 确定的顶点集 P_k 进行分组操作后得到的预测 POI 集合, 本文设定 k 的初始值为 0, 并从 0 到 1 按照 0.01 的间隔依次递增, 共进行 101 次分组操作, 设动态分组获得的预测 POI 集合为 I , 获得 I 的过程可以表示为

3 实验与结果分析

3.1 实验评估

3.1.1 数据集

本文实验采用 SHREC 2011 数据集^[28], 其为一个公开的三维模型数据集, 总共包含 30 类模型, 每类模型有 20 个样本, 每个样本事先由志愿者手工标记 POI. 在实验中, 将每类模型中 12 个样本作为训练集, 8 个样本作为测试集.

3.1.2 评价指标

在评价方法预测的 POI 是否检测到手工标记的 POI 时, 设定一个最大测地线距离百分比范围 λ , 若方法预测的 POI 落在手工标记的 POI 的 λ 之内, 则认为方法预测的 POI 是正确的 POI.

在三维模型上, 设手工标记的 m 个 POI 构成的集合为 $A = \{v_{a,1}, \dots, v_{a,m}\}$, 方法预测的 n 个 POI 构成的集合为 $B = \{v_{b,1}, \dots, v_{b,n}\}$; $d(v_i, v_j)$ 表示顶点 v_i 和顶点 v_j 之间的测地线距离, $D_{\max}(v_i)$ 表示所有顶点到顶点 v_i 中的最远测地线距离; $r(v_{a,i})$ 表示 A 中顶

点 $v_{a,i}$ 到 B 中所有预测的 POI 中最小的测地线百分比距离, $f(v_{a,i})$ 表示 $v_{a,i}$ 是否被预测的 POI 检测到; T 表示在给定的 λ 范围内手工标记的 POI 被检测到的数量. 本文采用 FPE 和 FNE 评价方法预测 POI 的准确性, 计算 FPE 和 FNE 的公式为

$$r(v_{a,i}) = \min_{1 \leq j \leq n} \left(\frac{d(v_{b,j}, v_{a,i})}{D_{\max}(v_{a,i})} \right),$$

$$f(v_{a,i}) = \begin{cases} 1, & \text{if } r(v_{a,i}) \leq \lambda \\ 0, & \text{otherwise} \end{cases},$$

$$T = \sum_{i=1}^m f(v_{a,i}),$$

$$\text{FNE} = 1 - \frac{T}{m},$$

$$\text{FPE} = 1 - \frac{T}{n}.$$

其中, FPE 和 FNE 分别从方法检测 POI 的错误率、遗漏率等方面评价 POI 检测方法; FNE 表示所有手工标记的 POI 中判断为阴性结果的比例, 其值越小, 说明手工标记的 POI 中被方法正确预测出来的比例越大; FPE 表示所有预测的 POI 中错误预测的比例, 其值越小, 说明方法预测的 POI 中正确的 POI 比例越大, 方法的错误率越小.

3.1.3 实验配置

本文实验在一台 Intel Core TM i7-11700K 3.60 GHz CPU, 64GBRAM 内存, NVIDIA GeForce GTX 3090Ti 显卡的设备上进行, 基于 Matlab 和 C++ 实现; 使用 SHREC 2011 数据集, 使用 FNE 和 FPE 评价指标, 设定的测地线距离百分比范围 λ 从 0.00 到 0.12, 按照 0.01 的间隔变化. 训练网络时, 设置的学习速率为 0.001、训练批大小为 64、训练轮次为 10, 由梯度下降法优化网络参数, 损失函数使用的超参数为 $\alpha = 0.6$, $\beta = 0.8$; 提取顶点邻域使用的参数 $K = 10$, 提取顶点语义序列时使用的中间维度 $D_{\text{mid}} = 64$, FPS 采样的全局顶点参数 $N_g = 512$.

3.2 实验结果

图 7 所示为本文方法与 Ground Truth 的一些对比结果. 可以看出, 本文方法预测的 POI 与 Ground Truth 的 POI 基本一致, 预测的概率分布与 Ground Truth 的概率分布也几乎一致. 图 8 所示为本文方法在 SHREC2011 数据集上绝大部分模型中

的 POI 预测结果. 可以看出, 预测的 POI 效果非常好, 基本上对应于三维模型上有特殊语义的部位.

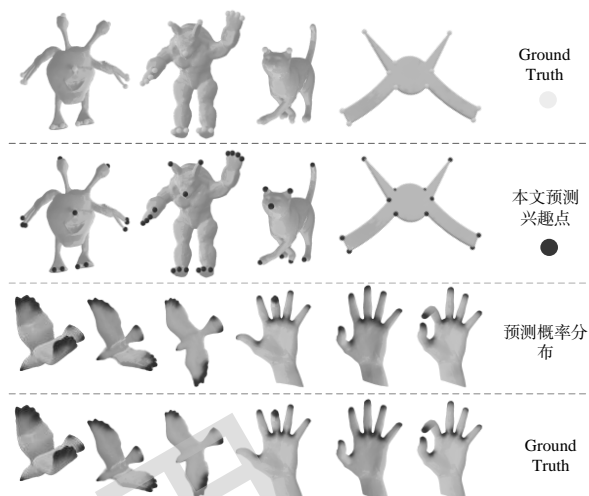


图 7 本文方法与 Ground Truth 的对比结果

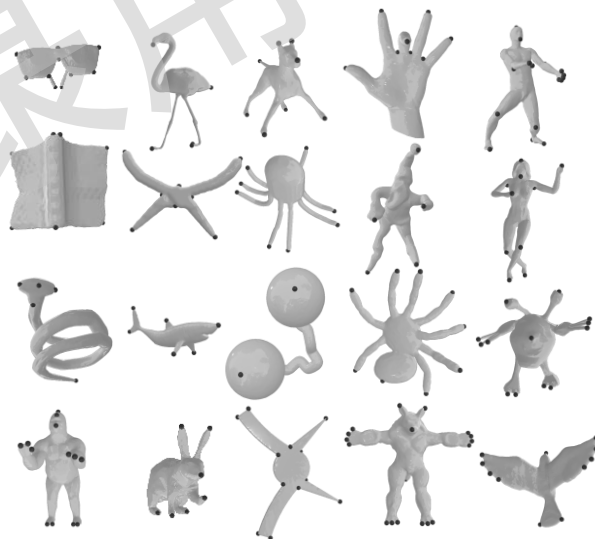


图 8 本文方法在 SHREC 2011 上部分模型的预测结果

表 1 所示为 $\lambda = 0.12$ 时, 本文方法在 SHREC2011 数据集上每类模型中的平均 FNE. 可以看出, 在多数模型上的 FNE 接近于 0, 有 17 类模型的 FNE 等于 0, 说明几乎所有模型上的手工标记 POI 被检测到的比例为 100%.

表 1 本文方法在 SHREC2011 上 30 类模型的 FNE

模型	FNE	模型	FNE	模型	FNE	模型	FNE	模型	FNE
Alien	0.0000	Cat	0.0000	Flamingo	0.0000	Man	0.0208	Scissor	0.0354

%

Ants	0.0498	Centaur	0.0156	Glasses	0.0875	Octopus	0.0000	Shark	0.0000
Armadillo	0.0000	Dino	0.0000	Gorilla	0.0000	Paper	0.0000	Snake	0.1250
Bird1	0.0000	Dinosaur	0.0035	Hand	0.1424	Pliers	0.0469	Spider	0.0139
Bird2	0.0139	Dog1	0.0000	Horse	0.0000	Rabbit	0.0000	Two_balls	0.1250
Camel	0.0000	Dog2	0.0000	Lamp	0.2500	Santa	0.0000	Woman	0.0000

表 2 所示为 λ 变化时, 本文方法在 SHREC2011 数据集上全部模型的平均 FNE 和 FPE 对比. 可以看出, 当 λ 为 0.03, 即在检测精度十分高的条件下, FNE 仍只有 0.1051, FPE 只有 0.2803; 当 λ 放宽至 0.12 时, FNE 为 0.0286, FPE 为 0.2206. 实验结果表明, 手工标记的 POI 均被检测出来, 本文方法能达到接近 80% 的正确率.

表 2 本文方法在 SHREC2011 上平均检测性能对比 %

λ	FNE	FPE
0.00	0.9306	0.9448
0.01	0.4542	0.5710
0.02	0.2059	0.3655
0.03	0.1051	0.2803
0.04	0.0665	0.2515
0.05	0.0455	0.2351
0.06	0.0427	0.2294
0.07	0.0344	0.2269
0.08	0.0314	0.2225
0.09	0.0292	0.2206
0.10	0.0286	0.2206
0.11	0.0286	0.2206
0.12	0.0286	0.2206

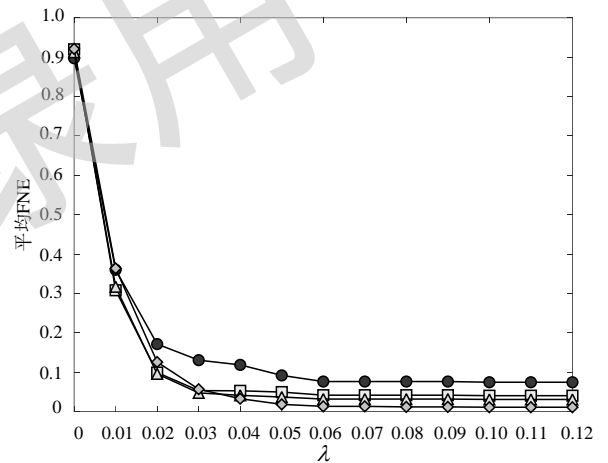
在 SHREC2011 数据集上的 Glasses 类模型中统计本文方法的平均运行时间消耗, 随机选择 5, 10, 15, 20 个模型进行 10 轮网络训练和动态分组并计算所用时间, 重复 3 次最后得到的平均时间开销如表 3 所示. 可以看出, 本文方法总体上运行时间较快, 网络部分平均 0.5 min 处理完一个模型, 动态分组平均 7.7s 完成一个模型的 POI 提取.

表 3 网络训练和动态分组的平均用时 min

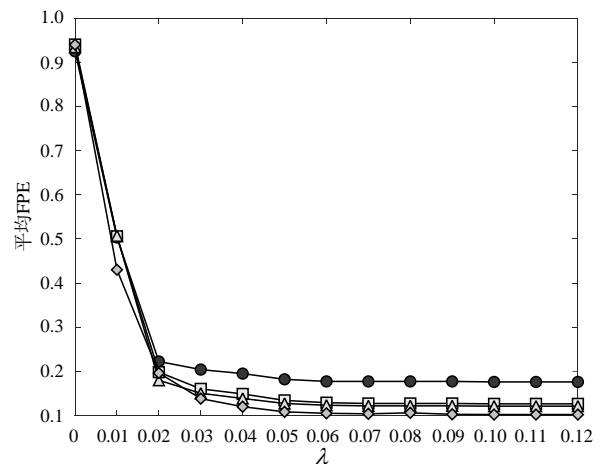
模型数量	网络训练	动态分组
5	2.56	0.67
10	4.85	1.28
15	7.13	1.96
20	9.32	2.57

通过设计不同训练/测试数据集划分比例对 POI 提取准确性的实验, 验证本文方法的泛化性能, 其中, 使用 SHREC2011 数据集上的 Ant, Cat 和 Glasses 这 3 类模型, 每类模型均使用训练/测试为 4/16, 8/12, 12/8 和 16/4 的 4 组比例划分. 图 9 所示为在 4 组比例下得到的平均 FNE 和 FPE 结果. 可以看出, 在 4 种划分比例的情况下, FNE 均能降到

0.1 以下; 比例为 4/16 时 FNE 稍高, 其他 3 种比例下 FNE 依次减少但基本相差不大; FPE 在 4 种划分比例的情况下均低于 0.2, 在比例为 4/16 时高于其他 3 种比例; 随着训练数据比例的增加, FPE 逐渐减少但减少幅度不大; 在不同划分比例下, 本文方法检测的正确率均能超过 80%, 检测出 90% 以上的人工标记 POI. 实验结果表明, 在不同的训练/测试的比例下, 本文方法均有出色的表现, 即使在训练数据较少的情况下, 依然能给出较好的结果, 具有较好的泛化性能.



a. 4 种比例下的平均 FNE



b. 4 种比例下的平均 FPE

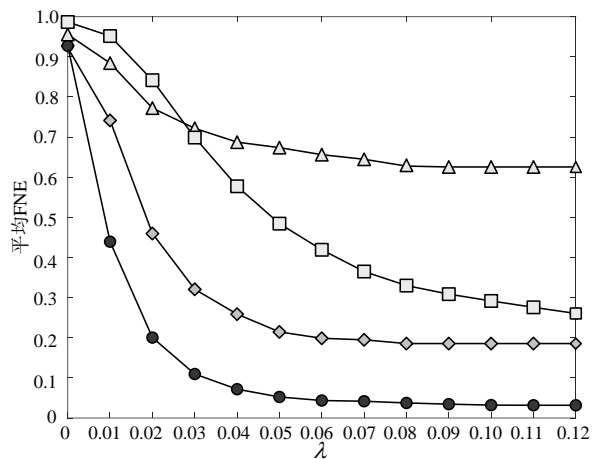
● 4/16; □ 8/12; △ 12/8; ◇ 16/4

图 9 4 种比例下平均 FNE 和 FPE 对比

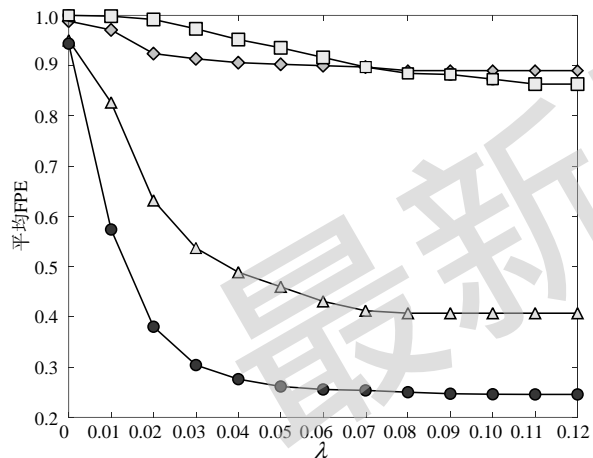
3.3 与其他方法对比

将本文方法与传统的 POI 提取方法, 如 HKS^[29], 3D-Harris^[30]和 3D-SIFT^[31]进行对比, 结果

如图 10 所示. 可以看出, 随着 λ 的增大, FNE 和 FPE 都逐渐减小, 本文方法在 2 项指标上明显优于其他 3 种方法.



a. 4 种方法的平均 FNE



b. 4 种方法的平均 FPE

● 本文; □ 3D-SIFT^[31]; △ HKS^[29]; ◇ 3D-Harris^[30]

图 10 4 种方法的 FNE 和 FPE 对比

在 KeypointNet 数据集^[32]的 10 类模型上, 将本文方法与基于深度学习的 POI 提取方法, 如 PointNet^[33], DGCNN^[34]和 UIRL^[23]进行对比, 并采用双向豪斯多夫距离(bidirectional Hausdorff distance, BHD)和倒角距离(chamfer distance, CD)作为评价指标, 结果如表 4 所示. 其中, BHD 衡量 2 个顶点集之间的相似性, CD 则衡量 2 个顶点集之间的距离, 其值均是越小表现越好. 假设 2 个顶点集为 V_1 和 V_2 , 则 $BHD(V_1, V_2)$ 和 $CD(V_1, V_2)$ 可表示为

$$BHD(V_1, V_2) = \frac{1}{2} \left(\max_{a \in V_1} \min_{b \in V_2} \|a - b\|_2 + \max_{a \in V_2} \min_{b \in V_1} \|a - b\|_2 \right),$$

$$CD(V_1, V_2) = \frac{1}{|V_1|} \sum_{a \in V_1} \min_{b \in V_2} \|a - b\|_2^2 + \frac{1}{|V_2|} \sum_{b \in V_2} \min_{a \in V_1} \|a - b\|_2^2.$$

表 4 4 种方法的平均 BHD 和 CD 对比

模型	PointNet ^[33]	DGCNN ^[34]	UIRL ^[23]	本文方法
Airplane	0.366/0.070	0.321/0.098	0.124/0.015	0.102/0.011
Bath	0.422/0.081	0.354/0.156	0.235/0.031	0.194/0.027
Chair	0.310/0.097	0.421/0.119	0.148/0.018	0.137/0.014
Car	0.474/0.198	0.247/0.023	0.121/0.009	0.128/0.013
Guitar	0.612/0.249	0.183/0.019	0.097/0.007	0.083/0.006
Knife	0.890/0.376	0.775/0.113	0.147/0.025	0.139/0.019
Laptop	1.022/0.552	0.635/0.433	0.097/0.015	0.086/0.013
Motor	0.542/0.276	0.474/0.166	0.194/0.017	0.237/0.024
Table	0.660/0.253	0.159/0.073	0.117/0.026	0.107/0.022
Vessel	0.360/0.110	0.320/0.054	0.260/0.053	0.213/0.043

注. 粗体为最优值.

从表 4 可以看出, 本文方法在 8 类模型上均优于其他 3 种方法, 仅在 Car 和 Motor 模型上为次优. 实验结果表明, 本文方法预测的 POI 集合与手工标记的 POI 集合 BHD 更高、CD 更小.

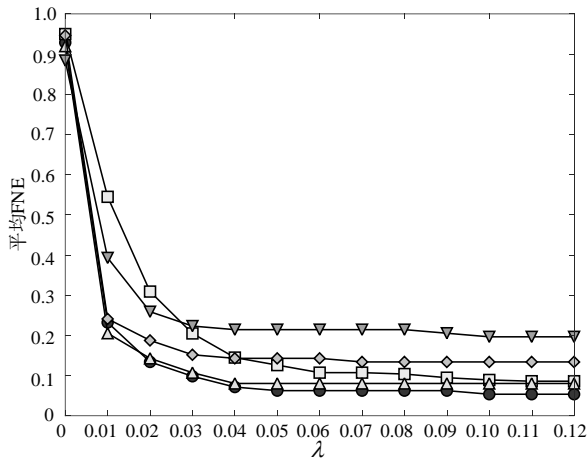
3.4 消融实验

本文设计了多组消融实验, 验证本文方法中的顶点编码模块、网络模块和动态分组模块的有效性. 其中, 验证每个模块的有效性时, 其他模块保持不变.

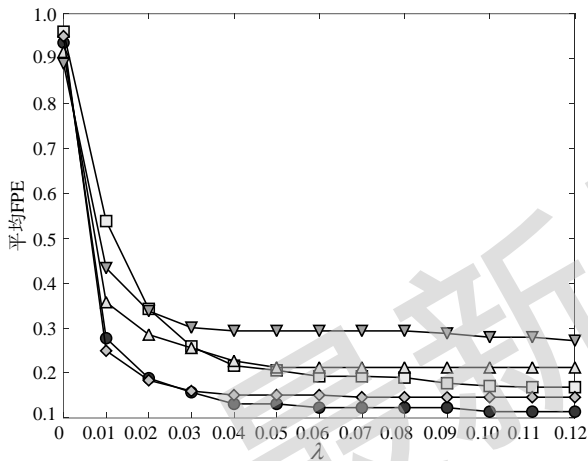
3.4.1 顶点编码模块

为了验证顶点编码模块的有效性, 在对比顶点编码模块提取顶点信息时, 只考虑利用顶点自身信息、融合顶点的邻域信息和本文方法 3 种情况, 结果如图 11 所示. 可以看出, 融合了邻域顶点信息的 FNE 和 FPE 指标明显低于只利用顶点自身信息的 FNE 和 FPE 指标; 融合了顶点的邻域和全局特征时, FNE 和 FPE 指标比在只融合邻域信息的基础上获得进一步提升.

表 5 所示为 3 种实验情况下 6 种三维模型的概率分布预测结果. 可以看出, 由于融合了顶点的邻域和全局信息, 预测的概率分布最接近 Ground Truth, 能够很好地区分出三维模型上具有特殊语义的部位; 仅利用顶点自身信息预测的概率分布精确性最差, 尤其是在手工标记的 POI 密集的情况下, 如在 Glasses 模型上, 概率分布的区分度最差, 不能很好地区分出三维模型上具有特殊语义的部位; 仅融合顶点邻域信息的结果介于仅利用顶点自身和融合顶点邻域与全局信息的结果之间. 实验结果表明, 融合顶点的邻域和全局信息能明显地提高预测概率分布的精确性和特殊语义部位的区分度, 进而提升 POI 提取效果.



a. 多组消融实验的平均 FNE



b. 多组消融实验的平均 FPE

● 本文; ◆ 邻域顶点; ▲ 顶点自身
□ 未使用投票机制; ○ 密度峰值聚类

图 11 多组消融实验的结果对比

表 5 顶点模块消融实验预测的概率分布结果对比

模型	Ground Truth	本文方法	仅邻域	顶点自身
Alien				
Rabbit				
Glasses				
Scissor				
Ants				



3.4.2 网络模块

图 11 中也展示了本文方法与未使用投票与置信度网络 2 种情况的实验结果. 可以看出, 本文方法的 FNE 和 FPE 指标明显好于未使用投票网络和置信度网络时的 FNE 和 FPE 指标, 说明使用投票网络和置信度网络确实能提高 POI 提取效果.

3.4.3 动态分组模块

为了验证动态分组模块的有效性, 本文对比了在相同的概率分布上, 分别使用动态分组和密度峰值聚类提取 POI 的情况, 结果如图 11 所示. 可以看出, 本文方法的 FNE 稍好于密度峰值聚类, 2 种方法都能检测出接近 90% 的 POI; 而在 FPE 上, 本文方法明显好于密度峰值聚类, 本文方法提取的 POI 正确率接近 90%, 而密度峰值聚类正确率接近 80%. 实验结果表明, 动态分组比密度峰值聚类具有提取的 POI 正确率更高、结果可信度更高的优势.

4 结 语

本文提出一种在三维模型上提取 POI 的方法. 在提取顶点特征时, 使用注意力机制融合顶点的邻域信息和全局信息能够更全面的提取顶点的特征; 在预测概率分布上, 提出投票网络和置信度预测网络共同协作来预测三维模型的概率分布, 与现有方法相比, 由于充分利用了其邻域信息, 因此能够更为准确地预测三维模型的概率分布; 针对在概率分布上提取 POI, 提出一种基于动态分组的方法.

本文方法比较依赖手工标记的 POI 准确性. 当手工标记的 POI 具有一定的随意性时, POI 概率预测网络不能准确地学习到顶点语义信息与 POI 概率之间的复杂映射关系, 导致检测结果的精度下降; 另一方面, 顶点编码器使用的邻域顶点序列的大小也会对结果有一定的影响, 太小将不能充分提取顶点的特征, 太大又会引入冗余信息带来干扰, 因此确定合适大小的邻域顶点序列是本文未来将要探索的一个方向.

参考文献(References):

[1] Zhen Q K, Huang D, Wang Y H, *et al.* Muscular movement model-based automatic 3D/4D facial expression recogni-

- tion[J]. *IEEE Transactions on Multimedia*, 2016, 18(7): 1438-1450
- [2] Tam G K L, Cheng Z Q, Lai Y K, *et al.* Registration of 3D point clouds and meshes: A survey from rigid to nonrigid[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(7): 1199-1217
- [3] Katz S, Leifman G, Tal A. Mesh segmentation using feature point and core extraction[J]. *The Visual Computer*, 2005, 21(8-10): 649-658
- [4] Xie J, Dai G X, Fang Y. Deep multimetric learning for shape-based 3D model retrieval[J]. *IEEE Transactions on Multimedia*, 2017, 19(11): 2463-2474
- [5] Chen J Y, Lin C H, Hsu P C, *et al.* Point cloud encoding for 3D building model retrieval[J]. *IEEE Transactions on Multimedia*, 2014, 16(2): 337-345
- [6] Bu S H, Liu Z B, Han J W, *et al.* Learning high-level feature by deep belief networks for 3-D model retrieval and recognition[J]. *IEEE Transactions on Multimedia*, 2014, 16(8): 2154-2167
- [7] Kim Y, Varshney A. Saliency-guided enhancement for volume visualization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2006, 12(5): 925-932
- [8] Valstar M, Martinez B, Binefa X, *et al.* Facial point detection using boosted regression and graph models[C] // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2010: 2729-2736
- [9] Shapira L, Shamir A, Cohen-Or D. Consistent mesh partitioning and skeletonisation using the shape diameter function[J]. *The Visual Computer*, 2008, 24(4): 249-259
- [10] Shapira L, Shalom S, Shamir A, *et al.* Contextual part analogies in 3D objects[J]. *International Journal of Computer Vision*, 2010, 89(2): 309-326
- [11] Gal R, Cohen-Or D. Salient geometric features for partial shape matching and similarity[J]. *ACM Transactions on Graphics*, 2006, 25(1): 130-150
- [12] Bronstein M M, Kokkinos I. Scale-invariant heat kernel signatures for non-rigid shape recognition[C] // *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2010: 1704-1711
- [13] Aubry M, Schlickewei U, Cremers D. The wave kernel signature: A quantum mechanical approach to shape analysis[C] // *Proceedings of the IEEE International Conference on Computer Vision Workshops*. Los Alamitos: IEEE Computer Society Press, 2011: 1626-1633
- [14] Sipiran I, Bustos B. Harris 3D: A robust extension of the Harris operator for interest point detection on 3D meshes[J]. *The Visual Computer*, 2011, 27(11): 963-976
- [15] Dutagaci H, Cheung C P, Godil A. Evaluation of 3D interest point detection techniques via human-generated ground truth[J]. *The Visual Computer*, 2012, 28(9): 901-917
- [16] Castellani U, Cristani M, Fantoni S, *et al.* Sparse points matching by combining 3D mesh saliency with statistical descriptors[J]. *Computer Graphics Forum*, 2008, 27(2): 643-652.
- [17] Lin X Y, Zhu C, Zhang Q, *et al.* 3D interest point detection based on geometric measures and sparse refinement[C] // *Proceedings of the IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*. Los Alamitos: IEEE Computer Society Press, 2016: 1-6
- [18] Creusot C, Pears N, Austin J. A machine-learning approach to keypoint detection and landmarking on 3D meshes[J]. *International Journal of Computer Vision*, 2013, 102(1-3): 146-179
- [19] Shu Z Y, Xin S Q, Xu X, *et al.* Detecting 3D points of interest using multiple features and stacked auto-encoder[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2019, 25(8): 2583-2596
- [20] Shu Zhenyu, Qi Chengwu, Xin Shiqing, *et al.* Unsupervised 3D shape classification algorithm using density peaks[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2016, 28(12): 2142-2150(in Chinese)
(舒振宇, 祁成武, 辛士庆, 等. 基于密度峰值的三维模型无监督分类算法[J]. *计算机辅助设计与图形学学报*, 2016, 28(12): 2142-2150)
- [21] Tonioni A, Salti S, Tombari F, *et al.* Learning to detect good 3D keypoints[J]. *International Journal of Computer Vision*, 2018, 126(1): 1-20
- [22] Shu Z Y, Yang S P, Xin S Q, *et al.* Detecting 3D points of interest using projective neural networks[J]. *IEEE Transactions on Multimedia*, 2022, 24: 1637-1650
- [23] Zhu X Y, Du D, Huang H B, *et al.* 3D keypoint estimation using implicit representation learning[J]. *Computer Graphics Forum*, 2023, 42(5): Article No. e14917
- [24] Li J X, Gim H L. USIP: unsupervised stable interest point detection from 3D point clouds[C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Los Alamitos: IEEE Computer Society Press, 2019: 361-370
- [25] Shu Zhenyu, Yi Shun, Yang Sipeng, *et al.* Detecting points of interest on 3D model via alternating optimization strategy[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2022, 34(7): 1095-1107(in Chinese)
(舒振宇, 易顺, 杨思鹏, 等. 采用交替优化策略的三维模型兴趣点提取算法[J]. *计算机辅助设计与图形学学报*, 2022,

34(7): 1095-1107)

[26] Graves A, Mohamed A R, Hinton G. Speech recognition with deep recurrent neural networks[C] //Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2013: 6645-6649

[27] Zhou P, Shi W, Tian J, *et al.* Attention-based bidirectional long short-term memory networks for relation classification[C] //Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 2016: 207-212

[28] Lian Z H, Godil A, Bustos B, *et al.* SHREC'11 track: shape retrieval on non-rigid 3D watertight meshes[C] //Proceedings of the 4th Eurographics Conference on 3D Object Retrieval. Aire-la-Ville: Eurographics Association Press, 2011: 79-88

[29] Sun J, Ovsjanikov M, Guibas L. A concise and provably informative multi-scale signature based on heat diffusion[J]. Computer Graphics Forum, 2009, 28(5): 1383-1392

[30] Sipiran I, Bustos B. A robust 3D interest points detector based on Harris operator[C] //Proceedings of the 4th Eurographics Workshop on 3D Object Retrieval. Aire-la-Ville:

Eurographics Association Press, 2010: 7-14

[31] Godil A, Wagan A I. Salient local 3D features for 3D shape retrieval[OL]. [2023-09-09]. <https://doi.org/10.1117/12.872984>

[32] You Y, Lou Y J, Li C K, *et al.* Keypointnet: A large-scale 3D keypoint dataset aggregated from numerous human annotations[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 13644-13653

[33] Qi C R, Su H, Mo K C, *et al.* PointNet: Deep learning on point sets for 3D classification and segmentation[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 77-85

[34] Wang Y, Sun Y B, Liu Z W, *et al.* Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2019, 38(5): Article No. 146