

## 采用交替优化策略的三维模型兴趣点提取算法

舒振宇<sup>1,2)</sup>, 易顺<sup>3)\*</sup>, 杨思鹏<sup>4)</sup>, 刘予琪<sup>5)</sup>, 隆威<sup>6)</sup>, 金海容<sup>4)</sup>, 辛士庆<sup>7)</sup>, 吴双卿<sup>8)</sup>

<sup>1)</sup> (浙大宁波理工学院计算机与数据工程学院 宁波 315100)

<sup>2)</sup> (浙江大学宁波研究院 宁波 315100)

<sup>3)</sup> (浙江大学机械工程学院 杭州 310027)

<sup>4)</sup> (浙江大学 CAD&CG 国家重点实验室 杭州 310058)

<sup>5)</sup> (浙江大学信息与电子工程学院 杭州 310027)

<sup>6)</sup> (浙江大学工程师学院 杭州 310015)

<sup>7)</sup> (山东大学计算机科学与技术学院 青岛 266237)

<sup>8)</sup> (浙大宁波理工学院信息科学与工程学院 宁波 315100)  
(21925143@zju.edu.cn)

**摘 要:** 针对三维模型的兴趣点提取问题, 提出一种基于交替优化的全监督检测算法. 第 1 步, 利用多种特征描述符对人工标注好的三维模型进行特征提取, 得到每个顶点的特征向量, 将其作为神经网络的输入; 第 2 步, 使用双调和距离场为模型表面顶点赋予概率标签, 并将顶点标签值作为神经网络的输出; 第 3 步, 通过神经网络学习输入特征与输出标签之间的复杂映射关系; 第 4 步, 将训练后的神经网络对训练集进行预测, 并把兴趣点提取结果与人工标签进行对比, 根据对比差异进一步优化顶点标签值, 然后将顶点标签值作为输出、顶点特征向量作为输入, 继续优化神经网络. 将第 3 步和第 4 步重复多次进行交替优化, 最终得到一个较优的神经网络模型. 在公开数据集 SHREC 2011 上的实验结果表明, 由于采用了交替优化的策略, 所提算法在三维模型表面兴趣点提取的关键评价指标 FNE 和 FPE 上均优于传统算法, 准确率实现了平均 11 个百分点以上的提升.

**关键词:** 三维模型; 兴趣点; 交替优化

**中图法分类号:** TP391.41 **DOI:** 10.3724/SP.J.1089.2022.19082

## Detecting Points of Interest on 3D Model via Alternating Optimization Strategy

Shu Zhenyu<sup>1,2)</sup>, Yi Shun<sup>3)\*</sup>, Yang Sipeng<sup>4)</sup>, Liu Yuqi<sup>5)</sup>, Long Wei<sup>6)</sup>, Jin Hairong<sup>4)</sup>, Xin Shiqing<sup>7)</sup>, and Wu Shuangqin<sup>8)</sup>

<sup>1)</sup> (School of Computer and Data Engineering, NingboTech University, Ningbo 315100)

<sup>2)</sup> (Ningbo Institute, Zhejiang University, Ningbo 315100)

<sup>3)</sup> (School of Mechanical Engineering, Zhejiang University, Hangzhou 310027)

<sup>4)</sup> (State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou 310058)

<sup>5)</sup> (College of Information Science & Electronic Engineering, Zhejiang University, Hangzhou 310027)

<sup>6)</sup> (Polytechnic Institute, Zhejiang University, Hangzhou 310015)

<sup>7)</sup> (School of Computer Science and Technology, Shandong University, Qingdao 266237)

<sup>8)</sup> (School of Information Science and Technology, NingboTech University, Ningbo 315100)

收稿日期: 2021-05-12; 修回日期: 2022-04-30. 基金项目: 国家自然科学基金(61872321, 62172356); 浙江省自然科学基金(LY22F020026, LQ17F030002); 宁波市自然科学基金(2017A610108); 宁波市“科技创新 2025”重大专项(2020Z005, 2020Z007, 2021Z012). 舒振宇(1979—), 男, 博士, 教授, 硕士生导师, 主要研究方向为计算机图形学、机器学习、数据挖掘; 易顺(1997—), 男, 硕士研究生, 论文通信作者, 主要研究方向为机器学习、计算机辅助设计; 杨思鹏(1995—), 男, 博士研究生, 主要研究方向为机器学习、计算机辅助设计与图形学; 刘予琪(1992—), 男, 博士研究生, 主要研究方向为图像处理; 隆威(1996—), 男, 硕士研究生, 主要研究方向为机器学习、计算机辅助设计; 金海容(1997—), 男, 博士研究生, 主要研究方向为机器学习、计算机辅助设计与图形学; 辛士庆(1979—), 男, 博士, 副教授, 博士生导师, 主要研究方向为计算机图形学; 吴双卿(1981—), 女, 博士, 讲师, 主要研究方向为机器学习.

**Abstract:** A supervised 3D points of interest (POI) detection algorithm is proposed based on alternating optimization. Firstly, the geometric features of a 3D shape are calculated from several hand-crafted feature descriptors and used as the input of the neural network. Secondly, the biharmonic distance field is utilized to assign a label to each vertex, which is regarded as the neural network's output. Thirdly, the complex mapping relationships between the feature vectors and the labels are learned through the neural network. Fourthly, predictions are made on the training set using the trained neural network. The differences between the predicted points of interest and ground truth are compared to further optimize the vertices' labels, which are then used as the output to train the neural network. The third and fourth steps for alternating optimization are repeated for several times, and a neural network is finally obtained. The experimental results on the SHREC 2011 dataset show that, due to the alternate optimization strategy, our algorithm is better than the traditional methods in the key evaluation indicators FNE and FPE, and the accuracy of proposed algorithm has achieved an average improvement of more than 11%.

**Key words:** 3D model; points of interest; alternate optimization

三维模型表面的兴趣点也被称为特征点,其通常具有特定的几何与语义特征,是三维模型表面视觉效果突出的点.提取三维模型上的兴趣点,有助于解决数字几何处理中的许多问题.例如,在三维模型分割问题中<sup>[1]</sup>,Katz 等<sup>[2]</sup>利用兴趣点将网格模型划分为不同的部分,然后沿着部位的边界生成一致性的网格分割;在人脸识别问题中,Blanz 等<sup>[3]</sup>通过在人脸图像上标注兴趣点,以此估计人脸的三维形状和纹理,从而建模出强大且通用的人脸表示.此外,提取兴趣点对于面部表情识别<sup>[4]</sup>、网络的变形和编辑<sup>[5]</sup>等问题也有较大的帮助.近年来,随着三维建模技术的不断进步,三维模型也变得越来越复杂,对模型兴趣点的提取算法也提出了更高的要求.然而,传统算法仍存在准确率不高,难以应用于复杂模型等问题,需要提出新的三维模型兴趣点提取算法,克服传统算法存在的问题.

## 1 相关工作

传统三维模型兴趣点检测算法通常根据多边形面片上几何特征的显著性提取兴趣点<sup>[6]</sup>.例如,Wang 等<sup>[7]</sup>提出了基于数据引力的三维模型特征点检测算法,其通过选择高斯曲率值<sup>[8]</sup>最大的顶点作为初始点,然后计算每个顶点的权重,并删除阈值之下的多余点,从而得到三维模型表面的兴趣点.Sipiran 等<sup>[9]</sup>提出的基于 Harris 算子的兴趣点检测器对三维模型的形变具有较强的鲁棒性.Hu 等<sup>[10]</sup>提出在 Laplace-Beltrami 谱域中提取几何特征点的算法,将频率与三维模型的尺度信息对应,实现了

全局和局部匹配,即使对形变的模型也十分有效.除了几何特征之外,还可以根据其他描述方式检测三维模型上的兴趣点.Wang 等<sup>[11]</sup>提出的多尺度图卷积网络,将小波能量分解特征转换为新的描述符,能够适应模型上多种不同曲面.Wang 等<sup>[12]</sup>提出的局部特征描述符,将顶点周围的光谱特征编码为几何图像,对二维图像进行傅里叶变换,进而计算区域的特征信息;使用该描述符计算局部特征信息时,模型的分辨率和面片的三角剖分方式并不影响实验结果.

此外,图像信息也对三维模型兴趣点提取有较大的帮助.Castellani 等<sup>[13]</sup>通过计算三维模型的局部特征,对三维模型表面的每个顶点训练隐马尔可夫模型,并利用相似度的差异匹配不同视图的兴趣点.Tonioni 等<sup>[14]</sup>将兴趣点检测问题与网格分割问题相结合,使用给定的描述符对特征点进行分类,并使用随机森林分类器提取兴趣点.

除了利用顶点的几何特征外,还能在顶点上附加如语义标签等额外的特征信息,从而提高三维模型兴趣点提取算法的鲁棒性.Teran 等<sup>[15]</sup>通过编码人工标签的主观认知标准,并由此提取模型上的兴趣点,使提取结果更符合人类的主观认知.Creusot 等<sup>[16]</sup>使用多个局部表面描述符提取特征,利用提取的特征拟合出一个最优函数,利用该函数分离出模型的普通顶点和兴趣点.Salti 等<sup>[17]</sup>利用一个预定义的特征描述符检测器,将三维模型上的顶点划分为特征点和非特征点,从而实现三维模型兴趣点的提取.

传统的兴趣点提取算法中,有的使用特征描述符提取特征<sup>[14-15]</sup>,有的使用模型的二维投影提

取特征<sup>[18-19]</sup>。然而, 对于复杂的模型, 在某些区域内模型的顶点特征差异较小, 点与点之间距离不大, 直接提取特征点可能会造成较多误判。而在某些区域, 点与点之间的特征变化明显, 提取这些顶点后, 可能会遗漏邻近区域特征变化平滑的兴趣点。

为了解决以上问题, 本文提出一种三维模型兴趣点提取算法, 通过交替优化的策略强化对复杂模型的提取能力。实验结果表明, 本文算法通过结合改进的密度峰值聚类算法提取三维模型兴趣点, 能够获得比传统算法更好的特征点提取效果。

## 2 本文算法

本文算法的主要流程如图1所示, 主要分为三

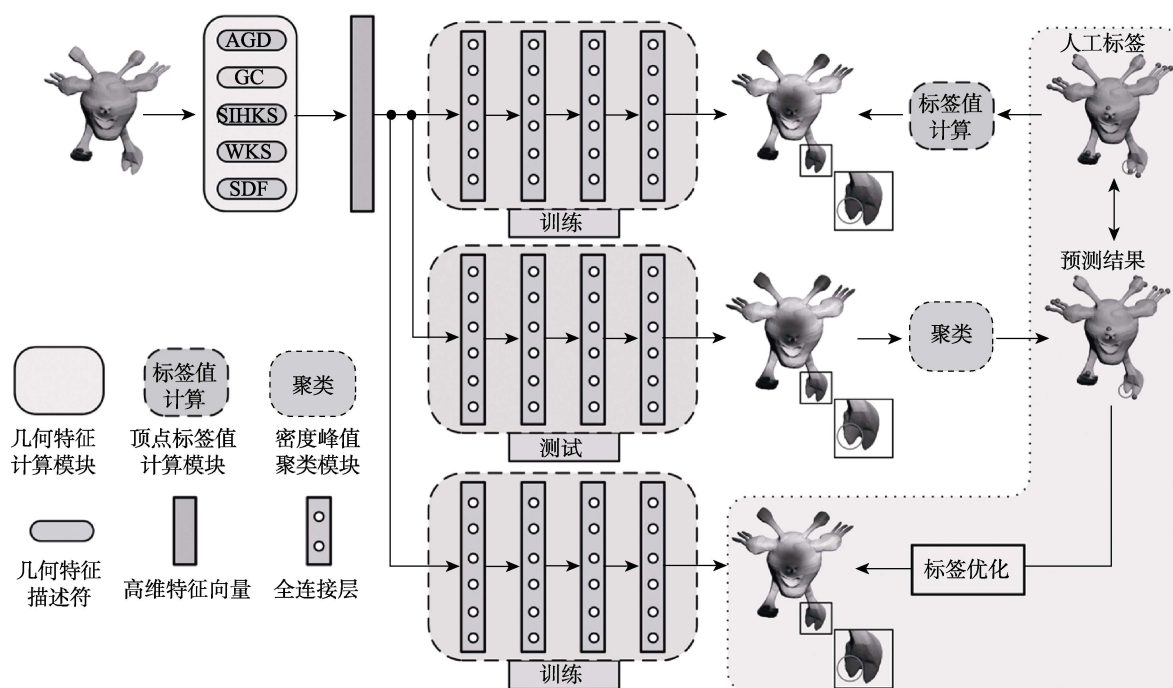


图1 本文算法流程示意

测试算法时, 提取需要测试的模型顶点的特征向量后, 将其作为测试数据输入经过训练的神经网络中, 从而得到模型表面顶点是兴趣点的概率, 并取得模型上所有顶点的概率分布; 然后结合改进的密度峰值聚类算法, 提取出概率峰值点作为测试集中模型的兴趣点。

### 2.1 模型表面顶点概率场的建立

在三维模型上, 兴趣点的数量相对较少。在如图2所示的猫模型兴趣点只有8个, 但非兴趣点有近万个。悬殊的正负样本比例给神经网络训练带来了较大的困难, 会导致神经网络模型泛化能力

维模型特征提取、训练神经网络和交替优化3部分。在三维模型特征提取阶段, 本文使用多种特征描述符<sup>[20]</sup>计算三维模型每个顶点的特征, 并组成一组特征向量, 将其作为神经网络的输入。在训练神经网络阶段, 本文算法将模型顶点的特征向量作为输入, 人工标记的数据作为标签, 通过神经网络学习三维模型几何特征与标签之间的映射关系。在交替优化阶段, 本文算法利用神经网络对输入数据预测, 使用改进的密度峰值聚类算法<sup>[21]</sup>提取兴趣点, 将提取出来的兴趣点与人工标记的标签进行对比, 根据差异对标签进行优化。在优化标签后, 本文算法将优化后的标签作为神经网络的输出, 继续训练神经网络, 最后得到一个预测三维模型兴趣点的分类器。

低下, 过拟合严重。

为了解决样本不平衡问题, 受到 Shu 等<sup>[6]</sup>算法启发, 本文利用双调和距离场<sup>[22]</sup>给顶点赋值, 利用高斯函数为兴趣点及其附近的顶点设置一个标签值, 从而在整体三维模型上构建一个概率场。通过对顶点赋值, 将提取兴趣点的问题转化为利用神经网络预测顶点是兴趣点的概率。在神经网络训练和测试阶段, 输入是每个顶点的特征向量, 输出是标签值。标签值越大的顶点是兴趣点的可能性越大, 标签值越小的顶点是兴趣点的可能性越小。

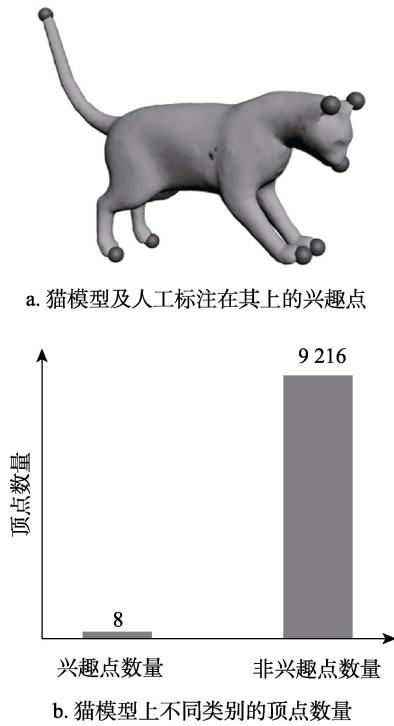


图2 猫模型上非兴趣点和兴趣点的数量对比

本文算法通过能量扩散的方式给每个顶点赋予标签值, 以此平衡数据样本的比例. 对于三维模型上的顶点  $v_i$ , 其标签值为

$$f_i = 1 - \frac{d_{\min}(c_k, v_i) - D_{\min}}{D_{\max} - D_{\min}} \quad (1)$$

其中,  $c_k$  表示模型上距离顶点  $v_i$  最近的兴趣点;  $d_{\min}(c_k, v_i)$  表示顶点  $v_i$  与兴趣点  $c_k$  的最小测地距离;  $D_{\max}$  和  $D_{\min}$  分别为  $d_{\min}(c_k, v_i)$  的最大值和最小值. 对于一个三维模型, 首先将其表面所有顶点代入式(1), 计算每个顶点的标签值, 然后得到此模型的概率场. 根据模型上人工标注的兴趣点与式(1), 在猫模型上得到的概率场如图3所示.

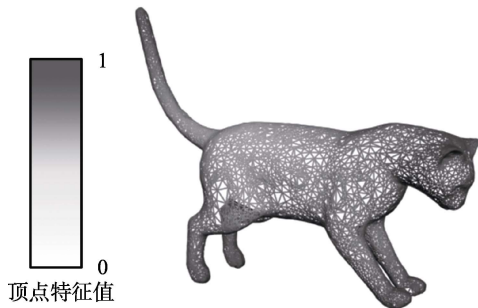


图3 由式(1)在猫模型上生成的概率场

由式(1)和图3可知, 根据测地距离大小给顶点赋值的方式是线性的, 大多数顶点被赋予了较

大的标签值, 不利于后续的聚类操作. 本文算法在式(1)的基础上将顶点标签值重新定义为

$$\rho_i = \varphi_{\mu, \sigma^2}(f_i) \quad (2)$$

其中,  $f_i$  为顶点  $i$  经过式(1)计算后的标签值; 函数  $\varphi_{\mu, \sigma^2}(x)$  是高斯分布函数, 其默认参数值为  $\mu=1$ ,  $\sigma=0.1$ .

本文算法在猫模型上最终生成的概率场如图4所示. 通过给模型表面顶点赋予标签值的方法, 将顶点的分类问题转化为预测顶点标签值的回归问题, 解决了训练样本不平衡的困难.

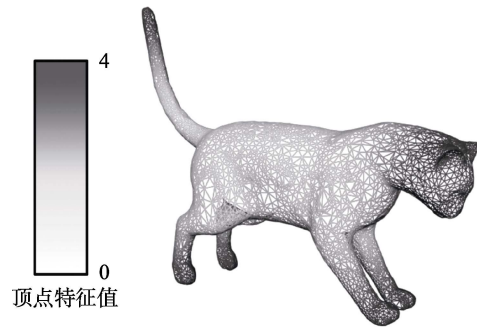


图4 本文算法在猫模型上生成的概率场

## 2.2 基于概率密度峰值聚类的兴趣点提取

利用神经网络预测每个顶点的标签值之后, 需要根据标签值从顶点中选择兴趣点. 由于本文使用能量扩散的方式给每个顶点赋予标签值, 因此不能仅由标签值的大小判断兴趣点. 对于每个顶点, 其标签值越大, 仅表示此点是兴趣点的可能性越大. 如图3所示, 用颜色的深浅描述模型顶点标签值的大小, 颜色越深表示顶点标签值越大. 但是, 标签值较大的顶点分布集中, 而且数量较多, 本文需要找到一种能够在标签值较大的众多顶点中提取兴趣点的方法.

### 2.2.1 密度峰值聚类算法

为了从三维模型的概率场中提取兴趣点, 受到 Rodriguez 等<sup>[21]</sup>和 Shu 等<sup>[6]</sup>算法的启发, 本文结合密度峰值聚类算法设计提取兴趣点的方法. 密度峰值聚类算法能够自动找出簇的中心, 实现任意形状数据的聚类. 该算法有2个前提假设: 密度峰值点的局部密度大于其周围邻居点的局部密度; 不同密度峰值点的距离相对较远.

本文在密度峰值聚类思想的基础上设计提取兴趣点的方法, 首先将三维模型顶点的标签值作为该点的密度, 然后将顶点间的测地距离作为聚类的参考距离, 最终得到一个关于距离和密度的

决策图. 在此决策图上, 横轴为每个点的密度值  $\rho$ , 纵轴为此点到距离最近的一个密度大于其顶点的测地距离值  $\delta$ . 构造的决策图如图 5 所示.

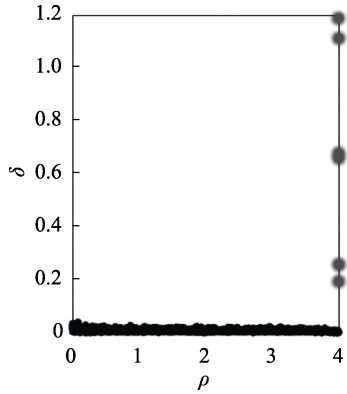


图 5 密度峰值聚类算法的决策图

构造决策图后, 算法将  $\rho$  值和  $\delta$  值都较大的点作为聚类的中心, 靠近决策图右上方的点  $\rho$  值较大, 表明该点密度值较高,  $\delta$  值也较大, 表明此点距离其他密度值高的点相对较远, 因此这些点就是局部密度峰值点, 也即需要提取的兴趣点. 利用该聚类的算法通过神经网络预测顶点的标签值后, 画出模型的决策图, 再通过决策图上点的分布提取出三维模型表面的兴趣点.

#### 2.2.2 在决策图上自动提取兴趣点

在决策图上准确找到兴趣点是本文算法的关键步骤. 在决策图上选择兴趣点有基于人工经验的手动提取和基于数据驱动的自动提取 2 种方法. 人工提取不仅耗时耗力, 而且在提取的过程中容易出现主观误差.

如图 6a 所示的鸟模型, 在神经网络预测顶点标签值后, 各个顶点在决策图上的分布情况如图 6b 所示. 若采用人工提取兴趣点的方法, 对于决策图右上方的顶点(图 6b 中深灰色区域), 由于它们  $\rho$  值和  $\delta$  值都相对较高, 因此能够正确地选择这些顶点; 而对于决策图右下角的部分兴趣点(图 6b 中浅灰色区域), 由于它们距离较近,  $\delta$  值较小, 因此其靠近横轴. 另外, 个人主观对这些点的判断不同, 所提取的结果会有一定的差异, 使一个或多个兴趣点被遗漏. 这种判断差异将导致提取结果不同, 进而影响后续交替优化过程及最终的实验结果. 因此, 找到一种自动提取兴趣点的方法十分必要.

本文采取数据驱动的方法自动提取决策图上的兴趣点. 在训练过程中, 本文算法从标签中学习每类模型的兴趣点提取阈值, 选择在阈值范围内的顶点作为兴趣点.

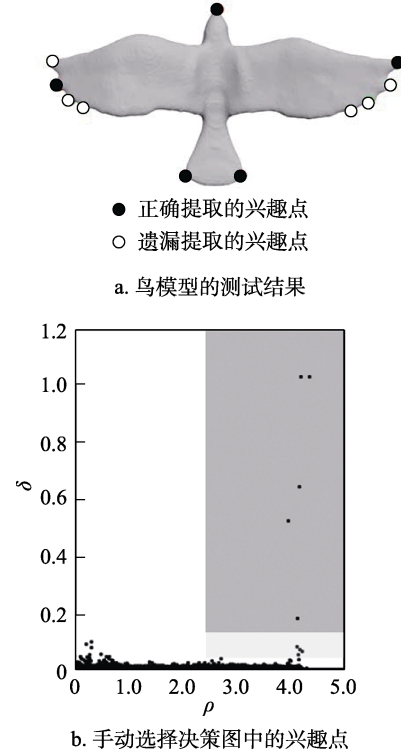


图 6 使用手动选择方法提取鸟模型的兴趣点

在神经网络训练过程中, 自动选择算法结合此类模型标签值在决策图上的分布, 计算选择兴趣点的阈值参数  $r_i$  和  $d_i$ , 即

$$\begin{cases} r_i = \frac{\rho_{\min}}{2} \\ d_i = \frac{\delta_k + \delta_{\min}}{2 \times \delta_{\min}} \end{cases} \quad (3)$$

其中,  $\rho_{\min}$  表示兴趣点中最小的  $\rho$  值;  $\delta_{\min}$  表示兴趣点之间的最小测地距离. 另外, 将非兴趣点中拥有最大  $\delta$  值的点命名为  $k$  点,  $\delta_k$  表示  $k$  点的  $\delta$  值.

如图 7 所示, 首先将一类模型的所有顶点都布置在决策图上, 然后利用式(3)确定出参数  $r_i$  和  $d_i$  的值. 在自动提取兴趣点的方法中, 利用参数  $r_i$  和  $d_i$  得到一个兴趣点选择区域, 选择此区域中的顶点作为兴趣点, 范围为

$$\begin{cases} \rho \geq r_i \times \rho_{\min} \\ \delta \geq d_i \times \delta_{\min} \end{cases} \quad (4)$$

由此, 得到兴趣点选择的提取范围. 实验证明, 此自动提取兴趣点的方式效果较好. 利用自动提取兴趣点的方法在鲨鱼模型上的自动提取效果如图 7 所示.

### 2.3 兴趣点的交替优化策略

传统的三维模型兴趣点提取算法大都从模型的整体出发, 寻找人工标记的兴趣点与非兴趣点

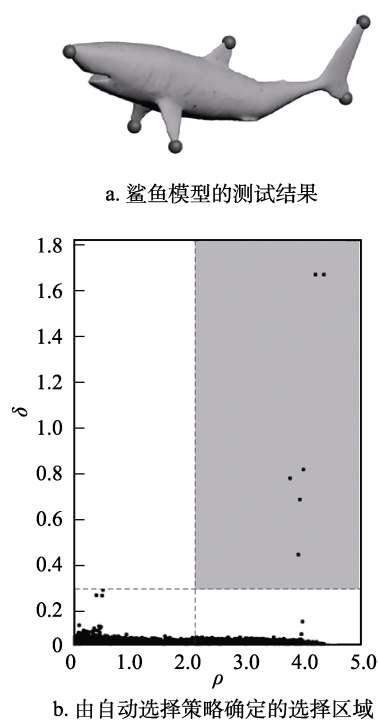


图7 使用自动选择策略提取鲨鱼模型的兴趣点

的集合特征差异,通过差异辨别新模型的兴趣点和非兴趣点.早期的三维模型扫描技术和建模技术还未发展完善,建立的三维模型结构较为简单,细节较为模糊,模型上人工标记的标签值位置差异较大,兴趣点提取算法只找出兴趣点的大致位置,提取结果不够准确.随着技术的进步,三维模型的结构越来越复杂,细节表达更加完善,这对三维模型相关处理算法要求也越来越高.

### 2.3.1 兴趣点错误提取问题

传统算法关注三维模型的整体特征,无法正确地提取特征区域内兴趣点;而改进之后的三维模型兴趣点提取算法中,单纯使用概率密度峰值聚类会受限于概率密度峰值之间的距离,导致兴趣点预测错误.如图8所示,在外星人模型中,触角部分具有突出的特征,在此区域中一般认为包含且仅包含一个人工标记的兴趣点.而在此区域中,各个顶点的特征向量十分相似,可能造成非兴趣点被错误提取的问题.同时,因为触角部分比较大,该部分局部峰值点之间的距离可能比脚尖的峰值距离更大,使用密度峰值聚类算法提取兴趣点时,会因为脚尖的峰值点距离较小,在决策图上 $\delta$ 值太小,较为靠近横轴,从而遗漏脚尖的兴趣点.

因此,本文提出采用交替优化的策略对模型的标签进行优化,使算法能正确地提取兴趣点.

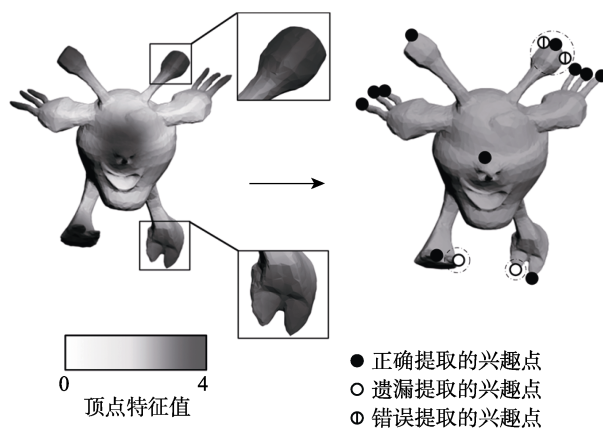


图8 外星人模型的提取兴趣点示例

### 2.3.2 利用交替优化提高显著性

本文算法在提取兴趣点时,特征突出的兴趣点容易被改进的密度峰值聚类算法选中,但部分兴趣点由于在决策图中位于右下角,选择兴趣点时可能出现误判.

为了准确地获得三维模型的兴趣点,本文采取交替优化的策略.在算法进行一轮神经网络训练之后,使用训练好的神经网络模型对训练模型进行一次预测,将预测结果体现在决策图上,再结合自动提取算法提取兴趣点,然后通过比较自动提取的兴趣点与人工标记的兴趣点之间的差异,对标签进行数值修正,从而提高神经网络的训练效果.

### 2.3.3 预测的点集划分

为了方便交替优化过程中对顶点标签的优化,本文算法首先对顶点的预测结果进行分类.与人工标记的标签相比,本文将自动提取出来的兴趣点  $A = \{A_1, A_2, \dots, A_k\}$  分为 2 大类:预测正确的点和预测错误的点.其中,预测错误的点又可以分为 2 类:遗漏的兴趣点和错误提取的非兴趣点.划分结果如下:预测正确的点为  $A_{\text{point}} = \{A_{p1}, A_{p2}, \dots, A_{pm}\}$ ,而遗漏的兴趣点为  $A_{\text{less}} = \{A_{l1}, A_{l2}, \dots, A_{ls}\}$ ,错误提取的非兴趣点为  $A_{\text{more}} = \{A_{m1}, A_{m2}, \dots, A_{mr}\}$ .其中,预测正确的点大多是在决策图中的靠近右上方的顶点,如图6b中深灰色区域的顶点.在决策图上选择顶点时,自动提取算法能够准确地框选出这些顶点.

遗漏提取的点大多是在决策图上特征不够突出的点,并且这些兴趣点的局部范围内有其他兴趣点.对于部分模型,如果某个兴趣点与另一个兴趣点之间的测地距离非常近,在决策图上则表现为 $\delta$ 值很小,靠近横轴,在提取兴趣点过程中,可

能会遗漏这些点, 如图 6b 中浅灰色区域的顶点。

错误提取的点大多分布在兴趣点附近, 它们是额外的局部峰值点。一个顶点越靠近兴趣点, 那么它的标签值  $\rho$  就越大。如果这个顶点与兴趣点的测地距离较大, 那么此点可能会成为局部峰值点, 在决策图上表现为  $\delta$  值较大。在决策图中, 如果  $\rho$  和  $\delta$  同时较大, 那么此点会靠近决策图的右上方。所以, 在利用密度峰值聚类算法进行兴趣点提取时, 这些局部峰值点会被错误提取。如图 8 所示, 外星人模型右边触角上形成了额外的局部峰值点, 从而造成了非兴趣点被错误提取。为了改善兴趣点提取效果, 找到一种能够减少非兴趣点被错误提取的方法十分必要。

### 2.3.4 交替优化的具体方法

本文算法自动提取兴趣点之后, 将提取实验结果与标签值进行对比, 对预测错误的顶点的标签值进行优化。对于预测正确的兴趣点, 它们大多分布在决策图右上角, 使用自动提取算法提取时出错的可能性很小, 本文算法对其标签值不作额外处理。

对于遗漏的兴趣点, 由于其  $\delta$  值太小, 与一些  $\rho$  值较大的非兴趣点一起分布在决策图的右下角, 造成提取时被遗漏。因此, 对遗漏的兴趣点采取的优化策略是利用

$$\tilde{\rho}_j = \rho_j - a \times d_{i,j} + b, 0 \leq d_{i,j} \leq t \cdot P_{\min} \quad (5)$$

线性提高此点周围邻域的  $\rho$  值。其中,  $\rho_j$  表示顶点  $p_j$  原来的  $\rho$  值;  $\tilde{\rho}_j$  表示线性修正后的  $\rho$  值;  $d_{i,j}$  表示遗漏的兴趣点  $p_i$  与当前点  $p_j$  之间的测地距离;  $P_{\min}$  表示该类模型训练集中兴趣点之间的最小测地距离;  $t$ ,  $a$  和  $b$  表示可以调整的线性提升参数, 本文根据实验结果对  $t$ ,  $a$  和  $b$  分别赋值为 0.40, 0.40 和 0.25。

对于分布在决策图右下角的兴趣点, 其本身的标签值  $\rho$  比较高, 但是其与其他兴趣点之间测地距离较近, 导致距离  $\delta$  值较小, 使得该兴趣点处在决策图右下角的位置。而使用本文算法提升其  $\rho$  值后, 该兴趣点到密度大于自身顶点之间的距离会增大, 达到  $\delta$  值增大的效果。在决策图上的具体表现为此点向决策图的右上角移动。通过线性修正的方法, 本文算法优化了提取遗漏的点及其附近的顶点的标签值。如图 9 所示, 用户可指定参数  $t$  调整测地距离, 从而控制修改标签值的顶点范围。

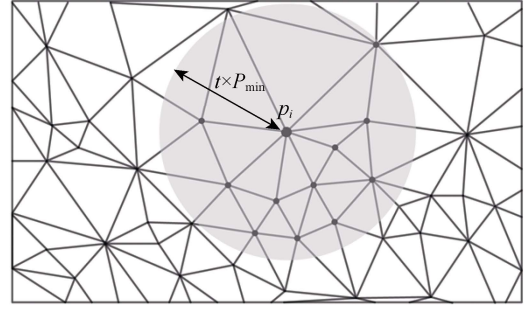


图9 指定参数  $t$  修正兴趣点  $p_i$  附近点的标签值

由于错误提取的点距离兴趣点较近, 本身  $\rho$  值较大, 而距离标签值大于自身的点有一定的测地距离, 因此它们会分布在决策图的右下角。在提取兴趣点时, 这些顶点容易被误识别成局部峰值点, 造成提取错误。由此, 本文算法对错误提取的点采取的优化策略是利用

$$\tilde{\rho}_j = \rho_j + a \times d_{i,j} - b, 0 \leq d_{i,j} \leq t \cdot P_{\min} \quad (6)$$

线性地降低此点周围顶点的  $\rho$  值。其中, 各参数的实际含义与式(5)相同。降低这些点的  $\rho$  值后, 它们到标签值大于自身顶点之间的测地距离会减小, 达到  $\delta$  值减小的效果。在决策图上的具体表现为这些点向决策图的左下角移动。

通过对标签值的优化, 算法改进了模型顶点标签值分布, 避免了产生部分局部峰值点, 突出了特征不明显的兴趣点, 从而提高了算法的准确率。

## 2.4 基于交替优化的兴趣点提取过程

### 2.4.1 特征描述符的选择

在神经网络训练之前, 本文算法需要对三维模型进行特征提取。目前学术界已有多种三维模型特征描述符可以进行模型的特征提取, 不同的特征描述符能从不同的角度刻画三维模型的特征。为了得到更好的特征描述效果, 本文使用了多种特征描述符, 分别为平均测地距离(average geodesic distance, AGD)的描述符<sup>[23]</sup>、形状直径函数(shape diameter function, SDF)<sup>[24]</sup>描述符、高斯曲率(Gaussian curvature, GC)描述符<sup>[25]</sup>、全局特征尺度不变热核特征描述符(scale-invariant heat kernel signatures, SIHKS)<sup>[26]</sup>和波核特征描述符(wave kernel signature, WKS)<sup>[27]</sup>。

本文算法使用以上 5 种三维形状特征描述符, 每种特征描述符对三维模型提取一定维度的特征, 得到相应的特征向量, 然后将特征向量串联成更高维的向量, 共同表示三维模型表面顶点的特征。实验表明, 这种组合方法能够很好地提取三维模型兴趣点。

### 2.4.2 算法流程

本文算法流程如下。

Step1. 对人工标注好的三维模型  $S = \{S_1, S_2, \dots, S_k\}$ , 利用 AGD, SDF, GC, SIHKS 和 WKS 这 5 种特征描述符提取三维模型表面所有顶点的特征向量, 并将得到的高维向量作为神经网络的输入  $L = \{g_1, g_2, \dots, g_n\}$ 。

Step2. 对三维模型上的所有顶点进行赋值, 得到每个顶点的标签, 将其作为神经网络的输出, 以此训练神经网络。

Step3. 经过一轮训练, 将得到的神经网络依次对每个训练集进行测试, 结合改进的密度峰值聚类算法画出决策图, 使用自动提取兴趣点的方法提取密度峰值点。

Step4. 将提取的密度峰值点与真实标签进行比较, 根据它们的差异对标签值进行优化。

Step5. 依旧将 5 个特征描述符提取的特征向量作为神经网络的输入, 将优化后的标签作为训练神经网络的真实值, 继续训练神经网络。

Step6. 重复 Step3~Step5, 直到 Step4 能够正确地提取训练集的兴趣点, 或者达到最大迭代次数 5 次, 从而得到一个较优的神经网络模型。

Step7. 输入测试模型, 利用 5 种特征描述符提取高维的特征向量作为神经网络的输入, 预测得到顶点的标签值。

Step8. 利用改进的密度峰值聚类算法自动提取峰值点, 得到测试模型的兴趣点。

## 3 实验与讨论

### 3.1 实验结果

本文算法在公开数据集 SHREC 2011<sup>[28]</sup>进行实验, 该数据集包含 30 类模型, 每类模型有 20 个, 共 600 个三维模型。部分三维模型如图 10 所示。

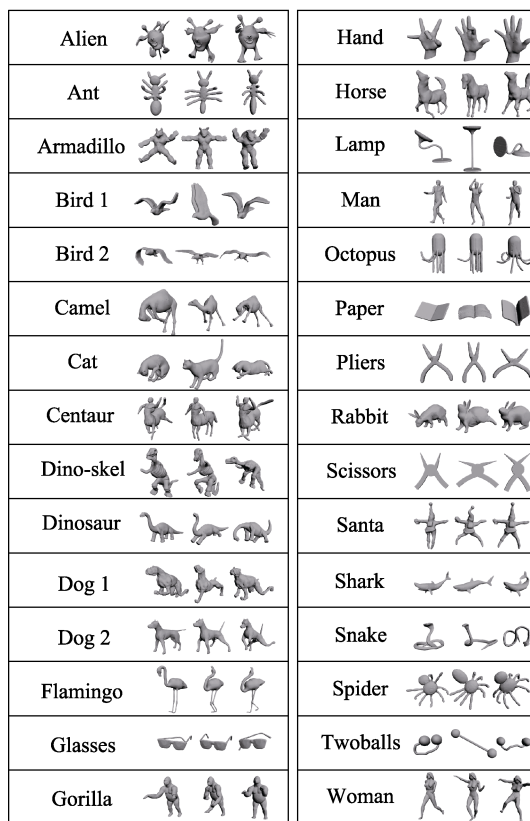


图 10 SHREC 2011 数据集中部分三维模型

本文算法在训练和测试的过程中, 随机从一类模型中抽取 70% 的样本, 即 14 个模型作为训练集; 另外 30% 的样本, 即 6 个模型作为测试集。图 11 所示为本文算法的实验结果, 与人工标记的标签对比可知, 本文算法预测出来的兴趣点与人工标记的兴趣点一致, 极少出现预测错误和遗漏的情况, 算法整体表现良好。训练神经网络时, 设置神经网络的学习率为 0.000 1, 使用随机梯度下降

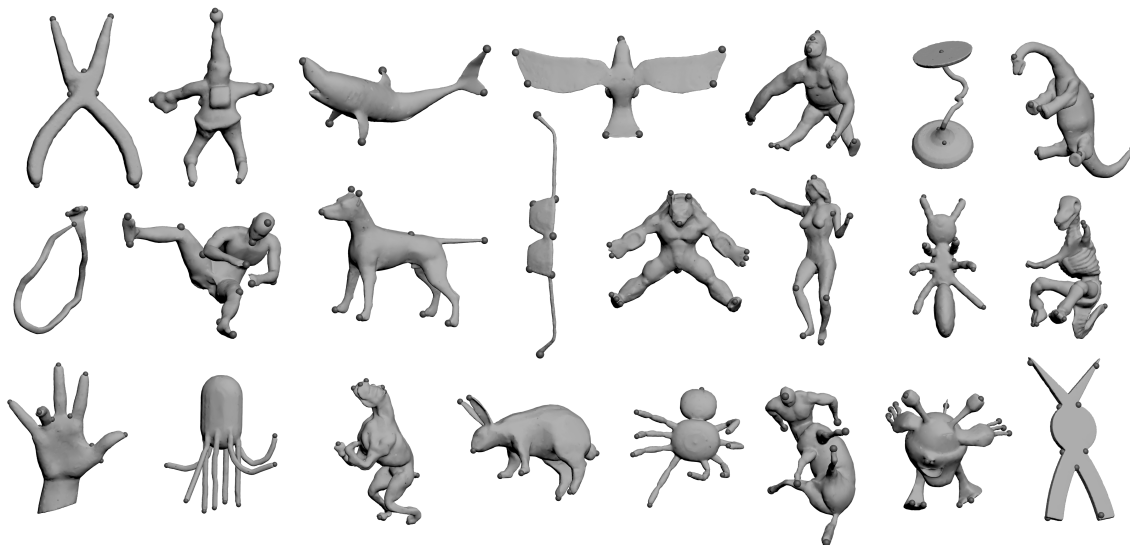


图 11 本文算法在 SHREC 2011 数据集上的部分测试结果

算法进行优化. 神经网络分为 4 层: 第 1 层神经元个数对应输入的特征向量的维数, 包含 122 个神经元; 第 2 层包含 50 个神经元; 第 3 层包含 25 个神经元; 第 4 层包含一个神经元. 实验结果表明, 这种神经网络的设计能得到较好的实验结果.

表 1 展示了本文算法在 SHREC 2011 数据集上的准确率. 在表 1 的数据中, 记录兴趣点之间的最大测地距离为  $d_{\max}$ , 在测地距离  $d \leq \lambda \times d_{\max}$  的范围内, 由本文算法提取到的顶点都被认为是正确兴趣点. 因为三维模型上的兴趣点并不是特指模型上某顶点, 而是代表某视觉效果突出区域的一个点. 例如, 手指尖视觉效果突出, 但构成手指尖的所有面片上有很多顶点, 每个顶点都可能被标注为兴趣点. 所以, 如果提取到的点距离人工标签的测地距离在一定的检测范围内, 则可被认为提取正确.

表 1 本文算法在 SHREC 2011 数据集上的表现 %

| $\lambda$ | 兴趣点准确率 | 兴趣点重复率 | 兴趣点遗漏率 |
|-----------|--------|--------|--------|
| 0.02      | 57.29  | 0.63   | 42.08  |
| 0.05      | 82.87  | 2.07   | 15.06  |
| 0.10      | 90.59  | 2.64   | 6.77   |

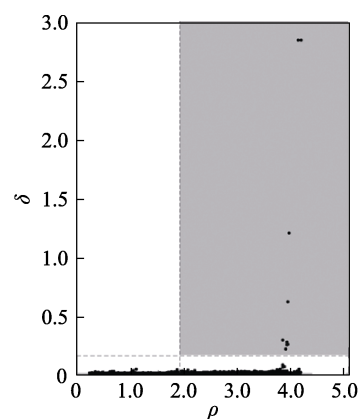
### 3.2 本文算法在复杂模型上的表现

本文算法通过交替优化的策略提取兴趣点的最大优势在于提升了对复杂三维模型兴趣点的提取能力, 尤其是在兴趣点数量较多且兴趣点特征不突出的模型上. 通过对标签的多轮优化和神经网络的多轮训练, 实现了算法在特征不明显的复杂模型中也能准确地提取到兴趣点. 图 12 展示了本文算法在复杂模型上的表现.

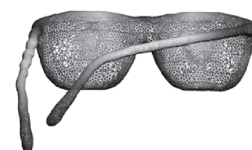
从图 12 中可以看出, 眼镜模型镜框部分兴趣点之间的测地距离很小, 在决策图中靠近右下方, 在提取过程中很有可能会被遗漏. 而交替优化的训练策略可以正确地提取这些兴趣点. 同时, 自动提取兴趣点的方法比人工提取方法更快, 神经网络预测出来的结果与人工标注基本一致.

### 3.3 与其他兴趣点提取算法的比较

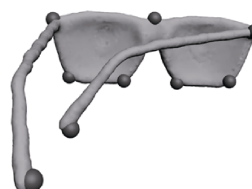
本文算法与 3D-SIFT<sup>[29]</sup>, 3D-Harris<sup>[30]</sup>和基于 HKS 的兴趣点提取算法<sup>[31]</sup>进行了比较, 比较结果如图 13 所示, 可以看出其性能比传统算法有一定的提升. 在神经网络预测出来的概率场中, 结合改进的密度峰值聚类算法提取概率峰值点作为算法结果. 由于在局部区域内的顶点特征向量相似, 因此算法提取到的兴趣点可能会在人工标签的附近, 而不是完全与人工标记的点重合. 随着检测范围



a. 决策图上的兴趣点分布

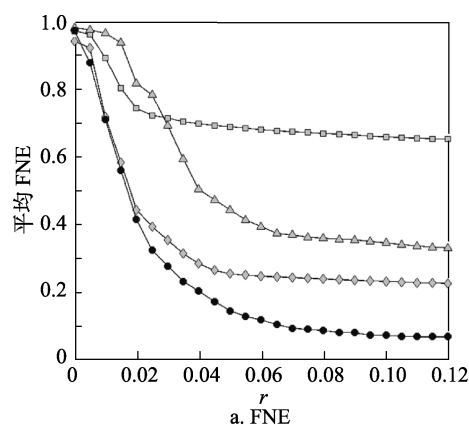


b. 在眼镜模型上预测的概率场

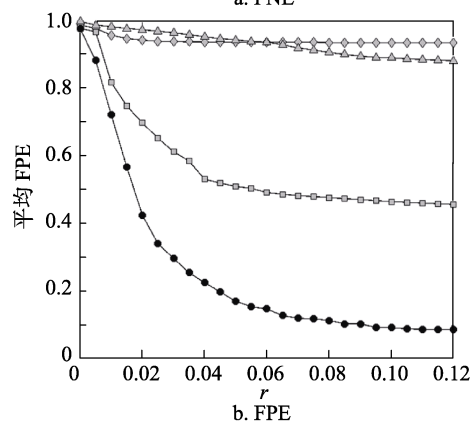


c. 兴趣点提取结果

图 12 本文算法在眼镜模型上的表现



a. FNE



b. FPE

—△— 3D-SIFT<sup>[29]</sup>; —□— 3D-Harris<sup>[30]</sup>; —◇— HKS<sup>[31]</sup>; —●— 本文

图 13 4 种算法在 2 个评价指标上的性能对比

增大, 各算法的误差都有减小的趋势, 到一定检测范围后趋于平稳.

本文采用 2 个评价指标对算法进行评价, 分别为假阴性误差(false negative error, FNE)和假阳性误差(false positive error, FPE), 它们的计算公式为

$$\begin{cases} \text{FNE} = 1 - \frac{K}{N} \\ \text{FPE} = 1 - \frac{K}{M} \end{cases} \quad (7)$$

其中,  $K$  表示在给定范围内判定为正确的兴趣点数目;  $N$  表示标签中的兴趣点数量;  $M$  表示算法提取的兴趣点数目. 根据式(7), FNE 表示没有被正确提取的兴趣点的比例, FPE 表示算法结果中不是正确兴趣点的比例. FPE 越小, 表示算法提取错误的点越少; FNE 越小, 表示算法提取兴趣点的准确率越高.

从图 13 的算法结果来看, 各算法随着检测范围增大, 遗漏的兴趣点比例会下降, 即 FNE 会下降. 同时, 各算法结果中错误提取的点的比例也会下降, 即 FPE 也会下降. 在给定检测范围  $r$  的情况下, 本文算法的准确率比传统算法平均高 11 个百分点以上.

结合 2 个评价指标分析可知, 3D-Harris 算法的 FNE 值较低, 但是 FPE 值较高, 这表明 3D-Harris 算法基本能够提取出所有兴趣点, 但同时在检测范围内错误提取了许多非兴趣点; 3D-SIFT 算法类似. 基于 HKS 的兴趣点提取算法 FPE 值较低, 但是 FNE 值较高, 这表明该算法错误率较低, 但是不能在模型上正确地提取出所有兴趣点, 造成兴趣点遗漏较多. 本文算法的 FNE 和 FPE 值都更低, 这表明相对于其他 3 种算法, 本文算法遗漏的兴趣点更少, 且错误提取的非兴趣点数量更少.

### 3.4 本文算法与未使用交替优化的算法比较

本文算法使用交替优化的方法提取三维模型的兴趣点, 与未使用交替优化的算法对比, 本文算法提取兴趣点的结果更好. 未使用交替优化算法的流程图如图 14 所示, 图中各模块的含义与图 1 相同. 未使用交替优化的算法训练过程和本文类似, 使用同样的三维模型和特征描述符, 采用相同的方法对神经网络进行训练. 其中, 未使用交替优化算法的神经网络迭代次数为使用交替优化算法的最大迭代次数. 2 种算法的对比结果如图 15 所示. 从图 15 的对比结果可知, 采用了交替优化的策略之后, 本文算法提取兴趣点的准确率更高.

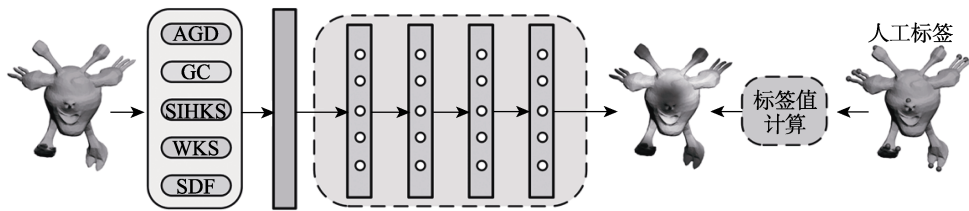


图 14 未使用交替优化的算法流程(其中各模块含义与图 1 相同)

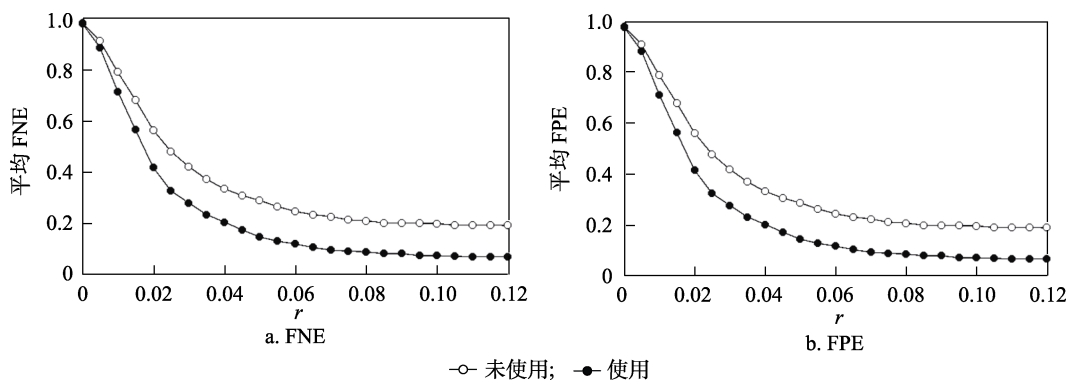


图 15 SHREC 2011 数据集上使用/未使用交替优化算法的性能对比

在 Intel Core TM i7 3.70 GHz CPU, 32 GB RAM 和 NVIDIA GeForce GTX 1080Ti GPU 的硬件设备上用 MATLAB 实现了本文算法. 在 SHREC 2011 数据集上的外星人模型类别中随机选取一定

数量的三维模型, 分别使用本文算法和未使用交替优化策略的算法训练神经网络, 重复 5 次后计算平均训练时间. 由以上过程得到算法效率对比结果, 如表 2 所示.

表 2 不同算法神经网络训练时间对比 min

| 算法      | 模型数量/个 |      |      |
|---------|--------|------|------|
|         | 2      | 5    | 10   |
| 本文      | 7.0    | 15.5 | 28.5 |
| 未使用交替优化 | 3.3    | 9.5  | 16.0 |

从表 2 可以看出, 在训练神经网络过程中, 因为涉及顶点标签值的交替优化过程, 所以本文算法的整体训练时间会比未使用交替优化算法的训练时间更长。

本文在公开数据集 SHREC 2007<sup>[32]</sup>上做了进一步实验. 该数据集包含 20 类模型, 每类模型有 20 个, 共 400 个模型, 此数据集与 SHREC 2011 数据集有较大不同, 能进一步体现实验效果. 在此数据集上使用交替优化进行模型兴趣点提取, 与没有进行交替优化的算法进行比较, 得到的算法结果对比如图 16 所示。

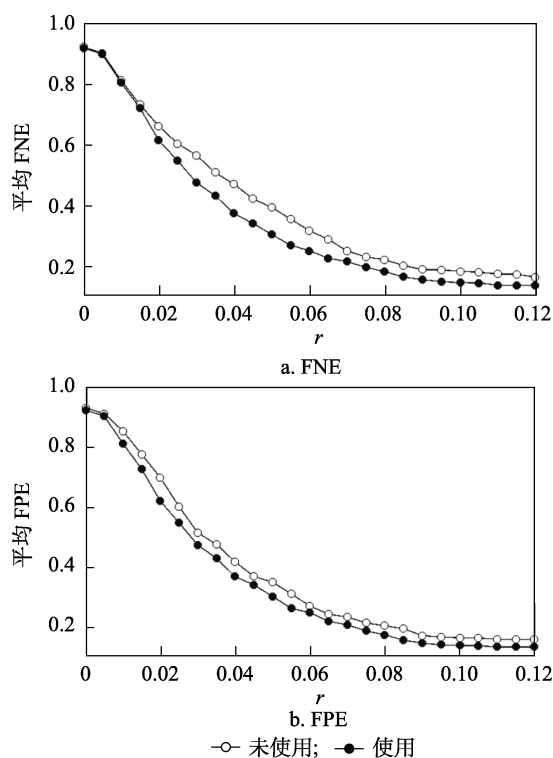


图 16 SHREC 2007 数据集上使用/未使用交替优化算法的性能对比

由图 15 和图 16 所示的结果可以看出, 在使用了 5 种特征描述符之后, 三维兴趣点提取的效果比较令人满意. 而使用了交替优化策略之后, 本文算法的效果能够在原有的基础上继续进行提升。

### 3.5 5 种特征描述符的对比实验

本文算法采用了 5 种特征描述符提取三维模型表面顶点的特征向量, 为了区分不同特征描述

符对算法性能的影响, 本文设计了 4 组对照实验. 在其他实验条件都相同的情况下, 分别使用不同的特征描述符组合进行性能测试, 得到的结果如图 17 所示。

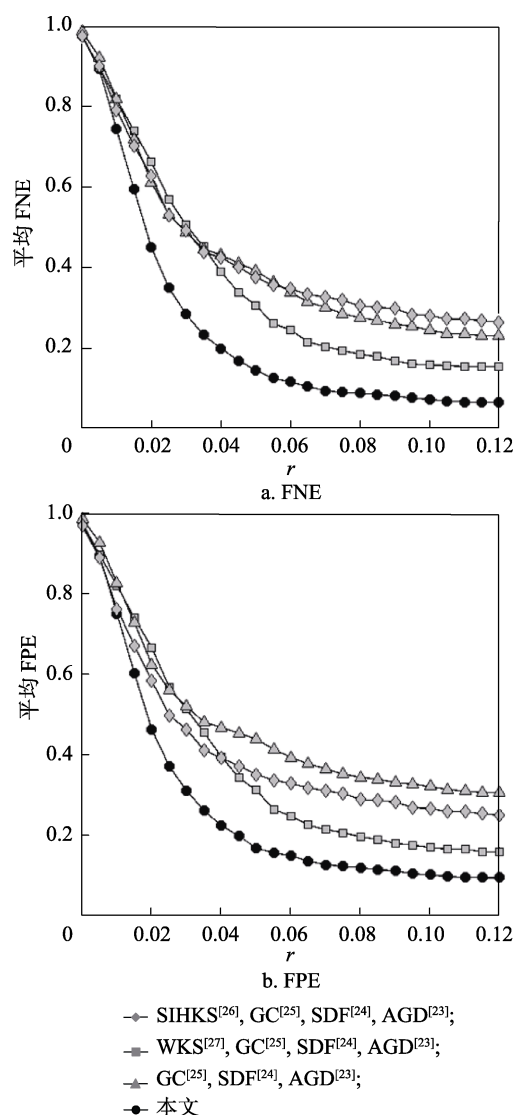


图 17 本文算法选取不同特征描述符对性能的影响

由图 17 可知, SIHKS 和 WKS 特征描述符比其他 3 种特征描述符对实验结果有着更大的贡献, 而这两者之间, WKS 描述符对算法性能提升的贡献更大. 综合考虑之下, 同时使用 5 种特征描述符时, 本文算法可以取得最好的性能。

### 3.6 本文算法的限制和不足

首先, 本文算法的性能依赖于人工标记数据的质量, 错误的人工标记将对算法性能产生较大的负面影响. 其次, 本文算法需要使用多种特征描述符提取模型顶点的特征向量, 其中, 大部分特征描述符只能应用于二维流形. 因此, 本文算法只适

用于二维流形。未来工作考虑如何将本文算法推广到非二维流形。

## 4 结 语

本文提出了一种基于交替优化的三维模型兴趣点提取算法。首先使用 5 种特征描述符提取三维模型表面顶点的特征向量, 将其作为神经网络的输入, 使用人工标注的标签对神经网络进行训练; 然后利用训练好的神经网络对训练集进行测试, 结合改进的密度峰值聚类算法提取三维模型表面的兴趣点; 再将得到的测试结果与人工标签进行对比, 利用其差异对人工标签进行优化; 完成标签的优化之后, 重新利用训练集对神经网络进行训练, 以此得到一个更好的神经网络模型。各类实验结果表明, 本文算法的效果明显优于传统算法, 能够准确地检测出三维模型表面的兴趣点。

## 参考文献(References):

- [1] 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
- [2] Katz S, Leifman G, Tal A. Mesh segmentation using feature point and core extraction[J]. *The Visual Computer*, 2005, 21(8/9/10): 649-658
- [3] Blanz V, Vetter T. Face recognition based on fitting a 3D morphable model[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, 25(9): 1063-1074
- [4] Valstar M, Martinez B, Binefa X, *et al.* Facial point detection using boosted regression and graph models[C] // *Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2010: 2729-2736
- [5] Kim Y, Varshney A. Saliency-guided enhancement for volume visualization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2006, 12(5): 925-932
- [6] 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
- [7] Wang C W, Kang D, Zhao X Y, *et al.* Extraction of feature points on 3D meshes through data gravitation[C] // *Proceedings of International Conference on Intelligent Computing*. Heidelberg: Springer, 2016: 601-612
- [8] Koenderink J J, van Doorn A J. Surface shape and curvature scales[J]. *Image and Vision Computing*, 1992, 10(8): 557-564
- [9] 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
- [10] Hu J X, Hua J. Salient spectral geometric features for shape matching and retrieval[J]. *The Visual Computer*, 2009, 25(5/6/7): 667-675
- [11] Wang Y Q, Ren J, Yan D M, *et al.* MGCN: descriptor learning using multiscale GCNs[J]. *ACM Transactions on Graphics*, 2020, 39(4): Article No.122
- [12] Wang Y Q, Guo J W, Yan D M, *et al.* A robust local spectral descriptor for matching non-rigid shapes with incompatible shape structures[C] // *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Los Alamitos: IEEE Computer Society Press, 2019: 6224-6233
- [13] 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
- [14] 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
- [15] Teran L, Mordohai P. 3D interest point detection via discriminative learning[C] // *Proceedings of European Conference on Computer Vision*. Cham: Springer, 2014: 159-173
- [16] 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(1/2/3): 146-179
- [17] Salti S, Tombari F, Spezialetti R, *et al.* Learning a descriptor-specific 3D keypoint detector[C] // *Proceedings of 2015 IEEE International Conference on Computer Vision*. Los Alamitos: IEEE Computer Society Press, 2015: 2318-2326
- [18] Yee H, Pattanaik S, Greenberg D P. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments[J]. *ACM Transactions on Graphics*, 2001, 20(1): 39-65
- [19] Guy G, Medioni G. Inference of surfaces, 3D curves, and junctions from sparse, noisy, 3D data[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(11): 1265-1277
- [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] Rodriguez A, Laio A. Clustering by fast search and find of density peaks[J]. *Science*, 2014, 344(6191): 1492-1496
- [22] Lipman Y, Rustamov R, Funkhouser T. Biharmonic distance[J]. *ACM Transactions on Graphics*, 2010, 29(3): Article No.27
- [23] Shapira L, Shalom S, Shamir A, *et al.* Contextual part analogies in 3D objects[J]. *International Journal of Computer Vision*, 2010, 89(2/3): 309-326
- [24] Shapira L, Shamir A, Cohen-Or D. Consistent mesh partition-

- ing and skeletonisation using the shape diameter function[J]. The Visual Computer, 2008, 24(4): 249-259
- [25] Gal R, Cohen-Or D. Salient geometric features for partial shape matching and similarity[J]. ACM Transactions on Graphics, 2006, 25(1): 130-150
- [26] Bronstein M M, Kokkinos I. Scale-invariant heat kernel signatures for non-rigid shape recognition[C] //Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2010: 1704-1711
- [27] Aubry M, Schlickewei U, Cremers D. The wave kernel signature: a quantum mechanical approach to shape analysis[C] //Proceedings of 2011 IEEE International Conference on Computer Vision Workshops. Los Alamitos: IEEE Computer Society Press, 2011: 1626-1633
- [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] Godil A, Wagan A I. Salient local 3D features for 3D shape retrieval[C] //Three-Dimensional Imaging, Interaction, and Measurement, 2011: 275-282
- [30] Sipiran I, Bustos B. A robust 3D interest points detector based on Harris operator[C] // Proceedings of the 4th Eurographics Conference on 3D Object Retrieval. Aire-la-Ville: Eurographics Association Press, 2010: 7-14
- [31] 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
- [32] Giorgi D, Biasotti S, Paraboschi L. Shape retrieval contest 2007: Watertight models track[J]. SHREC Competition. 2007: 7-8