

基于分层学习的三维模型兴趣点提取算法

舒振宇^{1,2)}, 杨思鹏^{3)*}, 辛士庆⁴⁾, 刘予琪⁵⁾, 龚梦航⁶⁾, 庞超逸^{1,2)}, 胡超⁷⁾

¹⁾ (浙江大学宁波理工学院计算机与数据工程学院 宁波 315100)

²⁾ (浙江大学宁波研究院 宁波 315100)

³⁾ (浙江大学机械工程学院 杭州 310027)

⁴⁾ (山东大学计算机科学与技术学院 青岛 266237)

⁵⁾ (浙江大学信息与电子工程学院 杭州 310027)

⁶⁾ (浙江大学工程师学院 杭州 310027)

⁷⁾ (浙江大学宁波理工学院信息科学与工程学院 宁波 315100)

(21825205@zju.edu.cn)

摘 要: 针对基于学习的三维模型兴趣点提取问题, 提出一种兴趣点分层学习的全监督算法. 提取三维模型表面所有顶点的特征向量后, 将人工标注的兴趣点分为稀疏点和密集点, 对于稀疏点使用整个三维模型进行神经网络训练, 对于密集点则找出兴趣点分布密集的区域进行单独的神经网络训练; 然后对 2 个神经网络进行特征匹配, 得到一个用于三维模型兴趣点提取预测的分类器. 测试时, 提取新输入的三维模型上所有顶点的特征向量, 将其输入到训练好的分类器中进行预测, 应用改进的密度峰值聚类算法提取兴趣点. 算法采用分层学习的策略, 解决了传统算法在模型细节处难以准确提取密集兴趣点的问题. 在 SHREC'11 数据集上的实验结果表明, 与传统算法相比, 该算法提取兴趣点的准确率更高, 出现的遗漏点和错误点更少, 对解决越来越精细的三维模型的兴趣点提取问题有较大帮助.

关键词: 三维模型; 三维模型兴趣点; 分层学习

中图法分类号: TP391.72 DOI: 10.3724/SP.J.1089.2020.17936

Detecting 3D Points of Interest Using Hierarchical Training Strategy

Shu Zhenyu^{1,2)}, Yang Sipeng^{3)*}, Xin Shiqing⁴⁾, Liu Yuqi⁵⁾, Gong Menghang⁶⁾, Pang Chaoyi^{1,2)}, and Hu Chao⁷⁾

¹⁾ (School of Computer and Data Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo 315100)

²⁾ (Ningbo Institute, Zhejiang University, Ningbo 315100)

³⁾ (School of Mechanical Engineering, Zhejiang University, Hangzhou 310027)

⁴⁾ (School of Computer Science and Technology, Shandong University, Qingdao 266237)

⁵⁾ (College of Information Science & Electronic Engineering, Zhejiang University, Hangzhou 310027)

⁶⁾ (Polytechnic Institute, Zhejiang University, Hangzhou 310027)

⁷⁾ (School of Information Science and Engineering, Ningbo Institute of Technology, Zhejiang University, Ningbo 315100)

Abstract: In this paper, we propose a novel supervised 3D points of interest (POIs) detection algorithm by using hierarchical training strategy. Firstly, the feature vectors of all the vertices of training shape are extracted, and the labeled POIs are divided into the part with sparse points and the part with dense points. Secondly, the two

收稿日期: 2019-06-23; 修回日期: 2019-07-23. 基金项目: 国家自然科学基金(61872321, 61572022); 浙江省自然科学基金(LQ17F030002, LQ19C060001); 宁波市自然科学基金(2019A610086); 宁波市领军和拔尖人才培养工程择优资助科研项目(NBLJ201801010); 宁波市面向生命健康的智能大数据工程应用创新团队(2016C11024); 宁波市科技计划项目(创新团队, 2014B82015); 浙江大学 CAD&CG 国家重点实验室开放课题(A1702). 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; 杨思鹏(1995—), 男, 硕士研究生, 论文通讯作者, 主要研究方向为机器学习、计算机辅助设计; 辛士庆(1979—), 男, 博士, 副教授, 主要研究方向为计算机图形学; 刘予琪(1992—), 男, 博士研究生, 主要研究方向为图像处理; 龚梦航(1994—), 男, 硕士研究生, 主要研究方向为机器学习、计算机图形学; 庞超逸(1960—), 男, 博士, 教授, 博士生导师, 主要研究方向为数据挖掘、数据压缩; 胡超(1960—), 男, 博士, 教授, 博士生导师, 主要研究方向为自动化控制、机器视觉.

grouped POIs are used to train neural networks. Finally, a 3D shape POIs classifier is obtained by matching the two neural networks with the feature vectors. In the testing process, the feature vectors of all the vertices are extracted and fed to the trained classifier for prediction. An improved density peak clustering algorithm is then used to detect the POIs. Our algorithm adopts the hierarchical training strategy, which can address the issue of accurately detecting the dense POIs in the model with details. The experimental results show that our method detects the POIs with higher accuracy when compared with the traditional algorithms. Both the false positive error and false negative error are greatly reduced by using our hierarchical training method.

Key words: 3D shape; 3D points of interest detection; hierarchical training

三维模型的兴趣点(points of interest)也被称为特征点(feature points), 是模仿人类的视觉感知特点在三维模型表面选取出来的一些具有几何和语义特征代表性的点. 三维模型的兴趣点被广泛用于三维模型分类分割^[1-2]、网格变形编辑^[3]、人脸识别^[4]、面部表情识别^[5]等多个方面. 近年来, 随着三维模型的应用范围越来越广, 模型的表达也越来越精细, 模型细节的提升对三维模型兴趣点提取算法的准确度和精细度提出了更高的要求.

一些兴趣点提取算法通过学习三维模型中人工标记的数据后, 在测试模型上找到与原模型中有相似特征的点作为新模型的兴趣点^[6]. 例如, Wang 等^[7]提出了基于测地距离的兴趣点提取算法, 首先根据曲面的高斯曲率^[8]选取所有可能成为兴趣点的顶点, 之后使用测地距离计算每个点的权重, 并与人工标记的点的权重作对比, 删去多余的与人工标记差异大的点, 就得到新模型上的兴趣点. 其特点是算法简单、易于实现, 但准确率不高. Sipiran 等^[9]提出一种基于 Harris 的三维模型兴趣点检测算法, 通过一个自适应的函数确定顶点的邻域以及邻域内的哈里斯响应, 使得兴趣点提取算法更具有鲁棒性. Hu 等^[10]将三维模型表面顶点的几何特征输入到由 Laplace-Beltrami 谱描述子构成的函数中, 之后选择函数响应最大的点作为新模型的兴趣点. 该算法没有直接使用从三维模型上提取出的几何特征, 而是将三维模型输入到一个屏蔽了尺度变化影响的函数中, 所以在它对不同形态的同一模型提取兴趣点的效果更好.

此外, Lowe^[11]提出了可以侦测和描述图像中的局部性特征的尺度不变特征变换(scale invariant feature transform, SIFT)算法, 其被应用于大量二维图像的特征点提取算法中^[12]. 也有一些学者将 SIFT 算法引入到解决三维模型的兴趣点提取的问

题上. 例如, Castellani 等^[13]用高斯差分(difference of Gaussian, DoG)函数构建了一个响应器, 用于测试三维模型上顶点的响应, 并将具有较高的 DoG 响应的顶点作为兴趣点. Zou 等^[14]提出了一个类似的算法, 首先构建一个三维模型顶点的测地距离的分布图, 再将 DoG 响应器应用到顶点的响应检测中, 以提取三维模型兴趣点. Tonioni 等^[15]将三维模型的兴趣点提取与三维模型网格分割相结合, 提出了一种使用随机森林分类器提取三维模型兴趣点的算法.

基于学习的三维模型兴趣点提取算法不仅可以快速地实现大批量三维模型兴趣点的提取, 还可以对输入训练的兴趣点附上语义标签, 使得输出预测的兴趣点上也能够带有相同的语义标签. Teran 等^[16]认为传统的兴趣点提取算法虽然在一定程度上描述了三维模型上兴趣点的几何特征, 但是提取方式的主观性不强, 不能处理多变的兴趣点提取任务, 所以他们提出基于全监督学习的兴趣点提取算法, 给兴趣点附上语义标签进行训练. Creusot 等^[17]提出了基于机器学习的三维模型兴趣点提取算法, 将三维模型兴趣点的几何特征与周围点的几何特征输入到一个待优化的函数中, 通过迭代训练学习区分兴趣点和非兴趣点. Salti 等^[18]将三维模型兴趣点的提取问题转换成顶点之间的分类问题, 并建立了一个基于三维形状特征描述符的兴趣点检测器, 以实现兴趣点和非兴趣点的分类.

传统的三维模型兴趣点提取算法已经从多个角度描述了兴趣点的提取, 有的使用三维模型的二维投影提取特征^[19-20], 也有的直接使用三维模型的特征描述符提取兴趣点的特征^[7-21]. 传统算法往往从三维模型的整体出发学习区分兴趣点和非兴趣点的特性, 然而, 由于模型细节处兴趣点的特

征差距很小, 很难从模型整体学习的角度找出这些细微的特征差距, 所以传统算法在模型细节处兴趣点密集的地方表现较差. 如对于一个人体的三维模型, 传统的算法能够很好地提取到手对应的兴趣点, 但是很难准确地提取到每一个指尖对应的兴趣点.

随着三维模型扫描技术和建模技术的不断进步, 三维模型的网格表达变得越来越精细, 模型包含的细节也越来越多, 传统的算法已经不能胜任越来越精细的兴趣点提取任务. 本文在三维模型兴趣点提取学习的训练过程中引入分层学习策略, 强化了兴趣点提取算法对三维模型细节处的处理能力, 并使用改进的密度峰值聚类算法提取出所有的兴趣点, 可以明显地提高提取三维模型兴趣点的准确率.

1 本文算法概述

本文提出一种基于分层学习的三维模型兴趣点提取算法, 其流程如图 1 所示, 主要分为训练过程和测试过程 2 个部分. 在训练过程中, 首先给定一批已经由人工标注过兴趣点的三维模型, 用特征描述符^[22]提取得到三维模型表面所有顶点的特征向量; 再使用一个带有自适应阈值的分层器, 将

人工标注的兴趣点分为细节兴趣点和普通兴趣点; 之后 2 组点通过兴趣点的能量扩散得到训练标签后, 分别输入到神经网络中训练, 再将 2 个神经网络训练的结果进行特征匹配, 得到一个用于三维模型兴趣点分类预测的分层器和一个用于细节兴趣点提取预测的检测器.

算法的测试过程对未标记的三维模型提取兴趣点, 首先对输入的三维模型的所有顶点提取特征向量, 并将这些向量输入到训练好的兴趣点分层器和兴趣点检测器中, 计算每一个顶点成为兴趣点的概率, 得到了模型上每一顶点的概率分布; 然后使用改进过的密度峰值聚类(density peaks clustering)算法^[23]提取出局部范围内概率突出的点作为兴趣点, 完成新模型兴趣点的提取.

2 兴趣点的分层学习策略

传统的基于学习的三维模型兴趣点提取算法往往从模型整体出发, 通过人工标记的兴趣点与非兴趣点的几何特征的差异来判断新模型上的哪些点是兴趣点. 早期的三维模型扫描技术和建模技术还处在起步阶段, 所以三维模型结构比较简单, 模型细节表达模糊, 兴趣点提取算法也只需要找出兴趣点大致的位置. 近年来, 随着技术的进步, 三维模型在细节表达越来越好, 对相关算法的精度也提出了更高的要求.

图 2 所示为一个人体的三维模型, 其中某些部位的兴趣点十分密集, 如指尖和指根等. 这些兴趣点之间相距很近, 所以特征向量也非常接近, 传统算法从三维模型整体出发学习, 难以学习到这些兴趣点的特征向量之间细微的差距.

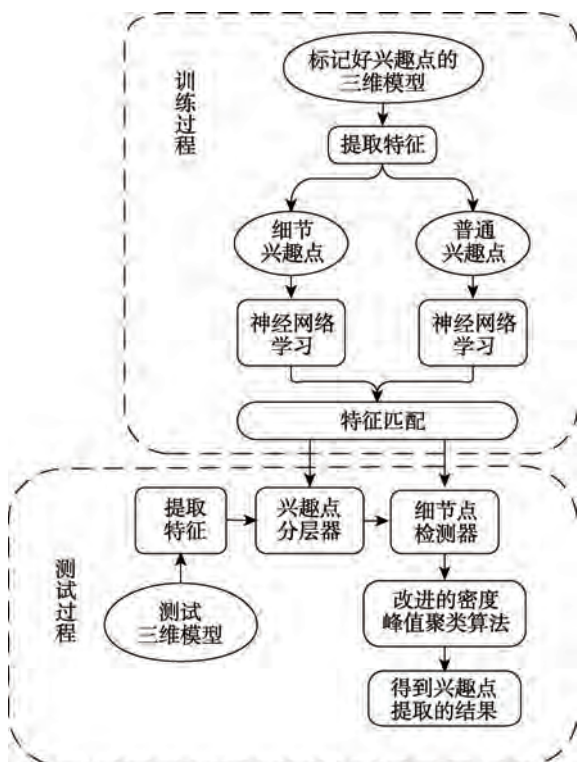


图 1 本文算法流程图



图 2 人体模型的兴趣点

因此, 本文算法采用兴趣点特征分层学习的策略, 将三维模型上所有的兴趣点 $P = \{p_1, p_2, \dots, p_n\}$ 分为图3所示黑色的普通兴趣点 $P_C = \{p_i, p_{ii}, \dots, p_s\}$ 和图3所示灰色的细节兴趣点 $P_D = \{p_1, p_{ii}, \dots, p_t\}$, 分别进行神经网络训练后再匹配到一起. 普通兴趣点 P_C 定义为兴趣点之间相距很远的点, 如图2中头、膝盖、手臂处的兴趣点; 细节兴趣点 P_D 定义为相距很近的兴趣点, 如手部和脚部的所有兴趣点. 应该注意的是, 某些兴趣点是相距很近的点群中的一个, 如手部的一个点代表了手部的所有兴趣点, 所以它们既属于普通兴趣点集 P_C , 又属于细节兴趣点集 P_D , 分组结果如图3所示. 具体的分层和神经网络训练策略如下.



图3 兴趣点分组结果

对于输入的人工标记好兴趣点的三维模型, 计算所有的兴趣点 P 之间的测地距离后将所有兴趣点 p_i 输入到分类器中, p_i 的分组方式为

$$p_i \in \begin{cases} P_C, \min_{i \neq j} d(p_i, p_j) > \delta \\ P_D, \min_{i \neq j} d(p_i, p_j) \leq \delta \end{cases} \quad (1)$$

其中, $d(p_i, p_j)$ 表示兴趣点 p_i 和 p_j 之间的测地距离; δ 表示一个自适应的距离阈值, 默认为 $\delta = 0.2 \times \max_{i \neq j} d(p_i, p_j)$, 即最大测地距离的 1/5. 分类器的作用是将某一区域内距离较远的兴趣点聚为一类, 距离较近的兴趣点聚为一类, 得到了比较稀疏的 $P_C = \{p_i, p_{ii}, \dots, p_s\}$; 将某一区域内密集的兴趣点分到 $P_D = \{p_1, p_{ii}, \dots, p_t\}$ 后, 还要从每一个密集兴趣点区域内随机选择一个点放入 P_C 内, 分

类的结果如图3所示.

$P_C = \{p_i, p_{ii}, \dots, p_s\}$ 均匀地分布在三维模型表面, 因此训练时要将三维模型上所有顶点输入到神经网络中, 得到一个用于整体的兴趣点提取的神经网络 N_C . 但是, 密集的 $P_D = \{p_1, p_{ii}, \dots, p_t\}$ 只分布在三维模型的部分位置, 所以训练时只提取细节兴趣点 P_D 及其附近的顶点输入到神经网络中, 得到用于局部网格提取兴趣点的神经网络 N_D .

在得到了 N_C 和 N_D 之后, 需要进行特征向量的匹配, 以实现 2 个神经网络网络的拼接, 最终得到一个可以准确地提取三维模型细节处兴趣点的分类器. 具体结果包括一个兴趣点的分层器和一个细节点检测器. 分层器可以提取出新输入的三维模型中所有可能的兴趣点, 并将其分为普通兴趣点 P'_C 和细节兴趣点 P'_D ; 之后细节点检测器再从 P'_D 附近找出全部的细节兴趣点 P''_D . 通过这种方式, 克服了传统算法在三维模型上兴趣点密集难以准确地提取兴趣点的困难. 理论上, 本文所提出的分层学习算法能够继续扩展, 将更精细的三维模型上的兴趣点分为 3 层甚至 4 层, 以实现三维模型局部细节处的兴趣点提取的学习和预测.

3 正负样本不均衡的神经网络训练

针对基于学习的三维模型兴趣点的提取算法, 学习训练时存在的最大问题是训练样本极度不均衡. 图4所示为初始正负数据样本比例, 正样本 $T_+ = \{x_i | l_i = 1\} (i=1, 2, \dots, m)$ 只有几个到十几个, 但是负样本 $T_- = \{x_i | l_i = 0\} (i=1, 2, \dots, n)$ 却有几千个, 这会导致训练出来的神经网络模型泛化能力较差, 容易发生过拟合.

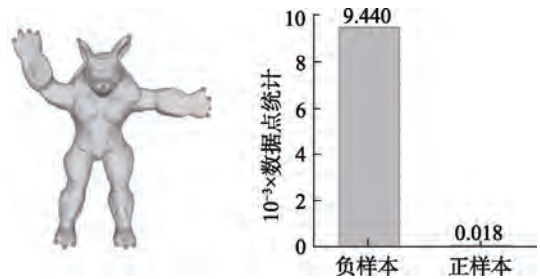


图4 初始正负样本

针对这个问题, 传统的基于学习的三维模型兴趣点提取算法提供了 2 种解决方案, 第 1 种方案首先计算三维模型表面所有顶点的特征向量, 然

后学习兴趣点与非兴趣点的特征向量之间的区分特性,最后利用训练好的分类模型对新输入的三维模型表面上的顶点进行分类,找到合适数量的兴趣点.另一种是采用双调和距离(biharmonic distance)^[24]为兴趣点及其附近的点设置一个逐渐衰减的值,通过这种方案扩大了训练中正样本数量,以保证神经网络训练的稳定性.

本文算法采用了第2种扩充正样本的方案,但是为了实现兴趣点能量扩散的自适应可调节,未使用双调和距离,而是设计了一个基于高斯函数的自适应激活函数 F ,其定义为

$$F_j = \varphi_{\mu, \sigma^2}(d(p_i, v_j)) \quad (2)$$

其中, $d(p_i, v_j)$ 表示顶点 v_j 到人工标记的兴趣点 p_i 的测地距离; $\varphi_{\mu, \sigma^2}(x)$ 表示关于 x 的高斯分布概率密度函数,默认 $\mu=1$, σ 表示模型最大测地距离的 $1/5$.

之后将包括兴趣点在内的所有顶点输入到激活函数 F 中,得到了一个自适应的从兴趣点开始扩散的标签分布,如图5所示,右侧展示了该模型上正负样本的直方图统计,对比图4可以明显看出,调整后的训练数据分布变得更加平滑,正样本 $T_+ = \{x_i | e_i > 0.5\} (i=1, 2, \dots, m)$ 数量也与负样本 $T_- = \{x_i | e_i \leq 0.5\} (i=1, 2, \dots, n)$ 数量达到了合适的比例.最后将处理过的数据点输入到神经网络中训练,训练测试表明,通过使用激活函数 F ,训练过程中验证准确率(accuracy)提升稳定,损失(loss)平稳地下降后趋于稳定.说明本文设计的激活函数 F 在不丢失兴趣点标签信息的情况下扩充了正样本,提升了神经网络训练的稳定性,这直接影响了之后对未标记的三维模型进行兴趣点提取预测结果的准确性和稳定性.

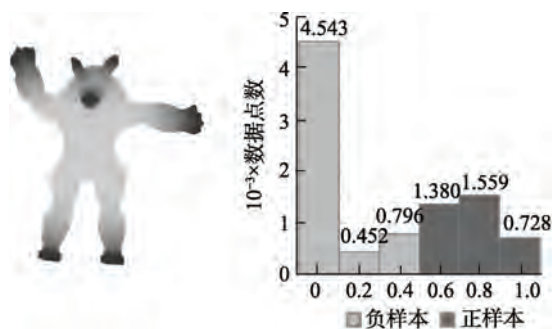


图5 修改后的正负样本

本文算法中神经网络的主要作用是学习从三维模型表面顶点 v_i 的特征描述符向量 I_i 到该顶点是兴趣点的概率 P_i 之间的映射关系.本文算法一

共用到2个神经网络,它们的学习率都设置为0.001.第1个用于学习特征描述符向量到一般兴趣点概率之间的映射关系;该神经网络共4层,第1层对应输入,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1个神经元.第2个用于学习特征描述符向量到细节兴趣点概率之间的映射关系.因为细节兴趣点之间的特征向量差距更小,所以第2个神经网络包含有更多的神经元;该神经网络设计为5层,第1层对应输入,第2层80个神经元,第3层40个神经元,第4层20个神经元,第5层1个神经元.实验结果表明,这2个神经网络都满足了本文算法的需求.

4 基于密度峰值聚类的兴趣点提取

对于输入待提取兴趣点的三维模型,神经网络预测结果是模型表面每一个顶点 v_i 是兴趣点的概率 P_i ,如图6所示,颜色越深,表示该顶点是兴趣点的概率越大.可以看出,通过神经网络预测之后,三维模型上的每一个顶点都被赋值了一个概率标签,并且越靠近兴趣点的顶点概率值越大.



图6 神经网络预测结果

为从图6所示的概率分布中提取出概率值高的区域的概率峰值点作为兴趣点,本文从 Rodriguez 等^[23]和 Shu 等^[6]算法中获得了启发.其中,密度峰值聚类(density peaks clustering)算法^[23]通过计算数据点在高维空间中的欧氏距离以及每个数据点在空间中的密度,得到一个横轴是数据点到密度大于自己的最近的数据点之间的距离 δ ,纵轴是每个数据点的密度 ρ 的决策图.决策图靠近右上角的数据点不仅密度大,而且距离密度比它大的数据点距离很远,这些点就可以作为聚类的中心,再将其他数据点按照聚类中心来分类.为了从概率图上提取峰值点,改进的密度峰值聚类算法^[6]将数据点的密度 ρ 改为三维模型上顶点的概率标签 ζ_i ,将数据点之间的距离 δ 改为三维模型顶点之间的测地距离 d ,使之可以在三维模型表面提取到局部范围内的概率峰值点.

本文采用了类似密度峰值聚类的算法提取三维模型表面概率峰值中心作为兴趣点, 但是在模型细节处, 兴趣点之间的测地距离 d 非常小, 导致在决策图上难以找到这些点. 为了能够在分层学习后从模型细节处准确提取到密集的兴趣点, 本文重新定义密度峰值决策图的横轴为

$$\sigma = \frac{1}{1 + \exp(-d(p_i, p_j) \cdot \zeta_i)} \quad (3)$$

其中, $d(p_i, p_j)$ 表示顶点 p_i 到概率值比它大的顶点 p_j 的测地距离; ζ_i 表示顶点 p_i 的概率值. 以此方式得到的概率峰值决策图如图7所示, 局部区域内的概率峰值点可以轻松地从决策图中找出来, 概率峰值点也就是算法最终要提取的兴趣点.

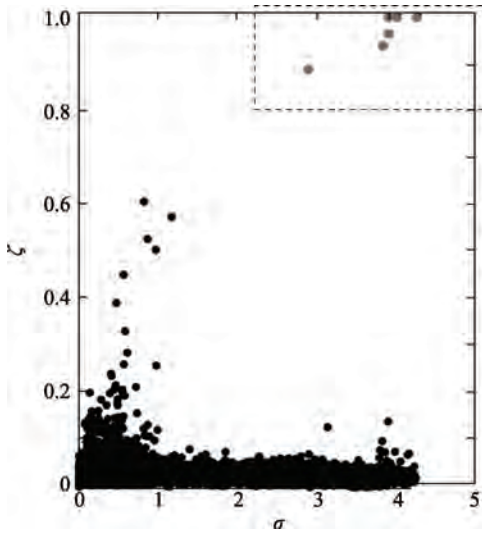


图7 概率峰值决策图

5 分层学习的兴趣点提取算法的实现

5.1 三维形状特征描述符选择

三维形状特征描述符可以刻画三维模型的形状特征, 一般将特征值赋值在模型表面的顶点或者三角面片上. 常用的特征描述符有: 刻画三维模型上顶点到其他顶点平均测地距离(vertex average geodesic distance, AGD)的描述符^[25]、刻画模型表面某一点到法向对面表面的欧氏距离的形状直径函数(shape diameter function, SDF)^[26]描述符、刻画模型表面曲率特征的高斯曲率(Gaussian curvature, GC)^[27]描述符、非刚性模型的全局特征尺度不变热核特征(scale-invariant heat kernel signatures, SIHKS)^[28]的描述符等. 不同的特征描述符从不同的角度描述三维形状的特征, 在应用过程中, 一般

同时使用多个特征描述符, 以达到更好的效果.

本文算法使用多种三维形状特征描述符串接成高维特征向量来提取模型表面顶点的特征, 但是在分层学习过程中, 不同的学习层需要不同的特征向量的组合. 普通兴趣点 P_C 分布在整个三维模型上, 所以在神经网络训练时, 本文算法选择 AGD, SDF, GC 和 SIHKS 串接成高维向量 l_i 作为普通兴趣点神经网络 N_C 的输入 $L_C = \{l_1, l_2, \dots, l_m\}$, 而细节兴趣点 P_D 只分布在三维模型的部分区域; 将 SDF 和 GC 串接成二维向量 k_i 作为细节兴趣点神经网络 N_D 训练的输入 $L_D = \{k_1, k_2, \dots, k_n\}$. 实验表明, 这种组合方法较好地适应了分层训练的特性.

5.2 实现步骤

给定同一类 m 个三维模型 $D = \{S_1, S_2, \dots, S_m\}$ 及每个模型上的人工标注兴趣点集 $P = \{P_1, P_2, \dots, P_m\}$, 本文算法通过对兴趣点进行分层训练, 实现了对精细的三维模型表面兴趣点的准确提取.

算法步骤如下:

Step1. 对于人工标注好兴趣点的三维模型 S_k , 利用多特征描述符提取三维模型表面所有顶点的特征描述符, 并将特征描述符 AGD, SDF, GC 和 SIHKS 组合成高维向量作为普通兴趣点 P_C 神经网络训练的输入 $L_C = \{l_1, l_2, \dots, l_m\}$; 再将特征描述符 SDF 和 GC 组合成二维向量作为细节兴趣点 P_D 神经网络训练的输入 $L_D = \{k_1, k_2, \dots, k_n\}$.

Step2. 将所有人工标记的兴趣点分为普通兴趣点 $P_C = \{p_1, p_{11}, \dots, p_s\}$ 和细节兴趣点 $P_D = \{p_1, p_{11}, \dots, p_t\}$, 使用激活函数 F 对兴趣点及兴趣点附近点的标签进行赋值, 将所有顶点的标签作为神经网络的输出, 以此训练 2 个神经网络.

Step3. 通过特征向量匹配将 2 个训练好的神经网络组合起来, 得到可以预测三维模型兴趣点的分层器和细节点提取器.

Step4. 输入一个新的三维模型, 利用多特征描述符提取三维模型所有顶点的特征向量, 将其输入到训练好的兴趣点分层器和细节点检测器中, 预测得到三维模型表面所有顶点是兴趣点的概率分布.

Step5. 利用改进的密度峰值聚类算法提取出顶点概率的峰值点, 将这些峰值点作为新模型的兴趣点.

6 实验与讨论

6.1 实验结果

本文算法选择 SHREC'11^[29] 三维模型数据集进行实验, 其中包含 30 类共 600 个三维模型, 每一

类有 20 个, 部分三维模型如图 8 所示.

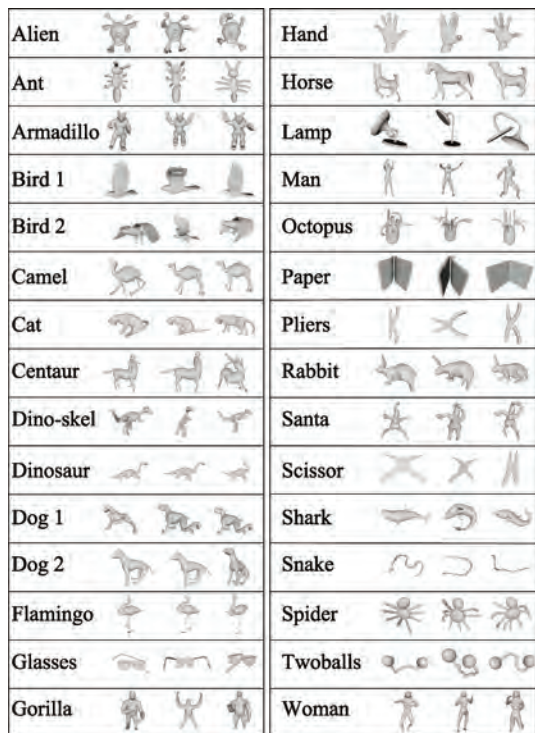


图 8 SHREC'11 数据集中部分三维模型

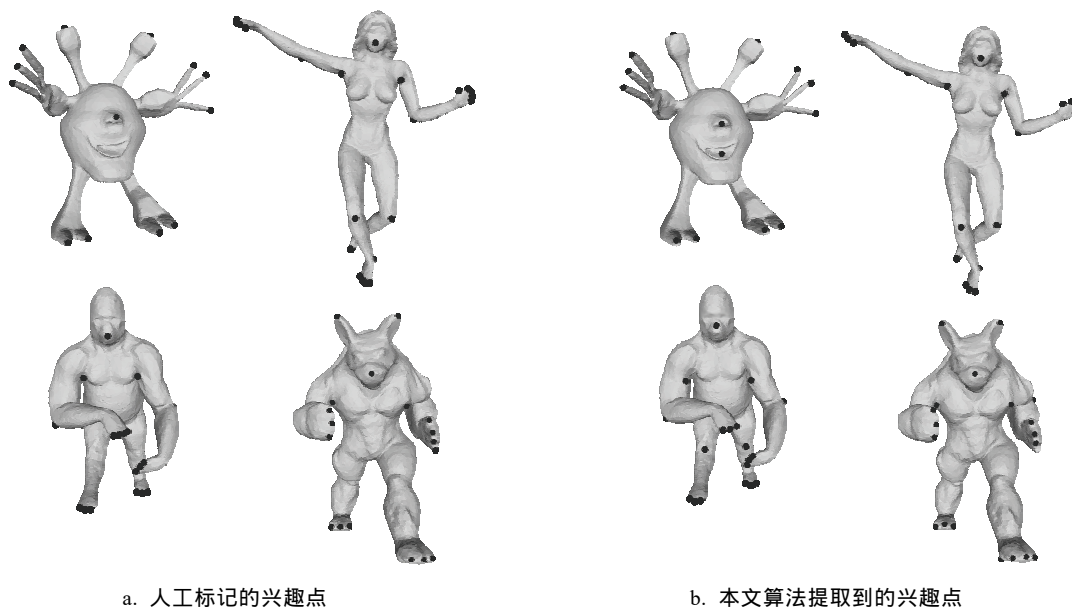


图 9 本文算法结果对比

阳性点准确率表示算法提取到的点是人工标记的兴趣点的比例, 阳性点重复率表示算法提取到的兴趣点中出现多个点对应一个人工标记的点的比例, 兴趣点遗漏率表示人工标注的兴趣点完全没有被算法提取到的比例. 表 1 中, 当检测范围为最大测地距离的 2% 时, 本文算法的阳性点准确

本文随机选择每一类中的 10 个模型作为训练样本, 另外的 10 个作为测试样本进行实验. 图 9 展示了本文算法兴趣点的提取结果与人工标记的兴趣点的对比, 可以明显地看出, 本文算法提取到的兴趣点与人工标记的兴趣点基本一致, 很少出现遗漏标记和错标记的兴趣点, 整体表现较好.

表 1 展示了本文算法在 SHREC'11 数据集上具体的准确率, 其中, 检测范围表示此统计是在人工标记的兴趣点周围一定范围内做的统计. 一般情况下, 由于在新模型上预测出的兴趣点几乎不可能与人工标记的点在同一位置, 大多数情况下只能提取到人工标记的兴趣点附近的一个点. 因此本文对不同范围内的点做了统计, 如 0.02 表示在人工标记兴趣点最大测地距离的 2% 范围内, 算法提取到的点都被认为是正确的兴趣点.

表 1 本文算法在选定数据集上的表现 %

检测范围	阳性点 准确率	阳性点 重复率	兴趣点 遗漏率
0.02	80.4	2.2	17.4
0.05	86.2	2.8	12.0
0.10	88.9	4.0	7.1

率已经到了 80% 以上, 同时阳性点重复率和兴趣点遗漏率都很低, 随着检测范围的扩大, 各项指标都有小幅提升, 本文算法提取三维模型表面兴趣点的结果整体表现较好.

6.2 本文算法在模型细节处的效果

本文算法提出的分层学习提取兴趣点的最大

优势在于提升了对三维模型细节处兴趣点的提取能力. 通过将模型上的某些部位单独提取出来训练, 实现了算法在特征差异较小的数据点中也能准确提取到兴趣点. 图 10 展示了本文算法在细节处的表现, 图 10a 是分层学习的神经网络预测的结果, 图 10b 是从概率值分布提取的兴趣点.

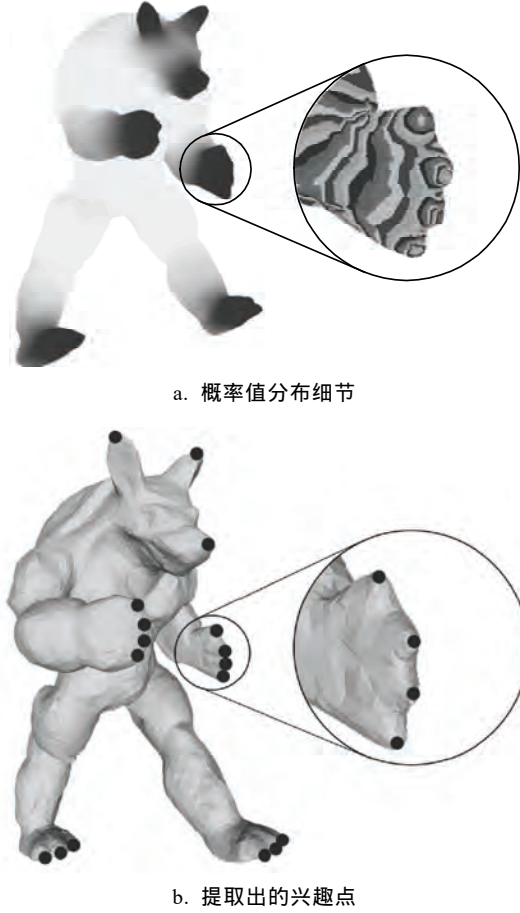


图 10 本文算法细节表现

从图 10 中可以看出, 分层学习训练可以使得兴趣点的概率值分布在密集的细节处依然保持较高区分度, 之后提取出的兴趣点准确率高, 并且可以明显地减少重复和遗漏的兴趣点数量, 达到更好的兴趣点提取效果.

6.3 噪声对神经网络鲁棒性的影响

三维模型可以由计算机软件进行建模获取, 也可以使用三维扫描仪扫描获取. 然而, 使用三维扫描仪进行扫描时, 采集到的信号会不可避免地掺入噪声. 为此, 本文设计了对模型添加噪声的鲁棒性实验, 以探究本文算法在含噪声的模型上的表现, 以及使用含噪声的模型进行训练对神经网络鲁棒性的影响.

在实验中, 本文对噪声三维模型添加了均值 $\mu=0$ 的高斯白噪声, 即对三维模型表面每一个顶点的 3 个坐标分量添加了一个随机的扰动. 添加噪声后的结果如图 11 所示. 左侧为原始模型, 表面较为光滑. 右侧为添加噪声后的模型, 模型表面变得十分粗糙.



图 11 模型添加噪声

本文设计了 4 组实验, 分别是: A 组噪声训练+无噪声测试, B 组无噪声训练+无噪声测试, C 组噪声训练+噪声测试, D 组无噪声训练+噪声测试. 选择狢狢模型进行测试, 测试结果依旧使用 FNE 和 FPE 来评价. 值得注意的是, 噪声训练是指在神经网络的训练过程中, 同时使用原始模型和添加了噪声的模型进行训练, 以测试添加噪声对神经网络鲁棒性的影响.

算法的评价指标使用假阴性误差(false negative error, FNE)、假阳性误差(false positive error, FPE)和加权后的遗漏误差(weighted miss error, WME)^[30], 它们分别定义为

$$\begin{cases} \text{FNE} = 1 - \frac{N_c}{G} \\ \text{FPE} = 1 - \frac{N_c}{P} \\ \text{WME} = 1 - \frac{\sum_{i=1}^G p_i \eta_i}{\sum_{i=1}^G p_i} \end{cases} \quad (4)$$

其中, N_c 表示在一定范围内判定为正确提取到的兴趣点数; G 表示人工标记的兴趣点数; P 表示由算法自动查找到的兴趣点数; p_i 表示多个人标记时顶点 v_i 的被标记为兴趣点次数. 当算法提取到的顶点 v_i 是兴趣点时, $\eta_i=1$; 否则 $\eta_i=0$. 按照

式(4)的计算方式, FNE 代表了兴趣点没有被算法提取到的比例, FNE 越小, 表示提取算法准确率越高; FPE 代表了算法提取到的点中不是兴趣点的比例, FPE 越小表示算法提取到的点错误越少; 而 WME 则是在统计中加入了人的主观评价, 被越多判定为是兴趣点的地方权重越大. 本文测试时只使用了一个人标记的点, 所以 WME 统计没有意义, 只统计了 FNE 和 FPE.

实验结果如图 12 所示, 图中展示了 4 组实验平均 FNE 和平均 FPE 在不同的检测范围 r 下的变化曲线. 从实验结果图可以发现, 在对带有噪声的粗糙的三维模型进行测试时, 算法表现没有在光滑的三维模型上表现好; 特别是 D 组在光滑的模

型上训练, 在粗糙的模型上测试, 算法结果表现最差. 对比 A 组与 B 组、C 组与 D 组可以发现, 在算法训练过程中加入带有噪声的训练样本可以获得更好的算法效果, 不仅平均 FNE 和平均 FPE 在检测的范围较小时下降较快, 而且最终稳定的 2 个误差值也相对较低. 因此, 在训练时加入带有噪声的训练模型可以明显提升神经网络的鲁棒性, 使本文算法在带有噪声的三维模型上测试表现更好.

6.4 与其他兴趣点提取算法的比较

本文与其他的 3 种传统的三维模型兴趣点提取算法进行了比较, 包括 3D-SIFT^[31], 3D-Harris^[32]和基于 HKS 的兴趣点提取算法^[30], 对比结果如图 13 所示. 因为算法提取到的点很难与人标记点重合,

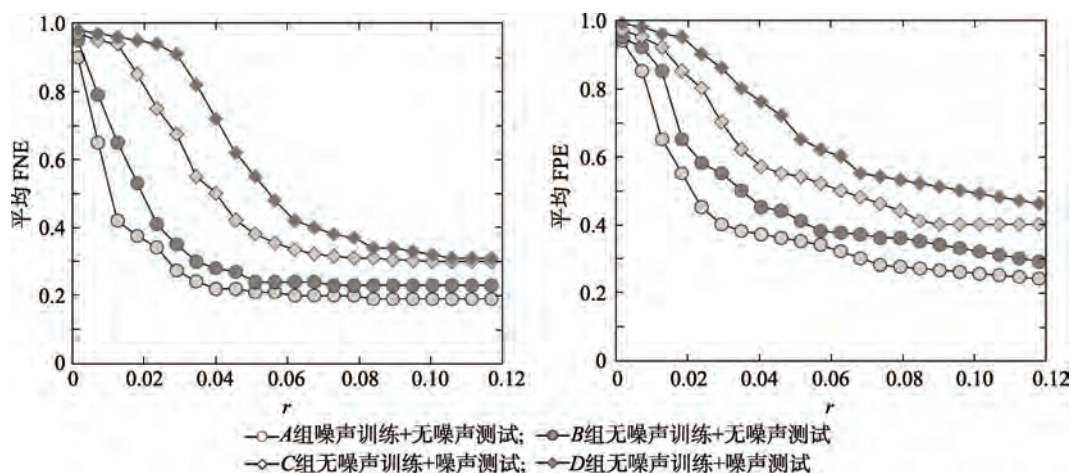


图 12 噪声鲁棒性实验结果对比

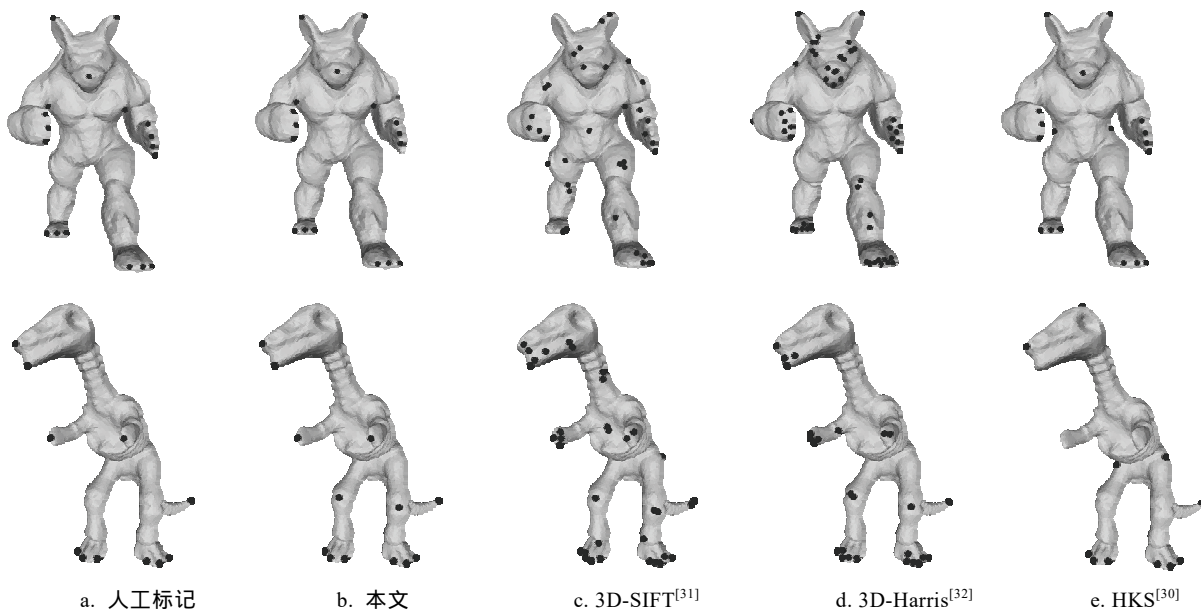


图 13 不同算法细节表现对比

一般是出现在人工标记点的附近, 所以随着判定范围的增大, 各算法的误差值都有减小的趋势, 之后又趋于稳定. 可以明显地看出, 本文算法的兴趣点提取结果的误差评价指标都低于传统算法. 图 14 展示了 2 个误差曲线, 其中 3D-Harris 算法的 FNE 很低但是 FPE 很高, 说明该算法提取到的兴趣点太多了, 虽然覆盖了较多的人工标记的点, 但是其中的错误点也很多. 类似地, HKS 算法 FPE 值很低, FNE 值很高, 说明该算法虽然提取到的点是

人工标记兴趣点的比例很高, 但是提取不够充分, 很多人工标记的兴趣点没有提取到. 这与图 14 展示的各算法的对比结果相符, 所以, 要同时获得较低 FNE, FPE 和 WME 是不容易的, 也进一步证明了本文算法提取兴趣点的高准确率.

值得指出的是, 本文算法结果的 2 个误差曲线在 $r \in (0, 0.03)$ 的范围内下降非常明显, 说明本文算法提取到的兴趣点与人标记的兴趣点之间的距离非常近, 提取到的兴趣点位置更加精准.

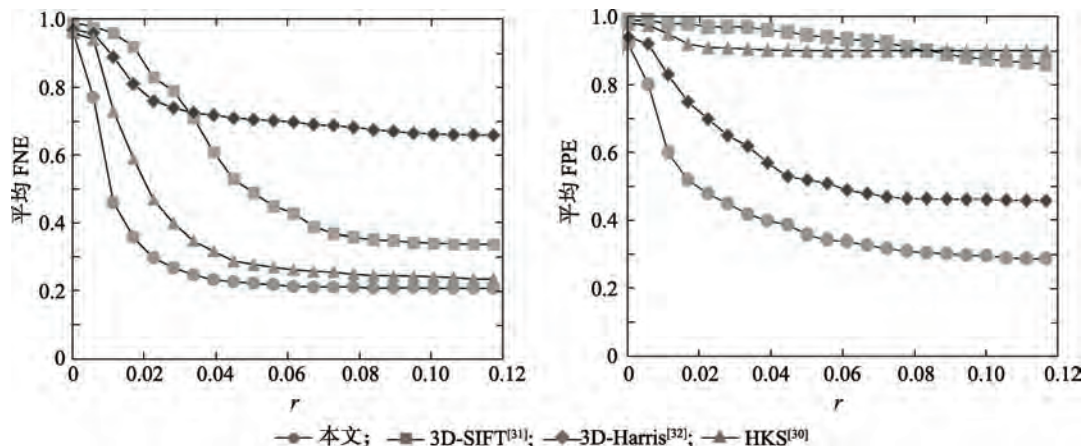


图 14 不同算法的误差曲线对比

6.5 本文算法的限制和不足

首先, 本文算法依赖三维模型特征描述符来提取三维模型每一个顶点上的特征向量. 然而, 许多特征描述符只能适用于二维流形, 因此, 本文提出的提取三维模型兴趣点的算法只能应用于二维流形.

其次, 本文算法对数据集内的所有模型都采用同一套特征描述符来提取几何特征, 但是不同类别的模型对不同特征描述符的敏感程度不同, 这也是限制本文算法提取到的兴趣点的准确率的主要原因. 未来, 可以考虑探索一种自适应的特征描述符选择方法来提升兴趣点提取的准确率.

7 结 语

本文提出一种分层学习的三维模型兴趣点提取算法. 首先将模型细节处密集的兴趣点提取出来, 得到普通兴趣点集和细节兴趣点集, 将其分别输入到神经网络中训练, 之后对两个训练好的神经网络进行特征匹配, 得到一个可以先提取三维模型上稀疏的兴趣点, 再提取三维模型部分区域内密集的兴趣点的分类器. 对于输入待提取兴趣

点的三维模型, 将提取到的模型表面所有顶点的特征输入到分类器中进行预测, 最后使用改进过的密度峰值聚类算法提取预测概率分布的峰值点, 得到输入模型的兴趣点. 实验结果表明, 本文算法的提取兴趣点的结果明显优于传统算法, 兴趣点遗漏率和重复率都很低. 并且模型局部区域概率分布图证实了本文分层学习算法的有效性, 提取到的兴趣点的位置也更加准确.

此外, 分层学习方法在模型处理细节问题上有很大的优势, 可以找出相近的特征向量之间细微的差别, 在实际的运用过程中, 可以考虑使用增强学习的训练方式, 理论上能够进一步发挥分层学习的优势. 对于兴趣点提取的分层学习问题, 使用增强学习训练方式能否提高算法预测结果的准确性和鲁棒性, 值得进一步探索.

参考文献(References):

- [1] Katz S, Leifman G, Tal A. Mesh segmentation using feature point and core extraction[J]. The Visual Computer, 2005, 21(8-10): 649-658
- [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] Kim Y, Varshney A. Saliency-guided enhancement for volume visualization[J]. IEEE Transactions on Visualization and Computer Graphics, 2006, 12(5): 925-932
 - [4] 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
 - [5] Valstar M, Martinez B, Binefa X, *et al.* Facial point detection using boosted regression and graph models[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2010: 2729-2736
 - [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 the 12th 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: Article No.963
 - [10] Hu J X, Hua J. Salient spectral geometric features for shape matching and retrieval[J]. The Visual Computer, 2009, 25(5-7): 667-675
 - [11] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110
 - [12] Guo Y L, Bennamoun M, Soheli F, *et al.* 3D object recognition in cluttered scenes with local surface features: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(11): 2270-2287
 - [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] Zou G Y, Hua J, Dong M, *et al.* Surface matching with salient keypoints in geodesic scale space[J]. Computer Animation and Virtual Worlds, 2008, 19(3/4): 399-410
 - [15] 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
 - [16] Teran L, Mordohai P. 3D interest point detection via discriminative learning[C] //Proceedings of the 13th European Conference on Computer Vision. Heidelberg: Springer, 2014: 159-173
 - [17] 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
 - [18] Salti S, Tombari F, Spezialetti R, *et al.* Learning a descriptor-specific 3D keypoint detector[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2015: 2318-2326
 - [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] 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
 - [21] Chen X B, Saparov A, Pang B, *et al.* Schelling points on 3D surface meshes[J]. ACM Transactions on Graphics, 2012, 31(4): Article No.29
 - [22] 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)
 - [23] Rodriguez A, Laio A. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492-1496
 - [24] Lipman Y, Rustamov R M, Funkhouser T A. Biharmonic distance[J]. ACM Transactions on Graphics, 2010, 29(3): Article No.27
 - [25] 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
 - [26] 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
 - [27] Gal R, Cohen-Or D. Salient geometric features for partial shape matching and similarity[J]. ACM Transactions on Graphics, 2006, 25(1): 130-150
 - [28] Bronstein M M, Kokkinos I. Scale-invariant heat kernel signatures for non-rigid shape recognition[C] //Proceedings of the 23rd IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2010: 1704-1711
 - [29] 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
 - [30] 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
 - [31] Godil A, Wagan A I. Salient local 3D features for 3D shape retrieval[J]. Proceedings of SPIE, 2011, 7864(1): Article No.78640S
 - [32] Sipiran I, Bustos B. A robust 3D interest points detector based on Harris operator[C] //Proceedings of the 3rd Eurographics Workshop on 3D Object Retrieval. Aire-la-Ville: Eurographics Association Press, 2010: 7-14