

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2025)08-2690-21

论文引用格式: Ding X Y, Zhang X F, Lin W S and Chen Z Z. 2025. Visual saliency detection for 3D object models. Journal of Image and Graphics, 30(8):2690-2710(丁晓颖, 张新峰, Lin Weisi, 陈震中. 2025. 三维模型视觉显著性检测研究综述. 中国图象图形学报, 30(8):2690-2710)
[DOI:10.11834/jig.240614]

三维模型视觉显著性检测研究综述

丁晓颖¹, 张新峰², Lin Weisi³, 陈震中^{4*}

1. 中南财经政法大学信息工程学院, 武汉 430073; 2. 中国科学院大学计算机科学与技术学院, 北京 100049;
3. 新加坡南洋理工大学计算机科学与工程学院, 新加坡 639798, 新加坡; 4. 武汉大学遥感信息工程学院, 武汉 430072

摘要: 三维模型视觉显著性检测通过模拟人类视觉系统, 定位模型中蕴含重要视觉信息的区域, 广泛应用于模型简化、分割以及压缩等相关任务, 是计算机视觉领域的研究热点。区别于图像和视频数据, 三维模型体量庞大、数据分布不均匀, 同时噪声数据较多, 使得现有图像和视频视觉显著性检测方法难以被拓展应用, 为三维模型视觉显著性检测任务带来了巨大的挑战。针对这一现状, 首先对近年来国内外公开发表的三维模型视觉显著性检测方法进行概述, 根据使用特征的不同将三维模型视觉显著性检测方法分为基于手工特征的方法和基于深度学习的方法, 并根据不同三维模型表示形式分别介绍面向点云模型的方法和面向网格模型的方法。对基于手工特征的方法, 根据特征尺度的不同将其细分为基于单尺度特征的方法和基于多尺度特征的方法, 重点介绍特征提取策略; 对基于深度学习的方法, 重点介绍设计思路。同时, 对现有三维模型视觉显著性检测数据集进行汇总, 介绍三维模型视觉显著性检测常用的性能评价指标, 并对部分方法进行性能对比。此外, 本文详细介绍视觉显著性检测结果在三维模型缩放、简化、降噪以及质量评价等相关领域的应用。最后, 基于国内外研究现状, 讨论三维模型视觉显著性检测领域亟需解决的问题, 并指出未来可能的发展方向。

关键词: 三维模型视觉显著性检测(3D-VSD); 三维模型处理; 视觉注意机制; 眼动追踪; 特征提取; 深度学习

Visual saliency detection for 3D object models

Ding Xiaoying¹, Zhang Xinfeng², Lin Weisi³, Chen Zhenzhong^{4*}

1. School of Information Engineering, Zhongnan University of Economics and Law, Wuhan 430073, China;
2. School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China;
3. School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, Singapore;
4. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China

Abstract: With the fast development of point cloud acquisition technology and the popularity of portable point cloud acquisition equipment, generating 3D model with high density and rich texture information is much easier than ever before. A three-dimensional (3D) model can represent the real world with rich information and has been widely applied in various applications in our daily life, such as smart cities, autonomous driving, virtual reality, and product design, making 3D

收稿日期: 2024-10-21; 修回日期: 2024-12-29; 预印本日期: 2025-01-05

* 通信作者: 陈震中 zzchen@whu.edu.cn

基金项目: 国家自然科学基金项目(62036005, 62071449); 湖北省自然科学基金项目(2022CFB984); 中国博士后科学基金项目(2024T171055, 2022M713503); 武汉大学测绘遥感信息工程国家重点实验室开放基金项目(22I03); 武汉市自然科学基金项目(2025040601020142)

Supported by: National Natural Science Foundation of China (62036005, 62071449); Natural Science Foundation of Hubei Province, China (2022CFB984); China Postdoctoral Science Foundation (2024T171055, 2022M713503); Open Fund of State Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University (22I03); Natural Science Foundation of Wuhan City, Hubei, China (2025040601020142)

model processing a hot research topic in the fields related to computer vision and 3D graphics. However, the growing number and size of 3D models bring great challenges to storage, transmission, and processing. With this huge amount of data, directly processing the whole 3D model in real time is impractical or uneconomical in many situations. Inspired by the selective mechanism of the human visual system (HVS), which indicates that the HVS selectively processes the huge amount of visual information that comes into our eyes and distributes more of the limited attention to regions with important visual information, researchers begin to investigate whether this selective mechanism can be applied to deal with 3D models. In essence, 3D visual saliency detection aims to detect the visually salient regions on 3D models and can help with various human-centered applications in 3D model processing, e. g., resizing, simplification, segmentation, and quality assessment. Different from 2D image and video data which have regular pixel distribution, a 3D model contains depth information and noise data and is usually with large data size. More often than not, points are distributed unevenly in 3D space, making it hard to extend existing image and video visual saliency detection methods to deal with 3D models straightforwardly. In addition, in contrast to 2D image and video visual saliency detection which have been studied by researchers for decades, research toward 3D visual saliency detection has just started during recent years. Considering the fact that humans are the final users for most small 3D object while machines are final users for large-scale/city-scale scenarios, this paper summarizes studies related to visual saliency detection for 3D object models to provide a comprehensive survey. First, we summarize existing 3D visual saliency detection methods. Considering the different kind of features used by researchers, these 3D visual saliency detection methods can be classified into two categories: handcrafted feature-based methods and deep learning-based methods. The former category can be further divided into two different types, namely, the single-scale feature-based methods and the multi-scale feature-based ones. In addition, according to different 3D representations, these methods can also be divided into 3D point-cloud based methods and 3D mesh based ones. We also investigate their feature extraction strategies and multi-feature integration frameworks. Compared with traditional handcrafted feature-based methods, deep learning-based methods achieve higher similarity with the ground-truth human labeling and fixation density maps. The reason is that deep learning-based features usually have stronger expressive power when compared with handcrafted features. Then, we summarize 3D visual saliency detection databases, which can be used to achieve and evaluate the performance of different 3D visual saliency detection methods. On the basis of different data collection approaches, these databases can be categorized into two different classes: mouse tracking-based database and eye tracking-based database. Limited by eye-tracking technology, earlier 3D visual saliency detection databases were mainly constructed using mouse-tracking technology. Compared with eye tracking, mouse tracking is cheaper and easier to implement. With development of eye-tracking technology in recent years, researchers tend to use eye-tracking technology to construct a 3D visual saliency detection database because recording the human visual behavior is more straightforward. During the eye-tracking experiment, several subjects are invited to observe 3D models or their projected images freely without any specific task. Their eye-movement data are collected during this observation and subsequently processed to construct the database. Moreover, we provide clear illustrations about the commonly used evaluation metrics for evaluating the performance of 3D visual saliency detection methods, such as the correlation coefficient score and the area under curve (AUC) score. In addition, we introduce how 3D visual saliency detection can be used as a guidance to improve the visual performance of other 3D model processing applications including 3D model resizing, simplification, denoising, and quality assessment. In 3D model resizing, uniformly adjusting the size of the entire 3D model may lead to unsatisfying visual performance because it could potentially cause excessive stretching of certain areas on the 3D model and lead to visual distortions, introducing visual saliency information can better preserve the structural features of the 3D model. In 3D model simplification, 3D visual saliency detection result identifies regions that are visually salient, which can guide the simplification procedure to preserve more details in visually salient regions and lead to more visual appealing results. In 3D model denoising, the visual saliency feature can help determine parameters of the denoising filter, allowing for enhanced preservation of detailed features in the denoised 3D model. In 3D model quality assessment, distortions in visually salient regions are more annoying than distortions appearing in nonsalient regions. Thus, accurately detecting visually salient regions on 3D model is crucial to 3D quality assessment and can help obtain more precise quality evaluation results. Finally, we provide an in-depth conclusion about 3D model visual saliency detection and discuss the problems that must be solved. Moreover, we

introduce the development trends of 3D model visual saliency detection from two perspectives: content of the 3D model and design of the algorithm, hoping that it can help with further improvement and trigger further exploration in the research community.

Key words: 3D model visual saliency detection (3D-VSD); 3D model processing; visual attention mechanism; eye-tracking; feature extraction; deep-learning

0 引言

三维模型能立体、真实地反映现实世界,广泛应用于智慧城市、工程制图与应急救援等领域(杨航等,2023)。随着激光点云技术的发展,三维模型数据获取变得越来越容易,不断增加的三维模型数量对模型的快速存储、传输与处理提出极高的要求。相较于图像与视频数据,三维模型体量庞大、数据分布不均匀,使得对其进行处理更具挑战性。如何快速、高效地提取三维模型关键信息,并对信息进行处理、分析与理解,具有重要研究意义。

人类每时每刻都在接受外界环境产生的视觉刺激,然而人脑无法对所有视觉刺激都进行处理,只能选择性处理部分重要视觉信息。这种视觉选择性处理机制能够极大程度提升人类对周围环境的感知能力,也是人类视觉系统的重要特性之一。受启发于这一机制,研究人员将其应用于图像与视频处理领域(孙耀等,2022;Chen等,2020;李贝等,2021),并取得了优秀的成果(王自全等,2022)。近年来,随着三维模型数量的不断增长及三维模型应用的逐渐普及,三维模型视觉显著性检测(3D visual saliency detection, 3D-VSD)逐渐成为计算机视觉领域的重要研究方向。3D-VSD旨在获取三维模型中蕴含重要视觉信息的区域,协助有限计算资源的合理分配,对于提升三维模型处理与分析效率具有重要意义,广泛应用于多种三维模型处理任务(Arvanitis等,2021;Nie等,2020),例如模型简化(Liu等,2023;Nousias等,2023)、模型压缩(Nousias等,2020a)以及模型质量评价等(Abouelaziz等,2020;Nouri等,2022),以获取更符合人眼视觉感受的三维模型处理结果。

根据表示形式的不同(杨航等,2023),三维模型可以分为:点云模型、网格模型和体素模型3类。点云模型是三维空间中一组点的集合,包含点的位置坐标、颜色等信息,是三维模型的原始表示形式;

网格模型由一系列相互连接的三角面组成,具备良好的拓扑结构并能够表示复杂的三维物体;体素模型将三维空间离散化为规则的网格,是2D像素在三维空间中的对应。点云模型与网格模型示意图如图1所示。此外,根据描述对象的不同,三维模型还可以分为小型物体模型与大型城市规模场景模型。人类通常是前者的终端用户群体,而后的终端用户往往是机器。因此,现有3D-VSD研究主要围绕小型物体的点云模型与网格模型开展。

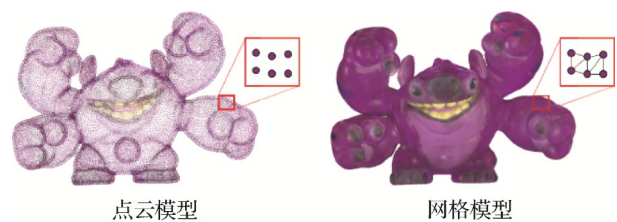


图1 点云模型与网格模型示意

Fig. 1 Examples of 3D point cloud and 3D mesh model

目前,国内外关于3D-VSD研究的综述文献有限,主要包括面向网格模型的3D-VSD方法与应用综述(Liu等,2016b)、面向网格模型与点云模型的3D-VSD与质量评价综述(Lin和Lee,2022)。前者主要总结基于手工特征的网格模型3D-VSD方法并讨论不同显著性检测类型间的区别与联系;后者主要介绍面向不同对象(人类/机器)的3D-VSD方法与质量评价模型的关联。此外,现有文献较少总结基于深度学习的3D-VSD方法和3D-VSD数据集。在此基础上,本文详细梳理2012—2024年发表于国内外权威期刊与会议的3D-VSD方法,以更好地整理与总结3D-VSD领域的研究进展和未来发展趋势。根据使用特征的不同,将现有方法分为基于手工特征的方法和基于深度学习的方法两类,并在每类方法中分别介绍面向点云模型和面向网格模型的方法。同时,考虑到手工特征具有不同尺度,对基于单尺度特征的方法和基于多尺度特征的方法分别介绍。此外,对3D-VSD常用数据集、性能评价指标与相关应用进行梳理,为后续相关研究提供参考。因此,本文

是上述综述文献的有力补充。本文中3D-VSD方法分类如图2所示。

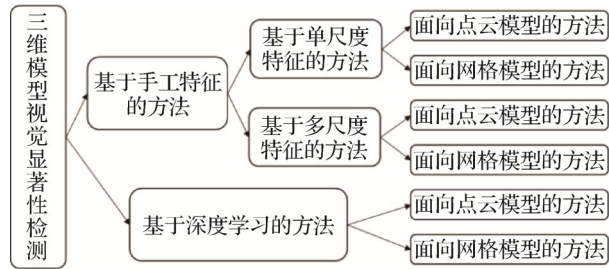


图2 3D-VSD方法分类

Fig. 2 Classification of 3D-VSD methods

1 基于手工特征的3D-VSD方法

基于手工特征的3D-VSD方法主要基于快速点特征直方图(Rusu等, 2009)、Zernike特征(Maximo等, 2011)和高斯曲率特征(Taubin, 1995a)等手工特征检测三维模型视觉显著性区域。根据使用特征尺度的不同,基于手工特征的3D-VSD方法可以分为基于单尺度特征的方法和基于多尺度特征的方法两类。区别于前者仅使用单一尺度特征,后者综合考虑不同尺度特征对视觉显著性检测的影响,小尺度特征有助于发现局部视觉显著性区域,而大尺度特征则能够协助检测全局视觉显著性区域。

1.1 基于单尺度特征的方法

单尺度特征主要在早期3D-VSD方法中应用,不同方法对于单尺度特征的选择也不尽相同。

1.1.1 面向点云模型的方法

点云模型作为3D扫描设备采集得到的原始数据形式,往往包含大量噪声数据,并且数据分布并不均匀,缺乏顶点间的连接关系。因此,对点云模型进行视觉显著性检测相较于网格模型更具挑战性。Tasse等人(2015)提出基于模糊聚类的3D-VSD方法,模拟Achanta等人(2012)提出的图像超像素分割方法,考虑欧氏距离与快速点特征直方图对点云模型进行模糊聚类,根据聚类结果构建簇邻接图与点—簇概率矩阵,随后计算簇级别唯一性特征与空间分布特征,结合点—簇概率矩阵获取点级别视觉显著性检测结果。Wang等人(2015a)提出基于法向量特征的3D-VSD方法,利用K均值聚类将点云模型分为若干个簇,利用最大簇质心计算点云模型的主

法向量,随后计算各点法向量与主法向量间的距离并对距离进行投影获取视觉显著性检测结果。受启发于协方差矩阵在三维模型处理任务中表现出的优异性能(Tabia和Laga, 2015),Guo等人(2018)引入协方差矩阵设计3D-VSD方法,该方法利用法向量特征与高斯曲率特征构建协方差描述符提取点云模型几何特征,通过对协方差矩阵进行映射并结合主成分分析法获得视觉显著性检测结果。

1.1.2 面向网格模型的方法

网格模型大多通过对点云模型进行表面重建获得,能够更加完整地呈现物体表面形状。曲率能够直观、准确地反映网格模型的弯曲程度,是描述其局部形状特征的有力工具,常用的曲率特征包括平均曲率、主曲率与高斯曲率等。Zhao等人(2012)提出基于曲率特征的3D-VSD方法,该方法首先对图像非局部均值滤波进行拓展获取网格模型形状指数,随后计算网格模型高斯曲率特征与平均曲率特征,结合Retinex理论(Elad, 2005)对不同特征进行融合,获取视觉显著性检测结果。Wei等人(2018)基于顶点高斯曲率特征构建曲率共生直方图,随后设计映射函数实现曲率共生直方图到顶点视觉显著性特征值的映射,进而获取视觉显著性检测结果。虽然曲率特征能够较好地描述三维模型局部形状,但是在面对曲率变化复杂的粗糙区域时,检测效果容易受到干扰。

网格模型Zernike特征具有良好的旋转不变性,能较好地描述顶点周围形状特征,在网格模型处理领域具有广泛应用。基于Zernike特征,Tao等人(2015)提出基于流形排序的3D-VSD方法,该方法将网格模型分为若干区域块,将各区域块内顶点的Zernike特征平均值作为区域特征描述符计算局部差异性特征,随后选择局部差异性特征较小的区域块作为查询块,并使用流形排序方法联合Laplacian平滑模块获取视觉显著性检测结果。Liu等人(2016a)利用吸收马尔可夫链对网格模型进行视觉显著性检测,该方法对网格模型进行分割获取各区域块Zernike特征值,利用区域块间特征差异获取种子区域块并构建图模型,随后将种子区域块中的节点作为吸收节点,通过计算转移节点吸收时间获得视觉显著性检测结果。由于Zernike特征计算复杂度较高,并且上述方法往往以区域块(簇)为单位进行视觉显著性检测,因此检测结果容易丢失模型表

面细节信息。

一些方法利用频谱特征、法向量特征等构建几何特征描述符。Hou 和 Qin(2013)提出基于小波扩散的 3D-VSD 方法,通过构建可允许扩散小波检测模型表面视觉显著性区域。Song 等人(2014)通过计算网格模型频谱残差进行视觉显著性检测,该方法首先利用局部均值滤波器对网格模型 Log-Laplacian 频谱进行处理,获取光谱偏差信息,随后利用多尺度转换将光谱偏差信息转换到空间域以获得视觉显著性检测结果。Tao 等人(2016)提出基于信息熵的 3D-VSD 方法,该方法将网格模型顶点法向量特征作为描述符,通过计算一定邻域范围内顶点法向量信息熵获取视觉显著性检测结果。Nouri 等人(2015a)利用顶点法向量特征与切平面特征构建自适应局部区域特征描述符,通过比较不同顶点区域特征的相似性实现 3D-VSD 任务。由于法向量特征难以有效捕捉模型表面细节特征,因此检测精度存在局限性。

基于单尺度特征的 3D-VSD 研究由于使用的特征尺度单一,难以同时检测三维模型表面尺寸相差较大的视觉显著性区域,检测精度提升困难,限制了其在三维模型处理任务中的应用。

1.2 基于多尺度特征的方法

随着多尺度特征在图像与视频视觉显著性检测任务中的优势日益凸显,研究人员逐渐将多尺度特征引入 3D-VSD 任务。

1.2.1 面向点云模型的方法

随着激光点云技术的发展,点云模型数量不断增长,数据处理需求也日益增加。Shtrom 等人(2013)提出面向点云模型的 3D-VSD 方法,首先使用卡方距离量化不同顶点的相似程度获取低层次差异性特征,并在此基础上计算顶点关联性特征值,随后在更大邻域范围内计算高层次差异性特征,通过多通道特征融合得到视觉显著性检测结果。Yun 和 Sim(2016)提出基于超体素分割的 3D-VSD 方法,该方法首先构建不同尺度的超体素簇,考虑几何、颜色特征计算各尺度下基于中心一周边对比度的视觉显著性特征值,通过多尺度特征融合实现视觉显著性检测。Leal 等人(2021)利用顶点法向量特征与位置特征等构建稀疏矩阵,提出基于稀疏表示的 3D-VSD 方法。

由于点云模型数据量庞大,直接进行特征提取

具有一定难度,研究人员参考图像数据处理方法,提出基于 2D 投影图特征的 3D-VSD 方法。Figueiredo 等人(2020)利用正交投影方法获取点云模型在不同角度下的 2D 投影图,并计算视觉显著性特征,随后对 2D 视觉显著性特征进行映射,结合低通滤波获取视觉显著性检测结果。Wang 等人(2023)考虑深度信息设计 3D-VSD 方法,该方法获取不同视角下点云模型 2D 投影图与深度图,利用 UHM (unsupervised hierarchical model)方法(Tavakoli 和 Laaksonen, 2017)计算 2D 投影图视觉显著性特征,并利用深度信息进行优化,通过多视角融合获取视觉显著性检测结果。这些方法通过提取投影图视觉显著性特征实现 3D-VSD 任务,检测精度极大程度依赖于图像 VSD 方法性能。

除上述方法外,部分方法引入随机游走算法协助视觉显著性检测。Ding 等人(2019)提出基于特征自适应融合的 3D-VSD 方法,考虑顶点间距离与相似度计算局部视觉显著性特征值,利用随机游走算法对全局视觉显著性特征值进行细化,随后考虑不同通道特征间内在联系设计自适应多通道特征优化融合框架获取视觉显著性检测结果。Jeong 等人(2020)提出基于随机游走的 3D-VSD 方法,该方法首先利用八叉树将彩色点云模型分为不同尺度的超体素簇,随后利用几何、颜色特征在各尺度上构建全连接图,引入随机游走算法获取各尺度视觉显著性特征值,通过对不同尺度计算结果求平均获取视觉显著性检测结果。基于随机游走的算法在视觉显著性特征扩散过程中能够充分考虑模型表面结构特征,检测结果的细节信息较为完善。

1.2.2 面向网格模型的方法

相较于点云模型,网格模型具有较少的噪声、相对均匀的顶点分布并且蕴含丰富的拓扑信息,能够为多尺度特征提取提供便利。在早期研究中,Lee 等人(2005)提出基于多尺度高斯加权曲率特征融合的 3D-VSD 方法,为后续基于多尺度特征的方法研究奠定了坚实的基础。为更好地捕捉视觉显著性信息,部分方法提取三维模型多层次特征,利用线性框架进行特征融合获取视觉显著性检测结果。受到视觉心理学研究启发,Wu 等人(2013)考虑局部对比度与全局唯一性进行视觉显著性检测,该方法首先对模型进行分割,获取区域块多尺度 Zernike 特征,考虑相邻区域块间距离及其在特征空间中的差异获取

局部特征值,随后利用聚类分析获取全局特征值,通过线性加权融合获取视觉显著性检测结果。Leifman 等人(2016)提出基于多维尺度分析的 3D-VSD 方法,该方法首先计算顶点间差异性,考虑顶点间测地距离获取多尺度差异性特征,随后利用多维尺度分析获取极值顶点,综合考虑差异性特征、关联性特征与极值特征计算视觉显著性检测结果。Song 等人(2018)提出局部与全局特征融合的 3D-VSD 方法,该方法首先利用 Laplacian 方法获取模型多尺度特征,计算局部视觉显著性特征图,随后基于池化方法和全局独特性计算全局视觉显著性特征图,通过线性加权融合获取视觉显著性检测结果。Arvanitis 等人(2021)提出面向工业建模应用的 3D-VSD 方法,对网格模型进行分割,分别获取几何视觉显著性特征与光谱视觉显著性特征,利用线性框架融合多通道特征得到视觉显著性检测结果。上述方法利用线性框架融合多层次特征以实现视觉显著性检测,在特征融合过程中难以确定不同特征的权重,检测精度存在局限性。

考虑到线性融合框架的局限性,部分研究基于非线性框架设计多层次特征融合方法。Song 等人(2012)提出基于条件随机场的 3D-VSD 方法,通过计算不同标准差下高斯滤波值间差异获取多尺度特征,随后将多尺度特征引入条件随机场,利用一致性约束提高视觉显著性检测结果的鲁棒性。Luan 等人(2020)提出基于特征动态融合的 3D-VSD 方法,该方法利用基于泰森多边形的特征描述子获取局部视觉显著性特征,通过分层聚类计算全局视觉显著性特征,随后考虑邻域内顶点特征获取视觉显著性检测结果。Nouri 等人(2024)提出基于球形邻域的 3D-VSD 方法,利用顶点间最大距离构建不同尺度球形邻域,并计算各尺度下邻域内顶点在 CIELab 色彩空间中的差异,随后考虑不同尺度特征间关联进行非线性融合,获取视觉显著性检测结果。

为高效捕捉三维模型不同特征的统计特性,部分研究基于信息熵构建 3D-VSD 方法。Nouri 等人(2015b)利用顶点法向量与切平面特征构建自适应区域块获取顶点局部特征,考虑区域块间相似性构建不同尺度视觉显著性特征图,并将特征图熵值作为权重指导视觉显著性特征融合。通过引入顶点颜色特征,该方法还可以被拓展应用于彩色网格模型视觉显著性检测任务(Nouri 等,2015c)。汪小东等

人(2017)提出基于局部区块加权曲率熵的 3D-VSD 方法,通过邻域累计体积定义自适应区块,考虑邻域内顶点加权曲率熵获取视觉显著性检测结果。Wang 等人(2022)考虑网格模型材料等信息进行视觉显著性检测,该方法首先通过色差图融合获取纹理视觉显著性特征,利用局部曲率熵获取几何视觉显著性特征,随后通过特征融合获取视觉显著性检测结果。dos Anjos 等人(2023)提出基于熵特征的 3D-VSD 方法,通过计算相邻顶点切平面法向量获取曲率角特征,随后利用曲率角熵特征计算视觉显著性检测结果。由于熵特征的计算极易受到噪声信息的干扰,因此上述方法检测得到的视觉显著性区域往往较为分散,检测精度难以进一步提升。

为进一步提高视觉显著性检测精度,研究人员考虑人类视觉认知的稀疏特性,开始研究基于稀疏表示的 3D-VSD 方法。Wang 等人(2015b)提出基于低秩分解和稀疏分析的 3D-VSD 方法,该方法结合高频信息和低频信息构建形状特征描述符捕捉网格模型结构特征,随后通过稀疏分析获取视觉显著性检测结果。Hu 等人(2020a)提出基于稀疏表示的 3D-VSD 方法,该方法对网格模型进行随机采样获取顶点形状特征,同时考虑稀疏原则与稀疏原则建立优化模型,稀疏原则帮助获取与其他区域相差较大的视觉显著性区域,而稀疏原则确保视觉显著性区域相对紧凑,随后引入高斯滤波获取视觉显著性检测结果。基于稀疏表示的方法能够充分获取模型表面结构特征,具有较好的 3D-VSD 精度。

此外,Zhao 等人(2016)提出基于区域块的 3D-VSD 方法,通过结合相对法线距离特征与基于 Retinex 理论的特征获得视觉显著性检测结果。Ding 和 Chen(2022)提出面向 6 自由度环境的 3D-VSD 方法,利用视觉中心偏向性特征对多尺度视觉显著性特征进行优化,进而实现特定 6 自由度条件下 3D-VSD 任务。随着机器学习方法的普及,Alfarasani 等人(2022)利用多视角 2D 投影图眼动数据构建训练数据集,以训练 3D-VSD 方法,考虑到数据集规模,该方法获取各顶点 SHOT(signature of histograms of orientations)特征(Tombari 等,2010)、HKS(heat kernel signature)特征(Sun 等,2009)等一系列几何特征,结合现有 3D-VSD 方法计算视觉显著性特征,随后利用机器学习算法进行训练,获取视觉显著性检测结果。虽然基于机器学习的 3D-VSD 方法能够充

分学习模型表面视觉显著性特征,但是检测精度易受不同机器学习方法影响。

基于多尺度特征的3D-VSD方法虽然能够较好地捕捉不同尺寸视觉显著性特征,但在多通道/尺度特征融合时大多采用线性加权融合方式,融合权重选择高度依赖于研究人员经验,并且在一组模型上融合效果良好的权重组合对于另一组模型并不一定适用,使得3D-VSD方法性能仍存在较大提升空间。如何充分利用多通道/尺度特征,值得深入挖掘。

1.3 小结

面对三维模型庞大的数据规模,部分方法对点云/网格模型进行聚类或分割,将其分为若干个区域块(簇),利用区域块特征获取视觉显著性检测结果。相较于基于点特征的方法,这些方法获取区域块特征进行视觉显著性检测能够更好地捕捉全局视觉显著性特征并降低计算复杂程度。然而,这些方法在进行聚类或分割时主要考虑顶点间欧氏距离与几何特征,可能将具有相同属性的顶点分配到不同区域块中,造成模型语义信息的损伤,在处理蕴含丰富语义信息的模型(例如人脸模型等)时视觉显著性检测效果往往不佳。

受限制于手工特征表达能力,基于手工特征的方法在处理复杂三维模型时视觉显著性检测效果与真值间仍存在较大差距,如何设计全新、高效的手工特征描述符值得进一步研究。此外,由于早期三维模型往往缺乏纹理信息,基于手工特征的方法大多利用模型几何特征进行视觉显著性检测。随着彩色三维模型数量的不断增长以及三维模型应用的逐渐普及,如何有效结合纹理特征实现高精度的视觉显著性检测值得深入探索。

表1对本节涉及的基于手工特征的3D-VSD方法从模型类型、特征类型及特点等方面进行总结。

2 基于深度学习的3D-VSD方法

深度学习技术的成熟极大程度推动了计算机视觉领域的发展,基于深度学习的三维模型分类、分割方法(Qi等,2017a,b)精度相较于传统方法也得到极大的提升,促使研究人员逐渐聚焦于基于深度学习的3D-VSD研究。相较于手工特征,基于深度学习技术获取的特征具有更强的表达能力,能够深入捕捉三维模型高层次语义信息,对于提升3D-VSD精

度具有重要意义。早期基于深度学习的3D-VSD研究受限于深度学习技术的发展以及大规模训练数据集的缺失,主要基于2D投影图特征进行视觉显著性检测,检测精度高度依赖于图像VSD方法性能;为降低对大规模训练数据集的依赖,部分方法基于一个假设:即相同类别的三维模型具有相似的视觉显著性分布特点,引入模型类别标签设计弱监督3D-VSD方法;随着深度学习技术的不断成熟,基于无监督学习的方法与基于端到端的方法相继提出,这些方法能够更高效地整合三维模型多级特征,使基于深度学习的3D-VSD方法性能得到进一步提升。

2.1 面向点云模型的方法

点云模型能够真实反映现实场景情况,具有较高的应用价值,如何提升其视觉显著性检测精度值得深入研究。受启发于图像视觉显著性检测方法的优异性能,Tliba等人(2024)提出基于投影变换的3D-VSD方法,通过正交投影获取点云模型2D投影图,利用SATSaI(Tliba等,2022)以及DeepGaze IIE(Linardos等,2021)等基于深度学习的图像视觉显著性检测方法提取2D投影图视觉显著性特征,通过特征映射获取视觉显著性检测结果。Ding等(2022)提出基于知识迁移的3D-VSD方法,该方法利用彩色图像视觉显著性分布知识构建点云模型伪视觉显著性数据集,从伪视觉显著性标签中学习点云模型视觉显著性分布,进而获取视觉显著性检测结果,随后该方法被拓展应用于点云分割任务中(Ding等,2024a),取得了优异的性能。这些方法主要利用2D投影图视觉显著性特征实现3D-VSD任务,缺乏对于三维模型表面几何特征的利用,检测性能提升存在局限性。

除上述方法外,针对点云模型的不同应用场景,研究人员提出多个3D-VSD方法。考虑到视觉显著顶点的缺失将对点云识别任务产生较大影响,Zheng等人(2019)提出基于点删除思想的3D-VSD方法,考虑到点删除的不可微性,该方法将顶点向点云质心进行移动以近似点删除操作,并利用球坐标下顶点梯度损失获取视觉显著性检测结果。考虑到点云模型前景区域与背景区域的不均匀性将对三维场景理解带来不利影响,An等人(2024)利用标签扩散缓解前景区域标签的缺乏问题,获取视觉显著性检测结果,以协助识别点云模型前景区域。

表 1 基于手工特征的 3D-VSD 方法总结

Table 1 Summary of hand-crafted feature based 3D-VSD methods

类别	方法	模型类型	特征类型	颜色	特点
基于特征扩散	Tasse 等人(2015)	点云模型	簇特征	无	检测精度易受聚类方法影响
基于法向量特征	Wang 等人(2015a)	点云/网格模型	点/簇特征	无	检测结果多集中于曲率变化区域
	Guo 等人(2018)				
	Nouri 等人(2015a)				
	Zhao 等人(2016)				
基于多层次特征	Shtrom 等人(2013)	点云/网格模型	点/簇特征	无/有	平坦区域检测精度较差
	Leifman 等人(2016)				
	Song 等人(2012)				
	Song 等人(2018)				
	Luan 等人(2020)				
	Nouri 等人(2024)				
基于投影图特征	Yun 和 Sim(2016)	点云模型	图特征	有	检测精度易受图像 VSD 方法影响
	Figueiredo 等人(2020)				
基于随机游走	Wang 等人(2023)	点云模型	点/簇特征	无/有	检测结果细节信息较为完善
	Ding 等人(2019)				
基于曲率特征	Jeong 等人(2020)	网格模型	点特征	无	粗糙区域检测精度较低
	Zhao 等人(2012)				
	Wei 等人(2018)				
基于 Zerinike 特征	Lee 等人(2005)	网格模型	簇特征	无	检测结果缺失细节信息
	Tao 等人(2015)				
	Liu 等人(2016a)				
基于频谱特征	Wu 等人(2013)	网格模型	点/簇特征	无	检测结果覆盖范围较大
	Hou 和 Qin(2013)				
	Song 等人(2014)				
基于熵特征	Arvanitis 等人(2021)	网格模型	点特征	无/有	检测结果较为分散
	Tao 等人(2016)				
	汪小东等人(2017)				
	dos Anjos 等人(2023)				
基于稀疏表示	Wang 等人(2022)	点云/网格模型	点特征	无	模型简单时检测性能较好
	Nouri 等人(2015b)				
	Wang 等人(2015b)				
基于机器学习	Hu 等人(2020a)	网格模型	点特征	无	检测精度易受机器学习方法影响
基于特征优化	Leal 等人(2021)				
	Ding 和 Chen(2022)	网格模型	点特征	无	仅能检测特定视角下的区域

此外,为克服大规模训练数据集缺失的问题,Jiang等人(2023)提出基于弱监督分类的3D-VSD方法,在训练过程中仅需使用点云类别标签信息。该方法从分类网络中获取视觉显著性特征,并引入多尺度点簇相似性矩阵进行视觉显著性特征扩散,得到点云模型视觉显著性检测结果。

2.2 面向网格模型的方法

网格模型具有规则的网格结构,能够便捷地进行数据处理与分析,具有广阔的应用前景。与基于深度学习的点云模型视觉显著性检测方法相类似,部分研究通过对网格模型进行投影变换获取2D投影图协助视觉显著性检测。Abid等人(2020)提出基于多视点融合的3D-VSD方法,从多个视点生成网格模型2D投影图,利用SALICON(saliency in context)(Jiang等,2015)图像视觉显著性检测方法提取2D投影图视觉显著性特征,考虑不同视点下可见区域的相关性获取视觉显著性检测结果。Ding等人(2024b)提出基于几何特征优化的3D-VSD方法,该方法首先利用图像视觉显著性检测方法获取2D投影图视觉显著性特征,通过投影变换获取初步视觉显著性检测结果,随后引入网格模型几何特征进行优化,实现视觉显著性检测。

由于缺乏大规模训练数据集,部分研究引入弱监督学习构建3D-VSD方法。Song等人(2021a)提出基于弱监督学习的3D-VSD方法,通过计算不同2D投影图间的差异获取视觉显著性特征,随后将视觉显著性特征引入弱监督分类网络,通过多视角特征融合获取视觉显著性检测结果。Song等人(2021b)还引入类别标签构建3D-VSD方法,该方法认为图像视觉显著性与网格模型视觉显著性具有一定相似性,并且相同类别的网格模型具有相似的视觉显著性分布特点。基于这两点认知,该方法首先获取网格模型2D渲染图,将带有类别标签的2D渲染图与带有注视点数据的自然图像作为网络输入,将图像视觉显著性特征迁移到网格模型中,进而获得网格模型视觉显著性检测结果。基于弱监督的3D-VSD方法取得了较为优异的检测性能,但是在训练过程中需要额外使用模型类别标签信息,限制了相关方法的应用场景。为进一步降低对训练数据的依赖,研究人员提出基于无监督学习的3D-VSD方法。Song等人(2022)考虑物体一视图一致性设计3D-VSD方法,该方法包含两条通道,前者获取2D投影

图与网格模型中的特征,后者充分利用视图内部与不同视图间的知识,通过推动两条通道生成一致的网格模型分类结果获取视觉显著性检测结果。

除上述方法外,部分方法构建训练数据集以支持模型的训练。Nousias等人(2020a)提出基于卷积神经网络的3D-VSD方法,该方法首先利用现有3D-VSD方法构建训练数据集,通过特征值融合获取局部与全局光谱视觉显著性特征,随后利用鲁棒主成分分析计算几何视觉显著性特征,结合卷积神经网络进行视觉显著性检测。Nousias等人(2023)对这一方法进行改进,计算效率进一步提升,但视觉显著性检测结果主要集中于模型的边缘区域。考虑PointNet++网络(Qi等,2017b)在模型分割中的优异性能,Martin等人(2024)在此基础上构建3D-VSD方法,首先开展眼动实验构建训练数据集,随后通过特征提取模块获取网格模型潜在特征,并对特征进行传播,结合PointNet++网络中的分割特征获取视觉显著性检测结果,该算法在处理相对简单的三维模型时具有较为优秀的检测性能。

此外,Chen等人(2023)提出端到端的3D-VSD方法,该方法采用多级特征连接弥补池化过程中的细节丢失,引入残差学习解决梯度消失问题,设计多尺度融合模块优化网络性能,实现视觉显著性检测。Hu等人(2020b)结合形状匹配任务设计3D-VSD方法,首先构建多层循环神经网络(recurrent neural network,RNN)结构,以有效整合不同顶点多尺度形状特征,随后设计软阈值算子以增强视觉显著性特征的稀疏性,进而获取视觉显著性检测结果。Rouhafzay等人(2023)也提出基于梯度类激活映射的3D-VSD方法,取得了优秀的视觉显著性检测结果。这些基于多层次特征的方法能够充分挖掘不同特征间的内在关联,检测结果细节信息较为完善。

2.3 小结

受限于开展三维模型眼动实验所需的高昂成本,大规模三维模型眼动数据集仍较为匮乏。为弥补大规模训练数据集的缺失,部分方法从不同视点获取三维模型2D投影图,利用大规模图像眼动数据集训练视觉显著性特征提取模型,并通过3D-2D映射结合多视角特征融合获取视觉显著性检测结果。然而,这类方法在投影过程中将不可避免地丢失深度信息,缺乏对于模型几何特征的利用并且更大程

度上依赖于模型的纹理特征,使得视觉显著性检测性能难以进一步提升。

表 2 从模型类型、训练集及特点等方面对本节涉及的基于深度学习的 3D-VSD 方法进行总结。

3 常用数据集与评价指标

3.1 3D-VSD 数据集

从 3D-VSD 问题提出至今,研究人员构建了一系列 3D-VSD 数据集。本节将介绍常用 3D-VSD 数据集,为研究人员开展相关研究奠定基础。

3.1.1 基于鼠标追踪的数据集

受限制于眼动追踪技术的发展,在早期视觉感知研究中,研究人员主要通过鼠标追踪技术获取志愿者浏览视觉刺激时的视觉行为并进行分析。相较于眼动追踪技术,鼠标追踪技术数据获取成本更低,广泛应用于构建大规模图像视觉显著性检测数据集(Tavakoli 等, 2017)。例如 SALICON 数据集(Jiang 等, 2015),该数据集包含 10 000 幅来自 MS COCO (Microsoft common objects in context)数据集(Lin 等, 2014)的图像,通过亚马逊众包平台 AMT(Amazon Mechanical Turk)采集鼠标追踪数据,并广泛应用于

基于深度学习的图像视觉显著性检测研究中。

基于鼠标追踪技术,研究人员邀请多名志愿者用鼠标点击三维模型表面感兴趣区域,并利用采集的鼠标追踪数据构建三维模型感兴趣点检测数据集。

1)美国普林斯顿大学于 2012 年建立的感兴趣点检测数据集(Chen 等, 2012)包含 19 种不同类型的网格模型(共 380 个),这些网格模型来自于 SHREC2007 形状检索比赛。在实验过程中,研究人员将“谢林点”拓展到网格模型中,邀请志愿者利用鼠标在网格模型表面标记他们认为其他志愿者可能选择的顶点。每个三维模型采集得到了约 25 名志愿者的鼠标标记数据。

2)土耳其埃斯基谢希尔大学于 2012 年建立的感兴趣点检测数据集(Dutagaci 等, 2012)包含两个子数据集。子数据集 A 包含 24 个网格模型,子数据集 B 包含 43 个网格模型。在实验过程中,志愿者可以自由旋转网格模型并用鼠标标记模型表面感兴趣的点。子数据集 A 采集 23 名志愿者的鼠标标记数据,子数据集 B 采集 16 名志愿者的鼠标标记数据。该数据集提供 6 种方法计算的感兴趣点检测结果,并使用两种性能评价指标,以协助感兴趣点检测方

表 2 基于深度学习的 3D-VSD 方法总结
Table 2 Summary of deep-learning based 3D-VSD methods

类别	方法	模型类型	训练集	颜色	特点
基于投影图特征	Tliba 等人(2024)	点云/网格模型	图像 VSD 数据集	有	性能依赖于图像 VSD 方法
	Ding 等人(2022)				
	Ding 等人(2024b)				
	Abid 等人(2020)				
基于点删除	Zheng 等人(2019)	点云模型	ModelNet 数据集	无	为点云识别任务设计
基于弱监督学习	Jiang 等人(2023)	点云/网格模型	ModelNet 数据集	无	需要模型类别标签信息
	Song 等人(2021a)				
	Song 等人(2021b)				
基于无监督学习	Song 等人(2022)	网格模型	自建数据集	无	可实现 3D 场景 VSD 任务
基于主成分分析	Nousias 等人(2020a)	网格模型	自建数据集	无	检测结果多集中于边缘区域
	Nousias 等人(2023)				
基于特征转换	Martin 等人(2024)	网格模型	自建数据集	无	模型简单时检测性能较好
基于多层次特征	Chen 等人(2023)	网格模型	Schelling 数据集	无	检测结果细节信息较为完善
	Hu 等人(2020b)				

法性能比较与评价。

上述数据集在早期 3D-VSD 任务中得到广泛使用,以量化评价不同视觉显著性检测方法性能。

3.1.2 基于眼动追踪的数据集

随着眼动追踪技术的发展日益成熟,研究人员开始利用眼动追踪技术采集多名志愿者浏览三维模型时的眼动数据,以分析志愿者的视觉行为特性,进而构建三维模型视觉显著性检测数据集。

1)爱尔兰圣三一大学于 2005 年建立的数据集(Howlett 等,2005)包含两个子数据集。子数据集 A 包括 37 个网格模型,子数据集 B 包括 30 个网格模型,眼动实验共采集 20 名志愿者数据。实验过程中,志愿者从不同的起始角度浏览屏幕中的网格模型,并可以旋转模型改变观测视角,每个模型将被展示 30 s。眼动追踪设备为 EyeLink 眼动仪,数据采集频率为 250 Hz。

2)美国马里兰大学于 2010 年建立的数据集(Kim 等,2010)包含 5 个网格模型,眼动实验共采集 18 名志愿者数据。实验过程中,志愿者浏览屏幕中的网格模型,每个模型将被展示 5 s。眼动追踪设备为 ISCAN ETL-500 眼动仪,数据采集频率为 60 Hz。

3)德国柏林工业大学于 2016 年建立的数据集(Wang 等,2016)包含 15 个网格模型,每个模型均由 3D 打印机打印。眼动实验共采集 30 名志愿者数据。实验过程中,每个模型提供 3 个观测视角,每个视角下模型将被展示 5 s。数据采集频率为 120 Hz。

4)德国柏林工业大学于 2018 年建立的数据集(Wang 等,2018)包含 16 个网格模型,每个模型使用两种不同材质打印,眼动实验共采集 70 名志愿者数据。实验过程中,志愿者从 7 个不同视角浏览模型,每个视角间隔 15°,每个模型将被展示 5 s。眼动追踪设备为 EyeLink 1000 眼动仪。

5)法国里昂大学于 2018 年建立的数据集(Lavoué 等,2018)包含 32 个网格模型,眼动实验共采集 20 名志愿者数据。实验过程中,模型材质、光照条件与观测视角都被纳入考虑以生成 2D 渲染图与渲染视频,每幅渲染图在屏幕中将被展示 7 s。眼动追踪设备为 Tobii TX-120 眼动仪,数据采集频率为 120 Hz。

6)瑞士洛桑联邦理工学院于 2019 年建立的数据集(Alexiou 等,2019)包含 12 个彩色网格模型。眼动实验共邀请 21 名志愿者参与。实验过程中,模型

分成两组进行实验,组间时间间隔为 5 min,模型放置在 HTC Vive 构建的虚拟房间中,志愿者可以自由移动以改变浏览视角。眼动追踪设备为 Pupil Labs 硬件设备,数据采集频率为 120 Hz。

7)法国南特大学于 2020 年建立的数据集(Abid 等,2020)包含 22 个彩色网格模型,眼动实验共采集 30 名志愿者数据。实验过程中,每个模型都从 4 个固定视角生成 2D 渲染图,每幅渲染图将在屏幕中展示 3 s。眼动追踪设备为 EyeLink 100 plus 眼动仪。

8)英国诺森比亚大学于 2020 年建立的数据集(Hu 等,2020a)包含 50 个网格模型,眼动实验共采集 8 名志愿者数据。实验过程中,每个模型从 12 个不同视点生成 3 s 渲染视频,视频总时长为 48 s,不同视点间切换时长为 1 s。眼动追踪设备为 SMI RED250 眼动仪,数据采集频率为 250 Hz。

9)武汉大学于 2022 年建立的数据集(Ding 和 Chen,2022)包含 32 个网格模型。眼动实验共采集 21 名志愿者数据。实验过程中,三维模型被放置在 HTC Vive 构建的虚拟房间中,志愿者可以自由移动以改变模型浏览视角,每个模型将被展示 10 s。眼动追踪设备为 aGlass 眼动模块,数据采集频率为 120 Hz。

10)西班牙萨拉戈萨大学于 2023 年建立的数据集(Martin 等,2024)包含 58 个网格模型。眼动实验共采集 32 名志愿者数据。实验过程中,每个模型被放置在 VR(virtual reality)环境中并被 360° 旋转展示 22 s。眼动追踪设备为 Pupil Labs 眼动模块。

11)中南财经政法大学于 2024 年建立的数据集(Ding 等,2024b)包含 17 个彩色网格模型,眼动实验共采集 34 名志愿者数据。实验过程中,志愿者从 12 个不同视角浏览模型,每个视角间隔 30°,每个模型被展示 10 s。眼动追踪设备为 HTC Vive Pro Eye,数据采集频率为 120 Hz。

网格模型眼动数据采集结果示意图如图 3 所示。图中第 1 行展示网格模型,第 2 行展示眼动实验采集得到的眼动注视点构建的注视点密度图,图中红色表示注视点密度值较高,蓝色表示注视点密度值较低。4 个展示的三维模型分别来自 4 个不同的 3D-VSD 数据集。

表 3 对 3D-VSD 常用数据集进行总结。在上述数据集中,Chen 等人(2012)提出的 Schelling 数据集由于提出时间最早,包含的模型数量较多并且类别丰富,广泛应用于多个 3D-VSD 方法以协助性能验



图3 网格模型注视点密度图

Fig. 3 Examples of 3D mesh fixation point density maps
(a) 3D mesh models; (b) fixation density maps

证与对比。Dutagaci 等人(2012)提出的感兴趣点检测数据集与 Lavoué 等人(2018)提出的 3DVA 数据集也在部分 3D-VSD 方法中得到应用,以评估对应方法的检测性能。

3.2 3D-VSD 评价指标

对 3D-VSD 方法性能进行评估需要将其计算得到的视觉显著性检测结果与眼动实验采集的视觉显著性真值图进行对比,并使用量化指标进行性能评价。不同的量化指标将从不同角度衡量 3D-VSD 方法性能,通过使用多个评价指标,能够获得更加全面的性能评价结果。

为方便对各评价指标进行详细介绍,令 $\mathbf{M}$ 表示

三维模型, $\mathbf{S}$ 表示视觉显著性检测结果, $\mathbf{Q}$ 表示眼动注视点数据, $\mathbf{G}$ 表示 $\mathbf{Q}$ 经过高斯函数处理后得到的视觉显著性真值图(注视点密度图)。

1) SE 值。Song 等人(2018)提出视觉显著性误差 SE 值以衡量 3D-VSD 结果 $\mathbf{S}$ 与真值图 $\mathbf{G}$ 间的差异。该评价指标首先计算三维模型表面任一顶点 i 的视觉显著性检测结果 S_i 与真值 G_i 间的差值,通过对所有顶点处差值求和得到 SE 值评价 3D-VSD 方法性能。具体为

$$SE(\mathbf{M}) = \frac{1}{N_t} \sum_{i=1}^{N_t} |S_i - G_i| \quad (1)$$

式中, N_t 表示三维模型顶点总数, S_i 和 G_i 分别表示 $\mathbf{M}$ 中第 i 个顶点的视觉显著性检测结果与真值。

SE 值越小,表示视觉显著性检测结果 $\mathbf{S}$ 与真值 $\mathbf{G}$ 差异越小,即该方法性能越好。

2) SIM 值。相似性测度 SIM 值在图像视觉显著性检测方法性能评价中得到广泛使用,也被拓展应用于 3D-VSD 方法性能评价中。该评价指标通过衡量视觉显著性检测结果 $\mathbf{S}$ 与真值 $\mathbf{G}$ 间的相似程度评价 3D-VSD 方法性能。具体为

$$\begin{aligned} SIM(\mathbf{M}) &= \sum_{i=1}^{N_t} \min(S_i, G_i) \\ \text{s.t. } \sum_{i=1}^{N_t} S_i &= \sum_{i=1}^{N_t} G_i = 1 \end{aligned} \quad (2)$$

表3 3D-VSD 常用数据集总结

Table 3 Summary of commonly used 3D-VSD databases

数据集	模型数量	刺激类别	展示时长/s	人数	观测视角	数据类型	颜色纹理	特点
Chen 等人(2012)	380	渲染图	-	25	自由视角	鼠标数据	单一颜色	模型数量较多,类别丰富
Dutagaci 等人(2012)	24+43	渲染图	-	23+16	自由视角	鼠标数据	单一颜色	提供感兴趣点检测真值
Howlett 等人(2005)	37+30	渲染图	30	20	自由视角	眼动数据	单一颜色	模型质量较差
Kim 等人(2010)	5	渲染图	5	18	固定视角	眼动数据	单一颜色	观测视角受限
Wang 等人(2016)	15	3D 打印模型	5	30	3个视角	眼动数据	单一颜色	观测视角较少
Wang 等人(2018)	16	3D 打印模型	5	70	7个视角	眼动数据	单一颜色	模型包含不同材质信息
Lavoué 等人(2018)	32	渲染图/视频	7	20	固定视角	眼动数据	单一颜色	提供静态/动态观测
Alexiou 等人(2019)	12	VR 三维模型	无限制	21	自由视角	眼/头动数据	彩色纹理	不限制观测时长
Abid 等人(2020)	22	渲染图	3	30	4个视角	眼动数据	彩色纹理	模型质量较高,纹理丰富
Hu 等人(2020a)	50	渲染视频	3	8	12个视角	眼动数据	单一颜色	志愿者人数较少
Ding 和 Chen(2022)	32	VR 三维模型	10	21	自由视角	眼/头动数据	单一颜色	允许自由移动
Martin 等人(2024)	58	VR 三维模型	22	32	360°旋转	眼动数据	单一颜色	模型缺乏纹理信息
Ding 等人(2024b)	17	VR 三维模型	10	34	12个视角	眼动数据	彩色纹理	观测视角较多

注:“-”表示数据未知。

SIM 值越大,表示视觉显著性检测结果 S 与真值图 G 越相似,即该方法性能越好。

3)CC 值。Ding 等人(2019)将相关系数 CC 值应用于 3D-VSD 方法性能评价中。该评价指标通过衡量视觉显著性检测结果 S 与真值图 G 间的相关程度评价 3D-VSD 方法性能。具体为

$$CC(\boldsymbol{M}) = \frac{\sigma(S, G)}{\sigma(S) \times \sigma(G)} \tag{3}$$

式中, $\sigma(S, G)$ 表示 S 和 G 的协方差。

CC 值越大,表示视觉显著性检测结果 S 与真值图 G 相关性越强,即该方法性能越好。

4)KL 值。Abid 等人(2020)将 KL (Kullback-Leibler) 散度值应用于 3D-VSD 方法性能评价中。该评价指标通过衡量视觉显著性检测结果 S 与真值 G 间的分布差异评价 3D-VSD 方法性能。具体为

$$KL(\boldsymbol{M}) = \sum_{i=1}^{N_i} G_i \log\left(\frac{G_i}{S_i}\right) \tag{4}$$

KL 值越小,表示视觉显著性检测结果 S 与真值图 G 越接近,即该方法性能越好。

5)NSS 值。NSS 评价指标首先对 3D-VSD 结果 S 进行归一化处理,随后获取眼动注视点图 Q ,通过计算眼动注视点处视觉显著性值的平均值评价 3D-VSD 方法性能。具体为

$$NSS(\boldsymbol{M}) = \frac{1}{N} \sum_{i=1}^{N_i} (\bar{S}_i \times Q_i) \\ \text{s.t. } \bar{S} = \frac{S - \mu(S)}{\sigma(S)} \tag{5}$$

式中, N 表示眼动注视点数量, $\mu(S)$ 表示均值, $\sigma(S)$ 表示标准差。

NSS 值越大,表示视觉显著性检测结果 S 与真值图 G 越接近,即该方法性能越好。

6)AUC 值。AUC 值是图像视觉显著性检测任务中常用性能评价指标,也可以应用于评价 3D-VSD 方法性能。计算 AUC 值首先需要利用不同阈值对视觉显著性检测结果 S 进行划分,分别计算假阳性率与真阳性率以绘制受试者工作特征曲线,通过计算该曲线线下面积获得 AUC 值,以评价 3D-VSD 方法性能。AUC 值越大,表示视觉显著性检测结果 S 与真值图 G 越接近,即该方法性能越好。

3.3 3D-VSD 方法对比

为客观评价不同 3D-VSD 方法性能,本文在 Chen 等人(2012)提出的 Schelling 数据集上采用 CC 值与 AUC 值作为评价指标,对不同方法进行对比。

受限于代码的可获得性,本文选取 11 种 3D-VSD 方法进行对比,包括基于手工特征的方法以及基于深度学习的方法,对比实验结果如表 4 所示。其中, σ 表示真值构建过程中使用的高斯滤波标准差, B 表示三维模型边界框对角线的长度。

表 4 Schelling 数据集上 3D-VSD 方法性能对比
Table 4 Experimental comparison of different 3D-VSD methods on Schelling database

方法	方法类别	$\sigma = 0.10 B$		$\sigma = 0.12 B$		$\sigma = 0.14 B$		$\sigma = 0.16 B$		$\sigma = 0.18 B$		$\sigma = 0.20 B$	
		CC	AUC	CC	AUC	CC	AUC	CC	AUC	CC	AUC	CC	AUC
Hou 和 Qin(2013)	单尺度手工特征	0.101	0.777	0.091	0.757	0.082	0.762	0.074	0.755	0.068	0.745	0.063	0.738
Song 等人(2014)	单尺度手工特征	0.324	0.811	0.323	0.814	0.313	0.814	0.301	0.810	0.293	0.804	0.284	0.805
Tao 等人(2015)	单尺度手工特征	0.301	0.788	0.294	0.791	0.287	0.790	0.280	0.786	0.273	0.783	0.267	0.785
Lee 等人(2005)	多尺度手工特征	0.223	0.700	0.213	0.701	0.202	0.707	0.193	0.705	0.186	0.700	0.179	0.696
Arvanitis 等人(2021)	多尺度手工特征	0.336	-	0.323	-	0.309	-	0.295	-	0.286	-	0.277	-
Leifman 等人(2016)	多尺度手工特征	0.437	0.855	0.421	0.861	0.402	0.863	0.376	0.870	0.360	0.861	0.340	0.858
Song 等人(2018)	多尺度手工特征	0.407	0.856	0.390	0.859	0.372	0.871	0.354	0.867	0.340	0.861	0.324	0.859
Nousias 等人(2020a)	深度特征	0.127	-	0.129	-	0.127	-	0.124	-	0.124	-	0.123	-
Song 等人(2021a)	深度特征	0.455	0.892	0.457	0.897	0.454	0.899	0.447	0.900	0.439	0.903	0.427	0.898
Liu 等人(2023)	深度特征	0.491	0.912	0.460	0.916	0.431	0.919	0.404	0.921	0.381	0.923	0.361	0.925
Song 等人(2023)	深度特征	0.476	-	0.498	-	0.510	-	0.515	-	0.516	-	0.510	-

注:加粗字体表示各列最优结果,“-”表示未测数据。

由表4可知,基于深度学习的3D-VSD方法检测精度相较于基于手工特征的方法往往更高,主要因为深度特征往往具有更强的特征表达能力,能够更好地捕捉模型表面视觉显著性区域。此外,基于多尺度手工特征的3D-VSD方法检测精度往往也高于基于单尺度特征的方法,主要因为基于多尺度特征的方法能够更加全面地捕捉不同尺寸的特征,进而提升检测精度。

在基于深度学习的方法中,Nousias等人(2020a)的方法检测性能较差,可能因为该方法使用的训练数据集中真值图通过融合光谱特征与几何特征获得,与Schelling数据集中的标记数据存在较大差异;Song等人(2023)的方法在特征提取与映射过程中综合考虑三维模型2D特征与3D特征,检测精度具有明显提升。在基于多尺度手工特征的方法中,Lee等人(2005)的方法检测性能较差,主要原因是该方法使用的高斯加权曲率特征表达能力存在局限性,在模型表面粗糙区域检测精度较低;Leifman等人(2016)与Song等人(2018)的方法充分考虑不同层次特征的关联,获得了较高的检测精度。在基于单尺度手工特征的方法中,Hou和Qin(2013)的方法在CC评价指标下得分较差,可能因为该方法为更多区域赋予了较高的视觉显著性值并且检测得到的区域较为分散;Song等人(2014)的方法具有较好的检测精度,主要由于该方法在特征映射过程中较好地保留了模型的细节特征。

4 3D-VSD应用

3D-VSD方法能够获取模型表面吸引人眼视觉注意的区域,在与视觉感知相关的三维模型处理任务中有着广泛的应用,包括三维模型缩放、简化、降噪以及质量评价等。

4.1 三维模型缩放

三维模型缩放指通过调整模型尺寸使其能够适应不同应用场景的需求,在工业设计与游戏建模等领域具有广泛应用,能够极大程度提升作业效率。

传统缩放方法往往使用均匀缩放方式或根据模型表面特征进行缩放(Kraevoy等,2008;Nozawa等,2017),虽然能够有效调整模型尺寸,但却忽略了人眼视觉感受在模型缩放任务中的重要意义,缩放后模型的视觉效果往往改变较大。考虑到人眼对模型

表面不同区域的视觉感受并不相同,引入3D-VSD方法能够更好地在模型缩放过程中降低视觉显著性区域形变程度(Miao和Lin,2013),尽可能保证缩放后模型的视觉效果。例如,Jia等人(2014)提出基于多尺度视觉显著性检测的模型缩放方法,通过融合多尺度特征获取视觉显著性检测结果,进而指导模型的缩放处理;朱建等人(2014)提出视觉显著性特征保持的模型缩放方法,基于视觉显著性检测结果获取最小能量曲线,通过删除曲线上顶点并进行局部拼接实现模型缩放。传统三维模型缩放结果(Wang和Zhang,2009)与基于视觉显著性的缩放结果(Jia等,2014)对比如图4所示。在引入视觉显著性特征后,狮子面部区域(视觉显著性区域)形变程度较低,毛发区域(非视觉显著性区域)形变程度较高,整体视觉效果优于传统缩放结果。



图4 三维模型缩放结果对比

Fig. 4 Comparison of 3D mesh resizing results
(a) original mesh; (b) traditional resizing result;
(c) 3D-VSD guided resizing result)

4.2 三维模型简化

激光点云技术的发展使得采集的三维模型细节信息日渐丰富,但也不可避免地为模型数据处理与存储带来巨大压力,对其进行简化能够在尽可能保留模型特征的同时降低模型内顶点数量,以提高计算机处理效率。

传统方法通常采用顶点聚类或边缘折叠操作进行模型简化(Garland和Heckbert,1998),虽然能在一定程度上保留模型细节特征,但却忽略了简化后模型在视觉效果上的改变。针对这一问题,研究人员将视觉显著性检测引入三维模型简化任务中(An等,2016),通过增加视觉显著性区域的采样数量保证简化后模型的视觉效果。例如,张亦芳等人(2021)提出基于特征显著性的模型简化方法,通过构建视觉显著性词典计算各区域采样率,进而实现模型简化;Ding等人(2024b)提出视觉显著性引导的

彩色模型简化方法,通过融合纹理、几何特征获取视觉显著性检测结果,并将其作为权重指导模型简化过程。传统三维模型简化结果(Garland 和 Heckbert, 1997)与基于视觉显著性的简化结果(Yang 等, 2016)对比如图 5 所示。视觉显著性引导的方法在模型眼睛、鼻子等区域(视觉显著性区域)保留更多的顶点数量,在脖子、肩膀等区域(非视觉显著性区域)保留更少的顶点数量,简化后模型视觉效果更好。

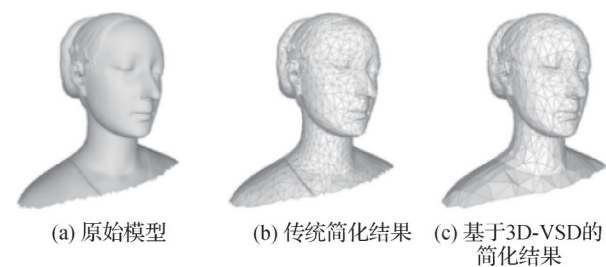


图 5 三维模型简化结果对比
Fig. 5 Comparison of 3D mesh simplification results
((a)original mesh;(b)traditional simplification result;
(c)3D-VSD guided resizing result)

4.3 三维模型降噪

三维模型在采集与传输过程中容易受到影响,导致数据包含噪声信息,对后续模型分割与理解任务造成困扰。因此,研究三维模型降噪方法具有重要意义。

传统降噪方法主要使用滤波对三维模型进行降噪处理,例如 Taubin (1995b)提出基于低通滤波的模型降噪方法。该方法具有较高的计算效率,但在降噪过程中并没有考虑模型表面特征,难以根据不同表面区域确定合适滤波参数。为解决这一问题,研究人员引入视觉显著性检测技术,以指导降噪过程中对模型中顶点位置的调整过程。例如,Nouri 等人(2017)提出基于视觉显著性特征图的彩色模型降噪方法,将视觉显著性特征值作为权重指导扩散过程,进而获取模型降噪结果;Li 等人(2019)提出基于特征可靠性的模型降噪方法,利用视觉显著性进行特征选择以实现可靠的模型降噪结果。传统三维模型降噪结果(Fleishman 等, 2003)与基于视觉显著性的降噪结果(Dutta 等, 2015)对比如图 6 所示。相较于传统降噪结果,基于视觉显著性的模型降噪方法处理得到的三维模型能够更好地抑制噪声信息,并较好地保留人物身体区域(视觉显著性区域)的重要特征。

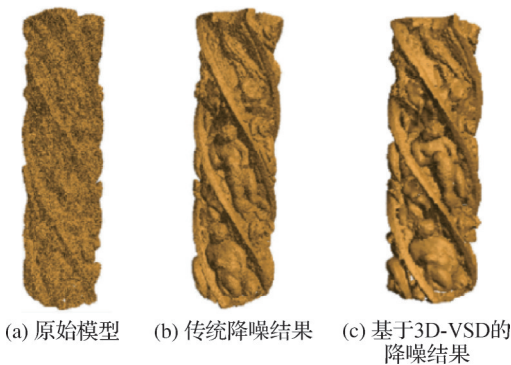


图 6 三维模型降噪结果对比
Fig. 6 Comparison of 3D mesh denoising results
((a)original mesh;(b)traditional denoising result;
(c)3D-VSD guided denoising result)

4.4 三维模型质量评价

在实际应用中,三维模型通常需要进行压缩、加密等处理,这些处理将不可避免地导致模型出现失真与质量退化(Feng 等, 2018; Zhang 等, 2022)。对模型质量进行客观、准确的评价能够协助优化模型压缩、加密等方法的性能。

传统质量评价方法通常使用均方根误差(Cignoni 等, 1998)或 Hausdorff 距离(Aspert 等, 2002)衡量模型间的差异。然而这些方法只从几何特征差异的角度衡量模型质量,评价结果与视觉主观感受存在较大差异。针对这一现状,Nouri 等人(2016)提出基于视觉显著性特征的三维模型质量评价方法,将视觉显著性检测结果与视觉掩蔽效应结合获得质量评价结果;Nehmé 等人(2021)将视觉显著性特征与几何、纹理特征结合,对三维模型质量进行评价。

除上述应用外,3D-VSD 方法也广泛应用于其他三维模型处理任务中,如形状匹配(Hu 等, 2020b)、三维重建(Li 等, 2020)和模型压缩(Nousias 等, 2020b)等。

5 结 语

近年来,研究人员在 3D-VSD 领域取得了一系列研究成果。在前文基础上,本文将对现有研究进行总结,并讨论未来发展方向。

5.1 当前研究总结

3D-VSD 方法分为基于手工特征的 3D-VSD 方法和基于深度学习的 3D-VSD 方法。

基于手工特征的 3D-VSD 方法视觉显著性检测

精度较大程度依赖于手工特征的表达能力。然而,现有手工特征表达能力往往不足,使得视觉显著性检测结果与真值间存在较大差距,难以满足现实应用需求。考虑到优异的手工特征设计高度依赖于研究人员经验,设计表达能力更强的人工特征较为困难。总体而言,现有方法主要存在两个局限:1)单尺度特征难以有效捕捉微小细节,在处理视觉显著性区域尺寸变化较大的三维模型时往往表现不佳;2)多尺度特征虽然能够较好地捕捉不同尺度视觉显著性区域,但是相较于单尺度方法,计算复杂度往往更高,并且在多通道/尺度特征融合过程中往往难以有效利用不同特征的信息,使得视觉显著性检测性能提升存在局限性。传统线性加权融合方法在融合过程中权重难以确定,并且固定的融合权重无法适用于所有三维模型,因此,需要深入挖掘不同通道/尺度特征间的内在关联,设计多特征自适应融合框架,以提升基于多尺度特征的3D-VSD方法性能。

基于深度学习的3D-VSD方法能够更高效地捕捉模型高层次语义信息,特征表达能力更强。该方法相较于基于手工特征的方法视觉显著性检测精度明显提升,但与真值间仍存在一定差距。受限于大规模训练数据集的缺乏,该类方法的视觉显著性检测性能难以进一步提升,主要解决思路在于增加训练数据数量。虽然部分方法通过投影变换方式利用成熟的图像视觉显著性检测数据集协助特征提取,但却不可避免地损伤(或丢失)三维模型深度信息,性能提升存在局限性。随着深度学习技术的发展,部分方法逐渐利用弱监督学习、迁移学习实现3D-VSD任务,这类方法利用三维模型分类标签协助视觉显著性检测,获得了较高的性能提升。此外,部分方法引入无监督学习思想或构建伪视觉显著性标签,以解决训练数据不足的问题,该类方法展现出较强的性能,值得进一步深入研究。

5.2 未来发展方向

3D-VSD方法在检测内容和检测方法方面均存在较大的发展空间。就检测内容而言,随着点云采集技术的不断进步,动态单体模型、复杂场景模型逐渐出现,未来需要投入大量精力对这类数据进行研究;就检测方法而言,基于深度学习的三维模型视觉显著性检测方法将是未来的研究重点,大规模训练数据集的构建也将是无法回避的问题。

1)检测内容。现有3D-VSD数据集大多由静态单体模型构成,模型类别较为有限并且单个模型包含的顶点数量较少。然而,在实际应用中,需要处理的三维模型种类往往复杂多样,并且数据量更大,这在一定程度上限制了现有数据集的应用。

动态单体模型包含丰富的纹理信息与运动信息,能够真实反映现实世界中物体的运动状态,对自动驾驶等领域具有重要意义(Liu等,2019)。相较于静态单体模型,动态单体模型往往包含更多的噪声信息,并且数据量巨大,为模型的快速处理带来巨大挑战。视觉显著性检测能够提取动态三维单体模型中的重要视觉信息,协助有限计算资源的合理分配,服务于动态单体模型的压缩、简化与分割等任务,具有较高的研究价值。目前,针对动态单体模型的视觉显著性检测数据集与方法还较为缺乏,如何在噪声影响下稳定、高效地提取动态单体模型视觉显著性特征值得进一步探索。

复杂场景模型能够逼真地展示现实世界,在工程设计等领域应用前景广阔。虽然一些复杂场景模型数据集已经被提出,例如Armeni等人(2017)提出大规模复杂场景模型数据集S3DIS,但是鲜有研究在此基础上构建视觉显著性检测数据集。区别于静态单体模型,复杂场景模型包含多个静态单体模型,模型类别丰富并且模型间的关联更加多样化,使得针对复杂场景模型进行视觉显著性检测更具有挑战性。现有针对静态单体模型设计的视觉显著性检测方法缺乏对于模型间关联的深入探索,难以直接拓展应用于面向复杂场景模型的视觉显著性检测任务。因此,在后续研究中,如何构建基于复杂场景模型的3D-VSD数据集值得进一步关注,如何设计不同模型间关联设计面向复杂场景模型的3D-VSD方法值得深入研究。

2)检测模型。受限于大规模训练数据集的缺乏,基于深度学习的3D-VSD方法性能仍具有较大提升空间。主要研究策略包括利用图像视觉显著性检测方法或引入弱监督学习辅助3D-VSD任务。前者主要通过对多视角2D投影图进行处理获取视觉显著性特征,并进行特征映射;后者主要利用弱监督方式结合2D投影图视觉显著性特征、三维模型类别标签等进行视觉显著性预测。此外,如何基于无监督学习思想设计3D-VSD方法值得关注,如何引入

迁移学习思想设计 3D-VSD 方法也值得探索。

参考文献 (References)

- Abid M, Da Silva M P and Le Callet P. 2020. Towards visual saliency computation on 3D graphical contents for interactive visualization// *Proceedings of 2020 IEEE International Conference on Image Processing*. Abu Dhabi, UAE: IEEE: 3448-3452 [DOI: 10.1109/ICIP40778.2020.9191064]
- Abouelaziz I, Chetouani A, El Hassouni M, Latecki L J and Cherifi H. 2020. 3D visual saliency and convolutional neural network for blind mesh quality assessment. *Neural Computing and Applications*, 32(21): 16589-16603 [DOI: 10.1007/s00521-019-04521-1]
- Achanta R, Shaji A, Smith K, Lucchi A, Fua P and Süsstrunk S. 2012. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11): 2274-2282 [DOI: 10.1109/TPAMI.2012.120]
- Alexiou E, Xu P S and Ebrahimi T. 2019. Towards modelling of visual saliency in point clouds for immersive applications//*Proceedings of 2019 IEEE International Conference on Image Processing*. Taipei, China: IEEE: 4325-4329 [DOI: 10.1109/ICIP.2019.8803479]
- Alfarasani D A, Sweetman T, Lai Y K and Rosin P L. 2022. Learning to predict 3D mesh saliency//*Proceedings of the 26th International Conference on Pattern Recognition*. Montreal, Canada: IEEE: 4023-4029 [DOI: 10.1109/ICPR56361.2022.9956579]
- An G M, Watanabe T and Kakimoto M. 2016. Mesh simplification using hybrid saliency//*Proceedings of 2016 International Conference on Cyberworlds*. Chongqing, China: IEEE: 231-234 [DOI: 10.1109/CW.2016.47]
- An P, Duan Y C, Huang Y L, Ma J, Chen Y F, Wang L H, Yang Y and Liu Q. 2024. SP-Det: leveraging saliency prediction for voxel-based 3D object detection in sparse point cloud. *IEEE Transactions on Multimedia*, 26: 2795-2808 [DOI: 10.1109/TMM.2023.3304054]
- Armeni I, Sax S, Zamir A R and Savarese S. 2017. Joint 2D-3D-semantic data for indoor scene understanding [EB/OL]. [2024-10-21]. <https://arxiv.org/pdf/1702.01105.pdf>
- Arvanitis G, Lalos A S and Moustakas K. 2021. Robust and fast 3-D saliency mapping for industrial modeling applications. *IEEE Transactions on Industrial Informatics*, 17(2): 1307-1317 [DOI: 10.1109/TII.2020.3003455]
- Aspert N, Santa-Cruz D and Ebrahimi T. 2002. MESH: measuring errors between surfaces using the Hausdorff distance//*Proceedings of 2002 IEEE International Conference on Multimedia and Expo*. Lausanne, Switzerland: IEEE: 705-708 [DOI: 10.1109/ICME.2002.1035879]
- Chen C L Z, Wang G T, Peng C, Zhang X W and Qin H. 2020. Improved robust video saliency detection based on long-term spatial-temporal information. *IEEE Transactions on Image Process-*
- ing, 29: 1090-1100 [DOI: 10.1109/TIP.2019.2934350]
- Chen G, Dai H, Zhou T, Shen J B and Shao L. 2023. Automatic schelling point detection from meshes. *IEEE Transactions on Visualization and Computer Graphics*, 29(6): 2926-2939 [DOI: 10.1109/TVCG.2022.3144143]
- Chen X B, Saparov A, Pang B and Funkhouser T. 2012. Schelling points on 3D surface meshes. *ACM Transactions on Graphics*, 31(4): #29 [DOI: 10.1145/2185520.2185525]
- Cignoni P, Rocchini C and Scopigno R. 1998. Metro: measuring error on simplified surfaces. *Computer Graphics Forum*, 17(2): 167-174 [DOI: 10.1111/1467-8659.00236]
- Ding G Q, Imamoglu N, Caglayan A, Murakawa M and Nakamura R. 2022. SalLiDAR: saliency knowledge transfer learning for 3D point cloud understanding//*Proceedings of the 33rd British Machine Vision Conference*. London, UK: BMVA: #584
- Ding G Q, Imamoglu N, Caglayan A, Murakawa M and Nakamura R. 2024a. Attention-guided LiDAR segmentation and odometry using image-to-point cloud saliency transfer. *Multimedia Systems*, 30(4): #188 [DOI: 10.1007/s00530-024-01389-7]
- Ding X Y, Chen Z, Lin W S and Chen Z Z. 2024b. Towards 3D colored mesh saliency: database and benchmarks. *IEEE Transactions on Multimedia*, 26: 3580-3591 [DOI: 10.1109/TMM.2023.3312924]
- Ding X Y and Chen Z Z. 2022. Towards mesh saliency in 6 degrees of freedom. *Neurocomputing*, 502: 120-139 [DOI: 10.1016/j.neucom.2022.06.088]
- Ding X Y, Lin W S, Chen Z Z and Zhang X F. 2019. Point cloud saliency detection by local and global feature fusion. *IEEE Transactions on Image Processing*, 28(11): 5379-5393 [DOI: 10.1109/TIP.2019.2918735]
- dos Anjos R K, Roberts R A, Allen B, Jorge J and Anjo K. 2023. Saliency detection for large-scale mesh decimation. *Computers and Graphics*, 111: 63-76 [DOI: 10.1016/j.cag.2023.01.012]
- Dutagaci H, Cheung C P and Godil A. 2012. Evaluation of 3D interest point detection techniques via human-generated ground truth. *The Visual Computer*, 28(9): 901-917 [DOI: 10.1007/s00371-012-0746-4]
- Dutta S, Banerjee S, Biswas P K and Bhowmick P. 2015. Mesh denoising using multi-scale curvature-based saliency//*Proceedings of 2015 Asian Conference on Computer Vision Workshops*. Singapore, Singapore: Springer: 507-516 [DOI: 10.1007/978-3-319-16631-5_37]
- Elad M. 2005. Retinex by two bilateral filters//*Proceedings of the 5th International Conference on Scale-Space Theories in Computer Vision*. Hofgeismar, Germany: Springer: 217-229 [DOI: 10.1007/11408031_19]
- Feng X, Wan W G, Xu R Y D, Perry S, Zhu S and Liu Z X. 2018. A new mesh visual quality metric using saliency weighting-based pooling strategy. *Graphical Models*, 99: 1-12 [DOI: 10.1016/j.gmod.2018.07.001]

- Figueiredo V F, Sandri G L, de Queiroz R L and Chou P A. 2020. Saliency maps for point clouds//Proceedings of the 2nd IEEE International Workshop on Multimedia Signal Processing. Tampere, Finland; IEEE: 1-5 [DOI: 10.1109/MMSP48831.2020.9287102]
- Fleishman S, Drori I and Cohen-Or D. 2003. Bilateral mesh denoising. *ACM Transactions on Graphics*, 22(3): 950-953 [DOI: 10.1145/882262.88236]
- Garland M and Heckbert P S. 1997. Surface simplification using quadric error metrics//Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques. Los Angeles, USA; ACM: 209-216 [DOI: 10.1145/258734.258849]
- Garland M and Heckbert P S. 1998. Simplifying surfaces with color and texture using quadric error metrics//Proceedings of Visualization. Research Triangle Park, USA; IEEE: 263-269 [DOI: 10.1109/VISUAL.1998.745312]
- Guo Y, Wang F and Xin J M. 2018. Point-wise saliency detection on 3D point clouds via covariance descriptors. *The Visual Computer*, 34(10): 1325-1338 [DOI: 10.1007/s00371-017-1416-3]
- Hou T B and Qin H. 2013. Admissible diffusion wavelets and their applications in space-frequency processing. *IEEE Transactions on Visualization and Computer Graphics*, 19(1): 3-15 [DOI: 10.1109/TVCG.2012.111]
- Howlett S, Hamill J and O'Sullivan C. 2005. Predicting and evaluating saliency for simplified polygonal models. *ACM Transactions on Applied Perception (TAP)*, 2(3): 286-308 [DOI: 10.1145/1077399.1077406]
- Hu S F, Liang X H, Shum H P H, Li F W B and Aslam N. 2020a. Sparse metric-based mesh saliency. *Neurocomputing*, 400: 11-23 [DOI: 10.1016/j.neucom.2020.02.106]
- Hu S F, Shum H P H, Aslam N, Li F W B and Liang X H. 2020b. A unified deep metric representation for mesh saliency detection and non-rigid shape matching. *IEEE Transactions on Multimedia*, 22(9): 2278-2292 [DOI: 10.1109/TMM.2019.2952983]
- Jeong S W, Yun J S and Sim J Y. 2020. Multiscale saliency detection for colored 3D point clouds based on random walk//Proceedings of 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. Auckland, New Zealand; IEEE: 1193-1200
- Jia S X, Zhang C M, Li X M and Zhou Y F. 2014. Mesh resizing based on hierarchical saliency detection. *Graphical Models*, 76(5): 355-362 [DOI: 10.1016/j.gmod.2014.03.012]
- Jiang M, Huang S S, Duan J Y and Zhao Q. 2015. SALICON: saliency in context//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA; IEEE: 1072-1080 [DOI: 10.1109/CVPR.2015.7298710]
- Jiang Z Y, Ding L Y, Tam G K L, Song C, Li F W B and Yang B L. 2023. C2SPoint: a classification-to-saliency network for point cloud saliency detection. *Computers and Graphics*, 115: 274-284 [DOI: 10.1016/j.cag.2023.07.003]
- Kim Y, Varshney A, Jacobs D W and Guimbretière F. 2010. Mesh saliency and human eye fixations. *ACM Transactions on Applied Perception (TAP)*, 7(2): #12 [DOI: 10.1145/1670671.1670676]
- Kraevoy V, Sheffer A, Shamir A and Cohen-Or D. 2008. Non-homogeneous resizing of complex models. *ACM Transactions on Graphics (TOG)*, 27(5): #111 [DOI: 10.1145/1409060.1409064]
- Lavoué G, Cordier F, Seo H and Larabi M C. 2018. Visual attention for rendered 3D shapes. *Computer Graphics Forum*, 37(2): 191-203 [DOI: 10.1111/cgf.13353]
- Leal E, Sanchez-Torres G, Branch-Bedoya J W, Abad F and Leal N. 2021. A saliency-based sparse representation method for point cloud simplification. *Sensors*, 21(13): #4279 [DOI: 10.3390/s21134279]
- Lee C H, Varshney A and Jacobs D W. 2005. Mesh saliency. *ACM Transactions on Graphics (TOG)*, 24(3): 659-666 [DOI: 10.1145/1073204.1073244]
- Leifman G, Shtrom E and Tal A. 2016. Surface regions of interest for viewpoint selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(12): 2544-2556 [DOI: 10.1109/TPAMI.2016.2522437]
- Li B, Yang Y and Liu Q. 2021. RGB-D video saliency detection via superpixel-level conditional random field. *Journal of Image and Graphics*, 26(4): 872-882 (李贝, 杨铀, 刘琼. 2021. 超像素条件随机场下的RGB-D视频显著性检测. *中国图象图形学报*, 26(4): 872-882) [DOI: 10.11834/jig.200122]
- Li H, Ye W C, Zhang G F, Zhang S Y and Bao H J. 2020. Saliency guided subdivision for single-view mesh reconstruction//Proceedings of 2020 International Conference on 3D Vision. Fukuoka, Japan; IEEE: 1098-1107 [DOI: 10.1109/3DV50981.2020.00120]
- Li T, Liu W, Liu H, Wang J and Liu L G. 2019. Feature-convincing mesh denoising. *Graphical Models*, 101: 17-26 [DOI: 10.1016/j.gmod.2018.12.002]
- Lin T Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P and Zitnick C L. 2014. Microsoft COCO: common objects in context//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland; Springer: 740-755 [DOI: 10.1007/978-3-319-10602-1_48]
- Lin W S and Lee S. 2022. Visual saliency and quality evaluation for 3D point clouds and meshes: an overview. *APSIPA Transactions on Signal and Information Processing*, 11(1): #28 [DOI: 10.1561/116.00000125]
- Linardos A, Kümmerer M, Press O and Bethge M. 2021. DeepGaze IIE: calibrated prediction in and out-of-domain for state-of-the-art saliency modeling//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada; IEEE: 12899-12908 [DOI: 10.1109/ICCV48922.2021.01268]
- Liu C M, Luan W N, Fu R H, Pang H B and Li Y H. 2023. Attention-embedding mesh saliency. *The Visual Computer*, 39(5): 1783-

- 1795 [DOI: 10.1007/s00371-022-02444-y]
- Liu X P, Tao P P, Cao J J, Chen H and Zou C Q. 2016a. Mesh saliency detection via double absorbing Markov chain in feature space. *The Visual Computer*, 32(9): 1121-1132 [DOI: 10.1007/s00371-015-1184-x]
- Liu X Y, Liu L G, Song W J, Liu Y P and Ma L Z. 2016b. Shape context based mesh saliency detection and its applications: a survey. *Computers and Graphics*, 57: 12-30 [DOI: 10.1016/j.cag.2016.03.001]
- Liu X Y, Yan M Y and Bohg J. 2019. MeteorNet: deep learning on dynamic 3D point cloud sequences//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, Korea (South); IEEE: 9245-9254 [DOI: 10.1109/ICCV.2019.00934]
- Luan W N, Liu C M and Pang H B. 2020. Skeleton-bridged mesh saliency and mesh simplification//*Proceedings of the 4th International Conference on Computer Science and Artificial Intelligence*. Zhuhai, China: ACM: 278-283 [DOI: 10.1145/3445815.3445861]
- Martin D, Fandos A, Masia B and Serrano A. 2024. SAL3D: a model for saliency prediction in 3D meshes. *The Visual Computer*, 40(11): 7761-7771 [DOI: 10.1007/s00371-023-03206-0]
- Maximo A, Patro R, Varshney A and Farias R. 2011. A robust and rotationally invariant local surface descriptor with applications to non-local mesh processing. *Graphical Models*, 73(5): 231-242 [DOI: 10.1016/j.gmod.2011.05.002]
- Miao Y W and Lin H B. 2013. Visual saliency guided global and local resizing for 3D models//*Proceedings of 2013 International Conference on Computer-Aided Design and Computer Graphics*. Guangzhou, China: IEEE: 212-219 [DOI: 10.1109/CADGraphics.2013.35]
- Nehmé Y, Abid M, Lavoué G, Da Silva M P and Le Callet P. 2021. Cmdm-Vac: improving a perceptual quality metric for 3D graphics by integrating a visual attention complexity measure//*Proceedings of 2021 IEEE International Conference on Image Processing*. Anchorage, USA: IEEE: 3368-3372 [DOI: 10.1109/ICIP42928.2021.9506662]
- Nie W Z, Qu L, Ren M J, Liang Q, Su Y T, Li Y Y and Jin H. 2020. Two-stream network based on visual saliency sharing for 3D model recognition. *IEEE Access*, 8: 5979-5989 [DOI: 10.1109/ACCESS.2019.2963511]
- Nouri A, Charrier C and Lézoray O. 2015a. Mesh saliency with adaptive local patches//*Proceedings of SPIE 9393, Three-Dimensional Image Processing, Measurement (3DIPM), and Applications*. San Francisco, USA: SPIE: 98-109 [DOI: 10.1117/12.2076656]
- Nouri A, Charrier C and Lézoray O. 2015b. Multi-scale mesh saliency with local adaptive patches for viewpoint selection. *Signal Processing: Image Communication*, 38: 151-166 [DOI: 10.1016/j.image.2015.08.002]
- Nouri A, Charrier C and Lézoray O. 2015c. Multi-scale saliency of 3D colored meshes//*Proceedings of 2015 IEEE International Conference on Image Processing*. Quebec City, Canada: IEEE: 2820-2824 [DOI: 10.1109/ICIP.2015.7351317]
- Nouri A, Charrier C and Lézoray O. 2016. Full-reference saliency-based 3D mesh quality assessment index//*Proceedings of 2016 IEEE International Conference on Image Processing*. Phoenix, USA: IEEE: 1007-1011 [DOI: 10.1109/ICIP.2016.7532509]
- Nouri A, Charrier C and Lézoray O. 2017. Global visual saliency: geometric and colorimetric saliency fusion and its applications for 3D colored meshes//*Proceedings of the 7th International Conference on Image Processing Theory, Tools and Applications*. Montreal, Canada: IEEE: 1-6 [DOI: 10.1109/IPTA.2017.8310142]
- Nouri A, Charrier C and Lézoray O. 2022. A genetically based combination of visual saliency and roughness for FR 3D mesh quality assessment: a statistical study. *The Computer Journal*, 65(3): 606-620 [DOI: 10.1093/comjnl/bxaa089]
- Nouri A, Lézoray O, Ibork Z, Charrier C and Touahni R. 2024. A fully automatic colorimetric saliency detection approach for 3D meshes//*Proceedings of the 12th IEEE International Symposium on Signal, Image, Video and Communications*. Marrakech, Morocco: IEEE: 1-5 [DOI: 10.1109/ISIVC61350.2024.10577926]
- Nousias S, Arvanitis G, Lalos A and Moustakas K. 2023. Deep saliency mapping for 3D meshes and applications. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(2): #71 [DOI: 10.1145/3550073]
- Nousias S, Arvanitis G, Lalos A S and Moustakas K. 2020a. Mesh saliency detection using convolutional neural networks//*Proceedings of 2020 IEEE International Conference on Multimedia and Expo*. London, UK: IEEE: 1-6 [DOI: 10.1109/ICME46284.2020.9102796]
- Nousias S, Arvanitis G, Lalos A S, Pavlidis G, Koulamas C, Kalogeras A and Moustakas K. 2020b. A saliency aware CNN-based 3D model simplification and compression framework for remote inspection of heritage sites. *IEEE Access*, 8: 169982-170001 [DOI: 10.1109/ACCESS.2020.3023167]
- Nozawa N, Fukusato T and Morishima S. 2017. 3D model partial-resizing via normal and texture map combination//*Proceedings of the ACM SIGGRAPH 2017 Posters*. Los Angeles, USA: ACM: #1 [DOI: 10.1145/3102163.3102203]
- Qi C, Su H, Mo K C and Guibas L J. 2017a. PointNet: deep learning on point sets for 3D classification and segmentation//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE: 77-85 [DOI: 10.1109/CVPR.2017.16]
- Qi C, Yi L, Su H and Guibas L J. 2017b. PointNet++: deep hierarchical feature learning on point sets in a metric space//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, USA: Curran Associates Inc.: 5105-5114 [DOI: 10.48550/arXiv.1706.02413]

- Rouhafzay G, Cretu A M and Payeur P. 2023. A deep model of visual attention for saliency detection on 3D objects. *Neural Processing Letters*, 55(7): 8847-8867 [DOI: 10.1007/s11063-023-11180-w]
- Rusu R B, Blodow N and Beetz M. 2009. Fast point feature histograms (FPFH) for 3D registration//*Proceedings of 2009 IEEE International Conference on Robotics and Automation*. Kobe, Japan: IEEE: 3212-3217 [DOI: 10.1109/ROBOT.2009.5152473]
- Shtrom E, Leifman G and Tal A. 2013. Saliency detection in large point sets//*Proceedings of 2013 IEEE International Conference on Computer Vision*. Sydney, Australia: IEEE: 3591-3598 [DOI: 10.1109/ICCV.2013.446]
- Song R, Liu Y H, Martin R R and Echavarria K R. 2018. Local-to-global mesh saliency. *The Visual Computer*, 34 (3): 323-336 [DOI: 10.1007/s00371-016-1334-9]
- Song R, Liu Y H, Martin R R and Rosin P L. 2014. Mesh saliency via spectral processing. *ACM Transactions on Graphics (TOG)*, 33(1): #6 [DOI: 10.1145/2530691]
- Song R, Liu Y H and Rosin P L. 2021a. Mesh saliency via weakly supervised classification-for-saliency CNN. *IEEE Transactions on Visualization and Computer Graphics*, 27(1): 151-164 [DOI: 10.1109/TVCG.2019.2928794]
- Song R, Liu Y H, Zhao Y T, Martin R R and Rosin P L. 2012. Conditional random field-based mesh saliency//*Proceedings of the 19th IEEE International Conference on Image Processing*. Orlando, USA: IEEE: 637-640 [DOI: 10.1109/ICIP.2012.6466940]
- Song R, Zhang W, Zhao Y T and Liu Y H. 2022. Unsupervised multi-view CNN for salient view selection and 3D interest point detection. *International Journal of Computer Vision*, 130 (5): 1210-1227 [DOI: 10.1007/s11263-022-01592-x]
- Song R, Zhang W, Zhao Y T, Liu Y H and Rosin P L. 2021b. Mesh saliency: an independent perceptual measure or a derivative of image saliency?//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville, USA: IEEE: 8853-8862 [DOI: 10.1109/CVPR46437.2021.00874]
- Song R, Zhang W, Zhao Y T, Liu Y H and Rosin P L. 2023. 3D visual saliency: an independent perceptual measure or a derivative of 2D image saliency? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45 (11): 13083-13099 [DOI: 10.1109/TPAMI.2023.3287356]
- Sun J, Ovsjanikov M and Guibas L. 2009. A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum*, 28(5): 1383-1392 [DOI: 10.1111/j.1467-8659.2009.01515.x]
- Sun Y, Chen C Y, Hu X J, Li L and Xing Q W. 2022. Saliency detection algorithm of panoramic images using joint weighting with observer's attention longitude. *Journal of Image and Graphics*, 27(4): 1322-1334 (孙耀, 陈纯毅, 胡小娟, 李凌, 邢琦玮. 2022. 观看经度联合加权全景图显著性检测算法. *中国图象图形学报*, 27(4): 1322-1334) [DOI: 10.11834/jig.200682]
- Tabia H and Laga H. 2015. Covariance-based descriptors for efficient 3D shape matching, retrieval, and classification. *IEEE Transactions on Multimedia*, 17 (9): 1591-1603 [DOI: 10.1109/TMM.2015.2457676]
- Tao P P, Cao J J, Li S H, Liu X P and Liu L G. 2015. Mesh saliency via ranking unsalient patches in a descriptor space. *Computers and Graphics*, 46: 264-274 [DOI: 10.1016/j.cag.2014.09.023]
- Tao P P, Zhang L N, Cao J J and Liu X P. 2016. Mesh saliency detection based on entropy//*Proceedings of the 6th International Conference on Digital Home*. Guangzhou, China: IEEE: 288-295 [DOI: 10.1109/ICDH.2016.065]
- Tasse F P, Kosinka J and Dodgson N. 2015. Cluster-based point set saliency//*Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago, Chile: IEEE: 163-171 [DOI: 10.1109/ICCV.2015.27]
- Taubin G. 1995a. Estimating the tensor of curvature of a surface from a polyhedral approximation//*Proceedings of 1995 IEEE International Conference on Computer Vision*. Cambridge, USA: IEEE: 902-907 [DOI: 10.1109/ICCV.1995.466840]
- Taubin G. 1995b. A signal processing approach to fair surface design//*Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*. Los Angeles, USA: ACM: 351-358 [DOI: 10.1145/218380.218473]
- Tavakoli H R, Ahmed F, Borji A and Laaksonen J. 2017. Saliency revisited: analysis of mouse movements versus fixations//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE: 1774-1782 [DOI: 10.1109/CVPR.2017.673]
- Tavakoli H R and Laaksonen J. 2017. Bottom-up fixation prediction using unsupervised hierarchical models//*Proceedings of 2017 Asian Conference on Computer Vision Workshops*. Taipei, China: Springer: 287-302 [DOI: 10.1007/978-3-319-54407-6_19]
- Tliba M, Kerkouri M A, Ghariba B, Chetouani A, Çöltekin A, Shehata M S and Bruno A. 2022. SATSal: a multi-level self-attention based architecture for visual saliency prediction. *IEEE Access*, 10: 20701-20713 [DOI: 10.1109/ACCESS.2022.3152189]
- Tliba M, Zhou X M, Viola I, Cesar P, Chetouani A, Valenzise G and Dufaux F. 2024. Enhancing immersive experiences through 3D point cloud analysis: a novel framework for applying 2D visual saliency models to 3D point clouds//*Proceedings of the 16th International Conference on Quality of Multimedia Experience*. Karlshamn, Sweden: IEEE: 307-313 [DOI: 10.1109/QoMEX61742.2024.10598254]
- Tombari F, Salti S and Di Stefano L. 2010. Unique signatures of histograms for local surface description//*Proceedings of the 11th European Conference on Computer Vision*. Heraklion, Greece: Springer: 356-369 [DOI: 10.1007/978-3-642-15558-1_26]
- Wang H Y, Luo H, Wen C L, Cheng J, Li P, Chen Y P, Wang C and Li J. 2015a. Road boundaries detection based on local normal saliency from mobile laser scanning data. *IEEE Geoscience and Remote Sensing Letters*, 12 (10): 2085-2089 [DOI: 10.1109/

- LGRS.2015.2449074]
- Wang K P and Zhang C M. 2009. Content-aware model resizing based on surface deformation. *Computers and Graphics*, 33 (3): 433-438 [DOI: 10.1016/j.cag.2009.03.004]
- Wang S F, Li N N, Li S, Luo Z X, Su Z X and Qin H. 2015b. Multi-scale mesh saliency based on low-rank and sparse analysis in shape feature space. *Computer Aided Geometric Design*, 35-36: 206-214 [DOI: 10.1016/j.cagd.2015.03.003]
- Wang X, Koch S, Holmqvist K and Alexa M. 2018. Tracking the gaze on objects in 3D: how do people really look at the bunny? *ACM Transactions on Graphics*, 37 (6): #188 [DOI: 10.1145/3272127.3275094]
- Wang X, Lindlbauer D, Lessig C, Maertens M and Alexa M. 2016. Measuring the visual salience of 3D printed objects. *IEEE Computer Graphics and Applications*, 36 (4): 46-55 [DOI: 10.1109/MCG.2016.47]
- Wang X D, Kang F J, Gu H and Liang H T. 2022. Regional importance detection of 3D mesh via fusion of local color difference and curvature entropy. *International Journal of Modeling, Simulation, and Scientific Computing*, 13 (6): #2250060 [DOI: 10.1142/S179396232250060X]
- Wang X D, Liang H T, Kang F J and Gu H. 2017. Multi-scale mesh saliency with local patch weighted curvature entropy. *Journal of System Simulation*, 29 (9): 1976-1983 (汪小东, 梁洪涛, 康凤举, 顾浩. 2017. 基于局部区块加权曲率熵的多尺度网格显著性(英文). *系统仿真学报*, 29 (9): 1973-1983) [DOI: 10.16182/j.issn1004731x.joss.201709014]
- Wang Z Q, Zhang Y S, Yu Y, Min J and Tian H. 2022. Review of deep learning based salient object detection. *Journal of Image and Graphics*, 27 (7): 2112-2128 (王自全, 张永生, 于英, 闵杰, 田浩. 2022. 深度学习背景下视觉显著性物体检测综述. *中国图象图形学报*, 27 (7): 2112-2128) [DOI: 10.11834/jig.200649]
- Wang Z Y, Zhang Y J, Yang Q, Xu Y L, Sun J and Liu S. 2023. Point cloud quality assessment using 3D saliency maps//*Proceedings of 2023 IEEE International Conference on Visual Communications and Image Processing*. Jeju, Korea (South): IEEE: 1-5 [DOI: 10.1109/VCIP59821.2023.10402711]
- Wei N, Gao K Y, Ji R R and Chen P. 2018. Surface saliency detection based on curvature co-occurrence histograms. *IEEE Access*, 6: 54536-54541 [DOI: 10.1109/ACCESS.2018.2872168]
- Wu J L, Shen X Y, Zhu W and Liu L G. 2013. Mesh saliency with global rarity. *Graphical Models*, 75 (5): 255-264 [DOI: 10.1016/j.gmod.2013.05.002]
- Yang B L, Li F W B, Wang X, Xu M L, Liang X H, Jiang Z Y and Jiang Y H. 2016. Visual saliency guided textured model simplification. *The Visual Computer*, 32 (11): 1415-1432 [DOI: 10.1007/s00371-015-1129-4]
- Yang H, Chen R, An S P, Wei H and Zhang H. 2023. The growth of image-related three dimensional reconstruction techniques in deep learning-driven era: a critical summary. *Journal of Image and Graphics*, 28 (8): 2396-2409 (杨航, 陈瑞, 安仕鹏, 魏豪, 张衡. 2023. 深度学习背景下的图像三维重建技术进展综述. *中国图象图形学报*, 28 (8): 2396-2409) [DOI: 10.11834/jig.220376]
- Yun J S and Sim J Y. 2016. Supervoxel-based saliency detection for large-scale colored 3D point clouds//*Proceedings of 2016 IEEE International Conference on Image Processing*. Phoenix, USA: IEEE: 4062-4066 [DOI: 10.1109/ICIP.2016.7533123]
- Zhang Y F, Li L and Liu G S. 2021. Adaptive simplification of point cloud based on feature saliency. *Computer Engineering and Design*, 42 (8): 2211-2217 (张亦芳, 李立, 刘光帅. 2021. 基于特征显著性的点云自适应精简. *计算机工程与设计*, 42 (8): 2211-2217) [DOI: 10.16208/j.issn1000-7024.2021.08.016]
- Zhang Z C, Sun W, Min X K, Wang T, Lu W and Zhai G T. 2022. No-reference quality assessment for 3D colored point cloud and mesh models. *IEEE Transactions on Circuits and Systems for Video Technology*, 32 (11): 7618-7631 [DOI: 10.1109/TCSVT.2022.3186894]
- Zhao Y T, Liu Y H, Song R and Zhang M. 2012. Extended non-local means filter for surface saliency detection//*Proceedings of the 19th IEEE International Conference on Image Processing*. Orlando, USA: IEEE: 633-636 [DOI: 10.1109/ICIP.2012.6466939]
- Zhao Y T, Liu Y H, Wang Y J, Wei B G, Yang J, Zhao Y F and Wang Y T. 2016. Region-based saliency estimation for 3D shape analysis and understanding. *Neurocomputing*, 197: 1-13 [DOI: 10.1016/j.neucom.2016.01.012]
- Zheng T H, Chen C Y, Yuan J S, Li B and Ren K. 2019. Pointcloud saliency maps//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, Korea (South): IEEE: 1598-1606 [DOI: 10.1109/ICCV.2019.00168]
- Zhu J, Yang L, Liu W X and Xiao C X. 2014. Salient features preserved resizing of point-based geometry. *Journal of Computer-Aided Design and Computer Graphics*, 26 (10): 1624-1632 (朱建, 杨龙, 刘维星, 肖春霞. 2014. 显著性特征保持的点云模型缩放. *计算机辅助设计与图形学学报*, 26 (10): 1624-1632)

作者简介

丁晓颖, 女, 副教授, 主要研究方向为计算机视觉、显著性检测和视觉质量评估. E-mail: dingxiaoying@zuel.edu.cn

陈震中, 通信作者, 男, 教授, 主要研究方向为多媒体信号处理、计算视觉和多媒体数据挖掘.

E-mail: zzchen@whu.edu.cn

张新峰, 男, 副教授, 主要研究方向为视频编解码技术、质量评价和视频增强处理、点云压缩和处理.

E-mail: xfzhang@ucas.ac.cn

Lin Weisi, 男, 教授, 主要研究方向为图像处理、感知建模、视频压缩、多媒体通信和计算机视觉.

E-mail: wslin@ntu.edu.sg