

密级: _____



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

三维场景的观察度量及其应用研究

作者姓名: 韩红雷

指导教师: 王文成 研究员

中国科学院软件研究所

学位类别: 工学博士

学科专业: 计算机应用技术

培养单位: 中国科学院软件研究所

2014 年 11 月

Research on Exploration Metrics
and
Their Applications of the 3D Scene

By
Honglei Han

A Dissertation Submitted to
The University of Chinese Academy of Sciences
In partial fulfillment of the requirement
For the degree of
Doctor of Computer Applications

Institute of Software
Chinese Academy of Sciences

November, 2014

摘要

近些年，随着三维模型、交互式虚拟环境及游戏场景等在内的数字三维内容不断丰富，对这些内容进行有效观察及管理的需求变得十分迫切。目前人们主要通过二维手段观察和操作这些三维内容。因此，如何将三维内容通过二维显示手段有效展示给用户就变得十分关键，而对场景中的虚拟摄像机的观察质量和关注度进行度量是解决这一问题的关键。在针对单个三维模型的处理方面，本论文主要研究如何对其观察视点的质量进行量化，并根据量化结果选择恰当的模型代表视点，或者将结果应用于模型摆正当中。在三维交互场景方面，本论文提出一种不依赖硬件的用户关注度测量方法，有利于解决关注度测量时准确度和复杂性之间的矛盾。本论文的主要贡献和创新点归纳如下：

(1) 提出一种基于显著度分割的三维模型视点质量评分方法。在众多三维模型视点评分方法当中，基于结构化分解的方法更容易获得可信的评分结果。这些方法都需要将模型进行分割，并使用经验值为每个分块赋予重要度权重，但这并不一定能真实反映分块的重要度。为此，本论文提出一种新的基于分块的视点评分方法来解决这个问题。将衡量人眼视觉感知的网格显著度引入进来，将其作为模型分块及分块重要度的计算依据。这样，可以只利用模型的几何结构来自动获取带有视觉重要度的模型分块结果，并将其作为视点评分的基本计算单元，实验结果表明，这种方法对视点的评分更符合人眼的观察习惯。

(2) 提出一种结合视觉感知与信息量的视点质量评分方法。传统的视点评分方法存在的问题是，要么计算时间冗长，要么结果很难符合人类观察习惯。对此，本文提出一种基于视觉感知信息量的快速视点评分方法。首先，将计算相对简单的平均曲率作为度量三维物体视觉特征的要素。然后，结合信息熵理论，对视点进行评分，使那些能看到尽可能多重要特征且所见特征分布较广的视点分数较高。最后，将这种评分策略应用于最优视点集的计算当中，利用信息理论对视点集获取信息进行量化，以确保利用最少数目的视点来有效认知三维物体。实验结果表明，该方法得到的视点分数及最优视点集质量与目前最好的方法相当，但本论文方法计算速度更快，且无需人工交互。

(3) 提出一种基于视点评分的三维模型摆正方法。三维模型摆正方向的确定，有利于模型对齐、功能恢复等应用。但已有方法只关注于人造模型的处

理，且多数还需要人工干预，工作效率不高。本论文提出一种可同时处理人造模型和自然模型的方法，具有很好的适应性。这主要是基于以下的观察：无论是人造模型还是自然模型，其底部往往是表面内容很贫乏的区域；而那些主要观察这些区域的视点，其视点评分一般很低。因此，根据视点评分，就能有效地找出模型的底部朝向，即模型的摆正方向。为提高计算的可靠性，进一步引入了物体摆放稳定性和人们观察物体习惯性的度量。实验表明，新方法能有效处理各种模型，即便在处理人造模型时，也能很好地处理一些已有方法难以处理的情况，并且有很好的计算效率。

(4) 提出一种虚拟场景漫游中用户关注度的内隐测量方法。使用眼动仪等硬件是测量用户关注度的主要方式，但其可靠性差、复杂度高且价格昂贵，导致其难以有效应用于三维漫游场景中。本论文提出了一种易于嵌入到虚拟场景中的用户关注度软件测量方法，将用户操作虚拟摄像机漫游场景看作关注度内隐表达过程，通过物体相对摄像机中心的距离、遮挡情况和投影面积，结合观察时间对用户的关注度进行量化。论文利用大量用户学习的方式验证了这一关注度测量方式的准确性，并将其应用于游戏场景物体的量化评估中，获得了较好的效果。

关键词： 视点评分,模型分割,摄像机控制,感知,最优视点集,信息熵,视觉关注度,虚拟漫游,场景设计,显著度

Abstract

Nowadays, there exists huge number of three-dimensional digital contents, such as three-dimensional models, interactive virtual environments, and game scenes. The efficient observation and management of these contents are becoming more urgent. However, the common way to display 3D models is still through two-dimensional means. Therefore, the highlight is that how to efficiently present the 3D content by means of the 2D display. How to measure the quality and gaze parameters of the virtual camera in a 3D scene is the key to solve this problem. While dealing with a single 3D model, this paper mainly studies how to quantify the quality of its viewpoints, and then we could select the best view to observe it. The result of viewpoints scoring can then be applied to search the upright orientation of 3D objects. In a 3D interactive scenario, the paper presents a no hardware measurement of user's attention while interacting with the 3D scene. And this helps to resolve the conflicts between the accuracy and complexity of the measurement of user's gaze attention.

The main contributions and innovations of this paper are summarized as follows:

(1) We propose a new viewpoint scoring method based on saliency segmentation. Generating preferred viewpoints are important to promote model investigation. Semantic based viewpoint ranking methods are more likely to get reasonable results as they have the properties related to cognitive laws. However, it is required to decompose the model and assign the decomposed parts with suitable semantic meaning, and the two steps are always difficult to handle with. In this paper, a new viewpoint ranking method is presented to challenge this problem, which performs saliency-based segmentation for viewpoint measurement. The trouble task of assigning semantic meaning will be solved by mesh saliency which has been verified consistent with human perception. So, we are able to get a reasonable viewpoint ranking result using geometric computation only, as the model features with perceptual meaning can be distinguished well by mesh

saliency. Experimental results show the effectiveness of our method, and attested by a user study.

(2) We propose an efficiency viewpoint scoring method based on perception and information entropy. With the rapid growing of 3D digital contents, there is an urgent necessity for effective retrieval and displaying of the data. Viewpoint selection via viewpoint scoring is an important approach to the problem. However, traditional viewpoint scoring either requires complex computation, or fails to provide results that conform to human visual habits. To this, we propose a hybrid viewpoint scoring measure using visual perception and information entropy. In this paper, mean curvature is regarded as a key factor for visual perception of 3D objects. We score the viewpoints by blending visual perception and information entropy. Viewpoints observing more important features with wider distribution can achieve higher scores. This scoring measure is then applied in the computation of the N best viewpoints, while the information acquired from the set of viewpoints are further quantized using information theory to ensure efficient perception of the 3D objects with minimum number of viewpoints. Experimental results show that the best viewpoints obtained by the proposed method can better conform to human visual habits. At the same time, the method has a higher speed and requires no user interaction.

(3) We propose an upright orientation selection method based on viewpoint scoring. Upright orientation of 3D object is conducive to applications such as model alignment, and function recovery. However, state of art methods mainly focus only on manmade models, some of which still require user interaction. Therefore, an upright method suited for any kind of models is proposed in this paper. This method is based on such a phenomenon: the bottom areas of 3D objects, no matter manmade or natural objects, barely contain rich features, so viewpoints that focus on these areas will be given lower scores. Thus, according to the viewpoint scores, we can effectively find the bottom of the model, and use it to get the upright orientation of the model. To improve the reliability of the result, stability and observing habit metrics are further involved. Experimental results show that the new method can effectively handle a variety of models, even some manmade models that cannot be handled by existing methods.

(4) We propose an implicit measure of user attention in virtual environment navigation. The eye-tracker device is the main tool to obtain visual attention. However, it is hard to be applied in the 3D virtual scene navigation because of its poor reliability, great complexity and high price. We propose a new attention measure metric which can be easily embedded in the virtual environment. Behaviors of virtual camera controlled by the user are regarded as an implicit expression of visual attention. The implicit expression is parameterized by the observed object's distance to the center of the camera, occlusion, projected area, and observation time. A sophisticated user study experiment is designed to verify the reliability of this kind of measurement. The proposed implicit measurement of user attention can be effectively applied to an application of user attention aided game design.

Keywords: Viewpoint scoring, Mesh segmentation, Camera control, Perception, N best viewpoints, Information entropy, Visual attention, Virtual navigation, 3D scene design, Saliency

目 录

摘要	v
Abstract	vii
目录	xi
第一章 绪论	1
1.1 研究背景及意义	1
1.2 视点评分	6
1.3 模型摆正	7
1.4 物体关注度测量	8
1.5 本文内容及结构	9
第二章 相关工作	11
2.1 视点评分	11
2.2 模型分割	15
2.3 模型摆正	18
2.4 关注度测量	21
第三章 基于显著度分块的视点评分方法	25
3.1 基于网格显著度的模型分割	25
3.1.1 高效网格显著度计算	26
3.1.2 基于图的分割	28
3.2 使用显著度分块进行视点评分	30
3.3 实验结果及分析	32
3.3.1 和其他工作的对比实验	33
3.3.2 用户学习	35
3.4 小结及将来工作	36

第四章 结合视觉感知与信息量的视点评分方法	37
4.1 曲率计算	38
4.2 视点评分	40
4.3 生成代表性视点	43
4.4 实验结果及分析	46
4.4.1 效率分析	46
4.4.2 最优视点	47
4.4.3 最优视点集	48
4.4.4 局限性	51
4.5 小结及将来工作	53
第五章 基于视点评分的三维模型摆正	55
5.1 基本思路	55
5.1.1 使用视点评分获得候选底部	55
5.1.2 模型底部判定准则	57
5.2 方法实现	58
5.2.1 视点评分方法	59
5.2.2 底部候选集生成	60
5.2.3 摆正方向确定	61
5.3 实验结果及分析	64
5.3.1 适用性	64
5.3.2 对比实验	65
5.3.3 效率分析	66
5.4 小结及将来工作	68
5.5 附录	68
第六章 虚拟场景漫游中用户关注度的内隐测量方法	71
6.1 基本思路	71
6.2 有效性验证	74

6.2.1 建立元型场景	74
6.2.2 用户学习及分析	77
6.3 游戏场景辅助设计	77
6.4 小结及将来工作	80
第七章 总结	81
参考文献	83
发表文章目录	97
简历	99
致谢	101

表 格

4.1 部分模型计算时间统计表（时间单位：秒）	46
5.1 不同视点评分方法判断底部视点为最差视点的比例	56
5.2 在3组模型上的摆正准确率对比	66
5.3 本章对不同复杂度模型的计算时间统计（单位：秒）	67
5.4 不同方法的效率比较(本章方法实验所用CPU能力弱于其它方法 所用的CPU)	67
6.1 志愿者关注度统计信息	76
6.2 志愿者对不同场景设计方案关注度统计信息	79

插 图

1.1	三维内容的组织结构	2
1.2	3Dwarehouse模型网站中有大量三维模型	3
1.3	不同角度观察三维模型所获取的信息不同	4
1.4	模型摆正示意图	5
2.1	各种视点描述子的效果（图片来自文献[113]）	14
2.2	文献[96]中使用用户学习的方式训练视点选择模型的部分结果 (图片来自文献[96])	15
2.3	对称层次关系能够表示物体各个部件的结构及组织关系（图片来 自文献[116]）	17
2.4	文献[87]中的视点评分结果（图片来自文献[87]）	20
2.5	文献[49]对某些模型的摆正效果并不好（图片来自文献[49]）	21
2.6	文献[12]中，针对三维动作类游戏得到重要度映射的过程（图片 来自文献[12]）	22
2.7	文献[42]中提出的客观预测计算方法主要步骤，红色表示该文的 主要贡献（图片来自文献[42]）	23
3.1	显著度模型和人眼关注度的对应关系（图片来自文献 [58]）	26
3.2	本章方法的主要步骤，色温越高数值越高	27
3.3	使用不同方法进行的网格分割（第一行）以及对应的视点选择结 果（第二行）	29
3.4	Armadillo模型在文献[76]中的分割和视点选择结果，以及本章方 法的视点选择结果	31
3.5	Santa模型使用文献[76]的方法选择的最优视点和本章方法的比较 ..	31
3.6	使用不同评分公式选择的Bunny模型最优视点	32
3.7	使用本章方法为观察球候选视点评分的结果	33

3.8 文献[68]中的最优视点(a, b, c)和本章方法最优视点(d, e, f)的比较 ······	34
3.9 文献[96]中部分模型的最优视点和本章方法最优视点的比较 ······	34
3.10 使用本章方法得到的5个不同质量级别视点, 从左到右质量依次降低 ······	35
3.11 模型库中28个模型的逆序对平均值及平均差 ······	35
3.12 本章视点评分方法的失效情况 ······	36
4.1 不同视点评分标准得到的Cow模型最优视点 ······	38
4.2 使用不同邻域半径得到Lion模型的平均曲率分布图 ······	39
4.3 部分模型的平均曲率、平均曲率直方图均衡化, 以及人眼关注度[58]之间的相关性示意图 ······	40
4.4 对Igea模型计算平均曲率, 并进行曲率直方图均衡化的结果 ······	42
4.5 分别使用3个不同评分公式得到的Bimba模型最优视点及其平均曲率示意图 ······	43
4.6 模型的候选视点及其质量 ······	44
4.7 需要大量人工交互的state-of-arts方法[96]和本章方法得到的最优视点比较 ······	47
4.8 $\delta_{kl} = 0.02$, $\delta_r = 0.8$ 时, 数据库中模型最优视点集的视点数量分布情况 ······	48
4.9 使用本章方法得到的Horse模型的不同数量最优视点集获取的信息分布和模型信息分布之间的KL距离 ······	48
4.10 KL阈值=0.02, 可见面积比阈值=0.8时, 部分模型得到的最优视点集 ······	49
4.11 使用文献[63]方法和本章方法得到的最优视点集结果 ······	50
4.12 使用文献[68]和本章方法得到的HappyBuddha模型的最优视点集 ······	50
4.13 使用文献[96]的方法和本章方法得到的Lucy模型最优视点集 ······	51
4.14 使用本章方法得到的Fish模型的最优视点(左) 和次优视点(右) ······	52
4.15 本章方法将后视角错误地作为最优视角(a); 而更符合人眼观察习惯并且能获得更多有用信息的是(b) ······	52

5.1 不同种类模型分别使用4种视点评分方法得到的最优视点和最差视点.....	56
5.2 模型的规范视点示意图.....	58
5.3 本章对模型摆正的主要过程	59
5.4 视点合适度准则示意图.....	62
5.5 稳定性准则示意图	62
5.6 支撑性准则示意图	63
5.7 清真寺模型的候选视点和使用公式5.1得到的摆正结果	64
5.8 本章对PSB数据库中不同种类模型的平均摆正准确度.....	65
5.9 本章方法能摆正其他方法失效的模型	67
5.10 本章方法失效的模型	68
5.11 PSB模型库中部分不同种类模型分别使用4种视点评分方法得到的最优视点和最差视点.....	69
5.12 PSB模型库中部分不同种类模型的摆正结果.....	70
6.1 游戏中用户操作化身进行探索过程	72
6.2 物体被关注参数示意图.....	73
6.3 本章使用的三维场景	75
6.4 元型场景中的不同显著度级别物体	75
6.5 物体显著度和关注度测量值之间关系	77
6.6 游戏的概览图.....	78
6.7 Dream Land游戏两套场景物体设计方案	78

第一章 绪论

随着三维模型及场景获取手段的多样化，数字化三维内容资源变得越来越丰富，对这些三维内容的研究热点已经转变为如何更加有效地对其进行使用。本论文将三维模型及由三维模型组成的场景作为研究对象，主要研究如何对三维模型进行有效观察及预处理，并进一步将这种研究思路扩展到三维交互场景中，研究如何准确测量物体的关注度。接下来，我们将详细介绍本论文所涉及的研究点的研究背景及意义。

1.1 研究背景及意义

横看成岭侧成峰，远近高低各不同。

不识庐山真面目，只缘身在此山中。

这首名为《题西林壁》的诗出自我国北宋著名诗人苏轼，它描绘了诗人游览庐山后的感受。这首诗也常用来借景说理，指出世间的事物复杂多变，而人的认知视角是受限的，因此对同样一个物体，由于观察角度的区别，不同的人可能得出不同的结论。这也提醒人们注意，要认识事物的真相与全貌，就必须客观全面地来进行观察。

当今，我们处于信息大爆炸时代，而数字化是目前信息表达的最主要形式，将信息进行数字化处理极大推动了信息化技术的发展。在众多数字化信息内容当中，数字三维内容在近些年引起了极大的关注。这主要得益于以三维动画电影、三维游戏、科学可视化、虚拟现实技术等产业的发展。然而，目前用户操作和浏览三维内容的方式主要是使用鼠标键盘等二维手段，以此进行交互来控制虚拟摄像机浏览单个模型，或者在三维场景中进行漫游。三维场景信息最终也需要藉由二维显示器传达给用户。二维界面难以全面的反映三维的内容，这个矛盾增加了用户理解三维内容的难度，很容易产生“不识庐山真面目”的困惑。本文以解决这一矛盾为出发点，研究三维场景（包括三维模型以及由若干模型组成的交互式三维场景）内容和观察视点之间的关系，将这种关系量化后，应用于视点选择、模型摆正以及三维交互场景辅助设计当中。

如图1.1所示，数字三维内容可以表示成一个层次化的结构。组成三维场景的底层是点线面这样的基本几何单元，它们依靠附带的法线、纹理、材质等信

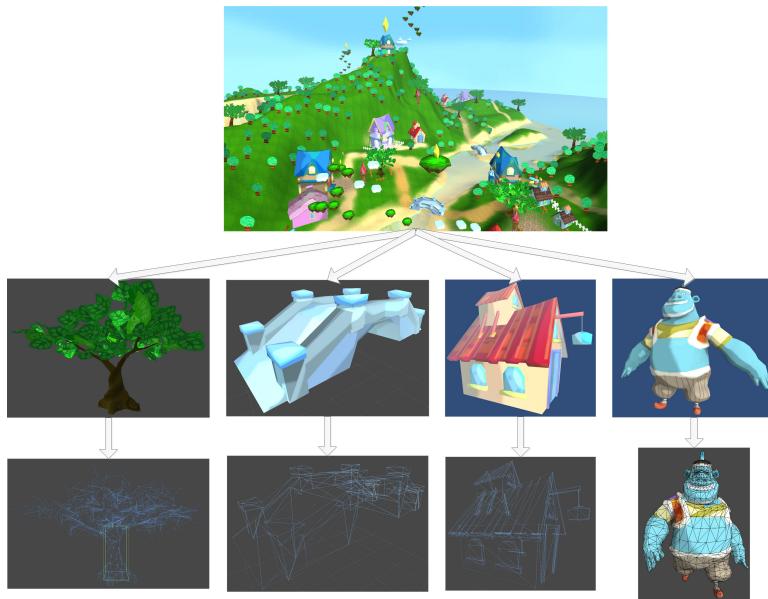


图 1.1: 三维内容的组织结构

息，可以描述三维模型的外观，动画电影及电子游戏等应用中常见的三维场景也是由这些模型经过设计师的组合来获得。本文以组成三维场景内容的层次结构为研究脉络，通过三维模型底层的点线面等几何信息（图1.1 的叶子结点），得到三维模型（图1.1 的第二层节点）表面所具备的特征信息，以此建立三维观察视角和所获取的模型信息之间的通道，从而对三维模型的观察质量进行量化分析。这种量化分析结果可以进一步被应用于数据库中三维模型的优化管理中。这种对观察视角进行量化分析的思路最终被推广到三维虚拟场景（图1.1中的根节点）的用户关注度测量中，通过分析用户在漫游过程中对虚拟摄像机的控制参数，获知用户对不同三维物体的观察意图，最终测量出用户对不同物体的关注度数值，该测量值可以被用于三维场景辅助设计等应用当中。

目前获取三维模型的途径很多，从面向专业三维艺术家的3ds Max^①及Maya^② 3D²，到更加易用的SketchUp³ 和Meshlab[22]之类的三维建模软件，再到将实体进行扫描转换为数字形式的三维扫描仪[94]，以及最近出现的使用普通摄像机或者手机进行扫描建模的软件等[61]。这些方法都极大提高了三维模型的获取效率，导致三维模型数量急剧增加。比如在三维模型搜索

¹<http://www.autodesk.com/products/3ds-max/overview/>

²<http://www.autodesk.com/products/maya/overview/>

³<http://www.sketchup.com/>

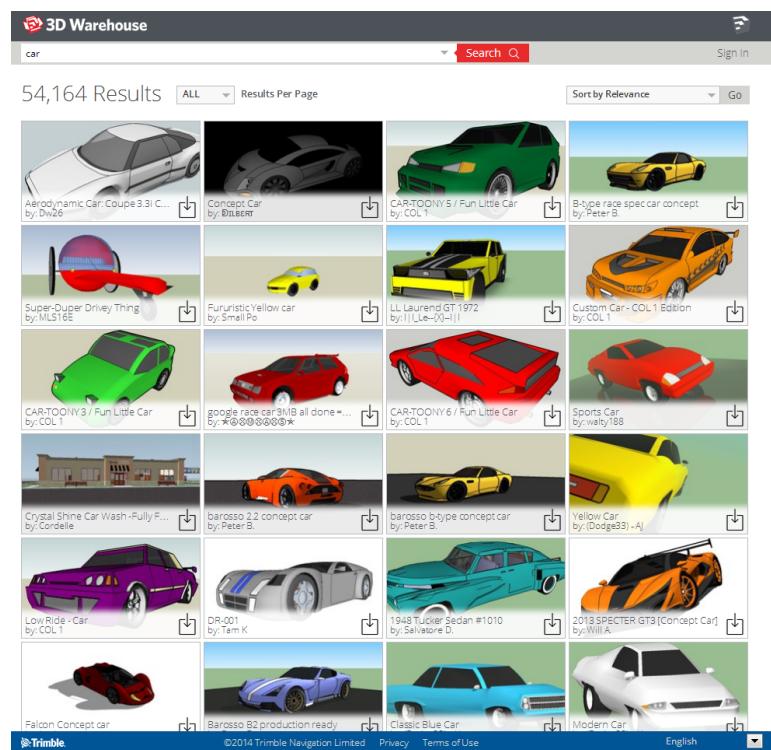
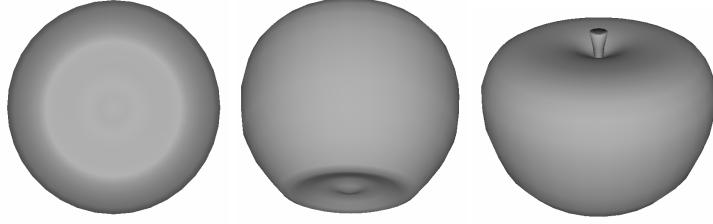


图 1.2: 3Dwarehouse模型网站中有大量三维模型

及共享网站3Dwarehouse中⁴，输入“model”关键字可以得到接近五十万个搜索结果，而普通的“car”关键字，则可以得到超过5万个搜索结果（如图1.2所示）。

大量三维模型素材的存在，可以让内容创作者更高效地进行三维场景的搭建。然而，目前我们观察及操作三维内容的方式仍然是二维的。比如在观察一个三维模型时，普通用户只能通过鼠标拖拽的方式进行交互，应用程序底层需要将这种在二维平面上的操作，采用某些方式映射为对三维空间的操作（比如采用Arcball方式[103]），转换三维空间的观察视角，进而让用户对三维模型进行多角度观察。观察结果也需要使用三维图形渲染流水线，通过扫描转换，将结果显示到二维屏幕上[110]。这种操作方式不是很直观，限制了很多初级用户的使用。因此，如何帮助用户快速选定便于认知三维模型的观察视角十分关键。图1.3展示了这种现象，用户从三个角度观察三维模型所获取的模型信息

⁴<https://3dwarehouse.sketchup.com/>



(a) 最差观察角度 (b) 较差观察角度 (c) 最优观察角度

图 1.3: 不同角度观察三维模型所获取的信息不同

并不相同。由于从图1.3c 中我们能够获取模型表面更加多样化的重要特征，更便于识别出这是一个“苹果”模型，因此它对应的视点质量最高，选择它作为这个三维模型的缩略图最合适。

为三维模型自动生成合适的代表视点对很多应用都十分关键，比如Gao 等在文献[32]中将这一思想应用于模型库索引中，文章将三维模型的最优二维渲染图像作为物体的代表，通过匹配请求物体的二维代表图像和数据库中其他物体的二维代表图像来完成检索。再比如图1.2所示的3Dwarehouse搜索结果中，系统将模型的代表图像作为缩略图展示给用户，从而方便用户快速浏览索引结果。这种缩略图一般是通过模型设计者上传得到的，这种手动上传方式效率较低，而且一些经验不足的设计者可能将不恰当的缩略图上传。比如图1.2中的第四行最后一个模型，缩略图是侧后方视角，据此缩略图只能了解到这大概是一辆具有黄色外观的小汽车，但车辆的造型，以及类似轮胎、车灯等关键部件的信息却无法从这个缩略图中获得。因此，本文将三维模型的自动视点评分作为主要研究方向，希望通过分析三维模型的几何数据得到其表面的特征信息，建立这些信息和观察视点之间的联系，据此对观察视点质量进行自动评价。

有了大量模型组成的模型数据库，对它们进行有效管理及再利用就变得十分迫切，近些年出现了较多这方面的工作。Jain等人在文献[47]中提出了一个形状混合系统，可以由数据库中已有形状进行混合得到新形状。他们提出的这套系统可以用来快速生成新模型，得到的模型和参考模型具有相似的对称性和连接结构。而文献[57]中关注了如何高效浏览模型库中形状的问题，作者采用模糊对应计算来寻找模型摆正，并允许用户指定感兴趣区域（ROI），以此来查看形状之间的异同。Kalogerakis等人在文献[50]中对基于部件的形状进行分析，得到部件之间关系的可能性分布，然后对其重新组合以得到新的结构。他们提



图 1.4: 模型摆正示意图

出的概率模型能表示形状部件属性之间关系的可能性，模型通过对数据库中的部件进行机器学习得到。与此应用相类似，Xu等人在文献[117]中由一定数量的初始模型作为输入，不断进行进化繁殖，最终得到大量新的模型。

为了使这些应用有效进行，一般都需要首先对数据库中的模型进行一定程度的预处理，比如将所有模型在坐标系中进行自动对齐。这是因为当前的三维模型获取手段很多，特别是很多非专业用户也能够参与模型的制作，导致了模型的存储方式不规范；此外，三维建模软件所使用的坐标系可能也有区别。这就导致模型库中的三维模型可能并不对齐（如图1.4 所示的Buddha模型），导致“远近高低各不同”的后果。因此，如何为不同种类的三维模型自动进行摆正计算也是本文主要的研究方向。为了使本文的摆正方法适用性更广，我们将前文提到的“视点评分”引入模型摆正当中，由于无论人造模型还是自然模型，底部视点一般为较差视点，因此，利用视点评分得到的较差视点可以将模型摆正方向的搜索范围大幅度缩小，同时避免了其他摆正方法存在的，只能适用于人造模型的问题。

三维场景设计是一门为展现虚拟环境、交互场所、故事情节、完成戏剧冲突，或者刻画人物的时空造型艺术。三维场景是动画角色表演的舞台，是假定性的空间，它在作品画面里大多处于第二视觉层次。成功的三维动画电影、游戏及虚拟现实程序都有较好的场景烘托，场景设计既要有高度的创造性，又要具有很强的艺术性。场景设计还要具有合理性，要依据策划方案需要，创造需要的构图与氛围。场景设计师通过将三维模型、地形、环境（天空体等）以及各种特效（比如光晕、雾等）在时空域进行组合来搭建三维场景。

正是由于三维模型数量的增加，将其用于搭建三维场景也变得更加容易。特别是随着游戏引擎技术的不断成熟，普通用户经过简单培训也能够利用这

些工具进行场景搭建。然而，三维场景搭建归根到底还是一个专业性较强的工作。搭建一个具备专业水平的三维场景除了需要场景师具有较好的审美能力之外，还需要具有较强的空间造型、色彩搭配、视觉印象概括等能力。特别是现在最常见的交互式虚拟场景（比如电子游戏），对设计人员的专业能力提出了更高的要求，他们需要以策划案及设计草图为依据进行场景搭建。由于前文所述的二维操作界面和三维内容之间的矛盾，在场景搭建过程中，很容易让设计人员陷入“只缘身在此山中”的困惑。因此，如何采用计算机辅助手段帮助设计人员提高工作效率，量化设计效果就变得十分关键。本文以此为研究目标，提出不使用硬件的用户关注度测量方法，可以辅助设计人员分析场景物体的各种设定是否能达到预期的设计目标。

综上，本文针对三维环境中的观察视角的质量及关注度进行量化分析的研究，无论对于三维模型浏览、索引、管理，还是三维场景设计，都具有重要的理论及应用价值。接下来，我们分别对本文研究密切相关的几个研究点的背景进行介绍。

1.2 视点评分

模型视点评分的目的是为了量化分析模型观察视点的质量，并据此选择评分最高的一个或若干个视点作为模型的代表视点，从而让用户能更好地获取模型信息。在Blanz等人的文献[15]以及Tarr等人的文献[108]中，都使用了心理学实验的方式，证明了三维模型不同视角对人认知物体有较大区别。这也说明我们可以藉由对模型观察视点质量评分的方式，获得认知效果最好的视点作为代表视点或者规范视点（canonical view）。这一研究可以被用于很多应用中，比如自动生成模型缩略图、模型索引、模型对齐，和三维场景的摄像机自动控制。

视点评分的研究涉及到计算机图形学、计算机视觉、认知心理学和信息论等多种学科。比如在图1.2中的网上三维模型库中，采用二维缩略图的方式可以较直观地向用户展示模型信息。从图中可以看出，绝大多数模型的缩略图质量较高，用户可以通过缩略图了解到模型的重要信息。但不好的缩略图会造成用户判断失误，影响模型库使用体验（如图1.2中第四行最后一个模型）。

然而，为三维模型进行自动视点评分是一个很有挑战性的研究课题，因为在衡量视点质量的时候需要考虑很多因素，而很多三维模型一般只包含几何信

息。此外视点质量很多时候还和具体的应用相关，而且由于视觉观察主观性很强，不同个体对视点质量评价可能存在明显的差异。

目前，对视点评分的研究主要是建立视点和其观察到的模型表面内容之间的信息通道，将其作为视点的描述子，描述子的量化结果作为视点得分。由于很多三维模型只包含几何信息，这些描述子只能通过它们来构建，比如模型的可见面积、轮廓线长度等。然而，人是在一个更高的抽象层次对物体进行认知的，因此，如何通过底层几何信息获取模型表面高层次的特征信息就变得十分关键。

本文将和人眼关注度密切相关的网格显著度作为模型分割依据，将带有显著度信息的分块作为计算模型表面重要度信息的基本单元，因此能够得到更符合人眼观察习惯的视点评分结果。此外，为了解决视点评分效率较低的问题，本文进一步将计算效率较高的平均曲率看做模型表面特征的测量依据，引入信息熵来对视点所见的模型特征进行量化，得到了较好的结果。

1.3 模型摆正

人们在观察物体时，被观察物体所处的姿态会极大地影响到人们的观察效果。一般情况下，物体都有一个或少数几个被大多数人认可的竖直摆正方向及对应的底部，物体底部向下摆正的姿态更符合人们的观察习惯，因而更容易被人们准确认知。这也解释为什么一些倒着或侧着拍摄的照片，需要在电脑中进行旋转摆正后才更容易被我们浏览。

其实，在三维模型管理方面也一直存在这样的问题。这是因为三维模型获取手段很多，比如通过三维扫描仪或者各种不同种类的建模软件，这使得模型在默认情况下并不一定都处于底部向下的摆正姿态。当我们在使用或浏览这些模型时，就需要通过手动旋转的方式将其进行摆正，这无疑极大影响了模型使用效率。因此近年来，关于三维模型摆正方向的自动计算得到了很大的关注。因为对模型进行自动摆正除了便于我们浏览及使用模型之外，还是模型数据库管理、分类及存储等诸多应用的重要基础[17][102]。但目前这方面的工作[31][48][69]主要是针对人造模型进行探讨，以有效利用人造模型几何规整、对称性强、具有明显底座等特点。而其它非人造模型，如人体、树木等，则缺乏相关的研究，因为它们不像人造模型那样有便于特征分析的几何特征。

我们在文中提出一种新的确定模型摆正方向的方法，利用视点评分方法找到主要观察表面内容比较贫乏区域的观察方向，再结合物体摆放的稳定性和人

眼观察习惯性进行优化计算来得到摆正方向。由于无论是人造模型还是自然形成的非人造模型，其底部往往是表面内容很贫乏的区域，满足底部视点是较差视点的规律。因此我们方法的适用对象将不仅限于其他相关工作所讨论的人造模型。另外，得益于一些视点评分方法的快速计算，我们方法在计算效率上也比已有摆正处理方法有明显的提高。

1.4 物体关注度测量

在三维交互式场景中（包括虚拟现实和电子游戏等应用），准确获知用户对不同物体的关注度信息十分重要，它可用于场景的加速绘制及辅助设计等诸多领域，现已成为计算机图形学研究的热点之一。获取关注度最直接的方法是利用眼动仪，这种主观测量方式能直接记录用户眼球运动的数据，通过分析眼动信息和屏幕元素的对应关系，预测用户对不同物体的关注度。然而，心理学研究表明，人脑认知会受用户所处环境、执行的观察任务等因素影响。这导致眼球的关注区域和实际上对人脑产生刺激的区域不一定完全对应。而且，使用眼动仪会造成用户的沉浸感变差，影响测量数据的真实性和有效性。此外，眼动仪的配置比较繁琐，特别是在基于物体的关注度测量中，需花费大量的工作将记录到的眼动数据映射到物体上。因此，很多研究者开始致力于利用客观计算的方法预测物体关注度，以视觉心理学等学科为研究的理论基础，将场景内容所固有的颜色、位置、运动以及虚拟摄像机的移动等客观底层属性看成外界对人眼的刺激，再考虑场景任务等高层属性，预测用户对不同物体的关注程度。然而，该方法仅是利用客观数据预测而非测量人眼的主观关注信息，其结果仍需要通过眼动仪等设备进行验证，大多普适性不高，并不能产生令人满意的预测结果。

通过观察，我们发现用户在三维虚拟场景漫游过程中，会将自身投射到虚拟世界中，通过操作虚拟世界中的化身主动对场景进行探索，用户的漫游过程就是用户自己对场景认知的自然表达过程。交互式虚拟场景可以通过记录虚拟摄像机的位置和朝向等信息来获取用户的漫游行为，进而对不同物体的观察行为进行量化。这一内隐的记录方式相比于使用眼动仪更能反映用户在漫游过程中的真实观察意图，进而能得到更准确的物体关注度。另一方面，本文采用测量，而非预测的方式获取用户主观关注度，保证得到的关注度信息可以真实反映不同个体的关注行为，避免了客观预测方法中预测结果难以符合不同个体关注习惯的问题。本文利用大量用户学习的方式来验证这一测量方法的有效性，

并将其应用于实际的游戏设计案例中。据我们所知，目前还没有一种类似于本文提出的用户关注度非硬件测量方法。

1.5 本文内容及结构

在介绍完本文工作的研究背景及主要研究方向之后，我们将紧接着在第二章中详细介绍和本文研究有关的工作，然后对我们在三维场景的观察度量及其应用方面的研究展开讨论。

在第三章中，我们将提出一种新的三维模型视点评分方法。该方法将和人眼感知密切相关的网格显著度作为模型分割及分块重要度的计算依据，并结合新的考虑分块可见度的视点评分公式对视点质量进行量化，能够获得符合人眼观察习惯的视点评分结果。

在第四章中，我们提出另外一种视点评分方法。第三章中提出的视点评分方法由于要计算网格显著度，导致计算效率不高，不适合于一些需要快速计算的场合。此外，这种方法只能得到符合人眼观察习惯的评分结果，而对于人眼据此获取的信息量化没有进行计算。因此，我们在第四章中将计算效率较高的模型表面曲率作为衡量模型表面重要度的依据，并结合信息论原理，对人眼观察到的模型表面信息进行量化。此外，这种方法还可以衡量多个观察视点组成的最优视点集获取的信息量，保证利用最少的视点来有效认知三维物体。

在第五章中，我们提出一种新的模型摆正策略，考虑人眼习惯于在规范视角而非模型底部来观察物体这一规律，利用视点评分方法找到观察视点质量最差的若干观察位置作为模型底部视点候选集合。结合底部稳定性、对称性和观察习惯准则，从候选底部集合中找到模型真正的底部视点，从而将模型进行自动摆正。和传统的摆正方法相比，第五章中提出的方法通用性更好，不局限于人造物体，而且计算速度快，无需人工干预。

在第六章中，我们提出一种三维场景中物体关注度的“软件”测量方法，用户在漫游场景过程中，程序记录物体被用户控制的虚拟摄像机关注的3个参数，据此可以在隐式条件下，真实记录用户对不同物体的关注程度。相比于传统的“硬件”方法，这种隐式的软件测量方法更容易获得真实的关注结果。并且，由于软件方法容易被嵌入到现有虚拟现实或者游戏程序当中，使其应用于实际项目中的难度更小。

最后，我们在第七章中对本文的主要工作进行总结。

第二章 相关工作

本论文主要研究如何对三维内容的观察进行度量，其中包括对三维模型的观察视点进行质量评分并将其应用于模型摆正当中，以及在三维交互场景中如何对场景物体进行关注度测量。本章将详细介绍和本论文这些研究点密切相关的工作，首先是三维模型的观察视点评分工作，将对这方面工作进行分类比较，并分析和本论文研究的关系；很多视点评分或其他的三维模型处理研究中都需要对模型进行结构化分析，即模型分割，本章也将对其展开介绍；对三维模型进行自动摆正是本论文视点评分研究的一个重要应用，它同时也是很多三维模型应用的重要预处理步骤，本章将对其相关工作进行介绍；最后本章将详细介绍关注度测量方面的相关工作，并着重分析针对三维场景的关注度测量。

2.1 视点评分

对三维模型的观察视点进行评分是很多应用的重要基础，比如模型向上方向确定[69]，三维场景中的摄像机自动控制[21][24][78]，以及三维模型搜索[32]等。比如在文献[18]中将视点评分应用于模型简化当中，这种基于信息度量的方法使用视点互信息(viewpoint mutual information)的概念，建立了从视点到场景的信道，视点互信息越小，则视点质量越好。通过这种方法可以找到较好的视点，然后利用这些视点作为衡量简化效果的依据。

早在上世纪80年代末，Kamada和Kawai就在工作[52]中探讨了三维物体视点的质量评价问题，他们将视点正交投影中模型退化面片的数量作为判断视点质量的依据，数目越少视点质量越高。Gómez等人在文献[35]中将这种思路扩展到了透视投影当中。

视点评分其实是将视点处可见的模型表面信息作为视点质量计算依据，形成以可见模型表面信息为依据的视点描述子。在视点选择应用当中，最优视点即为描述子处于最大极值点的情况。视点评分的研究集中于找到更加合理的视点描述子上面。

由于模型内容很多是通过几何来表达的，因此，许多方法通过度量一些几何特征进行视点评分。Plemenos和Benayada在文献[86]中将二维可视度准则

引入到视点评分当中，将模型未被遮挡的投影面积作为视点质量准则。而文献[27]认为模型轮廓包含了认知模型的重要信息，因此将其作为视点质量评判依据。有些研究者认为视点处观察到的模型表面深度变化越大表明模型信息越多样化，因此将可以观察到的深度信息作为视点质量计算依据[15][105][4]。Polonsky等人在文献[88]中介绍了多个视点描述子，比如投影面积、曲率、轮廓和拓扑复杂度，这些描述子最大化的结果可以获得不同的最优视点。Podolak等人在文献[87]中提出使用对称性的概念来标记模型表面的重要度，文章使用对称变换方式来定义两种几何属性：对称中心和主对称轴，并将其应用于模型对齐、分割和视点评分当中。在应用于视点评分时，作者主要考虑了模型对称性包含了冗余信息，所以，好的视点处于看到冗余信息少的位置，也即所见模型表面对称性小的位置。

在模型的几何信息当中，曲率在很多三维模型处理应用中都扮演了重要的角色，这是因为视觉重要度信息往往跟模型表面的曲率密切相关[8]。所以，很多视点选择算法直接将平均曲率或者高斯曲率作为物体表面特征量[88]，有些方法则使用了曲率在一定范围内的变化情况作为重要度计算依据[66]。

这些基于模型表面几何的视点评分方法计算简便，速度快，在很多实验中也取得了不错的成绩[96]。但由于它们主要依赖于一些基本的几何信息，对于模型内容的反映不是很好。很多研究者认为视点评分是一个和人眼感知密切相关的应用，选择的视点必须符合人眼视觉观察习惯。因此，他们期望从三维模型中得到视觉重要度信息，进而将视觉重要度作为衡量视点分数的标准[66][5][6][68]。文献[38]中通过显著度学习的方法，定义了更加复杂的模型表面区域区分度，以此为依据计算视点质量。Shilane和Funkhouser在文献[101]中认为计算模型表面区域的重要度不能孤立地使用几何特征，而应将其放入模型所处类别中进行考察，将那些能够将模型从其他种类中区别出来的区域设定为高重要度。文章采用基于形状搜索的方法，利用模型表面区域作为索引关键字，那些能够在同类中保持稳定并和其他分类不同的区域即是可区分区域，也即这个模型的高重要度区域。与它们类似的描述子还被用于生成三维场景的最优视点集应用当中，比如在文献[30]中使用基于模型表面可见度的描述子；而在文献[118]中，作者使用了基于显著度的视角稳定性概念。这些以度量模型表面视觉重要度为依据的视点评分方法的主要问题是计算效率不高。比如Lee等提出的网格显著度（Mesh saliency）算法[66]将网格模型顶点一定范围内的曲率变化作为计算依据，得到三维形状表面的视觉重要度信息，称为网格显著

度。他们将顶点显著度作为特征数值，并将可见元素的显著度总和作为视点评分依据。这种方法为具有 $2M$ 个顶点的模型计算显著度需要数小时[66]，导致很多应用只能将其放入预算阶段[96]。

很多研究工作引入了更加完备的信息理论来对视点获取的模型信息进行量化[82][88][112][4][98]。比如Vázquez等人在文献[111]中引入了香农熵理论，将模型面片作为信息基本单位，可见面片信息量大的视点质量较高。受这一工作启发，文献[95]和文献[6]中采用互信息理论在视点和模型表面之间建立信息通道来对视点质量进行量化分析。但这些工作仍然是将模型表面的面片作为信息测量依据，难以保证获取信息和人眼真实认知结果相匹配。而文献[88]中使用模型可见区域的高斯曲率熵作为视点描述子，这一定程度上可以解决只使用表面几何单元作为信息熵计算依据带来的问题。

还有一些研究引入了艺术和设计等学科的准则。Gooch等在文献[36]中将艺术美学的准则引入到视点评价当中，利用启发式的方法判断三维物体视点的质量。Blanz等在文献[15]中定义了若干影响人们对三维模型观察视点喜好的因素，他们得出结论，人们选择规范视角会受到经验、任务和模型几何的影响。文章使用实证手段来寻找人们对物体观察的规律，这种研究思路很有意义，因为视点评分归根到底需要由人来进行评判。这种将用户学习引入视点评分的策略也启发了后续很多的相关研究[113][5][96]。

Liu等人在文献[70]中提出一个极具创意的最优视点选择方法，作者将某种物体在互联网上的海量图片作为其最优视点的依据，文章的思路和用户学习方式类似，认为大多数人认可的拍摄角度能够作为物体的最优视角。这种方法直接从互联网图片中获取用户的视点喜好，比使用少量样本的用户学习方式普适性更好。

近些年，有些研究者考虑到人认识物体主要来源于物体表面的结构化信息[13]，所以在视点评分时应该以模型表面更加高层的结构化信息作为计算依据。有些工作引入语义信息来进行视点评分，文献[76]提出使用模型分块，并为分块赋予语义权重的方法来引导最优视点的选择。由于语义信息的加入，避免了单纯使用模型底层几何信息视点评分没有考虑人类感知习惯的问题[63][76]。然而，由于这些基于语义的方法对模型分块的语义重要度没有合理的赋值方式，导致视点评分结果的质量容易受到经验值的影响。

有些工作也将视点评分推广到三维场景当中，用于辅助摄像机路径的优化

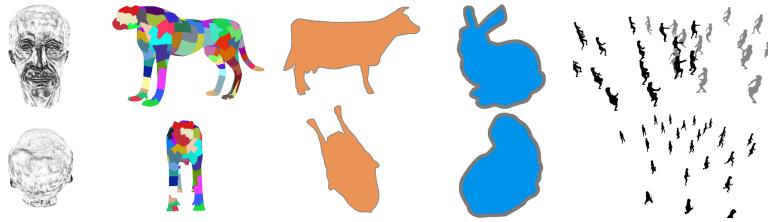


图 2.1: 各种视点描述子的效果 (图片来自文献[113])

自动生成。比如Sokolov等人在文献[104]中关注虚拟世界漫游中的摄像机自动控制问题，文章介绍了一种高层次的摄像机控制方法，首先计算能够最大化接收场景信息的若干好视点，然后将其作为关键点生成摄像机路径，最终定义场景物体之间语义距离的概念。

尽管利用语义信息或信息熵等方式可以极大提高最优视点和人类感知习惯的契合度，然而，视点质量高与否终归需要由人来衡量。所以，有些研究者将用户学习引入，通过学习大量志愿者的视点喜好来指导视点评分[113][5][96][70]。Vieira等人在文献[113]中详细分析了前人各种视点描述子的优劣，图2.1中给出了部分描述子的视点选择结果，图中的5个描述子分别是：平均曲率、可见分块、图像平面对齐、轮廓复杂度和动画领域的角色分组，第一行表示描述子值高的结果，而第二行是最低的视点，第一行更符合人们的认知结果。他们认为视点选择和具体应用及个体相关，前人的自动视点选择算法很难满足个性化需要。他们使用支持向量机（SVM）的机器学习方法得到个人对不同视点的喜好，利用学习结果指导其他视点选择。Secord等人在文献[96]中利用了下面5种视点描述子：(1) 面积，比如投影面积，模型表面可见度[86]和视点熵[112][98]；(2) 轮廓[27][113]；(3) 深度[15]；(4) 表面曲率[66][68][82][88][119]；(5) 语义[15][31][76][63]。考虑到他们单独使用不具有很高的鲁棒性，作者将这些自动视点选择技术结合来进行视点评分，利用用户学习的方法来得到它们的组合参数。然而，这些引入用户学习的方法质量严重依赖于所用样本的数量、种类，并且和具体应用密切相关。如图2.2所示，在a和f图中，用户对视点的偏好不明显，而其他图中用户都偏好右边的视点。第一行中的(a,b,c)模型不同视点能够被文中提出的训练模型所处理，而使用文章提出的视点评分模型却无法有效处理第二行中的不同视点(d,e,f)。加之这种方式计算复杂、投入巨大，并不能很好满足实际应用需要。

有些研究者认为只依靠一个视点很难全面地认知三维模型，需要在最

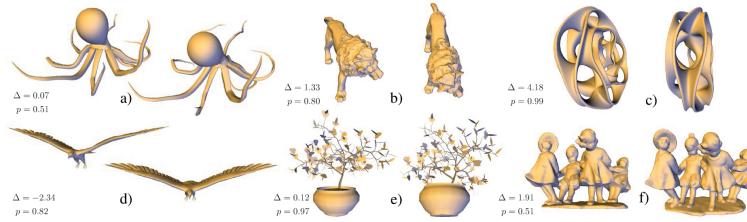


图 2.2: 文献[96]中使用用户学习的方式训练视点选择模型的部分结果 (图片来自文献[96])

优视点之外, 增加尽可能少的其他视点组成最优视点集 (N best viewpoints) [118][68][97]。然而, 如何对最优视点集中获取的信息进行量化分析是解决这类问题的关键。有些方法使用固定数目的视点[88][63]; 有些方法使用经验阈值, 比如文献[68]将最优视点集中可见顶点的显著度和模型表面显著度比值作为依据, 设定阈值为60%。本文将Kullback-Leibler距离 (相对熵) 引入, 用于度量最优视点集获取的信息分布和模型表面固有信息分布的差异程度。与以往利用KL距离衡量视点的方法不同, 我们基于不同级别的平均曲率来计算KL距离, 并且将结果用于最优视点集视点数量的有效控制上。

我们在论文中提出的“基于显著度分块的视点评分”方法属于基于语义方法的范畴, 然而, 在对模型表面进行分割和赋予每个分块重要度信息的时候, 我们采用了和人眼关注度密切相关的网格显著度属性, 从而解决了其他基于语义方法很难得到合理的分块重要度的问题。

但这种方法的主要问题是计算效率不高, 此外它还无法对所获取的信息进行量化分析。因此, 在文中, 我们提出了另外一种适用于需要快速计算场合的“结合视觉感与和信息量的视点评分方法”。这个方法将模型表面和视觉感知相关度较高, 且计算速度较快的平均曲率作为模型特征度量依据。并将信息熵引入视点质量评分公式中, 更容易准确度量人眼对模型表面特征的获取量, 还可以用于最优视点集的生成当中。

2.2 模型分割

关于三维模型的很多应用都需要首先抽取模型的高层结构信息, 比如模型编辑[122]、模型搜索[81], 以及上文中讨论的“视点评分”等应用。按照Shamir在文献[99]中的研究结果, 形状识别和基于语义的物体表达都依赖于

三维网格中的特征提取和结构提取。而三维模型分割是结构提取最为重要的一种方式，因此，在近些年来，针对三维网格模型的分割方法受到了很大的关注，很多文章也对这方面研究进行了详细综述[1][10][20][99]。针对三维模型的分割方法有全自动和交互式[123][25]两种，采取的策略又可以分为切分和聚类两种。

切分方法找到不同分块之间在某一层面上的分界线（比如几何或者功能差异），以此来对模型进行分割[34]。Shapira等在文献[100]中使用形状直径函数（shape diameter function）来对模型进行分割，而文献[10]使用了基于平面和法线方向的方法。在分水岭算法中[71][62]，根据输入三角网格曲面每个顶点的曲率信息，将局部最小值作为网格曲面的初始分割；然后，开始自下而上或者自上而下地合并分水岭高度低于指定阈值的区域；最后处理过分割问题。基于拓扑信息的网格分割算法根据网格几何及拓扑信息，将网格分割为有视觉意义或者物理意义的组件[39]。Bischoff等人在文献[14]中将三维形状使用粗糙层次的包围椭球几何进行表示，将三维模型表面映射到不同椭球上的方式对模型进行分割。Reniers等人在文献[90]中提出一种使用曲线骨架对网格模型进行分割的方法，可以自动判断或者手动分配骨架关键点，然后使用这些点作为分割依据，结合网格表面测地线对模型进行分割。由于采用测地线方法，文章可以得到平滑的分割边界。另外，由于骨架具有层次特点，所以文章方法可以对网格模型进行层次化划分。

而聚类方法将具有相似属性且位置临近的模型表面元素聚合为一类，得到模型分块结果[100]。比如Katz等在文献[53]中提出基于模糊聚类的层次分解方法，层次树的根表示整个网格模型，每个节点表示网格模型分片。算法可以对任意拓扑连接或无拓扑连接的、可定向的网格进行处理，避免了过分割和边界锯齿。由于考虑测地距离和凸性，使分割边界通过凹度最深的区域，从而能够得到更有意义的分割结果。

这些网格分割方法普遍需要分析模型表面的几何信息或拓扑关系，这些采用底层几何信息的方法容易处理表面比较规整的模型。和这些方法不同，我们在本文中需要得到和视觉关注相关的模型分块结果，这就需要引入网格显著度这种模型表面和视觉关注密切相关的参数来作为分割依据。

除此之外，还有一些方法使用用户学习结果来引导模型分割，比如在文献[51]中，提出了一种数据驱动的模型分割和标记的方法。文章构造一个目标

函数作为条件随机场模型，包括两个数据项：第一个评估标签面片的一致性；第二个反映相邻面片标签的关联性，目标函数从标记网格集合中学习得到。文中的算法使用成百上千的几何和语义标签，从中学习不同风格的模型分割方法，而不需要用户调整参数。普林斯顿模型分割Benchmark中的实验表明，文中的算法比当前最好的方法提高很多，分割以及标记结果和人为计算结果一致性较高。然而，这类方法的主要问题是算法复杂、需要大量的用户交互、要保证样本模型具有尽可能广泛的多样性，并且要求所选用的训练模型适用于特定的场合。

很多三维网格分割或评估算法都来自于二维图像处理领域。比如三维分割领域很常用的random walk分割方法[64]借用了二维图像分割中的思想[37]；而Chen等将二维图像分割中的评价标准[72]引入到三维模型分割的评价当中[20]。本文所使用的依据网格显著度对三维模型进行聚类分块的方法思想也借鉴自二维图像处理领域的基于图的分割方法[28]，我们将其进行改进，能够被用于三维网格表面分割中。

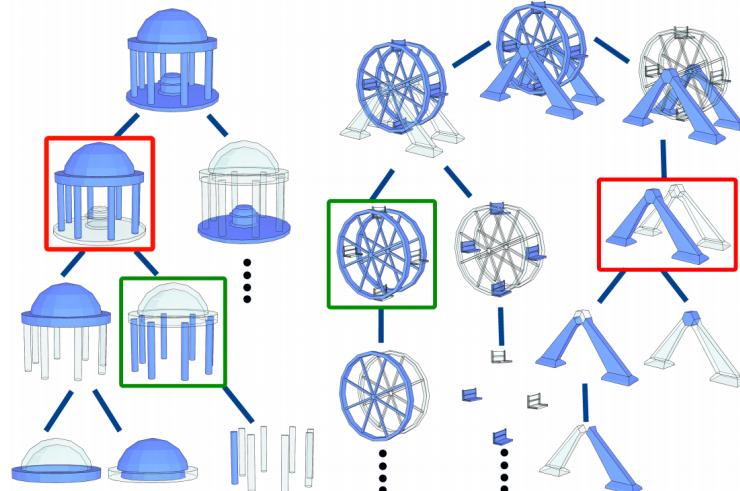


图 2.3: 对称层次关系能够表示物体各个部件的结构及组织关系（图片来自文献[116]）

近些年，随着模型库规模的提升，对模型管理及再利用的需求旺盛。在此应用的驱动下，很多研究者开始寻找在更高层次进行模型分割的方法，比如Wang等在文献[116]中提出了一种人造模型的对称层次图表示方法，它是三维模型的高层结构化表示，能够考虑模型部件的对称关系并表达为层次化结构。对于输入的网格模型，首先使用分割算法将其分割为部件形式，对部件关

系进行分析后得到所谓的初始图，这个图能够表示部件之间的对称和连接关系，也能够得到单个部件的自对称信息。利用初始图进行递归的节点合并可以得到对称层次图，合并过程是对称部件分组和连接相邻部件（如图2.3所示，图中每个节点代表的部件用蓝色表示，每个节点或者表示对称组（绿色框）或者表示组装关系（红色框））。对称层次图能够得到比单独使用局部几何准则方式更加有意义的层次化分割结果，用这种分割结果进行模型视点选择和模型摆正等操作可以得到更加符合人们认知习惯的结果。然而，这种方法仅适用于人造模型，并且计算复杂，目前还很难直接使用。

我们在“基于显著度分块的视点评分”中，将模型表面和人眼关注度相关的网格显著度信息作为分割依据，将相邻面片中显著度数值接近的部分聚类到一起，最终形成具备不同显著度的模型分块，并以此作为衡量视点优劣的基本单元，能够得到更加符合人眼观察习惯的视点评分结果。

2.3 模型摆正

目前许多关于模型摆正方向计算的方法，主要都基于主元分析法(principal component analysis，简称PCA)。这是因为主元分析法能找到致密包围模型的长方体包围盒，且一般情况下，这样的长方体能反映模型摆放的姿态。在PCA分析的基础上，一些方法进一步根据稳定性、对称性等准则来计算摆正方向[17][102][114][120]，在此，它们主要是以长方体包围盒的6个面的法向作为候选方向进行筛选，从中得到摆正方向。但这样的计算方法并不鲁棒，比如对于倾斜姿态的模型就难以得到摆正方向。为此，一些工作分析模型的几何特征与其摆正方向的关联性，以此作为依据来计算模型的摆正方向。比如，Fu等人就提出通过机器学习得到模型形状与摆放姿态之间的关联性[31]。但这样基于机器学习的方法依赖于所选择的学习样本，适应性不高；在他们的工作中，也主要是针对人造模型进行处理。

近年来，一些工作探讨如何通过几何特征的度量来提高模型摆正方向计算的效率。Chaouch等人在文献[19]中观察到模型在其摆正方向上一般具有很好的对称性，因此，在PCA的基础上，提出通过模型的反射对称性及局部平移对称性来计算摆正方向。但该方法对于对称性不明显的模型不太适应。文献[69]则提出对模型进行简化并计算其凸包，以方便利用几何稳定性快速得到候选的摆正方向，再根据基于几何、物理和可见性情况的评估函数来确定摆正方向。Jin等人在文献[48]中认为，模型沿着其摆正方向的投影可以看作一个低秩矩阵，

因此，可在PCA的基础上通过稳定性、对称性、可见性等的综合考虑得到摆正方向。和Jin等人的思路类似，Wang等人在最近的文献[115]中将三维模型对坐标轴的对齐问题看作低秩张量的优化问题，当找到6个候选摆正方向之后，作者也引入各种几何属性从中筛选出模型的底部方向，进而对模型进行摆正。姜玻等在文献[7]中则对模型上各个面片的朝向、模型的对称性及模型凸包上各个面片的朝向进行统计和综合分析，再根据静力学平衡原理、模型的可见性等准则进行度量，以得到摆正方向。综上，这些基于几何特征度量的方法，对于几何特征与摆正方向的关联性比较强的情况是比较适合的，因此，它们在处理人造模型时有较好的效果，这也是它们所限定讨论的问题。

本文工作根据视点评分方法以及模型摆放稳定性和人眼观察习惯性来计算摆正方向。虽然我们也基于一些几何特征的度量，但对几何特征的依赖性不像已有方法那样直接，而是能方便利用一些语义信息，如模型底部的表面内容比较贫乏、人眼一般习惯从一个规范视点（canonical view）来观察模型[15]等，以提高计算的准确性。因此，我们也能方便地处理非人造模型。由于一些视点选择方法计算速度快，不必像已有计算摆正方法那样要进行几何特征分割等复杂的处理，我们的方法也有很好的计算效率，并为实验所验证。

我们方法在使用视点评分时，不是要找到能观察到模型更多内容的好视点，而是寻找主要观察到表面内容贫乏的视点，也即差视点。而内容贫乏的表面区域，一般几何特征平淡，因此，我们可使用只需度量几何特征进行视点评分的方法，而不必使用基于语义信息的视点评分方法。这样，我们的方法也就有很好的适应性，便于处理各类模型。我们注意到有些视点选择的工作，是先进行模型摆正方向的确定，以提高好视点寻找的效率[69]。这与我们工作的方向正好相逆。我们还没有看到基于视点选择来计算摆正方向的工作。

目前已经有一些研究者注意到模型观察视点质量和物体跟摄像机的相对位置有关，比如在文献[15]中，作者认为较好的视点出现在模型斜上方位置。这个现象也可以从图2.4的第二列观察到，利用文献[87]自动找到的最优视点都处于模型斜上方的位置。实际上，绝大多数视点选择算法都是利用各种评分公式，将处于这些位置的最优视点自动找到。然而，对最差视点的规律却没有进行深入研究。我们观察到，和物体最优视点具有一定规律性类似，物体的最差视点往往出现于物体的底部或者顶部位置，这个现象也能从图2.4的最后一列中看出，除Bunny模型外，其他模型的最差视点都位于模型的底部或者顶部位

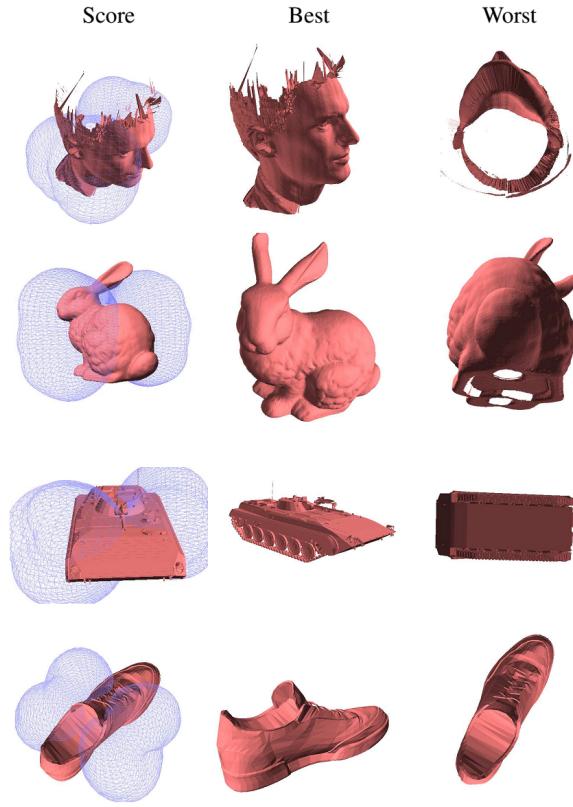


图 2.4: 文献[87]中的视点评分结果 (图片来自文献[87])

置。另外一个示例可以参考图2.1的Maxplanck模型和Bunny模型，利用平均曲率和轮廓复杂度作为视点描述符，得到的这两个模型的最差视点分别位于模型的顶部和底部区域。这种现象的存在就为模型摆正提供了一种新的思路，即将评分较差的位置作为模型底部的候选位置，然后通过筛选策略从中找到模型的真正底部来完成模型摆正。文献[49]中即利用了这一思想，作者将模型投影面积作为视点质量描述子，利用和最优视点选择相反的思路，基于最小投影面积找到能够对齐模型的3个主坐标轴。第一个主轴是平行投影面积最小的那个方向；第二个主轴垂直于第一个，并且投影面积最小；第三个主轴是前两个轴的叉乘结果。但在一些情况下，模型沿着摆正方向的投影面积却比较大，比如牛、马等四肢动物模型，因此，该方法的适应性也有限（如图2.5所示）。和文献[49]的思路类似，我们提出的模型摆正方法也基于视点观察质量，但和这篇文章不同的是，我们使用更加多样化的视点描述子来得到模型的候选底部集合，接着使用多个判定准则从中找到模型的真正底部，这使得我们方法的适用

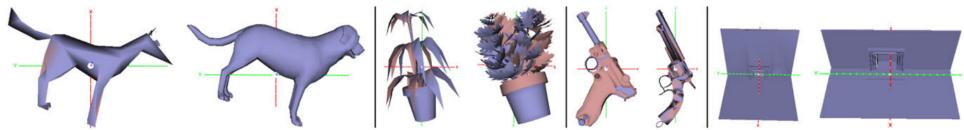


图 2.5: 文献[49]对某些模型的摆正效果并不好 (图片来自文献[49])

面更广，鲁棒性更高。

2.4 关注度测量

近十几年来，由于互联网广告、电子游戏、虚拟现实等应用的发展，针对人眼关注度的研究获得了极高的重视。获取到的人眼关注度信息可以帮助设计人员优化画面配置，将重要信息放置于用户关注度高的区域[9]；在三维环境中将用户不关注的物体进行简化，以提高场景渲染速度[67]；或者，利用人眼对物体运动关注的特点，实现特效绘制的加速[121]。在文献[73]中讨论了利用人眼关注度特点，如何使用尽可能少的模型和动画，来在虚拟现实程序中制作人群。文章通过实验发现，以下的元素会影响到用户对于重复的判断：外貌（皮肤、衣服纹理等）、动画（包括动画的种类、相同动画是否同步等）。外貌克隆会更容易被辨认出来，所以在制作人群效果的时候，要尽可能多一些外貌不同的模型。并且文章得出了定量化的数据，可以指导场景设计人员在制作场景时，外貌克隆的数目和动画克隆的数目应如何控制，以便用户在一定时间内无法察觉。

眼动仪[75]可以较准确地记录人的眼动轨迹，这为人眼关注度计算带来极大帮助。但是大多数眼动设备都需要穿戴到用户身上[40]，造成用户沉浸感变差，影响获取数据的准确性；而无需穿戴的设备往往价格昂贵。此外，眼动设备只能捕捉眼球在时域空间中的注视位置，而将这些位置信息映射回动态场景中仍然需要大量的处理工作[12]。更为重要的是，有研究表明，当观察者集中注意于某个物体或事件时，经常不能觉察到其他出现于视野内的意外物体或事件，这种现象被称为不注意视盲（inattentional blindness）[56][77][2]。因此，即便利用眼动仪设备获知用户看到了某些物体，也无法保证这些物体受到了用户的关注[60]。综上，眼动仪这种主观测量方式很难直接获取用户在虚拟漫游场景中对不同物体的准确关注度。

很多研究者试图通过分析画面内容所具备的分布特点，结合人眼关注规

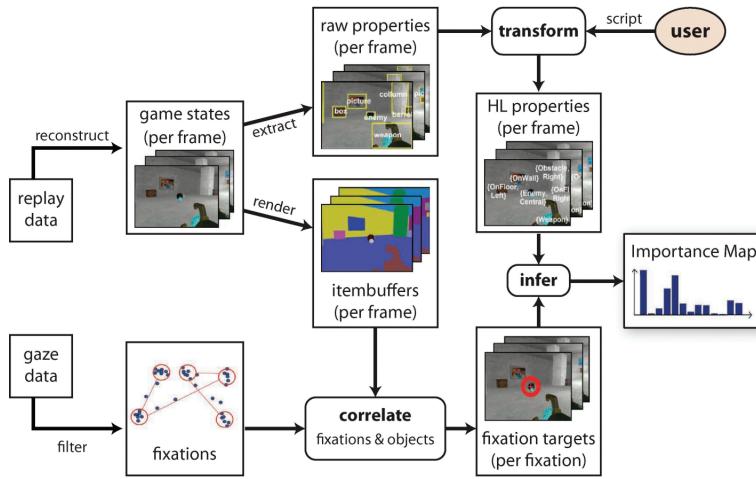


图 2.6: 文献[12]中, 针对三维动作类游戏得到重要度映射的过程 (图片来自文献[12])

律, 推测人们对不同区域的关注度大小。这种在没有眼动仪参与条件下获取人眼关注度的技术称为关注度预测, 是一种客观预测方法。在这里, 眼动仪只用于在用户学习阶段训练预测函数, 或者最后评估预测算法的性能[85][12][42]。关注度预测技术可以大致分为自底向上和自顶向下两种[43]。前者认为人们在无意识情况下对一些底层特征表现得敏感; 而后者指有导向的自我关注控制, 涉及到了高层次的过程如思考、推理和记忆。

大多数视觉关注预测方法均依赖于底层的特征, 因此只能采用自底向上的关注方法。第一个使用自底向上的生物启发式计算模型是由Koch和Ullman在文献[59]中提出的, 并在文献[46][45][83]中得到了进一步改进。该方法是创建一些底层特征的特征图(特别是颜色、强度、方向之间的对比)组合成一张显著图, 每个像素均为一个标量值(评分值)。在视频或者电脑游戏等非静态场合, 可以通过引入时域特性提高显著度计算准确性[44][85]。

然而, 在三维虚拟场景中, 需要计算关注度的往往是组成场景的物体, 而非屏幕像素, 所以需要找到屏幕显著度到物体显著度的映射关系[12]。再者, 自底向上的算法在自由视角的情况下可靠性最大, 而视频、三维虚拟场景中很少有自由视角的行为[107], 即便处于被动的观看情况下, 观察者也会假想一个任务或意图, 其注视目标会更集中于任务相关的物体上[92][84]。为此, 需要将更高层次的属性(比如任务)引入, 使用自顶向下的方法提高虚拟场景关注度预测的准确性[67][12][42][106]。Bernhard等人在文献[12]中认为由于三维透视和视点的多样性, 再加上引入任务以后带来的影响关注度分布的参数, 所以在交

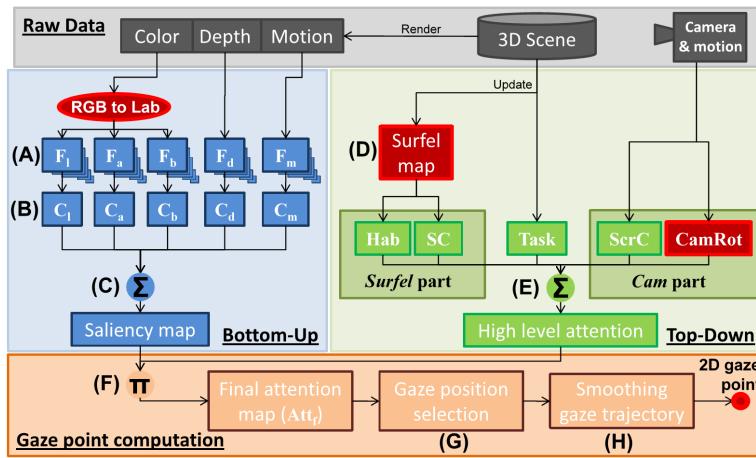


图 2.7: 文献[42]中提出的客观预测计算方法主要步骤, 红色表示该文的主要贡献 (图片来自文献[42])

互虚拟环境中的人眼关注分析和预测是一个很有挑战性的话题。文章展示了一个用于学习眼动数据的流水线, 流水线的最终结果是重要度映射, 可以记录每个物体上的关注度 (过程如图2.6所示)。这一重要度映射被用做启发式数据, 来实时预测用户对不同物体属性的视觉关注度。本文方法的创新点是所有的分析在物体空间中进行, 而重要度映射在高层属性的特征空间中进行。高层属性用来编码任务相关和个性度等属性。文章将这个流水线在第一人称射击游戏中进行了测试, 结果表明, 使用重要度映射得到的关注度信息和实际关注度之间关联性极强, 特别是玩家被游戏内任务吸引时。

在文献[42]中, 作者将自底向上和自顶向下的计算相结合, 在计算得到的场景显著度 (自底向上) 的基础上, 考虑任务、摄像机运动等高层属性 (自顶向下), 将二者结合来预测用户的视觉关注点。具体计算如图2.7所示, 其中主要计算元素有: 特征图 (A), 显眼图 (B), 自底向上的关注度 (显著度图) (C), 更新逐面元 (surfel) 数据 (D), 自顶向下关注度 (E), 计算屏幕空间的最终关注度 (F), 计算下一个可能关注点 (G), 对最终关注路径进行平滑滤波 (H)。

然而, 无论采用自底向上的方法, 还是自顶向下的方法, 计算得到的显著度都是在一定置信区间内对用户实际关注度的预测。其预测的效果及应用场合受限于所使用的计算策略, 需要使用眼动仪记录大量用户数据进一步检验。很多数据驱动的方法还依赖于前期的用户学习结果[12], 从而导致预测结果对学

习样本的规模和多样性敏感。

虚拟漫游场景和其他视觉关注的研究对象（比如图像、视频等）最大的不同是：在虚拟漫游场景中，用户可以通过控制化身来主动寻找自己感兴趣的物体，并通过虚拟摄像机渲染画面对物体进行观察[21][16]。因此，用户自由控制摄影机在虚拟场景中的位置、朝向和视野等参数的行为，就隐含了用户对场景中物体的观测意图。以此为依据，本文通过记录不同物体在虚拟摄像机中的观测参数来获取用户对不同物体的内隐关注度信息。

第三章 基于显著度分块的视点评分方法

合理的模型分割能够得到较高层次的模型结构信息，而人眼对物体的认知主要来自于对物体结构信息的认知。视点评分作为一个和人眼认知习惯密切相关的应用，应该将不同视觉重要度的模型分割结果作为计算依据。因此，如何为模型分割选择分割依据，以及如何为每个分块赋予合理的视觉重要度权重就成为解决视点评分问题的关键。已经有一些研究者使用模型分割为基础，通过为每个分块赋予语义权重的方式进行视点自动评分[76]，但在这些方法中，语义权重如何合理赋予一直没有一个令人信服的方法。

我们将网格显著度和网格结构化信息相结合，提出了显著度分块的概念，并将其应用到视点选择中。如图3.1所示，第一行和第二行分别是平均曲率和网格显著度，第三行是利用眼动仪获取的志愿者关注位置，从中可以看出人眼关注度和网格显著度具有较强的对应关系[58]。而对模型进行分解有助于得到三维模型的高层次全局信息[99][11]。我们将模型表面具有相似显著度的相邻元素进行聚类，可以将模型表面分解为具有不同显著度的分块集合。在此基础上，我们提出一个新的视点评分公式，为那些能看到更多高显著度分块的视点赋予更高分数。为了解决网格显著度计算较慢的问题，并提高网格显著度计算结果和人眼感知的一致度，我们将原始模型简化为非冗余表达后才进行后续计算。这种带有显著度的模型分块既能真实反映每个分块的视觉重要度，又具备三维物体表面的结构化信息，用它进行视点评分更容易得到合理的结果。

本章的主要步骤如图3.2所示。我们首先利用视觉感知作为指导，对原始模型进行简化，得到非冗余网格（图3.2a）。接着利用文献[66]的方法计算模型显著度，并对结果进行拉普拉斯平滑（图3.2b）。将模型表面位置相邻且显著度接近的视点进行聚类，以此对模型表面进行分块，每个分块的显著度是其成员显著度的平均值（图3.2c）。最终，我们使用带有显著度数值的分块作为视点观察单元，依据本章提出的视点评分公式对模型观察视点进行评分（图3.2d）。

3.1 基于网格显著度的模型分割

早在1987年，Biederman就提出了“通过部件识别理论”(Recognition by Component Theory)[13]，他认为可以通过结构分解的方式，来有效认知复杂物

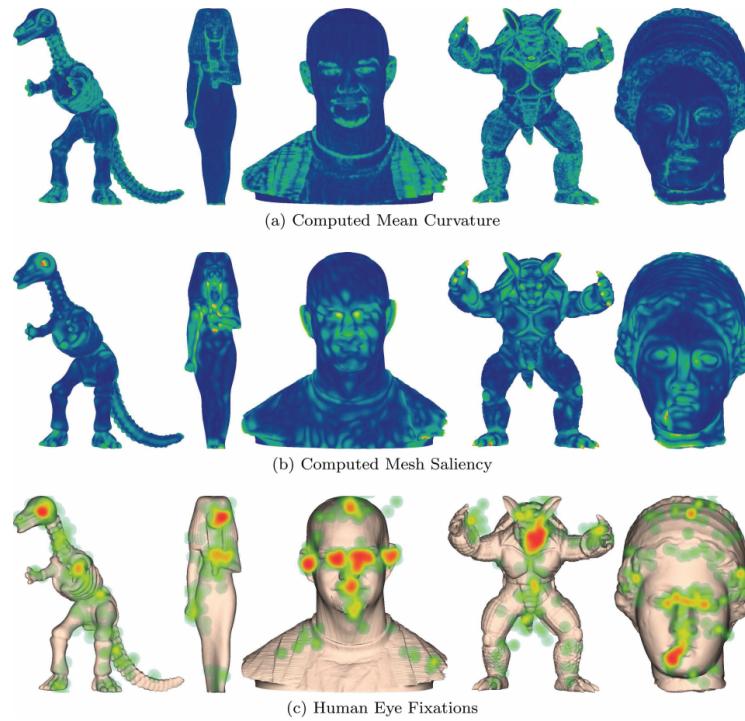


图 3.1: 显著度模型和人眼关注度的对应关系 (图片来自文献 [58])

体。以此理论为基础，很多基于网格分割的视点评分方法更有可能得到可信的结果[88][76]。然而，这些方法在进行分割时，使用底层的几何信息（比如折叠角、曲率等）作为分割依据，而且每个分块的视觉重要度采用经验值的方式赋予[76]，这导致视点选择结果很难完全符合人们认知物体的习惯。有鉴于此，我们提出采用网格显著度[66]作为分割依据，因为这个模型表面属性和人眼关注度相关性极高[58]。我们将带有显著度的分块结果作为接下来的视点评分依据，结合提出的视点评分方法，可以获得更加可信的视点评分结果。

我们的视点评分方法有效的前提是，三维物体的内容可以通过其表面的显著度分块得到体现。为了增加计算的有效性，我们采用下面两个策略：1.将原始模型进行简化，得到没有冗余信息的网格模型；2.将网格表面顶点显著度相近且相邻的顶点聚为一类。接下来，我们将对这两个步骤进行详细讨论。

3.1.1 高效网格显著度计算

我们发现人们在认知复杂物体时，物体表面的高频几何细节并不能被人眼所察觉，这些细节对认知过程并无贡献。而且，网格显著度计算的复杂性和网

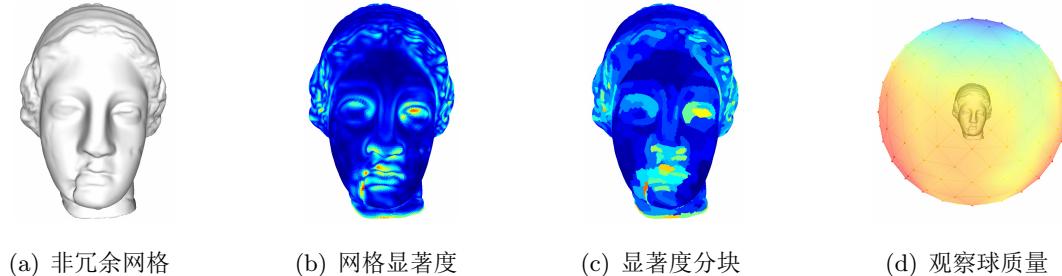


图 3.2: 本章方法的主要步骤, 色温越高数值越高

格模型分辨率正相关, 保留这些对认知并无贡献的几何细节将降低计算效率。我们通过精确控制的网格简化方法来将这些细节去除, 得到经过简化的不包含冗余信息的网格模型, 接下来的网格重要度计算将在这个非冗余模型上进行。相比于使用原始模型, 计算准确度和效率都将得到提升。

得到非冗余网格的关键是如何控制简化误差阈值, 即人眼恰好感觉不到差别的网格表面距离。我们将这个误差阈值 (E_m) 用于控制简化过程, 简化误差超过这个阈值则停止简化。

这里, 我们使用对比敏感度函数 (Contrast Sensitivity Function, 简称CSF) [55]来进行 E_m 的计算。依据文献[89]的结论, 人眼对小于60周/度 (cycles per degree, 简称CDP)的细节将无法察觉。因而我们可以将小于60CDP的细节删除而不会引起人们的视觉感知。在屏幕空间, 这个数值对应的长度约为1/120cm, 我们将其记为 E_s 。物体空间的误差阈值 E_m 可以通过下面的公式计算得到:

$$E_m = E_s * \frac{d}{d_n}$$

其中 d 和 d_n 分别表示视点到三维物体的距离以及到近裁剪面的距离。视点位于物体的观察球上, 朝向物体中心。

我们采用了Quadric Error Metrics (QEM)方法[33]进行网格简化, 这种方法速度快, 简化误差可控, 适合我们的应用场合。QEM方法依次选择开销最小的边, 将其两个端点进行合并, 并将合并后的顶点位置进行优化。合并操作一直进行, 直到合并误差超过一定阈值。这种方法中的合并误差定义为新生成的顶点和其相关平面之间距离的平方。为了保证简化不引起视觉注意, 我们将阈值设定为 E_m^2 , 保证原始网格简化为一个理想的级别 (称为非冗余网格), 不引起注意的几何细节已经被移除。接下来的网格显著度计算和视点评分计算都

使用这个非冗余网格来完成。我们采用文献[65]中提出的衡量两个网格之间视觉误差的方法Mesh Structural Distortion Measure (MSDM2)来进行实验，依靠我们这种误差控制策略，得到的非冗余网格和原始网格的MSDM2误差平均值为0.15，最大值不超过0.2，这表明这种误差控制策略引起的视觉变化是可以接受的。

在得到的非冗余网格上，我们使用文献[66]中提出的方法来计算网格显著度。这种方法使用人眼视觉系统具有center-surround机制的特点，将模型顶点及其周围一定范围内的曲率变化作为衡量显著度的依据。顶点显著度用下面的公式表示：

$$\varphi = \sum_i S(\varphi_i)$$

其中 S 是非线性的抑制操作符； φ_i 是不同尺度 i 下的显著度，其值依靠下面的公式得到：

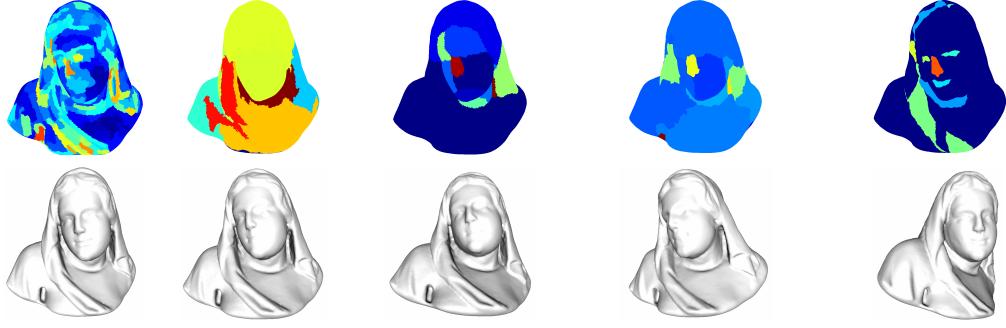
$$\varphi_i(v) = |G(\varepsilon(v), \sigma_i) - G(\varepsilon(v), 2\sigma_i)|$$

其中 G 表示对顶点 v 的平均曲率($\varepsilon(v)$)进行高斯加权求平均， σ_i 表示高斯滤波器在尺度 i 下的标准差，具体计算过程可以参考文献[66]。本章使用非冗余网格进行显著度计算，计算效率比使用原始网格提高6到7倍。

3.1.2 基于图的分割

目前有相当数量的考虑视觉感知的网格分割方法，Yamauchi等人在文献[118]中使用区域生长的技术来将网格中的顶点迭代地进行合并，最终得到网格分块。而Yang等人在文献[119]中同样采用区域生长的方法来进行模型分割，并在多尺度下计算模型最优视点。然而，这些分割方法的计算依据是物体表面的曲率这样的几何信息，并不能保证分割结果符合人眼观察习惯。

为得到和视觉重要度相关的模型分割结果，我们将网格显著度[66]作为分割准则，并引入已经在二维图像处理中得到成功应用的基于图的分割方法[28]来对三维网格进行划分。之所以选择这种分割方法进行我们的视点评分，是因为它具有以下2个针对性的优势：其中之一是它的分割过程考虑了分组或者区域的视觉重要度，另外一个是它的高效性。为了进一步消除噪声对模型分块的影响，提高分块的稳定性，在进行分割之前，我们将计算得到的模型显著度进行拉普拉斯平滑[29]。经过平滑以后，顶点显著度的一致性更高，有利于接下来的分块计算。



(a) 本章方法 (b) FitPrim[10] (c) RandCuts[34] (d) ShapeDiam[100] (e) Benchmark[20]

图 3.3: 使用不同方法进行的网格分割（第一行）以及对应的视点选择结果（第二行）

在三维网格模型上使用基于图的分割算法的主要步骤如算法1所示，其中 V 代表网格中的面片， E 代表面片之间的邻接关系。这里，我们将边权重 $w(v_i, v_j)$ 表示为相邻面片 i 和 j 之间的显著度差的绝对值，其中面片的显著度是其顶点显著度的平均值。参数 K 表示相邻分块之间显著度差异的阈值， R 用于避免过小的分块产生。

Algorithm 1: 对网格进行显著度分割的伪代码

Input: A mesh's topology graph $G(V, E)$ with n nodes and m edges
Output: A set of patches $S = \{P_1, \dots, P_r\}$

- 1 Sort E into $\pi = (o_1, \dots, o_m)$ by non-decreasing edge weight;
- 2 $S = \{P_1, \dots, P_n\}$, where $P_1 = \{v_1\}, \dots, P_n = \{v_n\}$;
- 3 Set threshold T of each patch to K ;
- 4 **for** $i = 1$ to m **do**
 - 5 $o_i = (v_a, v_b), (v_a \in P_a \text{ and } v_b \in P_b)$;
 - 6 **if** $P_a \neq P_b$ and $(w(v_a, v_b) < T_a \text{ and } T_b)$ **then**
 - 7 merge P_b to P_a , update $T_a = w(v_a, v_b) + K / \text{sizeof}(P_a)$;
- 8 **for** $i = 1$ to m **do**
 - 9 $o_i = (v_a, v_b), (v_a \in P_a \text{ and } v_b \in P_b)$;
 - 10 **if** $\text{sizeof}(P_a) < R \text{ or } \text{sizeof}(P_b) < R$ **then**
 - 11 merge P_a and P_b ;

分割计算完成之后，在同一分块中的面片比处于不同分块的面片在显著度上更相似，每个分块的显著度依靠其成员面片的显著度求平均得到。在算

法实现的过程当中，我们将参数 K 设置为网格中最大显著度和最小显著度之差的0.4%，而将 R 设为 G 中所有节点数目的0.05%。

从图3.3的对比中可以看出，相比于曲率及其他几何属性，网格显著度更适合作为分块依据，利用网格显著度得到的分块和人眼感知习惯相符度更高。图3.3(b)~(e)是文献[76]中使用不同模型分割方法得到的结果，我们将这些分块的重要度使用其面片重要度求平均来表示，而在文献[76]中采用的是经验值。从图中可以看出，很多重要的模型特征在这些基于几何的分割结果中被丢失了（比如在(b)(c)(d)中的眼睛、嘴和腋窝等，以及在(e)中的头发），而使用我们的分割方法则能够较好地保留这些高重要度的特征（图3.3(a)）。因此，使用我们的显著度分块作为视点评价依据得到的最优视点（图3.3(f)）相比于其他方法得到的最优视点（图3.3(g)~(j)）更加接近于规范视点[15]。

3.2 使用显著度分块进行视点评分

由于显著度分割结果相比其他分割方法更能真实反映模型的视觉感知相关特征，因此将其作为基本单元参与视点评分的话更容易获得接近人眼观察习惯的视点评分结果。接下来，我们首先讨论当前利用模型分块进行视点评分的方法中存在的主要问题，然后介绍如何使用显著度分块来构造评分公式。

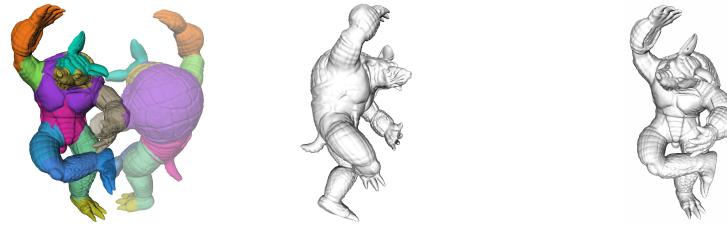
Mortara等人在文献[76]中引入基于语义分块的视点评分公式，他们认为不同类型的分块具有不同的语义重要度，视点的评分公式是考虑这些具备不同语义重要度分块可见性的结果：

$$Score(w) = N(w) \cdot \sum_c V(c, w)R(c)W(c) \quad (3.1)$$

其中 N 表示在视点 w 处可见的分块数目； c 是某个可见分块； V 是分块的可见度，使用投影面积和分块总面积的比例来表示； R 是分块面积占整个模型面积的比例； W 是依靠分块类型通过经验赋予的语义重要度值。

相比于只使用模型表面几何信息的传统视点评分方法，这种将语义分块作为计算单元的方法不容易受到局部高重要度信息的影响，更符合人眼通过结构化信息认知物体的规律。

然而，文献[76]中使用的公式3.1存在的主要问题是，其中的语义重要度 W 是依据所使用的分割算法以及分块的类型手动赋予的经验值，经验值不恰当就容易影响最终的视点评分效果。比如在文献[76]中，如果使用Fitting



(a) FP分割结果 (b) 文献[76]的最优视点 (c) 本章方法得到的最优视点

图 3.4: Armadillo模型在文献[76]中的分割和视点选择结果, 以及本章方法的视点选择结果



图 3.5: Santa模型使用文献[76]的方法选择的最优视点和本章方法的比较

Primitives视点评分方法, 平面形的分块重要度最低, 而球形、圆柱形和圆锥形的分块重要度相等, 这就很难反映这三种高重要度分块的差别。这种情况请参考图3.4中对Armadillo模型的处理, 在文献[76]中将其使用FP方法[10]分解为12个分块(图3.4(a)), 平面形分块并没有出现在分割结果当中, 所以其分块的语义重要度都相同, 导致重要度失去意义, 进而选择了并不符合人们习惯的侧下视角(图3.4(b)), 而本章得到的最优视点(图3.4(c))则更加接近规范视点[15], 一些类似眼睛等重要特征可以被观测到。

在图3.5中展示了另外一个例子表明通过经验值赋予语义重要度权重的问题, 文献[76]中使用Plumber分割方法对Santa模型进行分割, 并将管状分块赋予较低的权重, 而将主体分块赋予更高的权重(图3.5(a))。这种权重赋值方法并没有考虑模型表面更加多样化的特征(比如面部细节等), 因此, 得到的从后往前的最优观察视角, 明显不符合人们的观察习惯。而本章方法可以得到更加自然的前向观察视角(图3.5(b)), 这是因为通过显著度分割可以将模型表面的多样化特征区别开来(比如面部、胡须等高重要度特征在显著度分块中都得到了体现, 如图3.5(c))。

为了解决模型分块重要度很难有效赋值的问题，我们将上一小节中得到的显著度分割结果作为视点评分单元，将评分公式修改如下：

$$Score(w) = N(w) \cdot \sum_c V(c, w)R(c)S(c) \quad (3.2)$$

其中S表示分块的显著度，其他参数和公式3.2相同。在实现过程中，我们将模型表面的显著度归一化到区间 $[\epsilon, 1 + \epsilon]$ ，其中 ϵ 是远小于1的一个正数，这样可以保证显著度最低显著度的分块仍然具备一定的影响力。

然而，通过观察公式3.2我们可以发现，将分块面积直接作为权重并不合适，比如观察人体模型时，眼睛部分相对于身体面积很小，但其语义重要度要远大于其他分块。此外，可见分块数量直接出现在评分公式中也不合适，因为将可见分块的信息进行加权和已经隐含了可见分块数量信息。考虑到这两个因素，我们将公式3.2修改为如下形式：

$$Score(w) = \sum_c V(c, w)S(c) \quad (3.3)$$



图 3.6: 使用不同评分公式选择的Bunny模型最优视点

从图3.6中可以看出不同评分公式的区别，使用公式3.2更容易受到少数面积较大的高显著度分块的影响，而使用公式3.3则更容易获取全局最优的结果。图中利用公式3.2得到的Bunny模型最优视点是从斜后下方的观察视点（图3.6(a)），并不符合人眼观察习惯。而使用公式3.3则可以得到更接近于规范视点的结果（图3.6(b)）。

3.3 实验结果及分析

我们在一台具备Core i7-860 CPU (2.80GHz, 8CPUs), 12GB RAM和NVIDIA GeForce GTX 460 GPU (1GB RAM)的PC上进行了若干实验。为了衡量我们方法相比于其他工作的优势，我们将本章方法得到的结果和其他最先进的方法进行了详细比较，并使用一个用户学习实验来表明本章方法的有效性。

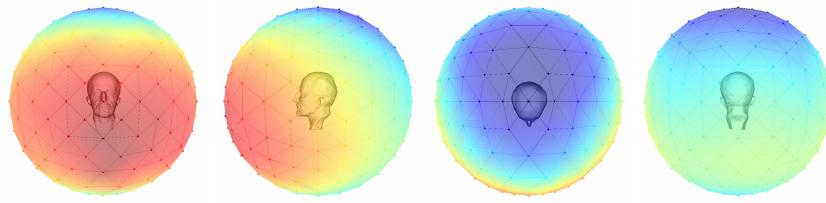


图 3.7: 使用本章方法为观察球候选视点评分的结果

3.3.1 和其他工作的对比实验

我们在每个测试模型的观察球上平均分布了258个候选视点，然后基于显著度分块，依靠公式3.3对这些视点进行评分。图3.7中给出了使用这种方法得到的MaxPlanck模型观察球质量分布图，从图中可以看出，较好的视点普遍分布于模型的前部，而较差的视点集中于模型后部及顶部，这和我们日常的观察习惯相符，说明本章方法的视点评分结果具有较好的合理性。

我们将本章得到的视点选择结果和文献[76]中引入的语义驱动方法进行了详细比较，从图3.3、3.4和3.6中可以看出，我们得到的最优视点更加符合人眼观察习惯，这主要得益于下面两个因素：

1. 使用显著度分块作为视点评分单元，分块已经具备了合理的视觉相关重要度；
2. 对文献[76]中的评分公式进行了合理的优化。

我们也和其他两个近期最先进的方法进行了比较。分别是Leifman在文献[68]中提出的利用模型表面观察兴趣度数值进行的视点评分，以及Secord 等人在文献[96]中使用的引入用户学习的方法。

图3.8给出了我们方法的结果和文献[68]的比较，从图中可以看出，在考虑对称性的情况下，二者对Dinosaur和Motorcycle模型得到的最优视点很接近。而Lion模型的最优视点有较大差别，文献[68]中的最优视角是正前方偏侧面（图3.8(b)），而我们的最优视点处于斜上方（图3.8(e)），更接近规范视角，从这个视角更容易获取Lion模型的身体及尾巴长度等信息。在另外一个目前最先进的文献[96]中，作者在由16个模型组成的库中，为每个模型选择20对视点（一共得到3840个渲染图像）。要求志愿者从每对图像中选择自己喜好的视点。他们分析了30到40个有效答卷，从中得到每个模型可以被大多数人认可的最优

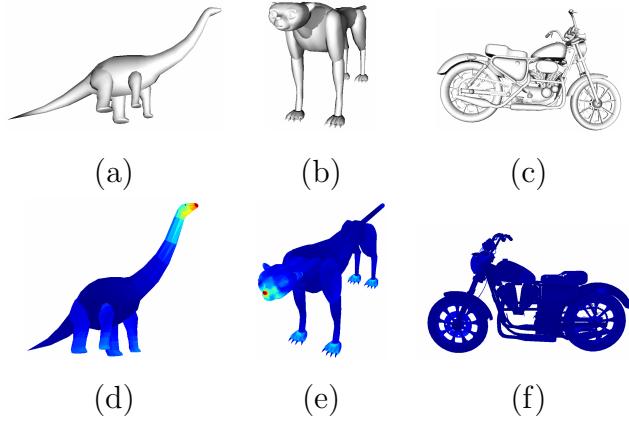


图 3.8: 文献[68]中的最优视点(a, b, c)和本章方法最优视点(d, e, f)的比较

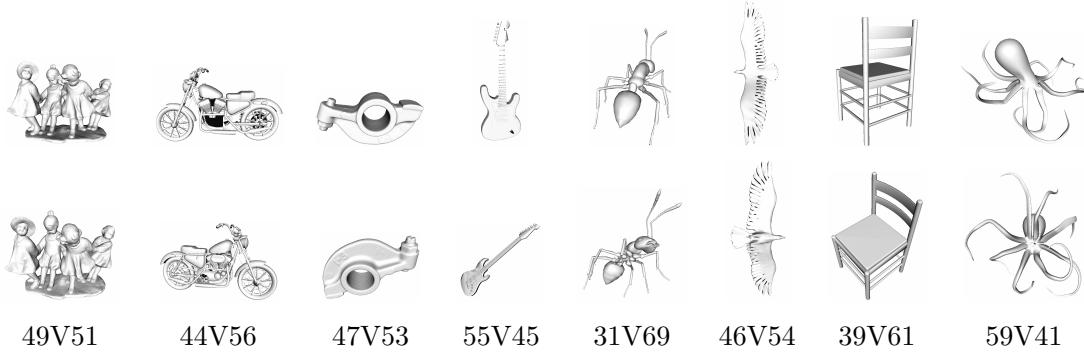
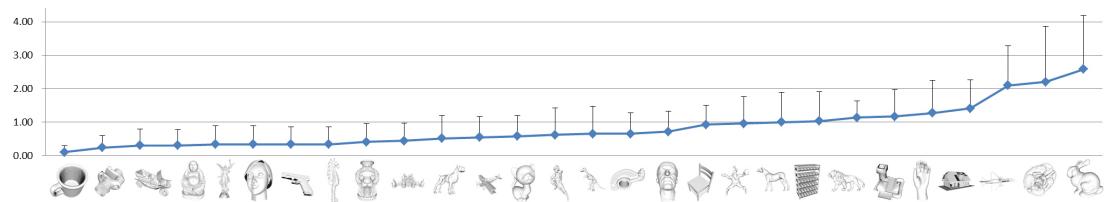


图 3.9: 文献[96]中部分模型的最优视点和本章方法最优视点的比较

视点（图3.9中给出了和本章方法进行比较的部分结果）。我们使用强制选择配对比较法（forced choice pair-wise comparison）来判断用户对文献[96]中的视点选择结果和本章方法的喜好程度。我们邀请了34名志愿者参与评判，图3.9中给出了用户选择结果，其中图片下方的两个数字表示该视点在我们的用户学习实验中的喜好程度百分比。考虑对称性的情况下，对于半数左右的数模型，本章方法自动产生的最优视点和文献[96]中使用大量用户交互选择得到的最优视点十分接近，因而用户对他们的喜好程度也比较接近（图3.9(a)(b)(c)(d)）。然而，其他模型的最优视点差别较大，因此用户对他们的喜好程度也较容易产生分歧（图3.9(e)(f)(g)(h)）。比如其中的Ant, Eagle和Chair模型，在文献[96]中将背后视点作为最优视点，这并符合人们的观察习惯，因而在我们的用户学习实验中这些视点相比于本章方法的最优视点得分较差。而对于Octopus模型，由

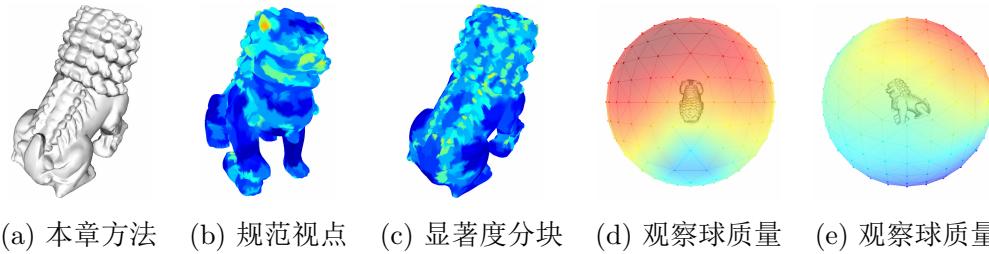


图 3.10: 使用本章方法得到的5个不同质量级别视点, 从左到右质量依次降低



越小，证明用户排序结果和本章视点评分的排序结果越接近。我们搜集了29份有效答卷，逆序对统计结果见图3.11（其中的缩略图是本章方法得到的每个模型的最优视点）。从图中可以看出，绝大部分模型的平均逆序对小于1，这表明我们方法对模型视点质量的识别和人眼的观察习惯基本相符。其中只有3个模型的逆序对大于2，它们是Fighter, Brain和Bunny，这可能和这些模型的表面特征分布有关，从不同视点观察这些模型得到的表面信息区别度不大。

3.4 小结及将来工作



(a) 本章方法 (b) 规范视点 (c) 显著度分块 (d) 观察球质量 (e) 观察球质量

图 3.12: 本章视点评分方法的失效情况

由于网格显著度和人眼关注度之间具有很强的对应关系，据此作为表面聚类条件得到的模型分块能够更好地反映模型表面不同视觉重要度的结构分布情况。本章将显著度分块作为视点评分的基本计算单元，解决了以往类似工作中很难有效获取分块重要度的问题。因此，本章提出的这种视点评分方法能够在只使用模型表面几何信息的条件下，得到基于语义方法的视点评分结果。

然而，本章对模型表面的分割以及分块重要度的计算依赖于网格显著度，如果显著度数值并不能很好反映模型表面的视觉重要度信息，则会导致得到的视点评分结果不符合人们的观察习惯。如图3.12所示，本章得到的ChinaDog模型最优视点是背向的(a)，而前向斜上方的规范视点(b)更加符合人们的观察习惯，(c)是本章最优视点处的显著度分块分布情况，(d)和(e)是观察球的质量分布。从图中可以看出，由于高显著分块大量分布于ChinaDog模型的后背，导致质量高的视点也普遍分布于观察球的后半部分，因而得到了从后向前的最优视点，一些重要的信息（比如脸部）无法观测到。这可能需要进一步探索将其他视觉感知相关的参数引入进来，比如最近关于模型表面兴趣区域的研究成果[68]。

第四章 结合视觉感知与信息量的视点评分方法

上一章中介绍的“基于显著度分块的视点评分方法”需要计算模型表面的网格显著度，效率较低，并不适用于需要快速计算的场合。此外，这个方法的视点评分公式主要受高显著度分块影响，可能忽略其他视觉重要度低但语义重要度高的区域，比如某些高细节的褶皱区域视觉重要度很高，却并不一定对认知物体起到十分重要的作用（参见图3.12）。

为了解决这两个问题，我们将测量几何表面粗糙程度的平均曲率作为特征信息度量值，并引入信息熵来衡量视点所获取的模型表面特征的数量。平均曲率计算速度较快，和网格显著度有很强的对应关系，特别适合于视觉感知相关的应用；另一方面，在视点评分公式中引入信息熵理论有助于量化视点所观察的模型信息。

在传统的采用信息熵作为视点评判依据的方法中[111][112][3][4][97][95][6]，由于包含信息量最大的视点是那些信息分布较为均衡的视点，因而得到的视点往往比较平淡。而直接使用视点可见区域重要度作为描述子又容易受到局部高重要度区域的影响。有鉴于此，本章的视点评分公式，在可见元素重要度之和的基础上，添加信息熵，用以调节局部高重要度所造成的人工因素，并避免产生过于平淡的最优视点。图4.1对比了文献[66]和本章方法得到的最优视点。其中，图4.1a和图4.1b中的第二个图分别是网格显著度和平均曲率热度图，颜色热度（红色到蓝色）代表重要度大小。图4.1a直接使用可见区域显著度之和作为视点质量判断依据，得到的最优视点关注了Cow模型肚子部位的高显著度区域，却失去了眼睛等重要特征。图4.1b是利用本章方法得到的最优视点。它既捕捉到了用于识别Cow模型牛角、牛乳、眼睛等主要特征，也更加符合人们的观察习惯。

接下来，我们将首先说明本章所使用的平均曲率计算方法及调整策略；接着讨论如何使用可见面片平均曲率结合信息熵的方法对观察视点进行评分；然后，介绍将这种评分方法如何应用于最优视点集的生成中；最后分析实验结果并得出结论。

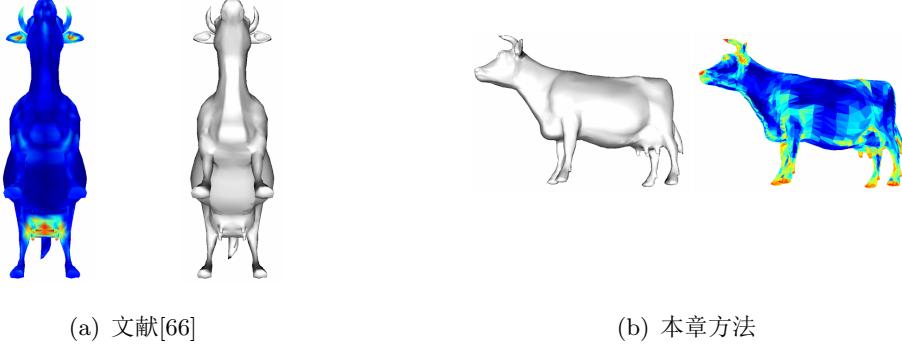


图 4.1: 不同视点评分标准得到的Cow模型最优视点

4.1 曲率计算

三维物体的局部结构特征信息依赖于曲率，更准确的说是平均曲率，因为平均曲率这种张量场可以很好地表现三维模型的视觉特征[8]。为了对物体表面的特征信息进行量化，本章将平均曲率作为顶点重要度依据，参与后面的视点评分运算。

目前存在很多计算曲率的办法（比如Normal Cycle算法[23]），然而其中多数算法结果严重依赖于模型表面的网格分辨率，这很大程度归因于这些算法都采用1环邻域的方法[93]。我们更希望得到三维物体表面较大尺度的外观特征，将那些在一定范围内顶点法线变化大的区域标记为高曲率，反之，标记周围顶点法线变化不大的区域为低曲率。

所以我们采用了Alliez等提出的曲率计算方法[8]。为了在网格模型表面得到一个连续的张量场，这种方法通过估算每个顶点处的曲率张量，来建立分段线性曲率张量场，然后在三角形上对这些值进行线性插值。一种自然的曲率张量可以通过沿着边的每个顶点来定义。顶点 v 处的曲率张量计算方法如下：

$$T(v) = \frac{1}{|B|} \sum_e \beta(e) |e \cap B| e_n e_n^t$$

其中 $|B|$ 表示待估算的顶点 v 周围区域的面积， $\beta(e)$ 表示 B 内的边 e 的两个邻接三角形的法线间的有符号夹角（凸面是正，凹面是负）， $|e \cap B|$ 表示 $e \cap B$ 的长度（介于0 和 $|e|$ 之间）， e_n 表示边 e 的单位方向向量。顶点 v 的相邻区域 B 可以通过以 v 为圆心，测地线距离 r 为半径来调整。张量矩阵的两个较大特征

值 k_{max} 和 k_{min} 分别是顶点 v 的两个主曲率，平均曲率就是二者的平均值。

这种曲率计算的一个重要特征是通过调整邻域半径 γ ，可以得到不同尺度下的曲率。如图4.2所示（其中三个图从左至右使用的半径 γ 分别是0.2%、0.6%、1%乘模型包围盒最长边），使用较小的半径可以得到更多更小的细节，而较大的半径得到的效果类似于对区域进行平滑[65]。本章采用了半径 $\gamma = 0.6\% \times B_{max}$ （ B_{max} 表示模型包围盒最长边的长度），因为它既能得到大多数中等尺度的表面特征，又可以保证较高的计算效率。

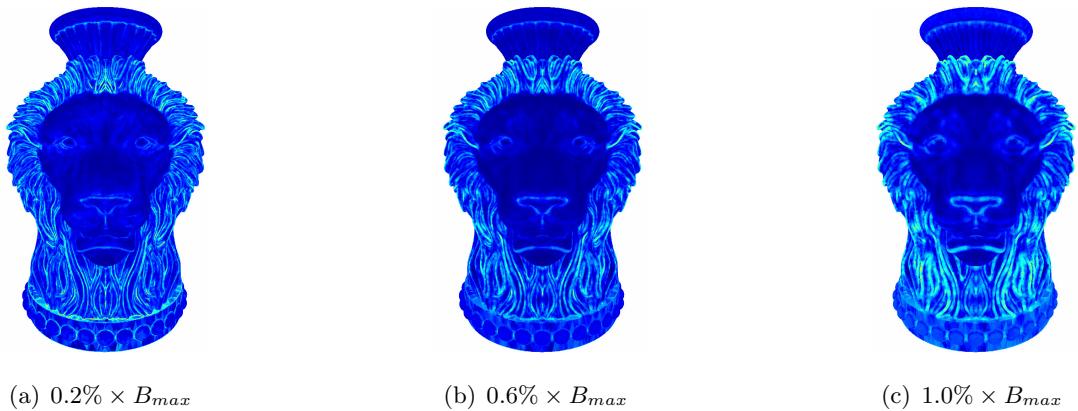


图 4.2: 使用不同邻域半径得到Lion模型的平均曲率分布图

由于使用文献[8]计算得到的曲率可以捕获顶点周围一定范围内的法线变化情况，而人眼对于特征变化的高频部分更具敏感性，故而得到采用上述方法得到的平均曲率可以反映出视觉重要度特征。这可以从图4.3中得到明显体现，其中a列为Armadillo、Dinosaur、Igea模型的平均曲率，b列是进行直方图均衡化后的结果，c列是文献[58]中实验得到的人眼关注度。计算得到的平均曲率(图4.3a)和实验得到的视觉关注度(图4.3c)具有很强的对应关系，比如Armadillo模型的脸部、腹部及四肢，Dinosaur模型的眼睛、鼻子、脖子、四肢和尾巴，以及Igea 模型的五官、左脸凿痕和头发部位。

为了避免模型表面噪声的干扰和网格上局部大曲率区域(对应于尖锐特征)的存在，我们对求得的平均曲率进行均衡化处理，以使曲率直方图近似为均匀分布。从图4.3中可以看出经过直方图均衡化以后(图4.3b)，平均曲率分布和视觉关注度(图4.3c)的对应关系进一步得到了加强。

为了保证均衡化过程能够更加真实反映模型表面曲率分布，我们将直方图

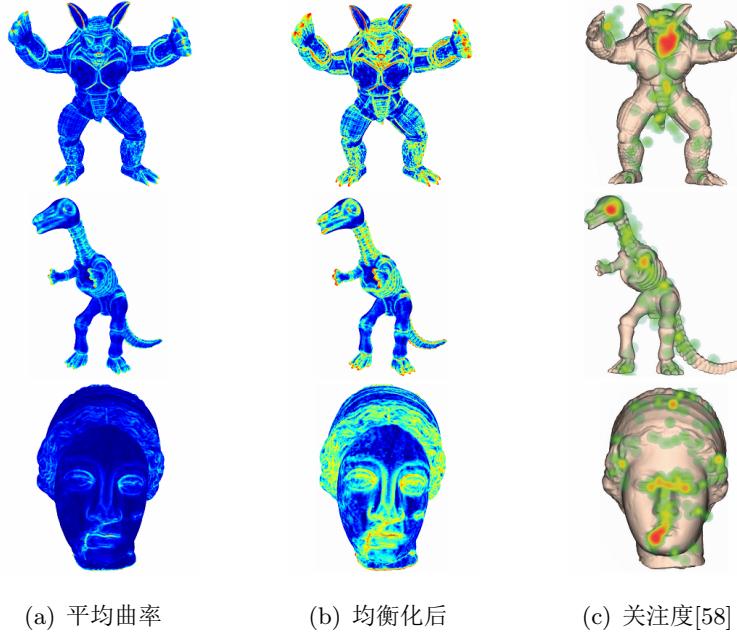


图 4.3: 部分模型的平均曲率、平均曲率直方图均衡化，以及人眼关注度[58]之间的相关性示意图

采样级别设定为256。对模型进行平均曲率直方图计算，并进行均衡化处理的伪代码表示在算法2中。

算法2中的 α 用于对原始平均曲率和均衡化的平均曲率进行融合。 α 取0到1之间的小数，值越大，对平均曲率均衡化效果越强。为了体现均衡化效果并且保留原始曲率分布特征，本章将 α 设定为0.5。

平均曲率直方图均衡化计算完毕之后，为了使特征分布更趋一致性，并便于后续的计算，我们将曲率分布再转换为32阶直方图。这个过程如图4.4所示，我们对Igea模型表面的256阶曲率分布直方图(图4.4a)进行均衡化处理(图4.4b, 4.4c)，最后将结果离散化为32阶直方图(图4.4d)。

4.2 视点评分

大多数视点评分方法将视点处的模型投影图像作为计算对象，模型表面的特征信息通过颜色来编码，最后使用图像处理技术来计算投影面积、轮廓等信息[96]。使用图像作为计算对象的问题是，计算结果往往对图像分辨率敏感，而选择较高的分辨率会对算法效率产生较大影响[3]。

Algorithm 2: 计算曲率直方图并进行均衡化

Input: 网格模型 $Mesh$

```

1 float bin[256];
2 for  $Mesh$  中的每个面片  $f$  do
3     int  $idx = \text{int}((\text{face}(f).curv - minC)/256);$ 
4      $bin[idx] += 1;$ 
5 for  $Mesh$  中的每个面片  $f$  do
6     float  $oldC = \text{face}(f).curv;$ 
7     int  $idx = \text{int}((oldC - minC)/256);$ 
8     float  $cdf = 0;$ 
9     for  $k = 1$  to  $idx$  do
10         $cdf += bin[k]/\text{face\_count}();$ 
11     float  $newC = cdf \cdot (maxC - minC) + minC;$ 
12      $\text{face}(f).curv = alpha \cdot newC + (1 - alpha) \cdot oldC;$ 

```

为了保证视点评分结果忠实反映三维模型表面特征，本章直接将三维模型面片作为计算对象。面片曲率依靠其顶点曲率求平均得到。面片对某一视点是否可见取决于其三个顶点是否都可见，或者其中心点是否可见。而顶点的可见性可以很容易地通过深度缓存数值和顶点实际深度比较来判断。

从上一小节的分析可知，高平均曲率区域也是人眼习惯关注的区域。因此，和传统的视点评分方法类似[66]，我们可以将可见面片平均曲率之和作为视点评分依据。可见平均曲率之和的计算公式为：

$$M'(w) = \sum_f C_f$$

为了反映透视投影导致的面片投影面积的变化，我们将面片法线和视线向量之间夹角作为调节参数，将上式修改如下：

$$M(w) = \sum_f C_f \cdot p_f \quad (4.1)$$

其中 $p_f = \text{abs}(\text{dot}(V, N_f))$ ，用来表示面片 f 在视平面的投影权重（ V 表示视线向量， N_f 表示面片 f 的法线向量）。

然而通过这个公式得到的视点分数会受到高平均曲率区域的影响，导致过多关注局部高特征的区域，而忽略物体全局特征分布。如图4.5a是使用公

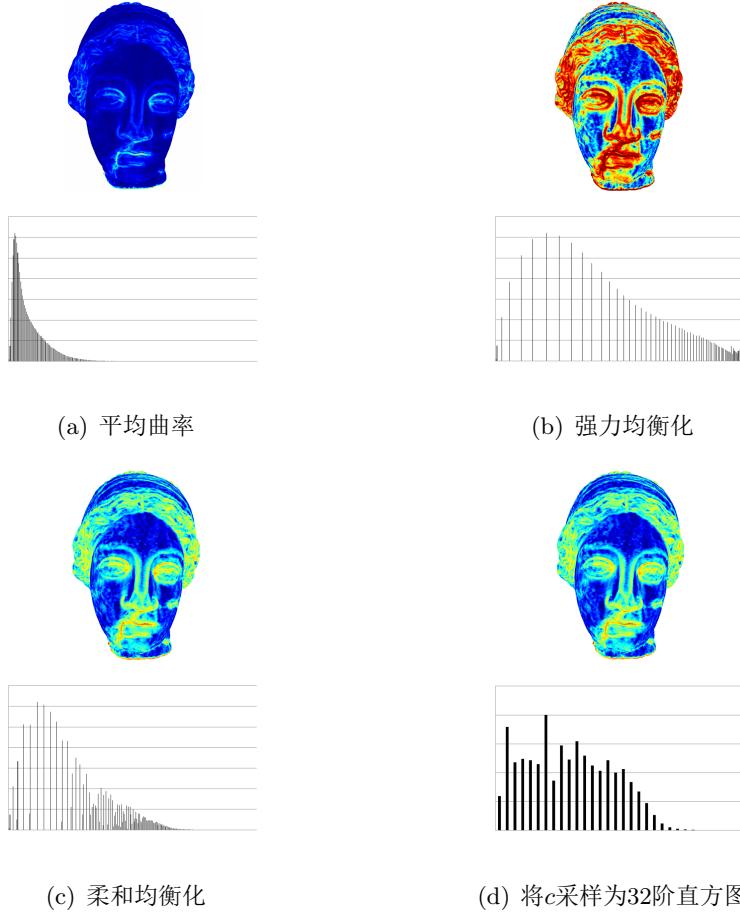


图 4.4: 对Igea模型计算平均曲率，并进行曲率直方图均衡化的结果

式4.1得到的模型最优视点，这个视点可以观察到模型五官、发型等重要信息，但却忽略了模型的底座。

为了解决这个问题，我们将信息熵引入。香农提出的“信息熵”理论，解决了对信息的量化度量问题，也被广泛应用于视点评分应用当中[27][97][112][96]。一个视点对于平均曲率分布的信息熵可以定义如下：

$$S(w) = - \sum_i \frac{R_i}{R} \log\left(\frac{R_i}{R}\right), R_i = \sum_{f \in H(i)} R_f p_f \quad (4.2)$$

其中 R_i 表示属于曲率直方图第*i*级的网格模型可见面片在视点*w*处的投影面积， R 表示模型总的投影面积。

利用信息熵，可以找到特征分布尽可能广且平均的视点。如图4.5b是使用

公式4.2得到的模型最优视点，从其平均曲率分布图可以看出，这个视角能够获取到多样化的模型表面特征，然而一些重要特征（比如五官）却丢失了。

我们期望获取的最佳视点不但能包含多样化的曲率特征信息，而且能包含更多的视觉重要特征区域。所以，我们将两个评分方式（公式4.1和公式4.2）进行结合：

$$Score(w) = M(w) \cdot S(w) \quad (4.3)$$

从图4.5c可以看出，利用公式4.3得到的模型最优视点，既很好地捕获到了五官、发型等重要特征区域，也尽可能地保留了底座等低曲率但语义重要的区域，用户通过这个视角可以更加全面地了解到这是一个微笑的、梳着发辫，底座水平的小女孩雕像。

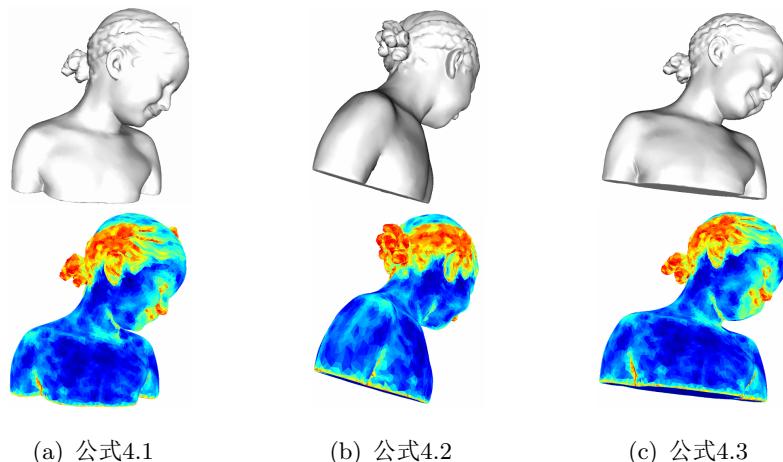


图 4.5: 分别使用3个不同评分公式得到的Bimba模型最优视点及其平均曲率示意图

4.3 生成代表性视点

为了得到最优视点，我们首先使用Bourke的方法¹生成258个平均分布于模型观察球上的候选视点。其中，观察球的球心位于模型中心，观察球半径R通过下面的公式计算得到：

$$R = b / \sin(F/2)$$

¹<http://paulbourke.net/geometry/>

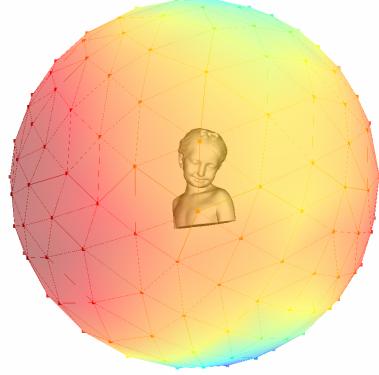


图 4.6: 模型的候选视点及其质量

其中 b 表示模型包围球半径, F 表示视点视野角度。使用公式4.3为这些候选视点计算得分, 分数最高的视点被称为最优视点, 可用于代表三维模型(如图4.6 所示)。

我们还可以继续添加其他最优视点来共同组成最优视点集。而新添加的视点必须满足如下两个条件:

1. 新加入的视点在剩余候选视点中评分最高;
2. 新加入的视点和最优视点集中已有视点区分度足够大。

为了满足这两个条件, 我们对视点 w 的平均曲率重要度公式(公式4.1)修改如下:

$$M_n(w) = \sum_f C_f \cdot \max(p_f - p_{fmax}, 0) \quad (4.4)$$

其中 $p_{fmax} = \max_{w \in W}(p_f(w))$, 表示目前选定的最优视点集合 W 中面片 f 的最大投影权重, 初始为0。

我们将此公式带入公式4.3中, 得到应用于最优视点集生成的视点评分标准如下:

$$Score_n(w) = M_n(w) \cdot S(w) \quad (4.5)$$

这个公式将赋予包含重要信息多、信息量大、且与集合中已有元素区别度大的视点较高评分。这很好地满足了我们上面提出的添加新的最优视点的两个条件。

我们利用公式4.5不断选定最优视点, 将其插入最优视点集中, 直到如下两个终止条件之一得到满足为止:

1. 最优视点集合中可见的模型面片已经占据了模型绝大部分区域[97];
2. 最优视点集中已经包含了足够多的用于认知这个三维模型的信息。

前者使用可见面片面积和模型总面积的比例来表示; 后者使用Kullback-Leibler距离(简称KL距离)来衡量。KL距离也被称为相对熵(Relative entropy), 是两个概率分布 p 和 q 差别的非对称性的度量。两个定义在 χ 上的概率分布 $p = \{p(x)\}$ 和 $q = \{q(x)\}$ 之间的KL距离计算可以表述如下:

$$KL(p|q) = \sum_{x \in \chi} \log(p(x)/q(x)) \quad (4.6)$$

为了连续性, 设 $0 \log 0 = 0$ 。 p 表示集合中可见面片的平均曲率分布, 而 q 表示模型表面平均曲率的分布。此处的KL距离表示藉由最优视点集所获取的模型表面特征分布和三维模型表面特征分布的接近程度。数值越小, 表明最优视点集越有助于观察者准确认知三维物体。值得注意的是, KL距离虽已被应用于衡量视点质量的计算[26][80], 但这些方法与我们的存在很大差异。以往的方法在计算KL距离时均使用最基本的面片作为衡量条件, 而在公式4.6中我们则使用不同级别的平均曲率(即不同级别的视觉重要度)进行计算, 具有更高层的视觉信息。此外, 我们将KL距离的计算应用于控制最优视点集中视点的数量, 而这种有效控制手段也是以往文献所欠缺的。

综上, 我们将选择最优视点集的算法用伪代码表示出来, 见算法3。其中, 用户可以调节 δ_{kl} 和 δ_r 来影响得到的最优视点集中的视点数量。

Algorithm 3: 生成最优视点集伪代码

Input: 观察球上的所有视点集合 W

Output: 最优视点集 N

```

1 for  $W$  中的候选视点 $w$  do
2   | 利用公式4.5计算视点 $w$ 得分;
3   | 将 $W$ 中得分最高的视点插入 $N$ 中;
4   | 依靠公式4.6计算 $N$ 的KL距离 $k$ , 并计算可见面片面积和模型表面积比值 $r$ ;
5   if ( $k < \delta_{kl}$  or 不收敛) or  $r > \delta_r$  then
6     | Return;
7   else
8     | Goto step 3;
```

4.4 实验结果及分析

为验证新视点评分方法的性能，我们在一台配备Intel Core i7-860 CPU (2.80 GHz, 8 CPUs), 12GB内存和NVIDIA GeForce GTX 460 显卡(1GB显存)的PC上进行了实验，包含效率分析、最优视点和最优视点集三个方面的内容。

4.4.1 效率分析

表 4.1: 部分模型计算时间统计表 (时间单位: 秒)

模型名称	面片数	文献[66]		Our	
		Saliency	最优视点	平均曲率	最优视点
pitbull	25,030	10.78	19.65	0.14	8.04
bunny	70,202	21.26	38.23	0.27	13.78
horse	96,966	31.92	53.02	0.9	16.54
santa	151,558	178.36	207.534	1.9	22.29
chambord	224,644	22.53	68.27	4.46	40.18
Armadillo	345,944	690.94	760.48	6.68	58.09
angel	474,048	3586	3680.31	37.37	113.61
dragon	871,306	3534.5	3698.36	46.06	159.41
blade	1,765,388	11452.46	11782.3	85.03	317.37

由于经典的网格显著度算法[66]需要计算多个层次的曲率信息，计算效率受限。而[68]中为了提高效率，将模型进行简化后计算显著度。但即便如此，对于具有540,000顶点的Buddha模型来说，只计算其显著度也需要102秒。利用本章方法为部分模型计算最优视点的时间统计如表1所示，从表中可以看出，本章方法的计算效率比[66]有了大幅提高。本章方法的计算时间和模型面片数成比例增长，这是因为所有计算都在模型表面完成。在使用图像处理技术计算视点评分的方法中，计算时间主要受投影图像分辨率影响，比如[3]中，使用512x512分辨率的图像，候选视点是100的条件下，计算模型最优视点的时间大约是110秒。虽然对于大复杂度模型（面片数大于500,000），本章方法时间耗费相对更多，但本章方法得到的最优视点结果不受投影图像分辨率影响，更加稳定。

4.4.2 最优视点

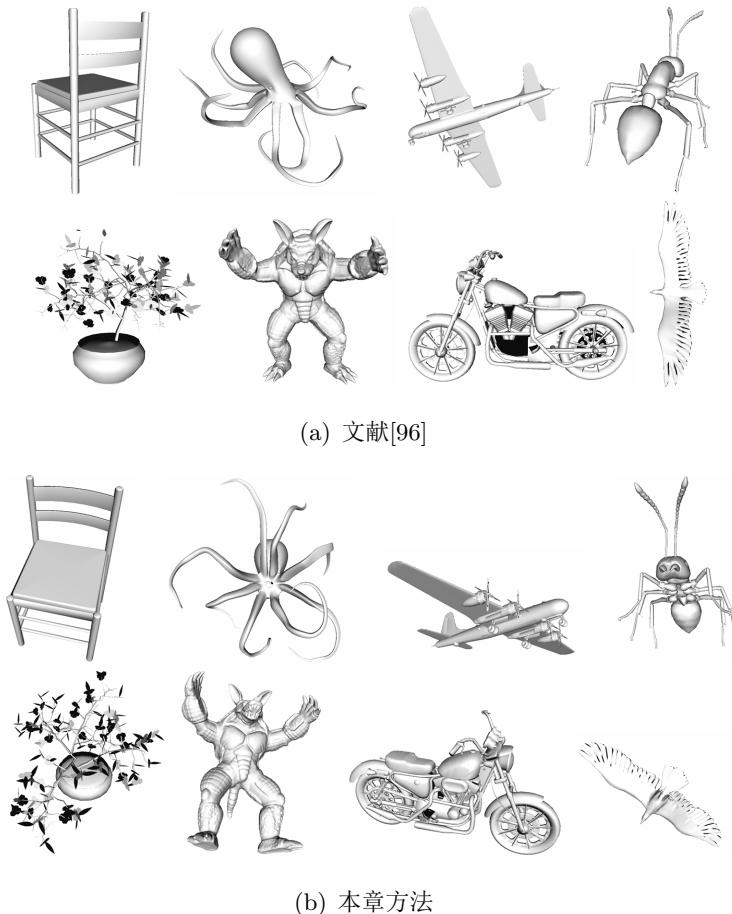


图 4.7: 需要大量人工交互的state-of-arts方法[96]和本章方法得到的最优视点比较

我们将本章方法得到的最优视点和目前最具权威的Secord等人的方法[96]进行了对比，从图4.7中可以看出，尽管本章方法是全自动的，但其大部分结果却和需要大量用户交互的[96]方法十分接近。这表明，使用本章的方法对视点的评分和用户观察习惯一致度极高。这主要得益于本章使用的平均曲率可以较好的反映模型表面的视觉特征，并且信息熵的引入使得视点所获取的信息量化成为可能。虽然对于蚂蚁、椅子、老鹰、八爪鱼等模型，两种方法得到的最优视点有比较明显的区别，但这些模型的最优视点往往表现出仁者见仁的特点，很难说哪个视点能够被多数人接受。本章方法得到的蚂蚁、椅子模型的最优视角是前方，从观察习惯上来说，要优于从后方的观察角度。而老

鹰、八爪鱼模型在本章中得到的仰视视角，从认知角度来说，更利于捕获这两种生物所具有的特性（如爪子、触角）。需要注意的是，本章方法是全自动的，而[96]方法需要大量的用户学习。

4.4.3 最优视点集

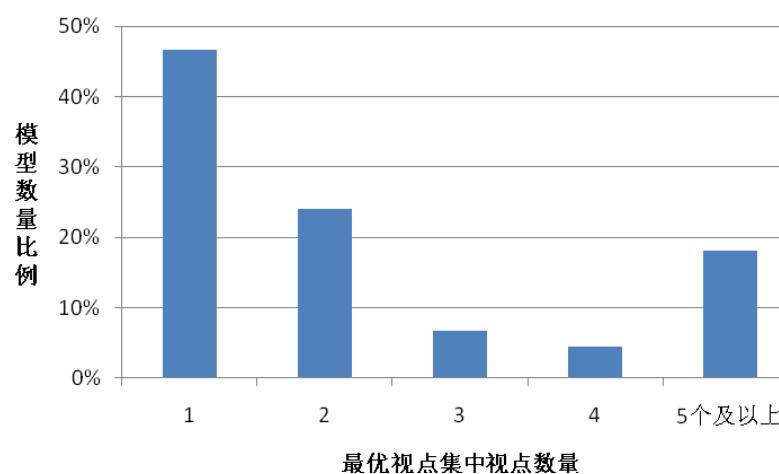


图 4.8: $\delta_{kl} = 0.02$, $\delta_r = 0.8$ 时，数据库中模型最优视点集的视点数量分布情况

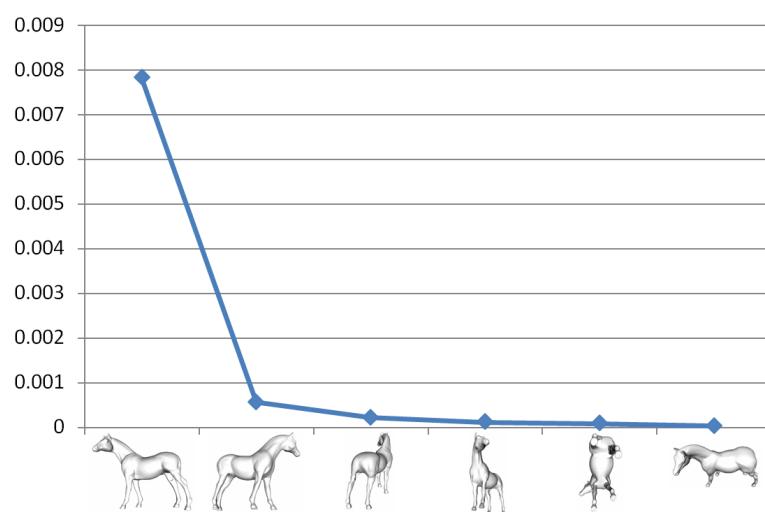


图 4.9: 使用本章方法得到的Horse模型的不同数量最优视点集获取的信息分布和模型信息分布之间的KL距离

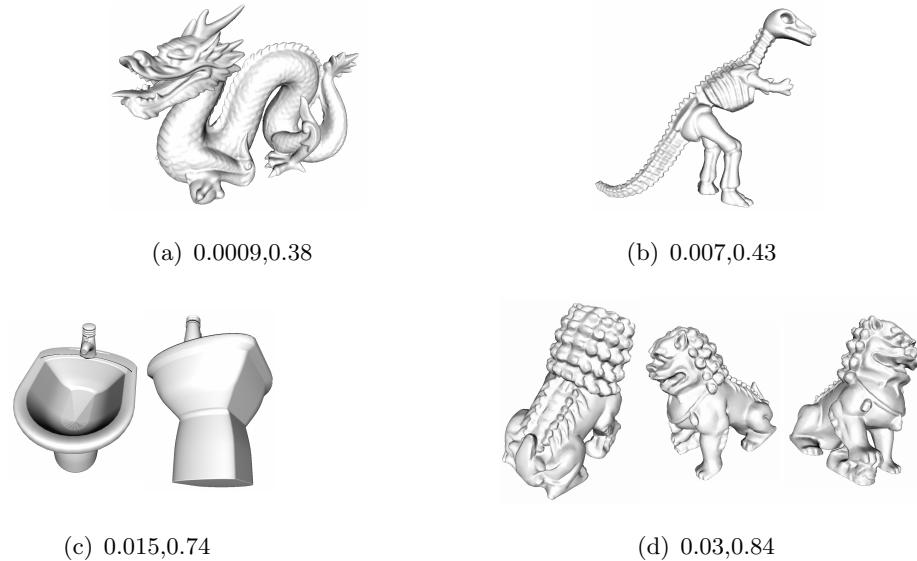
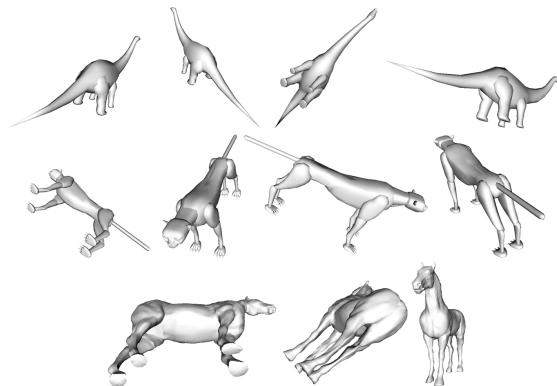


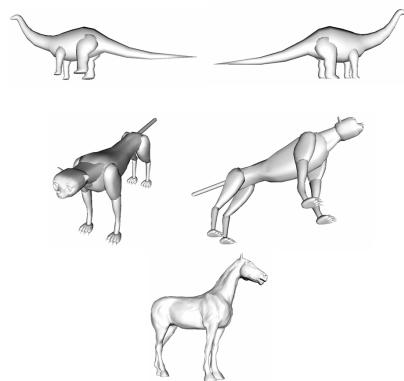
图 4.10: Kl 阈值=0.02, 可见面积比阈值=0.8时, 部分模型得到的最优视点集

从我们对大量不同种类模型（133个，包括扫描模型、人造、自然、动物、人体、家居、植被、交通工具、建筑、器官等）的统计结果来看（图4.8），接近一半的模型只需要1个最优视点即可满足认知需求。对于其他模型，最优视点获取的信息分布和原始信息分布的KL距离较大。对这一问题，我们一般可以通过构造最优视点集来解决。如图4.9所示，依次将最优视点插入最优视点集以后，视点集和Horse模型表面的信息分布之间的KL距离由最初的0.008左右下降到接近0。在最优视点选择过程中，由于模型的对称性，最优视点所观察到的模型快照可能重复，这可以很容易地通过深度图之间的图像对称性检测来解决[54]。

使用算法3，并将KL距离阈值 δ_{kl} 和所见面积比阈值 δ_r ，分别设定为0.02和0.8后，得到的部分模型最优视点集如图4.10所示（其中子标题的数字分别表示最优视点集的KL距离和可见面积比）。对个别模型（比如图中的ChinaDog模型），增加最优视点并没有使KL距离明显收敛，这是由于已有视点信息分布已经和模型表面比较接近了，增加了其他视点，反而导致信息分布偏离，这种情况下，通过可见面积比阈值 δ_r 可以阻止进一步添加过多视点。本章方法得到的模型视点集数量统计如图4.8所示，从图中可以看出，绝大多数模型仅需要1个或2个最优视点即可得到信息充分的缩略图。这表明本章方法得到的最优视点



(a) 文献[63]



(b) 本章方法

图 4.11: 使用文献[63]方法和本章方法得到的最优视点集结果



(a) 文献[68]



(b) 本章方法

图 4.12: 使用文献[68]和本章方法得到的HappyBuddha模型的最优视点集

极具代表性，并且更加紧凑、有效。

由于本章使用KL距离来定量表示最优视点集所信息分布和模型表面信息分布之间的差异程度，所以可以定量判断添加到最优视点集中的视点是否已经



图 4.13: 使用文献[96]的方法和本章方法得到的Lucy模型最优视点集

能够有效认知三维模型。而其他很多得到最优视点集的方法并不能对获取模型的信息进行定量分析。如图4.11所示，在[63]中使用固定数目为4或3的最优视点集，并没有考虑模型的特殊性，而我们的方法可以得到更具代表性且数目更少的视点集。在图4.12中，本章方法得到的HappyBuddha模型最优视点集包含2个最优视点，而[68]中是3个，从图中可以看出，其第3个视点中观察到的侧面衣服褶皱等信息利用前两个视点即可获取。此外，Secord等人也在文献[96]中探讨了最优视点集的生成，他们使用mean-shift聚类方法来得到视点评分方程中的峰值视点，组成最优视点集。但他们无法对最优视点集中包含的信息量和原始模型之间进行定量比较，导致最优视点集的冗余。图4.13a是[96]中得到的Lucy模型最优视点集。这7个视点得分分别为9.3、9.2、8.2、8.0、6.7、6.7、6.2。图4.13b是本章的结果，该最优视点集和模型的KL 距离是0.023，面积比为0.87。尽管文献[96] 方法得到的视点区分度足够大，但数量太多。和本章方法得到的结果相比，其后3个视点和前面的第2、1、4个视点很接近，属于冗余视点。而本章方法第一个视点和模型的KL 距离已经达到了0.035，表明第一个最优视点已经极具代表性。

4.4.4 局限性

尽管大量的实验表明本章提出的使用平均曲率并结合信息熵理论得到

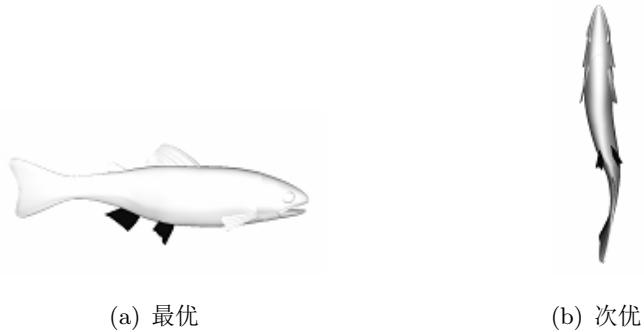


图 4.14: 使用本章方法得到的Fish模型的最优视点(左)和次优视点(右)

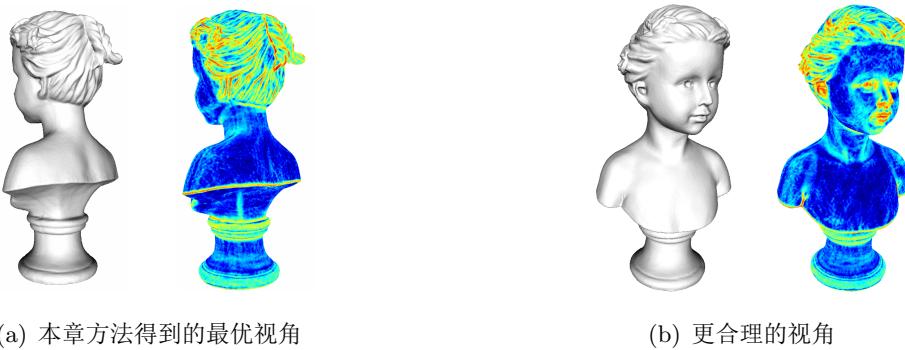


图 4.15: 本章方法将后视角错误地作为最优视角(a); 而更符合人眼观察习惯并且能获得更多有用信息的是(b)

的视点评价结果和人眼观察习惯基本相符, 但我们也发现了一些例外。比如图4.14中的Fish 模型, 利用本章的方法得到的最优视点是侧面, 并且此时视点包含信息和模型信息的KL距离已经小于0.003, 证明最优视点已经足够好。然而, 这个最优视角很难发现这条鱼其实是弯曲身体的状况, 需要添加第二个最优视点才能正确反映。

此外, 由于本章方法严重依赖于平均曲率, 尽管使用信息熵可以一定程度上调节大曲率区域对最优视点产生的影响, 但一些本身具备多样化曲率, 但语义上并不重要的区域会使得最优视点并不符合人类观察习惯。如图4.15中的模型, 使用本章方法得到的最优视点是斜后方(图4.15a), 因为后部的头发以及肩膀、底座等部位曲率分布范围广, 并且高曲率区域较多。很显然, 采用图4.15b这样的观察视角可以获得更多语义上重要信息(比如五官)。

4.5 小结及将来工作

本章提出一种能够定量评价视点获取模型表面特征信息量的视点评分方法，该方法将平滑滤波后的平均曲率值作为模型表面的特征分布数值，将能够获取更多高重要度特征并且特征分布范围比较广的视点评价为高分。这种视点评分方法除了可以进行最优视点选择之外，还可以被用于最优视点集的生成当中，能够产生有效认知模型的最少数目的视点集。

虽然本章使用平均曲率作为衡量模型表面特征的数值，但其他类似的数值，比如网格显著度[66]、兴趣区域（Interesting regions）[68]，以及语义分块[76]等都可以作为表面特征数值，并使用本章介绍的方法来评估视点质量。但这些数值计算复杂，需要提高效率，我们会在将来工作当中探索这种可能性。

第五章 基于视点评分的三维模型摆正

三维模型摆正属于形状分析问题。现实世界当中的物体，都存在稳定的或被人们所认可的摆正方向，比如动物一般以双足或者四足站立的方向为摆正方向；类似家具等人造模型以底部为摆正方向；而一些没有明显稳定底部的模型则具有人们习惯的摆正方向，比如吉他等乐器以及刀剑等武器人们习惯于将竖直方向或者平躺方向作为摆正方向。自动获得模型的摆正方向在模型的管理、分类、存储等诸方面有重要的应用[17][102]。但目前这方面的工作[31][48][69]，主要是针对人造模型进行探讨，以有效利用人造模型几何规整、对称性强、具有明显底座等特点。而其它非人造模型，如人体、树木等，则缺乏相关的研究，因为它们不像人造模型那样有便于特征分析的几何特征。

本章将提出一种新的确定模型摆正方向的方法，利用前面两章对模型观察视点进行评分的思想，将评分最差的若干视点作为模型的候选底部，并引入摆正稳定性及观察习惯性准则从中筛选出模型的真正底部，从而对模型进行自动摆正。实验表明，我们的方法对于人造模型和非人造模型都具有很好的适应性，即便在处理人造模型时，也能处理已有方法难以处理的一些情况。得益于一些视点评分方法的快速计算，我们方法在计算效率上也比已有摆正处理方法有明显的提高。

下面，我们先论述本章方法的思路及合理性，接着对具体步骤进行介绍和讨论，最后，结合大量的实验说明本章方法的有效性。

5.1 基本思路

5.1.1 使用视点评分获得候选底部

对于三维模型的摆正，需要确定模型的底部方位。为此，我们通过视点评分的方法来进行相关处理。模型底部的内容一般比较贫乏，因而相比于其他视点，观察底部的视点的观察质量较差。因此，我们可以采用视点评分方法，找到模型观察质量最低的区域，就很可能是模型的底部。对视点质量进行量化评分的方法很多，均使用一定的量化准则对观察视点质量进行评价，主要目的是为了找到观察质量最高的视点，所以也称为视点选择。我们选择了4种基于几

MA		
VE		
MC		
MS		
(a) 最优视点		(b) 最差视点

图 5.1: 不同种类模型分别使用4种视点评分方法得到的最优视点和最差视点

何特征度量的视点选择方法来进行对比实验，它们分别度量以下特征：投影面积 (MA) [86][49]、视点熵 (VE) [111]、平均曲率 (MC) [88][8]和网格显著度 (MS) [66]。然而，正如文献[96]中得出的结论：在选择最优观察视点时，没有一种视点选择方法适应于所有种类的模型。从图5.1中给出的部分实验结果可以看出，这4种常见的方法对模型的最优视点评价存在很大的差异（图5.1a）。然而，它们所得的最差视点中，却往往包含观察底部的视点（图5.1b中的红色框图）。这是因为内容贫乏的区域，无论怎么度量，其视点评分都是很低的。此外，在一些情况下，模型顶部的视点观察质量可能会比底部视点更差（图5.1b中的蓝色框图）。这也容易从自然界中找到例证，比如图5.1中的三脚椅模型，由于顶视角只能观察到椅面，特征很少，被大多数方法评价为最低分。这种同样可能被评价为差视点的顶部视点和模型底部视点关于模型中心是对称的（下文称其为底部视点的镜像点），因此，找到顶部视点也可以很容易地通过镜像点来定位底部视点。更多的实验结果见附录中的图5.11，其中红框表示最差视点为底部视点，而蓝框表示最差视点为底部视点的镜像点。

表 5.1: 不同视点评分方法判断底部视点为最差视点的比例

	底部视点为最差视点比例	底部视点或其镜像点为最差视点比例
MA	58.1%	86.3%
VE	61.2%	89.5%
MC	42.3%	70.5%
MS	40.9%	68.2%

我们对PSB数据库[102]中的模型进行了考察，在表5.1中列出了这4种视点评分方法判断底部视点为最差视点的情况。从表5.1中的“底部视点为最差视点比例”列中可以看出，这些方法能够正确将底部视点识别为最差视点的比例能够达到40%~60%左右，这表明最差视点有较大可能性就是模型底部视点。从表5.1的“底部视点或其镜像点为最差视点比例”列实验结果可以看出，模型底部视点的镜像点为最差视点的概率也很大。此外，还有一些情况下，模型的底部视点或其镜像点并非最差，但质量也较差，这时，可以通过调整合理阈值的方式得到若干差视点来将模型底部或其镜像点包含进来。

综上，我们可以通过视点评分方法找到若干差视点组成底部视点候选集，在此，考虑到差视点可能是顶视点的情况，我们将这些差视点的镜像点也加入这个集合当中。为了提高底部视点候选集中包含模型底部视点的几率，我们将多种视点评分方法得到的底部视点候选集合进行合并。然而，引入过多的视点评分方法会降低计算效率，并且可能导致候选底部视点过多，干扰后续的筛选计算。在实验过程中（如表5.1所示），我们发现MC和MS获得的候选底部准确度不高，并且两者都需要计算模型表面曲率，效率较低，因此，我们方法实现中将不使用这两种视点评分方法。对于VE方法而言，其考虑了视点可见信息量，故而一般情况下，使用VE可以正确判断出模型底部视点。但VE方法对于动物、植物这样的自然模型却不尽如人意，而计算简单的MA方法却在这两类模型上体现出了优势。这是因为自然模型表面特征复杂，并且不规整。而MA方法只考虑模型宏观层次的可见区域大小，避开了复杂的物体表面特征测量，反而容易获取较准确的候选底部信息。因此，VE方法和MA方法二者能够互为补充，它们获取到的差视点的并集中有90%以上的可能性包含了模型底部视点。因此，我们方法实现中，主要将MA和VE两种视点评分方法结合，将它们分别得到的底部视点候选集进行合并，以能处理更多种类的模型。

5.1.2 模型底部判定准则

依靠视点评分方法得到极可能包含模型底部视点的集合之后，我们需要依据一系列准则，从中筛选出模型真正的底部视点，从而达到对模型进行摆正的目的。已有工作探讨了许多相关准则[31][49][48][69][7][115]，它们大多是基于文献[31]提出的静态稳定性（较好的底部支撑可保证模型稳定站立）、对称性（关于模型底面的垂直方向上，物体对称性较强）、平行性（模型中和底面平行的面会比较多）、可见性（模型摆正后，其可见面比较多）等准则。这些准则普

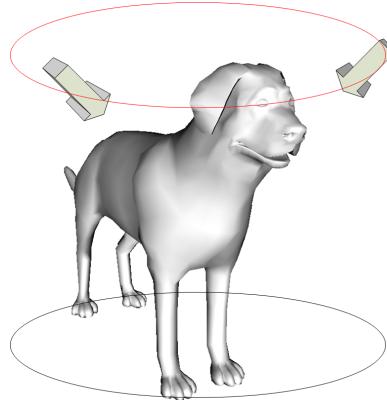


图 5.2: 模型的规范视点示意图

遍针对外形比较规整的人造模型，可以很好地在有限多个候选方向上筛选出物体的真正摆正方向。相较而言，本章得到候选集合中元素较少（对于库中的模型，有40%左右只有2个，有70%左右是4个以内，而其他方法往往有6个[115]，因此我们可用少量准则进行有效筛选。此外，本章方法的处理对象为任意种类的模型，这些针对人造模型的准则并不能很好地应用到其他外形不规整的非人造模型上（如图5.1中的潜水员模型，其底部的脚蹼就很难起到对物体进行稳定支撑的作用），并且使用过多的准则，会导致筛选效果变差。鉴于此，本章使用2种主要针对人造模型的准则：稳定性和支撑性。前者主要区分那些具有明显大底座的模型，而后者考虑了依靠支撑脚站立的模型。

为了进一步拓宽方法的适用性，我们提出视点合适度准则。这一准则主要基于文献[15]的研究结论：规范视点是观察模型的好视点，一般位于模型摆正姿态的上四分之三角度（three-quarter view）（图5.1a中的多数最优视点即为规范视点，比如MA中得到的所有最优视点）。这意味着，如果视点评分结果得当，则最优视点方向和摆正方向（模型底部观察视点到观察球中心连线方向）的夹角接近于 135° 。如图5.2所示，对于狗模型的规范视点（上部椭圆轨迹上的箭头所示），和模型摆正方向（竖直向上）夹角接近 135° ，符合这一规律。这一准则对人造模型和自然模型都是适用的。

5.2 方法实现

本章的模型摆正方基本流程如图5.3所示：

- 1.载入模型后（图5.3a），在模型的观察球面上均匀采样一些点作为候选视

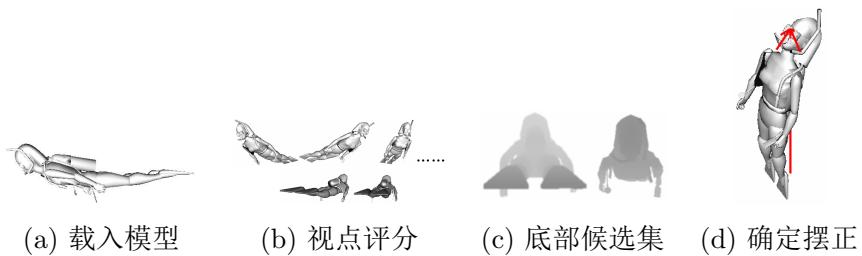


图 5.3: 本章对模型摆正的主要过程

点。

2. 分别使用MA和VE视点评分方法对观察球面上的候选视点进行评分(图5.3b)。

3. 对候选视点进行优化处理后，将MA和VE方法得到的差视点及其镜像点合并作为计算摆正方向的底部视点候选集(图5.3c)。

4. 对于底部视点候选集中的每个视点，我们考察其对于模型摆放的支撑性、稳定性、以及与规范视点的相关性的情况，从中选择最佳的底部视点，就可得到摆正方向(图5.3d)。

接下来，我们对主要过程分别进行详细介绍。

5.2.1 视点评分方法

我们选择MA和VE两种视点评分方法，因为二者得到的候选底部集合能够互为补充，并且它们很好地满足了以下几个条件：

1. 使用基于几何特征度量的视点选择方法，这些方法对于特征内容不多的底部方向是容易探测到的；
2. 它们的计算比较简单，结果比较稳定。这样，我们的方法就便于实践应用；
3. 计算过程是自动化的，不需要人工干预。

投影面积 (MA)

这种视点评分方法认为：能看到模型更大面积的视点，一般能获知模型更多的内容。因此，对于投影面积越大的视点，就给予更大的分值。最后，以分数高的视点作为好视点，反之为差视点。其关于视点w的分值 $Q_r(w)$ 计算式如

下：

$$Q_r(w) = \sum_f R_f \cdot p_f$$

这里， R_f 表示3D网格模型中在 w 处可见的面片 f 的面积； $p_f = \text{abs}(\text{dot}(V, N_f))$ 是投影计算，其中 V 表示视线向量， N_f 表示面片 f 的法线向量。

视点熵（VE）

该方法根据信息论进行视点评分，认为好视点看到的内容多，其信息熵的值会比较高。因此，该方法以一个面片在一个投影中所占比例作为该面片在这个视点观察中被看到的可能性，并据此进行视点熵的计算，其关于视点 w 的分值 $Q_i(w)$ 计算式如下：

$$Q_i(w) = - \sum_f \frac{R_f \cdot p_f}{R} \log\left(\frac{R_f \cdot p_f}{R}\right)$$

这里， R 表示模型在 w 处可见的面片投影面积和，其他参数和MA评分公式相同。

5.2.2 底部候选集生成

为兼顾计算的速度和计算结果的准确性，我们在实现时首先在模型观察球面均匀采样258个视点，然后使用上面介绍的视点评分方法对它们进行评分，以此得到差视点。为了避免观察球上相近视点得分接近，导致候选底部视点过多且位置接近的问题。我们依据评分值，对候选视点进行聚类计算。这样，位置及得分都接近的差视点会被划分到一个类中。

聚类计算，我们使用运算速度快的图像聚类算法[28]，该方法可以将得分差小于一定阈值的相邻视点聚为一类。聚类时，聚类阈值对于形成的类的数目有很大的影响。阈值比较大，则很多视点可放在一个类中，则分类的数目比较少；反之，则分类的数目比较多。如果聚类数目少，则每个类覆盖的观察球球面的范围比较大，会导致候选视点代表性变差；而聚类数目过多的话，每个类覆盖的观察球球面的范围比较小，导致得到的候选视点数量多且位置接近，视点候选集冗余。通过实验，我们设立聚类阈值为 $(S_h - S_l) \times 5\%$ ，一般能得到比较好的聚类结果，这里的 S_h 和 S_l 分别是观察球表面上采样视点的最高和最低得分。我们选择每个聚类中得分最低的视点作为代表点，模型底部视点候选集从这些代表点中产生。

采样视点进行聚类划分后，我们选定其中的若干类为差视点类。由于视点评分是对观察情况进行的一种逼近计算，摆正方向对应的视点不一定是得分最差的。因此，设立一个阈值，对于视点评分与最低视点评分的差距小于该阈值的视点，我们认为其为差视点。通过实验，该阈值也和聚类阈值一样，设定为 $(S_h - S_l) \times 5\%$ ，一般能得到比较好的结果。由于观察球上的视点是离散采样的，对应摆正方向的视点可能没有被采样到，但一般其附近的采样视点会是差视点。为此，我们以代表点为中心的一定范围内对它们视点评分值进行二次曲面拟合，将代表点位置调整到拟合曲面的极值点在观察球面上的对应点，就很可能是摆正方向对应的视点。相关计算步骤如下：

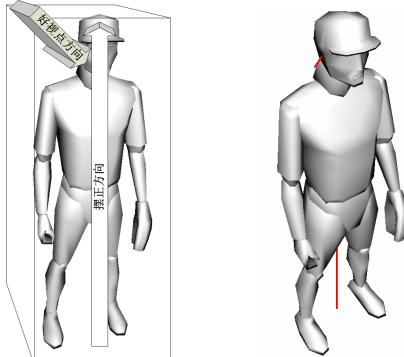
1. 以代表点为中心，1环邻域的离散视点作为边界，形成观察球面上的一个区域。
2. 对所形成的区域进行折半细分，得到一些新的采样视点，并对它们进行视点评分。
3. 在该区域内，根据所采样的视点及其视点评分进行二次曲面拟合（以球面区域作为定义域，视点评分值为值域），并找到该拟合曲面在此区域内具有最小视点评分值所对应的点，则该点即为所求的考察视点。之后，将考察视点的镜像点也作为考察视点一并加入视点候选集中。

使用上面的步骤，将MA和VE方法分别获得的经过优化调整以后的差视点及其镜像点作为模型底部视点候选集。

5.2.3 摆正方向确定

当底部视点候选集生成之后，接下来需要根据3个度量准则考察集合中的每个视点，从中筛选出模型真正的底部视点，从而确定模型的竖直方向，依据这一方向对模型进行摆正。由于不同准则对这些视点度量的协调性不便统一，且数值计算也可能有一定的偏差，我们在综合这些度量进行摆正方向的考察时，不是对这些度量数据值直接处理，而是按照下面的方式进行。首先，分别根据各个度量准则，对这些视点进行排序，得分高者排在后面，以得到较高的排序序列分（视点合适度、稳定性和支撑性3个准则排名分别设为： R_u, R_s, R_h ）。然后，对于一个视点，根据它在这个排序中的序列分，计算其作为底部视点的得分值，得分最高者为模型底部视点，计算公式为：

$$B(v) = w_u \times R_u(v) + w_s \times R_s(v) + w_h \times R_h(v) \quad (5.1)$$



(a) 好视点和候选摆正方向 (b) 摆正结果

图 5.4: 视点合适度准则示意图

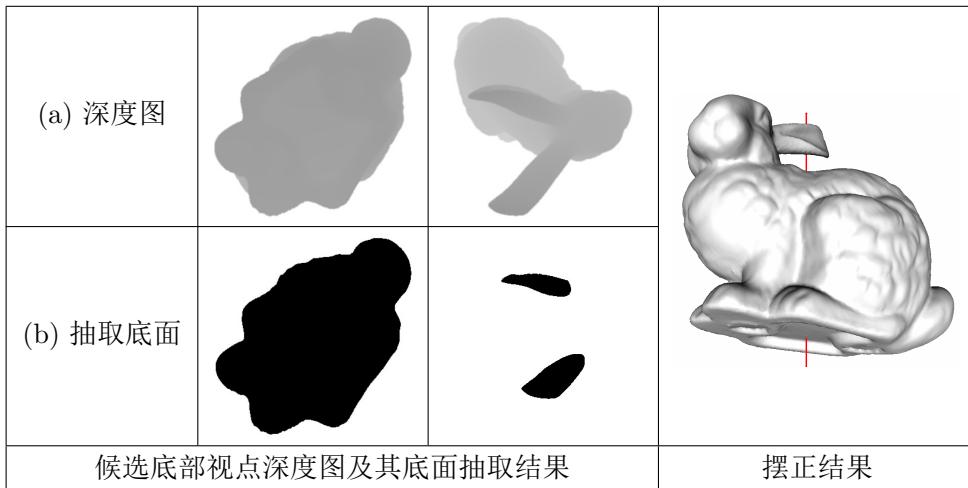


图 5.5: 稳定性准则示意图

这里 w_u, w_s, w_h 都是权值。考虑到本章方法的适用对象要求尽可能广泛，我们在测试时，为公式5.1中的3个准则赋予了相同的权重。如果最终程序面向自然模型，则可以为视点合适度权重 w_u 设定更高的数值（比如0.5），而为稳定性权重 w_s 和支撑性权重 w_h 设定更低的数值（比如0.25）。而面向人造模型时，则为 w_u 设定更低的数值（比如0.2），而为 w_s 和 w_h 设定更高的数值（比如0.4）。

接下来，我们介绍本章使用的这3个筛选准则具体的计算方法。

视点合适度：如前一小节所讨论，如果考察视点与最好视点的观察方向的夹角接近于135°，则考察视点的观察方向就很有可能是摆正方向。图5.4使用这一准则作为判定条件得到了正确摆正结果，图5.4a使用VE视点评分方

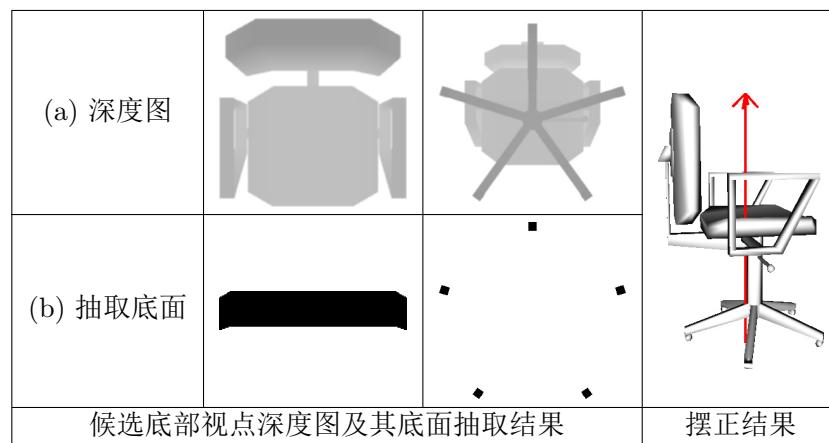


图 5.6: 支撑性准则示意图

法到了差视点和好视点，他们之间方向夹角更接近于 135° ，故而将其选定为底部视点，得到了正确的摆正结果（图5.4b）。视点合适度计算公式为： $U(v) = 1 - \text{abs}(\vec{V}_g \cdot \vec{V} - \cos(135))$ ，其中 \vec{V}_g 表示视点评分结果中的最好视点方向单位向量，而 V 表示视点 v 的观察方向单位向量。

稳定性：如果一个视点看到的模型部分，大部分位于离视平面最近的平面上，则表明模型沿着该视点的观察方向进行摆放，物体将有较大的接地面积，模型会比较稳定（考虑圆锥形的底部）；否则模型会不稳定（考虑圆锥形的顶部）。图5.5使用了这一准则，从图5.5a可以看出，从两个候选底部视点得到的深度图中抽取出的底面来看，第一个底面面积更大，故而稳定性更好，将其作为底面视点得到的摆正方向是正确的（图5.5b）。

支撑性：在一些情况下，沿着一个视点看到的模型部分，可能只有少部分位于离视平面最近的平面上，但这些部分分布均衡，可很好地支撑模型，形成摆正方向。图5.6使用了这一准则，从图5.6a可以看出，第2个候选底部处有5个支撑部位，支撑中心位于投影中心附近，而第1个候选底部得到的支撑底部是椅背顶端，偏离投影中心，支撑性差。故而选择第二个候选底部用于模型摆正是准确的（图5.6b）。

在计算稳定性和支撑性准则时，本章使用基于深度图像的方法，以获得更快的计算速度，伪代码表示在算法4中，其中第5行的 ϵ 表示一个冗余量，设为 $(Z_{max} - Z_{min}) \times 10\%$ 。

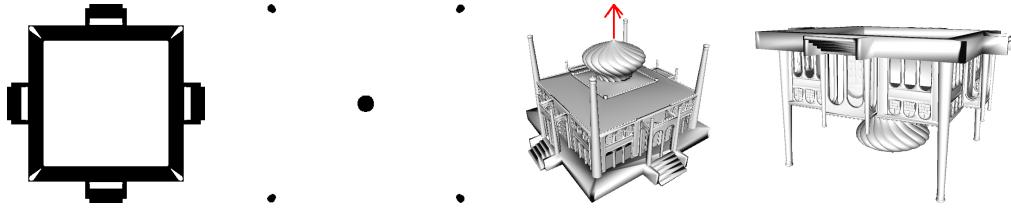
由于公式5.1既考虑了传统模型摆正方法针对人造模型的准则，也引入了更

Algorithm 4: 计算稳定性和支撑性准则伪代码

Input: 投影深度图 $depthMap$

Output: 支撑性得分 S , 稳定性得分 H

- 1 找到投影图像中心点 C ;
 - 2 计算 C 到离最近投影边界像素点的距离 D ;
 - 3 计算投影图像区域的面积 T ;
 - 4 计算到视平面距离最近的深度值 Z ;
 - 5 过滤, 删除深度值大于 $Z + \epsilon$ 的像素, 得到底座图;
 - 6 找到所有连通区域 $A_1 A_n$ 形成的图象块;
 - 7 计算每个图像块的面积 $T_1 T_n$;
 - 8 计算 $A_1 A_n$ 图像块的质心 C_a ;
 - 9 稳定性得分: $S = \sum_n (T_i / T)$;
 - 10 支撑性得分: $H = 1 - |C_a - C| / D$;
-



(a) 本章候选视点1 (b) 本章候选视点2 (c) 本章摆正结果 (d) 文献[31]摆正结果

图 5.7: 清真寺模型的候选视点和使用公式5.1得到的摆正结果

加通用的视点合适度准则, 所以对于有些已有工作难以处理的情况也能得到很好的结果。如图5.7所示, 利用本章的视点评分方法, 可以得到对称的2个候选视点(图5.7ab分别是两个视点处深度图过滤后得到的底座图), 分别位于模型的底部和顶部。由于这两个候选者的几何稳定性都很好, 难以区分, 所以文献[31]得到了错误的摆正结果(图5.7d); 而我们加入了视点合适度准则, 就能有效区分它们, 得到合适的摆正方向(图5.7c)。

5.3 实验结果及分析

5.3.1 适用性

我们在PSB模型库[102]上验证了本章提出的模型摆正方法, PSB模型库包含了1815个不同类型的3D形状(细分到162种)。我们依据PSB模型库的层次化分类方法, 将模型分为人造模型、自然模型(非人造模型)和杂项模型3种,

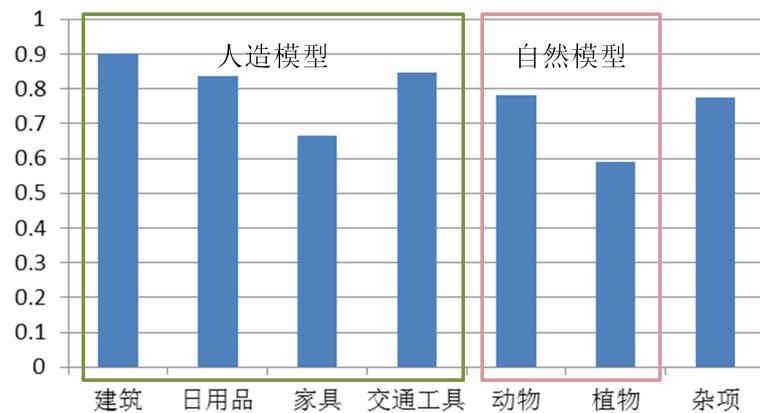


图 5.8: 本章对PSB数据库中不同种类模型的平均摆正准确度

其中杂项中既有人造模型也有自然模型。人造模型包括4个子类：建筑、日用品、家具和交通工具，自然模型包括2个子类：动物和植物。本章方法对各种类型模型的部分摆正结果见附录中的图5.12，从中可以看出，绝大多数模型都可以利用本章方法得到正确摆正结果。

我们方法在PSB模型库上的平均摆正准确度如图5.8所示。从图中可以看出，对于大多数类型的模型，我们的准确度已经达到75%以上。但对于植物类和家具类的准确度较低。这是因为植物模型规整性差，很难找到几何稳定的底部；而家具类模型中有很多只能在语义上区分的摆正方向，利用几何计算很难区分。比如图5.10中的平底桌模型，无论从几何稳定性还是视觉观察角度，依据正反两个方向都可以较好地摆正模型，只有通过人们的生活经验才能从中区别出日常的摆正方向。

5.3.2 对比实验

我们也把本章得到的结果和目前最好的方法进行了比较。得益于本章使用视点评分方法获取模型候选底部，无论是人造模型还是自然模型，本章方法都可以获取比较合理的候选底部视点。在筛选最终底部视点的过程中，本章使用的3个评分参数也充分考虑了所处理模型的普适性，通过高层次的视觉舒适度评分及模型稳定性、平衡性评分可以较准确地得到模型最终底部视点。

我们方法对PSB数据库模型的平均摆正准确率为79%左右，要优于文献[115]中平均70%左右的准确率，和文献[69] 中的摆正准确率相当，然而文献[69]只针对人造模型，并且统计样本只有120个。跟文献[31]中的87.5%的准确

度相比，我们的方法还有一定差距，但文献[31]需要大量的用户交互，适用的模型类别有限，且只针对人造模型。对于其他模型摆正方法主要处理的人造模型[31][48][69][7]，本章方法可以达到80%以上的准确度；对于其他方法难以处理的自然模型，平均准确度也能达到68%以上。

表 5.2: 在3组模型上的摆正准确率对比

模型种类	模型个数	摆正方法	摆正准确率
赛车	12	文献[69]	50
	15	文献[7]	73.3
	14	本章方法	73.3
显微镜	5	文献[69]	80
	6	文献[7]	83.3
	5	本章方法	100
花瓶	20	文献[69]	95
	12	文献[7]	91.7
	22	本章方法	86.4

本章方法和其他2种针对人造模型摆正方法在3组模型上的摆正准确率对比如表5.2所示（文献[7]和[69]公开了这3组模型的准确率），本章严格按照PSB数据库的模型分类进行实验，而2组对比方法增删了部分模型。对于赛车类模型，本章摆正准确率和文献[7]相当，都要高于文献[69]。对于显微镜类模型，我们可以达到100%的摆正准确率，远高于其他两类方法。对于花瓶类模型，本章方法准确度要差一些，其中颠倒错误占了很大比例，这是因为花瓶底部和顶部相似度很高，利用本章的3个优化参数，很难将花瓶的顶部和底部区别开来。

此外，对于[31]中难以处理的模型（图5.9a），利用本章方法可以得到令人满意的结果（图5.9b）；另外一个对比示例如图5.9中的床模型，在文献[48]和[7]中很难处理的模型（图5.9c），本章方法也能够得到准确的结果（图5.9d）。

5.3.3 效率分析

本章的实验在一台配备Intel Xeon E5405 CPU(2.0GHz, 2CPUs)，4GB内存的台式机上进行。本章方法对不同复杂度模型的处理时间如表5.3所示，从表中可以看出，主要的时间消耗集中于生成模型的视点候选集阶段。当获得视点候选集之后，由于本章使用基于图像的方法从中确定底部视点，和模型复杂度

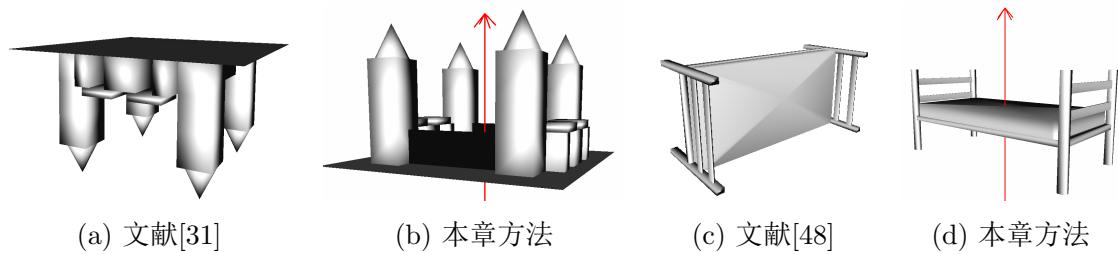


图 5.9: 本章方法能摆正其他方法失效的模型

无关，这个过程所需时间很少，比如在我们的实验中一般只需要200毫秒左右。总耗时方面，对于PSB模型库中绝大多数模型，整个摆正过程大约在8秒以内。

表 5.3: 本章对不同复杂度模型的计算时间统计（单位：秒）

顶点数	面片数	视点候选集生成	底部确定	总耗时
3k	6k	2.3	0.2	2.6
12k	25k	4.6	0.2	5.2
23k	46k	7.1	0.2	7.8

本章并未使用并行优化，也没有使用GPU参与运算，所以计算效率主要受CPU频率和模型自身复杂度的影响。和其他模型摆正方法的效率对比如表5.4所示，从表中可以看出，对于复杂度近似的模型，本章方法效率更高，即便我们所使用的CPU还是略差于其它方法所使用的CPU。这主要是由于其他方法需要分析组成模型部件的几何特征，用于确定模型支撑底部，而这个过程往往需要对模型进行复杂的分割计算。

表 5.4: 不同方法的效率比较(本章方法实验所用CPU能力弱于其它方法所用的CPU)

摆正方法	平均耗时	CPU频率 (GHz)
文献[31]	十几秒	2.13
文献[48]	十几秒	2.5
文献[69]	上百秒	2.33
文献[7]	几十秒	2.6
文献[115]	几十秒	3.3
本章方法	小于8秒	2.0

5.4 小结及将来工作

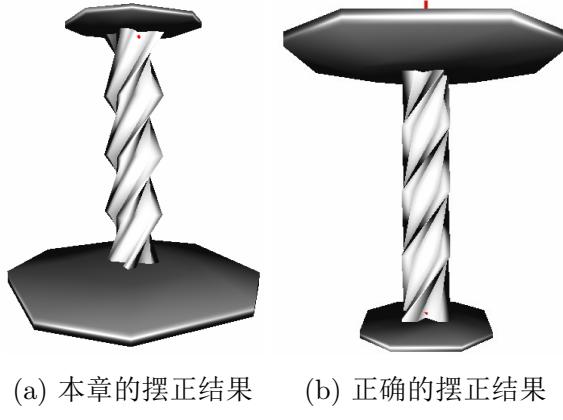


图 5.10: 本章方法失效的模型

本章介绍了一种新的模型摆正方法，利用模型底部视点往往评分较差这一规律，将候选底部的索引范围大幅度缩小。在3种筛选策略的配合下，我们可以较准确地从候选底部中找到模型的真正底部，从而对模型进行摆正。由于本章的候选底部及筛选规则对人造模型和自然模型都适用，因此相比于其他针对人造模型的方法更加鲁棒。

尽管绝大多数情况下，利用本章方法得到的视点候选集中包含了模型的真正底部视点，但在之后的优化计算过程中，仍然有一些错选情况。比如一些只能在高层语义上判断模型底部的情况，本章方法依旧无法处理。如图5.10所示，平底桌的桌面和底面都具备作为底面的特征，而桌面面积往往较大，导致其稳定性参数较大，从而将其误判为底面。此外，本章中更主要的失效情况是颠倒错误（如附录中图5.12的四腿桌、三脚凳模型），这表明本章的方法可以较好地将模型真正底部视点包含进视点候选集当中。然而，由于有些模型依靠水平底座摆放（花盆、杯子），有些模型则依靠支撑脚站立来保持平衡（桌椅、动物），而本章实验并没有考虑这些模型的特殊性，这需要在将来的工作中针对特定类型模型采用恰当的准则权重。

5.5 附录

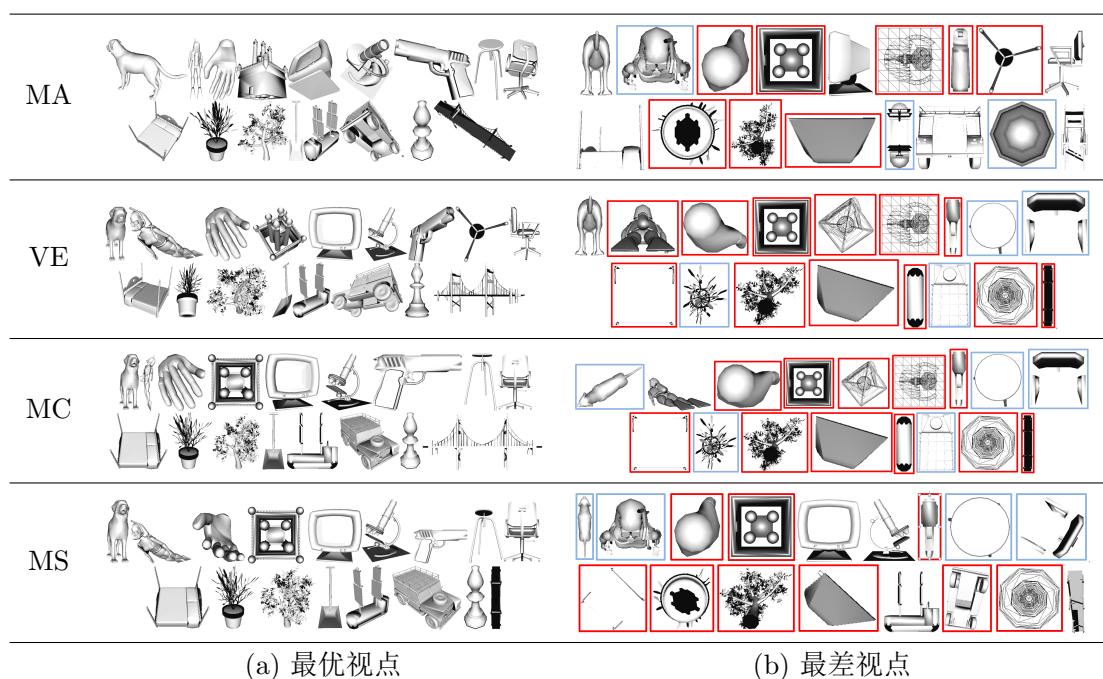


图 5.11: PSB模型库中部分不同种类模型分别使用4种视点评分方法得到的最优视点和最差视点

类型	摆正结果					
建筑						
家具						
日用品						
交通工具						
动物						
植物						
杂项						

图 5.12: PSB 模型库中部分不同种类模型的摆正结果

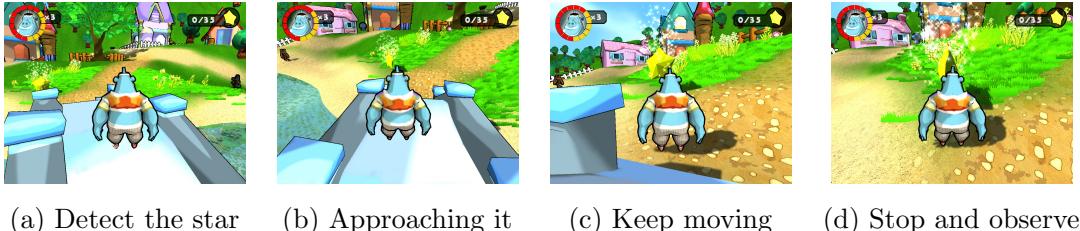
第六章 虚拟场景漫游中用户关注度的内隐测量方法

我们观察到，用户在进行电子游戏或虚拟现实场景漫游过程中，会将自身投射到虚拟世界中，通过操作虚拟世界中的化身主动对场景进行探索。在这种情况下，用户的漫游过程就是用户自己对场景认知的自然表达过程。因此，我们可以通过记录虚拟摄像机获取用户的漫游行为，进而对不同物体的观察行为进行量化。这一内隐的记录方式相比于使用眼动仪更能反映用户在漫游过程中的真实观察意图，从而能得到更准确的物体关注度。另外，本章采用测量，而非预测的方式获取用户主观关注度，保证得到的关注度信息可以真实反映不同个体的关注行为，避免了客观预测方法中预测结果难以符合不同个体关注习惯的问题。本章利用大量用户学习的方式来验证了这一测量方法的有效性，并将其应用于实际的游戏设计案例中。

6.1 基本思路

内隐是一个心理学术语，常用于注意捕获研究中，内隐的注意捕获指的是非期望刺激已经影响了个体的行为，但是却并没有进入意识[77]。与之相对的是外显的注意捕获，指的是一个非期望刺激引起了人的注意，并进入了人的意识当中。内隐的注意捕获通过研究被测试者的客观反应来获知当前的刺激如何影响用户的关注行为，无论这个刺激是否被测试者所意识到。而外显的注意捕获需要通过被测试者报告其主观意识的变化来实现注意捕获，这会造成某些内隐刺激的缺失。文献[77][74]为研究内隐方式测量而进行了一系列的对比实验，实验1会随机出现与测试任务无关的非期望刺激，而实验2则不出现这种刺激。结果显示，实验1中的很多被测试者并没有报告觉察到非期望刺激，但他们测试任务的完成情况却和实验2的测试者有显著差异。利用内隐的测量方式，更容易获取被测试者的真实意图。

在虚拟场景漫游过程中，用户的主观意识直接受到屏幕渲染图像的刺激，人脑会在任务等因素影响下对受到的刺激进行深层次加工，依据加工结果，用户做出相应行为，控制化身向着自己感兴趣的区域进行移动，最终形成完整的交互循环，推动用户对虚拟场景以自己所期望的方式进行漫游（这一过程如图6.1所示）。用户对化身的控制行为是其对外在刺激的反馈结果，所以通过对



(a) Detect the star (b) Approaching it (c) Keep moving (d) Stop and observe

图 6.1: 游戏中用户操作化身进行探索过程

用户控制化身行为的记录就可以测量出用户漫游的真实意图，从而得到用户针对不同场景物体的关注度。

虚拟漫游中，用户控制化身进行的观测最终要通过虚拟摄像机来完成。因此，用户关注度内隐测量的关键是如何定量表示摄像机对物体的关注程度。为此，我们根据已有的研究工作选择度量下面的3个参数来测量人对物体的关注度：离屏幕中心的距离(C)、遮挡率(O) 和投影面积(R)。

A. 离屏幕中心的距离(C)

这个参数主要是用来确定该物体是否为当前的观测主体。图像认知领域中的很多研究已经证实：人眼趋向于首先关注图像的中心区域[91]。在互动场景中，人们会控制摄像机将关注的物体调整到屏幕中心[109]。如图6.2a，火车处于屏幕中心，所以当前摄像机的观察主体是火车，而非远离屏幕中心的汽车和树木。这一结论也用在交互场景的关注度预测方面[41][42]。我们使用以下的公式计算这一参数：

$$C = \begin{cases} 1 - |x|/S_x, & x/S_x > y/S_y \\ 1 - |y|/S_y, & x/S_x \leq y/S_y \end{cases}.$$

其中 x 和 y 表示以屏幕中心点为原点的物体投影中心坐标， S_x 和 S_y 分别表示屏幕中心到屏幕右边界和上边界的像素数。这种计算方式将每个物体作为一个整体来看待，以物体投影中心到屏幕中心远近来测量物体当前的关注度。然而，只使用这个参数会忽略物体投影的大小和形状对关注度计算的影响。这个问题可以通过后面介绍的遮挡率和投影面积参数来处理。

B. 遮挡率(O)

这个参数主要考察当前物体是否被前面物体所遮挡，以及被遮挡的程度。在自由控制视角的交互场景中，用户通常会绕过遮挡物对感兴趣的物体进行观测。因此若屏幕中一个物体被严重遮挡，则它不是用户主要的观测对象。这也

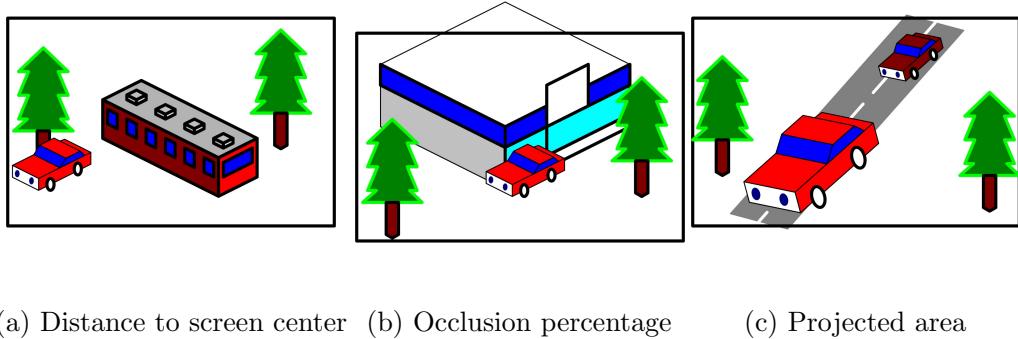


图 6.2: 物体被关注参数示意图

符合视觉认知领域方面的显著度理论, 即前景物体比其遮挡的背景物体受到更多关注[79]。如图6.2b, 房屋和汽车虽然都处于接近屏幕中心的观察位置, 然而, 汽车是前景, 而房屋多处被遮挡, 说明用户的观察主体是汽车。这个参数可以使用下面的公式来定义: $O = R_s/R_o$, 其中 R_s 表示模型实际投影面积, R_o 表示模型不遮挡情况下的投影面积。

C. 投影面积(R)

文献[67]指出: 用户习惯于近距离观察他所感兴趣的物体。由于直接使用观察距离会导致小物体的关注度测量值过高, 因此我们将物体的投影面积参数用于评估物体与观察者的距离。如图6.2c 所示, 两辆大小近似的车辆, 离屏幕中心距离相同, 并且都没有被遮挡, 但由于左下角车辆更接近观察者, 导致投影面积较大, 故而相对于右上角车辆来说, 这辆车受到的关注度更高。该参数计算公式为: $R = R_o/S_r$, 其中 S_r 表示屏幕面积。

我们对以上3个参数都进行了归一化并统一到屏幕空间中, 以便于后续计算。

用户在虚拟场景中漫游是一个持续的过程。其结果是, 用户会将重要物体逐步调整到舒适的观察状态, 并增加其观察时间[9]。为此我们将上述3个参数在时域空间累积, 以找到被观测物体在整个漫游过程而非某个时间点的被关注度, 计算公式如下:

$$G = \int [\lambda_C C(t) + \lambda_O O(t) + \lambda_R R(t)] dt. \quad (6.1)$$

其中 $C(t)$ 、 $O(t)$ 和 $R(t)$ 分别表示三个测量参数对时间 t 的函数，而 λ_C 、 λ_O 和 λ_R 分别表示这三个测量参数的权重。由于虚拟漫游场景是以帧为单位进行离散更新，所以，我们可以将公式6.1的积分形式利用下面公式6.2的逐帧叠加形式近似表示：

$$\begin{aligned} G_C &= \sum_{f \in F} C_f \Delta t, G_O = \sum_{f \in F} O_f \Delta t, G_R = \sum_{f \in F} R_f \Delta t. \\ G &= \lambda_C G_C + \lambda_O G_O + \lambda_R G_R. \end{aligned} \quad (6.2)$$

其中， G_C 、 G_O 和 G_R 分别记录漫游过程中的关注度参数 C 、 O 和 R ， F 表示物体处于可见状态的所有帧， C_f 、 O_f 和 R_f 分别表示处于帧 f 时三个关注度测量参数的数值， Δt 表示上一帧到当前帧的时间间隔。

6.2 有效性验证

为了验证上节提出的物体关注度内隐测量公式是否符合真实情况，我们进行了大量的用户体验调查，所采取的实验策略是：在漫游场景中随机放置若干区别度足够大的不同重要度物体，采集大量用户的浏览数据，以验证本章提出的物体关注度内隐测量公式(公式6.2)的计算结果是否和物体的重要度相符。

6.2.1 建立元型场景

本章使用了Unity3D 游戏引擎¹进行元型场景搭建，场景地图为 500×500 米，其中包含了平原、丘陵等地貌，并且具有河流、树木等自然元素（图6.3）。通过添加雾和真实感光照效果，力图营造一种清晨野外的自然场景效果，以便让测试者以一种自然的交互习惯进行场景漫游。用户操作采用第一人称视角，ASWD键控制角色的四个方向移动，空格键跳跃，鼠标控制视角旋转。主角行走速度为0.6米/秒，摄像机视野角度（FOV）为60 度，远近裁剪面距离分别是1000 和0.3，采用透视投影方式。限定浏览时间最长是10 分钟，大部分用户都能够在这段时间内完全浏览整个场景。

我们引入物体显著度的概念来表示虚拟场景中物体的视觉重要度[46]，将其从0 到1分为5个级别，显著度为0 的物体是背景物体，而显著度为1的物体具备最高的显著度。由于背景物体会对用户关注度的统计分析准确性产生影

¹<http://unity3d.com/>



图 6.3: 本章使用的三维场景



图 6.4: 元型场景中的不同显著度级别物体

响[12]，所以，我们在实验统计中将背景物体排除，只考虑其他4个显著度级别的物体（如图6.4所示），通过在物体上挂接脚本的方式来记录上一小节介绍的关注度测量值。我们邀请了10名志愿者在场景中漫游，并给出问卷，让他们为每种物体的显著度进行打分。对统计结果进行分析后，我们发现大部分志愿者的打分结果和我们的预期十分接近。

为保证用户在交互浏览场景时其观察行为只受到物体显著度影响，我们采取了以下措施：

- A. 场景中不出现明显的路径，弱化路线对于观察行为的影响。
- B. 将每个物体复制5份，随机分布于场景中的不同区域，并合并计算结果，弱化位置对观察行为的影响。
- C. 合并计算每个显著度级别的所有物体关注度，进一步保证采集数据的

表 6.1: 志愿者关注度统计信息

S	ID	G_C	G_O	G_R	G	G_C	G_O	G_R	G
0.25	1	47802.5	54320.4	9366.1	111489.0	43771.1	61503.7	5929.3	111204.1
	2	44042.9	66920.0	4803.2	115766.1				
	3	33854.4	44857.4	1610.3	80322.1				
	4	50777.2	74535.7	6711.6	132024.5				
	5	42378.5	66884.9	7155.3	116418.7				
0.5	6	46896.0	67491.2	5014.7	119401.9	46128.1	67295.8	6077.0	119500.9
	7	50087.7	78842.9	5279.1	134209.7				
	8	43632.6	59140.4	3927.1	106700.1				
	9	39884.8	62373.6	7797.2	110055.6				
	10	68765.0	94376.6	9320.3	172461.9				
	11	41304.8	58126.1	6902.6	106333.6				
	12	32325.7	50719.7	4297.9	87343.2				
0.75	13	53990.4	76299.3	9732.6	140022.4	51067.2	73615.7	9492.1	134175.0
	14	62862.6	94254.6	13940.3	171057.4				
	15	44840.3	63065.6	6117.6	114023.5				
	16	53339.0	80645.8	8713.3	142698.1				
	17	49399.3	72173.6	11721.9	133294.8				
	18	41971.4	55255.7	6727.0	103954.0				
1	19	76749.9	107515.0	9629.3	193894.2	88231.1	11276.8	161601.1	
	20	53904.5	71957.7	12141.5	138003.7				
	21	61453.7	85076.1	16827.9	163357.7				
	22	56775.0	86214.6	9681.2	152670.8				
	23	61583.3	90392.1	8104.0	160079.4				

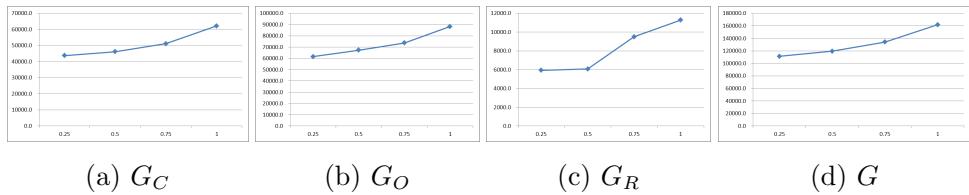


图 6.5: 物体显著度和关注度测量值之间关系

稳定性，降低志愿者个体对结果的影响。

6.2.2 用户学习及分析

我们邀请了70名志愿者参与此次用户学习，他们都是游戏设计相关专业的本科生，年龄在22岁左右，有多年3D虚拟场景操作经验。为了避免他们的浏览行为受到特定任务影响[12]，他们被提前告知这次的实验目的是为了进行一个3D虚拟场景稳定性研究的课题，要求他们在场景中随意浏览。

14个志愿者由于没有认真进行实验或者没有完整浏览整个场景，他们的结果被剔除。剩余的56个有效结果中，男性37名，女性19名。我们利用公式6.2测量得到的不同物体关注度如表6.1所示。在计算物体关注度 G 时，权重 λ_C 、 λ_O 和 λ_R 取相同数值。

我们将表6.1中物体显著度和关注度测量参数关系表示在图6.5中，横坐标表示不同显著度级别，纵坐标表示关注度测量值。从图中可以明显看出，物体显著度和关注度测量参数之间具有极强的相关性。这表明，本章提出关注度内隐测量方法可以有效地测量用户对物体的真实关注程度。

6.3 游戏场景辅助设计

在游戏设计过程当中，需要谨慎地设定物体的外观及各种属性，以便引导玩家以策划所预期的方式进行游戏。比如，将金币等奖励物品设计为闪闪发光属性，目的是吸引玩家进行拾取。然而，游戏场景设计完成以后，玩家漫游行为是否符合设计人员的预期(比如某个任务相关的重要物体是否受到了玩家的重点关注)仍然难以验证。利用本章提出关注度内隐测量方法，就可以较容易地得到玩家在游戏过程中对不同物体的关注度，从而帮助游戏设计人员评估设计方案效果，定位设计失败的物体。本章提出的这种内隐测量方法只需要在场景更新程序的每帧添加公式6.2即可实现，所以能够方便地嵌入到既有场景中。



图 6.6: 游戏的概览图

Scheme	Important objects	Background objects
1		
2		

图 6.7: Dream Land游戏两套场景物体设计方案

我们实验室的游戏创作团队制作了一款名为“Dream Land”的动作冒险类游戏²(如图6.6所示)。在场景中有一些比较重要的场景元素: NPC、敌人、健康之星，同时也有一些背景元素: 路标、蘑菇、石块、木桶。游戏设计人员希望玩家在游戏过程中主动和重要场景元素进行交互，以推动游戏剧情发展，而背景物体不需要得到玩家太多关注。在游戏的初始设计方案中(图6.7中的方案1)，有经验的游戏设计人员通过为重要物体使用鲜艳的材质、设定动画、添加特效等方式，吸引玩家对其进行交互。而他们为背景物体则设置了暗色材质，颜色接近于其所处环境，并且没有动画及特效，使得他们不容易被玩家察觉。我们要求设计人员对设定方案进行修改，在修改后的方案2中(图6.7)，有意将重要物体的材质亮度降低，并取消动画和特效，而对背景物体则增亮材质，并添加旋转动画和粒子系统特效。

²主要作者易恺铭和徐凯，游戏视频: <http://www.iqiyi.com/dongman/20110407/74b397d4092d6aff.html>

表 6.2: 志愿者对不同场景设计方案关注度统计信息

Scheme 1									
Saliency	ID	G_C	G_O	G_R	G	G_C	G_O	G_R	G
High	1	610.2	809.1	147.9	1567.2	629.4	871.7	121.9	1623.0
	2	615.4	836.5	140.8	1592.6				
	3	692.0	922.7	118.6	1733.3				
	4	600.1	918.3	80.3	1598.8				
Low	5	395.2	647.9	68.5	1111.5	463.3	688.5	70.1	1221.9
	6	612.5	868.0	75.7	1556.1				
	7	435.2	660.8	60.6	1156.6				
	8	410.2	577.5	75.7	1063.4				

Scheme 2									
Saliency	ID	G_C	G_O	G_R	G	G_C	G_O	G_R	G
High	1	371.7	591.9	76.8	1040.4	282.7	432.8	41.2	756.7
	2	264.0	334.3	27.8	626.1				
	3	252.8	432.6	37.6	723.0				
	4	242.4	372.4	22.6	637.3				
Low	5	233.1	308.8	29.9	571.8	331.9	426.1	61.6	819.6
	6	460.8	582.8	88.0	1131.6				
	7	419.0	509.4	79.1	1007.4				
	8	214.6	303.3	49.5	567.4				

我们共邀请了6名志愿者，每个方案均有3个不同志愿者参与测试。利用本章方法测量志愿者在不同场景方案中对物体的关注度结果如表6.2所示。从表中可以看出，方案1中，重要物体的关注度明显高于背景物体。而方案2中重要物体和背景物体的关注度数值差异并不明显。从平均关注度来看，重要物体关注度还小于背景物体。使用本章提出的关注度测量方法，可以较容易地得出方案1更符合游戏设计要求的正确结论。

6.4 小结及将来工作

本章介绍了我们提出的针对三维虚拟漫游场景物体关注度的内隐测量方法，并讨论了如何将其应用于电子游戏场景辅助设计当中。相比于使用眼动仪等硬件的主观测量方法，我们方法的准确度更高，并且很容易嵌入当前的虚拟漫游程序当中。

将来，我们拟从以下两个角度入手继续本章工作：

1) 关注度参数权重优化。本章在利用公式6.2的过程中，为3个关注度参数设定了相同的权重。尽管实验效果较好，但我们仍然无法从理论上证明3个参数对关注度测量是否具有相同的影响，这需要在后续工作中进行更进一步分析。

2) 更多领域的应用。本章将提出关注度测量方法应用于游戏场景辅助设计当中，取得了较好的效果。其实，物体关注度的量化分析还可以应用于其他更多领域中，比如利用关注度信息进行场景自适应优化、进行用户认知心理学分析等。

第七章 总结

本论文对三维场景中摄像机设置对其观察质量及关注度的影响展开研究，提出了两种适用于不同场合的三维模型观察视点质量评价方法，并将结果应用到模型摆正当中；本论文还提出以内隐方式测量用户关注度的方法，能够在虚拟漫游场景中高效地测量出用户对不同物体的关注度，并将结果应用于计算机游戏场景辅助设计当中。实验结果表明，这些研究成果可以较好地定量分析摄像机对三维场景的观察结果，从而帮助内容创作者有效展示或设计三维场景。本论文的具体工作总结如下：

(1) 论文在第三章中提出一种新的使用网格显著度为依据的模型分割方法，以此将模型表面分解为不同显著度的结构化分块。为了增强视点评分和人眼观察习惯的一致性，将这种显著度分块应用于视点评分当中，将那些可以看到更多重要分块的视点评价为高分。大量实验数据表明，这种视点的评分方法可以在只使用几何信息的条件下，得到包含语义信息的评分结果，特别适合于需要更符合人们观察习惯的视点评分结果的应用场合。

(2) 在第四章中，将计算效率较高的平均曲率作为度量模型表面特征的依据，并利用信息熵来量化视点所获取的特征信息。该章中使用的视点评分公式，可以将那些能够看到更多高平均曲率区域，并且所见信息量更大的视点，赋予更高的得分。这种方法可以提高视点评分计算效率，且不需要任何用户交互，特别适用于三维模型的在线预览、索引等需要快速计算的场合，以及最优视点集的生成中。

(3) 为了提高模型摆正方法的鲁棒性，论文在第五章中通过视点评分来指导模型候选底部的选择，并进一步引入人眼观察习惯及稳定性度量来提高计算结果准确性。从实验结果来看，该方法速度快，且能处理不同种类的模型，特别是对前人工作中难以处理的模型也能得到较好的结果。

(4) 在第六章中，提出了一种虚拟漫游场景中用户关注度的内隐测量方法。它能在没有硬件辅助的情况下，通过记录虚拟摄像机的观察参数在时间域上的变化情况，较准确地测量出用户在漫游过程中对不同物体的关注度数值。从应用效果来看，这种关注度测量方法极易和虚拟现实场景结合，能帮助场景设计人员快速地定量获取用户的关注度信息，从而提高场景设计质量。

参考文献

- [1] 孙晓鹏, 李华. 三维网格模型的分割及应用技术综述. 计算机辅助设计与图形学学报, 17(8):1647–1655, 2005.
- [2] 耿海燕, 蔡文菁. 不注意视盲的实验研究综述. 北京大学学报: 自然科学版, 43(2):282–287, 2007.
- [3] 杨利明, 王文成, 吴恩华. 基于视平面上特征计算的视点选择. 计算机辅助设计与图形学学报, 20(9):1097–1103, 2008.
- [4] 曹伟国, 胡平, 李华, 林宗楷. 基于距离直方图的最优视点选择. 计算机辅助设计与图形学学报, 22(9):1515–1521, 2010.
- [5] 杨利明, 韩红雷. 视点选择benchmark 的设计与应用. 软件学报, 21:165–172, 2010.
- [6] 缪永伟, 王洪军, 寿华好. 浮雕显著性信息熵驱动的三维模型最优视点选择. 计算机辅助设计与图形学学报, 23(12):2033–2039, 2011.
- [7] 姜波, 曾鸣, 刘新国. 利用形状特征的三维人造物体模型正朝向识别算法. 计算机辅助设计与图形学学报, 25(8):1099–1106, 2013.
- [8] P. Alliez, D. Cohen-Steiner, O. Devillers, B. Lévy, and M. Desbrun. Anisotropic polygonal remeshing. *ACM Transactions on Graphics (TOG)*, 22(3):485–493, 2003.
- [9] F. Alt, A. S. Shirazi, A. Schmidt, and J. Mennenöh. Increasing the user’s attention on the web: using implicit interaction based on gaze behavior to tailor content. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, pages 544–553. ACM, 2012.

- [10] M. Attene, B. Falcidieno, and M. Spagnuolo. Hierarchical mesh segmentation based on fitting primitives. *The Visual Computer*, 22(3):181–193, 2006.
- [11] M. Attene, S. Katz, M. Mortara, G. Patané, M. Spagnuolo, and A. Tal. Mesh segmentation-a comparative study. In *Shape Modeling and Applications, 2006. SMI 2006. IEEE International Conference on*, pages 7–7. IEEE, 2006.
- [12] M. Bernhard, E. Stavrakis, and M. Wimmer. An empirical pipeline to derive gaze prediction heuristics for 3d action games. *ACM Transactions on Applied Perception (TAP)*, 8(1):4, 2010.
- [13] I. Biederman. Recognition-by-components: a theory of human image understanding. *Psychological review*, 94(2):115, 1987.
- [14] S. Bischoff and L. Kobbelt. Ellipsoid decomposition of 3d-models. In *3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on*, pages 480–488. IEEE, 2002.
- [15] V. Blanz, M. J. Tarr, H. H. Bülthoff, and T. Vetter. What object attributes determine canonical views? *Perception-London*, 28(5):575–600, 1999.
- [16] P. Burelli. Virtual cinematography in games: Investigating the impact on player experience. *Foundations of Digital Games*, 2013.
- [17] B. Bustos, D. Keim, D. Saupe, T. Schreck, and D. Vranic. An experimental comparison of feature-based 3d retrieval methods. In *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*, pages 215–222. IEEE, 2004.
- [18] P. Castelló, M. Sbert, M. Chover, and M. Feixas. Viewpoint-driven simplification using mutual information. *Computers & Graphics*, 32(4):451–463, 2008.
- [19] M. Chaouch and A. Verroust-Blondet. Alignment of 3d models. *Graphical Models*, 71(2):63–76, 2009.

- [20] X. Chen, A. Golovinskiy, and T. Funkhouser. A benchmark for 3d mesh segmentation. *ACM Transactions on Graphics (TOG)*, 28(3):73, 2009.
- [21] M. Christie, P. Olivier, and J.-M. Normand. Camera control in computer graphics. *Computer Graphics Forum*, 27(8):2197–2218, 2008.
- [22] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia. Meshlab: an open-source mesh processing tool. In *Eurographics Italian Chapter Conference*, pages 129–136. The Eurographics Association, 2008.
- [23] D. Cohen-Steiner and J.-M. Morvan. Restricted delaunay triangulations and normal cycle. In *Proceedings of the nineteenth annual symposium on Computational geometry*, pages 312–321. ACM, 2003.
- [24] S. M. Drucker and D. Zeltzer. Camdroid: A system for implementing intelligent camera control. In *Proceedings of the 1995 symposium on Interactive 3D graphics*, pages 139–144. ACM, 1995.
- [25] L. Fan, M. Meng, and L. Liu. Sketch-based mesh cutting: A comparative study. *Graphical Models*, 74(6):292–301, 2012.
- [26] M. Feixas, E. Del Acebo, P. Bekaert, and M. Sbert. An information theory framework for the analysis of scene complexity. *Computer Graphics Forum*, 18(3):95–106, 1999.
- [27] J. Feldman and M. Singh. Information along contours and object boundaries. *Psychological review*, 112(1):243, 2005.
- [28] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.
- [29] D. A. Field. Laplacian smoothing and delaunay triangulations. *Communications in applied numerical methods*, 4(6):709–712, 1988.

- [30] S. Fleishman, D. Cohen-Or, and D. Lischinski. Automatic camera placement for image-based modeling. *Computer Graphics Forum*, 19(2):101–110, 2000.
- [31] H. Fu, D. Cohen-Or, G. Dror, and A. Sheffer. Upright orientation of man-made objects. *ACM Transactions on Graphics (TOG)*, 27(3):42, 2008.
- [32] Y. Gao, M. Wang, Z.-J. Zha, Q. Tian, Q. Dai, and N. Zhang. Less is more: efficient 3-d object retrieval with query view selection. *Multimedia, IEEE Transactions on*, 13(5):1007–1018, 2011.
- [33] M. Garland and P. S. Heckbert. Surface simplification using quadric error metrics. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 209–216. ACM Press/Addison-Wesley Publishing Co., 1997.
- [34] A. Golovinskiy and T. Funkhouser. Randomized cuts for 3d mesh analysis. *ACM Transactions on Graphics (TOG)*, 27(5):145, 2008.
- [35] F. Gómez, F. Hurtado, J. A. Sellàrès, and G. Toussaint. Nice perspective projections. *Journal of Visual Communication and Image Representation*, 12(4):387–400, 2001.
- [36] B. Gooch, E. Reinhard, C. Moulding, and P. Shirley. Artistic composition for image creation. In *Proceedings of the 12th Eurographics conference on Rendering*, pages 83–88. Eurographics Association, 2001.
- [37] L. Grady. Random walks for image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(11):1768–1783, 2006.
- [38] L. Hamid and M. Nakajima. Supervised learning of salient 2d views of 3d models. *芸術科学会論文誌*, 7(4):124–131, 2008.
- [39] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii. Topology matching for fully automatic similarity estimation of 3d shapes. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 203–212. ACM, 2001.

- [40] S. Hillaire, G. Breton, N. Ouarti, R. Cozot, and A. Lécuyer. Using a visual attention model to improve gaze tracking systems in interactive 3d applications. *Computer Graphics Forum*, 29(6):1830–1841, 2010.
- [41] S. Hillaire, A. Lécuyer, R. Cozot, and G. Casiez. Depth-of-field blur effects for first-person navigation in virtual environments. In *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, pages 203–206. ACM, 2007.
- [42] S. Hillaire, A. Lécuyer, T. Regia-Corte, R. Cozot, J. Royan, and G. Breton. Design and application of real-time visual attention model for the exploration of 3d virtual environments. *Visualization and Computer Graphics, IEEE Transactions on*, 18(3):356–368, 2012.
- [43] L. Itti. Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition*, 12(6):1093–1123, 2005.
- [44] L. Itti and P. Baldi. Bayesian surprise attracts human attention. *Vision research*, 49(10):1295–1306, 2009.
- [45] L. Itti and C. Koch. Computational modelling of visual attention. *Nature reviews neuroscience*, 2(3):194–203, 2001.
- [46] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 20(11):1254–1259, 1998.
- [47] A. Jain, T. Thormählen, T. Ritschel, and H.-P. Seidel. Exploring shape variations by 3d-model decomposition and part-based recombination. *Computer Graphics Forum*, 31(2pt3):631–640, 2012.
- [48] Y. Jin, Q. Wu, and L. Liu. Unsupervised upright orientation of man-made models. *Graphical Models*, 74(4):99–108, 2012.
- [49] H. Johan, B. Li, Y. Wei, et al. 3d model alignment based on minimum projection area. *The Visual Computer*, 27(6-8):565–574, 2011.

- [50] E. Kalogerakis, S. Chaudhuri, D. Koller, and V. Koltun. A probabilistic model for component-based shape synthesis. *ACM Transactions on Graphics (TOG)*, 31(4):55, 2012.
- [51] E. Kalogerakis, A. Hertzmann, and K. Singh. Learning 3d mesh segmentation and labeling. *ACM Transactions on Graphics (TOG)*, 29(4):102, 2010.
- [52] T. Kamada and S. Kawai. A simple method for computing general position in displaying three-dimensional objects. *Computer Vision, Graphics, and Image Processing*, 41(1):43–56, 1988.
- [53] S. Katz and A. Tal. Hierarchical mesh decomposition using fuzzy clustering and cuts. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 22(3):954–961, 2003.
- [54] M. Kazhdan, B. Chazelle, D. Dobkin, A. Finkelstein, and T. Funkhouser. A reflective symmetry descriptor. In *Computer Vision—ECCV 2002*, pages 642–656. Springer, 2002.
- [55] D. Kelly. Motion and vision. ii. stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America*, 69(10):1340, 1979.
- [56] C.-Y. Kim and R. Blake. Psychophysical magic: rendering the visible ‘invisible’ . *Trends in cognitive sciences*, 9(8):381–388, 2005.
- [57] V. G. Kim, W. Li, N. J. Mitra, S. DiVerdi, and T. Funkhouser. Exploring collections of 3d models using fuzzy correspondences. *ACM Trans. Graph.*, 31(4):54:1–54:11, July 2012.
- [58] Y. Kim, A. Varshney, D. W. Jacobs, and F. Guimbretiere. Mesh saliency and human eye fixations. *ACM Transactions on Applied Perception (TAP)*, 7(2):12, 2010.
- [59] C. Koch and S. Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. In *Matters of Intelligence*, pages 115–141. Springer, 1987.

- [60] M. Koivisto, J. Hyönä, and A. Revonsuo. The effects of eye movements, spatial attention, and stimulus features on inattentional blindness. *Vision Research*, 44(27):3211–3221, 2004.
- [61] K. Kolev, P. Tanskanen, P. Speciale, and M. Pollefeys. Turning mobile phones into 3d scanners. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 3946–3953. IEEE, 2014.
- [62] A. Koschan. Perception-based 3d triangle mesh segmentation using fast marching watersheds. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–27. IEEE, 2003.
- [63] H. Laga. Semantics-driven approach for automatic selection of best views of 3d shapes. In *Proceedings of the 3rd Eurographics conference on 3D Object Retrieval*, pages 15–22. Eurographics Association, 2010.
- [64] Y.-K. Lai, S.-M. Hu, R. R. Martin, and P. L. Rosin. Fast mesh segmentation using random walks. In *Proceedings of the 2008 ACM symposium on Solid and physical modeling*, pages 183–191. ACM, 2008.
- [65] G. Lavoué. A multiscale metric for 3d mesh visual quality assessment. *Computer Graphics Forum*, 30(5):1427–1437, 2011.
- [66] C. H. Lee, A. Varshney, and D. W. Jacobs. Mesh saliency. *ACM Transactions on Graphics (TOG)*, 24(3):659–666, 2005.
- [67] S. Lee, G. Joungyun Kim, and S. Choi. Real-time tracking of visually attended objects in virtual environments and its application to lod. *Visualization and Computer Graphics, IEEE Transactions on*, 15(1):6–19, 2009.
- [68] G. Leifman, E. Shtrom, and A. Tal. Surface regions of interest for viewpoint selection. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 414–421. IEEE, 2012.

- [69] C.-K. Lin and W.-K. Tai. Automatic upright orientation and good view recognition for 3d man-made models. *Pattern Recognition*, 45(4):1524–1530, 2012.
- [70] H. Liu, L. Zhang, and H. Huang. Web-image driven best views of 3d shapes. *The Visual Computer*, 28(3):279–287, 2012.
- [71] A. P. Mangan and R. T. Whitaker. Partitioning 3d surface meshes using watershed segmentation. *Visualization and Computer Graphics, IEEE Transactions on*, 5(4):308–321, 1999.
- [72] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 416–423. IEEE, 2001.
- [73] R. McDonnell, M. Larkin, S. Dobbyn, S. Collins, and C. O’Sullivan. Clone attack! perception of crowd variety. *ACM Transactions on Graphics (TOG)*, 27(3):26, 2008.
- [74] C. MOORE and H. EGETH. Perception without attention: Evidence of grouping under conditions of inattention. *Journal of experimental psychology. Human perception and performance*, 23(2):339–352, 1997.
- [75] J. D. Morgante, R. Zolfaghari, and S. P. Johnson. A critical test of temporal and spatial accuracy of the tobii t60xl eye tracker. *Infancy*, 17(1):9–32, 2012.
- [76] M. Mortara and M. Spagnuolo. Semantics-driven best view of 3d shapes. *Computers & Graphics*, 33(3):280–290, 2009.
- [77] S. B. Most, B. J. Scholl, E. R. Clifford, and D. J. Simons. What you see is what you set: sustained inattentional blindness and the capture of awareness. *Psychological review*, 112(1):217, 2005.

- [78] K. Mühler, M. Neugebauer, C. Tietjen, and B. Preim. Viewpoint selection for intervention planning. In *EuroVis*, pages 267–274, 2007.
- [79] M. B. Neider and G. J. Zelinsky. Searching for camouflaged targets: Effects of target-background similarity on visual search. *Vision Research*, 46(14):2217–2235, 2006.
- [80] L. Neumann, M. Sbert, B. Gooch, W. Purgathofer, et al. Viewpoint quality: Measures and applications. In *Proceedings of the 1st Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging. Aire-la-Vile: The Eurographics Association Press*, pages 185–192. Citeseer, 2005.
- [81] M. Ovsjanikov, W. Li, L. Guibas, and N. J. Mitra. Exploration of continuous variability in collections of 3d shapes. *ACM Transactions on Graphics (TOG)*, 30(4):33, 2011.
- [82] D. L. Page, A. Koschan, S. R. Sukumar, B. Roui-Abidi, and M. A. Abidi. Shape analysis algorithm based on information theory. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 1, pages I–229. IEEE, 2003.
- [83] D. Parkhurst, K. Law, and E. Niebur. Modeling the role of salience in the allocation of overt visual attention. *Vision research*, 42(1):107–123, 2002.
- [84] J. B. Pelz and R. Canosa. Oculomotor behavior and perceptual strategies in complex tasks. *Vision research*, 41(25):3587–3596, 2001.
- [85] R. J. Peters and L. Itti. Applying computational tools to predict gaze direction in interactive visual environments. *ACM Transactions on Applied Perception (TAP)*, 5(2):9, 2008.
- [86] D. Plemenos and M. Benayada. Intelligent display in scene modeling. new techniques to automatically compute good views. In *Proceedings of GraphiCon*, volume 3, 1996.

- [87] J. Podolak, P. Shilane, A. Golovinskiy, S. Rusinkiewicz, and T. Funkhouser. A planar-reflective symmetry transform for 3d shapes. *ACM Transactions on Graphics (TOG)*, 25(3):549–559, 2006.
- [88] O. Polonsky, G. Patané, S. Biasotti, C. Gotsman, and M. Spagnuolo. What’s in an image? *The Visual Computer*, 21(8-10):840–847, 2005.
- [89] M. Reddy. Perceptually optimized 3d graphics. *IEEE computer Graphics and Applications*, 21(5):68–75, 2001.
- [90] D. Reniers and A. Telea. Skeleton-based hierarchical shape segmentation. In *Shape Modeling and Applications, 2007. SMI’07. IEEE International Conference on*, pages 179–188. IEEE, 2007.
- [91] L. W. Renninger, P. Verghese, and J. Coughlan. Where to look next? eye movements reduce local uncertainty. *Journal of Vision*, 7(3):6, 2007.
- [92] C. A. Rothkopf, D. H. Ballard, and M. M. Hayhoe. Task and context determine where you look. *Journal of vision*, 7(14):16, 2007.
- [93] S. Rusinkiewicz. Estimating curvatures and their derivatives on triangle meshes. In *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*, pages 486–493. IEEE, 2004.
- [94] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy. Real-time 3d model acquisition. In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH ’02*, pages 438–446, New York, NY, USA, 2002. ACM.
- [95] M. Sbert, M. Feixas, P. Castelló, and M. Chover. Information theory tools for viewpoint selection, mesh saliency and geometry simplification. In *Intelligent Computer Graphics 2009*, pages 41–61. Springer, 2009.
- [96] A. Secord, J. Lu, A. Finkelstein, M. Singh, and A. Nealen. Perceptual models of viewpoint preference. *ACM Transactions on Graphics (TOG)*, 30(5):109, 2011.

- [97] E. Serin, C. Doger, and S. Balcisoy. 3d object exploration using viewpoint and mesh saliency entropies. In *Computer and Information Sciences II*, pages 299–305. Springer, 2012.
- [98] E. Serin, S. Sumengen, and S. Balcisoy. Representational image generation for 3d objects. *The Visual Computer*, 29(6-8):675–684, 2013.
- [99] A. Shamir. A survey on mesh segmentation techniques. *Computer graphics forum*, 27(6):1539–1556, 2008.
- [100] L. Shapira, A. Shamir, and D. Cohen-Or. Consistent mesh partitioning and skeletonisation using the shape diameter function. *The Visual Computer*, 24(4):249–259, 2008.
- [101] P. Shilane and T. Funkhouser. Distinctive regions of 3d surfaces. *ACM Transactions on Graphics (TOG)*, 26(2):7, 2007.
- [102] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The princeton shape benchmark. In *Shape modeling applications, 2004. Proceedings*, pages 167–178. IEEE, 2004.
- [103] K. Shoemake. Arcball: a user interface for specifying three-dimensional orientation using a mouse. In *Proceedings of the conference on Graphics interface'92*, pages 151–156. Morgan Kaufmann Publishers Inc., 1992.
- [104] D. Sokolov and D. Plemenos. Virtual world explorations by using topological and semantic knowledge. *The Visual Computer*, 24(3):173–185, 2008.
- [105] S. L. Stoev and W. Straßer. A case study on automatic camera placement and motion for visualizing historical data. In *Proceedings of the conference on Visualization'02*, pages 545–548. IEEE Computer Society, 2002.
- [106] V. Sundstedt, M. Bernhard, E. Stavrakis, E. Reinhard, and M. Wimmer. Visual attention and gaze behavior in games: An object-based approach. In *Game Analytics*, pages 543–583. Springer, 2013.

- [107] V. Sundstedt, E. Stavrakis, M. Wimmer, and E. Reinhard. A psychophysical study of fixation behavior in a computer game. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, pages 43–50. ACM, 2008.
- [108] M. J. Tarr and D. J. Kriegman. What defines a view? *Vision Research*, 41(15):1981–2004, 2001.
- [109] P.-H. Tseng, R. Carmi, I. G. Cameron, D. P. Munoz, and L. Itti. Quantifying center bias of observers in free viewing of dynamic natural scenes. *Journal of vision*, 9(7):4, 2009.
- [110] A. Van Dam and S. K. Feiner. *Computer graphics: principles and practice*. Pearson Education, 2014.
- [111] P.-P. Vázquez, M. Feixas, M. Sbert, and W. Heidrich. Viewpoint selection using viewpoint entropy. In *Proceedings of the Vision Modeling and Visualization Conference 2001*, pages 273–280. Aka GmbH, 2001.
- [112] P.-P. Vázquez, M. Feixas, M. Sbert, and W. Heidrich. Automatic view selection using viewpoint entropy and its application to image-based modelling. *Computer Graphics Forum*, 22(4):689–700, 2003.
- [113] T. Vieira, A. Bordignon, A. Peixoto, G. Tavares, H. Lopes, L. Velho, and T. Lewiner. Learning good views through intelligent galleries. *Computer Graphics Forum*, 28(2):717–726, 2009.
- [114] D. V. Vrancic, D. Saupe, and J. Richter. Tools for 3d-object retrieval: Karhunen-loeve transform and spherical harmonics. In *Multimedia Signal Processing, 2001 IEEE Fourth Workshop on*, pages 293–298. IEEE, 2001.
- [115] W. Wang, X. Liu, and L. Liu. Upright orientation of 3d shapes via tensor rank minimization. *Journal of Mechanical Science and Technology*, 28(7):2469–2477, 2014.

- [116] Y. Wang, K. Xu, J. Li, H. Zhang, A. Shamir, L. Liu, Z. Cheng, and Y. Xiong. Symmetry hierarchy of man-made objects. *Computer graphics forum*, 30(2):287–296, 2011.
- [117] K. Xu, H. Zhang, D. Cohen-Or, and B. Chen. Fit and diverse: set evolution for inspiring 3d shape galleries. *ACM Transactions on Graphics (TOG)*, 31(4):57, 2012.
- [118] H. Yamauchi, W. Saleem, S. Yoshizawa, Z. Karni, A. Belyaev, and H.-P. Seidel. Towards stable and salient multi-view representation of 3d shapes. In *Shape Modeling and Applications, 2006. SMI 2006. IEEE International Conference on*, pages 40–40. IEEE, 2006.
- [119] Y.-L. Yang and C.-H. Shen. Multi-scale salient features for analyzing 3d shapes. *Journal of Computer Science and Technology*, 27(6):1092–1099, 2012.
- [120] T. B. Zaharia and F. J. Preteux. 3d versus 2d/3d shape descriptors: A comparative study. In *Electronic Imaging 2004*, pages 47–58. International Society for Optics and Photonics, 2004.
- [121] M. Zhang, W. Wang, H. Sun, and H. Han. Perception-based model simplification for motion blur rendering. *Graphical Models*, 76(3):116–127, 2014.
- [122] Y. Zheng, H. Fu, D. Cohen-Or, O. K.-C. Au, and C.-L. Tai. Component-wise controllers for structure-preserving shape manipulation. *Computer Graphics Forum*, 30(2):563–572, 2011.
- [123] Y. Zheng and C.-L. Tai. Mesh decomposition with cross-boundary brushes. *Computer Graphics Forum*, 29(2):527–535, 2010.

发表文章目录

- [1] **Honglei Han**, Jing Li, Wencheng Wang, Huiwen Zhao, Miao Hua. Reasonable Viewpoint Ranking via Saliency-based Segmentation[J]. *International Journal of Modeling, Simulation, and Scientific Computing*, accepted. (EI)
- [2] 韩红雷, 徐婵婵, 费广正, 崔响, 王文成. 虚拟场景漫游中用户关注度的内隐测量方法[J]. 中国科学.信息科学, 2014, 44(11): 1398-1408. (一级学报)
- [3] 韩红雷, 王文成, 华淼. 基于视点评分的三维模型摆正[J]. 软件学报, 已录用. (EI 一级学报)
- [4] 韩红雷, 李静, 费广正. 结合视觉感知与信息量的视点评分方法[J]. 计算机辅助设计与图形学学报, 2014, 26(6): 939-947.(EI 一级学报)
- [5] Zhang M, Wang W, Sun H, **Han H**. Perception-based model simplification for motion blur rendering[J]. *Graphical Models*, 2014, 76(3): 116-127.(SCI)
- [6] 王光伟,王文成,华淼,韩红雷. 优化光源设置的自适应增强绘制[J]. 软件学报, 2013,24(8): 1885-1897. (EI 一级学报)
- [7] 杨利明, 韩红雷. 视点选择Benchmark的设计与应用[J]. 软件学报, 2010, 21: 165-172. (EI 一级学报)
- [8] 韩红雷, 王文成. 光栅化时的线性纹理插值[J]. 计算机辅助设计与图形学学报, 2011, 23(6): 999-1005. (EI 一级学报)

简 历

基本情况

韩红雷，男，汉族，1980年1月出生于山西太原，博士研究生。

教育背景

1999年9月-2003年7月，在北京物资学院，管理科学与工程系，获得管理学士学位。

2004年9月-2006年7月，中国传媒大学，计算机与软件学院，获得计算机应用技术工学硕士学位。

2009年9月考取中国科学院软件研究所，攻读博士学位，师从王文成研究员，研究方向为计算机图形学。

工作经历

从2006年6月至今，在中国传媒大学任教，主要承担计算机程序设计、计算机图形学及游戏开发技术相关的教学及科研工作。

联系方式

hanhonglei@sina.com

致 谢

硕士期间我开始接触计算机图形学，这个学科给了我很多有趣的工作体验。在拿到硕士学位的时候，我就立志将来要在这个专业工作。我有幸在硕士毕业之后留校到中国传媒大学工作，继续从事计算机图形学相关的教学和科研工作，实现了我硕士期间的理想。然而，面对讲台下一群求知若渴的大学生，我内心是忐忑的，因为这个专业发展变化日新月异，我只有时刻保持旺盛的学习及科研热情，才能将更多一线的技术传授给他们。因此，我在2009年决定报考博士研究生，继续在专业方面进行深造。在顺利通过中科院的博士入学考试之后，我从2009年7月开始，来到中国科学院软件研究所计算机科学国家重点实验室学习。

在五年的博士学习期间有过挫折、有过迷茫，最终在老师、同学、朋友及家人的帮助下，顺利发表了若干研究成果，并完成了毕业论文，最重要的是开阔了专业眼界，提高了科研能力。

首先，我要衷心地感谢我的导师王文成研究员。他知识渊博、思维活跃、治学严谨、工作兢兢业业、写作功底深厚，所有这些都让我受益匪浅。在学术上，他不仅在宏观上也在细节上指导学生；在生活上，他对学生关怀备至，宽容又无比真诚。无论在学术上还是在生活上碰到困难，总是能够得到导师的帮助。在软件所的五年里，我得到了王老师的认真指导和悉心栽培，他在学术上的严格要求，工作上的尽善尽美，才使得我取得现在的成绩。王老师严谨治学的科研态度必将成为我将来学习和工作的楷模和榜样，他对我人生道路的支持和指引也是我终生最大的一笔精神财富。在此谨以此文表示我衷心的感谢，祝福王老师身体健康、万事如意！

我还要把感激之情表达给中国传媒大学，这所学校治学严谨，有浓厚的艺术和技术结合的学术氛围。学校特别重视年轻教师的专业发展，这也最终使得我获得了这次攻读博士学位的机会。我要特别感谢我所在的动画学院的前后两任院长：廖祥忠教授和黄心渊教授，他们治学严谨，管理科学，为我们打造了完美的工作平台。我特别要感谢动画学院的费广正教授，他为人谦和，学术造诣深厚，并且具备深远的学术眼光。我是在他的鼓励下走上计算机图形学的研究道路，也正是在他的影响下，我最终报考了软件所进行博士学习。

感谢图形组已经毕业的杨利明和陈昕，他们的前期研究成果对我的帮助很大，我进行实验的程序所使用的库很多都是由他们开发的。我要特别感谢图形组的其他师兄弟们：谢国富、别晓辉、王光伟、华淼、张民英、高天豪、徐盼盼、郭诗辉、王东旭、李雷和宋滢，感谢他们给予我的合作、建议和帮助。每次与大家讨论，总能得到不少的提高。感谢李静师姐和我合作完成了2篇论文。正是由于图形组形成的良好学习和研究氛围，使得我克服了很多困难，顺利完成学业。

同时，感谢图形组的其他老师。作为国内计算机图形学的前辈级大师，吴恩华研究员儒雅温厚、和蔼谦逊，学术功底深厚，为我树立了做人和做学问的榜样。非常感谢吴老师在学术和生活上的关怀照顾，祝吴老师身体健康，事事顺心！感谢陈彦云研究员和刘学慧副研究员，在我研究方面的指导和帮助。

感谢实验室的林惠民老师、陈海明老师、郭菊卿老师、庄丽华老师、张丽老师、费腾老师、刘辉老师、孙守云老师以及其他老师为我们的科研提供的优质服务，谢谢你们的辛苦付出和默默奉献！感谢各部的李彩丽老师、韩智芳老师、安亚丽老师、刘娟老师和李晓云老师等在学生事务上对我的支持与帮助。特别感谢研究生部的张欢老师为我们论文答辩做了大量细致入微的准备工作。

感谢游戏设计系的石民勇教授、黄石副教授、高金燕老师、李晋老师、税琳琳老师、陈京炜老师、王巍寅老师，感谢他们在工作上对我的帮助。感谢中国传媒大学博士研究生徐婵婵，她除了跟我合作完成科研工作之外，还在论文撰写方面给了我很大帮助。感谢图形组已经毕业的杨猛，感谢他一直坚持和我打羽毛球锻炼身体，而且在锻炼间隙还不忘讨论科研项目的申报。

感谢我的父母和岳父母，他们对家庭的悉心照料使我能够全身心投入到工作当中，我要祝他们身体健康。我也要感谢我的妹妹、妹夫还有小外甥王予函，他们给予了我无尽的关怀。

感谢我的爱人李岩给我的鼓励和帮助，她总是在背后默默地奉献和全力支持，使我得以克服在博士期间遇到的种种困难和挫折。感谢我的宝贝女儿韩欣宜，在我工作烦闷，项目受阻的时候，她总能带给我快乐，缓解我烦躁的情绪，我要把这篇论文献给她。

最后，诚挚感谢为评审本论文而付出辛勤劳动的各位专家！

感谢所有给予我关心、支持和帮助的人们！