

基于视点评分的三维模型摆正^{*}

韩红雷^{1,2,3}, 王文成¹, 华 淼^{1,2}

¹(计算机科学国家重点实验室(中国科学院 软件研究所), 北京 100190)

²(中国科学院大学, 北京 100049)

³(中国传媒大学 动画与数字艺术学院, 北京 100024)

通讯作者: 韩红雷, E-mail: honglei@ios.ac.cn

摘 要: 3D 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, 且不少方法还需人工干预, 工作效率不高. 提出一种可处理任意模型的方法, 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; 而主要能观察到这些区域的视点, 其视点评分一般很低. 因此, 根据视点评分, 就能有效地寻找到模型的底部朝向, 即, 模型的摆正方向. 为了提高计算的可靠性, 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. 实验结果表明: 新方法能够高效地处理不同种类的模型, 包括一些已有方法中难以处理的情况, 并具有很好的计算效率.

关键词: 摆正方向; 视点评分; 三维模型

中图法分类号: TP391

中文引用格式: 韩红雷, 王文成, 华淼. 基于视点评分的三维模型摆正. 软件学报, 2015, 26(10): 2720–2732. <http://www.jos.org.cn/1000-9825/4742.htm>

英文引用格式: Han HL, Wang WC, Hua M. Getting upright orientation of 3D objects via viewpoint scoring. Ruan Jian Xue Bao/ Journal of Software, 2015, 26(10): 2720–2732 (in Chinese). <http://www.jos.org.cn/1000-9825/4742.htm>

Getting Upright Orientation of 3D Objects via Viewpoint Scoring

HAN Hong-Lei^{1,2,3}, WANG Wen-Cheng¹, HUA Miao^{1,2}

¹(State Key Laboratory of Computer Science (Institute of Software, The Chinese Academy of Sciences), Beijing 100190, China)

²(University of Chinese Academy of Sciences, Beijing 100049, China)

³(School of Animation and Digital Arts, Communication University of China, Beijing 100024, China)

Abstract: Upright orientation of a 3D object is conducive to applications such as model alignment and function recovery. However, state of art methods mainly focus on man-made models only, and some of them require much user interaction, thus lowering the processing efficiency. In this paper, a new method is proposed to effectively attain the upright orientation of an object, no matter whether it is man-made or natural. It is based on the observation that there are very few surface features on the base of the object, and so the views for watching the base will be given very low scores in view evaluation. Thus, using to the view scores, the base of the model can be found effectively to get the upright orientation of the model. To improve computation reliability, two metrics are further integrated into the method to measure the stability of model placement and the habitual viewing direction of human beings to effectively watch a model. Experimental results show that the presented method can effectively handle a variety of models, including some man-made models that cannot be well handled with existing methods, and it can run much faster than existing methods.

Key words: upright orientation; viewpoint scoring; 3D model

近年来,关于三维模型摆正方向的检测得到了很大的关注.这是因为模型的摆正方向在模型的管理、分类、存储等诸方面有着重要的应用^[1,2],如:根据模型的摆正方向进行陈列,可有效降低摆放的空间需求.但目前,这方

* 基金项目: 国家科技支撑计划(2012BAH62F03); 国家自然科学基金(61379087); 中国科学院知识创新工程领域前沿项目

收稿时间: 2014-03-27; 修改时间: 2014-07-16, 2014-10-09; 定稿时间: 2014-10-20

面的工作^[3-5]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.而其他非人造模型,如人体、树木等,则缺乏相关研究,因为它们不像人造模型那样有便于特征分析的几何特征.

本文提出一种新的检测模型摆正方向的方法,可有效处理多种模型,包括人造和非人造的自然模型.其出发点是: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;而视点选择方法在视点评分时,对于观察内容贫乏区域的视点,其评分是很低的.由此,我们可以利用视点评分方法找到评分低的视点,以有效探测模型底部,并进而提高模型摆正方向的检测效率.当然,模型表面内容贫乏的区域可能有多个.为此,我们进一步根据模型摆放的稳定性和人眼观察模型的习惯性,以提高摆正方向探测的可靠性.因为模型沿着摆正方向具有很好的摆放稳定性,且大量研究表明,人眼观察模型的习惯方向一般位于模型摆正姿态的上四分之三角度(three-quarter view)^[6].关于模型的底部一般表面内容比较贫乏的成因,大概如下:人造模型在制作过程中,其底部很少被处理,且在制作过程中要保持摆放的稳定性,因此底部一般是比较平坦的,表面内容不丰富;而非人造模型在其形成过程中,其底部一般较少受到各种形成因素的影响,如阳光、雨水、风沙等,因此其底部的表面内容一般也不如其他部位丰富.实验结果表明: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,我们的方法在计算效率上也比已有摆正处理方法有明显的提高.

下面我们首先简要介绍相关工作.然后讨论本文方法的可行性及实现的具体步骤.最后,结合大量的实验说明本文方法的有效性.

1 相关工作

1.1 模型摆正

许多关于模型摆正方向检测的方法,主要是基于主元分析法(principal component analysis,简称 PCA).这是因为主元分析法能够找到致密包围模型的长方体包围盒;且一般情况下,这样的长方体能够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、对称性等准则来检测摆正方向^[1,2,7,8].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加以筛选,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,以此作为依据来探测模型的摆正方向.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^[3].但这些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检测的效率.文献[9]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检测摆正方向.但该方法对于对称性不明显的模型不太适用.文献[5]则提出对模型进行简化并计算其凸包的方法,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[4]中指出,模型沿着其摆正方向的投影可以看作一个低秩矩阵.由此,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.与 Jin 等人的思路类似,Wang 等人在文献[10]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,他们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.文献[11]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.文献[12]认为:模型沿着摆正方向的投影面积较小,可据此得到一些候选的摆正方向,再依据文献[3]的方法来确定摆正方向.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.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检测摆正方向.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、人眼一般习惯于从一个规范视点(canonical view)来观察模型^[6]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,不必像已有检测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.

1.2 视点评分方法

关于视点选择的工作有很多,其中的主要任务是对采样的视点进行评分,度量它们各自观察模型的效果,以找到好的视点,可观察到模型尽可能多的内容.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,如模型投影面积或者可见度^[13]、可见轮廓^[14]、深度^[6,15,16]、表面曲率^[17]等.这些方法计算简便、速度快,但它们主要依赖一些基本的几何信息,对于模型内容的反映不是很好,在许多情况下不太适用.为此,许多工作研究度量一些视觉感知特征来提高视点评分的有效性^[18-20],如特征显著度(mesh saliency)^[21]等.进一步地,一些工作研究度量一些语义信息来进行视点评分^[22],甚至通过用户学习来进行统计分析模型特征与人眼观察习惯之间的关系^[23].这些基于语义的方法一般依赖于学习样本的情况,计算复杂,适用性不高.

我们的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,我们的方法也就具有了很好的适应性,便于处理各类模型.我们注意到:有些视点选择的工作是首先进行模型摆正方向的确定,以提高好视点寻找的效率^[5].这与我们工作的方向正好相逆.因此,目前还未看到基于视点选择来计算摆正方向的工作.

2 基本思路

2.1 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.如前文所述,底部的内容一般比较贫乏,因而相比于其他视点,底部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有很多,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)^[12,13]、视点熵(VE)^[24]、平均曲率(MC)^[17,25]和网格显著度(MS)^[21].然而,正如文献[23]中得出的结论所述:在选择最优观察视点时,没有一种视点选择方法适用于所有种类的模型.从图 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(如图 1(a)所示).然而在它们所得到的最差视点中,却往往包含观察底部的视点(如图 1(b)中的实线框图所示).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,在某些情况下,模型顶部的视点观察质量可能会比底部视点更差(如图 1(b)中的虚线框图所示).这也容易从自然界中找到例证,比如图 1 中的三脚椅模型,由于顶视角只能观察到椅面,特征很少,被大多数方法评价时给出最低分.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.

我们对 PSB 数据库^[2]中的模型进行了考察,在表 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 1 第 1 列中可以看出,这些方法能够正确地将底部视点识别为最差视点的比例达 40%~60%左右.这表明,最差视点有较大的可能性就是模型底部视点.从表 1 的最后一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,虽然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到这个集合中.

Table 1 Percentage of the base view can be judged as the worst view using different viewpoint scoring methods

表 1 不同视点评分方法判断底部视点为最差视点的比例

方法	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

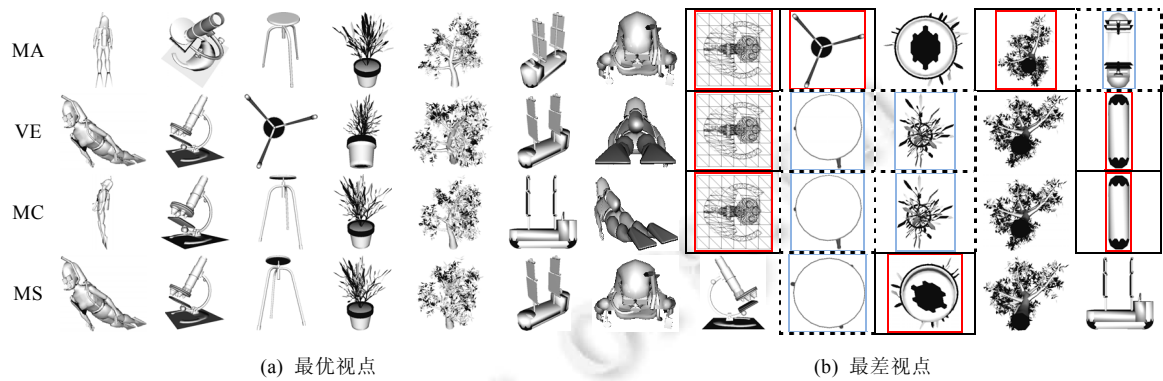


Fig.1 Best and the worst views of different types of models using 4 viewpoint scoring methods

图 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

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加以合并.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(见表 1)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,我们的方法实现中将不使用这两种视点评分方法.对于 VE 方法,它考虑了视点可见信息量,故一般情况下,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,反而容易获取较为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,我们的方法在实现过程中主要将 MA 和 VE 两种视点评分方法相结合,将它们分别得到的底部视点候选集进行合并,以便能够处理更多种类的模型.

2.2 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^[3-5,10-12],它们大多是基于文献[3]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.相比较而言,本文得到的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^[10]),因此,我们可用少量准则进行有效筛选.此外,本文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 1 所示的潜水员模型,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,本文使用两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[6]的研究结论:规范视点₁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(图 1(a)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 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),与模型摆正方向(竖直向上)的夹角接近 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.

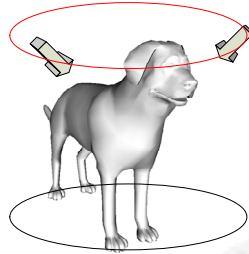


Fig.2 Canonical view of the dog

图 2 “狗”模型的规范视点示意图

3 方法实现

本文的模型摆正方法基本流程如图 3 所示.

- 1) 载入模型后(如图 3(a)所示),在模型的观察球面上均匀采样一些点作为候选视点;
- 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(如图 3(b)所示);
- 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(如图 3(c)所示);
- 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,即可得到摆正方向(如图 3(d)所示).

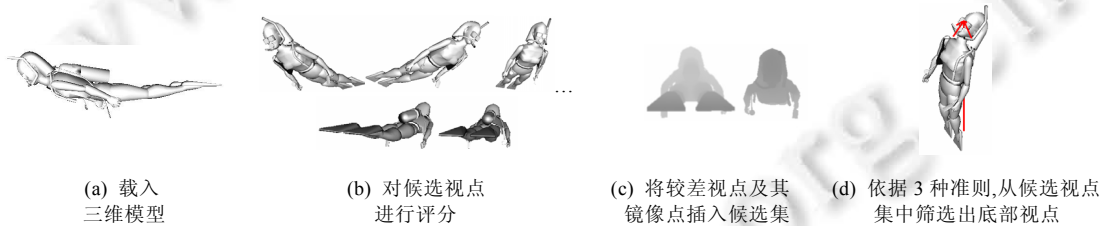


Fig.3 Main process of getting upright orientation of the model

图 3 本文对模型摆正的主要过程

接下来,我们对主要过程分别进行详细介绍.

3.1 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,无需人工干预.

3.1.1 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,一般便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 $Q_r(w)$ 的计算公式如下所示:

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

这里, R_f 表示 3D 网格模型中在 w 处可见的面片 f 的面积; $p_f = \text{abs}(V \cdot N_f)$ 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.

3.1.2 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更多,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 $Q_l(w)$ 的计算公式如下所示:

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

这里, R 表示模型在 w 处可见的面片投影面积之和,其他参数与 MA 评分公式相同.

3.2 视点候选集生成

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^[26],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.若阈值比较大,很多视点可放在一个类中,则分类的数目就比较少;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 $(S_h - S_l) \times 5\%$,一般都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,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,均认为其为差视点.通过实验,该阈值也和聚类阈值一样设定为 $(S_h - S_l) \times 5\%$,一般都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,我们在以代表点为中心的一定范围内对其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该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,将由 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.

3.3 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

$$B(v)=w_u \times R_u(v)+w_s \times R_s(v)+w_h \times R_h(v) \quad (1)$$

这里, w_u, w_s, w_h 都是权值.考虑到本文方法的适用对象要求尽可能地广泛,因此在测试时,为公式(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,我们介绍本文使用的这3个筛选准则具体的计算方法.

- 视点合适度

如第3.2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4使用这一准则作为判定条件得到了正确的摆正结果,图4(a)使用VE视点评分方法得到了差视点和好视点,它们之间的方向夹角更接近于 135° ,故而将其选定为底部视点,得到了正确的摆正结果(如图4(b)所示).视点合适度计算公式为 $U(v)=1-abs(\vec{V}_g \cdot \vec{V}-\cos(135))$,这里, \vec{V}_g 表示生成这个候选视点所使用的视点评分结果中的最好视点单位向量,而 \vec{V} 表示视点 v 的观察方向单位向量.

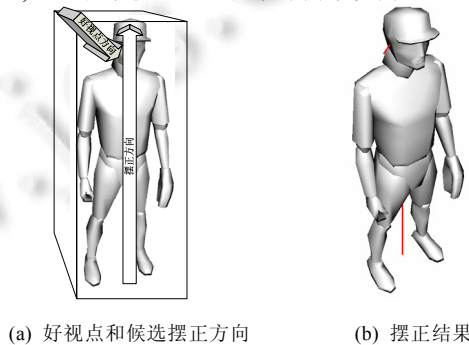


Fig.4 Viewpoint fitness metric

图4 视点合适度准则

- 稳定性

如果一个视点看到的模型,其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使用了这一准则,从图5(a)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1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(如图5(b)所示).

- 支撑性

在某些情况下,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,可以很好地支撑模型,形成摆正方向.图6使用了这一准则,从图6(a)可以看出:第2个候选底部有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2个候选底部用于模型摆正是准确的(如图6(b)所示).

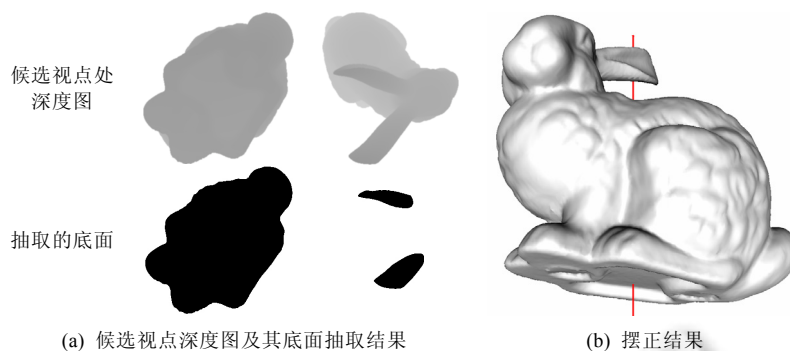


Fig.5 Stability metric

图 5 稳定性准则

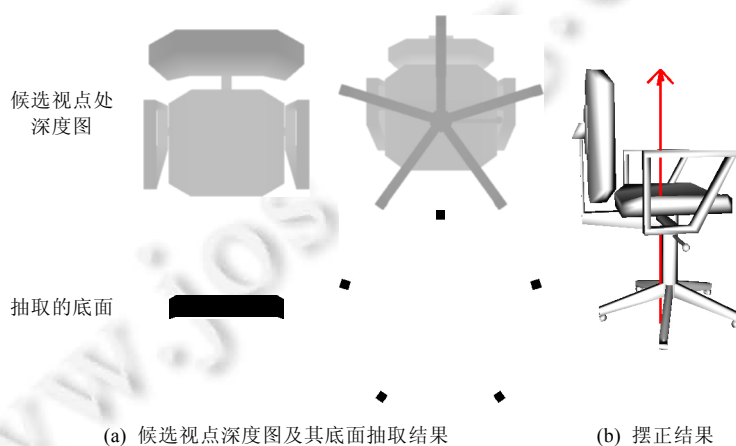


Fig.6 Supporting metric

图 6 支撑性准则

在计算稳定性和支撑性准则时,本文使用基于深度图像的方法,以获得更快的计算速度,伪代码如下:

输入:投影深度图 $depthMap$;

输出:支撑性得分 S ,稳定性得分 H .

在 $depthMap$ 中:

找到投影图像中心点 C ;

计算 C 到最远投影边界像素点的距离 D ;

计算投影图像区域的面积 T ;

计算到视平面距离最近的深度值 Z ;

过滤,删除深度值大于 $Z+e$ 的像素,得到底座图; //这里, e 是一个冗余量,一般设为 $(Z_{\max}-Z_{\min})\times 10\%$

找到所有连通区域 $A_1\sim A_n$ 形成的图像块;

计算每个图像块的面积 $T_1\sim T_n$;

计算 $A_1\sim A_n$ 图像块的质心 C_a ;

稳定性得分:

$$S = \sum_n (T_i / T);$$

支撑性得分:

$$H = 1 - |C_a - C| / D.$$

由于公式(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 7 所示,利用本文的视点评分方法可以得到对称的两个候选视点(图 7(a)、图 7(b)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[3]得到了错误的摆正结果(如图 7(d)所示).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(如图 7(c)所示).

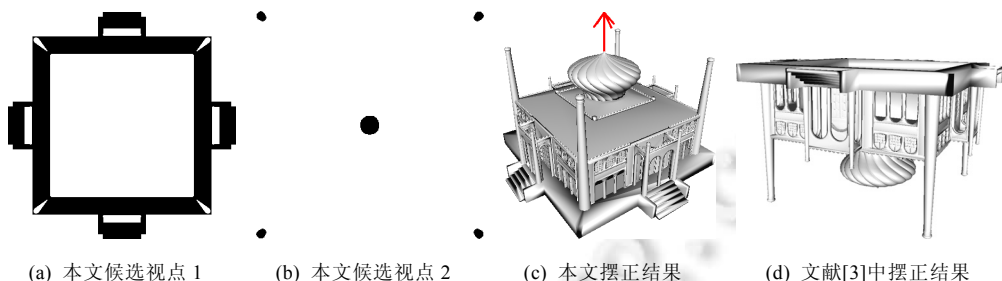


Fig.7 Candidate views of masjid model, and the upright orientation obtained using Eq.(1) compared with Ref. [3]

图 7 清真寺模型的候选视点和使用公式(1)得到的摆正结果以及与文献[3]的比较

4 实验分析

4.1 视点评分适用性

我们在 PSB 模型库^[2]上验证了本文提出的模型摆正方法,PSB 模型库包含了 1 815 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)和杂项模型这 3 种,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;自然模型包括两个子类:动物和植物.

在 PSB 模型库上得到我们的方法针对不同类型模型的平均摆正准确度,如图 8 所示.从图中可以看出:对于大多数类型的模型,我们的方法的准确度已经达到 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.比如后文图 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.

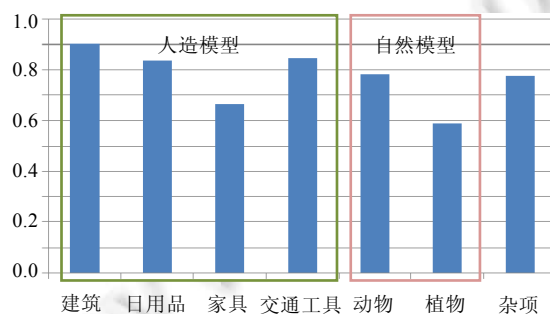


Fig.8 Average upright accuracy rating of different types of models in PSB database using our method

图 8 本文对 PSB 数据库中不同种类模型的平均摆正准确度

4.2 对比实验

我们也把本文得到的结果与目前最好的方法进行了比较.得益于本文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,本文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,本文使用的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,可以较为准确地得到模型最终底部视点.

利用本文方法得到PSB数据库中所有模型平均摆正准确率为79%左右,要优于文献[10]中平均70%左右的准确率,与文献[5]中的摆正准确率相当.然而文献[5]只针对人造模型,并且统计样本只有120个.与文献[3]中的87.5%的准确度相比,我们的方法还有一定差距,但文献[3]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.对于采用其他模型摆正方法主要处理的人造模型^[3-5,11],本文方法可以达到80%以上的准确度;对于其他方法难以处理的自然模型,采用本文方法,其平均准确度也能达到68%以上.

本文方法与其他两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见表2(文献[5]和文献[11]公开了这3组模型的准确率),本文严格按照PSB数据库的模型分类进行实验,而两组对比方法增删了部分模型.对于赛车类模型,本文方法的摆正准确率与文献[11]相当,都要高于文献[5];对于显微镜类模型,本文方法可以达到100%的摆正准确率,远高于其他两类方法;对于花瓶类模型,本文方法的准确度要差一些,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,利用本文的3个优化参数,很难将花瓶的顶部和底部区别开来.

Table 2 Average upright accuracy rating of 3 groups of modes
表 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[5]	50
	15	文献[11]	73.3
	14	本文方法	73.3
显微镜	5	文献[5]	80
	6	文献[11]	83.3
	5	本文方法	100
花瓶	20	文献[5]	95
	12	文献[11]	91.7
	22	本文方法	86.4

此外,对于文献[3]中难以处理的模型(如图9(a)所示),利用本文方法可以得到令人满意的结果(如图9(b)所示);另外一个对比示例如图7所示.在文献[4]和文献[11]中很难处理的模型(如图9(c)所示),本文方法也能够得到准确的结果(如图9(d)所示).

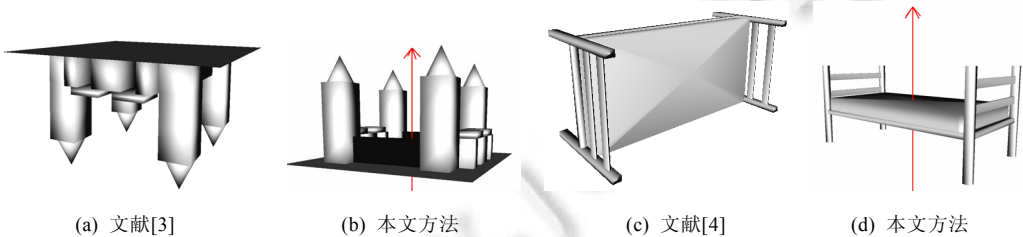


Fig.9 Proper upright orientation of some models, which cannot be handled by other methods, can be obtained using our method

图 9 本文方法能摆正其他方法失效的模型

4.3 效率分析

本文的实验在一台配置为 Intel Xeon E5405 CPU(2.0GHz,2CPUs),4GB 内存的台式机上进行.本文方法对不同复杂度模型的处理时间见表3,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,由于本文使用基于图像的方法从中确定底部视点,与模型复杂度无关,这个过程所需时间很少,比如在我们的实验中一般只需要200ms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8s以内.

本文并未使用并行优化,也未使用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.与其他模型摆正方法的效率对比情况可见表4,从表中可以看出:对于复杂度近似的模型,本文方法效率更高,即便我们所使用的CPU还是略差于其他方法所使用的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.

Table 3 Computing time of different complexity models using our method (in seconds)

表3 本文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

Table 4 Efficiency comparison of different methods

表4 不同方法的效率比较

摆正方法	平均耗时	CPU 频率(GHz)
文献[3]	十几秒	2.13
文献[4]	十几秒	2.5
文献[5]	上百秒	2.33
文献[11]	几十秒	2.6
文献[10]	几十秒	3.3
本文方法	小于8秒	2.0

注:本文方法实验所用CPU能力弱于其他方法所用的CPU

5 结论及未来工作

本文通过视点评分来指导模型候选底部的确定,进一步引入人眼观察习惯及稳定性度量来提高计算结果的准确性.从实验结果来看:本文提出的方法速度快,可以处理不同种类的模型,并且对其他方法中难以处理的模型也能得到较好的结果.考虑到不同视点评分方法适用的模型种类可能存在差异,我们将多种不同的方法加以结合来处理,文中对MA和VE相结合的摆正策略进行了验证,取得了较好的实验结果.将来的工作包括:引入新的更加恰当的视点评分方法以及选取更多模型数据库参与实验.

尽管绝大多数情况下利用本文方法得到的视点候选集中包含了模型的真实底部视点,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,本文方法依旧无法处理.如图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.

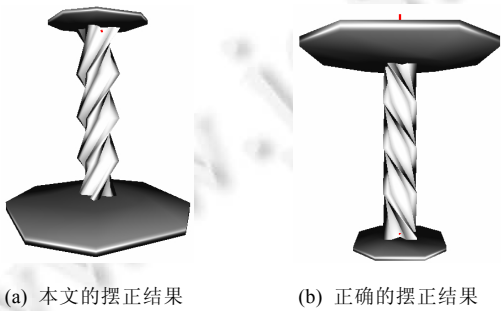


Fig.10 Failure result of our method

图10 本文方法失效的模型

此外,本文中更主要的失效情况是颠倒错误,这表明本文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),而本文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。

References:

- [1] Bustos B, Keim D, Saupé D, Schreck T, Vranic D. An experimental comparison of feature-based 3D retrieval methods. In: Proc. of the 2nd Int'l Symp. on 3D Data Processing, Visualization and Transmission (3DPVT 2004). 2004. 215–222. [doi: 10.1109/TDPVT.2004.1335197]
- [2] Shilane P, Min P, Kazhdan M, Funkhouser T. The Princeton shape benchmark. In: Proc. of the Shape Modeling Applications. 2004. 167–178. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=1314504> [doi: 10.1109/SMI.2004.1314504]
- [3] Fu H, Cohen-Or D, Dror G, Sheffer A. Upright orientation of man-made objects. In: Proc. of the ACM SIGGRAPH 2008. Los Angeles, 2008. 1–7. [doi: 10.1145/1399504.1360641]
- [4] Jin Y, Wu Q, Liu L. Unsupervised upright orientation of man-made models. *Graphical Models*, 2012,74(4):99–108. [doi: 10.1016/j.gmod.2012.03.007]
- [5] Lin CK, Tai WK. Automatic upright orientation and good view recognition for 3D man-made models. *Pattern Recognition*, 2012, 45(4):1524–1530. [doi: 10.1016/j.patcog.2011.10.022]
- [6] Blanzfíl V, Btlthoff HH. What object attributes determine canonical views. *Perception*, 1999,28(5):575–599.
- [7] Vranic DV, Saupé D, Richter J. Tools for 3D-object retrieval: Karhunen-Loeve transform and spherical harmonics. In: Proc. of the 2001 IEEE 4th Workshop on Multimedia Signal Processing. 2001. 293–298. [doi: 10.1109/MMSP.2001.962749]
- [8] Zaharia TB, Preteux FJ. 3D versus 2D/3D shape descriptors: A comparative study. In: Proc. of the SPIE Conf. on Image Processing: Algorithms and Systems III, Vol. 5298. San Jose, 2004. 47–58. [doi: 10.1117/12.533092]
- [9] Chaouch M, Verroust-Blondet A. Alignment of 3D models. *Graphical Models*, 2009,71(2):63–76. [doi: 10.1016/j.gmod.2008.12.006]
- [10] Wang W, Liu X, Liu L. Upright orientation of 3D shapes via tensor rank minimization. *Journal of Mechanical Science and Technology*, 2014,28(7):2469–2477. [doi: 10.1007/s12206-014-0604-6]
- [11] Jiang B, Zeng M, Liu XG. Upright orientation detection 3D man-made objects principal axis. *Journal of Computer-Aided Design & Computer Graphics*, 2013,25(8):1099–1106 (in Chinese with English abstract).
- [12] Johan H, Li B, Wei Y, Iskandarsyah. 3D model alignment based on minimum projection area. *The Visual Computer*, 2011,27(6-8): 565–574. [doi: 10.1007/s00371-011-0590-y]
- [13] Plemenos D, Benayada M. Intelligent display in scene modeling. In: Proc. of the Int'l Conf. on New Techniques to Automatically Compute Good Views (GraphiCon). 1996. 1–5. [doi: 10.1007/978-3-642-01259-4_2]
- [14] Feldman J, Singh M. Information along contours and object boundaries. *Psychological Review*, 2005,112(1):10. [doi: 10.1037/0033-295X.112.1.243]
- [15] Stoev SL, Strasser W. A case study on automatic camera placement and motion for visualizing historical data. In: Proc. of the Visualization 2002 (VIS 2002). IEEE, 2002. 545–548. [doi: 10.1109/VISUAL.2002.1183826]
- [16] Cao WG, Hu P, Li H, Lin ZK. Canonical viewpoint selection based on distance-histogram. *Journal of Computer-Aided Design & Computer Graphics*, 2010,22(9):1515–1521 (in Chinese with English abstract).
- [17] Polonsky O, Patané G, Biasotti S, Gotsman C, Spagnuolo M. What's in an image. *The Visual Computer*, 2005,21(8):840–847.
- [18] Sbert M, Feixas M, Castelló P, Chover M. Information theory tools for viewpoint selection, mesh saliency and geometry simplification. In: Plemenos D, Miaoulis G, eds. Proc. of the Intelligent Computer Graphics 2009. Berlin, Heidelberg: Springer-Verlag, 2009. 41–61. [doi: 10.1007/978-3-642-03452-7_3]
- [19] Yang LM, Han HL. Viewpoint selection benchmark design and application. *Ruan Jian Xue Bao/Journal of Software*, 2010,21: 165–172 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/10018.htm>
- [20] Miao YW, Wang HJ, Shou HH. Best viewpoint selection driven by relief saliency entropy. *Journal of Computer-Aided Design & Computer Graphics*, 2011,23(12):2033–2039 (in Chinese with English abstract).

- [21] Lee CH, Varshney A, Jacobs DW. Mesh saliency. ACM Trans. on Graphics, 2005,24(3):659–666. [doi: 10.1145/1073204.1073244]
- [22] Mortara M, Spagnuolo M. Semantics-Driven best view of 3D shapes. Computers & Graphics, 2009,33(3):280–290. [doi: 10.1016/j.cag.2009.03.003]
- [23] Secord A, Lu J, Finkelstein A, Singh M, Nealen A. Perceptual models of viewpoint preference. ACM Trans. on Graphics, 2011, 30(5):1–12. [doi: 10.1145/2019627.2019628]
- [24] Vázquez PP, Feixas M, Sbert M, Heidrich W. Viewpoint selection using viewpoint entropy. In: Proc. of the Vision Modeling and Visualization Conf. 2001. 273–280. <http://www.cs.ubc.ca/~heidrich/Papers/VMV.01.pdf>
- [25] Alliez P, Cohen-Steiner D, Devillers O, Lévy B, Desbrun M. Anisotropic polygonal remeshing. In: Proc. of the ACM SIGGRAPH 2003. San Diego, 2003. 485–493. [doi: 10.1145/1201775.882296]
- [26] Felzenszwalb P, Huttenlocher D. Efficient graph-based image segmentation. Int'l Journal of Computer Vision, 2004,59(2): 167–181. [doi: 10.1023/B:VISI.0000022288.19776.77]

附中文参考文献:

- [11] 姜玻,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,2013,25(8):1099–1106.
- [16] 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,2010,22(9):1515–1521.
- [19] 杨利明,韩红雷.视点选择 Benchmark 的设计与应用.软件学报,2010,21:165–172. <http://www.jos.org.cn/1000-9825/10018.htm>
- [20] 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,2011,23(12): 2033–2039.



韩红雷(1980—),男,山西太原人,博士,讲师,主要研究领域为计算机图形学,计算机动画技术.



华淼(1987—),男,博士,主要研究领域为图像处理,真实感绘制.



王文成(1967—),男,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为计算机图形学,虚拟现实,可视化计算.