

基于自适应几何分析的高效交互式导航

郝竹明 黄 惠

(中国科学院深圳先进技术研究院 深圳 518055)

摘 要 三维场景的虚拟导航需要同时保证相机视角的光滑性和智能性, 在算法设计上极具挑战。文章研发了一套实时自动生成光滑连续且高效的相机路径导航系统。首先通过线下模型几何分析和语义分析, 包含分析场景中建筑物模型、道路模型以及非建筑的重要模型等, 求得模型重要性值; 其次是自适应地获取路线采样点以及高效存储采样点的视角图片; 最后是利用动态规划的方式求出每个采样点的最佳视角并光滑地连接这些采样点形成相机运行轨迹。4 个不同三维场景的实验结果表明了所提算法的高效性和智能性, 同时文章所进行的用户调查也充分反映了所提方法具有的明显优势。

关键词 导航; 几何分析; 重要性值; 自适应

中图分类号 TP 391 **文献标志码** A

Efficient Interactive Navigation Based on Adaptive Geometry Analysis

HAO Zhuming HUANG Hui

(Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

Abstract Virtual navigation in 3D scenes requires both smooth and intelligent camera control, which is a very challenging task. In this paper, a real time navigation system that could automatically generate an effective and smooth camera flyby trajectory was presented. Firstly, offline geometric and semantic features of models were analyzed in the given 3D scene, including buildings, roads and other visible parts, to evaluate the importance of models. Secondly, the path from origin to destination was sampled adaptively, and visual images from different views on each sampled point were stored effectively. Finally, the best view on each sampled point was extracted and connected smoothly to get the camera flyby trajectory. The experiments done in four different 3D scenes clearly demonstrate the efficiency and intelligence of the proposed algorithm, and the user investigation also verifies its superiority over the existing methods.

Keywords navigation; geometry analysis; importance value; adaptively

收稿日期: 2015-06-13 修回日期: 2015-09-11

基金项目: 国家自然科学基金项目(61379091); 科技部 863 项目(2015AA016401); 广东省科技厅项目(2014B050502009); 深圳市科技创新委项目(CXB201104220029A, JCYJ20140901003938994, CXZZ20130322162845315)

作者简介: 郝竹明, 硕士研究生, 研究方向为计算机图形学; 黄惠(通讯作者), 研究员, 博士研究生导师, 研究方向为计算机图形学, E-mail: hhzhiyan@gmail.com。

1 引言

三维虚拟场景漫游和相机控制在最近十年内得到广泛的研究^[1-4]。各种各样的导航平台相继产生, 如谷歌地球, smart3D 浏览工具等。这些平台提供了通过卫星、航拍、手工等方式生成的大量三维模型, 同时集成了简单的虚拟漫游方式。这些漫游方式大多数是基于匀速、固定俯仰角的方式, 且每天都还在产出大量的三维场景模型, 这使得研究三维场景的虚拟导航成为可能, 也极其紧迫。

通常一个虚拟的三维场景包含了以下几个部分: 三维建筑物模型、雕塑等非建筑物模型、道路模型以及高速路地形模型。如果对这些模型不做任何分析, 而是简单地让相机沿路匀速运行, 将会导致虚拟漫游既耗时又低效。Grabler 等^[5]在研究地图导航的工作中就提出了一套很有用的三维场景中模型的分析方法, 包括分析了建筑物和路型等信息。在 Polonsky 等^[6]的工作中, 对场景分析所用的描述子包含了诸如建筑物表面可见性、目标显著性、曲率、外表轮廓和拓扑复杂度等。

现存许多方法是关于视点选择的, 其中基于信息论理论的有 Vazquez 等^[7]提出的作为度量具体视角下的信息量的视点熵概念。这种方法主要是利用了球体方向下建筑物投影面积比例: 当投影面积平均分布时, 熵最大。除了考虑视点下的几何特征外, 部分研究^[8,9]还考虑了语义特征, 如建筑物风格、结构类型或建筑位置, 从而更好地改进了计算最佳视点的方式。

本文关于建筑物分析和视点选择有借鉴前人研究的思路, 但又有所不同: 文中考虑到了一些具体描述子, 比如独特性, 而不是单独的单栋建筑的特性分析; 重要性值的加权权重通过学习获得, 不是简单经验观察。

相机控制一直是一个很具挑战性的问题: 需要解决视点计算、路径规划以及用户交互等任务。

Christie 等^[10]曾对有关虚拟世界的导航方法和动机进行调查研究。Drucker 等^[11]是第一次提出关于虚拟三维博物馆场景自动漫游方法的人, 他们使用了路径规划和图的搜索方法。其他一些方法^[12], 如基于运动规划也被大量研究。此外, 还有很多基于几何角度的相机研究方法, 如碰撞避免^[13,14]、目标可见性^[15]、光滑性^[16]等。但以上这些方法大都是仅关注了相机轨迹的生成, 而忽略了相机运行速度的计算, 而速度本身也是导航中一个很重要的问题。因此, 本文考虑了关于速度的一个解决方案。

在许多三维场景漫游的应用中, 如游戏寻路等, 需要大量用户交互, 但同时又需要自动导航的部分, 这样一个交互的需求给相机控制提出了极大的挑战。怎样在用户新设定的相机姿态下继续自动导航, 目前几乎很少研究涉及这个问题的解决。

综上所述, 本文提出了一种基于几何分析的理想的高效交互式三维导航方法, 希望能够做到实时自动生成光滑连续且高效的相机导航路径。为此, 需要做到以下几点: 一是如第 2 节所述的一种三维场景几何特征的分析方法, 能够有效分析出哪些场景模型是关键, 哪些相对没那么关键; 二是怎样高效利用这些分析结果进行实时运算, 得出当前采样点的最佳视点, 即第 3 节提出的一种基于全景图的采样存储和分析方式; 最后一点即第 4 节所描述的利用得到的最佳视点进行动态规划的最优求解得出光滑连续的相机路径。

文章最后在四种三维虚拟场景上证明了所提解决方法的有效性。同样也把该方案应用到了谷歌地球、智能三维浏览工具来生成真实卫星或航拍的三维场景下的导航。

2 场景几何特征分析

本研究开发了一个自动化的系统来评估场

景的几何特征重要性。所提导航系统的输入包含了城市场景中的三维几何模型(模型的经纬度坐标)、道路的几何模型和路型信息(如图1所示)。其中,路型信息包括普通街道、主干道、干线、匝道或者高速路。另外,交通图将道路分成四个交通层次;地形平面图将场景分成水路、公园和地面。对于关键的地标建筑或者标识性非建筑模型(雕塑、桥梁等)需要建立一定规模的数据库,这些数据可根据模型经纬度以及名字从网络获取。

分别对以上信息进行分析,最后将分析结果渲染成RGB图。其中,暖色调代表重要的信息;冷色调代表不重要信息。

2.1 建筑物特征分析

本文基于Appleyard^[17]提出的一个调查结果来决定分析建筑的哪些特征,这个调查主要是咨询了调查者记住了沿路的哪些建筑。

2.1.1 几何特征

在某一块区域内,如果建筑物有一些区分性很强的特征(如颜色、形状等),被调查者将会很容易记住。本文主要分析了四种最显著的几何特征:高度、不规则性、二面角、独特性。以下将分别介绍这四种特性以及如何将其关联。

(1) 高度

一些比较高的建筑从远处看会更加显眼,像摩天大厦这样的建筑往往都是一个城市的地标。

因此使用了建筑物高度作为一个几何特征。越高的建筑得到的分值越大。我们在给定的邻域内(本文统一采用150 m作为邻域半径),通过归一化的方式求得该项的值。

$$S_h(b) = \frac{h(b)}{\max_{c \in \Omega} h(c)} \quad (1)$$

其中, Ω 是给定邻域内建筑物集合; $h(b)$ 指建筑物***b***的高度; $S_h(b)$ 指建筑物***b***在高度一项的得分。

(2) 不规则性

许多建筑物都有一个简单的、规则的矩形形状。一栋建筑物的形状和规则的标准形状相离越远,这个建筑物就越显著。使用了这样一个几何特性做为度量之一,度量了建筑物的包围盒体积和真实建筑物的体积^[18]的比例,公式如下:

$$S_r(b) = 1 - \frac{V(b)}{V(BB(b))} \quad (2)$$

其中, $V(b)$ 指建筑物***b***的体积; $BB(b)$ 指建筑物的包围盒; $S_r(b)$ 指建筑物***b***在不规则性一项的得分。

(3) 二面角

二面角一直是一个很好的表征建筑物结构复杂度的特征。如图2所示,同样是接近标准包围盒的建筑物结构,高度虽然也相似,但是由于左图的二面角和更大,该建筑显然更容易被用户所记住。

这一项度量了根据二面角所在边的边长作为权重的加权二面角和。

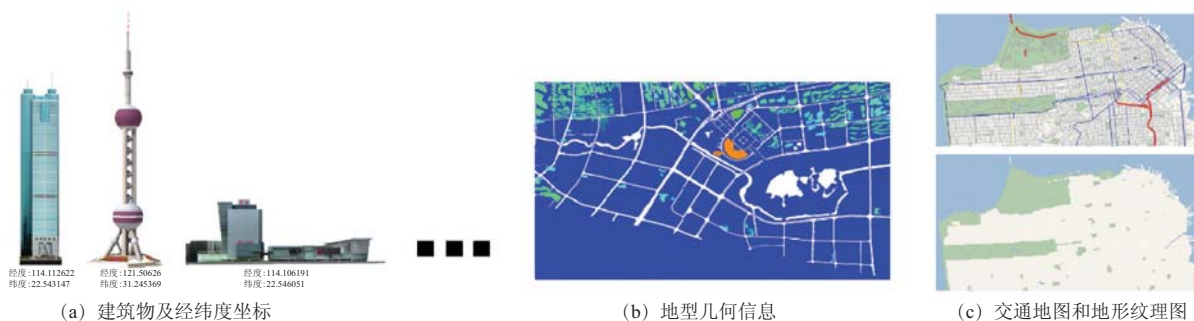


图1 系统的输入数据

Fig. 1 The input data of the system

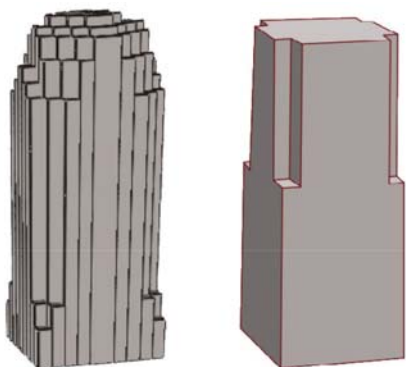


图2 二面角图示

Fig. 2 The diagrams of dihedral angle

$$S_a(b) = \sum_i \text{dihedral}(i) \text{len}(i) \quad (3)$$

其中, $\text{len}(i)$ 指第 i 个二面角所对应的边长; $\text{dihedral}(i)$ 指第 i 个二面角的角度; $S_a(b)$ 指建筑物 b 在二面角属性一项的得分。

(4) 独特性

如果一个建筑物和周边的相邻建筑物有很大的区别, 将会很显而易见地成为漫游中的一个标志性建筑。为了计算这样一个特征, 使用如下一个打分项:

$$S_u(b) = \frac{|x - \text{med}(x)|}{\text{med}(|x - \text{med}(x)|)} \quad (4)$$

其中, x 是指前面三项的值; $\text{med}(x)$ 是指中值。中值所采用的邻域和以上几项类似。

最终建筑物几何特征得分通过线性加权求和得到:

$$S_{geo} = w_1 S_h + w_2 S_r + w_3 S_a + w_4 S_u \quad (5)$$

为了得到这些权重, 随机抽取了场景中一些地标建筑, 并在网上(主要来源主流网站, 如雅虎、大众点评等)取得这些建筑的介绍以及一些评分, 最终通过上述权重加权来拟合这些评分, 从而得到理想的权重值。

2.1.2 位置结构特征

一些认知心理学的研究表明, 游览者更加容易关注到道路交叉口的建筑物或者城镇广场周边

的建筑^[19]。这样的空间位置使得这些建筑在导航中起到非常关键的作用, 因此本文将其作为一项重要度量。

(1) 交叉路口的结构

越是重要交叉路口的建筑越比其他建筑显著。因此, 需给所有道路的每一个交叉路口打一个分: 打分机制更倾向于给分支越多的交叉路口越高的得分; 最终使用一个高斯加权因子, 基于路口到建筑的距离来给建筑物一个得分; 对于每一栋建筑累加周边所有交叉路口对其的得分。

道路通常被分为一些离散的类型集合。在本文中, 总共分为五类, 主要是普通街道、主要街道、干道、匝道和高速路。这些分类可以根据经纬度从谷歌地球的 API 获取到。根据经验估计, 给予对应项的权重分布分别为 0.25, 0.5, 0.75, 0.8, 1.0。这样也就给了交叉路口一个得分评估。

(2) 广场结构

为了标识建筑物是坐落在广场周边, 首先从地形图上提取出广场和计算出广场的重要性值。这样一个重要性值使用一种类似于交叉路口处理的方式转化成建筑物的得分。

作为一个广场区域, 正常从地图上截取是没有一个很明显或精确的边界。所以我们根据建筑物密度谷值为中心, 采用均值移动算法, 不断调整密度最低点, 即广场中心点。从而给在这个广场周边的建筑一个得分, 见图 3 所示的移动过程。

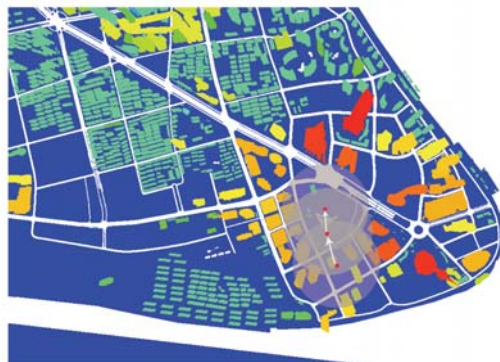


图3 广场中心点移动示意图

Fig. 3 Illustration of square center movement

2.2 非建筑物模型数据驱动分析

场景中一些非建筑物模型，包括了雕塑、重要路标、桥梁等模型，主要采用数据驱动的处理方式，即提前将重要小件离线标定出来，加入数据库。以这些小件的坐标和名称在网络上搜索其相关属性，然后将这些属性按特定结构存储起来。所保存的属性主要包含主流旅行网站给予这个模型的打分，以及这个模型坐落的区域、地址、所属类型、用户所倾向的得分。而大多数非建筑物模型是没有网络数据的，此时一律打分为1分。

图4所示为非建筑物模型存储结构和数据来源。得到这样一份存储结构后，根据其所属区域的建筑物得分，来给出新的关于坐落区域的打分。

$$S_d(m) = \frac{bldaffect(m)}{\max_{n \in \Omega} bldaffect(n)} \quad (6)$$

其中， Ω 指所有非建筑物模型集合； $bldaffect(m)$ 指建筑物对非建筑物模型的影响得分，建筑物的影响权重，是用建筑物的得分高斯加权作为模型的影响得分。

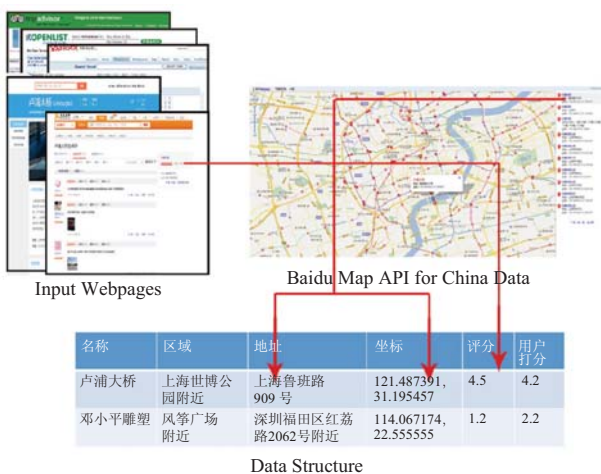


图4 非建筑物模型存储结构和数据来源

Fig. 4 Storage structure and data source for non-building models

3 视点的重要性值计算和存储

计算出每个单独模型的重要性得分(图5)，作为建筑物加权得分。然后把场景纹理模型根据重要程度渲染成对应的颜色模型。其中，暖色调代表重要的模型，冷色调则代表不重要的模型，如图6所示，也即得到了如下的一个模型变换。

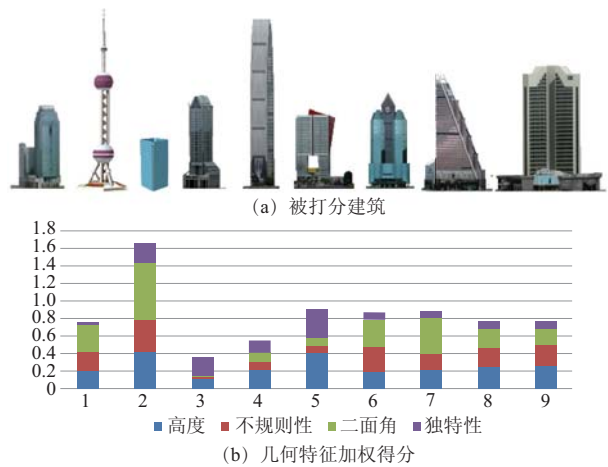


图5 建筑物几何特征评估得分

Fig. 5 The building geometrical characteristics assessment score

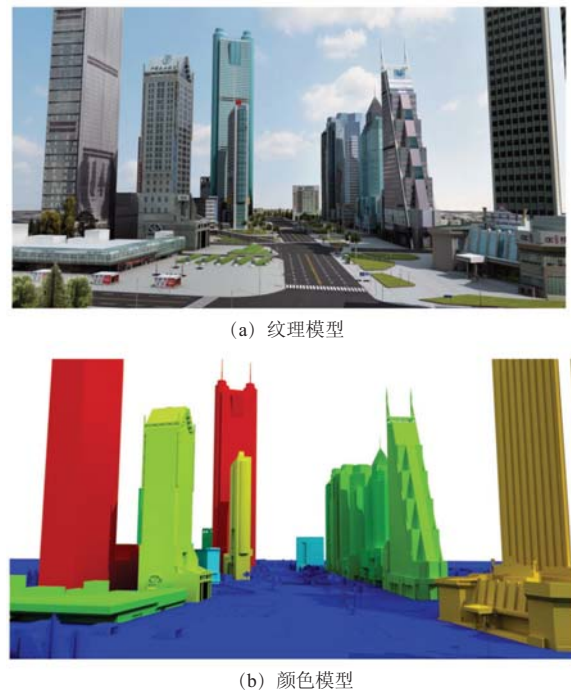


图6 纹理模型到颜色模型的映射

Fig. 6 Mapping between texture model and the color models

下面的工作将分为: 给定起始点和终点之后, 进行自适应采样路径点; 对于每个路径点, 计算各个高度的场景重要性值; 离线保存这些场景, 以保证能够实时响应用户的交互。

3.1 自适应采样

由于大多数三维场景组成包含了范围较小的城市区域以及范围较大的高速路区域。如果完全均匀对这些点采样, 过于稠密将会影响计算效率, 而过于稀疏将会影响城市区域的计算精确度, 所以需要有一个自适应的采样方式。同时, 在每个采样点, 由于要覆盖相机的 6 个自由度, 高度上采用了非线性采样, 而由于角度范围是固定的, 且视场角有一个固定覆盖值, 所以朝向和俯仰角采用了均匀采样。相机翻滚角这个自由度则默认不变。

3.1.1 路径点的自适应采样

如图 7 所示, 红色点即为自适应的采样点, 此处自适应采样主要考虑某个点附近建筑物的密集程度以及路型的复杂度这两个特征。



图 7 路径点的自适应采样

Fig. 7 Adaptively sampling path points

初始阶段先均匀密集采样整个路线, 对第一次的均匀采样点, 可以轻易计算出给定范围内建筑物的密集度, 同时根据第 2 节对于不同路型会有不同的得分, 从而得出每个均匀采样点的价值:

$$C(wp) = w_d d(wp) + w_r r(wp) \quad (7)$$

其中, $d(wp)$ 指的是采样的建筑物密度值; $r(wp)$ 指的是路型的加权值; w_d 和 w_r 分别为对应权重, 文中对应采用 0.7 和 0.3 这两个系数。

$$d(wp) = \sum_{b \in \Omega} \left(1 - \frac{\|b-p\|}{150} \right) \quad (8)$$

其中, Ω 指该采样点 150 m 范围内的建筑物集合; b 为单个建筑物的位置坐标; p 为采样点的位置坐标。

在计算出每个均匀采样点的价值后, 开始自适应地调整新的采样间隔, 保证价值越大的点所在的区域采样越密集。先累计求出所有采样点的价值总和 (设为 V)。将 V 值等分: 假设要采样 1000 个点, 那么从起始点开始累加均匀采样点的价值, 累加到 $V/1000$ 的时候, 保留该点; 如果超过 $V/1000$, 最后一段根据 V 值的差值按比例截取, 获得一个新的点。所以算法所提的自适应采样的本质是根据重要性值的均匀采样。

3.1.2 非线性的高度采样

由于在三维场景中还有高度这个维度, 而高度的范围是无限的, 但是根据现有研究表明, 高于 1000 m 的高度在虚拟漫游中已经没有太大意义。因此高度取值范围为 10~1000 m。

图 8 所示为在高度方向采用了非线性的采样方式, 具体计算如下:

$$h(i) = 10 \times \frac{1-a^i}{1-a} \quad (9)$$

其中, a 满足 $h(i)=1000$, 即可求得 $a=2.84$ 。求得 a 之后可以很方便地求出对应的 5 个采样高度。

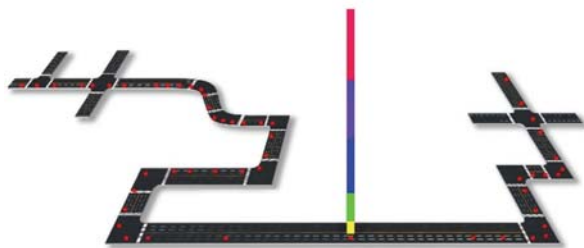


图 8 非线性高度采样

Fig. 8 Nonlinear sampled elevation

3.1.3 朝向和俯仰角的均匀采样

相机除了 x 、 y 、 z 三个位置坐标的自由度外, 还有围绕这三条轴旋转所产生的三个自由度。在本解决方案中侧滚角保持不变, 因为如果侧滚角产生变化, 会导致相机过于灵

活,从而使用户产生眩晕感;偏航角的范围是 $-180^{\circ} \sim 180^{\circ}$,而俯仰角的范围是 $-90^{\circ} \sim 90^{\circ}$ 。相机的视场角在水平和垂直方向设定有一个固定比例,本文采用的是 $60:36$ 。由于在水平方向的视场角是 60° ,故在水平方向分成6等分;同理,垂直方向分成5等分,结果如图9所示。

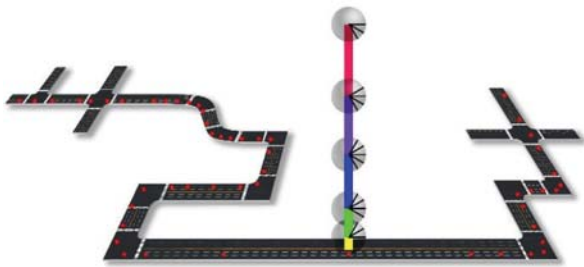


图9 相机俯仰角角度的采样

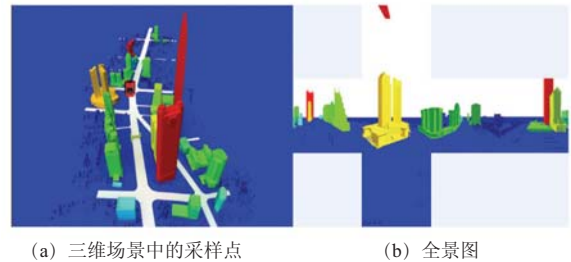
Fig. 9 Pitching angle sample for camera

观察到在海拔最高和最低的位置,仰视和俯视是没有任何意义的;同样在次高和次低的位置,完全的仰视(即仰视角度很大)和完全的俯视(即俯视角度很大)也是没有意义的,所以此处需要根据采样高度去掉一些方向。

3.2 全景图存储

整个自动的采样过程如上两节所述,由于需要满足后续更灵活的用户交互需求,需要动态获得某些视角的重要性值,此时为了响应用户的交互操作,需要提前采样出更多的视角。如果对于插值采样的视角也完全按照之前上一小节的精细处理方案,将导致存储量过于庞大,线下运算过于耗时。

本文采用全景盒的方式来存储视场图片,以视场角为 90° 定点采样六张图,分别为朝前、朝后、朝左、朝右、朝上和朝下。从而利用这六张图片来构成全景图,保证后续动态获得重要性值,而不需要再次花费大量时间来渲染出新视角下的图片求解。这样就可以达到用户要求的实时性需求,因为本算法线下预处理最耗时的部分就是视角采样渲染,几乎占据了90%的时间消耗。图10所示为本文保存的全景立方盒图片。



(a) 三维场景中的采样点 (b) 全景图

图10 全景立方盒图

Fig. 10 Panoramic cube image

3.3 重要性值计算

对于3.1中采样得到的图片,将通过累加其像素值来求得该视角下的重要性值。至此,整条路线上每个采样点都有一系列的重要性值。在累加的过程中考虑到图像的一些显著性特性,比如中心位置、图像分布均匀性等,添加了一些额外的权重。而对于存储的全景图,需要二次截取其对于视场角 60° 的位置,来累计像素求和。

4 相机轨迹优化求解

至此已经可以得到一系列采样点,同时得到了采样点对应的一系列视角下的重要性值。此时只需要以这些视角作为求解的离散值,然后在采样点间线性插值,即可得到最终的连续相机轨迹。

4.1 动态规划的轨迹优化方程

由上一小节,得到如下一个矩阵 M :

$$M = \begin{bmatrix} 0.223 & 0.234 & \cdots & 0.013 \\ 0.324 & 0.353 & \cdots & 0.015 \\ 0.321 & 0.336 & \cdots & 0.016 \\ \vdots & \vdots & \ddots & \vdots \\ 0.281 & 0.276 & \cdots & 0.018 \end{bmatrix}$$

该矩阵的行代表了某个采样点的 $(3+4+5+4+3) \times 6$ 个采样视角,6指在朝向上的6个均匀采样的方向;括号中数值代表不同高度在俯仰角上的采样数量,见图9;而列代表了各个采样点。

首先需要保证整个轨迹在每个采样点上都尽

量经过重要性值最高的视角, 其次要保证整个相机的视角变化是连续的。要保证视角变化光滑亦即保证如果从第 i 个点的第 j 个视角跳到第 $i+1$ 个点的第 k 个视角, 如果 k 离 j 越远会得到越大的惩罚。

为保证相机从一个舒适的角度开始运动, 动态规划的起始状态设置为 10 m 高度朝正前方的姿态。也就是 $S_{1,1} = M_{1,1}$ 。

动态规划的状态转换方程为:

$$S_{i+1,j} = \max(S_{i,j} + M_{i+1,j}, p_1 \times S_{i,j-1} + M_{i+1,j}, p_1 \times S_{i,j+1} + M_{i+1,j}, p_2 \times S_{i,j-a} + M_{i+1,j}, p_2 \times S_{i,j+a} + M_{i+1,j}, p_3 \times S_{i,j-b} + M_{i+1,j}, p_3 \times S_{i,j+b} + M_{i+1,j}) \quad (10)$$

其中, a 是指高度和朝向不变, 俯仰角变化时采用的间隔数; b 是指角度不变, 高度变化时的采样间隔数; p_1 和 p_2 是对朝向和俯仰角变化的惩罚因子; p_3 是对高度变化的惩罚因子。结果分析发现, 惩罚因子的变化会极大地影响整个相机的轨迹自由度, 见图 11。利用上式可以求解出一系列对应采样点的相机参数以及对应参数下的重要性值。为了保证较好的漫游体验, 对这些参数再次做拉普拉斯光滑, 进一步保证了相机轨迹的光滑性。

4.2 动态速度的线性求解

当求出最优化的轨迹之后, 也即可以得到每个采样点上所取的视角兴趣值。高效导航的目标是保证相机能够在兴趣值高的地方运动更长时间, 在兴趣值低的地方运行很少的时间。

设定需要求解的 t_i 即相机从采样点 p_i 运行到采样点 p_{i+1} 的时间。求解如下的优化方程:

$$\arg \min_{t_i} \sum_{i=0}^{n-1} f(t_i) V_i S_i, \quad s.t. \sum t_i = T \quad (11)$$

其中, T 是用户指定的漫游导航的总时长; V_i 指对应采样点的重要性值; S_i 是指 p_i 到 p_{i+1} 的路线长度。



图 11 惩罚因子的变化导致的完全不同的相机轨迹

Fig. 11 Different camera trajectory caused by the change of penalty factors

5 结 果

5.1 分析和对比

为了评估本文所提工作的有效性, 做了三组对比, 分别比较了几何得分、视角效果、运行时间在虚拟导航中最影响用户体验的方面。

图 12 比较了本文的几何分析得分结果和基于信息熵^[7]的重要性得分结果。其中, 左右两张图为以两个相似的视角截取的导航场景: 右图的建筑物地标性更强; 左图只是三维场景中一个小城镇的普通建筑群, 且这些建筑没有任何区分性。因而本方案在左图的得分明显低于右图。

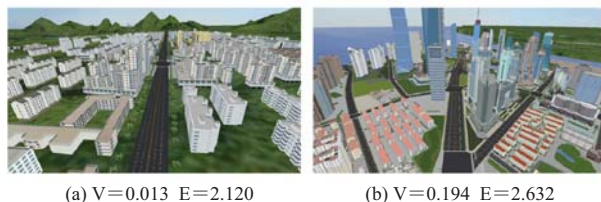


图 12 视角重要性值打分对比 (V 是本文重要性值得分, E 是信息熵得分)

Fig. 12 Importance value comparison (V is our score, E is entropy based score)

但是信息熵的方法给了两个视角基本差不多的得分。主要是因为信息熵的方法只考虑了相对投影面积，而本文的方法还考虑了一些额外的属性，比如独特性。

在视角效果这一块，主要是和谷歌地球^[20]内嵌的导航系统进行了对比。很容易发现由于谷歌的内嵌导航需要用户提前设定好相机自由度，然后整个过程都固定不变，这就使得不太可能设定出一个对全程都很好的视角，见图 13。下图右侧是本文的结果，首先在高速路部分本方案有一个很好的鸟瞰视角，帮助用户把握整体信息；同时在城区，能够以地面行车的模式运行，用户可以看到更多细节。而谷歌由于是固定视角，两者都不能很好地兼顾。

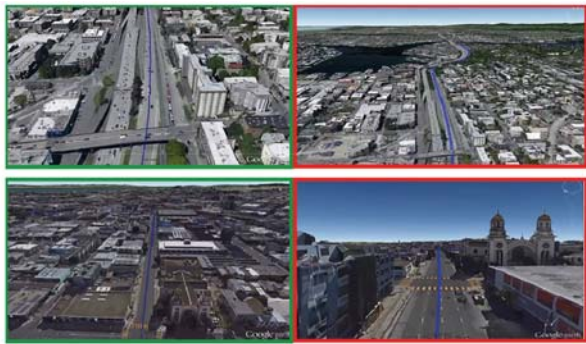


图 13 本文系统和谷歌导航的对比

Fig. 13 Comparison with google default navigation

最后和匀速导航对比了在时间分布上的差异。现在大多数导航平台都是基于匀速的简单导航方式，如谷歌地球^[20]、smart3D 的浏览平台^[21]、上海龙引擎公司的浏览平台^[22]，而本文的目标之一是希望在重要位置导航时间更长，不重要位置时间短一些。这样保证了导航了高效性，用户不至于失去兴趣，又能获得关键信息，见图 14。

本文在大量的三维虚拟场景和谷歌等真实扫描数据平台上执行了如前所述的算法。在试验中发现离线最耗时的部分主要是视点采样，即渲染某个采样视角下的图片最耗时。

图 15 展示了导航中一些代表性的截图。包含

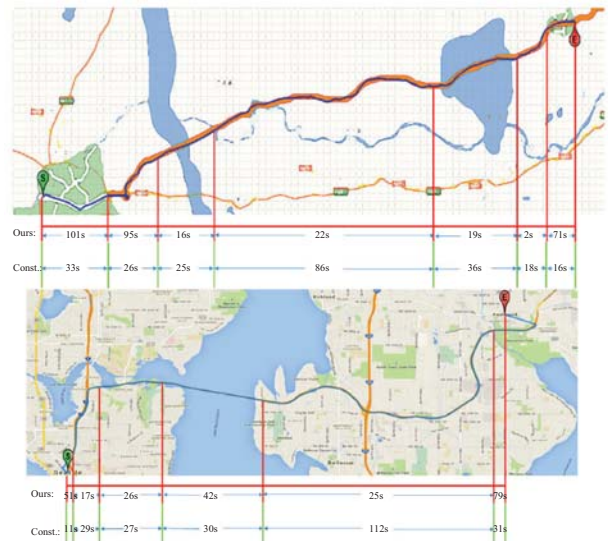


图 14 时间分布

Fig. 14 Time variation



图 15 导航效果截图

Fig. 15 Navigation effect screenshots

了虚拟平台和谷歌地球的平台，截取了其高速路段和城区部分的一些导航图。对应截图四个场景的数据见表 1。由表中数据可以发现，大量的时间

消耗都是发生在采集视角图片的过程。如果完全高密度均匀采样, 时间代价还会提高一个级数。

因而本文工作的主要贡献点: (1) 提出了一套合理完善的三维场景的模型分析方案, 兼顾考虑了包括建筑物、非建筑物以及道路模型; (2) 使用了自适应的获取采样点方法, 同时提出了使用全景图存储渲染结果的方案; (3) 解决了如何基于分析结果, 实现最佳视点下光滑连续导航的动态规划求解方案。

5.2 用户调查

为了评估算法的有效性, 我们做了用户调研, 并和一些当前可选方案做了对比, 分别是:

(1) 匀速导航^[22]; (2) 非连续的重点位置导航: 在关键位置匀速, 其他部分直接跳过; (3) Chen 等^[23]提出的导航方式; (4) 谷歌内置导航方式^[20]。

为了生成非连续的重点位置导航, 手动选择了沿路最关键的路段。在选出来的每一段, 以匀速、固定俯仰角运行, 跳过没有选出来的不重要路段, 保证有足够时间运行在关键部分。

有 58 位参与者参加了本次调研。这些参与

者事先并不清楚本次调查实验的目的, 也不知道我们的工作; 所有参与者视力正常或者纠正后是正常的; 用户事先都同意本次调查且本次调查遵循赫尔辛基声明。在调查过程中, 每一位参与者被展示了 8 对视频; 每一对包含了本工作的结果和上述的一种方案的结果; 8 对的出现顺序以及左右位置都是随机的; 整个实验时长大概在 30 分钟, 参与者可以重复观看视频, 然后需要回答如下问题:

问 1: 哪一个视频更好地概括了整个线路?

问 2: 哪一个视频看起来更舒服?

问 3: 哪一个视频提供了最多的信息?

对每一种情况, 样本数量是 116 (58 位参与者, 2 次重复), 调查结果见图 16。结果显示, 本文设计系统具有明显的优势。

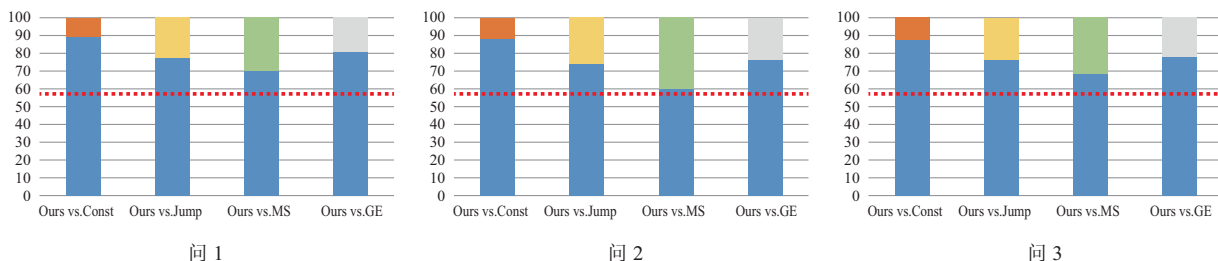
6 总结与展望

在本工作中我们解决了关于交互式高效导航的问题, 保证了轨迹既能兼顾整个场景的重要性

表 1 测试场景数据分析

Table 1 Statistics for test scene

名称	路长(km)	导航时长(s)	第一次采样时长(s)	自适应采样时长(s)	动归求解时长(s)
谷歌	21.06	240	351.0	7808.5	256.2
苏州场景 1	34.55	240	572.4	13125.6	414.4
上海场景	35.27	240	599.4	12385.7	458.2
苏州场景 2	21.64	240	360.2	9005.1	277.1



(Const、Jump、MS 和 GE 分别对应匀速、不连续、Chen 的方法、谷歌内置方法; 虚线代表用户的结果可能是由于随机选择的阈值)

图 16 对应三个问题的用户调研结果

Fig. 16 User study for three questions

值, 又能光滑连续。分别分析了场景的建筑、路型和一些关键的非建筑物的模型, 提出一套完整的分析方案。对于场景的路径采样, 本文提出了智能化自适应的采样方案, 包括采样相机的坐标和朝向共 6 个自由度。最后使用了动态规划的方案求解出了光滑连续的相机路径。

当前的解决方案也有一些限制, 首先是在高度变化上不能太大, 因为该方案不允许高密度、无限制的在海拔这个维度采样。一个可能的解决方法是由用户指定高度需求, 然后在该范围内高密度采样。其次是该方法主要是针对距离不是很长的导航, 因为要离线采样存储, 这就使得不太容易处理太长的路径, 否则会需要存储非常大量的数据。而在一个更长的路线上导航会是一个很有趣的问题。

参考文献

- [1] Mühler K, Neugebauer M, Tietjen C, et al. Viewpoint selection for intervention planning [C] // Proceedings of IEEE/ Eurographics Symposium on Visualization, 2007: 267-274.
- [2] Ji GF, Shen HW. Dynamic view selection for time-varying volumes [J]. IEEE Transactions on Visualization and Computer Graphics, 2006, 12(5): 1109-1116.
- [3] Klomann M, Milde JT. Semi autonomous camera control in dynamic virtual environments [M] // Virtual and Mixed Reality-Systems and Applications. Springer Berlin Heidelberg, 2011: 362-369.
- [4] Serin E, Adali S, Balcisoy S. Entropy assisted automated terrain navigation using traveling salesman problem [C] // Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry, 2011: 41-48.
- [5] Grabler F, Agrawala M, Sumner RW, et al. Automatic generation of tourist maps [J]. ACM Transactions on Graphics, 2008, 27(3): 100.
- [6] Polonsky O, Patané G, Biasotti S, et al. What's in an image? [J]. The Visual Computer, 2005, 21(8-10): 840-847.
- [7] Vázquez P, Feixas M, Sbert M, et al. Viewpoint selection using viewpoint entropy [C] // Proceedings of Conference on Vision Modeling and Visualization, 2001: 273-280.
- [8] Sokolov D, Plemenos D. Virtual world explorations by using topological and semantic knowledge [J]. The Visual Computer, 2008, 24(3): 173-185.
- [9] Yiakoumettis C, Doulamis N, Miaoulis G, et al. Active learning of user's preferences estimation towards a personalized 3D navigation of georeferenced scenes [J]. GeoInformatica, 2014, 18(1): 27-62.
- [10] Christie M, Olivier P, Normand JM. Camera control in computer graphics [J]. Computer Graphics Forum, 2008, 27(8): 2197-2218.
- [11] Drucker SM, Zeltzer D. Intelligent camera control in a virtual environment [C] // Graphics Interface. Canadian Information Processing Society, 1994: 190-190.
- [12] Nieuwenhuisen D, Overmars MH. Motion planning for camera movements [C] // Proceedings of IEEE International Conference on Robotics and Automation, 2004: 3870-3876.
- [13] Li TY, Lien JM, Chiu SY, et al. Automatically generating virtual guided tours [C] // Proceedings of Computer Animation, 1999: 99-106.
- [14] Salomon B, Graber M, Lin MC, et al. Interactive navigation in complex environment using path planning [C] // Proceedings of the 2003 Symposium on Interactive 3D Graphics, 2003: 41-50.
- [15] Oskam T, Sumner RW, Thuerey N, et al. Visibility transition planning for dynamic camera control [C] // Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2009: 55-65.
- [16] Hsu WH, Zhang YB, Ma KL. A multi-criteria approach to camera motion design for volume data animation [J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(12): 2792-2801.
- [17] Appleyard D. Why buildings are known: a predictive tool for architects and planners [J]. Environment and Behaviour, 1969, 1(2): 131-156.
- [18] Mirtich B. Fast and accurate computation of polyhedral mass properties [J]. Journal of Graphics Tools, 1996, 1(2): 31-50.
- [19] Michon P, Denis M. When and why are visual landmarks used in giving directions? [M] // Spatial information theory. Springer Berlin Heidelberg, 2001: 292-305.
- [20] Google. Google 地球. [2015-06-17]. <http://www.google.cn/intl/zh-CN/earth/>.
- [21] Acute3D. Smart3D 浏览平台. [2015-06-17]. <http://www.acute3d.com/>.
- [22] 上海龙引擎网络有限公司. 龙引擎三维虚拟现实平台. [2015-06-17]. <http://www.dragonengine.cn/>.
- [23] Chen B, Neubert B, Ofek E, et al. Integrated videos and maps for driving directions [C] // Proceedings of the 22nd annual ACM Symposium on User Interface Science and Technology, 2009: 223-232.