

硕 士 研 究 生 读 书 报 告



题目 城市3D建模系统

作者姓名 丁丹翔

作者学号 21960538

指导教师 李启雷

学科专业 计算机技术

所在学院 工程师学院

提交日期 二○一九 年 十二 月

PhotoSketch: a photocentric urban 3D modeling system

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Computer Technology

Advisor: Qilei Li

By

Danxiang Ding

Zhejiang University, P.R. China

2019

**摘要**

Google，Apple 和 Microsoft 的在线地图服务是探索 3D 城市的极受欢迎的应用程序。 它们的爆炸性增长为城市场景的逼真的 3D 建模提供了动力。 尽管诸如多视图立体和激光测距仪之类的经典算法是现有结构的详细 3D 模型的传统来源，但它们会生成不适用于这些导航应用程序利用的流数据的重量级模型。相反，由基于交互式图像的工具生成的轻量级模型更适合此领域。这项工作的贡献在于，它融合了多视图几何图形，直观的草图绘制界面和动态纹理贴图的优点，以生成轻巧的逼真的 3D 模型建筑物。 我们使用 PhotoSketch 系统呈现城市场景的实验结果。

**关键词**：基于图像的建模， 3D 纹理模型，结构和运动，多视图几何，3D 摄影，相机校准

**Abstract**

Online mapping services from Google, Apple, and Microsoft are exceedingly popular applications for exploring 3D urban cities. Their explosive growth provides impetus for photorealistic 3D modeling of urban scenes. Although classical algorithms such as multiview stereo and laser range scanners are traditional sources for detailed 3D models of existing structures, they generate heavyweight models that are not appropriate for the streaming data that these navigation applications leverage. Instead, lightweight models as produced by interactive image-based tools are better suited for this domain. The contribution of this work is that it merges the benefits of multiview geometry, an intuitive sketching interface, and dynamic texture mapping to produce lightweight photorealistic 3D models of buildings. We present experimental results from urban scenes using our PhotoSketch system.

**Keywords：**Image-based modeling · Phototextured 3D models · Structure and motion · Multiview geometry · 3D photography · Camera calibration

**1研究背景**

3D 纹理建筑模型的生产已经应用于导航，地图绘制，娱乐，虚拟旅游，城市规划和应急管理等多个领域。 Google，Apple 和 Microsoft 流行的导航和地图绘制工具已将城市重建的好处广泛传播给了公众。

现有城市结构的 3D 纹理建模问题催生了许多交互式技术和自动建模方法。交互式建模过程繁琐且耗时，而自动建模方法容易出错，且经常回出现产生嘈杂或不完整的结果。例如多视图立体声之类的自动建模方法通常因需要进行繁杂的编辑操作而受阻，而这些编辑操作对于修复生成的密集 3D 模型是必不可少的，这就一定程度上损害了自动建模的优势。已知自动建模方法会省略用户交互，在这个前提下，数据错误或者部分数据缺失通常不会产生令人满意的结果。

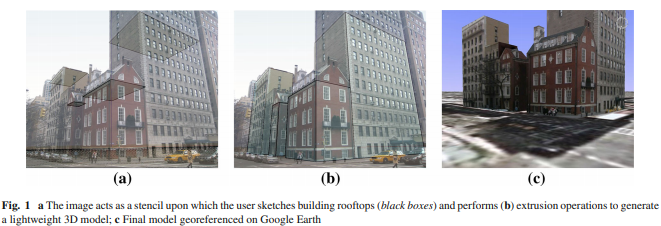
在本文中，作者设计出一个出色的交互式系统，可受益于相机姿势自动恢复和稀疏点云的生成，而在建模过程中保留了人员，以指导几何图形的完成。

**2 研究内容**

本文的主题涉及从照片生成轻量级模型，从而使这种方法能够吸引广泛的用户。作者提出了一个名为 PhotoSketch 的新系统，它是一个以光为中心的城市 3D 建模工具，允许用户直接在现有建筑物的照片上绘制草图。草图概述了可以通过一系列推拉挤压和锥度操作将其提升为 3D 形状的轮廓线。与其将照片视为模型生成后应用的后处理，不如将照片作为模型生成之前的起点。工作流程将照片视为描图纸，然后将其定义为 2D 形状，然后再将其压缩为 3D 模型。用作模型基础的照片也自动用作它们的纹理元素，从而促进了逼真的可视化。这种方法的特点是用户，对于交互式可视化来说，生成近似轻量级纹理模型至关重要。

PhotoSketch 系统面向主流用户，他们将轻松利用照片来创建 3D 模型。当前的技术水平不足以直接从照片容易地产生轻质的照片纹理模型。文中结合了运动结构（SfM），以从场景的一组重叠照片中自动恢复稀疏点云和一组相机姿势。这对于简化直观的用户界面以基于直接在照片上绘制的草图上的拉伸操作构建 3D 模型至关重要。 虽然用户有传统上在 3D 环境中使用拉伸和推拉工具，文中应用程序通过在 2D 图像空间中应用这些工具来寻求更加直观。

将 3D 图元定向为 3D 场景的 2D 照片这一艰巨的过程不再是笨拙的，而是可以沿着恢复的地面直接在图像上绘制图像。以这种方式，可以将绘图限制在场景的墙壁和地板上，以产生足迹，然后可以将其挤出以形成体积。 通过在照片上草绘 2D 足迹并将其捕捉到通过 SfM 恢复的 3D 点，将其拉伸到对象的适当高度来构建模型。 为此使用推挽图形界面。 图 1 给出了一个例子。



3 PhotoSketch工作流程

文章中介绍 PhotoSketch 建模工作流程，并演示其设计如何简化用于建模城市区域的用户体验。系统的输入是无序重叠图像的集合。然后使用运动结构（SfM）跟踪照片中的特征，以确定相机的姿势参数。这使我们能够将所有照片放入一个参考框架中，在该参考框架中构建 3D 模型。

一旦相机姿势恢复完成，在输入图像其中之一上绘制的任何用户绘图都将在其余图像中正确对齐。具有多个重叠图像的存在使存在遮挡的情况下促进整个场景的覆盖。由于可以将每个图像投影回场景中，所有 3D 表面的纹理都将来自无遮挡视图。

PhotoSketch 的基本前提是场景图像足以引导用户进行一系列草图绘制操作，并充当用户在其上跟踪建筑物占地面积的模具。 PhotoSketch 以简化城市模型建模的用户体验的方式设计系统。

这是通过提供一组 2D 草图绘制工具来实现的，这些工具只能在地平面和多边形面上运行。 工具由矩形，圆形 / 椭圆形，弧形，折线和样条线组成。 地平面作为绘制 2D 立面轮廓的草图板。由于可见性问题，有时需要在与地面不一致的平面上绘制足迹， 因此，允许用户以零偏移量为参考地面来更改草图的偏移量或高度。

PhotoSketch 工作流程包括以下步骤：

（1）通过多视图几何自动恢复稀疏的 3D 点云和相机姿态信息（第 3.1 节）；

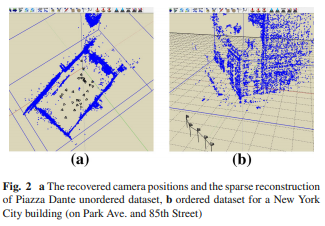
（2）摄像机相对于接地平面的对准（第 3.2 节）；

（3）基于素描 2D 足迹并应用一组由照片引导的拉伸和锥度操作的交互式建模（第 3.3、3.4 节）。

**3.1 Structure from motion (SfM)**

SfM 从一组重叠的场景图像中使用自动特征提取和跟踪来查找相机姿势并重建稀疏的 3D 点云。 需要照相机姿势的自动恢复才能将纹理准确地投影到模型上，并且稀疏的 3D 点云有助于帮助用户将拉伸或锥度操作捕捉到所需的高度。 重要的是要注意，恢复的结构稀疏且不完整。 尽管不足以对对象进行完全建模，但有助于用户构建模型还是很有用的。

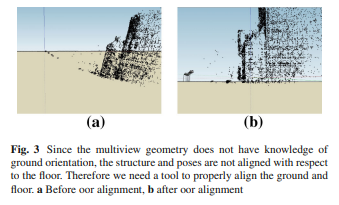
图 2 描绘了但丁广场和公园大道 / 第 85 街（NYC）数据集的摄像机姿势和稀疏重建。 图 2 中的视锥表示恢复的相机姿态。 这些结果来自我们自己的 SfM 实施。 可以应用其他开源解决方案，例如 Visual SfM [41]，Bundler [32] 或 OpenMVG [23]。 用户可以将他们的照片输入这些系统，然后我们可以导入和解析这些系统的输出，以获取相机的姿势和稀疏点云。



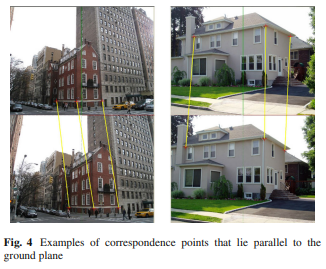
**3.2摄像机相对于地平面对齐**

由于初始相机的绝对位置和方向未知，因此我们将第一相机放置在世界坐标系的原点，即 K [I | 0]。 除非有 GPS 和 / 或 IMU 数据中可用的其他信息，否则大多数 SfM 系统都是以这种假设开始的，以设置其帧坐标系。关于此相机的坐标系，模型的稀疏 3D 点云的地板现在看起来是倾斜的，如图 3a 所示。如图 3b 所示，必须有一个地平面对准台才能正确旋转摄像机和稀疏点云。 这使地板与地面平行。

这种对齐是在所有视图中匹配建筑物轮廓或屋顶 2D 草图的关键步骤。除上述问题外，我们还假定挤压操作垂直于地板，与大多数建筑物的外墙一致。我们已经开发出一种自动方法来恢复第一台相机的未知旋转 Rg。这是通过让用户调用 3D 点选择系统中的套索工具来在地面上的平面上收集一组 3D 点来实现的。 我们使用 RANSAC 方法拟合通过这些点的平面，这对于异常值具有鲁棒性。恢复平面的法线 n 是 SfM 坐标系中的重力方向。我们求解旋转法线 n 的旋转 Rg，使其与我们的世界坐标系向上方向（0，0，1）对齐。



在实践中，我们观察到，在许多情况下，地面上的点被汽车，树木，行人遮挡，并且没有足够的平面 3D 点来推断一个平面。 另外，在实际情况下，嘈杂的选定 3D 点对于接地平面检测也不可靠，因为一些误差会严重破坏模型。 当将一面向上拉动并且绘制的体积的边缘在视觉上未看起来与图像对齐时，用户可以轻松观察到此错误。 在这种情况下，用户可以通过在两个视图中选择至少三个对应的图像点来激活我们的手动地平面检测工具，这些视图对应于图像中的地面或屋顶线（图 4）。



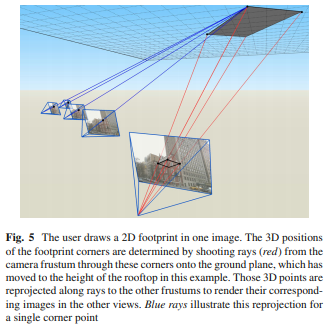
**3.3绘制 2D 轮廓**

恢复地板方向后，用户可以将该平面的高度捕捉到稀疏点云中的任何 3D 点。 如果可以看到地面上的一点，则最好将其捕捉到，以便可以从地面开始进行建模。 但是，通常情况下，地面点被遮挡，并且更容易捕捉到屋顶线上的可见点以建立覆盖区。 然后可以将该足迹向下压向地面。 本文的示例中使用了这种方法。

请注意，如果没有 3D 点可以将地平面捕捉到屋顶线或地板上，则用户可以调用系统的手动特征跟踪来建立场景图像中可见角的对应关系。由于已知相机的姿势，可以通过三角测量找到被跟踪要素的 3D 位置。

在摄像机和地板与地面对齐后，用户可以从输入集中选择图像，并通过各自的摄像机视锥台查看 3D 场景。然后，用户在地平面上绘制草图。 用户可以选择 2D 绘图工具（例如矩形，折线，圆形 / 椭圆形或样条线）并绘制建筑物的可见轮廓。此过程仅需要用户单击建筑物立面的角。为了帮助用户完成此过程，系统提供了一个虚拟放大镜，以帮助用户准确地确定顶点。

图 5 显示了该过程的实际操作。 用户在图中最右边的图像上单击屋顶的三个角，以获取沿图像中的屋顶放置的平行四边形。 为了确定这些 2D 角的 3D 点，我们从相机视锥的投影中心穿过每个 2D 点发射光线，并计算与 “地面” 平面的交点。 这些在图 5 中显示为红线。可以将生成的 3D 点重新投影到其他相机视锥上，并使生成的蓝光穿过其他图像视图中的相应角。 通过使用 SfM 计算出的相机姿态恢复，一切都可以实现。 结果，将一张图像中的任何草图正确投影到所有其余视图上。



我们的系统允许用户在草绘过程中从一个视点切换到另一个视点，以添加当前视图中被遮挡的角点 图 6 显示了用黑色绘制的脚印。作为 SfM 的结果，相机的位置和方向是已知的。因此，在一个视点中绘制的足迹将显示为在其他视点中对齐。由于通过捕捉到屋顶边缘上的 3D 点来选择绘图平面的高度，因此绘制足迹角的每个 3D 位置 Mi 是已知的。我们可以基于从 SfM 派生的已知外部和固有参数将 Mi 投影到视图 j 中，并使用 Eq （1）在视图 j 上获得其二维投影 mi。

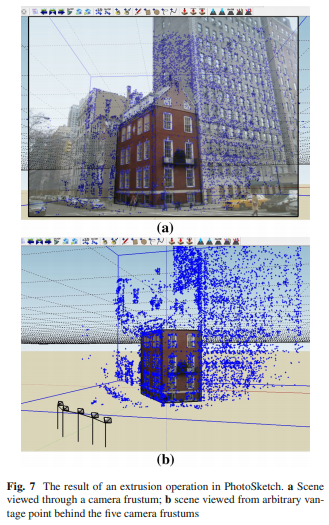
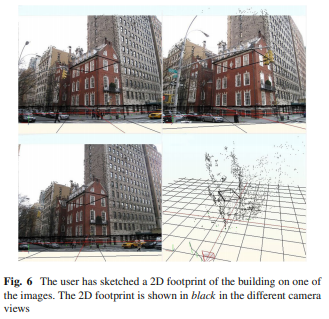
请注意，K j 是用于视图 j 的摄像机的本征 3×3 矩阵，外部参数是旋转 Rj 和平移 Tj，它们定义了用于视图 j 的摄像机姿态。



**3.4挤压，推拉和锥度操作**

文章的工作基础是假设一组简单的挤压和锥度操作足以模拟一组丰富的城市结构。这与过程建模中的相关工作是一致的。一套简单的规则足以生成整个虚拟城市。但是，过程建模的重点是根据语法创建模型。尽管此方法可以自动创建一般城市区域，但不适用于重建现有建筑物。逆过程建模的最新工作是找到现有对象或场景的过程表示，但是这种方法需要有关输入图像或建筑结构的先验知识。

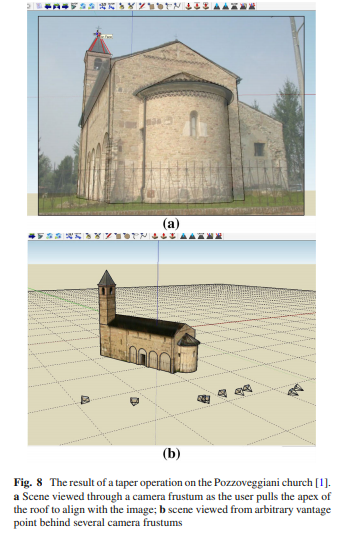
PhotoSketch 试图通过将人员置于循环中并建立一套简单的规则来超越这些限制，用户可以在其中根据现有照片快速高效地对建筑物建模。 在本文系统的工具集中，最简单的操作就是从脚印挤出。 用户只需要将足迹拖到所需的高度即可。 这可以通过捕捉到从 SfM 恢复的 3D 点的高度或基于动态纹理的图像上的视觉提示来完成。 在此需要强调的是，动态纹理化是系统的关键优势，它可以帮助用户基于实时纹理投影进行建模。通过将照片重新投影到模型上，任何建模错误都会以未对准的纹理和几何形状的形式迅速显现出来。实时动态纹理是使用 GPU 实现的。 图 7 显示了在图 6 的占位面积上进行挤压操作的结果。



用户可以使用推拉式界面执行拉伸。 通过将面对齐到代表平面的稀疏 3D 点，可以进一步完善。 草绘不仅限于在地平面上绘制足迹。 用户还可以在拉伸的面上进行绘制，并使用推拉界面完善模型。

用户可以通过渐缩到点，线或偏移来进一步编辑模型。 通常用于建模屋顶。 在这些情况下，用户可以捕捉到一个具有特色的 3D 点，该点代表

逐渐减小高度或动态调整高度，以在可见的表面上获得适当的纹理。 图 8 示出了挤压操作之后的锥度操作的结果。



**4.论文的创新点及成果可以应用的领域**

本文介绍的建模方法的贡献在于，它融合了多视图几何，直观的草图绘制界面以及动态纹理贴图的优点，以生成轻量级的逼真的建筑物 3D 建模。

这种方式也有一定的缺陷，不能处理不遵循传统的自上而下设计的当代建筑，其中包括悉尼歌剧院，或者火山及的沃尔特迪斯尼音乐厅等弗兰克·盖里建筑。但是，这些类型的建筑只占城市景观的一笑部分，本文的建模方法专注于绝大多数建筑物，包括程序建模技术针对的建筑物。参考文献

[1]Wolberg, G., & Zokai, S. (2018). PhotoSketch: a photocentric urban 3D modeling system. *The Visual Computer*, *34*(5), 605-616.