

# 一种基于单幅图像的3D场景建模方法\*

胡春雨<sup>1,2</sup>, 刘弘<sup>1,2</sup>, 陆佃杰<sup>1,2</sup>

(1. 山东师范大学信息科学与工程学院, 济南 250358; 2. 山东师范大学山东省分布式计算机软件新技术重点实验室, 济南 250014)

**摘要:** 针对计算机仿真平台上的三维场景建模问题, 提出了一种基于单幅图像的交互式三维场景建模方法。该方法考虑到场景建模的特性, 以三种基元为基础, 通过人机交互、基元构建以及基于图像的真实感处理, 快速得到逼真的三维场景。实验表明, 所提出的三维场景建模方法具有成本低、鲁棒性好的特点, 能够快速生成用户期望的三维场景重建效果。

**关键词:** 三维场景建模; 基元; 计算机仿真; 人机交互

**中图分类号:** TP391.9

**文献标志码:** A

**文章编号:** 1001-3695(2014)06-1909-05

doi:10.3969/j.issn.1001-3695.2014.06.073

## Monocular image based 3D-scene reconstruction method

HU Chun-yu<sup>1,2</sup>, LIU Hong<sup>1,2</sup>, LU Dian-jie<sup>1,2</sup>

(1. School of Information Science & Engineering, Shandong Normal University, Jinan 250358, China; 2. Shandong Provincial Key Laboratory of Distributed Computer Software Novel Technology, Shandong Normal University, Jinan 250014, China)

**Abstract:** In order to build 3D scene models on the platform of computer simulation, this paper put forward a monocular image based interactive 3D-scene reconstruction method. The algorithm took the characteristic of 3D scene modeling into account, and was based on three kinds of primitives. The implementation included human-computer interaction, primitive construction and realistic processing based on images. Finally, it got the realistic 3D scenes fast. Experimental results show that the algorithm is low cost and robust, and the user can get the desired 3D-scene quickly.

**Key words:** 3D scene modeling; primitives; computer simulation; human-computer interaction

三维场景建模是虚拟现实及计算机动画的基础, 随着计算机图形学以及仿真技术的不断发展, 人们对三维场景建模的真实感要求与日俱增<sup>[1]</sup>。传统的三维场景建模技术主要包括基于几何造型的建模技术和采用高精度设备获取场景几何信息的建模技术。前者虽然成本较低, 但构建的三维场景真实感差, 而后者虽然能够得到高精度的三维模型, 但成本很高。本文以单幅图像为基础进行三维场景建模, 在控制成本的同时实现了高逼真度的要求。图像是非常容易获得的灵感来源, 获取单幅图像的简单性也使得其在数量和丰富程度上远远优于其他的创作来源; 除此之外, 所拍摄到的场景以其最真实也是最自然的样貌呈现给使用者。因此以图像为基础所创建的三维场景模型能够更好地模拟现实场景, 给人以身临其境的感受。然而, 在仅有单幅图像的情况下, 构建一个复杂的三维场景是很困难的。文献[2]提出了一种长方体代理的思想, 并将其应用到了图像真实感处理中, 本文借鉴了这一思想, 并将其应用于三维场景建模中。针对三维场景建模的特点, 本文提出了长方体、圆柱、圆锥三种基元, 并利用这三种基元实现了对单幅图像中的场景实现三维恢复, 建模效果如图1所示。

## 1 相关研究

真实场景建模是虚拟现实以及计算机图形学领域内的一

个重要研究内容, 国内外许多学者都在这一领域取得了突出成就。

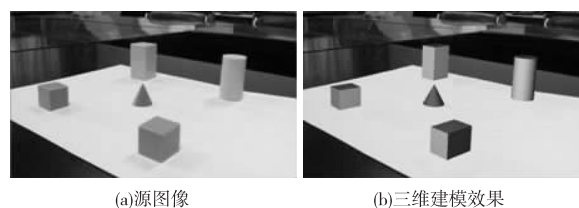


图1 三维几何建模效果

传统的三维建模方法通常需要专门的设备对实际场景进行数据采集, 如使用三维激光扫描仪获取场景的几何信息<sup>[3]</sup>, 这种建模方式虽然满足较高的精度要求, 但数据采集过程复杂、成本较高, 并且该方法的使用受到使用环境以及被测物体的限制, 影响了它的应用。另一种传统的建模方式是起源于20世纪70年代的基于几何造型的建模方式, 这种建模方式是利用点、线、面、体等几何元素, 对它们进行组合、相互作用, 最终产生用户所需要的三维造型。文献[4]提出了一种基于CAD电子图纸的复杂场景快速建模技术, 该方法具有快速、精确的特点, 但需要电子图纸作为输入, 不适合对缺少电子设计方案的场景进行场景重建。文献[5]利用已有CAD数据建立三维模型, 这种建模方法快速准确, 但是需要已有的CAD数据

收稿日期: 2013-07-18; 修回日期: 2013-08-24 基金项目: 国家自然科学基金资助项目(61202225, 61272094); 山东省高等学校科技计划资助项目(J13LN13)

作者简介: 胡春雨(1990-), 女, 山东人, 硕士研究生, 主要研究方向为3D虚拟建模(1032078953@qq.com); 刘弘(1955-), 女, 山东人, 教授, 博士, 主要研究方向为分布式人工智能、软件工程及计算机辅助设计; 陆佃杰(1981-), 男, 山东人, 讲师, 博士, 主要研究方向为可视化与三维建模。

作为支持,对建模的前期准备要求较高。除此之外,还有许多专门应用于各种紧急情形下进行场景仿真模拟的商业软件,如 FDS、Massive software、Exodus 等,但上述商务软件均没有开放技术方案,无法作为相关研究的基础。

随着数码相机、手机等电子产品的普及,图像的获取越来越容易,基于图像进行三维建模成为真实场景建模的一个新的研究方向。目前已存在很多使用图像驱动建模的方法,但多是以多幅图像为基础。文献[6]提出了一种基于图像序列的交互式三维几何建模方法,该方法能够通过对几幅图像的分析实现快速高精度的场景重建,但该方法对所使用的图像要求较高,需要输入的图像为连续的图像序列。文献[7]将二维图像与三维点云相融合,从而生成有层次感和质感的三维模型,但该方法需要高精度的激光雷达扫描仪作为实验的支撑。文献[8]使用基于单幅照片的三维几何建模方法,在该方法中,通过使用输入照片的剪影轮廓来引导现有模型进行保持结构的变形,使结果与照片中的物体相匹配,该方法的优点是能够实现复杂模型的构建,其不足之处在于需要使用现有模型作为基础,并且只能实现单个物体的建模。文献[9]提出了一种基于平行六面体中的正交性约束来进行相机校准以及三维建模的方法,该方法完全由计算机进行计算处理,而没有进行任何的人机交互,其建模时间较长。

本文针对场景建模的特性,提出了一种基于单幅图像的交互式 3D 场景建模方法,该方法使用单幅图像作为输入,降低了对输入的要求。在图像处理的过程中加入了用户交互,并利用交互信息指导二维图像与三维模型之间的对应关系计算,从而大大缩短了图像处理的时间。实验结果表明,该方法具有良好的鲁棒性。

## 2 基于单幅图像的 3D 场景建模

### 2.1 算法框架

本文采用单幅照片作为输入信息,输入的图片需满足以下几个条件:a)场景内的物体需要处于同一水平面上;b)场景内物体投影的出现是由单一点光源产生的;c)场景内的物体是不透明的。

本文的目标是实现虚拟场景的三维几何建模。具体地,能够以单幅图像为基础,通过使用基元构建出基本的三维物体或者建筑物的模型;然后对背景图片进行修复,对物体进行纹理贴图,最终实现基于二维图像的三维场景几何建模。除此之外,本文的方法还提供了对三维场景的编辑功能,能够使用户根据自身需要对场景进行编辑。具体流程如图 2 所示。

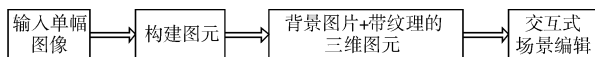


图 2 基于单幅图像的三维场景建模方法流程

### 2.2 虚拟场景中的三维几何基元

通过对现实中各种场景的观察,发现场景中的大部分对象都是由规则的图形构成的,如楼房多为长方体,花坛多为圆柱体等。为实现对虚拟场景的三维重建,本文提出三种几何基元:长方体基元、圆柱体基元、圆锥基元,并将这三种基元作为构建三维场景最基本的元素。

**定义 1** 几何基元是指构建和处理三维场景的基本单元,并且是场景中看得见的几何实体,具有自己的属性和方法。本文可以将基本三维几何物体的构建看成是几何基元的构建过程。

首先引入以下特殊符号: $\{R_1, R_2, \dots, R_n\}$  表示用户标注出的感兴趣的区域,每一个区域都与图像中的一个物体相对应; $H_i$  表示用户选取的六边形; $PH_{R_i}$  表示对  $R_i$  区域抽象出的几何基元。

### 2.3 基于单幅图像的几何基元构建

a) 图像标志。对于任意区域  $R_i$ , 都需要构建一个对应于该区域的几何基元  $PH_{R_i}$ 。本文采用人机交互的方式,对需要进行几何建模的对象进行六点标志,这种方式标志出对应于区域  $R_i$  的凸多边形,利用其拐点与三维几何体之间的对应关系来构建几何基元。拐点标注(参见电子版)及三维对应关系分别如图 3、4 所示。

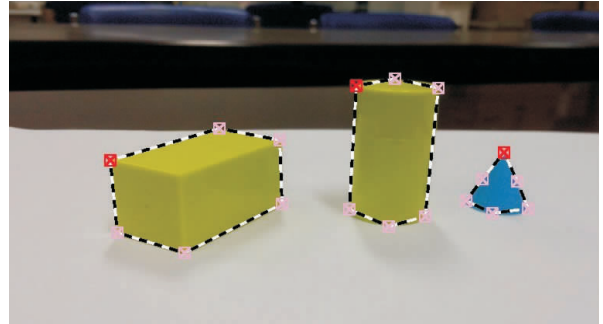


图 3 对各区  $R_i$  分别标注出六个标注点(图中的粉红色和红的十字)

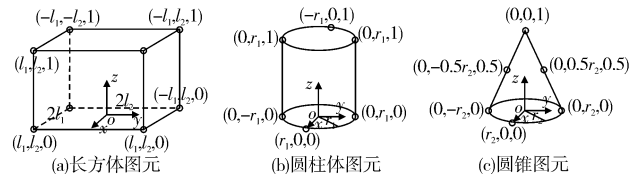


图 4 与图 3 中的各区相对应的几何基元的三维信息

b) 相机参数估计。本文采用文献[10]中介绍的标准相机模型,其被描述为  $P = K[R|t]$ ,  $K, R, t \in \mathbb{R}^3$ , 其中  $K$  为相机的标定矩阵,即相机内部矩阵,其参数为相机内部参数; $R$  和  $t$  分别为旋转矩阵和平移矩阵,它们共同组成了相机的外部矩阵,其参数决定了相机的方位和位置,并被称为相机外部参数。

$$K = \begin{bmatrix} a_x & s & x_0 \\ & a_y & y_0 \\ & & 1 \end{bmatrix} \quad (1)$$

在本文的方法中,假定扭曲参数  $s$  为 0,世界坐标系为均匀坐标系,即  $a_x = a_y$ 。在进行相机参数估计时,需要估计的参数主要包括相机内部参数(来自  $K$  的 3 个参数)、相机外部参数(来自  $R$  的 3 个参数和来自  $t$  的 3 个参数)以及  $l_1, l_2, r_1$  和  $r_2$ 。对于该均匀的世界坐标系,假定三维点  $X_i$  与二维点  $x_i$  为对应点,即  $X_i \leftrightarrow x_i$ ,那么对于投影矩阵  $P$ ,存在  $x_i = PX_i$ ,本文采用直接线性变换(DLT)算法估计相机参数以及  $l_1, l_2, r_1$  和  $r_2$ 。在长方体基元中需要估计的参数为 11 个,圆柱基元需要估计的参数为 10 个,圆锥基元需要估计的参数为 10 个。利用标注出的六对对应点以及线性规划,可以求解出上述未知参数。此时对每个基元都单独产生一个相机标定,后续通过联合优化将场景中的物体统一放在同一坐标系下进行重新标定。

得到相机的内部矩阵、外部矩阵以及基元参数之后,利用公式  $X_i = P^{-1}x_i$  计算出在世界坐标系下特征点的三维坐标,并利用三维点坐标构建三维模型。假设与  $R_i$  区域相对应的基元标记为  $T_j$ ,本文定义误差函数:

$$E = \eta \sum_j \left( \sum_i \|P(X_i) - x_i\|^2 \right) + \gamma S(T_1, T_2, \dots, T_n) \quad (2)$$

$$1 \leq i \leq 6, 1 \leq j \leq |T|$$

其中: $\eta$ 和 $\gamma$ 分别表示投影误差和语义约束的权重。语义约束主要包括对重复性和共面性的检测,重复性是指场景中具有相同几何信息的基元,共面性是指所构建的基元具有相同的高度。语义约束表示如下:

$$S(T_1, T_2, \dots, T_n) = \sum_{(m,n) \in T} R(T_m, T_n) + \sum_{(k,h) \in T} C(x_k, x_h) \quad (3)$$

关于语义约束的判断:(a)用户通过观察指定场景中存在的约束条件;(b)在计算过程中,计算出场景中用户指定的约束条件相符的越多, $S$ 的值就越小,相应的误差函数越小。

c)联合优化。联合优化是指对场景中的所有三维模型进行联合操作,最终得到最优的三维几何场景。

当完成相机的标定工作之后,对场景内的每个物体都单独生成一个世界坐标系。因此,要想在同一场景内显示标注出的所有物体,就需要对生成的所有基元进行联合操作,将其转换为同一坐标系下的物体。选定某一基元作为基准基元,假设将其标记为 $T_0$ ,此时为每一基元添加四个参数:基元高度 $h^i$ , $x$ 方向的平移参数 $t_x^i$ , $y$ 方向的平移参数 $t_y^i$ ,旋转角度 $\theta^i$ 。长方体基元具有六个参数 $\{t_x^i, t_y^i, h^i, t_x^i, t_y^i, \theta^i\}$ ,圆柱基元和圆锥基元都具有五个参数,分别为 $\{r_1^i, h^i, t_x^i, t_y^i, \theta^i\}$ 、 $\{r_2^i, h^i, t_x^i, t_y^i, \theta^i\}$ 。对于基准基元 $T_0$ , $h^0 = 1.0$ , $t_x^0 = t_y^0 = 0$ , $\theta^0 = 0$ ,其他基元 $T_i$ 的平移参数以及旋转角度均以基准基元 $T_0$ 为基准进行计算。在计算相机参数及基元参数时使用文献[11]中的非线性最优化算法 Levenberg-Marquardt 方法进行参数优化,求解出最优参数。分别以场景中的各基元作为基准基元计算出与其相对应的基元集合,并在结果中选出误差最小的一组基元集合作为最佳结果进行三维绘制显示。

上述联合优化的算法流程如下:

```

begin
    error ← max
    n ← 图元个数
    if n < 2
        print "图元个数小于2,不需要进行联合优化。"
    else
        for i ← 0 to n - 1
            以 primitives[i] 为基准图元计算出各图元的参数
            currerror ← 当前场景的误差值
            if currerror < error
                最佳场景 ← 当前场景
                error ← currerror
            i ++
        end
end

```

其中,error表示全局误差,max表示一个无穷大的值,primitive[i]表示第*i*个几何基元,currerror记录当前场景的误差值。三维建模效果如图5所示。

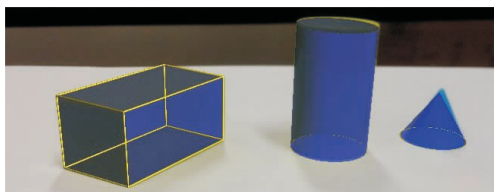


图5 三维建模效果

## 2.4 基于图像信息的虚拟场景真实感处理

为使所建场景更具真实感,需要对构建出的三维基元进行纹理贴图并将其放置在原图像背景之中。为实现该效果,在基于图像信息的虚拟场景真实感处理中需要利用图像信息来获取图像背景以及纹理信息。在本阶段,需要完成两个主要任

务:a)利用改进的GrabCut算法对图像进行前景与背景的分离,将用户感兴趣的区域作为前景分离出之后,对背景图像进行图像修补从而得到背景层;b)利用提取出的前景图像为构建出的三维基元贴上纹理信息,使其尽可能地逼真。在完成上述操作之后,会得到图像完整的背景层以及贴有纹理的三维基元。

### 2.4.1 前景背景分割

本文在文献[12]的基础上提出了一种基于GrabCut改进的图像分割算法,使其适用于对多个前景物体进行图像提取。GrabCut算法的本质是对graph cut算法的一个改进,其主要思想是利用图像中的纹理信息和边界信息来进行图像分割。本文在GrabCut的基础上,通过将图像分解为多个子图像,对多个子图像进行并行计算来提高其对具有多个前景对象的图像分割的速度。

GrabCut算法的基本思想为:首先,通过用户交互在图像中用矩形框标注出需要进行前景提取的区域,此时会获得一个硬分割;然后使用高斯混合模型(Gaussian mixture model, GMM)取代graph cut中的灰度直方图来初始化每个像素点;接下来使用迭代优化的方法最小化Gibbs能量函数:

$$E(\alpha, k, \theta, z) = U(\alpha, k, \theta, z) + V(\alpha, z) \quad (4)$$

其中: $\alpha$ 表示不透明度且 $\alpha \in [0, 1]$ ,0为背景,1为前景; $\theta$ 为图像的前景与背景的灰度直方图; $k$ 为每个像素独立的GMM参数。GrabCut算法最终通过最小化Gibbs能量,为图像中的每个像素点赋一个 $\alpha$ 值,从而实现图像分割的结果。图6给出了利用GrabCut算法对具有多个前景物体的图像进行图像分割的效果。

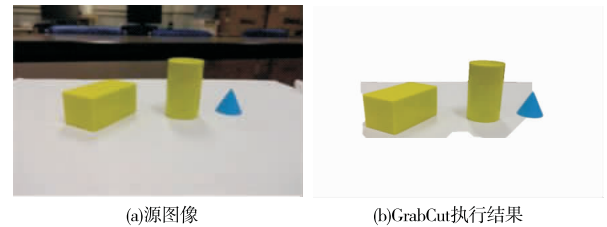


图6 GrabCut算法执行效果

从图6中可以观察到,GrabCut算法适用于提取单个前景对象,对于具有多个物体的场景进行前景提取时效果并不理想。虽然用户可以通过人机交互对图像进行前景背景标注,但要想获得理想的执行效果,需要复杂而精细的人机交互并且耗费很长时间。

针对以上问题,本文对GrabCut算法作了一定的改进,改进方法具体如下:

a)预分割。输入图像 $I$ ,根据 $I$ 中需要提取出的前景对象数目,由用户通过人机交互对 $I$ 进行分割,使得获得每个图像块 $M_i$ 只包含一个前景对象。

b)框选出前景对象。对于每个图像块 $M_i$ 使用矩形框框选出其对应的前景对象。

c)并行计算。对于所有的 $M_i$ 采用并行处理的方式进行图像分割。在并行处理的过程中,将一个图像块作为一幅单独的图像使用GrabCut对其进行图像分割。

$$E(\alpha, k, \theta, z) = \sum_{i=1}^n E_i(\alpha, k, \theta, z) \quad (5)$$

本文所提出的改进算法对每个图像块 $M_i$ 迭代优化能量函数 $E_i$ ,并对每个图像块的能量函数收敛到的最小值进行求和,将该值作为整幅图像能量函数的最优解。这种改进方式的优点在于将对场景图像的处理转换为对具有单个对象的图像



进行分割,并且采用并行计算的方式加快了图像处理的速度。图 7 给出了采用本文所提出的改进算法进行图像分割所得到的结果。

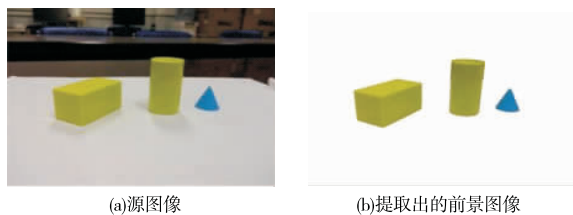


图 7 改进的 GrabCut 算法执行结果

#### 2.4.2 图像背景层的修补

本文采用 Telea 于 2004 年提出的基于快速行进法 (fast marching method, FMM) 的图像修补算法<sup>[13]</sup>对背景层进行修补。该算法采用了 FMM 方法中的水平集曲线进化算法来传递图像信息,具有实现简单、执行速度快、修补效果好等优点,这些特点完全满足三维几何场景重建的要求。

快速行进图像修补算法的基本思想是:对于标注出的待修复区域,先修复该区域边界上的像素点,然后层层向内推进,直到待修复区域内所有点都修复完成。图 8 给出了使用该算法进行图像修补的过程及结果。

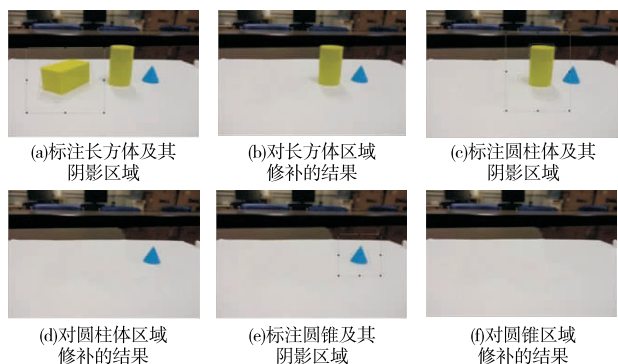


图 8 FMM 图像修补算法的执行过程及结果

#### 2.4.3 三维基元的纹理贴图

对提取出的前景图像进行相应处理,使其存储为长宽均为  $2^n$  的图像,利用该图像对构建出的三维基元进行纹理贴图。由于对三维基元的背面或者是遮挡区域的图像会产生缺失现象,本文采用来自同一基元的纹理对称修补的方法填补丢失的纹理,对比更复杂的纹理相似性检测,该方法具有更快的运行速度。由于三维重建时对纹理贴图的要求并不十分严格,同一基元纹理对称修补法完全满足本文的要求。基于图像信息的虚拟场景真实感处理结果如图 9 所示。

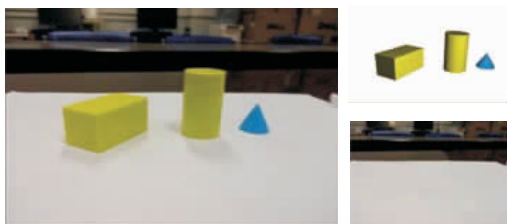


图 9 基于图像信息的虚拟场景真实感处理得到的结果

#### 2.5 虚拟场景的交互式编辑

将图像修补得到的背景图像设置纹理贴图结果的背景,就产生了一组基于单幅图像恢复得到的三维场景。对于得到的三维场景,通过对基元的每个面构建法向量从而实现对场景中单个基元的交互式操作,如平移、旋转、删除等,使其满足用户的需求。

### 3 三维几何建模实例

本文采用 VS2010、C# 进行程序设计,并使用 OpenCV 进行图像处理,使用 OpenGL 进行三维建模。下面以化工厂建模为例,介绍基于单幅图像进行三维场景建模的过程。

a) 将想要用于建模的图片导入到三维场景建模系统中,并对想要建模的物体进行标注,本文所提出的方法要求对各种基元统一标注出六个特征点。

b) 利用每一个基元的六个标注点,使用直接线性变换的方法,分别计算出各基元所对应的相机参数。接下来使用联合优化的方法计算出使整个场景误差最小的相机参数,并在该相机下得到一个相对于原图像误差最小的三维场景模型。

c) 对原图像进行图像处理:(a) 使用改进的 GrabCut 算法提取前景图像,用来对所建模型进行贴图;(b) 使用基于快速行进算法的图像修补算法进行图像修补,得到背景图像。

d) 使用步骤 c) 中得到的前景图像对所建模型进行纹理贴图,并将贴图后得到的三维模型置于背景图像中,从而得到与源图像相似度极高的三维场景。

图 10 给出了基于给定图片进行三维场景建模的过程及效果。表 1 给出了图 10 中场景的相机参数。

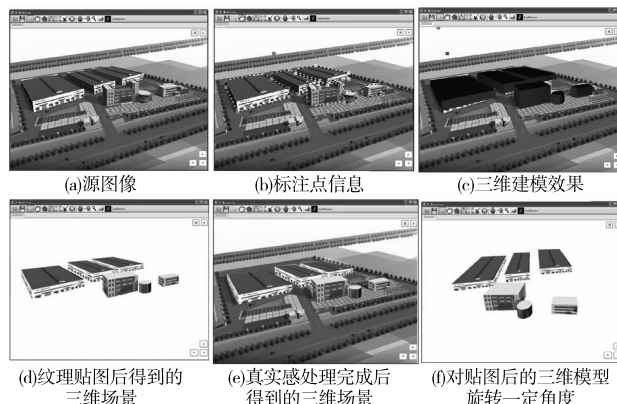


图 10 基于给定图片进行建模得到的化工厂效果

表 1 联合优化后得到的最佳相机参数

参数名称	参数值
相机内参矩阵	$K = \begin{bmatrix} 633.29351 & 0 & 780.02912 \\ 0 & 633.29351 & 168.99417 \\ 0 & 0 & 1 \end{bmatrix}$
相机旋转矩阵	$R = \begin{bmatrix} -0.54089 & 0.84049 & 0.03189 \\ 0.05240 & 0.07151 & -0.99606 \\ -0.83946 & -0.53709 & -0.08272 \end{bmatrix}$
相机平移矩阵	$t = [-1.74306835 \quad 4.43794 \quad 13.5724]$
相机位置	$W = [698.697 \quad 376.06981754 \quad 1]$

为验证本文方法的良好性能,以水立方照片作为原图像,分别采用本文所提出的方法、基于平行六面体的方法<sup>[9]</sup>、基于多幅图像的场景交互建模系统<sup>[1]</sup>以及基于几何形体的交互建模方法<sup>[14]</sup>进行三维重建,并分析这四种方法的性能。四种方法的建模效果如图 11 所示。

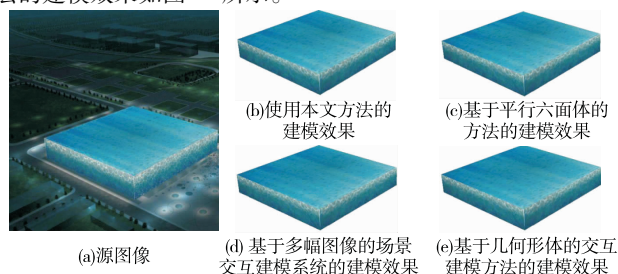


图 11 两种方法的建模效果

表2与图12中的方法1为本文所提出的方法,方法2为基于平行六面体的方法,方法3为基于多幅图像的场景交互建模方法,方法4为基于几何形体的交互建模方法。表2列出了四种建模方法的各项实验参数,图12对比了使用四种建模方式分别对具有不同对象个数的图像进行三维重建所使用的时间。通过四种建模方法的对比分析可以看出,通过本文所提出的方法进行三维建模得到的模型精度更高,与原图像的相似度更高。除此之外,在建模速度方面也有所提升。

表2 四种建模方法的各项参数对比

参数	原图像	方法1	方法2	方法3	方法4
长/像素	231	230	229	230	229
宽/像素	227	227	226	229	225
高/像素	53	53	51	54	52
误差 $E$		1.2511887	1.3251453	1.432566	1.3356632
建模时间/s		0.69	0.82	0.95	0.84
使用图像数目		1幅	1幅	多幅	多幅

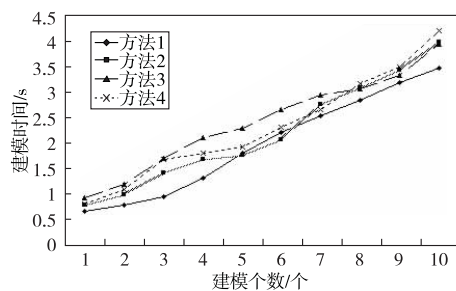


图12 四种建模方法的建模时间对比

## 4 结束语

基于图像进行三维几何建模目前已经成为计算机图形学领域内一个新的热点。本文提出了一种以单幅图像作为输入,对三维虚拟场景进行重建的算法。该算法以三种基元作为虚拟场景建模的基础,通过基元构建、基于图像信息的真实感处理以及交互式操作最终得到使用者所需要的三维场景。将该算法应用于化工厂的三维建模之中,得到了很好的3D建模效果;同时,实验验证该算法具有可行性高、执行速度快的特点。进一步的工作是对算法进行改进,使其适用于不规则对象以及复杂场景的三维模型重建。

## 参考文献:

- [1] 刘钢,彭群生,鲍虎军. 基于多幅图像的场景交互建模系统[J]. 计算机辅助设计与图形学学报, 2004, 16(10): 1419-1424,

1429.

- [2] ZHENG You-yi, CHEN Xiang, CHENG Ming-ming, *et al.* Interactive images: cuboid proxies for smart image manipulation [J]. *ACM Trans on Graphics*, 2012, 31(4): 99-109.
- [3] 沈方阳, 齐越. 基于互联网图像集的室外场景建模技术综述[J]. 计算机辅助设计与图形学学报, 2012, 24(1): 29-36.
- [4] LI Ting, SHU Bo, QIU Xian-jie, *et al.* Efficient reconstruction from architectural drawings[J]. *International Journal of Computer Applications in Technology*, 2010, 38(1): 177-184.
- [5] 陈谊, 孙悦红, 牛兵. 基于CAD数据的3D场景建模及演示技术研究[J]. 系统仿真学报, 2007, 19(7): 1504-1506.
- [6] 姜翰青, 章国锋, 董子龙, 等. 基于图像序列的交互式快速建模系统[J]. 计算机辅助设计与图形学学报, 2008, 20(9): 1196-1203.
- [7] LI Yang-yan, ZHENG Qian, SHARF A, *et al.* 2D-3D fusion for layer decomposition of urban facades [C]//Proc of IEEE International Conference on Computer Vision. [S. l.]: IEEE Press, 2011: 882-889.
- [8] XU Kai, ZHENG Han-lin, ZHANG Hao, *et al.* Photo-inspired model-driven 3D object modeling[J]. *ACM Trans on Graphics*, 2011, 30(4): 80-89.
- [9] WILCZKOWIAK M, STURM P, BOYER E. Using geometric constraints through parallelepipeds for calibration and 3D modeling[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2005, 27(2): 194-207.
- [10] HARTLEY A, ANDZISSERMAN A. Multiple view geometry in computer vision[M]. 2nd ed. Cambridge: Cambridge University Press, 2006.
- [11] LOURAKIS M. Levmar: Levenberg-Marquardt non-linear least squares algorithms in C/C++ [CP/OL]. (2011-11-29). <http://www.ics.forth.gr/~lourakis/levmar/>.
- [12] ROTHER C, KOLMOGOROV V, BLAKE A. GrabCut: interactive foreground extraction using iterated graph cuts[J]. *ACM Trans on Graphics*, 2004, 23(3): 309-314.
- [13] TELEA A. An image inpainting technique based on the fast marching method[J]. *Journal of Graphics Tools*, 2004, 9(1): 23-34.
- [14] DEBEVEC P E, TAYLOR C J, MALIK J. Modeling and rendering architecture from photographs: a hybrid geometry-and image-based approach [C]//Proc of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 1996: 11-20.

(上接第1900页)

- [5] KRAHNSTOEVEER N, YU T, PATWARDHAN K A, *et al.* Multi-camera person tracking in crowded environments [C]//Proc of the 12th IEEE International Performance Evaluation of Tracking and Surveillance Workshop. 2009:1-7.
- [6] MARCENARO L, FERRARI M, MARCHESOTTI L, *et al.* Multiple object tracking under heavy occlusions by using Kalman filters based on shape matching [C]//Proc of International Conference on Image Processing. 2002:341-344.
- [7] LIEN C C, HUANG Ya-lin, HAN C C. People counting using multi-mode multi-target tracking scheme [C]//Proc of the 5th Conference on IHH-MSP. 2009:1018-1021.
- [8] KHAN S, SHAH M. Tracking people in presence of occlusion [C]//Proc of Asian Conference on Computer Vision. 2000:1132-1137.

- [9] CHAN A, LIANG Z, VASCONCELOS N. Privacy preserving crowd monitoring: counting people without people models or tracking [C]//Proc of IEEE Conference on Computer-Vision and Pattern Recognition. 2008:1-7.
- [10] HARITAOGLU I, HARWOOD D, DAVIS L S. W4 real-time surveillance of people and their activities [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 809-830.
- [11] FEHR D, SIVALINGAM R, MORELLAS V, *et al.* Counting people in groups [C]//Proc of the 6th IEEE Conference on Advanced Video and Signal Based Surveillance. 2009:152-157.
- [12] YILMAZ A, JAVED O, SHAH M. Object tracking: a survey [J]. *ACM Computing Surveys*, 2006, 38(4): 13.
- [13] Van Der TUIN V. Computer-aided security surveillance design of the quo vadis object tracker [D]. Enschede: University of Twente, 2007.