

# 一种基于照片中纹理重构三维模型的方法<sup>\*</sup>

杨孟洲 石教英

(浙江大学 CAD&CG 国家重点实验室 杭州 310027)

E-mail: mzyang@cad.zju.edu.cn

**摘要** 如何从真实世界中获取具有真实感的三维场景模型一直是计算机图形学中的一个难点. 该文给出了一种从真实世界的照片中重建三维场景模型的算法. 算法根据在空间稀疏分布的不同视点处的真实场景照片中颜色纹理的一致性来建立达到照片级真实程度的三维场景模型, 可用于真实世界复杂形体真实感三维模型的建立.

**关键词** 基于照片建模, 三维模型, 一致性.

**中图法分类号** TP391

在计算机中如何表达真实世界一直是计算机图形学研究中的难点和热点. 目前主要有 3 类方法: 基于几何造型、基于图像绘制以及几何、图像混合的方法.

基于几何造型的方法通常利用成熟的造型软件手工搭建模型. 这种方法工作量大, 结果不够精确, 绘制效果也不够逼真. 基于图像绘制的方法<sup>[1,2]</sup>通常用场景的一个图像集合和与之对应的深度映射来表达场景模型. 基于图像绘制的主要优点是, 它提供了一种对结果图像中每个像素只需要固定计算量, 而与场景复杂度无关的绘制方法, 但对于大规模场景则需要很多照片, 使存储成为瓶颈, 而且, 系统能提供的场景交互手段有限, 对于很多应用领域不能满足要求. 把几何造型和基于图像绘制结合起来是一种非常有效的方法, 最常见的例子是纹理映射的应用.

如果能从真实世界直接获取几何模型, 将大大降低人工建模的工作量, 同时将大大提高绘制的逼真度. 很多学者对这种方法进行了研究. Debevec<sup>[3]</sup>给出了一种从建筑物的照片中提取有关信息重建三维模型的方法, 该方法由用户交互指出建筑物的组成构件、构件之间的相互关系以及构件的一些边在照片中的对应边, 根据建筑物的规整性及相互依赖性计算出照相机的参数和构件的位置、尺寸, 提取出相应的纹理, 搭建成具有照片级真实度的建筑物三维模型. Debevec 的照片建模算法只对具有规整性和相互关系明确的建筑物类场景有效, 对任意形状的场景则无能为力. 另外, 还有很多学者对如何从一组场景的照片中合成该场景在任意视点的新视图进行了研究, 提出了立体视觉匹配<sup>[4]</sup>、体元颜色赋值<sup>[5]</sup>等方法. Seitz 的体元颜色赋值模型重建方法根据场景在空间中稀疏分布的视点处照片中像素颜色之间的相关性来重建模型. Seitz 首先将建模空间离散成体元, 并对照片提出一个限制条件: 所有照片的相机位置形成一个空间凸包, 模型上任意一点都不能位于这个凸包之内, 否则将导致重建算法失败; 对于满足该条件的照片, 根据每个体元在输入照片上的投影覆盖的像素之间的相关性给这些体元赋以颜色值, 由这些有色体元来表示三维场景模型. Seitz 的模型重建算法的优点是, 对于模型外形的复杂程度没有限制, 同时重建过程不需要人工参与; 缺点是对相机位置有限制, 建立的场景模型只包含场景的一部分信息, 使得模型只能从有限的角度去观察, 同时用体元表示模型增加了计算中的复杂度.

本文给出的模型重建算法与 Seitz 的体元颜色赋值方法都是用像素的相似性来计算模型. 不同之处在于, 本算法对相机的位置没有任何限制, 这是优于体元颜色赋值方法之处, 同时解决了模型重建中的场景之间的遮挡

\* 本文研究得到国家自然科学基金(No. 69823003)资助. 作者杨孟洲, 1971 年生, 博士生, 主要研究领域为分布式虚拟现实技术. 石教英, 1937 年生, 教授, 博士生导师, 主要研究领域为计算机图形学, 多媒体计算技术, 分布式虚拟现实技术.

本文通讯联系人: 杨孟洲, 杭州 310027, 浙江大学 CAD&CG 国家重点实验室

本文 1999-01-18 收到原稿, 1999-04-14 收到修改稿

问题,并用面片来表示模型,使计算更为简单.

## 1 基于照片的三维模型重建

模型重建算法应该满足下面两个原则<sup>[5]</sup>.

(1) 照片完备性.将重建得到的三维模型重新投影,应该能精确地产生出输入图像,并且保持图像的颜色、纹理以及像素分辨率与输入图像相同.

(2) 视点覆盖范围宽.模型在一个宽范围内的视点处的重新投影应该精确,这就要求输入图像的视点在环境中稀疏分布.

将二维场景空间用3组分布于3个坐标平面上的平行平面分割开来.本算法要解决的问题就是给这些平面上分割成的面片赋以颜色值,使得对一组输入图像最大程度地满足照片完备性,即在相同视点处绘制出的图像应与原始图像尽可能地接近.

### 1.1 符号表示

假定场景和光照都是静止的,且场景的表面是朗伯表面.在这种条件下,场景表面上每一点的辐射度都具有各向同性特征,这样就可以用颜色来描述该点的辐射度.

我们以一个不透明、有色面片的有限集合  $S$  来表示三维场景表面,以  $F$  表示场景空间中由划分得到的所有面片的集合,以  $I$  表示一幅图像上所有像素的集合.同时,我们假设面片和像素无限小.

定义处于同一平面上的面片为一面片层  $F_i$ ,定义一个平行于该面片层的平面为参考平面  $s$ ,参考平面位于重建空间外,且面片的法向量指向该参考平面.  $f$  与  $s$  的距离为  $\|f\|$ ,则

$$F_i^d = \{f \mid \|f\| = d\}.$$

假设由重建空间中所有面片组成集合  $F$ ,即

$$F = \bigcup_{i=1}^r F_i^d, \quad d_1 < d_2 < \dots < d_r.$$

则组成场景空间表面的所有面片都包含在  $F$  中.

### 1.2 辅助定义

定义 1. 场景完备性.给定场景  $S$  和一幅视点在  $V$  的图像  $I$ ,存在一面片  $f \in S$ ,  $f$  对  $V$  可见,将  $f$  投影到  $I$  上得到像素  $p$ :  $f = S(p)$ . 记像素  $p \in I$  的颜色为  $color(p, I)$ ,面片  $f$  的颜色为  $color(f, S)$ . 假设对于一组图像  $I_1, I_2, \dots, I_n$  中的每一幅图像  $I_i$  及  $I_i$  上每一个像素  $p$ ,都存在一个面片  $f \in S$ ,使得  $f = S(p)$ ,我们就说场景  $S$  对这组图像完备.

定义 2. 场景一致性.假设对于一个完备的场景  $S$ ,有

$$color(p, I) = color(S(p), S), \quad \forall p \in I, \forall I \in \{I_1, I_2, \dots, I_n\},$$

我们就说场景  $S$  与图像  $I_1, I_2, \dots, I_n$  一致.

模型重建的任务是对于一组输入图像  $I_1, I_2, \dots, I_n$ ,求取一个面片集来表示三维场景模型:

$$S = \{f_p \mid p \in I_i, 1 \leq i \leq n\},$$

其中  $f_p$  为与输入图像一致的面片.

为了确保重建过程是逐步一致的,即重建的每一步都与输入图像一致,我们引入不完全面片集的弱一致性.

定义 3. 面片集弱一致性.假设一个面片集的投影与每一幅输入图像的重叠部分完全一致,我们就说它对输入图像弱一致.更确切地说,一个面片集  $S'$  中的每一个面片  $f \in S'$ ,假设对输入图像  $I_1, I_2, \dots, I_n$  上的像素点  $p \in I_i, q \in I_j$ ,面片  $f = S(p) = S(q)$ ,有

$$color(p, I_i) = color(q, I_j) = color(f, S),$$

那么  $S'$  对于图像  $I_1, I_2, \dots, I_n$  弱一致.

因此,对于面片集  $S_i^d = \bigcup_{i=1}^r \{f \mid f \in F_i^d \cap S, d_i < d\}$  来说,场景一致性与面片集弱一致性具有以下两个性质.

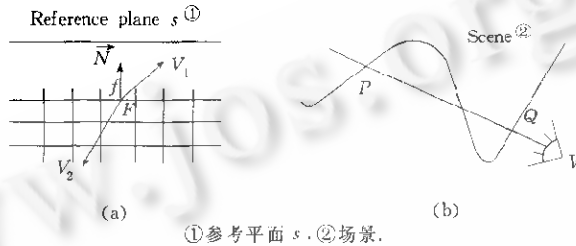
性质 1. 如果  $S$  是一致性场景,则对于每一个  $f \in S$ ,  $f$  与参考平面  $s$  的距离为  $d$ ,  $\{f\} \cup S_i^d$  是弱一致面片集.

性质 2. 假定场景  $S$  完备, 且对每一个  $f \in S, f$  与参考平面  $s$  的距离为  $d, \{f\} \cup S^d$  是弱一致面片集, 则  $S$  是一致性场景.

### 1.3 遮挡问题的处理

对于在某一视点可见的一部分场景, 当视点转移后可能会被另一部分场景遮挡, 即变为不可见. 在三维模型重建过程中, 如何判断场景是否被遮挡对重建成功与否至关重要. 如图 1(a) 所示,  $\vec{N}$  为面片  $f$  的法向量,  $V_1, V_2$  为视点,  $F$  为面片  $f$  的中点.  $\vec{N} \cdot \vec{FV}_1 > 0$ , 面片  $f$  对视点  $V_1$  有可能可见;  $\vec{N} \cdot \vec{FV}_2 \leq 0$ , 面片  $f$  对视点  $V_2$  不可见. 如果一个面片对一幅图像不可见, 则在重建该面片时不考虑该图像.

当  $\vec{N} \cdot \vec{FV} > 0$  时, 如图 1(b) 所示, 场景表面上两点  $P, Q$  和视点  $V$  位于一条直线上,  $P, Q$  对视点  $V$  都可能可见. 由于  $|PV| > |QV|$ , 因此对  $V$  来说  $P$  不可见, 这时, 我们就说  $Q$  遮挡住了  $P$ , 则在重建面片  $P$  时不考虑视点  $V$  的图像, 而在重建面片  $Q$  时要对该图像投影.



①参考平面  $s$ , ②场景.  
Fig. 1 Processing of occlusion  
图1 遮挡问题的处理

### 1.4 模型重建

重建算法根据归纳法逐步地计算出场景模型:

$$F_0 = F = \bigcup_{i=1}^r F_i^d, \quad d_1 < d_2 < \dots < d_r,$$

$$F'_k = \{f | f \in F_k^d, \{f\} \cup F_{k-1} \text{的投影中 } f \text{ 的投影与输入图像重叠部分不一致}\}, \quad 1 \leq k \leq r, \quad (1)$$

$$F_k = F_{k-1} - F'_k, \quad 1 \leq k \leq r, \quad (2)$$

$$F_{\text{result}} = \{f | f \in F_r, F_r \text{ 的投影中 } f \text{ 的投影与输入图像重叠部分一致}\},$$

则  $F_{\text{result}} = \bar{S}$  即为所求的场景模型.

证明: 首先我们定义  $\bar{S}_i = \{f | f \in S, \|f\|_s \leq d_i\}, \bar{S}'_i = \{f | f \in \bar{S}_i, \|f\|_s = d_i\}, F''_i = \{f | f \in F_i, \|f\|_s \leq d_i\}, F''_i = F_i^d - F'_i, 1 \leq k \leq r$ , 则有  $\bar{S}_i \subseteq F''_i$ , 且  $F''_i$  弱一致.

假设  $\bar{S}_{i-1} \subseteq F''_{i-1}, F''_{i-1}$  弱一致, 由一致性性质 1 有  $F''_i = F''_{i-1} \cup F''_i$  弱一致. 由式 (1), (2) 可知  $\bar{S}'_i \subseteq F''_i$ , 因此有  $\bar{S}_i - \bar{S}_{i-1} \cup \bar{S}'_i \subseteq F''_i$ .

由归纳法可得,  $F''_i$  弱一致,  $\bar{S} = \bar{S}_r \subseteq F''_r = F_r$ .

由于  $\bar{S} \cap \{f | f \in F_r, f \text{ 与输入图像不一致}\} = \emptyset$ , 因此  $\bar{S} \subseteq F_{\text{result}}$ , 并且  $F_{\text{result}}$  是弱一致面片集. 很明显, 由于重建过程中仅仅去掉了  $F$  中与输入图像不一致的面片, 因此  $F_{\text{result}}$  是完备的, 由一致性性质 2 可知,  $F_{\text{result}}$  是一致性场景, 因此  $F_{\text{result}} = \bar{S}$ . □

## 2 算法描述

本节给出了从一组图像  $I_1, I_2, \dots, I_n$  中重建三维场景模型的算法. 在一节, 我们假定将三维重建空间(空间六面体)用 3 组分布于 3 个坐标平面上的平行平面划分, 在这些平面上可以得到一组面片层  $F_1^d, \dots, F_6^d, d_1 < d_2 < \dots < d_6, 1 \leq i \leq 6$ . 如果一个面片  $f$  对于图像  $I_j$  没有被完全遮挡, 其投影将覆盖  $I_j$  上的一部分像素, 这些像素组成一个非空集合  $\pi_i$ . 在没有噪声及量化误差的情况下, 一致场景  $S$  中的面片  $f$  的投影将得到一组具有相同颜色的像素集合. 考虑到噪声及量化误差的影响, 就要用像素集  $\pi_1, \pi_2, \dots, \pi_i$  之间的相关性来判断一致性. 相关性由各个像素集中所有像素 RGB 的平均值的方差来衡量, 方差越小, 相关性越强, 反之亦然.

## 2.1 重建过程

重建空间中所有面片组成集合  $S_1$ , 取定重建空间某一面为参考平面, 首先对距离参考平面最近的面片层进行重建, 即从  $S_1$  中剔除该面片层中与输入图像不一致的面片, 然后对距离参考平面次近的面片层进行重建, 从  $S_1$  中剔除该面片层中与输入图像不一致的面片. 依次对所有平行于参考平面的面片层进行重建, 即剔除  $S_1$  中与输入图像不一致的面片. 取重建空间的另一面为参考平面, 对  $S_1$  中与之平行的面片层进行重建, 得到面片集  $S_2$ . 同样地, 依次将重建空间的另外 4 个表面作为参考平面, 按照距离参考平面由近及远的次序对面片层进行重建, 剔除面片层中与输入图像不一致的面片, 最后得到面片集  $S_6$ . 将  $S_6$  向每幅输入图像投影, 去掉其中对所有图像都不可见的片, 得到场景的三维模型  $\bar{S}$ .

## 2.2 重建算法

本节给出重建算法结构如下:

```

 $S_1 = F$ 
for  $k = 1, \dots, 6$ 
  for  $i = 1, \dots, r$  do
    for every  $f \in F_{i,k}^{d_1}$  do
      project to  $I_1, I_2, \dots, I_n$ , compute  $\lambda$ 
      if  $\lambda \geq thresh$  then  $S_1 = S_1 - \{f\}$ 
 $S = \emptyset$ 
for  $k = 1, \dots, 6$ 
  for  $i = 1, \dots, r$  do
    for every  $f \in F_{i,k}^{d_2} \cap S$  do
      project to  $I_1, I_2, \dots, I_n$ , compute  $\lambda$ 
      if  $\lambda < thresh$  then  $S = S \cup \{f\}$ 

```

$\lambda$  为  $\pi_1, \pi_2, \dots, \pi_j$  的方差,  $thresh$  是最大允许方差.  $thresh$  越小, 计算结果越精确, 但也可能越不完备; 反之,  $thresh$  越大, 计算结果越不精确, 但也越完备. 重建中遮挡的检测通过逐步远离参考平面的重建过程来解决, 即如果  $f_2$  被  $f_1$  遮挡, 则  $f_1$  必定先被重建. 具体到计算中, 首先对每一幅输入图像上的每一个像素设置标志为 0, 表示没有一致性面片投影到该像素上. 如果对面片  $f$  的投影有  $\lambda < thresh$ , 则将其投影在输入图像上的重叠部分的像素设置标志为 1. 以后再有面片投影到该像素上时, 就认为后者被遮挡.

## 3 实验结果

实验的输入图像由 3D MAX 在不同的照相机位置对一个三维犀牛模型加以渲染而获得, 即任意设定了 18 个照相机位置, 获得 18 幅图像. 其中相机位置符合在空间中稀疏分布的原则, 然后将这些图像输入重建程序, 同时给出相应的相机位置和其他相机参数, 开始重建. 考虑到在实际照片中照相机的位置不可能很精确地获得, 而且由 Tsai<sup>[6]</sup> 的照相机位置校正方法对真实照片校正后照相机的位置误差为 3%, 因此在重建时对输入图像照相机位置加了大约 3% 的随机扰动来模拟真实照片输入. 重建后输出一个彩色面片集来表示三维模型. 图 2 给出了重建的实验结果, 其中 (a) 是原始输入图像中的一幅, (b) 为与 (a) 在同一视点由重建后的模型绘制的图像, (c) 为重建模型在任意视点绘制的图像.

算法在 P.I 350 兼容机上实现, 软件环境为 Windows 95, VC++ 5.0.

## 4 结 论

本文给出了一种新的基于照片的三维场景重建算法, 根据不同照片中场景颜色纹理的相关性来建立达到照片级真实程度的三维场景模型. 算法突破了体元颜色赋值方法对照相机位置的限制, 使得可以对真实场景全方位拍摄照片以重建包含全部信息的三维场景模型, 解决了重建中场景之间的遮挡问题, 并且用面片代替体元来

表示模型,使计算更为简单.算法可用于建立真实世界复杂形体的真实感三维模型.

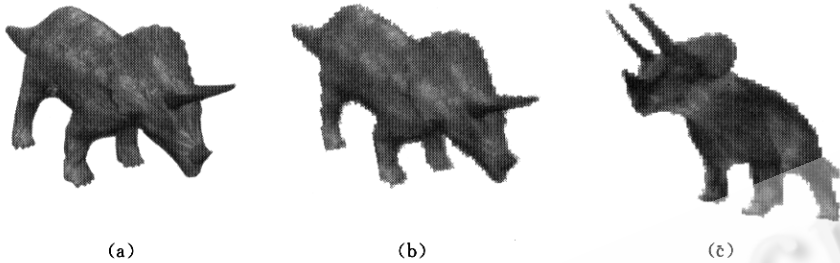


Fig. 2 The experimental results of rhinoceros model

图 2 犀牛模型实验结果

参考文献

- 1 Levoy M, Hanrahan P. Light field rendering. In: Rushmeier H ed. Computer Graphics Proceedings, Annual Conference Series. New York: ACM SIGGRAPH, 1996. 31~42
- 2 Gortler S J, Grzeszczuk R, Szeliski R *et al.* The lumigraph. In: Rushmeier H ed. Computer Graphics Proceedings, Annual Conference Series. New York: ACM SIGGRAPH, 1996. 43~54
- 3 Debevec P E, Taylor C J, Jitendra Malik. Modeling and rendering architecture from photographs: a hybrid geometry-and image-based approach. In: Rushmeier H ed. Computer Graphics Proceedings Annual Conference Series. New York: ACM SIGGRAPH, 1996. 11~20
- 4 Maximilian Ott, Lewis J P, Ingernar Cox. Teleconferencing eye contact using a virtual camera. 1993, ftp://ftp.nj.nec.com/pub/ingemar/papers/interchi.ps.Z
- 5 Setiz S M, Dyer C R. Photo realistic Scene Reconstruction by Voxel Coloring. 1997, http://www.cs.wisc.edu/computer-vision/cvpr97-seitz.ps
- 6 Tsai R Y. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf cameras and lenses. IEEE Transactions on Robotics and Automation, 1987,3(4):323~344

A 3D Model Reconstruction Algorithm Based on the Texture of Photos

YANG Meng-zhou SHI Jiao-ying

(State Key Laboratory of CAD & CG Zhejiang University Hangzhou 310027)

**Abstract** In computer graphics, it's difficult to acquire realistic 3D scenes from the real world. In this paper, a novel 3D scene model reconstruction algorithm is presented to deal with this problem. This method reconstructs photo-realistic scene models according to the color and texture of basic input photos, which are distributed in the environment. This method can be used to reconstruct photo-realistic 3D model of complex real objects.

**Key words** Photo-Based modeling, 3D model, consistency.