

# 基于双目立体视觉三维重建系统的研究与设计

于 勇, 张 晖, 林茂松

(西南科技大学 计算机科学与技术学院, 四川 绵阳 621010)

**摘 要:** 基于双目的立体视觉方法, 介绍了一套由双工业 CCD 构成的立体视觉系统, 设计了一套切实可行的三维重建系统方案, 其中包括图像获取模块, 基于 OpenCV 的摄像机标定系统, SIFT 算法实现特征点提取与立体匹配, 深度信息计算, OpenGL 三维模型重构几大模块。系统各模块经过试验测试和验证, 能够通过两幅图像恢复出物体三维可见表面几何形状, 充分发挥了 OpenCV 函数库的功能, 基本上能满足三维重建目标的要求, 尤其对城市景观的三维重建有较大应用价值。

**关键词:** 三维重建; 双目视觉; OpenCV; 张正友标定; SIFT 算法

**中图分类号:** TP18

**文献标识码:** A

**文章编号:** 1673-629X(2009)06-0127-04

## Research and Design of the 3D Reconstruction System Based on Binocular Stereo Vision

YU Yong, ZHANG Hui, LIN Mao-song

(Sch. of Computer Sci. & Tech., Southwest University of Sci. and Tech., Mianyang 621010, China)

**Abstract:** Based on binocular stereo vision methods, introduced a three-dimensional visual system constituted by a pair of industrial CCD, designed a practical system of three-dimensional programme, including image acquisition module, camera calibration system based on the OpenCV, feature extraction and matching based on SIFT algorithm, in-depth information calculated, 3D model reconstruction based OpenGL several modules. After the test and verify, it can restoration 3D visible surface geometry of the object through two images, give full play to the OpenCV library functions, it can basically meet the requirements of three-dimensional reconstruction objectives, in particular has a greater value in the urban landscape reconstruction.

**Key words:** 3D reconstruction; binocular; OpenCV; Zhengyou Zhang calibration; SIFT

## 0 引 言

用立体视觉方法进行三维重建<sup>[1]</sup>, 在计算机视觉中是指由两幅或多幅二维图像恢复物体三维可见表面几何形状的方法。对生物视觉系统, 人们早就注意到, 几乎所有具有视觉的生物都有两个眼睛。用两个眼睛同时观察物体时, 会有深度或远近的感觉。立体电影之所以有逼真的深度感, 也是仿照了立体视觉原理。在立体电影拍摄中, 用两个摄像机同时拍摄, 而在放映时, 将两个摄像机拍摄的图像同时投影到屏幕上, 并利用偏振光的原理, 使人的双眼分别看到左右摄像机拍摄的图像, 从而使人感到真实三维场景的立体感。在计算机立体视觉系统中, 也可以利用摄像机从不同角度获取同一景物的两幅图像, 然后利用三维重建原理,

由计算机重建景物的三维形状, 恢复出物体的空间位置信息。

## 1 三维重建原理

### 1.1 图像获取

立体图像的获取是立体视觉的基础, 图像获取<sup>[2]</sup>的方式很多, 主要取决于应用的场合和目的, 还要考虑视点差异、光照条件、摄像机性能以及景物特点等因素的影响, 以利于立体计算。

### 1.2 摄像机标定

摄像机标定<sup>[3]</sup>是为了建立成像模型、确定摄像机的位置和属性参数, 以确定空间坐标系中物体点同它的像点之间的对应关系。两个摄像机都要标定, 如果摄像机是固定的, 当从 2-D 计算机图像坐标推导 3-D 信息时, 只需一次标定。

摄像机的标定方法可分为两大类: 第一类是直接估计摄像机的位置、光轴方向、焦距等参数; 第二类是通过最小二乘法拟合, 确定三维空间点映射为二维图

收稿日期: 2008-09-18; 修回日期: 2008-12-05

基金项目: 国家自然科学基金(10676029)

作者简介: 于 勇(1985-), 男, 山东安丘人, 硕士研究生, 研究方向为机器视觉; 张 晖, 博士, 副教授, 研究方向为数据挖掘与模式识别; 林茂松, 硕士, 副教授, 研究方向为机器视觉与图像处理。

像点的变换矩阵。一个有效的摄像机模型,除了能够精确地恢复出空间景物的三维信息外,还有利于解决立体匹配问题。

### 1.3 特征提取

由多视点的视差确定 3-D 信息,关键是确定场景中同一物点在不同图像中的对应关系,解决该问题的方法之一是选择合适的图像特征<sup>[2]</sup>并进行匹配。特征是像素或像素集合或它们的抽象表达,常用的匹配特征主要有点状特征、线状特征和区域特征等。一般来讲,大尺度特征含有较丰富的信息,本身数目较少,易于得到快速匹配,但对它们的提取与描述相对复杂,定位精度也差;而小尺度特征本身定位精度高,表达描述简单,但数目较多,所含信息量却较少,因而在匹配时需要采用较强的约束准则和匹配策略。

### 1.4 立体匹配

立体匹配<sup>[2]</sup>是根据所选特征的计算,建立特征间的对应关系,将同一个空间点在不同图像中的映像点对应起来,并由此得到相应的视差图像,立体匹配是双目视觉最重要也是最困难的问题。当空间三维场景被投影为二维图像时,同一景物在不同视点下的图像会有很大不同,而且场景中的诸多因素,如光照条件、景物几何形状和物理特性、噪声干扰和畸变以及摄像机特性等,都被综合成单一的图像灰度值。因此,要准确地对包含了如此之多不利因素的图像进行无歧义匹配十分困难。

匹配方法需要解决以下问题:选择正确的匹配特征;寻找特征间的本质属性;建立能正确匹配所选特征的稳定算法。

### 1.5 三维重建

在得到空间任一点在两个图像中的对应坐标和两摄像机参数矩阵的条件下,即可进行空间点的重建<sup>[4]</sup>。通过建立以该点的世界坐标为未知数的 4 个线性方程,可以用最小二乘法求解得该点的世界坐标。实际重建通常采用外极线结束法。空间眯、两摄像机的光心这三点组成的平面分别与两个成像平面的交线称为该空间点在这两个成像平面中的极线。一旦两摄像机的内外参数确定,就可通过两个成像平面上的极线的约束关系建立对应点之间的关系,并由此联立方程,求得图像点的世界坐标值。

## 2 系统设计与实现

### 2.1 图像获取模块

基于应用场合的需要,采用北京微视新纪元科技有限公司提供的工业 CCD, V411 型号的图像采集卡,此采集卡支持 DirectShow,拥有私有 SDK,正是在此

SDK 基础上做二次开发,从不同角度采集图像<sup>[5]</sup>,其中包括单目图像采集、双目图像采集、多目图像采集(最多 4 个)几大功能模块。

### 2.2 摄像机标定模块

采用张正友的平面标定方法<sup>[6]</sup>,这种方法是介于传统标定方法和自标定方法之间的一种方法,它避免了传统方法设备要求高、操作繁琐等缺点,又较自标定方法精度高,符合办公、家庭使用的桌面视觉系统(DVS)的标定要求。

张正友标定算法<sup>[7]</sup>的基本原理可用式 1 表示。

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (1)$$

这里假定模板平面在世界坐标系  $Z = 0$  的平面上,其中  $K$  为摄像机的内参数矩阵,  $\tilde{M} = [X \ Y \ 1]^T$  为模板平面上点的齐次坐标,  $\tilde{m} = [u \ v \ 1]^T$  为模板平面上点投影到图像平面上对应点的齐次坐标,  $[r_1, r_2, r_3]$  和  $t$  分别是摄像机坐标系相对于世界坐标系的旋转矩阵和平移向量。

具体实现方法描述<sup>[8]</sup>:

- (1) 打印一张模板(如图 1 所示)并贴在一个平面上;
- (2) 从不同角度拍摄若干张模板图像;
- (3) 检测出图像中的特征点;
- (4) 求出摄像机的内参数和外参数;
- (5) 求出畸变系数;
- (6) 优化求精。

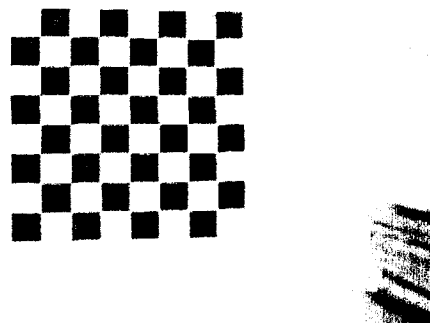


图 1 系统所采用的平面模板

立体视觉系统的两个摄像机的相对位置也可以通过对标定得到。但摄像机定时,分别得到两个摄像机的内外参数,设  $R_l, t_l$  和  $R_r, t_r$  分别表示左、右摄像机与世界坐标系的相对位置,利用下式<sup>[9]</sup>可以求出摄像机的相对位置,即

$$\begin{cases} R = R_l R_l^{-1} \\ t = t_r - R_l^{-1} t_l \end{cases} \quad (2)$$

### 2.3 特征点提取与立体匹配

采用相关特征匹配<sup>[10]</sup>算法。David G. Lowe<sup>[11]</sup>在2004年总结了现有的基于不变量技术的特征检测方法,并正式提出了一种基于尺度空间的、对图像缩放、旋转甚至仿射变换保持不变性的图像局部特征描述算子——SIFT算子,其全称是 Scale Invariant Feature Transform,即尺度不变特征变换。

SIFT算法<sup>[12]</sup>首先在尺度空间进行特征检测,并确定关键点(Keypoints)的位置和关键点所处的尺度,然后使用关键点邻域梯度的主方向作为该点的方向特征,以实现算子对尺度和方向的无关性。

一幅图像SIFT特征向量的生成算法包括4步:

(1) 尺度空间极值检测,以初步确定关键点位置和所在尺度。

(2) 通过拟和三维二次函数以精确确定关键点的位置和尺度,同时去除低对比度的关键点和不稳定的边缘响应点,以增强匹配稳定性、提高抗噪声能力。

(3) 利用关键点邻域像素的梯度方向分布特性为每个关键点指定方向参数,使算子具备旋转不变性。

通过上面三个步骤,图像的关键点就已经检测完毕,每个关键点有三个信息:位置、所处尺度、方向。由此可以确定一个SIFT的特征区域(在实验的时候通常用椭圆或箭头表示)。

(4) 生成SIFT特征向量(如图2所示)。

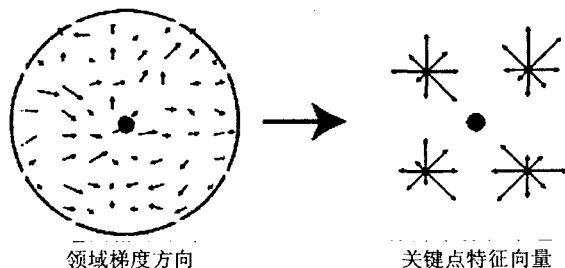


图2 由关键点邻域梯度信息生成特征向量

图中一个关键点由 $2 \times 2$ 共4个种子点组成,每个种子点有8个方向向量信息,实际计算过程中,为了增强匹配的稳健性,对每个关键点使用 $4 \times 4$ 共16个种子点来描述,这样对于一个关键点就可以产生128个数据,即最终形成128维的SIFT特征向量。

两幅图像的SIFT特征向量生成后,采用关键点特征向量的欧式距离来作为两幅图像中关键点的相似性判定度量。取图像1中的某个关键点,并找出其与图像2中欧氏距离最近的前两个关键点,在这两个关键点中,如果最近的距离除以次近的距离少于某个比

例阈值,则接受这一对匹配点。降低这个比例阈值,SIFT匹配点数目会减少,但更加稳定。

### 2.4 三维重建

在立体匹配和摄像机标的基础上进行三维重建<sup>[1,13,14]</sup>,利用SFM<sup>[14,15]</sup>(Structure from Motion)算法恢复出摄像机外参数,然后求出离散空间点的三维坐标,经过Delaunay三角划分和贴纹理,最终通过OpenGL编程显示出物体的三维模型。

SFM方法的算法步骤:

- (1) 从基础矩阵<sup>[16]</sup>中求出本质矩阵 $E$ ;
- (2) 求解摄像机外参数 $R, t$ 的候选值;
- (3) 判断 $R, t$ 的符号,确定唯一合理的一组值;
- (4) 计算空间的三维坐标值。

三维点的空间三角化采用基于Delaunay<sup>[17]</sup>代数的三角划分(Delaunay triangulation),划分后得到许多小三角形,每个三角形作为纹理的一个单元。实现中,主要是按照Jonathan Richard Shewchuk<sup>[18]</sup>提出的方法在二维图像平面进行基于匹配点的三角划分,再将此划分反投回三维空间,就得到了空间点集的划分。

三维重建结果的可视化过程如图3所示。

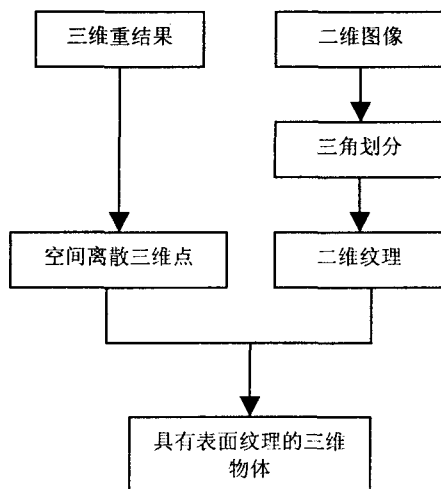


图3 三维重建结果的可视化

对图像进行二维三角化后,三角化的二维点列反投回三维空间,就得到空间点网格。在此过程中,三角化的二维点列将图像分割,形成很小的三角化的纹理,将对应的三角纹理贴回到对应的空间点网格,就得到了具有纹理的三维物体。最后通过OpenGL<sup>[19]</sup>的函数glTexImage2D进行纹理映射。

## 3 结束语

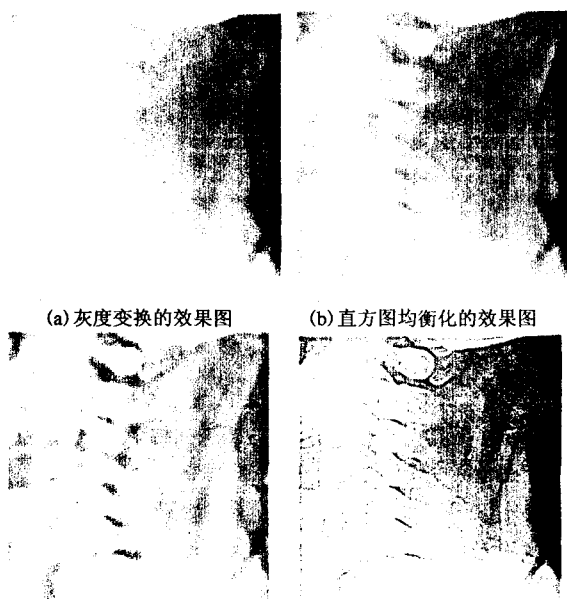
对基于双目立体视觉的三维重建这一热点问题及其子问题,进行了较为系统的研究,在分析和总结现有各种方法优缺点的基础上,提出了一套切实可行的方

案,设计了一套稳定可靠的三维重建系统,可以通过两幅图像恢复出物体的三维形状,结果较为满意,为系统的进一步改进和完善打下了基础,本系统将主要用于城市景观的三维重建,目的是能够对建筑物、绿地、河流等三维模型实现快速精确的重建,对“数字化城市”的建设有较为广阔的应用价值。

#### 参考文献:

- [1] 王新宇. 基于计算机立体视觉的三维重建[D]. 长沙:中南大学,2004.
- [2] 李健,史进. 一种基于 OpenCV 的三维重建实现方案[J]. 计算机与信息技术,2008(4):47-49.
- [3] 马颂德,张正友. 计算机视觉(计算理论与算法基础)[M]. 北京:科学出版社,2003:52-92.
- [4] 隋婧,金伟其. 双目立体视觉技术的实现及其进展[J]. 计算机应用,2004(10):4-6.
- [5] 何海,汤春林,孙华燕. 双目立体视觉在模型姿态监测中的应用研究[J]. 计算机技术与发展,2006,16(11):238-240.
- [6] Zisserman H R. A Multiple View Geometry in Computer Vision[M]. Cambridge: Cambridge University Press,2000.
- [7] Zhang Z. A Flexible New Technique for Camera Calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2000,22(11):1330-1334.
- [8] Zhang Zhengyou. Flexible Camera Calibration By Viewing a Plane From Unknown Orientations[D]. [s.l.]: ICCV,1999: 666-673.
- [9] 蔡健荣,赵杰文. 双目立体视觉系统摄像机标定[J]. 江苏大学学报:自然科学版,2006,27(1):6-9.
- [10] Barnard S T, Fischler M A. Computational stereo[J]. ACM Computing Surveys,1982,14(4):553-572.
- [11] Lowe D G. Distinctive image features from scale invariant key-point[J]. International Journal of Computer Vision,2004,60(2):91-110.
- [12] 胡志勇,张秀芬,蔚刚,等. 基于灰度图像的三维曲面重建系统设计[J]. 计算机技术与发展,2007,17(10):61-64.
- [13] 崔明. 基于图像序列的三维模型重建[D]. 杭州:浙江大学,2005.
- [14] Hartley R. Estimation of relative camera position from uncalibrated cameras[C]// In Proceedings of European Conference on Computer Vision. [s.l.]: Springer-Verlag,1992.
- [15] Zhang Z. A New Multistage Approach to Motion and Structure Estimation: From Essential Parameters to Euclidean Motion via Fundamental Matrix[R]. France: INRIA Sophia-Antipolis,1996.
- [16] Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections[J]. Nature,1981,293:133-135.
- [17] 管镭,孟宪琦,魏生民. Delaunay 三角网络化算法及实现[J]. 西北工业大学学报,1996,14(1):138-142.
- [18] Shewchuk J R. Delaunay Refinement Mesh Generation[D]. USA: School of Computer Science, Carnegie Mellon University,1995.
- [19] OpenGL 体系结构审核委员会. OpenGL 编程指南[M]. 第4版. 邓郑祥译. 北京:人民邮电出版社,2005.

(上接第 126 页)



(c) 对比度自适应直方图均衡化的效果图 (d) 文中的效果图

图 4 算法对比效果图

#### 参考文献:

- [1] 张淑骅. 基于 MATLAB 的 X 光图像处理[D]. 上海:华东师范大学,2005.
- [2] 尹立敏,刘艳滢,顾蕊,等. 一种可控的直方图均衡算法[J]. 微计算机信息,2005,21(12-3):147-148.
- [3] 罗军辉,冯平. MATLAB7.0 在图像处理中的应用[M]. 北京:机械工业出版社,2005.
- [4] Paranjape R B. Adaptive neighborhood histogram equalization for image enhancement[J]. CVGIP,1992,54(3):259-267.
- [5] 周德龙,潘泉,张洪才. 自动图像边缘检测方法的研究[J]. 计算机工程与科学,2001,23(4):11-14.
- [6] Kundu A, Zhou J. Combination media filter[J]. IEEE Transaction on Image Processing,1992(1):422-429.
- [7] 黄伟,周鸣争,李小牛. 基于滤波器叠加的彩色图像矢量中值滤波方法[J]. 计算机技术与发展,2008,18(1):143-145.
- [8] Gonzalez R C, Woods R E. Digital Image Processing (影印版)[M]. 北京:电子工业出版社,2002.
- [9] 吴燕,张晨燕. 基于目标特征的灰度直方图重分布图像增强算法[J]. 计算机与信息技术,2007(8):14-15.