

双目视觉的台球桌面的三维重构^①

张旭飞, 王朝立, 袁 伟

(上海理工大学 光电信息与控制工程学院, 上海 200093)

摘 要: 随着计算机技术的迅猛发展, 在越来越多的体育比赛中纷纷出现了电子裁判来辅助裁判进行判罚工作, 以实现奥林匹克的公平竞赛精神. 而在实现这种数字化裁判的过程中, 摄像机在计算机视觉中起着重要的作用, 三维重构就是根据单视图或者多视图的图像重建三维信息的过程. 主要以计算机视觉中的三维重构技术作为研究对象, 具体详述了有关双目视觉的一些理论知识, 充分利用计算机视觉库 OpenCV 提供的重建算法, 来完成对台球桌面的三维重构工作. 台球桌面的三维重构具有很强的现实意义, 为电子裁判进入桌球比赛打下了良好的基础.

关键词: 三维重建; 双目视觉; 台球桌面; OpenCV

3D Reconstruction of Desktop of Table Tennis Based on the Binocular Vision

ZHANG Xu-Fei, WANG Chao-Li, YUAN Wei

(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: At present, with the rapid development of the computer technologies which has grown dramatically, and there are more and more sports competition where some high and new technologies are employed as the electronic referee to assist referee to do the decisions, in order to achieve the spirit of fair play which respected by Olympic. During this process, the camera plays an important role in computer vision. 3D reconstruction is a process based on a single view or multi-view images to obtain the 3D information in the field of computer vision. In this paper, in order to complete a task about three-dimensional reconstruction of desktop of Table tennis, we take 3D reconstruction technologies in computer vision as the mainly research object, including specific details about the theoretical knowledge on the binocular vision and the reconstruction algorithm provided by the computer vision library OpenCV. The 3D reconstruction of billiards desktop have strong practical significance, which laid a good foundation for the electronic referee to work in the billiards competition.

Key words: 3D reconstruction; binocular vision; desktop of table tennis; OpenCV

双目视觉的优点是可以用来模仿人的眼睛, 它是由不同位置的两台摄像机 (CCD) 经过移动或旋转拍摄同一幅场景, 通过计算空间点在两幅图像中的视差, 获得该点的三维坐标值. 在得到场景的一系列空间点坐标以后, 利用这些坐标信息来实现对场景的三维重构. 三维重构技术的运用相当之广泛, 不仅仅是在工业方面, 建筑方面, 文化方面, 三维重建技术的应用更是在渗透到医学方面. 同样, 三维重构技术也是虚

拟现实最最基础的技术之一.

本文则是从实际出发, 将台球桌面作为切入点, 来实现台球桌面的三维重构^[1]. 其意义在于完成台球桌面的三维重构之后, 就可以随时准确定位球在桌面的具体位置. 这样就使得在台球比赛的过程中, 对于各种可能出现的肉眼无法进行准确判断的情况(台球被打出桌面重新放置等), 利用此技术来达到辅助裁判进行判罚工作, 从而也使得结果更为准确, 更无争议.

① 基金项目:国家自然科学基金(60874002);上海市研究生创新基金(JWCXSL1102);上海市教委重点科技创新项目

收稿时间:2012-08-29;收到修改稿时间:2012-09-27

本文的主要创新之处在于: 1. 台球比赛中, 经常会出现的是裁判来对台球进行重新放置, 这就不可避免的会引起裁判与队员之间看法上存在分歧, 从技术上解决这一问题, 是本文的主要目的. 有鉴于网球比赛中的“鹰眼”的技术, 可以很好的辅助裁判来对网球的落点是否在边线内来做出准确的判断. 因此, 我们可以仿效该项技术原理, 创造一套双目视觉系统来实现对台球的实时定位, 做到实时定位桌面上任一球在任意时刻的具体位置. 通过该系统不仅可以帮助运动员平时进行台球训练, 让运动员更加清晰的观察到台球撞击后的走向, 更多的是帮助裁判比赛时进行判罚, 从而使得比赛更加公开和公平. 2. 考虑到图像拍摄过程中环境因素的影响, 本文在图像处理方面进行了多方面的比较, 通过具体实验结果来选取图像处理的方法, 其中包括滤波、锐化、边缘提取等. 此外, 在进行特征提取之前, 还对二值图像进行了形态学处理, 使得特征更为明显, 从而更加有利于特征点的提取.

1 摄像机的标定

1.1 摄像机标定原理^[2,3]

摄像机标定的方法有很多, 下面以 OpenCV 标定为例, 简述其标定原理. 在描述相机内外参数之前, 首先引入最基本的摄像机成像模型来了解摄像机的成像原理, 如下图 1 所示. 其中 $O_F(U_P, V_P)$ 表示物理坐标系(用来描述点物理尺寸坐标)原点 O_F 在图像坐标系(用来描述点的像素坐标)下的坐标, $P_C(X_C, Y_C, Z_C)$ 和 $P_W(X_W, Y_W, Z_W)$ 分别表示空间任一点 P 在摄像机坐标系下的坐标和世界坐标系下的坐标. $P(X_P, Y_P)$ 和 $P(X_F, Y_F)$ 分别表示点 P 的图像坐标与物理坐标, f 表示摄像机的焦距, S_x 与 S_y 分别表示在 X 方向和 Y 方向上单位长度的像素个数. 通过一些列的几何推导以及坐标变换可得如下的几何关系:

$$\begin{bmatrix} X_P \\ Y_P \\ 1 \end{bmatrix} = \frac{1}{Z_C} \begin{bmatrix} f_x & 0 & U_P & 0 \\ 0 & f_y & U_V & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \frac{1}{Z_C} A P_C \quad (1)$$

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3 \times 3} & t_{3 \times 1} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} X_P \\ Y_P \\ 1 \end{bmatrix} = \frac{1}{Z_C} A \begin{bmatrix} R_{3 \times 3} & t_{3 \times 1} \\ 0^T & 1 \end{bmatrix} P_W \quad (3)$$

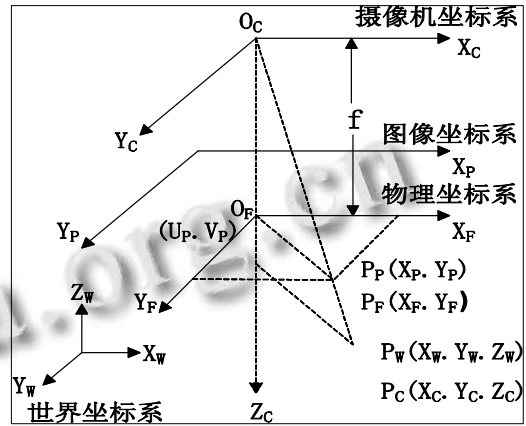


图 1 摄像机成像模型图

其中 $f_x = f * S_x$, $f_y = f * S_y$, U_P, V_P 为摄像机的内参数, 矩阵 A 称为内参数矩阵. $R_{3 \times 3}$ 是旋转矩阵, $t_{3 \times 1}$ 是平移向量, 二者综合称为摄像机的外参数.

通过摄像机的标定, 可以得出摄像机的内外参数. 基于 OpenCV 的摄像机标定通常是采用标定板, 在不同位置拍摄多幅图片后, 通过 OpenCV 提供的一系列标定函数相继完成图片角点的提取, 亚像素精确化, 计算出内外参数值以及畸变系数, 最后进行误差分析得出标定结果. 具体流程图如下图 2 所示:

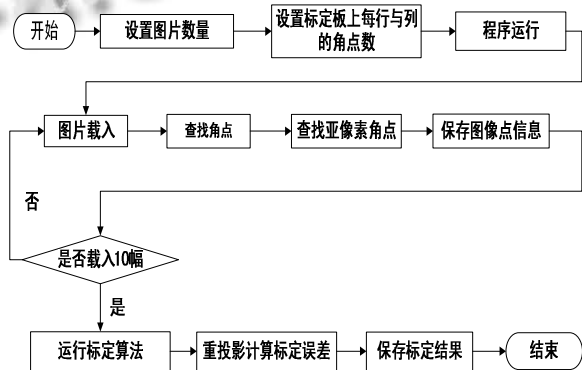


图 2 摄像机标定的流程图

OpenCV 中采用的标定方法是介于传统标定方法和自标定方法之间的一种方法需要一个特定的标定物以及一组已知的特征基元的坐标, 由于此种标定方法

精度不错, 避免了传统方法设备要求高、操作繁琐等缺点, 故被 OpenCV 所采纳. 其标定的结果见下表 1.

表 1 左右摄像机标定结果

| | 左边摄像机 | 右边摄像机 |
|-----------|-------------------------------------|--------------------------------------|
| 内参数 矩阵 | {8552.2,0,770.9} | {8375.15,0,778.2} |
| | {0,7031.7,730.5} | {0,7151.58,727.91} |
| | {0,0,1} | {0,0,1} |
| 畸变 系数 | {-2.829,14.293, 0.19297,-0.0357} | {-1.5798,17.282, -0.18569,0.0350} |

1.2 摄像机标定的意义

摄像机的标定作为立体视觉的基础部分, 它标定结果的好与坏将直接影响立体视觉的后续部分. 因此, 我们必须综合考虑多方面因素, 找到适合实验室进行的标定方法.

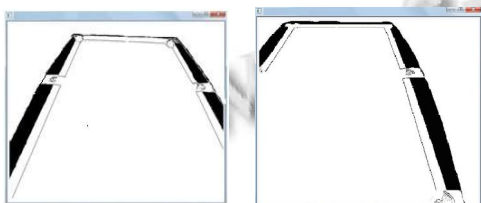
2 特征点匹配

2.1 图像处理^[4]

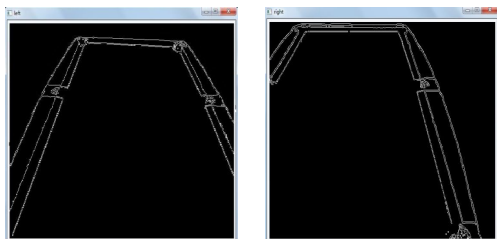
在进行特征点匹配之前, 我们必须对该图像对进行处理. 二维图像的处理包括图像的平滑滤波、对比度增强、图像特征提取等. 实验结果如下图 3 所示:



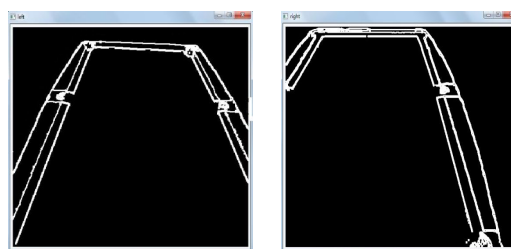
(a) 实验中用到的原始左右图像对



(b) 二值化后得到的左右图像对



(c) Canny 算子提取轮廓左右图像对



(d) 膨胀后得到的左右图像对

图 3 图像处理效果图

2.2 极线几何^[5]

对于空间的任一点 P , 它在两个摄像机成像平面上的投影点分别为 P_1 和 P_2 , 如图 4 所示. C_1 和 C_2 分别为两个摄像机的中心, 即摄像机坐标系的原点. 而在极线几何中, C_1 和 C_2 的连线又称之为基线. 基线和两摄像机成像平面的交点 e_1 和 e_2 则称之为两摄像机的极点, 它们分别为两个摄像机中心 C_1 和 C_2 在对应的摄像机成像平面上的投影坐标. 此外, P 、 C_1 和 C_2 组成的三角平面称为极平面 π . π 和两个摄像机成像平面的交线 I_1 和 I_2 则称之为极线, I_1 为点 P_2 对应的极线, I_2 为点 P_1 对应的极线, I_1 和 I_2 互为对应极线. 通过在极平面 π 上另取一点 P' , 从图中可以看到, 它在两摄像机平面上的投影点为 P_1 和 P_2' , 其中 P_2 和 P_2' 都在极线 I_2 上. 这就是所谓的极线约束, 由此, 可以将我们的搜索空间压缩到一维的直线, 即极线上. 从而大大提高了匹配的精度, 减少了匹配的时间.

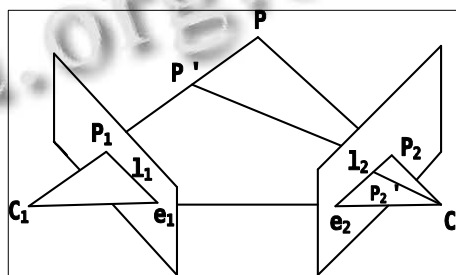


图 4 双摄像机成像模型

2.3 匹配方法简介

特征点领域的研究者提出了很多匹配方法, 究其本质主要有两种, 一种是稀疏点匹配, 另一种是密集点匹配.

稀疏点匹配主要是对图像中的一些强特征点进行匹配, 该方法匹配结果比较可靠, 而且速度快, 效率高, 缺点则是插值过程就非常盲目. 密集点匹配则是

基于图像对中的所有的像素点进行的匹配,最大限度的恢复物体的细节特征,但是容易发生误匹配.

权衡上述两种匹配方法的优缺点,结合具体的实验环境,在进行台球桌面三维重建的试验时选取的是稀疏点匹配,最后经过插值得到的视差图如下图 5 所示:

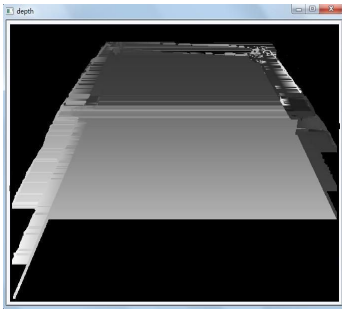


图 5 计算得到的视差图

3 三维重构

在进行了特征点的匹配以后,紧接着要进行三维重构了.空间点三维重构的方法描述如下:假设有一对匹配点 $m_1(u_1 v_1 1)^T$ 和 $m_2(u_2 v_2 1)^T$,则由二维图像点与三维图像点之间的映射关系可得如下两式:

$$m_1 = k_1 \cdot P_1 \cdot M \tag{4}$$

$$m_2 = k_2 \cdot P_2 \cdot M \tag{5}$$

其中 P_1 、 P_2 为对应图像的透视投影矩阵, $M(XYZ1)^T$ 为空间点三维坐标, k_1 、 k_2 为比例系数,将上述两式展开可得如下表达式:

$$\begin{bmatrix} u1 \\ v1 \\ 1 \end{bmatrix} = k_1 \cdot \begin{bmatrix} m_{11}^1 & m_{12}^1 & m_{13}^1 & m_{14}^1 \\ m_{21}^1 & m_{22}^1 & m_{23}^1 & m_{24}^1 \\ m_{31}^1 & m_{32}^1 & m_{33}^1 & m_{34}^1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{6}$$

$$\begin{bmatrix} u2 \\ v2 \\ 1 \end{bmatrix} = k_2 \cdot \begin{bmatrix} m_{11}^2 & m_{12}^2 & m_{13}^2 & m_{14}^2 \\ m_{21}^2 & m_{22}^2 & m_{23}^2 & m_{24}^2 \\ m_{31}^2 & m_{32}^2 & m_{33}^2 & m_{34}^2 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{7}$$

由上述两式可得到如下四个式子并联立,得到了如下关于 X,Y,Z 三个变量的如下方程组:

$$\begin{cases} (u_1 m_{31}^1 - m_{11}^1)X + (u_1 m_{32}^1 - m_{12}^1)Y + (u_1 m_{33}^1 - m_{13}^1)Z = m_{14}^1 - u_1 m_{34}^1 \\ (v_1 m_{31}^1 - m_{21}^1)X + (v_1 m_{32}^1 - m_{22}^1)Y + (v_1 m_{33}^1 - m_{23}^1)Z = m_{24}^1 - v_1 m_{34}^1 \\ (u_2 m_{31}^2 - m_{11}^2)X + (u_2 m_{32}^2 - m_{12}^2)Y + (u_2 m_{33}^2 - m_{13}^2)Z = m_{14}^2 - u_2 m_{34}^2 \\ (v_2 m_{31}^2 - m_{21}^2)X + (v_2 m_{32}^2 - m_{22}^2)Y + (v_2 m_{33}^2 - m_{23}^2)Z = m_{24}^2 - v_2 m_{34}^2 \end{cases} \tag{8}$$

由上述方程组取最小二乘意义上的一组解视为特征点的三维坐标.再将这些特征点的三维坐标导入到 OpenGL^[6]中,接着用坐标点绘制方法绘制出桌面的三维图像,实验结果如下图 6 所示:

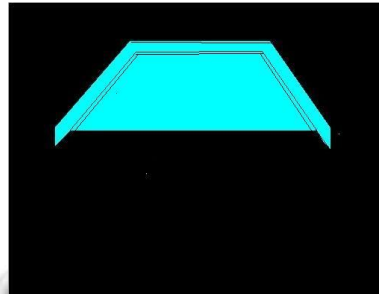


图 6 三维重建效果图

4 结束与展望

三维重构技术作为计算机视觉理论的一个重要分支,一直都是研究的热点与难点,应用领域^[7-8]不断拓宽.本文主要以双目视觉系统作为实验平台,通过 OpenCV 与 OpenGL 的混合编程来实现台球桌面的三维重构,程序运行效率较高,实验结果较为精确,简单易行.其不足之处在于对实验稳定性不高,抗外界因素干扰能力不强.考虑到实验环境的复杂性与随机性,以及对精度要求的不断提高,我们必须不断尝试,不断创新,努力在已有的技术基础上拓展出新思路新方法.

参考文献

- 1 李向荣.美式落袋球自动摆球系统的图像处理研究.仪器仪表学报,2008,29(4):551-553.
- 2 陈胜勇,刘盛等.基于 OpenCV 的计算机视觉技术实现.北京:科学出版社,2008.
- 3 Zhang ZY. A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, November, 2000,22(11):1330-1334.
- 4 张宁蓉.计算机图形学.西安:西安电子科技大学出版社,2011.
- 5 张晓明.基于双目立体视觉的三维重构.合肥:中国科学技术大学,2008.
- 6 尚游,陈岩涛.OpenGL 图形程序设计指南.北京:中国水利水电出版社,2001.10.
- 7 乐琴兰.对羽毛球赛场引入电子“鹰眼”的分析与思考.军事体育进修学院报,2011,30(4):123-125.
- 8 李志中,林波萍,张智.鹰眼系统对足球运动发展的影响.广州体育学院学报,2011,31(1):82-84.