

文章编号:1002-2082(2015)02-0177-06

一种目标空间运动分析系统

屠礼芬¹, 彭 祺¹, 仲思东²

(1. 湖北工程学院 物理与电子信息工程学院, 湖北 孝感 432000; 2. 武汉大学 电子信息学院, 湖北 武汉 430072)

摘 要:针对运动目标空间定位问题, 将立体视觉技术与单视频序列下的运动目标检测相结合, 设计了一种目标空间运动分析系统, 实现目标的空间定位, 扩展了单视频序列运动目标检测课题的应用范围。针对从单目向多目过渡过程中面临的问题, 从硬件上提出了相应的解决方案, 将立体镜头与高速相机相结合, 克服了芯片一致性、相机同步、运动拖影等影响。由于立体像对源于同一帧图像, 所以可以直接使用传统单序列运动目标检测方法获取运动掩模。使用该空间运动分析系统在光照充足的室外条件下进行单一目标的自由落体和斜坡运动 2 组实验, 实验结果表明:该系统运行稳定, 实现了运动目标的空间位置分析功能, 1 m 范围内, 定位绝对误差小于 0.01 m。

关键词:图像处理; 空间运动分析; 高速相机; 立体镜头

中图分类号: TN911.73; TP391

文献标志码: A

doi: 10.5768/JAO201536.0201003

Stereo motion analyzing system for object in space

Tu Lifan¹, Peng Qi¹, Zhong Sidong²

(1. School of Physics and Electronic Information Engineering, Hubei Engineering University, Xiaogan 432000, China; 2. School of Electronic Information, Wuhan University, Wuhan 430072, China)

Abstract: According to the problem of spatial location for the moving object, the stereo vision technique and moving object detection in a single video sequence were combined. A stereo motion analyzing system was designed to achieve the goal of spatial location. It can expend the application range for the moving object detection issue in a single video sequence. According to the problems in the course of transition from monocular to multi-view, a solution was proposed based on the hardware. Combined the stereo lens with the high-speed camera, the problems of chip consistency, camera synchronization and the motion ghost can be overcome. Because the stereo images come from the one image, the traditional moving object detection method in a single video sequence can be used directly to obtain motion mask. Two groups of real experiments were taken by the system in full bright daylight conditions, containing free fall and slope movement. The results show that the system works stably and realizes the stereo motion analyzing capabilities for object. The absolute location error is less than 0.01 m at a range of 1 m.

Key words: image processing; stereo motion analyzing; speed camera; stereo lens

引言

立体视觉^[1]是获取空间三维信息的关键,其

核心技术在于 2 个方面:1) 硬件上,需要对通过刚体连接的双相机进行标定^[2],获取各相机的内外

收稿日期:2014-08-18; 修回日期:2014-11-20

基金项目:国家“973 计划”资助项目(2012CB725301); 湖北工程学院科学研究项目(201512)

作者简介:屠礼芬(1986-),女,湖北孝感人,博士,讲师,主要从事图像测量与机器视觉研究工作。

E-mail: tulifen_0301@163.com

参数;2) 软件上,需要将真实目标在不同相机中对应起来,这个过程称为匹配^[3]。国内外学者已经对立体视觉相关技术进行了大量的研究,并且已在核心技术上取得许多成果,但研究对象主要是针对静态立体图像对。本文研究的对象是运动目标,即引入立体视觉相关理论,获取“运动”状态下目标的三维信息。在标定与匹配问题上,我们可以对当前方法进行相应改进。但要从静态过渡到动态,实现空间运动分析,还有2个必须要面对的难题:1) 属性一致的图像采集系统。双目立体视觉系统中的2个相机,要求参数、性能一致,但2个不同的个体,总会有微小的差别,这些差别在运动目标检测应用中,会导致运动掩模获取的误差,对后期定位的影响较大;2) 相机同步问题。由于场景中目标是运动的,不同时刻其位置是不一样的,尤其对于高速运动目标,差异会很大,这就要求立体视觉系统中的2个相机要“精确”同步。虽然目前关于相机同步^[4]的研究方法较多,精度较高,但依然存在误差,在空间运动分析系统中,这些很小的误差也是需要处理的。

针对以上问题,本文建立了一种空间运动分析系统,并针对其特点对当前的相机标定方法进行改进,用来获取其参数。该运动分析系统的核心部件由高速相机和3D镜头组成,可以从硬件上直接克服目标空间运动分析所面临的两个难题。更重要的是,由于只有单一的视频序列,其运动掩模的获取可以采用单视频序列下的运动目标检测算法^[5-6]。另外,受限于该系统的视场和基距,传统的3D标定场^[7]不再适用,所以本文还设计了一种改进的标定方法,可以适用于小视场、短基距立体视觉系统。

1 主要硬件设备

图1所示为空间运动分析系统核心硬件部分,由一个单一的高速相机和一个合二为一的立体镜头组成。

高速相机是 Optronis^[8]提供的,型号为 CL600X2/M,获取的图像是8位灰度图。它的图像传感器有较好的性能,其特点是具有高分辨率、高图像采集速率,全分辨率下,其有效像素在水平方向为1280像素,垂直方向为1024像素,最高图像采集帧率为500 f/s。其全局式快门对于捕捉运动目标有较大的优势,在目标高速运动条件下,模

糊、拖影等现象对于结果影响较小。

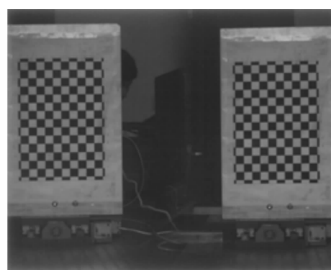


图1 空间运动分析系统核心硬件

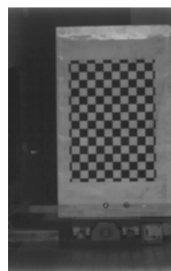
Fig. 1 Hardware core of stereo motion analyzing system

3D镜头是 Loreo^[9]提供的,型号为 Cap 9005,它获取的是一种平行模式的立体图像对。与普通镜头不一样的设计在于,3D镜头的前端由2个一样的镜头组成,2个镜头分别成像,并在后端进行拼接,成像在同一个CCD部件的左右2部分。由于2个镜头有视差,具有与传统立体相机一致的获取目标空间信息的功能,因为使用的是同一个CCD芯片,所以镜头属性的一致性较强,并且图像序列的采集是精确同步的。

图2(a)是用该空间运动分析系统获取的原始图像,图2(b)、2(c)为对原图左右对称分割后获取的左右像对。由于相机获取的原始图像需要平均分为左右2部分生成立体像对,图像分辨率下降了一半,为了保证足够的视场范围,相机模式选择为分辨率优先,原始图像分辨率为1280像素×1024像素,被分割后的立体像对分辨率为640像素×1024像素。



(a) 空间运动分析系统采集的原始图像



(b) 分离的左图



(c) 分离的右图

图2 样本图像

Fig. 2 Sample images

2 标定方案

相机标定的基本思路是结合已知空间坐标的三维空间点的世界坐标系坐标及其分别在左右像对中的图像坐标,通过矩阵运算获取各相机对应的内外参数模型,一旦获取了标定参数,立体相机的相对位置是不可以改变的。本系统的标定方法结合了传统标定方法^[10]的高精度和自标定方法^[11]的灵活性、低成本,标定所使用的主要硬件如图3所示,由平面标定板、光栅尺组成,光栅尺有与其配套的数显设备,可以直接获取标定板在光栅尺上移动的距离。

图3中黑白棋盘格标定板的尺寸是已知的,假设左下角角点坐标为(0,0),横轴为 X 坐标,向右为正方向,纵轴为 Y 坐标,向上为正方向,那么可以依次获得所有角点的坐标(X,Y)。棋盘格被垂直地固定于光栅尺上,光栅尺固定于光学台上,其

直线移动会使标定板随之前后移动,该“直线”即为 Z 坐标。按照右手坐标系原则,向前为正方向,光栅尺的移动距离即为深度 Z 值的变化,可以从数显表上精确读取。标定时,固定空间运动分析系统,将标定板移动到光栅尺上离系统最远的位置,设置数显设备为0,采集此时的图像,然后每移动一定的距离采集一帧图像,同时记录数显设备的值,记录此时标定板上所有点的 Z 标值。结合不同位置图像中各点在标定板上的平面坐标,可以生成各点的世界坐标系坐标(X,Y,Z),与传统标定方法中的三维控制场获取的点一样具有较高的精度,但硬件系统简单、设计灵活方便。最后可以使用传统标定方法,结合这些已知世界坐标系的坐标点分别在左右图像中的图像坐标,通过矩阵运算分别获取3D镜头和高速相机组成的单镜头立体视觉系统左右部分的内外参数。

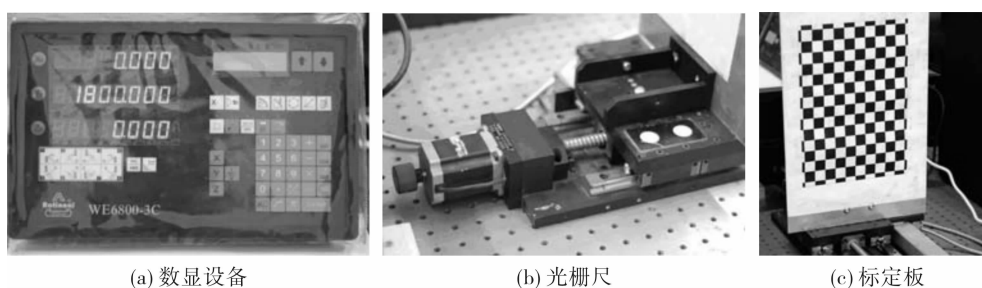


图3 标定相关硬件

Fig. 3 Calibration components

本文使用的相机模型中,每个相机有4个内参数: r_0 和 c_0 分别表示光心与图像中心的偏移,单位为像素; f_u 和 f_v 分别为行方向和列方向坐标变换的中间变量,单位也为像素。所指的外参数指旋转量 \mathbf{R} 和平移量 \mathbf{T} ,通常是依靠6个外部参数来计算,分别为:平移元素 t_1 、 t_2 和 t_3 ,单位为m;角方位

元素 φ 、 $\bar{\omega}$ 和 κ ,单位为rad。关于这些参数更详细的描述及计算方法可以见本课题组较早的研究文献^[7]。本文在标定时,将标定板置于4个不同的位置,即在4种 Z 坐标下分别获取标定板的图像,一共生成了660个可用匹配点对,通过矩阵运算,标定的左右相机内外参数如表1所示。

表1 空间运动分析系统参数

Table 1 Stereo camera parameters

参数	L(左相机)	R(右相机)
r_0/pixel	340.000 164 225 675 12	340.703 156 167 154 42
c_0/pixel	546.810 801 780 528 44	546.468 264 553 031 99
f_u/pixel	-3100.889 681 606 792 7	-3 130.020 840 982 391 3
f_v/pixel	3119.280 606 212 459 6	3 131.823 733 363 881 7
t_1/m	0.107 622 665 675 039	0.006 290 730 861 631
t_2/m	0.111 598 226 222 524	0.115 215 722 875 107
t_3/m	2.370 911 381 154 92	2.369 281 725 595 442
φ/rad	-0.072 901 439 053 671	0.089 053 572 023 792
$\bar{\omega}/\text{rad}$	-0.019 510 546 940 336	-0.026 172 121 512 121
κ/rad	-0.014 161 534 378 966	-0.009 792 244 279 391

3 实验结果分析

为了验证本文建立的空间运动分析系统的有效性,选用了小球作为感兴趣的运动目标来进行测试,由于小球具有空间对称性,所以只需对其重心进行定位即可获取空间位置信息。由于目标比较单一,相机的视场小,场景并不复杂,所以使用了传统的 GMM(Gaussian mixture model)算法^[12]来快速获取运动掩模。然后将检测结果分割成左右图像对,分别获取左右图像中掩模部分的重心位置并记录其图像坐标。

两组实验都是轨迹可预测、可重复的,分别为自由落体运动和斜坡运动。二者的初速度均为 0,且运动轨迹为直线,加速度为固定值,用 a 表示,单位为 m/s^2 ;用 h 表示距离,单位为 m ; t 表示时间,单位为 s 。球只有全部进入视场才开始定位,此时的时刻用 t_i 表示,那么不同时刻球的位置关系如图 4 所示。

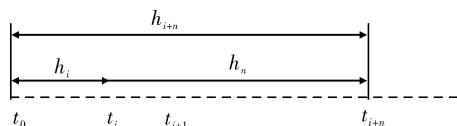


图 4 不同时刻球的位置关系

Fig. 4 Relations between different positions

图 4 中, t_0 为小球开始运动的时刻, t_i 、 t_{i+1} 、 t_{i+n} 分别表示后续不同的时刻,与 t_0 的时间间隔都是最小时间单位 Δt 的整数倍。其中 Δt 为帧间隔时间,此处为 $1/500 \text{ s}$; h_i 、 h_n 、 h_{i+n} 分别为 t_i 与 t_0 、 t_{i+1} 与 t_i 、 t_{i+n} 与 t_0 时间间隔内球的运动距离,它们的关系可以用(1)式~(4)式表示:

$$h_i = \frac{1}{2} a t_i^2 \quad (1)$$

$$h_{i+n} = \frac{1}{2} a t_{i+n}^2 \quad (2)$$

$$t_{i+n} = t_i + n \cdot \Delta t \quad (3)$$

$$h_n = h_{i+n} - h_i \quad (4)$$

依次获取球重心在每个时刻的空间坐标系坐标值,则可以描述其空间运动属性,假设 t_i 和 t_{i+n} 时刻球的空间坐标值分别为 (X_i, Y_i, Z_i) 和 $(X_{i+n}, Y_{i+n}, Z_{i+n})$,那么其距离 h_n 可以表示为

$$h_n = \sqrt{(X_{i+n} - X_i)^2 + (Y_{i+n} - Y_i)^2 + (Z_{i+n} - Z_i)^2} \quad (5)$$

类似地,对于有其他运动属性的目标,也可以

根据这些参数设计不同的数学模型来估计目标速度、加速度等参数。

3.1 自由落体运动

自由落体运动是一个初速度为 0 的匀加速运动,运动轨迹为直线,由于三维空间坐标是一个相对值,所以单一的空间坐标值信息无实际意义,需要分析目标一系列连续的坐标值,才能达到分析目标空间运动属性的目的。图 5 为用该空间运动分析系统对自由落体空间轨迹的记录结果。由于相机的位置不会与标定时一致,所以该自由落体轨迹不是一条仅沿 Y 轴下降的直线,在沿 X 和 Z 轴方向也有变化,并且由于重心提取存在误差,该轨迹还有一些轻微的抖动,尤其是在开始和结束的位置,由于接近图像的边缘,因此误差较大。但该轨迹的线性度依然是比较高的,可以通过线性拟合获取其运动属性。

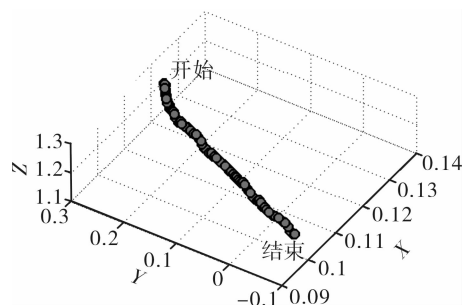


图 5 球的空间运动轨迹

Fig. 5 Ball's space trajectory

球完全进入视场的时刻是 t_i ,那么结合(1)式~(5)式及球在后续帧中的空间位置,就可以计算其值:

$$t_i = \frac{h_n}{a \cdot n \cdot \Delta t} - \frac{n \cdot \Delta t}{2} \quad (n=1, 2, \dots) \quad (6)$$

式中:由于是自由落体运动, $a = 9.8 \text{ m/s}^2$,不同的 n 表示与 t_i 有不同的时间间隔,但是所得 t_i 的值理论上是不变的,由于误差的存在,不同帧计算的 t_i 有一些差异。该实例选取了连续 41 帧来计算,那么可以获取 40 个不同的 t_i 值,结果如表 2 所示。

分析表 2 中的数据可知, t_i 的平均值为 $\bar{t}_i = 0.283 057$,标准差为 $\sigma_{t_i} = 0.007 137$,表 2 中所有值残余误差都在 3σ 以内,不含粗大误差,故该系统可以实现对目标的空间定位。

表 2 40 个 t_i 的取值Table 2 40 values for t_i by consecutive frames

连续 41 帧计算的 t_i			
0.266 424	0.276 321	0.279 741	0.286 343
0.285 446	0.276 682	0.280 507	0.287 347
0.274 611	0.276 466	0.281 195	0.287 676
0.270 623	0.276 992	0.281 814	0.288 684
0.286 008	0.277 486	0.282 784	0.289 752
0.286 847	0.279 610	0.283 293	0.290 637
0.299 074	0.278 241	0.284 180	0.291 558
0.288 931	0.277 989	0.284 146	0.293 060
0.277 009	0.279 491	0.284 969	0.295 042
0.272 845	0.279 217	0.286 024	0.297 198

表 3 距离测量结果

m

Table 3 Distances measurement results

测量值	距离差	真实值	绝对误差
0.077 998	0.080 09	0.08	0.000 09
0.158 088	0.076 513	0.08	-0.003 487
0.234 601	0.084 857	0.08	0.004 857
0.319 458	0.076 701	0.08	-0.003 299
0.396 159	0.081 578	0.08	0.001 578
0.477 737	0.081 277	0.08	0.001 277
0.559 014	0.087 794	0.08	0.007 794
0.646 808	0.077 363	0.08	-0.002 637
0.724 171	0.078 451	0.08	-0.001 549
0.802 622	/	/	/

3.2 斜坡运动

斜坡运动是依靠一个带凹槽的导轨实现的。小球平行于导轨滚动,导轨不同位置之间距离是可以较精确测量的。假设开始定位的时刻为 i , 间隔固定的时间 n 后,再次获取其位置,两位置的关系可以参照(1)式~(4)式,那么依次获取时间间隔为 n 的小球空间位置,记为 $s_i, s_{i+n}, \dots, s_{i+mn}$, 那么位置 s_{i+mn} 与位置 $s_{i+(m+1)n}$ 之间的距离可表示为

$$s_{i+(m+1)n} - s_{i+mn} = \frac{1}{2}a(n^2 + 2mn^2 + 2in) \quad (10)$$

导轨位置一旦固定,小球加速度 a 、初始时刻 i 为固定值,若固定时间间隔 n ,那么该距离与 m 呈线性关系,相邻 2 个距离差为常数。

依次获得球在世界坐标系中的位置,然后就可以根据(5)式获得小球在不同帧之间的距离。本实验采集了 11 个位置的图像,获取了 10 组距离值,检测结果如表 3 所示。表 3 中第 1 列为测量值,第 2 列为相邻 2 个距离差,第 3 列为手动测量导轨全长取平均后所获取的值,用来与本系统的测量结果对比,小球在所测量范围内的导轨总长为 $L=0.72$ m,由于距离做了 9 等分,则对应真实值为 $L/9=0.08$ m。

分析表 3 中的结果可以得出结论:绝对误差低于 0.01 m,并且系统无累积误差,相对来说,一个较大的测量值可以降低系统的相对误差。

4 结论

本文建立的一种空间运动分析系统引入立体视觉的相关理论思路,扩展了当前单相机视频序列运动目标检测课题的应用范围,可以用来获取感兴趣运动目标的空间位置并分析其运动属性,也使与其相关的跟踪、定位等领域更加符合人的认知习惯。针对从单视频到多视频扩展所遇到的相机属性一致性、精确同步问题,本文从硬件上提出了比较好的解决方案:通过前端的 3D 镜头将视场分割,在同一 CCD 芯片上成像,生成立体图像对,有效地解决了这 2 个问题。另外,由于立体像对序列实际上是从同一个图像序列分割而来的,所以可以使用传统单序列下的运动目标检测算法来获取运动掩模,再分割掩模图像生成有视差的立体图像对序列,易于实现。实验结果表明,该硬件系统可以有效地分析目标的空间运动属性,其定位较为准确,且数据的稳定性较好。

参考文献:

- [1] Hu X, Mordohai P. A quantitative evaluation of confidence measures for stereo vision[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2121-2133.
- [2] Zhang Zhengyou. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [3] Cigla C, Alatan A A. Information permeability for

- stereo matching[J]. *Signal Processing: Image Communication*, 2013, 28(9): 1072-1088.
- [4] Shrestha P, Barbieri M, Weda H, et al. Synchronization of multiple camera videos using audio-visual features [J]. *IEEE Transactions on Multimedia*, 2010, 12(1): 79-92.
- [5] Tu Lifan, Zhong Sidong, Peng Qi. Moving object detection based on image sequences[J]. *Journal of Applied Optics*, 2012, 33(5): 899-903.
屠礼芬, 仲思东, 彭祺. 基于序列图像的运动目标检测[J]. *应用光学*, 2012, 33(5): 899-903.
- [6] Tu Lifan, Zhong Sidong, Peng Qi. A method for detecting moving objects and eliminating shadow in natural scene[J]. *Journal of Xi'an Jiaotong University*, 2013, 47(12): 26-31.
屠礼芬, 仲思东, 彭祺. 自然场景下运动目标检测与阴影剔除方法[J]. *西安交通大学学报*, 2013, 47(12): 26-31.
- [7] Liu Yong. 3D scene acquiring and modeling based on images[D]. Wuhan: Wuhan University, 2004.
- 刘勇. 基于图像的空间三维数据获取及建模[D]. 武汉: 武汉大学, 2004.
- [8] Optronis GmbH. CamRecord CL 600 × 2 [S/OL]. [2013-4-10]: http://www.optronis.de/fileadmin/Upload/Product/CL/CL600x2_engl.pdf.
- [9] Loreo Asia Ltd. 3D Lens in a Cap 9005 [S/OL]. [2013-4-10]: http://www.loreo.com/pages/products/loreo_3dlenscap9005.html.
- [10] Sirianni M, Jee M J, Benitez N, et al. The photometric performance and calibration of the hubble space telescope advanced camera for surveys [J]. *Publications of the Astronomical Society of the Pacific*, 2005, 117: 1049-1112.
- [11] Ma Songde. A self-calibration technique for active vision systems[J]. *IEEE Transactions on Robotics and Automation*, 1996, 12(1): 114-120.
- [12] Kaewtrakulpong P, Bowden R. An improved adaptive background mixture model for real-time tracking with shadow detection[C]. USA: Kluwer Academic Publishers, 2001.