

## 单目平行线约束下的空间点坐标恢复

张元元<sup>1</sup>, 吴晓娟<sup>1</sup>, 李秀媛<sup>1</sup>, 阮秋琦<sup>2</sup>

(1. 山东大学 信息科学与工程学院, 山东 济南 250100; 2. 北京交通大学 信息科学研究所, 北京 100044)

**摘要:**提出一种从单目相机拍摄的二维投影图像中恢复空间点三维坐标的简化算法. 假定摄像机与空间中的一对平行线的位置关系已知, 同时摄像机坐标系下平行线所在平面中的任意空间点在像平面中的投影也已知, 通过射影几何可得图像坐标到目标点三维空间坐标的转换系数, 即坐标转换因数, 从而估计出三维空间点的坐标. 实验中通过对人体行走轨迹进行估计来测试算法性能, 结果表明算法简单有效, 计算开销小.

**关键词:**三维重建; 单目摄像机; 平行线约束; 坐标转换因数; 坐标恢复

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 1673-4785(2010)01-0024-06

## Extraction of 3-D coordinates using parallel lines in monocular images

ZHANG Yuan-yuan<sup>1</sup>, WU Xiao-juan<sup>1</sup>, LI Xiu-yuan<sup>1</sup>, RUAN Qiu-qi<sup>2</sup>

(1. School of Information Science and Technology, Shandong University, Ji'nan 250100, China; 2. Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China)

**Abstract:** A method was proposed to simplify the extraction of 3-D coordinates from spatial points in 2-D images captured by a monocular camera. Assume that two parallel lines could be known in the image, and also that their relative positions with respect to the video camera were known. In the camera coordinate system, the projection of any spatial point in the plane determined by parallel lines to the image plane of the camera could be measured. By applying the knowledge of projection geometry, a coordinate conversion factor (CCF) for the given point was then calculated; these were used to estimate 3-D spatial coordinates. Estimations of the paths of people walking were then done to analyze the effectiveness of this method. Results showed that the proposed algorithm is simple and effective, and has low computational costs.

**Keywords:** 3-D reconstruction; monocular camera; parallel lines constraint; coordinate conversion factor; coordinates recovery

视觉传感器如摄像机拍摄到的图像是三维空间的景物在二维空间中的投影, 通过投影点估算对应空间点的三维坐标称为空间点的三维重建. 它是计算机视觉<sup>[1]</sup>中的一个基本问题, 在相机标定、三维重建等很多方面有广泛的应用价值<sup>[2]</sup>. 然而正是因为视觉传感器只能获取三维世界的投影信息, 这个过程导致大量信息丢失, 使三维重建任务变得更加复杂, 至今仍是一个亟待解决的问题.

利用2幅或多幅图像恢复空间点的三维几何信息, 具体的做法有2类. 一类方法使用双摄像机或多摄像机拍摄空间点的同一点, 需要对摄像机进行定

标, 确定空间点在拍摄图像中的对应点位置以及摄像机的投影矩阵, 通过解析几何的方法进行求解. 卡内基梅隆大学的Kanade<sup>[3-5]</sup>和其研究小组在从侧影中恢复形状的研究中取得了较大的成果. Mikic<sup>[6]</sup>给出一种多相机获取人体运动的系统, 首先通过Shape From Silhouette得到人体的voxel数据, 然后通过扩展Kalman filter进行姿态估计. 另一类方法使用单摄像机, 通过移动摄像机对目标点进行多次拍摄, 求解其三维坐标. Sminchisescu等<sup>[7]</sup>采用单目序列图像进行人体运动估计, 人体模型采用超二次曲面加变形参数, 模型有30个关节, 选择边缘和灰度特征设计似然函数. 还有文献<sup>[8-9]</sup>提出了利用单目摄像机拍摄的图像序列对目标点的三维运动轨迹进行估计的算法. 这2类方法在本质上是一

收稿日期: 2008-12-24.

基金项目: 国家自然科学基金资助项目(60675024).

通信作者: 吴晓娟. E-mail: xiaojwu@sdu.edu.cn.

样的,需要通过2幅或多幅图像计算目标点的三维空间坐标,重建的过程通常要求对摄像机进行定标,并寻找目标点在不同图像上的对应点,即对应元匹配问题.在实际求解过程中,方法的性能受到2个因素的影响:定标精度和对应元的匹配精度.通常,较高的精度需要较复杂的定标和匹配过程.

一般而言,在没有任何约束条件的情况下,从单目摄像机拍摄的图像恢复空间点的三维位置信息是极其困难的.现有的三维重建算法均需满足一定的约束条件,约束条件随实际应用场所的不同而有所变化.例如,Hild<sup>[10]</sup>利用身高对人体的运动轨迹进行估计,通过人体实际身高以及图像中的人体像高确定转换系数,求解目标步行轨迹.为了准确地估计转换系数,该算法要求摄像机的光轴方向垂直于人体站立方向.然而,实际场景中摄像机的拍摄视角往往不能满足这一要求.本文提出了一种基于平行线约束的三维空间点的坐标恢复算法,通过空间中的一组平行线,确定二维图像坐标到三维空间坐标的转换关系.理论分析和实验结果表明,该算法在大多数视角下均能较准确地恢复目标点的三维空间坐标.

## 1 问题描述

在以 $O_c$ 为原点的三维坐标系 $XYZ$ 中,存在一对平行线 $l_1$ 和 $l_2$ ,它们所在的平面为 $S$ ,一条垂线 $A_1A_2$ 同时垂直于这2条直线.假定焦距为 $f$ 的摄像机 $C$ 位于原点 $O_c$ 处, $l_1$ 和 $l_2$ 在摄像机成像平面 $I$ 上的投影分别为 $l'_1$ 和 $l'_2$ ,如图1所示.为简化问题,本文将摄像机成像模型简化为小孔成像原理.

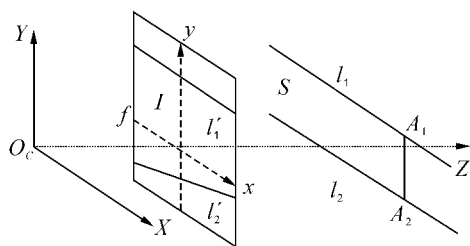


图1 空间平行线在像平面上的投影

Fig. 1 Projections of parallel lines in image plane  $I$

下面研究如何利用 $l_1$ 和 $l_2$ 求解平面 $S$ 中的一点 $P$ 在三维空间中的坐标.根据 $l_1$ 和 $l_2$ 及 $l'_1$ 和 $l'_2$ 这4条直线的位置关系可将问题分解为以下4种情况分别予以讨论,如图2所示.

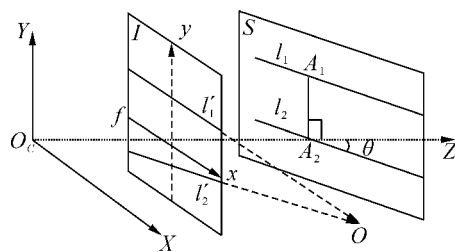
1)  $l_1$ 和 $l_2$ 与摄像机光轴 $O_cZ$ 的夹角为 $\theta$ ,且 $0^\circ < \theta < 90^\circ$ , $l'_1$ 与 $l'_2$ 相交于 $O$ ,垂线 $A_1A_2$ 同光轴 $O_cZ$ 垂直.

2)  $l_1$ 和 $l_2$ 与摄像机光轴 $O_cZ$ 的夹角为 $\theta$ ,且 $0^\circ < \theta < 90^\circ$ , $l'_1$ 与 $l'_2$ 相交于 $O$ ,垂线 $A_1A_2$ 与光轴

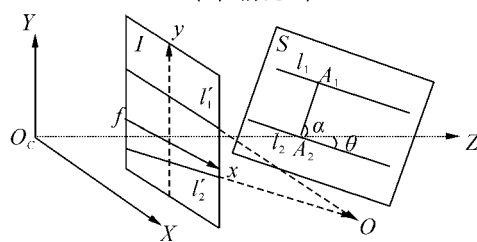
$O_cZ$ 存在一定的夹角 $\alpha$ ,且 $0^\circ < \alpha < 90^\circ$ .

3)  $l_1$ 和 $l_2$ 与像平面平行,且垂线 $A_1A_2$ 同光轴 $O_cZ$ 垂直,此时平面 $S$ 与像平面平行.

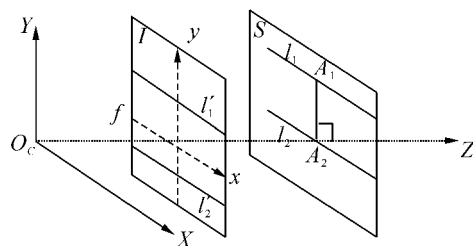
4)  $l_1$ 和 $l_2$ 在像平面上的投影 $l'_1$ 与 $l'_2$ 重合为一条直线.



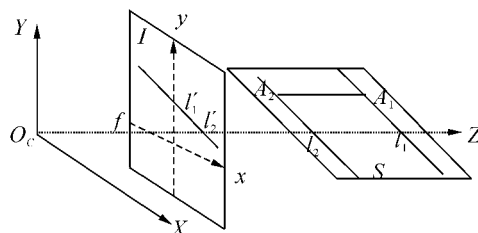
(a) 情况 1)



(b) 情况 2)



(c) 情况 3)



(d) 情况 4)

图2  $l_1$ 和 $l_2$ 在像平面中投影的4种情况

Fig. 2 Four cases of projection of parallel lines in image plane  $I$

## 2 特定视角下的坐标恢复

本节讨论特定摄像机拍摄视角即第1种情况下的三维坐标恢复问题.设平行线 $l_1$ 和 $l_2$ 之间的距离为 $H$ ,摄像机的焦距为 $f$ ,平面 $S$ 到摄像机的距离为 $Z_A$ . $P$ 为平面 $S$ 上介于 $l_1$ 和 $l_2$ 之间的任意一点,过 $P$ 作垂线 $P_1P_2$ 垂直于 $l_1$ 和 $l_2$ ,垂足为 $P_1$ 和 $P_2$ ,如图3所示.设 $l_1$ 和 $l_2$ 与光轴的夹角为 $\theta$ , $0^\circ < \theta < 90^\circ$ ,则 $l_1$ 和 $l_2$ 在像平面中的投影 $l'_1$ 和 $l'_2$ 相交于一点 $O$ .

$$\begin{cases} P_1 P = H \frac{y_1 (f \sin \alpha - y_3 \cos \alpha)}{(y_1 + y_2 + y_3) (f \sin \alpha + y_2 \cos \alpha)}, \\ PP_2 = H \frac{(y_2 + y_3) (f \sin \alpha + (y_1 + y_2) \cos \alpha)}{(y_1 + y_2 + y_3) (f \sin \alpha + y_2 \cos \alpha)}. \end{cases} \quad (5)$$

2) 计算  $P'_1P_2$  和  $P_1P'_2$  的长度. 类似于第一步, 首先计算正弦函数  $\sin \angle O_cP'_1P'_2$ 、 $\sin \angle PP'_1P'_1$  和  $\sin \angle PP_2P'_2$  的值, 分别用  $y_1$ 、 $y_2$ 、 $y_3$  和  $f$  来表示.

同样, 根据正弦定理, 可得到式(6):

$$\frac{P_1P}{\sin \angle PP'_1P'_1} = \frac{P'_1P}{\sin \angle PP_1P'_1},$$

$$\frac{PP_2}{\sin \angle PP'_2P'_2} = \frac{PP'_2}{\sin \angle PP_2P'_2}. \quad (6)$$

进一步可以得到

$$\begin{aligned} PP'_1 &= y_1 A^*, \\ PP'_2 &= (y_2 + y_3) A^*, \\ P'_1P'_2 &= PP'_1 + PP'_2 = \\ &= (y_1 + y_2 + y_3) A^*. \end{aligned} \quad (7)$$

式中:

$$A^* = \frac{H(f \sin \alpha - y_3 \cos \alpha)(f \sin \alpha + (y_1 + y_2) \cos \alpha)}{f(y_1 + y_2 + y_3)(f \sin \alpha + y_2 \cos \alpha)}.$$

3) 计算  $P$  点的三维坐标. 首先求出  $P$  点的坐标转换因数  $\eta_P$ :

$$\eta_P = \frac{P'_1P'_2}{P_1P_2} = \frac{(y_1 + y_2 + y_3) A^*}{y_1 + y_2 + y_3} = A^*. \quad (8)$$

最后, 求得  $P$  点的三维坐标为

$$[X_P \ Y_P \ Z_P]^T = \eta_P [x_P \ y_P \ f]^T \quad (9)$$

情况3)是情况1)的特例, 仍可通过式(3)来求解  $P$  点坐标. 情况4)时, 由于平行线的投影重合, 无法使用平行线约束来恢复  $P$  点的三维坐标, 此处不予讨论.

理论分析表明, 在大多数情况下, 在已知摄像机焦距的情况下, 通过测量图像中的某位置的像素坐标可估算其相应的三维空间中的坐标. 当然这一坐标是以摄像机位置为原点来建立的空间坐标系.

## 4 实验结果及分析

为了测试本文算法的有效性, 本文设计了一组实验进行验证. 为简单起见, 这里只考虑第一种情况, 即摄像机光轴与人体直立方向垂直的情况. 设置这样的场景: 将摄像机固定于一处较平坦的地面上, 光轴  $O_cZ$  与地面平行, 在地面上画网格建立相应的平面坐标系, 设地面与  $XO_cZ$  平面平行, 竖直方向为  $Y$  方向. 人在摄像机视野范围内沿直线或分段直线行走. 在左脚鞋底涂上特殊的物质, 这样左脚在着地时便在预先画好的网格内留下脚印. 记录这些脚印的位置坐标, 即左脚着地时刻的三维坐标中的  $X$ 、 $Z$

分量. 利用本文的算法来估计相应时刻左脚的空间坐标分量  $X'$  和  $Z'$ , 通过与真实值的对比来测试算法的性能.

### 4.1 摄像机焦距测定

像平面中各点的坐标是像素意义上的坐标, 并非常见的米制下的坐标值, 而焦距  $f$  通常是以米制的形式给出的. 因此, 需要通过实验来确定以像素形式表示的焦距值. 具体方法如下:

在场景中设置标志物  $P_i$  和  $P_j$ , 它们到像平面的距离相同均为  $M$ ,  $P_i$  和  $P_j$  之间的距离为  $D_{ij}$ ,  $p_i$  和  $p_j$  是  $P_i$  和  $P_j$  在像平面的投影, 两者的距离为  $d_{ij}$ , 如图6所示.

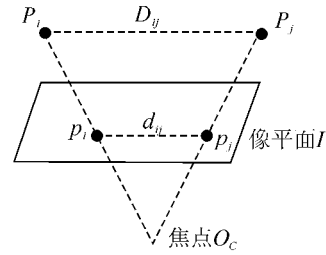


图6 估计摄像机焦距

Fig. 6 Estimate the focal length of camera

由几何关系易知:

$$P_iP_j \parallel p_ip_j.$$

由相似性原理可得

$$D_{ij}/d_{ij} = (M + f)/f.$$

整理可得

$$f = (M + f) d_{ij}/D_{ij}. \quad (10)$$

通常情况下,  $M \gg f$ , 故式(10)可以近似为

$$f \approx M \cdot d_{ij}/D_{ij}. \quad (11)$$

式中:  $M$  和  $D_{ij}$  的单位是 m,  $d_{ij}$  的单位是像素. 因此  $f$  的单位为像素. 设置多组标志物, 记第  $k$  组测得的结果为  $f_k$ , 将其均值作为摄像机的焦距  $f$ . 本实验中测得摄像机焦距  $f$  为 430 像素.

### 4.2 行走轨迹估计

人在行走过程中, 2 个相邻的相同姿态可以确定一对头部和脚部的连线作为空间中的平行线, 目标的真实高度  $H$  以及各时刻图像中人体的高度  $h$  都容易得到, 从而计算相应时刻的  $\eta$ , 最后利用本文算法对人体各时刻关键点的坐标位置进行估计. 通过检测可以得到左脚着地时刻的图像帧中脚部的像素坐标, 将其与预先记录的真实坐标相比来计算估计误差. 表1中给出了某行人行走过程中左脚着地时刻的估计值、真实值以及相应的误差. 计算各个点的误差的平均可得平均误差为 0.210 m.

表1 实验数据  
Table 1 Experimental data

图像中的脚部 位置/像素		真实值/m		估计值/m		误差/m
$y$	$x$	$Z$	$X$	$Z'$	$X'$	$((Z - Z')^2 + (X - X')^2)^{1/2}$
226	95	5.00	-1.00	4.85	-0.91	0.175
248	125	4.30	-0.70	4.09	-0.47	0.331
260	168	3.80	-0.10	3.78	-0.06	0.045
264	197	3.60	0.20	3.64	0.18	0.045
246	234	4.20	0.50	4.14	0.49	0.061
235	280	4.80	1.20	4.53	1.09	0.292
224	309	5.30	1.60	5.14	1.59	0.163
212	339	5.90	2.20	5.78	2.19	0.220
211	300	6.10	1.70	6.05	1.76	0.078
206	263	6.90	1.30	6.64	1.34	0.064
203	224	7.60	0.70	7.64	0.85	0.155
192	197	8.20	0.10	8.05	0.39	0.327
203	127	6.90	-1.00	6.52	-0.78	0.779

将脚部位置的三维坐标分量  $X$  和  $Z$  的真实值与估计值绘制并用线段连接起来,便得到目标行走的真实轨迹与估计轨迹,如图7所示.从图中也可以看出,本文的算法可以得到近似精确的结果.

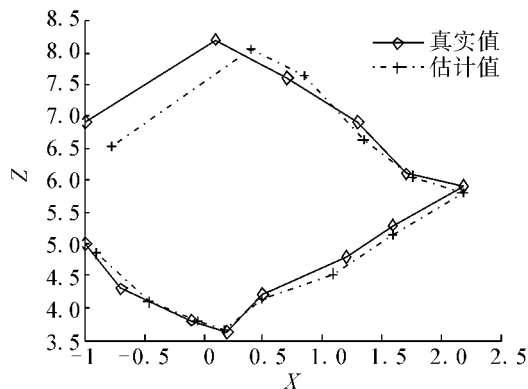


图7 真实轨迹与估计轨迹

Fig. 7 Real and estimated trajectories

统计目标离摄像机不同距离时的平均估计误差,可以获得误差同距离的关系,如表2所示.这里,目标离摄像机的距离定义为  $Z$  方向上的坐标差.从表中可以看出,估计误差随着目标同摄像机距离的增加而增大.由此,可以得出结论:在视频采集阶段,摄像机距离目标越近,则使用平行线约束的方法对目标进行估计的效果就越好.

表2 不同距离下的平均误差

Table 2 The average errors in difference distances m

距离	平均误差
3~4	0.045
4~5	0.107
5~6	0.191
>6	0.281

## 5 结束语

本文针对计算机视觉中的基本问题——空间点的三维坐标恢复进行了研究,提出一种单目摄像机单幅图像下空间点的三维坐标估计算法,该算法通过空间中的一对已知的平行线求解相应的坐标转换因数,有效地克服了视角因素的影响,无需处理复杂的定标和对应元匹配问题.本文通过估计行人行走时的脚部位置验证了算法的有效性,实验中所确立的平行线为目标在行走过程中相邻的两个相同姿态下头部和脚部的连线.在实际应用中,可根据实际情况确定所需平行线.此外,通过估计视频中人体各关键点的坐标可进行人体运动分析,进一步可进行基于步态的身份识别等研究.

## 参考文献:

- [1] 马颂德, 张正友. 计算机视觉[M]. 北京: 科学出版社, 1998: 52-98.
- [2] 李立冬, 刘教民. 单目视图下相机标定和平面测距研究[J]. 河北科技大学学报, 2005, 26(1): 47-50.  
LI Lidong, LIU Jiaomin. Research on camera calibration and plane measurement from a single view[J]. Journal of Hebei University of Science and Technology, 2005, 26(1): 47-50.
- [3] CHEUNG K M, BAKER S, KANADE T. Shape from silhouette of articulated objects and its use for human body kinematics estimation and motion capture[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Madison, USA, 2003, 1: 18-20.
- [4] CHEUNG K M, BAKER S, HODGINS J K, et al. Markerless human motion transfer[C]// Proceedings of 2nd International Symposium on 3D Data Processing, Visualization and Transmission. Thessaloniki, Greece, 2004: 373-378.
- [5] CHEUNG K M, BAKER S, KANADE T. Shape from silhouette across time, part I: theory and algorithms[J]. International Journal of Computer Vision, 2005, 62(3): 221-247.
- [6] MIKIC I, TRIVEDI M, HUNTER E, et al. Human body model acquisition and tracking using voxel data[J]. International Journal of Computer Vision, 2003, 33(3): 199-223.
- [7] SMINCHISESCU C, TELEA A. Human pose estimation from silhouettes: a consistent approach using distance level sets[C]// International Conference in Central Europe on Computer Graphics, Visualization. Plzen-Bory, Czech Republic, 2002: 413-420.
- [8] AVIDAN S, SHASHUA A. Trajectory triangulation: 3D reconstruction of moving points from a monocular image sequence[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(4): 348-357.
- [9] PONS D, LOPEZ A, LUMBRERAS F, et al. 3D vehicle sensor based on monocular vision[C]// Proceedings of IEEE

Conference on Intelligent Transportation Systems. Vienna, Austria, 2005: 1096-1101.

- [10] HILD M. Estimation of 3D motion trajectory and velocity from monocular image sequences in the context of human gait recognition[C]// Proceedings of the 17th International Conference on Pattern Recognition. Cambridge, UK, 2004: 231-235.

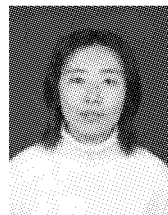
### 作者简介:



张元元,男,1984年生,博士研究生,主要研究方向为计算机视觉、图像处理 and 模式识别等。



吴晓娟,女,1944年生,教授、博士生导师,主要研究方向为智能信息处理、图像处理、模式识别、计算机视觉等。近年来主持国家自然科学基金、军工、国际合作、教育部、省自然科学基金、省科技攻关等项目20余项,发表学术论文100余篇。



李秀媛,女,1972年生,博士研究生,主要研究方向为无线通信、计算机网络、图像处理等。



阮秋琦,男,1944年生,教授,博士生导师。IEEE高级会员,IEEE北京分会技术委员会主席,中国通信学会会士,中国电子学会会士、理事、学术委员会副主任,中国图像图形学学会理事、北京图像图形学学会副理事长、中国图像图形学学会铁道专业委员会主任委员。主要研究方向为图像处理、计算机视觉等。承担和完成国家级和省部级科研项目50余项,发表学术论文350余篇,出版专著和教材8部。