

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 9
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年9月 第17卷 第9期(总第197期)

目次

图像处理和编码

- 自适应阈值 HMRF 模型超分辨率重建 朱虹, 刘薇, 姚杰, 欧阳光振, 刘小乾(1049)
- 亮度补偿变换矩阵的颜色恒常性算法 袁兴生, 王正志(1055)
- 小波变换和稀疏冗余表示的混合图像去噪 李慧斌, 刘峰(1061)
- 基于运动信息自适应的快速运动估计算法 李子印, 杨齐(1069)
- 融合纹理特征的两阶段聚类分割算法 王改华, 李德华(1075)
- 利用纹理和边缘特征的 Criminisi 改进算法 任澍, 唐向宏, 康佳伦(1085)

图像分析和识别

- 激光焊接对接拼缝测量的微景深方法 王平江, 黄雅婷, 齐江飞, 彭芳瑜, 李斌, 唐小琦(1092)
- 基于码本模型和多特征的早期烟雾检测 姜明新, 王洪玉, 蔡兴洋(1102)
- 结合局部特征和全局信息的自适应活动轮廓模型 朱晓舒, 孙权森, 夏德深(1109)
- 高斯颜色模型在瓷片图像分类中的应用 郑霞, 胡浩基, 周明全, 樊亚春(1115)
- 区域生长的轮对图像分割 史倩, 吴开华(1122)
- 图像搜索结果的滑动窗口多标记传播聚类 谷瑞军, 陈圣磊, 陈耿, 汪加才(1128)

图像理解和计算机视觉

- 中介相似性量度的图像匹配 周宁宁, 吕文杰(1135)
- 自适应字典学习的多聚焦图像融合 严春满, 郭宝龙, 易盟(1144)
- 摄像机几何约束及人体定位 陶霖密, 于亚鹏, 邸慧军, 孙洛(1150)

虚拟现实与增强现实

针对层次式累进网格模型的动态评估与调整 王丹华, 卢威, 潘金贵(1158)

医学图像处理

结合形状与纹理特征的椎间盘退化辅助诊断..... 郝世杰, 蒋建国, 郭艳蓉, 詹曙, 李鸿(1168)

改进的核磁共振图像分割与偏移场恢复耦合模型

..... 王顺风, 冀晓娜, 张建伟, 陈允杰, 方林, 詹天明(1175)

遥感图像处理

光学遥感图像低可观测区域舰船检测 周伟, 关键, 何友(1181)

非下采样 Contourlet 变换与脉冲耦合神经网络相结合的 SAR 与多光谱图像融合

..... 金星, 李晖晖, 时丕丽(1188)

“计算机视觉前沿论坛”专栏征文通知 (1196)

中国图象图形学报

刊名题字: 宋 健

月刊(1996 年创刊)

第 17 卷 第 9 期

2012 年 9 月 16 日出版

主管单位 中国科学院

主 办 中国科学院遥感应用研究所

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 李小文

编辑出版 《中国图象图形学报》编辑出版委员会

北京 9718 信箱 邮编 100101

电子信箱:jig@irsa.ac.cn

电话:010-64807995 010-82614429

网 址:www.cjig.cn

印刷装订 北京北林印刷厂

广告经营许可证 京朝工商广字第 0346 号

总 发 行 北京报刊发行局

订 购 全国各地邮局

国外发行 中国国际图书贸易总公司

(中国国际书店)

(北京 399 信箱 邮编 100044)

Superintended by Chinese Academy of Sciences

Sponsored by Institute of Remote Sensing Application,

CAS China Society of Image and Graphics

Institute of Applied Physics and Computational

Mathematics

Chief editor LI Xiaowen

Editor, Publisher Editorial and Publishing Board

of Journal of Image and Graphics

(P. O. Box 9718, Beijing 100101, China)

E-mail:jig@irsa.ac.cn

Distributed by Beijing Bureau for Distribution of Newspapers

and Journals

Domestic All Local Post Offices in China

Foreign China International Book Trading Corporation

(P. O. Box 399, Beijing 100044, China)

Printed by Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTXFZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 9 September 2012

Contents

Image Processing and Coding

- Adaptive threshold HMRF model for super-resolution reconstruction
..... Zhu Hong, Liu Wei, Yao Jie, Ouyang Guangzhen, Liu Xiaoqian(1049)
- Color constancy algorithms with luminance-compensation transform matrix Yuan Xingsheng, Wang Zhengzhi(1055)
- Hybrid image denosing method based on wavelet transform as well as on a sparse and redundant representations model
..... Li Huibin, Liu Feng(1061)
- Fast motion estimation algorithm based on motion information adaptation Li Ziyin, Yang Qi(1069)
- Texture feature fusion-based two-stage clustering segmentation algorithm Wang Gaihua, Li Dehua(1075)
- Improved Criminisi algorithm with the texture and edge features Ren Shu, Tang Xianghong, Kang Jialun(1085)

Image Analysis and Recognition

- Seam measurement based on micro depth of field lens for laser welding
..... Wang Pingjiang, Huang Yating, Qi Jiangfei, Peng Fangyu, Li Bin, Tang Xiaoqi(1092)
- Early smoke detection based on codebook model and multiple features Jiang Mingxin, Wang Hongyu, Cai Xingyang(1102)
- Adaptive active contour model integrating global and local image fitting energy Zhu Xiaoshu, Sun Quansen, Xia Deshen(1109)
- Porcelain shard images classification based on Gaussian color model Zheng Xia, Hu Haoji, Zhou Mingquan, Fan Yachun(1115)
- Image segmentation for wheel set measurement based on region growing Shi Qian, Wu Kaihua(1122)
- Sliding window multi-label propagation clustering for searched images
..... Gu Ruijun, Chen Shenglei, Chen Geng, Wang Jiakai(1128)

Image Understanding and Computer Vision

- Image matching based on the medium similarity measure Zhou Ningning, Lv Wenjie(1135)
- Multi-focus image fusion using adaptive dictionary learning method Yan Chunman, Guo Baolong, Yi Meng(1144)
- Geometric constraints of locating cameras and people Tao Linmi, Yu Yapeng, Di Huijun, Sun Luo(1150)

Virtual Reality and Augmented Reality

- Dynamic evaluation and modification methods against aggressive mesh hierarchy models
..... Wang Danhua, Lu Wei, Pan Jingui(1158)

Medical Image Processing

- Computer aided diagnosis on intervertebral disc degeneration with shape and texture features
..... Hao Shijie, Jiang Jianguo, Guo Yanrong, Zhan Shu, Li Hong(1168)
- Improved coupled model for MR images segmentation and bias restoration
..... Wang Shunfeng, Ji Xiaona, Zhang Jianwei, Chen Yunjie, Fang Lin, Zhan Tianming(1175)

Remote Sensing Image Processing

- Ship detection from low observable regions in optical remote sensing imagery Zhou Wei, Guan Jian, He You(1181)
- SAR and multispectral image fusion algorithm based on pulse coupled neural networks and non-subsampled Contourlet transform
..... Jin Xing, Li Huihui, Shi Pili(1188)

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2012)09-1150-08

论文引用格式: 陶霖密, 于亚鹏, 邸慧军, 孙洛. 摄像机几何约束及人体定位[J]. 中国图象图形学报, 2012, 17(9): 1150-1157.

摄像机几何约束及人体定位

陶霖密, 于亚鹏, 邸慧军, 孙洛

清华大学计算机系, 北京 100084

摘要: 人体定位一直是计算机视觉领域的热点课题之一, 是人机交互的基础。提出多摄像机条件下的投影关系的两类几何约束, 及基于几何约束的人体定位方法。人体定位算法的精度计算依赖于实验中人体在世界坐标系中精确位置的获取, 而这一直是计算机视觉实验中的难点。本文提出基于虚拟数据的新的人体定位实验方法, 可以方便准确地获取到人体在世界坐标系的位置。由于几何约束的完备性, 定位过程中无需对摄像机的位置和姿态进行假设, 也不需要限制人体在场景中的位置(遮挡关系)、行为等。采用虚拟数据与真实数据相结合, 可以更准确地验证方法的精确度和适应性。实验结果表明, 本文算法运行稳定且定位精度高。在定位结果的基础上, 又进一步给出了人体在世界坐标系中的运动轨迹。

关键词: 几何约束; 投影变换; EM 算法; 虚拟实验

Geometric constraints of locating cameras and people

Tao Linmi, Yu Yapeng, Di Huijun, Sun Luo

College of Computer Science and Technology, Tsinghua University, Beijing 100084, China

Abstract: Locating people has been one of the most popular subjects in computer vision, which is the basis for human-computer interaction. Locating people and motion tracking is the inverse problem of multi-camera imaging, which is gaining videos from one or more cameras, to compute the location or motion of people in videos. In this paper, we propose two types of geometrical constraints for multi-camera approaches of locating people. The accuracy of people locating algorithms relies on the exact location in the world coordinate system during experiments, which has been a great difficulty in computer vision experiments. In this paper, we present a new people locating experimental method based on virtual data, from which we can more easily get the accurate location of people in the world coordinate system. Because of the completeness of the geometric constraints, we do not need to assume the cameras' placement and restrict the location (occlusion) or motion of the people in the scene. Using virtual and real data in the experiments, it can be more accurate for verifying the accuracy and adaptability of the algorithm. The experimental results show that the algorithm works stable and accurate. Addition, the trajectory of the people can be acquired based on the results.

Key words: geometric constraint; homography mapping; EM algorithm; virtual experiments

0 引言

人体运动跟踪和定位一直是计算视觉研究的重

要问题之一, 人体运动和位置信息是行为计算的重要特征, 在人本计算和交互领域, 具有重要的应用。例如在智能家居、老年人看护等应用中, 必须首先知道场景中人体的目标位置和运动轨迹, 才能对人的

收稿日期: 2012-02-09; 修回日期: 2012-03-30

基金项目: 国家自然科学基金项目(60873266, 90820304)

第一作者简介: 陶霖密(1962—), 男, 副教授, 2001年清华大学获博士学位, 主要研究领域为计算机视觉、人机交互。E-mail: tao.

linmi@gmail.com

行为进行理解和交互。本文将就该问题提出基于几何约束的定位算法,并将之扩展应用于多目标定位的场合。

计算机视觉领域对基于多摄像机的人体目标的定位,已有多年的研究,主要方法可以分为两类:一类是图像特征法,即从每个单独的摄像机捕捉到的场景中挖掘出所需的特征信息,作为多摄像机联合定位的数据基础。文献[1-3]提出以人体的颜色信息为特征的方法无疑是最简单直接的特征提取方向,即通过颜色信息进行相关模型的建立,以达到目标识别和定位的目的。但是该方法对场景的环境变量和目标的本体特征过于敏感,在有光照、阴影、遮挡、衣着差异等各种噪声影响的情况下,会出现不可避免的误差或错误的识别或定位结果。文献[4]中根据目标的质地特性(刚性)选用角点和边缘特征实现目标的定位跟踪,相对于颜色来说,对环境的变化更加鲁棒。

Eshel^[5]和Khan^[6]分别利用基于头部和足部的特征完成对特定场景下目标的定位。Dockstader^[7]则提出更加关注于局域点阵的定位方法,可较为充分地利用人体行为的连续信息提高定位的精准度。这些基于区域模块点集的算法,对目标局域点特征过于敏感,在场景环境较为复杂,或遮挡的情况下,将会导致目标丢失情况发生。

另一类是基于前背景分割的方法,包括中心轴线法和直接像素法。其中文献[5,8]等采用中心轴线法,首先对单摄像机视频进行前背景分割,对提取到的前景点集进行预处理,求得相对最优化的中心轴线或质心点等,然后以这些处理过的信息作为多摄像机联合定位的数据集,求解得到人体的位置。这类方法的最大问题是,前背景分割的问题将会极大地影响定位的精度;文献[1,6]等采用直接像素法,提出将前背景分割得到的前景点直接作为多摄像机联合定位的数据集,通过优化的方法,求得人体的位置。这类算法尽可能多地保留了前背景分割数据所包含的信息,发挥了多摄像机联合定位的优势,相对前者避免了因预处理过程而产生的计算误差,使定位结果相对更加精准。但是这类算法一般都具有比较高的计算复杂度。

在前景点集的基础上,Hu^[8]提出了基于几何约束的算法,对当前场景的环境变量和人体目标的颜色、局部区域特征没有过多关注,在几何关系的基础上,对摄像机成像平面中的投影不进行敏感区域提取或划分(即人体特殊部位模型匹配,包括头部、脸

部、足部、手部等),利用的是全体投影点集。通过摄像机的平面坐标系变换关系,利用较为严谨的几何约束方法,定位人体目标在投影平面上的位置。但是文献[8]对摄像机的安装有约束,要求摄像机投影平面垂直于地平面,这在室外较开阔处可以基本成立,但在室内场景下则难以成立。文献[8]中对单摄像机的前景点进行轴线归一的预处理,降低了前景点信息的利用率。

人体运动跟踪和定位是人体目标在多摄像机成像中的逆问题,即从一台或多台摄像机所获取的视频中,通过分析人体目标特征和运动,得到人体在当前世界坐标系下的位置和运动轨迹。在物理时空中,人体目标的位置在时空坐标系中是唯一确定的,即在某一时刻人体目标的空间坐标是唯一确定的,相应地,人体目标在摄像机中的投影成像也是唯一确定的。因此,计算机视觉中人体运动跟踪和定位的实质就是通过分析由视觉设备获取的人在摄像机中的成像,根据摄像机的投影关系,求解得到该时刻人体目标在世界坐标系下的唯一位置。由于目标在摄像机上的投影,失去了目标和摄像机之间的距离信息,在实际计算中,根据目标和摄像机之间的投影几何关系,仅能求出目标所处位置和摄像机光心投影之间的连线。因此,理论上,两个非共光轴的摄像机,如目标位于两个摄像机的共同可视区域内,就可以根据投影关系求解得到两直线之间的交点,从而确定目标在世界坐标系中的唯一位置。

基于上述分析,本文扩展了文献[9]的工作,提出了整合利用所有摄像机前景点集的数据,根据纯粹的几何约束直接求得当前场景下多人目标的空间位置。该算法中不涉及人体模型,对人体的运动、动作没有约束,充分利用每个摄像机成像平面所得的投影,对整个投影前景点进行整体几何关系处理,直接实现人体目标的空间定位。且对摄像机位置、朝向等无约束,仅需目标能够被至少两个摄像机捕获即可。通过目标定位之后,可进一步获取到多目标的运动轨迹等后续信息。

1 方法概述

将坐标系设定为世界坐标系,即以地面为XOY平面,垂直地面向上为纵轴Z方向。根据环境需要,在场景中布置 $K(K \geq 2)$ 个摄像机,以保证场景中的人体目标至少出现在两个摄像机的视场中,对

摄像机的朝向、俯仰等没有要求。设定场景中有 $N(N \geq 2)$ 个人体目标在自由运动,此处对人体的运动方式(姿态、轨迹等)不加约束。这里对算法轻量级的假定,目的是保证自由运动的人体能够同时出现在多个视场中,在纯几何关系的条件下,能够得到深度信息,并减少遮挡等噪声的影响。

在同一时刻,所有摄像机开始拍摄,得到一组同步的视频或图片,设定每个相机所拍摄的视频或图片为 $C_k(k = 1, 2, 3, \dots, K)$ 。传统的基于区域的方法,则需要对 C_k 进行分析处理,得到关心的特征,例如颜色、亮度、足部等。本文算法则不涉及人体本体的特征,而对 C_k 进行背景模型处理后得到前景点,投影到地面坐标系,得到前景点簇 $H_k^i(k = 1, 2, 3, \dots, K, i$ 为对应的帧号)。

帧号 i 所对应的每一簇 H_k^i , 根据相关几何约束关系可以得到当前帧时刻下,每个人体目标在场景中所存在的坐标,即得到其所在的位置。而跟踪每一组帧所得到的位置点,就可进一步得到目标的运动轨迹。

由于人体目标体积的不可忽略性,理想情况下纯粹的几何约束关系所造成问题不可解,我们选用概率的方法进行进一步处理。

2 多摄像机环境下的约束关系

Homography 用于描述两个映射平面之间的变换关系,一平面上的直线经过 Homography 变换后在其对应的映射平面上仍为直线。Homography 投影变换为

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \mathbf{H}_g \cdot \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \quad (1)$$

式中,点 (x_i, y_i) 和 (u_i, v_i) 分别为两平面中的对应点, \mathbf{H}_g 即为两平面间的 Homography 变换矩阵。该矩阵中有 8 个未知量,可以通过场景中的 4 个已知点来求解得到。

基于上述变换关系,提出两类多摄像机的几何约束,即“多线单摄像机共点约束”与“单线多摄像机共点约束”。而文献[8]中提出的几何关系并不完备,这也导致其对摄像机镜头必须水平摆放方式的假定,而本文中在完备约束下无需对摄像机有过多假定。同时相较于文献[8]首先在单一的摄像机

平面中进行预处理(即在每个摄像机成像平面中将前景点集归一为一条代表轴线),在这里不进行预处理,而是直接将所有摄像机中得到的前景点全部投影到统一的地平面坐标系下,求最优解得到人体的位置,提高了定位准确性。

在上述投影变换的基础上, π 平面上的垂线存在如下两个共点约束(如图 1 所示,地平面 π 中两个人,处于两个摄像机视场内的投影关系约束):

约束 1 多线单摄像机共点约束:垂直于平面 π 的不同直线相对于同一个摄像机 C 在 π 平面上的 Homography 投影所构成的直线簇具有共点性,其交点为摄像机 C 的光心在平面 π 上的垂直投影 C_π 。如图 1 所示,二直线在 π 平面的投影 $\overline{P_{1\pi 1} Q_{1\pi 1}}$ 、 $\overline{P_{2\pi 1} Q_{2\pi 1}}$ 相交于摄像机光心的投影 $C_{1\pi}$ 。

约束 2 单线多摄像机共点约束:垂直于平面 π 的直线 L 经过不同摄像机在 π 上的 Homography 投影所构成的直线簇具有共点性,其交点即为直线 L 在 π 上的垂足。如图 1 所示,垂直于 π 平面的线段 $\overline{P_1 Q_1}$ 在两摄像机下的投影 $\overline{P_{1\pi 1} Q_{1\pi 1}}$ 和 $\overline{P_{1\pi 2} Q_{1\pi 2}}$ 相较于线段的垂足 $F_{1\pi}$ 。

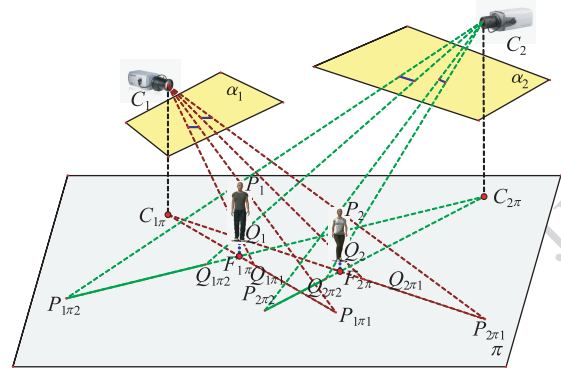


图 1 几何约束

Fig. 1 Geometric constraint

上述几何约束表明,人体目标在地面上的位置,可以通过以下两步求解得到:1)通过 Homography 投影变换,从图像计算得到人体目标在地面的投影;2)求解其地面投影轴的交点。该交点不但需要遵循上述几何约束,而且由于时间和空间的唯一确定性,即在同一时间同一地点只能存在单一的物体,所以使用多摄像机对多目标进行检测,可以将物体完全区分开。

3 多目标定位算法描述

根据第 2 节中描述的约束关系,可以很直观地

从几何关系的角度得到多目标定位算法,具体算法步骤如下:

1)已知垂线 $\overline{P_1Q_1}$ 在摄像机成像平面 α_1, α_2 中的像为线段 $\overline{P_{1\alpha_1}Q_{1\alpha_1}}, \overline{P_{1\alpha_2}Q_{1\alpha_2}}$, 其中 $(P_{1\alpha_1}, Q_{1\alpha_1}), (P_{1\alpha_2}, Q_{1\alpha_2})$ 为上述线段的顶点坐标;

2)根据式(1)定义的 Homography 变换,可以求解得到线段 $\overline{P_{1\alpha_1}Q_{1\alpha_1}}, \overline{P_{1\alpha_2}Q_{1\alpha_2}}$ 在平面上的投影线段 $\overline{P_{1\pi_1}Q_{1\pi_1}}, \overline{P_{1\pi_2}Q_{1\pi_2}}$, 其中 $(P_{1\pi_1}, Q_{1\pi_1}), (P_{1\pi_2}, Q_{1\pi_2})$ 为上述线段顶点的世界坐标;

3)同理可以求得垂线 $\overline{P_2Q_2}, \overline{P_3Q_3}$ 在 π 平面上的投影线段 $\overline{P_{2\pi_1}Q_{2\pi_1}}, \overline{P_{2\pi_2}Q_{2\pi_2}}, \overline{P_{3\pi_1}Q_{3\pi_1}}, \overline{P_{3\pi_2}Q_{3\pi_2}}$ 的世界坐标;

4)求3组线段所在直线的两两交点,即可得到上述3条垂线的垂足的世界坐标 $F_{1\pi}, F_{2\pi}, F_{3\pi}$ 。

但是上述对基于几何约束的分析均是在一个默认假设情况下进行的,即目标没有体积,如图2中将人体抽象为一条垂直于地面的直线,图2中各色连线表示为摄像机光心投影与目标当前位置之间的连线,通过人体投影点的最优位置。而实际的人体其体积在距离镜头非足够远的情况下不可忽略,目标在成像平面 α 上的像具有一定的面积,经过 Homography 投影到 π 平面后(图2),依然具有一定的面积,并非是一条理想的线段。这就造成了纯粹的几何关系在实际定位中无法真正实现。

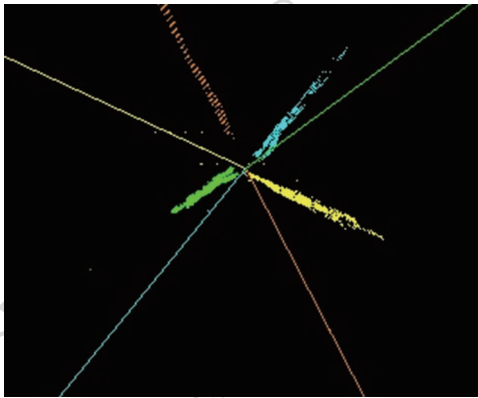


图2 人体地平面投影

Fig. 2 Body mapping on the ground

因此选择以概率模型的方式,并设计图3所示算法流程:从多台摄像机获取的数据必须首先进行同步处理,以保证帧间的对应关系;利用背景模型,对每一组视频帧作前背景分割,得到相应的前景点集;根据每个摄像机平面与地平面间的几何投影关系,将前景点集映射到地平面坐标系下,统一坐标

系;由 EM 算法(expectation-maximization algorithm, 即期望最大化算法)的思想,联合前景点集中的所有点,得到当前迭代下每个目标的最优化位置;继续迭代,直至结果相对稳定。图3中核心部分在于根据 EM 算法,求得目标在所有摄像机下最优化的坐标点,即使所有 Homography 投影点以最优化的方式,分布在目标坐标与摄像机光心投影点坐标所确定的直线周围,例如使所有投影点到连线距离最短。经推导论证,该最优化问题可以简化为一维线性优化问题,可以利用最小二乘法求解。

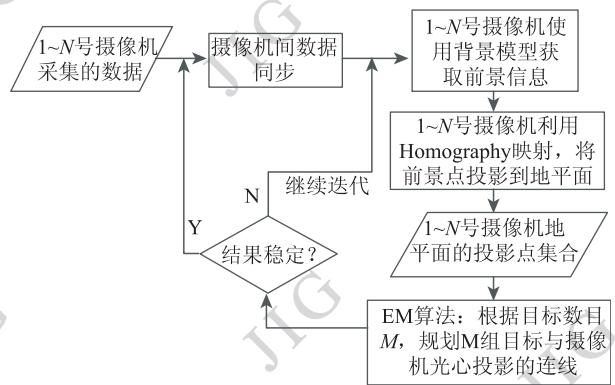


图3 算法流程图

Fig. 3 Flow chart of the algorithm

4 实验与分析

利用虚拟数据真实测定算法精确度。传统的实际场景数据,尤其是视觉领域的的数据,由于采集帧率较高,很难达到算法结果与摄像机当前帧的时间点同步,从而导致无法准确测定算法的精确程度或者误差率等。但是使用虚拟数据,可以完全获得任意所需时刻下的目标数据(本文中为人体体的空间位置),为算法结果的讨论提供了较好的数据准备。虚拟数据也可以完全真实反映实际场景下的所有情景,能够保证算法在真实场景下的可适用性。

选择以 Smith Micro Software 公司的 Poser 软件生成虚拟数据。场景中围绕地平面区域布置了10台状态各异的摄像机,保证在同一时刻目标能够被多台摄像机同时捕捉到。

实验均是在如下机器配置条件下完成的:单台 PC 台式机、处理器 Intel® Core™2 Duo CPU E8400@3.00 GHz、内存 4 G;程序以单进程模式设计。安排虚拟数据实验时,从图4中有目的地选取特定的摄

像机,以获得不同数目、不同角度条件下视场中的实验数据,分别设计了摄像机与人体共轴的特殊情况、不同数目摄像机(即数据量)情况下的定位实验,以及在产生误差(数据量逐层减少)情况下的定位结果等方面的实验。同时提供了在真实实验环境下的实验,以验证算法的适应性。

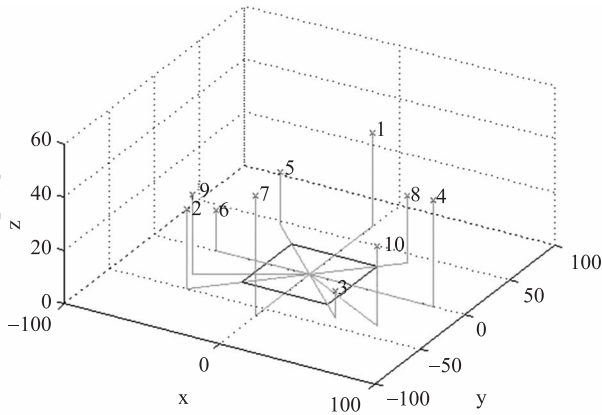
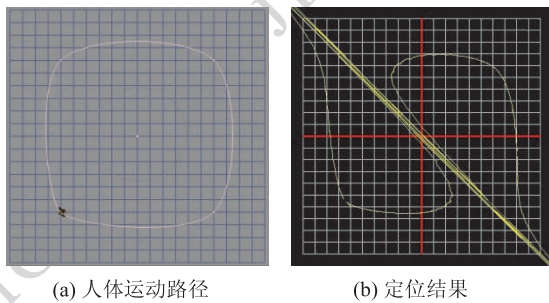


图 4 摄像机分布图

Fig. 4 Positions of the cameras

4.1 摄像机共轴

多摄像机共轴问题是在多摄像机定位问题中出现的特定问题,是指场景中的多个摄像机在分布时出现两个或两个以上光心轴线共线的情况。当目标运动到该轴线上时,将失去深度信息,出现如图 5 所示的情况。



(a) 人体运动路径

(b) 定位结果

图 5 共轴多摄像机结果跳变

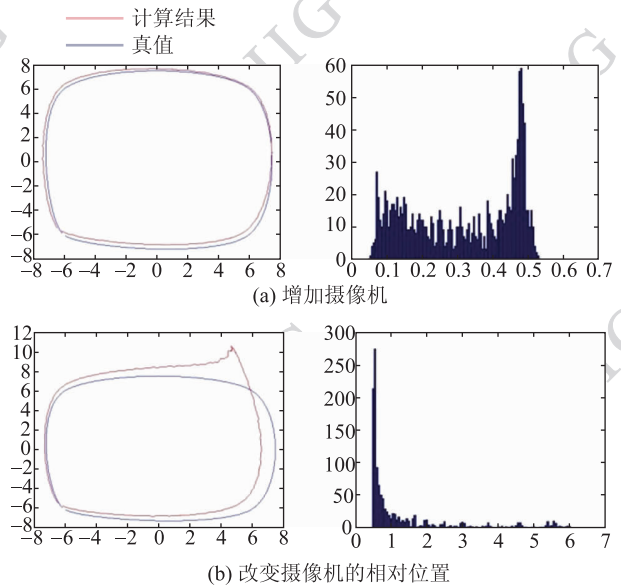
Fig. 5 The result of coaxial multi-cameras transitions

正如前文所述,使用虚拟数据可以得到确定时刻下的准确真值数据。图 5(a)为人体在视场中运动方式的俯视图,白色环行线为实际的人体运动轨迹,可以得到任意时刻下人体运动的精确位置。而图 5(b)黄色表示的是定位结果,即通过算法测定出的人体目标的运动轨迹线。通过对比左右两图中的真值与计算值,可以得到算法的精确

定量分析。

由于视场中仅布局的两个摄像机分别位于左上角和右下角的对角线方向(图 4 中的 3 号、5 号摄像机),导致目标运行到接近两摄像机光心连线时,丢失了几乎所有的深度信息,进而在中央区域定位结果出现跳跃抖动。但是过了光心连线后,本文算法可以进行迅速调整,使定位结果恢复至真值所在轨迹上。

要消除摄像机共轴所带来的影响,可以采取增加摄像机(添加 8 号摄像机)或调整两摄像机的位置(改为使用 1 号、3 号摄像机)等方式。如图 6 所示,图 6(a)为增加 8 号摄像机的运行结果,图 6(b)为改用 1 号、3 号摄像机的定位结果。左侧一列为算法计算结果(红色)和真值(蓝色)的对比图,而右侧一列则为误差统计结果。从统计结果可以看出,调整摄像机位置而得到的结果误差远远超过通过增加摄像机的方法。针对此情况,下面将在该实验基础上讨论算法误差等相关结果。



(a) 增加摄像机

(b) 改变摄像机的相对位置

图 6 共轴摄像机解决方案

Fig. 6 Solution for coaxial cameras; increase the camera or change the relative position of the cameras

4.2 误差分析

从图 4 中选取不同数目的摄像机,进行不同数目摄像机下定位结果分析。选取 1 号、9 号、10 号作为 3 台摄像机实验时的机位;2 号、3 号、5 号、8 号作为 4 台摄像机实验时的机位;1~8 号作为 8 台摄像机实验时的机位。

通过继续对上述实验模型进行不同数目摄像机下的实验测试,得到如表 1 所示结果。

表1 不同数目摄像机下定位结果

Table 1 Locating results with different numbers of cameras

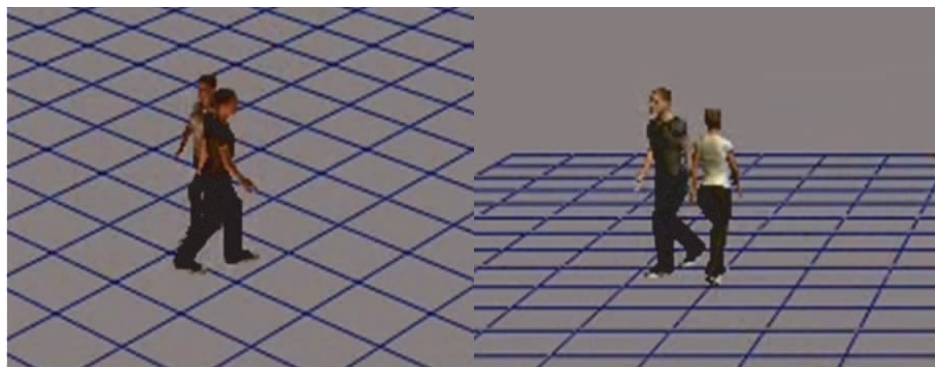
摄像机数	误差最大值	误差最小值	误差均值	误差值方差
3	0.372 7	0.016 0	0.221 9	0.008 3
4	0.375 5	0.002 4	0.214 6	0.008 1
8	0.342 6	0.013 5	0.207 9	0.006 9

从表1的统计结果可以看出,随着摄像机数目的增加,定位误差在逐渐降低。因为随着摄像机数目的不断增加,可利用的数据量在倍增,而本文算法不存在累计误差的问题,所以定位误差随着摄像机数目的增加稳定降低。

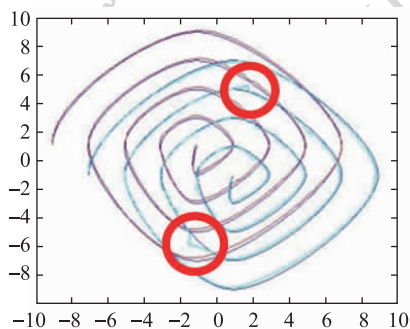
另一方面,多目标定位的情况下存在“接近误差”,即不同目标互相靠近达到一定距离的情况下,由于在特定摄像机视角下产生的目标遮挡重叠现

象,而使定位出现的误差。而在目标没有相互接近,即摄像机可以较为明确区分目标的情况下,可以视为多个单目标共同定位,仅仅只是增加了对不同目标的运算量。

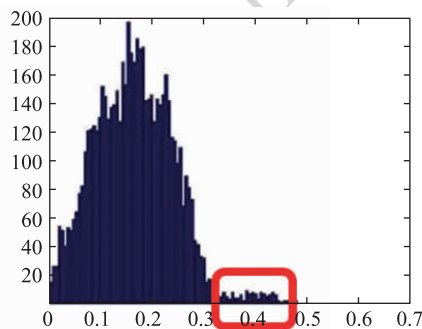
如图7所示实验,场景中两目标随螺旋轨迹运动过程中,目标间距离存在两次逐渐靠近与分离的过程。图中第一行中显示的是同一时刻两个视角下的视频截图,两个目标的相对位置明显不同;第二行则表示在多台摄像机下联合定位的结果。从结果对比和误差统计图中可以明显看出,随着目标之间距离的减少,由于不同目标在同一摄像机下数据集的相互覆盖叠加,致使目标间产生相互影响,如图7(a)(b)中红色圈定区域所示,即为两目标靠近后造成的误差增加。而在极限情况下,即两目标最终重叠,抑或真实场景下两目标完全贴合,则在本文基于纯几何关系定位的算法中,只能将之视为同一个目标。



(a) 目标运动视频数据



(b) 定位结果与真值对比



(c) 定位结果误差统计

图7 接近误差实验结果

Fig. 7 Experimental results of closing-error

4.3 算法稳定性及运行效率测试

为了测试算法稳定性,对之前实验数据的分割“产生质疑”,即假设采集到的数据进行前景点分割时分别产生10%、50%、90%的误差率,进行算法测试。

以前背景分割的数据量减少来模拟前景点分割错误等造成的数据损失。由表2可以看出,本文算法在前景点数据量骤减的情况下,对于目标的定位产生的误差几乎不变,即本文算法容错性较高,在真实场景中也取得较好的定位结果。

表 2 分割错误定位结果对比

Table 2 Results with segmentation fault contrast

分割错误率/%	误差最大值	误差最小值	误差均值	误差值方差
0	0.342 7	0.013 5	0.207 9	0.006 931
10	0.349 9	0.012 4	0.207 9	0.006 958
50	0.345 1	0.011 6	0.208 2	0.006 983
90	0.348 7	0.006 5	0.210 0	0.006 920

如表 3 所示,在不同摄像机数目即数据集情况下,对本文算法的运行效率的大致测定。该次测试是运行在单机 PC 环境下,对不同摄像机数据集及整合计算是采取单线程顺序化处理。在实际工程应用中,可以采取多线程并行化处理;抑或将不同摄像机数据处理分布化到不同机器中,处理结果传递给中央集成计算,则算法运行效率将成倍提高,完全可以满足工程中实时化定位运算的需求。

表 3 算法运算效率

Table 3 Efficiency of the algorithm computing

摄像机数	视频片段长度/帧	算法运算总时间/s	平均运算效率/(帧/s)
2	1 300	104.161	12.480 7
3	1 300	147.452	8.816 43
4	1 300	193.561	6.716 23
6	1 300	280.645	4.632 19
8	1 300	364.666	3.564 91

4.4 实际场景数据实验

经过虚拟数据情况下的实验及其结果统计分析,可以看出,本文算法在目标数目增加的情况下,仍然有较好的稳定性,且误差均处于较低的范围。在多摄像机定位情况下,通过增加摄像机数目或对摄像机覆盖区域的调整,本文算法可进一步提高定位的准确度。

如图 8 所示,为在一般实验室条件下进行真实场景数据拍摄的环境。场景中布置了高低、倾斜角度等均不相同的 4 台摄像机,每个摄像机的视场范围存在较大差异性。且场景中存在沙发桌椅等遮挡物,墙壁、地面等存在一定程度的镜面反射效果。场景中灯光强度等其他环境因素相对稳定。



图 8 真实场景数据部分截图展示

Fig. 8 Some screenshots of the real scene data

从图 9 中可以看出,在目标运动过程中定位结果一直较为平滑和稳定。在右下角区域起步的过程中,由于目标过于接近墙面而导致过多的镜面反射噪声的影响出现一定程度的跳变。而图 9 中 3 处环形区域所标示的部分,则是目标长期驻足于某个区域(沙发、冰箱)而导致前景分割后数据量减少,而产生的区域性抖动。但是在逐渐离开特殊区域后本算法得到较为稳定的定位结果。

经过上述对真实场景中数据的定性分析可以看出,本文算法对真实场景有较好的适应性和稳定性,且对障碍物遮挡等情况能够在摄像机覆盖的情况下较为完美的解决,发挥了多摄像机定位的优势。同时在经过结果平滑性处理的情况下,能够得到较为平滑的结果集,且没有过多的数据跳变。

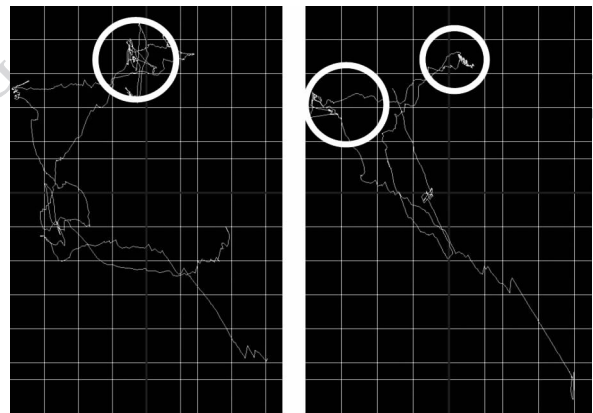


图 9 真实场景下定位结果

Fig. 9 Locating result in real scene

5 结 论

在多摄像机条件下对目标进行定位的稳定算法,基于几何投影关系,在对目标本身的外部特征没有任何关注的情况下,能够充分利用每个摄像机捕获到的前景点信息,通过联合所有摄像机直接得到目标的空间位置,同样避免了跨摄像机进行目标匹配的计算工作。大量的实验结果表明,本文算法具有定位精度高,在有障碍物遮挡、色变、光强等多种外部干扰的情况下运行稳定。

本文实验测定方法,利用虚拟数据的真实可读性,可获取到所需的任意时刻下的数据。较之实际场景数据,无疑对算法精确度和误差率的测定更为准确。同时虚拟数据可以在实验环境下得到所需的任意场景,对于计算机视觉领域所需的复杂而精确的数据测定尤为有效。

此外由于本文算法仅涉及目标前景点集,在实际应用中可以通过添加其他外部特征(颜色等),进一步提高定位的稳定性,和对多目标的区分度。

参考文献(References)

- [1] Mittal A, Davis L S. M2tracker: a multi-view approach to segmenting and tracking people in a cluttered scene using region-based stereo [J]. *International Journal of Computer Vision*, 2003, 51(3): 189-203.
- [2] Krumm J, Harris S, Meyers B, et al. Multi-camera multi-person tracking for easy living [C] // *Proc. of IEEE International Workshop Visual Surveillance*. Washington DC, USA: IEEE Computer Society, 2000:3-10.
- [3] Wei G, Petrushin V A, Gershman A V. Multiple-camera people localization in a cluttered environment [C] // *Proc. of the 5th international workshop on Multimedia Data Mining*. Seattle, WA, USA: Valery A. Petrushin, 2004: 52-60.
- [4] Martin F and Horaud T. Multiple-camera tracking of rigid objects [J]. *International Journal of Robotics Research*, 2002, 21(2): 97-113.
- [5] Eshel R, Moses Y. Homography based multiple camera detection and tracking of people in a dense crowd [C] // *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Alaska, USA: IEEE 2008, 5:1-8.
- [6] Khan S, Shah M. A multi-view approach to tracking people in crowded scenes using a planar homography constraint [C] // *Proc. of the 9th European Conf. on Computer Vision*. Graz, Austria: *Lecture Notes in Computer Science*, 2006, 6:133-146.
- [7] Dockstader S L, Tekalp A M. Multiple camera tracking of interacting and occluded human motion [J]. *Proceedings of the IEEE*. 2001, 89(10): 1441-1455.
- [8] Hu W M, Hu M, Tan T N, et al. Principal axis-based correspondence between multiple cameras for people tracking [J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2006, 28(4): 663-671.
- [9] Sun L, Di H J, Tao L M, et al. A robust approach for person localization in multi-camera environment [C] // *Proceedings of the 20th International Conference on Pattern Recognition*. Istanbul, Turkey: IEEE, 2010:4036-4039.