

DOI:10.19651/j.cnki.emt.2107977

基于计算机视觉的室内定位系统设计与实现

周宇辉 朱晓强 曾丹

(上海大学通信与信息工程学院 上海 200444)

摘要: 现在传统的室内定位技术大多需要在被定位物体上安装标签或者终端设备,在一些特定场合存在很大的局限性。为此提出了一种基于计算机视觉的室内定位系统。该系统通过目标检测算法检测出特定物体并获取其图像坐标,经过目标点判定算法判断该物体是否在待定位区域内,最后通过投影变换算法获取到该物体对应的地图坐标。在被定位物体无感知的情况下,完成对其定位且定位误差在1 m以内。为了方便用户直观地观察,采用网页的形式,将实际地图与物体位置显示在终端设备上。

关键词: 无感知;室内定位;计算机视觉;投影变换

中图分类号: TP391.7 **文献标识码:** A **国家标准学科分类代码:** 520.20

Design and implementation of indoor positioning system based on computer vision

Zhou Zihui Zhu Xiaoqiang Zeng Dan

(School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China)

Abstract: Now the traditional indoor positioning technology mostly needs to install labels or terminal equipment on the located object, which has great limitations in some specific occasions. Therefore, an indoor positioning system based on computer vision is proposed. The system detects a specific object and obtains its image coordinates through the target detection algorithm, judges whether the object is in the area to be located through the target point determination algorithm, and finally obtains the map coordinates corresponding to the object through the projection transformation algorithm. When the object to be located has no perception, the positioning is completed and the positioning error is within 1 m. In order to facilitate users' intuitive observation, the actual map and object position are displayed on the terminal device in the form of web page.

Keywords: no perception; indoor positioning; computer vision; projection transformation

0 引言

随着我国北斗卫星导航系统的全面建成与部署,室外导航定位系统趋于成熟^[1]。而室内定位系统越来越被人们所关注与研究。现在主流的室内定位系统按照定位方式可分为信号到达时间(time of arrival, TOA)^[2]、信号到达时间差(time difference of arrival, TDOA)^[3]、信号到达角(angle of arrival, AOA)^[4]、接受信号强度(received signal strength, RSS)^[5]等。按照定位技术又有超宽带^[6]、射频^[7]、WiFi^[8]、蓝牙^[9]等。

在室内定位技术上有许多学者做了大量的研究。文献[3]提出一种超宽带室内定位技术,该技术定位精度较高,但是设备成本高且定位范围受设备部署范围限制。文献[4]提出一种融合定位算法,该算法利用已知节点位置通

过算法融合来获取未知节点的坐标,但受节点位置影响较大对未知节点未知定位准确度不高。文献[5]提出一种基于信号强度定位算法。该算法通过获取信号强度建立指纹库来进行定位,但信号强度受环境影响明显。文献[7]和文献[9]分别通过射频和蓝牙信号来实现室内定位。同样受环境影响较大,信号强度衰减的厉害定位准确度大打折扣。

上述定位技术在一些特定场合中如封闭园区、室内停车场^[10]不仅定位精度不高而且可操作性低,尤其是有目标频繁进出的情况上述方案成本高昂。

基于以上情况,本文研究目的为提出一种用户无感知且不需要携带任何终端设备的定位方式即基于计算机视觉的定位系统^[11]。研究设想为仅通过若干路摄像头以及一台服务器就可以完成目标的定位。本实验设计为通过目前较为成熟的目标检测算法^[12-13]获取被定位物体的像素坐标

值,使用坐标检测算法判断目标是否在定位区域内并通过坐标映射算法将像素坐标转换成地图坐标,预期结果为最终将物体位置显示在网页界面中。从而实现在被定位物体无感知状态下的实时定位,具有重要的实际运用意义。

1 室内定位系统

整个室内定位系统主要分为 3 个部分:1)由多路摄像头及目标检测算法组成的目标定位模块;2)检测物体是否在定位区域内并将物体图像坐标转换成地图坐标的分析与映射模块;3)将定位数据显示在出来的显示模块。此系统将持续监测被定位区域,将进入该区域内的特定物体实时定位并显示在网页界面中。

1.1 系统整体框架

整个室内定位系统如图 1 所示,若干个摄像头采集到视频数据交由目标检测算法获取视频中特定物体的图像坐标,将计算出的坐标数据传送到分析模块中判断并获取该物体的地图坐标,最后在网页中实时显示出该物体在地图中的位置。

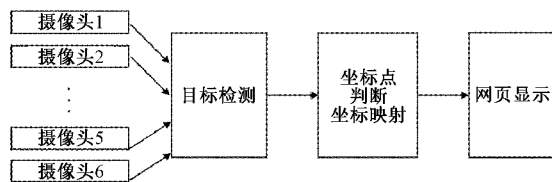


图 1 系统框图

1.2 目标定位模块

该模块主要分为 2 个部分:1)用于获取视频数据的摄像头;2)检测视频数据中特定物体的目标检测算法。要想完成对一片区域中特定物体的定位,需要在该区域部署一定数量的摄像头并尽量保证摄像头对该区域的全面覆盖。通过网线可以将摄像头获取到的视频数据传输到服务器,服务器通过目标检测算法可以实时检测出视频流中的特定物体并获取该物体在图像中的像素坐标,同摄像头编号一起组成该物体唯一的坐标编码。该数据将会传输到下一个模块中进行分析处理。本系统的目标检测算法使用的是 YoloV3 算法,该算法实时性好准确度高^[14]。

1.3 数据分析与映射模块

该模块主要分为 2 个部分:1)判断物体是否在被定位区域内;2)将图像坐标映射到地图坐标。由于摄像头拍摄到的图像存在畸变以及多个摄像头可能重复覆盖同一区域,这将导致图像中一部分区域无法被映射或存在重复映射的可能,因此需要对摄像头中的图像进行区域划分,并在地图坐标中相应位置进行标定。

完成上述操作以后,就可以使用投影变换将图像中的指定区域映射到地图坐标上相应的位置。投影变换法是本系统设计的一个非常重要的算法。使用该方法可以很好地完成图像坐标到地图坐标的转化。

具体的实施方案如下,首先将每路摄像头采集到的图像保存下来,然后使用 OpenCV 工具打开图片并在图片相应位置选取 4 个坐标点并记录下来如图 2 所示中实线方框位置,同样的在地图坐标中选取相对应的 4 个点并记录如图 3 所示中实线方框,之后使用投影变换法建立两个区域之间的映射关系,后续可以通过该映射关系将图像坐标转换成地图坐标。



图 2 摄像头采集图像

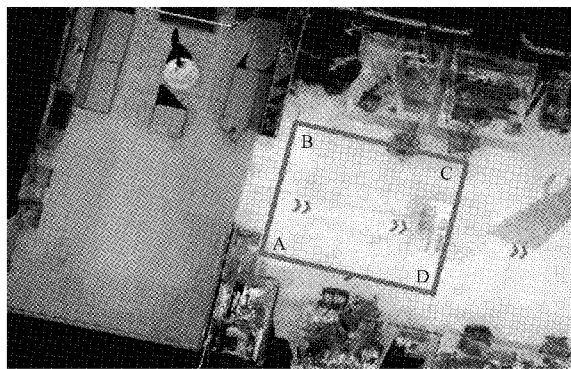


图 3 地图局部图像

1.4 显示模块

经过摄像头获取视频数据经由目标检测算法获取物体图像坐标,再经过投影变换将图像坐标转换成地图坐标后就可以在地图对应位置显示出相应物体。为了能够让用户操作方便无论是在移动端还是在电脑端都可以显示,本系统采用网页的形式展示出地图和对应物体定位坐标点。网页中将某公司实验室的俯视图加载进来,当服务端解析出物体坐标时发送至网页,网页中将在对应的位置以图标的形式显示出相应的物体,如图 4 所示。具体摄像头拍摄到的图像如图 5 所示。

2 系统实现相关算法

2.1 目标检测算法

目标检测算法的准确度直接影响到了本文系统的定位精度,因为本系统定位物体的前提条件就是能够将物体实时检测出来。由于神经网络技术的不断发展,现在有很多比较流行的目标检测算法,其中 YoloV3 算法有着较高的精度与

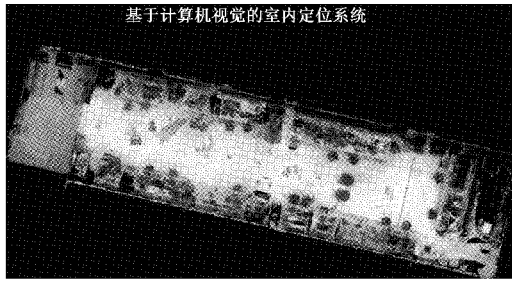


图4 网页显示定位位置



图5 摄像头拍摄图片

非常好的实时性。因此本系统的目标检测算法采用的是 YoloV3 算法。根据文献该算法的检测框解码方法如下:

$$b_x = \sigma(t_x) + c_x \quad (1)$$

$$b_y = \sigma(t_y) + c_y \quad (2)$$

$$b_w = p_w e^{t_w} \quad (3)$$

$$b_h = p_h e^{t_h} \quad (4)$$

其中, b_x, b_y 为预测框的中心点坐标; b_w, b_h 为预测框的宽和高大小; p_w, p_h 为先验框的宽和高; t_x, t_y, t_w, t_h 为预测框的中心点坐标和对应的宽和高; c_x, c_y 为对应特征图的维度值。由此可以得到待测物体的图像坐标。

2.2 目标点判定算法

在本文 1.3 节提到,需要在图像中选取 4 个坐标点形成一块区域。那么在实际使用过程中只有当物体出现在这片区域内才需要进行坐标映射。那么需要使用一个目标点判定算法,判定当前点是否在指定区域内。本文系统中使用的算法为 PNPoly 算法。该算法的核心如下:1)选取最小的横纵坐标与最大的横纵坐标,判断该点是否在这个大的矩形内,如果在矩形外则直接返回 false。2)如果在矩形内,则进一步通过判断该点纵坐标相邻两点的连线是否在该点的上方,统计个数,如果为偶数个则在区域内反之在区域外。该算法的流程如图 6 所示。返回结果为是否在指定区域内部。

2.3 投影变换算法

在投影变换算法中,引入一个概念称之为单应性(Homography)变换^[15],简单的说该变换可以用来描述图像坐标与地图坐标的映射关系。而对应的变换矩阵称之为单应性矩阵。具体表现形式为:

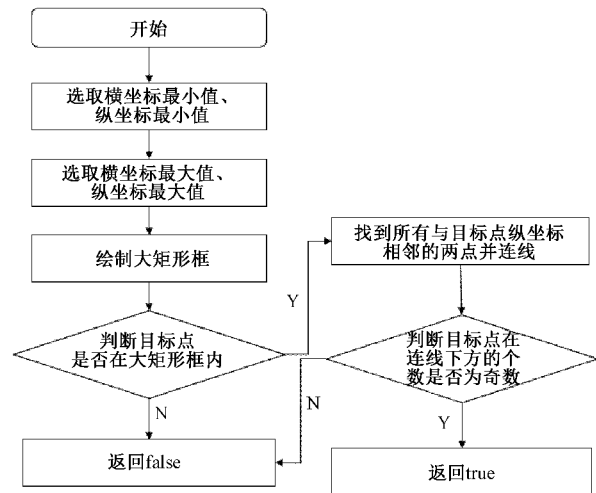


图6 算法流程

$$\mathbf{H}_{3 \times 3} = \begin{bmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ c_0 & c_1 & c_2 \end{bmatrix} \quad (5)$$

则由单应性矩阵将图像坐标与地图坐标的转换公式如下所示:

$$\mathbf{H}\mathbf{X} = \mathbf{X}' \quad (6)$$

将其展开则可以写成下式:

$$\begin{bmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ c_0 & c_1 & c_2 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad (7)$$

其中, \mathbf{X} 为图像坐标矩阵, \mathbf{H} 为单应性矩阵, \mathbf{X}' 为地图坐标矩阵。

则只需要求解出单应性矩阵 \mathbf{H} , 即可以实现图像坐标到地图坐标之间的转换。由于这里使用的是齐次坐标系, 可以将坐标进行任意尺度的缩放, 那么将单应性矩阵 \mathbf{H} 乘上一个非零常数也不会改变等式的结果, 所以可以得出单应性矩阵 \mathbf{H} 只有 8 个自由度。即需要至少 4 对坐标点来确定 \mathbf{H} 的具体值。

确定 \mathbf{H} 的值的具体做法如下所示:

$$\mathbf{A}\mathbf{h} = 0 \quad (8)$$

其中 \mathbf{A} 为 4 对坐标点矩阵, \mathbf{h} 为单应性矩阵的展开式具体形式如下所示:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x'_1 & -y_1x'_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y'_1 & -y_1y'_1 & -y'_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4x'_4 & -y_4x'_4 & -x'_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y'_4 & -y_4y'_4 & -y'_4 \end{bmatrix}_{8 \times 9} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ c_2 \\ c_3 \end{bmatrix}_{9 \times 1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}_{8 \times 1} \quad (9)$$

现在拿 1.3 节中提到的图片举例,其中图 2 中实线框为摄像头拍摄的待定位区域,对应实际地图图像图 3 中实线框。

分别获取对应四边形 4 个顶点坐标如表 1 所示。由上公式代入计算,可以解得单应性矩阵 H 值。

表 1 摄像头图像坐标与地图图像坐标

坐标信息	A 点	B 点	C 点	D 点
摄像头图像坐标	(598,263)	(1 306,280)	(1 696,990)	(226,1 001)
地图图像坐标	(234,213)	(263,110)	(388,140)	(364,243)

按照上述流程,可以解得每个摄像头对应的单应性矩阵。之后可以通过各自的单应性矩阵把每个摄像头中的目标映射到地图上。

实验结果,本实验选取 2 类特征明显的目标并完成定位,一类目标为人,另一类目标为 AGV 小车。测试人员分别在不同摄像头中心位置与同一摄像头不同位置做实验并计算出误差。

3 算法结果及性能分析

不同摄像头中心位置图像坐标及地图坐标和同一摄像头不同位置图像坐标及地图坐标分别如表 2、3 所示,同时根据该坐标计算出实际误差范围。以及误差计算平均值如表 4 所示。

3.1 定位实验及结果分析

本实验在实验室内布设 6 个摄像头,为了方便展示定位

表 2 不同摄像头中心点坐标与误差距离

坐标	摄像头 1	摄像头 2	摄像头 3	摄像头 4	摄像头 5	摄像头 6
图像坐标	(930,658)	(952,548)	(1 096,810)	(933,610)	(948,790)	(943,627)
地图坐标	(338,185)	(462,213)	(552,221)	(648,261)	(721,278)	(810,304)
误差距离/cm	75	52	30	32	25	28

表 3 相同摄像头不同点坐标与误差距离

坐标	采样点 1	采样点 2	采样点 3	采样点 4
图像坐标	(720,433)	(1 347,461)	(469,897)	(1 441,863)
地图坐标	(290,200)	(313,129)	(358,227)	(372,151)
误差距离/cm	30	20	70	70

表 4 平均误差

误差	不同摄像头	相同摄像头
	中心点	不同点
平均误差/cm	40.33	47.5

3.2 目标识别准确度与实时性分析

目标识别算法直接影响了本系统的定位精度,本次实验主要通过如下 2 个方面来进行分析。

1) 准确度

将采集到的视频图像分解成训练集与测试集,在完成数据训练之后通过对测试集的验证结果来评估算法准确度。具体的测试结果如表 5 所示。

表 5 测试集不同标签准确度 %

种类	查全率	查准率
人	98.90	100
agv 小车	97.86	100

2) 实时性

算法实时性也是一个衡量标准,如果实时性差则在

定位系统中有非常大的延迟极大的影响定位准确度。实时性主要分为 2 个部分:1)目标识别算法的实时性;2)服务器传递数据到网页的延迟。目标识别算法的实时性可以通过查看算法每秒处理的图像数量来衡量。在具体的神经网络计算时,将每 2 个摄像头数据打包成一个批处理,因此下表所示的处理数为 2 个摄像头合并后的数量,具体数据如表 6 所示。服务器传递数据到网页的延迟取决于网络环境在本次实验环境中使用的是内部局域网,延迟低于 50 ms。实验结果表明,目标识别的准确度优于 90%,延迟低于 500 ms 用户体验良好无明显卡顿实时性良好。

表 6 各摄像头不同时间点每秒处理图片数 fps

摄像头	时间 节点 1	时间 节点 2	时间 节点 3	平均值
摄像头 1~2	49.70	50.08	49.55	49.78
摄像头 3~4	49.94	50.22	49.41	49.86
摄像头 5~6	50.16	49.86	49.95	49.99

4 结 论

本文针对传统的室内定位系统需要被定位物体携带终端设备或标签这一不足之处,创造性的提出了基于计算机视觉的定位系统。本文研究的核心思路为使用图像坐标到地图坐标的转换从而达到定位的目的。本文的核心技术为通过目标识别算法识别出待定位目标,通过目标点判定算法以及投影变换算法等将待定位物体坐标标记在地图上,核心难点为投影变换算法。在未来的研究中,可以针对如何提高定位精度等问题展开优化。

参考文献

- [1] 张奋,贾小林,姬剑锋,等.北斗系统应用趋势分析[J]. 卫星应用,2021(8):43-47.
- [2] 石欣,梁妙珠,邱雷,等.基于LTE网络传输损耗的定位改进算法[J]. 仪器仪表学报,2018,39(5):57-64.
- [3] 陈浩,李起伟,王子龙.基于改进TDOA在煤矿井下超宽带定位算法的研究[J]. 电子测量技术,2021,44(6):96-102.
- [4] 杨晓倩.基于TOA-AOA数据融合的定位算法研究[J]. 电子测量技术,2020,43(16):104-108.
- [5] 李洪雷,冷文,王安国,等.基于信号强度的改进室内定位算法[J]. 电子测量技术,2016,39(3):136-140,150.
- [6] ZHANG K, SHEN CH, ZHOU Q, et al. A combined GPS UWB and MARG locationing algorithm for indoor and outdoor mixed scenario[J]. Cluster Computing, 2019, 22(3): 5965-5974.
- [7] CUI D, ZHANG Q. The RFID data clustering algorithm for improving indoor network positioning based on LANDMARC technology [J]. Cluster Computing, 2019, 22(3): 5731-5738.
- [8] 杨敏,刘思怡.一种基于先验信息的WiFi室内定位方法[J]. 电子测量与仪器学报,2020,34(6):163-168.
- [9] 陈伟.蓝牙定位技术的探讨及测试[J]. 国外电子测量技术,2020,39(10):143-146.
- [10] 谢宏兰.室内停车场车辆定位及导航系统的设计[J]. 电脑知识与技术,2020,16(12):219-220,224.
- [11] 叶盛清.视觉定位技术在行业场景中的创新应用[J]. 中国安防,2021(6):104-107.
- [12] 叶靓玲,李伟达,郑力新,等.结合目标检测与特征匹配的多目标跟踪算法[J]. 华侨大学学报(自然科学版), 2021,42(5):661-669.
- [13] 赵梓杉,秦玉英,李刚,等.基于深度学习的目标检测算法综述[J]. 汽车实用技术,2021,46(17):207-209.
- [14] 豆世豪.改进YOLOv3的道路场景目标检测方法[J]. 电脑知识与技术,2021,17(23):94-96.
- [15] TIAN J H, WU Y L, CAI Y H, et al. A novel mosaic method for spaceborne ScanSAR images based on Homography matrix compensation [J]. Remote Sensing, 2021, 13(15): 2866-2866.

作者简介

周宇辉,工学硕士,主要研究方向为计算机视觉与模式识别。

E-mail:shu_zzh@shu.edu.cn

朱晓强,博士,副教授,主要研究方向为计算机视觉与智能图形图像分析。

E-mail:xqzhu@shu.edu.cn

曾丹,博士,教授,主要研究方向为计算机视觉与模式识别。

E-mail:dzeng@shu.edu.cn