

基于视觉物体识别的抗差岭估计定位算法

徐昊玮* 廉保旺 邹晓军 岳哲 吴鹏

(西北工业大学电子信息学院 西安 710072)

摘要: 基于视觉物体识别的室内定位算法是一种新型的室内定位解决方案, 算法通过物体检测、位置匹配、定位方程解算等步骤确定用户位置。然而, 受到单目相机视域较小和物体检测算法精度较低的影响, 根据检测物体测距信息而构成的定位方程存在严重的病态性, 极大降低了算法的定位成功率和定位精度。因此, 该文提出一种抗差岭估计定位解算算法, 通过引入岭参数将无偏估计变为有偏估计, 实现均方误差最小约束条件下的最优位置估计, 并利用迭代选权降低了质量较差的观测量对定位精度的影响。实验结果表明, 与OLS (Ordinary Least Square), LM (Levenberg-Marquardt)和RR (Ridge Regression)算法相比, 该文提出的抗差岭估计定位解算算法能够有效提高基于视觉物体识别的室内定位方法的成功率和精度。

关键词: 室内定位; 岭估计; 深度学习

中图分类号: V249.32+8

文献标识码: A

文章编号: 1009-5896(2018)10-2453-08

DOI: 10.11999/JEIT170876

Visual Objects Detection Based Robust Ridge Regression Indoor Localization Method

XU Haowei LIAN Baowang ZOU Xiaojun YUE Zhe WU Peng

(School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: The indoor vision positioning algorithm based on object detection is a novel indoor positioning solution, which determines the position of the user through the process of objects detection, position matching, location equation calculation, etc. However, limited by the field-of-view of monocular camera and objects detection accuracy, the localization equation, which is constructed according to the detected objects range information, is seriously ill conditioned. Therefore, this paper proposes a novel localization method based on an improved robust ridge regression estimation, which reduces the influence of the lower accurate observations by iterative weight selection. The experimental results show that compared with Ordinary Least Square (OLS), Levenberg-Marquardt (LM) and Ridge Regression (RR) algorithms, the proposed improved robust ridge regression estimation algorithm can effectively improve the positioning success rate and positioning accuracy of the object detection based indoor navigation method.

Key words: Indoor localization; Ridge regression; Deep learning

1 引言

室内定位技术作为定位技术在室内环境的延续, 弥补了传统卫星定位技术的不足, 因而有着良好的应用前景^[1-3]。近年来, 随着标配单目摄像头的智能手机逐渐普及, 无需应用附加设备的室内视觉定位系统为室内导航应用提供了一种低成本、有效的解决方案。现有的视觉定位的主要方法是:

在区域中的不同位置、不同角度拍摄图像, 将图像与拍摄时相机的坐标、姿态一起存储于服务器, 当接收到用户拍摄图像后, 服务器通过多种特征点提取与比较方式, 将用户图像与存储在数据库中的图像相匹配, 得到用户位置。首尔大学应用视觉地标匹配实现了单目摄像头的即时定位与地图构建(SLAM)^[4]; 韩国产业技术大学^[5]和罗切斯特理工大学^[6]分别应用不同类型的图像特征实现了0~3 m精度的室内定位。但是, 构建本地数据库工作量大、提取特征信息计算量大、特征信息受外部因素影响严重等问题仍制约着视觉导航的精度和可用性。

受到近年来逐渐成熟的深度学习算法的启发, 文献^[7]开创性地提出了一种新型的基于物体的室内

收稿日期: 2017-09-18; 改回日期: 2018-07-23; 网络出版: 2018-07-26

*通信作者: 徐昊玮 xuhaoweiself@aliyun.com

基金项目: 国家自然科学基金(61473308, 61771393)

Foundation Items: The National Natural Science Foundation of China (61473308, 61771393)

视觉定位方法。文献[8]在此基础上通过提高物体检测精度、改进地图构建方式进一步对该定位方法进行了完善。与基于特征信息的视觉定位方法不同,该方法是根据图像中检测出的物体信息计算得到用户位置。相比于基于特征信息的视觉定位算法,基于物体的定位算法对硬件要求简单,且无需大量的布置工作,因此有着较好的发展前景。然而,当应用最小二乘法进行定位解算时^[8],一方面受到单目相机视域较小的影响,识别物体间的几何结构较差,容易引起法矩阵奇异化。另一方面受到物体识别算法的精度影响,物体伪距观测信息存在野值点。以上两方面因素会导致定位精度恶化,甚至出现在位置解算过程中迭代结果发散的情况。LM算法^[9]和RR算法^[10]虽然能在一定程度上改善了因共线带来的收敛率低的问题,但仍受到观测量野值点的影响。针对以上问题,本文首先将岭估计算法引入到位置解算中,在均方根误差最小约束条件下,将无偏估计变为有偏估计,缓解了因矩阵奇异而带来的定位发散问题,并在此基础上提出了一种迭代选权的算法,算法根据每次迭代后残差的大小自适应调整权值矩阵,逐渐降低质量较差的观测量对定位结果的影响,提高了基于物体的视觉室内定位方法的可用性和精确性。

2 基于物体的视觉定位算法

基于物体的室内视觉定位方法的工作流程如下:

(1)用户使用智能手机拍摄包含周围物体的照片,并上传服务器;(2)在服务器端,应用训练后R-CNN(Regional-Convolutional Neural Network)识别照片中的物体^[5];(3)提取图像中物体的空间关系,并与标记有物体实际位置的电子地图进行匹配,得到可能的用户位置;(4)根据物体信息建立非线性方程组,应用高斯牛顿算法,以步骤(3)中得到的位置解作为初始点,迭代求解方程组,最终将收敛得到的位置作为最终结果发送用户。以上步骤可以主要涉及两种算法:物体识别算法与定位算法。

2.1 物体识别算法

应用R-CNN算法作为物体识别算法,与传统

卷积神经网络算法(CNN)仅能为整张照片分类不同,R-CNN算法可以将图像分割成区域,因此可以对图像中多个对象进行目标提取和分类。R-CNN算法最初由Girshick等人^[11]在ILSV2014上提出,而本文所应用的是Girshick等人^[12,13]于2015年提出改进版本Faster-R-CNN,新版本已将单张照片的识别速度提升至0.3 s。Faster-R-CNN算法的输出是检测物体类别、覆盖该物体的矩形框和代表检测成功率的分数 s ,其中 $s \in [0, 1]$ 。应用以上Faster-R-CNN输出数据,结合已知的标记有物体位置的地图和物体的实际高度信息,设计了基于物体的视觉定位算法^[7,8]。

2.2 基于物体的视觉定位算法

视觉定位算法包括粗定位和精定位两个步骤。粗定位^[6]是通过将用户拍摄图像中提取出的物体关系与地图中物体实际空间关系相匹配而得到的。为了完成粗定位,首先需要在电子地图中的各个位置,以不同角度产生用于描述真实物体相对空间关系(例如,上、下、左、右)的拓扑图,即MG(Map Graph)图。现以图1为例说明MG图的产生方法:将虚拟相机放置在图1(a)所示的地图上虚线的交点处,可以根据虚拟相机的视域以及标记的物体位置得到关于物体的虚拟点,将物体的位置投影到图中横虚线所示的相机图像平面上,便可以得到如图1(b)所示的MG图。之后,以相同的拓扑图产生方式为用户拍摄照片中检测出的物体产生IG(Image Graph)图,如图1(c)所示。最后将IG图与生成的所有MG图根据规则进行匹配,相似度最高的MG图的位置和方向被选取为拍摄用户的粗定位、定向结果。粗定位的定位精度一般为5 m左右。

为了进一步优化定位精度,在得到粗定位结果与物体匹配结果后,进一步应用图像中物体的尺寸信息计算出物体伪距,建立并解算定位方程,实现精定位。

图2为单目相机成像原理图,其中 u 为镜头至物体所在垂直于光轴平面的距离,简称像面距, v 为像距, l 为物体中心至光轴的垂直距离, l_c 为成

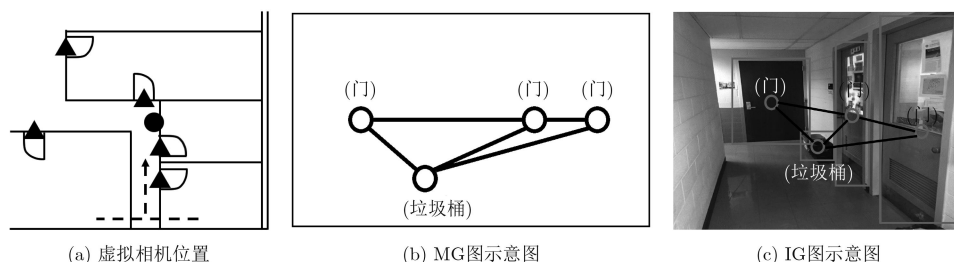


图1 MG与IG生成图示例

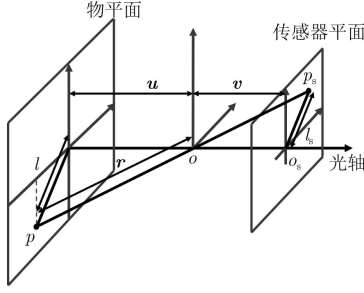


图2 单目相机成像原理

像物体中心至图像中心的距离。在精定位算法中，根据以上成像模型可以得到基于物体的定位模型如式(1)：

$$r^2 - \alpha \cdot v^2 = \rho \quad (1)$$

其中， $\rho = (l_c \cdot h_r \cdot c \cdot \Delta r / n_v)^2$ ， $\alpha = (h_r \cdot c \cdot \Delta r / n_v)^2$ ， Δr 为单位尺寸下CMOS中像素个数， n_v 为物体在图像中高度所占像素数， c 为数字变焦系数， h_r 为物体的实际高度。对于在同一张照片中检测出的物体， v 相等，但由于如IOS, Android等手机操作系统中没有直接获取 v 的API函数，因此 v 为未知量。由于镜头距离物体的实际距离与 v 相关，因此 ρ 被称为检测物体的伪距，它是一个与图像中物体位置与大小相关的变量。对于一张检测出 n 个物体的照片，可以建立如式(2)所示的非线性定位方程组：

$$\left. \begin{aligned} (x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2 - \alpha_1 \cdot v^2 &= \rho_1 - \varepsilon_1 \\ (x_2 - x)^2 + (y_2 - y)^2 + (z_2 - z)^2 - \alpha_2 \cdot v^2 &= \rho_2 - \varepsilon_2 \\ &\vdots \\ (x_n - x)^2 + (y_n - y)^2 + (z_n - z)^2 - \alpha_n \cdot v^2 &= \rho_n - \varepsilon_n \end{aligned} \right\} \quad (2)$$

其中， (x, y, z) 为未知的用户位置， (x_i, y_i, z_i) 为图像中检测出的第 i 个物体的位置， ε_i 为伪距测量噪声。式(2)可以进一步被改写为式(3)所示的向量形式：

$$\mathbf{f}(\mathbf{x}) = \boldsymbol{\rho} - \boldsymbol{\varepsilon} \quad (3)$$

式中， $\mathbf{x} = [x, y, z, v]$ 为未知量，因此当一张照片能够检测出的物体数大于等于4个时，如式(2)所组成的方程组正定或超定，有唯一解。需要说明的是，受到物体检测算法和物体空间分布的影响，实际应用中，检测出3个物体的情况较多，为了提高算法的可用性，对于此种情况，我们将用户所在位置的高程数据设置为常量，即 $z = h$ ，因此当检测物体数量大于等于3时就可以对位置进行求解。

高斯牛顿法是求解非线性方程组的常用方法，算法通过在 $\mathbf{x} = [x^{(k)}, y^{(k)}, z^{(k)}, v^{(k)}]^T$ 点进行1阶泰勒展开将非线性方程线性化，之后应用迭代算法逐渐逼近方程解。如式(4)所示为高斯牛顿算法的迭代方程：

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \Delta \mathbf{x}^{(k)} \quad (4)$$

$$\Delta \mathbf{x}^{(k)} = \mathbf{H} \mathbf{J}^T(\mathbf{x}^{(k)}) \mathbf{W} \mathbf{e}(\mathbf{x}^{(k)}) \quad (5)$$

其中， $\mathbf{x}^{(k)}$ 为第 k 次迭代时得到的方程解， $\mathbf{e}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) - \boldsymbol{\rho}$ 为式(2)中所示非线性方程的残差函数， $\mathbf{J}(\mathbf{x}^{(k)})$ 为雅克比矩阵，如式(6)所示。

$$\mathbf{J}(\mathbf{x}^{(k)}) = \begin{pmatrix} \frac{x_1 - x^{(k)}}{r_1(\mathbf{x}^{(k)})} & \frac{y_1 - y^{(k)}}{r_1(\mathbf{x}^{(k)})} & \frac{z_1 - z^{(k)}}{r_1(\mathbf{x}^{(k)})} & \frac{2a_1 v^{(k)}}{r_1(\mathbf{x}^{(k)})} \\ \frac{x_2 - x^{(k)}}{r_2(\mathbf{x}^{(k)})} & \frac{y_2 - y^{(k)}}{r_2(\mathbf{x}^{(k)})} & \frac{z_2 - z^{(k)}}{r_2(\mathbf{x}^{(k)})} & \frac{2a_2 v^{(k)}}{r_2(\mathbf{x}^{(k)})} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{x_n - x^{(k)}}{r_n(\mathbf{x}^{(k)})} & \frac{y_n - y^{(k)}}{r_n(\mathbf{x}^{(k)})} & \frac{z_n - z^{(k)}}{r_n(\mathbf{x}^{(k)})} & \frac{2a_n v^{(k)}}{r_n(\mathbf{x}^{(k)})} \end{pmatrix} \quad (6)$$

其中， $r_i(\mathbf{x}^{(k)}) = \sqrt{(x_i - x^{(k)})^2 + (y_i - y^{(k)})^2 + (z_i - z^{(k)})^2}$ 。

\mathbf{W} 为权值矩阵，由于Faster-R-CNN算法输出的物体检测分数 s_i 与物体的标记准确度成正比，而标记准确度进一步影响了伪距的测量误差，且同样满足 $\varepsilon_i \propto s_i$ ，因此设置权值矩阵为

$$\mathbf{W} = \text{diag} [s_1^2, s_2^2, \dots, s_n^2] \quad (7)$$

2.3 定位性能分析

基于式(2)的定位误差协方差矩阵如式(8)：

$$\text{Cov}([\varepsilon_x, \varepsilon_y, \varepsilon_z, \varepsilon_v]^T) = \mathbf{H} \delta_\rho^2 \quad (8)$$

其中， δ_ρ^2 为伪距测量误差。当应用OLS估计时有

$$\mathbf{H}_{\text{OLS}} = [\mathbf{J}^T \mathbf{W} \mathbf{J}]^{-1} \quad (9)$$

此时，定位算法的精度和可用性与以下两方面因素有关。

2.3.1 物体的几何分布 矩阵 $[\mathbf{J}^T \mathbf{W} \mathbf{J}]^{-1}$ 完全取决于检测出的物体个数以及它们相对于相机的几何分布。当检测出物体数量多，且位置分散时矩阵 $\mathbf{J}^T \mathbf{W} \mathbf{J}$ 会有较好的非奇异性，此时矩阵的特征根一般不会接近0，因此矩阵 $\mathbf{J}^T \mathbf{W} \mathbf{J}$ 求逆后元素值相对较小，测量误差被放大成定位误差的程度较低。然而在实际应用中，受到相机感光元件尺寸的影响，取景的视角范围受到限制，仅有位于相机正前方一定视域范围内的物体能够被相机记录。因此，即便图中检测出的物体数量较多且识别、标记相对准确，但由于物体集中，使得回归自变量之间存在着近似线性关系，即复共线性，最终导致定位精度较差。另外，在检测物体数量较少的情况下，复共线性会影响矩阵的正定，严重时甚至会导致在求解非线性方

程组过程中迭代发散。

2.3.2 伪距测量误差 在式(8)中,除 \mathbf{H} 外,伪距测量误差 ϵ 的方差 δ_ρ^2 也是影像定位精度的重要因素,且 δ_ρ^2 越大,定位误差方差越大。物体的伪距测量误差是几部分测量误差的总和。物体检测算法是影响物距测量误差的一个重要因素,标记矩形框高度误差会影响物体 n_i 的量测值;标记矩形框中心与实际物体中心误差,则会影响 l_c 的量测值。此外,图像失真、物体高度不确定性等都会为物体伪距测量值带来误差。以上这些误差是由许多复杂的、确定的或半确定的因素而组成的,因此很难进行精确的建模。伪距测量误差不但会影响定位结果的精度,当某个物体有较大的伪距测量误差产生时,即便物体数量多、几何结构好,仍有大概率导致迭代发散。

为解决上述两因素给定位性能造成的影响,本文引入并改进了抗差岭估计以取代OLS估计进行非线性方程组求解。

3 基于改进的抗差岭估计的室内定位算法

岭估计算法最早由Hoerl和Kennard提出^[4],是一种改进的最优化估计算法,它通过引入岭参数 $\mathbf{K} = \text{diag}[k_1, k_2, \dots, k_n]$ 将OLS算法的无偏估计变为有偏估计,从而获得更为符合观测量、更可靠的回归结果。也由于岭参数的引入, \mathbf{H} 矩阵的复共线性得到极大缓解,因此它对病态数据的拟合要强于OLS算法。如式(10)所示为岭估计算法中 \mathbf{H} 矩阵的表达式。

$$\mathbf{H}_{RR} = [\mathbf{J}^T \mathbf{W} \mathbf{J} + \mathbf{K}]^{-1} \quad (10)$$

著名的LM算法^[15]也是岭估计的一种,算法根据每次迭代后的残差与设定门限的比较结果对岭参数 \mathbf{K} 进行调整。LM算法虽然在一定程度上加快了求解非线性方程的收敛速度,但由于对岭参数 \mathbf{K} 的调整有一定随机性,因此估计精度没有得到较大提高。

3.1 岭参数选取

为了获得最优的岭参数 \mathbf{K} ,文献^[14]将岭估计均方误差对岭参数 \mathbf{K} 求导后令其等于0,经过推导得到均方误差最小条件下的岭参数取值如式(11):

$$k_i = \frac{\delta^2}{\Delta x_i^2} \quad (11)$$

其中, δ^2 表示式(2)中误差的方差, Δx_i 为 $\Delta \mathbf{x}$ 中的第 i 个元素。由于实际应用中 δ^2 和 Δx_i 均为未知参量,因此需要对它们进行估计,本文应用文献^[16]中提出的基于几何均值的岭估计参数估计量,如式(12)所示。

$$k_i = \hat{\delta}^2 / \left(\prod_{j=1}^n \Delta \hat{x}_j^2 \right)^{1/n} \quad (12)$$

其中, $\hat{\delta}^2$ 为LS解残差的均方估计值, $\Delta \hat{x}_j$ 为LS解中的第 j 个元素, $\hat{\delta}^2$ 和 $\Delta \hat{x}_j$ 均为对应未知量的无偏估计。

3.2 抗差估计算法

岭参数 \mathbf{K} 的引入使得由检测物体几何结构分布而产生的 \mathbf{H} 矩阵复共线性得到缓解,为了进一步减小伪距测量误差带来的影响,不同于传统的OLS算法中 \mathbf{W} 矩阵在初始确定后保持不变,本文提出了一种迭代选权的方法。抗差岭估计(Robust Ridge Regression, RRR)算法的 \mathbf{H} 矩阵如式(13)所示。

$$\mathbf{H}_{RRR} = [\mathbf{J}^T \widetilde{\mathbf{W}} \mathbf{J} + \mathbf{K}]^{-1} \quad (13)$$

$\widetilde{\mathbf{W}}$ 为权值矩阵,在每次解算后算法根据残差大小自适应调整观测值的权值,如式(14)所示为迭代选权公式:

$$\tilde{w}_i^{(k)} = \begin{cases} \frac{s_i}{\max(\theta, e(x^{(k-1)}))}, & n > 3 \\ s_i, & n = 3 \end{cases} \quad (14)$$

其中, n 为观测伪距数量, $\tilde{w}_i^{(k)}$ 表示第 k 次迭代时,迭代权矩阵 $\widetilde{\mathbf{W}}^{(k)}$ 中对角线上第 i 个元素, $\widetilde{\mathbf{W}}^{(k)}$ 中其他非对角线元素均为0。式中设置当且仅当 $n > 3$ 时,权值矩阵才开始迭代,这是由于当矩阵刚好正定时,即 $n = 3$ 时,迭代权的加入可能会使得某个观测量权值过低,从而造成方程组欠定,直接导致收敛失败。另外, θ 的引入防止了当残差等于0或接近于0权值过大的问题。

基于以上描述与分析,建立抗差岭估计(RRR)位置解算方案如下:

步骤1 计算式(2)的OLS估计和OLS估计残差;

步骤2 应用OLS估计和OLS估计残差分别根据式(12)和式(14)确定岭参数和迭代权值,并将得到的估计值代入式(13)中计算 \mathbf{H}_{RRR} 矩阵;

步骤3 将 \mathbf{H}_{RRR} 矩阵代入式(5)中,并应用式(4)进行解算,判断迭代是否收敛,如果没有收敛则返回步骤1进行下轮迭代。

4 实验测试与性能分析

岭参数的选取是抗差岭估计算法中最为重要的步骤之一,本文采用了岭参数的几何估计值,为了验证该岭参数估计值的准确性,首先设计了仿真实验。在仿真实验中,在水平 $\pm 60^\circ$ 、垂直 $\pm 80^\circ$ 的视野范围内,随机生成代表物体的点,之后根据点的位置分别计算到各点的距离,并人为添加均值为0、方差一定的误差作为距离观测值,利用点的位置与距离观测值建立非线性方程组,最后应用RRR算法计算岭参数并求解方程组,在计算过程中,记录 \mathbf{K} 的估计值 $\widehat{\mathbf{K}}$ 与 \mathbf{K} 的真实值 \mathbf{K}' ,并通过式(15)计算 $\widehat{\mathbf{K}}$ 与 \mathbf{K}' 的相似度。

$$\eta = \frac{1}{1 + \sqrt{\sum (\widehat{\mathbf{K}}(i) - \mathbf{K}'(i))}} \quad (15)$$

其中, $\eta \in [0, 1]$, 且值越大说明相似度越大。在实验中, 生成点的个数被分别设置为3, 4, 5, 6, 7, 而误差噪声方差则被分别设置为0.1 m, 0.3 m, 0.5 m, 每组参数共进行100次实验。图3(a)展示了仿真实验中相似度的结果, 图中的每个点均为100次实验的平均值, 从图中可以看出估计值 $\widehat{\mathbf{K}}$ 的准确性受到仿真点与距离测量值精度的影响, 且点数越多, 距离测量值误差越小, 得到的 $\widehat{\mathbf{K}}$ 值越准确。以上现象是由于较小的测距误差和更多的点测距信息都会使得 $\widehat{\mathbf{K}}$ 的值更加准确, 从而提高 $\widehat{\mathbf{K}}$ 的估计精度。

为了进一步说明 \mathbf{K} 的准确性对RRR算法的影响, 在仿真实验中, 通过人为地将 \mathbf{K} 的真实值 \mathbf{K}' 添加误差得到满足一定相似度的 $\widetilde{\mathbf{K}}$, 并应用 $\widetilde{\mathbf{K}}$ 进行解算, 最终比较解算的性能。在实验中, 距离误差噪声方差被设定为0.1 m, 生成点的个数被分别设置为3, 5, 7, 添加误差后的 $\widetilde{\mathbf{K}}$ 与真实值的相似度被分别设置为1.0, 0.9, 0.8, 0.7, 0.6, 每组参数同样进行100次实验。图3(b)–图3(d)分别展示了非线性方程解算过程的收敛速率、收敛精度和收敛率受 \mathbf{K} 值的影响, 从图中可以看出随着 $\widetilde{\mathbf{K}}$ 与 \mathbf{K}' 的相似度逐渐降低, 非线性方程解算所用的迭代周期数以及收敛精度都有较大幅度提升, 而收敛率也有较大程度

的降低。当相似度仅有60%、仿真点个数为5的情况下, 平均需要15个周期才能完成迭代解算, 而收敛精度仅为0.1 m, 收敛率则下降为85%。

为验证算法在定位解算中的效果, 本文在室内环境中进行了实验测试, 实验测试的范围包括两条尺寸分别为2.4 m×33.0 m和2.4 m×52.0 m的走廊以及一个7.0 m×7.0 m的房间。在测试环境中分布着包括饮水机、垃圾桶、桌子、柜子、窗户、灯等多种室内物体, 部分物体用不同的标志标记于图4的2维地图中。实验前对环境每个物体的高度和位置进行了精确的测量并将其标记于电子地图上, 应用该电子地图, 以2.0 m位置间隔和45.0°方向间隔生成了包含458张MG图本地数据库。选用Iphone6手机在室内环境中拍摄了500张包含物体的照片, 其中420张经过人为物体标记后用于Faster-R-CNN识别算法的训练, 另外80张照片用于定位算法的验证。图5(a)、图5(b)分别定量地展示了Faster-R-CNN对80张测试照片中物体的检测率与伪距测量精度。在80张照片中共包含有402个物体, 其中距离拍摄相机10.0 m内的物体数为303个, 大于10.0 m的物体数为99个。Faster-R-CNN算法共识别出289个物体, 总识别率达到72%, 根据物体标记框计算得到的伪距平均误差为0.7 m。识别算法共识别出10.0 m范围内的物体数为251, 识别率83%, 而对于10.0 m外的物体共识别出

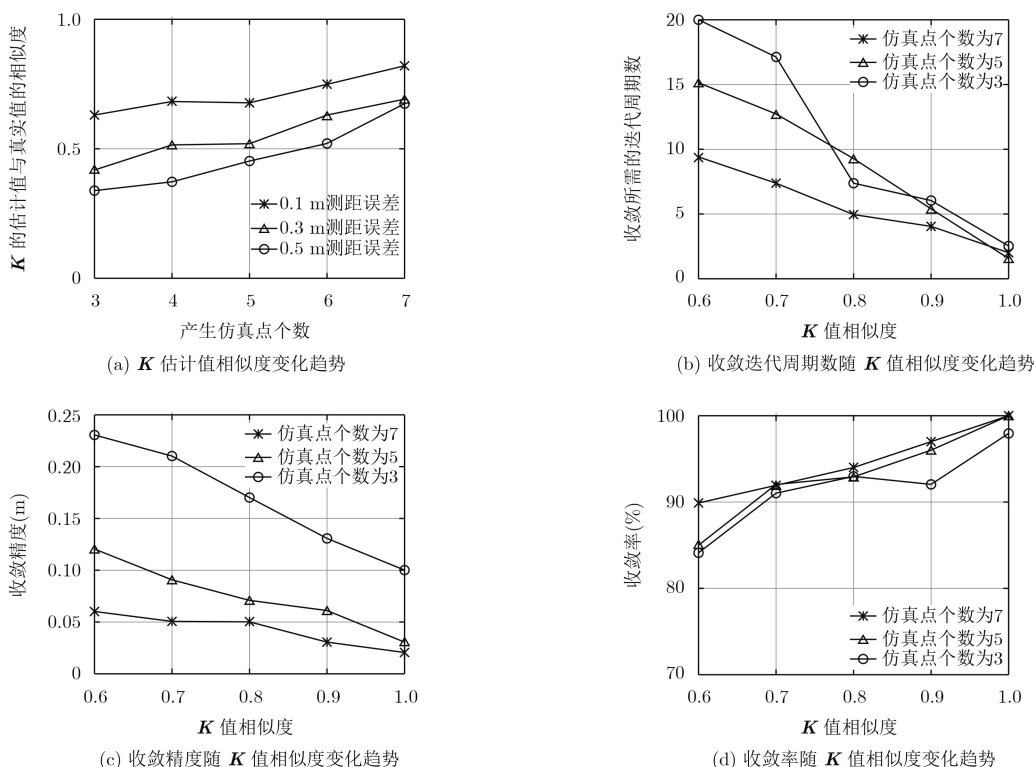


图3 岭估计参数 \mathbf{K} 对非线性方程解算的影响

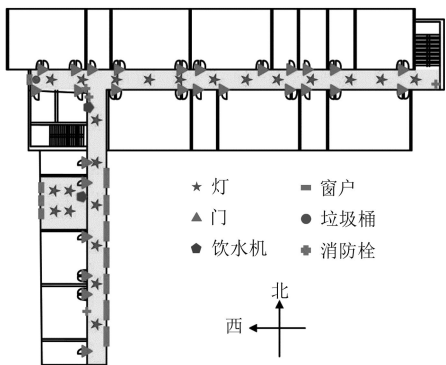


图4 实验环境示意图

45个，识别率仅为46%。物体的伪距误差也有相似趋势，除柜子外（拍摄到的柜子均为10.0 m外），其他所有物体在10.0 m以内的伪距测量精度都明显优于10.0 m外。

在检测得到物体信息后，先应用粗定位算法计算出可能的粗定位结果和对应的物体匹配结果，之

后分别应用OLS算法、LM算法、RR算法和本文提出的改进的RRR算法进行精定位求解，求解的未知量为 (x, y, v) ，用户的高程被设置为1.5 m，因此满足检测物体数大于等于3就可以进行位置求解。另外，4种算法的迭代收敛终止条件都设置为 $\Delta \mathbf{x}^{(k)} < 0.05$ ，当 $k > 30$ 时仍没有满足迭代收敛终止条件时则视为迭代发散。

图6展示了一种典型的伪距观测情况下的定位示例，其中图6(a)所示为得到的物体检测和标记结果，图中白框中为识别分数 s ；图6(b)–图6(e)为应用不同定位解算方法得到的定位结果比较。实验中，收敛性的比较是通过设置不同初始迭代位置而进行的，即：拥有较好收敛性的解算方法总能找到合适的路径收敛到最优位置，因此受初始位置影响较小。图6(b)–图6(e)中黑色点表示应用粗定位算法得到的位置解，其中黑色实心点表示精定位算法应用该点作为初始迭代点能够收敛到合理位置，而

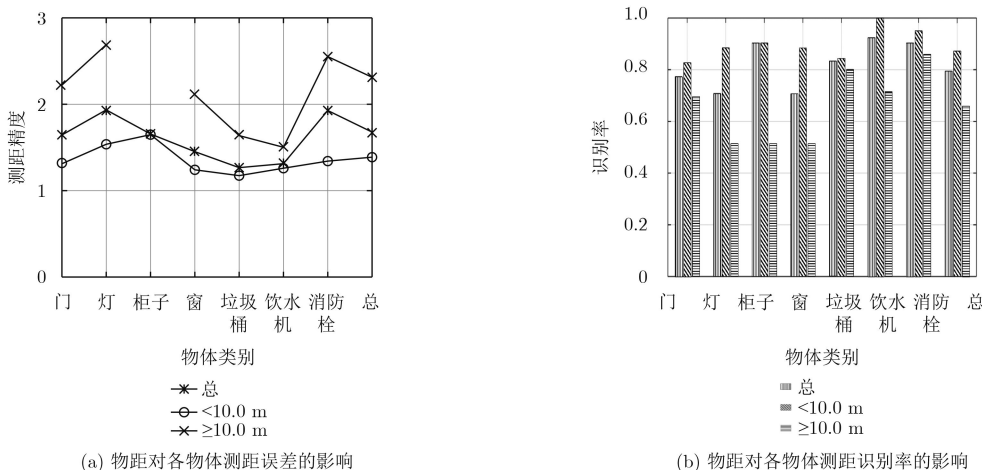


图5 Faster-R-CNN对于物体的检测示例以及对物体的识别率与伪距测量精度

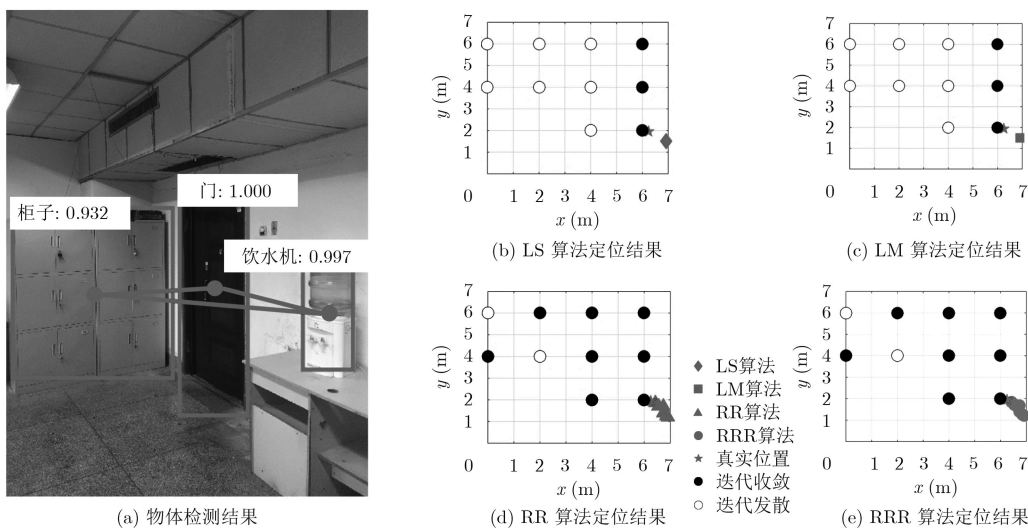


图6 物体几何结构较差时的定位示例

黑色空心点则代表精定位算法发散，或收敛于地图外某处；各精定位算法的最终定位结果用不同灰色记号标记于各个子图中；用户真实位置用五角星表示。

在图6所示的定位示例中，Faster-R-CNN在照片中共检测出3个物体，虽检测结果正确且标记准确，但物体在图像中互相临近，因此几何结构较差。应用粗定位可以计算得到如图中黑色空心点所示的10个粗定位结果，将各粗定位结果作为起始点应用不同估计算法进行迭代解算，结果如图6(b)–图6(e)所示。OLS算法和LM算法都仅能根据其中3个初始位置迭代收敛，经算术平均后的3维定位精度均为0.8 m；由于检测物体数为3，在RRR算法中并没有进行迭代选权，所以在此定位示例中RR算法与RRR算法的定位结果一致：能够收敛的初始位置数提升为8，定位误差降低到0.6 m。另外，在图中可以看出，相比于OLS和LM算法，RR算法和RRR算法的定位结果并不集中于同一点，这是由于RR算法和RRR算法在迭代过程中持续对H矩阵的改变而造成的，LM算法虽然也会改变H矩阵，但仅在迭代过程中出现残差增大时才对H矩阵进行调整。

表1为应用不同解算方法对80张照片的定位结果，在表中分别比较了各算法的定位成功率、定位精度、收敛率和迭代收敛次数。从表中可以看出，在定位成功率与定位精度方面，随着物体检测数量的增长，OLS、LM和RR 3种算法的定位误差与定

位精度都有较为明显的下降趋势。虽然在理论上，更多的识别物体，意味着有更高的概率会得到较好的几何关系，但是一旦某个物体的伪距存在较大的测量误差，则可能会引起解算无法收敛，而增加的物体检测数量和物体距离的增大都增加了产生较大测量误差的可能性，所以对于传统解算方法来说，更多的识别物体反而使位置解算结果恶化。但是RRR算法中的迭代选权机制使得算法可以随着迭代的进行控制权值将误差较大的观测量剔除，因此物体数量的增加给RRR解算过程提供了更多的观测量选择。表中显示当检测到3个物体时，RRR算法的平均定位误差为1.3 m，定位成功率82.5%。RRR算法1.6 m的总平均定位误差和67.5%的总定位成功率也优于RR算法、LM算法和LS算法。在收敛率方面，RRR算法61.1%的总收敛率明显高于其余3种算法。而在迭代次数方面，LM算法通过调节K可以达到梯度下降法的收敛速度，因此LM算法收敛最快。而RRR算法虽然也对K进行调节，但却是以均方误差最小为条件，而且在迭代开始时，由于误差较大，导致计算的各观测量权值普遍较低，因此缩小了步长，影响了收敛速度，相比于其他3种算法，RRR算法的平均迭代次数为9，是4种算法中最高的。

5 结束语

针对基于物体识别的室内定位方法中成功率

表 1 4种定位解算方法的定位结果比较

		检测出物体数	3个	4个	5个	6个	7个	总
		照片数	47	22	7	3	1	80
OLS算法	定位成功率(%)	61.7	50.0	42.8	0.0	0.0	0.0	53.8
	定位误差(m)	1.8	2.1	2.4	–	–	–	1.9
	收敛率(%)	42.1	37.5	23.1	0.0	0.0	0.0	39.6
	迭代次数	8	8	9	–	–	–	8
LM算法	定位成功率(%)	66.0	68.2	42.8	33.0	0.0	0.0	62.5
	定位误差(m)	1.6	2.0	2.0	1.2	–	–	1.7
	收敛率(%)	45.7	42.5	34.6	9.0	0.0	0.0	43.3
RR算法	定位成功率(%)	70.2	63.6	71.4	66.0	0.0	0.0	67.5
	定位误差(m)	1.5	1.9	1.9	1.0	–	–	1.6
	收敛率(%)	56.0	50.9	53.8	36.4	0.0	0.0	53.7
	迭代次数	7	7	9	9	–	–	7
RRR算法	定位成功率(%)	80.9	81.8	85.7	100.0	100.0	100.0	82.5
	定位误差(m)	1.4	1.3	1.3	0.5	0.6	0.6	1.3
	收敛率(%)	62.9	62.1	57.7	45.5	40.0	40.0	61.1
	迭代次数	8	9	11	13	12	12	9

低、定位精度差的问题,本文提出了一种抗差岭估计定位解算算法,实验结果表明,新算法能够有效提高基于物体的室内视觉定位方法的稳定性和精确性。在之后的研究中,计划对基于物体识别的室内定位方法进行如下扩展和改进:(1)进一步提高物体识别算法的识别率和识别精度;(2)选用更加丰富的图像数据库训练识别模型,使得识别算法无需重复训练就能应用于其他场景;(3)与Wi-Fi定位结合,将MG图与Wi-Fi信号强度一同加入到本地“指纹”数据库中,实现更加准确、鲁棒性更强的“粗定位”结果。

参考文献

- [1] KOPPANYI Z and TOTH C K. Indoor ultra-wide band network adjustment using maximum likelihood estimation[J]. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2014, 2(1): 31–35. doi: [10.5194/isprannals-II-1-31-2014](https://doi.org/10.5194/isprannals-II-1-31-2014).
 - [2] 陈增强, 国峰, 张青. 基于模糊神经网络建模的RFID室内定位算法[J]. *系统科学与数学*, 2014, 34(12): 1438–1450. CHEN Zengguo, GUO Feng, and ZHANG Qing. An RFID indoor location algorithm based on fuzzy neural network model[J]. *Journal of Systems Science & Mathematical Sciences*, 2014, 34(12): 1438–1450.
 - [3] 周牧, 唐云霞, 田增山, 等. 基于流形插值数据库构建的WLAN室内定位算法[J]. *电子与信息学报*, 2017, 39(8): 1826–1834. doi: [10.11999/JEIT161269](https://doi.org/10.11999/JEIT161269). ZHOU Mu, TANG Yunxia, TIAN Zengshan, et al. WLAN indoor localization algorithm based on manifold interpolation database construction[J]. *Journal of Electronics & Information Technology*, 2017, 39(8): 1826–1834. doi: [10.11999/JEIT161269](https://doi.org/10.11999/JEIT161269).
 - [4] CHAE H, PARK C, and SONG J B. Sliding window based landmark extraction for indoor 2D SLAM using forward-looking monocular camera[C]. *Control, Automation and Systems (ICCAS)*, Busan, Korea, 2015: 336–341.
 - [5] SHIM J H and CHO Y I. A Mobile robot localization using external surveillance cameras at indoor[J]. *Procedia Computer Science*, 2015, 56: 502–507. doi: [10.1016/j.procs.2015.07.242](https://doi.org/10.1016/j.procs.2015.07.242).
 - [6] LU Guoyu, YAN Yan, SEBE N, et al. Indoor localization via multi-view images and videos[J]. *Computer Vision and Image Understanding*, 2017, 32: 121–132. doi: [10.1016/j.cviu.2017.05.003](https://doi.org/10.1016/j.cviu.2017.05.003).
 - [7] XU Haowei, KOPPANYI Z, TOTH C K, et al. Indoor localization using region-based convolutional neural network[C]. *Proceedings of the 2017 International Technical Meeting of ION*, Monterey, USA, 2017: 314–323.
 - [8] TOTH C K, KOPPANYI Z, XU Haowei, et al. Localization using region-based convolution neural network: A comparison study[C]. *Proceedings of the 10th International Conference of Mobile Mapping Technology*, Cairo, Egypt, 2017: 134–142.
 - [9] UEDA K and YAMASHITA N. On a global complexity bound of the Levenberg-Marquardt method[J]. *Journal of Optimization Theory and Applications*, 2010, 147(3): 443–453.
 - [10] CHEN B H, DA R, and TEKIN I. Location determination using weighted ridge regression[P]. USA, Patent, US6587692, 2003-7-1.
 - [11] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, USA, 2014: 580–587.
 - [12] GIRSHICK R. Fast R-CNN[C]. *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 2015: 1440–1448.
 - [13] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031).
 - [14] HOERL A E and KENNARD R W. Ridge regression: Biased estimation for nonorthogonal problems[J]. *Technometrics*, 1970, 12(1): 55–67.
 - [15] MORÉ J J. *The Levenberg-Marquardt Algorithm: Implementation and Theory*[M]. Berlin, Heidelberg: Springer, 1978: 105–116.
 - [16] KIBRIA B M G. Performance of some new ridge regression estimators[J]. *Communications in Statistics-Simulation and Computation*, 2003, 32(2): 419–435. doi: [10.1081/SAC-120017499](https://doi.org/10.1081/SAC-120017499).
- 徐昊玮: 男, 1990年生, 博士生, 研究方向为卫星导航、惯性导航、LiDAR导航、组合导航、深度学习。
 廉保旺: 男, 1962年生, 教授, 研究方向为导航、通信系统、雷达。
 邹晓军: 男, 1984年生, 博士生, 研究方向为卫星导航、跟踪环路设计。
 岳 哲: 男, 1992年生, 博士生, 研究方向为卫星导航、多源组合导航。
 吴 鹏: 男, 1986年生, 博士, 研究方向为卫星导航、多径抑制。