

doi: 10.7690/bgzd.2024.05.008

基于四叉树法和 PROSAC 算法改进的视觉 SLAM 技术

杜 根, 张志安

(南京理工大学机械工程学院, 南京 210094)

摘要: 为解决在同时定位与地图构建 (simultaneous localization and mapping, SLAM) 的前端进行特征点匹配时, 随机抽样一致法 (random sample consensus, RANSAC) 存在的迭代次数高、实时性较差、鲁棒性不稳定等问题, 提出一种基于四叉树法和渐进一致采样法 (progressive sample consensus, PROSAC) 算法融合改进的图像匹配算法。实现四叉树法+PROSAC 算法的误匹配剔除算法, 在 EuRoC 数据集上对改进后的 ORB-SLAM2 算法进行实验。结果表明: 相比于 ORB-SLAM2 系统, 该算法在 Vicon Room 1 03 数据集上总体绝对轨迹误差平均值减小了 39.28%, 总体相对位姿误差减小了 35.45%, 具有更高的建图精度。

关键词: 四叉树编码; 特征点匹配; PROSAC 算法; SLAM

中图分类号: TP391.9 **文献标志码:** A

Improved Visual SLAM Technology Based on Quadtree Method and PROSAC Algorithm

Du Gen, Zhang Zhian

(School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: In order to solve the problems of RANSAC, such as high number of iterations, poor real-time performance and unstable robustness in the front end of simultaneous localization and mapping (SLAM), an improved image matching algorithm based on the fusion of quadtree method and PROSAC algorithm is proposed. The mismatching elimination algorithm of quadtree method + PROSAC algorithm is implemented, and the improved ORB-SLAM2 algorithm is tested on EuRoC data set. The results show that compared with ORB-SLAM2 system, the proposed algorithm reduces the average absolute trajectory error by 39.28% and the relative pose error by 35.45% on Vicon Room 1 03 dataset, and has higher mapping accuracy.

Keywords: quadtree coding; feature point matching; PROSAC algorithm; SLAM

0 引言

同步定位与地图构建 (SLAM) 是移动机器人实现导航运动功能的重要组成部分之一。所谓的 SLAM 技术^[1]是指机器人通过在自身搭载各种传感器器件, 并利用这些传感器获取周围环境的观测信息, 在进行大量复杂的信息处理运算后, 完成对自身在环境中的定位与对周边环境的建图任务, 是机器人学中一个热门研究方向^[2]。

在视觉 SLAM 前端中^[3], 传感器采集的数据会受到噪声的干扰, 导致对前后 2 帧图像特征点的误匹配, 从而影响相机的定位。当前最常用的特征点匹配算法有 SIFT、SURF、BRISK、ORB 等, 文献[4]利用改进的 Harris 算法检测具有仿射变换的特征点, 优化了匹配的精确度。文献[5]在图像受到一定缩放、旋转和光照的情况下, 改进了 SIFT 算法的匹配准确度^[6]。文献[7]提出了 ORB 算法, 该算法具有运行速度快, 且对图像旋转具有较强的鲁棒性。

文献[8]剔除了样本集中不符合图像特性的部分外点, 解决了传统的随机抽样一致法 (RANSAC) 算法迭代次数多、匹配精度较差的问题。

上述方法解决了视觉 SLAM 前端中图像匹配的一些问题, 但也存在一定的局限性^[9]。SIFT 算法需要构造高维的特征描述向量^[10], 计算过程复杂且需要较长运行时间; ORB 算法特征点具有局部聚集性^[11], 无法利用到整张图像上的信息; RANSAC 算法迭代次数不稳定, 要得到准确的模型参数只能通过不断提高迭代次数, 时间开销过大。

针对上述问题, 笔者提出一种基于四叉树法^[12-13]和渐进一致采样法 (PROSAC) 算法^[14]融合改进的图像匹配算法, 在利用四叉树法均匀化特征点分布后, 使用 PROSAC 算法进一步降低图像的误匹配对。笔者在 EuRoC 数据集^[15]上进行仿真实验分析, 确定该方案有效地降低了图像的误匹配概率, 并显著地提升了视觉 SLAM 系统的建图精度。

收稿日期: 2024-01-23; 修回日期: 2024-02-22

第一作者: 杜 根(1998—), 男, 江西人, 硕士。

1 二叉树法

二叉树也被称作 Q 树(Q-Tree)，常用于在储存稀疏数据、图像信息处理等领域。二叉树将每个节点划分为 4 个子节点，在 2 维空间中即将该空间划分为 4 个区域，并将区域中的信息存入该二叉树节点中。二叉树通常由单个节点开始生长，当有多个对象被添加至该单节点中时，该单节点就会被划分为 4 个子节点，若某个被划分的节点中仍存在多个对象，二叉树将持续这个划分过程，直到二叉树达到一定的深度，或者是满足了某种要求后才会停止分割。

图像的特征点一般分布在灰度变化明显的角点、边缘等位置附近，且由关键点和描述子 2 部分组成。主流的特征提权算法有如 Harris、SIFT、SURF、ORB 等，笔者从尺度不变性、旋转、稳定性和速度 4 方面综合考量上述特征提取算法的性能，ORB 在稳定性和速度方面具有非常明显的优势，且对尺度和旋转的变化不明显，所以笔者使用 ORB 进行研究。笔者使用的关键点检测框架如图 1 所示。

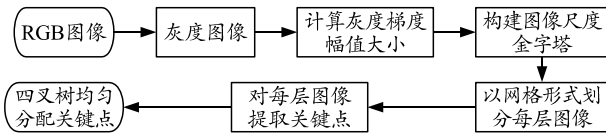


图 1 关键点检测框架

假设图像的灰度函数为 $f(x, y)$ ，则在图像中的某个像素点 $p(x, y)$ 在该位置的梯度可表示为：

$$\nabla f(x, y) = [G_x, G_y]^T = \left[\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y} \right]^T \quad (1)$$

式中 G_x 和 G_y 为像素点 p 在 x 和 y 轴方向上的梯度，可用差分方式近似微分为：

$$G_x = f(x, y) - f(x-1, y); \quad (2)$$

$$G_y = f(x, y) - f(x, y-1). \quad (3)$$

计算像素点 $p(x, y)$ 的梯度幅值为：

$$\text{mag}[\nabla f(x, y)] = \sqrt{G_x^2 + G_y^2}. \quad (4)$$

对图像中的每个像素点重复上述操作过程，最后对所有像素点的梯度幅值进行求和并计算平均值：

$$\text{mag}[\nabla f]_{\text{average}} = \frac{1}{n} \sum_{(x, y) \in f} \text{mag}[\nabla f(x, y)]. \quad (5)$$

提取图像中关键点数量的公式为：

$$N = (k / \text{mag}[\nabla f]_{\text{average}}) C. \quad (6)$$

式中： k 为根据实验来设定的比例系数，笔者选用 k 值为 10； C 以一般 ORB 特征点检测中提取的关键

点数量作为参考，本文中选取为 1 000，通过上述方法动态的调整提取过程中关键点的数量，能有效提升算法在纹理不丰富的场景中所提取到的特征点数量。

构建如图 2 所示图像金字塔，并对金字塔的层数进行设定。

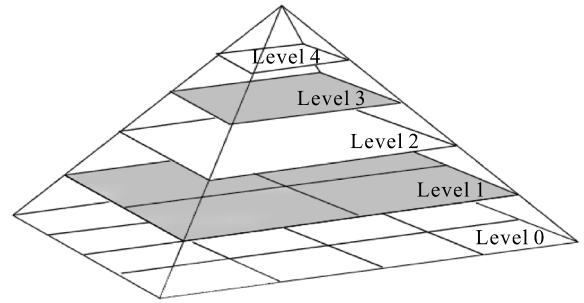


图 2 图像金字塔

在图像中选取一个像素点 p ，并假设其灰度值为 I_p ，以像素点 p 为圆心，以 3 个像素点大小的长度为半径选取 p 在此圆周上的 16 个像素点，如图 3 所示。以圆周上位于 p 正上方的像素点作为起始点，分别假设 16 个像素点灰度值为 I_{p_i} ， $i=1, 2, 3, \dots, 16$ 。当圆周上超过 N 个像素点的灰度值全部 $< I_p - \epsilon$ 或全部 $> I_p + \epsilon$ 时，由于中心点 p 与其邻域像素点灰度值差别过大，则判断 p 为关键点。通常 N 取 9 或者 12，笔者经过实验验证，当 N 选取 9 时识别准确率更高；因此，本文中 N 的取值为 9，阈值 ϵ 参考 I_p 选取 $0.2I_p$ 。

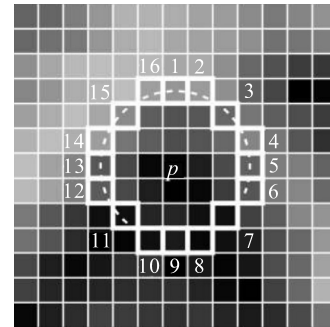


图 3 人工选取关键点

接着以像素点 p 为几何中心，选取其周边一个小的图像块 S ，定义图像块 S 的矩为：

$$m_{pq} = \sum_{(x, y) \in S} x^p y^q I(x, y), p, q = \{0, 1\}. \quad (7)$$

通过矩可以找到图像块的质心 C ：

$$C = (m_{10}/m_{00}, m_{01}/m_{00}). \quad (8)$$

连接像素点 p 与质心 C ，得到一个方向向量 \overline{OC} ，定义特征点方向为：

$$\theta = \arctan(m_{01}/m_{10}). \quad (9)$$

对图像金字塔的每个分层进行网格划分操作，

通过调整阈值 ε 的大小，确保每个网格内至少包含 5 个关键点。以四叉树形式重新划分图像中的关键点分布位置，若划分后的某个区域中所含有的关键点数量仍大于 4 个时，需对其进行重新处理，直到所有区域都不满足该条件。最后在筛选最终保留的关键点时，仅选取 Harris 响应值最大的关键点，将除此以外的关键点全部丢弃。图 4 为改进前后的关键点分布情况。



(a) 改进前

(b) 改进后

图 4 改进前后关键点分布对比

2 PROSAC 算法

PROSAC 是 RANSAC 的一种优化版本，相比于 RANSAC 在整个集合中均匀的进行采样，PROSAC 是从持续增大的最佳对应点集合中进行采样的，假设出发点为：具有更高相似性的数据点更有可能是真正的内点。PROSAC 在采样阶段就过滤掉了部分不相关幅度较大的点集，所以它的匹配精确度更高、迭代时间更短。

记 N 个数据点的集合为 U_N ，评价函数为 q ，根据评价函数 q 可将 U_N 内的数据点排序为：

$$u_i, u_j \in U_N : i < j \Rightarrow q(u_i) \geq q(u_j). \quad (10)$$

式中评价最好的 n 个数据点集合记作 U_n 。

从数据 U_N 中的采样点集合记作 M ，其元素个数为 $|M| = m$ ，采样点集合的评价函数定义为其所有元素评价函数值中最小的评价函数值：

$$q(M) = \min_{u_i \in M} q(u_i). \quad (11)$$

PROSAC 先对数据点进行排序，然后在其数据点子集中选取评价函数最高的子集进行随机采样，这个被选中的子集称为假设生成集，且其中包含的元素应该是持续增加的，所以在使用 PROSAC 算法的过程中，需处理确定假设生成集的大小和确定采样过程的停止策略 2 个问题。

2.1 确定假设生成集的大小

生长函数用于确定假设生成集的大小，需要有

2 个作用：1) 不能过分依赖评价函数的排序作用；2) 不能对所有数据赋以相同的权重。本文中用 $P\{u_i\}$ 表示 u_i 为正确点的概率，并和评价函数之间取一个单调性假设：

$$i < j \Rightarrow q(u_i) \geq q(u_j) \Rightarrow P(u_i) \geq P(u_j). \quad (12)$$

用 $\{M_i\}_{i=1}^{T_N}$ 表示 T_N 次采样点集合的序列，且每次采样都是从 N 个数据点中选取其中的 m 个，即 $M_i \in U_N$ ，用 $\{M_{(i)}\}_{i=1}^{T_N}$ 表示根据评价函数排序之后的序列：

$$i < j \Rightarrow q(M_{(i)}) \geq q(M_{(j)}). \quad (13)$$

用 T_n 表示 $\{M_i\}_{i=1}^{T_N}$ 中所有元素都来源于 U_N 的采样点集合 M_i 的平均个数：

$$T_n = T_N \binom{n}{m} / \binom{N}{m} = T_N \prod_{i=0}^{m-1} \frac{n-i}{N-i}; \quad (14)$$

$$\frac{T_{n+1}}{T_n} = \prod_{i=0}^{m-1} \frac{n+1-i}{N-i} \prod_{i=0}^{m-1} \frac{N-i}{n-i} = \frac{n+1}{n+1-m} \Rightarrow T_{n+1} = (n+1)T_n / (n+1-m). \quad (15)$$

T_n 个采样点集合只包含 U_n 的数据点，则 T_{n+1} 个采样点集合只包含 U_{n+1} 的数据点。由于 $U_{n+1} = U_n \cup u_{n+1}$ ，所以 $T_{n+1} - T_n$ 个采样点集合包括一个数据点 u_{n+1} 和 $m-1$ 个来源于 U_n 的数据点。当 n 从 m 一直取值到 $N-1$ 时，即可高效且随机地得到采样点集合 $M_{(i)}$ 。由于 T_n 的值一般不为整数，令 $T'_n = 1$ ，且有：

$$T'_{n+1} = T'_n + \lceil T_{n+1} - T_n \rceil. \quad (16)$$

式中 $\lceil \cdot \rceil$ 括号为 \geq 括号中值的最小整数，所以最终定义生长函数为：

$$g(t) = \min\{n : T'_n \geq t\}. \quad (17)$$

式中 n 为采样次数 $\geq t$ 次的最小 n 值。

2.2 停止策略

当集合 U_n 中的内点数 I_n 满足以下 2 个条件时，PROSAC 算法就会停止：

1) 极大性：存在一个解 U_n ，这个解中存在比 I_n 中更多的内点，且此种情况在 k 次采样后仍未发现的概率 $< \eta_o$ (一般设为 5%)。

2) 非随机性： n 个数据点的类内点恰好也是任意一个错误模型内点的概率 $< \psi$ (一般设为 5%)。

极大性定义了保证解置信度的采样次数，从 U_n 中采样得到的 m 个点都来自于类内点的概率为：

$$P_{I_n} = \binom{n}{m} / \binom{N}{m} = \prod_{j=0}^{m-1} \frac{I_n - j}{n - j} \approx \varepsilon_n^m \quad (18)$$

式中： I_n 为 U_n 中类内点的数量； $\varepsilon_n = I_n/n$ 为类内点数的比例，则 k 次采样全是外点的概率为：

$$\eta = (1 - P_{I_n})^k \quad (19)$$

在给定了阈值 η_0 之后，采样次数满足如下条件：

$$k_{n^*}(\eta_0) \geq \log(\eta_0) / \log(1 - P_{I_{n^*}}) \quad (20)$$

非随机性可以防止由外点产生的模型带来的误差，随机内点集合的元素个数呈二项分布：

$$P_n^R(i) = \beta^{i-m} (1 - \beta)^{n-i+m} \binom{n-m}{i-m} \quad (21)$$

式中 β 为与错误模型相符合的点不在采样集合中的概率，且该采样点集合包含一个类外点，则最少类内点数 I_n^{\min} 为：

$$I_n^{\min} = \min \{ j : \sum_{i=j}^n P_n^R(i) < \psi \} \quad (22)$$

最终 n^* 的大小是在满足 $I_{n^*} \geq I_n^{\min}$ 的条件下最小化 $K_{n^*}(\eta_0)$ 得到的。

3 实验分析

3.1 实验环境

采用 EuRoC 数据进行单目和双目场景的实验验证，所选用的数据为 Vicon Room 1 03 序列，该数据时长 104.655 s，轨迹长度为 78.982 m，所使用的平台环境为 AMD R7 5800H CPU、RTX 3060 GPU 和 16 GB 内存的笔记本。笔者使用绝对轨迹误差 (absolute trajectory error, ATE) 和相对位姿误差 (relative pose error, RPE) 作为系统最终误差评判的标准：

1) 绝对轨迹误差：利用官方所公布的相机真实运动轨迹与 SLAM 系统运行所估计的轨迹进行差值计算，这种对比能表现出全局轨迹一致性。

2) 相对位姿误差：用于计算 2 个相同时间戳上相机位姿的真实值与 SLAM 系统的估计值之间每隔一段时间位姿变化量之间的差值，也可以理解为位姿的真实值与估计值的实时比较。

3.2 对照实验

3.2.1 四叉树法实验

使用 OpenCV 中自带的 ORB 算法和四叉树法分别进行特征提取，该对照实验共做了 30 组，由于

实验结果基本相似，本文中只展示其中一组对照实验结果，如图 5 所示。



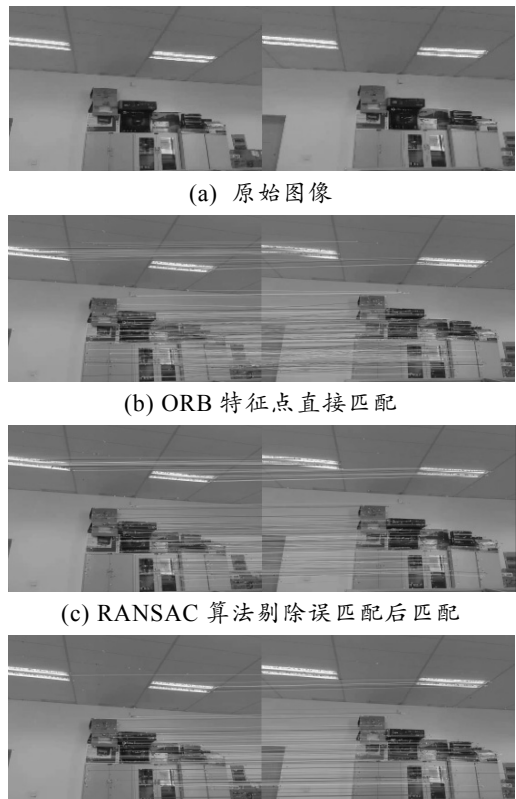
(a) ORB 算法提取的特征点 (b) 四叉树法提取特征点

图 5 ORB 与四叉树提取特征点分布

可以看到，传统 ORB 提取算法在图像中提取关键点时，容易出现在角点、边缘等位置，具有聚集性，不能充分利用整张图像的信息，无法给后续位姿估计提供良好的初始值。四叉树法通过网格划分，从每个区域中选取 Harris 响应值最大的角点作为最终的关键点，使得关键点均匀分布到整张图像中。四叉树法解决了 ORB 算法关键点分布过于集中的问题，充分利用了整张图像的信息。

3.2.2 PROSAC 算法实验

为验证四叉树与 PROSAC 算法结合的有效性，使用 OpenCV 工具进行对比实验，实验结果如图 6 所示。



(a) 原始图像 (b) ORB 特征点直接匹配 (c) RANSAC 算法剔除误匹配后匹配 (d) PROSAC 算法剔除误匹配后匹配

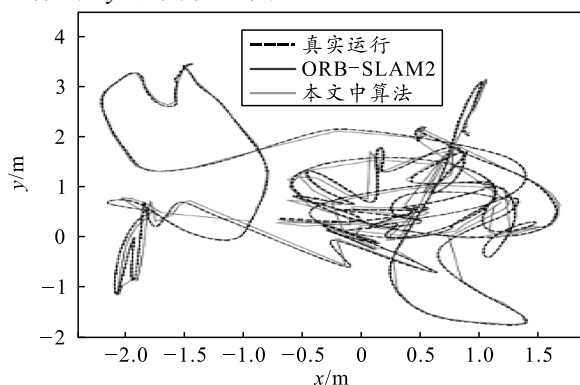
图 6 实验结果

对比实验结果可知，RANSAC 算法和 PROSAC 算法的匹配效果都优于 ORB 直接匹配的效果，直接匹配不仅耗时还存在大量的错误匹配对。而 PROSAC 算法效果比 RANSAC 算法更好，其内点数更少，提高了匹配命中率，且缩短了匹配时间。

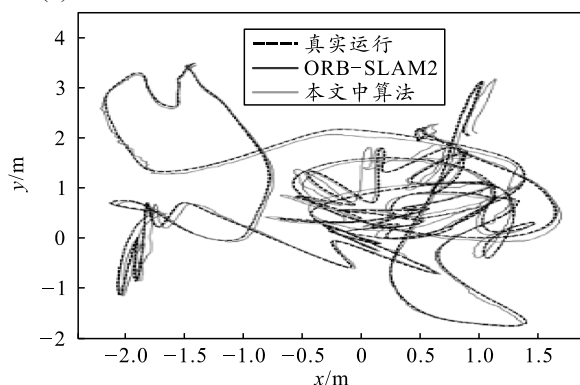
3.3 实验结果分析

本文中利用 EVO 工具对绝对轨迹误差和相对轨迹误差进行分析，EVO 可以用于评估 SLAM 系统输出的轨迹的精度，自动生成轨迹、误差等信息的图表形式。选取其中 5 项作为误差评价指标：mean 表示平均误差，median 表示误差中位数，rmse 表示均方根误差，sse 表示误差平方和，std 表示标准差。

如图 7 所示，图 7(a)是使用 ORB-SLAM2 系统与利用本文中方法改进的 ORB-SLAM2 算法在单目模式下运行的 xy 平面轨迹图，所使用的数据集为 Vicon Room 1 03 序列；图 7(b)则是两者在双目模式下运行的 xy 平面轨迹图。



(a) 单目模式 ORB-SLAM2 与本文中算法轨迹



(b) 双目模式 ORB-SLAM2 与本文中算法轨迹

图 7 不同模式与本文中算法轨迹

可以看出两者都能很好地拟合真实轨迹，但利用本文中算法改进后的轨迹更贴合真实轨迹。

选择对双目模式下的绝对轨迹误差和相对位姿误差进行分析对比，表 1 和 2 分别是 ORB-SLAM2 系统与本文中系统在 ATE 和 RPE 上的对比分析。

表 1 绝对轨迹误差 m

系统	Mean	Median	Rmse	Sse	Std
ORB-SLAM2	0.079	0.071	0.090	0.842	0.043
本文中系统	0.052	0.036	0.024	0.685	0.034

表 2 相对位姿误差 m

系统	Mean	Median	Rmse	Sse	Std
ORB-SLAM2	0.040	0.033	0.049	5.102	0.028
本文中系统	0.026	0.020	0.035	0.404	0.016

从实验结果可以看出，本文中算法在 Vicon Room 1 03 数据集上总体绝对轨迹误差平均值减小了 39.28%，总体相对位姿误差减小了 35.45%，即本文中算法在对真实轨迹的拟合上拥有更高的精确度，这对促进 SLAM 系统的最终建图与定位效果有着极大提升。

4 结论

笔者针对 ORB 提取特征点具有不均匀性以及 RANSAC 剔除误匹配速率与精度较差的缺点，提出使用四叉树提取特征点与 PAOSAC 算法剔除误匹配的改进 SLAM 技术。笔者在 ORB-SLAM2 系统的基础上，融入改进思想，在 EuROC 公开数据上进行了单、双目模式实验。经过实验验证可知：PROSAC 算法剔除误匹配对的精度远高于 RANSAC 算法，在保证 ORB-SLAM2 系统整体运行速率的情况下，相比于 ORB-SLAM2 系统，本文中算法在 Vicon Room 1 03 数据集上总体绝对轨迹误差平均值减小了 39.28%，总体相对位姿误差减小了 35.45%，有效地降低了 SLAM 系统的建图精度。

参考文献：

- [1] 权美香, 朴松昊, 李国. 视觉 SLAM 综述[J]. 智能系统学报, 2016, 11(6): 768-776.
- [2] 邱凯昌, 万文辉, 赵红颖. 视觉 SLAM 技术的进展与应用[J]. 测绘学报, 2018, 47(6): 84-93.
- [3] 高翔, 张涛. 视觉 SLAM 十四讲: 从理论到实践[M]. 北京: 电子工业出版社, 2017: 133-182.
- [4] 朱凯, 刘华峰, 夏青元. 基于单目视觉的同时定位与建图算法研究综述[J]. 计算机应用研究, 2018, 35(1): 1-6.
- [5] BENNASAR M, HICKS Y, SETCHI R. Feature selection using joint mutual information maximization[J]. Expert Systems with Applications, 2015, 42(22): 8520-8532.
- [6] FISCHLER M A, BOLLES R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Communications of the ACM, 1981, 24 (6): 381-395.
- [7] MURARTAL R, TARDÓS J D. ORB-SLAM2: an

open-source slam system for monocular, stereo, and RGB-D cameras[J]. IEEE Transactions on Robotics, 2017, 33(5): 1255-1262.

[8] 席志红, 王洪旭, 韩双全. 基于 ORB-SLAM2 系统的快速误匹配剔除算法与地图构建[J]. 计算机应用, 2020, 40(11): 3289-3294.

[9] 王巍. 惯性技术研究现状及发展趋势[J]. 自动化学报, 2013, 39(6): 723-729.

[10] 茹淑慧, 王红旗, 唐浩. 基于自适应多特征匹配的移动机器人 SLAM 研究[J]. 传感器与微系统. 2019, 38(12): 33-35, 39.

[11] FORSTER C, PIZZOLI M, SCARAMUZZA D. SVO: fast semi-direct monocular visual odometry[C]//Proc of IEEE

International Conference on Robotics and Automation. Piscataway, NJ: IEEE Press, 2014: 15-22.

[12] 张美玉, 王洋洋, 侯向辉, 等. 基于 ORB 和改进的 RANSAC 图像拼接算法[J]. 计算机科学, 2019, 46(S2): 294-298.

[13] 张阳, 许钢, 张星宇. 一种改进的 ORB 特征点匹配算法[J]. 重庆工商大学学报(自然科学版), 2018, 179(3): 73-78.

[14] 李伦. 基于 PROSAC 算法的室内定位算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2018: 36-50.

[15] BURRI M, NIKOLIC J, GDHI P, et al. The EuRoC micro aerial vehicle datasets[J]. International Journal of Robotics Research, 2016, 35(10): 1157-1163.

(上接第 36 页)

图 6 所示为变电站巡检人员不规范佩戴安全帽检测部分结果展示。可以看出，所提模型可以有效检测变电站巡检人员是否携带安全装备(如安全帽)，从而发出预警，降低安全事故发生概率。



图 6 变电站巡检人员不规范佩戴安全帽检测部分结果展示

4 结论

笔者对变电站巡检人员安全行为监控进行分析，建立一种基于改进 YOLO 轻量化的变电站巡检人员不规范佩戴安全帽检测模型。首先，对基础 YOLO 模型相关概念进行介绍。其次，详细设计了改进 YOLO 轻量化模型的结构。最后，基于实验分析以及部分结果展示，验证了所提模型的有效性。该模型为变电站巡检安全隐患排查提供了一定借鉴作用。

未来工作可对巡检时其他不规范行为进行增量补充，包括持续收集未佩戴安全手套、未进入指定安全区域等数据集，从而增加不规范行为检测类别。此外，可对嵌入式、巡检机器人等领域进行研究，如引入无人机、无人车等技术提高变电站巡检的可靠性及效率，进一步降低巡检人力成本。

参考文献:

[1] 李默煊. 变电站运维管理中人工智能技术的应用[J]. 光源与照明, 2022(11): 125-127.

[2] 赵爽, 王文娟, 钱纹. 云南末端电网安全运行风险分析及对策[J]. 云南电力技术, 2020, 48(6): 39-41.

[3] 毛龙灿, 杨南. 基于大数据背景的皮革人才培养优化研究[J]. 中国皮革, 2021, 50(9): 38-41.

[4] 杨涛. 互联网时代下皮革行业电商物流体系研究分析[J]. 中国皮革, 2021, 50(8): 82-85.

[5] 钟建棚, 余少锋, 廖崇阳, 等. 基于云计算的电力设备智能监测系统[J]. 云南师范大学学报(自然科学版), 2022, 42(3): 37-41.

[6] 朱建宝, 马青山, 俞鑫春, 等. 基于 PSPNet 的变电站人员巡检危险行为检测[J]. 自动化与仪表, 2021, 36(11): 35-39.

[7] 朱建宝, 许志龙, 孙玉玮, 等. 基于 OpenPose 人体姿态识别的变电站危险行为检测[J]. 自动化与仪表, 2020, 35(2): 47-51.

[8] 李金鑫, 徐润, 陈家辉, 等. 基于视频识别和 UWB 定位的变电站作业现场安全识别方法[J]. 电气技术与经济, 2022, 30(6): 123-126.