

基金项目论文

基于改进 ORB 算法的视觉里程计特征匹配方法

殷新凯¹, 茅 健¹, 周玉凤¹, 陈晓平²

(1. 上海工程技术大学 机械与汽车工程学院, 上海 松江 201620; 2. 宁波工程学院 机械学院, 浙江 宁波 315016)

摘 要: 为了满足排爆机器人自主导航要求, 利用视觉 SLAM 技术, 为排爆机器人路径规划做准备。本文主要对用特征点法搭建的前端视觉里程计进行优化, 将 FAST 角点法与 Harris 角点法进一步结合, 同时引入 SIFT 算法关键点检测, 然后与 BRIEF 描述子结合, 用暴力匹配算法完成不同图像的匹配。通过仿真实验验证算法的可行性, 为后续的路径规划奠定基础。

关键词: 视觉里程计; ORB 算法; 特征点检测; 特征匹配

中图分类号: TP242.3 **文献标识码:** A **DOI:** 10.3969/j.issn.1003-6970.2020.04.014

本文著录格式: 殷新凯, 茅健, 周玉凤, 等. 基于改进 ORB 算法的视觉里程计特征匹配方法[J]. 软件, 2020, 41 (04): 57-62

Visual Odometry Feature Matching Method Based on Improved ORB Algorithm

YIN Xin-kai¹, MAO Jian¹, ZHOU Yu-feng¹, CHEN Xiao-ping²

(1. School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China;
2. School of Mechanical Engineering, Ningbo University of Technology, Ningbo 315016, China)

【Abstract】: In order to meet the requirements of the autonomous navigation of the explosive ordnance disposal robot, the Visual SLAM technology is used to prepare for the path planning of the explosive ordnance disposal robot. This paper mainly optimizes the front-end visual odometry constructed by the feature point method, further combines the FAST corner method and the Harris corner method, and introduces the key point detection of the SIFT algorithm, and then combines with the BRIEF descriptor to quickly approximate the nearest neighbor algorithm. Complete the matching of different images. The feasibility of the algorithm is verified by simulation experiments, which lays a foundation for subsequent path planning.

【Key words】: Visual odometry; ORB algorithm; Feature point detection; Feature matching

0 引言

SLAM 是同步定位与地图构建, 是指携带传感器的移动设备, 在部分未知或全部未知的环境中, 在运动过程中建立环境的模型, 同时估计自身运动^[1]。一般有视觉 SLAM 和激光 SLAM 两种, 激光 SLAM 技术已经相对成熟, 广泛应用到移动式机器人中, 而视觉 SLAM 大多还在理论研究阶段。本文利用视觉 SLAM 技术, 研究排爆机器人自主导航定位与路径规划可行性, 对视觉 SLAM 前端视觉里程计中特征匹配进行优化, 为排爆机器人完成作业任

务奠定基础。

1 视觉 SLAM

SLAM 问题是为了同时解决移动设备的两个问题: 一是确定自身所处位置; 二是了解自身所在环境的信息。SLAM 问题实质上就是一个估计状态问题, 移动设备通过自身传感器获得位置信息, 估计自身位置和地图环境。选择相机作为机器人的主要传感器, 运用 SLAM 的方法, 就是视觉 SLAM。我们往往用以下数学模型来视觉 SLAM 的过程, 在一段时间($t_0 \sim t_n$)内, 机器人自身位置用 x_1, x_2, \dots, x_n 表

基金项目: 浙江省自然科学基金(批准号: LY20A020004宁波工程学院)

作者简介: 殷新凯, 男, 研究生, 主要研究方向: 计算机视觉与图像处理; 茅健, 男, 教授, 主要研究方向: 精度检测与质量控制、先进制造技术、智能物流装备与系统; 周玉凤, 女, 副教授, 主要研究方向: 机械制造及其自动化; 陈晓平, 男, 教授, 主要研究方向: 机械疲劳寿命与检测。

示, 设定环境中存在一些路标, 记做 y_1, y_2, \dots, y_m , 那么机器人通过自身传感器获得位置信息就被分成两步, 一是运动, 主要通过机器人自身位置变化表示机器人运动; 二是观测, 通过机器人感知路标来表示机器人观测。我们可以用数学模型表示为:

$$x_n = f(x_{n-1}, u_n, w_n) \quad (1)$$

上式中, u_n 是机器人传感器采集的数据, w_n 是噪声, 方程 f 是运动方程, 表示第一步。第二步的方程表示为:

$$z_{n,j} = h(y_j, x_n, v_{n,j}) \quad (2)$$

式中 y_j 表示机器在 x_n 处观测到的路标, $v_{n,j}$ 是观测中的噪声, 方程 h 是观测方程。上述两个式中 x, y, z 都可根据自身携带的不同传感器来用具体的参数信息来表示, 例如常用的平面坐标 (x, y) , 转角 θ 等, 视觉 SLAM 一般都是通过相机来观测。我们可以将视觉 SLAM 总结为:

$$\begin{cases} x_n = f(x_{n-1}, u_n, w_n) \\ z_{n,j} = h(y_j, x_n, v_{n,j}) \end{cases} \quad (3)$$

式 (3) 是视觉 SLAM 的状态估计方程, 求解这一方程之前的工作就是前段视觉里程计的内容, 之后就是后端优化的过程。

状态估计方程的求解, 一般按照运动和观测方程是否为线性, 噪声是否服从高斯分布分为线性与非线性和高斯与非高斯系统^[3]。线性高斯系统是比较简单的状态估计问题, 此类问题时比较理想的情况, 可以由卡尔曼滤波器求出, 而非线性非高斯系统是实际情况中出现较多的情况, 此类问题一般由扩展卡尔曼滤波和非线性优化去求出。视觉 SLAM 成熟的今天, 我们更倾向用图优化方法进行状态估计^[2]。

在明确状态估计方程的求解后, 视觉 SLAM 需要考虑的就是回环检测, 其目的是解决位置估计随时间漂移产生的误差问题。

视觉 SLAM 的整个过程可以概括为传感器信息读取、视觉里程计、后端优化、回环检测、建立地图五步。其中视觉里程计又称为前端, 是状态估计的过程, 而后端优化是状态估计的求解和优化过程。本文主要对前端里程计中的特征提取和匹配算法进行优化, 提高匹配精确率, 更好的匹配排爆机器人的导航定位过程, 视觉 SLAM 框架如图 1 所示。

视觉里程计目前分为特征点法和光流法, 两者的区别在于是否需要提前特征^[3]。基于特征点法的前端视觉里程计, 运行稳定, 识别能力较强, 是主流的里程计实现方法。

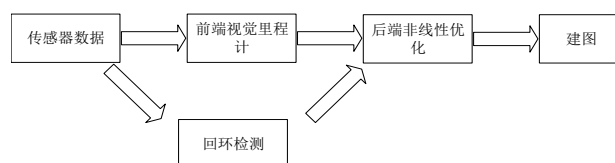


图1 经典视觉 SLAM 框架

Fig.1 Classic visual SLAM framework

2 特征点法

特征点是指相机在运动过程中, 某一区域中, 能反应自身特点与其他区域区别开, 而且运动过程中未发生明显化的点或者区块。

特征点法的两帧之间里程计研究过程包括^[4]:

(1) 通过相机所采集的图像, 对图像中特征点进行提取; (2) 对提取好的特征点进行处理, 完成特征匹配; (3) 估计相机两帧之间的相机运动和场景构建。

特征点分为关键点和描述子两部分。关键点是图像中能表示图像局部特征的点, 描述子是描述关键点附近像素信息的参数。当两帧图像中, 某一处特征点的关键点和描述子相似时, 那么这两个特征点相似, 特征点法要做的就是找出所有图像中的关键点, 计算相应的描述子, 进行特征匹配, 进而估计相机运动和场景构建。常见的关键点提取算法有 Harris 角点、FAST 角点等。

2.1 FAST 角点法

FAST 角点是 E.Rosten 和 T.Drummond 在 2006 年提出^[5], 是一种快速的角点提取算法。其检测过程如图 2 所示。FAST 角点是根据图像像素亮度大小来检测角点, 在左图中选择一像素 P , 亮度为 I_p , 在其周围选择半径为 3 的圆, 将圆周上的区域标记为 1~16, 设定阈值 T , 若圆周上存在连续大于设定值 N 个点的亮度满足:

$$|I_x - T| > 0, x \in (1, 16) \quad (4)$$

那么该点就是 FAST 角点, N 的取值一般有 9、

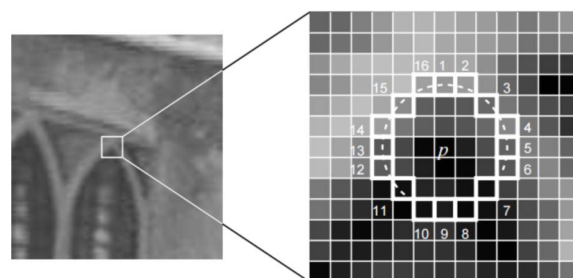


图2 FAST 角点检测示意图

Fig.2 FAST corner detection diagram

11、12 三种，ORB 算法的关键点检测技术是利用 FAST 角点完成。

2.2 Harris 角点

Harris 角点是 C Harris 在 1988 年提出^[6]，通过描述滑动窗口图像像素灰度变化，可表示为：

$$E(u, v) = \sum_{(x, y)} w(x, y) [I(x+u, y+v) - I(x, y)]^2 \quad (5)$$

式中 $E(u, v)$ 是误差函数， $w(x, y)$ 是窗函数，一般取高斯函数， $I(x, y)$ 和 $I(x+u, y+v)$ 表示窗口平移前后的灰度变化值。将 $I(x+u, y+v)$ 进行泰勒展开得：

$$I(x+u, y+v) = I(x, y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + O(u^2, v^2) \quad (6)$$

将式 (6) 化简得：

$$E(u, v) = [u \ v] \sum_{(x, y)} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \quad (7)$$

取 M 表示：

$$M = \sum_{(x, y)} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (8)$$

有：

$$E(u, v) = [u \ v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad (9)$$

用二级矩阵 M 的两个特征值 λ_1, λ_2 来判断是否是角点。若特征值 λ_1, λ_2 值都较大，则该点是角点；若特征值 λ_1, λ_2 都较小，则该点附近都较为平坦；若特征值 λ_1, λ_2 一大一小，则该点为边缘点。给定阈值 C 大小，判断角点。

$$C = \text{Det}(M) - k(\text{Trace}(M))^2 \quad (10)$$

其中， $\text{Det}(M)$ 是矩阵 M 的行列式， $\text{Trace}(M)$ 是矩阵 M 的迹，系数 k 一般取 0.04 ~ 0.06。

在完成角点检测后，需要对角点进行描述子表示，构建出误差较小的特征向量。常见的描述子有两类：一类是基于梯度直方图；一类是基于二进制位串。SIFT 算法中的 SIFT 描述子和 BRIEF 描述子分别对应两类。

2.3 SIFT 描述子

SIFT 描述子是在 SIFT 关键点检测后，对这些关键点局部特征用梯度表示^[7]。首先计算 DOG 空间图像 $L(x, y)$ 内各点的梯度幅值 m 和方向 θ ：

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (11)$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (12)$$

用直方图统计领域像素的梯度方向，接着构造关键点特征向量图，如图 3 所示：图 a 所示区域中，每个小格代表关键点所在尺度空间的一个像素，每个向量代表关键点附近的像素点 (i, j) 的梯度幅值 $m_{i,j}$ 和方向 $\theta_{i,j}$ 的高斯加权值，向量长度为 $G(\sigma, i, j) * m_{i,j}$ ，方向为 $\theta_{i,j}$ 。图 a 中蓝色圆为高斯加权区域，黑色圆点为关键点，图 b 中 4×4 窗口中都含有 8 个方向的向量，将所有的矢量结合起来构成最终的 SIFT 描述子。

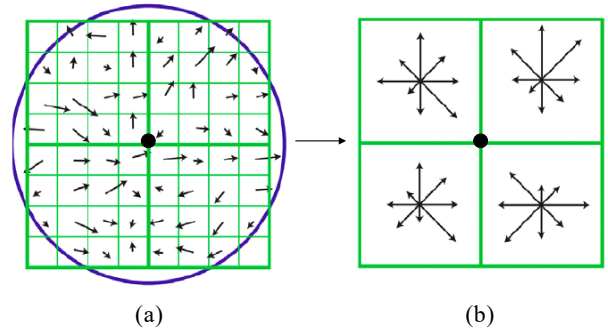


图 3 关键点特征向量图

Fig.3 Key point feature vector diagram

2.4 BRIEF 描述子

BRIEF 描述子是 M. Calonder 在 2010 年提出^[9]，是一种基于二进制位串的描述子方法。其原理是对图像进行分块，比较灰度值，将比较结果进行 0/1 分类。其步骤如下：

(1) 对采集到的图像进行预处理，使用方差为 σ ，窗口为 $q \times q$ 的高斯滤波；

(2) 在图像上选取大小为 $S \times S$ 的矩形区域，在 P 上进行 BRIEF 描述过程；

(3) 定义测试 τ ，其表达式为：

$$\tau(p, x, y) = \begin{cases} 1 & I(p, x) < I(p, y) \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

式中 (x, y) 是区域 P 内的一对像素点， $I(p, x)$ 、 $I(p, y)$ 是其对应的灰度值， τ 的输出值为 0 或 1；

(4) 一个 BRIEF 描述是由无数个 τ 组成的二进制比特串，将像素点排序后进行 τ 测试，代入公式：

$$f_{n_d} = \sum_{1 \leq i \leq n_d} 2^{i-1} \tau(p, x, y) \quad (14)$$

式中， n_d 对像素点 (x, y) 是有序的，经实验验证，

采用 $n_d = 128$ ，方差 $\sigma = \frac{1}{5}S$ 的高斯分布时，BRIEF

描述子的特征向量最优,在特征提取和之后的特征匹配过程,此顺序是不变的。

2.5 ORB 算法

ORB 特征算法则是将 FAST 角点法和 BRIEF 描述子结合得到的。针对边缘反应这一问题,计算了 Harris 响应值,选取了前 N 个具有最大响应值的角点作为角点集合;针对不具有方向性这一问题,引入灰度质心法,质心 C 是将图像灰度值作为权重的中心,步骤如下^[10]:

(1) 在图像 B 中,定义图像块的矩:

$$m_{pq} = \sum_{x,y \in B} x^p y^q I(x,y), p,q = \{0,1\} \quad (15)$$

(2) 通过矩可以找到图像块的质心:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (16)$$

(3) 连接图像块的几何中心 O 与质心 C , 得到一个方向向量 \overrightarrow{OC} , 特征点的方向定义为:

$$\theta = \arctan \left(\frac{m_{01}}{m_{10}} \right) \quad (17)$$

2.6 优化 ORB 算法

本文将 FAST 角点法与 Harris 角点法进一步结合,同时引入 SIFT 算法关键点检测,然后与 BRIEF 描述子结合,用暴力匹配算法完成不同图像的匹配。FAST 角点法提取出来的角点往往分布集中,忽视了其他区域出现关键点概率,因此引入 Harris 角点,然后考虑 FAST 角点计算迅速,但是无法保证尺度不变性,用 SIFT 检测中的高斯金字塔构建尺度不变空间,结合 BRIEF 描述子,完成算法的优化,优化后的算法更加适合排爆机器人对算法精度的高要求的同时,减小了对自身所需 CPU 的高要求,减小了机器人自身负载,保障了危险物处理工作的有效解决。算法的具体流程如下:

(1) 如上式(4)所示,设定阈值对图像进行 FAST-12 角点检测,得到总数为 N 的初始点集合 P_0 ;

(2) 对于集合 P_0 中的每一点 (x,y) , 判断在该点附近半径为 3 的圆周内是否存在 Harris 响应值大于该点,提出后保留剩余点集 P_1 ;

(3) 随后引入 SIFT 关键点检测方法,根据剩余点集 P_1 , 建立图像金字塔,由式(16)、(17)确定 P_1 内所有点对应的质心和方向;

(4) 用优化后的 BRIEF 描述算法,在上一步的基础上,为了保证描述子的方向不变性,用 $2 \times n_d$ 的矩阵表示确定方向的点集 P_1 为:

$$P_2 = \begin{pmatrix} x_1 & x_2 & \cdots & x_{n_d} \\ y_1 & y_2 & \cdots & y_{n_d} \end{pmatrix} \quad (18)$$

本文取 $n_d = 128$ 。根据方向角 θ , 确定旋转矩阵:

$$R_\theta = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (19)$$

将点集 P_2 与 R_θ 相乘,有:

$$P_3 = R_\theta P_2 \quad (20)$$

则有描述子:

$$g_n(t, \theta) = f_{n_d}(t) | (x_i, y_i) \in P_3 \quad (21)$$

产生的描述子既保留了关键点方向不变性与少边缘响应,同时有 ORB 算法中 rBRIEF 算法对噪声的抑制;

(5) 用汉明距离进行匹配,特征匹配方法和仿真结果将在下一节详细描述。

3 特征匹配

特征匹配是视觉里程计模块中重要的一步,在特征点法的基础上,通过图像与图像之间的描述子进行准确匹配,为后续后端优化铺垫。特征匹配的算法已经成熟,常用的有暴力匹配与快速近似最近邻算法^[11]。

暴力匹配是较简单的特征匹配方法,其原理是对每一个特征点 x_i^m 与所有的 x_{i+1}^n 测量描述子的距离,然后排序,取最近的一个作为匹配点。这里的距离采用汉明距离,而不是欧式距离,本文所用的描述子方法是基于 BRIEF 算法,属于二进制串的描述子方法,往往采用汉明距离,在特征匹配模块中指的是不同位数的个数。

4 仿真分析

4.1 仿真环境

本文所选的场景图是实验环境,如图4所示。



图4 实验环境图

Fig.4 Experimental environment diagram

仿真用的两图如图 5 所示。

4.2 仿真过程

分布用 SIFT 算法、ORB 算法与优化 ORB 算法对实验图进行特征匹配，对比图如图 6 所示。

其中，图(a)、(b)、(c)对应着 SIFT 算法、ORB 算法与优化后的 ORB 算法的结果图，比较三个算法图，结果如表 1 所示，SIFT 算法耗 1.63 s，ORB 算法耗时 1.14 s，优化后的 ORB 算法耗时 0.93 s。



图 5 仿真实验图
Fig.5 Simulation experiment diagram

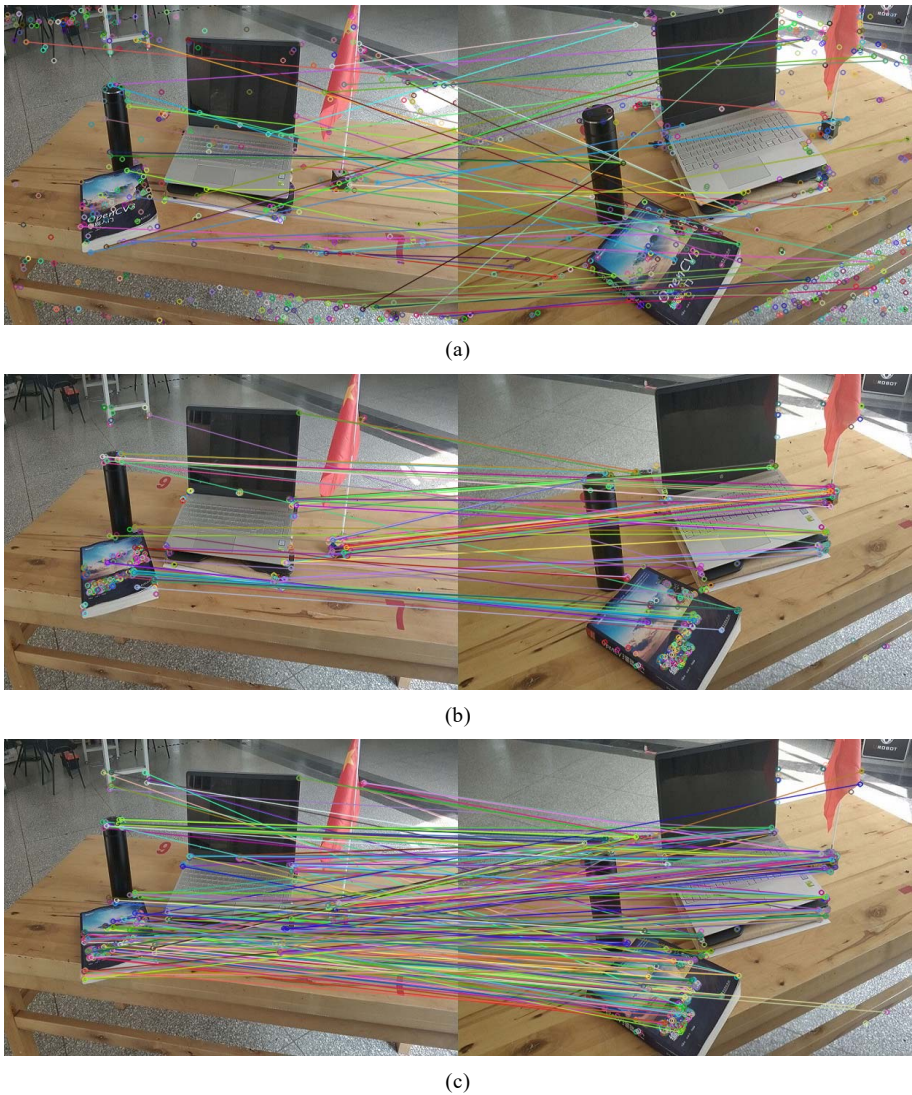


图 6 特征匹配算法匹配结果图
Fig.6 Feature matching algorithm matching result graphs

表 1 特征匹配算法对比
Tab.1 Feature matching algorithm comparison

	SIFT	ORB	优化 ORB
特征点数	850	500	500
匹配数	122	92	359
匹配耗时 (s)	1.63	1.14	0.93

由表 1 可知, SIFT 算法可提取较多的特征点, ORB 算法的特征点虽然少, 但是匹配率更高, 同时 SIFT 算法耗时相对较多。由图 6(a)可知, SIFT 算法存在许多错误, 检测到的特征点不能大概率地匹配上, 而对 ORB 算法来说, 由图 6(b)、(c)及表 1 可以看出, 优化后的 ORB 算法在匹配率和耗时上都得到了一定优化, 为排爆机器人下一步的路径规划做铺垫, 保证了机器人行走的安全性。

5 总结

本文结合 FAST 角点法、Harris 角点法、SIFT 关键点检测、BRIEF 描述子, 对 ORB 算法进行优化, 以便得到适用于排爆机器人的视觉里程计特征匹配方法。通过仿真, 验证了算法的可行性, 为下一步排爆机器人路径规划奠定了基础。

参考文献

- [1] Davison A, Red I, Molton N, Stasse O. Monoslam: Real-time single camera SLAM[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6): 1052-1067.
- [2] Strasdat H, Montiel J M, Davison A J. Visual slam: Why filter?[J]. Image and Vision Computing, 2012, 30(02): 65-77.
- [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.
- [4] 余杰. 基于ORB关键帧闭环检测算法的SLAM方法研究[D]. 浙江工业大学, 2016.
- [5] Rosten E, Drummond T. Machine learning for high-speed corner detection[C]. In European Conference on Computer Vision, 2006.
- [6] Harris C, Stephens M. A combined corner and edge detector[A]. Proceedings of the 4th Alvey Vision Conference[C], Manchester, UK 1988: 147-151.
- [7] Lowe D G. Distinctive image features from scale-invariant key points[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [8] 任忠良. 一种基于SIFT特征的快速图像匹配算法[J]. 软件, 2015, 36(6): 53-57.
- [9] Calonder M, Lepetit V, Fua P. BRIEF: Binary Robust Independent Elementary Features[C]. ECCV, 2010: 778-792.
- [10] Rosin P L. Measuring corner properties. Computer Vision and Image Understanding[J], 1999, 73(2): 291-307.
- [11] 张明军, 俞文静, 袁志, 等. 视频中目标检测算法研究[J]. 软件, 2016, 37(4): 40-45.