

# 弱纹理环境下视觉里程计优化算法研究

张易<sup>1,2</sup>, 项志宇<sup>1,2</sup>, 陈舒雅<sup>1,2</sup>, 顾淑霞<sup>1,2</sup>

<sup>1</sup>浙江省信息处理与通信网络重点实验室, 浙江 杭州 310027;

<sup>2</sup>浙江大学信息与电子工程学院, 浙江 杭州 310027

**摘要** 针对视觉里程计在弱纹理的道路环境下定位精度骤降的问题, 提出一种双目视觉里程计优化算法。首先, 通过提取局部特征平面, 为立体匹配缺失的特征提供平面约束, 增加有效立体特征的数量; 其次, 在特征追踪过程中, 使用匀加速运动模型, 提高特征追踪的数量和质量; 最后, 在位姿计算和优化过程中, 使用考虑特征置信度的光束法平差算法来减少远距离特征引入的误差影响, 提高算法的精度和稳健性。数据集实验和实际场景实验表明, 该算法在占有极少计算资源的情况下对弱纹理环境下的定位精度有较明显的优化效果, 在其他场景也具有较好的适应性。

**关键词** 机器视觉; 视觉里程计; 弱纹理; 局部特征平面; 特征置信度

中图分类号 TP242.6

文献标识码 A

doi: 10.3788/AOS201838.0615001

## Optimization on Visual Odometry under Weak Texture Environment

Zhang Yi<sup>1,2</sup>, Xiang Zhiyu<sup>1,2</sup>, Chen Shuyi<sup>1,2</sup>, Gu Shuxia<sup>1,2</sup>

<sup>1</sup>Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking,  
Hangzhou, Zhejiang 310027, China;

<sup>2</sup>College of Information Science & Electronic Engineering, Zhejiang University,  
Hangzhou, Zhejiang 310027, China

**Abstract** To solve the problem that the accuracy of visual odometry was decreased in the weak texture environment, an optimization algorithm of binocular visual odometry is proposed. Firstly, the planar constraints are provided for the features without stereo matching by extracting the local feature plane, to increase the quantity of effective stereo features. Secondly, in feature tracking, the uniform acceleration motion model is used during feature tracking to increase the quantity and improve the quality of tracked features. Finally, in pose estimation and optimization, the bundle adjustment method which considers the feature confidence is adopted to reduce the influence of distant feature, and improve the accuracy and robustness of the algorithm. The experimental results based on the datasets and actual scene show that the proposed algorithm has obvious optimization effect on the positioning accuracy under weak texture environment with little computational resources, and has adaptability in other scenes.

**Key words** machine vision; visual odometry; weak texture; local feature plane; feature confidence

**OCIS codes** 330.1400; 110.2970; 110.4153; 150.0155

## 1 引 言

视觉里程计是视觉定位的方法之一, 通过行人、车辆或机器人上装备的单个或多个相机的输入即可估计其运动<sup>[1]</sup>。相对于成本较高、易受环境影

响的全球定位系统和漂移较大的惯性导航系统, 视觉里程计具有低成本、高精度的优势。在野外作业、室内导航、水下导航和航空航天等领域, 视觉里程计发挥着重要的作用<sup>[2]</sup>, 是对传统定位方式的有效补充<sup>[3]</sup>。

收稿日期: 2017-10-11; 收到修改稿日期: 2017-12-30

基金项目: 国家自然科学基金(61571390)、NSFC-浙江两化融合联合基金(U61709214)

作者简介: 张易(1996—), 男, 硕士研究生, 主要从事计算机视觉方面的研究。E-mail: 15868103993@163.com

导师简介: 项志宇(1973—), 男, 教授, 博士, 主要从事计算机视觉、智能移动机器人的环境感知和导航控制等方面的研究。E-mail: xiangzy@zju.edu.cn(通信联系人)

视觉里程计由 Nister 等<sup>[4]</sup>于 2004 年提出,主要由图像输入、特征提取、特征的立体匹配和追踪、运动估计和局部优化等部分构成。其中,特征处理的算法对系统性能有着很大影响。为了提高系统在不同环境下的稳健性和精确度,对特征点的选取和内点数量应作深入研究<sup>[5]</sup>。文献[6]针对无人自主车和非轮式机器人实现了一种立体视觉里程计系统,该系统采用 Harris 和基于加速分割测试的特征(FAST)特征点保证实时性,采用特征点匹配法寻找前后帧图像中的对应特征点。文献[7]使用速度限制作为先验信息,在特征追踪中采用速度滤波器滤除外点,提高特征点集中内点的比例。文献[8]建立像素坐标的矫正模型,以提供更精确的特征提取。Kitt 等<sup>[9]</sup>使用图像分格的方法,使得提取的特征在图像上分布均匀。Buczko 等<sup>[10]</sup>使用绝对差之和算法(SAD)进行双目视图的立体匹配,并使用归一化重投影误差作为内点判别的依据,优化了高速场景的定位精度。文献[11]通过循环匹配的方式滤除误匹配点,筛选准确追踪的特征。文献[12]使用绝对差之和(SAD)匹配算法,完成双目视图左右帧特征的立体匹配,并通过循环匹配的方式进一步筛选,基于归一化互相关(NCC)算法作再次检验,从而滤除大部分外点。文献[13]使用特征的多帧融合算法,在多帧追踪中对特征点的像素位置进行优化处理,提高了算法的精度。

在特征丰富的场景中,上述方法均取得了较好的效果。但在无人驾驶时,多有公路、草地等较难提取特征或具有较多重复性纹理的环境,视觉里程计通常因为特征强度较弱而提取较少的特征或产生较多的误匹配。使用上述方法进行外点滤除,则使得作为内点参与位姿估计的特征数量较少,对系统的精度和稳健性产生较大的影响,算法性能在骤降。

针对这类弱纹理的场景,本文提出一种基于局部特征平面以及特征置信度的双目视觉里程计优化算法,以有效提高无人车在弱纹理环境下视觉导航的精度和稳健性。

## 2 视觉里程计总体框架

本文实现的双目视觉里程计系统主要流程如图 1 所示,分为计算线程和优化线程两部分。

在计算线程中,为了使视觉里程计在普通 CPU 上满足实时性的要求,使用效率较高的 ORB(oriented FAST and rotated BRIEF)特征检测和描述算法<sup>[14]</sup>进行特征的提取和处理。利用基于局部特征平面的

优化算法进行特征的匹配追踪,并通过前后帧间的光束法平差算法<sup>[3]</sup>最小化误差函数得到相机的位姿。同时将追踪特征点数较多的帧作为关键帧,计算该帧特征点在世界坐标系的立体坐标作为地图点。

在优化线程中,寻找和当前帧具有较多共同特征的关键帧作为相邻关键帧<sup>[15]</sup>,与这些帧的地图点共同构建局部地图,并对相机位姿和匹配的地图点进行局部地图的光束法平差优化<sup>[16]</sup>。

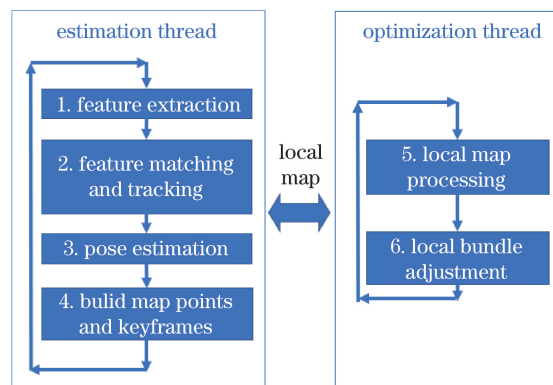


图 1 整体系统框架

Fig. 1 Framework of system

## 3 基于局部特征平面的优化算法

为了提高双目视觉里程计在弱纹理环境下的稳定性,在特征的匹配追踪阶段提出基于局部特征平面的优化算法(图 2)。首先进行特征的立体匹配<sup>[17]</sup>,通过特征的尺度信息减少匹配搜索范围,并通过相互一致性检测得到精度较高的立体匹配点。然后进行图像分块,对每个图像分块使用随机采样一致性(RANSAC)算法<sup>[3]</sup>分别提取可能存在的局部特征平面,为图像块内匹配缺失的特征提供局部平面约束信息,并以此进行特征的再匹配,减少每个

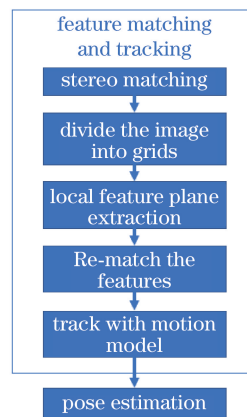


图 2 优化算法的流程

Fig. 2 Flowchart of optimization algorithm

特征的匹配范围,提高匹配的准确率。最后,通过匀加速运动模型进行更精确的特征追踪。

### 3.1 左右图像立体匹配

对于标定过的双目相机,双目图像平面的极线<sup>[18]</sup>互相平行,世界坐标系中的同一个特征点在左右相机的输入图像中有相同的  $y$  像素坐标。在左右图像的立体匹配中,对于提取的每个特征点,根据其尺度信息确定  $x$  坐标搜索范围,提高匹配的速度和精度<sup>[15]</sup>。在匹配过程中使用相互一致性检测,令左右图中的匹配点都是互相的最佳匹配点,并且和次匹配点相似度较小<sup>[19]</sup>。图 3(a)为当前帧双目图像的左图,图 3(b)为对应的右图,其中红色点代表立体匹配的特征,蓝色点表示在此次立体匹配中被筛选的特征。在弱纹理环境下,大部分相似性特征会被筛选,因此得到的立体匹配点数目较少,但为之后的局部特征平面提取提供了精度较高的立体匹配点。

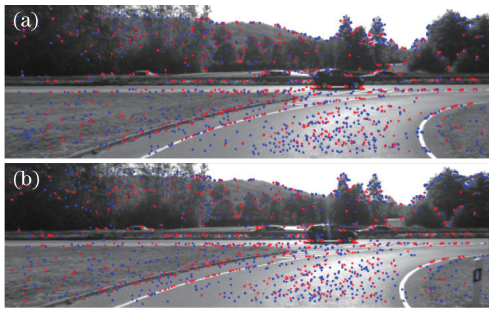


图 3 KITTI01 序列中立体匹配示例。

(a)左目图像特征;(b)右目图像特征

Fig. 3 Example of stereo matching in KITTI01 sequence.

(a) Features in left image; (b) features in right image

### 3.2 图像分块

为了保证算法的速度和稳健性,需要对图像进行分块处理。在无人车的视觉导航中,弱纹理环境多为平坦地面、公路、草坪等场景,在无人车的行进过程中,这些区域主要分布在图像下半部分。考虑到在实际场景中远处特征具有较大的不确定性,且平面较少,仅对包含大量近处特征的图像下半部分进行分块处理。为了较好地提取局部特征平面,图像分块应保持分布均匀,且包含一定数量的特征。本文使用的分块方法如图 4 所示,将相机采集的图像下半部分分成 10 块,其中,中部两块因为在视角中心,面积略大于其他分块。

### 3.3 局部特征平面提取

对于当前帧的每一个图像分块,可以得到具有立体坐标的特征点集,任取 3 个特征点,坐标分别为

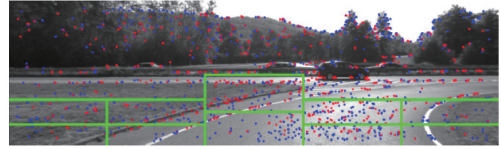


图 4 KITTI01 序列中的图像分块示例

Fig. 4 Example of image segmentation in KITTI01 sequence

$P_1(X_1, Y_1, Z_1)$ 、 $P_2(X_2, Y_2, Z_2)$ 、 $P_3(X_3, Y_3, Z_3)$ , 平面方程使用  $AX + BY + CZ + D = 0$  表示,则:

$$\begin{cases} A = (Y_2 - Y_1)(Z_3 - Z_1) - (Z_2 - Z_1)(Y_3 - Y_1) \\ B = (Z_2 - Z_1)(X_3 - X_1) - (X_2 - X_1)(Z_3 - Z_1) \\ C = (X_2 - X_1)(Y_3 - Y_1) - (Y_2 - Y_1)(X_3 - X_1) \\ D = -(AX_1 + BY_1 + CZ_1) \end{cases} \quad (1)$$

使用点到平面的距离  $d$  来判定块内特征点  $P'(X_0, Y_0, Z_0)$  是否在平面  $AX + BY + CZ + D = 0$  上:

$$d = \left| \frac{AX_0 + BY_0 + CZ_0 + D}{\sqrt{A^2 + B^2 + C^2}} \right| \quad (2)$$

对于每个图像分块,通过 RANSAC 算法迭代计算包含最多共面立体特征的平面,并设定最大迭代次数以保证算法速度。通过迭代得到具有较多共面立体特征的图像分块,即成功提取了局部特征平面。如图 5 所示,红色点即代表在一个含有路边草地干扰的图像块中经过 RANSAC 算法得到特征平面的特征点集。



图 5 一个图像块中的特征平面提取

Fig. 5 Feature plane extracted in an image block

### 3.4 特征再匹配

对于成功提取局部特征平面的图像分块,如图 6 所示,红色点为构成特征平面的特征点集,蓝色点为立体匹配缺失的特征。对于左目图像,该分块中每一个立体匹配缺失的特征点  $p(x_1, y_1)$ ,使用该特征平面  $AX + BY + CZ + D = 0$  的约束可以估算其对应的视差:

$$\hat{d}_{\text{cal}} = -\frac{AB_c(x_1 - x_0) + BB_c(y_1 - y_0) + CB_c f}{D} \quad (3)$$

式中  $B_c$  为相机的基线长度,  $f$  为相机焦距,  $x_0, y_0$  代表相机的图像像素中心坐标。从而得到该特征在对应右目图像中的估计匹配点的像素坐标  $\hat{p}(\hat{x}_r, \hat{y}_r)$ :

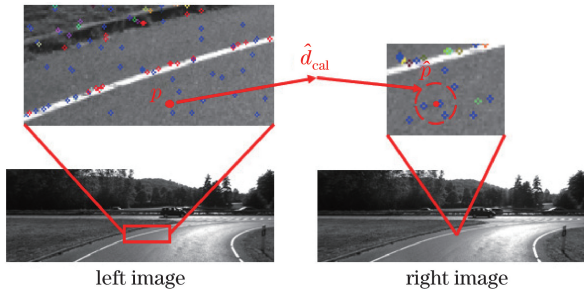


图 6 特征再匹配示例

Fig. 6 Example of feature re-matching

$$\begin{cases} \hat{x}_r = x_l - \hat{d}_{cal} \\ \hat{y}_r = y_l \end{cases} \quad (4)$$

在该估计匹配点周围 8 个像素范围内通过比较特征描述子的相似度进行匹配。若只有一个相似特征,该特征即为最佳匹配点;若有多个相似特征,则以距离估计匹配点最近的相似特征点作为最佳匹配点;若没有相似的特征点,直接根据平面约束得到该特征的三维坐标,作为一个新的立体特征点。

### 3.5 基于匀加速运动模型的特征追踪

对于弱纹理环境,使用局部特征平面算法重匹配的特征具有较弱的特征强度,在追踪时较易丢失,所以使用匀加速运动模型来减少前后帧特征追踪匹配的搜索范围,提高追踪的精度和速度。

对于当前帧  $k$ ,前几帧的相机位姿  $T$  是已知的,设定前几帧世界坐标系相对于相机坐标系的变换分别为  $T_{k-1}^{wc}$ 、 $T_{k-2}^{wc}$ 、 $T_{k-3}^{wc}$ ,则相机坐标系相对于世界坐标系的变换分别为  $T_{k-1}^{cw}$ 、 $T_{k-2}^{cw}$ 、 $T_{k-3}^{cw}$ 。可以得到  $k-1$  帧坐标系相对于  $k-2$  帧坐标系的变换:

$$V = T_{k-1}^{cw} * T_{k-2}^{wc}, \quad (5)$$

$$\Delta V = T_{k-1}^{cw} * T_{k-2}^{wc} - T_{k-2}^{cw} * T_{k-3}^{wc}. \quad (6)$$

根据匀加速运动模型,可以认为当前帧  $k$  相对于  $k-1$  帧相机坐标系的变换:

$$V_{cur} \approx V + \Delta V. \quad (7)$$

因此可以估计当前帧的相机位姿:

$$\hat{T}_k^{cw} = V_{cur} * T_{k-1}^{cw}. \quad (8)$$

根据估计的当前帧位姿  $\hat{T}_k^{cw}$  和  $k-1$  帧的立体特征点对应的世界坐标  $P_w = (X_w \ Y_w \ Z_w)^T$  可以得到其相对于当前帧的相机的立体坐标  $P_c = (X_c \ Y_c \ Z_c)^T$ :

$$P_c = \hat{R}_{cw} * P_w + \hat{t}_{cw}, \quad (9)$$

式中  $\hat{T}_{cw} = \begin{bmatrix} \hat{R}_{cw} & \hat{t}_{cw} \\ 0 & 1 \end{bmatrix}$ ,从而得到特征在当前帧的

投影像素坐标  $p_c(x_c, y_c)$ ,减小特征追踪的匹配范围:

$$\begin{cases} x_c = \frac{fX_c}{Z_c} + x_0 \\ y_c = \frac{fY_c}{Z_c} + y_0 \end{cases}, \quad (10)$$

式中  $f$  为相机焦距,  $x_0, y_0$  为图像中心坐标。

## 4 位姿估计和优化

通过特征点的三角计算可以得到特征的立体坐标<sup>[19]</sup>。图 7 中的  $B_c$  代表左、右相机光心距离的基线长度,  $f$  为相机的焦距。相机坐标系中的三维特征点  $P(X_c, Y_c, Z_c)$  在左、右相机中的投影分别为  $(x_l, y_l)$ 、 $(x_r, y_r)$ ,  $(x_0, y_0)$  为图像的中心点。

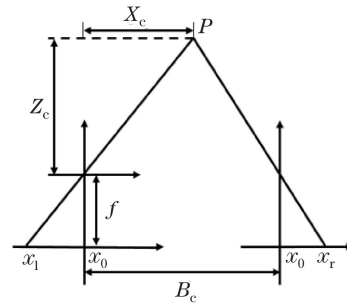


图 7 双目立体成像模型

Fig. 7 Binocular stereoscopic imaging model

根据三角形近似关系可得:

$$\begin{cases} X_c = \frac{B_c(x_l - x_0)}{x_l - x_r} \\ Y_c = \frac{B_c(y_l - y_0)}{x_l - x_r} \\ Z_c = \frac{B_c f}{x_l - x_r} \end{cases}. \quad (11)$$

假设基线长度  $B_c$  和相机焦距  $f$  没有测量误差,则特征点立体构建的主要误差即为特征点像素坐标的测量误差,根据误差传播定理<sup>[20]</sup>可得:

$$\begin{cases} S_{X_c} \approx \frac{Z_c}{f} S \\ S_{Y_c} \approx \frac{Z_c}{f} S \\ S_{Z_c} \approx \frac{2Z_c^2}{B_c f} S \end{cases}, \quad (12)$$

式中  $S$  为特征点的像素坐标的测量误差。

在相同的像素测量误差下,近距离的特征具有更高的置信度。为了在近距离特征缺失时减小远距离特征引入的误差,提高系统的稳健性,设定重投影误差为

$$\boldsymbol{\varepsilon}_i^k = \frac{\|\mathbf{p}_i^k - \pi(\hat{\mathbf{R}} * \mathbf{P}_i^{k-1} + \hat{\mathbf{t}})\|_2}{|x_i^l - x_i^r|}, \quad (13)$$

式中  $x_i^l - x_i^r$  表示特征的视差,  $\pi$  表示标准平面投影,  $\mathbf{p}_i^k$  指当前帧特征的像素坐标,  $\mathbf{P}_i^{k-1}$  指上一帧特征的立体坐标,  $\hat{\mathbf{R}}$  和  $\hat{\mathbf{t}}$  分别为当前帧相对于上一帧估计的旋转矩阵和位移矩阵。

因此在位姿估计中,对于前后帧追踪到的特征点,通过最小化如下误差函数可以获得相机的位姿估计:

$$(\hat{\mathbf{R}}, \hat{\mathbf{t}}) = \arg \min_{\hat{\mathbf{R}}, \hat{\mathbf{t}}} \sum_i \boldsymbol{\varepsilon}_i^k = \arg \min_{\hat{\mathbf{R}}, \hat{\mathbf{t}}} \sum_i \frac{\|\mathbf{p}_i^k - \pi(\hat{\mathbf{R}} * \mathbf{P}_i^{k-1} + \hat{\mathbf{t}})\|_2}{|x_i^l - x_i^r|}. \quad (14)$$

上述得到的当前帧相机的位姿转换为绝对位姿,作为初值输出到优化线程中进行局部地图范围内的优化,得到优化后的最终位姿:

$$(\mathbf{R}_j^g, \mathbf{t}_j^g) = \arg \min_{\mathbf{R}_j^g, \mathbf{t}_j^g} \sum_{j=k-M}^k \sum_{i=1}^N \boldsymbol{\varepsilon}_i^j, \quad (15)$$

式中  $\boldsymbol{\varepsilon}_i^j = \frac{\|\mathbf{p}_i^j - \pi(\mathbf{R}_j^g * \mathbf{P}_i^j + \mathbf{t}_j^g)\|_2}{|x_{li}^j - x_{ri}^j|}$ ,  $M$  代表当前帧在局部地图中共同匹配特征较多的相邻关键帧的个数,  $N$  代表局部地图中相邻关键帧匹配的地图点的个数。  $\mathbf{R}_j^g, \mathbf{t}_j^g$  为第  $j$  帧世界坐标系相对于相机坐标系的旋转和位移矩阵,  $|x_{li}^j - x_{ri}^j|$  为第  $j$  帧第  $i$  个特征的视差,  $\mathbf{P}_i^j$  指局部地图中和当前帧匹配的地图点。整个优化过程采用 g2o 软件包实现<sup>[21]</sup>。

## 5 结果验证与分析

在实验中,整个系统在 i5 6300HQ 2.3 GHz 的 CPU 上测试运行。首先使用 KITTI<sup>[22]</sup> 数据集进行实验,与排行榜中算法进行对比分析,并使用 BumbleBee2 双目相机进行实际场景实验。

### 5.1 KITTI 数据集实验

KITTI 数据集是自动驾驶领域较为著名的公开数据集,通过在车上装载双目摄像机和高精度全球定位系统(GPS)提供 21 个真实场景下长距离的双目图像序列和对应的真值轨迹。其中:00~10 这 11 个序列公开了轨迹的真值,用于在本地对系统进行测试;11~21 这 10 个序列仅提供图片序列,不提供轨迹真值,用于在线提交算法结果,从而评价视觉里程计算法的综合性能。整套数据集具有较高的客观性和参考价值。

KITTI 数据集 01 序列中的高速公路场景即为

比较典型的弱纹理环境,如图 8 所示,白色点即为优化前算法用于计算位姿追踪到的特征点,因为环境较为单一,近处场景多为纹理较弱且相似的路面,很难提取足够数量的特征用于位姿估计,使得算法精度在类似序列出现骤降。



图 8 KITTI01 序列的场景

Fig. 8 Scenes of KITTI 01 sequence

在弱纹理环境下,使用基于局部特征平面的算法对提取的特征提供平面约束,为每个特征提供视差估计,减少特征匹配搜索的范围,提高有效立体匹配的数目。如图 9(b),白色特征点表示优化算法新增的立体匹配点。在具有纹理缺失和重复性纹理特点的地面上,该算法可明显增加有效的立体匹配数量。

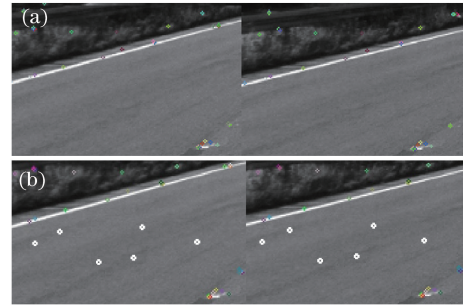


图 9 (a)优化前和(b)优化后左右图像局部立体匹配点数目对比

Fig. 9 Comparison of the number of local stereo match points in left and right images (a) before and (b) after optimization

对具有较多弱纹理环境的 KITTI01 序列,统计特征立体匹配数目如图 10 所示。  $F$  指序列的帧数,  $n$  指每一帧的图片分块区域提取到的立体匹配点个数,立体匹配点指优化前提取的匹配点,新增匹配点指使用局部特征平面算法新增的立体匹配点。可以发现,整个序列中,该算法对于增加有效立体匹配点的数量具有明显的优化效果。

通过计算算法估计结果和 KITTI01 序列真值轨迹的平移误差,发现该算法在重复性纹理及特征缺失环境下,对视觉里程计的性能有较明显的提升。如图 11 所示,  $L$  指测试距离,  $e$  为在 100~800 m 长的轨迹下对应的平均平移误差。

使用 KITTI 数据集所有序列对算法进行测试并提交结果,和排行榜中其他算法<sup>[15,23-26]</sup>的精度比

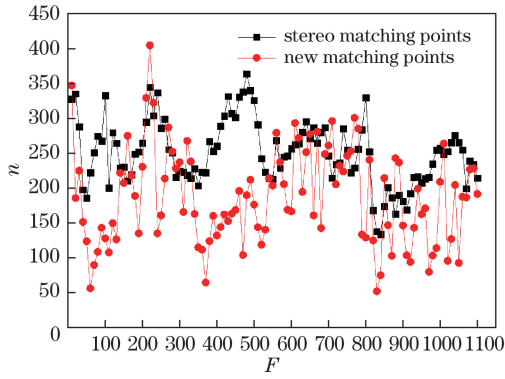


图 10 KITTI01 序列立体匹配点数目统计  
Fig. 10 Number of stereo matching points in KITTI 01 sequence

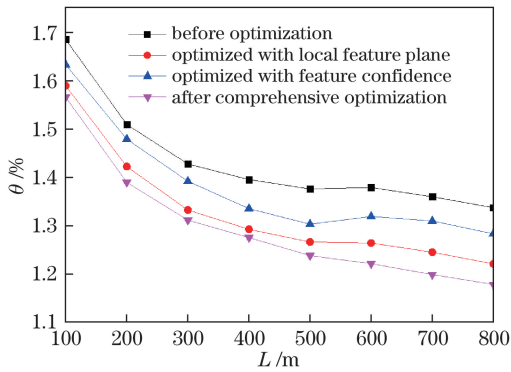


图 11 KITTI01 序列误差比较

Fig. 11 Comparison of error of KITTI 01 sequence

较如表 1 所示。其中基于特征选择和追踪的立体视觉里程计(SOFT2)<sup>[24]</sup>、基于稳健的外点检测的视觉里程计(LG\_SLAM)、RotRocc+<sup>[25]</sup>为 KITTI 视觉里程计排行榜中前三位的算法,FPVO 为本文算法,ORB\_SLAM2<sup>[15]</sup>、基于立体视觉的并行跟踪与建图(S-PTAM)<sup>[26]</sup>是较为著名的开源 SLAM 算法。

表 1 KITTI 排行部分算法精度对比

Table 1 Precision comparison of different algorithms in KITTI benchmark

Method	Average translation error / %	Average rotation error / [(°)/m]
SOFT2	0.66	0.0014
LG_SLAM	0.82	0.0020
RotRocc+	0.83	0.0026
FPVO	1.10	0.0023
ORB_SLAM2	1.15	0.0027
S-PTAM	1.19	0.0025

在 KITTI 数据集的性能测试中,本文算法的平均平移误差为 1.10%,平均旋转误差为 0.0023 (°)/m,相较 ORB\_SLAM2 和 S-PTAM 性能更优,平均旋

转误差也优于排名第 3 的 RotRocc+ 算法,具有较高的综合性能。

因为 KITTI 数据集中包含特征缺失及重复性纹理环境的序列较少,使得本文算法在该环境下的优化对平均误差的影响较小,因此本文算法在综合性能上相较于 SOFT2 等排名顶尖的算法仍有差距。但根据排行榜中提供的误差分析可以发现,对本文关注的弱纹理环境,如具有较多弱纹理及重复性纹理环境的 11 序列,相对于 SOFT2、LG\_SLAM、RotRocc+ 等算法,本文算法也有较好的精度和稳健性(图 12)。

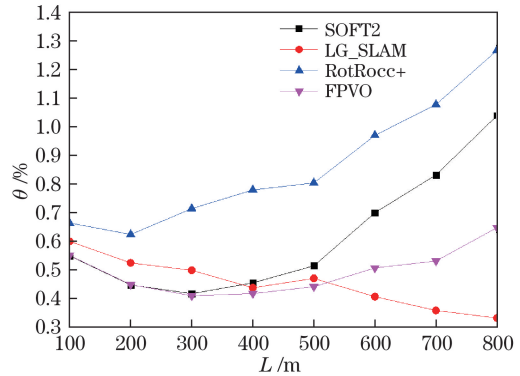


图 12 KITTI11 序列不同算法精度比较

Fig. 12 Accuracy comparison of different algorithms in KITTI 11 sequence

统计系统的各处理流程在 i5 2.3 GHz 的 CPU 上的平均运行时间(表 2),平均耗时仅为 82 ms·frame<sup>-1</sup>。由于使用了较低的硬件配置,相较排名前 3 位的算法(SOFT2 算法为 100 ms·frame<sup>-1</sup>, LG\_SLAM 算法为 200 ms·frame<sup>-1</sup>, RotRocc+ 算法为 250 ms·frame<sup>-1</sup>)具有更优的实时性,适合应用于无人车。

表 2 系统运行时间统计

Table 2 System runtime statistics

Operation	t / s
Feature extraction and matching	0.0390
Optimized by local feature	0.0052
Feature tracking and pose optimization	0.0380
Total	0.0820

## 5.2 实际场景实验

实际场景实验采用电动打猎车作为机器人移动平台,如图 13(a)所示,使用 BumbleBee2 双目相机,如图 13(b)所示,在校园环境中采集双目图像序列。采集的图像分辨率为 640 pixel×480 pixel,基线长度为 12 cm。

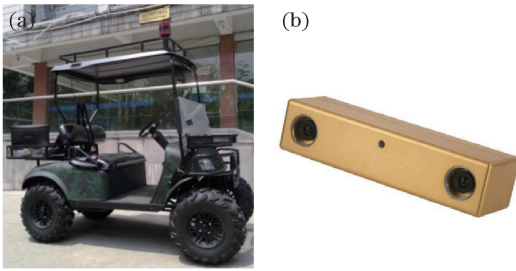


图 13 (a) 电动打猎车; (b) BumbleBee2 双目相机  
Fig. 13 (a) Electric hunting cart;  
(b) BumbleBee2 binocular camera

实验采集的实际轨迹距离约为 1030 m, 对应 10000 帧双目图像序列, 具有较长的距离。采集的图片中包含弱纹理、光照变化、运动物体等多种较为复杂的环境, 较为真实地体现了无人车在实际应用中的情况, 如图 14 所示。



图 14 部分实验场景。(a) 弱纹理环境;

(b) 移动的行人和车辆; (c) 阴影、曝光等光照变化

Fig. 14 Part of the experimental scene. (a) Weak texture environment; (b) moving pedestrians and vehicles;  
(c) shadow, exposure and other light changes

因为实际场景测试较难获得位姿真值, 所以本文把重建得到的轨迹和卫星地图进行比照。如图 15 所示, 红线表示算法重建的轨迹图, S 点为起始点, E 点为终止点。本文算法的结果能与实际运动轨迹较好地吻合, 具有较好的精度和稳健性。

## 6 结 论

提出一种基于局部特征平面和特征置信度的优化算法。实验结果表明, 本算法可明显增加有效立体匹配的数量, 对弱纹理道路环境下的定位精度有明显提升。在 KITTI 数据集和实际场景的测试中, 本算法具有较好的精度和稳健性, 同时在较低的硬件配置下亦具有较好的实时性, 适合于无人车视觉导航的应用。



图 15 实际场景实验结果

Fig. 15 Results of actual scene experiment

目前在视觉里程计研究中, 典型的特征环境下的算法已经达到了很高的精度, 但是对于复杂多变的环境, 如大量光照变化、大规模运动环境等, 通常出现精度骤降的问题。复杂环境下的稳定估计作为算法实用化必需解决的难题, 将是之后的重要研究方向。

## 参 考 文 献

- [1] Scaramuzza D, Fraundorfer F. Visual odometry Part I: the first 30 years and fundamentals[J]. IEEE Robotics & Automation Magazine, 2011, 18(4): 80-92.
- [2] Yuan M, Li A H, Zheng Y, *et al.* Point-line feature fusion in monocular visual odometry[J]. Laser & Optoelectronics Progress, 2018, 55(2): 021501. 袁梦, 李艾华, 郑勇, 等. 点线特征融合的单目视觉里程计 [J]. 激光与光电子学进展, 2018, 55(2): 021501.
- [3] Fraundorfer F, Scaramuzza D. Visual odometry Part II: matching, robustness, optimization, and applications[J]. IEEE Robotics & Automation Magazine, 2012, 19(2): 78-90.
- [4] Nister D, Naroditsky O, Bergen J. Visual odometry[C]. Washington, DC: CVPR, 2004: 652-659.
- [5] Lin Z L, Zhang G L, Yao E L, *et al.* Stereo visual odometry based on motion object detection in the dynamic scene[J]. Acta Optica Sinica, 2017, 37(11): 1115001. 林志林, 张国良, 姚二亮, 等. 动态场景下基于运动物体检测的立体视觉里程计 [J]. 光学学报, 2017, 37(11): 1115001.
- [6] Howard A. Real-time stereo visual odometry for autonomous ground vehicles[C]// International Conference on Intelligent Robots and Systems, 2008: 3946-3952.
- [7] Escalera A D L, Izquierdo E, Martín D, *et al.* Stereo

- visual odometry in urban environments based on detecting ground features[J]. *Robotics and Autonomous Systems*, 2016, 80: 1-10.
- [8] Krešo I, Šegvić S. Improving the Egomotion Estimation by correcting the Calibration Bias[C]// *International Conference on Computer Vision Theory and Applications*. Berlin: VISAPP, 2015: 347-356.
- [9] Kitt B, Geiger A, Lategahn H. Visual odometry based on stereo image sequences with RANSAC-based outlier rejection scheme[C]// *Intelligent Vehicles Symposium*. La Jolla: IEEE, 2010: 486-492.
- [10] Buczko M, Willert V. How to distinguish inliers from outliers in visual odometry for high-speed automotive applications[C]// *Intelligent Vehicles Symposium*. Gothenburg: IEEE, 2016: 478-483.
- [11] Fanfani M, Bellavia F, Colombo C. Accurate keyframe selection and keypoint tracking for robust visual odometry[J]. *Machine Vision and Applications*, 2016, 27(6): 833-844.
- [12] Cvišić I, Petrović I. Stereo odometry based on careful feature selection and tracking[C]// *European Conference on Mobile Robots*. Lincoln: ECMR, 2015: 1-6.
- [13] Badino H, Yamamoto A, Kanade T. Visual odometry by multi-frame feature integration[C]// *IEEE International Conference on Computer Vision Workshops*. Sydney: ICCV, 2013: 222-229.
- [14] Rublee E, Rabaud V, Konolige K, *et al.* ORB: an efficient alternative to SIFT or SURF[C]// *IEEE International Conference on Computer Vision*. Barcelona: ICCV, 2011: 2564-2571.
- [15] Mur-Artal 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.
- [16] Triggs B, McLauchlan P F, Hartley R I, *et al.* Bundle adjustment: a modern synthesis[C]// *International Workshop on Vision Algorithms*. Springer Berlin Heidelberg, 1999: 298-372.
- [17] Liu J, Zhang J X, Dai Y, *et al.* Dense stereo matching based on cross-scale guided image filtering[J]. *Acta Optica Sinica*, 2018, 38(1): 0115004.
- 刘杰, 张建勋, 代煜, 等. 基于跨尺度引导图像滤波的稠密立体匹配[J]. *光学学报*, 2018, 38(1): 0115004.
- [18] Zhang Q, Dong Y, Li H B, *et al.* Underwater curve constraint model based on non-parallel binocular vision[J]. *Acta Optica Sinica*, 2016, 36(3): 0315001. 张强, 董玉, 李海滨, 等. 非平行双目视觉系统水下成像的极线约束模型[J]. *光学学报*, 2016, 36(3): 0315001.
- [19] Lu W. Research on key techniques of high-precision and real-time visual localization[D]. Hangzhou: Zhejiang University, 2015: 9-10. 卢维. 高精度实时视觉定位的关键技术研究[D]. 杭州: 浙江大学, 2015: 9-10.
- [20] Luhmann T, Robson S, Kyle S, *et al.* Close-range photogrammetry and 3D imaging[M]. 2nd ed. Berlin: De Gruyter, 2013.
- [21] Kümmerle R, Grisetti G, Strasdat H, *et al.* g2o: a general framework for graph optimization[C]// *IEEE International Conference on Robotics and Automation*. Wellington: ICRA, 2011: 3607-3613.
- [22] Geiger A, Lenz P, Stiller C, *et al.* Vision meets robotics: the KITTI dataset[J]. *The International Journal of Robotics Research*, 2013, 32(11): 1231-1237.
- [23] Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite[C]// *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
- [24] Cvišić I, Česić J, Marković I, *et al.* SOFT-SLAM: Computationally efficient stereo visual simultaneous localization and mapping for autonomous unmanned aerial vehicles[J]. *Journal of Field Robotics*, 2018, 35(4): 578-595.
- [25] Buczko M, Willert V. Flow-decoupled normalized reprojection error for visual odometry[C]// *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016: 1161-1167.
- [26] Pire T, Fischer T, Civera J, *et al.* Stereo parallel tracking and mapping for robot localization[C]// *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015: 1373-1378.