

文章编号:1005-0523(2020)04-0048-06

DOI:10.16749/j.cnki.jecjtu.2020.04.008

基于高斯金字塔的视觉里程计算法研究

刘瑞,徐雪松,曾昱

(华东交通大学 电气与自动化工程学院,江西 南昌 330013)

摘要:针对视觉里程计常用角点提取算法因角点分布不均匀而导致运动信息偏差较大的问题,提出一种基于高斯金字塔的角点提取算法。该算法在角点提取过程中先采用高斯金字塔算法对图片进行尺度压缩,纹理丰富区域压缩纹理,纹理稀疏区域聚集纹理,得到小尺度顶图;然后采用 Shi-Tomasi 算法提取小尺度顶图角点特征以实现角点粗定位,最后将粗定位信息映射回细节丰富的原图进行角点精准定位,得到图片特征信息。最后,利用金字塔 LK 光流法追踪角点,根据相机对极几何约束模型恢复运动信息。论文采用 KITTI 数据集,与原 Shi-Tomasi 算法、Harris 算法、Fast 算法进行了对比实验,结果表明本算法可有效改善角点分布均匀性,提高视觉里程计运动信息恢复的精度。

关键词:高斯金字塔;纹理丰富;角点定位;分布均匀;视觉里程计

中图分类号:TP391 **文献标志码:**A

本文引用格式:刘瑞,徐雪松,曾昱. 基于高斯金字塔的视觉里程计算法研究[J]. 华东交通大学学报,2020,37(4):48-53.

Citation format:LIU R, XU X S, ZENG Y. Research on visual odometry based on Gaussian Pyramid[J]. Journal of East China Jiaotong University, 2020, 37(4): 48-53.

视觉里程计算法^[1-2]是视觉定位的方法之一,通过在车辆、机器人等载体上装备的单个或多个相机获取视频信息,然后通过对连续帧图像分析计算运动情况^[3]。相比全球定位系统或惯性导航等定位方式,视觉里程计^[4-5]成本更低、累积误差更小,因而在室内和航空航天等领域发挥着重要的作用^[6],是对传统定位方式的有效补充^[7]。

视觉里程计的主要环节包括图片获取、特征提取与匹配追踪、运动估计等^[8]。图片获取可通过相机获得。特征提取与追踪^[9-10]是匹配或追踪前后帧的同一特征点。运动估计则是根据特征点变换信息并结合相机内外参数恢复其运动信息。在特征提取环节中,由于角点检测算法简单、实时性好,为目前常用的提取算法,主要分为基于图像边缘和基于图像灰度的方法两大类^[11]。基于灰度的方法运行速度快、精度高,为主流算法。代表性算法有 Moravec、Fast、Harris、Shi-Tomasi 等。但这些算法存在一个共同问题:容易出现角点聚簇现象。

为解决特征分布不均匀的问题,一些学者进行了研究。如:文献[12]利用相邻匹配特征点之间的空间距离不变性,将特征点分为静态与动态区域,最终只提取静态背景区域的特征点。文献[13]将图片均匀等分成几个区域,然后对每个区域采用不同的阈值进行特征提取。这些方法的基本思路都是先将图像分块,然后针对不同块选择不同阈值,以保证每个部分都保留基本的角点信息。对于这类方法,合理分块和选择角点阈值都是很重要的问题。

针对角点分布聚集问题,本文将高斯金字塔^[14-15]方法应用到视觉里程计中的角点特征提取过程中,提出了一种新的角点提取方法。该方法通过对金字塔顶层图像进行角点提取,对特征位置进行粗定位;然后映射到金字塔底层图像进行精确定位,从而获取均匀的角点信息。文章采用 KITTI 数据集进行了算法验证,结果表明该方法能够有效改善所提取的角点分布均匀性和所恢复得运动信息的精确度。

收稿日期:2020-01-16

基金项目:国家自然科学基金(61763012)

作者简介:刘瑞(1994—),女,硕士研究生,主要研究方向为图像处理,视觉里程计。

通信作者:徐雪松(1970—),男,教授,博士,硕士生导师,主要研究方向为无人机控制,智能交通,机器视觉等。

1 基于高斯金字塔的角点提取算法

1.1 促进角点信息均匀化的思路

造成角点信息分布不均匀的根本问题在于角点特征反差太大。对于纹理特征分布不均衡的图像,角点特征都聚集在纹理丰富区域,而纹理稀疏区域则数量很少。针对该问题,应该对各个区域精准定位,单独提取特征。基本思路是通过高斯金字塔缩小原图尺度,压缩原纹理丰富区域细节,增加纹理稀疏区域纹理。通过尺度压缩可以自适应的促使整个图像特征分布均匀化,通过控制每个区域的特征数量来促使不同的图像特征信息保持在一个稳定的范围内。

本文先采用 Shi-Tomasi 方法提取小尺度图片角点,这些角点信息可以提示原图像对应位置特征信息。不过,由于高斯金字塔变换会模糊图像细节,故小尺度图像的角点信息精确度不够。因此,可将该角点位置映射回原图,在原图相对应位置的周围区域重新分块搜索精确的角点位置,这样一来,既在宏观上保持了角点分布的均匀性,在微观上也能获得更精确的角点位置信息。

1.2 角点提取算法流程

基于高斯金字塔的角点提取算法主要有以下几个步骤:

Step1:用高斯金字塔对原图像进行尺度压缩。

设原图为 I ,高斯金字塔有 N 层,从塔底到塔顶图像为 I^1, I^2, \dots, I^N ,则高斯金字塔可以用公式(1)生成

$$I^l(i, j) = I(i, j) \\ I^l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) I^{l-1}(2i+m, 2j+n) \quad (1)$$

其中: l 为第 l 个图层; $w(m, n)$ 为高斯滤波窗口函数; $1 \leq l \leq N, 0 \leq i \leq r, 0 \leq j \leq c; r, c$ 分别是第 l 图层的高和宽。

Step2:提取高斯金字塔顶层图像角点信息。

角点提取算法很多,如 Sift、Fast、Shi-Tomasi 等。考虑到 Shi-Tomasi 方法稳定性好、自适应能力强、实时性好,本文选用 Shi-tomasi 方法对金字塔顶层图像进行角点提取。具体计算方法如下:

Step2.1:逐个像素点计算金字塔顶层图像 I^N 的横向梯度 I_x 和纵向梯度 I_y 。

Step2.2:按公式(2)计算每个像素点的 M 矩阵

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

其中: $w(x, y)$ 是窗口滤波函数,通常选择为高斯函数; x 和 y 是所选择的平滑窗口内的像素点坐标索引。

Step2.3:逐个像素点计算相应像素点 M 矩阵的特征值 λ_1, λ_2 ,若两个特征值中较小的一个大于最小阈值,则会得到角点,并标记相应位置。所得角点位置数据集为 corners1

$$\text{corners1} = \{(i_k, j_k) | k = [1, 2, \dots, n_1]\} \quad (3)$$

其中: i, j 是当前像素点位置索引; n_1 为角点数阈值。

Step3:将顶层角点位置信息映射回原图,确定原图的角点搜索区域。该过程分成如下几步:

Step3.1:按公式(4)计算 corners1 对应原图像的位置 corners2

$$\text{corners2} = 2^{N-1} \text{corners1} = \{(2^{N-1} i_k, 2^{N-1} j_k) | k = [1, 2, \dots, n_1]\} \quad (4)$$

其中: N 为金字塔层数。

Step3.2:以 corners2 中各点为中心,重新确定角点搜索窗口。由第 k 个点构成的窗口范围为

$$\left. \begin{aligned} 2^{N-1} i_k - W/2 \leq x \leq 2^{N-1} i_k + W/2 \\ 2^{N-1} j_k - H/2 \leq y \leq 2^{N-1} j_k + H/2 \end{aligned} \right\} \quad (5)$$

其中:角点搜索窗口的宽度为 $W+1$,高度为 $H+1$ 。

Step3.3:针对每个搜索窗口,按特定数量阈值 Q ,采用 Shi-Tomasi 方法搜索角点信息,并记录下来。具体

计算方法见 Step2.1-Step2.3。第 k 个窗口搜索到的角点坐标信息构成角点集 corners3_k 。

Step3.4: 将所有窗口搜索结果合成角点集 corners , 也就是最终结果

$$\text{corners}=\{\text{corners3}_k|k=[1,2,\dots,n_1]\} \quad (6)$$

2 实验结果及分析

为了验证算法的有效性, 本文进行了相关实验。实验主要分两步进行: 第 1 步是角点提取算法的比较。分别采用 Shi-Tomasi、Fast、Harris 算法和本文算法对相同图像进行角点提取, 用来反映四种算法在保证提取信息均匀化方面的性能; 第 2 步是采用不同的角点提取方法计算载体运动信息精确度的对比, 主要是将几种算法应用于视觉里程计, 计算载体运动信息, 用来观察本文角点提取方法对载体运动信息计算精确度的影响。

采用的硬件平台是 Windows10 64 位 NVIDIA GeForce MX150, 软件平台是 vs2017+opencv3.4, 用做实验的数据集采用通用的 KITTI^[16]数据集。

2.1 角点提取算法比较

图片选取的是 KITTI 数据集中一张典型的路面信息图, 环境单一, 路面纹理较弱, 图片中一般只有天空、草丛和树叶等信息。因为路旁树木繁茂, 纹理丰富; 路面光滑, 纹理较少, 两者反差较大, 再加上光线的不确定因素, 特征很难提取均匀。在高速公路这种典型的纹理分布不均匀场景中行驶, 树叶和天空形成的斑驳印记一般都是角点聚簇的区域, 存在大量相似、重复、聚集的角点, 这种类型的角点在前后帧追踪时难度较大。如遇转弯或角点聚集在图片边缘等情形, 前后帧画面差别较大, 大量聚集在边缘的角点易大量瞬间滑出画面, 无法保证图片角点个数稳定在一个常数周围。而且, 图像中的角点相对集中, 利用不够离散的角点求解位姿不能够代表相邻帧之间的位姿变化。

本文分别采用 Shi-Tomasi、Harris、Fast 以及本文算法对图片进行角点提取, 每种方法均提取 300 个角点, 其中本文算法参数为 $N=3, W=H=30, n_1=147, Q=2$ 。4 种算法角点分布情况见图 1~图 4。



图 1 Shi-Tomasi 算法提取结果
Fig.1 Extraction result of Shi-Tomasi algorithm



图 2 Fast 算法提取结果
Fig.2 Extraction results of Fast algorithm



图 3 Harris 算法提取结果
Fig.3 Extraction results of Harris algorithm



图 4 本文算法提取结果
Fig.4 Extraction results of the text algorithm

从图 1 与图 4 可以看出,Harris 与 Shi-Tomasi 算法角点聚簇现象比较明显,主要聚集在光线明亮、树叶、草丛等强纹理区域。Fast 算法则强烈依赖阈值,角点聚集程度很高,而本文算法检测到的角点,原本聚簇、相似的角点被分隔开来,原本光线较暗区域也有角点被提取出来,整张图片各个角落、各个区域都有角点分布,分布均匀、离散。

为验证本文算法可满足角点提取实时性要求,采用本文算法与 Shi-Tomasi 算法、Fast 算法、Harris 算法分别对 2 张图片进行不同角点阈值的特征提取,角点检测时间如下表 1 所示。计算中,感兴趣区域的获取以及每个区域进行二次角点提取是导致运行速度慢的主要原因。但是尺度化处理缩小了需处理图片的大小,这在一定程度上弥补了感兴趣区域耗时的缺点,满足实时性要求。

表 1 计算时间比较
Tab.1 Comparison of calculation times

图片信息	计算方法	角点个数	运行时间/s
图片 1 (1 241*376)	本文算法	220	0.23
	Shi-Tomasi 算法	220	0.20
	Fast 算法	220	0.76
	Harris 算法	220	0.19
图片 2 (1 241*376)	本文算法	500	0.27
	Shi-Tomasi 算法	500	0.23
	Fast 算法	500	0.80
	Harris 算法	500	0.21

为了检验本文算法对追踪信息的保持能力,采用 KITTI 数据集 01 序列与 10 序列,用金字塔 LK 光流法追踪前 20 帧图片,并与综合性能表现最好的 Shi-Tomasi 方法进行对比。金字塔 LK 光流法是在原光流法的基础上加入了图像金字塔算法,用于解决单纯光流法只适用于相机的微小运动,而在实际情况中经常出现较大的运动幅度的问题。金字塔 LK 光流法为常用的角点追踪算法。

01 序列与 10 序列采集的是转弯处数据集,前后帧画面差别较大。采用召回率来反应光流追踪性能,召回率=追踪角点个数/角点个数,角点个数即第一帧图片检测到的总角点个数,追踪角点个数为之后每一帧追踪到的角点个数。为反映召回率,没有设置追踪角点个数低于某个阈值即重新检测。召回率反映了角点的可区分性,召回率高说明角点可区分性好。

具体情况如图 5 所示,随着物体的运动,一些角点会逐渐跑出画面外,所以召回率是逐渐降低的。从图中可看出,均匀化后的本文算法召回率优于原 Shi-Tomasi 算法,更有利于维持不同图像特征点数量的稳定性。

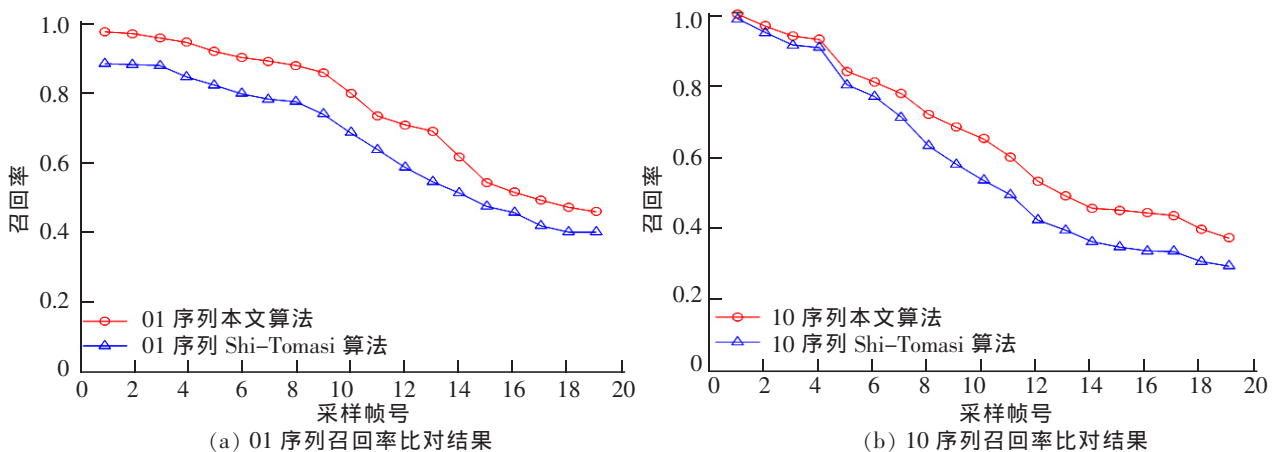


图 5 01 序列与 10 序列召回率计算结果

Fig.5 Recall rate calculation results of Sequence 01 and Sequence 10

2.2 载体运动信息计算精确度对比

为了比较角点检测结果对载体运动信息计算的影响,将本文方法用于视觉里程计计算。角点提取环节用本文算法与 Shi-Tomasi 算法、Harris 算法与 Fast 算法对比,然后用金字塔 LK 光流法追踪角点,最后根据相机对极几何约束模型恢复运动轨迹。求出前两帧图片之间的运动尺度 k 与相对位姿信息 (R_t, T_t) , 后续位姿信息通过 $P_t=R_{t-1}P_{t-1}+T_{t-1}k$ 计算得到。

数据集采用 KITTI 数据集 02 左序列前 500 张。其真实轨迹由 KITTI 数据集 groundtruth 提供。图 6 所示为分别采用 Shi-Tomasi 算法、本文算法、Harris 算法、Fast 算法对比所得到的角点信息恢复的运动信息。从实验结果中可看出,四种算法随着时间的推移轨迹都有一定程度的发散现象,这是由于算法本身只计算两帧图片之间的运动信息导致的误差累计。而本文算法计算得到的运动信息的精度优于另外三种算法。

为了更好的评测视觉里程计的效果,本文没有加入回环检测环节。如加入回环检测累计误差会小很多。同时,本文进一步计算了运动信息的误差曲线。如图 7 所示,误差曲线是计算实验值与真实值之间的欧氏距离。横坐标为图像采样帧号,纵坐标为欧式距离。从图中可以看出本文算法恢复的运动信息比 Shi-Tomasi 算法误差要小,恢复的位姿信息准确度更高。而且随着图像采样帧号的增加,因为本文算法的召回率表现更优良,故角点流失也保持在一个稳定的状态,因此误差曲线逐渐趋于稳定,恢复的位姿信息准确度更高。

3 结论

针对视觉里程计算法中因角点提取不均匀导致恢复的运动信息误差较大问题,本文提出了一种基于高斯金字塔的角点提取算法。实验表明,本文算法能够准确提取图片中的特征信息,有效解决重复纹理与丰富纹理特征聚簇问题,角点分布均匀、离散,提高了算法的鲁棒性。在同等条件下运行本文算法,在视觉里程计中可以得到更加准确的运动信息。

参考文献:

- [1] LI H, HU Z, CHEN X. Robust monocular visual odometry using optical flows for mobile robots[C]//Control Conference. IEEE, 2016.
- [2] 袁梦, 李艾华, 崔智高, 等. 基于改进的 3 维 ICP 匹配的单目视觉里程计[J]. 机器人, 2018, 40(1): 56-63.
- [3] DAVIDE S, FRIEDRICH F. Visual odometry[J]. Robotics & Automation Magazine IEEE, 2011, 18(4): 80-92.
- [4] 郑驰, 项志宇, 刘济林. 融合光流与特征点匹配的单目视觉里程计[J]. 浙江大学学报(工学版), 2014, 48(2): 279-284.
- [5] XU P, SU K, HONG C, et al. Simultaneous localization and mapping technology based on project tango[J]. Wuhan University Journal of Natural Sciences, 2019, 24(2): 176-184.
- [6] 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): 372-379.

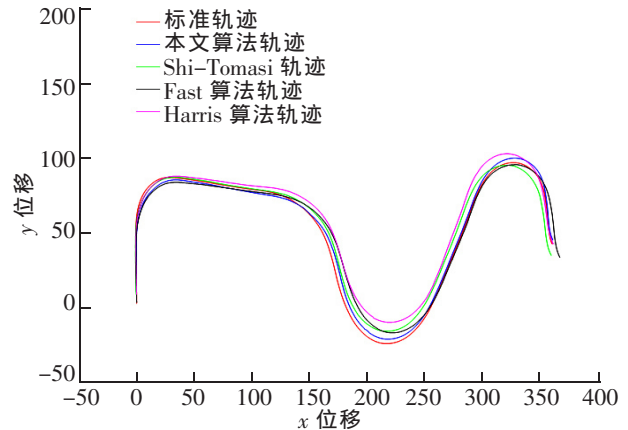


图 6 数据集 02 运动信息计算结果
Fig.6 Calculation results of motion information in Dataset 02

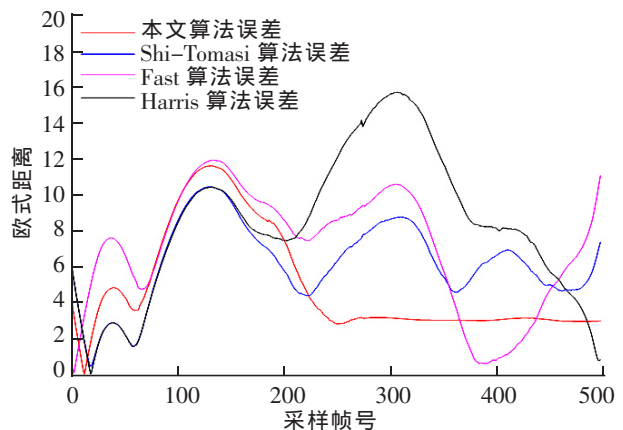


图 7 02 序列误差曲线
Fig.7 Error curve of Sequence 02

- [7] FRAUNDORFER F, SCARAMUZZA D. Visual odometry; Part II: Matching robustness optimization and applications[J]. IEEE Robotics & Automation Magazine, 2012, 19(2): 78–90.
- [8] 张易, 项志宇, 陈舒雅, 等. 弱纹理环境下视觉里程计优化算法研究[J]. 光学学报, 2018, 38(6): 226–233.
- [9] 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.
- [10] PIRE T, FISCHER T, CIVERA J, et al. Stereo parallel tracking and mapping for robot localization[C]//2015 IEEE/RSJ International Conference on Intelligent Robot and System (IROS), 2015: 1373–1378.
- [11] 张立亭, 黄晓浪, 鹿琳琳, 等. 基于灰度差分与模板的 Harris 角点检测快速算法[J]. 仪器仪表学报, 2018, 39(2): 218–224.
- [12] 彭真. 动态环境下基于视觉的自运动估计与环境建模方法研究[D]. 杭州: 浙江大学, 2017.
- [13] 赵泽壹. 无人驾驶汽车视觉里程计算法研究[D]. 西安: 陕西师范大学, 2018.
- [14] LAI W, HUANG J, AHUJA N, et al. Fast and accurate image super-resolution with deep Laplacian Pyramid networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 41(11).
- [15] 陶攀, 付忠良, 朱锴, 等. 空间金字塔分解的深度可视化方法[J]. 哈尔滨工业大学学报, 2017, 49(11): 60–65.
- [16] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: The KITTI dataset[J]. International Journal of Robotics Research, 2013, 32(11): 1231–1237.

Research on Visual Odometry Based on Gaussian Pyramid

Liu Rui, Xu Xuesong, Zeng Yu

(School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: Aiming at the problem that many feature extraction algorithms for visual odometry have large deviation of motion information due to uneven corner distribution, a corner extraction method based on Gaussian Pyramid is proposed. Firstly, the algorithm adopted the Gaussian Pyramid algorithm to compress the image scale, compressed the texture in the area of rich texture and aggregated texture in the area of sparse texture. Then, the small-scale top diagram was obtained and the Shi-Tomasi algorithm was used to extract the corners of the top graph for rough feature location. The information was mapped back to the original graph with rich details for accurate corner positioning, and the feature information of the picture was obtained. Finally, the Pyramid LK optical flow method was used to track the corner points, and the motion information is obtained according to the geometric constraint model of camera polar line. In this paper, the KITTI data set was used to compare with the original Shi-Tomasi, Harris, Fast Algorithms. The results show that the algorithm can effectively improve the uniformity of corner distribution and improve the accuracy of motion information recovery of visual odometry.

Key words: Gaussian Pyramid; rich texture; corner location; uniform distribution; visual odometry