

基于改进 SURF 算法的单目视觉里程计*

冉峰^{1,2} 李天² 季渊² 刘万林²

(1. 上海大学新型显示技术及应用集成教育部重点实验室 上海 200072;

2. 上海大学微电子研究与开发中心 上海 200072)

摘要: 针对传统单目视觉里程计在特征提取过程中误匹配点过多,匹配精度低、运算量大、提出了一种基于改进 SURF 算法的单目视觉里程计模型,首先使用 SURF 算法对单目摄像头采集的图像的相邻两帧进行特征点的检测与匹配,然后用 RANSAC 算法对误匹配点进行剔除,提高匹配的精度,减少运算量,最终求出相邻两帧图像特征点匹配的旋转矩阵 R 和平移向量 T ,完成运动估计。实验结果表明,该模型在预估曲线运动和直线运动时的运算速度分别提高了 11.2% 和 10.38%。

关键词: SURF 算法;RANSAC 算法;单目视觉里程计;旋转矩阵;平移向量

中图分类号: TP242;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Monocular visual odometry based on the improved SURF Algorithm

Ran Feng^{1,2} Li Tian² Ji Yuan² Liu Wanlin²

(1. Key Laboratory of New Display Technology and Application Integration of Ministry of Education, Shanghai University, Shanghai 200072, China; 2. Microelectronic R&D Center, Shanghai University, Shanghai 200072, China)

Abstract: In traditional monocular visual odometries, there are many false match points, low match accuracy, large amount of computation during the feature extraction process of the traditional monocular visual odometry. This paper presents an improved monocular visual odometry model based on the SURF algorithm. The feature points between two adjacent frames are detected and matched with the SURF algorithm. The RANSAC algorithm is applied to remove the error feature points. Then, the rotation matrix R and shift vector T between two adjacent frames are calculated to accomplish the motion estimation. The experiment results show that the computing speed is accelerated by 11.2% and 10.38% in the curve motion and the straight motion, respectively, with the proposed visual odometry estimating model.

Keywords: SURF; RANSAC; monocular visual odometry; rotation matrix; shift vector

1 引言

随着社会的发展,实现较为精确的自主定位显得愈发重要,而传统的定位方式受到的限制较多,比如 GPS 不能实现室内的定位,惯导价格昂贵。随着图像处理的不断发展,利用视觉里程计的方法实现定位已经获得了极大的发展。视觉里程计是仅利用单个或多个相机的输入信息估计运动信息的过程^[1],单目视觉里程计花费小的同时精度较高,应用相对比较广泛。视觉里程计具有精度较高,隐蔽性较好,不受地理和空间限制等优点。目前,视觉里程计主要的研究方法有两种,分别为基于特征跟踪和基于光流这两种方法^[2],随着目前特征点匹配方面的研究不断发展,使得基于特征点的视觉里程计成为目前较为主流的方法,但

是在特征点匹配过程中,由于外部环境的以及其他因素的影响,经常会出现误匹配的点多,这样会使得运算量过大,导致最终得到的定位精度不高,实时性较差。

本文研究了基于改进 SURF 算法的单目视觉里程计模型,首先采用 SURF 算法对采集图像的相邻两帧进行特征点的检测和匹配^[3],然后使用 RANSAC 算法对误匹配点进行剔除,构建单目视觉里程模型,求出匹配点的旋转矩阵 R 和平移向量 T ,完成运动估计^[4]。通过实验数据表明,本文所提出的单目视觉里程计模型具有定位精度较高,成本较低,应用价值较大等优点。

2 SURF 算法的特征提取与匹配

经常采用的特征点提取的算法有 SUSAN^[5]、Harris、

收稿日期:2016-10

* 基金项目:国家自然科学基金(60777018)资助项目

SIFT^[6], SURF^[7]等,但是 Harris 和 SUSAN 算法具有一定的缺陷,比如通过这两种算法提取出来的特征点,尺度和旋转不变性比较差。而 SIFT 算法具有的这两种特性比较好,但是这种算法的计算时间长,计算的复杂性较高^[8]。SURF 算法是 SIFT 算法的加速版,一般来说,标准的 SURF 算法比 SIFT 算法快好几倍^[9],而且对多幅图片进行检测时,该算法具有更好的稳定性。SURF 算法最大的特性在于采取了 harr 特征以及积分图像的观点,程序的运行速率获得了较大的改善。因而,本文决定采取 SURF 算法提取出具有高鲁棒性的特征点。SURF 算法主要由两部分组成,分别是特征点检测以及特征点匹配,其流程如图 1 所示。

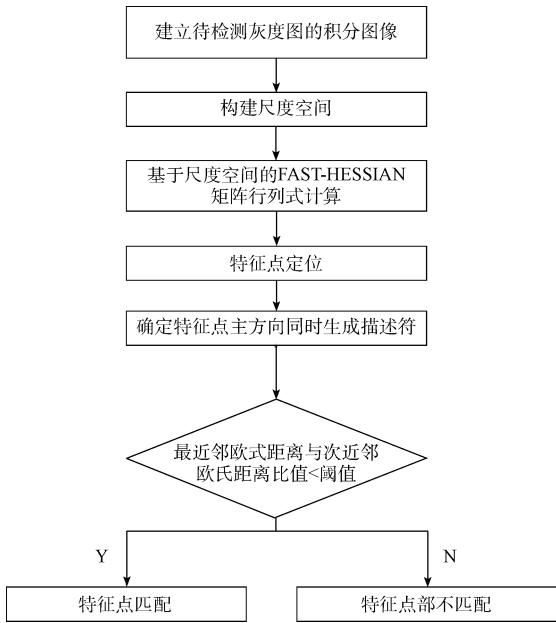


图 1 SURF 算法流程

2.1 特征点检测

该算法对特征点的检测是以 Hessian 矩阵为基础,从积分图像 I 中可以看出,尺度是 σ 的点 $X(x, y)$ 的 Hessian 矩阵定义为:

$$H(X, \sigma) = \begin{pmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{pmatrix} \quad (1)$$

式中: $L_{xx}(X, \sigma)$ 是高斯滤波二阶导数 $\frac{\partial^2 g(\sigma)}{\partial x^2}$ 与图像 I 的卷积,其中 $g(\sigma)$ 为:

$$g(\sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right) \quad (2)$$

其中: $L_{xy}(X, \sigma)$ 、 $L_{yy}(X, \sigma)$ 具有类似含义。

在具体的操作中,高斯二阶滤波的近似, SURF 算法使用的是盒子滤波器(box filter)来实现的。采用 D_{xx} 、 D_{xy} 、 D_{yy} 分别对应通过箱式滤波器得到的 L_{xx} 、 L_{xy} 、 L_{yy} 的近似,然后 Hessian 矩阵的行列式能够近似的计算成^[10]:

$$\det(H_{\text{approx}}) = D_{xx}D_{yy} - (\omega D_{xy})^2 \quad (3)$$

其中 ω 为权重系数,值为 0.9。根据式(3)的 Hessian 矩阵求出图像点在尺度 σ 处的极值。计算的结果若是负的话,则这个点不属于极值点,如果是正,那么这个点就是极值点。极值点确定之后,对该点周围 $3 \times 3 \times 3$ 的空间相邻域内作非极大值抑制,然后将该极值点和它邻近的 26 个点进行值的比较,假设所得到的极值点比其他 26 个点的响应值都要大的话,那么,认为所计算得到的极值点是区域的特征点。

对特征点的描述主要通过特征点主方向,让其能够保持一定的旋转不变性,与此同时一并生成特征点描述符。

2.2 特征点匹配

前后两帧图像的特征点得到之后,就需要从一帧图像的特征点序列中找到另一帧图像与之对应的特征点,此为特征点匹配定义,之后就可以获得这两帧图像之间的联系。

SURF 算法的特征点匹配以欧氏距离作为基准,欧氏距离指的是特征点描述符之间的距离,所采用的方式是最相邻匹配方法,如式(4)所示。从已经设定好的参考图像中依次选取每个点,根据式(4)来计算同待匹配图像中点的欧氏距离,根据所得到的计算结果不断的更新最小的欧氏距离,而更新的依据是最小的欧氏距离与次小距离的比值在某一指定的阈值范围之内,当参考图像里所有点都进行遍历之后,如果此时不再更新的话,那么就可以认为获得了正确的匹配特征点对。

$$Lis_{ij} = \left[\sum_{k=0}^{k=n} (X_{ik} - X_{jk})^2 \right]^{1/2} \quad (4)$$

式中: Lis_{ij} 指的是两个匹配点之间的欧氏距离, n 表示的是特征向量数, X_{ik} 和 X_{jk} 分别标的是待匹配图像以及参考图像中第 i, j 个特征点描述子的第 k 个元素。

3 RANSAC 算法进行误匹配剔除

特征点检测与匹配完成后,因为种种外部环境的影响,会存在一些误匹配点对,决定采取 RANSAC (random sample consensus)算法对个中的误匹配点进行剔除^[11]。

RANSAC 算法假设采集的样本里不仅有合理的数据,也有异常的数据。异常的数据可能是来自于外部环境,测量误差,计算错误等因素产生的。在本文中,出现异常的数据主要来源于 SURF 算法的误匹配。当然除了这个之外,这种算法还假定了,比如现在有一组合理的数据,是通过计算得到匹配该数据的特定模型参数的方式。

RANSAC 的主要算法过程描述如下:

1) 对一个最小抽样集合的势为 n (最小样本数) 的模型以及样本集合 P 进行分析,集合 P 的样本数是比 n 大的,从 P 里随机的抽出具有 n 个样本的 P 的子集和 S ,用于初始化模型 M 。

2) 余集 $SC = P/S$ 中与模型 M 的误差小于预定阈值的 t 的样本集合与 S 共同组成 S^* 。则把 S^* 认为是一个内

点集,而它们整体构成 S 的一致集。

3)如果 $\#(S^*) \geq N$,那么就认为获得了正确的模型参数,之后集合 S^* 采用最小二乘法重新计算新的模型 M^* ,重新随机的抽取新的 S ,不停的重复上述过程。

4)当到达一定的抽样次数之后,若是没有在找到一致集就认为是失败的,不然选择抽样之后获得的最大一致集来判断是否为内外点,此时该算法运算结束。

4 单目视觉里程计的定位

单目视觉里程计主要指的是通过单目摄像机得到图像集合,通过特征的提取、匹配以及坐标的转化之后,实现运动的估计,图 2 所示为单目视觉里程计的设计流程。

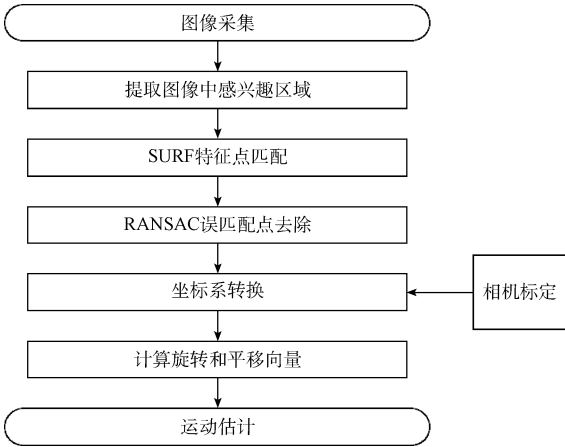


图 2 视觉里程计设计流程

特征点匹配且通过标定之后,可以得到匹配点在世界坐标系下的坐标,因而需要把空间坐标系转变成运动物体坐标系来估计人的运动参数。

为了能够得到行人的运动轨迹,这时候需要引入人体坐标系,这样有利于完成人体的定位任务。首先,人在行进过程中,在两个不同地点进行图像采集,会产生两组坐标系,两组坐标系之间只含有旋转和平移。

如图 3 所示,点 P 为三维世界坐标系下的一个点,在某一个观测点的坐标是 p_1 ,当运动到另一个观测点时,在第二个坐标系下面,观测到 p 点的坐标是 p_2 ,这时可以通过对旋转和平移向量计算,实现对整个运动过程进行描述:

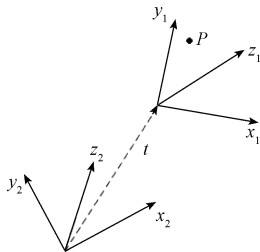


图 3 帧间坐标系转换

$$p_2 = Rp_1 + T \tag{5}$$

式中: R 表示坐标系之间的旋转, T 表示坐标系之间的平移,通过 R 和 T 来完成运动估计。采用最小平方中值定理对运动信息进行估计。

令 r_i 为第 i 个数据的残差,即 $r_i = \|P_{e_j} - (RP_{e_j} + T)\|$ 。最小平方中值定理可以通过下式来完成参数的估计:

$$(R, T) = \min med r_i^2 \tag{6}$$

式中: r_i 是回归方差, $med(\cdot)$ 是中值运算。式(6)的求解步骤如下。

1)通过使用蒙特卡罗法在 n 个匹配对里随机性的抽出 m 个样本,每一个样本是由 q 个点匹配而成的。此处取 $q=7$ 。

2)可以使用奇异值分解的方法计算其中一个样本的运动 $E_j = (R_j, T_j)$ 。

3)对于 E_j ,计算出这个点匹配集的回归方差,同时找出中值 $M_j = med r_i^2 (i = 1, \dots, n)$ 。

4)重复步骤 3) m 次,从 m 个中值中找到最小的中值 M_{min} 以及它对应的 E_j ,然后即可得到最终的运动估计 (R, T) 。

5 实验结果

图 4 所示为本次实验所采用的硬件平台,单目 CMOS 摄像头用于视频的采集,采集的图像的像素大小为 720×480 ,便携式计算机用于对视频进行处理,型号为三星 R439,CPU 频率为 2.4 GHz,运行内存为 2 G,软件编程环境采用 VS2010,OpenCV。OpenCV 是一个开源的计算机视觉库,由一系列 C 函数和 C++类所组成的库构成。

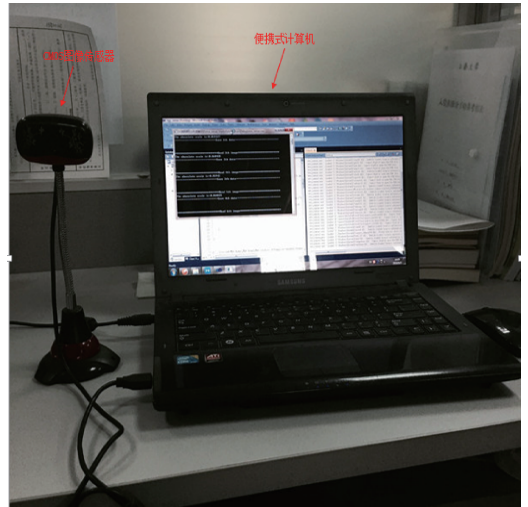


图 4 实验平台

由于运动主要分为两种,一种是直线运动,一种是转弯运动,因此本次实验主要对这两种运动进行研究。图 5 所示为图 6 转弯运动中 A 点到 B 点 15 帧之间图像减少的误

匹配点和匹配减少时间之间的关系。图 5 中第一个点表示 A 点到 B 点这 15 帧图像中第 1 帧和第 2 帧之间剔除的匹配点数量以及对应匹配减少的时间,第二个点表示第 1 帧到第 3 帧之间剔除的匹配点数量以及对应匹配减少的时间,以此类推,最后一个点表示第 1 帧到第 15 帧之间剔除的误匹配点总和和以及对应的匹配减少时间,减少的总时间为 0.661 s,经过测试,这 15 帧图像处理的总时间为 5.9 s,减少时间占 15 帧运行时间的 11.2%。图 6 则表示了转弯运动的定位结果比较。真实轨迹为采用 GPS 得到的真值。从该图中可以看出转弯运动时,曲线发生了一些偏离,这主要是由于大角度转弯时,特征匹配的数量减少,会产生偏差,当转弯运动结束时,曲线及时的进行了矫正,基本沿着真值定位曲线运动。图 7 和 8 所示为直线运动时的曲线图,分析同曲线运动一样,图 8 中从 C 点到 D 点 15 帧图像减少的时间总和为 0.405 s,C 点到 D 点运行的总时间是 3.9 s,减少时间占 15 帧运行时间的 10.38%。直线运动中的实验轨迹与真实轨迹基本重合。通过以上分析,说明改进 SURF 算法可以有效的提高单目视觉里程计的匹配精度和实时性。

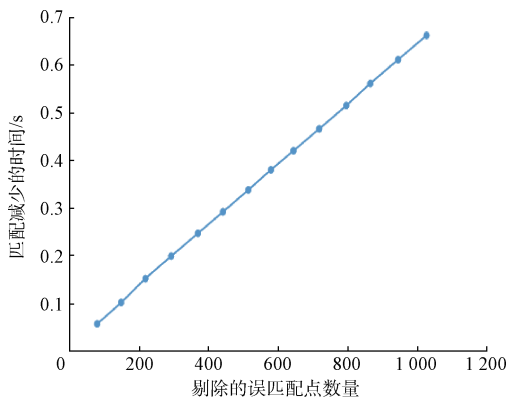


图 5 A 点到 B 点剔除误匹配点数量与匹配减少时间关系

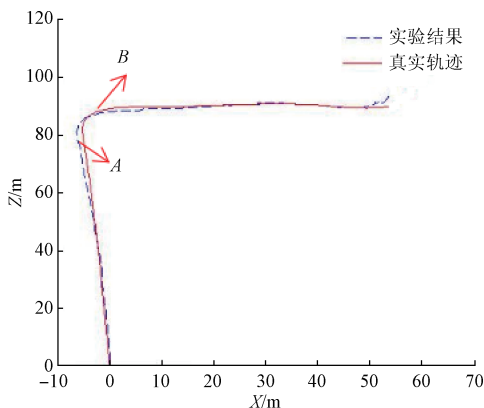


图 6 转弯运动定位结果比较

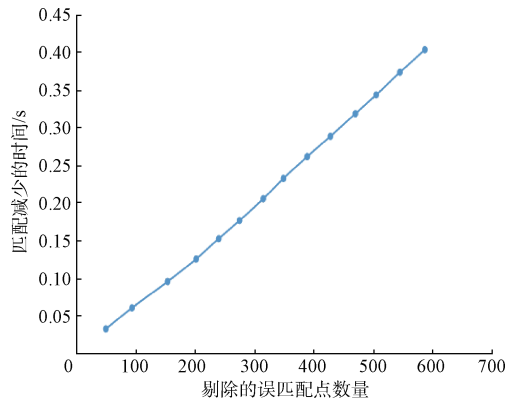


图 7 C 点到 D 点剔除误匹配点数量与匹配减少时间关系

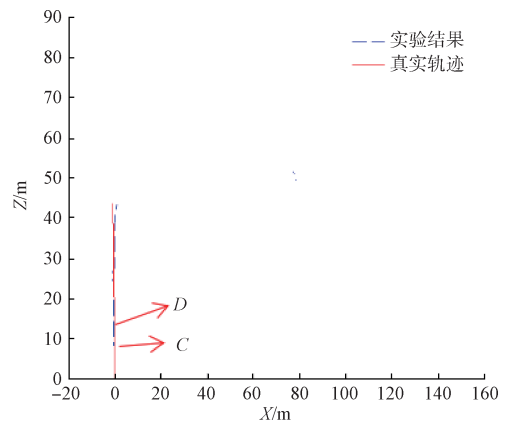


图 8 直线运动定位结果比较

6 结 论

本文提出了一种改进 SURF 算法的单目视觉里程计模型,采用 RANSAC 算法对特征点匹配过程中的误匹配点进行剔除,实验结果表明本模型具有较高的定位精度,同时有效的减少了匹配的时间,曲线运动和直线运动减少的时间比分别为 11.2%和 10.38%,提高了模型的实时性,而且所需硬件设备的成本极低,仅需普通的单目图像传感器即可,而随着目前虚拟现实和增强现实的火爆,该模型可以方便的应用于可穿戴设备,增强现实定位系统等,具有较广阔的应用前景。

参 考 文 献

- [1] SCARAMUZZA D, FRAUNDORFER F. Visual odometry [J]. IEEE Robotics & Automation Magazine, 2011, 18(4):80-92.
- [2] 阳吉斌,胡访宇,朱高. 基于改进 SURF 算法的遥感图像配准[J]. 电子测量技术, 2012, 35(3):69-72.
- [3] 傅为平,秦川,刘佳,等. 基于 SIFT 算法的图像目标匹配与定位[J]. 仪器仪表学报, 2014,31(4):45-49.

(下转第 200 页)