

双目立体视觉中的图像匹配方法研究

原思聪, 刘金颂, 张庆阳, 江祥奎

YUAN Si-cong, LIU Jin-song, ZHANG Qing-yang, JIANG Xiang-kui

西安建筑科技大学 机电工程学院, 西安 710055

College of Mechanical and Electronic Engineering, Xi'an University of Architecture & Technology, Xi'an 710055, China

E-mail: ysc_3c@163.com

YUAN Si-cong, LIU Jin-song, ZHANG Qing-yang, et al. Research on image matching method in binocular stereo vision. Computer Engineering and Applications, 2008, 44(8): 75-77.

Abstract: The image matching technique based on Scale Invariant Feature Transform is studied, and the fake matching points are eliminated effectively by adding epipolar restriction in this method. A new matching method which combines feature matching and area matching, edge feature and corner feature, is depicted, and the experiments show that the matching time can be reduced, and precision can be improved by using this method.

Key words: binocular stereo vision; image matching; Scale Invariant Feature Transform (SIFT); corner detection

摘 要: 对目前匹配能力很强的基于 SIFT 特征的图像匹配方法进行研究, 并在该方法中加入极线约束, 有效去除了大部分虚假匹配。提出以特征匹配与区域匹配相结合、边缘特征与角点特征相结合的立体匹配方法。实验证明该方法不仅能够有效地缩短匹配时间, 还能达到较高的匹配精度。

关键词: 双目立体视觉; 图像匹配; SIFT; 角点检测

文章编号: 1002-8331(2008)08-0075-03 **文献标识码:** A **中图分类号:** TP391

1 引言

图像匹配是立体视觉中最重要、最困难的问题, 也是三维信息恢复的先决条件, 一直是立体视觉领域研究的焦点。目前, 还没有一种方法能够完美地解决图像对应点匹配问题。因此, 如何实现多幅图像间快速、准确地匹配对立体视觉的研究具有重要意义。基于 SIFT^[1]特征的匹配算法是目前国外图像匹配研究领域的热点与难点, 其匹配能力较强, 可以处理两幅图像之间发生平移、旋转、仿射变换情况下的匹配问题, 甚至在某种程度上对任意角度拍摄的图像也具备较为稳定的特征匹配能力。国内对 SIFT 特征匹配算法的研究比较少, 应用也比较少。本文在分析 SIFT 特征匹配算法主要过程的基础上, 将极线约束加入其中, 滤除了大部分虚假匹配, 改善了匹配效果。但在实验过程中发现, 很多 SIFT 特征点并不是所想要的特征点, 既不是边缘点, 也不是角点。为此, 本文提出以特征匹配与面积匹配相结合、边缘特征与角点特征相结合的图像匹配方法。实验证明, 该方法不仅能够有效地缩短匹配时间, 还能达到较高的匹配精度。

2 SIFT 匹配算法主要思想与步骤

基于 SIFT 特征的图像匹配算法由 Lowe D G 1999 年提出, 2004 年完善总结, 其全称是 Scale Invariant Feature Trans-

form, 即尺度不变特征变换。SIFT 算法是一种提取局部特征的算法, 在尺度空间寻找极值点, 提取位置、尺度、旋转不变量, 生成关键点特征描述符, 然后根据这些不变量特征进行匹配。SIFT 算法的主要步骤如下:

(1) DOG 尺度空间的构建: 为了有效地在尺度空间检测到稳定的关键点, Lowe 提出了高斯差分尺度空间 (DOG scale-space), 利用不同尺度的高斯差分核与图像卷积生成。即:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1)$$

DOG 算子计算简单, 是归一化 LoG (Laplacian-of-Gaussian) 算子的近似。

(2) 尺度空间极值点检测: 对于图像上的点, 计算其在每一尺度下 DOG 算子的响应值。为了寻找尺度空间上的极值点, 每一个采样点要和它所有的相邻点比较, 看其是否比它的图像域和尺度域的相邻点大或者小, 即每个检测点要和它同尺度的 8 个相邻点和上下相邻尺度对应的 9×2 个点共 26 个点比较, 以确保在尺度空间和二维图像空间都检测到极值点。

(3) 精确确定极值点位置: 通过拟和三维二次函数以精确确定关键点的位置。所用的拟合函数为 DOG 算子的二阶泰勒展开式。并通过检测高斯差分算子的主曲率来滤除边缘响应。

(4) 关键点方向分配: 利用关键点邻域像素的梯度方向分布特性为每个关键点指定方向参数, 使算子具备旋转不变性。

作者简介: 原思聪 (1955-), 男, 教授, 博士生导师, 主要研究方向为现代设计理论与方法、计算机辅助技术; 刘金颂 (1981-), 女, 硕士, 助理工程师, 主要研究方向为智能测控技术; 张庆阳 (1981-), 男, 硕士研究生; 江祥奎 (1979-), 男, 博士研究生。

收稿日期: 2007-07-05 **修回日期:** 2007-10-25

(5)特征点描述符生成:实际计算过程中,为了增强匹配的稳定性,Lowe 建议对每个关键点使用 4×4 邻域内共 16 个种子点来描述,这样对于一个关键点就可以产生 128 个数据,即最终形成 128 维的 SIFT 特征向量。

当两幅图像的 SIFT 特征向量生成后,下一步采用关键点特征向量的欧氏距离来作为两幅图像中关键点的相似性判定度量。取左摄像机图像中的某个关键点,并找出其与右摄像机图像中欧式距离最近的前两个关键点,在这两个关键点中,如果最近的距离除以次近的距离少于某个比例阈值,则接受这一对匹配点。降低这个比例阈值,SIFT 匹配点数目会减少,但匹配更加准确,这样增加匹配点数目与提高匹配准确率之间就存在着矛盾,结果导致要得到一定数量的匹配点,总会有一些虚假匹配点存在。为此,本文再将极线约束加入其中。

3 极线约束中的基本矩阵 F

极线约束是双目立体视觉系统的一个重要性质,对于双目立体视觉中的对应点匹配具有重大贡献。要用极线约束,就必须计算基本矩阵。由文献[2]可知:设 p_l, p_r 是空间点 P 在两个图像上的投影点, p_l, p_r 是其对应的图像齐次坐标,则 p_l 在图像 I_l 上对应极线方程为: $p_r^T F p_l = 0$, 同样有 p_r 在图像 I_r 上的极线方程 $p_l^T F^T p_r = 0$, F 即为基本矩阵。

如果已知左右摄像机的内参数矩阵 A_l, A_r 和两个摄像机之间的结构参数,旋转矩阵 R 和平移向量 t ,则基本矩阵 F 可表示为:

$$F = A_r^{-T} [t]_x R A_l^{-1} \quad (2)$$

式中 $[t]_x$ 为平移向量 t 的反对称矩阵。

从式(2)可以看出,基本矩阵 F 实际上包括双目立体视觉系统的所有参数,即两摄像机的内参数 A_l, A_r 和视觉系统的结构参数 R, t 。这表明 F 只与视觉系统的参数有关,与外部场景无关,是双目立体视觉内在的一种约束关系。 F 除可用式(2)求出外,还可以用 8 点算法求出。

4 匹配实验

运用 MATLAB7.1 对前面所述的 SIFT 匹配算法进行编程实验。图 1 为左摄像机拍摄的图像检测出的 847 个关键点,箭头表示 SIFT 特征向量,右图像检测出 956 个关键点。

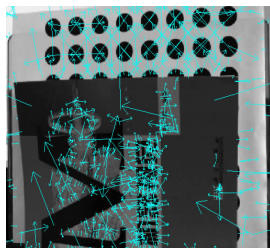


图1 检测出的关键点的 SIFT 特征向量

这里采用 8 点算法来计算 F 。可以随机选取 8 对以上匹配点来计算 F 。但是,采用随机选取的办法,有可能使选取的 8 个点集中在某一个区域内,也可能选中初始匹配中的虚假匹配点,那么,计算出的 F 就会有较大误差。这里采用交互方式,用鼠标在初始匹配点中,选取 8 对以上匹配点进行计算 F ,选取的匹配点应尽量均匀分布在初始匹配点区域内。选取 13 对初

始匹配点计算出的基本矩阵 F 为:

$$F = \begin{bmatrix} -0.000 & 0 & -0.000 & 0 & 0.000 & 3 \\ -0.000 & 0 & 0.000 & 0 & -0.005 & 1 \\ 0.000 & 3 & 0.005 & 8 & -0.216 & 2 \end{bmatrix}$$

计算出了基本矩阵 F , 就可以得到极线方程 $p_r^T F p_l = 0$, 但是,由于误差的存在,左右图像中的匹配点并不能完全满足该极线方程,也就是不一定就在极线上,而是与极线有一段很小的距离,那么判断点是否在极线上就变成判断点与极线的距离是否小于某个阈值。图 2 是经过极线约束后的匹配结果,共 263 对匹配点,为使图像不至于太模糊,图中仅画出了所有匹配点对应极线总数量的 1/8 部分。椭圆内的点为极线约束剔除的虚假匹配点,对应虚假匹配点用直线连接。

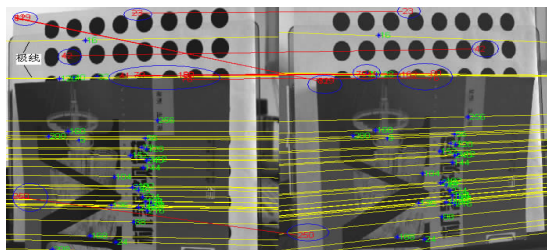


图2 经过极线约束后的匹配结果

经过匹配实验可知,在基于 SIFT 特征的匹配方法中加入极线约束能有效剔除大部分虚假匹配。但是,从本次实验中,也能看出 SIFT 方法的不足。从图 1 中可以看出,图像中的上方圆列部分检测出的关键点也不少,但是由图 2 看出,这一部分匹配成功率比较低,而在下方的图像部分,关键点数量很充足,匹配成功率也比较高,这说明该方法对于具有较强特征区别的图像非常有效,但是对于不具有明显特征区别的图像匹配效果不是很好;再则就是成功匹配点很多不是立体视觉中所需要的特征点(如边缘点、角点)。为此,提出以特征匹配与区域匹配相结合、边缘特征与角点特征相结合的图像匹配方法。也就是下面将要论述的基于 Harris 角点特征的匹配方法。

5 基于 Harris 角点特征的匹配方法

5.1 基于 Harris 角点特征的匹配方法主要过程

Harris 在文献[3]中提出了一个用于角点提取的“Corner detector”,被称为 Harris 角点检测算法,该算法以其较好的稳定性和方便性得到广泛应用。这里,考虑以 Harris 角点作为对象点进行匹配。提取左右两幅图像中的 Harris 角点特征,计算左图像以某一特征点为中心的一个小区域内的像素与另一幅图像中各个潜在对应特征点为中心的同样的小区域内的像素之间的相关值,具有最大相关值的特征就是匹配特征。但是仅仅以此进行匹配,是非常耗时的。由第 3 章可知,双目立体视觉系统经过标定之后,两个摄像机的内部参数以及视觉系统的结构参数已知,就可以按式(2)计算出基本矩阵,即能够获得双目立体视觉系统的极线约束。这样,匹配点的搜索范围大大减少,仅与对应极线的两旁一个窄带内的候选特征点计算相关值。考虑角点都是在边缘上,因此,先进行 Canny 边缘检测,在边缘检测的基础上进行角点检测。实验证明,该方法可以节约很多时间。

采用归一化的相关系数作为两特征点的相似性判定度量,来提高匹配算法的稳定性。考虑两幅图像 f_1 和 f_2 , 设待匹配的

候选特征点对的视差为 (d_x, d_y) , 则以特征点为中心的区域之间相似性测度可由归一化相关系数 $r(d_x, d_y)$ 定义为:

$$r(d_x, d_y) = \frac{\sum_{(x,y) \in S} [f_1(x,y) - \bar{f}_1][f_2(x+d_x, y+d_y) - \bar{f}_2]}{\sqrt{\sum_{(x,y) \in S} [f_1(x,y) - \bar{f}_1]^2 \sum_{(x,y) \in S} [f_2(x+d_x, y+d_y) - \bar{f}_2]^2}}^{1/2} \quad (3)$$

这里的 \bar{f}_1 和 \bar{f}_2 是被匹配的两个区域中的所有像素灰度平均值, 和式符号是在以特征点为中心的小窗函数中的所有像素上进行的。具体过程如图3所示。

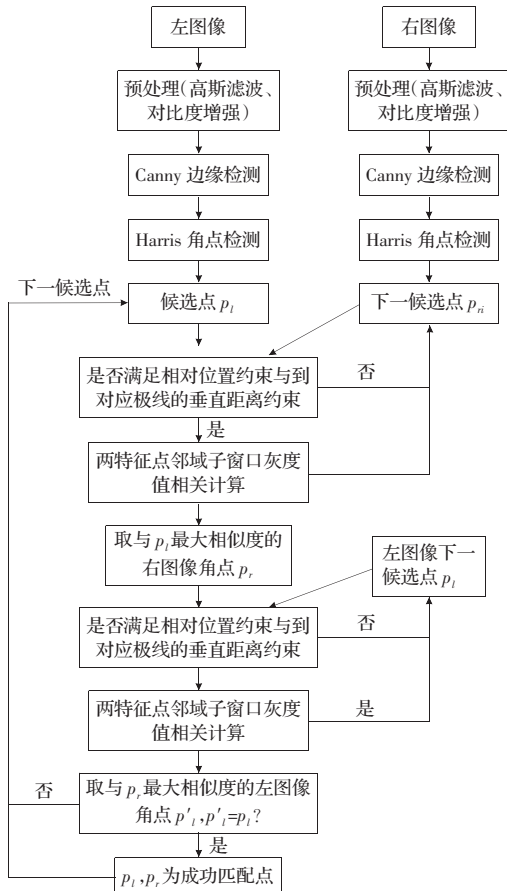


图3 基于 Harris 角点的匹配过程

首先对左右图像进行平滑滤波、对比度增强, 然后用 Canny 边缘检测器提取被测物体的边缘, 在边缘基础上提取 Harris 角点。基于相对位置约束(若右图像中特征点 p_r 与左图像中特征点 p_i 为正确匹配点对, 则在右图像中 p_r 应该在以 p_i 坐标为中心的的一个子区域内)、极线约束建立初始候选匹配关系, 然后基于区域匹配方式对特征点邻域内的灰度值进行相似性测度, 并进行对称性测试, 将最后的匹配对应点作为正确的匹配特征点。

5.2 匹配实验

下面按照 5.1 节所述方法, 对两摄像机拍摄的图像对中的角点进行匹配。图 4 为用 Canny 边缘检测器检测出的边缘(右

图像)。用 Harris 角点检测算法对图像进行角点检测, 左图像检测出 698 个角点, 右图像检测出 647 个角点, 图 5 为右图像角点检测结果。图 6 为匹配结果, 共获得 547 对匹配点, 对应匹配点用直线连接。为了进行比较, 同样对两幅图像用前述基于 SIFT 特征的匹配方法进行匹配, 左图像中检测出 2 269 个 SIFT 特征点, 右图像共检测出 1 919 个 SIFT 特征点, 未加入极线约束前获得 555 对匹配点, 加入极线约束后, 获得 541 对匹配点。

由实验可知, 两种匹配方法得到匹配点数量差不多。但是, 满足 SIFT 特征的关键点有些时候并不是所想要的特征点, 既不是边缘点又不是角点, 并没有太大意义, 而这里想要的特征点很多却没有匹配出来。相反, 基于角点特征的匹配方法, 以角点作为匹配基元, 能够获得立体视觉中用于三维信息获取的重要的图像特征。

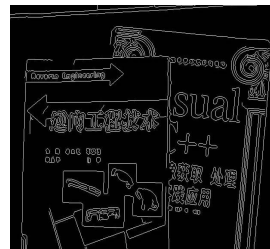


图4 Canny 边缘检测
结果(右图像)



图5 Harris 角点检测
结果(右图像)



图6 基于 Harris 角点特征的匹配结果

6 结论

基于 SIFT 特征的匹配方法对有很多区别性特征的图像具有很强的匹配能力, 但是其检测出的特征点很多并不是立体视觉匹配所需的特征点。本文提出的基于角点特征的匹配方法可以弥补这一缺点, 并且可以获得大量的角点特征, 但是其对缩放、旋转的适应能力不如 SIFT 算法。实际应用中, 根据具体情况决定选用哪种算法。可以考虑将基于角点特征或边缘特征的匹配方法和基于 SIFT 特征的匹配方法相结合, 两者互补长短, 就应该能够取得更好的结果, 这有待进一步的研究。

参考文献:

- [1] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [2] 张广军. 机器视觉[M]. 北京: 科学出版社, 2005.
- [3] Harris C, Stephens M. A combined corner and edge detector[C]// Proceedings 4th Alvey Vision Conference, 1988: 189-192.