

基于离线标定的全景视频快速拼接算法设计

石绍鹏¹, 王 军^{1,2}

¹苏州科技大学电子与信息工程学院, 江苏 苏州 215009;

²中国科学院长春光学精密机械与物理研究所, 长春 130033

摘 要: 针对实际条码识别系统中全场景监控和传统图像拼接算法速度慢的问题, 提出了一种基于离线标定的快速全景视频拼接算法。在实际应用中多台相机位置固定, 采用离线标定计算出图像拼接的单应性矩阵, 在实时拼接中直接加载该矩阵进行计算, 从而省去了大量的特征提取和配准时间。为了提高图像特征的配准精度, 设计了一种改进的 SIFT (Scale-Invariant Feature Transform) 特征匹配算法。实验结果表明, 该算法对比传统算法显著减少了错误特征匹配的影响, 提高了图像拼接速度, 具有更好的特征匹配效果。

关键词: 图像拼接; 全景监控; 特征提取; 图像融合

中图分类号: TN27 **文献标识码:** A **doi:** 10. 14016/j.cnki.jgzz. 2021. 01. 113

Design of panoramic video fast stitching algorithm based on offline calibration

SHI Shaopeng¹, WANG Jun^{1,2}

¹School of Electronic and Information Engineering, Suzhou University of Science and Technology, Suzhou Jiangsu 215009, China;

²Fine Mechanics and Physics, Changchun Institute of Optics, Chinese Academy of Science, Changchun 130033, China

Abstract: A fast video stitching algorithm is proposed based on the problem of slow speed of panoramic monitoring and traditional image stitching algorithm in actual barcode recognition system. In practical applications, multiple cameras are fixed in position, and the homography matrix of image stitching is calculated by offline calibration. The matrix is directly loaded in real-time stitching for calculation, thereby a lot of time feature extraction and feature matching time are eliminated. In order to improve feature matching accuracy, an improved SIFT (Scale Invariant Feature Transform) feature matching algorithm is designed. The results show that the proposed algorithm can significantly reduce the influence of error feature matching, improve the image stitching speed and have better feature matching effect.

Key words: image stitching; panoramic monitoring; feature extraction; image fusion

1 引言

图像拼接技术是将一组具有一定重叠部分的图像进行配准、融合之后生成一张宽角度、高分辨率图像的技术。图像拼接技术广泛应用于安防监控、战场态势感知、虚拟现实技术和医学图像等多种领域。实际应用场景中, 由于设备的限制很多情况下需要通过图像拼接来得到一幅高质量的广角图像。

传统图像拼接过程一般包含特征提取、图像配

准、图像映射和图像融合等步骤^[1]。目前主流的特征提取算法有 SIFT^[2]、SURF^[3]、Harris^[4]等算法。SIFT 算法对图像缩放、旋转甚至仿射变换均能保持不变性。刘文亮等人将 LPB 特征引入 SIFT 特征点以进行特征点提取^[5], 采用隔点计算特征点周围像素点的旋转不变均匀 LBP 特征值, 并替代一阶梯度幅值生成改进的 SIFT 特征点描述向量。赵岩等人提出一种结合投影误差校正的快速 SIFT 图像拼接方法^[6]。SURF 特征提取算法基于 Hessian 矩阵, 通过积分图像减少运算时间, 在图像平移、旋转、缩放和光照变化等方面都具有较好的鲁棒性^[7]。Harris 算法通过提取图像焦点作为特征, 具有速度快、稳定性高等优点, 但是 Harris 算法对尺度敏感, 需要人为设定阈值^[8]。许佳佳通过构建高斯尺度空间, 提取具有尺度不变性的角

收稿日期: 2020-09-15

基金项目: “十三五”江苏省重点学科项目 (No.20168765)

作者简介: 石绍鹏 (1994-), 男, 硕士生, 主要研究数字图像处理 and 光电信息检测。E-mail: 936047044@qq.com

通讯作者: 王军 (1979-), 男, 副教授, 硕士生导师, 主要研究图像处理技术与光电测控技术与仪器。E-mail: wjyhl@126.com。

<http://www.laserjournal.cn>

点特征改善 Harris 算法^[8-9]。冯宇平等采用阈值排序的方法寻找局部最佳阈值,最后通过加权得到图像的总阈值,实现 Harris 算子对阈值的自适应^[10]。Harris 算法虽然特征提取效果较好,但是计算量较大。主要的图像融合算法有加权平均融合、泊松图像融合、最佳缝合线等算法^[11-22]。罗永涛等人提出改进渐入渐出的图像拼接算法改善渐入渐出图像融合算法,将重叠区域划分 3 部分,对不同部分采取不同的融合方式实现对加权平均融合算法的优化^[23]。谷雨等人提出最佳缝合线和多分辨率融合算法,先得到最佳缝合线区域,再在该区域内通过图像金字塔实现对重叠区域的融合,对重叠区域内物体具有更好的过渡效果^[24],但是该方法计算量仍然较大,很难满足实时图像融合的需求。

传统的图像拼接算法是基于输入的每一帧图像进行特征提取和计算配准的拼接方法。具体步骤是先提取特征点,然后再进行特征匹配与融合等步骤。但是对于实时的图像拼接场景,传统方法需要消耗大量的时间,因此,若是对每一帧图像都进行特征提取和匹配,是无法满足实时拼接的要求。考虑到实际应用场景中相机的角度、位置和重叠的区域固定,所以采用离线标定的方法。在离线标定阶段计算出相机的映射关系矩阵,然后保存到本地。在每次拼接的时候加载映射矩阵,无需再重复进行特征提取和匹配,可以省去拼接过程中大量的前期工作,直接进入图像对齐、融合步骤。采用效果最好的 SIFT 算法,并改进 SIFT 算法,提出通过二次 NCC 提高特征点匹配精度的方法。最后采用图像融合速度最快的加权平均融合算法实现视频图像快速拼接。

2 离线标定

2.1 SIFT 算法原理

David G. Lowe 教授在 1999 年在总结基于不变量特征的特征检测方法的基础上,提出了 SIFT 算法。SIFT 算法可以保证特征点的提取不受光照、仿射变换和噪声等外在因素的影响。SIFT 算法利用一组连续的高斯卷积核与原图像进行卷积,生成一系列尺度空间的图像,并在所有尺度上识别对于尺度和旋转不变的兴趣点。在每个候选位置上,通过曲面拟合的方法确定关键点的位置和尺度,并剔除一些对比度较低的点。SIFT 算子在局部梯度方向为每个关键位置分配一个或多个方向,之后对图像的操作都相对于关键点方向、尺度和位置进行变换。SIFT 特征描述符通过这种梯度位置和方向的量化使得对 SIFT 算子对图像中细小的几何畸变和外在因素的变化具有良好的抗干

扰性。

2.1.1 SIFT 特征点检测

基于 SIFT 算子的特征点检测,可以通过改变空间尺度的参数找到不同的特征点,这代表在不同尺度空间选取的特征点是独立的,从而使得特征点满足视觉不变性。空间尺度 $L(x, y, \sigma)$ 的表达式为:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\frac{m}{2})^2 + (y-\frac{n}{2})^2}{2\sigma^2}} \quad (1)$$

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2)$$

其中, $G(x, y, \sigma)$ 为高斯函数, σ 为尺度空间因子。尺度空间点的极值点就是在高斯差分金字塔中检测的特征点,比较不同尺度下图像之间的差别寻求稳定的特征点。

2.1.2 SIFT 特征匹配

传统 SIFT 特征匹配使用最近邻算法,通过计算特征点之间的欧式距离判断相似度。计算基准图像中某一个特征 R_i 和另外一个图像中所有 n 个特征 $S_1, S_2, \dots, S_n$ 的欧式距离 $d_1, d_2, \dots, d_n$ 后,找出最小值 d_k 和次小值 d_p ($1 \leq k \leq n, 1 \leq p \leq n, k \neq p$)。当 $d_k/d_p < T$ 时,则认为 R_i 和 S_k 匹配, T 为阈值。

假设 R_i 和 S_k 是一对配准点的 SIFT 描述符, SIFT 特征向量为 128 维,即包含 128 个数据^[8],其中 $R_i = (r_{i1}, r_{i2}, \dots, r_{i128})$, $S_k = (s_{k1}, s_{k2}, \dots, s_{k128})$, 则 R_i 和 S_k 的欧氏距离计算公式如下:

$$d(R_i, S_k) = \sqrt{\sum_{i=1}^{128} (r_{ij} - s_{kj})^2} \quad (3)$$

2.2 改进的 SIFT 特征配准算法

传统的 SIFT 算法虽然能够提取出大量的特征点,但存在错误匹配的情况,由于这种错误的存在,会造成拼接后的图像无法达到理想的效果,因此必须去除错误匹配点。使用了一种利用灰度信息的归一化互相关 NCC (Normal Cross Correlation) 技术,在 SIFT 算法配准的基础上再一次进行配准和提纯。由于 2 台相机拍摄是在同一地点和时间,所以对于一个正确的 SIFT 匹配点对,在图像中同样大小的模板上的灰度值应该非常接近。因此,可以通过计算归一化互相关系数来剔除错误匹配点。

NCC 算法在图像灰度值变化和图像几何畸变不大的情况下,算法具有较高的匹配精度。NCC 算法是一种遍历算法,因此计算量很大。但本文使用的是离线标定,而配准过程是在离线标定阶段,因此,不需要考虑配准消耗,只需要关注精度问题。算法流程如下:

步骤 1 使用传统的 SIFT 算法进行特征提取和

配准,假设提取出的配准点共 t 个,并且这些特征点均保存在集合 P 中, $P = \{ (R_{i1}, S_{k1}), (R_{i2}, S_{k2}), \dots, (R_{it}, S_{kt}) \}$, 令 $i = 1$;

步骤 2 对步骤 1 中提取的配准点再次使用 NCC 算法进行提纯。NCC 计算公式如下:

$$NCC = \frac{\sum_i [I_1(x_i, y_i) - \bar{I}_1] [I_2(x_i, y_i) - \bar{I}_2]}{\sqrt{\sum_i [I_1(x_i, y_i) - \bar{I}_1]^2} \sqrt{\sum_i [I_2(x_i, y_i) - \bar{I}_2]^2}} \quad (4)$$

其中, I_1 和 I_2 分别是两个特征点在各自图像中对应的相关窗口。 $\bar{I}_1$ 和 $\bar{I}_2$ 为各自窗口内像素点的平均灰度值。若计算出 $NCC < T_1$, 则将该匹配点对从集合 P 中去除, T_1 为阈值;

步骤 3 $i = i + 1$, 若 $i \leq t$, 转到步骤 2, 否则转到步骤 4;

步骤 4 算法结束。

通过以上步骤可以对传统的 SIFT 特征匹配集进一步的提纯, 实现图像特征点的更精确匹配。

2.3 单应性矩阵计算

在摄像机数学模型中, 一个三维空间中的点 $P = [X, Y, Z, 1]^T$ 映射到摄像机二维平面上的点 $p = [x, y, 1]^T$ 满足以下关系:

$$p = MP \quad (5)$$

假设 2 个摄像机的相机映射矩阵分别为 M_1 和 M_2 , 通过这 2 个相机拍摄的图像中有相同点 P , 而点 P 在两幅图像中的齐次坐标分别为 p_1 和 p_2 , 则有下列公式成立:

$$p_1 = M_1^{-1} p_1 = M_2^{-1} p_2 \quad (6)$$

$$p_1 = M_1^{-1} M_2^{-1} p_2 = H p_2 \quad (7)$$

这样, 摄像机中相应点的变换就可以用一个 3×3 的单应性矩阵 H 来表示了。2 幅图像之间的单应性矩阵 H 可以表示成如下形式:

$$H = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix} \quad (8)$$

H 为非奇异矩阵, 因此, 对于点 p_1, p_2 用非齐次坐标代替齐次坐标, 设 p_1 和 p_2 的非齐次坐标为 $p_1 = [x_1, y_1, 1]^T$ 和 $p_2 = [x_2, y_2, 1]^T$, 则二维投影变换的关系可表示为:

$$x_1 = \frac{a_{11}x_2 + a_{12}y_2 + a_{13}}{a_{31}x_2 + a_{32}y_2 + 1} \quad (9)$$

$$y_1 = \frac{a_{21}x_2 + a_{22}y_2 + a_{23}}{a_{31}x_2 + a_{32}y_2 + 1} \quad (10)$$

由于一对匹配点只能确定 2 个线性方程, 因此, 理论上需要至少 4 对匹配点才可以求出投影变换矩阵 H 。对于实际工程中有 n 个点 ($n > 4$) 时, 最终可利用最小二乘法计算出投影变换矩阵 M 。

2.4 配准算法结果

在 NCC 的每一步迭代中, 几组匹配点对被随机地选择用来生成一个最佳的变换矩阵, 并且和该矩阵一致的点对被看成是内点。在一定量的迭代后, 拥有最多内点对的变换被看作最后的结果。



图 1 原始图像



图 2 传统 SIFT 算法特征配准结果



图 3 改进的 SIFT 算法提纯后的结果

3 快速拼接

3.1 摘要

在柱面投影模型中, 可以减少因视差引起的误匹配, 降低重合区域内出现错位、断裂等不良情况。将图像投影到投影柱面再进行拼接可以获得很好的拼

接效果。柱面投影公式如下:

$$x' = f \cdot \arctan \frac{x-W/2}{f} + f \cdot \arctan \frac{W}{2f} \quad (11)$$

$$y' = \frac{f(y-H/2)}{\sqrt{(x-W/2)^2 + f^2}} + \frac{H}{2} \quad (12)$$

其中, W 、 H 分别为原图像宽、高, f 为相机焦距。通过上述公式可将原图映射到一张空图像上, 由此确定空像素上对应位置像素点的灰度值, 即可获得柱面投影图像。

3.2 图像融合算法

在柱面投影完成之后, 就是图像融合^[10-14]。图像融合是将图像配准之后的待拼接图像在空间上进行叠加。本文使用加权平均融合算法^[15]对图像进行平滑过渡。假设 I_1 和 I_2 是 2 幅待拼接的图像, 则融合后的图像 I 像素为:

$$I(x, y) = \begin{cases} I_1(x, y), & (x, y) \in I_1 \\ m_1 I_1(x, y) + m_2 I_2(x, y), & (x, y) \in I_1 \cap I_2 \\ I_2(x, y), & (x, y) \in I_2 \end{cases} \quad (13)$$

其中, m_1 、 m_2 表示权重值, 其值一般与重叠区域宽度相关, 且 $m_1 + m_2 = 1$, $0 < m_1 < 1$, $0 < m_2 < 1$ 。权值 m_1 逐渐从 1 变为 0, m_2 逐渐从 0 变为 1, 因此, 可实现重叠区域逐渐过渡并完成两幅图像的拼接。

3.3 拼图阶段整体算法流程

在实际拼图之前必须先进行离线标定获取单应性矩阵, 拼图中使用该单应性矩阵计算并完成拼图流程。但是拍摄的多台相机需要保证位置固定, 且有一定的重合面积, 具体拼接过程如下:

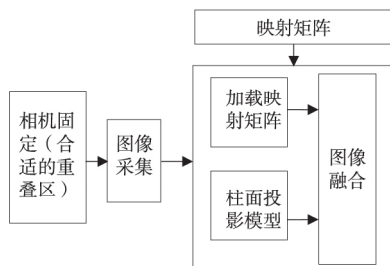


图4 拼接算法流程图

4 实验结果

为验证本文算法的效果, 进行了实验验证, 系统参数为 Win10, 64 位, 软件部分使用 opencv4.0 + vs2017 平台。

实验对比了 SIFT 算法和本文改进的 SIFT 算法图像配准效果以及多种图像融合算法的效果, 并列出了相应的数据表格。

(1) 离线标定阶段:

离线标定阶段主要包括特征提取、特征匹配和单应性矩阵的计算。



图5 待拼接图像



图6 传统 SIFT 算法特征配准结果



图7 改进的 SFIT 算法特征配准结果

表1 二者剔除率剔除率对比

	筛选特征点/对	误匹配/对	剔除率/%
改进前	87	13	14.94
改进后	87	33	37.93

上述是离线标定阶段, 从图中可以看出, 改进后的 SIFT 算法, 剔除了很多无效的匹配点, 具有更好的匹配效果。

(2) 实时拼接阶段:

在离线标定的基础上, 图像拼接的速度取决于图像融合算法。以下是对比加权平均融合、改进渐入渐出算法和多分辨率图像融合算法的效果和速度。



(a) 待拼接左图柱面投影 (b) 待拼接右图柱面投影

图 8 柱面投影



(a) 加权平均图像融合算法结果 (b) 改进渐入渐出图像融合算法结果 (c) 多分辨率图像融合算法结果

图 9 拼接后的图像

表 2 算法效率对比

图像大小	加权平均融合/ms	改进渐入渐出/ms	多分辨率融合/ms
1 080* 1 440	94	151	2 029
540* 720	36	70	1 331

5 结论

在传统全景图像拼接的基础上,提出了先离线标定后实时拼接的方法。由于相机的位置固定,因此,标定过程只需要进行一次,实时图像拼接中图像对齐和图像融合 2 个步骤。不需要再对待拼接图像进行特征提取和配准,可以节省大量的时间。在标定阶段采用 NCC 算法对传统 SIFT 算法特征匹配结果进行二次配准,提高了匹配精度。采用加权平均融合算法可以实现对图像快速拼接,从实验结果可以看出,加权平均融合算法效果好,计算量小,图像融合速度最快,能够满足实时性的要求。

参考文献

- [1] 杨春德,成燕菲. 基于拼接缝自适应消除和全景图矫正的快速图像拼接算法 [J]. 计算机应用,2019,39(10): 3053-3059.
- [2] Lowe DG. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004,60(2): 91-110.
- [3] 罗久飞,邱广,张毅,等. 基于自适应双阈值的 SURF 双目视觉匹配算法研究 [J]. 仪器仪表学报,2020,41(3): 240-247.

- [4] 周志艳,闫梦璐,陈盛德,等. Harris 角点自适应检测的水稻低空遥感图像配准与拼接算法 [J]. 农业工程学报, 2015,31(14): 186-193.
- [5] 刘文亮,王正勇,卿粼波,等. LBP 算法在岩石薄片图像拼接中的应用 [J]. 计算机与数字工程,2016,44(2): 3263-330.
- [6] 赵岩,陈月,王世刚. 结合投影误差矫正的快速 SIFT 图像拼接 [J]. 光学精密工程,2017,25(6): 1645-1651.
- [7] 张翔,王伟,肖迪. 基于改进最佳缝合线的图像拼接方法 [J]. 计算机工程与设计,2018,39(7): 1964-1970.
- [8] 许佳佳. 结合 Harris 与 SIFT 算子的图像快速配准算法 [J]. 中国光学,2015,8(04): 574-581.
- [9] 卢嘉铭,朱哲. 基于 GPU 加速的实时 4K 全景视频拼接 [J]. 计算机科学,2017,44(8): 18-21.
- [10] 冯宇平,戴明,孙立悦,等. 图像自动拼接融合的优化设计 [J]. 光学精密工程,2010,18(02): 470-476.
- [11] 朱炼,孙枫,夏芳莉,等. 图像融合研究综述 [J]. 传感器与我系统,2014,33(2): 14-18.
- [12] 杜往泽,宋执环,闫文博,等. 单摄像头旋转监控下的快速图像拼接 [J]. 中国图像图形学报,2016,21(2): 245-254.
- [13] 杨思燕,贺国旗,刘如意. 基于 SIFT 算法的大场景视频拼接算法及优化 [J]. 计算机科学,2019,46(7): 286-291.
- [14] 李倩,江泽涛. 二值化的 SIFT 特征描述子及图像拼接优化 [J]. 中国图像图形学报,2016,21(12): 1593-1601.
- [15] 雍玖,王阳萍,雷晓妹,等. 基于改进 ORB 特征检测的全景视频拼接 [J]. 计算机应用与软件,2017,34(5): 182-188.
- [16] 徐鑫,孙韶媛,沙钰,等. 一种基于改进 RANSAC 的红外图像拼接方法 [J]. 激光与光电子学进展,2014,51(11): 135-140.
- [17] 宋建辉,闫蓓蕾. 基于 SIFT 的无人机航拍图像快速拼接技术研究 [J]. 计算机应用与软件,2018,35(2): 230-234.
- [18] 王民,郑余杰,权宇寰,等. 基于斜率改正比的遥感图像拼接缝消除算法 [J]. 计算机工程与应用,2015,51(18): 165-168.
- [19] 仵建宁,郭宝龙,冯宗哲. 一种基于兴趣点匹配的图像拼接方法 [J]. 计算机应用 2006,26(3): 610-612.
- [20] 李冰,李范鸣,谢江荣,等. 基于红外面阵扫描图像的配准和融合 [J]. 激光与红外,2017,47(2): 246-251.
- [21] 郑悦,程红,孙文邦. 邻域最短距离法寻找最佳拼接缝 [J]. 中国图像图形学报,2014,19(2): 227-233.
- [22] 黄立勤,陈财淦. 全景图拼接中图像融合算法的研究 [J]. 电子与信息学报,2014,36(6): 1292-1298.
- [23] 罗永涛,王艳,张红民. 结合最佳缝合线和改进渐入渐出的图像拼接算法 [J]. 红外技术 2018,40(4): 382-387.
- [24] 谷雨,周阳,任刚,等. 结合最佳缝合线和多分辨率融合的图像拼接 [J]. 中国图像图形学报,2017,22(6): 842-851.