

图像拼接技术在双摄像机监控中的应用

魏 川¹, 张功国^{1,2}, 吕晓萌²

(1. 重庆信科设计有限公司, 重庆 401121;

2. 重庆邮电大学 重庆 400065)

摘 要:针对单摄像机监控范围的局限性,给出一种双摄像机图像拼接方法。首先检测出视频序列图像的兴趣点,并引入改进的尺度不变特征变换算法构建64维匹配描述子;在匹配过程,采用BBF算法进行匹配,提高搜索效率;匹配完成后,采用随机抽样一致性消除误匹配;最后采用加权平均法进行视频序列图像融合,实现序列图像的无缝拼接。实验结果表明,本文方法有较高的匹配效率和匹配准确率,可以明显扩大视频监控的范围,其在视频监控、信息收集及司法取证等方面有一定的实际应用价值。

关键词:双摄像头;图像拼接;搜索效率;随机抽样一致性

中图分类号:TP391 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2014.04.021

Application of image mosaic in binocular cameras surveillance

WEI Chuan¹, ZHANG Gong-guo^{1,2}, LÜ Xiao-meng²

(1. Chongqing Information Technology Designing Co. Ltd., Chongqing 401121, China;

2. Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Aiming at the disadvantage of the small monitoring range of single camera, an image mosaic method based on binocular cameras is presented. Firstly, the interesting points from video sequence image are detected, and improved scale invariant feature transform algorithm is introduced to build the 64-dimensional descriptors. In the process of matching, BBF algorithm is adopted, which can improve the matching efficiency. After that, the error matching is eliminated by using random sample consensus method. Finally, the image mosaic is done with the weighted average method. The experimental results show that this method has a higher matching efficiency and accuracy, and it can effectively extend monitoring range. It has a practical application value in the fields of video surveillance, information collection and judicial forensics.

Key words: binocular cameras; image mosaic; detecting efficiency; random sample consensus

1 引言

在广场等大场景使用的监控系统中,由于视频质量的高低和视频监控范围大小二者相互制约,为了得到较高的监控视频质量,一般都采用多个监控摄像头同时使用。但是,当需要从监控视频中获取有用的信息,如进行人群密度估计^[1],由于多个监控的分别使用,得到的视频图像无法从整体角度来进行评估,造成信息采集人员的工作

不便。

图像拼接是将有共同部分的待拼接图像通过计算机技术进行处理,拼接成一幅整体图像。

基金项目: 国家科技重大专项资助项目(No. 2009ZX03003-01); 科技型中小企业技术创新基金(No. 11C26215113601)资助。

作者简介: 魏 川(1981-),男,工程师,主要研究方向为通信新技术及应用。E-mail: lvxiaomeng13@126.com

收稿日期: 2013-08-07

1996年, Richard Szeliski 通过获取图像间的几何变换完成图像配准, 进而完成图像拼接的全景图像拼接模型^[2]。2010年, 吴锦杰, 刘肖琳提出了一种基于双目相机的图像拼接方法^[3], 该方法引入自适应的非极大值抑制技术, 检测重合区域的 Harris 角点, 对图像进行快速匹配; 但其在图像匹配阶段的运算时间较长, 实时性不强。

本文提出一种基于双目摄像机的图像拼接方法, 在图像配准阶段借鉴文献^[3]中所采用基于特征点的图像匹配方法的思想, 采用基于特征点的图像匹配方法, 通过对匹配阶段描述子的维数降低来减少运算时间, 采用 BBF 匹配算法来提高匹配效率, RANSAC 消除错误匹配对, 最后用加权平均融合法消除拼接缝隙, 达到平滑拼接缝隙的目的。

2 改进的图像拼接方法

图像拼接的一般流程如图 1 所示。

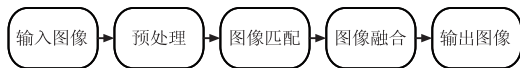


图 1 图像拼接流程

图像预处理就是在进行图像匹配之前, 对图像进行噪声消除、几何校正等工作, 以此来提高拼接的成功率。图像匹配是图像拼接阶段最重要的一个步骤, 匹配结果的好坏直接影响后期拼接的效果。图像融合是图像拼接的最后一步, 图像匹配解决图像相同部分的对齐问题, 图像融合解决图像对齐后的缝隙问题。

本文方法的适宜应用场景大多采用固定摄像头, 本文使用水平固定的两个小型摄像头来进行视频的采集, 采集到的图像受到的几何畸变对后期的匹配和配准影响不大, 因此可以对所获取的图像可以直接进行匹配操作。

SIFT (Scale Invariant Feature Transform) 算法是 David G. Lowe 教授提出的一种图像局部特征描述算子, 它是基于图像特征尺度选择的思想, 在尺度空间寻找极值点, 获取特征点的位置、尺度、方向等信息。

2.1 改进的 SIFT 算法

SIFT 算法主要步骤如下:

(1) 极值点检测

基于特征的图像匹配的关键是稳定的特征点^[4]。为了使最终的描述子具有尺度及缩放不变

性, 要对图像进行多次的高斯模糊和降采样处理, 形成高斯金字塔。1994年, Lindeberg^[5]发现高斯差分函数 (Difference of gaussian) 与尺度归一化的高斯-拉普拉斯函数 $\sigma^2 \nabla^2 G$ 非常相似, 而尺度归一化的高斯-拉普拉斯函数 $\sigma^2 \nabla^2 G$ 的极大值和极小值和特征提取函数, 如 Hessian 矩阵比较, 能够产生最稳定的图像特征。

在生成的 DOG 尺度空间中, 将任意一个像素点与同尺度周围 8 个像素点及上下尺度对应的 18 个点比较其像素值, 如果为极大值或极小值时, 标注为极值点。

(2) 特征点方向确定

为了使生成的 SIFT 特征点具有旋转不变性, 利用特征点邻域像素的梯度方向分布特性为每个特征点分配方向参数, 特征点的梯度的模和方向计算公式为:

$$m(x, y) = [(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2]^{\frac{1}{2}} \quad (1)$$

$$\theta(x, y) = \tan^{-1} \left[\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right] \quad (2)$$

其中, L 为对应特征点所在的尺度; (x, y) 为特征点的位置。

(3) 改进的描述子

以特征点为中心, 计算 16×16 邻域内的像素梯度值的大小和方向, 将特征点邻域划分为 4×4 的种子点, 统计每个种子点 8 个方向的梯度方向直方图。16 个种子点, 每个种子点 8 个方向, 形成 $16 \times 8 = 128$ 维描述子。生成的 SIFT 描述子有较好的旋转不变性, 光照不变性, 当检测到的特征点数量较多的时候, 128 维描述子会产生庞大的计算量, 实时性不强。对此, 本文给出一种改进方法, 降低计算量, 提高实时性。

以检测到的特征点为中心, 构造分别以半径为 2, 4, 6, 8, 10, 12, 14, 16 个像素的 8 个圆, 计算这 8 个半径圆内各像素的梯度信息, 统计出每个圆域的 8 个方向的梯度值。将特征点的邻域高斯加权, 这样可以避免由于图像旋转而造成的特征点发生位移导致生成描述子偏差。

接下来将 2 个像素邻域内的 8 个方向的梯度累加值作为生成描述子的前 8 个向量, 然后对第 2 到

第4像素邻域内的梯度累加值作为描述子的9~16个向量,然后是计算第4到第6像素邻域内的梯度累加值,生成17~24个向量。依次地,形成8组邻域的向量,总共64维描述子。

与改进前相比较,特征点的邻域范围为 16×16 ,但特征点的描述子维数降低到了64维,描述子维数明显降低。维数的降低将会直接导致计算速率的提高,运算时间的减少。

(4) 匹配和误匹配的消除

对描述子的匹配常采用的方法是穷举法,穷举法的优点是不需要进行任何数据预处理,但其搜索效率不高。本文通过BBF查寻机制来提高匹配点的搜索效率。

BBF(Best-Bin-First)查寻机制,是KD-Tree(k dimension tree)的一种扩展^[6]。KD-Tree是一个二叉树,它是用分而治之的思想对数据点在K维空间划分的一种数据结构,在检索的数据集的维数高于20维时,其检索效率明显下降^[7]。BBF是在KD-Tree的基础上加入了查找优先级的概念,它的检索总是从优先级最高(Best Bin)的节点开始,通过设置一个运行超时限定来确保在所有节点检查完毕或超出时间限制后算法能返回当前找到的最好结果,其主要流程如图2所示。

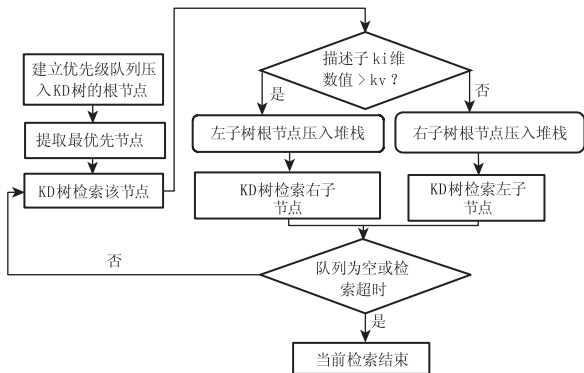


图2 基于BBF的KD-Tree检索流程

采用BBF查寻机制,可以将KD-Tree扩展到高维数据集上,通过建立优先队列,其很好地控制了‘最佳点’查寻的进程,可以在任何时候中断并退出查询进程,并且能得到比较好的结果。

初始匹配对中,存在部分匹配错误,消除匹配错误的方法有最小二乘法和随机抽样一致性^[8]。随机抽样一致性(RANSAC RANdom SAmple Consensus)是一种数学模型的参数迭代算法,它的主要特点是随着迭代次数的增加,正确的几率会逐渐提高。它通过将

数据分为内点和外点,然后判断出内点,取出外点来达到消除错误匹配的目的。在该方法中有3个关键参数可以对其效果产生影响,分别是:估计次数,内外点距离判定阈值和一致性集合的大小阈值。

本文中的随机抽样一致性消除错误匹配分为3步来进行:①随机选取3对匹配对,然后根据待匹配图像的变换模型估计出变换参数。②利用估计的参数剩余的匹配点进行判断,通过内外点的判定阈值区分出内外点,然后重复步骤1。③当内点数目最大时,得出变换模型的最佳估计。

2.2 图像融合

图像融合是将采集到的包含有同一目标的图像数据经过处理,综合到同一幅图像中。由于图像采集视角的不同,拼接的重叠部分会产生模糊和重影等。为了提高拼接效果,就需要对拼接后的图像进行融合处理。

图像融合技术分为像素级、特征级和决策级3个层次,多数应用场合都需要进行像素级融合,原因是像素级融合能尽可能多地保留场景的原始信息^[9]。本文采用加权平均融合法进行像素级图像融合。

3 实验结果与分析

本文使用水平固定的两个小型摄像头来进行视频的采集,水平夹角为 40° ,采集的视频分辨率为 640×480 ,帧率为30 f/s。获取的原始图像帧如图3所示。



图3 待拼接图像

SIFT算法匹配结果如图4,采用本文的改进方法进行匹配操作,结果如图5,图中出现的交叉线表示匹配错误。与改进前相比,使本文改进方法进行匹配,错误的匹配点明显减少,匹配准确率明显提高。



图4 SIFT配准



图5 本文方法配准

由于图像拼接的关键在图像匹配过程,为进一步检验本文改进方法的匹配性能,本文分别将本文方法、改进前 SIFT 算法及其改进—SURF 算法进行仿真对比,采用图像尺寸大小不同的 10 组图像进行匹配,并对其匹配时间进行统计,结果如图 6 所示。

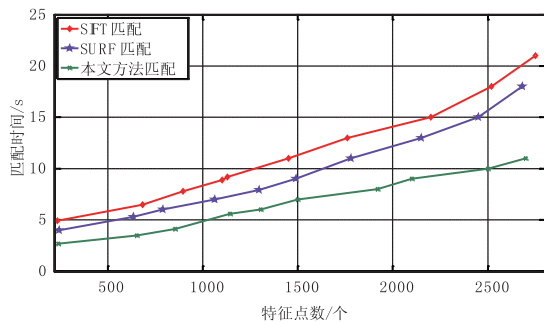


图6 特征点数目与运算时间关系

如图 6 中统计结果,本文方法经过对描述子的降维处理和匹配阶段 BBF 匹配策略的引入,大幅减少了匹配时间;在检测到的特征点数目较少的时候,与 SIFT 算法和 SURF 算法相比,本文方法有较少的匹配时间,并且在一定的特征点数目范围内保持稳定匹配性能;随着检测到的特征点数目的继续增加,SIFT 算法和 SURF 算法的匹配时间大幅增加,而本文方法的匹配时间趋于稳定增加,这是因为 BBF 匹配策略利用优先级队列,针对某查询点,搜索整个 KD-Tree 的根节点和树节点,然后从这些节点中通过超时限制提取出优先级最高的节点,并退出搜索进程。

为验证本文采用加权平均融合法对拼接缝隙的处理性能,本文采用目前普遍使用的直接平均法与本文方法作对比处理,实验结果分别如图 7 和图 8 所示。



图7 直接平均法



图8 本文方法

可以看出,直接平均法对拼接缝隙的处理不够,存在较明显的拼接缝隙,而本文采用的加权平均法,能够根据权值对拼接的重合区域很好的处理。

综合来看,本文方法提高了图像的匹配效率,大幅减少了匹配时间,并且有较高的匹配准确率,实现了图像的无缝拼接,经过实验证明,本文方法是有效的。

4 结束语

针对但摄像头和双摄像头拍摄范围各自的局限性,给出了一种图像拼接方法,通过该方法可以对双摄像头的拍摄视频图像进行快速拼接,有较高的实时性,在相关的监控视频信息收集中有一定的应用价值。本文方法是正常的自然光照条件下进行的,若遇突发情况,如较强的光照影响下(如摄像头受到灯光照射),采集到的视频资料受到影响较大,如何在影响较大的情况下继续进行准确快速拼接有待做进一步的研究。

参考文献:

- [1] GUO Sen, LU Xin. Application of the binocular camera to crowd density estimation [J]. Computer & Digital Engineering, 2012, 40(8): 116-118. (in Chinese)
郭森, 卢鑫. 应用双目摄像机进行人群密度估计 [J]. 计算机与数字工程, 2012, 40(8): 116-118.
- [2] Richard Szeliski. Video mosaics for virtual environments [J]. IEEE Computer Graphics and Applications. 1996. 16(2): 22-30.
- [3] WU Jinjie, LIU Xiaolin. Image mosaicing based on binocular cameras [J]. Computer Engineering, 2010, 36(12): 209-212. (in Chinese)
吴锦杰, 刘肖琳. 基于双目相机的图像拼接 [J]. 计算机工程, 2010, 36(12): 209-212.
- [4] ZHANG Jiao, LI Junshan, ZHANG Shaohua. Algorithm for IR/Visual images matching based on line features and SIFT points [J]. 2013, 43(5): 554-558. (in Chinese)

- 张姣,李俊山,张少华. 线特征和 SIFT 点特征结合的异源图像匹配算法[J]. 激光与红外, 2013, 43(5): 554 - 558.
- [5] Lindeberg T. Scalespace theory: A basic tool for analyzing structures at different scales[J]. Journal of Applied Statistics, 1994, 21: 224 - 270.
- [6] ZHAO Qinghua, CHEN Luo, JING Ning. Distributed spatial join query based on kd-tree recursive partitioning[J]. Computer Engineering & Science, 2011, 33(8): 167 - 173. (in Chinese)
- 赵清华, 陈萃, 景宁. 基于 Kd 树递归区域划分的分布式空间连接查询[J]. 计算机工程与科学, 2011, 33(8): 167 - 173.
- [7] WANG Peizhen, CHEN Ping, ZHOU Fang. Method of feature extraction and matching for complex image[J]. J. of Anhui University of Technology, 2012, 29(1): 73 - 77. (in China)
- 王培珍, 陈平, 周芳. 复杂图像特征点提取与匹配方法[J]. 安徽工业大学学报, 2012, 29(1): 73 - 77.
- [8] CAO Hongxing, LIU Jiahang, RUAN Ping. Stitching image based on SIFT[J]. Video Engineering, 2008, 32(S1): 146 - 148. (in Chinese)
- 曹红杏, 柳稼航, 阮萍. 基于 SIFT 特征的图像自动拼接[J]. 电视技术, 2008, 32(S1): 146 - 148.
- [9] SONG Baosen, FU Yongqing, SONG Hailiang. New efficient image fusion algorithm for image mosaic[J]. Computer Science, 2011, 38(2): 260 - 264. (in Chinese)
- 宋宝森, 付永庆, 宋海亮. 一种消除图像拼接痕迹的新方法[J]. 计算机科学, 2011, 38(2): 260 - 264.