

# 基于特征点的图像拼接方法的研究与应用

汪前进,朱 斌,李存华

(淮海工学院计算机工程学院,江苏 连云港 222005)

**[摘要]** 图像拼接技术是图像处理的热点领域,图像配准和图像融合是图像拼接的两个关键技术.本文研究了基于特征点的配准方法,使用 SURF 算法完成特征点的检测和描述.在此基础上,调用 OpenCV 的 API,实现了图像拼接,完成了柱面全景图的拼接和分析.实验证明,该方法在自动化程度、运算速度和鲁棒性方面效果理想.

**[关键词]** OpenCV, 图像拼接, 图像分析, SURF 算法

**[中图分类号]** TP391 **[文献标志码]** A **[文章编号]** 1672-1292(2016)03-0048-06

## Research and Application of Image Stitching Method Based on Feature Points

Wang Qianjin, Zhu Bin, Li Cunhua

(School of Computer Engineering, Huaihai Institute of Technology, Lianyungang 222005, China)

**Abstract:** Image mosaic technology is a hotspot in the field of image processing, image registration and fusion are two key technologies of image mosaic. This paper proposes a method based on feature points, it detects and describes feature points using SURF. It is realized by API on OpenCV and completes the cylindrical panoramic mosaic and achieves an ideal effect in automation, computing speed and the robustness by experiment.

**Key words:** OpenCV, image stitching, image analysis, SURF algorithm

图像拼接技术是指将多张具有重合或相似部分的图像拼接成一幅分辨率不变且没有拼接缝的图像的技术.其目的是将一系列真实世界的图像拼接成一幅更宽视野的大型场景图像.通过图像拼接,可以剔除图像中冗余的信息,减少图像的存储量.图像拼接技术被广泛应用于虚拟现实技术、数字视频压缩、运动分析、医学图像分析、遥感图像处理等领域,具有很好的应用前景.图像拼接技术的关键是图像配准和图像融合.目前,图像拼接技术主要有基于相位、基于几何区域和基于特征点3类.1999年,剑桥大学的 Brown 概述了图像配准技术的基本理论以及主要方法<sup>[1]</sup>, Zitová 等人综述了随后十年的图像配准领域的相关方法<sup>[2]</sup>, Love 提出的传统 Sift 算法<sup>[3]</sup>在拼接效果上取得成功,但在特征点的描述、匹配上却耗时过长,效率不高.

本文研究了基于特征点的图像拼接技术,提出并实现了多幅图像的拼接方法,其主要步骤为:(1)特征点检测;(2)特征点描述;(3)特征点匹配;(4)变换矩阵估计;(5)图像映射和插值;(6)图像融合.在实现过程中,使用近似 Hessian 矩阵检测特征点;使用 SURF 描述子描述特征点;对于特征点匹配结果,通过最近距离比次近距离和 RANSAC 算法剔除误匹配点;使用单应性变换矩阵估计图像之间的几何变换模型;使用多频段融合算法完成图像融合.在拼接过程中,对于没有对应映射值的点,采用后向插值的方法进行图像插值.

## 1 特征点检测与描述

### 1.1 特征点检测

特征点检测是指在具有独特性的位置检测出特征点,具有独特性的位置包括角点(corners)、斑状结构

收稿日期:2016-07-20.

基金项目:江苏省高等教育教改研究立项课题(2015JSJG126)、连云港市应用研究课题(LSYZ16409).

通讯联系人:汪前进,副教授,研究方向:智能信息处理、智慧教育. E-mail:747359918@qq.com

(blob)、T 型结合处(T-junctions)等. 重复度是特征点检测算法的重要指标. 所谓重复度就是检测算法在不同视角拍摄的图像中找到相同特征点的能力.

本方法使用近似 Hessian 矩阵检测特征点. 图像中,某点  $P(x,y)$  的像素可以用函数值  $F(x,y)$  表示,使用标准高斯函数二阶导数作为滤波器,用特定的卷积模板计算二阶偏导数,通过计算出 3 个矩阵元素  $L_{xx}, L_{xy}, L_{yy}$ , 从而计算出 Hessian 矩阵:

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

由于所使用的特征点需要具备尺度无关性,所以在进行 Hessian 矩阵构造前,需要对其进行高斯滤波,其公式如下:

$$L(x, t) = G(t)I(x, t).$$

$L(x, t)$  是一幅图像在不同尺度下的表现形式,通常可以使用高斯核  $G(t)$  与图像函数  $I(x)$  在点  $x$  的卷积来表示,高斯核  $G(t)$  可表示为:

$$G(t) = \frac{\partial^2 g(t)}{\partial x^2},$$

其中,  $t$  为高斯方差,  $g(t)$  为高斯函数. 图像中每个像素点的  $H$  行列式决定值都能通过该方法求出,然后,使用该值来判断是否为特征点. 为方便使用, Herbert Bay 提出了使用近似值来代替  $L(x, t)$ . 但使用近似值会产生误差,为了平衡这种误差,引入了权值. 权值是随着尺度变化而变化的,则  $\mathbf{H}$  矩阵判别式可表示为:

$$\det(\mathbf{H}_{\text{approx}}) = D_{xx}D_{yy} - (0.9D_{xy})^2,$$

其中, 0.9 是本文所给出的经验值.

图 1 所示为检测点检测效果图,图中圆心是特征点的位置,圆的半径代表该特征点的尺度,圆中的横线代表方向.

## 1.2 特征点描述

特征点描述是指用一个特征向量描述一个兴趣点邻域的信息,使该特征点能够被唯一识别. 此外,特征点描述算子还应具有鲁棒性,不受噪声、光照变化及拍摄视角变化等影响.

SURF (Speeded Up Robust Feature) 是一种高鲁棒性的局部特征点检测器. SURF 算法使用一阶的 Haar 小波构建特征点的特征向量,通过其在  $X$  和  $Y$  两个方向的响应表示特征点分布信息. SURF 算法统计特征点邻域内的 Harr 小波特征,即在特征点的邻域(例如,半径为  $6s$  的圆内,  $s$  为该点所在的尺度)内,统计  $60^\circ$  扇形内所有点的水平 Haar 小波特征和垂直 Haar 小波特征总和, Haar 小波的尺寸为边长  $4s$ , 这样可得到一个值. 然后将  $60^\circ$  扇形以一定间隔进行旋转,最后将最大值那个扇形的方向作为该特征点的主方向<sup>[4-6]</sup>. 特征点主方向的确定如图 2 所示.



图 1 特征点检测效果图

Fig.1 Effect map of Feature point detection

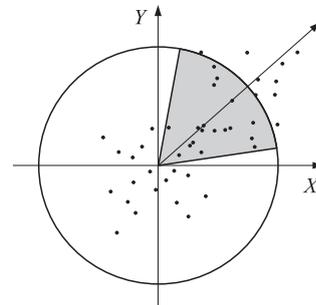


图 2 特征点主方向确定图

Fig.2 The map of feature points in the main direction

OpenCV 的 API 封装了 SURF 的实现过程. 在使用 SURF 算法时,先创建一个 SurfFeatureDetector 类的对象. 在初始化时,需要传入一个参数作为阈值,这个阈值可以控制检测到的特征点的数量,阈值越大,检测到的特征点越多. 用该对象调用 detect 方法,用以检测特征点. detect 方法需要两个参数,一是要检测的图像,另一个是存放检测到的特征点的变量. 最后调用 compute 方法以完成检测点描述,此

处有3个参数,一是图像本身,二是上一步检测到的特征点,其三为存放特征向量的变量.其代码如下:

```
SurfFeatureDetector detector(300);
detector.detect(img[i],feature[i].keypoints);
detector.compute(img[i],feature[i].keypoints,feature[i].descriptors).
```

## 2 特征点匹配

特征点匹配的目标是找出两幅图像中正确匹配的特征点.当检测出图像的特征点之后,需要完成两幅图像之间特征点的匹配,也即将反映相同物理位置的特征点建立起对应关系,计算出两幅图像间的几何变换模型.

特征点的定位误差一般比较小,但是没有误匹配点是求解高精度的变换参数的前提.因此程序必须先剔除误匹配点.剔除误匹配点的方法主要有:(1)最近距离比次近距离.假设ND表示最近距离(Nearest Distance),NND表示次近距离(Next Nearest Distance),则距离之比(Ratio of distance) $Rod=ND/NND$ .可以给定一个阈值Threshold,该阈值是一个小于1的正数,当 $Rod<Threshold$ 时成功,否则即为失败.根据实验结果,Threshold在区间[0.5,0.7]内较为比较理想.(2)RANSAC(Random Sample Consensus)算法,又称随机抽样一致性算法,在OpenCV中能很好地剔除错误的匹配点对<sup>[7]</sup>.

### 2.1 基础矩阵

在使用RANSAC算法时,需选取一个几何约束模型,基础矩阵可以充当这种模型.基础矩阵是一个自由度为7的 $3\times 3$ 矩阵,行列式为0,秩为2.基础矩阵不可逆.

### 2.2 特征点匹配与分析

本文基于VS2010开发平台搭建实验环境,将OpenCV和Qt配置在一起,对两幅同一建筑物不同角度的图像进行实验.图3所示为经过一系列误匹配点剔除算法处理之后产生的图像.初始,两幅图检测到的特征点数分别为274和438,使用最近距离比次近距离算法剔除误匹配点之后,两幅图像特征点数分别为65和60,分别剔除了209和378个错误特征点,然后将左侧65个特征点和右侧的60个特征点逐一比较,确定是否存在指向同一物理位置的点对,处理之后,初步确定了两幅图像之间存在19对相互匹配的特征点对.可以看出,只是根据简单的最近距离比次近距离算法就能剔除大部分误匹配点,并初步确定匹配点对.



图3 剔除误匹配点后的匹配对

Fig.3 Matched pairs of eliminating the mismatch points

为使匹配点对更精确,使用RANSAC算法剔除一些误匹配点对,调用OpenCV提供API,将上面初步确定的匹配点对传入API函数.函数内部根据这些点对产生一个供该算法使用的变换矩阵,然后将右侧图上的19个匹配点根据该矩阵映射到左侧图上,判断是否和左侧的19个匹配点吻合,若不吻合则剔除这一对匹配点.经过该步处理之后,匹配点对剩下16对,剔除了3对.

为使匹配点对更精确,使用RANSAC算法剔除一些误匹配点对,调用OpenCV提供API,将上面初步确定的匹配点对传入API函数.函数内部根据这些点对产生一个供该算法使用的变换矩阵,然后将右侧图上的19个匹配点根据该矩阵映射到左侧图上,判断是否和左侧的19个匹配点吻合,若不吻合则剔除这一对匹配点.经过该步处理之后,匹配点对剩下16对,剔除了3对.

为确保特征点匹配万无一失,本方法还采用根据尺度空间的平均误差来剔除误匹配对.该方法的原理是,每个特征点都有一个尺度(就是特征点圆的半径),每个匹配点对的尺度相减就是一个尺度差,然后求出上面16对匹配点对的平均尺度差,若某对匹配点对的尺度差比这个平均尺度差大很多,则剔除该匹配点对,经过这一步处理又剔除了1对误匹配点,最后剩下15对匹配点对.

## 3 变换矩阵估计

### 3.1 单应性变换矩阵

当两幅图像相互对应的特征点之间完成匹配之后,可通过单应性变换矩阵这种对应关系估计其间的几何变换模型.在单应性变换矩阵的估计过程中,使用直接线性变换算法.

### 3.2 图像映射

完成变换矩阵的估计之后,利用得到的变换矩阵  $H$  把源图像中的每一个像素点映射到参考图像的坐标系中.待转换图像中的某点  $m(x_1, y_1)$  与矩阵相乘即可将  $m$  点映射到参考图像的坐标系中,得到新的坐标点  $m'(x_2, y_2)$ :

$$\begin{aligned} x_2 &= \frac{H_{11}x_1 + H_{12}y_1 + H_{13}z_1}{H_{31}x_1 + H_{32}y_1 + H_{33}z_1}, \\ y_2 &= \frac{H_{21}x_1 + H_{22}y_1 + H_{23}z_1}{H_{31}x_1 + H_{32}y_1 + H_{33}z_1}. \end{aligned}$$

图像之间只存在平移、旋转、缩放等变换时,  $H_{31}$  和  $H_{32}$  的理论值均等于 0,  $H_{33}$  理论值为 1. 此时,该公式可表示为:

$$\begin{aligned} x_2 &= H_{11}x_1 + H_{12}y_1 + H_{13}, \\ y_2 &= H_{21}x_1 + H_{22}y_1 + H_{23}. \end{aligned}$$

### 3.3 图像插值

在图像映射的过程中,会出现有些点没有对应的映射值的情况,以致出现亮点或暗点,这时需要进行图像插值.从映射方向上来说,图像插值可以分为两类:后向插值和前向插值.

前向插值是将源图像中的某个像素点映射到目标图像中的某 4 个像素点之间,再按照某种方式把源图像的灰度值分配到目标图像相邻的 4 个像素点上.后向插值是将目标图像中的某个像素点映射到源图像的某 4 个像素点之间,然后根据源图像上的 4 个像素点的值来计算目标图像的像素点值.

本文采用后向插值法,这是因为在前向插值方法中每个目标图像的像素点要经过多次运算,而后向插值每个目标图像的像素点只需一次运算.

图 4 是根据变换矩阵变换后所得的图像.实验程序每次处理 2 张图像,假设有 3 幅图像  $A, B, C$ ,先处理  $A$  和  $B$ ,找出  $B$  到  $A$  的变换矩阵  $H_1$ ,然后处理  $B$  和  $C$ ,找出  $C$  到  $B$  的变换矩阵  $H_2$ ,由于  $B$  和  $C$  都需要变换到  $A$  的坐标系,拼接成一张完整的图像,所以,要求一个矩阵  $H_3$ ,使图像  $C$  可以直接变换到图像  $A$ ,矩阵  $H_3$  可根据矩阵  $H_1$  和  $H_2$  求出.



图 4 变换后的图像

Fig.4 The image after transform

## 4 图像融合

图像融合方法有直接覆盖、加权平均融合和多频段融合等.图像直接覆盖会产生拼接缝.加权平均融合方法也存在两个缺点:第一,因为加权平均融合方法假设的前提是重合区域内进行加权叠加的像素点必须一一正确对应,若配准的精度不高,存在几何错位时,必然会使图像模糊,文献[8]中就介绍了这种情况;第二,若配准精度过高,但图像的重叠区域宽度很大时,图像的高频细节会损失.多频段融合方法(Multi-band blending)是由 Burt 和 Adelson 提出的,该方法的基本思想为:先把待拼接的两幅图像依据其重叠区域分别分解成不同频段的两组图像,构造出两个 Laplacian 金字塔,然后在金字塔的各层进行融合,融合后重构出图像,最后得到完整的拼接图像.在本程序中,使用多频段融合算法来完成图像的拼接.

## 5 应用实例

例 1 图 5 中,校园全景图(d)是由 3 张校园场景图(a)、(b)和(c)拼接而成的,可以看出,3 张校园场景图远近不同,拍摄角度有较大差异,本方法能够有效进行拼接,效果较好,显示了该方法的鲁棒性.程序能够自动完成多幅图片拼接任务,在一台 3G 主频、2G 内存的 PC 机上实验,拼接两幅 1 024×768 的图像,

需要时间不超过 1 s,效率能够较好满足地实际应用的需求.但对于多幅图像,由于拼接后图像体积较大,速度下降明显.

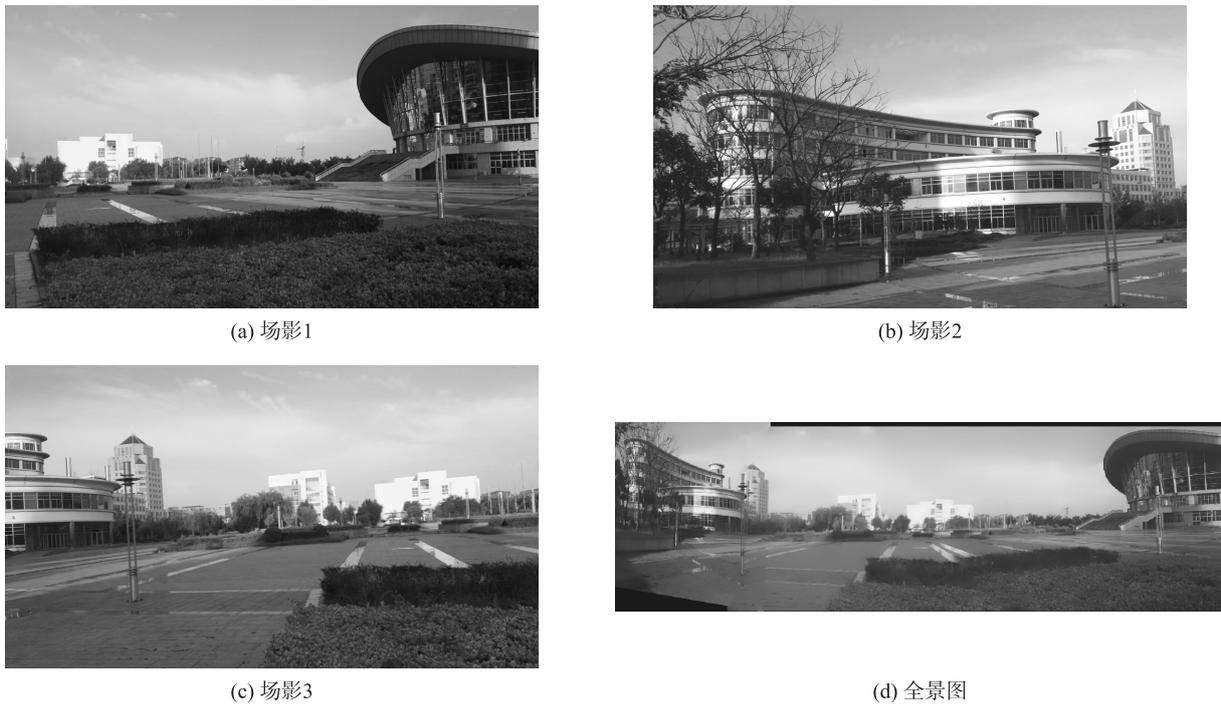


图5 由3幅场景图像拼接的全景图

Fig.5 The panoramic image according to three images

例2 图6所示为连云港周边地区卫星图片的拼接.图6中,连云港全景图(d)是由3张周边地区卫星图(a)、(b)和(c)拼接而成的.可以看出,本方法很好地实现了全景图像的拼接和融合.

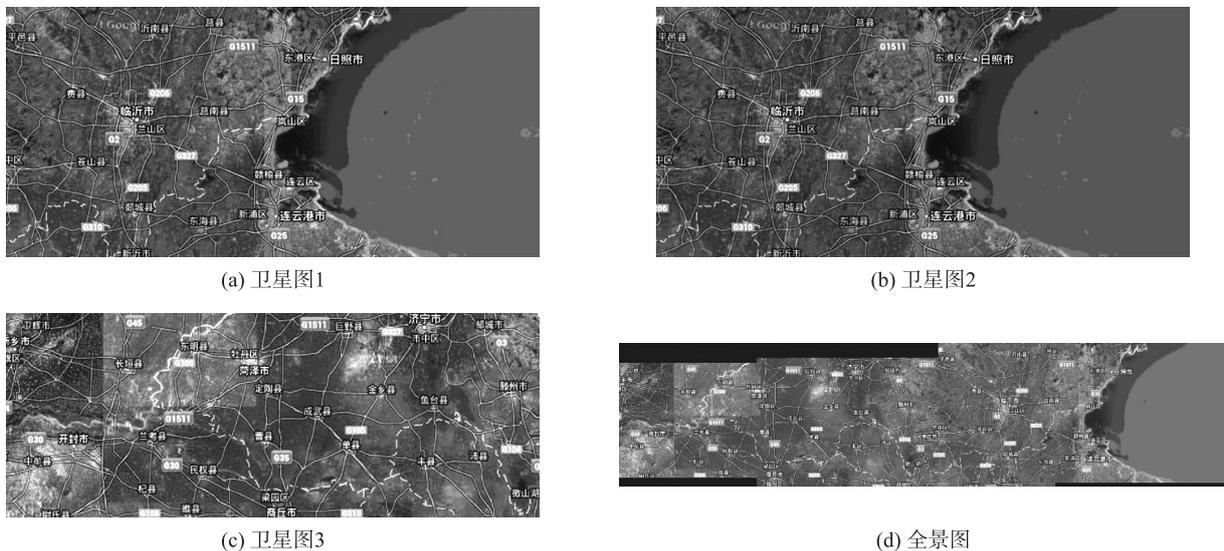


图6 依据3幅卫星图产生的全景图

Fig.6 The panoramic image according to three satellite images fuse

## 6 结语

本文研究了基于特征点的配准方法,使用 SURF 算法完成特征点的检测和描述,并进一步研究了图像拼接技术和图像融合技术,完成了柱面全景图的拼接,并结合连云港及周边地区卫星地图进行了实际应用.该方法在自动化程度、运算速度和鲁棒性方面取得了理想的效果.但图像体积较大时,速度下降明显,效率有待进一步提高.在云计算、大数据环境下,可以通过并行计算,进一步提高算法效率.展望未来,在

虚拟现实、计算机视觉等诸多领域,基于特征点的图像拼接技术具有广泛的应用前景。

### [参考文献](References)

- [1] BROWN L G. A survey of image registration techniques[J]. ACM computing surveys, 1999, 24(4): 325-376.
- [2] ZITOVA B, FLUSSER J. Image registration methods: a survey[J]. Image and vision computing, 2003, 21(11): 977-1 000.
- [3] LOWE D G. Distinctive image features from scale in variant key points[J]. International journal of computer vision, 2004, 60(2): 91-110.
- [4] 阮芹, 彭刚, 李瑞. 基于特征点的图像配准与拼接技术研究[J]. 计算机与数字工程, 2011, 39(2): 141-144.  
RUAN Q, PENG G, LI R. Study on image registration and mosaic technology based on surf feature[J]. Computer & digital engineering, 2011, 39(2): 141-144. (in Chinese)
- [5] 王君本, 卢选民, 贺兆. 一种基于快速鲁棒特征的图像匹配算法[J]. 计算机工程与科学, 2011, 33(2): 112-115.  
WANG J B, LU X M, HE Z. An Improved algorithm of image registration ration based on fast robust feature[J]. Computer engineering & science, 2011, 33(2): 112-115. (in Chinese)
- [6] HERBERT B, ANDREAS E, TINNE T, et al. Speeded-up robust features (SURF) [J]. Computer vision and image understanding, 2008, 110(3): 346-359.
- [7] MARTIN A F, ROBERT C B. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Communications of the ACM, 1981, 24(6): 381-395.
- [8] BROWN M, LOWE D G. Automatic panoramic image stitching using invariant features[J]. International journal of computer vision, 2007, 74(1): 59-73.

[责任编辑: 严海琳]

---

(上接第 47 页)

- [18] OWOPUTI O, O'CONNOR B, DYER C, et al. Improved part-of-speech tagging for online conversational text with word clusters[C]//Proceedings of NAACL, USA, 2013: 380-390.
- [19] BROWN P, DESOUZA P, MERCER R, et al. Classbased  $n$ -gram models of natural language[J]. Computational linguistics, 1997, 18(4): 467-479.
- [20] NAKOV P, KOZAREVA Z, RITTER A, et al. SemEval-2013 task 2: sentiment analysis in Twitter [C]//Proceedings of SemEval2013, USA, 2013: 312-320.
- [21] POURSEPANJ H, WEISSBOCK J, INKPEN D. uOttawa: System description for SemEval 2013 task 2 sentiment analysis in Twitter [C]//Proceedings of SemEval2013, USA, 2013: 380-383.

[责任编辑: 严海琳]