

基于特征点的图像拼接方法^①

张 东, 余朝刚

(上海工程技术大学 城市轨道交通学院, 上海 201620)

摘 要: 提出了一种基于特征点匹配的全景图像拼接方法. 该方法首先利用 sift 算法提取各图像中的特征点并利用 Harris 算法对图像特征点提取进行了优化, 然后采用基于 K-d 树的 BBF 算法查找和确定初始匹配点对, 完成特征点的粗匹配, 再根据图像配准结果使用稳健的 RANSAC 算法对粗匹配的特征点进行筛选, 计算出图像间变换矩阵 H, 最后采用渐入渐出的加权平均的融合算法对两幅图像进行无缝拼接, 形成一幅完整的全景画面. 实验结果验证了该方法的有效性, 拼接效果较好.

关键词: 图像拼接; sift 特征点; 图像配准; RANSAC; 变换矩阵

Image Mosaic Technology Based on Feature Point Match

ZHANG Dong, YU Chao-Gang

(College of Urban Railway Transportation, Shanghai University of Engineering Science, Shanghai 201620, China)

Abstract: This paper proposes a panoramic image mosaic method based on feature points matching. This method firstly uses the sift algorithm to extract the image feature points and uses Harris algorithm to optimize the image feature extraction. Then the BBF algorithm based on K-d tree is used to find and determine the initial matching points and complete the coarse matching of the feature points. Then according to the result of image registration, robust RANSAC algorithm is utilized to filter coarse matching feature points and calculate the transformation matrix H. Finally, the gradually fading out fusion algorithm of the weighted average is used in the seamless Mosaic of two images, form a complete panoramic view picture. Experimental results verify the effectiveness of the proposed method and the splicing effect is better.

Key words: image matching; sift feature points; image registration; RANSAC; transformation matrix

随着计算机图像处理技术的不断发展, 全景图在生活中有着广泛的应用, 通常由于各种条件的制约, 人们往往很难拍出宽视角、高分辨率的全景图像, 因此, 可实现多幅图像无缝拼接的图像拼接技术应运而生.

图像拼接就是将同一幅场景中有相互重叠的多幅图像拼接成大幅面、宽视角、失真小且没有明显缝合线的高分辨率图像^[1]. 目前, 图像拼接算法主要可以分为两类: 一类是基于区域相关的图像拼接算法; 另一类是基于特征相关的图像拼接算法. 第一类是从图像的灰度值出发, 通过计算两幅图像的灰度相关性来确定相似程度, 得到拼接图像重叠区域的位置和范围,

实现图像拼接; 而基于特征的拼接方法是通过提取图像特征点来对图像重叠部分的对应特征区域进行搜索匹配, 该类算法有比较高的稳定性, 应用也比较广泛^[2]. 近年来, 国内外研究人员提出了多种图像拼接算法, Lowe 等人在 2003 年提出了完整的 sift 算法^[3], 该算法将特征点检测、描述和匹配整合为一个统一的过程; Yanke 等^[4]针对 sift 算法计算量大、耗时长这一缺点, 提出了 PCA-SIFT 改进算法, 但计算量并未减少, 且对原有的特征点提取部分并没有改进; 赵向阳、杜立民^[5]提出了一种基于特征点匹配的拼接算法, 其中使用 Harris 算法^[6]提取角点并进行匹配, 该算法使用鲁棒变换估计技术, 一定程度提高匹配算法的稳健性, 但计

① 基金项目: 上海市研究生教育创新计划学位点引导布局与建设培育项目(13sc002)

收稿时间: 2015-06-11; 收到修改稿时间: 2015-07-24

算速度较慢。

本文结合 sift 算法和 harris 算法各自的特点, 首先提出了改进的 sift 算法来提取待拼接图像的特征点; 然后通过 BBF 算法搜索查找匹配点对, 进行粗匹配, 并结合 RANSAC 算法实现图像特征点的精确匹配, 同时估计出图像间的变换矩阵; 最后根据变换矩阵采用加权平均的融合方法进行图像的无缝拼接得到全景图像。

1 特征点的提取

1.1 sift 特征提取算法

尺度不变特征(Scale Invariant Feature Transform, SIFT)是 1999 年由 Lowe 提出的一种提取局部特征的算法, 它在尺度空间中寻找极值点, 并提取出其位置、尺度、旋转不变量。其实现主要包括 4 个步骤: 尺度空间的构建; 空间极值点检测; 确定特征点方向; 生成特征描述符。下面将详细介绍 sift 特征提取的流程。

① 尺度空间的构建

在不同的尺度下观察一幅图像, 成像是不同的, 我们寻找的局部点就是那些图像尺度发生变化, 但其自身的相对位置仍保持不变的特征点。而高斯核被证明是唯一的线型卷积核, 因此采用高斯卷积核对初始图像进行一系列的尺度变换。

首先我们使用高斯函数与图像进行卷积运算对图像进行若干次的连续滤波处理建立了第一个尺度组。然后对图像进行降采样运算将图像尺寸减小到原来的一半, 进行同样的高斯滤波形成了第二个尺度组, 之后重复上述操作直到图像小于某一给定的阈值为止。最后对每个尺度组的高斯图像进行差分运算得到高斯差分图像, 即 DOG 图像^[7]。

DOG 可由下式求得:

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

其中, $G(x, y, \sigma)$ 是尺度可变的高斯卷积核,

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad I(x, y) \text{ 为二维图像,}$$

* 为卷积运算, (x, y) 是空间坐标, σ 是尺度空间因子。

② 空间极值点检测

得到了高斯差分图像以后, 接下来就要对特征点进行检测。本文选取这些高斯差分图像中的局部极值

点作为图像的特征点, 在尺度空间极值点的检测过程中, 每一个采样点要和它所有的相邻点进行比较, 为了确保检测出来的特征点不仅是二维空间上的局部极值而且还是尺度空间中的极值, 除底层和顶层外中间层次的每个检测点都要和它同尺度的 8 个点和上下相邻尺度的 9×2 个点共 26 个相邻像素点进行比较, 如果该点为极值点, 就认为其是图像在该尺度下的一个特征点。

此时得到的极值点中有不稳定的点, 这些点对噪声敏感, 这将影响后续匹配的稳定性, 因此要对极值点进行精确定位, 去除那些对比度低的特征点和不稳定的边缘响应点。根据 Lowe 的研究, 通过三维二次方函数可以确定采样点的位置。尺度函数的二阶泰勒展开式为:

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (2)$$

式中 $X = [x, y, \delta]^T$ 表示特征点与采样点之间的尺度和位置偏移, 对上式求导并另其等于 0, 得到了极值点位置 $\hat{X} = -\frac{\partial^2 D^{-1} \partial D}{\partial X^2 \partial X}$, 将 \hat{X} 的值带入式(2)运算求得极

值的大小为 $D(\hat{X}) = D + \frac{1}{2} \frac{\partial D}{\partial X} \hat{X}$, 判断其绝对值大小,

如果小于某阈值, 判定该点不稳定将其舍弃, 保留下来的极值点即作为特征点。

③ 确定特征点方向

利用特征点邻域像素的梯度方向分布特征为每一个特征点指定方向参数, (x, y) 处的梯度值和方向分别为:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3)$$

$$\theta(x, y) = \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (4)$$

以特征点为中心, 对它的邻域像素采样, 并用梯度方向直方图进行统计, 直方图每 10° 为一个柱将 0° 到 360° 分为 36 个柱, 直方图的峰值则代表了该像素特征点处邻域梯度的主方向, 即为该特征点的主方向。当存在另一个与主峰值能量相当的峰值时, 将这个方向作为该特征点的辅方向。

④ 生成特征描述符

得到特征点主方向以后, 旋转特征点主方向, 使之与坐标轴方向重合, 然后在旋转后区域内取以特征点为中心的 16×16 像素大小的邻域, 并将它分成 16

个 4×4 的子窗口区域, 每个子窗口区域形成一个8维向量种子点, 这样就构成了一个 $16 \times 8 = 128$ 维的特征向量, 即为sift特征向量描述符^[7].

1.2 改进sift特征提取算法

根据Lowe的研究, 使用sift算法对物体进行识别时, 如果能够匹配的关键点达到3个以上, 则可以确定该图像中存在目标物体, 因此在图像拼接过程中只需要少量的特征点就可以完成对重叠部分的匹配, 然而, 对于一幅 512×512 的图像, 一般可以检测出1000个左右的特征点, 在增强鲁棒性处理之后, 仍可以达到200到500个特征点, 在拼接时相当大部分的时间都用在了检测冗余的特征点上.

本文利用Harris算法对sift算法进行了改进, 上文中假设采用sift算法共获取了 N 个特征点, 每个特征点包含3个信息: 位置、尺度和方向, 可以表示为 $P(x_n, y_n, \sigma_n, \theta)$, $n=1, 2, 3, \dots, N$, 其中, x_n 、 y_n 表示特征点的位置信息, 表示特征点的横坐标和纵坐标, σ_n 和 θ 分别表示特征点的尺度和方向信息. 利用这些信息可以计算Harris角响应值 $R_{\sigma_n}(x_n, y_n)$, σ_n 表示计算自相关矩阵 M 时的高斯加权函数的标准差. M 矩阵表示为:

$$M = \sum_{(x,y) \in \mathcal{N}(x_n, y_n)} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (5)$$

式中 $w(x,y)$ 为高斯加权函数, I_x 为 x 方向的梯度, I_y 为 y 方向的梯度, 自相关矩阵 M 具有两个特征值 λ_1 和 λ_2 , 它同矩阵 M 的曲率成正比, 因此可以得到Harris角点的响应函数为:

$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2) = \det(M) - k(\text{trace}M)^2 \quad (6)$$

k 通常取值0.04到0.06, 计算所有特征点响应值并计算其绝对值的累加, 然后求其均值的一半作为阈值 H , 表达式如下所示:

$$H = \frac{1}{2N} \sum_{n=1}^N |R_{\sigma_n}(x_n, y_n)| \quad (7)$$

若 $R_{\sigma_n}(x_n, y_n) > H$ 则剔除该特征点, 保留下来的点则作为改进sift算法提取出的特征点

2 图像特征匹配

2.1 基于K-d树的匹配算法

特征点的匹配可以采用穷举法来完成, 但是这样的话会消耗大量的时间, 因此, 一般采用基于K-d树

的数据结构来完成匹配搜索. 搜索的内容是与匹配目标图像特征点最邻近和次邻近的原始图像特征点. K-d树是分割 k 维数据空间的一种数据结构, 每个数据结点表示 k 维空间的一个点. 每一层都是根据该层的判决器来对相应对象做出分枝决策, 顶层结点按由判决器决定的一个维度来进行划分, 下一层则按照该层的判决器决定的一个维进行划分, 同理在余下各维之间不断地划分. 直至一个结点中的点数少于给定的最大点数时, 结束划分^[10].

当用到高维数据搜索问题时, 这种方法的效率明显下降, 针对这一问题, 提出了基于K-d树的BBF算法^[11], 这种算法的实现依靠一个优先级队列, 队列成员是按照搜索节点与目标节点之间距离的升序排列的. 在搜索中, 当依靠K-d树判决器值的决策沿某个分支方向搜索某个节点时, 一个元素会被加入到优先级队列, 此元素记录了该节点未被搜索分支的信息, 主要包括当前节点与目标节点之间的距离和当前节点在K-d树中的位置. 当一个叶节点搜索到后, 由优先级队列中队首元素的信息按相同方法搜索包含下一近邻节点的其他分支, 并删除队首元素. 在指定数量的节点搜索完后, 结束搜索, 将按距离升序排列的指定数量的近似近邻结果返回.

利用基于K-d树的BBF算法进行匹配: 首先, 采用BBF算法找出目标点与待匹配点欧式距离最近的 K 个点, 因sift特征向量的维数很高, 使很多非匹配点与待匹配点之间的距离集中在一起. 因此使用最近邻中的最小距离点和次小距离点比较, 并设置一个阈值 T , 如果两者的比值大于阈值 T , 则说明待匹配特征点没有匹配点, 否则, 则匹配成功. 具体做法如下:

在BBF搜索中, 假设找到了 K 个待匹配特征点 p 的近似最邻近特征点, 则 K 个特征向量表示为: $v_i \in V, i=0, 1, 2, \dots, k$, 且 $D(v_j, p) \leq D(v_{j+1}, p)$, $j=0, 1, 2, \dots, k-2$ (距离按升序排列), 对于匹配点的判断可用下式进行:

$$\frac{D(v_i, p)}{D(v_{i+1}, p)} > T_r, i=1, 2, \dots, k-2 \text{ 且 } \frac{D(v_0, p)}{D(v_1, p)} \leq T_r.$$

其中, $D(v_i, p)$ 表示为向量 v_i 与 p 之间的距离, 若上式成立, 则 v_0 为匹配的特征点, 否则, 匹配失败.

使用上述基于K-d树的BBF算法搜索匹配获得的匹配点对通常存在一定数量的误匹配. 为此, 本文采用了RANSAC算法进一步完成匹配工作.

2.2 RANSAC算法

RANSAC(Random Sample Consensus)算法^[12],即随机抽样一致性算法,是在矩阵估计、模式识别中最常用的对特征点进行提纯的方法.根据一组包含错误数据的样本集计算出数据的数学模型参数,得到有效样本数据,也就是说首先提出假设的模型,再通过已知数据进行验证,判定出该模型是否成立.通过最优模型的建立,可以判断出不符合模型的外点,也就查找出了特征点粗匹配中的误匹配,然后剔除误匹配,达到提纯的效果.本文采用RANSAC算法对图像进行精确匹配,同时得到图像间变换关系,具体步骤如下:

① 待拼接的两幅图像之间由于存在一部分的重叠区域,它们之间是存在相应变换的,本文采用的是投影变换模型来准确的描述图像之间的关系,投影变换模型可以准确描述图像间存在的变换情况,不仅可以对图像的旋转、缩放、平移和规则性的变换准确描述,也可以对一些不规则的变换进行描述.变换公式为:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (8)$$

式中, h_{11} 、 h_{12} 、 h_{21} 、 h_{22} 为缩放及旋转因子, h_{13} 、 h_{23} 为平移因子, h_{31} 、 h_{32} 为仿射因子.

② 在粗匹配的匹配点集 S 中随机抽取 4 对匹配点,判断其中是否有 3 点共线的,若有则舍弃重新随机抽取,直至 4 对匹配点中无 3 点共线的情况.得到 4 对匹配点后计算其对应的变换模型.

③ 上一步确定了一个模型之后,需要匹配点集 S 中所有的数据来验证该模型的有效性,即计算所有的粗匹配点是否满足该模型.能被模型描述的点定义为内点,相反,不能被模型描述的点定义为外点.统计出内点数量,分别求出内点匹配点和变换点之间的欧式距离,再进行求和处理.距离之和公式如下:

$$sum = \sum_{i \leq num} \sqrt{(x_i^j - x_i^i)^2 + (y_i^j - y_i^i)^2} \quad (9)$$

式中 num 为满足模型的内点数量.

④ 通过设定阈值来判断该距离之和是否满足该阈值来选择模型参数.

⑤ 重复②③两步,更新变换模型,直至得到最优模型为止,根据最优模型的内点来计算最终的模型参

数,即为所要求的变换矩阵 H ,并用 H 实现图像的配准^[13].

3 图像融合

经过图像匹配后,根据图像间变换矩阵 H 可以对相应图像进行变换来确定图像间重叠区域,对图像进行拼接形成一幅全景图像.需要注意的是,由于相机角度和曝光时间的各种差异导致拍摄到的图像可能存在亮度上的差异,图像边缘可能出现失真现象,这样会导致拼接后的图像缝合线处出现明显的明暗变化.为了消除图像拼接缝隙,实现无缝拼接,要对图像拼接处的缝合线进行平滑处理^[14].本文采用渐入渐出的加权平均融合法.

加权平均融合法类似于直接平均法,但其重叠区域的像素值不再是简单的叠加,而是先进行加权后再叠加平均.假设 f 为融合之后的图像, f_1 、 f_2 是待拼接的图像,拼接后图像如公式(6)所示为:

$$f(x,y) = \begin{cases} f_1(x,y) & (x,y) \in f_1 \\ \omega_1 f_1(x,y) + \omega_2 f_2(x,y) & (x,y) \in f_1 \cap f_2 \\ f_2(x,y) & (x,y) \in f_2 \end{cases} \quad (10)$$

式中 ω_1 和 ω_2 分别是第一幅和第二幅图像重叠区域对应像素的权值,并且满足 $\omega_1 + \omega_2 = 1$, $0 < \omega_1, \omega_2 < 1$,选择合适的权值可以使重叠区域实现平滑过渡,同时可以消除拼接缝隙.

4 实验结果及分析

本文主要实现的是基于 sift 特征点匹配的全景图像拼接功能,实验是在参数为 Intel Core i3-350M 2.27G/2G/Visual Studio 2012 的 PC 机上加以实现.首先利用相机拍摄两幅不同视角的图片,提取其特征点,然后采用基于 K-d 树的 BBF 算法对特征点进行粗匹配;再利用 RANSAC 算法对粗匹配进行提纯,删除“外点”,得到筛选后的匹配点对;利用这些匹配点对计算图像间变换矩阵,采用渐入渐出的加权平均法进行融合得到拼接图像,如图 1 所示.

图 1 中,(a)、(b)是两幅原始图像,(c)、(d)为特征点的提取图像,其中(c)图检测到特征点 473 个,(d)图检测到特征点 398 个,(e)为采用 BBF 算法匹配后的图像,共有 178 个匹配点,匹配时间为 0.125 秒,可以看出,在这 178 个匹配点中存在的误匹配对数比较多,(f)为 RANSAC 算法匹配后的图像,匹配时间为 0.082 秒,

可以看出经过 RANSAC 算法提炼后基本无误匹配点对, 计算出变换矩阵后, 对图像进行拼接, (g)为直接进行拼接效果图, 重叠区域有明显明暗变换, (h)为使用渐入渐出的加权平均融合算法得到的拼接图像, 由图 (h)可知, 图像拼接效果较好, 没有明显的亮度差异, 视觉效果自然。



图 1 图像拼接过程

为了验证该方法的有效性, 使用数码相机拍摄了 4 幅图片进行拼接实验, 每幅图片之间有大约 30%到 50%的重叠区域, 如图 2 所示上下左右四幅图像, 图 3 为拼接的效果图, 由结果可以看出处理后拼接痕迹消除, 实现了重叠区域的平滑过渡, 效果良好, 得到了高品质的全景图像。

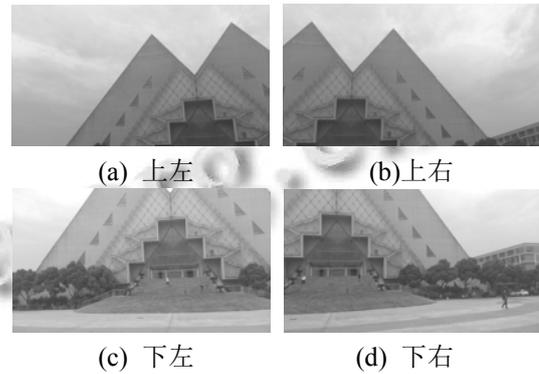


图 2 原始图像



图 3 拼接效果图

5 结论

本文主要对基于特征点的图像拼接算法进行了研究, 利用改进的 sift 算法提取图像的特征点, 并结合 RANSAC 算法实现了相邻帧图像特征点的精确匹配, 利用变换投影模型估计出两幅图像之间的变换矩阵, 最后利用渐入渐出的加权平均融合算法消除了图像拼接处的缝合线和色彩差异, 实现了全景图像的高质量拼接。

参考文献

- 1 Brown M, Lowe DG. Automatic panoramic image stitching using invariant features. IJCV. 2007. 59-73.
- 2 汪华琴. 基于特征点匹配的图像拼接方法研究[学位论文]. 武汉: 华中师范大学, 2007.
- 3 Lowe DG. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91-110.
- 4 Ke Y, Sukthankar R. PCA-SIFT: A more distinctive

- representation for local image description. Proc. of IEEE Computer Vision and Pattern Recognition Conference. 2004, 2. 506–513.
- 5 赵向阳,杜立民.一种全自动稳健的图像拼接融合算法.中国图形图像学报,2004,9(4):417–422.
- 6 Harris C, Stephen M. A combined Corner and edge detect. Proc. of 4th Alvey Vision conference. UK. 1988. 15–50.
- 7 李寒,牛纪帆,郭禾.基于特征点的全自动无缝图像拼接方法.计算机工程与设计,2007,28(9):2083–2085.
- 8 张恒.基于 sift 的图像拼接算法研究[学位论文].天津:河北工业大学,2012.
- 9 郭晓冉,崔少辉.基于局部特征点配准的图像拼接算法.半导体光电,2014,35(1):89–94.
- 10 王俊秀,孔令德.基于特征点匹配的全景图像拼接技术研究.软件工程师,2014,17(11):10–13
- 11 Jeff B, Lowe DG. Shape indexing using approximate nearest-neighbor search in high-dimensional spaces. Conference on Compute Vision and Pattern Recognition. 1997. 1000–1006.
- 12 Hartley R, Zisserman A. Multiple View Geometry in Computer Vision. London: 2nd Cambridge University Press, 2004.
- 13 宋宝森.全景图像拼接方法研究与实现[学位论文].哈尔滨:哈尔滨工程大学,2012.
- 14 林诚凯,李惠,潘金贵.一种全景图生成的改进算法.计算机工程与应用,2004,40(35):69–71,159.