

基于改进 SIFT 特征点匹配的图像拼接算法研究

张 勇^{1,2}, 王志锋¹, 马 文¹

(1 重庆邮电大学 通信新技术应用研究中心, 重庆 400065; 2 重庆信科设计有限公司, 重庆 400065)

摘 要: 基于传统的 SIFT 特征点匹配方法和图像像素级融合的思想, 提出一种改进的 SIFT 特征点匹配的图像拼接算法. 首先该算法在特征提取的过程中加入 Harris 角点算子, 筛选出鲁棒性较强的点; 其次采用基于圆形窗口的 48 维向量来进行特征降维, 并借助匹配特征点对之间的几何一致性来对特征点进行粗提纯, 提高算法运行效率; 最后采用重叠区域线性过渡融合算法对图像进行平滑过渡, 消除拼接缝隙. 实验验证了该算法的鲁棒性和快速性.

关键词: 图像拼接; Harris 角点; SIFT 特征; RANSAC 算法; 图像融合

中图分类号: TP391.4

文献标识码: A

文章编号: 1000-7180(2016)03-0060-05

Research Image Mosaic Algorithm Based on Improved SIFT Feature Matching

ZHANG Yong^{1,2}, WANG Zhi-feng¹, MA Wen¹

(1 Research Centre for Application of New Communication Technology, Chongqing

University of Posts and Telecommunications, Chongqing 400065, China;

2 Chongqing Information Technology Designing Co., Ltd, Chongqing 400065, China)

Abstract: Based on traditional SIFT feature point matching method and image pixel level fusion, an improved SIFT feature matching algorithm is proposed. Firstly the algorithm in the feature extraction process with Harris corner detector and picks out robust points; Followed use of 48 dimensional vector based on a circular window to dimension reduction and the feature points are refined by the geometrical consistency of the matching feature points, to improve the efficiency of the algorithm. Finally, the overlapped area linear transition fusion algorithm of image smoothing the transition to eliminate the splicing gap. The experiment verifies the robustness and rapidity of the proposed method.

Key words: image mosaic; Harris corner; SIFT feature; RANSAC algorithm; image fusion

1 引言

图像拼接就是将多幅具有一定重叠部分的图像进行融合, 合成一张宽视野、没有明显拼接缝隙的高分辨率图像. 目前图像拼接技术广泛应用于计算机图像处理、海底勘探、医学图像、虚拟现实、遥感影像处理等领域^[1].

图像拼接的算法基本可以分为以下两种^[2]: 基于灰度信息的方法和基于特征的方法. 基于灰度信息的方法是通过使用待拼接图像的像素值来计算两幅图像的相似程度, 从而确定拼接的重叠区域, 实现

图像的拼接, 这种方法计算量大, 并且鲁棒性较差. 而基于特征的方法是通过提取图像上有用的特征信息, 然后对两幅图像的特征进行匹配来获取图像之间的映射关系, 该方法具有较高的鲁棒性, 目前大多数拼接算法都采用这种方法. David Lowe 在 2004 年提出了尺度不变特征变换(SIFT)的算法^[3]. 该算子对两幅图像之间的尺度变化、旋转变化、亮度变化、仿射变化等具有较为稳定的匹配性能. 但该算法提取的特征点过多, 容易造成误匹配, 并且该算法计算耗时过长, 难以满足实时性.

2 图像拼接的流程

图像拼接过程一般包含图像预处理、图像匹配、图像融合这三个步骤. 图像预处理一般包括图像去噪和图像修正,其目的是更好地进行图像匹配,降低匹配的难度;图像匹配就是在有重叠部分的两幅图像中寻找找到相重合的点,然后利用寻找到的空间变换完成图像的对齐操作,匹配的精度对于最终的拼接效果影响很大;图像融合就是将完成图像匹配的对齐图像拼接到一张图像中去,并消除拼接后的颜色差异和拼接缝隙. SIFT 特征点匹配的图像拼接算法流程如图 1 所示.

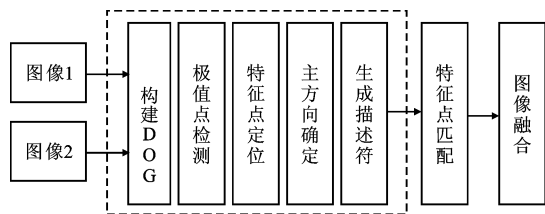


图 1 SIFT 图像拼接算法流程图

3 SIFT 算法

SIFT 特征是基于图像尺度空间理论的局部特征描述子,它利用高斯差分算子构建尺度空间,并在尺度空间中寻找极值点,提取位置信息,然后通过统计其梯度信息和方向,生成 SIFT 描述子,然后对 128 维描述子进行匹配,减小因图像发生旋转、光照、尺度、噪声等变化造成的错误匹配概率.

3.1 尺度空间极值点检测

SIFT 算法通过高斯差分尺度空间对极值点进行检测,高斯卷积核是进行尺度变换的唯一线性核^[4]. 若 $I(x, y)$ 为原始输入图像, $G(x, y)$ 是一个尺度变换的高斯函数,那么将一个二维图像的尺度空间函数 $L(x, y, \sigma)$ 定义为

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

其中,

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

式中, (x, y) 是图像像素点的坐标, σ 是尺度因子, σ 的大小决定了图像平滑程度,图像的概貌可以通过大尺度来表示,而图像的细节可以通过小尺度来表示.

为了能够更好地检测到相对稳定的极值点, Lowe 提出了更为高效的 DOG 算子. DOG 是高斯金字塔中上下相邻两层图像的差值,表示如下:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (3)$$

得到图像中的高斯差分金字塔之后,就可以计算尺度空间中的极值点了. 需要将尺度空间中每一个像素点与它同尺度的 3×3 邻域内的 8 个相邻点进行比较,同时,还要和相邻尺度中对应的 18 个像素点进行比较,即总共与 26 个点进行比较. 如果该点的灰度值是 26 点中的最大或最小值,那么该点就被定义为候选极值点.

检测到极值点之后,还需要对其进行进一步筛选. 第一步是利用尺度空间 $D(x, y, \sigma)$ 函数的泰勒二次展开式进行曲线拟合,第二步是获取该点的 2×2 Hessian 矩阵,并计算其主曲率.

Taylor 展开式为

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (4)$$

Hessian 矩阵为

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (5)$$

通过该步骤的处理,对比度较低的极值点和一些不稳定的边缘点可以部分去除,这样处理的好处就是算法的匹配稳定性增强了,同时抗噪声能力也提高了.

3.2 计算特征点的幅值和方向

为了使最终生成描述子在图像发生旋转的情况下依然能够很好的匹配, SIFT 算法利用统计直方图信息的方式来对特征点的主方向进行确定. 对于在高斯差分金字塔中找出的极值点,计算其所在的高斯金字塔邻域窗口内像素梯度的模值 $m(x, y)$ 和方向 $\theta(x, y)$,计算公式为

$$\begin{aligned} m(x, y) &= ((L(x+1, y) - L(x-1, y))^2 + \\ &\quad (L(x, y+1) - L(x, y-1))^2)^{\frac{1}{2}} \end{aligned} \quad (6)$$

$$\theta(x, y) = \arctan \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (7)$$

式中, L 为特征点所在的尺度. 梯度直方图将 $0 \sim 360$ 度划分为 36 个柱,其中每 10 度为一个柱. 直方图的峰值代表该特征点邻域的主方向. 为了增强算法的鲁棒性,原算法还给每一个极值点定义了一个辅方向.

3.3 生成特征点描述符

通过以计算,每一个特征点都拥有位置、方向、尺度这三个信息. 为了保证特征点的旋转不变性,需

要将坐标轴旋转到关键点所在的主方向。以特征点为中心取一个 16×16 的矩形窗口,把窗口均匀地分为 16 个 4×4 的子窗口。然后在每一个 4×4 的子窗口上计算 8 个方向的梯度累加值。每一个 4×4 的子窗口可形成一个种子点,因此每一个特征点都由 16 个种子点构成,每个种子点都含有 8 个方向的向量信息,所以一个特征点最终可以形成 128 维的特征向量。为了减少光照变化带来的影响,还需要对生成的特征向量进行归一化,以期提高匹配的稳定性。

3.4 特征向量的匹配

特征向量的匹配一般采用欧式距离对 128 维的特征描述子进行相似性度量。SIFT 特征点匹配采用的是穷举法,这样会导致算法耗费的时间过多,因此本文采用基于 k-d 树的 BBF 搜索策略。从待匹配图像中选择一个特征点,从原始图像中找出与该点最近距离的特征点和次近距离的特征点,从而得到最近邻与次近邻的比值,若比值小于规定的阈值,则认为最近距离的点为匹配点。

4 改进的 SIFT 算法

4.1 SIFT-Harris 算法

通过计算特征点的幅值和方向,每一个特征点都包含有位置、尺度、方向的信息。假如总共提取了有 n 个特征点,每个特征点的信息可以表示为 $I(x_n, y_n, \sigma_n, \theta_n)$ 。 x_n, y_n 是特征点的横坐标和纵坐标, σ_n 为特征点的尺度, θ_n 为特征点的方向。Harris 角点检测的自相关矩阵可以表示为:

$$\mathbf{M} = \sum_{(x,y) \in W(x_n, y_n)} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (8)$$

$$w(x, y) = e^{-\frac{(x^2+y^2)}{\sigma^2}} \quad (9)$$

式中, I_x, I_y 表示图像在 x 方向和 y 方向的偏导数, $w(x, y)$ 是以 (x_n, y_n) 为中心的高斯加权窗口函数, 标准差为 σ_n 即特征点的尺度。设自相关矩阵 $\mathbf{M}$ 的两个特征值为 λ_1, λ_2 , 当 λ_1, λ_2 都很小时, 则认为该点对应于图像中的平滑区域; 当 λ_1, λ_2 都很大时, 对应于图像中的角点; 当 λ_1, λ_2 一个很大一个很小时, 对应于图像中的边缘点。由于特征值的计算需要很大的计算量, 所以使用如下函数来判断角点:

$$R = \lambda_1 \lambda_2 - k (\lambda_1 + \lambda_2)^2 = \det(\mathbf{M}) - k \text{trace}^2(\mathbf{M}) \quad (10)$$

式中, k 是一个常数, 取值为 $0.04 \sim 0.06$, $\det(\mathbf{M})$ 是矩阵 $\mathbf{M}$ 的行列式, $\text{trace}(\mathbf{M})$ 为矩阵 $\mathbf{M}$ 的迹也为矩阵 $\mathbf{M}$ 的两个特征值之和。然后设定一个阈值 T , 当角

点量 R 大于 T 时, 则判定此点是候选角点, 否则就不是候选角点^[5]。本文采用一种自适应阈值的方法, 阈值 T 的计算公式如下:

$$T = \frac{1}{2N} \sum_{n=1}^N |R_{\sigma_n}(x_n, y_n)| \quad (11)$$

即计算出所有特征点的角点值 R , 并求其平均值的一半作为阈值 T 。剔除 R 值小于阈值的点, 则保留下来的特征点具有较强的鲁棒性。

4.2 改进的特征描述子

图像在发生旋转变换时, 特征点周围的区域会发生变化, 考虑到圆具有良好的旋转不变性, 因此本文采用圆形区域来构造特征描述符。

Lowe 指出, 在特征点周围 16×16 的邻域内包含有足够的信息, 并且不会造成太大的计算量, 因此以特征点为中心, 取一个半径为 8 的同心圆环区域。如图 2 所示, 圆内总共有 4 环, 采用不同的步长, 即采用 3, 2, 2, 1 的环数。因为离特征点越近的环形区域对特征点的影响越大, 如果每一环半径都取 1, 当圆环旋转时, 仍然会有一部分离散的像素点不在原来的环中, 为了降低图像旋转引起的误差, 本文采用递减的环数来表示特征点的邻域区域。统计每个圆环上的 12 个方向上的梯度累计直方图, 并采用高斯加权窗口对梯度值进行加权。然后对每个圆环内 12 个梯度累加值进行排序, 并将特征向量进行归一化处理, 最终得到 48 维的特征描述符。

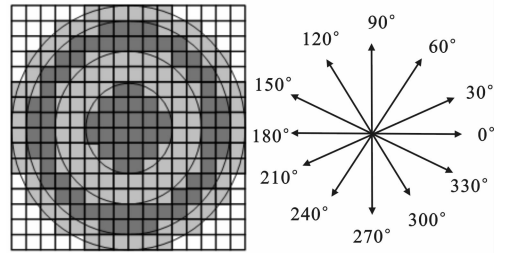


图2 改进的 SIFT 特征描述符

4.3 添加约束条件进行粗提纯

采用欧式距离进行匹配往往会有很多的误匹配点, 通常采用 RANSAC 算法来剔除误匹配点^[6]。本文针对 RANSAC 算法内点概率小, 导致算法计算效率低的问题, 提出了一种约束条件的方法来对匹配点对进行筛选, 提高了算法的运算效率。

在研究 SIFT 特征时, SIFT 特征包括 4 个部分, 特征点的位置、尺度、方向以及 128 维的特征向量。在研究 SIFT 特征时, 对图像大小为 300×211 的 beaver.png 图像进行旋转 $20^\circ, \dots, 140^\circ, 160^\circ, 180^\circ$; 同时将图像缩小为原图的 $90\%, 80\%, \dots, 10\%$; 然后

与原图进行 SIFT 特征提取与匹配,可以得到 9 组匹配结果,图 3 所示为图像旋转 40° 并且缩小 80% 后的 SIFT 特征匹配结果,带箭头的粗线为 SIFT 特征点的主方向和尺度,箭头指向的方向为特征点的主方向,箭头的长度为特征点的尺度。

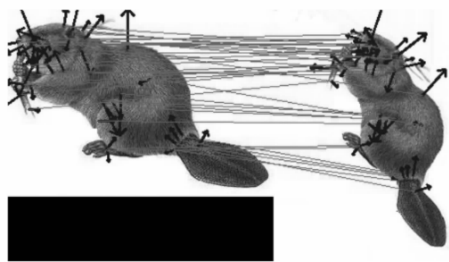


图 3 旋转 40° 且缩小 80% 的 SIFT 匹配结果图

通过对 9 组图像匹配结果进行比较发现,当图像发生尺度和旋转变换时,对应的正确匹配点的主方向和尺度满足一定的规律。即正确匹配的特征点尺度的比值近似等于两幅图像尺度缩放的比值;正确匹配的特征点主方向角度的差值近似等于两幅图像旋转变化的角度。假设 $(a_1, b_1), (a_2, b_2), (a_3, b_3), (a_4, b_4)$ 为正确匹配的四个点,则上述结论可以用下面两个公式来表示:

$$\frac{\sigma_{a_1}}{\sigma_{b_1}} \approx \frac{\sigma_{a_2}}{\sigma_{b_2}} \approx \frac{\sigma_{a_3}}{\sigma_{b_3}} \approx \frac{\sigma_{a_4}}{\sigma_{b_4}} \quad (12)$$

$$|\theta_{a_1} - \theta_{b_1}| \approx |\theta_{a_2} - \theta_{b_2}| \approx |\theta_{a_3} - \theta_{b_3}| \approx |\theta_{a_4} - \theta_{b_4}| \quad (13)$$

因此本文借助匹配特征点对之间的几何一致性来对粗匹配的特征点进行粗提纯。假设图像 A 和 B 表示待匹配的图像,特征点集分别取图像 A 和图像 B 粗匹配点的集合 $a = (a_1, a_2, \dots, a_n)$ 和 $b = (b_1, b_2, \dots, b_n)$ 。并设粗匹配点集的尺度比值集合和角度差值集合为 M, N 则:

$$M = \left\{ \frac{\sigma_{a_1}}{\sigma_{b_1}}, \frac{\sigma_{a_2}}{\sigma_{b_2}}, \dots, \frac{\sigma_{a_n}}{\sigma_{b_n}} \right\} \quad (14)$$

$$N = \{ |\theta_{a_1} - \theta_{b_1}|, |\theta_{a_2} - \theta_{b_2}|, \dots, |\theta_{a_n} - \theta_{b_n}| \} \quad (15)$$

并规定 m_σ 为计算出的图像尺度缩放的比例, n_θ 为计算出的图像旋转的角度。为了选取最合适的阈值,对 20 组图像分别进行缩放,尺度变换,并对正确的匹配特征点对进行分析,得出最合适的阈值为 0.06,因此本文将阈值 T 设为 0.06。文献[2]采用的是相邻匹配特征点的距离大致相等的特性来对匹配点进行筛选,当图像发生大尺度的变化时,文献[2]

中的阈值也要进行相应的调整,阈值不具有稳定性。算法的改进步骤如下,

从 $i = 1$ 开始,依次求出 $\left| \frac{\sigma_{a_i}}{\sigma_{b_i}} - m_\sigma \right|$ 和

$|\theta_{a_i} - \theta_{b_i} - n_\theta|$ 的值,如果这两个值均少于阈值 T ,保留此匹配点。否则,剔除此点。最后采用 RANSAC 算法对筛选出来的特征点进行提纯。

5 图像融合

在图像拼接的过程中,当待拼接的图像受到噪声、几何校正差异、光照不均等影响时,最后的拼接图容易产生拼接缝隙^[7]。为了得到一幅大尺寸的、无缝的图像,同时保证拼接后的图像灰度的连续性,就需要采用适合的图像融合方法来进行处理,消除拼接缝隙。目前常用的融合算法有平均值融合和加权融合算法。当两幅图像光照差别较大时,采用平均值融合算法很难消除因为图像光照强度或色彩不连续而造成的拼接缝隙,加权融合算法需要适当地选择权值才能实现重叠区域平滑过渡,如果权值选用不当也会对拼接效果产生影响,因此本文采用重叠区域线性过渡融合的方法进行图像拼接。

如图 4 所示,假设两幅图像重叠区域宽度为 L ,过渡因子为 σ , σ 是个在 0 和 1 之间的一个数。如果两幅图像的重叠区域在水平轴方向上的最大值和最小值分别是 $x_{\max}$ 和 $x_{\min}$,那么过渡因子

$$\sigma = \frac{x_{\max} - x}{x_{\max} - x_{\min}} \quad (16)$$

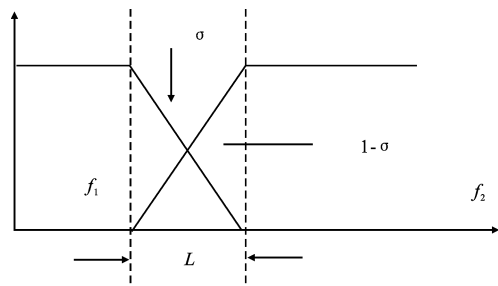


图 4 图像线性过渡图

令 $f_1(x, y), f_2(x, y)$ 分别表示待拼接的图像 1、图像 2 的灰度值, $f_3(x, y)$ 表示融合两幅图像之后的图像 3 的灰度值,那么重叠区域 $f_3(x, y)$ 的计算公式为

$$f_3(x, y) = \sigma f_1(x, y) + (1 - \sigma) f_2(x, y) \quad (17)$$

由于 σ 的值由 1 渐变为 0, $1 - \sigma$ 的值由 0 渐变为 1,权值在不断的变化,因此式(17)即可实现两幅图像在重叠区域的平滑过渡,达到无缝拼接的效

果.

6 实验结果及分析

实验平台为 Windows 7 系统,2.6 GHz 主频,4 GB 内存,利用 vs2010 + opencv2.4.4 平台进行仿真实验.

本文分别用 SIFT 算法、PCA-SIFT^[8-9] 算法及文献[2]采用的筛选算法和本文改进的算法,对 4 组不同类型的图像进行实验测试,并对实验结果的匹配准确率和匹配时间进行统计,结果如表 1 所示.

表 1 实验数据统计					
类别		实验 1	实验 2	实验 3	实验 4
SIFT	匹配准确率/%	83.5	76.8	84.3	86.1
	匹配时间/s	6.5	5.3	7.6	6.8
PCA-SIFT	匹配准确率/%	84.2	79.9	85.5	88.4
	匹配时间/s	4.4	3.2	5.4	3.5
文献[2]	匹配准确率/%	88.7	82.1	89.9	87.2
	匹配时间/s	5.2	4.5	6.1	5.3
本文算法	匹配准确率/%	96.3	88.4	95.2	96.8
	匹配时间/s	4.2	2.7	4.7	3.2

通过表 1 中的数据可以发现,在匹配正确率方面,PCA-SIFT 算法的匹配准确率稍微高于 SIFT 算法,文献[2]的匹配准确率高 SIFT 算法,本文方法在匹配正确率方面的优势高于前面的三种算法.在匹配时间方面,SIFT 算法的效率低于 PCA-SIFT 算法和文献[2]提出的算法,本文算法的匹配时间明显优于前面的三种算法.

对特征点进行匹配之后,接下来就是确定两幅图像的重叠区域,并对重叠区域进行平滑融合达到无缝拼接的效果.

可以看出,本文的算法能够有效的平滑图像,消除拼接缝隙,实现两幅图像的无缝拼接.

7 结束语

提出了一种 Harris-SIFT 算法,该算法可以提取出鲁棒性更好的特征点.为了提高计算效率,加入了基于圆形窗口的 48 维特征描述子,进行特征降维;

并根据匹配特征点对之间的几何一致性进行粗提纯,减少了 RANSAC 算法的时间;最后,利用重叠区域线性过渡融合算法消除图像之间的拼接缝隙,达到了两幅图像高质量的拼接效果.实验结果表明,该方法是一种高精度、高效率且拼接效果良好的无缝拼接算法.

参考文献:

[1] 宋佳乾,汪西原. 基于改进 SIFT 特征点匹配的图像拼接算法[J]. 计算机测量与控制,2015,23(2):512-515.

[2] 焦丽龙,韩燮,李定主. 改进的基于特征点匹配的图像拼接融合算法[J]. 计算机工程与设计,2014,35(3):918-922.

[3] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004,60(2):91-110.

[4] 王田甲,刘国荣. SIFT 改进算法在图像配准中的应用[J]. 微电子学与计算机,2011,28(5):184-188.

[5] 尚明珠. 基于改进 SIFT 特征匹配的快速图像拼接算法[J]. 微电子学与计算机,2014,31(1):64-67.

[6] 王小平,王建勇,杨坝. 基于 Sobel 算子和改进 SURF 算法的图像拼接方法[J]. 电视技术,2014,38(13):43-46.

[7] 胡同喜,牛雪峰,谭洋,等. 基于 SURF 算法的无人机遥感影像拼接技术[J]. 测绘通报,2015,60(1):55-58.

[8] Ke Y, Sukthankar R. PCA-SIFT: A more distinctive representation for local image descriptors [C]// Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition. Washington DC, USA, 2004: 511-517.

[9] 王鹤,谢刚. 基于 PCA-SIFT 特征的目标识别算法[J]. 电视技术,2013,37(15):30-32.

作者简介:

张 勇 男,(1970-),高级工程师,硕士生导师. 研究方向为通信新技术应用.

王志锋(通讯作者) 男,(1988-),硕士研究生. 研究方向为数字图像处理. E-mail:278711005@qq.com.

马 文 男,(1988-),硕士研究生. 研究方向为数字图像处理.