

基于特征不变描述的图片无缝镶嵌技术研究

杨争争, 李良福

(陕西师范大学 计算机科学学院, 陕西 西安 710062)

摘要:在图像镶嵌算法中,SIFT 算法本身就具有非常强的匹配能力,它可以处理两幅图像之间存在着的平移、旋转、尺度缩放等仿射变换,因此它是现今图像镶嵌中比较常用的方法。但在图像镶嵌中,如果直接使用 SIFT 算法会使运算量非常庞大,影响速度。针对这一问题,文中首先利用相位相关法来确定图像之间大致的重叠区域,再在大致的重叠区域提取 SIFT 特征点;其次利用特征点实现配准,并且在全局参数优化时考虑了径向畸变;最后用金字塔融合算法对图像进行融合。实验结果表明该算法能有效提高图像镶嵌的速度和精度,速度约提升三分之一,实现了一幅无缝的全景图像。

关键词:图像镶嵌;相位相关;金字塔融合;径向畸变

中图分类号:TN911.73

文献标识码:A

文章编号:1673-629X(2015)01-0096-04

doi:10.3969/j.issn.1673-629X.2015.01.022

Seamless Image Mosaic Technology Research Based on Characteristics of Same Description

YANG Zheng-zheng, LI Liang-fu

(College of Computer Science, Shaanxi Normal University, Xi'an 710062, China)

Abstract: In the image mosaic algorithm, SIFT matching algorithm has strong ability to handle the presence of translation, rotation, scaling affine transformation between two images, so it is now more commonly used image mosaic approach, but in the image mosaic, if using directly SIFT algorithm makes computation very large, impacting the speed. Aiming at this problem, first determine the approximate region of overlap between the images by phase correlation method, and then extract the SIFT feature points in the overlapping area roughly, followed by the use of feature points to achieve alignment, and consider the radial distortion when optimizing the global parameters, finally pyramid fusion algorithm is used for image fusion. The experimental results show that the algorithm can be very effective in improving the speed and accuracy of image mosaic, the speed increased to about third, to achieve a seamless panoramic image.

Key words: image mosaic; phase correlation; pyramid fusion; radial distortion

0 引言

图像镶嵌是计算机视觉领域的一个非常重要的组成部分,此技术被非常广泛地用于许多方面,例如:计算机视觉、遥感图像处理、医学图像处理、军事等各个方面。图像镶嵌技术是把许多图像序列组成一幅全景图像,图像镶嵌大致由图像预处理、图像配准、图像融合三部分组成。其中图像配准是整个图像镶嵌的非常关键的部分。

目前图像配准一般分为基于图像灰度的配准方法、基于特征配准方法和基于变换域的配准方法。基于图像灰度的配准方法主要有相关匹配算法^[1]、Mean Shift 算法等;基于特征的配准方法主要有 Harris 特征

点算法^[2]、SIFT 尺度不变特征描述算法^[3-8]等;基于变换域的配准方法主要有相位相关法^[9]等。SIFT 算法有较强的匹配能力,对图像的旋转、尺度缩放及仿射变换有较好的鲁棒性,针对从任意角度拍摄的图像也有较好的特征匹配效果。

近年来针对 SIFT 的图像镶嵌方法得到了广泛研究,但该算法由于特征提取的比较复杂,因此导致此算法具有运算量大、计算时间长的缺点。文中针对 SIFT 算法的不足,提出一种新的基于特征点的图像拼接方法。实验结果表明该算法构成一幅宽视野的、完整的、高分辨率的全景图像,能有效提高图像拼接的速度和精度。

收稿日期:2014-01-18

修回日期:2014-04-24

网络出版时间:2014-11-17

基金项目:国家自然科学基金资助项目(61201434)

作者简介:杨争争(1988-),男,硕士研究生,研究方向为计算机视觉;李良福,副教授,通信作者,研究方向为计算机视觉。

网络出版地址:<http://www.cnki.net/kcms/detail/61.1450.TP.20141117.2205.014.html>

1 基于相位相关的图像配准方法

如果直接对多幅图像运用 SIFT 图像镶嵌算法,不仅计算量庞大而且浪费时间,可以根据相位相关法估计出图像平移量,从而得到两幅图像之间大致重叠区域,然后再在此大致重叠区域运用 SIFT 图像镶嵌算法。相位相关法是 Kuglin 等^[9]一些研究者在 1975 年提出来的,利用了傅里叶变换的平移性质,即它的互功率谱的相位等价于图像之间相位差,因为一般信号的卷积是对应频率上的乘积。

该相位相关算法的原理如下:设两幅图像 $I_1(x, y)$ 和 $I_2(x, y)$, $I_2(x, y)$ 是 $I_1(x, y)$ 经过平移 (x_0, y_0) 之后的图像,它们之间明显具有的关系如下:

$$I_2(x, y) = I_1(x - x_0, y - y_0) \quad (1)$$

根据已知的傅里叶变换有关知识,如果有这么两幅图像 $I_2(x, y)$ 和 $I_1(x, y)$ 对应的傅里叶变换分别为 $G(u, v)$ 和 $F(u, v)$, 则满足以下关系:

$$G(u, v) = F(u, v) e^{-j2\pi(ux_0 + vy_0)} \quad (2)$$

即经过平移变换后的图像之间其具有的傅里叶变换将有相同的振幅,通过计算两幅图像之间的互功率谱可以得到相位差,互功率谱定义如下:

$$\frac{F(u, v) G^*(u, v)}{|F(u, v) G^*(u, v)|} = e^{-j2\pi(ux_0 + vy_0)} \quad (3)$$

其中, $G^*(u, v)$ 是 $G(u, v)$ 的复共轭。互功率谱的相位将等价于图像之间的相位差,通过对式(3)做傅里叶逆变换,在 (x_0, y_0) 处将形成一个脉冲函数 $\delta(x - x_0, y - y_0)$, 该函数其在平移量的位置它的取值将为 1, 其余的地方其取值为 0, 相关的峰值大小是一种反映两幅图像的相似度大小的参考,在脉冲峰值位置就指的是两幅图像之间的大致相对平移量 x_0 和 y_0 。这种方法是一种对图像先进行粗匹配,根据估计出的平移量可以估计出大致重叠区域,然后提取特征点,这种方法对进行基于特征配准提供了先验知识。

2 基于 SIFT 特征的图像拼接

Lowe^[3-8]提出了里程碑意义的 SIFT 著名的图像特征描述算子,此算法得到了国内外广泛的研究,是一种非常经典的算法。它是一种非常有效的基于尺度空间的、基于图像局部的特征描述算子,这种特征描述子能有效地解决图像之间可能包含的一系列变换并且对其保持不变性。例如:图像之间可能包含尺度变换、光照变换、平移、旋转等一系列的变换,对 SIFT 提取是通过高斯金字塔模拟多尺度下的图像,并构建差分金字塔 DOG, 然后检测 DOG 金字塔下的极值点来获取稳定的特征点。SIFT 算法提取出的特征点对尺度变换、平移、旋转等具有不变性,同时对光照的变化也有一定的适应性,到现在为止,一般在图像镶嵌中 SIFT 算法提

取主要包括以下几方面的内容。

2.1 SIFT 算法提取

2.1.1 计算 SIFT 特征极值点

在日常生活中,人们所看到的信息中都是有尺度的,所以在图像中引入一个尺度因子,这样就可以使得图像信息转换到其尺度空间,使得因为尺度改动的原因对特征产生损失。级联滤波器是一种非常有效的方法能得到候选的位置和尺度下的特征点。

Lindeberg 在 1994 通过大量研究,通过了一系列合情合理的假设,实验证明,高斯核函数是所有函数中唯一一个可以用来对尺度空间有效地进行表示,所以,应该用在不同的尺度因子情况下的高斯函数和原有的图像进行卷积运算来得出尺度空间的函数 $L(x, y, \delta)$:

$$L(x, y, \delta) = G(x, y, \delta) * I(x, y) \quad (4)$$

为了有效检测出空间中稳定的关键点位置,可以使用高斯差分函数与原来的图像进行卷积得到:

$$D(x, y, \delta) = (G(x, y, k\delta) - G(x, y, \delta)) * I(x, y) = L(x, y, k\delta) - L(x, y, \delta) \quad (5)$$

可以通过其中一个点所在的 DOG 尺度空间的本身及其在它的上下两层的 26 个邻域中作出比较,如果其是最大值或者是最小值时,就认为它是图像在该尺度空间下的一个特征点。

2.1.2 消除掉一些不确定的精确极值点

当得到了关键点的坐标之后,下一步应该做的是需要更加精确地确定出图像所在的方向、图像所在的尺度等参数。这样的话除去一些对比度较低的点,可有效解决一些在边缘等情况下对比度非常低的点。在这里用 $D(x)$ 来得出上面所获得的点是不是很稳定。使用 $D(x)$ 在该点处进行展开后可得以下结果:

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (6)$$

令公式(6)等于 0, 可得 $\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x}$, 则有

$D(x)$ 为:

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (7)$$

对于 $D(x) < 0.03$ 的情况下的点是一个应该丢弃的点,这个点应该被省略。

2.1.3 为关键点进行方向的分配

为了使图像能更好地保持其不变性, SIFT 算法将会为每一个特征点提供一个主方向。在特征点描述子进行构造的时候要考虑该点的主方向的一些信息,而在对图像进行匹配的时候能消除因为摄像机拍出的图像由于旋转所带来的影响。

利用高斯函数 $L(x, y, s)$ 来得到其有效的梯度直方图,为其提供一些主方向。首先是要为特征点计算

出其值,并且给出其方向。

$$m(x, y) =$$

$$\sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (8)$$

$$\theta(x, y) = \tan^{-1} \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (9)$$

其中, $m(x, y)$ 为梯度的值; $\theta(x, y)$ 是梯度的方向; L 函数为关键点的所在尺度。

用梯度直方图统计出其方向的一些信息, 这些信息被分为 36 份。

2.1.4 确定其特征向量

通过以上的研究, 得出了特征点其所在的方向、在哪个尺寸下及大小后, 为了得到其特征描述向量, 首先, 用图像中心的 8×8 区域对其进行采样, 然后对这个区域建立一个方向直方图, 让其分别在 8 个方向计算其直方图, 下来将该区域分成 4 个部分来计算区域的直方图, 这代表着用 2×2 表示梯度直方图, 但是在实际情况下, 是对 16×16 分成 4×4 计算直方图, 这样每个格子代表了 8 个方向的梯度信息将形成 128 维的特征向量。

2.2 特征匹配

为了得到两个特征相匹配, 运用 k-d 树对其进行搜索, k-d 树能很好地提高其匹配点的搜索速度。找出每个特征点的最近邻特征点和次近邻特征点并计算出它们之间的相应欧氏距离, 根据特征点与最近邻域的点除以次近邻域的特征点得到结果, 如果结果小于 0.6, 则认为它们已经是一个正确匹配到的特征点, 否则认为其没有匹配。

3 变换矩阵计算

3.1 RANSAC 匹配点对提纯

计算出单应矩阵, 用 RANSAC^[10-13] 算法进行精确估计, 这种方法能有效地消除外点, 只剩下满足要求的内点, 通过这些内点就能很好地计算出两幅图像之间的变换矩阵。RANSAC 算法是 1981 年 Fischler 等人提出的一种非常有效的消除误匹配的技术, 它现在已经很好地用在了计算机视觉中图像匹配的各个方面, 是计算机视觉中的一种非常有效的方法。它能很好地得出一些最大匹配的匹配特征集, 它将匹配的特征点分为内点(正确满足要求的匹配的特征点)和外点(反之, 错误的匹配特征点), 下来就是应该选择一个其中所有内点集中最多的一个内点集, 然后在这个内点集中选择四对匹配内点重新估计变换矩阵。它的思想是先对特征匹配的点对集进行分配, 随机选中其中的四对特征点, 但是对特征点有一个要求, 即选取的特征点其中任何三点它们都不能共线, 否则重新选取四对特征

点, 下来就是根据特征对进行变换模型, 求出其变换参数。根据计算出来的参数估计出匹配对点之间的几何距离, 在这里给出一个得到内点的阈值, 如果这个匹配对之间的距离小于此阈值的话, 认为其是一个内点, 如果它的距离大于阈值, 则认为其为外点, 即使一个错误匹配对, 也对其进行删除。对上面的步骤进行不断计算, 可以得到一个内点数量最大的点集, 最后, 通过此内点集, 计算出模型的变换矩阵, 这个矩阵就是所要求的结果。

3.2 L-M 优化计算透视矩阵

为了消除变换参数之间以及由于消费者的摄像机中可能出现畸变(这里只考虑径向畸变)而引起的累积误差, 设 k_1 和 k_2 为径向畸变的参数, 畸变点 $u^d = (x^d, y^d)$ 和未畸变点 $u^u = (x^u, y^u)$ 之间的关系如下:

$$u^u = u^d (1 + k_1 u^2 + k_2 u^4) \quad (10)$$

$$u^2 = \left(\frac{x^d}{f}\right)^2 + \left(\frac{y^d}{f}\right)^2 \quad (11)$$

其中, f 是焦距长度。

假设畸变中心为图像中心, 即坐标系原点。设图像 I_1 的实际坐标 $U = (u, v)^T$, 校正后的理想坐标 $U' = (u', v')^T$ 。设图像 I_2 的实际坐标 $V = (u_1, v_1)^T$, 校正后的理想坐标 $V' = (u_1', v_1')^T$ 。两个图像坐标之间的变换是投影变换, 即

$$V' = M U' \quad (12)$$

其中, M 为变换矩阵。设有 n 个匹配点对 (I_1, J_1) , $(I_2, J_2), \dots, (I_n, J_n)$, 假设 k_1, k_2 初始为零, 即 I_i 和 $J_i (i = 1, 2, \dots, n)$, 用 n 个匹配点对来计算出 M 的初始估计, 然后用式(12) 计算出 I_i 对应的 J_i' , 令 $E = \sum_{i=1}^n (J_i' - J_i)^2$ 最小, 文中用 Levenberg-Marquardt 方法计算出 M, k_1, k_2 。

4 图像融合

在上面, 已经计算完了图像的变换矩阵, 通过计算后的变换矩阵, 就能通过运算求出两幅图像的重叠区域, 如果对图像只做一些简单的叠加, 这样会导致镶嵌后的结果会存在明显的镶嵌痕迹, 在这种情况下就需要用到图像融合技术, 图像融合的目标就是为了让图像变得更加光滑, 去掉其镶嵌痕迹。文中采用金字塔融合算法, 它的方法如下:

该方法是 1983 年由 Burt 等^[14] 提出来的, 这种方法在计算机视觉中有着很广泛的应用。其基本思想就是首先对两幅待镶嵌的图像进行一系列的多分辨率的分解, 从而得到一系列分解后的图像, 接下来是对这些图像让其在各自的分辨率下进行镶嵌, 最后用重构算法构造来得到分辨率下的图像。

图1和图2分别是使用渐入渐出算法和文中算法得到的示例图。



图1 使用渐入渐出算法示例图



图2 使用文中算法示例图

5 结束语

近些年,图像镶嵌技术已经引起了全世界的广泛关注,由于其有很好的应用价值已经成为了数字图像处理 and 计算机视觉的重点研究领域。文中提出了一种对 SIFT 特征点提取的方法进行改进的全景镶嵌方法。该方法主要针对在镶嵌的过程中直接使用 SIFT 算法运算量庞大以及全局参数优化时考虑了径向畸变方面进行改进,提高了图像镶嵌的精度和速度(速度约提升三分之一)。使用金字塔融合算法能有效地消除缝隙,使成为一幅全景图像。图像镶嵌领域将会是一个非常活跃的方向,而且对人类的日常生活起到了非常大的帮助,未来将会有许多新的算法出现,但也会伴随许多问题。图像镶嵌是一个非常活跃的研究方向,文中虽然在这些方面取得了一些成果,但是还有许多方面需要进一步的研究。

参考文献:

- [1] Kontitsis M, Valavanis K. A cost effective tracking system for small unmanned aerial systems[J]. *Journal of Intelligent Robotic Systems*, 2010, 57(1): 171-191.
- [2] Zagrouba E, Barhoumi W, Amri S. An efficient image-mosaicing method based on multifeature matching[J]. *Machine Vision and Applications*, 2009, 20(3): 139-162.
- [3] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 90-110.
- [4] Brown M, Lowe D G. Automatic panoramic image stitching using invariant features[J]. *International Journal of Computer Vision*, 2007, 74(1): 59-73.
- [5] 程咏梅, 张绍武, 花永强, 等. 基于 SIFT 联合描述子的航拍视频图像镶嵌[J]. *西北工业大学学报*, 2010, 28(1): 51-56.
- [6] 鲁恒, 李永树, 何敬, 等. 一种基于特征点的无人机影像自动拼接方法[J]. *地理与地理信息科学*, 2010, 26(5): 16-19.
- [7] 蹇森, 朱剑英. 基于改进的 SIFT 特征的图像双向匹配算法[J]. *机械科学与技术*, 2007, 26(9): 1179-1182.
- [8] 李柏林. 基于特征点图像拼接的配准算法研究[D]. 天津: 天津大学, 2008.
- [9] Kuglin C D, Hines D C. The phase correlation image alignment method[C]//*Proc of international conference on cybernetics society*. San Francisco, USA: [s. n.], 1975: 163-165.
- [10] Fischler M A, Bolles R C. 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.
- [11] 周剑军, 欧阳宁, 张彤, 等. 基于 RANSAC 的图像拼接方法[J]. *计算机工程与设计*, 2009(24): 5692-5694.
- [12] 钟慧湘. 基本矩阵计算方法的研究[D]. 长春: 吉林大学, 2005.
- [13] 单欣, 王耀明, 董建萍. 基于 RANSAC 算法的基本矩阵估计的匹配方法[J]. *上海电机学院学报*, 2006, 9(4): 66-69.
- [14] Burt P J, Adelson E H. A multiresolution spline with application to image mosaics[J]. *ACM Transactions on Graphics*, 1983, 2(4): 217-236.

基于特征不变描述的图像无缝镶嵌技术研究

作者: [杨争争](#), [李良福](#), [YANG Zheng-zheng](#), [LI Liang-fu](#)
作者单位: [陕西师范大学 计算机科学学院, 陕西 西安, 710062](#)
刊名: [计算机技术与发展](#) 
英文刊名: [Computer Technology and Development](#)
年, 卷(期): 2015(1)

引用本文格式: [杨争争](#), [李良福](#), [YANG Zheng-zheng](#), [LI Liang-fu](#) [基于特征不变描述的图像无缝镶嵌技术研究](#)[期刊论文]-[计算机技术与发展](#) 2015(1)