

一种基于小波变换的快速 SIFT 图像拼接算法^{*}

左 毅¹, 陈 勇², 游 华³

(1. 重庆科技学院 信息化处, 重庆 401331; 2. 重庆师范大学 计算机与信息科学学院, 重庆 401331;

3. 重庆师范大学 数学学院, 重庆 401331)

摘要: SIFT 特征匹配算法是当前图像拼接研究领域的热点。为改进 SIFT 算法在图像拼接过程中特征点提取环节计算量大、耗时较长的缺点。本文在研究 SIFT 算法的基础上, 提出一种改进的图像拼接算法。此算法先通过小波变换对图像进行预处理, 减少在 SIFT 尺度不变特征提取过程中的数据计算量, 以提高图像拼接速度。对两幅待拼接图像, 分别采用经典 SIFT 算法和本文提出的改进算法做仿真实验, 当匹配点控制系数 $\alpha=0.5$, 经典 SIFT 算法在确定特征点耗时 1.192 891 s, 改进后的算法仅用了 0.856 712 s。仿真对比试验的结果验证了该算法的有效性, 同时表明该算法能够在保持图像拼接效果的基础上, 提高图像拼接算法的速度和准确性。

关键词: 图像拼接; 小波变换; SIFT

中图分类号: TP311

文献标志码: A

文章编号: 1672-6693(2014)03-0071-05

随着计算机技术和多媒体技术的迅猛发展, 图像拼接技术^[1]受到了研究者和数字图像应用领域的广泛关注, 成为了计算机视觉和数字图像^[2]处理领域的研究热点。由于图像拼接技术具有十分广阔的应用前景和重要的研究价值。Kuglin 和 Hines 在 1975 年提出了相位相关法^[3], 1996 年, Richard Szeliski 提出了一种基于运动的全景图像拼接模型^[4], 1999 年, Lowe David G 提出 SIFT (Scale invariant feature transform) 特征点匹配算法^[5], 并于 2004 年完善总结^[6]。该算法采用待拼接图像关键点的 SIFT 特征向量进行匹配, 对图像的平移、旋转、尺度关系、噪声、光照变化、仿射乃至透射变换都具有较高的鲁棒性。SIFT 特征配准是目前国内外特征点配准领域研究的主要方向。

本文在小波变换^[7]的多尺度分析基础上, 结合 SIFT 特征提取与匹配算法, 提出一种快速高质的图像拼接算法。不仅保留了经典 SIFT 算法匹配能力强, 即使两幅待拼接图像之间存在平移、旋转、仿射变换的复杂情况下, 也能保留稳定匹配的原有特性, 而且通过小波变换的预处理, 减小了搜索范围, 使得算法在计算速度方面有较大提高。

1 SIFT 算法描述

SIFT 算法的实质是在不同的尺度空间上查找关键点, 并计算出关键点的方向。该算法在图像匹配中不受目标的旋转、缩放、平移, 图像仿射、投影变换和光照等因素的影响, 具有较强的匹配能力。该方法包括检测尺度空间极值点、精确定位特征点、确定特征点方向参数和生成特征向量描述等步骤。

1.1 建立尺度空间及特征点定位

为了能够确定稳定的特征点的位置, 需要在所有尺度上对图像进行搜索, 利用高斯微分函数来识别潜在的不随尺度和旋转改变的稳定点。

文献[8]证明了一个图像的尺度空间^[9], $L(x, y, \sigma)$ 定义为一个变化尺度的高斯函数 $G(x, y, \sigma)$ 与原图像 $I(x, y)$ 的卷积。

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

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

其中, $*$ 表示卷积运算, σ 是图像的尺度空间因子, (x, y) 代表图像的像素位置。

将两个不同尺度高斯核的差分同图像进行卷积运算得到图像的 DOG 算子,DOG 算子的公式定义如下:

$$D(x,y,\sigma)=(G(x,y,k\sigma)-G(x,y,\sigma))\ast I(x,y)=L(x,y,k\sigma)-L(x,y,\sigma)$$

将 DOG 尺度空间中的每个像素点(最底层和最顶层除外)跟同一层的相邻 8 个像素点以及它的上一层和下一层的 9 个相邻像素点总共 26 个相邻像素点进行比较,求得 DOG 函数在尺度空间和二维图像空间的极值点,与此同时记下符合条件点的位置和对应尺度。并记为一个局部特征点。

为了获取到稳定的特征点,对上述方法得到的极值点,用 Hessian 矩阵 $\mathbf{H}$ 计算该点的稳定性。对用 DOG 算子得到的 $D(x,y,\sigma)$,它的 Hessian 矩阵为 $\mathbf{H}=\begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$

稳定的特征点须满足如下条件:稳定度 $=\frac{(D_{xx}+D_{yy})^2}{(D_{xx}D_{yy}-D_{xy}^2)}\leqslant\frac{(r+1)^2}{r}$ 。这里, r 是最大特征与最小特征值的比率。一般可取 $r=10$ 。

1.2 稳定的特征点的主方向计算

主方向可通过如下梯度公式计算:

$$m=\sqrt{(L_{x+1,y}-L_{x-1,y})^2+(L_{x,y+1}-L_{x,y-1})^2},\theta=\tan^{-1}\left(\frac{L_{x,y+1}-L_{x,y-1}}{L_{x+1,y}-L_{x-1,y}}\right)$$

其中 L 表示该特征点的尺度,当 m 取最大值时, θ 方向为特征点的主方向。为了确定特征点的主方向,一般以特征点为中心,在它的邻域窗口内用直方图统计邻域像素的梯度方向,一般将 $0^\circ\sim360^\circ$ 按每 10° 计算一个方向梯度,形成 36 个柱。直方图中的峰值对应的角度区间可近似为该特征点的主方向。

1.3 生成 SIFT 特征向量

通过以上步骤,求得每一个稳定特征点都具有的位置、尺度以及方向等 3 个参数。为保持拼接图像时不受到图像旋转的影响,首先将稳定特征点的方向设置为坐标轴方向,然后以特征点为中心建立 4×4 的领域子块,在 4×4 的领域中,对每个像素点定义 8 个方向向量。因此,每个特征点就可得到 $4\times4\times8$ 个方向特征描述,即这 128 个向量能够分别描述 Sift 算法的所有特征点。

1.4 匹配

在得到两幅图像各自的特征点集合和它们对应的 SIFT 特征向量后,可以通过欧式距离法进行特征点的匹配。任取一幅图像中的某个特征点,分别计算这点和另一幅图像中的所有特征点的欧式距离,找出距离最小的前两个特征点,若这两个距离中小小的一个与大的一个之比小于某个给定的阈值时,则可视距离最小的特征点为第一幅图中特征点的匹配特征点。

2 基于小波变换的 SIFT 特征点的提取的匹配算法描述

由于经典 SIFT 特征向量高达 128 维,采用 SIFT 算法匹配图像时,必须计算一幅图像中的特征点与另外一幅图像中的所有特征点的距离,而每一个距离涉及 128 维数据,计算的复杂度是可想而知的,占用了大量系统时间,严重影响算法的实时性。本文提出的改进的算法思想可用图 1 进行描述。

2.1 利用小波变换对图像进行预处理

这里小波变换将图像分解为 4 个子带图像,如图 2 所示:子图像 LL 是低频分量,为原始图像的近似图像;子图像 HL 是水平方向低频、垂直方向高频的分量,表现原图的水平边缘;子图像 LH 是水平方向为高频、垂直方向为低频的分量,表现原图的垂直边缘;子图像 HH 是高频分量,表现原图的斜边缘。

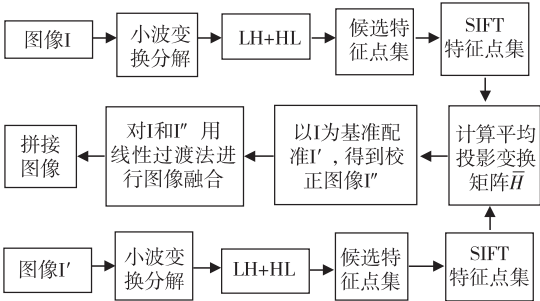


图 1 改进后算法流程图

从上面的分析中和图 3 Lena 图像小波分解示例可以得出,不同细节图像有一定的方向性。所以对待拼接的两幅图像,分别进行小波变换,获取图像的横向和纵向细节,然后提取反映图像横向边缘轮廓特征 HL 和纵向边缘轮廓特征的子带图 LH,并对边缘轮廓特征(LH+HL)进行预处理,剔出灰度值小于阈值 50 的点,余下点均为候选的特征点,分别在各自的候选特征点集中利用 SIFT 算法来提取两幅(LH+HL)图像的 SIFT 特征点。由于 LH 和 LH 叠加(LH+LH)得到边缘轮廓特征图的大小为原图的四分之一,其搜索范围也减小到原来的四分之一。因此减少了后续算法的计算量。计算速度也得到了较大提高。

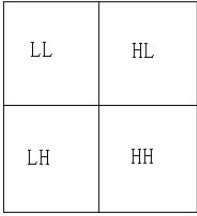


图 2 二维图像小波分解后的 4 个分量



图 3 Lena 图像小波分解示例

2.2 图像配准

将经过小波变换图像预处理以后的图像分别采用传统的 SIFT 算法确定各自的特征点集。然后通过 8×8 参数投影变换模型来描述图像对应特征匹配点间的变换关系。

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & h_1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

这里, $(x, y, 1)^T$ 和 $(x', y', 1)^T$ 分别为两幅图像 I 和 I' 一对匹配点的坐标。依据第一幅图 I, 得到对应点坐标为

$$x' = \frac{h_0 x + h_1 y + h_2}{h_6 x + h_7 y + 1}, y' = \frac{h_3 x + h_4 y + h_5}{h_6 x + h_7 y + 1}$$

这里需要建立 8 个方程来求解投影变换矩阵 8 个未知的参数 $h_i, i=0, 1, 2, \dots, 7$, 即联合可求得 $h_i, i=0, 1, 2, \dots, 7$ 。首先在两幅图像中选取 $k(\geq 4)$ 对特征匹配点, 然后通过求解方程组来确定变换矩阵。利用变换矩阵提供的对应关系, 将一幅图像上的匹配点映射到另一幅图像上, 就能够生成拼接图像。

一般, 两幅图像中的特征对应点还是比较多。可以进一步借助统计方法消除变化矩阵的误差。利用 n 组对应特征点对照, 分别求解方程组, 可得到 n 个变换矩阵 $\mathbf{H}_i, i=1, 2, \dots, n$, 然后利用平均投影变换矩阵可进一步消除匹配的误差。

2.3 图像融合

为了消除图像拼接后因明暗强度的差异形成的拼缝, 缓解拼接后图像的视觉异样感, 让拼接的效果更加无缝平滑。同时减少算法的计算复杂度。本文将采用性能较好的线性过渡法图像融合^[9]技术, 它属于一种渐入渐出式的图像平滑过渡算法。记两幅图像的重叠区域的 x 轴方向的区间为 $[a, b]$, 则重叠区域宽度为 $d=b-a$, 对重叠区域中的某个 $x \in [a, b]$, 可计算过渡因子 $\lambda = \frac{b-x}{d}$, 则拼接图像在重叠区域的灰度值为 $g(x, y) = \lambda g_1(x, y) + (1-\lambda)g_2(x, y)$ 。这里, $g(x, y), g_1(x, y)$ 和 $g_2(x, y)$ 分别为拼接融合后的图像, 原图像 I 和 I' 的在点 (x, y) 处的灰度值。当 λ 从 1 过渡到 0 时, 重叠区域的图像就会从图像 I 逐渐过渡到图像 I' 配准后的图像, 这样在实现图像拼接的同时, 还消除了拼接带来的拼缝, 实现了自然过渡。

3 实验与结论

3.1 实验 1

为了验证改进算法的有效性, 本文选取了在自然光的情况下拍摄的一个大型建筑物的左右两部分带重叠区域的两幅图片对本文提出的图像拼接算法进行验证。建筑物的左侧图像大小为 423×285 (图 3), 建筑物的右侧图像大小为 395×276 (图 4)。图 5 为 $\alpha=0.5$ 的特征点匹配图, 图 6 为拼接后全景效果图。该实验环境为: 测试软件 Matlab 7.1, 运行配置为 Genuine Intel(R) CPU I5-3450 3.50 GHz 和 8GB RAM 平台。



图 3 建筑物左侧图像图



图 4 建筑物右侧图像图

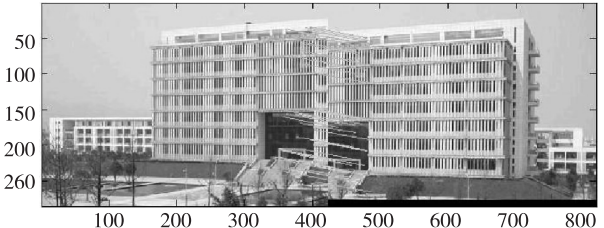


图 5 两幅图片的匹配点对应图

经过实验得到两种算法的实验数据如表 1 所示,其中表 1 中参数 α 为控制匹配点数量的系数,判断匹配的准则是欧氏距离小于 α 倍的下一个最近匹配的距离。

实验数据表明:1)改进后的算法,在 α 取值相同的情况下,与经典的 SIFT 算法相比,在两幅图中分别找到的特征点数,要比原 SIFT 算法有所减少,但最后的两幅图的特征点匹配数是一致的。这说明改进后的算法,提取特征点的精度有了提高。

2)两种算法的特征点匹配数量都受到 α 取值的影响,当 α 值越大,匹配点数就越多,需要的时间也越多。当 α 相同的时候,改进后的算法与传统的 SIFT 算法相比,虽然在小波变换环节消耗了一定的时间,但是在后面的特征点提取的过程中也节约了很多时间,表 1 中,显示了两种算法从图像处理开始到图像提取完成所消耗的时间,结果表明改进后的算法在速度上明显提高。

3)在 SIFT 算法和本文提出的算法中,当 $\alpha=0.2$,匹配点数只有一个,需要时间为 1.061 001 s;当 $\alpha=1$,匹配点数为 950,需要时间为 1.275 743 s;当 $\alpha>1$,匹配点无意义。经过多次实验证明, α 取值在 0.5~0.8 是比较合理的。

表 1 经典 SIFT 算法和本文算法实验数据对照表

α (匹配点控制系数)	经典 SIFT 算法		改进后算法	
	0.5	0.8	0.5	0.8
图书馆左图特征点数量	950	950	936	942
图书馆右图特征点数量	1 140	1 140	1 127	1 132
匹配点数	74	152	74	152
小波变换时间/s	0	0	0.053 702	0.053 702
特征点提取时间/s	1.192 891	1.202 388	0.856 712	0.896 381

3.2 实验 2

为了进一步验证本文算法的优越性,现对本文算法和 SIFT 算法进行图像拼接时的误差进行统计对比,结果如图 7 所示。

由图 7 可知:传统算法进行图像拼接时,测量误差在 50%左右,本文算法建立的模型进行图像拼接时误差在 15%以内,图像拼接的准确性得到很大提高。

3.3 实验 3

为了证明本文算法在速度上的优越性,实验结果更加科学合理。现采用与实验 1 的相同环境,且设匹配点控制系数 $\alpha=0.5$ 。将利用本文算法、经典的 SIFT 算法以及文献[10]中的 Harris-sift 算法,对 100 组不同的图片进行反复实验,计算不同算法的极值点检测时间、生成特征点描述子耗时、匹配时间等,对比结果如表 2 所示,其中消耗时间为计算出 500 个特征点所需的平均时间。

Harris-sift 算法首先利用改进的 Harris 算法提取图像特征点,再使用 SIFT 算法来描述特征点,然后利用欧氏距离对所得的特征向量进行匹配,最终实现图像的自动拼接。虽然 Harris 算法能反映图形的结构。但是该算法需要对图像中的每个点都计算兴趣点,然后在领域内选择最优点,所以该算法使得在检测尺度空间极值点时耗费的时间增加,计算量随着图片像素的增大而增大。Harris 算法并不具有 SIFT 算法应有尺度不变性,所以利用 Harris 算子和 SIFT 提取的特征点对相同的一组图片提取的特征点并不具有一致性,从而导致在使用 Harris 算法提取图像角点的过程中,损失部分 SIFT 特征点信息。本文提出的小波变换预处理能够将图片大量无用信息丢弃保留 SIFT 算法所需要的相关信息。从而保证特征点的准确性和拼接速度上的优越性。

表 2 本文算法与经典 SIFT 算法 Harris-sift 算耗时对比表 ms

	本文改进算法	Sift 算法	Harris-sift 算法
极值点检测耗时	256	203	284
生成特征向量	802	1467	1358
匹配时间	275	311	281
总体时间	1 330	1 981	1 923



图 6 拼接后的建筑物全景图

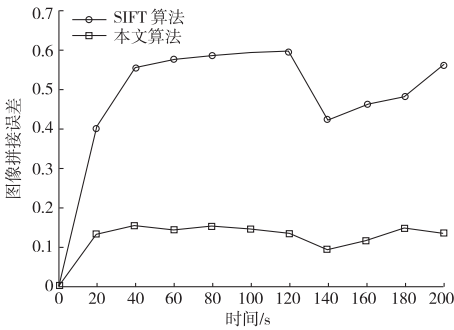


图 7 图像拼接测量误差对比

4 结束语

通过使用本文设计的图像拼接算法,即使在两幅图像之间存在平移、旋转、仿射变换的情况,小波分解也能够有效地丢弃无关的图像信息,保留对拼接有用的特征信息,减少配准时的信息处理量,提高拼接速度。SIFT算法消除了配准时对尺度和方向的依赖性,从而实现稳定的特征点匹配。最后通过渐近式融合,实现了图像的拼接,试验结果验证了改进后算法的有效性和优越性。需要指出的是伪匹配点的排除算法需深入研究,将极大地提高图像拼接中配准环节的可靠性和准确性。

参考文献:

- [1] 席海锋,田超.基于SVR的宽基线图像匹配方[J].重庆邮电大学学报:自然科学版,2013,25(2):197-202.
Xi H F, Tian C. Wide baseline image matching using support vector regression[J]. Journal of Chongqing University of Posts and Telecommunications: Natural Science Edition, 2013, 25(2): 197-202.
- [2] 于海鹏,文政颖.一种基于混沌理论的图像 Hash 算法[J].重庆师范大学学报:自然科学版,2013,30(3):99-102.
Yu H P, Wen Z Y. An image hash algorithm based on chaos theory[J]. Journal of Chongqing Normal University: Natural Science, 2013, 30(3): 99-102.
- [3] Brown L G. A survey of image registration techniques[J]. ACM Computer Surveys, 1992, 24 (4): 325-376.
- [4] Zitova B, Flusser J. Image registration methods: a survey, [J]. Image and Vision Computing, 2003, 21: 997-1000.
- [5] Lowe D G. Object recognition from local scale invariant features[C]//International Conference on Computer Vision. Corfu, Greece, 1999: 1150-1157.
- [6] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91- 110.
- [7] Lindeberg T. Feature detection with automatic scale selection[J]. International Journal of Computer Vision, 1998, 30 (2): 79-116.
- [8] Lindeberg T. Scale2space theory in computer vision[M]. The Kluwer International series in engineering and computer science, dordrecht: Kluwer Academy Publishers, Netherlands, 1994.
- [9] 汪强,尹峰,刘钢钦.基于小波的彩色图像融合技术[J].计算机仿真,2005,22(11):201-204.
Wang Q, Yin F, Liu G Q. The color image fusion algorithm using wavelet transform[J]. Computer Simulation, 2005, 22 (11): 201-204.
- [10] 赵钦君,赵东标,韦虎. Harris-SIFT 算法及其在双目立体视觉中的应用[J].电子科技大学学报,2010,39(4):546-550.
Zhao Q J, Zhao D B, Wei H. Harris-SIFT algorithm and its application in binocular stereo vision[J]. Journal of University of Electronic Science and Technology of China, 2010, 39(4): 546-550.

A Fast SIFT Image Mosaic Algorithm Based on Wavelet Transformation

ZUO Yi¹, CHEN Yong², YOU Hua³

- (1. Division of Information, Chongqing University of Science & Technology, Chongqing 401331;
2. College of Computer and Information, Chongqing Normal University, Chongqing 401331;
3. College of Mathematics Science, Chongqing Normal University, Chongqing 401331, China)

Abstract: Recently, SIFT feature matching algorithm is becoming the focus of the image mosaic. In traditional SIFT algorithm, the mosaic procedure is computationally intensive and time-consuming. For solving this problem, an improved SIFT algorithm is presented in this paper. The proposed algorithm combing wavelet transform into SIFT to simplify the scale-invariant feature extraction process, and speed up the image mosaic. The tests of two images mosaicing with the classical SIFT algorithm and the algorithm proposed in this paper have done respectively. With the control parameter $\alpha = 0.5$, the classical SIFT algorithm took 1.192 891 seconds to extract the feature points, while the improved algorithm took 0.856 712 seconds. The contrast simulation tests demonstrate the validity of the proposed algorithm, and show that the speed and accuracy of the proposed algorithm is improved, at the same time the effect of the image mosaic is maintained.

Key words: image mosaic; wavelet transformation; SIFT