基于 SIFT 特征检测的图像拼接优化算法研究 *

党建武, 宗 岩, 王阳萍
(兰州交通大学 电子与信息工程学院, 兰州 730070)

摘 要: 针对复杂场景下图像拼接,误匹配点比例较大时,传统匹配优化算法效率低,合成图像易产生鬼影等问题,在 SIFT 算法基础上,采用一种新的聚类方法预筛选特征点对,再用 RANSAC 算法精确提纯,减少算法迭代次数;并提出了改进的基于特征点的最佳缝合线与多分辨率样条法相结合的融合方法,提升了融合图像质量。实验结果表明,经过对以上两部分的改进,算法效率有较大提高,并能有效去除鬼影现象。

关键词: 图像拼接; 聚类; 最佳缝合线; 鬼影; 多分辨率融合
中图分类号: TP391.41 **文献标志码**: A **文章编号**: 1001-3695(2012)01-0329-04
doi:10.3969/j.issn.1001-3695.2012.01.091

Research on image mosaic optimization algorithm based on SIFT feature detection

DANG Jian-wu, ZONG Yan, WANG Yang-ping
(*School of Electronics & Information Engineering, Lanzhou Jiaotong University, Lanzhou* 730070, *China*)

Abstract: In complex scenes, the traditional image match optimization algorithm was low efficiency, and motion ghosting was often contained in the merged image. In order to solve these problems, on the basis of the SIFT algorithm, this paper adopted a new method of clustering to pre-screening feature points, then used RANSAC algorithm for purifying feature points accurately. About image fusion, based on feature point, it proposed an improved methods which combined the best suture line and multi-resolution method. Experiments prove that, the improved algorithm which is efficiency, can effectively remove the ghosting.

Key words: image mosaics; clustering; best suture; ghosting; multi-resoltion fusion

图像拼接是将一组相互间存在重叠部分的图像序列进行配准融合,形成一幅包含各序列图像信息的宽视角场景的新图像。通过图像拼接技术,可以扩展图像的分辨率,压缩冗余信息。在虚拟现实领域,拼接生成的全景图可代替场景建模,进而用于虚拟漫游等。

目前,国内外已经提出了多种图像拼接算法,其中尺度不变特征变换[1](scale invariant feature transform,SIFT)匹配算法以其良好的基于尺度、旋转、光照等不变特性被广泛应用于目标识别、图像拼接等领域。SIFT 算法不仅能提取出大量稳定的特征点,而且其独特性较高的特征描述符大多数情况下也能保证其较高的匹配率。但当待匹配图像含有大量的相似结构或包含运动物体时,由于 SIFT 特征描述符仅利用了特征点的局部邻域信息,因此对散落在这些相似区域中的点极易发生误匹配。目前去除误匹配的方法多采用极几何约束、迭代求精,如 M-estimators[2]、RANSAC[3] 等方法,但这些方法受初匹配内点(正确匹配的点)比例影响较大,如何提高内点的比例常常决定匹配图像的配准精度和迭代效率。文献[4]利用欧氏距离通过调整最近邻(NN)与次近邻(SCN)距离的比值阈值,可以减少一些误匹配,但同时也易损失一部分原本正确的匹配点,不能在真正意义上提高正确匹配率。文献[5]提出一种基于中值滤波的特征点对匹配算法,能部分但不能完全剔除错误匹配的特征点对,但该方法执行效率较低。文献[6]提出把中值滤波用于检测 RANSAC 的初始迭代特征点对,并没有考虑排除错误的特征点对,因此对 RANSAC 的执行效率没有实质的改进。图像融合是图像拼接的另一关键部分,目前的问题主要是鬼影现象。效果较好的融合算法有基于图像分割的消鬼影算法[7] 和最佳缝合线算法,但前者复杂度较高。文献[8]首次提出最佳缝合线思想,但其利用复杂度很高的 Dijkstra 算法来查找最佳缝合线。文献[9]提出了一个寻找缝合线的准则,并用动态规划的思想寻找最佳缝合线,降低了计算复杂度。但其仅考虑了颜色差值,实际中不易求出最佳缝合线。文献[10]对上述算法进行了改进,给出了新的缝合线求解准则,加入了对结构差异的考虑,并限制了缝合线的扩展范围,但大量实验表明改进后局部小范围还有可能出现鬼影现象,因为分析可知它得到的只是一局部最优解。本文在上述算法的基础上,针对误匹配造成外点比例高时,使用 RANSAC 算法效率低、配准稳定性差的问题,提出了一种新的聚类方法预筛选待匹配点以提高内点的纯度,再使用 RANSAC 算法精确提纯,以提高配准效率。图像融合部分,在原有算法的基础上提出了基于特征点对的最佳缝合线与多分辨率样条法相结合的融合算法。

1 聚类法预筛选特征点

1.1 算法的基本思路

RANSAC(random sample consensus,随机抽样一致性算法)

收稿日期: 2011-05-05; **修回日期**: 2011-06-23 **基金项目**: 国家自然科学基金资助项目(60962004);甘肃省科技攻关计划资助项目(0708GKCA047);甘肃省自然科学基金资助项目(0803RJZA015)

作者简介:党建武(1963-),男,陕西富平人,教授,博士,主要研究方向为智能信息处理(552408549@qq.com);宗岩(1987-),男,河南南阳人,硕士研究生,主要研究方向为图形图像处理;王阳萍(1973-),女,甘肃定西人,副教授,博士,主要研究方向为医学图像处理.

是一种鲁棒性很强的模板估计方法,其实质是一个反复测试、不断迭代的过程。在图像拼接中,RANSAC 算法的模板是 3×3 变换矩阵(选择 8 参数投影变换,至少需要四个特征点对来生成),而目标函数是计算经变换矩阵变换后的特征点与其匹配特征点之间的误差。算法中,迭代的时间取决于抽样点的数量和线内点所占的比例。可以用式(1)在理论[11]上进行估算:

$$N=\frac{\log(1-p)}{\log[1-(1-\varepsilon)^4]} \tag{1}$$

其中:N 为采样的次数;p 为取样 N 次得到正确样本的概率,一般要求必须大于 95%;ε 为外点的比例。随着 ε 的增加需要随机采样次数的变化情况如表 1 所示。

表 1 随机采样次数与集合中外点比例关系

	集合中外点的比例 ε/%					
	20	30	40	50	60	70
随机采样数 N/次	9	17	34	72	178	567

从表 1 数据可以看出,当线外点的比例为 70% 时,随机采样次数高达 567 次。实际中由于图像噪声等原因,线外点的比例可能会非常高,这样就大大增加了算法迭代次数。为此本研究提出用聚类的方法预筛选数据,减少误匹配,以提高线内点的比例,进而提高算法效率。

预筛选的基本思路是根据拼接图像待匹配点之间的连线斜率应相同或相近的特性,将两幅待匹配图像置于同一坐标系下,然后以某一斜率值为中心,设定一邻域阈值,若在较小邻域阈值内包含最多的斜率相近的特征点对,则这些点就是最精确匹配的相关点对。据此可以排除大部分误差较大的错误匹配。上述较小邻域根据经验值一般可取 0.01 ~0.05。

1.2 算法的主要步骤

SIFT 算法提取特征点后,以一幅图像的特征点为参考基由 BBF(best-bin-first)搜索算法在另一幅图像中查找每个特征点相对应的最近邻与次近邻点,通过判定欧式距离比值得到初匹配,这时每个匹配的特征点对都包含有位置、尺度、方向及 128 维的特征描述符信息,初匹配中含有大量的外点。将两幅待匹配图像变换到同一坐标系中并排列在一起,然后把两幅图像中的匹配特征点用直线连接,则这些直线的斜率应相同或相近(当镜头造成的图片变形被忽略时)。运用斜率聚类的方法,由式(2)(3)求得一个包含最多斜率值的小邻域:

$$M=\max\{\sum_{i=1}^{n}\sum_{j=1}^{n}f(|d_i-d_j|)\} \tag{2}$$

$$f(d)=\begin{cases}1 & d\leqslant t\\ 0 & d>t\end{cases} \tag{3}$$

其中:d_i、d_j 分别代表第 i 和第 j 条直线的斜率;t 为设定邻域的阈值。具体算法分为以下四个部分:

a)计算初匹配数据集中所有直线的斜率值。

b)根据经验值,设定邻域阈值为 $t=0.01$,依据式(2)(3)通过循环计算得到 M 的值,并同时记下此时 i 的值,然后依此计算出此时领域内的点对集,作为预选出的待匹配点。删除在邻域外的点对集。

c)使用 RANSAC 算法对预筛选出的点对进行二次筛选,得到的最大结果集即为匹配的特征点对。

d)使用匹配的特征点对计算变换矩阵,并采用 Levenberg-Marquardt(L-M)算法进行优化。

2 基于特征点的最佳缝合线多分辨率图像融合

2.1 最佳缝合线的确定

2.1.1 最佳缝合线准则

一条理想的缝合线应是在两幅图像重叠区域的差值图像上颜色、结构强度差值最小的一条线,因此定义缝合线求解准则如下:

$$E(x,y)=E_{\text{color}}(x,y)^2+E_{\text{gemetry}}(x,y) \tag{4}$$

式(4)中:$E_{\text{color}}(x,y)$ 表示重叠像素点的颜色值之差,而 $E_{\text{gemetry}}(x,y)$ 表示其结构差值,它是通过修改梯度计算 Sobel 算子实现的。考虑到像素点周围的相似性,梯度计算时,在 x 和 y 方向的梯度分别采用新的模板:$S_x=\begin{pmatrix}-2&0&2\\-1&0&1\\-2&0&2\end{pmatrix}$ 和 $S_y=\begin{pmatrix}-2&-1&2\\0&0&0\\-2&1&2\end{pmatrix}$。假定两幅原始图像为 I_1 和 I_2,则 $E_{\text{gemetry}}(x,y)=\text{diff}(I_1(x,y),I_2(x,y))$,diff 的求解是通过计算两幅图像 I_1 和 I_2 在 x 和 y 方向的梯度差的积得到的。

2.1.2 确定最佳缝合线

根据上述准则,运用动态规划的思想,从重叠区域差值图像的第一行开始,建立以该行上每一个像素为起始点的缝合线,最后从所有的缝合线中寻找一个最优解作为最佳缝合线。

具体步骤如下:

a)初始化。第一行各列像素点为每一条缝合线的起点,各个点处的强度值为按式(4)计算得到的准则值,向下一行扩展。

b)扩展缝合线。已经计算过缝合线强度值的一行向下扩展,直到最后一行为止。扩展的方法是将每一条缝合线的当前点与下一行中该点紧邻的三个像素准则值进行比较,取对应的强度值最小的像素点作为该缝合线的扩展方向,更新此缝合线的强度值,并将缝合线的当前点更新为得到最小强度值所在的下一行中的紧邻像素值所在的列。

c)选择最佳缝合线。从所有的缝合线中选择强度值最小的一条作为最佳缝合线。

图 2 为最佳缝合线算法搜索示意图,黑点表示像素点,实线段代表当前像素点的扩展方向。上述算法有其结构合理性,在一定程度上避开了场景中运动物体,减少了鬼影出现的概率。但仔细分析会发现,它只是一个局部最优解,还有可能出现鬼影现象。因为可能某条缝合线中间有个别误差较大的点,但其整体强度值最小,而这几个特殊的点往往就是造成鬼影现象的客观原因,则这条缝合线就被错误地当做是最佳缝合线。

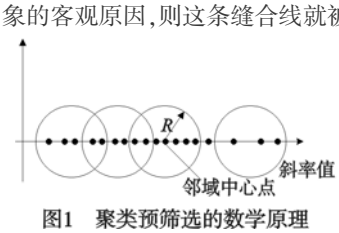

图1 聚类预筛选的数学原理

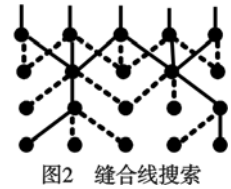

图2 缝合线搜索

2.2 基于特征点的最佳缝合线多分辨率融合

2.2.1 算法思路

为了使缝合线在搜索过程中有效避开两幅图像上的差异

对象或运动物体,避免产生一些误差较大的点,可以适当对缝合线的扩展条件加以限制。由前面对 SIFT 算法的介绍知道,经过匹配、预筛选、提纯之后的匹配特征点对,肯定不会出现在两幅图像中的差异对象上。结合特征点的性质,可以利用特征点对来限制缝合线的扩展方向,并对特征点处的准则值适当加权(权值应大于0 小于 1),这样可加大包含特征点的缝合线被选择的概率。同时,为了能够使缝合线包含尽可能多的特征点,可以适当拓宽缝合线的扩展范围,采用将下一行与该点相邻的 5 个像素点作为缝合线的扩展方向。求出缝合线后,因为两边图像往往还存在着曝光差异,为此对合成图像采用多分辨率样条算法[12]二次融合,以实现平滑自然的无鬼影拼接。

2.2.2　算法的实现步骤

算法的具体实现过程描述如下:

a)初始化。第一行各列像素点为每一条缝合线的起点,其强度值取其准则值,向下扩展。

b)扩展缝合线。已经计算过缝合线强度值的一行向下一列扩展,直到最后一行为止。扩展的方法是如果与当前点相邻的下一行的 5 个像素点内有特征点,则将此特征点作为这条缝合线的扩展方向;如果有多个特征点,则取像素点强度值最小的一个特征点作为缝合线的扩展方向,并将该点的准则值乘以相应的权值(权值 $0<d<1$,根据经验值本文取 $d=0.5$)。如果没有特征点,则取像素点的颜色结构差值最小的点作为缝合线的扩展方向。更新每条缝合线的强度总值。

c)选择最佳缝合线。从所有的缝合线中选取颜色结构差异总值最小的一条缝合线作为最佳缝合线;如果有多个相同的最小值,则选取包含特征点对最多的一条作为最佳缝合线。

d)求出缝合线后,以拼接后的图像尺寸大小,生成一幅区域图像 R_1,R_1 上缝合线的右边用0(黑色)来表示,左边用 255(白色)表示,如图 3 所示。将两幅原始图像 A、B 分别扩展至区域图像大小,扩大了的部分取原图像相应部分填充。

e)分别对步骤 d)中的两幅扩大了的原始图像求解拉普拉斯图像金字塔 L_A、L_B。对区域图像 R_1 求解高斯图像金字塔 GR,为了降低求解过程降采样对图像边缘的影响,引入高斯滤波能减少对缝合线上像素点的影响。

f)生成合成的拉普拉斯金字塔 S,S 的第 L 层按式(5)[12]进行特征融合,N 表示金字塔的总层数。

$$LS_L(i,j)=G_{Rl}(i,j)L_{AL}(i,j)+(1-G_{Rl}(i,j))L_{BL}(i,j)$$
$$0\leqslant L\leqslant N-1 \qquad (5)$$

其中:L_{AL}表示 L_A 的第 L 层,而 L_{BL}表示 L_B 的第 L 层,(i,j)为像素点的位置。

g)将 S 从高层到底层不断执行扩展和叠加操作以完成合成,最终就能得到没有曝光差异的图像。

3　实验结果及分析

根据上述算法思想,用 Visual C++2008 实现了一个图像拼接程序,通过在校园里拍摄多组照片,并对其进行了测试实验。下面是其中一组照片,原始图像用佳能 500D 相机加鱼眼 15 mm 广角镜头旋转拍摄,如图 4 所示(大小已放缩)。其图像重叠部分是一个电子显示屏,下面是滚动的字幕,重叠部分信息不完全一样,并且由于旋转拍摄等原因,两幅图像背景光照强度存在差异。

图3　区域图像R_1

图4　原始图像

未用预筛选前,设定欧式距离比例阈值为 0.6。初匹配结果如图 5 所示,图中含有 106 对匹配点,但包含有大量外点(平行的连线为正确的匹配,交叉的连线为错误的匹配)。

设定聚类法邻域阈值 $t=0.01$,用聚类法预筛选后,得到如图 6 所示结果,图中包含有 36 对匹配特征点对。可以看出,经过预筛选已大大过滤掉了错误的匹配点,有效提高了特征点对内点的比例。虽然预筛选后匹配点的数量有很大的减少,但是得到的匹配点数量已经能够求得图像变换矩阵。改进后算法与原算法一些参数相比,预筛选滤除不匹配点达到了66.57%。

图5　未使用预筛选前匹配特征点对

图6　使用聚类法预筛选后匹配特征点对

在图像融合部分,用经典缝合线算法多分辨率拼接结果如图 7 所示。从拼接结果可以看出,算法一定程度上实现了无缝拼接,但仍存在明显的瑕疵。图 7 下面包含运动物体部分没有出现模糊鬼影现象,说明算法有效避开了差异物体,但图的上面部分却出现了重影现象,说明算法虽然找到了一条最佳路径,但结果是一局部最优解。

改进算法拼接结果如图 8 所示。可以看出,改进后的算法完全消除了重影,平滑了亮度差异,实现了图像的无缝拼接,取得了预期效果。

为了进一步验证本文算法的有效性,选取了 20 组不同场景下拍摄的包含运动物体的照片进行了实验,并统计了其运行结果,如表 2 所示(实验 PC 为 Pentium® Dual-Core CPU E5300 @2.60 GHz 2.59 GHz,2.00 GB 的内存,XP SP2 操作系统,每组照片含两张存在相互重叠区域的图片,重叠区域大概为原图像的 1/3,每张原始图片大小为 400×600)。

表 2　实验结果统计对比

对比项	RANSAC + 传统最佳缝合线算法	聚类法 + RANSAC + 传统最佳缝合线算法	RANSAC + 基于特征点的最佳缝合线算法	聚类法 + RANSAC + 基于特征点的最佳缝合线算法
平均运行时间/s	5.32	4.72	5.36	4.84
平均迭代特征点/对	112	44	114	43
无鬼影合成数/组	12	13	15	16
无鬼影融合率/%	60	65	75	80

表 2 中平均运行时间 = 运行总时间/总组数,无鬼影融合

率 = 无鬼影合成数/总组数。从表 2 数据对比可以看出:a)采用聚类的方法,对初匹配进行提纯,有效减少了 RANSAC 算法初始迭代特征点数,提高了系统运行效率;b)预筛选法可以消除误匹配,尤其是选取在相似区域或运动物体上的伪匹配,在提高算法效率的同时,一定程度上也提升了配准的精度,从而降低了配准鬼影出现的概率,提高了合成图像无鬼影融合率;c)改进后的基于特征点对的最佳缝合线多分辨率融合算法较传统的缝合线融合算法虽然时间效率上相差无几,但无鬼影融合的比率却大大提高,性能优于传统算法;d)整体来看,本文算法不管从时间上还是图像融合质量上,性能较传统算法都有较大提升。

图7　经典最佳缝合线方法拼接结果　图8　改进后算法拼接结果

尽管如此,改进的算法仍不能 100% 保证合成图像不出现鬼影,原因是多方面的。但手动去添加一些特征点,可增加本文融合算法的鲁棒性,提升无鬼影融合率。

4　结束语

本文算法鲁棒性较强,适应于各种复杂场景下包含运动物体的图像的拼接,支持大图片,尤其是旋转拍摄图片的拼接,稍作推广后,也可用于 360° 全景图的拼接,有极大的推广应用前景。文中算法也有不足之处,如计算时间稍长,不能用于实时的图像拼接系统。

参考文献:

[1] BROWN M, LOWE D G. Automatic panoramic image stitching using invariant features [J]. International Journal of Computer Vision, 2007,4(1):59-73.

[2] CHEN Jiun-hung, CHEN Chu-song, Chen Yong-sheng. Fast algorithm for robust template matching with M-estimators [J]. IEEE Trans on Signal Processing, 2003,51(1):230-243.

[3] CHEN Fu-xing, WANG Run-sheng. Fast RANSAC with preview model parameters evaluation [J]. Journal of Software, 2006, 16(8):1431-1437.

[4] KASAR T, RAMAKRISHNA A G. Block-based feature detection and matching for mosaicing of camera-captured document images [C]// Proc of IEEE Region 10 Conference, Symposium on Applications of Holography in Mechanics. New York:ASME, 2007:1-4.

[5] 邹北骥,阮鹏,向遥. 一种精确匹配的全景图自动拼接算法[J]. 计算机工程与科学, 2010, 32(8):60-63.

[6] FANG Xian-yong, ZHANG Ming-min, PAN Zhi-geng, *et al*. A new method of manifold mosaic for large displacement images [J]. Journal of Computer Science and Technology, 2006,21(2):218-223.

[7] AGARWALA A, DONTCHEVA M, AGRAWALA M, *et al*. Interactive digital photomontage [J]. ACM Trans on Graphics, 2004,23(3):294-302.

[8] DAVIS J. Mosaics of scenes with moving objects [C]//Pro of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 1998: 354-360.

[9] EFROS A, FREEMAN W. Image quilting for texture synthesis and transfer[C]// Proc of Annual Conference Series on Computer Graphics Proceedings. 2007:227-239.

[10] 赵亮. 柱面全景图全自动拼接系统研究与实现[D]. 北京:北京化工大学,2009.

[11] 张媛,高冠东,贾克斌. 运用特征点匹配的柱面全景图像快速拼接算法[J]. 中国图象图形学报,2009,3(3):247-250.

[12] 范鹏轩. 虚拟现实系统关键技术和算法的研究及应用[D]. 重庆:重庆交通大学,2010.

(上接第 328 页)

[2] McCALLJOEL C, TRIVEDIMOHAN M. An integrated, robust approach to lane marking detection and lane tracking [C]//Proc of IEEE Intelligent Vehicles Symposium. 2004:533-537.

[3] KIM Z W. Robust lane detection and tracking in challenging scenarios [J]. IEEE Trans on Intelligent Transportation Systems, 2008,9(1):16-26.

[4] DANESU R, NEDVSCHI S. Probabilistic lane tracking in dificult road scenarios using stereovision [J]. IEEE Trans on Intelligent Transportation Systems,2009,10(2):272-282.

[5] SATZODA R K, SATHYANARYNA S, SRIKANTHAN T. Hierarchical additive Hough transform for lane detection [J]. IEEE Embedded Systems Letters,2010,2(2):23-26.

[6] KREUCHER C , LAKSHMANAN S. LANA:a lane extraction algorithm that uses frequency domain features [J]. IEEE Trans on Robotics and Automation,1999,15(2):343-349.

[7] DAGAN E, MANO O, STEIN P, *et al*. Forward collision warning with a single camera [C]//Proc of IEEE Intelligent Vehicles Symposium. 2004:37-42.

[8] WANG Rong-ben, YU Tian-hong, JIN Li-sheng, *et al*. Edge extraction method study based on maximum entropy for linear lane identifying and tracking[C]//Proc of IEEE Intelligent Vehicles Symposium. 2005:849-854.

[9] 张荣,王勇,杨榕. TM 图像中道路目标识别方法的研究[J]. 遥感学报,2005,9(2):220-224.

[10] 郭磊,李克强,王建强,等. 应用方向可调滤波器的车道线识别方法[J]. 机械工程学报,2008,44(8):214-218.

[11] JUNG C R, KELBER C R. A robust linear-parabolic model for lane following[C] //Proc of the 17th Brazilian Symposium on Computer Graphics and Image Processing. Washington DC:IEEE Computer Society,2004:72-79.

[12] GORMAN F O, CLOWES M. Finding picture edges through collinearity of feature points[J]. IEEE Trans on Computers,1976,25(4):449-456.

[13] 贾阳,王荣本,余天洪,等. 基于熵最大化边缘提取的直线型车道标识线识别与跟踪方法[J]. 吉林大学学报:工学版,2005,35(4):420-425.