图像拼接技术综述

裴红星， 刘金达， 葛佳隆， 张　斌

(郑州大学 物理工程学院　河南 郑州 450001)

摘要: 图像拼接技术是图像处理技术中的研究热点.对图像拼接技术进行分析与总结,基于图像拼接技术的发展历史与现状,介绍了图像拼接技术的背景与应用;概括了图像拼接的基本概念与拼接流程,主要研究了图像拼接中的图像配准技术,对图像配准中的经典算法进行分类概述,并研究了各算法的基本思想与优缺点,对其中的尺度不变特征变换算法进行了重点介绍,概述了图像融合的基本类别,详细介绍了图像融合中最常用的加权平均法,最后指出了图像拼接中的当前问题与发展趋势.

关键词: 图像拼接; 图像配准; 图像融合; SIFT

中图分类号: TP391　　**文献标志码:** A　　**文章编号:** 1671-6841(2019)04-0001-10

DOI: 10.13705/j.issn.1671-6841.2019004

0 引言

普通相机拍摄图像时,无法兼顾相机视场与视场中单个物体的分辨率问题,而全景相机普遍价格昂贵,不适用于低成本的一般性场景.为了使用普通相机获取宽视角,甚至是 360°全景图像,人们提出了图像拼接技术.图像拼接技术是指将含有重叠部分的两幅或多幅图像,通过图像预处理、图像配准和图像融合技术,拼接成一幅包含各图像信息的高分辨率、宽视角图像的技术.该技术广泛地应用在医学成像[1]、遥感技术[2]、虚拟现实[3]、视频编辑[4]等方面,随着计算机视觉等技术的发展,其重要性日益凸显.此外,图像拼接技术还用于视频压缩[5],将视频逐帧拼接为高分辨率图像,去除重复部分,降低存储空间,提高传输速度,采用相位相关法搜索给定目标,实现图像索引[6].

图像拼接技术最先应用在遥感技术领域,其核心是图像配准,产生了两个分支,一个分支是基于区域的图像配准.早期主要采用全局搜索法进行图像配准,计算量极大. 模板匹配法[7]相对于全局搜索法操作简单,容易实现,但都只能解决平移问题.Reddy 于 1996 年提出扩展相位相关法[8],将相位相关法[9-10]和对数极坐标变换相结合,解决旋转和尺度缩放问题,但该方法计算较为复杂.另一个分支是基于特征的图像配准,包括角点、边缘、斑点等.1988 年,Harris 提出了经典的 Harris 角点检测算法[11],将图像中具有旋转、平移不变性的角点作为特征点进行图像配准.Lowe 于 1999 年[12]提出,并在 2004 年[13]完善尺度不变特征变换(scale-invariant feature transform,SIFT)算法,该算法对平移、旋转、尺度缩放和光照不均都有很好的适应性,是目前的研究热点.

1 图像拼接流程

图像拼接包括 3 部分:图像预处理、图像配准和图像融合.

图像预处理的目的是提高配准精度、降低配准难度,包括调整灰度差异、去噪、几何修正以及将两幅图像投影到同一坐标系等基本操作.

图像配准是计算出两幅图像间的空间变换模型并进行空间变换,使两幅图像的重叠部分在空间上对准,是图像拼接的关键.图像之间的空间变换关系包括:平移、旋转、尺度缩放、仿射变换、投影变换,其中投影变

收稿日期:2019-01-03
基金项目:国家自然科学基金项目(81171410).
作者简介:裴红星(1975—),男,河南安阳人,副教授,主要从事机器视觉、图像处理方向研究,E-mail:phxlee@zzu.edu.cn;通信作者:张斌(1967—),男,河南唐河人,副教授,主要从事机器视觉、图像处理、医学图像处理方向研究,E-mail:zb1967@zzu.edu.cn.

换更具有普遍性.

假设图像 $f_1(x,y)$、$f_2(x,y)$ 存在投影变换关系,则用齐次方程(1)表示:

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \boldsymbol{M} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}, \tag{1}$$

其中:m_0、m_1、m_3 和 m_4 共同表示旋转角度和缩放尺度;m_2 和 m_5 分别表示 x 方向与 y 方向上的平移量;m_6 和 m_7 分别表示 x 方向和 y 方向上的变形量.图像配准的关键是用式(1)确定空间变换模型 $\boldsymbol{M}$ 的参数.

图像融合的目的是得到无缝的高质量图像.在不损失原始图像信息的前提下,消除接缝与亮度差异,实现拼接边界的平滑过渡.

2 图像配准技术

2.1 基于区域的图像配准

基于区域的图像配准分为基于空间域的图像配准和基于变换域的图像配准.

2.1.1 基于空间域的图像配准　基于空间域的图像配准法有全局搜索法和模板匹配法[7].直接利用图像像素的灰度信息,采用某种搜索方法,如以单个像素为单位平移的方法,确定当两图像重叠部分之间的相似性测度达到最大时的变换模型,常见的相似性测度有平均绝对差(mean absolute deviation,MAD)、误差平方和(sum of squared difference,SSD)、归一化互相关 (normalized cross correlation,NCC)等.

全局搜索法是使待配准图像在参考图像上平移,比较平移过程中两幅图像的配准程度,在进行所有平移之后选择相似性测度最优的平移量,这种方法的缺点是计算量极大.模板匹配法以两幅图像的重叠部分为基准,选取待配准图像的一块矩形区域为模板,在参考图像中搜索相似区域,相似性测度最优的位置即为最佳配准位置,如图 1 所示.

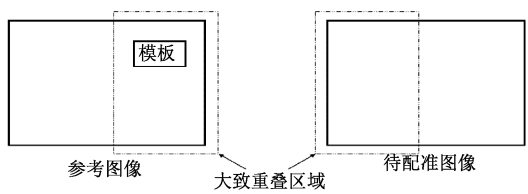

图 1 模板匹配法

Fig.1 Template matching method

模板匹配法操作简单、匹配精度高,在图像拼接[14-15]、目标识别[16-17]等方面得到广泛应用.但全局搜索法和模板匹配法都要求图像不能包含较大噪声,不能有旋转、尺度缩放等问题,计算量仍然较大,并且要求参考图像中有完整的模板匹配目标.

2.1.2 基于变换域的图像配准

1) 相位相关法

根据傅立叶变换的平移不变性,相位相关法[9-10]将空间域上像素的平移转换为频率域上相位的平移.

设 $f_1(x,y)$ 和 $f_2(x,y)$ 是具有平移关系的两幅图像,其空间变换关系为

$$f_2(x,y) = f_1(x - x_0, y - y_0), \tag{2}$$

其中:(x_0,y_0) 为两幅图像间的平移量. 则图像 f_1 与 f_2 对应的傅立叶变换 F_1 和 F_2 之间关系为

$$F_2(\mu,\nu) = \mathrm{e}^{-\mathrm{j}(\mu x_0 - \nu y_0)} F_1(\mu,\nu). \tag{3}$$

由式(3)得出,图像 $f_1(x,y)$ 和 $f_2(x,y)$ 在频率域中幅值相同,相位不同.两幅图像间的互功率谱定义为

$$\frac{F_1^*(\mu,\nu) F_2(\mu,\nu)}{|F_1^*(\mu,\nu) F_2(\mu,\nu)|} = \mathrm{e}^{-\mathrm{j}(\mu x_0 - \nu y_0)}, \tag{4}$$

其中:* 是复共轭运算符.

对式(4)右侧进行傅里叶逆变换得冲击响应函数 $\delta(x-x_0, y-y_0)$,搜索使冲击响应函数 δ 最大的点(x_0, y_0),则(x_0,y_0)即为图像 $f_1(x,y)$ 和 $f_2(x,y)$ 之间的最佳平移量.

快速傅里叶变换(fast fourier transform,FFT) [18]提升了基于变换域的相位相关法的算法性能,效率高、计算简单,相对于基于空间域的图像配准,该方法对光照变化不敏感、抗噪能力强、稳定性好. 该方法只对有平移关系的两幅图像有良好的效果,不适用于有旋转和尺度缩放关系的图像.二维快速傅里叶变换的时间复

杂度可以用 $O(M\cdot N\cdot \log(M\cdot N))$ 表示,其中 M、N 为图像尺寸,所以其时间复杂度随图像尺寸增大呈指数增长.

2) 扩展相位相关法

Castro 等[19]解决了相位相关法不适用于平移和旋转的问题.Reddy[8]又在相位相关法的基础上提出了将相位相关法和对数极坐标变换相结合的扩展相位相关法,通过对数极坐标变换将两幅图像间的旋转和尺度缩放关系转换为该坐标系下的平移关系.

设图像 $f_1(x,y)$ 和 $f_2(x,y)$ 间存在平移、旋转和尺度缩放关系,则其空间变换关系为

$$f_2(x,y)=f_1[s(x\cos\theta_0+y\sin\theta_0)-x_0,s(-x\sin\theta_0+y\cos\theta_0)-y_0],\tag{5}$$

其中:θ_0 为旋转角度;s 为尺度因子;(x_0,y_0) 为平移量.

则图像 f_1 与 f_2 对应的傅立叶变换 F_1 和 F_2 之间关系如式(6),

$$F_2(u,v)=\mathrm{e}^{-\mathrm{j}2\pi(ux_0+vy_0)}\,s^{-2}\,F_1[s^{-1}(u\cos\theta_0+v\sin\theta_0),s^{-1}(-u\sin\theta_0+v\cos\theta_0)].\tag{6}$$

取 F_1、F_2 的幅度谱 M_1、M_2,则 M_1 与 M_2 关系如式(7),

$$M_2(u,v)=s^{-2}\,M_1[s^{-1}(u\cos\theta_0+v\sin\theta_0),s^{-1}(-u\sin\theta_0+v\cos\theta_0)].\tag{7}$$

忽略比例因子 s^{-2} 的影响,对 M_1 与 M_2 进行极坐标变换得

$$M_2(r,\theta)=s^{-2}\,M_1(s^{-1}r,\theta-\theta_0).\tag{8}$$

沿极坐标半径 r 方向取对数可得

$$M_2(\log r,\theta)=s^{-2}\,M_1(\log r-\log s,\theta-\theta_0).\tag{9}$$

设 $\xi=\log r$,$d=\log s$ 可得

$$M_2(\xi,\theta)=s^{-2}M_1(\xi-d,\theta-\theta_0).\tag{10}$$

对式(10)用相位相关法可以计算出旋转角度 θ_0 和尺度因子 s,对待配准图像进行旋转和尺度缩放后,两幅图像间只剩平移关系,再次使用相位相关法可以确定两幅图像之间的平移量(x_0,y_0).

扩展相位相关法具有相位相关法的效率高、对光照变化不敏感、抗噪能力强、稳定性好等优点,同时又解决了旋转和尺度缩放问题,较相位相关法适用范围更广.该方法还可以与边缘检测相结合[20-22],进一步提高鲁棒性和计算效率.但是算法复杂,并且要求两幅图像之间有较高的重合度.

2.2　基于特征的图像配准

与基于区域的图像配准相比,基于特征的图像配准只使用图像的部分信息,如轮廓、角点等特征,主要有角点检测法、基于轮廓特征的图像配准、基于 SIFT 的图像配准等.

2.2.1　*角点检测法*　角点是指灰度图像中局部灰度梯度变化较大的点,或轮廓线上局部范围内的曲率极大值点.角点检测法基本思想是提取特征点,匹配两幅图像中的特征点作为特征点对,再通过随机抽样一致性(random sample consensus, RANSAC)算法[23],根据式(1)使用特征点对估计空间变换模型的参数.

1) 角点提取常用 Harris 角点检测法[11]、加速分割测试特征(features from accelerated segment test, FAST)算法[24-25]、最小核值相似区(smallest univalue segment assimilating nucleus, SUSAN)算法[26-27]等提取图像中的角点.

a) Harris 角点检测法基本思想是计算某像素邻域内各像素的 x 和 y 方向上的梯度 I_x、I_y,用角点响应函数(corner response function, *CRF*)表示梯度 I_x、I_y 的特征,如果某像素邻域内 I_x、I_y 的幅值都较大,*CRF* 取局部区域极大值,则为特征点,如图 2(a)所示;如果只有一个方向的梯度幅值较大,另一方向的梯度几乎为 0,则为边缘点,不是特征点,如图 2(b)所示;如果 I_x、I_y 的幅值几乎都为 0,则位于平坦区域,不是特征点,如图 2(c)所示.

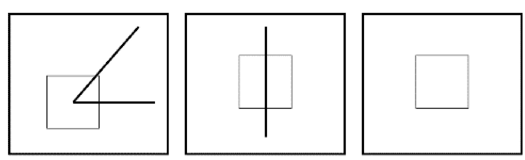
(a)　(b)　(c)

图 2　不同类型的点

Fig.2　Types of different points

Harris 算法使用图像的梯度特征检测角点,具有光照不变性,还可以解决平移、旋转问题,但是像素邻域的窗函数对尺度变换敏感,不具备尺度不变性;Harris 需要人工指定阈值,适应性较差,但可以与自适应阈值算法相结合[28]提高算法效率.

b) FAST算法基本思想是如果某像素的灰度值与足够多的邻域像素灰度值的差值较大时,认为该点为特征点.

以点 $P(x,y)$ 为圆心,以3为半径作圆,如图3所示,则有16个像素 $p_1,p_2,\cdots,p_{16}$ 分布在圆上,计算 $p_1,p_2,\cdots,p_{16}$ 与中心点 P 的灰度值差值,如果有超过8个像素灰度值的差值的绝对值大于给定阈值 s,则点 P 为潜在特征点.

为了去除干扰点,需要剔除潜在特征点中的非极大值点.若潜在特征点 P 的邻域 N 内,如5×5邻域,有多个潜在特征点,则用式(11)计算邻域 N 内每个潜在特征点的 $score$ 值,若点 P 是邻域 N 内 $score$ 值最大的潜在特征点,则保留点 P,否则,剔除点 P.

$$score = \max\begin{cases}\sum |f(i,j) - f_p(x,y)| & f(i,j) - f_p(x,y) > t \\ \sum |f_p(x,y) - f(i,j)| & f_p(x,y) - f(i,j) > t\end{cases}, \tag{11}$$

其中:$f(i,j)$ 是潜在特征点的邻域像素灰度值;$f_p(x,y)$ 是潜在特征点的灰度值;t 为给定阈值.

FAST算法操作简单,实时性好,但对包含噪声的图像鲁棒性不好,检测结果受给定阈值 s 和 t 的影响较大,不具备尺度不变性和旋转不变性.研究表明,适当增加图像对比度并滤波去噪能显著提升FAST的稳定性[29].

c) SUSAN算法与FAST算法基本思想类似,如果某像素灰度值与足够多的邻域像素的差值较大,认为该点为特征点,并可以根据邻域特征剔除干扰点.

以点 $P(x,y)$ 为圆心,以 r 为半径的圆形区域作为点 P 的邻域,邻域内共有 R 个像素.若邻域内有 n 个像素的灰度值与中心像素 P 灰度值之差小于给定阈值,则认为这 n 个像素与中心像素 P 灰度值相同,n 个像素共同构成核值相似区(univalue segment assimilating nucleus, USAN).

当中心像素 P 位于目标或背景中时,如图4(a)、(b)所示,n 与 R 大致相等;当中心像素 P 位于边缘上时,如图4(c)所示,n 约为 R 的1/2;当中心像素 P 位于角点上时,如图4(d)所示,n 远小于 R;当中心像素 P 位于孤立点上时,如图4(e)所示,n 约为0.

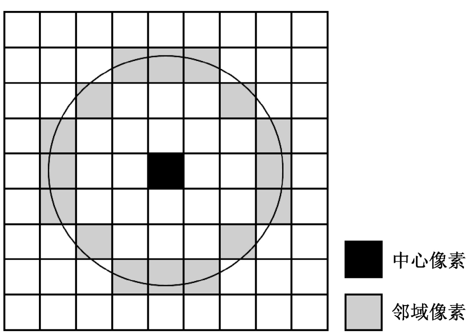

图3 FAST算法中点的邻域

Fig.3 Neighborhood of a feature point in FAST

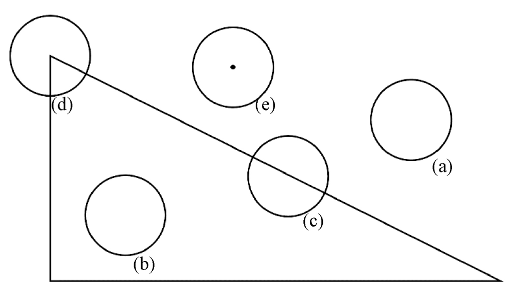

图4 SUSAN算法中点的邻域

Fig.4 Neighborhood of a feature point in SUSAN

然后根据USAN区域的特征去除干扰点.若USAN区域的质心与模板中心 P 的距离小于给定阈值,则剔除;若存在USAN区域的质心与模板中心 P 的连线上的像素不在USAN区域内,则剔除.

SUSAN算法与FAST算法类似,操作简单,实时性好.该算法还根据USAN区域中像素数量 n 与邻域内像素总数量 R 的相对大小实现特征点检测[30]或边缘检测[31],功能性强,可以解决旋转和一定程度的尺度缩放问题,不涉及微分等操作,不易受单个噪声点的影响,鲁棒性好.但是该算法需要调节多个阈值,算法灵活但自适应性较差,对图像亮度敏感,不具备光照不变性.文献[32]提出,SUSAN与高斯-拉普拉斯算子(Laplace of Gaussian, LoG)相结合可以显著改善SUSAN算法对噪声和光照的敏感性.

2) 角点配对

设图像 $f_1(x,y)$ 和 $f_2(x,y)$ 中角点的集合分别为 P 和 Q:$P=\{p_1,p_2,\cdots,p_A\}$;$Q=\{q_1,q_2,\cdots,q_B\}$.

a) 邻域特征法

用角点的邻域特征构成的特征向量描述角点,如邻域灰度平均值 $\bar{f}$、邻域灰度值标准差 σ、熵 e 等.

$$\boldsymbol{p}_a=[\bar{f}_{pa},\sigma_{pa},e_{pa},\cdots],\ a=1,2,\cdots,A;\boldsymbol{q}_b=[\bar{f}_{qb},\sigma_{qb},e_{qb},\cdots],\ b=1,2,\cdots,B.$$

计算角点 p_1 的特征向量 $\boldsymbol{p}_a$ 与 $\boldsymbol{q}_b$ 间的欧式距离,距离最近且小于给定阈值的角点与 p_1 组成特征点对,否则 p_1 无配对特征点.在 Q 中依次为 P 中的其余角点配对.

b) 相关系数法

取各角点的 $(2K+1)\times(2K+1)$ 邻域,其中:$K=1,2,\cdots$;R_b 为角点 p_1 与 Q 中各角点之间的相关系数,则有

$$R_b=\frac{\sum_{i=-N}^{K}\sum_{j=-N}^{K}(f_1(x_1-i,y_1-j)-\bar{f}_1)(f_2(x_b-i,y_b-j)-\bar{f}_2)}{\sqrt{\sum_{i=-N}^{K}\sum_{j=-K}^{K}(f_1(x_1-i,y_1-j)-\bar{f}_1)^2\sum_{i=-N}^{K}\sum_{j=-K}^{K}(f_2(x_b-i,y_b-j)-\bar{f}_2)^2}},\tag{12}$$

其中:$b=1,2,\cdots,B$;f_1 和 f_2 是两幅图像的灰度值;$\bar{f}_1$ 和 $\bar{f}_2$ 是两幅图像中各像素点的灰度均值.R_m 最大且大于给定阈值的角点与 p_1 组成特征点对,否则角点 p_1 无配对特征点.在 Q 中依次为 P 中的其余角点配对.

角点检测法在提取图像特征时,有效降低了信息量,明显降低了时间复杂度,其准确性和实时性较好,还可以解决部分目标物被遮挡、几何变形等问题.但该类算法在特征提取过程中会损失大量的图像信息,对特征不明显、噪声较多的图像适应性较差.研究表明,与去噪算法、图像增强算法相结合可以提升该类算法性能[33-34].

2.2.2　基于轮廓特征的图像配准　基于轮廓特征的图像配准的基本思想是根据轮廓特征对所提取的轮廓进行匹配,再将轮廓匹配对中的某些点作为特征点对,并根据式(1)用特征点对估计空间变换模型.

Li 等[35]提出的轮廓匹配算法,能够处理包含有较大噪声但具有清晰边缘的图像,且两图像中相同目标之间有形变的情况,具有很好的鲁棒性和可靠性,可以处理图像间平移、旋转和尺度缩放的问题,对目标形变也有较好的适应性.

1) 轮廓提取

通常采用二值化、卷积或形态学处理等方法提取图像中的目标轮廓,Li 用高斯-拉普拉斯(LoG)算子对图像进行卷积运算来提取图像轮廓.拉普拉斯算子是二阶微分算子,卷积后图像中连续过零点即为轮廓线;为了克服拉普拉斯算子对噪声与离散点的增强作用,在处理前使用高斯函数对原图像进行降噪处理.提取的轮廓使用链码[36]编码描述,并对链码进行归一化处理,对归一化链码进行平滑滤波,使链码表示的轮廓更加平滑.

2) 轮廓匹配

通过轮廓的形状特征,包括周长 c、轮廓到质心的最大与最小距离 $d_{\max}$ 和 $d_{\min}$、Hu 不变矩[37]中的一阶不变矩 h_1 和二阶不变矩 h_2,构成的特征向量间的欧氏距离对两幅图像上的轮廓配对,构成轮廓匹配对.

3) 将轮廓匹配对的质心作为特征点对,根据式(1)估计空间变换模型.

基于轮廓特征的图像配准的关键是准确地提取轮廓特征,由于轮廓特征包含信息量较少,所以计算简单、效率高,相对于角点检测法,不易受单个噪声点的影响,鲁棒性较好,但是该类算法的配准精度依赖于轮廓提取精度,所以要求图像具有明显的轮廓特征.

2.2.3　基于 SIFT 的配准算法　SIFT 算法[12]能提取具有尺度不变、旋转不变、光照不变的局部极值点作为特征点.其基本思想是使用降采样和高斯函数构建图像尺度空间,提取极值点作为潜在特征点并剔除干扰点,将特征点邻域像素的梯度作为特征向量来描述特征点,计算特征向量间的欧氏距离进行特征点配对,根据公式(1)估计空间变换模型的参数.

1) 构建尺度空间

如图 5 所示,用高斯函数 $G(x,y,\sigma)$ 对原图像进行平滑处理,构建图像尺度空间 $L(x,y,\sigma)$,

$$L(x,y,\sigma)=I(x,y)G(x,y,\sigma),\tag{13}$$

其中:$\sigma(t,r)=\sigma_0 2^{t+\frac{r}{s}}$,$\sigma_0$ 是高斯正态分布的标准差初始值,建议值为 1.6;$t\in[0,1,\cdots,T-1]$;T 是高斯图像金字塔的总级数,且 $T=\log[\min(X,Y)]-2$;X、Y 为原图像尺寸;$r\in[1,2,\cdots,s-1]$,s 为每级金字塔的尺度空间层数,且 $s>3$.

T 级金字塔的第 $t(t>0)$ 级第 1 张图像由第 $t-1$ 级中第 $s-2$ 层的图像降采样得到.最后构建出级数为

T,每级包含 s 层图像的高斯图像金字塔.

在 T 级高斯图像金字塔的每一级中,相邻尺度图像两两相减得到高斯差分(difference of Gaussian,DoG)图像金字塔,即 $D(x,y,\sigma)=\{f_1-f_0,f_2-f_1,\cdots,f_s-f_{s-1}\}$,其中 $f_i,i\in[0,s]$ 是尺度空间图像.最后构建出级数为 T、每级包含 $s-1$ 层图像的 DoG 图像金字塔.

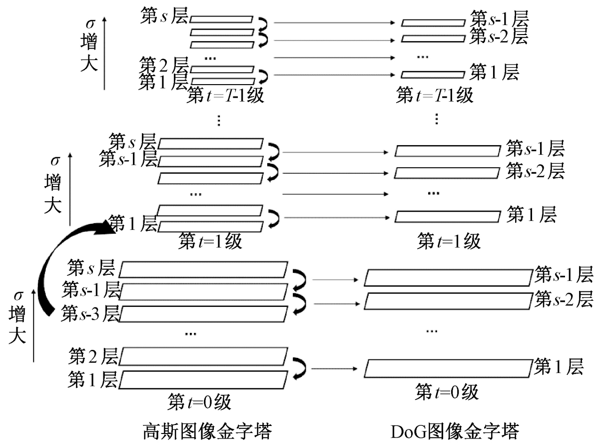

图 5 构建 DoG 图像金字塔的过程

Fig.5 The process of producing the DoG images

2)特征点提取

在 DoG 图像金字塔中,任意像素都可以用 (t,r,x,y) 表示,其中 t 为该像素所在 DoG 图像金字塔的级数,r 为该像素所在第 t 级 DoG 图像金字塔的尺度空间层数,x 和 y 为像素的坐标.当某像素 $d(t,r,x,y)$ 的灰度值为其尺度空间 r 内 8 邻域内的极值,且为第 $r-1$ 和 $r+1$ 层尺度空间图像相同坐标的 18 个像素的灰度值极值时,该点为其尺度空间中 3×3×3 邻域内的潜在特征点,如图 6 所示.

将位于曲率较低的边缘上的潜在特征点剔除,如图 2(b)所示,并通过拟合特征点所在行和列的像素灰度值曲线,确定特征点的实际位置,最终得到特征点,如图 7 所示.

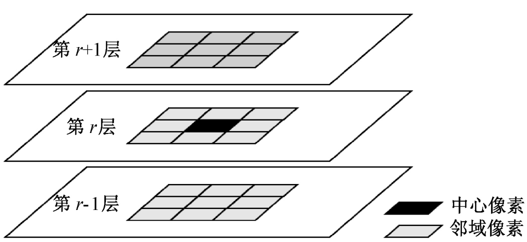

图 6 SIFT 特征点检测

Fig.6 The detection of feature points in SIFT

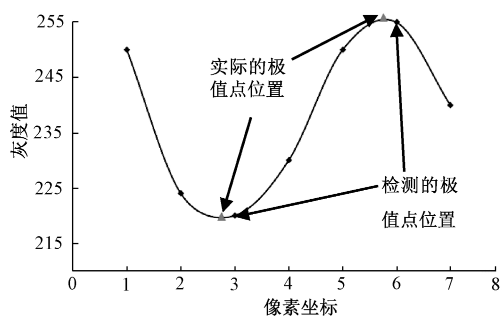

图 7 SIFT 特征点检测

Fig.7 The accurate location of feature points in SIFT

3)特征点描述

利用邻域像素的梯度特征为特征点指定主方向,具体如下.

确定以特征点 $p(t,r,x,y)$ 为圆心、$3\times1.5\sigma$ 为半径的邻域,计算各邻域像素梯度的幅值 M_f 和角度 δ,

$$M_f(x,y)=\sqrt{(L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1))^2},\tag{14}$$

$$\delta(x,y)=\arctan\left(\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)}\right).\tag{15}$$

以 45°的间隔将 360°分为 8 个角度区间:

$$D_\theta=(\theta-22.5°,\theta+22.5°],\ \theta=k\cdot45°,k=0,1,2,\cdots,7.\tag{16}$$

计算特征点 $p(t,r,x,y)$ 邻域内各像素梯度角度在 8 个角度区间内的分布情况,将各角度区间包含的像素的梯度幅值相加,幅值之和最高的角度区间 D_θ 即为特征点的主方向 θ,如图 8 中的 270°.

为了确保特征点的旋转不变性,将特征点的邻域旋转 θ; 取特征点的 16×16 邻域,如图 9,计算该邻域内各点梯度的幅值与角度,并使用高斯窗口对各梯度的幅值与角度高斯加权,再将 16×16 邻域划分为 16 个 4×4 子邻域,在各 4×4 子邻域内统计 8 个角度区间内的梯度幅值之和作为子邻域的种子点;16 个种子点各有 8 个方向,共构成 128 维的特征向量,最后对 128 维特征向量归一化,去除光照影响,保证光照不变性.

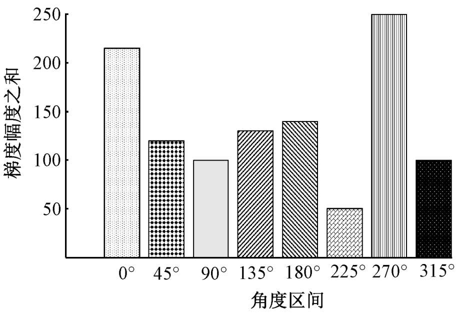

图 8　8 个 45°的角度区间与梯度幅度直方图

Fig.8　Orientation intervals and the gradient magnitude histogram

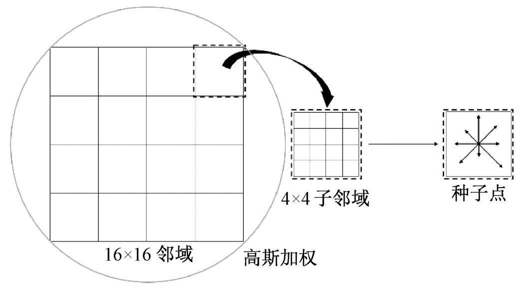

图 9　SIFT 特征点描述

Fig.9　The feature description in SIFT

4) 根据两幅图像特征点的特征向量之间的欧氏距离对特征点配对,并根据式(1)用特征点对估计空间变换模型的参数.

SIFT 算法适用于平移、旋转、尺度缩放、光照不一等情况,适应性强,并且对视角变化、仿射变换也保持一定程度的稳定性;少量目标就可产生大量特征点,特征点的特征向量信息丰富,匹配精度高.该算法可以与其他形式的特征向量相结合,具有一定程度的可扩展性,是图像配准领域中的研究热点,其衍生算法加速稳健特征[38-39]基本满足了实时性要求.但 SIFT 算法复杂,构建图像金字塔、确定特征点的准确位置、估计主方向等操作可能会带来累积误差,造成特征点稳定性较差;对纹理较少、较为平滑的图像,128 维特征向量可能没有显著差异,造成特征点匹配精度低.

3　图像融合技术

由于图像亮度差异等因素,配准后图像可能存在明显接缝,需要使用图像融合技术来消除拼接缝隙.根据图像的表征层可将图像融合分为:像素级融合、特征级融合和决策级融合[40-41].像素级融合是目前常用的融合方法,指对像素灰度值进行直接平均、加权平均等操作,计算量较大.特征级融合是首先对图像提取轮廓、纹理、角等特征,再对这些特征分析、处理与整合,得到融合后的特征,速度较快.决策级融合是在特征级融合的基础上,对特征信息进行识别和分类,再基于认知模型和判别系统对决策的可信度量进行判断,融合效果较好,速度快.但是特征级融合和决策级融合的算法复杂,融合过程中丢失较多的原始图像信息,并且可能造成融合后图像的失真,应用较少.

像素级融合中最简单的方法是直接平均法,直接对配准后图像重叠区域的灰度值取平均值,避免了灰度值直接相加造成的溢出.设 $f(x,y)$ 是 $f_1(x,y)$、$f_2(x,y)$ 融合后图像,则直接平均法可以表示为

$$f(x,y)=\begin{cases} f_1(x,y), & (x,y)\in f_1 \\ [f_1(x,y)+f_2(x,y)]/2, & (x,y)\in(f_1\cap f_2). \\ f_2(x,y), & (x,y)\in f_2 \end{cases} \tag{17}$$

该算法简单易行,但融合质量较差,融合后图像仍然可能存在明显拼接缝隙.

加权平均法是对配准后两幅图像重叠区域的灰度值按一定规则加权后再计算平均值.设 $f(x,y)$ 是 $f_1(x,y)$、$f_2(x,y)$ 融合后图像,则加权平均法表示为

$$f(x,y)=\begin{cases} f_1(x,y), & (x,y)\in f_1 \\ w_1 f_1(x,y)+w_2 f_2(x,y), & (x,y)\in(f_1\cap f_2). \\ f_2(x,y), & (x,y)\in f_2 \end{cases} \tag{18}$$

w_1、w_2 分别为图像 f_1、f_2 对应像素灰度值相加时的权值,且 $w_1 + w_2 = 1, 0 \leqslant w_1, w_2 \leqslant 1$, 常使用渐入渐出法对权值 w_1、w_2 进行动态选取, $w_1 = x_2 - x_i / x_2 - x_1, w_2 = 1 - w_1 = x_i - x_1 / x_2 - x_1$, 其中: x_1 为待融合部分的左边界; x_2 为待融合部分的右边界; x_i 为待融合像素的横坐标,且 $x_1 \leqslant x_i \leqslant x_2$. 加权平均法计算简单、效率高,融合效果好,因此应用更为广泛.另外,还可以使用拉普拉斯图像金字塔[42]、小波变换[43]等方法实现图像融合,但操作复杂、耗时较长.

4 图像拼接面临的问题与发展

在各个领域,图像拼接都面临通用性、鲁棒性与实时性的问题.由于图像种类繁多,尺寸、亮度、图像特征多变,尚没有一种算法可以适用所有情况的图像拼接,找到通用性强的图像拼接方法是目前重要的研究方向.在大尺度缩放、旋转、平移的情况下,几何畸变、重叠区域面积等问题会对图像拼接的效率与精度产生很大影响,多数算法在上述条件差别较小时可以取得良好效果,但当上述条件差别较大时,可能存在拼接效果差、甚至拼接失败的情况,提高算法的适应性和鲁棒性仍是一个重要课题.另外,诸多实际场景,如军事卫星的成像、无人汽车对路况和行人的成像与识别等,对图像拼接的实时性提出了更高的要求.

目前,研究人员通常从算法和硬件两方面解决这些问题.在算法方面,常使用不同算法相结合的方式提升整体效率,如文献[44]用相位相关法对图像粗配准后,再用角点检测法进行精配准,在不影响实时性的前提下提高了对红外图像的配准精度;文献[45]将 Harris 的高效率与 SIFT 描述子的稳定性相结合,先用 Harris 算法提取特征点,再用改进的 SIFT 特征向量对特征点配对实现图像拼接;文献[46]先根据图像的梯度和不变矩确定两幅图像的大致重叠区域,再用 SIFT 对重叠区域实现配准,极大提高了图像配准的鲁棒性和实时性;文献[47]使用 Harris 算法对 SIFT 提取的特征点进一步筛选,提高了特征点的准确性;文献[48]根据噪声点、边缘点、平滑点位于不同子空间的性质,用谱聚类的思想将像素进行分类并检测边缘,特别适用于包含有噪声的图像.在硬件方面,文献[49]用图形处理器(graphics processing unit, GPU)实现硬件加速,将配准速度提高了约 10 倍;文献[50]用基于现场可编程门阵列(field-programmable gate array, FPGA)并行处理的 Harris 算法,实现了视频的实时拼接.当前的图像处理理论和硬件发展水平都影响着图像拼接技术的发展与应用.

未来几年,深度学习可能会对图像拼接技术产生巨大推动作用.深度学习理论诞生于 20 世纪 60 年代,但由于硬件技术的限制,该技术在 21 世纪初才得到广泛应用,尤其是在计算机视觉和自动语音识别领域.文献[51]在仿真条件下,用全连接神经网络实现了医学上计算机断层扫描(computed tomography, CT)与超声成像(ultrasound scan, US)图像平移关系上的配准,但还不能应用在实际中.文献[52]用同一卷积神经网络结构分别实现了对手写数字、心脏核磁共振图像两种数据集的高效配准,但单个卷积神经网络对多种场景图像配准的适应性还有待进一步研究.文献[53]首次用卷积神经网络实现了无人汽车上多模态动态视频的实时配准,但对更多平台上、更多模态的图像配准是另一个待解决的问题.目前,深度学习极大促进了图像拼接领域的发展,对图像拼接的通用性、鲁棒性与实时性都产生了积极影响,但是深度学习在图像拼接领域的应用还处在起步阶段.更多方面,如软硬件结合的方法、对多个场景的适应性、与传统图像拼接方法相结合的理论框架等,还有待进一步研究.

5 结语

本文研究了图像拼接技术中的核心算法,首先介绍了图像拼接的技术背景、发展历史和应用领域,指出了图像拼接技术的目的与意义;然后概括了图像拼接的主要流程,包括图像预处理、图像配准和图像融合,并分析了图像之间的变换关系;着重分析对比了图像配准中的核心方法及各方法的特点,尤其对 SIFT 技术做了重点剖析,简述了图像融合的分类与常用技术,最后,分析了图像拼接当前面临的主要问题与未来的可能发展趋势.

目前,图像拼接技术是图像处理技术的重要研究领域,已经在卫星遥感、医学图像分析、虚拟现实、增强现实等方面得到了广泛应用,也是机器人技术、自动驾驶、太空探索等前沿技术领域的重要组成部分.随着计

算机性能的发展和数字图像处理理论的完善,以及近年来各技术领域对图像处理技术需求的提高,图像拼接技术无疑将会更进一步发展.

参考文献:

[1] GUPTA V, BUSTAMANTE M, FREDRIKSSON A, et al. Improving left ventricular segmentation in four-dimensional flow MRI using intramodality image registration for cardiac blood flow analysis[J]. Magnetic resonance in medicine, 2018, 79(1): 554-560.

[2] 詹斌. 遥感图像拼接系统[J]. 计算机系统应用, 2014, 23(5): 31-36.

[3] 杨迪航. 基于多全景相机拼接的虚拟现实和实景交互系统[D]. 杭州: 浙江大学, 2017.

[4] 朱云芳. 基于图像拼接的视频编辑[D]. 杭州: 浙江大学, 2006.

[5] VIGUIER R, LIN C C. Automatic video content summarization using geospatial mosaics of aerial imagery[C]//IEEE International Symposium on Multimedia. Miami, 2015: 249-253.

[6] SAUR G, KRUEGER W. Change detection in UAV video mosaics combining a feature based approach and extended image differencing[J]. The international archives of the photogrammetry, remote sensing and spatial information sciences, 2016, 41(B7): 557-562.

[7] TOYAMA F, SHOJI K. Image mosaicing from a set of images without configuration information[C]//The 17th International Conference on Pattern Recognition. Cambridge, 2004: 899-902.

[8] REDDY B S. An FFT-based technique for translation, rotation, and scale-invariant image registration[J]. IEEE transaction on image processing, 1996, 5(8): 1266-1271.

[9] 邹进贵, 杨义辉, 李琴, 等. 局部相位相关法在密集匹配中的应用及精度分析[J]. 测绘通报, 2018, 1 (5): 25-28.

[10] RENUKA DEVI S M. Image mosaicing using phase correlation and feature based approach: a review[J]. International journal of engineering research, 2016, 4(1): 270-275.

[11] HARRIS C. A combined corner and edge detector[C]//The Fourth Alvey Vision Conference. Manchester, 1988: 147-151.

[12] LOWE D G. Object recognition from local scale-invariant features[C]//The Seventh IEEE International Conference on Computer Vision. Kerkyra, 1999: 1150-1157.

[13] LOWE D G. Distinctive image features from scale-invariant key-points[J]. International journal of computer vision, 2004, 60(2): 91-110.

[14] ASCHENBRENNER K P, GUTHIER C V. Feasibility of using single photon counting X-ray for lung tumor position estimation based on 4D-CT[J]. Zeitschrift für medizinische physik, 2017, 27(3): 243-254.

[15] 贾迪, 杨宁华, 孙劲光. 像对匹配的模板选择与匹配[J]. 中国图像图形学报, 2017, 22 (11): 1512-1520.

[16] MOSTAFAVI H, SLOUTSKY A, JEUNG A. Detection and localization of radiotherapy targets by template matching[C]//The 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. San Diego, 2012: 6023-6027.

[17] ZHAO T S. Small target detection and recognition based on template matching[C]//The 4th International Conference on Machinery, Materials and Information Technology Applications. Xi′an, 2016: 964-968.

[18] COOLEY J W, TUKEY J W. An algorithm for the machine calculation of complex fourier series[J]. Mathematics of computation, 1965, 19(90): 297-301.

[19] CASTRO E D, MORANDI C. Registration of translated and rotated images using finite fourier transforms[J]. IEEE transactions on pattern analysis and machine intelligence, 1987, 9(5): 700-703.

[20] 邵向鑫. 数字图像拼接核心算法研究[D]. 长春: 吉林大学, 2010.

[21] XU J, HONG Y X, XIN S X, et al. Medical image mosaic technology based on image phase correlation[C]//The 4th International Conference on Digital Home. Guangzhou, 2012: 274-277.

[22] WANG W, CHEN H J, LI J P, et al. A registration method of fundus images based on edge detection and phase-correlation[C]//The 1st International Conference on Innovative Computing, Information and Control. Beijing, 2006: 572-576.

[23] FISCHLER M A, BOLLES R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Graphics and image processing, 1981, 24(6): 381-395.

[24] ROSTEN E, DRUMMOND T. Machine learning for high-speed corner detection[J]. European conference on computer vision, 2006, 3951(1): 430-443.

[25] ROSTEN E, PORTER R, DRUMMOND T. Faster and better: a machine learning approach to corner detection[J]. Computer science, 2008, 32(1): 105-119.

[26] CHEN X, LIU L, SONG J X, et al. Corner detection and matching for infrared image based on double ring mask and adaptive

SUSAN algorithm[J]. Optical and quantum electronics, 2018, 50(4): 193-202.
[27] SMITH S M, BRADY J M. SUSAN—a new approach to low level image processing[J]. International journal of computer vision, 1997, 23(2): 45-78.
[28] WANG Z C, LI R, SHAO Z H, et al. Adaptive Harris corner detection algorithm based on iterative threshold[J]. Modern physics letters B, 2017, 31(15): 1750181.
[29] WU T, ZHANG Z L, ZHAO J Y, et al. An adaptive FAST corner detector based on high contrast grid image[C]//The 2nd International Conference on Mechatronics Engineering and Information Technology. Dalian, 2017: 591-597.
[30] LI H L, GUO L H, CHEN T, et al. The corner detector of the teeth image based on the improved SUSAN algorithm[C]//The 3rd International Conference on Biomedical Engineering and Informatics. Yantai, 2010: 609-612.
[31] 曾俊, 李德华. 彩色图像 SUSAN 边缘检测方法[J]. 计算机工程与应用, 2011, 47 (15): 194-196.
[32] 吕冀, 汪渤, 高洪民. 快速 SUSAN-Laplace 角点算法[J]. 光学技术, 2008, 34(6): 873-876.
[33] MANORANJITHAM R, DEEPA P. Efficient invariant interest point detector using Bilateral-Harris corner detector for object recognition application[J]. Multimedia tools and applications, 2018, 77(8): 9365-9378.
[34] 高华. Log-Gabor 梯度方向下的角点检测[J]. 中国图像图形学报, 2017, 22 (6): 797-806.
[35] LI H, MITRA S K. A contour-based approach to multi-sensor image registration[J]. IEEE transactions on image processing, 1995, 4(3): 320-334.
[36] 靳秋霞. 基于 Freeman 链码的二维曲线匹配[J]. 计算机应用, 2015, 35 (S1): 247-251.
[37] HU M K. Visual pattern recognition by moment invariants[J]. Information theory, 1962, 8(2): 179-187.
[38] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features (SURF)[J]. Computer vision and image understanding, 2008, 110(3): 346-359.
[39] 巨刚, 袁亮, 刘小月, 等. 基于改进 SURF 算法的移动目标实时图像配准方法研究[J]. 通信学报, 2017, 38 (1): 177-186.
[40] 朱炼, 孙凤, 夏芳莉, 等. 图像融合研究综述[J]. 传感器与微系统, 2014, 33 (2): 14-18.
[41] 徐书欣,葛洪央.形态人波和对比敏感度函数结合的图像融合算法[J]. 信阳师范学院学报(自然科学版),2017,30(3): 449-453.
[42] DOGRA A, GOYAL B, AGRAWAL S. Osseous and digital subtraction angiography image fusion via various enhancement schemes and Laplacian pyramid transformations[J]. Future generation computer systems, 2018, 82(1): 149-157.
[43] LI H, MANJUNATH B S, MITRA S K. Multi-sensor image fusion using the wavelet transform[J]. Graphical models and image processing, 1995, 57(3): 235-245.
[44] 袁梦笛, 顾国华, 钱惟贤, 等. 红外搜索跟踪系统中实时图像配准的研究和实现[J]. 红外技术, 2012, 34 (9): 497-502.
[45] ZHANG N, HOU J, WANG R W. Harris-SIFT algorithm optimization and its applications on nonwovens[C]//Advances in Materials, Machinery, Electronics. Wuhan, 2017: 123-135.
[46] AGRAWAL V, BHEDA D. Optimized image mosaicing with moment invariants and SIFT features[J]. Oriental journal of computer science and technology, 2017, 10(1): 66-75.
[47] 张震, 邵星星. 一种 SIFT 虹膜匹配算法[J]. 郑州大学学报(理学版), 2017, 49(3): 14-19.
[48] 郭新, 徐明, 张众. 基于谱聚类的边缘检测算法[J]. 郑州大学学报(理学版), 2018, 50(3): 83-87.
[49] CHEN S F, QIN J, XIE Y M, et al. CUDA-based acceleration and algorithm refinement for volume image[C]//International Conference on Future Bio-medical Information Engineering. Sanya, 2009: 544-547.
[50] OKUMURA K, RAUT S, GU Q Y, et al. Real-time feature-based video mosaicing at 500 fps[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems. Tokyo, 2013: 2665-2670.
[51] SUN Y Y, MOELKER A, NIESSEN W J, et al. Towards robust CT-ultrasound registration using deep learning methods[C]//Understanding and Interpreting Machine Learning in Medical Image Computing Applications. Granada, 2018: 43-51.
[52] VOS B D, BERENDSEN F F, VIERGEVER M A, et al. End-to-end unsupervised deformable image registration with a convolutional neural network[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Quebec, 2017: 204-212.
[53] GIERING M, VENUGOPALAN V, REDDY K. Multi-modal sensor registration for vehicle perception via deep neural networks[C]//IEEE High-performance Extreme Computing Conference. Waltham, 2015: 1-6.

(下转第 29 页)

[17] 梁军, 柴玉梅, 原慧斌,等. 基于极性转移和 LSTM 递归网络的情感分析[J]. 中文信息学报, 2015, 29(5):152-159.

[18] LI D, QIAN J. Text sentiment analysis based on long short-term memory[C]// 1st IEEE International Conference on Computer Communication and the Internet.Wuhan, 2016: 471-475.

Research on the Prediction of Group User Sentiment Trend Based on Commodity Comment

ZHOU Junpeng[1], GAO Ling[1,2], CAO Rui[1], GAO Quanli[2], ZHENG Jie[1], WANG Hai[1]

(1. *School of Information Science and Technology*, *Northwest University*, *Xi'an* 710027, *China*;

2. *School of Computer Science*, *Xi'an Polytechnic University*, *Xi'an* 710048, *China*)

Abstract: The group user sentiment trend prediction method which was based on commodity comment was proposed. Firstly, the analysis method of emotional characteristic words based on BosonNLP was put forward, to select the commodity comment information in the time dimension. Secondly, MLP was constructed by using group user multi-dimensional feature vector to analyze user's sentiment. Finally, the group user time sentiment tendency sequence was constructed by integrating the comment time and the user's sentimental trendency value. The sequence sentimental trend prediction was performed by using the LSTM model. Experimental results of large-scale real datasets indicated that the MLP model had a good classification effect. Compared with the existing autoregressive (AR) model, the average *MSE* of the LSTM was reduced by 79.06%, which could achieve a more accurate prediction result.

Key words: group user; commodity comment; sentiment analysis; time series; trend prediction

(责任编辑:孔　薇)

(上接第 10 页)

A Review on Image Mosaicing Techniques

PEI Hongxing, LIU Jinda, GE Jialong, ZHANG Bin

(*School of Physics and Engineering*, *Zhengzhou University*, *Zhengzhou* 450001, *China*)

Abstract: Image mosaicing was analyzed and summarized comprehensively. Based on the development of this technology the background and applications were described generally. The definition and steps of image mosaicing were indicated initially, and the algorithms of image registration were examined by classifying them into several groups. For each algorithm, especially the scale-invariant feature transform (SIFT), the basic theory and characteristics were demonstrated in detail. In the last part, the classifications of image fusion were illustrated and methods of this stage, specially the weighted average, were explained briefly. In addition, the challenges and tendencies of image mosaicing were also pointed out.

Key words: image mosaicing; image registration; image fusion; scale-invariant feature transform(SIFT)

(责任编辑:王浩毅)