

DOI:10.13203/j.whugis.20170038



文章编号:1671-8860(2019)02-0260-08

基于重叠区域相关系数的视频影像 关键帧提取算法

卢小平¹ 卢遥^{2,3} 焦金龙^{1,4} 童小华² 张继贤³

1 河南理工大学矿山空间信息技术国家测绘地理信息局重点实验室,河南 焦作,454003

2 同济大学测绘与地理信息学院,上海,200092

3 国家测绘产品质量检验测试中心,北京,100830

4 河南省遥感测绘院,河南 郑州,450003

摘要:提出了3种基于影像重叠区相关系数的视频影像关键帧快速提取算法。该算法首先采用多项式快速、准确地拟合出视频影像相关系数的变化趋势,然后利用相邻关键帧影像重叠区的高相关特性实现关键帧帧位的快速、准确提取,并通过实例与基于重叠度计算关键帧算法的时效性进行对比。结果表明,该算法单次计算相关系数的时间相对于单次重叠度计算时间缩短了近4倍,准确度达到88%,且针对不同拍摄手段下的视频影像,提取的关键帧影像精度较高。同时,还对基于视频提取的关键帧影像进行全景图拼接,结果显示未出现明显场景缺失,且可靠性高,适用性强。

关键词:视频流数据;关键帧;多项式拟合;重叠度;相关系数

中图分类号:P231;TP37

文献标志码:A

近年来,视频数据在诸多领域得到了广泛应用,尤其是随着视频监控技术及搭载平台的快速发展,视频监控已成为应急指挥的主要技术手段。但由于视频影像相邻帧间重叠度大、存在大量数据冗余,将所有视频数据作为处理对象会显著增加数据处理的工作量,从而降低数据的处理效率。因此,将视频影像中满足一定重叠度要求的部分帧(称为关键帧)作为处理对象,是减少数据的处理量、提高计算效率的重要技术手段。

目前,视频影像的关键帧提取方法大多是基于视频内容的检索^[1-7],即首先建立当前关键帧的特征(聚类)空间,然后计算下一帧与当前关键帧的特征距离,并通过设置特征距离阈值对关键帧进行筛选提取,计算效率较高。但如果将视频影像用于测绘方面,相邻关键帧之间就必须满足一定的重叠度要求,而采用上述算法获取的关键帧影像之间的重叠度浮动差异较大,无法满足测绘精度要求。针对上述问题,许多学者提出了新的关键帧提取算法,如Li等^[8]、Szeliski^[9]、De Haan等^[10]、张剑清等^[11]利用同名点匹配方法,通过逐

帧计算并筛选关键帧,准确得到了满足重叠度要求的关键帧,但这些算法由于采用逐帧计算,处理效率低,难以满足对时效性要求较高的应用领域。刘永等提出了一种分层式自适应帧采样的关键帧提取算法,对关键帧进行二次筛选,提高了拼接的效率^[12]。Fadaeieslam等^[13]、林宗坚等^[14]通过Kalman滤波预测影像4个角点的运行轨迹,实现计算相邻帧图像之间的重叠度,但由于受飞行姿态不稳定的影响,该方法难以通过数学模型进行准确预测。李朝奎等^[15]提出基于POS(positioning and orientation system)的关键帧选取算法,但其仅适应于飞行姿态相对稳定的视频影像。刘善磊等根据摄像机的规格参数、飞行平台速度、地面相对高程等先验知识,计算出相邻帧之间的理论重叠度,并按照固定的帧幅间隔筛选关键帧,但该方法在地形起伏较大地区难以保证达到对重叠度的要求。邢诚^[17]提出了一种简化SIFT(scale-invariant feature transform)计算重叠度的方法,提高了关键帧提取的效率。柳长安等^[18]提出一种基于影像空间地理位置的快速提取算法。

收稿日期:2018-04-11

项目资助:国家重点研发计划(2016YFC0803103);航空遥感技术国家测绘地理信息局重点实验室开放课题(2016B12)。

第一作者:卢小平,博士,教授,博士生导师,主要研究方向为摄影测量与遥感。hpuluxp@163.com

通讯作者:卢遥,博士生,工程师。156291029@qq.com

上述关键帧提取算法均是通过计算当前关键帧与上一帧之间的重叠度来提取关键帧,即设置初始关键帧及重叠度的阈值,如果当前帧符合重叠度阈值要求,则将其视为关键帧;否则,递归到下一帧继续进行计算。上述算法虽然都能较为准确地提取出关键帧,但必须循环计算上一关键帧影像与后续每一帧影像的航向重叠度,直至满足重叠度要求的下一关键帧出现。由于重叠度无法直接量测,只能通过同名点计算得出,而提取同名点并进行匹配需要大量的计算工作,从而降低了数据处理的效率。因此,本文提出一种基于重叠区域相关系数的关键帧快速提取算法,该算法依据相邻关键帧重叠区域的高相关性计算相关系数,并采用多项式拟合的方法拟合出视频影像的相关系数变化趋势,实现对关键帧帧位的准确定位,从而快速准确地提取出关键帧。

1 基于重叠区域相关系数的关键帧快速提取

1.1 算法原理

视频影像相邻视频帧之间的重叠度通常高达 90%,相邻帧之间存在较大的相关性(两幅影像的重叠区域越大,其相关性就越大),因此可以利用相邻视频帧之间的相关系数替代重叠度进行关键

$$r = \frac{\sum_{i=1}^m \sum_{j=1}^n (I_1(i,j) - E(I_1)) \times (I_2(i,j) - E(I_2))}{\sqrt{\sum_{i=1}^m \sum_{j=1}^n (I_1(i,j) - E(I_1))^2 \times \sum_{i=1}^m \sum_{j=1}^n (I_2(i,j) - E(I_2))^2}} \quad (3)$$

由式(3)可以看出,解算相关系数 r 实质上就是两幅灰度图像的矩阵运算。由于图像的灰度矩阵可以通过计算机直接读取,避免了对重叠度进行点特征提取及影像匹配,从而减少了计算量,提高了数据处理的效率。在视频帧图像处理过程中,考虑到相邻关键帧之间的不相关区域(非重叠区域)约占整幅影像的 40%(本文以 60%作为对相邻关键帧的重叠度要求),如果将相邻两幅帧图像直接进行相关性计算,势必会增加算法的错判率,降低处理结果的准确性。为提高计算精度,本文将相邻两幅图像的相关区域作为处理对象,并对式(1)进行了改进和优化,即:

$$r = \frac{\sum_{i=1}^m \sum_{j=1}^{n \times P} (I'_1(i,j) - E(I'_1)) \times (I'_2(i,j) - E(I'_2))}{\sqrt{\sum_{i=1}^m \sum_{j=1}^{n \times P} (I'_1(i,j) - E(I'_1))^2 \times \sum_{i=1}^m \sum_{j=1}^{n \times P} (I'_2(i,j) - E(I'_2))^2}} \quad (6)$$

改进后的式(6)虽然与式(3)结构相似,但实

际参与计算的数据仅为影像重叠区域,数据量压

帧提取。
相关系数是标准化的协方差函数,协方差函数除以两个信号的方差即得相关系数。当目标影像的灰度与搜索影像的灰度之间存在线性畸变时,仍然能较好地评价其相似性程度。相关系数可通过两幅影像的灰度矩阵直接计算求解,即:

$$r = \frac{D_{I_1 I_2}}{\sqrt{D_{I_1} \times D_{I_2}}} \quad (1)$$

式中, r 为相关系数;参数 D_{I_1} 、 D_{I_2} 、 $D_{I_1 I_2}$ 的计算式为:

$$\begin{cases} D_{I_1} = \frac{\sum_{i=1}^m \sum_{j=1}^n (I_1(i,j) - E(I_1))^2}{m \times n} \\ D_{I_2} = \frac{\sum_{i=1}^m \sum_{j=1}^n (I_2(i,j) - E(I_2))^2}{m \times n} \\ D_{I_1 I_2} = \frac{\sum_{i=1}^m \sum_{j=1}^n (I_1(i,j) - E(I_1)) \times (I_2(i,j) - E(I_2))}{m \times n} \end{cases} \quad (2)$$

式中, m 、 n 为灰度矩阵的总行数、总列数; $E(I)$ 为灰度矩阵灰度均值; $I(i,j)$ 为灰度矩阵 i 行 j 列的灰度值。

将式(2)代入式(1),可得到相关系数的计算公式:

$$r = \frac{D((I_1 \cap I_2)_1, (I_1 \cap I_2)_2)}{\sqrt{D(I_1 \cap I_2)_1 \times D(I_1 \cap I_2)_2}} \quad (4)$$

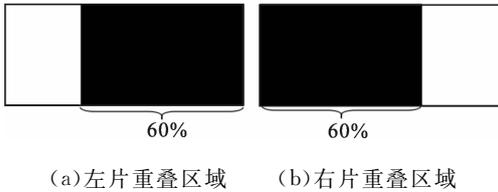
式中, $(I_1 \cap I_2)_1$ 、 $(I_1 \cap I_2)_2$ 为左右影像重叠区域的灰度矩阵(如图 1 中黑色区域),分别用 I'_1 、 I'_2 替换,其计算公式如下:

$$\begin{cases} I'_1(i,j) = I_1(i,j + n \times (1 - P) + 1) \\ I'_2(i,j) = I_2(i,j) \\ (i = 1, 2, \dots, m; j = 1, 2, \dots, n \times P) \end{cases} \quad (5)$$

式中, $I'_1(i,j)$ 、 $I'_2(i,j)$ 为影像重叠区域灰度矩阵; P 为影像重叠度。

结合式(2)可将式(4)变换为如下形式:

实际参与计算的数据仅为影像重叠区域,数据量压



(a)左片重叠区域 (b)右片重叠区域

图1 左右相邻影像的重叠区域

Fig. 1 Overlapping Region of Left-Right Adjacent Images

缩为原数据的 P 倍,显著提高了计算效率。同时,由于该算法仅有重叠区域参与计算,在一定程度上抑制了不相关区域的影响,提高了相关系数法判别重叠区域的准确性。

1.2 多项式曲线拟合算法

本文提出的关键帧提取算法是以相关系数作为判断准则,建立当前帧与上一关键帧之间相关系数的算法模型,无需逐帧检测影像之间的相关系数,从而提高计算效率。在计算相关系数时,若将当前关键帧中的重叠区域视为运动目标,像幅视为静止对象,就会发现重叠区域范围会随着拍摄时间的延续逐渐向像幅后方移动,直至完全离开。由图1可以看出,当像幅重叠区域从图1(a)移动到图1(b)位置时,左右相邻影像的重叠度达到60%,此时相关系数最大,即随着重叠区域的移动,相关系数呈先增大、后减小的变化趋势,而最大值对应的影像则为下一关键帧影像。

将少量影像作为实验数据进行多项式拟合,能够近似拟合出影像相关系数变化趋势,尤其是当数据量大且具有一定规律时,多项式拟合能够快速、真实地呈现数据的动态变化趋势,具有计算效率高、拟合真实度高、定位准确等特点。因此,如何选择最佳的多项式拟合形式以及最优的数据采样间隔,能准确反映结果的细微变化(即曲线拟合函数返回的极值位置与实际位置差值较小),实现数据极值位置的快速定位,是本文重点研究的问题。

多项式拟合就是利用实验数据 $A=(x, y)$ 求取近似函数 $y=f(x)$,反映离散数据的分布规律,从而实现对数据的分析与预测。其一般形式为:

$$y = f(x) = a_1 x^m + a_2 x^{m-1} + \dots + a_m x + a_{m+1} \quad (7)$$

式中, x 为输出量; y 为被测物理量; $a_1, a_2 \dots a_n$ 为多项式系数。

为保证近似函数的准确性,以最小二乘法为约束条件,使实验数据的 y 值与函数 $f(x)$ 差值的平方和最小,则数学表达式为:

$$\varphi = \varphi(a_1, a_2 \dots a_{m+1}) = \sum_{i=1}^{m+1} \left(\sum_{k=0}^m a_{m-k+1} x_i^k - y_i \right)^2 \quad (8)$$

式中, m 为多项式的阶数。为筛选出相关系数最大值所在的帧位,需要尽量多的影像参与解算,以反映出相关系数的变化趋势,但参与解算的影像过多会降低计算效率。因此,本文通过设定关键帧位置,并在该关键帧设定的邻域内按照一定的采样间隔选取取出4帧影像,将这4帧影像、理论关键帧与初始关键帧联合解算相关系数,得到离散点分布图,然后通过曲线拟合获得关键帧的准确帧位。

为了真实描述视频流影像的变化趋势,本文结合参与拟合的影像数量采用4次多项式对参与拟合影像的相关系数进行曲线拟合,则式(7)可化简为:

$$y = a_1 x^4 + a_2 x^3 + a_3 x^2 + a_4 x + a_5 \quad (9)$$

1.3 实验方法

为了验证本文算法的适用性及准确性,分别对水平拍摄、垂直拍摄以及航拍正视获取的视频影像进行关键帧提取,其中视频帧率为30帧/s,水平拍摄视频影像像幅为 1080×1920 ,垂直拍摄视频影像像幅为 1280×720 ,航拍正视视频影像像幅为 1920×1080 。基于帧率、像幅及拍摄状态等先验知识,将获取初始关键帧与第2关键帧的理论帧间距 T 作为本文关键帧提取算法的初始帧间距。然后以最佳的采样间隔 Δ 进行多项式拟合,得到当前关键帧帧位,并将当前帧间距作为与下一关键帧的理论帧间距,直至所有关键帧提取完毕。其中参与实验的视频影像如图2所示。

采用 Matlab 编程实现对视频数据进行分割并获取序列影像,得到与初始关键帧帧间距为 T 的帧影像,在其设定邻域内以特定的采样间隔 Δ 选取4帧影像,分别与初始关键帧计算相关系数,直至所有相关系数计算完毕;利用多项式对相关系数进行拟合,将拟合函数最大值处的影像帧设为关键帧;计算与初始关键帧的实际帧间距,将该帧间距作为与下一关键帧的初始帧间距,重复上述过程直至关键帧选取完成。本文总体技术流程如图3所示。

2 实验与分析评价

为验证本文所提算法的适用性,将本文算法选取的关键帧影像进行重叠度计算,并对关键帧影像选取的正确性进行判定;然后与基于重叠度的关键帧选取算法的计算效率、准确性两项指标进行对比;最后将关键帧拼接的全景影像与视频全景拼接影像进行对比,作为判断场景信息是否缺失的依据。

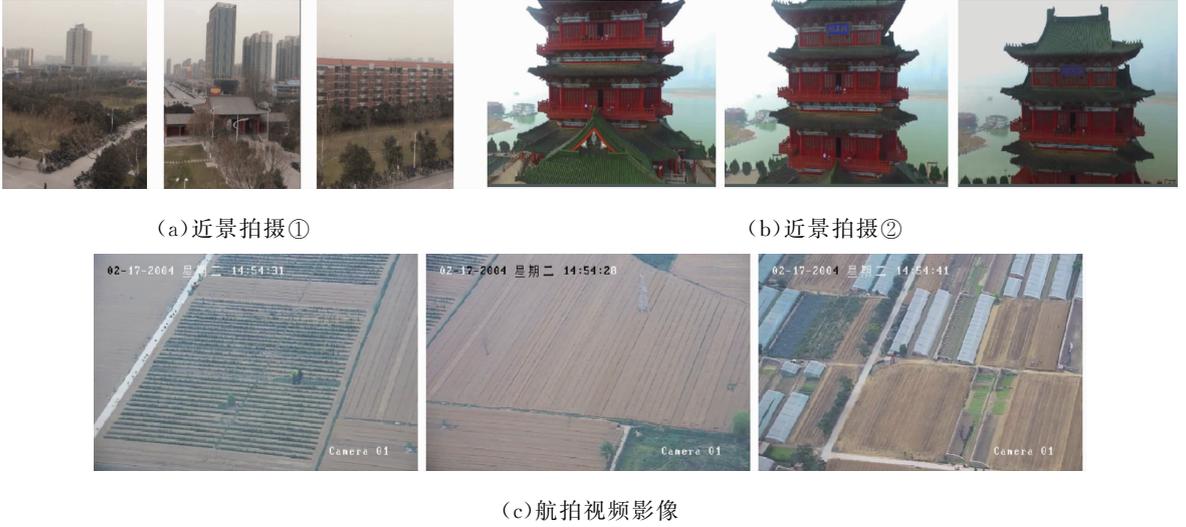


图 2 视频影像

Fig. 2 Video Images

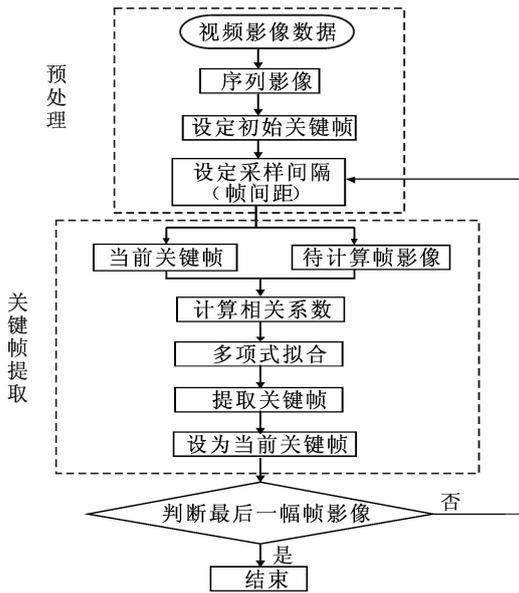


图 3 相关系数法关键帧提取流程

Fig. 3 The Process of Key Frame Extraction Based on Correlation Coefficient Algorithm

2.1 多项式拟合精度分析

多项式拟合精度取决于拟合曲线能否近似地反映出视频流的实际变化趋势。拟合曲线是由 5 个离散点(按照设定的采样间隔)通过式(9)拟合得到,因此采样间隔的确定是多项式拟合准确性的重要前提。

为确定最佳采样间隔 Δ , 分析不同采样间隔的多项式拟合精度, 本文采用一组总数为 200 帧的序列影像进行实例验证。设第 1 帧影像为关键帧, 第 100 帧预选为下一理论关键帧, 并由该帧位向左右两侧分别以 $\Delta = 5, 10, 15, 20, 30$ 为采样间隔各获取 5 幅帧影像, 分别计算每张帧影像与关键帧的相关系数, 以确定不同采样间隔的关键帧

最佳帧位。得到的离散点分布与拟合的曲线如图 4 所示, 不同采样间隔选取的关键帧帧位如表 1 所示。

表 1 曲线拟合帧位对应表

Tab. 1 Correspondence Table of Frame Positions Based on Curve Fitting

采样间隔 Δ	关键帧帧位	拟合值(峰值)	实际值	重叠度/%
5	97	0.903 6	0.912 4	61.1
10	98	0.894 4	0.913 2	60.4
15	99	0.892 7	0.904 5	59.6
20	103	0.893 7	0.859 1	58.3
30	107	0.898 1	0.836 3	57.6

从表 1 可以看出, 采样间隔 Δ 分别为 5、10、15 时, 关键帧帧位相差很小; 当采样间隔 Δ 分别为 20、30 时, 关键帧帧位变化明显。多项式拟合时, 较小的采样间隔对局部变化趋势拟合效果最佳。由于相机在拍摄过程中难以持续保持平稳, 骤然抖动或倾斜造成影像序列重叠度的陡增或骤减, 都会对较小采样间隔的拟合结果造成影响, 如出现峰值前移、后退甚至丢失等情况, 因此在确保拟合精度的同时, 应尽量加大采样间隔。由表 1 可知, 采样间隔 Δ 分别为 5 和 10 的拟合结果相似, 本文以 $\Delta = 10$ 进行多项式拟合, 图 5 为拟合出的关键帧影像与初始关键帧影像重叠区域, 由此看出, 多项式拟合选取的帧影像与关键帧影像重叠区域相似度高, 具有较高的定位精度。

2.2 对比分析

利用本文算法分别筛选出关键帧影像 8 帧、3 帧、26 帧, 对应序列影像数分别为 366 帧、310 帧以及 340 帧, 显著减少了参与计算的数据量。为验证本算法的有效性, 以准确度、时效性及信息量 3 个指标对关键帧选取结果进行对比分析。

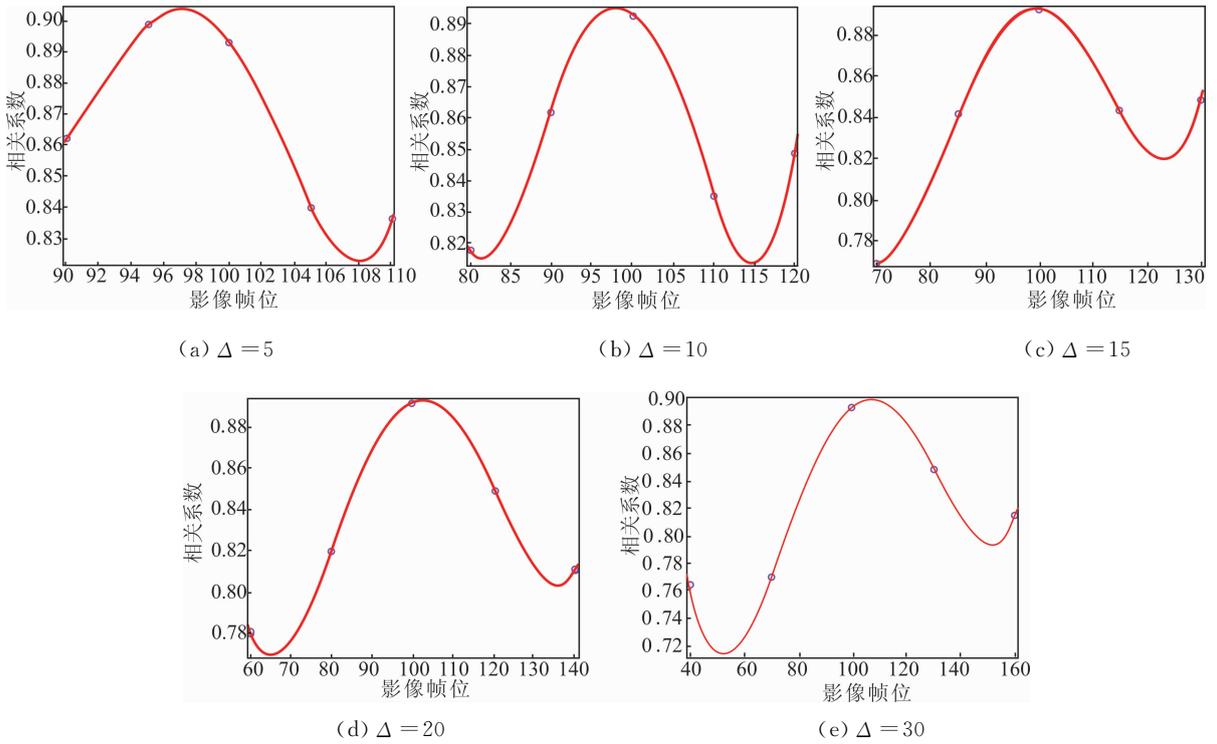
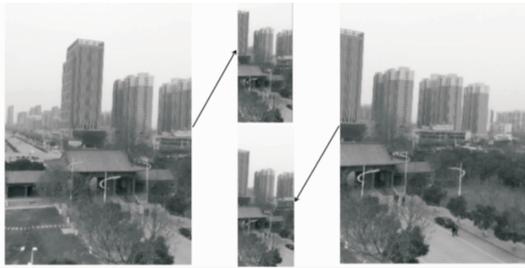


图4 离散点分布与拟合曲线

Fig. 4 Distribution of Discrete Points and Fitting Curves



关键帧影像 重叠区域 提取影像

图5 关键帧影像与提取影像的重叠区域

Fig. 5 The Overlapping Region of Key Frame Image and Selected Images

1) 准确度。准确度是指提取的关键帧之间的重叠度能够满足规定重叠度的比例,为验证本算法提取的准确度及对不同拍摄方式下的视频影像的适用性,依次进行关键帧提取。受拍摄过程中设备及环境影响,实际的影像重叠度不可能完全达到理论值,而是在该值附近摆动。因此,将计算重叠度与理论重叠度差值在2%内的关键帧视为准确提取,不同重叠度的提取结果准确度如表2所示。

表2 不同拍摄方式下相关系数法提取的准确度

Tab. 2 The Accuracy Extracted by Coefficient Correlation Algorithm of Different Shooting Modes

拍摄方式	最大值	最小值	平均值	关键帧帧数	准确度
水平	62.9%	60.3%	61.6%	8	88.0%
垂直	61.2%	60.2%	60.7%	3	100.0%
航拍	63.4%	57.3%	60.0%	26	84.5%

由表2可知,不同拍摄方式下提取的关键帧均在60%上下浮动,其中垂直拍摄准确度最高,水平拍摄次之,航拍最低。经分析发现,其主要因素在于拍摄过程中速度大小不一,随着拍摄速度的增大,姿态变化对相邻影像的重叠度也增大,而航拍无人机飞行速度较快,使得拍摄瞬间的轻微抖动也能造成影像之间重叠度的显著变化。此外,从表2中还可以发现,本文算法针对不同拍摄方式的视频影像,关键帧提取准确度均优于80%,具有较高适用性。

2) 时效性。本算法以相关系数作为判别关键帧影像依据,与基于重叠度的提取算法相比,缩短了提取关键帧的时间进度,如表3所示。

表3 相关系数法与其他算法对比结果

Tab. 3 The Comparison of Correlation Coefficient Algorithm and Other Algorithms

关键帧提取算法	平均单次相关系数(重	总计算	准确率
	叠度)计算时间/s	时间/s	/%
相关系数法	1.892	83.52	88
逐帧重叠度算法	7.423	1723	100
先验知识重叠度算法	7.423	103.79	53

从表3可以看出,相关系数与逐帧重叠度算法相比,虽准确率降低至88%,但单次计算相关系数的时间相对于单次重叠度计算时间缩短了近4倍,整体时效性提高近21倍;与先验知识重叠度算法相比,准确度及时效性均显著提高。

3)全景拼接效果对比。将关键帧影像拼接效果与逐帧影像拼接的效果进行对比,检查场景信

息有无损失。二者对比结果如图 6、图 7、图 8 所示。



(a)逐帧影像拼接效果



(b)关键帧影像拼接结果

图 6 近景拍摄视频影像①拼接图

Fig. 6 The Panchromatic Image① After Mosaic with Close-Range Photography



(a)逐帧影像拼接效果



(b)关键帧影像拼接结果

图 7 近景拍摄视频影像②拼接图

Fig. 7 The Panchromatic Image② After Mosaic with Close-Range Photography



(a)逐帧影像拼接效果



(b)关键帧影像拼接结果

图 8 航拍视频影像拼接图

Fig. 8 The Panchromatic Image After Mosaic with Aerial Photography

从图 6、图 7、图 8 可以看出,关键帧影像与逐帧影像拼接结果基本一致,信息量缺失并不明显,能够满足后续数据的处理需要。

3 结 语

针对视频影像信息冗余度高、处理效率低等问题,本文算法在实验中取得了比较满意的结果,在技术路线上主要作出以下创新。

1)提出了基于重叠区域相关系数的关键帧提取算法,以相邻帧的相关系数替代重叠度作为选

取关键帧的依据。

2)通过多项式拟合实现关键帧位置快速、准确定位。在提高了视频影像关键帧提取的计算速度的同时,利用少量关键帧有效保留了视频数据的大量场景信息,能够满足后续的数据处理。

参 考 文 献

- [1] Wolf W. Key Frame Selection by Motion Analysis [C]. IEEE International Conference on Acoustics, Speech & Signal Processing, Atlanta, Georgia, United States, 1996

- [2] Zhang H J, Wu J, Zhong D, et al. An Integrated System for Content-based Video Retrieval and Browsing[J]. *Pattern Recognition*, 1997, 30(4): 643-658
- [3] Gresle P O, Huang T S. Gisting of Video Documents: A Key Frames Selection Algorithm Using Relative Activity Measure[C]. The 2nd IntConf on Visual Information Systems, San Diego California, United States, 1997
- [4] Zhuang Y, Yong R, Huang T S, et al. Adaptive Key Frame Extraction Using Unsupervised Clustering[C]. IEEE International Conference on Image Processing, Chicago, Illinois, USA, 1998
- [5] Ferman A M, Tekalp A M. Multiscale Content Extraction and Representation for Video Indexing[J]. *Proceedings of SPIE - The International Society for Optical Engineering*, 1997, 3 229:23-31
- [6] Lu Weiyan, Xia Dingyuan, Liu Yi. An Approach of Key Frame Extraction Based on Mutual Information [J]. *Microcomputer Information*, 2007, 23(33): 298-300(陆伟艳, 夏定元, 刘毅. 基于内容的视频检索的关键帧提取[J]. 微计算机信息, 2007, 23(33):298-300)
- [7] Zhu Dengming, Wang Zhaoqi. Extraction of Key-frame from Motion Capture Data Based on Motion Sequence Segmentation[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2008, 20(6): 787-792(朱登明, 王兆其. 基于运动序列分割的运动捕获数据关键帧提取[J]. 计算机辅助设计与图形学学报, 2008, 20(6):787-792)
- [8] Li J, Pan Q, Yang T, et al. Automated Feature Points Management for Video Mosaic Construction [C]. IEEE International Conference on Information Technology & Applications, Sydney, Australia, 2005
- [9] Szeliski R. Video Mosaics for Virtual Environments [J]. *IEEE Computer Graphics & Applications*, 1996, 16(2):22-30
- [10] De Haan G, Biezen P W A C, Huijgen H, et al. True-motion Estimation with 3-D Recursive Search Block Matching[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 1993, 3(5): 368-379, 388
- [11] Zhang Jianqing, Pan Li. Photogrammetry[M]. 2nd ed. Wuhan: Wuhan University Press, 2012(张剑清, 潘励. 摄影测量学)[M]. 2版. 武汉:武汉大学出版社, 2012)
- [12] Liu Yong, Wang Guijin, Yao Anbang, et al. Video Mosaicking Based on Adaptive Sampling[J]. *J Tsinghua Univ(Sci & Tech)*, 2010, 50(1):108-112(刘永, 王贵锦, 姚安邦, 等. 基于自适应帧采样的视频拼接[J]. 清华大学学报(自然科学版), 2010, 50(1):108-112)
- [13] Fadaeieslam M J, Soryani M, Fathy M. Efficient Key Frames Selection for Panorama Generation from Video[J]. *Journal of Electronic Imaging*, 2011, 20(2):2 763-2 769
- [14] Lin Zhongjian, Zhang Yonghong. Measurement of Information and Uncertainty of Remote Sensing and GIS Data[J]. *Geomatics and Information Science of Wnhan University*, 2006, 31(7):569-572(林宗坚, 张永红. 遥感与地理信息系统数据的信息量及不确定性[J]. 武汉大学学报·信息科学版, 2006, 31(7):569-572)
- [15] Li Chaokui, Zhao Huibing, Yang Gang. Research on Separating Technology of UAVs Video Image Stream Based on the POS System[J]. *Journal of Hunan University of Science & Technology(Natural Science Edition)*, 2009, 24(3):51-55(李朝奎, 赵会兵, 杨刚. 基于POS系统的无人机视频影像分割技术研究[J]. 湖南科技大学学报(自然科学版), 2009, 24(3):51-55)
- [16] Liu Shanlei, Zhao Yingdi, Wang Ganghui, et al. A Method for Automatic Extraction of Key Frames [J]. *Science of Surveying and Mapping*, 2012, 37(5):110-115(刘善磊, 赵银娣, 王光辉, 等. 一种关键帧的自动提取方法[J]. 测绘科学, 2012, 37(5): 110-115)
- [17] Xing Cheng. Overlapping Degree Analysis of Images from an Unmanned Aerial Vehicle Based on a Reduced Scale-Invariant Feature Transform (SIFT) Algorithm[J]. *Journal of Harbin Engineering University*, 2012, 33(2):221-225(邢诚. 基于简化SIFT算法的无人机影像重叠度分析[J]. 哈尔滨工程大学学报, 2012, 33(2):221-225)
- [18] Liu Chang'an, Sun Jin, Wu Hua, et al. Research on Keyframe Extraction Pretreatment of Power-tower in Flying Robot Inspection Video[J]. *J Huazhong Univ of Sci & Tech (Natural Science Edition)* 2015, 43(s1):477-480(柳长安, 孙金, 吴华, 等. 飞行机器人电塔巡检视频关键帧提取预处理[J]. 华中科技大学学报(自然科学版), 2015, 43(s1):477-480)

Key Frame Extraction Algorithm for Video Images Based on Correlation Coefficient of Overlap Regions

LU Xiaoping¹ LU Yao^{2,3} JIAO Jinlong^{1,4} TONG Xiaohua² ZHANG Jixian³

1 Key Laboratory of Mine Spatial Information Technologies of National Administration of Surveying, Mapping and Geoinformation, Henan Polytechnic University, Jiaozuo 454003, China

2 School of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China

3 National Quality Inspection and Testing Center for Surveying and Mapping Products, Beijing 100830, China

4 Institute of Remote Sensing and Surveying and Mapping Henan, Zhengzhou 450003, China

Abstract: This paper presents an effective algorithm to extract key frames from video images based on the correlation coefficients derived from the overlapping regions of adjacent images. Firstly, the polynomial model is employed to fit the change trend of correlation coefficients. Secondly, the positions of key frames are located and extracted rapidly and correctly via the high correlation of overlapping regions of adjacent key frame images. Finally, the time-effectiveness of extracting key frames based on the degree of overlap is tested in a stance. The result shows that the proposed algorithm takes only near 1/4 of time consumed by single computing of the degree of overlap, with an accuracy up to 88% and it can achieve higher accuracy for extracting key frames with different requires of degree of overlap between video images. The panchromatic image generated from the mosaic of extracted key video frames shows no apparent missing of scenes, and the new algorithm demonstrates high reliability and transferability.

Key words: video stream data; key frame; polynomial fitting; degree of overlap; correlation coefficient

First author: LU Xiaoping, PhD, professor, specializes in photogrammetry and remote sensing. E-mail: hpuluxp@163.com

Corresponding author: LU Yao, PhD candidate, engineer, E-mail: 156291029@qq.com

Foundation support: The National Key Research and Development Program, No. 2016YFC0803103; the Key Laboratory for Aerial Remote Sensing Technology of NASG, No. 2016B12.

(上接第 259 页)

ne, we firstly estimate the spatial layout of the room in the image based on 3D indoor scene understanding. Then, according to phone's direction and position, we find the triangular patches of the targeted wall in the 3D geometric model. At the end, the (u, v) coordinates of the image are calculated by matching the convex points of the layout box with the corresponding patches of the 3D model. The conducted experiments demonstrate that the proposed method can automatically map a single image onto the 3D indoor model, which is meaningful for the augmented reality in the indoor space.

Keywords: mobile phone image; texture updating; 3D indoor scene understanding; structured prediction; fusion visualization

First author: MA Wei, PhD, specializes in the 3D GIS and spatial data fusion. E-mail: maweiweiweiwei@163.com

Corresponding author: XIONG Hanjiang, PhD, professor. E-mail: xionghanjiang@163.com

Foundation support: The National Key Research and Development Program of China, No. 2016YFB0502203; Mapping Geographic Information Industry Research Projects of Public Interest Industry, No. 201512009; the Special Research Funding of Stae Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing.