

doi: 10.6046/zrzyg.2021035

引用格式: 任超锋, 蒲禹池, 张福强. 顾及地理空间信息的无人机影像匹配像对提取方法[J]. 自然资源遥感, 2022, 34(1): 85–92. (Ren C F, Pu Y C, Zhang F Q. A method for extracting match pairs of UAV images considering geospatial information[J]. Remote Sensing for Natural Resources, 2022, 34(1): 85–92.)

顾及地理空间信息的无人机影像匹配像对提取方法

任超锋, 蒲禹池, 张福强

(长安大学地质工程与测绘学院, 西安 710054)

摘要: 针对无人机影像三维重建中匹配像对提取适应性差、效率低、需准确的先验知识等问题, 提出一种顾及无人机影像地理空间信息的匹配像对提取方法。首先利用主成分分析(principal component analysis, PCA)方法将高维度的特征降到低维特征, 提高检索词典构建效率; 其次, 通过计算查询影像之间的反距离权重因子, 构建综合检索因子, 提高相似影像之间的可区分性; 最后, 通过计算检索阈值, 舍弃阈值后的无效匹配像对, 进一步提高了影像查准率。实验结果表明, 与传统脚印图法和 128 维特征检索方法相比, 该方法获得了更高的处理效率及更全面的稀疏重建结果, 尤其针对海量无人机数据更具优势。

关键词: 匹配像对; 视觉词袋; 无人机; 三维重建; 影像检索

中图法分类号: TP 79 **文献标志码:** A **文章编号:** 2097-034X(2022)01-0085-08

0 引言

无人机(unmanned aerial vehicle, UAV)低空摄影测量具有现势性强、分辨率高、采集方式灵活、传感器多样化等优点, 近年来在应急测绘^[1]、城市三维重建^[2]、滑坡地形重建^[3]、文物三维重建^[4]等领域得到了越来越广泛的应用。然而, 由于低空 UAV 一般搭载非量测相机, 且受平台不稳定性及传感器多样化等诸多因素影响, 获取的影像尺度不一致, 且数据量庞大, 为影像匹配带来极大挑战。此外, 近年来新兴的仿地飞行、环绕飞行、贴近飞行等数据获取方式, 更使传统的影像匹配方法难以满足需求。

影像匹配作为三维自动重建的基础环节, 其效率与稳健性对重建结果起着决定性作用。尤其针对复杂地形条件下的三维重建, 能够获取的有效匹配像对越多, 影像连接成功的概率也越高。总体上可将其分为 2 个环节: ①提取匹配像对, 即从大数据量的遥感影像中按照一定方法提取具有重叠区的一对影像; ②按照影像匹配算法对提取的像对进行双像匹配提取匹配像对, 获取同名点列表。近年来, 针对多角度、大倾角的无人机影像, 文献[5–7]在尺度不变特征变换(scale-invariant feature transform,

SIFT)^[8]和仿射尺度不变特征变换(affine-scale-invariant feature transform, ASIFT)^[9]算法基础上提出了相应的匹配策略, 同时借助图形处理器(graphics processing unit, GPU)^[10]并行运算, 这一类方法均可快速完成像对的同名点匹配。然而, 如何从大量影像中提取匹配像对的研究还处于比较初级的研究阶段。

在 UAV 影像匹配像对提取方面, 文献[11–12]采用穷举遍历策略, 对影像集中任意两两影像进行匹配。该方法可靠性最高, 但存在大量盲目的无效运算, 效率太低; 文献[13]利用影像初始地理位置信息, 计算当前影像与其相邻影像的空间距离, 进而采用固定阈值范围内的像对进行匹配; 文献[14]则通过动态搜索的方法确定阈值, 提高了匹配的可信度, 但当测区内存在多层次、不同分辨率的 UAV 影像时, 其获取的搜索阈值无法保证为全局最优值; 文献[7, 15–16]则利用影像的定位定向系统(positioning and orientation system, POS)测量数据、传感器结构设计参数及内方位元素、测区地形信息等先验知识, 计算每张影像的脚印图, 进而利用脚印图的拓扑关系, 判定像对是否具有重叠区。由于该类方法计算量小, 且对常规数据获取方式具有较好的适应性, 因此是目前低空 UAV 影像匹配像对提取的主要方法。然而, 从原理可知, 该类方法高度依赖

收稿日期: 2021-02-07; 修订日期: 2021-04-15

基金项目: 国家自然科学基金项目“非常规摄影测量模式下无人机影像匹配像对的提取方法研究”(编号: 41801383)、国家重点研发计划项目“岩溶山区特大滑坡早期识别与监测预警”(编号: 2018YFC1504805)、中央高校基本科研业务费“海量无人机影像三维自动重建关键技术研究”(编号: 300102269206)和“三维模型智能重建”(编号: 300102269304)共同资助。

第一作者: 任超锋(1984–), 男, 博士, 讲师, 研究方向为低空无人机摄影测量。Email: ren_cf@163.com。

先验知识的准确性,因此,当测区地形信息不确定(如滑坡、山谷、独立地物等)、或者传感器设计参数无法准确得知时(如多镜头倾斜相机、组合倾摆相机),该类方法便无法准确计算出像对的相关性。文献[17]则从影像内容信息出发,利用提取的特征信息构建视觉词袋(bag of visual words, BoW)模型,进而利用影像检索方式确定待匹配像对;文献[18]在生成 BoW 模型过程中,计算海明嵌(hamming embedding, HE)来提升影像检索的准确度;文献[19]则提出了一种霍夫投票算法加速影像检索过程。由于这类方法一般面向的都是无序、无地理信息的网络图像数据,其检索结果与影像是否具有重叠区没有明确关系,且计算量过大,难以直接使用。

针对上述问题,本文提出一种顾及影像地理空间信息的 BoW 模型方法来确定待匹配像对,对地形条件、影像获取方式、传感器类型均无限制条件,进而减少影像匹配过程中的冗余计算,以解决 UAV 影

像匹配像对的高效、准确提取问题。

1 方法原理

本文提出的 UAV 影像匹配像对提取方法流程为:①按照文献[8]和[10]所述方法,依次完成测区所有影像的 SIFT 特征提取;②为了提高影像检索效率,对提取的 SIFT 特征进行降维,降低生成视觉词汇树的运算量;③采用文献[20]方法,利用降维后的特征向量构建视觉词汇树;④检索所有影像,并计算词汇树内单词的检索权重;⑤在词汇树内,查询与当前影像最相似的影像列表,并计算其与查询列表内影像的空间距离指数,综合相似指数与空间距离指数对检索列表进行排序,最后利用综合指数计算查询深度阈值,将阈值之前的查询影像与当前影像组合形成匹配像对。具体方法流程如图 1 所示。

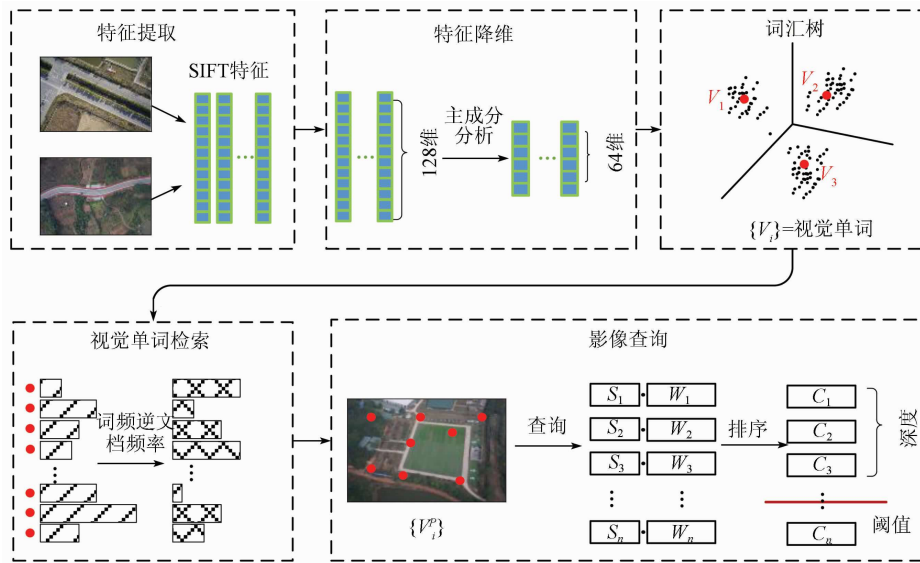


图 1 顾及地理空间信息的 UAV 影像匹配像对提取方法流程

Fig. 1 Flowchart of UAV image pairs extraction method referring to image geospatial information

1.1 特征向量降维

SIFT 算法以其尺度、旋转不变性并能克服一定程度仿射变形和光照变化得以在影像匹配领域得到广泛使用^[7],但原始的 SIFT 特征包含 128 维特征向量,若将其直接用于影像检索,会产生大量的高维度运算,造成影像检索效率过低。因此,本文采用主成

分分析(principal component analysis, PCA)方法对高维度的 SIFT 特征向量进行降维。

将构建视觉词汇树的 m 个 SIFT 特征组成矩阵 $X_{m \times 128}$,按照 PCA 原理对其进行奇异值分解(singular value decomposition, SVD),即

$$X_{m \times 128} = U_{m \times m} W_{m \times 128} V^T_{128 \times 128} = U \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \\ & & & & 0 \end{bmatrix} V^T, \quad (1)$$

式中: U 和 V 分别为 m 阶和 128 阶正交矩阵;

$W_{m \times 128}$ 为 r 个降序排列的特征值 $\sigma_i (i = 1, 2, \dots, r)$

构成的 $m \times 128$ 矩形对角矩阵。

若将 SIFT 特征向量降维至 d 维,则取右奇异值向量 V 的前 d 个特征向量组成矩阵 $V_{d \times 128}$,则降维后的 SIFT 特征向量 $X'_{m \times d}$ 为:

$$X'_{m \times d} = X_{m \times 128} V_{d \times 128}^T \quad (2)$$

1.2 视觉单词检索

将降维后的 SIFT 特征进行聚类,聚类的过程即为构建视觉词汇树过程。聚类之后,每一个聚类中心表示为一个视觉单词,一幅影像可以表示为多个视觉单词的无序集合,此时,UAV 影像之间的相似性判定即可转变为视觉单词之间的相似性判定。

本文采用文献[20]的方法构建层次词汇树,同时采用词频逆文档频率 (term frequency - inverse document frequency, TF-IDF) 评价某一个视觉单词对于视觉词汇树中某一影像的重要程度,其定义为:

$$t_i = f_i f_{id} = \frac{n_{ip}}{n_p} \ln \frac{N}{n_i} \quad (3)$$

式中: f_i 为词频; f_{id} 为逆文档频率; n_{ip} 为影像 p 中出现视觉单词 i 的数量; n_p 为影像 p 中出现的所有视觉单词数量; N 为影像总体数量; n_i 为包含视觉单词 i 的影像数量。

词频 f_i 表达了某个视觉单词在影像中的出现频率,而逆文档频率 f_{id} 则表达了该视觉单词在其他影像上的重复频率,两者组合之后可将视觉单词的重要性随着它在影像文件中出现的频率呈正比增加,同时也会随着它在视觉词典中出现的频率呈反比下降。

当视觉词汇树创建完成后,依次检索数据集中每一幅影像中视觉单词的出现频率。检索完成之后,即可计算其 TF-IDF 因子。此时,每一幅影像均可表示为一组带不同权重的视觉单词组合。评价 2 幅影像是否相似,即可通过计算 2 幅影像的单词向量点积完成,公式为:

$$s_i^j = V_i \cdot V_j^T \quad (4)$$

式中: $V_i = (t_1, \dots, t_i, \dots, t_k)$ 为第 i 幅影像的 k 个视觉单词所组成的 TF-IDF 因子向量; s_i^j 为第 i 幅影像与第 j 幅影像之间的相似因子,用以评价两者之间的相似度。

1.3 影像查询

相似因子只是评价 2 幅影像所含视觉单词的相似性,在大部分情况下,具有相似内容的影像一般也具有重叠区。然而,当地表类型比较单一时(如大片田地、灌木、裸露山地),相似的区域并非一定具有重叠区。此时,若将影像之间的空间距离作为影响因素参与评价,则可大大提高两者之间的相关性。图 2 为反距离权重因子示意图。

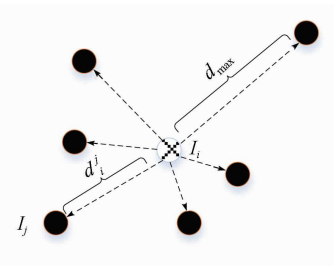


图 2 反距离权重因子

Fig. 2 Inverse distance weighting factor

如图 2 所示,当前影像 I_i 与查询影像 I_j 之间空间距离越近,则其存在重叠区的可能性越高。因此,本文计算当前影像与相似影像列表之间的反距离权重因子,用以评价两者之间的空间相关性,公式为:

$$w_i^j = 1 - \frac{d_i^j}{d_{\max}} \quad (5)$$

式中: $d_{\max}$ 为查询影像 I_i 与相似影像集之间最大的欧式距离; d_i^j 为查询影像 I_i 与相似影像 I_j 之间的欧式距离; w_i^j 为查询影像 I_i 与相似影像 I_j 之间的空间距离因子。将相似因子与空间距离因子组合起来形成综合权重因子,用以综合评价查询影像与相似影像之间的空间相关性 c_i^j ,公式为:

$$c_i^j = s_i^j w_i^j \quad (6)$$

1.4 查询深度阈值

查询深度是指以综合权重因子为依据,在影像集合中查询出与当前影像相似性最高的前 Q 张影像,组成待匹配像对,如图 3 所示。

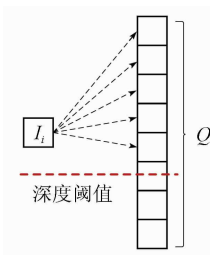


图 3 查询深度阈值

Fig. 3 Query depth threshold

图 3 中,查询影像 I_i 与查询深度 Q 共组成 Q 对像对进行匹配。实际处理过程中, Q 值过小会造成漏检,而 Q 值过大则会引入大量无效匹配像对,降低匹配效率。因此,本文采用查询深度阈值的方式对查询深度进行分割,仅将阈值前的影像与查询影像组成匹配像对进入匹配环节。阈值计算公式为:

$$\begin{cases} w_1 = N_1/Q \\ w_2 = 1 - N_1/Q \\ \mu_1 = \sum_{i=0}^l iC_i \\ \mu_2 = \sum_{i=l+1}^Q iC_i \\ g = w_1 w_2 (\mu_1 - \mu_2)^2 \end{cases} \quad (7)$$

式(7)以类间方差最大为原则,将 Q 内的影像分为前景与背景 2 部分。式中: N_1 为属于阈值 t 之前的影像数量; w_1 和 w_2 分别为前景和背景的影像频率; μ_1 为阈值 t 之前影像的综合因子平均值; μ_2 为阈值 t 之后的影像综合因子平均值; g 表示前景影像与背景影像之间的类间方差。在查询深度内,类间方差最大对应的位置即为查询阈值 t^* , 公式为:

$$g(t^*) = \max_{0 \leq t < Q} [g(t)] \quad (8)$$

表 1 试验无人机数据集

Tab. 1 Experimental UAV images dataset

数据	传感器	采集方式	数据类型	位置	主要地形	影像大小/像素	影像数量/幅	采集时间	地面分辨率/cm
A	单相机	等高	正射	陕西省华阳县	高山峡谷	7 952 × 5 304	1 659	2018-02-12	4
B	单相机	变高	正射	贵州省鸡场镇	高山滑坡	5 472 × 3 648	465	2019-07-25	10
C	单相机	等高/贴近	正射/倾斜	贵州省紫岭镇	高山滑坡	5 472 × 3 648	2 137	2019-07-29	8/1
D	双相机	等高	倾斜	山西省平陆县	丘陵	7 952 × 5 304	2 728	2019-04-07	2
E	五相机	等高	倾斜	陕西省大同市	城市	6 000 × 4 000	8 585	2019-08-26	2

基于 Windows10 64 位操作系统,采用 VC++ 2015 开发了海量 UAV 影像自动空三处理软件 MRI,用于测试本文方法的适应性。硬件平台为 Dell Precision 3630 工作站,CPU i7-8700K 3.7 GHz,内存 64 G DDR4,硬盘 512 G SSD,显卡为英伟达 Titan XP 12 G。

2.2 结果与分析

为了评价低维特征对影像检索效率及精度造成的影响,对表 1 中 5 组影像提取的原始 128 维特征向量进行降维,分别降至 96 维、64 维、32 维,并按照文献[19]方法对影像进行检索。检索过程中,生成词汇树的聚类中心数量统一设置为影像数量的 200 倍,检索深度统一设置为 100,构建词汇树的影像从数据集中随机提取,其数量设置为数据集影像数量的 20%,且最大影像数量不超过 500 幅。评价检索方法精度时,首先采用穷举法完成 5 组影像匹配,然后以其匹配结果为基准,评价其他检索方法的查询精度。

评价影像查询精度通常采用查准率和查全率(图 4)。查准率通过计算查询深度内正确的查询影像与查询深度的比值构成,它反映了查询过程中正

2 实验与分析

2.1 实验数据与配置参数

为验证本文方法在多采集方式、多传感器类型、多地形条件下提取匹配像对的可行性、精度与效率,共收集了 5 组实验数据进行实验。数据集的详细信息如表 1 所示。

确像对的比例,查准率越低,意味着匹配环节引入错误匹配像对数量越高,相应的匹配耗时也越高。查全率则通过计算查询深度内正确的查询影像与穷举法匹配中得到的所有正确影像数量比值构成,它反映了当前特征条件下,能提取到的匹配像对的完整度,查全率越低,意味着稀疏重建时的可靠性越低,容易丢片。因此,查准率与查全率之间相互制约,在查全率相当的前提下,查准率越高,则算法的效率越高。除了计算 128 维、96 维、64 维、32 维特征的检索精度外,按照本文方法分别计算 32 维和 64 维的综合检索因子检索精度,分别用 32G 和 64G 表述。从图 4(a)中可知,总体上,影像查询的查准率随着特征维度降低而递减,当特征维度降到 32 维时,其影像查准率明显降低。反观图 4(b),此时的查全率却最高。其原因主要是 32 维的影像特征丢失了过多的细节信息,使特征之间的可分性降低,在引入大量错误匹配像对基础上,也将大量弱连接的像对引入匹配环节。虽然提升了整体的查全率,但后续的匹配环节效率太低。因此,综合考虑效率与精度,将特征降低至 64 维进行影像检索是合适的。

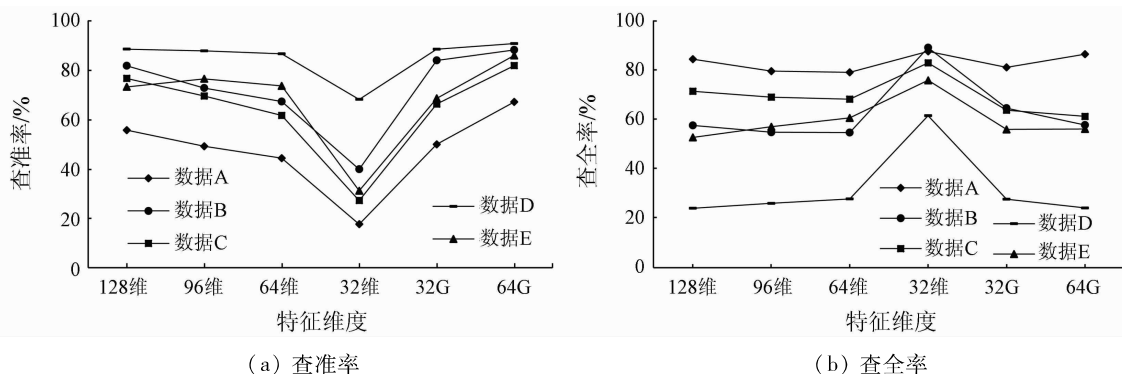


图 4 不同维度模式的影像查准率和查全率

Fig. 4 Image precision and recall of different dimensional models

表 1 中的 5 个测试数据集,数据 A 和 B 为正射类型,数据 C,D,E 可归为倾斜类型。正射类型中,随着特征维度降低,其查准率和查全率均缓慢降低,而本文的综合查询因子方法在相同特征维度条件下,均获得了最高的查准率和查全率。为了分析综合查询因子的计算过程,从数据集 B 中提取一张影像在 64 维和 64G 模式下的查询因子曲线进行说明,其结果如图 5 所示。

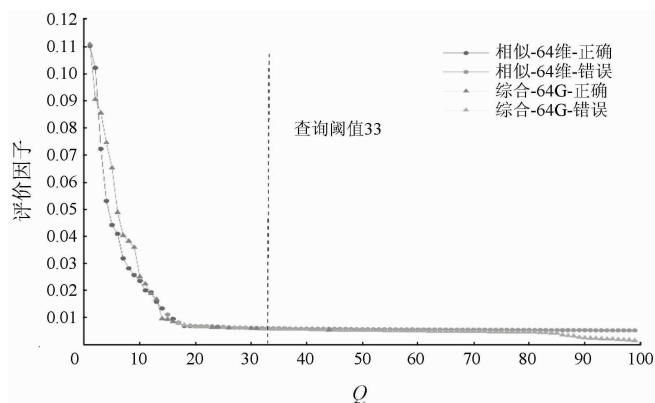


图 5 相似因子与综合因子曲线
Fig. 5 Similarity factors and comprehensive factor curves

图 5 中,相似因子曲线在前端具有明显差异,后端则趋于平坦,不具有明显的可分性。而且,在常规相似因子曲线中,查询曲线后端还包含大量正确的检索影像。而采用本文综合因子的曲线,正确的检索影像大部分集中在查询曲线的前端。

图 6 即为图 5 的部分查询影像,其中图 6(a)为当前查询影像,图 6(b)为采用 64 维传统相似因子的查询影像,在曲线中索引位置为 21,按照传统相似因子判定依据,图 6(a)和(b)将会组成待匹配像对进入匹配环节。然而,从影像内容分析,虽然两者之间存在大量相似的林木区域,但却不具有重叠区,因此,内容相似的影像并非一定具有重叠区。通过计算综合查询因子,图 6(b)所代表的影像已不在综合查询曲线内,而原本不在传统查询曲线内的图 6(c)影像进入了综合查询曲线。因此,综合查询因子不仅可以有效剔除错误像对,还能将遗漏的部分影像纳入影像匹配环节。此外,通过式(7)计算得到的查询阈值,将大量错误的影像剔除匹配环节,进一步提高了查询过程的查准率。



(a) 当前查询影像 (b) 采用 64 维传统相似因子的查询影像 (c) 遗漏的待匹配影像

图 6 部分查询影像
Fig. 6 Partial query images

与正射类型数据不同,图 4 中的倾斜影像数据 C,D 和 E 呈现超高的查准率和超低的查全率,且查全率随着特征维度降低不降反升。通过分析实验数据发现,其原因主要在于数据类型的差异。通常情况下,正射摄影方式获取的 UAV 数据,其影像重叠度一般不超过 50,即一个地物点可在 50 张影像上

成像。而倾斜影像不同,其影像重叠度一般超过 200。此时,将 Q 设置为 100 将使查询出的大部分影像都为正确的匹配影像,即查询结果表现为超高的查准率。而大量的正确匹配像对由于查询深度限制并未提取出来,从而造成了超低的查全率。

表 2 中依次设置 Q 为 100,200 和 300 时,对数

表 2 不同查询深度的查准率和查全率
Tab. 2 Precision and recall of different query depths (%)

数据	查询深度 Q	查准率				查全率			
		128 维	96 维	64 维	64G	128 维	96 维	64 维	64G
数据 C	100	88.66	87.95	86.78	90.91	23.82	25.79	27.60	23.92
	200	82.91	79.26	75.77	83.01	46.41	48.96	50.56	44.26
	300	74.43	64.91	60.27	75.49	63.67	61.39	62.87	64.49
数据 D	100	99.22	99.03	98.81	99.79	29.76	33.44	37.03	31.98
	200	91.47	88.87	86.75	95.30	60.96	64.27	68.20	63.34
	300	77.83	72.94	68.20	83.99	78.42	78.91	80.76	82.01
数据 E	100	73.36	76.64	73.76	86.11	52.52	56.84	60.43	55.93
	200	35.64	44.86	46.14	59.40	55.72	69.19	77.56	78.02
	300	29.06	35.00	33.09	43.74	66.21	79.23	84.34	85.77

据 C,D,E 的查询精度进行统计,随着查询深度的增加,倾斜类型数据的查准率逐渐降低,而查全率得到大幅提升,并且本文方法均取得最高的查准率和查全率。因此,针对倾斜类型的查询深度设置应不小于 200。

查询过程中会产生很多重复查询像对,而最终进入匹配环节的则是剔除重复像对之后的像对列表。因此,为了综合评价本文方法效率,特将匹配像

对提取耗时归入影像匹配环节。此外,为便于比较分析,分别实现了基于脚印图的影像检索^[7]、基于 128 维传统影像检索^[17]以及本文顾及地理空间信息的 64 维特征影像检索 3 种匹配像对提取方法,针对倾斜影像数据 C,D,E,其查询深度统一设置为 200,稀疏重建采用增量式重建,统计结果如表 3 所示。采用本文方法进行稀疏重建的结果如图 7 所示。

表 3 稀疏重建的效率、完整性及精度

Tab. 3 Efficiency, integrity and accuracy of sparse reconstruction

数据	方法	词汇树/min	索引/min	检索/min	双像匹配/min	稀疏重建/min	连接点	成功数	总数	投影误差/像素
数据 A	脚印图	—	—	—	4.20	28.81	362 822	1 591	1 591	0.70
	128 维	5.34	1.70	1.38	6.66	30.36	332 736	1 581	1 659	0.70
	64G	3.55	1.17	2.12	1.26	28.24	364 783	1 591	1 659	0.70
数据 B	脚印图	—	—	—	4.64	5.17	121 527	465	465	0.70
	128 维	5.29	0.49	0.31	1.54	4.94	119 183	465	465	0.70
	64G	3.46	0.24	0.35	0.63	5.20	115 048	465	465	0.70
数据 C	脚印图	—	—	—	11.28	61.69	299 041	1 634	2 137	0.72
	128 维	5.12	1.35	1.27	8.12	58.60	323 224	2 100	2 137	0.73
	64G	3.44	0.82	1.52	3.79	58.49	338 049	2 123	2 137	0.72
数据 D	脚印图	—	—	—	16.67	103.17	367 247	2 728	2 728	0.75
	128 维	5.23	0.94	0.62	26.01	105.41	367 238	2 728	2 728	0.75
	64G	3.44	0.46	1.54	7.35	103.38	375 990	2 728	2 728	0.75
数据 E	脚印图	—	—	—	83.50	3 096.26	1 342 650	8 585	8 585	0.71
	128 维	11.80	11.89	9.92	61.10	2 169.59	1 344 968	8 585	8 585	0.71
	64G	7.20	9.07	11.56	17.89	2 085.62	1 488 024	8 585	8 585	0.70

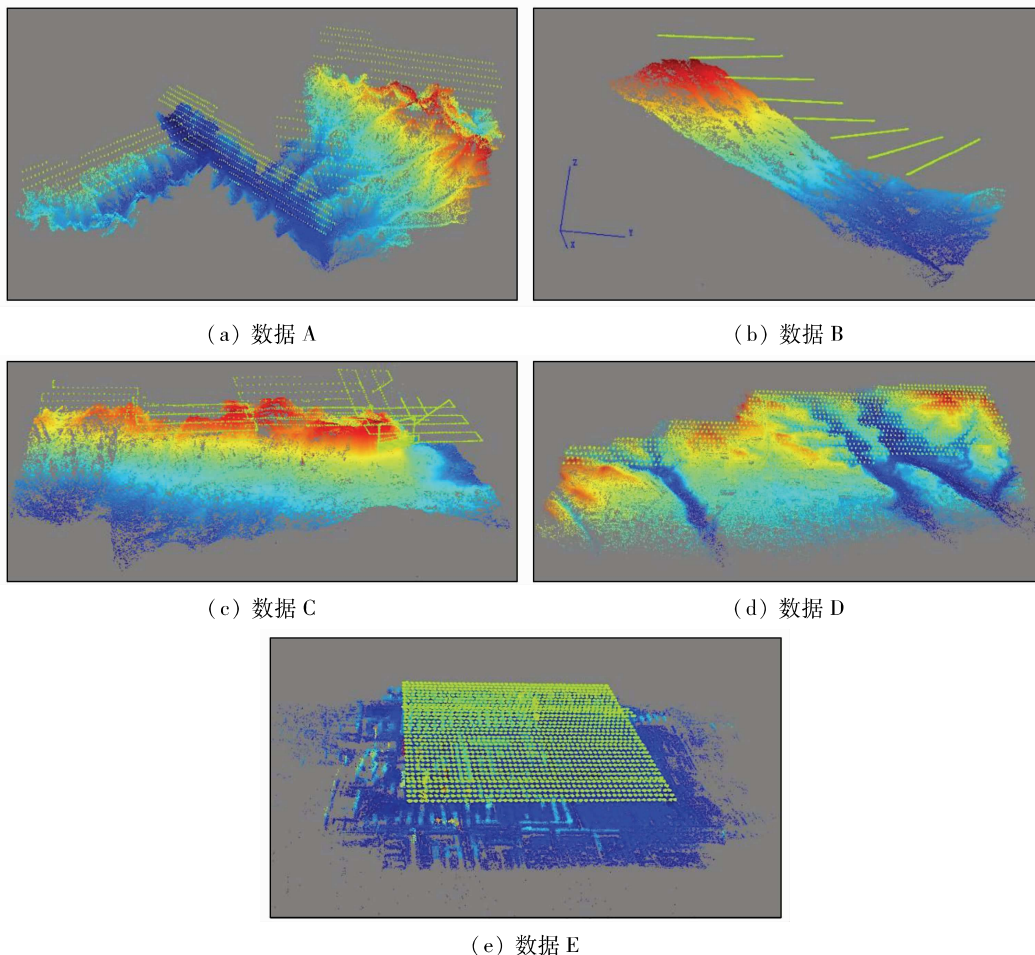


图 7 本文方法提取像对的稀疏重建结果

Fig. 7 Sparse reconstruction results of image pairs extracted by the proposed method

为了综合评价不同方法之间的匹配效率,本文将构建词汇树、构建索引、影像检索及双像匹配耗时相加,统称为匹配时间。从表3中可知,针对正射影像数据A和B而言,传统的脚印图法匹配效率最高。而128维特征检索方法效率最低,且稳定性也不足。比如数据A中,相较于其他2种方式,128维特征的稀疏重建结果丢失了10幅影像,而丢失的影像基本都处于影像纹理匮乏的测区边缘区域,说明单纯依靠纹理特征的相似性判定像对是否具有重叠区并不严密。而针对倾斜类型的数据C,D,E而言,本文方法的效率、稳定性最高。如图7(c)所示,数据C由于地形条件限制,其航线设计比较混乱,且包含了正射、贴近2种摄影方式。传统的脚印图方法没法准确估算贴近摄影方式的脚印图,所以最终获取的稀疏重建结果丢片比较严重,且主要集中在测区右上角,即贴近方式获取的影像区域,而128维、64G检索方法则成功将大部分贴近方式获取的影像连接成功。此外,本文的64G模式获得了最完整的稀疏重建结果。当数据量增多时,本文方法的匹配效率优势更加明显。与脚印图方法相比,数据C,D,E的匹配效率分别提升了15.16%,23.27%,45.25%,而与传统128维检索方法相比,其匹配效率分别提升了39.66%,61.00%,51.73%。

综上所述,本文方法的优势可归纳为以下3点:第一,适应性最高。不需要传感器的先验知识,对场地类型及数据获取方式也无限制。第二,通过综合查询因子,提高了匹配效率及精度,尤其适合海量UAV数据的匹配像对提取。第三,检索深度与影像类型相关。正射类型的影像数据,查询深度设置为100即可,倾斜类型的数据,查询深度应不小于200。结合处理效率、重建结果的完整性、算法的适应性等结果对比,本文提出的顾及地理空间信息的UAV影像匹配像对提取方法更具优势。

3 结论

本文针对UAV影像匹配像对的提取问题,将高维度的特征降维至低维特征,同时引入影像之间的空间信息构建综合查询因子,并通过计算检索阈值,舍弃检索深度内的无效匹配像对,获得了较高的效率及全面的重建结果。

利用5种不同类型的数据进行实验与分析,结果表明,与前人已提出的脚印图法相比,前2种常规类型的单相机正射影像数据匹配效率并未提高,而后3种多相机倾斜影像数据匹配效率分别提升了15.16%,23.27%和45.25%。与传统128维检索

方法相比,5种数据的效率分别提升了46.29%,38.66%,39.66%,61.00%和51.73%。此外,本文方法仅需影像的空间位置信息,更适合数据量较大的倾斜影像,具有更好的适应性。

本文方法还需改进的地方在于,遗漏了少量正确的待匹配像对,其对三维重建结果的影响还需进一步评定。

参考文献(References):

- [1] 雷添杰,李长春,何孝莹. 无人机航空遥感系统在灾害应急救援中的应用[J]. 自然灾害学报,2011,20(1):178-183.
Lei T J, Li C C, He X Y. Application of aerial remote sensing of pilotless aircraft to disaster emergency rescue[J]. Journal of Natural Disasters, 2011, 20(1): 178-183.
- [2] 李明,张卫龙,范丁元. 城市三维重建中的自动纹理优化方法[J]. 测绘学报,2017,46(3):338-345.
Li M, Zhang W L, Fan D Y. Automatic texture optimization for 3D urban reconstruction[J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(3): 338-345.
- [3] 黄海峰,易武,张国栋,等. 引入小型无人机遥感的滑坡应急治理勘察设计方法[J]. 防灾减灾工程学报,2017,37(1):99-104.
Huang H F, Yi W, Zhang G D, et al. Method of using micro-VAVRS in engineering geological investigation and design for emergency control of landslides[J]. Journal of Disaster Prevention and Mitigation Engineering, 2017, 37(1): 99-104.
- [4] 于丙辰,陈刚,段森然,等. 无人机遥感在大型不可移动文物三维重建中的应用[J]. 测绘通报,2017(5):43-46.
Yu B C, Chen G, Duan M R, et al. UAV RS applied in 3D reconstruction of huge immovable cultural relics[J]. Bulletin of Surveying and Mapping, 2017(5): 43-46.
- [5] 何海清. 低空摄影测量航带重构及数据处理方法研究[J]. 测绘学报,2014,43(40):440.
He H Q. The study of strips reconstruction oriented data processing methodology in low-altitude photogrammetry[J]. Acta Geodaetica et Cartographica Sinica, 2014, 43(4): 440.
- [6] 肖熊武,郭丙轩,李德仁,等. 一种具有仿射不变性的倾斜影像快速匹配方法[J]. 测绘学报,2015,44(4):414-421.
Xiao X W, Guo B X, Li D R, et al. A quick and affine invariance matching method for oblique images[J]. Acta Geodaetica et Cartographica Sinica, 2015, 44(4): 414-421.
- [7] 闫利,费亮,叶志云,等. 大范围倾斜多视影像连接点自动提取的区域网平差法[J]. 测绘学报,2016,45(3):310-317.
Yan L, Fei L, Ye Z Y, et al. Automatic tie-points extraction for triangulation of large-scale oblique multi-view images[J]. Acta Geodaetica et Cartographica Sinica, 2016, 45(3): 310-317.
- [8] David G L. Distinctive image feature from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [9] Morel J M, Yu G. ASIFT: A new framework for fully affine invariant image comparison[J]. SIAM Journal on Imaging Sciences, 2009, 2(2):438-469.
- [10] Wu C. SiftGPU: A GPU implementation of scale invariant feature

- transform(SIFT)[EB/OL]. (2013-03-11)[2021-02-07].
https://github.com/pitzner/SiftGPU.
- [11] Agarwal S, Snavely N, Simon I, et al. Building rome in a day[C]//12th International Conference on Computer Vision. IEEE, 2009: 72-79.
- [12] Snavely N, Seitz S, Szeliski R. Photo tourism: Exploring photo collections in 3D[C]//Special Interest Group on Computer Graphics and Interactive Techniques Conference. ACM, 2006, 25(3): 835-846.
- [13] Johannes L, Schonberger, Frahm J M. Structure-from-motion revisited[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 1-10.
- [14] 任超锋, 彭晓东. 低空无人机影像匹配像对的提取方法[J]. 测绘工程, 2019, 28(6): 36-41.
Ren C F, Peng X D. Extraction method of image pairs for low-altitude UAV image matching[J]. Engineering of Surveying and Mapping, 2019, 28(6): 36-41.
- [15] 李劲澎, 姜挺, 肖达, 等. 基于关系图的无人机影像三维重建[J]. 光学精密工程, 2016, 24(6): 1501-1509.
Li J P, Jiang T, Xiao D, et al. On diagram-based three-dimensional reconstruction of UAV image[J]. Optics and Precision Engineering, 2016, 24(6): 1501-1509.
- [16] 姜三, 许志海, 张峰, 等. 面向无人机倾斜影像的高效 SfM 重建方案[J]. 武汉大学学报(信息科学版), 2019, 44(8): 1153-1160.
Jiang S, Xu Z H, Zhang F, et al. Solution for efficient SfM reconstruction of oblique UAV images[J]. Geomatics and Information Science of Wuhan University, 2019, 44(8): 1153-1160.
- [17] Sivic J, Zisserman A. Video Google: A text retrieval approach to object matching in videos[C]//International Conference on Computer Vision. IEEE, 2003: 1470-1477.
- [18] Jegou H, Douze M, Schmid C. Hamming embedding and weak geometric consistency for large scale image search[C]//European Conference on Computer Vision, 2008: 304-317.
- [19] Schönberger J L, Price T, Sattler T, et al. A vote-and-verify strategy for fast spatial verification in image retrieval[C]//Asian Conference on Computer Vision. Springer, 2016: 321-337.
- [20] Nister D, Stevenis H. Scalable recognition with a vocabulary tree[C]//Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2006: 2161-2168.

A method for extracting match pairs of UAV images considering geospatial information

REN Chaofeng, PU Yuchi, ZHANG Fuqiang

(College of Geological Engineering and Geomatics, Chang'an University, Xi'an 710054, China)

Abstract: To overcome the shortcomings such as poor adaptability, low efficiency, and the demand for prior knowledge in the 3D reconstruction using UAV images, this study proposed a method for extracting match pairs of UAV images considering geospatial information. The steps of this method are stated as follows. Firstly, reduce high-dimensional features of the images to low-dimensional features using the principal component analysis (PCA) method to improve the construction efficiency of the retrieval vocabulary. Secondly, construct a comprehensive retrieval factor by calculating the inverse distance weighting factor between query images to improve the distinguishability between similar images. Finally, discard invalid match pairs by calculating the retrieval threshold to improve the query precision of images. The experimental results show that, compared to the traditional footprint map method and 128-dimensional feature retrieval method, this method enjoys higher processing efficiency and more comprehensive sparse reconstruction results, especially for the massive UAV data.

Keywords: match pairs; bag of visual words; UAV; 3D reconstruction; image retrieval

(责任编辑: 张仙)