

张博伦,杨化超,薛鹏,等.基于WebGIS的三维场景与无人机视频流的融合方法[J].湖南科技大学学报(自然科学版),2025,40(2):71-79.doi:10.13582/j.cnki.1672-9102.2025.02.009

ZHANG B L, YANG H C, XUE P, et al. Fusion Method of 3D Scene and Unmanned Aerial Vehicle Video Stream Based on WebGIS [J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2025, 40(2): 71-79. doi: 10.13582/j.cnki.1672-9102.2025.02.009

基于 WebGIS 的三维场景与无人机视频流的融合方法

张博伦,杨化超*,薛鹏,梅洋

(中国矿业大学 环境与测绘学院,江苏 徐州 211116)

摘要:针对智慧城市应用中的视频融合方向存在融合环境及使用平台的可重复性不高、融合方法及结果缺乏足够精度的地理信息支撑等缺陷,提出一种基于 Cesium 的三维地球引擎框架,利用数字正射影像图(Digital Orthophoto Map, DOM)将视频实时影像与实景三维模型融合的新方案.首先采用无人机获取目标区域的影像数据生成高精度地理空间信息数据(DOM和实景三维模型数据等),然后基于 Cesium 框架在浏览器中构建实景三维场景,最后根据 DOM 所包含的位置信息,使其与无人机视频影像进行匹配定位,基于定位结果将视频嵌入三维场景中实现深度融合.通过试验,对比其他融合方法的视频定位精度和融合效果,验证视频与三维模型的融合精度的优势.

关键词:视频融合;数字正射影像图;实景三维模型;单应性变换;Cesium

中图分类号:TP391.41 文献标志码:A 文章编号:1672-9102(2025)02-0071-09

Fusion Method of 3D Scene and Unmanned Aerial Vehicle Video Stream Based on WebGIS

ZHANG Bolun, YANG Huachao, XUE Peng, MEI Yang

(School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou 211116, China)

Abstract: A novel approach is proposed for fusing real-time video images with realistic 3D models in the direction of video fusion in the context of smart city applications, addressing low repeatability of fusion environment and platform, and lack of sufficient accuracy of geospatial information support for fusion methods and results. The proposed solution is based on the Cesium 3D Earth engine framework and uses the DOM to fuse video and 3D model data. Firstly, high-precision geospatial information data, such as DOM and realistic 3D model data, are generated by using UAV to acquire image data of the target area. Secondly, a realistic 3D scene is constructed in the web browser based on the Cesium framework. Lastly, the position information contained in the DOM is used to match and locate the UAV video image, and the video is embedded into the 3D scene based on the positioning result to achieve deep fusion. Comparative experiments are conducted to evaluate the video positioning accuracy and fusion effect of different fusion methods, and the proposed approach demonstrates significant advantages in terms of video and 3D model fusion accuracy.

Keywords: video fusion; digital orthophoto map; real-world 3D model; homography; Cesium

无人机低空倾斜摄影测量技术因其成本低、效率高等优势,为近景测量提供了巨大便利.利用该技术可获得高空间分辨率、高精度及高现势性的基础地理空间信息产品,如实景三维模型、数字正射影像图(Digital Orthophoto Map, DOM)等.实景三维模型作为其获得的较为重要的地理空间信息产品之一,已经在很多领域得到了广泛的应用.其中,在网络地理信息系统(Web-Geographic Information System, WebGIS)中,作为智慧城市概念的基础数据^[1-2],通过叠加多源异构的其他时空及属性信息,可极大地增强各类数据的利用率、实现数据可视化和增强现实等.因此,多源数据的深度融合成为主要的研究方向之一^[3].

视频影像数据具有高实时性、动态性及易理解等优点,将其与静态的地理空间信息数据进行有效融合,对于增强现实及场景理解具有重要的作用.近年来,国内外学者基于不同的应用平台及技术框架进行了研究. NEUMANN 等^[4]利用多种传感器进行场景模型构建,基于 OpenGL 的动态纹理投影技术实现了一种增强虚拟环境(Augmented Virtual Environment, AVE)系统;SAWHNEY 等^[5]提出基于浸入式模型的可视化系统,通过固定相机的位姿与非实景模型进行融合,为视频监控的应用提供了增强现实的解决方案;刘振东等^[6]基于视频监控数据,并利用监控摄像头之间位置的拓扑关系实现其与三维场景的融合;马原野等^[7]利用 OsgGIS 渲染引擎,将监控视频与 3D 模型相互融合并实现多路视频的实时渲染.上述方法均从不同角度实现了视频影像数据在三维虚拟空间中的融合,但仍存在可视化平台的可重复性不高,静态的视频采集方式和融合方法及结果缺乏足够精度的地理空间数据支撑等缺陷.针对上述问题,本文提出在 WebGIS 的三维地球引擎 Cesium 框架下,无人机动态采集的实时视频影像通过 DOM 赋予其高精度的地理空间数据,将具有空间属性的视频影像与实景三维模型嵌入式融合的新方案,为智慧城市等相关领域的应用提供借鉴.

1 SIFT 算法实时影像定位流程

依靠 Cesium 三维地球引擎框架^[8]实现融合环境和 Web 端浏览器的可视化平台,系统的总体流程如图 1 所示.首先,采用倾斜摄影测量技术生成目标区域的实景三维模型,利用光束法区域网平差理论,校正和拼接目标区域的每一张正射图,得到具有地理坐标的整体 DOM 影像.基于 Cesium 框架构建 Web 浏览器平台下的三维场景,实现实景三维模型的可视化.在此基础上,无人机的图传数据从飞机端发送到遥控器端,再通过安卓设备进行实时渲染视频数据.安卓设备接收到无人机的视频和 POS 数据需要发送到服务器,实现其他 Web 端的访问,在视频流传输至视频服务器后提取关键帧数据.然后,利用尺度不变特征变换(Scale-Invariant Feature Transform, SIFT)匹配算法将关键帧图像与 DOM 影像金字塔匹配定位,定位结果为关键帧影像角点像素在地理坐标系下的坐标.最后,在三维场景中结合无人机位姿信息,对含有定位信息的视频影像进行投影纹理渲染,实现实时影像与实景三维模型的高度融合.

1.1 基于 Cesium 的三维场景构建

作为一个三维地球与地图可视化开源 JavaScript 库, Cesium 可用于地理空间数据可视化与海量数据的高效渲染,具有跨平台、跨浏览器、高效的三维模型加载及纹理渲染效率高等优点.基于 Cesium 的三维场景构建包含 3 个关键步骤:

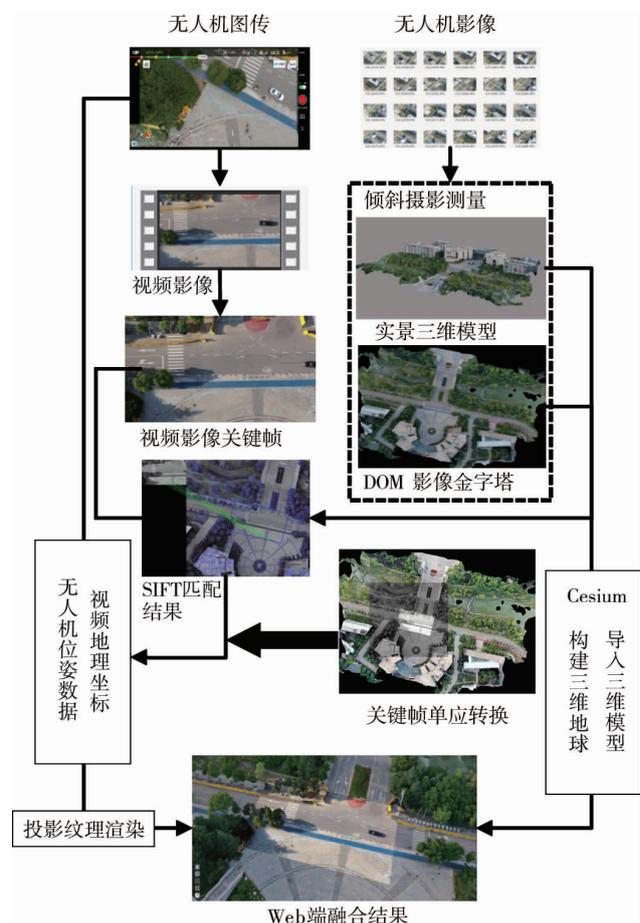


图 1 系统总体流程

(1) 创建并配置 1 个场景对象 (Viewer), 该场景指 1 个模拟地球的椭球体, 叠加数字高程模型 (Digital Elevation Model, DEM) 并加载全球遥感影像; (2) 创建本地模型对象 (Cesium3DTileset), 通过统一资源定位系统 (Uniform Resource Locator, URL) 读取实景三维模型索引文件; (3) 添加模型对象到场景中. 配置三维场景对象如图 2 所示, 此场景基础地图构建参数包括底图影像、3D 地形数据、参考椭球标准、投影类型、初始相机视角、地球背景、天空纹理和光照等, 交互组件包括获取地理坐标、模型量算、图层选择和时间轴等.



图 2 配置三维场景对象

1.2 视频影像的匹配及地理空间定位

1.2.1 视频影像与 POS 数据的获取及传输

视频影像与位姿系统数据 (Position and Orientation System, POS) 的获取是通过对无人机飞控系统 with 通信链路系统进行安卓移动端的二次开发实现的, 2 种数据依据不同的网络协议进行实时数据传输, 数据传输流程如图 3 所示. 服务器端使用公网地址, 主要目的为转发数据, 实现任意设备访问上述数据.

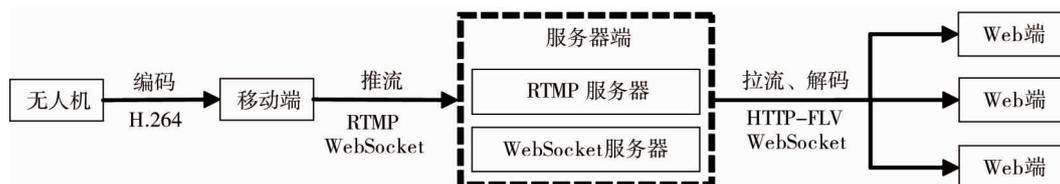


图 3 数据传输流程

视频流数据根据无人机相机从原始格式 (YUV) 编码为 H.264 的格式进行传输, 可以压缩传输数据量, 提高传输效率. FFmpeg 是专用于处理音视频的跨平台开源库, 使用 FFmpeg 对无人机获取的 H.264 视频流进行复用打包为 FLV 格式, 从安卓移动端通过 RTMP 协议将视频流推送到服务器中, 数据经由服务器端转码后, Web 端使用 HTTP-FLV 协议进行拉流^[9], 浏览器实现解码并播放.

POS 数据以字节流的形式, 通过监听无人机的实时位置与姿态信息, 构建 json 格式数据, 使用 WebSocket 协议从移动端发往服务器端再转发至 Web 端. 安卓端的关键代码如下, 可以实现将 POS 数据发送到服务器端.

```
FlightController flightController =
droneHelper.fetchFlightController(); // 获取飞行控制实例
flightController.setStateCallback(flightController -> { // 飞控状态回调函数
String message = (“构建回传数据字符串”);
client.send(message); } }); // 向服务器发送数据
```

1.2.2 视频影像的地理空间匹配定位

DOM 影像与视频影像的关键帧相互匹配, 通过对关键帧单应性变换找到其在 DOM 影像中的准确位置, 即可实现关键帧的地理空间定位, 由无人机采集的高分影像利用光束法区域网平差理论得到具有地理空间属性的 DOM 影像. 在视频的传输过程中, 依据 H.264 编码格式的特点, 其中的 I 帧定义为关键帧, 是一段视频流中数据量最为完整的数据, 依据二者的特点即可实现视频影像的地理空间匹配定位. 在影像匹配的过程中, 首先通过粗略定位确定关键帧影像的范围, 以达到减少匹配面积、提高效率的目的, 然后使用

DOM 金字塔进行匹配.经过试验验证,针对低分辨率 DOM 影像数据,该匹配策略通过由粗到精的分层匹配机制,能够快速定位目标区域并降低计算量,在保证匹配准确率的同时显著提升效率,可以有效平衡时间成本与精度需求.

为提高关键帧的定位效率,提出由粗到精的视频影像地理空间匹配定位.具体步骤如下:

1) 粗略定位

(1) 确定粗略定位区域的中心点.将 POS 坐标转换到 DOM 投影坐标系下,在 DOM 影像中确定无人机的像素坐标即粗略定位区域的中心.

(2) 确定定位区域的范围.粗略定位区域是以无人机位置为圆心的圆外切矩形,圆的半径 R 的计算公式如式(1)和式(2)所示,参数几何关系如图 4 所示.

$$R = H \tan(\omega + \alpha); \quad (1)$$

$$\alpha = \arctan\left(\frac{w/2}{f}\right). \quad (2)$$

式中: H 为航高; ω 为俯仰角; α 为垂直方向视场角的二分之一; w 为相机成像元件的宽度; f 为相机焦距.

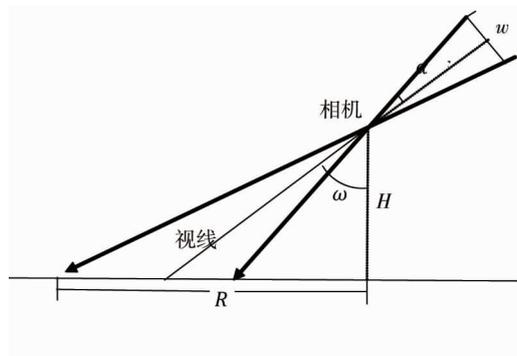


图 4 几何关系

2) 基于 DOM 金字塔的关键帧快速精确定位

影像地理空间位置如图 5 所示.以粗略定位区域的 DOM 金字塔影像作为匹配对象,设各角点在 DOM 上的像素坐标为 $P_i(u_i, v_i)$, $i = 1, 2, 3, 4$. O -XYZ 为地面测量坐标系, $S(x, y, z, \varphi, \omega, \kappa)$ 为某摄站 S 的初始外方位元素.

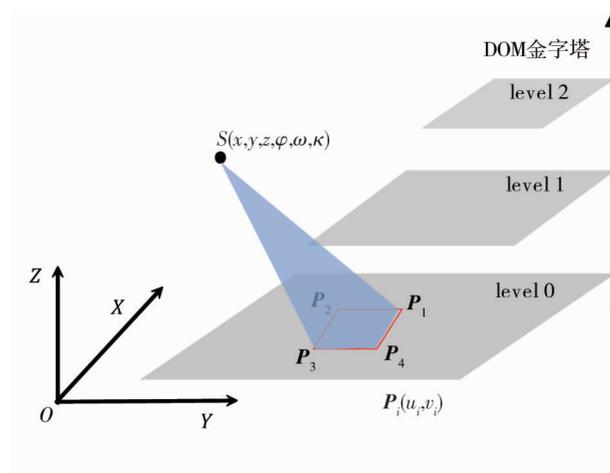


图 5 影像地理空间位置

基于影像金字塔的视频影像快速精确定位的步骤如下:

(1) 按式(3)和式(4)采用像元平均法构建 DOM 影像金字塔.

$$g(x, y) = \sum_{k, l} f(x + k, y + l) h(k, l); \quad (3)$$

$$h(3,3) = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}. \quad (4)$$

式中: $g(x,y)$ 为均值滤波结果; $f(x,y)$ 为原图像; $h(k,l)$ 为滤波算子,算子尺寸为 3×3 时, $k,l \in \{-1,0,1\}$.

(2)由粗到精匹配.在 DOM 金字塔的基础上,利用 SIFT 算法采用基于影像金字塔的由粗到精的匹配策略在视频影像和金字塔影像间匹配特征点^[10-11].匹配结束后,记获得的匹配点坐标为 $Q_j(u'_j, v'_j)$ 和 $P_j(u_j, v_j)$, $j = 1, 2, 3, \dots, m(m \geq 4)$.

(3)基于单应性变换的视频影像空间定位.单应性变换是由一个视角平面的点映射到另一个视角平面内对应点的透视投影转换.根据相机内外参数组成的模型矩阵,视频影像像素 $Q(u', v')$ 与正射影像像素 $P(u, v)$ 通过 M 相联系,通过第(2)步获得的匹配特征点按式(5)计算 M ,然后通过计算得到的 M 同样利用式(5)将视频影像角点坐标 $Q(u', v')$ 转换到 DOM 影像中获得 $P(u, v)$ 坐标,进一步基于 DOM 内的视频交点坐标 $P(u, v)$ 通过双线性插值获得其地理空间位置.

$$P_i = MQ_i. \quad (5)$$

式中: $P_i = (u_i \ v_i \ 1)^T$; $Q_i = (u'_i \ v'_i \ 1)^T$; M 为单应性变换矩阵.

1.3 视频影像与实景三维场景融合

基于 Cesium 的投影纹理映射融合技术是按照 WebGL 纹理渲染的深度计算与矩阵转换等方法进行投影变换与纹理渲染^[12-13].纹理渲染是对视频影像的像素坐标进行渲染,要将其投影到三维场景中,需要构建视锥体如图 6 所示.根据相机图像的传感器尺寸和图像比率,设置水平张角和垂直张角,根据无人机传输的 POS 数据,视点坐标为无人机 POS 坐标,3 个姿态角一一对应,保持实时更新位置和姿态.投影面的角点 ABCD 坐标使用 1.2 节匹配定位的结果予以修正.当视频流数据传入浏览器时,视频将在投影平面进行播放.

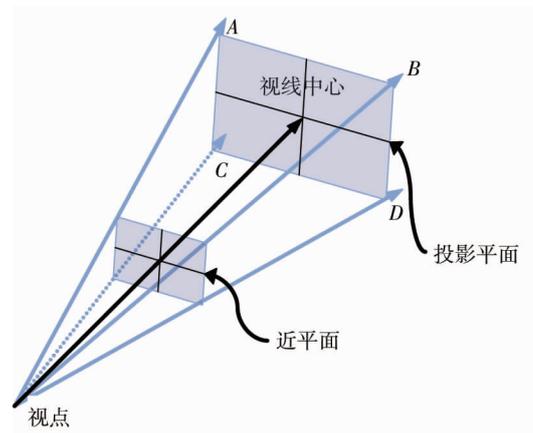


图 6 视锥体

2 试验结果与分析

2.1 试验数据获取与精度分析

2.1.1 DOM 和实景三维模型数据构建

采用六旋翼无人机搭载五镜头倾斜摄影相机,照片分辨率为 $7\ 952 \times 5\ 304$.传感尺寸宽为 $35.9\ \text{mm}$,高为 $24\ \text{mm}$,焦距为 $40\ \text{mm}$.获取了目标区域地面分辨率约为 $2.5\ \text{cm}$ 的倾斜数字影像,影像数量为 200 张,航向及旁向重叠度为 80% ,区域面积约为 $1.89\ \text{km}^2$.通过三维建模软件 ContextCapture 构建目标区域的 DOM 和实景三维模型.

构建的三维地球场景具有真实的地理坐标系,这就要求 DOM 与实景三维模型和真实世界相比具有较低的误差,并且 DOM 影像的分辨率也间接决定视频影像的定位精度,所构建的 DOM 影像的地面分辨率为 $3\ \text{cm}$.为验证二者的误差,在区域内随机设立 5 个检验点,通过载波相位差分技术(Real Time Kinematic, RTK)测量各点获取的真实地理坐标,对比 DOM 影像与实景三维模型上相同点位的坐标值,两者的平面误差在 $0 \sim 3\ \text{cm}$.

2.1.2 视频影像与 POS 数据获取

采用四旋翼无人机获取分辨率为 $1\ 920 \times 1\ 080$ 、帧率为 30 帧/s的视频影像,相机等效焦距为 $28.620\ 3\ \text{mm}$.使用无人机二次开发组件开发具有视频推流功能的安卓 APP,无人机的飞控数据通过遥控器传入安卓设备,在 APP 中将向服务器发送无人机的 POS 数据和视频数据.视频流编码格式为 H.264,安卓设备与服务器的传输协议选择实时消息传输协议(Real Time Messaging Protocol, RTMP),这是专门用来进行实时数据通信的网络协议.RTMP URL 的格式与 HTTP URL 相似,格式如下:

```
rtmp://<host>[<port>]/<appName>/<streamName>[? param1 = value1][&param2 = value2]
```

利用 RTMP 协议进行安卓设备与服务器之间的数据交换.POS 数据使用 WebSocket 协议进行传输,此时,服务器获取视频影像与 POS 数据.

2.2 试验平台搭建

硬件环境: Intel(R) Core(TM) i9-10900K 3.70 GHz, 4×16 GB 内存, 显卡型号为 NVIDIA GeForce RTX 3060, 12 GB 显存. 软件环境: Windows 10 Professional x64, VS2019, Python 3.6.4 以及由 Cesium, GDAL 3.1.4 和 OpenCV-python 等组成的开源库.

2.3 三维场景构建

所构建的三维场景主体是基于 WebGIS 平台坐标系统为 EPSG:4326 (WGS 84) 球面地理投影的椭球体来模拟真正的地球空间. 所构建的实景三维模型与 DOM 影像都是在 WGS 84 地理坐标系中的数据. 采用 Visual Studio 2019 编写 HTML, JavaScript 和 CSS 代码, 用到的工具库有 Cesium.JS, flv.min.js 等. 主要实现: (1) 在椭球体表面, 以天地图公布的卫星影像作为底图; (2) 地形数据为美国地质调查局发布的 GTOPO30 全球数字高程模型; (3) 添加构建的目标实景三维模型; (4) 常用插件包括时间轴、更换底图和鼠标悬停显示三维地理空间坐标等. 三维场景构建结果如图 7 所示, 部分伪代码如下:

```
var viewer = new Cesium.Viewer("CesiumContainer", { // 创建场景
    terrainProvider: new Cesium.CesiumTerrainProvider({ // 在场景中添加全球数字高程模型
        url: "数字高程模型索引文件本地路径"}));
var imageryProvider = new Cesium.webMapTileServiceImageryProvider({ // 创建底图图层
    url: "天地图影像网络地址",
    credit: "天地图"});
viewer.imageryLayers.addImageryProvider(imageryProvider); // 添加卫星影像到场景中
const tileset = await Cesium.Cesium3DTileset.fromUrl( // 创建三维模型
    "实景三维模型的本地网络路径", {
    show: true, // 显示模型
    shadows: ShadowMode.ENABLED, // 开启阴影模式
    level of detail: true, }); // 模型细节完全加载
scene.primitives.add(tileset); // 添加三维模型到场景中
```



图7 三维场景构建结果

2.4 无人机视频与三维场景实时融合处理

2.4.1 无人机视频帧定位

视频帧定位是将关键帧图像与包含空间坐标的 DOM 影像金字塔进行匹配定位, 得到图像在 DOM 中的相对位置, 从而得到图像的地理坐标. 为了提高匹配效率, 使用由粗到精的匹配定位方法, 主要用到的库有 GDAL 3.1.4 和 OpenCV-Python 等. 首先根据安卓设备传输的 POS 数据, 对粗略区域的 DOM 影像金字塔进行裁剪, 不同等级的影像同时与关键帧图像进行匹配, 当得到匹配结果时, 停止本次匹配. 然后将匹配结果发送到浏览器, 以修正无人机视频融合位置.

关键帧匹配定位过程如图 8 所示, 主要包括: (1) SIFT 算法对 2 幅影像进行特征点提取与描述; (2) 利用快速最近邻搜索算法、K 最邻近算法 (K-Nearest Neighbor) 和随机样本一致性算法 (RANSAC) 对 2 组特

征点进行匹配与筛选;(3)多组对应特征点计算单应性矩阵;(4)关键帧图像经过单应性变化,根据角点的行列号检索得到空间坐标.

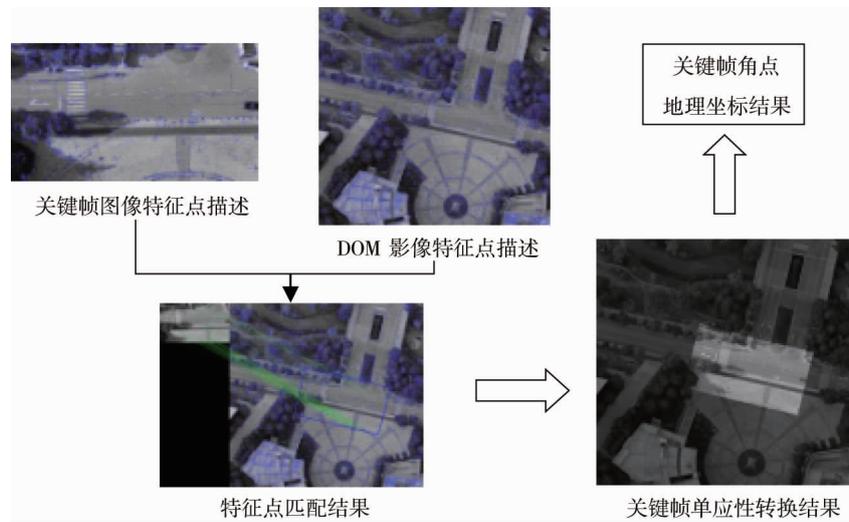


图 8 关键帧匹配定位过程

2.4.2 视频与三维场景融合

在 2.3 节构建的三维场景的基础上构建视锥体模型,将视频数据投影在模型表面,实现视频与三维场景的融合.视锥体的主要属性包括视点坐标、航向角、俯仰角、投影内容和投影面角点坐标等.从服务器传入的 POS 及无人机影像数据不断更新视锥体属性与投影内容.移动中的无人机影像融合结果如图 9 所示.图 9 中的图片按照顺序,每 3 张为一组,各组图片来自不同角度的连续移动的无人机.

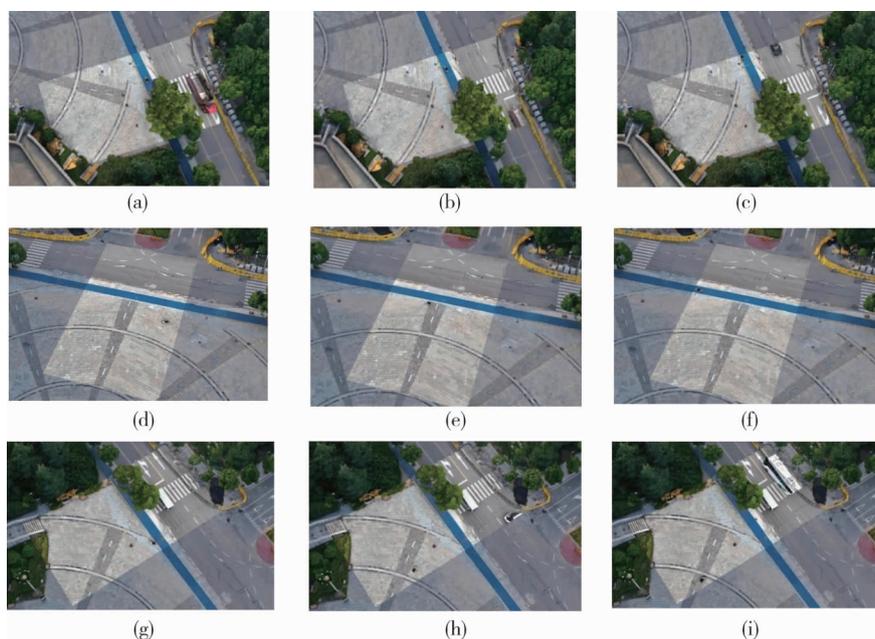


图 9 融合结果

利用 HTML5 提供的 WebSocket 协议接口,实现浏览器与服务器之间的双向通信,通过此协议接收来自服务器的无人机 POS 数据,伪代码如下:

```
WebSocket = new WebSocket("服务器 ip 及接口") //连接服务器,创建对话
WebSocket.onmessage = function (event) { //连接对话执行的方法
    json = json.parse(event.data) //json 数据解包
    video.position = json.position // 更新视锥体的视点坐标
```

```

video.heading = json.Compass //更新视锥体的航向角度
video.points = json.points //更新视锥体的投影面角点坐标
video.pitch = json.GimbalPitch //更新视锥体的俯仰角度

```

视频流数据拉取使用 HTTP-FLV 协议,视频播放器使用 flv.js 库.在 HTML 中创建 video 标签并设置视频来源,视锥体投影内容为 video 标签.播放器与视锥体构建伪代码如下:

```

var flvPlayer = window.flvjs.createPlayer({ //创建 FLV 拉流播放器
  type: "FLV", //视频复用类型
  url: HTTP-FLV URL, //视频流数据服务器地址
  enableStashBuffer: true}) //是否启用缓存区
flvPlayer.attachMediaElement(videoElement) //绑定 video 标签
var video = new view.graphic.Video3D({
  position: 视点坐标,
  container: videoElement,
  angle: ((Math.atan(6.4 / 28.6203) * 180) / Math.PI) * 2, //相机水平张角推算
  angle2: ((Math.atan(4.8 * (7 / 8) / 28.6203) * 180) / Math.PI) * 2, //相机垂直张角推算
  heading: 相机航向角,
  pitch: 相机俯仰角})

```

2.4.3 精度分析

在 DOM 影像和视频影像间隔帧上选取若干具有明显特征的同名地物点,分别测量其平面坐标 (X, Y) 和 (X', Y') 并计算二者的差值,统计二者差值的中误差(式(6))作为视频关键帧空间定位的精度指标,并将本文的视频空间定位算法与文献[14]的算法进行精度对比.类似地,通过在视频影像和实景三维模型上选取同名地物点,按式(6)统计中误差作为视频融合精度的评价指标.精度统计结果如图 10 所示,视频融合效果对比如图 11 所示.由图 10 可知:本文方法优于文献[14]方法,其中第 20 帧的中误差最大为 0.150 m,最大中误差不超过 0.545 m.由图 11 可知:视频影像空间定位后的视频融合结果与文献[15]利用无人机位姿信息进行视频融合的结果相比,前者的融合精度优于后者.

$$m_{XY} = \pm \sqrt{\sum_{i=1}^n (\Delta X_i^2 + \Delta Y_i^2) / n}. \quad (6)$$

式中: $\Delta X_i = X_i - X'_i$; $\Delta Y_i = Y_i - Y'_i$; n 为某帧选取的检查点个数.

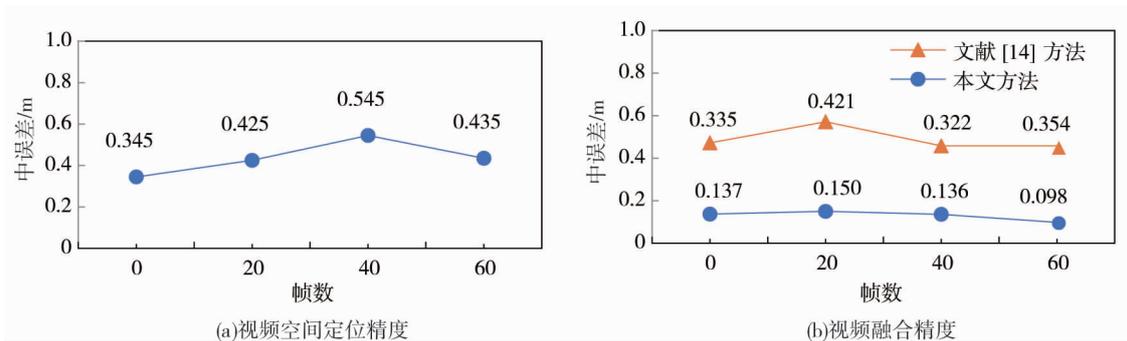


图 10 精度统计



图 11 视频融合效果对比

3 结论

1) 本文方法在 3 cm 级 DOM 分辨率支撑下可实现 0.150 m (RMS) 的地面空间定位精度,其单应性变换模型经验证具有参数稳定性,适用于动态飞行姿态下的实时地理配准。

2) 与传统监控视频融合方法相比,本文方法在核心机理与应用范式上存在显著差异。二者虽均需解决视频-模型空间配准问题,但传统方法多依赖固定摄像头与局部坐标系(误差>1 m),通过人工标定实现静态场景融合,而本文基于无人机动态视频流,以 DOM 地理空间基准(3 cm 分辨率)与 Cesium 的 WGS84 椭球面坐标系构建全局映射,并依托 WebGIS 架构突破传统 C/S 系统的封闭性限制,支持跨平台的实时融合。

3) 随着低空经济的发展,会有大量低空视角的视频数据产生,这类数据与地理信息的深度融合将有助于视频数据的进一步应用,其目标是通过 WebGIS 平台实现“动态感知-空间建模-实时服务”的技术闭环,为数字孪生城市、元宇宙等新兴领域提供低成本、高精度的时空数据基底。

参考文献:

- [1] 朱庆. 三维 GIS 及其在智慧城市中的应用[J]. 地球信息科学学报, 2014, 16(2): 151-157.
- [2] 李德仁, 邵振峰, 杨小敏. 从数字城市到智慧城市的理论与实践[J]. 地理空间信息, 2011, 9(6): 1-5.
- [3] JEANSOULIN R. Multi-source geo-information fusion in transition: a summer 2019 snapshot[J]. ISPRS International Journal of Geo-Information, 2019, 8(8): 330.
- [4] NEUMANN U, YOU S, HU J H, et al. Visualizing Reality in an Augmented Virtual Environment[J]. Presence, 13(2): 222-233.
- [5] SAWHNEY H S, ARPA A, KUMAR R, et al. Video flashlights-Real time rendering of multiple videos for immersive model visualization[C]//Proceedings of the 13th Eurographics Workshop on Rendering. 2002: 157-168.
- [6] 刘振东, 戴昭鑫, 李成名, 等. 三维 GIS 场景与多路视频融合的对象快速确定法[J]. 测绘学报, 2020, 49(5): 632-643.
- [7] 马原野, 赵刚, 何彬. 3D GIS 与多视频融合系统的设计与实现[J]. 计算机应用与软件, 2012, 29(6): 109-112.
- [8] 杨菁, 陈冰凌, 王文鹏, 等. 基于 Cesium 的三维可视化场景建设及发布技术的研究[J]. 测绘通报, 2021(增刊1): 50-53.
- [9] 朱倩. JavaScript 在 HTML5 视频和音频控制中的应用[J]. 无线互联科技, 2021, 18(18): 89-91.
- [10] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [11] MUJA M, LOWE D G. Fast approximate nearest neighbors with automatic algorithm configuration[C]//Proceedings of the Fourth International Conference on Computer Vision Theory and Applications. Lisboa, Portugal, February 5-8, 2009: 331-340.
- [12] 宋碧波, 卢小平, 孙晓峰, 等. OpenGL 虚拟相机模拟摄影测量二次成像[J]. 测绘通报, 2012(7): 21-23.
- [13] KOUICHI M, RODGER L. WebGL 编程指南[M]. 谢光磊, 译. 北京: 电子工业出版社, 2014.
- [14] 赵维淞, 钱建国, 汤圣君, 等. 一种移动视频与地理场景的融合方法[J]. 测绘通报, 2020(12): 11-16.
- [15] 王磊, 周鑫鑫, 吴长彬. 无人机实时视频与三维地理场景融合[J]. 测绘通报, 2021(12): 33-37.