



哈尔滨工业大学

Harbin Institute of Technology

无人机摄像及应用

课程结课报告

组长 : 刘宇轩 25L0212 2025112195

成员 1 : 何家震 25L0502 2025110903

成员 2 : 刘一诺 25L0205 2025111917

成员 3 : 张芮嘉 25L0209 2025111819

成员 4 : 石艾琳 25L0207 2025112232

成员 5 : 齐露雅 25L0216 2025111932

指导教师 : 李顺龙, 郭亚朋

一、报告简介

1.1 报告背景

中国古代木桥、石桥和铁索桥都长时间保持世界领先水平，在世界桥梁发展史上曾占据重要地位，为世人所公认。21 世纪伴随着科技水平日新月异的迅猛发展，新材料、新结构、新技术、新工艺不断涌现，必将进一步推动桥梁工程建设水平快速发展[1]。现如今，我国拥有 100 余万座桥梁，桥梁数量世界第一。其中公路桥梁 80 余万座，铁路桥梁 20 余万座，均位于世界第一。对于社会大众，他们普遍关注的焦点之一主要在于桥梁工程施工建设和运营维护过程中，施工装备、人员、病害等常见目标导致的安全事故。



图 1:我国著名桥梁

与此同时，无人机被广泛应用于摄影测量和遥感应用当中，其飞行性能灵活而稳定、具有良好的开发改装潜力、便于人们操作与控制。在此之前，进行倾斜摄影的时候需要使用多个镜头，无人机影像采集技术出现以后，进行倾斜影像采集的时候凭借单镜头相机就可以达成目标。无人机摄像的背景源于相关技术的成熟与普及。它凭借独特的空中视角，突破了传统拍摄的物理限制。从最初的军事用途，迅速扩展到影视航拍、测绘勘察、农业植保及应急救援等民用领域。这种技术以成本低、效率高、灵活性强的特点，正深刻变革着我们观察和改造世界的方式。因此，无人机摄像不仅仅是一种新的拍摄工具，更是一种颠覆性的生产力工具和创意媒介，正在深刻改变着我们观察、理解和改造世界的方式。



多旋翼平台

固定翼平台

直升机平台

浮空翼平台

1.2 报告目的及任务

基于面向桥梁工程安全重大需求，响应新工科建设学科交叉、新旧融合需要，精准定位交通强国战略中智慧土木的发展要求，将无人机图像采集技术、深度学习、人工智能算法等技术与桥梁工程方面有机结合，可以为全面排查桥梁工程隐患，保障桥梁工程安全提供一条重要途径。因此，正确掌握无人机测绘技术，努力将自己打造为桥梁工程交叉融合的人才，响应国家号召，投身到国家需要的领域去我们责无旁贷。报告主要阐述图像拼接原理以及三维重建原理两方面技术，并应用与二校区校园二维图像拼接及校图书馆的三维重建中，制定实验方案，得出结论并进行结果的对比与分析。**总结来说**，无人机摄影的目的是作为人类视觉和行动的延伸，以实现更高效、安全、全面的认知与改造世界。而其具体**任务**，则是将这一目的落地，转化为在测绘、巡检、农业、影视、救援等众多行业中不可替代的解决方案。

二、报告内容

2.1 图像拼接原理

2.1.1 图像拼接的概念

图像拼接[2]是指将一组具有重叠区域的序列图像通过相应的图像处理技术，最终融合成一幅包含完整序列拍摄图像信息的宽视角、高分辨率新图像的技术。换句话说，就是把图像集中的所有图像映射到一个坐标系下，使图像集中的图像重叠部分能够映射到一起，最后创建出可以一张显示图像集拍摄区域的全部信息的全景图像。

2.1.2 图像拼接的基本流程

图像拼接的基本流程包含三个核心步骤。首先是图像预处理，进行畸变校正与滤波去噪，为后续处理提供优质输入。接着是关键环节——图像配准：通过特征提取（SIFT 等算法）寻找关键点，进行特征匹配后，采用 RANSAC 方法剔除误匹配点，最终估计出描述图像间变换关系的单应性矩阵。最后是图像融合阶段，将配准后的图像重投影至统一坐标系，通过颜色均衡处理曝光差异，并运用多频段融合等先进技术消除接缝，生成无缝的高质量全景图或大画幅图像，完成整个拼接过程。

2.1.3 图像配准方法

1. 基于特征的配准方法

这是目前最主流、应用最广泛的方法。其核心思想是从图像中提取显著且稳定的特征，然后通过特征匹配来建立图像间的对应关系。该方法对光照变化、遮挡和噪声具有一定的鲁棒性。其流程通常包括：

特征提取：从每幅图像中检测出关键点（如角点、边缘点）并生成特征描述符。常用的算法有 SIFT（尺度不变特征变换）、SURF（加速稳健特征）和 ORB（定向 FAST 和旋转 BRIEF）。其中，SIFT 特征在尺度和旋转不变性方面表现优异，被视为标杆。

特征匹配：通过计算特征描述符之间的相似度（如欧氏距离），为一张图像中的特征点在另一张图像中寻找最相似的对应点。

误匹配剔除：匹配过程中会产生大量错误对应关系。通常采用 RANSAC（随机抽样一致）算法来鲁棒地估计变换模型参数，并同时剔除不符合模型的误匹配点对。

变换模型估计：利用正确的匹配点对，计算最终的变换矩阵（最常用的是单应性矩阵 H ），该矩阵定义了从一幅图像到另一幅图像的透视变换关系。

2. 基于区域的配准方法

该方法不提取显式特征，而是直接利用图像像素的灰度信息进行配准。它通过在参考图像上滑动模板，并在待配准图像中寻找与模板最相似的区域来完成匹配。相似性度量通常使用互相关、互信息等准则。该方法在特征稀疏或纹理缺失的场景下可能效果更好，但其计算量通常较大，且对光照变化和旋转较为敏感。

2.1.4 图像融合方法

图像融合是图像拼接流程中保证最终输出质量的关键步骤。其核心任务是将经过几何配准的多幅图像合成为一幅视觉一致、无缝的高质量全景图。该过程需要解决因拍摄条件差异导致的颜色、亮度不一致以及重叠区域的拼接痕迹两大核心问题。

融合过程首先进行**图像重投影与坐标变换**，依据图像配准环节计算出的精确变换矩阵（如单应性矩阵），将所有输入图像的像素映射到统一的输出全景坐标系下，形成一个初步对齐但存在重叠的中间结果。

随后，**颜色均衡**处理被应用于解决图像间的曝光与色彩差异。通过分析重叠区域的像素统计信息，采用如直方图匹配、增益补偿或基于最优缝合线的颜色传递等方法，对每一幅图像进行全局或局部调整，使所有图像单元的亮度和色调协调统一，避免出现明显的边界。

最后，也是最核心的环节是接缝消除。直接合成会在重叠区域产生明显的接缝。目前最先进且广泛应用的技术是多频段融合。该算法将每幅图像分解为不同空间频率的子带序列：低频子带主要包含颜色和光照信息，高频子带则捕获边缘和纹理细节。融合时，在重叠区域对低频子带进行加权平均或渐变过渡，以平滑地融合颜色与光照；而对高频子带则采用“最佳缝合线”策略，优先选择梯度变化最小的路径来保留最多的细节信息。最终将所有融合后的子带重构，即可得到一幅既无缝过渡又最大限度保持细节清晰度的全景图像。

2.2 三维重建原理

2.2.1 三维重建简介

三维重建（3D Reconstruction）是通过技术手段将现实中的三维物体或场景，在虚拟计算机环境中还原并构建数学模型的过程，本质是“照相机逆操作”——照相机将三维现实映射为二维图像，而三维重建则从二维图像中提取信息，复现三维空间形态。

当前主流的三维重建方法可分为三类：

1. 基于图像的三维重建：通过多角度相机拍摄获取图像，依托计算机视觉技术（如特征匹配、几何计算）构建三维模型，无需专用扫描设备，适配无人机航拍等场景。
2. 传统几何建模：采用 Nurbs、布尔建模等技术，人工定义几何参数构建模型，精度依赖人工设计，更适用于规则几何体的精细化建模。
3. 基于点云的三维重建：通过激光扫描、结构光扫描等设备直接获取物体表面点云数据，再对数据进行处理生成模型，精度高但对硬件设备要求较高。

三维重建技术的分类可按“数据采集方式”进一步细化，具体如下：

采集类型	技术方向	具体方法
接触式	数据测量	连续式数据测量、触发式数据测量、三坐标测量
非接触式	主动视觉法	激光扫描法、结构光法、阴影法、TOF 技术、雷达技术、Kinect 技术
非接触式	被动视觉法 - 按相机数目	单目视觉法、双目视觉法、多目视觉法
非接触式	被动视觉法 - 按匹配方法	区域视觉法、特征视觉法、因子分解法、运动恢复结构法、多视图几何法
非接触式	被动视觉法 - 按应用方法	统计学习法、机器学习法、神经网络法、深度学习与语义法

表 1：三维重建技术的分类

2.2.2 三维重建核心流程

基于图像的三维重建（适配无人机航拍场景）需遵循标准化流程，确保从数据到模型的精度与完整性，具体步骤如下：

1. 数据采集：使用无人机、相机或激光扫描设备，从不同角度、位置采集目标场景（如建筑物、校园区域）的图像数据。若为无人机航拍，需规划航线确保图像重叠率（通常不低于 60%），避免信息缺失。
2. 图像预处理：对采集的原始图像进行校正（消除镜头畸变）、去噪（采用高斯滤波、中值滤波等）、色彩统一（校正不同光照下的色差）处理，为后续特征提取提供高质量图像数据。
3. 特征提取：从预处理后的图像中提取具有唯一性和稳定性的特征点与特征描述符，常用算法包括 SIFT（尺度不变特征变换）、SURF（加速稳健特征）、ORB（快速旋转不变特征），这些特征可表征图像中的角点、边缘等关键信息，是图像匹配的基础。
4. 特征匹配：通过 FLANN（快速最近邻搜索库）、RANSAC（随机抽样一致性）等算法，对不同图像中的特征点进行匹配，确定同一物理点在不同图像中的对应位置，筛选出正确匹配对，剔除误匹配点。
5. 初始模型建立：基于匹配的特征点对，利用三角剖分原理或几何约束关系（如基础矩阵、本质矩阵计算），估计目标场景的初始三维结构，生成粗略的点云或简化多边形模型，确定物体的大致空间形态。
6. 稠密重建：在初始模型基础上，引入更多图像数据，采用多视图立体重建技术（如立体匹配、光流法），补充大量细节点云，提升模型的密度与细节丰富度，使模型更贴近真实场景。
7. 网格生成：通过点云三角化（如泊松重建、Alpha 形状算法）或体素网格化，将离散的点云数据转换为连续的网格面，形成可编辑、可视化的几何模型框架。
8. 纹理映射：将原始图像中的纹理信息（如建筑物墙面颜色、材质纹理）投影到重建的网格模型上，通过纹理融合、平滑处理，增强模型的真实感，使其视觉效果与现实场景一致。
9. 优化与后处理：对重建模型进行噪声去除（剔除孤立点、平滑网格）、空洞填补（修复数据缺失区域）、几何形状调整（校正模型变形），确保模型的精度与完整性。
10. 可视化与应用：将优化后的模型导出至 CAD、MeshLab 等三维软件，用于虚拟现实（VR）、增强现实（AR）、建筑设计验证、基础设施巡检等场景，实现模型的实际应用价值。

三、实验方法及过程

3.1.1 基于 OpenCV 的开源实现

(1). OpenCV 工具基础

OpenCV（Open Source Computer Vision Library）是基于 BSD 许可的开源计算机视觉库，支持 Linux、Windows、Android、Mac OS 多操作系统，以轻量级、高效性为核心优势——核心代码由 C 函数与 C++ 类构成，同时提供 Python、MATLAB 等语言接口，可直接调用图像处理、特征提取、几何计算等通用算法，无需重复开发底层模块，是实现三维重建（尤其是基于图像的重建）的

高效工具。

(2). 基于 OpenCV 的三维重建流程

依托 OpenCV 实现三维重建，需结合 “图像预处理 - 特征匹配 - 相机参数估计 - 点云生成 - 模型优化” 全链路，具体步骤如下：

1. 输入图像与参数初始化

导入采集的多角度图像（如无人机航拍的建筑物图像），统一调整图像尺寸（如缩放至相同分辨率），减少后续计算量；

初始化相机内参（如焦距、主点坐标），若未知内参，可通过 OpenCV 的 `calibrateCamera` 函数进行相机标定，获取内参矩阵与畸变系数，消除镜头畸变对重建的影响。

2. 特征提取与匹配

调用 OpenCV 的 `cv2.SIFT_create()`、`cv2.ORB_create()` 函数，提取每张图像的特征点（`KeyPoint`）与特征描述符（`Descriptor`）；

使用 `cv2.FlannBasedMatcher` 或 `cv2.BFMatcher` 进行特征匹配，得到初始匹配对；

通过 `cv2.findHomography` 结合 RANSAC 算法，筛选出鲁棒性强的匹配对（剔除误匹配），计算图像间的单应矩阵，确定图像间的空间几何关系。

3. 相机姿态估计与基础矩阵计算

基于匹配的特征点对，调用 `cv2.findFundamentalMat` 函数计算图像对的基础矩阵，再结合相机内参矩阵，通过 `cv2.computeEssentialMat` 得到本质矩阵；

对本质矩阵进行 SVD（奇异值分解），求解相机的旋转矩阵（ R ）与平移向量（ t ），确定不同图像对应的相机姿态，建立图像间的三维空间约束。

4. 三角剖分生成点云

利用已求得的相机内参、旋转矩阵与平移向量，调用 `cv2.triangulatePoints` 函数，对匹配的特征点进行三角剖分，计算特征点在三维空间中的坐标，生成初始稀疏点云；

通过 `cv2.convertPointsFromHomogeneous` 函数，将齐次坐标转换为三维笛卡尔坐标，得到可直接处理的点云数据。

5. 点云优化与稠密重建辅助

对初始稀疏点云进行滤波处理：使用 `cv2.filter2D` 进行高斯滤波，剔除孤立噪声点；通过距离阈值筛选，保留场景内有效点云；

若需提升点云密度，可结合 OpenCV 的 `calcOpticalFlowPyrLK`（光流法），追踪相邻图像中的像素运动，补充更多特征点，生成稠密点云。

6. 网格生成与可视化

将优化后的点云数据导出至 open3d 等库（OpenCV 需结合第三方库实现网格生成），通过三角化算法构建网格模型；

调用 OpenCV 的 `cv2.drawKeypoints`、`cv2.drawMatches` 函数，可视化特征点与匹配结果；结合 `matplotlib` 或 `open3d` 的可视化接口，展示三维点云与网格模型，验证重建效果。

7. 误差补偿与结果输出

计算重建模型的重投影误差：将三维点云投影回二维图像，通过 `cv2.projectPoints` 函数得到投影点，与原始图像特征点对比，若误差超出阈值，重新优化相机参数与点云坐标；

将最终的三维模型导出为 PLY、OBJ 等通用格式，用于后续纹理映射、模型

编辑或实际应用场景。

(3). 关键技术要点

相机标定准确性：若无人机相机无预设内参，需提前使用棋盘格标定板，通过 cv2.calibrateCamera 获取精准内参，否则会导致相机姿态估计误差，影响重建精度；

特征匹配鲁棒性：优先选择 ORB 算法（兼顾速度与旋转不变性），配合 RANSAC 算法剔除误匹配，确保图像间几何关系计算的准确性；

点云滤波处理：稀疏点云需进行噪声去除，避免孤立点影响网格生成，可通过设置“点云邻域密度阈值”或“距离均值阈值”实现筛选。

3.1.2 基于 cc 的闭源实现

Context Capture 是一款基于数字影像处理技术的实景三维建模软件，能够自动生成高精度三维模型，广泛应用于智慧城市、数字孪生、BIM 设计等领域。它可以实现：

多源数据支持：可处理来自智能手机、无人机、专业航摄仪等设备的影像数据，支持倾斜摄影测量技术；

高精度建模：通过自动空中三角测量和几何优化算法，生成厘米级精度的三维模型，模型细节可实现毫米级还原；

有 GIS 兼容性：支持生成符合 GIS 标准的数据格式（如 OBJ、OSGB、Esri I3S 等），可直接用于地理空间分析；

进行多场景应用：适用于建筑规划、城市监测、考古研究等场景，支持红外多光谱影像和激光点云数据混合建模。

技术特点

采用 GPU 加速计算和多核处理技术，支持大规模数据集处理；

通过地面控制点或 GPS 标签实现高精度地理定位，可测量坐标、面积和体积；

提供可视化编辑工具，允许用户调整相机参数、姿态修正等。

适用场景：

包括近景物体（如文物修复）、中小型空间（如仓储、办公区）、城市级建模等，覆盖从厘米级到公里级的多种尺度需求。

“Context Capture 是基于图形运算单元的快速三维场景运算软件，仅需少量人工干预，依靠连续的二维影像就能还原出达到原始影像像素分辨率的实景三维模型 [2-3]。利用该软件生产三维模型具有工作量小、效率高、成果效果好等优点。” [3]

3.2 实验过程

本次实验的硬件平台为：笔记本电脑等

软件平台为：Context Capture 等 手机软件 DJI Pilot、DJI Fly 等

本文所采用的实验数据场景是依靠大疆四旋翼无人机，摄像头镜头方向朝下。拍摄的是哈工大二校区图书馆及其周边的一部分，一共有 88 张建模图片

名称	创建日期	分辨率	类型	大小
DJI_20250929154725_0095_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	26,812 KB
DJI_20250929154730_0096_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	25,876 KB
DJI_20250929154737_0097_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	25,760 KB
DJI_20250929154741_0098_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	26,032 KB
DJI_20250929154744_0099_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	26,084 KB
DJI_20250929154746_0100_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	26,408 KB
DJI_20250929154749_0101_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	26,288 KB
DJI_20250929154752_0102_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	26,152 KB
DJI_20250929154755_0103_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	25,852 KB
DJI_20250929154758_0104_D.JPG	2025/9/29 16:40	8064 x 4536	JPG 文件	26,352 KB
DJI_20250929154800_0105_D.JPG	2025/9/29 16:40	8064 x 4536	JPG 文件	26,648 KB

名称	创建日期	分辨率	类型	大小
DJI_20250929155033_0172_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	27,580 KB
DJI_20250929155035_0173_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	27,556 KB
DJI_20250929155037_0174_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	27,528 KB
DJI_20250929155040_0175_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	27,564 KB
DJI_20250929155042_0176_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	27,504 KB
DJI_20250929155045_0177_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	27,296 KB
DJI_20250929155047_0178_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	27,076 KB
DJI_20250929155049_0179_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	26,848 KB
DJI_20250929155051_0180_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	26,572 KB
DJI_20250929155053_0181_D.JPG	2025/9/29 16:41	8064 x 4536	JPG 文件	26,540 KB
DJI_20250929155055_0182_D.JPG	2025/9/29 16:39	8064 x 4536	JPG 文件	26,220 KB

图 3,4:部分建模图片

移动设备上提前下载 DJI Fly APP 和 DJI poilt APP 并注册账号。使用 DJI poilt 可以进行最佳测绘和建模。打开手机软件 DJI poilt 并同时连接无人机，选择航线模式。DJI poilt 有多种航线模式，可以使用矩形双重网格用于图书馆的三维建模。但本次使用了 DJI Fly 的自动焦点环绕和手动拍照。设置好参数，若飞行区域合适，即规划的航迹符合无人机性能（电量）要求，点击 START，准备进行规划航迹飞行。待飞行任务执行完后，无人机拍摄的照片存储在无人机 SD 卡中，将拍摄照片下载到手机和笔记本电脑检查拍摄效果。检查无误存储准备下一步使用。打开 Context Capture 新建任务，导入无人机拍的图片，调整分辨率，开始重建。

四、结果分析

三维重建

三维重建，利用无人机拍摄的图书馆图片直接计算生成三维建模。通过旋转图书馆建模，我们可以清晰地看到图书馆的样貌和与其他设施的位置关系。但缺点是有一小部分分辨率低且发生了扭曲现象，导致没有完全还原样貌。整体来讲，效果还是不错的。



(a) 二区图书馆三维重建



(b) 二区图书馆三维重建图



(c) 二区图书馆三维重建图

图书馆空间立体感较强。软件较好，笔记本处理时间较短(约 40 分钟)，并且效果不错。

五、课程感想及建议

有幸参加了无人机摄像及应用这门课程，我们不但收获了平日里很少涉及的无人机技术的专业知识，而且能够以实际操作体会无人机的独特魅力。

首先，在知识层面，有老师谆谆善诱引导我们主动思考和开拓未知之境、耐心细致地解答和处理我们的疑难与问题、以小组讨论报告的方式敦促我们集思广益、每次课程开始前提问带领我们回顾与复习从而帮助我们加深记忆……我们不但认识到了无人机比我们以为的广泛的多的应用场景，也深入学习了需要创新与理解的航迹规划与图像拼接等，这些都是不可多得的精神财富，让我们受益匪浅。

其次，在实践层面，一堂理论一堂实践的课堂规划让我们能够即时巩固新习得的知识。HIT 为这门课程提供了先进的无人机设备，让我们得以学以致用、有效分配课堂时间规划任务与活动，在乐趣中培养我们的实操能力、每个学生都有自己的独立操作时间，让课程中的每一个人都满载而归。

最后，我们衷心感谢开设这堂课程的学校，以及悉心教导的老师与热心的学长，让我们在轻松愉快的氛围中对无人机摄像及应用有了一个全面的认识。同时，为了让这门课得到更多人的青睐，我建议提供更多的线上理论资源，吸引学生；同时在授课中适当加入更多吸引人的实例和当下热点，紧跟时代的脚步；将相关成果进行公众号推文及宣传，让更多人有渠道认识到这门课程的价值，从而吸引学生们踊跃参加。

六、参考文献

- [1] 中国桥梁领域开拓国际工程市场的现状与对策[J]. 中外公路, 2005, 25(5):140 - 146.
- [2] 蔡健. 图像拼接技术综述[J]. 计算机光盘软件与应用, 2011(11):87-87.
- [3] 杜金莉. 基于 ContextCapture 的三维建模研究[J]. 地理空间信息, 2021, 19(5): 53-54. DOI: 10.3969/j.issn.1672-4623.2021.05.014.