# 1绪论

# 1．1研究背景和研究意义

## 1.1.1研究背景

人类通过双眼来探索与发现世界，在接收外部信息的方式中，有不到三成来自于听觉、触觉、嗅觉等感受器官，而超过七成、最丰富、最复杂的信息则通过视觉进行感知的。计算机视觉便是一种探索给计算机装备眼睛（摄像头）与大脑（算法）的技术，以使计算机能够自主独立的控制行为、解决问题，同时感知、理解、分析外部环境。

20世纪60年代，计算机视觉得到了最初的发展，该阶段的研究重心主要体现在如何从二维图像中恢复出如立方体、圆柱体等立体化的三维形状，解释各个物体的空间位置关系。1982年David Marr从信息处理的角度对数学、神经生理学、计算机图形学等学科的研究成果进行了归纳总结，并在此基础上提出了一系列计算机视觉理论，经典Marr视觉信息处理过程如图所示。得益于这个完整明确的理论体系，计算机视觉得到了蓬勃的发展，它的核心思想是从二维图像恢复三维结构。

近年来，图像的三维重建在计算机视觉中发挥了很大的作用，并且在质量和性能上有了较大的提升。其主要应用是自动地对于难以建模的对象建模，加快了图像运用的建模过程。这种技术需要处理大量的数据，可以使用于室内和室外的场景, 而不受控制的环境通常影响室外场景，如密集建筑群，或者复杂的原始森林等。对于这些场景，虚拟现实和计算机模拟可以被用来分析工作环境和工作难度等方面，三维图像重建技术本身被视为一个生成3D模型的技术。快速有效完整重建类似于雕塑三维物体目标的三维模型成为目前的研究方向。由连续图像的三维重建主要是指从二维图像序列中的获取物体的信息并进行三维重建。然而,这个领域并没有引起人们足够的重视，因此本文将对三维重建的具体原理以及改善展开讨论。

当前在一些工业场景中，需要对其中场景进行测量，主要包括场景的高度，面积，体积，甚至于温度，湿度分布等，传统方式中主要采用雷达扫描场景的方式来进行，但考虑到该雷达本身成本较高，受场景的限制也很大，在室外环境或者大尺度环境下，就难以发挥作用。现在更多采用计算机视觉的方式来解决这些问题，可以根据三维重建的点云结果测量以上所描述的几何特性，这样做只需要结合摄像机和测量算法即可实现，可以很简易的复现在多种场合中。

当前基于无人机自主飞行采集图像数据进行三维重建以及体积测量的方案

本项目研究目标是解决目前煤堆体积测量过程中存在的成本高、效率低、工作强度大、自动化程度低等主要问题。本项目根据燃料煤堆场半露天基础设施及煤堆成形分布的结构特点，探讨利用无人机控制及视觉智能感知技术进行煤堆体积及温度场快速测量的方法。通过研究无人机的定位、路径规划、飞控及防撞等技术，实现无人机对煤堆的全覆盖、高效率视觉扫描；通过研究视觉三维重建及体积计算技术，实现对煤堆存量的准确估计；通过融合可见光和热成像相机的数据，实现对煤堆温度场的测量，以确保火力发电厂对煤堆存量实现科学准确管理。

## 1.1.2研究意义

本文的研究目标是让无人机能够在全封闭或者半封闭的环境中，完全基于视觉的方法完成自主飞行任务，并且在飞行的过程中，采集待测物体的实拍图像，进行三维重建，最终根据三维重建的点云结果测算待测物体的体积值。

传统无人机的飞行都需要依靠操作员手动控制，或者是基于GPS定位结果进行巡航，本文提出了一种完全基于纯视觉的方法来进行无人机定位的导航得到的方法，该方法可以在全封闭无GPS的环境中，为无人机提供定位信息，解决了场景限制的问题，并且整个无人机的飞行过程可以完全自动化的运行，并对场景进行图像采集。

从二维图像重建三维立体具有重要的研究价值和潜在经济社会价值，其核心技术是通过运动来恢复结构。三维重建系统在不同的应用领域有着不同的预设条件和技术要求，主要包括医学领域的重建系统，机器人导航相关实时重建系统，工业领域包括3D打印在内的室内高精度重建系统，以及摄影测量领域实景三维重建系统。 三维重建已能提供完整方案，但传统三维重建算法得到的点云结果存在精确度不高，场景无法闭合等问题，本文提出一种结合SLAM结果的三维重建方法，拟生成一个高精度，强鲁棒的点云地图。

有关实景物体三维测量，传统方法常采用激光雷达，声波等方案，这些方案成本较高，且对场景有限制要求要求，本文提出了一种基于三维重建后的点云结果进行体积估计的方法，可以极大降低成本且能够在多重场合下复用算法。

因此基于无人机自主飞行采集图像数据的三维重建和体积测量有着十分重要的工程意义，并且在很多场合下，整体流程可靠性和可行性都得到充分验证。

1.2国内外研究现状

## 1.2.1 SLAM的研究现状

即时定位与地图构建是让机器人在未知环境中持续地构建环境地图，并同时在地图中给自己定位。最早的SLAM技术还不是使用视觉的方法，而是使用声波传感器或者激光以及惯性测量单元实现环境建模和自身定位。直到21世纪，Stephen Se等人首次使用图像的特征点实现视觉SLAM[1],之后由Davison使用EKF框架实现了最早的单目实时SLAM系统[2]，奠定了单目系统的基础；Davison在2007年成功实现基于单相机的纯视觉SLAM系统，算法的关键是在线建立2D点到3D点的映射关系，并且使用实时运动模型估计相机的位置[3]；Mur-Artal使用ORB特征点作为地图构建特征点，大幅度降低了点云的数量，并且使用回环检测的方法使定位与建图的精度都大幅提升[4]；随着硬件计算能力和数据储存的提升，提取目标深度信息的技术得到了很大的发展，戚传江等人使用2D slam的解决方案，采用多传感器数据融合的方法，完成多自由度位姿检测[5]，拓展了SLAM的应用场景；Whelan的实验通过使用体积融合的方法实现了实时大范围的稠密RGB-D的SLAM系统[6]。

近些年来，SLAM的发展更多的开始和机器人领域相结合，Durrant-Whyte和Bailey首先对前20年里SLAM的发展做出了详细的历史回顾。

在两次调查[8]中对SLAM问题的前20年进行了详尽的历史回顾[70]。这些主要包括我们所说的古典时代(1986-2004);在经典时代，SLAM引入了主要的概率公式，包括基于扩展卡尔曼滤波器(EKF)、Rao-Blackwellized粒子滤波器和最大似然估计的方法;此外，它还描述了与效率和健壮数据关联相关的基本挑战。另外两篇优秀的参考文献描述了古典时代的三种主要的SLAM公式，分别是Thrun, Burgard, and Fox[240]和Stachniss等人的章节[234,Ch. 46]。接下来的阶段我们称之为算法分析时代(2004-2015)，在[65]中，Dissanayake等人对此进行了部分阐述。算法分析阶段研究了SLAM的基本性质，包括可观测性、收敛性和一致性。在此期间，稀疏性对于高效的SLAM求解器的关键作用也得到了理解，并开发了主要的开源SLAM库如表所示，回顾了近些年来SLAM领域的发展情况。

应用到目标跟踪领域，单纯点云集还是无法满足要求，因此需要将点云数据语义化，Reiger使用关系树的方法实现物体的语义识别[7]，这项技术对于目标跟踪是很重要的；之后Sarkar在Reiger的研究基础上结合FastSLAM的方法，使得识别速度更快，鲁棒性更强[8]；Zhang, G等人使用基于线条的SLAM算法[9]提高物体识别的准确率，该方法能够对物体的边沿与轮廓进行稳定的识别。

## 1.2.2 三维重建的研究现状

根据图3生成的地图，通过对地图的解析，可以获取每一个二维码的三维

## 1.2.3 基于视觉体积测算的研究现状

根据图3生成的地图，通过对地图的解析，可以获取每一个二维码的三维

## 1.2.4 语义结构化地图的研究现状

根据图3生成的地图，通过对地图的解析，可以获取每一个二维码的三维

# 1.3待解决问题

当前基于无人机自主飞行采集图像数据进行三维重建以及体积测量的方案还存在很多的待解决问题，如：

1. 无人机在完全封闭环境中依靠纯视觉进行定位和建图，难以获取高精度的飞行位姿和地图信息。
2. 基于三维重建算法对场景进行三维重建时，面临整个流程耗时长，输出点云噪音点大，需要对传统三维重建进行提升，以获取高精度强鲁棒性的大尺度地图。
3. 基于三维点云的体积测算，点云中缺少水平面信息，尺度信息，导致无法直接获取到体积真值。

# 1.4主要研究内容和技术路线

## 1.2.1 SLAM研究现状