

媒体与认知 （2022 秋）

Media and Cognition (2022 Autumn)

期末大作业要求及评分细则

1. 双目相机搭建-标定-深度计算 （50 分）

1.1 任务描述

- a. 利用两个单目相机，设计并构建双目相机系统
 - a1. 设计相机固定支架并交付 3D 打印（考虑结构强度，材料用量，固定方式，基线选择；baseline：两相机不轻易掰弯、不晃动的固定，走线不大角度弯曲）（10 分）；
 - a2. 使用棋盘格图标定配准双目相机系统，需展示标定的过程，相机矩阵，与校正后的左右视角图（此步骤为标准化过程，不区分 baseline 与额外得分项）（15 分）；
- b. 使用搭建的双目相机系统完成深度图计算 （25 分）

1.2 考察方式

- a. 形式为项目答辩形式，展示双目相机实物图片，标定过程，与深度计算结果。
 - a1. 对于标定过程，请简述实际操作步骤与每步的中间结果；
 - a2. 对于深度计算结果，请以视频形式，展示双目相机对实际场景的实时深度计算结果。
- b. 考察重点
 - b1. 计算速度：深度计算算法的运行速度，是否可以以视频帧率实时流畅输出深度图；
 - b2. 重建质量：输出深度图是否边缘清晰，平滑，是否有明显空洞；
 - b3. 重建精度：是否可以准确计算场景中细小物体与细节部分的深度；
 - b4. 鲁棒性：对于双目深度估计中常见的有难度场景，如无纹理区域、重复纹理区域、遮挡区域 等，是否有较好的应对。

* 对于 b2、b3、b4，若算法在这些方面有优势，请在答辩时结果展示视频中着重标记。

1.3 其他注意事项

若需修改支架设计、提交新设计方案、零件换新等，需提供邮寄地址与联系方式。

2. 复杂场景目标检测 （50 分）

2.1 任务描述

- a. 实现目标检测算法：

探索并实现至少两类目标检测算法，并分别对两种输入图（双目相机自行拍摄的 RGBD、PANDA 数据集的亿像素图像），探讨分析两类算法在不同场景下的表现。

* 两类算法指的是 two-stage、one-stage、anchor-free、无监督等目标检测算法。

* 算法可自行设计实现，也可采用现有算法及预训练模型。

2.2 考察方式

- a. 形式为项目答辩形式，展示算法流程、目标检测结果。

- a1. 简述所实现两类算法的基本原理和特点；
 - a2. 用双目相机捕获的 RGB 图像作为输入，分别测试所实现两类算法的效果，要求场景内物体较为多样（至少含有三类物体），要求算法准确检出目标，输出边界框参数，并分析两类算法的表现；将双目相机计算出的深度图 Depth 也作为检测算法的输入，分析探讨深度图的引入对检测带来的影响；
 - a3. 用 PANDA 数据集的亿像素图像作为输入，要求检测出图中的物体，输出边界框参数，分析和比较两类算法的表现，具体格式及评测指标详见如下链接 <https://cloud.tsinghua.edu.cn/d/4c89f841250a44458627/>。
 - a4. 对所实现的每一类算法，分析和比较该算法在 a2、a3 两种输入图像上的表现。
- b. 考察重点
- b1. 检测精度：算法输出边界框的精度和类别精度；
 - b2. 检测速度：目标检测算法的运行速度；
 - b3. 鲁棒性：对于复杂场景，如：目标数量众多、遮挡明显、目标尺度变化大等情形的检测效果。

注：要求各组上传项目全部代码，并保证可读性且可复现。

联系方式：

- Zhihao Xu 徐智昊, xu-zh20@mails.tsinghua.edu.cn, 15802245434
- Guangyu Wang 王光宇, wanggy21@mails.tsinghua.edu.cn, 15336429761
- Haiyang Ying 营海洋, yinghy21@mails.tsinghua.edu.cn, 15013026001