基于 RGB-D 图像的三维物体识别算法的研究与实现

李勇奇　　指导教师：陈启军

2018 年 3 月 19 日

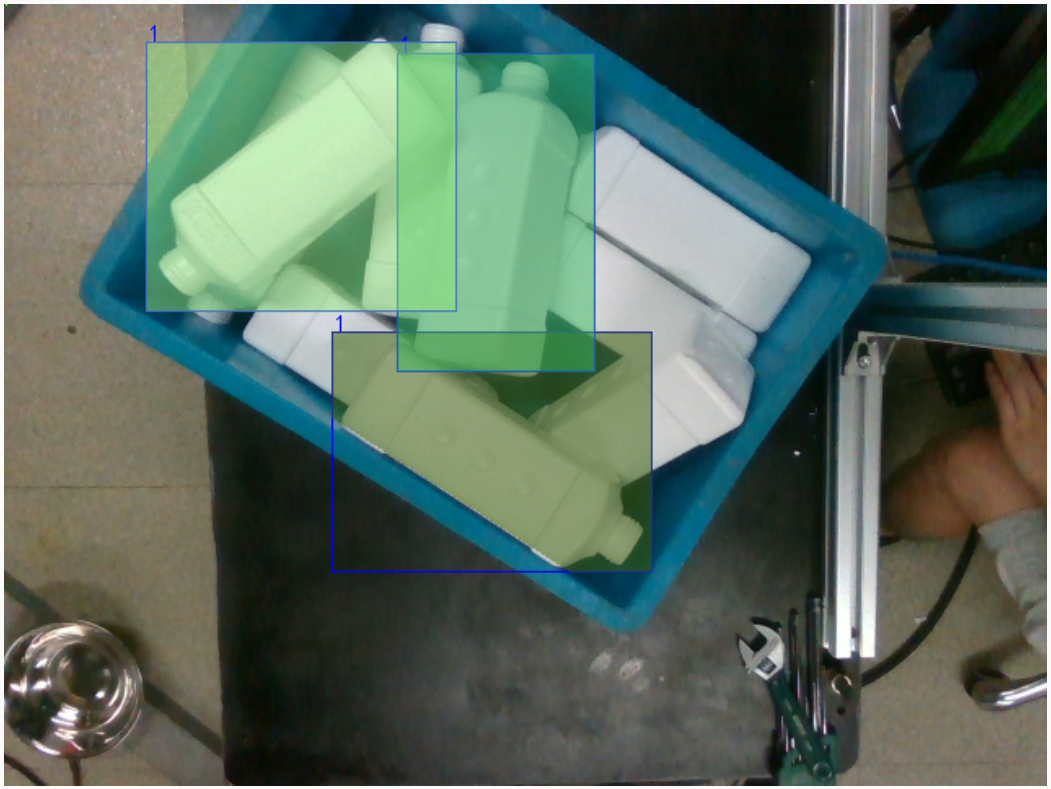
同济大学，控制科学与工程系

Table of contents

Introduction

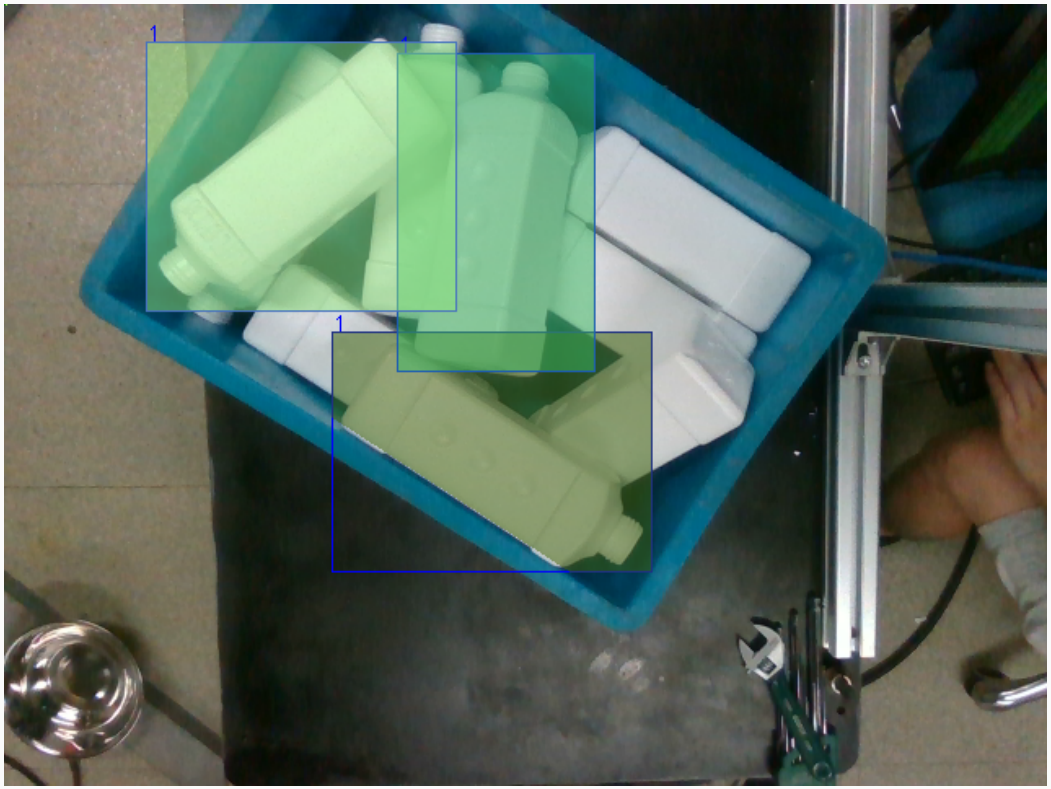
Problem

2D Detection

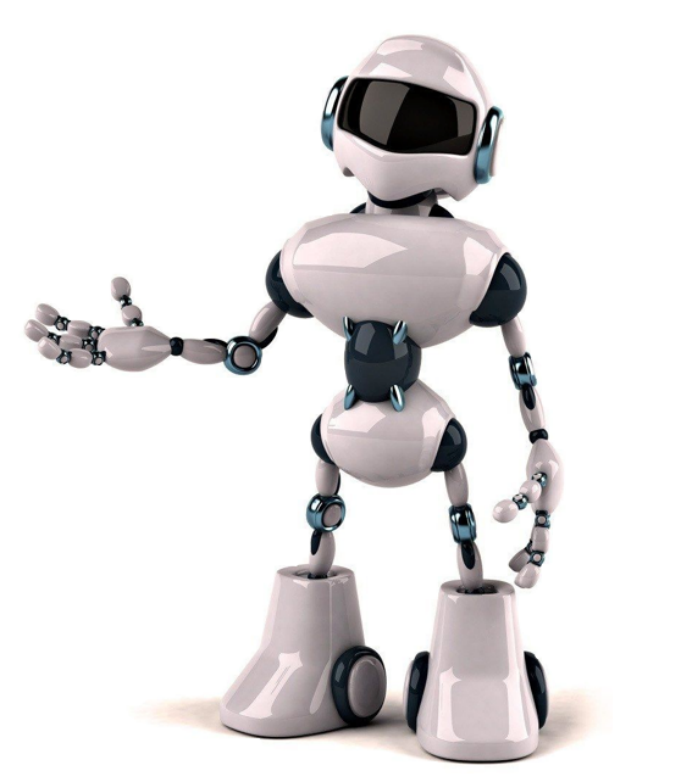


Problem

2D Detection

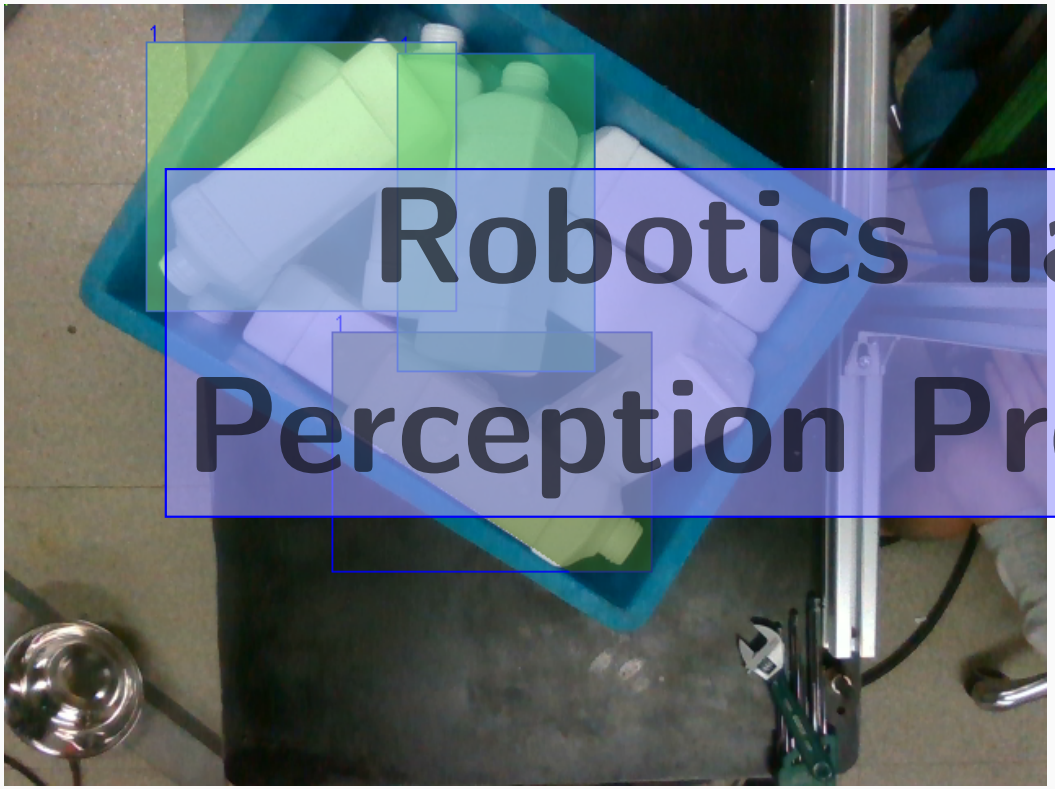


Where to pick?

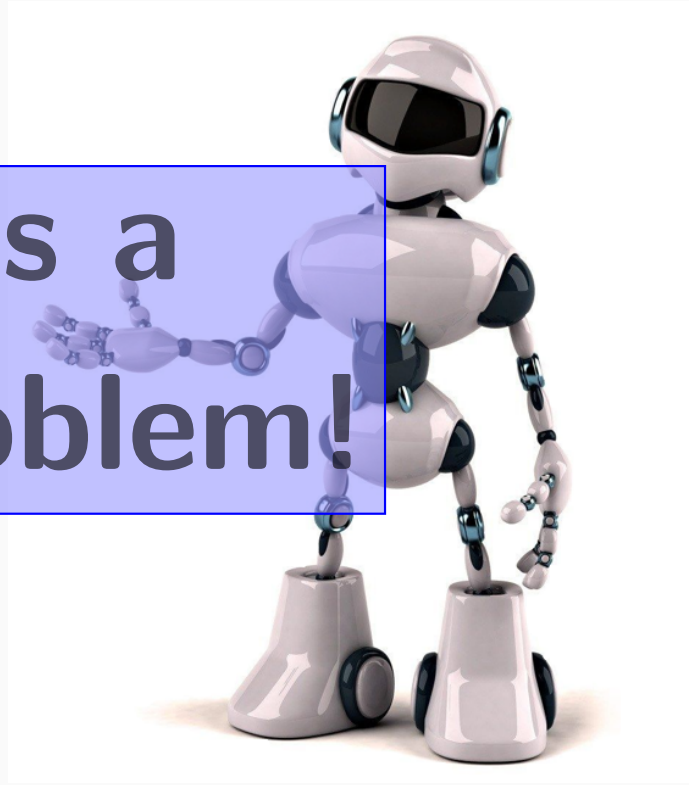


Problem

2D Detection



Where to pick?

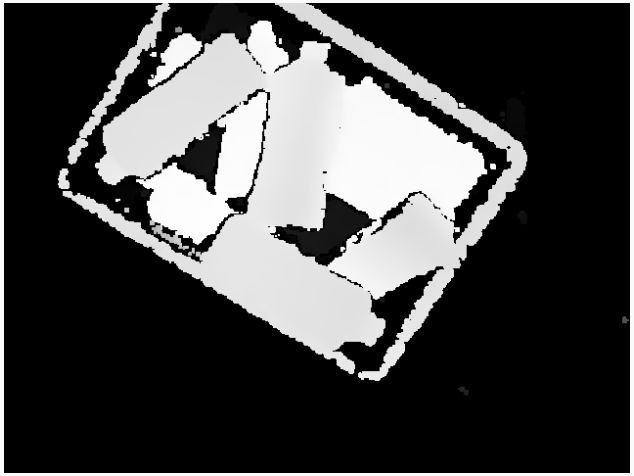


Robotics has a Perception Problem!

3D Detection

3D Detection

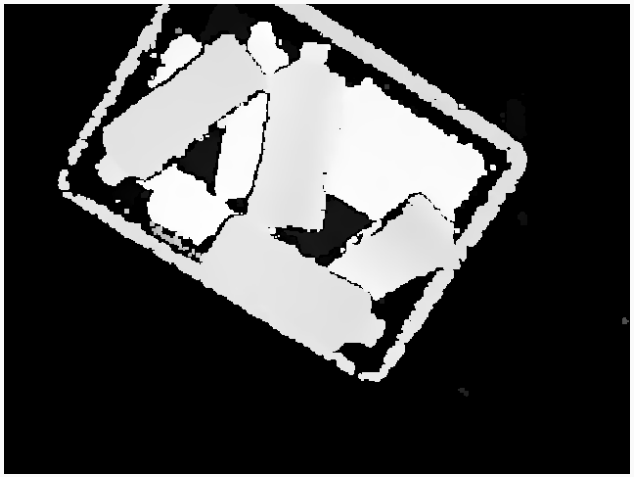
RGB Image



Depth Map

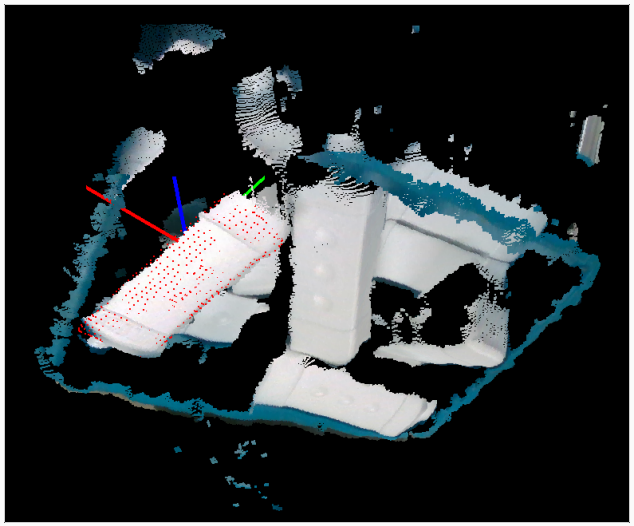
3D Detection

RGB Image



Depth Map

3D Detection Results

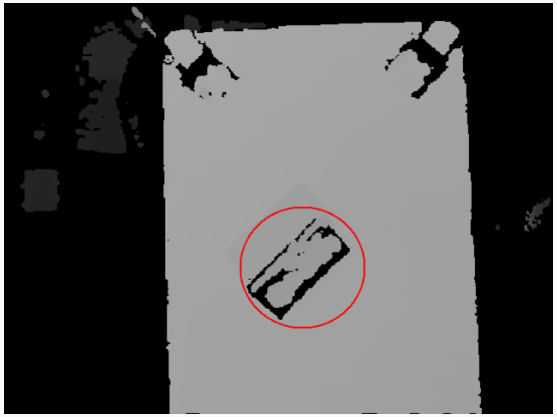


Dual RGB-D Camera

Problem & Key

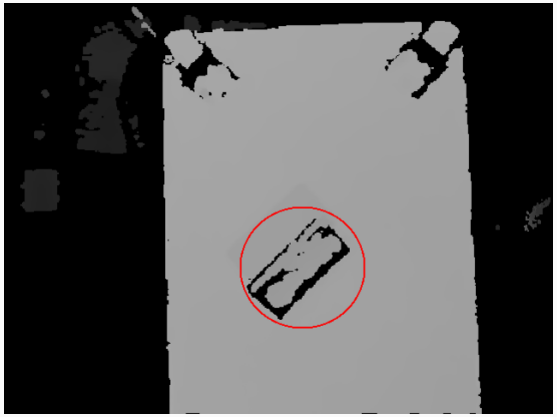
Problem & Key

Problem: SR300 相机对反光物体在某些角度下深度信息有严重的缺失

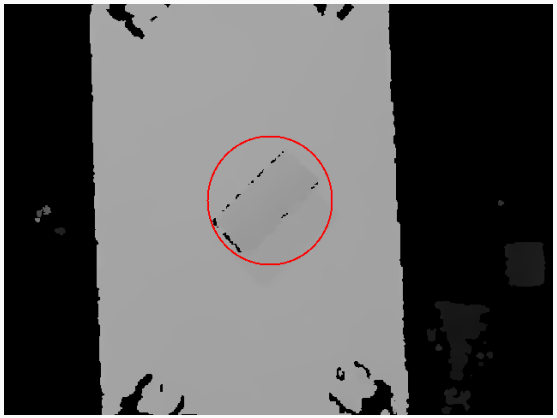


Problem & Key

Problem: SR300 相机对反光物体在某些角度下深度信息有严重的缺失

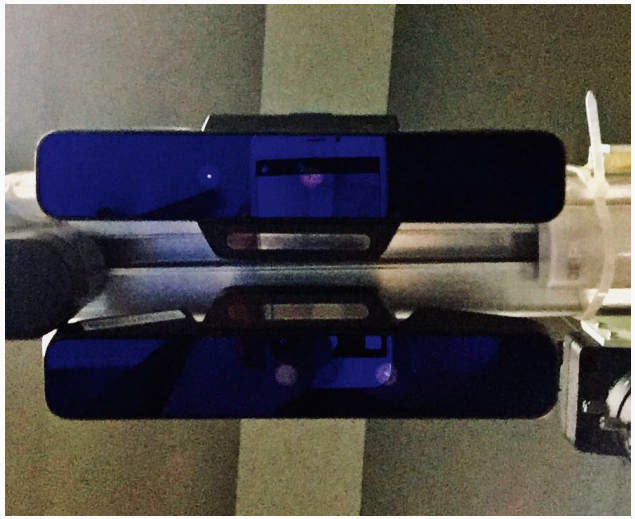


Key: 改变相机的拍摄角度



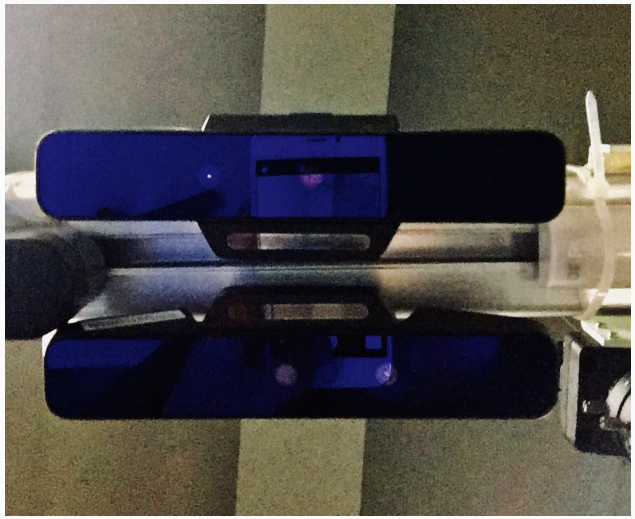
Structure of Dual RGB-D Camera

Structure of Dual RGB-D Camera

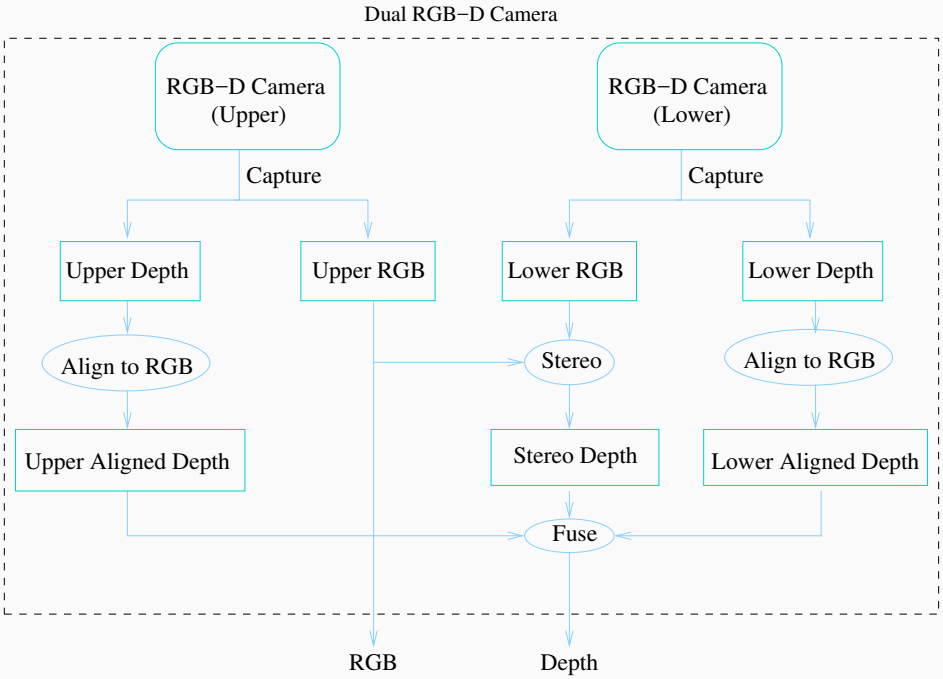


对偶 RGB-D 相机物理结构

Structure of Dual RGB-D Camera



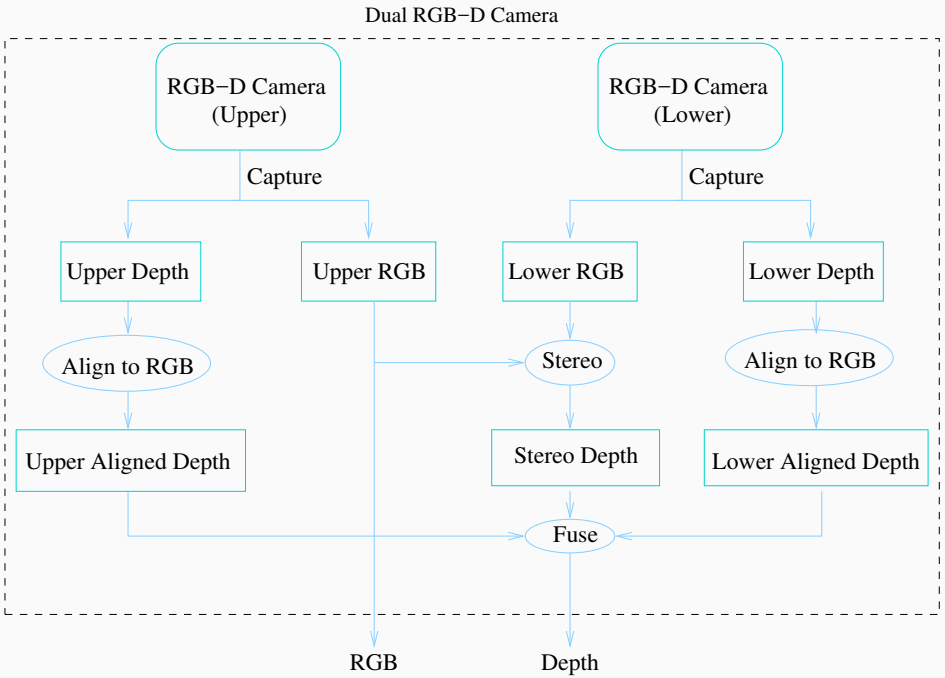
对偶 RGB-D 相机物理结构



对偶 RGB-D 相机内部原理图

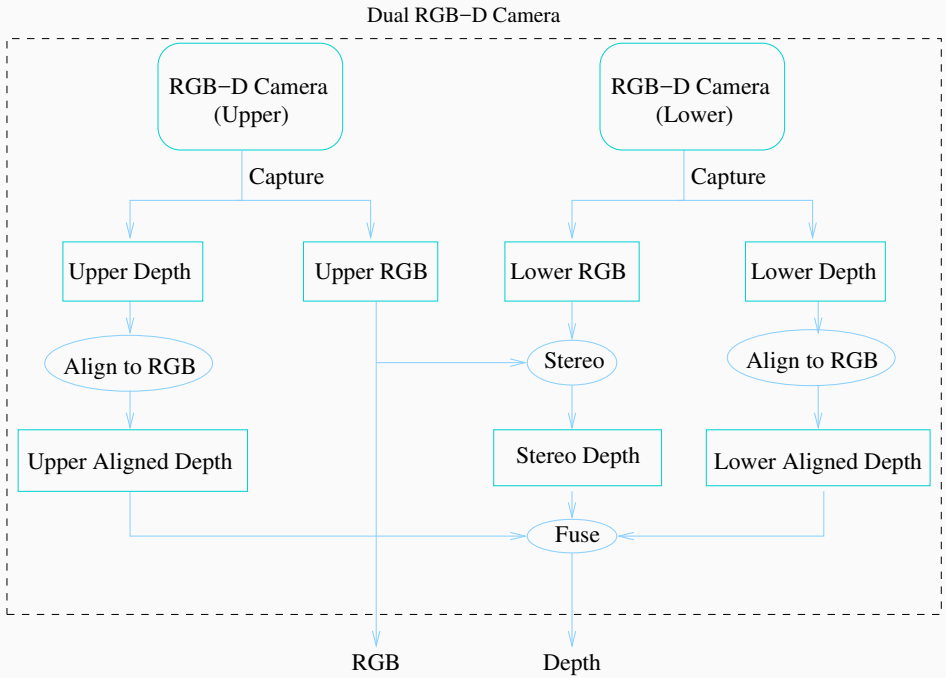
Implementation

Implementation



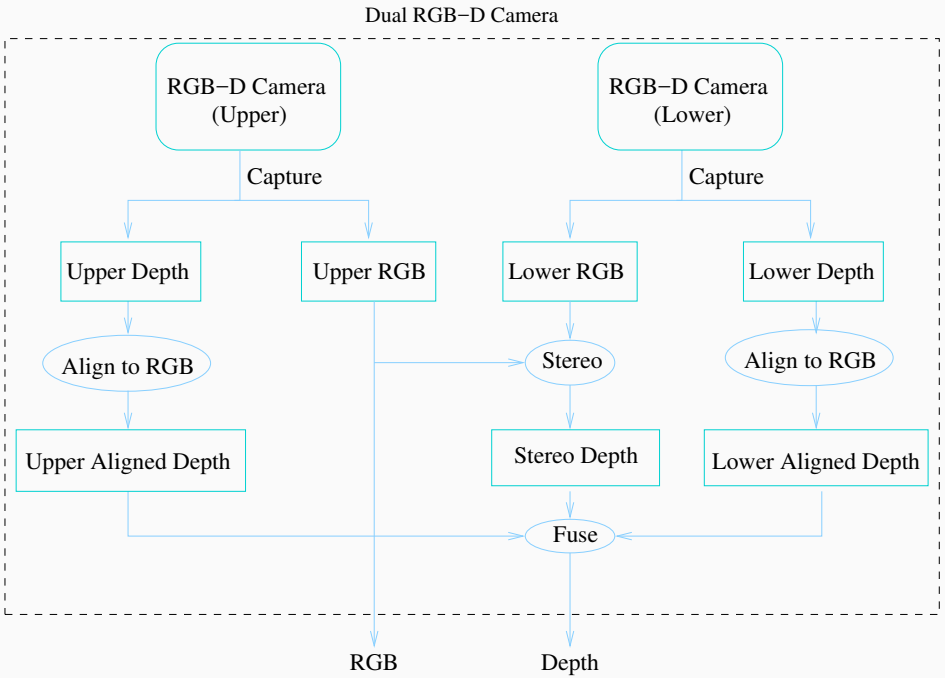
Implementation

- 将深度图与彩色图对齐



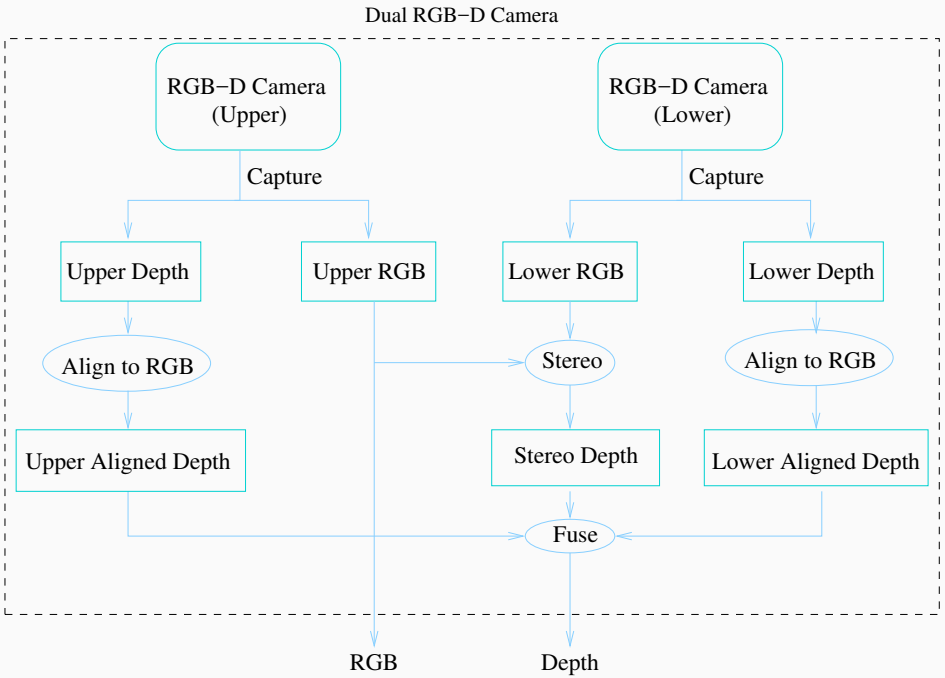
Implementation

- 将深度图与彩色图对齐
- 通过双目匹配算法 (ELAS) 形成一张新的深度图



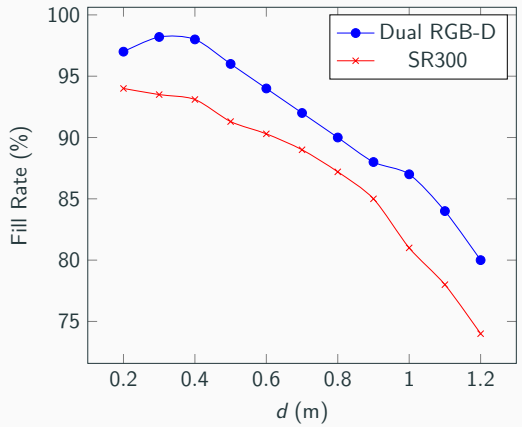
Implementation

- 将深度图与彩色图对齐
- 通过双目匹配算法 (ELAS) 形成一张新的深度图
- 融合三张深度图



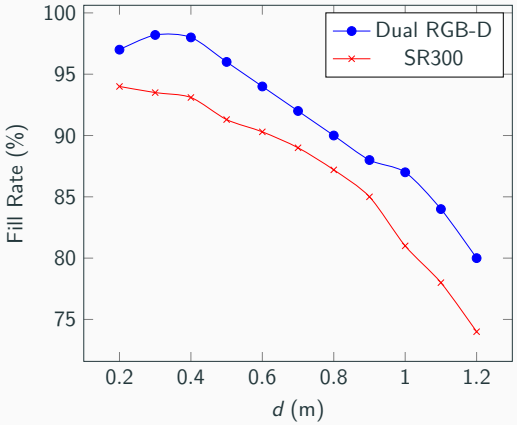
Experiment & Results

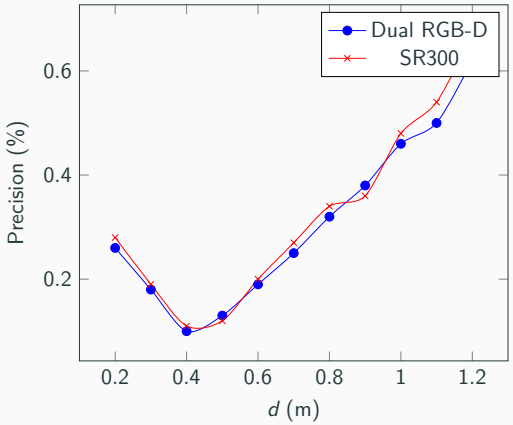
Experiment & Results



$$\frac{M'}{M} \times 100\%$$

Experiment & Results

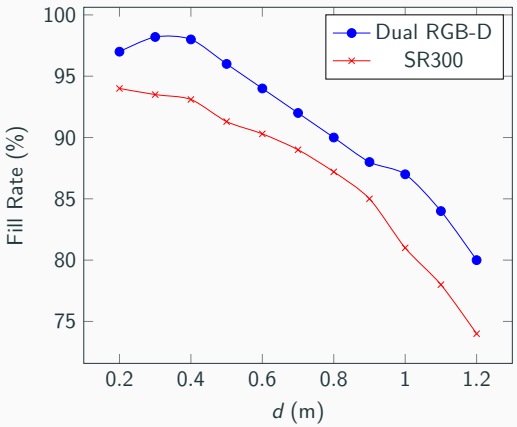




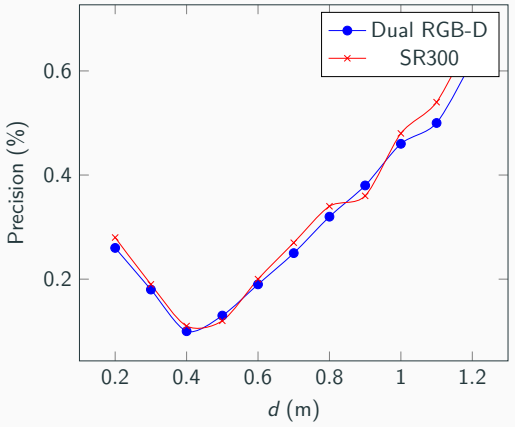
$$\frac{M'}{M} \times 100\%$$

$$\frac{|\hat{d}-d|}{d} \times 100\%$$

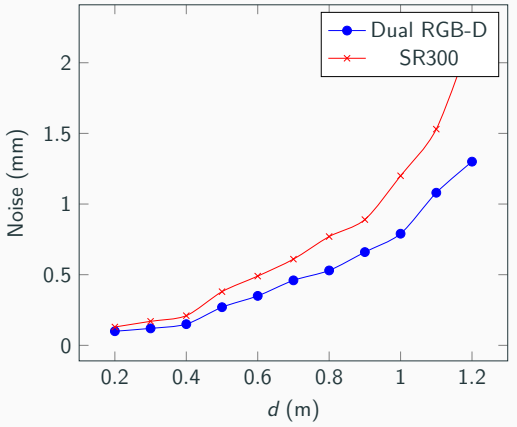
Experiment & Results



$$\frac{M'}{M} \times 100\%$$



$$\frac{|\hat{d}-d|}{d} \times 100\%$$

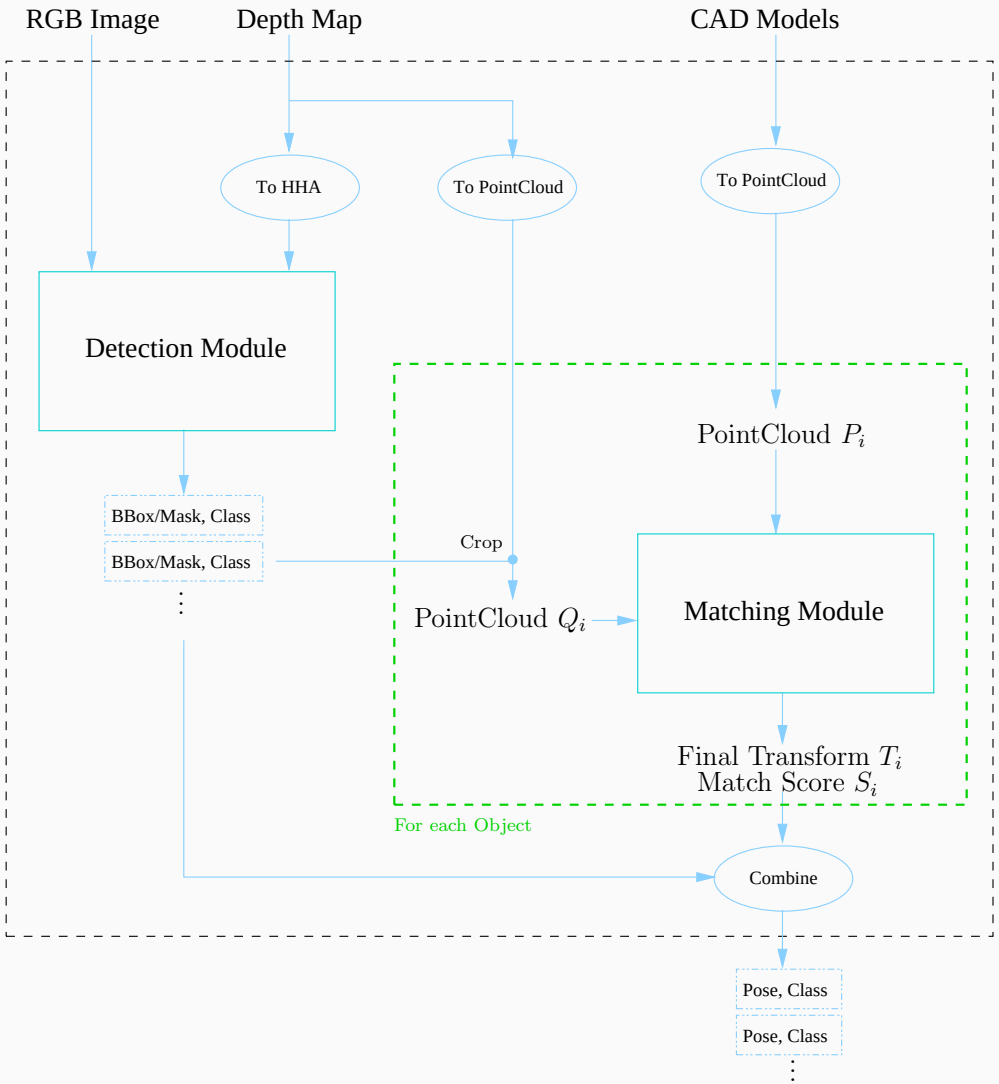


$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2}$$

3D-MRAI Algorithm

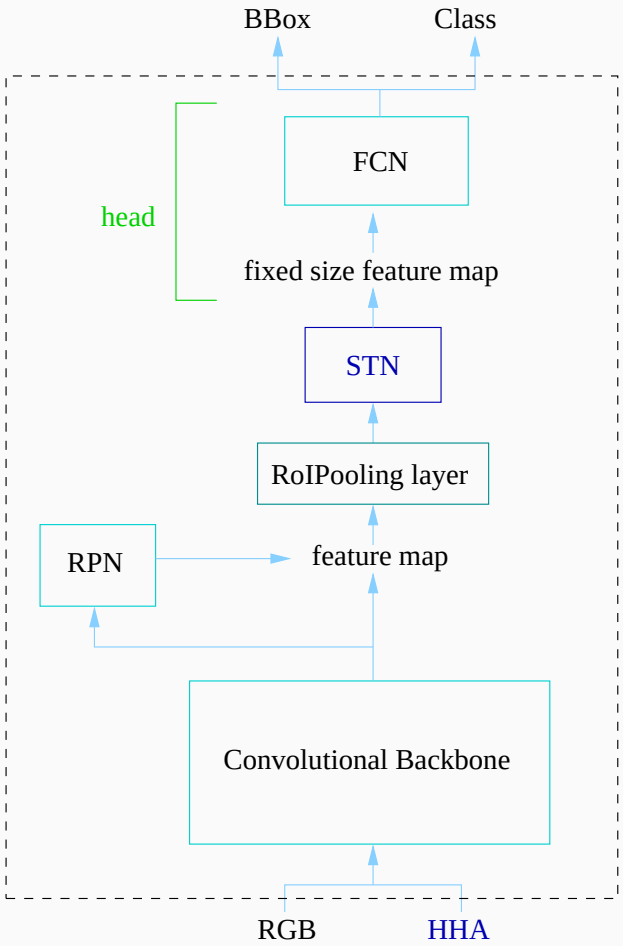
3D-MRAI

3D-MRAI (3D Mask R-CNN with Angle-fixed-4PCS and ICP)



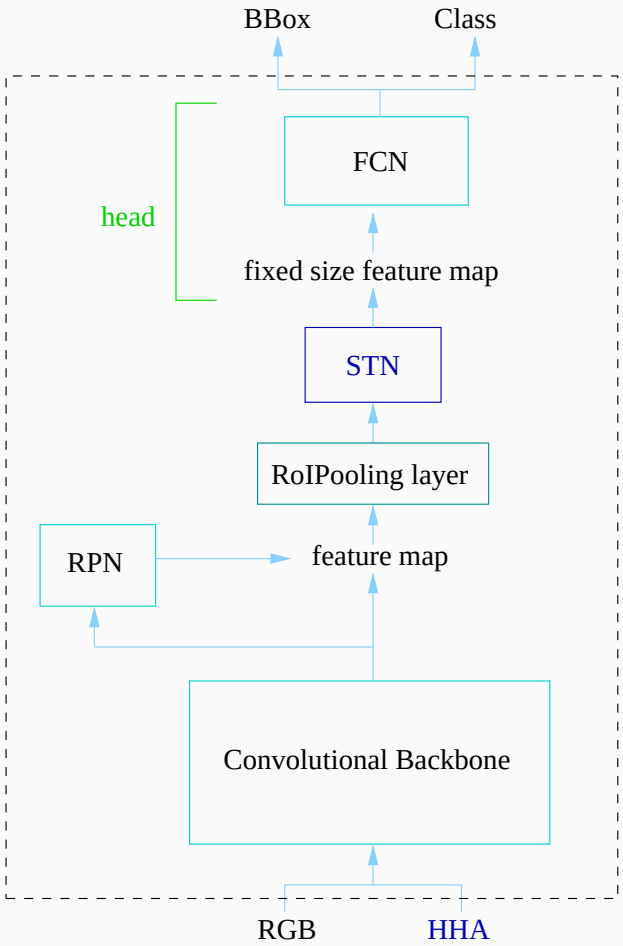
Detection Module

Detection Module

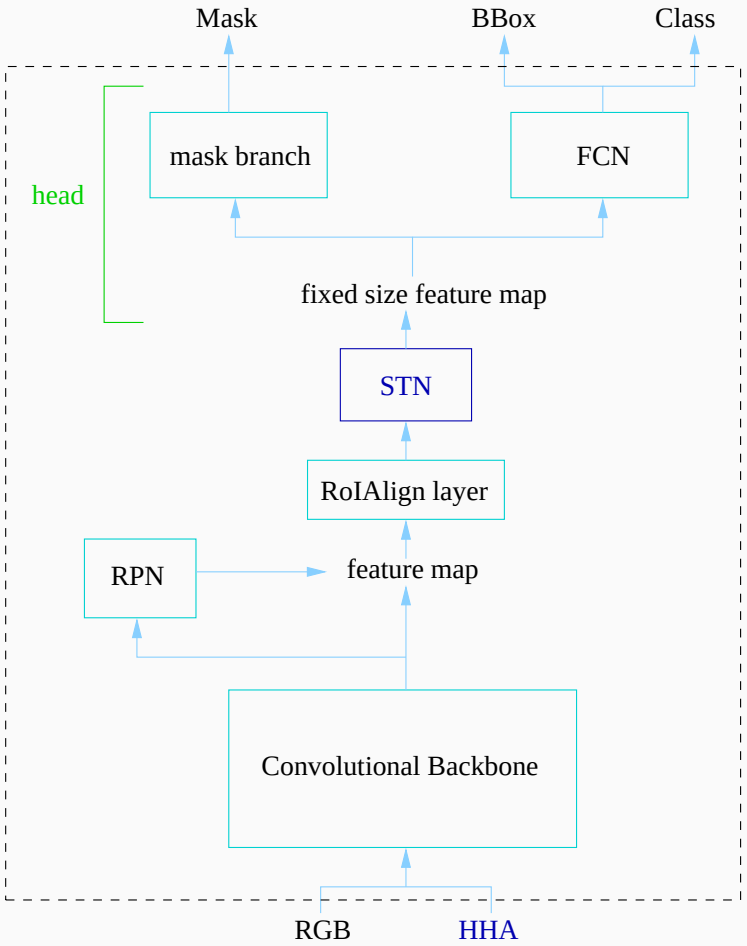


Based on Faster R-CNN

Detection Module



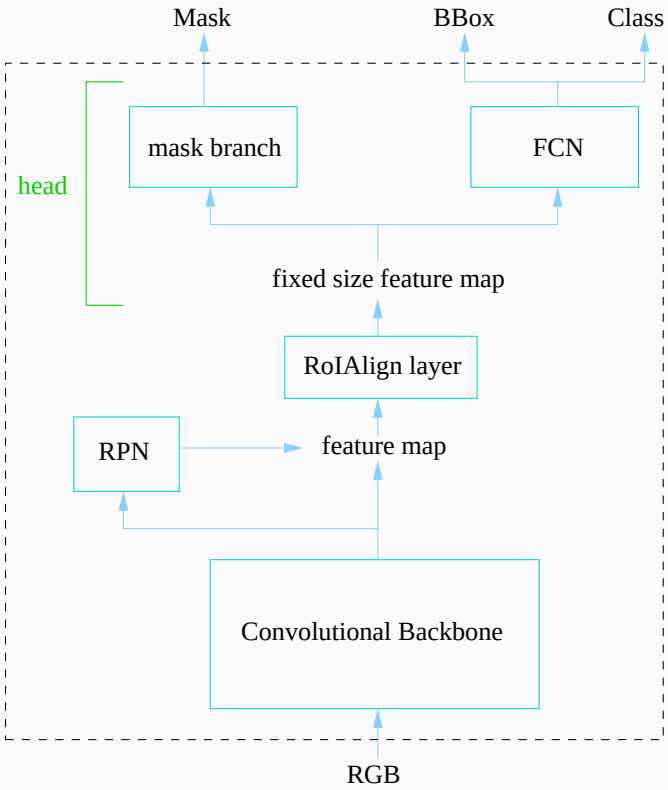
Based on Faster R-CNN



Based on Mask R-CNN

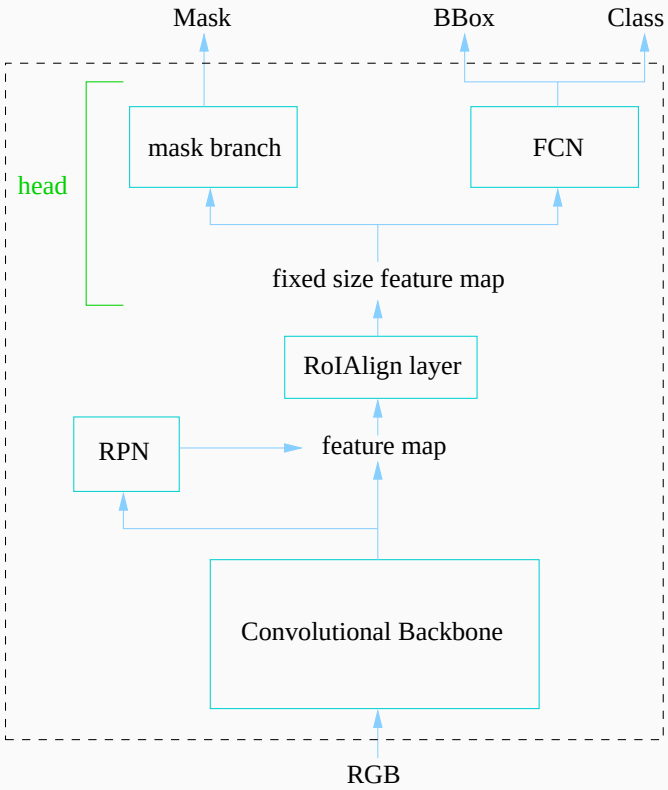
Mask R-CNN Module

Mask R-CNN Module

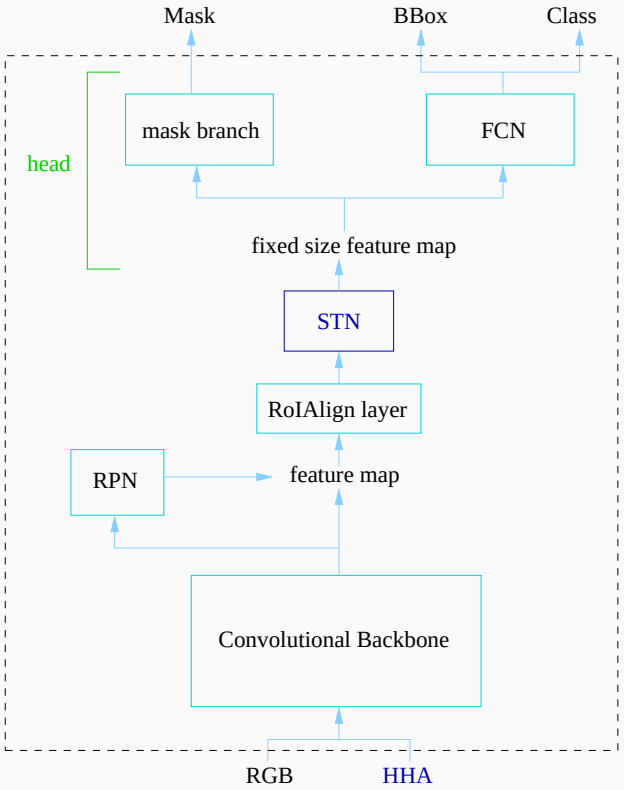


Original Mask R-CNN

Mask R-CNN Module

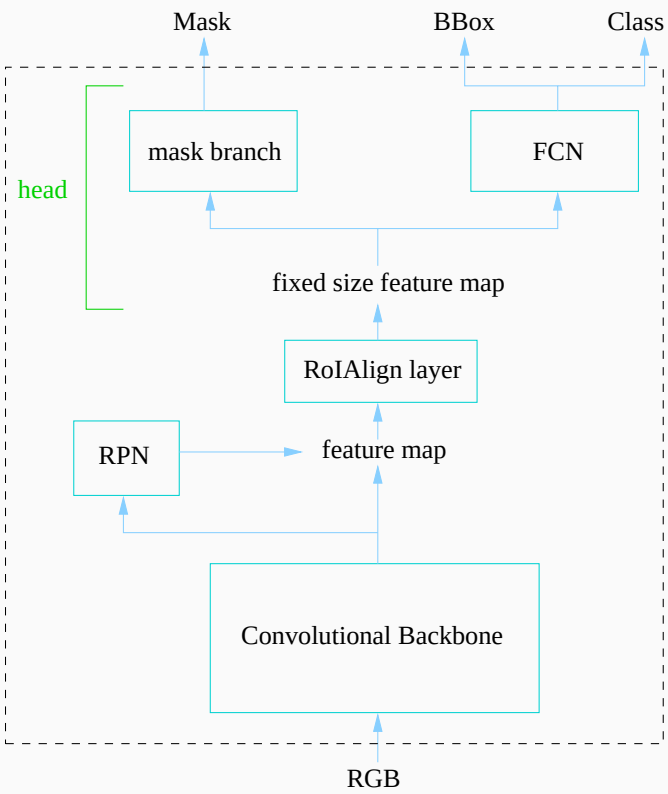


Original Mask R-CNN

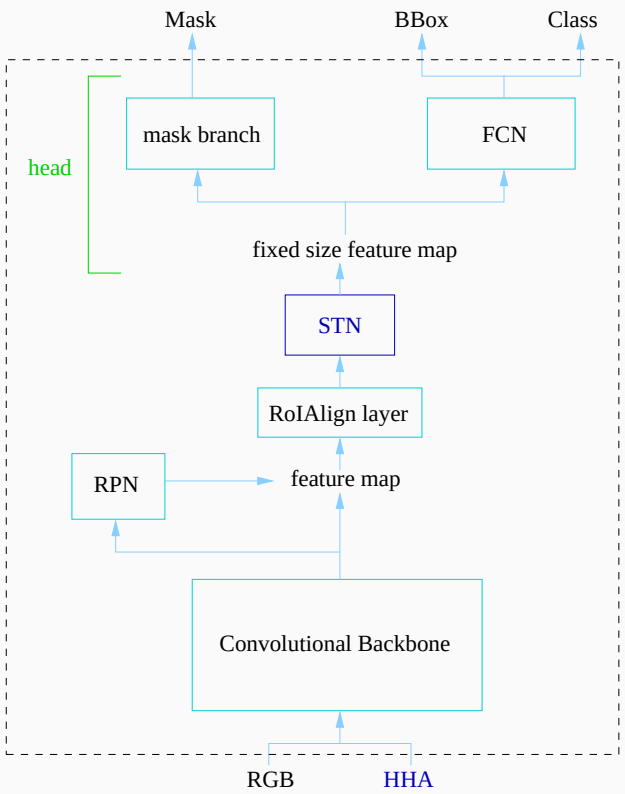


Mask R-CNN Detection Module

Mask R-CNN Module



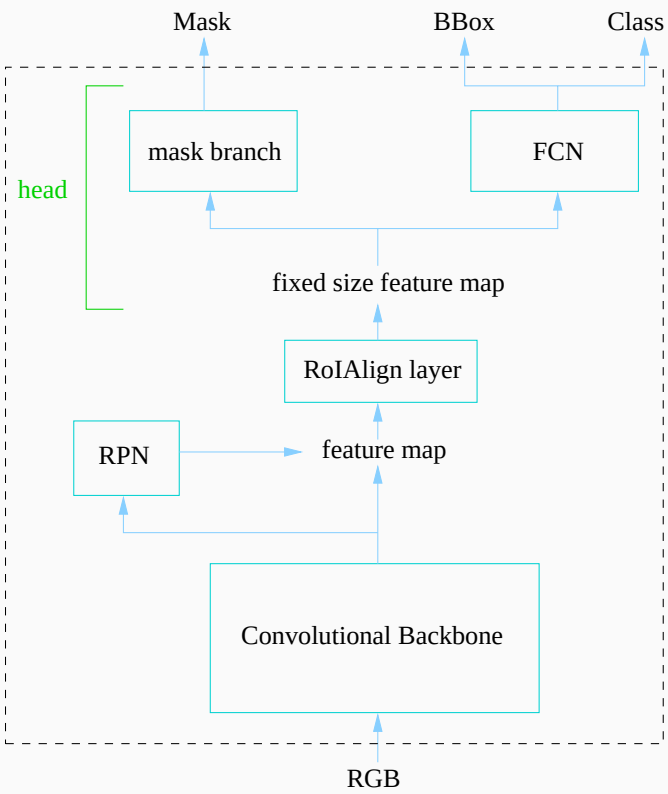
Original Mask R-CNN



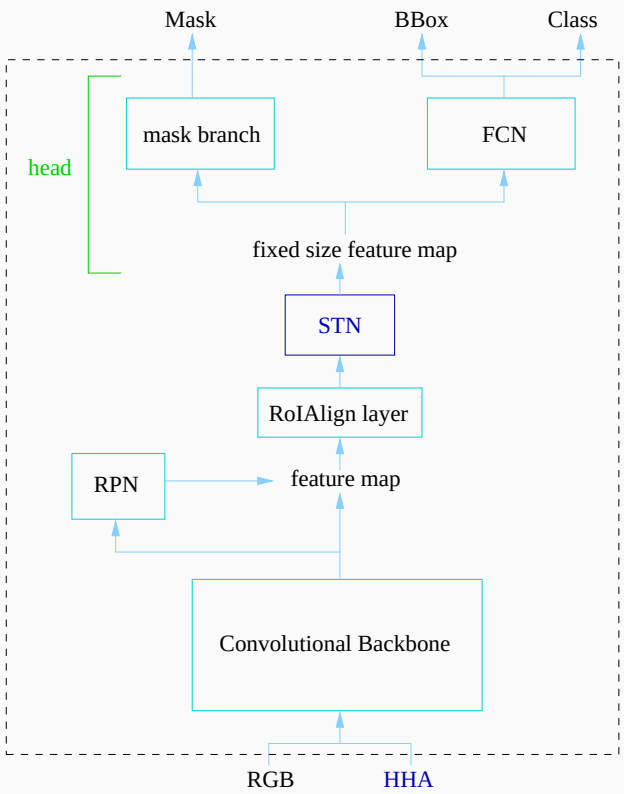
Mask R-CNN Detection Module

- 引入 HHA，有效利用三维信息，解决原算法难以检测纹理缺少的物体

Mask R-CNN Module



Original Mask R-CNN

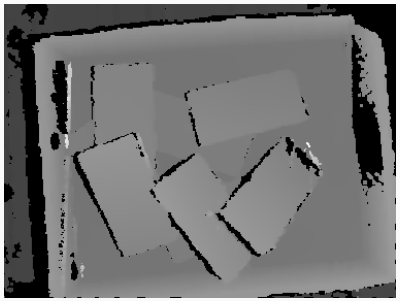


Mask R-CNN Detection Module

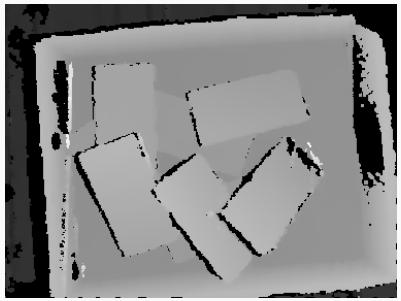
- 引入 HHA，有效利用三维信息，解决原算法难以检测纹理缺少的物体
- 在网络中增加 STN(Spatial Transformer Network)，使得提取的特征具有旋转不变性，增加了算法检测准确度

Detection Module

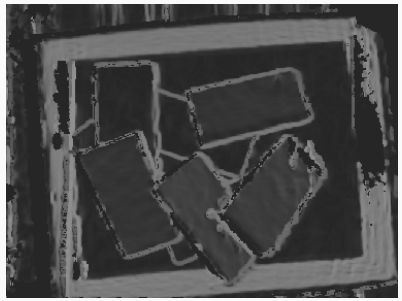
Detection Module



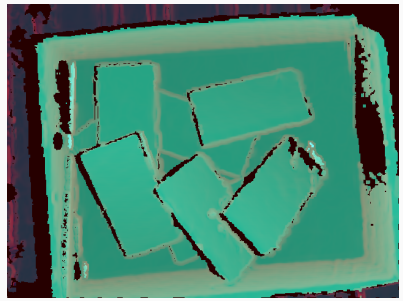
Horizontal disparity frame



Height above ground frame

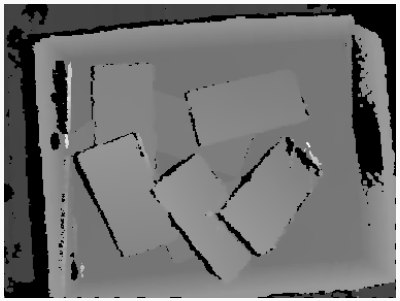


Angle with gravity frame

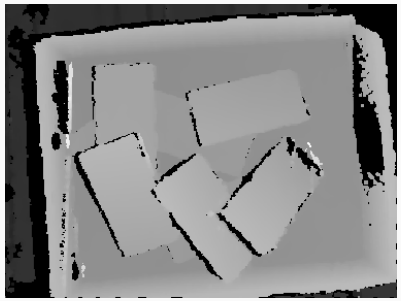


HHA frame

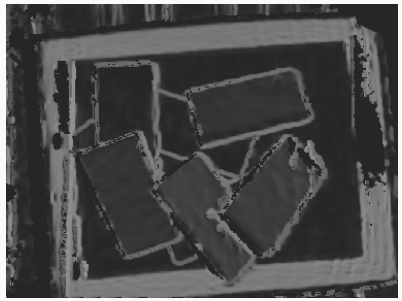
Detection Module



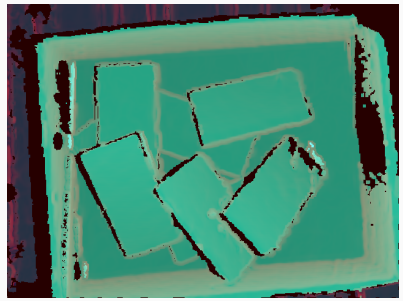
Horizontal disparity frame



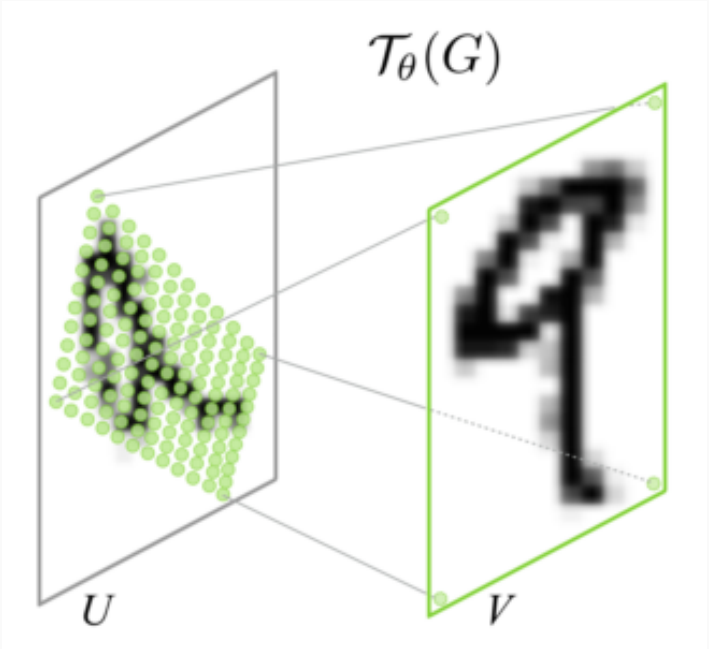
Height above ground frame



Angle with gravity frame



HHA frame



Detection Module Results

Detection Module Results

表 1: APC 数据集上的精确度

	input	output	AP	$AP_{0.5}$	$AP_{0.75}$
Faster R-CNN	RGB	bbox	33.26	56.29	34.03
Our method(based on Faster R-CNN)	RGB+HHA	bbox	**34.55**	**57.99**	**34.69**
Mask R-CNN	RGB	mask	32.34	55.78	33.12
Our method(based on Mask R-CNN)	RGB+HHA	mask	**33.94**	**56.45**	**33.99**

Detection Module Results

表 1: APC 数据集上的精确度

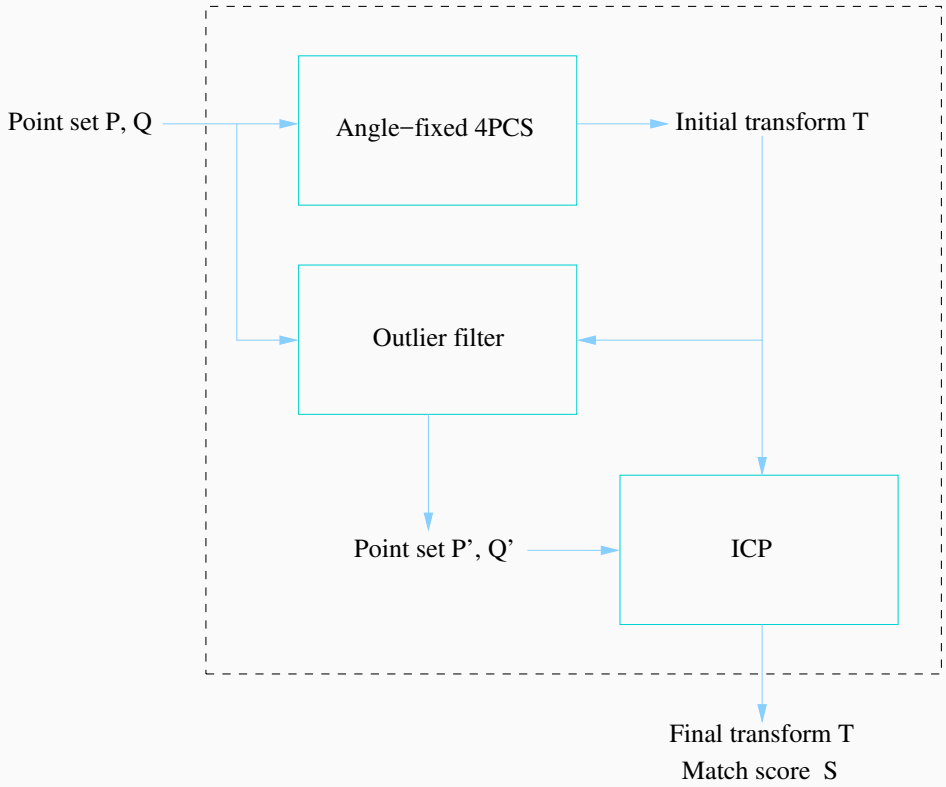
	input	output	AP	$AP_{0.5}$	$AP_{0.75}$
Faster R-CNN	RGB	bbox	33.26	56.29	34.03
Our method(based on Faster R-CNN)	RGB+HHA	bbox	**34.55**	**57.99**	**34.69**
Mask R-CNN	RGB	mask	32.34	55.78	33.12
Our method(based on Mask R-CNN)	RGB+HHA	mask	**33.94**	**56.45**	**33.99**

表 2: workpiece 数据集上的精确度

	input	output	AP	$AP_{0.5}$	$AP_{0.75}$
Faster R-CNN	RGB	bbox	18.78	37.49	19.46
Our method(based on Faster R-CNN)	RGB+HHA	bbox	**32.39**	**56.37**	**33.54**
Mask R-CNN	RGB	mask	16.12	35.95	18.74
Our method(based on Mask R-CNN)	RGB+HHA	mask	**30.98**	**53.74**	**32.19**

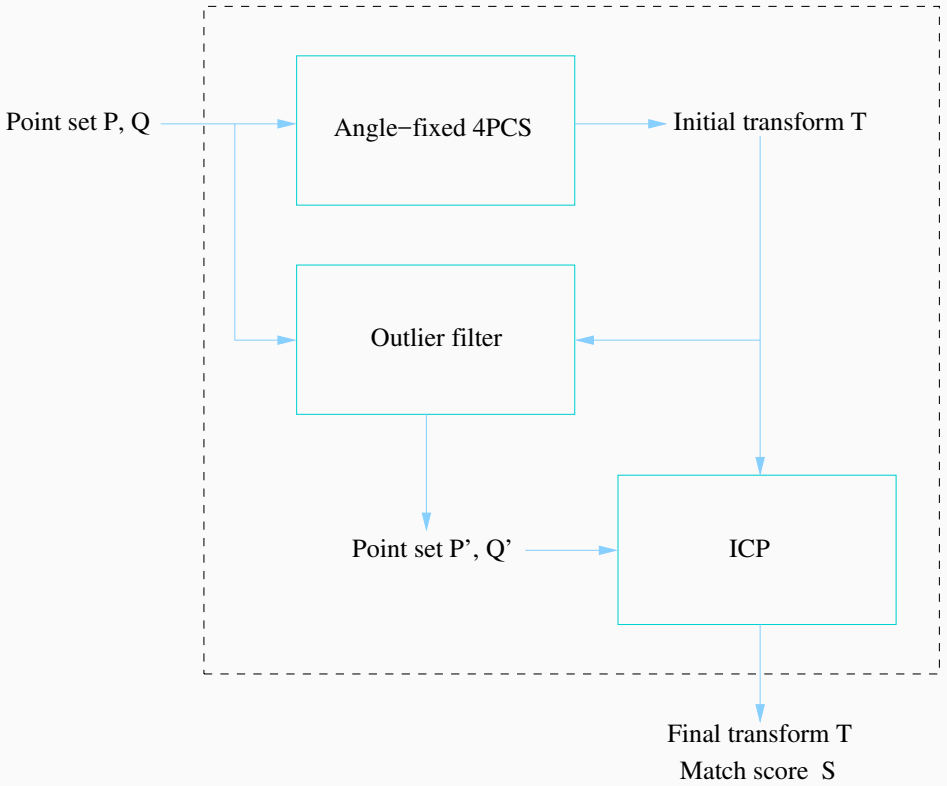
Matching Module

Matching Module



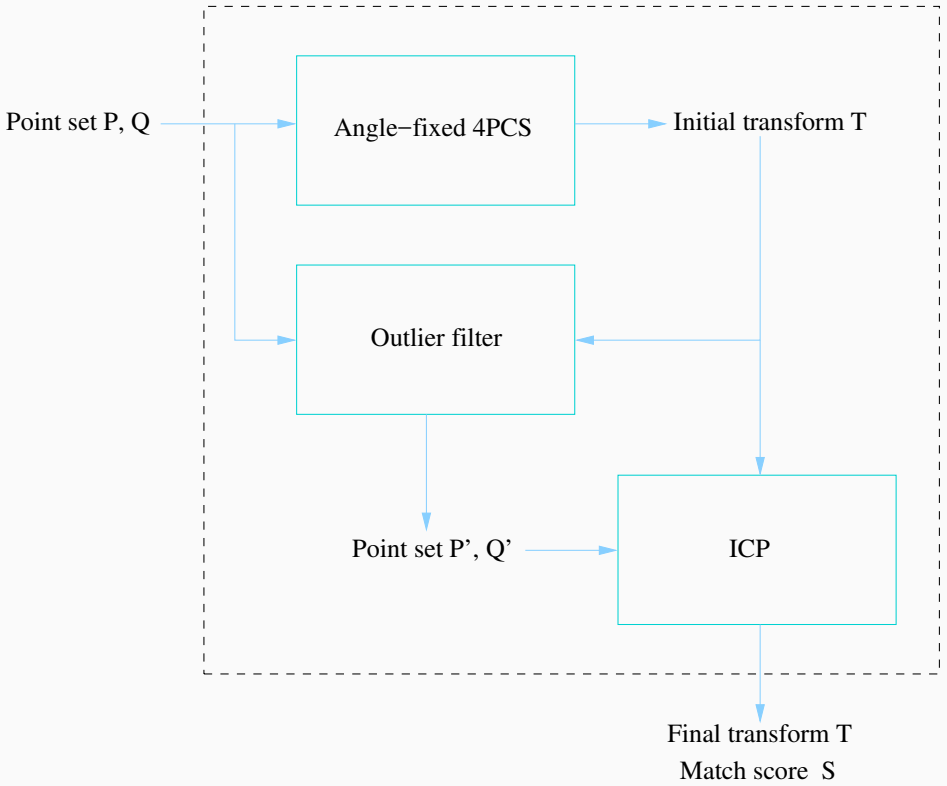
Matching Module

- 去除 4PCS 算法中角度不相等的基，减少了算法运算时间



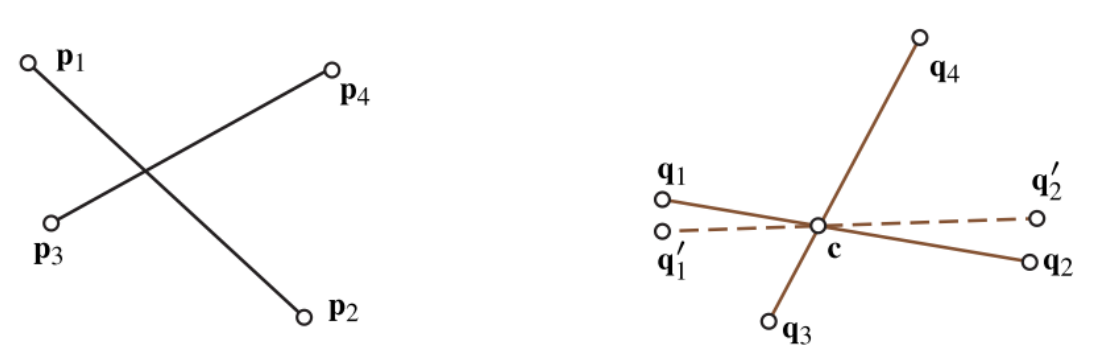
Matching Module

- 去除 4PCS 算法中角度不相等的基，减少了算法运算时间
- 通过增加滤波和 ICP 算法，增加了匹配精度



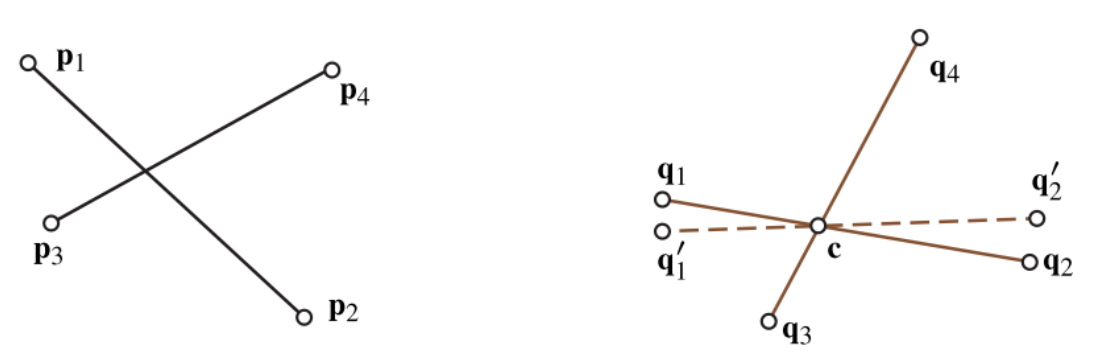
Matching Module

Matching Module

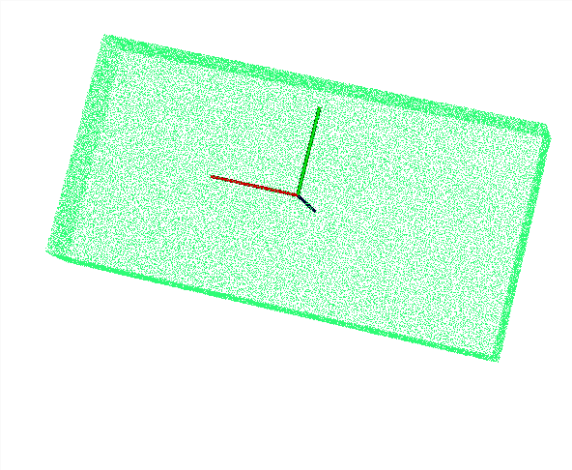


4PCS 存在的问题

Matching Module



4PCS 存在的问题

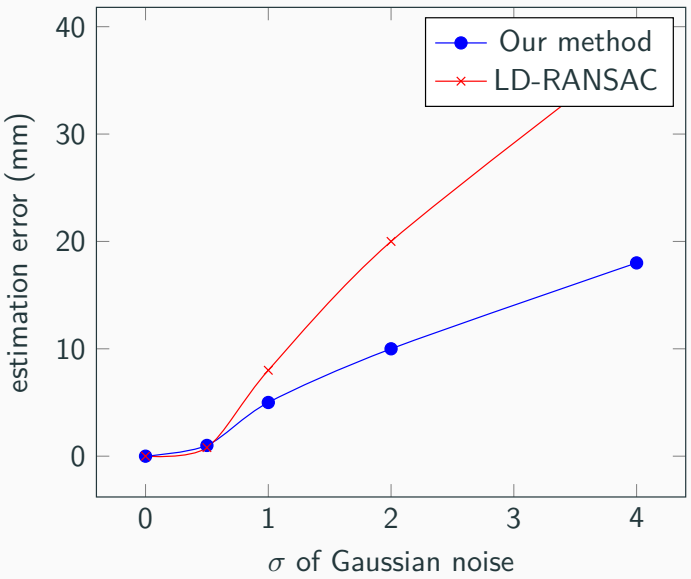


目标 3D 模型转化的点云

从 BBox 裁剪得到的点云

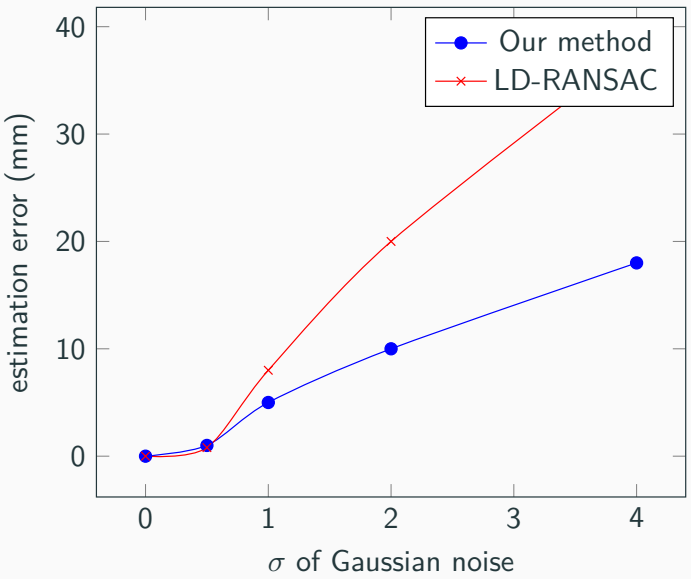
Matching Module Results

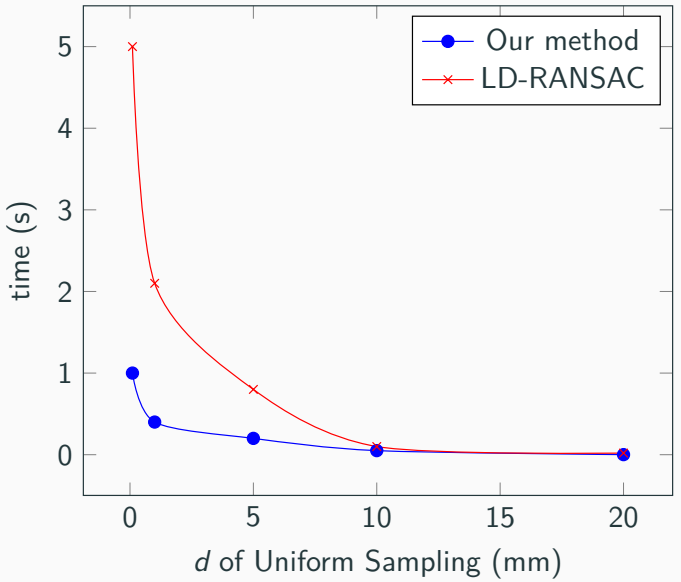
Matching Module Results



$$E = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \min_{p_j \in T(P)} \|q_i - p_j\|^2}$$

Matching Module Results

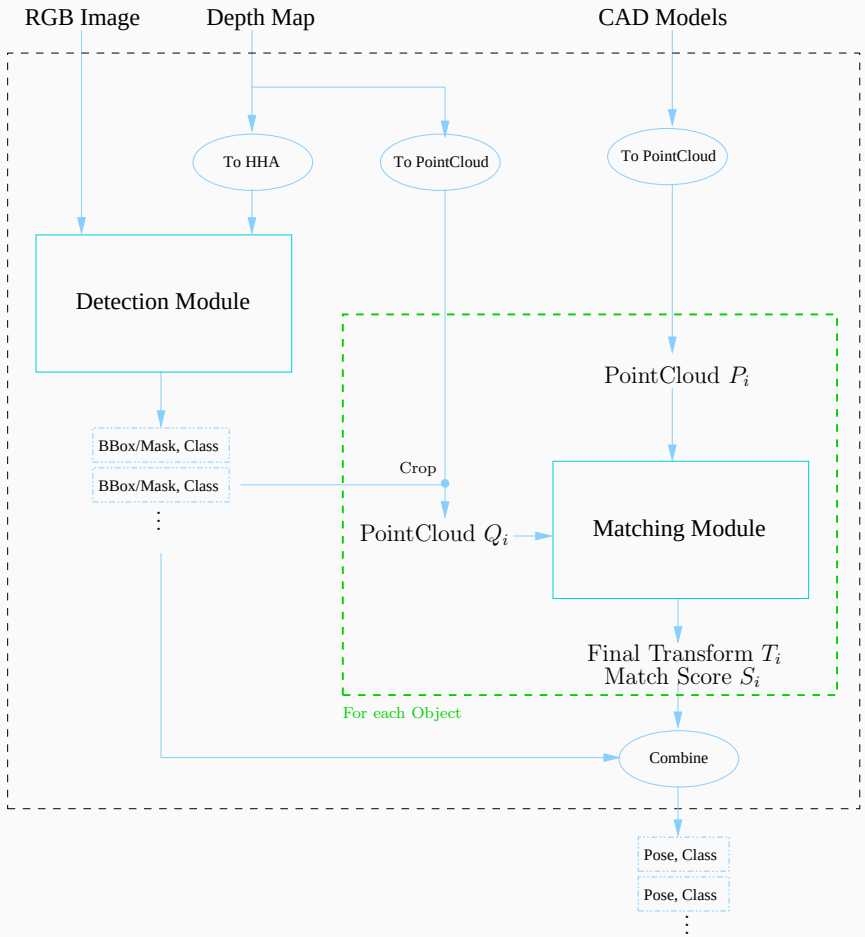




$$E = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \min_{p_j \in T(P)} \|q_i - p_j\|^2}$$

$d \uparrow$, the size of point cloud $\downarrow$

Combine Everything



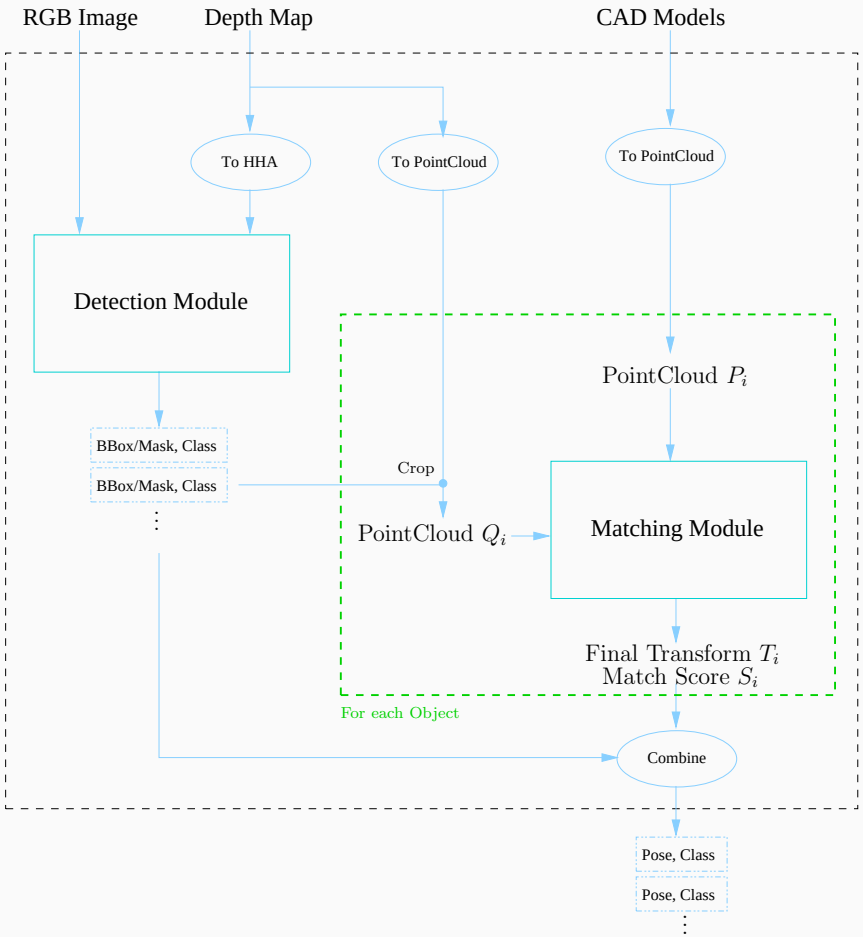
Combine Everything

算法 2: 3D-MRAI 算法

Input: RGB Image I, Depth Map D, CAD Models M

Output: Set of Pose and Class Res

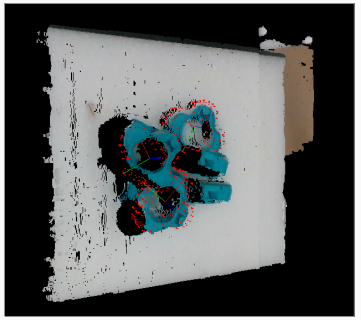
$Res \leftarrow \varnothing$;
$P \leftarrow \varnothing$;
forall $M_i \in M$ **do**
    $P \leftarrow \{P, CAD2PointCloud(M_i)\}$;
$H = Depth2HHA(D)$;
$Q = Depth2PointCloud(D)$;
$Mask, Class \leftarrow DetectModule(I, H)$;
forall $m_i \in Mask, c_i \in Class$ **do**
    $Q_i \leftarrow Crop(Q, m_i)$;
    $P_i \leftarrow P(c_i)$;
    $T_i, S_i \leftarrow MatchModule(P_i, Q_i)$;
    if $S_i > S_{min}$ **then**
        $Res \leftarrow \{Res, [T_i, c_i]\}$;
return Res



3D-MRAI Results

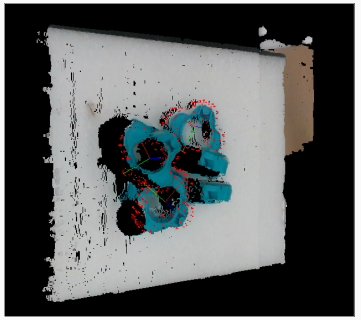
3D-MRAI Results





3D-MRAI Results

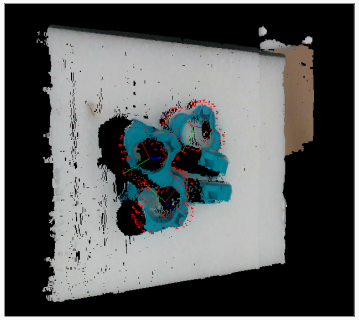




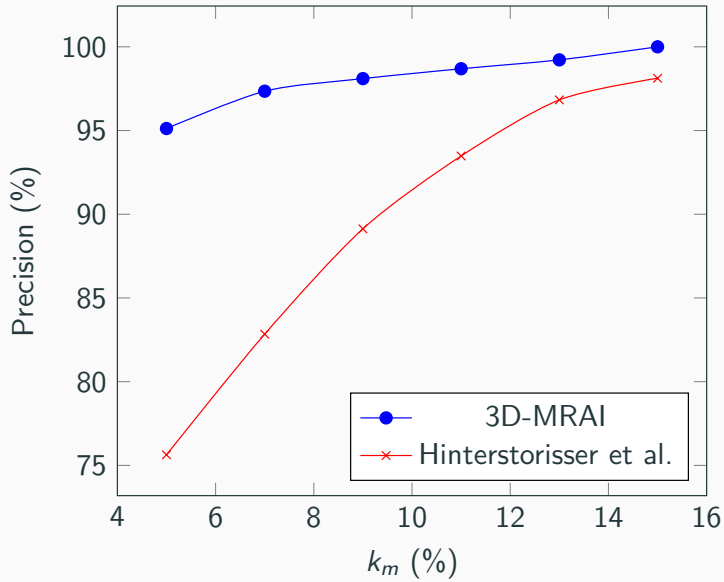
k_m[%]	5	7	9	11	13	15
Hinterstorisser et al.	75.63	83.84	89.13	93.48	96.83	98.12
3D-MRAI	95.12	97.35	98.10	98.69	99.22	100.00

3D-MRAI Results



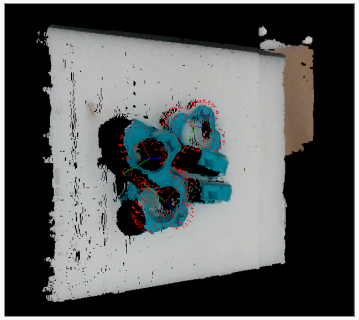


k_m[%]	5	7	9	11	13	15
Hinterstorisser et al.	75.63	83.84	89.13	93.48	96.83	98.12
3D-MRAI	95.12	97.35	98.10	98.69	99.22	100.00



3D-MRAI Results

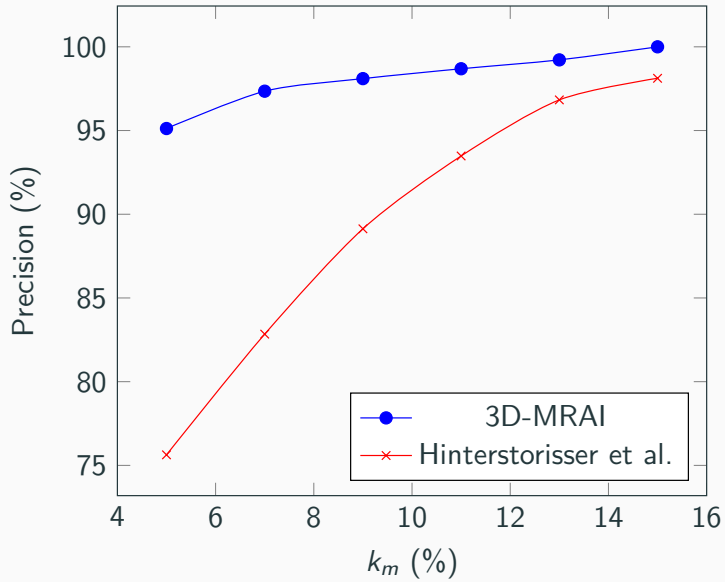




$k_m[\%]$	5	7	9	11	13	15
Hinterstorisser et al.	75.63	83.84	89.13	93.48	96.83	98.12
3D-MRAI	95.12	97.35	98.10	98.69	99.22	100.00

$$m = \underset{\mathbf{x} \in M}{avg} \parallel (R\mathbf{x} + t) - (\tilde{R}\mathbf{x} + \tilde{t}) \parallel$$

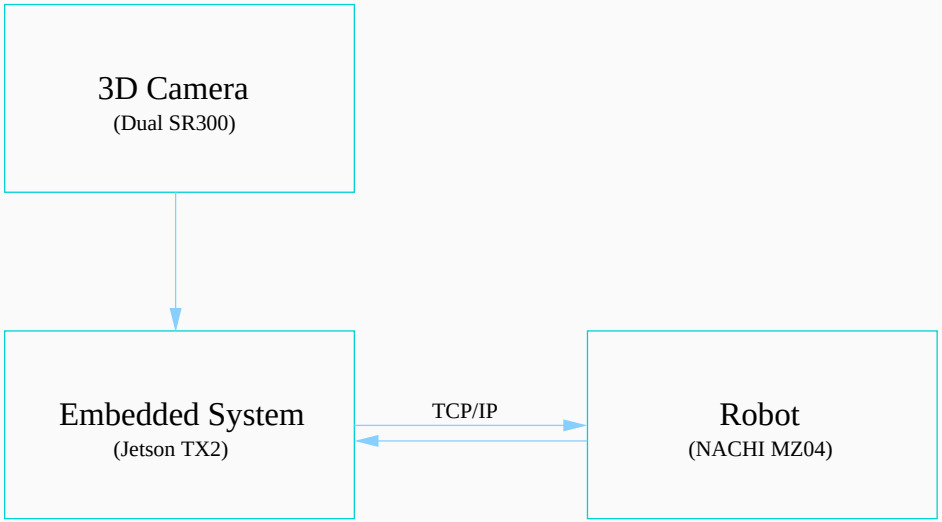
如果 $m < k_m d$，算法输出位姿正确



Application: Bin-picking

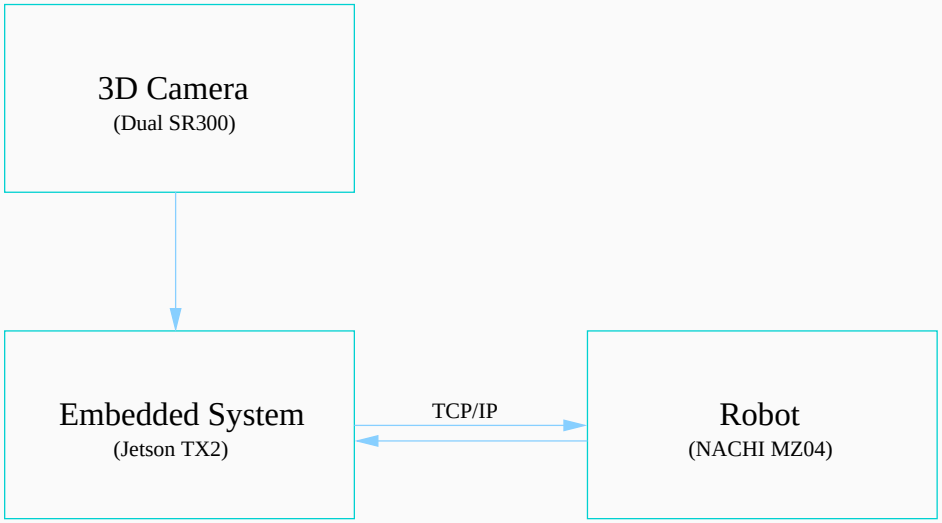
System Structure

System Structure

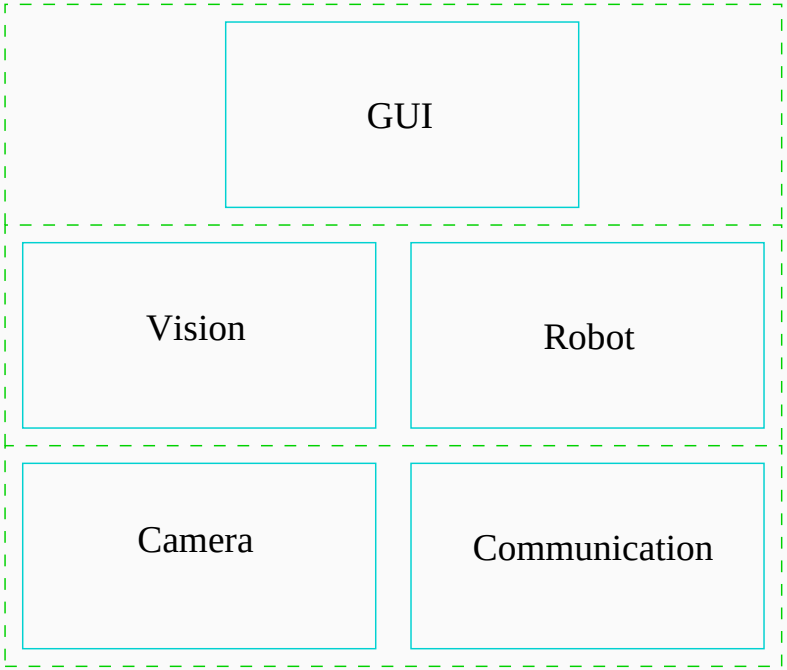


硬件架构

System Structure



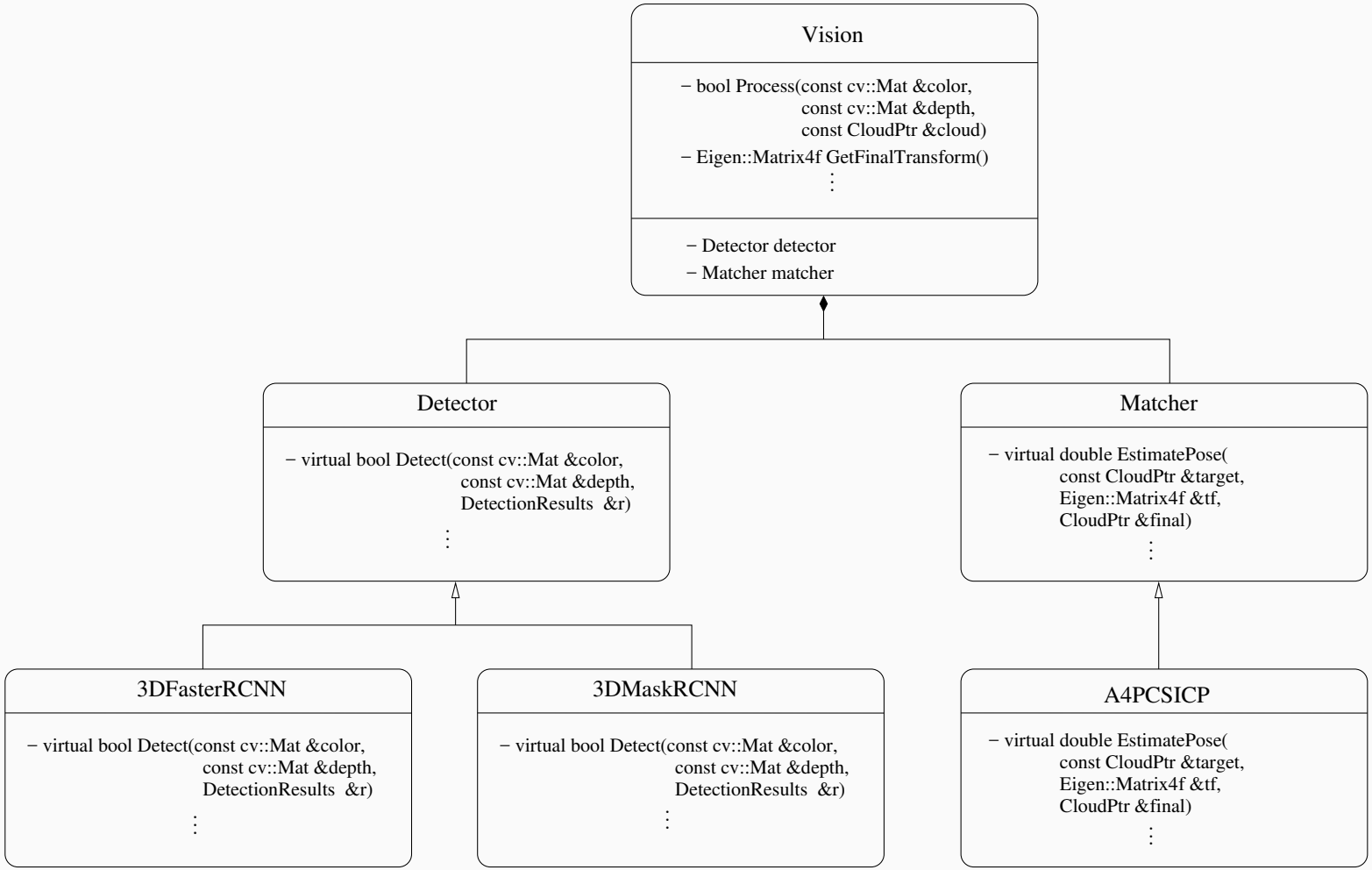
硬件架构



软件架构

Design Example

Design Example



Experiment

Experiment

实验环境

Experiment

实验环境

物料箱

Experiment & Results

Experiment & Results

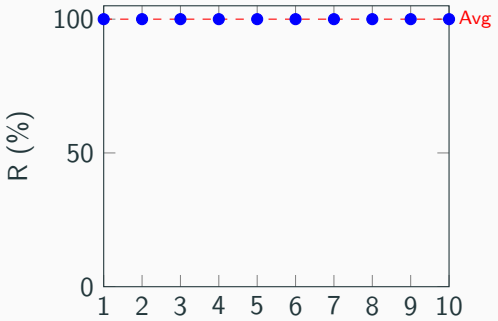
	成功率 R	响应时间 T_r	抓取时间 T_1	放置时间 T_2	工作周期 T
1	100%	711ms	7.6s	4.2s	11.9s
2	100%	729ms	6.8s	4.2s	11.0s
3	100%	708ms	6.5s	4.2s	10.7s
4	100%	701ms	8.1s	4.2s	12.3s
5	100%	713ms	9.3s	4.2s	13.5s
6	100%	722ms	6.6s	4.2s	10.8s
7	100%	693ms	7.9s	4.2s	12.1s
8	100%	732ms	9.1s	4.2s	13.3s
9	100%	718ms	8.9s	4.2s	13.1s
10	100%	723ms	6.9s	4.2s	11.1s
Avg.	**100%**	**715ms**	**7.77s**	**4.20s**	**11.98s**

Experiment & Results

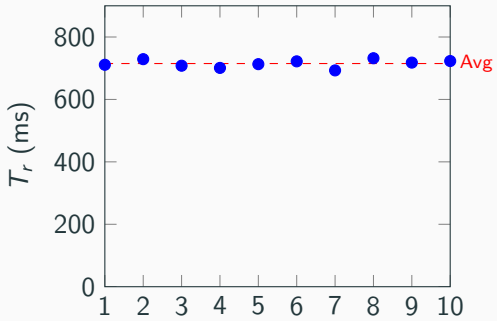
	成功率 R	响应时间 T_r	抓取时间 T_1	放置时间 T_2	工作周期 T
1	100%	711ms	7.6s	4.2s	11.9s
2	100%	729ms	6.8s	4.2s	11.0s
3	100%	708ms	6.5s	4.2s	10.7s
4	100%	701ms	8.1s	4.2s	12.3s
5	100%	713ms	9.3s	4.2s	13.5s
6	100%	722ms	6.6s	4.2s	10.8s
7	100%	693ms	7.9s	4.2s	12.1s
8	100%	732ms	9.1s	4.2s	13.3s
9	100%	718ms	8.9s	4.2s	13.1s
10	100%	723ms	6.9s	4.2s	11.1s
Avg.	**100%**	**715ms**	**7.77s**	**4.20s**	**11.98s**

$$T = T_1 + \max(T_r, T_2)$$

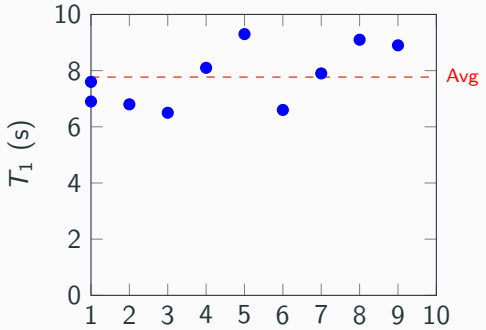
Experiment & Results



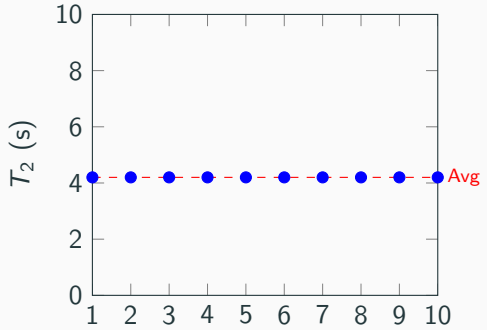
(a) 抓取成功率



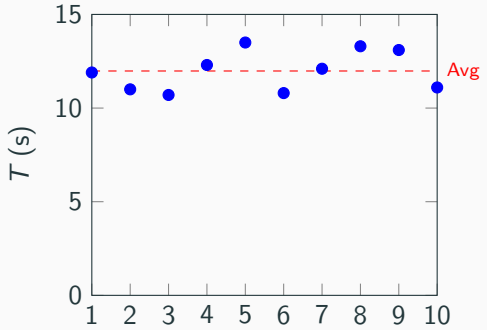
(b) 响应时间



(c) 抓取时间



(d) 放置时间



(e) 工作周期

Conclusion

Summary

Summary

- 对偶 RGB-D 相机结构相比单个 RGB-D 相机有更高的填充率，更低的噪声，但对精度的提升不明显

Summary

- 对偶 RGB-D 相机结构相比单个 RGB-D 相机有更高的填充率，更低的噪声，但对精度的提升不明显
- 3D-MRAI 算法相比传统算法表现出了较高的检测准确率，估计的位姿精度也更高，但 FPS 较低

Summary

- 对偶 RGB-D 相机结构相比单个 RGB-D 相机有更高的填充率，更低的噪声，但对精度的提升不明显
- 3D-MRAI 算法相比传统算法表现出了较高的检测准确率，估计的位姿精度也更高，但 FPS 较低
- 基于 3D-MRAI 算法设计的 Bin-Picking 视觉系统具有更高的抓取成功率、更快的响应速度，以及更低的成本

Questions?