

视觉SLAM方案整理及硬件选型调研

2018年05月09日 19:08:17 Mega_Li 阅读数 8710 更多

目前个人初步接触视觉SLAM开发相关工作，现在就相关学习做一些总结以加深个人理解，同时也希望能给其他网友提供一些帮助。这篇文章主要于视觉SLAM方案和硬件选型调研的总结，文中有关的视频是从youtube上收集的，上传到了百度网盘（链接：<https://pan.baidu.com/s/1-5B4q1rtL3XcCEIDzou0JA> 密码：lfm1），有需自取。由于个人能力有限，不保证文中说法的准确性，更多的是互相交流学习。

一、SLAM的引入

1.1定义

SLAM 是 Simultaneous Localization and Mapping 的缩写，中文译作“同时定位与地图构建”。它是指搭载特定传感器的主体，在没有环境分况下，于运动过程中建立环境的模型，同时估计自己的运动。如果这里的传感器主要为相机，那就称为“视觉 SLAM”。

1.2开发背景

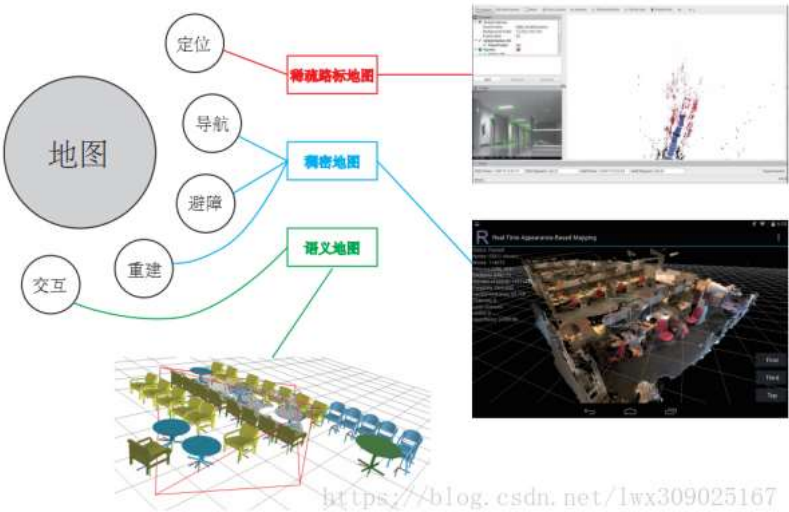


图1-1 SLAM中建图的分类与作用

- 引入SLAM的主要目的如下：
- 1) 建图。使用SLAM可在传感器具有良好表现的环境下建立精度较高的全局地图，建立好的地图会为后面的定位、导航等功能提供服务。
 - 2) 定位。视觉SLAM中可通过帧间数据匹配计算相机的相对变换，对应的就是机器人的位姿信息，不过该计算结果中存在累计误差的影响；利用全局地图，通过相机采集到的环境信息与地图进行匹配可以减小累积误差的影响，获得更加精准的机器人位姿。
 - 3) 导航。如果我们建立的地图中包含了“哪些地方可以通过，哪些地方不能通过”的信息，那么我们可以以此实现机器人在地图中从某一起点到径规划与跟踪，并能够对地图中固定障碍物实现避障。但这对我们能够建立的地图有要求，需要是“稠密”地图。

二、视觉SLAM

2.1 经典视觉SLAM框架

SLAM至今已历经三十多年的研究，这里给出经典视觉SLAM框架，这个框架本身及其包含的算法已经基本定型，并且已经在许多视觉程序库和机提供。



图2-1 经典视觉SLAM框架

- 我们把整个视觉 SLAM 流程分为以下几步：
- 1) 传感器信息读取。在视觉 SLAM 中主要为相机图像信息的读取和预处理。如果在机器人中，还可能有码盘、惯性传感器等信息的读取和同步。
- <https://blog.csdn.net/lwx309025167/article/details/80257549#commentBox>

- 2019/8/6(2条消息)视觉SLAM方案整理及硬件选型调研 - On my way - CSDN博客
- 2) 视觉里程计 (Visual Odometry, VO)。视觉里程计任务是估算相邻图像间相机的运动，以及局部地图的样子，VO 又称为前端 (Front End) 。

3) 后端优化 (Optimization) 。后端接受不同时刻视觉里程计测量的相机位姿，以及回环检测的信息，对它们进行优化，得到全局一致的轨迹和接在 VO 之后，又称为后端 (Back End) 。

4) 回环检测 (Loop Closing) 。回环检测判断机器人是否曾经到达过先前的位置。如果检测到回环， 它会把信息提供给后端进行处理。

5) 建图 (Mapping) 。它根据估计的轨迹， 建立与任务要求对应的地图。

某些使用场合中，我们只截取SLAM的部分功能应用到实际场景中。举例来说只使用VO部分我们可以得到连续的相机相对运动信息，虽然该运动误差，但应用中对此要求不高甚至不做要求，譬如VR头显中计算头显设备运动时的位姿。

不过一般在机器人应用场景中，个人认为建图功能也是不可或缺的。因为前端包括后端优化得到的运动信息始终包含累积误差，该累积误差在SLAM过回环检测（机器人回到某一处曾经经过的地方且系统识别出来）或者与事先建立好的具有较高精度的全局地图匹配来消除。但是机器人在实际运动中局路径一定会有重叠处，也就是说在SLAM计算中很可能出现不存在回环的情况，此时累积误差只能通过与全局地图匹配来消除，因此SLAM的建图功能缺。

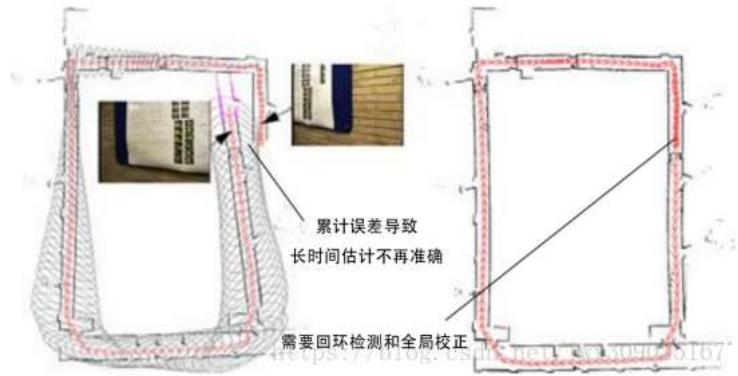


图2-2 回环检测消除累积误差

这一点在VINS开源项目（香港科技大学团队基于单目+IMU开发的开源SLAM方案）中的测试视频（见“视频/VINS/[Open Source] VINS-Monocular Visual-Inertial System in EuRoC MAV Dataset (MH_05 V1_03).mp4”）中也有体现。在回环检测前，SLAM计算得到的位姿与真实位姿生了很大的偏差，如图2-3所示；该偏差只有在相机回到了曾经经过的地方且SLAM成功执行了回环检测后才得到了消除，如图2-4所示。

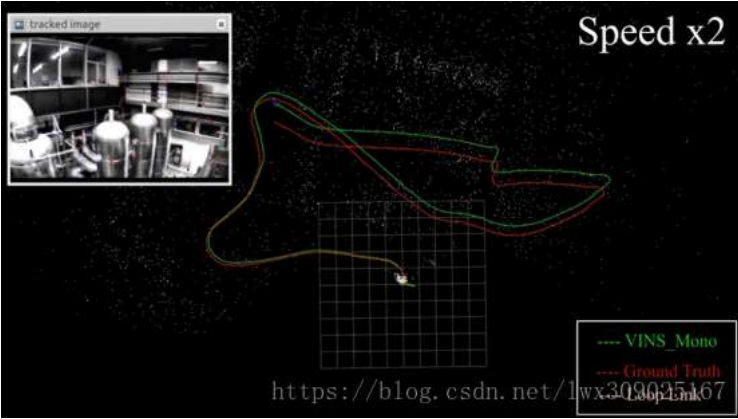


图2-3 未进行回环检测前的位姿累积误差

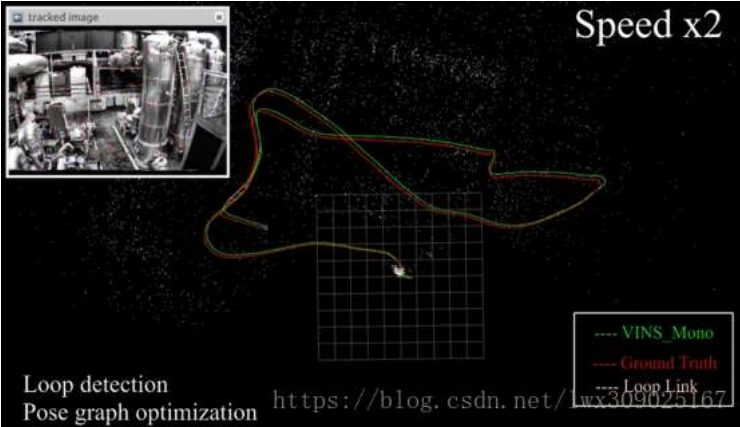


图2-4 回环检测消除累计误差

2.2视觉SLAM方案的分类

视觉SLAM方案可按照传感器的不同（单目、双目、RGBD、与IMU的组合等）、前端方法的不同（主要分为直接法和特征点法）、后端优化方案或者非线性优化）、生成地图形式的不同（稀疏地图、稠密地图等）具有不同的划分。这里主要以**传感器的不同对slam方案进行简单的介绍。**

- 1）单目slam。只使用一个摄像头采集信息，在尺度完成初始化的情况下（即相机初始阶段进行了一段位移且确定了该段位移的实际大小作为参考）完成连续图像间相机位姿的计算与建图。优点在于设备简单，成本低。缺点在于存在尺度漂移现象；图像的深度信息只能通过三角测量计算得到，对环境相机位移较小时计算误差较大，建图精度不高。
- 2）双目slam。使用双目相机采集信息，双目相机可以通过立体视觉原理计算得到像素深度，这样就不存在单目slam中深度未知的情况。优点在于尺度漂移现象要高于单目slam，且能够计算得到像素真实深度；缺点在于像素深度计算量较大，一般通过FPGA或者GPU加速实现实时计算输出。
- 3）RGBD SLAM。RGBD相机是指能够同时输出RGB图像和对应的深度图的相机，其测量像素深度不再通过耗时的被动双目匹配计算，而是**通过图像纹理加速计算或者硬件测量（结构光、TOF等）实现**。因此它可以看做减小了计算量的双目SLAM，但是RGBD相机普遍在室外表现效果不佳，更适用于室内环境。
- 4）近年来有学者提出单目/双目+IMU的slam方案，其中IMU主要起到的作用包括（1）解决单目slam的初始化尺度问题（2）追踪中提供较好的运动估计（3）提供重力方向（4）提供一个时间误差项以供优化。理论上来说IMU提供了冗余的运动信息，通过数据融合可以得到更加精确的运动估计。

从实现难度上来看：单目SLAM>双目SLAM>RGBD SLAM。

2.3 开源视觉SLAM方案汇总

目前比较流行的开源视觉SLAM方案如表2-1所示：

方案名称	传感器形式	地址
MonoSLAM	单目	https://github.com/hanmekim/SceneLib2
PTAM	单目	http://www.robots.ox.ac.uk/~gk/PTAM/
ORB-SLAM2	单目/双目/RGBD	https://github.com/raulmur/ORB_SLAM2
LSD-SLAM	单目为主	http://vision.in.tum.de/research/vslam/lstdslam
SVO	单目	https://github.com/uzh-rpg/rpg_svo
DTAM	RGB-D	https://github.com/anuranbaka/OpenDTAM
DVO	RGB-D	https://github.com/tum-vision/dvo_slam
DSO	单目	https://github.com/JakobEngel/dso
RTAB-MAP	双目/RGB-D	https://github.com/introlab/rtabmap
RGBD-SLAM-V2	RGB-D	https://github.com/felixendres/rgbdslam_v2
Elastic Fusion	RGB-D	https://github.com/mp3guy/ElasticFusion
OKVIS	多目+IMU	https://github.com/ethz-asl/okvis
ROVIO	单目+IMU	https://github.com/ethz-asl/rovio
VINS	单目+IMU	https://github.com/HKUST-Aerial-Robotics/VINS-Mono

表2-1 开源SLAM方案汇总

三、视觉SLAM设备选型

3.1 设备选型的重要性

说了这么多，终于到了设备选型这一部分。设备选型的重要性不言而喻，毕竟**“好模型架不住坏数据”**，SLAM模型建立得再好，如果设备选型太差，计算结果必定也不够理想。先以VINS项目为例，根据论文内容他们的设备型号和具体信息如下。

相机模块：单目相机，型号为MatrixVision的mvBlueFOX-MLC200w，具有全局快门；cmos型号为MT9V034，单色，分辨率为752X480。

IMU：该模块使用的是大疆的集成飞控模块A3的内置IMU模块，芯片型号为ADXL278和ADXRS290（都为工业级IMU芯片）；可以确定该模块使用了**卡尔曼滤波和零偏补偿**等处理算法。



Fig. 15. The device we used for the indoor experiment. It contains one forward-looking global shutter camera (MatrixVision mvBlueFOX-MLC200w) with 752×480 resolution. We use the built-in IMU (ADXL278 and ADXRS290, 100Hz) for the DJI A3 flight controller.

图3-1 VINS中设备型号

由此看来VINS选用的硬件是具有一定要求的，其采集数据的精度也对SLAM算法最终的效果产生正向作用。所以如果我们选用了精度没那么高的单目相机采集数据时，VINS的输出结果就不一定能够达到论文中的精度了。

3.2 设备类型选择

我们的项目中工作环境为室外，对传感器选型提出了较高要求。以RGBD相机为例，很多基于结构光或者TOF方案的深度摄像头在室外表现不佳，室外自然光的影响。个人初步倾向于选用双目或者RGBD（双目某些情况下可视为RGBD）+IMU的方案，主要理由如下：

- 1) 双目/RGBD+IMU的硬件，可在此基础上尝试基于单目/双目/RGBD/单目+IMU/双目+IMU等多种开源SLAM方案；反之单目+IMU的设备对制较大（只能是单目/单目+IMU）
- 2) 单目SLAM在建图方面，尤其是深度估计方面，对场景适应性不好且精度较差。双目/RGBD因为可以计算得到深度，在建图方面相对具有优势“稠密”地图
- 3) SLAM定位实现中，基于单目的方案其计算量也要比基于双目/RGBD要大，且因为引入了三角测量部分，对于环境适应性较差

所以在选型方面，我主要聚焦在双目/RGBD类型；同时为了加快开发进程，若选择能够提供开发SDK等工具的厂家，可以省去对于相机的标定、开发工作。

经过一些筛选后，个人列出如下备选设备。有关设备的参数信息在下方链接中都有详细说明，我只列出一些重要参数。

3.2.1 ZED



图3-2 zed双目相机

- 1) 官网：<https://www.stereolabs.com/zed/>
- 2) 基本参数：基于双目原理匹配计算深度信息，分辨率最高可达2.2K@15fps (4416X1242) ；Field of View: 90°(H) X 60°(V) X 110°(D)；室内外场景感知；使用GPU加速运算（支持TX1、TX2）
- 3) SDK：提供SDK，支持包括Windows、Linux、ROS等开发平台；使用SDK能够获得双目图像和深度图像、机器人位姿跟踪（官网称频率可达100Hz，精度达到1mm，角度精度达到0.1°）、3D重建等功能
- 4) 应用：在网上搜到一些评测和应用视频，主要内容总结如下。

A.视频1（见“视频/zed/Realtime depth test using ZED stereo camera.mp4”）表明zed在深度计算方面室外表现较好，距离方面应该可以度方面官方未给出参数，从双目原理上来说深度测距精度与测量距离的平方成正比，仅从zed输出的深度图上来看，深度图数据比较连续，未出



图3-3 zed在室外测量深度的表现

B. 3D重建也是SLAM的一种应用，使用SLAM输出的全局世界坐标结合RGB信息可实现真实世界中场景的三维绘制，所以3D重建的效果好坏能够在一定程度上反映设备的深度计算精度与SLAM效果。从官方视频和youtube上视频（见“视频/zed/Introducing ZED for Live 3D Mapping.mp4”，“视频/zed/Output for Graph-based RGB-D SLAM using ZED camera on UGV and UAV.mp4”，“视频/zed/ZEDfu - Real-time 3D Mapping using ZED stereo camera.mp4”）的测试效果来看，3D重建效果还是不错的，物体轮廓比较连续，未出现明显的边界不重合的情况。

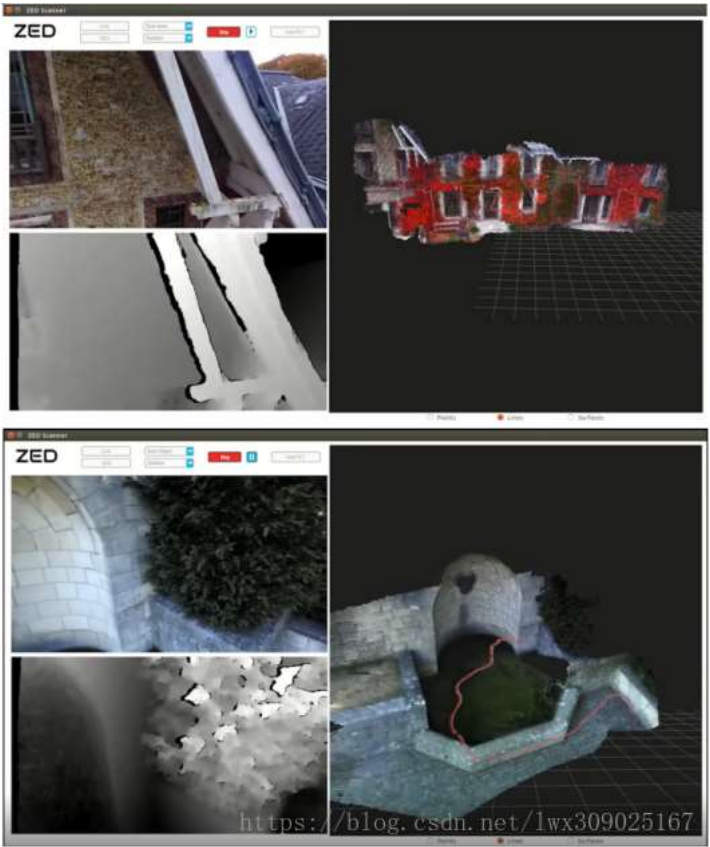


图3-4 使用zed进行三维重建

C.有研究人员在室外汽车上使用zed采集信息接入ORB-SLAM2计算车辆的位姿（见“视频/zed/Evaluation of ORB-SLAM2 in outdoor urban environment using stereo camera.mp4”）。测量结果表明（1）在停车场环境和街道环境下大都能够实现闭环检测（2）非极端情况下相机采集的图像满足特征跟踪要求。对于相机帧率要求越高，zed在WVGA分辨率下帧率可到100Hz，满足了图像采集频率要求。



图3-5 使用zed接入ORB_SLAM2进行室外定位

5) 价格：国外官网上为\$449，国内淘宝上价格3200-3800元不等

3.2.2 Intel D415/D435系列



图3-6 Intel D415/D435

- 1) 官网：<https://software.intel.com/zh-cn/realsense/d400>
- 2) 基本参数：基于主动红外测距（激光散斑增加纹理后立体匹配测距）；RGB分辨率和帧率为1920X1080@30fps；内置视觉处理器可直接输出计算输出分辨率和帧率为1280X720@90fps；**室内外检测范围为0.16m-10m**；D415为卷帘快门，D435为全局快门

	英特尔实感深度摄像头 D415	英特尔实感深度摄像头 D435
环境	室内和室外	室内和室外
深度技术	主动红外 (IR) 立体	主动 IR 立体
图像传感器技术	卷帘快门：像素大小为 1.4 微米 x 1.4 微米	全局快门：像素大小为 3 微米 x 3 微米
主要英特尔® 实感™ 产品	英特尔® 实感™ 视觉处理器 D4 英特尔® 实感™ 模块 D410	英特尔® 实感™ 视觉处理器 D4 英特尔® 实感™ 模块 D430
深度视场角 (FOV) — (横向 × 纵向)，适用于高清 1080p	63.4° x 40.4° (+/- 3°)	65.2° x 58° (+/- 3°)
深度流输出分辨率	达 1280 x 720	达 1280 x 720
深度流输出帧速率	达 90 fps	达 90 fps
最小深度距离 (Min-Z)	0.16 米	0.11 米
最大范围	约 10 米 准确度随校准、场景和光照条件而变	约 10 米 准确度随校准、场景和光照条件而变
RGB 传感器分辨率和帧速率	30 fps 时为 1920 x 1080	30 fps 时为 1920 x 1080
RGB 传感器视场角 (横向 × 纵向)	69.4° x 42.5° (+/- 3°)	69.4° x 42.5° (+/- 3°)
摄像头尺寸 (长度 × 厚度 × 高度)	99 毫米 x 20 毫米 x 23 毫米	90 毫米 x 25 毫米 x 25 毫米
连接器	C 型 USB®	C 型 USB®
挂载机构	一个 1/4-20 UNC 螺纹挂接点 两个 M3 螺纹挂接点	一个 1/4-20 UNC 螺纹挂接点 两个 M3 螺纹挂接点

图3-7 D415/D435详细参数对比

- 3) SDK：支持Linux/windows/Mac OS，可获得彩色图像与深度图像，设置相机参数等
- 4) 应用：根据网上相关测评视频整理如下

A.有人对比了D415和D435在室内的表现（见“视频/Intel/Intel RealSense D435 vs D415.mp4”），摄像头如图3-8所示放置。分别比较了彩色图、近距离深度图的效果，具体如图3-9所示。结果表明D435在远距离深度图上效果要优于D415，后者出现较大面积的空洞（计算不出深度信息）。**摄像头的测量环境并不完全一致，同时存在发射的红外光互相干涉影响的可能性，该结论是否成立不予保证。**



图3-8 D435与D415对比

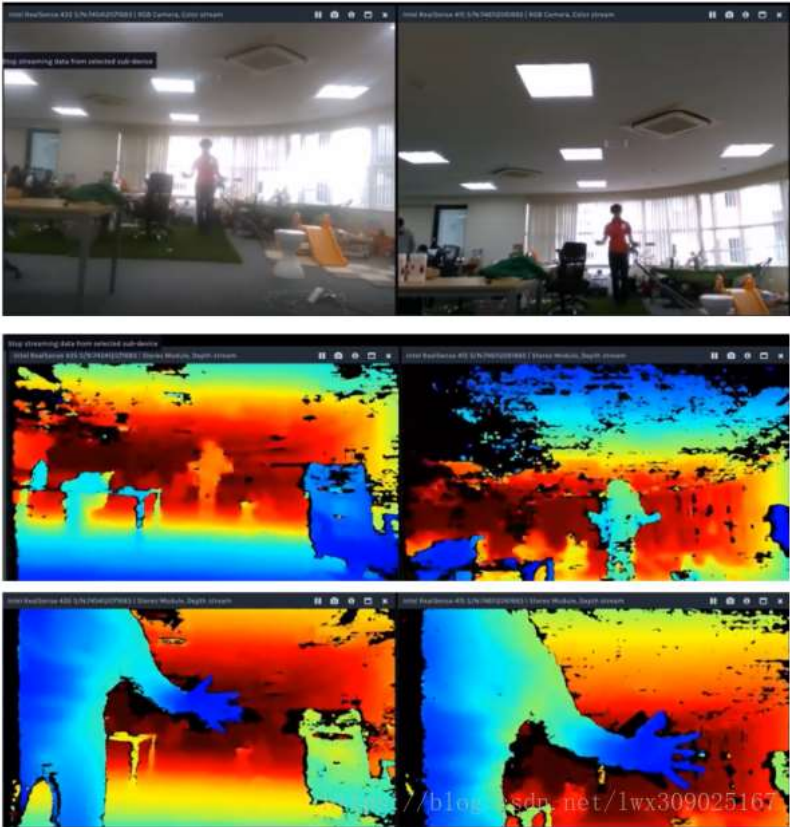


图 3-9 D435（左边）与D415（右边）室内表现效果对比

B.有人在室外街道中使用D415观察其深度信息（见“视频/Intel/Review on Intel Realsense D415 RGBD Camera Part 2_ Outdoor test.mp4”），3-10所示。结果表明：

- (1) 室外D415的深度检测距离应该可以达到10m
- (2) 相对于室内，室外的深度图噪点较多，这应该是自然光对于D415发射的红外光的影响，造成图像匹配失败或者误匹配

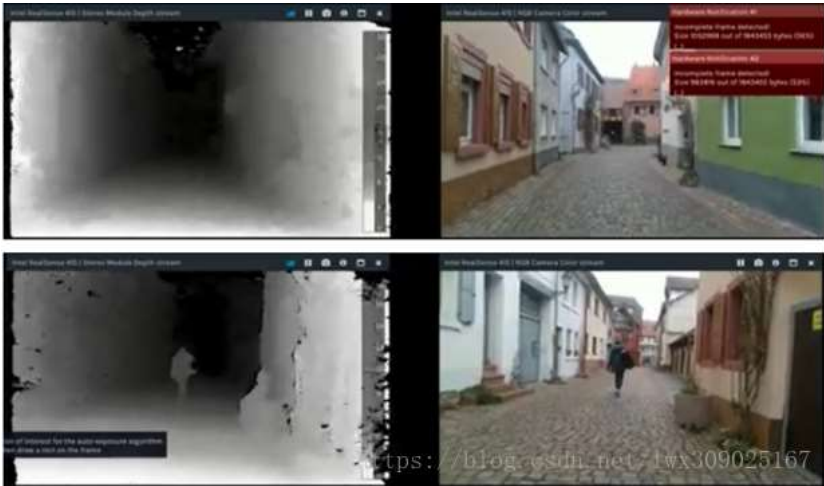


图3-10 D415在室外的深度测距表现

2019/8/6

(2条消息)视觉SLAM方案整理及硬件选型调研 - On my way - CSDN博客

C.有人将D435固定在汽车上查看道路的深度信息（见“视频/Intel/Realsense D435 - Outdoor test on road.mp4”），这里截取几幅图像如图3-11出：

(1) D435在室外能够获得比较丰富的深度信息

(2) D435深度计算错误/失败的情况多于zed

5) 价格：官网上D435为\$179.00，国内价格¥1500不等；官网上D415为\$149.00，国内淘宝价格¥1300不等

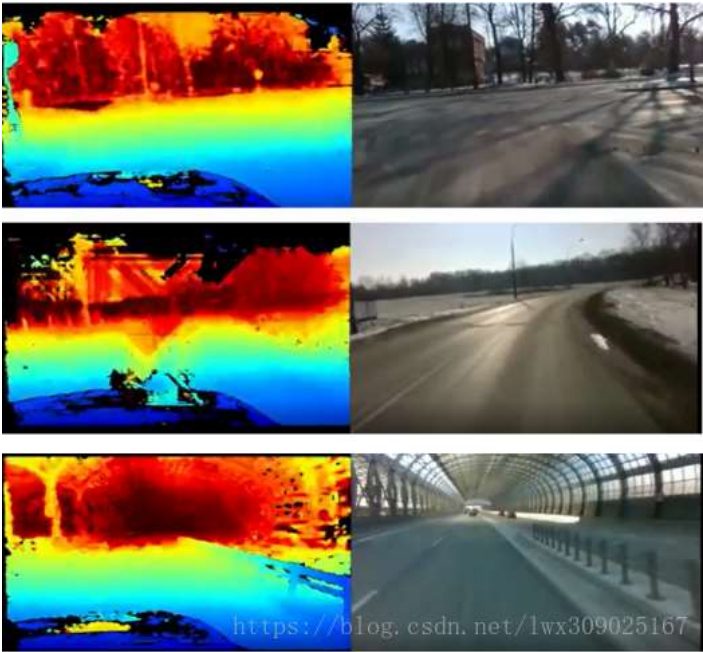


图3-11 D435在室外深度测距表现

3.2.3小觅双目摄像头

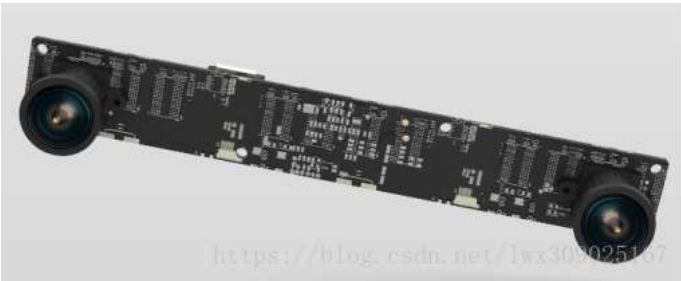


图3-12 小觅双目摄像头外观

- 1) 官网链接 <http://www.myntai.com/camera>
- 2) 基本参数：基于双目测距，内置6轴IMU（ICM2060，消费级IMU）。有常规版（即被动双目）和IR增强版（原理应该是激光散斑增加纹理后立体黑白CMOS，分辨率752X480@50FPS。

双目摄像头产品详情

产品规格		
型号	MYNT EYE S1010	MYNT EYE S1010-IR
尺寸	PCB dimension145mm*20mm; total dimension 145*20*28.6 (depend on lens dimension)	PCB dimension145mm*20mm; total dimension 145*20*28.6 (depend on lens dimension)
帧率	50FPS	50FPS
分辨率	752x480@50 FPS	752x480@50 FPS
深度分辨率	Base on CPU/GPU Up to 752x480@50FPS	Base on CPU/GPU Up to 752x480@50FPS
像素尺寸	6.0 x 6.0μm	6.0 x 6.0μm
快门速度	17 millisecc	17 millisecc
基线	120.0 mm	120.0 mm
镜头	Replacable Standard M12	Replacable Standard M12
IR主动光	无	有
视角	D:120° H:96° V:61°	D:92° H:65° D:40°
焦距	2.6mm	4.2mm
滤镜	Tavg>=88%@440-620nm T=50%@648±10nm Tavg<=3% @700-1000nm T<5%@1050nm	无
运动感知	6 Axis IMU ICM2060	6 Axis IMU ICM2060
色彩模式	Monochrome (单色)	Monochrome (单色)
控制功能	Exposure/Shutter/Brightness	Exposure/Shutter/Brightness
扫描模式	Global shutter	Global shutter
功耗	1.4W @ 5V DC from USB	2.45W @ 5V DC from USB
同步精度	<1ms	<1ms
深度图	室外效果更好	室内效果更好
最高图像帧率	50FPS	50FPS
最高IMU频率	500 Hz (Adjustable)	500 Hz (Adjustable)
输出数据格式	Raw data	Raw data
接口	USB 3.0	USB 3.0
重量	46.7g	48.9g
UVC MODE	YES	YES

<https://blog.csdn.net/lwx309025167>

图3-13 小觅双目摄像头详细参数

3) SDK: 支持Windows、Linux、TX1、TX2; 基于CPU/GPU计算深度信息; 提供接入了OKVIS、ORB-SLAM2、VINS、VIORB (都是开源SLAM工sample

4) 应用: 目前尚未找到相关评测信息, 不过据知乎上一些问题的回答, 该模块的售后支持不错。深度图效果方面, 只有天猫旗舰店上提供的一小段室内(见“视频/MYNT-EYE/室内景深测试.mp4”), 截图如图3-14所示。个人感觉深度效果计算较差, 有较多的计算错误点出现; 无法判断深度范围大, 称的20m。



图3-14 MYNT-EYE室内深度测试效果

5) 价格: 天猫¥1999.00

现在对以上几款设备进行一个信息的对比整理, 如表3-1所示。

型号	CMOS	分辨率/帧率	视角（角度制）	快门类型	IMU	深度
Zed	彩色	side-by-side 4416X1242@15fps 3840X1080@30fps 2560X720@60fps 1344X376@100fps	90(H)X60(V)X110(D)	未知	无	被z
D435	彩色	1920X1080@30fps	85.2(H)X58(V)	全局快门	无	近
D415	彩色	1920X1080@30fps	63.4(H)X40.4(V)	卷帘快门	无	近
MYNT-EYE	黑白	side-by-side 1504X480@50fps	96(H)X61(V)X120(D)	全局快门	ICM2060	被z
型号	深度范围	深度分辨率/帧率	SDK	其他		价
zed	0.5m-20m	Baed on CPU/GPU 最高与图像分辨率/帧率一致	支持	机器人位姿跟踪（官网称频率可达100Hz，位置精度达到1mm，角度精度达到0.1°）		\$4
D435	10m	分辨率最高720P；帧率最高90fps	支持	内置运算芯片直接输出深度数据		\$1
D415	10m	分辨率最高720P；帧率最高90fps	支持	内置运算芯片直接输出深度数据		\$1
MYNT-EYE	20m（未验证）	Baed on CPU/GPU 最高与图像分辨率/帧率一致	支持	提供接入了OKVIS、ORB-SLAM2、VINS、VIO RB（都是开源SLAM工程）的sample		¥1

表3-1 三款设备主要信息对比

3.3 个人意见

整体看来，个人**首先推荐zed**，**其次推荐DM435和MYNT-EYE**。若选用zed或者D435后准备测试融合了IMU的开源SLAM方案时，可使用IMU模
息，之后自己做数据同步。

