

关于深度相机的精度问题

CASIA

常用的三种类型的深度相机^[1]，大致分为以下三种：基于主动投射结构光的深度相机（如Kinect 1.0, Intel RealSense, Enshape, Ensensio等）、被动双目相机（如STEROLABS推出的 ZED 2K Stereo Camera, Point Grey 公司推出的 BumbleBee）以及ToF^[2]相机（如微软的Kinect 2.0, MESA 的 SR4000，Google Project Tango 中使用的PMD Tech 的TOF相机，Intel 的 SoftKinect DepthSense, Basler基于松下的芯片开发的TOF相机以及国内一些初创公司基于TI的方案开发的TOF相机等等）

- 目前市面上常有的 3D 相机方案就就是下面3种，对应上面的：
- (1) 结构光(Structured-light)，代表公司有奥比中光（国产，比较推荐，性价比可以，也有高分辨率的款），苹果(Prime Sense)，微软 Kinect-1，英特尔RealSense, Mantis Vision 等。
 - (2) 双目视觉(Stereo)，代表公司 Leap Motion，ZED，大疆;
 - (3) 光飞行时间法(TOF)，代表公司微软 Kinect-2，PMD，SoftKinect，联想Phab。

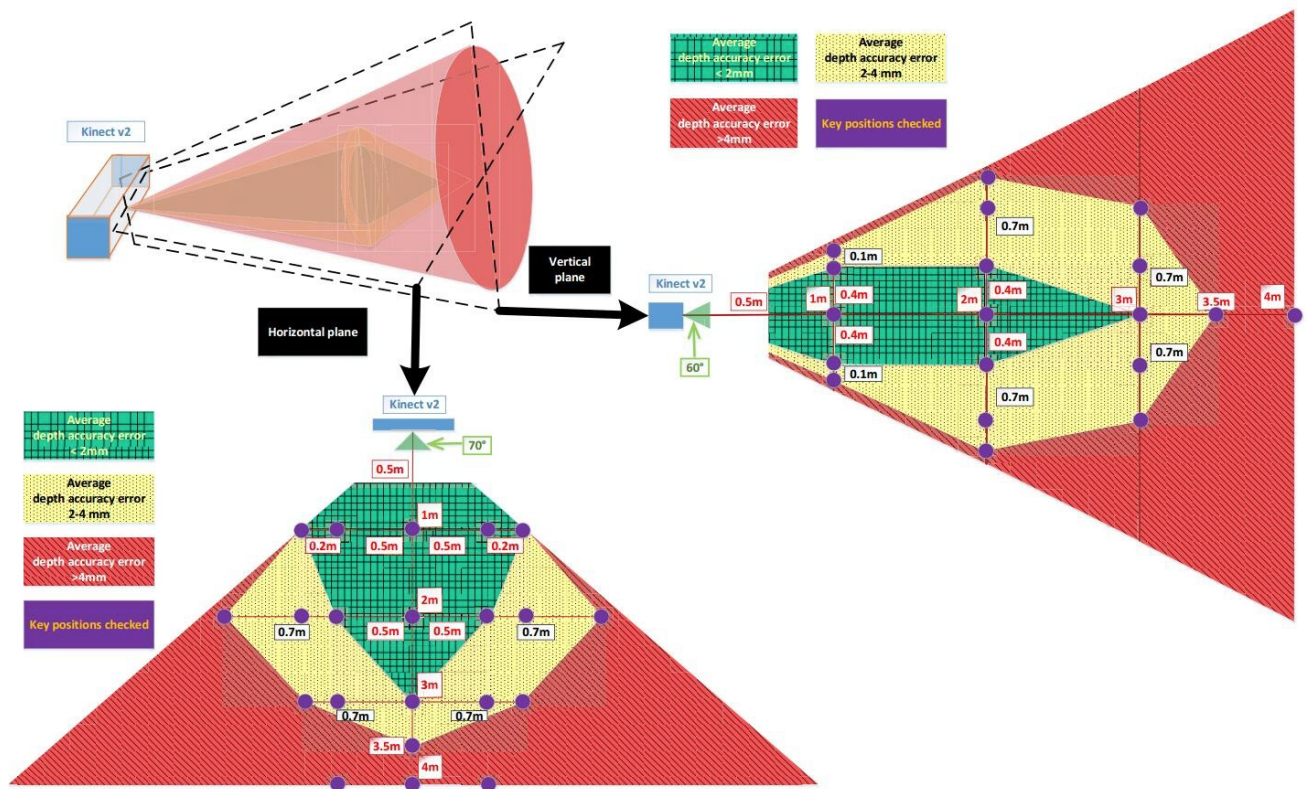
性能指标	主动投射结构光	被动双目视觉	TOF
测量范围	0.1m ——10m	中距	0.1m ——100m
精度	短工作范围内能够达到高精度 0.01mm——1mm	短工作范围能够达到高精度 0.01mm——1cm	典型精度(1cm)
软件复杂度	中	非常高	低
帧率	较低，几十Hz	低到中 <百Hz	可以做到非常高
户外工作情况	影响较大，功率小的时候基本无法工作	无影响	有影响但较低，功率小的时候影响较大
黑暗环境能否工作	可以	不可以	可以
价格	随精度价格不同： 1mm级精度千元量级； 0.1mm级万元量级； 0.01mm级几十万量级；	非常便宜：千元量级	依测量范围、帧率不同： 几千-几百万
典型应用场景	1.近距体感识别、VR、AR； 2.近距物体识别、姿态检测、测量等；	尚不明朗	1.近距体感识别、VR、AR； 2.近距物体识别、姿态检测、测量等； 3.远距环境建模、物体识别、测量等；

1. Kinect

先给出结论，KinectV2的精度在2mm~4mm左右，距离越近精度越高，越远精度稍微差点；kinectV1误差约2mm~30mm。

Kinectv2 for Mobile Robot Navigation: Evaluation and Modeling

- Kinect v2在不同位置的精度问题



如上图所示，右侧大三角是KinectV2在纵向（垂直于地面）的精度分布，下侧大三角是KinectV2在水平面（平行于地面）上的精度分布。在绿色区域精度最高，误差小于2mm，在黄色区域误差在2~4mm，红色区域误差大于4mm。所以在设计交互场景时，在黄色区域以内会达到最好的效果（3.5m内）。如果对精度要求很高，如控制机械，最好在绿色区域进行交互。

2 Kinect v2和Kinect v1

		Kinect v1	Kinect v2预览版
颜色 (Color)	分辨率 (Resolution)	640×480	1920×1080
	fps	30fps	30fps
深度 (Depth)	分辨率 (Resolution)	320×240	512×424
	fps	30fps	30fps
人物数量 (Player)		6人	6人
人物姿势 (Skeleton)		2人	6人
關節 (Joint)		20關節／人	25關節／人
手的開閉狀態 (Hand State)		△ (Developer Toolkit)	○ (SDK)
检测範圍 (Range of Detection)		0.8～4.0m	0.5～4.5m
角度 (Angle) (Depth)	水平 (Horizontal)	57度	70度
	垂直 (Vertical)	43度	60度
(Tilt Motor)		○	× (手動)
複数的App		× (单一的App)	○

Kinect v2的rgb视场 (FOV^[3]) 是84.1 x 53.8 , 关于FOV的建模和模型可以[参考](#)。

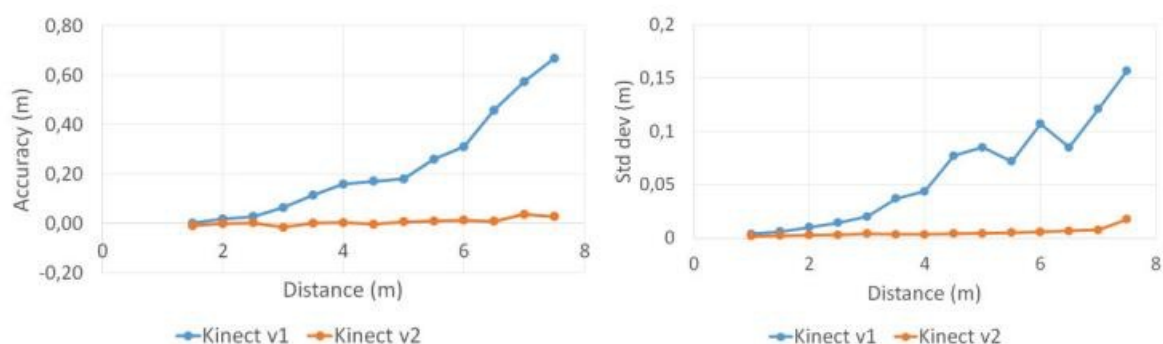


Fig. 9. a) Depth accuracy and b) depth standard deviation of Kinect v1 and v2 at all ranges.

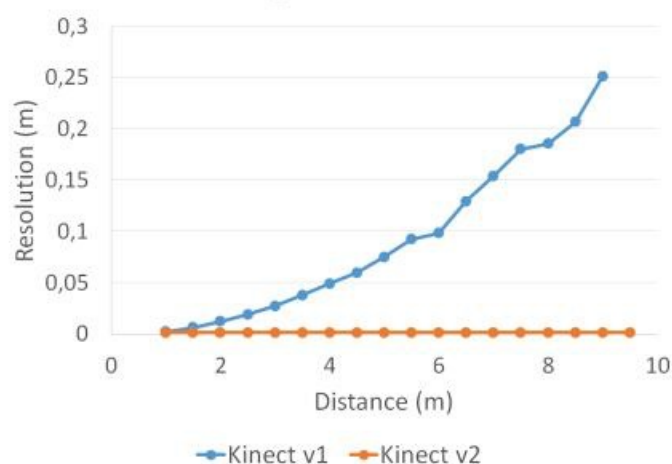


Fig. 10. Depth resolution of Kinect v1 and v2 at all ranges.

如图所示，KinectV1随着距离增加，误差指数性增加，在距离达到4m时，kinectV1误差接近**0.2m**。而KinectV2的误差几乎不随距离增加而变化。V2比V1追踪准确度好20%。V2可以在户外进行人体跟踪，最远到4m。V2在近距离有比V1高2倍的精度，在6m有高数十倍的精度。

- kinectv1和kinectv2比较

-	KinectV1	KinectV2
检测范围(Range of Detecton)	0.8–6.0m	0.5 –4.5m
深度误差(depth Uncertainty)	2mm-30mm	<0.5% of range
角度(Angle , horizontal-vertical)	57-43	70-60

3. LeapMotion

LeapMotion的精度平均下来是**0.7mm**的精度，也是达不到所谓的0.01mm的。

Analysis of the Accuracy and Robustness of the Leap Motion Controller

上面的论文对初步版本中的Leap Motion控制器进行研究，分别在静态与动态设置下的精度和准确性，考虑到人手的可达到的平均约为0.4mm，实验用设备使用参考笔，位置精度可达0.2mm，且参考笔对仪器精度测量无可观察到的影响。在基于静态设置的测量下，获得了期望的3D位置与**小于0.2mm**的测量位置之间的与轴无关的偏差。在动态情况下，独立于平面，可以获得**小于2.5mm**的精度（平均1.2毫米）。重复性**平均小于0.17毫米**。在基于姿势的用户界面方面，在实际条件下不可能实现0.01mm的理论精度，而是高精度（总平均精度为**0.7mm**）。

最后比较一下以上设备的优缺点

1. Microsoft Kinect



优点：

- 可以获取深度数据 (320 * 240)、rgb
- 数据 (640 * 480)、声音、骨骼节点 (20个)
- 拥有三套 SDK : 微软 SDK、OpenNI、libfreenect
- 后两个 SDK 是跨平台, 支持各种开发语言
- 价格便宜
- 社区成熟, 开发资源丰富

缺点:

- 传感器分辨率不够, 看不清手指
- 由于使用结构光技术, 深度传感器的可视范围无法重叠
- OpenNI 和 libfreenect 虽然跨平台, 但是功能远不如微软 SDK
- 设备尺寸大, 需要一坨电源线
- 致命缺点, 微软已宣布停止生产 Kinect 一代

2. Microsoft Kinect One



优点:

- 分辨率更大、可以看到更广阔的场景

- 可以检测到的人体关节更多（25个），能看清手指
- 拥有两套 SDK：微软 SDK、libfreenect2
- 可以开发 Windows Store 应用

缺点：

- libfreenect2 基本不能检测骨骼，功能缺太多，同时 OpenNI 也不支持它，因此局限于 Windows 平台
- 设备尺寸比一代更大，需要一坨电源线，比一代贵一些
- 致命缺点：只能运行在 64 位 Windows 8 系统上，必须使用 USB 3.0 端口

3. Intel / Creative / SoftKinetic



优点：

- 小巧，普通 USB 摄像头的尺寸
- 不需要外界电源线
- 近距离使用，可实现表情分析和手势识别

缺点：

- 不适合远距离交互，也无法检测完整的身体

- 只能在中高端的 Intel CPU 上才能运行

4. Leap Motion



优点：

- 小巧，一根 usb 线就可以使用
- 跨平台
- 支持的开发语言比较多，甚至通过 WebSocket
- 实现了浏览器中的 JavaScript API
- 跟踪手指和手掌，精度较高

缺点：

- 检测范围小，手臂酸疼（见上图）
- 不能检测身体和脸部
- 作为生产力工具，完全无法替代鼠标键盘
- 致命缺点：找不到合适的使用场景

5. PrimeSense / Apple / 华硕 (ASUS)



和 Kinect 一代的优缺点类似，

[1] <https://zhuanlan.zhihu.com/p/28274727> ↵

[2] <https://baike.baidu.com/item/TOF/19952376?fr=aladdin> ↵

[3] <http://www.coloreye.cn/wiki/doc-view-716.html> ↵