51CTO WOTI

当人工智能遇上虚拟现实

马赓宇 博士

uSens 凌感技术研发副总裁

uSens Inc

VR/AR中的核心技术

VR/AR中的核心技术

VR/AR Interaction

- HCI
  - Gesture
  - View Direction
  - Position Tracking
- Display
  - AR
  - VR

VR/AR未来需要解决的三个问题

01 产品碎片化

02 内容缺乏

03 技术进步

显示技术

01 FOV

02 分辨率

03 硬件尺寸

交互技术

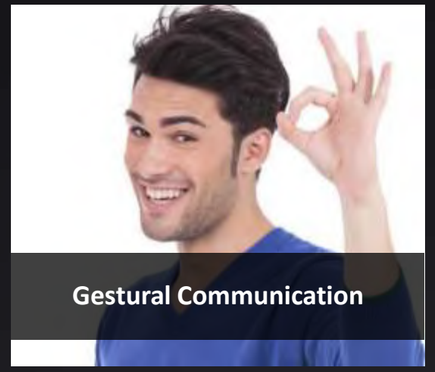
Type of Interaction

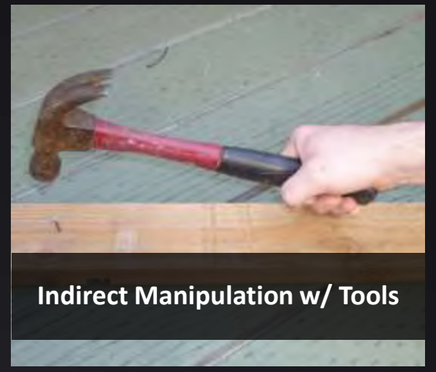
Gestural
Communication

indirect Manipulation
w/tools

Direct
Manipulation

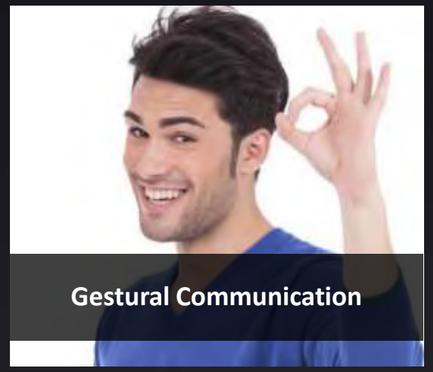
交互技术

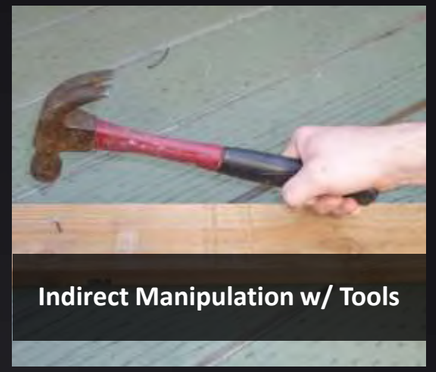




Game Pad	↓	↓	↓
6DOF Controller	↓	↑	↓
3D Hand Tracking	↑	↑	↑

交互技术

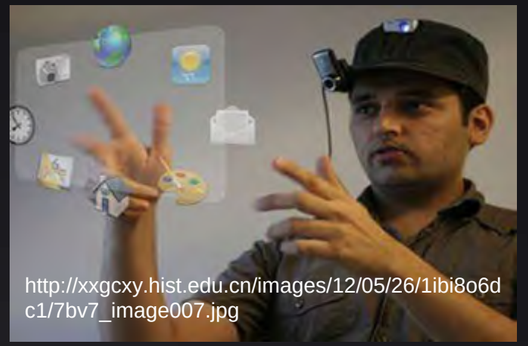




HoloLens	↑	↓	↓
uSens Fingo	↑	↑	↑

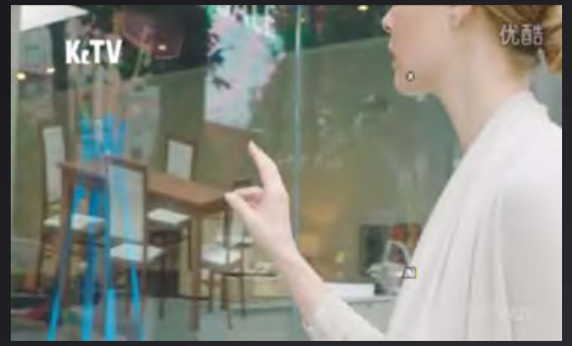
交互技术

Game



3D UI

Drive



Public Display

Medical

Intelligent Home

基于人工智能的交互技术

Deep Learning

- Use CNN network, input is depth map, output heat maps of every joint.

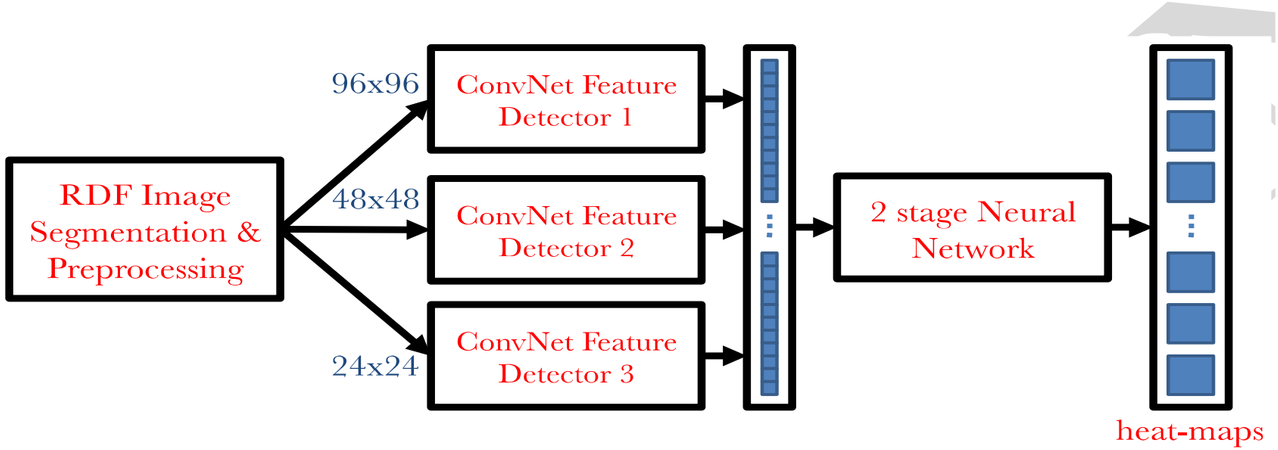


Fig. 6: Convolutional Network Architecture

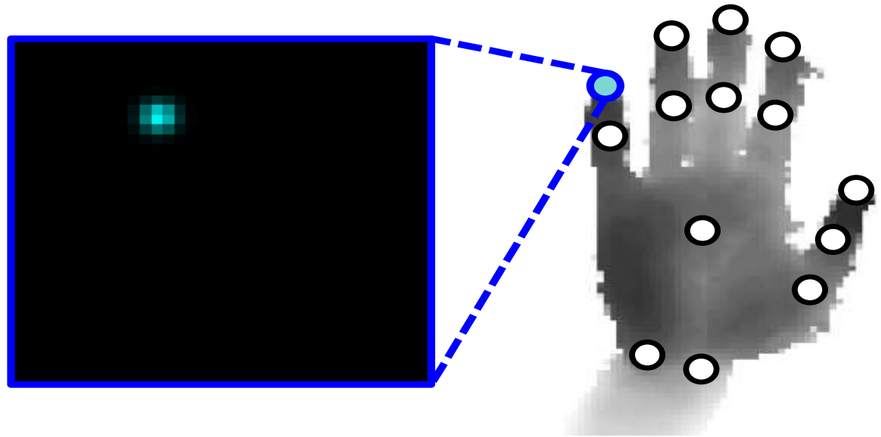


Fig. 5: Depth image overlaid with 14 feature locations and the heat-map for one fingertip feature.

Jonathan Tompson, etc. Real-Time Continuous Pose Recovery of Human Hands Using Convolutional Networks. Siggraph 2014

空间定位技术

空间定位技术

- Google Tango
- Microsoft Hololens
- Qualcomm VRSDK
- Apple ARKit
- Snapchat
- Facebook
- uSens

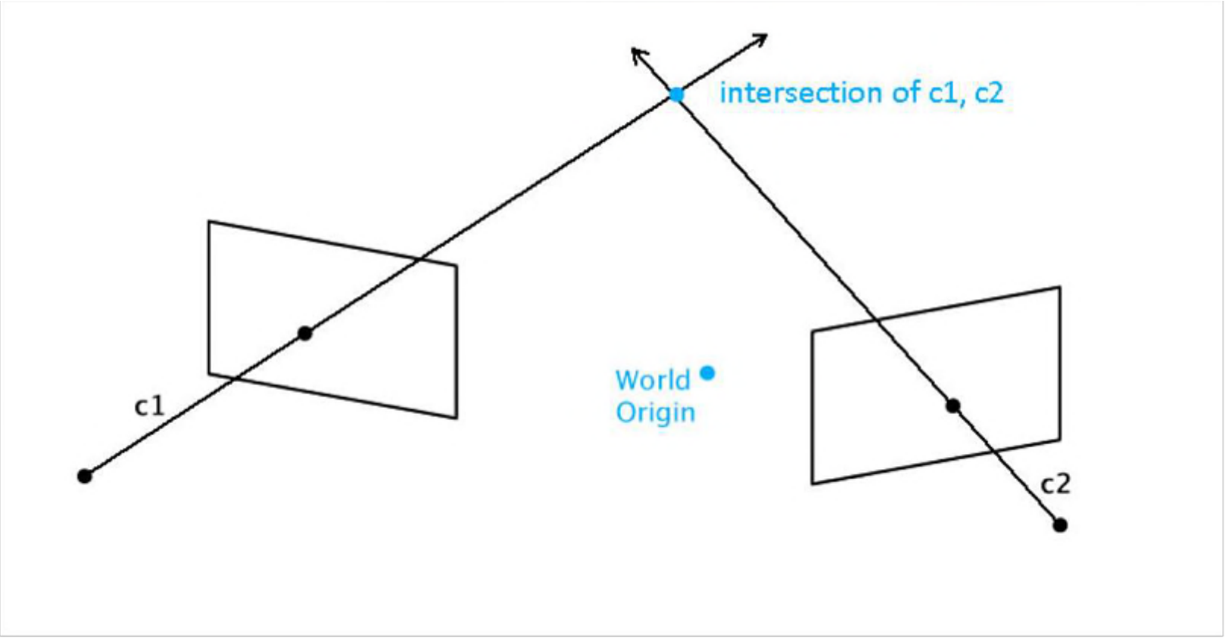
空间定位技术/视觉方法

空间定位
- 已知：3D点 (x,y,z)，2D投影 (u,v)
- 未知：相机姿态 (R，T)

三角化
- 已知：相机姿态 (R, T)
- 未知：3D点 (x,y,z)

SLAM
- 已知：2D投影 (u,v)
- 未知：R、T、(x,y,z)

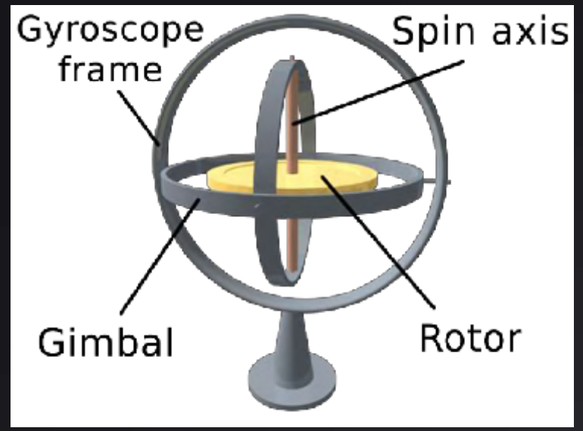
$$\begin{pmatrix} u \\ v \end{pmatrix} \sim \begin{pmatrix} f & & cx \\ & f & cy \\ & & 1 \end{pmatrix} \cdot \left[R \cdot \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} - T \right) \right]$$



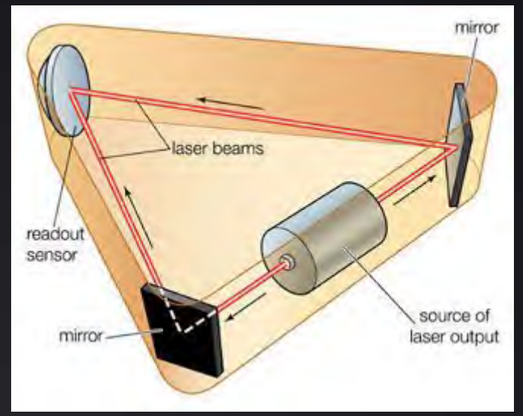
空间定位技术/视觉方法

视觉SLAM方法包含两个模块	• Tracking：已知3D点位置，求解每帧图像的相机pose • Mapping ：更新3D点的位置
两种视觉特征	• 基于图像特征点的方法，如PTAM, ORB • 直接法，比较像素灰度差，如LSD-SLAM，DSO-SLAM
两种优化求解方法	• 基于滤波的算法 • 基于优化的算法

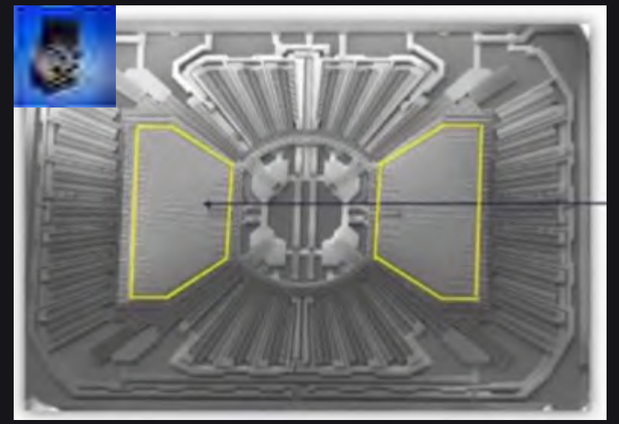
空间定位技术/传感器



Mechanic Gyroscope

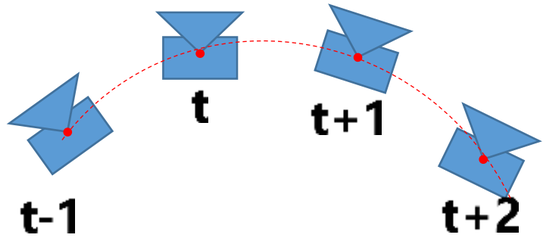


Laser Gyroscope



MEMs Gyroscope

空间定位技术/传感器

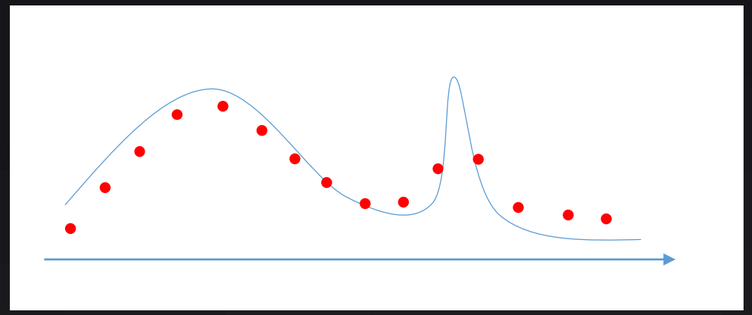


- IMU中的陀螺仪，输出的是相邻时刻相机旋转的角度
- IMU中的加速度计，输出的是相邻时刻相机的加速度，即速度的变化率

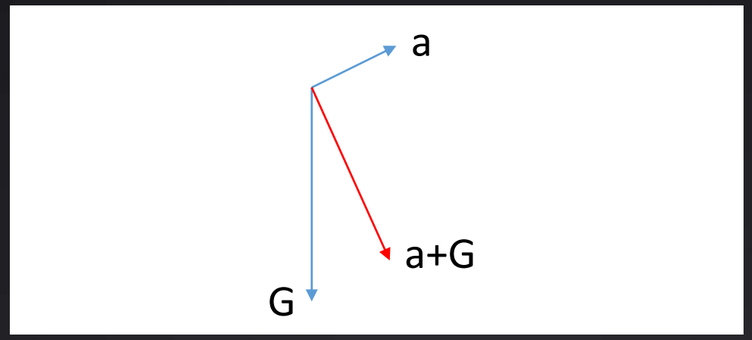
$$
\begin{aligned}
\mathrm{R}(t+\Delta t) &= \mathrm{R}(t)\,\mathrm{Exp}\left(\left(\tilde{\boldsymbol{\omega}}(t)-\mathbf{b}^{g}(t)-\boldsymbol{\eta}^{gd}(t)\right)\Delta t\right)\\
\mathbf{v}(t+\Delta t) &= \mathbf{v}(t)+\mathbf{g}\Delta t+\mathrm{R}(t)\left(\tilde{\mathbf{a}}(t)-\mathbf{b}^{a}(t)-\boldsymbol{\eta}^{ad}(t)\right)\Delta t\\
\mathbf{p}(t+\Delta t) &= \mathbf{p}(t)+\mathbf{v}(t)\Delta t+\frac{1}{2}\mathbf{g}\Delta t^{2}\\
&\quad+\frac{1}{2}\mathrm{R}(t)\left(\tilde{\mathbf{a}}(t)-\mathbf{b}^{a}(t)-\boldsymbol{\eta}^{ad}(t)\right)\Delta t^{2}, \qquad (31)
\end{aligned}
$$

States：R（Rotation），p（Position），v（Velocity），b（bias）

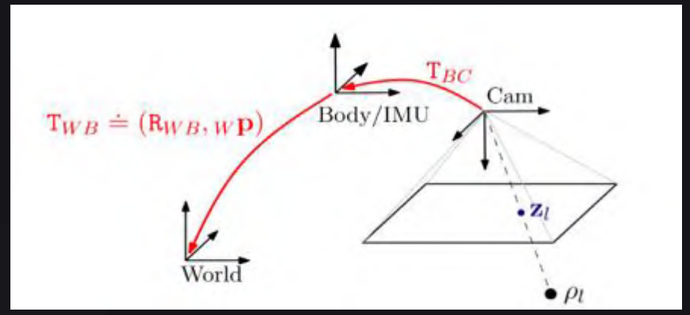
空间定位技术/传感器



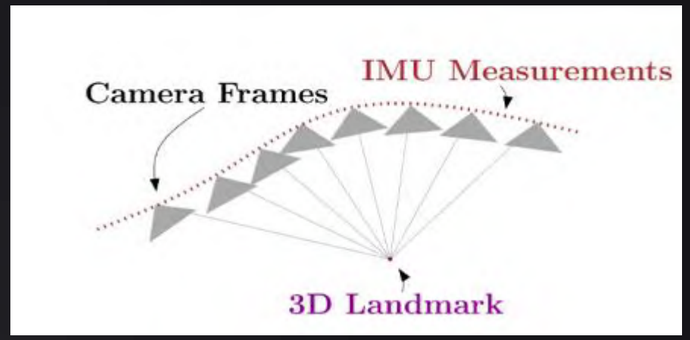
1. IMU数据是有偏差和漂移的



2. IMU得到的加速度是包含重力的



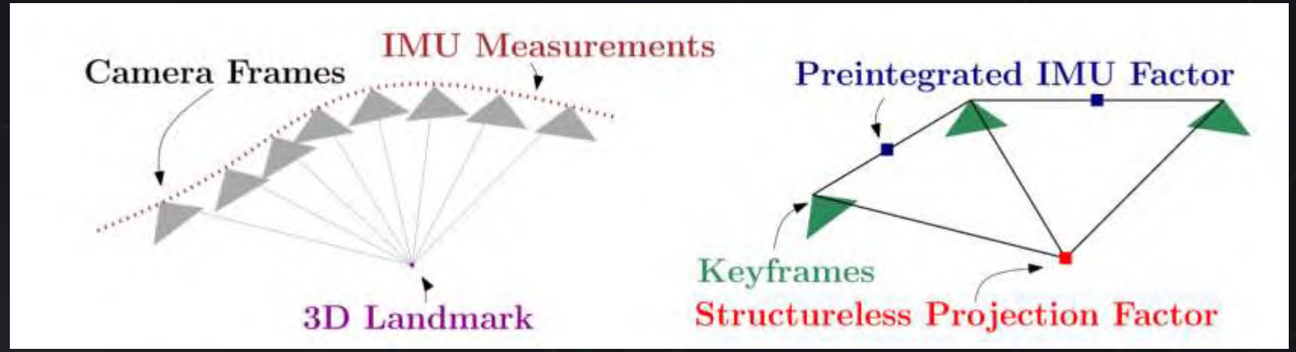
3. IMU和相机的相对位置和方向影响很大



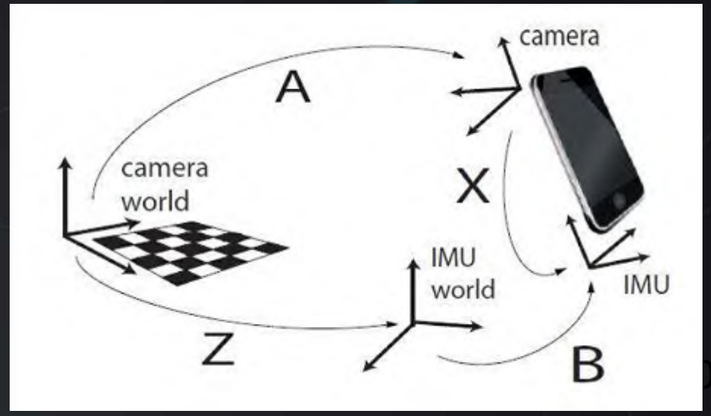
4. IMU和图像的采集时间不一致

空间定位技术/传感器+视觉融合

- IMU frequency is much higher than camera. Apply Pre-Integration technology

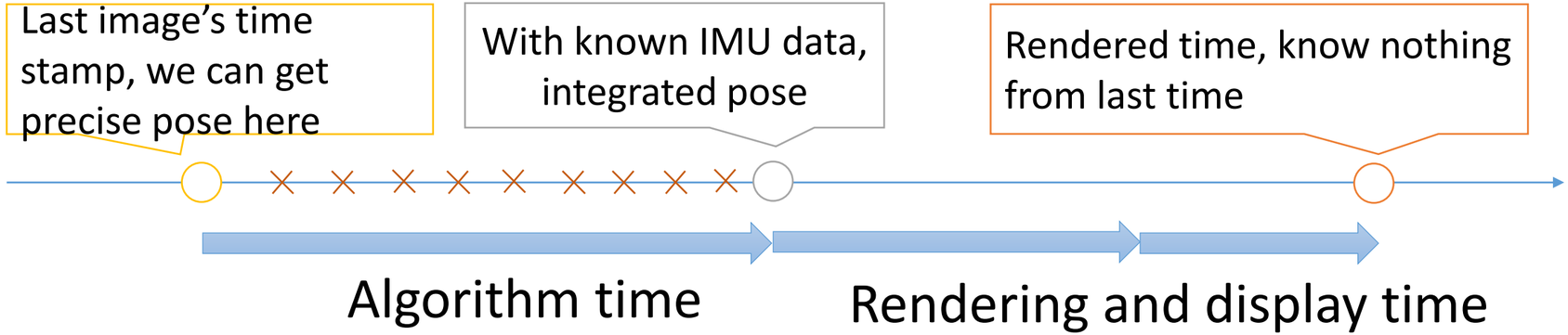


- IMU-Camera calibration is VERY important



空间定位技术/ATW

- 相机图像采集, SLAM算法, 渲染, 这些过程都需要处理时间，统称为“Motion to Photon Latency”。VR应用中，需要小于20ms，才能保证用户不会眩晕
- ATW 是一种有效减小延迟的技术, 它是通过预测未来的双眼位置来提前渲染



Hardware

2013

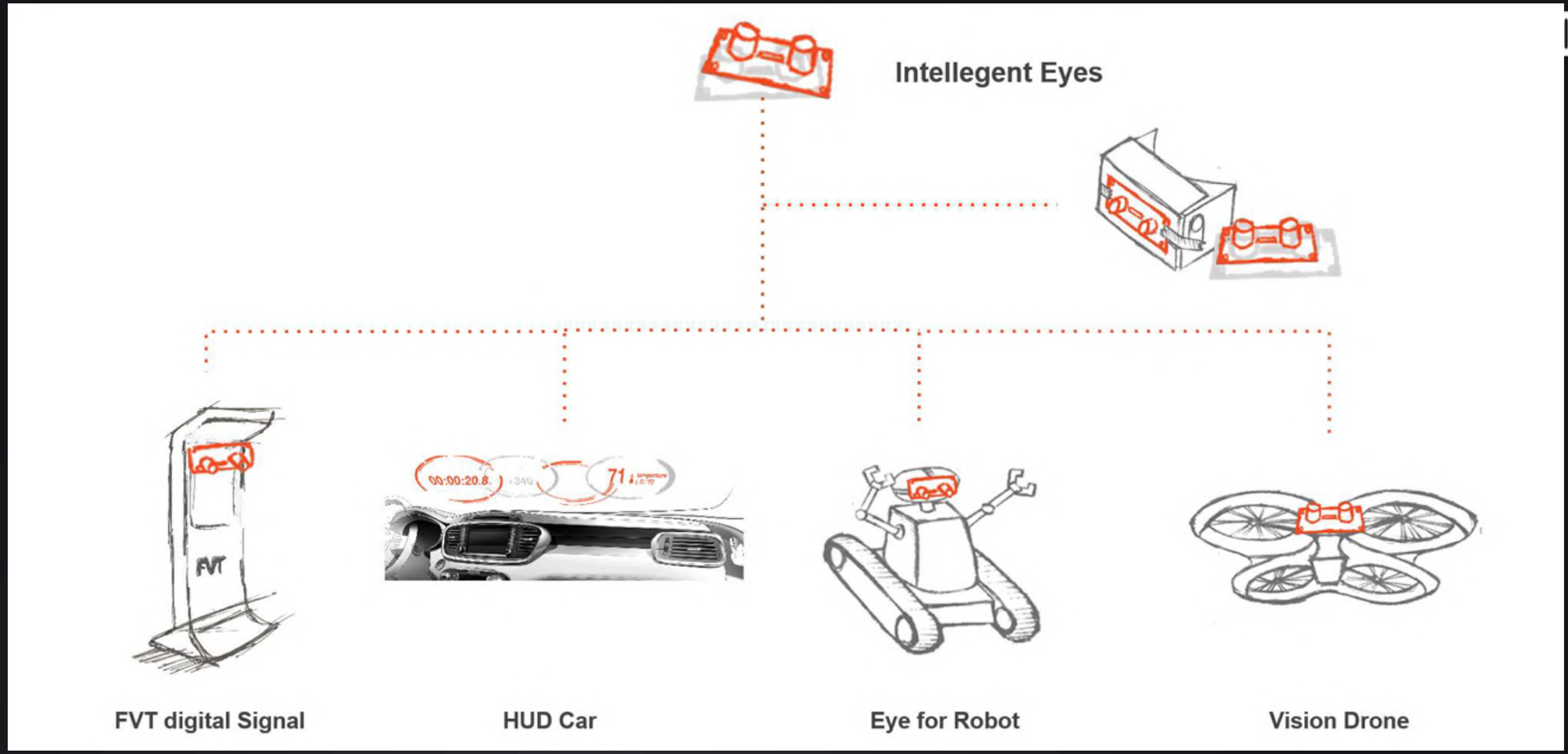
2015

2016

2017

Hardware Spec.

Product Dimensions	• V1 12.9mm(H) × 37.5mm(W) × 102.4mm(L) • V2 6mm x 15mm x 70mm
PCB Dimensions	• V1 8.4mm(H) × 24mm(W) × 85mm(L) • V2 5mm(H) × 12mm(W) × 65mm(L)
Image Sensor	• VGA, Global Shutter • Hardware Sync. • Frame Rate: 60 fps ~ 120 fps • Shutter Speed: 0.01ms ~ 10ms
Viewing Angle	• Lens: Horizontal: 140° Vertical: 120°
Power Consumption	• Active Motion Mode: 1.0W • Non-tracking Mode: 0.6W
Data Interface	• Micro-B SS • USB 2.0 / USB 3.0
Illumination	• 3 × 850nm LED • 3 × 33mW average


Intellegent Eyes
FVT
00:00:20.8
71
FVT digital Signal
HUD Car
Eye for Robot
Vision Drone

51CTO WOTI

THANK YOU

uSens Inc