3D Sensor

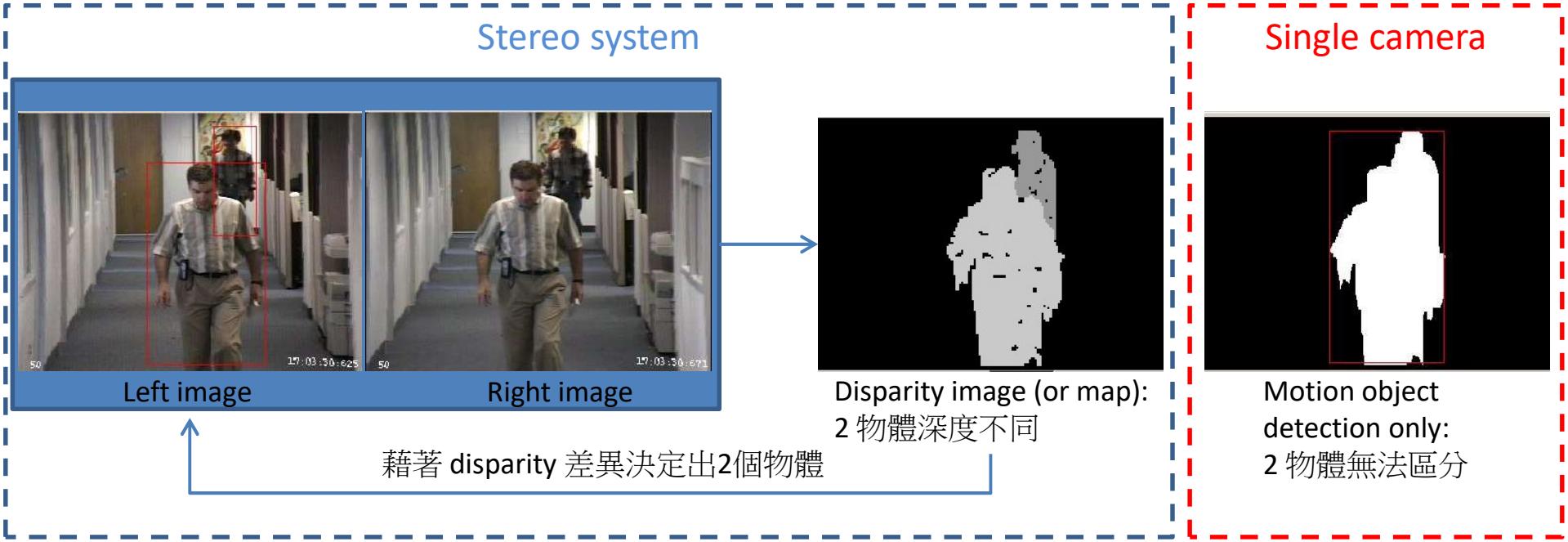
Major Issues

1. 3 kinds of 3D sensors:
   1) Stereo,
   2) Time-of-Flight (ToF), and
   3) Structured light

2. Detection Vs. Segmentation: Example cancer
   (1) Detection
      > Region Of Interest
      > Faster R-CNN
   (2) Segmentation =>
      > Pixel-Wise
      > Mask R-CNN

1. Introduction - Motivation

- Motivation:
  1. Single camera:
     i. 僅能拍攝場景物體的位置(2D).
  2. Stereo system:
     i. 藉由左右2顆攝影機的視差所求得的對應位置的 disparity，可用來估計拍攝場景中的物體的深度差異(3D)。
     ii. 利用 disparity image 來實現場景中的 object segmentation



1. Demo Video:

1) 3d_ArrayCamera_3dPrint_Pelican_20130620.fly

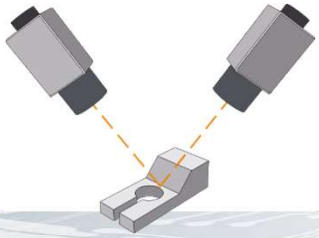
2.0.1 3D Vision Technology

- Three 3D Vision Technologies:
  1. Stereo: Can embed the stereo system into TI DM8148/8168.
  2. Time of Flight (TOF): Supported by TI DM365 VICP (current: 200x200 TOF sensor).
     - Kinect 2
  3. Structured Light
     - Kinect1, DLP

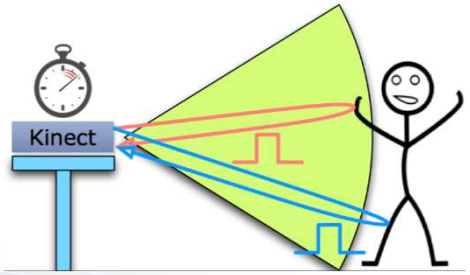
Stereo Vision

Time of Flight (TOF)

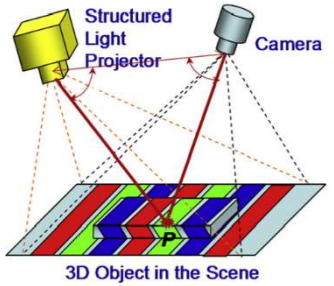
Structured Light



1) Stereo Vision



2) Time of Flight



3) Structured Light

2.0.1 3D Vision Technology – Comparison

- Three 3D Vision Tech. Comparison (Showed in TI confidential info.).

	Stereo Vision	Structured Light	Time of Flight (TOF)
SW Complexity	Middle	High	Low
Material cost	Low < NT$8K	High > NT$120K	Middle NT$4K~5K
Response Time	Middle 30 fps (33 ms/f)	Slow (1 sec/frame)	Fast 30~60 fps
Accuracy Depth resolution	Low 10 cm or (cm sometimes)	High <= 10 um	Middle 1 mm
Low light	~~Weak~~ Good	Light source dep.	~~Good~~ (IR, laser) Ok
Outdoor	OK	Weak	Weak
Game Control /HCI (1~5m)			✓
3D Movie /ADAS (>5m)	✓ + LiDar		
3D Scanning /3D printing or 3D AOI (<1mm)		✓	

2.0.2 Hardware Specification – RGB-D

1. Camera: Intel RealSense D415

	W	H
(1) RGB Resolution (pixels)	1920	1080
(2) RGB Angle of View	69.4˚	42.5˚
(3) RGB Frame Rate (fps)		30
(4) RGB Focal Length		1.88 mm
(5) Depth Resolution (pixels)	1280	720
(6) Depth Angle of View	63.4˚	40.4˚
(7) Depth Frame Rate (fps)		Up to 90
(8) Depth Distance Range (m)		0.3~10m
(9) Depth Accuracy (mm)		≤ 2% < 20mm @ 1m

Need Fixed

Robotiq 2F-85 gripper

2) Robot Arm: Yaskawa MH5LF

(1) Controlled axes	6
(2) Payload (kg)	5
(3) Repeatability (mm)	±0.03
(4) Weight (kg)	29
(5) Opened Gripper Width(mm)	85

Yaskawa MH5LF

Controller

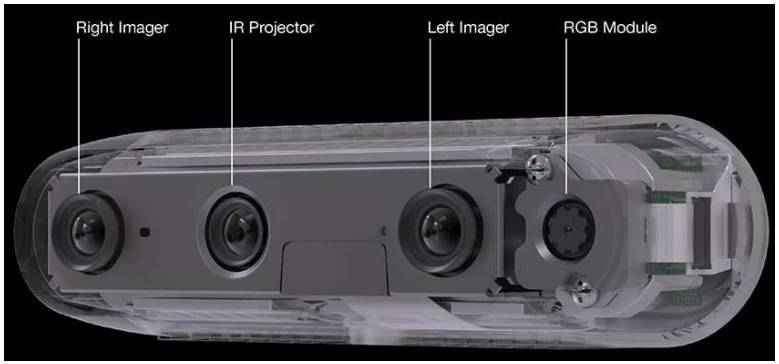
2.0.2 Hardware Specification – RGB-D Camera

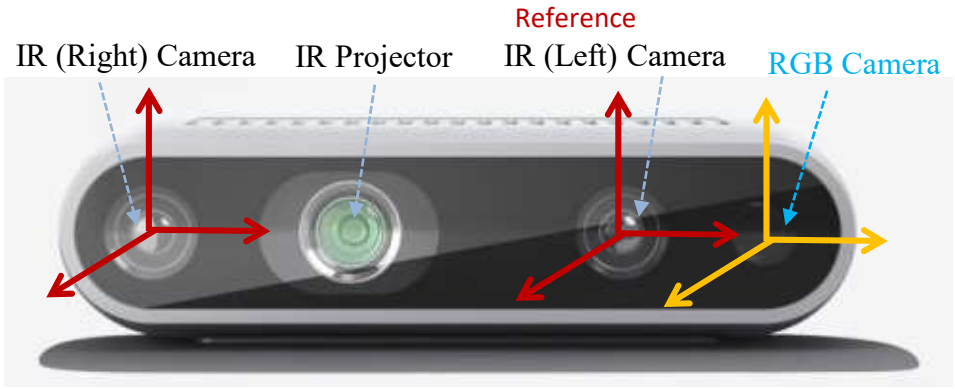
2. RGB-D Camera: Intel RealSense Depth Camera-D435

RGB Sensor	W	H
Resolution	1920	1080p
pixel size	3 um	3 um
FOV	69.4° x 42.5° x 77°	
Frame Rate	30 fps	

Depth Sensor	W	H
Resolution (pixel*pixel)	1280	720p
FOV	85.2° x 58° x 94° (+/- 3)	
Frame Rate	30 fps	
Z-axis Accuracy	1 mm (increased by depth)	
Depth Technology	Active IR stereo	
Minimum Distance	0.11m	
Maximum Distance	Approx. 10 meters	

Source : https://click.intel.com/intelr-realsensetm-depth-camera-d435.html



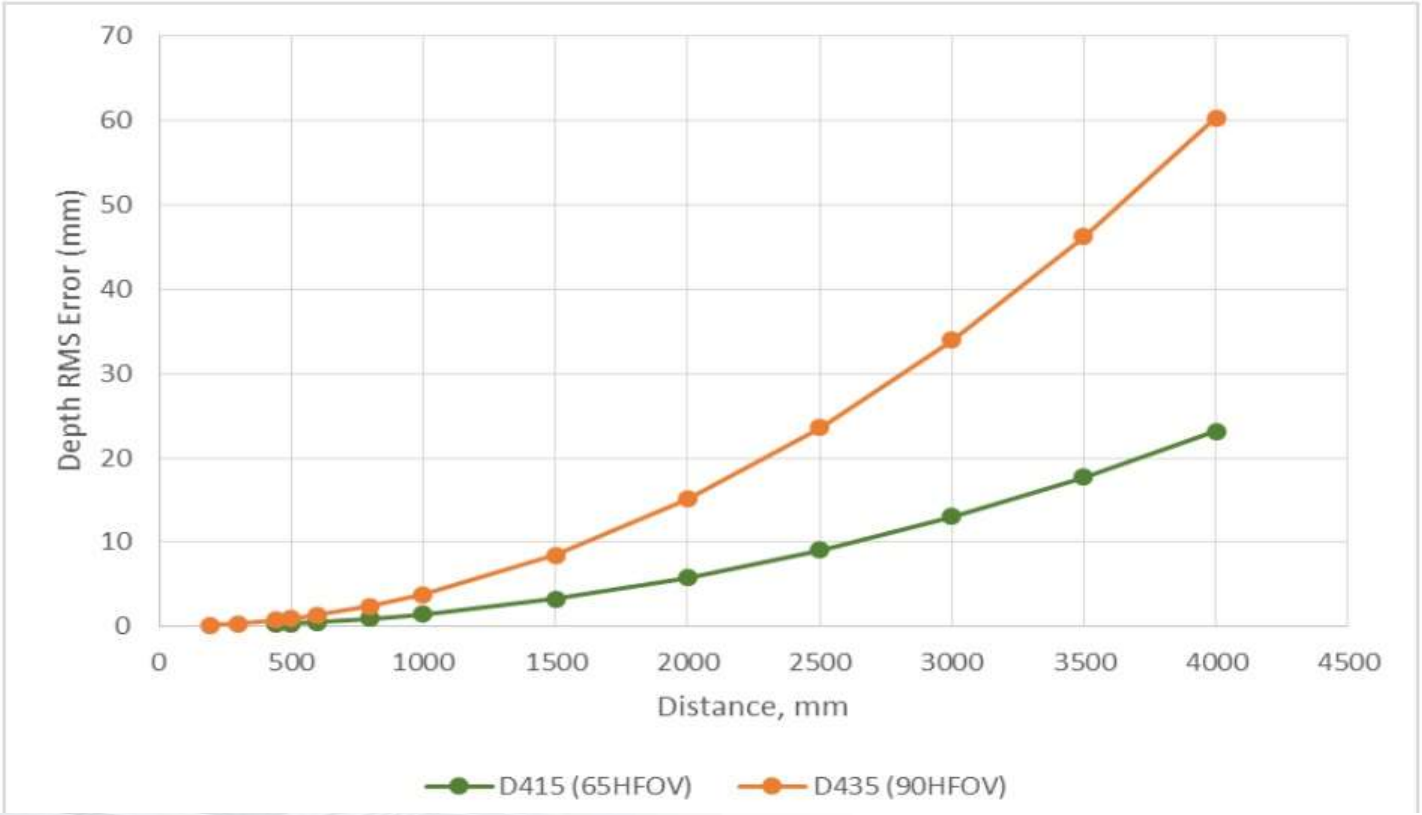


2.0.2 Hardware Specification – RGB-D Camera

2. RGB-D Camera – Intel RealSense D435

UNDERSTAND THEORETICAL LIMIT

The graph is obtained using subpixel=0.08:
D415 with HFOV=65deg, Xres=1280, and baseline=55mm,
D435 with HFOV=90deg, Xres=848, and baseline=50mm

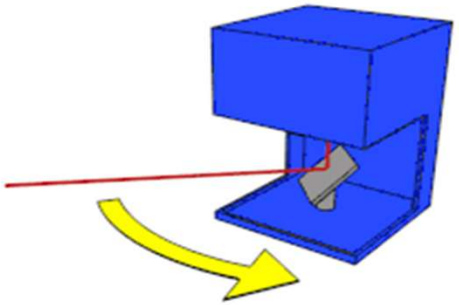


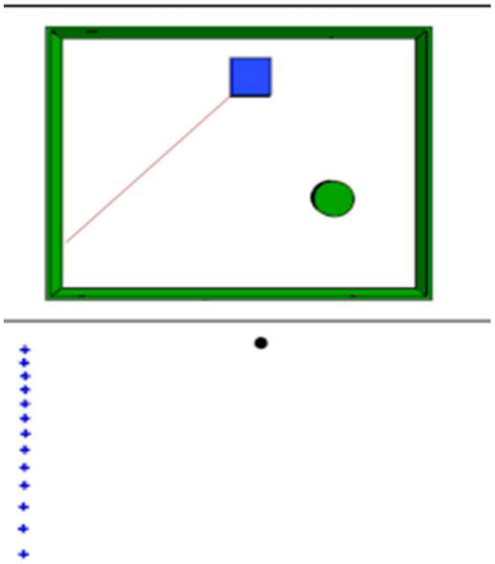
Reference: https://realsense.intel.com/wp-content/uploads/sites/63/BKM-For-Tuning-D435-and-D415-Cameras-Webinar_Rev3.pdf

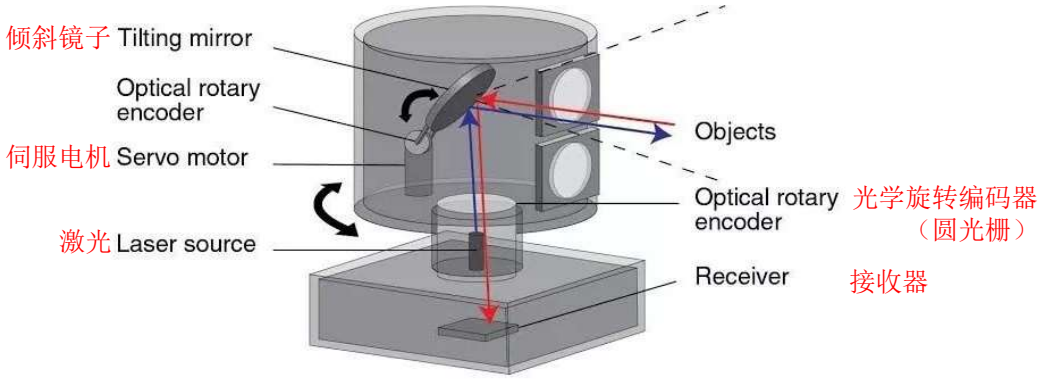
2.0.2 Hardware Specification L515

3. Realsense L515

	W	H
(1) RGB Resolution (pixels)	1920	1080
(2) RGB Angle of View	69˚ +/-1˚	42˚+/-1˚
(3) RGB Frame Rate (fps)		30
(4) RGB Focal Length		1.88 mm
(5) Depth Resolution (pixels)	1024	768
(6) Depth Angle of View	70˚	55˚
(7) Depth Frame Rate (fps)		Up to 30
(8) Depth Distance Range (m)		0.25 - 6.5m (640 x 480分辨率下0.25 - 9m)
(9) Depth Accuracy (mm)		< 5mm @ 1m < 14mm @ 9m







激光雷达LiDAR，全称为Light Detection and Ranging激光探测和测距，又称光学雷达。它通过在目标上发射光束来测量物体的距离，并使用反射光束的时间和波长来估计距离，是基于光束（激光）的测距方法。

realsense l515 有四个微镜（MEMS，意法半导体PM56A）

（动图，幻灯片放映即可播放）

2.0.3 Hardware Specification – 2D Lidar

1. 2D Lidar: EAI-Flash Lidar F4

~~Price: 19,000 NTD~~

Flash Lidar F4	
Size	Φ70 x 60 mm
Sample Rate	4000 Hz
Scan Rate	6 ~ 12 Hz
Degree Precision	0.72 degree
Maximum Distance	8~10 m
Distance accuracy	0.5 mm (< 2m)
	Distance * 1% (> 2m)

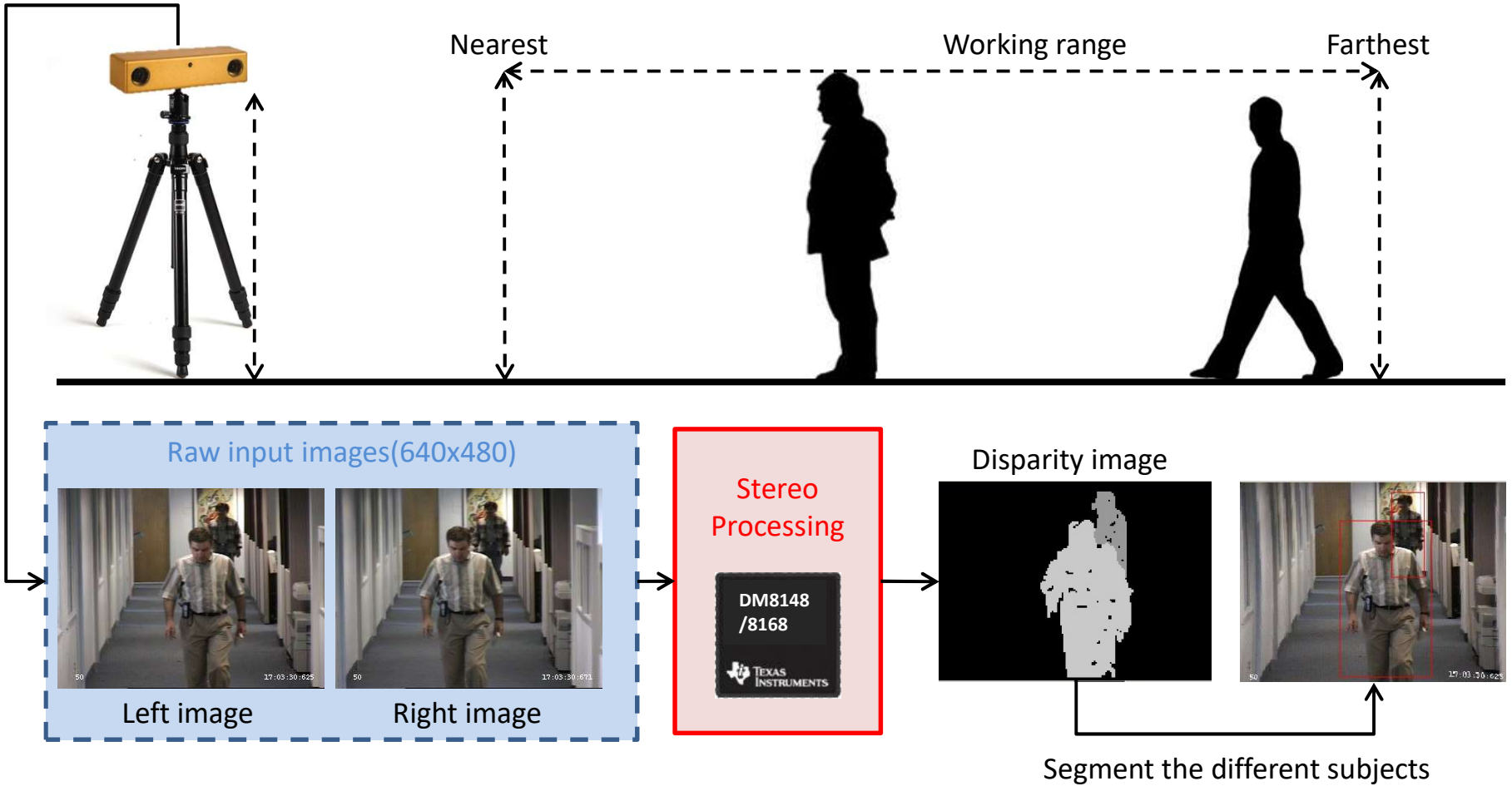
2. 2D Lidar: SLAMTEC-RPLiDAR A3

Price: 24,675 NTD

RPLiDAR A3	
Size	Φ76 x 41 mm
Sample Rate	16000 Hz
Scan Rate	10~20 Hz
Degree Accuracy	0.225 degree
Maximum Distance	25 m
Distance accuracy	0.5 mm (< 2m)
	Distance * 1% (> 2m)

2.1 Stereo System – Hardware:
Stereo System Diagram

- System Diagram:



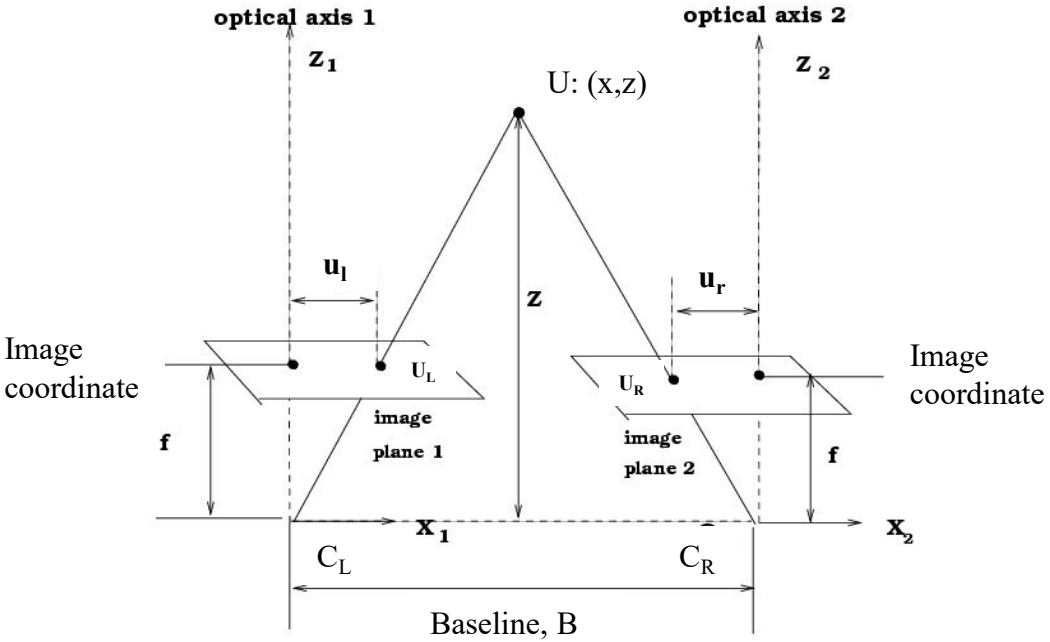
2.1 Stereo System – Hardware: Embedded Hardware

- Processing Unit of Embedded Based Stereo System (預計):

Chip	DM8148	DM8168
CPU	Cortex A8 1GHz DSP C674x 750MHz	Cortex A8 1.2GHz DSP C674x 1GHz
RAM	DDR2 or DDR3	DDR2 or DDR3
Feature	Vision Coprocessor (VCoP) 500MHz Accelerator HD VICP 2.0 320 MHz *2 Video Processing Subsystem (VPSS) 3D Graphics Engine Fixed/Floating Point	Vision Coprocessor (VCoP) 500MHz Accelerator HD VICP 2.0 320 MHz *3 Video Processing Subsystem (VPSS) 3D Graphics Engine Fixed/Floating Point
市場定位	HD Video Conferencing Video Surveillance DVR IPCam	HD Video Conferencing Video Surveillance DVR IPCam

2.1 Stereo System – Software: Model

- System Model:



U: point (X,Y,Z) in real world.
C_L, C_R: left/right cameras. (Left is the reference)
f: focal length of both cameras.
B: baseline, distance between 2 cameras
U_L, U_R: 點 U 在左右 2 cameras 的成像點
Disparity: $\Delta d = u_l - u_R$.
Base on 三角測量，depth, Z(m):

$$\frac{\Delta d\ (pixel)}{B\ (m)} = \frac{f\ (pixel)}{Z\ (m)} \Rightarrow Z(m) = \frac{f\ (pixel) \times B\ (m)}{\Delta d\ (pixel)}$$

2.1 Demo Video: Stereo

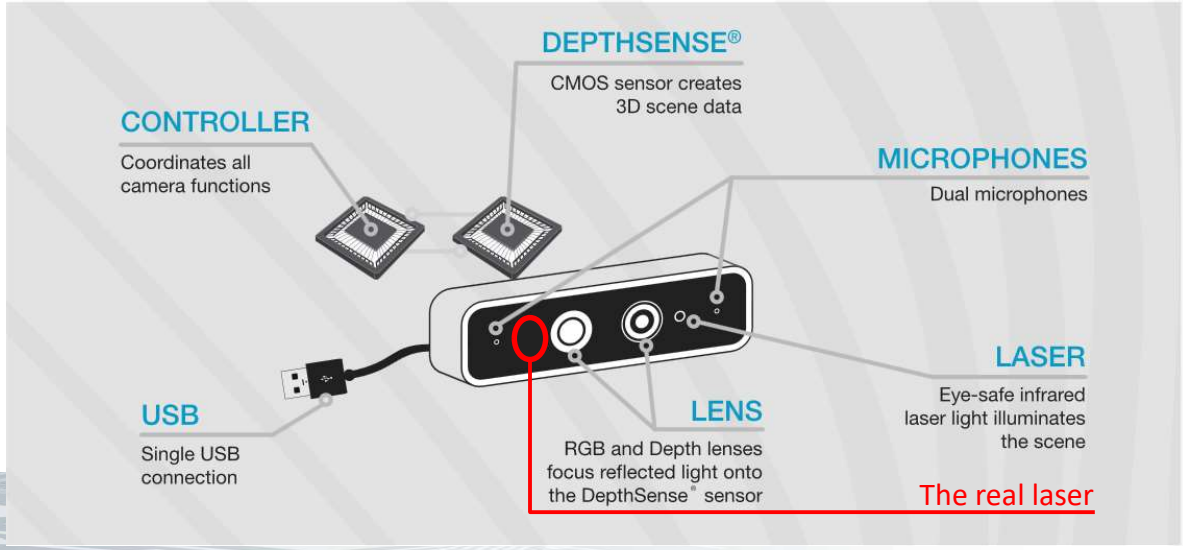
1) 3d_Stereo_Visionics.mpg
2) 3d_Stereo_Skeleton_Gesture_Etron_20130422.flv
3) 3d_Stereo_MagicBody_SClass_Benz_2013.flv

2.2 ToF: Introduction

- 3D Time-of-Flight
  - Hardware
    - Active modulated light source: Infrared rays
    - CMOS pixel array
  - Applications
    - 3D scanning/printing
    - Automotive driving/parking
    - Hand gesture
    - Etc.

Gesture Recognition

Automotive

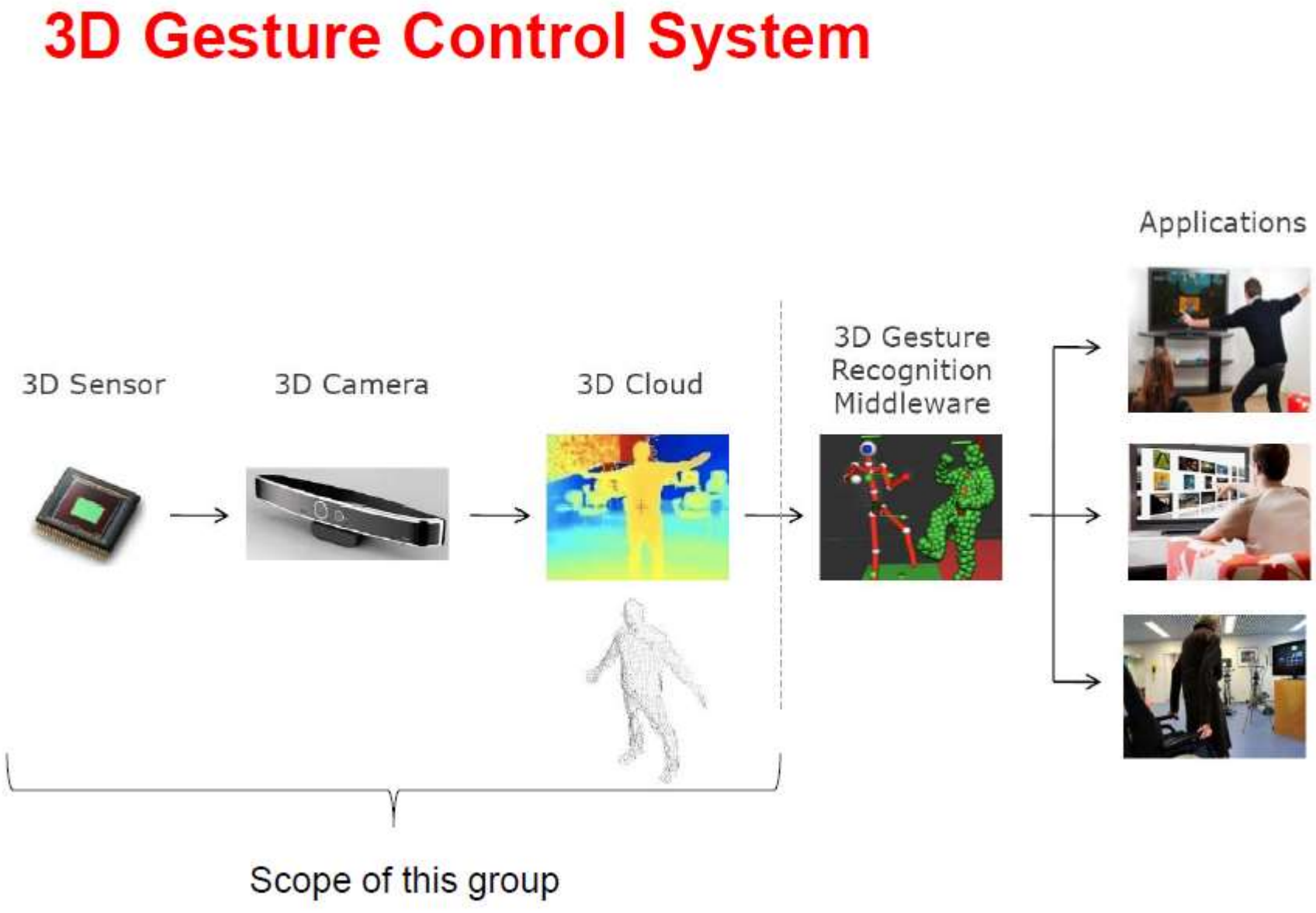


Technical Specifications

- **Technology**: DepthSense®
- **Depth field of view**: 74° x 58° x 87° (H x V x D)
- **Depth resolution:** 320 x 240 | QVGA
- **Frame rate**: 25 fps – 30 fps | QVGA
  50 fps – 60 fps | QVGA
- **Nominal operating range**: 0.15m – 1.0m
- **Depth noise**: < 1.4cm at 1m (50% reflectivity)
- **Illumination type**: Diffused laser
- **Ambient light**: Typical indoor
- **RGB resolution**: HD 720p
- **RGB field of view**: 63.2° x 49.3° x 75.2° (H x V x D)
- **Accelerometer**: 3 axis
- **Microphones**: 2
- **Connectivity**: Single USB
- **Operating temperature**: 10°C to 40°C
- **Power**: < 2.5W
- **Size**: 10.5cm (W) x 3cm (H) x 2.3cm (D)

Specification of SoftKinetic DS325

2.2 ToF: Introduction - One Example jj



2.2 ToF: Theory of Operation (1/5) jj

- How to get distance?
  - Detect the phase shift between pulsed source and reflection.

1) Pulsed source
   - Continuous-wave: Square wave
   - Can be easily realized using digital circuits.

2) Reflection
   - Sensor receives two components
     (1) Reflected component: The reflected light of pulsed source
     (2) Ambient component: Any other light source, cause noise, need to reduce
     - High ambient component reduces SNR (increase noise)

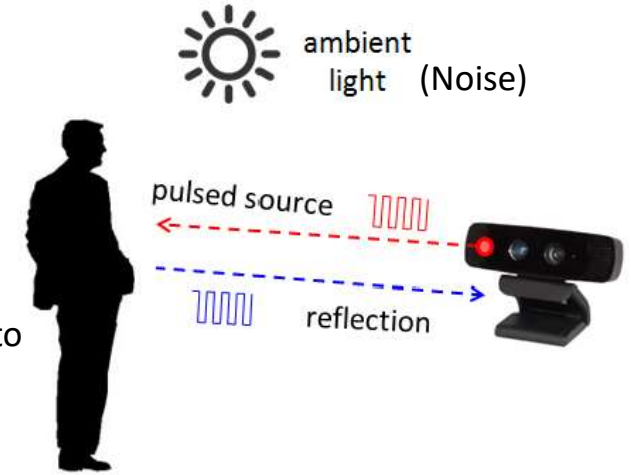
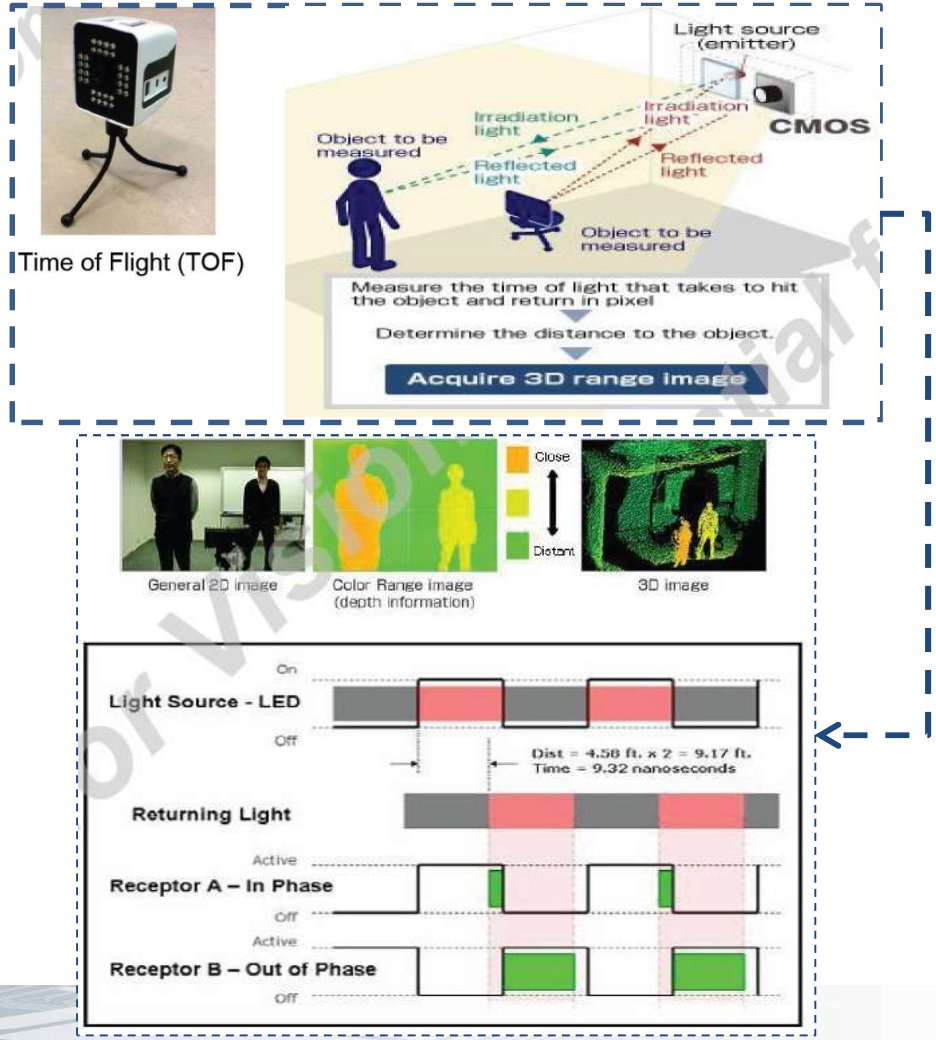


Figure 1: 3D time-of-flight camera operation.

- How to detect phase shift and phase angle?
  - Two methods:

1) Pulsed method: Phase shift
2) Continuous-wave method (used): Phase angle

2.2 3D Vision Technology – Time of Flight (TOF) (2/5)



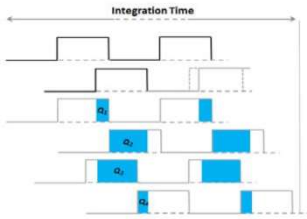
- Application (kinect2):
  1. Gesture control: Body or hand
  2. People counter:
     Partial occlusion in multiple people situation.

2.2 3D Vision Technology – Time of Flight (TOF) (3/5)

3D Tof camera
SoftKinetic DS325
Resolution:
Color - [640X480]
Depth - [320X240]

Phase map

1. Convert Phase φ to Depth Z



$Z = f(\varphi)$

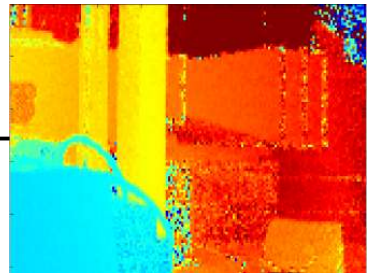
$$\varphi = \arctan\left(\frac{Q_3 - Q_4}{Q_1 - Q_2}\right)$$

Depth map Z [320X240]

Color Image [640X480]

Depth map Z

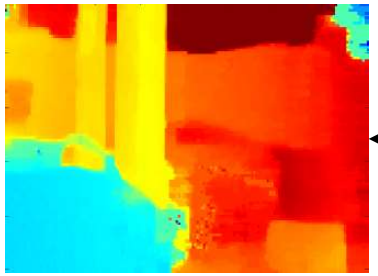
2. Depth Map Refinement (預計)



Depth map Z [320X240]

Color Image [640X480]

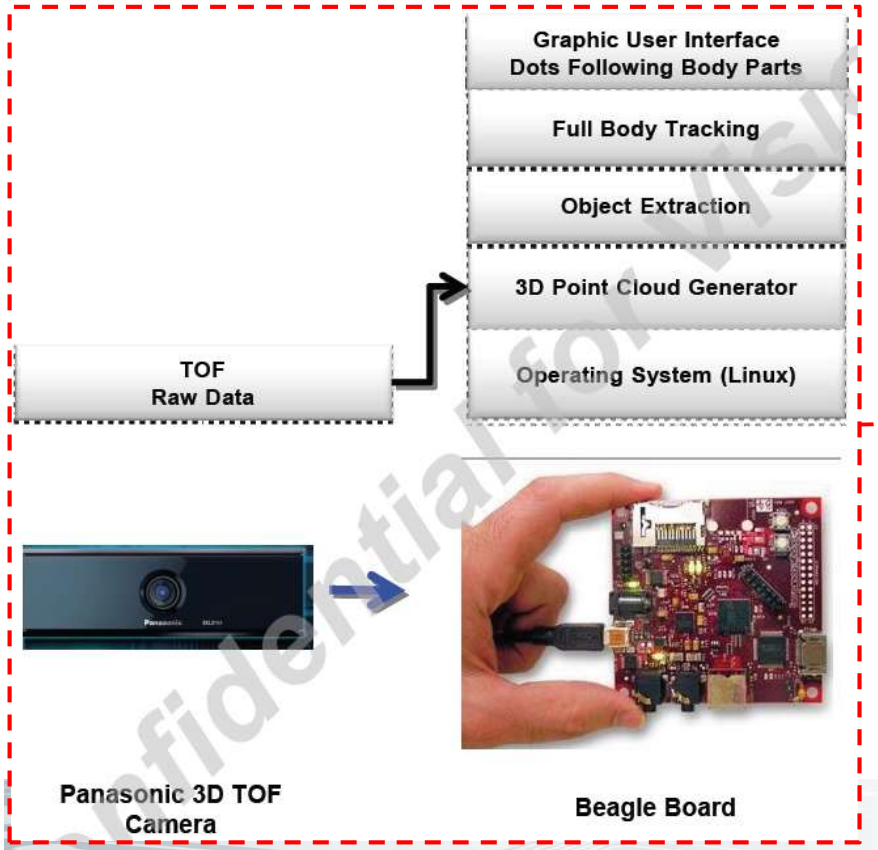
Denoise



Output: (預計)
Refined depth map Z' [320X240]
預計完成日期: 2014/07

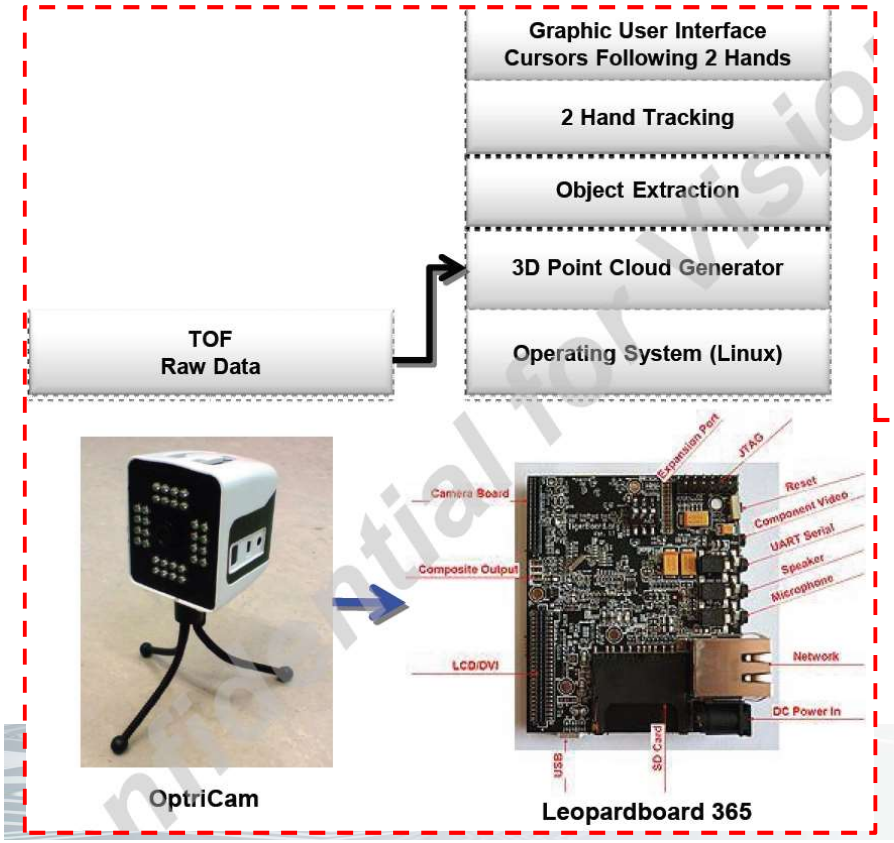
2.2 3D Vision Technology – Time of Flight (TOF) : Sensor Demo 01 (4/5)

- Full body Tracking (TI confidential info.):
  1. Detect body parts: head, shoulder, elbow, hand and torso.
  2. Kinect2 like user experience.



2.2 3D Vision Technology – Time of Flight (TOF) : Sensor Demo 02 (5/5)

- 2 hand tracking (TI confidential info.):
  1. Detect 2 hands motion.
  2. iPhone like user experience.



2.2 ToF: TI Solution

- 3-chip solution
  1) 3D TOF sensor array
     - Addressable CMOS pixel array
     - High pixel modulation frequency (>50MHz)
     - Up to 5x increase in SNR
     - Respond to specific optical spectrum (850-870nm)

  2) Analog front-end (AFE)
     - Supports up to 4 differential inputs
     - Sample-and-hold front-end that helps reject common-mode noise
     - High-speed, low-power 12-bit ADC samples

  3) TOF Controller (TFC)
     - Synchronizes the operation of TOF sensor, AFE and illumination
     - **Calculates the depth for each pixel**
     - Performs dealiasing, de-noising
     - Frequency tuning and temperature compensation

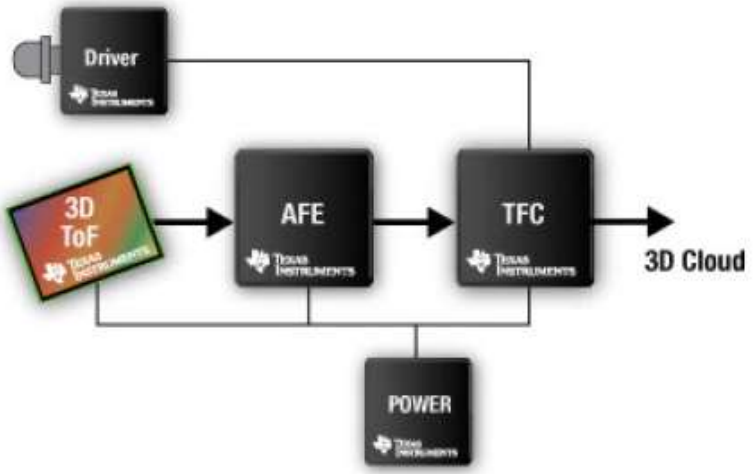
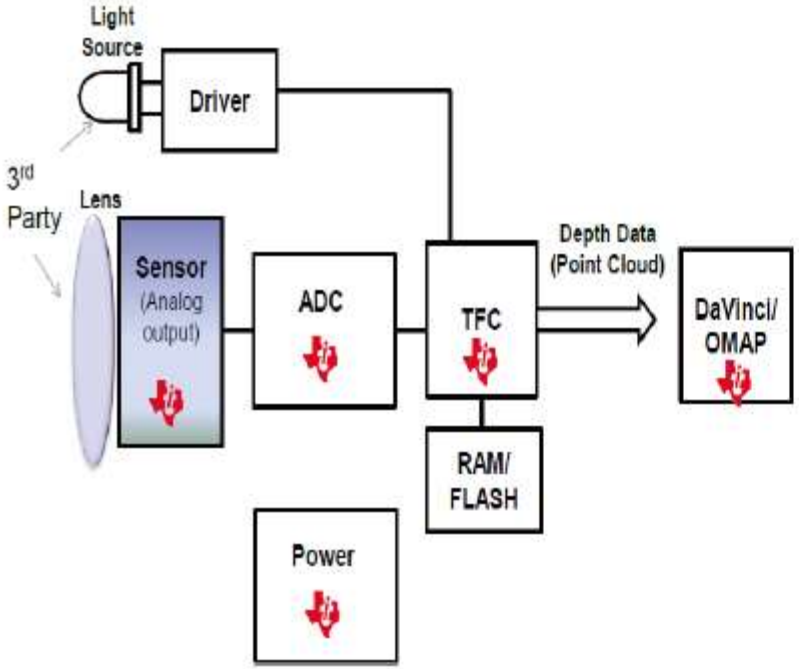


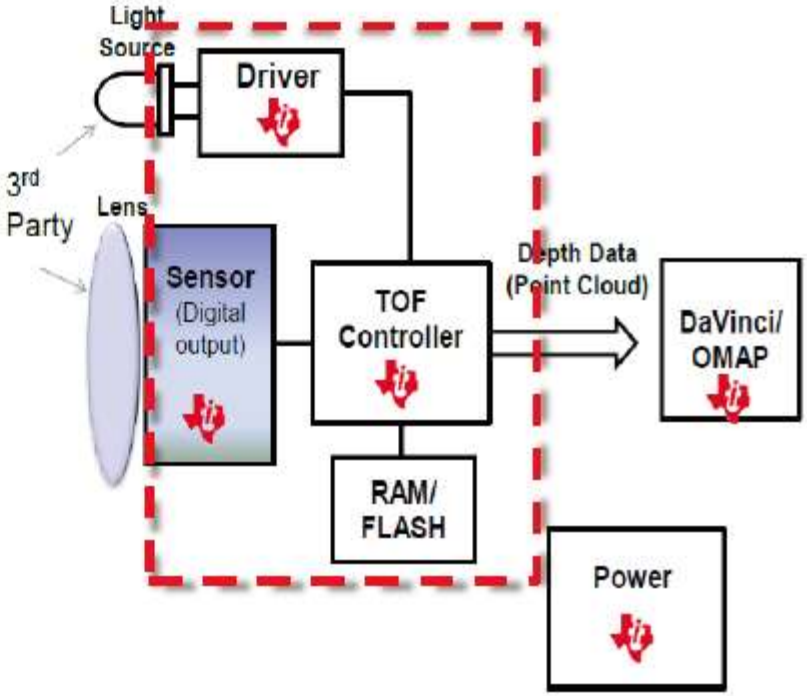
Figure 12: TI 3D-TOF chip set.

2.2 ToF: TI Solution (1/3)

TI Content in TOF System



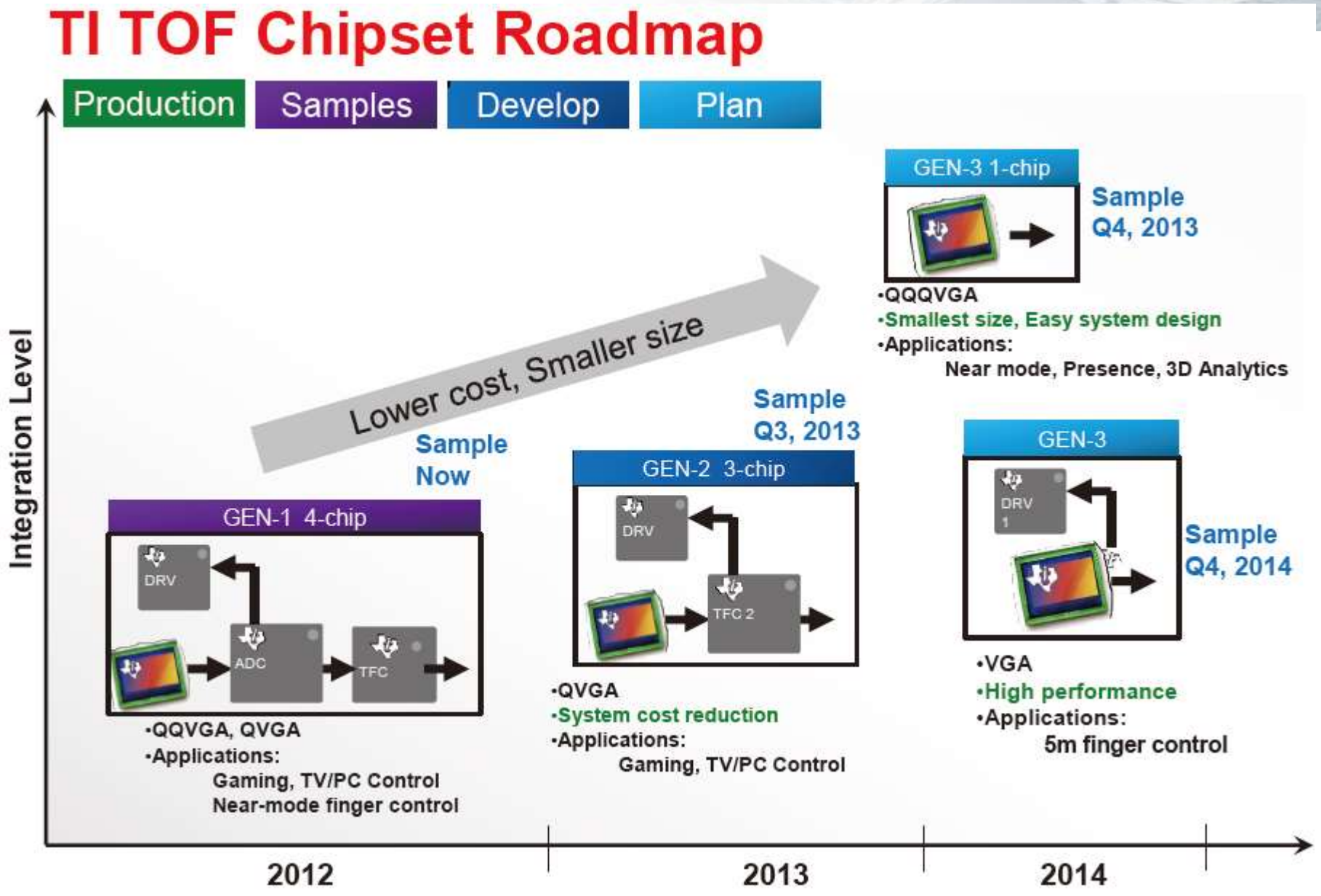
Today: Analog Out Sensor



2013: Digital Out Sensor

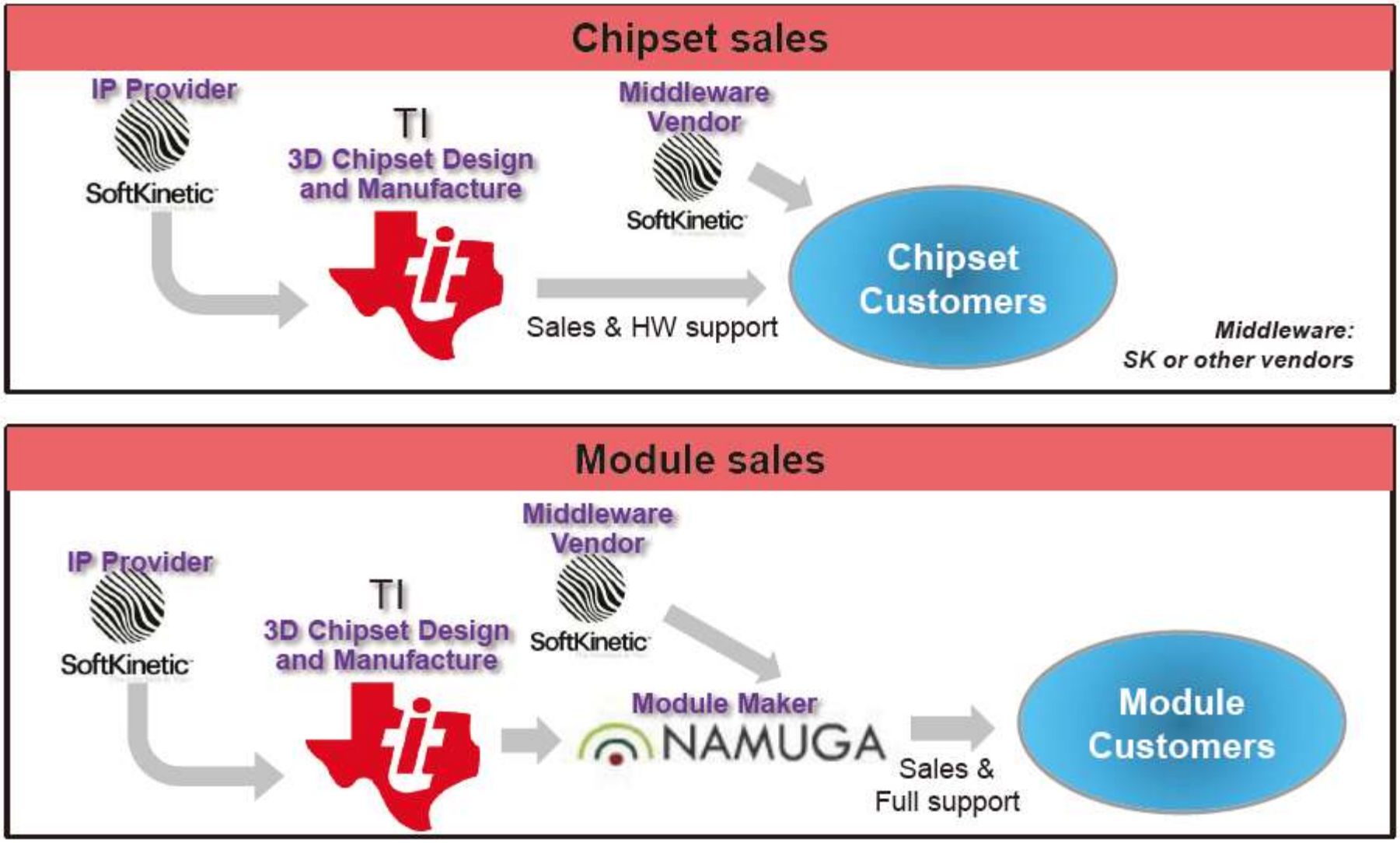
2014: Fully integrated QQQVGA Sensor

2.2 ToF: TI Solution (2/3)

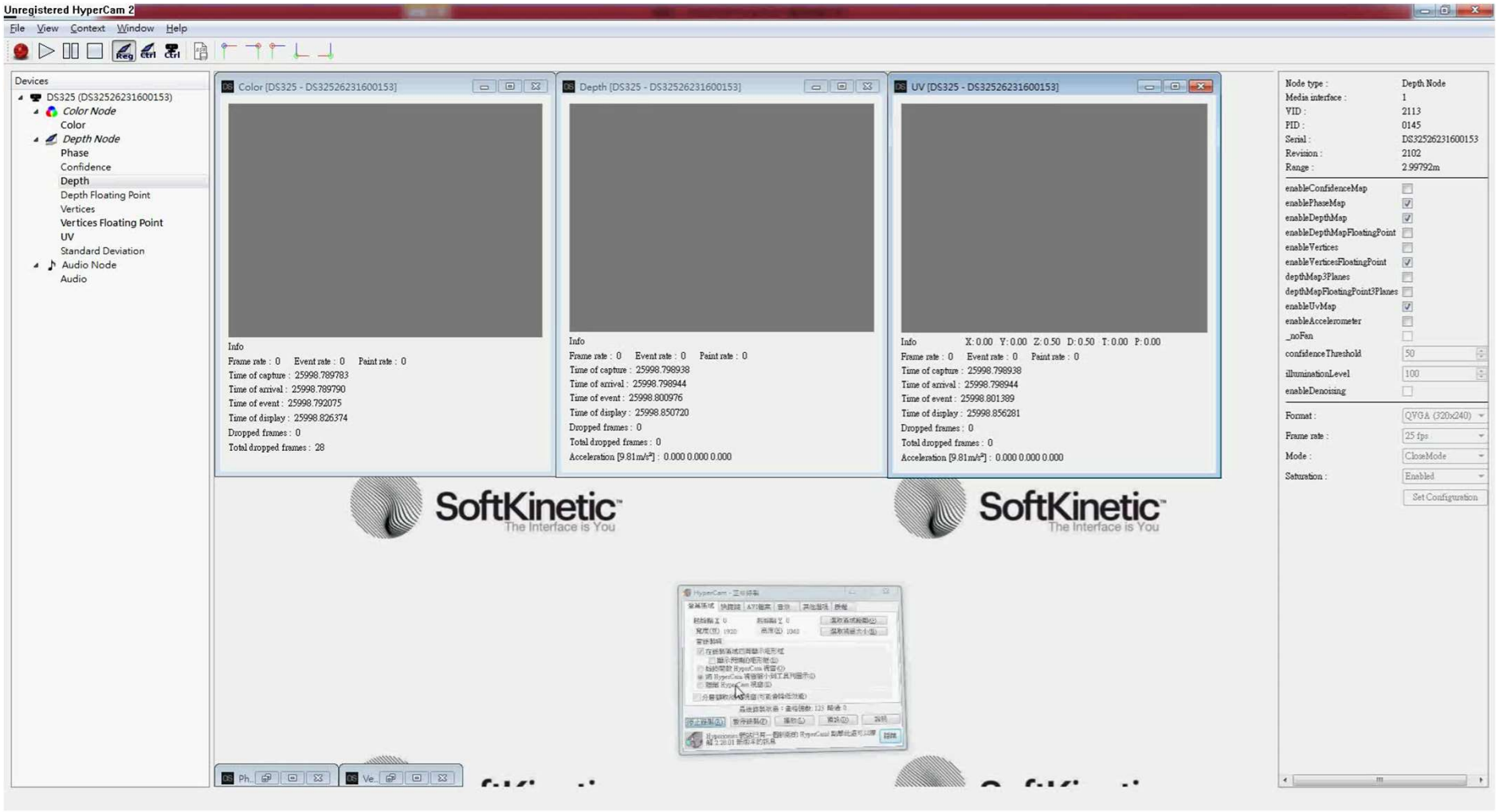


2.2 ToF: TI Solution (3/3)

TI 3D-TOF Solutions Delivery Model



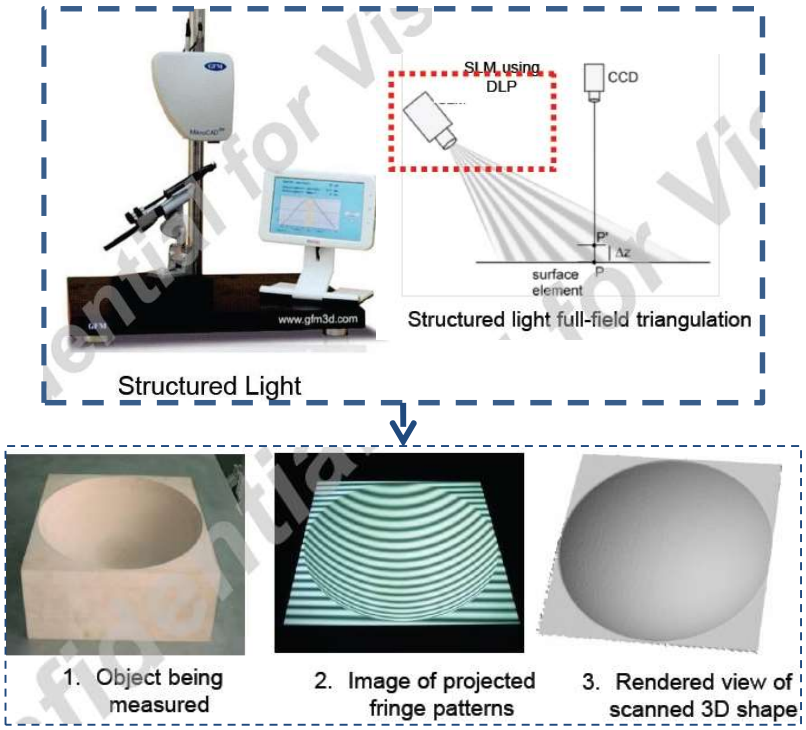
2.2 ToF: Demo Video 1



2.2 ToF: Demo Video 2

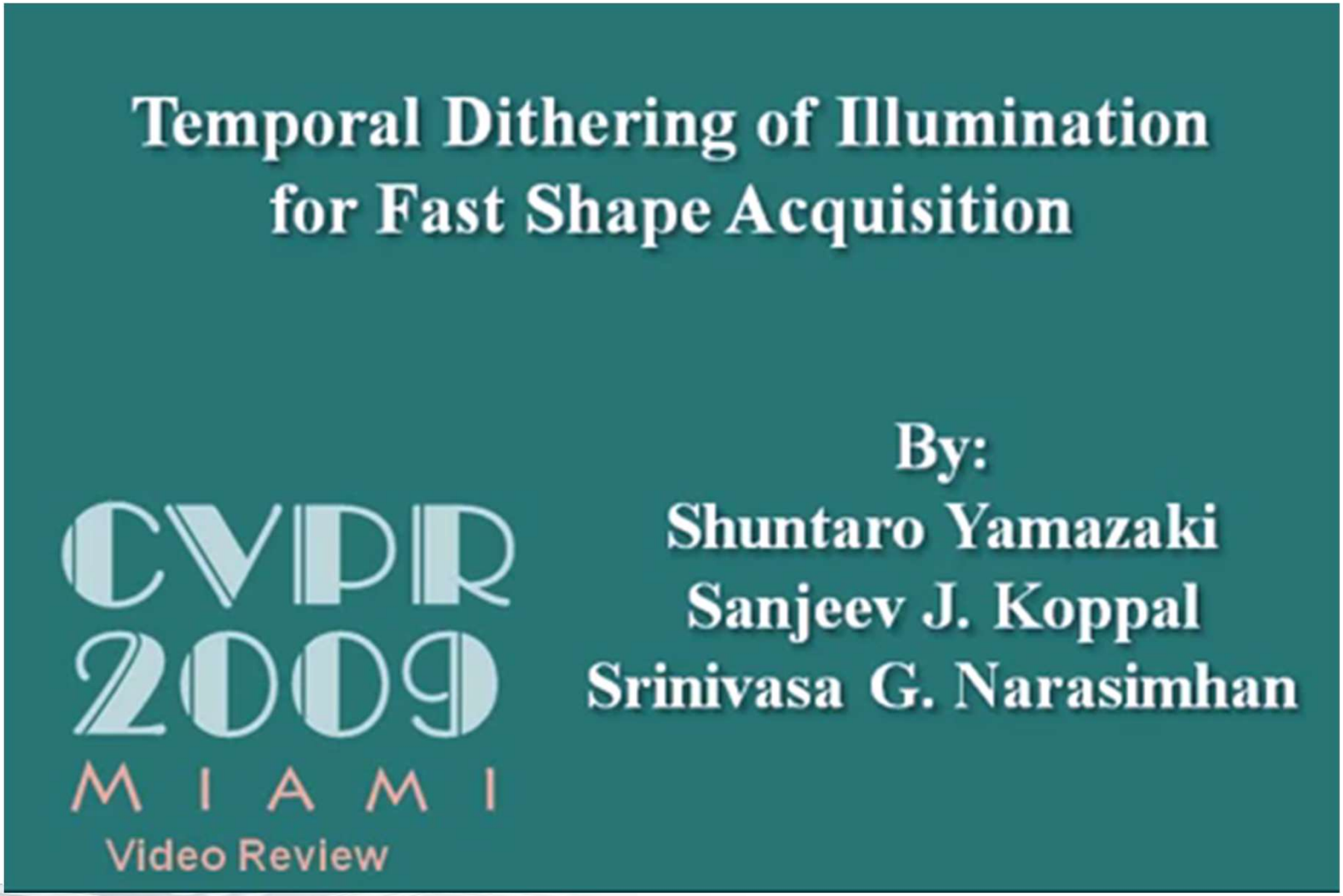
1) 3d_Tof_20140730_YuSh01.avi

2.3 3D Vision Technology – Structured Light



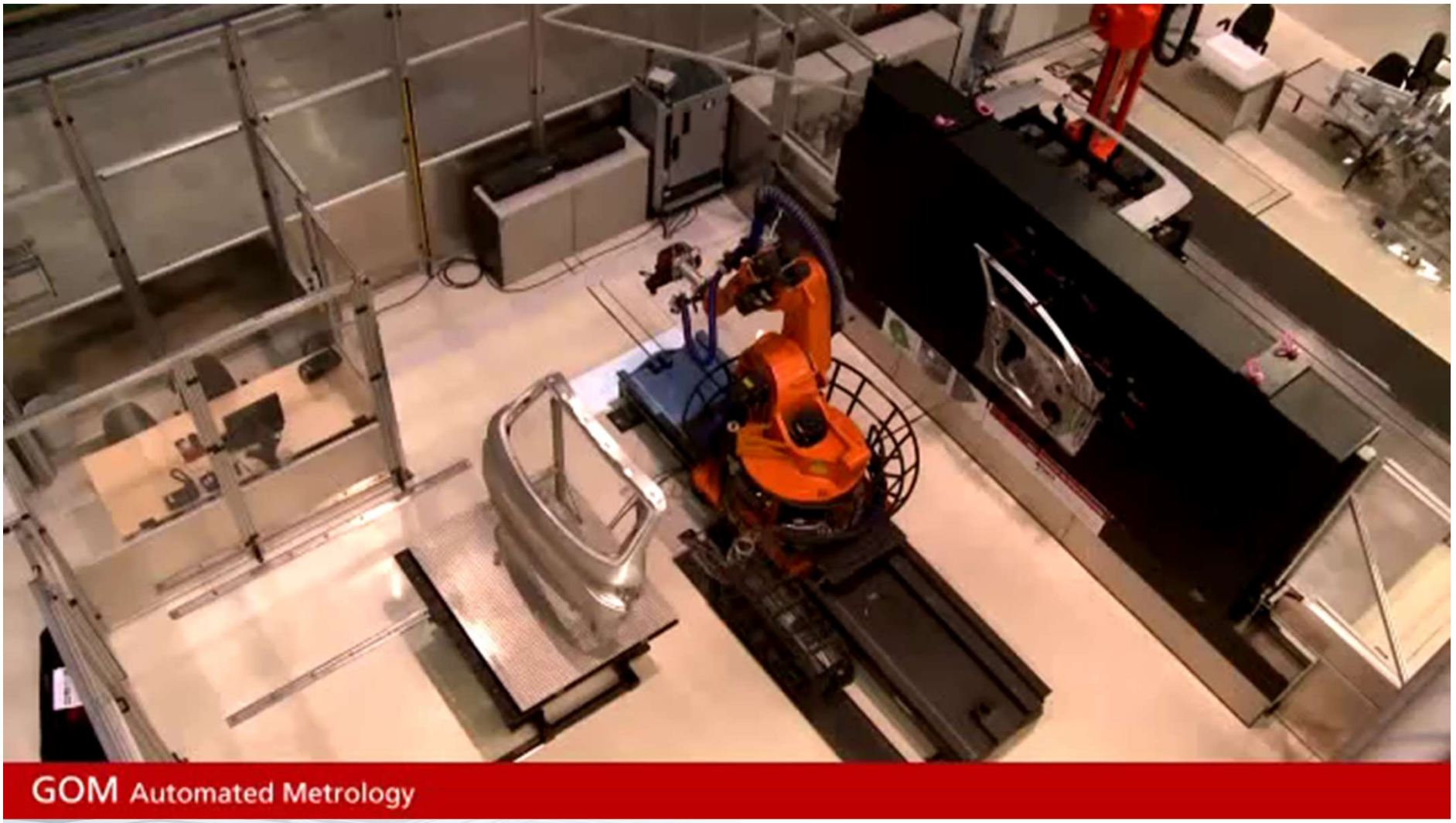
Kinect 1, DLP

2.3 Structured Light: Demo Video - Theory



2.3 Structured Light: Robot Arm 1

2.3 Structured Light: Robot Arm 2



2.3 Structured Light: Demo Video

1) 3d_StructureLight_Kinect_HowWork1_20140922.flv
2) 3d_StructureLight_Kinect_HowWork1_20140922.flv

2.3 Structured Light: Robot Arm + Pile + 3D Vision

3. Applications (1/3)

- Application filed
  - Automotive
    - Autonomous driving
    - Surrounding awareness
  - Industrial
    - HMI (Human-Machine Interface)
  - Healthcare
    - Gesture
  - Smart advertising
    - Gesture
    - Human recognition
  - Gaming
  - Entertainment

3. Applications (2/3)

Figure 10: TOF technology applies to a wide range of applications.

3. Applications (1/3)

- Gesture applications
  - Channel surfing
    - Can be done by waving of hands
  - Presentation
    - Scrolls using finger flickering

- Non-gesture applications
  - Automotive
    - Alerting the driver when it detects people and objects in the vicinity of the car
  - Robotics and automation
    - Detect product defects and enforce safety envelopes
  - 3D printing/ 3D scanning
    - Enable “3D copier” capability

Figure 11: Gesture recognition using a 3D-TOF sensor.

* Used with permission