



CEVA TECHNOLOGY  
SYMPOSIUM SERIES

# Sensing The World With Depth Cameras

## An introduction to 3D Sensing

Luke Liu, Founder and CEO, LIPS

[www.ceva-dsp.com](http://www.ceva-dsp.com)

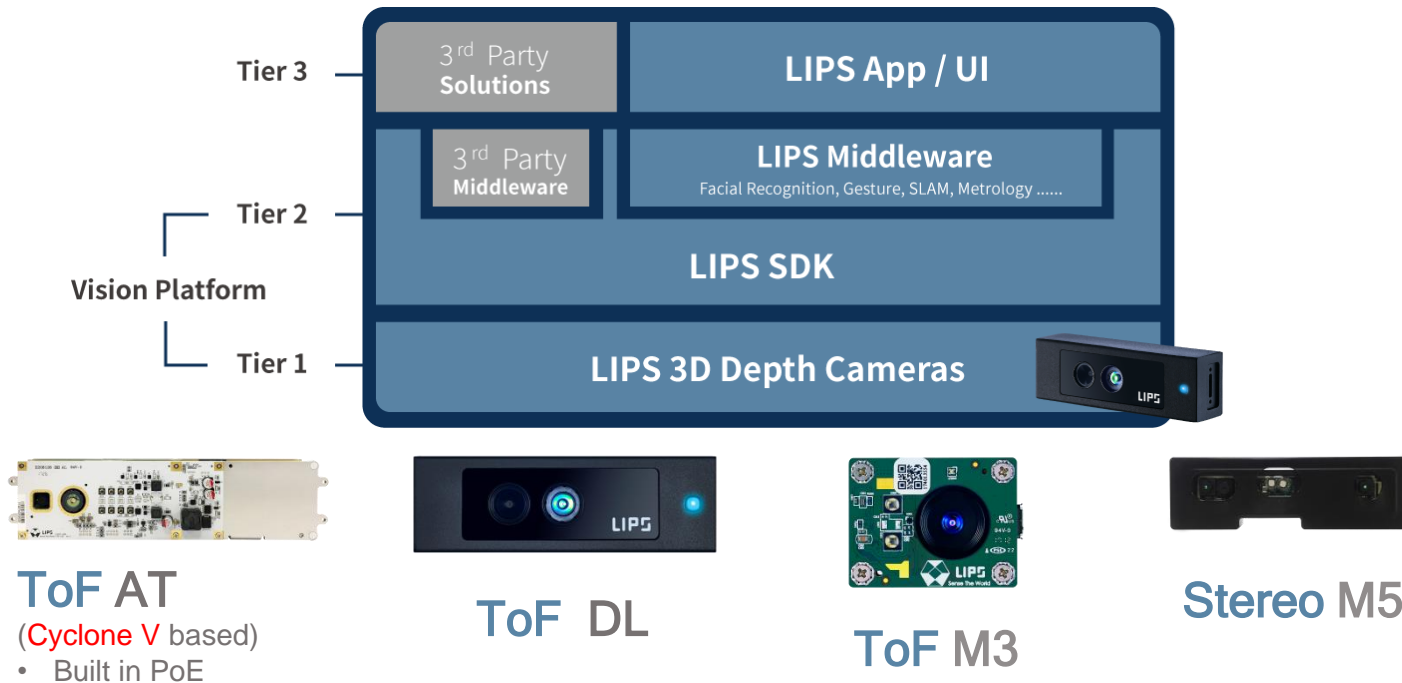


# Agenda

- ▶ Company Overview
- ▶ What's 3D/Depth Information
- ▶ The Benefit of Depth for CV and ML
- ▶ Case Study, Motion Gesture Recognition
- ▶ Some Tips for XM4 & 3D ML
- ▶ *Overall Market on 3D Sensing*

# Company Overview

LIPS Corp. is a 3D Sensing Solution Provider



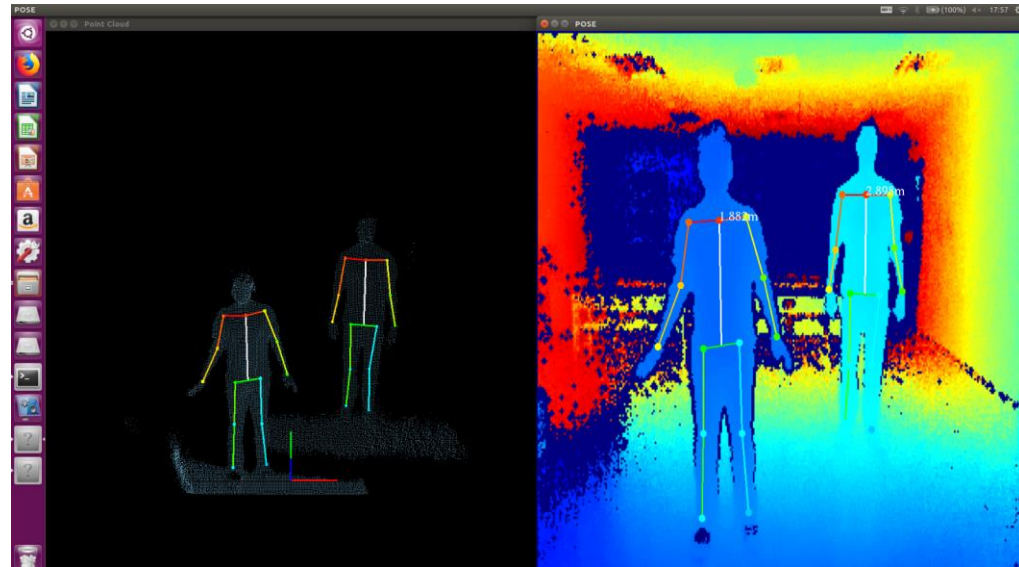
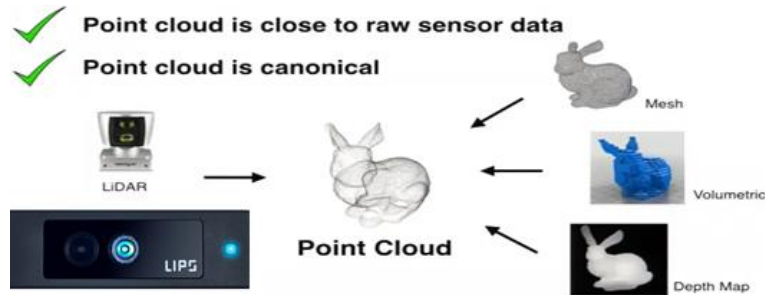
# Core Competency



- ▶ HW: State-of-art 3D depth camera
  - ▶ 3D depth cameras includes ToF, Stereo and Structured Light
  - ▶ Edge Computing for IoT
- ▶ MW, AL & SW
  - ▶ Parallel processing(DSP, OpenCL, CUDA and FPGA)
  - ▶ Machine vision recognition algorithms base on depth and RGB information for ML & traditional programming
- ▶ Customization

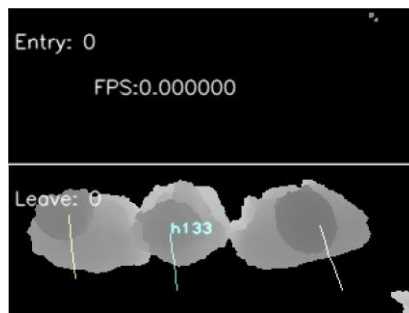
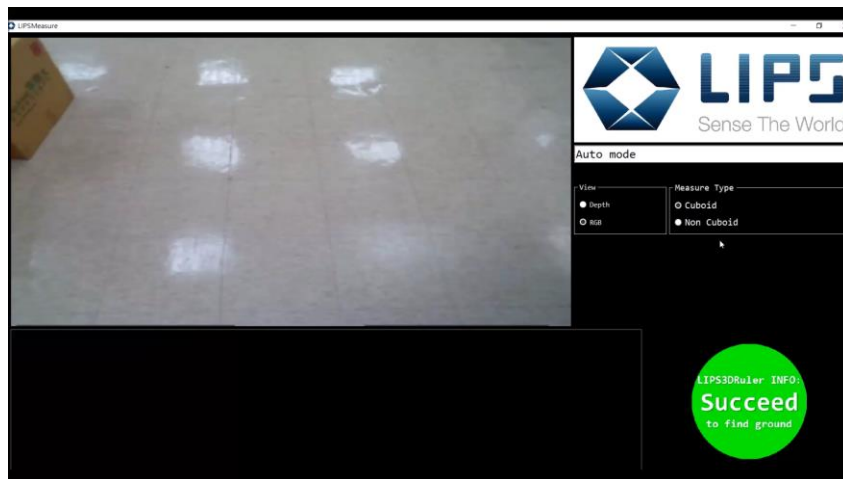
# What's 3D Information

## Point Cloud, Depth and RGB



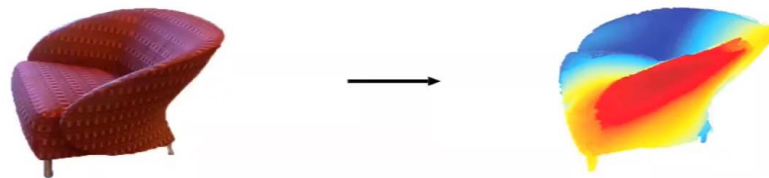
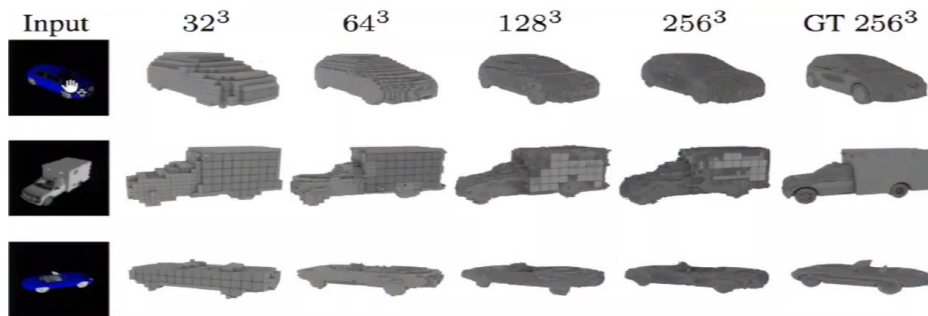
# The Benefit of Depth for Traditional program

- ▶ Easy to segmentation, compare with RGB information
- ▶ Get the real dimension of the object
- ▶ Get the relative distance and angle
- ▶ Not a projected image plane, but multi view object
- ▶ And the most important, not effective by light or others



# The Benefit of Depth for Machine Learning

- ▶ Easy to segmentation, and feature extraction
- ▶ Get the real dimension of the object, shape with different view
- ▶ Get the relative distance and angle
- ▶ Not a projected image plane, but multi view object
- ▶ And the most important, not effective by light or others



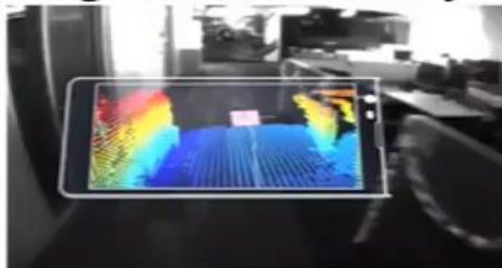
# Big Data + Deep Representation Learning

## Robot Perception



source: Scott J Grunewald

## Augmented Reality



source: Google Tango

## Shape Design



source: solidsolutions

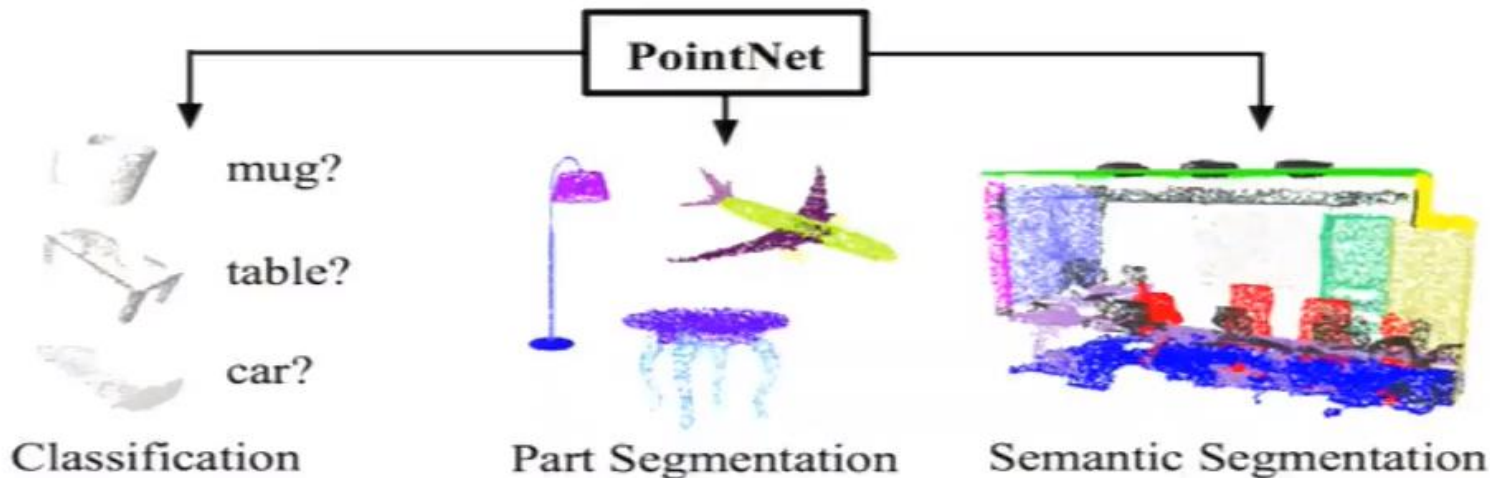
**Need for 3D Deep Learning!**



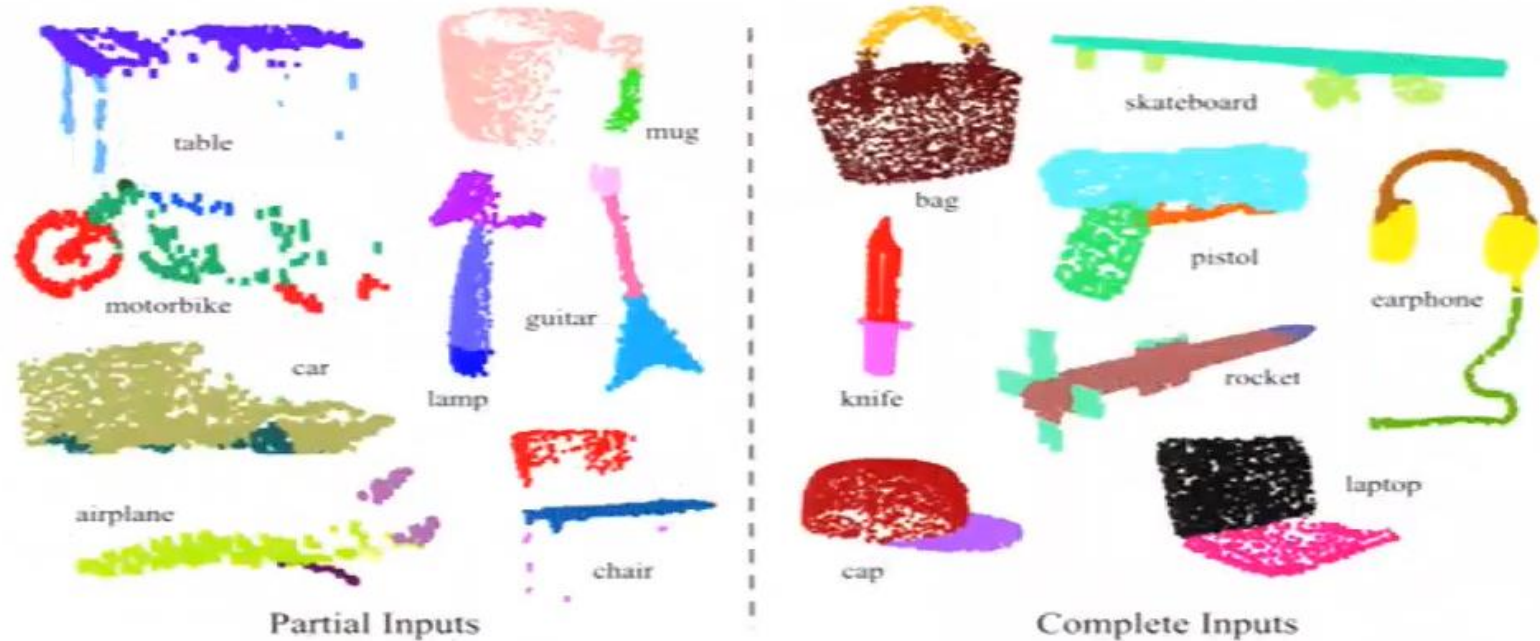
# Neuro Network

End-to-end learning for **scattered, unordered** point data

**Unified** framework for various tasks



# Results on Object Part Segmentation



# Motion Gesture Recognition Demo

The screenshot displays a motion gesture recognition demo. On the left, a terminal window titled "Display0 view" shows the following output:

```
FPS:28.255444  
none : 0.934063
```

Below the terminal, a log of system messages and gesture data is visible:

```
INFO: in LoadCalibrationData, function is  
startCameraIR  
INFO: Copy Protection HW detected  
WARNING: 8bytes mode is not support  
INFO: Process GT2(24f2:0214) init...  
Flash CT = 2, Reg CT = 1  
INFO: in setIRResolution  
INFO: in setDepthFPS  
INFO: in setIRResolution  
INFO: in setDepthFPS  
IR video mode - FPS=30, X=80, Y=60  
controlGesture = up0.798373  
controlGesture = right0.616597  
controlGesture = right0.825191  
controlGesture = right0.616933  
controlGesture = right0.775261  
controlGesture = right0.624958  
controlGesture = down0.51681  
controlGesture = right0.630413
```

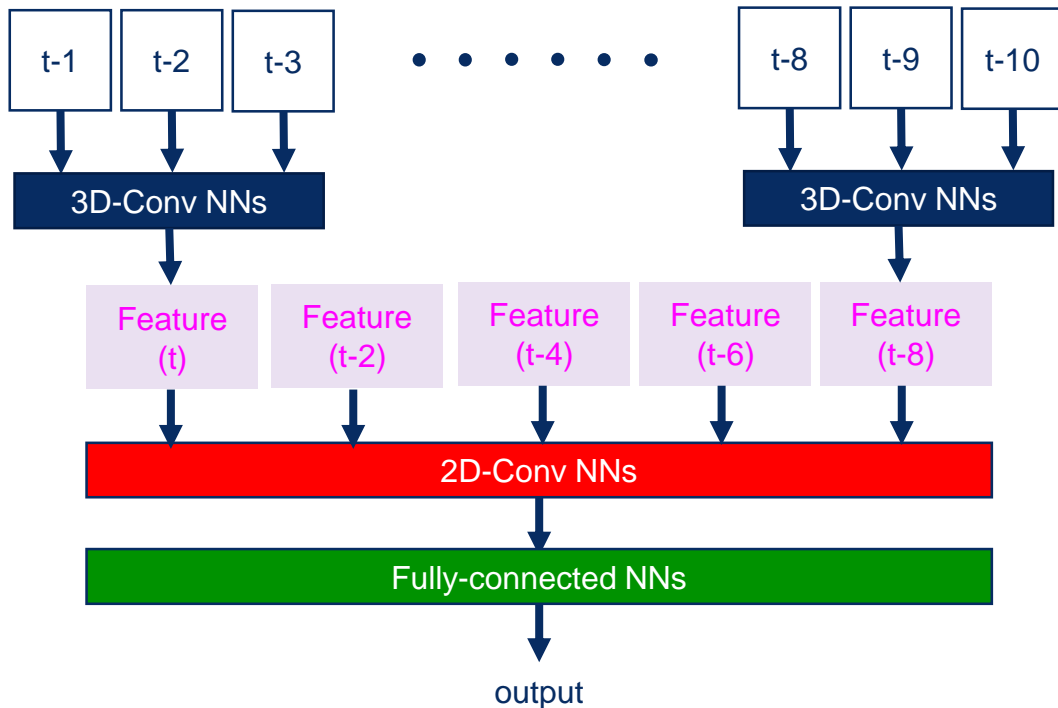
On the right, a browser window shows a YouTube page titled "Algorithm Weekly Report 2017". The page displays a grid of music videos, including:

- 周杰伦 Jay Chou 【聽見下雨的聲音 Rhythm of the Rain】 Official (4:44)
- 周杰伦 Jay Chou (with 楊瑞代) 【等你下課 Waiting For You】 歌 (4:35)
- G.E.M. 【光年之外 LIGHT YEARS AWAY】 MV (電影《太空潛航者》) (3:56)

The browser's address bar shows the URL: <https://www.youtube.com/tv#/surface?c=FEtopics&para...>

# Motion Gesture Recognition base on 2D

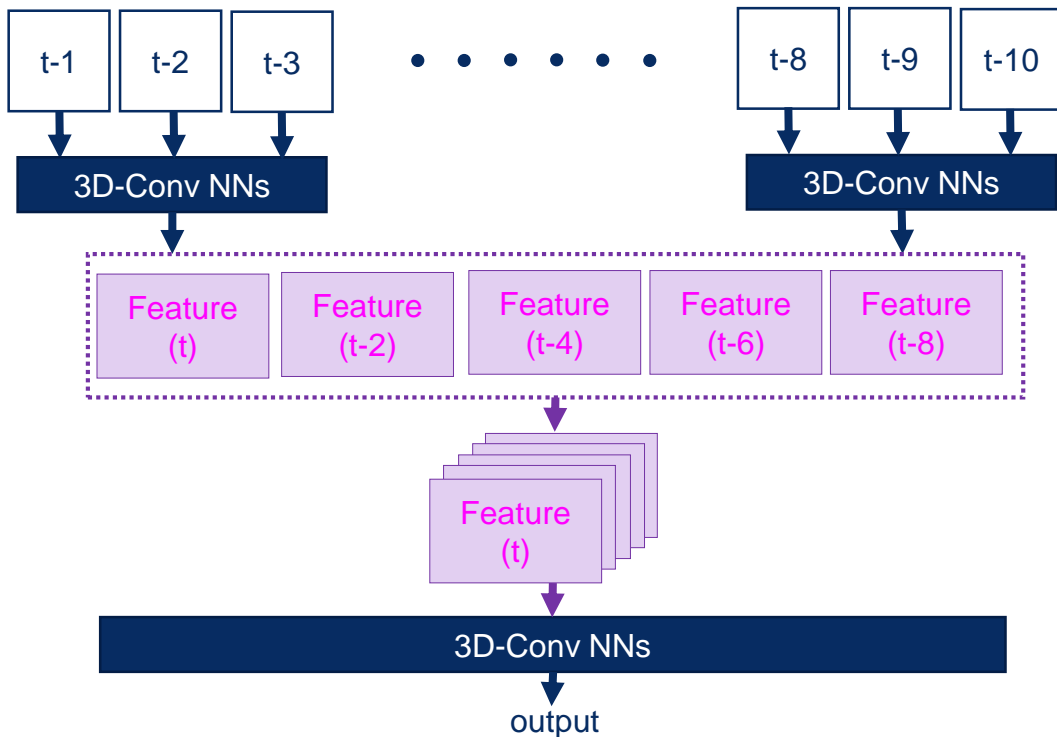
## LocalNet(3D-Conv) + GlobalNet(2D-Conv)



```
Accuracy: 0.906
AveragePrecision: 0.907681888775
AverageRecall: 0.911635380279
Precision:-----
up      :0.911
down    :0.897
+       :0.919
-       :0.925
right   :0.925
left    :0.944
click   :0.863
open    :0.878
none    :0.906
Recall:-----
up      :0.891
down    :0.972
+       :0.971
-       :0.925
right   :0.925
left    :0.850
click   :0.917
open    :0.857
none    :0.896
```

# Motion Gesture Recognition base on 3D

LocalNet(3D-Conv) + GlobalNet(3D-Conv)



```
Accuracy: 0.98
AveragePrecision: 0.985719261753
AverageRecall: 0.97697826508
Precision:-----
up      :1.000
down    :1.000
+       :0.971
-       :1.000
right   :1.000
left    :1.000
click   :0.958
open    :0.976
none    :0.966
Recall:-----
up      :1.000
down    :0.972
+       :0.971
-       :0.950
right   :1.000
left    :1.000
click   :0.958
open    :0.952
none    :0.988
```

# Some Tips for CEVA-XM4 & 3D ML



- ▶ Suggest float type input
  - ▶ CEVA-XM4 enabled
- ▶ Suggest float accuracy inference at least FP16, better FP32
  - ▶ CEVA-XM4 enabled
- ▶ Besides ML, what you need?
  - ▶ Good performance on DSP for pre-processing or post-processing
- ▶ If you only need traditional programming, instead of ML
  - ▶ Built-in OpenCV functions shorten development time

# Overall Market on 3D Sensing

Grow at CAGR 48% over the next decade to reach approximately \$64.5 billion by 2025



## Technology

Stereo

Structured Light

Time of Flight

Active Stereo

active



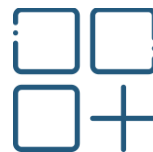
## Host

Standalone Professional

Laptop/Computer

Smartphone

Tablet



## Verticals

Automotive

Robotics

3D Metrology

Retail

Security

Healthcare

3D Scanning/Printing

Gaming

AR/VR

Smart Home

# Thank You

Luke Liu

lukeliu@lips-hci.com



[www.ceva-dsp.com](http://www.ceva-dsp.com)

[www.lips-hci.com](http://www.lips-hci.com)