NCKU MiinWu School of Computing
敏求智慧運算學院

AI Robotics

How to Learn Deep Learning? (1/2)
(如何學習人工智慧-深度學習?)

National Cheng Kung University
Dept. of Computer Science and Information Engineering
Robotics Lab.
國立成功大學 資訊工程系 機器人實驗室

(Jenn-Jier James Lien / 連震杰)
http://robotics.csie.ncku.edu.tw
jjlien@csie.ncku.edu.tw

國立成功大學

Department of Computer Science and Information Engineering
Robotics Lab

I. Course Syllabus

Computer Vision and Deep Learning: FTP Site and Website

- **FTP site to download and upload:**
  - IP: 140.116.154.1
  - Port: 21
  - ID: cvdl2020
  - Password: cvdl2020

- **Download lectures: After 17:00 every Wednesday**

- **Website:**
  - http://robotics.csie.ncku.edu.tw/course.html

- **Office Hour: CSIE 9F Robotics Lab**
  - Monday – 19:00~21:00
  - Wednesday – 09:00~11:00

OpenCV and Deep Learning: FTP Site and Website

- **FTP site to download and upload:**
  - IP: 140.116.154.1
  - Port: 21
  - ID: opencvdl2020
  - Password: opencvdl2020

- **Download lectures: After 17:00 every Thursday**

- **Website:**
  - http://robotics.csie.ncku.edu.tw/course.html

- **Office Hour: CSIE 9F Robotics Lab**
  - Monday – 19:00~21:00
  - Wednesday – 09:00~11:00

R.1 Profile in Academia – Education: ~ 1998

RI: Robotics Institute
SCS: School of Computer Science
CMU: Carnegie Mellon U.
UPMC: U. of Pittsburgh Medical Center

EMG: Electromyography
ECE: Electrical and Computer Engineering

- **Education:**
  - 1985/09 ~ 1989/05 - B.S. Dept. of Biomedical Engineering, Chung Yuan Christian University, Taiwan.
  - 1991/08 ~ 1993/05 - M.S. Program of Biomedical Engineering, ECE, Washington U., St. Louis, MO.
  - 1993/08 ~ 1998/04 - Ph.D. ECE, U. of Pittsburgh, Pittsburgh, PA.
    - Research Assistant: Research conducted at the RI, SCS, CMU.
    - Dissertation Title: "Automatic Recognition of Facial Expressions Using Hidden Markov Models and Estimation of Expression Intensity"
    - Advisor: Takeo Kanade, RI, SCS, CMU.
      > Member of National Academy of Engineering
      > Fellow of the American Academy of Arts and Sciences

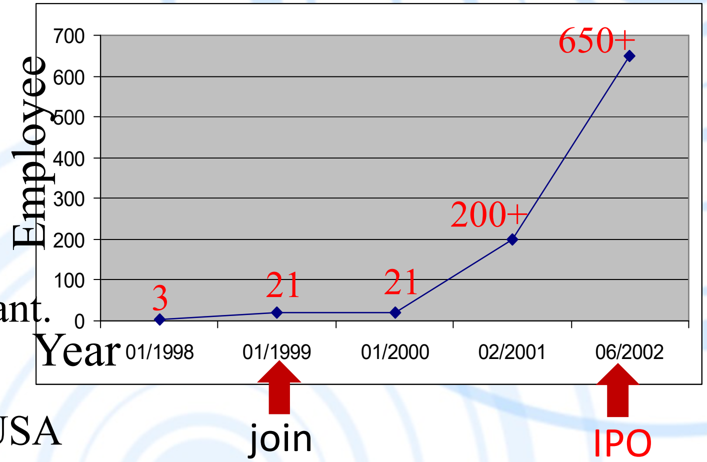
- **Apply facial expression analysis dissertation work to plastic surgery and EMG at UPMC**

1) Original motion trajectory uses attached circle dots and manually mark dot center frame by frame

2) Facial expression analysis uses computer vison with optical flow tracking for motion trajectory

R.1 Profile at Face Recognition Company – Visionics (IPO VSNX): 1999~2002, USA

- **Biometrics technology: Face and fingerprint recognition**
  - Capital: US$40M
  - Award: DARPA FERET Face Recognition Competition: No.1 in 2002 & 2004
  - Grant: Project leader for US$5M DARPA surveillance grant.
  - Business: Government - Surveillance in Birmingham, UK
    - Surveillance for Customs & Border Protection, USA
  - Strategy: - Merge small (200+ employees) and then big (400+) fingerprint companies
    - Partnership with NEC, Japan – made face detection chips in 2005.
  - Stock: IPO VSNX at Nasdaq in 2002



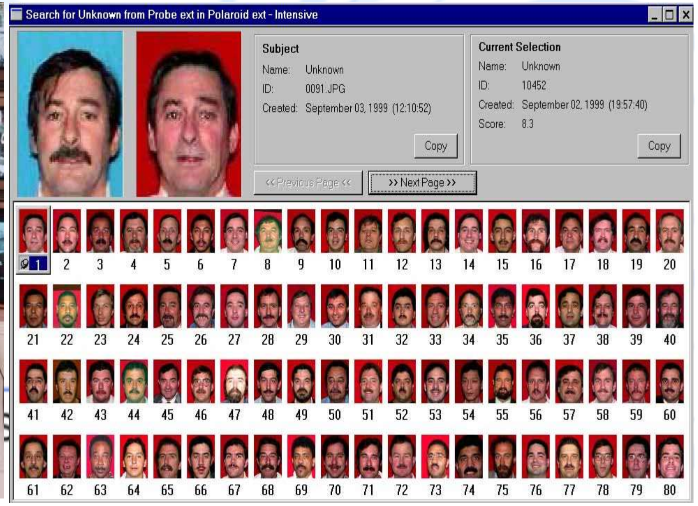
1) Surveillance and security: Preprocessing and display

2) Central management system (CMS) and database center: Recognition process and data management




Face and fingerprinting images

Recognition result: ID, profile, record…

Media: Newspaper about Human-Computer Interaction

- **1998/08/03: *Washington Post, Washington Times*, and *Pittsburgh Post-Gazette* –**

  "Look Closely: Computer program reads deep into our true feelings by analyzing our facial expressions.“

- **2001: ABC, NBC, and CBS TV Channels and Discovery Channel in USA: “Surveillance – Security via CCTV”**

BIG UNIT MENACES BUCS/SPORTS, C-1

Pittsburgh Post-Gazette

ONE OF AMERICA'S GREAT NEWSPAPERS

MONDAY, AUGUST 3, 1998 VOL. 72, NO. 3 8/3/98

Lie detector

Is this woman truly surprised? A computer can tell whether facial expressions are genuine.

Science, Page A-6

LOOK CLOSELY

Computer program reads deep into our true feelings by analyzing our facial expressions

By Sharon Voas
Post-Gazette Staff Writer

If Detective Sipowicz of "NYPD Blue" had this computer program, he wouldn't have to threaten so many suspects with bad things involving big men in prison.

He'd *know* if they were telling the truth most of the time.

A computer program just developed here can tell the difference between a spontaneous, or genuine expression, and one that is deliberate, but not necessarily false. In another year of study it's expected the program will be able to detect a true statement from a lie.

When computer analyses of suspects' expressions are used along with lie detector tests, which are accurate about 90 percent of the time, a detective would know most of the time whether a suspect is telling it straight.

This new computer program, developed by researchers at the University of Pittsburgh and Carnegie Mellon University, can discern thousands of facial expressions. The prototype is so precise that it deciphers the order and speed different facial muscles move in to create a genuine expression. That would be virtually impossible for someone to fake.

The new program is so far superior to existing methods of analyzing facial expressions that Pitt researcher Jeff Cohn says it's like inventing a new microscope.

"You can see so much more, and you can see things you couldn't see before," Cohn said.

The program has potential in mental health research as well, by revealing true feelings and helping doctors learn if a patient is really suicidal.

The program may make a big step toward developing computers users can actually talk to, as if talking to another human, instead of using keyboards and mice. But to have a conversation like two humans would, the computer would have to read the user's expressions. Also, the program may help movie animators draw more realistic expressions.

Automated Face Analysis, the official name of the program, was developed by Cohn, an associate professor of psychology and psychiatry at Pitt, Takeo Kanade, director of The Robotics Institute at Carnegie Mellon, James Lein, who recently received his doctorate in electrical engineering from Pitt, and Adena Zlochower, a doctoral candidate in psychology at Pitt.

Only one other group of researchers from the Salk Institute and the University of California at San Francisco is using a similar program of computer vision and facial analysis. Marian Stewart Bartlett, of the Salk Institute, said these programs will be especially useful for psychological research because researchers will no longer have to rely on inferring what the patient might feel. They will know.

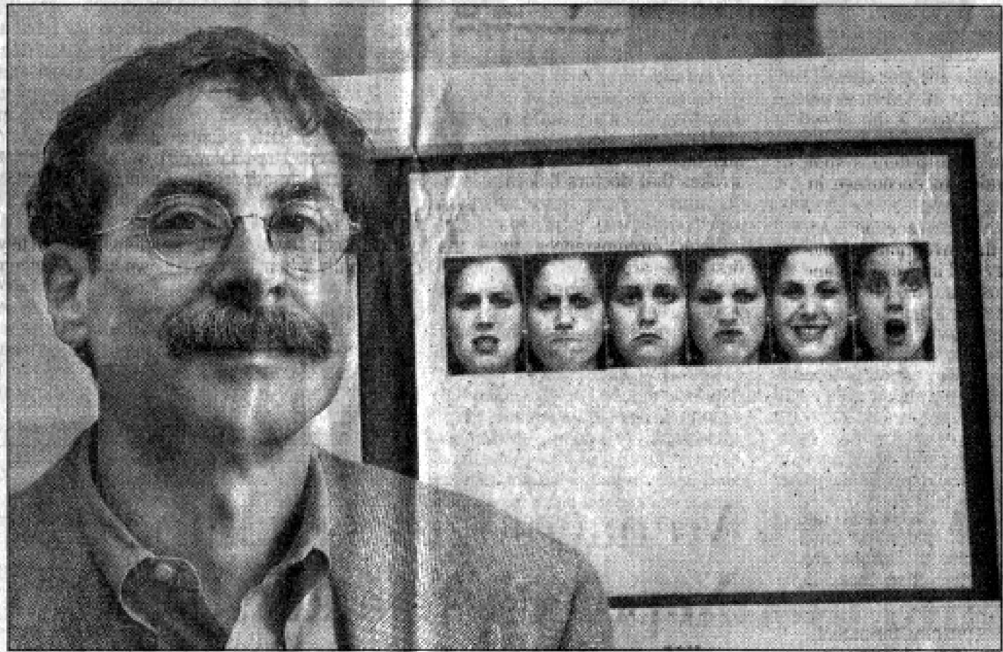
"Facial expression is an important mode of human expression," Cohn said. "It regulates social behavior. It's important to people's self-presentation and to how they are perceived."

Facial expressions tell so much that we sometimes have "gut feelings" that something just doesn't feel right, even though we can't explain why.

For instance, President Clinton's chin boss is giving him trouble these days.

When a person smiles a genuine smile, his whole face moves, with the muscles around the eyes contract-

SEE **FACES,** PAGE A-7



Tony Tye/Post-Gazette

Jeff Cohn, associate professor of psychology and psychiatry at Pitt, says of the new program: "You can see so much more, and you can see things you couldn't see before."

SCIENCE & ENVIRONMENT

PITTSBURGH POST-GAZETTE ■ MONDAY, AUGUST 3 , 1998

Jeff Cohn/University of Pittsburgh

Adena Zlochower, a Pitt researcher working on a computer analysis of facial expressions, shows how muscles move in some basic expressions.
ABOVE LEFT: In joy, the muscles around her eyes tighten, making her cheeks rise and pulling up the corners of her lips and opening her mouth.
ABOVE CENTER: In anger, her lips press together and pull toward the middle, making a ball-shaped area called the "chin boss" push up. Eyes widen, the lower eyelid tightens and upper eyelids rise. Eyebrows draw together sharply.
ABOVE RIGHT: In surprise, the jaw drops, stretching her mouth down, while her eyes widen and her brows rise.

PITTSBURGH POST-GAZETTE ■ MONDAY, AUGUST 3, 1998 A-7

SCIENCE & ENVIRONMENT

Computer deciphers genuine, deliberate expressions

FACES FROM PAGE A-6

ing and the cheeks raising the mouth. The top of Clinton's face smiles, Cohn says, but Clinton pushes his lower lip up into what could be taken for a pout or some expression of sadness. Pulling that lip up tight creates a tennis-ball shaped area on the chin called the chin boss.

"That gives people a mixed message," Cohn said. "They're not quite sure about his smile."

Or take Susan Smith, the South Carolina woman who told police an attacker had kidnapped her 1- and 3-year-old sons, when she had actually belted them into her car and rolled it into a lake where they drowned.

"I knew right off she was lying," Cohn said. "When she was interviewed, she feigned sadness. In sadness, a characteristic is that the inside corners of the brows are pulled together and raised in a triangle, and the lip corners pull down. Very few people can do that voluntarily."

Cohn, 51, leans back in his office chair near a computer screen full of emoting faces. He's slender and wears the moss green cable knit sweater, khaki shorts and hiking shoes of an intellectual who spends a lot of time hiking or biking in the woods. A slender nose leads from his large blue eyes — that look larger because of thick-lensed, wire-rimmed glasses — to his thick brown and gray thatch of a mustache.

He swivels and sets the faces on the computer screen in motion, or, one could say, in emotion.

The program tracks each and every pixel — the tiniest unit of light on a computer or TV screen — of colleague Zlochower's face all the way through an expression. Little tracker lines on her face show not only which muscles move on different parts of her face, but in what order, with what intensity and how long each part lasts.

Those things weren't known before the development of this program because it uses video.

"Imagine Susan Smith has done some homework and was able to display sadness," Cohn said. "She would still have trouble performing that action in a realistic way because she'd have to pull her eyebrows up and her lip corners down in the right order and fast enough."

Most other methods of facial expression analysis focus on prototypes of a few basic expressions such as anger, joy, disgust, fear, sadness, surprise and so forth.

"But people very rarely show these peak expressions," Cohn said. "There are probably thousands of common expressions. So Automated Face Analysis examines components of expressions — a furrowed brow, widened eyes, tightened lips. We can [recognize] extreme expressions and subtle expressions."

Automated Face Analysis breaks components of facial analysis down into AUs, or action units, the smallest unit of change in an expression. How big is an AU?

"Think of Clint Eastwood giving someone a dirty look by narrowing his eyes by tightening his lower eyelids," Cohn said. "That's an AU-7."

AUs give researchers a much larger vocabulary to describe expressions.

"If we only have the terms for basic expressions, then [researchers] have an inadequate vocabulary to describe expression."

So, instead of saying someone looks happy, which may be deceiving, researchers might say she is displaying an AU-6 + 12 or an AU-6 + 12 + 20, types of smiles.

Other methods of facial analysis also use photos of people looking straight ahead and not moving. But in reality, you may close your eyes and shake your head.

Cohn turns back to the computer and calls up a series of photos of an apple-cheeked baby that start with the glimmer of a smile and move into a full-blown beaming smile that would elicit "oooohs."

As the baby's smile progresses, he throws his head back and to the right. This computer program can account for that as part of the expression.

And babies' expressions are always spontaneous.

Jeffrey Cohn

These images from the Automated Face Analysis Program show how the program tracks the movement of muscles around the mouth and eyes when the subject is registering sadness, fear and disgust.

R.2 Robotics Lab. at NCKU: 2002~ (1/2)

- 約24名研究生 + 4名大學專題生

- 2014 研究組別: AI智慧製造、AI半導體、AI醫材、AI運動分析

  1. 深度學習組
     (Deep Learning Team)

  2. 自走車 + 機械手臂組
     (AGV + Robot Arm Team,
     Automatic Guided Vehicle + Visual-Guided Robot

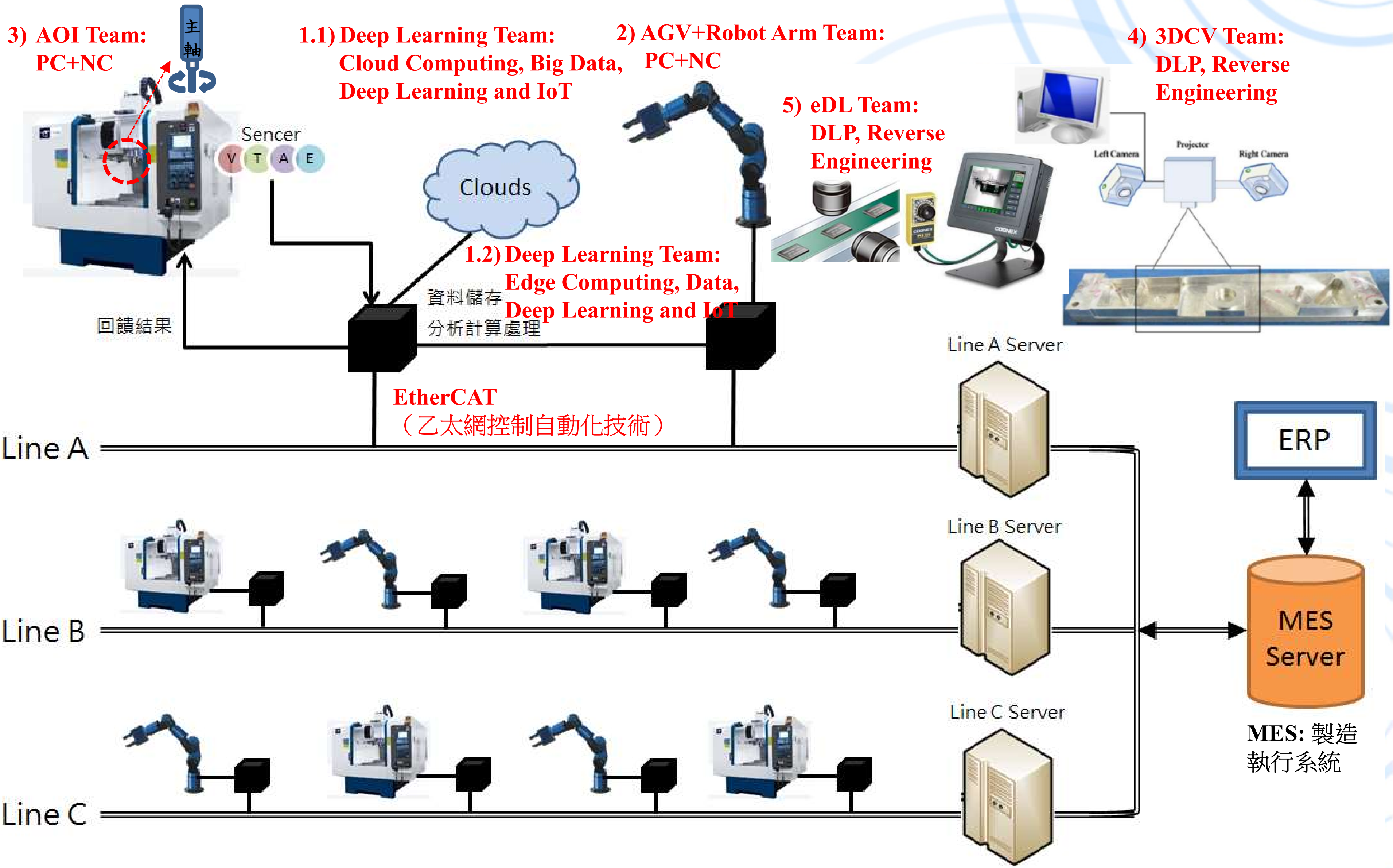
  3. 自動光學檢測組
     (AOI Team,
     Automatic Optical Inspection)

  4. 3D電腦視覺組
     (3D Sensor and CV Team,
     3D Sensor and Computer Vision)

  5. 嵌入式深度學習組
     (eDL Team,
     Embedded Deep Learning)

  6. 生物醫學工程 + 運動分析組
     (BME + Sports Analytics Team,
     Biomedical Engineering + Sports Analytics)

R.2 Relationship between Teams

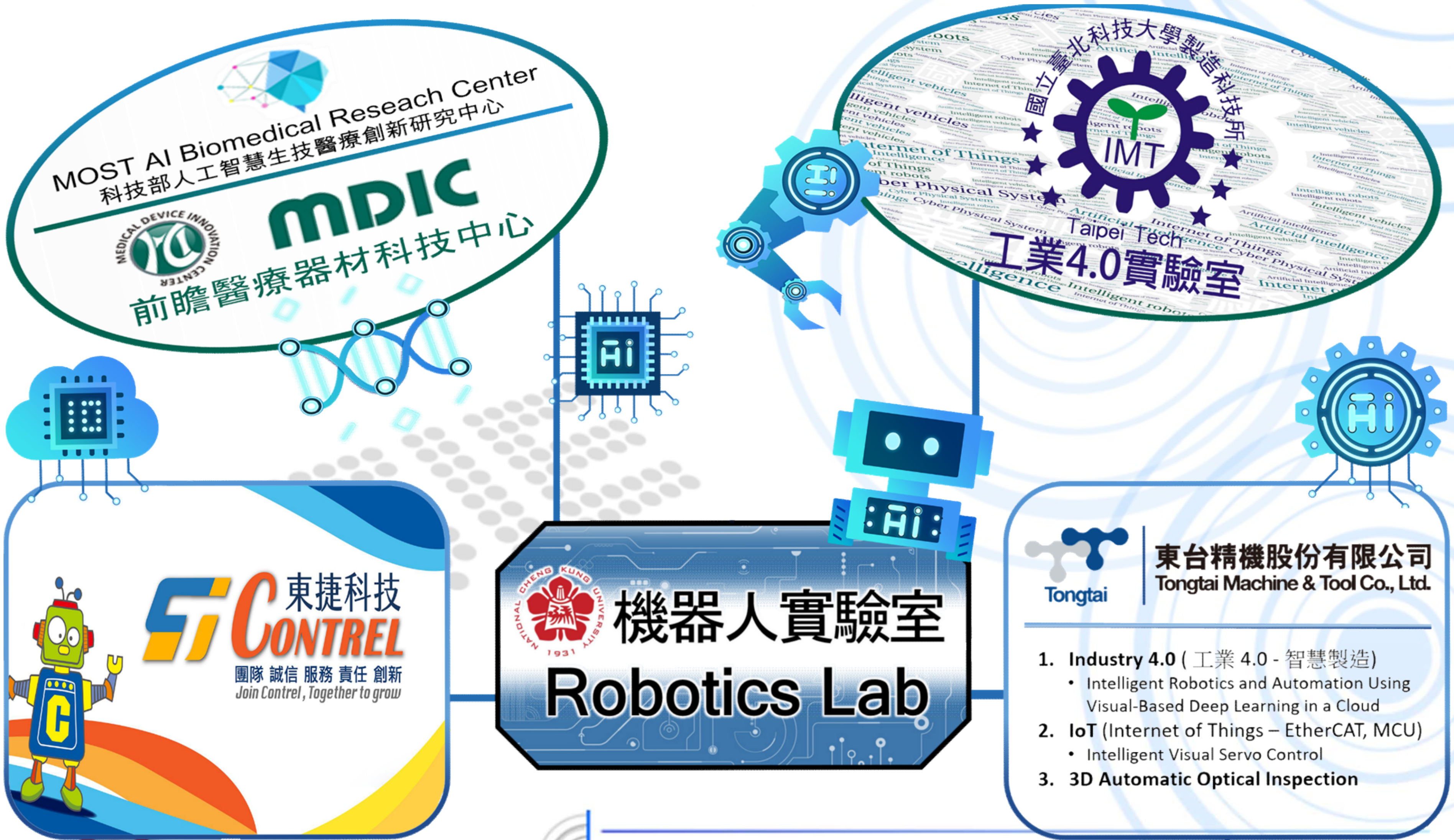


MES: 製造
執行系統

ERP（Enterprise Resource Planning）:企業資源計劃或稱企業資源規劃簡稱，應具備以下功能：基本資料與管理維護、庫存管理、採購進貨管理、配銷管理、財務管理、人資/事務管理、生產管理與決策支援管理等系統功能..

R.2 Robotics Lab. at NCKU: 2002~ (2/2)

- ❑ 2014 合作夥伴與領域: AI智慧製造、AI半導體、AI醫材（、AI農業）



1.0 History: Computer Vision and Deep Learning

1. Industrial Revolution:

3)1969, Industry 3.0:
Computing and Automation

4)2013, Industry 4.0:
Smart Manufacturing

1)1981, 生產力1.0：程式化

2)1991, 生產力2.0：整線化　3)2001, 生產力3.0：電子化　4)2011, 生產力4.0：智慧化

1990　2000　2010　2020

2. OpenCV: Image Processing, Computer Vision, Machine Learning, and Deep Learning

0)1999, OpenCV0.0 Started　1)2007, OpenCV1.0: IP+CV　3)2016, OpenCV3.0:
IP+CV+ML

2)2009, OpenCV2.0:
IP+CV

4)201x, OpenCV 4.0:
IP+CV+ML+DL?

1)1999, OpenCV Supported by
Intel

2)2010, OpenCV Supported by NVidia
3)2011, OpenCV Supported by Google

3. Deep Learning:

1)1989, LeCun tPami1989　**2)1998, LeNet IEEE1998**　**3)2012, AlexNet NIPS2012**

4. Robotics Lab.: Computer Vision and Machine Learning – 1991 ~ 2013

1)1993, Facial Expression Recognition at CMU, USA + DARPA Surveillance (VSAM) Project

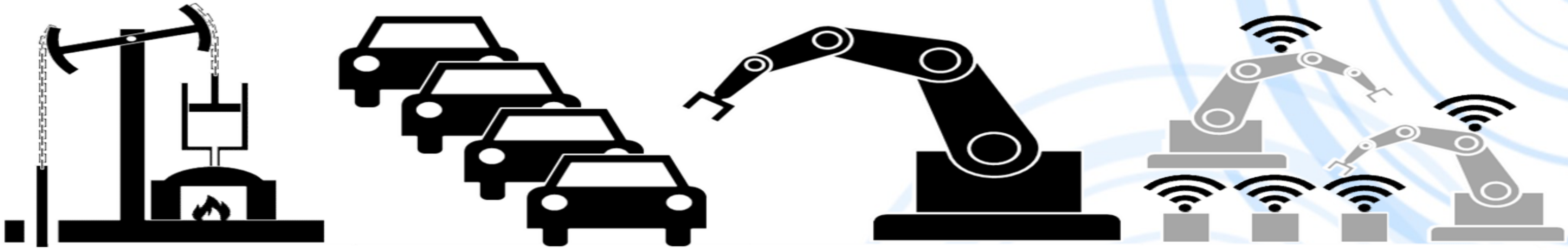
2)1999, Face Detection and Recognition Using Stereo at Visionics, USA + DARPA Project

3)2004, Automatic Optical Inspection (AOI System Integration):
TFT-LCD with 奇美、友達 + Solar Cell with 益通

4)2009, eCV Surveillance + ADAS with 美國德儀 (TI DSP), 智原(ASIC) + Stereo Sensor with 鈺創 + Embedded AOI with 研華

5. Robotics Lab.: Computer Vision and Deep Learning - Since 2014: AI智慧製造+AI半導體+AI醫材+AI

1.1 Industrial Revolution 1.0~4.0: History



1st (1765)	2nd (1870)	3rd (1969)	4th (2013)
Mechanization, water power, steam power	Mass production, assembly line, electricity	Computer and automation	Cyber Physical Systems

(105 years) (99 years) (44 years)

商業周刊

國立成功大學

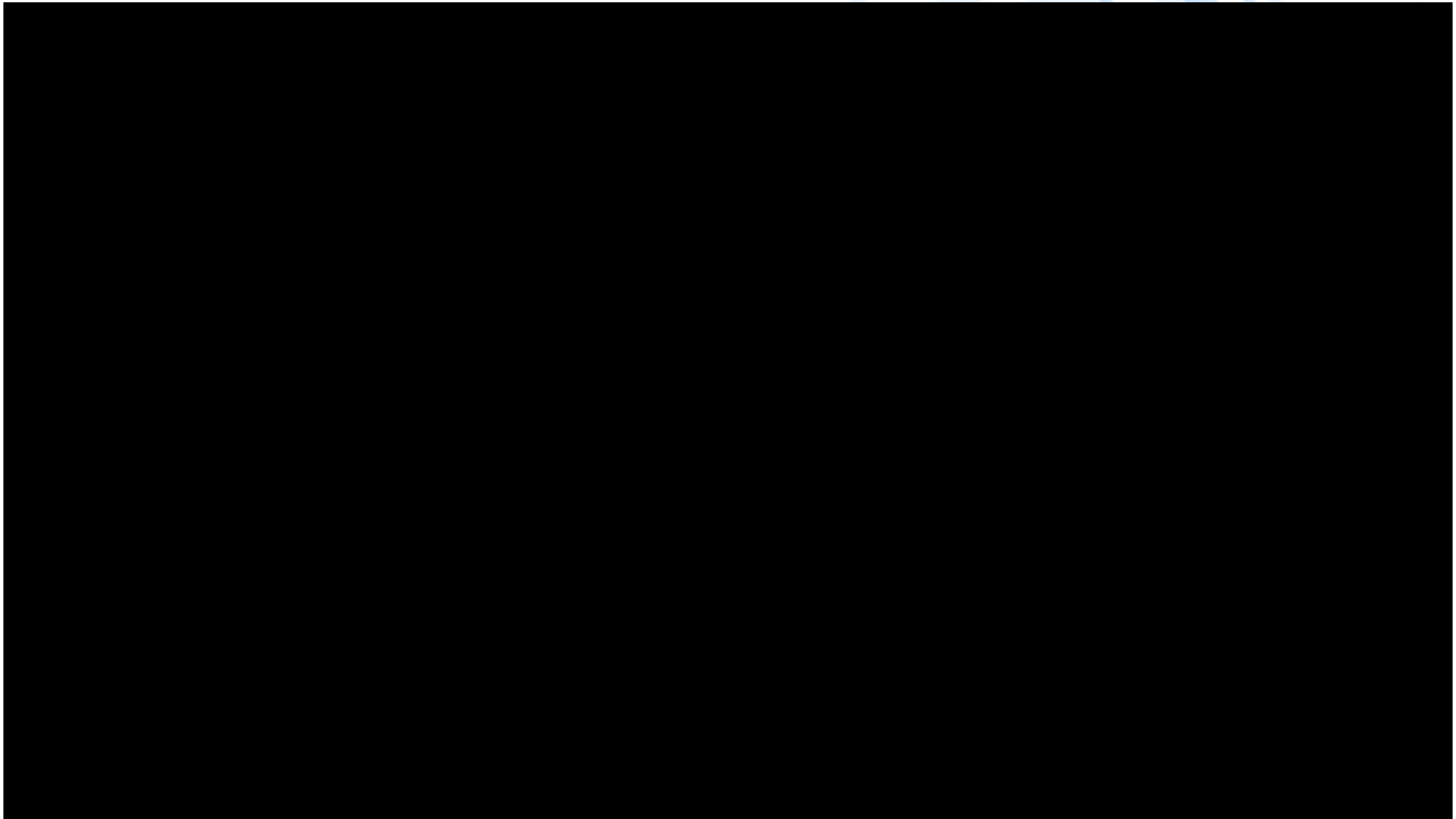
1.1 Industrial Revolution 1.0~4.0: Industry 4.0 in 2013

- Industry 4.0 – Smart Manufacturing:
  - To connect all <u>means of production</u> to enable their interaction in real-time

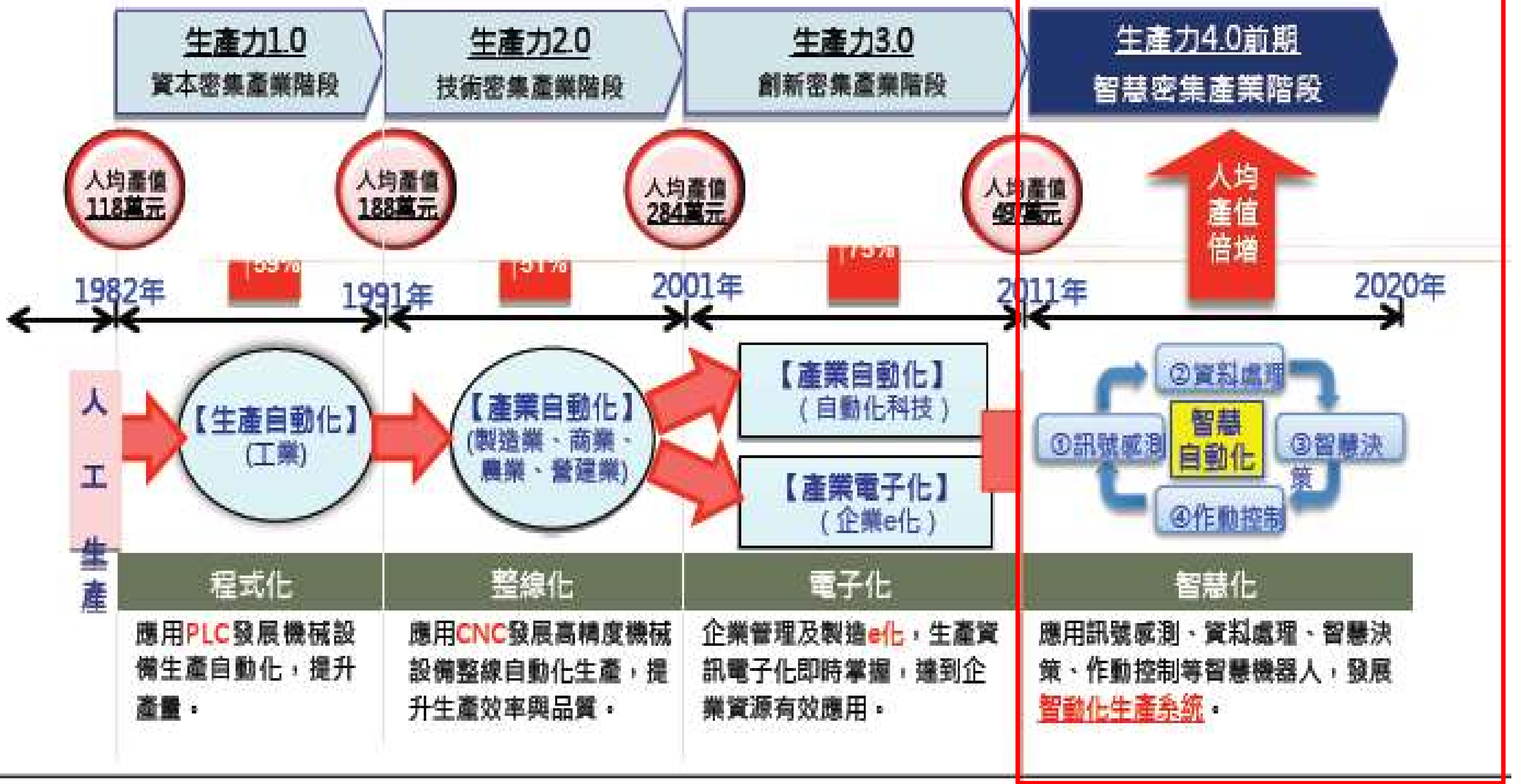
- Technology consists of:

1) Cloud: Big Data Analysis
2) Internet of Things (IoT) : 5G + 3D (RGB-D)
3) Cyber Physical System (CPS,網絡實體系統,人機物融合系統)
   (1) Simulator
       CAD, VR, ….
   (2) Reality
       a) Computing (algorithm),
       b) Sensor
       c) Actuator (controller)

1.1 Industrial Revolution 1.0~4.0: Industry 4.0 in 2013



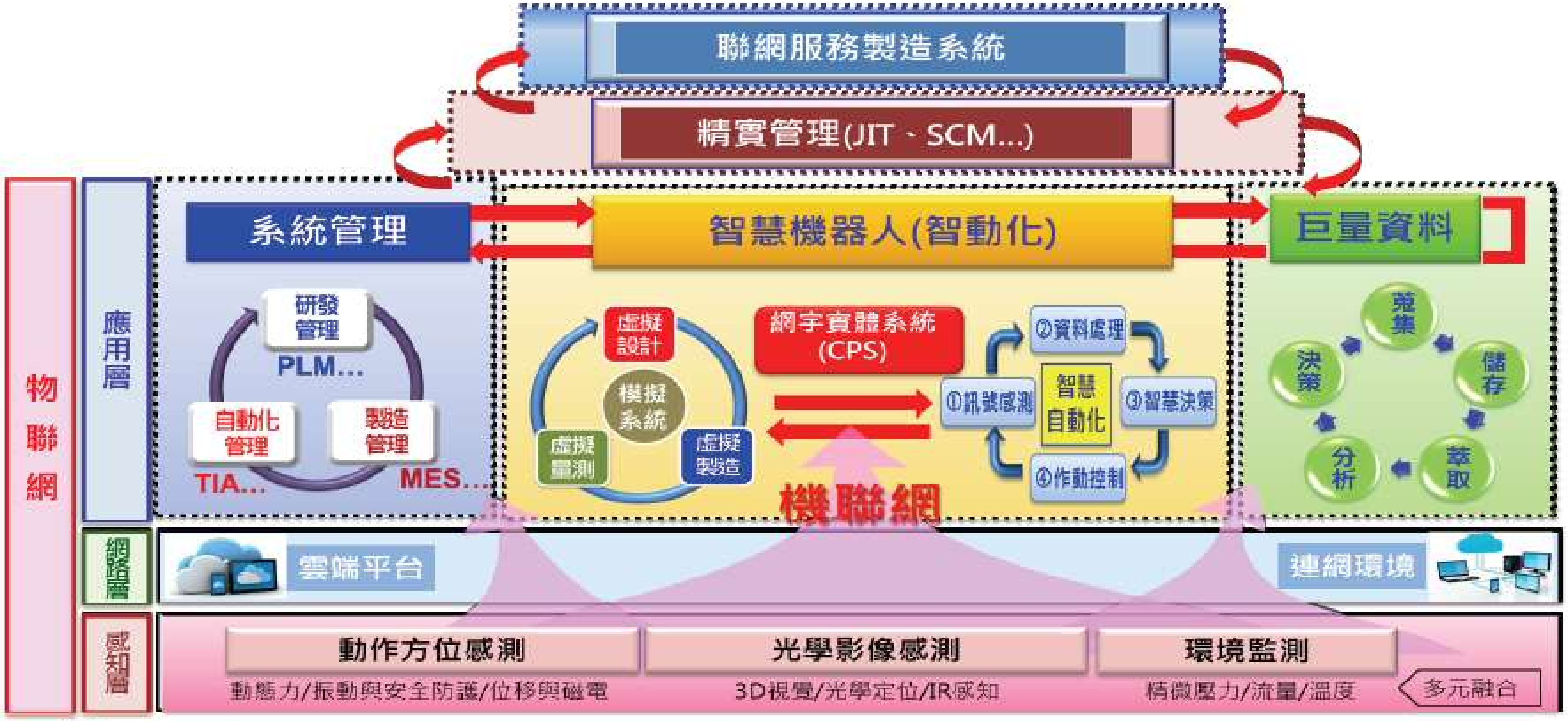
1.2 生產力 4.0: IoT (Internet of Things) and Automation – History



1.2 生產力 4.0: IoT (Internet of Things) and Automation

台灣競爭力優勢分析(生產力4.0系統架構)

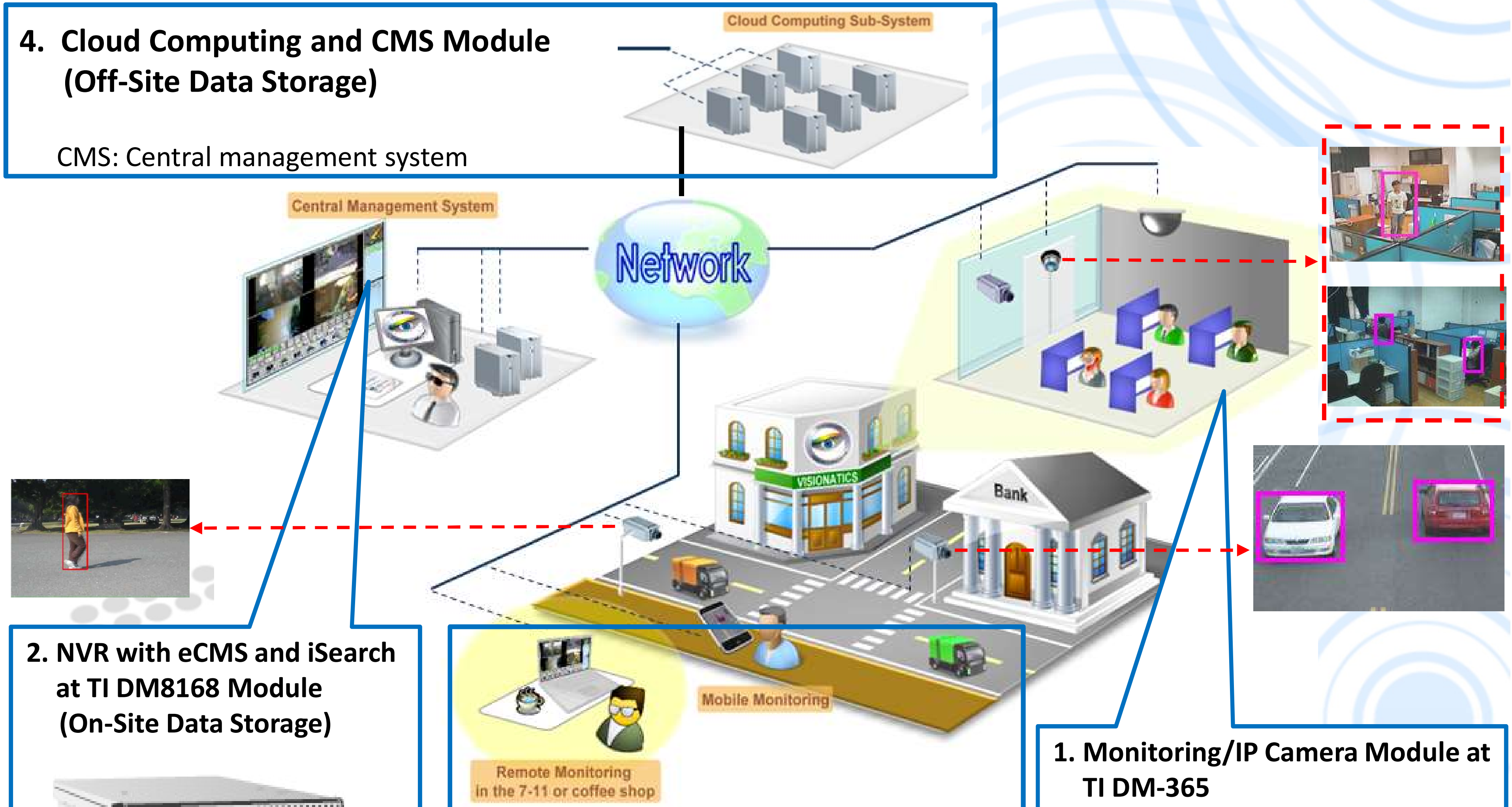
– 透過物聯網將生產資訊數位化，並延伸至機器端形成機聯網，再藉由系統管理、巨量資料(製造+服務)技術、以及精實管理，達成聯網服務製造系統之創新營運模式。



備註：1.JIT(Just In Time)；SCM(Supply Chain Management)
2.PLM (Product Lifecycle Management)；MES (Manufacturing Execution System)；TIA (Totally Integrated Automation)

1.2.1 Application 1: eCV Surveillance System 2009~2013

❑ **TI（美國德儀）Embedded Surveillance System: NVR (Network Video Recorder) + IP Cam + CMS**



4. Cloud Computing and CMS Module (Off-Site Data Storage)

CMS: Central management system

2. NVR with eCMS and iSearch at TI DM8168 Module (On-Site Data Storage)

- 32-Ch iNVR with eCMS and iSearch

3. Remote Monitoring Module

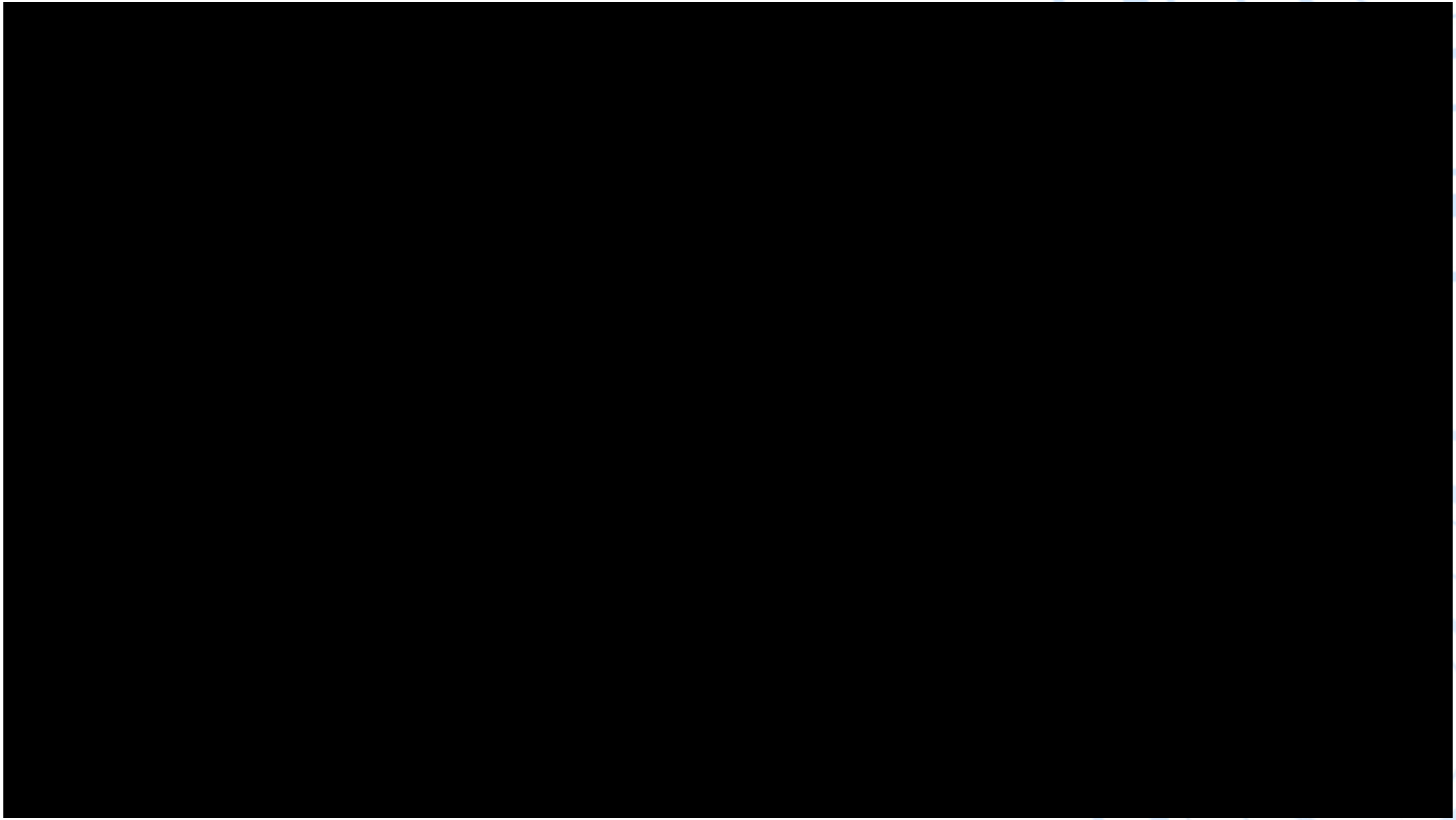
- Mobile Phone: Android, iOS, WinPhone
- Tablet: Android, iOS…
- NoteBook: Windows

1. Monitoring/IP Camera Module at TI DM-365

- IP CAM, PTZ IP CAM or
- Analog Camera + Video Server Box

21

1.2.2 Application 2: Scraping System (5/5) 01:34.77



2.0 History: Computer Vision and Deep Learning

1. Industrial Revolution:

3)1969, Industry 3.0:
Computing and Automation

4)2013, Industry 4.0:
Smart Manufacturing

1)1981, 生產力1.0：程式化

2)1991, 生產力2.0：整線化

3)2001, 生產力3.0：電子化

4)2011, 生產力4.0：智慧化

1990 **2000** **2010** **2020**

2. OpenCV: Image Processing, Computer Vision, Machine Learning, and Deep Learning

0)1999, OpenCV0.0 Started

1)2007, OpenCV1.0: IP+CV

2)2009, OpenCV2.0:
IP+CV

3)2016, OpenCV3.0:
IP+CV+ML

4)201x, OpenCV 4.0:
IP+CV+ML+DL?

1)1999, OpenCV Supported by
Intel

2)2010, OpenCV Supported by +NVidia

3)2011, OpenCV Supported by +Google

3. Deep Learning:

1)1989, LeCun tPami1989

2)1998, LeNet IEEE1998

3)2012, AlexNet NIPS2012

4. Robotics Lab.: Computer Vision and Machine Learning – 1991 ~ 2013

1)1993, Facial Expression Recognition at CMU, USA

2)1999, Face Detection and Recognition at a Distance Using Stereo at Visionics, USA

3)2004, Automatic Optical Inspection (AOI System Integration):
TFT-LCD, Solar Cell…

4)2009, eCV Surveillance with TI & Faraday

5. Robotics Lab.: Computer Vision and Deep Learning - Since 2014: AI智慧製造+AI半導體+AI醫材+AI

2.0.0 Computer Vision and Deep Learning – Basic Courses and Books

- **Undergraduate:**
  - **Linear Algebra**
  - **Introduction to Probability and Stochastic**
  - **Signals and Systems**

    - Alan V. Oppenheim and Alan S. Willsky, *Signals and Systems*, 2nd, 1996.
- **Graduate**
  - **Probability and Stochastic**

  - **Digital Signal Processing**

    -A.V. Oppenheim, R.W. Schafer, and J.R. Buck, *Discrete-Time Signal Processing*, Prentice Hall, 2nd, 1999. ISBN: 0137549202.

  - **Digital Image Processing**

    -R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, Prentice Hall, 3rd, 2007. ISBN: 013168728X.

  - **OpenCV**

    -G. Bradski and A. Kaebler, *Learning OpenCV, Computer Vision with the OpenCV Library*, O'Reilly, 2008. ISBN-10: 0596516134 or ISBN-13: 978-0596516130.

  - **Computer Vision**

    -R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010. ISBN-10: 1848829345 or ISBN-13: 978-1848829343.

    -R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd, 2004. ISBN: 0521540518.

  - **Machine Learning and Pattern Recognition**

    -C.M. Bishop, *Machine Learning and Pattern Recognition*, Springer, 2007. ISBN: 0387310738.

  - **Deep Learning**

    -I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT, 2016. ISBN: 0262035618

2.0.0 Computer Vision and Deep Learning – Relative Conferences and Journals

- **Conference – CV and Deep Learning:**
  - CVPR: Computer Vision and Pattern Recognition
  - ICCV: International Conference on Computer Vision
  - ECCV: European Conference on Computer Vision
  - ACCV: Asian Conference on Computer Vision

  - NeurIPS: Advances in Neural Information Processing Systems
  - ICML: International Conference on Machine Learning
  - arXiv

  - ICME:
  - ICPR:
  - ICIP:

  - eCVW: Embedded Computer Vision Workshop

- **Conferences - Robot:**
  - IROS: IEEE International Conference on Intelligent Robots and Systems
  - ICRA: International conf. on Robotics and Automation
  - SMC: IEEE International Conference on Systems, Man, and Cybernetics

- **Journal – CV and Deep Learning:**
  - IJCV: International Journal on Computer Vision
  - IEEE tPAMI:
  - Image and Vision Computing

  - IEEE tVisualization and Computer Graphics
  - SIGGRAPH:

  - Pattern Recognition
  - IEEE tIP
  - IEEE tCSVT
  - IEEE tMM

- **Journal – Robot:**
  - IEEE tRobotics and Automation

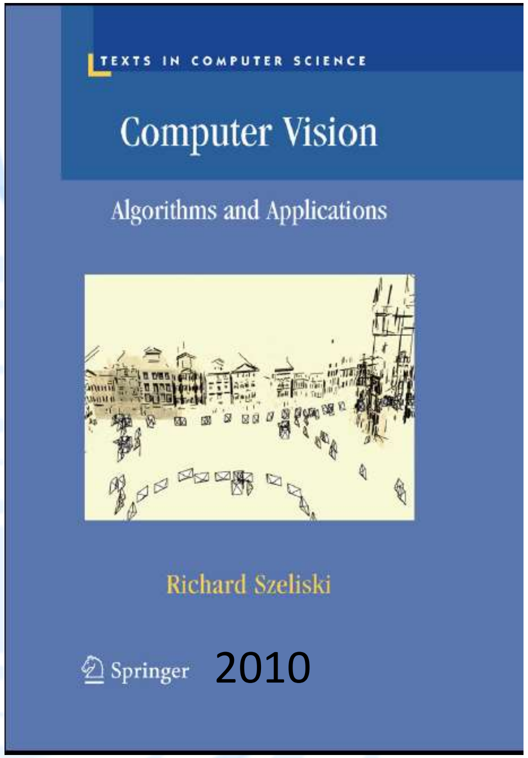
2.0.0 Computer Vision and Deep Learning - Functions

Visual Servo Control:

Robotics / Automatic Optical Inspection (AOI)

2D/3D Sensor (Optics) + Vision + Control + Mechanics + Power

Computer Vision: From Theory to Application, 2002, Jenn-Jier James Lien



Computer Vision and Pattern Recognition / Deep Learning

1. Image Signal Processing (ISP) (Preprocessing or Low-Level Image Processing)

2. Detection / Extraction / Segmentation

3. Tracking / Registration

4. Recognition / Classification

5. Stereo / 3D Reconstruction

6. Synthesis

Computer Graphics: Augmented Reality, Virtual Reality, …

3D Printing

Reverse Engineering

7. Multimedia: Content-Based Image/Video Retrieval (CBVR)

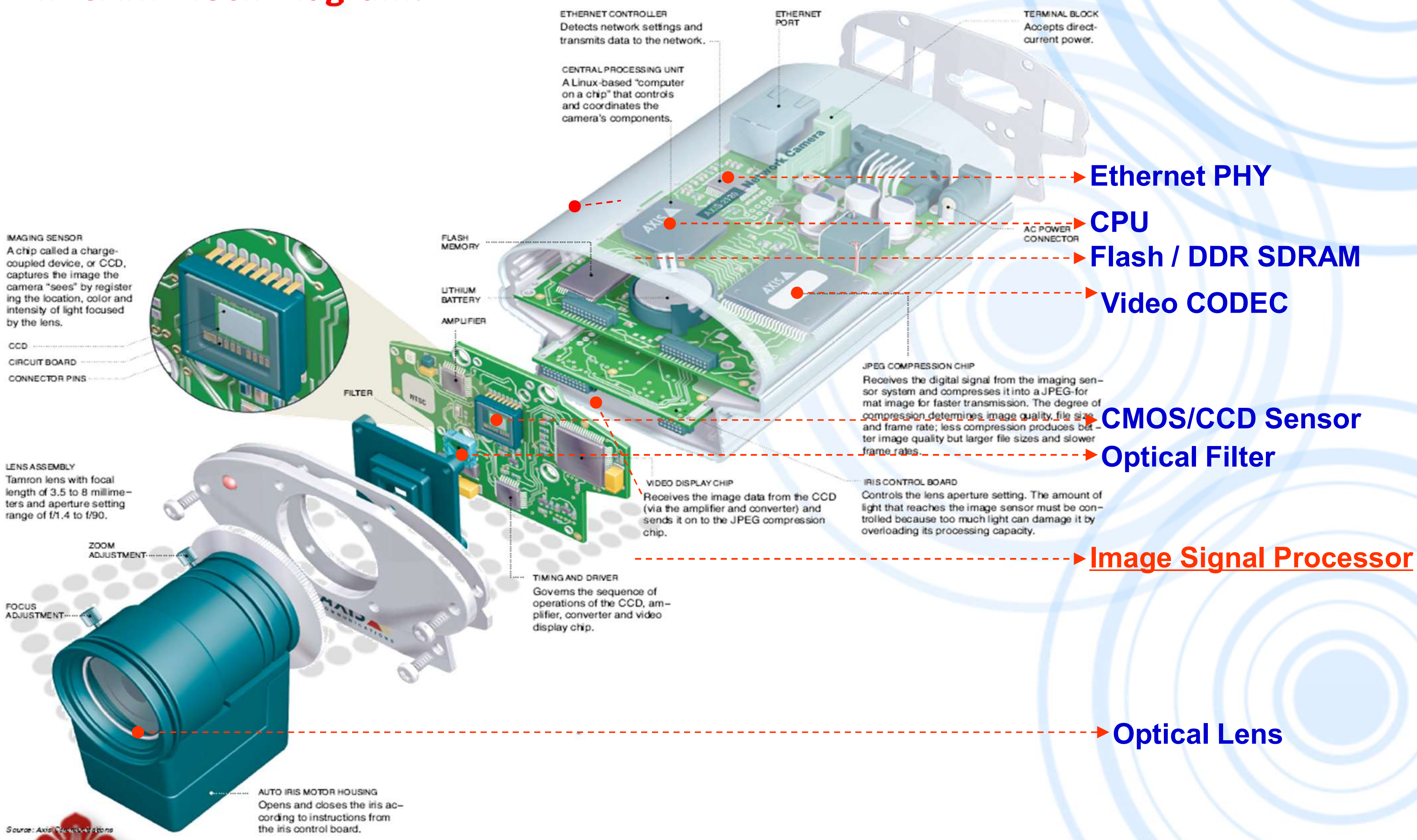
8. Human-Computer Interaction (HCI) or Human-Robot Interaction

ISP (Image Signal Processing): Function Block 1
Embedded IVA (Intelligent Video Analytics): Function Blocks 2~7

ARM: Function Blocks 1~3
DSP: Function Blocks 3~7

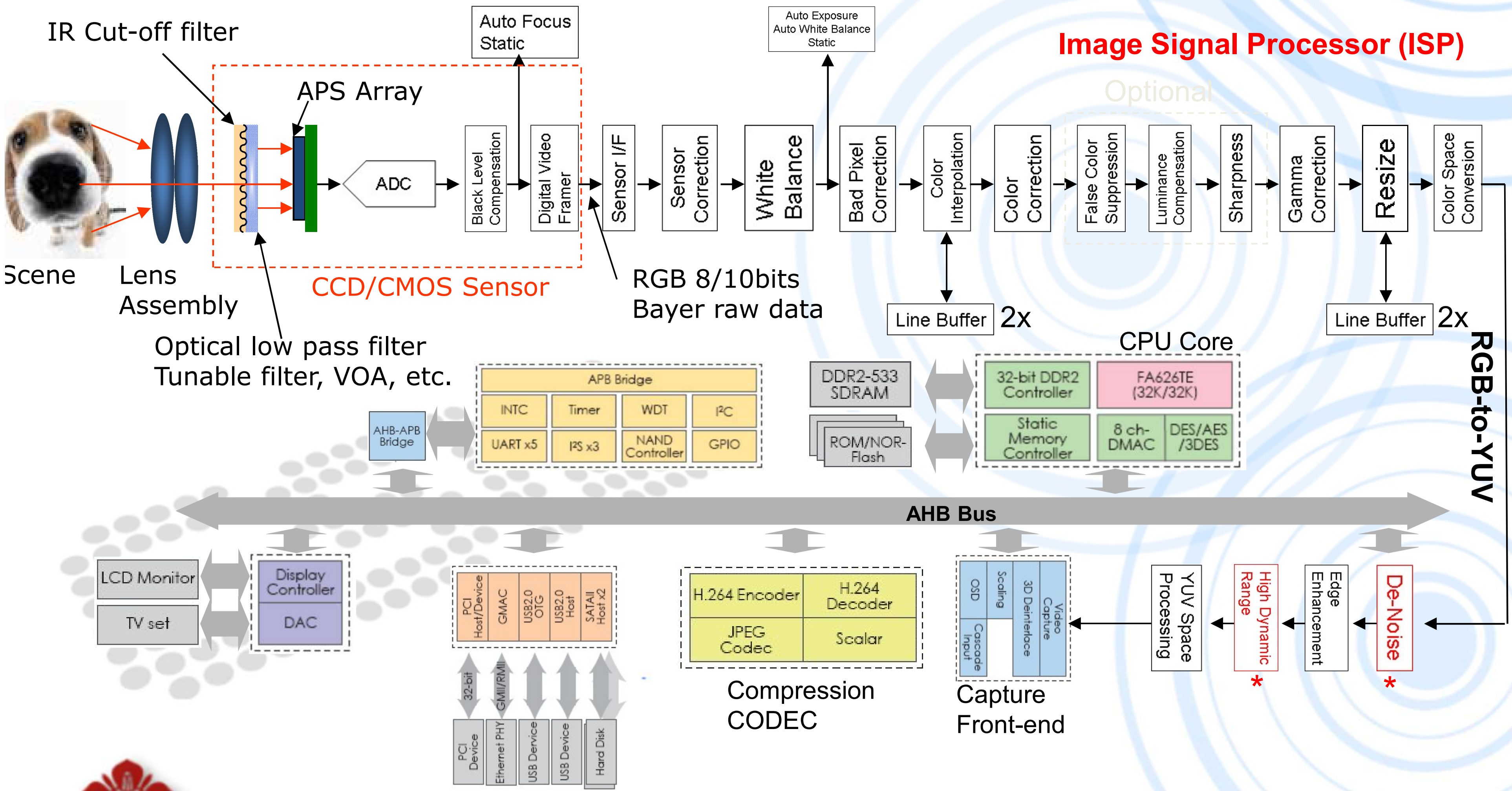
2.0.1 Image Signal Processing (ISP): Sensor – IP Camera (1/2)

- **IP CAM Block Diagram:**



2.0.1 Image Signal Processing (ISP): Sensor – IP Camera (2/2)

❑ IP CAM Function Block Diagram:



國立成功大學
Department of Computer Science and Information Engineering
Robotics Lab

2.0.1 ISP: High Dynamic Range (HDR)

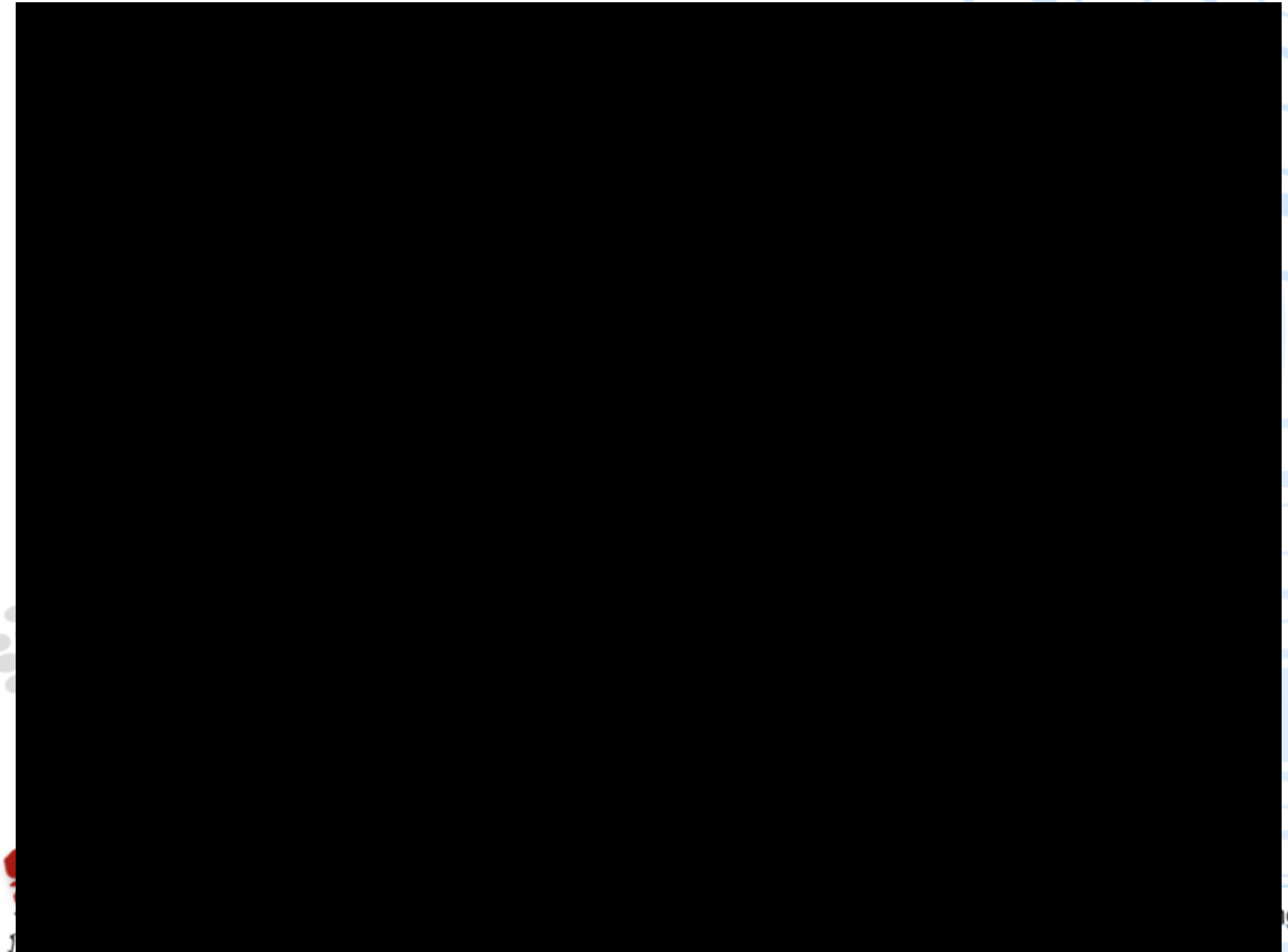
2.0.1 ISP: Stabilization – Fixed Scene Vs. Panning Scene

1) Fixed Scene Stabilization

2) Panning Scene Stabilization

國立成功大學

2.0.1 ISP: Panorama



2.0.2 Face Detection Using AdaBoost Under Various Views

- Face detection – Pose estimation

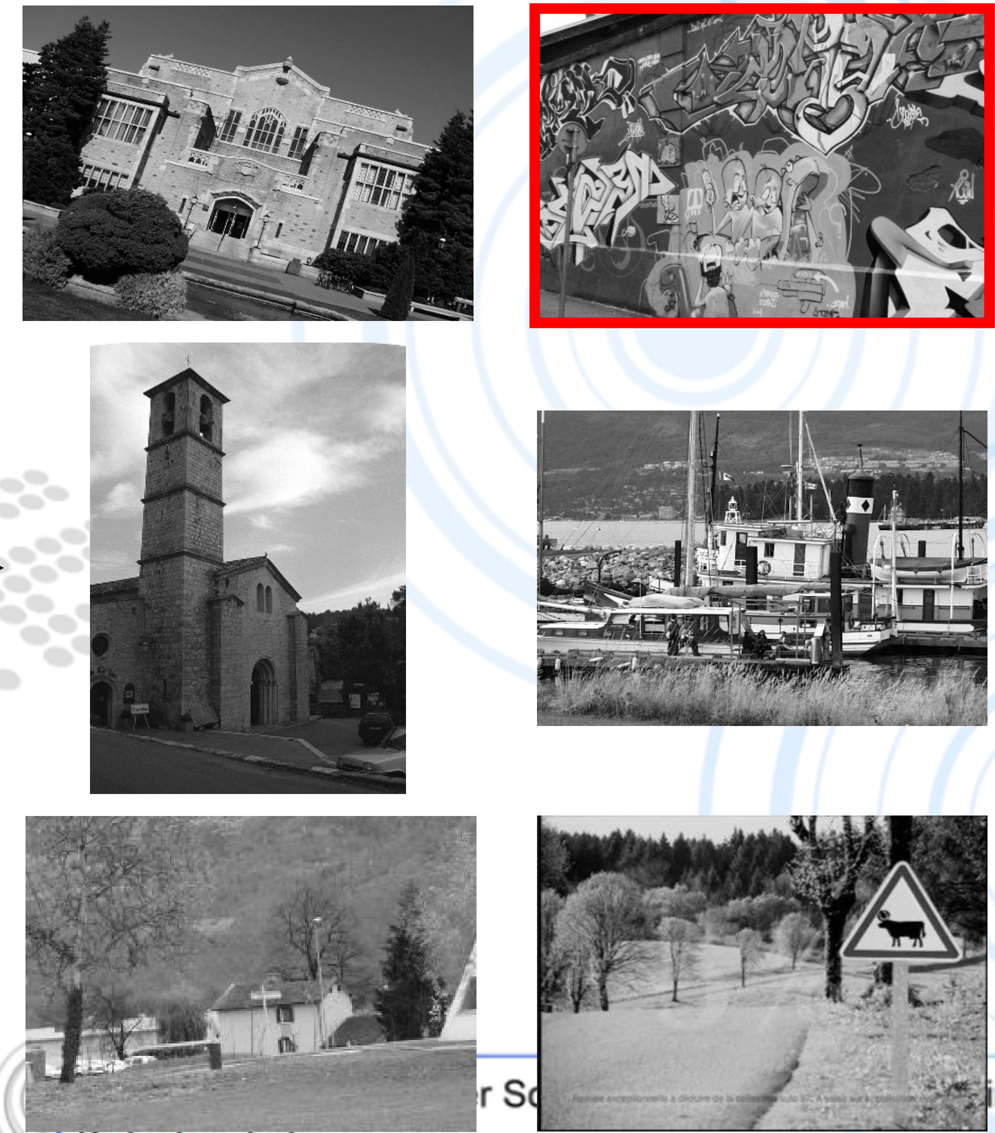
2.0.2 Point Detection and Tracking: Shape Extraction Using DCM (Direct Combined Model)

- From Lucas-Kanade optical flow to ASM/AAM to DCM

2.0.2 Feature Extraction and Tracking: SIFT

2.0.2 Feature Extraction: SIFT - Image Retrieval

change in viewing angle
+ scale change

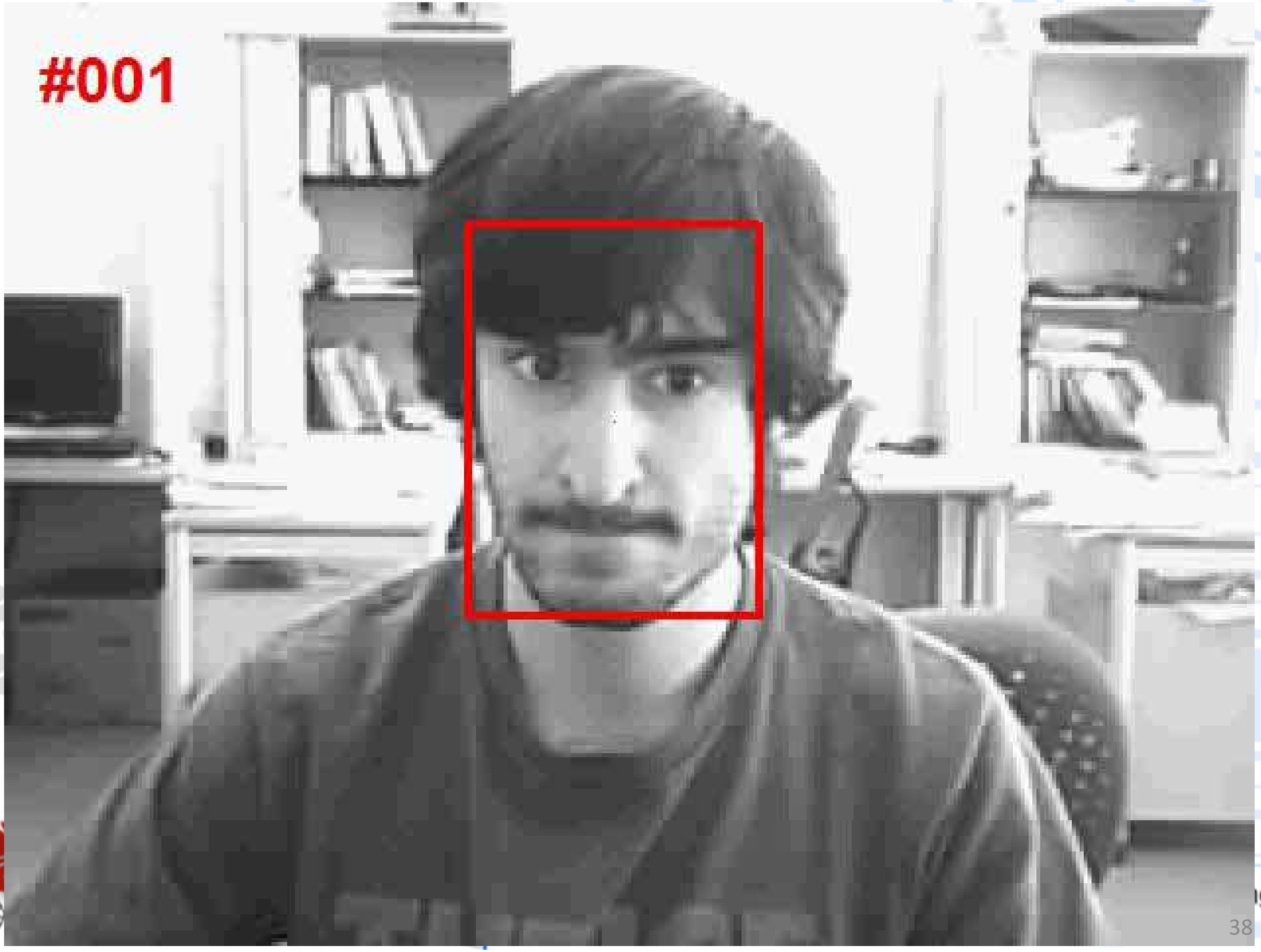


• • •
> 5000
images

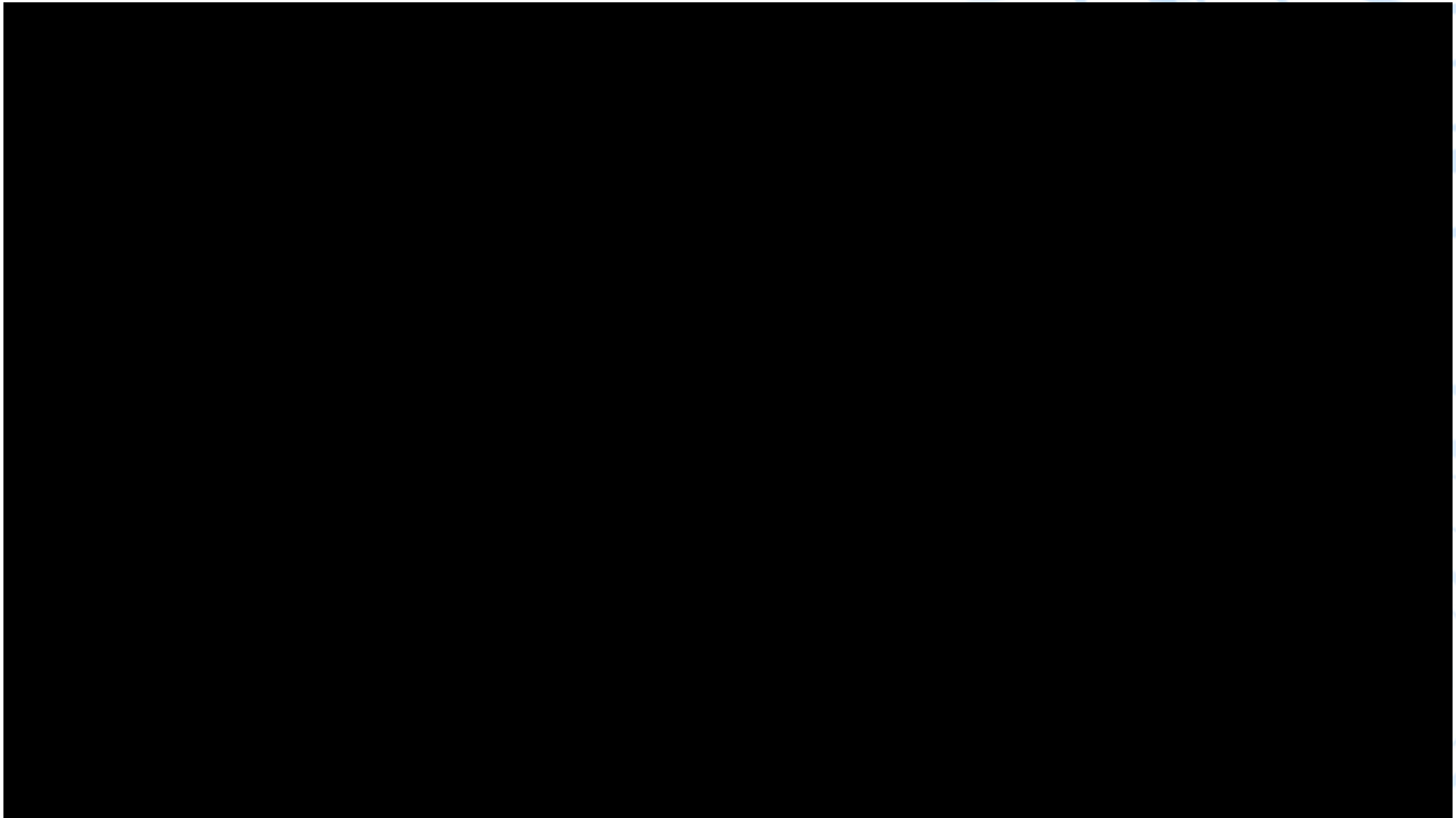
2.0.2 Point Tracking and Synthesis:

2.0.3 Tracking: Moving Vehicle

2.0.3 Tracking: Occlusion



2.0.3 Feature Tracking and Panorama: Slow Motion in Sports



2.0.3 Skeleton Detection and Tracking

2.0.3 Skeleton Detection and Tracking

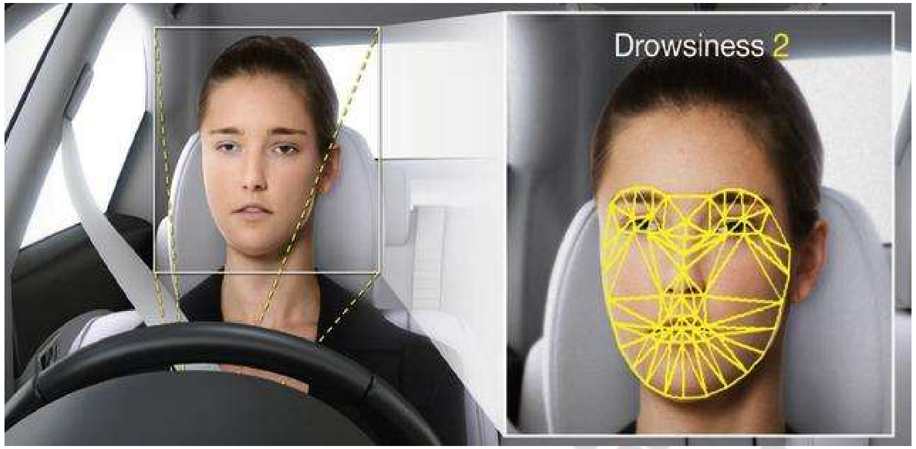
國立成功大學
Department of Computer Science and Information Engineering
Robotics Lab

*Pictures and videos are from the internet (Google and YouTube).

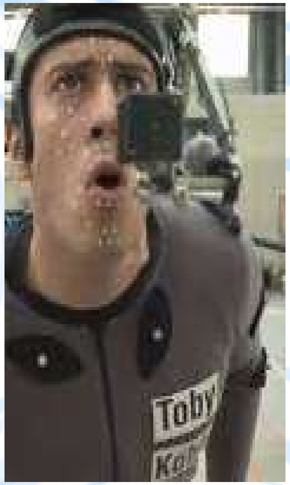
2.0.3 Point Tracking and AR: Feature Points and Mesh

- Technology: Facial Feature Point Detection and Tracking

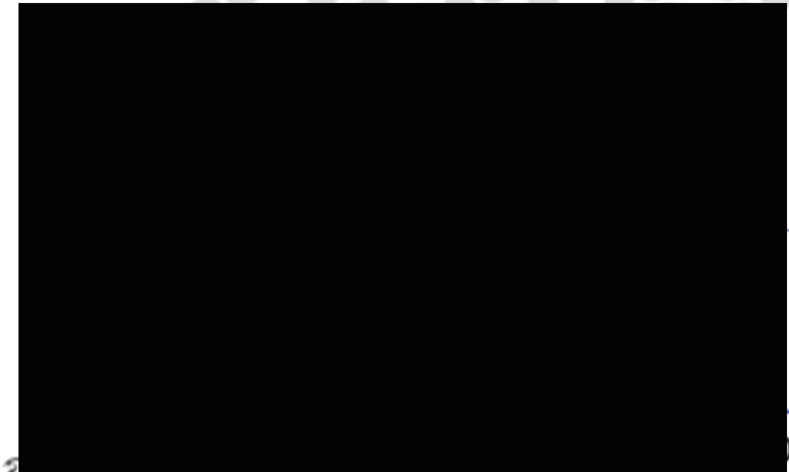
1. Drowsy Driving Awareness System (Eye and Mouth)



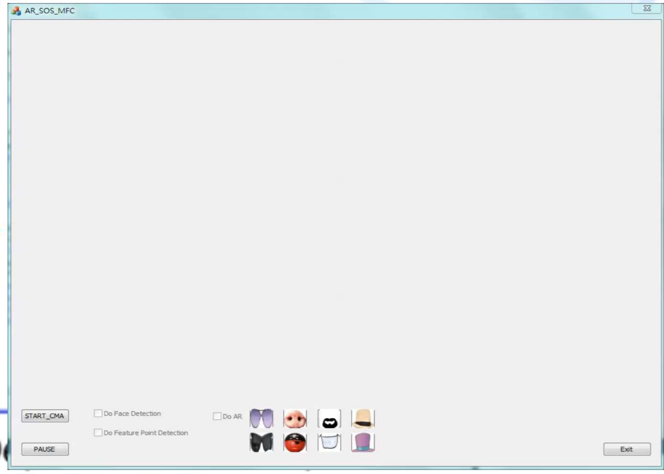
2. Animation of Movie Virtual Effects (Facial contour, eyes, eyebrows, mouth and nose)



3. Virtual Makeup (Facial contour, eyes, eyebrows, mouth and nose)



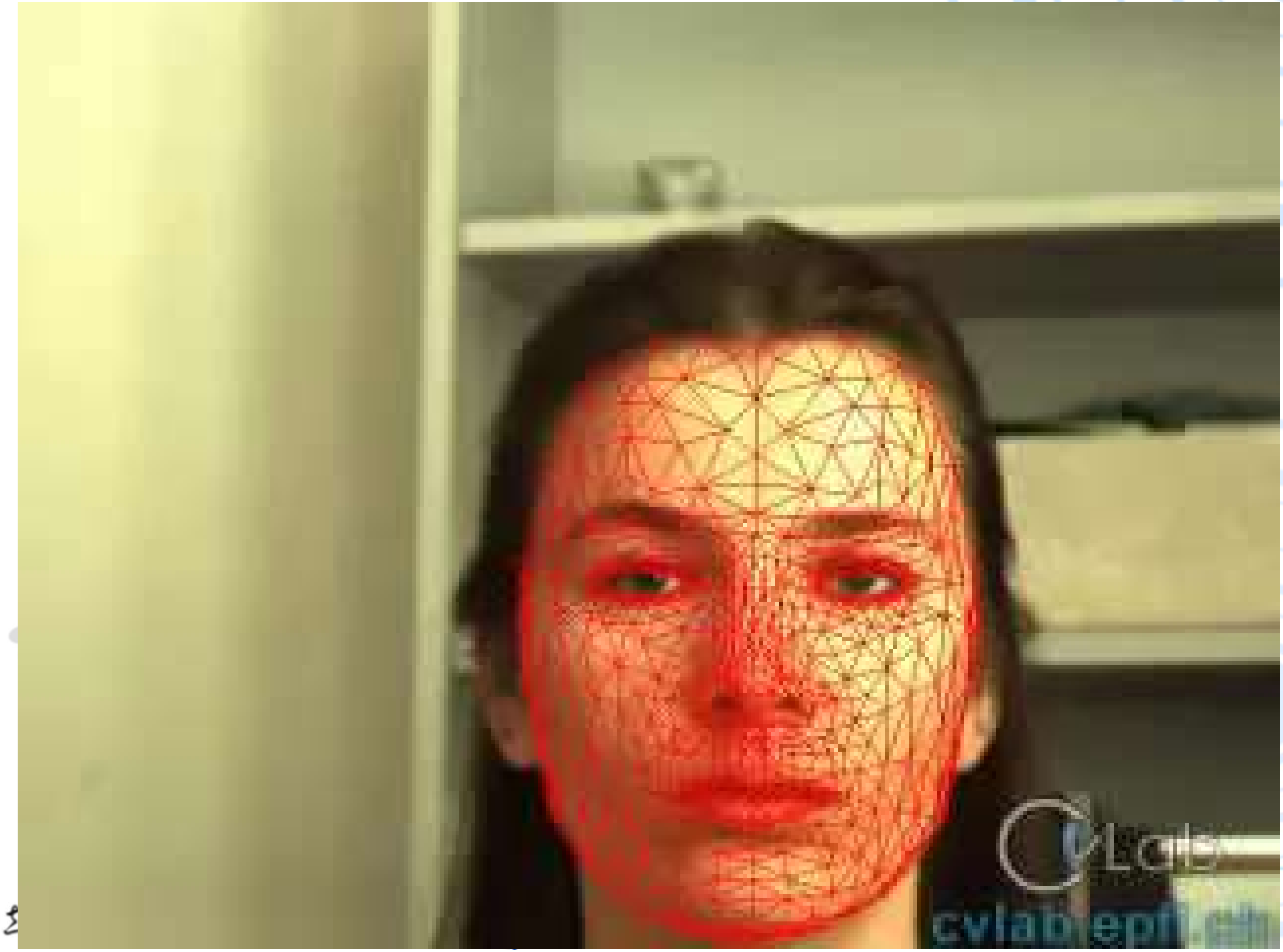
1) Digital Makeover Mirror



2) Our initial experimental result

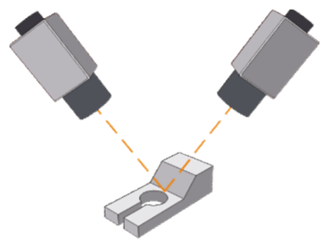
3) Android Virtual Makeup App

2.0.3 Point Tracking and AR: Feature Points and Mesh

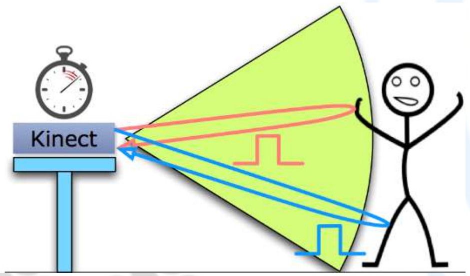


2.0.4 3D Sensors

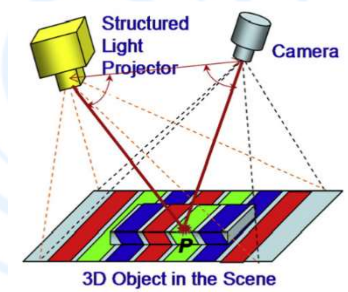
1. 3D Shape Measurement Technology
   1) Stereo Vision:
      - 已知兩個相機間的距離及角度，透過兩個影像對應點的位置計算出待測物的外型輪廓與高度。
   2) Time of Flight (TOF):
      - 對待測物體發射光線，計算該光線到達待測物體表面後反射回來的時間差。
   3) Structured Light:
      - 將結構光投射至待測物體表面，根據相機擷取變形結構光圖像訊息，求出待測物的外型輪廓與高度。



1) Stereo Vision



2) Time of Flight



3) Structured Light

Table. 3D Shape Measurement Technology

Method / Item	1) Stereo Vision	2) Time of Flight	3) Structured Light
a. 設備費用	Middle < NT$80K	Low NT$8K	High > NT$ 150K
b. 精準度: 深度	Low $10\ mm$	Middle $1\ mm$	High $< 1\ mm$
c. 軟體演算法複雜度	Middle	Low	High
d. 工作距離	Far $\geq 5m$	Middle $1m \sim 5m$	Near $\leq 1m$

*: DLP (1920 x 1080) 與相機 (5472 x 3648)不同，校正難度較高

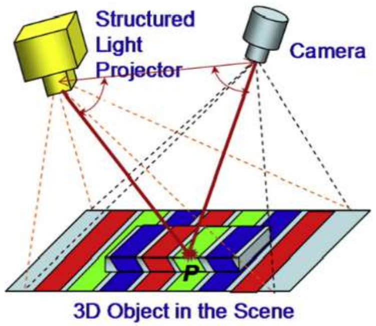
2.0.4 3D Sensors

2. Structured light System Comparison
   1) Single Camera + DLP Projector:[2]

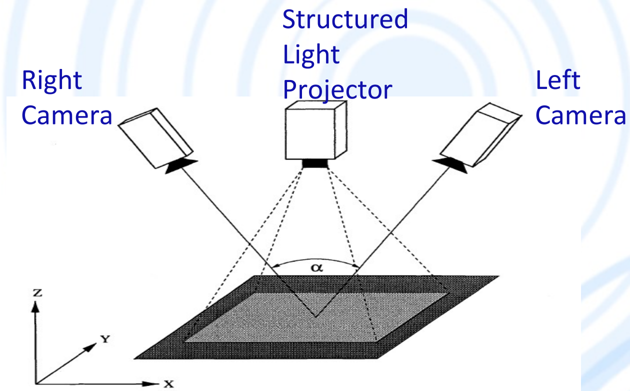
      - 將投影機視為逆相機(inverse camera)進行校正，並由投影機投影的結構光對待測物進行編碼建立相機及投影機間影像對應點關係，計算出待測物的外型輪廓與高度。

   2) Stereo Camera + DLP Projector:

      - 只校正兩顆相機，並由投影機投影的結構光對待測物進行編碼建立相機及相機間影像對應點關係，計算出待測物的外型輪廓與高度。



1) Single Camera + DLP Projector



2) Stereo Camera + DLP Projector

Table. Structured light System Design

Method / Item	1) Single Camera + DLP Projector [2]	2) Stereo Camera + DLP Projector
a. 設備費用	Low > NT$ 40K (DLP) + NT$ 50K (Camera + Lens)	High > NT$ 50K (DLP) + NT$ 70K (Camera * 2 + Lens * 2)
b. 精準度: 深度 (mm)	0.14	0.01
c.1 軟體演算法複雜度 c.2 Calibration or Rectification	Higher DLP Calibration *校正較複雜	High Stereo Rectification
d. 穩定度: 標準差 (mm)	0.0800	0.0039

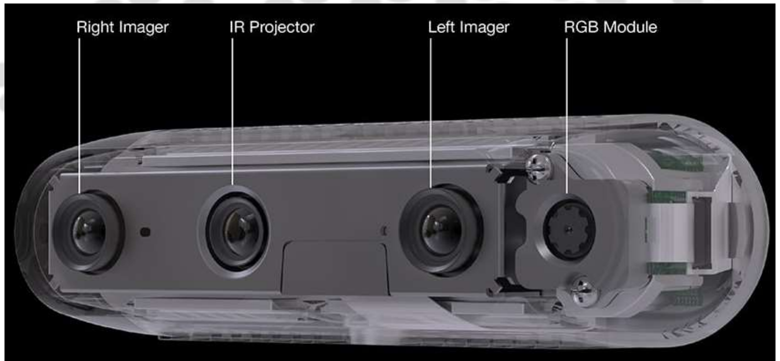
2.0.4 Stereo: RGB-D Camera

3. 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

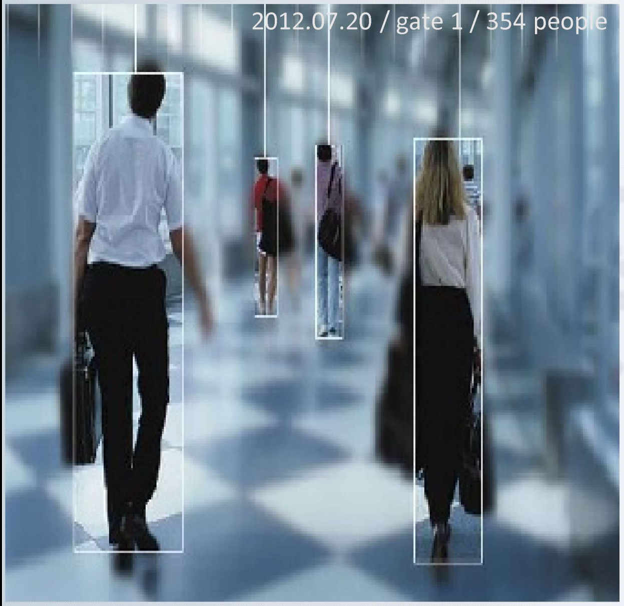


Reference

2.0.4.1 Stereo: Target Applications

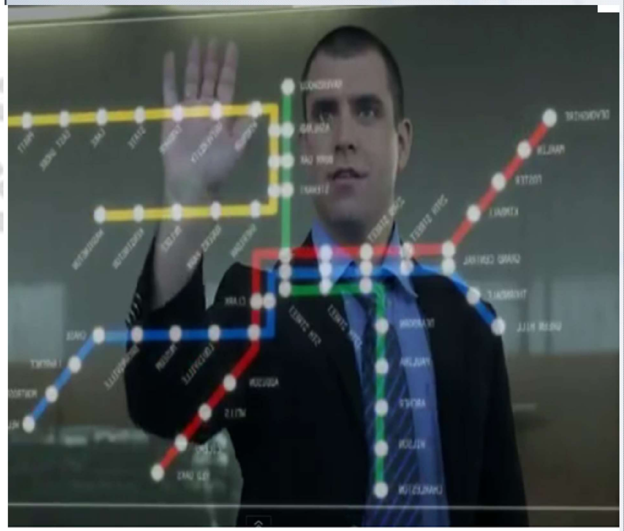
1) Surveillance

3D People Counting



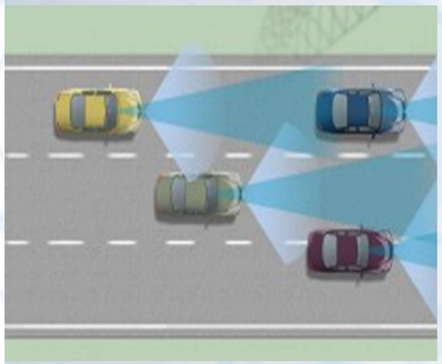
2) Human-Computer Interaction

Hand Gesture Recognition

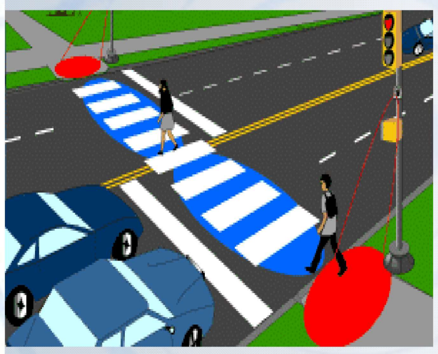


3) Intelligent Transportation System

Forward Collision Warning



Pedestrian Detection

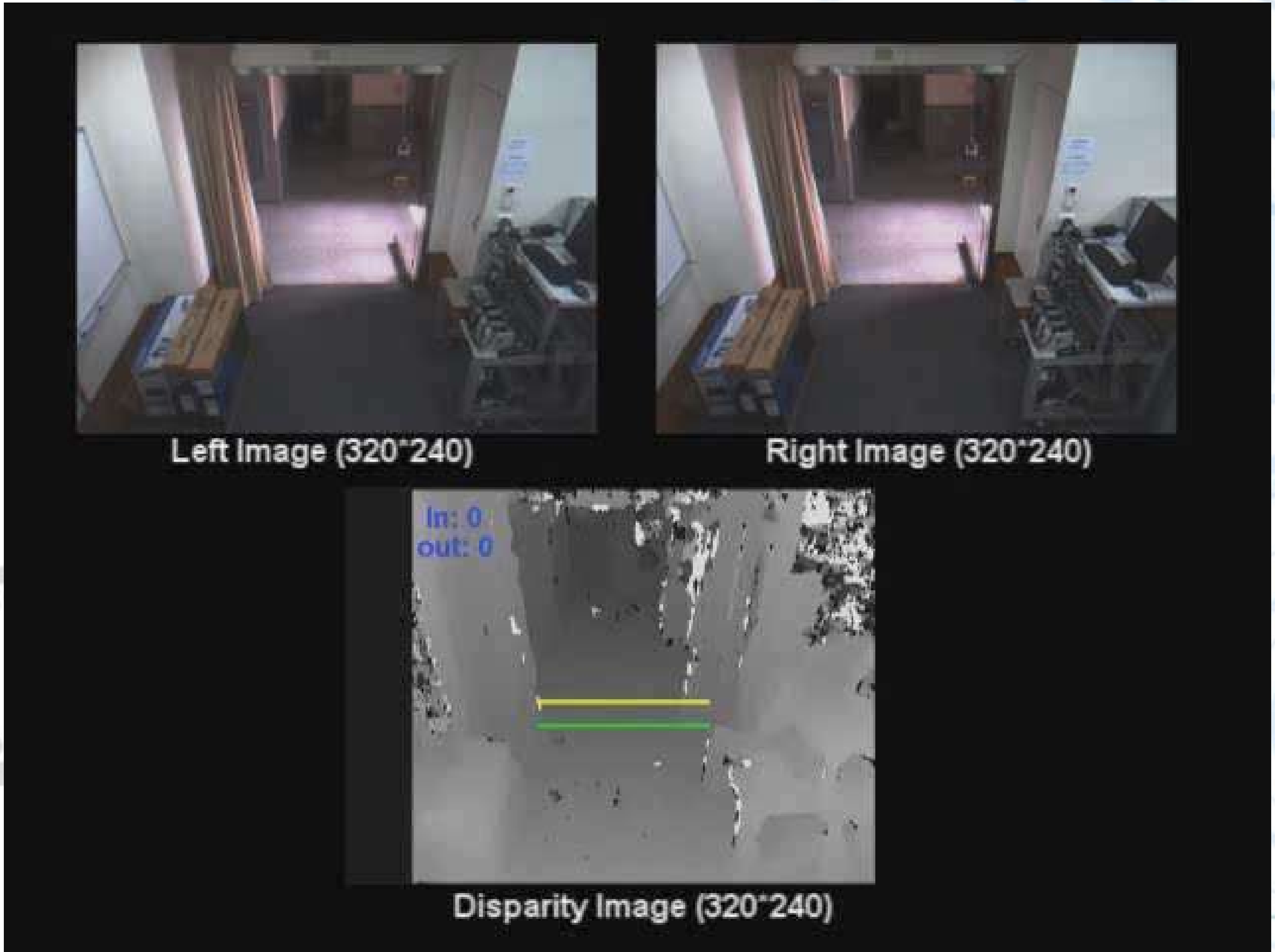


國立成功大學

Tongtai

2.0.4.1 Stereo: Surveillance

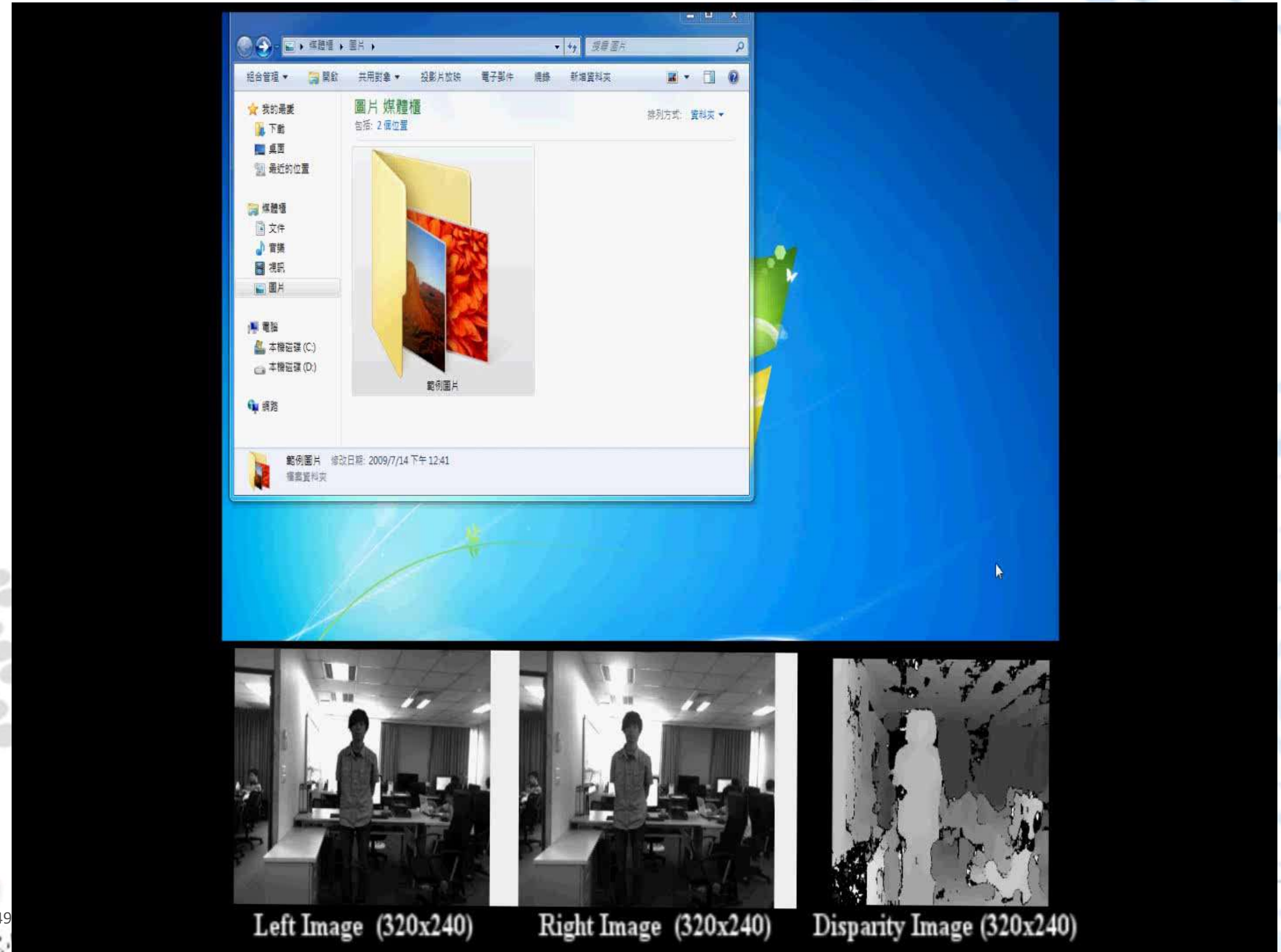
3D People Counting at TI DM8168



國立成功大學 Tongtai

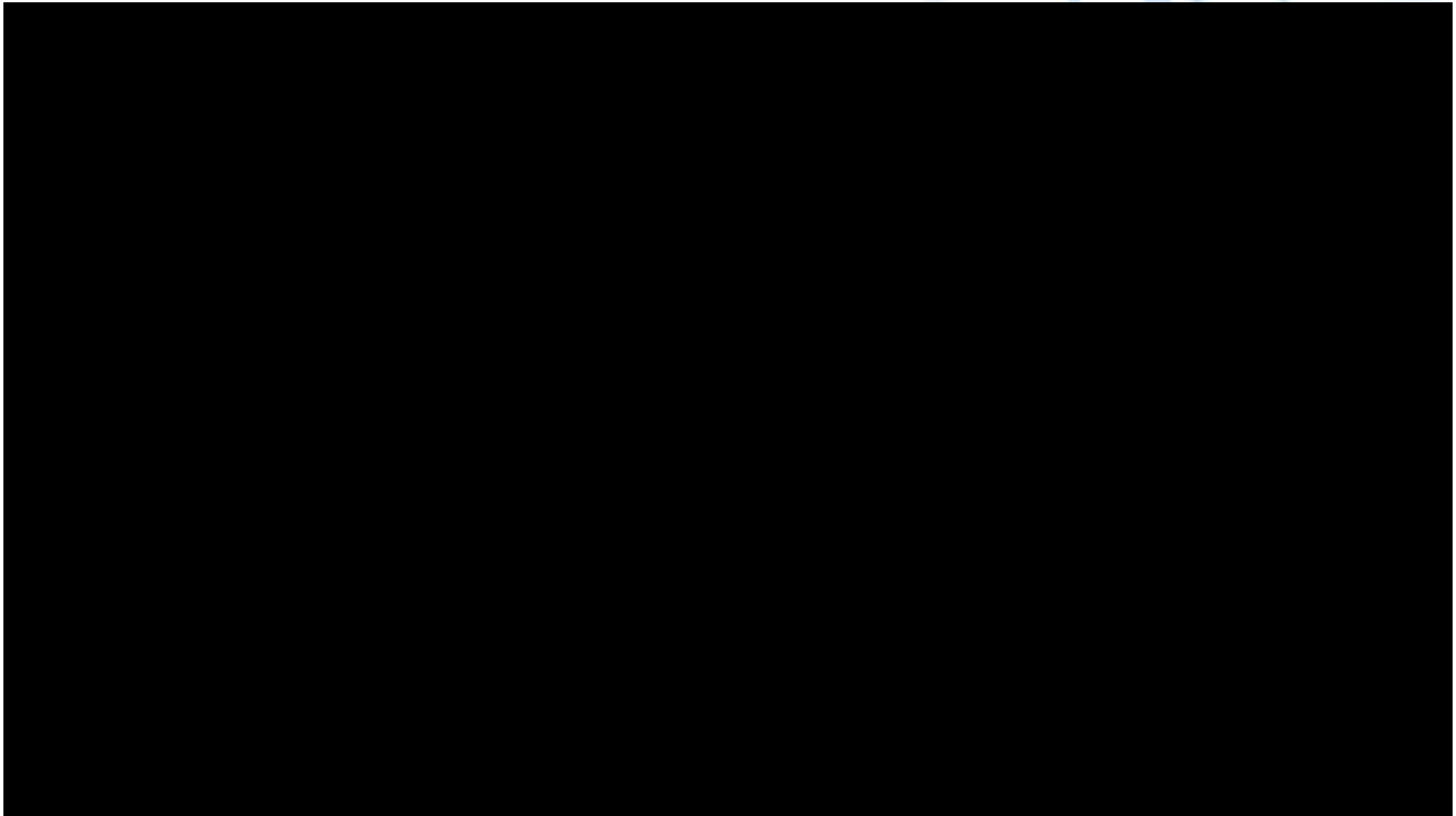
2.0.4.1 Stereo: Human-Computer Interaction

Hand Gesture Recognition at TI DM8168

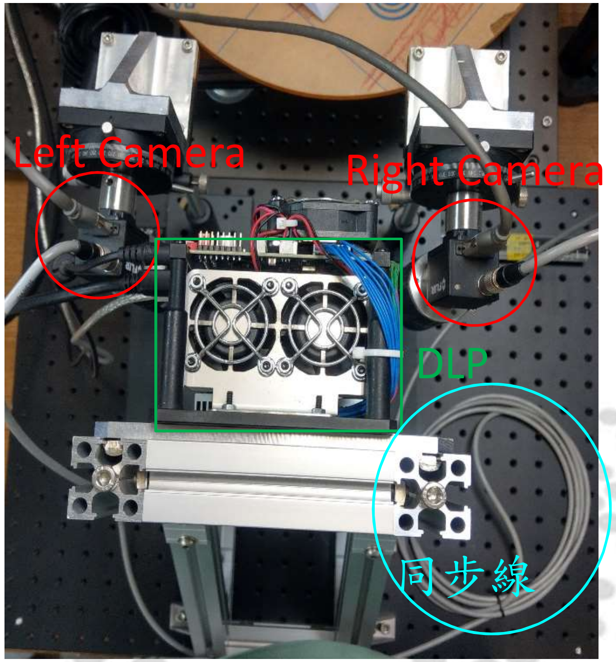


2.0.4.1 Stereo: Embedded Robot Vision and Human-Computer Interaction – MG+4C: 玉創

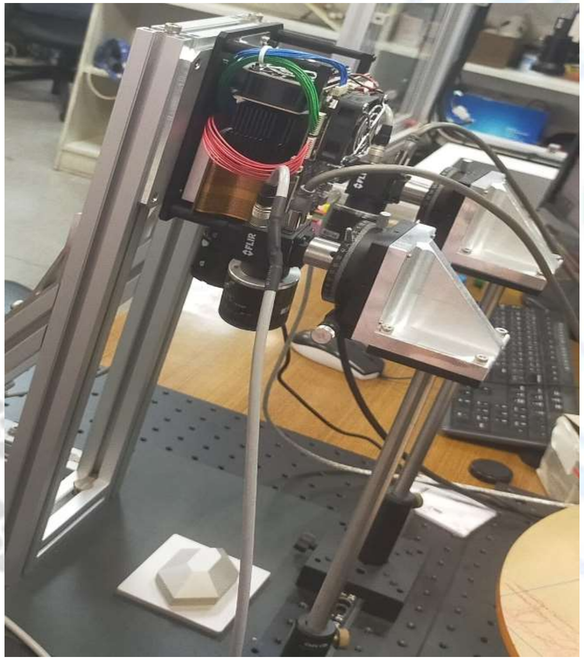
2.0.4.1 Stereo: Embedded Robot Vision and Human-Computer Interaction - Augmented Reality



2.0.4.2 Stereo-Based DLP



Top View



Side View

Bottom View

2.0.4.2 Stereo-Based DLP

3D Object Inspection System Using Deep Learning-Based Structured Light, and System Stability and Speedup

（使用深度學習結構光於3D物體檢測系統與其系統穩定性及加速）

Student : 葉家瑋

Advisor : 郭淑美 教授

Advisor : 連震杰 教授

Robotics Lab, CSIE NCKU
2020/07/24

2.0.4.3 RGB + 2D/3D LiDar

4.1 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)

4.2 2D Lidar: SLAMTEC-RPLiDAR A3

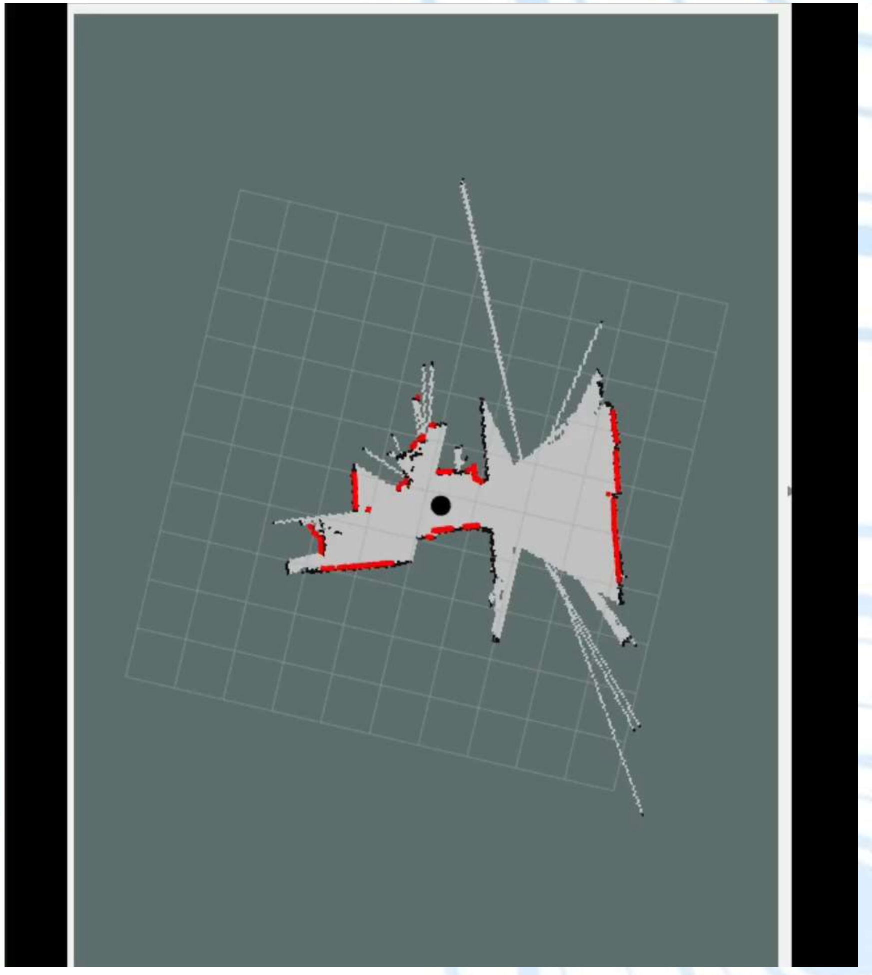
Price: 24,675 NTD

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)

國立成功大學 Tongtai

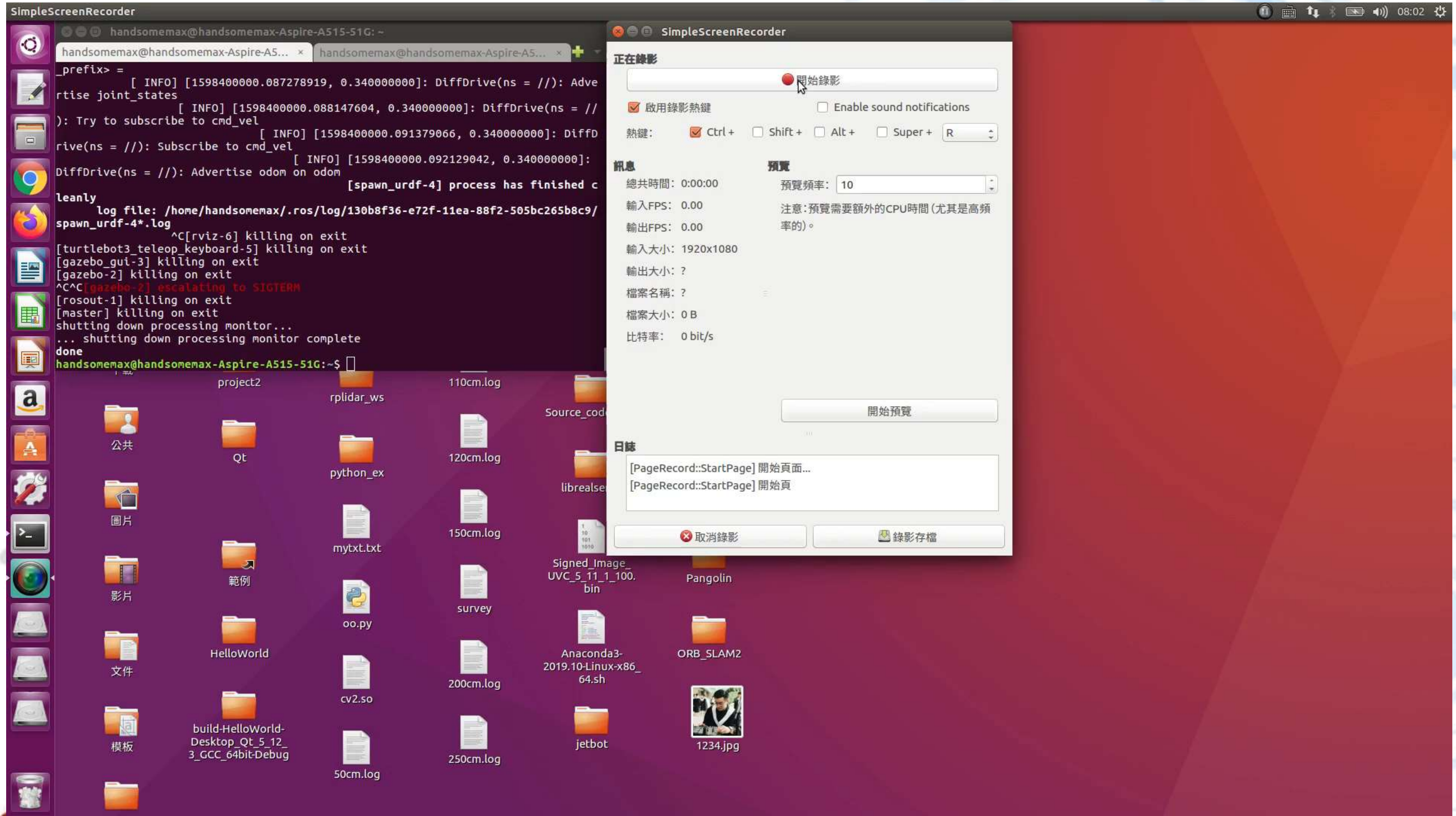
2.0.4.3 Building LiDAR Map: Mapping Demo Video

-



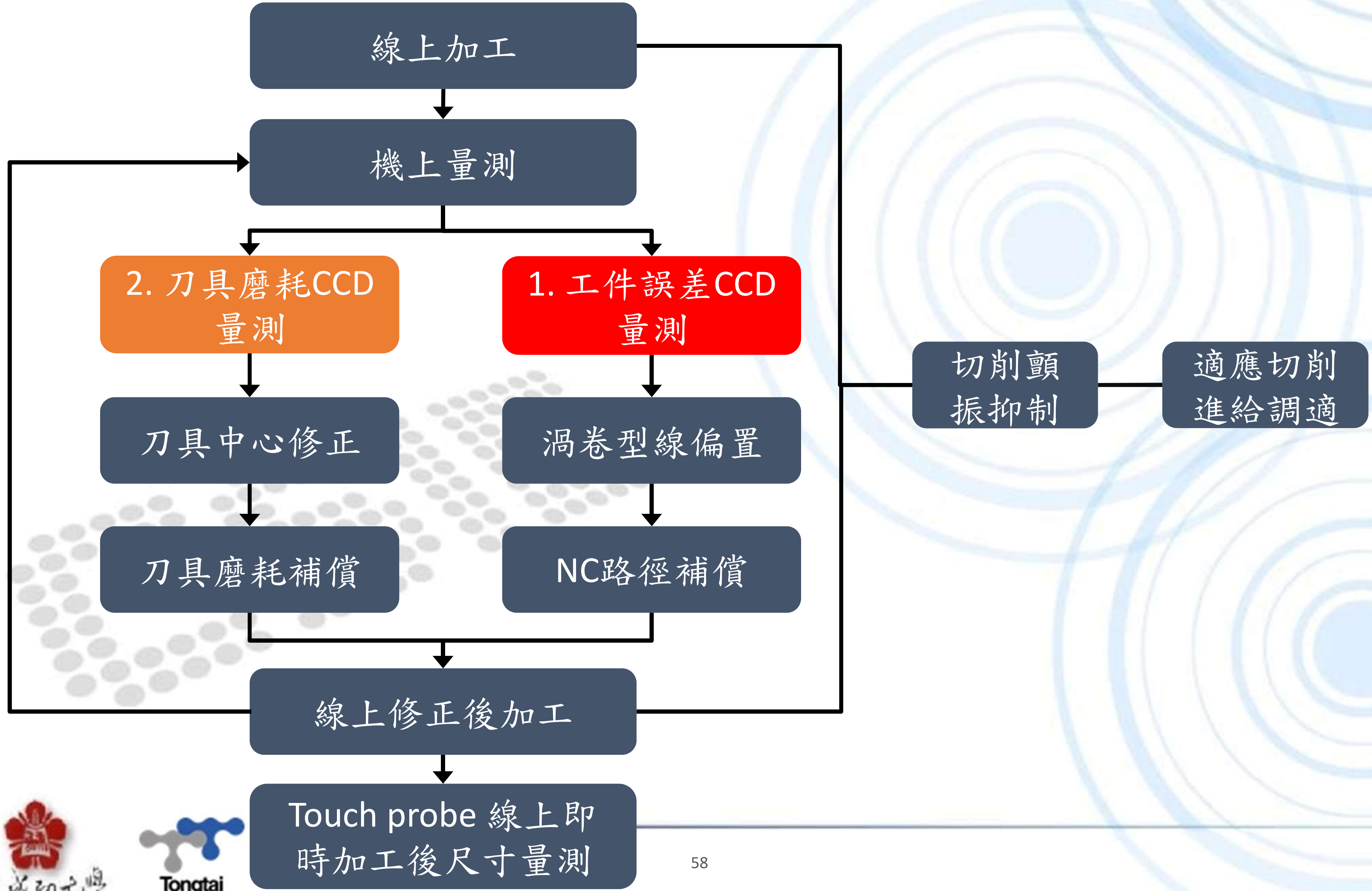
2.0.4.3 ROS Simulation with Gazebo: ROS + 2D Lidar – Demo Video3

- Gazebo: Simulator
- Rviz: GUI offered by ROS



2.0.5 AOI：渦卷曲線智慧化加工及檢測技術開發

2.0.5 AOI: 總計畫 Framework

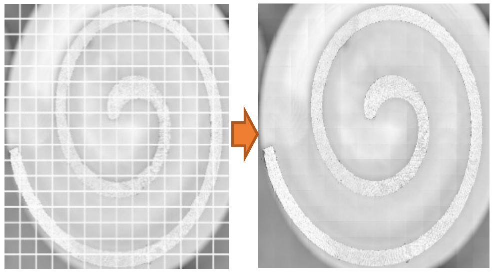


2.0.5 AOI: 子計畫B.2.2: 工件輪廓誤差量測 – Framework

$pixel\ size$ = 1.67

1. Input

I_m

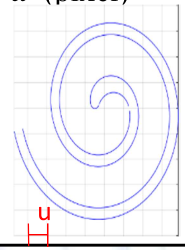


- 影像張數: 238 張 (17 * 14)
- Resolution (h x v): 53760 x 46716 Pixels
- FOV (h x v) : 89.78 x 78.02 mm
- File Size: 2.34 GB

I_m

3. CAD Coordinate Transformation

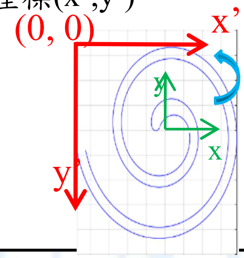
1) 將CAD單位 u (μm)換算為 u' (pixel)



$$u' = \frac{u}{pixel\ size}$$

CAD Image

2) CAD座標(x,y)轉至影像座標(x',y')

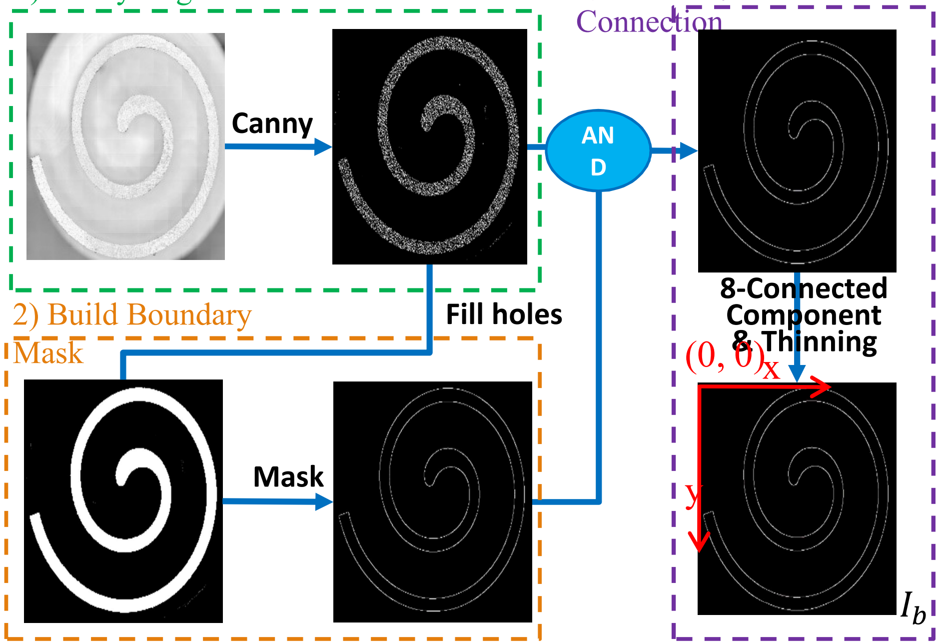


I_c

I_c

(Boundary = Profile)

2. Scroll Boundary Detection

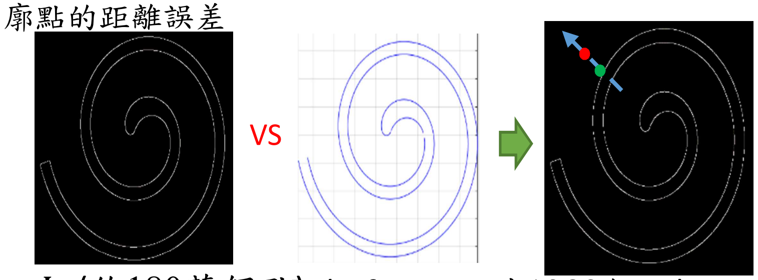


I_b

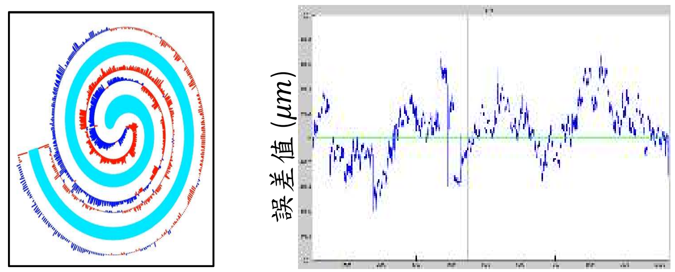
I_b

4. Error Estimation btw CAD and Boundary Image

- 誤差計算: 計算CAD點與沿其法向量上最接近的輪廓點的距離誤差



I_b (約180萬個點) I_c (reference,共1382個點)

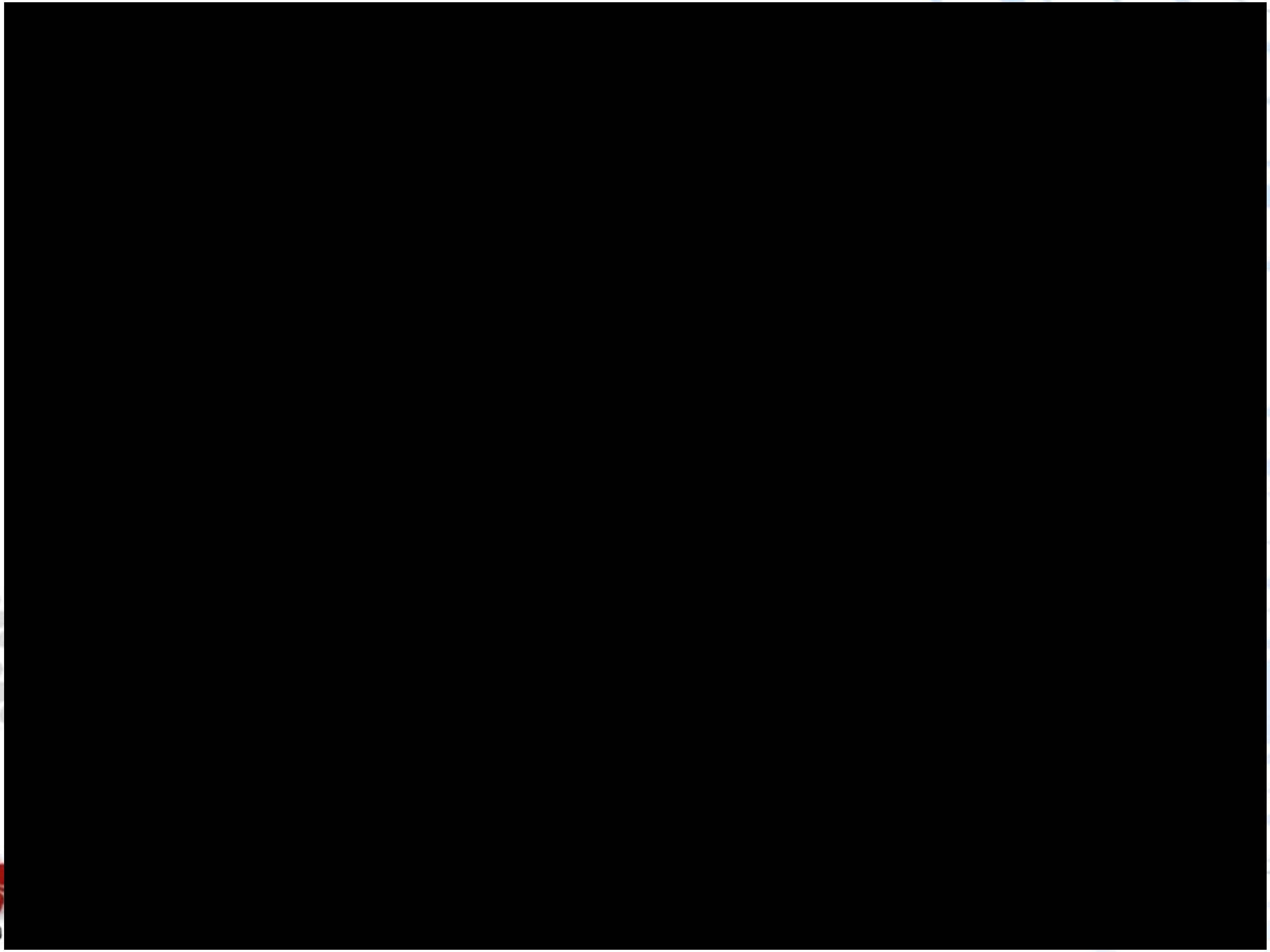


型線誤差圖

誤差分布圖

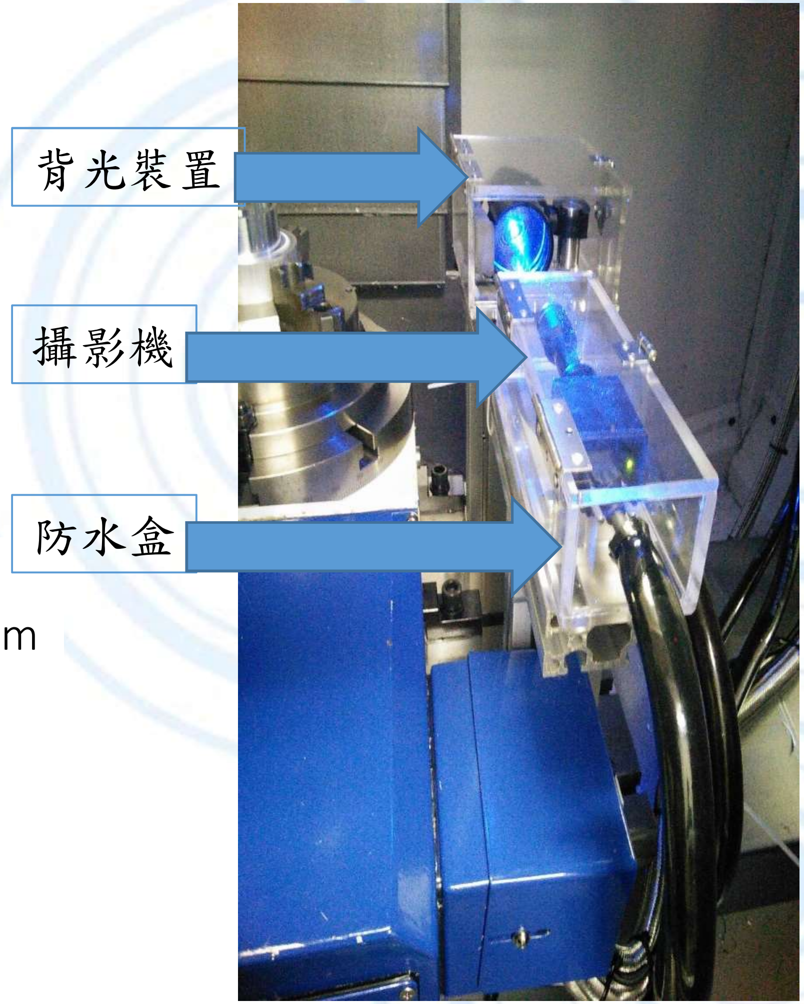
CAD point

2.0.5 AOI: 子計畫B.2.5: 工件輪廓誤差量測 – DEMO (02:41:69)



2.0.6 AOI: 子計畫B.3.1:刀具磨耗檢知 –相機規格與機上架構

1. Camera Spec.	
Resolution (H*V)	2048 * 2048 = 4.2M pixels
Pixel size (H*V)	5.5 * 5.5 µm
2. Lens Spec.	
Magnification	0.8 ± 5%



圖、機上架構 (攝影機+背光

- 影像精度：為 5.5/0.8 = 6.9 (µm/pixel)
- 上機刀具轉速：上機測試，轉速目前為2000 rpm
- 上機量測總花費時間　：約17秒
  (演算法花費時間2+2.8+0.4= 5.2秒)
  1) 移動機台至拍攝位置：約6秒
  2) 刀具取像90張圖　：約2秒 (平均一張0.022秒)
  3) 顯示90張圖　：約3.1秒
  4) 疊圖時間　：約2.8秒
  5) 量測數據　：約0.4秒
- 適合量測轉速：60~6000 rpm

2.0.6 AOI: 子計畫B.3.5:刀具磨耗檢知 (01:47:69)

2.0.6 AOI:刀具磨耗檢知 - Stitching ()

1. Feature Extraction Using SIFT → 2.1 Feature Matching: Initial Feature Matching → 2.2 Feature Matching: Good matches → 3. Panoramic Image Stitching → 4. Result of Panoramic Image Stitching

.......共90張

2.0.6 Multi-Task: Application – Tool Defect Detection

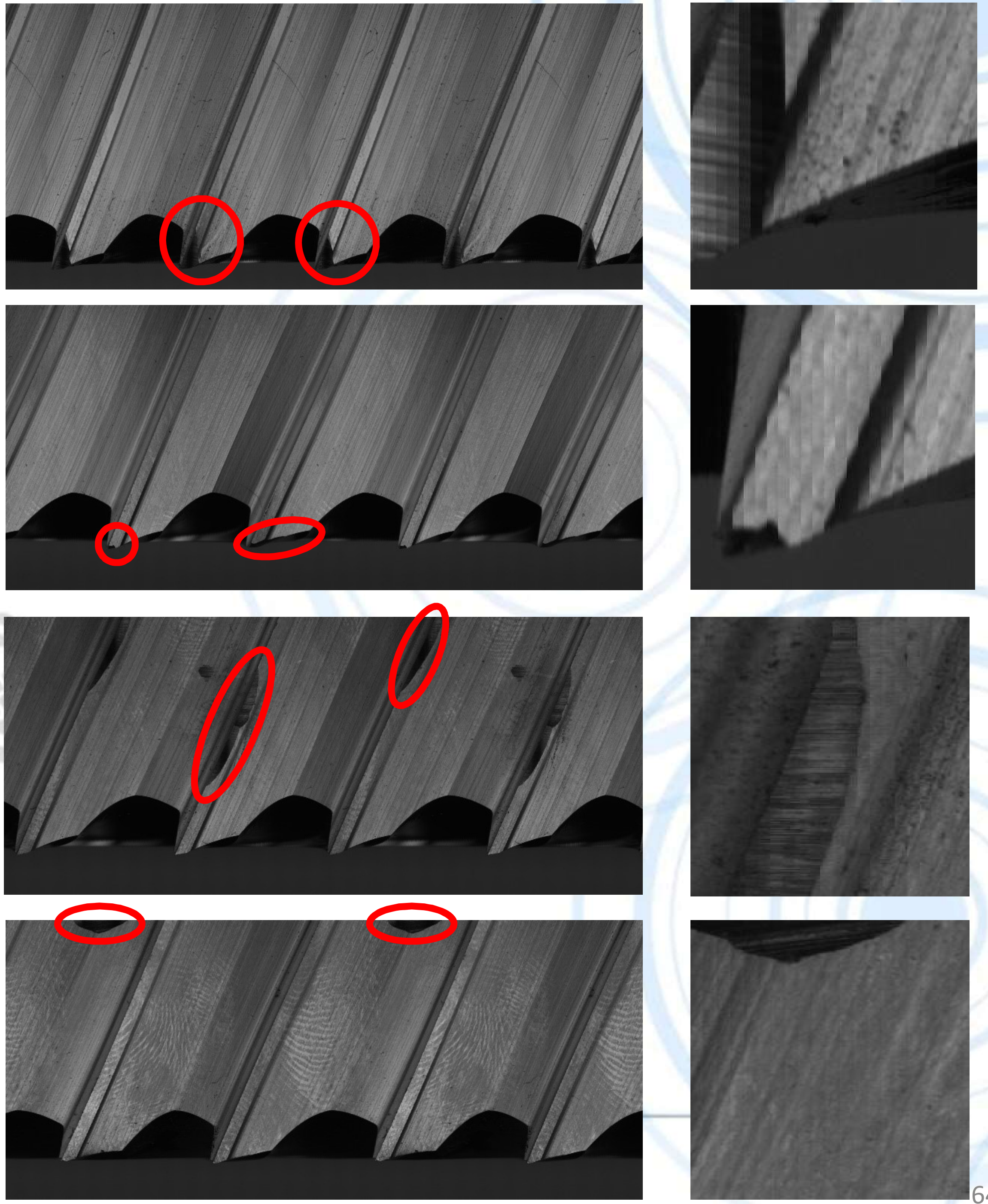
➢ 與中正/台大蔡孟勳教授及東台精機合作

Defect

1) Tip

2) Flank

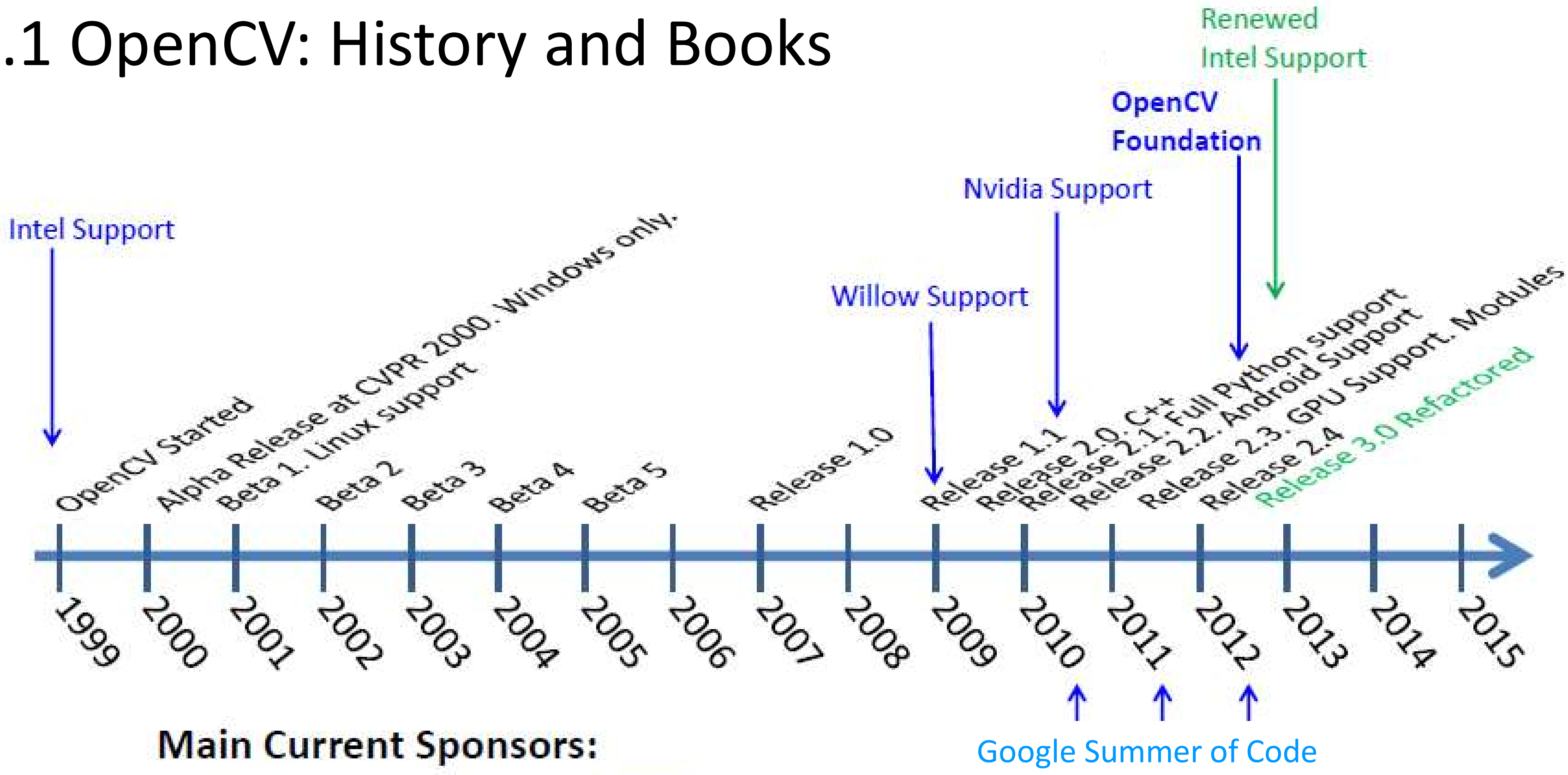
3) Plane



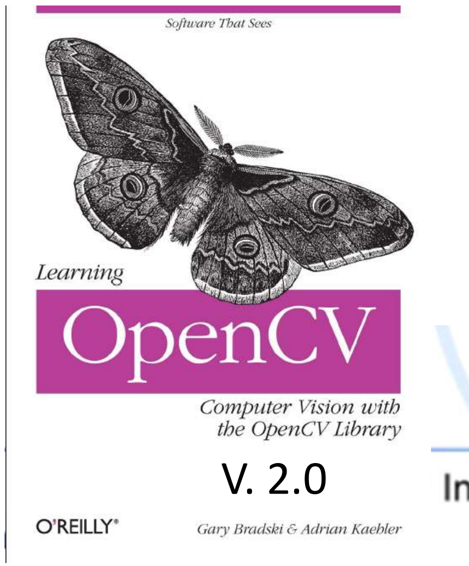
- Image Resolution: 2054 pixels * 4581 pixels
- Patch size: 224 pixels * 224 pixels

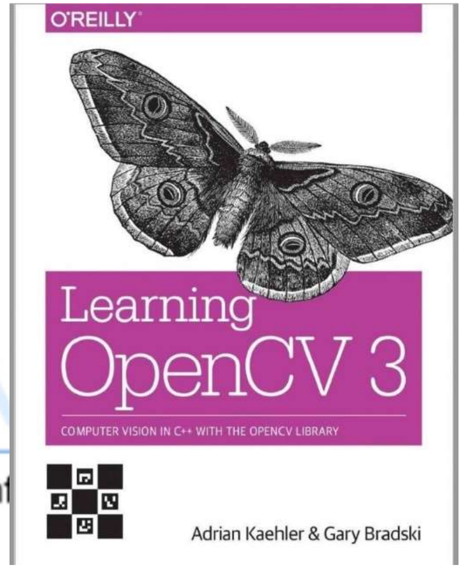
2.0.6 AOI:刀具磨耗預測 – 深度學習 ()

2.1 OpenCV: History and Books



Main Current Sponsors:





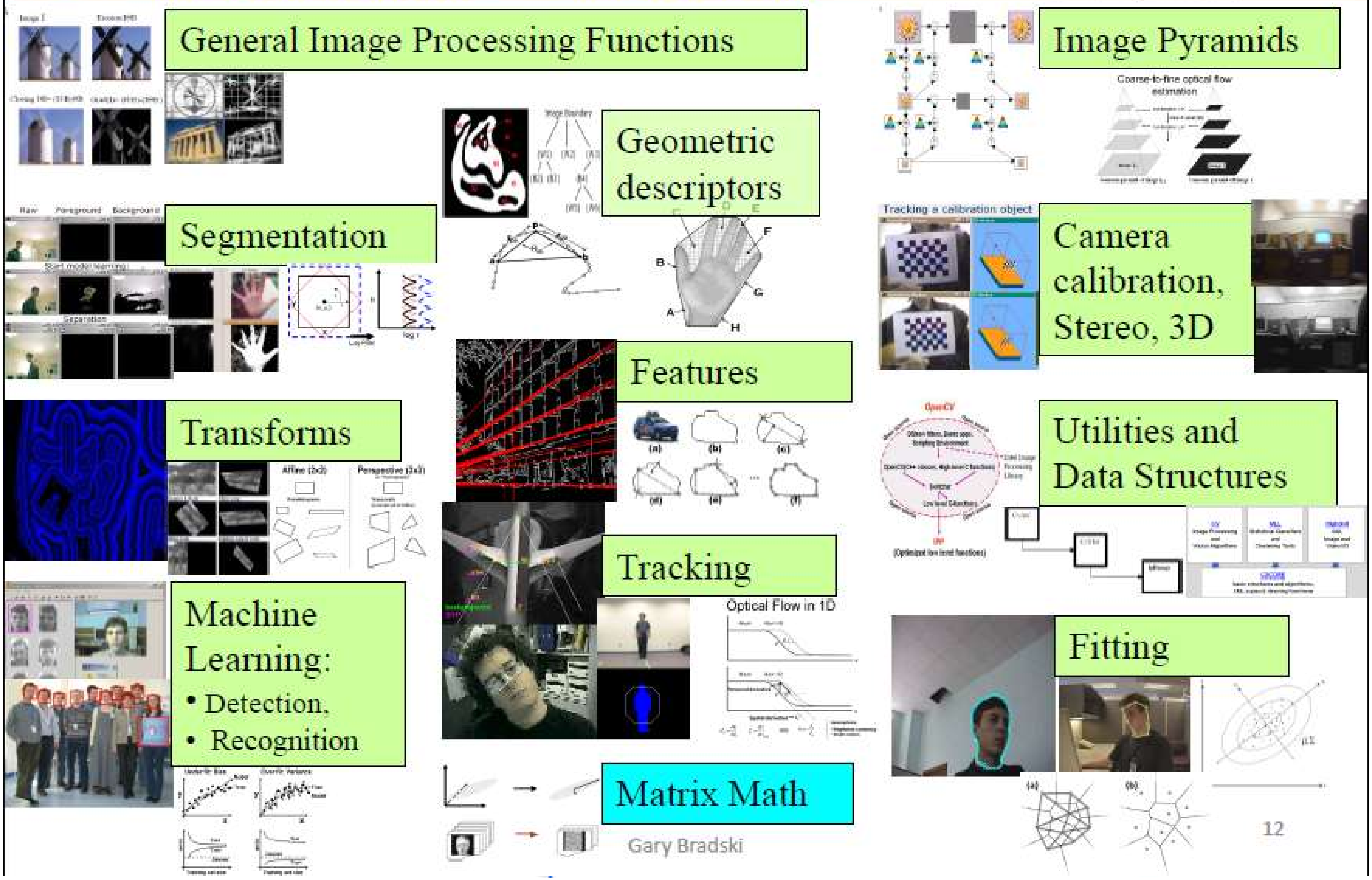
OpenCV Books V2 in 2008 and V3 in 2016
by Adrian Kaehler & Gary Brdski

2.2 OpenCV Functions:

OpenCV

Developer http://code.opencv.org; User: http://opencv.org

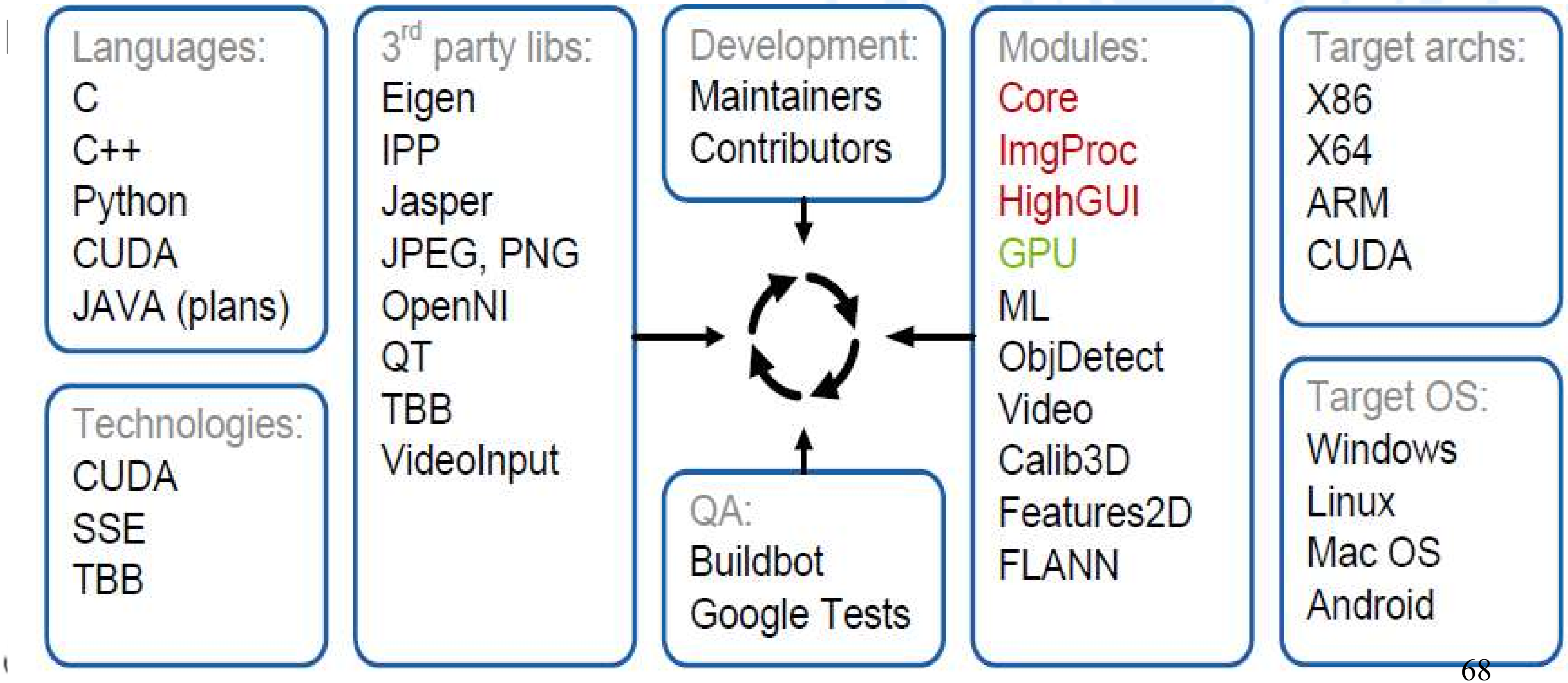
> 2500 algorithms



2.3 OpenCV Environments & Platforms, and Architecture & Development

❑ Environment & Platforms

- Languages:
  - C, C++, C#, Java, Python, Matlab
- Platform:
  - GPU/Cuda, Android, iOS,



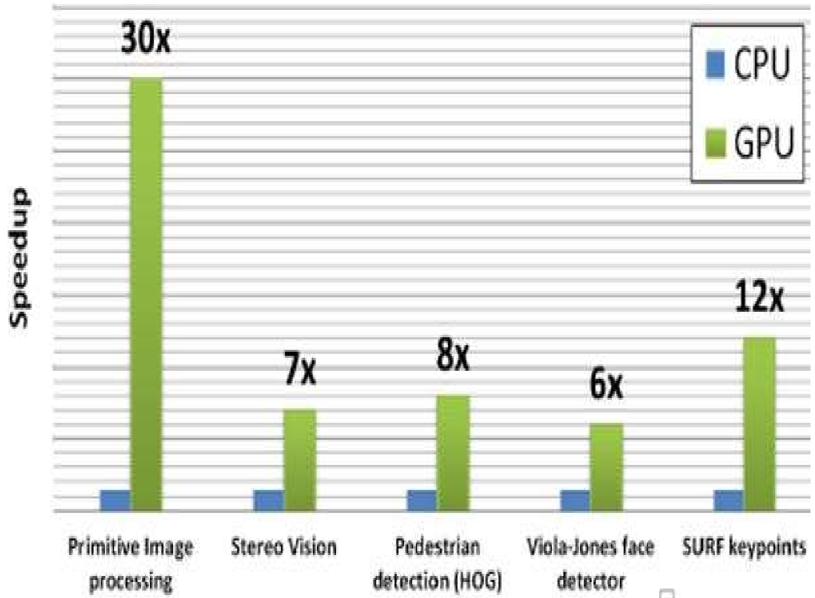
2.4 OpenCV GPU Module Performance

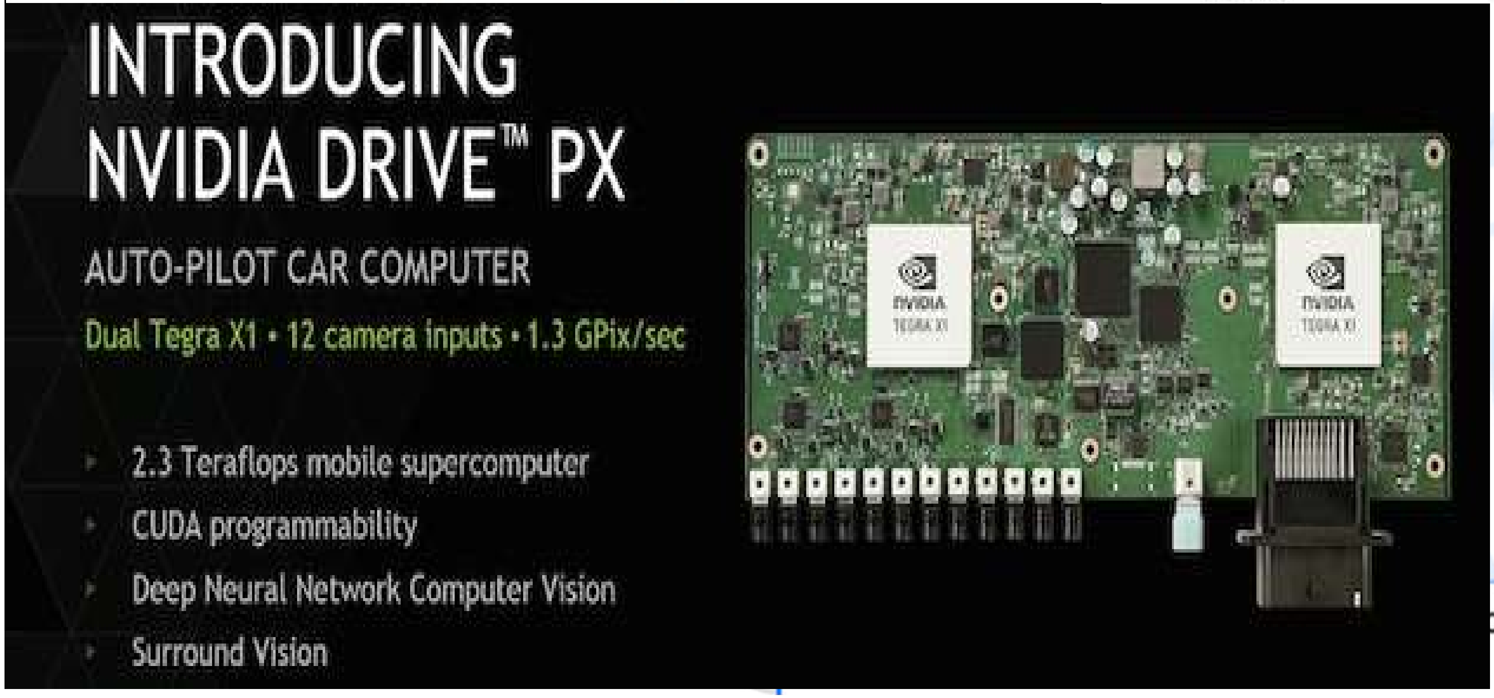
Tesla C2050 (Fermi) vs. Core i5-760 2.8GHz (4 cores, TBB, SSE)

- Average speedup for primitives: 33×
  - For "good" data (large images are better)
  - Without copying to GPU

What can you get from your computer?

- opencv\samples\gpu\perfomance





2.4 OpenCV GPU Module Performance: Embedded Module

1. OpenVX:
2. OpenMP (Open Multi-Processing: Parallel Processing):
   - ➢ Multi-Cores CPUs, GPUs or DSPs
3. OpenCL (Open Computing Language):
   - ➢ Cross CPU + GPU + DSP + FPGA

- khronos 發表電腦視覺API 標準：**OpenVX**. 在2011 年的時候，Khronos 曾經發表過一個名為「Vision」的API 標準，希望可以為電腦視覺（Computer Vision）的處理、定義一套標準的介面，作為硬體加速的抽象層；當時，基本上只是剛開始的階段，並沒有完整的介面出來。2014年10月21日

- Khronos group - OpenVX:
  - Connecting software to silicon
  - OpenVX is an open, royalty-free standard for cross platform acceleration of computer vision applications.
  - It is designed by the Khronos Group to facilitate portable, optimized and power-efficient processing of methods for vision algorithms.
  - This is aimed for embedded and real-time programs within computer vision and related scenarios. It uses a connected graphics representation of operations.

OpenVX (Khronos HAL)