



Multimedia Big Data Analytics in AI Era 人工智慧時代的多媒體大數據分析

Wen-Huang Cheng (鄭文皇)

Associate Professor / Associate Research Fellow

Multimedia Computing Lab (MCLab),
Research Center for Information Technology Innovation (CITI),
Academia Sinica, Taipei, Taiwan
whcheng@citi.sinica.edu.tw

關於我

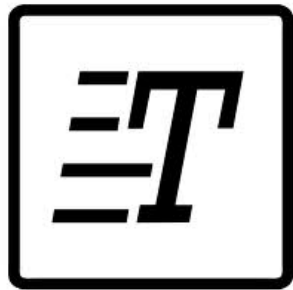


- 現職
- 副研究員，中央研究院資訊科技創新研究中心
 - 合聘副研究員，中央研究院資訊科學研究所
 - 兼任副教授，國立台灣科技大學資訊工程系
 - 監事，臺灣金融科技協會
 - 理事，臺灣人機互動學會
 - 技術顧問，資訊工業策進會
- 經歷
- 主任工程師，HTC宏達電
 - 技術顧問，歐勝科技
- 學歷
- 博士，國立台灣大學資訊網路與多媒體研究所
 - 碩士，國立台灣大學資訊工程學研究所
 - 學士，國立台灣大學資訊工程學系



What is Multimedia?

- Multimedia refers to content that uses a combination of different content forms (**multimodality**), such as text, audio, still image, animation, video, and interactivity.



text



audio



still image



animation



video



interactivity

Multimedia: “the biggest big data”

- Multimedia makes up 60% of internet traffic
- 70% of mobile phone traffic and 70% of all available unstructured data
- Web users are uploading 72 video-hours to YouTube per minute
- On an average day, social media users are posting 300 hundred million photos to Facebook.
- Consumers using mobile phones and digital cameras are taking 500B photos per year, which is 78 per person on the planet.
- Medical institutions are acquiring 1B radiology images per year.

(John R. Smith @ SIGIR'13)

By 2018, 84% of Internet traffic will be videos!



Multimedia Research?



ACM Multimedia Conference



ACM MULTIMEDIA 2015

26 - 30 October 2015, Brisbane, Australia

[Home](#)

[Call for Contributions](#)

[Authors](#)

[Organisation](#)

[The Venue](#)

[Registration & Visa Information](#)



Research Center for Information Technology Innovation

2015 ACM Multimedia:

15 topic areas grouped into 4 themes

- Theme 1: Systems
 - Def: Research in multimedia systems is generally concerned with **understanding fundamental tradeoffs between competing resource requirements**, developing practical techniques and heuristics for realizing complex optimization and allocation strategies, and demonstrating innovative mechanisms and frameworks **for building large-scale multimedia applications**.
 - Mobile Multimedia
 - Multimedia Systems and Middleware
 - Multimedia Telepresence and Virtual/Augmented Reality
 - Multimedia Transport and Delivery



2015 ACM Multimedia:

15 topic areas grouped into 4 themes

- Theme 2: Experience
 - Def: One of the core researches is that multimedia data contributes to the user experience in a rich and meaningful manner. The topics organized under this theme are concerned **with innovative uses of multimedia to enhance the user experience**, how this experience is manifested in specific domains, and metrics for qualitatively and quantitatively measuring that experience in useful and meaningful ways.
 - Multimedia for Collaboration in Education & Distributed Environments
 - Multimedia HCI and Quality of Experience
 - Music, Speech and Audio Processing in Multimedia
 - Multimedia Authoring and Enrichment



2015 ACM Multimedia:

15 topic areas grouped into 4 themes

- Theme 3: Engagement
 - Def: The engagement of multimedia with society as whole requires research that addresses **how multimedia can be used to connect people with multimedia artifacts that meet their needs in a variety of contexts.**
 - Multimedia Art, Entertainment and Culture
 - Multimedia Search and Recommendation
 - Social Multimedia
 - Emotional and Social Signals in Multimedia



2015 ACM Multimedia:

15 topic areas grouped into 4 themes

- Theme 4: Understanding
 - Def: Multimedia data types by their very nature are complex and often involve intertwined instances of different kinds of information. We can **leverage this multi-modal perspective in order to extract meaning and understanding of the world, often with surprising results.**
 - Deep Learning for Multimedia
 - Multimodal Analysis and Description
 - Multimedia and Vision



Evolution of the Topics in ACM MM

2013

- Art, Entertainment, and Culture
- Authoring and Collaboration
- Crowdsourcing
- Media Transport and Delivery
- Mobile & Multi-device
- Multimedia Analysis
- Multimedia HCI
- Music & Audio
- Search, Browsing, and Discovery
- Security and Forensics
- Social Media & Presence
- Systems and Middleware

2014

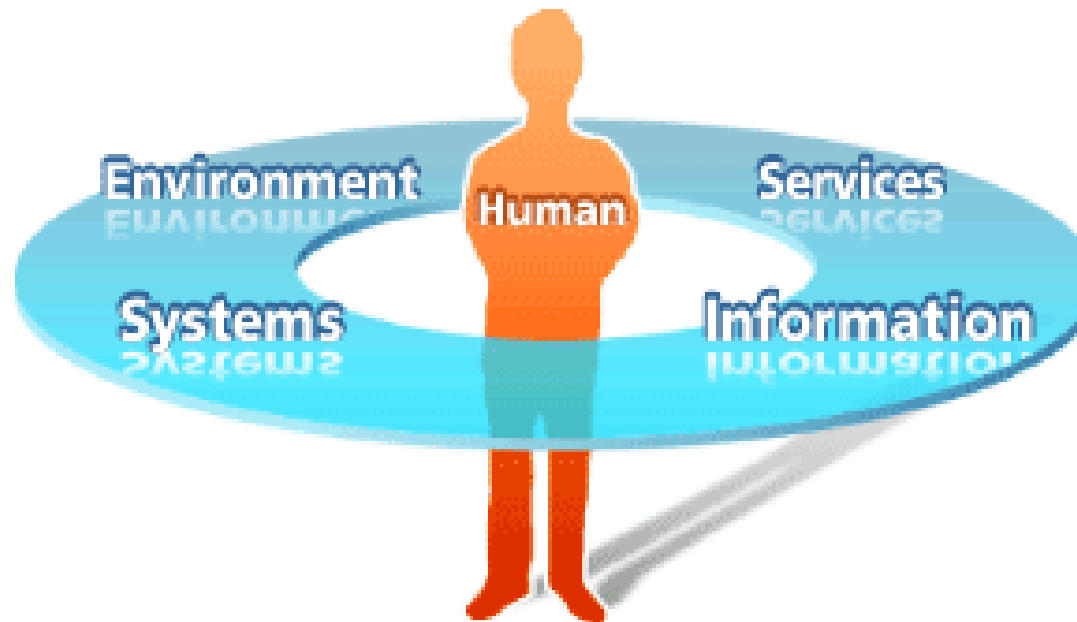
- Media Transport and Delivery
- Multimedia Systems and Middleware
- Big and Broad Multimedia
- Deep Learning for Multimedia
- Emotional and Social Signals in Multimedia
- Multimedia and Society
- Multimedia and the Crowd
- Multimedia Art, Entertainment and Culture
- Multimedia HCI and QoE
- Multimedia Search and Recommendation
- Multimedia Security, Privacy and Forensics
- Multimodal Analysis and Description
- Music, Speech and Audio Processing in Multimedia
- Social Media and Collective Online Presence

2015

- Mobile Multimedia
- Multimedia Systems and Middleware
- Multimedia Telepresence and Virtual/Augmented Reality
- Multimedia Transport and Delivery
- Multimedia for Collaboration in Education & Distributed Environments
- Multimedia HCI and Quality of Experience
- Music, Speech and Audio Processing in Multimedia
- Multimedia Authoring and Enrichment
- Multimedia Art, Entertainment and Culture
- Multimedia Search and Recommendation
- Social Multimedia
- Emotional and Social Signals in Multimedia
- Deep Learning for Multimedia
- Multimodal Analysis and Description
- Multimedia and Vision



Human Centered Design

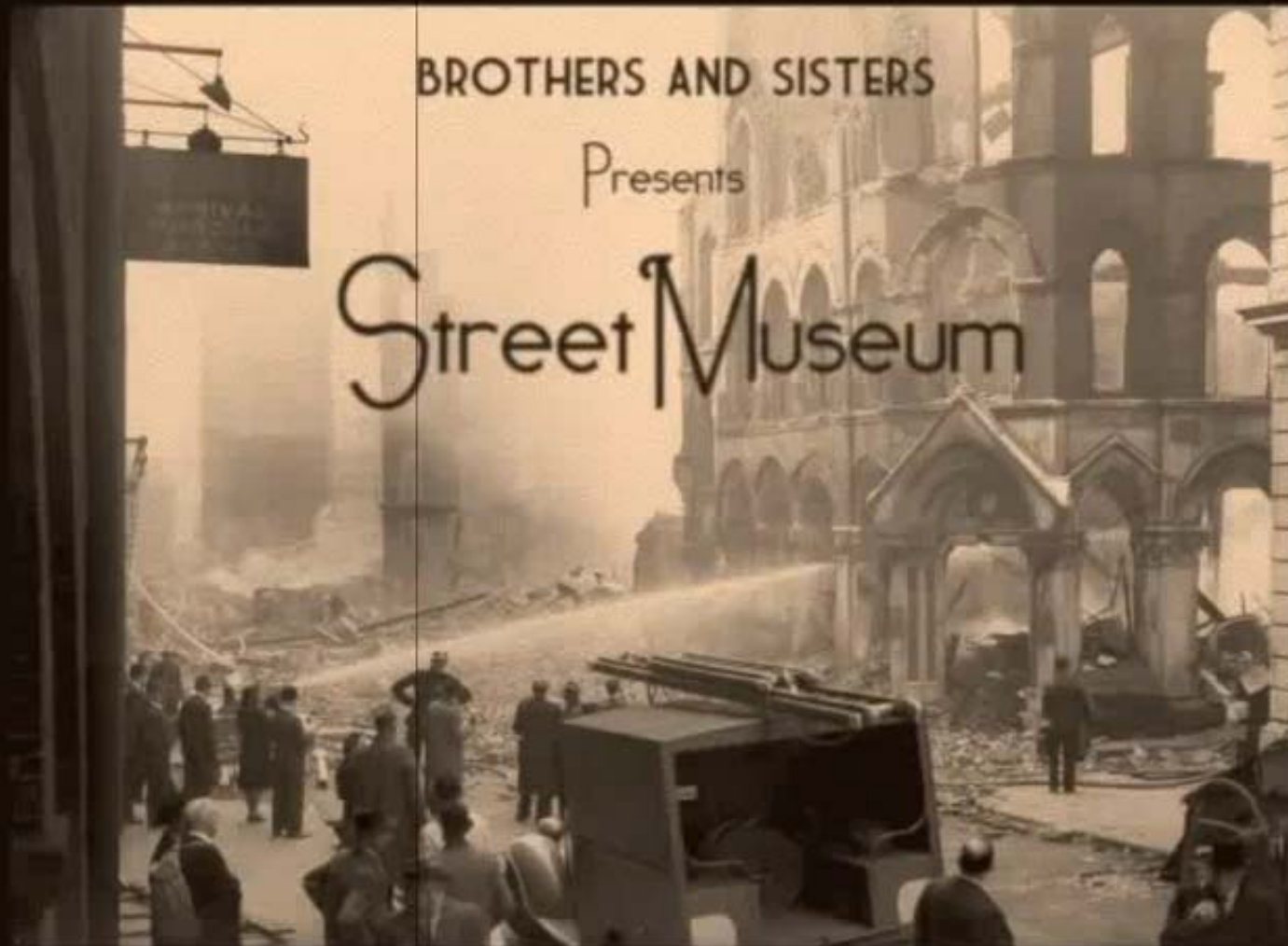




BROTHERS AND SISTERS

Presents

Street Museum



youku



Intelligent Multimedia Computing



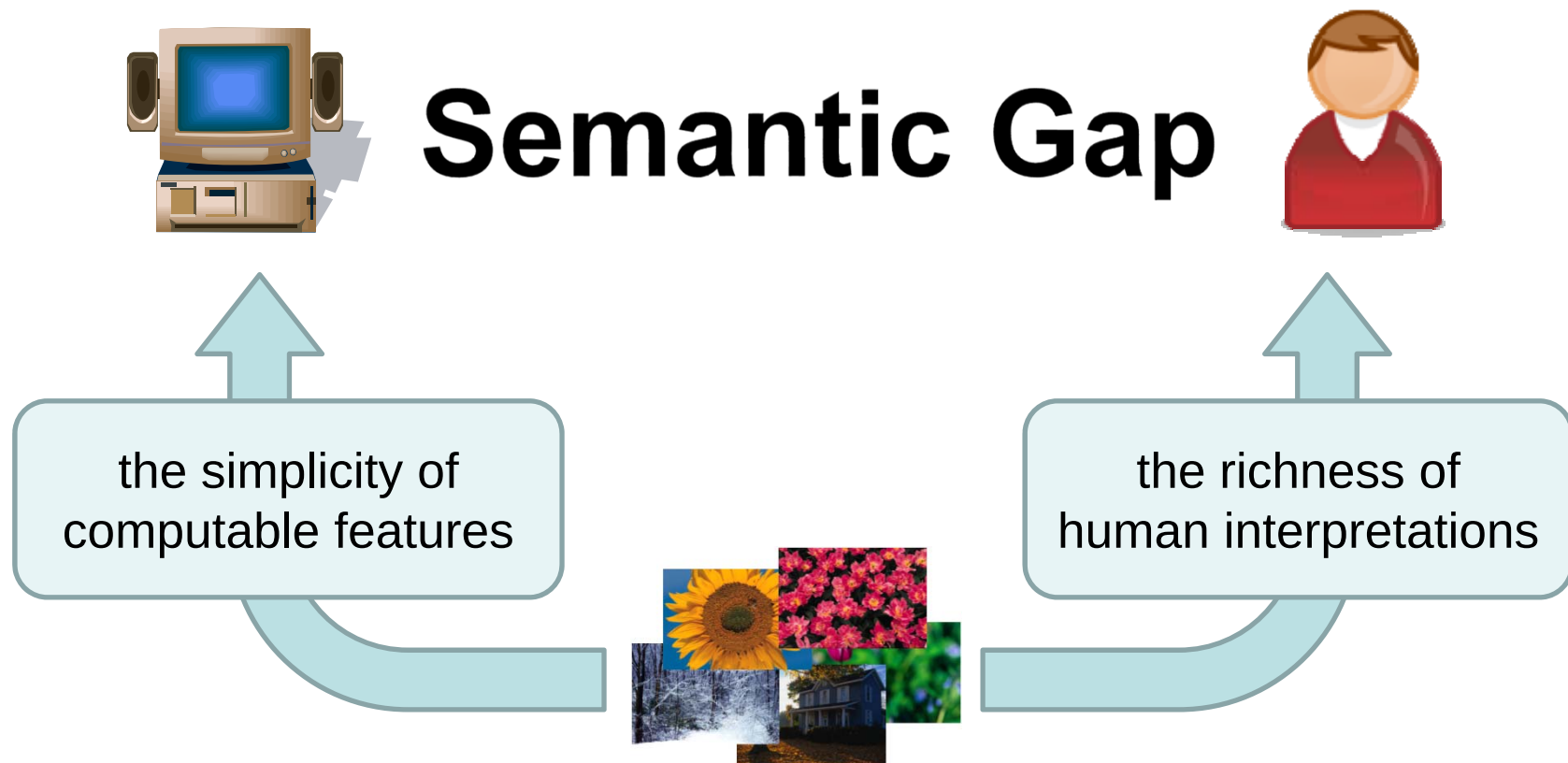
A Classical Painting of the Northern Sung



“Listening to the Qin” (partial) by Ji Zhao, 1082-1135.

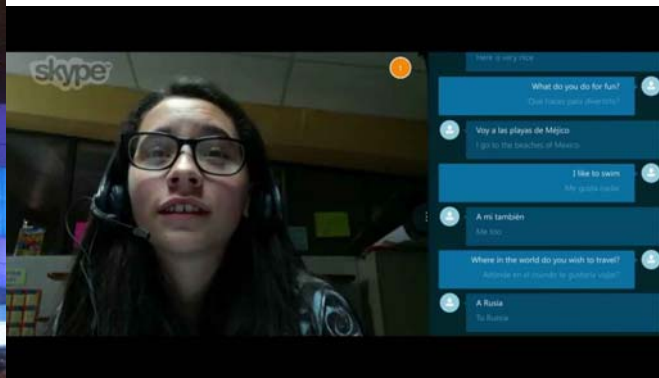
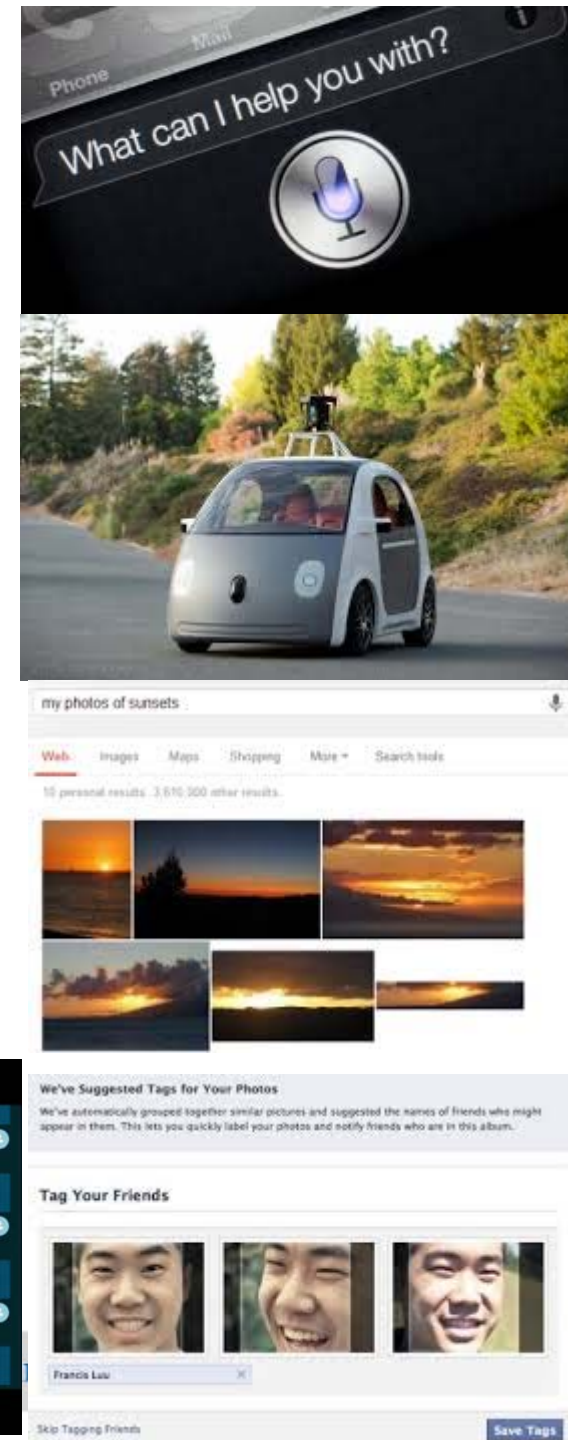


A Big Challenge in Multimedia Research



“Smart” computers behave like humans

- Recent intelligent products in the market
 - Google : Self-driving car, Google Glasses
 - Facebook : face recognition, user modelling
 - Apple : virtual assistant Siri
 - Microsoft : Skype translator
 - Yahoo : image recognition for Flickr
 - IBM's Watson : Jeopardy contest
 - Baidu : speech recognition system



Artificial Intelligence

Intelligence



Computer

versus

Human



Research Center for Information Technology Innovation

Academia Sinica (中央研究院)

- The highest national research institute in Taiwan
 - with about 1,000 researchers/professors (50 in EE/CS)

located in Nangang, Taipei





Multimedia Computing Lab (MCLab)



Welcome to the Multimedia Computing Laboratory (MCLab)!



[A view around the MCLab in Academia Sinica]



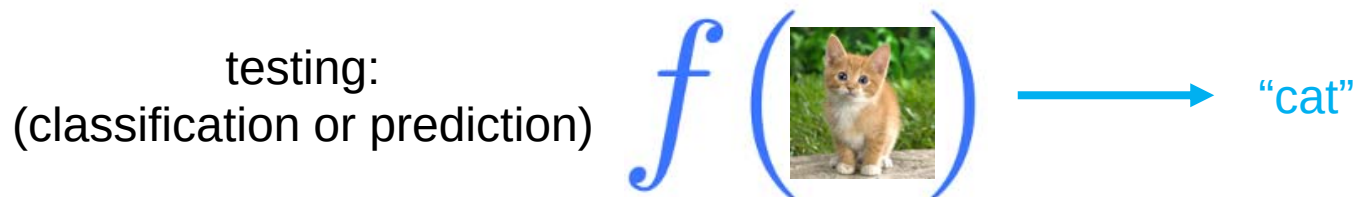
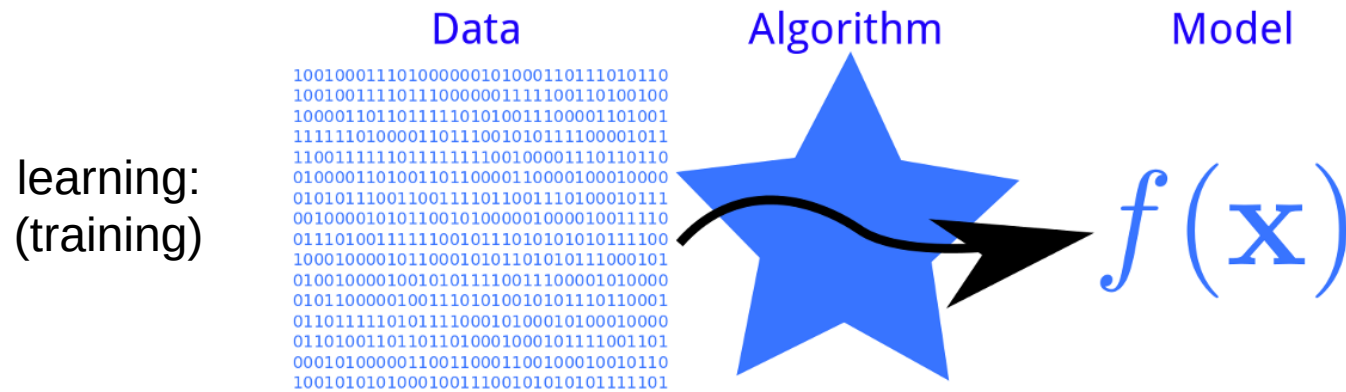
<http://mclab.citi.sinica.edu.tw>



Research Center for Information Technology Innovation

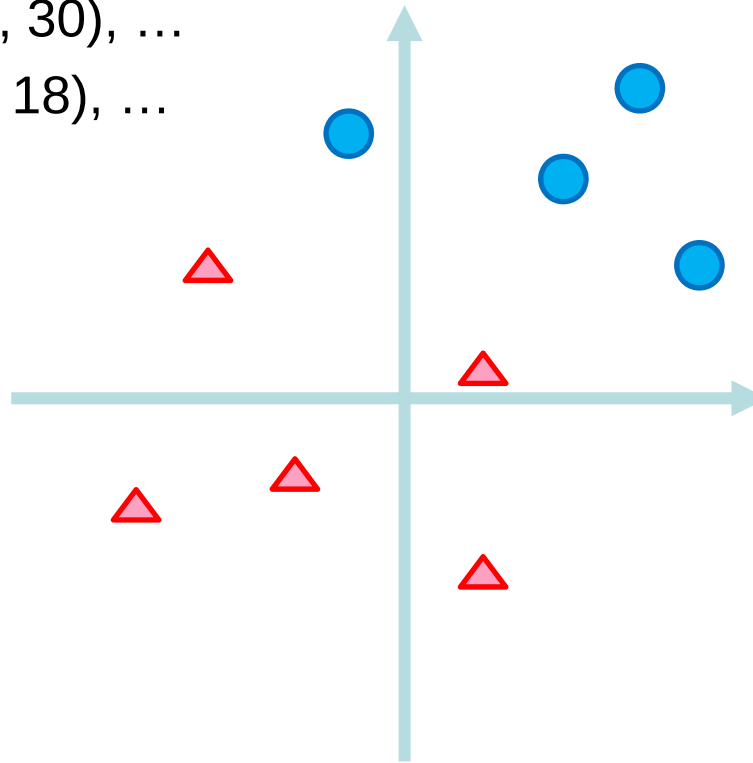
Machine Learning in Artificial Intelligence

- Machine learning is a method of data analysis that automates analytical model building.



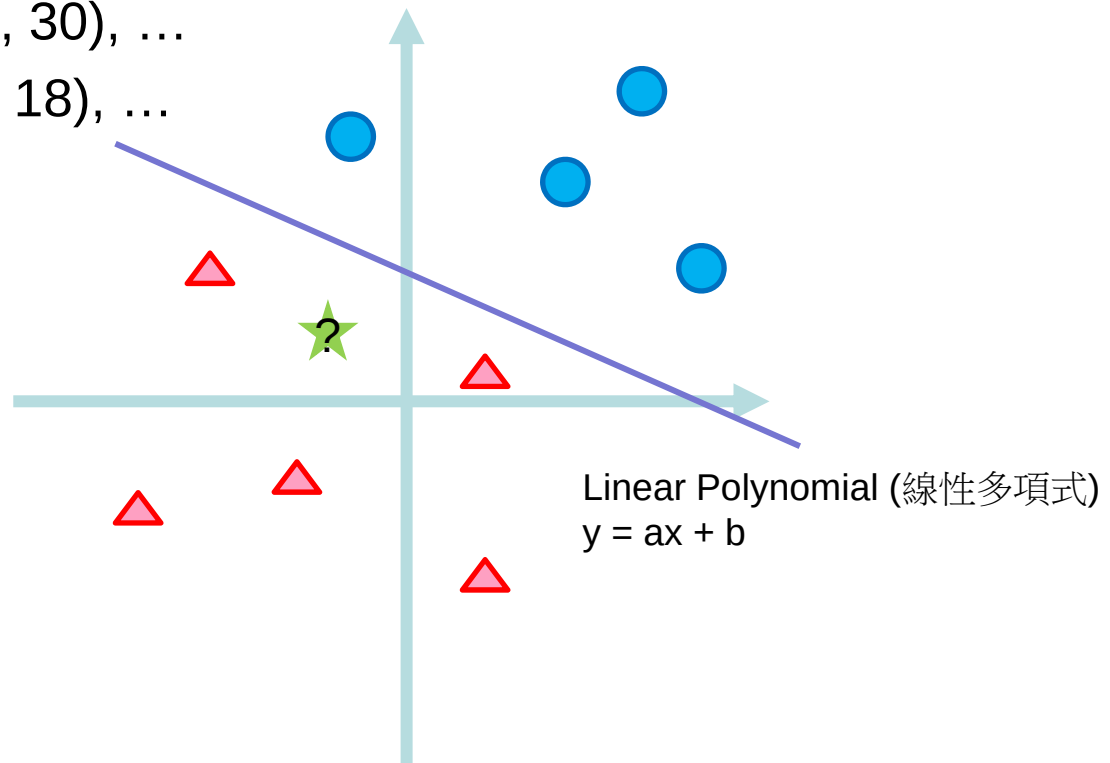
A Demonstrative Example

- Algorithm (SVM, Support Vector Machine)
- Training Data
 - ● Dog: (60, 45), (38, 30), ...
 - ▲ Cat: (30, 25), (25, 18), ...

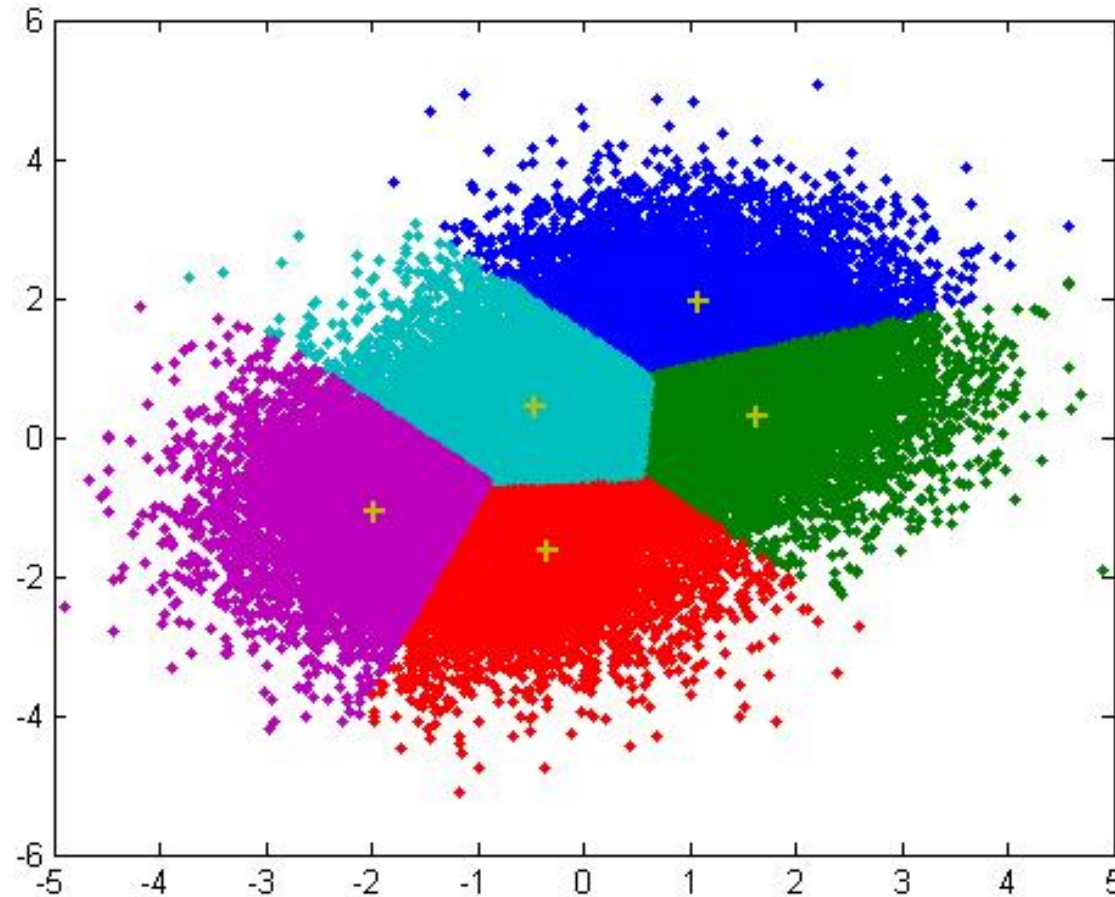


A Demonstrative Example

- Algorithm (SVM, Support Vector Machine)
- Training Data
 - ● Dog: (60, 45), (38, 30), ...
 - ▲ Cat: (30, 25), (25, 18), ...



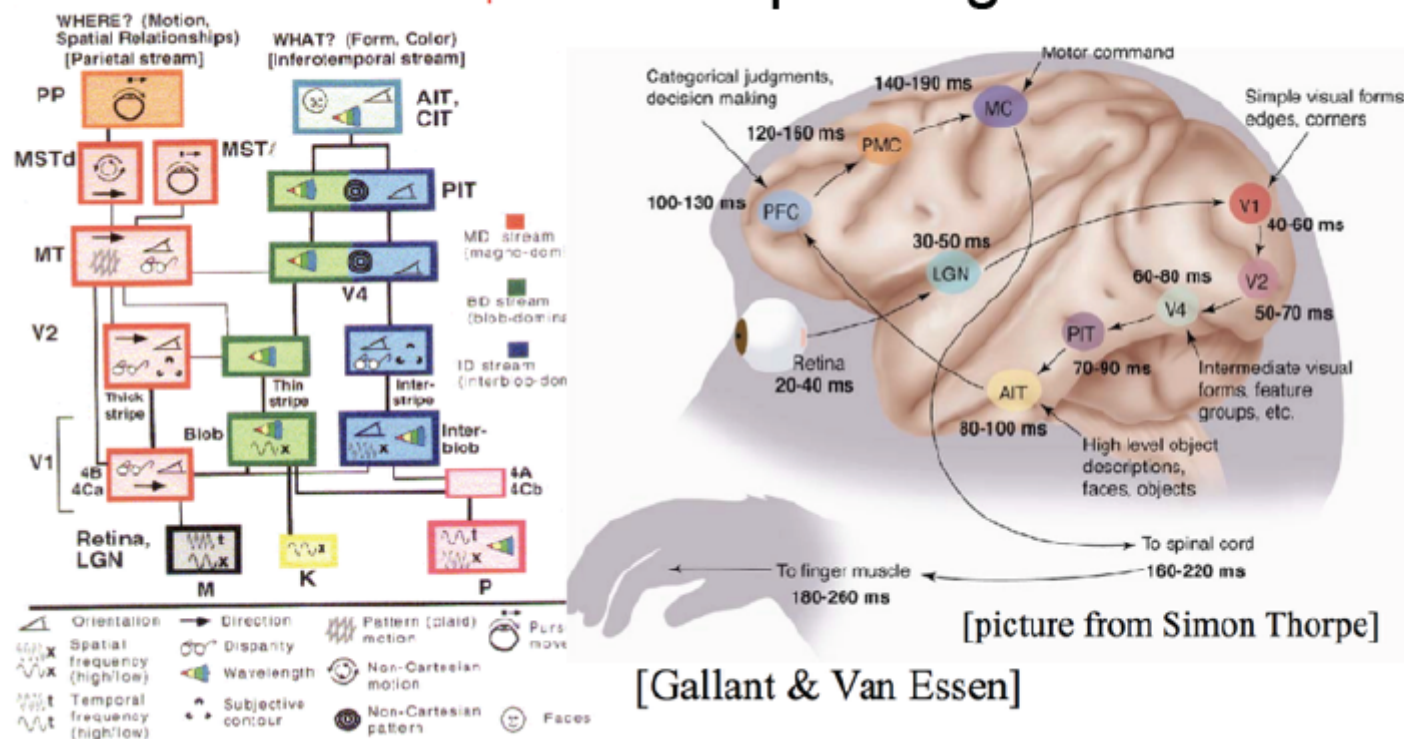
Real Data are Complex



Deep Learning (深度學習)

- Using brain-inspired mechanics to achieve brain-like function

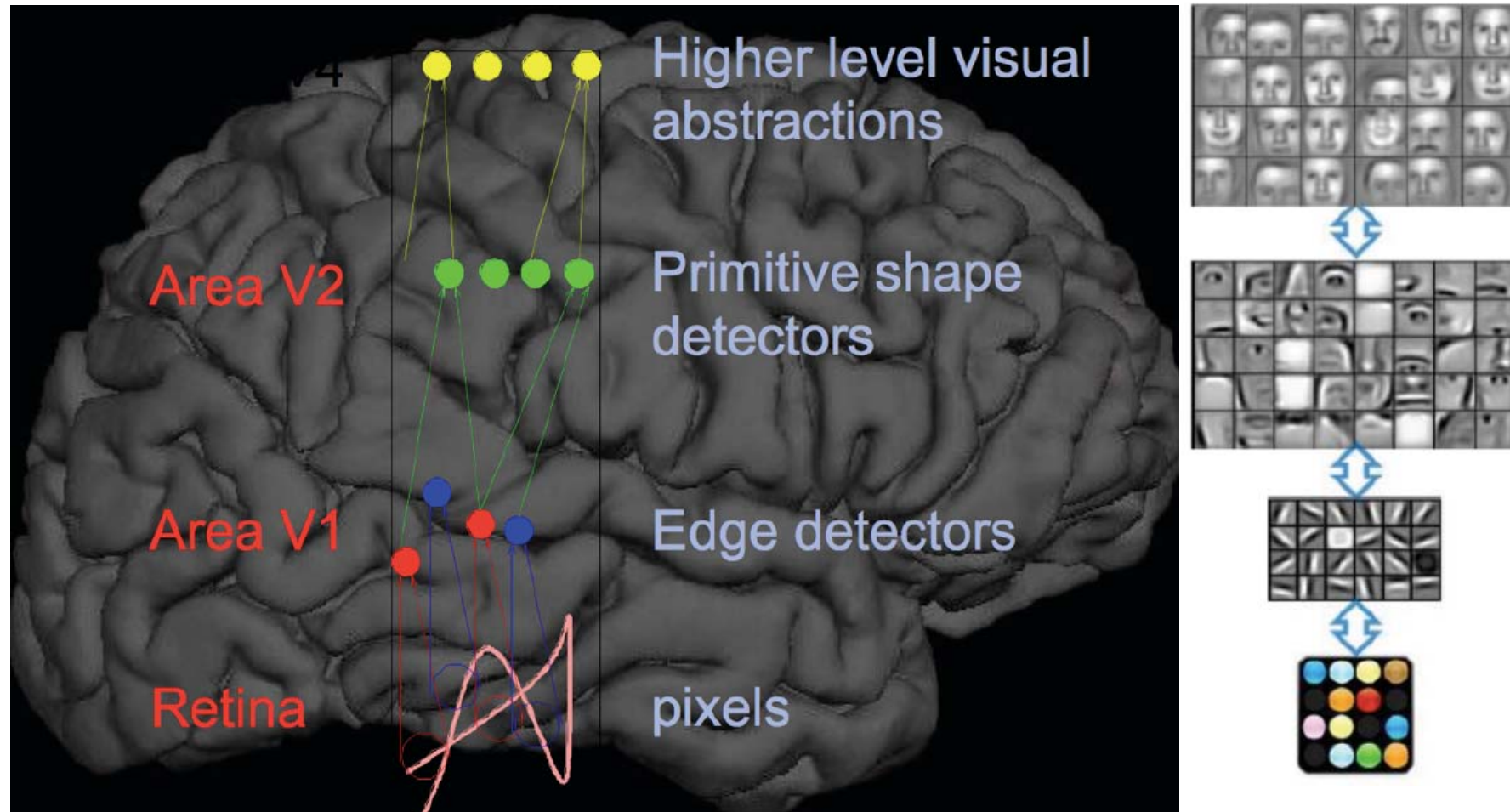
Audio/Visual Cortex has multiple stages == Hierarchical



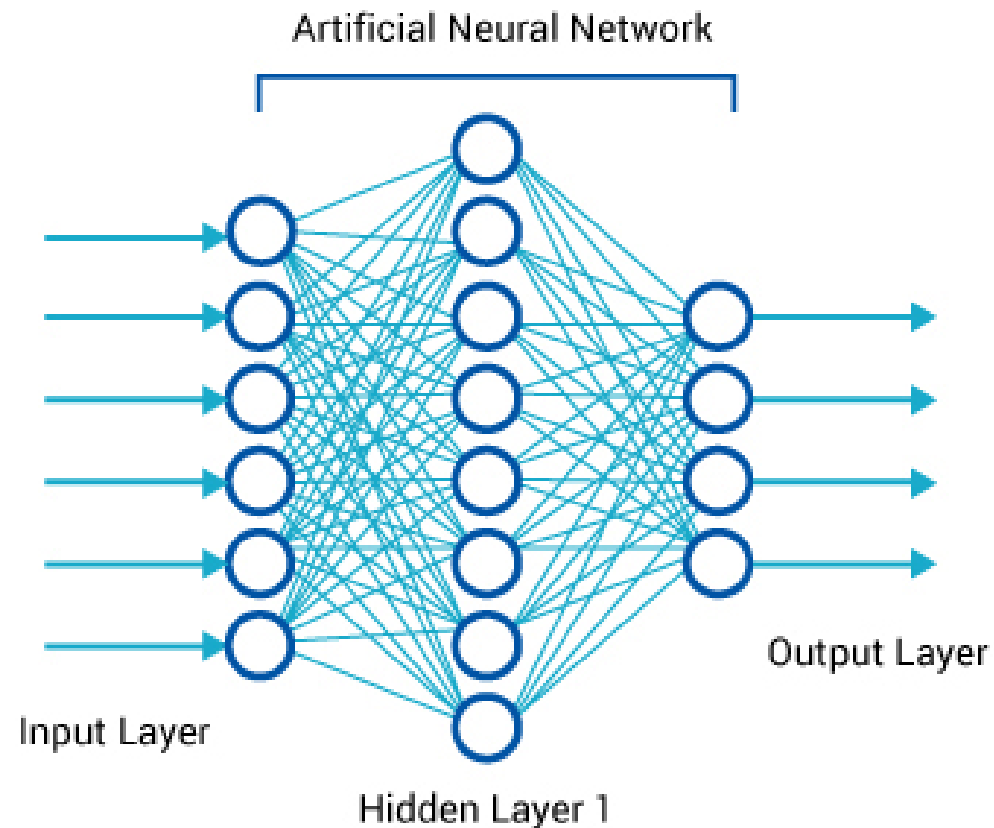
[Ref] <http://www.slideshare.net/roelofp/041114-dl-nlpwordembeddings>



Different Levels of Abstraction



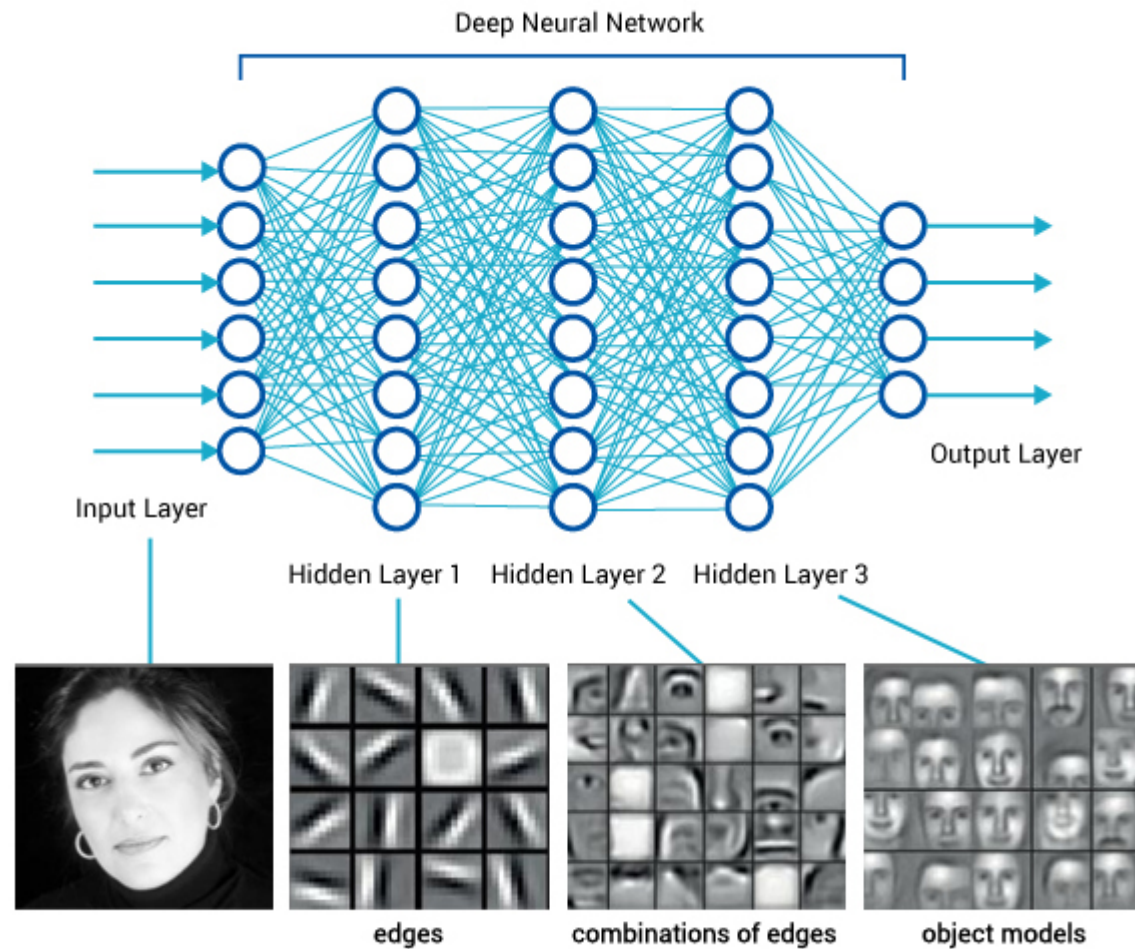
Basic Neural Network



[Ref] <http://www.amax.com/blog/?p=804>



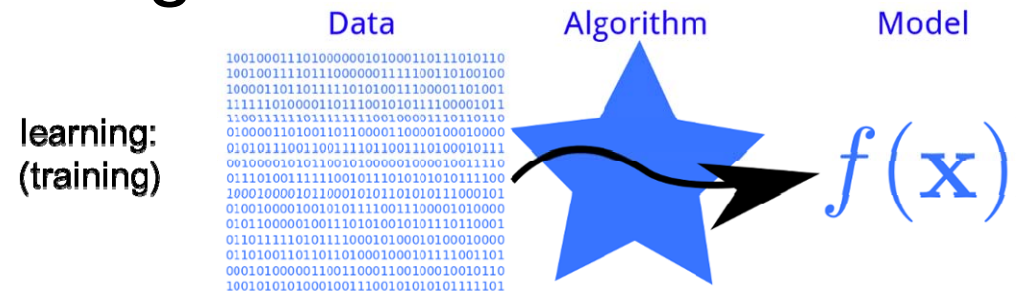
Deep Neural Network



Example:

Learning a classifier for numbers

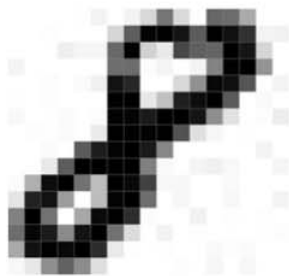
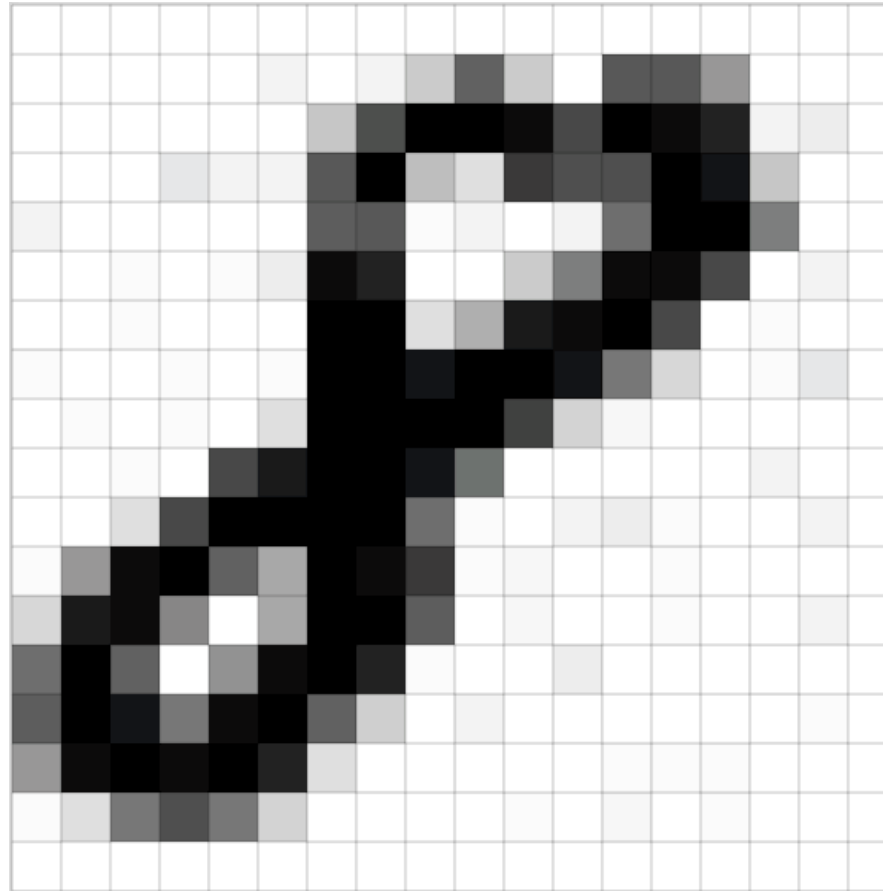
- Preparation of Training Data



8 2 7 7 5 7 7 2 8 8 5 7 0 7 1 7 5 9 3 1 0 2 7 9 9 6 9 4 2 4 1 1 4 4 8 8 0 2 6 3
0 0 7 6 3 4 4 4 3 4 2 3 2 8 0 8 2 9 7 6 7 9 0 0 4 2 0 6 6 4 3 3 9 0 4 7 3 2 2 0
2 6 4 6 4 7 5 9 8 7 1 9 0 6 8 7 7 1 9 8 6 5 7 1 0 1 0 8 3 4 7 7 1 3 0 9 6 0 3 8
0 2 8 3 6 5 7 6 6 7 2 6 1 0 2 6 9 7 1 9 5 8 7 0 0 6 1 6 4 4 8 6 2 3 3 1 3 9 9 4
5 1 0 2 1 4 2 2 0 9 9 9 3 1 3 4 1 9 5 5 4 3 9 3 3 5 8 5 0 6 5 1 8 2 6 8 9 2 2 8
4 7 2 7 5 5 0 7 2 2 1 3 5 8 4 8 8 5 2 5 7 1 6 1 8 3 8 0 0 1 0 3 6 2 4 0 8 6 6 2
1 3 3 9 0 4 9 7 6 4 9 5 5 2 6 9 5 3 4 7 3 0 4 6 2 9 4 0 6 2 7 1 0 3 9 1 2 6 0 6
3 4 1 1 9 0 8 2 1 1 9 0 7 5 7 4 2 3 9 9 0 0 2 5 2 1 3 8 3 3 1 6 7 6 0 7 2 0 0 5
7 1 3 1 2 8 8 2 9 4 4 2 4 7 9 8 4 8 0 3 0 7 8 8 3 9 4 7 3 3 1 6 0 8 7 2 1 1 6 2
6 0 1 7 2 3 6 1 6 5 0 7 8 7 8 6 9 2 3 8 8 6 5 1 1 3 2 6 0 6 0 5 9 9 1 0 2 2 1 9

[Ref] <https://medium.com/@ageitgey/machine-learning-is-fun-part-3-deep-learning-and-convolutional-neural-networks-f40359318721#.ws86qjlwk>





=

```
[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 12, 0, 11, 39, 137, 37, 0, 152, 147, 84, 0, 0, 0, 0,
0, 1, 0, 0, 0, 41, 160, 250, 255, 235, 162, 255, 238, 206, 11, 13, 0, 0, 0, 0, 16, 9, 9, 150, 251, 45, 21, 184, 159, 154, 2
55, 233, 40, 0, 0, 10, 0, 0, 0, 0, 0, 145, 146, 3, 10, 0, 11, 124, 253, 255, 107, 0, 0, 0, 0, 3, 0, 4, 15, 236, 216, 0, 0,
38, 109, 247, 240, 169, 0, 11, 0, 1, 0, 2, 0, 0, 0, 253, 253, 23, 62, 224, 241, 255, 164, 0, 5, 0, 0, 6, 0, 0, 4, 0, 3, 252
, 250, 228, 255, 255, 234, 112, 28, 0, 2, 17, 0, 0, 2, 1, 4, 0, 21, 255, 253, 251, 255, 172, 31, 8, 0, 1, 0, 0, 0, 0, 0, 4,
0, 163, 225, 251, 255, 229, 120, 0, 0, 0, 0, 0, 11, 0, 0, 0, 0, 21, 162, 255, 255, 254, 255, 126, 6, 0, 10, 14, 6, 0, 0, 9
, 0, 3, 79, 242, 255, 141, 66, 255, 245, 189, 7, 8, 0, 0, 5, 0, 0, 0, 0, 26, 221, 237, 98, 0, 67, 251, 255, 144, 0, 8, 0, 0
, 7, 0, 0, 11, 0, 125, 255, 141, 0, 87, 244, 255, 208, 3, 0, 0, 13, 0, 1, 0, 1, 0, 0, 145, 248, 228, 116, 235, 255, 141, 34
, 0, 11, 0, 1, 0, 0, 0, 1, 3, 0, 85, 237, 253, 246, 255, 210, 21, 1, 0, 1, 0, 0, 6, 2, 4, 0, 0, 0, 6, 23, 112, 157, 114, 32
, 0, 0, 0, 0, 2, 0, 8, 0, 7, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
```



Challenges: Observed Real-World Objects from Mobile Cameras are in Uncontrolled Settings

Jersey numbers



- In real-world scenes, an observed visual object varies highly in size, position, pose, and even appearance.
- The background is also complex with different lighting conditions, shadowing, and occlusions.

Google's Deep Learning AI Knows Where You Live And Can Crack CAPTCHA

Written by Harry Fairhead

Monday, 21 April 2014



<http://www.i-programmer.info/news/105-artificial-intelligence/7206-googles-deep-learning-ai-knows-where-you-live-and-can-crack-captcha.html>



Business Signs (BSs)

- A common form of advertisement often attached to or is part of a building
- Unlike logos or trademarks, a BS is not a fixed pattern but a mixture of texts (e.g. the business's name) and images (e.g. the business's logo).



Why Business Signs?

- A common form of advertising signs, often attached to or is part of a building, especially in Eastern Asia



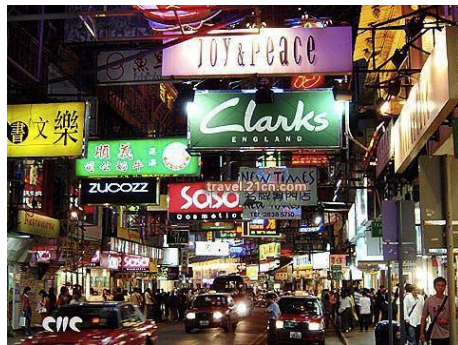
USA (New York)



Taiwan (Taipei)



Japan (Tokyo)



Hong Kong



China (Beijing)



Korea (Seoul)



Enabling Mobile Sensing for BS Objects

Position the phone camera to view the target scene



Select a recognized store as the preference



Show the store information, e.g. discounts or promotions



Image Databases for Mobile Visual Search

ZuBuD



Oxford



INRIA



UKY



Image Nets



SMVS



Vijay Chandrasekhar et. al. , MMSys 2011.



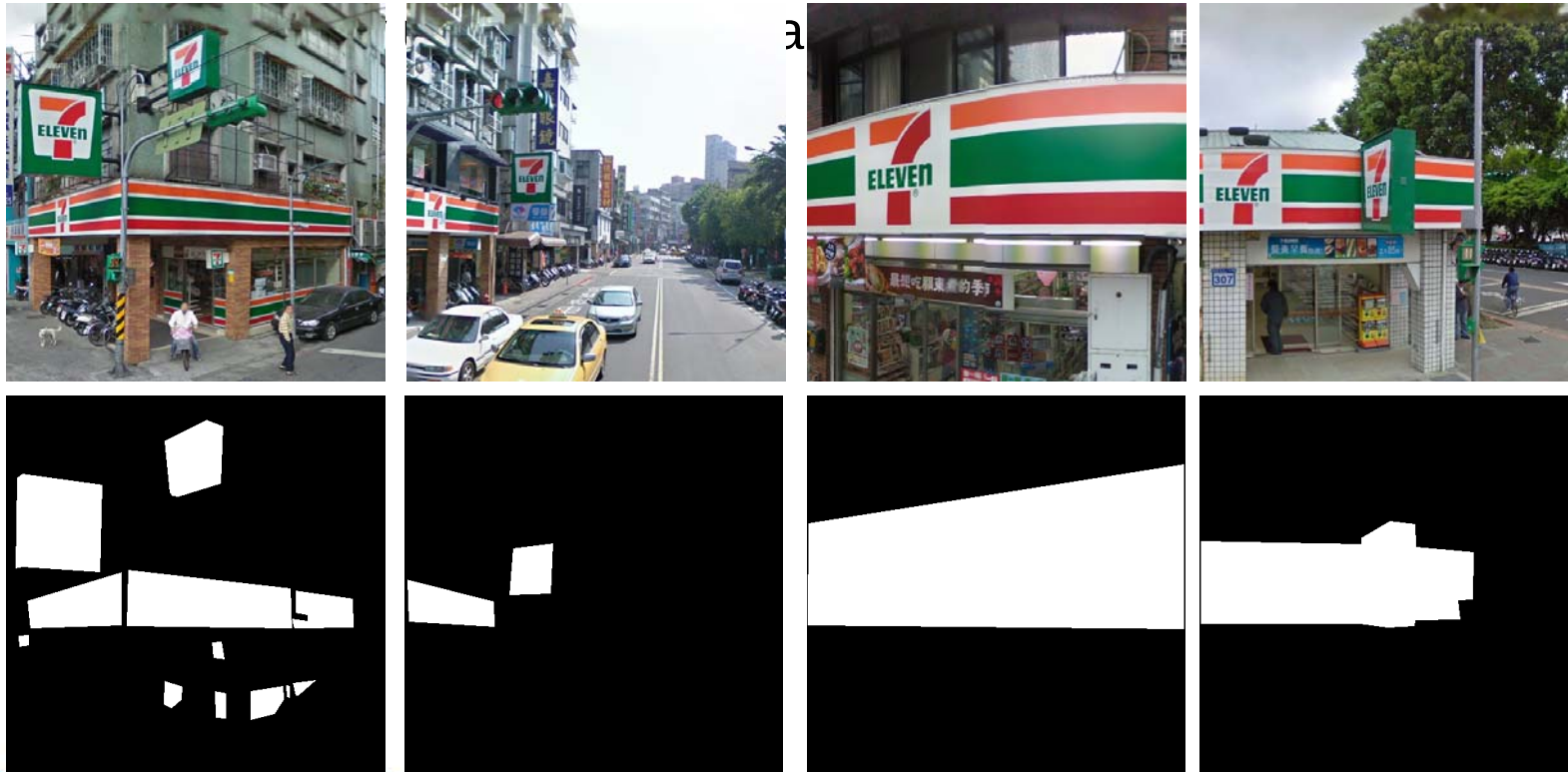
What is Lacking in these Databases?

- Objects in real-world images are often
 - in arbitrary size
 - in arbitrary position
 - in arbitrary viewing angles
 - accompanied with perspective distortion
 - accompanied with foreground and background clutter
 - accompanied with varying lighting conditions



OPS-62 Data Set

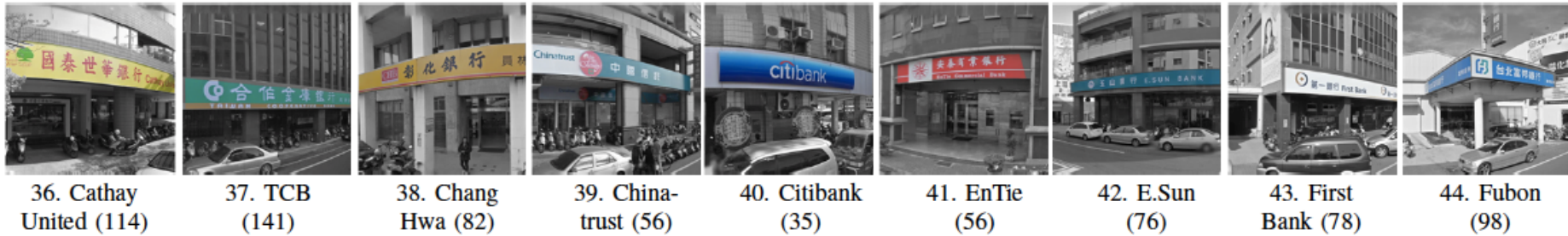
- 4,649 OPS (On-Premise Sign) images in total
 - now with 62 different categories (chain stores)
 - each OPS image with a pixel-level label (binary mask) and GPS tag



Business type	Specialty Stores				Convenience Stores			
<i>Sample image</i>								
OPS category (# of images)	1. Aso Shoes (135)	2. BlockBuster (32)	3. Elifemall (75)	4. Hsin Tung Yang (38)	5. 7-Eleven (200)	6. Family (55)	7. Hi-Life (85)	8. OK (101)
	Soft-Drink Stores				Pharmacy Stores		Supermarkets	
								
	9. 50 Lan (84)	10. Ching Shin (70)	11. Dante Coffee (55)	12. Ikari Coffee (33)	13. Orange House (34)	14. Starbucks Coffee (55)	15. Cosmed (85)	16. Watsons (116)
	Supermarkets		Telecomm Service Centers			Outlets		
								
	18. Pxmart (60)	19. Wellcome (54)	20. Chunghwa Telecom (76)	21. Fareaston (84)	22. Taiwan Mobile (68)	23. Vibo (91)	24. Aurora (112)	25. Arcoa (108)
	Restaurants							
								
	27. Bafang Yunzi (82)	28. Domino's Pizza (91)	29. Overseas Dragon (54)	30. Mc- Donald's (43)	31. MOS Burger (102)	32. KFC (79)	33. Pizza Hut (32)	34. Subway (50)
								35. Sushi Express (55)



Commercial Banks



Commercial Banks



Gas Station



54. CPC (52)

Glasses Shops



55. Formosa Optical (127)



56. Kobayashi Optical (86)



57. Century 21 (63)



58. CREC (90)



59. Eastern Realty (109)



60. Sinyi Realty (84)



61. Taiwan Realty (84)

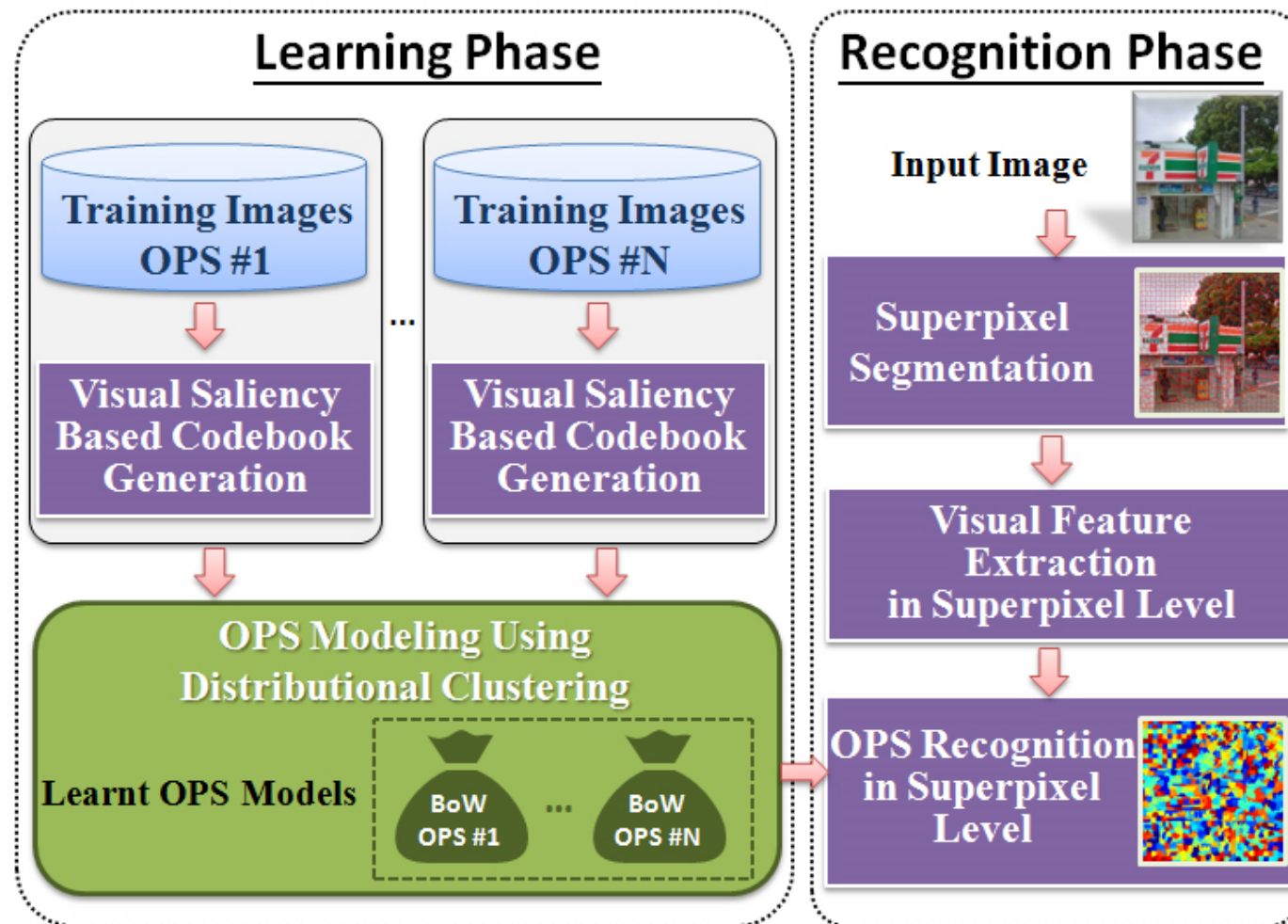


62. Yung Ching (90)

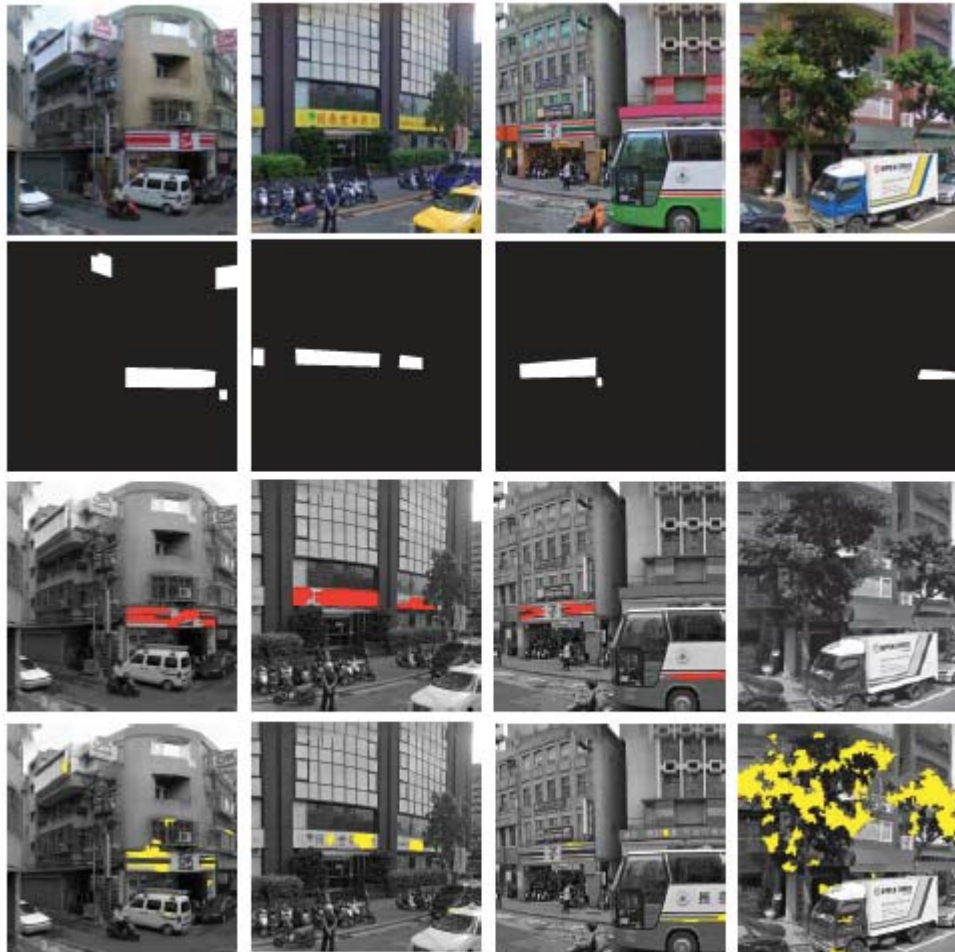
Realtors Offices



Overview of the proposed system framework



Sample Results



images

ground truth

our approach

F-pLSA



NVIDIA Self-Driving Car @ CES 2016

AI Learns How to Driving by Playing Video Games

NEC Face Recognition

Nanit Baby Monitor

AR Smart Glasses

Continuous and Real-time Object Recognition

- Applications **locate** objects in a video stream and **label** them with information associated with the objects.
- Challenges
 - The required computer-vision algorithms are computationally intensive and must run at the real-time rate (e.g., 30fps)



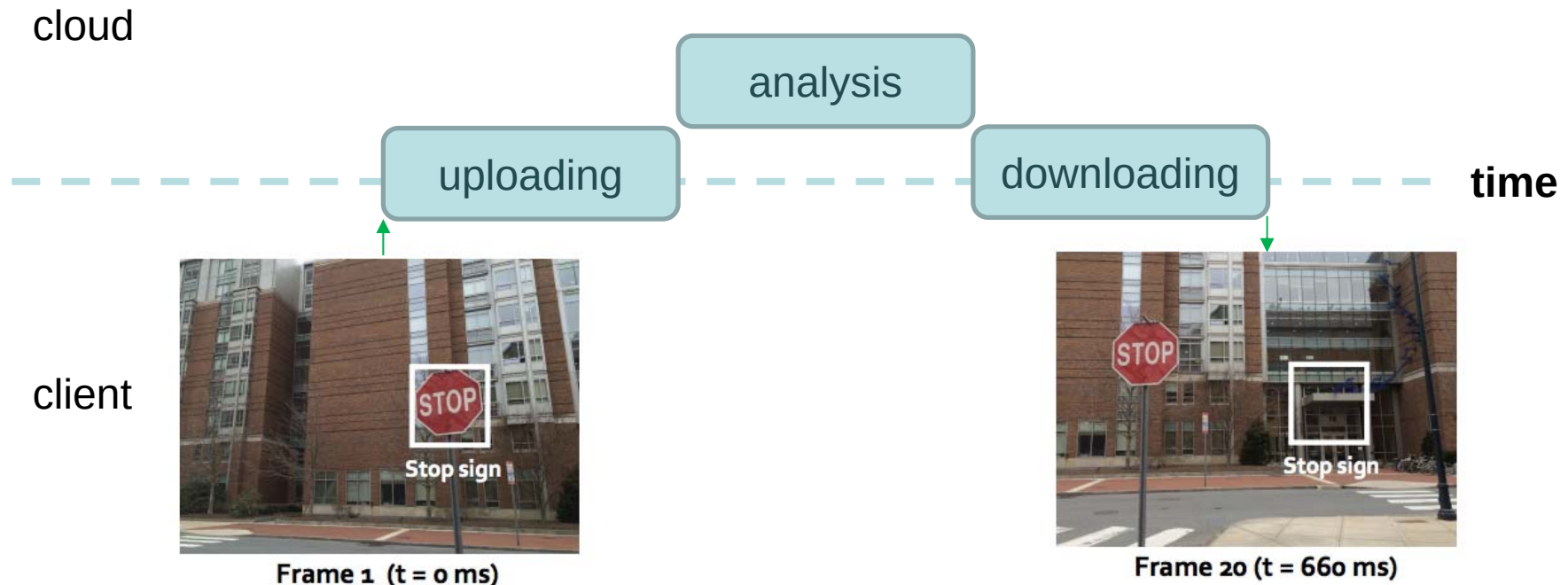
Cloud Computing

- Offloading the Computation to More Powerful Servers



However, there is no free lunch...

- Issue: wireless network latencies are high

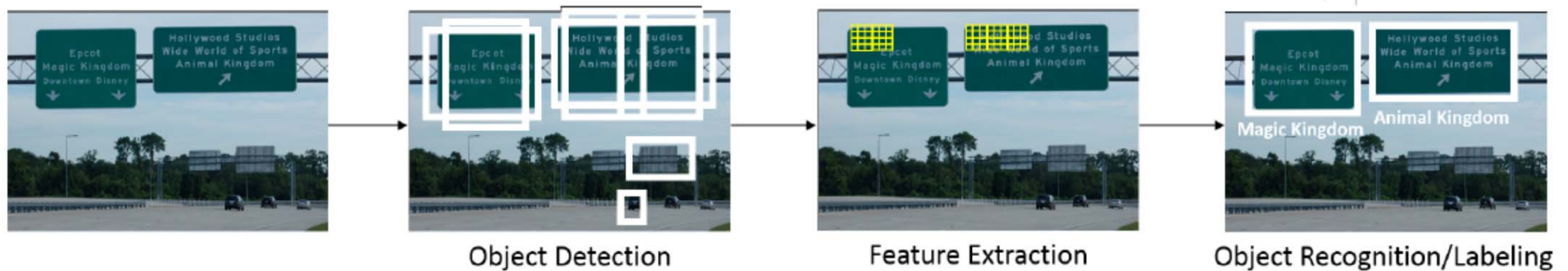


- Offloading every frame to a server reduces trackability: the stop sign's location is wrong

[Ref] "Glimpse: Continuous, Real-Time Object Recognition on Mobile Devices," ACM SenSys, 2015.

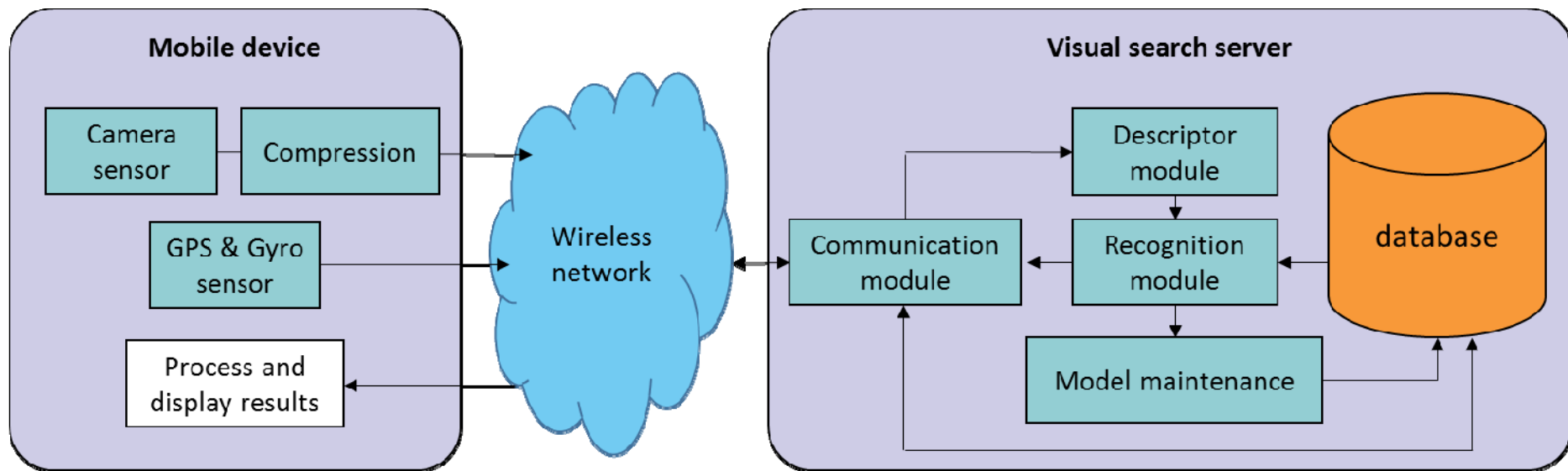


Object Recognition Pipeline



Mobile-Cloud Architecture (I)

- The mobile device transmits the compressed image, while analysis of the image is done entirely on a remote server.

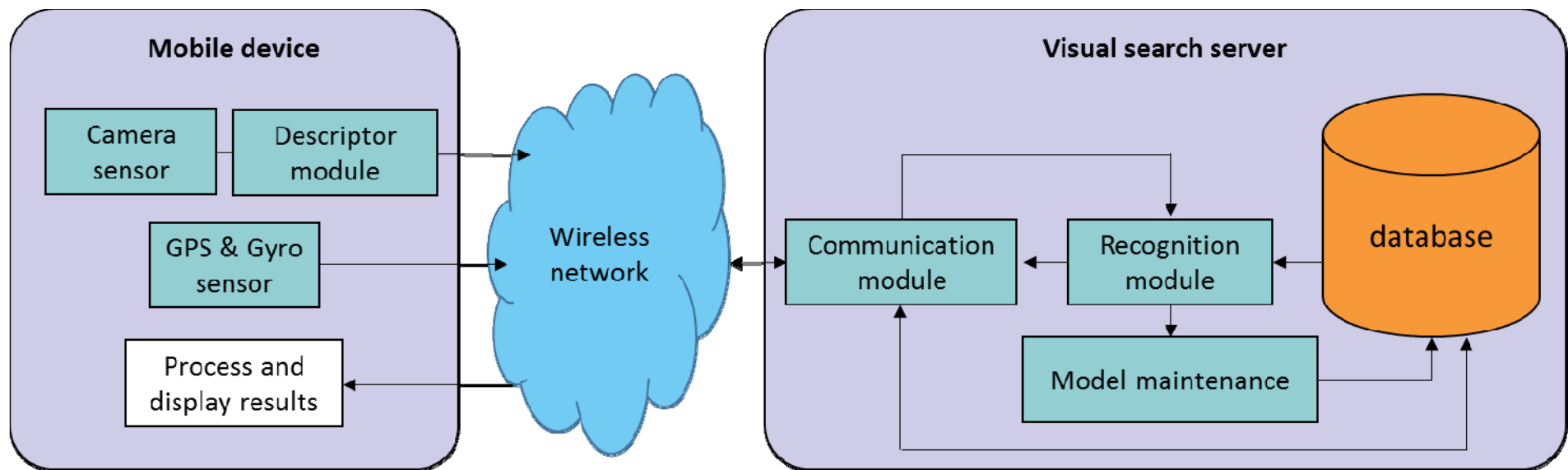


[Ref] "Mobile visual search: Architectures, technologies, and the emerging MPEG standard," IEEE Multimedia, 2011.



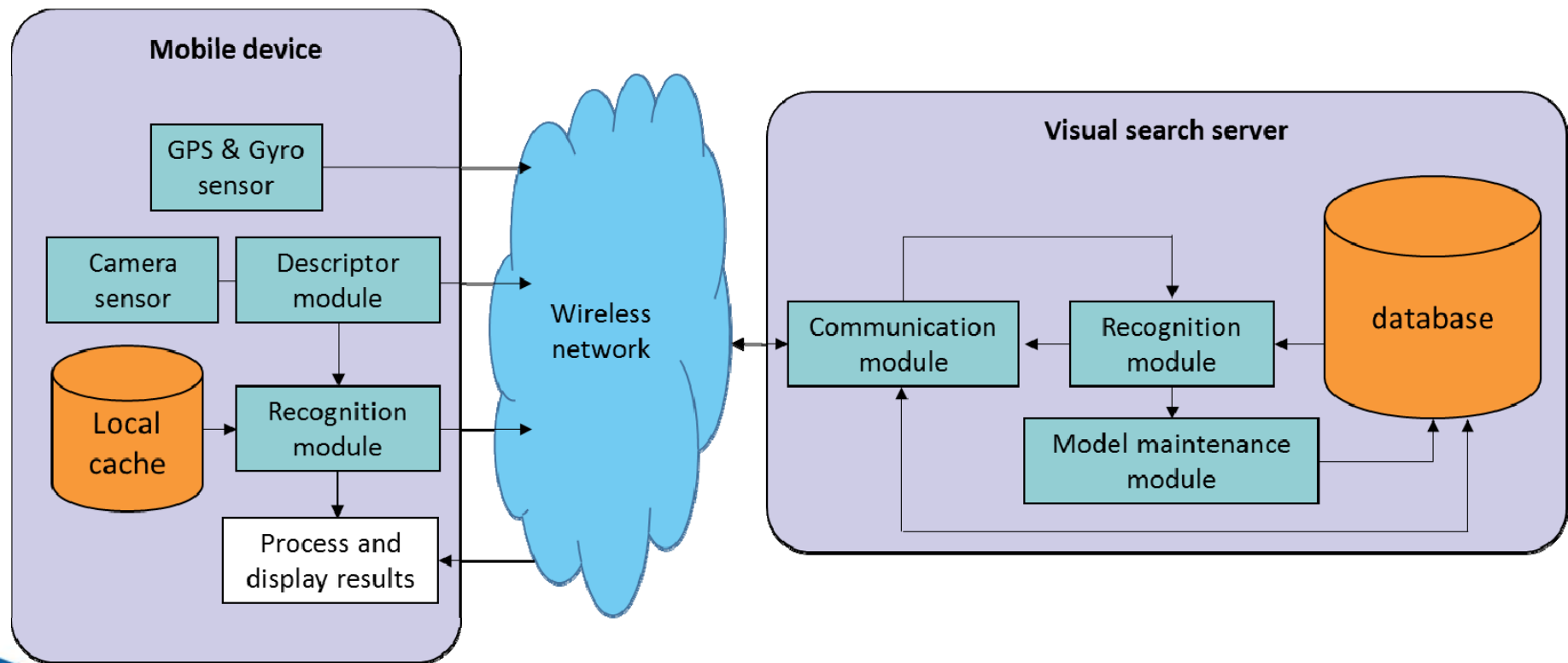
Mobile-Cloud Architecture (II)

- The local image features (descriptors) are extracted on a mobile phone and then encoded and transmitted over the network. Such descriptors are then used by the server to perform the analysis.



Mobile-Cloud Architecture (III)

- The mobile device maintains a cache of the database and sends analysis requests to the remote server only if the object of interest is not found in this cache, further reducing the amount of data sent over the network.

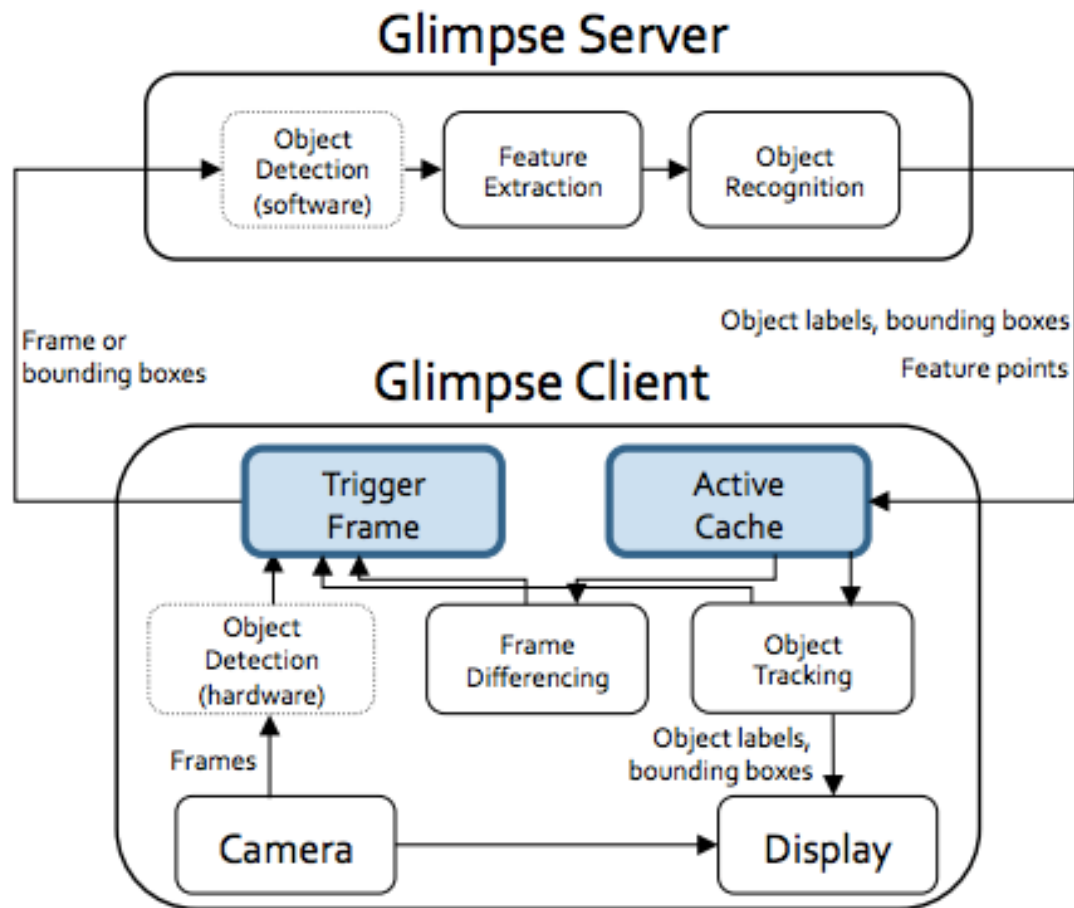


Glimpse by MIT+Microsoft

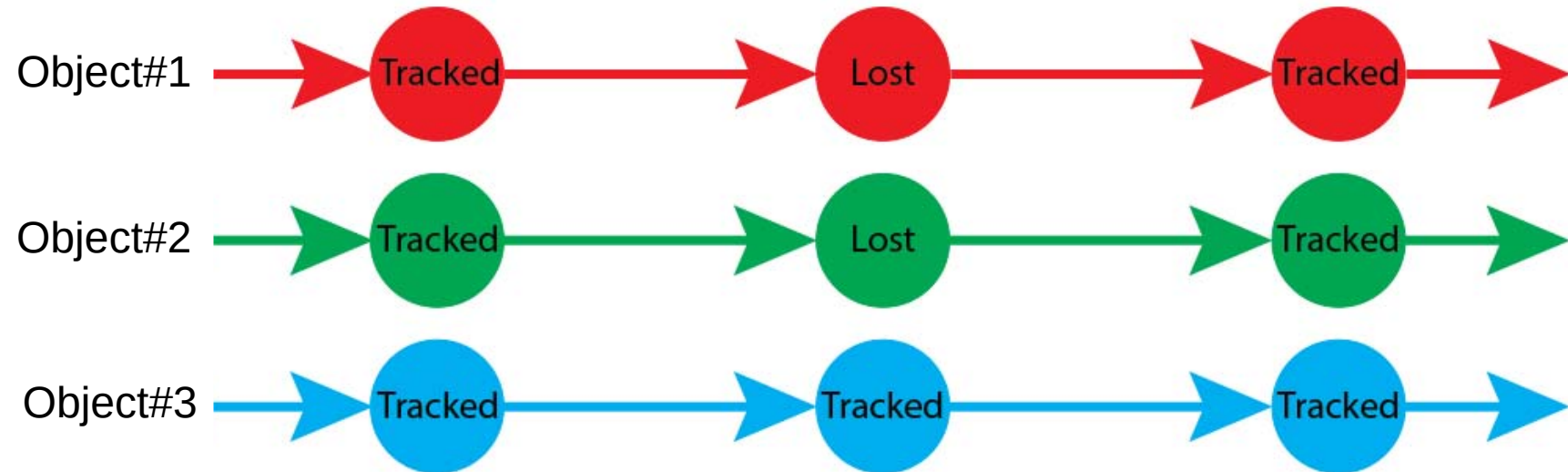
- Glimpse is a continuous, real-time object recognition system for camera-equipped mobile devices. Glimpse captures full-motion video, locates objects of interest, recognizes and labels them.
 - **Active cache**: Glimpse reduces bandwidth consumption by strategically sending only certain **trigger frames** to the server to obtain object recognition hints.



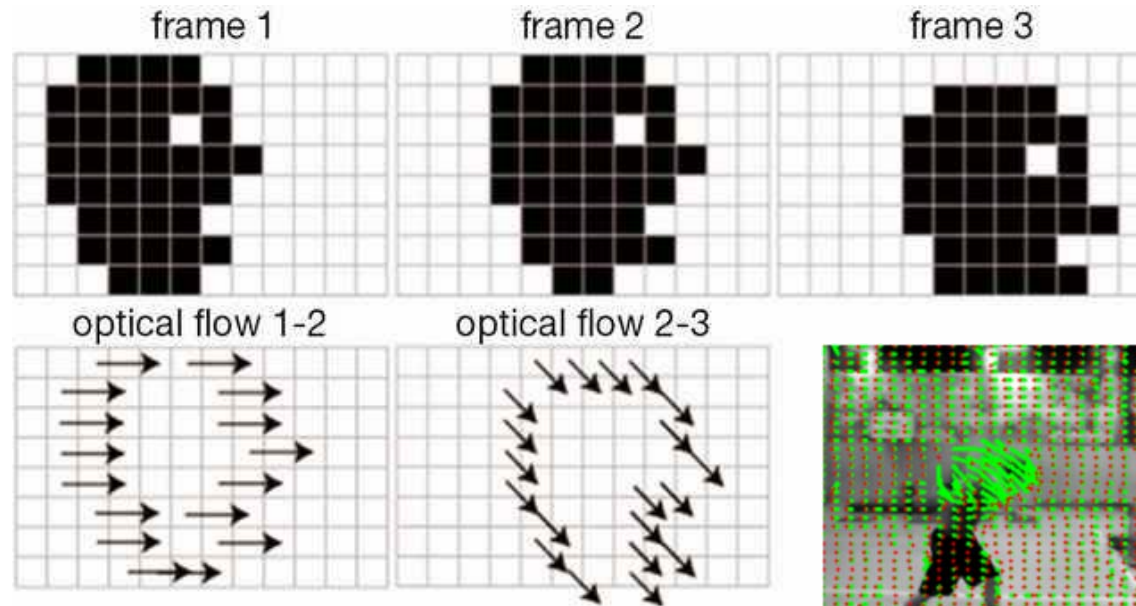
Glimpse Architecture



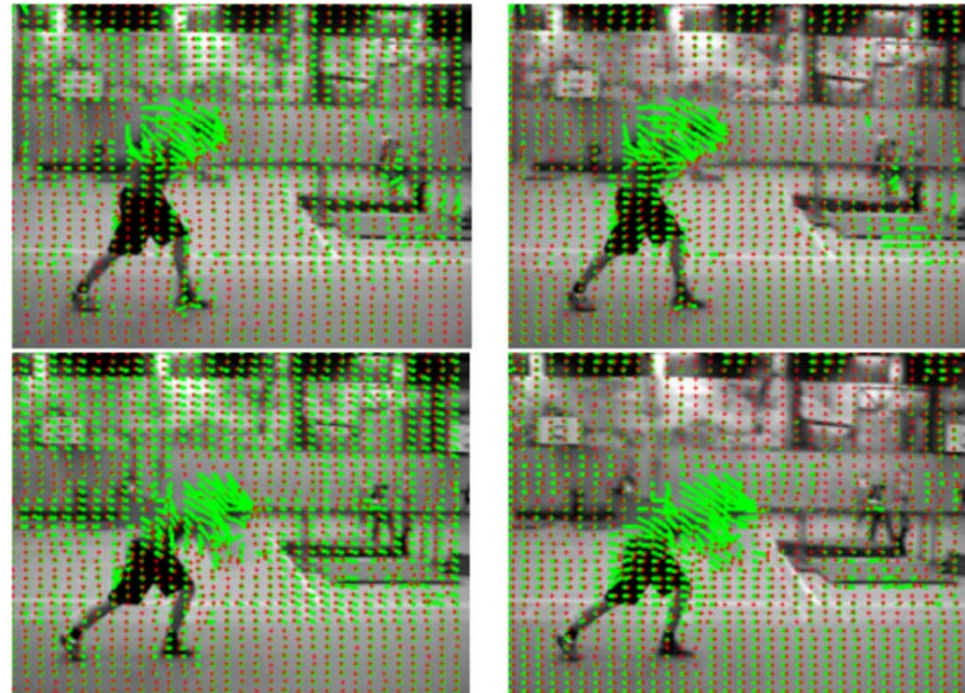
Visual Object Tracking



Optical Flow



The pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer and a scene



Demonstration: Server-only

Network: Verizon Wireless's LTE

- The client sends frames to the server and outputs the responses as they arrive.



[Ref] <http://people.csail.mit.edu/yuhan/glimpse/>



Research Center for Information Technology Innovation

Demonstration: No Active Cache

Network: Verizon Wireless's LTE

- The client tracks between the frame sent to the server (processed frame) and the current frame, without an active cache. It sends a new frame to the server once we receive a response for the previous frame.



Demonstration: Active Cache

Network: Verizon Wireless's LTE

- The client maintains an active cache and tracks through frames using the technique of trigger frames.



Mahimahi: A Lightweight Toolkit for Reproducible Web Measurement

<http://mahimahi.mit.edu/>

Mahimahi

A set of lightweight tools for browser developers, website authors, and network protocol designers that provides accurate measurements when recording and replaying HTTP content over emulated network conditions.

Mahimahi is free software and is available on Ubuntu (version 14.04 or higher).

[Getting Mahimahi »](#)

Record and replay HTTP traffic

Using Mahimahi's RecordShell, you can completely record and store real HTTP content (such as Web pages) to disk. Recorded sites can then be repeatably replayed using ReplayShell. ReplayShell accurately models a Web application's multi-server structure by locally mirroring a page's server distribution.

Emulate numerous network conditions

Mahimahi's network emulation tools can be used to emulate many different link conditions for replaying recorded HTTP traffic. Mahimahi supports emulating fixed propagation delays (DelayShell), fixed and variable link rates (LinkShell), stochastic packet loss (LossShell). LinkShell also supports a variety of queueing disciplines such as DropTail, DropHead, and active queue management schemes like CoDel. Each of Mahimahi's

Use unmodified client applications

Recording and replaying sites can be done using commercial browsers such as Google Chrome, Mozilla Firefox, and Safari. Additionally, other client applications including virtual machines and mobile-device emulators can be run unmodified within each of Mahimahi's tools.

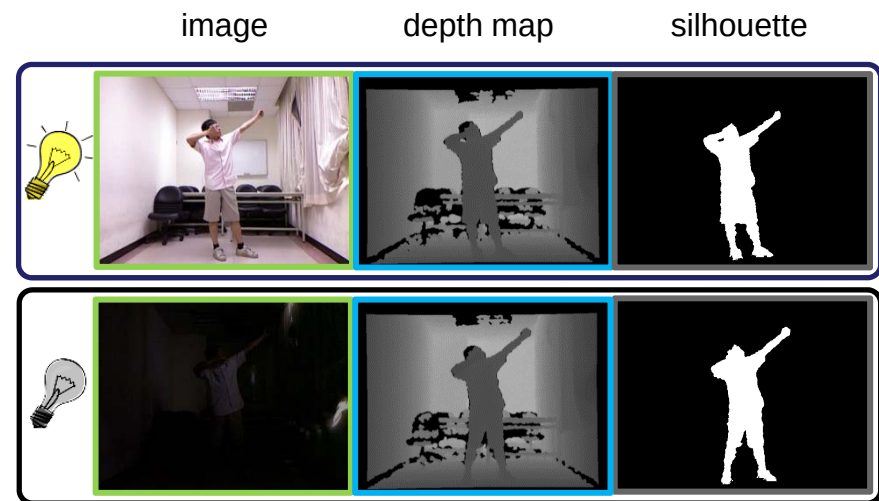


Research Center for Information Technology Innovation

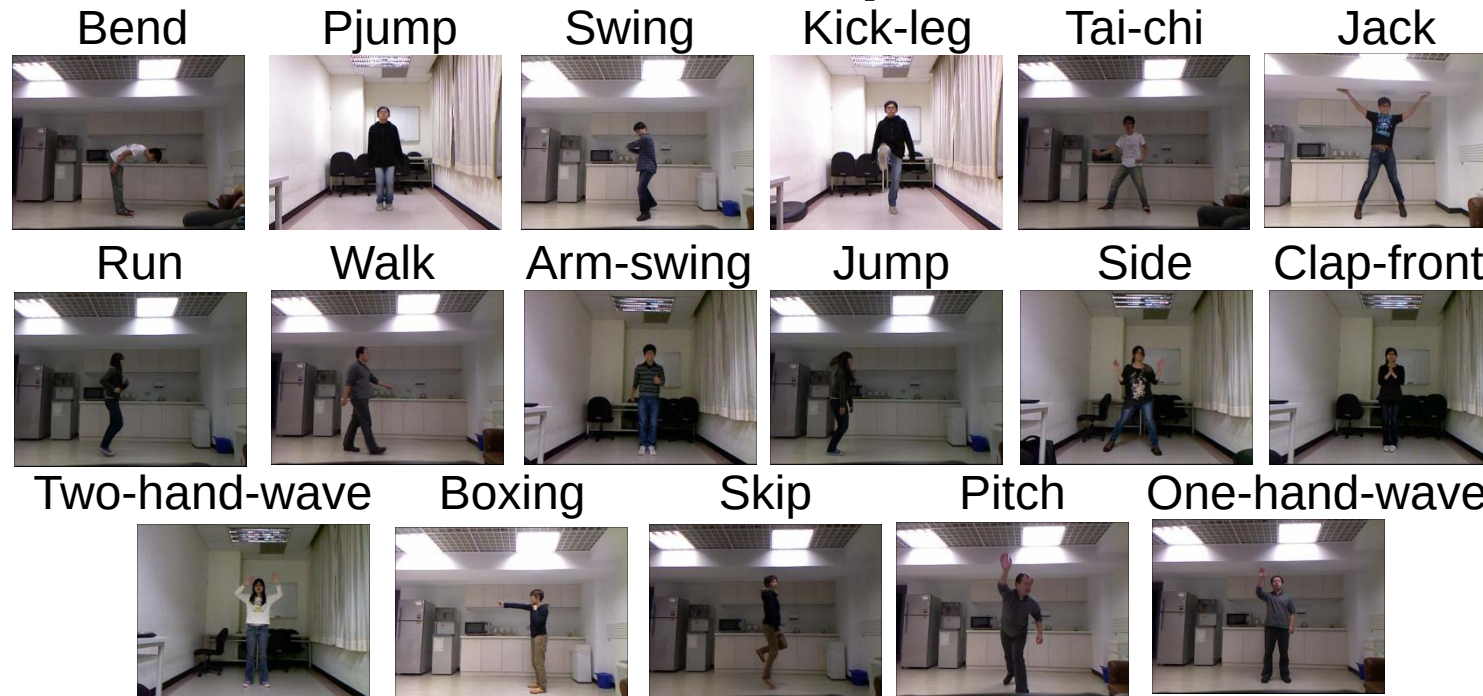
Human Posture/Gesture Analysis

- Witnessing the trend of rapid increase in consumer 3D sensors (e.g. Kinect), we go beyond the traditional color-based approaches by **using the depth information only**

- Simple but effective novel depth features are proposed
- Showing competitive performances
- Pioneering work in the related fields



Benchmark Development



- Depth-included **H**uman **A**ction video dataset
 - 17 categories, each with 21 recordings from different users
 - One of the largest depth-included video datasets of human actions in the literature

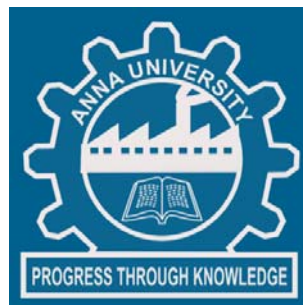
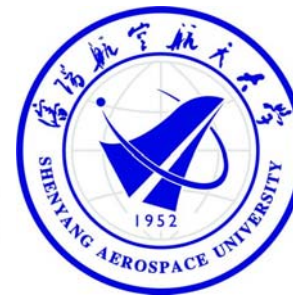




Universitat d'Alacant
Universidad de Alicante



Inria
INVENTEURS DU MONDE NUMÉRIQUE



北京大学
PEKING UNIVERSITY



北京邮电大学

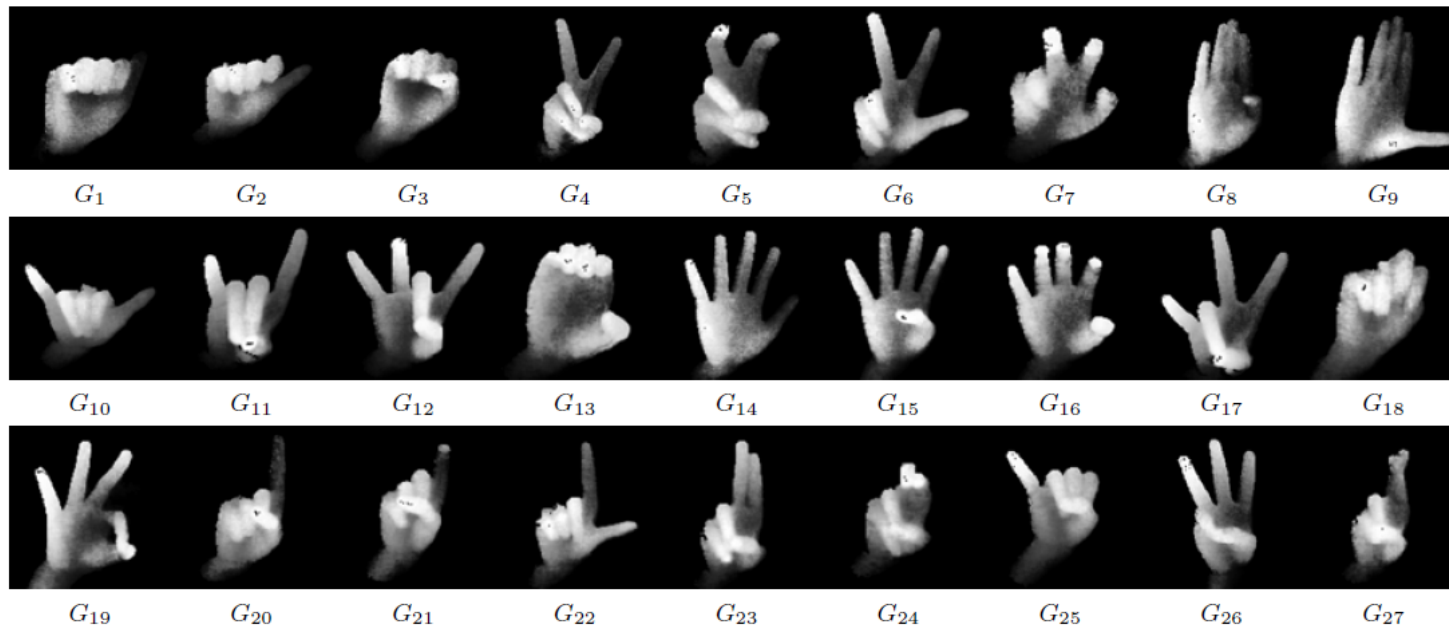
BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS



Research Center for Information Technology Innovation

Hand Gesture Tracking

- Hand gesture is a natural way for HCI
 - challenges: high-dimensionality problem, self-occlusions, processing speed, uncontrolled environments, and rapid hand motion



Human Action Recognition in EMSS (Emergency Medical Services System)

科技部
Ministry of Science and Technology

Outstanding Young Scholar Award, 2012-2015

Collaborator: NTU-Hospital



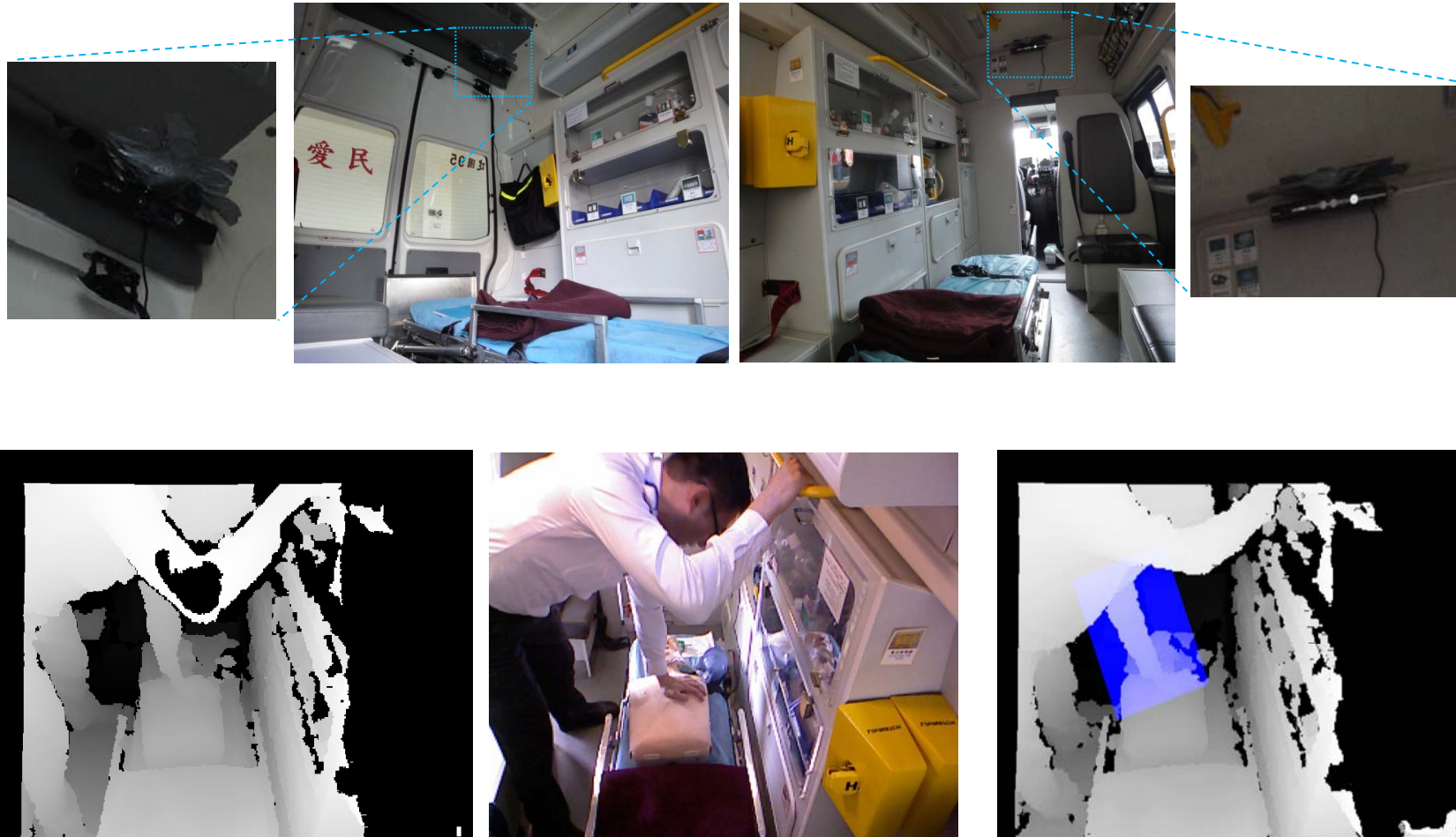
- EMSS is critical to public health system and the survival of out-of-hospital cardiac arrest (OHCA) would best reflect the level of EMSS in a community
 - Calls for regular monitoring of CPR quality
 - i.e. ensuring high-quality CPR:
adequate rate, depth, and recoil
 - Our participation
 - To help develop computer-aided technologies by deploying 3D sensors in ambulance to assess operator-related factors affecting CPR quality during ambulance transport



Paramedics performing CPR
on patient in ambulance



CPR Tracking with Depth Camera



Heart Bus

Collaborators:  Taipei City Government
國立臺北藝術大學 Taipei National University of the Arts
 中國信託慈善基金會 CTBC Charity Foundation

An Experience Bus for Real-time Interactive Training of Rescuer Skills

- The dissemination of rescuer knowledge is crucial in public education
- The world's first experience bus for real-time interactive training of rescuer skills (CPR, AED, Heimlich), by leveraging the motion-sensing techniques



The bus is now circulating in Taipei City for public experience.



CPR



AED



Heimlich



Research Center for Information Technology Innovation



iCPR Demo



E-Certificate

iCPR電子學習證明



姓名：H. T. Kung
性別：男生
出生年次：2014

學習日期
2014.08.11

中央研究院
資訊科技創新研究中心
Research Center for Information Technology Innovation
Academia Sinica

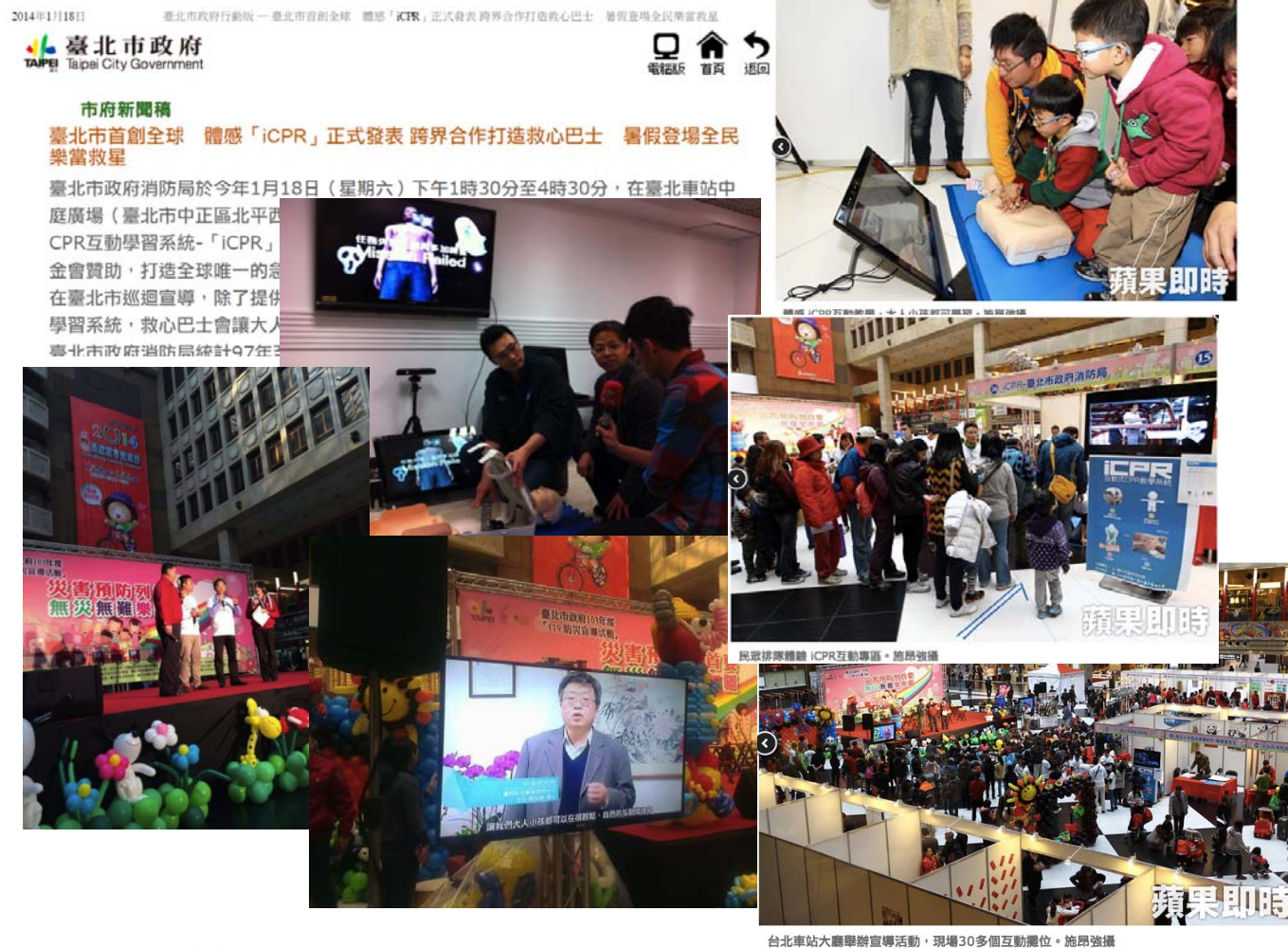


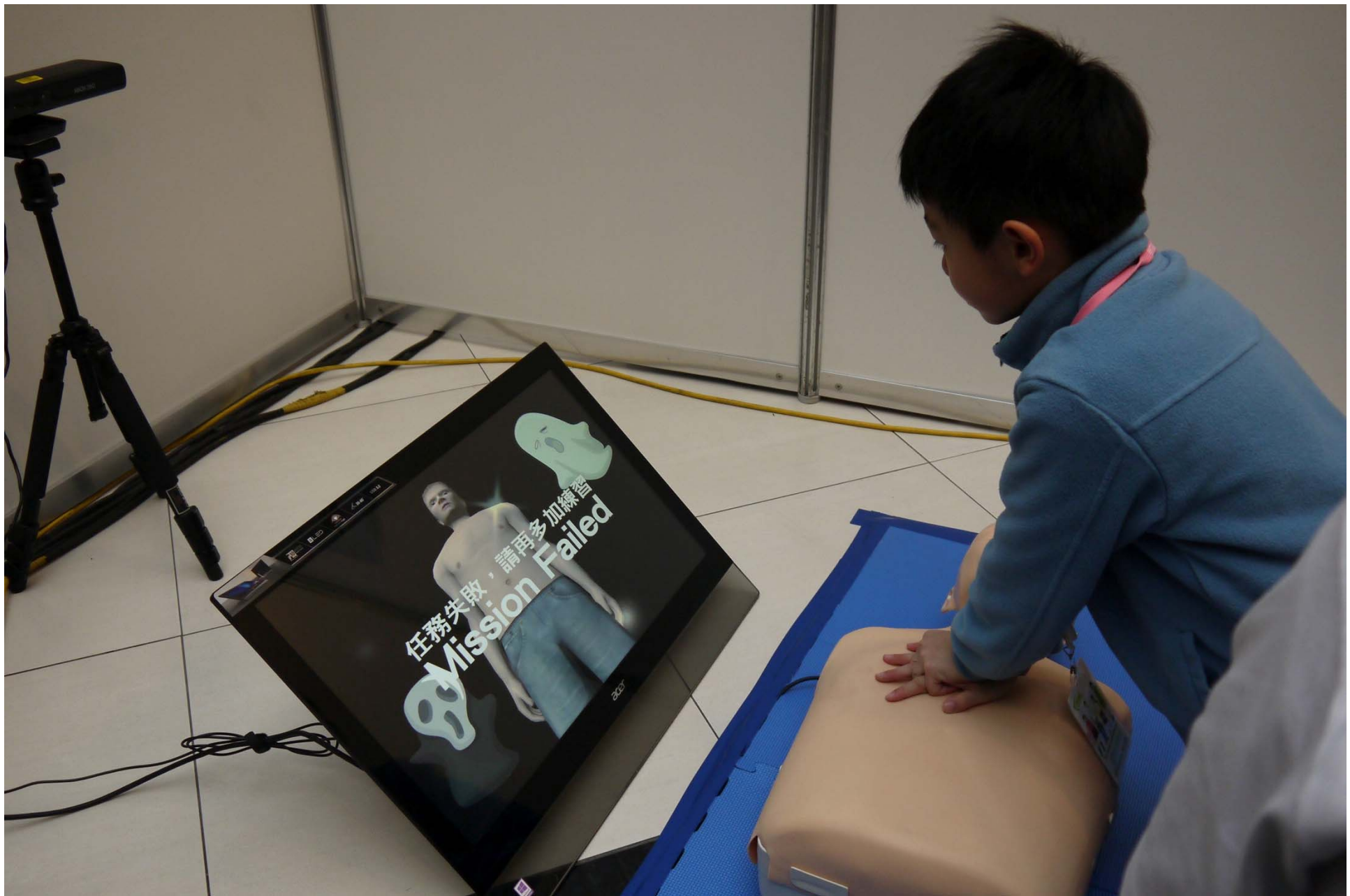
Professor H.T. Kung, Harvard University



Research Center for Information Technology Innovation

The first debut of the CPR training system in the Taipei Main Station on 18 January, 2014







HeartBus opening by the Taipei City Mayor in the Taipei City Hall on 24 July, 2014



National Disaster Prevention Day on 20 Sept, 2014



HeartBus in Taipei 101 on 15 Oct, 2014



Tokyo Fire Department coming to Visit HeartBus on 12 Dec, 2014



HeartBus in Smart City Expo on 20-23 Jan, 2015



HeartBus in Chinese Culture University on 18 Mar, 2015



HeartBus in DaHu Park on 21 Mar, 2015



HeartBus in Taiwan Presidential Palace on 29 Apr, 2015



Fire Safety Museum



Remarks

- Be creative for building your own multimedia applications!



Facial Expression Transfer



Face2Face: Real-time Face Capture and Reenactment of RGB Videos

*Justus Thies¹, Michael Zollhöfer²,
Marc Stamminger¹, Christian Theobalt²,
Matthias Nießner³*

¹University of Erlangen-Nuremberg

²Max-Planck-Institute for Informatics

³Stanford University

CVPR 2016 (Oral)

Remarks

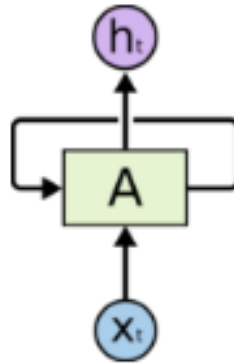
- Keep your eyes on the latest research trends!

Please get more information at
<http://mclab.citi.sinica.edu.tw/>

Thank you!

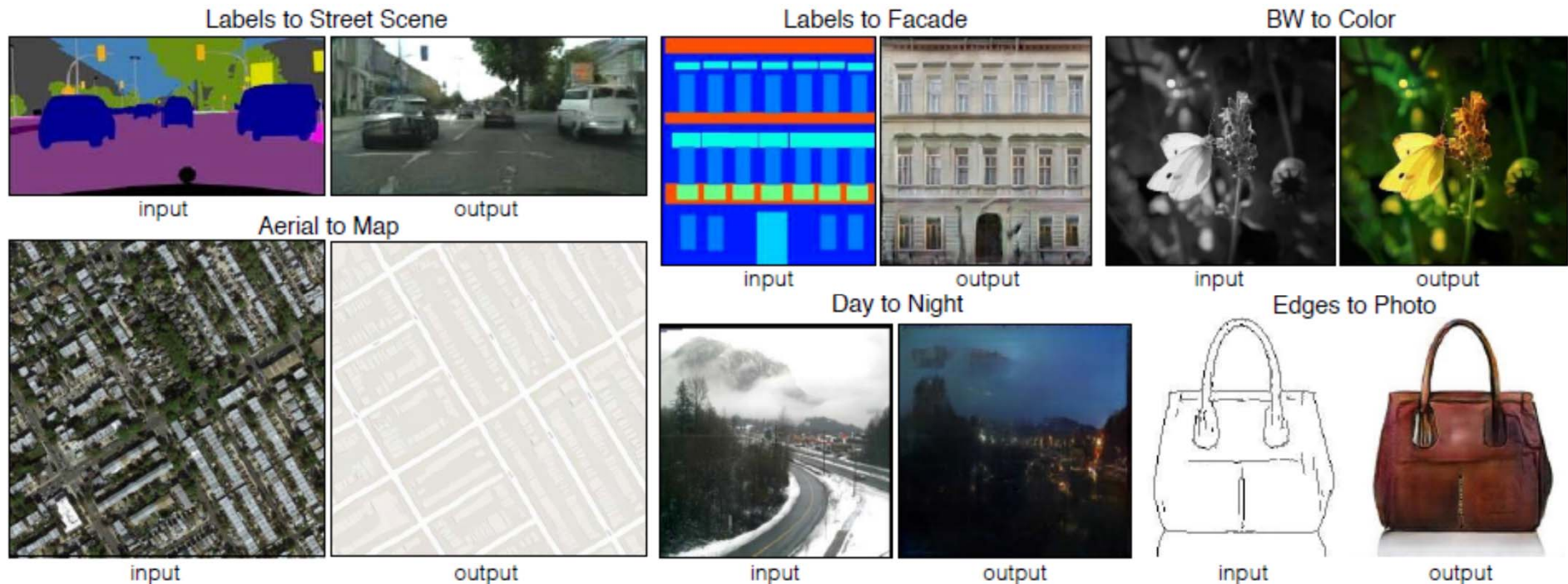


- Recurrent Neural Networks (RNN)



Recurrent Neural Networks have loops.

- GAN (Generative Adversarial Networks)



IEEE International Conference on Visual Communications and Image Processing

IEEE VCIP 2018

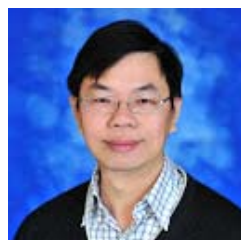
9 - 12 December 2018, Taichung, Taiwan

vcip2018.org

General Chairs



Wan-Chi Siu
Hong Kong Polytechnic University



Chia-Wen Lin
National Tsinghua University

Program Chairs



Wen-Huang Cheng
Academia Sinica



Gene Cheung
National Institute of Informatics



Research Center for Information Technology Innovation

ACM Multimedia 201

<http://www.acmmm.org/2017/>



Grand Challenge

Social Media Prediction (SMP):

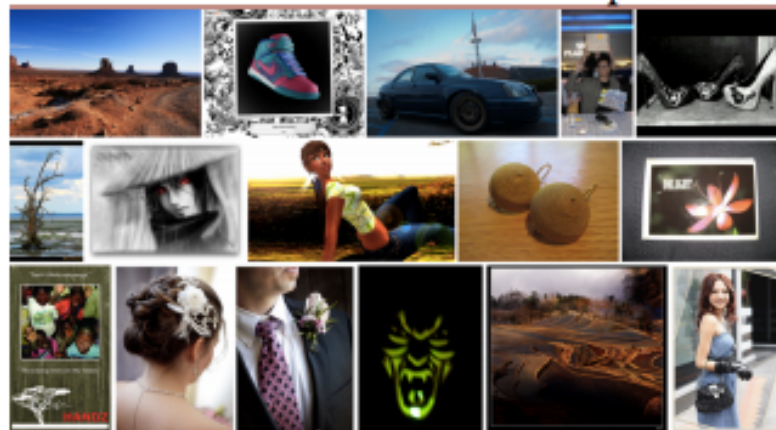
Predicting the “Stars of Tomorrow” on Social Media

Overview

People are interested in predicting the future. For example, which films will bomb or who will win the upcoming Grammy awards? Making predictions about the future in many aspects is not only fun matters but can bring real value to those who correctly predict the course of world events, such as which stocks are the best purchases for short-term gains. Predictive analytics is thus a field that has attracted major attention in both academia and the industry.

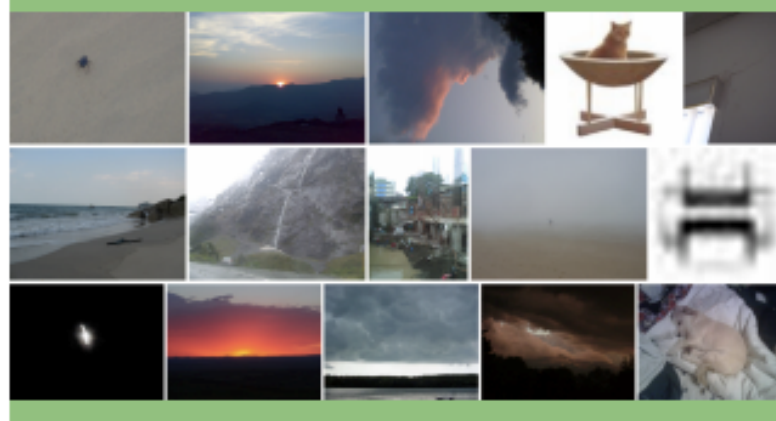


估測吸引力值
predicted popularity



ground truth popularity

真實吸
引力值



資料來源: "What Makes an Image Popular?" WWW, 2014.



Research Center for Information Technology Innovation

Let's exchange ideas!



Wen-Huang Cheng



wenhuangcheng



whcheng@citi.sinica.edu.tw

