高等電腦視覺

Introduction

黃正民
Cheng-Ming Huang

EE, TaipeiTech

Lecture Information

- Time
  - Thr. 1:10pm – 4:00pm
- Room
  - 先鋒402
- Website: 北科ｉ學園+
  - 上課講義、錄影
  - 作業繳交、成績公告

Instructor

- 黃正民 Cheng-Ming Huang
- E-mail

  cmhuang@ntut.edu.tw
- Office

  綜科館412B
- Phone

  (02)27712171 #2170
- Office Hours
  - by e-mail appointment

Syllabus

- Week 1. Introduction
- Week 2. Fundamental Knowledge
- Week 3. Thresholding, Segmentation, and Region Analysis
- Week 4. Mathematical Morphological Processing
- Week 5. Neighborhood Operator
- Week 6. Color Space and Transformations
- Week 7. Point Operators
- Week 8. Image Stitching
- Week 9. Contour Operators
- Week 10. Texture Analysis
- Week 11. Detection, Matching, and Tracking
- Week 12. Midterm Exam
- Week 13. Recognition and Classification
- Week 14. Calibration, Stereo Correspondence
- Week 15. 3D Reconstruction, SLAM
- Week 16. Visual Servoing
- Week 17. Presentation of Final Project
- Week 18. Presentation of Final Project

Syllabus

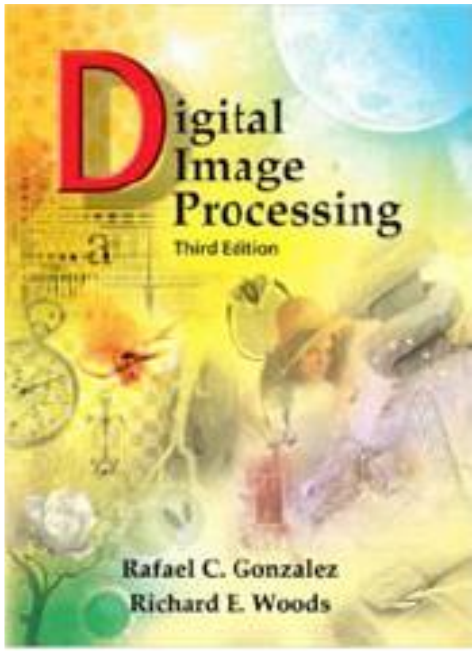
- Grading
  - Homework (with C/OpenCV programming) 70%
  - Midterm Exam (closed book) 10%
  - Final Project 20%
- Textbook
  - Richard Szeliski, Computer Vision: Algorithms and Applications , 2nd ed, 2020. (http://szeliski.org/Book/)
- Reference
  - D. A. Forsyth and John Ponce, Computer Vision: A Modern Approach, Pearson Education, 2003.
  - Richard Hartley and Andrew Zisserman, Multiple View Geometry in Computer Vision, 2nd ed, Cambridge University Press, March 2004.

Slide Credits

- Richard Szeliski, at University of Washington
- Trevor Darrell, at Berkeley
- Antonio Torralba, at MIT
- Michael Black, at Brown
- Kristen Grauman, at UT Austin
- Chiou-Shann Fuh, at NTU
- Yung-Yu Chuang, at NTU

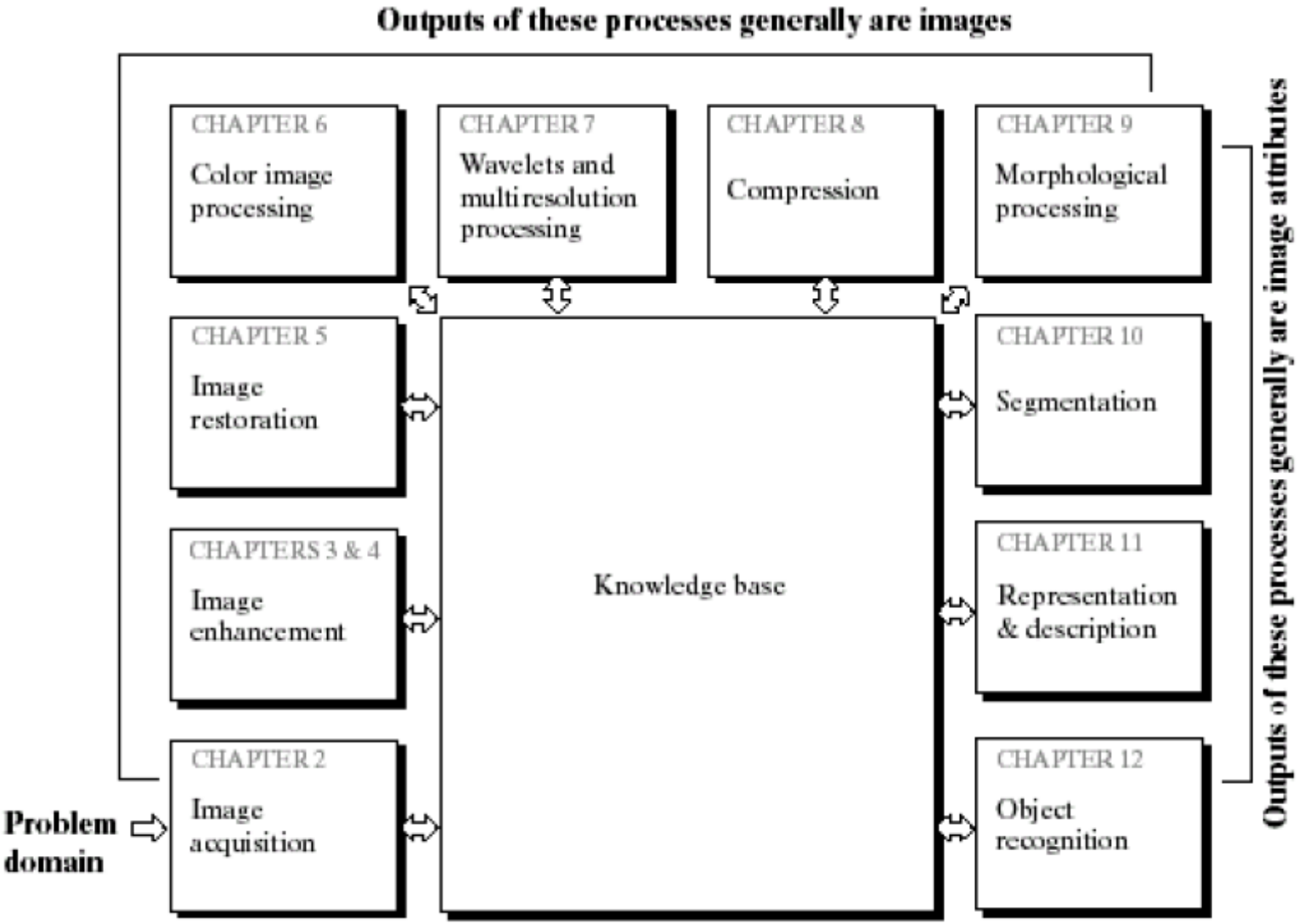
Related Courses

- Digital image processing
- Computer vision
- Digital video technology
- Information theory
- Pattern recognition
- Artificial intelligence
- Robotics



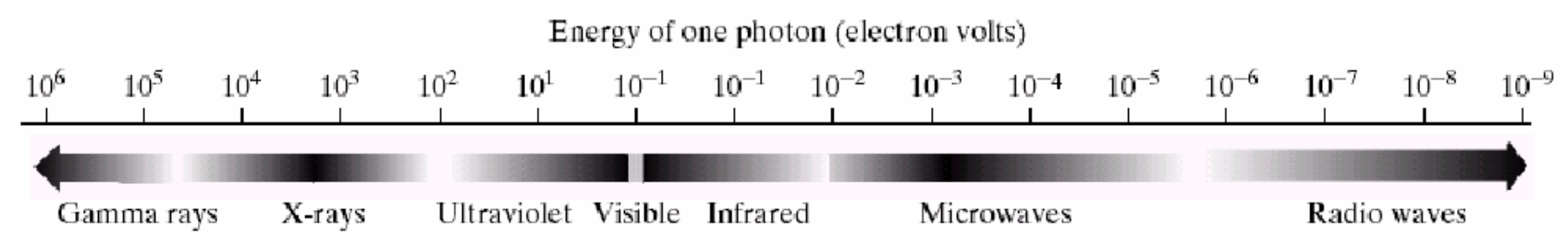
Gonzale, R. C. and R.E. Woods. 2008. "Digital Image Processing.", 3rd Ed., Prentice Hall, Inc.Upper Saddle River, New Jersey

Related Topics (Digital Image Processing)



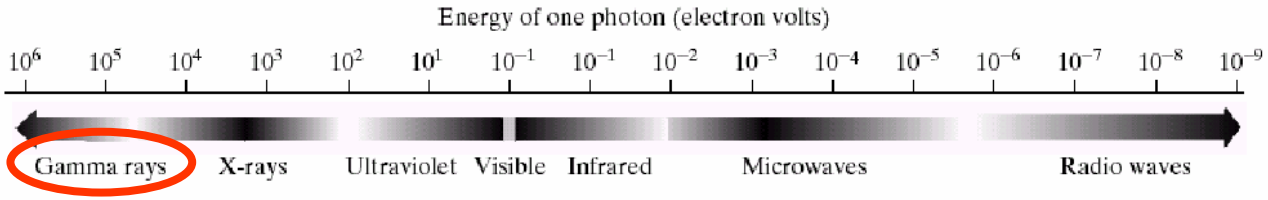
Other visual sensor

- Electromagnetic spectrum arranged according to energy per photon



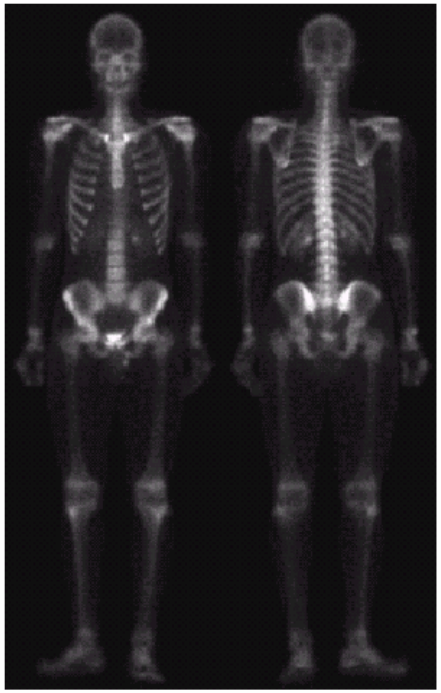
- Imaging machines cover almost the entire EM spectrum

Gamma Rays



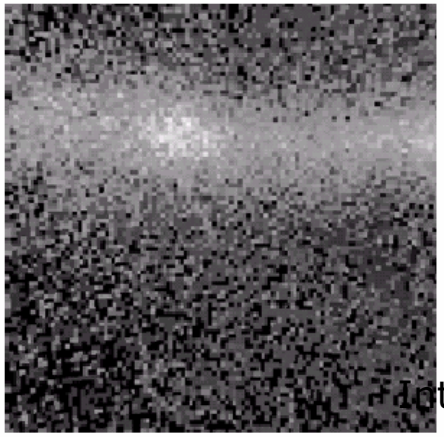
- 醫學、天文

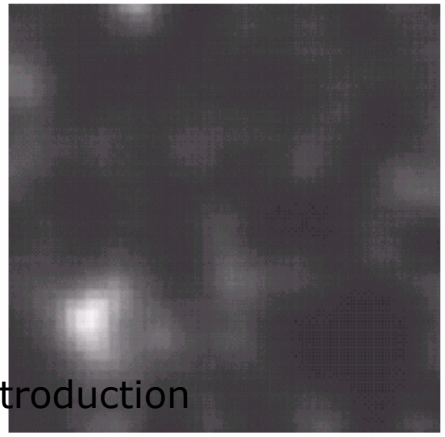
骨骼掃瞄



Positron Emission Tomography
正子電腦斷層造影術

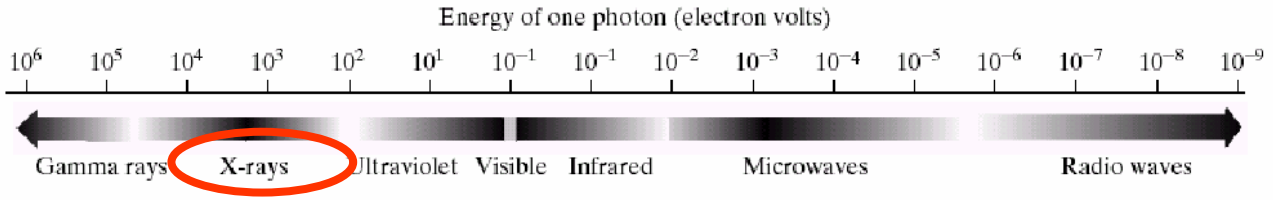
天鵝座環狀星雲

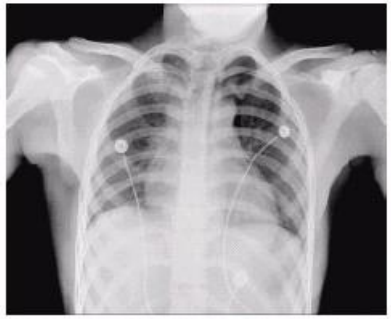


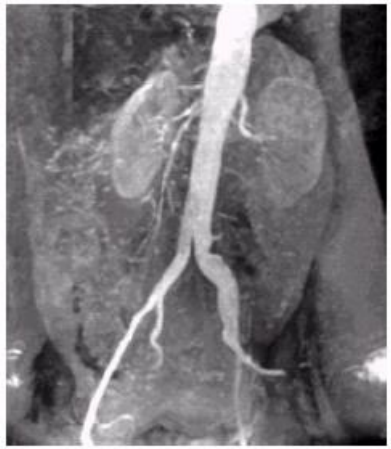


核子反應爐閥
Gamma 輻射

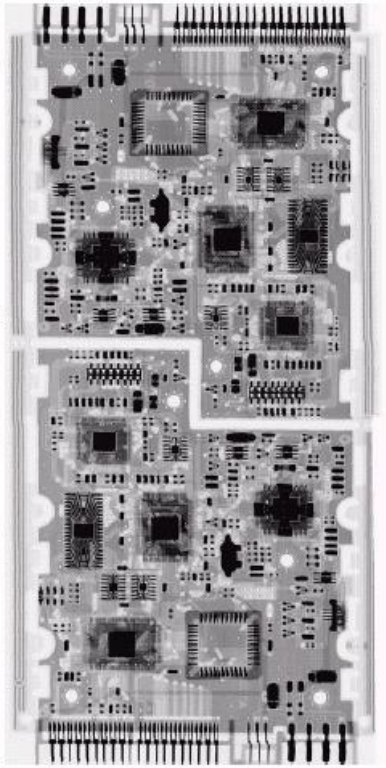
X-rays

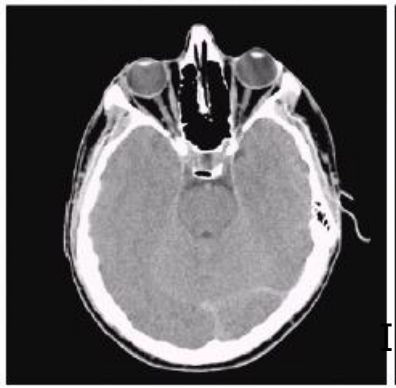




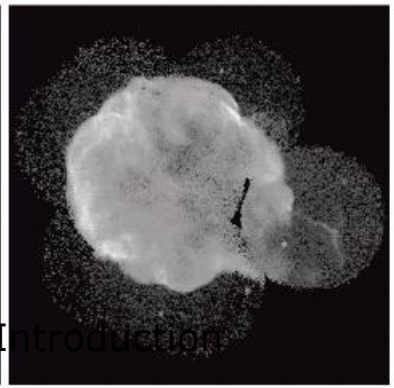


血管造影術



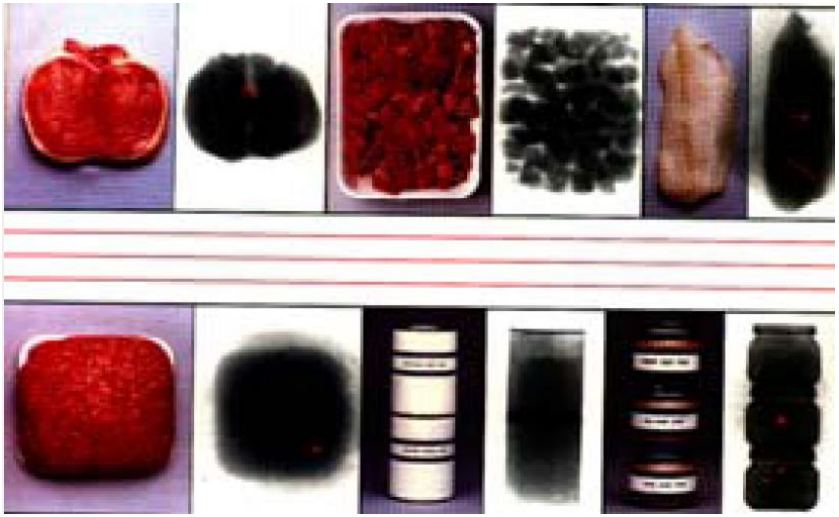


Computerized Tomography



天鵝座環狀星雲

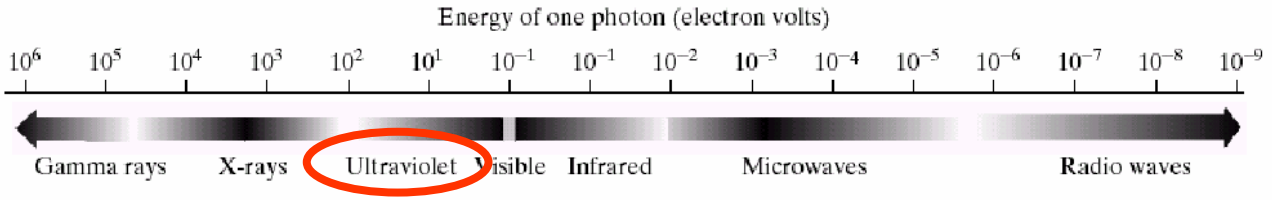
X-rays



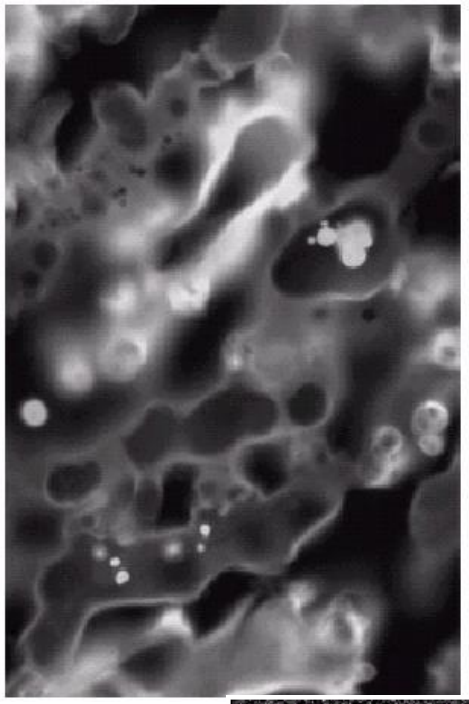
食品安全檢測

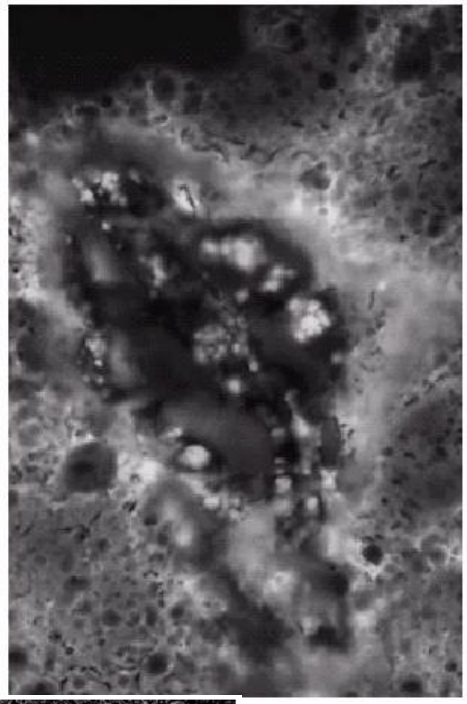
機場行李安全檢測

Ultraviolet

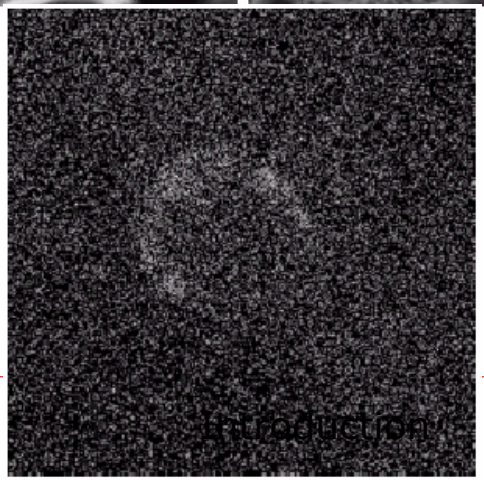


螢光顯微技術
—正常玉米



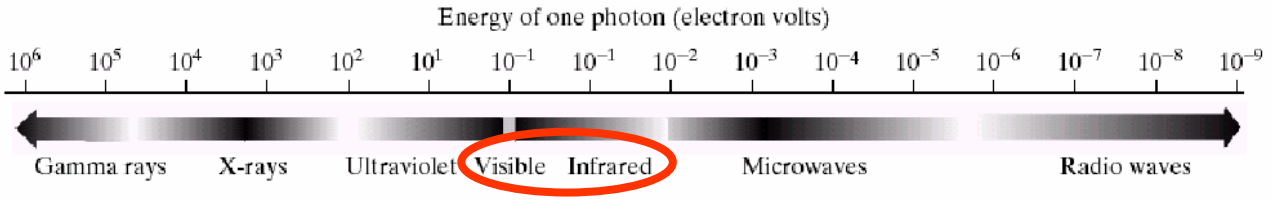


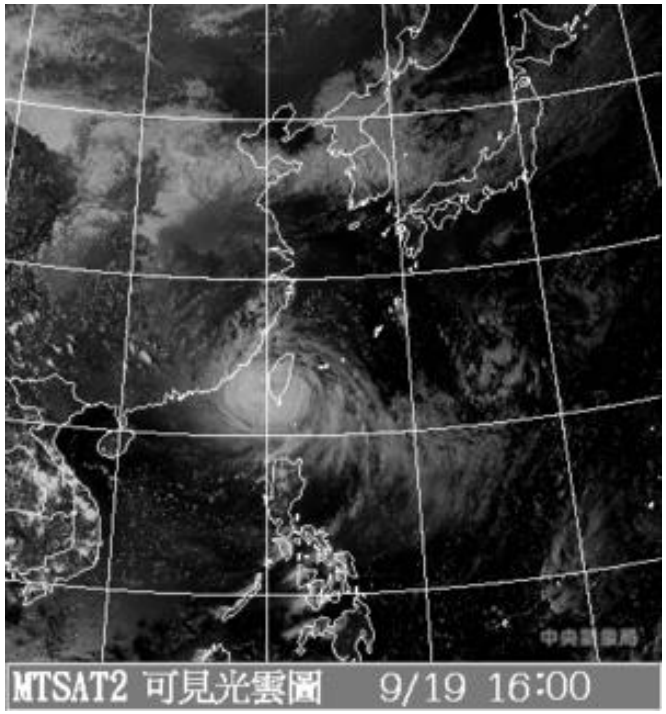
螢光顯微技術
—染病玉米



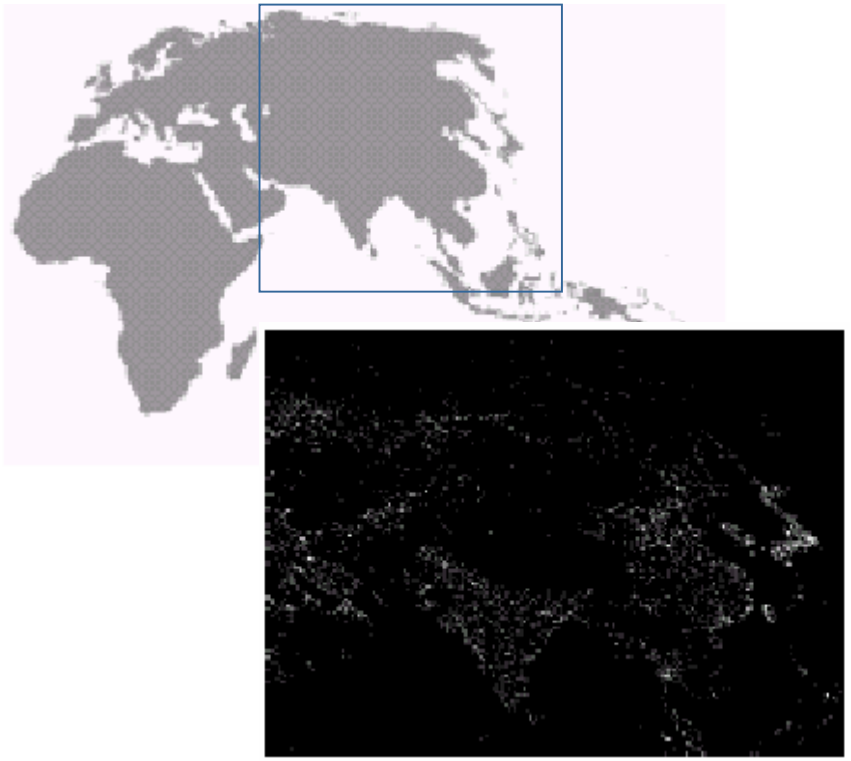
天鵝座環狀星雲

Visible, Infrared



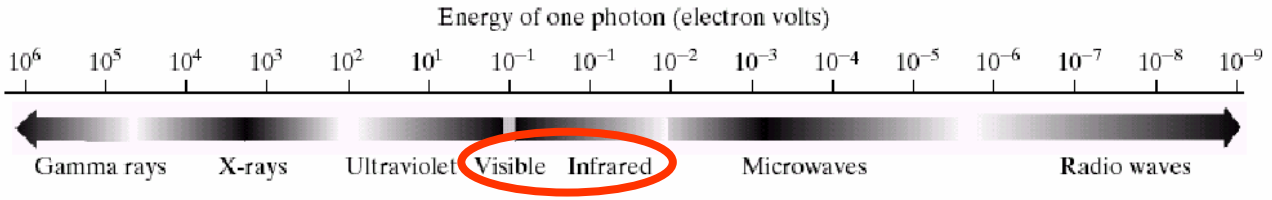


multispectral image



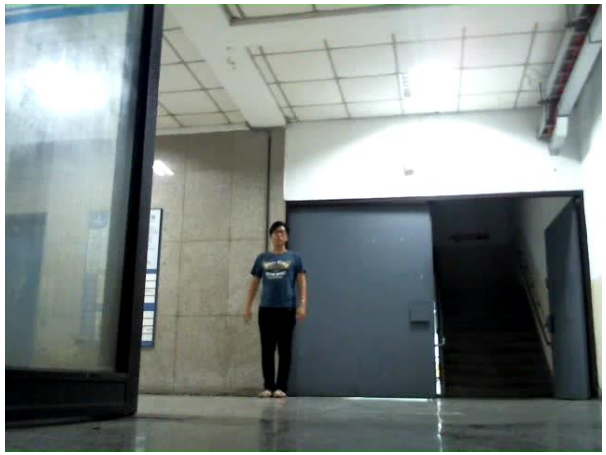
infrared satellite image

Visible, Infrared

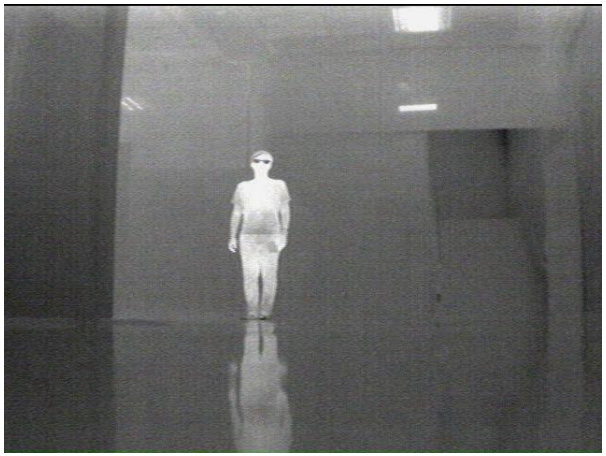


- Thermal camera

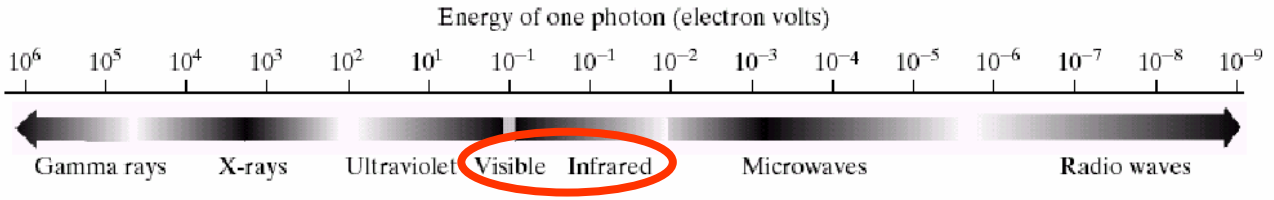
Color image



Thermal image

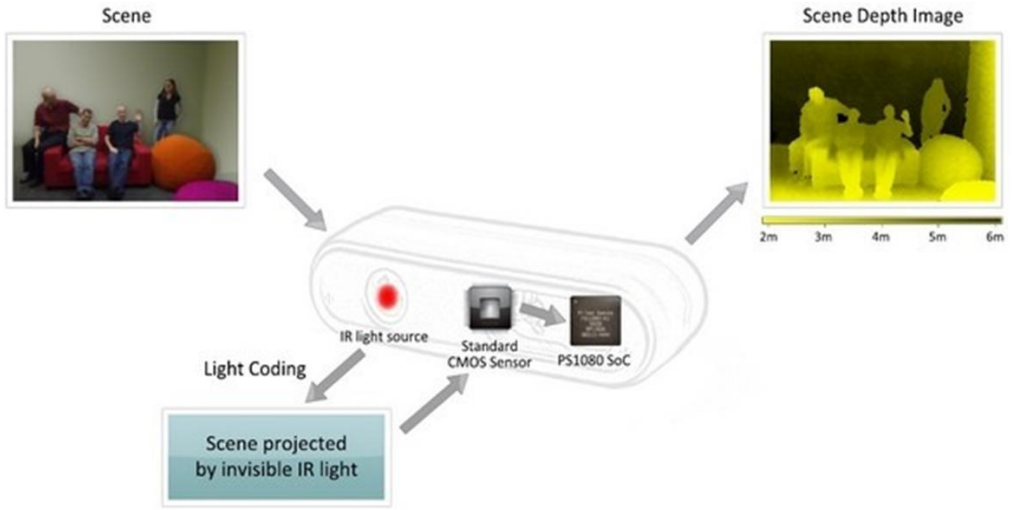


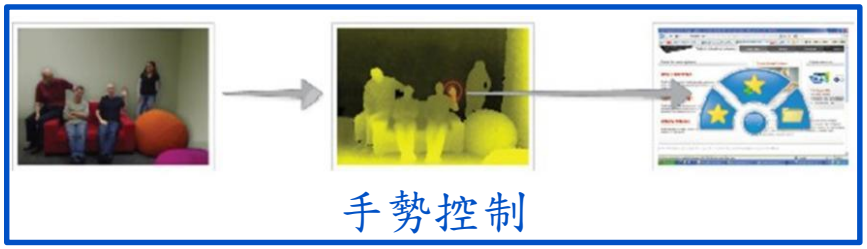
Visible, Infrared

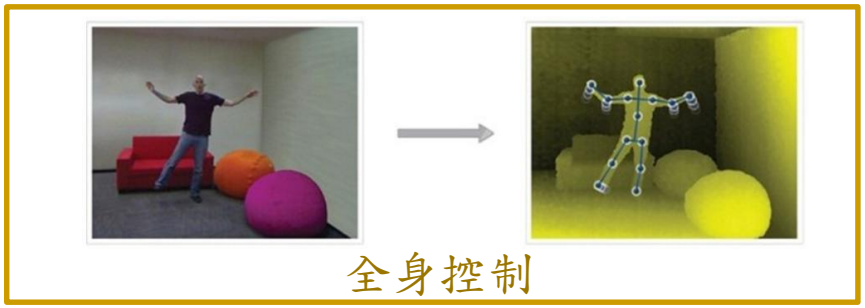


- PrimeSense PrimeSensor
- Microsoft Kinect
- Asus Xtion
  - IR 光源, IR CMOS 接收器, CMOS camera

http://pointclouds.org/documentation/tutorials/openni_grabber.php

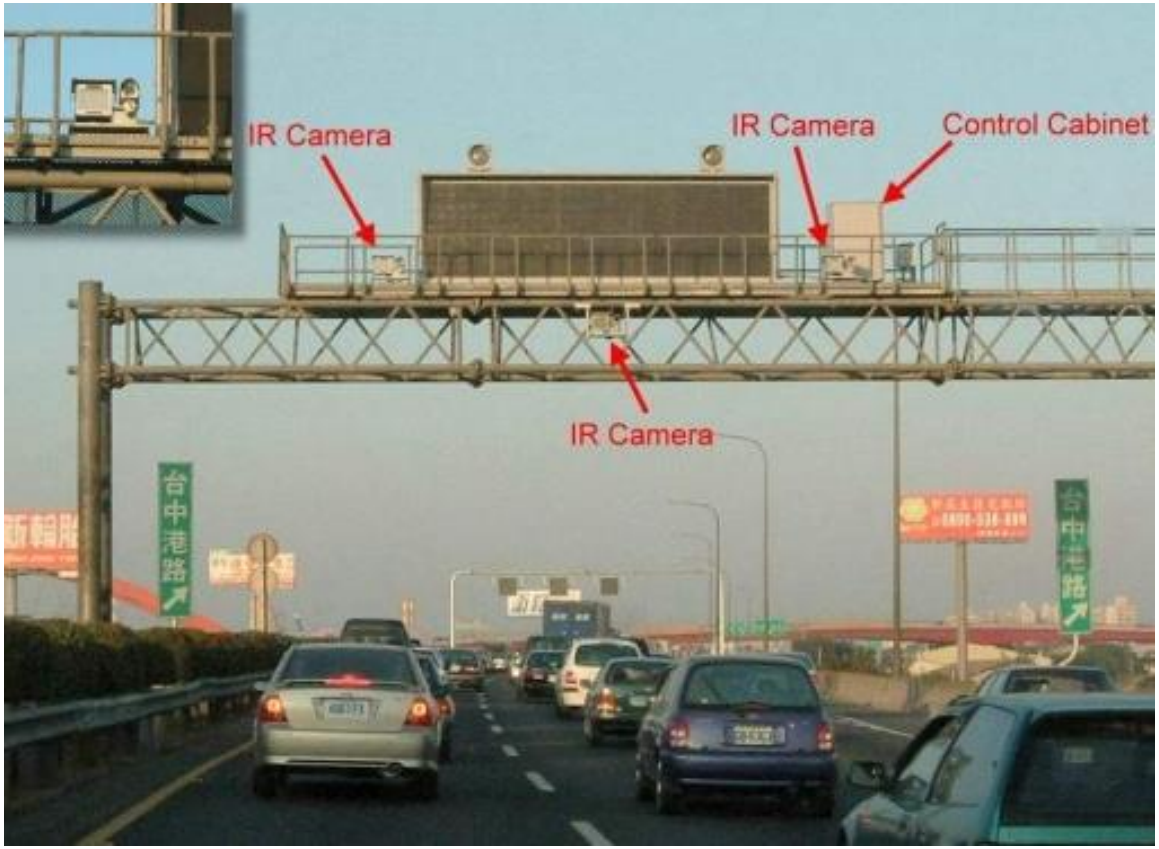




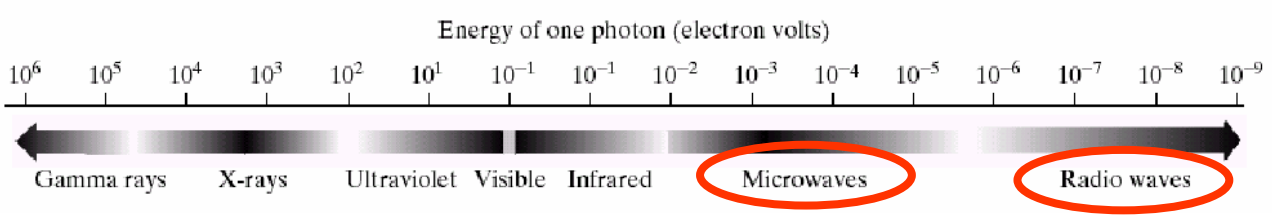


Visible, Infrared

Infrared illuminator

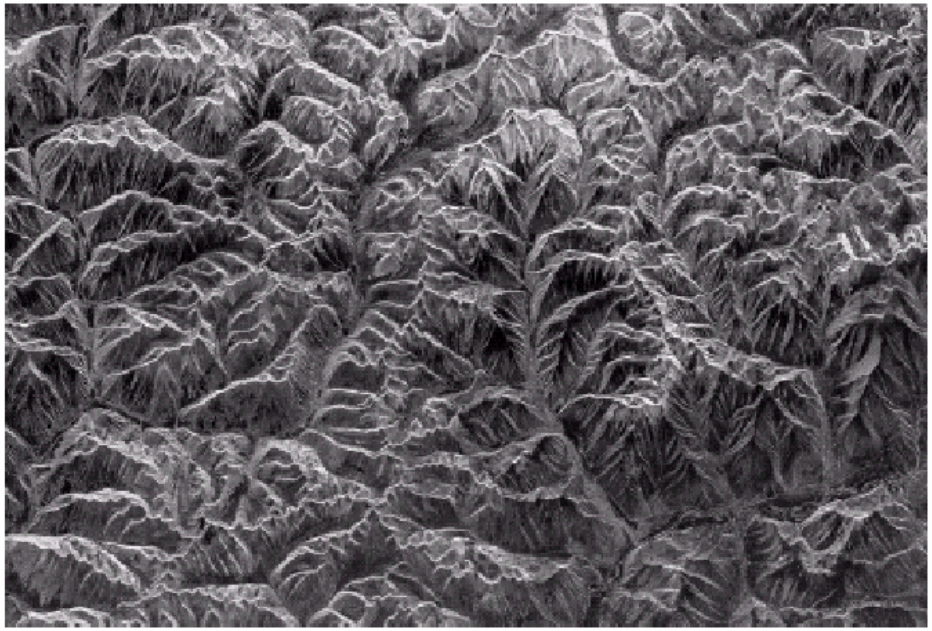


http://www.komoto.com.tw

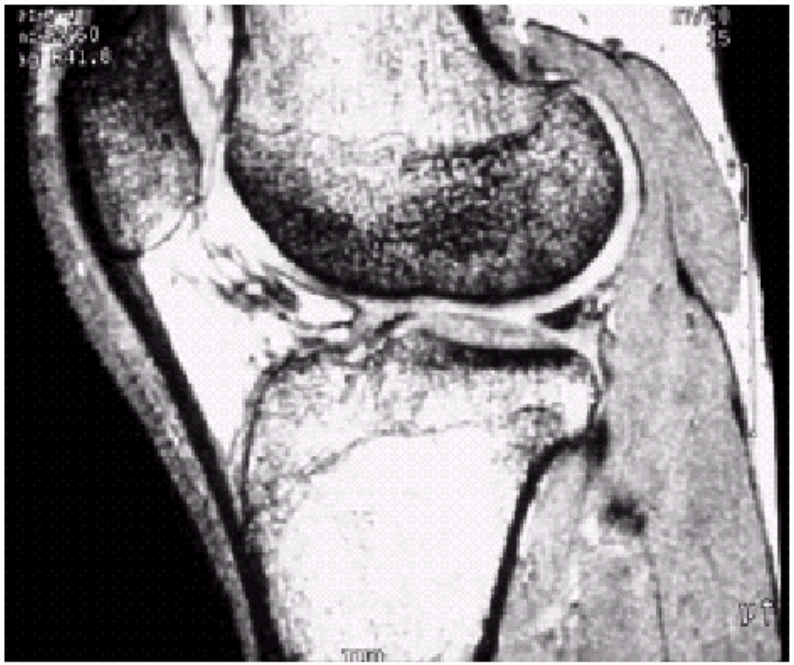


- Microwaves

- Radio waves



微波雷達影像

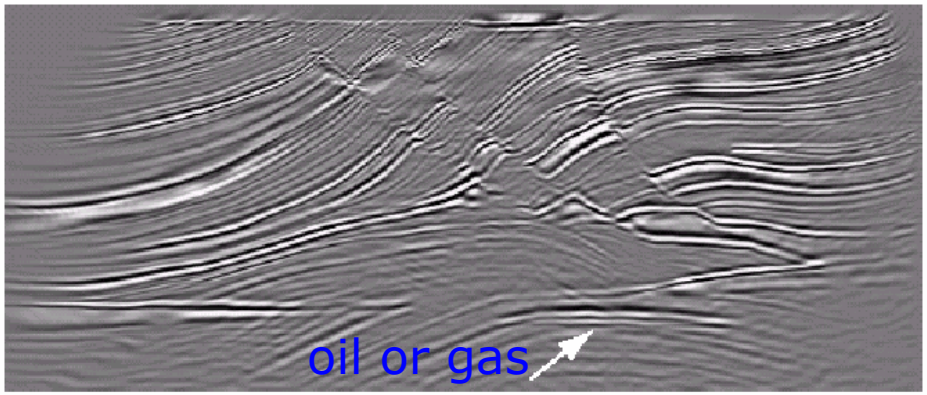


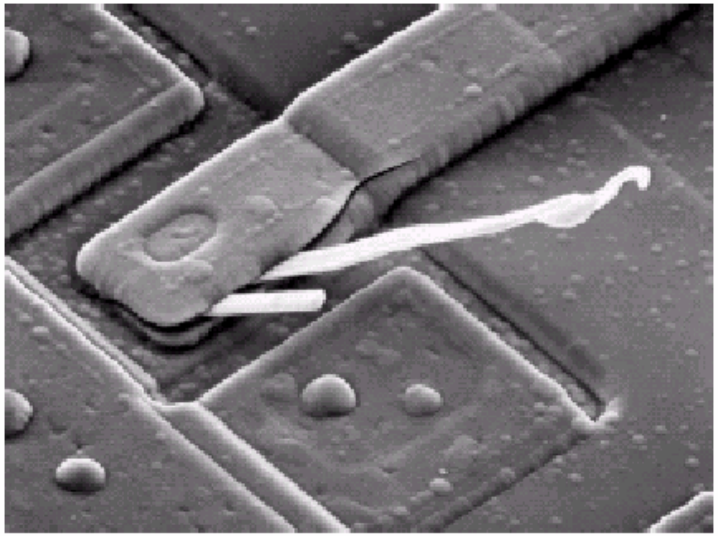
Magnetic Resonance Imaging
磁共振造影

Other Imaging Modalities (other than EM)

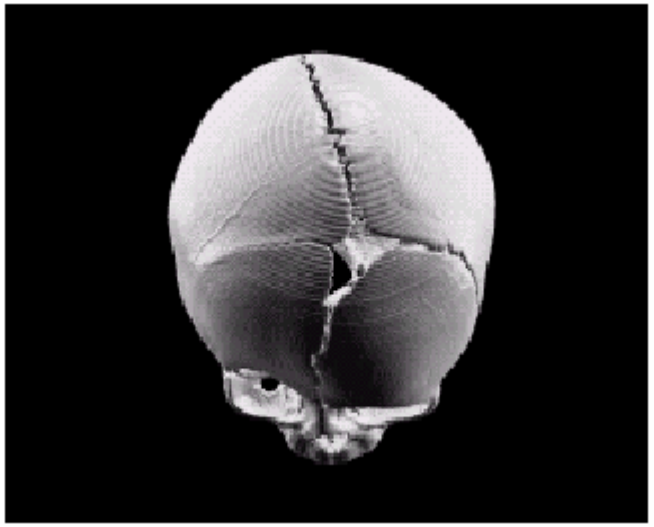
超音波影像 Ultrasound Imaging

震測影像 Seismic Imaging





掃描式電子顯微鏡 (2500x)
Scanning Electronic Microscope (SEM)



Computer Generated Images

Rough Timeline: 1970

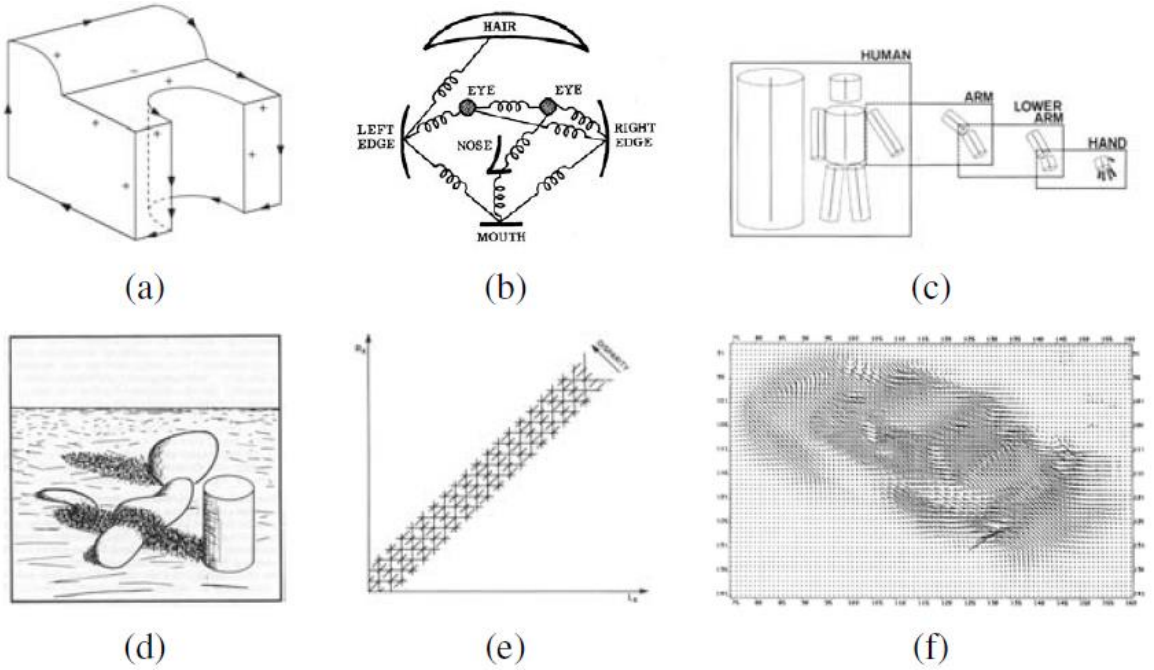
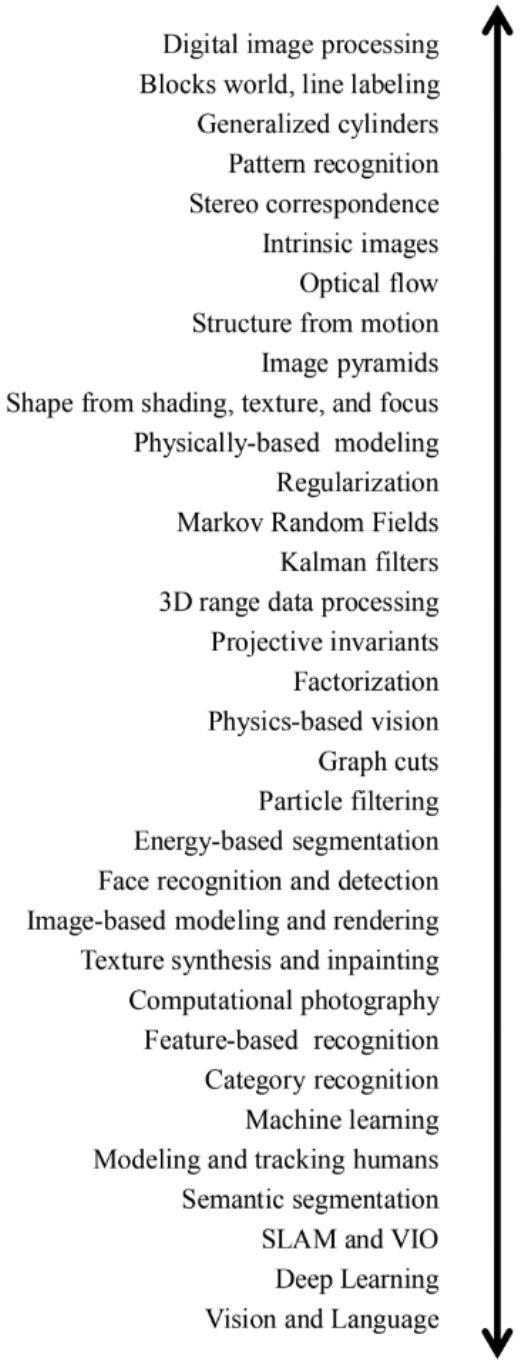


Figure 1.7 Some early (1970s) examples of computer vision algorithms: (a) line labeling (Nalwa 1993) © 1993 Addison-Wesley, (b) pictorial structures (Fischler and Elschlager 1973) © 1973 IEEE, (c) articulated body model (Marr 1982) © 1982 David Marr, (d) intrinsic images (Barrow and Tenenbaum 1981) © 1973 IEEE, (e) stereo correspondence (Marr 1982) © 1982 David Marr, (f) optical flow (Nagel and Enkelmann 1986) © 1986 IEEE.



Rough Timeline: 1980

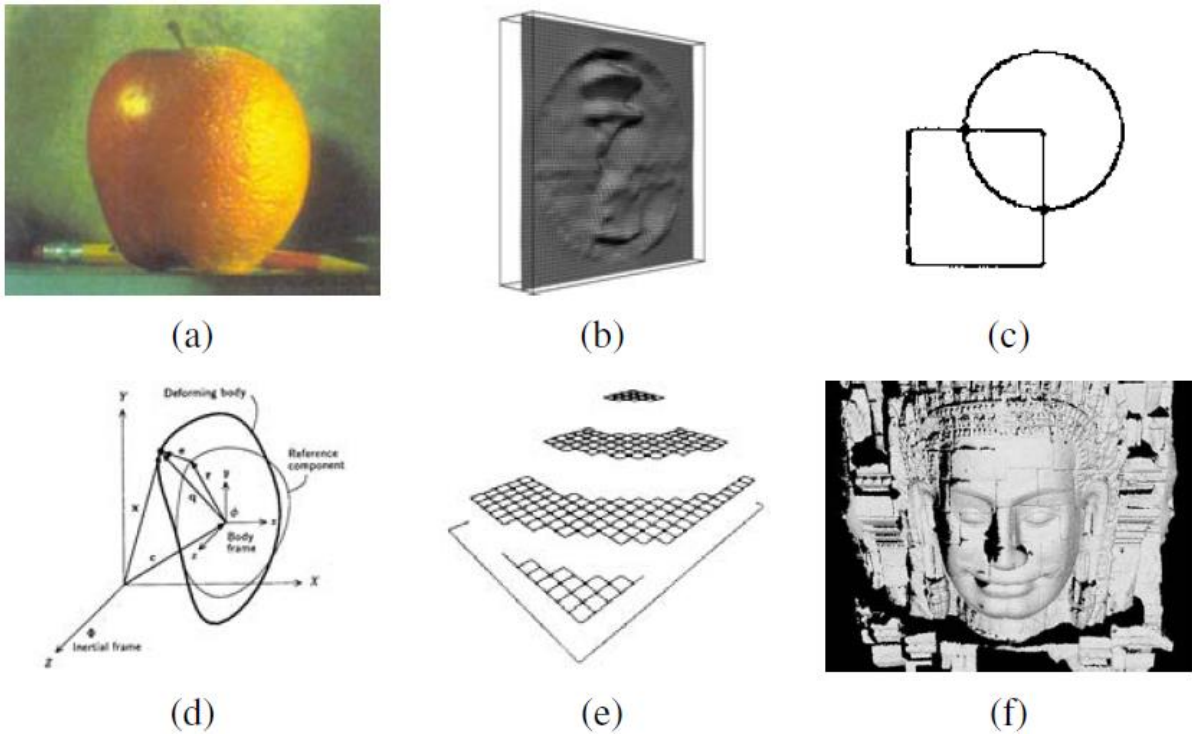


Figure 1.8 Examples of computer vision algorithms from the 1980s: (a) pyramid blending (Burt and Adelson 1983b) © 1983 ACM, (b) shape from shading (Freeman and Adelson 1991) © 1991 IEEE, (c) edge detection (Freeman and Adelson 1991) © 1991 IEEE, (d) physically based models (Terzopoulos and Witkin 1988) © 1988 IEEE, (e) regularization-based surface reconstruction (Terzopoulos 1988) © 1988 IEEE, (f) range data acquisition and merging (Banno, Masuda, Oishi *et al.* 2008) © 2008 Springer.

1970
Digital image processing
Blocks world, line labeling
Generalized cylinders
Pattern recognition

1980
Stereo correspondence
Intrinsic images
Optical flow
Structure from motion
Image pyramids
Shape from shading, texture, and focus
Physically-based modeling
Regularization
Markov Random Fields

1990
Kalman filters
3D range data processing
Projective invariants
Factorization
Physics-based vision
Graph cuts
Particle filtering

2000
Energy-based segmentation
Face recognition and detection
Image-based modeling and rendering
Texture synthesis and inpainting
Computational photography

2010
Feature-based recognition
Category recognition
Machine learning
Modeling and tracking humans
Semantic segmentation
SLAM and VIO
Deep Learning

2020
Vision and Language

Rough Timeline: 1990

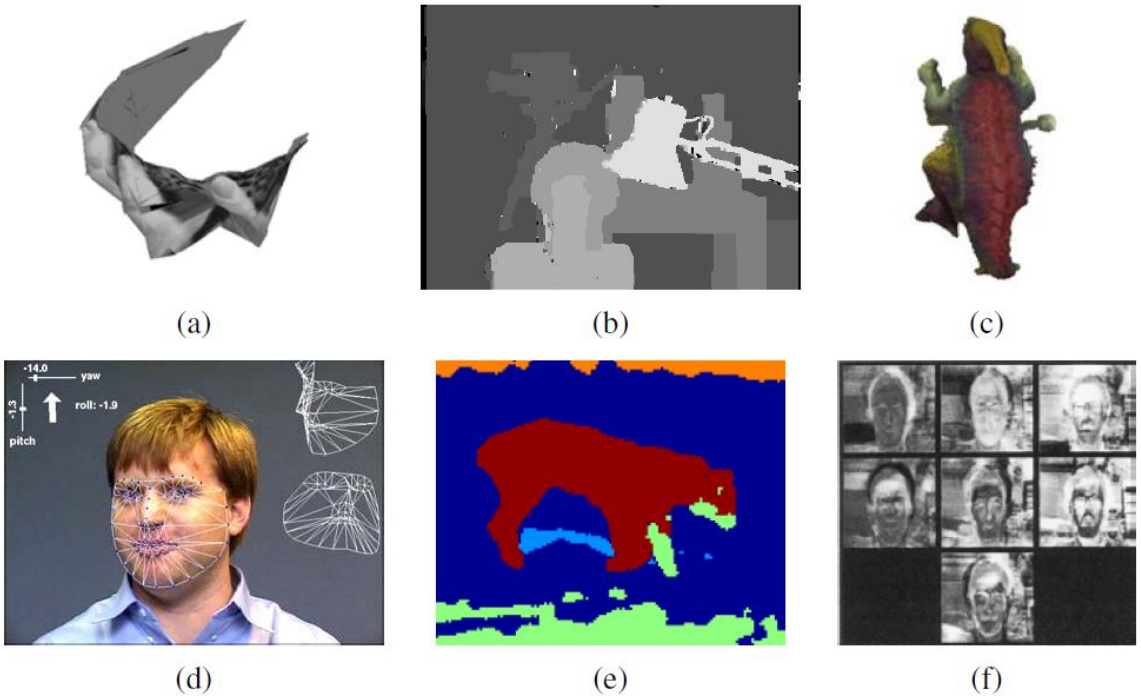
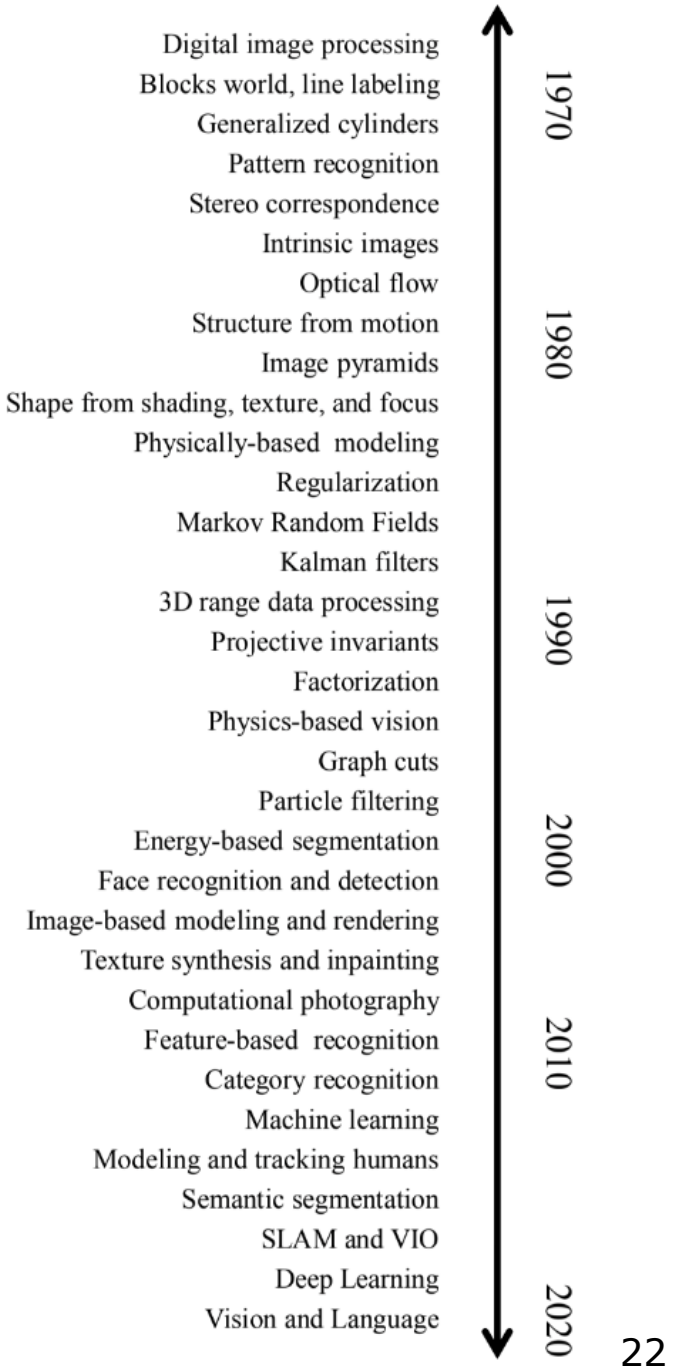
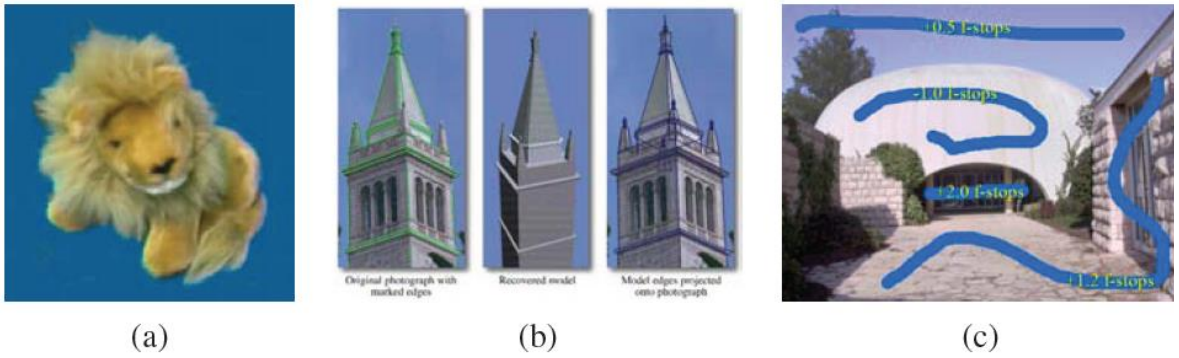


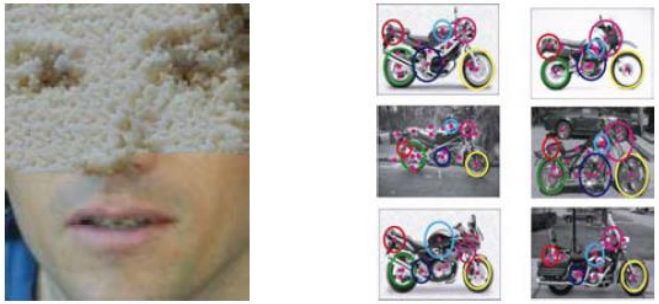
Figure 1.9 Examples of computer vision algorithms from the 1990s: (a) factorization-based structure from motion (Tomasi and Kanade 1992) © 1992 Springer, (b) dense stereo matching (Boykov, Veksler, and Zabih 2001), (c) multi-view reconstruction (Seitz and Dyer 1999) © 1999 Springer, (d) face tracking (Matthews, Xiao, and Baker 2007), (e) image segmentation (Belongie, Fowlkes, Chung *et al.* 2002) © 2002 Springer, (f) face recognition (Turk and Pentland 1991a).



Rough Timeline: 2000

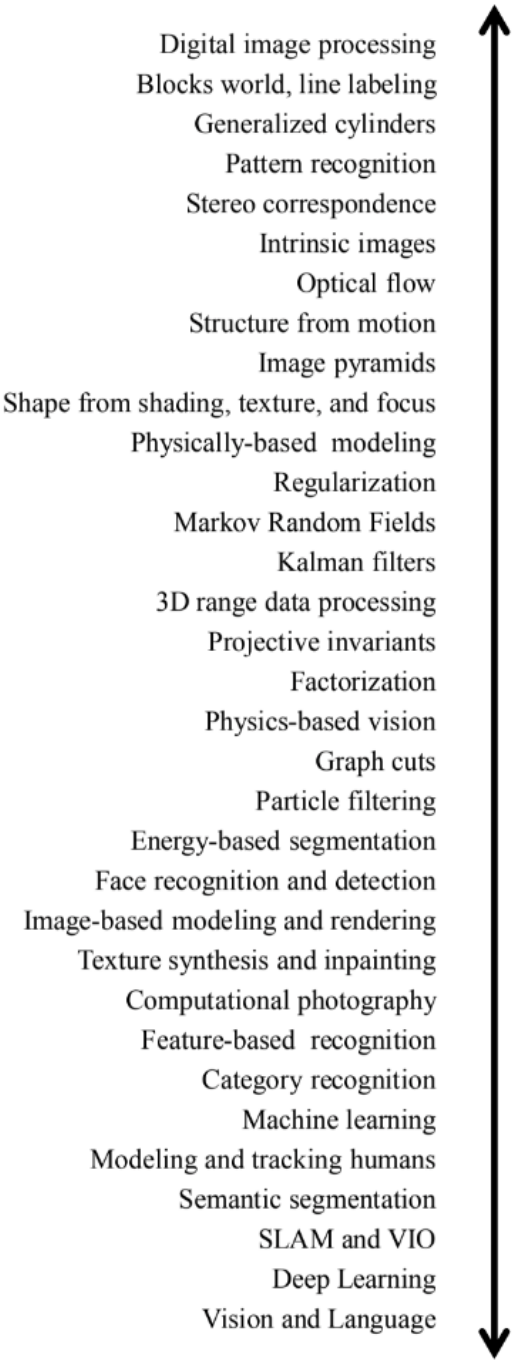


(a) (b) (c)



(d) (e) (f)

Figure 1.10 Recent examples of computer vision algorithms: (a) image-based rendering (Gortler, Grzeszczuk, Szeliski *et al.* 1996), (b) image-based modeling (Debevec, Taylor, and Malik 1996) © 1996 ACM, (c) interactive tone mapping (Lischinski, Farbman, Uyttendaele *et al.* 2006a) (d) texture synthesis (Efros and Freeman 2001), (e) feature-based recognition (Fergus, Perona, and Zisserman 2007), (f) region-based recognition (Mori, Ren, Efros *et al.* 2004) © 2004 IEEE.



Rough Timeline: 2010

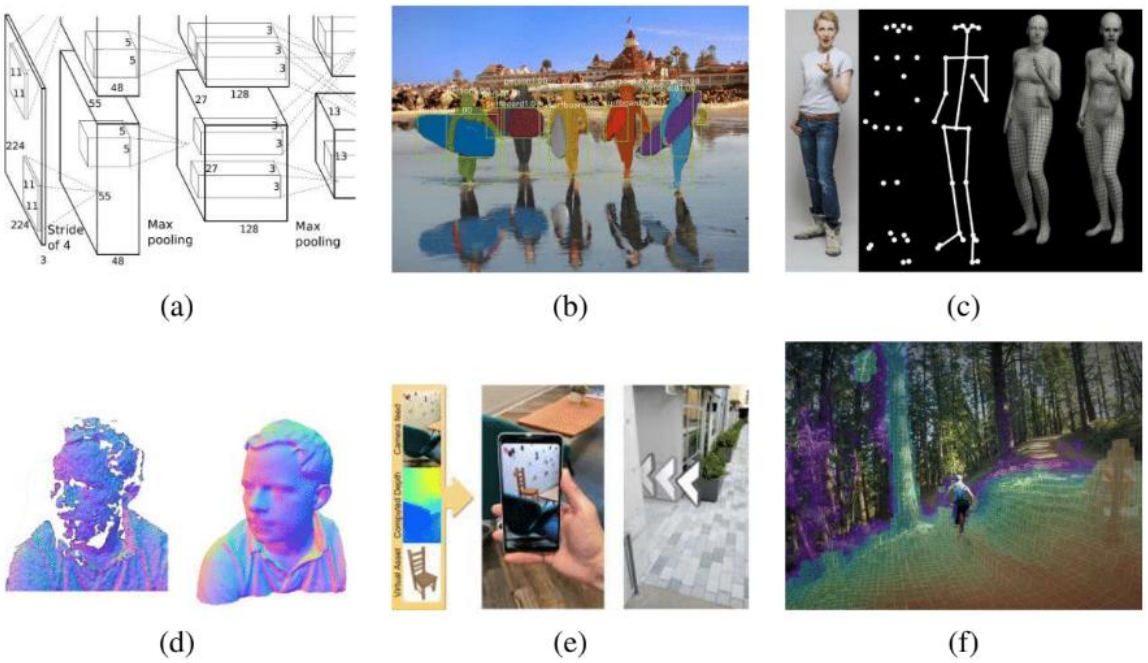


Figure 1.11 *Examples of computer vision algorithms from the 2010s: (a) the SuperVision deep neural network © Krizhevsky, Sutskever, and Hinton (2012); (b) object instance segmentation (He, Gkioxari* et al. *2017) © 2017 IEEE; (c) whole body, expression, and gesture fitting from a single image (Pavlakos, Choutas* et al. *2019) © 2019 IEEE; (d) fusing multiple color depth images using the KinectFusion real-time system (Newcombe, Izadi* et al. *2011) © 2011 ACM; (e) smartphone augmented reality with real-time depth occlusion effects (Valentin, Kowdle* et al. *2018) © 2018 ACM; (f) 3D map computed in real-time on a fully autonomous drone (Cross 2019)* [TODO: copyright]

1970
- Digital image processing
- Blocks world, line labeling
- Generalized cylinders
- Pattern recognition

1980
- Stereo correspondence
- Intrinsic images
- Optical flow
- Structure from motion
- Image pyramids
- Shape from shading, texture, and focus
- Physically-based modeling
- Regularization
- Markov Random Fields

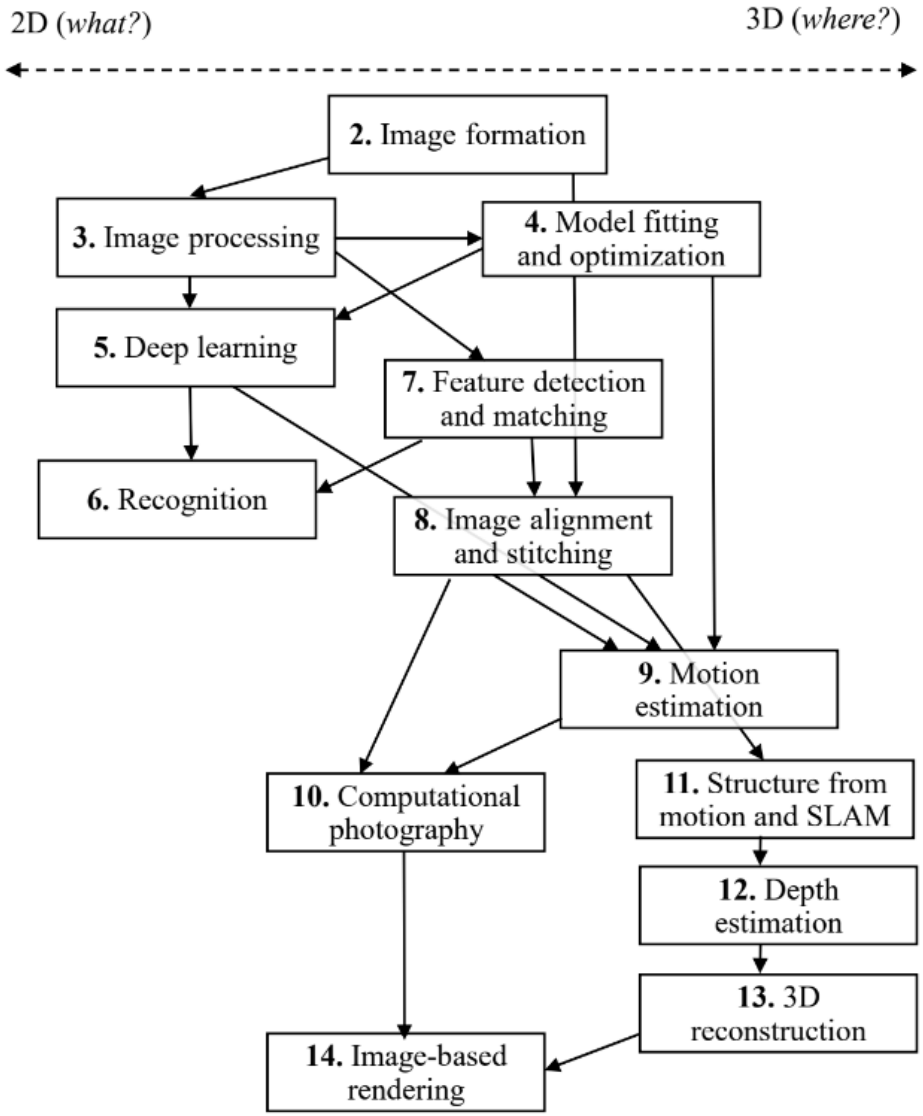
1990
- Kalman filters
- 3D range data processing
- Projective invariants
- Factorization
- Physics-based vision
- Graph cuts

2000
- Particle filtering
- Energy-based segmentation
- Face recognition and detection
- Image-based modeling and rendering
- Texture synthesis and inpainting
- Computational photography

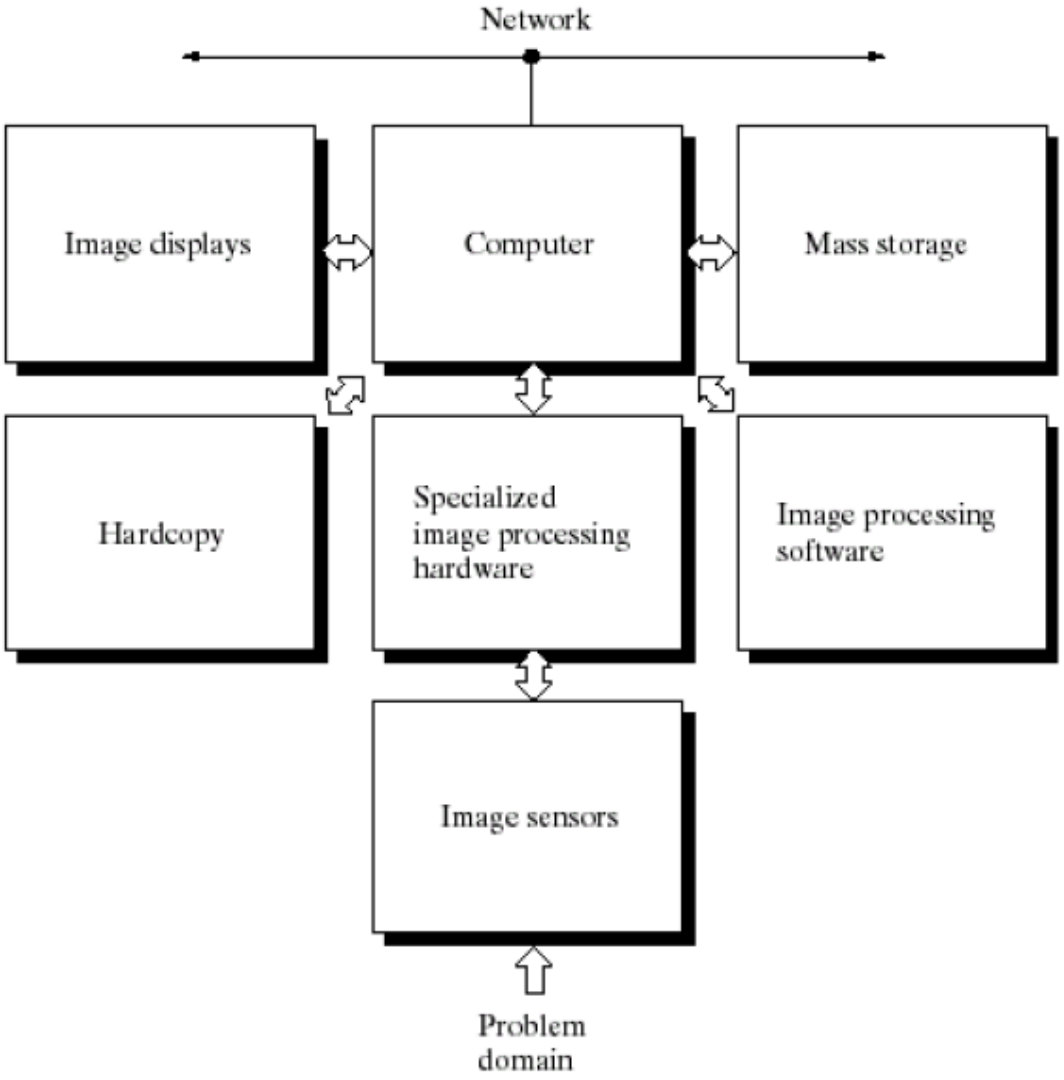
2010
- Feature-based recognition
- Category recognition
- Machine learning
- Modeling and tracking humans
- Semantic segmentation
- SLAM and VIO
- Deep Learning
- Vision and Language

2020

Related Topics (Computer Vision)



Components of Vision System



Components of Vision System

- Acquisition
- Processing
- Storage
- Display
- Accessory
  - Illumination, transmission, etc.
- Motion actuator
- Control interface

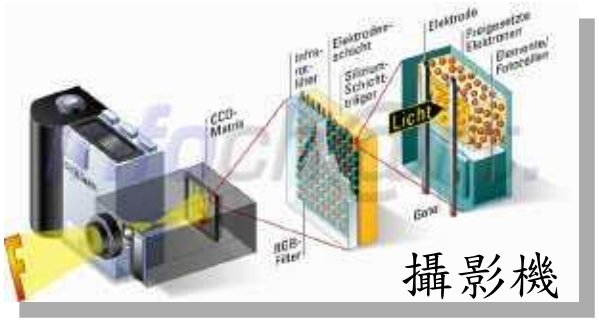
可動機構

傳輸儲存

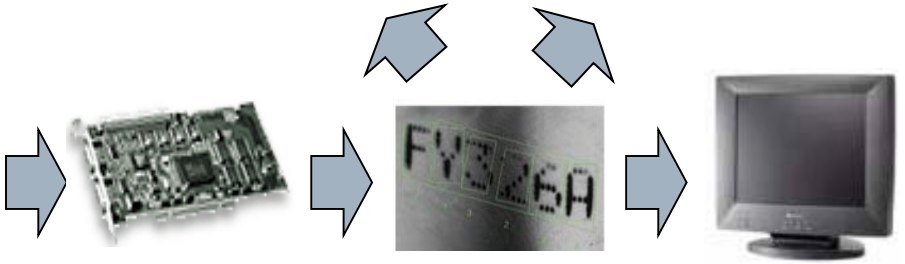
照明

F

物體



攝影機



擷取

處理

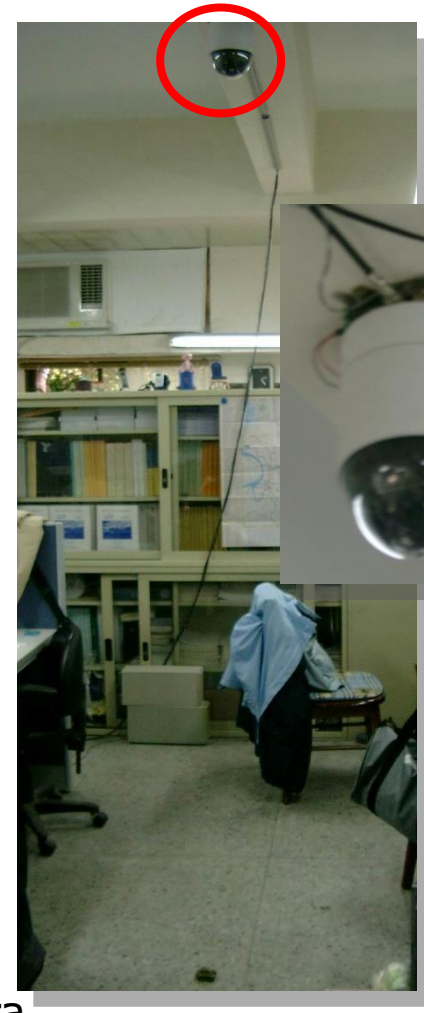
顯示

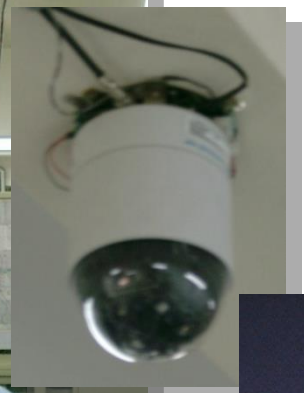
Camera

- 2 degree-of-freedom moving platform, Pan-Tilt-Zoom (PTZ), Dome

DirectedPerception

Dynacolor 彩富科技





Logitech

webcam

Sony

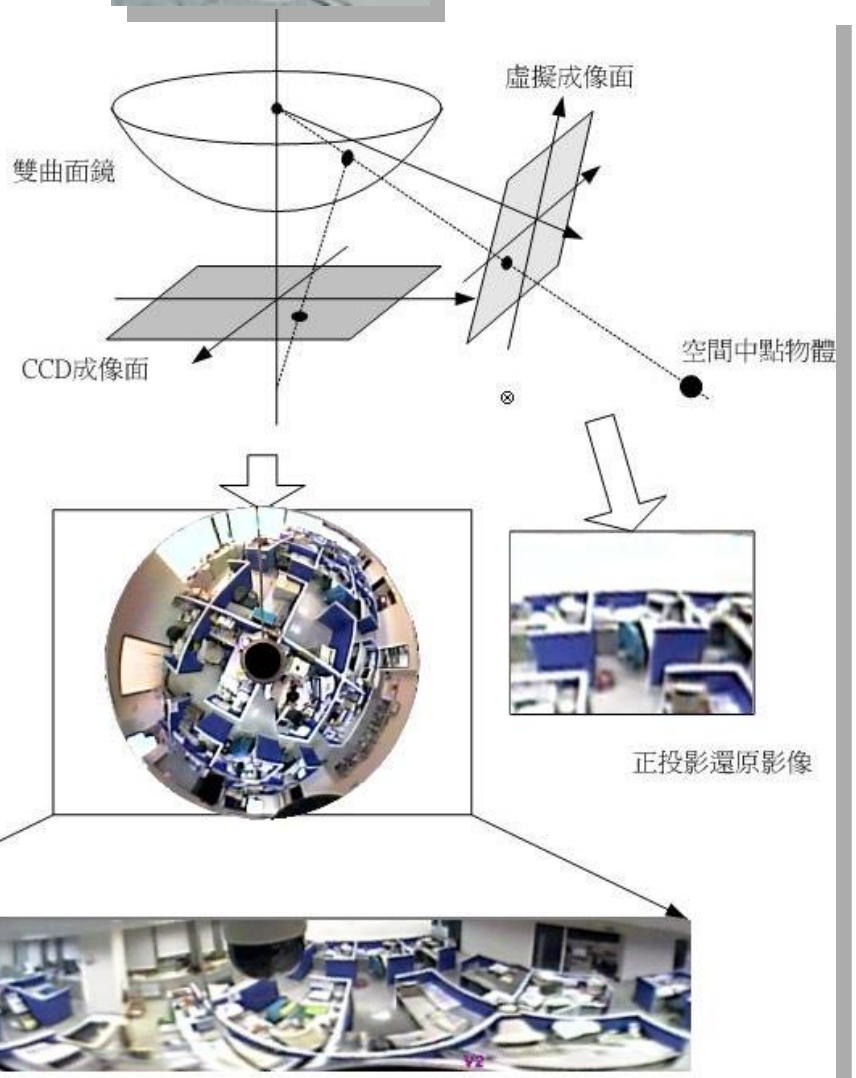
IP camera
(network camera)

Vivotek
晶睿通訊

wireless camera

Camera

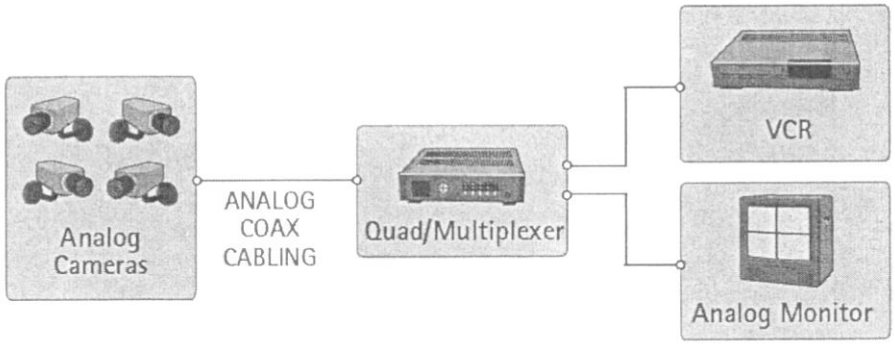
- Omnidirectional



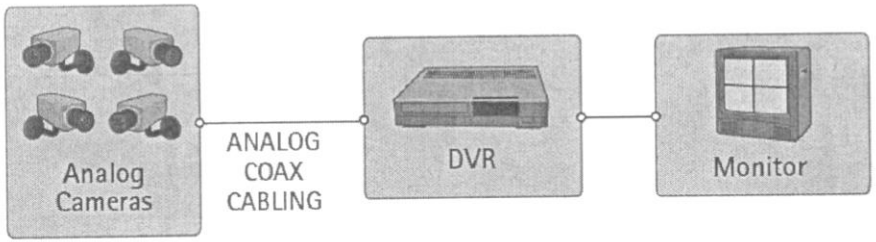
EeRise 宜昇科技

Evolution of Vision System

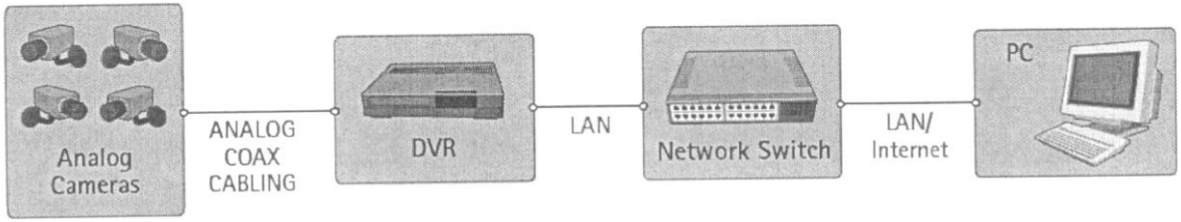
- Generation 1: Analog CCTV systems using VCR



- Generation 2: Analog CCTV systems using DCR

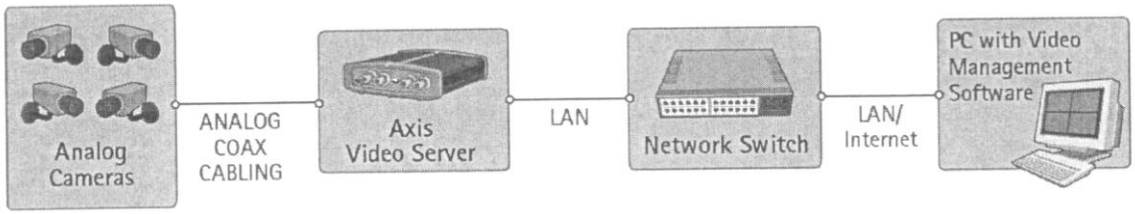


- Generation 2: Analog CCTV systems using network DCR

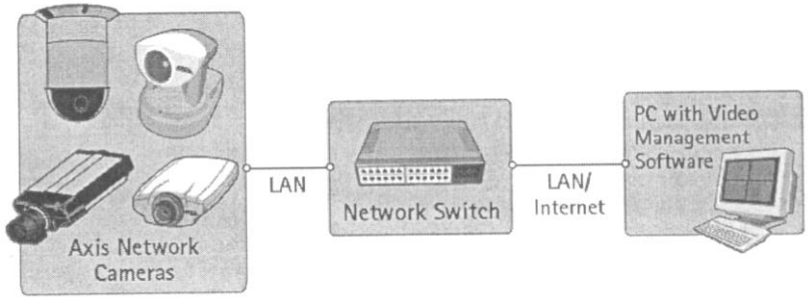


Evolution of Vision System

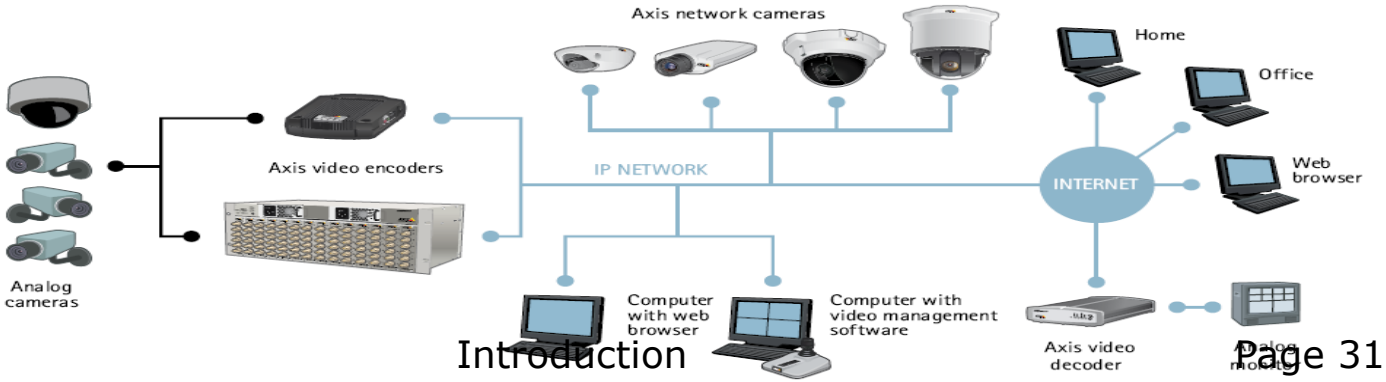
- Generation 3: Network video systems using video server



- Generation 3: Network video systems using network cameras



- Integration

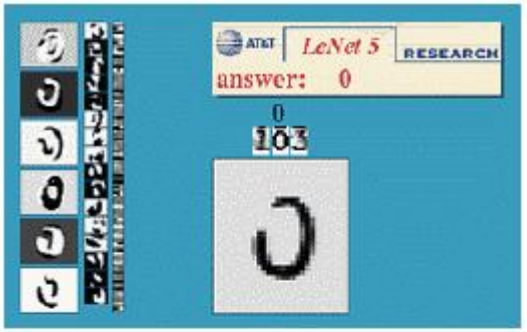


Comparison of VCR, DVR, NVR

	VCR Video Cassette Recorder	DVR Digital Video Recorder	NVR Networked Video Recorder
Period	1970 ~	1990 ~	2000 ~
Scene	Analog	Analog	Analog
Signal	Analog	Analog	Digital
Interface	Cable	Cable	RJ485
Storage	Tape	HDD	HDD / Streaming
Indexing	Manual	Multimedia database	Multimedia database

Some Industrial Applications

optical character recognition



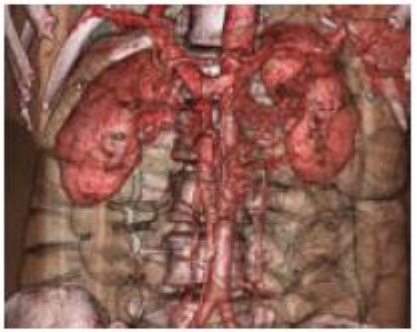
(a)

(b)

mechanical inspection

retail

(c)

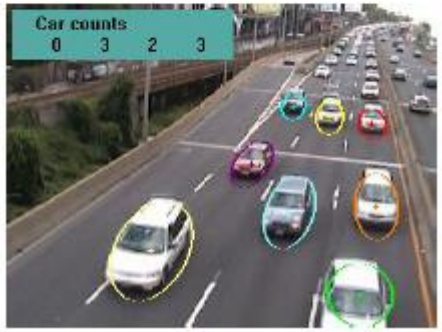


(d)

medical imaging

automotive

(e)



(f)

surveillance and traffic monitoring

Some Consumer Applications

(a)

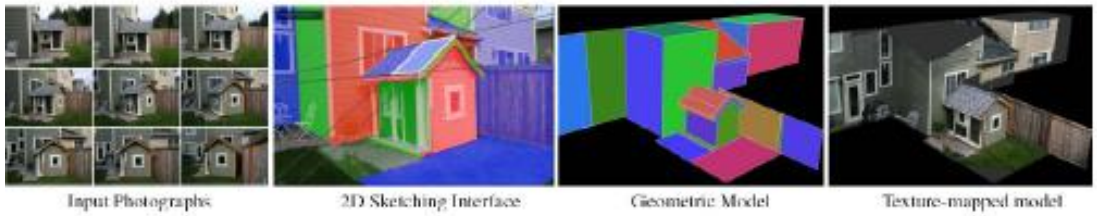
image stitching: merging different views

(b)

exposure bracketing: merging different exposures

(c)

morphing: blending between two photographs

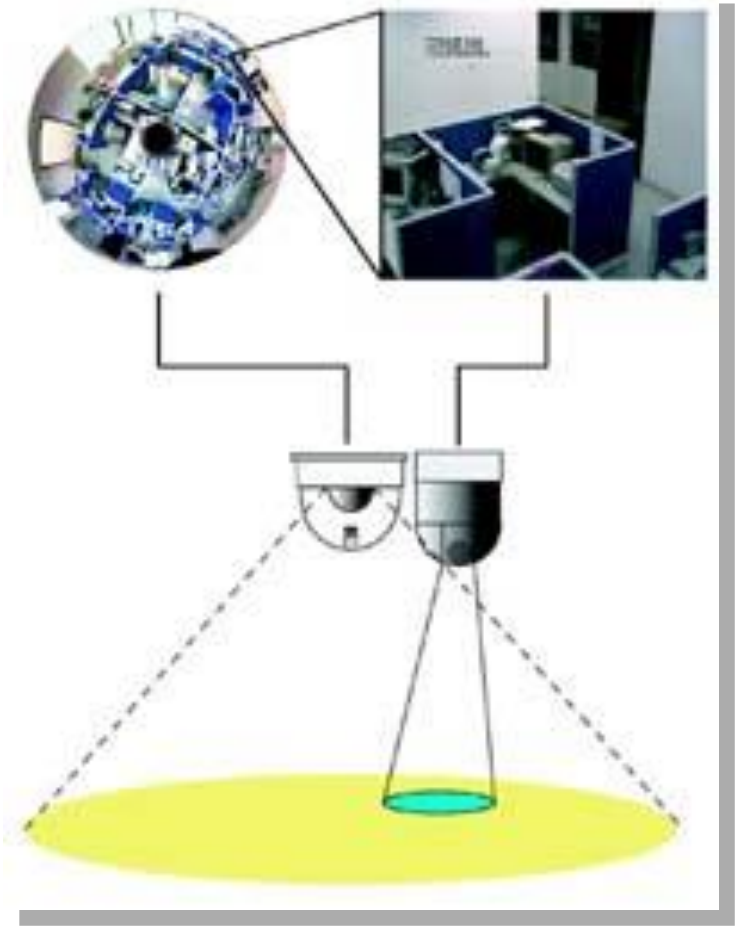


(d)

turning a collection of photographs into a 3D model

Applications

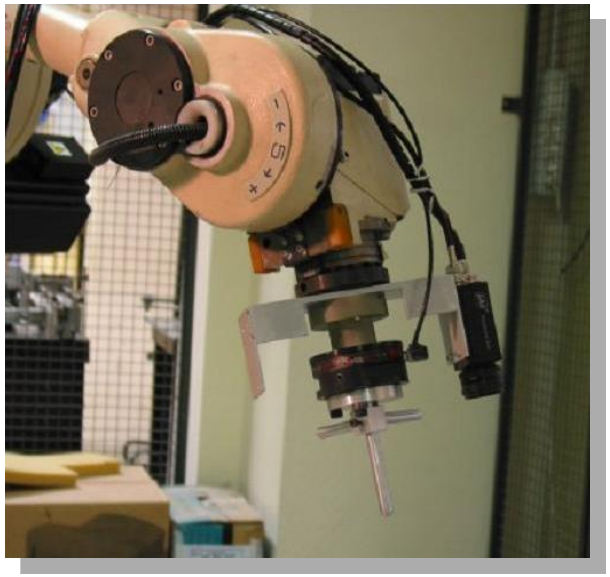
- Surveillance



EeRise 宜昇科技

Applications

- Industrial robot
  - 即時目標影像擷取
  - 座標定位
  - 影像伺服控制
  - 檢測、排程



- Aerial robot
  - 無人駕駛飛行器 (unmanned aerial vehicle, UAV)
  - 影像追蹤鎖定目標物
  - 光學獨立瞄準具(light of sight, LOS)
  - 影像伺服控制飛行器
  - 重建地形地貌

Applications

- Underwater robot
  - 影像地標搜尋選取
  - 追蹤鎖定目標物
  - 依照地標地形線等軌跡航行
  - 多種感測器資料融合
  - 影像伺服控制

- Entertainment robot
  - 影像識別、追蹤，色彩區塊分割及辨識
  - 建構立體視覺及定位、閃避障礙物

Applications

- Home, surveillance robot
  - 即時影像傳輸、動態影像擷取
  - 即時定位及環境重建
  - 人形辨識判斷以供居家看護
  - 人臉辨識、影像追蹤鎖定侵入者
  - 影像辨識依照指定軌跡行進
  - 全域式相機增廣監控的視野

SecurityBot

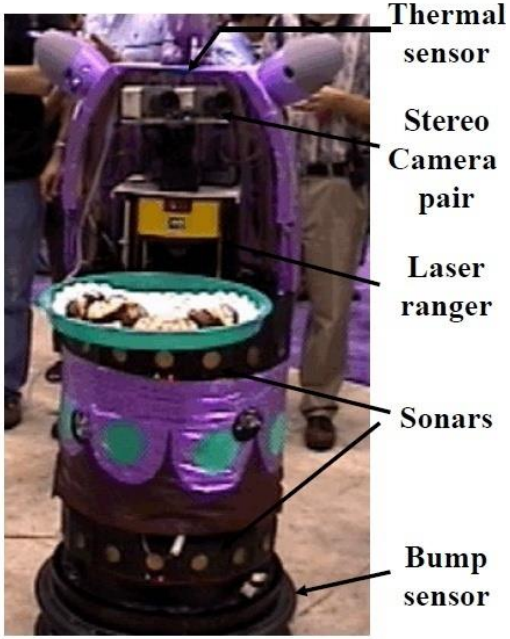
EMIEW 2

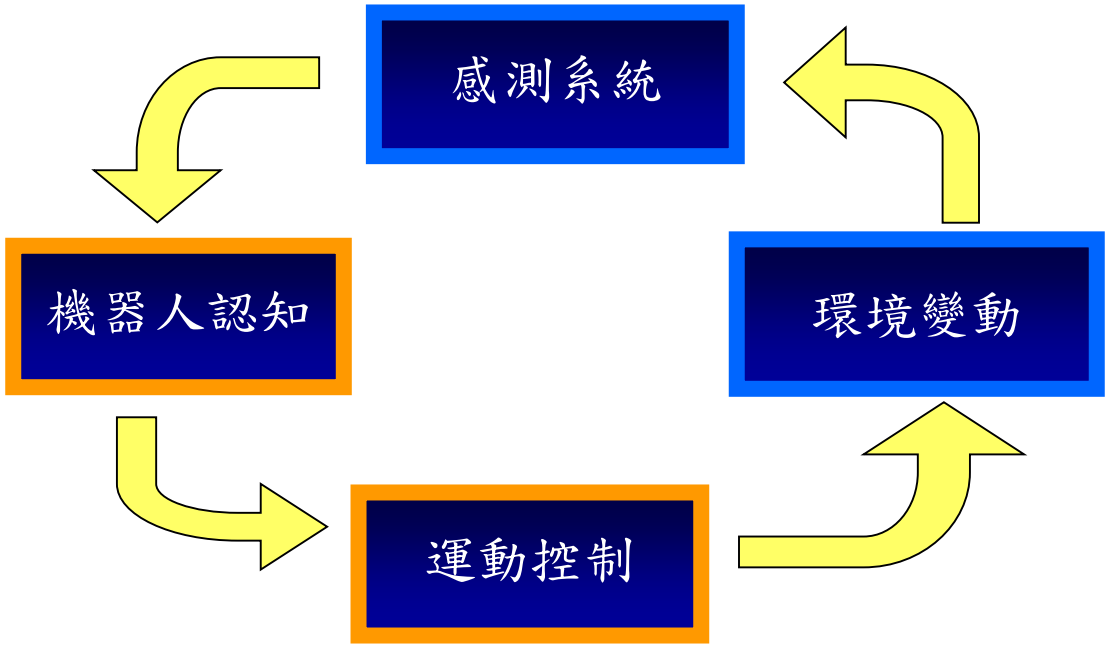
Mahru-Z

新保中正一號

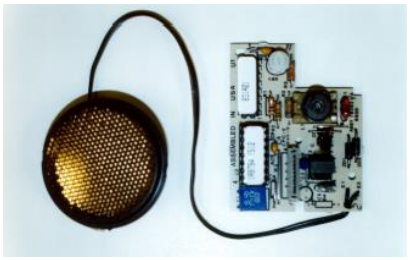
感知系統對機器人的重要性

- 智慧型機器智能能對外在的改變進行認識
- 機器人能增加學習資訊
- 機器人得以藉此規劃出特定的動作
- 機器人能產生對應的行動模式迴授給環境或人類



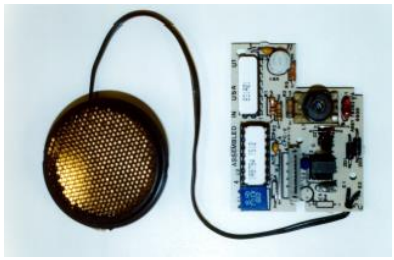


感測器比較



	超音波感測器	雷射感測器	視覺感測器
實體大小	小	大	可微小化
對資料解釋的可靠度	低	最好	好
資料運算量	少， 可處理的資訊量少	不多， 但堆疊時運算量大	與影像擷取大小相關
感測區域範圍	近	遠	與鏡頭相關

感測器比較



	超音波感測器	雷射感測器	視覺感測器
準確度、 可重複性	低 (約0.5inch)	極高	高
能量消耗	高	最高	低
在目標領域的反應	易受環境、 目標物的影響	較佳	與光源、對比相關
硬體成本	低	高	低

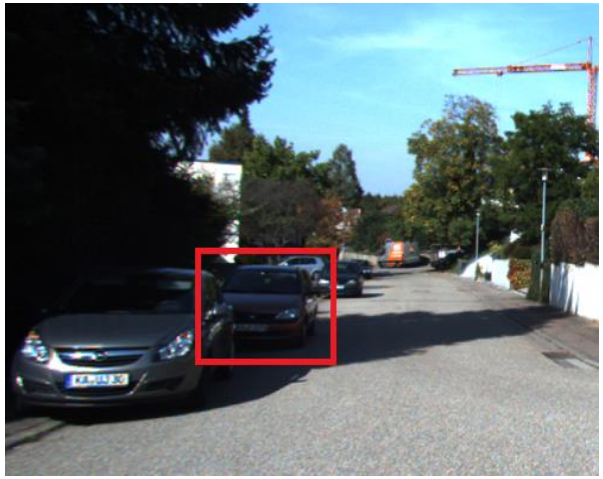
感測器比較

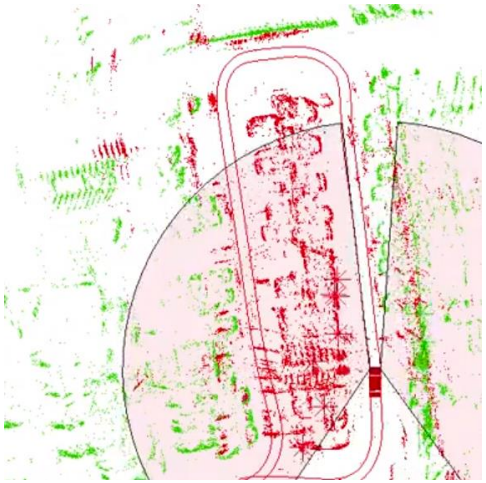
Camera

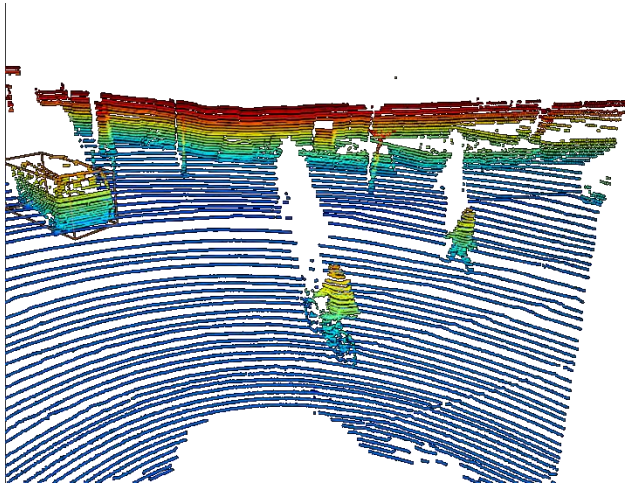
Radar / 2D Laser

Lidar

感測資料







視覺感測器的功能特性

- 最直觀的感測資訊
- 大量多元的環境資料
- 可細部分析：
  - 物體大小、物體形狀、邊緣輪廓
  - 明暗、顏色、多種頻譜與色彩空間
  - 材質、紋理
  - 轉動方向、姿態