ระเบียบวิธีวิจัย

Research Methodology

Introduction to Computer Vision

ภิญโญ แท้ประสาทสิทธิ์

(pinyo at su.ac.th, pinyotae at gmail dot com)

ภาควิชาคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร

Topic Outline

- ภาพและการมองเห็น
  - อะไรคือภาพ
  - อะไรคือการมองเห็น (ด้วยคอมพิวเตอร์) ?
- ความเกี่ยวข้องของวิชาการมองเห็นของคอมพิวเตอร์และศาสตร์อื่น ๆ
- ทำไมถึงต้องมีการเรียนรู้และวิจัยเกี่ยวกับการมองเห็นของคอมพิวเตอร์
- การประยุกต์ใช้คอมพิวเตอร์ในงานที่เกี่ยวกับการมองเห็น
- รู้จักกับวิชาการมองเห็นของคอมพิวเตอร์ให้มากขึ้น

ภาพและการมองเห็น

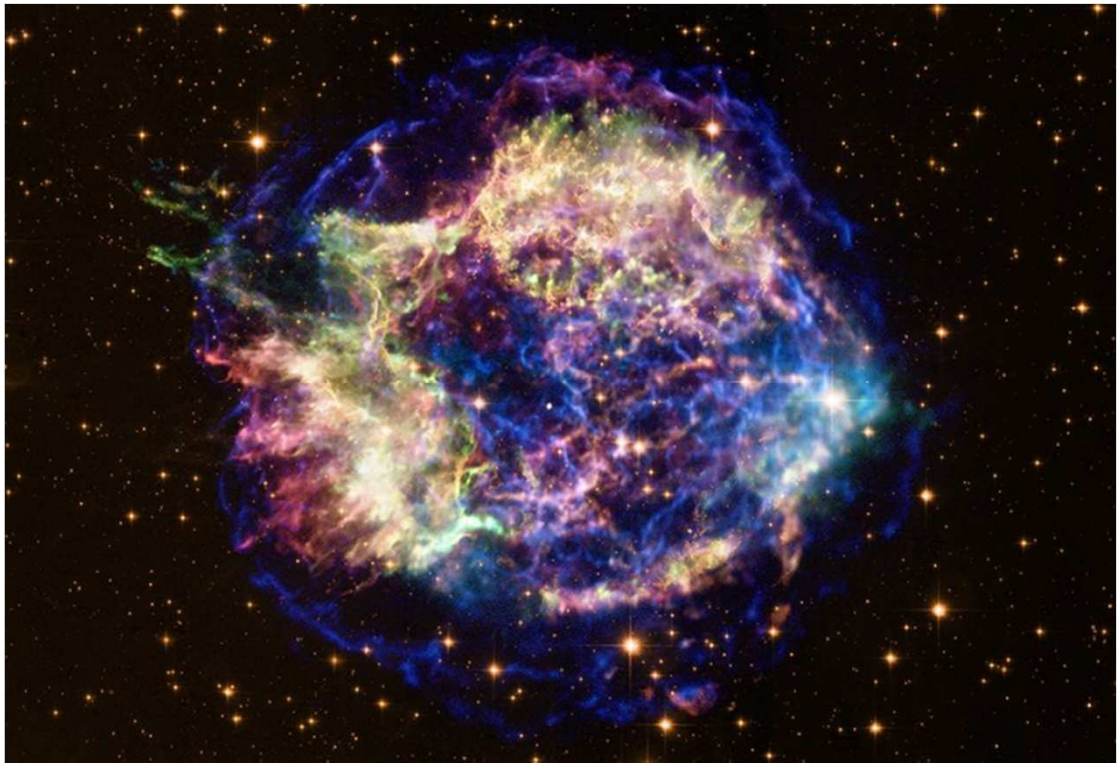
- ภาพ (image) คือ สิ่งที่ถูกเห็น
  - จะเป็นภาพหนึ่งมิติก็ได้ (มักถูกนำมาให้ประกอบการอธิบายทฤษฎีต่าง ๆ)
  - ถ้ามีเวลามากำกับก็อาจจะถือเป็นภาพสี่มิติ เช่น วีดีโอของภาพสามมิติ
  - มากกว่าสี่มิติมักจะไม่มีแม่แบบทางด้านการมองเห็น ข้อมูลที่เข้ามาประมวลผลมักจะถูกเรียกว่าสัญญาณหลายมิติมากกว่าที่จะเป็นภาพหลายมิติ
  - นิยามของภาพหลายมิติมักจะขึ้นอยู่กับระบบ เช่น MeVisLab ถือว่าระนาบสีแต่ละอันเป็นมิติที่แยกออกมาต่างหาก
- การมองเห็น (Vision) คือการมองเห็นภาพ รับรู้ได้ผ่านสัญญาณอิเล็กโทรนิกหรือสัญญาณชีวภาพโดยมากจะเป็นสองมิติหรือสามมิติ
- การมองเห็นของคอมพิวเตอร์ (Computer Vision) จะเน้นที่ความสามารถในการดึงข้อมูลออกมาจากภาพที่เครื่องมองเห็น มักจะเกี่ยวข้องกับ

  การรับรู้ (perception), การประทับจำ (registration), และความฉลาด (intelligence)

ภาพดิจิตอลอาจได้มาจากสัญญาณหลายประเภท

ภาพมักถูกสร้างจากสัญญาณอย่างใดอย่างใดอย่างหนึ่ง

- ส่วนใหญ่มาจากแสงที่ตามองเห็น (visible spectrum)
- ภาพไม่จำเป็นต้องมาจากแสงที่ตามองเห็นแต่เพียงอย่างเดียว



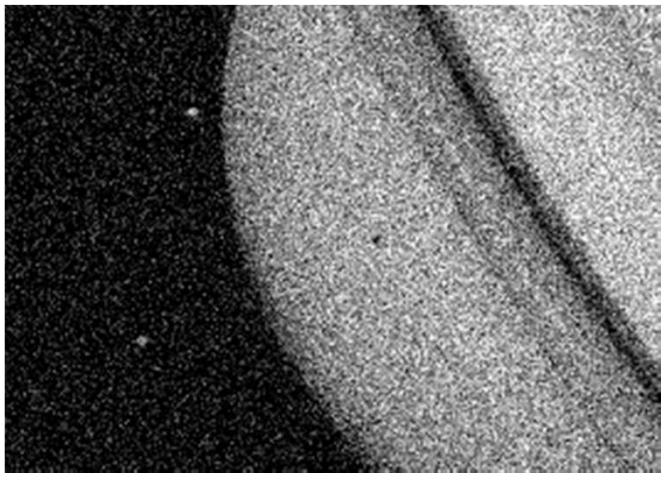
Supernova Remnant Cassiopeia A -- Blend of X-ray and visible spectrum

(image courtesy: NASA, Chandra, Hubble)

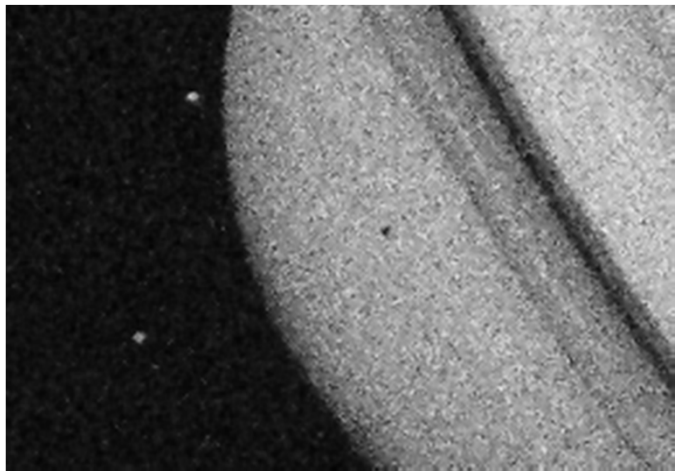
การประมวลผลภาพและการมองเห็นของคอมพิวเตอร์

การประมวลผลภาพเน้นที่การทำให้ภาพมีคุณสมบัติตามความต้องการของผู้ใช้

- แต่ในปัจจุบันขอบเขตของงานประมวลผลภาพก็รวมไปถึง
  - การแยกพื้นที่ในภาพ (image segmentation)
  - การรู้จำภาพ (image recognition)
  - การประทับจำภาพ (image registration)
- งานบางส่วนของการประมวลผลภาพกับการมองเห็นของคอมพิวเตอร์มีลักษณะเหมือนกัน



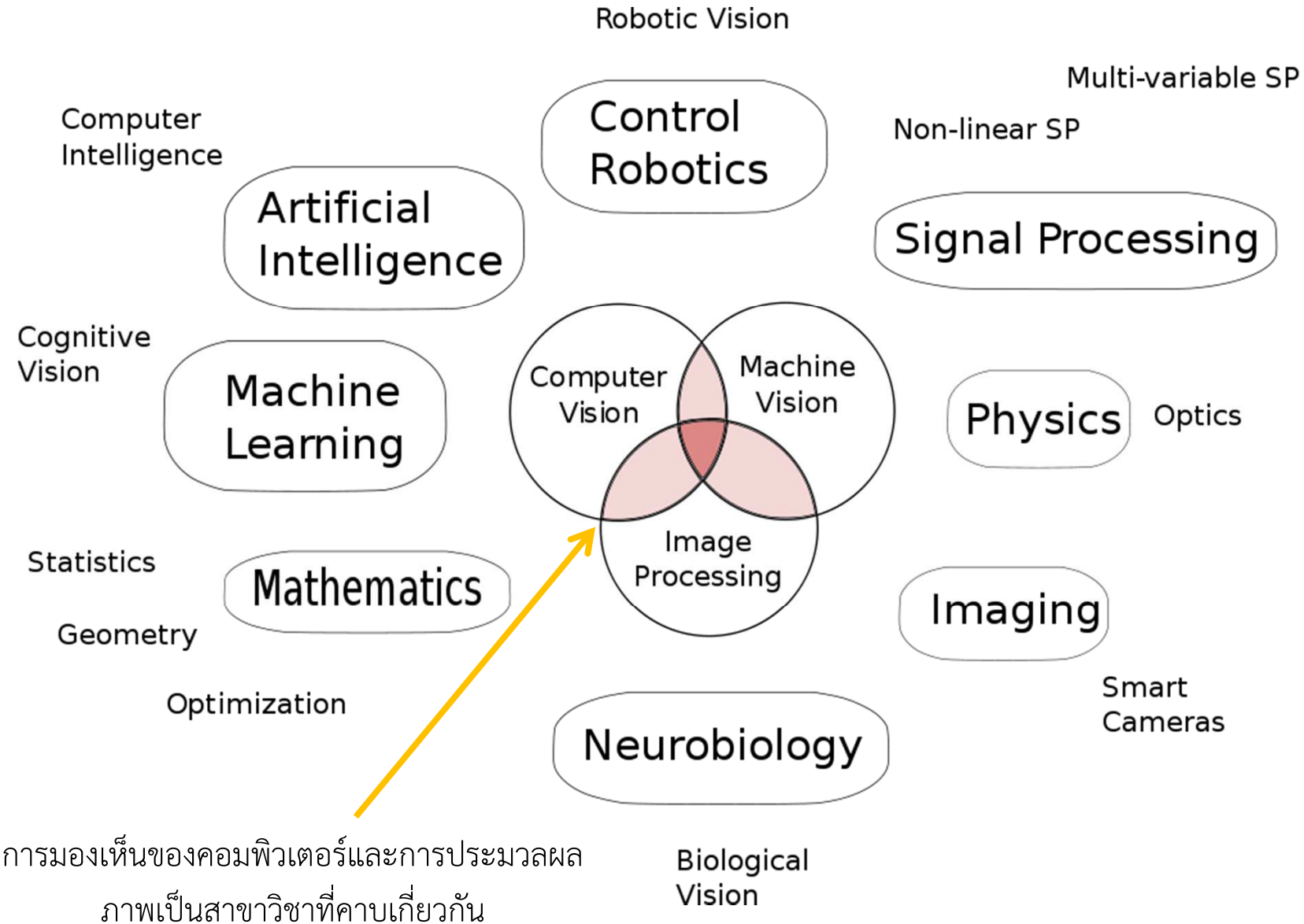
ภาพดาวเสาร์ที่เต็มไปด้วยสัญญาณรบกวน



ภาพที่ผ่านการลดสัญญาณรบกวน

ภาพจาก MathWorks (สัญญาณรบกวนเป็นการสังเกตุการณ์ขึ้นมา)

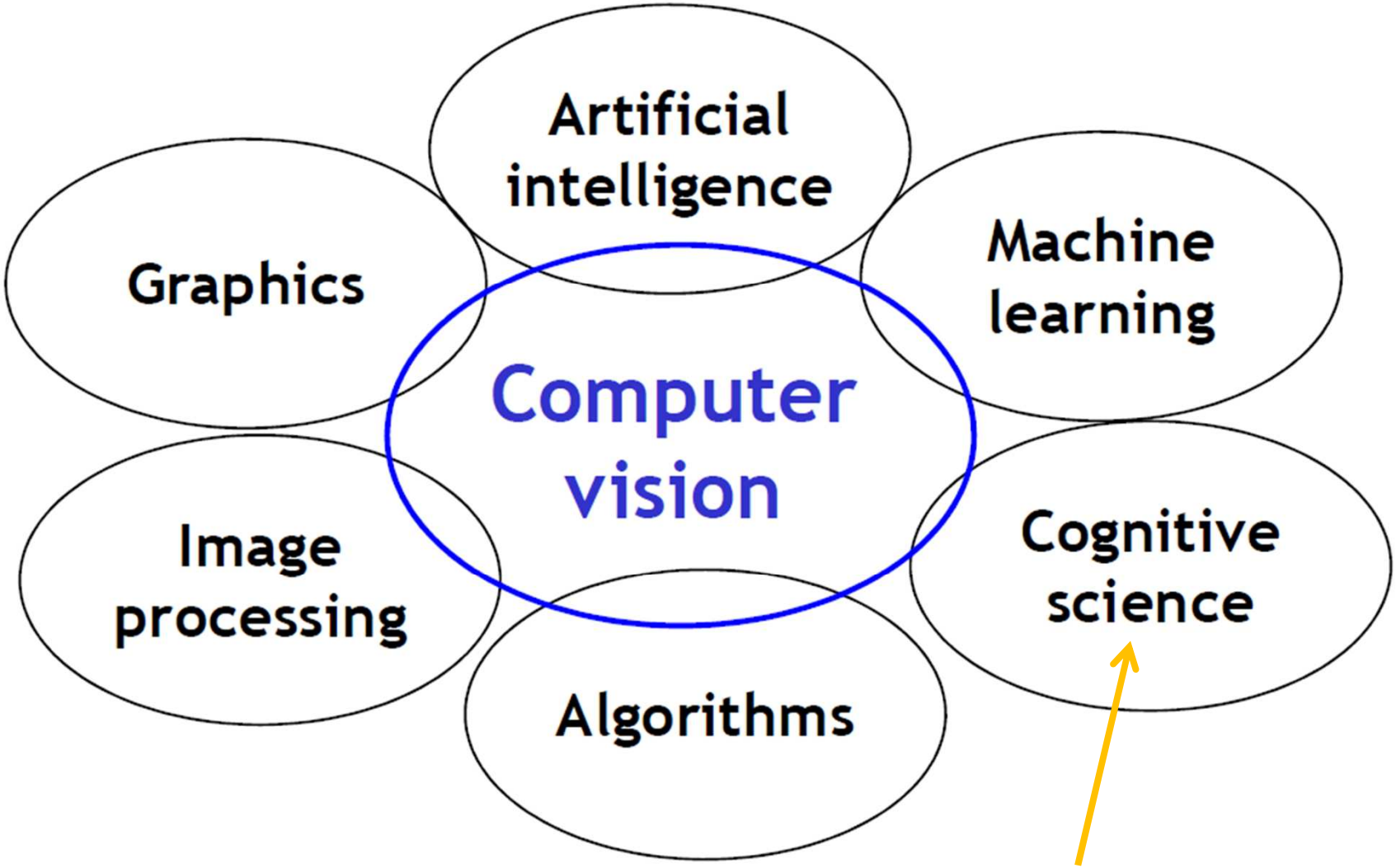
ความเกี่ยวข้องกับศาสตร์แขนงอื่น



การมองเห็นของคอมพิวเตอร์และการประมวลผลภาพเป็นสาขาวิชาที่คาบเกี่ยวกัน

ภาพจาก wikipedia.org

ความเกี่ยวข้องกับสาขาอื่นแบบย่อ ๆ สไตล์ภาคคอม



มีแค่วิชาเดียวที่ไม่อยู่ในหลักสูตรภาคคอมทั่วไป

Slide credit: Kristen Grauman and Bastian Leibe

ทำไมต้องศึกษาคอมพิวเตอร์วิชัน

เพราะการมองเห็นผ่านสัญญาณอิเล็กทรอนิกมีอยู่ทั่วไปในปัจจุบัน มีการใช้งานที่แพร่หลายในสาขาต่าง ๆ

Slide credit: Kristen Grauman

ทำไมต้องศึกษาคอมพิวเตอร์วิชัน (2)

- กล้อง ภาพนิ่ง และ ภาพเคลื่อนไหวมีอยู่ทุกที่

- การประยุกต์ใช้ที่มีประโยชน์เพิ่มขึ้นอย่างรวดเร็ว
  - การแสดงผลสามมิติจากการนำภาพสองมิติมารวมกัน
  - การตรวจตราความปลอดภัยแบบอัตโนมัติ (ค้นหากิจกรรมที่ต้องสงสัยของคนในภาพ)
  - การตัดต่อภาพยนตร์

- ทำให้เข้าใจการทำงานของการมองเห็นของมนุษย์ดียิ่งขึ้นด้วย

- ยังมีสิ่งลี้ลับทางด้านวิทยาศาสตร์ที่น่าสนใจอีกมาก
  - การรู้จำวัตถุในบุคคลเกิดขึ้นได้อย่างไร

Slide credit: Robert T. Collins

Topic Outline

- ภาพและการมองเห็น
  - อะไรคือภาพ
  - อะไรคือการมองเห็น (ด้วยคอมพิวเตอร์) ?
- ความเกี่ยวข้องของวิชาการมองเห็นของคอมพิวเตอร์และศาสตร์อื่น ๆ
- ทำไมถึงต้องมีการเรียนรู้และวิจัยเกี่ยวกับการมองเห็นของคอมพิวเตอร์
- **การประยุกต์ใช้คอมพิวเตอร์ในงานที่เกี่ยวกับการมองเห็น**
- รู้จักกับวิชาการมองเห็นของคอมพิวเตอร์ให้มากขึ้น
- เราจะเรียนอะไรบ้างในวิชานี้

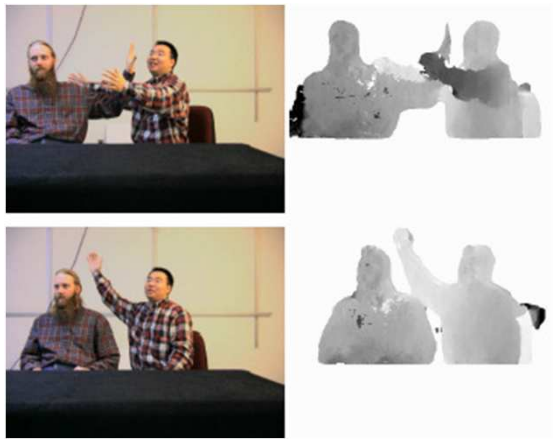
การประยุกต์ใช้คอมพิวเตอร์ในงานที่เกี่ยวกับการมองเห็น

- **การทำความเข้าใจภาพโดยอัตโนมัติผ่านระบบคอมพิวเตอร์**
  - การคำนวณคุณสมบัติต่าง ๆ เช่น ระยะทาง จากโลกสองหรือสามมิติจากข้อมูลทางภาพ (การวัดคุณสมบัติ / measurement)
  - การรู้จำวัตถุ บุคคล สถานที่ และ กิจกรรม ตลอดจนการหาความสัมพันธ์ของสิ่งเหล่านี้ (การรับรู้และการแปลความหมาย / perception and recognition)
  - การแบ่งขอบเขตรูป (segmentation)

- **การประยุกต์ใช้ที่เจาะจงในแขนงใดแขนงหนึ่ง**
  - การทหาร
  - การกีฬา
  - การแพทย์

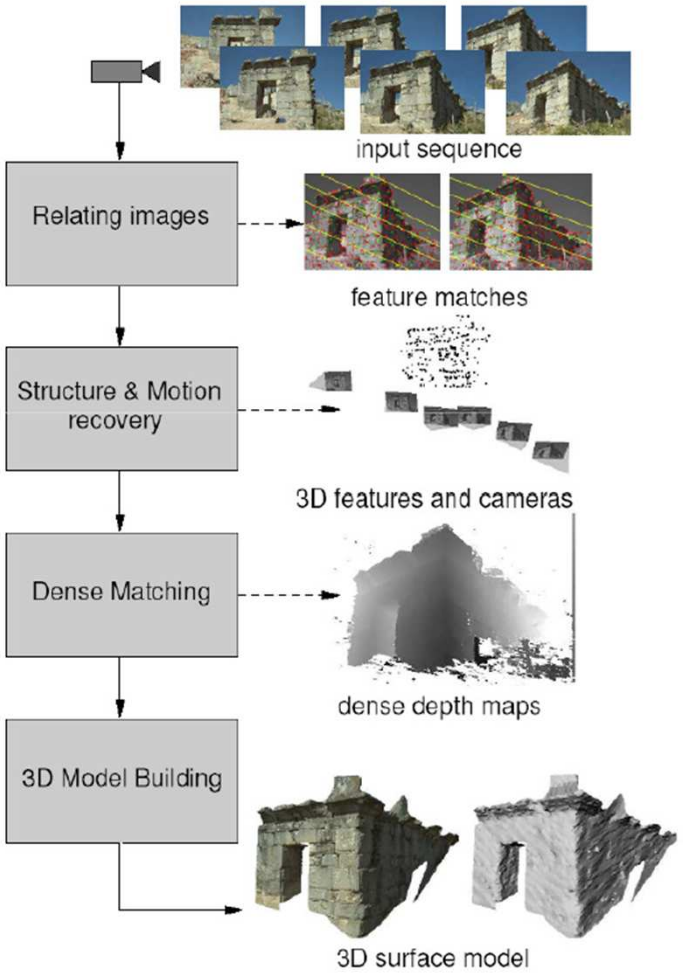
การวัดคุณสมบัติในภาพ

Real-time Stereo



Pollefeys et al.

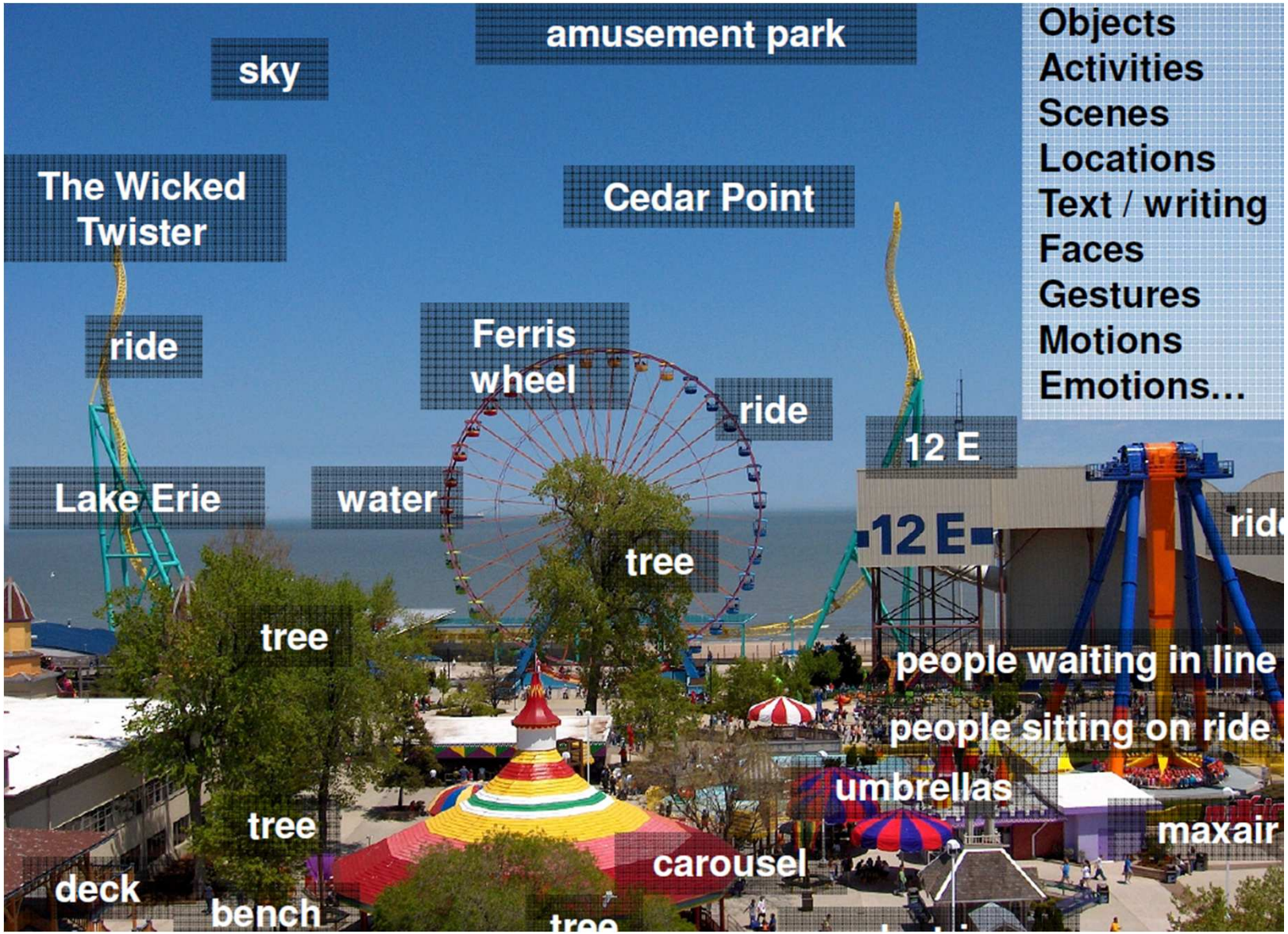
Structure from Motion



วัดระยะได้จากการใช้กล้องหรือเซนเซอร์สองตัวขึ้นไป

Slide credit: Kristen Grauman and Bastian Leibe

การมองเห็นทำให้เกิดการรู้จำและตีความหมาย



Slide credit: Kristen Grauman

Face Detection and Recognition

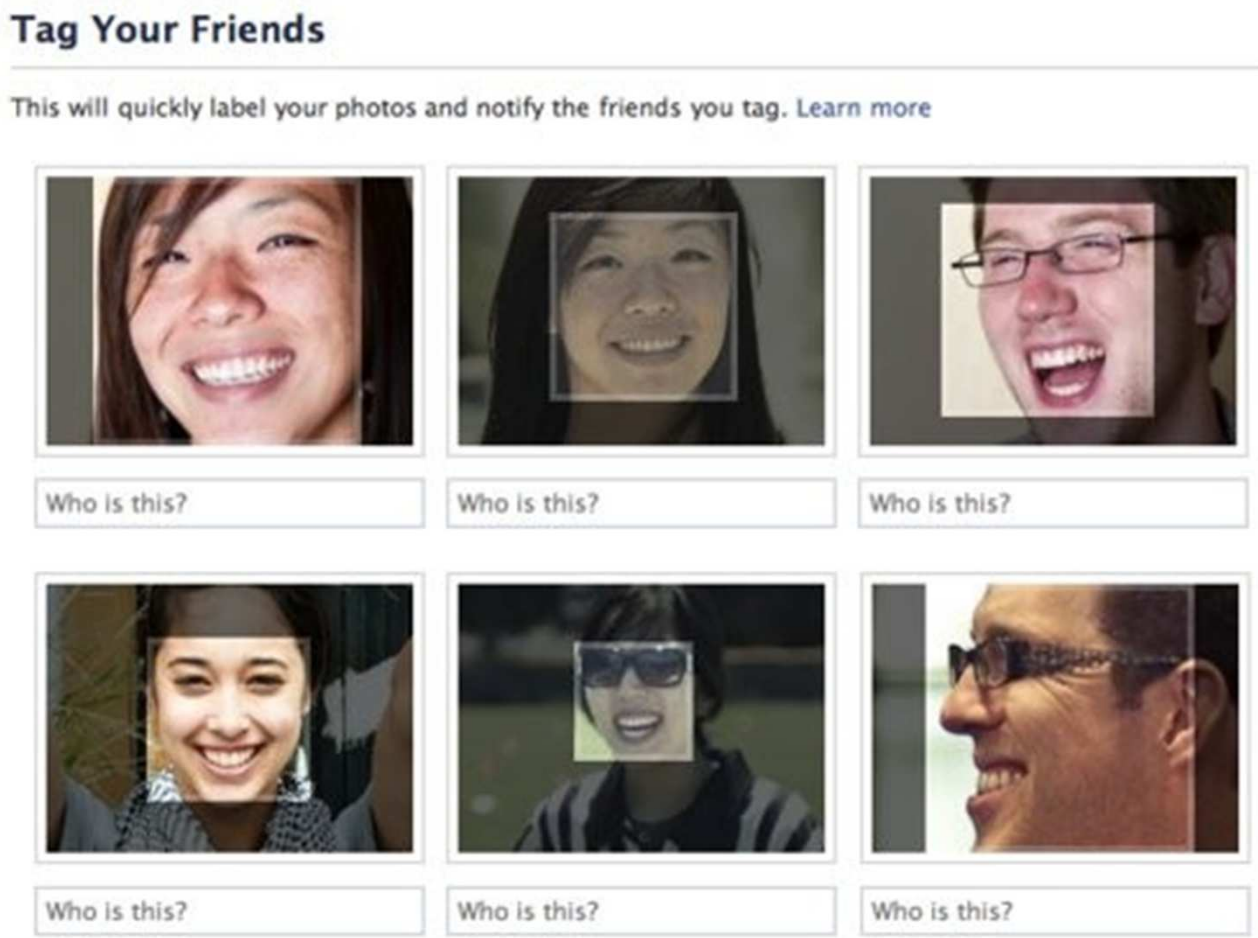


Image: Facebook

การแบ่งขอบเขตของรูป

- กรรไกรอัจฉริยะ (Intelligent Scissors or Live-wire Segmentation)

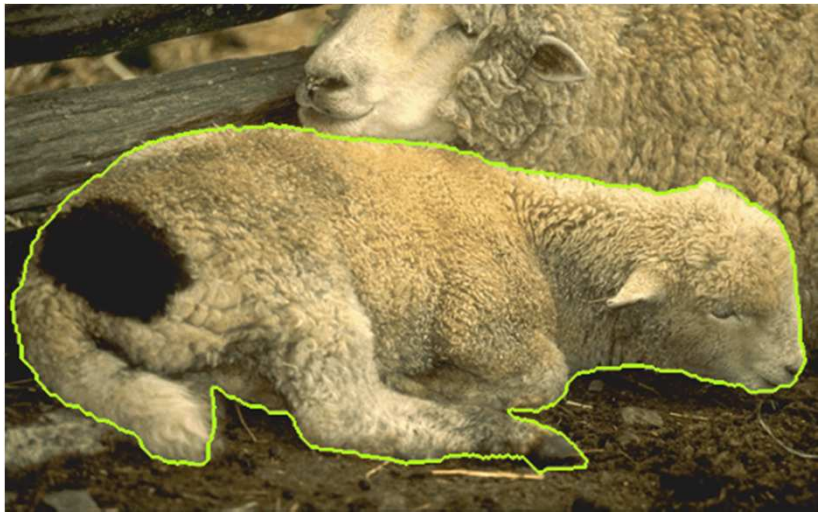
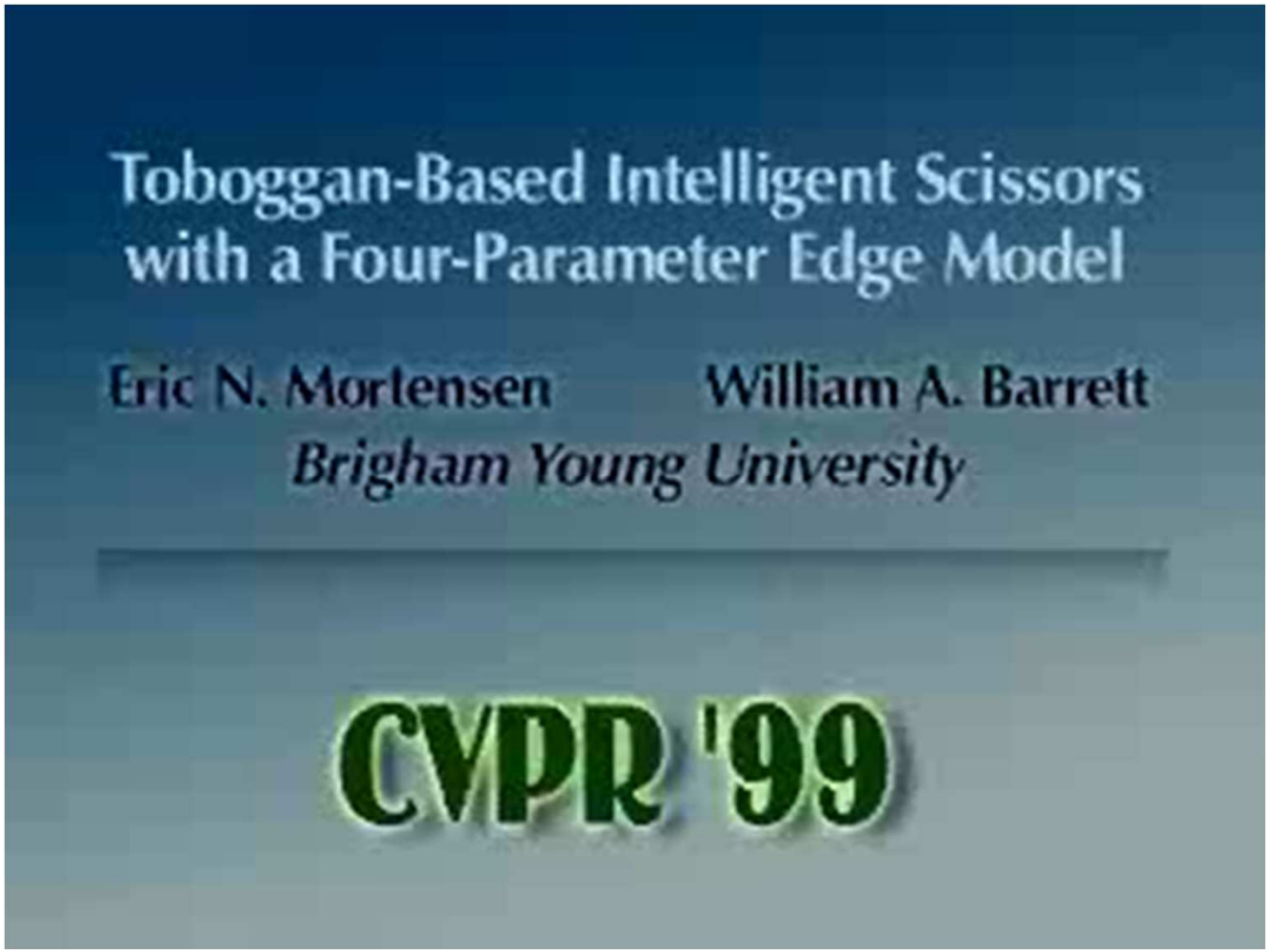


Image: Mortensen SIGGRAPH 1995

Intelligent Scissors in Action



http://web.engr.oregonstate.edu/~enm/publications/CVPR_99/toboggan_scissors.mov

การประยุกต์ใช้คอมพิวเตอร์ในงานที่เกี่ยวกับการมองเห็น

- การทำความเข้าใจภาพโดยอัตโนมัติผ่านระบบคอมพิวเตอร์
  - การคำนวณคุณสมบัติต่าง ๆ เช่น ระยะทาง จากโลกสองหรือสามมิติจากข้อมูลทางภาพ (การวัดคุณสมบัติ / measurement)
  - การรู้จำวัตถุ บุคคล สถานที่ และ กิจกรรม ตลอดจนการหาความสัมพันธ์ของสิ่งเหล่านี้ (การรับรู้และการแปลความหมาย / perception and recognition)
  - การแบ่งขอบเขตรูป (segmentation)

- **การประยุกต์ใช้ที่เจาะจงในแขนงใดแขนงหนึ่ง**
  - การทหาร
  - การกีฬา
  - การแพทย์

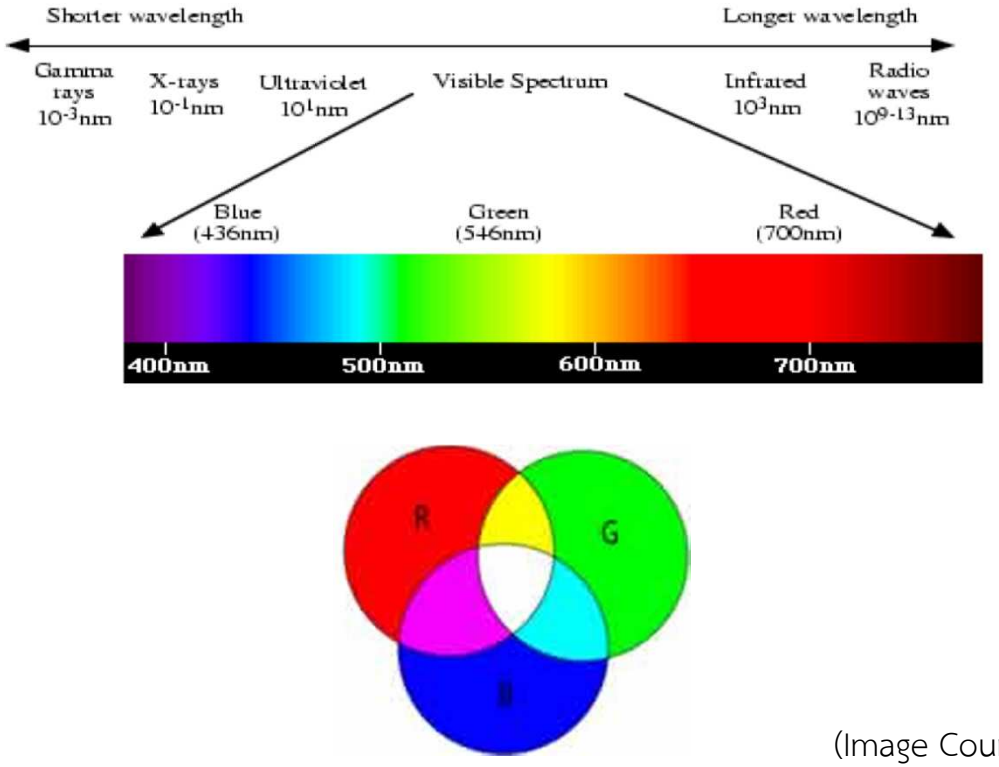
ตัวอย่าง: การประยุกต์ใช้การมองเห็นสีในการทหาร

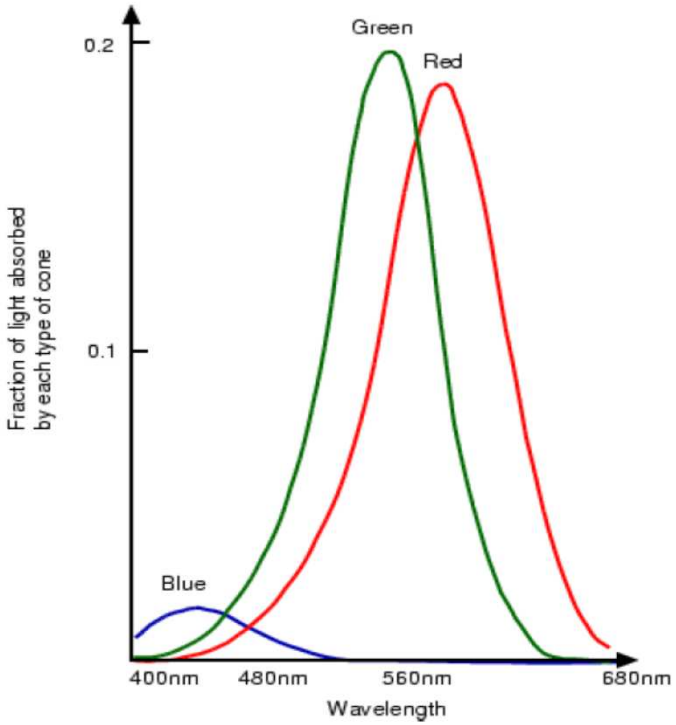
- เป็นการประยุกต์เชิงวิศวกรรมที่มีพื้นฐานมาจากความรู้ทางระบบประสาท
- **การมองเห็นสี** คือ ความสามารถของสิ่งมีชีวิตหรือเครื่องกลในการแยกแยะวัตถุโดยอาศัยความยาวคลื่นที่วัตถุเหล่านั้นสะท้อนหรือปล่อยออกมา

(Image Courtesy: Wikipedia.org)

การรับรู้สีของตามนุษย์

- ตาของมนุษย์สามารถรับรู้คลื่นความถี่แม่เหล็กไฟฟ้าได้ในช่วงความถี่ประมาณ 400 – 750 nm
- ตาของมนุษย์ไม่ได้ไว้ต่อสีแต่ละอันเท่ากันหมด แต่ไว้ต่อสีเขียวมากที่สุด

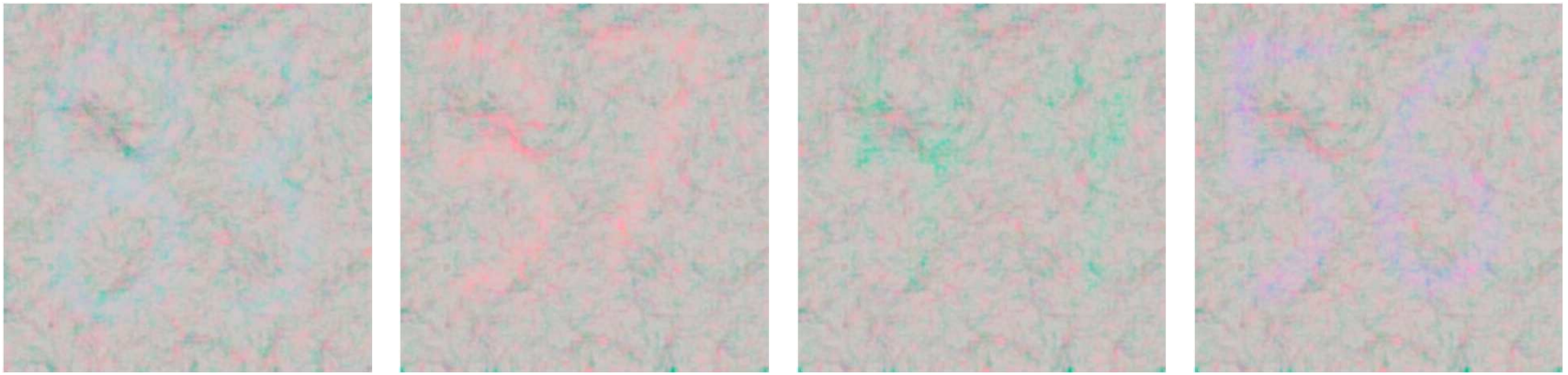




(Image Courtesy: Gonzalez)

ตาบอดสี

- **ตาบอดสี** (**Color blindness** or **color vision deficiency**) ในมนุษย์คือความไม่สามารถรับรู้ความแตกต่างระหว่างสีบางอย่างหรือทั้งหมดที่คนทั่วไปรับรู้ได้
- โดยมากเป็นผลจากกรรมพันธุ์ แต่ก็มีบางกรณีเกิดขึ้นจากความเสียหายที่ตา ระบบประสาท หรือ สมอง หรือการได้รับสารเคมีบางชนิด
- ตาบอดสีมีอยู่หลายประเภท คนจำนวนมากตาบอดสีแค่สีเดียว



(Image Courtesy: Wikipedia.org)

การมองเห็นสีกับการประยุกต์ใช้ทั่วไป

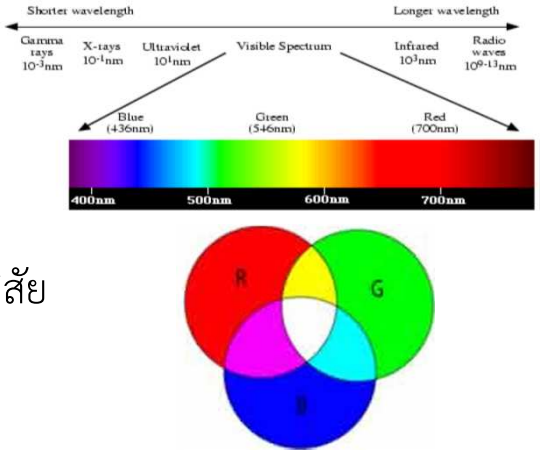
- Better graphic design: avoid using only color coding or color contrasts to express information.
- Traffic light system design.
- Better equipment design: some emergency equipments have only shades of gray. Why?
- **Human's vision enhancement.**

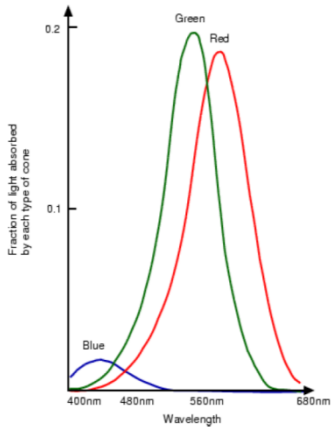
แว่นมองกลางคืน (Night-Vision Goggles)

การมองเห็นในกลางคืน (Night Vision) คือ ความสามารถในการมองเห็นผ่านขบวนการทางชีวีภาพหรือเทคโนโลยีในสภาพแวดล้อมที่มืดหรือมีแสงน้อย

(Image Courtesy: U.S. Army)

เทคนิคพื้นฐานที่เกี่ยวข้อง:
เป็นการผนวกของสองวิธี คือ การขยายพิสัยคลื่นความถี่และการขยายพิสัยความเข้มแสง





(Image Courtesy: Gonzalez)

คำถามที่น่าสนใจ

1. ทำไมเราเลือกที่จะใช้สีเขียวในการแสดงผลที่กล้อง
2. ทำไมเราไม่เพิ่มความเข้มแสงขึ้นไปตรง ๆ ทำไมต้องขยายพิสัย

ข้อสังเกต: การแก้ปัญหาทางเทคนิคบางทีก็ต้องอาศัยความรู้จากหลาย ๆ แขนงประกอบกัน

การประยุกต์ใช้ทางการกีฬา (1)

- การวางซ้อนภาพในการถ่ายทอดกีฬาโดยอัตโนมัติ

(Video courtesy: NFL SuperBowl 2010, obtained from YouTube)

http://www.youtube.com/watch?v=j3L09liVopc

การประยุกต์ใช้ทางการกีฬา (2)

Hawk-Eye: ระบบการติดตามลูกบอลและบอกตำแหน่งเชิงสามมิติ

(Ball-Tracking and 3D-Positioning System)

(Video courtesy: Australian Open 2009, obtained from YouTube)

http://www.youtube.com/watch?v=E5-abgnmGIU

การประยุกต์ใช้ทางการกีฬา (3)

- EyeVision: 3D Dynamic Video Panorama

(Video courtesy: Takeo Kanade, Carnegie Mellon University; NFL SuperBowl XXXV)

http://www.ri.cmu.edu/events/sb35/eyevision_best_of.mpg

การประยุกต์ใช้งานด้านการแพทย์ (1)

NECTEC's DentiPlan: ซอฟต์แวร์ช่วยวางแผนผ่าตัดสำหรับงานทันตกรรมรากฟันเทียม

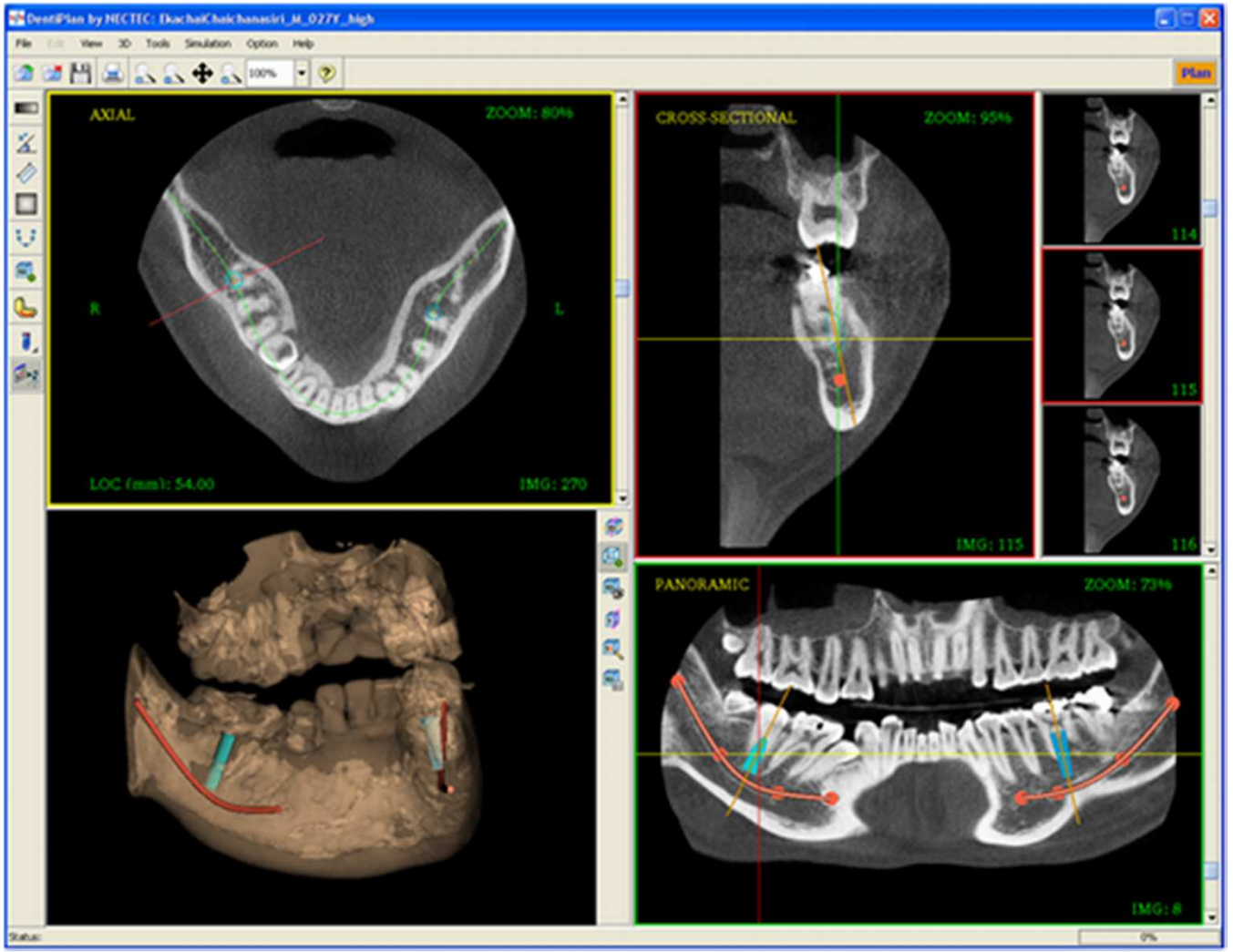


Image Courtesy: NECTEC

DentiPlan's Features

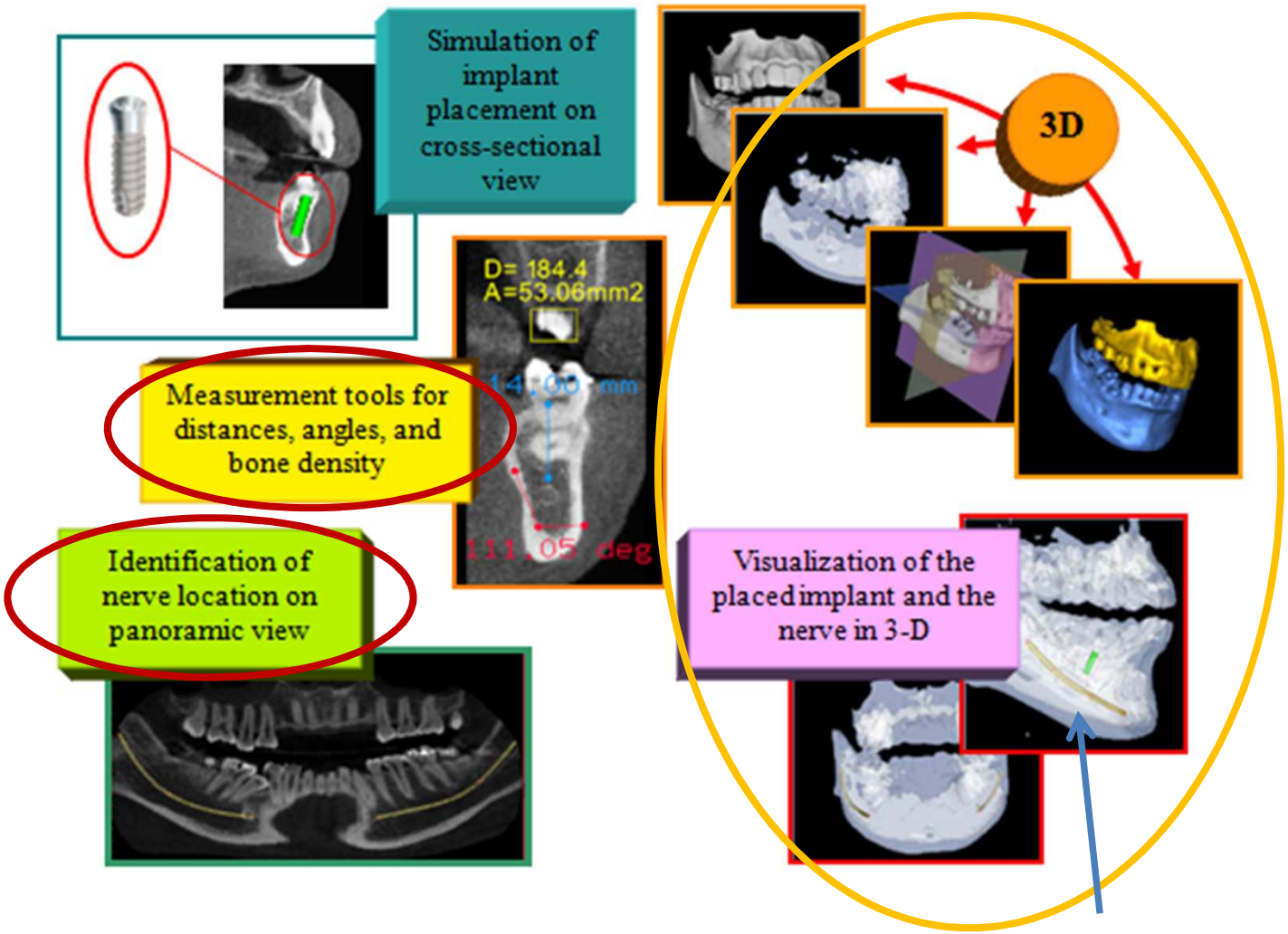


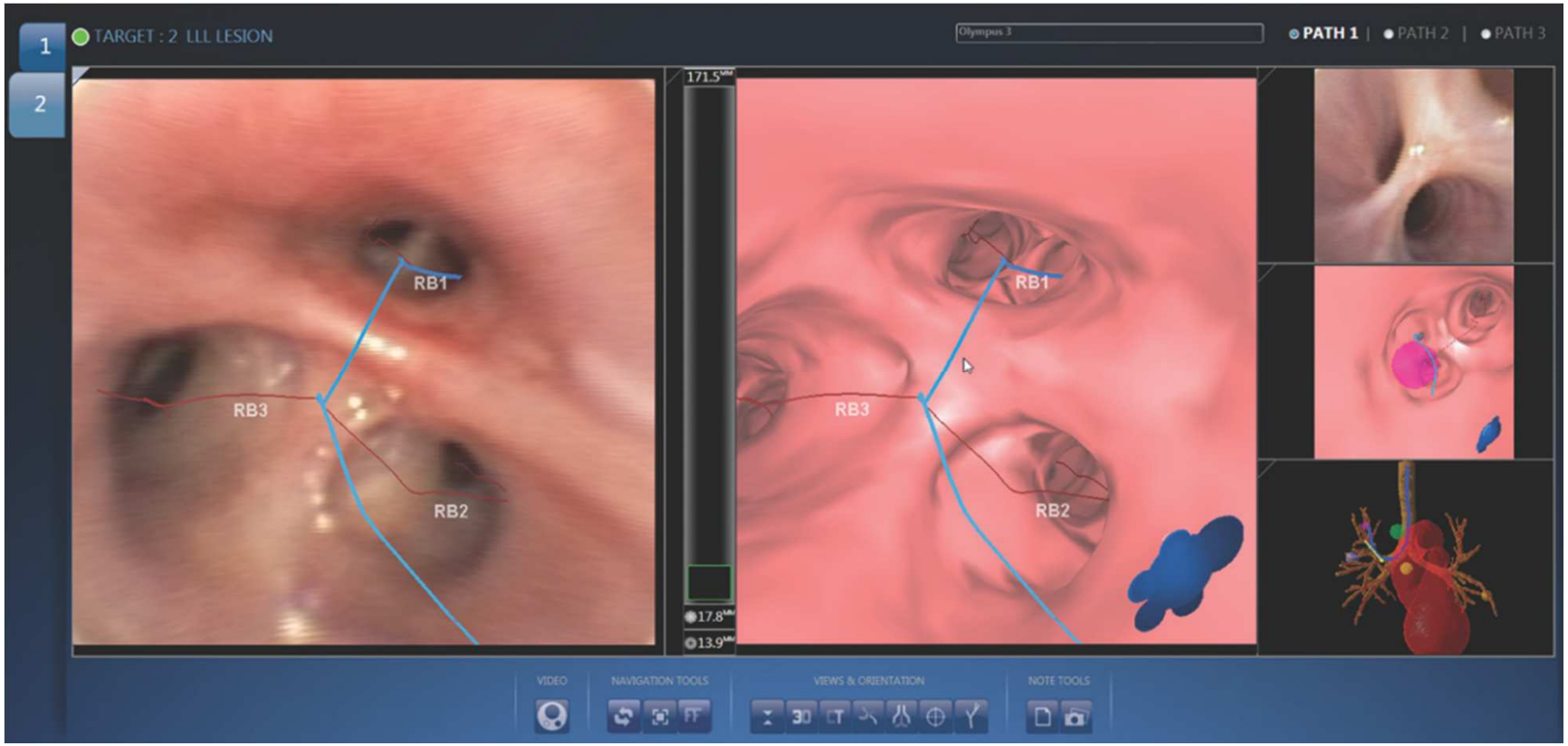
Image Courtesy: NECTEC

Computer graphics is usually an integral part of a computer vision system

การประยุกต์ใช้งานด้านการแพทย์ (2)

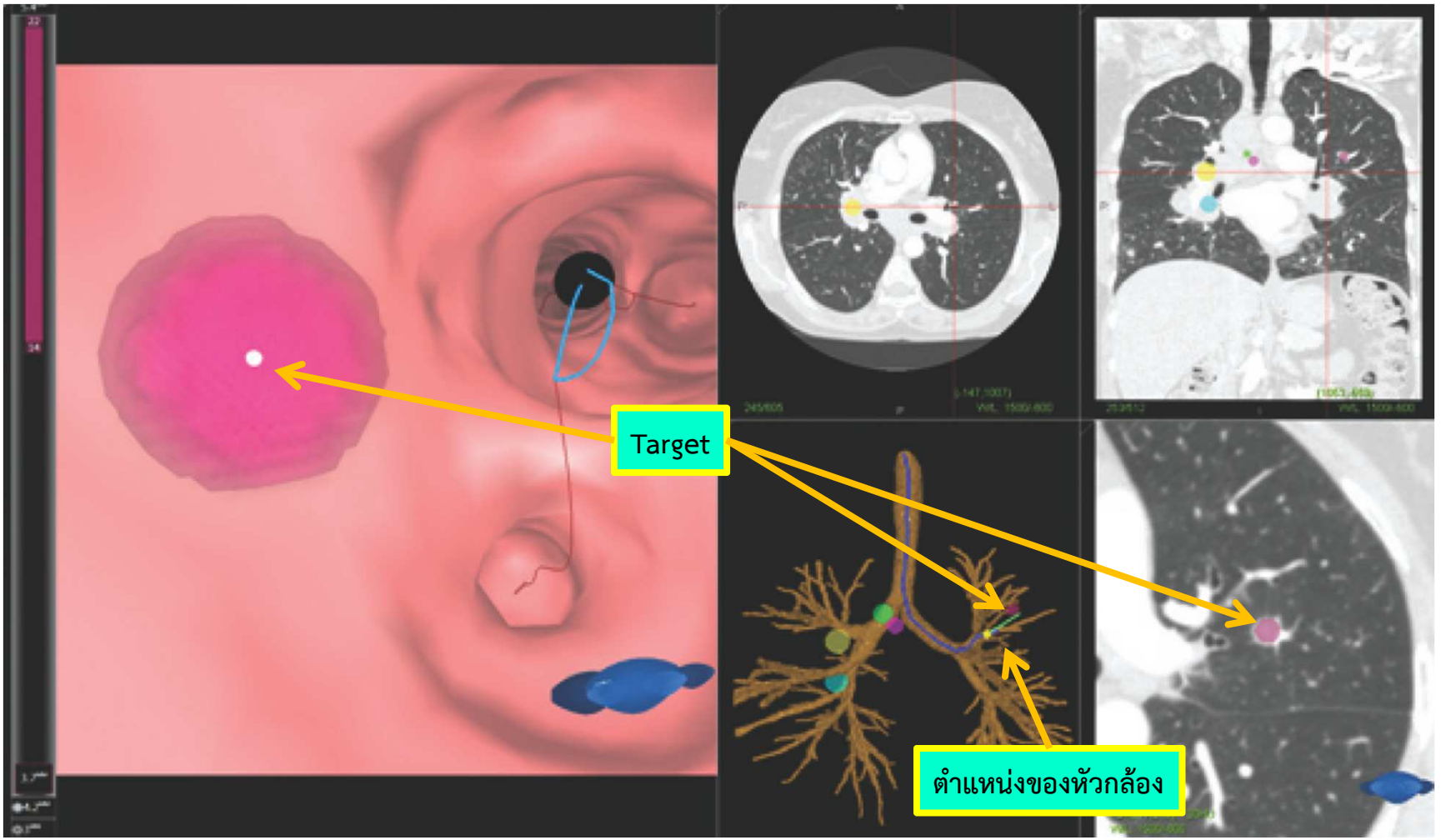
การวินิจฉัยมะเร็งปอด

- Penn State's Virtual Navigator and Tree Analyzer
- Broncus' LungPoint® Virtual Bronchoscopic Navigation



(Image Courtesy: Broncus' LungPoint Virtual Bronchoscopic Navigation)

LungPoint Virtual Bronchoscopic Navigation



(Image Courtesy: Broncus' LungPoint Virtual Bronchoscopic Navigation)

The system allows the physician to see a target behind the airway wall

Topic Outline

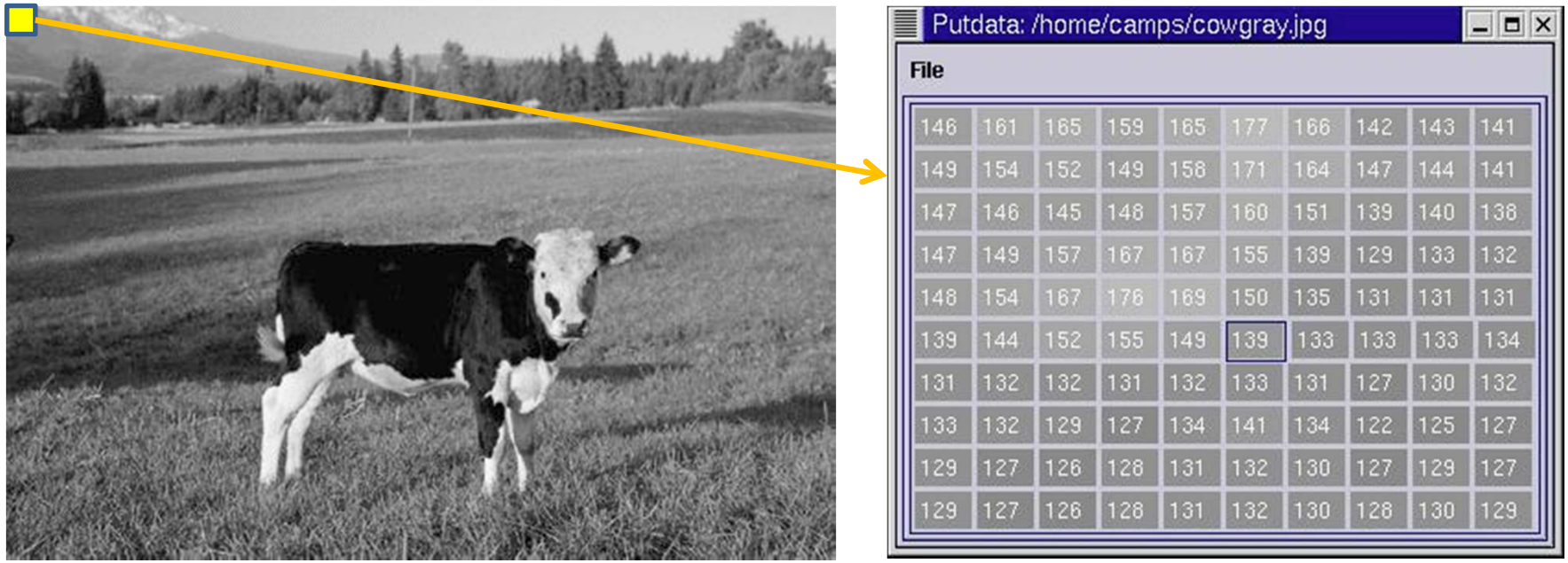
- ภาพและการมองเห็น
  - อะไรคือภาพ
  - อะไรคือการมองเห็น (ด้วยคอมพิวเตอร์) ?
- ความเกี่ยวข้องของวิชาการมองเห็นของคอมพิวเตอร์และศาสตร์อื่น ๆ
- ทำไมถึงต้องมีการเรียนรู้และวิจัยเกี่ยวกับการมองเห็นของคอมพิวเตอร์
- การประยุกต์ใช้คอมพิวเตอร์ในงานที่เกี่ยวกับการมองเห็น
- **รู้จักกับวิชาการมองเห็นของคอมพิวเตอร์ให้มากขึ้น**
- เราจะเรียนอะไรบ้างในวิชานี้

Topic Outline สำหรับการทำความรู้จักวิชา

- อินพุตของปัญหาคืออะไร
- เรื่องวุ่น ๆ ในคอมพิวเตอร์วิชัน
- คณิตศาสตร์กับขีดจำกัดในการแก้ปัญหา
- ตามหาข้อมูลทางเรขาคณิตเพื่อแก้ปัญหา
- ผลของแสงกับภาพที่เกิดขึ้น
- การรับรู้ที่คลาดเคลื่อนของตามนุษย์ (‘imperfection’ of human vision)
- บทสรุปของคอมพิวเตอร์วิชันกับการแก้ปัญหา

อินพุต: ภาพดิจิทัล

- สำหรับภาพสองมิติก็คืออาเรย์สองมิติของตัวเลข

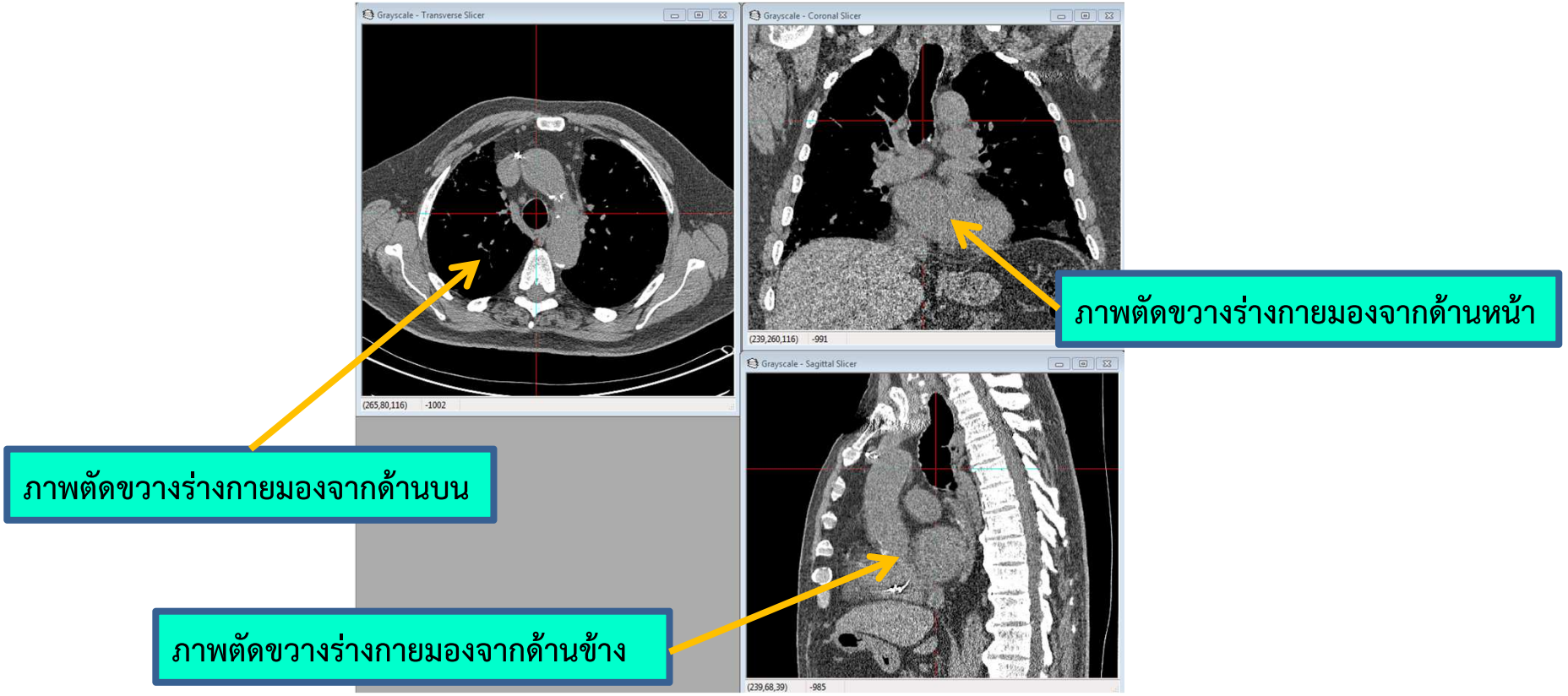


- ในกรณีที่เป็นภาพสี เราก็มักจะต้องจัดการกับอาเรย์สามตัวที่เก็บข้อมูลของสี แดง เขียว และ น้ำเงิน (RBG color format)
- องค์ประกอบย่อยแต่ละอันในภาพเรียกว่าพิกเซล (pixel, มาจาก picture + cell)
- ภาพวีดีโอทั่วไปก็สามารถจัดให้เป็นอาเรย์สองมิติหลาย ๆ เฟรมมาประกอบกัน

Image: Robert T. Collins

อินพุต: ภาพดิจิทัลสามมิติ

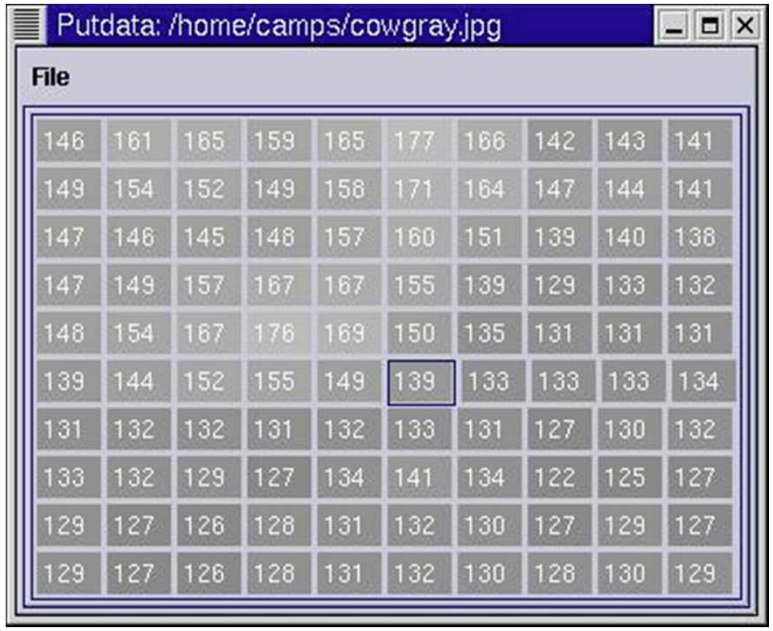
- ในกรณีที่เป็นภาพสามมิติก็เป็นอาเรย์สามมิติมาแต่แรก โดยมากจะจัดภาพเป็นแผ่นสองมิติ (2D slice) แล้วนำมาประกอบกันเป็นชั้น ๆ เพื่อให้ได้ภาพสามมิติ



- องค์ประกอบย่อยของภาพเรียกว่าวอกเซลล์ (voxel, มาจาก volume + cell)

ทำไมคอมพิวเตอร์วิชันถึงเป็นเรื่องยาก

พวกเรากำลังพยายามทำการอนุมานสิ่งต่าง ๆ ในโลกจากอาเรย์ของตัวเลข



ภาพบริเวณหลังของวัว?

ปัญหาในที่นี้ก็คือว่า: ข้อมูลเจาะจงกับพื้นที่มากเกินไป, ขาดข้อมูลเชิงบริบท, เลือกระดับการวิเคราะห์ข้อมูลที่ไม่เหมาะสม

แต่เดี๋ยวก่อน ที่จริงเรื่องที่มันวุ่นวายกว่านี้ก็มีอยู่อีก ...

Image: Robert T. Collins

เรื่องวุ่น ๆ ในคอมพิวเตอร์วิชัน

- เรามักจำแนกวัตถุจากข้อมูลเชิงเรขาคณิตที่เรารู้มาก่อนหน้า
- ถ้าเรารู้มาแต่แรกว่าวัตถุแต่ละอันมีลักษณะทางเรขาคณิตเป็นยังไง มีพื้นผิวทำจากอะไร แสงในสภาพแวดล้อมเป็นอย่างไร มุมมองของผู้สังเกตอยู่ที่ไหน

  ➔ มันก็เป็นที่เข้าใจเป็นอย่างดีในโลกของคอมพิวเตอร์กราฟฟิกว่าเราจะกำหนดให้พิกเซลแต่ละอันมีค่าเท่าไหร่ (เหมือนที่เราเห็นในเกม)

  ➔ เราก็น่าจะตอบได้โดยง่ายว่าของในภาพคืออะไร (หรือว่าไม่ใช่)?

- ปัจจัยทางด้านกราฟฟิกเมื่อมารวมกันแล้วเกิดเป็นพิกเซลจริง แต่เราไม่สามารถจำแนกปัจจัยนั้นจากค่าของพิกเซลได้ เพราะกระบวนการกำเนิดภาพเป็นกระบวนการที่ย้อนกลับไม่ได้ (irreversible process)

  ➔ การจะตอบได้ว่าของในภาพคืออะไร ยังไงก็เป็นเรื่องยากอยู่ดี

แล้วการมองเห็นของมนุษย์ล่ะ

คุณบอกได้มั้ยว่าในภาพนี้มีอะไรอยู่บ้างสักสามอย่าง

Slide credit: Robert T. Collins

Congratuations!

ยินดีด้วยคุณสามารถทำสิ่งที่เป็นไปไม่ได้ในทางคณิตศาสตร์

มีอะไรเกิดขึ้น?

คุณได้ใช้สมมติฐานที่มีรากฐานอยู่บนความรู้ก่อนหน้า (prior knowledge) และประสบการณ์ (experience) เกี่ยวกับวิถีและความเป็นไปในโลก

ทำไมล่ะ?

การกู้โครงสร้างเชิงสามมิติกลับมาจากภาพเพียงภาพเดียวเป็นปัญหาที่ผิดรูปทางคณิตศาสตร์

สรุปก็คือ ...

เราไม่สามารถแก้ปัญหาบางอย่างในคอมพิวเตอร์วิชันด้วยคณิตศาสตร์ ☹

Slide credit: Robert T. Collins

มารู้จักกับปัญหาทางคณิตศาตร์ที่แก้ได้และไม่ได้กันก่อน

ปัญหาที่ถูกรูปแบบ (Well-posed problem) มาจากนิยามโดย Jacques Hadamard ซึ่งเขาเชื่อว่า แม่แบบทางคณิตศาสตร์ (mathematical models) ของปรากฏการณ์ทางกายภาพควรจะมีคุณสมบัติสามข้อต่อไปนี้

1. มีผลเฉลย
2. มีผลเฉลยเฉพาะ (Unique solution)
3. ผลเฉลยขึ้นอยู่อย่างต่อเนื่องกับข้อมูลในโทโพโลยีที่สมเหตุผล

ถ้าแม่แบบทางคณิตศาสตร์ขาดคุณสมบัตินี้ไปแม้แต่เพียงอย่างเดียวจะถือว่าเป็นปัญหาที่ผิดรูปแบบทางคณิตศาสตร์ (mathematically ill-posed)

สาเหตุหลักของการเกิด ill-posed problem

โดยทั่วไปแล้ว ill-posed problem มักจะเกิดขึ้นจากการที่เราพยายามทำการอนุมาน กฎเกณฑ์หรือผลเฉลยจากจำนวนตัวอย่างที่จำกัด

➔ แต่อันที่จริงกฎเกณฑ์หรือผลเฉลยเฉพาะอาจจะไม่มีก็ได้

เราสามารถสรุปสาเหตุหลักของการเกิดปัญหา ill-posed ได้ดังนี้

- ขอบเขตของสมมติฐานกว้างเกินไป
- ข้อมูลไม่พอ
- มีปัจจัยที่เกี่ยวข้องกับปัญหาที่เราไม่รู้จัก (hidden variables)

แล้วเราจะแก้ปัญหาพวกนี้ยังไงดี?

Prior Knowledge

เราสามารถแก้ปัญหา ill-posed ได้ด้วยการผนวก prior knowledge เข้าไปในการแก้ปัญหา

ความเป็นไปได้ในการใช้ Prior knowledge กับการแก้ปัญหา ill-posed

1. จำกัดขอบเขตของสมมติฐานทำให้มีผลเฉลยเฉพาะ
2. เลือกเฉพาะผลเฉลยที่ดีที่สุดจากผลเฉลยทั้งหมด

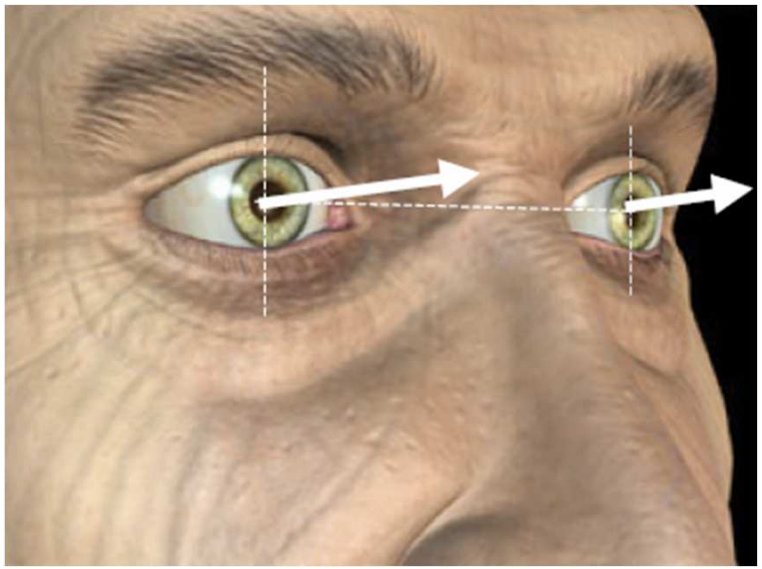
- เช่น หากเราระบุขอบเขตไปว่าสิ่งที่ยืนสองขาลำตัวตั้งตรงได้คือ มนุษย์ เท่านั้น และห้ามนับหุ่นขี้ผึ้ง หุ่นยนต์ (คนเหล็ก iRobot) ฯลฯ ผลเฉลยเฉพาะก็อาจจะมีก็ได้

- ถ้าเราบอกว่าเลือกเอาเฉพาะที่คิดว่าใช่ที่สุดให้เป็นผลเฉลยเฉพาะไปเลยก็ได้เหมือนกัน

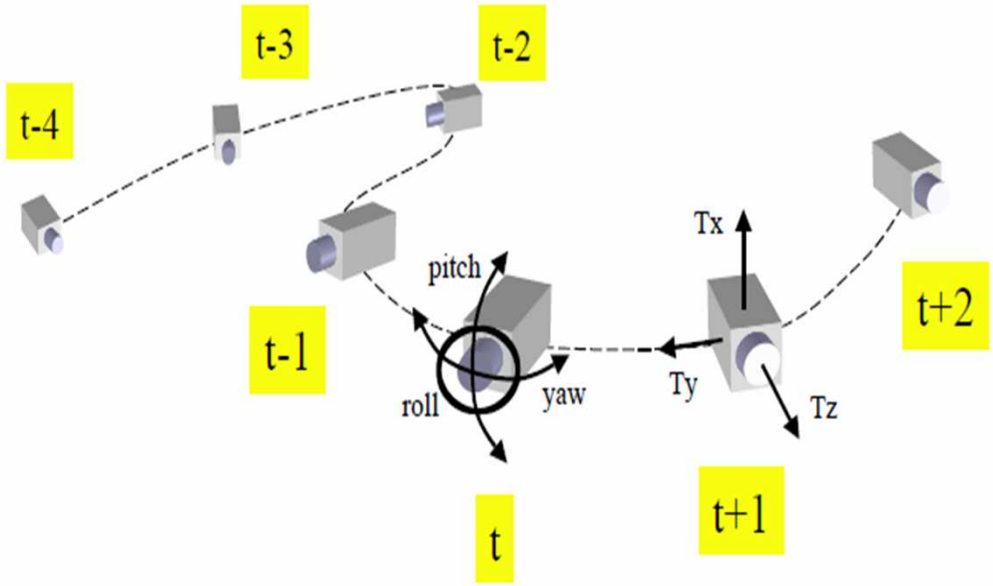
- Prior knowledge แก้ปัญหาทางการคำนวณได้ แต่ไม่ได้บอกว่าเราจะได้คำตอบที่ถูกต้องที่สุดออกมาจากการทำปัญหา ill-posed ให้เป็น well-posed

ตามหาข้อมูลทางเรขาคณิต

- เราสามารถเก็บข้อมูลเชิงสามมิติมาวิเคราะห์เป็นข้อมูลทางเรขาคณิตได้
- แต่กล้องเรารับภาพเป็นสองมิติ แล้วเรื่องนี้มันจะเป็นไปได้อย่างไร?
- มีอยู่สองวิธีหลัก ๆ ที่ใช้กันอยู่มาก คือ stereo vision และ structure from motion



Stereo vision ใช้กล้องสองตัวขึ้นไป



Structure from motion เป็นการเคลื่อนกล้องตัวเดียวเพื่อจับภาพหลาย ๆ มุม

Slide credit: Robert T. Collins

แล้วต้องเคลื่อนกล้องแค่ไหนถึงจะพอ

- เรื่องของ Illusion ทำให้เราถูกหลอกได้ง่าย ๆ
- การหมุนกล้องไปเพียงเล็กน้อยใช่ว่าจะช่วยได้ ซ้ำยังอาจจะเข้าใจผิดไปด้วย

Video courtesy: http://www.grand-illusions.com/opticalillusions/dragon_illusion/

เรื่องวุ่น ๆ เกี่ยวกับการหาข้อมูลจากหลายมุมกล้อง

Slide credit: Robert T. Collins

Human Perception: Color Constancy

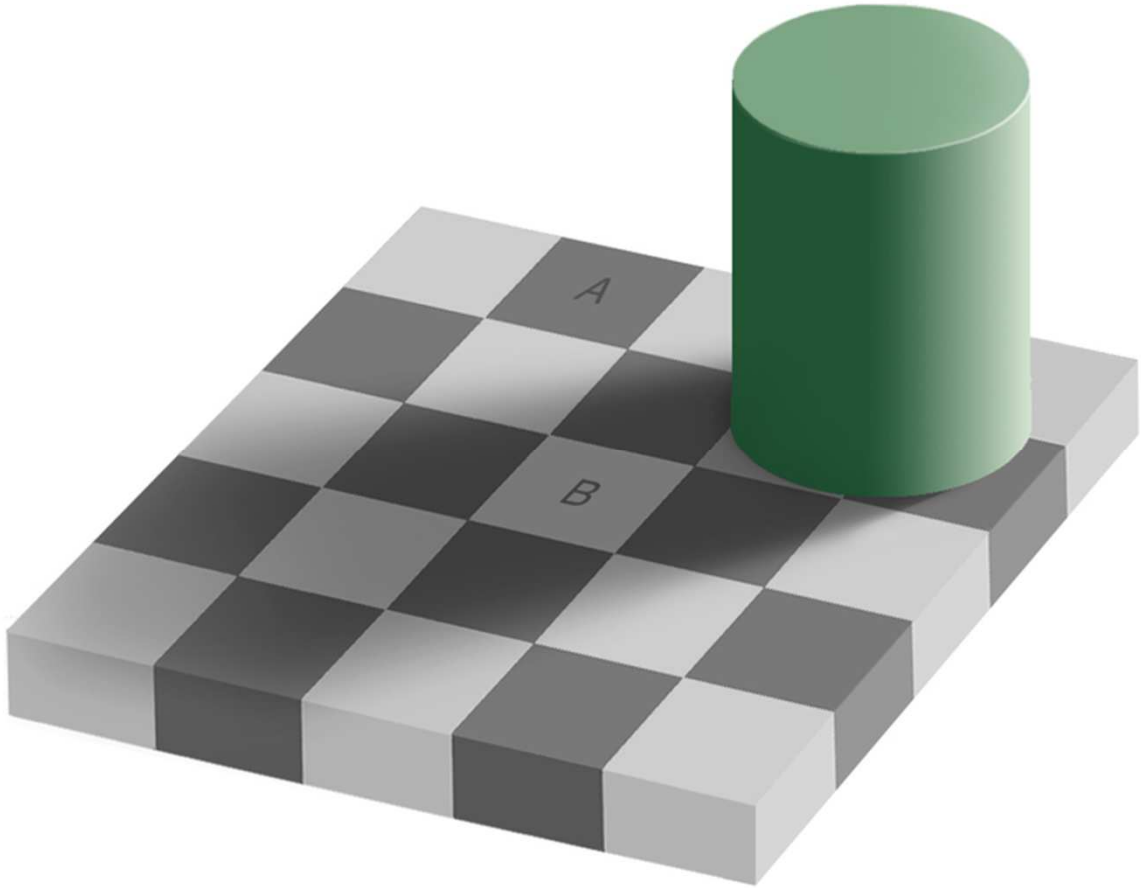
- กลไกทางการมองเห็นของมนุษย์มีจุดอ่อนที่ทำให้เราเข้าใจผิดได้ง่าย โดยเฉพาะเรื่องของ color constancy

- **Color constancy** is an example of subjective constancy and a feature of the human color perception system which ensures that the perceived color of objects remains relatively constant under varying illumination conditions [wikipedia].

- คอมพิวเตอร์ดูเหมือนจะเข้าใจอะไรได้ยาก (ต้องเขียนโปรแกรมสอนมันยาวมาก) แต่มันก็ไม่ถูกหลอกด้วยเรื่อง color constancy

- แล้วเราควรจะเชื่อตาของเราหรือตาของคอมพิวเตอร์?

Color Constancy Example

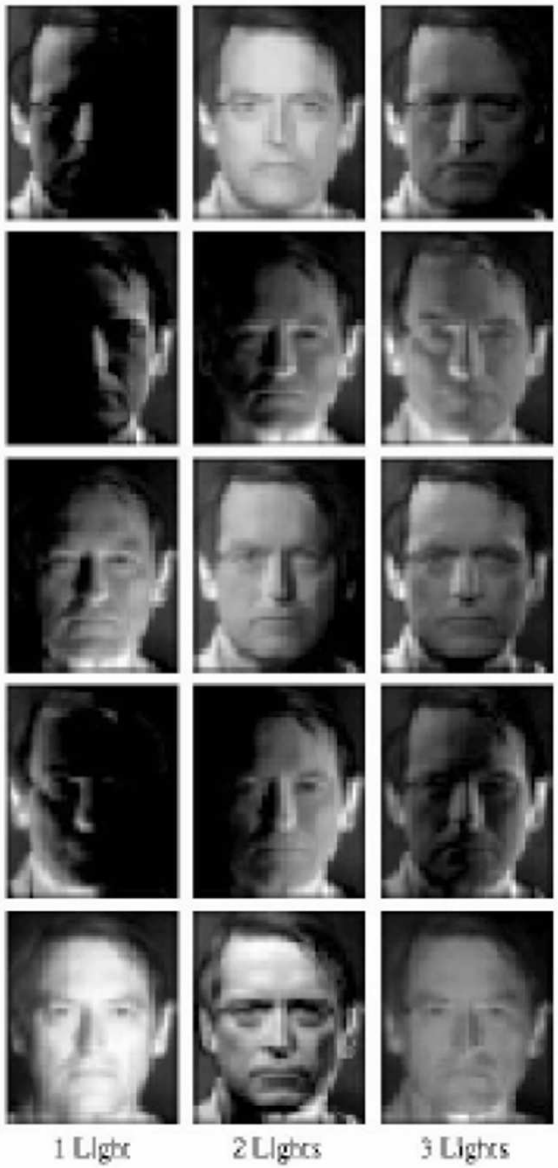


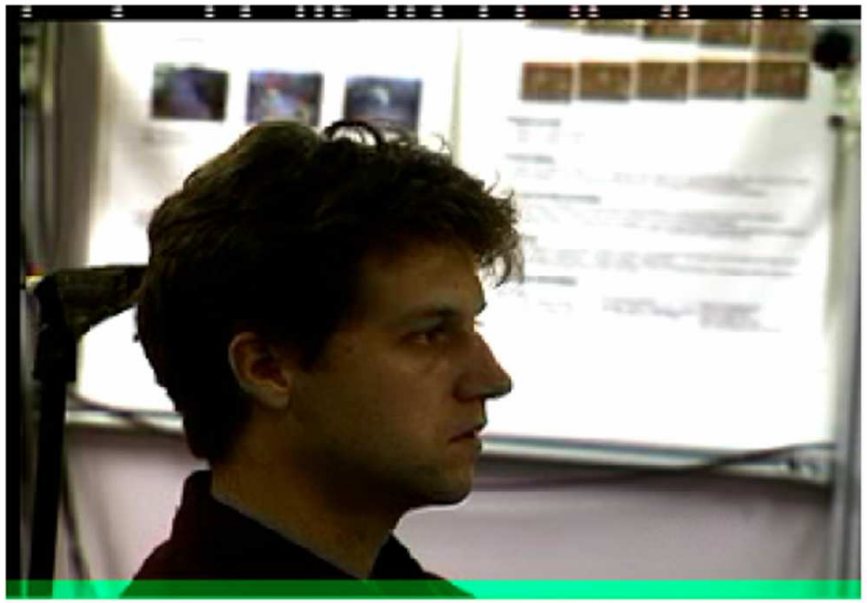
ถ้ามีคนบอกว่าสีในช่อง A และ B เป็นสีเดียวกัน คุณจะเชื่อหรือไม่?

พิสูจน์ได้ด้วยตัวเองผ่านการใช้โปรแกรม XnView (กด Ctrl+Shift+I เพื่อแสดงข้อมูลเกี่ยวกับสี)

Image credit: Edward H. Adelson (obtained from Wikipedia)

อิทธิพลของแสงกับภาพที่ปรากฏ





Object appearance also varies with respect to lighting magnitude and direction

Slide credit: Robert T. Collins

บทสรุปของคอมพิวเตอร์วิชันกับการแก้ปัญหา

- แม้ว่าคณิตศาสตร์ไม่สามารถนำมาใช้แก้ปัญหาบางอย่างในคอมพิวเตอร์วิชันได้ มันก็เป็นพื้นฐานของคอมพิวเตอร์วิชันที่สำคัญมาก

- ข้อมูลทางเรขาคณิตเป็นองค์ประกอบที่สำคัญมากในการแก้ปัญหาด้วยคอมพิวเตอร์วิชัน ไม่ต้องแปลกใจถ้าเราจะเห็นคณิตศาสตร์อยู่ทั่วไปในวิชานี้

- ถ้าเรารู้จักปัญหาดีพอ เราสามารถใช้ prior knowledge จำกัดขอบเขตที่โปรแกรมด้านคอมพิวเตอร์วิชันจะทำงาน ปัญหาก็อาจจะถูกแก้ไขได้ไม่ยากนัก

- Prior knowledge ที่เลือกใช้เป็นตัวกำหนดความสามารถของวิธีการในคอมพิวเตอร์วิชันที่สำคัญมาก ปัญหาจะแก้ได้ไม่ได้ส่วนหนึ่งขึ้นอยู่กับสมมติฐานที่เราเลือกใช้

สรุปภาพรวมเกี่ยวกับการมองเห็นของคอมพิวเตอร์

- “**Vision is the act of knowing what is where by looking.**” –Aristotle
- คอมพิวเตอร์วิชันมีการประยุกต์ใช้งานที่แพร่หลายในแทบทุกวงการ
- อินพุตมักเป็นภาพสองมิติ แต่ภาพสามมิติก็แพร่หลายสำหรับการใช้งานบางอย่าง
- เราต้องการข้อมูลทางเรขาคณิตมาตอบคำถามบ่อย ๆ แต่กระบวนการสังเคราะห์ภาพสองมิติทำให้ข้อมูลเรขาคณิตบางอย่างหายไป
- เราหาข้อมูลเชิงสามมิติได้จากการเปลี่ยนมุมมองในการรับภาพหรือใช้กล้องหลายตัว
- ข้อมูลทางด้านแสงและสีของวัตถุเป็นสิ่งที่ช่วยได้มากในหลาย ๆ ปัญหา แต่การจัดแสงในการเก็บภาพตัวอย่างทำให้ข้อมูลประเภทนี้มีความสำคัญน้อยลง
- ตาของมนุษย์ที่จริงก็มีปัญหาในการจำแนกวัตถุเหมือนกันโดยเฉพาะเรื่องสี คอมพิวเตอร์ก็มีแต่มักจะเป็นปัญหาคนละแบบ

เรื่องอื่น ๆ

- คอมพิวเตอร์วิชันเป็นวิชาที่ต้องใช้ไหวพริบและการเป็นคนช่างสังเกตในการแก้ปัญหา
- เป็นวิชาที่ผสมผสานระหว่างคณิตศาสตร์และความคิดสร้างสรรค์
- ปัญหาทางคอมพิวเตอร์วิชันจำนวนมากที่ยังไม่มีวิธีการใดที่โดดเด่นถึงขั้นที่ผู้คนรู้สึกพอใจในผลลัพธ์ที่ได้ เช่น การหาวัตถุที่มีความสมมาตร
- มีวิธีการทางคณิตศาสตร์บางอย่างสมองมนุษย์ไม่สามารถทำได้โดยง่าย ทำให้อัลกอริทึมด้านคอมพิวเตอร์วิชันบางอันมีความสามารถมากกว่ามนุษย์ได้