Computer Vision

• Computer Vision deals with recovery and use of information about objects present in a scene from images of the scene.

Computer Vision emerged from:

- Image Processing
- Pattern Recognition

Fairy Tale
• Computer Vision started as an AI problem.

AI

• Artificial Intelligence is the study of mental faculties through the use of computational models.
  – Search
  – NLU
  – Speech Recognition
  – Games
  – Computer Vision
  – Expert Systems

Image Understanding

• To understand a single image of a scene, locate and identify objects, their structure, and spatial arrangements, and relationships with other objects.

Different Levels

• Low Level: Extraction of symbolic information
• Intermediate Level
• High Level: Interpretation
High Level Vision

- Image Understanding
- Scene Interpretation
- Line Drawings

Interpretation of Line Drawing

MIT Copy Demo

What happened?

- In order to do line interpretation, need to extract lines from images
  - Horn-Binford line finder
  - Solve low level problems before high level problems can be solved.
Horn: Physics Based Vision

- Optics
- Reflectance
- Illumination

Marr Approach

- Human vision system
- Shape from X: Recover 3-D from 2-D
- Quantitative vs Qualitative

Shape from X

- Shading
- Stereo
- Texture
- Motion
- Contours

Shape from Texture
Shape from Shading

Shape from Stereo

Marr’s Three Levels

• Primal Sketch
  – Marr-Hildreth edge detector
• 2.5 Sketch
  – Marr-Poggio stereo algorithm
  – Grimson’s stereo algorithm
  – Ullman’s structure from motion
  – Pearsall, Witkin, Kass
  – Terzopoulos: surface reconstruction
• 3-D
  – Generalized Cylinders, Nishihara

After 30 Years of Research

• Stereo is almost a solved problem
• Structure from motion is very hard
• Shape from shading is not interesting/applicable
• Range images did not help much
• Not much progress in understanding/recognition/interpretation
Motion-Based Recognition

- A longer sequence leads to recognition of higher level motions, like walking or running, which consist of a complex and coordinated series of events that cannot be understood by looking at only a few frames.
  - 3-D is not necessary for recognition
  - Use motion directly for recognition vs
    - Recognition followed by reconstruction

Video Understanding

- Gestures
- Activities
- Facial expressions
- Visual Speech

- Applications
  - Video Surveillance and Monitoring
  - Perceptual User Interface
  - Model-based Video Compression
  - Augmented Reality and Video Games
  - Synthesis of Video Sequences

Copy Demo Using A Video Sequence:

Making a Sandwich

[bread, lettuce, ham, bread]

A picture is worth a thousand words.
A word is worth a thousand pictures.

Computer Vision: Outlook
- Video surveillance and monitoring
- Multi-media
- Computer Graphics

Computer Vision Text Books
History
• Started in August 1986
• Developed four courses
  – Intro to Robot Vision
  – Computer Vision
  – Computer Vision Systems
  – Advanced Computer Vision
• Graduated first Ph.D. student in 1989
• Dr. Lobo joined in 1992

Vision Books Used at UCF
Perceptual Organization and Visual Recognition
David Lowe
Vision Ph.D. Graduates
Donna J Williams, 1989


Krishnan Rangarajan, 1990


Ping-Sing Tsai, 1995

- Ping-Sing Tsai, Mubarak Shah, Katharine Keiter, and Takis Kasparis. "Cyclic Motion Detection for Motion Based Recognition", Pattern Recognition, Vol. 27, No. 12, 1994.

Ruo Zhang, 1995

Yu Tian, 1997


Niels Haering, 1999


Current Ph.D. Students

- Sohaib Khan
- Cen Rao
- Omar Javed
- Zeeshan Rasheed
- Alper Yilmaz
- Orkun Alatas
- Lisa Spencer
- Yaser Shaikh
- Jinruiyan Xue
- Yun Zhai

End of Story