COSC579: Computer Vision

Jeremy Bolton, PhD

Assistant Teaching Professor

A very special thanks to those who have contributed to this area of research over the years. Some slides used are from their research and efforts: the following professor and researchers Yung-Yu Chuang, Fredo Durand, Alexei Efros, William Freeman, James Hays, Svetlana Lazebnik, Andrej Karpathy, Fei-Fei Li, Srinivasa Narasimhan, Silvio Savarese, Steve Seitz, Noah Snavely, Richard Szeliski, and Li Zhang.

GEORGETOWN UNIVERSITY

Outline

I. Welcome!
II. Course Overview and Administration
III. Topics and Goals
IV. Fun Examples

GEORGETOWN UNIVERSITY

Welcome!

COSC-579: Computer Vision

Instructor: Jeremy Bolton, Ph.D.
Assistant Teaching Professor
Department of Computer Science
Email: jeremy.bolton@georgetown.edu

Office Hours: Daily hours will be entered on Canvas calendar (or by appointment)

TAs: TBD (see Canvas calendar for office hours)

GEORGETOWN UNIVERSITY

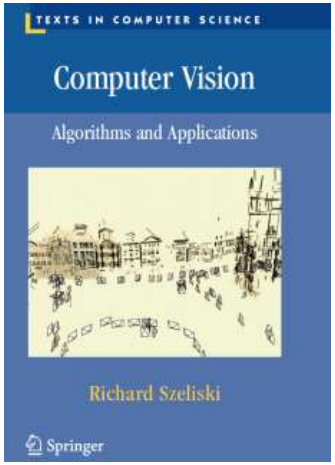
Course Summary

- **Course Description:** This course provides a comprehensive introduction to computer vision including image acquisition, low-level vision, and high-level vision. Image acquisition topics may include camera geometry, radiometry, illumination, noise, stereopsis, and affine transformations. Low-level vision topics may include, convolution, Fourier Transform, filters, operators, and feature generation. High-level vision topics may include detection, classification, segmentation, spatial relations, spatio-temporal models, object tracking, deformable models, and graph-based models.

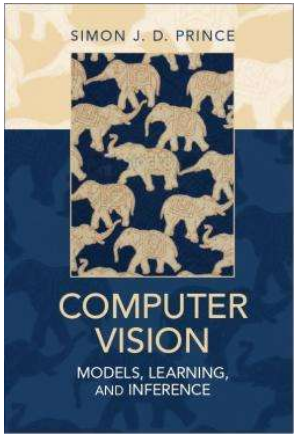
- **Required Prerequisites:** Mathematical Statistics and Linear Algebra

- *Some existing knowledge of Machine Learning and Image Processing is preferable, but a brief review will be provided.*

GEORGETOWN UNIVERSITY



Texts



- Szeliski, Computer Vision: Algorithms and Applications, Springer, 2010
  - http://szeliski.org/Book/
- Prince, Computer Vision: Models, Learning, and Inference, Cambridge, 2012
  - www.computervisionmodels.com
- Goodfellow, Bengio and Courville, Deep Learning, MIT Press, 2016.
  - http://www.deeplearningbook.org
- Nielsen, Neural Networks and Deep Learning, 2015.
  - http://neuralnetworksanddeeplearning.com/
- Horn, Berthold K.P. Robot Vision. The MIT Press, 1986.

Other References

- Forsyth and Ponce, Computer Vision: A Modern Approach, Prentice Hall, 2011.
- Bishop, Pattern Recognition and Machine Learning, Springer 2006.
- Vapnik, Statistical Learning Theory, Wiley, 2006.
- Duda, Hart, Stork, Pattern Classification, 2nd edition, Wiley, 2000.

GEORGETOWN UNIVERSITY

Course Website

- Link can also be found on Canvas or my department page.

http://jeremybolton.georgetown.domains/courses/cv/

GEORGETOWN UNIVERSITY

Exercises and Final Project

- **Exercises**
  - **Theory (math and stats) and Application (coding).**
- **Final Project: Implement Computer Vision Solution**
  - **Report**
  - **Presentation**

GEORGETOWN UNIVERSITY

Notes about Coding

- **Recommended Languages: Matlab or Python**
  - **Request approval for use of another language.**

- Coding Exercises are an integral part of this course! It is assumed that you have a proficient understanding of a programming language. Students are responsible for learning and/or reviewing, as needed, the programming language chosen.

- Matlab and Python have many packages which perform Computer Vision tasks. Depending on the exercise, you may or may not be permitted to use these built-ins (pre-existing code). Details will be provided in exercise instructions. If you have any questions about a built-in, simply ask.

- Cheating will not be tolerated. Any form of cheating will be reported to the GU honor council. Please read the following guidelines for project submissions:
  - Discussion among students pertaining to project content and general methodology is encouraged; however, students are NOT PERMITTED to share code, copy code, or use code composed by others.
  - A student may be asked to present, demonstrate, or explain a project submission at any time, without notice. At my sole discretion, a student's project grade can be adjusted based on this presentation, demonstration, and/or explanation. If a student does not sufficiently understand or explain their submission, further action may be taken.
  - Due Dates will be posted in Canvas or announced in class.

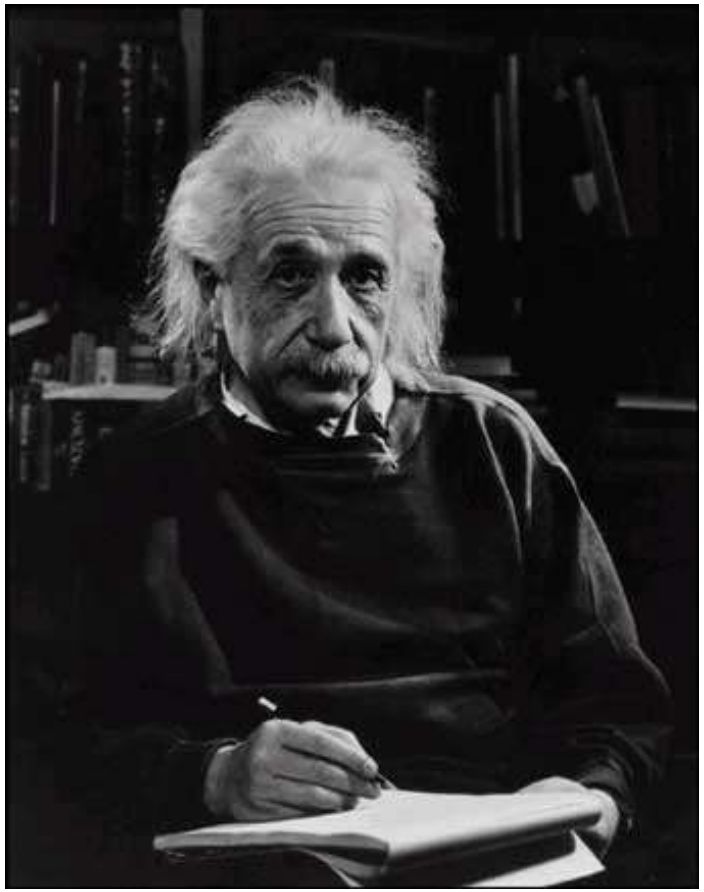
A Picture is Worth 100 Words

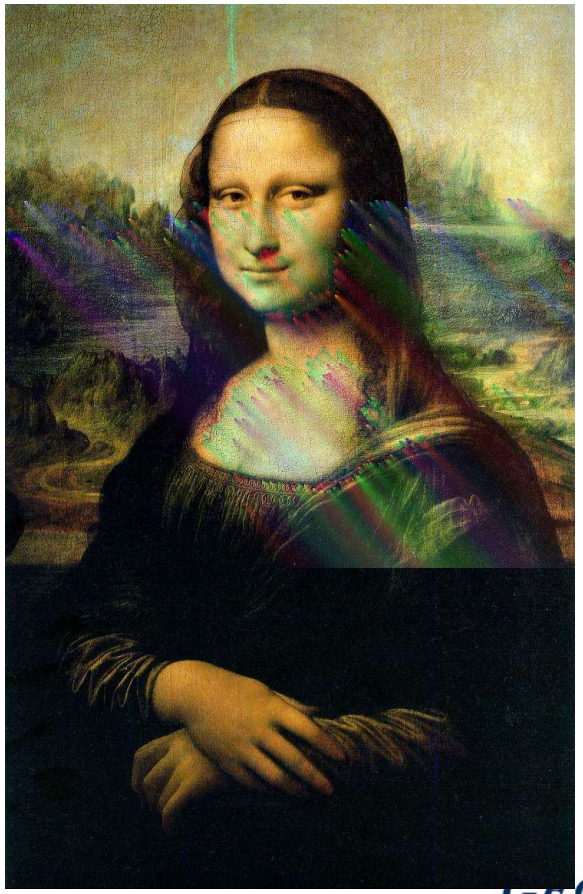
GEORGETOWN UNIVERSITY

A Picture is Worth 10,000 Words

GEORGETOWN UNIVERSITY

A Picture is Worth a Million Words



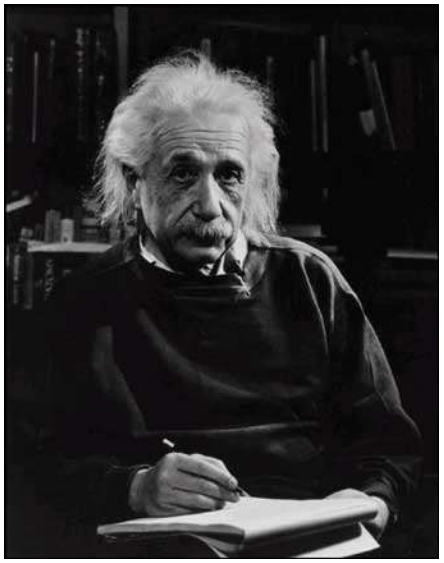


GEORGETOWN UNIVERSITY

Human Vision

- Can do amazing things like:
  - Recognize people and objects
  - Navigate through obstacles
  - Understand mood in the scene
  - Imagine stories
- But still is not perfect:
  - Suffers from Illusions
  - Ignores many details
  - Ambiguous description of the world
  - Doesn’t care about accuracy of world

Computer Vision



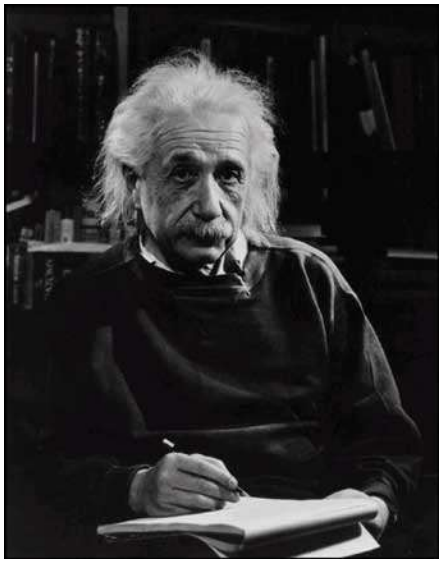
What we see

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

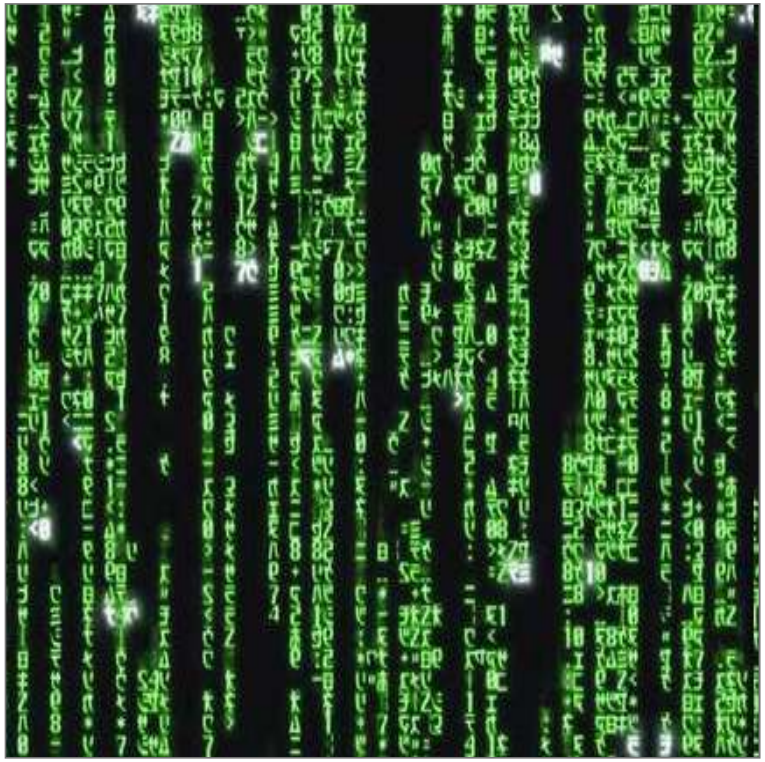
What a computer sees

GEORGETOWN UNIVERSITY

Computer Vision



What we see



What a computer sees

GEORGETOWN UNIVERSITY

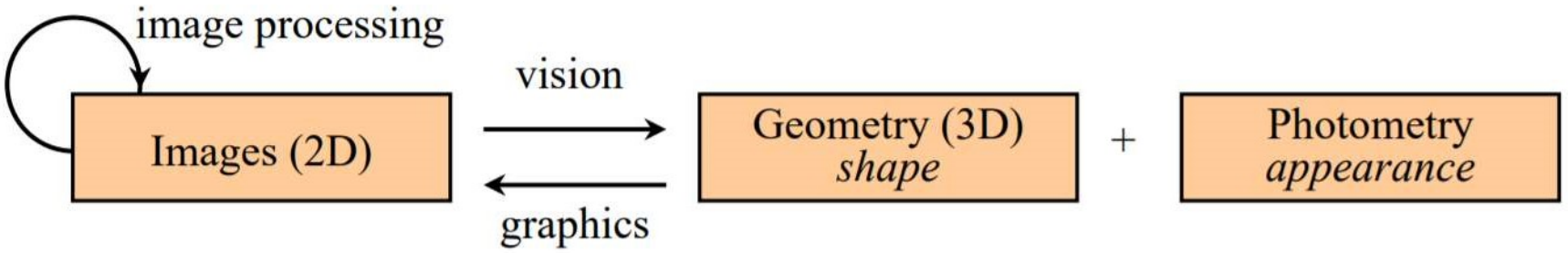
What is computer vision?

Terminator 2

GEORGETOWN UNIVERSITY

What is Computer Vision?

- Computer Vision
  - Inverse problem: Hard
  - Image Processing
- Graphics
  - Forward problem



GEORGETOWN UNIVERSITY

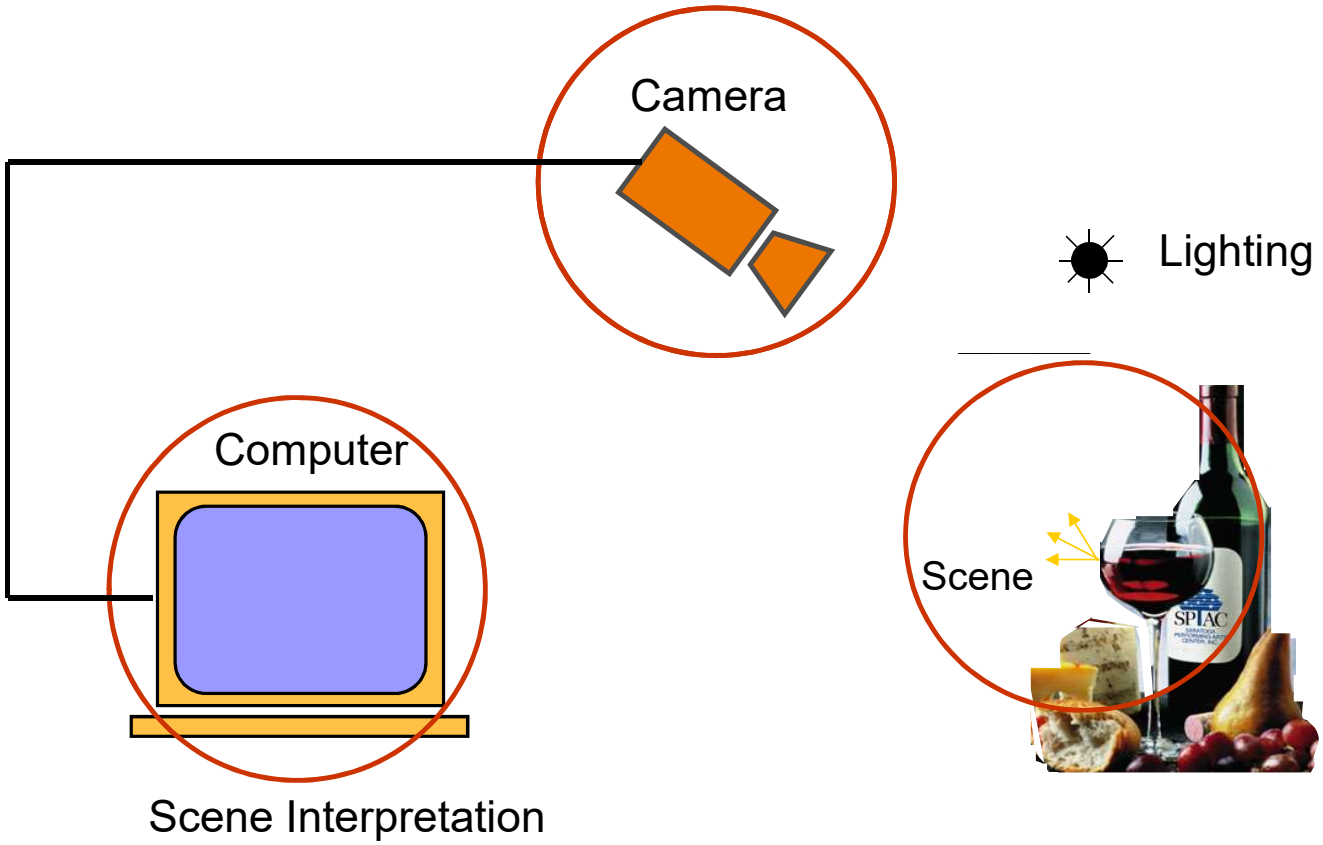
What is Computer Vision?

- Inverse Optics
- Intelligent interpretation of Imagery
- Building a Visual Cortex

- No matter what your definition is…
  - Vision is hard.
  - But is fun...

GEORGETOWN UNIVERSITY

Components of a Computer Vision System



GEORGETOWN UNIVERSITY

Every picture tells a story

- Goal of computer vision is to write computer programs that can interpret images

GEORGETOWN UNIVERSITY

Can computers match (or beat) human vision?

- Yes and no (but mostly no!)
  - humans are much better at “hard” things
  - computers can be better at “easy” things

GEORGETOWN UNIVERSITY

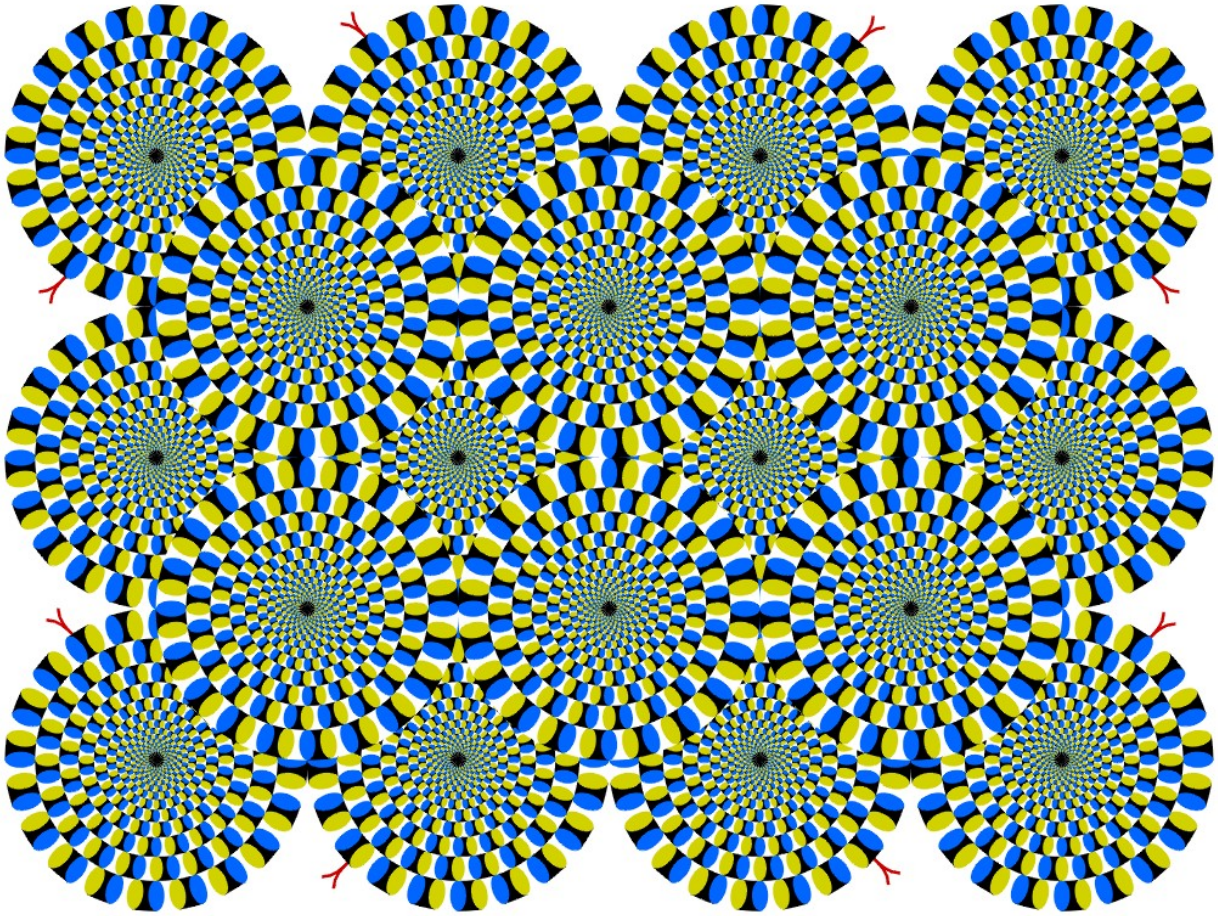
Human perception has its shortcomings…

Sinha and Poggio, *Nature*, 1996

GEORGETOWN UNIVERSITY

Is this image in motion?

- Look at the image as a whole. Are the wheels in motion?
- Now focus on one wheel in particular. Is that one in motion?



GEORGETOWN UNIVERSITY

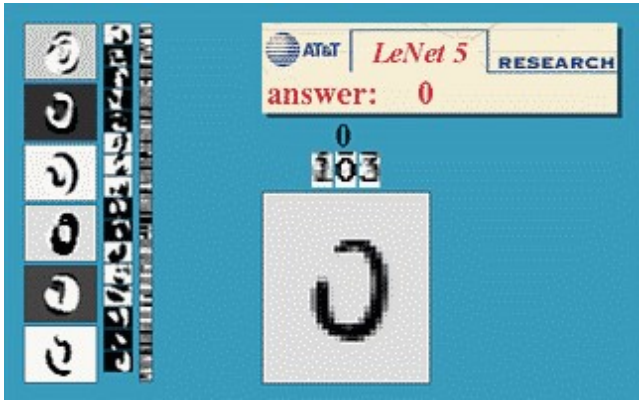
Some Topics in Computer Vision

GEORGETOWN UNIVERSITY

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

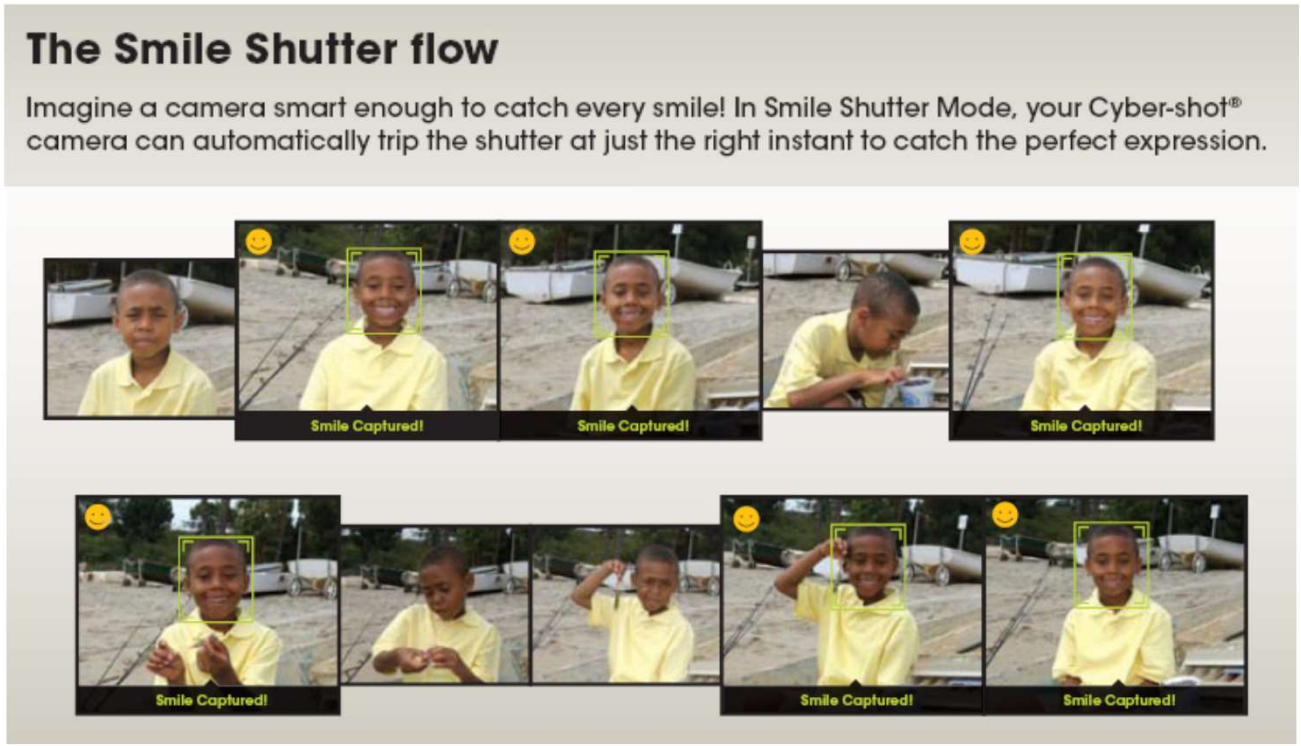
GEORGETOWN UNIVERSITY

Face detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

GEORGETOWN UNIVERSITY

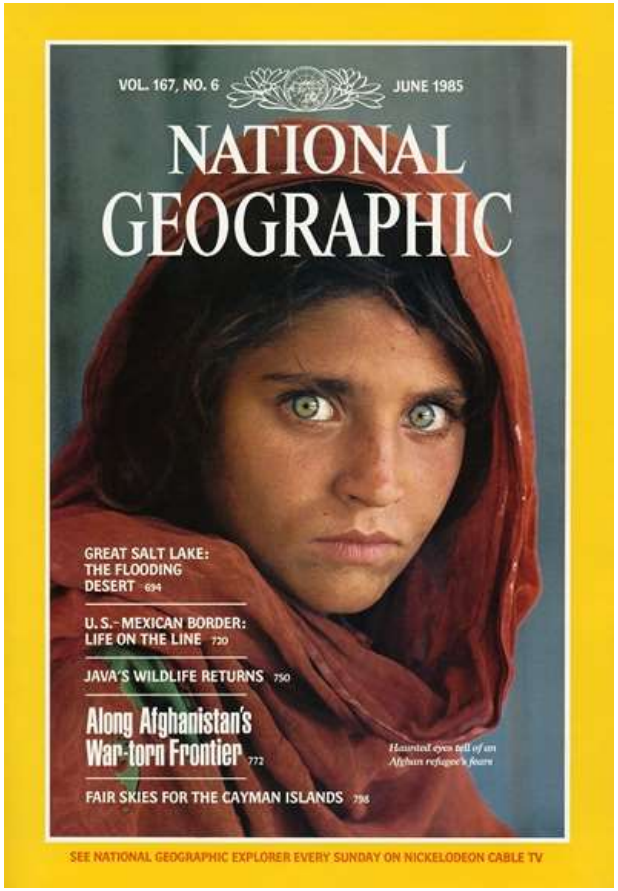
Smile detection?



Sony Cyber-shot® T70 Digital Still Camera

GEORGETOWN UNIVERSITY

Face recognition

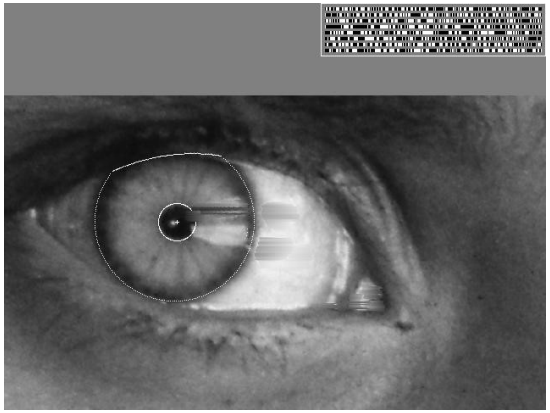


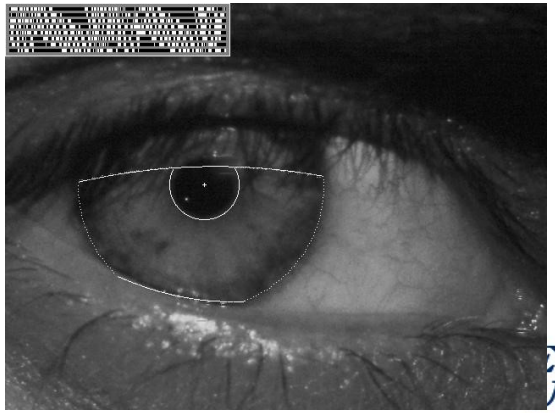
Who is she?

GEORGETOWN UNIVERSITY

Vision-based biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story





GEORGETOWN UNIVERSITY

Login without a password…

Fingerprint scanners on many new laptops, other devices

Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

GEORGETOWN UNIVERSITY

Object recognition (in mobile phones)

- This is becoming real:
  - Lincoln Microsoft Research
  - Point & Find, Nokia
  - https://www.google.com/intl/en_us/insidesearch/features/images/searchbyimage.html

GEORGETOWN UNIVERSITY

Shape and Morphology: Deformable Models

The Matrix movies, ESC Entertainment, XYZRGB, NRC

GEORGETOWN UNIVERSITY

Slide content courtesy of Amnon Shashua

Smart cars

- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers.
  - Video demo

GEORGETOWN UNIVERSITY

Vision in space

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

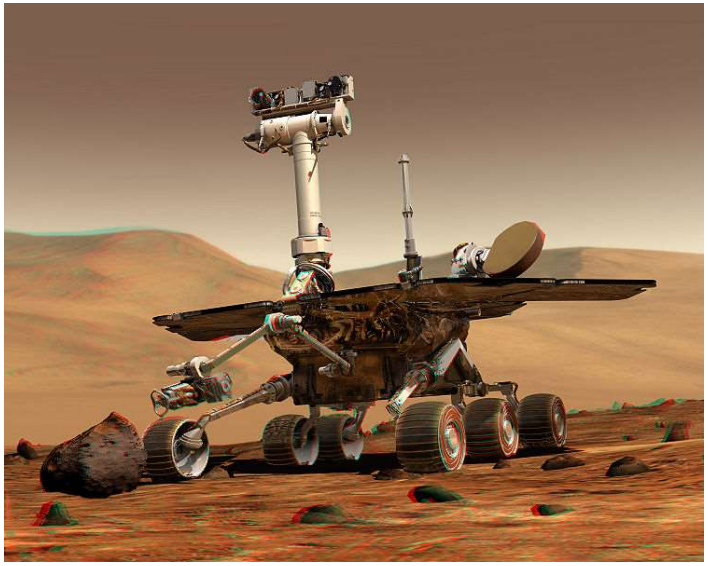
GEORGETOWN UNIVERSITY

Delivery Robots: Starship Technologies and more

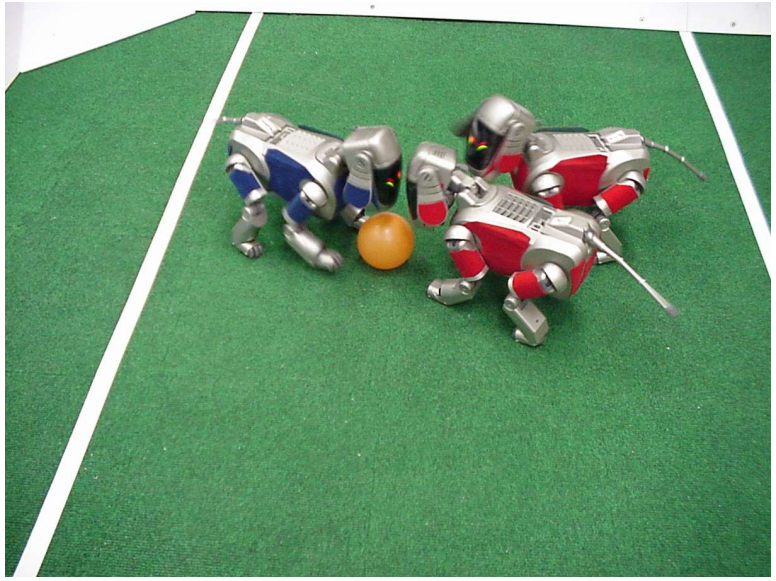
- **Computer Vision will transform the delivery industry**

GEORGETOWN UNIVERSITY

Robotics



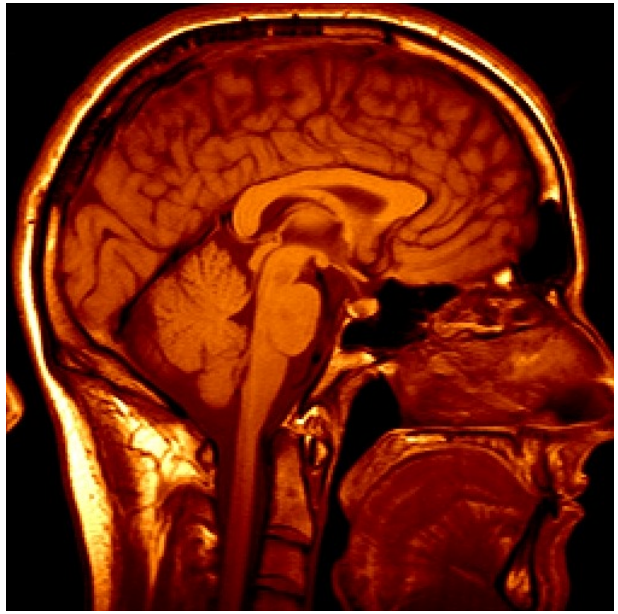
NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

GEORGETOWN UNIVERSITY

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

GEORGETOWN UNIVERSITY

Other Computer Vision Applications

- Some industrial applications of computer vision:
  - automotive monitoring, car counting http://www.mobileye.com/;
    - https://youtu.be/Y3ac5rFMNZ0?t=283
  - Surveillance. Fight, Flight Detection / Target Tracking
    - https://youtu.be/QcCjmWwEUgg
    - https://youtu.be/InqV34BcheM
  - Sports Data: https://www.secondspectrum.com/
  - Morphing:
    - https://youtu.be/pqpS6BN0_7k
    - https://youtu.be/nUDIoN-_Hxs

GEORGETOWN UNIVERSITY

Current state of the art

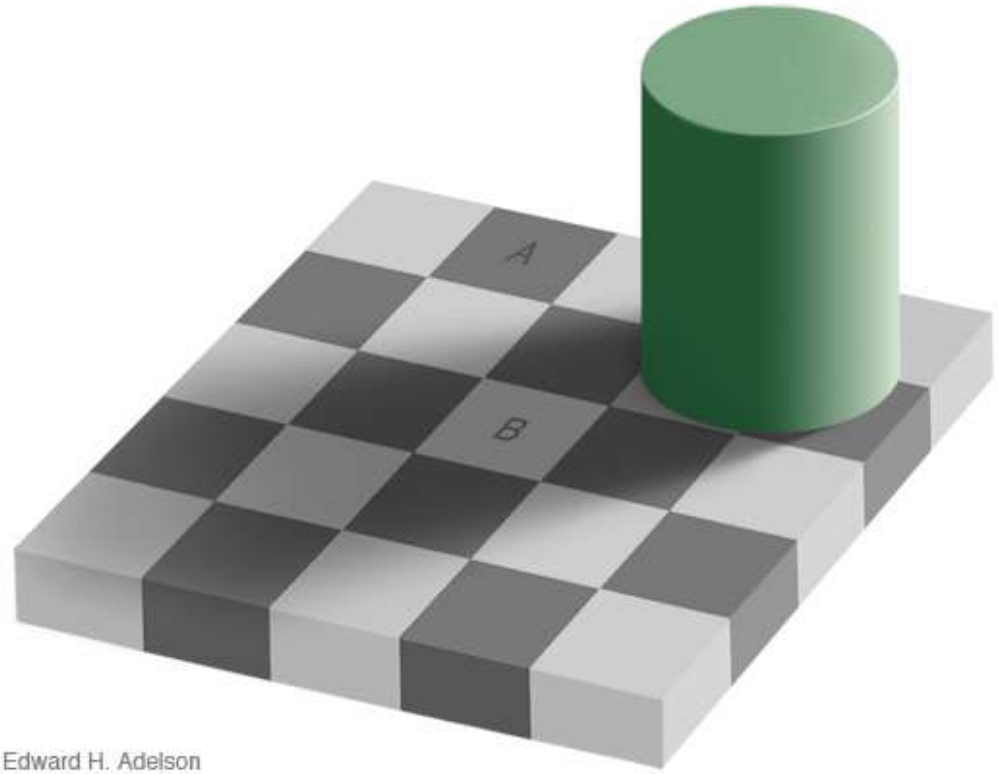
- You just saw examples of current applications.
  - Many of these are less than 5 years old

- This is a **very active** research area, and **rapidly changing**
  - Many new apps in the next 5 years

- To learn more about vision applications and companies
  - David Lowe maintains an excellent overview of vision companies
    - http://www.cs.ubc.ca/spider/lowe/vision.html

Topics covered

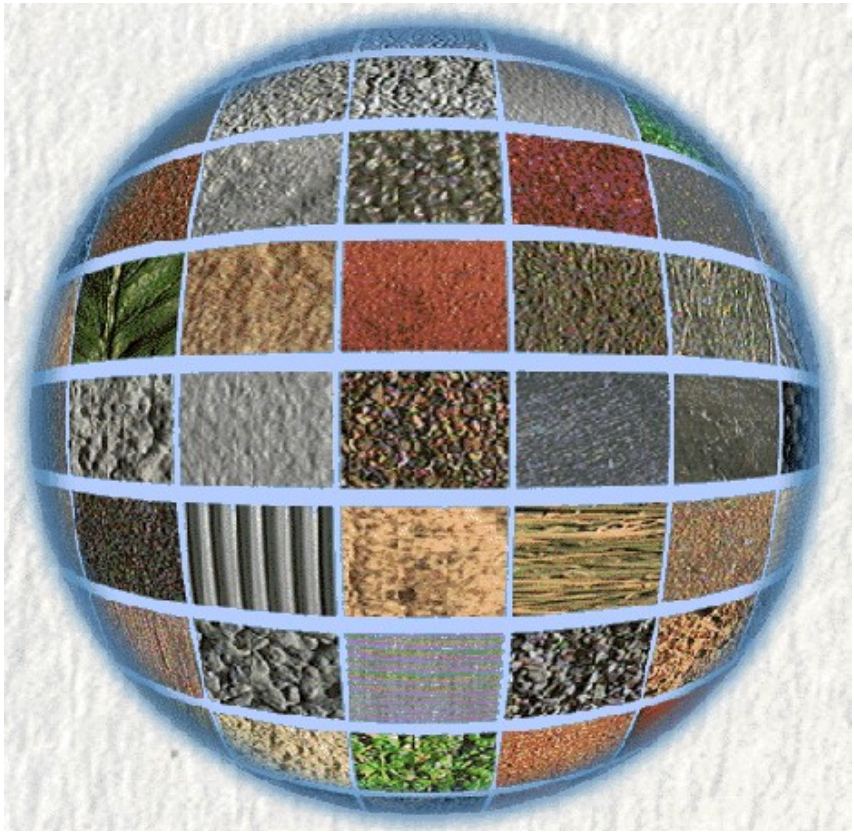
Lightness and Perception



Edward H. Adelson

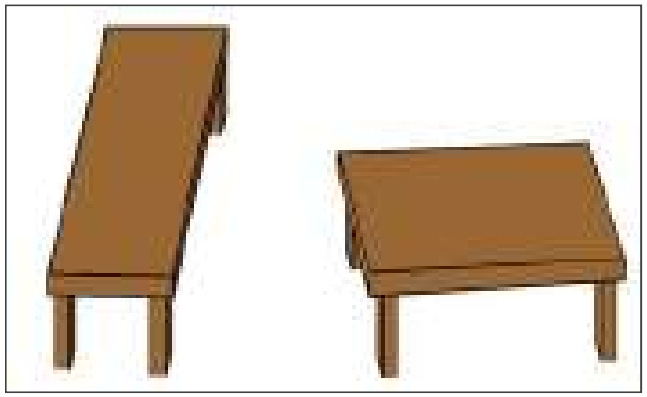
GEORGETOWN UNIVERSITY

Surface Reflectance

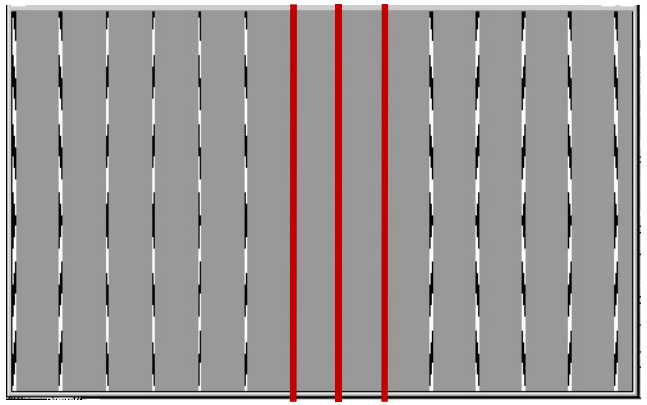


[CURET]

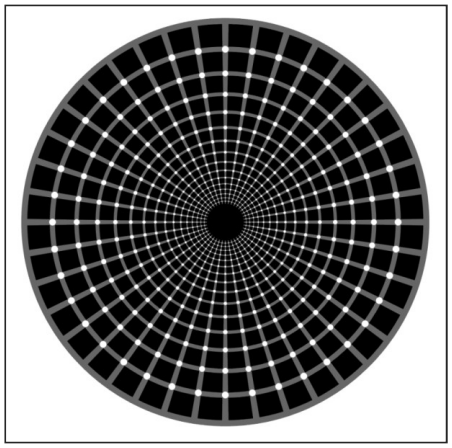
Human Vision: Optical Illusions



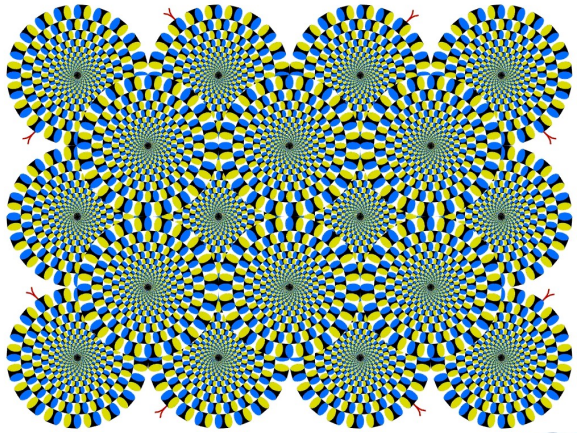
Which is bigger?



Straight Lines?



Dots White? Or Black?

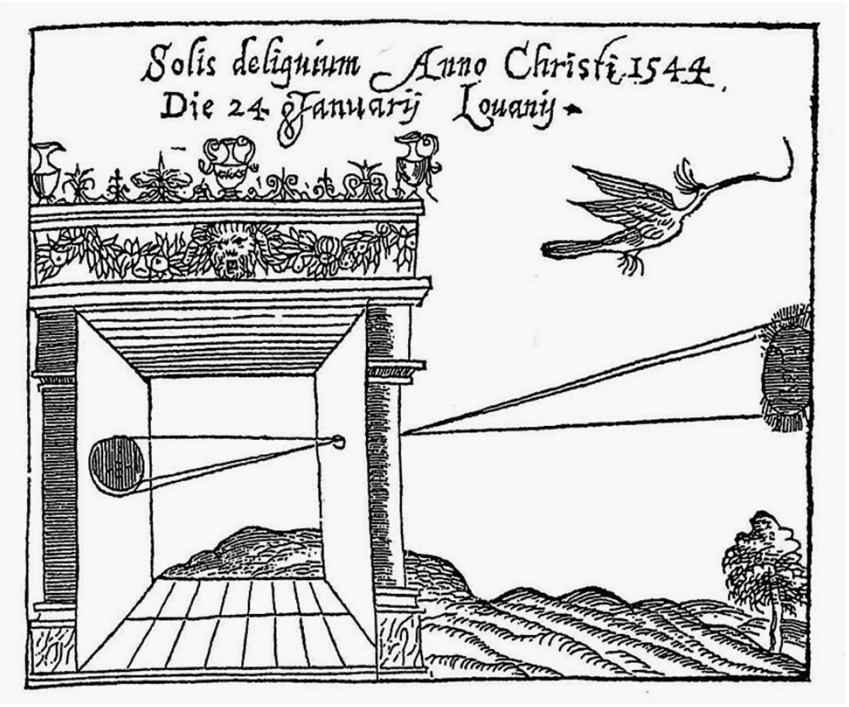


Spinning Wheels?

GEORGETOWN UNIVERSITY

Cameras and their Optics

Today's Digital Cameras



The *Camera Obscura*

GEORGETOWN UNIVERSITY

Binocular Stereo

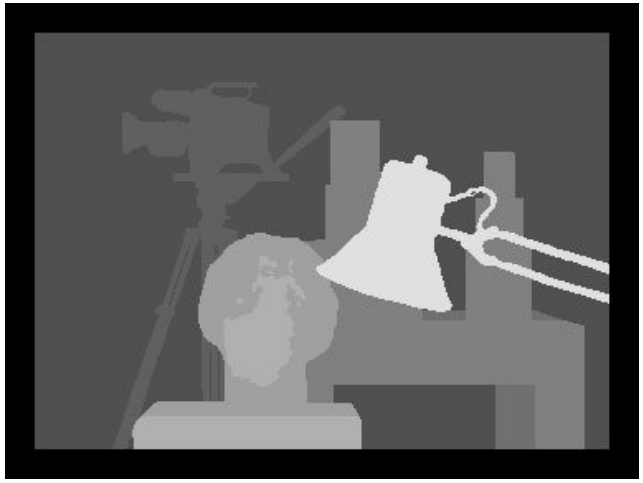
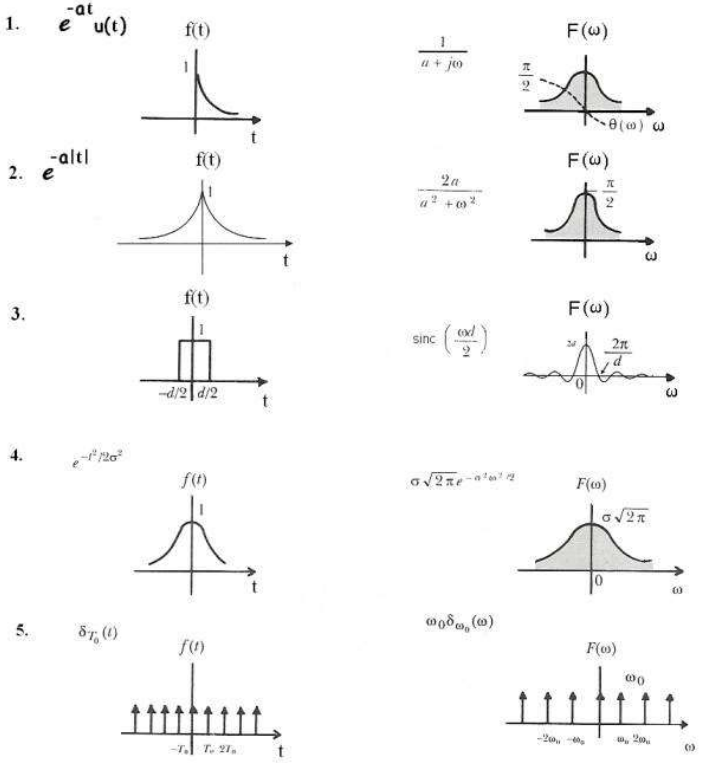


Image Processing

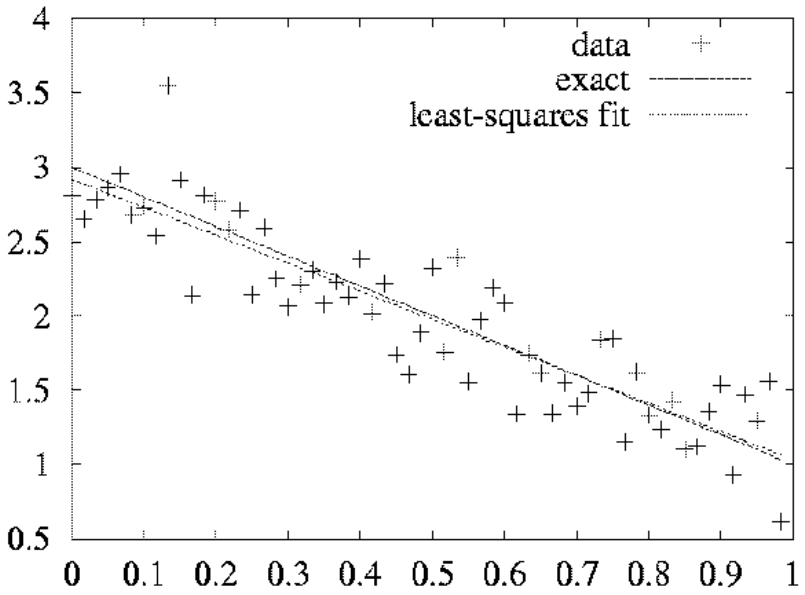


Fourier Transform
Sampling, Convolution

Image enhancement
Feature detection

GEORGETOWN UNIVERSITY

Statistical Techniques



Least Squares Fitting

GEORGETOWN UNIVERSITY

Face detection

GEORGETOWN UNIVERSITY

Face Recognition

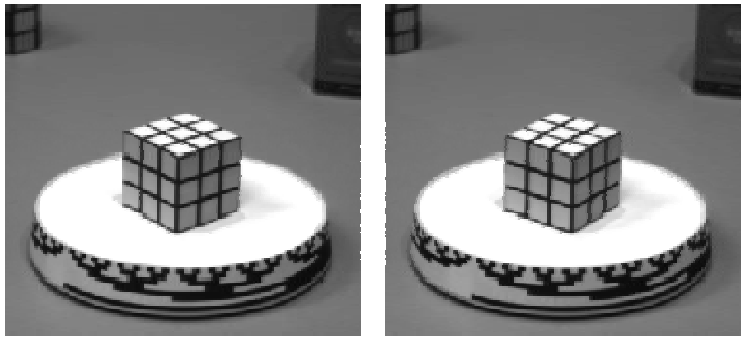
- Principle Components Analysis (PCA)
- Face Recognition

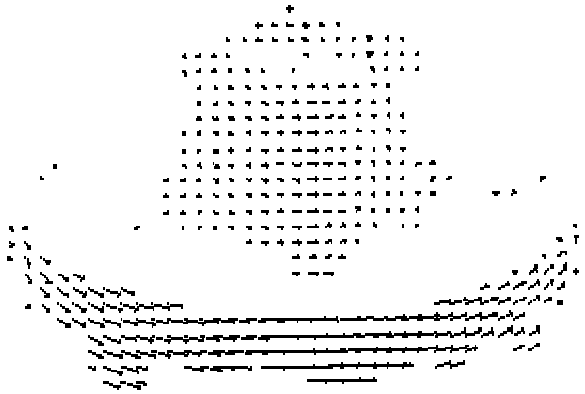
GEORGETOWN UNIVERSITY

Tracking

GEORGETOWN UNIVERSITY

Optical Flow



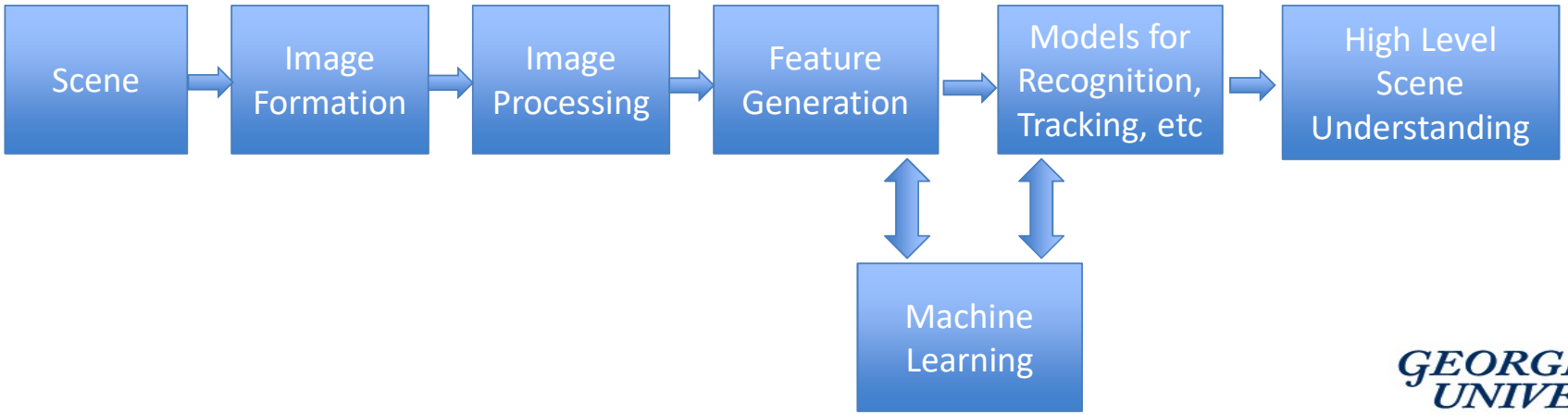


Some Recent Trends in Vision

GEORGETOWN UNIVERSITY

Computer Vision … the journey

- Attaining scene understanding is a long journey
- Our course will largely develop these concepts in their intuitive order



GEORGETOWN UNIVERSITY

Appendix

GEORGETOWN UNIVERSITY