Computer Vision

ITCS 4152 / 5152

- Me: Shaoting Zhang
  - Office: 210F Woodward
  - Email: szhang16@uncc.edu
  - Course materials: Adapted from Dr. Richard Souvenir (was in UNCC for ten years, now in Temple Univ.)
- Grader: Xiaofan Zhang
  - Email: xzhang35@uncc.edu

Introduction to Computer Vision

Slides adapted from James Hays, Steve Seitz, Amnon Shashua

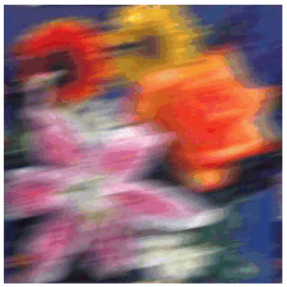
What is computer vision?

- Trucco and Verri: computing properties of the 3D world from one or more digital images
- Stockman and Shapiro: To make useful decisions about real physical objects and scenes based on sensed images
- Ballard and Brown: The construction of explicit, meaningful description of physical objects from images
- Forsyth and Ponce: Extracting descriptions of the world from pictures or sequences of pictures

Related Fields

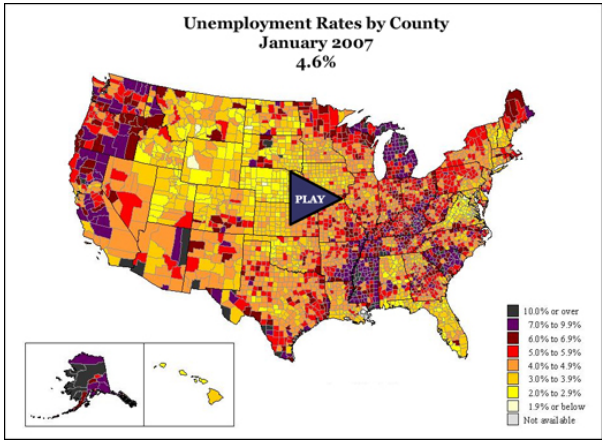
- Image processing
  - images → images
- Visualization
  - data → images
- Graphics
  - models → images

- *Computer Vision*
  - *images → models*



Blurred image

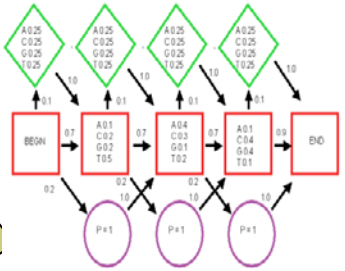
Image restored by Wiener filter



Images / Video

Vision

Graphics

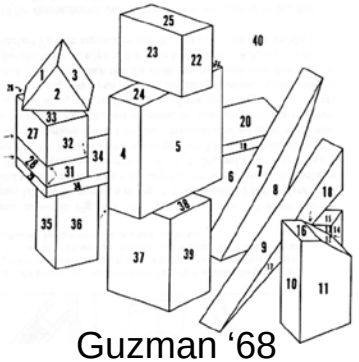


Models

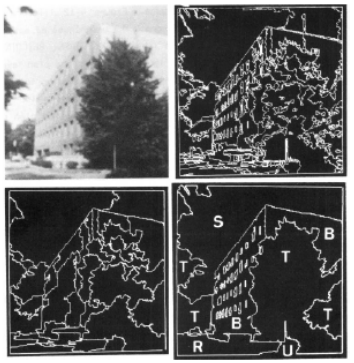
Ridiculously brief history of computer vision

- David Marr (1945-1980): integrated results from psychology, artificial intelligence, and neurophysiology into new models of visual processing. (Marr Prize in ICCV)

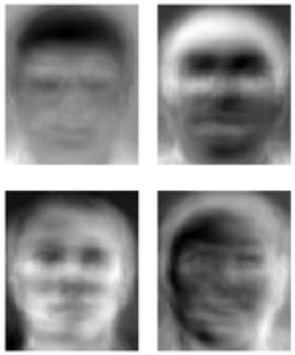
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- Now: Deep learning such as deep CNN, Examples of state-of-the-art



Guzman '68



Ohta Kanade '78

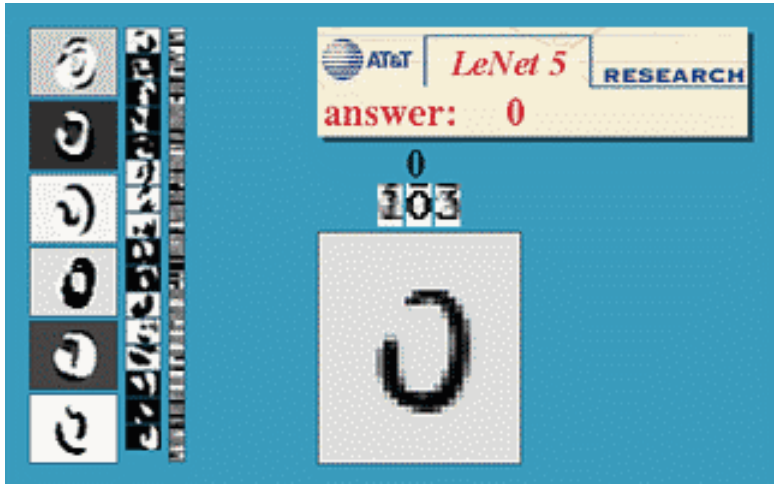


Turk and Pentland '91

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

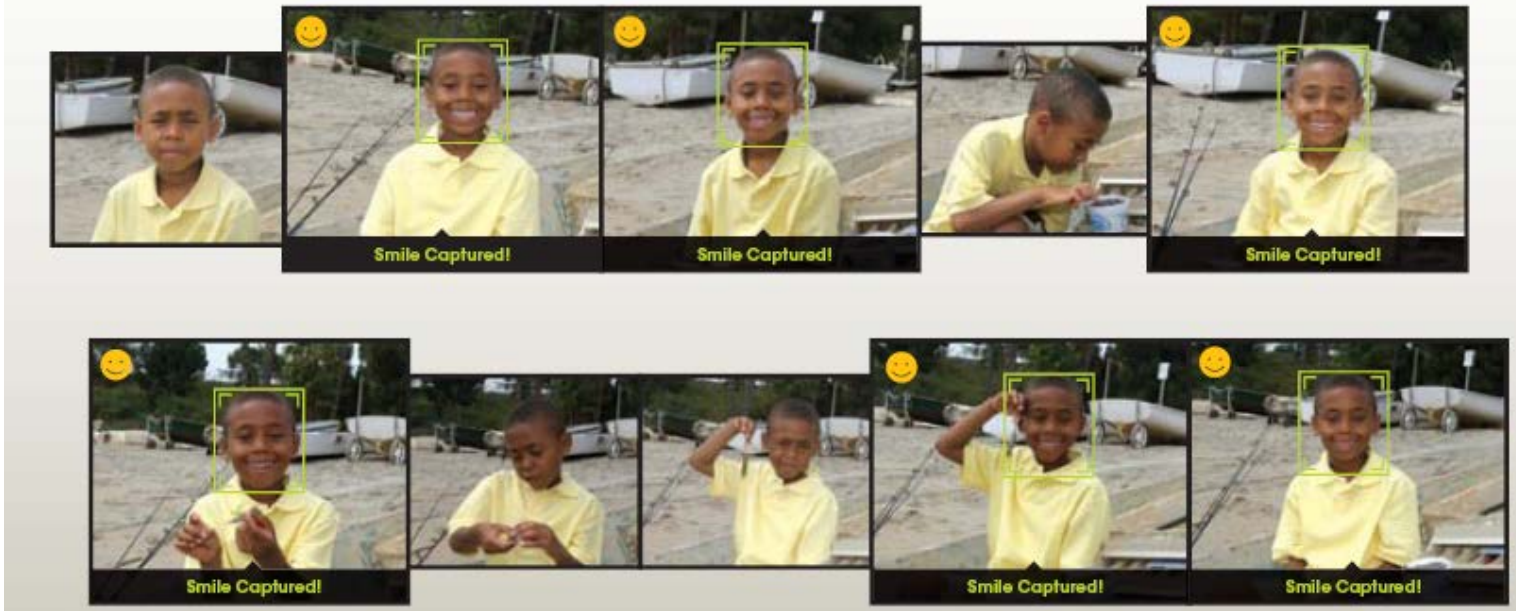
Face Detection

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

Smile detection

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



[Sony Cyber-shot® T70 Digital Still Camera]

3D from thousands of images

Building Rome in a Day: Agarwal et al. 2009

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/

Object recognition (in mobile phones)

[Point & Find](), [Nokia]()

[Google Goggles]()

Sports

Sportvision first down line
Nice explanation on www.howstuffworks.com

http://www.sportvision.com/video.html

Business Analytics



Reveal (Auckland, New Zealand). Systems for counting and tracking pedestrians using overhead cameras.

Smart Cars

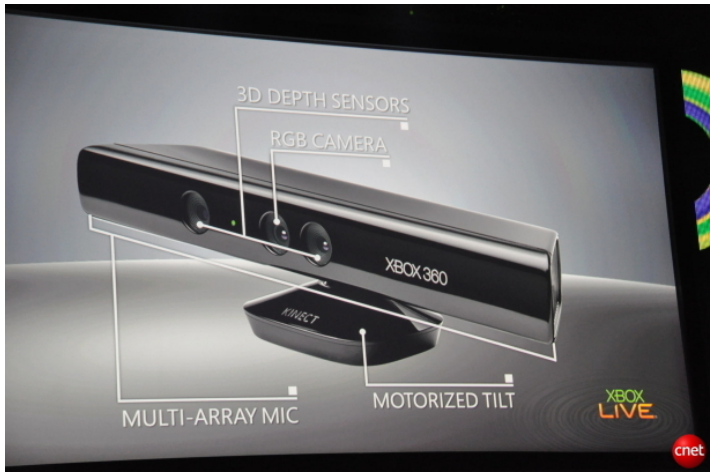
- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models
  - 2010: ~70% of car companies use smart cams

Google Self-Driving Cars

http://www.nytimes.com/2010/10/10/science/10google.html?ref=artificialintelligence

Interactive Games: Kinect

- Object Recognition: http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o
- Mario: http://www.youtube.com/watch?v=8CTJL5lUjHg
- 3D: http://www.youtube.com/watch?v=7QrnwoO1-8A
- Robot: http://www.youtube.com/watch?v=w8BmgtMKFbY



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

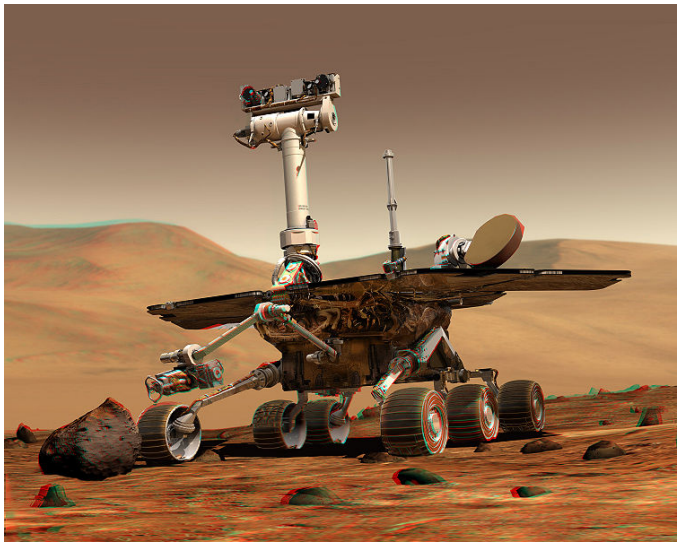
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Industrial Robots

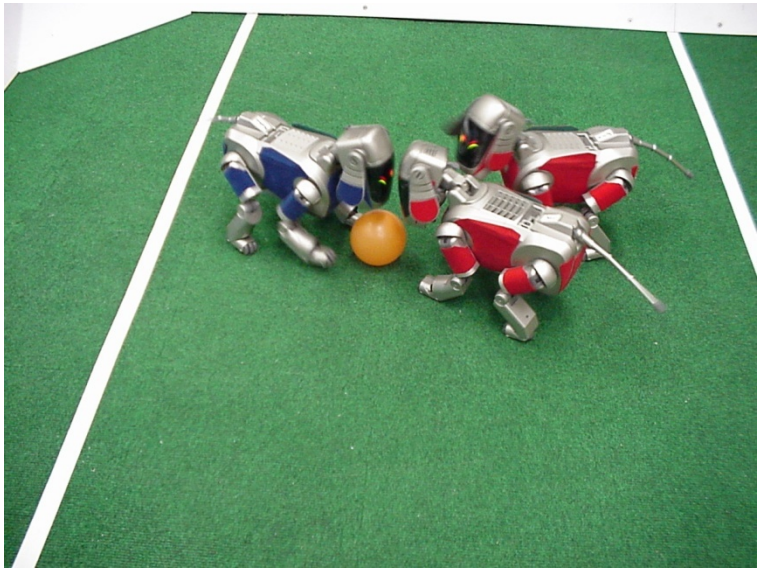
(Often called Machine Vision)

Vision-guided robots position nut runners on wheels

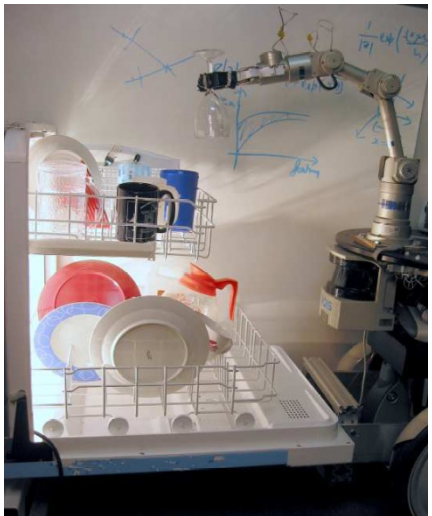
Mobile Robotics



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

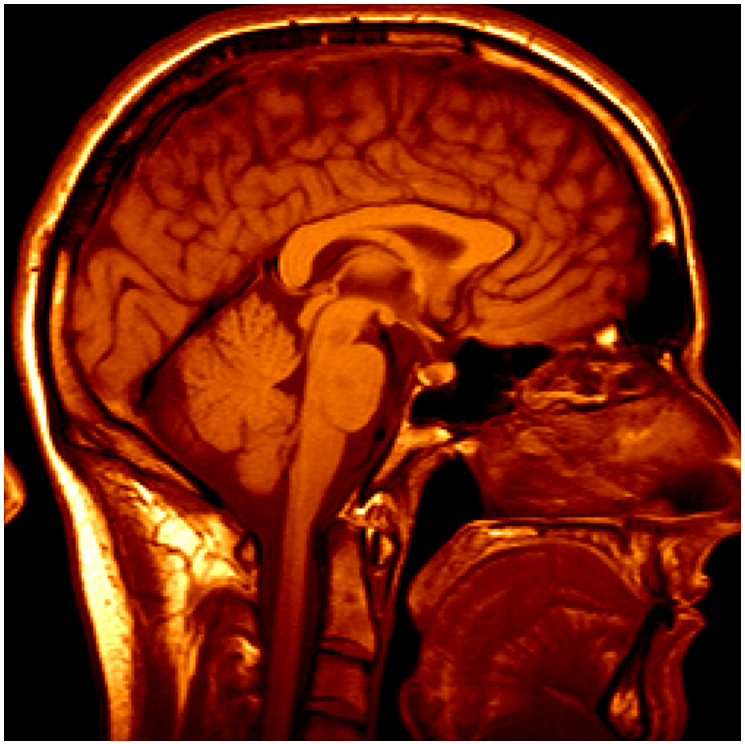


http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

Medical imaging



3D imaging
MRI, CT

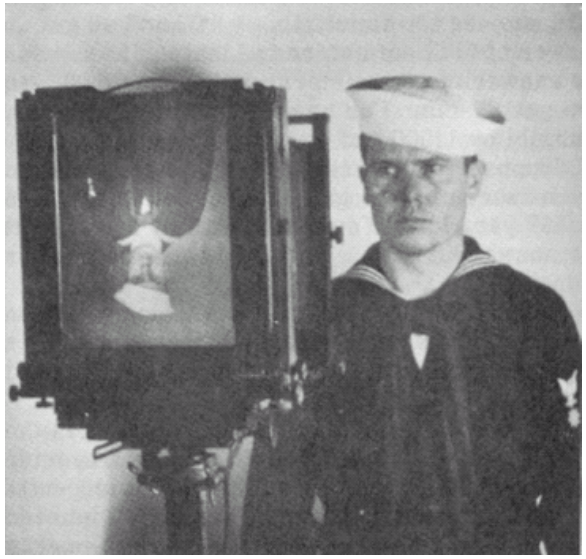
Image guided surgery
Grimson et al., MIT

Sounds like fun…

- What math background do I need for this course?
- What CS background do I need for this course?
- What are we going to learn?

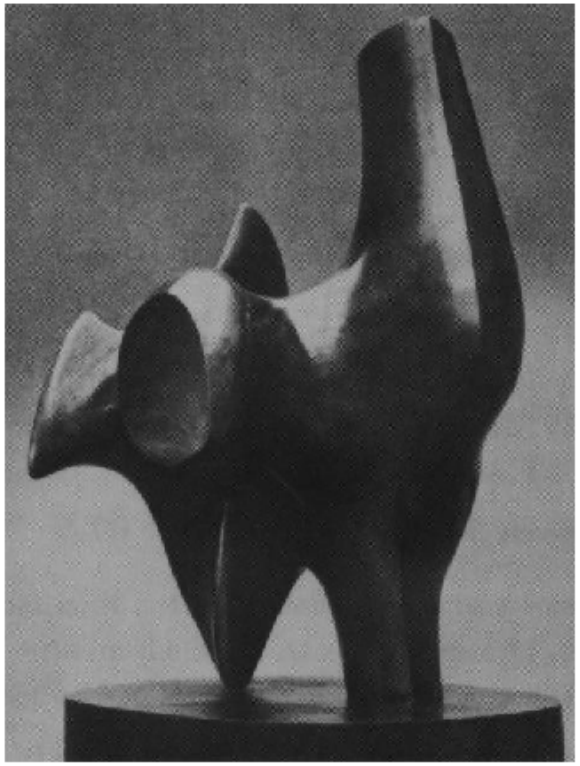
Image Formation

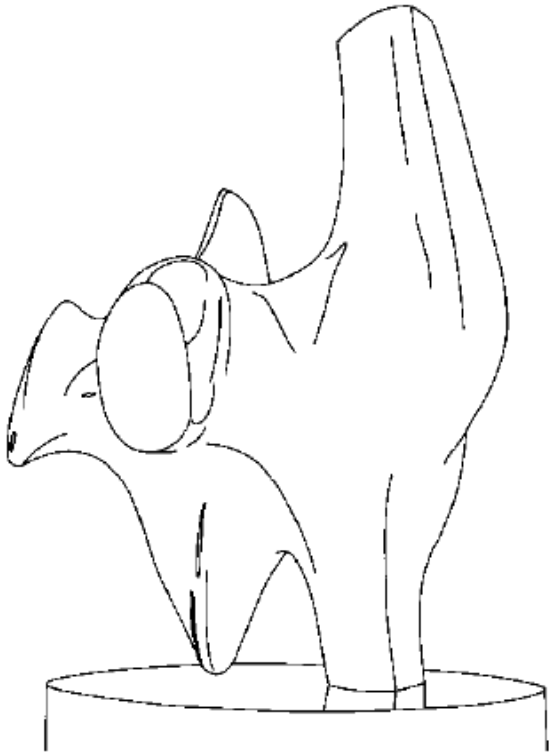
- Brief history & progression different cameras and optics.
- This is not a hardware course, but we need to know how images are generated



Geometry, Linear Algebra

Finding “Interest Points”

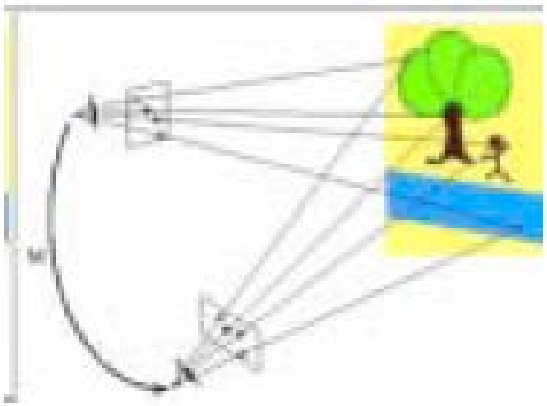




Derivatives and Convolutions

Stereo

- Humans use stereo… so, it must be a good thing
- What can you do with 2 cameras that you cannot do with one?

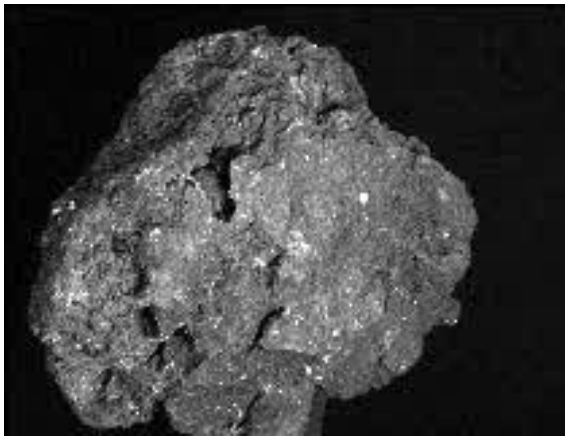


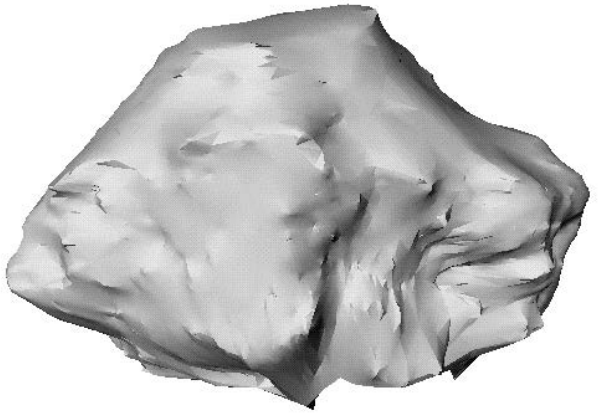
- Linear Algebra,
- Geometry,
- Robust Statistics

Motion & Video

- Interpreting motion in video
- Measuring motion
- Relate motion to 3D shape

Linear Algebra,
Geometry,
Statistics





Detection and Classification

Machine Learning + Computer Vision

What kind of scene?

Where are the cars?

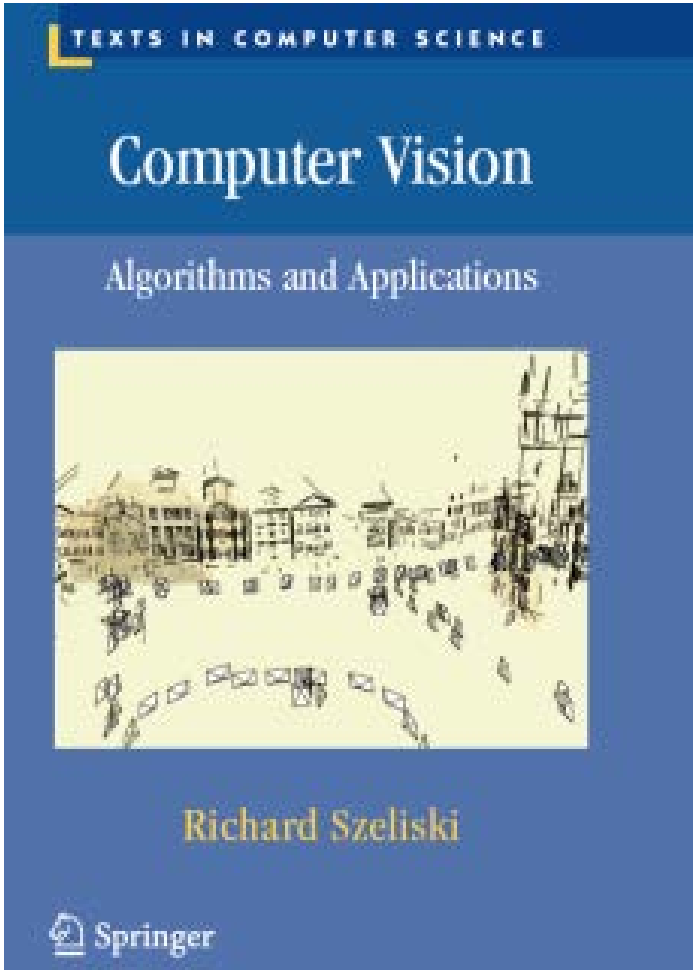
Which part of the image corresponds to “sky”?

…

Statistics

Course Textbook

- Should I buy it?
  - If you hate reading PDFs that you can get online for free, sure.
- Computer Vision: Algorithms and Applications
  by Richard Szeliski
  - PDF available online
  - You can also buy hardcopy
- $60 on Amazon.com



Slides, Lectures, Attendance

- Attendance is required
  - Most of the material is in the book
  - You are responsible for any announcements made in class.
  - Productive participation in class is a criteria for resolving borderline final grades
- Slides are posted on Moodle *after** class

Homework & Exams

- The 3-5 homework assignments will generally take the form of programming assignments where you implement methods we cover
- The 1-2 exams will cover the theoretical / mathematical aspects of the methods
- There may be a final project
  - More details later

Programming Language

- Matlab is the programming language for this course*
  - Pros: Rapid prototyping, widely used by the CV community, lots of available code, familiar syntax
  - Cons: Expensive

- Installed on UNCC computer lab machines
- Open-source alternative: Octave

**One exception is for the final project if another language is more suitable (e.g., mobile, real-time)*

How can I learn Matlab for this course?

- It's not too hard. Syntax should be familiar to Java / C / C++ user. (You can make direct Java calls, in fact.)
- Lots of tutorials on the web
  - Some resources on Moodle page

- *There will be one Recitation on Matlab primer*

Additional Course Information

- For hw0, you must read the syllabus on UNCC Policy on Academic Integrity
- Cheating, plagiarism are taken very seriously and dealt with harshly.
- Be familiar with the course policies as stated on the syllabus available online.

How should I go about getting help?

- Depends on the problem
- Ex: “How do I open an image using Matlab?”
  - Google
- Ex: “I don’t understand why we are multiplying this matrix by this vector.”
  - Ask TA, or come to my office hours
- Ex: “I would like to patent / publish this great idea!”
  - Ask the TA for my cell phone number and contact me immediately!

Next Steps

- 2 Assignments posted
  - Hw0: read the syllabus and academic integrity policy
  - Hw1: Image Processing, Computer Vision, “Grand Challenge”: Eye localization (due in two weeks).