C280, Computer Vision

Prof. Trevor Darrell

trevor@eecs.berkeley.edu

Today

- Administrivia
- “What is vision?”
- Elementary Image formation
- Schedule
- Introductions

Prerequisites

- This course is appropriate as a first course for graduate students with a EECS background, which should have prepared the students with these essential prerequisites:
  - Data structures
  - A good working knowledge of MATLAB programming (or willingness and time to pick it up quickly!)
  - Linear algebra
  - Vector calculus
- The course does not assume prior imaging experience, computer vision, image processing, or graphics

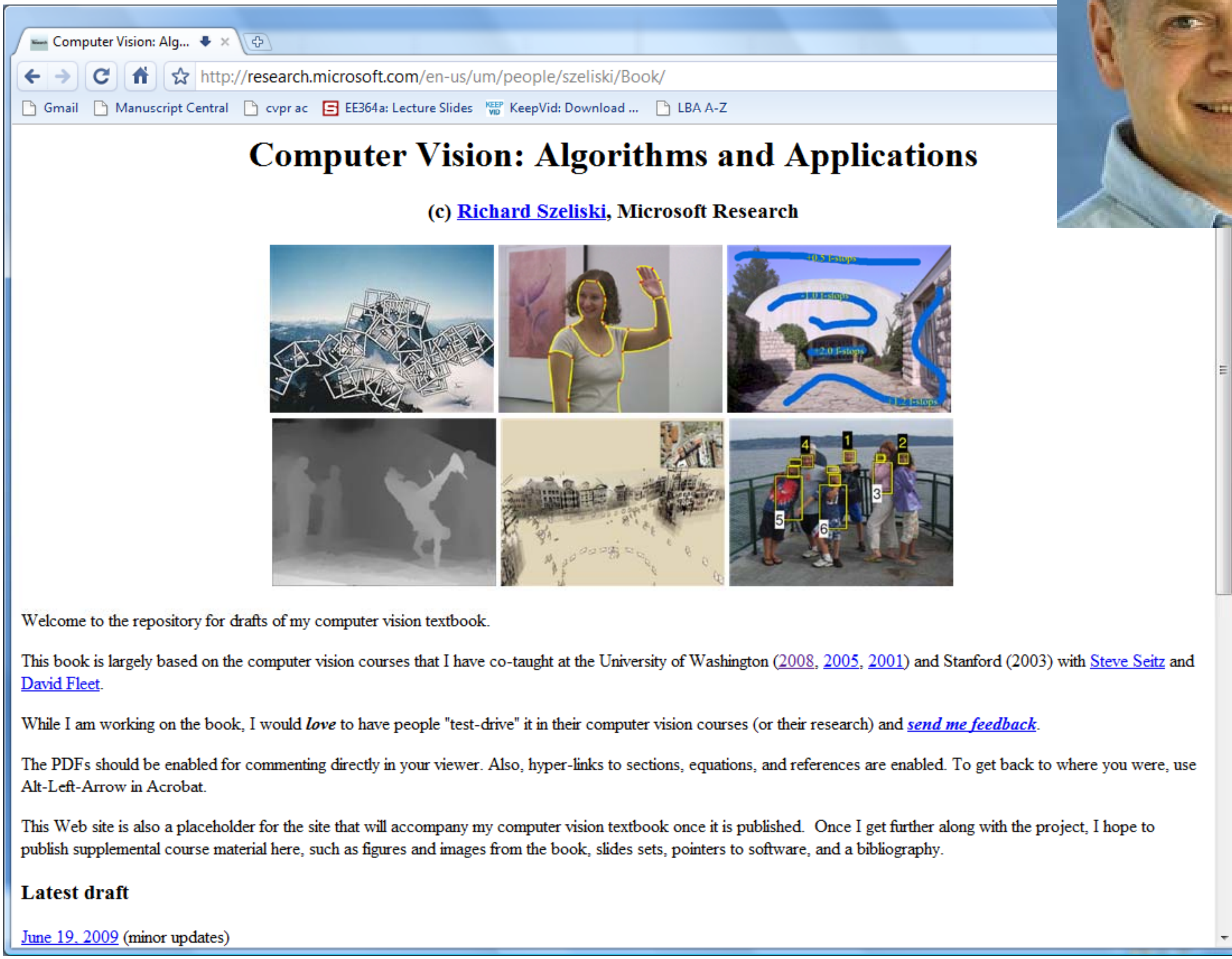
Grading

- There will be three equal components to the course grade
  - Five problem sets
  - A take-home exam
  - Final project (including evaluation of proposal document, in-class presentation, and final report)
- In addition, strong class participation can offset negative performance in any one of the above components.

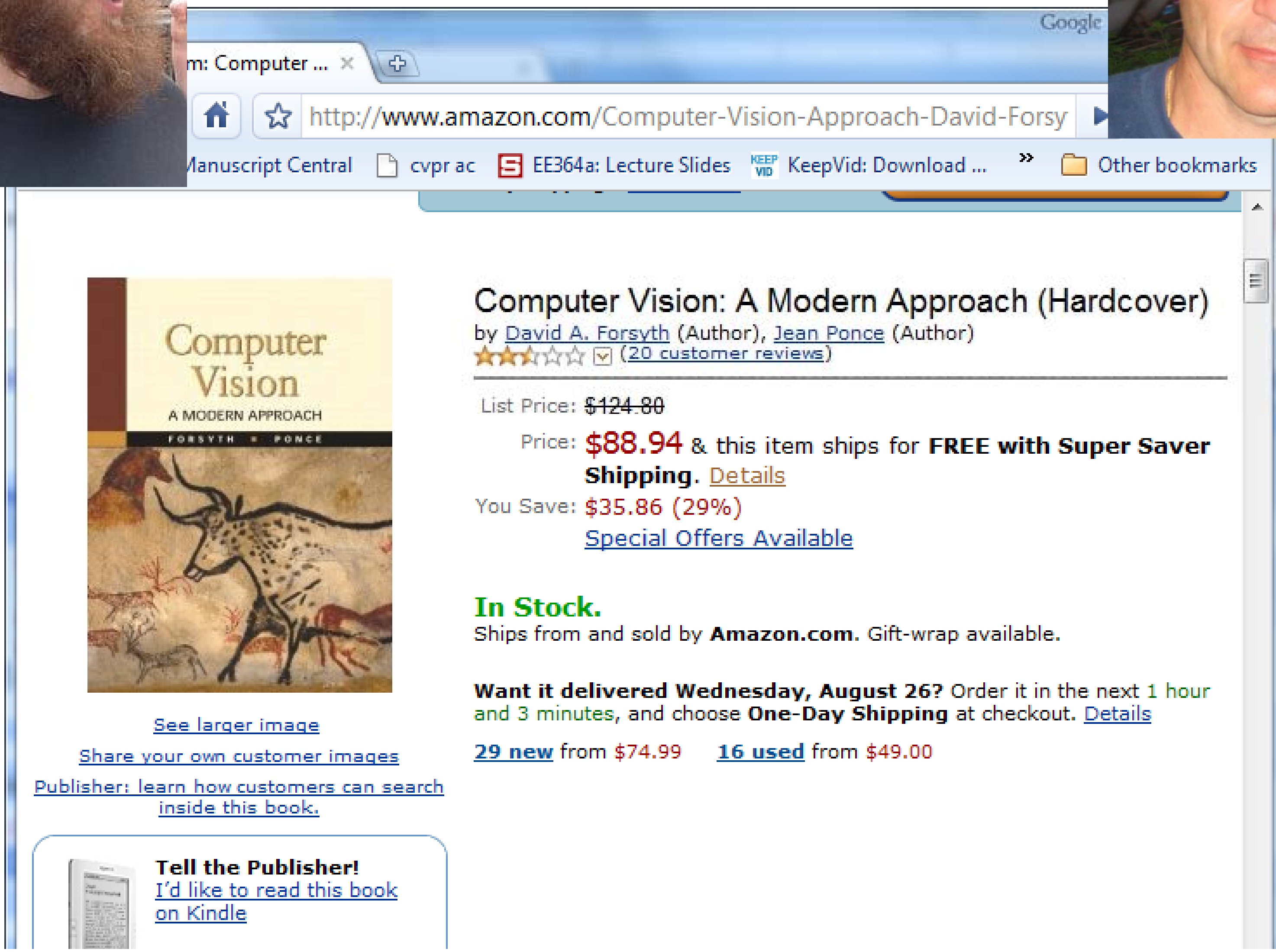
Text

- The primary course text will be Rick Szeliski's draft Computer Vision: Algorithms and Applications; we will use an online copy of the June 7th draft.
- The secondary text is Forsyth and Ponce, Computer Vision: A Modern Approach.

Primary Text



Secondary Text



Matlab

- Problem sets and projects will involve Matlab programming (you are free to use alternative packages). Matlab runs on all the Instructional Windows and UNIX systems. Instructions and toolkits are described in http://inst.eecs.berkeley.edu/cgi-bin/pub.cgi?file=matlab.help.
- CS280 students can use their existing EECS Windows accounts in EECS instructional labs, and they can request new accounts (for non-majors) or additional access to Instructional resources by following the instructions about 'named' accounts in http://inst.eecs.berkeley.edu/connecting.html#accounts. They can logon remotely and run it on some of our servers: http://inst.eecs.berkeley.edu/connecting.html#labs

Problem sets

- Pset0 – Basic Image Manipulation in Matlab
- Pset1 – Filtering and Features
- Pset2 – Geometry and Calibration
- Pset3 – Recognition
- Pset4 – Stereo and Motion
- *Can discuss, but must submit individual work*

Take-home

- Limited time: 3 days
- Covers everything through hand out date
- Little programming
- **No discussion or collaboration allowed**

Final project

- Significant novel implementation of technique related to course content
- Teams of 2 encouraged (document role!)
- Or journal length review article (no teams)
- Three components:
  - proposal document (no more than 5 pages)
  - in class results presentation (10 minutes)
  - final write-up (no more than 15 pages)

Class Participation

- Class participation includes
  - showing up
  - being able to articulate key points from last lecture
  - having read assigned sections and being able to "fill in the blank" during the lecture
- I won't cold call, but will solicit volunteers
- Strong in-class participation can offset poor performance in one of the other grade components.

Course goals....(broadly speaking)

- principles of image formation
- convolution and image pyramids
- local feature analysis
- multi-view geometry
- image warping and stitching
- structure from motion
- visual recognition
- image-based rendering

What is computer vision?

Done?

What is computer vision?

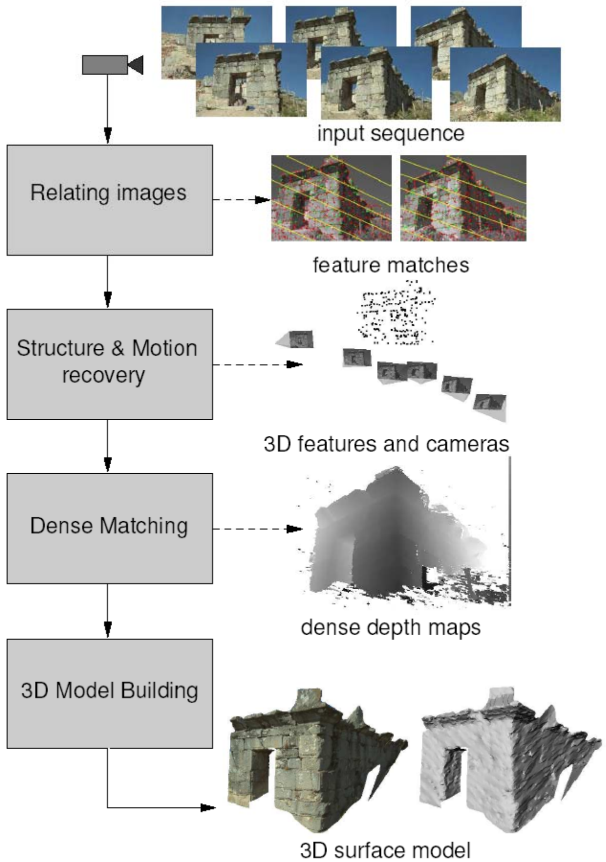
- Automatic understanding of images and video
  - Computing properties of the 3D world from visual data *(measurement)*
  - Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. *(perception and interpretation)*

Vision for measurement

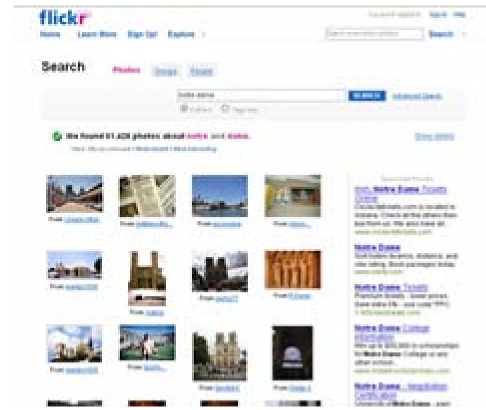
Real-time stereo

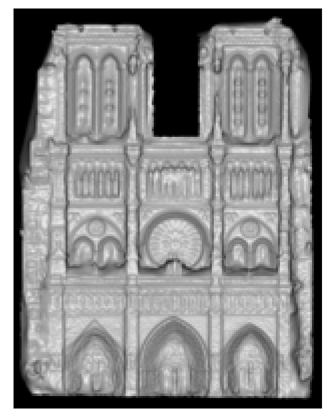
Pollefeys et al.

Structure from motion



Multi-view stereo for community photo collections

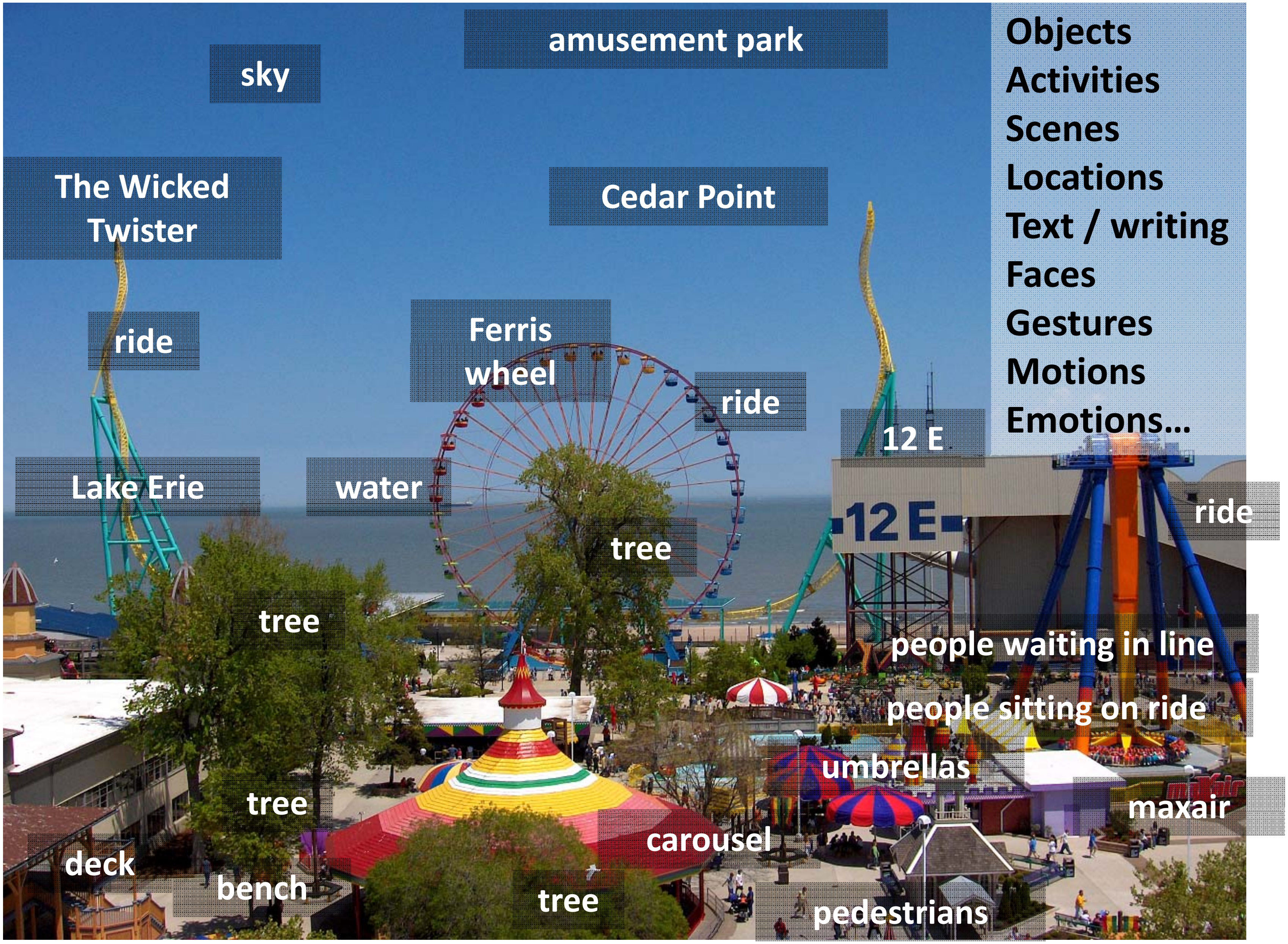




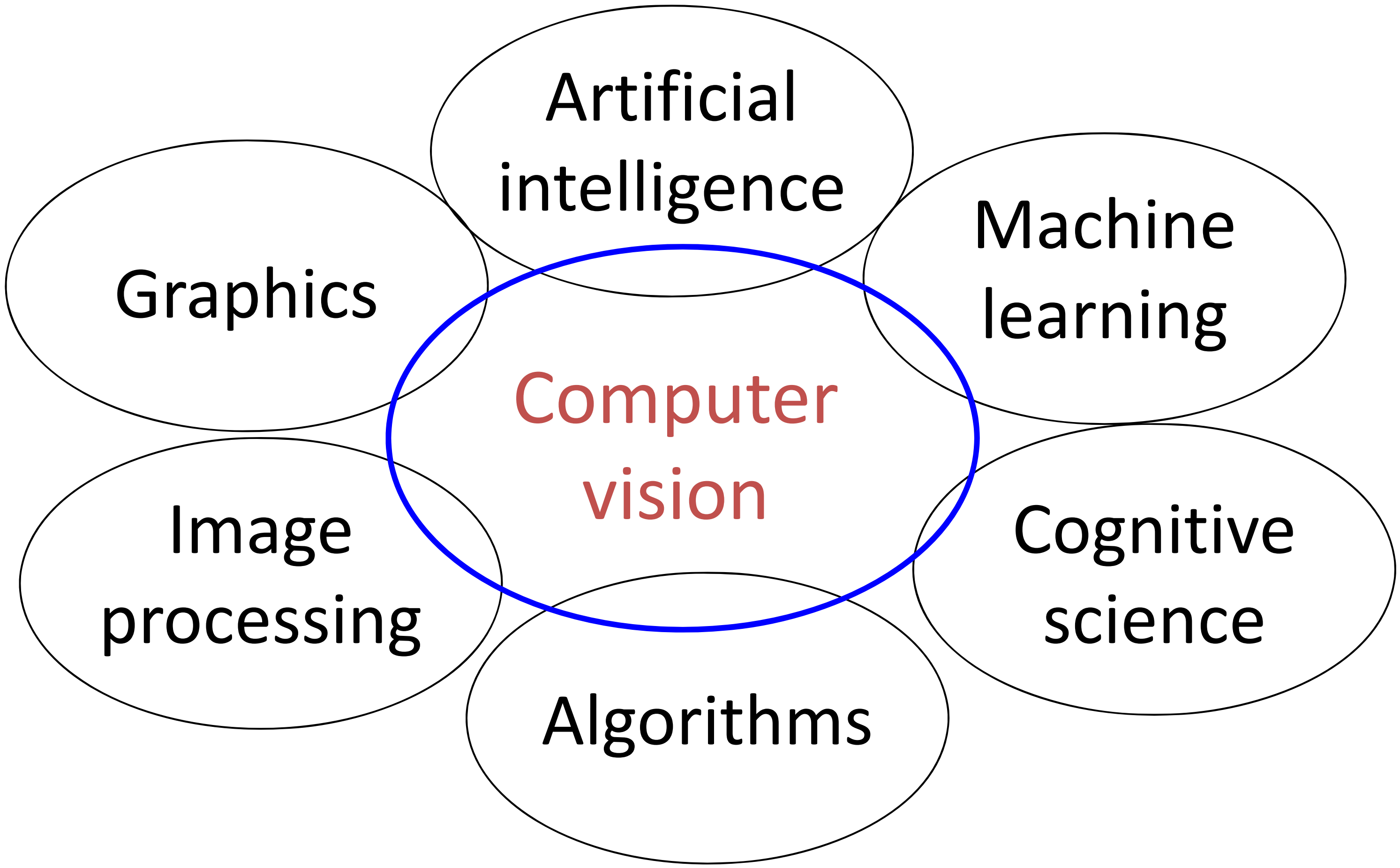
Goesele et al.

Slide credit: L. Lazebnik

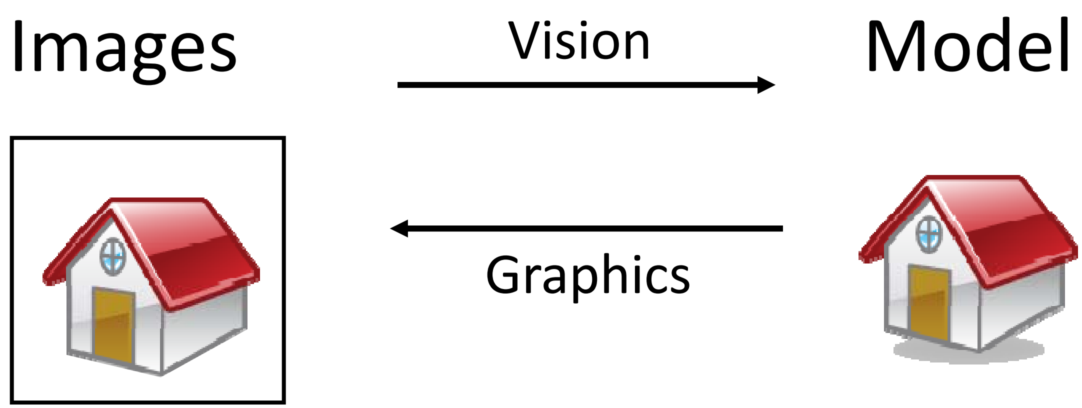
Vision for perception, interpretation



Related disciplines



Vision and graphics



Inverse problems: analysis and synthesis.

Why vision?

- As image sources multiply, so do applications
  - Relieve humans of boring, easy tasks
  - Enhance human abilities: human-computer interaction, visualization
  - Perception for robotics / autonomous agents
  - Organize and give access to visual content

Why vision?

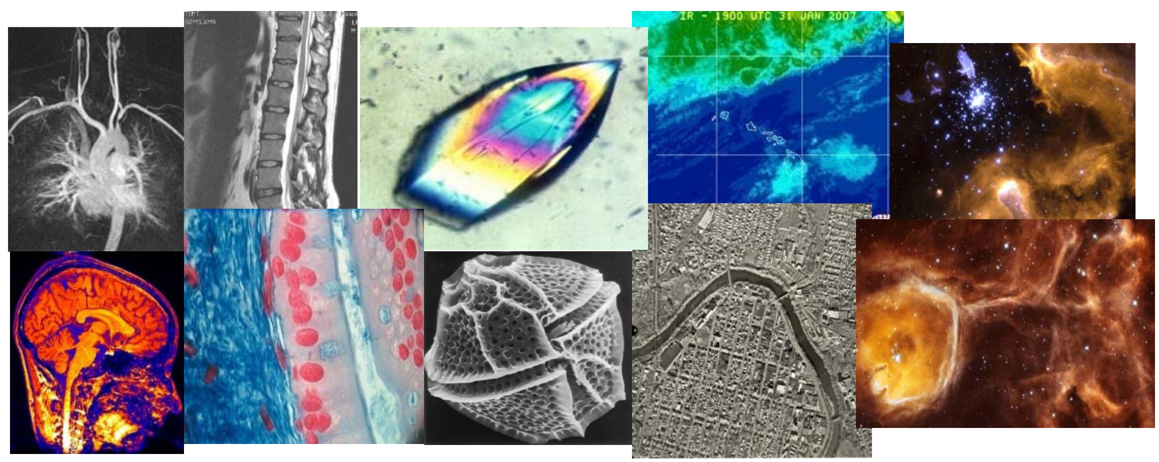
- Images and video are everywhere!

Personal photo albums

Movies, news, sports

Google Image Search Picasa flickr webshots picsearch YouTube Broadcast Yourself

Surveillance and security



Medical and scientific images

Slide credit; L. Lazebnik

Again, what is computer vision?

- Mathematics of geometry of image formation?
- Statistics of the natural world?
- Models for neuroscience?
- Engineering methods for matching images?
- Science Fiction?

Vision Demo?

Terminator 2

we're not quite there yet….

Every picture tells a story

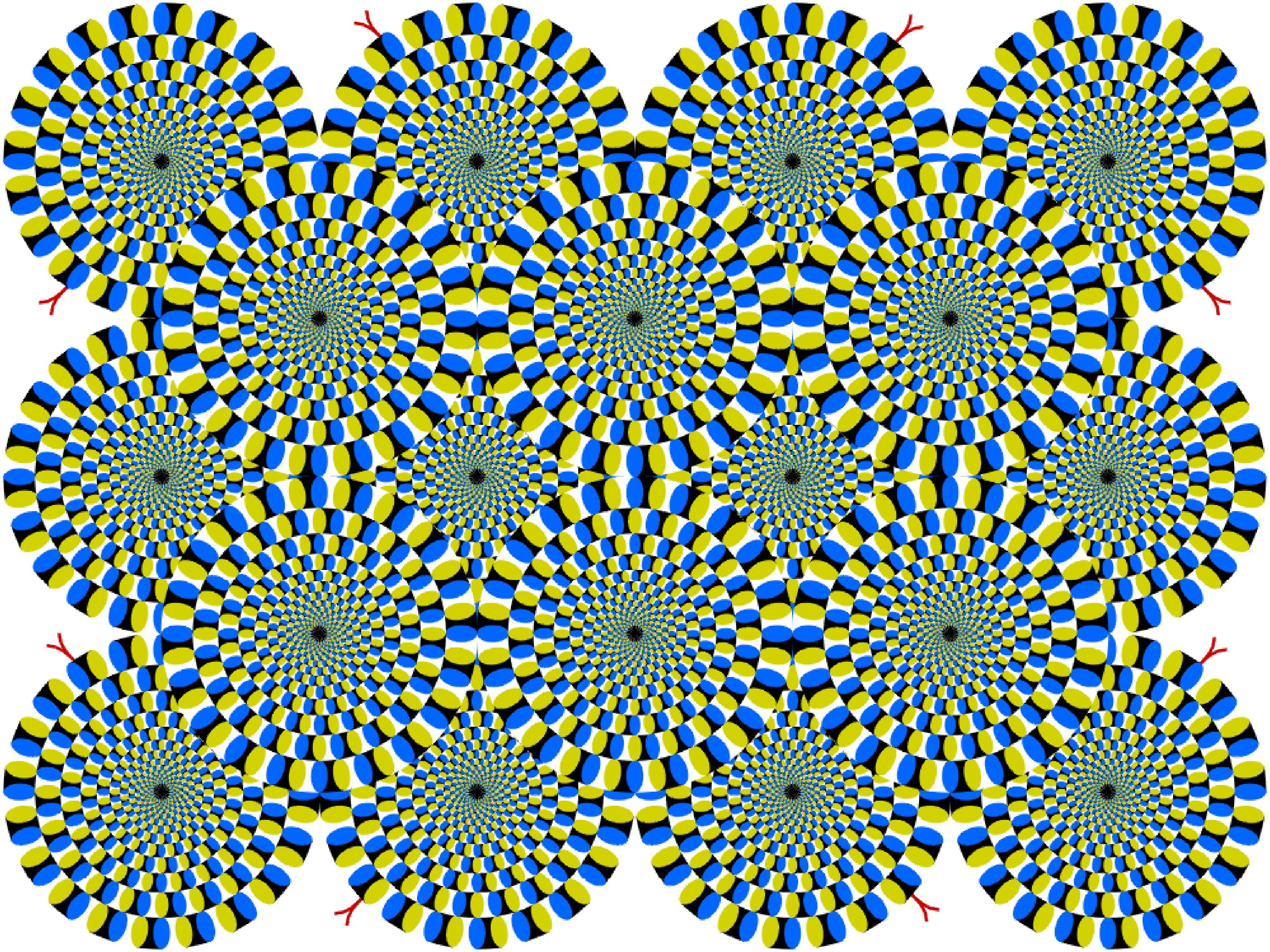
- Goal of computer vision is to write computer programs that can interpret images

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

Human perception has its shortcomings…

Sinha and Poggio, *Nature*, 1996



Current state of the art

- The next slides show some examples of what current vision systems can do

Earth viewers (3D modeling)

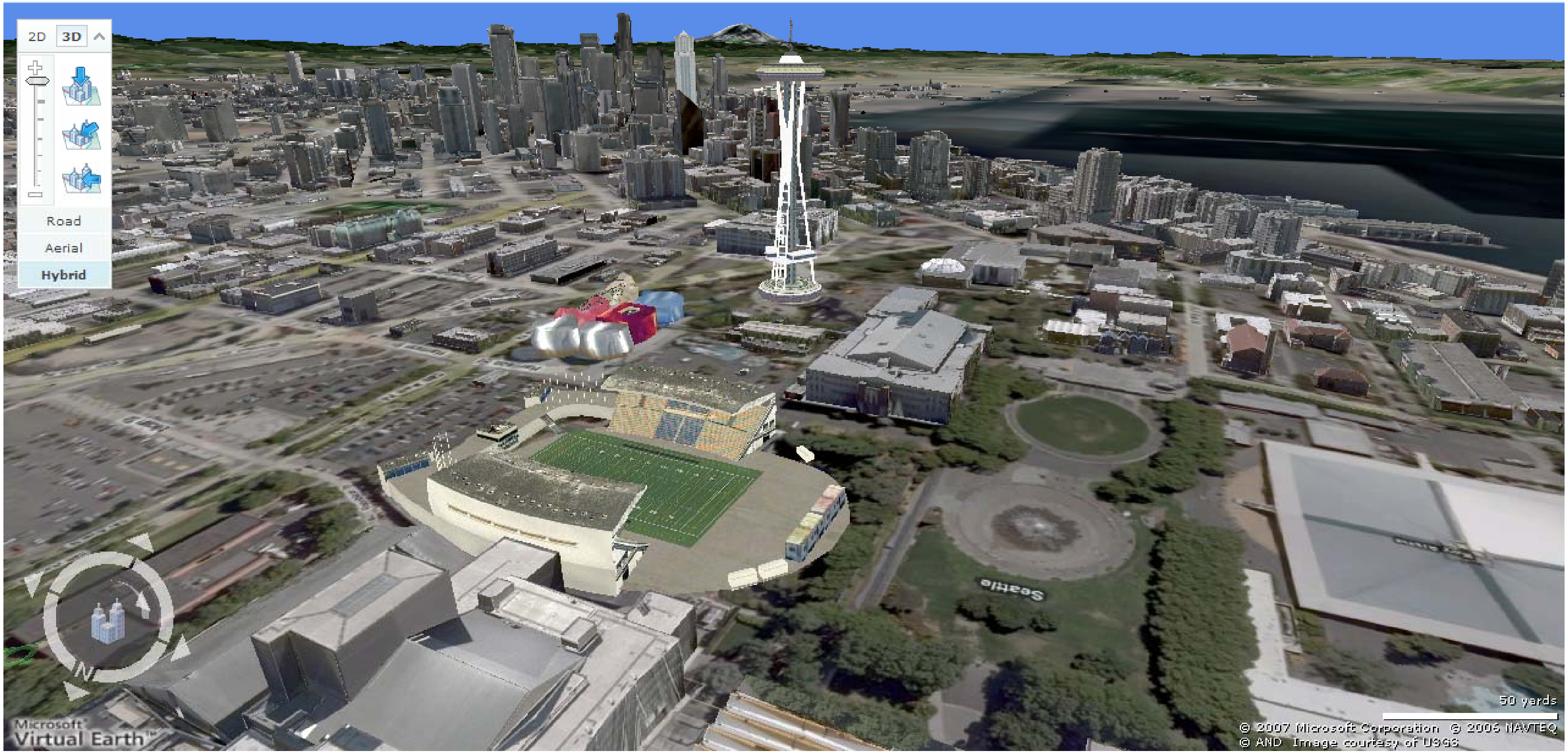
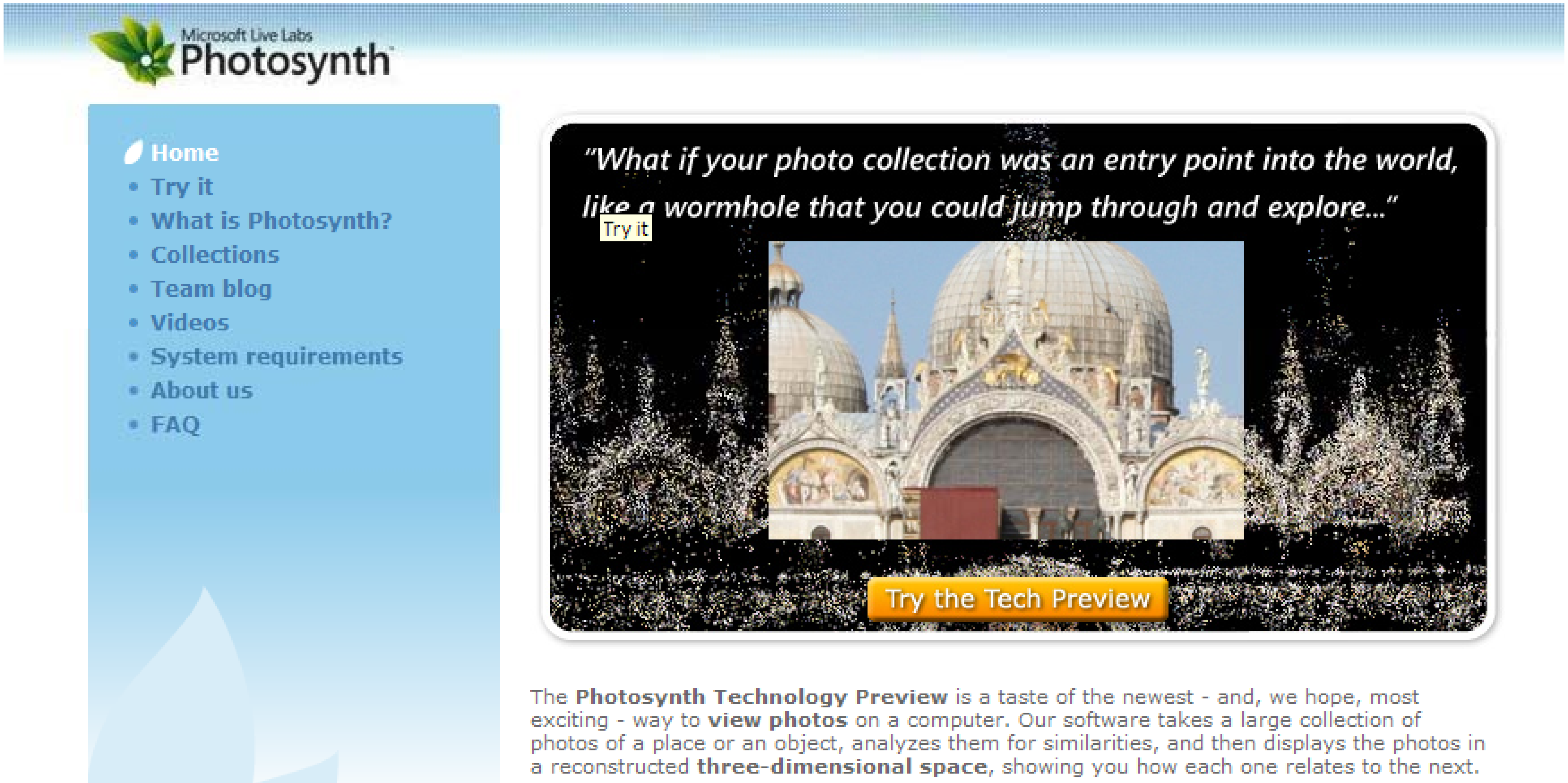


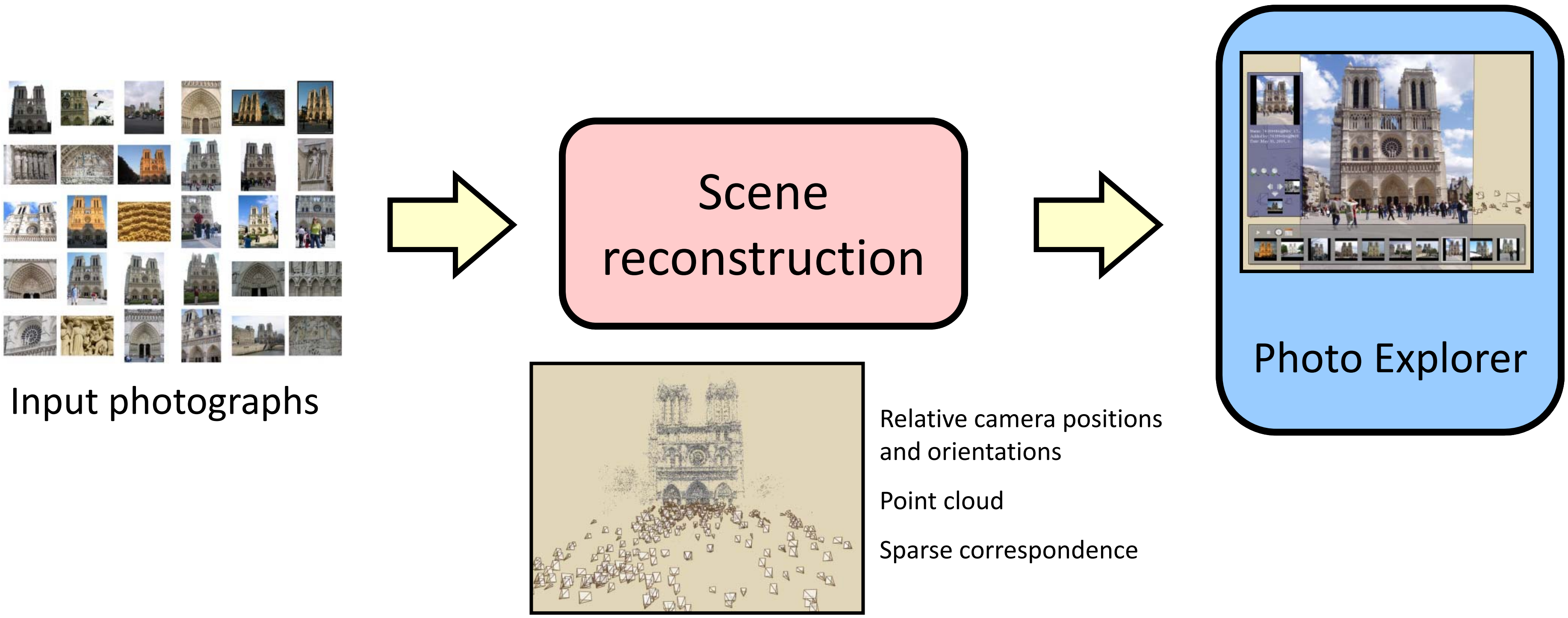
Image from Microsoft's [Virtual Earth](#)
(see also: [Google Earth](#))



http://labs.live.com/photosynth/

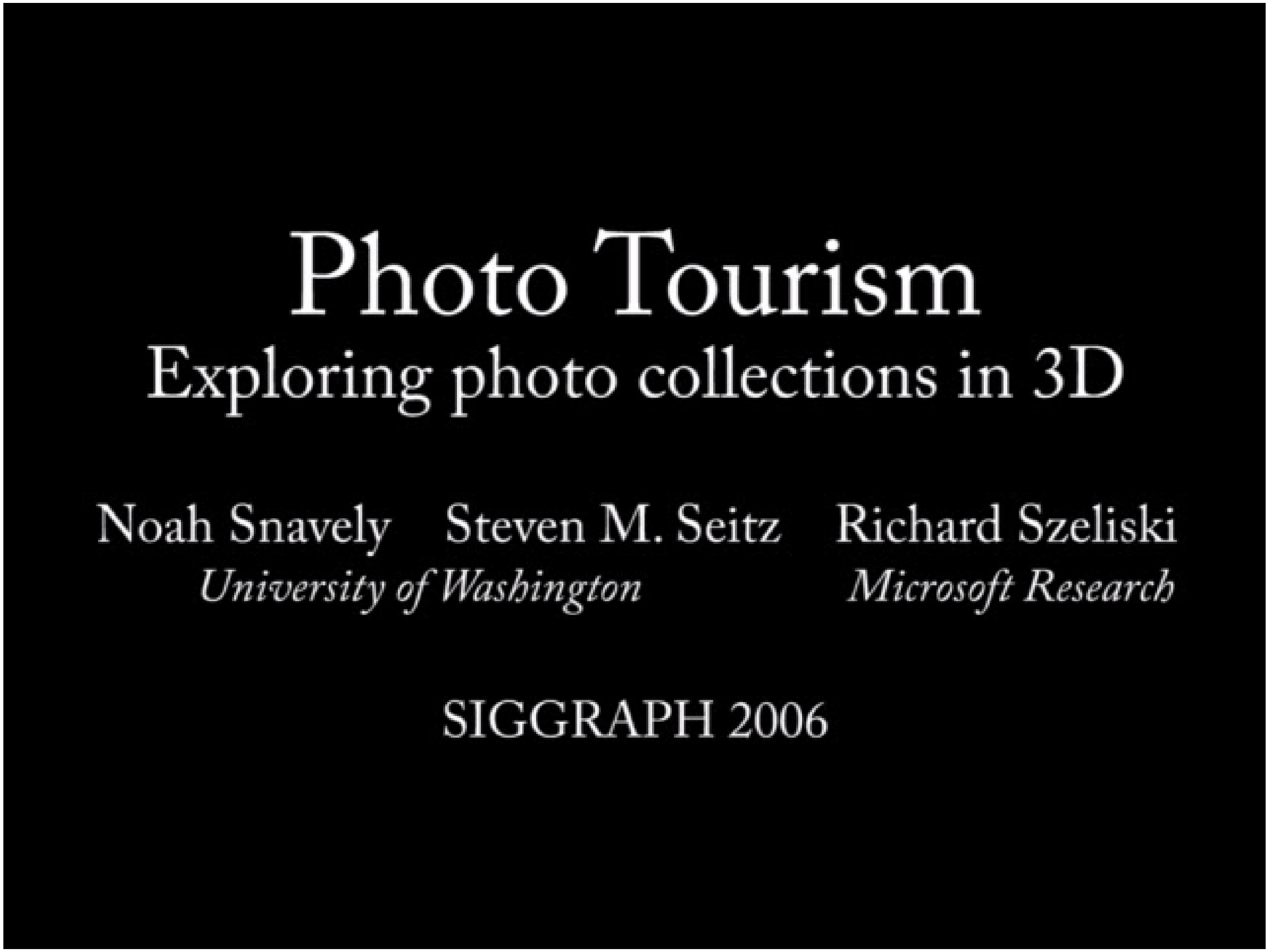
Based on Photo Tourism technology developed by Noah Snavely, Steve Seitz, and Rick Szeliski

Photo Tourism overview



System for interactive browsing and exploring large collections of photos of a scene. Computes viewpoint of each photo as well as a sparse 3d model of the scene.

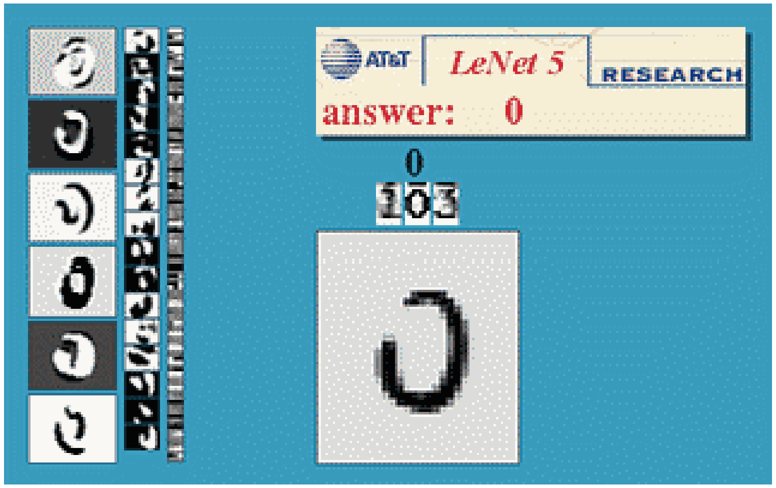
Photo Tourism overview



Optical character recognition (OCR)

Technology to convert scanned docs to text

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



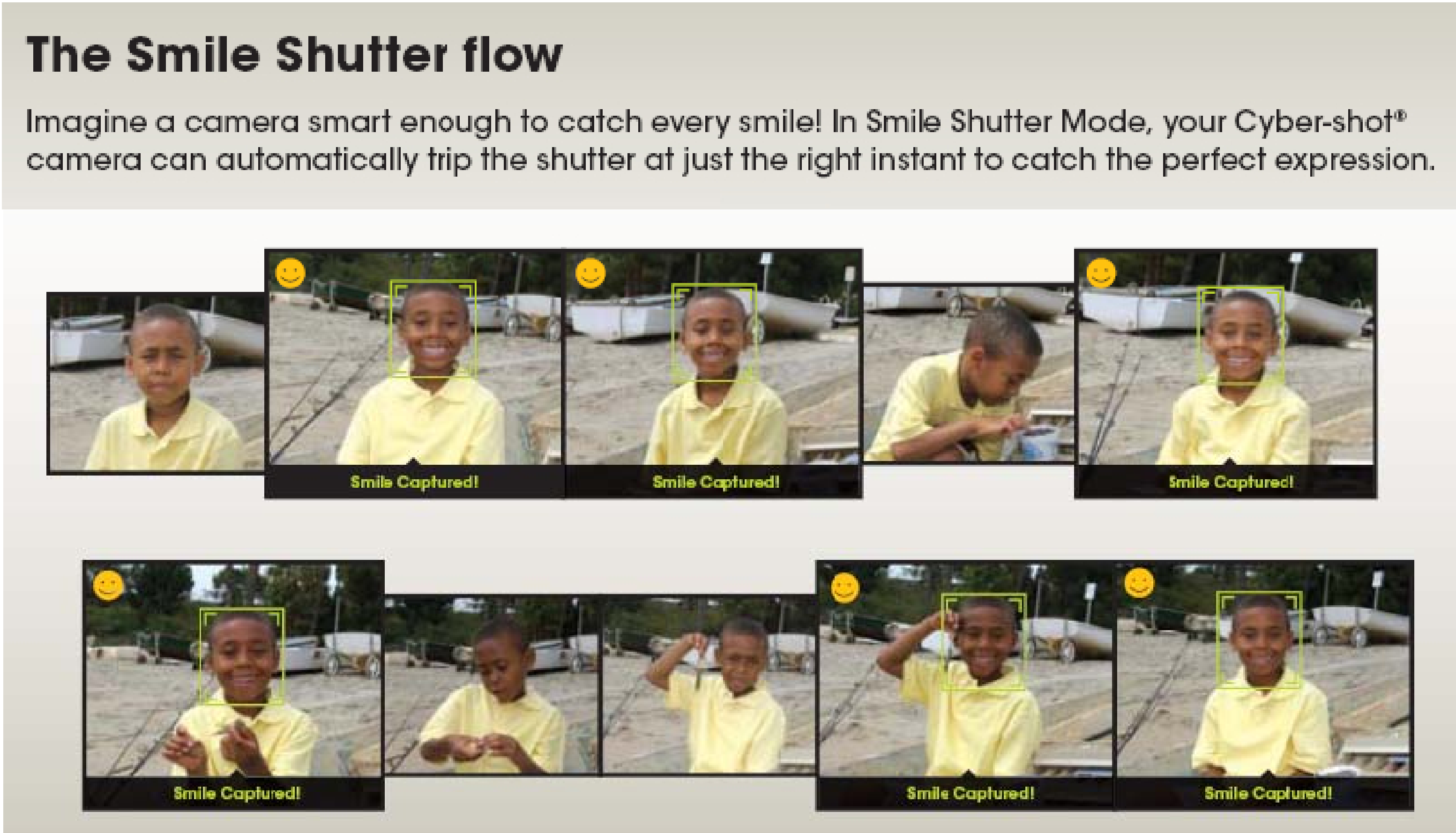
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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?



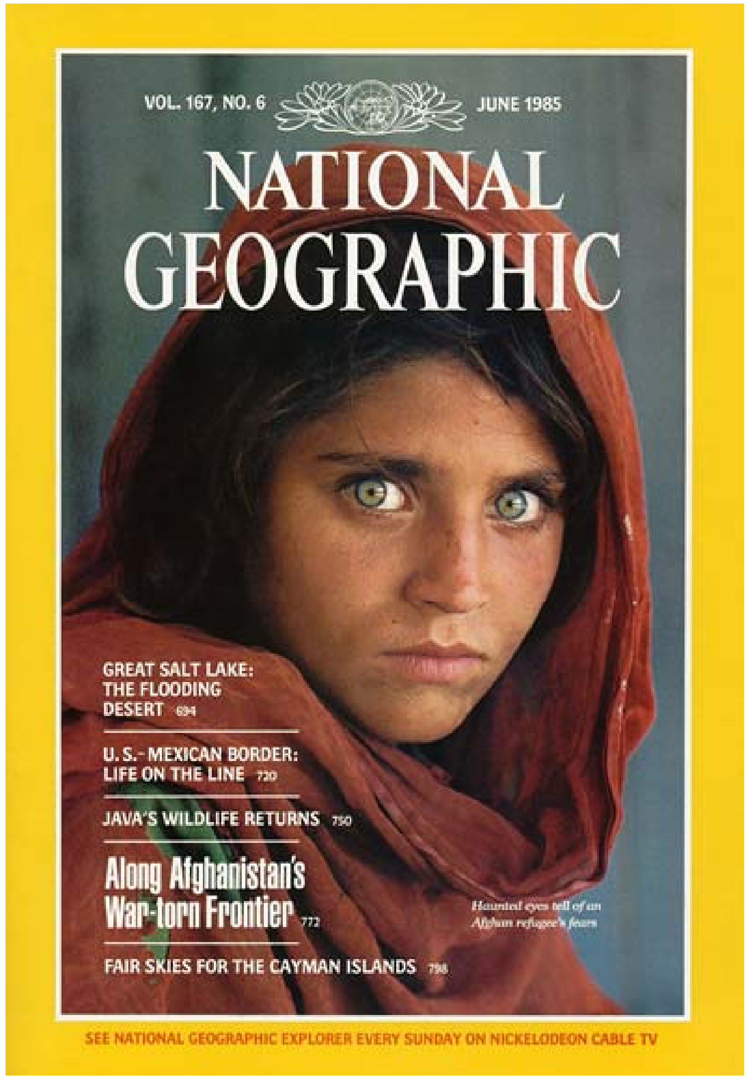
[Sony Cyber-shot® T70 Digital Still Camera](#)

Object recognition (in supermarkets)

LaneHawk by EvolutionRobotics

"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you are assured to get paid for it… "

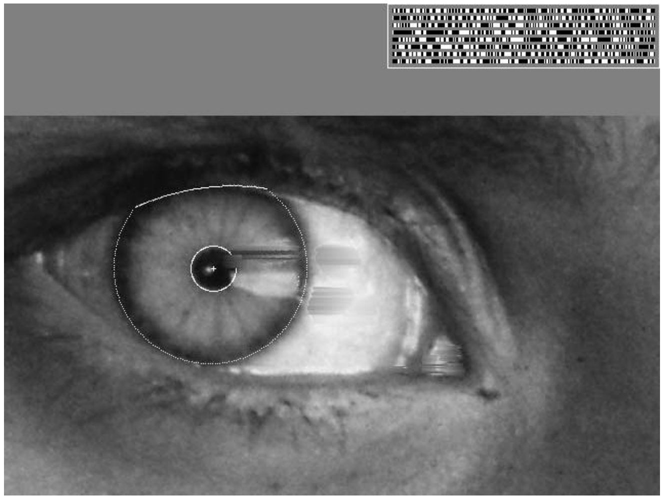
Face recognition

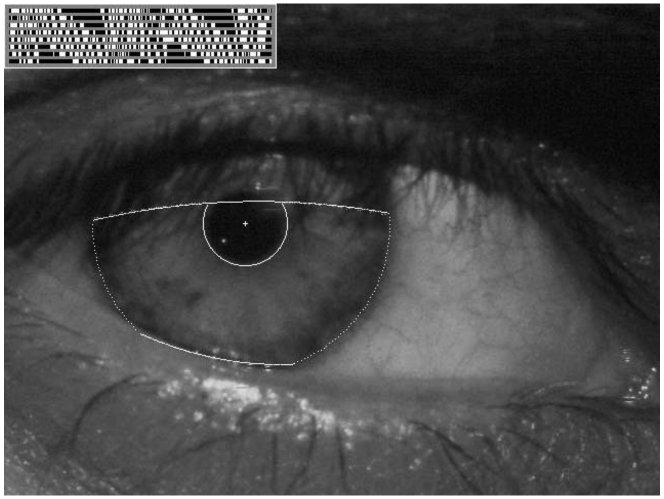


Who is she?

Vision-based biometrics

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





Login without a password…

Fingerprint scanners on
many new laptops,
other devices

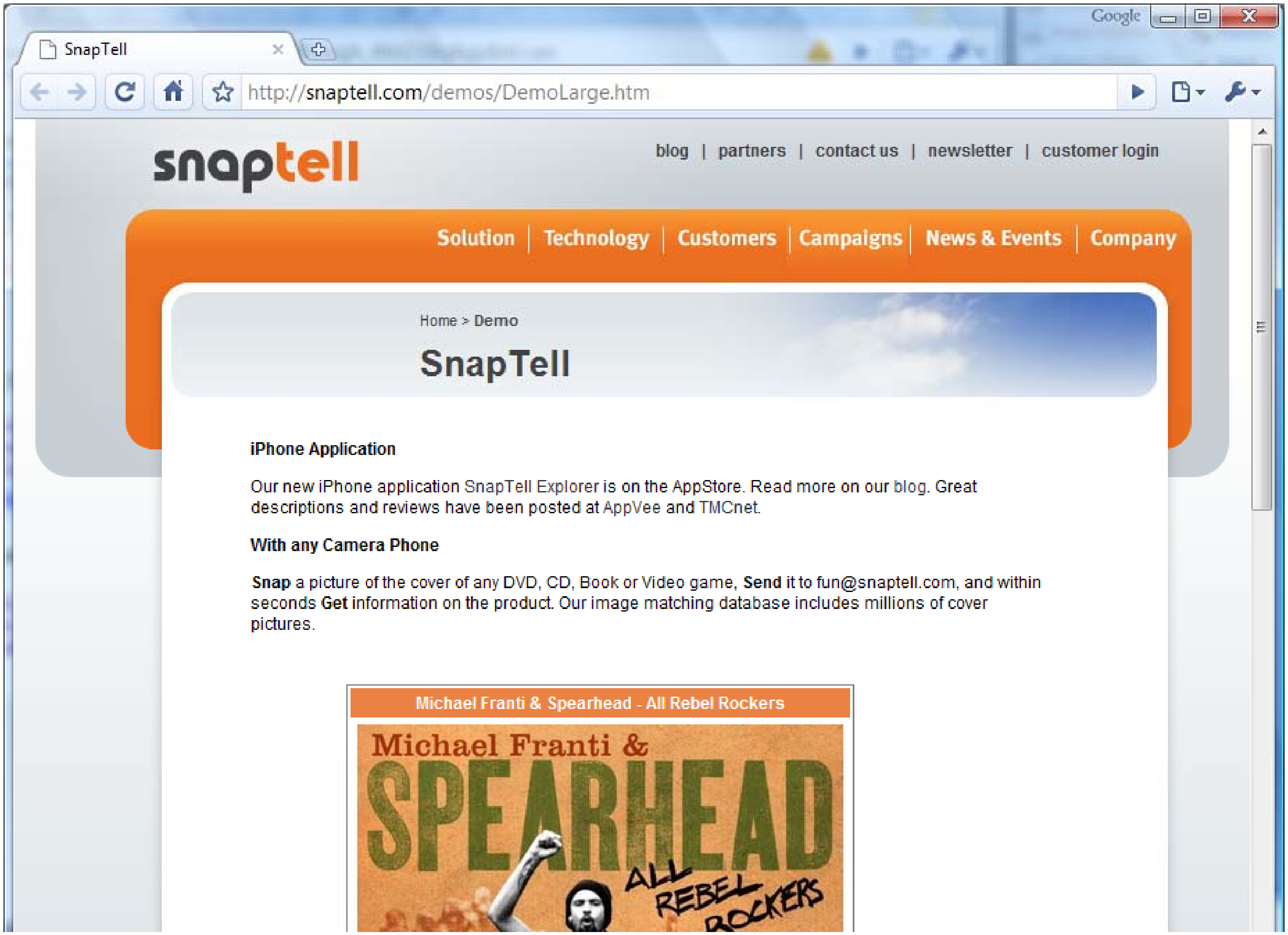
Face recognition systems now
beginning to appear more widely
http://www.sensiblevision.com/

Object recognition (in mobile phones)

- This is becoming real:
  - Lincoln Microsoft Research
  - Point & Find, Nokia
  - SnapTell.com (now amazon)

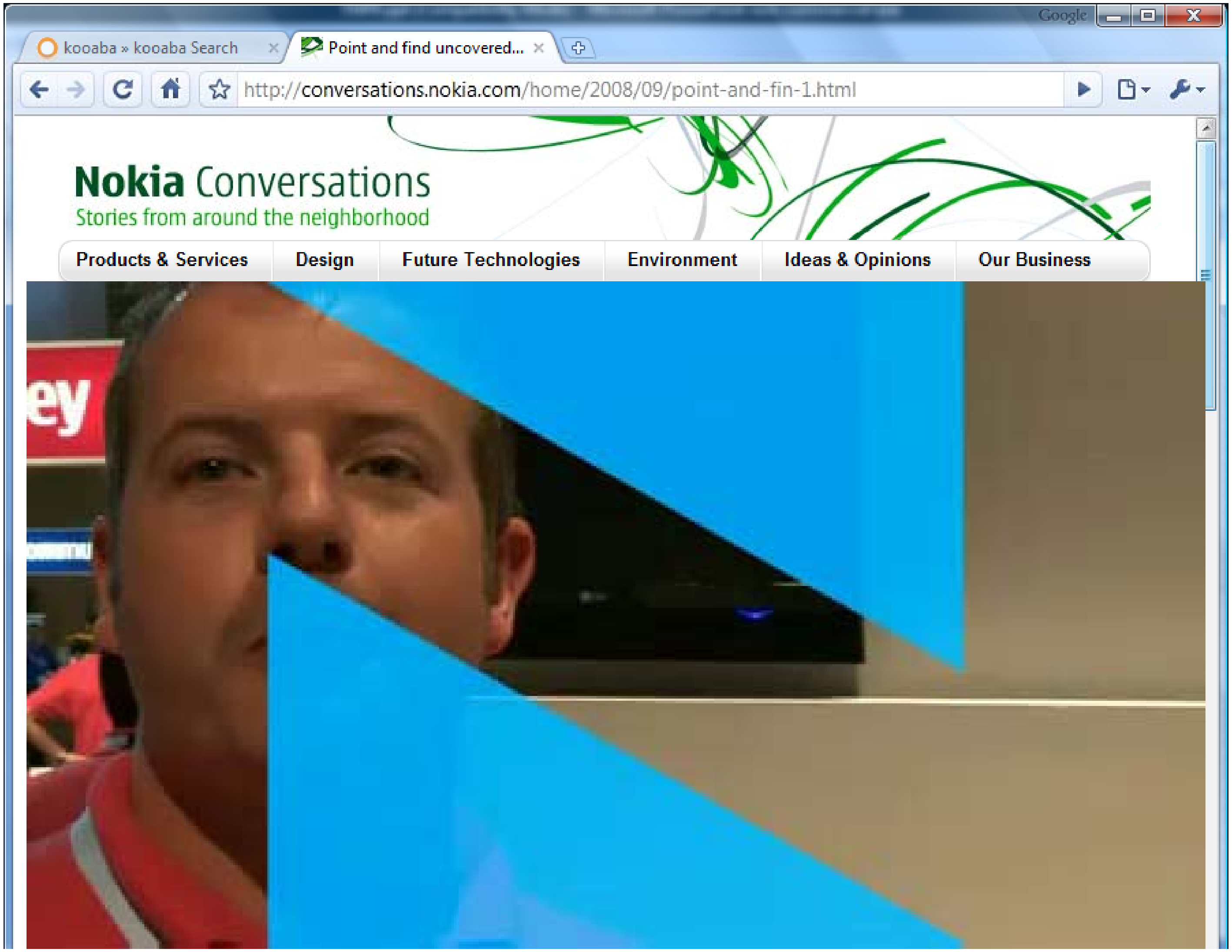
Snaptell

http://snaptell.com/demos/DemoLarge.htm



Nokia Point and Tell…

http://conversations.nokia.com/home/2008/09/point-and-fin-1.html



Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic
[Click here for interactive demo]

Sports

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

Smart cars

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

Slide content courtesy of Amnon Shashua

Smart cars

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

Slide content courtesy of Amnon Shashua

Vision-based interaction (and games)

Nintendo Wii has camera-based IR tracking built in. See Lee's work at CMU on clever tricks on using it to create a multi-touch display!

Digimask: put your face on a 3D avatar.

"Game turns moviegoers into Human Joysticks", CNET
Camera tracking a crowd, based on this work.

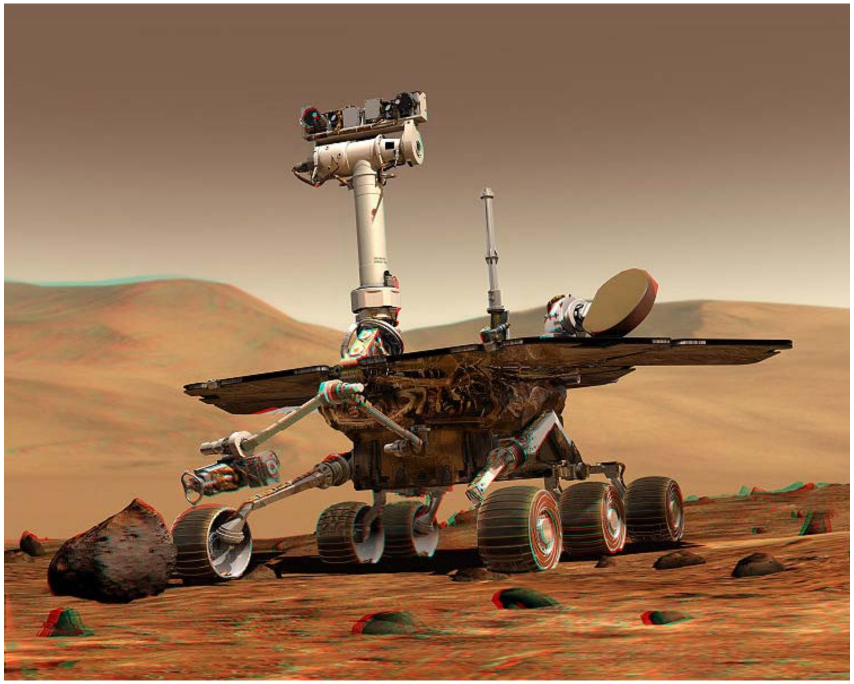
Vision in space

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

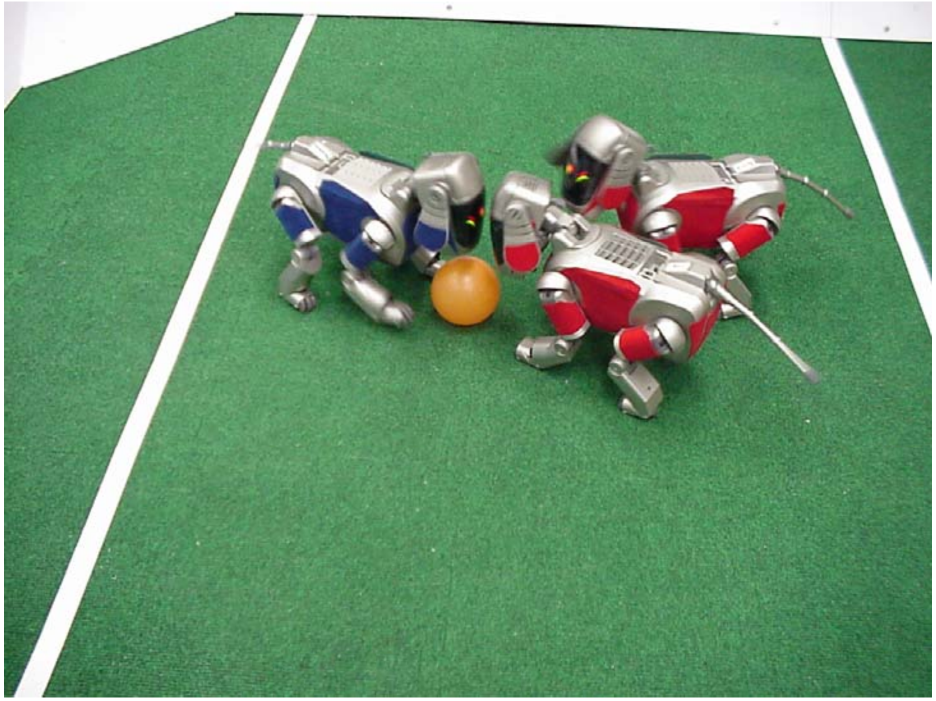
Vision systems (JPL) used for several tasks

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

Robotics

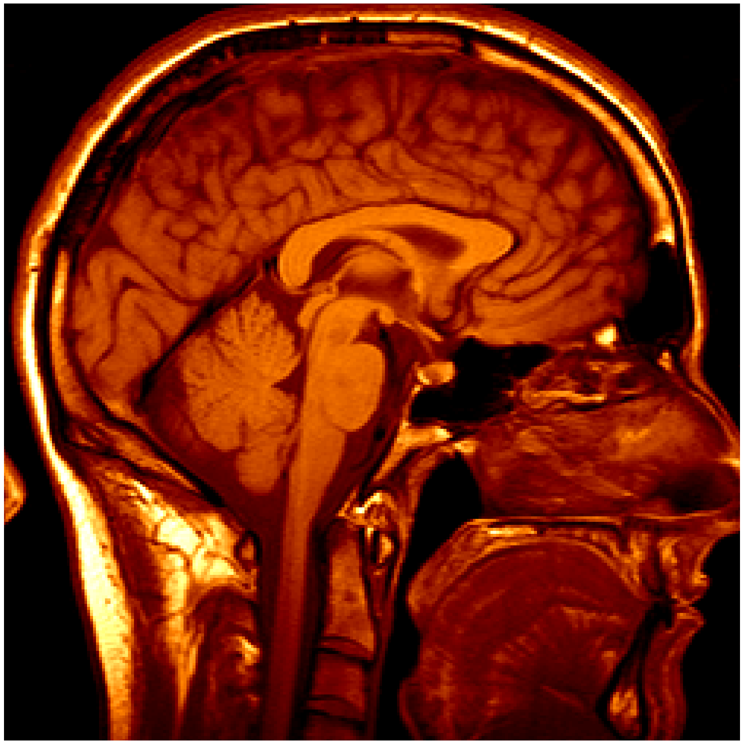


NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

Medical imaging



3D imaging
MRI, CT

Image guided surgery
[Grimson et al., MIT]

Current state of the art

- You just saw examples of current systems.
  - 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

Syllabus / Schedule (see handout)

http://tinyurl.com/UCBC280CAL

- Image Formation
- Image Filtering
- Pyramids & Regularization
- Feature Detection and Matching
- Geometric Alignment
- Calibration
- Geometric Image Stitching
- Photometric Image Stitching
- Recognition
- Stereo
- Optic Flow
- Dense Motion Models
- Shape from Silhouettes
- Shape from Shading and Texture
- Surface Models
- Segmentation
- SFM
- IBR & HDR…

And now, who are you?

- And what do you expect to get out of this class?
- Previous experience in vision, learning, graphics?
- Research agenda?
- (Project topics?)

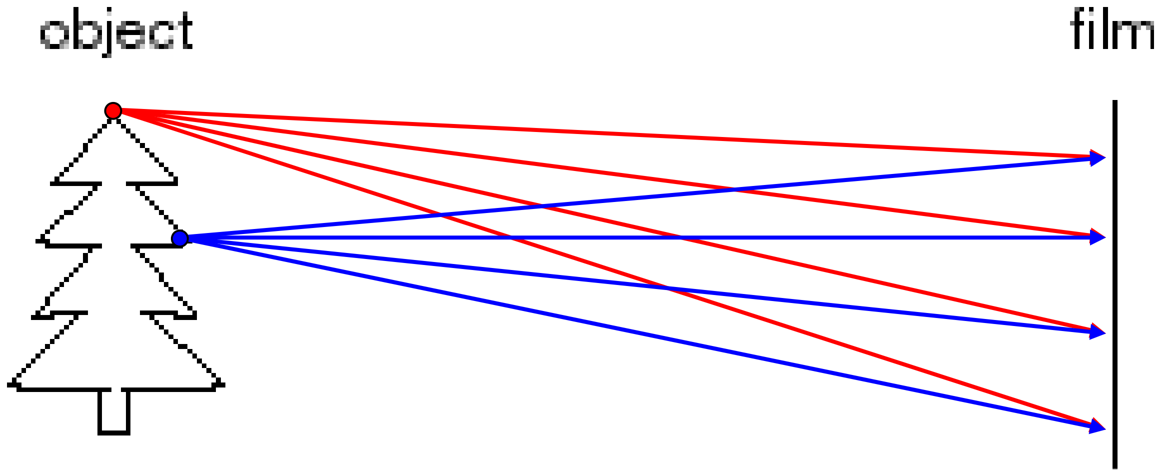
Let’s get started: Image formation

- How are objects in the world captured in an image?

Physical parameters of image formation

- Geometric
  - Type of projection
  - Camera pose
- Optical
  - Sensor's lens type
  - focal length, field of view, aperture
- Photometric
  - Type, direction, intensity of light reaching sensor
  - Surfaces' reflectance properties

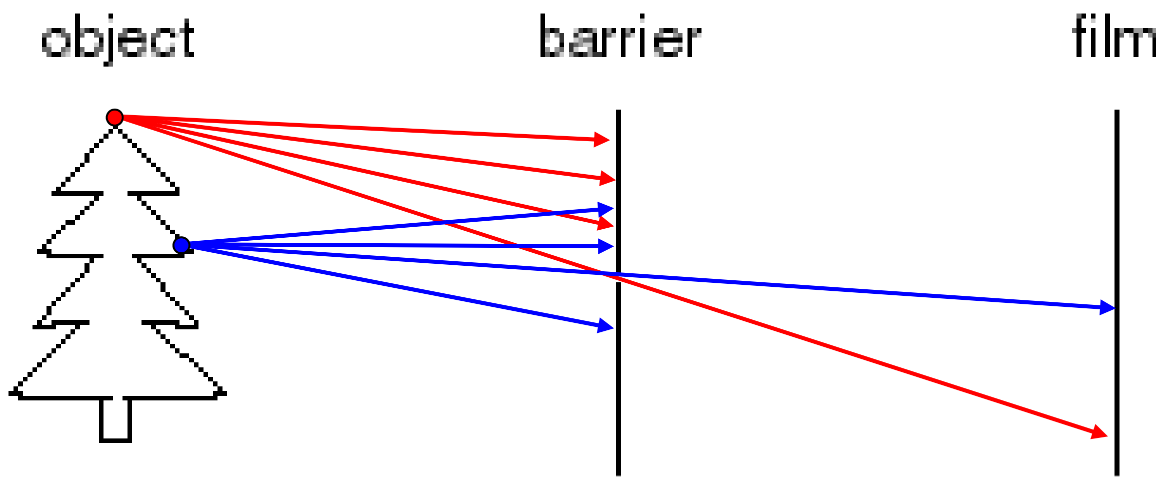
Image formation



- Let's design a camera
  - Idea 1: put a piece of film in front of an object
  - Do we get a reasonable image?

Slide by Steve Seitz

Pinhole camera

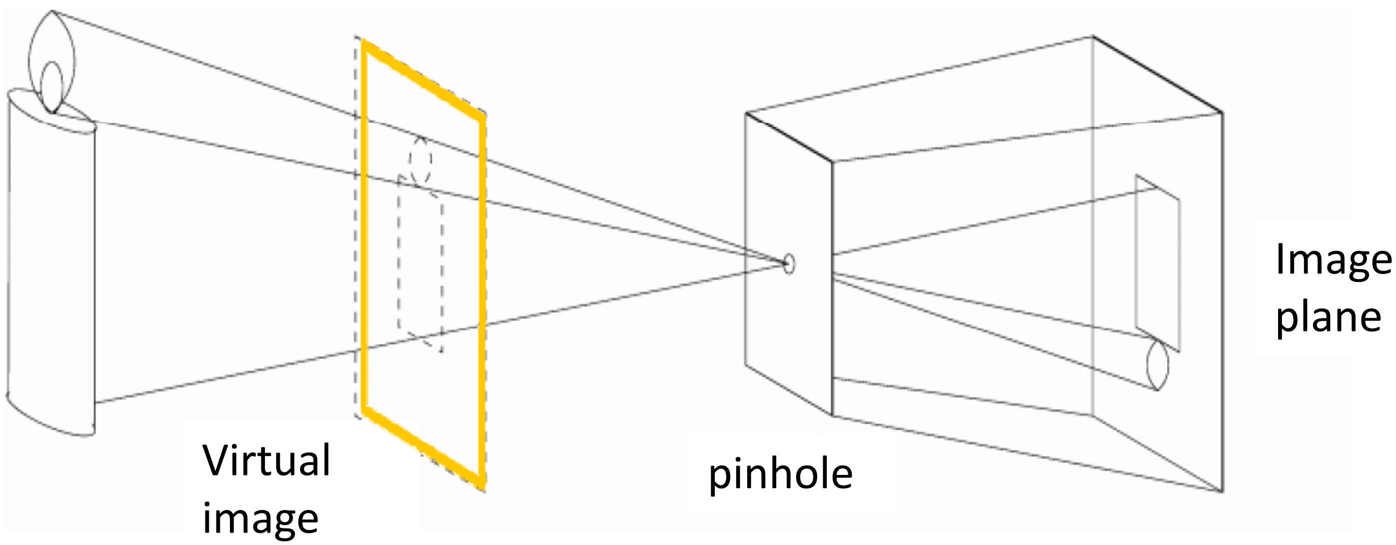


- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the **aperture**
  - How does this transform the image?

Slide by Steve Seitz

Pinhole camera

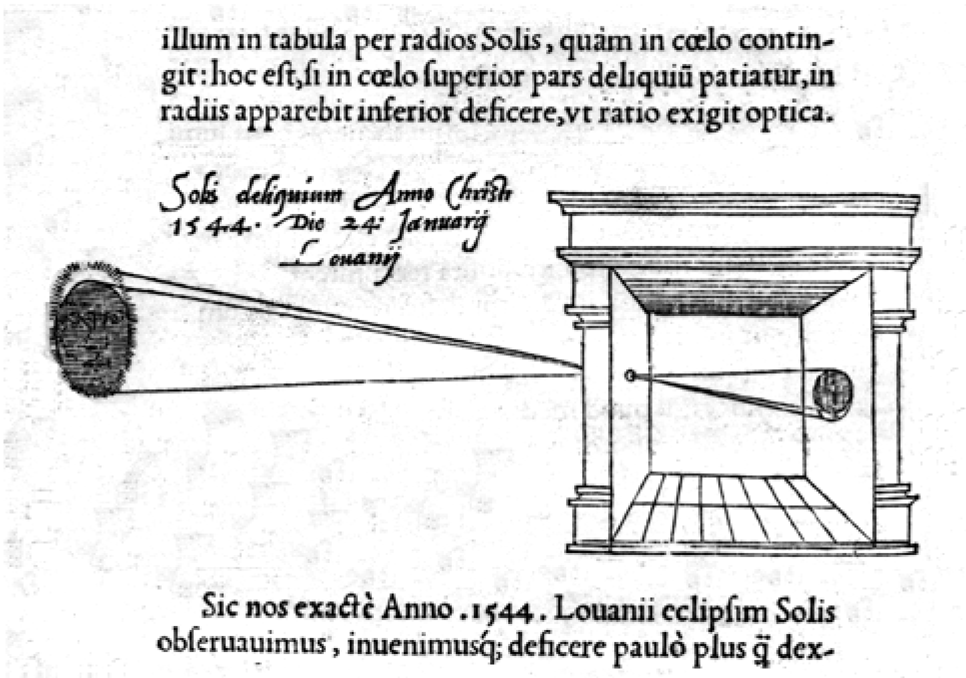
- Pinhole camera is a simple model to approximate imaging process, perspective **projection**.



If we treat pinhole as a point, only one ray from any given point can enter the camera.

Fig from Forsyth and Ponce

Camera obscura



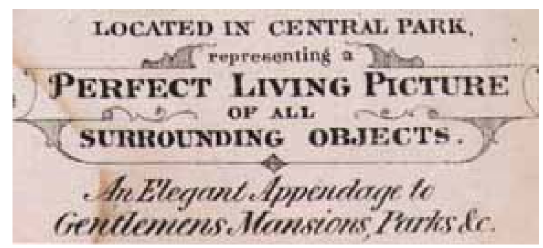
In Latin, means 'dark room'

"**Reinerus Gemma-Frisius**, observed an eclipse of the sun at Louvain on January 24, 1544, and later he used this illustration of the event in his book <u>De Radio Astronomica et Geometrica</u>, 1545. It is thought to be the first published illustration of a camera obscura..." Hammond, John H., <u>The Camera Obscura, A Chronicle</u>

http://www.acmi.net.au/AIC/CAMERA_OBSCURA.html

Camera obscura

Jetty at Margate England, 1898.



Around 1870s

An attraction in the late 19^{th} century

http://brightbytes.com/cosite/collection2.html

Adapted from R. Duraiswami

Camera obscura at home

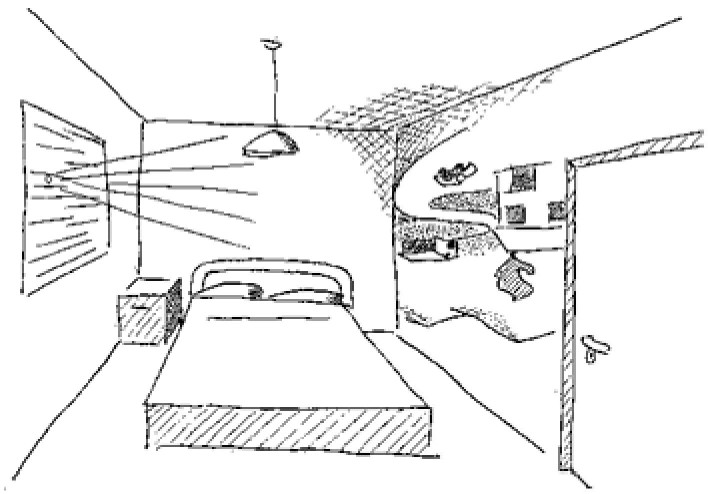


Figure 1 - A lens on the window creates the image of the external world on the opposite wall and you can see it every morning, when you wake up.

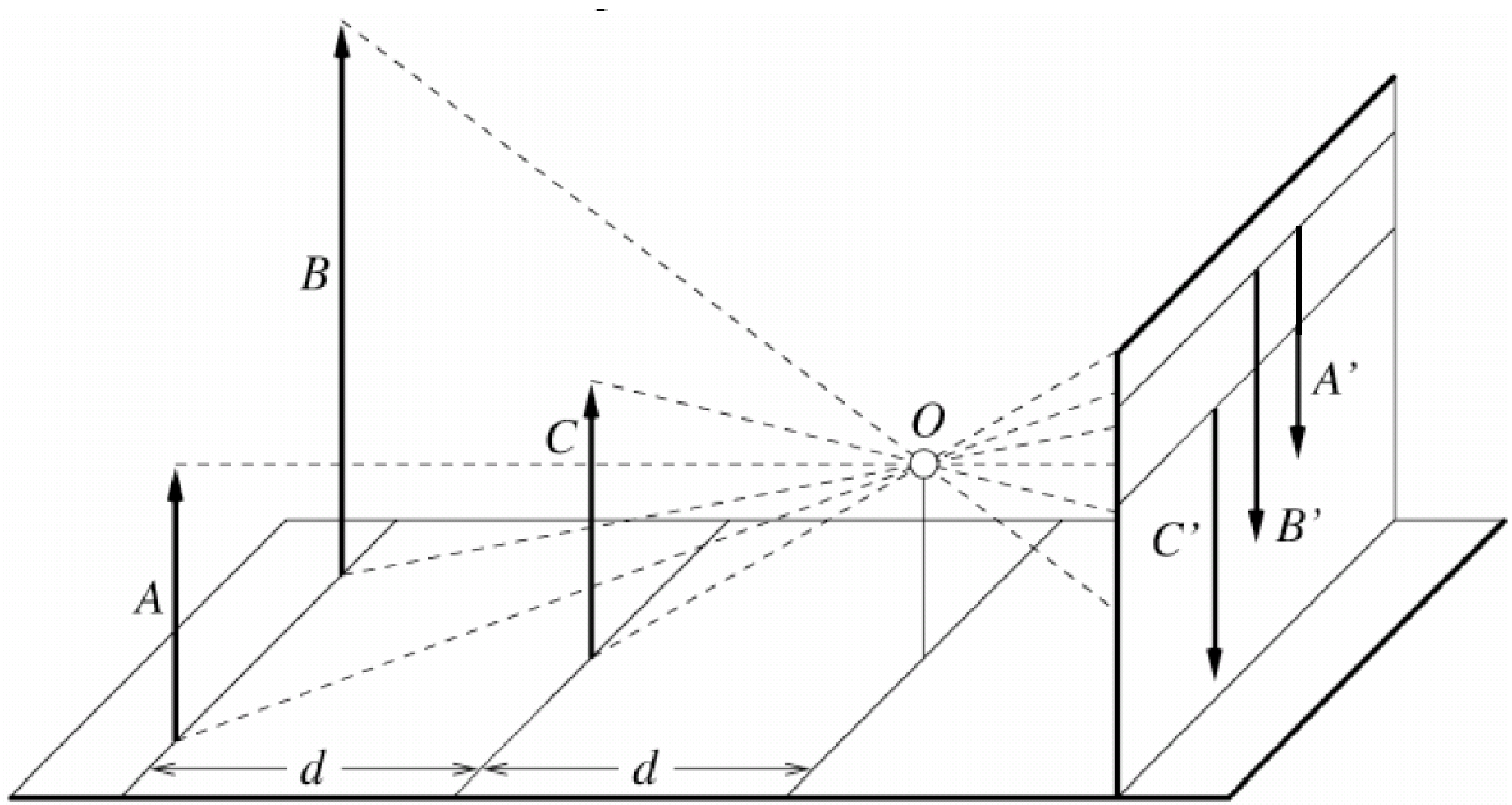
Sketch from http://www.funsci.com/fun3_en/sky/sky.htm

http://blog.makezine.com/archive/2006/02/how_to_room_sized_camera_obscu.html

Perspective effects

Perspective effects

- Far away objects appear smaller

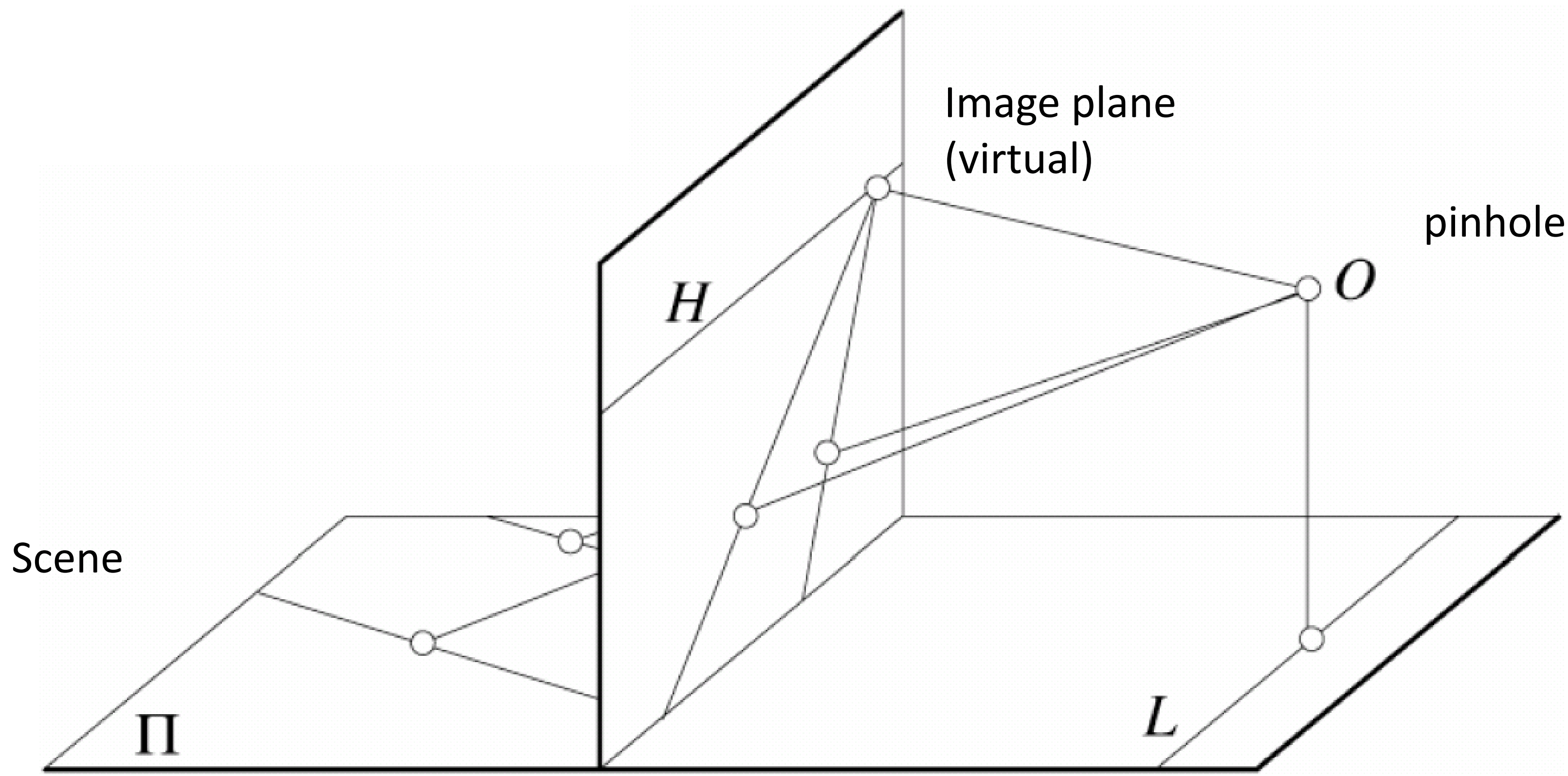


Forsyth and Ponce

Perspective effects

Perspective effects

- Parallel lines in the scene intersect in the image
- Converge in image on horizon line



Slide Credits

- Slides 14-21, 55-66: Kristen Grauman
- Slides 23-40,43-52: Steve Seitz
- and others, as marked…

Next time

- Continue with Image Formation
- Readings for today: Szeliski, Ch. 1
- Readings for next lecture: Szeliski 2.1-2.3.1, Forsyth and Ponce 1.1, 1.4 (optional).
- Pset 0 released tomorrow, due following Friday