CS 6476 Computer Vision
GTL Version

Introduction

Cédric Pradalier -- Aaron Bobick
School of Interactive Computing

Who are we?

Professor:

- Cédric Pradalier
  cedric.pradalier@georgiatech-metz.fr
- Office: GTL 222
- Office hours:
  - Anyday 10:00-17:00 by appointment (email is **much** better)

Original material from:

- Aaron Bobick
- ~~afb@cc.gatech.edu~~
- ~~Office: CCB 316~~

Where are you?

- **CS 6476 Computer Vision**

- **Piazza:** You will need to register to CS 6476 **GTL** on Piazza. If not, send me an email.
  - **Announcements** will likely be done through Piazza
  - Full syllabus in Piazza

- **Slides, Problems, Resources:** PDFs will be posted on the resource page on Piazza.

- **Matlab access (on windows)**: if you don’t know how to get Matlab access, first ask a friend. Then come see me. And ask about Python/OpenCV.

Link with other classes

- CS 7495 in Atlanta
  - Advanced Computer Vision (Grad Only)
  - Prof. Frank Dellaert
  - Pre-Req: CS 4495

- Our class: starting with material from CS 4495
  - Moving towards advanced topics depending on you

Who are you?

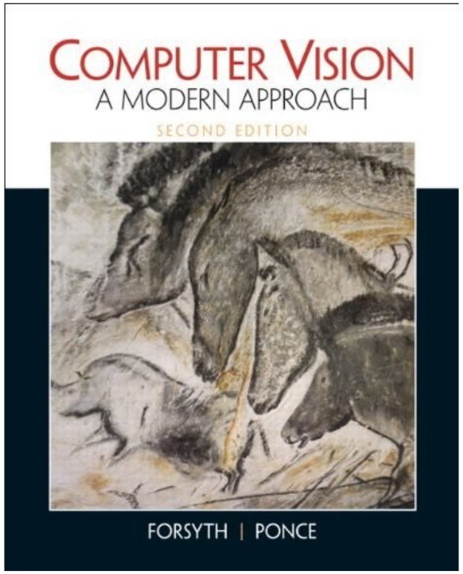
- *GTL:*
  - *Wide variety of prior experience depending on your original school, some of you might have had computer vision classes before.*
  - ***Your feedback will make sure we're moving at the appropriate speed, with the appropriate level of details.***

- What do you know?
  - Data structures – you'll be writing code that builds representations of images, features, and geometric constructions
  - Matlab (or Python with OpenCV or C++/OpenCV)
  - More math than most CS courses: Linear algebra, Vector Calculus, Linear algebra, Probability, Linear algebra
  - No CV assumed but maybe you've hacked an image?

What should you read?

Forsythe and Ponce:

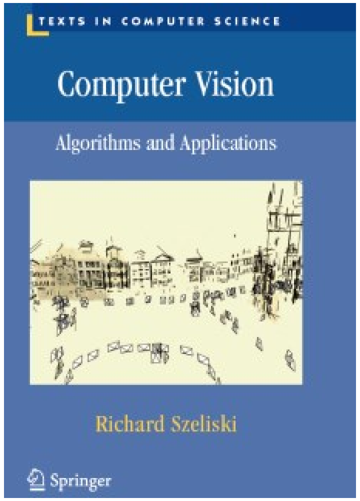
Computer Vision: A Modern Approach (2nd Ed)



(no pdf)

- Rick Szeliski's book:

  Computer Vision: Algorithms and Applications



(pdf)

Class format

- Flipped Classroom
  - Read the lecture **BEFORE** the class
  - Prepare 3 questions on tearable piece of paper
    - Good Exam Question
    - Hard point
  - Distribute the questions around the class.
  - Work in small groups to build answers to the questions
- What if you can't find the time to read the lecture and prepare the questions:
  - Case 1: you have a good excuse and prior approval. Good
  - Case 2: your cat ate your hard-drive.
    - **Don't come to class, you won't be accepted.**

What you will do?

- The grade is mostly problem set based (80%). Currently 8 tentatively scheduled:
  0. Due Sun Jan 12th midnight. This is just to make sure you can read, numerically manipulate, and display images. ***This one is no work. Others are serious work.***
  1. Edges, lines, Hough transform
  2. Stereo Matching
  3. Absolute calibration – where is the camera given the points.
  4. Relative motion – SIFT features to find where the cameras are
  5. Motion flow
  6. Tracking
  7. Motion History Images (action recognition)

- And there will be an exam – 20%

Project Grading

- Three point scale:
  - Below Average
  - Satisfactory: You did your best to answer all the questions. There may still be a bug, but you know where it is and have showed some insight about it.
  - Above Expectations: Satisfactory + everything works and you explored the parameter space, found you own examples that challenge the algorithm, optimized your code, …
- You must be in average Satisfactory to hope for an A (see next slide).
- Projects cannot be accepted late without prior approval.
- Each non-returned project costs one letter grade point.

Exam and other grades

- Exam form:
  - Selected 30 questions out of the ~600 seen in class.
  - Answers fit in 3-4 cm.
  - Closed-book
  - Questions will be collected on a google spreadsheet (no surprise).

- How not to get an A in this class (OR-list):
  - Don’t return a project.
  - Crash the exam.
  - Don’t prepare the lecture questions twice.

When will you do them?

- They will all be due on Sunday at 11:59pm.
- Late submissions will only be accepted at full credit with prior approval. Otherwise 50% (yes half) reduction.
- Prof ***not*** obligated to get back to you about permission the weekend it is due! At your own risk.

With whom will you do it?

- Honesty/Integrity policy (from web site):

  Problem sets are to be done individually but you may collaborate at the "white board level" helping each other with algorithms and general computation, BUT YOUR CODE MUST BE YOUR OWN.

- Do not hand in other people's code unless you (1) say you are, and (2) you want no credit for that section. We will be explicit about what previous or provided code you can use.

How will you do it?

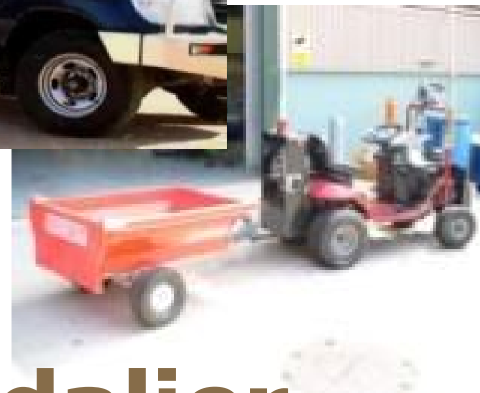
- We will mostly support Matlab
  - In class, most of the time we will use Matlab to handle the image processing (though we'll learn some of that) and then your data structures for the vision.

- If you want be even more "job ready"
  - OpenCV is a Python (and C/C++) package for computer vision (Linux or Windows)
  - Will make sure problem sets are Python-doable.
  - If possible, link with real-time algorithm in ROS (see CS 7630)

Any questions so far…

Who am I…

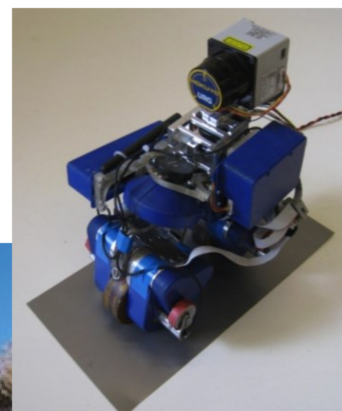
- Background: degrees in Math, CS (ENSIMAG), PhD in Robotics, Imaging and Graphics (Inria); interest in robotic systems and field application
- Post-Doc at CSIRO, Australia
- Deputy director, Autonomous Systems Lab, Zurich, Switzerland
- Faculty at GTL since Jan 2013.

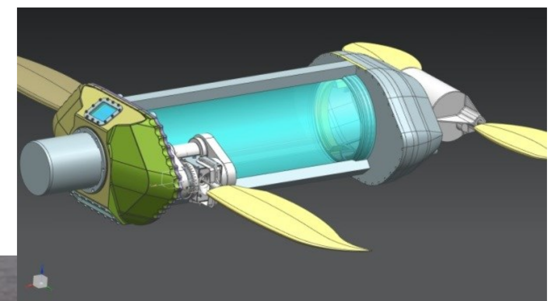


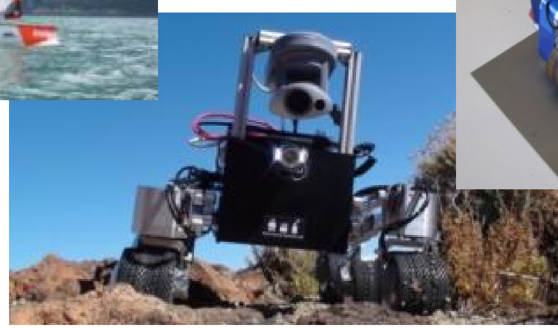
Prof. Cédric Pradalier

cedric.pradalier@georgiatech-metz.fr

DREAM lab: Data-driven Robotic for Environment Assessment and Monitoring



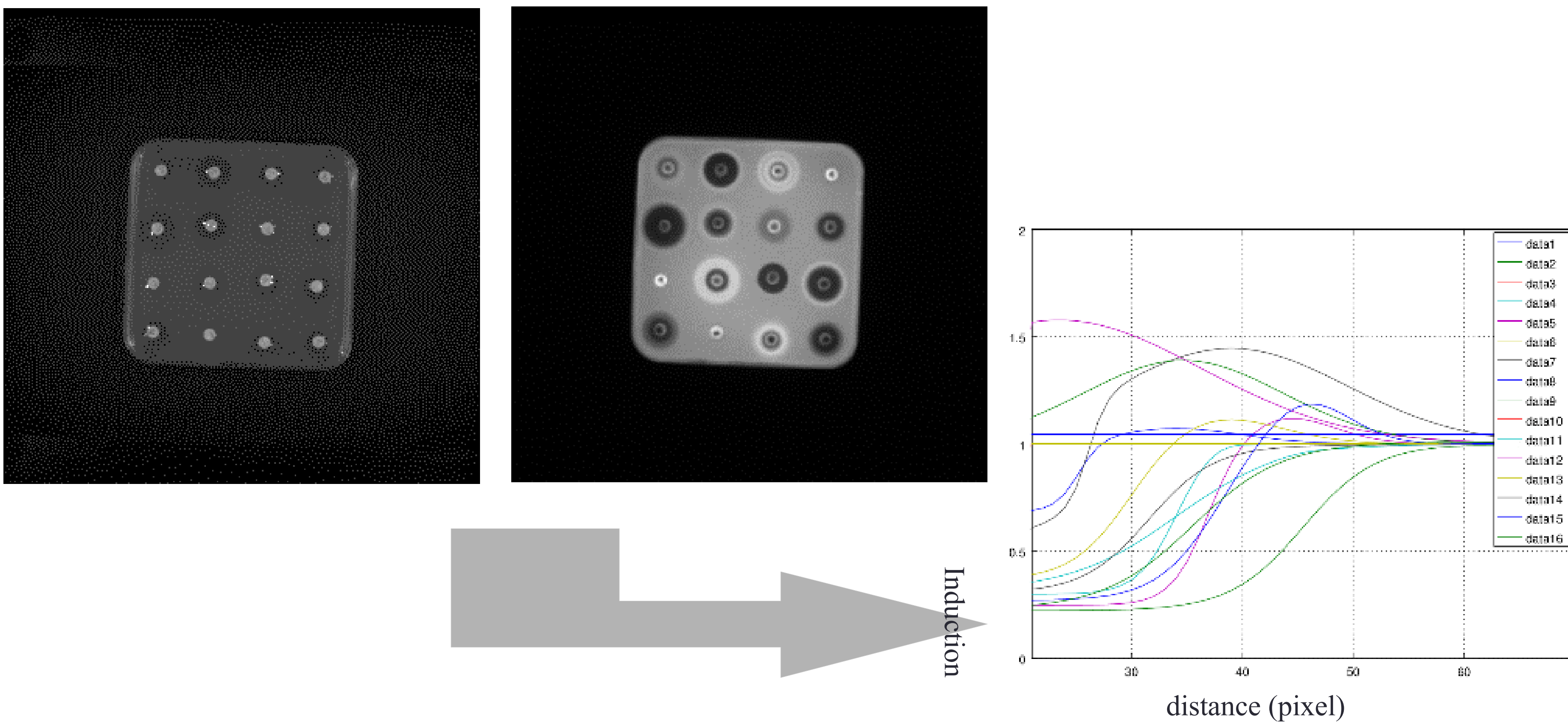




Example: Image Processing for Feature Detection.

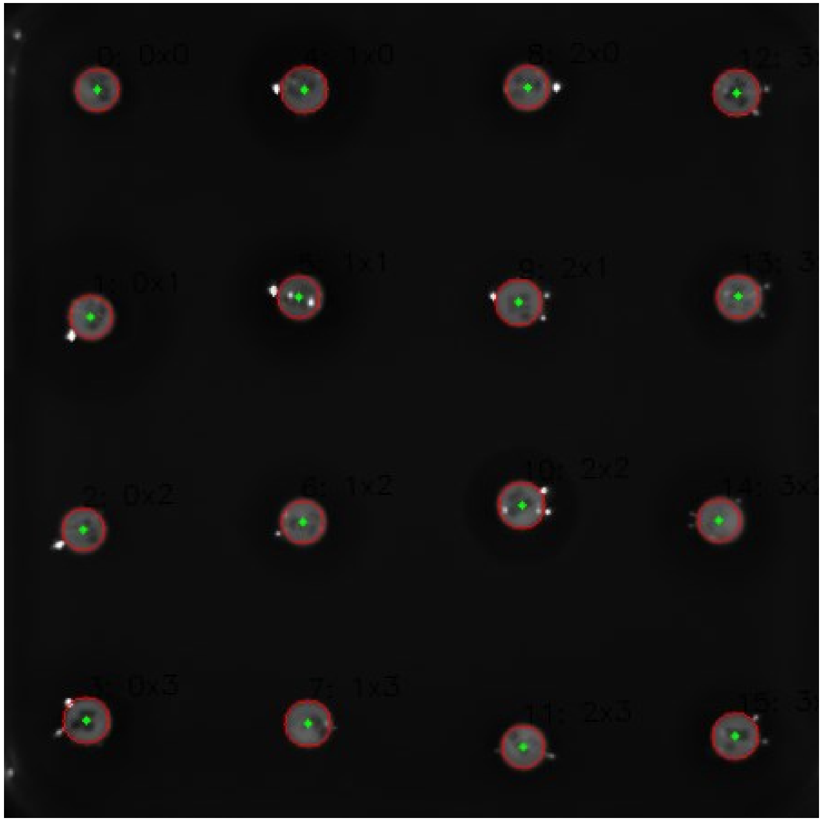
Objectives

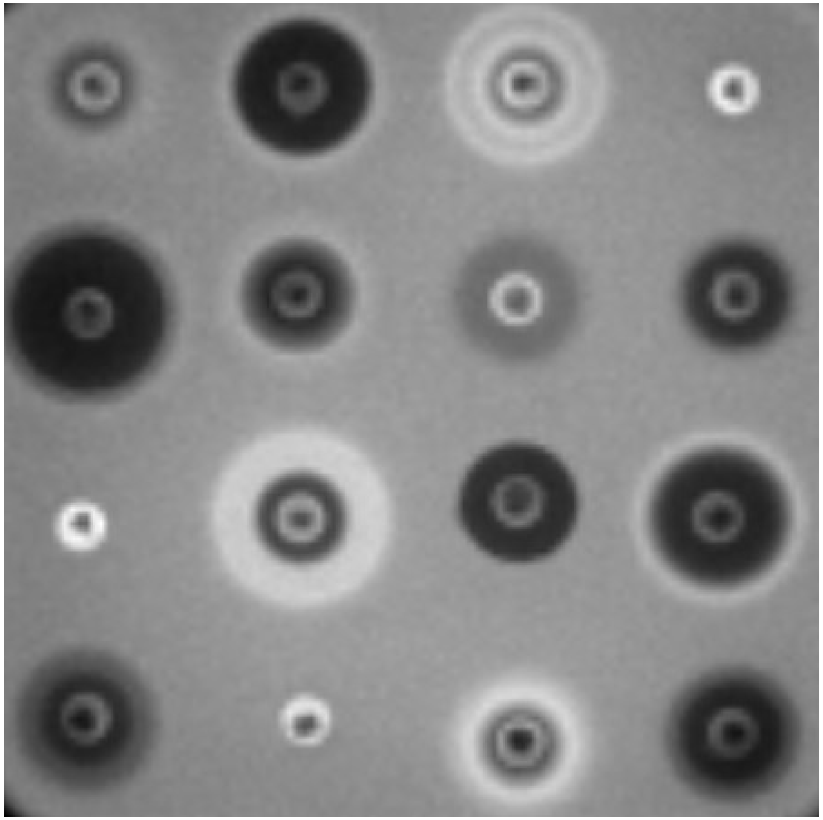
- From standard and luminescence images to antibiotic induction profiles



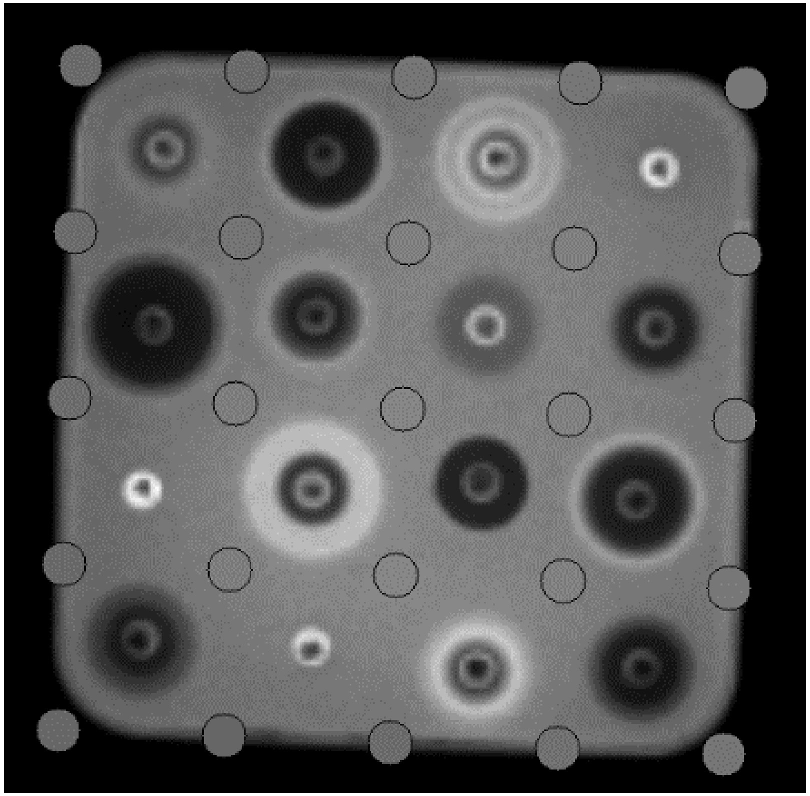
Disk detection

- After extraction of aligned image around 4x4 mesh (standard and luminescence):





Background normalisation

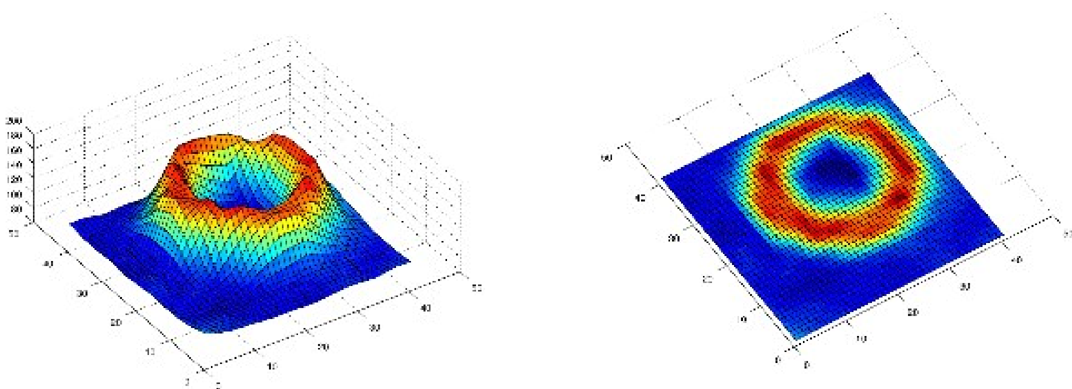


5x5 background mesh

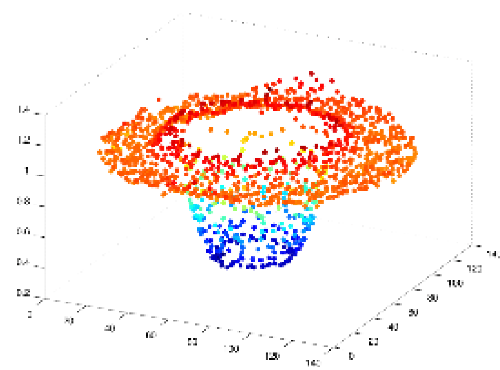
Resulting synthetic background

Data collection

- Take all pixels outside of 1.5 disk radius and less than 75% of the inter-disk distance
  - Allows accounting for synergies
  - Special care around box borders to avoid artefacts: ignore any measurement between disk center and box border.



Raw luminescence (3D / bird's eye view)

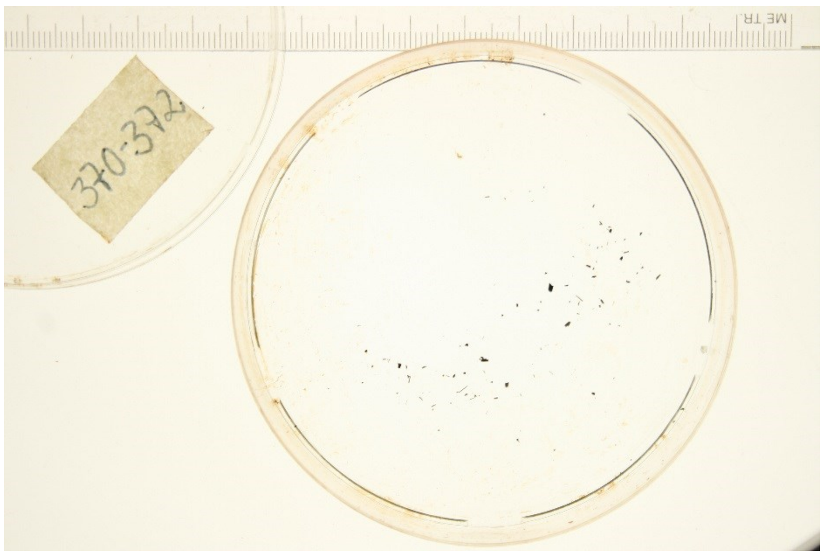


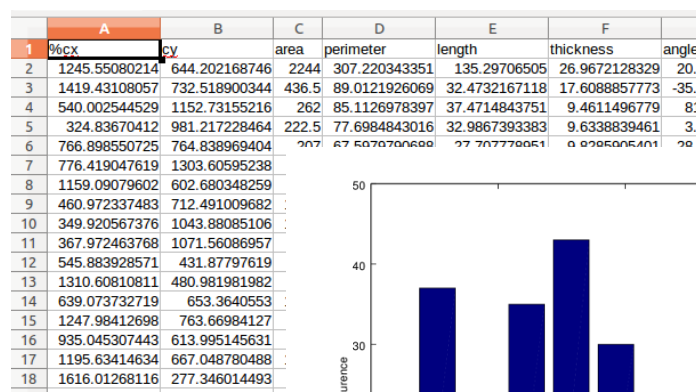
Induction (after division by backgrounc

Example: Image Processing for Feature Detection.

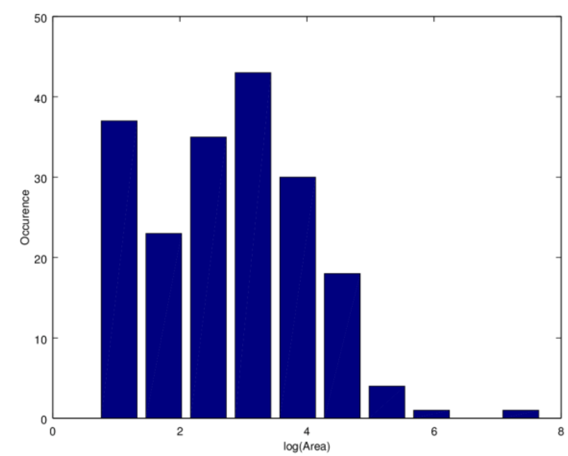
Objectives

- Quantification of coal particle volume for paleo-ecology



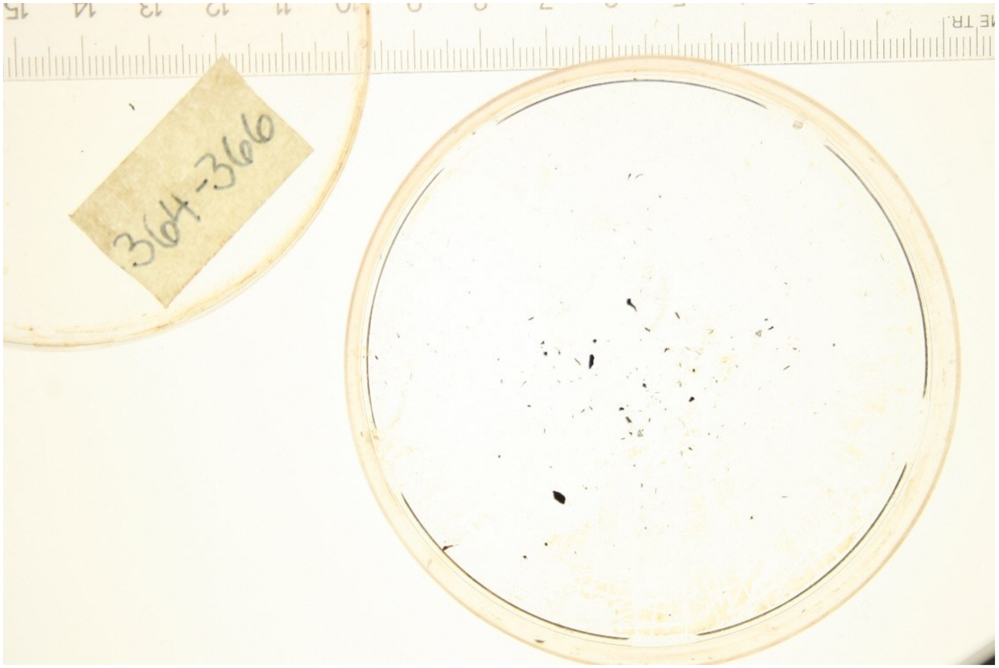


	A	B	C	D	E	F	
1	%cx	cy	area	perimeter	length	thickness	angle
2	1245.55080214	644.202168746	2244	307.220343351	135.29706505	26.9672128329	20.0
3	1419.43108057	732.518900344	436.5	89.0121926069	32.4732167118	17.6088857773	-35.1
4	540.002544529	1152.73155216	262	85.1126978397	37.4714843751	9.4611496779	81.
5	324.83670412	981.217228464	222.5	77.6984843016	32.9867393383	9.6338839461	3.9
6	766.898550725	764.838969404	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
7	776.419047619	1303.60595238					
8	1159.09079602	602.680348259					
9	460.972337483	712.491009682					
10	349.920567376	1043.88085106					
11	367.972463768	1071.56086957					
12	545.883928571	431.87797619					
13	1310.60810811	480.981981982					
14	639.073732719	653.3640553					
15	1247.98412698	763.66984127					
16	935.045307443	613.995145631					
17	1195.63414634	667.048780488					
18	1616.01268116	277.346014493					



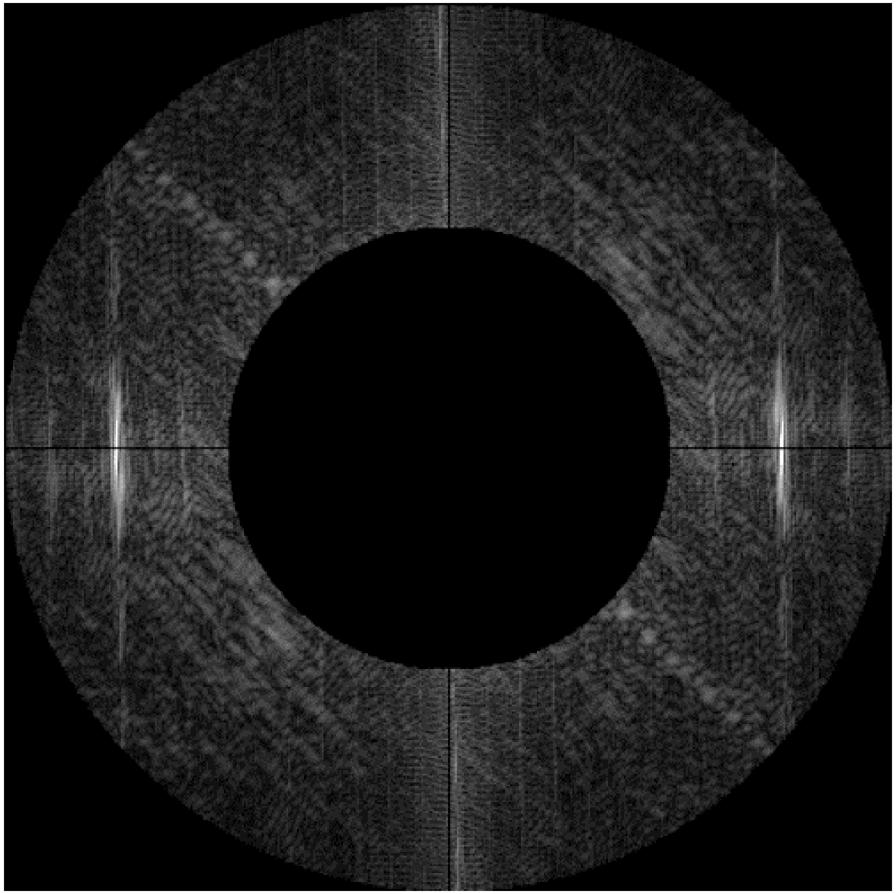
Particle Detection

Detection of Petri dish, thresholding, connected components, contours



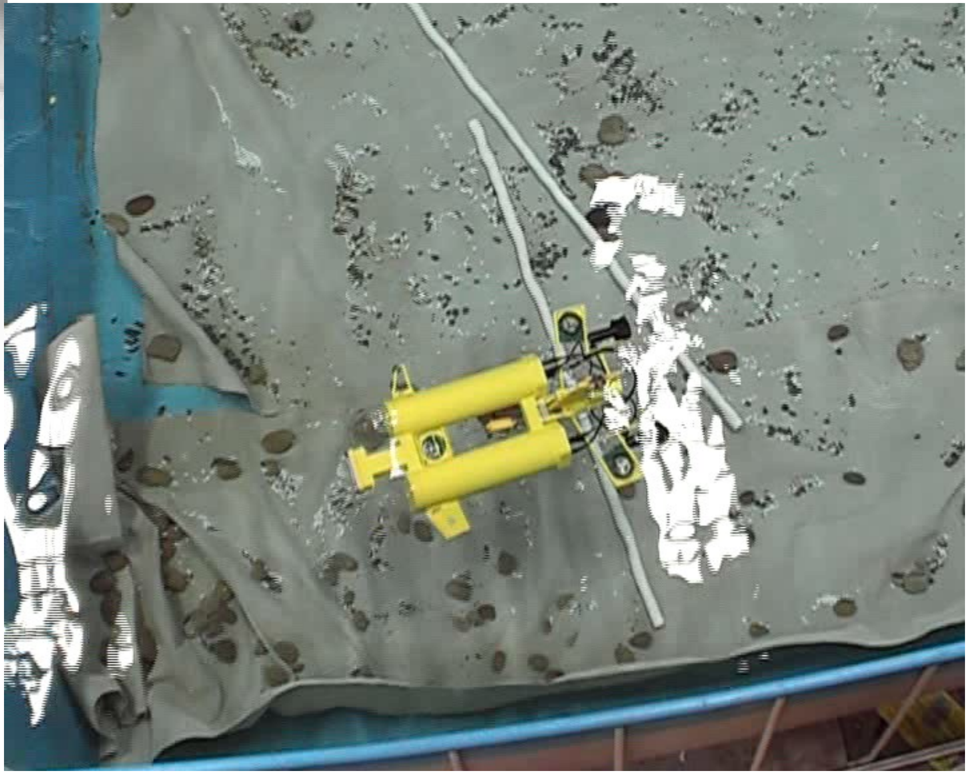
Scale detection

- FFT, ruler edge detection, period detection

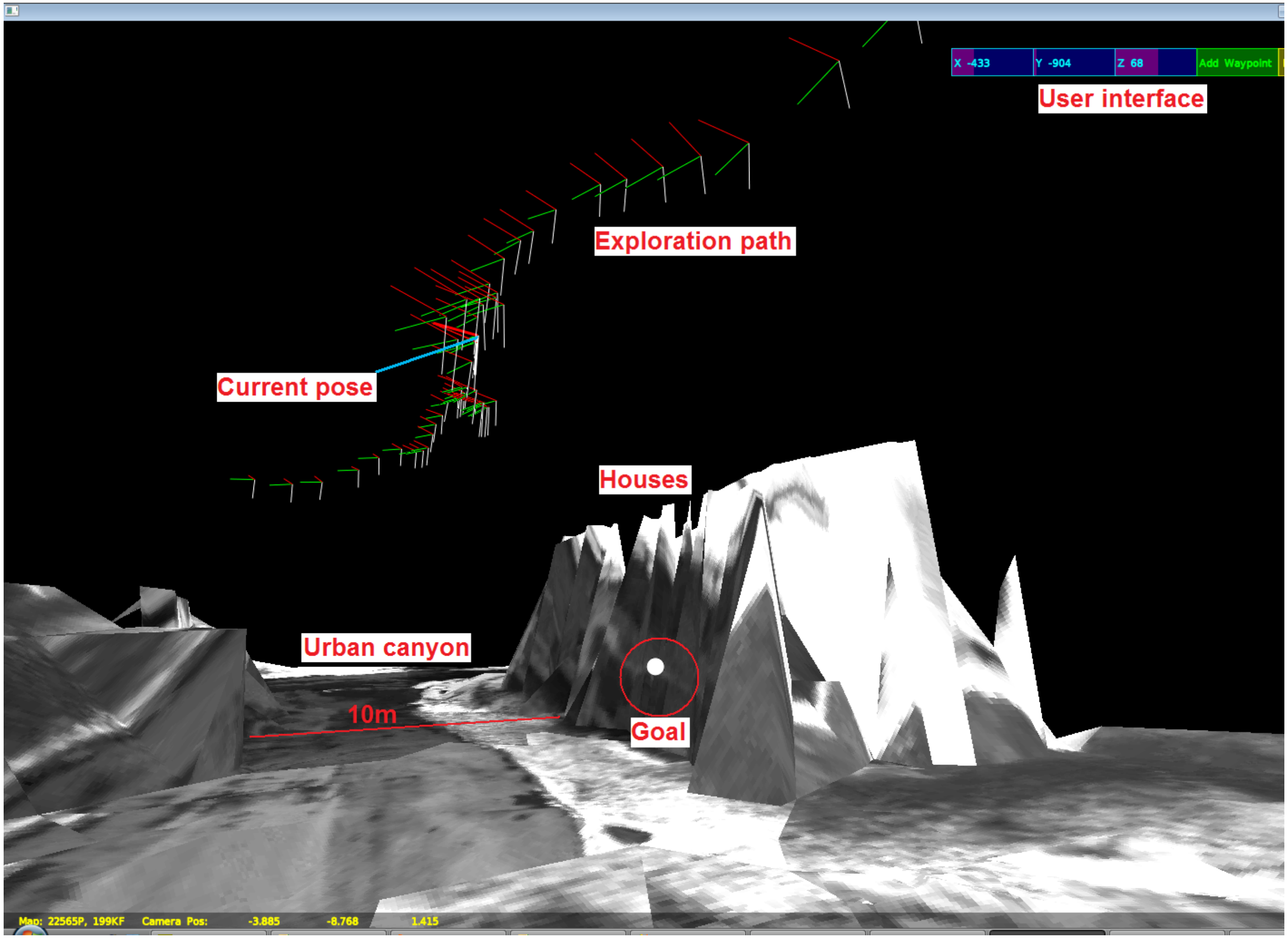


Other examples from robotics

Underwater Robotics



Micro-Aerial Vehicle

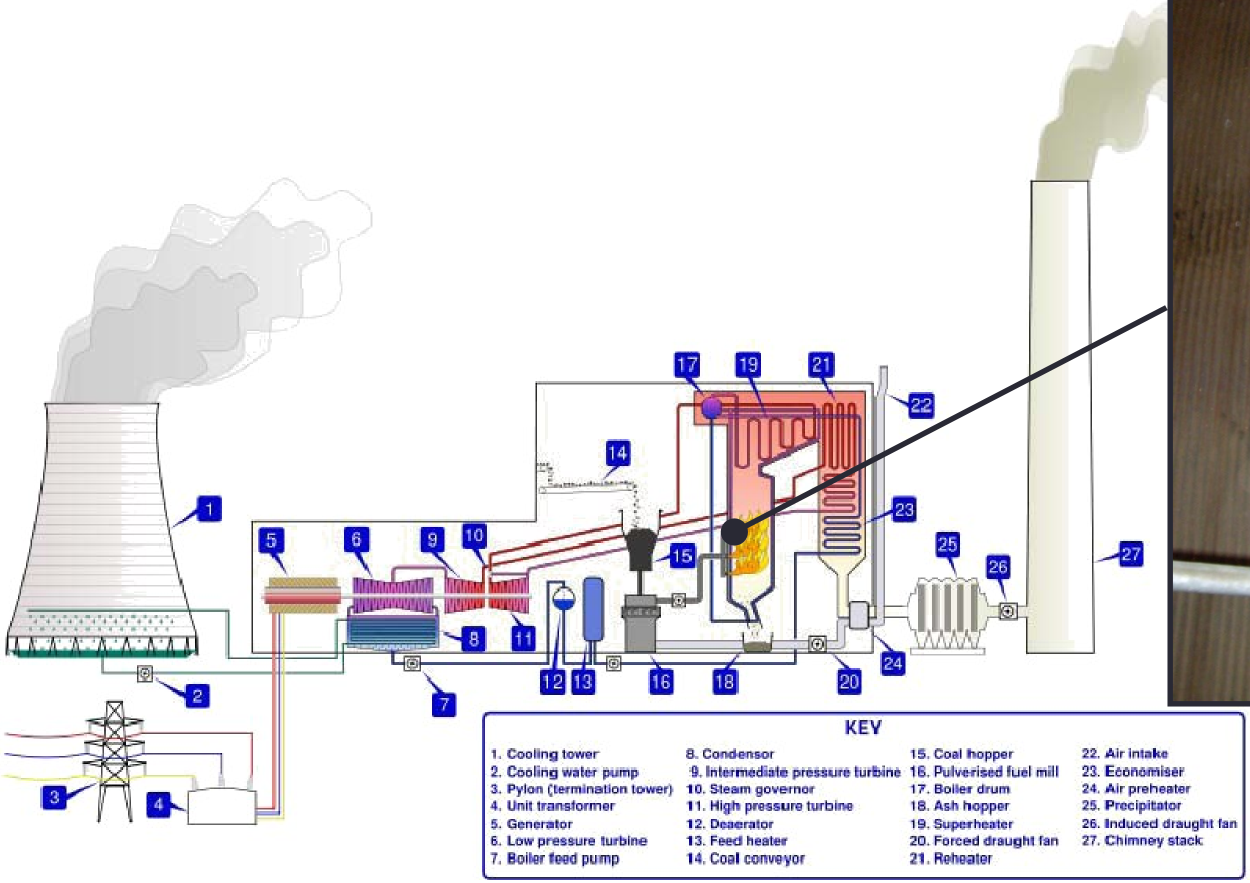


Micro-helicopters for Inspection

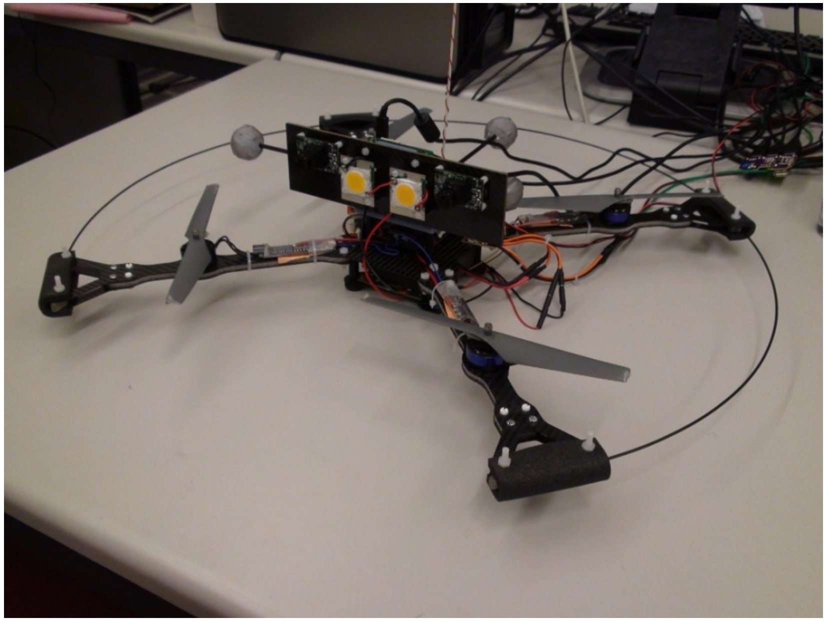
Vision, (GPS)-INS
High accuracy requirements
High dynamics
Strong energy constraints

AIRobots – Navigation with Stereo

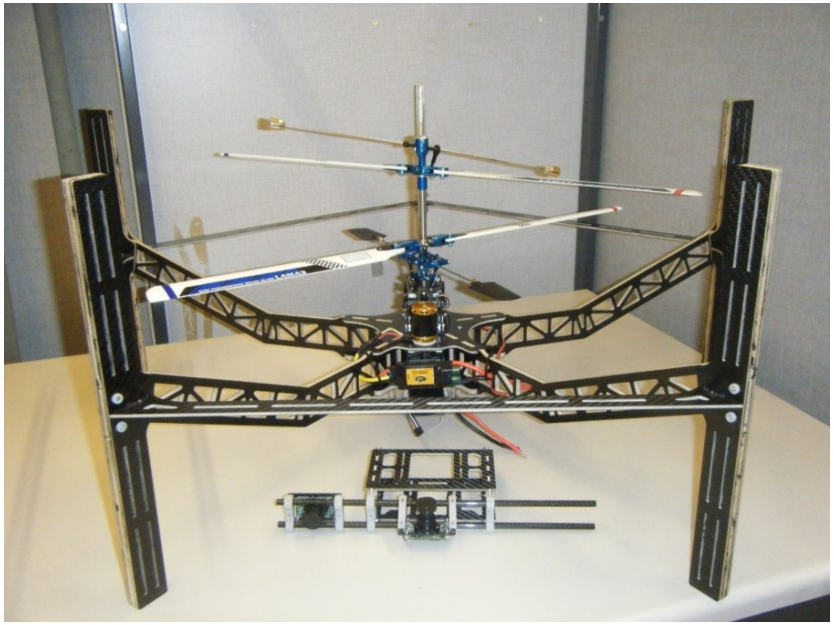
- Inspection of a coal boiler using aerial vehicles



Micro-helicopters for inspection



First Prototype



Second Prototype

Characteristics

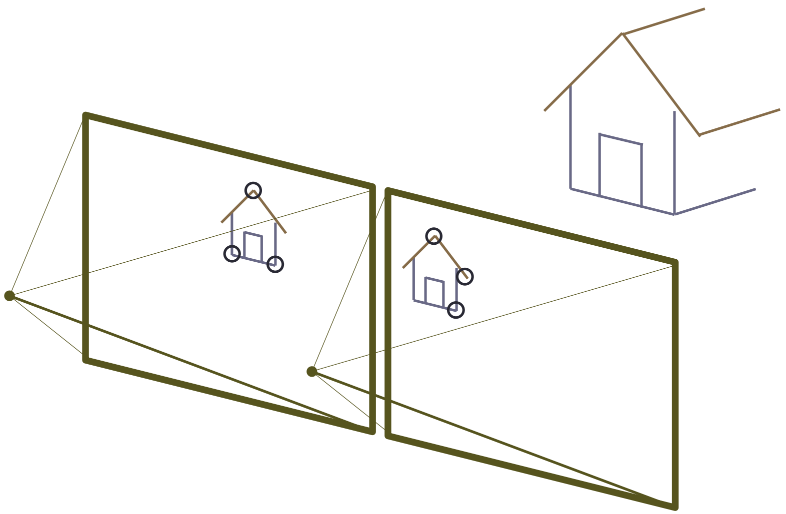
- Localisation:
  - Stereo camera + IMU (+ onboard lights): computationally and energetically expensive
- Mapping:
  - Online SLAM
- Path Planning:
  - Predefined path segments (for now)
- Task Scheduling:
  - Simple: take-off, fly segments, land…
- Obstacle Avoidance:
  - Using the Stereo Cam (not addressed yet)
- Control:
  - Complex flight dynamic + controlled contact with the wall surfaces

Challenges

- 3D environment
  - Dark and very self-similar
- Low computational resources (weight/energy)
- Complex control and obstacle interaction
  - Cluttered environment
  - Contact with the walls

Visual Egomotion – Keypoints, Descriptors

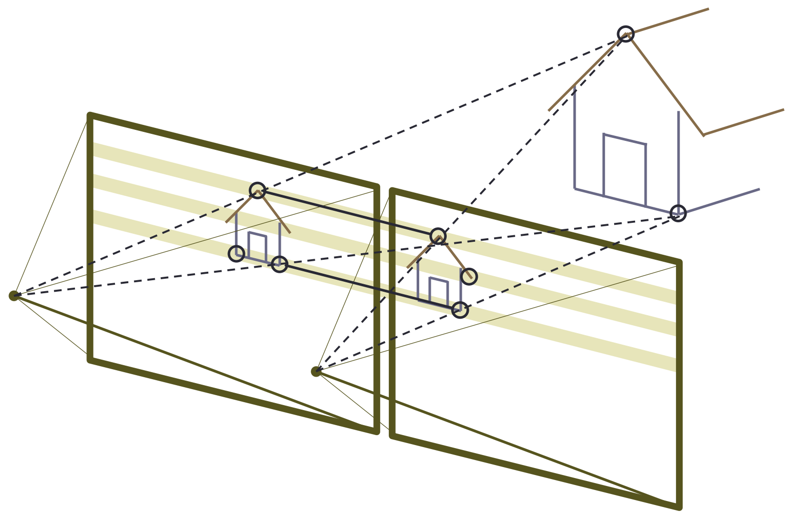
- Capture stereo shot



- Extract key points
  - FAST corner detector
  - Adaptive thresholding
- Compute key point descriptors
  - BRIEF feature descriptor

Visual Egomotion – Stereo Triangulation

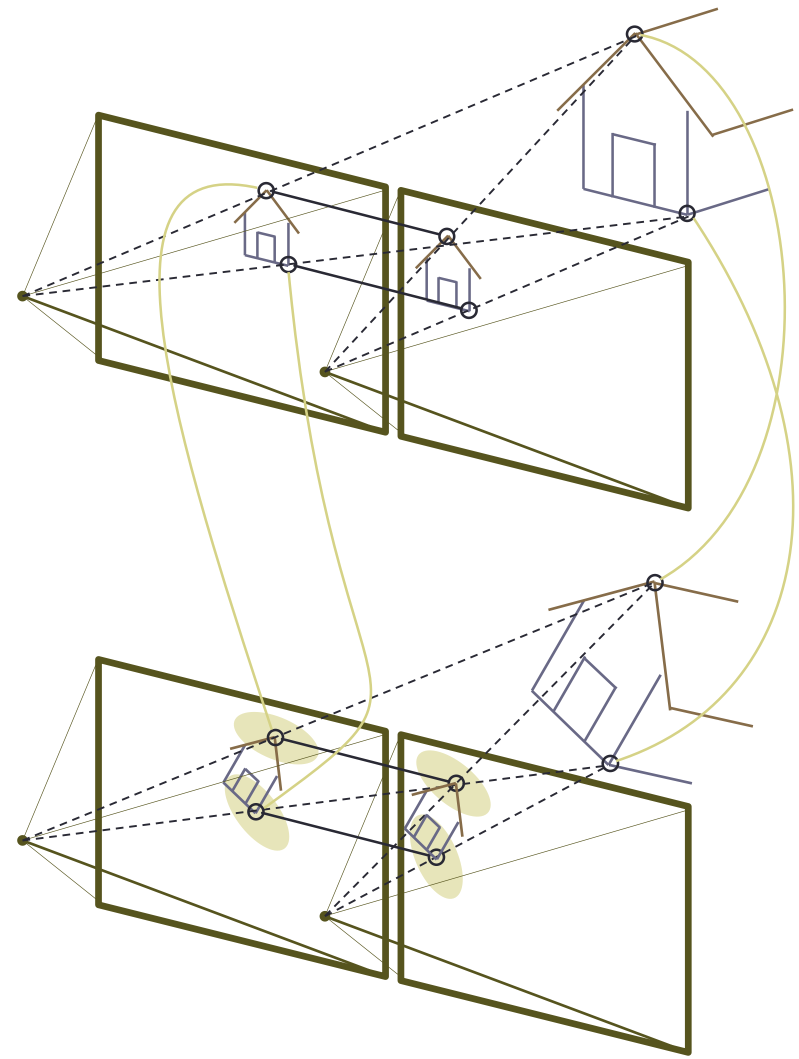
- Associate features of left and right image
  - Epipolar constraint
  - Descriptor matching



- Triangulate associated features to obtain 3D points

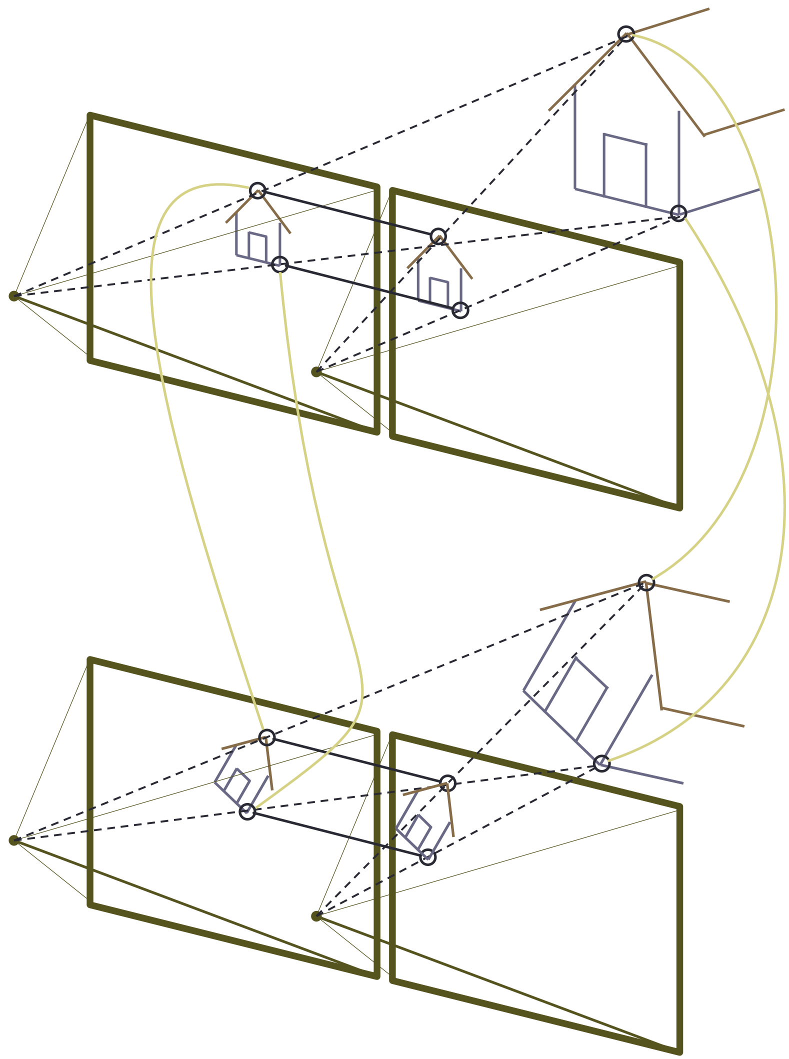
Visual Egomotion – 3D Feature Matching

- Capture next stereo shot
- Compute key points, descriptors and 3D points as before
- Associate features
  - Descriptor matching
  - IMU motion constraints



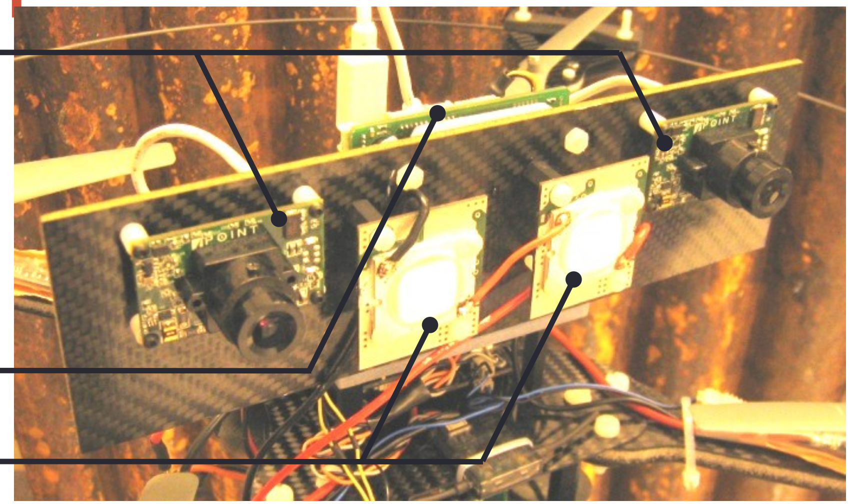
Visual Egomotion – Pose Estimation

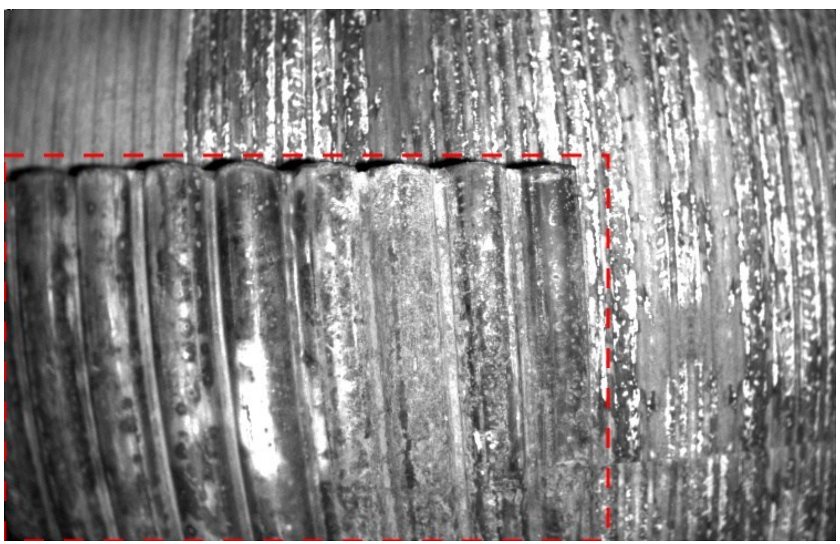
- RANSAC outlier rejection
  - P3P motion hypotheses
  - Apply density filter before counting hypothesis inliers
- Refinement via bundle adjustment

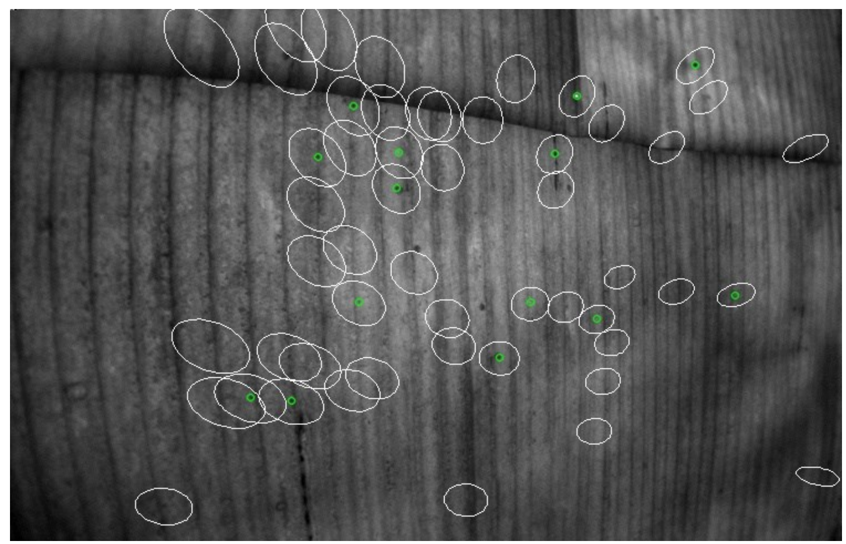


Results – Boiler Experiments

- Stereo Rig
- Mockup
- Uncleaned boiler surface

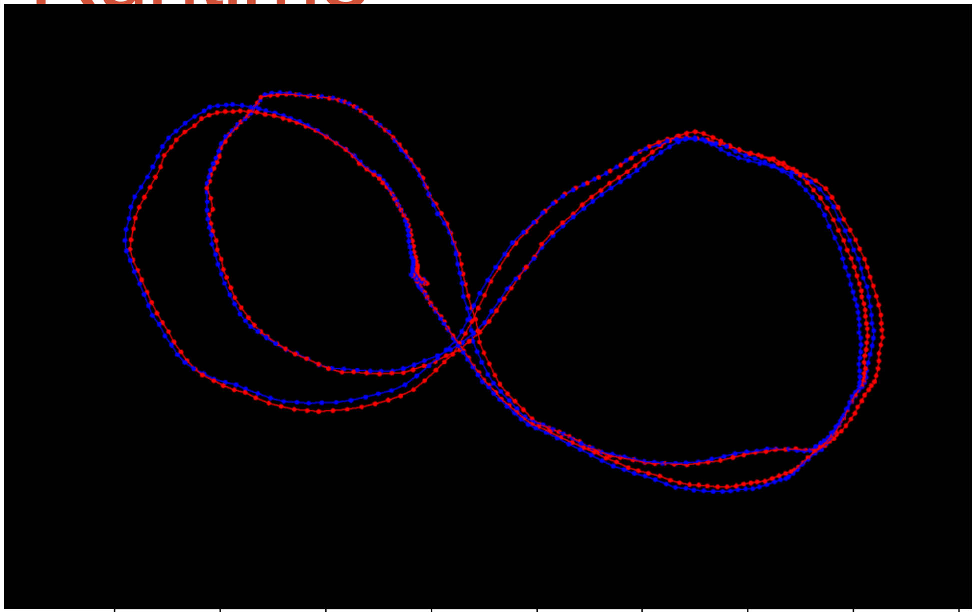




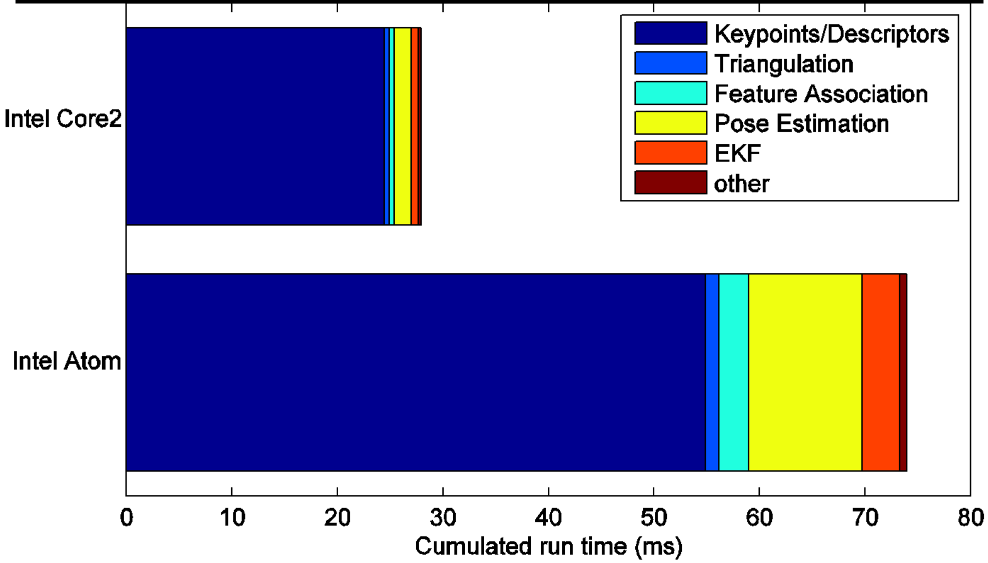


Results – Accuracy, Runtime Performance

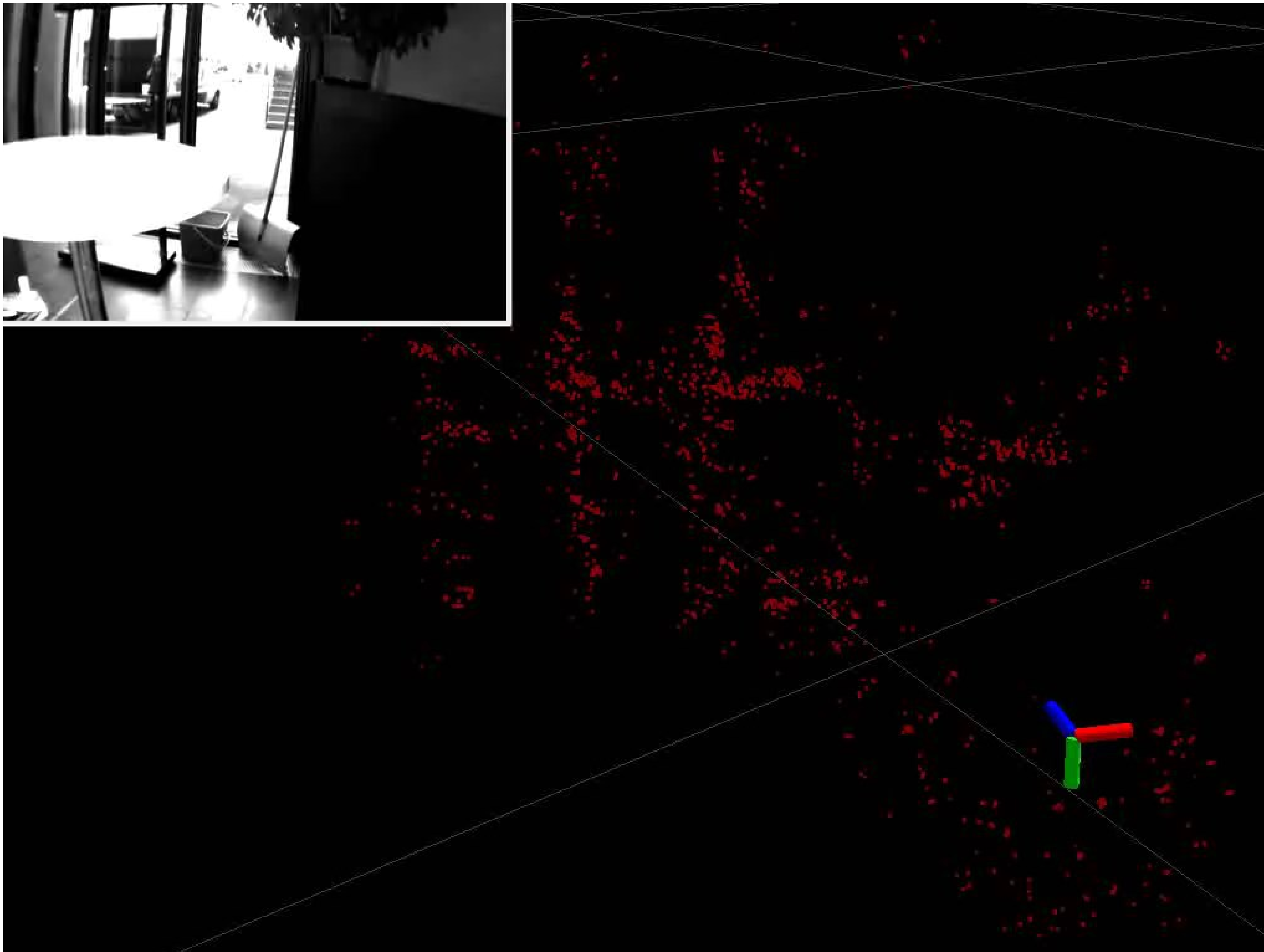
- Final error ~0.1% to Vicon ground truth



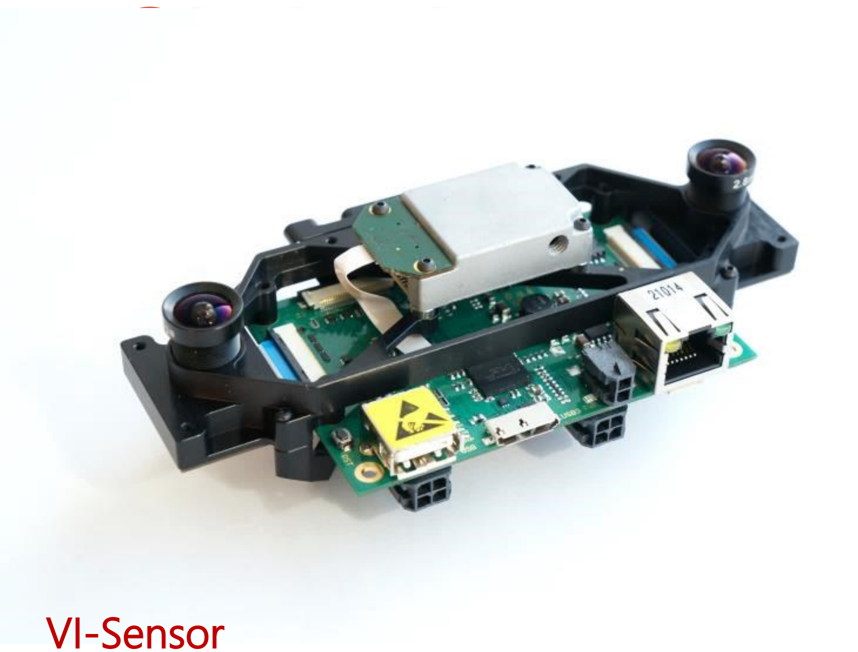
- Runs at 10Hz – 15Hz on single core Intel Atom



Results: Indoor Odometry



Results: Helicopter Control



VI-Sensor
- Visual-inertial Sensing Platform
- Stereo-Camera
- Industrial-grade IMU

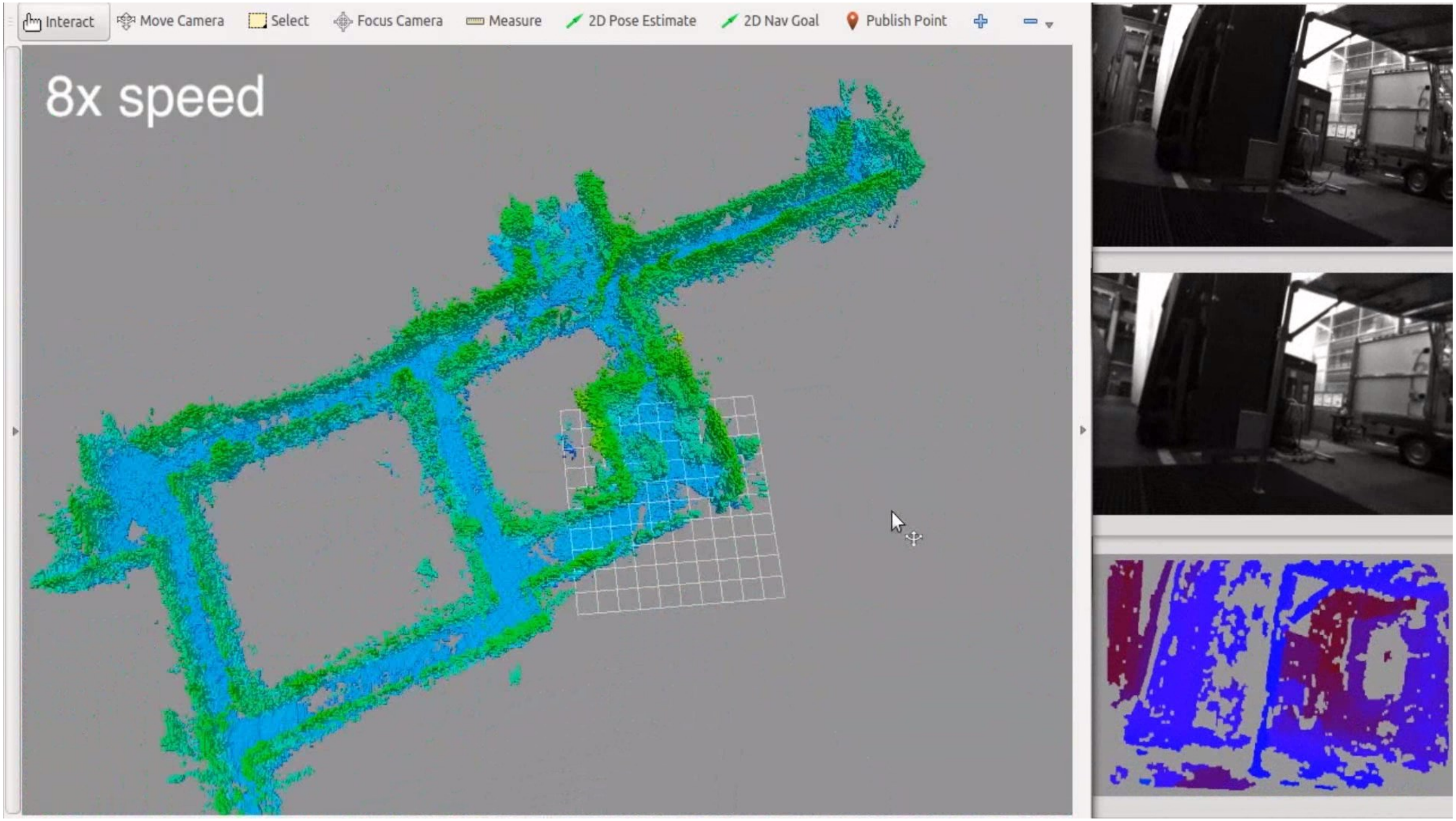
Datasheet at www.skybotix.com

VI-Sensor Firefly
- UAV: Ascending Technologies Firefly
- Skybotix VI-Sensor
- Skybotix Visual Positioning

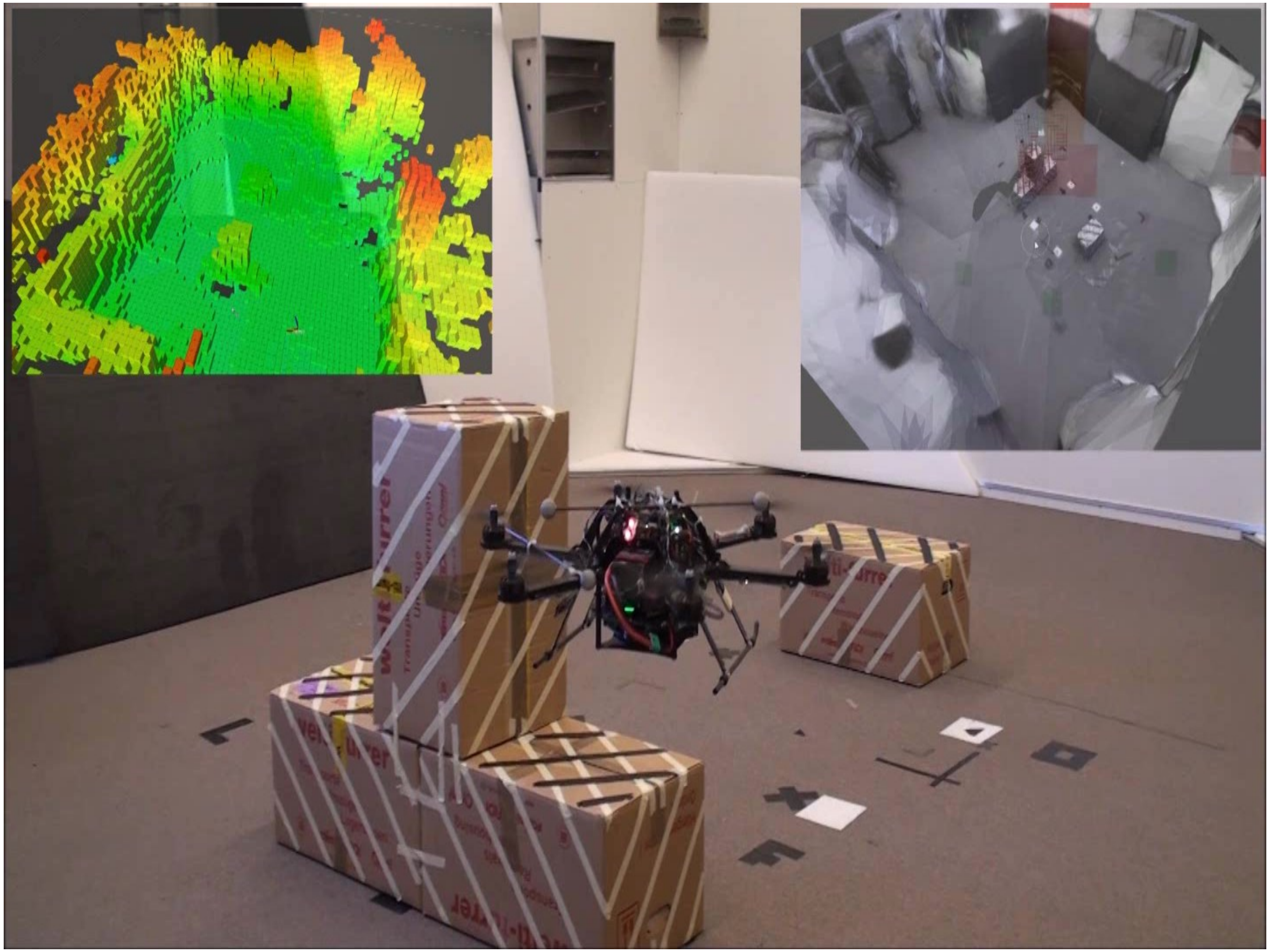
UAV Partner

ASCENDING TECHNOLOGIES

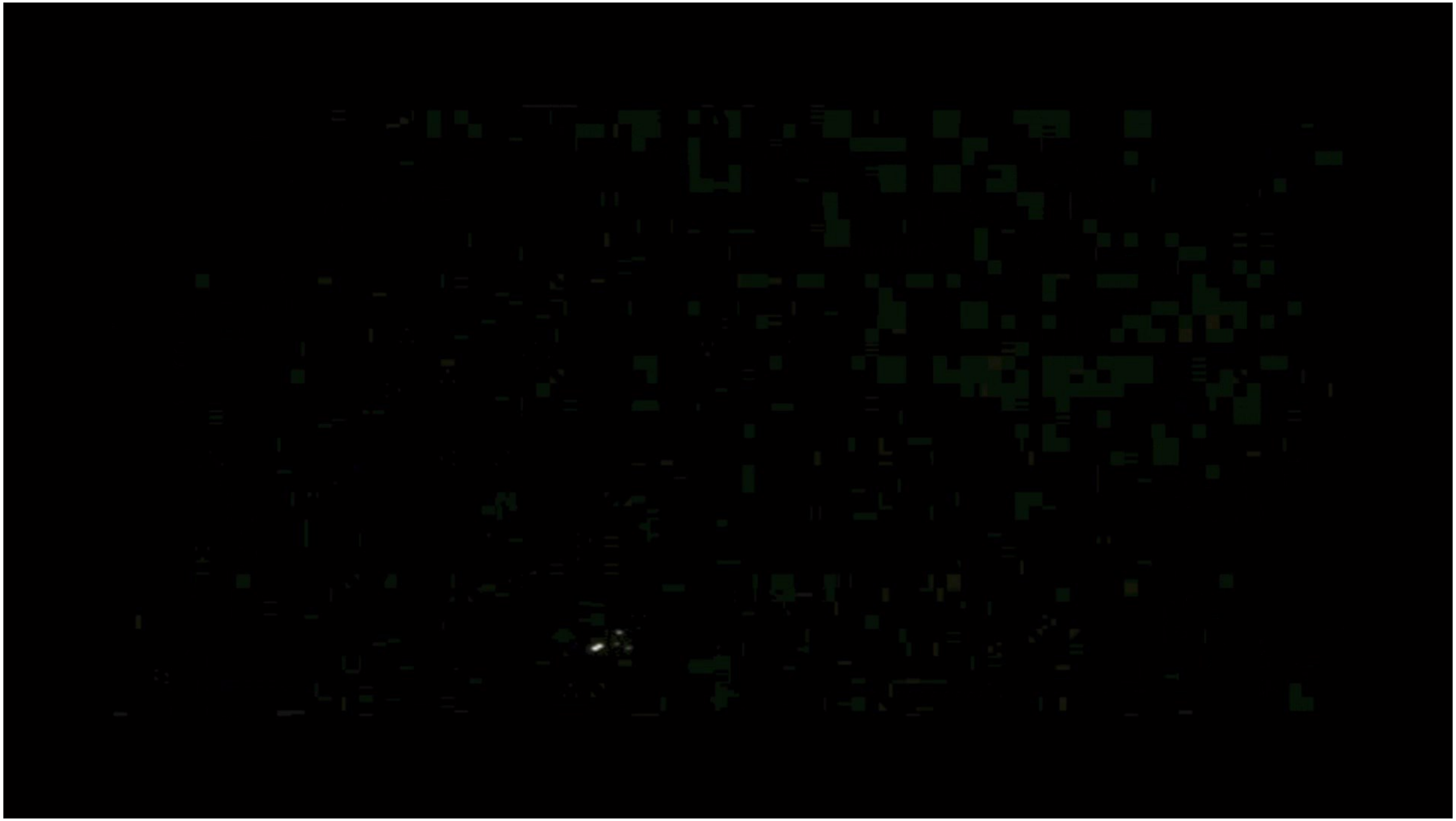
Visual Mapping



Obstacle Avoidance



Boiler Inspection



Search and rescue

- External Video

Façade Inspection

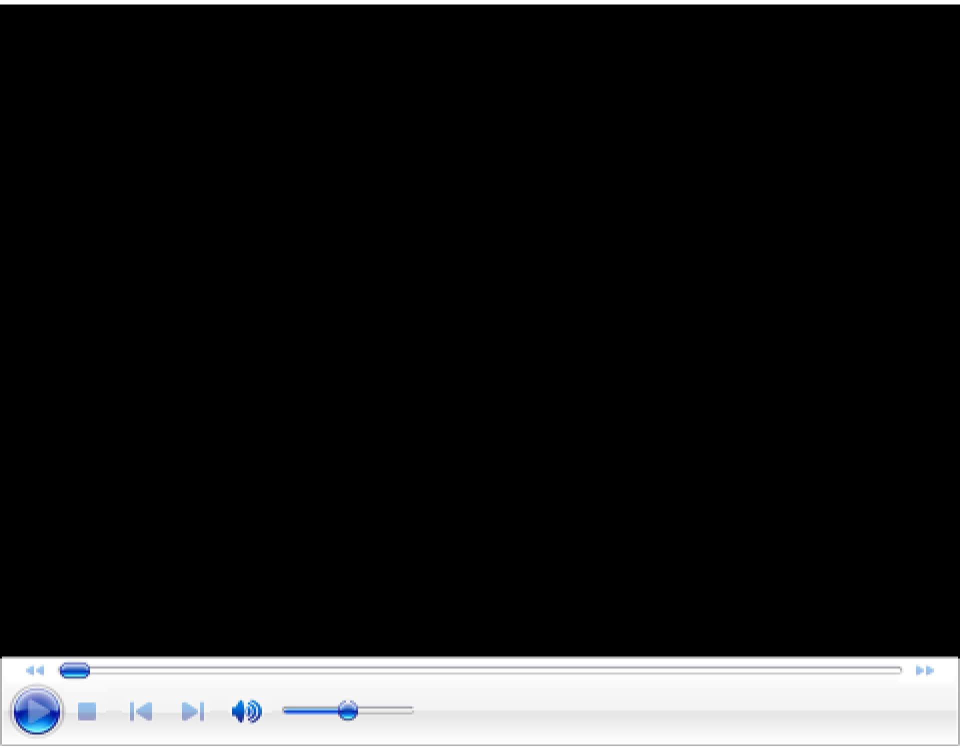
- External Video

UMI 2958 GT-CNRS

Main contributors: Shane Griffith, Paul Drews

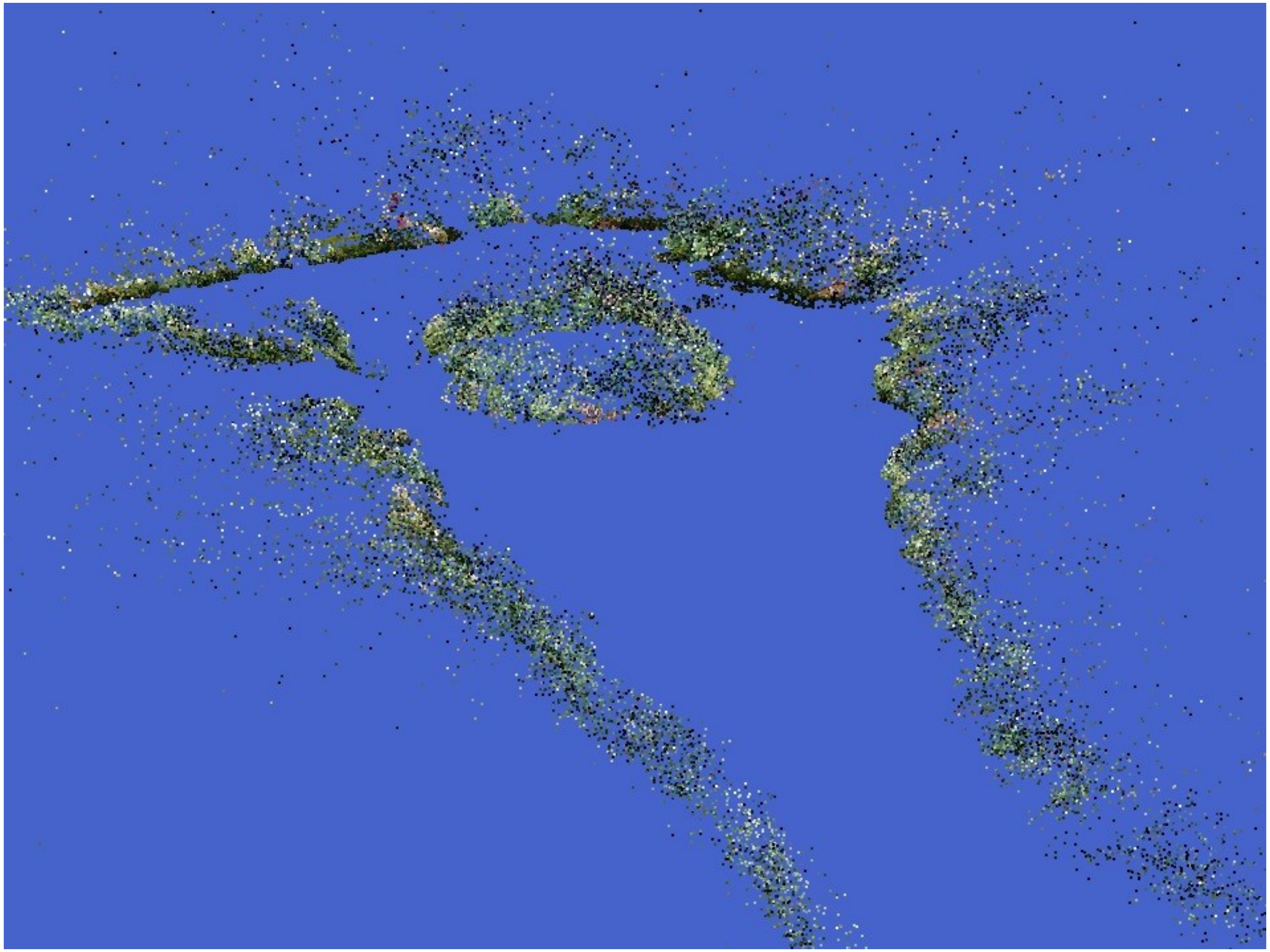
REGULAR INSPECTION OF NATURAL ENVIRONMENTS

Autonomous Lakeshore Monitoring



Dataset Overview

The Geometry of a Lakeshore



Natural Scene Variation Across Surveys

Natural Scene Variation Across Surveys

Vehicles

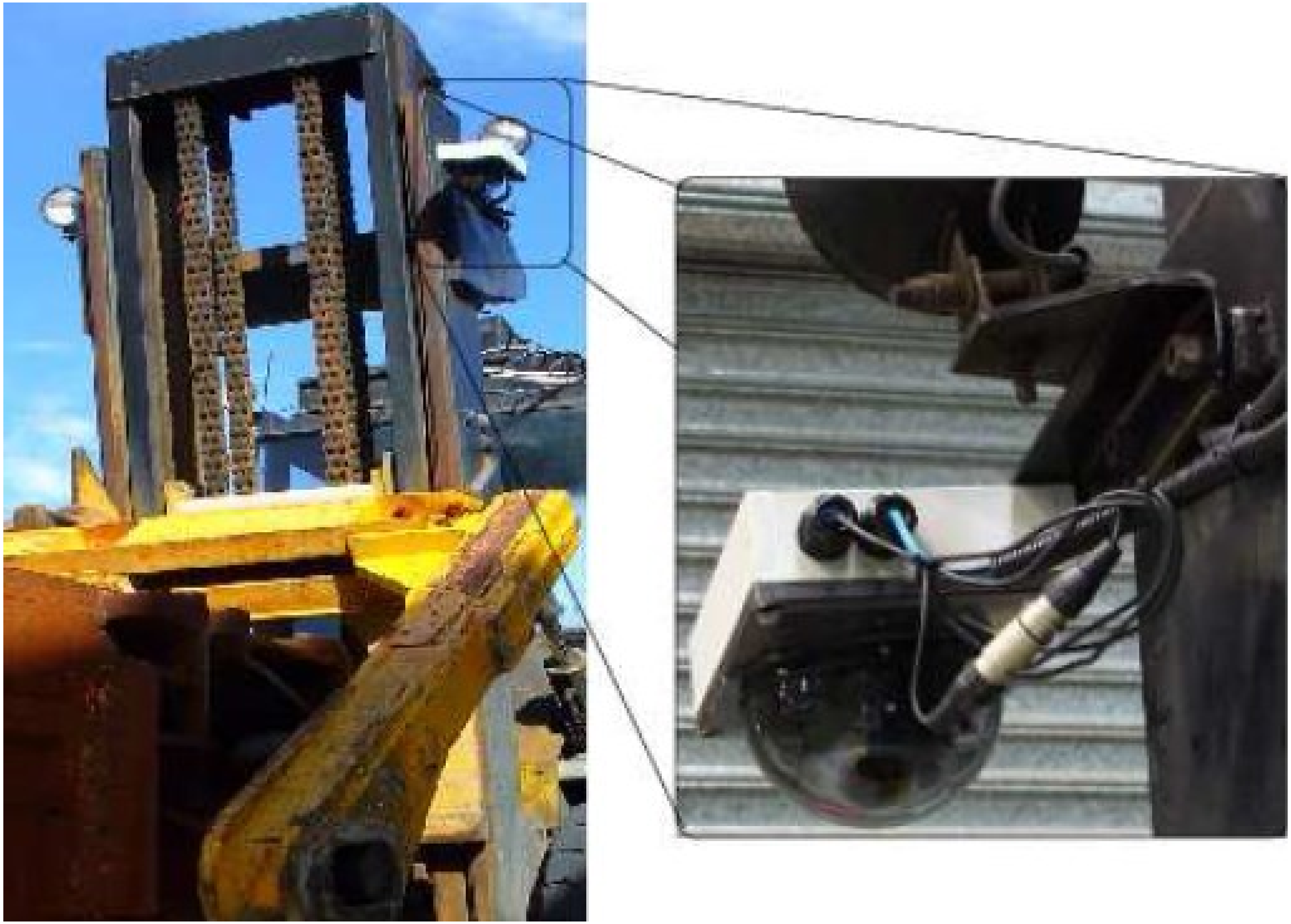
V-Charge

V-Charge

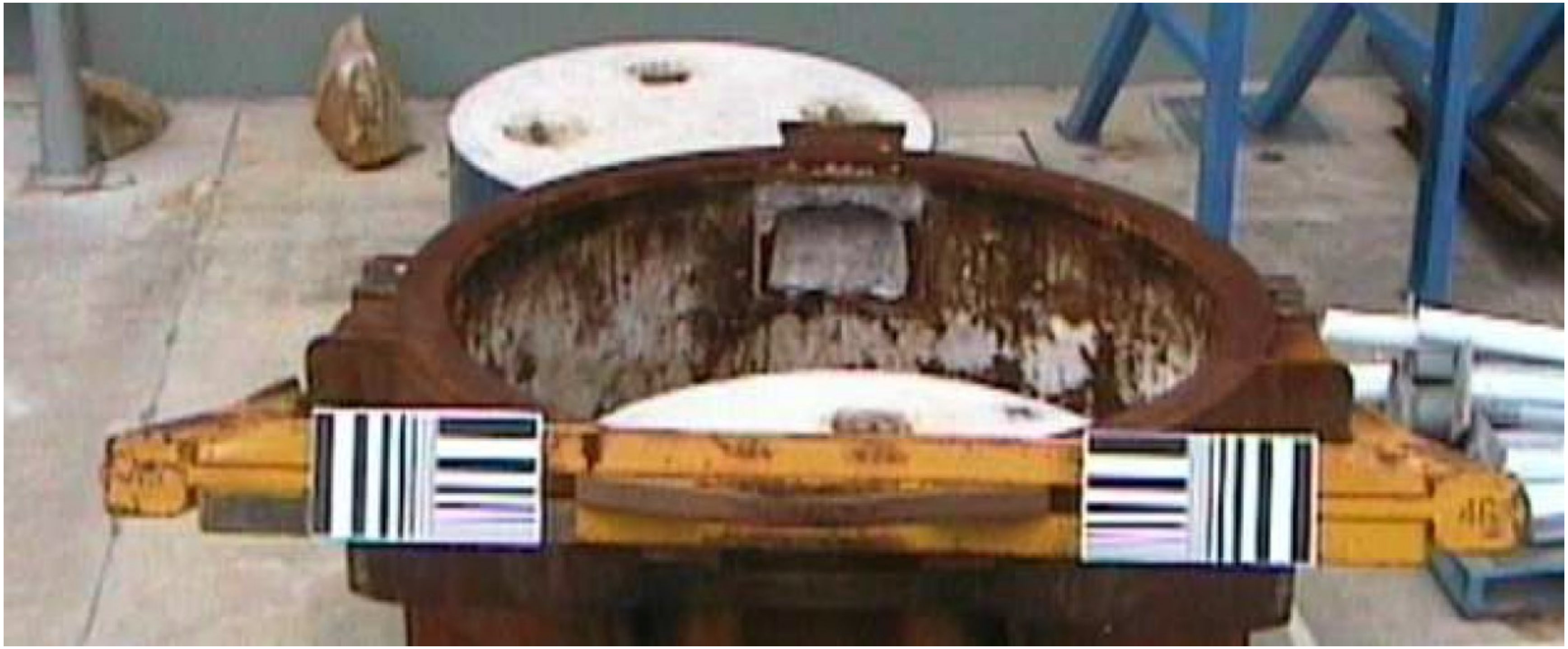
- External Video

Hot Metal Carrier

Hot Metal Carrier



Hot Metal Carrier



Hot Metal Carrier

Why study Computer Vision?

- Images (and movies) have become ubiquitous in both production and consumption.
- Therefore application to manipulate images (movies) are becoming core.
- As are systems that *extract information* from imagery
  - Surveillance
  - Building 3D representations
  - Motion capture assisted
- But most of all…

It is a really deep and cool set of problems!

What is computer vision?

Terminator 2

Steve Seitz

Every picture tells a story

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

Steve Seitz

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

Steve Seitz

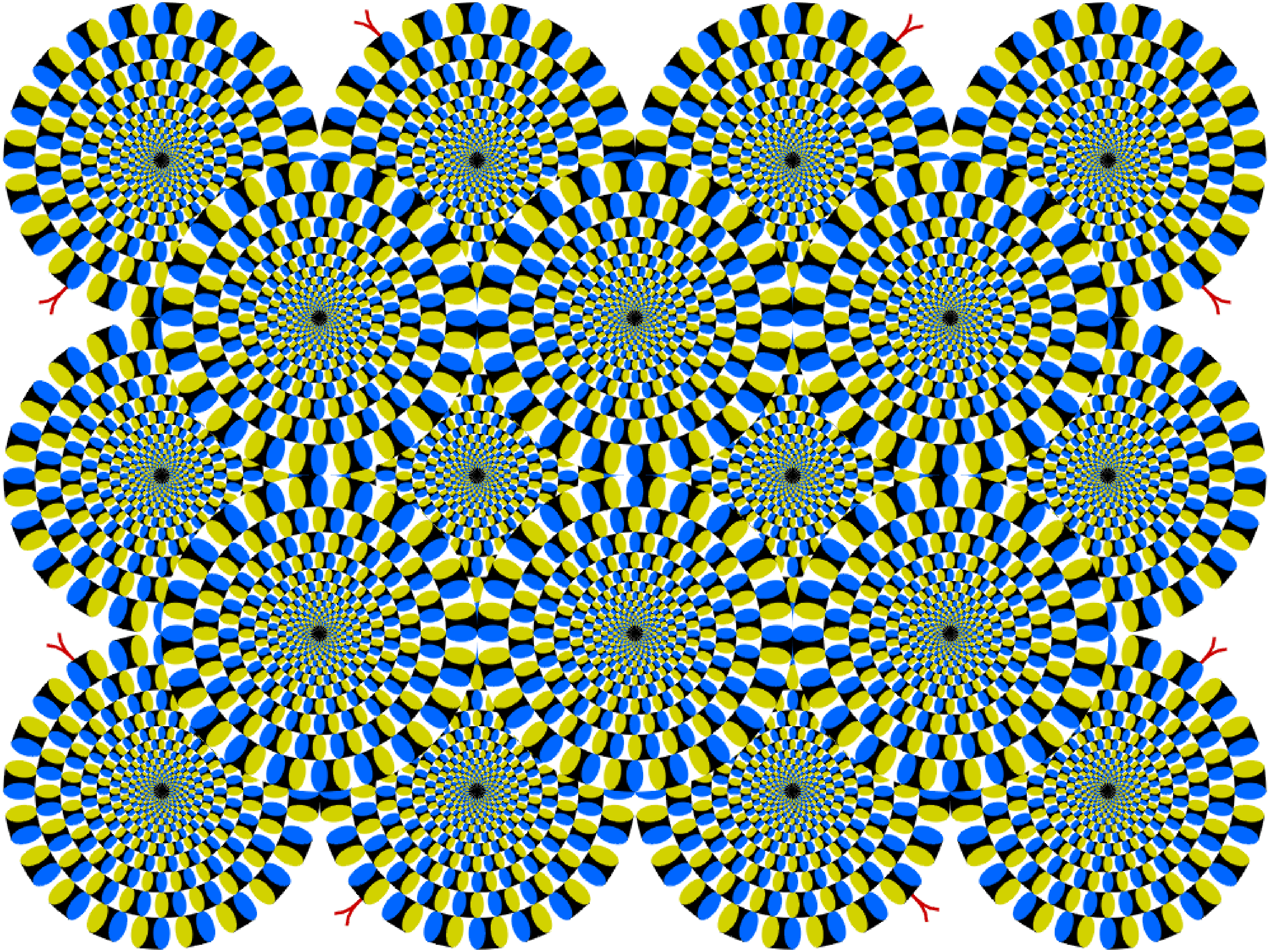
Human perception has its shortcomings…

[Sinha and Poggio, *Nature*, 1996]

Steve Seitz

Human perception has its shortcomings…

Sinha and Poggio, *Nature*, 1996



Steve Seitz

Current state of the art

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

Earth viewers (3D modeling)

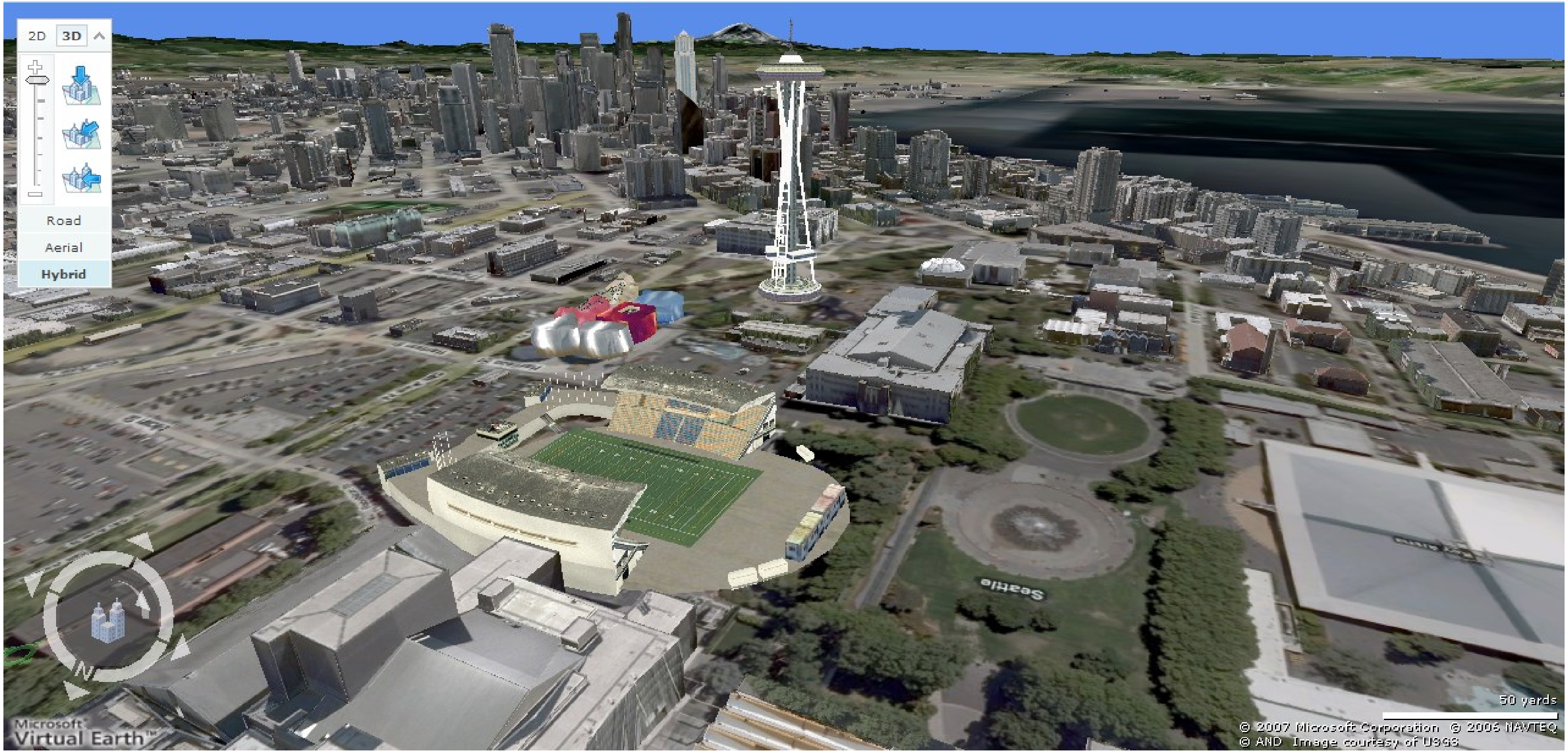
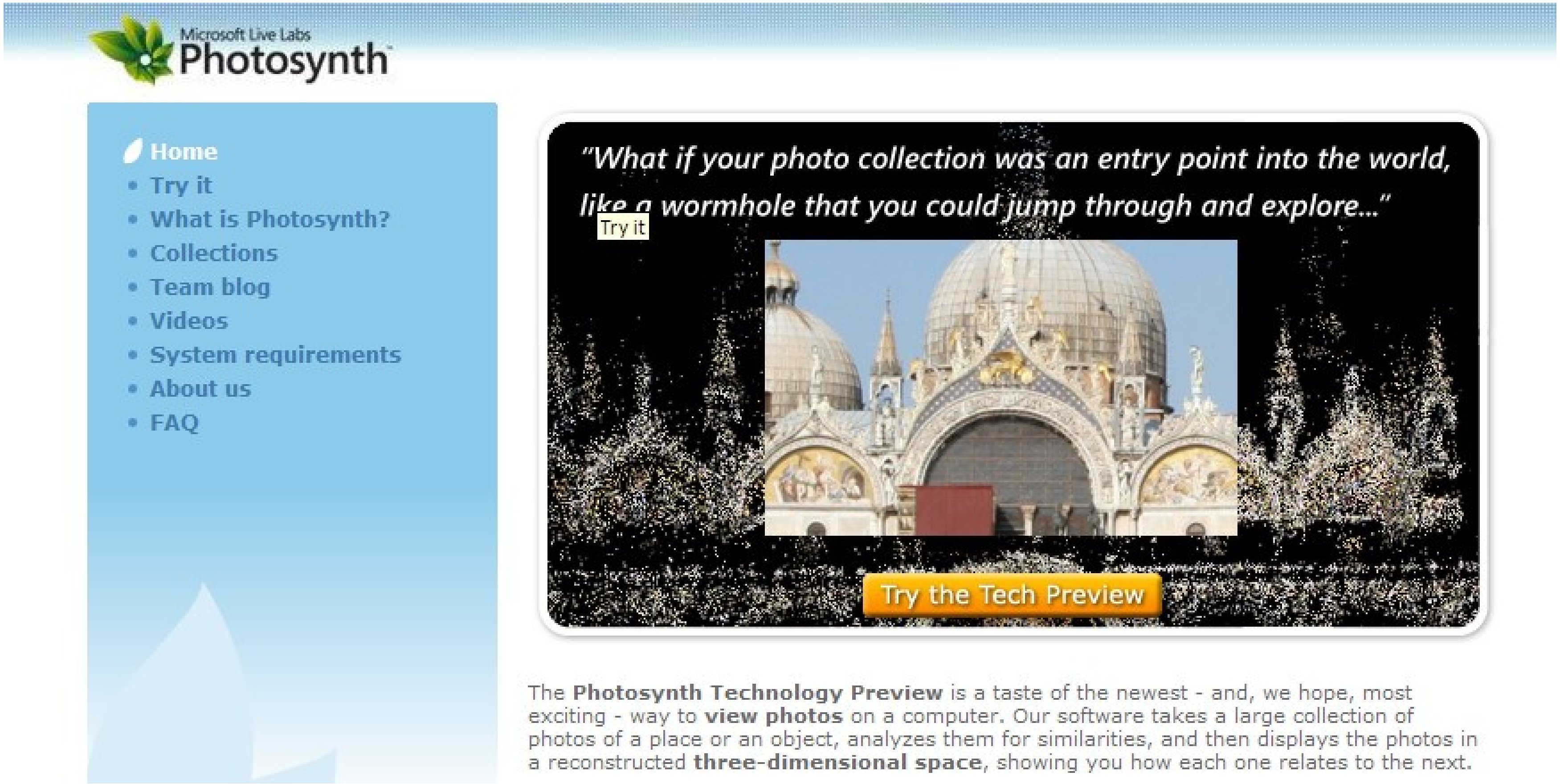


Image from Microsoft's Virtual Earth
(see also: Google Earth)

Steve Seitz



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

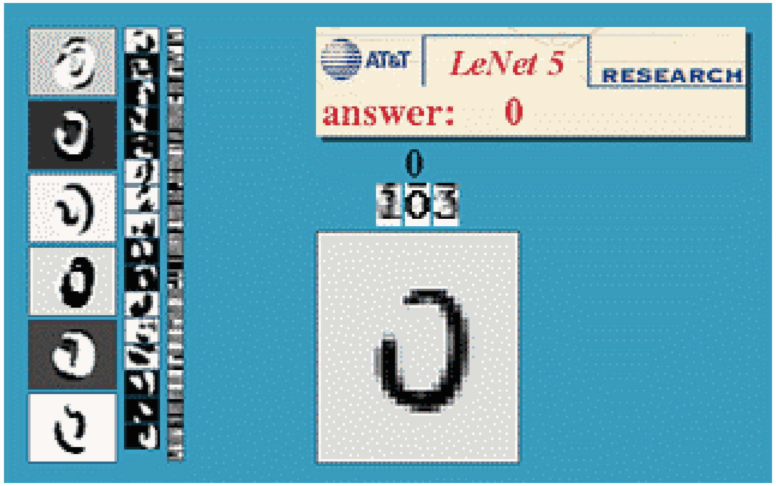
Based on Photo Tourism technology developed here in UW.
by Noah Snavely, Steve Seitz, and Rick Szeliski

Steve Seitz

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

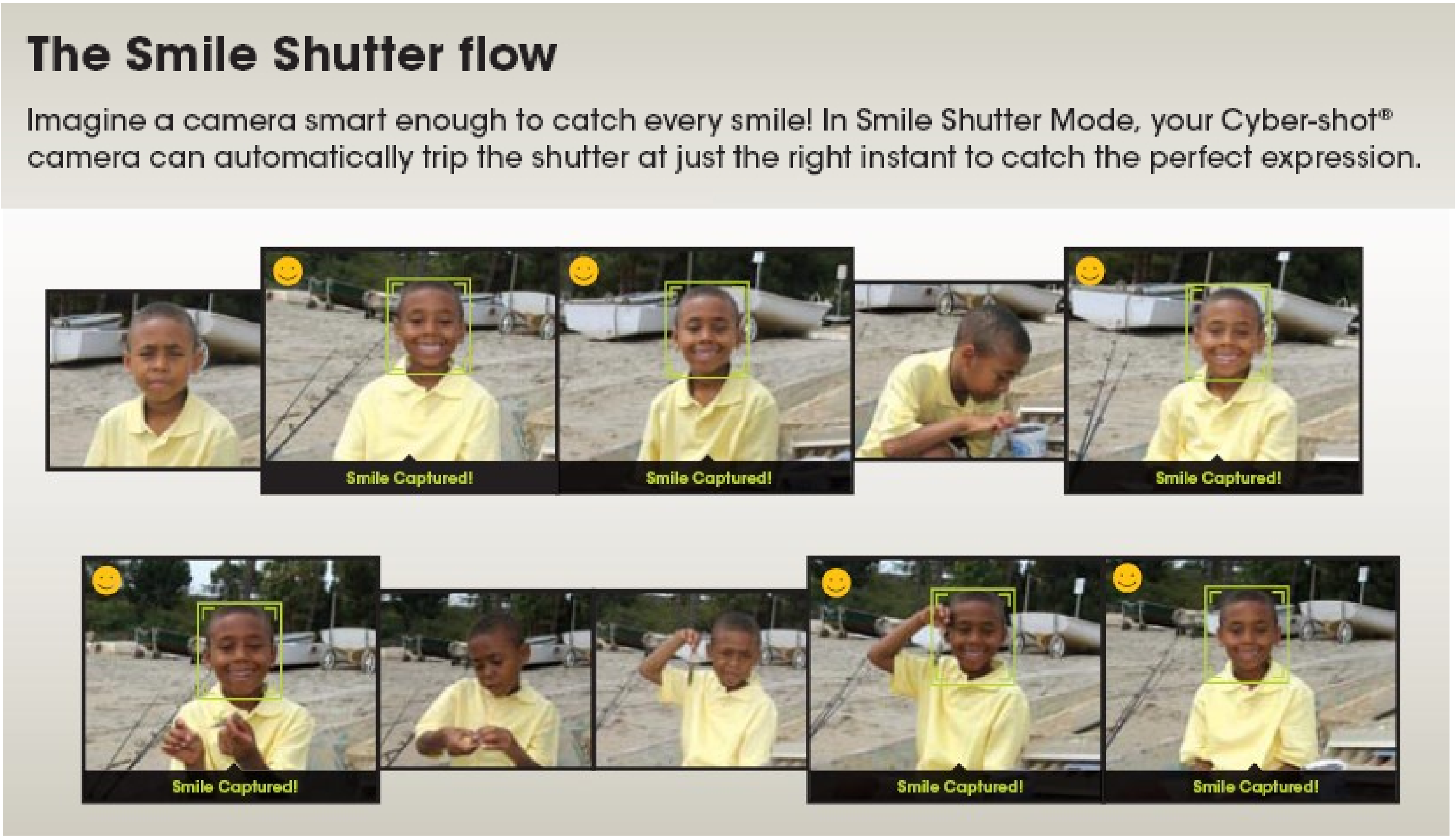
Steve Seitz

Face detection

- Most digital cameras now detect faces
  - Canon, Sony, Fuji, …

Steve Seitz

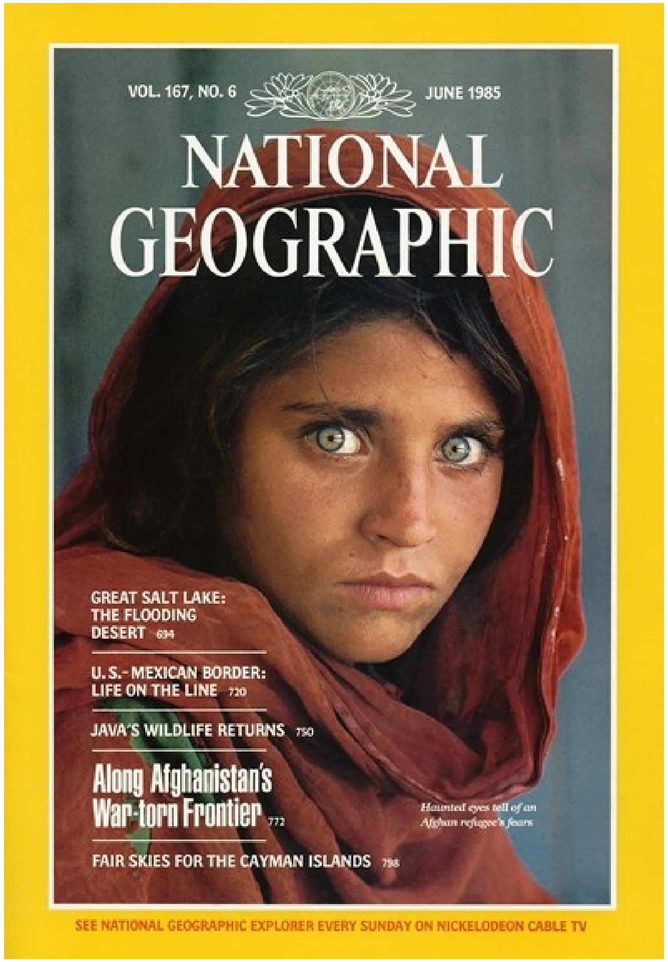
Smile detection?



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

Steve Seitz

Face recognition

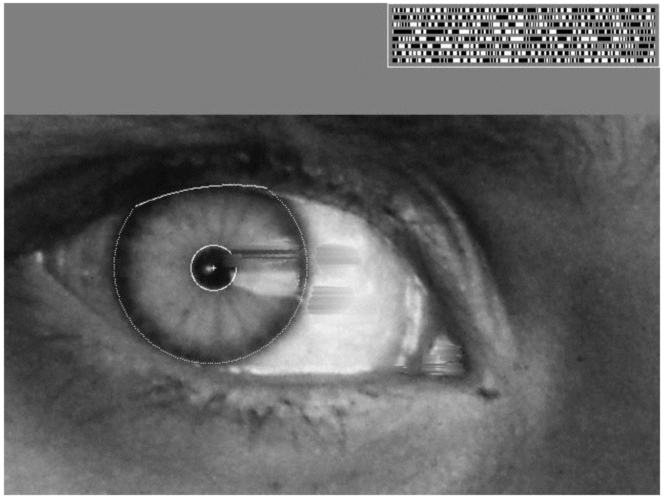


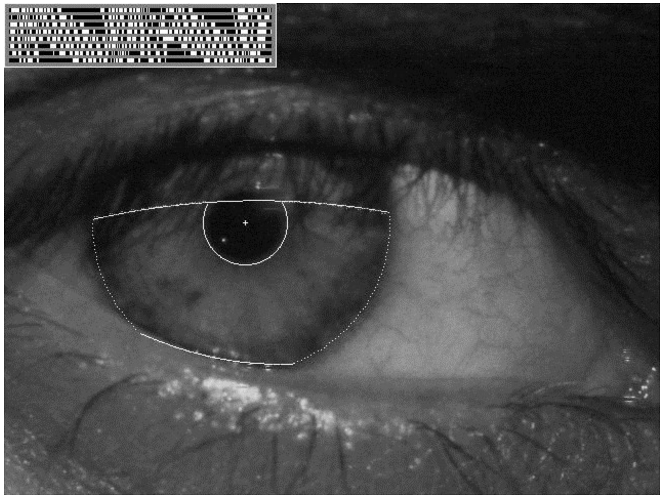
Who is she?

Steve Seitz

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/

Steve Seitz

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… "

Object recognition *(in mobile phones!)*

- This is becoming real:
  - Google Goggles
  - Lincoln Microsoft Research
  - [Point & Find](), [Nokia]()

Steve Seitz

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Steve Seitz

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic
Click here for web site

Steve Seitz

Sports

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

Steve Seitz

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.

Smart cars *are here!*

Nevada embraces the future, approves self-driving cars

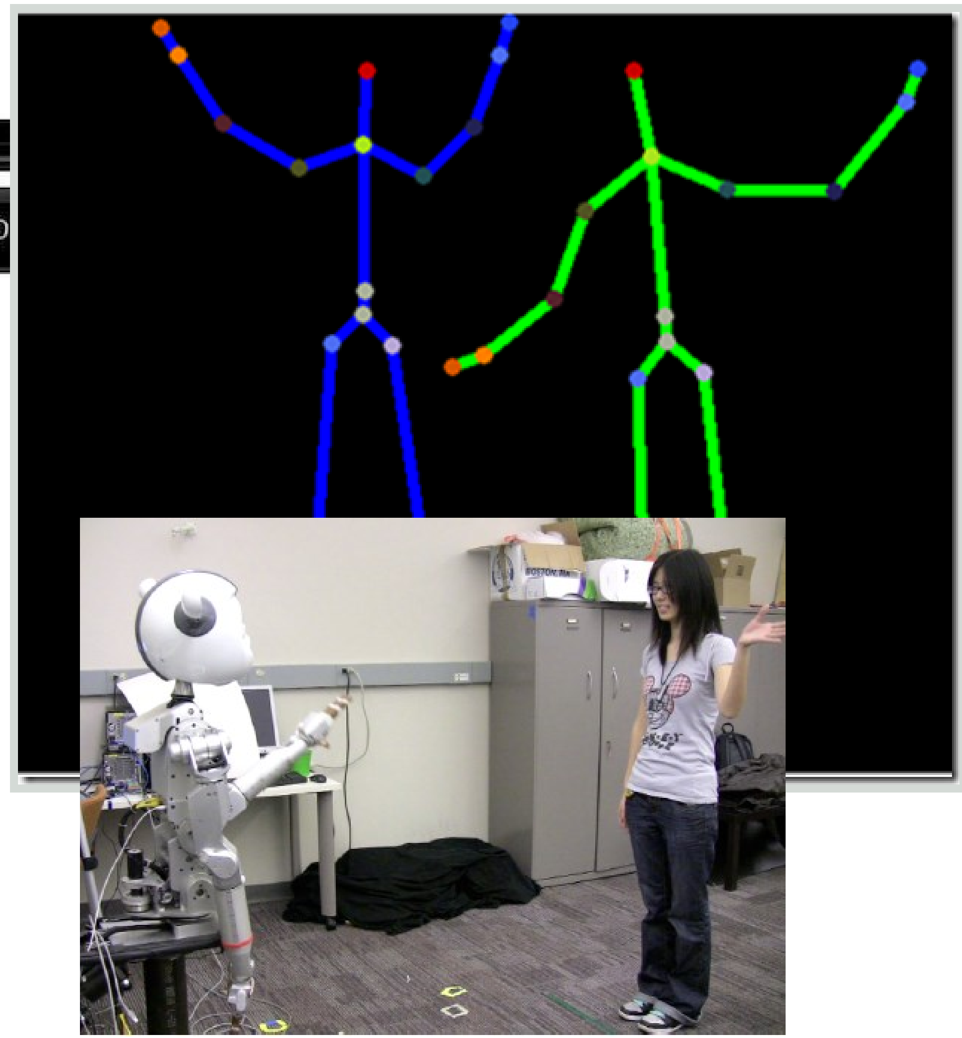
By Bill Howard on February 20, 2012 at 9:00 am | 1 Comment

Vision-based interaction (and games)

Nintendo Wii has camera-based IR tracking built in.

Steve Seitz

But the game changer:



Vision in space

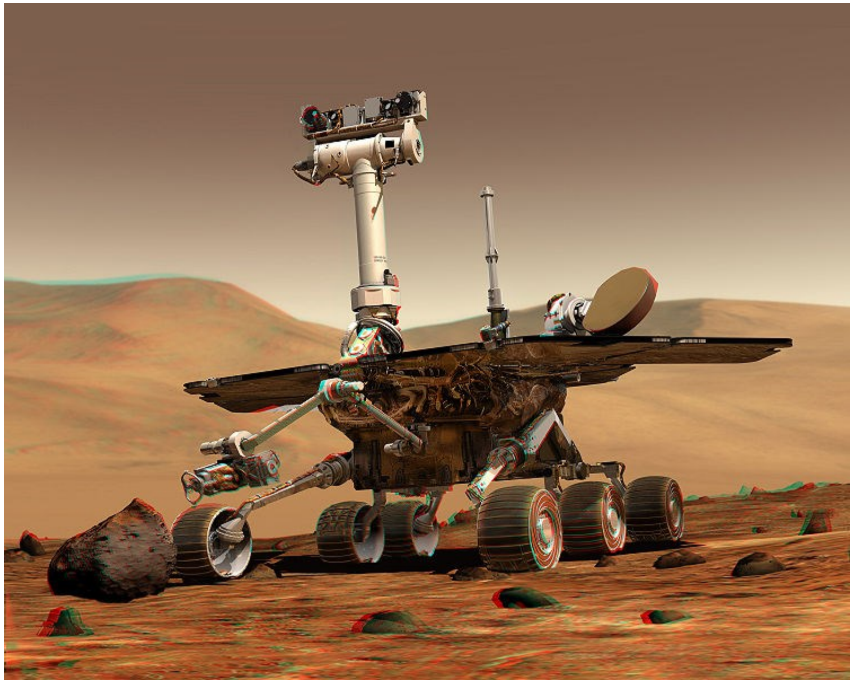
NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

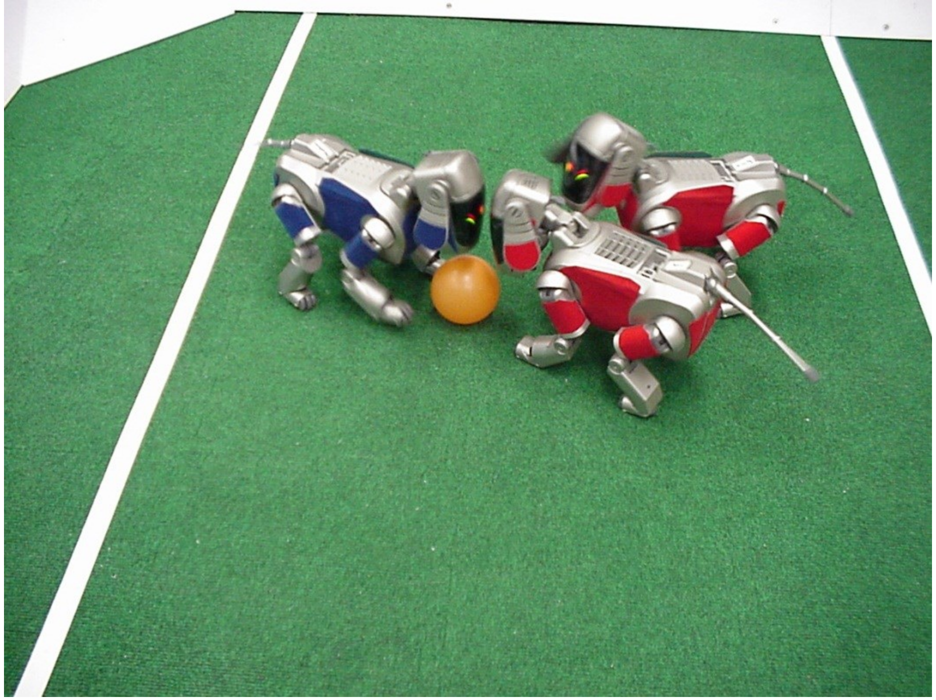
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Steve Seitz

Robotics



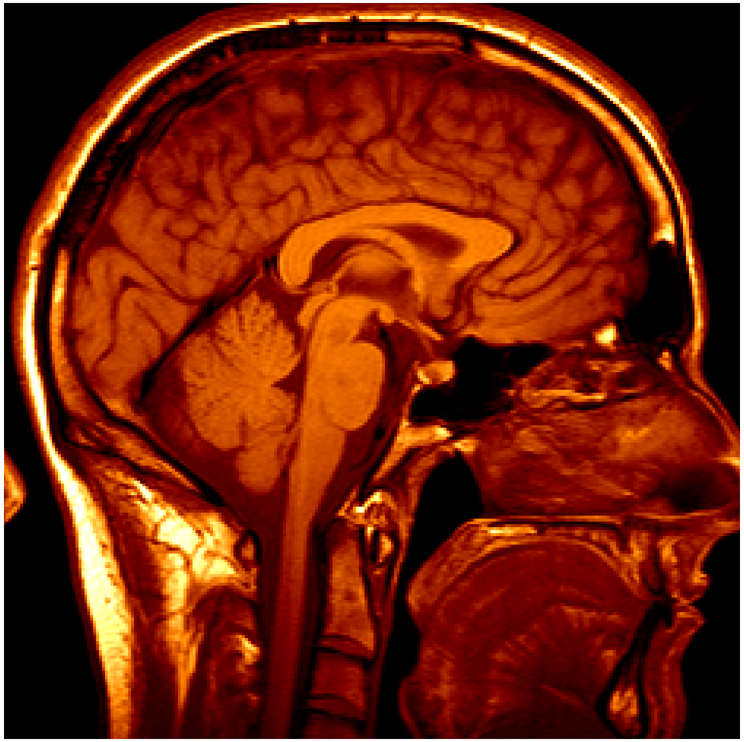
NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

Steve Seitz

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Steve Seitz

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

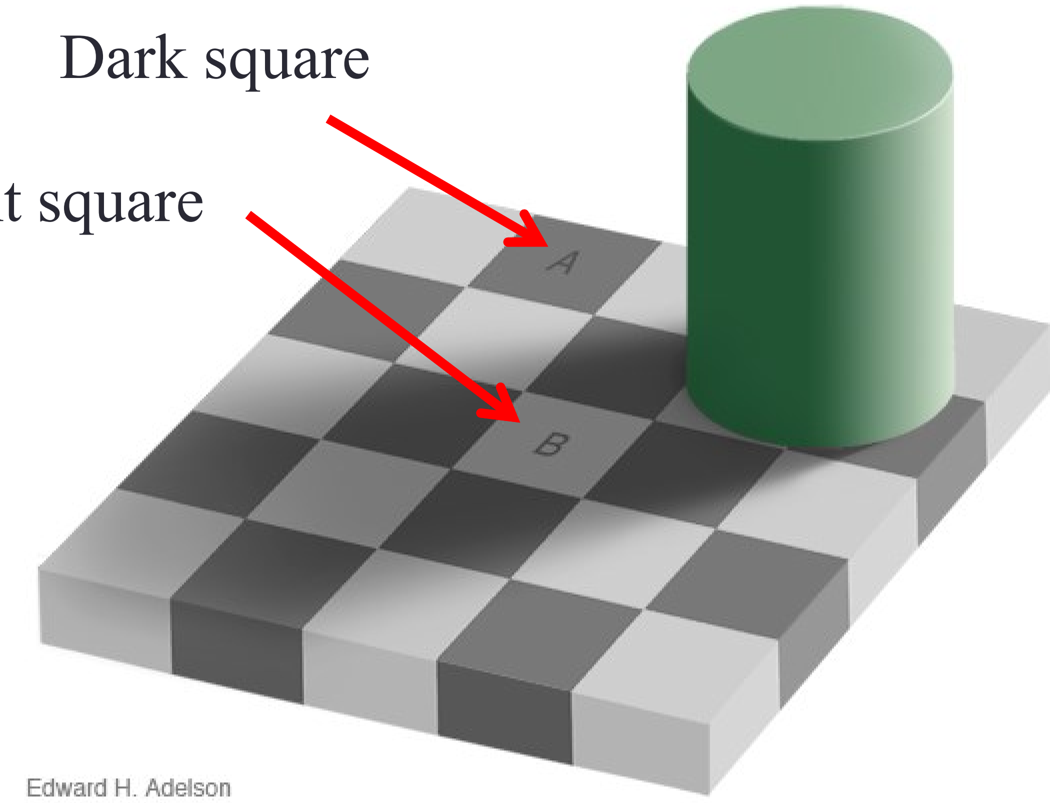
Steve Seitz

Why is this hard?

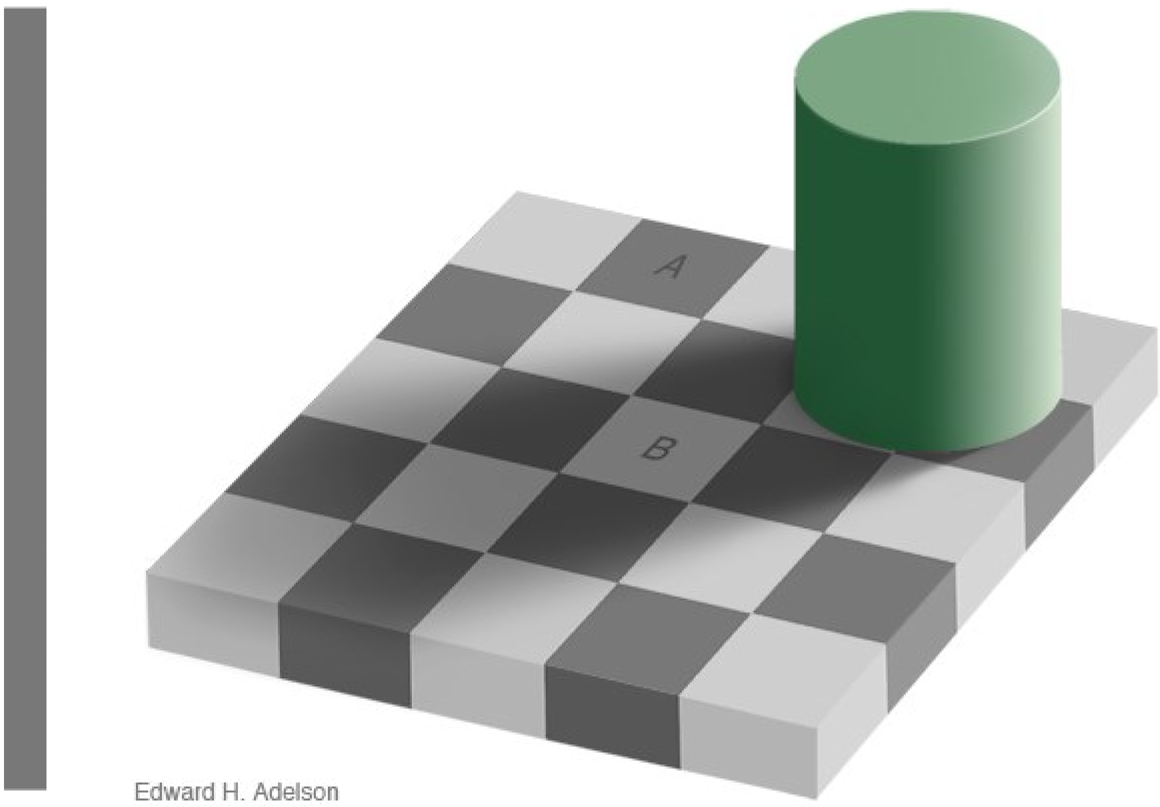
Simple scene right?

Dark square

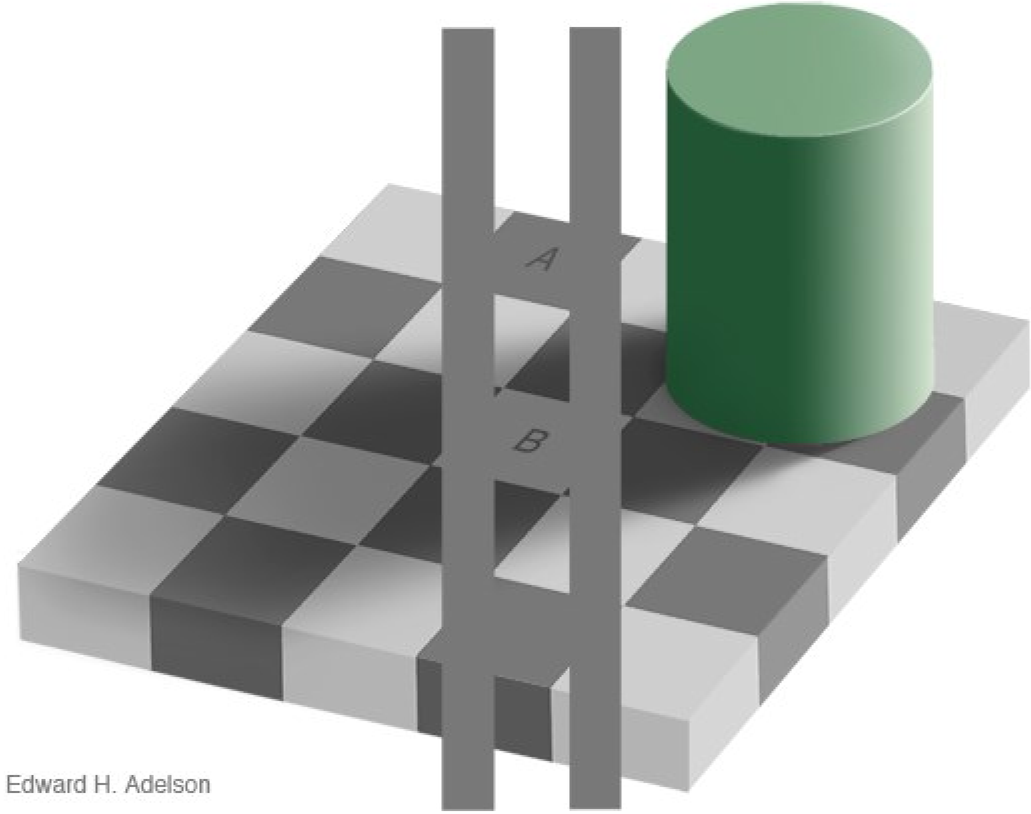
Light square



Really?



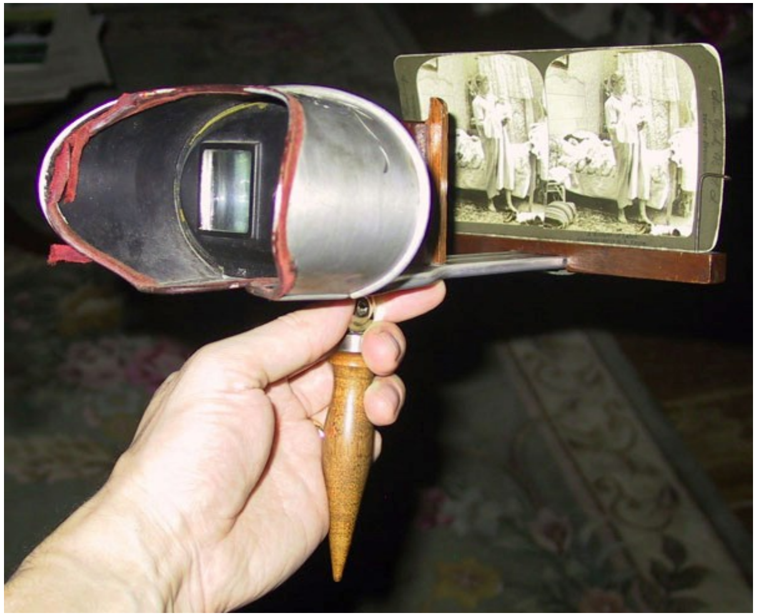
Really!



Vision is NOT Image Processing

- In the previous example, the two squares have exactly the same *measurement* of intensity.
- So, seeing is not the same as measuring properties in the image.
- Rather, "seeing" is building a *percept* of what is in the world based upon the measurements made by an imaging sensor.

Building models from change (1)



Michael Black

Going forward

- For coming lectures:
  - FP: Chapter 4 (all), 5.1, 5.2

- Get yourself Matlab (and/or Python/OpenCV)

- PS0 – yes it's easy but you need to do it and need to submit it in the format requested.