CIS 849: Autonomous Robot Vision

Instructor: Christopher Rasmussen

Course web page:

www.cis.udel.edu/~cer/arv

University of Delaware

Purpose of this Course

- To provide an introduction to the uses of visual sensing for mobile robotic tasks, and a survey of the mathematical and algorithmic problems that recur in its application

What are “Autonomous Robots”?

- Mobile machines with power, sensing, and computing on-board
- Environments
  - Land (on and under)
  - Water (ditto)
  - Air
  - Space
  - ???

What Can/Will Robots Do?

- Near-term: *What People Want*
  - Tool analogy
  - Never too far from human intervention, whether physically or via tele-operation
  - Narrow tasks, limited skills
  - "3-D": Dirty, Dangerous, and Dull jobs

What Can/Will Robots Do? Task Areas

- Industry
- Transportation & Surveillance
- Search & Science
- Service

What Might Robots Do?

- Long-term: *What They Want*
  - "Mechanical animal" analogy may become appropriate
  - Science fiction paradigm
    - On their own
    - Self-directed generalists

Industry

- Ground coverage
  - Harvesting, lawn-mowing (CMU)
  - Snow removal
  - Mine detection
- Inspection of other topologies
  - MAKRO (Fraunhofer): Sewer pipes
  - CIMP (CMU): Aircraft skin

MAKRO

CIMP

UNIVERSITY OF DELAWARE

CMU Demeter

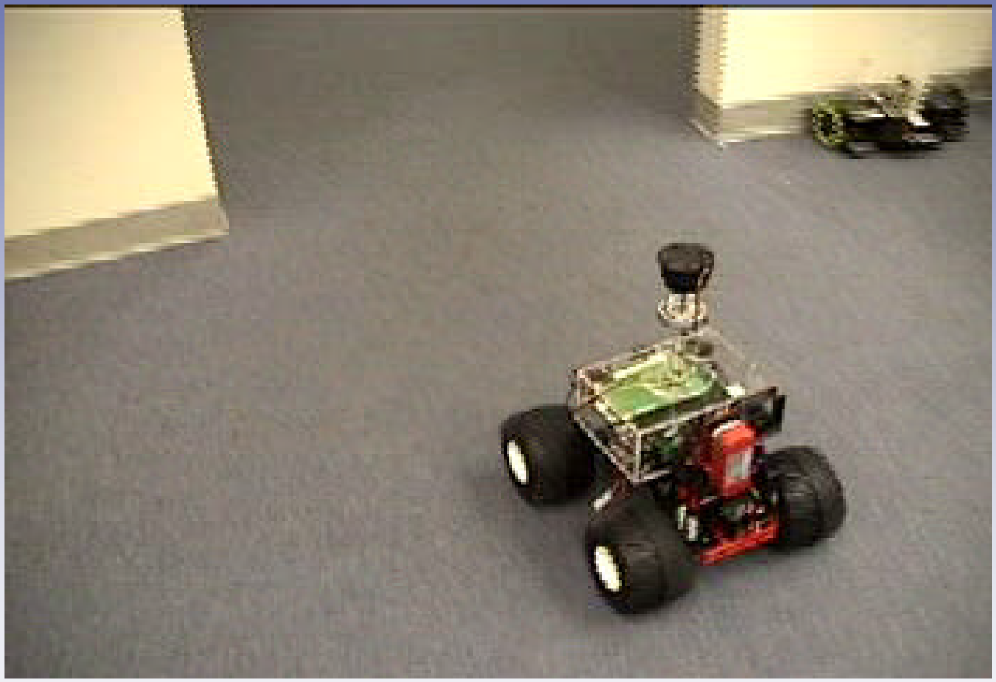
Transportation & Surveillance: Ground

- Indoors
  - Clodbusters (Penn)
  - Many others
- Highways, city streets
  - VaMoRs/VaMP (UBM)
  - NAVLAB/RALPH (CMU)
  - StereoDrive (Berkeley)
- Off-road
  - Ranger (CMU)
  - Demo III (NIST, *et al.*)

VaMoRs

Ranger

Penn Clodbuster



Obstacle avoidance with omnidirectional camera

UBM VaMoRs

Detecting a ditch with stereo, then stopping

Transportation & Surveillance: Air

- Fixed wing (UBM, Florida)
- Helicopters (CMU, Berkeley, USC, Linkoping)
- Blimp (IST, Penn)

UBM autonomous landing aircraft

Florida MAV

USC Avatar

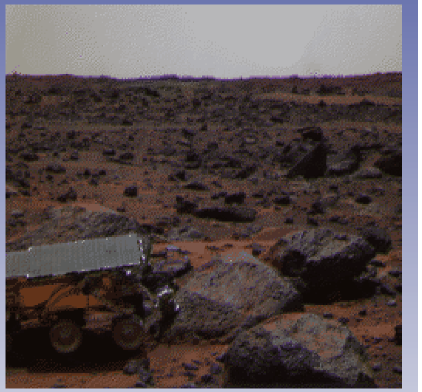
Landing on target (mostly)

University of Delaware

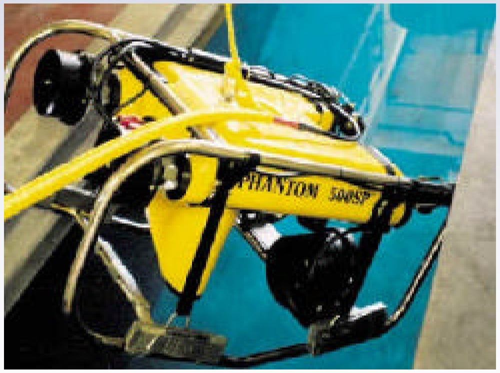
Search & Science

- Urban Search & Rescue
  - Debris, stairs
  - Combination of autonomy & tele-operation
- Hazardous data collection
  - Dante II (CMU)
  - Sojourner (NASA)
  - Narval (IST)

Dante II



Sojourner



Narval

UNIVERSITY OF DELAWARE

USF at the WTC

courtesy of CRASAR

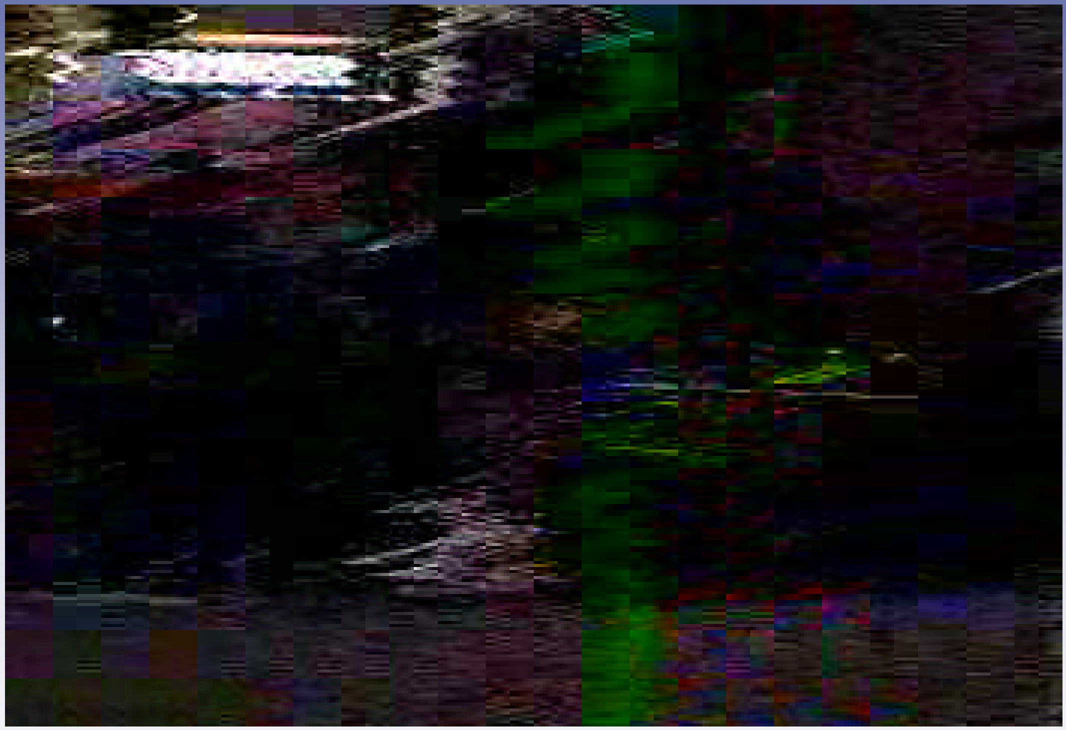
Urbot & Packbot reconnoiter surrounding structures

Service

- Grace (CMU, Swarthmore, *et al.*): "Attended" AI conference
  - Register, interact with other participants
  - Navigate halls, ride elevator
- Guides
  - Polly (MIT): AI lab
  - Minerva (CMU): Museum
- Personal assistants
  - Nursebot (CMU): Eldercare
  - Robotic wheelchairs

Grace

CMU Minerva



In the Smithsonian

What Skills Do Robots Need?

- **Identification**: *What/who is that?*
  - Object detection, recognition
- **Movement**: *How do I move safely?*
  - Obstacle avoidance, homing
- **Manipulation**: *How do I change that?*
  - Interacting with objects/environment
- **Navigation**: *Where am I?*
  - Mapping, localization

University of Delaware

Why Vision?

- Pluses
  - Rich stream of complex information about the environment
  - Primary human sense
  - Good cameras are fairly cheap
  - Passive ? stealthy
- Minuses
  - Line of sight only
  - Passive ? Dependent on ambient illumination

Aren't There Other Important Senses?

- Yes—
  - The rest of the human "big five" (hearing, touch, taste, smell)
  - Temperature, acceleration, GPS, etc.
  - Active sensing: Sonar, ladar, radar
- But...
  - Mathematically, many other sensing problems have close visual correlates

The Vision Problem

How to infer salient properties of 3-D world from time-varying 2-D image projection

University of Delaware

Computer Vision Outline

- Image formation
- Image processing
- Motion & Estimation
- Classification

Outline: Image Formation

- 3-D geometry
- Physics of light
- Camera properties
  - Focal length
  - Distortion
- Sampling issues
  - Spatial
  - Temporal

Outline: Image Processing

- Filtering
  - Edge
  - Color
  - Shape
  - Texture
- Feature detection
- Pattern comparison

University of Delaware

Outline: Motion & Estimation

- Computing temporal image change
  - Magnitude
  - Direction
- Fitting parameters to data
  - Static
  - Dynamic (e.g., tracking)
- Applications
  - Motion Compensation
  - Structure from Motion

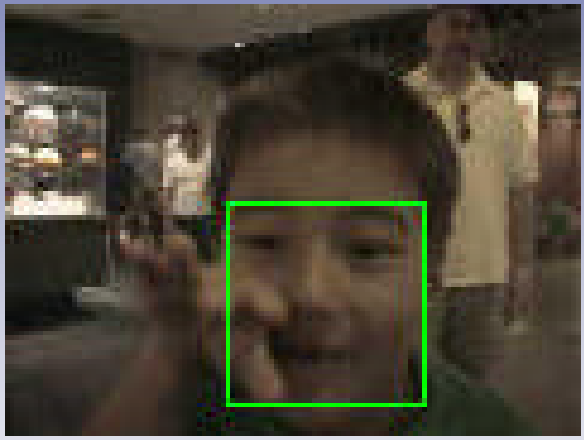
Outline: Classification

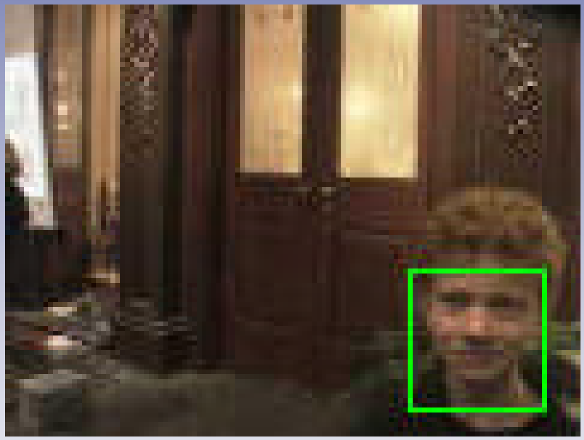
- Categorization
  - Assignment to known groups
- Clustering
  - Inference of group existence from data
  - Special case: Segmentation

Visual Skills: Identification

- Recognizing face/body/structure: *Who/what do I see?*
  - Use shape, color, pattern, other static attributes to distinguish from background, other hypotheses
- Gesture/activity: *What is it doing?*
  - From low-level motion detection & tracking to categorizing high-level temporal patterns
- Feedback between static and dynamic

Minerva Face Detection





Finding people to interact with

UNIVERSITY OF DELAWARE

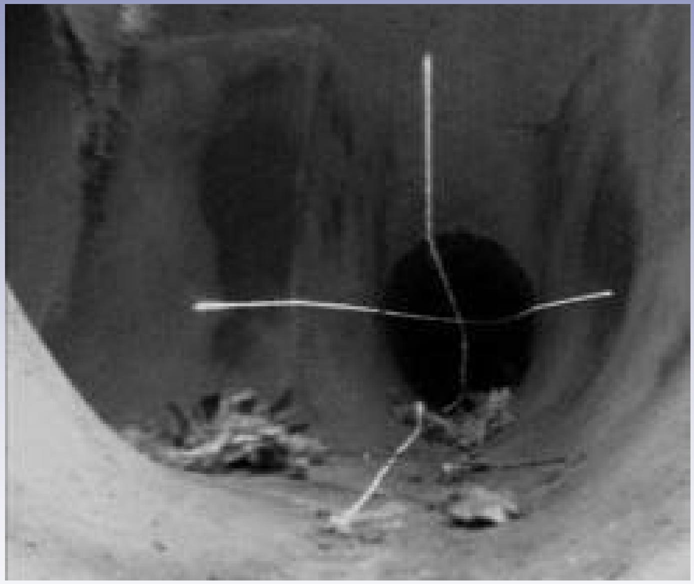
Penn MARS project

Blimp, Clodbusters

Airborne, color-based tracking

Visual Skills: Movement

- Steering, foot placement or landing spot for entire vehicle



MAKRO sewer shape pattern

Demeter region boundary detection

Florida Micro Air Vehicle (MAV)

Horizon detection for
self-stabilization

UBM Lane & vehicle tracking (with radar)

Visual Skills: Manipulation

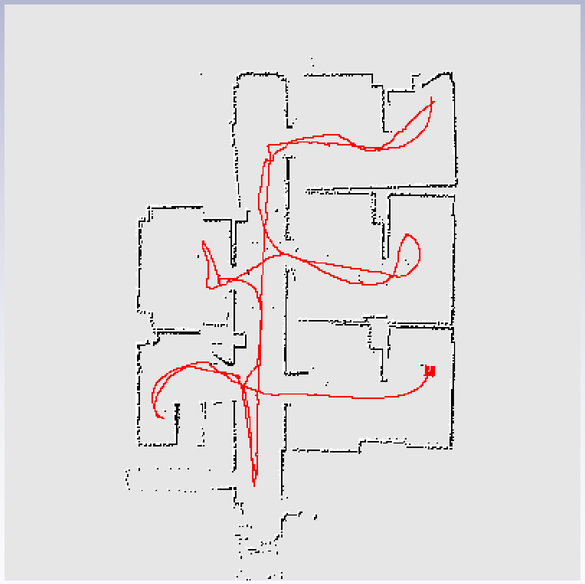
- Moving other things
  - Grasping: Door opener (KTH)
  - Pushing, digging, cranes

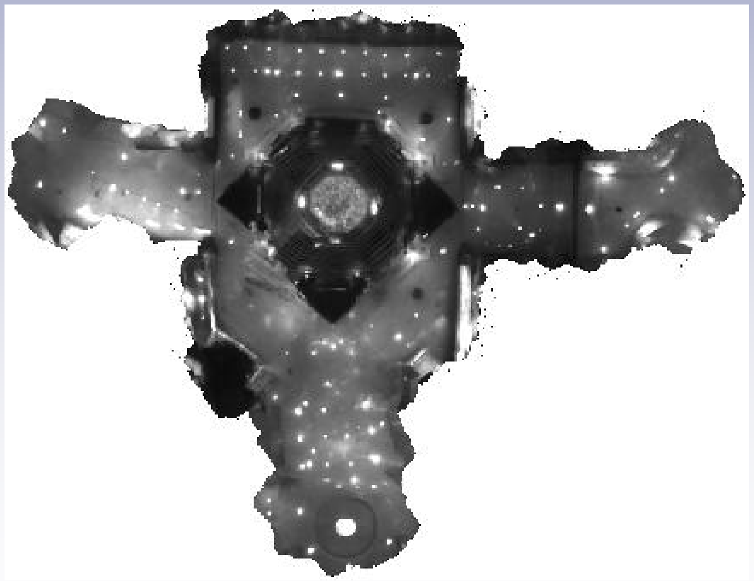
KTH robot & typical handle

Clodbusters push a box cooperatively

Visual Skills: Navigation

- Building a map [show “3D.avi”]
- Localization/place recognition
  - *Where are you in the map?*





Laser-based wall map (CMU)   Minerva’s ceiling map

Course Prerequisites

- Strong background in/comfort with:
  - Linear algebra
  - Multi-variable calculus
  - Statistics, probability
- Ability to program in:
  - C/C++, Matlab, or equivalent

Course Details

- First 1/3 of classes: Computer vision review by professor
- Last 2/3 of classes: Paper presentations, discussions led by students
- One major programming project
- Grading
  - 10%: Two small programming assignments
  - 30%: Two oral paper presentations + write-ups
  - 10%: Class participation
  - 50%: Project

Readings

- All readings will be available online as PDF files
- Textbook: Selected chapters from pre-publication draft of *Computer Vision: A Modern Approach*, by D. Forsyth and J. Ponce
- Web page has other online vision resources
- Papers: Recent conference and journal articles spanning a range of robot types, tasks, and visual algorithms

Presentations

- Each student will submit short written analyses of two papers, get feedback, then present them orally
- Non-presenting students should read papers ahead of time and have some questions prepared. I will have questions, too :)

Project

- Opportunity to implement, test, or extend a robot-related visual algorithm
- Project proposal due in October; discuss with me beforehand
- Data
  - I will provide "canned" data, or gather your own
  - We will have a small wheeled robot to use for algorithms requiring live feedback
- Due Wednesday, November 27 (just before Thanksgiving break)

More questions?

Everything should be on the web page:

www.cis.udel.edu/~cer/arv

or e-mail me at

cer@cis.udel.edu