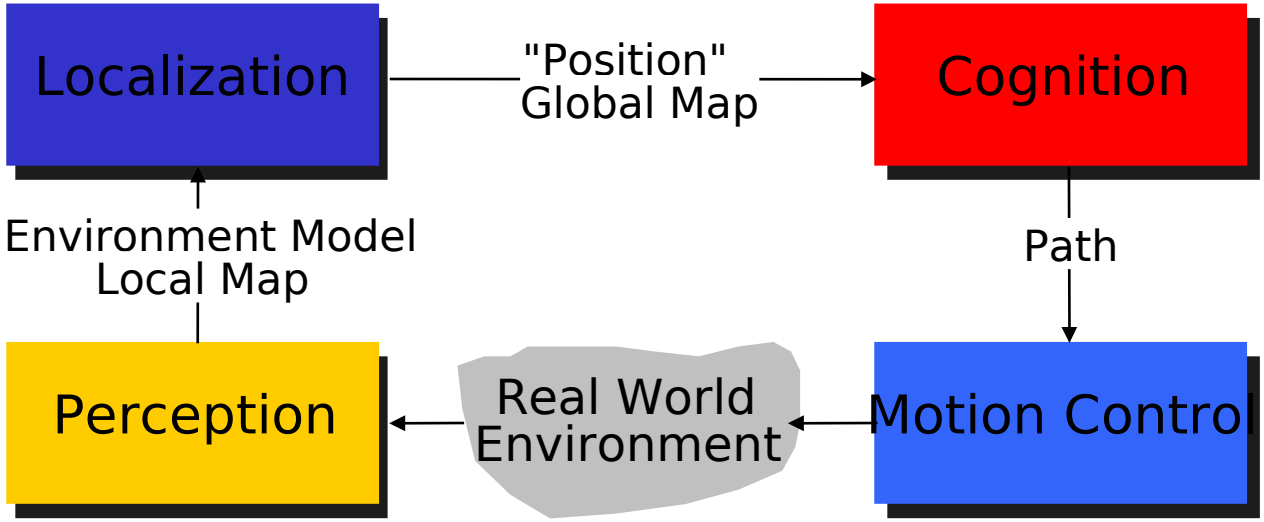
Perception

CS485
Autonomous Robotics
Amarda Shehu
Fall 2013

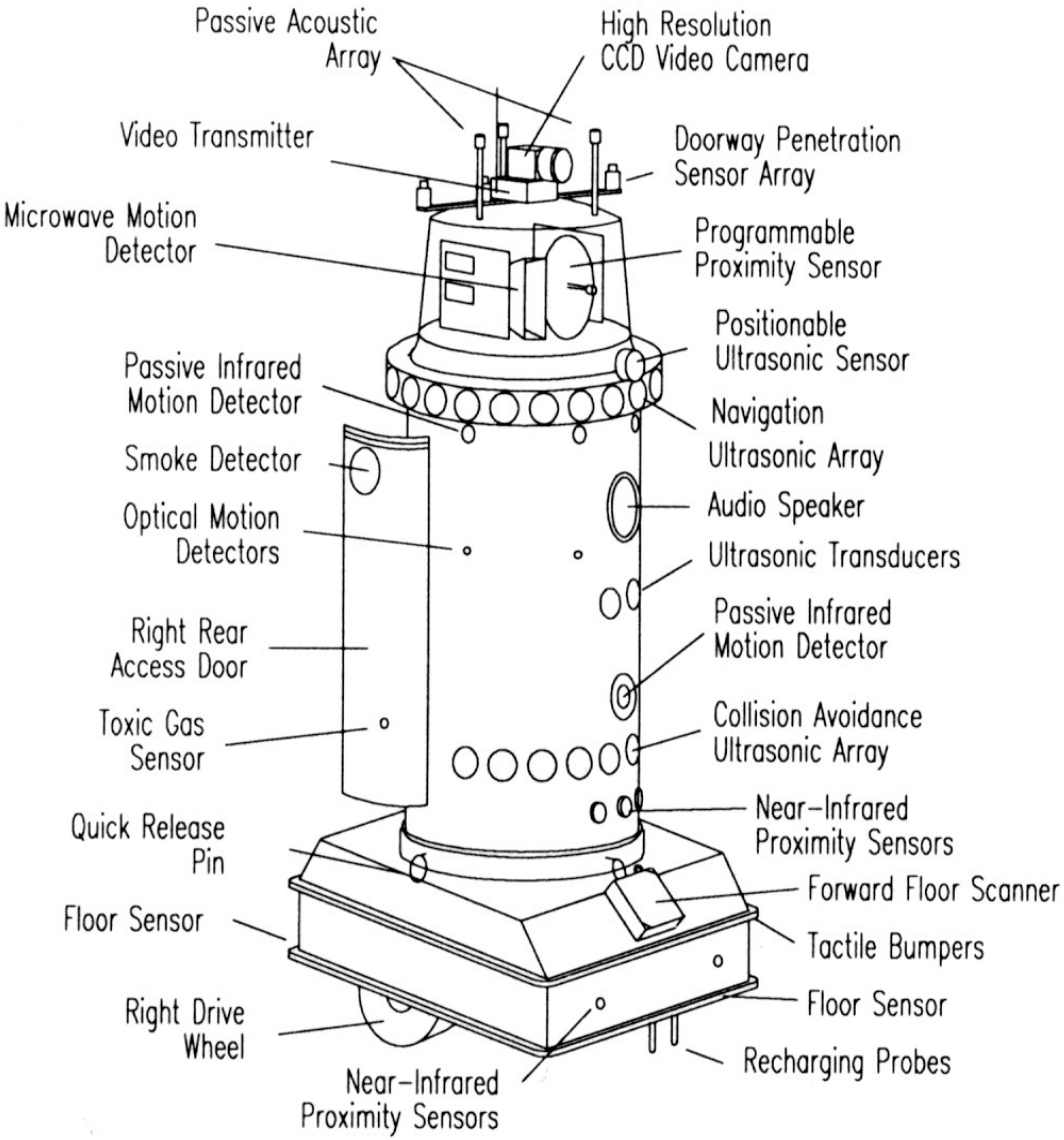
Notes modified from Jana Kosecka, George Mason University

Perception

- Sensors
- Uncertainty
- Features
- Introduction Chapter 4 [Nourbaksh & Siegwart]
- Introductory slides (courtesy [Nourbaksh & Siegwart])



Example Robart II, H.R. Everett



BibaBot, BlueBotics SA, Switzerland

Omnidirectional Camera

Pan-Tilt Camera

IMU
nertial Measurement Unit

Sonar Sensors

Emergency Stop Button

Laser Range Scanner

Wheel Encoders

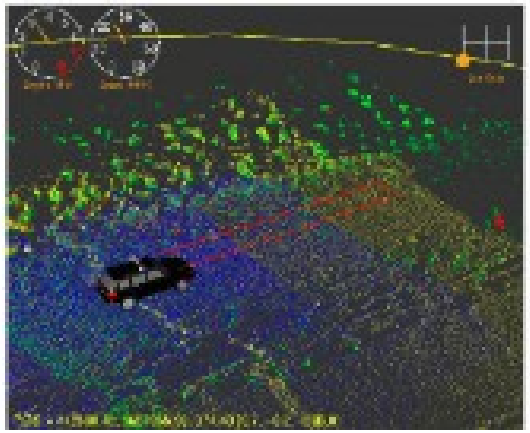
Bumper

Robotic Navigation

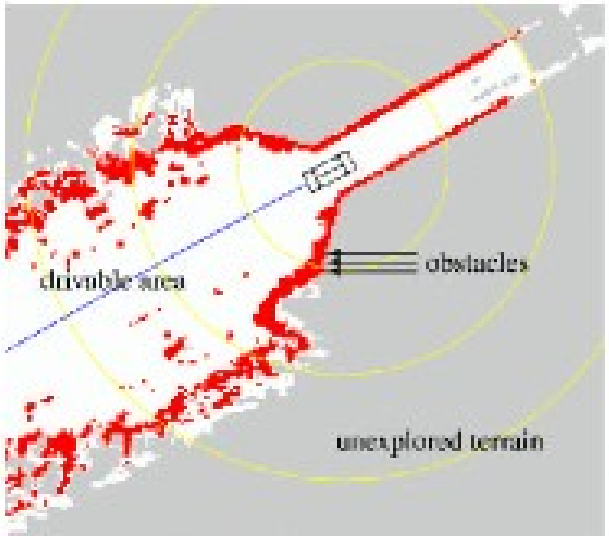
- Stanford Stanley Grand Challenge
- Outdoors unstructured env., single vehicle

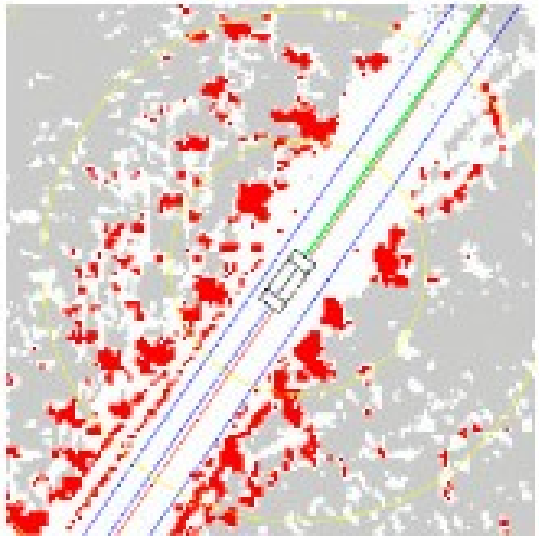
- Urban Challenge
- Outdoors structured env., mixed traffic, traffic rules

- Terrain mapping using lasers



- Determining obstacle course





Classification of Sensors

- Proprioceptive sensors
  - measure values internally to the system (robot),
  - e.g. motor speed, wheel load, heading of the robot, battery status
- Exteroceptive sensors
  - information from the robots environment
  - distances to objects, intensity of the ambient light, unique features.
- Passive sensors
  - energy coming for the environment
- Active sensors
  - emit their proper energy and measure the reaction
  - better performance, but some influence on environment

Role of Perception in Robotics

- Where am I relative to the world?
  - sensors: vision, stereo, range sensors, acoustics
  - problems: scene modeling/classification/recognition
  - integration: localization/mapping algorithms (e.g. SLAM)

- What is around me?
  - sensors: vision, stereo, range sensors, acoustics, sounds, smell
  - problems: object recognition, structure from x, qualitative modeling
  - integration: collision avoidance/navigation, learning

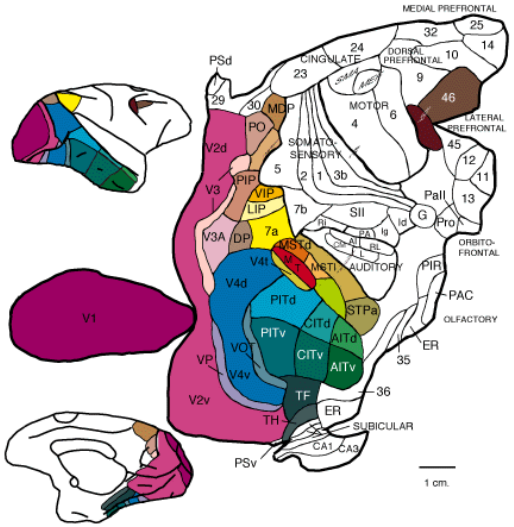
Role of Perception in Robotics

- How can I safely interact with environment (including people!)?
  - sensors: vision, range, haptics (force+tactile)
  - problems: structure/range estimation, modeling, tracking, materials, size, weight, inference
  - integration: navigation, manipulation, control, learning

- How can I solve “new” problems (generalization)?
  - sensors: vision, range, haptics, undefined new sensor
  - problems: categorization by function/shape/context/??
  - integrate: inference, navigation, manipulation, control, learning

Challenges/Issues

- About 60% of our brain is devoted to vision
- We see immediately and can form and understand images instantly



- Detailed representations are often not necessary
- Different approaches in the past Animate Vision (biologically inspired), Purposive Vision (depending on the task/purpose)

Visual Perception Topics

Techniques

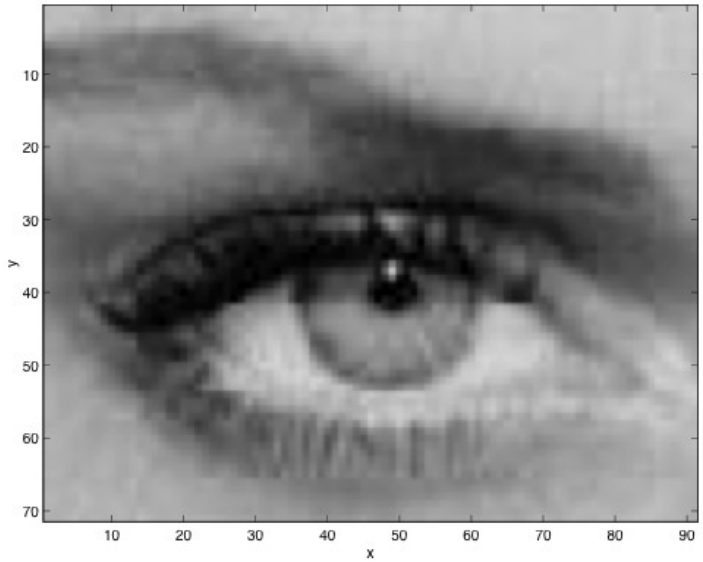
- Computational Stereo
- Feature detection and matching
- Motion tracking and visual feedback

Applications in Robotics:

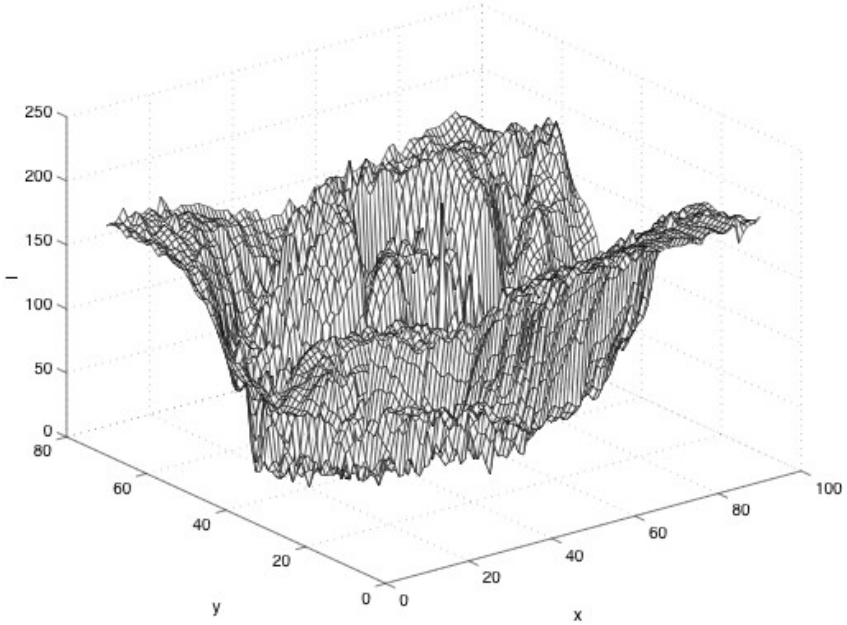
- range sensing, Obstacle detection, environment interaction
- Mapping, registration, localization, recognition
- Manipulation

Image - Apperance

Image



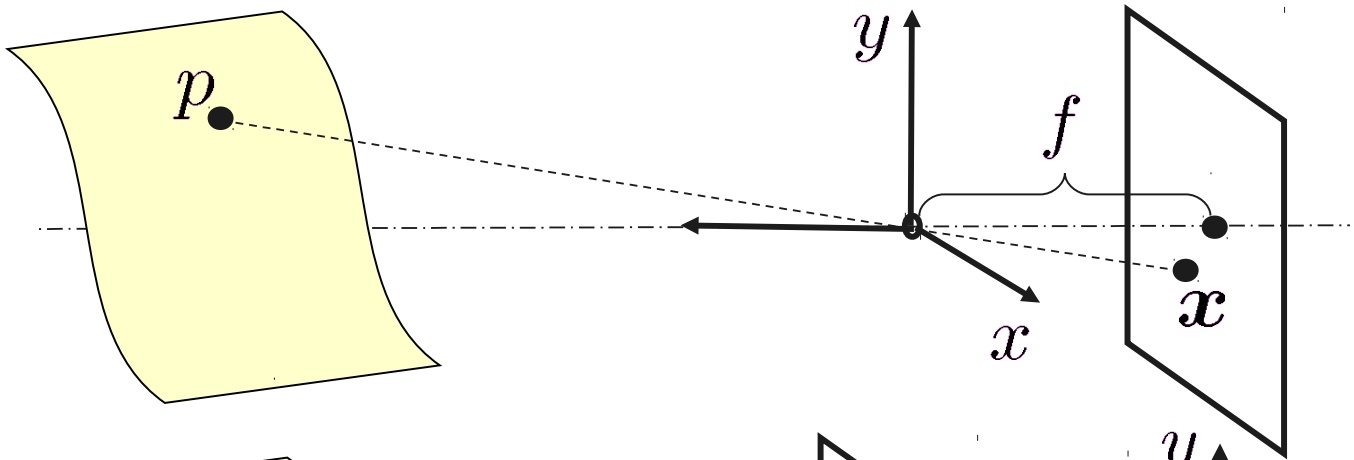
Brightness values



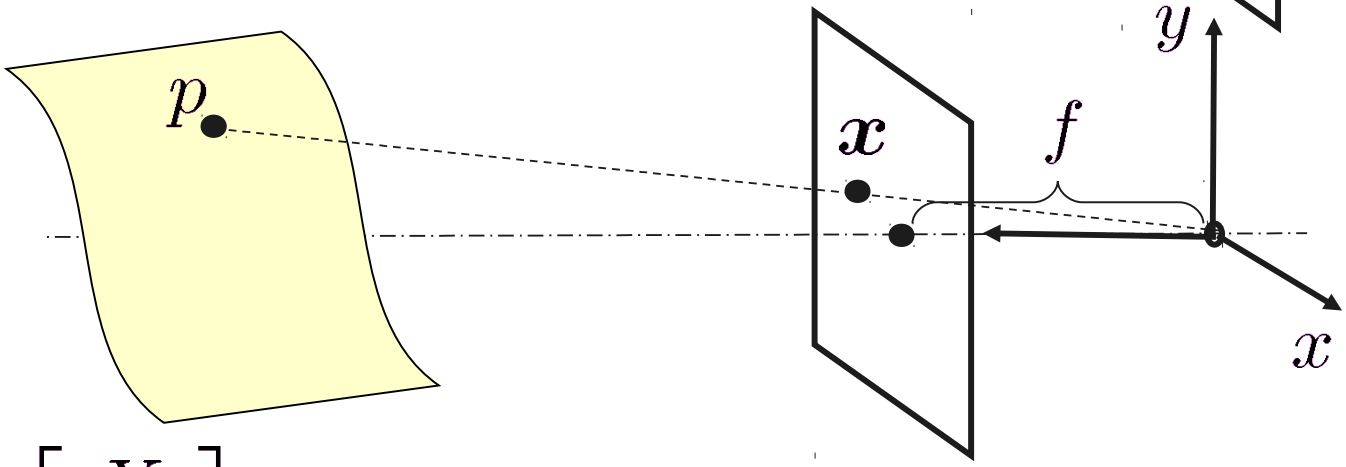
I(x,y)

Image Formation

Pinhole



Frontal pinhole



$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Pinhole Camera Model

- Image coordinates are nonlinear function of world coordinates
- Relationship between coordinates in the camera frame and sensor plan

2-D coordinates $$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Homogeneous coordinates

$$\boldsymbol{x} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}, \qquad \boldsymbol{X} \rightarrow \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

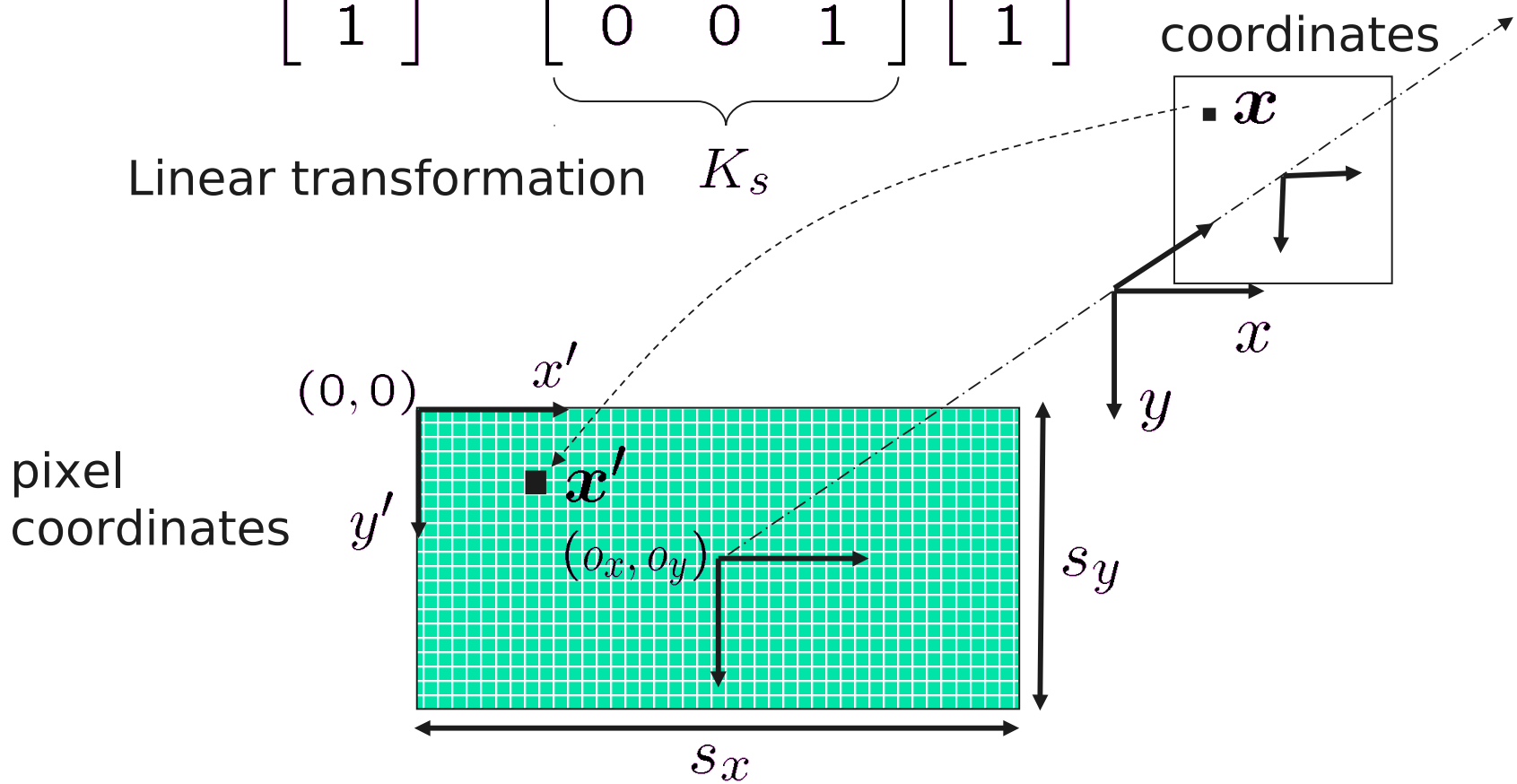
$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{K_f} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Image Coordinates

- Relationship between coordinates in the sensor plane and image

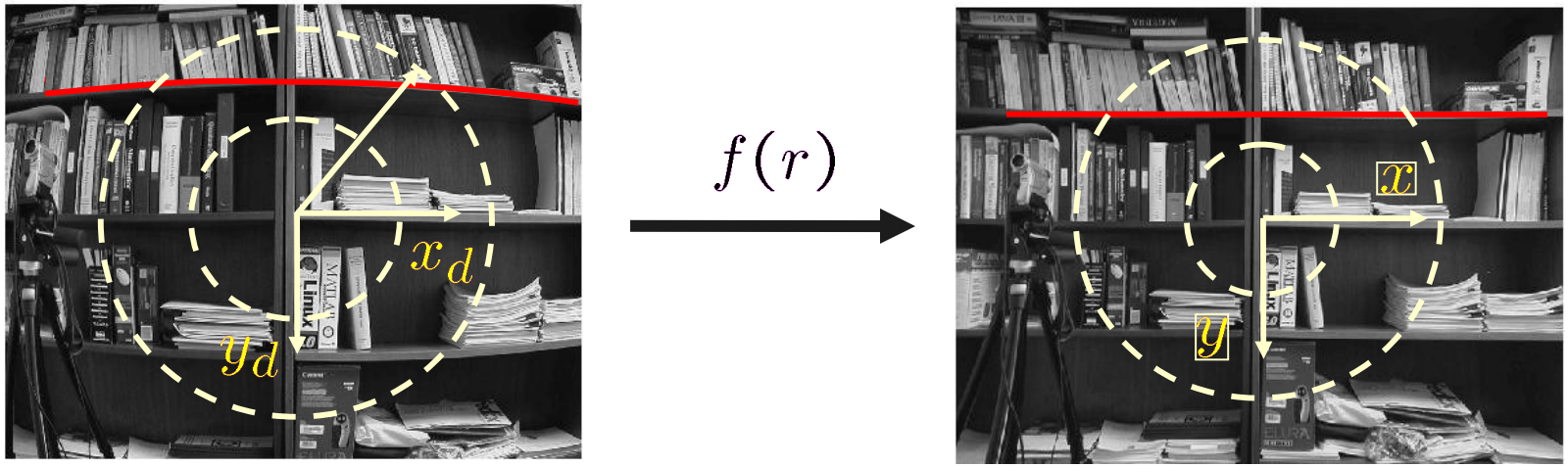
$$\boldsymbol{x}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{K_s} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Linear transformation



Camera parameters – Radial Distortion

Nonlinear transformation along the radial direction



$$\begin{aligned} \boldsymbol{x} &= c + f(r)(\boldsymbol{x}_d - c), \quad r = \|\boldsymbol{x}_d - c\| \\ f(r) &= 1 + a_1 r + a_2 r^2 + a_3 r^3 + a_4 r^4 + \cdots \end{aligned}$$

Distortion correction: make lines straight

Calibration Matrix and Camera Model

- Relationship between coordinates in the world frame and image
- Intrinsic parameters

Pinhole camera

$$\lambda \boldsymbol{x} = K_f \Pi_0 \boldsymbol{X}$$

Pixel coordinates

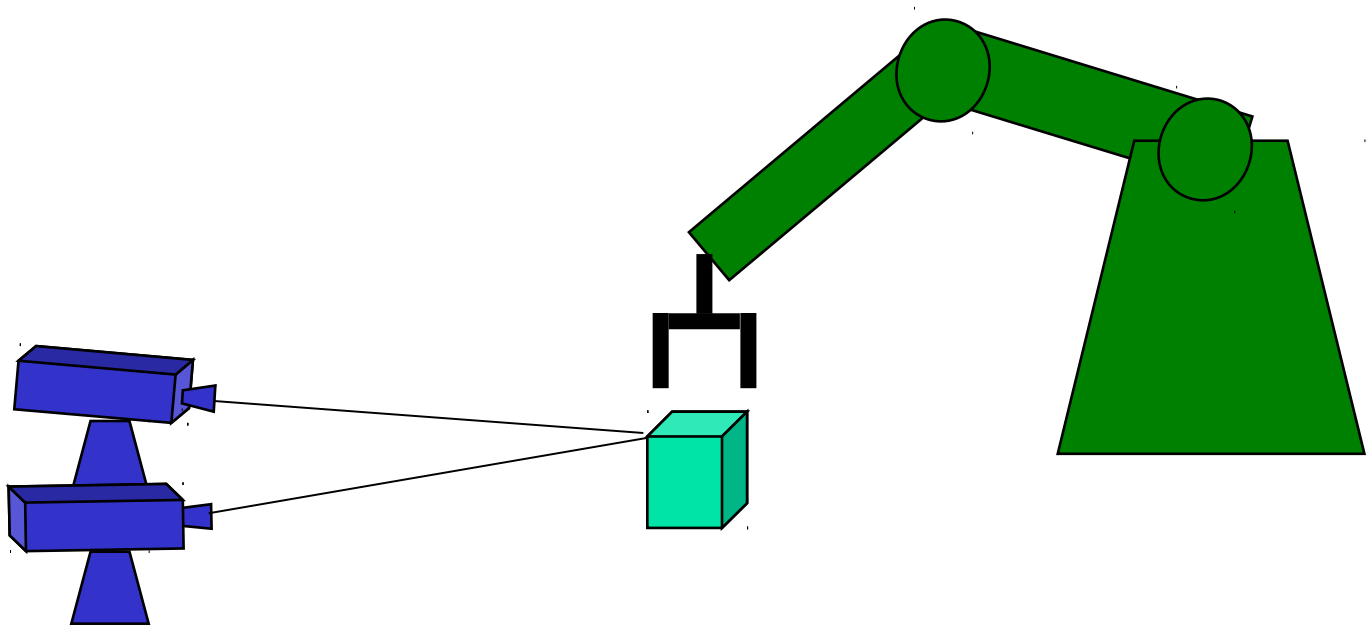
$$\boldsymbol{x}' = K_s \boldsymbol{x}$$

- Adding transformation between camera coordinate systems and world coordinate system
- Extrinsic Parameters

$$\lambda \boldsymbol{x}' = \begin{bmatrix} fs_x & fs_\theta & o_x \\ 0 & fs_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\lambda \boldsymbol{x} = K_f \Pi_0 g \boldsymbol{X} = \Pi \boldsymbol{X}$$

What is Computational Stereo?



Viewing the same physical point from two different viewpoints allows depth from triangulation

Computational Stereo

- Much of geometric vision is based on information from 2 (or more) camera locations
- Hard to recover 3D information from a single 2D image without extra knowledge
- Motion and stereo (multiple cameras) are both common in the world

- Stereo vision is ubiquitous in nature
  (oddly, nearly 10% of people are stereo blind)

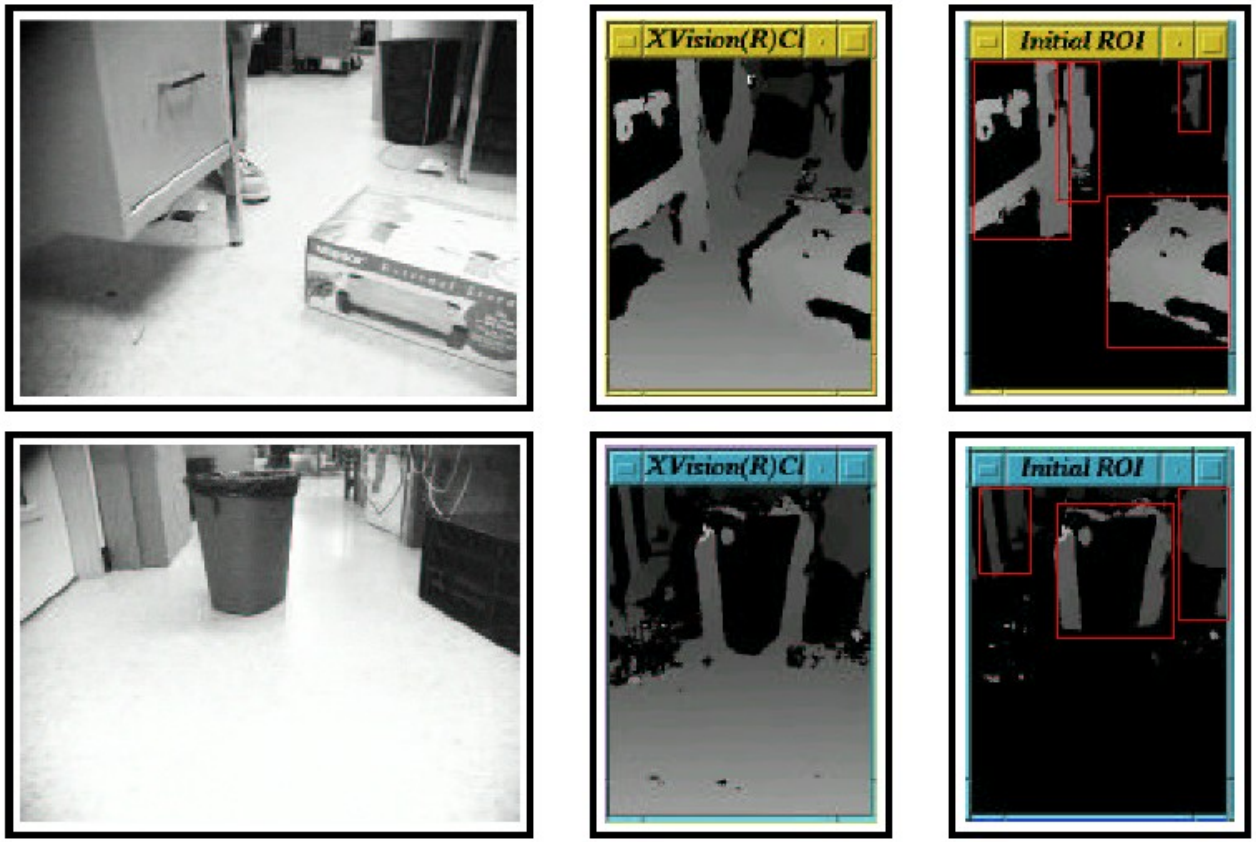
- Stereo involves the following *three problems*:
  - calibration
  - matching (correspondence problem)
  - reconstruction (reconstruction problem)

Applications of Real-Time Stereo

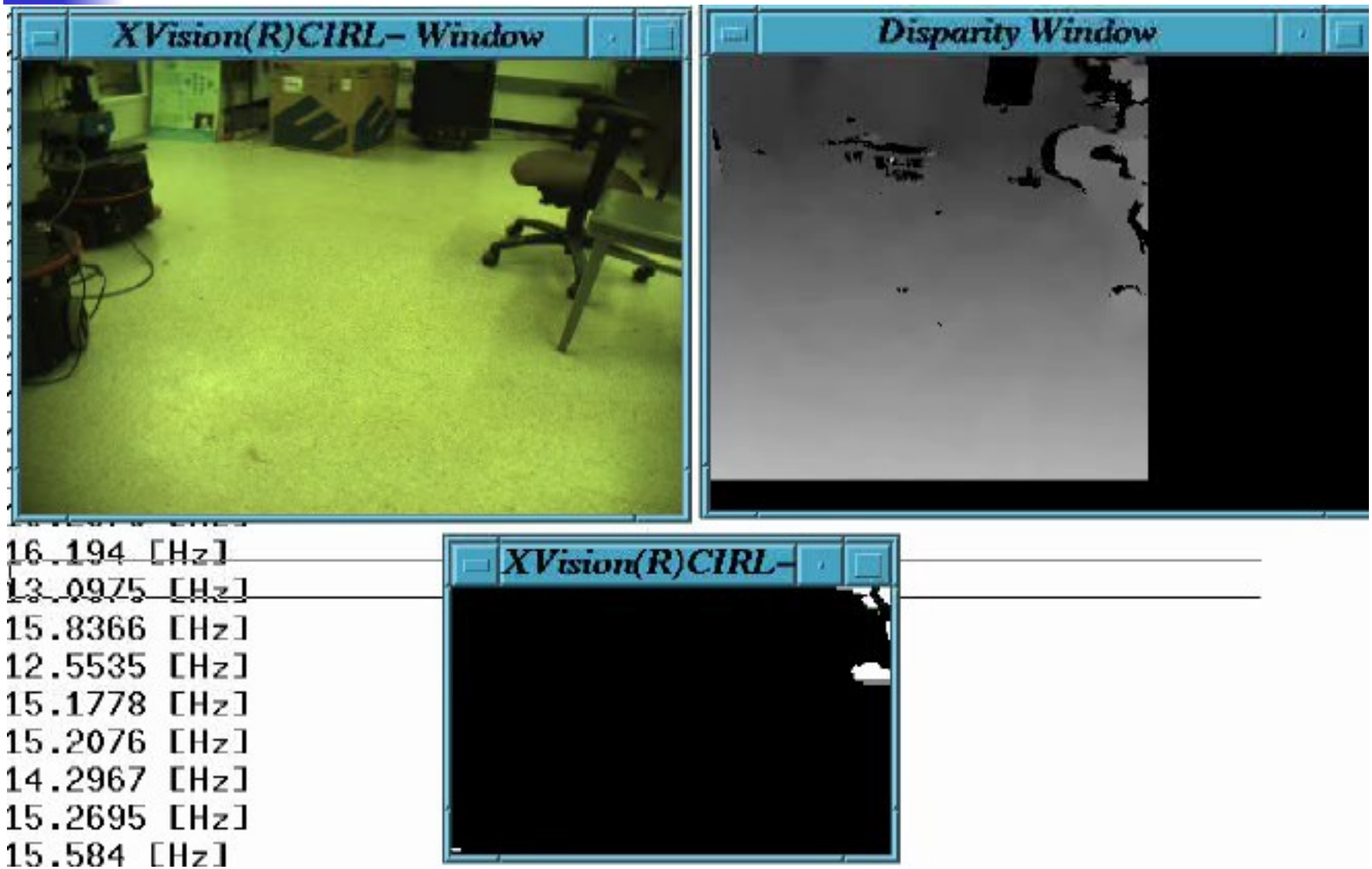
- Mobile robotics
  - Detect the structure of ground; detect obstacles; convoying
- Graphics/video
  - Detect foreground objects and matte in other objects (super-matrix effect)
- Surveillance
  - Detect and classify vehicles on a street or in a parking garage
- Medical
  - Measurement (e.g. sizing tumors)
  - Visualization (e.g. register with pre-operative CT)

Obstacle Detection (cont'd)

Observation: Removing the ground plane immediately exposes obstacles



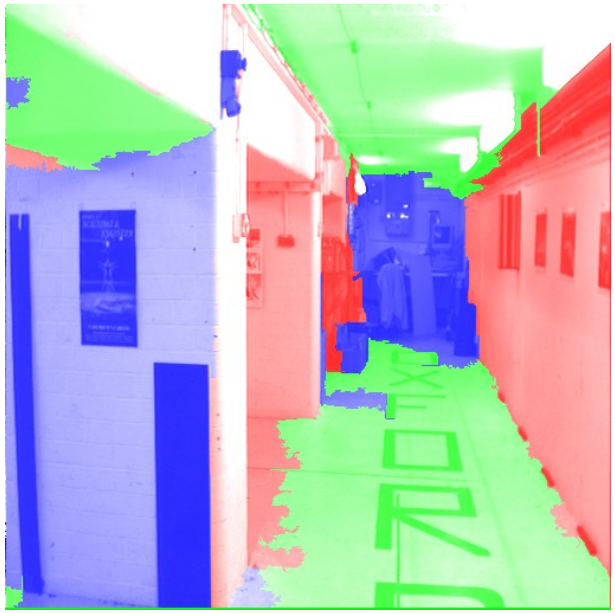
Applications of Real-Time Stereo



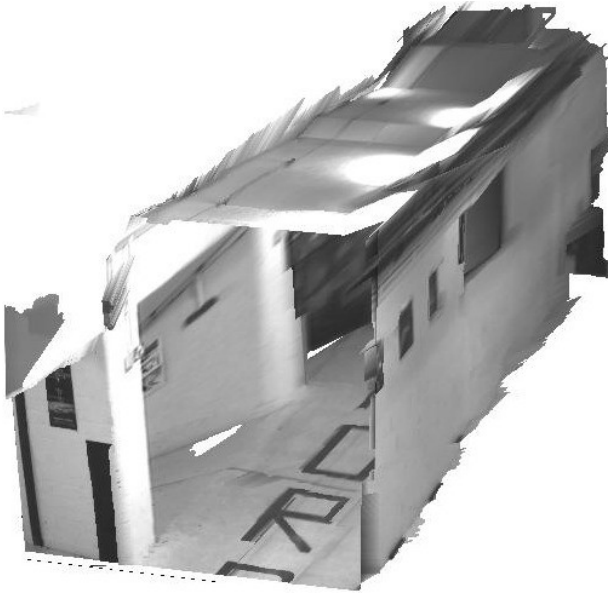
GMU building




Oxford corridor



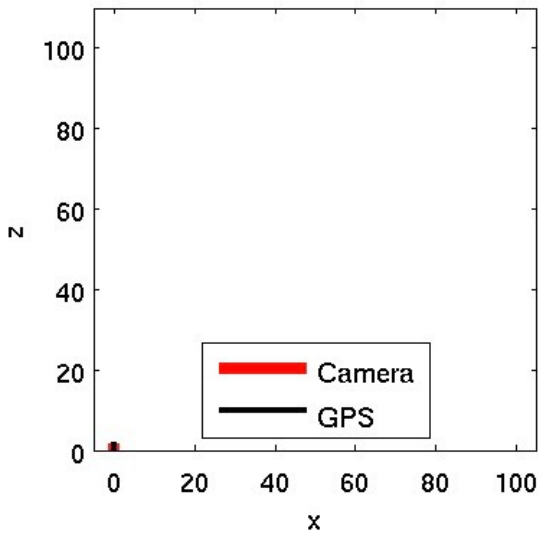
using 6 images

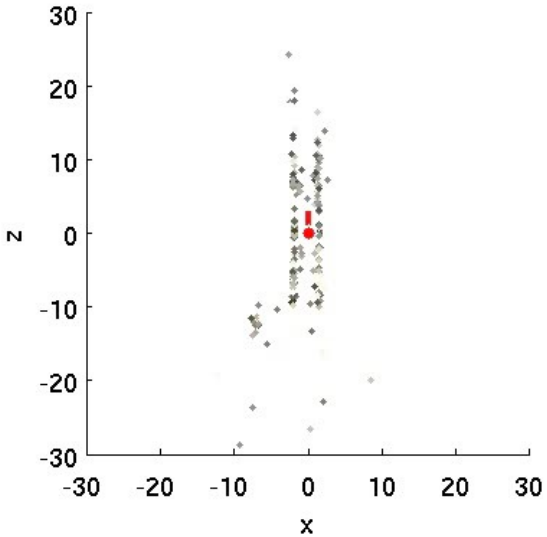


3D model

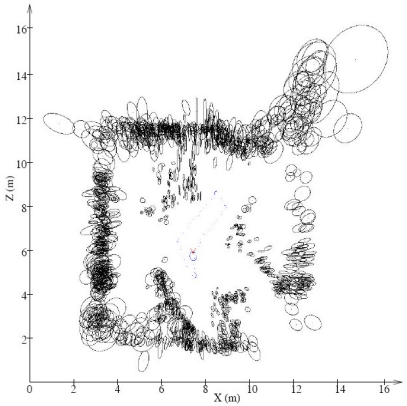
Visual Odometry

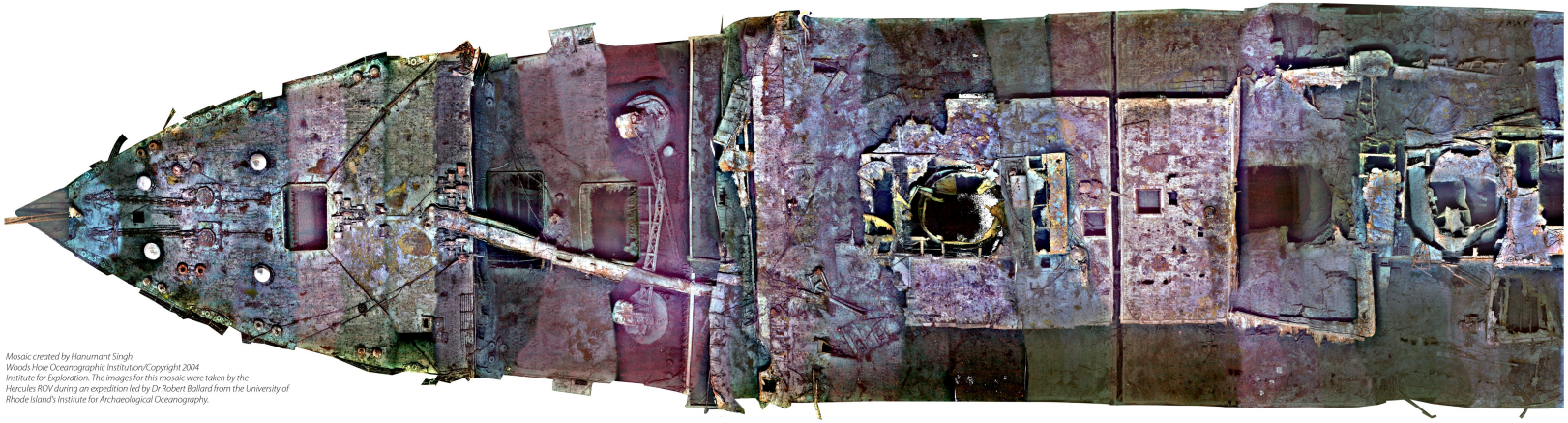
estimate motion from image correspondences





Mapping, Localization, Recognition





Mosaic created by Hanumant Singh,
Woods Hole Oceanographic Institution/Copyright 2004
Institute for Exploration. The images for this mosaic were taken by the
Hercules ROV during an expedition led by Dr Robert Ballard from the University of
Rhode Island's Institute for Archaeological Oceanography.