Perception

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

![Diagram of Perception Process]

1. Perception
2. Environment Model
   - Local Map
3. Localization
4. "Position" Global Map
5. Cognition
6. Path
7. Motion Control
   - Real World Environment
   - Environment Model
Bibabot, BlueBotics SA, Switzerland

- Omnidirectional Camera
- Pan-Tilt Camera
- Sonar Sensors
- Laser Range Scanner
- Bumper
- IMU (Inertial Measurement Unit)
- Emergency Stop Button
- Wheel Encoders
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 - Appearance

Image

Brightness values

\[ I(x, y) \]
Image Formation

Pinhole

Frontal pinhole

\[
X = \begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} \rightarrow \mathbf{x} = \begin{bmatrix}
x \\
y
\end{bmatrix} = \frac{f}{Z} \begin{bmatrix}
X \\
Y
\end{bmatrix}
\]
Pinhole Camera Model

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

**2-D coordinates**

\[ \mathbf{x} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \]

**Homogeneous coordinates**

\[ \mathbf{x} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}, \quad \mathbf{X} \rightarrow \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \]

\[ Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \]

\[ K_f \quad \mathbf{\Pi}_0 \]
Image Coordinates

- Relationship between coordinates in the sensor plane and image

\[
\begin{bmatrix}
  x' \\
  y' \\
  1
\end{bmatrix}
= \begin{bmatrix}
  s_x & s_\theta & o_x \\
  0 & s_y & o_y \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix}
\]

Linear transformation

metric coordinates

pixel coordinates

\( (0, 0) \)
Camera parameters – Radial Distortion

Nonlinear transformation along the radial direction

\[ x = c + f(r)(x_d - c), \quad r = \|x_d - c\| \]

\[ f(r) = 1 + a_1r + a_2r^2 + a_3r^3 + a_4r^4 + \cdots \]

Distortion correction: make lines straight
Calibration Matrix and Camera Model

- Relationship between coordinates in the world frame and image
- Intrinsic parameters
  
  Pinhole camera  
  \[
  \lambda \mathbf{x} = K_f \Pi_0 \mathbf{X}
  \]
  
  Pixel coordinates  
  \[
  \mathbf{x}' = K_s \mathbf{x}
  \]

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

\[
\lambda \mathbf{x}' = \begin{bmatrix}
  f_s x & f_s \theta & o_x \\
  0 & f_s y & o_y \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
  R & T \\
  0 & 1
\end{bmatrix}
\begin{bmatrix}
  X_w \\
  Y_w \\
  Z_w \\
  1
\end{bmatrix}
\]

\[
\lambda \mathbf{x} = K_f \Pi_0 g \mathbf{X} = \Pi \mathbf{X}
\]
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
Oxford corridor

using 6 images

3D model
Visual Odometry
estimate motion from image correspondences
Mapping, Localization, Recognition