Robot Perception

Nart Shawash
ELEE 5810

April 19, 2021

Table of contents

Sensors and Measurement

- Sensors are either laser, or ultrasound (sonar suffers from specular reflection).
- Inability of sonar to reliably measure range to near by object is often modeled as sensor noise.
- Inaccuracies of sensor models are described using conditional pdf: $p(z_t|x_t)$.
- Many sensors (eg. cameras) generate more than one numerical measurement $z_t = \{z_t^1, z_t^2, \cdots, z_t^K\}$.
- Assuming independence of z_t^i and z_t^j $(i \neq j)$ for convenience, we will use

$$p(z_t|x_t, m) = \prod_{k=1}^{K} p(z_t^k|x_t, m)$$

Maps: Feature based and Location based

Definition

A *map* m of the environment is a *list of objects* in the environment and their *locations*. $\boxed{m = \{m_1, m_2, m_3, \ldots, m_N\}}$.

- Feature based maps.
  - m_n contains: a) properties of a feature, b) cartesian location of a feature.
  - Specifies the shape of the environment at specific location.
  - Constructed from sensor data.
- Location based maps.
  - Uses $m_{(x,y)}$.
  - Occupancy grid assigns to each (x, y) a binary occupancy value.

§6.3.1 Basic Measurement Algorithm

- Beam finders measure the range to near by objects
  - along laser beams.
  - with a cone of ultrasonic sensor.
- Four types of measurement errors:
  1. Small measurement noise (truncated gaussian $p_{\text{hit}}(z_t^k|x_t, m)$)
  2. Errors due to unexpected objects (truncated exponential $p_{\text{short}}(z_t^k|x_t, m)$)
  3. Errors due to failure to detect objects (point mass at $z_{\max}$ $p_{\max}(z_t^k|x_t, m)$)
  4. Random unexplained noise ($p_{\text{rand}}(z_t^k|x_t, m)$ is uniform on $[0, z_{\max}]$)
- $p(z_t|x_t, m)$ is a "mixture" of 4 densities.

$$p(z_t^k|x_t, m) = \begin{pmatrix} z_{\text{hit}} & z_{\text{short}} & z_{\max} & z_{\text{rand}} \end{pmatrix} \begin{pmatrix} p_{\text{hit}}(z_t^k|x_t, m) \\ p_{\text{short}}(z_t^k|x_t, m) \\ p_{\max}(z_t^k|x_t, m) \\ p_{\text{rand}}(z_t^k|x_t, m) \end{pmatrix}$$

Correct range with local measurement noise. $p_{\text{hit}}(z_t^k | x_t, m)$

- Let z_t^{k*} be the true range of the object, and let z_t^k be the measurement.
- Error $z_t^k - z_t^{k*}$ is normally distributed, with mean z_t^{k*} and variance σ_{hit}^2.
- In practice measured values are limited to $[0, z_{\text{max}}]$. Where z_{max} is the sensor range. So need to truncate the gaussian, so it becomes
- $p_{\text{hit}}(z_t^k | x_t, m) = \begin{cases} \eta \mathcal{N}(z_t^k; z_t^{k*}, \sigma_{\text{hit}}^2) & \text{if } 0 \leq z_t^k \leq z_{\text{max}} \\ 0 & \text{otherwise} \end{cases}$
- Recall that $\mathcal{N}(z_t^k; z_t^{k*}, \sigma_{\text{hit}}^2) = \frac{1}{\sqrt{2\pi\sigma_{\text{hit}}^2}} e^{-\frac{1}{2}\frac{\left(z_t^k - z_t^{k*}\right)^2}{\sigma_{\text{hit}}^2}}$
- $\eta = \frac{1}{\int_0^{z_{\text{max}}} \mathcal{N}(z_t^k; z_t^{k*}, \sigma_{\text{hit}}^2)\, dz_t^k}$

Unexpected objects. $p_{\text{short}}(z_t^k | x_t, m)$

- Objects (people) in the environment not modeled by the map cause range finders to produce short ranges.
- Treated as sensor noise, with the property that they cause ranges to be shorter than z_t^{k*}.
- Two people (with $\{r_1, r_2\}$) appear in laser beams' way
  - will measure r_1 is both people or just 1st present.
  - will measure r_2 if only 2nd is present.
- Modeled by exponential distribution.
- Need to trancate, since $z_t^k \in [0, z_{k*}]$. And we obtain
- $p_{\text{short}}(z_t^k | x_t, m) = \begin{cases} \eta\, \lambda_{\text{short}}\, e^{-\lambda_{\text{short}} z_t^k} & \text{if } 0 \leq z_t^k \leq z_t^{k*} \\ 0 & \text{otherwise} \end{cases}$
- Here $\eta = \dfrac{1}{\int_0^{z_t^{k*}} \lambda_{\text{short}}\, e^{-\lambda_{\text{short}} z_t^k}\, dz_t^k} = \dfrac{1}{1 - e^{-\lambda_{\text{short}} z_t^{k*}}}$

Failures to detect objects. $p_{\max}(z_t^k | x_t, m)$

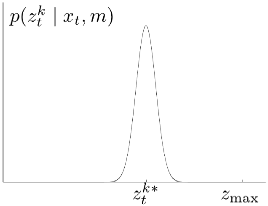
- Objects are not detected because
  - Specular reflection in sonar.
  - Black, light absorbing objects in laser detection.
  - Measuring objects in bright sun light using lasers.
- Typical result of sensor failure is max-range measurement.
- $p_{\max}(z_t^k | x_t, m) = \mathcal{I}(z = z_{\max}) = \begin{cases} 1 & \text{if } z = z_{\max} \\ 0 & \text{otherwise} \end{cases}$
- This distribution is referred to as *point mass* distribution.

Random unexplained noise measurements. $p_{\text{rand}}(z_t^k|x_t, m)$

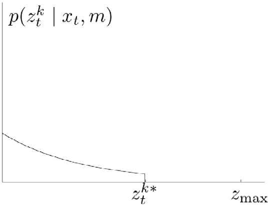
- Examples of random "unexplained" measurements could be
  - Phantom reading in sonars (bouncing off walls).
  - Crass-talk between sensors.
- Modeled using uniform distribution
- $p_{\text{rand}}(z_t^k|x_t, m) = \begin{cases} \frac{1}{z_{\text{max}}} & \text{if } 0 \le z_t^k \le z_{\text{max}} \\ 0 & \text{otherwise} \end{cases}$

Four densities

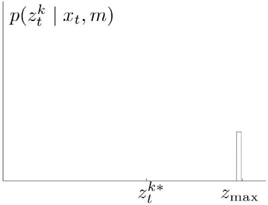
(a) Gaussian distribution p_{hit}



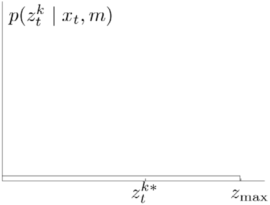
(b) Exponential distribution p_{short}



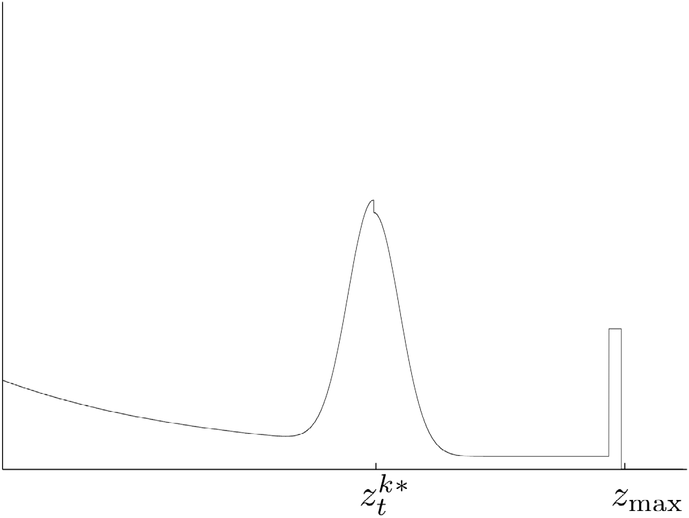
(c) Uniform distribution $p_{\max}$



(d) Uniform distribution p_{rand}



$p(z_t^k | x_t, m) = z_{\text{hit}} p_{\text{hit}} + z_{\text{short}} p_{\text{short}} + z_{\max} p_{\max} + z_{\text{rand}} p_{\text{rand}}$



Computing $p(z_t|x_t, m) = \prod_{k=1}^{K} p(z_t^k|x_t, m)$

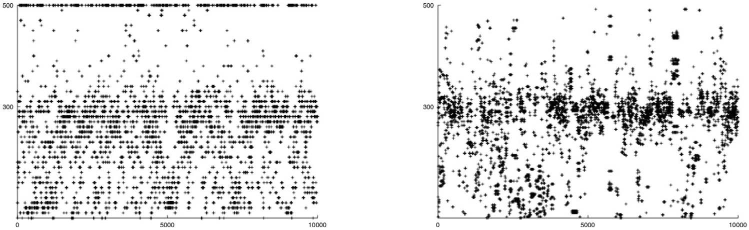
Recall that the robot has K sensors.

1. Algorithm beam-range-finder-model(z_t, x_t, m):
2. $q = 1$
3. for $k = 1$ to K do
4. compute z_t^{k*} for measurement the z_t^k using ray casting
5. $$p = z_{\text{hit}} \cdot p_{\text{hit}}(z_t^k|x_t, m) + z_{\text{short}} \cdot p_{\text{short}}(z_t^k|x_t, m) + z_{\text{max}} \cdot p_{\text{max}}(z_t^k|x_t, m) + z_{\text{rand}} \cdot p_{\text{rand}}(z_t^k|x_t, m)$$
6. $q \leftarrow q \cdot p$
7. return q

Adjusting intrinsic model parameters $\{z_{\text{hit}}, z_{\text{short}}, z_{\text{max}}, z_{\text{rand}}, \sigma^2_{\text{hit}}, \lambda_{\text{short}}\} = \Theta$

(a) Sonar data

(b) Laser data



Algorithm learn-intrinsic-parameters(Z, X, m)

  repeat until converges

    for all $z_i \in Z$ do

      $\eta = \frac{1}{p_{\text{hit}}(z_i|x_i,m)+p_{\text{short}}(z_i|x_i,m)+p_{\text{max}}(z_i|x_i,m)+p_{\text{rand}}(z_i|x_i,m)}$

      calculate z_i^*

      $e_{i,\text{hit}} = \eta\, p_{\text{hit}}(z_i|x_i, m)$

      $e_{i,\text{short}} = \eta\, p_{\text{short}}(z_i|x_i, m)$

      $e_{i,\text{max}} = \eta\, p_{\text{max}}(z_i|x_i, m)$

      $e_{i,\text{rand}} = \eta\, p_{\text{rand}}(z_i|x_i, m)$

    $z_{\text{hit}} = \left(\sum_i e_{i,\text{hit}}\right)/|Z|$

    $z_{\text{short}} = \left(\sum_i e_{i,\text{short}}\right)/|Z|$

    $z_{\text{max}} = \left(\sum_i e_{i,\text{max}}\right)/|Z|$

    $z_{\text{rand}} = \left(\sum_i e_{i,\text{rand}}\right)/|Z|$

    $\sigma_{\text{hit}}^2 = \frac{\sum_i e_{i,\text{hit}}(z_i - z_i^*)^2}{\sum_i e_{i,\text{hit}}}$

    $\lambda_{\text{short}} = \frac{\sum_i e_{i,\text{short}}}{\sum_i e_{i,\text{short}} z_i}$

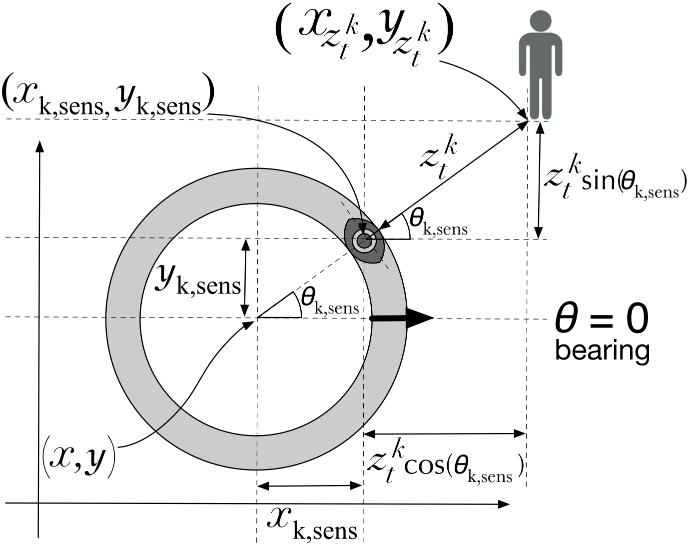
  return $\Theta = \{z_{\text{hit}}, z_{\text{short}}, z_{\text{max}}, z_{\text{rand}}, \sigma_{\text{hit}}^2, \lambda_{\text{short}}\}$

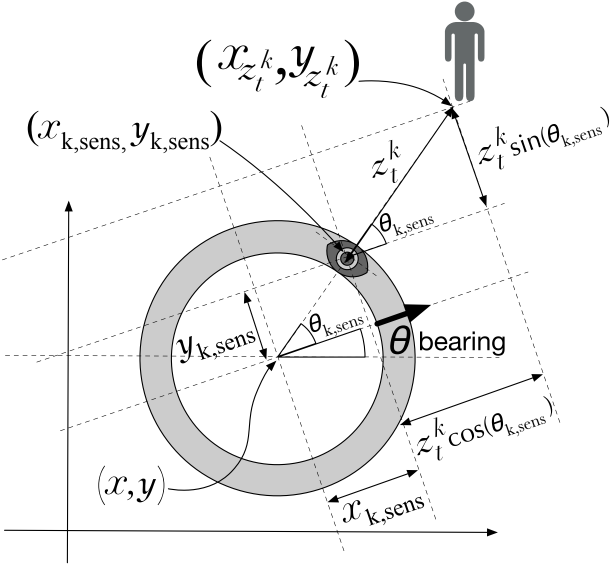
§6.4 Likelihood Fields and Range Finders

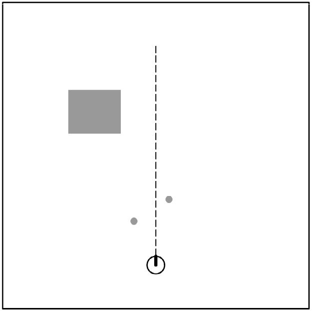
Recall: z_t^* is range measurement, and $x_t = (x \; y \; \theta)^T$ is robot pose.

- $\begin{pmatrix} x_{k,\text{sens}} \\ y_{k,\text{sens}} \end{pmatrix}$ is the relative location of the sensor in robot's fixed coordinate system.
- $\theta_{k,\text{sens}}$ is the angular orientation of sensor beam relative to robot's heading direction.
- The end point of measurement z_t^k, is mapped into global coordinate system using

$$\begin{pmatrix} x_{z_t^k} \\ y_{z_t^k} \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \underbrace{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_{k,\text{sens}} \\ y_{k,\text{sens}} \end{pmatrix}}_{\text{vector from robot's center to } k\text{th sensor}} + z_t^k \begin{pmatrix} \cos(\theta + \theta_{k,\text{sens}}) \\ \sin(\theta + \theta_{k,\text{sens}}) \end{pmatrix}$$


$(x_{z_t^k}, y_{z_t^k})$
$(x_{\text{k,sens}}, y_{\text{k,sens}})$
z_t^k
$z_t^k \sin(\theta_{\text{k,sens}})$
$\theta_{\text{k,sens}}$
$y_{\text{k,sens}}$
$\theta_{\text{k,sens}}$
$\theta = 0$
bearing
(x, y)
$z_t^k \cos(\theta_{\text{k,sens}})$
$x_{\text{k,sens}}$


$(x_{z_t^k}, y_{z_t^k})$
$(x_{k,sens}, y_{k,sens})$
z_t^k
$z_t^k \sin(\theta_{k,sens})$
$\theta_{k,sens}$
$\theta_{k,sens}$
$y_{k,sens}$
θ bearing
$z_t^k \cos(\theta_{k,sens})$
(x, y)
$x_{k,sens}$



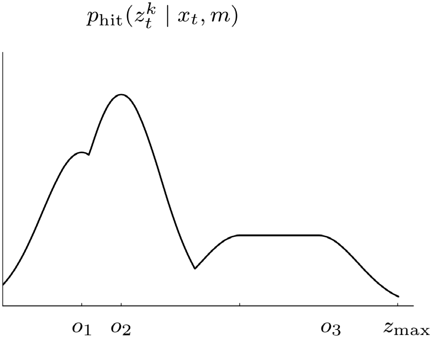


Figure: A cross-section through likelihood field.

Three types of sources of noise and uncertainty.

1. **Measurement noise**. Modeled using 2D (actually 1D) Gaussians.
   1. $dist = \min_{(x',y')} \left\{ \sqrt{\left(x_{z_t^k} - x'\right)^2 + \left(y_{z_t^k} - y'\right)^2} \,\middle|\, (x', y') \text{ occupied in } m \right\}$, and measurement coordinates $\left(x_{z_t^k}, y_{z_t^k}\right)$.
   2. Probability of sensor measurement is zero-centered Gaussian $p_{\text{hit}}(z_t^k | x_t, m) = \epsilon_{\sigma_{\text{hit}}}(dist)$
2. **Failures**. Max-range readings, modeled by point-mass distribution $p_{\max}$.
3. **Unexplained random measurements**. Uniform distribution p_{rand} used to model random noise in perception.

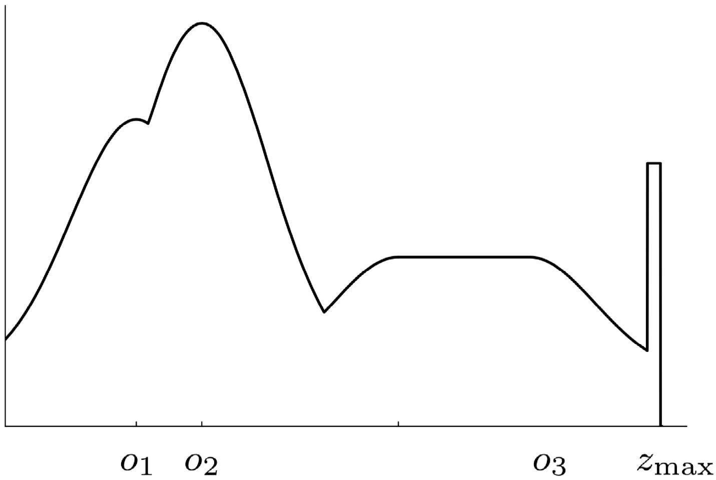


Figure: $p\left(z_k^t | x_t, m\right) = z_{\text{hit}} \cdot p_{\text{hit}} + z_{\text{rand}} \cdot p_{\text{rand}} + z_{\max} \cdot p_{\max}$

1. Algorithm likelihood-field-range-finder-model(z_t, x_t, m)
2. $q = 1$
3. for all k do
4. if $z_t^k \neq z_{\max}$
5. $x_{x_t^k} = x + x_{k,\text{sens}} \cos\theta - y_{k,\text{sens}} \sin\theta + z_t^k \cos(\theta + \theta_{k,\text{sens}})$
6. $y_{x_t^k} = y + y_{k,\text{sens}} \cos\theta + x_{k,\text{sens}} \sin\theta + z_t^k \sin(\theta + \theta_{k,\text{sens}})$
7. $dist = \min_{(x',y')} \left\{ \sqrt{\left(x_{z_t^k} - x'\right)^2 + \left(y_{z_t^k} - y'\right)^2} \right\}$
8. $q \leftarrow q \cdot \left(z_{\text{hit}} \cdot \mathbf{prob}(dist, \sigma_{\text{hit}}) + \frac{z_{\text{random}}}{z_{\max}} \right)$
9. return q

Comments on likelihood-field-range-finder-model

- $dist = \min_{(x',y')} \left\{ \sqrt{\left(x_{z_t^k} - x'\right)^2 + \left(y_{z_t^k} - y'\right)^2} \right\}$ is costly.
- Euclidean distance is smooth, so $p(z_t^k | x_t, m)$ is smooth.
- Disadvantages:
  1. Does not model people making short readings.
  2. Treats sensors as if they see through walls.
  3. Can not handle unexpected areas.

§6.5 Correlation-Based Measurement Models

- Goal is to transform consecutive scans (measurements) into occupancy maps or local maps m_{local}.
- *Map matching* uses correlations between measurements (scans) and the map m.
- The robot is at $x_t = (x\ y)^T$.
- $m_{x,y,\text{local}}(x_t)$ is the grid cell in local map containing $x_t = (x\ y)^T$.
- Compare both maps (measured and m) using

$$\rho_{m,m_{\text{local}},x_t} = \frac{\sum_{x,y}(m_{x,y} - \overline{m}) \cdot (m_{x,y,\text{local}}(x_t) - \overline{m})}{\sqrt{\sum_{x,y}(m_{x,y} - \overline{m})^2 \sum_{x,y}(m_{x,y,\text{local}}(x_t) - \overline{m})^2}}$$

- where $\overline{m} = \frac{1}{2N}\sum_{x,y}(m_{x,y} + m_{x,y,\text{local}})$ and N umber of overlaps between local and global maps.

- $-1 \leq \rho_{m,m_{\text{local}},x_t} \leq +1$.
- *Map matching* interprets $p(m_{\text{local}}|x_t, m) \stackrel{\text{as}}{=} \max\{\rho_{m,m_{\text{local}},x_t}, 0\}$.

Properties of *Map matching*:

1. Easy to compute.
2. The resulting $p(m_{\text{local}}|x_t, m)$ is not smooth as a function of x_t.
3. To smooth out the dependence on x_t, can convolve m with a Gaussian. This will give approximate likelihood field.

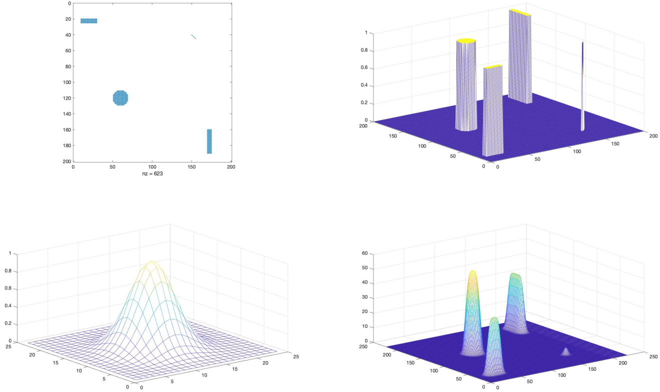


Figure: Convolving non-smooth map with a gaussian kernel.