Chapter 9 Localization and Mapping

Part 3

9.3 Simultaneous Localization and Mapping

Outline

- 9.3 Simultaneous Localization and Mapping
  - 9.3.1 Introduction
  - 9.3.2 Global Consistency in Cyclic Maps
  - 9.3.3 Revisiting
  - 9.3.4 EKF SLAM for Discrete Landmarks
  - 9.3.5 Example: Auto surveying of Laser Reflectors
  - Summary

Outline

- 9.3 Simultaneous Localization and Mapping
  - 9.3.1 Introduction
  - 9.3.2 Global Consistency in Cyclic Maps
  - 9.3.3 Revisiting
  - 9.3.4 EKF SLAM for Discrete Landmarks
  - 9.3.5 Example: Auto surveying of Laser Reflectors
  - Summary

9.3.1 Introduction

- The common acronym is S-L-A-M but C-L-M (for Concurrent Localization and Mapping) is used occasionally.
- Some say, the most fundamental problem in mobile robots.
- SLAM is a circular problem. The vehicle …
  - uses the landmark positions to determine its position and then
  - uses its position to update all of the landmarks.
- It is not possible to determine absolute position from such an arrangement, but...
- It is possible to use statistical modeling to remove most of the inconsistency between:
  - where a landmark is predicted to be based on measured vehicle motion and
  - where it now appears to be based on sensor readings.

Introduction

- Until a loop closes, SLAM is often akin to visual odometry.
- When a loop closes, things can get hard quickly.
  - How do you know it closed?
  - Does the entire map need to be updated?
- I will discuss the original Kalman filter formulation.
- I will assume bearing/range landmark measurements but bearing-only or range-only (or anything else) can be done similarly.

9.3.1 Introduction

(Map Quality Spectrum)

- A spectrum of increasing degrees of quality for maps includes the following stages
  - persistent storage (maps) enables point repeatability:
    - points in the map are fixed.
  - local smoothness enables tracking:
    - otherwise, discontinuities in the map cause loss of lock.
  - global (internal) consistency enables free-ranging:
    - minimal discontinuities associated with loop closure.
  - external consistency enables external programming:
    - E.g. map which is consistent with a CAD drawing.

Outline

- 9.3 Simultaneous Localization and Mapping
  - 9.3.1 Introduction
  - 9.3.2 Global Consistency in Cyclic Maps
  - 9.3.3 Revisiting
  - 9.3.4 EKF SLAM for Discrete Landmarks
  - 9.3.5 Example: Auto surveying of Laser Reflectors
  - Summary

9.3.2 Global Consistency in Cyclic Maps
(Setup)

- Suppose:
  - some kind of images overlap each other
  - the goal is to register them in all regions of overlap.
- Could be camera or rangefinder images.

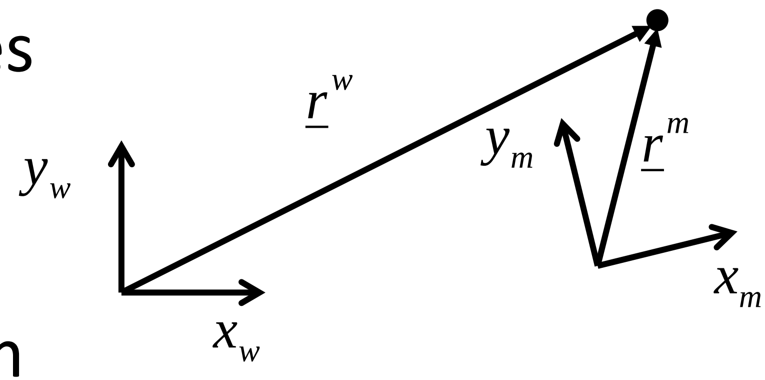
9.3.2 Global Consistency in Cyclic Maps

(Desiderata)

- Optimization.
  - Would like residuals in overlap regions to be as small as possible (zero is probably not possible due to distortion and feature localization error).
- Constraint.
  - Would like redundant degrees of freedom (if any) to be geometrically consistent.

9.3.2.1 Absolute Pose Formulation

- Suppose the vector $\underline{\rho}_m^w$ describes where image m is w.r.t a world frame.



- Residuals can be generated from images m and k:

$$\underline{r} = \underline{r}^w\big|_m - \underline{r}^w\big|_k = T_m^w(\underline{\rho}_m^w)\underline{r}^m - T_k^w(\underline{x}_k^w)\underline{r}^k$$

- This is of the form:

$$\underline{r} = h_m(\underline{\rho}_m^w) - h_k(\underline{\rho}_k^w)$$

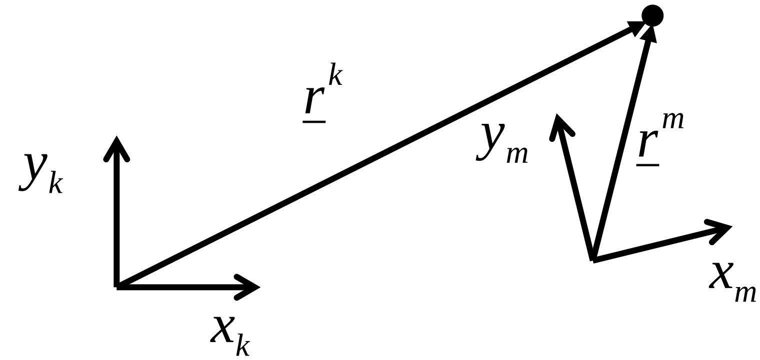
- Where both x's are the vector of all image absolute poses.

$$\underline{r}^w = T_m^w(\underline{\rho}_m^w)\underline{r}^m$$

$$\begin{bmatrix} x^w \\ y^w \\ 1 \end{bmatrix} = \begin{bmatrix} c\psi & -s\psi & a \\ s\psi & c\psi & b \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^m \\ y^m \\ 1 \end{bmatrix}$$

9.3.2.2 Relative Pose Formulation

- Suppose a vector $\underline{\rho}_m^k$ describes where image m is w.r.t. image k.



- Residuals can be generated from:

$$z = \underline{r}^k - \underline{r}^k\big|_m = \underline{r}^k - T_m^k(\underline{\rho}_m^k)\underline{r}^m$$

$$\underline{r}^k\big|_m = T_m^k(\underline{\rho}_m^k)\underline{r}^m$$

$$\begin{bmatrix} x^k \\ y^k \\ 1 \end{bmatrix} = \begin{bmatrix} c\theta & -s\theta & a \\ s\theta & c\theta & b \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^m \\ y^m \\ 1 \end{bmatrix}$$

- This is of the form:

$$\underline{r} = \underline{z} - h(\underline{\rho}_m^k) = \underline{z} - h(\underline{x})$$

- Where $\underline{x}$ is the vector of all image relative poses.

9.3.2.3 Unconstrained Optimization

- We might want to minimize something like:

$$\text{minimize:}_{\underline{x}} \qquad f(\underline{x}) = \frac{1}{2}\underline{r}(\underline{x})^T \underline{r}(\underline{x})$$

- Where $\underline{r}(\underline{x})$ is the composite of all residuals.
- Two Approaches:
  - Minimization
    - More robust, slower
  - Rootfinding
    - Assumes small residuals
    - Less robust, faster

9.3.2.3 Unconstrained Optimization

(Rootfinding Approach)

- Use the Jacobian from the linearized observer:

$$\Delta \underline{r} = H \Delta \underline{x} \quad \underline{x}$$

- When system is fully or even overdetermined, (more than enough features), we use the LPI and "solve" for the <u>whole residual</u> $\underline{r}$ in a single iteration.

$$\Delta \underline{x} = [H^T H]^{-1} H^T \underline{r}$$

- Unfortunately, $\Delta \underline{x}$ can have 10,000 elements and $\underline{r}$ can be larger so the matrices are huge.
  - Works only for small problems.

9.3.2.3 Unconstrained Optimization
(Minimization Approach)

- Gradient Descent:

$$\underline{d} = -\frac{\partial}{\partial \underline{x}} f(\underline{x}) \qquad \text{Follow gradient}$$

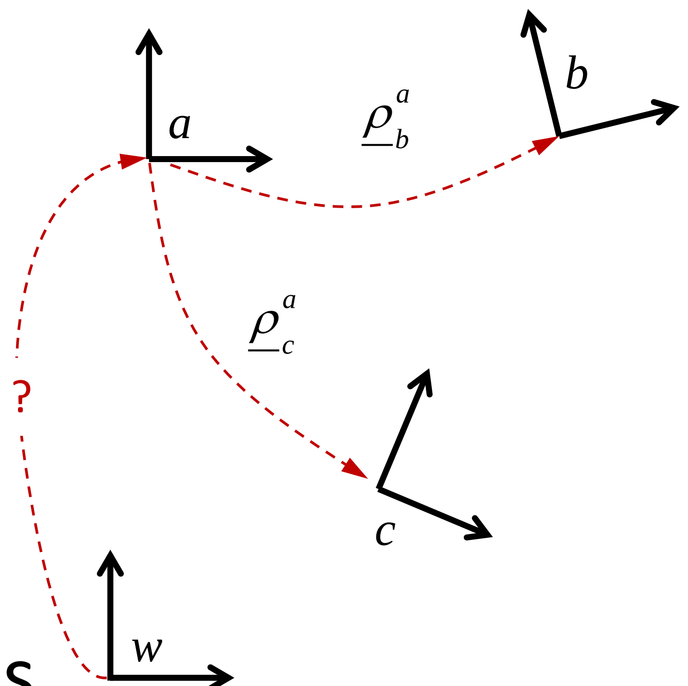
- Newton-Raphson:

$$\underline{d} = -\left[\frac{\partial^2 f}{\partial \underline{x}^2}\right]^{-1} \frac{\partial f}{\partial \underline{x}}(\underline{x}) \qquad \text{Find gradient roots}$$

- In either, there is never an issue getting enough equations to determine the step (always enough derivatives).

9.3.2.4 Constraint Satisfaction with Relative Poses

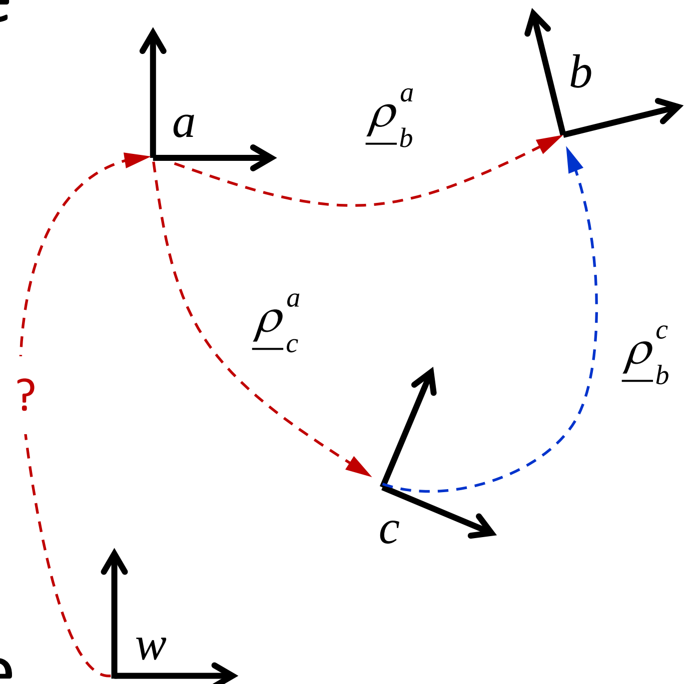
- Relative poses raise a different consistency issue.
- n images have n-1 degrees of pose freedom
  - Must fix one to fix the map location
- For n images, there are R possible, distinct, relative poses where:



$$R = \sum_i i = 1 + 2 + \ldots + n = \frac{n(n-1)}{2}$$

9.3.2.4 Constraint Satisfaction with Relative Poses

- If the state vector is larger than n-1 (i.e. if you have more relative poses than necessary)
  - the possibility of inconsistency arises
  - system of unconstrained poses is underconstrained because all the elements are not truly independent.
- There are natural consistency (i.e. loop closure) constraints that should be imposed of the form: $\underline{g}(\underline{x}) = \underline{b}$

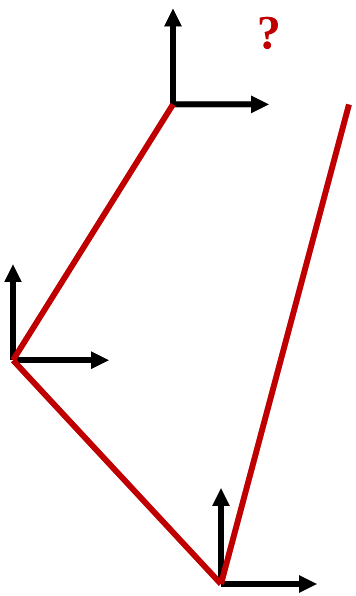


Carnegie Mellon
THE ROBOTICS INSTITUTE

9.3.2.4 Constraint Satisfaction with Relative Poses

- Again: There are natural constraints which should be imposed to adequately constrain the system:

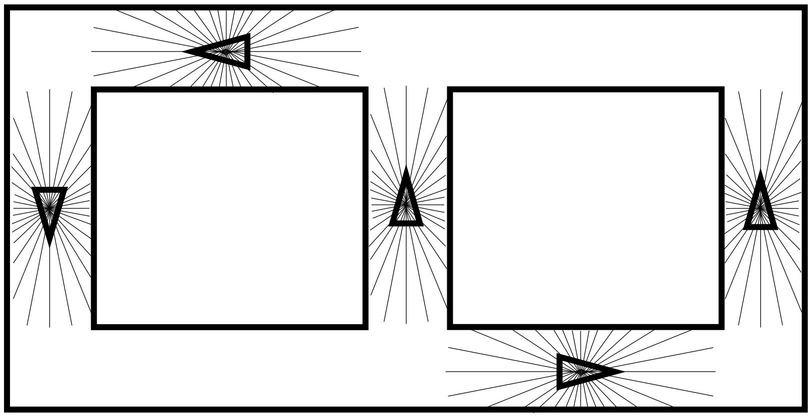
$$\underline{g}(\underline{x}) = \underline{b}$$



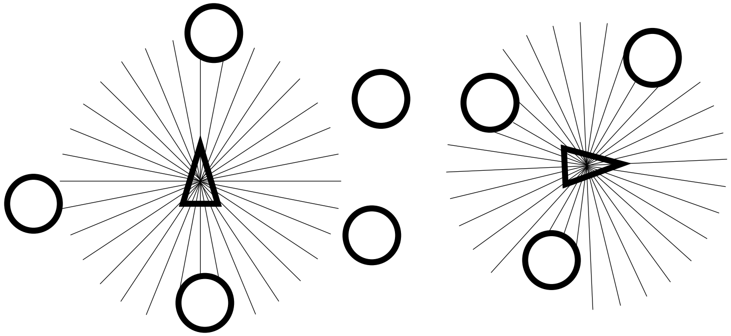
- These are <u>entirely separate from</u> the feature residual.
- Unless loops close, there is no issue of consistency.

9.3.2.4.1 Loop Constraints – Sparse Case

- Constraint equations are needed when loops close.
- Relatively few constraints are needed in sparse networks.
- Leads to use of a slightly redundant state vector and a few constraints.



Hallway Environment

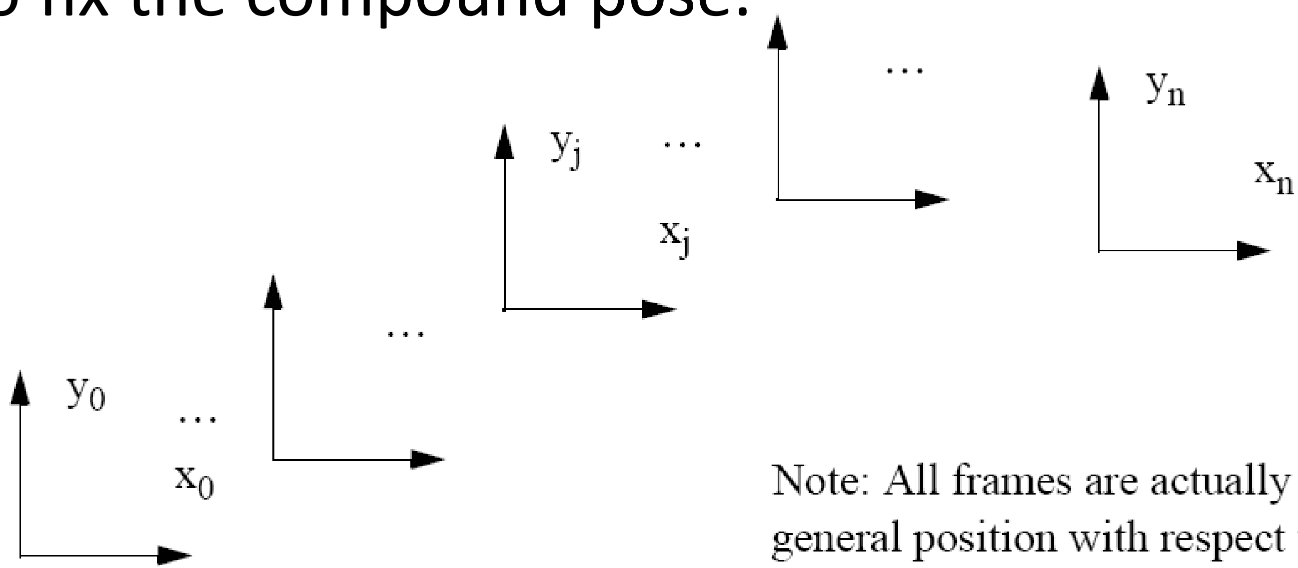


Forest Environment

Carnegie Mellon THE ROBOTICS INSTITUTE

9.3.2.4.2 Warping for Constraint Enforcement

- Suppose:
  - n poses which relate each of n+1 frames in an ordered sequence to its predecessor.
  - the relationship of the first frame with respect to the last has just been determined to be <u>slightly wrong</u>.
- Problem is to change ALL of the intermediate poses in order to fix the compound pose.



Note: All frames are actually in general position with respect to their predecessors.

9.3.2.4.2 Warping for Constraint Enforcement
(Total Differential)

- We can write the total differential <u>in terms of left and right pose Jacobians</u>:

$$\Delta\underline{\rho}_n^0 = \begin{bmatrix} \frac{\partial \underline{r}_n^0}{\partial \underline{\rho}_1^0} & \dots & \frac{\partial \underline{r}_n^0}{\partial \underline{\rho}_n^{n-1}} \end{bmatrix} \begin{bmatrix} \Delta\underline{\rho}_1^0 \\ \dots \\ \Delta\underline{\rho}_n^{n-1} \end{bmatrix}$$

- This is an underdetermined system which can be solved with the <u>right</u> pseudoinverse

$$\Delta\underline{\rho} = J^T[JJ^T]^{-1}\Delta\underline{\rho}_n^0$$

9.3.2.4.2 Warping for Constraint Enforcement

(Closing Loops)

- Special case of warping (last slide).
- When the pose $\underline{\rho}_n^0$ is associated with a loop, we have a constraint which is something like:

$$\underline{\rho}_1^0 * \underline{\rho}_2^1 \cdots \underline{\rho}_n^{n-1} * \underline{\rho}_0^n = \underline{0}$$

- This is notation for:

$$T_1^0 T_2^1 \ldots T_n^{n-1} T_0^n = I$$

- Which is of the form:

$$\underline{g}(\underline{x}) = \underline{b}$$

Often, maps are so smooth locally that enforcing loop constraints is enough to make a good map.

However, the extra computation needed to get an optimal map is also trivial once the loops close.

9.3.2.5 Constrained Optimization

- Can also do both optimization and constraint enforcement <u>at the same time</u>.
- Formulation is:

$$\text{minimize:} \quad f(\underline{x}) = \frac{1}{2}\underline{z}(\underline{x})^T \underline{z}(\underline{x})$$

$$\text{subject to:} \quad \underline{g}(\underline{x}) = \underline{b}$$

9.3.2.5 Constrained Optimization

(Penalty Function)

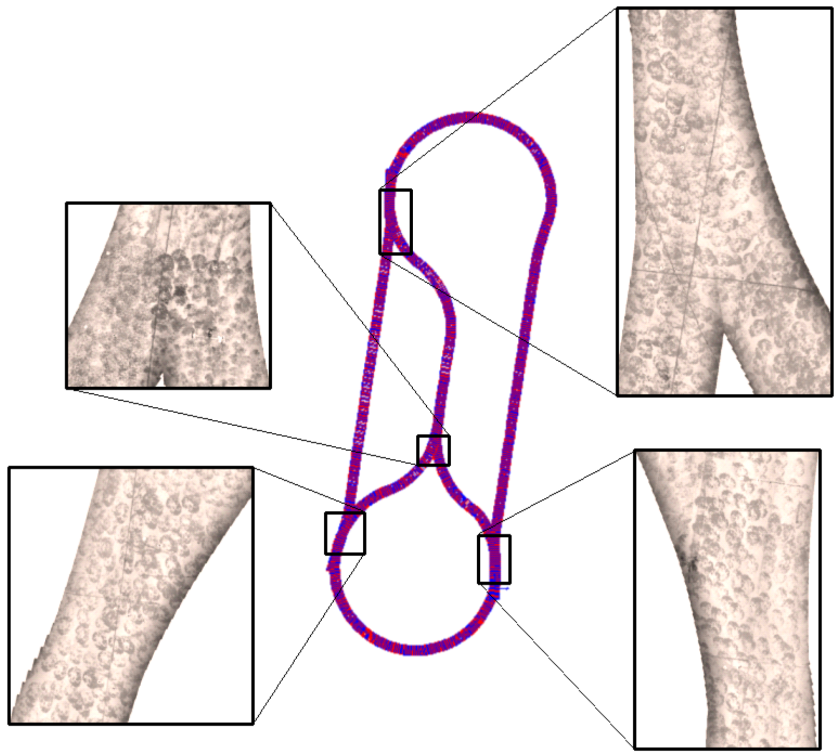
- Can also form a "constraint residual":

$$z(\underline{x}) = \underline{b} - \underline{g}(\underline{x})$$

- Add this to the overlap residual and find the minimum overall residual.
- This is the penalty function approach to constrained optimization.

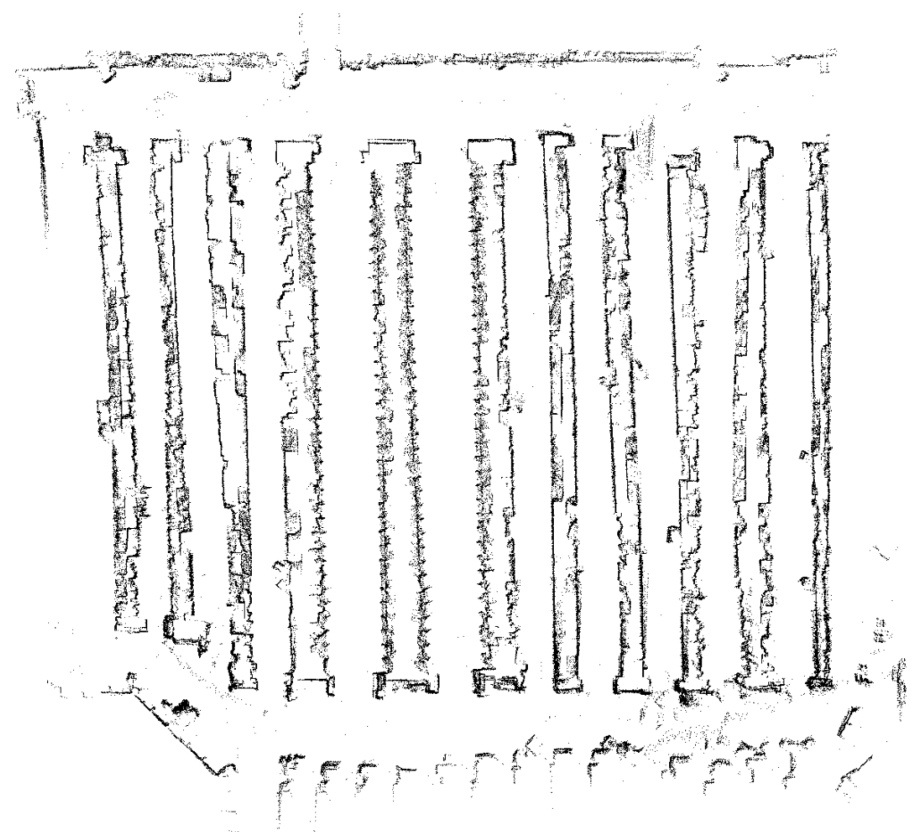
9.3.2.6 Example Floor Imagery Mosaics as Maps (AGV Guidance Maps)

- Guidance based on Mosaics of floor imagery.



9.3.2.6 Example: Large Scale Lidar Maps (Grocery Store)

- 10,000 images have been rendered globally consistent.
- About 10 seconds of computation.



Outline

- 9.3 Simultaneous Localization and Mapping
  - 9.3.1 Introduction
  - 9.3.2 Global Consistency in Cyclic Maps
  - 9.3.3 Revisiting
  - 9.3.4 EKF SLAM for Discrete Landmarks
  - 9.3.5 Example: Auto surveying of Laser Reflectors
  - Summary

9.3.3.1 Issues

- One of the most difficult problems in mobile robots.
  - connected to unsupervised object recognition
  - unsupervised because the system has to generate the models to be matched.
- Uniqueness/Aliasing:
  - If multiple places actually <u>look the same</u>, the difficulty is more serious.
- Omnidirectional sensors matter – Why?

9.3.3.1 Issues

- To do SLAM, you potentially have to check every image against every historical image to see if there is a match.
  - Maybe you need to match entire submaps of neighborhoods if one image is not unique.
  - Position estimates can be used to reduce search.
  - Multi-hypothesis approaches are in vogue.
  - Others have used correlation schemes.

9.3.3.1 Issues

(Global Data Association)

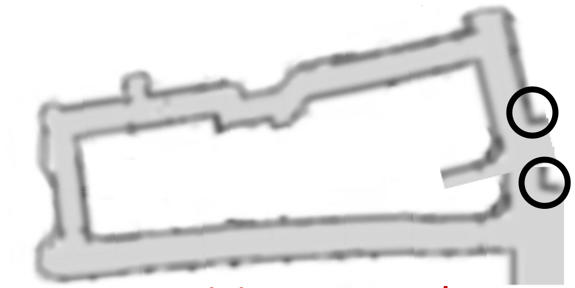
- Global level data association problems involve heuristic rather than brute-force search.
- Computer vision always has to solve this problem as part of the object recognition problem. Some examples include…
  - View/aspect recognition - know which piece of an object you are looking at.
  - Mosaicking / Global Registration - know which pieces belong together.
- Mobile robotics defines two instances of this problem:
  - Place recognition: determining that your sensor readings are consistent with a particular place in some map.
  - Revisiting problem: determining that your sensor readings are consistent with being in a place you have been before before a map is built.

9.3.3.2 Example: Revisiting from Lidar

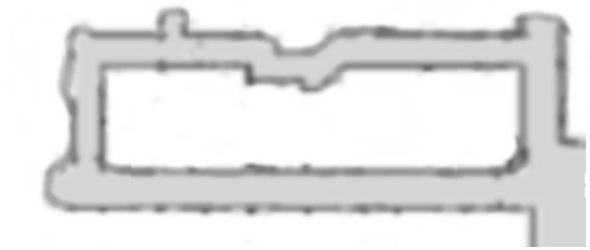
- Suppose the map is a certainty grid m(I,j).
- Correlate incoming image grids l(I,j) with the map m(I,j) so far using H/W acceleration.

$$p(\rho|l, m) = k\sum_{i \in I}\sum_{j \in J} l[\rho, i, j]m[i, j]$$

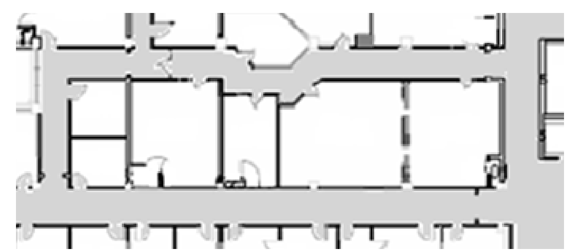
- Different approach: lidar keypoints-based revisit detection has been done <u>on the scale of entire cities</u> (Brisbane).



Revisit Detected



Repaired Map

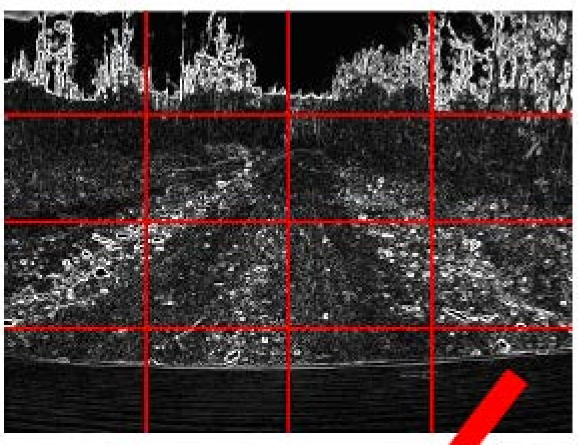


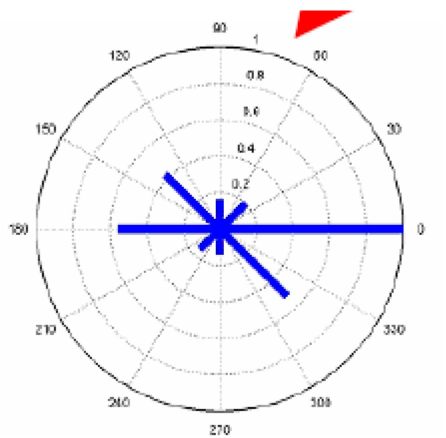
True Floor Plan

9.3.3.3 Example: Revisiting from Video

(Weighted Gradient Oriented Histograms)

- Divide image into 4 X 4 regions.
- Compute polar histogram of gradient magnitudes.
  - 8 bin histogram of …
  - gradients at each point …
  - weighted by distance from center …
  - and weighted by gradient magnitude





8 bin Histogram

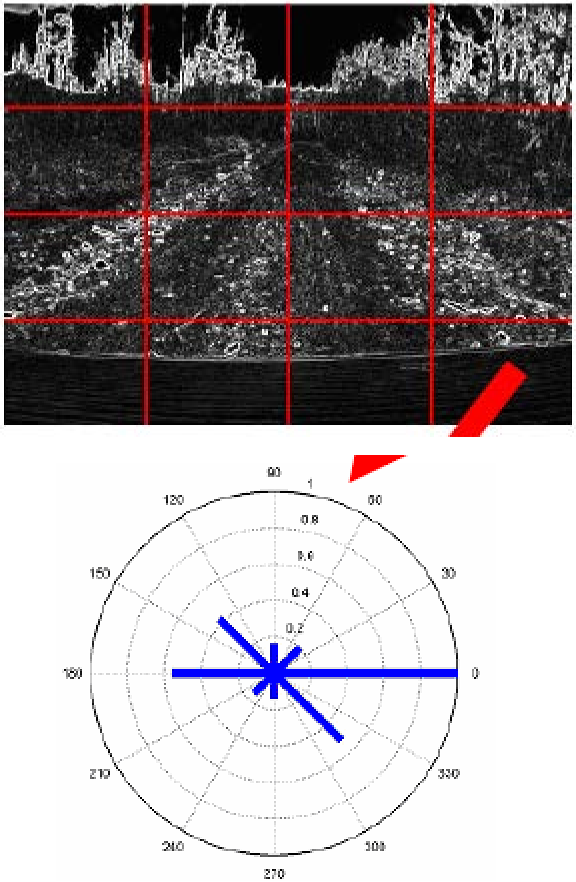
Carnegie Mellon
THE ROBOTICS INSTITUTE

9.3.3.3 Example: Revisiting from Video
(Weighted Gradient Oriented Histograms)

- Concatenate all 16, 8 bin histograms into a 128-vector.
- Normalize to unit length.
- These 128 numbers encode the (image at) the place.
- Comparison of images to training set based on dot product feature space

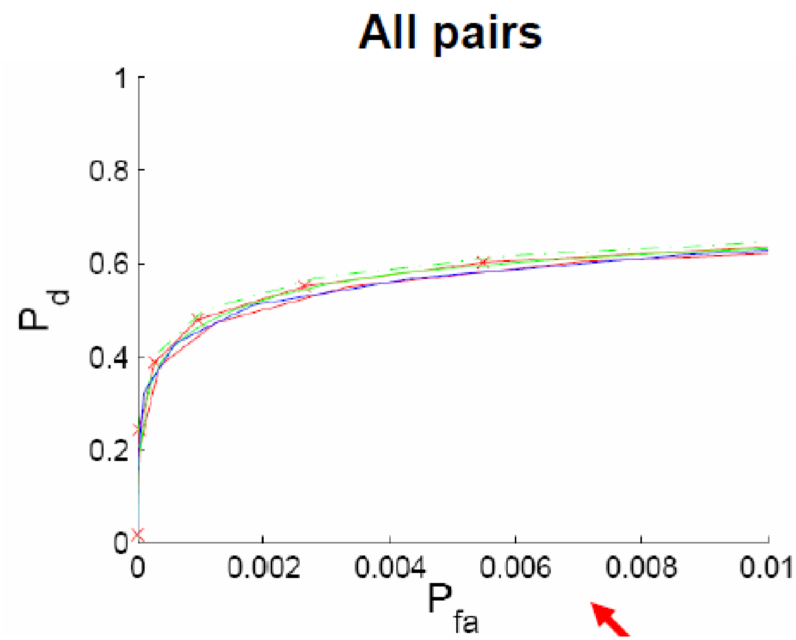
$$d(X_i, X_j) = 1 - X_i X_j^T$$

0 = parallel
1= orthogonal



Example: Weighted Gradient Oriented Histograms

- Nearest neighbor classifier trained on 4700 prior images.



- 80 % probability of detection
- With 6% false alarm rate

Outline

- 9.3 Simultaneous Localization and Mapping
  - 9.3.1 Introduction
  - 9.3.2 Global Consistency in Cyclic Maps
  - 9.3.3 Revisiting
  - <u>9.3.4 EKF SLAM for Discrete Landmarks</u>
  - 9.3.5 Example: Auto surveying of Laser Reflectors
  - Summary

9.3.4.1 System Model

(Vehicle)

- As we have seen before…
- State vector: $\underline{x} = \left[x \;\; y \;\; \theta \;\; V \;\; \omega \right]^T$
- Dynamics: $\dot{\underline{x}} = \left[-V\sin\theta \;\; V\cos\theta \;\; \omega \;\; 0 \;\; 0 \right]^T$
- "Transition matrix":

$$\hat{\Phi} \approx \begin{bmatrix} 1 & 0 & 0 & -s\theta dt & 0 \\ 0 & 1 & 0 & c\theta dt & 0 \\ 0 & 0 & 1 & 0 & dt \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

9.3.4.1 System Model
(SLAM)

- State vector includes landmarks:

$$\underline{x} = \begin{bmatrix} x & y & \theta & V & \omega & x_1 & y_1 & \dots & x_n & y_n \end{bmatrix}^T$$

- Dynamics for landmarks …

$$\dot{\underline{x}}_L = \frac{d}{dt}\begin{bmatrix} x_1 \\ y_1 \\ \dots \\ x_n \\ y_n \end{bmatrix} = \underline{0}$$

Landmarks don't move with time.

9.3.4.2 State Covariance Propagation

- Partition the state vector thus:

$$\underline{x} = \begin{bmatrix} \underline{x}_v^T & \underline{x}_L^T \end{bmatrix}^T$$

- Transition matrix:

$$\Phi = \begin{bmatrix} \Phi_{vv} & 0 \\ 0 & I \end{bmatrix}$$

- Partition State Covariance:

$$P = \begin{bmatrix} P_{vv} & P_{vL} \\ P_{Lv} & P_{LL} \end{bmatrix}$$

9.3.4.2 State Covariance Propagation

- Recall Covariance Propagation:

$$P^-_{k+1} = \Phi_k P_k \Phi_k^T + \Gamma_k Q_k \Gamma_k^T$$

- First term:

$$\Phi_k P_k \Phi_k^T = \begin{bmatrix} \Phi_{vv} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_{vv} & P_{vL} \\ P_{Lv} & P_{LL} \end{bmatrix} \begin{bmatrix} \Phi_{vv}^T & 0 \\ 0 & I \end{bmatrix}$$

$$\Phi_k P_k \Phi_k^T = \begin{bmatrix} \Phi_{vv} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_{vv}\Phi_{vv}^T & P_{vL}I \\ P_{Lv}\Phi_{vv}^T & P_{LL}I \end{bmatrix}$$

$$\boxed{\Phi_k P_k \Phi_k^T = \begin{bmatrix} \Phi_{vv}P_{vv}\Phi_{vv}^T & \Phi_{vv}P_{vL}I \\ IP_{Lv}\Phi_{vv}^T & IP_{LL}I \end{bmatrix}}$$

9.3.4.2 State Covariance Propagation

- Recall Covariance Propagation:

$$P^-_{k+1} = \Phi_k P_k \Phi_k^T + \Gamma_k Q_k \Gamma_k^T$$

- Second term:

$$\Gamma_k Q_k \Gamma_k^T = \begin{bmatrix} \Gamma_{vv} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} Q_{vv} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Gamma_{vv}^T & 0 \\ 0 & I \end{bmatrix}$$

Landmark uncertainty does not grow with time..

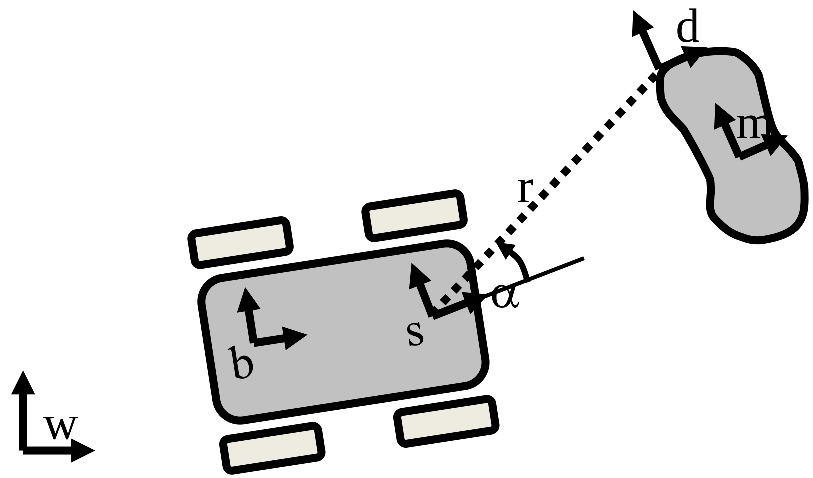
$$\Gamma_k Q_k \Gamma_k^T = \begin{bmatrix} \Gamma_{vv} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} Q_{vv}\Gamma_{vv}^T & 0 \\ 0 & 0 \end{bmatrix}$$

$$\boxed{\Gamma_k Q_k \Gamma_k^T = \begin{bmatrix} \Gamma_{vv} Q_{vv} \Gamma_{vv}^T & 0 \\ 0 & 0 \end{bmatrix}}$$

Independent of number of landmarks.

9.3.4.3 Measurement Model

- Covered in KF Slides



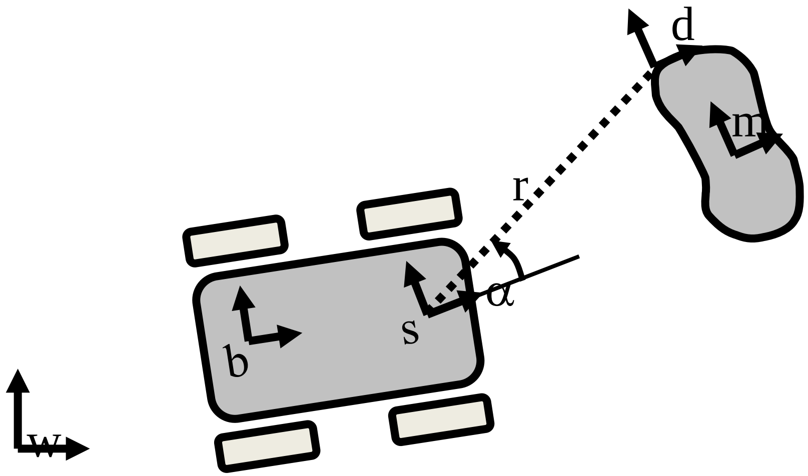
$$\underline{\rho}_d^s = \underline{\rho}_b^s * \underline{\rho}_w^b * \underline{\rho}_m^w * \underline{\rho}_d^m$$

9.3.4.3 Measurement Model

(Landmark Jacobian)

- Covered in KF Slides..
- Jacobian w.r.t veh pose.

$$H_x^z = \left(\frac{\partial \underline{z}}{\partial \underline{\rho}_d^s}\right)\left(\frac{\partial \underline{\rho}_d^s}{\partial \underline{\rho}_d^b}\right)\left(\frac{\partial \underline{\rho}_d^b}{\partial \underline{\rho}_b^w}\right) = H_s^z H_b^s H_x^b$$



- Jacobian w.r.t landmark pose:

$$\underline{\rho}_d^s = \underline{\rho}_b^s * \underline{\rho}_w^b * \underline{\rho}_m^w * \underline{\rho}_d^m$$

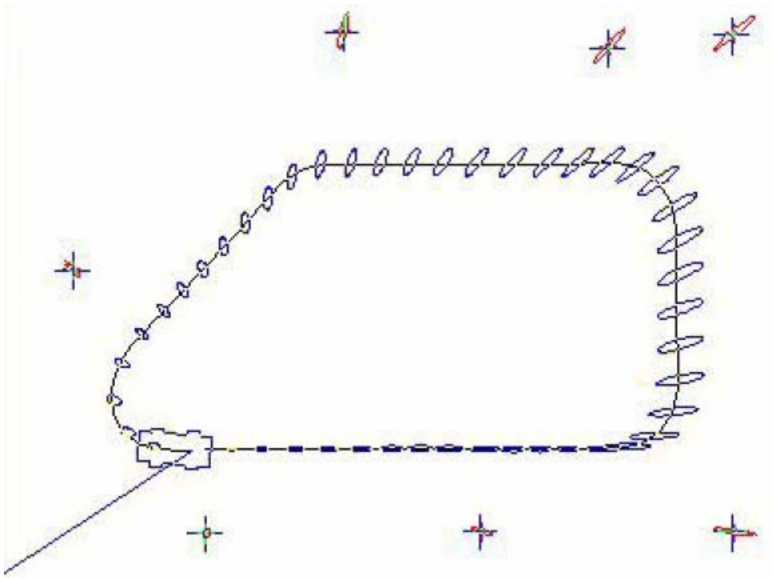
$$H_m^z = \left(\frac{\partial \underline{z}}{\partial \underline{\rho}_d^s}\right)\left(\frac{\partial \underline{\rho}_d^s}{\partial \underline{\rho}_d^b}\right)\left(\frac{\partial \underline{\rho}_d^b}{\partial \underline{\rho}_d^w}\right)\left(\frac{\partial \underline{\rho}_d^w}{\partial \underline{\rho}_m^w}\right) = H_s^z H_b^s H_w^b H_m^w$$

Outline

- 9.3 Simultaneous Localization and Mapping
  - 9.3.1 Introduction
  - 9.3.2 Global Consistency in Cyclic Maps
  - 9.3.3 Revisiting
  - 9.3.4 EKF SLAM for Discrete Landmarks
  - 9.3.5 Example: Auto surveying of Laser Reflectors
  - Summary

Initialization

- The uncertainty in SLAM is bounded from below by the uncertainty of the initial conditions.
- The map can never be more accurate than the error in the initial position.
- In bearing-only SLAM, downrange localization is poor. When the robot gets near landmarks, its uncertainty takes on the character of that of nearby landmarks.

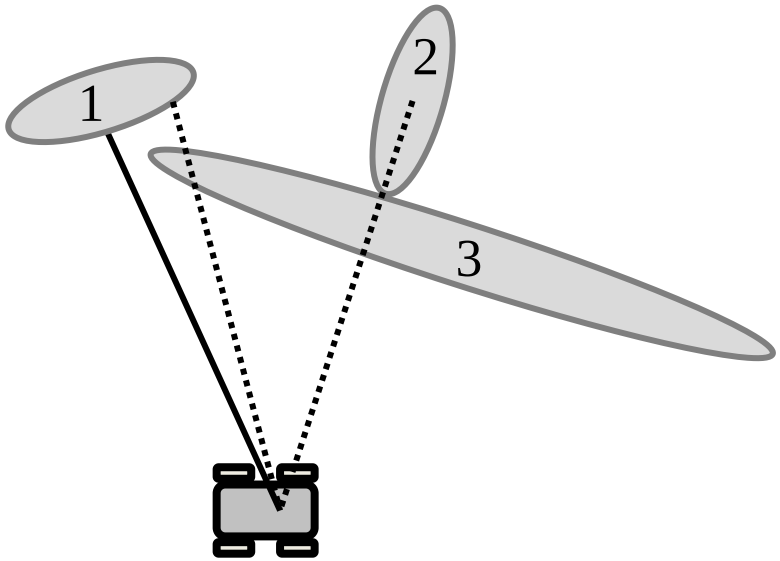


Initialization

- In the most general case, the landmarks are not known beforehand in number or location.
- In scenarios where measurements do not fully constrain landmark positions.
  - need some kind of structure from motion.
  - Data association may be pretty hard to do. Some kind of visual tracking may be called for.

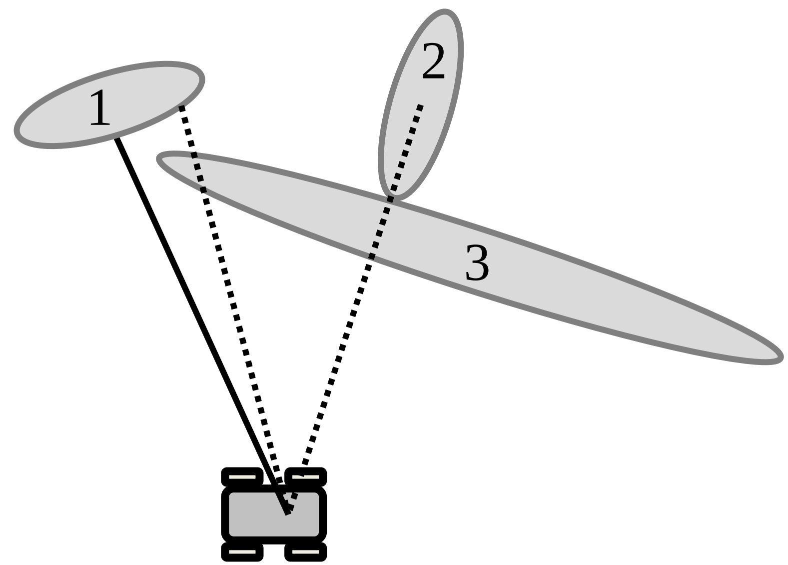
9.3.5 Example: Auto surveying of Laser Reflectors

- Motivation:
  – Surveying laser reflectors in factories costs a lot of money.
  – They move around and change in visibility as the plant is altered.
- It is possible to drive the robot around in a factory and use the laser guidance system to survey the reflector positions.
- Assume that the number and approximate location of landmarks is known beforehand.



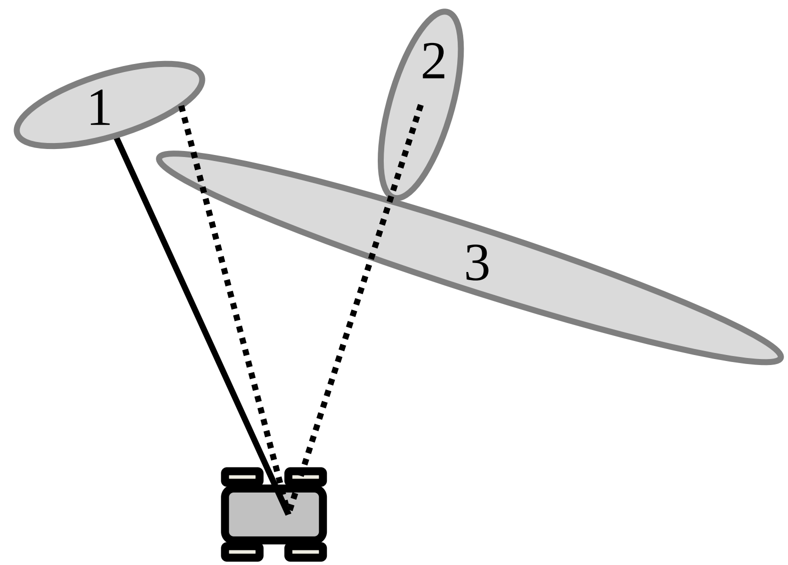
9.3.5 Example: Auto surveying of Laser Reflectors

- 3+ well-known coordinates (1-1/2 landmarks) visible initially makes a huge difference.
- Reasonableness tests:
  - is reflective side of landmark facing the laser?



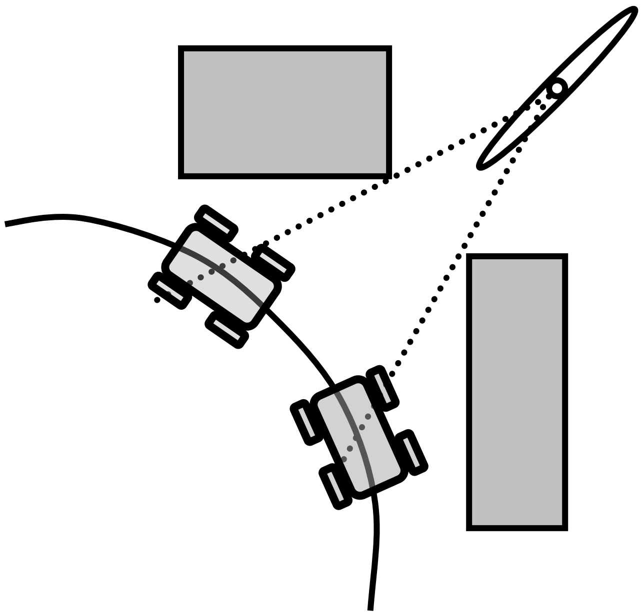
9.3.5.2 Data Association

- Sometimes uncertainty ellipses may overlap when projected onto the sensor space (bearing):
  - impossible to associate any readings unambiguously
  - impossible to locate the landmark
- Enough large ellipses and the system cannot work.

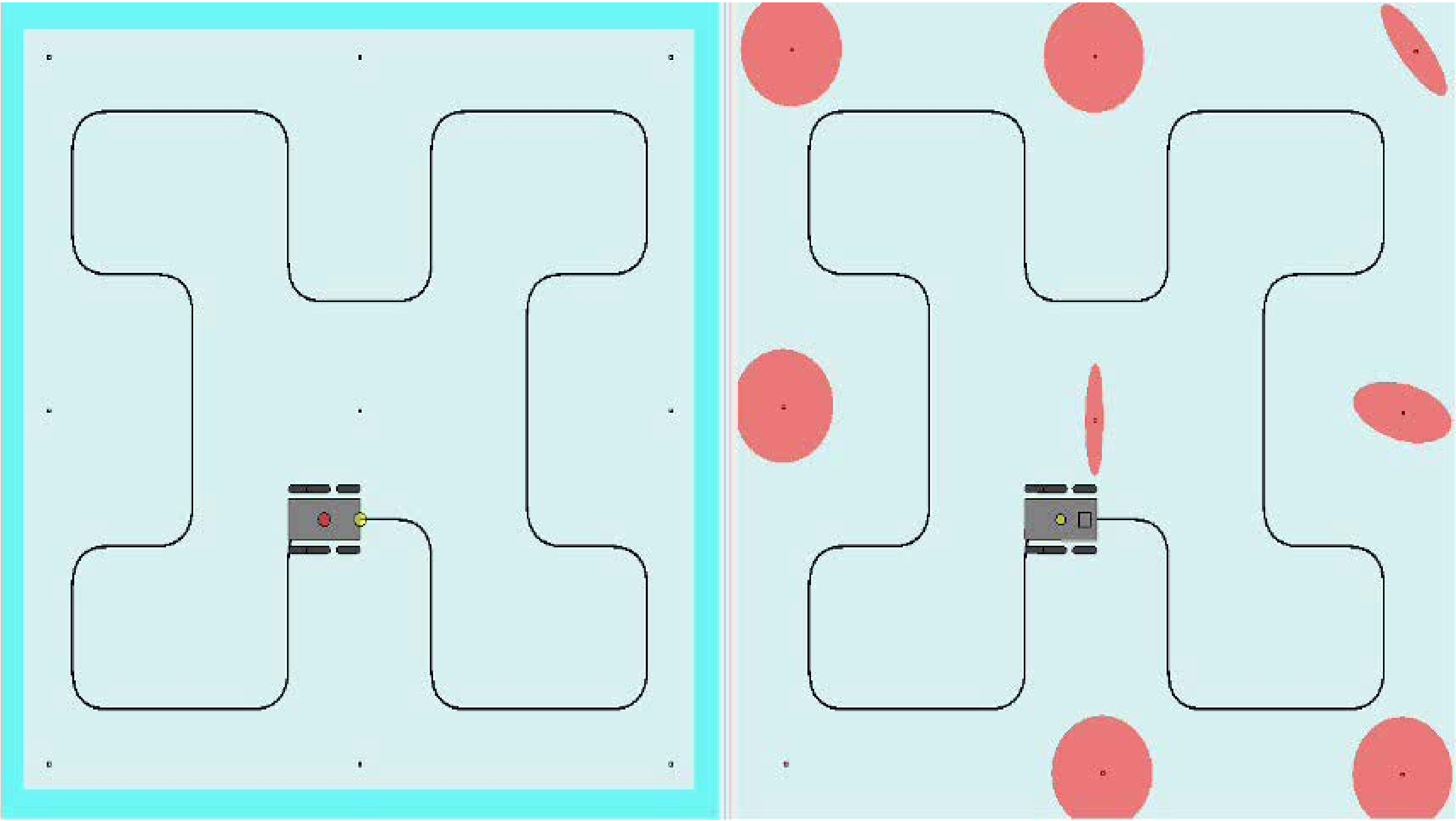


9.3.5.3 View Conditioning

- Reflector is viewed over a narrow range of viewing angles
  - its position along the direction of the laser cannot be resolved well.
- Does not too negatively affect the robot pose
  - Pose is insensitive to depth variation.
- Two sides of the same coin.



Video



9.3.5.4 Brittleness

- SLAM with Kalman filters is a house of cards.
- One incorrect positive association has great potential to break everything.
- False negatives (not using data that you could have used) is much less of a problem
  - unless they amount to a significant fraction.
- Hence
  - YOU CAN AFFORD TO BE CONSERVATIVE.
- Errors are assumed to be unbiased.
  - Systematic errors of any significant size can cause filter divergence.

Outline

- 9.3 Simultaneous Localization and Mapping
  - 9.3.1 Introduction
  - 9.3.2 Global Consistency in Cyclic Maps
  - 9.3.3 Revisiting
  - 9.3.4 EKF SLAM for Discrete Landmarks
  - 9.3.5 Example: Auto surveying of Laser Reflectors
  - Summary

Summary

- A spectrum of degrees of quality exists for maps in terms of metric accuracy.
  - Self consistency and external consistency are the two highest.
- State vector consistency is only an issue if there are more states than the degrees of freedom of the system.
  - For sparse systems, state consistency can be enforced very efficiently.
- Its easy to do it all automatically - except for one thing - the revisiting problem.
- Using these techniques a sparse system with 30,000 degrees of freedom can be rendered consistent in a few seconds.

Summary

- SLAM is an ambitious problem to tackle but some instances are harder than others.
- The amount and quality of initial information matters a lot.
- The degree of constraint generated by a single sensor reading matters a lot.
- Basically, its shape-from-motion. It cannot determine absolute location.