Autonomous Navigation and Perception for Aerial Vehicles

Assistant Prof. Vadim Indelman

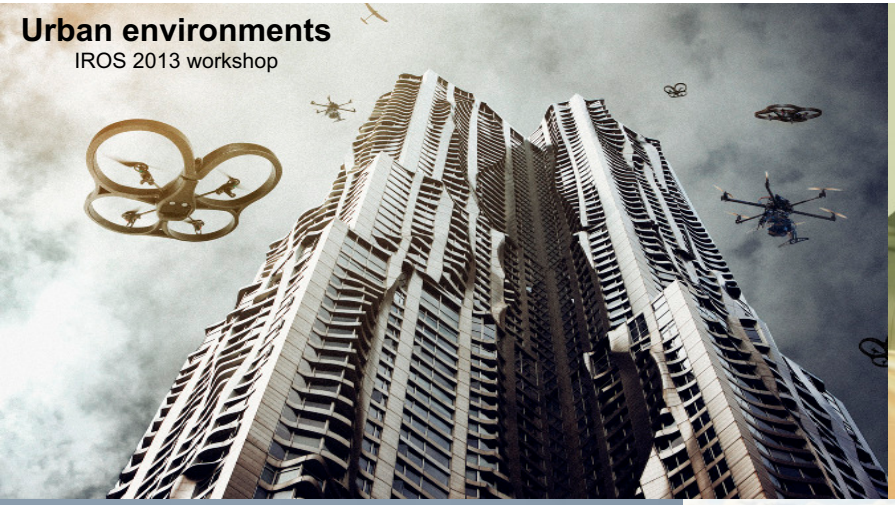
TECHNION Israel Institute of Technology

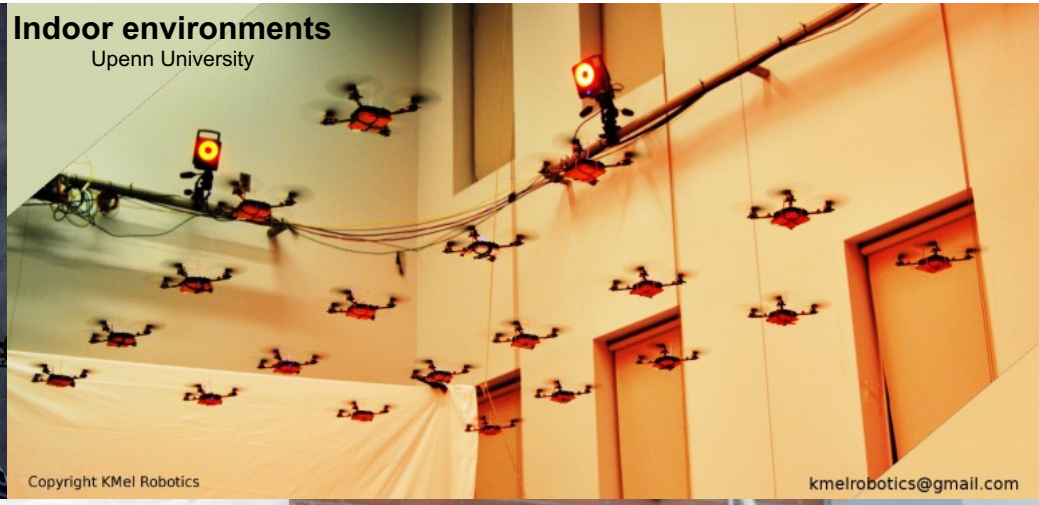
ANPL | Autonomous Navigation & Perception Lab

IBM Research Cognitive Computing Colloquium on Computer Vision and Video Technologies, November 2016

Aerial Autonomy

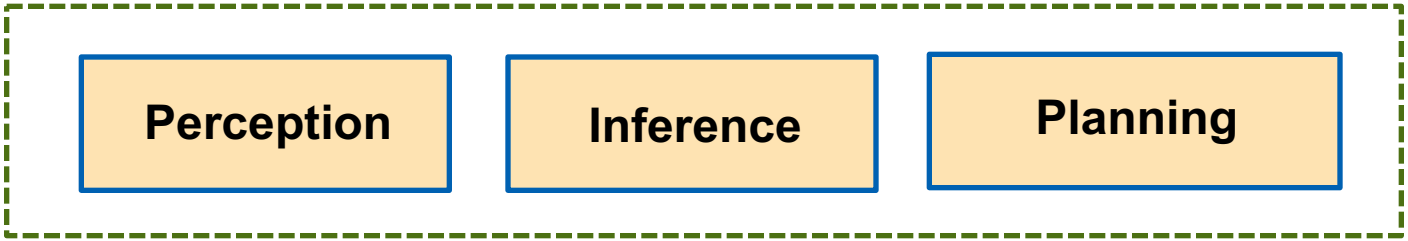
- Autonomous operation in complex, uncertain and ambiguous scenarios
  - **Independent of infrastructure and prior information**
  - **Within static and dynamic environments**





Aerial Autonomy

- Key required components
  - **Perception & inference**: Where am I? What is the surrounding environment?
  - **Planning**: What to do next?



- Required online
- Additional complexity for multiple robots

Perception and Inference

- Where am I?
- What is the surrounding environment?

Perception

Inference

- What if neither is known?
  - Need to estimate **both** robot pose and observed structure
  - Robotics: simultaneous localization and mapping (SLAM)
  - Computer vision: structure from motion, bundle adjustment

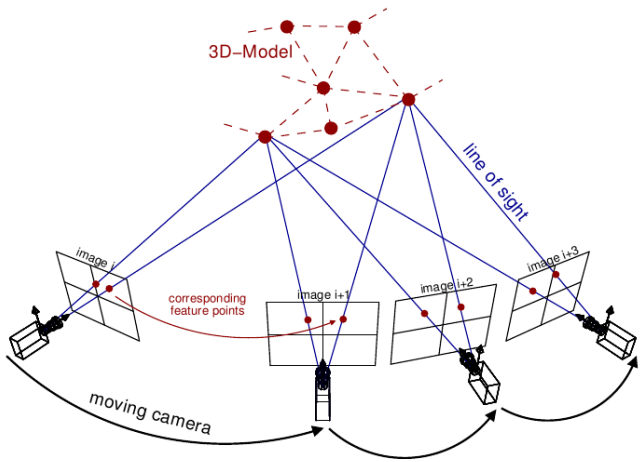


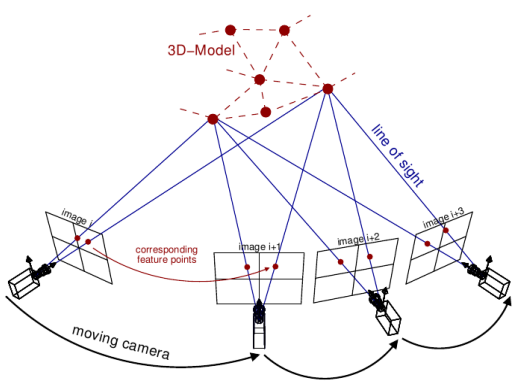
Image from: http://www.tnt.uni-hannover.de/project/motionestimation

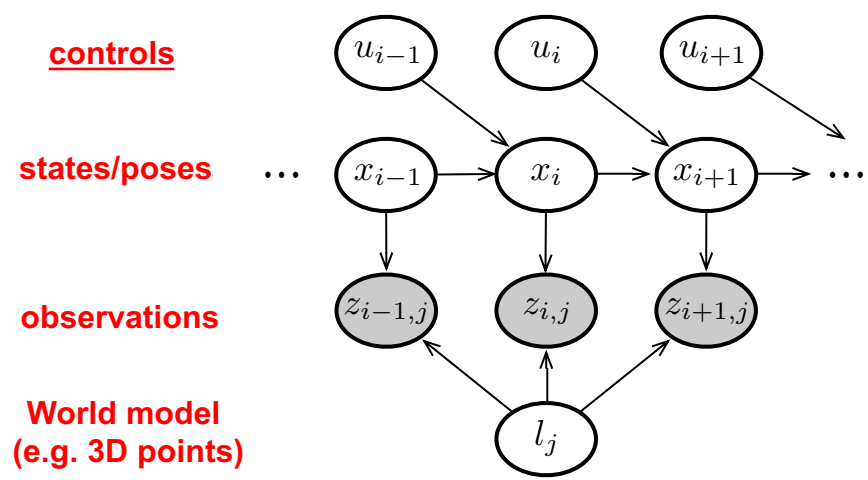
Notations and Formulation

- Joint state vector $$X_k \doteq \{\underbrace{x_0, \ldots, x_k}_{\text{Past \& current robot states}}, \underbrace{L_k}_{\text{Mapped environment}}\}$$

- Joint probability distribution function $p\left(X_k | \mathcal{Z}_k, \mathcal{U}_{k-1}\right)$

$$p\left(X_k | \mathcal{Z}_k, \mathcal{U}_{k-1}\right) = priors \cdot \prod_{i=1}^{k} p\left(x_i | x_{i-1}, u_{i-1}\right) \underbrace{p\left(z_i | X_i^o\right)}_{\text{General observation model } X_i^o \subseteq X_i}$$





Notations and Formulation

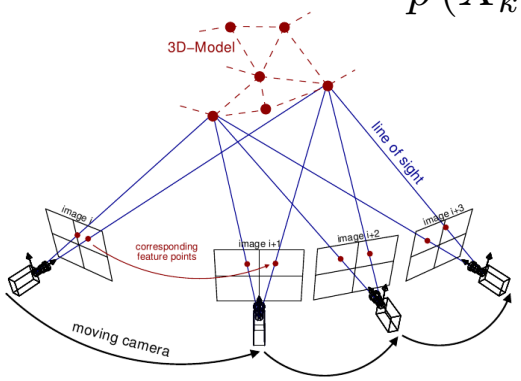
- Joint state vector $$X_k \doteq \{\underbrace{x_0, \ldots, x_k}_{\text{Past \& current robot states}}, \underbrace{L_k}_{\text{Mapped environment}}\}$$

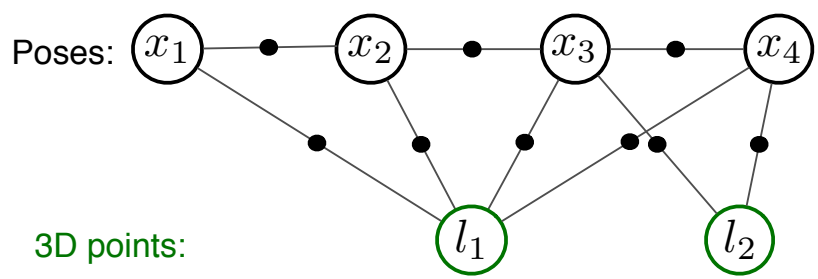
- Joint probability distribution function $p\left(X_k | \mathcal{Z}_k, \mathcal{U}_{k-1}\right)$

$$p\left(X_k | \mathcal{Z}_k, \mathcal{U}_{k-1}\right) = priors \cdot \prod_{i=1}^{k} p\left(x_i | x_{i-1}, u_{i-1}\right) \underbrace{p\left(z_i | X_i^o\right)}_{\text{General observation model } X_i^o \subseteq X_i}$$

- Computationally-efficient maximum a posteriori inference e.g. [Kaess et al. 2012]

$$p\left(X_k | \mathcal{Z}_k, \mathcal{U}_{k-1}\right) \sim N\left(X_k^*, \Sigma_k\right)$$

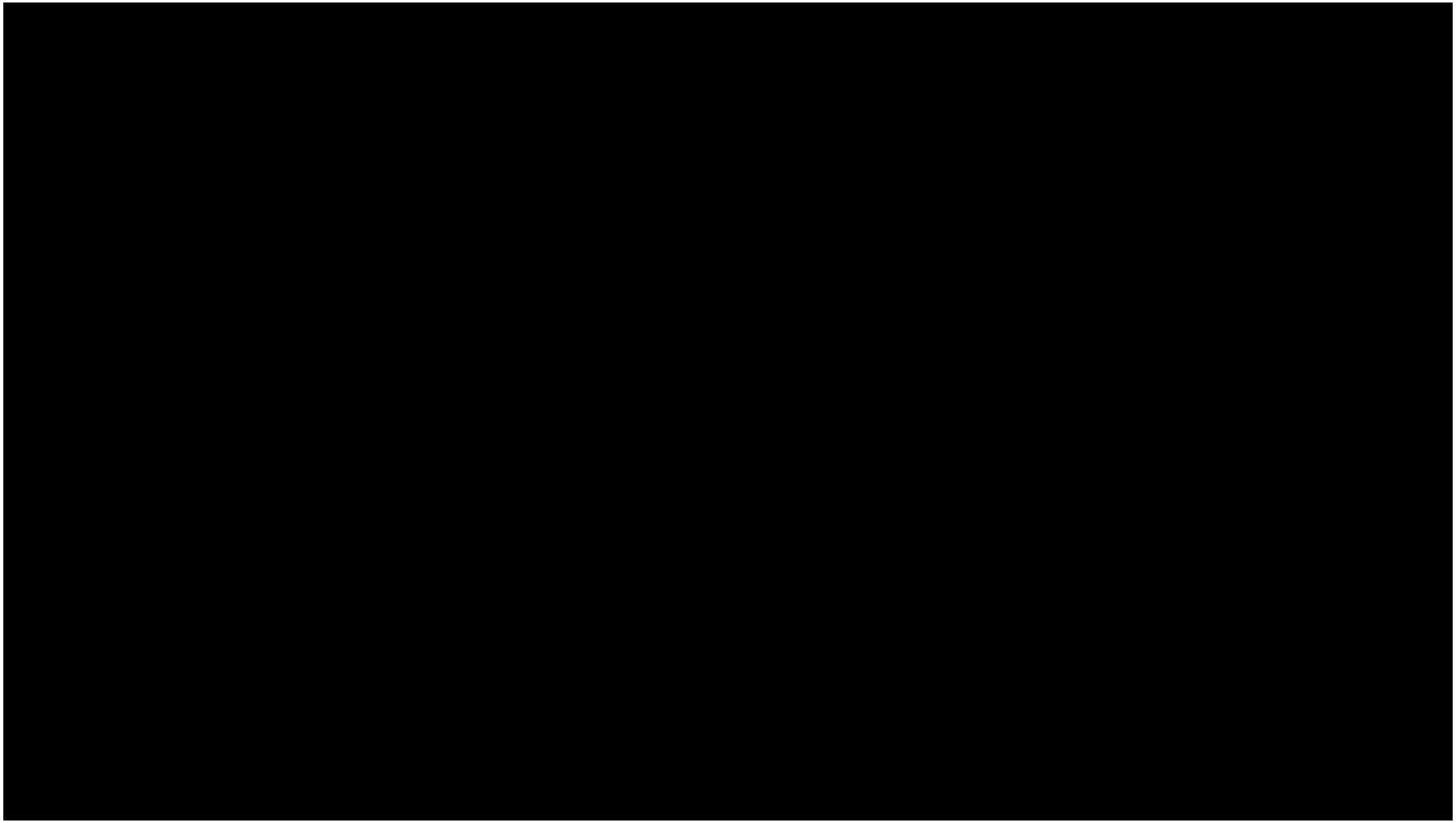




Vision Aided Quadrotor Navigation

MARS Lab, University of Minnesota

RGBD Mapping

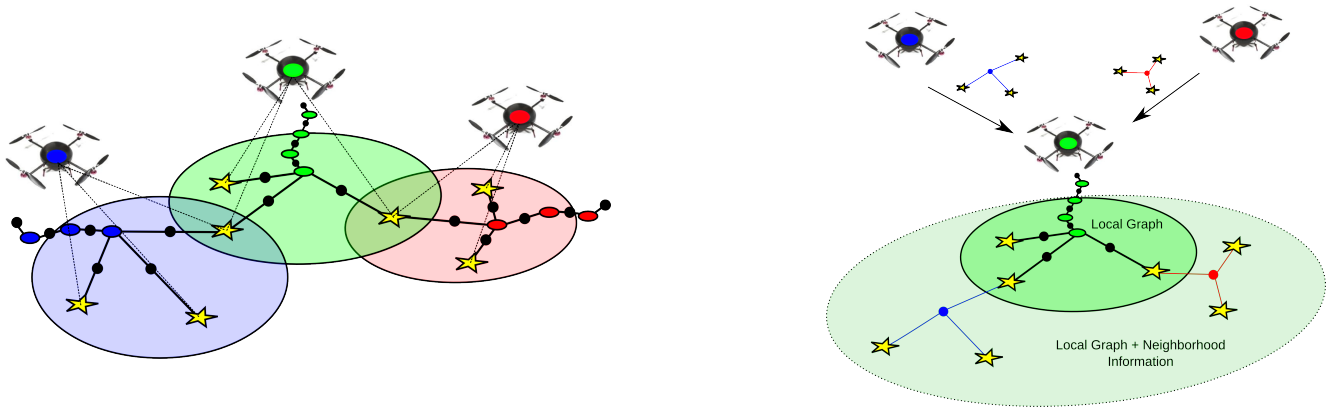


ANPL, Technion

Collaborative Navigation & Mapping

Collaborative Inference & Perception - Why?

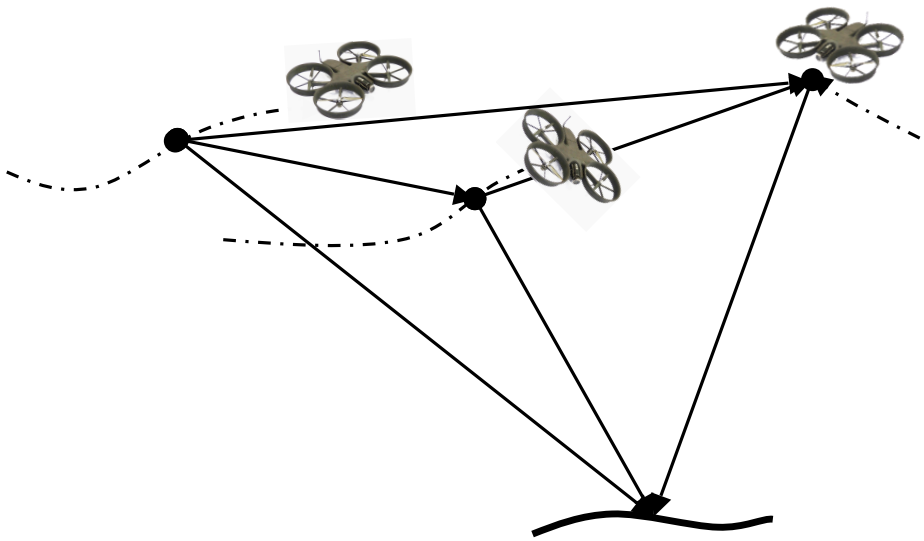
- Key capability:
  - By sharing information between robots and formulating multi-robot constraints, performance of individuals in the group can be greatly improved
  - Additional advantages, according to application (e.g. mapping - extend sensing horizon)



Images from “DDF-SAM 2.0: Consistent Distributed Smoothing and Mapping”, *ICRA*, 2013

Distributed Multi-View Geometry

- Exploit mutually observed scenes:
  - Share image observations and appropriate navigation information
  - Distributed information fusion using **multi-view geometry constraints**



[Indelman et al., IJRR 2012]

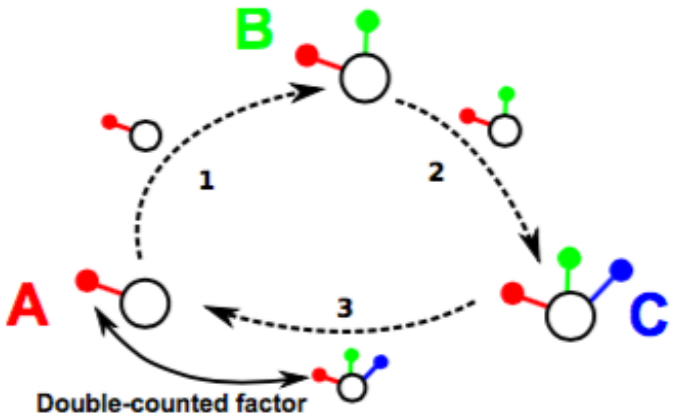
Challenges

- How to guarantee **consistent** information fusion?
- How to determine **data association**?
- How to determine next actions **autonomously & online**?

Consistent Decentralized Estimation

- Intuitive example:
  - Consider 3 robots: **A**,**B** and **C**, and a cyclic communication
  - Each robot estimates the variable x based on available data
  - Assume **A** transmits to **B** message $p(x|Z_A)$
  - **B** then passes to **C** the message $p\left(x|Z_B, Z_A\right)$
  - **C** sends to **A** the message $p\left(x|Z_C, Z_B, Z_A\right)$

- If **A** treats $p\left(x|Z_C, Z_B, Z_A\right)$ as independent wrt its local belief - it will double count information



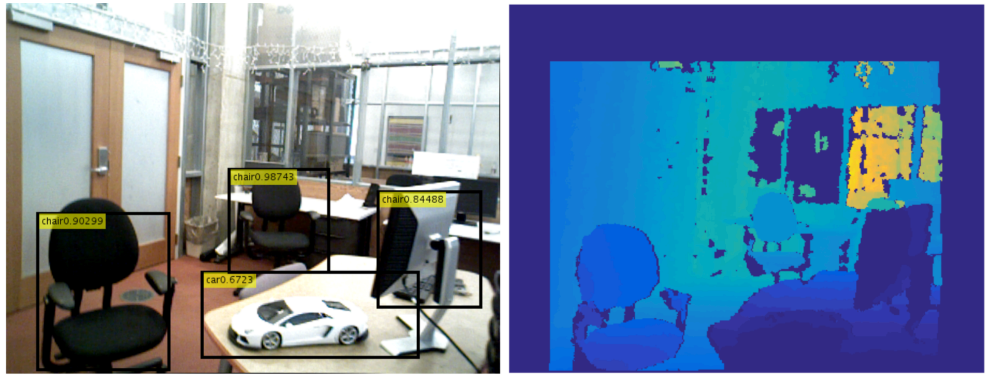
Images from "DDF-SAM 2.0: Consistent Distributed Smoothing and Mapping", *ICRA*, 2013

Challenges

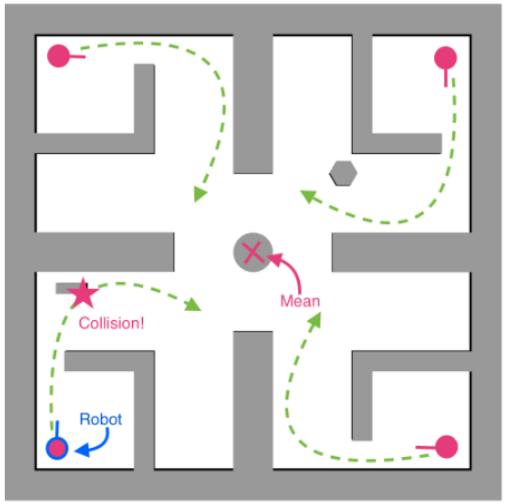
- How to guarantee **consistent** information fusion?
- How to determine **data association**?
  - Example: two matched images, same scene?
- How to determine next actions **autonomously & online**?

Ambiguous Environments

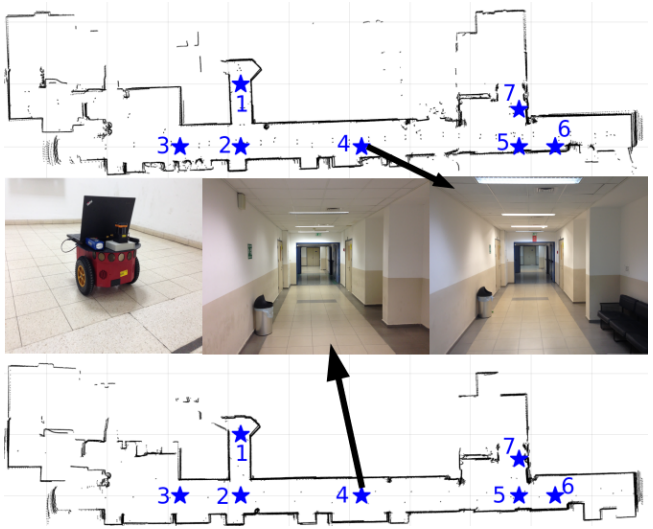
Angeli et al., TRO'08



Mu et al., IROS'16



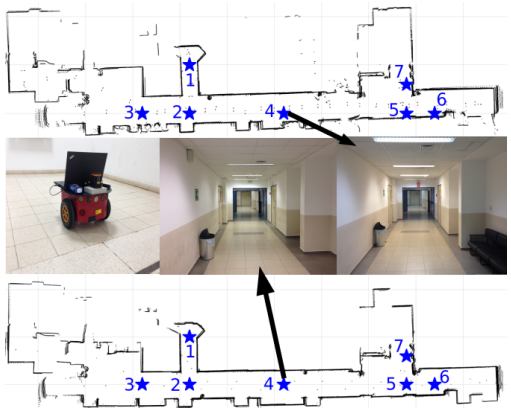
Agarwal et al., arXiv 2015



Pathak et al., arXiv 2016

Ambiguous Environments

- Ambiguity can be due to (combination of):
  - Perceptually aliased scenes (two similar objects)
  - Limited/imperfect sensing (limited sensing range)

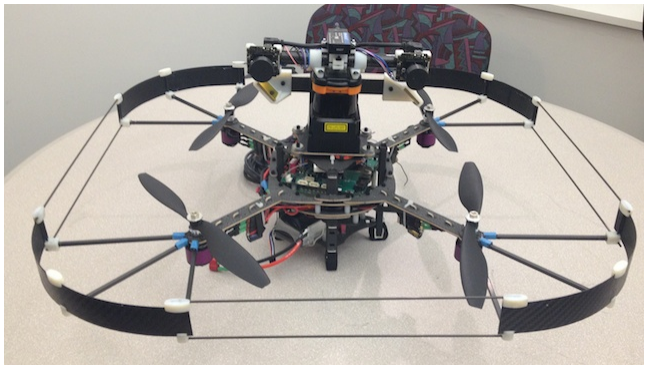


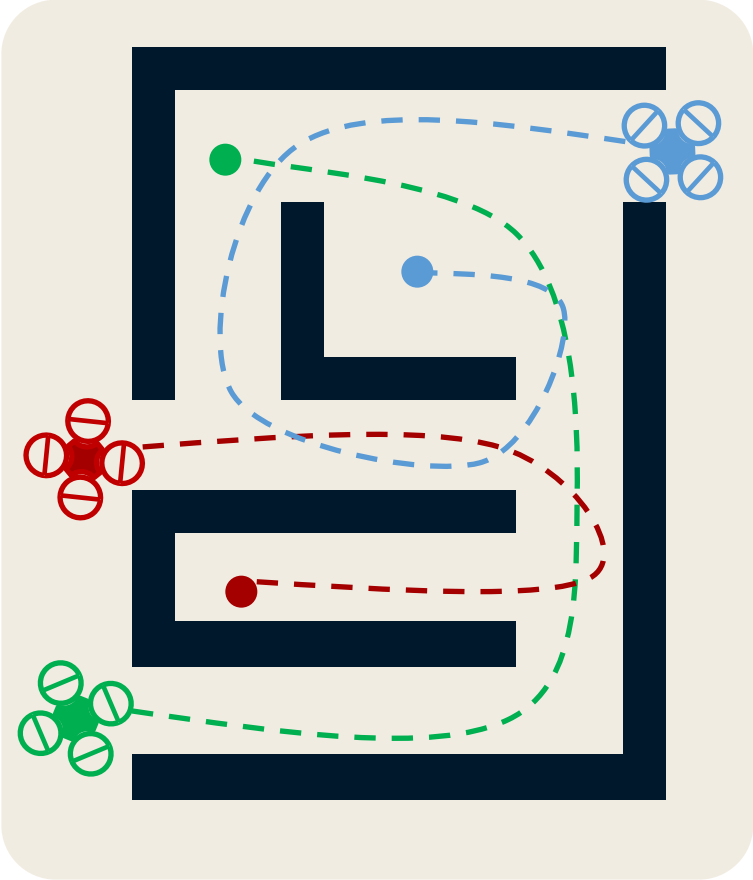
- Data association and SLAM/localization are inherently coupled
- Can we incorproate these aspects within **inference** and **decision making**?

Robust Perception

Possible Application

- Robots scattered in an unknown **perceptually aliased** environment
- Initially unaware of each others' location
- How to establish collaboration?
  - Unknown multi-robot data association
  - Unknown initial relative poses between robots





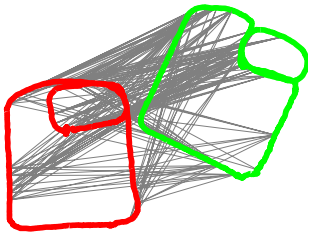
Robust Collaborative Perception

- Multi-robot collaborative localization and mapping – **distributed, online**
- **Robust perception** (cope with perceptual aliasing)

Indelman et al., CSM 2016

Probabilistic Formulation

- Notations:
  - $\mathcal{F}$: Multi-robot correspondences set
  - $\mathcal{J}$: Latent variables to indicate inliers/outliers



- Joint pdf over robot trajectories **and** multi-robot data association:

$$p\left(X,\mathcal{J}|Z\right) \propto \prod_{r} p\left(X^{r}|Z^{r}\right) \prod_{(r_1,r_2,k,l)\in\mathcal{F}} p\left(j_{k,l}^{r_1,r_2}\right) p\left(u_{k,l}^{r_1,r_2}|x_k^{r_1},x_l^{r_2},j_{k,l}^{r_1,r_2}\right)$$

Only local measurements

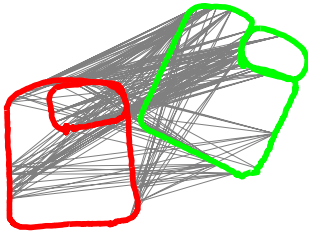
Data association

Multi-robot measurement likelihood, given data association

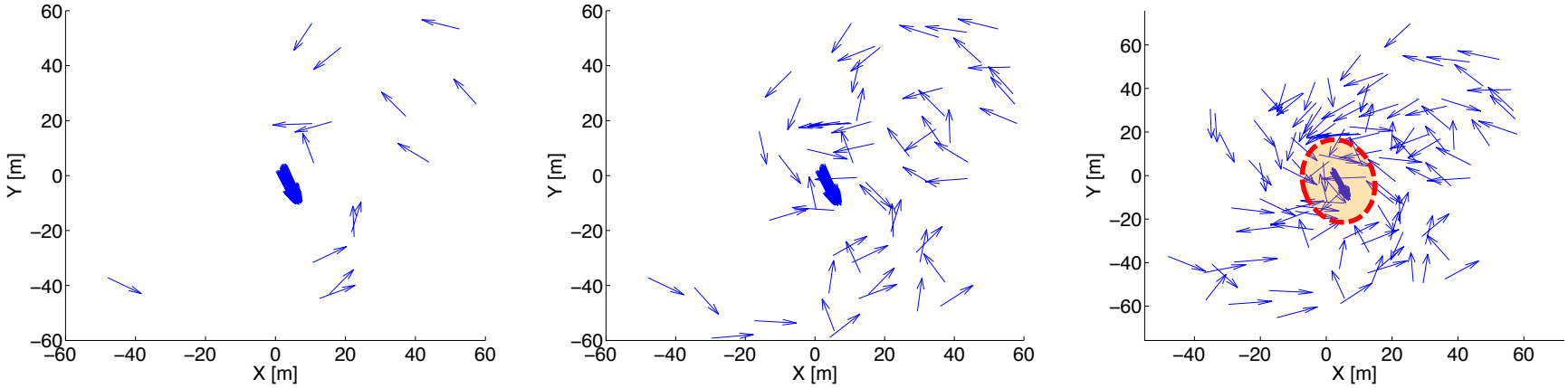
Each multi-robot correspondence

Key Observation

- Given robot local trajectories, relative initial pose can be calculated from each candidate multi-robot correspondence
  - **Only** inliers produce similar transformations
  - Objective: identify cluster



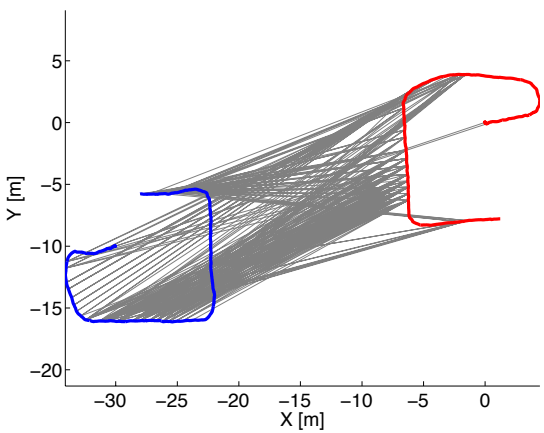
Initial relative pose between two robots (planar case: x, y, θ) [synthetic data]

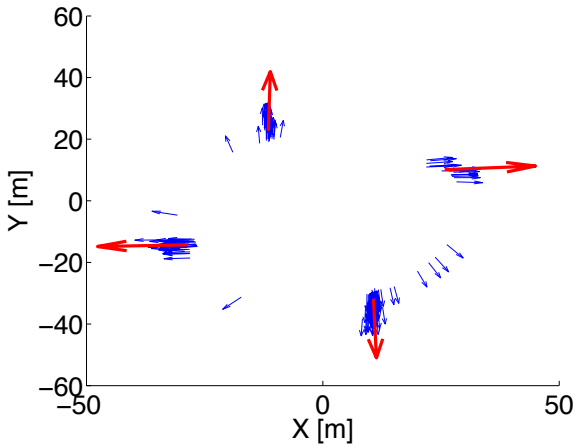


10% outliers 40% outliers 85% outliers

In Presence of Perceptual Aliasing

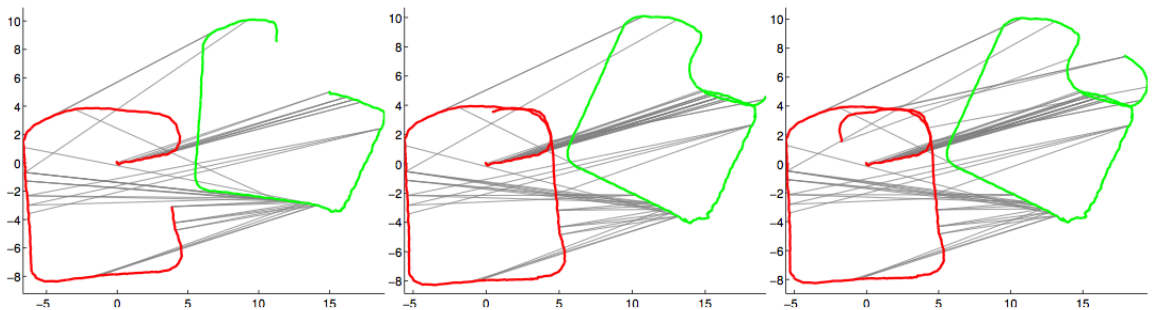
- Multiple hypotheses/clusters
- Which one to choose?



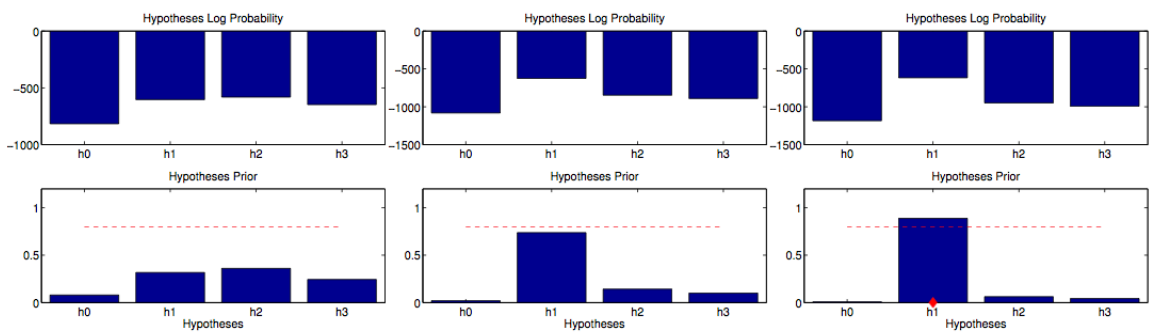


Incorporating Chinese Restaurant Process

time

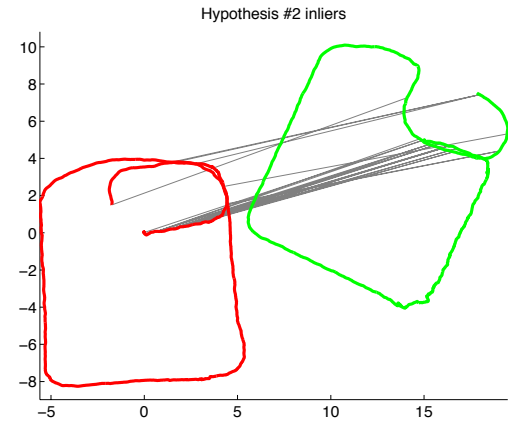


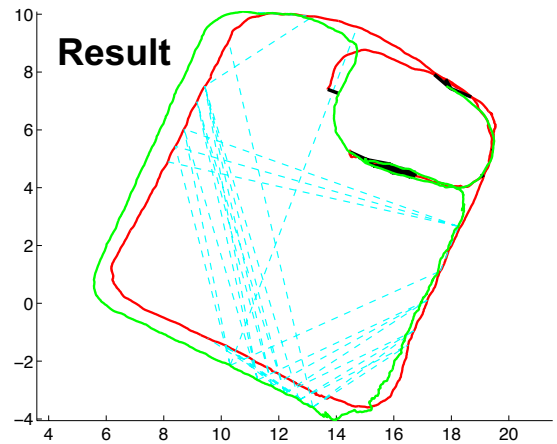
(a)

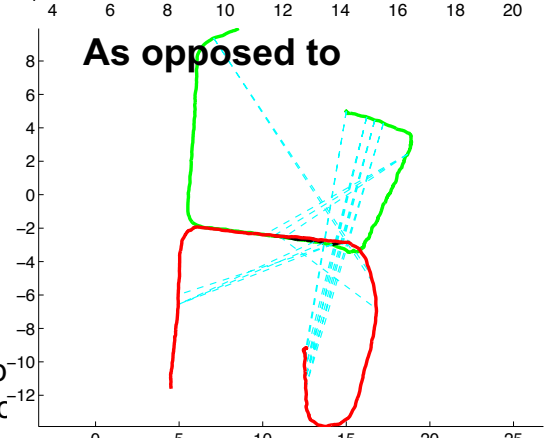


(b)

Inliers of chosen hypothesis







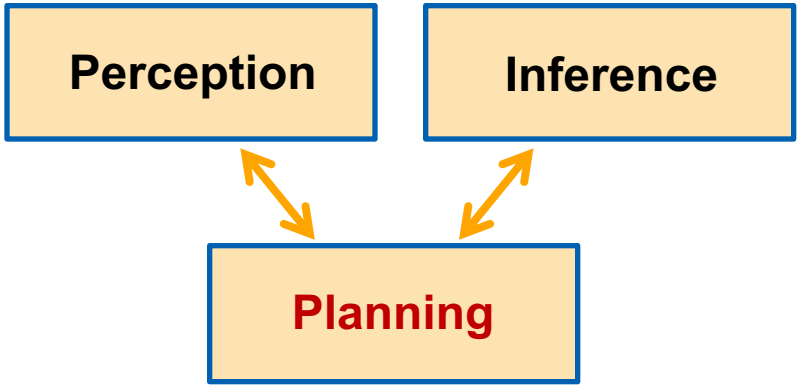
Challenges

- How to guarantee **consistent** information fusion?
- How to determine **data association**?
- How to determine next actions **autonomously & online**?

Active Robust Perception - Data Association Aware Belief Space Planning

Planning and Decision Making Under Uncertainty

- Given partial and uncertain knowledge, determine:
  - Next actions, given a task
- Also known as (or related to):
  - Active perception/vision
  - Belief space planning
  - Autonomous navigation

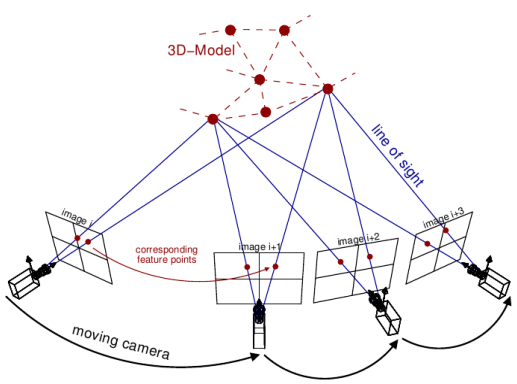


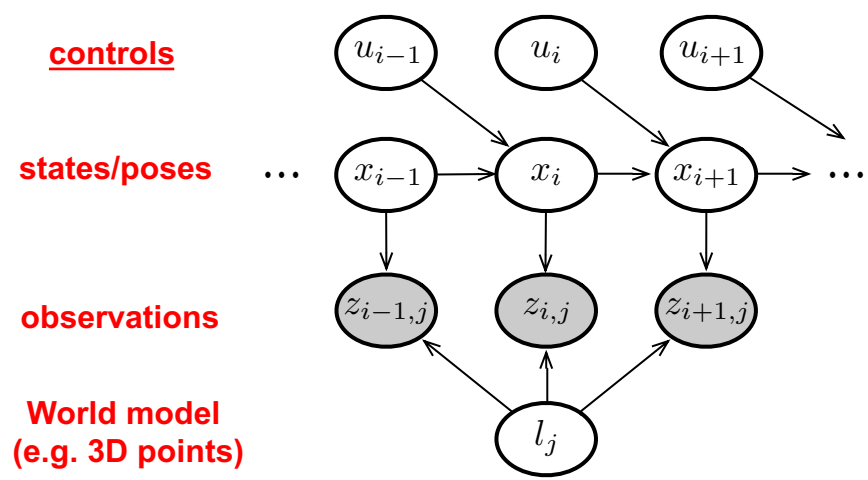
Notations and Formulation

- Joint state vector $X_k \doteq \{\underbrace{x_0, \ldots, x_k}_{\text{Past \& current robot states}}, \underbrace{L_k}_{\text{Mapped environment}}\}$

- Joint probability distribution function $p\left(X_k | \mathcal{Z}_k, \mathcal{U}_{k-1}\right)$

$$p\left(X_k | \mathcal{Z}_k, \mathcal{U}_{k-1}\right) = priors \cdot \prod_{i=1}^{k} p\left(x_i | x_{i-1}, u_{i-1}\right) \underbrace{p\left(z_i | X_i^o\right)}_{\text{General observation model } X_i^o \subseteq X_i}$$





Belief Space Planning

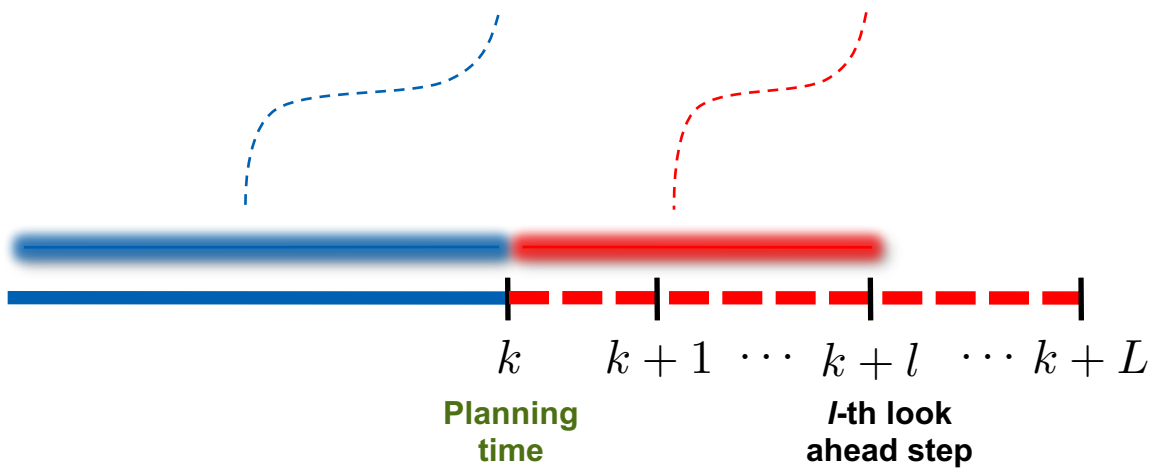
- Belief at the l-th look-ahead step
  - Describes the joint pdf (e.g. robot/camera poses and environment states) at that future time

$$b[X_{k+l}] \doteq \mathbb{P}(X_{k+l} | \mathcal{Z}_k, \mathcal{U}_{k-1}, Z_{k+1:k+l}, u_{k:k+l-1})$$

Joint state at the *l*-th look ahead step

Past controls & measurements

Actions & measurements at the first *l* look-ahead steps



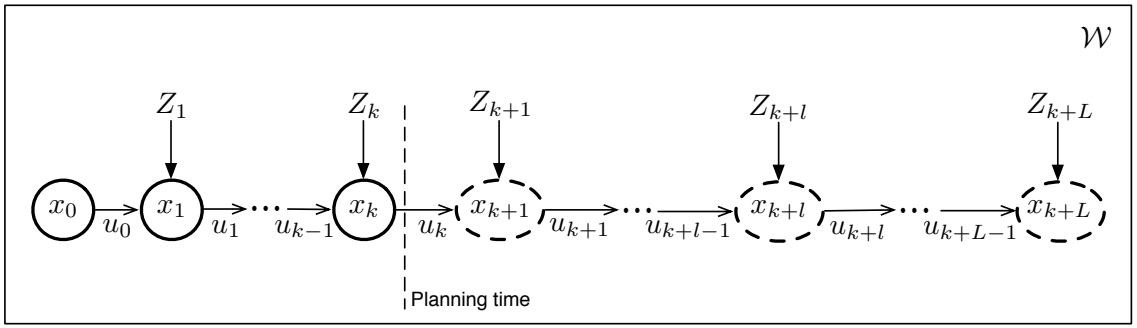
Belief Space Planning

- Belief at the l-th look-ahead step
  - Describes the joint pdf (e.g. robot/camera poses and environment states) at that future time

$$b[X_{k+l}] \doteq \mathbb{P}(X_{k+l} | \mathcal{Z}_k, \mathcal{U}_{k-1}, Z_{k+1:k+l}, u_{k:k+l-1})$$

- Objective function can now involve **uncertainty** aspects

$$J_k\left(u_{k:k+L-1}\right) \doteq \underset{Z_{k+1:k+L}}{\mathbb{E}} \left\{ \sum_{l=0}^{L-1} c_l\left(b\left[X_{k+l}\right], u_{k+l}\right) + c_L\left(b\left[X_{k+L}\right]\right) \right\}$$



Active Robust Perception

- Incorporate data association (DA) aspects within belief space planning
- Can be used for active disambiguation (for example)
- Relax common assumption in belief space planning that DA is given and perfect

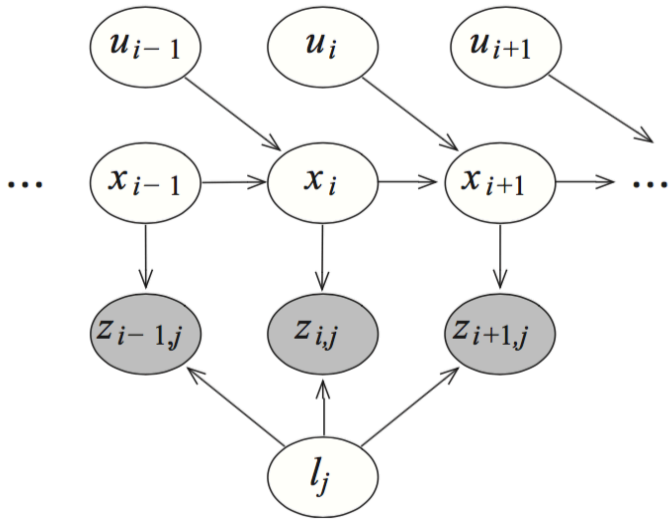


Image from Indelman et al., IJRR'15

Active Robust Perception

- Intuition

Distinct scenes

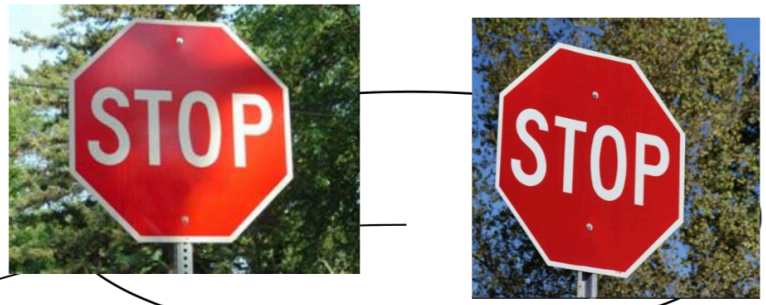
$p(x_{k+1}|z_{0:k}, u_{0:k})$

$p(x_k|z_{0:k}, u_{0:k-1})$

Active Robust Perception

- Intuition

Perceptually aliased scenes

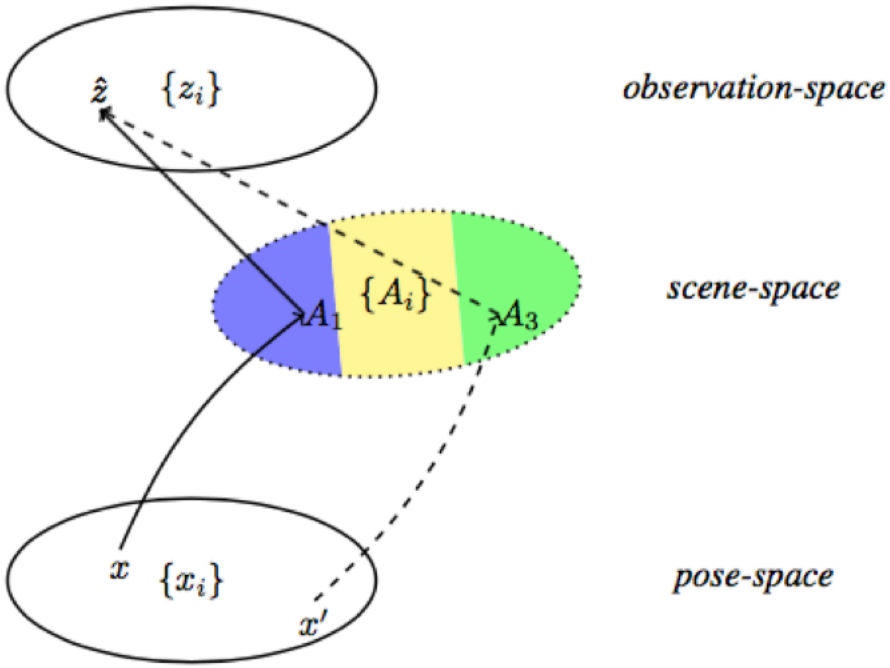


$p(x_{k+1}|z_{0:k}, u_{0:k})$

$p(x_k|z_{0:k}, u_{0:k-1})$

Active Robust Perception

- In presence of perceptual aliasing, the **same observation** could be obtained from **different poses** viewing **different scenes**



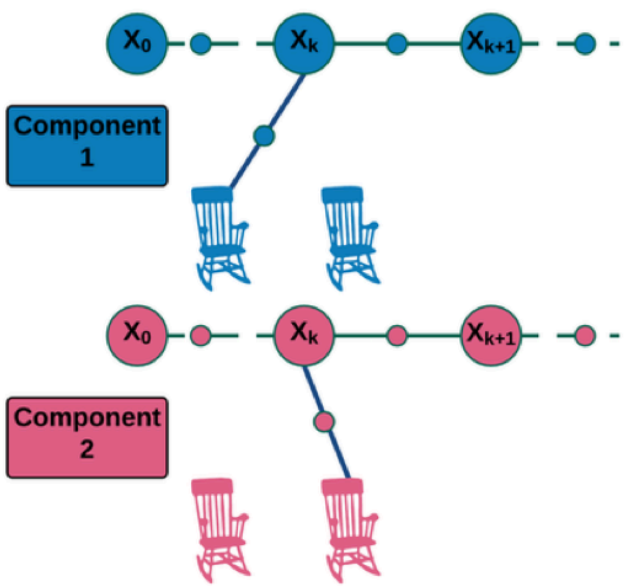
How to capture this fact within belief space planning?

Active Robust Perception

- Belief is represented by a Gaussian Mixture Model (GMM)

$$b[X_k] = \mathbb{P}(X_k|\mathcal{H}_k) = \sum_{j=1}^{M_k} \xi_k^j \mathbb{P}(X_k|\mathcal{H}_k, \gamma = j)$$

Weight — **Conditional Gaussian, represented by a factor graph**



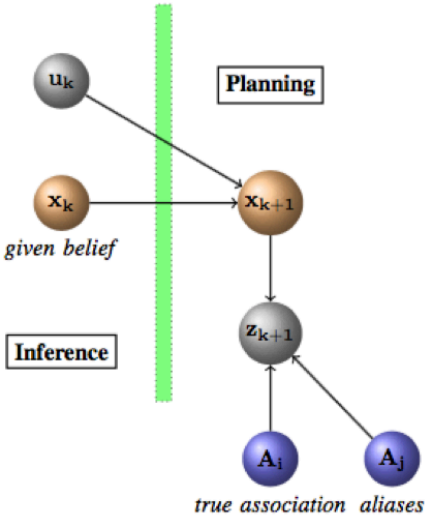
Pathak et al., submitted

Active Robust Perception

- Belief is represented by a Gaussian Mixture Model (GMM)

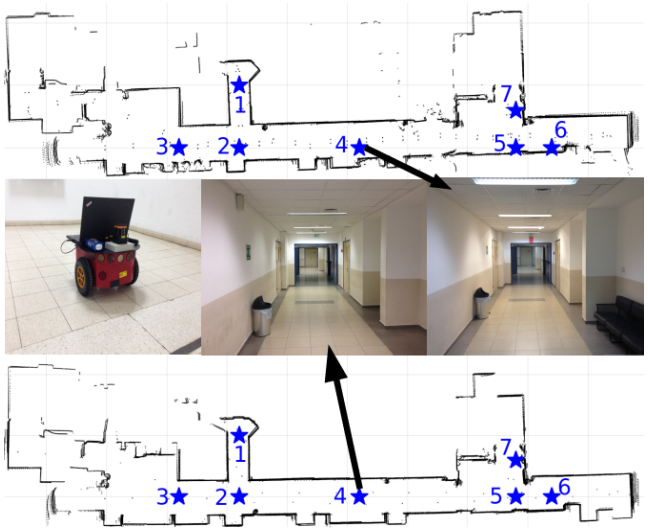
$$b[X_k] = \mathbb{P}(X_k|\mathcal{H}_k) = \sum_{j=1}^{M_k} \underbrace{\xi_k^j}_{\text{Weight}} \underbrace{\mathbb{P}(X_k|\mathcal{H}_k, \gamma = j)}_{\text{Conditional Gaussian, represented by a factor graph}}$$

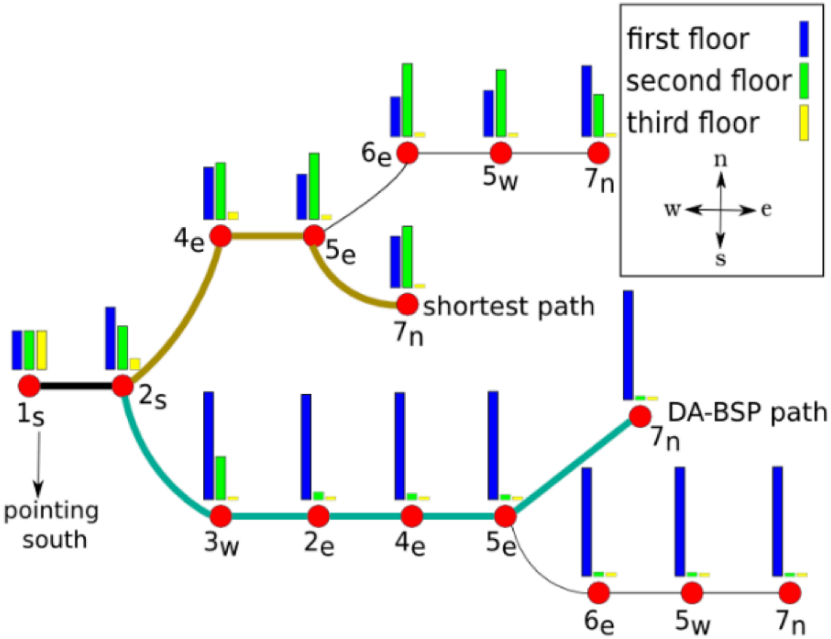
- **Main idea**: Reason how GMM belief will evolve for different candidate actions
- Number of modes can go down, and go up (!)



Active Robust Perception

- Multi-floor aliased environment, kidnapped robot problem





Pathak et al., arXiv 2016

Summary

- Autonomous Navigation and Perception
  - Collaborative multi-robot inference
  - Robustness to ambiguity and perceptual aliasing
  - Active robust perception via belief space planning