DA-BSP: Towards Data Association Aware Belief Space Planning for Robust Active Perception

Shashank Pathak and **Antony Thomas** and **Asaf Feniger** and **Vadim Indelman**[1]

Belief space planning (BSP) and decision-making under uncertainty are fundamental problems in robotics and artificial intelligence, with applications including autonomous navigation, object grasping and manipulation, active SLAM, and robotic surgery. In the presence of uncertainty, such as in robot motion and sensing, the true state of variables of interest (e.g. robot poses), is unknown and can only be represented by a probability distribution over possible states, given available data. This distribution, the belief space, is inferred using probabilistic approaches based on incoming sensor observations and prior knowledge. The corresponding BSP problem is an instantiation of a partially observable Markov decision problem (POMDP) [4].

Existing BSP approaches (e.g. [2,5,9,11]) typically assume data association to be given and perfect, i.e. the robot is assumed to correctly perceive the environment to be observed by its sensors, given a candidate action. However, this assumption, denoted for brevity as `DAS`, can be harder to justify while operating in ambiguous and perceptually aliased environments (see Figure 1), and in the presence of different sources of uncertainty (uncertainty due to stochastic control and imperfect sensing).

Indeed, in the presence of ambiguity, `DAS` may lead to incorrect posterior beliefs and as a result, to sub-optimal actions. More advanced approaches are thus required to enable reliable operation in ambiguous conditions, approaches often referred to as (active) robust perception. Yet, existing robust perception approaches (e.g. [1, 3, 6, 10]) focus on the passive case, where robot actions are externally determined and given.

In this work we develop a general data association aware belief space planning (DA-BSP) framework capable of better handling complexities arising in a real world, possibly perceptually aliased, scenarios. We rigorously incorporate reasoning about data association within belief space planning (and inference), while also considering other sources of uncertainty (motion, sensing and environment). In particular, we show that due to perceptual aliasing, the posterior belief becomes a mixture of probability distribution functions, and design cost functions that measure the expected level of ambiguity and posterior uncertainty. Using these and standard costs (e.g. control penalty, distance to goal) within the objective function, yields a general framework that reliably represents action impact, and in particular, capable of active disambiguation. Our approach is thus applicable to robust active perception and autonomous navigation in perceptually aliased environments. In this short paper, we provide a concise overview of the DA-BSP approach, referring the interested reader to [7, 8] for full details.

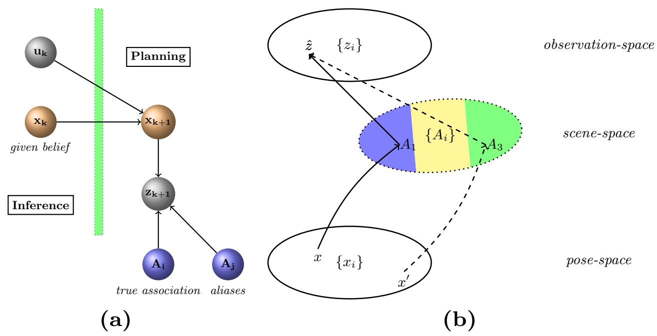

Figure 1: (a) Generative graphical model. Standard BSP approaches assume data association (DA) is given and perfect (`DAS`). We incorporate data association aspects within BSP and thus can reason about ambiguity (e.g. perceptual aliasing) at a decision-making level. (b) Schematic representation of pose, scene and observation spaces. Scenes A_1 and A_3 when viewed from perspective x and x' respectively, produce the same nominal observation $\hat{z}$, giving rise to *perceptual aliasing*.

Concept and Approach Overview

Given some candidate action u_k and the belief at planning time k, we can reason about a future observation z_{k+1} (e.g. an image) to be obtained once this action is executed; its actual value is unknown. All the possible values such an observation can assume should be thus taken into account while evaluating the objective function, which can be written as:

$$J(u_k) \doteq \int_{z_{k+1}} \overbrace{\mathbb{P}(z_{k+1} \mid \mathcal{H}_{k+1}^{-})}^{(a)} \; c \left(\overbrace{\mathbb{P}(X_{k+1} | \mathcal{H}_{k+1}^{-}, z_{k+1})}^{(b)} \right), \quad (1)$$

where X_{k+1} denotes the past and current robot poses $X_k \doteq \{x_0, \ldots, x_k\}$, and $\mathcal{H}_{k+1}^{-} \doteq \{u_{0:k}, Z_{0:k}\}$.

The two terms (a) and (b) in Eq. (1) have intuitive meaning: for each considered value of z_{k+1}, (a) represents how likely is it to get such an observation when both the history $\mathcal{H}$ and control u_k are known, while (b) corresponds to the posterior belief *given* this specific z_{k+1}.

Existing BSP approaches typically consider data association is solved (`DAS`), i.e. given and perfect. In other words, `DAS`

[1]Department of Aerospace Engineering, Technion - Israel Institute of Technology, Israel, emails: {shashank, antony.thomas, asaff, vadim.indelman}@technion.ac.il. This work was partially supported by the Israel Science Foundation (ISF) and the Technion Autonomous Systems Program (TASP).

means we can correctly associate each possible measurement z_{k+1} with the corresponding scene A_i it captures and write the corresponding measurement likelihood term $\mathbb{P}(z_{k+1}|x_{k+1}, A_i)$. Yet, it is unknown from what future robot pose x_{k+1} the actual observation z_{k+1} will be acquired, since the *actual* robot pose x_k at time k is unknown, the control is stochastic and sensing is imperfect. In inference, we have a similar situation with the key difference that the observation z has been acquired.

Rather than assuming DAS, in this work we incorporate within BSP (and similarly within inference) reasoning about possible scenes or objects that the future observation z_{k+1} could be generated from, see Figure 1. While this may seem computationally expensive, realistic scenarios typically exhibit *parsimonious data association*: If the environment has only distinct scenes or objects, then for each specific value of z_{k+1}, there will be only one scene A_i that can generate such an observation. In the case of perceptually aliased environments, there could be several other scenes (or objects) that are either completely identical or have a similar visual appearance when observed from appropriate viewpoints. They could equally well explain the considered observation z_{k+1}. Thus, there are several possible associations $\{A_i\}$ and due to localization uncertainty determining which association is the correct one is not trivial. As we show in [7], in these cases the posterior (term (b) in Eq. (1)) becomes a Gaussian mixture with appropriate weights that we rigorously compute. Additionally, the weight updates are capable of discriminating against unlikely data-associations, during the planning steps.

We now briefly summarize how terms (a) and (b) in Eq. (1) are calculated while reasoning about data association, referring the reader to [7] for full details.

Computing the term (a): $\mathbb{P}(z_{k+1}|\mathcal{H}^{-}_{k+1})$: Applying total probability over non-overlapping scene space $\{A_{\mathbb{N}}\}$ and marginalizing over all possible robot poses, yields

$$\mathbb{P}(z_{k+1}|\mathcal{H}^{-}_{k+1}) \equiv \sum_{i}^{|A_{\mathbb{N}}|} \int_x \mathbb{P}(z_{k+1}, x, A_i \,|\, \mathcal{H}^{-}_{k+1}) \doteq \sum_{i}^{|A_{\mathbb{N}}|} w^i_{k+1}. \quad (2)$$

As seen from the above equation, to calculate the likelihood of obtaining some observation z_{k+1}, we consider separately, for each scene $A_i \in \{A_{\mathbb{N}}\}$, the likelihood that this observation was generated by scene A_i. This probability is captured for each scene A_i by a corresponding weight w^i_{k+1}; these weights are then summed to get the actual likelihood of observation z_{k+1}. As shown in [7], these weights naturally account for perceptual aliasing aspects for each considered z_{k+1}.

In practice, instead of considering the entire scene space $\{A_{\mathbb{N}}\}$ that could be computationally costly, the availability of the belief from the previous time step, $b[X^{-}_{k+1}]$, enables us to consider only those scenes that could be actually observed from the viewpoints with non-negligible probability according to $b[X^{-}_{k+1}]$.

Computing the term (b) : $\mathbb{P}(X_{k+1}|\mathcal{H}^{-}_{k+1}, z_{k+1})$: The term (b), $\mathbb{P}(X_{k+1}|\mathcal{H}^{-}_{k+1}, z_{k+1})$, represents the posterior probability conditioned on observation z_{k+1}. This term can be similarly calculated, with a key difference: since the observation z_{k+1} is given, it must have been generated by *one* specific (but unknown) scene A_i according to an appropriate measurement model. Hence, also here, we consider all possible such scenes and weight them accordingly, with weights $\tilde{w}^i_{k+1}$ representing the probability of each scene A_i to have generated the observation z_{k+1}.

As shown in [7], the term (b) in Eq. (1) is a GMM with M_{k+1} components, $\mathbb{P}(X_{k+1}|\mathcal{H}^{-}_{k+1}, z_{k+1}) = \sum_{r=1}^{M_{k+1}} \xi^r_{k+1} b[X^{r+}_{k+1}]$, where $b[X^{r+}_{k+1}]$ represents the rth component of the belief, and the weights ξ^r_{k+1} are defined recursively (see full details in [7]). Interestingly, the number of components can not only go down $M_{k+1} \leq M_k$ (as a result of a partially or fully disambiguating action), but could also go up, i.e. $M_{k+1} > M_k$.

To summarize the discussion thus far, we have shown that for the myopic case, the objective function (1) can be re-written as

$$J(u_k) = \int_{z_{k+1}} (\sum_{i}^{|A_{\mathbb{N}}|} w^i_{k+1}) \cdot c \left(\sum_{r}^{M_{k+1}} \xi^r_{k+1} b[X^{r+}_{k+1}] \right). \quad (3)$$

In [7], we present the other ingredients of our approach, including sampling-based simulation of future observations $\{z_{k+1}\}$ given $b[X^{-}_{k+1}]$, and the design of suitable cost functions to quantify ambiguity level. We also show that DA-BSP considers data-association parsimoniously and a simple thresholding is enough for a scalable application of data-association aware belief space planning, and demonstrate key aspects basic and realistic simulations. Potential directions for future research include extension to non-myopic planning as well as proving the general theoretical properties of DA-BSP.

References

[1] L. Carlone, A. Censi, and F. Dellaert. Selecting good measurements via l1 relaxation: A convex approach for robust estimation over graphs. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, pages 2667–2674, 2014.

[2] V. Indelman, L. Carlone, and F. Dellaert. Planning in the continuous domain: a generalized belief space approach for autonomous navigation in unknown environments. *Intl. J. of Robotics Research*, 34(7):849–882, 2015.

[3] V. Indelman, E. Nelson, J. Dong, N. Michael, and F. Dellaert. Incremental distributed inference from arbitrary poses and unknown data association: Using collaborating robots to establish a common reference. *IEEE Control Systems Magazine (CSM), Special Issue on Distributed Control and Estimation for Robotic Vehicle Networks*, 36(2):41–74, 2016.

[4] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1):99–134, 1998.

[5] A. Kim and R. M. Eustice. Active visual slam for robotic area coverage: Theory and experiment. *Intl. J. of Robotics Research*, 2014.

[6] E. Olson and P. Agarwal. Inference on networks of mixtures for robust robot mapping. *Intl. J. of Robotics Research*, 32(7):826–840, 2013.

[7] S. Pathak, A. Thomas, A. Feniger, and V. Indelman. Robust active perception via data-association aware belief space planning. *arXiv preprint: 1606.05124*, 2016.

[8] S. Pathak, A. Thomas, A. Feniger, and V. Indelman. Towards data association aware belief space planning for robust active perception. In *AI for Long-term Autonomy, workshop in conjunction with IEEE International Conference on Robotics and Automation (ICRA)*, May 2016.

[9] S. Prentice and N. Roy. The belief roadmap: Efficient planning in belief space by factoring the covariance. *Intl. J. of Robotics Research*, 2009.

[10] N. Sunderhauf and P. Protzel. Towards a robust back-end for pose graph slam. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*, pages 1254–1261. IEEE, 2012.

[11] J. Van Den Berg, S. Patil, and R. Alterovitz. Motion planning under uncertainty using iterative local optimization in belief space. *Intl. J. of Robotics Research*, 31(11):1263–1278, 2012.