Noname manuscript No.
(will be inserted by the editor)

Grasp driven active touch

Grasp driven active touch

Claudio Zito · Marek Kopicki · Valerio Ortenzi · Jeremy L. Wyatt

Received: date / Accepted: date

Abstract Keywords Grasping · active vision

1 Introduction

Robot grasping is typically affected by uncertainty associated with the location of the object to be grasped. This is a challenging problem because it requires the robot to find collision-free trajectories that are robust in the face of such uncertainty. There are two fundamentally different approaches. First, a trajectory may be formulated that is robust to current uncertainty, but does not reason about how future information may reduce that uncertainty. Second, the robot may plan a trajectory to gather information that will reduce the uncertainty, so as to make a final grasping trajectory more reliable. Previous work typically separates these two aspects, separately planning information gathering trajectories and grasping trajectories. The two can be theoretically joined in a continuous state and action POMDP, but this leads to an infinite dimensional belief space planning problem that is hard to solve. In this paper we propose and validate a way to mix information gathering and reach to grasp trajectories. Our main insight is that to avoid the full complexity of belief state planning we can instead embed the value of information in the much lower dimensional physical space. This gives a well posed but tractable problem for reach to grasp planning under uncertainty. The specific contributions of this work are to:

1. plan information gain whilst simultaneously attempting to grasp the target object.
2. encode expected information gain by warping distances in the workspace, creating a non-Euclidean metric that is information sensitive.
3. employ a hierarchical planning approach to reduce planning complexity in this space.
4. update the belief about the object's pose using a tactile observation model for a multi-finger hand palpating the object.
5. evaluate the methods, proving that our approach improves reach to grasp planning for a dexterous robot.

Features of the work are that we use a state of the art grasp generation algorithm to obtain a target grasp (i.e. a set of finger contacts) on the fly, recomputing it as we update information about the object pose. One assumption is that a shape model of the object is also previously obtained from sensing, but this may be incomplete. The work is demonstrated in trials in simulation and on Boris, a half-humanoid robot platform. Empirical results confirm that sequential re-planning achieves a greater success rate than single grasp attempts, and that the information gain approach requires fewer iterations before a grasp is achieved.

In the rest of the paper we describe the problem formulation, and the core information gain planning algorithm (Section 2). We then describe the implementation on a dexterous robot platform (Section 3), and the

We gratefully acknowledge support of FP7 grant IST-600918, PacMan.

Zito, Kopicki, Ortenzi and Wyatt
CN-CR,
University of Birmingham,
Edgbaston, Birmingham
United Kingdom, B15 2TT
Tel.: +44-121-4144788
Fax: +44-121-4144281
E-mail: {claudio.zito.81,jeremy.l.wyatt}@gmail.com
E-mail: {marek.kopicki,valerio.ortenzi}@gmail.com

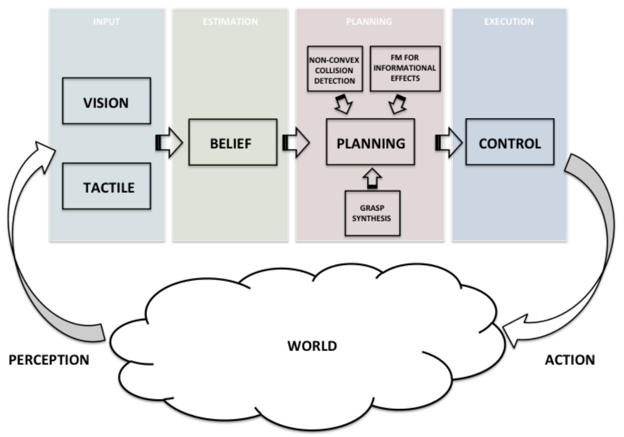

Fig. 1 The system architecture comprises sensory input, pose estimation, motion planning and active control.

experimental results (Section 4). We finish with related work (Section 5), and concluding remarks (Section 6).

2 Problem Formulation

The grasp planning problem is composed of four subproblems: state estimation, grasp synthesis, grasp planning and control. A typical approach is to represent the belief state using prior distributions (often a Gaussian), select a grasp robust in the face of uncertainty (be it pose uncertainty or shape incompleteness) and finally to use tactile feedback to adjust the grasping trajectory, see e.g. (Nikandrova et al. 2013). The reach-to-grasp trajectory is typically computed using some conventional sampling-based techniques which minimise the cost, in Euclidean space, of moving the robot's end effector to the selected grasp configuration. Comparatively little work has explored the more complex problem of reasoning about uncertainty while planning this dexterous reach-to-grasp trajectory. This is mainly due to the high dimensionality of the configuration space of a dexterous manipulator. In this work we consider precisely this problem. First we formulate the problem of reducing uncertainty in object pose, and then show how to formulate a reach to grasp problem that incorporates this measure.

2.1 Problem formulation

This section is concerned with the problem of planning control actions to reach a goal state in the presence of incomplete or noisy observations. Let us consider a discrete-time system with continuous non-linear deterministic dynamics,

$$x_{t+1} = f(x_t, u_t)$$

where $x_t \in \mathbb{R}^n$ is a configuration state of the robot and $u_t \in \mathbb{R}^l$ is a action vector, both parametrised over time $t \in \{1, 2, ...\}$. Let $p \in SE(3)$ describe the object pose, given an initial prior belief state b_1 and let us define a set of k hypotheses as $\{p^i\}_{i=1}^k$, where $p^1 = \arg\max b_1$ and $p^i \sim b_1, i \in [2, k]$.

We know that in general the problem of planning in belief space is intractable. Instead let us consider a method to search for a sequence of actions, $u_{1:T-1} = \{u_1, \dots, u_{T-1}\}$, that distinguish between observations that would occur if the object were in some pose p^1, from those that would occur in some other poses p^i, with $i \in [2, k]$. At each time step, t, the system will make an observation, $y \in \mathbb{R}^m$, that is a non-linear stochastic function of the joint state of the robot and some object state. Without losing generality, we define y_t to be a column vector of binary values. Each of these values represents whether or not a contact is observed between a given link of the robot and the object pose hypothesis p^i. However, binary values have been shown to be not very informative during the planning phase. Therefore let us define,

$$h(x, p^i) = Pr(y = 1 | x, p^i)$$

as a column vector of scores identifying the likelihood of observing a contact, $y = 1$, as a function of the joint robot and object state. More generally, let $F_t(x, u_{1:t-1})$ be the robot configuration at time t if the system begins at state x and takes action $u_{1:t-1}$. Therefore the expected sequence of observations over a trajectory, $u_{1:t-1}$, is:

$$h_t(x, u_{1:t-1}, p^i) = (h(F_2(x, u_1), p^i)^T, h(F_3(x, u_{1:2}), p^i)^T, \dots, h(F_t(x, u_{1:t-1}), p^i)^T)^T \quad (1)$$

a column vector which describes the likelihood of observing a contact at any time step of the trajectory $u_{1:t-1}$. We then need to search for a sequence of actions which maximise the difference between observations that are expected to happen in the sampled states, $p^{2:k}$, when the system is actually in the most likely hypothesis, p^1. In other words, we want to find a sequence of actions, $u_{1:T-1}$, that minimises

$$\mathcal{J}(x, u_{1:T-1}, p^{1:k}) = \sum_{i=2}^{k} N(h(x, u_{1:T-1}, p^i) | h(x, u_{1:T-1}, p^1), \mathbb{Q}) \quad (2)$$

where $N(\cdot|\mu, \Sigma)$ denotes the Gaussian distribution with mean μ and covariance Σ and $\mathbb{Q}$ is the block diagonal of the measurement noise covariance matrix. Rather than optimising (2) we follow the suggested simplifications described in (Platt et al. 2001) by dropping the normalisation factor in the Gaussian and optimising the exponential factor only. Let us define for any $i \in [2, k]$

$$\Phi(x, u_{1:T-1}, p^i) = ||h_t(x, u_{1:T-1}, p^i) - h_t(x, u_{1:T-1}, p^1)||^2_{\mathbb{Q}},$$

then the modified cost function is

$$\mathcal{J}(x, u_{1:T-1}, p^{1:k}) = \frac{1}{k} \sum_{i=2}^{k} e^{-\Phi(x, u_{1:T-1}, p^i)} \quad (3)$$

it is worth noting that when there is a significant difference between the sequence of expected observations, $h_t(x, u_{1:T-1}, p^i)$ and $h_t(x, u_{1:T-1}, p^1)$, the function $\Phi(\cdot)$ is large and therefore $\mathcal{J}(\cdot)$ is small. On the other hand if the sequence of expected observations are very similar to each other, their distance measurement tends to 0 and $\mathcal{J}(\cdot)$ tends to 1. Equation (3) can be minimised using different planning techniques such as Randomly-exploring Random Trees (RRTs) (LaValle 1998), Probabilistic Roadmap (PRM) (Kavraki and Svestka 1996), Differential Dynamic Programming (DDP) (Jacobson and Mayne 1970) or Sequential Dynamic Programming (SDP) (Betts 2001). We next use this mesaure to define a new non-Euclidean cost function that can optimised using any of these methods.

2.2 Belief state estimation

We employ a non-parametric representation of the belief state, in this case a particle filter, to model multi-modal uncertainty in object pose. Each particle is the result of a sample-based model-fitting procedure similar to the one presented in (Hillenbrand and Fuchs 2011). This procedure samples a random pair of features from the query point cloud (such as a pair of points with their relative normals) and computes the rigid body transformation to the closest pair of features in the model. Once this set of particles is computed, it is possible to calculate the object's pose estimate by using a clustering algorithm and taking a representative pose from the largest cluster.

When tactile observations occur the algorithm refines the current belief state using this particle filter. We think of the reach-to-grasp trajectory as composed of two parts: i) the approach trajectory which leads to a pre-grasp configuration of the robot in which the fingers generally cage the object to be grasped without generating any contact, and ii) a finger closing trajectory which moves the fingers into contact and generate a force closure grasp. In this way any contact which occurs during the approach trajectory is considered as an unexpected observation. Similarly an insufficient number of contacts for a force closure at the end of a grasping trajectory can also be used as an observation. In our implementation, the belief is updated assuming deterministic dynamics.

2.3 Tactile observation model

The manipulator is composed of rigid links organised in a kinematic chain and tree, comprising an *arm* and a set of $\mathbf{M}$ multi-joint *fingers*. In our robot only the final finger phalanges (finger tips) are able to detect a contact with the object. Let $\mathbf{M}$ be the ordered set of parts which compose the manipulator, then $x(j)$ is the configuration in joint space of the j^{th} part, with $j \in M$. In other words, j is the index of a specific chain. We also define $\bar{\mathbf{M}}$ to be the set of indices such that the respective part is used in the observational model. In addition, we use the operator $\mathcal{W}(x(j))$ to refer to the workspace coordinates in $SE(3)$ of the j^{th} joint with respect to a given reference frame.

The likelihood of observing a contact for each finger of the robot is an exponential distribution over the Euclidean distance, d_{ji}, between the finger tip's pose, $\mathcal{W}(x(j))$, and the closest surface of the object assumed to be in pose p^i. The observation model is limited to contacts which may occur on the internal surface of fingers. This directly affects the planner which rewards trajectories that would generate contacts on the finger tips rather than on the back side of the fingers. Therefore for any $j \in \bar{M}$ we write

$$Pr(y(j) = 1|x(j), p^i) = \begin{cases} \eta \exp(-\lambda d_{ji}) & \text{if } d_{ji} \leq d_{max} \\ & \text{and } \langle n_{xj}, \hat{n_{p^i}} \rangle < 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\langle n_{xj}, \hat{n}_{p^i} \rangle$ is the inner product of, respectively, j^{th} finger tip's normal and the estimated object surface's normal, and d_{max} describes a maximum range in which the likelihood of reading a contact is not zero, η is a normaliser. That allows us to rewrite the likelihood of reading a contact on the force/torque sensors of the robot, $h(x, p^i), i \in [1, k]$ with $j_1, \ldots, j_m \in \bar{M}$ as follows,

$$h(x, p^i) = [Pr(y(j_1) = 1|x(j_1), p^i), \ldots, Pr(y(j_n) = 1|x(j_m), p^i)]^T$$

2.4 Planning a trajectory to maximise information gain

The implementation of this planner uses a modified version of Probabilistic Roadmap (PRM), (Kavraki and Svestka 1996), to plan trajectories and detect collisions.

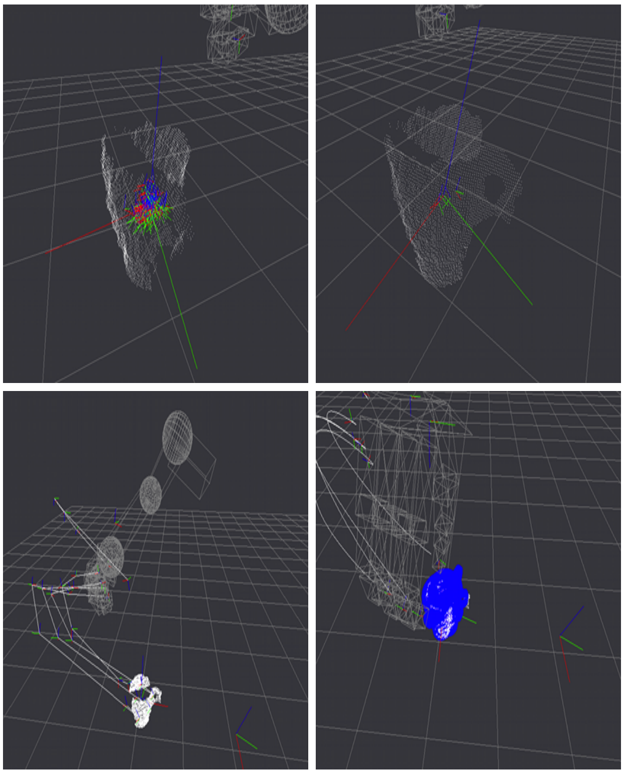

Fig. 2 Top row: The high dimensional belief state used to track object pose (left); the low dimensional filter used for planning (right). Bottom row: The unoptimised PRM plan for fingers and wrist (left); the optimised and smoothed trajectory (right).

As discussed in Sec. **??**, the PRM algorithm is composed of two phases: i) *learning phase*, in which a connected graph **G** of obstacle-free configurations of the robot is generated and, ii) *query phase*, in which a path is searched for a given pair of configurations x_{root}, x_{goal}. However the computational cost for the learning phase grows very fast with respect to the dimensionality of the problem. This planner therefore incrementally builds connections between neighbouring nodes during the query phase. Given a pair $\langle x_{root}, goal\rangle$ which describes the root state in configuration space, $\mathbb{R}^n$, and goal state in workspace, $SE(3)$, of the trajectory, this planner uses an A* algorithm to find a minimum cost trajectory in obstacle-free joint space according to:

$$c(x) = c_1(x, x_{root}) + c_2(x, x', \hat{x}_{goal}) \tag{4}$$

where $x, x' \in \mathbb{R}^n$ and $x' \in Neighbour(x)$, $\hat{x}_{goal}$ is a reachable goal configuration for the robot computed by inverse kinematics, $c_1(x, x_{root})$ is the cost-to-reach x from x_{root} and $c_2(x, x', \hat{x}_{goal})$ is a linear combination of the cost-to-go from x to a neighbour node x' and the expected cost-to-go from x' to the target. I implemented $c_1(\cdot)$ as a cumulative discounted and rewarded travelled distances. More specifically, I define

$$c_2(x, x', \hat{x}_{goal}) = \alpha d_{bound}(x, x') + \beta d(x', \hat{x}_{goal}) + \gamma d_{cfg}(x) \tag{5}$$

where $\alpha, \beta, \gamma \in \mathbb{R}$, $d(\cdot)$ is a distance function in $SE(3)$ which linearly combine rotational and translational distances in workspace[1]. For $d_{bound}(\cdot)$, let $\mathcal{B}_n(r) = \{x \in \mathbb{R}^n | x^T x \leq r^2\}$ and $\mathcal{B}(r_l, r_a) = \{A = [\begin{smallmatrix} R & p \\ 0 & 1 \end{smallmatrix}] \in SE(3) | p^T p \leq r_l^2$ and $1 - \langle Q(R), Q(R)\rangle \leq r_a^2\}$[2] denote repectively the r-ball in $\mathbb{R}^n$ and in $SE(3)$, then $b_{bound}(x, x')$ is defined as

$$d_{bound}(x, x') = \begin{cases} \psi(x, x') & \text{if } W(x) - W(x') \in \mathcal{B}(r_l, r_a) \\ & \text{and } x - x' \in \mathcal{B}_n(r) \\ +\infty & \text{otherwise} \end{cases}$$

where $Q(\cdot)$ is the Quaternion operator for $R \in SO(3)$, $\langle q_1, q_2\rangle$ is the inner product of two quaternions, $r_l, r \in \mathbb{R}$, $r_a \in (0, 1)$, and $\psi(x, x') = \zeta d(x, x') + (1 - \zeta)||x - x'||_2$ with $\zeta \in (0, 1)$. Finally, $d_{cfg}(\cdot)$ is a function which penalises dangerous configurations of the robot (i.e. close to joint limits).

I redefine the heuristic $c_2(\cdot)$ in order to reward informative tactile explorations while attempting to reach the goal state (described as a target configuration of the manipulator).

$$\bar{c}_2(x, x', \hat{x}_{goal}, A, p^{1:k}) = \alpha \mathcal{J}(x, x', p^{1:k}) d_{bound}(x, x') + \beta d_A(x', \hat{x}_{goal}) + \gamma d_{cfg}(x') \tag{6}$$

where A is the diagonal covariance matrix of the sampled states, for any column vector $a, \mu \in \mathbb{R}^n$, $d_A(a, \mu) = \sqrt{(a - \mu)^T A^{-1}(a - \mu)}$ is the Mahalanobis distance centered in μ and $\mathcal{J}(x, x', p^{1:k}) \in (0, 1]$ is a factor which rewards trajectories with a large difference between expected observations if the object is at the expected location, p^1, versus observations that would be expected if the object is at other poses, $p^{2:k}$, sampled from the distribution of poses associated with the object's positional uncertainty:

$$\mathcal{J}(x, x', p^{1:k}) = \frac{1}{k-1} \sum_{i=2}^{k} e^{-\Phi(x, x', p^i)} \tag{7}$$

where:

$$\Phi(x, x', p^i) = ||h_t(x, x', p^i) - h_t(x, x', p^1)||_2$$

for each $i \in [2, k]$ and $h_t(x, x', p^i)$ is sequence of probability of reading a contact travelling from state x to

[1] For the sake of simplicity, I reduce the mathematical notation by writing $d(x, x')$ instead of d(W(x),W(x')).

[2] I simplified the notation $\mathcal{B}_{SE(3)}(\cdot)$ in $\mathcal{B}(\cdot)$ for pratical reasons.

x'. In this implementation $h_t(x, x', p^i) = h(x', p^i)$. In other words, I evaluate the likelihood of making a contact while moving from state x to x' as the likelihood of making a contact only in the next state x'. Note that this observational model is designed to conserve (6) as in (5) when the likelihood of observing a tactile contact is zero. In fact, for robot configurations in which the distance to the sampled poses is larger than a threshold, d_{max}, the cost function $\mathcal{J}(\cdot)$ is equal to 1. However I also encode uncertainty in the second factor of the heuristics, $d_A(\cdot)$, which evaluates the expected distance to the goal configuration. In this way the planner also copes with pose uncertainty at the early stages of the trajectory, when the robot is still too far away from the object to observe any contacts.

2.5 Planning for Dexterous manipulator

In order to compute a dexterous trajectory which allows us to plan movement for both arm and fingers we need to break down the curse of dimensionality or, equivalently, increase the number of sampled configurations to properly cover the configuration space.

The proposed solution is to build a hierarchical planner. First a PRM is constructed only in the arm configuration space in order to find a global path between the $x_{root}, \hat{x}_{goal}$. It is worth noticing that in this phase the rest of the joints of the manipulator are interpolated in order to have a smooth passage from x_{root} to $\hat{x}_{goal}$. Then the planned trajectory is refined by constructing a new PRM in the entire configuration space of the manipulator (e.g. arm + hand joint space) along the global path. In other words, this approach limits the new PRM to explore only the subspace nearby the configurations which compose the global path. Subsequently an optimisation procedure is executed along the trajectory to generate a smoother transition from one configuration to the next.

This approach enable us to plan dexterous reach-to-grasp trajectories up to 21 DoF with only 1,000 sampled configurations. Note that this is the same order of magnitude that it is used in practise for planning trajectories of much simpler 6 DoF robot manipulators.

2.6 Belief update

Once a trajectory is executed and a real (unexpected) observation y is detected, the belief state is updated according to the Bayes' rule. The belief state is represented as a set of N particles $b_t = \{b_t^z\}_{z=1}^N$. In a particle filter fashion the weight of each particle b_t^z, $z \in \{1, \ldots, N\}$ is updated as follows

$$Pr(y|x, b_t^z) = \prod_{j \in \hat{M}} Pr(y_t(j)|x_t(j), b_t^z)$$

and then re-sampling is performed to generate a posterior distribution b_{t+1} as new set of particles $\{b_{t+1}^z\}_{z=1}^N$.

In simulation this approach assumes that there are no false detections. However it is possible to distinguish whether or not a contact occurs between the object to be grasped and the robot's end-effector. For example, in case a contact with the environment is detected, the algorithm skips the belief update step and moves the robot back to a safe configuration before triggering the re-planning.

3 Implementation

4 experiments

5 background

This work is related to the approach of the approach of Platt et al. (Platt et al. 2001), (Platt et al. 2012). That work plans a sequence of actions that will generate observations that distinguish a hypothesised state from competing hypotheses while also reaching a goal position. Platt et al. applied this to planning for a two degree of freedom manipulator using a laser range finder for observations, and employed an optimisation framework for planning. The algorithm is proved to localise the true state of the system in one dimension and to reach a goal region with high probability. In contrast to (Platt et al. 2001), our approach encodes information gathering actions to localise an object to be grasped in six dimensions while simultaneously attempting to achieve the task of grasping. Similarly to (Platt et al. 2001), (Platt et al. 2012), this method is guaranteed to converge to the true state of the system in which a reach-to-grasp trajectory succeeds with high probability, under the assumption that the system is not perturbed by previous grasping attempts (e.g. the robot contacts the target object without changing its configuration).

6 conclusion

References

Betts, J.: (2001). *Pratical methods for optimal control using nonlinear programming.* Siam

Ferrari, C., Canny, J.: (1992). Planning optimal grasps. In: *IEEE Proc. Int. Conf. on Robotics and Automation*

Hillenbrand, U., Fuchs, A.: (2011). An experimental study of four variants of pose clustering from dense range data. In: *Computer Vision and Image Understanding*

Jacobson, D., Mayne, D.: (1970). *Differential dynamic programming*. Elsevier

Kavraki, L., Svestka, P.: (1996). Probabilistic roadmaps for path planning in high-dimensional configuration spaces. In: *IEEE Trans. on Robotics and Automation*

LaValle, S.M.: (1998). Rapidly-exploring random trees: A new tool for path planning. Tech. rep., Computer Science Dept, Iowa State University

Nikandrova, E., Laaksonen, J., Kyrki, V.: (2013). Towards informative sensor-based grasp planning. *Robotics and Autonomous Systems* pp. 340–354

Platt, R., Kaelbling, L., Lozano-Perez, T., Tadrake, R.: (2001). Simultanious localization and grasping as a belief space control problem. Tech. rep., CSAIL, MIT

Platt, R., Kaelbling, L., Lozano-Perez, T., Tedrake, R.: (2012). Non-gaussian belief space planning: Correctness and complexity. In: *IEEE Proc. Int. Conf. on Robotic and Automation (ICRA)*

Suarez, R., Roa, M.A., Cornella, J.: (2006). Grasp quality measures. Tech. Rep. IOC-DT-P-2006-10, Technical University of Catalunya (UPC)