Grasp Affordances from Multi-Fingered Tactile Exploration using Dynamic Potential Fields

Alexander Bierbaum, Matthias Rambow, Tamim Asfour and Rüdiger Dillmann
Institute for Anthropomatics
University of Karlsruhe (TH)
Karlsruhe, Germany
{bierbaum,rambow,asfour,dillmann}@ira.uka.de

Abstract—In this paper, we address the problem of tactile exploration and subsequent extraction of grasp hypotheses for unknown objects with a multi-fingered anthropomorphic robot hand. We present extensions on our tactile exploration strategy for unknown objects based on a dynamic potential field approach resulting in selective exploration in regions of interest. In the subsequent feature extraction, faces found in the object model are considered to generate grasp affordances. Candidate grasps are validated in a four stage filtering pipeline to eliminate impossible grasps. To evaluate our approach, experiments were carried out in a detailed physics simulation using models of the five-finger hand and the test objects.

I. INTRODUCTION

Robotic grasping using multi-fingered hand constitutes a complex task and introduces challenging problems. For well known scenes a grasping or other manipulation process may be pre-programmed when using today’s robots. On the other hand, adaptation of a grasping algorithm to formerly unknown or only partially known scenes remains a difficult task, to which different approaches have been investigated. A classical approach consists in grasp analysis and planning, based on a geometric scene model. In force based grasp planning the forces and moments at selected grasping points are analyzed and matched against a grasp quality criterion considering e.g. force closure. This approach is usually independent of the hand kinematics. In contrast mere geometry based algorithms are tailored to specific gripper designs, especially in the context of multi-fingered hands. Comprehensive overviews on grasp planning are given in [1], [2]. Using grasp planning for previously unknown objects consequently introduces the difficulty of model building from sensor data which is delivered by robot perception. As alternatives to the mere planning approach online control algorithms driven by tactile information have been developed, which make use of a priori assumptions on the object to grasp, and control the grasping process by displacing robot fingers. Different control goals have been formulated for grasping convex objects in [3], [4] and later [5], where contact displacements are calculated in order to minimize a grasp quality cost function. The function values are computed using estimation of local surface parameters from haptic feedback, thus resulting in an online control scheme. A further extension capable of dealing with concavities on an object’s surface was presented in [6].

Online grasping approaches using a discrete set of hand postures or motions have also been presented [7], [8]. Beside vision based methods tactile exploration may be used for 3D reconstruction of an unknown object, as tactile sensing solves some severe limitations of computer vision, such as sensitivity to illumination and limited perspective. A reconstructed 3D object model may be used for grasp planning and execution as shown e.g. in [9]. Single finger tactile exploration strategies for recognizing polyhedral objects have been presented and evaluated in simulation, see [10] and [11]. In [12] a method for reconstructing shape and motion of an unknown convex object using three sensing fingers is presented. In this approach friction properties must be known in advance and the surface is required to be smooth, i.e. it must have no corners or edges. Further, multiple simultaneous sensor contacts points are required resulting in additional geometric constraints for the setup.

In general, previous approaches in robot tactile exploration for surface reconstruction did not cover the problem of controlling multi-finger robot hands during the exploration process. Also, real world constraints such as manipulator limits or robustness over measurement errors have not been considered. In [13] we have presented first results on the application of a dynamic potential field control technique for guiding a multi-finger robot hand across the surface of an unknown object and simultaneously building a 3D model from contact data.

In this paper we extend our approach in tactile exploration to serve the purpose of extracting grasp affordances for a previously unknown object. Therefore, we have added modifications to our exploration strategy which lead to a homogenous exploration process and prevent sparsely explored regions in the acquired 3D model. We have added a grasp planning system based on a comprehensive geometric reasoning approach as initially reported in [14]. We chose a geometric reasoning approach here as object modelling from tactile exploration currently does not deliver the details required for force analysis and contact modelling, as it is performed in force-based grasp planners, e.g. [15]. As we believe that robustness and applicability of tactile exploration
and robotic grasping algorithms depend significantly upon the deployed hardware configuration, we have evaluated our approach in the framework of a physical simulator, reflecting non-neglectable physical effects such as manipulator kinematics, joint constraints or contact friction. As in related approaches we initially limit our scope to the exploration of static scenes, which means the objects are fixated during exploration and may not move during interaction, although we wish later to develop means of pose estimation and tracking for objects in dynamic scenes. This paper is organized as follows. In the next section a short introduction to the potential field technique is given and the relevant details of the robot model are described. In Sec. IV-A we present the tactile exploration process and in Sec. IV-B grasp planning and execution. We give details on our simulation scenario and exploration results in Sec. V. Finally, our conclusions and outlook on our future work may be found in Sec. VI.

II. POTENTIAL FIELD CONTROL

Artificial potential fields have originally been introduced for the purpose of on-line collision avoidance in the context of robot path planning [16]. In the original approach, real-time efficiency was emphasized over obtaining a complete planner. The basic idea is that the robot behaves like a particle influenced in motion by a force field. The field is generated by artificial potentials $\Phi_i$, where obstacles are represented as repulsive potentials $\Phi_r(x) > 0$ and goal regions are represented as attractive potentials $\Phi_a(x) < 0$. The superposition property allows to combine potentials in an additive manner,

$$\Phi(x) = \sum_i \Phi_{r,i}(x) + \sum_j \Phi_{a,j}(x).$$

The force vector field or potential field $F$, which influences a Robot Control Point (RCP) at position $x$ is defined as

$$F = -\nabla\Phi(x).$$

A major drawback of potential fields is the existence of local minima outside the goal configurations in which the imaginary force exerted on an RCP is zero. By applying harmonic potential functions it is possible to construct potential fields without spurious local minima for point-like robots. This is not the case with robots that can not be approximated by a point, e.g. a manipulator arm. These are likely to exhibit structural local minima which need to be treated by dedicated escaping strategies [17].

III. ROBOT HAND KINEMATICS, CONTROL AND SENSORS

For exploration and grasping we consider a setup comprising a 6-DoF manipulator arm with a five finger robot hand attached to its Tool Center Point (TCP). The manipulator arm was modelled according to the Mitsubishi RM-501 five axis small-scale industrial manipulator, which is currently used as a research platform for dexterous haptic exploration in our lab. The model was augmented with a sixth DoF before the TCP to provide a larger configuration space. In our exploration control scheme we apply controller outputs to a set of five RCPs, located at the fingertips of the robot hand and to the TCP of the manipulator. The kinematic model of the robot hand is shown in Fig. 1. The hand model provides nine degrees of freedom and is modelled according to the FRH-4 anthropomorphic robot hand presented in [18].

During haptic exploration we are interested in controlling the velocity vectors of the RCP’s, which is a different task compared to trajectory control. In trajectory control the end-effector is commanded to follow a desired trajectory with the motion control goal of asymptotic tracking. Yet, the given exploration task does not induce specific trajectories due to the uncertainty in the environment. In our approach we compute the velocity vector applied to an RCP directly from the dynamic potential field, which guides the exploration process. In order to evaluate our concept in a physics simulation environment it was not required to develop a solution to the multipoint end-effector inverse kinematic problem. Instead we chose to take advantage of the physical model of the robot system and directly specify velocity vectors to the RCPs by using a virtual actuator which is commonly available in physics simulation frameworks. The joint angles are then determined by solving the constrained rigid body system and a stable and consistent configuration of the robot hand is maintained. In general, this approach is known as Virtual Model Control (VMC), which is described in detail in [19]. In our case we specify joint constraints and joint friction for the robot model for achieving an appropriate force distribution over the joint serial paths, while we do not model a compliant behavior. The physics simulation is solved by using the Inventor Physics Modeling API (IPSA) which was introduced in [20].

We also make use of the dynamic potential field concept during initialization and grasp execution by placing attractive sources at desired target locations.

For haptic exploration and contact sensing during grasping, tactile sensors are required which we have modelled in our physics simulation. Of course the simulation environment itself may be regarded as omniscient and therefore it is
possible to query all contact locations and force vectors during the interaction of modelled physical bodies. We have restricted contact sensing to dedicated sensor areas which cover the fingertips and the palm of the robot hand, see also Fig. 1. Further, we did not consider the contact force vector but only the contact location on the sensor area to provide a more realistic sensor model. This complies with current tactile sensor technology which in general can not provide both types of information. It is also possible to model more specific sensor characteristics such as a certain resolution in contact location or contact force thresholding, which we did not yet consider in our experiments.

IV. Exploration and Grasping System
The goal of our work is a system enabling a robot with a multi-fingered hand to explore an unknown object using tactile sensing and subsequently find suitable grasps. Therefore, our system comprises a module for tactile exploration as depicted in Fig. 2. In the following we will describe the exploration and grasp planning process and transition between both modes of operation. Tactile exploration is executed in closed-loop and online in simulation. In contrast, the extraction of grasp affordances is an offline planning process executed subsequently to exploration. Please note that major details of the dextrous tactile exploration process have been reported in [13]. Therefore we will summarize the basic concept and point out the improvements to the original algorithm.

A. Dexterous tactile exploration
As a prerequisite the system requires a rough initial estimate about the objects position, orientation and dimension. In simulation we introduce this information to the system, while this information will be provided by a stereo camera system in the real application. From this information an initial potential field containing only attractive sources is constructed. The trajectories for the RCPs are continuously calculated from the field gradient, while contact point locations and normals are sensed and stored as oriented 3D point set. The normal vectors are estimated by averaging the finger sensor orientations within a spherical neighborhood around a contact point. The RCP trajectories are constrained depending on the contact state of the sensor associated with each RCP, which aims to produce tangential motion during contact.

The potential field is updated from the tactile sensor information as follows. If a contact is detected, a repelling source is inserted at the corresponding location in the potential field. Otherwise, if no contact is found in the circumference of an attractive source, this source becomes deleted from the field. The robot system is likely to reach structural minima during potential field motion. We therefore introduced a reconfiguration observer which detects when the TCP velocity and the mean velocity of all RCPs fall below predefined minimum velocity values. This situation leads to a so called small reconfiguration which is performed by temporarily inverting the attractive sources to repulsive sources. This forces the robot into a new configuration from which previously unexplored goal regions may be explored. As this method is not guaranteed to be free of limit cycles we further perform a large reconfiguration if subsequent small reconstructions remain ineffective, i.e. the robot does not escape the structural minimum. During a large configuration the robot is moved to its initial configuration.

Our approach to extract grasp affordances relies on identifying suitable opposite and parallel faces for grasping. Therefore, we needed to improve the tactile exploration process as described above to explore the object surface in a dense scheme and prevent sparsely explored regions. The faces become extracted after applying a triangulation algorithm [21] upon the acquired 3D point set. Triangulation naturally generates large polygons in regions with a low contact point count. We use this property to introduce new attractive sources and guide the exploration process to fill the contact information gaps. Within fixed time step intervals we execute a full triangulation of the point cloud and rank the calculated faces by their size of area. We then add an attractive source at the centers of the ten largest faces. This
leads to preferred exploration of sparsely explored regions, i.e. regions that need further exploration, and consequently to a more reliable estimate for the objects surface. We apply a similar scheme to isolated contact points, i.e. contacts that have no further contact points in their immediate neighborhood. We surround these by eight cubically arranged attractive charges. This leads to the effect that once an isolated contact is added, the according RCP now explores its neighborhood instead of being repelled to a more distant unexplored region.

B. Grasping Phase
As an exemplary application for our exploration procedure we have implemented a method for identifying grasp affordances from the oriented point set. We did not choose a traditional force-based grasp planning algorithm as this would require to calculate a triangulated geometric object model from the 3D point set. The point set delivered by tactile exploration is inherently sparse and irregular and we found that most triangulation algorithms would fail to produce results in a usable way. Instead we found that extraction of local features from the point set is more robust than triangulation. We therefore chose a subset of a geometric reasoning approach as proposed in [14] in order to compute grasp affordances based on the acquired object information.

1) Extraction of grasping features: A grasp affordance contains a pair of object features from which the grasping points are determined in subsequent steps. In general, planar faces, edges and vertices of a polygonal object representation may be used as object features. We only consider planar faces in our implementation, as estimation and extraction of planar faces from the given 3D point set is much more reliable than that of edges or vertices. Therefore, we investigate the oriented 3D point set for neighboured contact points with similar normal vectors. Using a region growing method the contact points in adequate dense regions are assigned to faces. The original method is designed for parallel robot grippers therefore the grasp affordances found are consequently of a parallel type with opposing planar faces for grasping. We apply a mapping scheme as described below in Sec. IV-B.3 to compute the five finger tip target locations for the robot hand within each face.

2) Geometric feature filters: Initially every possible face pairing is considered as a potential grasp affordance. In a sequential geometric filtering process all grasps unlikely to be executed successfully with the given robot hand are eliminated from the set of all pairings. The filter parameters are chosen for the FRH-4 hand. We use a four stage filtering pipeline in our approach. The results of the filter stages are summed up to a score for each grasp affordance. Each filter is designed to return a value of 0 when disqualified a pairing and value \( 1 \leq o \leq 1.1 \) for accepting a pairing. As only grasp affordances with filter score \( \geq 4 \) are considered valid this automatically implies that valid grasps have to pass all filter stages successfully.

- **Parallelism:** This filter tests the two faces for parallelism. Let \( \vec{n}_1 \) and \( \vec{n}_2 \) be the normal vectors of the two faces \( f_1 \) and \( f_2 \), \( \phi \) the angle between \( \vec{n}_1 \) and \( \vec{n}_2 \) and \( \phi_{\text{max}} \) the maximum angle for acceptance. The output of the filter is:
  \[
  o = \begin{cases} 
  0, & \text{if } \phi > \phi_{\text{max}} \\
  1 + \frac{\phi_{\text{max}} - \phi}{\phi_{\text{max}}}, & \text{otherwise.}
  \end{cases}
  \]

- **Minimum Face Size:** This filter tests the two faces for adequate size of area. Let \( a_1 \) and \( a_2 \) be the areas of the faces \( f_1 \) and \( f_2 \). The minimum area for acceptance is \( a_{\text{min}} \), \( k_a \) is a normalization factor. Then the output of this filter is:
  \[
  o = \begin{cases} 
  0, & \text{if } (a_1 < a_{\text{min}}) \lor (a_2 < a_{\text{min}}) \\
  1 + \min(\min(\frac{a_1}{k_a}, \frac{a_2}{k_a}), 0.1), & \text{otherwise.}
  \end{cases}
  \]

- **Mutual Visibility:** With this filter the two faces are projected into the grasping plane \( gp \), which is the plane with the mean normal vector \( \vec{n}_{gp} \) situated in the middle of the two faces \( f_1 \) and \( f_2 \). So let \( f_{1gp} \) and \( f_{2gp} \) be the projections of \( f_1 \) and \( f_2 \) onto \( gp \). Then, \( a_{\text{int}} \) is the intersection area of \( f_{1gp} \) and \( f_{2gp} \). The minimum intersection area for acceptance is \( a_{\text{min}} \), \( k_{\text{mv}} \) is a normalization factor. The filter’s output is:
  \[
  o = \begin{cases} 
  0, & \text{if } a_{\text{int}} < a_{\text{min}} \\
  1 + \min(\frac{a_{\text{int}}}{k_{\text{mv}}}, 0.1), & \text{otherwise.}
  \end{cases}
  \]

- **Face Distance:** The last filter incorporates the characteristics of the used manipulator tool, i.e. the robot hand. The filter checks if the robot hands spreading capability matches the distance of the faces. Let \( d \) be the distance between the centers of the faces \( f_1 \) and \( f_2 \), \( d_{\text{min}} \) and \( d_{\text{max}} \) are the minimum respectively maximum admitted distance values. Then the filters output is:
  \[
  o = \begin{cases} 
  0, & \text{for } d \notin [d_{\text{min}}, d_{\text{max}}] \\
  1, & \text{otherwise.}
  \end{cases}
  \]

3) Grasp execution: The grasp affordance with the highest score is used as the candidate for grasp execution. In a first step we compute the grasping position \( \vec{p}_{\text{tcp,a}} \) of the TCP and the grasping approach direction as depicted in Fig. 3.

Initially we estimate the centers \( c_1, c_2 \) of the two faces \( f_1, f_2 \) as the centers of gravity of all contact points assigned to each face. From this we determine the center point \( g\vec{p} = \frac{c_1 + c_2}{2} \) on the line connecting the centers of the two faces. Then we analyse the first principle component \( \vec{p}\hat{c} \) of the acquired 3D point cloud and calculate the grasping position as

\[
\vec{p}_{\text{tcp,a}} = g\vec{p} + (\vec{n}_{gp} \times \vec{p}\hat{c}) \cdot d,
\]

where \( d \) is a distance which considers the fingers length of the robot hand. The cross product \( (\vec{n}_{gp} \times \vec{p}\hat{c}) \) becomes the approach direction. We only consider grasping the object from top. Therefore, in the case the coordinate \( \vec{p}_{\text{tcp,a}} \) is below the object to grasp, we mirror its location across the center.
between \( c_1 \) and \( c_2 \) and along the approach direction to a location above the object. Clearly, we make an assumption about the object's extension here. From the face pair of the grasp affordance finger tip target locations need to be computed. This is achieved by the following mapping scheme.

The target for the thumb \( \vec{p}_{\text{thumb}, a} \) is set to be the center of the smaller of the two faces. We choose target locations \( \vec{p}_{\text{index}, a}, \vec{p}_{\text{middle}, a}, \vec{p}_{\text{ring}, a} \) and \( \vec{p}_{\text{pinkie}, a} \) for the opponent fingers around the center and in the plane of the larger of the two faces. The arrangement is chosen, so that it is perpendicular to the approach direction in the plane of target face. If the target location of ring finger or pinkie is not situated within the face area the fingers will not be used for grasping. This way the number fingers involved during grasping is automatically adapted.

Motion execution starts with the hand in an initial pose, as it is always reached after a large reconfiguration. From here we apply the potential field control to the RCPs and the TCP. Unlike during the exploration phase, the TCP and the RCPs share the set of repulsive potential sources while having individual attractive potential sources as mentioned above. Repulsive sources located in the target planes become deleted.

As long as the TCP is distant from its target \( \vec{p}_{\text{tcp}, a} \) the potential field velocity control is only applied to the TCP while the finger joints remain open via direct joint control. When the TCP is close to its target we additionally apply the potential field control to the RCPs. If an RCP is not in use because the finger is not involved in grasping, the associated finger joints are still kept open. Further, the palm normal \( \vec{n} \) is aligned towards \( g \vec{p} \) by controlling forces acting on the hand’s pitch and roll DoFs.

If the RCPs in use have approached the finger target locations, the fingers are closed and the corresponding sensors are checked for contact. Once all assigned RCP sensors are in contact with the object, potential field control is turned off and the finger joints are closed directly. The virtual fixture of the object then becomes disabled in the simulation and the robot arm moves back to its initial position with the object grasped and lifted.

---

**V. Simulation Results**

We evaluated our exploration and grasping system in several virtual scenes using our physics simulator with standard earth gravity \( g_N = 9.81 \) applied. For contacts Coulomb friction with a friction coefficient \( \mu = 0.5 \) is considered. The virtual scenes were set up with different rigid objects of suitable size for grasping by the hand: a sphere, a telephone receiver and a rabbit. The objects are placed approximately in the center of the robots workspace. All objects are fixed floating above the simulators virtual ground to avoid interference, as we currently do not differ between contact between the object of interest and any other obstacle in the workspace.

As described in Sec. IV-A the cubical bounding box of the object is computed from position and space occupancy estimates and used to initialize the exploration potential field. Grasp affordances are extracted after a fixed number of 2000 control time steps, whereby each control time step comprises ten simulation time steps with a temporal resolution of \( T = 0.04s \).

Fig. 4 shows typical results. Here figures in column (c) show the 6 best candidate faces for grasping. The color indicates score ranking in following order: red, green, blue, magenta, cyan, yellow. Black dots indicate the center of a face, which is calculated as mean value of all points in the face. Colored lines connect corresponding centers of corresponding faces.

In column (d) the grasp affordance with the highest score is shown. Purple dots indicate grasping points for index, middle, ring and pinkie finger. Ring and pinkie grasping points are only plotted if they are used. The red dot marks the location of the attractive potential source for the TCP at start of the approaching phase. Naturally, the algorithm performs worse with objects exposing curved regions as the algorithm searches for planar faces. Therefore, only one grasp affordance was found for the sphere in the given exploration interval. The exploration of the rabbit shows similar results. Still, successful grasps can be performed with the grasp affordances identified.

In contrast, several affordances could be identified with the model of the telephone receiver consisting of large polygons. In general, the number of found grasp affordances increases with exploration time. The video accompanying this paper shows examples of tactile exploration and grasp execution for the rabbit.

Beside experiments with different objects we also investigated performance of the system with objects placed at different positions and orientations in the workspace. For the experiments a grasp is considered successful if the manipulator can grasp and lift the object in simulation.

We believe this is still a good approximation for reality as the simulator only calculates with rigid body dynamics and assumes point contacts. In reality such a robot system would be equipped with deformable rubber finger tips which will provide a significant larger contact area leading to higher tangential forces. Therefore we assume that a real robot system could execute the simulated successful grasps.

In a first experiment we placed the sphere, which is naturally
Fig. 4. Typical simulation results from top to bottom: Sphere, telephone receiver, rabbit. Column (a) shows a virtual scene snapshot during exploration, (b) final point cloud, (c) grasp affordances, (d) best grasp and grasping points.

Fig. 5. Number of identified grasp affordances. Blue: successful grasp execution, red: failed grasp execution with best candidate.

invariant to rotations, at different distances ranging from minimum to maximum reaching distance for the manipulator arm in the workspace. Fig. 5(a) shows the number of found grasp affordances after $N = 2000$ exploration steps. After generation the grasp affordance with the highest score is executed as described in Sec. IV-B.3. In the figure a red bar indicates a failed grasp execution, a blue bar indicates a successful grasp execution, both with the best candidate grasp applied. The failed grasps may be deduced to the error between the estimated grasping plane and the local tangential plane of the sphere in combination with an inappropriate situation of the sphere within the robot’s workspace. This could be improved by increasing the exploration time in order to collect more contact data points.

In a second experiment we investigated the scheme with the robot model for sensitivity towards different orientations of an elongated object as the telephone receiver. Therefore, the receiver is placed in the scene with different orientations
around the Y-axis (direction of gravity). The initial configuration can be seen in the mid image of Fig. 4(a). The receiver was situated in the workspace center area of the manipulator arm. The results are depicted in Fig. 5(b) and indicate that the receiver provides less features to extract grasp affordances from with its longer axis pointing toward the manipulator. The reasons for the failed grasp agree with those from experiment 1. Note that the receiver is not a symmetric object, therefore the number of grasping candidates is also not symmetric over rotation.

VI. CONCLUSIONS

In this paper we have presented a control scheme for tactile exploration and subsequent extraction and execution of grasp affordances for previously unknown objects using an anthropomorphic multi-fingered robot hand. Our approach is based on dynamic potential fields for motion guidance of the fingers. We have shown that grasp affordances may be generated from geometric features extracted from the contact point set resulting from tactile exploration. The complete control scheme was evaluated in a detailed physics simulation of the robot system with test objects of different shape and presented the results of the grasp planner based on the exploration data. Finally, we tested the best grasp candidate by executing the grasp within the physics simulation. In further experiments we have reported results for different object locations and orientations in the manipulator workspace.

For the future we are working on an extension of the presented set of geometric filters in order to further improve the success rate upon grasp execution with our robot hand. Further we will consider the incorporation of the palm during grasp execution, which would enable power grasps. Concluding, we are confident that the dynamic potential field based approach presented may be used for real world tactile exploration and grasping with an anthropomorphic robot hand, as it appears robust enough to autonomously control interaction of the robot hand with a previously unknown object using tactile information. We assume that the proposed scheme is transferable to different manipulator and robot hand kinematics by adapting filter parameters, number of RCPs and RCP locations. We further plan to investigate possibilities of combination with exploration methods based on sensors of different modalities than haptics, e.g. vision based object exploration. The developed control scheme based on VMC and dynamic potential fields is currently subject to implementation on our real world robot system equipped with five-finger hands [22].

ACKNOWLEDGEMENT

The work described in this paper was conducted within the EU Cognitive Systems projects PACO-PLUS (FP6-027657) and GRASP (FP7-215821) funded by the European Commission.

REFERENCES