Planning Hand-Arm Grasping Motions with Human-Like Appearance

Néstor García, Raúl Suárez and Jan Rosell

Abstract—This paper addresses the problem of obtaining human-like motions on hand-arm robotic systems performing grasping actions. The focus is set on the coordinated movements of the robotic arm and the anthropomorphic mechanical hand, with which the arm is equipped. For this, human movements performing different grasps are captured and mapped to the robot in order to compute the human hand synergies. These synergies are used to both obtain human-like movements and to reduce the complexity of the planning phase by reducing the dimension of the search space. In addition, the paper proposes a sampling-based planner, which guides the motion planning following the synergies and considering different types of grasps. The introduced approach is tested in an application example and thoroughly compared with a state-of-the-art planning algorithm, obtaining better results.

I. INTRODUCTION

Nowadays, robots are becoming essential in more and more fields and applications but they are also becoming more sophisticated and complex. The humanoid robots equipped with anthropomorphic dexterous hands are one of the most representative examples. These anthropomorphic mechanical hands are devices that concentrate in a compact volume a high number degrees of freedom (DOFs), ranging usually from 12 to 25 joints, as well as several different sensors. Obtaining a satisfactory performance of these hands requires the automatic planning of their movements, which is still an arduous and non-evident task since the complexity of the planning problem increases exponentially with the number of DOFs. Furthermore, sometimes not only a feasible path is required but also one that optimizes some quality metric, for instance, the motion planning of humanoid robot must not only focus on the efficient search of a valid solution but also on the search of robot movements that mimic the human motions to enhance the human-robot collaboration [1].

Motion planning of complex systems has been addressed using different planning algorithms, being the sampling-based planners [2] and, especially among them, the Probabilistic Roadmap planners, PRM [3], and the Rapidly-exploring Random Trees, RRT [4], the most commonly used. These algorithms have been extensively studied and, hence, several variants exist, for instance to deal with constraints [5], or to bias the sampling towards better regions of the configuration space by using potential fields [6] or retraction-based methods [7].

The authors are with the Institute of Industrial and Control Engineering (IOC), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain (nestor.garcia.hidalgo, raul.suarez, jan.rosell)@upc.edu. This work was partially supported by the Spanish Government through the project DPI2016-80077-R and N. García is supported by the Generalitat de Catalunya through the grant FI-DGR 2018.

On the other hand, the robot joints must be properly coordinated in order to obtain human-like motions [8]. Thereby, real human movements are commonly used as a reference [9], either pursuing a direct on-line teleoperation of the robot [10], or with the aim of analyzing these movements and getting some valuable information to be applied later in a planning phase [11]. Some relevant pioneering works dealt with the grasping problem analyzing the correlations of the finger joints when the human hand was grasping objects [12]. These correlations were called hand postural synergies and mapped into a mechanical hand [13]. The synergies existing in the human hand were also used for other objectives such as the analysis and design of robotic hands in order to mimic human grasps [14], the design of specific hand control systems [15], or the identification of the hand pose using low-cost gloves [16]. Nevertheless, there exist other approaches that, instead of studying the hand synergies while grasping an object, compute the synergies from hand movements when the human tries to cover the whole hand configuration space in an unconstrained way [17]. More recently, a compliant model, called soft synergies, was also introduced and used in the selection of grasping forces, in their control, and in the control of the motion of the grasped object [18], [19]. In addition, synergies were also used in a dual-arm anthropomorphic system while performing manipulation tasks [11], [20]. The works mentioned above dealt with synergies involving correlations between joint positions. More recent works extended the concept of synergies to the velocity space (i.e. the space of the first derivative of the configuration trajectories) calling them first-order synergies [21], [22] (in contrast with the synergies in the configuration space, that were called zero-order synergies).

This work proposes to characterize the synergies existing in the human grasping motions, considering the different grasp types [23], [24]. A method to identify the different phases in the grasping motions is presented and used to obtain an all-purpose pre-grasp set of synergies and a set of grasping synergies for each grasp type. The work pursues efficiency as well as human-likeness in the obtention of hand-arm movements and, for this, it introduces a new sampling-based motion planner that considers different potential grasps at the same time and steers the path towards the direction of the synergies.

After this introduction, Section II presents the problem statement and gives an overview of the proposed approach. Section III details the proposal, the approach is validated in Section IV and finally Section V presents the conclusions and future work.
II. PROBLEM STATEMENT AND APPROACH OVERVIEW

The goal of this work is to plan the motions of a hand-arm robotic system trying to mimic the hand-arm movements that a human does to pick an object performing different types of grasp. To this end, a sampling-based planning algorithm is designed and the movements of a human operator are used to guide the motion planning. The main key points of the proposed approach are the following:

1) The motions of a human operator performing different types of grasp on several objects are captured and then mapped to the robot whose motions are aimed to be planned (see Fig. 1-left).
2) From the captured information, the different grasping phases in the demonstrations are identified and the set of synergies existing in the human grasping motions are computed (one per grasp type). Thereby, the complexity of the different DOF-couplings in the human motions is accurately explained in a simple manner.
3) A bidirectional sampling-based planner is designed to bias the tree growth towards the directions of the computed synergies and to reduce the dimension of the search space, being this dimension-reduction process dependent on the synergies of grasp type to perform. Hence, human-like movements are obtained efficiently.

III. PLANNING PROCEDURE

A. Motion capture and mapping

In this work, human motions are used as a reference to obtain human-like movements of a hand-arm robotic system picking a given object. Many types of human grasps are gathered in the grasp taxonomy of Cutkosky [23], which classifies the grasps depending on the object size and on the task to perform. Although this classification is not complete, and there exist more extensive grasp classifications (e.g. [25]), it is detailed enough for the considered purposes. Besides, Dai, Sun and Qian [24] updated the taxonomy of Cutkosky and analyzed, from a different perspective, the entire grasping trajectory and not only the grasping configuration (i.e. the final snapshot), proving that the grasp types can be grouped naturally into a set of four consistent grasp families (see Fig. 2). This family-grouping is used here to adapt the planning process according to the grasp being performed, even several potential grasp types can be considered simultaneously.

Thereby, using a Cyberglove sensorized glove with a 50 Hz sampling frequency, the motions of a human operator are recorded performing 15 different grasp types on 9 objects, with 12 repetitions per grasp type and starting off from a comfortable position in front of the object (see Fig. 1-middle and Fig. 2). This implies 180 demonstrations and more than 15000 configuration samples (where each sample contains 22 measurements read from the glove describing the positions of the finger joints). Once the samples have been captured, they are mapped to the robotic hand. This mapping depends on the kinematic structure and particularities of the used robotic system. Nevertheless, slight differences in the mapping strategy do not affect excessively the overall performance of the proposed approach. In this work, a robotic hand-arm system composed of a 6-DOF UR5 robotic arm equipped with a 16-DOF Allegro Hand is used (see Fig. 1-right). Since the robotic hand has only three fingers besides the thumb, the information regarding the little finger is discarded. Then, the captured movements are mapped as follows. First, the values of the flexion/extension joints of the fingers and the thumb are computed with a joint-to-joint mapping. Next, the values of the remaining joints (i.e. the thumb-opposition joint and the abduction/adduction joints of the fingers and the thumb) are computed with a fingertip-position mapping.

B. Motion analysis

The synergies (i.e. couplings between DOFs) are obtained running a Principal Component Analysis (PCA) over the set of hand configurations once mapped from the human movements (instead of computing directly the human synergies and then having to map them to the robotic system in a non-trivial way, e.g. [26]). This returns a new basis of the hand configuration space, with the axes sorted in decreasing order of the associated sample variance (i.e. they are ordered with the first axis marking the direction with maximum sample variance). Each axis represents a synergy and the motion along it, equivalent to a single DOF, implies the movement of several (or all) joints. Although nonlinear approaches to obtain synergies have been also proposed (e.g. [27]), the simple linear approximation of the PCA is enough to capture the subspace where the demonstrated motions lie, having been proved to be useful and implementable by a drive mechanism [28] and a real-time algorithm [15].

Two phases are observed in the mapped grasping motions (see Fig. 3). During the first phase, called pre-grasp phase,
the trajectories of the hand joints are common motions opening the hand similarly in all the executions, regardless of the grasp type performed. Then, there is a moment in which the demonstrated trajectories begin to differ and specialize according to the type of grasp being carried out. This is the grasp phase itself. Nevertheless, the transition from one phase to the other is diffuse and does not occur at the same time for all the demonstrations. Hence, the transition time is obtained as follows. Let $Q$ be the set of hand configurations mapped from a given grasping demonstration, and, for a given time instant $t$, let $Q_A^t$ and $Q_B^t$ be the sets of configurations in $Q$ captured before and after $t$, respectively. In addition, let the likeness $L$ of two given sets $Q_A$ and $Q_B$ of hand configurations be defined as the overlapping of the distributions of the configurations in the sets, which is a measure of the similarity between $Q_A$ and $Q_B$ [20]. The index $L$ can be computed as

$$L(Q_A, Q_B) = \frac{\mu_A - \mu_B}{\sqrt{(2\pi)^{2n} |\Sigma_A + \Sigma_B|}}$$

where $\mu_A$ and $\mu_B$ are the barycenters and $\Sigma_A$ and $\Sigma_B$ are the covariance matrices of the configurations in $Q_A$ and $Q_B$, respectively. Then, the time $t$ marking the transition between the two phases is defined as the one minimizing $L(Q_A^t, Q_B^t)$. Thereby, the pre-grasp and grasp phases have been identified in the 180 mapped trajectories. On the one hand, all the pre-grasp phases have been grouped and used to compute the pre-grasp synergies. On the other hand, the grasp phases have been grouped according the grasp family which each demonstrated grasp belongs to, and, then, a set of grasp synergies has been computed for each grasp family (see Fig. 2). In this way, the pre-grasp synergies explain the hand motions in the pre-grasp phase in all the grasps, and each set of grasp synergies model the hand motions in the corresponding grasp family (Table I shows the accumulated sample variance for the obtained set of synergies).

For a robotic hand with $n$ DOFs, the synergies define an $n$-dimensional box centered at the barycenter of the configurations used to obtain the synergies and with each side aligned with a synergy [29]. In order for the box to contain the $(100 - \alpha)$% of the configuration distribution for a given $\alpha$ (i.e. any hand configuration inside the box would be then similar to the ones used to compute the synergies), each side of the box is set to $2\sqrt{2} \text{erf}^{-1}(\sqrt{1 - \alpha})$ times the standard deviation of the configurations in the corresponding direction (synergy). The dimension of the box can be reduced by using only $k < n$ synergies (picking them in order) such that $k$ is the minimum value making the accumulated variance be above a confidence level of $(100 - \beta)$% for a given $\beta$. In this work, $n = 16$ and $\alpha = \beta = 5\%$ is considered. Thus, the dimension $k$ of the resulting lower-dimensional boxes, called $B_k$, is 4 or 5 for the grasp phase, depending on the grasp family (see bold values in Table I). For the pre-grasp phase, 6 synergies are needed (a little bit greater, as it was expected, since the movements of all the grasp families are included). Despite the simplification, the boxes $B_k$ still represent accurately the mapped hand motions. Thereby,
configuration $q = (q^a, q^h) \in \mathcal{C}$ be a robot configuration defined as an arm configuration $q^a$ concatenated with a hand configuration $q^h$, and let $G = (q^h, \chi_h)$ be a grasp composed of the hand configuration $q^h$ and the object pose $\chi_h$ relative to the hand at the grasping time. Thereby, the proposed planner, outlined in Algorithm 1, is supplied with a collision-free start configuration $q_{\text{start}}$ of the whole robot, the object pose $\chi_r$ relative to the robot, and a set $\{G_i\}$ of grasps. The introduced planner is based on the RRT-Connect [4], which is widely used in motion planning since it obtains good results even on robots with a high number of DOFs and with cluttered environments. However, it has been modified here to:

a) **Extend the trees following the synergies:** In order to integrate the synergies into the motion planning, the standard function extending the tree in RRT-based planners is replaced here by the function STEER, described in Algorithm 2. As in the classic method, a single step is performed from $q_{\text{near}}$, the configuration in the graph closest to the desired target configuration $q_{\text{target}}$ (Line 1), reaching a new configuration $q_{\text{new}}$. If the segment connecting $q_{\text{near}}$ and $q_{\text{new}}$ is collision-free, $q_{\text{new}}$ is returned (Line 6). Otherwise, $\emptyset$ is returned (Line 7). However, here, $q_{\text{new}}$ is computed differently, i.e. following the synergies instead. A step, with a maximum length $\epsilon$, is taken not towards the desired $q_{\text{target}}$ (as it would be done in the standard procedure) but towards its projection $q_{\text{proj}}$ onto the lower-dimensional box spanned by the synergies (Lines 4-5), see Fig. 4. It should be remarked that, in the computation of $q_{\text{proj}}$, the arm component of $q_{\text{target}}$ remains the same and the hand component is projected onto the lower-dimensional box $B_k$ of synergies associated with the root of the tree containing $q_{\text{near}}$, i.e. if $q_{\text{near}}$ belongs to the tree rooted at $q_{\text{start}}$, $q_{\text{near}}$ is projected onto the box of pre-grasp synergies; otherwise, $q_{\text{near}}$ belongs to a tree rooted at a certain $G_i$ and, hence, $q_{\text{near}}$ is projected onto the box of synergies associated with $G_i$. However, if $q_{\text{target}}$ is close to $q_{\text{near}}$, $q_{\text{new}}$ is $q_{\text{target}}$ (Line 2), so that in the event that the two trees are close to be connected, the guideline to follow the synergies may be relaxed.

**C. Motion planning**

![Fig. 4. Motion planning representation in $\mathcal{C}$-space: sample trees rooted at the start configuration $q_{\text{start}}$ and the grasps $G_1$ and $G_2$, growing close to the associated synergy lower-dimensional boxes (0 to 2), while steering a given configuration $q_{\text{near}}$ towards a random $q_{\text{rand}}$ and reaching $q_{\text{new}}$.](image)

if the planning of the hand motions is performed in the corresponding $B_k$, the planning complexity is reduced and the obtained motions are still similar to the movements mapped from the human.

**Algorithm 1: PLANNER**

**Input**: Start configuration $q_{\text{start}} \in \mathcal{C}$, object pose $\chi_r$, and set of grasps $\{G_i\}$

**Output**: Collision-free path $P$ connecting $q_{\text{start}}$ and one $G_i$

1: $(E_a, V_a) \leftarrow (\emptyset, q_{\text{start}})$
2: $(E_b, V_b) \leftarrow (\emptyset, \emptyset)$
3: forall $G_i$ do
4:  $q_{\text{goal}} \leftarrow $ INVKIN$(G_i, \chi_r)$
5:  if $q_{\text{goal}} \neq \emptyset$ and COLLISIONFREE$(q_{\text{goal}})$ then $V_b \leftarrow V_b \cup q_{\text{goal}}$
6:  while not ENDCONDITION() do
7:     $q_{\text{rand}} \leftarrow $ RANDCONF()
8:     $q_{\text{new}} \leftarrow $ STEER$((E_a, V_a), q_{\text{rand}})$
9:     $(E_a, V_a) \leftarrow (E_a \cup \{q_{\text{near}} : q_{\text{new}}\}, V_a \cup q_{\text{new}})$
10:    while $q_{\text{new}} \neq \emptyset$ do
11:        if $V_a \cap V_b \neq \emptyset$ then return PATH$((E_a, V_a), (E_b, V_b))$
12:        else
13:            $SWAP((E_a, V_a), (E_b, V_b))$
14:            $q_{\text{new}} \leftarrow $ STEER$((E_a, V_a), q_{\text{new}})$
15:            $(E_a, V_a) \leftarrow (E_a \cup \{q_{\text{near}} : q_{\text{new}}\}, V_a \cup q_{\text{new}})$
16:        $SWAP((E_a, V_a), (E_b, V_b))$
17:    return $\emptyset$

**Algorithm 2: STEER**

**Input**: Graph $(E, V)$ and configuration $q_{\text{target}}$

**Output**: Configuration $q_{\text{new}}$

1: $q_{\text{near}} \leftarrow $ NEARESTCONF$(V, q_{\text{target}})$
2: if $\|q_{\text{target}} - q_{\text{near}}\| \leq \epsilon$ then $q_{\text{new}} \leftarrow q_{\text{target}}$
3: else
4:  $q_{\text{proj}} \leftarrow $ PROJECT$(q_{\text{target}}, (E, V))$
5:  $q_{\text{new}} \leftarrow $ NEAREST$(\{q_{\text{near}} : q_{\text{proj}}\}, (q_{\text{proj}} - q_{\text{near}}) / \|q_{\text{proj}} - q_{\text{near}}\|)$
6: if COLLISIONFREE$(q_{\text{near}}, q_{\text{new}})$ then return $q_{\text{new}}$
7: else return $\emptyset$

b) **Deal with multigoal queries:** The planner maintains two graphs, each one denoted by a pair formed by a set of edges $E$ and a set of vertices $V$. One of this graphs represents a tree rooted at the start configuration $q_{\text{start}}$ (Line 1), and the other contains the trees rooted each one at a grasp configuration (Lines 2-5), see Fig. 4. These grasp configurations are computed by, first, solving the arm inverse kinematics (IK) given the grasps $G_i$ to be performed and the object pose $\chi_r$ (Line 4) and, then, rejecting those cases that do not have an IK-solution or that imply collisions (Line 5). Thus, in each iteration, one of the graphs is steered towards a random configuration $q_{\text{rand}}$ (uniformly d in $\mathcal{C}$), reaching a configuration $q_{\text{new}}$ (Lines 7-9). Note that the STEER method, explained above, returns $\emptyset$ a collision is found, and consequently the graph is not extended. Next, the connection between the graphs is attempted.

c) **Connect the trees in a less greedy fashion:** In the classic RRT-Connect, the trees are connected greedily by extending one of trees directly until the other tree is reached or a collision occurs. However, here, in order to obtain a smoother connection, both graphs are, in alternation and successively, extended towards the last added configuration in the other graph (Lines 13-15), until the graphs are connected and, then, the found solution path is returned (Line 11). In case the steering process fails (Line 10), the graphs swap their roles (Line 16) and the whole process is repeated.
until a solution is found or some termination condition is satisfied (Line 6), e.g. surpassing a maximum planning time, number of iterations or memory allocation.

IV. VALIDATION OF THE APPROACH

The motions of an anthropomorphic dual-arm robot have been planned for illustrative purposes (see Fig. 1-right). The robot is located in front of a bookshelf and, starting off from a natural standing pose, it must grasp a cylinder standing on one of the shelves (see Fig. 5). Besides, the robot must perform human-like motions while avoiding the collisions with itself, the bookshelf and the cylinder. For this, the planning algorithm is provided with the exact position of the cylinder and with a set of different force-closure grasps $G_i$ (from different grasp types and grasp families, see Fig. 6). This information can be obtained, for instance, from the vision system on the robot and a grasp generator, respectively. In order to evaluate and compare the performance of the proposed approach, three planners have been benchmarked:

a) A standard RRT-Connect, modified to tackle multi-goal queries, planning without using synergies.

b) The proposed approach, planning using the proper grasp synergies related to the grasp type to be performed.

c) The proposed approach, planning with the grasp synergies and the grasp families intentionally mismatched (i.e. each grasp family has been randomly associated with the synergies of another grasp family).

Note that, at each planner execution, the orientation and position of the robot are lightly modified at random and that 8 different grasps $G_i$ are randomly selected from the set of precomputed grasps. Thereby, the planners are provided with a single start configuration and a set of 8 different goal configurations, i.e. one per each of the selected grasps.

The experiments introduced above have been implemented within The Kautham Project [30], a simulation and motion planning framework developed at the Institute of Industrial and Control Engineering (IOC-UPC) for teaching and research purposes, and run in a 2.13-GHz Intel 2 PC. A maximum planning time of 100 s is considered for each planning instance. Thereby, if a path is not found within this time, the execution is marked as a failure. After 100 executions, Table II shows the average values of the success rate, the planning time, the number of iterations and collision checks, the rate of valid segments (i.e. the ratio of iterations in which the trees actually grow), the path length (measured in $C$ as the weighted sum of accumulated joint movements along the path), and the path human-likeness. The human-likeness index computes the misalignment of a path with respect to some given reference human movements [29]. Here, natural free-movements of the operator while moving freely the fingers in an unconstrained way (i.e. without performing any specific task), trying to cover the whole hand workspace, are used as a reference.

On the one hand, it can be noticed from the simulation results that effectively the proposed planning approach is several times faster than the standard RRT-Connect algorithm (up to an order of magnitude). In fact, the motion planning can be solved within the time restrictions for the 100% of the executions only when the proposed approach is used, either when the grasp synergies are properly associated with the selected grasps or when they are mismatched. It

<table>
<thead>
<tr>
<th>Case</th>
<th>Success rate</th>
<th>Planning time</th>
<th># iterations</th>
<th># collision checks</th>
<th>Valid segments</th>
<th>Path length</th>
<th>Human-likeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>97%</td>
<td>51.80 s</td>
<td>1834</td>
<td>32231</td>
<td>68.3%</td>
<td>14.18 rad</td>
<td>73.6%</td>
</tr>
<tr>
<td>b</td>
<td>100%</td>
<td>6.21 s</td>
<td>274</td>
<td>10649</td>
<td>80.0%</td>
<td>7.79 rad</td>
<td>83.1%</td>
</tr>
<tr>
<td>c</td>
<td>100%</td>
<td>11.79 s</td>
<td>484</td>
<td>13667</td>
<td>75.3%</td>
<td>8.35 rad</td>
<td>81.9%</td>
</tr>
</tbody>
</table>

Fig. 5. Snapshots of paths obtained with an standard RRT-Connect (top), and the proposed procedure using the proper synergies for each grasp type (bottom).

Fig. 6. Examples of different grasp types used within the motion planning, each one of a different grasp family: Thumb-2 Finger, Thumb-3 Finger, Medium Wrap and Lateral Pinch (from left to right, respectively).

TABLE II
AVERAGE RESULTS OF THE MOTION PLANNING WHEN RUNNING THE CLASSIC RRT-CONNECT (a) AND THE PROPOSED APPROACH WITH THE PROPER (b) AND WITH MISMATCHED GRASP SYNERGIES (c).
can be stated that the use of synergies clearly reduces the planning time since the solution is enforced to lie close to the lower-dimensional boxes $D_k$. This focuses the search efforts close to the demonstrated movements (which belong to a set of demonstrated feasible solutions), thus accelerating the connection of the trees and, thereby, reducing the number of iterations and collision checks to find a solution. In addition, since the grasp synergies are obtained from feasible movements, the probability of obtaining collision-free robot configurations increases when using synergies (see valid segments rate in Table II), reducing greatly the computation time. The results also show that even when no correct grasping synergies are used, i.e. case (c), the benefits of using synergies are still evident. In this case, the planning time is slightly penalized, however, it is still a better option than not using synergies at all.

On the other hand, the proposed planning procedure produces movements of the robotic system that look more natural and human-like (see human-likeness in Table II), since the grasp synergies are obtained from human demonstrations and the human-likeness is preserved within the planning process. Besides the numerical results, the higher human-likeness of the proposed approach can be noticed in Fig. 5, which shows representative solution paths for cases (a) and (b).

V. CONCLUSIONS AND FUTURE WORK

This paper has proposed a procedure to efficiently obtain human-like hand-arm movements to grasp a given object. To this end, the movements of a human operator performing different grasps on different objects have been captured and mapped to the robot. These grasp movements have been classified according to a grasp taxonomy, and for each grasp family a set of human-demonstrated synergies (couplings between DOFs) have been computed. In addition, a pre-grasp set of synergies has also been computed, common for all the grasp families. Finally, a motion planner profiling from these synergies has been presented and compared against a state-of-the-art planner planning the motions of a real anthropomorphic dual-arm robot. The effect of using the grasp synergies, even when they are not the ones associated with the grasp being performed, has also been investigated, producing good results in both cases. Besides, the proposal opens several interesting potential research lines, such as its extension to the velocity space and the coordination of the robot base, arms and hands all at the same time.

REFERENCES