Introducing Geometric Constraint Expressions Into Robot Constrained Motion Specification and Control
Gianni Borghesan, Enea Scioni, Abderrahmane Kheddar, Herman Bruyninckx

To cite this version:

HAL Id: lirmm-02155096
https://hal-lirmm.ccsd.cnrs.fr/lirmm-02155096
Submitted on 13 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Introducing Geometric Constraint Expressions into Robot Constrained Motion Specification and Control

Gianni Borghesan¹, Enea Scioni¹,², Abderrahmane Kheddar³, Herman Bruyninckx¹,⁴

Abstract—

The problem of robotic task definition and execution was pioneered by Mason, [1], who defined setpoint constraints where the position, velocity, and/or forces are expressed in one particular task frame for a 6-DOF robot. Later extensions generalized this approach to constraints in i) multiple frames, ii) redundant robots, iii) other sensor spaces such as cameras, and iv) trajectory tracking. Our work extends tasks definition to i) expressions of constraints, with a focus on expressions between geometric entities (distances and angles), in place of explicit set-point constraints, ii) a systematic composition of constraints, iii) runtime monitoring of all constraints (that allows for runtime sequencing of constraint sets via, for example, a Finite State Machine), and iv) formal task descriptions, that can be used by symbolic reasoners to plan and analyse tasks. This means that tasks are seen as ordered groups of constraints to be achieved by the robot’s motion controller, possibly with different set of geometric expressions to measure outputs which are not controlled, but are relevant to assess the task evolution. Those monitored expressions may result in events that trigger switching to another ordered group of constraints to execute and monitor.

For these task specifications, formal language definitions are introduced in the JSON-schema modeling language.

Index Terms—Software, Middleware and Programming Environments; Motion and Path Planning; Behaviour-Based Systems.

I. INTRODUCTION

Specifying what one wants a robot to do, instead of having to code manually how it has to do it, is still a challenge for the research community; for example, in Fig. 1, where a simple open-the-drawer task is represented, what the robot does is to open the drawer.

However, the task specification concerns the satisfaction of some geometric constraint (distance to zero, correct alignment, etc.), in order to grasp the uppermost handle. The how-to-do task specification approach can be made easy if regarded from the perspective of recent constrained optimization based control. Instantaneous motion of the robot’s joints is computed as the solution of a constrained optimization problem in which:

- the objective function defines the general behavior (e.g. time, energy consumption...) and can even include weight-prioritized tasks and/or utility functions;
- the constraints expressed as equality as well as inequality functions, define the sub-space in which the solution belongs to in terms of the task and robot “state variables”.

Moreover, the switching between different sets of objective functions and constraints is ruled on the basis of monitoring, the extent to which some functions (not necessary constraint functions) assume given values during the task execution. Previous work such as, e.g., Stack of Tasks (“SoT”, [2]), instantaneous Task Specification and Control (“iTaSC”, [3]), Whole Body Control (“WBC”, [4]), or a revised versions of the Task Frame Formalism, (“TFF”, [5], [6]), are examples of such a trend. However, a large part of these formalisms, a robot can be driven to achieve complex tasks that require the definition of many constraints in different spaces and in different reference frames. However, large part of the research has focused on maturing the control aspects.

Figure 1. The representation of the a grasping task for the (topmost) handle and open-the-drawer task needs the definition of many geometric features upon which constraints must be enforced.

Manuscript received: August 12, 2015; Revised October 6, 2015; Accepted November 27, 2015.

This paper was recommended for publication by Editor Tamim Asfour upon evaluation of the Associate Editor and Reviewers’ comments. The authors gratefully acknowledge the support from the European Commission's FP7 project "RoboHow" (FP7-ICT-288533), and KU Leuven’s Geconcerteerde Onderzoeks-Actie “Global real-time optimal control of autonomous robots and mechatronic systems”.

¹Mechanical Engineering, University of Leuven, Belgium.
²Engineering Department (ENDIF), University of Ferrara, 44122 Ferrara, Italy.
³CNRS-AIST Joint Robotics Laboratory (JRL), UMI3218/RL, Japan, and CNRS-Université Montpellier, LIRMM, Interactive Digital Human, France.
⁴Mechanical Engineering, Eindhoven University of Technology, the Netherlands.

Digital Object Identifier xxxxxxxxxxxxxxxxxxxxxxx
(namely the numerical solver part), leaving the way tasks are designed and formalized untreated.

Summing up, the state of the practice is that:

- current frameworks do not follow the rule of the separation of concerns (see e.g. [7]): they do not separate the configuration of the hardware platform, of the data communication and synchronization between software components, etc., from the task definition;
- a formal semantics is still missing, reducing the opportunities to use task planning at the symbolic level to automatically generate tasks.

Looking at the above considerations’ state-of-the-art, we are convinced that a fundamental step to ease the employment of constraint-based systems and task-based programming is to separate the task description from its implementation and describe the first with a Domain Specific Language that follows semantic rules. Our main contribution is therefore the design of a DSL that allows abstraction of: i) the space where tasks are expressed, which is defined in terms of a symbolic expressions, and ii) the type of control to be enforced on such space, which is defined in terms of the achieved behaviour.

II. Tasks as sets of constraints: fundamentals

Our goal is to describe a task constraint with a minimal set (of properties, relations, equations...), in such a way that the specification is abstracted (as much as possible) by the implementation of how to solve for the constraint during the execution of the task. We start considering the simplest task description formalism, the Task Frame Formalism, [1], [5], where each constraint is defined employing a minimal set of specifications, namely: i) the controlled and task frames, ii) selection matrix $S$ that expresses which, among the six components of the Cartesian space, are controlled in position and which are controlled in force, and iii) the reference values. Thus, the description of a task (in the Task Frame Formalism) consists of at most six non-conflicting constraints, that are applied in the controlled frame (that normally corresponds to the robot end effector) and expressed in a frame where position and orientation are chosen to represents elements of interest for the task itself. Employing the Task Frame Formalism approach, Bruyninckx and De Schutter, [5], already proposed a generic way of defining a tasks as a set of constraints (an example is given in Listing 1).

The Task Frame Formalism separates the task description from its implementation: indeed, Listing 1 fully defines a task, with no explicit dependencies to robot information, such as: i) the kinematics (redundancy as well as mechanical and actuator constraints), ii) the control laws employed in the motion or force control, and iii) the perception capabilities (estimation, observation...).

Similar considerations can be applied to more recent frameworks (e.g. SoT, [2], and WBF, [4]), that, while posing the emphasis on different aspects, maintain the same approach in task definition.

Our extensions to this task description approach are:

- a wider choice of the constrained variables, which, instead of being related to the Cartesian space representation (not always suited to easily describe a generic relationship between two bodies), will be represented as (one of an enumerable set of functions between) geometric primitives. In this regard, Sec. IV focuses on the description of the geometric expressions, and on the geometric primitives used by these geometric expressions;
- the substitution of the type of control (e.g., force and position) with the type of behaviour that we want to achieve; which includes the description of the type of control and the type of constraint, Sec. V;
- the specification of measurement expressions that are employed to monitor the task execution state and to trigger changes in the robot behaviour.

The combination of the previous elements (expressions, behaviours and monitors) formally defines a Domain Specific Language (DSL) for robotic tasks. Task instances, as well as the specification itself, have been encoded as JavaScript Object Notation (JSON) [9] and JSON-Schema [10] documents, respectively. This allows us to decouple the language-dependent task grounding from the formal description, not only to guarantee portability and legibility, but also to exploit existing tools and model validators. The validation of the model semantic is the first step to allow automatic reasoning systems that generate symbolic plans, [11], [12], or symbolic task scheduling, [13]. Further numerical post-check operations have been implemented in the host language of the controller. For sake of brevity, we will report snippets of the models written in JSON-Schema only for the model point and the entity.

III. Analogies with symbolic planning

While we develop this formalism as a way to describe tasks at an abstract level, we later realized that this effort has overlapping goals with works that investigate how to bridge the gap between symbolic planning and continuous action planning. In the latter, actions are often the result of the execution of pre-computed joint trajectories; these trajectories are in turn generated from one or more target configurations that are then checked

Other works followed this line, either focusing on the specification (e.g. [8]), or on the flexibility of constraint definition (e.g. [3]).

Listing 1. Example of a guarded-motion task definition from [5].

```
move compliantly |
  with task frame directions
  xt: velocity @ mm/sec
  yt: velocity @ mm/sec
  zt: velocity @ mm/sec
  zt: velocity @ mm/sec
  zt: velocity v_des @ mm/sec
  ax: velocity @ rad/sec
  ay: velocity @ rad/sec
  az: velocity @ rad/sec
  until zt force <= f_max N
```

1All the models described can be retrieved at https://github.com/gborghesan/jgeom_constr and accompanying material.
for feasibility and reachability (e.g. in [14], [15] RRT-based methods are used). In [16], object geometry is leveraged to reduce computation time of RTT, while in [17] the Task Space Regions are described: TSRs allow to combine constraints expressed in Cartesian space and compute joint trajectories by means of a RRT-based algorithm. Generally, path planning is solved in joint space: as a consequence, if the state of the environment changes, replanning is preferred to active control (e.g. [15]), and it is not easy to deal with under-constrained goals. Moreover, planning deals with joint trajectories, and interaction tasks cannot be detailed.

We favour (constraint-based) control rather than full-planning. Constraint-based control is known to be compliant to changes in the environment (as shown in [3]) and allows for the exploitation of unconstrained degrees of freedom (doF). It is naturally suited for hybrid (force-position) task specification. On the other hand, complex environments can cause a task-based control to fall in local minima that a global trajectory planner can avoid. From these considerations, the reader can see that the two approaches are in fact complementary and cover different requirements.

IV. FROM GEOMETRIC ENTITIES TO EXPRESSIONS

This section introduces the geometric concepts of i) geometric entities (a geometrical feature such as a point, a versor, a line, a plane, etc.), ii) geometric primitives (entities expressed in a frame), and iii) geometric expressions (a scalar function that relates primitive). Expressions can be non-geometric as well (e.g. joint expressions).

Expressions are a relationship that maps joint-space values and measurements, to task-space values (controlled, measured, etc.). The use of expressions is two-fold: i) to compute “deviations” from (task-space) positions, or ii) to compute the “Jacobian” (the partial derivative of the expression w.r.t. joint angles) that relates joint velocities (forces) to generalized velocities (torques). Since our focus is task specification, no assumption is made on how values and Jacobian are computed (e.g. analytically, numerically, with solution of implicit equations, etc.).

A. Geometric entities

The elementary entities considered are the point and the versor; both have three parameters, that can be represented by triples (x, y, z). Additionally, the versor must comply with unitary norm constraint. Combinations of points and versors yield the other two common geometric entities, line and plane, each with five free parameters in its representation. (Choosing a non-minimal set of representation parameters often simplifies specification.) Since a line is oriented and has an origin, projection relationships can be defined (see below).

B. Geometric primitives

Entities need a frame to ground their coordinate representations: so, a geometric primitive (Listing 3) associates a geometric entity with a frame. The four possible geometric primitives are reported in Table I, along with the associated mathematical symbol. Formal meta models of point entity and primitive are given in Listing 2–Listing 3; a meta model represents all the formal constraints that every individual instance must conform to.

C. Geometric expressions

The list of primitives in Table I is not exhaustive, but suffices for most scalar expressions that describe posi-
The paper considers the following behaviours:

- **Positioning**, used in position regulation problems.

V. The behaviour

The “output space” of a task specification is described by geometric expressions, but what kind of constraints should be realised in such space must still be specified.

The latter specification is what we call the behaviour; a behaviour is a way to specify what we want to achieve in a given space while delegating the determination of the control law to the underlying solver.

The paper considers the following behaviours:

- **Positioning**, used in position regulation problems.

---

Table II

Summary of geometric expressions between primitives.

<table>
<thead>
<tr>
<th>point</th>
<th>point-point distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>line</td>
<td>line-point distance</td>
</tr>
<tr>
<td></td>
<td>projection of point on line</td>
</tr>
<tr>
<td>plane</td>
<td>point-plane distance</td>
</tr>
</tbody>
</table>

(a) Geometric expressions on distances.

<table>
<thead>
<tr>
<th>versor</th>
<th>angle btw versors</th>
</tr>
</thead>
<tbody>
<tr>
<td>plane</td>
<td>angle btw planes</td>
</tr>
</tbody>
</table>

(b) Geometric expressions on angles.

The geometric expressions between pairs of objects. For example, Tables IIa and IIb give distance and angle expressions. Since most of the table entries are self-explaining, we focus only on the explanation of the line-point and line-line distances.

The line-point entry (Fig. 2a) has two expressions: the line-point distance, that is the distance between the point \( p_2 \) and the point of shortest distance \( f_1 \), and the projection of point on line, that is the (signed) distance between the points \( p_1 \) and \( f_1 \).

The line-line entry (Fig. 2b) has three distance expressions: the distance from lines, distance \( d \) between the two lines, the signed distance between the point \( p_1 \) and the minimum distance point \( f_1 \), and the signed distance between the origin of the line \( p_1 \) and the minimum distance point \( f_2 \).

D. Expressions in joint space

Many cases need to express constraints in a robot’s joint space, the most obvious cases being limits in torque, position, or velocity.

E. Composite expressions

Many tasks require more complex expressions, either as a combination of the elementary expressions above, or based on more complex geometrical shapes, or specified directly as geometric “curves”. The approach introduced in this paper is to enumerate all models that one requires in a specific task context, and with this approach it is straightforward to extend the enumeration list. Here are a number of such higher complexity expressions.

- **Non-scalar expressions**: multi-dimensional expressions, as (3D) rotations. Rotations are often used when full orientation is constrained. But in many practical cases, analogous results can be achieved using three angle between versors expressions.

- **Sensor-space expressions**: in some applications constraints are directly expressed in a sensor space, e.g. in visual servoing with eye-in-hand camera. If this kind of relation cannot be expressed by means of geometric expression, a new expression must be formalized, and the underlying implementation realized:

\[
y = f(x_0, q, \ldots),
\]

where \( x_0 \) is the relative position of the measured object w.r.t. the camera, \( y \) the measured output, and \( q \) the robot configuration.

- **Virtual Fixtures and Mechanisms**: are geometric constraints modelled as kinematic chains. These computer generated geometric features were pioneered by [18] to overlay real sensory feedback in telerobotics. Later on [19] extended virtual fixtures to generate control primitives in robot teleprogramming of remote robots in the presence of delay. Recently, they are largely exploited in teleoperation and shared control scenarios, especially in the robotic medical field (see [20] and cited works). In this case, more complex primitives are employed, as often the goal of the task is to constrain the end effector to move in a sub-space defined on top of generic curves, and leaving few free directions of motion. These motion primitives are simple and their combination for complex tasks is not always easy or even possible. The concept of virtual mechanism proposed in [21] and also in [22] allows a more systematic description of the task and embed a task controller with desired properties (e.g. passivity).

Virtual fixtures and mechanisms that rely on linear, rotational, or spherical degrees of freedoms models can be achieved with geometric expressions. More complex scenarios (e.g. moving on a generic curve or surface) require an extension to the entity and primitive sets, with new expression models.
Designing the behaviour of the system when constraints are conflicting cannot be done in an exhaustive way: the system developers must choose what the application requires or varies between executions, on- and off-line, depending on the task as described in Sec. VI-A, these gains can be used to assess a smooth convergence toward the limits following the control equation (2).

Lastly, in case of limiting-type behaviours, parameters are considered optional, since in many cases (e.g. joint position or effort limits), the desired interval should not be violated (and no convergence rule toward limits is needed). On the contrary, in cases where limiting behaviours are used to define “soft tasks” (e.g. an auxiliary task as described in Sec. VI-A), these gains can be used to extend the enumerated list with additional behaviours (or add additional parameters to the proposed ones), provided that controllers exist that are able to execute them.

VI. CONSTRAINTS, MONITORS, AND TASK DEFINITIONS

A. Constraint definition

A constraint is defined as a set including the following combination: i) an expression, ii) a behaviour, and (optionally) iii) a trajectory generator, that produces the desired setpoint, as described in Table III.

In all but the simplest tasks, several constraints are enforced together. Since the initial conditions, the environment, and other aspects can be unknown at the time of defining the application, or vary between executions, tasks could results in conflicting objectives that cannot be achieved simultaneously.

For this reason, it is necessary to express explicitly in which way conflicts should be handled at run-time. Designing the behaviour of the system when constraints are conflicting cannot be done in an exhaustive way without peering to the underlying implementation and the possible options that it provides. To the best of the authors’ knowledge, two methods are used: i) weighting of “deflection” of constraints from their nominal values, employed in velocity- and acceleration- resolved schemes, and ii) prioritization of constraints, achieved mostly with null space projectors.
In our experience, we found out that three priority levels suffice, in most applications (see [13], [23] for use cases). We named these three levels as follows:

1) Safety constraints: constraints that are necessary for the robot platform or critical surroundings integrity, (e.g. hardware limitations, non-desired collision avoidance, sustain balance in humanoid robots, etc.).

2) Primary constraints: the actual constraints to be executed.

3) Auxiliary constraints: constraints that facilitate the execution of primary constraints; an example is the optimization of the robot pose configuration with respect to the manipulability index, [24], another one is gazing, etc.

B. Monitor definition

Driven by the necessity to change the system behaviour once a task is fulfilled, we include monitors in the task description. Monitors observe expression variables, raising an event once a particular logical condition is fulfilled on the monitored expression value. Monitors can be used to verify pre-, per- as well as post-conditions, e.g. [13]. A monitor can either observe:

1) a controllable variable (for example a variable that is used as constraint), or
2) a measured quantity that is influenced by the robot actions, but cannot be directly controlled. This quantity can be measured in terms of: i) a variable that lives in the space described by an expression, or ii) an external monitored value.

For the cases 1) and 2), a monitor is defined as: i) an expression (that specifies the space where the variable lives), ii) an event name (e.g. finished, failed), iii) the (monitored) variable type (POSITION, VELOCITY, FORCE), that is expressed in the space defined by the expression, iv) a comparison type (<, >, ∈, ∉), v) reference value(s). The external monitor (case 2ii) allows for composability with external sources, for example, time-out or safety events.

C. Task definition

Finally, the formal definition of a Task wraps the previous things together; it includes i) at least one primary constraint, and ii) at least one (end-of-task) monitor. Optionally, it can have iii) safety constraints, and iv) auxiliary constraints.

Thus, the minimum specification of the task corresponds to a constraint (an expression, a behaviour that rules which kind of control must be used, and a reference value) and one monitor (that dictates the end of the task execution). Having a well defined end-of-task monitor is instrumental to define the post-condition(s) of the task to be used in plan reasoning and scheduling algorithm (see [13] and references).

VII. TASK SPECIFICATION EXAMPLE

This section illustrates the proposed approach by means of an open-a-drawer operation composed of the following tasks: i) Approach the handle of the drawer, ii) Grasp the handle, and iii) Open the drawer. In order to show the results of task specification (without introducing the disturbances of controllers, uncertainties, etc.), we report the execution of this example run with kinematic simulation, using the software described in [25].

In addition, the accompanying multimedia shows the same experiment executed by the real robot, as well as a “spreading tomato” example, that makes use of force constraints and monitors.

In the following, we briefly describe the tasks in a narrative way. The simulation consist of 3 tasks, whose switching is managed by a finite state machine that reacts upon the event generated by the monitors.

A. Geometric Primitives

The object frames that are involved in the Geometric Primitives are the following:

- \(\{a^\text{1a}\}\), attached to the handle center, with the z-axis along the handle itself.
- \(\{o^\text{ib}\}\), attached to the chest of drawers (fixed in the world).
- \(\{o^\text{1}\}\), the grasp frame of the robotic hand, i.e. the frame that is the center of the grasp once the hand closes.

Specifications are available in the examples/app folder of accompanying material.
On such frames, the following geometric primitives are defined:

- $hp_{\omega}$ (handle_position): point in origin of $\{\omega^0\}$.
- $ha_{\omega}$ (handle_axis_direction): versor aligned with the $z$-axis of $\{\omega^0\}$.
- $\overline{hx}_{\omega}$, $\overline{hy}_{\omega}$, $\overline{hz}_{\omega}$ handle_axis_(x-y-z): three lines with origin in $\{\omega^0\}$, and aligned with the x-, y-, or z-axis, respectively.
- $\overline{od}_\omega$ opening_axis: line with origin in $\{\omega^0\}$, and aligned with the opening direction.
- $gp_\omega$ (grasp_position): point in origin of $\{\omega^2\}$.
- $\overline{gn}_\omega$ (grasp_normal_direction): versor aligned with the $x$-axis of $\{\omega^2\}$.
- $\overline{ga}_\omega$ grasping_axis: line with origin in $\{\omega^2\}$, and aligned with its $z$-axis.

B. “Open the drawer” tasks description

The example application is ruled by a simple three-states machine, whose state transitions are executed in response to the event raised by monitors (described in Sec. VII-C). The states (approach the tray, grasp the handle, and open the drawer) are ordered sequentially, and, for each state, a different set of task is enforced. In each task a set of primary constraint is enforced; in addition, in each task, a safety constraint is enforced in order to limit the joint positions in their range, (Position Limit behaviour).

a) Approach the tray (S.1): In this phase, the robot should bring its end effector in such a position that can conveniently grasp the handle. To do so, the robotic hand should be: i) oriented toward the object, ii) rotated along its $z$- (approach-) axis in the “correct” way, and then iii) bring the distance between the grasping point and the handle center to zero. This description translates in a set of Positioning behaviours, ruled by the following expressions:

- line-point distance between $\overline{ga}_\omega$ and $hp_{\omega}$ should be zero,
- angle btw versors $\overline{gn}_\omega$ and $ha_{\omega}$ should be zero,
- while three line-point projections ($\overline{hx}_{\omega}$, $gp_\omega$), ($\overline{hy}_{\omega}$, $gp_\omega$), and ($\overline{hz}_{\omega}$, $gp_\omega$), should go to zero.

These set of constraints allow for a one full degree of freedom (angle of approach to the handle around its normal axis). Note that expressions c) dictate that the distance between the origins of the frames $\{\omega^0\}$ and $\{\omega^2\}$ be zero. However, we express three constraints in place of a point-point distance as the derivative of such expressions is ill-defined at the desired value, thus causing local instability around such point.

b) Grasp the handle (S.2): While the positioning constraints are held, we close the gripper.

c) Open the drawer (S.3): At this point, the drawer is opened, so we command to increase the distance along the line direction of opening, by increasing the point-line projection between $\overline{od}_\omega$ and $gp_\omega$. As we do not want the handle grasp to be broken, we continue to enforce previous constrains.

C. Monitors and trajectory generator

For this example, we limit a sketch of possible monitors that can dictate the success of the action, leaving out the management of other occurrences.

a) Approach the tray: the final goal is to reach the center of the handle. We use the point-point distance between the origins of frames $\{\omega^0\}$ and $\{\omega^2\}$. In this case, the monitored expression is different to the controlled ones (the three line-point projections).

b) Grasp the handle: success can be achieved by force direct/indirect measurement in the grasp space. However, we rely on position information (the distance between fingertips).

c) Open the drawer: success can be acknowledged by monitoring when the distance between the gripper position and the initial drawer position is above a given threshold.

All the constraints that are described are commanded to the final value by means of trapezoidal trajectory generator. The results can be appreciated in Fig. 4, where the time evolution of all expressions are reported. In these figures we reported also which is the monitored expression, and the transitions between tasks.

VIII. Conclusion

This paper formalises task representations, using a strictly specified set of symbolic elements (and very few numerical values). The aim is to help the unification of the various state-of-the-art approaches that deal with complex tasks on complex robotic systems. The benefit of this process is manifold: by defining constraints with a limited number of formally described elements, it becomes possible to design reasoning algorithms to sort out the best plan for a given action (e.g. [11], [13]) and check (semantically) whether a task is ill- or well-defined, without prior knowledge of the solver specifics, and/or with a lot less involvement of human developers. This progress helps in the “ease of use” of task specification, but also in “verification & validation” of robotic applications. The set of primitives and relations can be extended, both for the geometric expressions (e.g. segments, planar polygons, solids, oriented curves to provide virtual guidance, …), or any other expressions that meets the above-mentioned criteria.

References

Figure 4. Values of constrained and monitored expressions during simulation. Vertical lines shows the transitions between states (labelled from S.1 to S.3, each one corresponding to a set of primary constraints). In the first state, the gripper is opened (Fig. 4c), aligned (Fig. 4b), and brought to the handler (Fig. 4a). State transitions is triggered when total distance (blue flat line in Fig. 4a) decreases under a given threshold. Distance (blue flat line in Fig. 4a) decreases under a given threshold.

(a) $x$, $y$, $z$ positions of gripper w.r.t. the handle (initial, fixed) frame, and total distance between frame origins.

(b) Point-line distance between the gripper grasping direction and the handle frame origin, and misalignment between the handle axis and normal grasping directions.

(c) Distance between gripper fingers.