A Tele-operated Swarm of UAVs for Cooperative Grasping and Manipulation

Guido Gioioso[1,2], Antonio Franchi[3], Gionata Salvietti[2], Stefano Scheggi[1], Domenico Prattichizzo[1,2]

I. Introduction

Aerial mobile manipulation is a growing area of robotics research that claims to bring in the air the results obtained on the terrain by the mobile manipulation community using UGVs. Different approaches have been pursued to let flying robots grasp and manipulate objects. A solution is to equip the vehicles with an arm and a gripper or more sophisticated end-effectors like in [1] .

Formations of UAVs has been also employed to grasp and transport objects using docking systems [2] or cables [3]

In this extended abstract we consider cooperative UAVs grasping an object, where each UAV makes a single contact with the object through a rigidly attached tool. Practically, we can assume each UAV acting as a *finger* of an N-fingered hand, applying a contact force on an object, and collaborating with the other UAVs to grasp and move it. Some preliminary results in this promising direction have been presented in [4].

Differently from the approach presented in [4], in this work the formation of UAVs is teleoperated by a human hand. As far as grasp and manipulation are concerned, autonomous robot navigation could fail due to the complexity of grasp planning and maintenance. The proposed teleoperation framework is described in the next section.

II. Teleoperation Framework

The main goal of the system is to use the hand/fingertip motion to remotely manipulate an object denoted by $\mathscr{O}$, changing its position $p_o \in \mathbb{R}^3$ in an inertial frame $\mathscr{W}$: $\{O_w, \vec{x}_w, \vec{y}_w, \vec{z}_w\}$ and orientation represented by the rotation matrix $R_o \in SO(3)$ (defining the orientation of a moving frame $\mathscr{B}$: $\{O_b, \vec{x}_b, \vec{y}_b, \vec{z}_b\}$ that is rigidly attached to $\mathscr{O}$), by exploiting the action of the N UAVs denoted by $\mathscr{U}_1 \dots \mathscr{U}_N$ as if they were a *remote flying hand*.

Each UAV possesses a rigidly attached tool whose tooltip acts as a finger to exert the needed interaction forces with the object. According to hard contact model [5] no torsional moments can be exchanged at the contact point and the contact force at the tooltip of $\mathscr{U}_i$ can be expressed in $\mathscr{B}$ and denoted with $\lambda_i \in \mathbb{R}^3$.

We assume that a grasp planner (that runs in a preliminary phase) returns N contact points, whose positions in $\mathscr{B}$ are denoted by $\bar{p}_1 \dots \bar{p}_N$ and a set of *nominal internal force*, denoted by $\bar{\lambda} = (\bar{\lambda}_1 \dots \bar{\lambda}_N)^T$ expressed in $\mathscr{B}$. The nominal internal forces guarantee a stable grasp in the static case and with nominal environment conditions.

As a second goal of the system, the human operator must be able to regulate the intensity of the internal forces, i.e. the component of the contact forces that does not contribute to the motion of the object but only affect the grasp stability. Defining $G \in \mathbb{R}^{6\times 3N}$ as the grasp matrix [5] and $G^\#$ as the Moore-Penrose pseudoinverse of G, the internal forces can be expressed as $\lambda_{int} = (I - G^\# G)\lambda$.

In order to achieve these goals, we propose a three-layer structured approach described in the following.

A. Hand Interpreter

Denote by $\mathscr{H}$: $\{O_h, \vec{x}_h, \vec{y}_h, \vec{z}_h\}$ a moving frame that is rigidly attached to the tracked hand, and by $\mathscr{H}_0$: $\{O_h^0, \vec{x}_h^0, \vec{y}_h^0, \vec{z}_h^0\}$ a fixed frame that coincides with $\mathscr{H}$ at t_0. Denote by $p_h \in \mathbb{R}^3$ the position of O_h expressed in $\mathscr{H}_0$, and by $R_h \in SO(3)$ the rotation matrix that expresses the orientation of $\mathscr{H}$ w.r.t. $\mathscr{H}_0$. Furthermore, denote by $(\phi_h, \theta_h, \psi_h)$ the roll-pitch-yaw (RPY) representation of such orientation. Finally, define with r_h the radius of the circle that best fits the detected positions of the fingertips, and with Δr_h the difference between the current value of r_h and its value at t_0.

The *hand parameters*: p_h, $\eta_h = (\phi_h\ \theta_h\ \psi_h)^T$, and Δr_h are used to represent the configuration of the hand, and are provided as input to the Virtual Point Mapping Layer.

In our specific implementation we tracked the hand movements using a RGB-D camera and computed the positions of the fingertips via cluster analysis. A RANSAC-based plane fitting is performed on the fingertips and the finger point clouds are projected on such plane to compute the related bounding circle, and its radius.

B. Virtual Point Mapping

The Virtual Point Mapping layer generates the N virtual attraction points on the basis of the human hand parameters provided by the Hand Interpreter (Fig. **??**). Denote by $\tilde{y}_i \in \mathbb{R}^3$ the i-th virtual attraction point position expressed in $\mathscr{W}$. At t_0

$$\tilde{y}_i = p_o + R_o \bar{p}_i,$$

i.e., the virtual points coincide with the grasping points.

At the generic time instant t we set

$$\tilde{y}_i(t) = \underbrace{\tilde{p}(t)}_{translation} + \underbrace{\tilde{R}(t)\bar{p}_i}_{rotation} + \underbrace{\alpha_r \Delta r_h \tilde{R}(t)\bar{\lambda}_i}_{grasping}, \qquad (1)$$

[1]Department of Information Engineering and Mathematical Sciences, University of Siena, via Roma 56, 53100 Siena, Italy. {gioioso, scheggi, prattichizzo}@dii.unisi.it

[2]Department of Advanced Robotics, Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova, Italy gionata.salvietti@iit.it

[3]Centre National de la Recherche Scientifique (CNRS), Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS), 7 Avenue du Colonel Roche, 31077 Toulouse CEDEX 4, France. antonio.franchi@laas.fr

$i = 1 \dots N$, where $\tilde{p}_h \in \mathbb{R}^3$ and $\tilde{R} \in SO(3)$ are generated through the following two independent dynamical systems

$$\dot{\tilde{p}} = \alpha_p p_h, \qquad \tilde{p}(t_0) = p_o(t_0) \tag{2}$$

$$\dot{\tilde{R}} = S(\omega_h)\tilde{R}, \qquad \tilde{R}(t_0) = R_o(t_0) \tag{3}$$

in which

$$\omega_h = \alpha_\omega T(\tilde{\eta})\eta_h$$

being $\tilde{\eta}$ the RPY angles associated to $\tilde{R}$ and $T(\cdot) \in \mathbb{R}^{3\times 3}$ the Jacobian matrix mapping RPY angle rates to angular velocity. The positive scalars $\alpha_p, \alpha_\omega, \alpha_r$ represent suitable scale factors.

It can be noted that if $\Delta r_h < 0$ and if p_h and η_h are identically zero, then the virtual points at time t are a compressed version of the contact points $\bar{p}_1 \dots \bar{p}_N$ only in the direction of the internal forces $\bar{\lambda}_i$.

Now, if we consider $\tilde{y}_1, \dots, \tilde{y}_N$ to be the attraction points for the UAV tooltips and if the tooltips behave like decoupled linear spring-damper systems, under the action of the UAV Force Control Layer, then the object will be moved accordingly to the operator directives.

C. UAV Force Control

The proposed tele-operation architecture works with different possible UAVs and implementations of the Force Control Layer. In order to show its practicability in a specific case we consider in this section the case of typical underactuated VTOL (Vertical Take-Off and Landing) vehicles, such as, e.g., quadrotors.

In [6] we presented a control law designed to allow the tooltip to exert a force on the environment. We derived the analytical expression of the force $-f_i^e$ exerted by the tooltip on the environment for any given $\tilde{y}_i - y_i$. The found relation for typical values of the mechanical and control parameters can be approximated with a linear map (see [6] for more details), thus showing that the tooltip behaves similarly to a linear spring. Therefore the UAV tooltip, under the action of the proposed controller is a compliant system whenever the tooltip is in contact with the object $\mathscr{O}$, justifying the design of the virtual points as attraction positions that indirectly generate the contact forces, implementing the action commanded by the human operator.

III. Simulations

In the simulations that we conducted, the motion of the human hand was tracked by a Kinect sensor communicating via UDP/IP with a Matlab/SIMULINK application which implements the hand-to-quadrotors mapping algorithm as well as the physical simulation of the environment and the object to be grasped. The user was asked to remotely move a wooden cubic box from its starting position to a target location, requiring the object to be lifted during the transport to overcome a wall. In the following some plots regarding a prototypical HIL simulation are presented with the aim of demonstrating the applicability of our framework in a realistic scenario.

Fig. 1 (left) represents the human translational command (top) and the corresponding object linear velocity (bottom). The dashed black lines delimit a dead zone introduced to facilitate the task. First a positive vertical velocity is commanded to lift the object. Then the user commands a horizontal velocity along the x axis to overcame the wall. Finally a negative vertical velocity is commanded. It is easy to see that the object velocity follows the commanded one. In Fig. 1 (right) the rotation rate commanded by the user (top) and the yaw rate of the object (bottom) manipulated by the swarm of UAVs are represented. Finally, in Fig. 2 the commanded radius (left) and the corresponding internal forces exerted by the UAVs onto the object (right) are plotted.

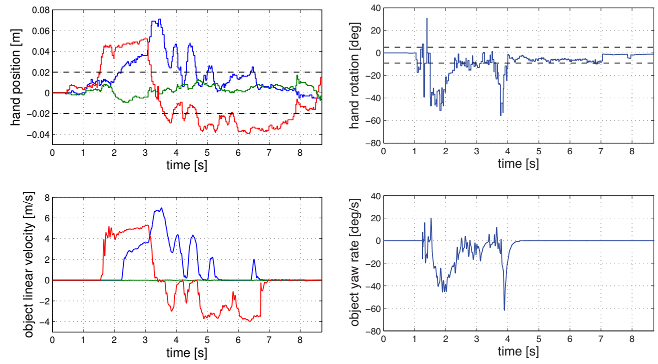

Fig. 1: Top left: displacement vector o_h of the hand centroid with respect to the rest position. Horizontal dashed black indicate the dead zone. Bottom left: linear velocity of the teleoperated box. The following color convention was used: blue$\rightarrow x$, green$\rightarrow y$, red$\rightarrow z$. Top right: rotation angle ω_h of the fingers with respect to the rest position. Horizontal dashed black indicate the dead zone. Bottom right: yaw rate of the teleoperated box.

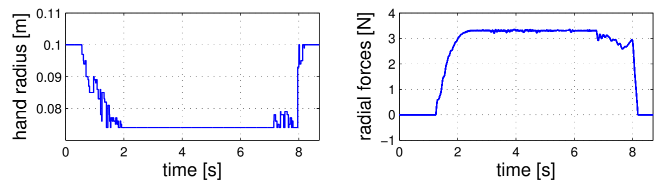

Fig. 2: Hand radius (left) and internal forces exerted by the quadrotors on the object (right).

References

[1] R. Spica, A. Franchi, G. Oriolo, H. H. Bülthoff, and P. Robuffo Giordano, "Aerial grasping of a moving target with a quadrotor UAV," in *2012 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Vilamoura, Portugal, Oct. 2012, pp. 4985–4992.

[2] D. Mellinger, M. Shomin, N. Michael, and V. Kumar, "Cooperative grasping and transport using multiple quadrotors," in *10th Int. Symp. on Distributed Autonomous Robotic Systems*, Lausanne, Switzerland, Nov. 2010.

[3] M. Manubens, D. Devaurs, L. Ros, and J. Cortés, "Motion planning for 6-D manipulation with aerial towed-cable systems," in *2013 Robotics: Science and Systems*, Berlin, Germany, May 2013.

[4] V. Parra-Vega, A. Sanchez, C. Izaguirre, O. Garcia, and F. Ruiz-Sanchez, "Toward aerial grasping and manipulation with multiple UAVs," *Journal of Intelligent & Robotics Systems*, vol. 70, pp. 575–593, 2012.

[5] D. Prattichizzo and J. C. Trinkle, "Grasping," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer, 2008, pp. 671–700.

[6] G. Gioioso, M. Ryll, D. Prattichizzo, H. H. Bülthoff, and A. Franchi, "Turning a near-hovering controlled quadrotor into a 3D force effector," in *2014 IEEE Int. Conf. on Robotics and Automation*, Hong Kong, China, May. 2014.