Teleoperating a mobile manipulator and a free-flying camera from a single haptic device*

Josep Arnau Claret, Isiah Zaplana and Luis Basañez
Institute of Industrial and Control Engineering
Technical University of Catalonia, Barcelona
Barcelona, Catalonia, Spain
Email: josep.arnau.claret@upc.edu, isiah.zaplana@upc.edu, luis.basanecz@upc.edu

Abstract—The paper presents a novel teleoperation system that allows the simultaneous and continuous command of a ground mobile manipulator and a free flying camera, implemented using an UAV, from which the operator can monitor the task execution in real-time. The proposed decoupled position and orientation workspace mapping allows the teleoperation from a single haptic device with bounded workspace of a complex robot with unbounded workspace. When the operator is reaching the position and orientation boundaries of the haptic workspace, linear and angular velocity components are respectively added to the inputs of the mobile manipulator and the flying camera. A user study on a virtual environment has been conducted to evaluate the performance and the workload on the user before and after proper training. Analysis on the data shows that the system complexity is not an obstacle for an efficient performance. This is a first step towards the implementation of a teleoperation system with a real mobile manipulator and a low-cost quadrotor as the free-flying camera.

I. INTRODUCTION

Robot usage in catastrophe scenarios has increased in the last years. From the September 11 attacks [1] to the more recent earthquake in Fukushima [2], a growth can be noted in the use of robots in service and rescue scenarios [3].

Due to the dangerousness for human workers, it is desirable to command the robots from a secure area. This concept is at the core of teleoperation since its inception for the manipulation of nuclear material [4]. Recent works deal with teleoperation in hazardous scenarios: [5] presents a system control architecture suitable for remotely operated platforms; a method that integrates the wireless signal strength into control architecture suitable for remotely operated platforms; [6] the combination of an UAV with a mounted camera opens a wide range of applications like exploration of disaster scenarios [11].

The irruption of low-cost Unmanned Air Vehicles (UAV) has led to an explosion in its interest, further embraced by the service and rescue robotics community [8]. Multiple UAV coordination [9] and planning [10] are nowadays active research areas. Remarkably, the combination of an UAV with a mounted camera opens a wide range of applications like exploration of disaster scenarios [11].

Both UAVs and ground mobile manipulator (MM) are widely used in teleoperation [12], [13]. Works addressing the coordination with or without teleoperation of combined ground mobile robots and UAVs can be found for group coordination [14], visual servoing [15], and UAV landing on a mobile robot with [16] and without the tracking of a ground mobile robot [17], [18]. To the authors knowledge, no work combines a ground mobile manipulator with an UAV, not to mention that none use the UAV camera to guide the teleoperation of a robot end-effector.

Workspace mapping algorithms are a key component in teleoperation systems. They map the input commands of the operator, usually from a haptic device or a joystick, to the robot workspace. This mapping is usually done in the position dimension, using a mapping from the haptic to the robot position (position-position mapping), a position-linear velocity mapping, or an hybrid mapping, and it has to account for the scaling and the unbound workspace for mobile robots [19]– [22]. Dynamic point of view deals with the orientation and can also be found for robot teleoperation [23] and in the exploration of virtual scenarios [24].

The contribution of this work is twofold. The first one is to present a novel decoupled position and orientation workspace mapping algorithm that allows an operator to command the Tool Center Point (TCP) of a MM while simultaneously and continuously changing the point-of-view of a free-flying camera from which the operator monitors the scene. Second, the feasibility of this approach has been evaluated by a study conducted in a virtual framework where users have been asked to perform a highly demanding task; the performance and workload results after four days of training are presented.

The rest of this paper is organized as follows. Section 2 introduces the workspace mapping. In section 3 the robot used in this work is presented. The user study is analysed in Section 4. Finally, the conclusions are exposed in Section 5.

II. THE TELEOPERATION APPROACH

A. Nomenclature

Several frames are defined in this work: 
mo, the master inertial frame, a.k.a., the frame of the origin of the haptic workspace; 
mi, the frame of the tip of the haptic device; 
s, the remote inertial frame, a.k.a., the MM workspace frame; 
sr, the frame of the origin of the haptic workspace in the remote inertial frame; 
si, the frame of the haptic tip in the remote inertial frame; 
c, the frame attached to the free-flying camera; and 
r, the frame of the robot TCP.

Let \( T_{i}^{j} \in SE(3) \) be an homogeneous transformation matrix, composed of a rotation matrix \( R_{i}^{j} \in SO(3) \) and a translation \( p_{i}^{j} \in \mathbb{R}^{3} \), that expresses the frame \( i \) w.r.t. the frame \( j \), and can...
be used to represent a position vector expressed in a frame $i$ into a new frame $j$, by pre-multiplication.

During the teleoperation, $T_{mi}^{so}$ contains the information of the user command to the system through the haptic. $p_{mi}^{so}$ is scaled by $K_S = \text{diag}(k_{Sx}, k_{Sy}, k_{Sz})$ as $p_{mi}^{si} = K_S p_{mi}^{so}$. The different scalings depending on the $X$, $Y$ or $Z$ dimension is convenient since allow the adjustment between the haptic and manipulator workspaces. $T_{si}^{so}$ corresponds to the user haptic input in the MM workspace, composed of the translation $K_S p_{mi}^{so}$ and the rotation $R_{mi}^{so}$.

Noticing that $T_{si} = T_{so} T_{si}^{so}$ then:

$$
p_{si}^{so} = p_{so}^{si} + R_{so}^{si} K_S p_{mi}^{so}
R_{si}^{so} = R_{so}^{si} R_{mi}^{so}
$$

(1)

Also from Fig. 1 note that:

$$
T_{ci} = T_{so} T_{si}^{so}
$$

(2)

where $T_{ci}^{so}$ is constant and couples the frames $c$ and $so$.

Finally, $T_{ci}^{so}$ is obtained by computing the inverse kinematics of the mobile manipulator, as explained in Sect. III.

B. Proposed solution

The proposed approach adapts the solution in [25] to the geometry of a mobile manipulator. The novelty is that this technique is applied not only to the position, but also to the orientation by using a free-flying camera.

1) The mobile manipulator: Regarding the position, when the user commands the tip of the haptic device inside the haptic workspace, TCP of the mobile manipulator is commanded using a position-position relation between the haptic tip position and the robot TCP. When the boundary of the haptic workspace, roughly defined as a thick front bordering the external workspace boundary, is reached, a position-linear velocity component is added to the previous position-position. This allows the user to command the mobile platform of the robot in an unbounded workspace by using a haptic device (which has a bounded workspace) while simultaneously commanding the robot TCP.

To adapt [25] to the workspace of mobile manipulator a "cylindrical bubble" is used as the haptic workspace (Figs. 1-2), the Z position is commanded by a position-position map (Eq. 1), while the position-linear velocity is activated in the XY plane when the distance from the haptic tip to the Z axis of the frame $mo$, $D_V$, is higher than a pre-set radius $R_V$. This can be imposed by:

$$
\dot{p}_{si} = \left(1 - \frac{R_V}{D_V}\right) k_V u_{i_{XY}}^{so} = \left(1 - \frac{R_V}{D_V}\right) k_V R_{so}^{si} u_{i_{XY}}^{mo}
$$

(3)

with $k_V > 0$, the maximum allowed velocity; $u_{i_{XY}}^{mo}$, the unit vector of $p_{mi}^{mo} = [p_{mi}^{mo}, p_{mi}^{mo}, 0]^T$; and $D_V = \|p_{mi}^{mo}\|$. The term $(1 - R_V/D_V)$ allows for velocities from zero ($D_V = R_V$) to the maximum defined in $k_V$ ($D_V \rightarrow \infty$). Thus, the farther away the haptic tip is from the vertical axis at its workspace, the fastest the MM will move in the same direction.

Differentiating and unifying translation in Eq. 1 and Eq. 3:

$$
\dot{p}_{si}^{so} = \left(1 - \frac{R_V}{D_V}\right) k_V R_{so}^{si} u_{i_{XY}}^{mo} x_{[R_V, \infty)}(D_V) + R_{so}^{si} K_S \dot{p}_{mi}^{so}
$$

(4)

where $x_C(x)$ is the indicator function.

Additionally, a force $F$ is applied to the user in position-linear velocity mode, with $F$ pointing towards the Z axis of the haptic workspace, and proportional to the distance from the Z axis. A damping has been added so that if the user releases the haptic its tip moves towards the position-position area, thus stopping the platform motion. Then $F$ becomes:

$$
F = - \left(1 - \frac{R_V}{D_V}\right) K_f R_{so}^{si} \left(K_S u_{i_{XY}}^{mo} - K_d \dot{u}_{i_{XY}}^{mo}\right) x_{[R_V, \infty)}(D_V)
$$

2) The free-flying camera: The free-flying-camera permits the use of the position-linear velocity mode on the orientation, enabling an orientation-angular velocity map.

The orientation-angular velocity control is activated when the haptic orientation surpasses a predefined boundary of its workspace, defined as the ZYX intrinsic Euler angles:

$$
R_{mi}^{so} = (r_{ij}) = \begin{pmatrix}
 c_\alpha c_\gamma & s_\alpha s_\gamma & -c_\alpha c_\gamma + s_\alpha s_\gamma c_\delta \\
c_\alpha s_\gamma & c_\alpha c_\gamma s_\alpha + s_\alpha c_\gamma & c_\alpha s_\gamma s_\alpha - c_\gamma c_\delta \\
-s_\beta & c_\beta s_\gamma & c_\beta c_\gamma
\end{pmatrix}
$$

(5)

with $s_a = \sin a$ and $c_a = \cos a$. The Euler angles can be retrieved as $\alpha = \arctan2(r_{32}, r_{33})$, $\beta = \arctan2\left(r_{31}, \sqrt{r_{11}^2 + r_{21}^2}\right)$, and $\gamma = \arctan2(r_{21}, r_{11})$. 

\[\text{Figure 1. The teleoperation system frames. Top: the haptic (a Phantom Omni) and the screen with the image from the UA V (a Parrot AR.Drone 2.0) camera, as seen by the operator. Bottom: The haptic workspace (in yellow) is mapped into the robot workspace. The red arrows depict the user command. The Barcelona Mobile Manipulator 1 (BMM1) can be seen at the right side.}\]
The orientation-angular velocity is triggered when $|\gamma| > \gamma_b$. Then, the frame $R^s_{so}$ is rotated around the vertical axis, $\hat{Z}$, according to the rotation direction specified by the user and with an angular velocity proportional to the amount of penetration into the orientation-angular velocity area (pink area in Fig. 2) in the haptic workspace.

By differentiating the orientation part of Eq. 2:

$$\omega^s_{so} = \omega^s_{so} + R^s_{so}\omega^{so}$$

with each $\omega$ corresponding to an angular velocity, it can be seen that the camera orientation can be commanded through the platform frame, $so$. Thus, the orientation-angular velocity mode can be imposed by:

$$\omega^s_{so} = \text{sign}(\gamma) k_R \left( 1 - \frac{\gamma_b}{|\gamma|} \right) \hat{Z} \chi_{[\gamma_b,\infty)}(|\gamma|)$$  \hspace{1cm} (6)

where \(\text{sign}(\gamma)\) accounts for the rotation direction around the $\hat{Z}$ axis, the term \((1 - \gamma_b/|\gamma|)\) allows for angular velocities from zero to the maximum defined by $k_R$ ($|\gamma| \rightarrow \infty$), and the indicator function, $\chi_{[\gamma_b,\infty)}(|\gamma|)$, becomes one when $|\gamma| \geq \gamma_b$.

Finally, the operator, by varying the values of $R_V$ in Eq. 4 and $\gamma$ in Eq. 6 through the haptic device can change the platform linear and angular velocities, $p^s_{so}$ and $\omega^s_{so}$. Integrating both variables, $T^s_{so}$ can be computed, and so can the new camera position and orientation, $T^o_c$, using Eq. 2.

C. Implementation

The algorithm can be summarized as:

1) $p^m_{so} = [p^m_{sox}, p^m_{soy}, p^m_{soz}]^T$
2) $D_V = ||p^m_{so}||$
3) if $D_V > R_V$:
   $$p^s_{so} = p^m_{so} + \left( 1 - \frac{R_v}{D_V} \right) k_V R^s_{so} p^{mso}_{XY} \frac{\Delta t}{D_V}$$

III. THE BARCELONA MOBILE MANIPULATOR

The Barcelona Mobile Manipulator 1 (BMM1) is an omnidirectional mobile platform with spherical wheels carrying an industrial serial robot (Fig. 1); thus, it allows a holonomic behaviour. It is composed by a main body of 138 kg and 1x0.78x0.708 m³. The platform possesses 3 degrees of freedom (DoF) in the plane: two independent translations and a rotation around its vertical axis. The arm manipulator is a Kuka LWR 4+, which has 7 DoF. Thus, overall, the BMM1 has 10 DoF. EtherCat is used for the low level control of the wheels while Orocos is used for the high level control in a Linux Xenomai environment over a PC with four Intel Core i5 CPUs at 3.1GHz. The BMM1 can be commanded from an external PC through a local Wi-Fi network that relies on ROS for the communications. [20] presents a detailed kinematic description of the platform.

The idea behind this work is to ultimately implement the proposed approach to teleoperate the BMM1 in real-time while the camera of a Parrot AR.Drone 2.0 is sending the video of the scene to the operator. The camera can be commanded through the quadrotor inputs: the desired linear and $Z$ angular velocities. The AR.Drone also outputs its estimated position and orientation.

Inverse kinematics

A CLIK algorithm [27] has been used to command the TCP:

$$\dot{q} = \mathcal{J}^+ \begin{bmatrix} p^n_{o} + K_P e_p \\ L^{-1} (L^T \omega^n_{so} + K_o e_o) \end{bmatrix}$$  \hspace{1cm} (7)

where:

- $p^n_{o}$: Desired position
- $K_P$: Position gain
- $K_o$: Orientation gain
- $L$: Link gain matrix
- $\omega^n_{so}$: Desired orientation
- $\omega^n_{so}$: Actual orientation
- $\Delta t$: Time step
- $\mathcal{J}$: Jacobian matrix
- $\mathcal{J}^+$: Pseudoinverse of $\mathcal{J}$
• \( e_p = p_s^t - p_r^t. \)
• \( e_n = \frac{1}{2} \left[ n_R(q) \times n_D(q) + s_R(q) \times s_D + a_R(q) \times a_D \right] \); with \( n, s, \) and \( a \) the columns of \( R = R_r^t \) and \( D = R_r^t. \)
• \( L = -\frac{1}{2} \left[ S(n_D)S(n_R) + S(s_D)S(s_R) + S(a_D)S(a_R) \right]. \)

Once \( q \) is obtained by integrating \( \dot{q} \) from the initial conditions, \( T^p_r(q) \) can be computed through the direct kinematics. This algorithm allows an online fast tracking of the TCP in position and orientation. Due to the robot redundancy, local minima is not a problem, even while considering the joint limits. The SVD decomposition has been used to compute the pseudoinverse in Eq. 7, which allows the removal of the limits. The SVD decomposition has been used to compute the joint velocities within the set of available solutions has minimum norm.

IV. THE USER STUDY

A study has been conducted in order to evaluate:
• The dexterity in the use of the teleoperation system.
• The load of the teleoperation system on the operator.
• The progression in the learning and improvement in the teleoperation skills.

A. Description

A virtual scenario has been designed where the operator has had to track a moving sphere with the TCP of a virtual BMM1 while simultaneously keeping a yellow spot on the sphere facing towards the camera (Fig. 4). In order to test the proposed teleoperation system in ideal conditions, time delays between the MM and the UAV and noise in the quadrotor pose measurement have not been considered. As the sphere has been randomly and smoothly translating and rotating around the Z axis (as the spot orientation), a high degree of coordination between the TCP and the camera view orientation is needed. An index has been obtained for each trial, which value is 0 for an insufficient performance and 1 when perfect. The index has been computed as the mean of:

\[
I = \begin{cases} 
0 & \text{if } \alpha > \alpha_0 \\
\left(1 - \frac{\alpha}{\alpha_0}\right)e^{-2(D-D_0)} & \text{if } \alpha \leq \alpha_0 \text{ and } D > D_0 \\
1 - \frac{\alpha}{\alpha_0} & \text{if } \alpha \leq \alpha_0 \text{ and } D \leq D_0
\end{cases}
\]  

(8)

at every sample time during the experiment, and where:
• \( \alpha \) is the angle between the normal on the center of the spot and the projection on the XY plane of the vector from the center of the sphere to the robot TCP.
• \( \alpha_0 \) defines the range of values of \( \alpha \) where its contribution on the performance of the teleoperation is computed. If \( \alpha \) is bigger than \( \alpha_0 \) the teleoperation performance is considered insufficient (\( I = 0 \)).
• \( D \) is the euclidean distance from the center of the sphere to the TCP position of the mobile manipulator.
• \( D_0 \) is a radius of a sphere which center coincides with the tracked sphere center. If the TCP of the MM is inside this sphere the performance of the teleoperation is considered perfect regarding the translation.

The values \( \alpha_0 = \frac{\pi}{2} \) rad and \( D_0 = 0.4 \) m have been used. \( p^o_c \) has been set to \((3,0,5)\) and \( R^o_c \) with the Z axis pointing towards the so frame origin and the X axis as horizontal.

The random motion of the sphere has been achieved using a potential field. An attractive potential field has been randomly generated every 5 to 8 seconds by placing an attractive pole in the scenario, defined as a position in the XY plane and an orientation around its Z axis. A repulsive potential field has been set at the scenario limits, so that the sphere would stay inside the limits of the scenario.

The experiment has been conducted by 8 persons (average age, 27.6 years) with little experience with haptic devices. Four sessions have been conducted per user in consecutive days. In each session five experiments have been executed. Each experiment lasted 1 min. After each set of experiments the users have completed a NASA Task-Load-Index (NASA-TLX) [28] test to measure the perceived workload.

The first session has included a description of the teleoperation system and the experiment with a duration of 20 minutes. The overall duration of the next sessions has been around 15 minutes per participant and day.

B. Implementation

The teleoperation system has been tested in a Intel Core i7 at 2.80GHz with Debian. The software has been implemented in C++ using ROS and the Kautham Project [29].

C. Results

Two main set of results have been obtained, one regarding the dexterity in the teleoperation of the users and the others regarding the workload on the operators during the task.

The dexterity performance scores

Regarding the dexterity level during the teleoperation, computed using Eq. [8] first the Saphiro-Wilk normality test [30] has been performed on the data (Fig. [5]). The data consisted on 40 values (5 experiments per participant) per day. The data has been found to be non normal for each of the days, mainly due to the skewness of the data (0.13 for the first training day and -1.63 for the fourth), the bounds on the underlying distribution and some isolated values.

Considering the operators on their first day as the population and on the fourth as the sample, the null hypothesis "the
operators do not improve their teleoperation skills after four days of training in the teleoperation of the combined mobile manipulator and free-flying camera teleoperation system” can be rejected by doing a Wilcoxon Rank Sum Test [31], suitable for non normal data. The ranksum of the Wilcoxon test is 937, so the p-value is almost zero.

The difference between two measures of the discrete distributions has been compared to have an intuitive idea of the magnitude of the increase in dexterity. The averages of the scores (41.77 for day 1, and 72.80 for day 4) are compared also against the dexterity of an expert user. Taking the mean of the score of the expert user (90.12) as an upper bound, and the lowest score (9.70) as a lower, the percentage of dexterity gain obtained can be computed as:

$$\frac{72.80 - 41.77}{90.12 - 9.70} = 38.58\%$$

The NASA-TLX scores

The second evaluation has focused on the operator workload and the difference between untrained and trained operators (Fig. 6). The workload has been evaluated using the NASA-TLX [28], a subjective assessment tool to rate perceived workload in order to assess a task effectiveness.

A Saphiro-Wilk test has been computed on the NASA-TLX scores using the data of the first and fourth training days to verify its normality. With a p-value of 0.10 the W values obtained have been 0.9079 and 0.9446 for day 1 and day 4, thus the data can be treated as normal. The NASA-TLX scores mean and standard deviation on the first day have been 52.54 (from a minimum of 0 to a maximum of 100) and 15.75, respectively. After the training, the average and standard deviations decreased to 41.79 and 6.18 (Fig. 7).

With this data the null hypothesis ”the workload during the teleoperation on the operators does not decrease after four days of training in the teleoperation of the combined mobile manipulator and free-flying camera teleoperation system” can be rejected with a probability of 97.32% (Z = -1.93). The mean value of the workload is between 30.88 and 52.70 with a 95% confidence, almost in its totality under the average workload of the untrained operators (52.54).

Contrary to the dexterity scores, the underlying normality of the NASA-TLX data permits its interpretation in a simple manner through the effect size (ES) [32], [33]. The ES eases the interpretation of the results of a statistical analysis and has been widely studied in the statistics community. The ES of the difference in the workload during the teleoperation between the untrained operators and after 4 days of training is -0.91, corresponding to a 81.79% probability that the average trained user will feel less workload than a random untrained operator, and between -1.94 and 0.12 with a 95% confidence (45.15% to 97.36%, respectively). Two interpretations can light up its significance, the rank interpretation and the Common Language Effect Size (CLES) [34]:

- **Rank** = 18.21. In a untrained group of 100 members the rank 18 indicates that only eighteen untrained members feel less workload than the average trained operator. In an hypothetic group of 100 operators, after the training, half of the operators feel less workload than 82 of the 100 untrained operators, thus improving a 64%.
- **CLES** = 73.78%: is a 73.78% probability that a trained operator will perceive less workload than an untrained.

V. CONCLUSIONS

A novel teleoperation system that permits the simultaneous command of a ground mobile manipulator and a free flying camera using a single haptic device is proposed.
A user study has been conducted to evaluate the dexterity in the use of the teleoperation system and the workload on the operator before and after proper training. The data analysis shows that training increases the operator performance while also lowering the workload demand on the operator.

The results point out that the system complexity is not an obstacle to command it; thus, the proposed solution can be regarded as a suitable mean to simultaneously command a mobile manipulator and a free-flying camera.

Future work will contemplate a study with more users and a comparison of the current experiment with a two stick control.

ACKNOWLEDGMENT

This work was partially supported by the Spanish Government through the projects DPI2013-40882-P and DPI2014-57757-R and the predoctoral grants BES-2012-054899 and BES-2012-059884.

REFERENCES


