

Vision-Based Teleoperation of Unmanned Aerial and Ground Vehicles

ChangSu Ha and Dongjun Lee

Abstract—We present a novel vision-based teleoperation control framework for a team of an unmanned aerial vehicle (UAV) and an unmanned ground vehicle (UGV). Our control law allows a remote human user to teleoperate the team with some useful haptic feedback, while also ensuring the UAV-UGV coordination via the camera installed on the UAV (and seeing the UGV) and the velocity limitation of the UGV. For this, we first elucidate a geometric condition for the UAV and UGV velocities to ensure the UAV-UGV coordination by driving the image feature of the UGV to converge to a desired one on the image surface, while also guaranteeing that the UGV's velocity, which is often much slower than that of the UAV, is under a certain specified bound. The UAV is then tele-controlled to track the teleoperation command as close as possible, yet, only to the extent permissible by this geometric condition. Simulation is performed to illustrate the theory.

I. INTRODUCTION

Vision-based control or visual servoing of robots has grown in its importance, since, in many applications, cameras provide affordable and efficient means for recognizing and sensing the surrounding environments and/or the manipulated/tracked objects. For mobile robots, the usage of the camera is even more promising (and thus demanded), particularly for outdoor applications, for which the frequently-used, yet, ground-fixed and expensive, motion capture systems (e.g., VICON®) are usually infeasible to deploy. Recently, visual-servoing of the unmanned aerial vehicles (UAVs) has also received much attention [1], [2] to truly make the UAV as a flexible robotic platform untethered from ground-bound measurement systems.

In this paper, we propose a novel teleoperation control framework for a team of an UAV and an UGV (unmanned ground vehicle), with a camera attached on the UAV to enforce the UAV-UGV coordination. More precisely, adopting the spherical camera description of [1], [3], we first elucidate a geometric condition for the UAV's and UGV's velocities to ensure that the image feature p of the UGV converges to a desired one p_d on the image surface (i.e., UAV-UGV coordination), even when the UGV's velocity is limited by a certain bound. We then dictate the UAV to track the (velocity) teleoperation command from the remote user as close as possible, yet, only to the extent permissible by this geometric condition, thereby, maintaining the UAV-UGV coordination with a higher-priority while still allowing the user to teleoperate the UAV-UGV team. We also provide the remote human user with some

useful haptic feedback on top of visual information (either from another front-facing onboard camera attached on the UAV or from fixed external cameras providing global vision information of the operation site). For this, we utilize our recently proposed passive set-position modulation (PSPM [4], [5]) to guarantee the passivity (i.e., stability) of the (bilateral) haptic feedback loop.

There are numerous results on the vision-based control of the UAVs for the pose stabilization and some tracking control relative to a fixed target (or ground). A backstepping stabilization controller using an onboard camera was proposed in [1], while two onboard cameras, one on the UAV and the other on the ground, were used in [2]. Optical flow are used to estimate horizontal velocity for stable hovering of eight rotor uav in [6]. Four kinematic IBVS (image-based visual-servoing) controllers were experimentally compared for the Cartesian positioning task in [7]. A (destabilizing) positive image feature feedback with (stabilizing) virtual spring approach was proposed in [8] to control the position and orientation of the UAV. All these results [1], [2], [6], [7], [8], yet, are about controlling the UAV relative to a fixed ground (or objects).

On the other hand, there are also many results for controlling the UAV relative to a moving target. Visual servoing for a UGV was proposed with an overhead camera that may be mounted on a UAV in [9]. However, the camera is assumed to be wide enough not to move to track the UGV so that camera motion is able to be neglected. In [10], position optimal estimation was suggested for cooperative strategy of multi-UAVs to tracking moving target, yet, vision sensor is only used for finding direction vector from UAVs to the target. Similarly, Particle filter and extended Kalman filter are used for estimation of moving target's location on camera surface in [11] and [12], respectively. A vision-based control law for hovering and autonomous landing on a moving platform using optical flow was proposed in [13], while a vision-based algorithm using a density-based object representation was proposed in [14] to chase a moving target. In [15], a mobile robot is used as a moving target and an UAV tracks this mobile robot and lands on it autonomously. However, in these results [9], [10], [11], [12], [13], [14], [16], [15], the UAV *unilaterally* tracks the UGV, with no feedback from the UAV to the UGV. Differently to these, our teleoperation control relies on the interplay between the UAV's and the UGV's motions to maintain the UAV-UGV coordination while also taking into account the slower dynamics of the UGV (i.e., velocity bound).

In contrast to these *autonomous* vision-based control results for the UAV [1], [2], [7], [8], [13], [14], [16], [15], in this paper, we advocate vision-based *teleoperation* for the UAV-UGV team operation, since: 1) when the operation site is uncertain, unknown, or unexplored (e.g., search and rescue

The authors are with the School of Mechanical & Aerospace Engineering and IAMD, Seoul National University, Seoul, Republic of Korea, 151-744. Email: {changsuha,djlee}@snu.ac.kr

Research supported in part by the National Research Foundation (NRF) funded by the Ministry of Education, Science & Technology (MEST) of Korea (2012-R1A2A2A0-1015797) and the Basic Science Research Lab program of the National Research Foundation (NRF) by the Korea Government (MEST) (2009-0087640).

in a disaster area), a fully-autonomous control is typically infeasible and teleoperation is often only a viable option; and 2) some tasks, which would be quite difficult when performed autonomously (e.g., SLAM [17], or motion/task planning), can be done relatively easily by the remote human operators using their sensory data (e.g., visual information and haptic feedback) and intelligent decision. There are several (recent) results on the teleoperation of a single or multiple UAVs (e.g., [5], [18]), yet, no UGVs are considered there. To our knowledge, our result in this paper is one of the very first results on the vision-based teleoperation of the UAV-UGV team, which would allow for many useful/interesting applications by exploiting their heterogeneous and complementary capabilities in an integrative way. The rest of this paper is structured as follows. Sec. II contains some preliminary materials on the UGV/UAV description and the spherical camera model [1], [3]. Sec. III presents our main result: a vision-based haptic teleoperation control law for the UAV-UGV team while ensuring the UAV-UGV coordination and the UGV's velocity bound. Simulation results are then given in Sec. IV, and some concluding remarks in Sec. V.

II. PRELIMINARY

A. Unmanned Aerial and Ground Vehicles

Our vision-based teleoperation control law specifies the high-level desired velocity commands for the UAV and UGV. We then assume that the UAV and UGV possess some (arbitrary) well-functioning low-level controllers so that these velocity commands can be faithfully tracked by them. For instance, we may use the schemes [19], [20] for the quadrotor-type UAVs, while that in [21] for the unicycle-type wheeled UGV (with some modification/simplification). Similar separation of the high-level and low-level controls and the availability of the low-level controls for UAV and UGV were also used in [14].

With this low-level controls assumed both for the UAV and UGV, let us denote the position of the UGV relative to the inertial frame $\{\mathcal{I}\}$ by $\bar{P} := [\bar{P}_1, \bar{P}_2, \bar{P}_3]^T \in \mathbb{R}^3$, and that of the UAV by $x := [x_1, x_2, x_3]^T \in \mathbb{R}^3$. In this paper, we assume that UGV is moving on a flat ground so that $\bar{P}_3 = 0$. We also assume that the camera is installed on the bottom of the UAV, which is directed downward from the UAV's body-frame $\{\mathcal{B}\}$ (i.e., along $e_3^{\mathcal{B}}$ -direction), through which the UAV can see the UGV. See Fig. 1. The pose of this camera then is given by $R \in \text{SO}(3)$, the rotation matrix of the body-frame $\{\mathcal{B}\}$ relative to the inertial-frame $\{\mathcal{I}\}$. Let us denote the angular rate of the camera represented in the body-frame $\{\mathcal{B}\}$ by $w \in \text{so}(3)$. This w is related to R s.t.

$$\dot{R} = RS(w) \quad (1)$$

where $S(\star)$ is the skew-symmetric operator s.t. $S(x)y = x \times y$, $\forall x, y \in \mathbb{R}^3$.

B. Spherical Camera

Eyes are very effective and the main sensor for humans to avoid obstacles, recognize environments, and navigate in environments. To imitate the function of human eyes, cameras have been used. Some types of camera are: fish-eye, catadioptric,

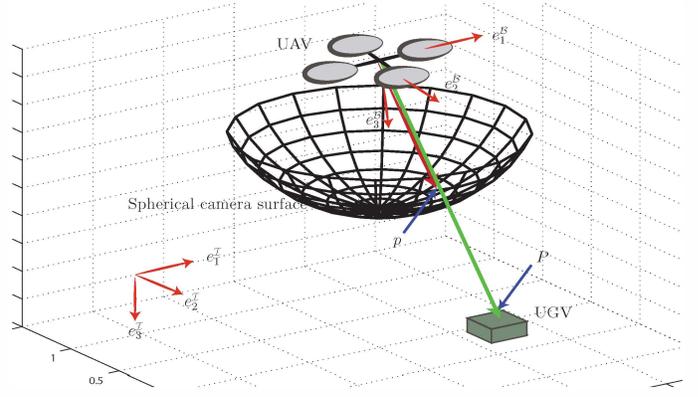


Fig. 1. UAV, UGV and spherical camera image surface.

and spherical camera. Among them, in this paper, we choose the spherical camera [1], since it provides a natural geometry when associated with the $\text{SO}(3)$ rotational motion of the UAV as shown in the following Sec. III.

With this spherical camera, we then have the image geometry as shown in Fig. 1, where the image surface is given by the sphere with the radius of the focal length f . In this paper, we set this $f = 1$ (i.e. unit radius sphere). On this image surface, we then have the image feature $p := [p_1, p_2, p_3]^T \in \mathbb{R}^3$ with $\|p\| = 1$ of the UGV as measured by the camera in the body-frame $\{\mathcal{B}\}$. This p is then given by

$$p := \frac{P}{r(P)} \quad (2)$$

where $P := [P_1, P_2, P_3]^T \in \mathbb{R}^3$ is the position of the UGV as seen from the UAV in the body-frame $\{\mathcal{B}\}$, and $r(P)$ is the relative depth of the spherical camera defined by $r(P) := \|P\|/f = \|P\|$, with $f = 1$.

The following facts will be used later: from $p^T p = 1$,

$$p^T \dot{p} = 0, \quad \frac{\partial r}{\partial P} = p^T \quad (3)$$

where, for the second equality, we use $p^T P = r(P)$ from (2). We can also show that

$$P = R^T(\bar{P} - x), \quad V_P = R^T(\dot{\bar{P}} - \dot{x}) \quad (4)$$

where \bar{P} and x are the UGV's and UAV's positions measured in the inertial frame $\{\mathcal{I}\}$; and also $V_P \in \mathbb{R}^3$ is the relative velocity between the UGV and the UAV as measured in the body-frame $\{\mathcal{B}\}$.

Using (1)-(3), we can write the evolution of the image feature p on the image surface s.t.

$$\dot{p} = \frac{d}{dt} \left(\frac{P}{r(P)} \right) = -w \times p + \frac{(I - pp^T)}{r(P)} V_P \quad (5)$$

where we use (3) and also the fact that $w \times p$ is orthogonal to p . For details, please see [1]. Here, observe that the right hand side of (5) is contained in the null-space of p (i.e., satisfying $p^T \dot{p} = 0$).

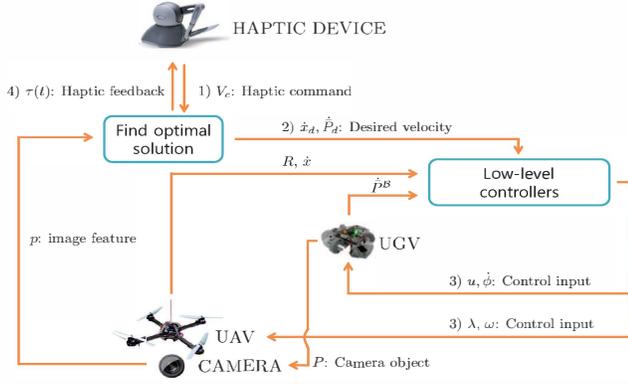


Fig. 2. Vision-based teleoperated coordination control architecture

III. VISION-BASED TELEOPERATION CONTROL DESIGN

In this section, we design our vision-based teleoperation control law, which enforces the coordination between the UAV and UGV with the highest priority (e.g., to prevent the UGV from being lost in the UAV’s camera view) using the camera installed on the UAV, while allowing a human user to (haptically) teleoperate the UAV, whose teleoperation command will be modulated (or compromised) so as not to violate the UAV-UGV coordination given the possible bound of the UGV’s velocity (see (13)).

Fig. 2 shows our control architecture and the information flow within it, where: 1) a human user sends a teleoperation velocity command to the UAV; 2) our vision-based control law computes the desired velocity commands for the UAV and UGV, which allow these UAV and UGV to move according to the teleoperation command as much as possible, yet, only to the extent permissible by the UAV-UGV coordination requirement and the UGV’s velocity bound; 3) the low-level controls then drive the UAV and the UGV to track their respective velocity commands; and 4) the human perceive the state of the UAV and its interaction with UGV via some form of haptic feedback on top of vision information provided by extra cameras. Let us start first with the UAV-UGV coordination control.

A. UAV-UGV Coordination Constraint

For the UAV-UGV coordination without losing generality, we want the UGV to move right below the UAV along the e_3^T -direction in the inertial frame $\{\mathcal{I}\}$. That is, if $R = I$, we want the image feature p of the UGV on the image surface to converge to the desired image feature $p_d = [0; 0; 1]$. However, if the UAV’s attitude changes, which is typically required to incur the UAV’s Cartesian velocity \dot{x} due to its under-actuation [20], [19], the camera view would also rotate, thus, although the UGV is still positioned directly under the UAV along e_3^T -direction as desired, enforcing $p_d = [0; 0; 1]$ w.r.t. the body-frame $\{\mathcal{B}\}$ would require the UAV to deviate from its original position. Thus, following [1], we define p_d to be time-varying in the body-frame $\{\mathcal{B}\}$ s.t.

$$p_d = R^T e_3^T = R^T [0 \ 0 \ 1]^T \quad \text{with} \quad \dot{p}_d = -w \times p_d \quad (6)$$

where we use (1).

Now, define the image feature tracking error $e := p - p_d$. Then, by enforcing $p \rightarrow p_d$, we will be able to achieve the UAV-UGV coordination. This can be achieved if we can drive \dot{p} in (5) to behave according to following desired error dynamics

$$\dot{p} = \dot{p}_d - \gamma(p - p_d) =: u_p.$$

This desired dynamics, however, is not always achievable. This is because, although \dot{p} should be contained within the null-space of p (i.e., $\dot{p} \in \text{null}(p)$ - see (5) or $p^T \dot{p} = 0$ (3)), in general, $u_p \notin \text{null}(p)$. To address this constraint, we modify the above control law u_p to be

$$\dot{p} = \frac{1}{p^T p_d} (I - pp^T) (\dot{p}_d - \gamma(p - p_d)) \quad (7)$$

which is now contained within the null-space of p , with $I - pp^T$ spanning the 2-dimensional $\text{null}(p)$ (i.e., $\text{rank}(I - pp^T) = 2$ with $p^T(I - pp^T) = 0$). This also implies that the modified control law (7) is the projection of u_p on $\text{null}(p)$ with the scaling $1/(p^T p_d)$, which will become 1 when $p \approx p_d$.

Proposition 1 Consider the dynamics of p in (7). Then, if $p^T(0)p_d(0) > 0$, $e(t) = p(t) - p_d(t) \rightarrow 0$.

Proof: First, let us define the error e projected on the null-space $\text{null}(p)$ s.t.

$$e_p := (I - pp^T)(p - p_d) = (I - pp^T)e$$

and define the Lyapunov function W_p s.t. $W_p := \frac{1}{2} e_p^T e_p$. We can then achieve that:

$$\begin{aligned} \dot{W}_p &= e_p^T \dot{e}_p = [(I - pp^T)(p - p_d)]^T \frac{d}{dt} [(I - pp^T)(p - p_d)] \\ &= -\gamma e_p^T e_p \end{aligned} \quad (8)$$

where, for this result, we use the facts that $(I - pp^T)pp^T = 0$ and $(I - pp^T)(I - pp^T) = (I - pp^T)$ from (3) with $p^T p = 1$.

This (8) then shows that the projected error $e_p \rightarrow 0$. Let us then see if the real error $e = p - p_d$ also converges to the origin. For this, note that, from the definition of e_p above, $e_p \rightarrow 0$ means that $p - p_d \rightarrow \lambda p$ with some $\lambda \in \mathbb{R}$, since $\text{rank}(I - pp^T) = 2$. However, since p and p_d are both on the unit image sphere, this condition $(1 - \lambda)p - p_d \rightarrow 0$ can be attained only with $p \rightarrow p_d$ or $p \rightarrow -p_d$ (i.e., antipodal equilibrium).

The antipodal equilibrium (i.e., $p \rightarrow -p_d$), yet, we can rule out. This is because, with $p^T(0)p_d(0) > 0$, to achieve $p \rightarrow -p_d$, p is required to rotate away from p_d such that, at some point, it must become orthogonal to p_d with $p^T p_d = 0$. At this point of $p^T p_d = 0$, the projected error e_p on the null-space of p will attain the maximum value with $\|e_p\| = 1$. Yet, with $p^T(0)p_d(0) > 0$, we have $\|e_p(0)\| < 1$, and, moreover, the above Lyapunov analysis (8) shows that $\|e_p(t)\|$ is non-increasing. This then implies that $p \rightarrow -p_d$ is impossible and we only have $p \rightarrow p_d$, i.e., $e(t) \rightarrow 0$. This completes the proof. ■

This Prop. 1 then shows that the control law (7) will guarantee the image feature tracking $p \rightarrow p_d$, thereby, enforce the UAV-UGV coordination. We now convert this control law

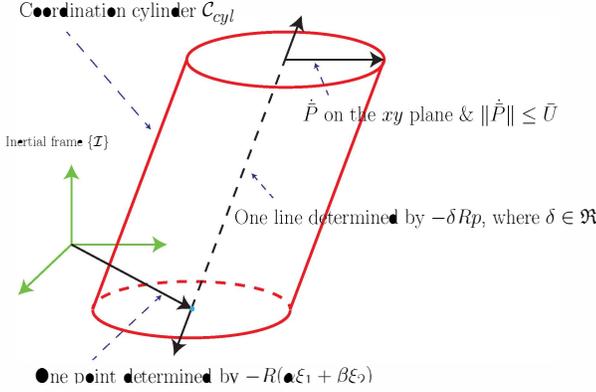


Fig. 3. Geometry of \dot{x}_d , \dot{P}_d and the coordination cylinder C_{cyl} .

(7) into a condition on the UAV velocity \dot{x} and the UGV velocity \dot{P} . For this, equating (7) with (5), we can obtain

$$(I - pp^T)V_P = \underbrace{r(P) \left[\omega \times p + \frac{1}{p^T p_d} (I - pp^T)(\dot{p}_d - \gamma(p - p_d)) \right]}_{=: f(p, p_d, \omega, r(P))} \quad (9)$$

where $V_P = R^T(\dot{P} - \dot{x})$ as defined in (4) and $\dot{p}_d = -w \times p_d$ as stated in (6).

Here, since $(I - pp^T)$ spans the 2-dimensional null-space of p and the right hand side of (9) is contained within this $\text{null}(p)$, we can write the solution V_P for (9) by

$$V_P = \alpha \xi_1 + \beta \xi_2 + \delta p \quad (10)$$

where $\xi_1, \xi_2 \in \mathbb{R}^3$ are the orthogonal unitary basis of $\text{null}(p)$ with $\|\xi_i\| = 1$ and $\xi_1^T \xi_2 = 0$, and $\delta \in \mathbb{R}$ can be any arbitrary number. By injecting this expression of V_P into (9), we then have: with $p^T \xi_i = 0$,

$$\alpha \xi_1 + \beta \xi_2 = f(p, p_d, \omega, r(P))$$

from which we can obtain α, β s.t.

$$\alpha = \xi_1^T f(p, p_d, \omega, r(P)), \quad \beta = \xi_2^T f(p, p_d, \omega, r(P)) \quad (11)$$

since ξ_i is the orthonormal basis for $\text{null}(p)$.

Then, using (10) with (4), we can decode the control action (7) into the coordination requirement for the UAV's velocity \dot{x} and that of the UGV \dot{P} in the inertial frame $\{\mathcal{I}\}$ s.t.,

$$\dot{x} = -R(\alpha \xi_1 + \beta \xi_2 + \delta p) + \dot{P} \quad (12)$$

where ξ_i is again the basis for $\text{null}(p)$, $\delta \in \mathbb{R}$ can be any arbitrary number, and R is the rotation matrix of the UAV. For (12), let us also assume that the possible velocity of the UGV \dot{P} is bounded s.t.,

$$\|\dot{P}\| \leq \bar{U}. \quad (13)$$

Recall also that the UGV's motion is planar, that is, $\dot{P} = [\dot{P}_1; \dot{P}_2; \dot{P}_3] \in \mathbb{R}^3$ with $\dot{P}_3 = 0$.

Taking these into account, we can then construct *coordination cylinder* C_{cyl} as shown in Fig. 3 in the inertial

frame $\{\mathcal{I}\}$, where the oblique cylinder's center is located at $-R(\alpha \xi_1 + \beta \xi_2)$, its center-axis along the vector Rp , and its radius on the inertial-frame's (e_1^T, e_2^T) -plane given by \bar{U} . Note that, as long as \dot{x} of the UAV is contained within this coordination cylinder C_{cyl} (i.e., satisfying (12)), we can find the UGV's velocity \dot{P} in the inertial-frame $\{\mathcal{I}\}$ (i.e., given by the (e_1^T, e_2^T) -planar vector from the center-line to this \dot{x}), with which the UGV can follow the UAV (flying with \dot{x}) to maintain the UAV-UGV coordination, under the UGV's velocity bound (13). Here, note that both \dot{x} and \dot{P} are in the inertial frame $\{\mathcal{I}\}$, with \dot{P} being the (e_1^T, e_2^T) -planar vector in $\{\mathcal{I}\}$.

This then says that, if the human's teleoperation command dictates the UAV to fly with $\dot{x} \in C_{cyl}$, we can achieve the desired teleoperation behavior and also the UAV-UGV coordination at the same time under the UGV's velocity bound constraint (13). Yet, if the human's command requires the UAV to fly with the velocity outside of this coordination cylinder C_{cyl} , the intended teleoperation behavior and the UAV-UGV coordination cannot be achieved at the same time. In the next Sec. III-B, we consider this problem, i.e., how to incorporate the teleoperation while ensuring the UAV-UGV coordination (i.e., $p \rightarrow p_d$) under the UGV's velocity limitation (13).

B. Teleoperation Control Design

Let us denote the teleoperation velocity command for the UAV by $\dot{x}_c \in \mathbb{R}^3$. This teleoperation command \dot{x}_c may or may not be within the coordination cylinder C_{cyl} . To address the possible conflict between this teleoperation command and the UAV-UGV coordination, here, with a higher-priority given on the UAV-UGV coordination (i.e., to keep $p \rightarrow p_d$ in the camera view), we define the (high-level) desired velocity commands for the UAV (i.e., \dot{x}_d) and the UGV (i.e., \dot{P}_d) in such a way that \dot{x}_d is designed as close to \dot{x}_c as possible, yet, only to the extent allowable by the UAV-UGV coordination (12) and the UGV's velocity bound (13). More precisely, given the teleoperation command \dot{x}_c , we choose \dot{x}_d and \dot{P}_d s.t.

$$\begin{aligned} \min_{\dot{x}_d} \quad & \|\dot{x}_c - \dot{x}_d\| \\ \text{subj.} \quad & \dot{x}_d = -R(\alpha \xi_1 + \beta \xi_2 + \delta p) + \dot{P}_d \\ & \|\dot{P}_d\| \leq \bar{U}, \quad [\dot{P}_d]_3 = 0, \quad \delta \in \mathbb{R} \end{aligned} \quad (14)$$

where $[\star]_3$ is the e_3 -component of $\star \in \mathbb{R}^3$ (i.e., $[\star]_3 := \star_3$ for $\star = [\star_1, \star_2, \star_3]$), the second line ensures that $\dot{x}_d \in C_{cyl}$ (i.e., UAV-UGV coordination), and the third line the UGV's velocity bound constraint (13). Here, since the cost function and the constraints are all convex, the solution (\dot{x}_d, \dot{P}_d) exists and is unique. In the following, we provide explicit solution of (14) when 1) $\dot{x}_c \in C_{cyl}$; and $\dot{x}_c \notin C_{cyl}$.

1) When $\dot{x}_c \in C_{cyl}$: Since $\dot{x}_c \in C_{cyl}$, we simply set $\dot{x}_d = \dot{x}_c$ (i.e., full accommodation of teleoperation command \dot{x}_c). We can also obtain \dot{P}_d from (14), for which δ can be computed by

$$\delta = \frac{-[\dot{x}_c + R(\alpha \xi_1 + \beta \xi_2)]_3}{[Rp]_3} \quad (15)$$

from (14) by using the fact that $[\dot{P}_d]_3 = 0$ with $\dot{x}_d = \dot{x}_c$, where $[\star]_3$ is the e_3 -component of \star . Here, note that $[Rp]_3 \neq 0$ (i.e., e_3 -component of the vector p relative to the inertial frame

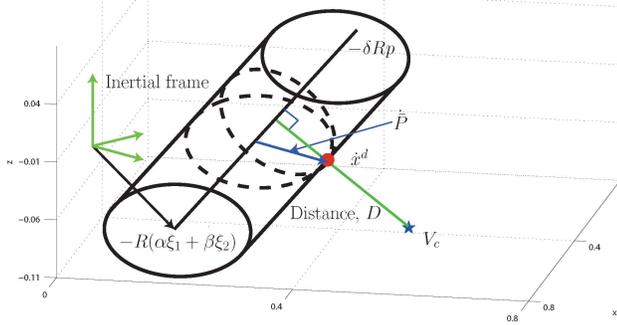


Fig. 4. Computation of \dot{x}_d on the coordination cylinder \mathcal{C}_{cyl} to be the closest point from the given teleoperation command \dot{x}_c .

$\{\mathcal{I}\}$), unless the UGV is located on the same height as the UAV, which we assume not to happen in this paper. With this δ , \dot{P}_d is then given from (14) by

$$\dot{P}_d = \dot{x}_c + R(\alpha\xi_1 + \beta\xi_2 + \delta p).$$

Note that, in this case, choosing \dot{x}_d and \dot{P}_d as above, we can fully realize the desired teleoperation behavior (i.e., \dot{x}_c), while also satisfying the UAV-UGV coordination (12) and the UGV's velocity limitation (13).

2) When $\dot{x}_c \notin \mathcal{C}_{\text{cyl}}$: In this case, if the UAV flies with the command \dot{x}_c , the UAV-UGV coordination would be compromised, which may also incur the image feature p of the UGV to be lost from the UAV's camera view. To avoid this, the algorithm (14) modulates the teleoperation command \dot{x}_c by choosing the UAV's desired velocity command \dot{x}_d to be the closest point on the coordination cylinder \mathcal{C}_{cyl} from \dot{x}_c . See Fig. 4. Once this \dot{x}_d is given, we can then solve \dot{P}_d similar as before, by using (15) with \dot{x}_c replaced by this \dot{x}_d . Note that, in this case, \dot{x}_d is on the surface of the coordination cylinder \mathcal{C}_{cyl} , implying that, to preserve the UAV-UGV coordination, the UGV should "catch up" the UAV with its maximum speed \bar{U} (13).

C. Human Haptic Interface

Although other interacting modalities are also possible, to allow a remote human user to intuitively teleoperate the UAV-UGV team, following [5], [20], we design the human haptic teleoperation interface as follows. First, the teleoperation velocity command \dot{x}_c for the UAV is command by the remote human user through their haptic device s.t. $\dot{x}_c(t) := \eta q(t)$ where $q(t) \in \mathbb{R}^3$ is the position of the haptic device and $\eta \in \mathbb{R}$ is some scaling to match the workspace size of the haptic device and the UAV's velocity. This velocity command $\dot{x}_c(t)$ then allows us to circumvent the issue of the master-slave kinematic dissimilarity (i.e., master device workspace is bounded, while that of the the UAV-UGV team unbounded [22]).

On the other hand, on top of the vision information on the motion of the UAV (either via cameras attached on the UAV (i.e., body-fixed perspective) or installed in its environment (i.e., global perspective)), we also provide haptic perception of the UAV's velocity to the human user. For this, we first define the haptic feedback signal $y(t)$ s.t. $y(t) := \frac{1}{\eta} \dot{x}_c(t)$ where

$\dot{x}(t)$ is the UAV's velocity and $1/\eta$ is the scaling compatible with $\dot{x}_c(t)$. This haptic signal $y(t)$ is then sent to the human user through some discrete-time communication network (e.g., Internet). Let us denote its reception by $y(k)$ at the reception time t_k . Then, we haptically present this information $y(k)$ to the human via the following haptic device control torque: for $t \in [t_k, t_{k+1}]$

$$\tau(t) = -B\dot{q} - K_1q - K_0(q - \bar{y}(k))$$

where $B, K_1, K_0 \in \mathbb{R}^{3 \times 3}$ are the symmetric and positive-definite gain matrices, and $\bar{y}(k)$ is the modulation of $y(k)$ through the passive set-position modulation (PSPM [4]). This PSPM can then enforce passivity (thus, robust interaction stability) of the master-side, even if the communication channel is imperfect (e.g., Internet), the device is engaged by a wide-range of human users, or other forms of the haptic signal $y(t)$ is used. See [5], [4], [20] for more details.

IV. SIMULATION

We consider a quadrotor-type UAV with a camera facing down to see the UGV for achieving the UAV-UGV coordination. The human operator can see and teleoperate the team of UAV and UGV from outside (i.e., global information). Also, for the UGV, we assume its evolution can be represented by a kinematic unicycle-type wheeled mobile robot [21], with the image feature p corresponding to the marker attached on its rotation center (i.e., axle center). For the UAV's low-level control (i.e., for $\dot{x} \rightarrow \dot{x}_d$), we utilize the backstepping control scheme of [20], while for the UGV, we use the following control to make $\dot{P} \rightarrow \dot{P}_d$:

$$\begin{bmatrix} \dot{u} + \rho u \\ \dot{\phi}/u \end{bmatrix} = \begin{bmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{bmatrix} [\ddot{P}_d + \rho \dot{P}_d] \quad (16)$$

where $u, \phi \in \mathbb{R}$ are the linear velocity input and the angle of the UGV, $\rho > 0$ is the control gain, and \ddot{P}_d is computed numerically. During our simulation, $u \neq 0$. This control is derived similar to [20], the details of which will be reported in a future publication. Of course, other low-level controllers may also be used instead of (16) (e.g., [21]).

We assume that the UAV is equipped with IMU (inertial measurement unit) and we can obtain R, w from this IMU (and also $\dot{p}_d = -w \times p_d$ for (9)). The relative depth $r(P)$ can also be computed by $r(P) = h/\cos\theta$, where θ is the angle between Rp and $e_3^T := [0 \ 0 \ 1]^T$ in the inertial frame and $h \in \mathbb{R}$ is the UAV's height. These θ and h can also be measured by using the camera and barometer. Using R and a number of markers on the UGV with the known geometry among them, we can also convert \dot{P}_d into the UGV's body-frame, which can then be tracked by the UGV using its local controller and local velocity sensing (e.g., encoder, LIDAR) both typically formulated in the UGV's body-frame.

The simulation results are shown in Figs. 5-7, where we can see that the human can teleoperate the UAV-UGV team while the UAV-UGV coordination is achieved by using the camera (Fig. 5). In this simulation, the human command is circular trajectory ; 2) the image feature p converges to the desired one p_d after some initial transient and also with $R(t) \rightarrow I$ (Fig. 6); and 3) both the projected and un-projected image tracking

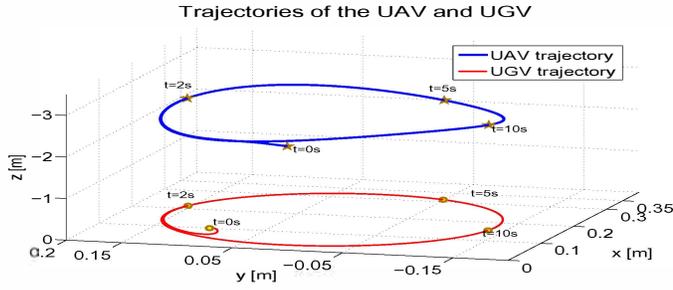


Fig. 5. Position trajectories of the UAV and UGV.

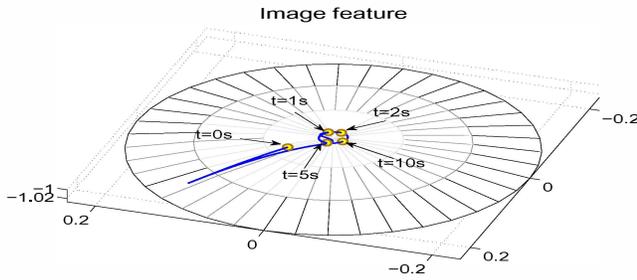


Fig. 6. Evolution of the image features on the image sphere.

errors (i.e., e_p, e) converge to zero exponentially while the UAV's motion stabilized (Fig. 7).

V. CONCLUSION

We present a novel vision-based teleoperation control framework for a UAV/UGV team, which allows a remote human user to teleoperate the UAV, while guaranteeing the UAV-UGV coordination using a camera attached on the UAV and seeing the UGV. A certain geometric condition is derived to ensure the UAV-UGV coordination and the UGV's velocity limitation, while the teleoperation command is modulated if it demands the violation of this condition, while also minimizing the deviation of the UAV's velocity from this teleoperation command. Simulation is performed to illustrate the theory. Some possible directions for future research include: extension of the proposed framework to the case of multiple UGVs and and UGVs on non-flat ground; and inclusion of the low-level UAV and UGV dynamics directly into the control design.

REFERENCES

- [1] T. Hamel and R. Mahony. Visual servoing of an under-actuated dynamic rigid-body system: an image-based approach. *IEEE Transactions on Robotics and Automation*, 18(2):187–198, 2002.
- [2] E. Altug, J. P. Ostrowski, and C. J. Taylor. Control of a quadrotor helicopter using dual camera visual feedback. *International Journal of Robotics Research*, 24(5):329–341, 2005.
- [3] N. Guenard, T. Hamel, and R. Mahony. A practical visual servo control for an unmanned aerial vehicle. *IEEE Transactions on Robotics*, 24(2):331–340, 2008.
- [4] D. J. Lee and K. Huang. Passive-set-position-modulation framework for interactive robotic systems. *IEEE Transactions on Robotics*, 26(2):354–369, 2010.
- [5] D. J. Lee, A. Franchi, P.R. Giordano, H.I. Son, and H.H. Bulthoff. Haptic teleoperation of multiple unmanned aerial vehicles over the internet. In *Proc. IEEE Int'l Conf. on Robotics and Automation*, pages 1341–1347, 2011.

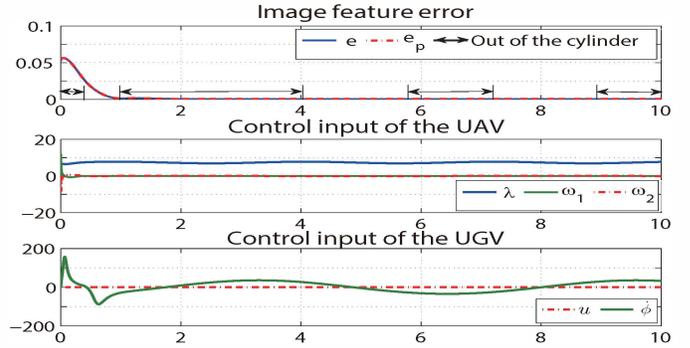


Fig. 7. Image feature tracking error, $\|e\|$ and $\|e_p\|$; the UAV control inputs λ, w ; and the UGV control inputs u, ϕ .

- [6] H. Romero, S. Salazar, and R. Lozano. Real-time stabilization of an eight-rotor uav using optical flow. *IEEE Transactions on Robotics*, 25(4):809–817, 2009.
- [7] O. Bourquardez, R. Mahony, N. Guenard, F. Chaumette, T. Hamel, and L. Eck. Image-based visual servo control of the translation kinematics of a quadrotor aerial vehicle. *IEEE Transactions on Robotics*, 25(3):743–749, 2009.
- [8] R. Ozawa and F. Chaumette. Dynamic visual servoing with image moments for a quadrotor using a virtual spring approach. In *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 5670–5676, 2011.
- [9] R. Rao, V. Kumar, and C. Taylor. Visual servoing of a ugv from a uav using differential flatness. In *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 743–748, 2003.
- [10] B. Bethke, M. Valenti, and J. How. Cooperative vision based estimation and tracking using multiple uavs. *Advances in Cooperative Control and Optimization*, pages 179–189, 2007.
- [11] R. He, A. Bachrach, M. Achtelik, A. Geramifard, D. Gurdan, S. Prentice, J. Stumpf, and N. Roy. On the design and use of a micro air vehicle to track and avoid adversaries. *The International Journal of Robotics Research*, 29(5):529–546, 2010.
- [12] B. Ludington, E. Johnson, and G. Vachtsevanos. Augmenting uav autonomy. *IEEE Robotics and Automation Magazine*, 13(3):63–71, 2006.
- [13] B. Hérissey, T. Hamel, R. Mahony, F. Russotto, et al. Landing a vtol unmanned aerial vehicle on a moving platform using optical flow. *IEEE Transactions on Robotics*, (99):1–13, 2012.
- [14] C. Teuliere, L. Eck, and E. Marchand. Chasing a moving target from a flying uav. In *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 4929–4934, 2011.
- [15] W. Li, T. Zhang, and K. Kuhnlenz. A vision-guided autonomous quadrotor in an air-ground multi-robot system. In *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 2980–2985, 2011.
- [16] K.E. Wenzel, A. Masselli, and A. Zell. Automatic take off, tracking and landing of a miniature uav on a moving carrier vehicle. *Journal of Intelligent and Robotic systems*, 61(1):221–238, 2011.
- [17] K. Celik, S-J. Chung, M. Clausman, and A. K. Somani. Monocular vision slam for indoor aerial vehicles. In *Proc. IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 1566–1573, 2009.
- [18] S. Stramigioli, R. Mahony, and P. Corke. A novel approach to haptic tele-operation of aerial robot vehicles. In *Proc. IEEE Int'l Conf. on Robotics and Automation*, pages 5302–5308, 2010.
- [19] M.D. Hua, T. Hamel, P. Morin, and C. Samson. A control approach for thrust-propelled underactuated vehicles and its application to vtol drones. *IEEE Transactions on Automatic Control*, 54(8):1837–1853, 2009.
- [20] D. J. Lee, C. Ha, and Z. Zuo. Backstepping control of quadrotor-type uavs: trajectory tracking and teleoperation over the internet. In *Proc. Int'l Conf. on Autonomous Systems*, pages 217–225, 2012.
- [21] R. Fierro and F. L. Lewis. Control of a nonholonomic mobile robot: backstepping kinematics into dynamics. *Journal of Robotic Systems*, 14(3):149–163, 1997.
- [22] D. J. Lee and D. Xu. Feedback r -passivity of lagrangian systems for mobile robot teleoperation. In *Proc. IEEE Int'l Conf. on Robotics and Automation*, pages 2118–2123, 2011.