Toward Coordinated Manipulator-Host Visual Servoing for Mobile Manipulating UAVs

Todd W. Danko[1] and Paul Y. Oh[2]

***Abstract*— Manipulating objects using arms mounted to unmanned aerial vehicles (UAVs) is attractive because UAVs may access many locations that are otherwise inaccessible to traditional mobile manipulation platforms such as ground vehicles.**

Historically, UAVs have been employed in ways that avoid interaction with the environment at all costs. The recent trend of increasing small UAV lift capacity and the reduction of the weight of manipulator components make the realization of mobile manipulating UAVs imminent. Despite recent work, several major challenges remain to be overcome before it will be practical to manipulate objects from UAVs. Among these challenges, the constantly moving UAV platform and compliance of manipulator arms make it difficult to position the UAV and end-effector relative to an object of interest precisely enough for manipulation. Solving this challenge will bring UAVs one step closer to being able to perform meaningful tasks such as infrastructure repair, disaster response, law enforcement, and personal assistance.

Toward a solution to this challenge, this paper describes an approach to coordinate the redundant degrees of freedom of an aerial manipulation system. The manipulator's six degrees of freedom will be visually servoed using an eye-in-hand camera to a specified pose relative to a target while treating motions of the host platform as perturbations. Simultaneously, the host platform's degrees of freedom will be servoed using kinematic information from the manipulator. This will drive the UAV to a position that allows the manipulator to assume a joint-space configuration that maximizes reachability, thus facilitating the arm's ability to compensate for undesired UAV motions without the need for an external pose estimation system.

I. Introduction

A. Mobile Manipulating Unmanned Aerial Vehicles

Unmanned Aerial Vehicles (UAVs) were originally deployed as target drones for combat pilot training but have evolved over time to provide valuable roles in intelligence, surveillance and reconnaissance for both civilian and military operations. Historically, UAVs were built and operated in ways to avoid interacting with their environment at all costs, affording them the ability to quickly and efficiently travel large distances. The ability for aerial vehicles to manipulate or carry objects that they encounter could greatly expand the types of missions achievable by unmanned aerial systems.

This work was supported in part by the National Science Foundation (NSF) award CNS-1205490. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF.

[1]Todd W. Danko is a Lead Research Scientist at Lockheed Martin's Advanced Technology Laboratories, Cherry Hill, NJ 08002, USA and is also with the Drexel Autonomous Systems Laboratory, Drexel University, Philadelphia, PA 19104, USA twd25@drexel.edu

[2]Paul Y. Oh is with the Drexel Autonomous Systems Laboratory, Drexel University, Philadelphia, PA 19104, USA pyo22@drexel.edu

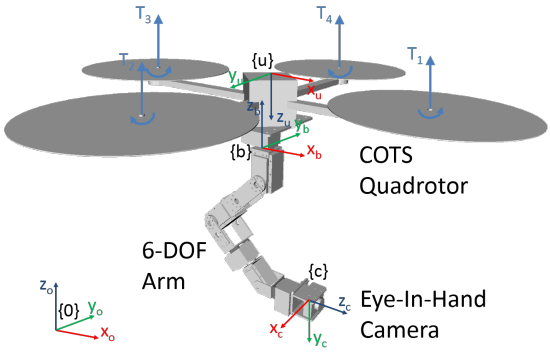


Fig. 1. Quadrotor coupled to a six degree of freedom manipulator with eye-in-hand camera

High degree of freedom robots with dexterous arms could lead to transformative applications such as infrastructure repair, law enforcement, disaster response, casualty extraction, and personal assistance, leading to a paradigm shift in the way UAVs are deployed. Such aerial manipulation systems have been coined Mobile Manipulating Unmanned Aerial Vehicles (MM-UAV).

B. Biological Inspiration

MM-UAV efforts are inspired by nature to implement truly dexterous manipulation from aerial vehicles in ways that are similar to how an octopus can use its tentacles to manipulate objects like seashells while hovering and "flying" over the ocean floor, vectoring jets of water to maintain dynamic stability.

The ability to steadily track an object before contact is made is seen in nature as shown by the ability for a chicken to maintain a steady head pose despite large perturbations to its body. Humans also seamlessly combine multiple degrees of freedom to track objects of interest. Small, quick motions are tracked by panning and tilting the eye, head motions often follow, allowing the eye to recenter to its straight ahead position while still tracking the object. A human can even become mobile to help track an object that is moving beyond the range of what is possible or comfortable using only eye or head motions.

C. Furthering the MM-UAV Paradigm

Problem Statement: Dynamically stabilizing a UAV while flying close enough to an object to manipulate it is a major challenge. Ground effect and turbulence act to destabilize the hovering UAV while a manipulator's dynamic motions and the shifting center of gravity make controlling

978-1-4799-4605-1/14/$31.00 ©2014 IEEE

the position of a MM-UAV non-trivial. One could affix a six degree of freedom arm to a quadrotor UAV, and control the arm in such a way that UAV motions are treated as perturbations, but how would one manage the redundant degrees of freedom of this MM-UAV system? Could one take advantage of these redundant degrees of freedom in such a way that the end-effector can visually servo to a desired pose relative to a target, compensating for undesirable platform motions? Could the motions of the arm, in turn be use as a kinematic sensor to provide feedback to the UAV's position controller, mitigating the need for external positioning components such as external motion capture systems?

Toward realizing this vision of MM-UAV, this paper describes initial steps in the direction of using visual and kinematic sensing to coordinate motions of the redundant degrees of freedom of an aerial manipulation system. Limitations in current UAV lift capacity and available manipulator components drive the use of a surrogate gantry to emulate the motions of a UAV. This gantry offers the additional benefits of experimental repeatability and the reduced potential for damaging equipment. The ultimate goal is to implement the algorithms described here on a flyable system such as the Ascending Technologies Pelican quadrotor.

II. Related Literature

A. Aerial Mobile Manipulation

With improvements in mobile manipulation techniques, particularly with ground robots, these methods are now being applied to aerial vehicles as well [1], [2] and [3]. The Yale Aerial Manipulator can grasp and transport objects using a compliant gripper attached to the bottom of a T-Rex 600 RC helicopter (Figure 2(b)) [4]. Researchers at the GRASP Lab at the University of Pennsylvania are using multiple quadrotors to transport payloads in three dimensions using cables or a gripper [5] and have also investigated dynamic maneuvers with single degree of freedom manipulators (Figure 2(a)) . Previous research at Drexel has produced a prototype UAV pickup mechanism with a hook to deliver and retrieve cargo [6] using image based visual servoing.

1) Aerial Multi-Link Manipulators: Recently, some groups that have mounted multiple-degree of freedom manipulators to UAVs. CATEC [7] described an approach to closely coordinate control of a quadrotor with a 3-link manipulator arm using a Variable Parameter Integral Backstepping (VPIB) controller. This paper shows that the VPIB controller outperforms standard PID control for platform stability, and the addition of a model based arm compensator further improves stability using VPIB.

2) Aerial Grasping: [8] describes the analysis of grasping, load stability, and hover control using a one degree of freedom gripper mounted to a traditional swash plate-equipped helicopter. This paper details recent advances in UAVs involving ground interaction and in particular, grasping of objects.

B. Partitioned Degrees of Freedom

Visual servoing approaches generally calculate the desired rotational and translational velocities for the camera in a coordinated manner. This is based at the assumption that the degrees of freedom of the manipulator are equally adept at translating and rotating the camera in space, which is not true for most manipulators, and is especially not true when considering the combined degrees of freedom of a UAV-manipulator system. Oh [9] devised a partitioning framework that uses frequency analysis to aid in the control law synthesis process, exploiting the kinematic and dynamic properties of each degree of freedom.

This paper describes an initial effort toward the use of partitioning to coordinate the redundant degrees of freedom between the UAV and manipulator arm for MM-UAV based on the assumption that degrees of freedom closer to the end-effector will have a higher frequency response than those closer to the robot's base. This is a reasonable assumption considering the serial configuration of an aerial manipulation system, where the closer a degree of freedom is to the end-effector, the less inertia it must overcome to achieve a goal position. The degrees of freedom of the host platform, on the other hand, must move not only the host platform, but all of the degrees of freedom of the manipulator, assuming a higher inertial burden.

III. MM-UAV Modeling

The model of the MM-UAV system (Figure 1) consists of two main components, the host platform, as described in III-A and the manipulator arm, fit with an eye-in-hand camera as described in III-B.

A. Host Platform

This work focuses on the execution of camera positioning tasks in near-hover conditions and makes use of a surrogate six degree of freedom gantry system to emulate such UAV motions. While this gantry is not a perfect substitute for an actual flying UAV, it is programmed to mimic the dynamics of a quadrotor UAV. The programmed x and y accelerations of the gantry are proportional to $roll$ and $pitch$ angles respectively. The gantry's $roll$ and $pitch$ axes are torque controlled providing an approximation of the behaviors that a UAV would experience when an attached manipulator arm moves or interacts with an object.

Quadrotors similar to the Ascending Technologies Pelican [10] are the target platform for this MM-UAV work because they are readily available and have a high lift capacity of up to 1 kg depending on configuration. This provides a target weight budget for the design of an arm as described in III-B.

B. Manipulator

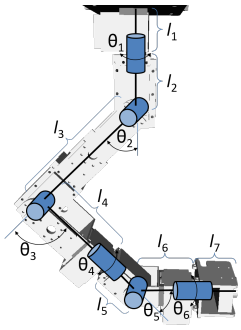
1) Arm Description: The arm is assembled from off-the-shelf Dynamixel servo motors and a mixture of off-the-shelf and custom brackets as shown in Figure 3. The Denavit-Hartenberg parameters that represent the six links of this arm are listed in Table I. This model is used in Matlab [11] to rapidly develop and test motion controllers without an

(a) GRASP quadrotor

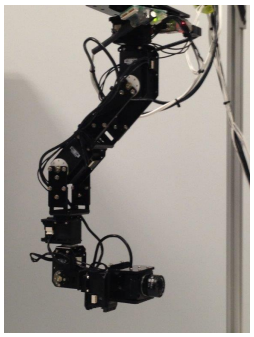
(b) Yale aerial manipulator

(c) CATEC Aerial robot with multi-link arm

Fig. 2. Aerial grasping implementations



(a)



(b)

Fig. 3. Manipulator constructed from off-the-shelf Dynamixel servos and brackets

TABLE I

DENAVIT-HARTENBERG PARAMETERS FOR THE MANIPULATOR

Link Number	θ (rad.)	d	a	α (rad.)
1	0	0	0	$-\pi/2$
2	0	0	L_2	0
3	0	0	0	$\pi/2$
4	0	L_3	0	$-\pi/2$
5	0	0	0	$\pi/2$
6	0	L_4	0	0

initial need for a motion capture system and accurate torque sensors at each joint to measure ground truth.

A lightweight camera is mounted to the arm's end-effector to provide an eye-in-hand visual servoing capability.

The link lengths used in the Denavit-Hartenberg model of the arm listed in Table II and are combinations of the lengths of the physical length of each joint-separated arm segment as listed in Table III.

Intermediate homogeneous transforms that represent each link are recreated from the Denavit-Hartenberg parameters using 2.

$$ {}^{b}T_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & L_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1} $$

TABLE II

MANIPULATOR LINK LENGTHS

Meta-Link	Length (m)	Composed of
L_1	0.096	$l_1 + l_2$
L_2	0.148	l_3
L_3	0.145	$l_4 + l_5$
L_4	0.127	$l_6 + l_7$

TABLE III

MANIPULATOR PHYSICAL LINK PROPERTIES

Physical Link	Length (m)	Mass (kg)
l_1	0.041	0.072
l_2	0.055	0.158
l_3	0.148	0.158
l_4	0.103	0.072
l_5	0.042	0.072
l_6	0.072	0.072
l_7	0.055	0.045

$$ {}^{n-1}A_n(\theta_n) = \begin{bmatrix} \cos\theta_n & -\sin\theta_n\cos\alpha_n & \sin\theta_n\cos\alpha_n & a_n\cos\theta_n \\ \sin\theta_n & \cos\theta_n\cos\alpha_n & -\cos\theta_n\sin\alpha_n & a_n\sin\theta_n \\ 0 & \sin\alpha_n & \cos\alpha_n & d_n \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2} $$

$$ {}^{6}T_E = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3} $$

2) Forward Kinematics: The position of the end-effector can be described relative to the arm's base as a series of transforms:

$$ T = {}^{b}T_E = \begin{array}{l} {}^{b}T_0\,{}^{0}A_1(\theta_1)\,{}^{1}A_2(\theta_2)\,{}^{2}A_3(\theta_3) \\ {}^{3}A_4(\theta_4)\,{}^{4}A_5(\theta_5)\,{}^{5}A_6(\theta_6)\,{}^{6}T_E \end{array} \tag{4} $$

3) Inverse Kinematics: Inverse kinematics calculations are used to identify positions for each joint of the manipulator (q) that when executed, result in the end-effector reaching a desired pose. A standard closed-form inverse kinematics solver is used for this anthropomorphic manipulator arm.

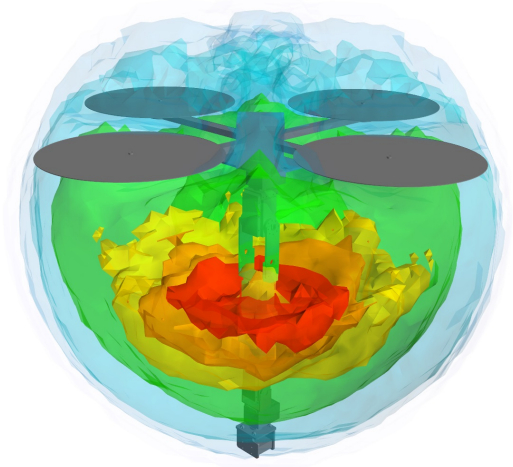


Fig. 4. Kinematic reachability of six degree of freedom manipulator attached to quadrotor UAV. End-effector poses that are most reachable are shown in hotter colors

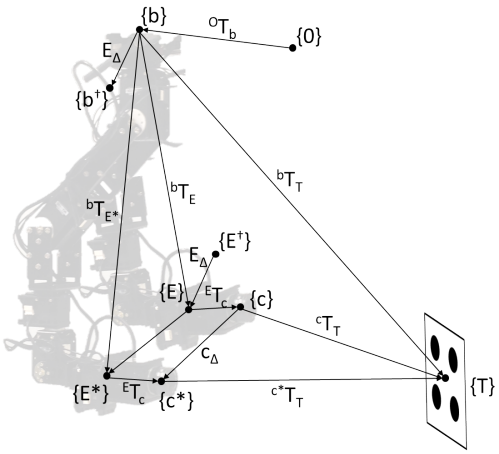


Fig. 5. Transforms between the world ($\{0\}$), arm base ($\{b\}$), end-effector ($\{E\}$), camera ($\{c\}$) and target ($\{T\}$)

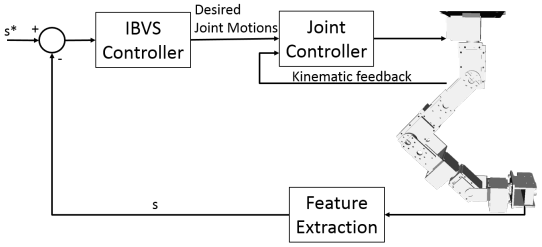


Fig. 6. Components inside of the red box are included in the dynamic look-and-move eye-in-hand framework

C. Manipulator Dynamics

The dynamics of the manipulator are described in Equation 5.

$$Q = M(q)\ddot{q} + C(q,\dot{q})\dot{q} + F(\dot{q}) + G(q) + J(q)^T f \quad (5)$$

Where Q is a vector of actuator forces at each joint of the manipulator given q,$\dot{q}$ and $\ddot{q}$, which are respectively the arm's joint-space angular positions, velocities and accelerations. M is the arm's inertial matrix, C, the Coriolis and centripetal coupling matrix, F, the friction force and G is the gravity loading for the link state. The final term includes J, which is the manipulator Jacobian that translates the wrench f at the arm's end-effector into generalized actuator forces.

D. Kinematic Reachability

Every manipulator has limitations in terms of the poses that the end-effector can reach. Analysis was performed to test how well this manipulator can reach finely spaced poses relative to the arm's base as shown in Figure 4.

This reachability analysis shows that the arm has a "sweet-spot" ($^{b}T_{E^{\dagger}}$) where the arm is more likely to be able to achieve a desired pose. The likelihood of achieving a given pose decreases as distance from $^{b}T_{E^{\dagger}}$ increases. This information is used in the coordination of manipulator and UAV motions. The manipulator is controlled to position the end-effector relative to a target, while the UAV's velocity is modulated to minimize the error $E_{\Delta} = {}^{b}T_{E}^{-1}\,{}^{b}T_{E^{\dagger}}$.

IV. Visual Servoing

The goal of visual servoing is to control the position of the robot relative to a target using visual features extracted from an image sensor.

A. Eye-In-Hand Visual Servoing Framework

Eye-in-hand visual servoing controllers as shown in Figure 5 make use of a camera mounted to a robot's end-effector to minimize the error (c_{Δ}) between a desired ($\{c^*\}$) and observed ($\{c\}$) camera pose relative to a target ($\{T\}$).

B. Image Feature Extraction

For this initial implementation, targets consisting of four dark circles with known radii, spaced in a square pattern with known edge length were used. The relevant features to extract from imagery using a Hough circle detector are the center points from each of the circles in image coordinates (u, v) which correspond to the row and column that the feature was found in the image. The coordinates of the detected circles are reported as s.

C. Image Based Visual Servoing

For IBVS, the task is to move features in image coordinates (s) to desired positions (s^*) through motions of the camera. This is shown in Figure 6.

The result of these visual servoing algorithms is the generation of a camera velocity vector (v) that will be executed by the manipulator. This velocity must therefor be converted into joint space motions using an inverse kinematics algorithm.

To accomplish this task, an image coordinate error that is a function of robot motion over time is defined as follows:

$$e(\vec{r}(t)) = C(s^* - s(\vec{r}(t))) \quad (6)$$

Where:

- s^* represents the desired positions of image features. The desired feature positions are calculated using a

camera model and the desired pose of the camera relative to the target
- $s(\vec{r}(t))$ represents the observed positions of image features. The observed feature positions are driven by the position of the camera relative to the target, as viewed by the camera
- C is the pseudo-inverse of the interaction matrix that converts image feature motions into desired camera velocities

For this work, a square target with a known edge length of l is assumed. To position the camera such that its optical axis is perpendicular to the target plane, centered a distance z^* away, a synthetic target model s^* is created by calculating the positions of four points in image coordinates ($s^* = (u_0 - a, u_0 + a, u_0 + a, u_0 - a, v_0 + a, v_0 + a, v_0 - a, v_0 - a)$) where a accounts for the projection of the target onto the focal plane assuming a pin-hole camera model. Knowledge of the focal length (f), detector pitch (ρ) and center detector position (u_0, v_0) is assumed.

$$a = \frac{fl}{2\rho z^*} \tag{7}$$

The interaction matrix, $L^T_{|s=s^*}$, associated with s^* is then calculated:

$$L^T_{|s=s^*} = \begin{bmatrix} l_1 & 0 & -a/z^* & -\rho a^2/f & -l_2 & a \\ l_1 & 0 & a/z^* & \rho a^2/f & -l_2 & a \\ l_1 & 0 & a/z^* & -\rho a^2/f & -l_2 & -a \\ l_1 & 0 & -a/z^* & \rho a^2/f & -l_2 & -a \\ 0 & l_1 & a/z^* & l_2 & \rho a^2/f & a \\ 0 & l_1 & a/z^* & l_2 & -\rho a^2/f & -a \\ 0 & l_1 & -a/z^* & l_2 & \rho a^2/f & -a \\ 0 & l_1 & -a/z^* & l_2 & -\rho a^2/f & a \end{bmatrix} \tag{8}$$

Where:

$$l_1 = \frac{-f}{\rho z^*}, l_2 = \frac{f^2 + \rho^2 a^2}{\rho f} \tag{9}$$

The pseudo inverse of C is such that $CL^T_{|s=s^*} = I_6$.
v is calculated:

$$v = C(s^* - s) = \begin{bmatrix} \delta x \\ \delta y \\ \delta z \\ \delta Rx \\ \delta Ry \\ \delta Rz \end{bmatrix} \tag{10}$$

An additional processing step is used to reformat the velocity vector v into a homogeneous transform c_Δ that represents the desired incremental motion of the camera.

$$c_\Delta = \begin{bmatrix} 1 & -\delta Rz & \delta Ry & \delta x \\ \delta Rz & 1 & -\delta Rx & \delta y \\ -\delta Ry & \delta Rx & 1 & \delta z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{11}$$

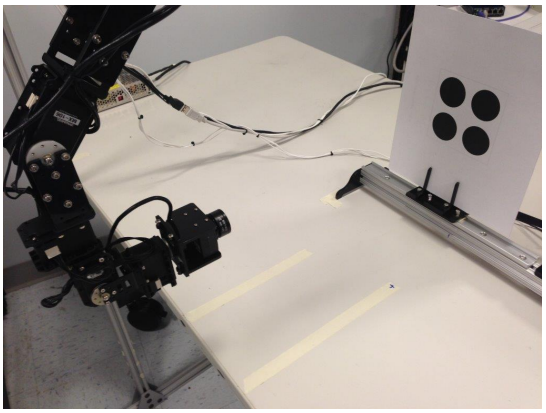


Fig. 7. Experimental setup for evaluating visual servoing algorithms

D. Conversion to Joint Space

The desired pose of the camera relative to the manipulator's base is simply an update to the camera's current pose modified by c_Δ.

$${}^bT_c(k+1) = {}^bT_c(k)\lambda c_\Delta \tag{12}$$

Where λ is a fraction between 0 and 1 that represents the size of step toward the calculated goal to move during each time step.

Knowing the desired motion of the camera, the desired pose of the end-effector (${}^bT_E(k+1)$) relative to the manipulator's base must be calculated to command the manipulator.

$${}^bT_E(k+1) = {}^bT_E(k)\,{}^ET_c\lambda c_\Delta\,{}^ET_c^{-1} \tag{13}$$

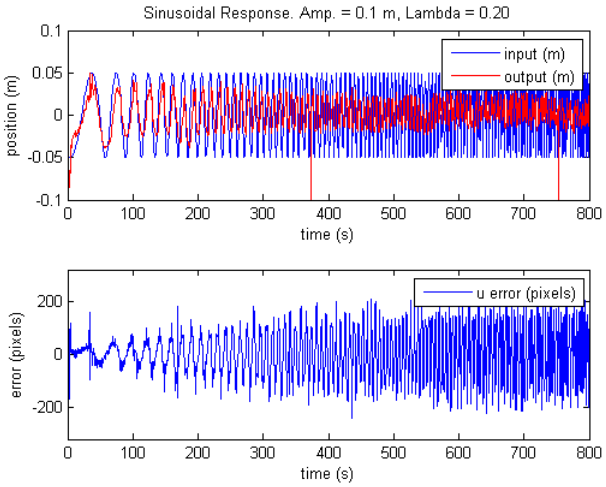
Where ${}^bT_E(k)$ is the current position of the end-effector and is found using equation 4. ${}^ET_c$ is the fixed transform between the end-effector and camera.

The inverse kinematics of the arm are used to solve for joint motions that will move the camera to the location specified by ${}^bT_E(k+1)$, thus executing the camera motion specified by λc_Δ.

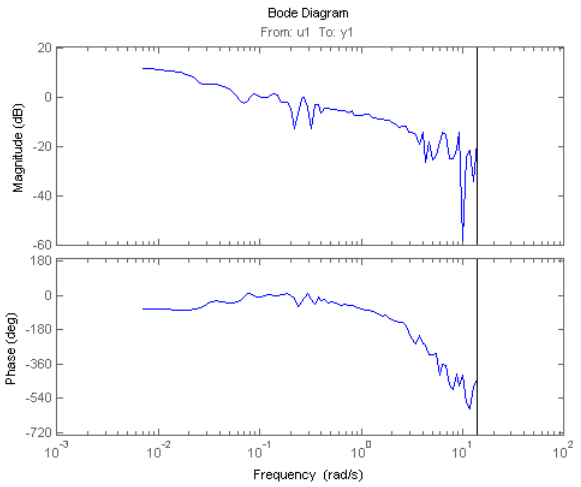
V. Implementation and Conclusions

For this initial evaluation, the visual servoing performance of the manipulator was evaluated using a target mounted to a linear motion platform. The linear motion platform translated the target a distance of $0.1\ m$ in an approximate sinusoidal motion at increasing frequencies. Truth information about the pose of the camera and position of the target was collected by reading encoder values from the servos that drove both the arm joint angles and the crank arm on the target motion platform.

The plots shown in Figure 8 show the visual servoing system's response to a sinusoidal pattern with a motion along the arm's y axis. The motion of the target is plotted in blue while the system response is plotted in red. Pixel error is also shown for each run. As expected, it can be seen that the system's output phase lag, which is attributed to latency in image capture and feature extraction and arm dynamics, increases along with attenuation with input frequency as



(a) Arm only chirp response data



(b) Bode plot

Fig. 8. IBVS tracking 0.1 m sinusoidal motion at various frequencies

shown in the accompanying Bode plot (Figure 8(b)). The system's bandwidth (frequency at $-3dB$) is approximately 0.5 rad/s.

This characterization of a visual servoing system is a single step toward the realization of a partitioning approach to coordinate the redundant degrees of freedom of an aerial manipulation system.

The visual servoing system frequency response data will be compared with the frequency response of various UAV platforms to aid in the selection of an appropriate aerial-manipulation host platform. Additionally, this information may be used as a benchmark for comparison when improvements are made the visual servoing processing chain.

These initial results will be expanded upon to ultimately create a flyable mobile manipulation system that uses visual servoing and kinematic feedback to position both the end-effector and UAV relative to a target of interest. The next steps include the synthesis of a controller that modulates the host vehicle's velocity based on kinematic feedback from the manipulator and its evaluation of a flying aerial manipulation system.

REFERENCES

[1] C. M. Korpela, T. W. Danko, and P. Y. Oh, "Mm-uav: Mobile manipulating unmanned aerial vehicle," *Journal of Intelligent & Robotic Systems*, vol. 65, no. 1-4, pp. 93–101, 2012.

[2] T. W. Danko and P. Y. Oh, "Design and control of a hyper-redundant manipulator for mobile manipulating unmanned aerial vehicles," *Journal of Intelligent & Robotic Systems*, pp. 1–15, 2013.

[3] M. Orsag, C. Korpela, and P. Oh, "Modeling and control of mm-uav: Mobile manipulating unmanned aerial vehicle," *Journal of Intelligent & Robotic Systems*, vol. 69, no. 1-4, pp. 227–240, 2013.

[4] P. E. Pounds and A. Dollar, "Hovering stability of helicopters with elastic constraints," in *ASME 2010 Dynamic Systems and Control Conference (DSCC2010)*, vol. 2. American Society of Mechanical Engineers, 2010, pp. 781–788.

[5] D. Mellinger, M. Shomin, N. Michael, and V. Kumar, "Cooperative grasping and transport using multiple quadrotors," in *Distributed autonomous robotic systems*. Springer, 2013, pp. 545–558.

[6] N. Kuntz and P. Y. Oh, "Towards autonomous cargo deployment and retrieval by an unmanned aerial vehicle using visual servoing," in *Proceedings of 2008 ASME Dynamic Systems and Controls Conference*, 2008.

[7] A. Jimenez-Cano, J. Martin, G. Heredia, A. Ollero, and R. Cano, "Control of an aerial robot with multi-link arm for assembly tasks," in *IEEE Int. Conf. Robotics and Automation (ICRA), Karlsruhe, Germany*, 2013.

[8] P. E. Pounds, D. R. Bersak, and A. M. Dollar, "Grasping from the air: Hovering capture and load stability," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2491–2498.

[9] P. Y. Oh and K. Allen, "Visual servoing by partitioning degrees of freedom," *Robotics and Automation, IEEE Transactions on*, vol. 17, no. 1, pp. 1–17, 2001.

[10] A. Technologies, "Pelican," *Online: www.asctec.de/uav-applications/research/products/asctec-pelican/*, 2013.

[11] P. I. Corke, "A robotics toolbox for matlab," *Robotics & Automation Magazine, IEEE*, vol. 3, no. 1, pp. 24–32, 1996.