# Visual Servoing of an ROV for Servicing of Tethered Ocean Moorings

Aaron M. Plotnik* and Stephen M. Rock*†
*Aerospace Robotics Laboratory, Stanford University
Durand Bldg, Rm 028
496 Lomita Mall
Stanford, CA 94305-4035
Email: {aplotnik,rock}@stanford.edu

†Monterey Bay Aquarium Research Institute
7700 Sandholt Road
Moss Landing, CA 95039-9644

*Abstract*— **This paper presents an approach for shared control of a Remotely Operated Vehicle (ROV) to assist the vehicle's pilots in servicing of moored underwater platforms. By providing precise automatic control of the ROV with respect to the mooring, a shared control system is established such that the ROV hovers automatically with respect to the mooring and the pilot is free to focus only on the manipulation tasks. The positioning system uses a single calibrated camera to measure bearings to several fiduciary markers on a test mooring of known position in the mooring's coordinates. An Unscented Kalman Filter (UKF) fuses these bearing measurements with vehicle sensors such as compass, inclinometers and rate gyros to estimate the relative position and orientation of the mooring with respect to the ROV. The approach leverages technology that has been presented previously for vision-based automatic tracking and observation of deep ocean animals. Results from simulation and field trials of this positioning system using the ROV *Ventana* in Monterey Bay are presented.**

## I. Introduction

Several deep ocean cabled observatory networks are nearing deployment. These networks will enable data sampling of the ocean at high temporal and vertical resolutions. Via cabling to shore, the networks will be remotely monitored and directed in real-time.

One component of these observatories is moored platforms which are to be deployed and serviced using Remotely Operated Vehicles (ROVs). One possible approach to servicing a mooring is to attach the ROV to it, followed by servicing or changeout of components using other manipulators of the ROV. However, attaching to such massive structures presents significant risks of damage to both the ROV and the mooring, as well as the potential for losing the ability to detach.

This paper discusses an alternate approach for servicing the moorings with an ROV by providing precise automatic control of the ROV with respect to the mooring requiring no fixed attachment. That is, the approach provides a shared control system in which the ROV hovers automatically with respect to the mooring, and the pilot is free to focus only on the manipulation tasks. Such automation requires a robust computer vision sensing system that can estimate the relative position and orientation of the mooring with respect to the ROV, along with an automatic positioning control system to achieve precise station-keeping of the vehicle with respect to a face of the mooring.

This paper presents the design of such a mooring-relative automatic positioning system that has been tested in the field using the ROV *Ventana* in Monterey Bay. The vision system incorporated in this study uses a single calibrated camera to measure bearings to several fiduciary markers on a test mooring of known position in the mooring's coordinates. An Unscented Kalman Filter (UKF) fuses these bearing measurements with vehicle sensors such as compass, inclinometers and rate gyros to estimate the relative position and orientation of the target with respect to the vehicle. Control loops are closed using the vehicle's thrusters to position the vehicle such that a constant range to the mooring is maintained while pointing directly at the marked portion of the surface. The vision system is designed to recognize when any of the markers is obscured by the ROV's manipulators, and redundancy in the number of markers allows one marker to be obscured without losing observability of the mooring's orientation.

The approach leverages technology that has been presented previously for vision-based automatic tracking and observation of deep ocean animals [1]–[8]. The animal observation system was designed to sense the relative position of the specimen of interest and servo the vehicle to maintain a constant range to the animal while pointing directly at it (thus absolute heading control is not a core requirement). Several extensions to that system are necessary to enable station-keeping with respect to the mooring. The primary differences are the needs (a) to estimate the orientation (and rotational velocity) of the tracked object, (b) to servo the vehicle such that it is always pointed at a specific face of the object as it rotates and (c) to handle intrusions of vehicle appendages (such as manipulators) in the camera's view.

The remainder of the paper is organized as follows: Section II describes the hardware and software architecture used for sensing and control. Section III describes the design of the

1-4244-0115-1/06/$20.00 ©2006 IEEE

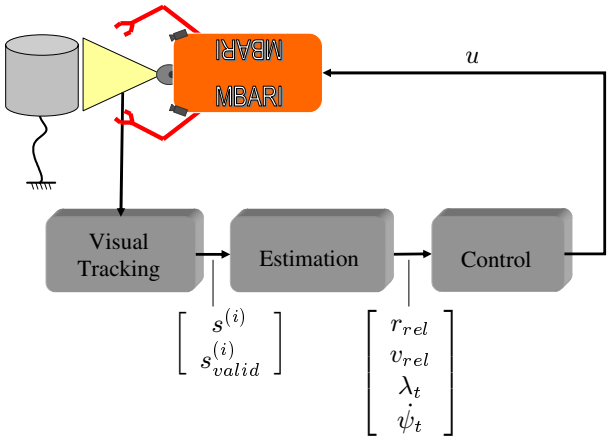


Fig. 1. Block diagram of top-level architecture. All software elements run at a 10 Hz rate.

vision system. Section IV covers the design of the estimator that computes the relative position, orientation, and velocity of the target with respect to the vehicle. Section V describes the architecture of the control system. Section VI presents results from simulated cases and documents field trials in Monterey Bay, California. Finally Section VII presents conclusions and the near-term future directions of development for the system.

## II. System Architecture

The high level architecture of the system for mooring-relative positioning of an ROV is shown in Figure 1. The key hardware components are the ROV and the thrusters that servo the vehicle, the camera used to view the mooring, and a gyrocompass. For field testing, an additional hardware component is the test target representing a simplified moored platform. The software side of the architecture is partitioned into three major components: visual tracking, estimation, and control.

### A. Hardware

The ROV *Ventana* is a hover-capable vehicle that can dive to 1500 m and carries a suite of instruments, cameras, a high definition camera, manipulators for instrument deployment and retrieval, and devices for sampling the ocean floor and animals in the midwater. The ROV is actively controlled in axial, lateral, vertical and yaw axes, but passively stabilized in pitch and roll. The ROV *Ventana* is pictured in Figure 2.

The camera used to view the mooring is the Pegasus model by Insite Pacific, Inc., a pressure-housed color camera with auto focus. This model has a horizontal field-of-view of 48 degrees, and a vertical field-of-view of 37 degrees, in water. The camera is mounted on a pan/tilt servo mechanism positioned at the front of the vehicle, near the top.

The vehicle is equipped with an Octans gyrocompass and Attitude and Heading Reference System (AHRS) (a product of Photonetics), which includes fiber optic gyros that measure angular rates and computes vehicle angles with respect to the earth. The angular rates and vehicle angles from this gyrocompass are used to assist the estimation process, which will be described in Section IV.

A test target was constructed to provide a portable "mooring" substitute that can be deployed temporarily and recaptured after experiments, and is shown in Figure 3. The target consists of a white metal plate with checkerboard fiduciary markers placed at known locations on the plate. The plate is weighted down from its two lower corners with lead weights on strings and a float is strung to its top center. Thus the plate is suspended vertically in the water column with only limited translations and rotations possible. A "handle" extends from the bottom center of the plate as a test point for manipulation by the ROV pilot during stationkeeping experiments.

With a single camera that can view at least three non-collinear points of known position in the frame of the mooring, the mooring's position and orientation with respect to the camera are observable. On the test target here, four fiduciary markers are present. Hence, there is redundancy in measuring its relative position and the system can operate at full capability with one marker obscured (e.g., by one of the robot's manipulators).

### B. Software

The software components include a set of algorithms that achieve visual tracking of the fiducial markers on the test mooring. These leverage the vision algorithms used in [4] to segment out the fiducial markers from the background of the test mooring and the water. The visual tracking algorithms also perform correspondence of the segmented patches from the current image to each marker, and maintain filtered predictions of the image patches for correspondence in the next cycle. Furthermore, the algorithms are capable of recognizing when a marker is obscured by a manipulator or otherwise not visible.

The estimator component estimates relative position and orientation of the mooring and the vehicle's velocities using the measured pixel locations from the visual tracking algorithms. The estimator implemented is an Unscented Kalman Filter (UKF) which handles nonlinear dynamics and sensor models naturally. (While a UKF might not be required for a

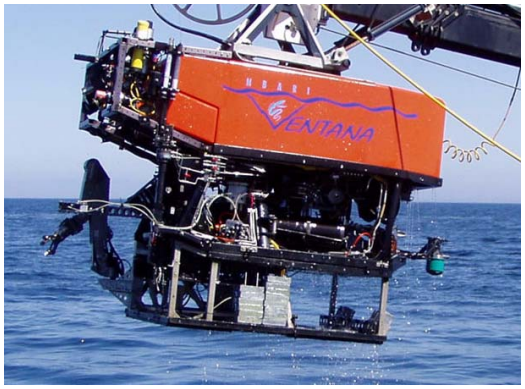


Fig. 2. The ROV *Ventana*.

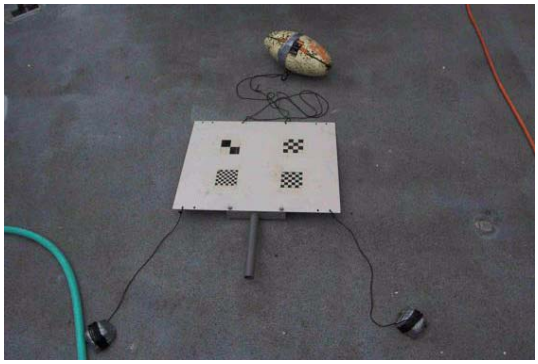

Fig. 3. Photo of test target used for relative positioning tests.

perturbational stationkeeping system, it was chosen here to enable the extension to systems executing larger and/or faster motions.)

The control portion of the software accepts estimates of the states of the vehicle and mooring to compute commanded positions and velocities. The control objectives are separated into four independent loops: a pointing loop that seeks to point the control camera at the center of the mooring, a ranging loop that seeks to maintain a constant range to the center of the mooring, a loop that regulates relative altitude between the vehicle and target, and an absolute heading control loop that seeks to match the vehicle's heading to the estimated heading of the mooring's face. Position commands are then computed by projecting errors into each of the vehicle's controllable axes, and are implemented by a set of independent PID control loops corresponding to each axis.

## III. Vision System Design

The vision sensing system receives camera images as input, and outputs the pixel locations of the centroids, in image coordinates, of each fiducial marker of the target mooring. Also output is a validity flag that identifies whether or not the vision system has successfully matched a pixel group to that marker.

The architecture is partitioned into three components: segmentation, correspondence, and prediction. Segmentation partitions the current image from the camera into background pixels and groups of pixels that may be the target's fiducial markers. Correspondence compares each pixel group to the expected characteristics for each marker and matches markers to pixel groups (blobs). The prediction step utilizes the corresponded marker/blob pairs to update the expected characteristic vector.

### A. Segmentation

The segmentation subsystem of vision processing is the first step in processing the input image. The segmentation process for the current implementation of the mooring-tracking application is identical to the process used in the tracking of deep ocean animals [1], [5]. The algorithms consist of a series of fixed parameter filters that remove noise and marine snow, followed by a morphological gradient filter. The output of this gradient filter is thresholded to extract image regions that contain high local gradients. This has been found to segment out successfully a wide range of ocean animals from the background at a low computational cost and with fixed thresholds. The test target used in the mooring tracking experiments is specially designed to trigger these same filters by placing checkerboard patterns as fiducial markers on a flat background.

### B. Correspondence

Correspondence is carried out by scoring each candidate pixel region, or "blob", with the predicted profile associated with each marker. The feature vector used for this comparison includes blob characteristics such as blob centroid in image coordinates, pixel area, aspect ratio, average intensity and average gradient. Multiple markers are matched to incoming measurements from the current image. First, each blob is compared to each profile using a weighted difference between the blob's features and the marker's predicted features, resulting in a score for each possible pairing. Then, the best scores for each marker are selected and matches are annunciated. If, for a given marker, no blob's score is sufficiently good (as occurs when a marker is obscured for instance), then no match is made and that marker's measurement is reported as unavailable to the estimator and the prediction step of vision processing.

### C. Prediction

The prediction step of vision processing uses the current matched blob measurements to update the profiles associated with the fiducial marker set. The update uses a weighted average of the previous profile vector and the current measurement, effecting a low pass filter on the measurements. Future implementations of this system will most likely leverage the information available from the estimator about motions of the vehicle and target to predict more accurately the expected changes in the metrics of the blob vector, similar to the image coordinate prediction step of [9].

## IV. Estimator Design

The states estimated are the relative position between the vehicle and target, the velocity of the vehicle relative to the target, the orientation of the target and its heading rate. This estimator is designed to handle nonlinearities in both the equations of motion and the observation models. Furthermore, its structure is flexible to handle losses of one or more observations dynamically as the vision system flags individual fiducial markers as unmeasured from time to time. The estimator is designed to have bandwidth properties such that noise is filtered adequately while still providing a fast enough response to allow tight control of the vehicle's relative position to the mooring.

### A. Nomenclature and Equations of Motion

Figure 4 illustrates the coordinate systems and vectors that define this problem, drawn from an overhead view. An

assumption is made that the target's center is stationary and therefore the water frame may be treated as inertial.

The estimator tracks the following state vector using a process model made up of the equations of motion in Equation 2:

$$\mathbf{x} = \begin{bmatrix} \mathbf{r_v^T} & \dot{\mathbf{q}}_\mathbf{v}^\mathbf{T} & \boldsymbol{\lambda}_\mathbf{t}^\mathbf{T} & \dot{\psi}_t \end{bmatrix}^T \tag{1}$$

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u^v}, \boldsymbol{\lambda}, \boldsymbol{\omega}, \mathbf{n^f}, \mathbf{n}^{\boldsymbol{\lambda}_\mathbf{t}}, {}^z n^{\omega_t}) = \begin{cases} \frac{d}{dt}\mathbf{r_v} &= -\dot{\mathbf{q}}_\mathbf{v} - \boldsymbol{\omega} \times \mathbf{r_v} \\ \frac{d}{dt}\dot{\mathbf{q}}_\mathbf{v} &= g\left(\mathbf{u^v} + \mathbf{n^f}, \dot{\mathbf{q}}_\mathbf{v}\right) - \boldsymbol{\omega} \times \dot{\mathbf{q}}_\mathbf{v} \\ \frac{d}{dt}\boldsymbol{\lambda}_t &= \begin{bmatrix} \mathbf{n}^{\boldsymbol{\lambda}_\mathbf{t}^T} & \dot{\psi}_t \end{bmatrix}^T \\ \frac{d}{dt}\dot{\psi}_t &= {}^z n^{\omega_t} \end{cases} \tag{2}$$

In Equation 2, coordinate systems are indicated with subscripts, $t$ and $v$, indicating the target frame and vehicle frame body coordinates respectively. The function $g(\cdot)$ represents the translational dynamics of the vehicle in the water, which can contain terms accounting for the relationships between the thrusters, the vehicle's dry and added masses and damping forces.

$\boldsymbol{\omega}$ is the angular rate of the vehicle, expressed in body coordinates. Note that both $\boldsymbol{\lambda}$ and $\boldsymbol{\omega}$ are not treated as states of the estimator, but rather are used directly as known quantities in the process model. These quantities are sensed by a gyrocompass which includes fiber optic gyros that measure angular rates and also computes vehicle angles with respect to the earth. The choice to remove them from the state vector was made to reduce the order of the estimator, and is possible due to the reliability and quality of the sensors used on the vehicle to measure these quantities, as well as their availability at a high sample rate.

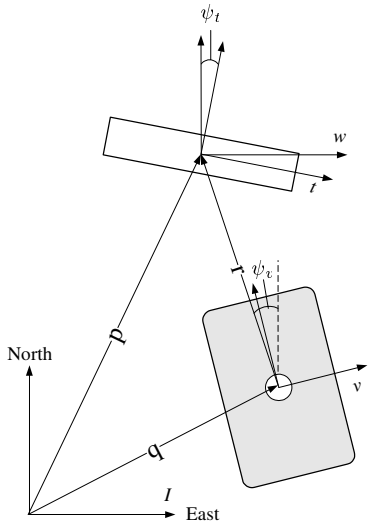


Fig. 4. Definition of vectors and coordinate frames.

The function $g(\cdot)$ in Equation 2, which represents the translational dynamics of the vehicle in the water is defined here in Equation 3 to be a purely kinematic model driven by white noise acceleration:

$$g\left(\mathbf{u^v} + \mathbf{n^f}, \dot{\mathbf{q}}_\mathbf{v}\right) = \mathbf{n^f} \tag{3}$$

The target is modeled as stationary in its translational degrees of freedom, but able to rotate about any axis. However, only the heading angle is assumed to experience any sustained rates of change. Thus, the equations of motion of the target in Equation 2 show the heading rate, $\dot{\psi}_t$, as an explicit state, whose derivative is driven by the white noise angular acceleration term, ${}^z n^{\omega_t}$. The dynamics in the target's pitch and roll angles (the first two terms of $\boldsymbol{\lambda}_\mathbf{t}$) are driven by white noise velocity inputs, $\mathbf{n}^{\boldsymbol{\lambda}_\mathbf{t}}$.

### *B. Observation Model*

The observations are expressed as a camera pixel measurement, $\mathbf{s^{(i)}}$, of the $i^{th}$ marker at the position in camera frame of $\mathbf{l_c^{(i)}}$:

$$\begin{aligned} \mathbf{s^{(i)}} &= h_{cam}\left(\mathbf{l_c^{(i)}}\right) + \mathbf{n_{cam}^{(i)}} \\ i &= \{1, 2, \ldots, N_{mrk}\} \end{aligned} \tag{4}$$

with,

$$h_{cam}(\cdot) = d^c\left\{n^*\ p(\cdot)\right\} \tag{5}$$

where, $p(\cdot)$ is the projection function for a pinhole camera model, and $d^c\{\cdot\}$ is a lens distortion model as calibrated for camera $c$ [10], [11]. The camera was calibrated in air and the underwater refraction effect is handled by scaling the bearings by $n^*$, the sea water/air refraction index of 1.339 [12].

Each position vector $\mathbf{l_c^{(i)}}$ is computed as follows:

$$\begin{aligned} \mathbf{l_c^{(i)}} &= {}^c\mathbf{T}^t \cdot \mathbf{l_t^{(i)}} \\ &= {}^c\mathbf{T}^v \cdot {}^v\mathbf{T}^t \cdot \mathbf{l_t^{(i)}} \\ &= {}^c\mathbf{T}^v \cdot {}^v\mathbf{T}^w(\boldsymbol{\lambda}, \mathbf{r_v}) \cdot {}^w\mathbf{T}^t(\boldsymbol{\lambda}_\mathbf{t}) \cdot \mathbf{l_t^{(i)}} \end{aligned} \tag{6}$$

where ${}^a\mathbf{T}^b$ represents the homogeneous transformation matrix converting vectors expressed in frame $b$ to frame $a$ [13]. The positions of the markers in the target frame, $\mathbf{l_t^{(i)}}$, are assumed known, as is the transform from vehicle frame to camera frame, ${}^c\mathbf{T}^v$, derived from measured pan and tilt angles of the camera mount and its known location with respect to the vehicle. The other transform matrices are computed from the sensed vehicle angles, $\boldsymbol{\lambda}$, (which are used directly) and the state estimates of $\mathbf{r_v}$ and $\boldsymbol{\lambda}_\mathbf{t}$.

### *C. Estimator Implementation*

A Sigma-Point Kalman Filter (SPKF), using the Unscented Kalman Filter (UKF) implementation was chosen here [14], [15]. The SPKF is an estimator that provides the ability to naturally handle nonlinearity in both the process and sensor models without linearizing those models and without incurring significant computational costs compared to an Extended Kalman Filter (EKF), which requires model linearization about the mean state estimate.

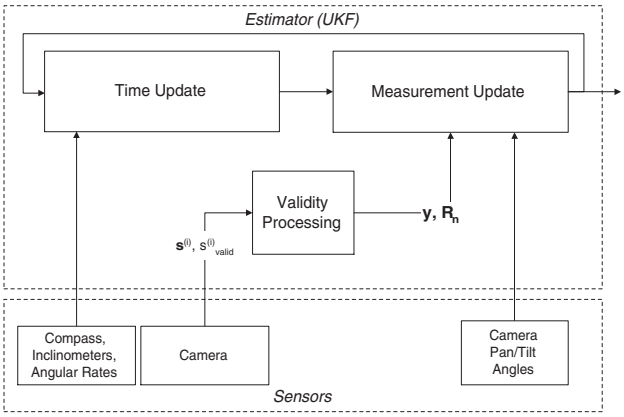


Fig. 5. Block diagram of estimator architecture.

The architecture relating the process and sensor models and their interactions with the system's sensors is illustrated in Figure 5. This diagram illustrates a number of features of this estimator. The handling of sensor validities is represented by the block labeled "Validity Processing". This process dynamically assembles available sensor measurements into the vector, $\mathbf{y}$, and constructs a measurement covariance matrix, $\mathbf{R_n}$ based on the subset of measurements available. Via this block, the measurement update dynamically adjusts its structure to prevent sensor readings that have been flagged as "not valid" from being applied to the estimation process. This happens often with the vision system, whose correspondence component is designed to reject false tracking matches and also occurs if a portion of the target leaves the view of the camera. The flexibility of the Kalman filter architecture that is carried through to the SPKF allows this adjustment to happen naturally without modification to the core estimation algorithms.

Because the UKF estimator used here is nonlinear, bandwidth was approximated using the rise time of responses to steps in sensor readings as a proxy for bandwidth [16]. The final bandwidths, in Hz, of the estimator are 0.7 Hz in vehicle $x$ axis (which dominates range estimates), 1.15 Hz in vehicle $y$ axis, 1.43 Hz in vehicle $z$ axis (vertical) and 0.7 Hz in estimating the target's heading.

## V. Control Design

The automatic positioning control system has a goal of maintaining a constant range to the center of the target, while facing directly at the face in heading.The control architecture is illustrated in Figure 6. This architecture is nearly identical to that used in automatic tracking of animals [7]. The main modification made is that instead of the heading reference being set by the operator, the reference heading is now commanded to match the estimated heading of the target, $\hat{\psi}_t$.

## VI. Results

### A. Simulated Performance

Step responses of the coupled estimator/control system were tested in simulation for the mooring-tracking application. The results for each loop are shown in Figure 7.

### B. Field Experiments

The system has been tested in Monterey Bay, California using the test target shown in Figure 3. The task was to demonstrate shared control with the ROV pilot, such that the automatic positioning system holds the ROV steady with respect to a face of the target while the pilot operates a manipulator to place a ring around the tube that descends from the bottom of the target. This task was achieved repeatedly in field testing, demonstrating that the relative positioning control system achieves stationkeeping performance that is sufficiently steady to enable finely controlled intervention tasks using the ROV's manipulators. A photograph taken from the control room during these experiments is shown in Figure 8.

## VII. Conclusions and Future Work

The concepts of outfitting a moored target with fiducial markers and sensing its position and orientation using a single camera, and thereby enabling precise relative positioning control have been proven out by the system and experiments presented in this paper. Future development of this system will address visual issues that are expected when dealing with a real moored instrument that is deployed for long periods of time and can be expected to incur bio-growth. Also, the disturbance levels due to currents and tether forces on the ROV in the environment of future moored instruments must be analyzed to ensure that the automatic positioning system is tuned to be able to reject them sufficiently.

## Acknowledgment

The authors would like to thank the Monterey Bay Aquarium Research Institute for their support in performing this research, including seatime on both the *Ventana* and *Tiburon* ROVs.

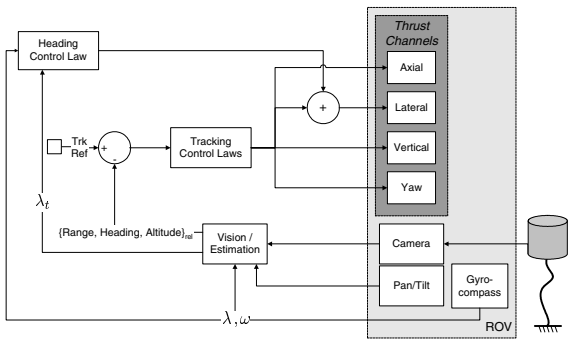


Fig. 6. Control architecture.

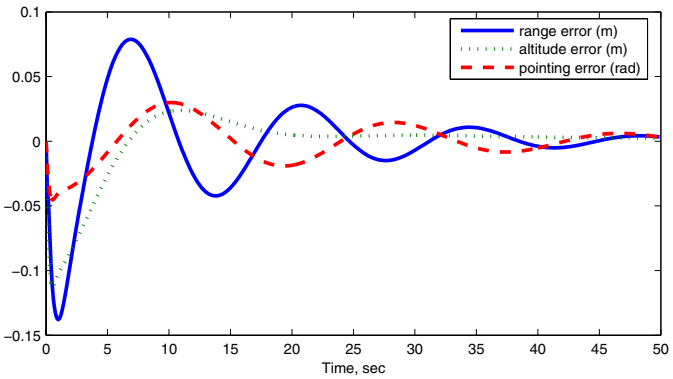


Fig. 7. Control loop responses, simulated. Step response errors in ranging, relative altitude, and pointing loops.

Fig. 8. A photograph of the control room during experiments in shared control of the ROV to manipulate a moored target.

## REFERENCES


[1] J. Rife and S. M. Rock, "Visual tracking of jellyfish in situ," in *Proc. of the 2001 International Conference on Image Processing*. Thessaloniki, Greece: IEEE, October 2001.

[2] ——, "A pilot-aid for ROV based tracking of gelatinous animals in the midwater," in *Proceedings of the OCEANS 2001 Conference*, Honolulu, HI, November 2001.

[3] ——, "Field experiments in the control of a jellyfish tracking ROV," in *Proceedings of the IEEE OCEANS Conference*, 2002, pp. 2031–2038.

[4] J. Rife, "Automated robotic tracking of gelatinous animals in the deep ocean," Ph.D. dissertation, Stanford University, Stanford, California, December 2003.

[5] J. Rife and S. M. Rock, "Segmentation methods for visual tracking of deep-ocean jellyfish using a conventional camera," *IEEE Journal of Oceanic Engineering*, vol. 28, no. 4, pp. 595–608, October 2003.

[6] ——, "Design and validation of a robotic control law for observation of deep-ocean jellyfish," *IEEE Transactions on Robotics*, vol. 22, no. 2, pp. 282–291, April 2006.

[7] A. M. Plotnik and S. M. Rock, "Relative position sensing and automatic control for observation in the midwater by an underwater vehicle," in *Proceedings of the Unmanned Untethered Submersible Technology Conference (UUST)*. Durham, NH: AUSI, Aug 2005.

[8] ——, "A multi-sensor approach to automatic tracking of midwater targets by an ROV," in *Proceedings of the AIAA Guidance, Navigation and Control Conference*, San Francisco, CA, Aug 2005.

[9] J. Langelaan and S. Rock, "Towards autonomous UAV flight in forests," in *Proceedings of the AIAA Guidance, Navigation and Control Conference*, San Francisco, CA, Aug 2005.

[10] J. Heikkilä and O. Silvén, "A four-step camera calibration procedure with implicit image correction," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, 1997, pp. 1106–1112.

[11] J.-Y. Bouget, "Camera calibration toolbox for matlab," http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

[12] J.-M. Lavest, G. Rives, and J.-T. Lapreste, "Underwater camera calibration," in *European Conference on Computer Vision*, vol. 2, 2000, pp. 654–668.

[13] J. J. Craig, *Introduction to Robotics: Mechanics and Control*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989.

[14] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proceedings of the American Control Conference*, 1995, pp. 1628–1632.

[15] R. van der Merwe, E. Wan, and S. Julier, "Sigma-point Kalman filters for nonlinear estimation and sensor-fusion: Applications to integrated navigation," in *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Providence, RI, Aug 2004.

[16] G. F. Franklin, M. L. Workman, and D. Powell, *Digital Control of Dynamic Systems*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1997.