Proceedings of the 2002 IEEE/RSJ
Intl. Conference on Intelligent Robots and Systems
EPFL, Lausanne, Switzerland • October 2002

Designing Visually Servoed Tracking to Augment Camera Teleoperators

Rares Stanciu, Paul Y. Oh *

Abstract

Keywords: visual-servoing, tracking, biomimetic, redundancy, degrees-of-freedom Abstract - There are many features to take into consideration when designing servoed vision for robot control, especially if there are redundant degrees-of-freedom (DOF). When human teleoperated platforms like rovers, booms, gantries, aircraft and submersibles are involved, visual-servoing is particularly challenging. Along with the kinematics and dynamics characteristics in the robot DOF one must consider the human-in-the-loop. Another factor that must be considered is potential motion conflicts arising from the shared man-machine control of the camera and can lead to unstable performance. When the DOF are redundant, the tracking problem becomes more complicated because there are many ways to perform tracking. This paper illustrates these issues and suggests control approaches to resolve them.

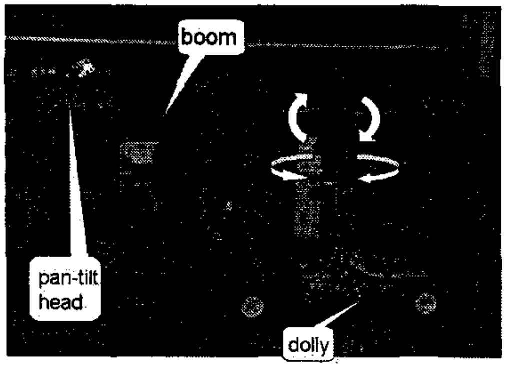
1 Introduction

Vision-based tasks like broadcasting and acquiring situational awareness often demand using a motion platform-mounted camera. A human operator pilots motion platforms like rovers, booms, gantries, aircraft or submersibles to position and orient the camera to capture desired fields-of-view. Typically such platform-camera systems possess redundant degrees-of-freedom (DOF) to overcome joint limits, avoid collisions with the environment and ensure occlusion-free views. Tracking a moving subject is a particularly challenging task because one or more highly skilled operators coordinate the platform's many DOF to keep the subject's image centered in the camera's field-of-view and invoke redundant joints when necessary. Tracking performance is thus limited by how quickly the operator manipulates and coordinates multi-DOF. Visually servoed tracking of eye-in-hand systems, where a camera is mounted a robot's end-effecter, is possible. For example, pose regulating simple objects like blocks [1], [2], [6], [13], [16] or keeping a target's image centered in the camera's field-of-view with a pan-tilt unit or camera head have been long demonstrated [5], [7], [16]. However, lacking are analytical methods to both visually servo robot-camera systems characterized by redundant DOF and handle visual-servoing for human-in-the-loop systems. Our interests in visual-servoing are in applying machine vision to augment the tracking performance of human camera operators. For example, a boom-mounted camera shown in Figure 1 is often used in sports broadcasting. One or more operators constantly push and steer the dolly, pan and tilt the boom and pan, tilt, and zoom the camera to keep the moving athlete's image centered in the camera. Automated image-centering can be achieved by visually servoing the pan-tilt camera and would thus reduce the number of DOF the operator(s) must manipulate. The net effect is a man-machine platform-camera system; the human operator controls some DOF while the computer controls others. This paper describes the design issues of such a system and some preliminary results with visual servoing, human-in-the-loop control and joint redundancies. Section 1 describes the boom-camera system in more detail and visual-servoing implementation is given in Section II. Section III gives some experimental results highlighting the stability issues of coordinating redundant joints and Section IV concludes with a map of our future work.

2 Section II: System Description

The boom-camera system is composed of a 4-wheeled dolly, boom, motorized pan-tilt head (PTH) and camera as shown in Figure 1. The 1.22 m long by 0.76 m wide dolly has four wheels and thus can be pushed and steered. The

*Drexel University Mechanical Engineering, & PRISM Lab, Email: ris22.drexel.edu, paul@coe.drexel.edu



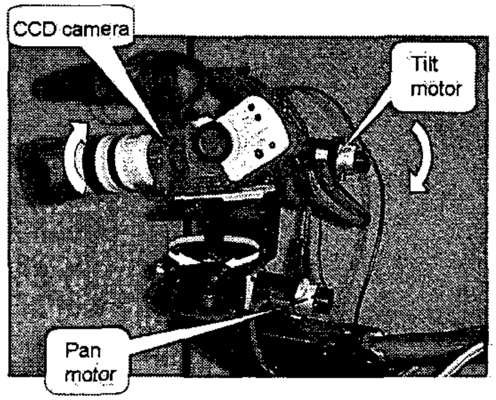


Figure 1: The left photo shows the boom, dolly and pan-tilt head camera system typically used in the broadcasting. Boom motions swing the pan-tilt head and camera horizontally and vertically. The right photo is a close up of the motorized pan-tilt head.

1.2 m long boom is linked to the dolly via a 1.04 m cylindrical pivot, which allows the boom to sweep motions horizontally (pan) and vertically (tilt). Mounted on one end of the boom is a 2-DOF motorized pan-tilt head and video camera weighing 21-lbs. The motors allow an operator to both pan and tilt the camera 360 deg at approximately 90 deg/sec. The pan-tilt head and camera are counterbalanced by 29.5 kg of dumbbell plates mounted on the boom's opposite end. Normally broadcast use of this boom-camera system entails one or more skilled personnel: (1) With a joystick, the operator servos the pan-tilt head's DC motors to point the camera. A PC or small board computer interfaced motion control card, ISA or PC-104 bus respectively, allow for accurate and relatively fast camera rotations. (2) The operator physically pushes on the counterweighted end to boom the camera horizontally and vertically. This allows one to deliver a diverse range of camera views (e.g. shots looking down at the subject), *overcome pan-tilt head joint limits and capture* occlusion-free views. (3) The operator can push and steer the dolly in case the boom and pan-tilt head are not enough to keep the target's image in the camera's field-of-view. These three actions are portrayed in Figure 2. Our augmentation interests are to use machine vision to visually servo the pan-tilt camera and integrate computer control in the human-in-the-loop system. For the former, the target's image centroid can be measured from the real-time frame data to visually servo the 2-DOF pan-tilt head and camera. This can automatically keep the image centered in the cameras field-of-view and allow the operator to just focus on boom swings and dolly translations.

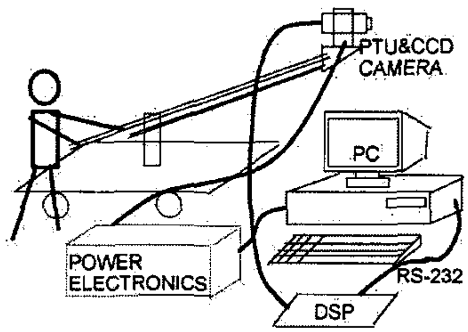


Figure 2: The *human operator pushes and steers* the dolly, swings the boom and joystick-controls the pan-tilt unit (PTU) to achieve a camera position and orientation.

For the latter, ultimately the pan-tilt head and boom motions redundantly orient the camera and can be problematic. For example during the visually servoing of the pan-tilt camera the operator conceivably can boom in the opposite direction. To compensate, the visually servoing must rotate the camera faster and if the two motions are out of phase by 180 degrees, they can conflict and visual-servoing will be unstable. The control aspects of this latter problem are particularly interesting to us; the pan-tilt head is a fast bandwidth actuator but has limited range-of-motion whereas the boom can swing the camera over large arcs but its inertia limits swinging speed. Fine/course

$$R = \begin{bmatrix} c\theta_1 c2\theta_2 c\theta_4 - s\theta_1 s\theta_4 & -c\theta_1 s2\theta_2 & -c\theta_1 c2\theta_2 s\theta_4 - s\theta_1 c\theta_4 & l_1 c\theta_1 c\theta_2 + l_2 c\theta_1 c2\theta_2 - l_3(c\theta_1 c2\theta_2 s\theta_4 - s\theta_1 c\theta_4) \\ \theta_1 c2\theta_2 c\theta_4 - c\theta_1 s\theta_4 & -s\theta_1 s2\theta_2 & -s\theta_1 c2\theta_2 s\theta_4 + c\theta_1 c\theta_4 & l_1 s\theta_1 c\theta_2 + l_2 s\theta_1 c2\theta_2 - l_3(s\theta_1 c2\theta_2 s\theta_4 - c\theta_1 c\theta_4) \\ s2\theta_2 c\theta_4 & -2\theta_2 & -s2\theta_2 s\theta_4 & d - l_1 s\theta_1 c\theta_2 - l_2 s2\theta_2 - l_3 s2\theta_3 s_4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(1) arm matrix; ($c\theta_i = \cos\theta_i$ and $s\theta_i = \sin\theta_i$)

actuation, if properly tuned can leverage the best each has to offer. Such fine/course schema characterize many motion platform-mounted camera systems like pan-tilt cameras mounted on helicopters and rovers. The vehicle provides large range of motion but fine pan-tilt motions are required to ensure the image remains centered in the camera. To underscore the man-machine control issues of visually servoing redundant DOF systems, we use a simple but reliable vision system. Real-time image centroid measurements

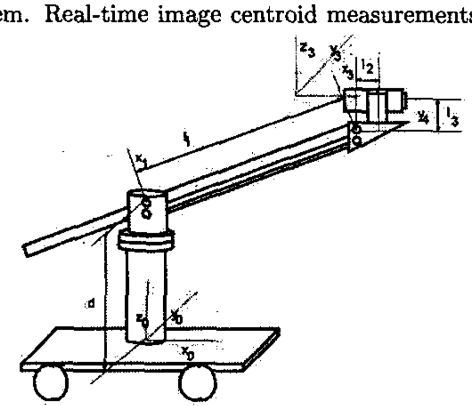


Figure 3: Denavit-Hartenberg Reference Frames

Joint number	α	d	a	θ
1	$-\frac{\pi}{2}$	d	0	θ_1
2	0	0	l_1	θ_2
3	$+\frac{\pi}{2}$	0	l_2	θ_3
4	$-\frac{\pi}{2}$	l_3	0	θ_4

Table 1: Denavit-Hartenberg link and joint parameters

are performed using a Newton Cognachrome color tracker, which is an embedded microprocessor that serially transmits the centroid's pixel location of a colored target. Additionally a joint encoder to measure horizontal booming was installed. To summarize, the boom-camera system's Denavit-Hartenberg reference frames and arm matrix are given in Figure 3, Table 1 and equation 1.

3 Section III: Technical Approach

Redundant joints can improve tracking. For example, people coordinate head-eye motions to track a moving target. Neuroscientists have long studied visual behaviors like saccades, gaze, and vestibular-ocular reflex to model the underlying coordination [9]. Corke illustrated their analogies, especially in terms of delay handling, to feedback and feed-forward compensators in a robot-camera context [3]. Beyond the biological underpinnings of coordinated redundant joints, an active research area pursued by stereo camera head developers, Oh and Allen showed that redundant joints can be coupled in the underlying control law they call partitioning [13]. They combined image and joint encoder data to visually and kinematically servo a 5-DOF gantry robot that tracks targets moving in a large workcell. The approach is appealing because of its analytical design framework; visual-servoing gain tuning for stable and desired transient response is based on the joint dynamic bandwidth and kinematic range-of-motion. We intend to expand their work by tuning gains that consider the human-in-the-loop aspects of our problem. Controlling two redundant DOF (horizontal booming and panning) in the presence of human-in-the-loop and visual-servoing is our key interest. We thus focus on only two DOF to gain insight on the stability issues. Our target only translates horizontally thus to keep its image centered in the camera's field-of-view, just horizontal booms and/or camera panning is needed and illustrated in Figure 4. Visually servoed pan is accomplished under proportional control as illustrated schematically by the following block diagram in Figure 5. Basically, it contains three principal blocks: the regulator (block), the pan-tilt head (PAN) and, the camera (CAM). As input, in the block diagram we have position.

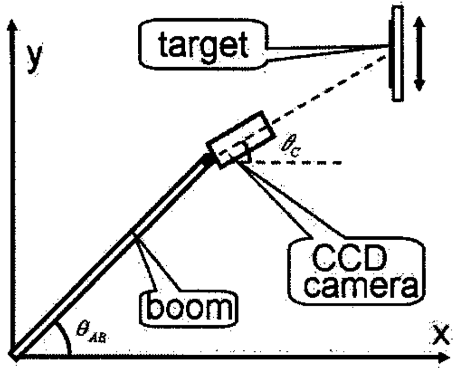


Figure 4: Top view of broadcast boom-camera system (gravity in z-direction). Horizontal booming and/or camera panning can keep a horizontally traveling target centered in the camera's field-of-view.

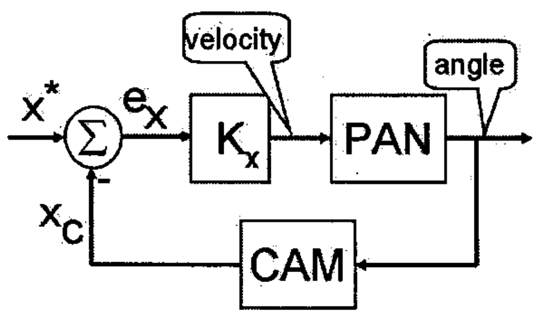


Figure 5: Closed-loop visually servoed camera pan Then the errors about image center will be:

$$e_x = x_c - x_{CI} \quad (1)$$

$$e_y = y_c - y_{CI} \quad (2)$$

x_{CI} nd y_{CI} are the values for the image center.

$$v_x = K_x * e_x \quad (3)$$

$$v_y = K_y * e_y \quad (4)$$

The net effect is velocity control and is called steering or piloting [1], [3].

4 Section IV: Experimental Results

In our tracking experiment, both the target moves and the operator booms. Figure 6 is a sequence of image stills grabbed while videotaping the experiment. The top and middle rows show the operator booming and the visually servoing of the camera head. The bottom row shows that the target is kept in the camera's field-of-view. Figure 7 is the time response plots of the pan and boom encoder and the pixel error. As the boom moves (right) the visually servoed pan compensates by performing a counter-rotation (left). The net effect is that the pixel error remains bounded between 80 pixels (roughly 1 cm), thus tracking the target. If gains are not properly tuned, ($K_x = 100$) unstable performance can arise as shown in Figure 8 (gain). Here, visual-servoing and boom motions are out-of-phase which leads to the growing oscillations in camera panning (left) and pixel error (right). Essentially the booming and visually-servoed panning are conflicting rather than the previous cooperation shown in Figure 7. Such out-of-phase performance could also arise if the target suddenly changes direction. To measure time, we are using a high resolution timer. It is easy to see that pan motor's encoder curve try to compensate the boom arm's encoder curve.

5 Section V: Conclusions and Future Work

This paper integrated visual-servoing to augment the tracking performance of camera operators. By reducing the number of DOF that needs to be manually manipulated, the operator can concentrate of course motions. A simple proportional control law was used to illustrate augmentation but its performance depends largely on proper gain tuning; man-machine shared control should cooperate and out-of-phase motions must be avoided. We are currently modeling and identifying the system dynamics to design a multi-input-multi-output controller to optimize the use of redundant DOF in the visual-servoing loop. In addition, Fitts law time-motion studies will be performed to incorporate the human-in-the-loop dynamics needed to successfully control camera motions. Adding somehow, the information offered by boom-arm's encoder will ensure an increase of stability. Also, we are currently working to develop such an algorithm.

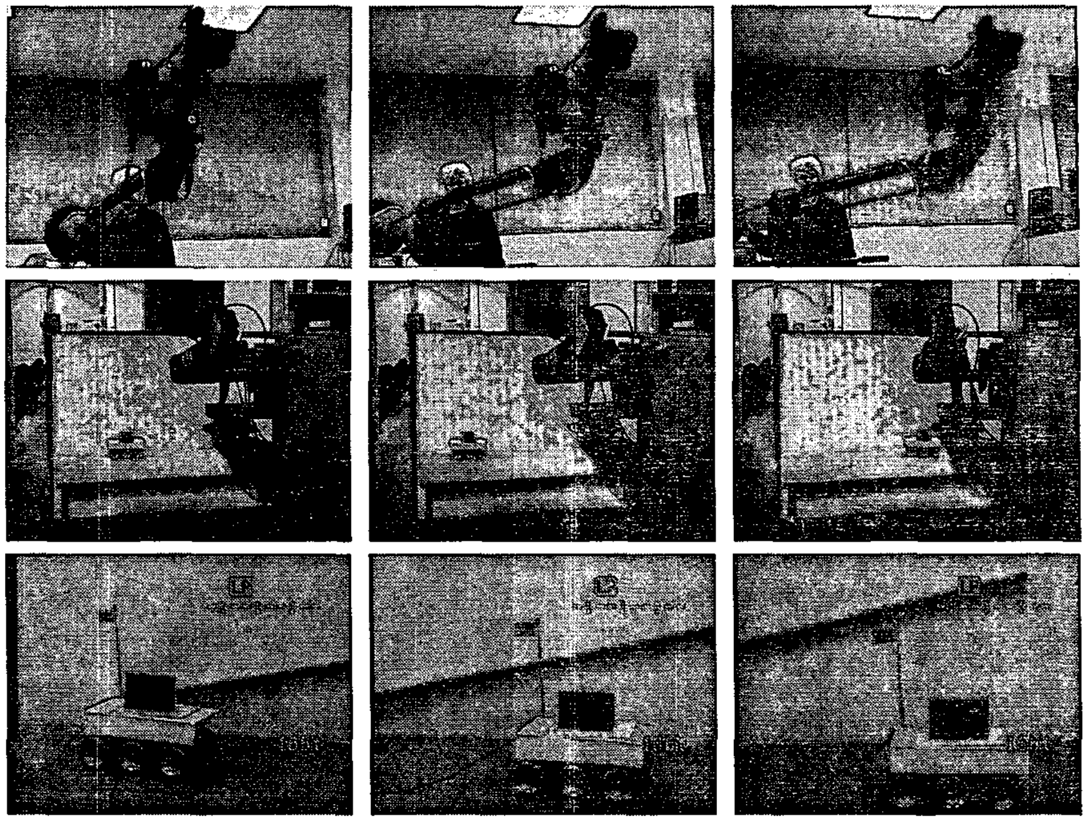


Figure 6: Several pictures taken during tracking

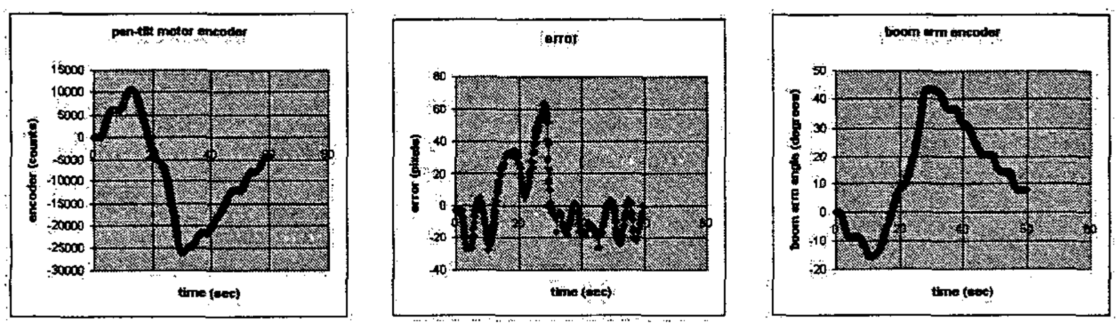


Figure 7: $K_x = 100$, a) encoder b)pixel-error c)boom-arm encoder

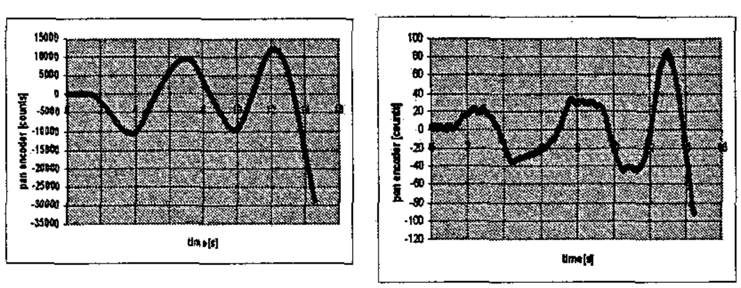


Figure 8: $K_x = 100$, a) encoder b)pixel-error

References

[1] Chaumette F., Rives P., Espiau B. *"Positioning of a robot with Respect to an Object, Track it and Estimating its Velocity by Visual Servoing"*, , Proceedings of the 1991 IEEE International Conference Robotics and Automation, Sacramento, CA, April 1991.

[2] Chaumette F., Rives P., *"Vision Based Control for Robotic Task "*,, IEEE International Workshop on Intelligent Motion Control, Bogazici University, Istanbul, August 1990

[3] Corke P. *"Design, Delay and Performance in Gaze Control: Engineering and Biological Approaches"*, in "The Confluence of Vision and Control" Springer Verlag, pp. 146-158, 1998

[4] Corke P., Hutchinson S. A. *"Real-Time Vision, Tracking and Control"*, Proceedings of the 2000 IEEE International Conference on Robotics Automation, San Francisco CA o April 2000,

[5] Dzialo K. A., Schalkoff R. J. *"Control Implications in Tracking Moving Objects Using Time-Varying Perspective-Projective Imagery"*, IEEE Transactions on Industrial Electronics Vol. IE-33, No. 3, August 1986.

[6] Feddema J. T., Lee G. C. S. *"Weighted Selection of Image Features for Resolved Rate Visual Feedback Control"*, IEEE Trans Robotics and Automation, V7 N1 2/91.

[7] Ferrier N. *"Achieving a Fitts Law Relationship for Visual Guided Reaching"*, Int. Conference Computer Vision (ICCV) Bombay India, pp 903-910, January 1998

[8] Hutchinson S., Hager Gregory D., Corke P. I. *"A Tutorial on Visual Servo Control"*, IEEE Transactions on Robotics and Automation vol. 12 no. 5 October 1996

[9] Krauzlis R., Lisberger S. *A model of visually-guided smooth pursuit eye movements based on behavioral observations.*, J. Computational Neuroscience, 1:265-283, 1994.

[10] Nagahama, K. Hashimoto K., Noritsugu *"Visual Servoing based on Object Motion Estimation"*, Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems

[11] Nagahama K., Kimoto T., Ebine T., Kimura H. *"Manipulator Control with Image Based Visual Servo"*, Proceedings of the 1991 IEEE International Conference Robotics and Automation, Sacramento, CA o April 1991

[12] Hashimoto K., Ebine T., T. Kimura T. *"Dynamic Visual Feedback Control for a Hand-Eye Manipulator"*, Proceedings of the 1992 IEEE / RSJ International Conference on Intelligent Robotics and Systems, Raleigh, NC o July 1992

[13] Oh, P. Y. Allen P. K. *"Visual Servoing by Partitioning Degrees of Freedom"*, IEEE Transactions on Robotics Automation vol. 17 no. I, February 2001

[14] Koivo A. J., Houshangi N. *"Real-Time Vision Feedback For Servoing Robotic Manipulator With Self-Tuning Controller"*, IEEE Transactions on Systems, Man and Cybernetics, vol. 21 no.1, January/February 1991.

[15] Papanikoloupolous N. P., Khosla P. K. *"Adaptive Robotic Visual Tracking: Theory and Experiments"*, IEEE Trans Automatic Control V38 N3 3/93.

[16] Sharkey P. M., Murray D. W., Vandevelde I. D. McLauchlan P.F. *"A Modular Head/Eye Platform for Real-Time Reactive Vision"*, Mechatronics, Vol. 3, No. 4, pp. 517-535, 1993

[17] Weiss L.E., Sanderson A. C. *"Dynamic Sensor-Based Control of Robots with Visual Feedback"*, IEEE J of Robotics and Automation V RA-3 N5 p404, 10/87.