Design and Control of a Two-Axis Gimbal System for use in Active Vision

by

Nitish Swarup

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Bachelor of Science in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1993

Author ....
Department of Mechanical Engineering
May 7, 1993

Certified by ....
Dr. J. Kenneth Salisbury, Jr.
Principal Research Scientist
Thesis Supervisor

Accepted by ....
Prof. Peter Griffith
Chairman, Department Committee

Design and Control of a Two-Axis Gimbal System for use in Active Vision

by

Nitish Swarup

Submitted to the Department of Mechanical Engineering
on May 7, 1993, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Mechanical Engineering

Abstract

This thesis describes the design and control of a two degree of freedom gimbal system for rapidly positioning a miniature CCD video camera. The gimbal system will be used to perform visual tracking experiments for the object-catching architecture built around the MIT Whole Arm Manipulator (WAM) at the Artificial Intelligence Laboratory. The gimbal device has ± 90 degrees of range of motion in the pan axis, and ± 45 degrees of range of motion in the tilt axis. Positioning accuracy is 4.36×10^{-4} radians on the pan axis and 7.14×10^{-4} radians on the tilt axis. Both of these accuracies approximate the maximum resolution of a 1/2" CCD camera fitted with a 16 mm lens. Two controllers were compared for their tracking performance on object position data. Although the computed torque controller resulted in better trajectory following at low frequencies, the bandwidth of the computed torque controller was much less than that of the position-derivative controller due to its poor robustness to unmodelled high-frequency dynamics. Experiments performed using a simulated camera mass on the device show peak accelerations of 1700 rad/sec^2 and peak velocities of 30 rad/sec on each axis.

Thesis Supervisor: Dr. J. Kenneth Salisbury, Jr.
Title: Principal Research Scientist

Design and Control of a Two-Axis Gimbal System for use in Active Vision

by

Nitish Swarup

Submitted to the Department of Mechanical Engineering
on May 7, 1993, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Mechanical Engineering

Abstract

This thesis describes the design and control of a two degree of freedom gimbal system for rapidly positioning a miniature CCD video camera. The gimbal system will be used to perform visual tracking experiments for the object-catching architecture built around the MIT Whole Arm Manipulator (WAM) at the Artificial Intelligence Laboratory. The gimbal device has ± 90 degrees of range of motion in the pan axis, and ± 45 degrees of range of motion in the tilt axis. Positioning accuracy is 4.36×10^{-4} radians on the pan axis and 7.14×10^{-4} radians on the tilt axis. Both of these accuracies approximate the maximum resolution of a 1/2" CCD camera fitted with a 16 mm lens. Two controllers were compared for their tracking performance on object position data. Although the computed torque controller resulted in better trajectory following at low frequencies, the bandwidth of the computed torque controller was much less than that of the position-derivative controller due to its poor robustness to unmodelled high-frequency dynamics. Experiments performed using a simulated camera mass on the device show peak accelerations of 1700 rad/sec^2 and peak velocities of 30 rad/sec on each axis.

Thesis Supervisor: Dr. J. Kenneth Salisbury, Jr.
Title: Principal Research Scientist

Acknowledgements

Acknowledging each person who contributed to the research behind this thesis would be a herculean task indeed. People from Kaydon Bearing, Maxon Motor, the Artificial Intelligence Laboratory, my colleagues in mechanical engineering, Graphic Arts, and Photo Quick in Central Square all provided invaluable assistance. My thanks to them all. I interrupted Akhil Madhani several times to ask for advice on eyeball building. Brian Eberman, John Morrell, and David Brock were also bothered many a time.

Special thanks go to my advisor, Ken Salisbury for being an excellent advisor and mentor during the past two years. His generosity and kindness have yet to be matched by any scientist I've seen. His keen robotics insight sparked many creative ideas during the design process.

My parents have always fully supported me in my endeavors. Thanks Mom and Dad for always being there.

Tarzen Kwok helped in the design and analysis of appropriate control algorithms for the gimbal system. Professor Haruhiko Asada and Professor Jean-Jacques E. Slotine also provided valuable advice.

I thank MIT for providing an outstanding environment in which to mature scientifically as well as personally. I am still amazed by the people and opportunities here.

And of course, thank *you*, God.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

1.1.1 Active Vision

Active visual sensing, that is, behavior in which the visual sensing and processing of information plays an integral part, has received increasing attention in the vision community due to its potential for providing high-resolution task-relevant information about the environment. An active vision system can be integrated into a robotic or computer vision system to provide visual feedback information to assist in task execution. In contrast to passive vision systems, an active vision system is active in that it can dynamically change visual parameters such as focusing, magnification, aperture, vergence, and field of view based on information from higher-level processing of the visual information. As a result, active vision systems maximize the visual efficiency of the system by directing computational resources to areas of interest within the task domain. Furthermore, several aspects of the spatial world can be determined more readily and accurately through active sensing, as opposed to passive sensing, when sensor dynamics and parameters can be controlled to disambiguate stationary objects.

Applications of active vision systems cover several task domains. Research applications abound in the robotics and vision fields. Researchers implementing robotic systems to explore closed-loop contact with the environment benefit from improved sensing information content, which results in more robust manipulation behavior. Vision researchers may explore the improvement in visual sensing achieved by induced motion, which may clarify objects in the visual domain. Real-world applications abound as well. Outer space operations benefit from active sensing in the maintenance of satellites, repair of structures, and assembly tasks. Manufacturing processes which involve tracking such as insertion, welding, painting, and finishing; inspection tasks that cover large areas with high resolution (such as checking rivets on an airplane exterior) and material handling all benefit from improved visual information. In general, real-time tasks which require a large domain to be perceived with high resolution are potential applications of active vision research.

Indirectly, active vision research is a rich medium for exploring the limitations of mechatronic systems. The highly integrated coupling between computation required by vision processing, high-performance mechanical systems, and control systems pushes the boundary of electro-mechanical information processing systems that employ different sensing and output modalities.

1.1.2 WAM Group Needs

The Whole Arm Manipulation (WAM) group at the MIT Aritificial Intelligence Laboratory is currently developing an active vision system for integration into and improvement of an adaptively controlled object-catching architecture [6] built around the high-performance MIT WAM manipulator [10], [11] . The current architecture is based on a passive vision system. Two cameras fixed at approximately 2 m away from each other, both 4 m away from the WAM, are being used stereoscopically to locate objects within the domain space of the WAM. Due to the fixed orientation of the cameras and inability to adapt to changing requirements, there are several limitations in the current architecture. These limitations include insufficient position-sensing accuracy, limited object perception, strict color requirements on the domain to filter object location, suboptimal utilization of camera resolution due to fixed magnification lenses, and a limited field of view. The new active vision system is being designed to overcome these limitations as well as to achieve a modular, robust design for use in other task-oriented domains.

As an intermediate step in the design of an active vision system, the WAM group decided it was wise to first construct a high-performance two-axis gimbal system for moving a small CCD camera to explore the tightly coupled issues involved in closed-loop visual servoing. Specifically, these issues involve the effects of finite image processing delays, joint sensing resolution, transmission friction, joint inertias, and controller choice on peak bandwidth and dynamic performance. Furthermore, higher level computational decisions that involve tracking algorithms, object perception, and image filtering can be implemented to quantify their performance. Once an optimal list of parameters is chosen based on these experiments, these parameters can then be more intelligently integrated into the design of a complete, stereoscopic active vision system. A possible solution may include using a pair of gimbal systems described in this thesis to provide the stereoscopy.

1.2 Thesis Overview

The purpose of this research was to design a two-axis gimbal system capable of servoing a miniature CCD camera to high precision with high dynamic performance. Design criteria were to keep the design modular, lightweight, compact, and portable. A set of control experiments were then performed to determine an optimal object-tracking controller for the gimbal system.

Chapter 2 describes the design of the gimbal system. Design criteria are motivated and described based on roughly defined WAM group requirements. Design options for the gimbal system are described with respect to kinematics, transmission design, actuation choice, and sensing. Finally, design choices and components are justified and described.

Chapter 3 discusses the gimbal dynamics and control system theory. Kinematic frames for the gimbal system are shown and the equations of motion are derived using a recursive computation algorithm. Both a position-derivative controller and computed torque controller are presented for robust tracking control.

Chapter 4 presents the results of the control system experiments. Each controller discussed in Chapter 3 is evaluated in response to an elliptical trajectory input. The bandwidth and tracking performance of each controller is compared and discussed.

Chapter 5 summarizes the results of this thesis and presents recommendations for future research.

Chapter 2

Design

2.1 Criteria

The design goals for the gimbal system were left fairly open-ended by the WAM group. The primary specification was that the device be able to move a CCD camera fast enough to track objects moving in the laboratory with sufficiently high precision and bandwidth. The mechanical bandwidth of the device should be larger than the computational bandwidth imposed by that of image processing and higher level decision making. Given these constraints, much time was spent trying to quantify design requirements to motivate design options and decisions.

2.1.1 CCD Camera Selection

Although closed-loop vision experiments were not conducted in this thesis, a CCD camera was selected prior to the gimbal design to guide design decisions. High modularity, however, to adapt to other cameras and optical devices, such as laser pointers, was still kept as a design criterion. Although a few smaller micro-miniature CCD cameras exist on the market, a miniature C-mount 8290-1000 Cohu color CCD camera shown in Figure 2-1 was chosen for several reasons. The addition of optical attachments such as zoom lenses and color filters will be facilitated given the much larger percentage of equipment on the market designed with C-mount adapters. The compromise in dynamic performance due to the larger inertial mass of the camera was deemed less important than the added modularity gained by adopting to the C-mount standard. A color camera was chosen because the WAM group is interested in exploring the use of color filtering to extract object information. Furthermore, the Cohu camera is designed to operate at shutter speeds of up to 10,000 Hz. The high shutter speed is mandatory in preventing image blur at high speeds. Finally, a lightweight HF16B-SND4-1 Fujinon auto-iris lens with focal length of 16 mm was purchased to give a field of view of 22°.

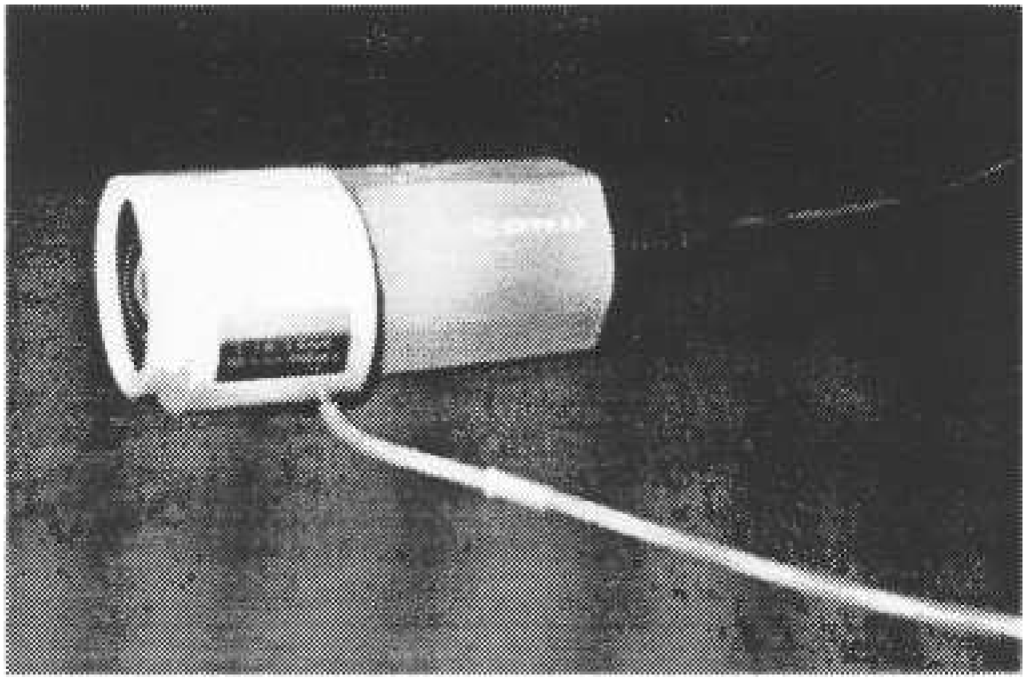

Figure 2-1: Cohu camera fitted with lens

2.1.2 Comparison with Human Eye

In many aspects, the gimbal system design simulates the human eyeball. The human eye is a system of muscles designed to position the retina on objects of interest rapidly and precisely. Because of the human eye's naturally high dynamic performance, the eye was used as a model to follow throughout the design process. Characteristics of the human eyeball are shown in Table 2.1 [3].

Characteristic	Value
Eye pan range of motion	$\pm$ 0.78 rad
Eye tilt range of motion	$\pm$ 0.78 rad
Peak acceleration	611 rad/sec^2
Peak velocity	10.5 rad/sec

Table 2.1: Characteristics of the human eye

Similar to the eye's construction, a design criterion was to position the focal center of the camera at the rotational center of the gimbal system. The visual and robotic kinematics are greatly simplified if the focal center lies coincident with the intersection of the pan and tilt axes. Furthermore, derivation of the equations of motion is simplified if only angular motions are involved. The range of motion was desired to be at least $\pm$ 0.8 radians on the tilt axis and $\pm$ 1.5 radians on the pan axis in order to achieve a large sensing domain.

2.1.3 Dynamic Performance

Dynamic performance was the key factor to be maximized in the design of the gimbal system. It is important for the dynamics of the pan/tilt device to be faster than those of the vision data acquisition and processing. As vision processing algorithms improve, the mechanical bandwidth of the gimbal system should not be the limiting factor in performance. The current standard sampling delay for image acquisition in CCD based video systems is 33 msec. Although the actual delay between new object position data is larger due to image processing computation time needed to threshold the image and calculate centroid position, new position information is calculated at approximately 30 Hz. Consequently, a goal was made for the position bandwidth to be larger than 15 Hz such that the gimbal system would reach the object's old position in approximately two image frames.

According to Table 2.1, peak accelerations of the human eye are 611 rad/sec^2. Peak velocities during saccadic motions reach 10.5 rad/sec. As a performance goal, peak acceleration and velocity of the gimbal system were desired to be $\sim$50% larger than those of the human eyeball. Lastly, given that objects in general achieve higher velocities in the horizontal direction than the vertical direction, position bandwidth on the pan axis should be given priority over position bandwidth on the tilt axis when applicable.

With respect to object speeds, current experimental speeds of objects in the WAM lab are between 3 and 5 m/sec [6]. In order to track on to a moving object with velocity v at 1 m away from the camera within the sampling time of two image frames, t_f, a necessary acceleration a can be found as follows. Assuming a linear case for small rotations, an object's position s_o can be described by

$$s_o(t) = vt \tag{2.1}$$

Assuming zero initial angular velocity and acceleration of the gimbal system, the gimbal's orientation s_g can be described by

$$s_g(t) = \frac{1}{2}at^2 \tag{2.2}$$

Setting $t_f = 2 \times 0.033$ sec and $s_o(t_f) = s_g(t_f)$,

$$a = \frac{2v}{t_f} \tag{2.3}$$

For object velocities of 5 m/sec, the necessary angular acceleration is 150 rad/sec^2. Hence, the peak acceleration specified of ~ 1000 rad/sec^2 easily meets this criterion.

2.1.4 Position Sensing Accuracy

One camera in itself cannot provide range information without apriori knowledge of object dimensions. An object is translationally located in three-dimensional space by three generalized variables; however, the gimbal system is only able to sense two rotational degrees of freedom. Nevertheless, it is important to have high position sensing accuracy for the following reasons. Primarily, inclusion of the gimbal system into a stereoscopic system will require that each subsystem have high positioning accuracy in order for triangulation algorithms to be accurate. A second justification is that controller performance will benefit from higher precision sensing. Although precise formulation of required position sensing accuracy for stereoscopic vision was not relevant to this thesis, the following derivation explains the maximum angular resolution obtainable from one camera due to the discretized nature of the CCD array. Consequently, the required joint sensing resolution on both pan and tilt axes will be naturally constrained.

A simplified one pixel model of a CCD camera is shown in Figure 2-2. The pixel lies on the camera's image plane a distance a behind a lens with focal length f . The pixel geometry is assumed to be square with width c. The object in focus lies a distance b in front of the lens. The relationship between the image plane distance and object distance from the lens is given by the lens law:

$$\frac{1}{f} = \frac{1}{a} + \frac{1}{b} \tag{2.4}$$

Using the law of similar triangles,

$$\frac{\delta}{b} = \frac{c}{a} \tag{2.5}$$

where δ is the width of the square image sensed by one pixel and $\frac{\delta}{b}$ is the arc subtended by one pixel. This arc is the maximum sensing resolution of the CCD array because displacement within this arc is undetectable by the camera. Substituting Equation 2.4 into 2.5,

$$\frac{\delta}{b} = c(\frac{1}{f} - \frac{1}{b}) \tag{2.6}$$

Given that object distances b are at least 1 m away from the camera, $\frac{1}{f} \gg \frac{1}{b}$ and Equation 2.6 reduces to

$$\frac{\delta}{b} = \frac{c}{f} \tag{2.7}$$

For the Cohu camera and Fujinon lens used, $c = 8 \times 10^{-6}$ m and $f = 16 \times 10^{-3} m$, resulting in a maximum angular visual resolution of 5×10^{-4} radians. The effective image resolution is less due to introduction of error from image processing algorithms. Thus, the joint position sensing resolution should approximate the maximum resolution of the CCD array. Larger sensing resolution may assist in smoother control, but

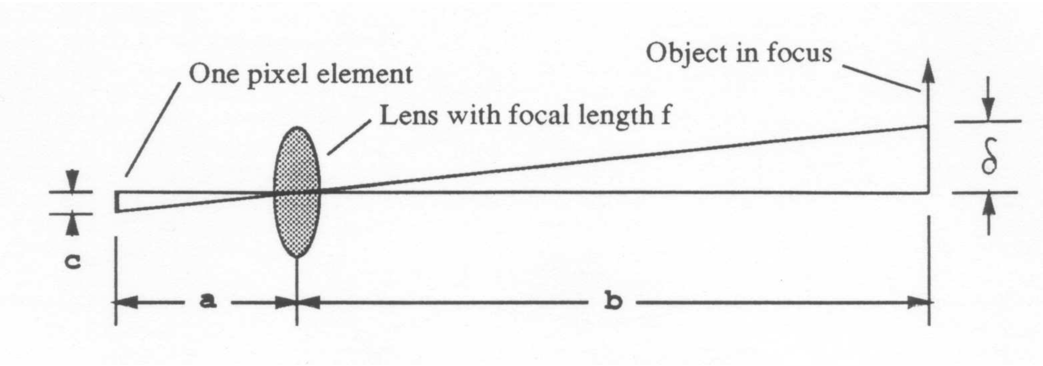

Figure 2-2: A simplified one pixel ccd camera model.

will not extract more precise object position information.

2.1.5 Compactness

A compact design was desired for several reasons. Primarily, high compactness in design minimizes design weight, encouraging minimal rotational inertias. A compact, lightweight design allows the device to be portable to several dynamic environments where weight is of primary consideration. A plausible application is to mount the gimbal system onto a robot arm to perform multi-degree of freedom visual servoing experiments or endpoint sensing. Given the importance of high-performance manipulator dynamics, lightweight space-saving sensors and effectors are important design requirements. Furthermore, a compact design encourages a smaller number of parts and a less complicated design, hence supporting design robustness. Lastly, a compact design is more aesthetically pleasing when considering an active "eyeball."

Table 2.2 summarizes the design criteria described above in Section 2.1.

Criteria	Required value
Position bandwidth	> 15 Hz
Peak acceleration	> 1000 rad/sec^2
Peak velocity	> 15 rad/sec
Positioning accuracy	$\sim 5 \times 10^{-4}$ rad
Pan range of motion	$\pm$ 1.5 rad
Tilt range of motion	$\pm$ 0.8 rad

Table 2.2: Summary of design criteria

2.2 Options and Choices

2.2.1 Kinematics

There were three design requirements to be met by the kinematic structure of the device. First, the gimbal system is to provide two rotational degrees of freedom: a pan motion, and a tilt motion. Secondly, the axis of pan rotation, axis of tilt rotation, and the focal center of the camera must intersect. Finally, the pan axis range of motion must be at least $\pm$ 85 degrees and the tilt axis range of motion must be at least $\pm$ 45 degrees with *unobstructed* visual sensing throughout the range of motion. The choice of kinematics is highly coupled with the transmission design and actuator choice. Throughout the design process, options were reconsidered due to inconsistencies with the transmission or actuators.

The initial aim was to minimize rotational inertia of the device by minimizing the number of moving parts in addition to the camera. An initial attempt was made to keep all actuators stationary, despite two independent degrees of camera motion. After several iterations, it was discovered that the most appropriate method for achieving stationary actuator position is to develop an $n + 1$ cable transmission involving three actuators [8]. This precludes a transmission with fixed length parameters. An $n + 1$ cable transmission, however, has a serious drawback with respect to sensing. Recalibration of the gimbal system would be required each time power to the system is turned off because the cables would become non-repeatably slack.

Given that an $n + 1$ transmission was unsatisfactory, the next best solution was to have one actuator move in a joint's coordinate frame while keeping one actuator stationary. Driven by the higher bandwidth priority for the pan axis, an initial decision was made to keep the tilt actuator stationary, and the pan actuator move with the camera in the tilt joint coordinate frame. Thus, the pan actuator would only be servoing the inertia of the camera, as opposed to the combined inertia of the camera, bearings, and actuator seen by the tilt actuator. This decision, however, disagreed with the third kinematic requirement. Since the tilt axis must intersect the focal point of the camera and perpendicularly intersect the pan axis, the motion of the pan axis is limited by the tilt joint support structure. Meeting the range of motion requirement for the pan axis would require expanding the plate supporting the pan actuator. This was deemed to be dynamically poor and aesthetically unsatisfactory given the added inertia. Consequently, given that the range of motion requirements of the tilt axis were less strict, the inverse solution was opted for despite an increase in the inertia seen by the pan joint. With this solution, the range of motion on the pan axis is a full $\pm$ 90 degrees, and the range of motion on the tilt axis is $\pm$ 45 degrees.

2.2.2 Actuation

The primary requirement in transmission design was to keep backlash minimal. Secondly, the transmission should be stiff to keep open-loop poles at higher frequencies than the frequency content of the controller. Several transmission designs were analyzed for overall performance and satisfaction of design requirements. At the top level, choices were split between direct-drive designs and transmission designs.

Direct Drive

A direct-drive design is appealing because problems due to backlash and transmission compliance are removed. Advantages of direct-drive actuators are high torque-to-inertia ratios, compactness in design, high stiffness, and zero backlash.

A subclass of direct-drive motors considered were limited angle torquers (LAT). Although LAT's provide excellent torque/current linearity, the range of linearity is typically limited to $\pm$ 35 degrees since only one phase is being driven. After this range, the torque output falls cosinusoidally, providing an undesirable nonlinear torque constant. The limited range of motion of LATs was considered to be unsatisfactory in meeting the kinematic range of motion requirements. The other class of direct-drive actuators considered was 3-phase brushed motors. 3-phase motors provide unlimited range of travel, yet they suffer from ripple torques of 5-7% of peak torque due to irregular reluctance paths and the discrete nature of commutation switching. Although a high servo rate controller could compensate for the torque ripple, 3-phase motors were eventually rejected due to limitations on dynamic performance imposed by structural design as explained below.

Transmission Options

Transmission designs are advantageous because a much smaller, lighter motor can be more ideally impedance matched with the joint load. Sensing resolution is magnified by the transmission ratio if the sensor is placed on the motor. Furthermore, the actuator can be located optimally to minimize link rotational inertia about the joint axis. The inertia, J_t, seen by a motor in a transmission can be expressed by the following equation:

$$J_t = J_m + \frac{J_l}{n^2} \tag{2.8}$$

where J_m is the rotor inertia of the motor, J_l is the inertia of the load by itself, and n is the transmission ratio. The maximum power is transmitted from the motor to the load when the optimal transmission ratio, n_o, is used, where

$$n_o = \sqrt{\frac{J_l}{J_m}} \tag{2.9}$$

A minimal backlash, high stiffness transmission was desired. A cable-driven transmission was not initially considered due to stretching over repeated use, hence requiring bothersome recalibration. The transmission selected was a fiberglass-reinforced rubber HTD belt drive consisting of drive and load sprockets because of its near zero backlash and high stiffness. For actuation, a high-performance lightweight dc servo motor with minimal brush friction was desired. Several motors were compared by torque constant, peak torque, brush friction, and package geometry. The end selection was a Maxon rare-earth magnet RE-025-055-35 dc servo motor due to its excellent torque constant and brush friction characteristics. Because of its rare-earth magnet construction, it achieves a torque constant of 3.34 oz-in/amp and a peak torque of 34.2 oz-in. Furthermore, due to its graphite brush design, brush friction is imperceptible by hand.

Transmission vs. Direct Drive

Both actuators for the design were intended to be similar. Although a direct-drive motor for one axis and a dc servo motor for the other is feasible, differences in motor design add complexity from a control standpoint in dealing with different motor dynamics.

The primary reason for deciding against a direct-drive design was poor dynamic performance. Since the kinematic structure described in Section 2.2.1 was fixed to having the tilt module being distal to the pan module, a direct-drive design would require that a wide diameter 3-phase motor lie coaxially with the camera. Furthermore, with the addition of an encoder and structural support, the rotational inertia seen by the pan actuator was deemed to be excessively large, especially since the actuating and sensing elements would have to be placed far off-axis to make room for the camera. This inertia was thought too large to meet the dynamic requirements specified in Section 2.1.3. In addition, given that compactness could be achieved by the belt drive design, sensing resolution was much higher, and dc servo motors were much less expensive, the HTD belt drive and Maxon motor combination was the actuation of choice.

2.2.3 Drive Electronics

Although selection of appropriate motor amplifiers is not directly relevant to the design of the gimbal system, amplifier choice is important when considering the entire active vision system. Conventional pulse-width-modulated (PWM) electronics may release high enough frequency drive signals to interfere with the video signal, severely degrading image quality. Linear amplifiers are better suited for an active vision system because the drive signal isn't modulated; however, they are generally more expensive due to their more complex analog output electronics and limited demand. In this thesis, PWM amplifiers were used because the vision processing was not incorporated into the control loop.

2.3 Description

The gimbal system design is depicted in Figure 2-3 complete with the chosen camera and lens. Figures 2-4, 2-5, and 2-6 display various orientations of the gimbal system. All structural parts were made of 6061T6 Aluminum for ease of machining, strength, and light weight.

2.3.1 Transmission

After estimating the rotational inertia of the device seen by the first joint, a transmission ratio of 7.2 was used to most closely impedance match the load to the motor given sprocket and belt size availability. Similarly, the tilt inertia was estimated and a tilt transmission ratio of 4.4 was used. The larger sprocket in each transmission was easy to mount given that it had a large surface area. The manufacturer's hub was machined off and mounting holes were used to affix the sprocket at the joint load. The drive sprockets presented a more difficult design problem. Given that the 3 mm motor shaft diameter was only 0.007" less than the sprocket inner diameter (I.D.) of 0.125", the sprocket I.D. was enlarged to 0.1875" to accommodate an aluminum shaft sleeve. A steel split hub and a steel gear clamp was used to rigidly, but removably, affix the sprocket assembly to the shaft sleeve. Permanent LocTite was used to glue the shaft sleeve and split hub to the motor shaft and drive sprocket, respectively. In order to tension the belt, slots for the motor faceplate holes were machined into the motor mounting plates.

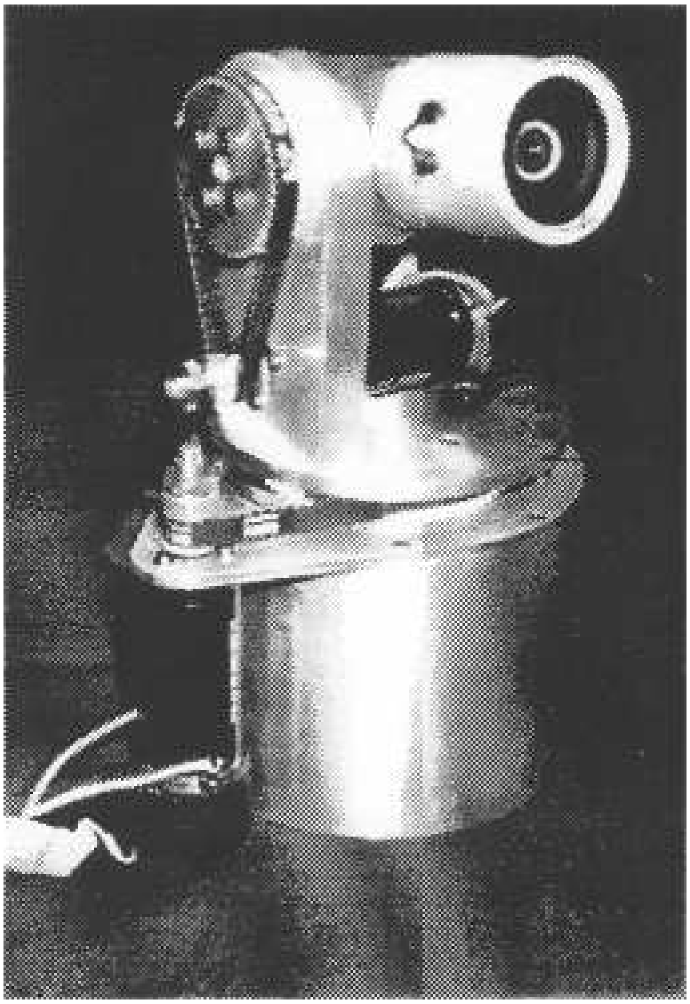

Figure 2-3: Gimbal system fitted with camera and lens

2.3.2 Minimized Rotational Inertia

An important goal throughout the design process was to minimize the rotational inertia of each axis to maximize dynamic performance. A challenging design problem was the design of the support structure for the tilt transmission, camera, and tilt axis motor. Given the minimal inertia requirement, a two-end supported structure was unsatisfactory. Hence, a cantilever plate was designed to mount the camera, motor, and transmission compactly. A front view and cross-section of the plate are shown in Figure 2-7. One Kaydon 4-pt contact bearing was used instead of two conventional angular contact bearings to minimize space. The cantilever plate was designed with

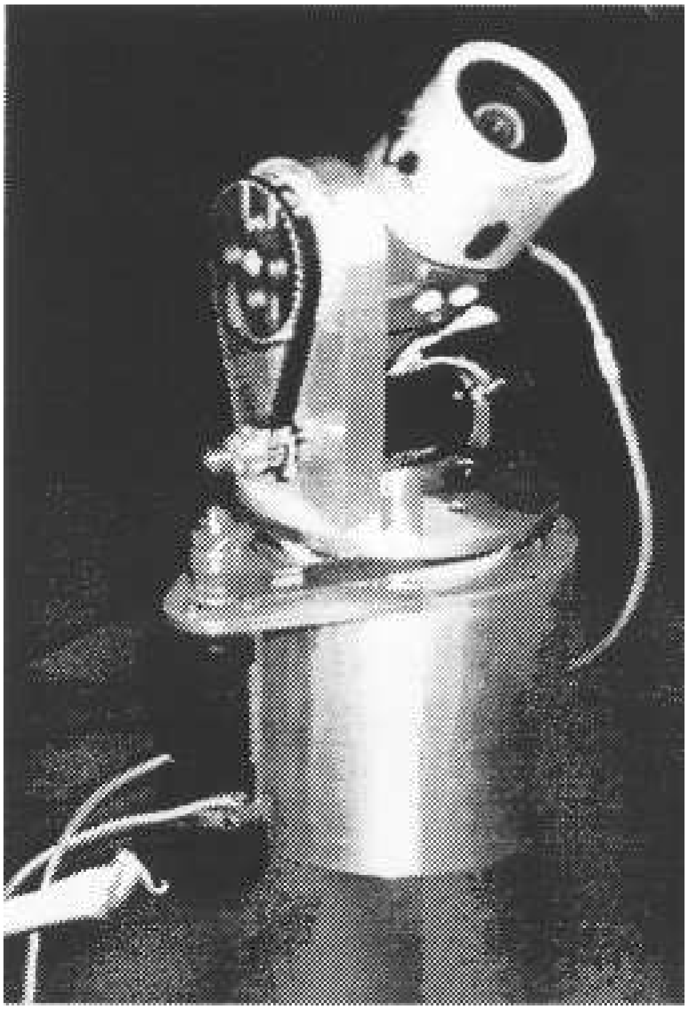

Figure 2-4: Gimbal system with camera pointing upward

side flanges to maintain high stiffness. Material was removed wherever possible to minimize part weight without sacrificing stiffness and structural strength.

2.3.3 Sensing

Optical encoders mounted on the motor are used to measure joint position. The resolution of the encoders is multiplied by the transmission ratio to give the joint position resolution. Given that each axis has a different transmission ratio, each axis has a correspondingly different resolution. Since each encoder has a line resolution of 500 pulses/revolution, the pan and tilt joints have position resolutions of 14400

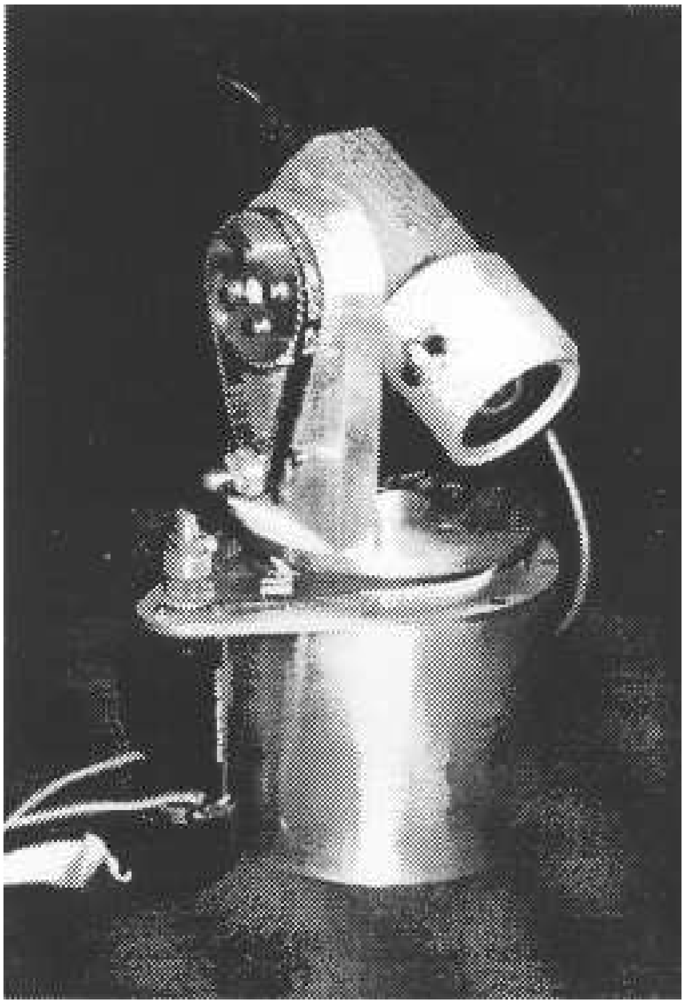

Figure 2-5: Gimbal system with camera pointing downward

and 8800 counts/revolution respectively. Converting to radians, the resolutions are 4.36×10^{-4} rad (0.87 pixels) and 7.14×10^{-4} rad (1.43 pixels) respectively. Both of these resolutions approximate the order of resolution provided by the CCD camera.

2.3.4 Cable Routing

Given the high speed motions of the camera, the camera cable must be routed such that the effects of its dynamics on the gimbal performance are minimized. Routing the camera cable along the outside of the gimbal results in a high stiffness seen by the camera as it rotates on the pan axis. If the cable is routed through the pan axis

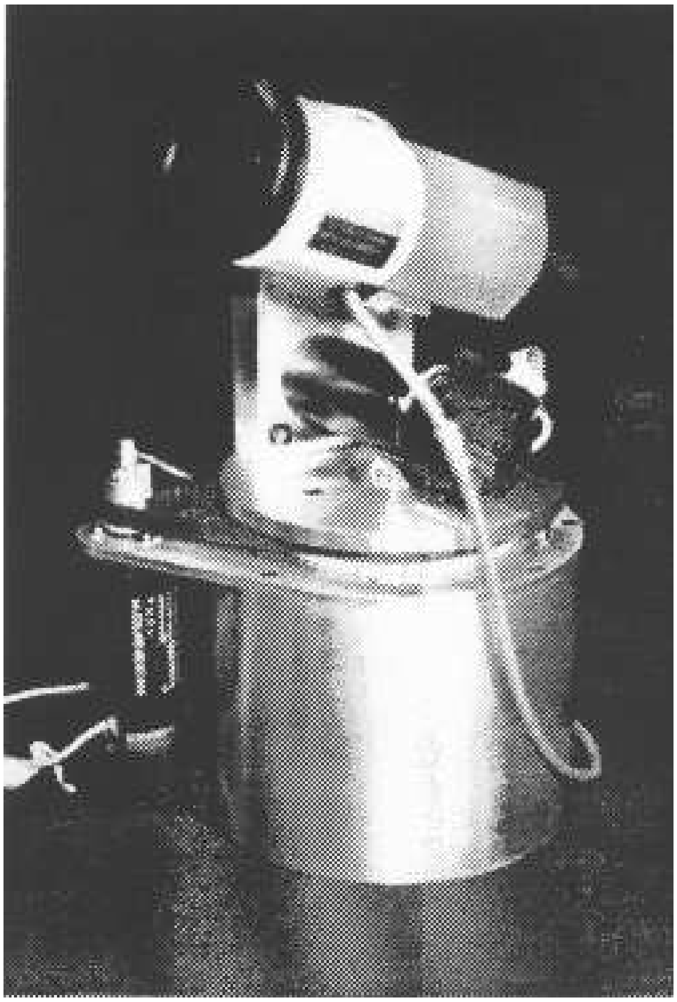

Figure 2-6: Gimbal system with camera pointing right

of rotation, then a much lower stiffness is seen by the camera because bending stress is almost eliminated and the torsional stress induced on the cable is spread out over the height of the gimbal. To meet this goal, parts that rotate about the pan axis were designed with a 1" clearance hole to allow for camera and motor cables to pass through. Ideally, the cable would be routed through the tilt axis of rotation as well; however, given the limited space, only the pan axis could be designed with a routing pathway.

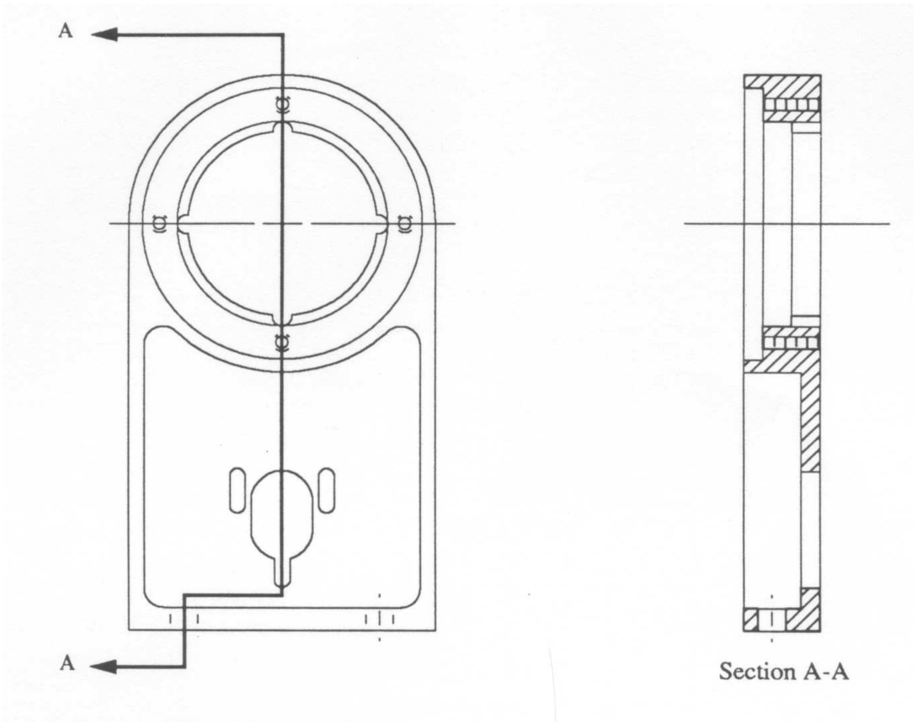

Figure 2-7: Cantilevered support structure for tilt axis.

Chapter 3

System Modelling

3.1 Kinematics

The kinematic frames of the gimbal system shown in Figure 3-1 are derived using the traditional Denavit-Hartenberg approach [1, p. 31]. The two generalized variables, θ_1 and θ_2, completely and independently describe the rotational motion of the gimbal system. The frames were chosen such that θ_2 is measured from the horizontal plane and θ_1 is measured from the location of the pan axis motor. Axis x_2 lies colinear with the optical axis of the camera. The origin of joint coordinate frames 1 and 2 intersect with the focal center of the camera.

3.2 Dynamics

The dynamics of the gimbal system were symbolically derived using the Luh-Walker-Paul algorithm [1, p. 126]. This algorithm greatly simplifies the derivation of the equations of motion by using a recursive computation of the Newton-Euler dynamic equations. A Lagrangian formulation of the closed-form dynamic equations is more complicated because all vectors and inertia tensors are defined with reference to the base coordinate system. For modelling purposes, the camera is modelled as a cylindrical mass whose center of mass coincides with the center of rotation. The equations of motion for the pan and tilt axis respectively are:

$$\tau_1 = (I + s_2^2 J_2 + c_2^2 J_1)\ddot{\theta}_1 + (J_2 - J_1)s_2 c_2 \dot{\theta}_1^{\,2} + (J_2 - J_1)s_2 c_2 \dot{\theta}_1 \dot{\theta}_2 \tag{3.1}$$

$$\tau_2 = J_1 \ddot{\theta}_2 + (J_1 - J_2)s_2 c_2 \dot{\theta}_1^{\,2} \tag{3.2}$$

where I is the rotational inertia of link 1 about y_1, J_2 and J_1 are the rotational inertias of link 2 about x_2 and z_2, $s_2 \equiv \sin(\theta_2)$, $c_2 \equiv \cos(\theta_2)$, and τ_1 and τ_2 represent the generalized forces for the pan and tilt axes. Note that the rotational inertia of link 2 about y_2 is also J_1.

The nonlinear terms in equations 3.1 and 3.2 arise from the cross-coupling in dynamics. Due to the asymmetry of the camera inertial tensor, centripetal and coriolis forces affect the dynamics. As a result of this cross-coupling, a multivariable methodology must be used when designing an appropriate control system.

3.3 Tracking Controller Design

There is a hierarchy of control processes in an active vision system, ranging from the primitive process, which servos the joint motors to desired positions, to the high level processes which interpret the visual data for object information that leads to task execution. This thesis is concerned with the device level control process, that of optimizing the tracking performance of the gimbal system given desired position data. Tracking controller design is an important issue for high performance rapid servoing. A high performance controller utilizes as much of the maximum mechanical bandwidth inherent in the device. Accurate tracking is necessary to maintain maximum visual information flow from objects of interest. Given the foveal nature of the optics to be used, only a limited field of view can be sensed. The maximum visual information is obtained when the object of interest lies on the optical axis of the camera. Hence, it is vital that a high performance controller be designed to rapidly position the field of view in response to fast moving objects. Steady-state positioning error is less a concern given that the lens' field of view allows for a small window of error. Once a robust device-level tracking controller is designed, higher level control can be designed by regarding the device level as an abstraction barrier.

3.3.1 Position-Derivative Control

A decoupled position-derivative (PD) controller for n axes is defined by

$$\tau_j := -k_{jD}\dot{\tilde{q}}_j - k_{jP}\tilde{q}_j \quad j = 1, ..., n \tag{3.3}$$

where τ_j represent generalized joint forces, $\tilde{q}_j$ are joint displacement errors, and k_{jP} and k_{jD} are position and velocity gains respectively. A PD controller is robust in that the controller will eventually servo the motors to the desired position, although a steady-state error is possible without integral feedback. Position and derivative gains can be tuned until a desired response is obtained. Large enough gains may cause instability, however, if they excite high-frequency unmodelled dynamics. Tracking performance of this controller is limited however because the control law does not take into account the configuration dependent dynamics of multi-degree of freedom systems.

3.3.2 Computed Torque Control

If the dynamics of a nonlinear system are represented by

$$\tau = \mathbf{H}\ddot{\mathbf{q}} + \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) \tag{3.4}$$

where τ are the generalized forces, $\mathbf{q}$ are the generalized coordinates, $\mathbf{H}$ is the system inertia tensor, and $\mathbf{h}$ represents the nonlinear terms, then a nonlinear controller can be designed to linearize the system dynamics. Specifically, the following decomposition

$$\tau = \mathbf{H}\ddot{\mathbf{u}} + \mathbf{h} \tag{3.5}$$

linearizes the system to one of controlling the desired acceleration, $\ddot{q_{dj}}$. Hence, the linearized system is

$$\ddot{\mathbf{q}} = \mathbf{u} \tag{3.6}$$

Position and derivative feedback terms can be added to the control law to account for inaccuracy in modelling of $\mathbf{H}$ and $\mathbf{h}$. Thus the controller is decomposed into n decoupled double integrators of the form

$$u_j := \ddot{q_{dj}} - k_{jD}\dot{\tilde{q}}_j - k_{jP}\tilde{q}_j \quad j = 1, ..., n \tag{3.7}$$

By explicitly accounting for the nonlinear dynamics of the system, a computed torque controller linearizes the *controlled* system and improves tracking performance.

Limitations of computed torque control arise primarily from inaccuracy in estimating the parameters involved in $\mathbf{H}$ and $\mathbf{h}$, accomplished through off-line mass parameter identification. This identification is done by exciting the system with various trajectories, recording generalized torque and variable data, and then performing a least squares fit between the two sets to predict mass parameters. In order for the results to be meaningful, the trajectory must be carefully tailored to sufficiently excite all the dynamic modes of the system. Since acceleration is not physically measurable, numerical methods must be used to differentiate position and velocity data. These computations, however, attenuate magnitude and introduce phase lag into the calculated acceleration, introducing error. Secondly, an accurate model of the system must be known in order for the computed torque controller to be effective. Time delays in the actuators, structural frequencies, and friction are often unmodelled and can reduce the control system bandwidth [1, p. 140].

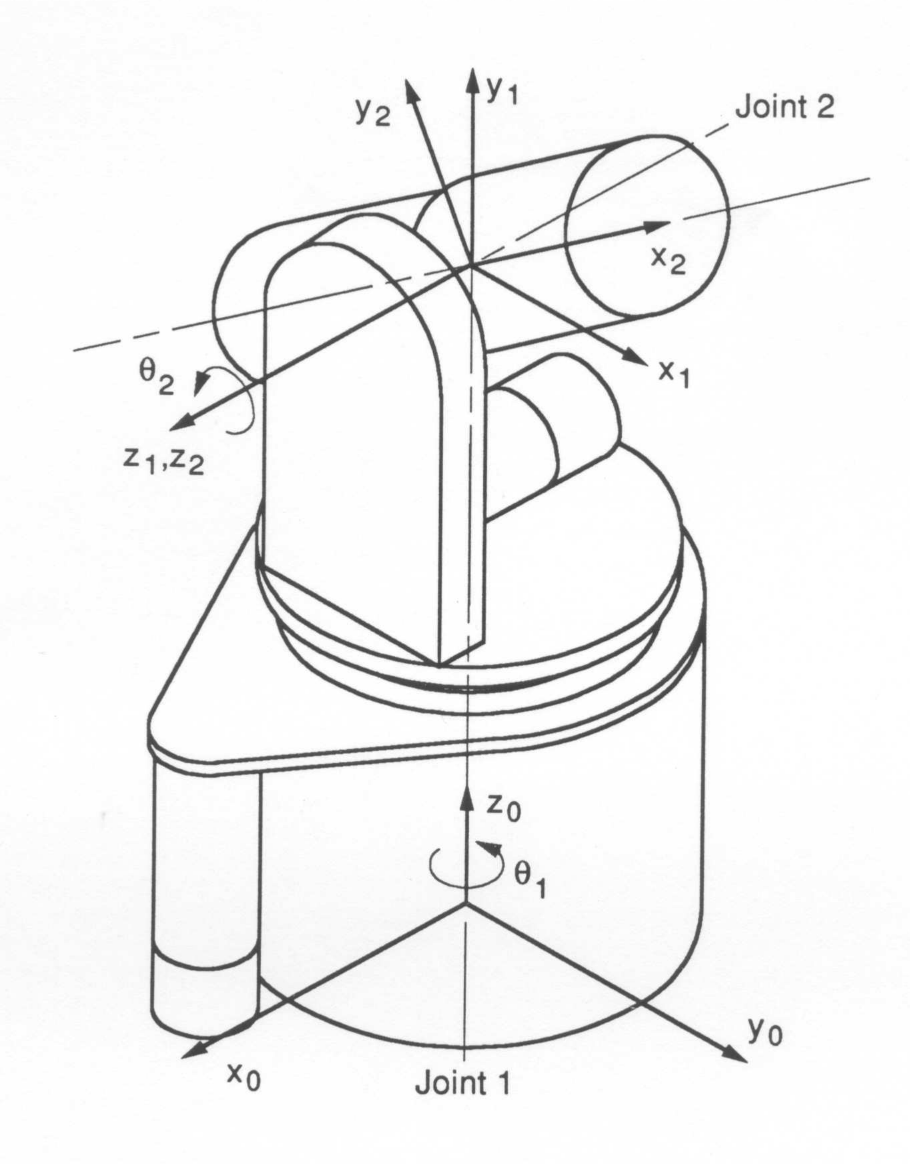

Figure 3-1: The gimbal system kinematic frames.

Chapter 4

Results

4.1 Controller Performance

In order to quantify the performance of both tracking controllers, an elliptical trajectory defined by Equations 4.1 and 4.2 was fed in to each controller and resulting joint position data was recorded.

$$\theta_{1d} = 0.75\sin(wt) \tag{4.1}$$

$$\theta_{2d} = 0.50\cos(wt) \tag{4.2}$$

An elliptical trajectory was chosen in order to quantify each controller's performance with respect to the coupled nonlinear dynamics. The drive frequency w was varied in order to observe the bandwidth limits of each controller. Although sensing of joint position is not measured exactly due to the noncollocated nature of the system, the critical frequency of the belt transmission was assumed to be much higher than that of the highest drive frequency tested.

During test runs, a cylindrical piece of nylatron with similar weight and size to the camera was mounted on the gimbal system to prevent camera damage. A PC-based control station was used to perform all control tests. The hardware consists of a 33 Mhz 386DX PC, a 4-axis Motion Engineering Motor Control board, and 2 Copley 303 amplifiers. Software was written using the Borland C++ programming environment under DOS. 2000 sample points at a servo rate of 1165 Hz were taken for each data run.

4.1.1 PD Controller

Prior to inputting the elliptical trajectory, the position and derivative gains for the PD controller were tuned by inputting a step function and modifying the gains. After several iterations, a response to a step function with $\pi/6$ amplitude characterized by less than 5% overshoot and 50 msec rise time for each axis was obtained. Using $\frac{1}{rise\ time}$ as an estimate of the controller bandwidth, a 20 Hz bandwidth was achieved. Faster rise times were obtainable; however, in order to prevent the system from saturating the amplifiers at high saccades, a compromise was made by reducing the gains to allow the gimbal to achieve larger saccades although at shorter rise times.

The response of the PD controlled system to the desired elliptical trajectory is shown in Figure 4-1. Although the tracking performance decreases marginally with increasing w, the overall shape of the trajectory closely follows that of the desired ellipse. The orientation of the trajectory, however, is that of an ellipse rotated slightly counterclockwise. As w increases, the actual trajectory becomes narrower and the overshoot outside of the desired trajectory near the upper left and bottom right hand areas of the curve worsen.

4.1.2 Computed Torque Controller

Mass parameter estimation

In order to estimate the mass parameters for $\mathbf{H}$ and $\mathbf{h}$ as described in Section 3.3.2, a triple sinusoid trajectory of the following form was inputted to the system to excite the dynamics. Figure 4-2 displays the trajectory.

$$\theta_{1d} = \frac{\pi}{6}\sin(6\sin(2\sin(0.6wt))) \tag{4.3}$$

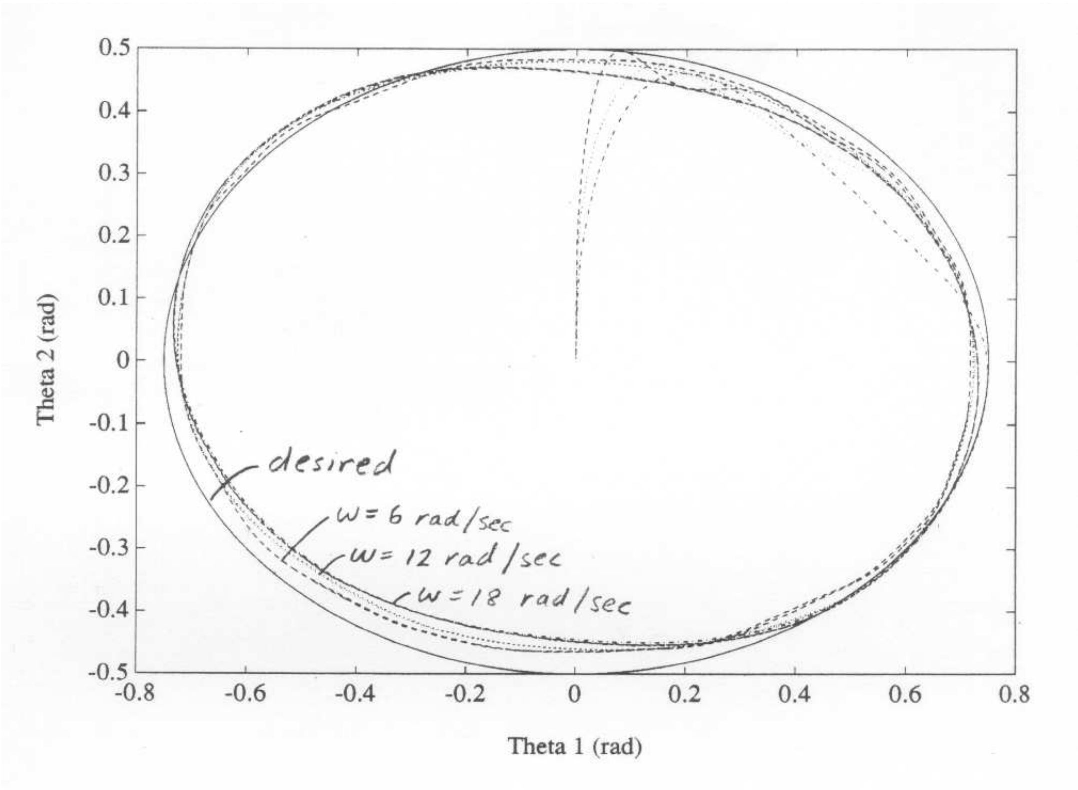

Figure 4-1: Response to elliptical trajectory under PD control for various w.

$$\theta_{2d} = \frac{\pi}{6}\sin(6\sin(2\sin(wt))) \tag{4.4}$$

These functions were chosen in an attempt to excite each of the dynamic modes comparably. From the figure, it can be seen that there are several periods where the two joints are experiencing different accelerations and velocity with respect to each other.

From this data, a least squares analysis was performed to find the mass parameter estimates. A comparison of the signal energies in the least squares analysis showed that the peak signal energies of the nonlinear terms were only 0.56% of the peak energies due to the individual joint acceleration terms. Hence, it was concluded that within the range of motion being tested on the gimbal system, the nonlinear dynamic modes were not being excited enough to be of sufficient concern to the controller

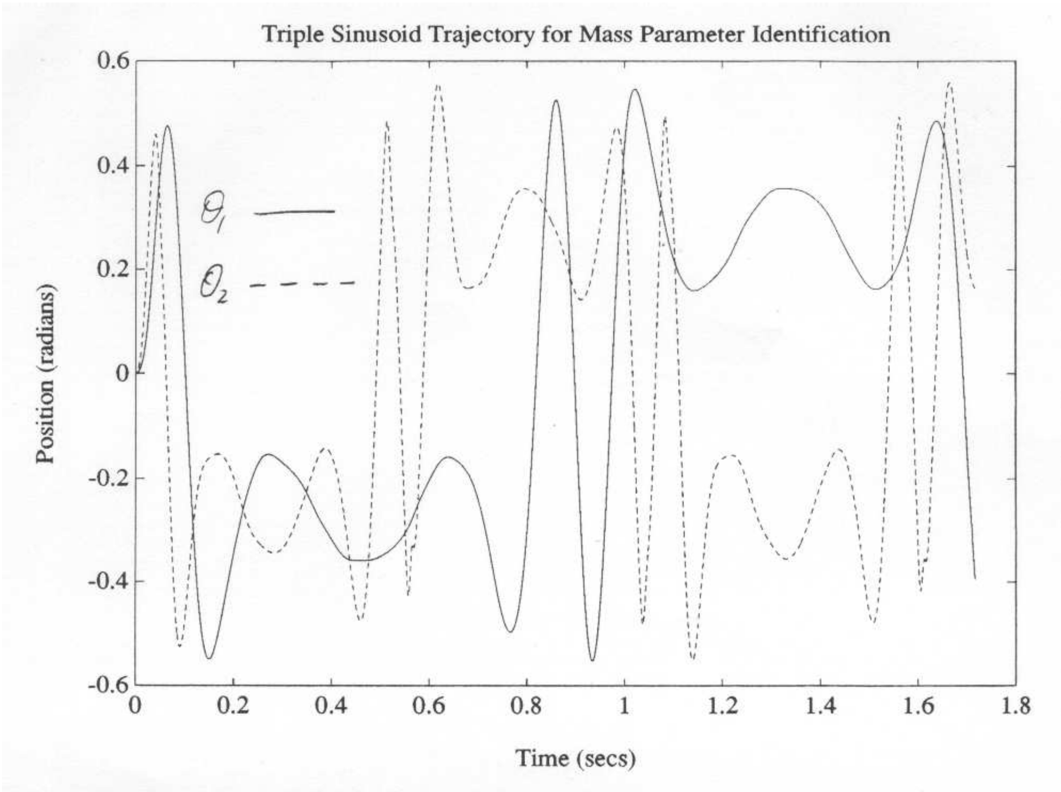

Figure 4-2: Triple sinusoid trajectory used for mass parameter identification

structure. Consequently, a simplified decoupled model of the following form

$$\tau_1 = I\ddot{\theta}_1 \tag{4.5}$$

$$\tau_2 = J\ddot{\theta}_2 \tag{4.6}$$

was used for the mass parameter estimation and computed torque control. The ratio of I/J obtained from the least squares analysis closely approximated the ratio estimated by hand calculation.

Once mass parameters were obtained, the controller according to Equation 3.7 was used to track the elliptical trajectory described by Equations 4.1 and 4.2 The results for the gimbal tracking performance at various w are shown in Figure 4-3. With increasing frequency, the computed torque controller's performance degrades rapidly. At w=18 rad/sec, the attenuation in amplitude for the pan axis is 90%.

Attenuation on the tilt axis, although less than that of the base, still approaches 30%. The bandwidth of the computed torque controller is obviously much lower than that of the PD controller. Although the bandwidth is lower, the overall shape of the trajectory is followed much more accurately than that of the PD controller. The response at w=6 rad/sec more closely resembles an elliptical path than the path followed by the PD controller at the same drive frequency.

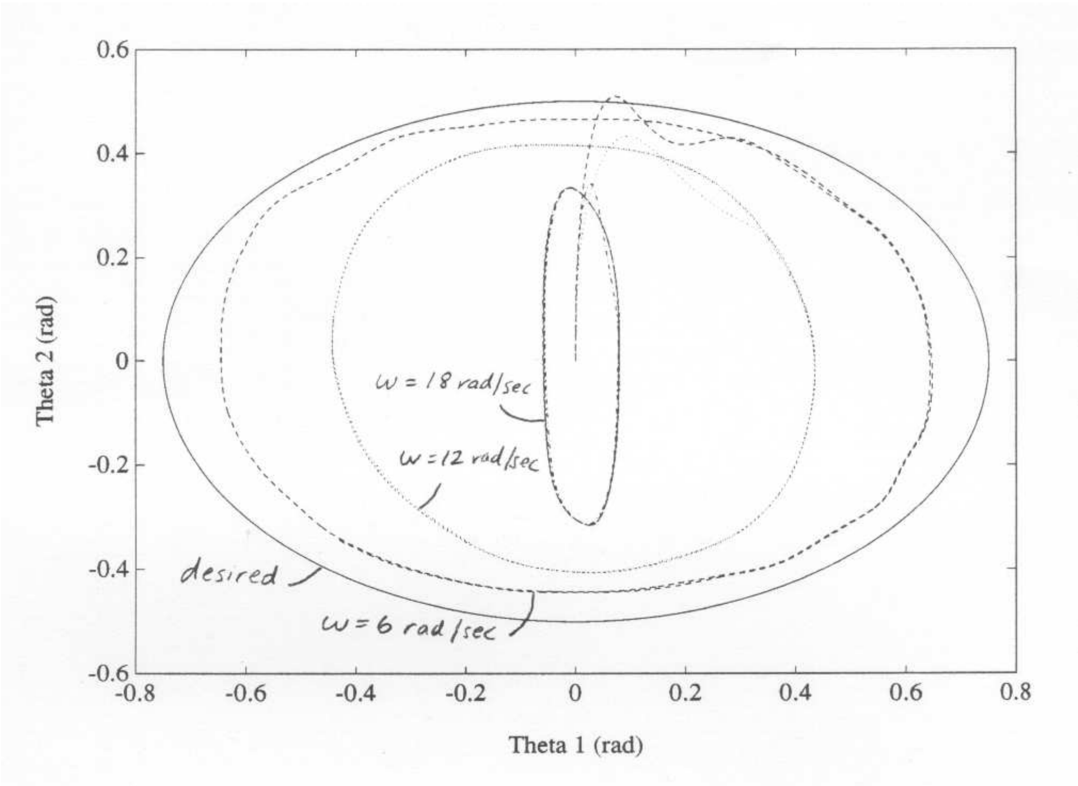

Figure 4-3: Response to elliptical trajectory under computed torque control for various w.

4.2 Discussion

In evaluation, the PD controller performs much better than the computed torque controller in tracking performance. Although the computed torque controller feeds forward the inverse plant dynamics in an attempt to linearize the system, the controller is characterized by low robustness to modelling uncertainty and high frequency unmodelled dynamics. The contribution of inaccuracy in mass parameter estimation to the low bandwidth is assumed to be minimal. Initial assumptions would indicate that the gyroscopic effects associated with an asymmetric mass rotating about two axes would affect the dynamics greatly. However, given the constrained range of motion of the gimbal system, joint acceleration isn't given enough time to accumulate the nonlinear velocity term. Thus, the constrained range of motion in effect decouples the dynamics to that of two individual Single Input Single Output (SISO) systems. Rather, the frequency component of the computed torque controller was high enough to excite dynamics in the system that were not explicitly accounted for in the system model. A possible element that may be excited is the motor dynamics, given that the mechanical time constant for the Maxon RE-025-055-35 motor is only 4.42 msec. With respect to quantitative performance benchmarks, peak accelerations and peak velocities reached during test runs were on the order of 1700 rad/sec^2 and 30 rad/sec respectively.

Chapter 5

Conclusions

In summary, a two-degree of freedom gimbal system was designed and built to provide a robust tool for integration into an active vision system. The design is modular to several small CCD cameras and other optical devices. Furthermore, positioning accuracy is on the order of the maximum resolution of a standard 1/2" CCD camera with a 16 mm lens. Peak accelerations of 1700 rad/sec^2 and peak velocities of 30 rad/sec surpass the dynamic performance of the human eye by an order of three, making the system well adaptable to faster vision processing systems as they become available. Table 5.1 summarizes key quantitative characteristics of the gimbal system.

Characteristic	Value
Peak acceleration	$\sim$1700 rad/sec^2
Peak velocity	$\sim$30 rad/sec
PD position bandwidth	20 Hz
Pan positioning accuracy	4.36×10^{-4} rad
Tilt positioning accuracy	7.14×10^{-4} rad
Pan range of motion	$> \pm 1.5$ rad
Tilt range of motion	± 0.8 rad

Table 5.1: Characteristics of the gimbal system

Control experiments show that a high servo rate PD controller satisfactorily tracks moving objects with minimal overshoot. Least squares analysis shows that the gyroscopic coupling effects of the gimbal system are insignificant within the motion space necessary for visual sensing. Although the computed torque controller feeds forward the inverse plant dynamics, the high frequency content of the controller resulted in a lower overall system bandwidth due to excitation of unmodelled dynamics and delays. As a result, the PD controller is more robust and has much higher bandwidth.

5.1 Recommendations for Future Work

Design improvements of the gimbal system center around the transmission. The small drive sprockets used on both axes caused visible ripple in the motion of each joint due to discontinuous belt transition on and off to the sprocket teeth. Although the ripple was not significant, it contributes static and dynamic friction to torque noise seen by the joint, reducing positioning and dynamic performance. A steel cable transmission characterized by a very smooth, low friction drive might be implemented to compare performance results. Dynamic improvements may outweigh the disadvantages of cables, specifically cable stretching, with respect to positioning requirements.

Further research on high performance tracking controllers needs to be done to examine algorithms which don't excite the high frequency dynamics of the gimbal system, yet achieve higher performance than regular PD. An adaptive controller might be implemented to examine its robustness and ability to adapt parameters in the control law. Furthermore, more extensive modelling of system elements, such as motor dynamics, may need to be done to obtain a more accurate system model and upper bound on controller frequency content.

Bibliography

[1] Asada, Haruhiko and Jean-Jacques E. Slotine. *Robot Analysis and Control*. John Wiley & Sons, Inc., New York, NY. 1986.

[2] Ballard, Dana Harvey and Christopher Harvey Brown. *Computer Vision*. Prentice-Hall, Englewood Cliffs, NJ. 1982.

[3] Fiala, John C., et al. "TRICLOPS: A Tool for Studying Active Vision." National Institute of Standards and Technology, Gaithersburg, MD. 1993.

[4] Guyton, Arthur C. *Basic Neuroscience: Anatomy and Physiology*. W.B. Saunders, Philadelphia, PA. 1991.

[5] Horn, B. K. P. *Robot Vision*. MIT Press, Cambridge, MA, 1990.

[6] Kimura, Hiroto. "Adaptive Visual Tracking Algorithms for 3-D Robotic Catching." Master's thesis, Dept. of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, August, 1992.

[7] Krotkov, Eric Paul. *Active Computer Vision by Cooperative Focus and Stereo*. Springer-Verlag, New York, NY. 1989.

[8] Mason, M.T. and J.K. Salisbury. *Robot Hands and the Mechanics of Manipulation*. MIT Press, Cambridge, MA, 1985.

[9] Ogata, Katsuhiko. *Modern Control Engineering*. Prentice Hall, Englewood Cliffs, NJ, 1990.

[10] Salisbury, J.K., W.T. Townsend, B.S. Eberman and D. DiPietro. "Preliminary Design of a Whole-Arm Manipulation System (WAMS)," *proc. 1988 IEEE International Conference on Robotics and Automation*, Philadelphia, PA, April 1988.

[11] Townsend, W. T. "The Effect of Transmission Design on Force-controlled Manipulator Performance." Ph.D. Thesis, Dept. of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 1988.