A LARGE VOLUME CLOSE-RANGE PHOTOGRAMMETRIC SYSTEM

by

PETER KETIL MANSFIELD

B.S. Columbia University (1979)
B.S. Saint Lawrence University (1979)
S.M. Mechanical Engineering, Massachusetts Institute of Technology (1983)

SUBMITTED TO THE DEPARTMENT OF
MECHANICAL ENGINEERING
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN MECHANICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 1990

Signature of Author_____
Department of Mechanical Engineering
May 1, 1990

Certified by_____
Professor Robert W. Mann
Thesis Supervisor

Accepted by_____
Professor Ain A. Sonin
Chairman, Departmental Committee on Graduate Studies

A LARGE VOLUME CLOSE-RANGE PHOTOGRAMMETRIC SYSTEM

by

PETER KETIL MANSFIELD

Submitted to the Department of Mechanical Engineering
on April 12, 1990 in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

Close-range analytical photogrammetry is a common and accepted method for remotely measuring spatial movement. Conflicting goals for high measurement accuracy and a large measurement volume often limit photogrammetric system utility. This is particularly true of applications to human movement measurement in biomechanics research. A technique for enlarging close-range photogrammetric measurement volumes has been designed and implemented. Two standard close-range photogrammetric cameras (SELSPOT) are each fixed in a vertical position. A flat circular mirror is mounted above each camera at 45 degrees to a rotational servomotor axis colinear with the camera axis. The principal axis of each camera's mirror image, or virtual camera, sweeps out a horizontal plane as the mirror is rotated. Real-time control of mirror position permits spatial tracking of objects. Novel calibration techniques identify fundamental camera-to-mirror space and global-to-mirror space transformations. Camera and mirror position data input to these transformations permits rapid estimation of three dimensional object information with a mean absolute static accuracy of approximately 0.2% of a tested measurement range of more than 20 cubic meters.

Thesis Committee:

Robert W. Mann (Chrmn.)	Whitaker Professor of Biomedical Engineering (MIT)
William K. Durfee	Assistant Professor of Mechanical Engineering (MIT)
Woodie C. Flowers	Professor of Mechanical Engineering (MIT)
Zvi Ladin	Assistant Professor of Biomedical Engineering (Boston U.)
Derek Rowell	Associate Professor of Mechanical Engineering (MIT)

Acknowledgements

I probably would not have pursued a Doctorate had it not been for my thesis advisor, Professor Robert W. Mann. He created a unique learning environment in the Newman Laboratory that was just too hard to leave after completing my Master's. For that, along with his insights and his determination to stand up for me and provide me with precious financial support, I am grateful.

The other members of my thesis committee, Professors William Durfee, Woodie Flowers, Zvi Ladin and Derek Rowell all contributed in their own ways. The insights, encouragements, and directions were much appreciated.

Having one of the world's foremost authorities in the technological aspects of photogrammetry applied to human motion studies review my work was very important to me. For this reason, special thanks goes to Dr. Herman J. Woltring, of Eindhoven, The Netherlands, for his last-minute effort in reviewing and commenting on my thesis manuscript.

I am convinced this thesis would have required another two years to complete had it not been for the help I received from fellow student Pat Lord. His knowledge of C and the UNIX operating environment and his willingness to take over the time consuming, and often thankless, tasks of maintaining the SUN system and TRACK provided me with precious freedom to help me finally get finished. I hope he is as fortunate when his time comes.

Fellow graduate student and apartment mate, Mike Murphy, arrived at MIT the same term as I and is finishing the same term as well. He's seen all of my work and has been an often needed supporter and critic. There is no doubt this thesis reflects his inputs.

Thanks also to my UROP students, Sheila Eglostein, Andy Karduna, and Alethea Morris. I only hope they learned as much from me as I from them.

To all my other friends in the lab, students, faculty and staff, past and present — you've helped maintain my sanity. Sometimes you just have to bounce your frustrations

off of someone. You know I've had a few. It was great to have the opportunity to know you all and I hope we'll keep in touch in the future.

To all my friends outside of the lab — what can I say, I'm finally done! You'll have to think of something else to ask me now. I hope to start having a chance to see you all again on a regular basis.

And of course to my parents, Sonja and Jack, and the rest of the Mansfield Family. I know this has probably seemed like a long drawn-out process to you and knowing that you don't fully understand what it is I've been doing probably made it seem even more so. But through it all I know you sensed the importance this has had to me and I thank you for always being there with your love and support.

This research was sponsored and student support was provided for in part by the U.S. Department of Education, National Institute for Disability Research and Rehabilitation (award no. H133E80024), and a Fairchild Fellowship. This work was performed in the Eric P. and Evelyn E. Newman Laboratory for Biomechanics and Human Rehabilitation at the Massachusetts Institute of Technology.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

For as long as there have been humans, there have been questions. Initially, questions were easily answered though simple experimentation. However, it was soon discovered that every answer only generated more questions. Through an evolving process, humans developed logical, analytical and experimental techniques for dealing with this expanding number of questions. This eventually became known as *science.*

Answering questions led to an accumulation of knowledge. Using their knowledge, humans developed "tools" for making their life easier. They designed and constructed tools for specific purposes, to fill certain needs, based on their "scientific" understanding of the world about them. This eventually became what we now call *engineering.*

Often, addressing a need through engineering, in turn, defines a set of scientific questions which must first be answered. Alternatively, attempting to answer a scientific question often creates needs requiring engineering solutions. Scientific discoveries in one discipline are often used by engineers to develop experimental apparatus and instru-

mentation for addressing scientific questions in other unrelated disciplines. In this way science and engineering have evolved a symbiotic relationship — each, in critical ways, dependent on the other for advancement.

In keeping with this relationship, this thesis involves the design, construction and implementation of sophisticated measurement instrumentation that evolved from a desire to answer fundamental questions relating to human rehabilitation. A basic question in human rehabilitation examines how we may provide or restore a degree of functional capability to a person who has some form of functional deficit. Of particular interest to this thesis are issues of human mobility. For "normal" humans the ability to walk or run is taken for granted. Free mobility, or the ability to move about in, and adapt to, an evolving environment, is so much a part of the normal everyday experience that the underlying complexity is rarely considered. That is, until we experience a loss of one or more of the functional components which make free mobility possible.

One of these functional components has nothing to do with our feet or legs. It is our vision system. Without eyesight, humans lack a critical source of information required for free mobility (we have all, at one time or another, fumbled in the dark). Substitutes for vision, such as a cane or seeing eye dog, only restore a very small portion of the missing information. Research into methods for improving the amount and quality of vision substitute information requires, for example, scientific investigations into the nature of "information" and how it is perceived, and engineering development of methods of acquiring, interpreting and "displaying" the vision substitute data.

Central to the investigations is the need to simulate and evaluate various approaches to

vision substitution. Filling this need requires the ability to sense and record the mobility of blind subjects as they negotiate a set of computer generated virtual obstacles while receiving information from some type of vision aid whose parameters are also defined by a computer. A system is required for accurately identifying the subject's position and orientation in real-time over a volume of motion encompasing several stride lengths in any direction. Developing a position sensing system to fill this need is the topic of this thesis.

The study of blind subject mobility is not the only envisioned need for such a system. Human mobility rehabilitation is also involved with investigating deficits in the lower extremities which restrict or alter normal mobility. Issues such as joint mechanics, muscle activation, and control strategies must be investigated not only in the pathological subject but also in the "normal" subject. Clinical mobility studies of pathological subjects seek to quantify specific actions during specific tasks to assist in evaluating the degree of pathology and to document response to corrective surgery or physical therapy. On the other hand, mobility research studies are often concerned with the nature of "normal" motion and investigate the fine details that make normal mobility possible.

To provide data for studying these types of questions a system for providing detailed information pertaining to the kinematic aspects of mobility is required. This need provides a major motivation for this thesis.

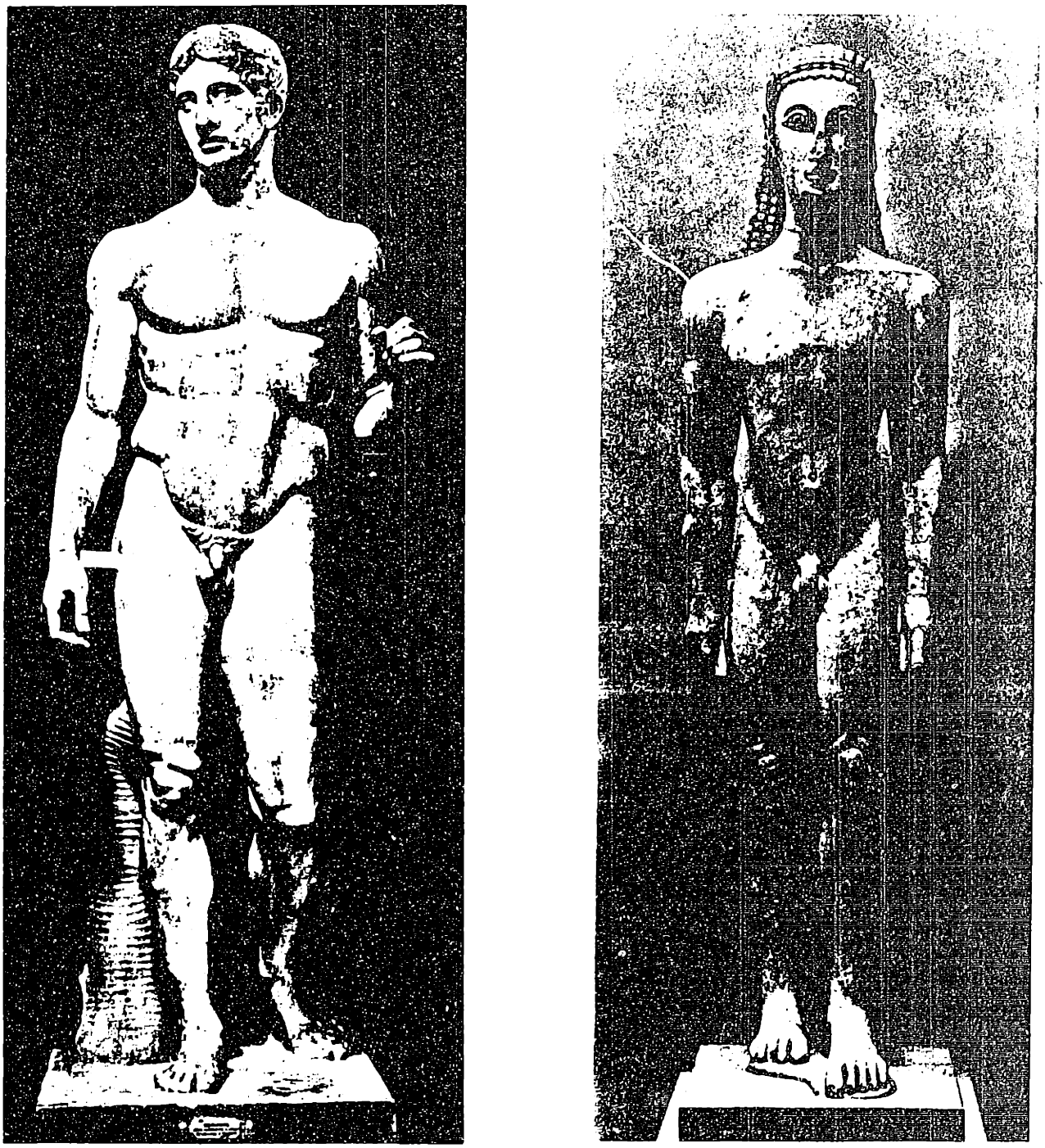
1.2 The Evolving Perception of Human Mobility

The earliest records of an awareness of human mobility are contained in cave paintings and sculptures dating from 15,000 to 10,000 B.C. Whether they were created for reasons of ritual, as documentation of events, or simply as decoration, it is clear the "artist" was required to observe the human "subject" and record the human form with as much detail as observational skill, artistic talent and resources permitted.

Representation of human form underwent major evolution in the period of Greek sculpture from about 600 to 440 B.C. During this 160 year period a new awareness of form and posture that could only have evolved from carefull observation of human movement is noted. The introduction of *contrapposto* (it.) or counterpoise was a fundamental addition to human sculptures [17]. The contrasting styles are shown in Figure 1.1. The *Spear Bearer* appears very natural as he walks forward exemplifying the sculptor's awareness of weight distribution and balance. *Standing Youth* on the other hand appears very rigid and unnatural. One has the impression that if *Standing Youth* were nudged on the shoulder he would tip over whereas *Spear Bearer* would brush it off and continue walking.

Early images of human motion were created solely from observations of the motion while in progress. The introduction of photography forever changed this my making it possible to record images of motion, frozen in time. Muybridge [31] developed techniques for acquiring sequential images of a variety of human activities. A series of Muybribge photographs, such as the high jumper shown in Figure 1.2, allowed for observation of human motion in a way never before possible.

Muybridge's work inevitibly spawned questions of movement which demanded more



(a) (b)

Figure 1.1: (a)*Doryphorus (Spear Bearer).* Roman copy after an original of c.450-440 B.C. by Polyclitus. Marble, height 6'6". National Museum, Naples [from Janson, H.W., "History of Art"]. (b) *Standing Youth (Kouros).* c.600 B.C. Marble, height $6'1\frac{1}{2}"$. The Metropolitan Museum of Art, New York (Fletcher Fund, 1932) [from Janson, H.W., "History of Art"]

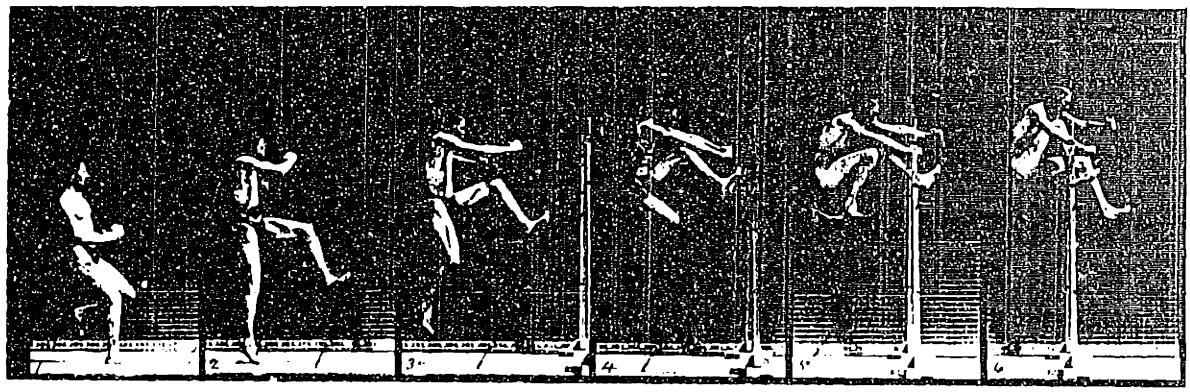


Figure 1.2: Man performing running straight high jump, [from Muybridge, 1955]

detailed data than the qualitative information available in his photographs provided. This generated the need for systems that were capable of measuring various parameters of the observed motion. As techniques evolved so did the questions. The resulting application of measurement systems primarily employing photogrammetric techniques has had a tremendous impact on our ability to investigate the details of human mobility. These developments have been summarized by others [18] [49].

1.3 Global Objective

This thesis is concerned with advancing the capabilities of photogrammetry applied to human mobility analysis a step further. In a typical photogrammetric study of human mobility a pair of cameras are placed so as to have overlapping views. A human subject passing through this viewing volume is detected in some manner by the cameras. The resulting information is processed using photogrammetric techniques to determine three dimensional or rigid body kinematic information relating to the subject's motion.

A limitation with this arrangement is often the viewing volume size and the demand that the subject perform a specific task while passing through it. The measurement volume may be increased by moving the cameras farther from the viewing area or using wide angle lenses but this sacrifices measurement resolution. The global objective of this thesis is to develop a high resolution photogrammetric mesurement system which permits subject mobility within a large measurement volume.

1.4 History of Human Mobility Quantification at MIT

Interest in quantifying human movement in the Newman Laboratory for Biomechanics and Human Rehabilitation at MIT dates back many years. In order to better understand how this thesis relates to previous research efforts in the Newman Laboratory a review of these efforts is presented here.

An important event occured in 1975 with the introduction of a SELSPOT camera system to the laboratory. This marked the beginning of an onging research effort into methods of acquiring human movement data in a rapid and precise manner in a computer controlled environment. Several MIT theses devoted to research and application of this technology have been completed and are reviewed in this section.

SELSPOT cameras, provided by SELSPOT AB of Mölndal, Sweden, are based on infra-red sensing lateral photo-effect diode detectors. These detectors are square, flat semiconductor devices ([23], [45]) placed at the image plane of a camera (ie. in lieu of film in a typical photographic camera). A lens focuses infra-red light from an infra-red light emitting diode (LED) marker onto the detector surface. Analog electronics associated

with the detector identify the *image* coordinates *(u,v)* of the intensity weighted centroid of infra-red light incident upon it. Additional electronics digitize the analog coordinates enabling direct computer interfacing for data acquisition. Detection of multiple LED markers is accomplished by pulsing individual markers in sequence at a total pulse rate of 10 KHz. A *frame* of SELSPOT data consists of one sequential pass through all LED markers used in an experiment.

The original SELSPOT camera consisted of a lens, detector and analog circuitry as described above, with digitizing electronics located in a remote central chassis. Each of two cameras observed the same LED marker pulse and allowed for observation of up to 30 LED markers at a frame rate of 312.5 Hz. Data were digitized to 10 bits or 1 part in 1024 (0.1%) across a 30 degree field of view in both the u and v image coordinates.

Prior to the introduction of SELSPOT to the Newman Lab, the potential capabilities and requirements of a blind mobility aid simulator had been investigated [38]. This identified the need for a system to monitor a blind subject's position and orientation continuously and automatically in real-time throughout a large measurement volume. Other research into the development and eventual implantation of an instrumented hip prosthesis for measuring intra-articular hip pressure *in vivo* required a means for monitoring the relative orientation of the pelvis and femur during normal human activities (e.g. gait). This combined with previous experience in stereophotogrammetry [24] provided motivation for acquiring the SELSPOT cameras as the data acqusition component of a proposed real-time spatial monitoring system.

The first effort toward developing such a system in the Newman Laboratory was

performed by Conati, as documented in his MS thesis [7] in 1977. He initiated development of the *TRACK* system (for Telemetered Real-time Acquisition and Computation of Kinematics). The hardware comprised the original SELSPOT cameras interfaced to a Digital Equipment Corporation (DEC) PDP-11/40 computer through a DEC-kit 11-D DMA Interface. A DEC VT11 interactive graphics display handled data plots with a Gould 5000 printer/plotter for hardcopy. Software, written in FORTRAN and the PDP MACRO-11 assembler, was developed in the Newman Laboratory for acquiring and processing data with the goal being to estimate the time varying position and orientation of rigid objects in space (ie. spatial kinematics).

Conati developed an algorithm for computing the three dimensional (3D) position of LED markers in space based on a two camera configuration. Cameras were arranged level, at the same height, with intersecting fields of view. The angles of camera convergence, camera baseline displacement, non-linearities in the lens, and camera "focal length" were estimated. Using these parameters, the intersection of a ray from each camera, passing through a detected image point and the camera "focal point" identified the 3D point position of an LED marker.

Conati then went on to investigate the problem of estimating object position and orientation in space. A technique investigated by Lenox [19], using an algorithm developed by Schut [35], provided the starting point for this work. In Conati's implemetation, rigid arrays of three or more LED markers were constructed. Each LED marker was located in an embedded cartesian body coordinate system (BCS). Global 3D positions of each marker were determined with the cameras and used as input to the Schut algorithm. Output consisted of a global position vector to the BCS origin and a rotation matrix

representing the transformation from global coordinates to BCS coordinates.

Data from the original SELPSOT cameras contained significant amounts of noise. To combat this, Conati introduced the use of a sixth order, two-pass Butterworth filter to the off-line processing.

Fixed cameras are limited in the spatial region they can observe. For blind mobility studies it was evident that a method was needed for moving the cameras to follow the blind subject throughout a large environment. This would require not only sophisticated apparatus to move the camera but also a means of processing data in real-time. Conati also explored the issues of real-time SELSPOT data processing which he was able to demonstrate at a low sampling frequency. He also proposed and described a system for moving cameras during data acquisition to enlarge the measurement volume.

Tetewsky, in his MS thesis [41], went further with the development of real-time aspects of TRACK and made improvements to the data processing. Conati's algorithm for 3D point estimation contained a singularity when LED markers were level with the cameras. Tetewsky corrected this problem via an alternative geometric formulation. Tetewsky also addressed a singularity present in Schut's algorithm whenever rotations approached 180 degrees and modified the algorithm to produce special case solutions. Internal camera calibration routines were developed further by Tetewsky.

Tetewsky put most of his efforts into developing the real-time capability with TRACK. Highly specialized programs for working in integer arithmetic with relatively small available program space were designed for the PDP-11/40. More detail of the progression of real-time TRACK capabilities may be found in a recent MS thesis by Lord [22].

Concurrent with Tetewsky's work, Antonsson, for his MS thesis, worked on developing a technique for using the kinematic information from TRACK to compute the dynamics of linkages [2]. This marked the beginning of the application of TRACK to general movement studies in biomechanics. In principle, inertial parameter estimates of linkage elements (eg. the foot, shank and thigh of the human leg) combined with accurate measurements of linkage element kinematics, permit estimation of the forces and torques responsible for producing the observed motions.

Antonsson pursued this further in his Doctoral thesis [3] and in the process saw TRACK through a substantial amount of development. Central to this development was a shift in computer processors from the PDP-11/40 to the faster PDP-11/60. Antonsson wrote interface software for SELSPOT communication through a customized DR11-B DMA interface to the PDP's Unibus operating under the RSX11M environment. To pursue dynamics estimation, a Kistler force platform was installed and interfaced for simultaneous data acquisition with SELSPOT. A new version of the TRACK software, having many "user-friendly" features lacking in previous versions was developed by Antonsson for use on the PDP-11/60 and the RSX11M operating system. To differentiate this new hardware/software system from previous versions, Antonsson called it TRACK3.

TRACK3 software was designed to process data off-line. Ottenheimer, in his MS thesis [32], modified TRACK3 to operate in a real-time processing mode. He also made improvements to the algorithms used in TRACK3. One improvement concerned the redundant information present in the 3D point estimates. The fundamental problem of 3D point reconstruction involves finding the intersection of two spatial rays, one from each camera to the point source in space. Camera data noise, error in camera position

and orientation and the inherent discrete nature of the data, combine to prevent these rays from intersecting at any point. Ottenheimer implemented a method for identifying the midpoint of the common perpendicular to these two rays as the desired 3D point estimate. He also realized that the length of the common perpendicular to the two camera rays gives an indication of the amount of error in the reconstruction. The 3D estimation and the "skew ray error" checking technique were incorporated into TRACK3 by Antonsson.

Ottenheimer also addressed the singularity in the Schut algorithm and reformulated the problem in terms of a rotation vector rather than a Rodrigues vector. This development was also incorporated into TRACK3 software by Antonsson.

Research employing TRACK3 was conducted in areas of blind mobility [42] [6], knee prostheses [36] [37], human knee kinematics [29] and the golf swing [25].

In 1982, SELSPOT-II cameras were introduced. Based on the same technology as the original camera, SELSPOT-II cameras featured onboard digitizing circuitry to reduce transmitted signal noise and a digitizing resolution of twelve bits, or 1 part in 4096 (0.025%), across the 30 degree field of view. Additional flexibility was added allowing for selection of up to 16 cameras and 128 LED markers sampled at user selectable frame rates. An LED marker intensity servo loop which adjusted LED intensity to prevent detector saturation at close camera-to-LED ranges was also added. Camera detected intensity levels were resolved to 4 bits and presented as status bits along with camera data. Camera control was handled through a 19 inch rack mountable SELSPOT Administrating Unit but was later replaced with a VMEbus based controller.

Antonsson designed and constructed an interface for this system to the PDP-11/60 for the developing Biomotion Laboratory at Massachusetts General Hospital. An interface of similar design, allowing for use of either the original SELSPOT cameras or the SELSPOT-II cameras each with optional simultaneous Kistler forceplatform data acquisition, was built by the author for the Newman Laboratory installation on the PDP-11/60 computer.

The author set up the SELSPOT-II system in a manner similar to the original SELSPOT system with cameras mounted on an optical bench with assumed camera positions and orientations. System performance was investigated, resulting in electronic adjustments to the SELSPOT-II intensity servo loop to realize maximum marker intensity at most operating ranges. Additionally, it was found that data noise levels could be reduced without sacrificing data reliability by employing markers comprising a cluster of three LEDs rather than the single LED used previously. A modified version of TRACK3, called TRACK4, was written by the author for operating the SELSPOT-II cameras through the new interface to the PDP-11/60 and to handle processing of higher resolution data.

Internal calibration[1] of the SELSPOT-II cameras was approached in a manner different from that used before. Shortly before his departure from the Newman Laboratory, Antonsson discovered an old discarded Gerber X-Y plotter and proposed using it to calibrate SELSPOT cameras. Berger [5] and the author reconditioned the plotter to run via stepper motors from the stepper motor control system associated with an experiment on another PDP-11/60. The SELSPOT interface was moved temporarily to this machine and Berger proceeded to calibrate the cameras as reviewed in detail in Chapter 4.

[1] identification of internal camera geometric parameters and linearisation of the image. See Chapter 4

In order to gain further understanding of the internal calibration problem and to test additional ideas it became apparent that the plotter stepper motors required their own interface to the PDP-11/60 from where SELSPOT was normally operated. The author proceeded to design and build a hardware interface to stepper motor control from an MDB DR11C Unibus I/O card. Additional internal calibration experiments were performed [26] and are described further in Chapter 4.

TRACK4 was fully operational in early 1986. The TRACK4 User's Manual was written and research was performed in areas of ankle kinematics [20], arm kinematics [16], gait kinematics display [21], and ballet [52].

During 1986 several factors contributed to an ultimate decision to move the host computer for SELSPOT from the PDP-11/60 to a SUN Microsystem 3/160. It was clear that a Large Volume TRACK system would require a substantial amount of real-time processing and control. The SUN 3/160, based on the VMEbus architecture, had been used succesfully as a host system for real-time control applications at the MIT Artificial Intelligence Laboratory. Additionally, SELSPOT introduced a VMEbus based interface for the cameras facilitating potential interfacing to the SUN VMEbus. The PDP-11/60 was rapidly becoming old technology lacking sufficient addressable memory for more advanced applications and becoming costly to maintain.

A SUN 3/160 fileserver and workstation along with a SUN 3/50 client workstation were installed by the author in late 1986. In order to facilitate access to the VMEbus during application developement, a 21 slot VMEbus expansion chassis was installed on the SUN 3/160 communicating with the SUN VMEbus through an HVE Repeater 2000

bus repeater.

Moving SELSPOT/TRACK to this system required several steps. Lord ported the TRACK data processing software from FORTRAN into C and organized the overall TRACK package to perform well in the UNIX environment. The author wrote the hardware interface programs for the SELSPOT VMEbus board and introduced use of the Autograph plot package (NCAR) for generating output plots. The Kistler forceplatform and stepper motor interfaces were also ported to the SUN 3/160 by the author. Murphy [30] investigated methods of smoothing noisy kinematic data with the subsequent replacement of the Butterworth filter with a spline smoother in TRACK processing. The new software package, TRACK5, was largely completed by mid 1988.

Concurrent with the above development, the author began investigations into alternative methods for determining external calibration parameters[2] for photogrammetric systems. This work included design and construction of a rotating calibration frame which was driven by stepper motors through the existing stepper motor interface. Details of this work are discussed in Chapter 4. This work resulted in implementation of a straightforward, and automatic external calibration procedure for determining camera positions and orientations whenever the cameras are repositioned for a new experimental setup. In addition, the author identified and implemented a unique algorithmic procedure for 3D point computation (see Chapter 2).

Full scale development of Large Volume TRACK was initiated by the author in late 1988 and is the subject of this thesis. Concurrent work with TRACK was conducted by Lord who explored and developed the real-time processing capabilitites of the VME-

[2]position and orientation of each camera relative to a global frame of reference. See Chapter 4.

bus development system. His work included the implementation of a modified SUN C compiler to generate code for downloading to additional processors on the expansion VMEbus. Lord's MS thesis [22], describes a real-time software system for processing TRACK kinematic data and generating 3D displays on a Silicon Graphics Iris Workstation in real-time.

1.5 Document Organization

The central development of this thesis, Large Volume TRACK, is the end result of years of work which in large part involved developing reliable and accurate methods of calibrating fixed field photogrammetric camera systems. As TRACK evolved, improvements were made to camera calibrations and kinematic estimation algorithms. Knowledge gained from these improvements contributed greatly to the eventual conceptualization and implementation of Large Volume TRACK. This tends toward a thesis document organization which is somewhat chronological. However, a strict chronological presentation would be extremely tedious to read, much less write, due to the iterative nature of scientific investigations. For this reason, discussions of error in the camera system, which where discovered and/or more thoroughly understood during various stages of system development are presented in Chapter 3.

The context for this thesis is developed in Chapter 1 through discussion of the problem in general and a review of research in the Newman Laboratory by the author's predecessors. Chapter 2 introduces photogrammetric concepts and definitions as needed for discussions in subsequent chapters. Also presented are background concepts and algorithmic operations comprising TRACK data processing in general. Both sources and effects

of systematic and stochastic error in the camera system are investigated and discussed in Chapter 3. Minimization of error propogation through accurate camera calibration is central to realizing a high degree of precision and accuracy with TRACK. Developments in internal and external camera calibration are presented in Chapter 4. The conceptualization of the Large Volume TRACK system including the algorithmic developments for 3D point estimation and system calibration is presented in Chapter 5. Finally, in Chapter 6, Large Volume TRACK implementation, testing, results and conclusions are presented along with suggestions for further work and improvements.

Chapter 2

Analytic Close-Range Stereophotogrammetry

Photogrammetry is the process of making measurements of three dimensional (3D) objects based on two dimensional (2D) images of the object. When two or more 2D images yield measurements of a 3D object and are obtained from different spatial perspectives, it is termed *stereophotogrammetry*. The term *close-range* photogrammetry distinguishes the problem of making measurements relatively near the imaging device from the related problem of measurements made at long range (eg. aerial photogrammetry). Close-range stereophotogrammetric systems are most often characterised by two or more imaging devices placed at fixed locations in a reference coordinate system with the imaging devices' axes forming mutually oblique angles. The development of *analytic* mathematical models representing the transformation of information from 1D or 2D detectors to estimates of corresponding 3D object locations is fundamental to computational photogrammetric techniques required for use with digital computers.

This chapter provides a background to analytic close-range stereophotogrammetry, relevant to specific applications and developments presented in subsequent chapters of

this thesis. Certain aspects of this background introduction (internal and external calibration in particular) are saved for chapters dealing with those particular issues in detail. A thorough introduction to most topics discussed may be found in Moffit and Mikhail [27] and Ghosh [11].

Also presented here are the specific algorithms contained in the latest version of TRACK. These algorithms reflect an accumulation of knowledge obtained through the development of TRACK by the author and his predecessors (see Chapter 1).

2.1 The Photogrammetric Camera Model

A typical camera is comprised of hundreds of individual parts. For photogrammetric purposes of camera modeling only two of these parts, the lens and detector, are relevant. While the camera body, which holds the lens and detector in a fixed relative position, is important to the operation of the camera, the assumption of its existence is sufficient for the photogrammetric camera model.

2.1.1 The Lens

The lens gathers and focuses light. Thin lenses which are idealized as having no thickness are discussed in the optics section of most introductory physics texts and therefore will not be discussed here. Thick lenses, the more common case in cameras, are described here and a thorough discussion on all lens types may be found in Chapter 2 of Moffit and Mikhail [27].

As with thin lenses, thick lenses are characterized by front and back foci (F and

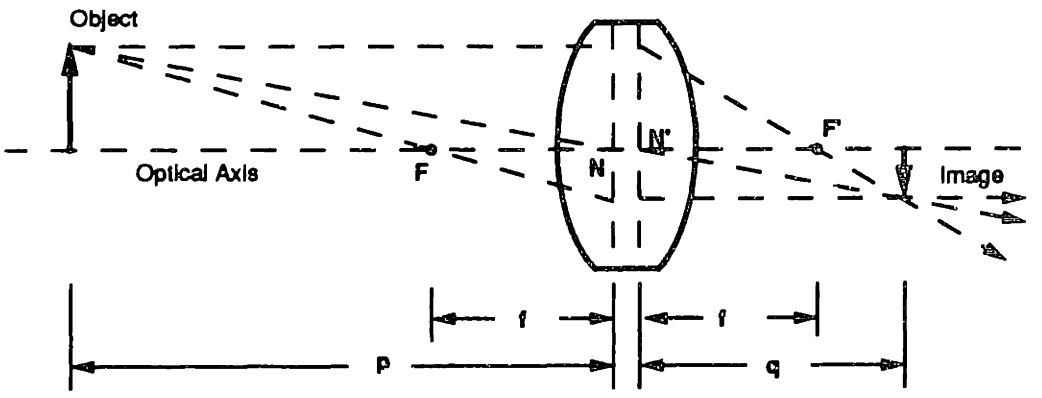


Figure 2.1: Basic thick lens (from Moffit and Mikhail pg. 34)

F') located on the lens *optical axis* at a *focal length* (f) from the effective plane of light refraction. In thick lenses, however, the plane of light refraction, or *principal plane* must be modeled as two parallel planes intersecting the optical axis at *nodal points* N and N' (see Figure 2.1). Ray tracing proceeds as with thin lenses with the exception that a ray entering the front nodal plane is refracted through the rear nodal plane.

Light from objects located a distance p ($> f$) from the front lens node N will be focused at a distance q ($> f$) from the rear lens node N'. The values of f, p and q obey the relationship

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f} \tag{2.1}$$

and the lens magnification (M), may be deduced using the principle of similar triangles from Figure 2.1 as

$$M = \frac{image\ size}{object\ size} = \frac{q}{p} \tag{2.2}$$

In the typical camera, focusing a lens involves displacing the lens along the optical axis relative to the detector surface until the image and detector are coincident. Note that for an object located at $p = \infty$ focusing the lens will put the focal point F' at the

detector surface (ie. $f = q$).

2.1.2 The Detector

The detector registers the image projected by the lens. Photographic film, charge coupled devices, image orthicon tubes, and lateral photoeffect diodes are common types of detectors. Detectors are usually planar and are mounted perpendicular to the lens optical axis at the focused image distance q from the lens rear node.

2.1.3 Idealized Camera Geometry

In order to facilitate study of the geometric relationship of objects to images, simplifications must be made to the geometry (see Figure 2.2). A bundle of rays from an object point strike the lens and are focused to a point in the image. For geometric purposes, it is sufficient to consider only the ray which passes through the lens nodes, called the *chief ray*. The additional rays add light intensity to the image point but add no additional information pertaining to image point location.

The lens nodes are collapsed to a single point called the *projection center*, or center of projection (c). The projection center serves as the origin of the camera coordinate system $\mathbf{x}_c$.

The detector is idealized as a planar surface, called the *image plane*. Image coordinates $\mathbf{u} = (u, v)^T$ identify the position of an image point on the image plane. The camera z-coordinate axis (also called the *principal axis*) is defined along the perpendicular to the image plane passing through c. Note a subtle difference in terminology in that the optical axis is a property of the lens alone while the principal axis is related to the relative

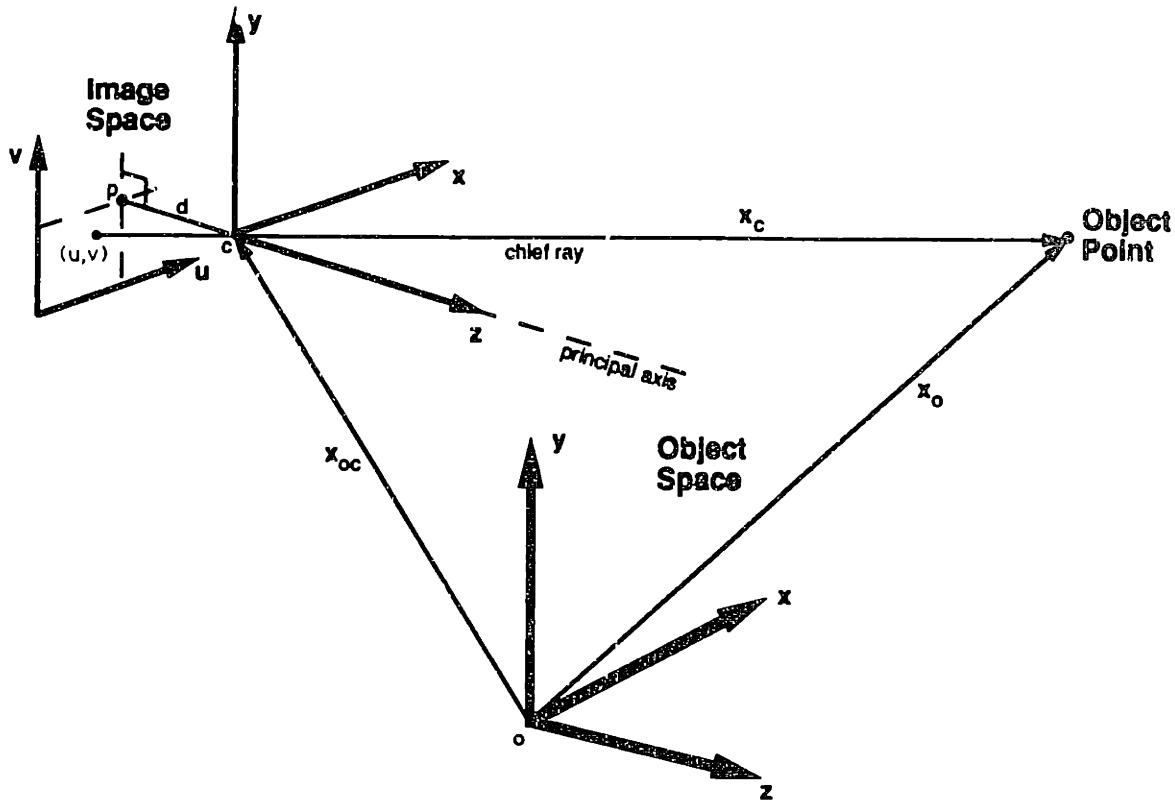


Figure 2.2: Photogrammetric camera model

mounting of the lens and detector. The intersection of the principal axis with the image plane at $\mathbf{u_p} = (u_p, v_p)^T$ is termed the *principal point* (p) and the displacement between p and c is termed the *principal distance* (d).

Relating the idealized detector and lens geometry to real data requires introduction of non-linear factors accounting for image distortions generated by the lens, detector and signal amplifying electronics. These functions are represented as corrections $(\Delta u, \Delta v)$ where both Δu and Δv are, in general, non-linear functions of the detected image coordinate (u, v).

To simplify notation, corrected image coordinates $(\bar{u}, \bar{v})$ are introduced where

$$\bar{u} = u - u_p + \Delta u$$

$$\bar{v} = v - v_p + \Delta v$$

Camera coordinates x and y are parallel to image coordinates u and v respectively. The area between the projection center and the image plane is termed *image space* while the area from the projection center extending to the object is termed *object space.*

A reference or *global* cartesian coordinate system $\mathbf{x}_o$ with origin o is defined for relating the position of points in object space to points in image space. A vector defining the position of the camera projection center as seen from the global coordinate system is defined as $\mathbf{x}_{oc} = (x_{oc}, y_{oc}, z_{oc})^T$. A 3 by 3 orthogonal rotation matrix M having components

$$\mathrm{M} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \tag{2.3}$$

defines the orientation of the camera coordinate system to the global coordinate system. The transformation between camera coordinates and global coordinates may be written as

$$\mathbf{x}_c = \mathrm{M}[\mathbf{x}_o - \mathbf{x}_{oc}] \tag{2.4}$$

2.1.4 The Colinearity Equations

With the definitions given above it is desired to derive an algebraic formula relating an object point to its image in a particular camera. This is approached by applying the *colinearity condition*, which states that a ray connecting an object point to its respective image point must pass through the projection center. Referring to Figure 2.2 and applying the principle of similar triangles the following equations relating a camera defined object point $\mathbf{x}_c = (x_c, y_c, z_c)^T$ to a corrected image point $\bar{\mathbf{u}} = (\bar{u}, \bar{v})^T$ may be written as

$$\frac{\bar{u}}{d} = \frac{x_c}{z_c} \quad ; \quad \frac{\bar{v}}{d} = \frac{y_c}{z_c} \tag{2.5}$$

Combining equations 2.5 and 2.4 yields

$$\begin{bmatrix} \bar{u} \\ \bar{v} \\ d \end{bmatrix} = \frac{d}{z_c}\mathbf{M}[\mathbf{x}_o - \mathbf{x}_{oc}] \tag{2.6}$$

If M, $\mathbf{x}_o$ and $\mathbf{x}_{oc}$ are written in terms of their components the above equation becomes

$$\begin{bmatrix} \bar{u} \\ \bar{v} \\ d \end{bmatrix} = \frac{d}{z_c}\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}\begin{bmatrix} x_o - x_{oc} \\ y_o - y_{oc} \\ z_o - z_{oc} \end{bmatrix} \tag{2.7}$$

Expanding this to remove z_c yields a form of the photogrammetric *colinearity* equations,

$$u' = \frac{\bar{u}}{d} = \frac{m_{11}(x_o - x_{oc}) + m_{12}(y_o - y_{oc}) + m_{13}(z_o - z_{oc})}{m_{31}(x_o - x_{oc}) + m_{32}(y_o - y_{oc}) + m_{33}(z_o - z_{oc})} \tag{2.8}$$

and

$$v' = \frac{\bar{v}}{d} = \frac{m_{21}(x_o - x_{oc}) + m_{22}(y_o - y_{oc}) + m_{23}(z_o - z_{oc})}{m_{31}(x_o - x_{oc}) + m_{32}(y_o - y_{oc}) + m_{33}(z_o - z_{oc})} \tag{2.9}$$

that allow for the direct computation of an image point u from any globally defined object point $\mathbf{x}_o$ provided the parameters M, $\mathbf{x}_{oc}$, $\mathbf{u}_p$, d and functions Δu and Δv are known. Further discussion in this chapter assumes these parameters are known. Methods for determining these parameters are discussed in Chapter 4.

2.2 Estimating 3D Object Points from 2D Image Points

While the colinearity equations 2.8 and 2.9 do map a 3D object point to a unique 2D image point, the reverse is not true. By rearranging the colinearity equations as

$$(m_{11} - m_{31}u')x_o + (m_{12} - m_{32}u')y_o + (m_{13} - m_{33}u')z_o = b_u \tag{2.10}$$

$$(m_{21} - m_{31}v')x_o + (m_{22} - m_{32}v')y_o + (m_{23} - m_{33}v')z_o = b_v \tag{2.11}$$

where

$$b_u = (m_{11} - m_{31}u')x_{oc} + (m_{12} - m_{32}u')y_{oc} + (m_{13} - m_{33}u')z_{oc} \tag{2.12}$$

$$b_v = (m_{21} - m_{31}v')x_{oc} + (m_{22} - m_{32}v')y_{oc} + (m_{23} - m_{33}v')z_{oc} \tag{2.13}$$

it is observed that each equation represents the general equation for a plane in global space. The intersection of these two planes defines the ray of colinearity between the 3D object point, projection center, and 2D image point. The colinearity equations therefore, map 2D image points to a ray in space. The desired 3D object point location lies somewhere along that ray.

The key to identifying 3D object points relies on multiple 2D images made from various perspectives. If the object point is fixed relative to the global frame then multiple 2D images may be made with the same imaging device at different times provided estimates of the transformation from global to image space may be made for each image used. In the case of non-stationary object points, multiple imaging devices must generate simultaneous images. The discussion here will assume the latter situation and it will be shown that two imaging devices, in this case optoelectronic cameras, are sufficient to estimate a 3D object point location.

There are two approaches to this problem. The first approach uses the colinearity equations generated from each camera and employs a least-squares solution technique to find the desired object space point. This method is readily expandable to n cameras where $n \geq 2$. The second approach, which the author has not seen published before, can be shown to be algebraically identical to the first method, but solves for the 3D point in two steps and in the process identifies a parameter useful in checking the validity of the

data. This method is most useful when only two cameras are used and is the method currently implemented in the TRACK software.

2.2.1 Colinearity Equation Approach

The colinearity equation approach to 3D object point estimation is basic to analytical photogrammetric techniques. It will be compared with another technique, the Direct Linear Transform, in Chapter 4.

Start by rewriting the colinearity equations in general plane form as in 2.10 - 2.13 for a specific camera, i, as

$$(m_{11_i} - m_{31_i}u_i')x_o + (m_{12_i} - m_{32_i}u_i')y_o + (m_{13_i} - m_{33_i}u_i')z_o = b_{u_i} \tag{2.14}$$

and

$$(m_{21_i} - m_{31_i}v_i')x_o + (m_{22_i} - m_{32_i}v_i')y_o + (m_{23_i} - m_{33_i}v_i')z_o = b_{v_i} \tag{2.15}$$

where

$$u_i' = \frac{\bar{u}_i}{d_i} \tag{2.16}$$

$$v_i' = \frac{\bar{v}_i}{d_i} \tag{2.17}$$

$$b_{u_i} = (m_{11_i} - m_{31_i}u_i')x_{oc_i} + (m_{12_i} - m_{32_i}u_i')y_{oc_i} + (m_{13_i} - m_{33_i}u_i')z_{cc_i} \tag{2.18}$$

$$b_{v_i} = (m_{21_i} - m_{31_i}v_i')x_{oc_i} + (m_{22_i} - m_{32_i}v_i')y_{oc_i} + (m_{23_i} - m_{33_i}v_i')z_{oc_i} \tag{2.19}$$

If n cameras ($n \geq 2$) each view the same object space point, $\mathbf{x}_o$, then the matrix equation

$$\mathbf{A}\mathbf{x}_o = \mathbf{b} \tag{2.20}$$

may be written, where

$$\mathbf{A} = \begin{bmatrix} m_{11_1} - m_{31_1}u'_1 & m_{12_1} - m_{32_1}u'_1 & m_{13_1} - m_{33_1}u'_1 \\ m_{21_1} - m_{31_1}v'_1 & m_{22_1} - m_{32_1}v'_1 & m_{23_1} - m_{33_1}v'_1 \\ m_{11_2} - m_{31_2}u'_2 & m_{12_2} - m_{32_2}u'_2 & m_{13_2} - m_{33_2}u'_2 \\ m_{21_2} - m_{31_2}v'_2 & m_{22_2} - m_{32_2}v'_2 & m_{23_2} - m_{33_2}v'_2 \\ \vdots & \vdots & \vdots \\ m_{11_n} - m_{31_n}u'_n & m_{12_n} - m_{32_n}u'_n & m_{13_n} - m_{33_n}u'_n \\ m_{21_n} - m_{31_n}v'_n & m_{22_n} - m_{32_n}v'_n & m_{23_n} - m_{33_n}v'_n \end{bmatrix}_{2n\times 3} \tag{2.21}$$

and

$$\mathbf{b} = \begin{bmatrix} b_{u_1} & b_{v_1} & b_{u_2} & b_{v_2} & \cdots & b_{u_n} & b_{v_n} \end{bmatrix}^T \tag{2.22}$$

These $2n$ equations in 3 unknowns may be solved using the following standard linear least squares approach. Premultiplying both sides of 2.20 by the transpose of A yields the *normal equations* [39]

$$\mathbf{A}^T\mathbf{A}\mathbf{x}_o = \mathbf{A}^T\mathbf{b} \tag{2.23}$$

The normal matrix, $\mathbf{A}^T\mathbf{A}$, is square (3×3), and nonsingular. Multiplying both sides of 2.23 by the inverse normal matrix yields

$$\mathbf{x}_o = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \tag{2.24}$$

which is the least-squares solution for the object point position in global coordinates.

Physical interpretation of the solution involves visualizing pairs of equations each defining a ray directed from respective cameras nominally through the object point being estimated. Errors in camera data and parameters (to be discussed in further chapters) prevent these rays from actually intersecting. The solution yields the point which most closely represents the intersection of the rays in a least squares sense.

Note that a minimum of two cameras are required to determine a unique solution, giving four equations in three unknowns. In the case of two cameras the resulting 3D object point can be shown to be the mean of the common perpendicular to the two camera rays.

Also note that if two cameras are used but only a total of three image coordinates are applied (eg. u_1, v_1, u_2) then three equations in three unknowns can be formed. The resulting solution is the intersection of three planes. This solution technique is used by 3D imaging systems consisting of three 1D linear detectors each generating data which resolves to a plane that passes through the object space point analagous to the ray of colinearity in the 2D detector case.

The overdetermined equations are useful in many situations particularly where relatively large measurement uncertainties are present. Variance in the 3D estimate, as propagated from the image data, is reduced when levels of redundancy are added to the solution. Overdetermined equations also allow for checking data for unusually high levels of error indicating a possible systematic problem.

2.2.2 Two Camera Approach with Error Checking

An alternative formulation can be shown to be algebraically identical to the above approach when applied to two cameras only. The 3D point is solved for in two steps and in the process a parameter relevant to the quality of the reconstruction is computed. To the author's knowledge, this formulation is new and is now unique to TRACK software processing.

The formulation starts with the image to global transformation, 2.6, rewritten here

using the notation for a particular camera:

$$\begin{bmatrix} u_i' \\ v_i' \\ 1 \end{bmatrix} = \frac{1}{z_{c_i}} \mathbf{M}_i [\mathbf{x}_o - \mathbf{x}_{oc_i}] \tag{2.25}$$

This may be rewritten to bring $\mathbf{x}_o$ to the left side of the equation as

$$\mathbf{x}_o = z_{c_i} \mathbf{M}_i^{-1} \begin{bmatrix} u_i' \\ v_i' \\ 1 \end{bmatrix} + \mathbf{x}_{oc_i} \tag{2.26}$$

Equation 2.26 solves for a unique $\mathbf{x}_o$ in terms of the data coordinates $\mathbf{u}_i$ for camera i and the z-camera coordinate for the object point, z_{c_i}. If two cameras view the same object point, 2.26 may be written for each camera ($i = 1, 2$) and combined to form

$$\left(\mathbf{M}_1^{-1} \begin{bmatrix} u_1' \\ v_1' \\ 1 \end{bmatrix} \quad \Big| \quad -\mathbf{M}_2^{-1} \begin{bmatrix} u_2' \\ v_2' \\ 1 \end{bmatrix} \right)_{(3\times 2)} \begin{bmatrix} z_{c_1} \\ z_{c_2} \end{bmatrix} = \mathbf{x}_{oc_2} - \mathbf{x}_{oc_1} \tag{2.27}$$

The three linear equations in two unknowns, z_{c_1} and z_{c_2}, may be solved easily using the linear least squares approach used in Section 2.2.1.

Once z_{c_1} and z_{c_2} have been estimated the globally defined points $\mathbf{x}_{o_1}$ and $\mathbf{x}_{o_2}$ may be determined from 2.26. Points $\mathbf{x}_{o_1}$ and $\mathbf{x}_{o_2}$ will never be identical as they each lie on different camera rays that approach each other but never quite intersect due to errors discussed in Chapter 3. It can be shown that $\mathbf{x}_{o_1}$ and $\mathbf{x}_{o_2}$ represent the respective points of intersection of each ray with their mutual perpendicular. The spatial mean of $\mathbf{x}_{o_1}$ and $\mathbf{x}_{o_2}$ represents an identical solution to that found for $\mathbf{x}_o$ in Section 2.2.1.

Skew in the rays is defined as the distance between $\mathbf{x}_{o_1}$ and $\mathbf{x}_{o_2}$ and in TRACK is termed *skew ray error*. Large skew ray error indicates problems with system parameter identification during calibation or, for active marker optoelectronic camera systems, the presence of marker reflections which corrupt the data. In TRACK processing, skew error

is normalized to camera units based on the mean marker displacement from the cameras. This permits comparison of skew levels near the cameras with skew levels more distant from the cameras.

2.3 Estimating Position and Orientation of Rigid Bodies

In many photogrammetric applications it is not sufficient to obtain only 3D position estimates of points in space. Often one must obtain an estimate for spatial position and orientation of objects. A convenient means of representing this information is through the identification of a 3D position vector $\mathbf{r}$ and orthogonal rotation matrix $\mathbf{R}$ [8].

Two methods of approaching this problem will be introduced here. Each method requires previously determined estimates of the 3D locations of at least three points on an object relative to a global coordinate system at both a reference (known) and arbitrary (to be determined) position and orientation. The details of these methods are not required for development of this thesis and the reader is directed to the cited references for additional information. The general approach of each method is discussed so that the reader has an appreciation for the overall problem and can compare these methods with a new approach, discussed in Appendix B, which does not require 3D point estimation at arbitrary object orientation. Definitions required for both methods are presented first and then each method is outlined.

Assume a system of n points ($n \geq 3$), located on a rigid object in a reference position and orientation, where each point P_i is defined by known global vector $\mathbf{a}_i$ (see Figure 2.3). The mean of $\mathbf{a}_i$ ($i = 1, 2, ...n$) vectors is defined as $\mathbf{a}$ such that $\mathbf{a} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{a}_i$. Next, define

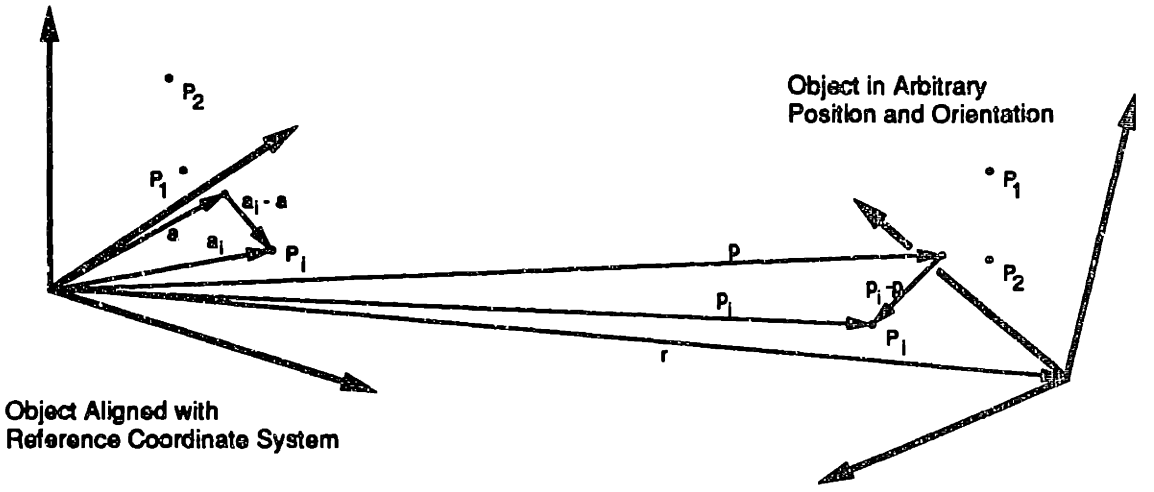


Figure 2.3: Position and orientation of objects in space

global vector $\mathbf{p}_i$ as the position of point P_i on the same object in an arbitrary (to be determined) position and orientation. The 3D vectors $\mathbf{p}_i$ have been estimated using the techniques described in previous sections of this chapter. The mean of $\mathbf{p}_i$ $(i = 1, 2, ...n)$ vectors is defined as $\mathbf{p}$ such that $\mathbf{p} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{p}_i$.

The specific problem is to find a rotation matrix $\mathbf{R}$ and position vector $\mathbf{r}$ such that any $\mathbf{p}_i$ is related to $\mathbf{a}_i$ as follows

$$\mathbf{p}_i = \mathbf{R}\mathbf{a}_i + \mathbf{r} \tag{2.28}$$

Equation 2.28 holds equally for the mean points as

$$\mathbf{p} = \mathbf{R}\mathbf{a} + \mathbf{r} \tag{2.29}$$

If 2.29 is subtracted from 2.28 then $\mathbf{r}$ is eliminated and the problem is reduced to determining the rotation matrix $\mathbf{R}$ in

$$(\mathbf{p}_i - \mathbf{p}) = \mathbf{R}(\mathbf{a}_i - \mathbf{a}) \tag{2.30}$$

Once $\mathbf{R}$ has been determined, $\mathbf{r}$ may be found from rewriting 2.29 as

$$\mathbf{r} = \mathbf{p} - \mathbf{R}\mathbf{a} \tag{2.31}$$

One method for determining $\mathbf{R}$ was developed by Veldpaus et.al. [44]. In their approach equation 2.30 is rearranged to form

$$\hat{\mathbf{p}}_i = \mathbf{a} + \hat{\mathbf{r}} + \mathbf{R}(\mathbf{a}_i - \mathbf{a}) \tag{2.32}$$

where vectors $\hat{\mathbf{p}}_i$ represent the measured (inexact) $\mathbf{p}_i$ vectors and $\hat{\mathbf{r}}$ is defined as

$$\hat{\mathbf{r}} = \hat{\mathbf{p}} - \mathbf{a} \tag{2.33}$$

They then seek a least-squares minimum of the function

$$f(\hat{\mathbf{r}}, \hat{\mathbf{H}}) = \frac{1}{n}\sum_{i=1}^{n}\left[(\hat{\mathbf{p}}_i - \mathbf{a} - \hat{\mathbf{r}} - \hat{\mathbf{H}}(\mathbf{a}_i - \mathbf{a}))^{\mathbf{T}}(\hat{\mathbf{p}}_i - \mathbf{a} - \hat{\mathbf{r}} - \hat{\mathbf{H}}(\mathbf{a}_i - \mathbf{a}))\right] \tag{2.34}$$

where $\hat{\mathbf{H}}$ is composed of a scalar "error" s and an estimate of the rotation matrix $\hat{\mathbf{R}}$ such that $\hat{\mathbf{H}} = s\hat{\mathbf{R}}$.

Using variational analysis Veldpaus et.al. show that if $rank(\hat{\mathbf{G}}) \geq 2$ then

$$\hat{\mathbf{G}} = \hat{\mathbf{R}}\hat{\mathbf{B}} \tag{2.35}$$

where

$$\hat{\mathbf{G}} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mathbf{p}}_i - \hat{\mathbf{p}})(\mathbf{a}_i - \mathbf{a})^{\mathbf{T}} \tag{2.36}$$

$$\hat{\mathbf{B}} = \hat{s}\mathbf{A} - \frac{1}{2\hat{s}}(\mathbf{L} + \mathbf{L}^T) \tag{2.37}$$

$$\mathbf{A} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{a}_i - \mathbf{a})(\mathbf{a}_i - \mathbf{a})^{\mathbf{T}} \tag{2.38}$$

and $\mathbf{L}$ is a matrix of Lagrange multipliers. Veldpaus et.al. demonstrate polar decomposition of $\hat{\mathbf{G}}$ yielding the symmetric matrix $\hat{\mathbf{B}}$ and orthogonal rotation matrix $\hat{\mathbf{R}}$.

An alternative to the above method is used by TRACK. Thompson [43] and Schut [34] initiated investigation into this approach which saw further development by Lenox [19], Conati [7], Ottenheimer [32] and Antonsson [4].

In this method the rotation matrix is solved via the estimation of four parameters of a quaternion. These parameters may be thought of as the three components of a unit vector pointing along the axis of rotation and a fourth parameter equal to the cosine of one half of the rotation angle about that axis. Schut represented these parameters as (a, b, c, d) and formed the following representation of the rotation matrix $\mathbb{R}$:

$$\mathbf{R} = (d\mathbf{I} - \mathbf{S})^{-1}(d\mathbf{I} + \mathbf{S}) \tag{2.39}$$

where the matrix $\mathbf{S}$ is defined as

$$\mathbf{S} = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \tag{2.40}$$

Schut demonstrated the use of 2.39 in 2.30 to form

$$(d\mathbf{I} - \mathbf{S})(\mathbf{p}_i - \mathbf{p}) = (d\mathbf{I} + \mathbf{S})(\mathbf{a}_i - \mathbf{a}) \tag{2.41}$$

When 2.41 is expanded and terms of (a, b, c, d) are collected, three linear equations in the four unknowns are formed for each point defined in both the reference and arbitrary coordinate systems. Three or more non-colinear points defined in the reference and arbitrary coordinate systems will form a unique nullspace solution for (a, b, c, d). In principle, fixing any of the four parameters allows for solution of the remaining three via linear least-squares. In practice, the potential for singularities in the solution is avoided through judicious selection of the fixed parameter [4].

Chapter 3

Photogrammetric Error with SELSPOT Cameras

Error in photogrammetric measurements may be traced to numerous sources. The ideal camera model presented in the previous chapter is only an abstraction of a real camera. Distortion caused by the lens and film or detector are but a few of the effects distinguishing a real camera from the ideal. This chapter discusses error in close-range photogrammetric measurements, specifically for applications involving SELSPOT[1] cameras. It begins with a review of measurement terminology.

3.1 Measurement Terminology

The goal of any measurement is to quantify a specific physical property and, with non-stationary systems, to do so at a specific instant in time. However, no measurement is exact. The difference between a measured property and its exact value is measurement *error*. Since quantifying a physical property always involves some form of measurement

[1]Cameras used for collecting data in this Chapter were MIT SELSPOT cameras, serial numbers 104 (camera 1) and 105 (camera 2). All other data in this thesis was collected using MGH SELSPOT cameras serial numbers 148 (camera 1) and 178 (camera 2). See conclusions in Chapter 6 for more details.

it follows that the exact value is never known and hence the error is never known. The most that can be done is to estimate the error. The error estimate is called measurement *uncertainty*. Measurement uncertainty indicates the boundaries or *tolerance* within which the physical property presumably must lie.

Measuring instruments are specified according to their resolution, accuracy and precision. *Resolution* refers to the finest increment that may be observed in the measurement. With computer based instruments this is often defined by the number of digital bits used to approximate an analog signal representing the property being measured. Instrument *accuracy* refers to the difference between an instrument reading and a "known" input [15]. Accuracy is usually expressed as a precentage of the full measurement range of the instrument. If a measurement having a particular level of accuracy is made repeatedly, the resulting cluster of measurement values indicates the *precision* or *repeatability* of the measurement. *Calibration* of the instrument involves making adjustments (if possible) to bring the mean value of the cluster of measurements in line with the the known input. In this manner a measurement's accuracy may be improved up to the limit imposed by the measurement precision.

The following discussion will address two basic classifications of error: stochastic and systematic. Stochastic error may be attributed to random influences on the measurement and influences instrument precision. Such errors are typically non-repeatable and cannot be predicted except in a statistical sense. In contrast, systematic error is repeatable for a particular measurement configuration. This characteristic repeatability allows for instrument calibration to play a role in reducing systematic error.

3.2 Stochastic Error in SELSPOT Camera Data

The physical sources of stochastic error in SELSPOT cameras are numerous but from the standpoint of analysis it suffices to treat the camera as a "black box" and compare inputs with outputs. From this evaluation approach two distinct types of stochastic error have been identified.

The first type of stochastic error consists of a broad band additive noise in the data. It is considered to have a zero mean distribution across the signal. Consequently, averaging many data samples under static input conditions allows for better approximation of the real signal.

Noise in the data has been found to be inversely correlated with the detected intensity of the active LED marker. Intensity varies with the inverse square of the camera to marker displacement, and also varies with the relative camera to marker angular displacement. Different types of markers are available and may be classified according to intensity output but even within one marker type a range of intensities may be observed.

An experiment was performed for the purpose of relating the intensity values detected at the camera to the noise level in the data. Cameras were mounted horizontally on tripods side-by-side a height of 1.7 meters from the floor. Two types of markers [2], having different intensities were mounted at camera level on a positioning mechanism, described later in this chapter, which was in turn mounted on a wheeled cart. A plumb bob extended to the floor beneath the markers to identify the marker location in the horizontal plane. The positioning mechanism allowed the markers to be angularly positioned about

[2]One marker was a single LED, the other a triad of LEDs where each LED was a Litronix LD242-3 infra-red emitter.

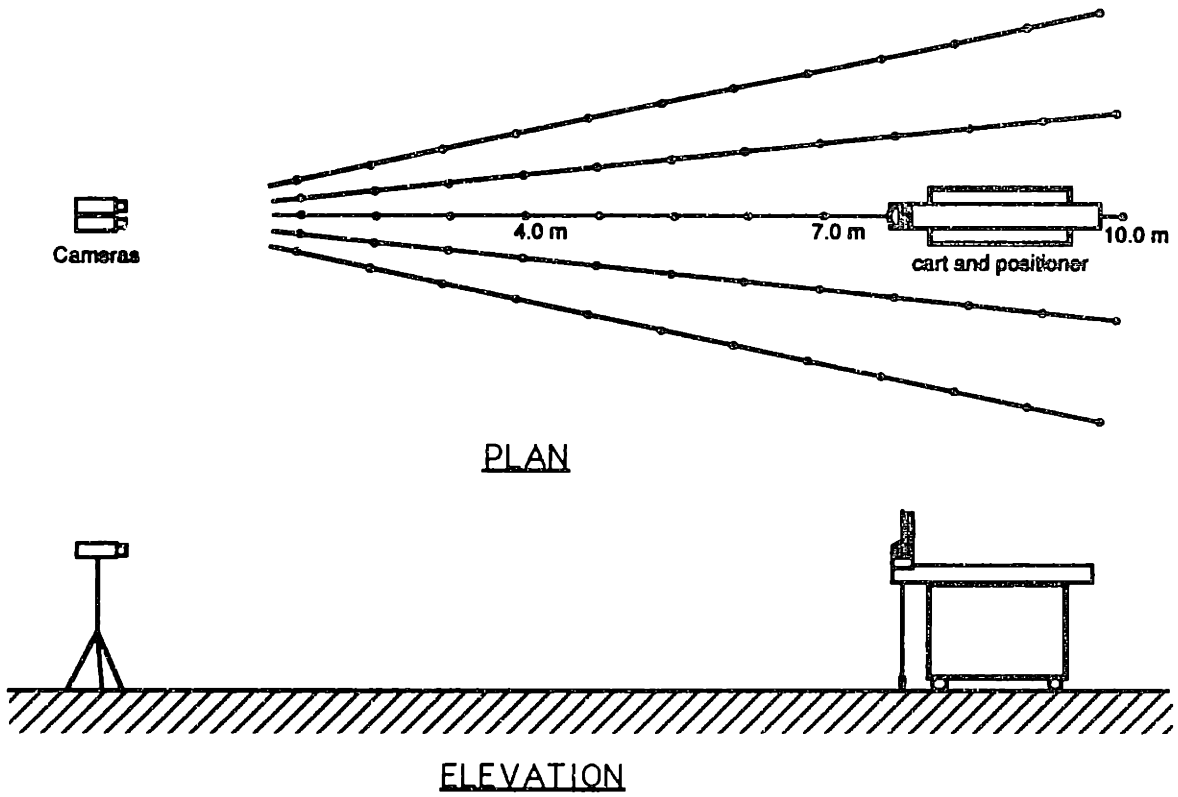


Figure 3.1: Intensity versus noise test layout

a vertical axis. Five radial lines emanating from the cameras in the horizontal plane were drawn on the floor. One line was drawn parallel to the cameras' principal axes, while the other four were drawn at angular intervals of ± 6 and ± 12 degrees (see Figure 3.1). Radial increments of 0.75 meters starting at 1.75 meters from the camera and ending at 10 meters were marked along each radial line. At each of the sixty defined positions, data was collected from each marker at each of six rotation angles equally spaced from zero to 75 degrees away from the radial line. The measure of noise level used was the standard deviation in the raw camera data over one thousand samples. The raw camera data consists of the digitized position (to 12 bits of resolution) of the marker image on the image plane. Each discrete increment of raw data is termed a SELSPOT unit (SU).

A plot relating mean standard deviation in all data at specific detected intensity levels is shown in Figure 3.2. One interesting aspect of the data is the non-uniformity

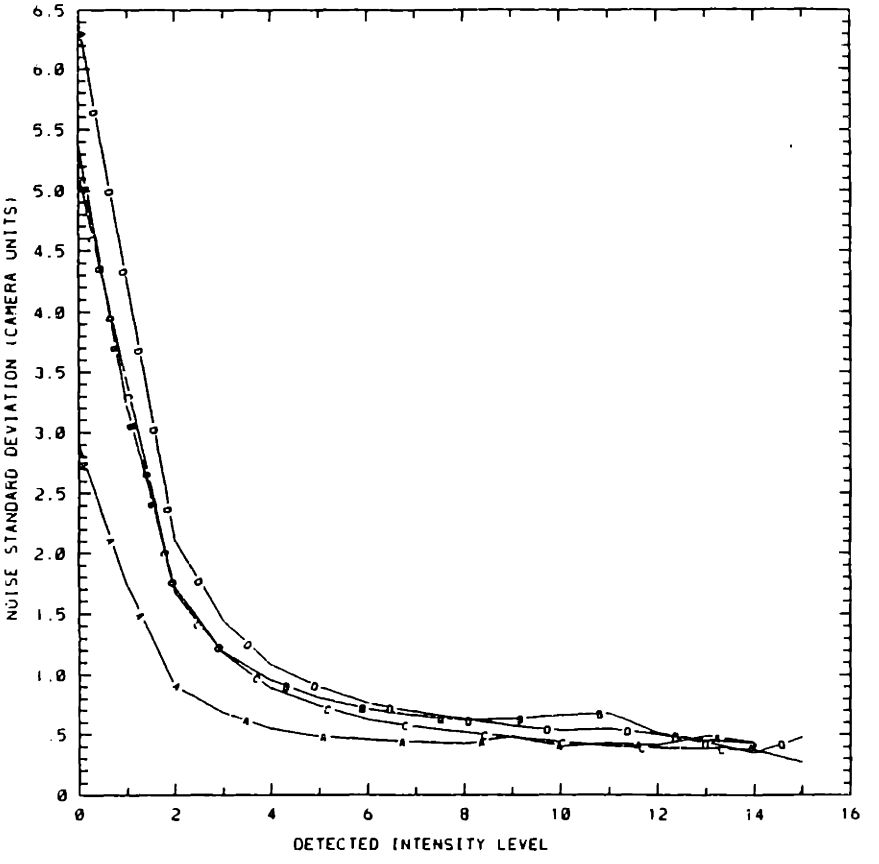


Figure 3.2: Mean standard deviation of SELSPOT data as a function of detected intensity. (A) Camera 1, u coord; (B) Camera 1, v coord; (C) Camera 2, u coord; (D) Camera 2, v coord

of intensity/noise across different camera coordinates. From the Figure 3.2 it is clear that the v coordinate of camera 2 contains a higher level of noise at the lower intensity levels than any of the other three coordinates. This information could be used to assign weighting factors to the four inputs to the colinearity condition used to estimate the 3D output coordinates for each marker.

A plot relating marker intensity to camera displacement at zero marker rotation from the radial line is shown in Figure 3.3.

Physical intuition for the effects of noise on reconstructed 3D data may be gained by computing the resultant magnitude of the error across the camera field at various

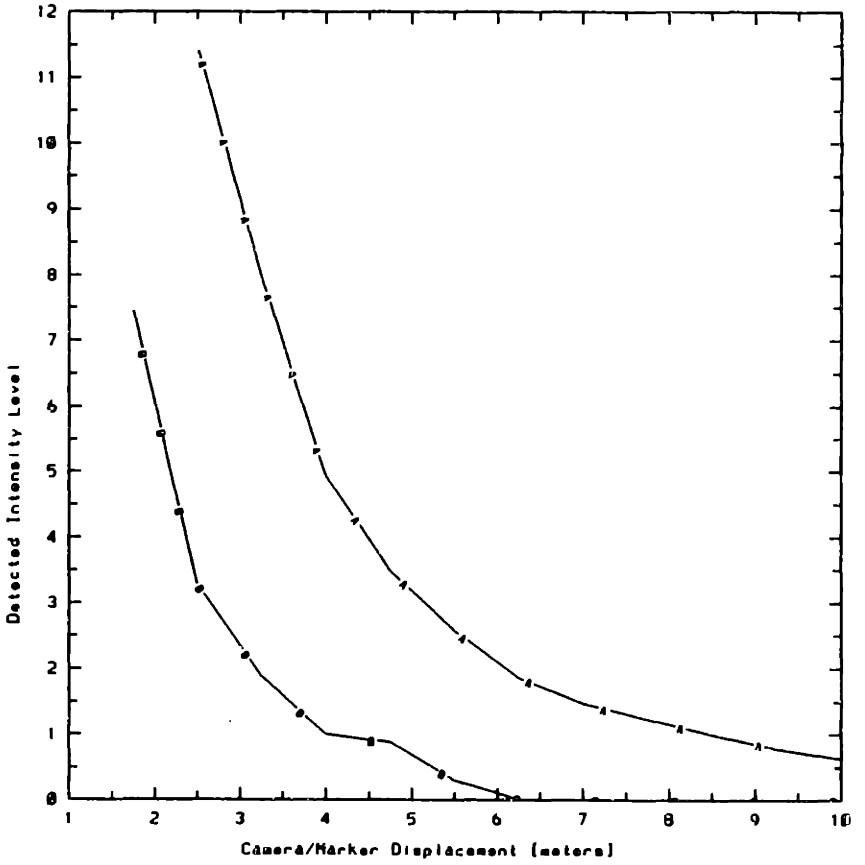


Figure 3.3: Marker detected intensity as a function of marker/camera displacment. (A) LED triad; (B) Single LED

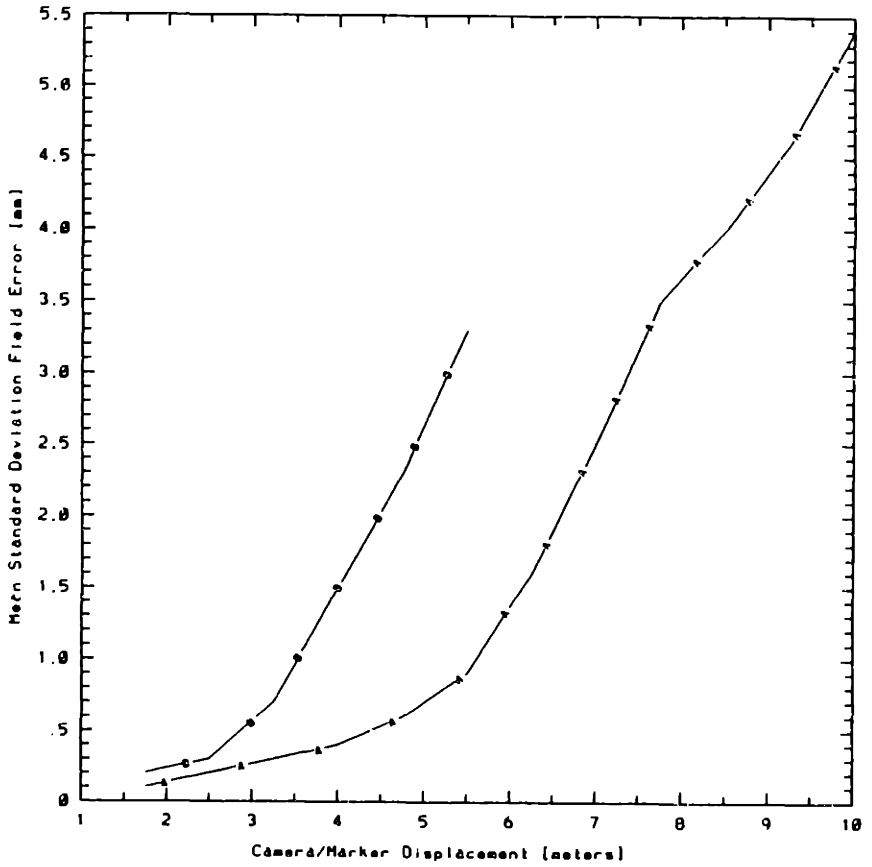


Figure 3.4: Physical error in the camera field from signal noise as a function of camera/marker displacment. (A) LED triad; (B) Single LED

displacements from the camera. The plot of figure 3.4 relating camera displacement and physical error in the camera field, was computed from the raw signal noise data using markers having zero rotation relative to the radial line. In the 3D computations, error in camera data propagates though the calculations to produce error in each of the 3D components. These errors in turn affect estimates of position and orientation of rigid bodies. Error propagation to 3D estimates has been investigated by Woltring [48]. Error propagation to position and orientation estimates has been studied by Morris [28] and Murphy [29].

A second type of stochastic error is caused by reflection of infra-red light from the active marker. A lateral photo-effect diode is only capable of detecting the intensity

weighted centroid of infra-red light incident upon its surface at a given instant. If light from the marker is reflected from a nearby surface in view of the camera then reflected light entering the camera will distort the perceived direction to the source.

For the most part reflections are a problem that can be reduced but never completely eliminated. Large reflections are often easy to detect as they usually lead to large skew in the 3D reconstruction rays (from colinearity condition). However, even if reflections are detected in a data set there is no way to recover the uncorrupted data. The most that can be done is to make some guesses as to where the reflections are coming from, try to correct the problem, and rerun the experiment.

Controlling reflections involves either physically obstructing the camera's view of the reflecting surface or applying a light absorptive material to the surface. Obstructing the camera's view of the reflecting surface is only feasible when desired movement in the camera field will not also be obstructed. A typical application of this type of reflection control is used during human gait studies. A low profile light baffle (figure 3.5) is used to obstruct the cameras' view of the floor between the subject and the camera. A specular floor finish, rather than a diffuse finish, allows for better prediction of reflection locations for optimal baffle placement.

The most insidious problem with reflections results from light bouncing from surface to surface creating a general background illumination of the laboratory space. Insight to this problem was gained by performing repeated external camera calibrations (to be described in Chapter 4) under different "passive" conditions in the lab. Repeated calibrations under identical conditions were performed to establish a baseline of variation

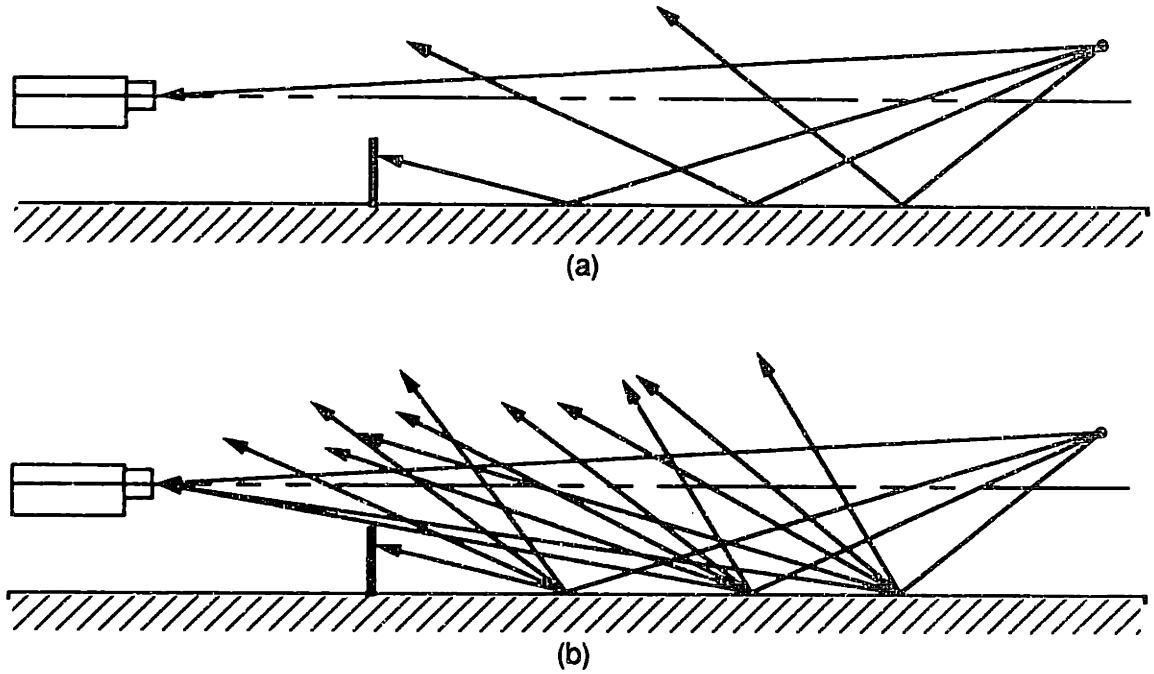


Figure 3.5: Use of light baffle to selectively block floor reflections. (a) Specular floor surface; (b) Diffuse floor surface

in the calibrations themselves. This was followed by repositioning various objects in the lab that were not in the cameras' field of view. Repeated calibrations were again performed and compared against the original calibrations. The difference indicates a shift in global origin position relative to the cameras on the order of a few millimeters depending on the severity of the changes made. The metallic venetian blinds behind the cameras were particularly troublesome, causing a significant change in the calibration when opened and then closed.

3.3 Systematic Error in SELSPOT Data

The dispersion of repetitive measurements about a mean was indicated above to represent the stochastic component of measurement error. The displacement of the mean measurement from a known reference will here be called the systematic component of measurement error. Under identical operating conditions the systematic error is consid-

ered to remain constant. Thus compensation for systematic error is possible through calibration.

3.3.1 Camera Nonlinearities

With SELSPOT cameras, systematic errors are observed in transmission and transduction of the marker signal through the lens, detector and internal electronics. These errors vary with marker position in the camera field of view and the intensity of marker light reaching the camera. The net result of these errors makes straight lines in object space appear curved in image space.

Calibration (linearization) of the camera image based on marker position in the field of view is addressed in Chapter 4. The additional nonlinear effect caused by fluctuation in marker intensity is discussed here. The effect may be described as follows: if a marker is placed in a fixed location and observed over a range of intensity values, not only will the noise (stochastic error) level change but the mean detected position (systematic error) will vary in a repeatable manner. The magnitude and direction of the variation depends on the general location of the image on the image plane. Experiments to identify this effect consisted of placing a single LED marker at close range to the camera on the crosshead of a large X-Y plotter (described in Chapter 4). The camera was positioned having its coordinate axes parallel with the coordinate axes of the plotter. A circular, variable density filter masked the marker allowing variation in the detected marker intensity level. The marker was moved with the plotter to different regions of the camera field and mean detected position was recorded at a range of intensity levels. Variations were generally on the order of one or two camera units toward the center of the field with variations of

10 or more camera units typical at the edges.

3.3.2 Marker Center Uncertainty

Another source of systematic error is the difference between the assumed and actual position of the point source of light emission from the infra-red markers. It is normally assumed that the center of light emission is coincident with the geometric center of the marker. In practice, slight errors are incurred by assuming this. In multi-LED markers, if one LED is emitting at a lower intensity, the observed center of light emission will be shifted away from that LED. Since these errors are repeatable they have the potential to be calibrated out of the system.

A series of experiments designed to identify typical magnitudes for this type of error were performed. The marker tested comprised a series circuit of three Litronix LD242-3 LEDs arranged with parallel axes in an equilateral triangle having 6 millimeters between LED axes.

A five degree-of-freedom positioning test stand was assembled from Klinger optical mount components (see Figure 3.6). A marker mounted on the distal end of the positioner could be aligned for rotation about any axis parallel to the marker axis or about any vertical axis perpendicular to the marker axis.

Two experiments were performed. The first sought to identify the center of light emission displacement from the marker axis. The positioner was placed approximately 3 meters from a camera. At this range one SU corresponds to approximately 0.25 millimeters of displacement perpendicular to the camera axis. A pointer fixed in global space was used to align the positioner's horizontal rotational axis with the marker axis.

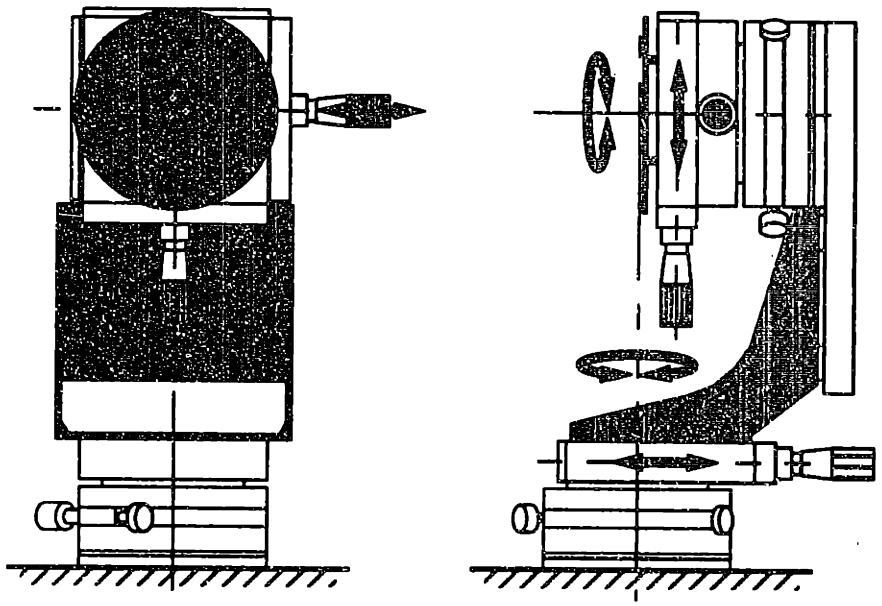


Figure 3.6: Five degree-of-freedom LED positioning stand

The marker was rotated through 360 degrees about its axis while (u,v) camera position readings were recorded. The readings indicated no more than ± 1 SU variatation in the detected marker location through the full rotation. This indicates the center of light from the marker deviates no more than 0.25 millimeters from the marker axis.

The second experiment attempted to identify the location of the center of light emission on the marker axis of symmetry above the marker surface. The positioning stand and marker were arranged as before with the vertical rotation axis passing through the point at which the marker axis intersects the LED plane. Camera data was recorded as the marker was rotated to $\pm$ 60 degrees about the vertical axis. This test indicated the center of light emission was located approximately 1 millimeter from the top surface of the LEDs, outward along the marker axis.

Chapter 4

Close-Range Photogrammetric Camera Calibration

Chapter 2 introduced fundamental analytical relationships required for estimating the position of points in 3D space based on 2D camera images. Those relationships contained parameters which were assumed known. Those parameters were: the global-to-camera rotation matrix M and position vector $\mathbf{x}_{oc}$, the camera principal distance d and principal point $\mathbf{u}_p$, and the internal corrections Δu and Δv. In general these parameters are different for each camera and constant only for a particular camera in a fixed geometric configuration. Quantification of these parameters, either explicitly or implicitly, is the end product of camera calibration.

This chapter assumes that cameras, once placed, remain fixed in the global coordinate system and that if cameras are moved then one or more aspects of calibration must be performed to reestablish the calibration parameters before reducible data may be acquired. This is in contrast to the more complex task of calibrating a camera which moves with respect to the global coordinate system. The solution to that problem, required for Large Volume TRACK, builds on the discussion in this chapter and is developed further

in Chapter 5.

This chapter presents a new approach to calibration that effectively treats the identification of internal and external orientation parameters as separate problems. The first section of this chapter contains a review of previously available calibration techniques. The Direct Linear Transform is reviewed in detail as it forms the basis from which to discuss the new technique for external calibration and is also compared with the new technique for performance. The remaining sections of this chapter discuss details of the new technique including theory, implementation, hardware components, and performance.

4.1 Review of Analytic Calibration Techniques

Historically, photogrammetric investigations were primarily carried out using *metric* cameras for which the elements of interior orientation were designed into the camera system. Precise (and expensive) design and construction provided instruments having internal parameters known to a high level of accuracy. Applications for these devices were typically in aerial photogrammetry and cartography.

In response to photogrammetric applications in a variety of fields in which the specialized construction of metric cameras was prohibitivly expensive, various approches for identifying the elements of internal and external orientation through experimental means have been explored. In all methods a set (or sets) of reference markers are placed in the field of view of one (or all) cameras. Methods vary in assumptions pertaining to reference marker position definitions in 3D space, number of marker sets used and algorithmic reduction of image data to the camera orientation parameters.

The most widely used method for calibrating non-metric cameras was developed by Abdel-Aziz and Karara [1]. This technique, called the Direct Linear Transform (DLT), allows for identification of elements of internal and external orientation in a convenient implicit form. The DLT method has been observed to exhibit poor accuracy when extrapolating to estimate 3D positions outside the volume of reference markers used [51]. This method therefore requires that reference markers fill the desired workspace. This requirement becomes cumbersome if a large workspace is desired. Improvements to the DLT accuracy have been made [14] but the extrapolation problem persists.

Woltring developed the Simultaneous Multiframe Analytical Calibration (SMAC) technique [46] [47] which provides explicit solutions for the internal and external orientation parameters. A planar distribution of markers is arbitrarily placed in at least three unique postions and orientations. One such planar position defines the global reference frame while the other positions are arbitrary and unknown. SMAC employs a sophisticated, statistical approach for reducing the information viewed by each camera of each planar position to the desired orientation parameters.

Dapena [9] and Fijan [10] developed similar calibration methods which eliminated the need for known reference markers. Each method employs separate procedures for internal and external parameter identification. In Fijan's case, the internal calibration is the same method described in Section 4.2.1. The external calibration consisted of identifying the camera position and orientation which minimizes in a least-squares sense the skew error between rays of colinearity of arbitrarily located markers observed by two cameras. In practice the method finds a solution to within a scale factor. If the camera baseline displacement or the constant displacement of two markers on a rod is known,

then the scale factor may be estimated and the solution is complete. The limitation with this method is that it does not directly solve for the camera positions and orientations with respect to a defined global reference frame.

The reasons for development of an alternative method of calibration are based on several factors which occured during TRACK development. One factor is that all of the above described calibration methods share a common problem concerning verification. No provision is made for verfying that markers not used in the original calibration may be accurately estimated in 3D. This fact led to developement of a novel calibration structure which provides an unlimited number of reference markers for calibration and/or verification.

Another factor concerns the history of TRACK calibration. Prior to the author's involvement with TRACK only internal parameter calibration had been applied. External calibration was assumed unnecessary due to careful mounting of cameras on optical bench equipment. As described below the internal calibration technique quickly evolved to the point where rapid and repeatable preformance and verification was possible. In contrast, the assumed external calibration was found to be severly lacking in performance.

This naturally led to treatment of external calibration as a separate problem and to a new overall calibration approach involving separate internal and external parameter identification. This approach was designed to be computationally efficient and to avoid the stated problems inherent in the methods described above. In particular its performance will be compared with the DLT method. For this reason the DLT is described in detail below.

The Direct Linear Transform

As originally developed, the DLT method simultaneously solved for the comparator-to-image[1] and image-to-object transformations. In this presentation it is assumed an opto-electronic camera is used which provides image point information directly relative to an image coordinate system fixed in camera coordinates. In effect the DLT method assumed a camera model similar to the model presented in Chapter 2. The only difference is in the treatment of the elements of interior orientation. Abdel-Aziz and Karara used linear functions of image coordinates (u, v) to represent image distortion and optical misalignment. These functions were of the form:

$$\bar{u} = c_1 + c_2 u + c_3 v \tag{4.1}$$

$$\bar{v} = c_4 + c_5 u + c_6 v \tag{4.2}$$

where $c_1, \ldots, c_8$ are constants representing the parameters $\mathbf{u}_p$, d and functions Δu and Δv as presented in the previous chapter. This results in a transformation identical in form to equation 2.6 with the substitution of the above $\bar{u}$ and $\bar{v}$. Algaebraic manipulation of this equation leads to

$$u = \frac{l_1 x_o + l_2 y_o + l_3 z_o + l_4}{l_9 x_o + l_{10} y_o + l_{11} z_o + 1} \tag{4.3}$$

and

$$v = \frac{l_5 x_o + l_6 y_o + l_7 z_o + l_8}{l_9 x_o + l_{10} y_o + l_{11} z_o + 1} \tag{4.4}$$

where the 11 parameters l_i (the DLT parameters) are nonlinear combinations of the orientation parameters $\mathbf{M}, \mathbf{x}_{oc}, d, c_1, \ldots, c_6$. Note the similarity of 4.3 and 4.4 to the colinearity equations 2.8 and 2.9.

[1] A comparator is a device for digitising film image points in a fixed reference coordinate system

The main feature of the DLT is that the 11 DLT parameters may be treated as constants for a given camera in a fixed orientation relative to global coordinates. No reduction to exact internal and external parameters is required. The exact parameters are implicitly solved for in knowing the DLT parameters.

The linearity of 4.3 and 4.4 permits their expression as two equations with either the eleven DLT parameters unknown ($\mathbf{x}_o$, u, v known) or with $\mathbf{x}_o$ unknown (DLT parameters and u, v known). In the first case the DLT parameters may be solved for by placing reference markers in each camera's field of view. If at least six non-coplanar reference markers are observed by the camera then twelve or more equations are available for reduction via least-squares to the eleven unknown parameters.

In the second case, once the DLT parameters have been determined for each camera the equations may be written in terms of an unknown $\mathbf{x}_o$ vector by combining the DLT parameters into constant coefficients of two general equations for a plane in space. The intersection of the two planes defines a ray of colinearity as discussed in Chapter 2. If n cameras are used to observe a single arbitrarily located marker, then $2n$ equations in terms of the three unknown components of $\mathbf{x}_o$ may be solved for via linear least-squares.

4.2 Separate Internal and External Calibration

As explained previously, limitations in existing calibration methods and previous MIT experience with internal camera calibration led this research investigation toward the new calibration method described below. Also providing a push was the development of Large Volume TRACK for which it was anticipated that explicit identification of all orientation

parameters would be required. The internal calibration problem was developed first as it was needed immediately in order to bring the SELSPOT-II cameras on line with existing TRACK software. Subsequent effort in external calibration began with investigation of the DLT method with an eye toward explicit parameter identification. The resulting new method evolved as elements of the DLT approach, not required for a system in which the elements of internal calibration are already known, were discarded and new techniques implemented.

4.2.1 Internal Camera Calibration

This section describes the technique developed for determining elements of interior orientation, d, u_p,v_p, Δu, and Δv, of a particular camera. The procedure involves mapping systematic error (distortion) in the camera field of view into a lookup table array of correction values. In the process, the principal point (u_p,v_p) is zeroed and an effective principal distance d is computed. In the case of SELSOPT-II cameras, sources of systematic error are in the lens, in the lateral photo-effect diode detector, and in signal processing electronics. All of these effects may be considered intrinsic to a particular camera and have been shown to be constant over time.

In Chapter 2 general transformation equations, 2.8 and 2.9, were presented which relate an object point in 3D to its respective 2D image point for a particular camera arbitrarily located in the global reference frame. If we now assume the camera is aligned such that global object space axis z is colinear with the camera axis z_c, and the axes x_o and y_o are parallel with respective axes x_c and y_c (see Figure 4.1) then the rotation

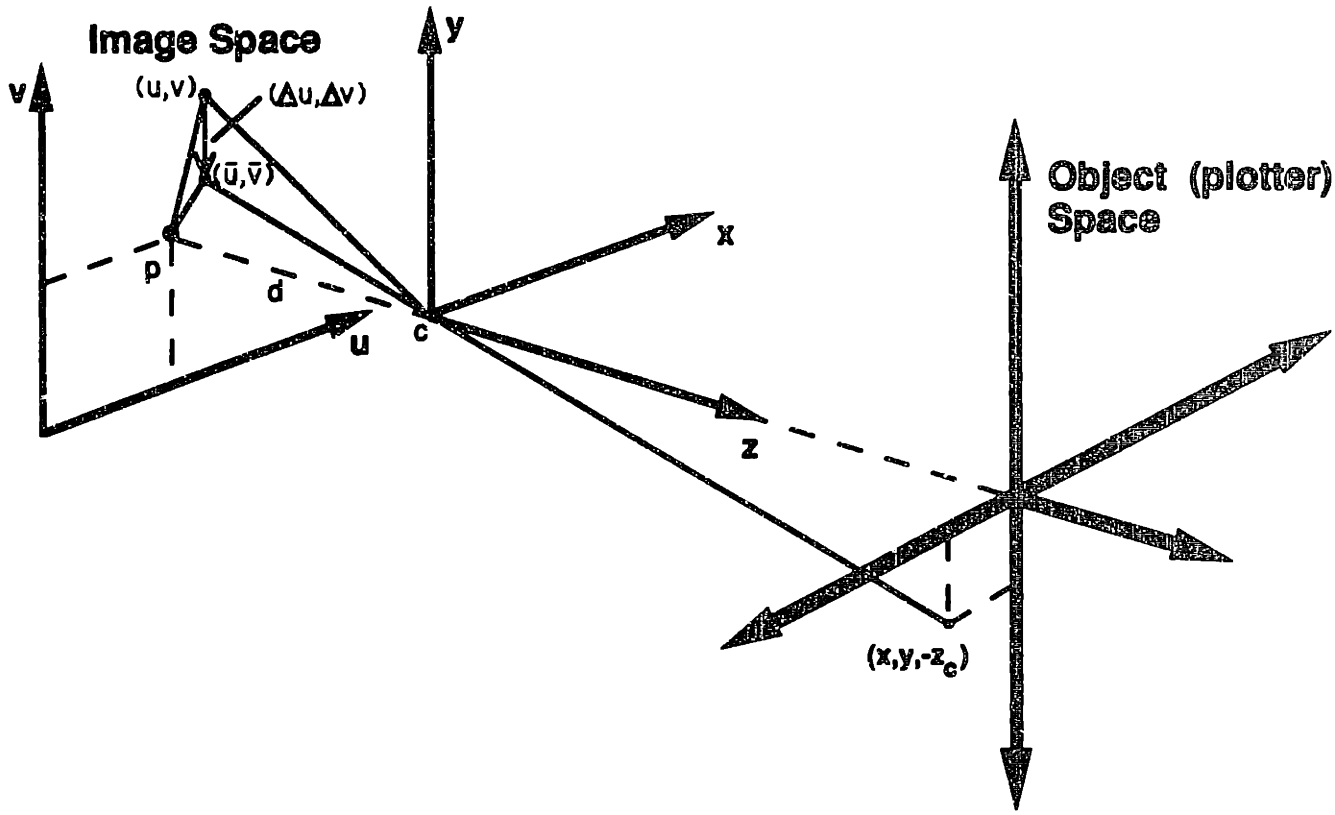


Figure 4.1: Internal calibration idealized camera and plotter geometry

matrix M becomes the identiy matrix I and the general equations 2.8 and 2.9 reduce to

$$\frac{\bar{u}}{d} = \frac{x_o}{z_o + z_{oc}} \tag{4.5}$$

$$\frac{\bar{v}}{d} = \frac{y_o}{z_o + z_{oc}} \tag{4.6}$$

where $\bar{u} = u - u_p + \Delta u$ and $\bar{v} = v - v_p + \Delta v$. Equations 4.5 and 4.6 are fundamental to relating a point in object space coordinates to a point in image space coordinates during internal calibration.

Internal Calibration Methodology

The general goal of internal calibration is to obtain a correction, Δu and Δv, for each discrete (u, v) coordinate in image space. In practice it is impractical to store a unique

correction pair for each discrete coordinate pair. A camera with 12 by 12 bit resolution would require at least $2 * (2^{12})^2$ bytes of memory to hold a correction pair for each coordinate pair. Continuous functions of image coordinates describing image distortion have been widely used and are typically described in standard photogrammetry texts. These functions are however limited in their ability to describe what may be highly non-linear and anti-symmetric image distortions. As an alternative, if correction pairs are stored for coordinate pairs at regular intervals across the camera field of view, interpolation between corrected coordinates will suffice provided the correction field is sufficiently dense.

The method described here for use with SELSPOT-II cameras defines a grid of 51 by 51 coordinate pairs. Each square of the grid spans 81 discrete camera units with a border of 23 camera units remaining outside the grid. Correction value pairs applied to each (u, v) data pair are derived from a look-up table of known points using a finite element technique involving the computation of *interpolation functions*. The use of interpolation functions in this manner was first described by Antonsson [3] but, for completeness, the technique is reviewed here.

First, the grid square containing the (u, v) camera coordinate being corrected is identified and the corner (node) positions, in camera coordinates, are computed. The positions of u and v within the selected grid square are scaled to lie in a range of from -1 to $+1$ in proportion to their position within the boundaries of the square. Four interpolation functions, h_i of the form

$$h_1 = \frac{1}{4}(1+r)(1+s) \quad ; \quad h_2 = \frac{1}{4}(1-r)(1+s)$$

$$h_3 = \frac{1}{4}(1-r)(1-s) \quad ; \quad h_4 = \frac{1}{4}(1+r)(1-s)$$

are computed, where (r, s) is the scaled (u, v) coordinate position. Correction values c_{ri}, c_{si} at each node i are read from the look-up table and corrections, Δu and Δv, to be added to (u, v) are computed via

$$\Delta u = \sum_{i=1}^{4} h_i c_{ri} \quad ; \quad \Delta v = \sum_{i=1}^{4} h_i c_{si}$$

During typical data processing, this procedure is repeated for every data coordinate pair processed in each camera.

Identification of the camera principal point (u_p, v_p) involves simply defining it to be the center of the camera view (2048,2048). This definition is used during the calibration process so that any error in this definition will be accounted for in the correction data.

The camera principal distance is selected so as to minimize the magnitude of the correction values used. This approach is valid since the principal distance is always used as a ratio to the corrected camera data and is decided upon before final correction values are determined. This is explained in more detail below.

Internal Calibration Apparatus

The main piece of apparatus used in internal calibration is a reconditioned Gerber X-Y plotter. The plotter was originally designed to generate large mechanical drawings. Early technology stepper motors and drivers were used to position its crosshead throughout a 60 by 60 inch surface area. Internal gearing allowed for stepping resolution to 0.001 inches in each axis.

Plotter reconditioning involved motor and drive replacement, and gear and bearing cleaning and lubrication [5]. All existing electronic components were removed and

discarded. The motors were replaced with Superior Electric Co. Slo-Syn Model M111-FD12E stepper motors with model MPS 3000A Power Supply and model STM103 Translator. Since it was desired to control two motors with one power supply and translator, a relay arrangement was used to allow selection of either motor from the same translator.

Two limitations were imposed by this:

1. Only one axis could be driven at a time. This was not considered a problem since speed of overall movement is not a factor during calibration.

2. The relay switches the signals to the four motor windings. There are eight combinations (or phases) of motor winding activation which when proceeded through in sequence, under control of the translator, produce normal motor rotation. Each phase represents a discrete motor step. If the motors are not positioned at the same phase when the relay switches, then the motor being switched to will make a sudden jump to the current translator defined phase. Since motor position is assumed to correlate to the number of pulses sent out, switching motors out of phase results in corruption of position information. The solution to this is to only move the motors in multiples of eight steps, thereby assuring the motors will be at the same phase when switching between them. This has the drawback of effectively limiting stepping resolution to 0.008 inches (eight times the plotter step resolution of 0.001 inches)

Internal Calibration Procedure

Calibration setup involves ensuring that the plotter plane of motion is perpendicular to the camera principal axis z_c. While this alignment is not critical [1] it is desireable to maintain small correction values and gradients for purposes of minimizing storage requirements and maximizing interpolation function accuracy. To perform the alignment, a ring of eight markers was mounted about the camera lens, pointing outward parallel to the camera principal axis. A mirror mounted in the plotter's planar bed reflected the marker images back to the camera. The plotter and camera were adjusted until the ensemble average (u, v) position of all markers read at the principal point.

To collect a calibration data set, the plotter marker was sequentially positioned at each of the 2601 look-up table points in the camera's field of view. At each point, the plotter position, (x_o, y_o), was recorded.

To obtain z_{oc}, an additional calibration set is obtained at a different plotter displacement d'_p. Since each calibration set relates the same camera coordinates to a unique set of (x_o, y_o) plotter coordinates, the distance between planes $(d_p - d'_p)$ may be used to construct spatial rays which intersect each plotter plane at identical camera coordinates. All such rays must converge at the camera center of projection. The best intersection point of a torroidal selection of approximately 1000 rays provides the estimate for z_{oc}.

Knowing the camera object space coordinates (x_o, y_o, z_{oc}), the principal point (u_p, v_p), and the camera image space coordinates (u, v) of each marker in a calibration set allows for computation of a best principal distance d using Equations 4.5 and 4.6 with correction

[1] Any misalignment will be accounted for automatically in the correction values.

values Δu and Δv assumed to be zero. Once a d has been determined, Equations 4.5 and 4.6 are again applied using (x_o, y_o, z_{oc}), d , (u_p, v_p) and (u, v) to compute Δu and Δv for each grid node. The values Δu and Δv are placed in respective positions of a look-up table file.

Internal Calibration Results

Results of calibration are easily visualized from plots of constant correction lines in each of the u and v camera coordinates and from vector field plots of the combined correction in (u, v) across image space. Plots for two SELSPOT cameras[2] are presented in Figures 4.2 through 4.7.

Note that the two outer elements on all sides of the look-up table were eliminated due either to nonexistent camera coordinates or unacceptable error. During actual experiments, data found in this area is discarded.

Look-up table and calibration technique validity was checked by programming the plotter to move through a series of 2000 randomly generated positions in the camera field of view. At each plotter position, 100 camera data samples were averaged to eliminate noise. Camera data was corrected using the generated look-up tables and compared to the expected data value based on the known plotter to camera transformation. The comparison indicated that, aside from a field of view border of 250 camera units, the maximum difference between the observed position and the predicted position was 2.5 camera units with a standard deviation difference of ± 0.7 camera units. When combined

[2]The SELSPOT cameras used were borrowed from the Biomotion Laboratory at the Massachusetts General Hospital, serial numbers 146 (camera 1) and 178 (camera 2).

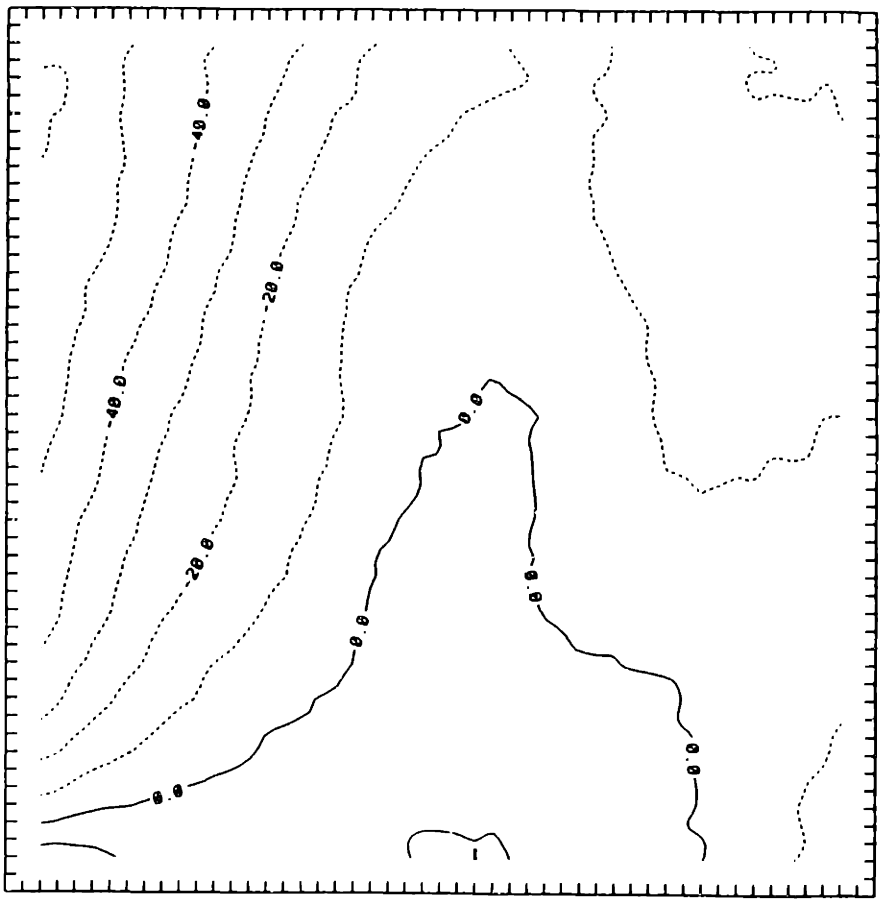


Figure 4.2: Correction Δu over (u, v) field: Camera 1 (sn. 148) [image units]

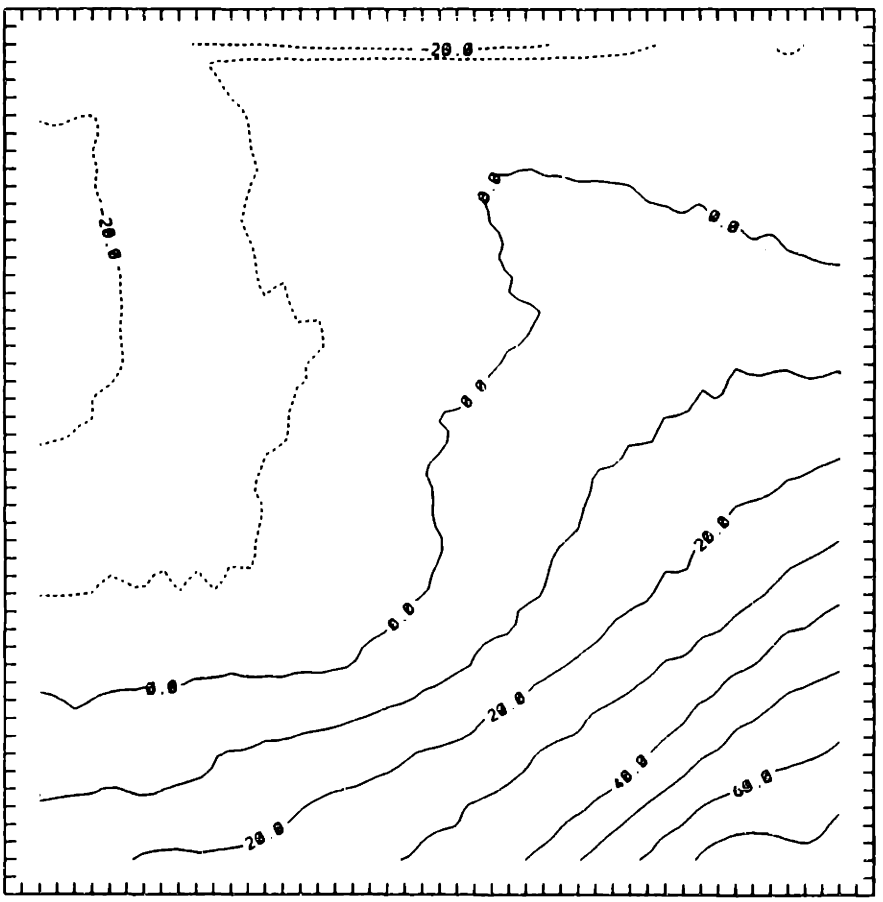


Figure 4.3: Correction Δv over (u, v) field: Camera 1 (sn. 148) [image units]

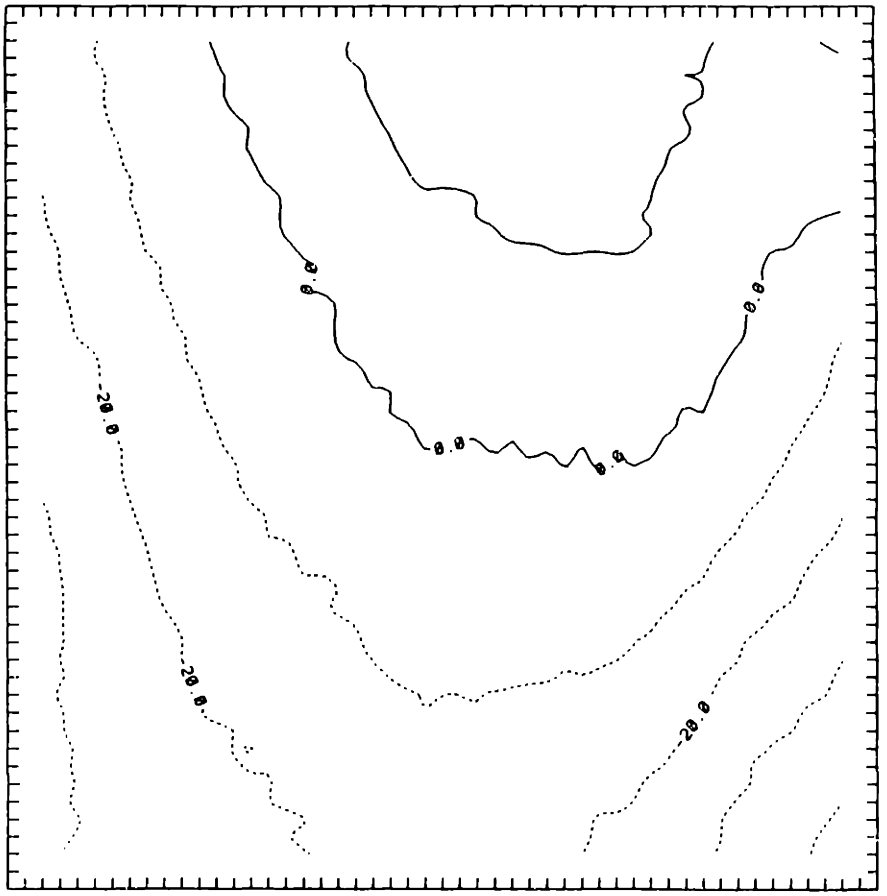


Figure 4.4: Correction Δu over (u, v) field: Camera 2 (sn. 178) [image units]

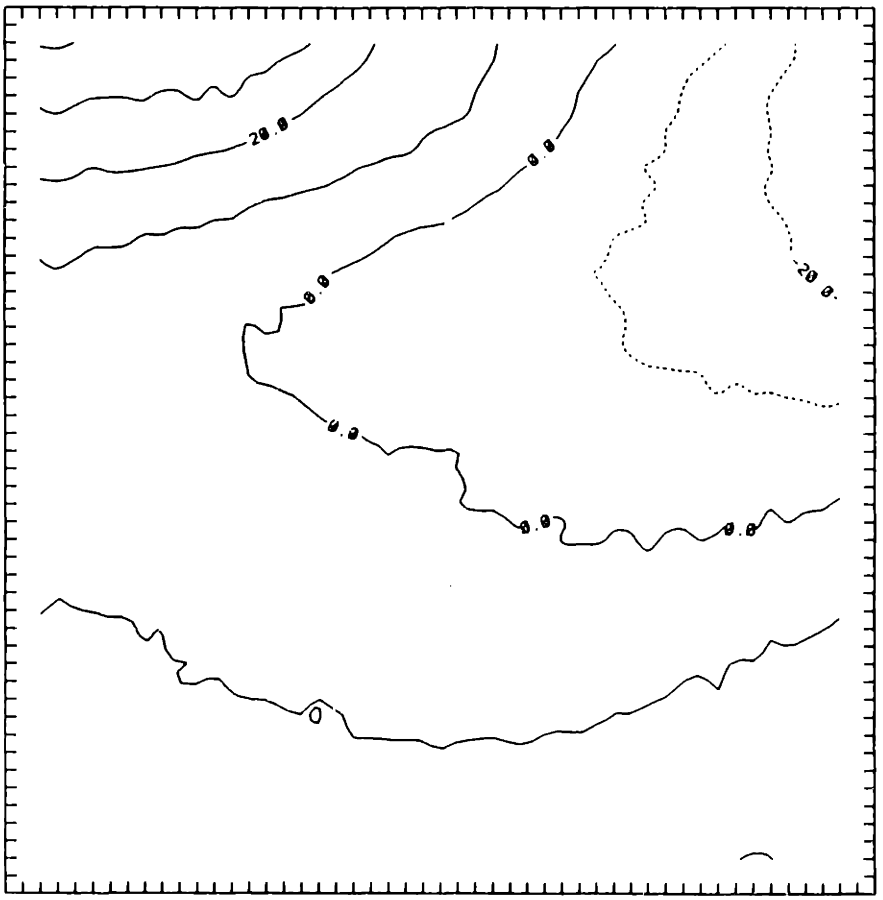
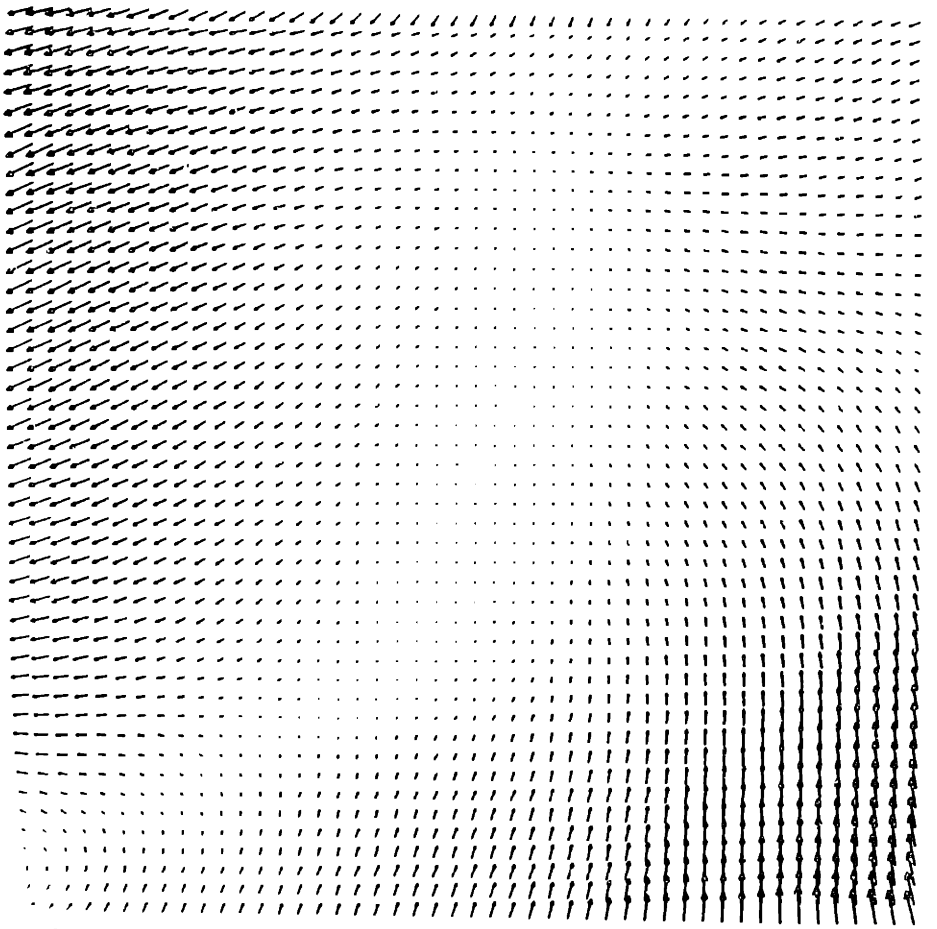


Figure 4.5: Correction Δv over (u, v) field: Camera 2 (sn. 178) [image units]



Figur 4.6: Correction vectors $(\Delta u, \Delta v)$ over field: Camera 1 (sn. 148) [image units]

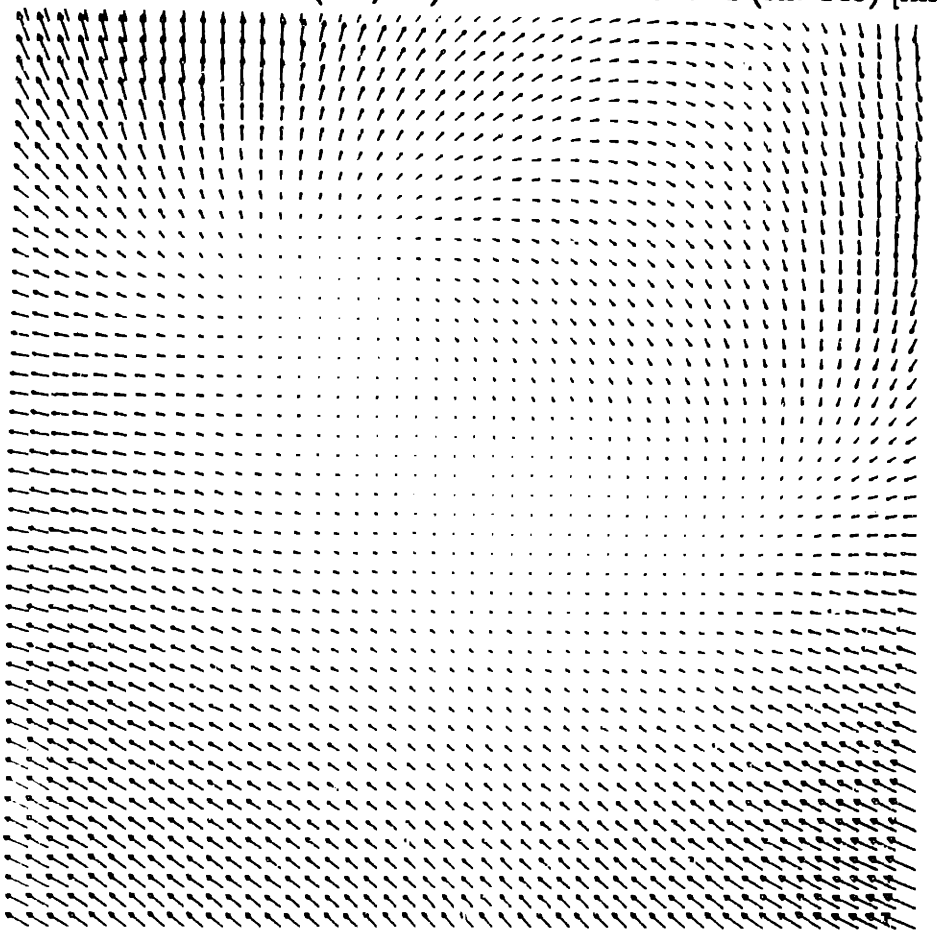


Figure 4.7: Correction Vectors $(\Delta u, \Delta v)$ over field: Camera 2 (sn. 178) [image units]

with an average uncertainty in plotter positioning of ± 0.2 camera units the resulting average overall uncertainty remaining in corrected data was ± 0.9 camera units. Expressed as a percentage of the reduced viewing field the overall worst case uncertainty was 0.07 percent while the average uncertainty was 0.025 percent.

This experimental protocol has been repeated several times over 18 months with similar results indicating no variation of overall camera non-linearities or any setup artifacts.

Further Work

The possibility of correcting for slight changes in the effective principal distance and calibration changes due to distance from the camera and intensity changes should ultimately be addressed. Calibration at a multiplicity of ranges from the camera and corrections for intensity could be accomplished and programmed readily given today's relatively inexpensive computer memory.

Accessing and interpolating camera data corrections is comparatively time consuming compared to the overall data processing time [22]. Tetewsky [41] first suggested implementing internal calibration correction values in hardware but lamented the lack of available memory to do so. This possibility has been considered for the current TRACK implementation. The approach would involve precomputing the corrections for all points in the camera field and storing the values in RAM. If corrections for all 4096 by 4096 points were precomputed, over 16 Megabytes of RAM would be required provided correction values do not exceed byte storage size. A more likely alternative would be to precompute alternate nodes, requiring approximately 4 Megabytes of RAM, and use a

nearest node criterion to determine the correction value to be used if the data does not fall on a precomputed node.

4.2.2 External Camera Calibration

In the previous section a method for estimating explicit values for the elements of interior orientation of a camera was presented. In this section the method for estimating explicit values for the external camera parameters, M and x_{oc}, will be presented. The general procedure is similar to other procedures discussed above in that a control volume of known 3D points is observed by individual cameras. The difference is in the way the data is processed, and in the ability of the calibration structure to generate additional control points, not used in the original calibration, for checking the accuracy of the calibration.

Algorithmic and Software Development

The algorithmic approach has three stages. It begins by generating a least squares solution to an overdetermined set of parameters, similar to the DLT method, but including the elements of interior orientation as known parameters. This generates an approximate, non-orthogonal, solution for the nine element rotation matrix and three element position vector. In the next step a best orthogonal matrix is derived from the approximation using polar decomposition (more generally, singular value decomposition). This is followed by successive small angle adjustment iterations to a best fit solution in a least-square sense. Each of the three steps is described in detail below.

Step 1: Initial Approximation Following the development of the colinearity equations presented in Chapter 2 the additional assumption is made that the elements of internal camera orientation are known. Writing the colinearity equations 2.8 and 2.9 in terms of the position vector of the global origin relative to the camera, $\mathbf{x}_{co}$, where $\mathbf{x}_{co} = -\mathbf{M}\mathbf{x}_{oc}$, the following 2 linear equations may be formed:

$$\begin{aligned} u' &= \frac{m_{11}x_o + m_{12}y_o + m_{13}z_o + x_{co}}{m_{31}x_o + m_{32}y_o + m_{33}z_o + z_{co}} \\ v' &= \frac{m_{21}x_o + m_{22}y_o + m_{23}z_o + x_{co}}{m_{31}x_o + m_{32}y_o + m_{33}z_o + z_{co}} \end{aligned} \tag{4.7}$$

A set of control markers, of known $\mathbf{x}_o$, may be observed by the camera being calibrated. The observed image points are combined with the interior orientation elements to produce u' and v'. Equations 4.8 now represent two equations in 12 unknown parameters.

$$\begin{aligned} m_{11}x_o + m_{12}y_o + m_{13}z_o + x_{co} - m_{31}x_o u' - m_{32}y_o u' - m_{33}z_o u' - z_{co}u' = 0 \\ m_{21}x_o + m_{22}y_o + m_{23}z_o + y_{co} - m_{31}x_o v' - m_{32}y_o v' - m_{33}z_o v' - z_{co}v' = 0 \end{aligned} \tag{4.8}$$

Using at least 6 control markers will result in at least 12 equations in the 12 unknowns.

Attempting to solve for the 12 unknowns directly in the form given yields a null space solution. A particular solution may be found by dividing through both equations by one of the unknowns and solving for the remaining unknowns in an inexplicit manner. In this case the unknown component z_{co} is selected and the following set of equations for n

control markers represented in matrix notation becomes:

$$\begin{bmatrix} x_{o_1} & y_{o_1} & z_{o_1} & 1 & 0 & 0 & 0 & 0 & -x_{o_1}u'_1 & -y_{o_1}u'_1 & -z_{o_1}u'_1 \\ 0 & 0 & 0 & 0 & x_{o_1} & y_{o_1} & z_{o_1} & 1 & -x_{o_1}v'_1 & -y_{o_1}v'_1 & -z_{o_1}v'_1 \\ x_{o_2} & y_{o_2} & z_{o_2} & 1 & 0 & 0 & 0 & 0 & -x_{o_2}u'_2 & -y_{o_2}u'_2 & -z_{o_2}u'_2 \\ 0 & 0 & 0 & 0 & x_{o_2} & y_{o_2} & z_{o_2} & 1 & -x_{o_2}v'_2 & -y_{o_2}v'_2 & -z_{o_2}v'_2 \\ \vdots & & & & & & & & & & \vdots \\ x_{o_n} & y_{o_n} & z_{o_n} & 1 & 0 & 0 & 0 & 0 & -x_{o_n}u'_n & -y_{o_n}u'_n & -z_{o_n}u'_n \\ 0 & 0 & 0 & 0 & x_{o_n} & y_{o_n} & z_{o_n} & 1 & -x_{o_n}v'_n & -y_{o_n}v'_n & -z_{o_n}v'_n \end{bmatrix} \begin{bmatrix} \bar{m}_{11} \\ \bar{m}_{12} \\ \bar{m}_{13} \\ \bar{x}_{co} \\ \bar{m}_{21} \\ \bar{m}_{22} \\ \bar{m}_{23} \\ \bar{y}_{co} \\ \bar{m}_{31} \\ \bar{m}_{32} \\ \bar{m}_{33} \end{bmatrix} = \begin{bmatrix} u'_1 \\ v'_1 \\ u'_2 \\ v'_2 \\ \vdots \\ u'_n \\ v'_n \end{bmatrix}$$

where the "barred" elements represent division by z_{co}.

Once the eleven "barred" elements have been estimated, an explicit estimate for the original 12 unknowns may be found by constraining the magnitude of the determinant of M to be unity[1] and z_{co} to be positive (the global origin is never behind the camera during calibration). Defining $\bar{\mathrm{M}} = (1/z_{co})\mathrm{M}$ then imposing the constraint $\|\mathrm{M}\| = \pm 1$ yields

$$z_o = |(\|\bar{\mathrm{M}}\|)^{1/3}| \tag{4.9}$$

The results of this step are estimates for the m_{ij} elements of M and estimates for the components of $\mathbf{x}_{co}$. In the following discussion the estimate for M will be refered to as A and the estimate for $\mathbf{x}_{co}$ as a.

Step 2: Generating an Orthonormal Fit The matrix A was estimated without fully imposing orthogonal constraints. It does not, therefore, represent a proper rotation. A best fit orthogonal matrix may be extracted from A via the general technique of Singular Value Decomposition tailored to the specific task of decomposing an arbitrary square

[1]this is a necessary, but not sufficient, requirement of all orthogonal transformations

matrix A into a symmetric matrix S multiplied by an orthogonal matrix Q. This is called Polar Decomposition (see Appendix A).

It is assumed that A will be "close" to orthogonal, meaning that the symmetric matrix resulting from the polar decomposition will be nearly the identity matrix, with diagonal elements close to unity and off diagonal elements close to zero. Eigenvalues of the symmetric matrix are stretching factors [40] and indicate distortion in the overdetermined solution. The result of this step is an orthogonal matrix Q which represents only the rotational component of A.

Step 3: Small Angle Adjustment Once a valid orthogonal rotation matrix Q and position vector a have been estimated, a best fit solution to six degrees-of freedom may be pursued. This method was suggested by Robert Fijan, a fellow PhD candidate. The goal is to determine M and $\mathbf{x}_o$ in

$$\mathbf{x} = \mathbf{M}\mathbf{x}_o - \mathbf{x}_{co} \tag{4.10}$$

Estimates Q and a will produce an approximate x called $\mathbf{x}_a$ where

$$\mathbf{x}_a = \mathbf{Q}\mathbf{x}_o - \mathbf{a} \tag{4.11}$$

The problem is now reduced to finding a small correction rotation B and translation b such that

$$\mathbf{x} = \mathbf{B}\mathbf{x}_a - \mathbf{b} \tag{4.12}$$

Combining 4.12 and 4.11 and comparing with 4.10 the desired result is

$$\mathbf{M} = \mathbf{B}\mathbf{Q} \tag{4.13}$$

and

$$\mathbf{x}_{co} = \mathbf{Ba} + \mathbf{b} \tag{4.14}$$

If **B** is assumed to represent a small rotation, it may be approximated in terms of three independent direction cosine elements (α, β, γ) [8] and represented as

$$\mathbf{B} = \begin{bmatrix} 1 & \gamma & -\beta \\ -\gamma & 1 & \alpha \\ \beta & -\alpha & 1 \end{bmatrix} \tag{4.15}$$

Rewriting 4.12 with 4.15 in colinearity equation form yields

$$\begin{aligned} u' &= \frac{x_a + \gamma y_a - \beta z_a + b_x}{\beta x_a - \alpha y_a + z_a + b_z} \\ v' &= \frac{-\gamma x_a + y_a + \alpha z_a + b_y}{\beta x_a - \alpha y_a + z_a + b_z} \end{aligned} \tag{4.16}$$

which is a function of only 6 unknowns $(\alpha, \beta, \gamma, b_x, b_y, b_z)$ and may be solved for in a manner similar to obtaining the initial estimate **A** in the first step.

In this case the set of $2n$ equations becomes

$$\begin{bmatrix} -u'_1 y_{c_1} & (u'_1 x_{a_1} + z_{a_1}) & -y_{a_1} & -1 & 0 & u'_1 \\ -(v'_1 y_{a_1} + z_{a_1}) & v'_1 x_{a_1} & -x_{a_1} & 0 & -1 & v'_1 \\ -u'_2 y_{a_2} & (u'_2 x_{a_2} + z_{a_2}) & -y_{a_2} & -1 & 0 & u'_2 \\ -(v'_2 y_{a_2} + z_{a_2}) & v'_2 x_{a_2} & -x_{a_2} & 0 & -1 & v'_2 \\ \vdots & & & & & \vdots \\ -u'_n y_{a_n} & (u'_n x_{a_n} + z_{a_n}) & -y_{a_n} & -1 & 0 & u'_n \\ -(v'_n y_{a_n} + z_{a_n}) & v'_n x_{a_n} & -x_{a_n} & 0 & -1 & v'_n \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ b_x \\ b_y \\ b_z \end{bmatrix} = \begin{bmatrix} x_{a_1} - u'_1 z_{a_1} \\ y_{a_1} - v'_1 z_{a_1} \\ x_{a_2} - u'_2 z_{a_2} \\ y_{a_2} - v'_2 z_{a_2} \\ \vdots \\ x_{a_n} - u'_n z_{a_n} \\ y_{a_n} - v'_n z_{a_n} \end{bmatrix} \tag{4.17}$$

The 6 unknowns in 4.17 may be solved for via linear least-squares.

In practice, elements (α, β, γ) of 4.15 are not always very close to zero. This results in **B** being non-orthogonal to a significant degree. In this case polar decomposition of **B** must be performed before it may be multiplied into the solution. If polar decomposition results in a symmetric matrix substantially different from **I** then the orthogonal approxi-

mation is used instead of B in 4.13 and 4.14. The process of adjustment is then repeated until B is sufficiently close to orthogonal without polar decomposition.

At this point an exact least-squares solution to the six degree-of-freedom external camera calibration problem has been obtained in terms of the rotation matrix M and position vector x_{co}.

Reference Structure Design

Implementation and testing of the external calibration approach described above was carried out using a specially designed reference structure. The structure was designed with several goals in mind. Primary among those goals was to minimize the potential for reflections. With active marker systems there is always the potential for objects (eg. a structure member) lying in the space between a marker and the camera to cast reflected marker light into the camera. This results in corruption of the camera data.

Calibration procedures are often checked by using the data acqusition system to estimate the position of reference markers after calibration and compare the obtained 3D position estimates with the assumed known positions. However, if reference markers used to perform the calibration are also used to check the calibration then systematic errors will be less apparent. For testing purposes it was desired to create a structure which would allow for the presentation of an alternate set of reference markers not previously used in the calibration. In doing this, no sacrifice in assumed 3D marker position accuracy should be made.

To realize these goals a planar structure of markers which could be accurately posi-

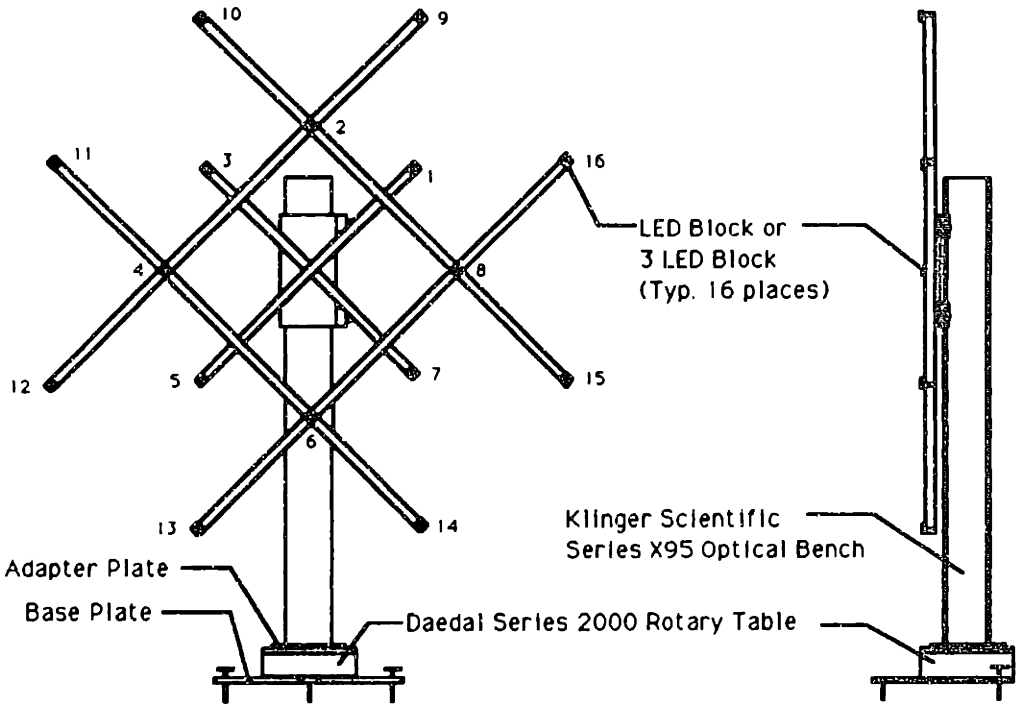


Figure 4.8: Calibration structure

tioned at different angles about a vertical rotational axis was designed and constructed. The proposed calibration procedure would involve sequentially positioning the marker structure at several known angular positions. Using the known structure angle, simple trigonometric relationships would allow for computing a known 3D marker position relative to an embedded reference coordinate system. This would have the effect of generating a unique volumetric reference structure for the external calibration process.

As designed and constructed the calibration structure consists of 16 markers arranged in two coplanar concentric circles of 8 markers each. The inner and outer circle diameters are 60 and 120 centimeters respectively. Square aluminum tubing forms a lightweight but rigid support structure for the markers (Figure 4 8).

Rotational positioning is accomplished using an 8 inch Daedal Series 2000 rotary positioning table having a positioning repeatability and accuracy to 0.2 and 3.0 arc minutes respectively. The worst case positioning accuracy occurs in the four markers

on the outer circle of markers at a radius of roughly 60 centimeters from the rotational axis. The rotatary table selected gives a positioning accuracy of 0.5 millimeters for these markers. With table calibration the inherent table repeatability could be used if needed to improve the worst case marker positioning accuracy.

The rotary table is driven by a Superior Electric M062-FD09 stepper motor having 400 steps per revolution. This combined with the 90:1 rotary stage drive ratio yields a rotary stage step count of 36000 steps per stage revolution.

The marker structure is mounted on a 150 centimeter long section of Klinger Scientific model X95 optical bench which in turn is screwed vertically into the rotary table surface. Klinger X95 optical bench carriages mounted on the marker structure allow for manual vertical positioning of the marker structure along the length of the optical bench section. Shims were used between the optical bench and the rotary table surface during assembly to ensure parallel alignment of the rotary table axis and optical bench axis.

The base of the structure carries the rotary table and consists of a triangular aluminum plate having pointed thumb screws at the rear corners and a fixed pointed screw at the front corner. Adjustment of the thumb screws permits leveling of the structure on the three pointed screw ends. The laboratory floor is surfaced with 12 inch linolium tiles. The base plate was sized so that placing the front point of the base at the point of intersection of four tiles allows for sighting along the rear edge of the plate to align it with the tiles. A mark on the stage and motor allow for repeatable positioning of the structure in the prefered laboratory coordinate system orientation. For portability, two small wheels are mounted at the back corners of the base plate and come into contact

with the floor when the thumb screws are retracted.

This design provides the following features:

- Minimum potential for reflections since at any one time all markers are in a planar configuration.
- Ability to provide a virtually unlimited combination of reference marker positions and alternate verification marker positions.
- Ability to point directly at camera being calibrated, thereby maximizing active marker intensity.
- Ability for all cameras located throughout 360 degrees to be calibrated with markers accurately referenced to a common coordinate system.

Reference Structure Geometric Verification

Once the calibration structure was constructed it was desired to verify the structure geometry as designed against the geometry as constructed and observed by the cameras. This involved identifying not only the relative position of markers in the plane of the structure but also identifying the position and orientation of the structure rotational axis in the structure coordinate system. Most errors in marker positions were created during manufacturing as a result of poor choice of materials. The thin walled aluminum tubing was observed to warp slightly during machining. The geometry seen by the cameras will in general also vary from any externally measured geometry due to uncertainty as to the exact point of light emission from the marker. While the difference between

marker geometric position and its camera observed position is small (see Chapter 3) it was deemed important to attempt to identify this error as well. With prior internal calibration the cameras themselves were deemed the best instruments for measuring the observed structure geometry. In addition, a direct measurement of the structure geometry was made. Both methods are described below.

Camera Measurement Based Structure Geometry Verification The procedure for verifying the calibration structure geometry using camera measurements involved primarily three steps. In the first step the structure was positioned in the view of both cameras and aligned with laboratory coordinates in a zero rotation angle. Data were collected from several rotated positions at which both cameras could view all markers at all times. Structure postions were equally distributed about a neutral position. The second step applied the method of Fijan [10] to determine a relative calibration of the two cameras based on a minimization of skew ray error from commonly observed points.

In the third step, data from individual markers observed at different structure angles are reduced to 3D points defining arcs of concentric circles distributed along the actual structure rotational axis. A least squares fit to all arc centers yields the 3D position of the structure rotational axis. The structure Y-axis was defined to lie along the rotational axis. The origin of the structure was defined to be the mean of all points used to estimate the axis. To determine X and Z axes a least-squares planar fit is made to structure data in the neutral position. The cross-product of the plane normal and the Y axis defines the X-axis. The Z-axis is then orthogonal to the X-Y plane in a right handed coordinate sense. Once the position and orientation of the structure coordinate system were

identified with respect to the relative calibration, the 3D marker positions estimated in relative calibration coordinates were transformed to structure coordinates. Ideally, these estimated structure marker coordinates will match the nominal "as designed" coordinates exactly. In reality, error will be identified in structure marker coordinates. It is the consistent and repeatable error that is of interest in this data.

In an effort to obtain consistent and repeatable error, fifteen data sets were collected. Cameras were placed with a baseline displacement of 2.5 meters at a range of 4 meters from the calibration structure. Five camera angle combinations at each of three structure height adjustments allowed for acqusition of fifteen data sets originating from different locations within the combined camera viewing volume. Each data set was composed of samples from both cameras at three structure rotational positions, -30.0, 0.0 and +30.0 degrees. It was anticipated that this approach would minimized the effect of error propagation from residual nonlinearity in each camera.

The nominal (as designed) geometry, adjusted geometry and geometry errors are displayed in Table 4.1. The maximum adjustment applied to any of the sixteen markers was 6.6 millimeters. The average adjustment was 3.9 millimeters. It is interesting to note that the majority of correction was seen in the z coordinate direction for almost all markers and all of the z coordinate errors are positive. This will be discussed further in section 4.2.3.

Direct Measurement-Based Structure Geometry Verification The calibration structure was measured directly with the use of the X-Y positioning calibration plotter used for internal camera calibration. A dial jaw caliper set was clamped to the plotter

Table 4.1: Nominal and camera measured calibration structure geometry (millimeters).

Marker	*Nominal*			*Measured*			*Difference*		
No.	*X*	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
1	-220.6	220.6	-102.5	-222.2	218.1	-104.4	1.6	2.5	1.9
2	0.0	324.7	-102.5	-1.4	323.1	-105.0	1.4	1.6	2.5
3	220.6	220.6	-102.5	220.9	219.7	-104.3	-0.3	0.9	1.8
4	324.7	0.0	-102.5	325.7	-0.5	-105.3	-1.0	0.5	2.8
5	220.6	-220.6	-102.5	221.3	-222.7	-103.9	-0.7	2.1	1.4
6	0.0	-324.7	-102.5	0.2	-326.1	-104.9	-0.2	1.4	2.4
7	-220.6	-220.6	-102.5	-220.7	-222.8	-105.0	0.1	2.2	2.5
8	-324.7	0.0	-102.5	-324.8	-2.3	-105.7	0.1	2.3	3.2
9	-220.6	545.3	-102.5	-222.8	541.6	-107.6	2.2	3.7	5.1
10	220.6	545.3	-102.5	219.2	545.0	-104.2	1.4	0.3	1.7
11	545.3	220.6	-102.5	548.5	221.0	-106.2	-3.2	-0.4	3.7
12	545.3	-220.6	-102.5	548.5	-221.1	-107.4	-3.2	0.5	4.9
13	220.6	-545.3	-102.5	222.2	-546.6	-107.0	-1.6	1.3	4.5
14	-220.6	-545.3	-102.5	-219.7	-547.4	-105.7	-0.9	2.1	3.2
15	-545.3	-220.6	-102.5	-544.9	-223.3	-104.3	-0.4	2.7	1.7
16	-545.3	220.6	-102.5	-545.0	216.9	-107.8	-0.3	3.7	5.3

crosshead with hole depth gauge pointed normal to and away from the plotter plane. A needle fixed to the end of the hole gauge was used as a pointer to make measurements of up to 6 inch range in the Z plotter direction. Moving the plotter crosshead in X and Y allowed for full coverage of the calibration structure planar marker distribution. Adjusting the caliper mounted pointer allowed for making out of plane measurements as well as identifying the calibration structure rotational axis.

Setup consisted of first identifying the point at which the structure rotational axis passes through the top cover surface of the central post of the calibration structure. This was done by using a fixed pointer to identify the point on the top surface that does not move when the post is rotated. Two wires fixed at the identified point pass over adjacent edges of the top cover and are fixed at the central post base at the midpoint

of the mounting. Together, these wires indicate the direction of the rotational axis. The calibration plotter was then brought within measurement range and the level of the calibration structure base was adjusted until the plotter pointer did not deviate from following each wire from bottom to top during a vertical crosshead movement. This operation assured that the vertical axis of the plotter was aligned with the calibration structure rotational axis.

Following this the calibration plotter coordinates were zeroed at the identified point on the top surface of the central post and all structure marker positions were sequentially identified. To assist in identification a mask was constructed having engraved crosshairs which when placed over a structure marker assisted in identifying the geometric center of the marker.

Measurement results are presented in Table 4.2. Note that now the nominal structure geometry has been adjusted to reflect the change in structure origin to the top center of the central post. The rather large positive and negative errors in the Z coordinate direction may be attributed in part to an offset in the rotation of the structure plane with respect to the zero reference mark used. This zero reference mark is fixed for all time so the offset will not have a detrimental effect.

4.2.3 Testing and Evaluation of Overall Internal and External Calibration

An evaluation of the calibration was performed using a few straightforward and repeatable tests. The goal of the evaluation was two fold: to identify what reconstruction error existed after calibration and to compare the calibration performance with the DLT

Table 4.2: Nominal and directly measured calibration structure geometry (millimeters).

Marker	*Nominal*			*Measured*			*Difference*		
No.	*X*	*Y*	*Z*	*X*	*Y*	*Z*	*X*	*Y*	*Z*
1	-220.6	102.1	-102.5	-223.5	103.6	-100.9	-2.9	1.5	1.6
2	0.0	206.2	-102.5	-2.0	207.2	-103.9	-2.0	1.0	-1.4
3	220.6	102.1	-102.5	218.4	102.2	-105.1	-2.2	0.1	-2.6
4	324.7	-118.5	-102.5	321.3	-118.9	-107.2	-3.4	-0.4	-4.7
5	220.6	-339.1	-102.5	216.8	-340.4	-105.2	-3.8	-1.3	-2.7
6	0.0	-443.2	-102.5	-4.3	-443.4	-103.8	-4.3	-0.2	-1.3
7	-220.6	-339.1	-102.5	-224.3	-338.8	-101.0	-3.7	0.3	1.5
8	-324.7	-118.5	-102.5	-327.8	-117.3	-101.0	-3.1	1.2	1.5
9	-220.6	426.8	-102.5	-222.5	427.7	-104.8	-1.9	0.9	-2.3
10	220.6	426.8	-102.5	219.3	426.9	-105.3	-1.4	0.1	-2.8
11	545.3	102.1	-102.5	543.6	100.5	-109.6	-1.7	-1.6	-7.1
12	545.3	-339.1	-102.5	541.9	-340.4	-111.2	-3.4	-1.3	-8.7
13	220.6	-663.8	-102.5	216.0	-663.9	-107.5	-4.5	-0.1	-5.0
14	-220.6	-663.8	-102.5	-225.4	-663.9	-101.0	-4.8	-0.1	1.5
15	-545.3	-339.1	-102.5	-549.7	-337.2	-97.9	-4.4	1.9	4.6
16	-545.3	102.1	-102.5	-547.4	104.8	-101.5	-2.1	2.7	1.0

method.

The test protocol in all cases consisted of mounting the cameras on a horizontal beam with a baseline displacement of 3.3 meters and a convergence angle of 25.0 degrees for each camera. This located the viewing volume center at a perpendicular distance from the baseline of 3.5 meters. The calibration structure height was adjusted to its highest possible setting and the camera mount heights adjusted to place the structure in the center of each cameras view.

Calibration data was obtained by generating an effective 3D structure for each camera individually. Each 3D structure was composed of three unique structure positions. The first position pointed the structure directly at the camera while the other two positions were at ± 30 degree rotations from the first. In each case data from all 16 structure

markers were collected and individually averaged over 300 frames.

Verification data was obtained by placing the calibration structure in its neutral position, pointing directly perpendicular to the camera baseline, and collecting data from both cameras simultaneously. Again, data from all 16 structure markers were collected and averaged over 300 frames. Note that this particular structure position was not the same as any position used in generating the calibration data and therefore defines 3D points for verification not used in the calibration.

Three different calibration methods were implemented and compared. In the first method the DLT was used without any preprocessing of the data. It was assumed this would help identify the raw performance of the system. In the second method the DLT was used with both calibration and verification data sets corrected using the internal calibration technique discussed in Section 4.2.1. This method, compared with the first, would identify any performance increase in applying the internal calibration. In the third method the full six degree-of-freedom (6 DOF) reduction was applied using internal and external calibration as discussed in Sections 4.2.1 and 4.2.2. This method was to be compared with the other two methods for performance.

Four sets of performance comparisons were processed. In each case one subset of markers from the calibration data was used to calibrate the system and then verify the 3D positions of a subset of markers in the verification data. In the first case all 16 markers from the calibration data were used to calibrate the system and verify 3D positions of all 16 markers in the verification data set. This test most closely resembles the tests commonly performed in calibration verifications.

Table 4.3: Mean error/(standard deviation) of predicted marker 3D position vs. known 3D position using different calibration methods with camera measured structure geometry (millimeters).

Calibrated Markers	Verified Markers	DLT *without* Internal Calib.	DLT *with* Internal Calib.	6 DOF Method
all 16	all 16	2.4/(0.6)	2.1/(0.8)	3.1/(0.8)
inner 8	outer 8	2.9/(1.5)	2.8/(1.4)	3.3/(1.7)
upper 7	lower 7	2.7/(1.4)	2.6/(0.9)	2.9/(0.5)
lower 7	upper 7	5.3/(2.1)	2.6/(0.8)	2.3/(0.6)

The next three tests were designed to identify the extrapolation capabilities of the calibration methods. The first of these employed calibration data from the inner 8 markers (1-8) to calibrate the system and used data from the outer 8 markers (9-16) to verify the calibration. The second extrapolation test used the upper 7 markers (1,2,3,9,10,11,16) to calibrate and the lower 7 markers (5,6,7,12,13,14,15) for verification while the third such test reversed this to use the lower 7 markers for calibration and the upper 7 markers for verification.

The four tests using three calibration methods each were repeated twice using first the camera measured structure geometry and second the direct measured structure geometry. For each method in each test the mean error and standard deviation in the predicted 3D position of the markers being verified against 3D marker positions from the assumed structure geometry was computed. The results of these computations are tabulated in Tables 4.3 and 4.4.

It is clear from the data that in both structure geometry cases the DLT methods are superior to the 6 DOF method when all 16 markers were used to calibrate and verify

Table 4.4: Mean error/(standard deviation) of predicted marker 3D position vs. known 3D position using different calibration methods with directly measured structure geometry (millimeters).

Calibrated Markers	Verified Markers	DLT *without* Internal Calib.	DLT *with* Internal Calib.	6 DOF Method
all 16	all 16	1.8/(0.7)	1.4/(0.5)	2.1/(1.1)
inner 8	outer 8	2.5/(1.0)	1.6/(0.5)	2.3/(1.5)
upper 7	lower 7	5.3/(1.4)	4.6/(1.6)	3.5/(0.4)
lower 7	upper 7	5.1/(1.7)	2.4/(0.9)	2.1/(0.6)

all 16 markers. The fact that the internally corrected DLT method is superior to the 6 DOF method indicates a likely persistence of non-linearity in the data even after internal correction. This slight amount of non-linearity is accounted for with the extra degrees-of-freedom in the DLT method of solution. The fact that the uncorrected DLT method is also better than the 6 DOF method attests to the ability of the DLT to "mold" itself to the non-linear properties of the raw camera system. In the second comparison the results were similar to the first.

Those results were initially very surprising. It was felt that in extrapolation one would witness a relative improvement in the 6 DOF method over the DLT methods. On the contrary it appears that the DLT methods were able to perform quite well under this test. The probable explanation for this lies in the nature of the distortion present in the tested system. If the distortion at the periphery of the viewing volume is well approximated by extrapolating distortion detected at the center of the volume then this would explain the observed results. In this case it is proposed that calibration with the inner ring of 8 markers captured the characteristics of the distortion to the extent that extrapolating to the outer ring of eight markers resulted in reasonable approximations

to the true 3D marker locations.

In the third and fourth tests the DLT method was required to extrapolate distortion observed at one side of the viewing volume to distortion on the opposite side. In these cases the uncorrected DLT method is clearly unable to perform as well as the 6 DOF method. The 6 DOF method is now comparable to the corrected DLT method and actually shows better performance in most cases. In keeping with the previous argument, in this case, calibrating with the upper (lower) markers did not capture the characteristics of the distortion to the point that extrapolation to predict the lower (upper) marker locations was any more accurate than the 6 DOF method.

The presented data argues for the internally corrected DLT method when 3D data originates within or near the periphery of the calibrated volume and for the 6 DOF method when data lies substantially outside of the calibrated volume. This argument would likely change depending on the amount of distortion present in the system. As the amount of distortion is decreased the 6 DOF method would more closely match the performance of the DLT. However, as will be seen in the chapter on Large Volume TRACK, the 6 DOF will be the method of choice there as it performs the required reduction to explicit camera position and orientation parameters.

A decision still must be made as to which structure geometry set is most valid. In general the data indicates better results from the directly measured structure geometry than the camera measured geometry. It is suspected that since a small amount of distortion evidently remains in the camera data that the camera based measurement of the calibration structure geometry is more likely in error than the directly measured struc-

ture geometry. This was backed up by a test in which 2.5 millimeters was added to the z-component of all marker locations in the camera based measurement data. This was done since it was noticed that on average all nominal marker locations were apparently in error by this amount (Table 4.1). Reprocessing the data led to much improved results across all tests indicating the camera measurement technique was not yielding results that made overall sense.

Additional measurements of the point of light emission from an LED marker were made indicating the marker geometric center is a very good approximation for the point of perceived light emission (see Chapter 3). For these reasons the directly measured data was adopted as the correct data for use with the calibration structure in the remaining studies.

Obviously additional tests and comparisons with the DLT and other calibration approaches could, and should, be made. Time and keeping toward the primary objective in this thesis prevented any such further investigations at present. The directly measured calibration structure geometry and the 6 DOF calibration method were integrated into the TRACK package for ultimate use with the Large Volume TRACK system discussed in Chapters 5 and 6. Internal calibration tables were loaded in software and used to correct all raw data acquired with SELSPOT. The DLT method was not integrated into TRACK but perhaps should be available for use in situations where the calibration structure fills the desired working volume. Implementation of the 6 DOF external calibration algorithm on the SUN 3/160, including the polar decomposition, is computationally efficient. Once calibration data have been acquired a solution (rotation matrix and position vector) for each camera is obtained in approximately 3 seconds.

Chapter 5

Large Volume TRACK (LVT)

The previous chapters, dealing with close-range photogrammetric calibration and error, have been leading up to the central issue of this thesis: the development of a close-range photogrammetric system having a large viewing volume. This system, called Large Volume TRACK (LVT), is a natural extension of the TRACK system as previously described. This chapter begins by reviewing the system's objectives and the various alternative approaches considered for meeting those objectives. The specific approach selected is then developed theoretically both from the standpoint of photogrammetric reconstruction and calibration

5.1 System Objectives and Alternative Approaches

As discussed in Chapter 1 the motivation for LVT comes from several perspectives, with the underlying objective to develop a system that maintains the accuracy and precision of the TRACK system while additionally being capable of acquiring data over a much larger volume than TRACK.

The desired measurement volume dimensions were not strictly specified. For human

gait measurement, a volume permitting acquisition of 3 or 4 consecutive gait cycles was considered important. This implied a 3 to 4 fold increase of the current volume in the object space X dimension to between 4 and 5 meters. For blind mobility studies freedom to move in the two dimensions parallel with the floor was desired. Increasing the measurement volume in the vertical direction was not considered important to any of the studies envisioned making use of the LVT system.

Other desired characteristics of the completed system were that it should be easy to use from a standpoint of setup, calibration and data acquisition. Flexibility in setup was also a key issue. Minimizing the number of constraints on selecting a working measurement volume was considered essential to allow application of LVT to the multi-faceted research environment typical of the Newman Lab.

5.1.1 Multiple Camera Method

One proposed method for enlarging the view of TRACK was to incorporate additional cameras into the overall system. For example, three pairs of converging axis cameras placed along a common baseline would permit an overall measurement volume approximating that shown in Figure 5.1.1. The advantage to this method would be the overall ease in extending what was already known about two camera photogrammetric methods to multi-camera methods. The software processing would be essentially identical. Additional cameras viewing the same LED marker would improve system accuracy by providing additional colinearity rays as input to the least-squares solution for the 3D marker location.

The principal disadvantage to this approach is a reduced overall sampling frequency.

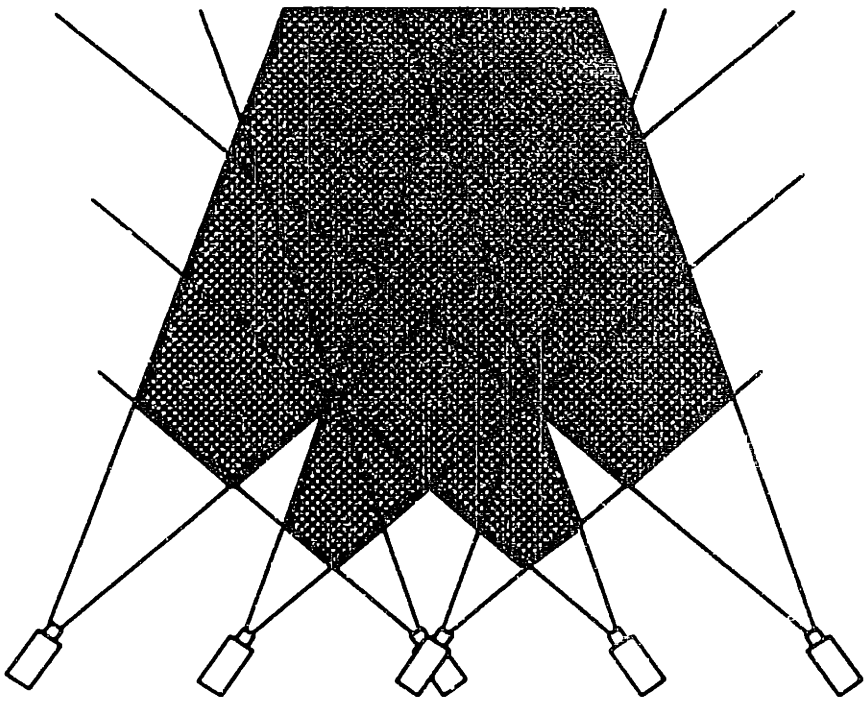


Figure 5.1: Possible multi-camera arrangement. Shaded area indicates measurement volume.

Opto-electronic systems perform best if the LED is pulsed independently for each camera thereby allowing independent LED pulse intensity adjustment. Each pulse cycle is 100 μseconds. Therefore, two cameras viewing 32 LED markers attain a maximum frame rate of approximately 150 Hz. Six cameras would reduce this rate to 50 Hz. This sampling frequency is on the borderline of acceptability for most human movement studies and would limit other types of studies. A solution to this problem continuously selects only cameras positioned for viewing current marker locations. This is not an option with commercially available opto-electronic systems. Expense is also an issue since additional cameras and interfacing are costly.

5.1.2 Moving Camera Methods

Various methods employing combinations of real-time active positioning of two cameras in translation and/or rotation were also proposed. The first approach (Figure 5.2) involves

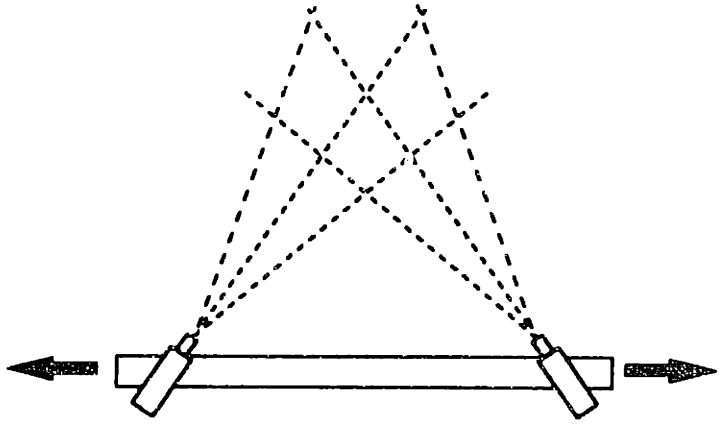


Figure 5.2: Common camera base translation (1 DOF)

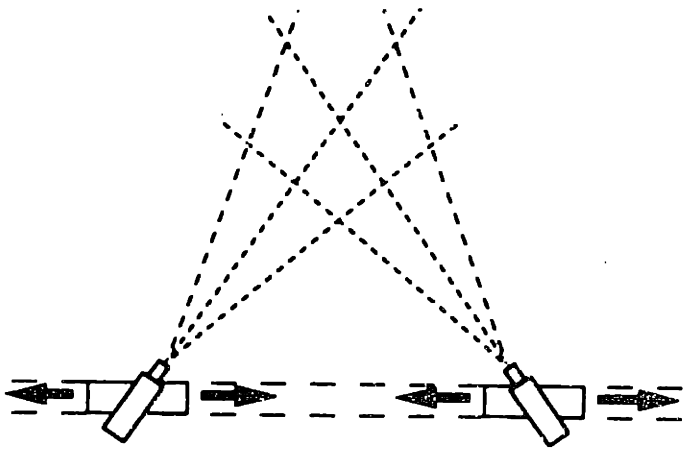


Figure 5.3: Independent camera base translation (2 DOF)

simple translation of two cameras on a common horizontal translational stage. This one degree-of-freedom (DOF) solution was attractive in that relative camera transformations remain constant thereby simplifying calibration and reconstruction algorithms. However, the measured volume would only increase in the dimension along the common base direction of travel; no added depth of measurement would be provided.

In the second and third methods (Figures 5.3 and 5.4) the problem of no increase in depth of measurement range is resolved. In the second method each camera is mounted on independent translational stages on a common horizontal guideway. In the third method cameras remain on a common translational stage but may rotate independently about

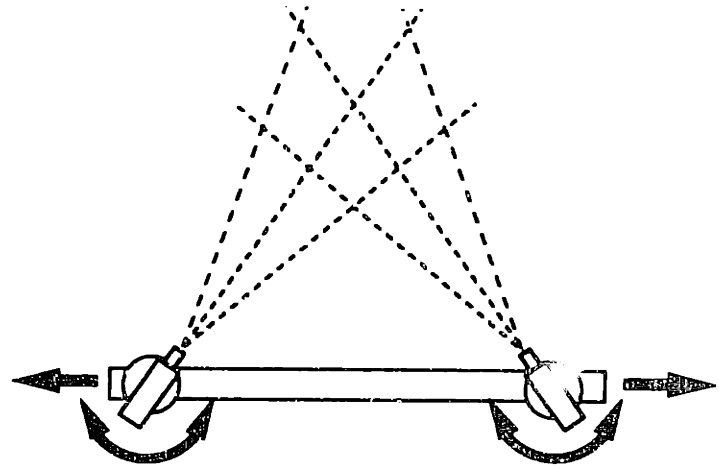


Figure 5.4: Common camera base translation with independent rotation (3 DOF)

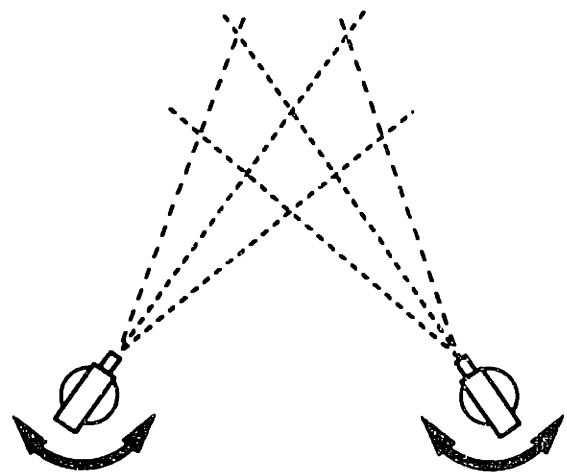


Figure 5.5: Independent camera rotation only (2 DOF)

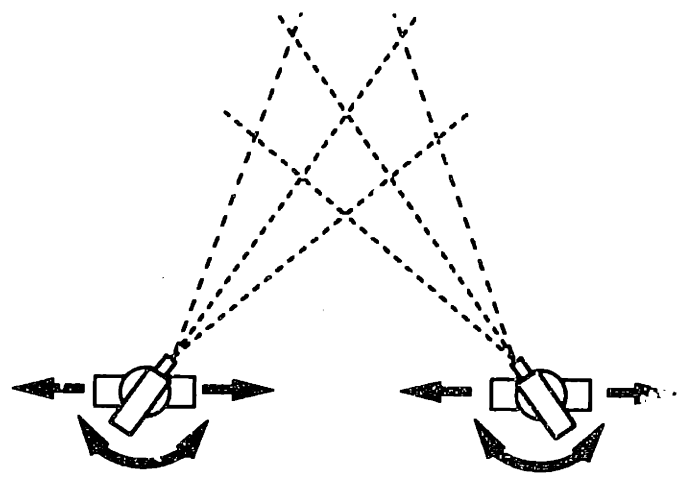


Figure 5.6: Independent camera base translation and rotation (4 DOF)

a vertical axis. With both of these methods, added complexity in calibration and 3D reconstruction is incurred due to the additional degrees of freedom compared with the first method. Choosing between these two methods would require analysis of requirements for measurement resolution in the horizontal dimension perpendicular to the camera baseline and measurement accuracy as effected by detected marker intensity. For example as the target marker moves away from the baseline, in method 3 (Figure 5.4) the camera axes become more parallel resulting in a decreased relative measurement resolution in that dimension. Alternatively, method 2 (Figure 5.3) maintains the camera axes relative angle and the relative 3D measurement resolution is not decreased. However, detected LED intensity level, which affects measurement accuracy (see Chapter 3), is maintained at a higher level with method 3 than method 2.

In the fourth approach (Figure 5.5) each camera rotates about a fixed vertical axis. This relatively simple 2 DOF method of camera positioning allows for adjustment for measurement depth but does not achieve adjustment of measurement range laterally as well as previously presented methods.

In the fifth approach (Figure 5.6) each camera can independently rotate and translate. This 4 DOF method, while technically the most complex, allows for optimal adjustment of camera convergence angle on the target to optimize some desired function of measured depth resolution and detected marker intensity (measurement relative accuracy).

All of these methods share the drawback that an expensive camera would be mounted on some type of servo positioning mechanism raising the possibility of damage to the camera or of calibration misalignment due to shock or vibration. This is especially

critical in a working laboratory environment where the typical system user will not be familiar with most of the underlying system components and is likely to initially make operational errors. In addition, running cables to the cameras and providing travel limit switches would increase design complexity.

The methods employing linear travel introduce unique problems. Linear bearings typically are available in lengths of approximately three meters. They may be joined at their ends to provide longer lengths of travel but vibration caused when traversing the joint would be likely. To maintain overall system accuracy, torsion in the guideway resulting in camera rotation about the pitch axis of more than approximately 0.1 milliradians could not be tolerated. Maintaining the required specifications for a very true, linear, guideway was deemed difficult over the long term. In addition, the range of experiments envisioned require the camera to be mounted anywhere from 0.5 meters to 1.5 meters from the floor. A linear guideway would be very inflexible in this application.

The design problem then became one of incorporating the principal of method 4 (Figure 5.5) without having to move the camera. The use of mirrors was suggested. As a result a sixth method was proposed and eventually adopted. This method is shown in Figure 5.7. Each camera is fixed and mounted vertically. Directly above each camera, a round, planar mirror, mounted at 45 degrees to a vertically rotating servo motor, generates a virtual camera image that sweeps a horizontal plane when the mirror/motor unit is rotated. This gives the effect of a rotating camera without actually moving the camera. Theorectical details of this method are described in the following sections of this chapter.

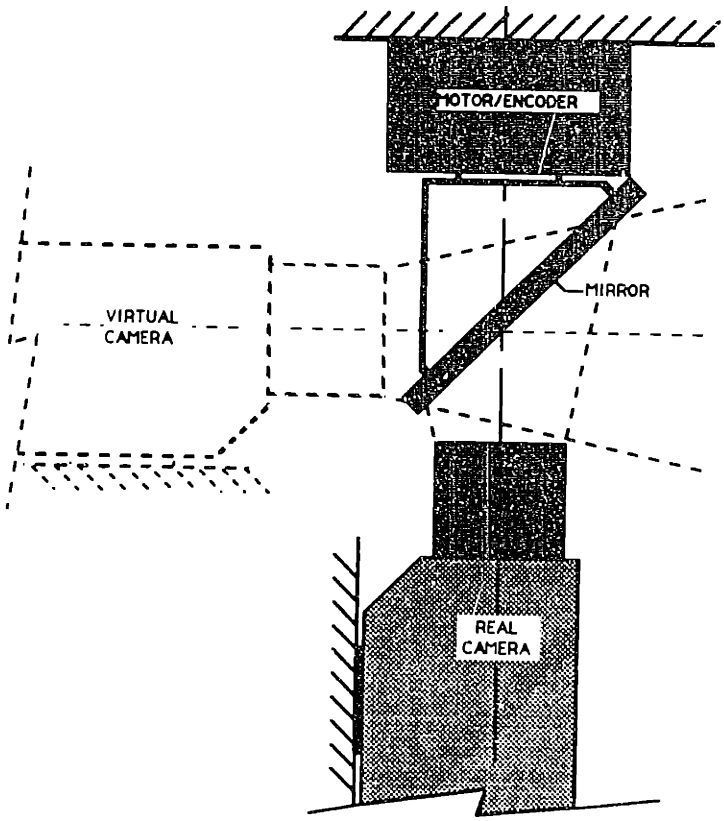


Figure 5.7: Camera/mirror arrangement

5.2 LVT Theoretical Development

A mirror reflecting the real world into a camera's view presents special problems from a photogrammetric perspective. The details of these problems will be presented in this section. The discussion begins by reviewing how reflections are handled mathematically and then addresses the specific case under consideration.

5.2.1 Basic Reflection Relationships

Elementary optics introduces the concept of reflections from a planar surface. Real objects lying a certain distance from the reflecting plane have corresponding inverted images lying at an equal distance on the opposite (virtual) side of the plane. Figure 5.8 illustrates this and coordinate conventions used throughout this development. A right handed coordinate system, **m**, having axes (X, Y, Z), is attached to the real mirror surface such that X and Y are coplanar with the mirror surface and Z is normal to the surface.

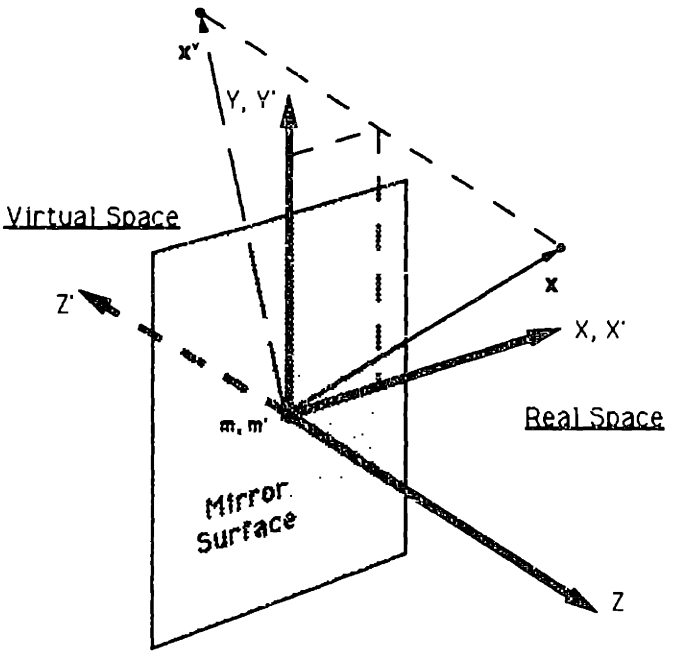


Figure 5.8: Planar mirror reflection transformation

A left handed coordinate system, $\mathbf{m}'$, having axes (X', Y', Z'), is defined as the reflection of (X, Y, Z) where X' and Y' are colinear with X and Y respectively and Z' is directed normal to the mirror surface into virtual space. At this stage of the development the origin of $\mathbf{m}$ and $\mathbf{m}'$ is arbitrarily fixed on the reflecting surface.

A vector $\mathbf{x}$ and its virtual image $\mathbf{x}^v$ may be represented in real mirror coordinates respectively as $\mathbf{x}_m$ and $\mathbf{x}^v_m$ or in virtual mirror coordinates as $\mathbf{x}_{m'}$ and $\mathbf{x}^v_{m'}$. The relationships between these vectors involve either vector reflections or coordinate basis reflections, in direct analogy to vector rotations or basis rotations commonly used in kinematic analyses. Note that vectors $\mathbf{x}_m$ and $\mathbf{x}^v_{m'}$ have identical components as do vectors $\mathbf{x}^v_m$ and $\mathbf{x}_{m'}$. Reflected vector pairs $\mathbf{x}_m$ and $\mathbf{x}^v_m$ or $\mathbf{x}_{m'}$ and $\mathbf{x}^v_{m'}$ differ only in the sign of their z components.

The common operator to both vector and basis reflections is the 3x3 *reflection matrix*,

R, defined as

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{5.1}$$

Note that R is *improper* orthogonal ($det|\mathbf{R}| = -1$) and is symmetric. This implies, along with other properties, that the inverse of R is identical with R. This can be seen intuitively in that the output vector of any vector multiplied by R will replicate the input vector with an inverted z component. This operation is invertible using R again on the output vector to reproduce the original input vector. This can be stated algebraically as

$$\mathbf{x} = \mathbf{R}[\mathbf{R}\mathbf{x}] \tag{5.2}$$

or simply

$$\mathbf{R}\mathbf{R} = \mathbf{I} \tag{5.3}$$

Vector reflections within a fixed coordinate basis may be represented as

$$\mathbf{x}_m = \mathbf{R}\mathbf{x}_m^v \tag{5.4}$$

or

$$\mathbf{x}_{m'} = \mathbf{R}\mathbf{x}_{m'}^v \tag{5.5}$$

Alternatively, a fixed vector may be represented in reflected coordinate bases as

$$\mathbf{x}_m = \mathbf{R}\mathbf{x}_{m'} \tag{5.6}$$

or

$$\mathbf{x}_m^v = \mathbf{R}\mathbf{x}_{m'}^v \tag{5.7}$$

Observe that these relationships are completely consistent with the properties of the real and reflected vectors mentioned above.

5.2.2 Mirror/Camera Transform Through a Fixed Planar Mirror

The transformation from object space (global) coordinates to camera coordinates is described in this section. In the specific situation being studied, a planar mirror reflects a virtual global image into the real camera's view. In visualizing two or more camera/mirror systems viewing the same global space it is more convenient (and intuitive) to think of the problem from the standpoint of the mirror generating a virtual *camera* image to view the real global coordinates. Along with other concepts it will be shown below that these two ways of looking at the problem are identical algebraically due to the invertible and reversible nature of the reflection matrix described in the previous section. This developement makes no assumptions regarding the precise alignment of camera optical components relative to the mirror or mirror rotational axis. The transformations developed are therefore general in nature.

In Figure 5.9 the real and virtual global coordinate systems, o and o', are shown along with the real and virtual camera coordinate systems, c and c'. Also shown are real and virtual mirror coordinates, m and m'.

Assuming the mirror is fixed, there exist constant transformations between the real camera, mirror and global coordinate frames. Employing a coordinate basis transformation and translation, any point, $\mathbf{x}_c$, defined with respect to the real camera may be represented as a point $\mathbf{x}_m$ with respect to the real mirror as

$$\mathbf{x}_m = \mathbf{M}_{cm}[\mathbf{x}_c - \mathbf{x}_{cm}] \tag{5.8}$$

where $\mathbf{M}_{cm}$ and $\mathbf{x}_{cm}$ are defined as a transformation matrix and position vector respec-

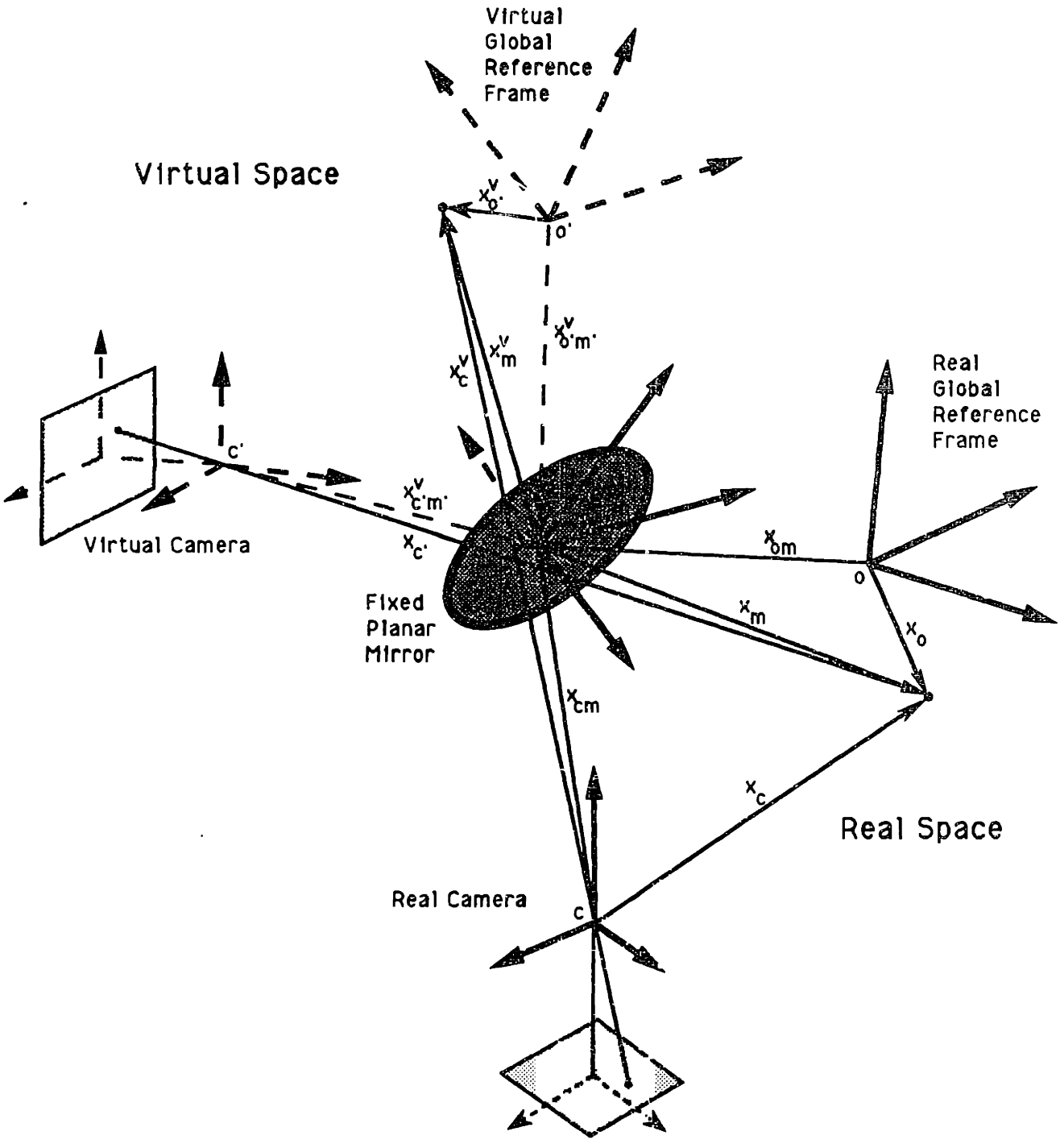


Figure 5.9: Camera/mirror transformations

tively of the real mirror relative to the real camera. Similarly, any point, $\mathbf{x}_o$, defined in the real global frame may be represented as a point $\mathbf{x}_m$ where

$$\mathbf{x}_m = \mathbf{M}_{om}[\mathbf{x}_o - \mathbf{x}_{om}] \tag{5.9}$$

and $\mathbf{M}_{om}$ and $\mathbf{x}_{om}$ are defined as a transformation matrix and position vector respectively of the real mirror relative to the real global frame. Combining the above two equations allows the relationship between the real camera and mirror to be formed as

$$\mathbf{x}_c = \mathbf{M}_{cm}^{-1}\mathbf{M}_{om}[\mathbf{x}_o - \mathbf{x}_{om}] + \mathbf{x}_{cm} \tag{5.10}$$

The real camera is, however, not viewing the real point $\mathbf{x}_c$. Therefore, equation 5.10 does not apply to data observed by the camera through the mirror. Instead, the camera is viewing a virtual point $\mathbf{x}_c^v$ as shown in Figure 5.9. Rewriting 5.8 for the virtual point as

$$\mathbf{x}_m^v = \mathbf{M}_{cm}[\mathbf{x}_c^v - \mathbf{x}_{cm}] \tag{5.11}$$

and combining with 5.9 and 5.4 gives the virtual point to real camera transformation as

$$\mathbf{x}_c^v = \mathbf{M}_{cm}^{-1}\mathbf{R}\mathbf{M}_{om}[\mathbf{x}_o - \mathbf{x}_{om}] + \mathbf{x}_{cm} \tag{5.12}$$

In the alternative derivation, similarly notated and algebraically identical [1] expressions exist for relating a point in virtual space to the virtual mirror, virtual global frame and virtual camera. These by analogy to equations 5.8 and 5.9 are

$$\mathbf{x}_{m'}^v = \mathbf{M}_{c'm'}^v[\mathbf{x}_{c'}^v - \mathbf{x}_{c'm'}^v] \tag{5.13}$$

[1]The manipulation of virtual vectors or virtual bases exclusively in virtual space involves operators identical to those used to perform similar manipulations to corresponding real vectors or bases exclusively in real space. It is only when relating vectors or bases across the real/virtual boundary that additional operators are required.

$$\mathbf{x}^{v}_{m'} = \mathbf{M}^{v}_{o'm'}[\mathbf{x}^{v}_{o'} - \mathbf{x}^{v}_{o'm'}] \tag{5.14}$$

In this case the virtual camera is viewing a real point and 5.13 must be rewritten as

$$\mathbf{x}_{m'} = \mathbf{M}^{v}_{c'm'}[\mathbf{x}_{c'} - \mathbf{x}^{v}_{c'm'}] \tag{5.15}$$

Combining this with 5.14 and 5.4 gives the real point to virtual camera transformation as

$$\mathbf{x}_{c'} = \mathbf{M}^{v}_{c'm'}{}^{-1}\mathbf{R}\mathbf{M}^{v}_{o'm'}[\mathbf{x}^{v}_{o'} - \mathbf{x}^{v}_{o'm'}] + \mathbf{x}^{v}_{c'm'} \tag{5.16}$$

By definition, all components on the right side of 5.16 are equivalent to corresponding right side components of 5.12. The two conceptual approaches are therefore identical algebraically. It is however most convenient to visualize the second conceptualization; that of a virtual camera viewing the real space. This is particularly evident when more than one virtual camera is used to reconstruct a point in 3D space. The concept of virtual cameras viewing a common real space allows for visualization of the intersecting colinearity rays. If the real camera viewing virtual space concept is used then colinearity rays are projected into independent virtual spaces and intersection occurs only in the algebra.

It is also convenient from this point forward to eliminate unnecessary notation when describing this transformation. Toward this end, notations of vectors or matrices that are equivalent in real and virtual space will be referred to using only their real space representations. As a result, equation 5.12 will be the fundamental equation governing the relationship between a real point in space being viewed by a virtual camera.

5.2.3 Mirror/Camera Transform Through Angularly Displaced Mirror

Thus far the reflecting surface in the camera/mirror model has been assumed to be fixed relative to the camera and global space. The goal of providing an enlarged measurement volume requires that the mirror be angularly displaced about a vertical axis, approximately aligned with the camera axis, so as to continuously point the virtual camera at the markers being measured. This development is general in that no assumptions pertaining to accurate mechanical alignments of the camera, mirror, or mirror axes have been made. This section will detail how this angular displacement is incorporated into the model described thus far.

This development begins by incorporating an axis of rotation for the mirror that fixes the position and orientation of the mirror coordinate system m on the mirror surface. Refering to Figure 5.10 an axis of rotation passing through the mirror at a fixed tilt angle (γ) defines the origin of mirror coordinates to be at the intersection of the axis and the reflecting surface. The z axis is parallel with the mirror normal directed into real space. Mirror rotation is defined to be positive in the clockwise direction when viewed from the camera position, or positive upward in the sense of a right hand screw. The cross-product of the positive mirror rotation unit vector and z_m defines the orientation of x_m. The axis y_m is then defined by the intersection of the mirror surface and the plane formed by the rotational axis and z_m.

A fixed *reference* orientation for the mirror is now defined arbitrarily with respect to the global and camera coordinate frames. The matrices $\mathbf{M}_{cm}$ and $\mathbf{M}_{om}$ and vectors

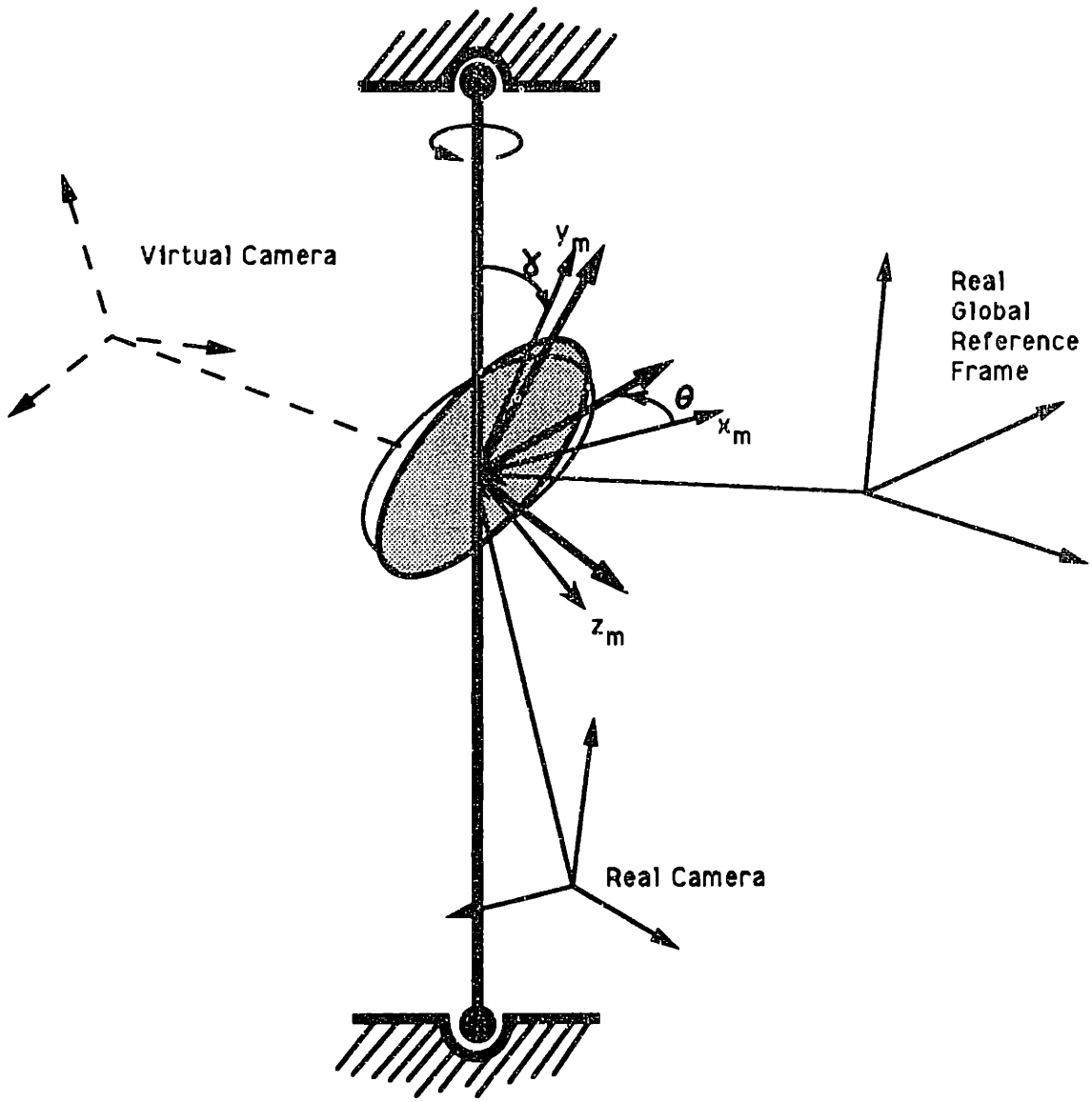


Figure 5.10: Model incorporating angular displacement of the mirror

$\mathbf{x}_{cm}$ and $\mathbf{x}_{om}$ of the previous section are now redefined to represent the transformation components from the camera or global frames to the mirror reference orientation and *not* to the actual current mirror orientation. Note that these matrices and vectors are constants for a particular camera/mirror setup in global space as long as the mirror reference orientation remains fixed.

Transforming from the reference orientation to the actual mirror orientation requires an additional coordinate system transformation corresponding to the angular displacement of the mirror rotation axis from the reference position. Note that since the rotation axis passes through the origin of mirror coordinates no translation of the mirror coordinate system is involved. An orthogonal transformation matrix is sufficient for describing the rotational displacement. This transformation matrix, $\mathbf{M}_\Theta$ is a function of the constant axis tilt angle, γ, and the axis rotation angle from reference, θ. It may be reduced to a series of three rotations about coordinate frame axes as shown in Figure 5.11 and described here:

1. From the reference orientation rotate through an angle $-\gamma$ about the reference x axis thereby aligning the reference y axis with the mirror rotation axis.

2. Next rotate about the current y axis through an angle θ corresponding to the measured current angular displacement of the rotational axis from the neutral position.

3. Finally, rotate through an angle γ about the current mirror x axis resulting in the actual mirror orientation.

By inspection steps 1 and 3 above are inverse operations and may be represented by

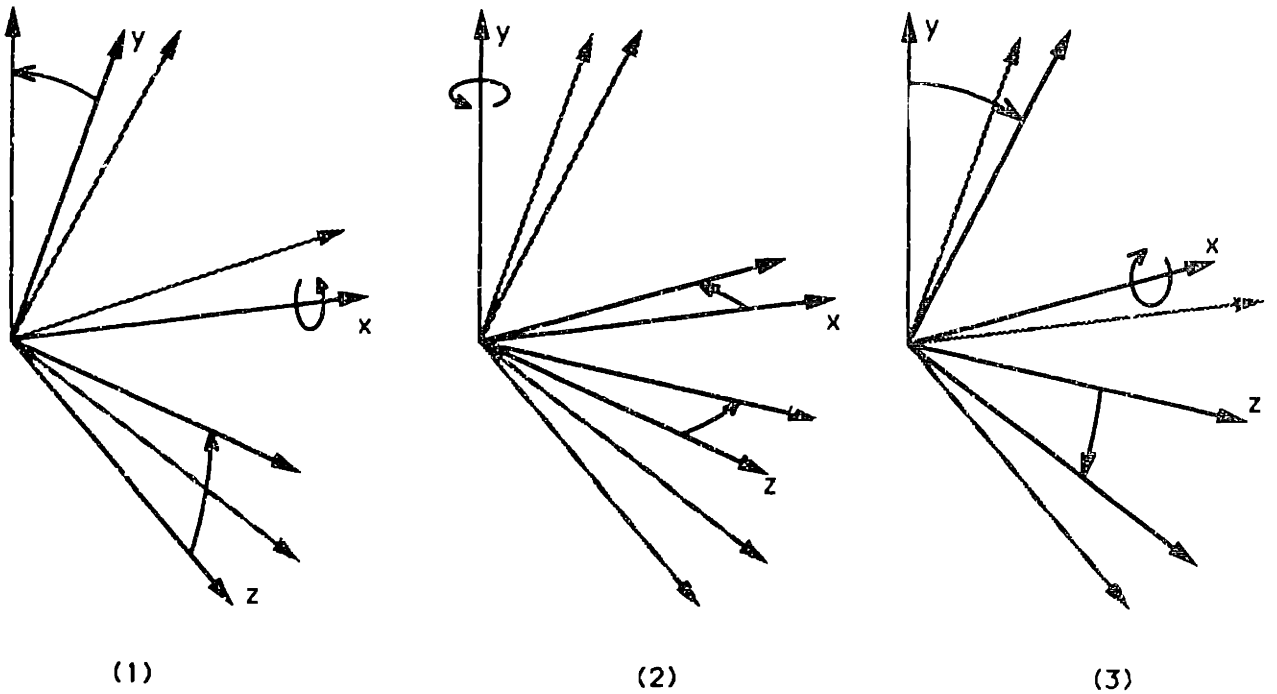


Figure 5.11: Three successive basis transformations comprising $\mathbf{M}_\Theta$. (1) negative γ rotation about initial x axis; (2) θ rotation about intermediate y axis; (3) positive γ rotation about final x axis

transformation matrices $\mathbf{M}_\gamma^{-1}$ and $\mathbf{M}_\gamma$ respectively where

$$\mathbf{M}_\gamma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\gamma) & \sin(\gamma) \\ 0 & -\sin(\gamma) & \cos(\gamma) \end{bmatrix} \tag{5.17}$$

Step 2 above may be represented by the transformation matrix $\mathbf{M}_\theta$ where

$$\mathbf{M}_\theta = \begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \tag{5.18}$$

The complete transformation, $\mathbf{M}_\Theta$, is then composed of a sequential transformation through $\mathbf{M}_\gamma^{-1}$, $\mathbf{M}_\theta$ and $\mathbf{M}_\gamma$ expressed algebraically as

$$\mathbf{M}_\Theta = \mathbf{M}_\gamma \mathbf{M}_\theta \mathbf{M}_\gamma^{-1} \tag{5.19}$$

The complete transformation matrix from camera (or global) coordinates to the current mirror coordinates is therefore now composed of two discrete transformation matrices: from camera (or global) to mirror reference coordinates $\mathbf{M}_{cm}$ (or $\mathbf{M}_{om}$), followed by

the transformation from mirror reference coordinates to the current mirror coordinates, $\mathbf{M}_\Theta$. Therefore if $\mathbf{M}_{cm}$ and $\mathbf{M}_{om}$ in 5.12 are replaced by $\mathbf{M}_\Theta\mathbf{M}_{cm}$ and $\mathbf{M}_\Theta\mathbf{M}_{om}$ respectively, the full transformation for a virtual camera viewing real global space through an arbitrary known mirror orientation may be represented as

$$\mathbf{x}_c^v = \mathbf{M}_{cm}^{-1}\mathbf{M}_\Theta^{-1}\mathbf{R}\mathbf{M}_\Theta\mathbf{M}_{om}[\mathbf{x}_o - \mathbf{x}_{om}] + \mathbf{x}_{cm} \tag{5.20}$$

The overall mirror transformation 5.20 may now be related to the theory presented in Chapter 2. If the matrix M and vector $\mathbf{x}_{oc}$ of 2.4 are replaced by

$$\mathbf{M} = \mathbf{M}_{cm}^{-1}\mathbf{M}_\Theta^{-1}\mathbf{R}\mathbf{M}_\Theta\mathbf{M}_{om} \tag{5.21}$$

and

$$\mathbf{x}_{oc} = \mathbf{M}_{cm}^{-1}\mathbf{M}_\Theta^{-1}\mathbf{R}\mathbf{M}_\Theta\mathbf{M}_{om}\mathbf{x}_{om} - \mathbf{x}_{cm} \tag{5.22}$$

then the colinearity equations 2.8 and 2.9 may be generated and applied as presented previously. The difference now is that M and $\mathbf{x}_{oc}$ are no longer constants and will change according to 5.21 and 5.22 as the mirror rotation angle θ and, therefore, $\mathbf{M}_\Theta$ change. There are, however, constant terms required for generating the revised M and $\mathbf{x}_{oc}$. These terms, $\mathbf{M}_{cm}$, $\mathbf{M}_{om}$, $\mathbf{x}_{cm}$, and $\mathbf{x}_{om}$, do not change for a specific camera/mirror setup in global space and must be accurately identified before 3D reconstructions can be performed. This problem falls under the heading of LVT Calibration and is addressed in the next section.

5.3 LVT Calibration

In Chapter 4 the concept of calibrating a fixed camera photogrammetric system was introduced and developed. This section will extend the concepts developed there through

application to the problem of LVT calibration.

An important point to note with LVT is that if mirrors in the two camera/mirror systems remain stationary (ie. the virtual camera remains fixed), then calibration and data acqusition may be performed exactly as presented in Chapter 4 with no changes in the algorithms or software. The transformation matrix M will be improper orthogonal as it includes the reflection matrix **R** in its makeup (5.21) but this poses no constraints on any of the algorithms presented previously. The polar decomposition technique used in Chapter 4 will perform equally well decomposing a nearly improper orthogonal matrix to the product of a symmetric matrix and an improper orthogonal matrix.

Two approaches to LVT calibration are presented in this section. Both make use of a series of fixed mirror (virtual camera) calibrations obtained at various mirror axis positions in an attempt to extract the constant terms $\mathbf{M}_{cm}$, $\mathbf{M}_{om}$, $\mathbf{x}_{cm}$, and $\mathbf{x}_{om}$ of 5.20 from the resultant M's and $\mathbf{x}_{oc}$'s. The first approach determines a solution for the constants using a direct closed-form algebraic approach. Problems encountered with implementation of this approach will be discussed in Chapter 6. The second approach employs a numerical iterative technique to identify the constants. This approach was implemented successfully but also has problems which will be discussed in Chapter 6.

Initially it may seem unnecessary to require an experimental method for obtaining the camera to mirror transformation. Why not design these transformations into the system mounting apparatus and be done with it? The problem with such an approach is that the camera center of projection and principal axis would require identification relative to some externally defined set of coordinates fixed in the camera body. Internal calibration

makes no provision for such definitions. The center of projection and principal axis move with the camera but are not explicitly referenced to a physical camera-based set of landmarks. The attractive aspects of a full calibration are that it allows for misalignments during camera/mirror relative mounting and is elegant overall.

5.3.1 Direct Closed-Form LVT Calibration

In this approach standard calibrations are performed as presented in Chapter 4. As with the standard calibration, each of the two camera/mirror systems is calibrated independently. The following discussion will refer to only one camera/mirror system with the understanding that the other system (and all such systems) would be handled in a similar manner.

The calibration structure defines a single fixed global coordinate frame for all calibrations. The mirrors are moved to a series of three positions from which the virtual camera views the calibration structure. At each virtual camera position a full calibration is performed resulting in three unique M's ($\mathbf{M}_A, \mathbf{M}_B, \mathbf{M}_C$) and three unique $\mathbf{x}_{oc}$'s ($\mathbf{x}_A, \mathbf{x}_B, \mathbf{x}_C$). At this point the solution technique for $\mathbf{M}_{cm}$ and $\mathbf{M}_{om}$ proceeds separately from the solution technique for $\mathbf{x}_{cm}$ and $\mathbf{x}_{om}$.

Solving for $\mathbf{M}_{cm}$ and $\mathbf{M}_{om}$

If the three matrices $\mathbf{M}_A$, $\mathbf{M}_B$ and $\mathbf{M}_C$ are expanded according to 5.21

$$\mathbf{M}_A = \mathbf{M}_{cm}^{-1}\mathbf{M}_{\Theta_A}^{-1}\mathbf{R}\mathbf{M}_{\Theta_A}\mathbf{M}_{om} \tag{5.23}$$

$$\mathbf{M}_B = \mathbf{M}_{cm}^{-1}\mathbf{M}_{\Theta_B}^{-1}\mathbf{R}\mathbf{M}_{\Theta_B}\mathbf{M}_{om} \tag{5.24}$$

$$\mathbf{M}_C = \mathbf{M}_{cm}^{-1}\mathbf{M}_{\Theta_C}^{-1}\mathbf{R}\mathbf{M}_{\Theta_C}\mathbf{M}_{cm} \tag{5.25}$$

then relative matrix transformations $\mathbf{M}_{AB} = \mathbf{M}_A\mathbf{M}_B^{-1}$ and $\mathbf{M}_{BC} = \mathbf{M}_B\mathbf{M}_C^{-1}$ may be formed eliminating one of the unknown components $\mathbf{M}_{cm}$ or $\mathbf{M}_{om}$. Choosing arbitrarily to eliminate $\mathbf{M}_{om}$, the relative transformation matrices become

$$\mathbf{M}_{AB} = \mathbf{M}_{cm}^{-1}\mathbf{M}_{\Theta_{AB}}\mathbf{M}_{cm} \tag{5.26}$$

$$\mathbf{M}_{BC} = \mathbf{M}_{cm}^{-1}\mathbf{M}_{\Theta_{BC}}\mathbf{M}_{cm} \tag{5.27}$$

where $\mathbf{M}_{\Theta_{AB}}$ and $\mathbf{M}_{\Theta_{BC}}$ are composed of known mirror axis position elements in the form

$$\mathbf{M}_{\Theta_{AB}} = \mathbf{M}_{\Theta_A}^{-1}\mathbf{R}\mathbf{M}_{\Theta_A}\mathbf{M}_{\Theta_B}^{-1}\mathbf{R}\mathbf{M}_{\Theta_B} \tag{5.28}$$

$$\mathbf{M}_{\Theta_{BC}} = \mathbf{M}_{\Theta_B}^{-1}\mathbf{R}\mathbf{M}_{\Theta_B}\mathbf{M}_{\Theta_C}^{-1}\mathbf{R}\mathbf{M}_{\Theta_C} \tag{5.29}$$

Notice equations 5.26 and 5.27 take the form of similarity transforms with orthogonal component matrices. The orthogonal transform matrix in both cases, $\mathbf{M}_{cm}$, is the unknown component being sought. Similarity transforms involving orthogonal components have unique properties beyond those normally recognized. Some of these properties are reviewed here.

A general property of similarity transforms is that the eigenvalues of the transformed matrices remain invariant [40]. Applied to 5.26 and 5.27 the eigenvalues of $\mathbf{M}_{\Theta_{AB}}$ ($Eig\mathbf{M}_{\Theta_{AB}}$) are equal to $Eig\mathbf{M}_{AB}$ and $Eig\mathbf{M}_{\Theta_{BC}} = Eig\mathbf{M}_{BC}$. In the specific case of 3×3 orthogonal rotation matrices the three eigenvalues are generally unique. One eigenvalue is always equal to $+1$ and the other two are complex conjugates of form $(e^{i\psi}, e^{-i\psi})$,

where ψ is the angle of rotation [2] represented by the transformed matrices [12]. This implies the total angle of rotation for the transformed matrices must also be invariant. The total angle of rotation represented by a 3×3 orthogonal rotation matrix M obeys the relationship

$$\mathbf{Trace(M)} = \mathbf{1} + \mathbf{2}\cos(\psi) \tag{5.30}$$

This has interesting implications for the problem at hand as it provides a means to check the transformed matrix pairs of 5.26 and 5.27 for identical total rotations. Since each matrix of the pair is estimated independently this can serve as an accuracy check.

A second property specific to similarity transforms involving 3×3 orthogonal rotation matrices is that the similarity transform has the property of rotating the unit rotation vector of the initial matrix into the unit rotation vector of the transformed matrix. This may be stated algebraically with reference to 5.26 as

$$\mathbf{M}_{cm}\mathbf{v}_{AB} = \mathbf{v}_{\Theta_{AB}} \tag{5.31}$$

where $\mathbf{v}_{AB}$ and $\mathbf{v}_{\Theta_{AB}}$ are the unit direction cosine vectors of $\mathbf{M}_{AB}$ and $\mathbf{M}_{\Theta_{AB}}$ respectively. Similarly

$$\mathbf{M}_{cm}\mathbf{v}_{BC} = \mathbf{v}_{\Theta_{BC}} \tag{5.32}$$

holds for 5.27.

The problem is now reduced to determining a unique rotation matrix $\mathbf{M}_{cm}$ which

[2] Any orthogonal transformation matrix M may be represented in terms of a rotation ψ about a directed line segment having unit vector representation $(\ell, m, n)^T$ such that [8]

$$\mathbf{M} = \cos(\psi)\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + (1-\cos(\psi))\begin{bmatrix} \ell^2 & \ell m & \ell n \\ \ell m & m^2 & mn \\ \ell n & mn & n^2 \end{bmatrix} + \sin(\psi)\begin{bmatrix} 0 & n & -m \\ -n & 0 & \ell \\ m & -\ell & 0 \end{bmatrix}$$

satisfies the constraints imposed by 5.31 and 5.32. The solution to either 5.31 or 5.32 defines a family of possible rotation vector directions which lie in a common plane (see Figure 5.12). The magnitude of the rotation will vary depending on the rotation direction selected. The smallest possible rotation of $\mathbf{v}$ into $\mathbf{v}_\Theta$ is about an axis normal to the plane including $\mathbf{v}$ and $\mathbf{v}_\Theta$ and has a magnitude of $(\mathbf{v} \times \mathbf{v}_\Theta)$. The largest possible rotation occurs about the same axis in the opposite direction. An additional possible rotation direction lies in the plane formed by the rotated vectors and is directed along the sum of the two vectors. The magnitude of this rotation is π radians. Since each solution vector direction is paired with an oppositely directed solution vector then limiting admissible solutions to only those directions having magnitudes less than or equal to π radians provides a unique solution for any rotation axis. The range of magnitues for the admissible solution vectors describes a half ellipse as shown in Figure 5.12. Combining the requirements of both 5.31 and 5.32 is accomplished by determining the direction and magnitude of the line (axis) of intersection of these two planar ellipses.

Defining for the first relationship, 5.31, a unit vector, $\hat{\mathbf{p}}_{AB}$, along the major axis of the ellipse as

$$\hat{\mathbf{p}}_{AB} = \frac{\mathbf{v}_{AB} + \mathbf{v}_{\Theta_{AB}}}{| \mathbf{v}_{AB} + \mathbf{v}_{\Theta_{AB}} |} \tag{5.33}$$

and a unit vector, $\hat{\mathbf{q}}_{AB}$, along the minor axis as

$$\hat{\mathbf{q}}_{AB} = \frac{\mathbf{v}_{AB} \times \mathbf{v}_{\Theta_{AB}}}{| \mathbf{v}_{AB} \times \mathbf{v}_{\Theta_{AB}} |} \tag{5.34}$$

the unit normal to the ellipse, $\hat{\mathbf{n}}_{AB}$, may then be defined as

$$\hat{\mathbf{n}}_{AB} = \hat{\mathbf{p}}_{AB} \times \hat{\mathbf{q}}_{AB} \tag{5.35}$$

Figure 5.12: Plane of possible vectors representing a rotation of v into v_Θ

Similarly for 5.32 a unit normal to the ellipse, $\hat{\mathbf{n}}_{BC}$, may be defined as [3]

$$\hat{\mathbf{n}}_{BC} = \hat{\mathbf{p}}_{BC} \times \hat{\mathbf{q}}_{BC} \tag{5.36}$$

The direction of the line of intersection of these two planes is computed from the cross-product of the two unit normal vectors. A unit vector, $\hat{\mathbf{r}}$, directed along this line of intersection may be obtained from

$$\hat{\mathbf{r}} = \frac{\mathbf{n}_{AB} \times \mathbf{n}_{BC}}{|\, \mathbf{n}_{AB} \times \mathbf{n}_{BC} \,|} \tag{5.37}$$

At this point the direction of $\hat{\mathbf{r}}$ has been arbitrarily defined based on the order of taking the cross-product in 5.37. We are interested in the unit rotation vector that is within the admissible region of the ellipse. For this reason the unit intersection vector is defined to be positive such that its projection on $\hat{\mathbf{q}}_{AB}$ and $\hat{\mathbf{q}}_{BC}$ is in the admissible region of the ellipse.

The magnitude of the rotation represented by $\hat{\mathbf{r}}$ may now be computed from either of the original two planes. The cross products of each of $\mathbf{v}$ and $\mathbf{v}_\Theta$ with $\hat{\mathbf{r}}$ define vectors in the rotation plane of $\hat{\mathbf{r}}$. The angle between these vectors is the rotation angle of interest. This is shown in Figure 5.3.1. The rotation angle, ψ, may be computed from

$$\sin(\psi) = \left| \left[\frac{\mathbf{v}_{AB} \times \hat{\mathbf{r}}}{|\, \mathbf{v}_{AB} \times \hat{\mathbf{r}} \,|} \right] \times \left[\frac{\mathbf{v}_{\Theta_{AB}} \times \hat{\mathbf{r}}}{|\, \mathbf{v}_{\Theta_{AB}} \times \hat{\mathbf{r}} \,|} \right] \right| \tag{5.38}$$

[3] A alternative derivation for $\hat{\mathbf{n}}_{AB}$ and $\hat{\mathbf{n}}_{BC}$ is simply [50]

$$\hat{\mathbf{n}}_{AB} = \frac{\mathbf{v}_{AB} - \mathbf{v}_{\Theta_{AB}}}{|\, \mathbf{v}_{AB} - \mathbf{v}_{\Theta_{AB}} \,|}$$

and

$$\hat{\mathbf{n}}_{BC} = \frac{\mathbf{v}_{BC} - \mathbf{v}_{\Theta_{BC}}}{|\, \mathbf{v}_{BC} - \mathbf{v}_{\Theta_{BC}} \,|}$$

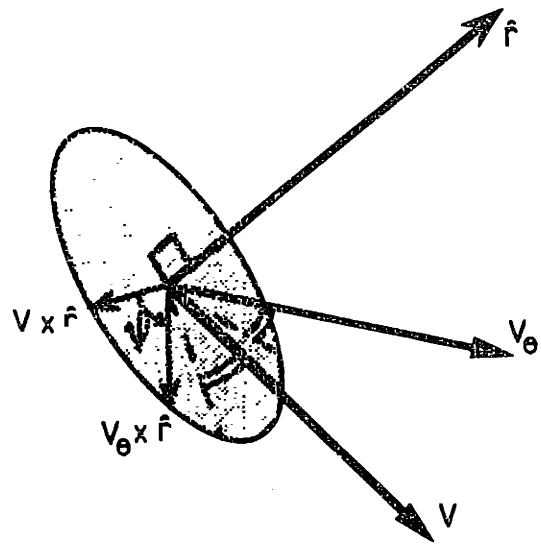


Figure 5.13: Estimating the rotation magnitude of $\hat{\mathbf{r}}$

The unit vector $\hat{\mathbf{r}} = (\ell, m, n)^T$ and rotation ψ may now be used to contruct the required rotation matrix $\mathbf{M}_{cm}$ (see footnote 2 of this Chapter).

An alternative approach to solving for $\mathbf{M}_{cm}$ has been suggested [50]. Returning to Equations 5.31 and 5.32 we may define matrices $\mathbf{V} = \left[\mathbf{v}_{AB} \vdots \mathbf{v}_{BC}\right]$ and $\mathbf{V}_\Theta = \left[\mathbf{v}_{\Theta_{AB}} \vdots \mathbf{v}_{\Theta_{BC}}\right]$ such that $\mathbf{M}_{cm}\mathbf{V} = \mathbf{V}_\Theta$ and $\mathbf{M}_{cm}\mathbf{V}\mathbf{V}^T = \mathbf{V}_\Theta\mathbf{V}^T$. Since $\mathbf{V}\mathbf{V}^T$ is symmetric and $\mathbf{M}_{cm}$ is orthogonal then this represents a polar decomposition of $\mathbf{V}_\Theta\mathbf{V}^T$.

With $\mathbf{M}_{cm}$ identified it is now straightforward to proceed to identify $\mathbf{M}_{om}$. Using the resulting transfomation matrix $\mathbf{M}$ from a standard 6 DOF calibration with the mirror at a known position and rearranging 5.21 yields the required solution:

$$\mathbf{M}_{om} = \mathbf{M}_\Theta^{-1}\mathbf{R}\mathbf{M}_\Theta\mathbf{M}_{cm}\mathbf{M} \tag{5.39}$$

With $\mathbf{M}_{cm}$ and $\mathbf{M}_{om}$ known it is now possible to generate the required overall transformation matrix $\mathbf{M}$ for any known position of the mirror using 5.21. The full transformation

also includes identifying the translation vector $\mathbf{x}_{oc}$ at any mirror position according to 5.22. This problem is addressed next.

Solving for $\mathbf{x}_{cm}$ and $\mathbf{x}_{om}$

The approach used in solving for $\mathbf{x}_{cm}$ and $\mathbf{x}_{om}$ recalls the form of the colinearity equations 2.8 and 2.9. Substituting 5.21 into 5.20 and following an approach identical to the colinearity equation derivation presented in Chapter 2 the following modified colinearity equations are formed:

$$u' = \frac{\bar{u}}{d} = \frac{m_{11}(x_o - x_{mo}) + m_{12}(y_o - y_{mo}) + m_{13}(z_o - z_{mo}) + x_{mc}}{m_{31}(x_o - x_{mo}) + m_{32}(y_o - y_{mo}) + m_{33}(z_o - z_{mo}) + z_{mc}} \tag{5.40}$$

and

$$v' = \frac{\bar{v}}{d} = \frac{m_{21}(x_o - x_{mo}) + m_{22}(y_o - y_{mo}) + m_{23}(z_o - z_{mo}) + y_{mc}}{m_{31}(x_o - x_{mo}) + m_{32}(y_o - y_{mo}) + m_{33}(z_o - z_{mo}) + z_{mc}} \tag{5.41}$$

This is similar to the general problem of external camera calibration except that the elements of M have been previously determined. Remaining to be determined are the components of $\mathbf{x}_{mc}$ and $\mathbf{x}_{mo}$. This may be done by rearranging 5.40 and 5.41 to form two linear equations in the six unknown components $(x_{mc}, y_{mc}, z_{mc})^T$ and $(x_{mo}, y_{mo}, z_{mo})^T$.

The method is then reduced to performing standard calibrations at a series of mirror positions to generate overall transformation matrices M and camera data (u', v'). Components of $\mathbf{x}_{mc}$ and $\mathbf{x}_{mo}$ are solved for using a least squares approach based on the matrix

equation

$$\begin{bmatrix} (m_{11}-m_{31}u_i')_j & (m_{12}-m_{32}u_i')_j & (m_{13}-m_{33}u_i')_j & -1 & 0 & (u_i')_j \\ (m_{21}-m_{31}v_i')_j & (m_{22}-m_{32}v_i')_j & (m_{23}-m_{33}v_i')_j & 0 & -1 & (v_i')_j \\ \vdots & & & & & \vdots \end{bmatrix} \begin{bmatrix} x_{mo} \\ y_{mo} \\ z_{mo} \\ x_{mc} \\ y_{mc} \\ z_{mc} \end{bmatrix} = \begin{bmatrix} b_{ij}^u \\ b_{ij}^v \\ \vdots \end{bmatrix}$$

incorporating data from the ith marker at the jth mirror position where

$$b_{ij}^u = (m_{11}-m_{31}u_i')_j x_{oi} + (m_{12}-m_{32}u_i')_j y_{oi} + (m_{13}-m_{33}u_i')_j z_{oi}$$

and

$$b_{ij}^v = (m_{21}-m_{31}v_i')_j x_{oi} + (m_{22}-m_{32}v_i')_j y_{oi} + (m_{23}-m_{33}v_i')_j z_{oi}$$

5.3.2 Numerical LVT Calibration

A second technique for identifying the LVT calibration elements employs an extension of the basic method developed by Fijan [10] to generate a relative external calibration for a two camera system based on minimizing a skew ray 3D reconstruction error function. The extended technique is used here to identify the constant camera to mirror transformation. Following this, the global to mirror transformation is derived in a similar manner to the direct method presented previously.

The development begins by assuming a pair of virtual camera/mirror systems as shown in Figure 5.14. The mirror coordinates m_1 and m_2 are shown in their neutral orientations (dashed vector triads) and arbitrarily rotated orientations (solid vector triads). The virtual cameras generated at the arbitrary mirror positions are represented by the triads c_1 and c_2.

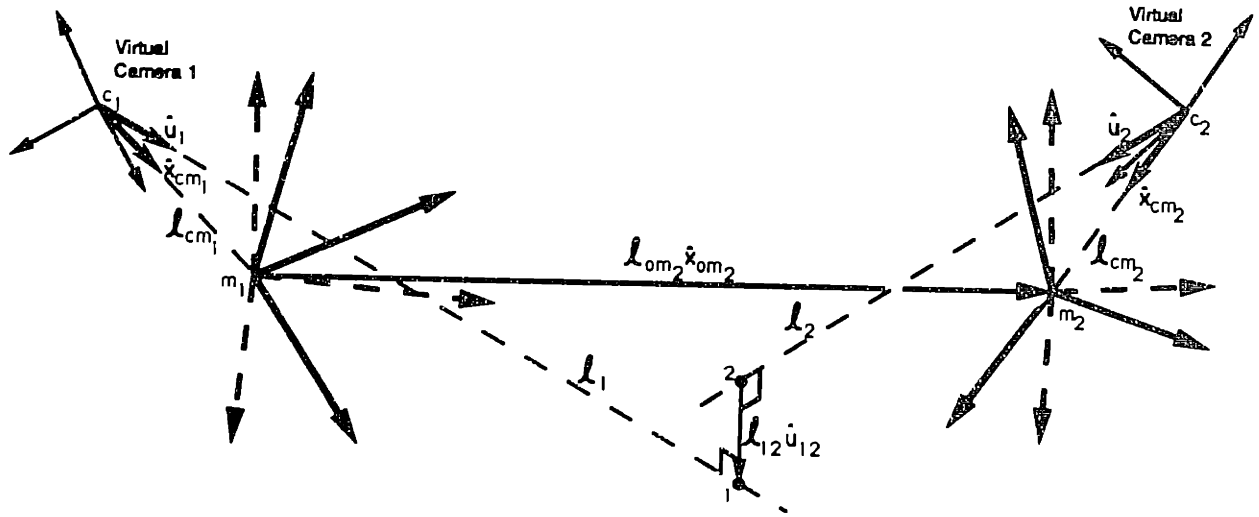


Figure 5.14: Arrangement of virtual camera/mirror system for numerical calibration

Keeping in mind that the transformation from each camera to its respective neutrally positioned mirror is independent of choice of global coordinates the global coordinates for this development will be defined to be identical with the neutral position and orientation of mirror 1. This gives an overall transformation equation from 5.20 for camera/mirror 1 as

$$\mathbf{x}^v_{c_1} = \mathbf{M}_1[\mathbf{x}_{o_1}] + \mathbf{x}_{cm_1} \tag{5.42}$$

and for camera/mirror 2 as

$$\mathbf{x}^v_{c_2} = \mathbf{M}_2[\mathbf{x}_{o_2} - \mathbf{x}_{cm_2}] + \mathbf{x}_{cm_2} \tag{5.43}$$

where

$$\mathbf{M}_1 = \mathbf{M}^{-1}_{cm_1}\mathbf{M}^{-1}_{\Theta_1}\mathbf{R}\mathbf{M}_{\Theta_1}$$

and

$$\mathbf{M}_2 = \mathbf{M}^{-1}_{cm_2}\mathbf{M}^{-1}_{\Theta_2}\mathbf{R}\mathbf{M}_{\Theta_2}\mathbf{M}_{cm_2}$$

Note that the above equations do not yet assume the two camera/mirror systems are observing the same globally defined point.

Redefining components of 5.42 and 5.43 with respect to global coordinates and introducing a unit vector notation (see Figure 5.14) gives the following relations:

$$\ell_{cm_i}\hat{\mathbf{x}}_{cm_i} = \mathbf{M}_i^{-1}\mathbf{x}_{mc_i}$$

$$\ell_{cm_i}\hat{\mathbf{u}}_i = \mathbf{M}_i^{-1}\mathbf{x}_{c_i}^{v}$$

for each camera/mirror system i.

The general 3D point reconstruction problem involves determining where in global space the colinearity rays intersect. Here the colinearity rays are defined along unit vectors $\hat{\mathbf{u}}_i, (i = 1, 2)$. As explained in Chapter 2, in general these rays do not intersect. The task here is to find the camera/mirror systems' position and orientation that minimizes the amount of skew in the rays constructed from the points observed.

The skew in each ray pair is equivalent to the length of the common perpendicular to each ray. A unit vector, $\hat{\mathbf{u}}_{12}$, defining the orientation of the common perpendicular is defined from the cross-product of the colinearity unit vectors as

$$\hat{\mathbf{u}}_{12} = \frac{\hat{\mathbf{u}}_1 \times \hat{\mathbf{u}}_2}{|\,\hat{\mathbf{u}}_1 \times \hat{\mathbf{u}}_2|} \tag{5.44}$$

Refering to Figure 5.14 and summing vectors results in the following vector relation:

$$\ell_1\hat{\mathbf{u}}_1 = \ell_{cm_1}\hat{\mathbf{x}}_{cm_1} + \ell_{om_2}\hat{\mathbf{x}}_{om_2} - \ell_{cm_2}\hat{\mathbf{x}}_{cm_2} + \ell_2\hat{\mathbf{u}}_2 + \ell_{12}\hat{\mathbf{u}}_{12} \tag{5.45}$$

Taking the dot product of all terms with respect to $\hat{\mathbf{u}}_{12}$ and rearranging gives a nondimensional representation for the skew error, e as

$$e = \frac{\ell_{12}}{\ell_{om_2}} = \left[\left(\frac{\ell_{cm_2}}{\ell_{om_2}}\right)\hat{\mathbf{x}}_{cm_2} - \hat{\mathbf{x}}_{om_2} - \left(\frac{\ell_{cm_1}}{\ell_{om_2}}\right)\hat{\mathbf{x}}_{cm_1}\right]\cdot\hat{\mathbf{u}}_{12} \tag{5.46}$$

Note that e may be negative due to the dot product on the right of the above equation. For this reason a numerical routine that minimizes the RMS value of e over all observed markers must be implemented.

A total of 18 unknown parameters may be deduced in this manner: the position and orientation of each camera relative to its respective mirror accounts for 12 parameters while the position and orientation of camera/mirror system 2 relative to global coordinates (camera/mirror system 1) accounts for the additional 6 parameters. The position vector for camera/mirror system 2 relative to global coordinates becomes a unit vector in the minimized function. As a unit vector may be represented by only 2 independent parameters (provided a nominal direction may be assumed to assign the sign), then the full minimization is only over 17 independent parameters.

The 18 unknown parameters are known only to within a scaling factor. The scale factor, ℓ_{om_2}, may be deduced by incorporating known physical displacements present in the input data. Rearranging 5.45 and using the results of the minimization allows estimation of scaled representations of ℓ_1, ℓ_{12} and ℓ_2 for each marker. An estimate for a scaled 3D position, $\mathbf{p}_{s_i}$, of each observed marker i is obtained from

$$\mathbf{p}_{s_i} = \left(\frac{\ell_{1_i}}{\ell_{om_2}}\right)\hat{\mathbf{u}}_{1_i} - \left(\frac{\ell_{cm_1}}{\ell_{om_2}}\right)\hat{\mathbf{x}}_{cm_1} - \frac{1}{2}e_i\hat{\mathbf{u}}_{12_i} \tag{5.47}$$

The real displacement between a pair of markers divided by the scaled displacement between the same pair of markers yields an estimate for the scale factor. The mean of all such estimates over all marker pairs gives the final estimate for ℓ_{om_2} and a unique solution for the full 18 parameter problem.

Chapter 6

LVT Implementation and Testing

This chapter begins by detailing LVT implementation from the standpoint of hardware and software components. Issues of mirror system design and rotational position control are addressed. This leads into actual testing of the system beginning with calibration. Issues of calibration sensitivities are discussed and suggestions for future improvements are made. A static test of system absolute accuracy is also made.

6.1 LVT Hardware and Control Design and Implementation

6.1.1 Mirror Drive and Angular Measurement System

This section presents details of the LVT mirror and positioning hardware design. A photograph of the constructed system is presented in Figure 6.1. Additional hardware information pertaining to overall LVT system operations is presented in Appendix C.

Design and selection of a mirror drive/measurement system involved several interrelated issues. During data acquisition the mirror position must be known to sufficient accuracy for minimum error in 3D estimations. The angular displacement of the mirror

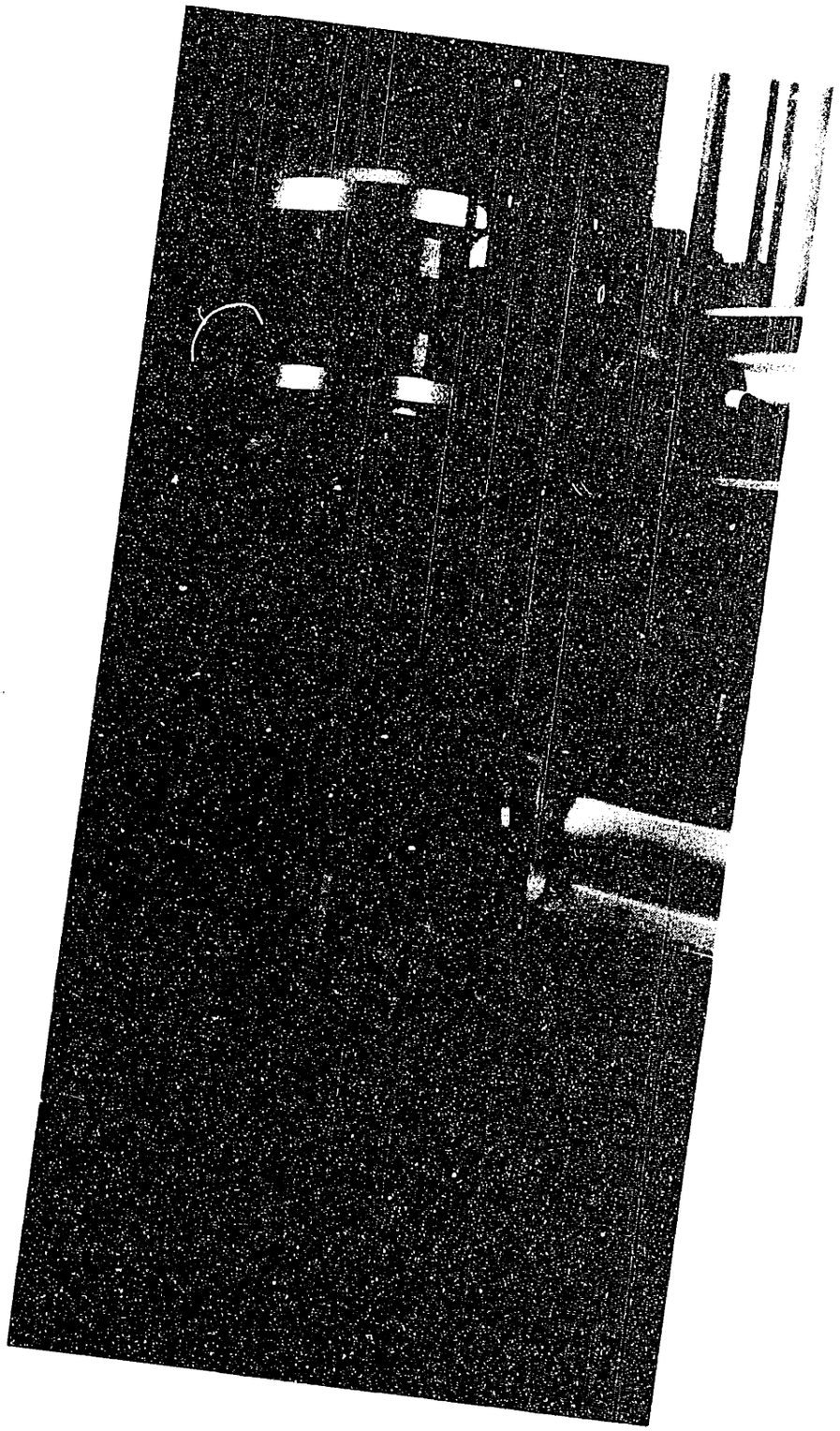


Figure 6.1: Photograph of mirror/motor mounting apparatus

from the neutral position must be measured accurately for the duration of SELSPOT data acquisition. It must also be possible to assume the rotational axis position and orientation is fixed. The drive system must be capable of rotating the mirror smoothly with sufficient acceleration to meet the forseeable demands of motion studies in the Newman Laboratory. Maintaining high accuracy angular position measurements in the presence of high angular accelerations requires that the angular measurement system be linked with the drive system with a maximum torsional stiffness. The issues of maintainability, reliability, and cost also factored into the design.

The basic unit of measurement with SELSPOT is called (in the Newman Lab) a SELSPOT unit (SU). Each SU is equivalent to one bit in 12, 1 part in 4096, or 0.025% of the camera field across each of the u and v camera coordinate directions. The camera lens provides an approximate 30 degree viewing angle. This implies a physical angular resolution for the camera of 0.00025×30 degrees or about 27 arc-seconds. Since the mirror drive rotates the virtual camera principal axis with the potential of adding error it was decided that the angular measurement system must be at least as accurate as the angular resolution of the camera and preferably an order of magnitude more accurate.

Acceleration demands of the system were estimated from envisioning the mirrors tracking a world class sprinter (approx. 10 meters/second) passing the two camera/mirror systems at a range of 3 meters. The maximum mirror angular acceleration required for constant tracking under these conditions is approximately 7 radians/second^2. As an additional safety factor and to be sure of adequate performance under all operating conditions, particularly during startup and mirrors locking on target, it was desired to have the system capable of rotating from stopped position at 0 degrees to stopped

position at 90 degrees in 0.5 seconds. This implies an angular acceleration requirement of approximately 25 radians/second^2.

Preliminary design of the mirror and mount apparatus predicted the rotated load moment of inertia to be approximately 5.9×10^{-3} kg-m^2. This inertia added to the inertia of the selected drive rotor is the total rotating inertia of the system.

After reviewing a range of rotary encoders and motors a decision was made to use a Yokogawa Corporation model DM1015B direct drive servo actuator system called *Dynaserv*^{TM}. The Dynaserv package includes a direct drive DC servo motor with an integral high resolution encoder. The maximum continuous output torque rating for the DM1015B of 15 N-m, rotor inertia of 12×10^{-3} kg-m^2 and maximum rated speed of 2 revolutions per second easily meets the stated requirements.

The encoder resolution of 655,360 pulses per revolution with a measurement repeatability (precision) of $\pm$ 2 arc-sec and measurment accuracy to $\pm$30 arc-sec is on the borderline of acceptability without calibration. The manufacturer offers calibration services to those customers requiring measurement accuracy equivalent to the precision. This would provide angular measurement accuracy well within the design specifications and could be done at a later date if required.

The motor employs pre-loaded crossed-roller bearings permitting large axial and radial loading (not required for this application). The manufacturer provided an estimate for maximum radial runout of 7 microns with no load based on their measurements. This combined with a rated radial stiffness of 1.0×10^{-4} rad/N-m ensures a very stable axis of rotation may be assumed.

The Dynaserve package also includes a servo driver power supply and control circuitry permitting closed-loop speed control with an analog ±6 Volt signal. The price of this system was substantially less than the combined price of separate items of comparable performance.

6.1.2 Mirror Surface and Optical Substrate

Planar mirrors are specified according to optical substrate material, flatness of surface finish, type of reflective coating and overall dimensions. The dimensions were defined by geometric constraints of the application and standard available components. To permit the virtual camera a full view without obstruction from the real camera lens (see Figure 5.7) the minimum length required along the mirror y axis was approximately 12.7 centimeters. A smaller minimum length along the x axis would be adequate implying an elliptically shaped mirror. As elliptical mirrors of that size are not standard components a circular mirror was considered the best option. To permit mirror mounting and optical component misalignment a 15.24 centimeter circular mirror was selected.

A common mirror component material in laboratory applications is Pyrex which provides a stable base for grinding and is relatively inexpensive. To investigate bending of a pyrex mirror under dynamic loading a 2.54 centimeter thick, 15.24 centimeter diameter pyrex optical flat angularly accelerated at 25 radians/second/second about a radial axis was estimated to undergo a maximum total tangential deflection of 125 nanometers. This corresponds to a mean angular deflection over the 7.62 centimeter radius of 0.34 arc-seconds. This was considered negligible since the actual axis of rotation is 45 degrees to the mirror plane normal yielding a much stiffer system.

Ealing Electro-Optics provides 15.25 centimeter diameter pyrex optical flats ground to $\lambda/10$ flatness using a reference wavelength, λ, of 632 nanometers. Since the lens is mounted close to the mirror and markers are typically several meters from the mirror, light entering the lens is reflected from an area roughly the size of the projected lens aperture on the mirror. In worst case scenario, in which a relative high point of $\lambda/10$ existed a distance of 20 millimeters from a relative low point of $-\lambda/10$, the mean slope of 1.3 arc-sec between extremes was deemed insignificant. An overcoated gold mirror surface finish was selected for its combined durability and infra-red reflectance of 95%.

6.1.3 Mirror Mount

Critical design considerations of the mirror mounting apparatus were that it be stiff in transmission of torque from the motor to the mirror and that it fixate the mirror without applying twisting stresses that might result in mirror warping. The mirror mount design drawings are presented in Appendix E. The design features a 7.5 centimeter diameter 10.8 centimeter long solid cylindrical aluminum shaft supporting a mirror back plate and mounting ring. The mount employs clamps acting normal to the mirror surface at three equally spaced locations on the mirror perimeter. In addition three nylon tipped set screws are located at equally spaced intervals about the mirror perimeter applying a radial load thereby preventing mirror shifting in the mount. Total angular deflection of the shaft under maximum acceleration is approximately 2.0×10^{-3} arc-sec.

6.1.4 Real-Time Mirror Control

The generation of control commands to the Dynaserv is accomplished with a Creonics VMEbus Motion Control Card (MCC) installed in the expansion VMEbus system. The MCC is capable of continuously accepting high level instructions from an active process specifying particular explicit speed and acceleration requirements for each of the two controlled axes. High speed quadrature encoder registers continuously maintain a position reading for each axis which may be requested by any active VMEbus process.

Both the MCC and Dynaserv offer a variety of control options. The option giving the smoothest motion at slow rotation speeds while still capable of meeting the acceleration objectives is shown in block diagram form in Figure 6.2. A high level jog command to the MCC specifying a particular velocity and acceleration is translated by the MCC into a series of position commands P_c. Encoder information, P_s, is subtracted from P_c generating a position error P_e. An appropriately scaled gain, P, and deadband compensation (for eliminating effects from static friction at low speeds) generates a command velocity signal, V_c, used as input to the Dynaserv. Dynaserv control subtracts an internally generated velocity signal, V_s, from V_c to generate a velocity error V_e. An appropriate gain, G, multiplies V_e to generate the actual voltage command to the motor V_o. This relatively simple proportional control is sufficient as requirements for zero positioning or following error are not stringent. Provided the target is in the camera field of view and provided the mirror position is known with sufficient accuracy, a 3D solution may be computed. There are also no external load disturbances requiring a more complex control strategy. All motions occur about a vertical axis so gravity effects are constant.

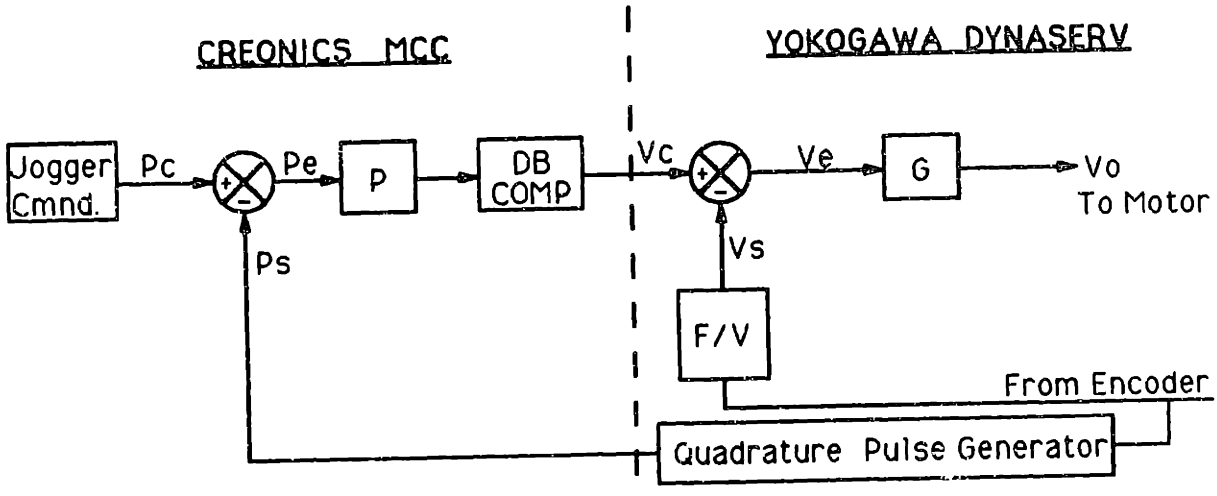


Figure 6.2: Creonics MCC and Yokogawa Dynaserv control diagram

A real-time algorithm generating speed control (jog) commands for the MCC operates on feedback from the cameras and mirror position encoders. In principal the goal is to keep the mean position of all currently active markers as close to the center of the camera view as possible. If the camera were mounted horizontally and rotated about a vertical axis the strategy would compute the mean u component of all active markers and use this as an indication of camera targeting error. In such a scenario only camera feedback information would be needed for generating the error signal.

In the actual mirror system, the virtual camera not only rotates about a vertical axis but also rolls about an axis aligned approximately with the camera principal axis. Assuming a nominal geometric configuration for the camera/mirror relative positions, a two dimensional transformation of the mean detected camera coordinates (u,v) based on the current mirror angular position (Θ) will eliminate the virtual camera roll effect and

produce the desired error, e, as

$$e = u\cos(\Theta) + v\sin(\Theta) \tag{6.1}$$

Scaling e appropriately leads to a suitable velocity input to the MCC jog command.

6.2 LVT Calibration Implementation and Testing

6.2.1 Direct Closed-Form Method Implementation

Implementing the method of Section 5.3.1 for determining the transformation matrices proved to have unforseen problems. The rotation vector determined from finding the intersection of the solution ellipse planes was at the center of the problem. The solution ellipse planes had a tendency to intersect very near the limit of admissible vectors along the $\pm\pi$ rotation line. Slight error in the data often caused plane intersections to occur in such a way that the line of intersection was never in common admissible regions of the two solution elipses. What was an admissible vector in one plane was inadmissible in the other.

An attempt at solving this problem proposed to reduce the sensitivity to data error by providing larger mirror excursions between relative full calibration comparisons. Two full calibrations with fixed global coordinates may be used to generate a similarity transform as described in Section 5.3.1. Moving global coordinates (the calibration structure) to a new position allows for generation of an equally valid similarity transform that may be used in comparison with the first. In principle, if only two full calibrations are performed at a given global coordinate definition then larger relative mirror rotations may be used while maintaining the calibration structure in the camera field of view. This method was

implemented but resulted in similar problems from noise.

Upon further analysis, the problem became more apparent. The nominal transformation from the camera to the mirror, $\bar{\mathbf{M}}_{cm}$, involves an orthogonal matrix of form

$$\bar{\mathbf{M}}_{cm} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \sin(\gamma) & \cos(\gamma) \\ 0 & \cos(\gamma) & -\sin(\gamma) \end{bmatrix}$$

where $\gamma = \frac{\pi}{4}$ radians. The unit rotation vector components $(\ell, m, n)^T$ of this matrix may in general be computed from the relations [8]

$$\ell = \frac{L}{R}; m = \frac{M}{R}; n = \frac{N}{R}$$

where (L, M, N) are determined from

$$\bar{\mathbf{M}}_{cm} - \bar{\mathbf{M}}_{cm}^{-1} = \begin{bmatrix} 0 & N & -M \\ -N & 0 & L \\ M & -L & 0 \end{bmatrix}$$

and R scales (L, M, N) to generate a unit vector representation from $R^2 = L^2 + M^2 + N^2$. The magnitude of the rotation ψ may is found from $\sin(\psi) = \frac{1}{2}R$ and $\cos(\psi) = \frac{1}{2}\left(Trace\left[\bar{\mathbf{M}}_{cm}\right] - 1\right)$. For the nominal rotation matrix $\bar{\mathbf{M}}_{cm}$ it is evident from the above relations that the implied rotation vector has a magnitude (ψ) of π but is undefined in direction $(L, M, N = 0)$.

A possible solution to this problem would involve the following. The actual matrix sought, $\mathbf{M}_{cm}$, may be thought of as the product of $\bar{\mathbf{M}}_{cm}$ and an additional unknown small rotation $\mathbf{M}_s$. By replacing $\mathbf{M}_{cm}$ with $\bar{\mathbf{M}}_{cm}\mathbf{M}_s$ in 5.26 and 5.27 and including the "known" nominal rotation in $\mathbf{M}_\Theta$, the similarity transform sought will be $\mathbf{M}_s$.

This may have sensitivity problems of its own due to the small rotation angle involved. The vectors v and v_Θ used to generate the solution ellipse will be very nearly parallel.

Small error in vector direction will tend to generate very different ellipse orientations. The ultimate solution may be to use a nominal transformation that only transforms part way through the full nominal transformation thereby avoiding the sensitivities at either extreme.

The second aspect of the direct closed-form LVT calibration, estimating the position vectors, proved to be sensitive to data error also. As the mirror is rotated through the range of positions that keep the calibration structure in the camera view, the virtual camera center of projection should sweep out an arc of a circle centered at the mirror coordinate system origin. The circle radius should be equivalent to the length of the camera/mirror position vector and the circle should lie approximately in the horizontal plane defined by the global X-Z axes.

A test was performed to help isolate this sensitivity. The virtual camera position vector from full calibrations at various mirror positions was used to identify the path followed by the virtual camera center of projection. Both camera/mirror systems were tested by placing them side-by-side with a mirror rotation axis separation of approximately 40 centimeters and placing the calibration structure at a range of 3.8 meters. This allowed a mirror excursion from side to side of approximately ± 5 degrees. Figure 6.3 shows a schematic plan view of one virtual camera in three mirror rotation positions and its relative displacement from the calibration structure. The center of projection is approximately indicated as a black dot in the camera lens. The virtual cameras were calibrated at each of three mirror positions. The resulting camera global positions (centers of projection locations) are presented in Table 6.1.

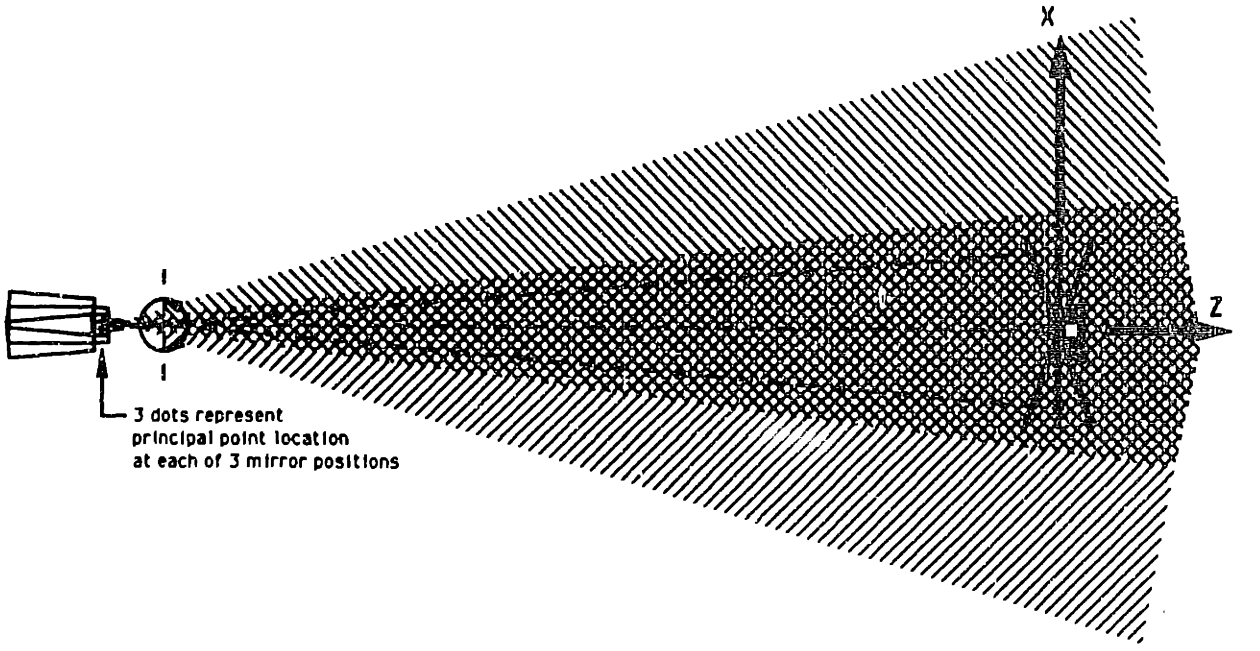


Figure 6.3: Three positions of virtual camera center of projection for three mirror positions

Table 6.1: Virtual camera global positions at 3 mirror positions

Mirror Position (Degrees)	Camera/Mirror 1			Camera/Mirror 2		
	X (m)	Y (m)	Z (m)	X (m)	Y (m)	Z (m)
-5.0	-0.153	-0.024	-3.754	0.182	-0.001	-3.920
0.0	-0.178	-0.022	-3.744	0.217	0.002	-3.888
5.0	-0.193	-0.024	-3.739	0.213	0.004	-3.900

If the system were ideal then for each camera/mirror system, the X positions would be equally distributed at ±5 degrees on either side of 0 degrees, the Y positions would all be equal, and the Z positions would indicate the greatest negative displacement at 0 degrees. This is not seen in the actual data of Table 6.1. In the case of camera/mirror 2 the X coordinate at ±5 degrees is less than the value at 0 degrees in both cases indicating significant error. It is clear that the effect of these errors could be reduced if larger angular excursions were possible. This however can not be accomplished with the 30 degree camera viewing angle limit. Dsiplacing the calibration structure farther from the cameras would allow for larger angles but the reduced marker intensity and resolution would likely offset potential accuracy gain from the increased angle.

The likely cause of this problem is a residual nonlinearity present in the internal camera calibrations and the ever present reflection problem. At the range used above, a 0.5 millimeter (1 SU) error in perceived calibration structure dimensions will lead to an approximate 5.0 millimeter error in the computed virtual camera vector length. Reflections and residual nonlinearity in different portions of a camera's field of view could easily account for this problem.

6.2.2 Numerical Method Implementation

The skew minimization was implemented using Powell's direction set method [33] for multidimensional spaces. The calibration structure was positioned at 5 locations (9,17,18,19,13) relative to the camera/mirror systems as shown in plan in Figure 6.4. At each position, three structure angular positions at which both cameras could view all 16 markers were used as individual data collection points. In this technique, knowing the relative posi-

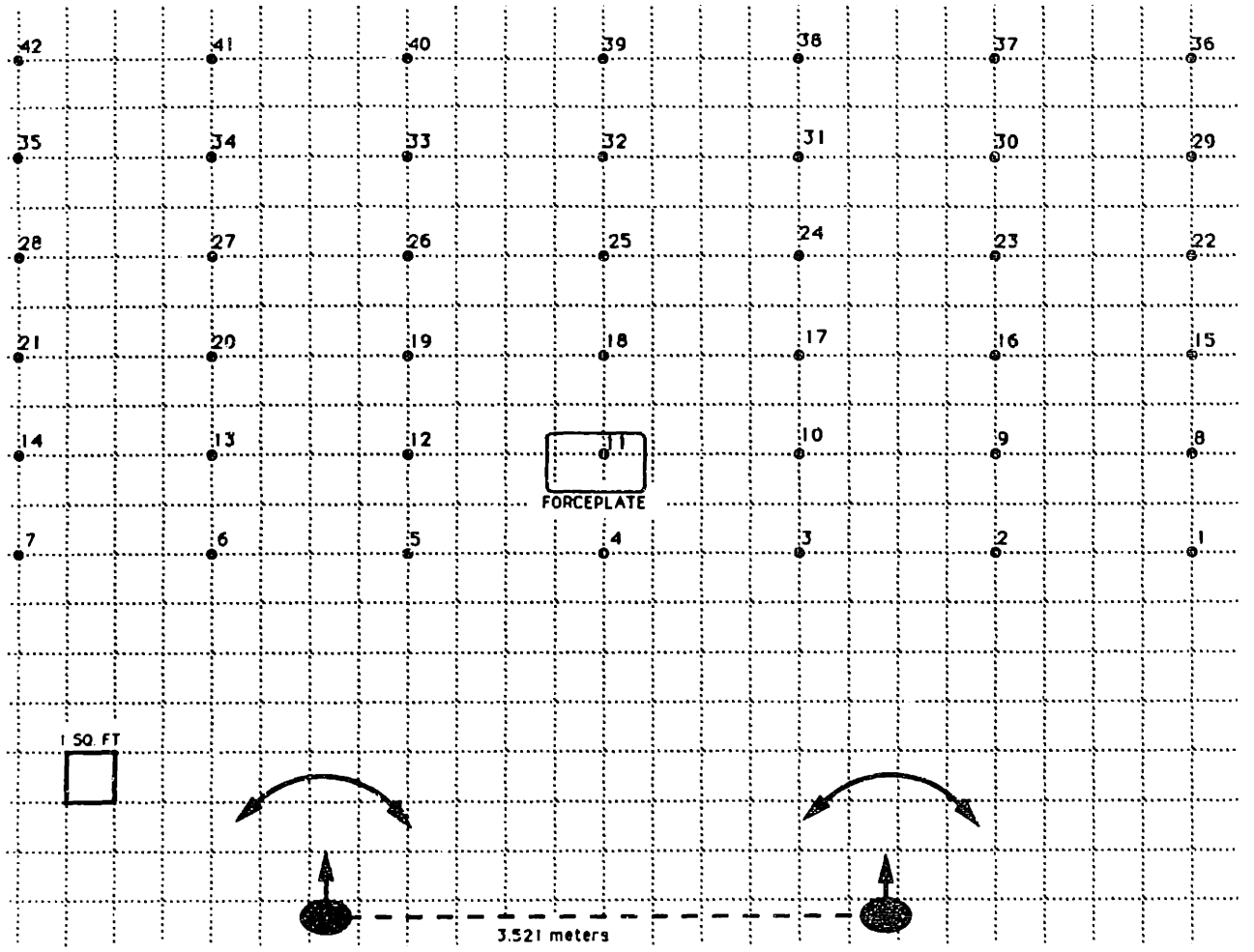


Figure 6.4: Experimental setup for numerical LVT calibration testing

tions of structure angles and structure positions is not important. All that is required is that both virtual cameras simultaneously view each marker and that a scale reference is available. The scale reference used was the known displacement of each marker relative to every other marker on a given structure position and angle. The number of structure positions and structure angles used provided 240 individual data points to be minimized.

The first attempt at implementing the skew-ray error minimization technique failed to find a reasonable and admissible solution. Analysis of the problem indicated convergence to an inappropriate minimum. The technique, in attempting to minimize skew-ray error in the reconstruction, was finding it easier to place the virtual camera between the mirror

and the global reference. This was obvious from an observed large negative z component in the camera-to-mirror vectors $\mathbf{x}_{cm_i}$. Since there was no fixed scale, the solution in effect was trying to shorten the camera baseline. This was aggravated by a poor choice of target positions for the input calibration data. The choices made did not significantly vary the virtual camera baseline displacement. A better set of positions would have the virtual cameras converge sharply on a set of targets at one position and then view targets at alternate extreme rotation positions of each camera.

The solution to this problem was to initially fix the geometry scale thereby allowing the calibration parameters to converge into a region that was admissible. Further unscaled skew minimization converged to an admissible solution ultimately yielding a smaller minimum than found previously. In a manner similar to the skew-ray minimization, the relative camera/mirror positions and orientations were used to reconstruct an estimated 3D location of each marker at each mirror position. The displacement between each marker and every other marker on a given structure position and angle was computed. The technique attempted to minimize the RMS difference in the perceived inter-marker displacement and the reference inter-marker displacement. The inter-marker error method cannot be used alone as rays do not have to intersect to provide perfect 3D displacement reconstruction. The entire technique required approximately 4 hours of computation time on the SUN 3/160 computer.

A future hybrid minimization of both inter-marker errors and skew errors may offer better overall performance. In addition, computing the partial derivatives of the minimized function over all degrees of freedom in the solution may improve the solution performance by allowing the use of maximum gradient descent techniques in the min-

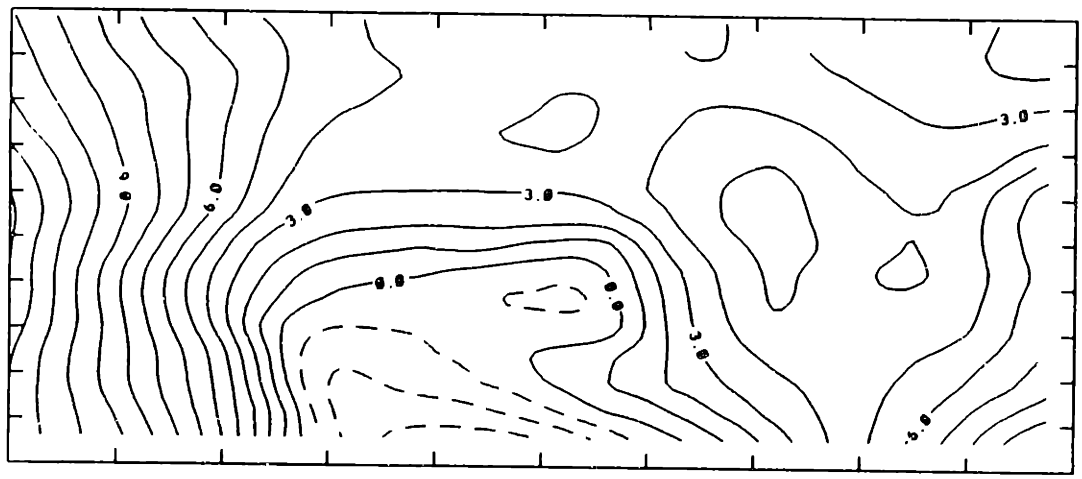


Figure 6.5: Topographical mapping of lab floor surface in measured region (millimeters)

imization. However, the additional overhead in computing the derivatives may negate any time gained through reducing the number of steps required to find the minimum.

Test of Calibration

A test of this calibration involved placing the calibration structure at 42 positions on the laboratory floor spanning an area of approximately 7.3 meters by 3 meters as indicated in Figure 6.4. At each location the calibration structure was carefully positioned square to the floor coordinates and leveled using a two axis leveling bubbie. A MilliBeam laser leveling system was used to mark relative heights of the calibration structure support post at each position to account for fluctuations in the level of the floor. The measured the floor topography is shown in Figure 6.5.

At each structure location static data was collected from each camera using the standard TRACK5 data acquisition program. Each data set was processed to determine the

6 DOF position and orientation of the structure using the calibration determined above for both the camera/mirror and global/mirror parameters. To provide a reference to the lab coordinates, the structure position and orientation at location 11 (see Figure 6.4) was used to reference all other structure location positions and orientations.

In comparing the predicted position and orientations of the calibration structure at each sampled location to the actual position and orientation, it was evident that a slight discrepancy in the assumed orientation of the reference structure with respect to the floor coordinate system existed. Apparently the calibration structure had been placed consistently out of proper orientation with respect to the floor defined coordinates over all 42 placements. This was corrected with a reference orientation adjustment of 0.75 degrees about global Z and 1.25 degrees about global Y.

The resulting computed errors in global (X, Y, Z) position components and global (α, β, γ) orientation components are shown in Figures 6.6 through 6.11.

Evaluating the plots of position error indicates approximate mean static absolute position error of 10 millimeters over the two dimensional area tested. This corresponds to an absolute position accuracy on the order of 0.2 percent over the tested area. The errors in orientation are very small having an approximate combined mean error over the tested range of 0.6 degrees.

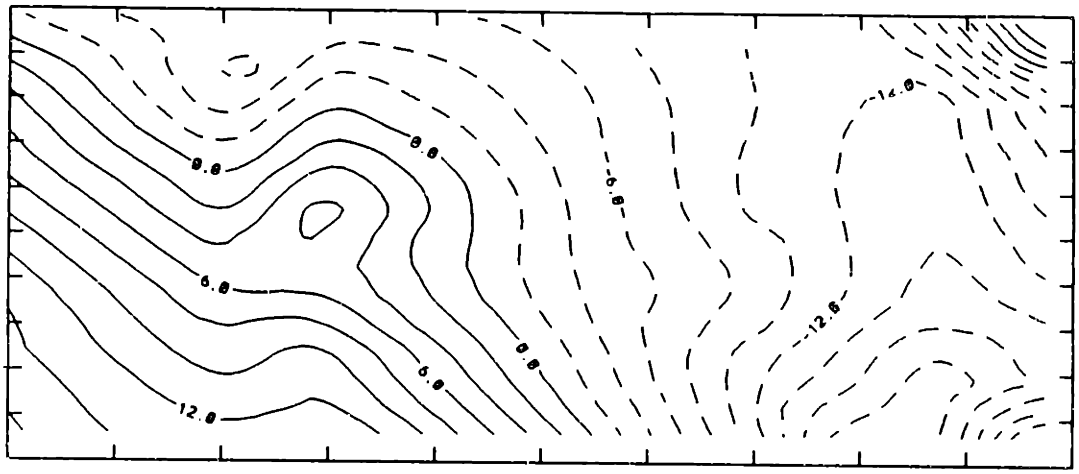


Figure 6.6: Global X error over tested measurement range (millimeters)

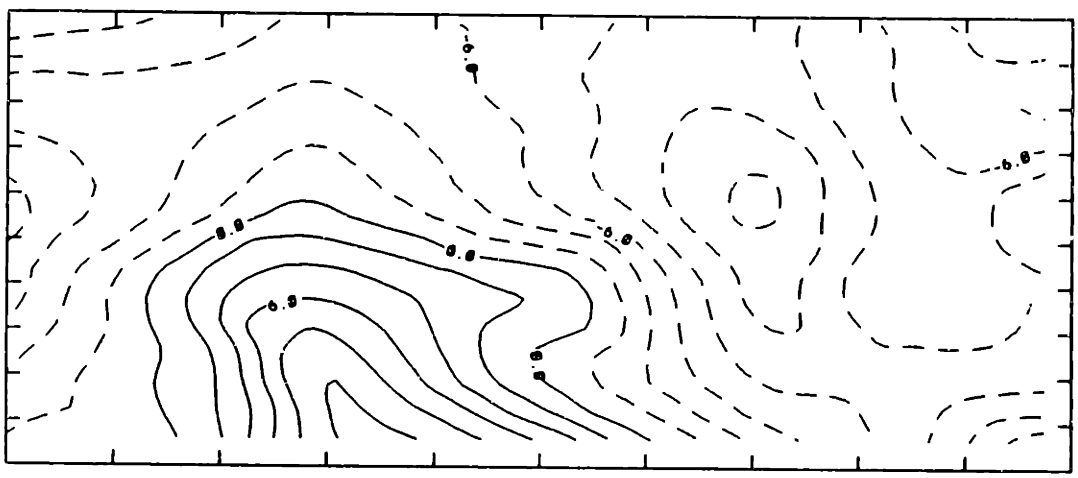


Figure 6.7: Global Y error over tested measurement range (millimeters)

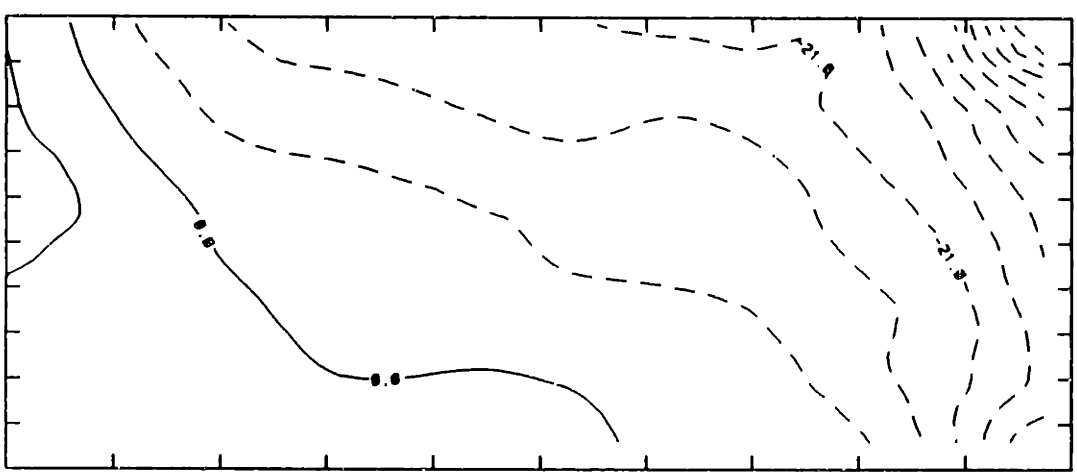


Figure 6.8: Global Z error over tested measurement range (millimeters)

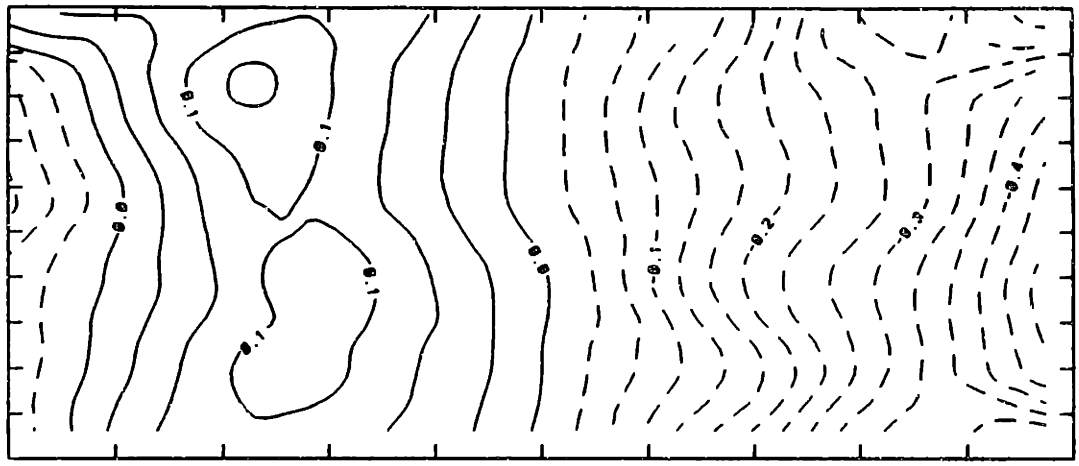


Figure 6.9: Global α error over tested measurement range (degrees)

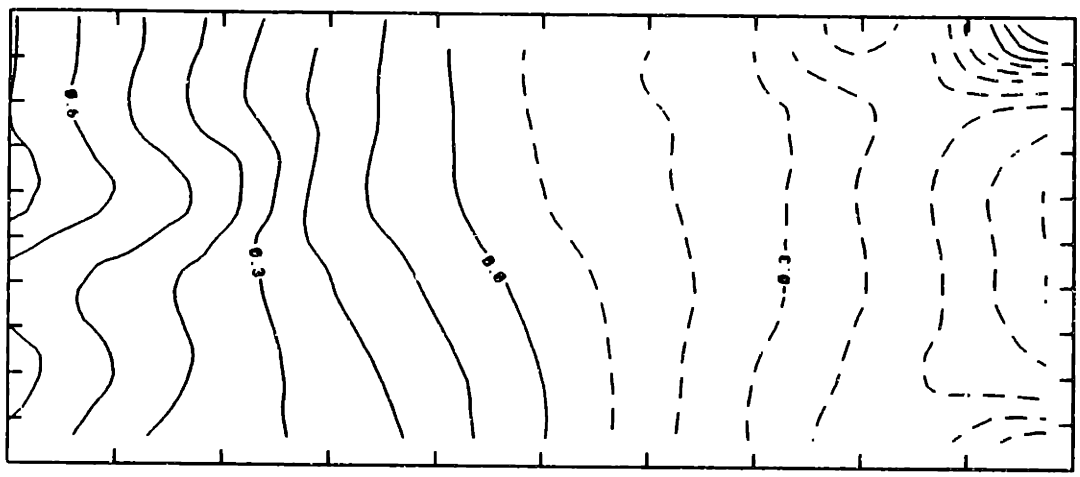


Figure 6.10: Global β error over tested measurement range (degrees)

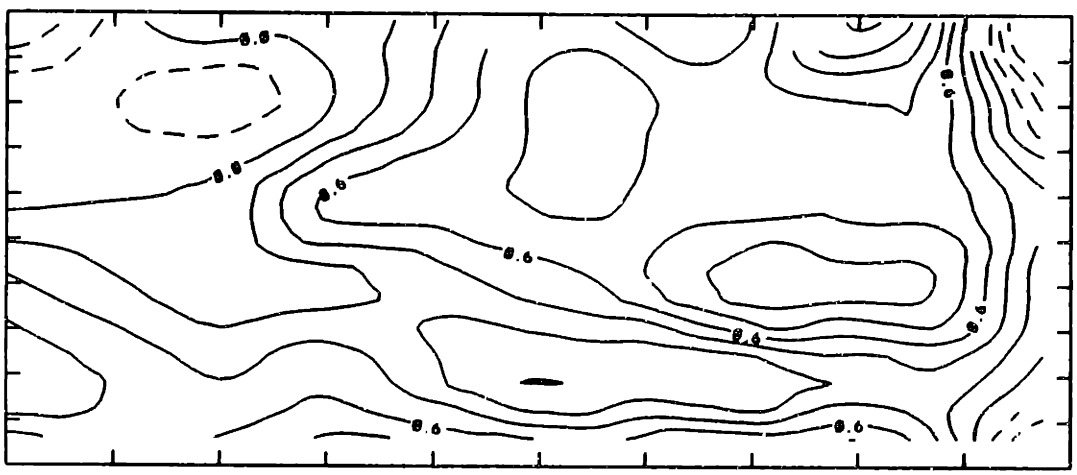


Figure 6.11: Global γ error over tested measurement range (degrees)

6.3 LVT Calibration Conclusions and Suggestions for Further Work

It is clear from the results presented above that the LVT system potentially can be very accurate. What is not clear at this point is how to realize that accuracy in a reliable and repeatable manner. Central to this issue is the apparent difficulty in identifying the camera-to-mirror transformation parameters, x_{cm} and M_{cm}. This problem was illustrated by a second attempt at calibrating the system using the two stage numerical iterative minimimization mentioned above. In this attempt, even the two stage approach failed to find an admissible minimum. It is likely that the number of degrees-of-freedom in the solution space combine to form geometrically inadmissible solutions that are acceptable to the algorithm. The minimized solution is obviously not properly constrained.

Further investigation into the problem was not pursued at this point for several reasons. Approximately 6 months before completion of this thesis, after numerous experiments attempting to identify residual error in MIT's SELSPOT system, SELSPOT cameras, with serial numbers 148 (camera 1) and 178 (camera 2) and loaned from the Biomotion Laboratory at Massachusetts General Hospital (MGH) were tested. The difference in performance was remarkable. There was almost an order of magnitude reduction in the typical skew-ray errors seen in data after calibration using the MGH cameras when compared to the MIT cameras. The difference could not be explained but it was clear the MGH cameras would be used in all final tests for documentation purposes. Due to the late discovery, many tests that were already considered complete had to be re-run to provide a uniform and consistent set of data to work with. While MGH was extremely

generous in making their camera system available to MIT, progress was slowed considerably. All data in this thesis baring that of Chapter 3 was collected using the MGH cameras.

The announced availability of an improved technology camera, Optotrack, manufactured by Northern Digital, Warterloo, Ontario, also contributed to a desire to postpone further work on this problem. The new cameras employ dual linear detectors arranged orthogonally with respective cylindrical lenses that take the place of the former two axis detector and spherical lens. This configuration permits location of the peak image intensity on the detector thereby eliminating the reflection problem plaguing the current system. It is also capable of much higher image resolution than the current SELSPOT cameras. As this represents an important developmental step in this technology it made pursuing the problems inherent in old technology cameras much less pertinent.

In addition to the issues mentioned above, it is clear that additional design and a restructuring of the calibration technique is required. In particular, since identifying the camera-to-mirror transformation has proved difficult and time consuming, a mount fixing the camera relative to the mirror system should be designed and constructed. This would allow the camera and mirror to be repositioned as a unit without changing the camera/mirror transformation parameters. A proposed configuration is shown in Figure 6.12. The camera and motor/mirror mount are fixed to a common base that is in turn mounted to a vertical support. A cable, pully, and counter weight arrangement permits easy vertical positioning of the unit. A base for the vertical post would have retractable wheels for easy horizontal repositioning of the unit on the laboratory floor.

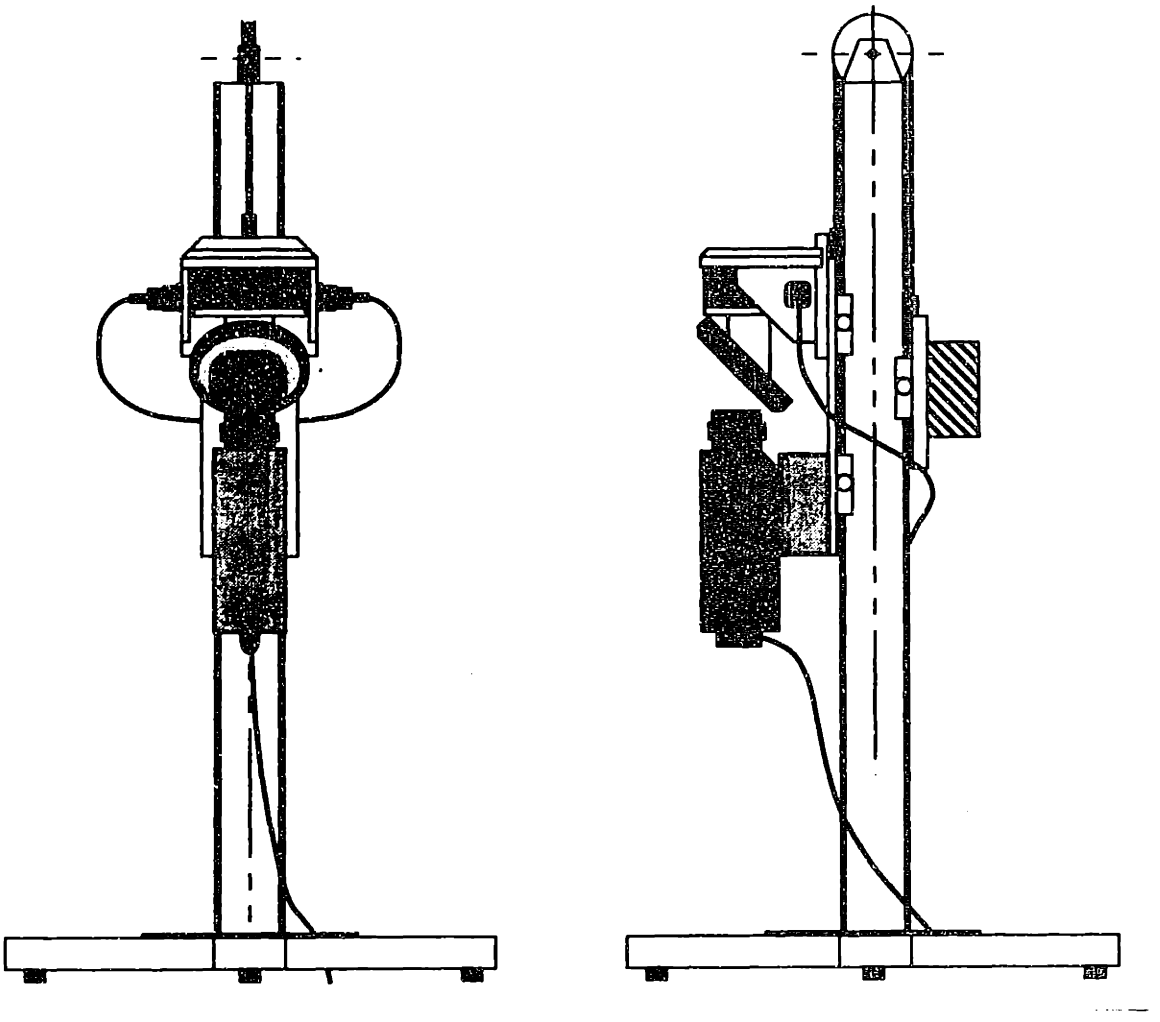


Figure 6.12: Proposed camera/mirror mounting on a common base

With a fixed camera/mirror transformation a more elaborate approach to calibration could be pursued. This would likely involve identifying a set of reference locations in the lab similar to the layout used to test the calibration that was successful (Figure 6.4). This would provide an effectively very large structure for the system to observe, thereby permitting large mirror angular excursions resulting in a reduced sensitivity to the error conditions mentioned. A slightly reworked direct method of calibration, relating the different calibration structure positions to a common coordinate system, would likely perform properly using this approach.

This modified camera/mirror calibration would only be required if the mirror or camera mounting was disassembled or if a possible missalignment was suspected due to mishandling of the system. Otherwise, calibration for a specific setup would be as straightforward to perform as the fixed camera system and would require no additional knowledge of the underlying operations by the typical system user.

Bibliography

[1] Y.I. Abdel-Aziz and H.M. Karara. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. In *Symposium on Close-Range Photogrammetry.* American Society of Photogrammetry, Jan 1971. University of Illinois, Urbana.

[2] E.K. Antonsson. The derivation and implementation of a dynamic three-dimensional linkage analysis technique. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, May 1978.

[3] E.K. Antonsson. *A Three-Dimensional Kinematic Acquisition and Intersegmental Dynamic Analysis System for Human Motion.* PhD thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, June 1982.

[4] E.K. Antonsson and R.W. Mann. Automatic 6-D.O.F. kinematic trajectory acquisition and analysis. *J. Dynamic Systems, Measurement, and Control*, 111:31–39, 1989.

[5] R.A. Berger. An automated calibration scheme for the Selspot/TRACK electro-optical transducer. Bachelor's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, June 1985.

[6] R. Chan. Real-time simulation of an auditory spatial display with selective scanning. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, September 1982.

[7] F.C. Conati. Real-time measurement of three-dimensional multiple rigid body motion. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, June 1977.

[8] S.H. Crandall, D.C. Karnopp, E.F. Jr. Kurtz, and D.C. Pridmore-Brown. *Dynamics of Mechanical and Electromechanical Systems.* McGraw Hill, New York, 1968.

[9] J. Dapena, E.A. Harman, and J.A. Miller. Three-dimensional cinematography with control object of unknown shape. *J. Biomechanics*, 15(1):11–19, 1982.

[10] R.S. Fijan, P.K. Mansfield, and R.W. Mann. Estimation of camera positions and orientations using a control distribution of unknown relative geometry. In *108th Winter Annual Meeting.* American Society of Mechanical Engineers, Dec 1987. Boston.

[11] S.K. Ghosh. *Analytical Photogrammetry.* Pergamon Press, New York, 1979.

[12] H. Goldstein. *Classical Mechanics, 2nd Edition.* Addison-Wesley Publishing Company, Reading, Massachusetts, 1981.

[13] S. Granapathy. Decomposition of transformation matrices for robot vision. In *Proc. IEEE International Conference in Robotics and Automation, Atlanta*, 1984.

[14] H. Hatze. High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT-approach. *J. Biomechanics*, 21(7):533–538, 1988.

[15] J.P. Holman. *Experimental Methods for Engineers, Second Edition.* McGraw Hill, New York, 1971.

[16] M.M. Ijaz. Asymmetric arm trajectory formation: A kinematic approach. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, June 1985.

[17] H.W. Janson. *History of Art.* Prentice-Hall, Inc. and Harry N. Abrams, Inc., Englewood Cliffs, N.J. and New York, 1973.

[18] M.H. Krag. Quantitative techniques for the analysis of gait. *Automedica*, 6:85–97, 1985.

[19] J.B. Lenox. *Six Degree of Freedom Human Eyeball Movement Analysis Involving Steriometric Techniques.* PhD thesis, Stanford University, 1976.

[20] M.A. Lesczynski. A kinematic analysis of the human ankle complex. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, June 1986.

[21] P.J. Lord. 3D-gait: A three dimensional computer graphic display for human motion analysis from track gait data. Bachelor's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, May 1987.

[22] P.J. Lord. Real-time analysis and display of human movement. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, June 1989.

[23] G. Lucovsky. Photoeffects in nonuniformly irradiated p-n junctions. *Journal of Applied Physics*, 31(6), June 1960.

[24] R.W. Mann. Stereophotogrammetry applied to hydraulic analogue studies of unsteady gas flow. *Photogrammetric Engineering*, September 1962.

[25] P.K. Mansfield. The trajectory of the golf club. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, September 1983.

[26] P.K. Mansfield, R.S. Fijan, and R.W. Mann. Internal calibration of opto-electronic cameras. In *Proceedings of IEEE Ninth Annual Conference of the Engineering in Medicine and Biology Society*. IEEE, 1987.

[27] F.H. Moffit and E.M. Mikhail. *Photogrammetry, 3rd edition*. Harper and Row, New York, 1980.

[28] T. Morris and M. Donath. Error analysis of 6 D.O.F. rigid body motion tracking systems. In *Winter Annual Meeting*. American Society of Mechanical Engineers, Nov 1988. Chicago.

[29] M.C. Murphy. *Geometry and the Kinematics of the Normal Human Knee*. PhD thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, expected, June 1990.

[30] M.C. Murphy and R.W. Mann. A comparison of smoothing and digital filtering/differentiation of kinematic data. In *Proceedings of IEEE Ninth Annual Conference of the Engineering in Medicine and Biology Society*. IEEE, 1987.

[31] E. Muybridge. *The Human Figure in Motion.* Dover Publications Inc., New York, 1955.

[32] D.E. Ottenheimer. Dynamic simulation of auditory spatial displays. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, Jar 1982.

[33] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes.* Cambridge University Press, Cambridge, 1987.

[34] G.H. Schut. Construction of orthogonal matrices and their applications in analytical photogrammetry. *Photogrammetria*, XV(4):149–162, 1958-59.

[35] G.H. Schut. On exact linear equations for the computation of the rotational elements of absolute orientation. *Photogrammetria*, XVII(1):34–37, 1960-61.

[36] J.L. Stein. *Design Issues in the Stance Phase Control of Above-Knee Prostheses.* PhD thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, January 1983.

[37] J.L. Stein and W.C. Flowers. Stance phase control of above-knee prostheses: Knee control versus sach foot design. *J. Biomechanics*, 20(1):19–28, 1987.

[38] D.R. Stoutemyer. Systems study and design of a blind mobility aid simulator. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, Jan 1965.

[39] G. Strang. *Introduction to Applied Mathematics.* Wellesley Cambridge Press, Wellesley, Massachusetts, 1986.

[40] G. Strang. *Linear Algebra and its Applications.* Harcourt Brace Jovanovich, San Diego, 1988.

[41] A.K. Tetewsky. Implementing a real time computation and display algorithm for the Selspot system. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, May 1978.

[42] T. Theil. A comparison of active and passive sound localization. Master's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, May 1983.

[43] E.H. Thompson. An exact linear solution of the problem of absolute orientation. *Photogrammetria*, XV(4):163–179, 1958-59.

[44] F.E. Veldpaus, H.J. Woltring, and L.J.M.G. Dortmans. A least-squares algorithm for the equiform transformation from spatial marker co-ordinates. *J. Biomechanics*, 21(1), 1988.

[45] H.J. Woltring. Single- and dual-axis lateral photodetectors of rectangular shape. *IEEE Transactions on Electronic Devices*, pages 581–590 (corr. pg. 1101), 1975.

[46] H.J. Woltring. Simultaneous multiframe analytical calibration (S.M.A.C) by recourse to oblique observations of planar control distributions. *SPIE Applications of Human Biosteriometrics (NATO)*, 166, 1978.

[47] H.J. Woltring. Planar control in multi-camera calibration for 3-D gait studies. *J. Biomechanics*, 13:39–48, 1980.

[48] H.J. Woltring. Estimation and precision of 3-D kinematics by analytical photogrammetry. In *Computing in Medicine, University of Strathclyde Bioengineering Seminars*, Aug 1981.

[49] H.J. Woltring. One hundred years of photogrammetry in biolocomotion. In *Proceedings of the Symposium on Biolocomotion: a Century of Research Using Moving Pictures, Formia, Italy*, 1989.

[50] H.J. Woltring, 1990. Personal Communication.

[51] G.A. Wood and R.N. Marshall. The accuracy of DLT extrapolation in three-dimensional film analysis. *J. Biomechanics*, 19(9):781–785, 1986.

[52] L.J. Wormley. A kinematic analysis of the grand-plie. Bachelor's thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, May 1985.

Appendix A

Polar Decomposition of Nearly Orthogonal Matrices

A technique for estimating the components of a pseudo transformation matrix was demonstrated in Chapter 4. The elements of the matrix were not constrained to have properties of an orthogonal transformation. This appendix provides a description of the method developed in this thesis for reducing the matrix found to a valid orthogonal transformation. The problem of obtaining the best orthogonal transformation matrix represented by a matrix estimate from experimentally obtained data (containing error) is not unique to this thesis and has been explored using alternative techniques [13].

A.1 Theory

Linear Algebra provides a theorm that states: given any real square matrix A there exists a symmetric matrix S and orthogonal matrix Q such that

$$\mathbf{A} = \mathbf{QS} \tag{A.1}$$

This is known as *polar decomposition* of A. Strang [39], (pg. 446) refers to its application in studying the deformation of solids:

> In any deformation it is important to separate stretching from rotation, and that is exactly what QS achieves. The orthogonal matrix Q is a rotation, and possibly a reflection. The material feels no strain. The symmetric matrix S has eigenvalues $\sigma_1, \ldots, \sigma_r$, which are the stretching factors (or compression factors).

In this case the "stretching" has been caused by error in the experimentally obtained data. The goal is to obtain a representation for the rotation alone.

Given that the inverse of an orthogonal matrix is equal to its transpose and that the transpose of a symmetric matrix is equal to itself, it is easy to show that

$$\mathbf{S}^2 = \mathbf{A}^T\mathbf{A} \tag{A.2}$$

Thus if we are given a matrix A and wish to decompose it into its symmetric and orthogonal components, a suitable S which makes A.2 true must be found.

If the assumption is made that A tends toward orthogonality it follows that S will tend toward the identity matrix, I. In this case S may be represented as the sum of I and a symmetric matrix L having all elements close to zero:

$$\mathbf{S} = \mathbf{I} + \mathbf{L} \tag{A.3}$$

where

$$\mathbf{L} = \begin{bmatrix} \ell_1 & \ell_2 & \ell_3 \\ \ell_2 & \ell_4 & \ell_5 \\ \ell_3 & \ell_5 & \ell_6 \end{bmatrix} \; ; \; |\ell_1, \ldots, \ell_6| \ll 1 \tag{A.4}$$

Squaring A.3 yields the following:

$$\mathbf{S}^2 = \mathbf{I} + 2\mathbf{L} + \mathbf{L}^2 \tag{A.5}$$

Since the elements of L are all close to zero, eliminating all second order terms of L from A.5 provides a good linear approximation for $\mathbf{S}^2$. Applying this to A.5 and expanding yields

$$\mathbf{S}^2 \approx \begin{bmatrix} 1+2\ell_1 & 2\ell_2 & 2\ell_3 \\ 2\ell_2 & 1+2\ell_4 & 2\ell_5 \\ 2\ell_3 & 2\ell_5 & 1+2\ell_6 \end{bmatrix} \tag{A.6}$$

A.2 Procedure

Decomposing a given A into its symmetric S and orthogonal Q components proceeds as outlined below.

1. Starting with A, compute elements ℓ_i of L using A.2 and A.6. Compute $\mathbf{S}_1$ from A.3.

2. Compute estimate $\mathbf{Q}_1$ from A.1 and A.3. $\mathbf{Q}_1$ is a matrix which is closer to orthogonal than A.

3. Repeating steps 1 and 2 substituting $\mathbf{Q}_1$ for A will yield $\mathbf{S}_2$ and $\mathbf{Q}_2$.

4. Repeating this procedure n times until $\|\mathbf{Q}_n\|$ is sufficiently close to 1 yields a total solution of $\mathbf{S} = \mathbf{S}_1\mathbf{S}_2...\mathbf{S}_n$ and $\mathbf{Q} = \mathbf{Q}_n$.

A.3 Discussion

This technique has been found to converge to a solution rapidly. A typical rotation matrix estimate A found using the techniques of Chapter 4 will decompose into a valid Q and S with $n \leq 4$. The convergence criterion was specified as $|\,\|\mathbf{Q}\| - 1| < 0.000001$.

For initial A matrices that tend toward improper orthogonality ($\|\mathbf{A}\| \approx -1$), as in the case of transformations through a mirror, this method performs equally well. The convergence criterion must however be changed to reflect a convergence of $\|\mathbf{Q}_n\|$ to -1.

It has also been found, though not investigated thoroughly, that arbitrary A matrices, having no presumed orthogonal tendencies, may be decomposed rapidly with this technique. Iterations on the order of $n = 10$ have been observed.

Appendix B

Direct Six Degree-of-Freedom Object Estimation from Photogrammetric Data

This appendix describes a technique for estimating the position and orientation of a rigid structure of markers in a global reference frame directly from photogrammetric image space data. The Direct 6 Degree-of-Freedom (D6DOF) method does not require the preliminary step of determining 3D positions of markers in object space common to most other object orientation algorithms. This algorithm will generate a solution when at least three markers, known in a common structure coordinate system, are seen by two cameras. Ideally, the cameras need not view the same markers provided a minimum of six camera/marker combinations are viewed and that all markers have a known position in the structure coordinate system. A preliminary investigation into the implementation of this algorithm indicates that noise in image data adds the constraint that at least one marker be viewed by two cameras.

B.1 Algorithmic Development

Algorithmic development of the D6DOF method requires combining concepts presented in Chapter 2 concerning 3-D object-point estimation and rigid body reconstruction. The approach is to restructure the 3D point estimation algorithm to use image space data from multiple points on a common rigid body to estimate the object position and orientation directly. As with the external calibration method described in Chapter 4 the approach requires three steps: initial approximation, reduction to admisible elements (polar decomposition), and iterative refinement of the solution.

In the initial approximation it is desired to obtain estimates for twelve dependent linearly related parameters defining the position and orientation of a body coordinate system relative to a reference system. These parameters consist of the 9 elements of a rotation matrix R and 3 elements of a position vector r as seen below (identical to equation 2.28).

$$\mathbf{p}_i = \mathbf{R}\mathbf{a}_i + \mathbf{r} \tag{B.1}$$

The problem is illustrated in Figure B.1 which combines the concepts presented in figures 2.2 and 2.3.

The camera to global coordinate system transformation is represented as in equation 2.4 as

$$\mathbf{x}_c = \mathbf{M}[\mathbf{x}_o - \mathbf{x}_{oc}] \tag{B.2}$$

It is assumed that external calibration has been performed so that the external calibration parameters, $\mathbf{M}$ and $\mathbf{x}_{oc}$, of equation B.2 are considered known.

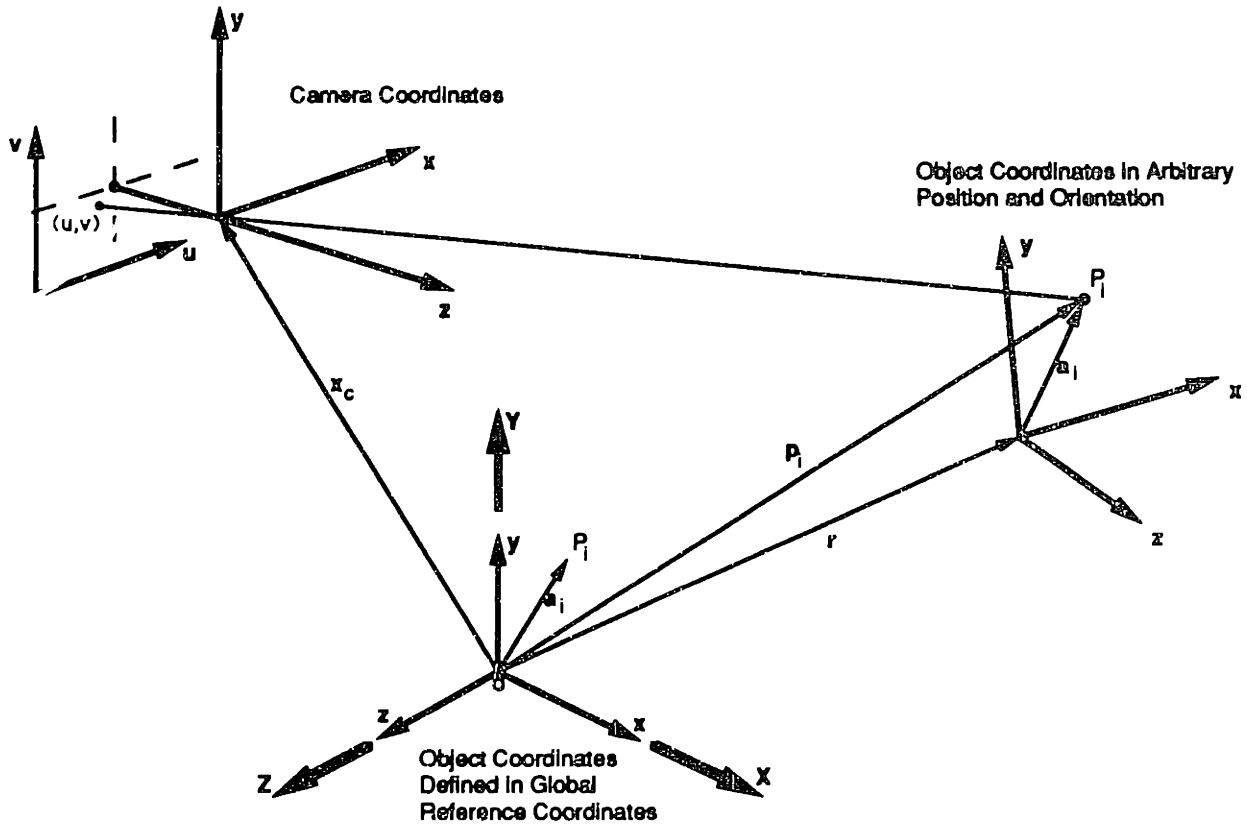


Figure B.1: The direct rigid body position and orientation problem

From this point the derivation proceeds from the camera colinearity equations in their rearranged form of equations 2.10 and 2.11. Since (x_o, y_o, z_o) in equations 2.10 and 2.11 represent any point in global space then they must also apply to the components of $\mathbf{p}_i$ in equation B.1. By defining components of $\mathbf{R}$ as

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

the components of $\mathbf{r}$ as

$$\mathbf{r} = [r_x r_y r_z]^T$$

the object points in the reference coordinate system as

$$\mathbf{a}_i = [x_i y_i z_i]^T$$

and the camera to global origin vector $\mathbf{x}_{co}$ in camera space as

$$\mathbf{x}_{co} = -\mathbf{M}\mathbf{x}_{oc}$$

expansion of equation B.1 followed by substitution of components of $\mathbf{p}_i$ into equations 2.10 and 2.11 yields the following set of equations for object point i viewed by camera j:

$$c_{ij}(x_i r_{11} + y_i r_{12} + z_i r_{13} + r_x) + d_{ij}(x_i r_{21} + y_i r_{22} + z_i r_{23} + r_y) + e_{ij}(x_i r_{31} + y_i r_{32} + z_i r_{33} + r_z) = (z_{co} u'_i - x_{co})_j \quad \text{(B.3)}$$

$$f_{ij}(x_i r_{11} + y_i r_{12} + z_i r_{13} + r_x) + g_{ij}(x_i r_{21} + y_i r_{22} + z_i r_{23} + r_y) + h_{ij}(x_i r_{31} + y_i r_{32} + z_i r_{33} + r_z) = (z_{co} v'_i - y_{co})_j \quad \text{(B.4)}$$

where

$$c_{ij} = (m_{11} - m_{13} u'_i)_j$$

$$d_{ij} = (m_{12} - m_{23} u'_i)_j$$

$$e_{ij} = (m_{13} - m_{33} u'_i)_j$$

$$f_{ij} = (m_{21} - m_{13} v'_i)_j$$

$$g_{ij} = (m_{22} - m_{23} v'_i)_j$$

$$h_{ij} = (m_{23} - m_{33} v'_i)_j$$

Equations B.3 and B.4 represent two linear equations in the twelve unknown components of $\mathbf{R}$ and $\mathbf{r}$ for camera j viewing object point i. If at least six camera/marker combinations are employed then a minimum of twelve equations will be available for reduction to the twelve unknowns using a linear least-squares approach.

As with the external calibration method, presented in Chapter 4, the $\mathbf{R}$ and $\mathbf{r}$ found above are only initial approximations. No constraint has been imposed to make $\mathbf{R}$ orthogonal. As in Chapter 4 the second step is to perform polar decomposition on $\mathbf{R}$ to find the best fit orthogonal matrix.

Once an orthogonal $\mathbf{R}$ has been found additional small transformation corrections may be computed and implemented in a manner similar to the presentation in step three of Chapter 4. Following this proceedure and using the definitions described above two linear equations in six unknowns $(\alpha, \beta, \gamma, b_x, b_y, b_z)$ may be written for each camera/object point (j, i) combination as

$$[d_{ij}\tilde{z}_i - e_{ij}\tilde{y}_i]\,\alpha + [e_{ij}\tilde{x}_i - c_{ij}\tilde{z}_i]\,\beta + [c_{ij}\tilde{y}_i - d_{ij}\tilde{x}_i]\,\gamma -$$

$$c_{ij}b_x - d_{ij}b_y - e_{ij}b_z = (z_{co}u'_i - x_{co})_j - c_{ij}\tilde{x}_i - d_{ij}\tilde{y}_i - e_{ij}\tilde{z}_i \qquad \text{(B.5)}$$

$$[g_{ij}\tilde{z}_i - h_{ij}\tilde{y}_i]\,\alpha + [h_{ij}\tilde{x}_i - f_{ij}\tilde{z}_i]\,\beta + [f_{ij}\tilde{y}_i - g_{ij}\tilde{x}_i]\,\gamma -$$

$$f_{ij}b_x - g_{ij}b_y - h_{ij}b_z = (z_{co}v'_i - y_{co})_j - f_{ij}\tilde{x}_i - g_{ij}\tilde{y}_i - h_{ij}\tilde{z}_i \qquad \text{(B.6)}$$

where $(\tilde{x}_i, \tilde{y}_i, \tilde{z}_i)$ are the current estimates of the object point i location in the reference coordinate system based on the current estimates for $\mathbf{R}$ and $\mathbf{r}$.

In practice it was found that the initial estimation step contains certain properties making a solution at times impossible. If the object points are arranged in the x-y plane of the object coordinate system then the solution technique fails since all z_i elements in equations B.3 and B.4 are zero. It was also found that there was no guarantee the polar decomposition in the second step would provide a proper orthogonal matrix.

For these reasons and to obtain some results promptly it was decided to eliminate the first and second steps of the method described above until the problems could be studied in detail. Instead, only the third step was used.

In processing successive frames of data the approximation used as input to the third step was provided by the previous frame's results. The identity matrix and zero vector were used as the approximation for the first frame of a data set.

A numerical minimum gradient search technique was implemented to verify the results of this approach. The current estimate of the six degrees of freedom allows for estimation of the object point positions in global space. The colinearity equations then allow for back calculation of the image space points. The technique attempted to minimize the RMS differences between each image space coordinate as acquired (and internally corrected) and its corresponding coordinate as back calculated. Although this approach is computationally intensive, the results agreed in fine detail with the D6DOF method proposed above.

B.2 Experimental Tests

A series of tests were performed using real SELSPOT data. Three of these tests will be discussed here as they illustrate both the advantages and disadvantages of this algorithm. Note that this presentation is not intended to represent an exhaustive analysis of the algorithm, but does demonstrate some of its interesting properties.

The cameras were placed on a horizontal beam roughly 1 meter from the floor with principal axes forming a plane parallel to the floor and a baseline displacement between principal points of roughly 3.5 meters. The camera angles of convergence were each 45 degrees resulting in camera principal axes forming a mutual 90 degree angle of intersection. The camera external orientation parameters were calibrated for as described in Chapter 4.

The rigid object studied consisted of a 25 by 8 centimeter strip of 4 millimeter plexiglass sheet. Two 45 degree bends across the width formed three facets of equal surface

Marker	*Facet*	*X*	*Y*	*Z*
1	1	0.0804	-0.0450	-0.0146
2	1	0.1086	0.0000	0.0136
3	1	0.0804	0.0450	-0.0146
4	1	0.0662	0.0000	-0.0288
5	2	0.0300	0.0300	-0.0500
6	2	0.0300	-0.0300	-0.0500
7	2	-0.0300	0.0300	-0.0500
8	2	-0.0300	-0.0300	-0.0500
9	3	-0.0662	0.0000	-0.0288
10	3	-0.0804	0.0450	-0.0146
11	3	-0.1086	0.0000	0.0136
12	3	-0.0804	-0.0450	-0.0146

Table B.1: Location of markers in segment body coordinate system

area. The facets on each end formed a mutual angle of 90 degrees. Four LED markers were mounted on each facet resulting in a total of 12 markers. The origin of the body coordinate system was at the intersection of the three center normal vectors from each facet. The body z-axis was normal to the center facet positive away from the facet. The body y-axis was parallel to the plane of all facets. Table B.1 lists the markers in their body coordinate locations. Held at neutral orientation with respect to global coordinates the outer facets were perpendicular to the principal axes of respective cameras (see Figure B.2). For reference, the facets were numbered from 1 to 3, with facet 1 containing +x markers 1-4, facet 2 in the center with markers 5-8, and facet 3 containing -x markers 9-12. In the neutral position, camera 1 could view facets 2 and 3, and camera 2 could view facets 1 and 2.

Data acqusition consisted of the following: Starting in the neutral position, the object was rotated by hand about the global y-axis repeatedly through approximately $\pm$ 45

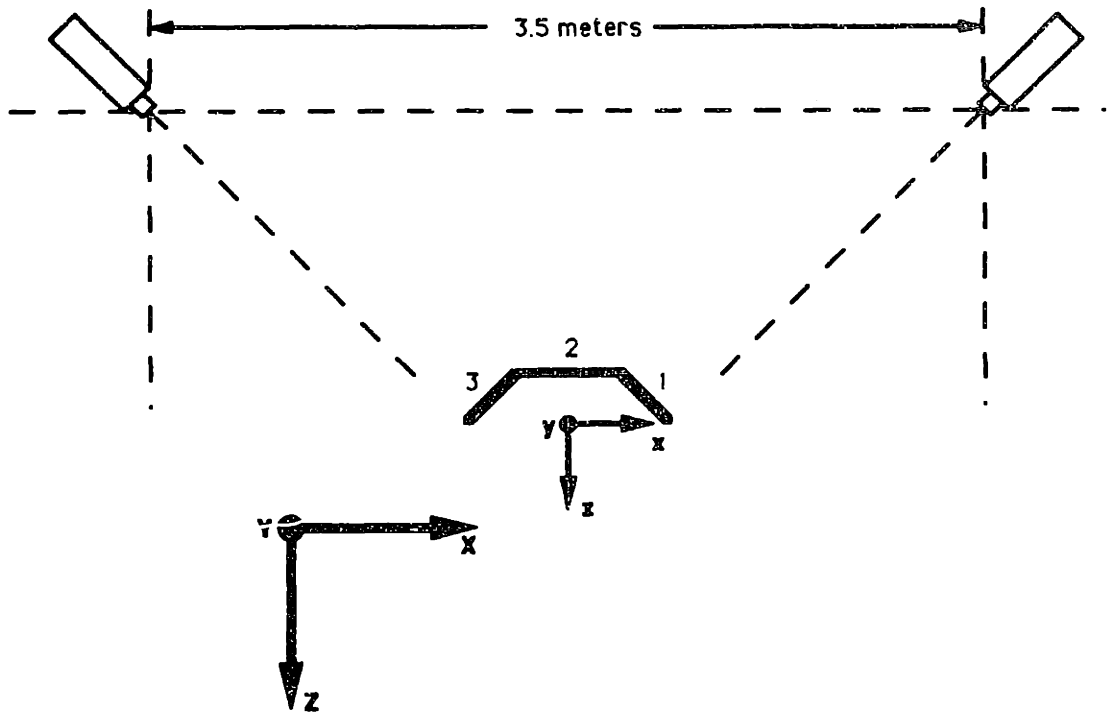


Figure B.2: D6DOF experimental setup

degrees. At +45 degree rotation angle, camera 1 could view facet 3 only, while camera 2 could view all facets. At -45 degree rotation angle, camera 1 could view all facets, while camera 2 could view facet 1 only. The twelve markers were sampled at a frame rate of 200 Hertz over a 2 second interval.

B.3 Data Processing and Results

It was originally desired to compare performance of the direct orientation algorithm with Schut's algorithm (see Chapter 2) under two different data cases. In case 1, data from all 12 markers was available for processing. In this case there were always at least four markers simultaneously visible to the two cameras.

In case 2 markers 5-8 (facet 2) were eliminated from the data available for processing. In this case there were periods (near zero degrees rotation) when no camera viewed a common marker, but each camera always viewed at least four markers.

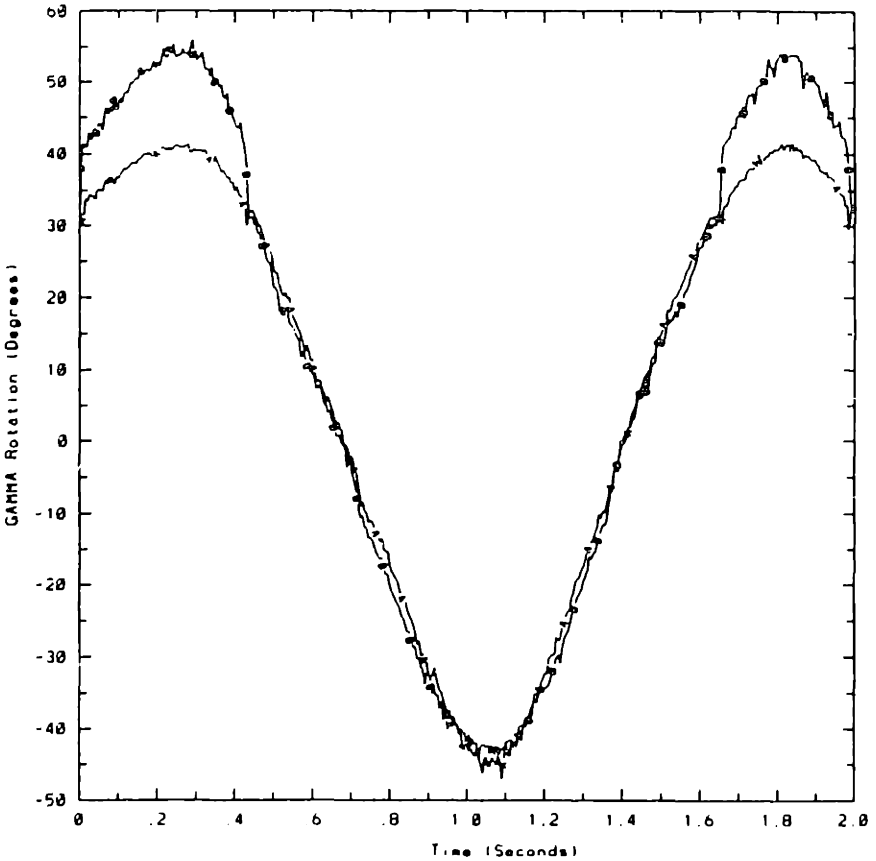


Figure B.3: Case 1: All 12 markers processed. D6DOF (A), Schut (B)

Once the data had been processed, it was decided to investigate a third data case. In case 3, markers 6-8 were eliminated from the data available for processing. In this case there was always at least one marker simultaneously visible to the two cameras.

In each case the data was processed using both the Direct 6 DOF algorithm (A) and Schut's algorithm (B). Superimposed plots of the resulting rotation information for cases 1-3 are presented in figures B.3 through B.5, respectively.

The results of case 1 indicate close agreement between the two algorithms except at the extreme positive rotation angle. It was observed that in this area, only one facet (facet 1) was visible to the two cameras. The sudden jump in the Schut algorithm trace (B) at 0.45 and again at 1.65 seconds indicates the transition to this state. It is possible that

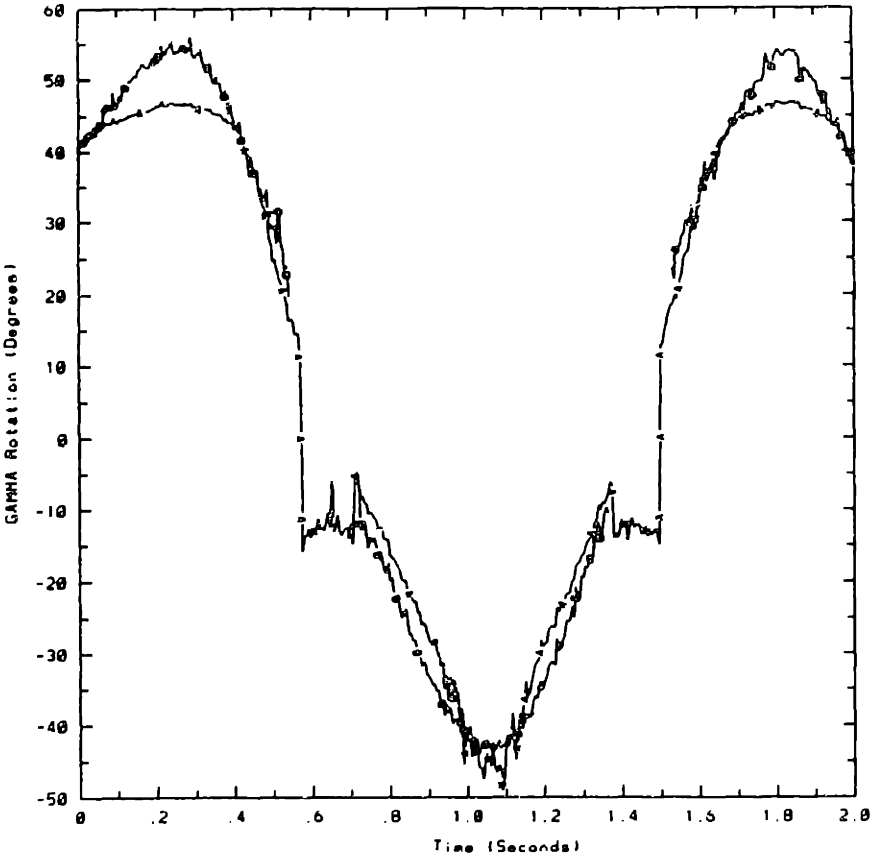


Figure B.4: Case 2: Markers 5-8 not processed. D6DOF (A), Schut (B)

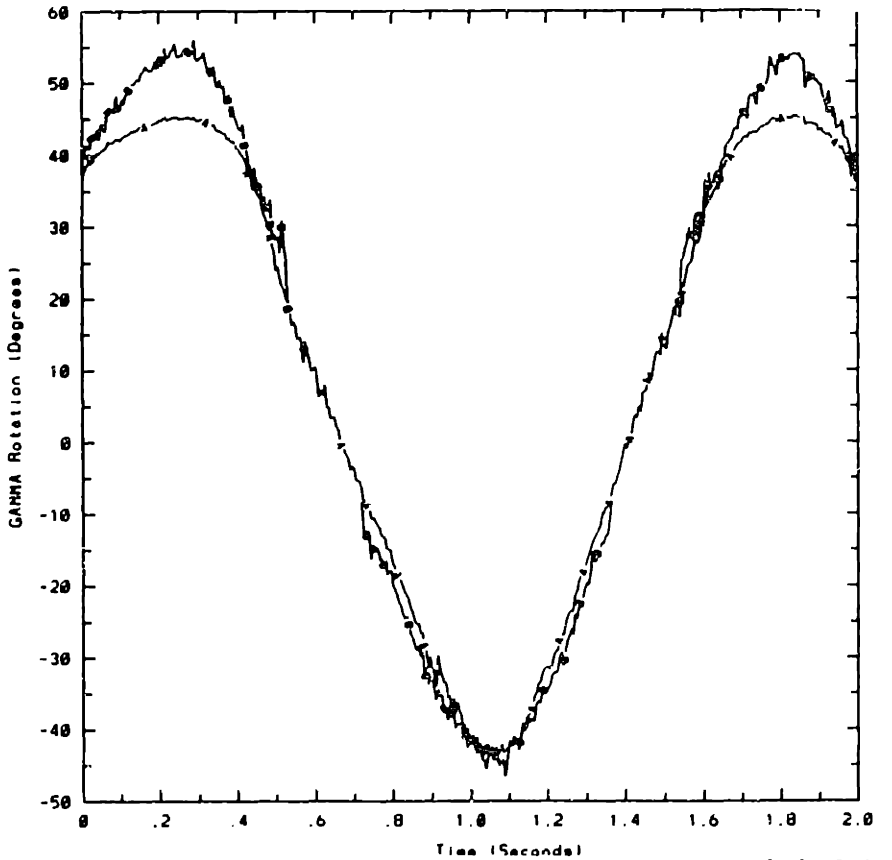


Figure B.5: Case 3: Markers 6-8 not processed. D6DOF (A), Schut (B)

errors in the marker body coordinates are contributing to this problem, but at this time the source of this discrepency has not been positively identified. Note that the D6DOF trace (A) appears somewhat smoother than the Schut trace (B). This is probably due to the greater number of image data points used in the D6DOF solution having a greater averaging effect on random image data noise.

In case 2, during the intervals of 0.52 to 0.72 seconds and 1.35 to 1.52 seconds no more than two markers are viewed simultaneously by both cameras. During these intervals the Schut algorithm fails to find a solution at all as it requires at least 3 markers seen simultaneously by both cameras (see Chapter 2).

The D6DOF method finds a valid solution up until the point where no cameras see the same marker. At this point the D6DOF method results in poor data reconstruction. This problem has been attributed to the condition that facet sizes are small relative to displacement from the cameras resulting in nearly parallel colinearity rays from a particular camera. This allows the solution to rotate and slide along the rays from each camera while maintaining a reasonably good fit. In this experiment this problem would manifest itself as error in the y-rotation and x-displacement estimates. This was verified with plots of other angles and displacements showing reasonable estimates over the same interval, except for x-displacement which, along with y-rotation, was poorly estimated in this interval.

Case 3 was studied as a result of observations made in case 2. With a minimum of one marker simultaneously visible to two cameras the D6DOF method provides a reasonable estimation even in regions where the Schut algorithm fails. Following the explanation in

the discussion of case 2 above, it appears that 1 commonly viewed point between two cameras is enough to lock in the position and orientation of the object.

B.4 Discussion

Based on the results described above, it appears the D6DOF method may be a reasonable algorithm to apply in situations where markers are occasionally hidden from a camera's view. For example, in human gait studies the hand and forearm often interfere with data acqusition from structures mounted on the pelvis. The D6DOF method would include data from a particular camera in its solution even when the other camera could not view that marker.

Also the potential use of marker structures designed for use in large angular excursion applications is much greater with D6DOF. A structure such as the one used in the above experiment would only require two markers per facet thereby reducing the number of markers required per structure.

There are also limitations to the D6DOF method. Techniques which compute the 3D object space position of markers are capable of checking for bad data (see Chapter 2). The inability of D6DOF to make these same checks could, under some conditions, cause problems. This is especially true for active marker systems such as SELSPOT which are notoriously prone to spurious data from marker reflections.

In terms of processing speed the D6DOF method requires roughly an order of magnitude longer to process data than the Schut algorithm. The D6DOF method is however roughly 2 orders of magnitude faster than a typical nonlinear numerical search algorithm.

This indicates the D6DOF method would not be the method of choice for real-time processing applications, but would be reasonable for off-line processing applications.

Further work with this method should include an analysis of what causes the sensitivities seen in the data. It was noticed that the D6DOF algorithm had difficulty in converging on a solution when the available markers all resided in a common plane. A mathematical error analysis of the algorithm would provide insights to this and other potential problems.

It would also be interesting to see how the D6DOF method would react to information from an additional camera having its principal axis oblique to the plane formed by the other two cameras' principal axes. This should eliminate the sensitivity described in case 2 above and the additional data would serve to smooth the results.

Appendix C

Hardware and Software Background

This appendix provides an overview of the hardware features and real-time data acquisition and processing software developed for the Large Volume TRACK (LVT) system. The first section gives an overview of the hardware while the second section provides a description of the multi-processor data acqusition and real-time processing routine used for LVT.

This appendix is not intended to be a detailed description of the system's operation. If you wish to know the system in detail, start by reading this appendix. Then locate the manuals referred to. Finally, examine the author's code (not listed in this document). Learning from the manuals how the VMEbus and VMEbus devices work and communicate will make sense of notations in the code.

C.1 Hardware Components

This section describes the hardware components of the LVT system. The block diagram in figure C.1 provides an overall layout of the various components in schematic form.

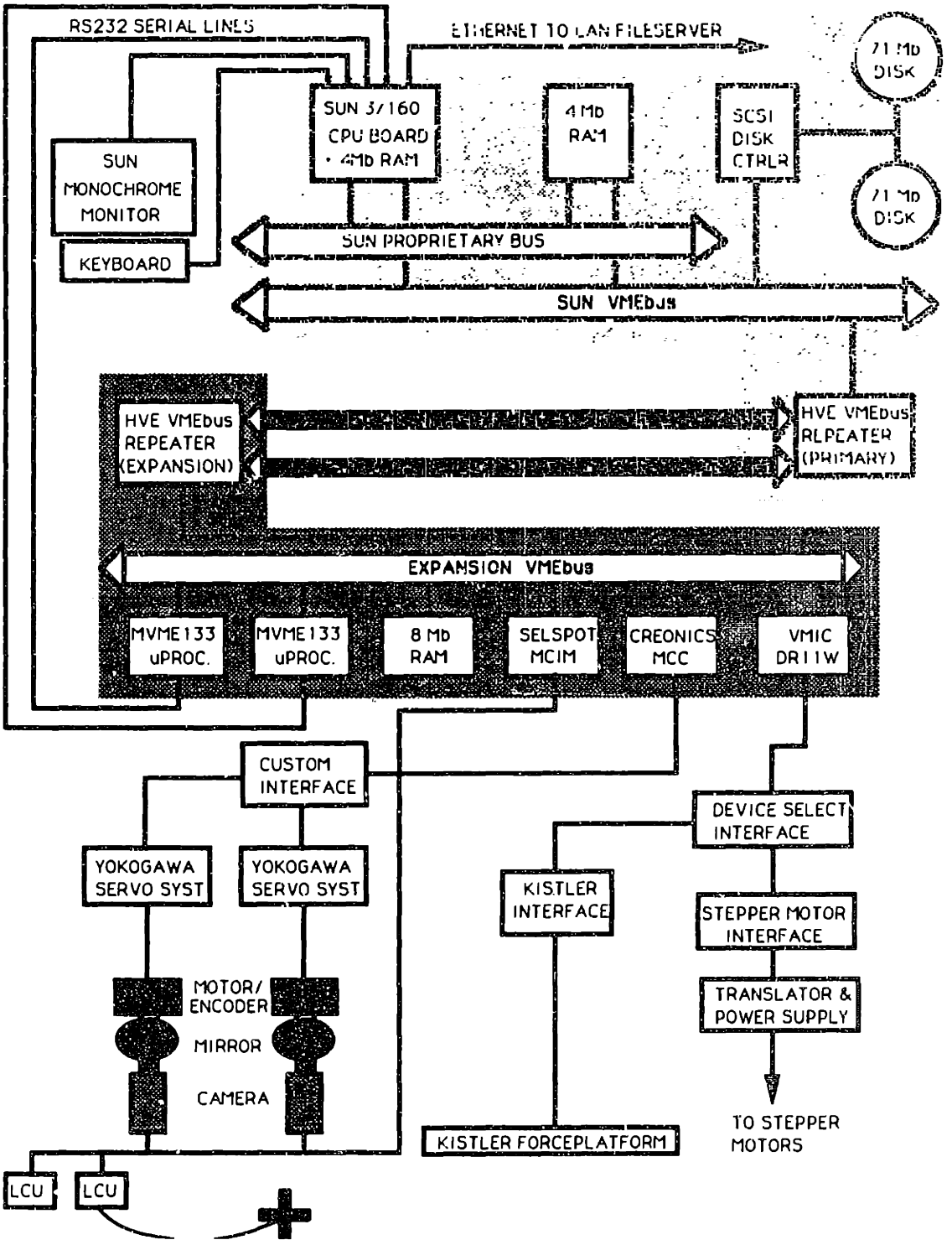


Figure C.1: Block diagram of Large Volume TRACK hardware

The host processor for the LVT system is a SUN 3/160 microcomputer. The 3/160 features a Motorola 68020 microprocessor with a 68881 floating point coprocessor, both running at 16 MHz, and 8 megabytes of RAM. In addition, two local 71 megabyte disk drives provide local file storage which augments storage available on the local area network fileserver. The Sun's operating system is Unix based. A 19 inch monochrome monitor permits viewing multiple windows with SUN's Sunview window environment.

Data acqusition and real-time processing are handled through a VMEbus expansion chassis. The expansion chassis communicates with the SUN through a pair of HVE Repeater 2000 VMEbus repeater cards. The repeater cards are configured to pass all VMEbus information between the two systems with the exception of interrupts.

Two microprocessor boards (Motorola MVME133 and MVME133-1) comprise the processing capabilities on the expansion chassis. Each board features a Motorola 68020 microprocessor with a 68881 floating point coprocessor, both running at 16.67 MHz (133-1) or 12.5 MHz (133), and one megabyte of dual-ported RAM. With an additional 8 megabytes of RAM on the expansion chassis, these processors handle the real-time data acqusition and processing demands of the LVT system.

Each processor carries a set of Motorola MVME133Bug EPROMs which "boot" a low level operating environment. Serial port communication to the two MVME133Bug monitors is accomplished using both SUN serial ports (ttya and ttyb). Kermit software on the SUN allows direct communication to each processor through a dedicated window in the Sunview window environment. This allows the user, among other options, to open and inspect memory locations, dissasemble programs in memory, and execute programs

loaded in dual-ported memory from the SUN.

Data from the SELSPOT cameras are obtained through the SELSPOT MCIM interface. During data acqusition the MCIM uses a Motorola MVME333 Serial Communications Board to perform DMA writes to VMEbus memory.

The Yokogawa Servo System, driving the mirrors, is controlled and interfaced to the VMEbus with a Creonics VMEbus Motion Control Card (MCC). The MCC accepts high level control commands and performs feedback and loop stabilization operations internally and transparently to the user.

The Kistler Forceplatform and Stepper Motor controller were originally designed to communicate with the PDP11/60 through DR11B and DR11C communications boards, respectively. With the implementation of these devices on a VMEbus it was desireable to avoid large amounts of interface reconstruction. A solution was found in the VMIC DR11W board which is capable of emulating all the functions of the DR11B and DR11C boards. Since the Kistler Forceplatform and Stepper Motor Controller are never used simultaneously, a device selection interface was designed and constructed to allow both peripherals to communicate through the same board.

A list of relevant hardware manuals follows:

- The VMEbus Specification Rev C.1
  (Motorola number HB212).
- Hardware Installation Manual for the SUN-3/160 SunStation
  (SUN Microsystems number 800-1314-01)

- Cardcage Slot Assignments and Backplane Configuration Procedures (SUN Microsystems number 813-2004-14).
- MVME133 VMEmodule 32-Bit Monoboard Microcomputer User's Manual (Motorola number MVME133/D1) and related MC68901 Multi-Function Peripheral (Motorola)
- MVME133BUG 133Bug Debugging Package User's Manual (Motorola number MVME133BUG/D1).
- VMEbus Repeater 2000 VMEbus High Speed Repeater System (HVE Engineering, Inc. product specification AS90172000A)
- Selspot MCIM User's Manual (SELSPOT) including Update (SELSPOT) and related MCIM hardware components:
  - MVME333 Intelligent Communications Controller User's Manual (Motorola number MVME333/GD2)
  - MC68010 16-Bit Virtual Memory Microprocessor (Motorola number ADI-942-R1).
  - MC68450 Direct Memory Access Controller (DMAC) (Motorola number ADI1216).
  - AmZ8030/8530 Serial Communications Controller (Advanced Micro Devices).
- VMIVME-DR11W VMEbus to DR11W Interface, Instruction Manual (VME Microsystems International Corp. number 500-000DR11-00).

- VMEbus Compatible Motion Control Card, User's Manual
  (Creonics number 999-041).

- DRAM16-16M 16 Mbyte VMEbus Dynamic RAM Memory Module User's Manual
  (VME Specialists number DRAM16-16M/D).

- Dynaserv Direct Drive Servo Actuator Instruction Manual
  (Yokogawa Precision number IM A101-03E).

C.2 Software Development and Operation

This section describes the software features of the LVT system. Most of the discussion will be general in nature with the exception of a detailed view of the software behind the real-time data acqusition and mirror servo control used with LVT.

C.2.1 LVT Software Background

As described in 1.4 TRACK has had a history of development spanning several Master's Theses and two Doctoral Theses over a period of fifteen years. During this time computer hardware and programming languages of preference have changed. Prior to the current implementation of TRACK, the prefered programming language on the DEC PDP machines was FORTRAN in its variety of forms (FORTRAN-4, FORTRAN-4PLUS, FORTRAN-77). Assembly level routines, required for real-time programming of TRACK peripherals, were written out of neccessity in DEC's MACRO-11 language.

With the introduction of the Unix based SUN System as the host computer, the

language of preference naturally changed to C (Unix is written in C). Aside from its ability to work well with the operating system, C brings the high level programming efficiency of FORTRAN together with many of the features that previously forced programs to be written in an assembly language (eg. direct register addressing for peripheral hardware programming).

Programming the expansion chassis processors also proved to be relatively straightforward to do in C. Patrick Lord, as a MIT design class project, developed a special set of compiling commands for use in writing and executing programs on the Motorolla MVME133 processors. These programs may be downloaded from the SUN to dual-ported RAM on the processor boards and executed there using commands provided with the MVME133Bug as explained in the hardware section above.

C.2.2 Real-Time Software Example

Real-time programming of the LVT system requires a substantial knowledge of the operation of each of the peripheral devices and their intercommunication over the VMEbus. The following example illustrates how the SUN processor and two expansion MVME133 processors may be configured and programmed to realize high frequency data acqusition from the SELSPOT and Creonics MCC (for mirror axis position),with simultaneous estimation of mirror position error for real-time mirror control.

A unique piece of software is driving each of the three processors. In the case of th SUN processor the software was written in C and compiled using the standard cc command in Unix. The primary functions of the SUN software are to

- Set up the sampling parameters
- Begin single cycle sampling and mirror tracking to ensure mirrors are on target when high frequency sampling begins.
- Initiate high frequency sampling.
- Wait for flags on the expansion VMEbus to be set prompting initialization of a new SELSPOT READ or signaling that all sampling is complete.
- Clean up after sampling is complete or aborted by deactivating devices and writing memory buffers in expansion VMEbus RAM to disk.

These functions are described in the flow diagrams of figures C.2 and C.3.

The two expansion chassis processor boards are refered to in the software as M133A and M133B. M133A, the MVME133-1 processor, is the only device configured to service interrupts from devices on the expansion VMEbus [1]. To ensure quick interrupt servicing and to facilitate interrupt vector loading, the software running on M133A is written entirely in the SUN assembly language for the Motorolla 68020 processor. Software for M133B is written in C and compiled using the modified compiler refered to above.

The primary function of M133A is to service interrupts from the SELSPOT MCIM, and DR11W (if active) thereby performing the synchronous operations of the real-time system. Flow diagrams of the software operations are provided in figures C.4 through C.6.

[1]The HVE Repeater 2000 is configured to *not* pass interrupts back to the SUN VMEbus chassis. This avoids the complexity of writing device drivers for real-time proceedures in a Unix environment.

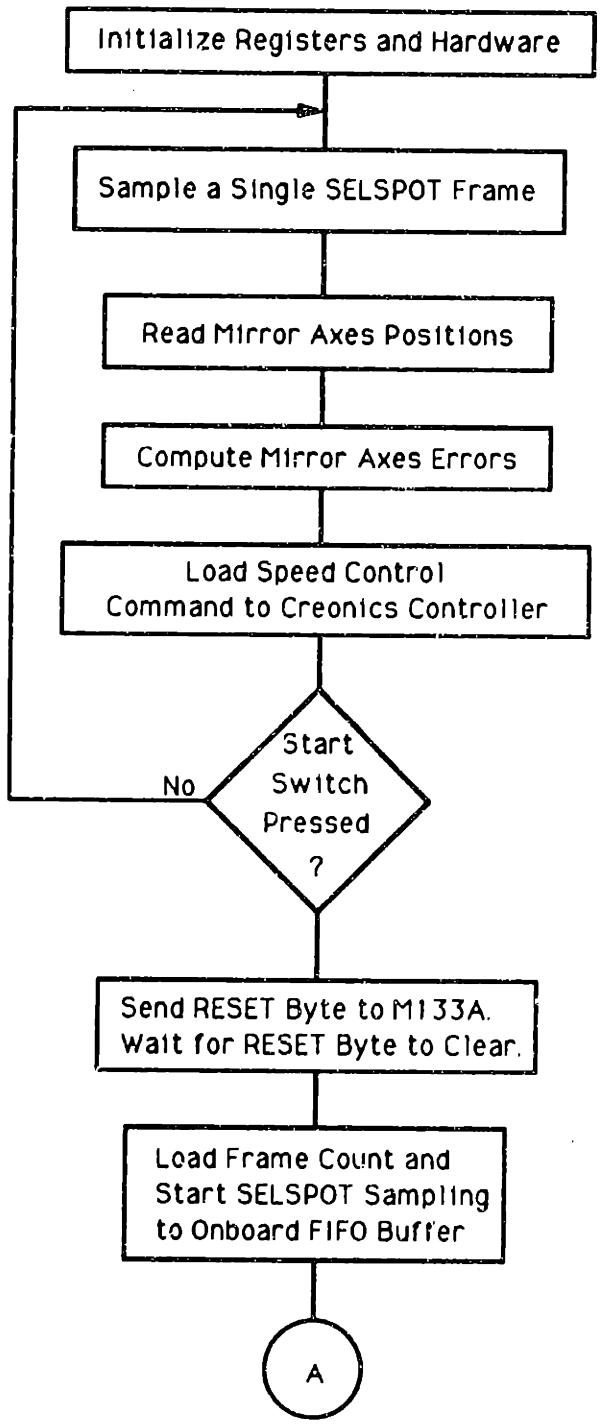


Figure C.2: SUN flow diagram Part 1: System initialization.

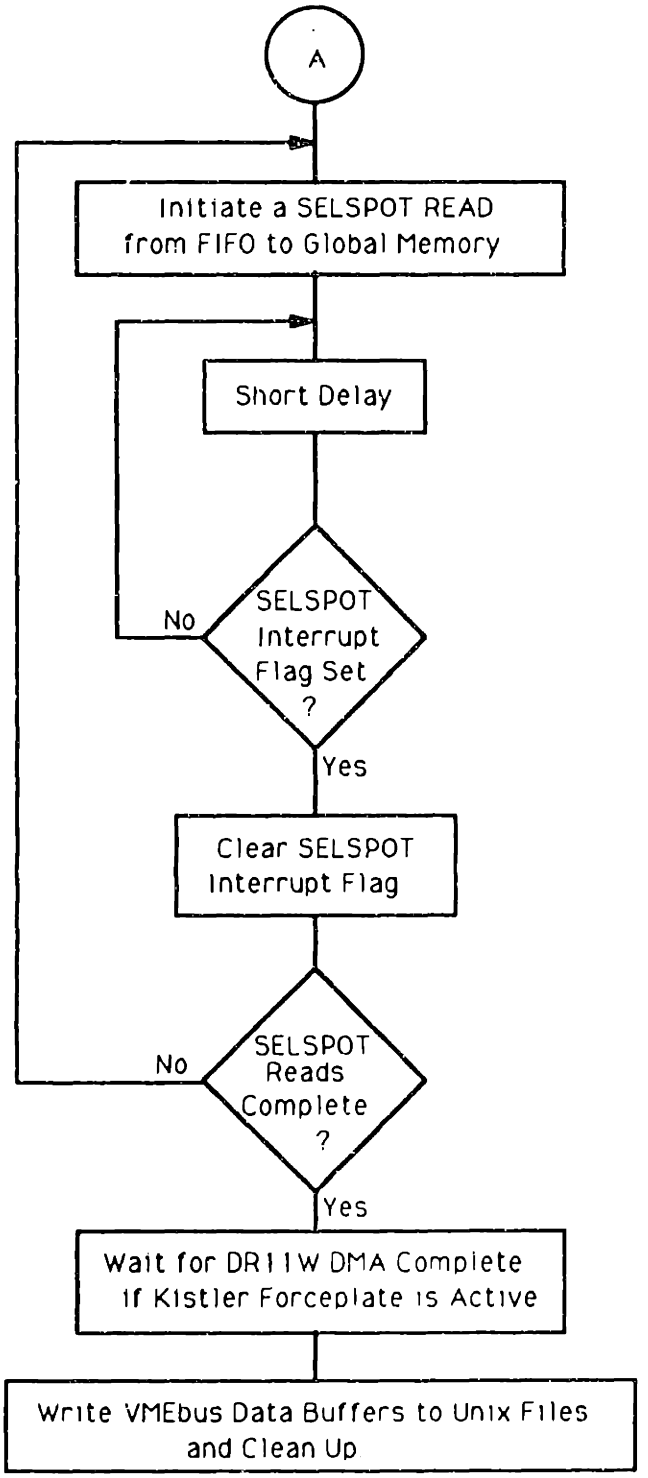


Figure C.3: SUN flow diagram Part 2: Sampling and clean up.

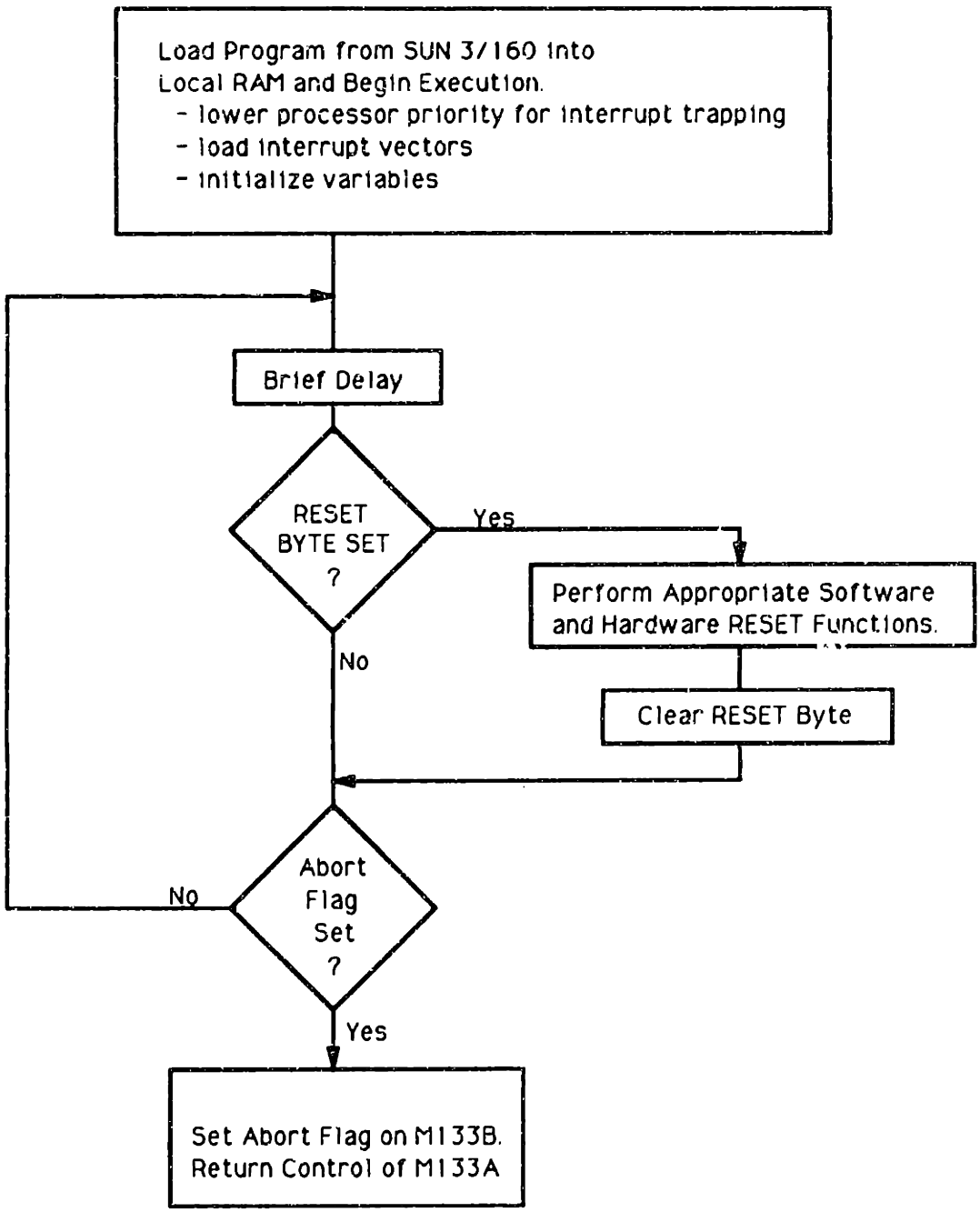


Figure C.4: Flow diagram of background software operation on M133A

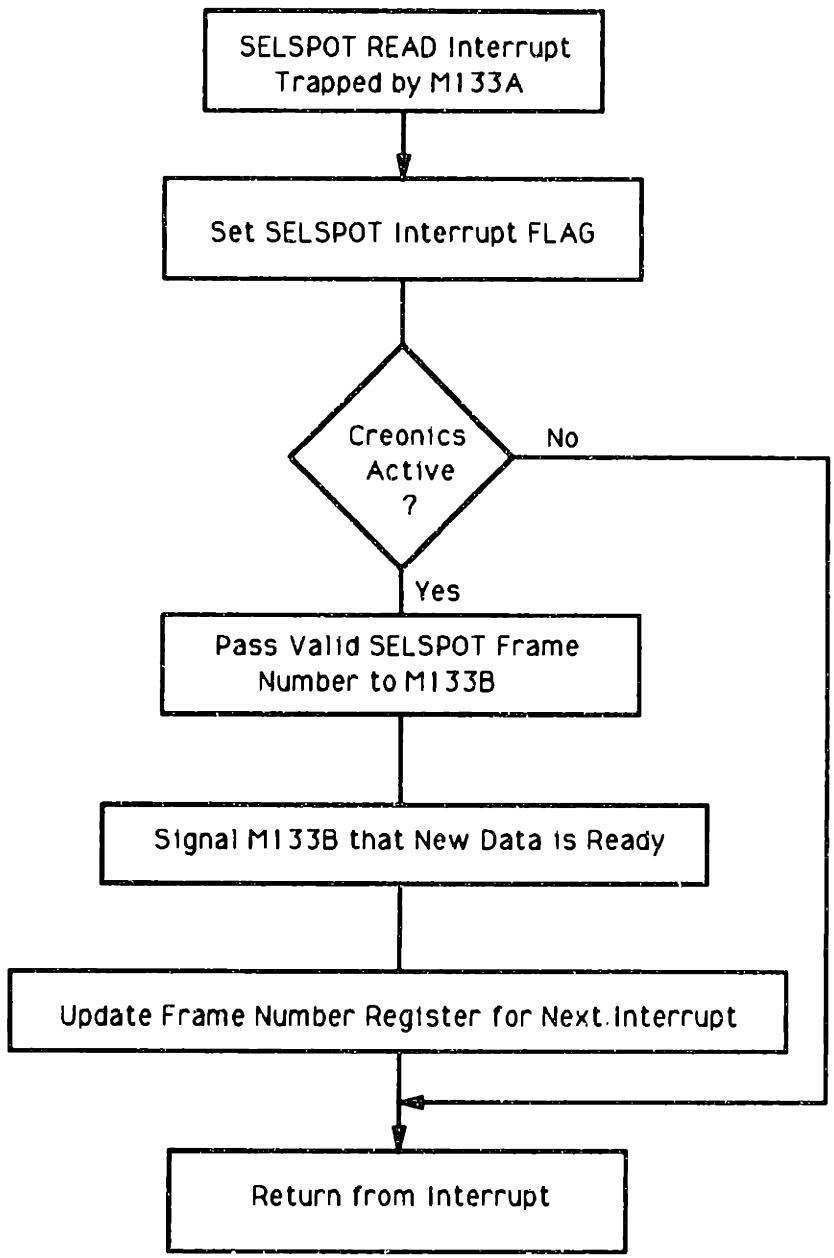


Figure C.5: Flow diagram of slow interrupt loop operation on M133A

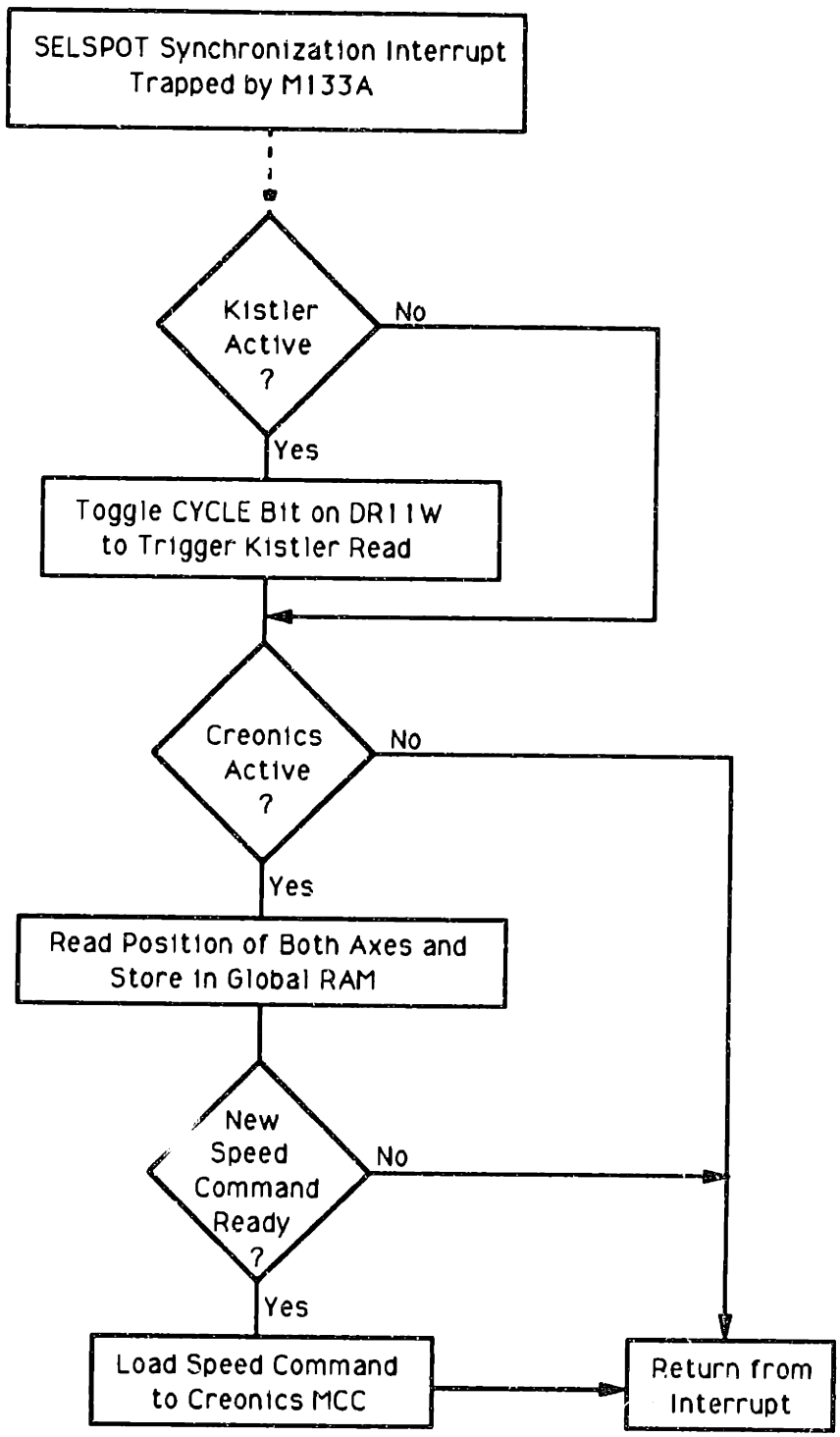


Figure C.6: Flow diagram of fast interrupt loop operation on M133A

A background routine on M133A performs initialization operations whenever a reset flag is set by the SUN. In addition it watches for SUN signaled abort flags in which case the abort is cascaded to M133B.

Upon M133A program startup, two interrupt service routines are loaded. The corresponding interrupts are both generated by the SELSPOT MCIM. A fast loop is generated by the SELSPOT synchronization interrupt to synchronize data acquisition from the Creonics encoder registers (mirror positions) with the SELSPOT frame rate. The interrupt service routine executes a Creonics postion read command and writes the response information to memory. If new mirror position information is ready from M133B then the appropriate speed control commands are sent to the Creonics MCC. Note that this is the only loop which executes real-time commands on the Creonics MCC. This is to avoid the problem of simultaneous access of the MCC by two different VMEbus devices.

A slower loop (approximately 20 Hertz) signals completion of a SELSPOT DMA write to memory of the SUN requested number of frames. The interrupt service routine raises a flag for the SUN and signals M133B that new data is available in memory for real-time processing.

M133B is programmed to perform the asynchronous task of computing the current mirror position error and associated mirror speed command value and passing the information back to M133A. A flow diagram describing this operation is provided in figure C.7. When a new data ready flag is sensed on M133B, an error computation loop is entered on M133B to generate a new mirror position error value and an updated speed command. Once the new command is available it is loaded

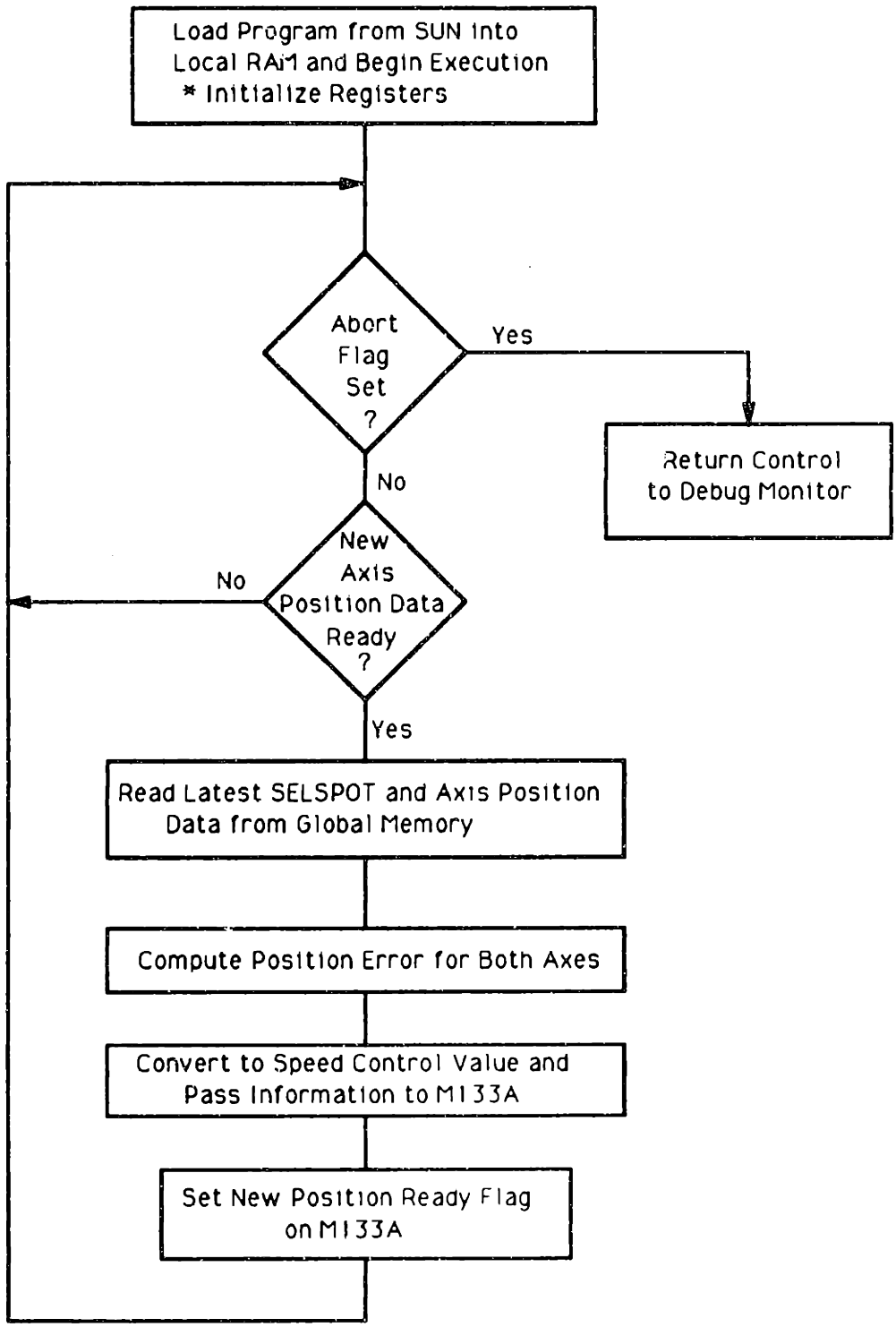


Figure C.7: Flow diagram of software operation on M133B

onto M133A's local RAM and a flag is raised signaling M133A that a new speed command is ready.

C.3 Future Hardware and Software Improvements

The hardware and software components described thus far are extremely effective in accomplishing the required real-time tasks. However, if changes to hardware or software are necessary or if new applications require development, the time to write, modify and debug programs can be great. This is largely due to the requirement that the programmer must know what all devices are doing, particularly with respect to VMEbus cycles, so that system performance may be optimized and to avoid bus overloading. With the three processor example given, this task was not too burdensome. The addition of another processor would make development much more complicated.

To facilitate application development on VMEbus based multiprocessor systems, third party software development systems are now available. These systems allow the programmer to develop all code in the SUN environment with compilers and debug tools provided for a host of popular VMEbus processor board models. Having ability to access a debug window on a remote processor would be a major improvement to the program development environment.

An early version of such a system, called CONDOR, was developed in MIT's Artificial Intelligence Laboratory as an outgrowth of the Utah/MIT Dexterous Hand project. A number of commercial real-time software products for interfacing to the UNIX environment are also currently available.

Appendix D

Calibration Structure Mechanical Drawings

This appendix presents the mechanical design drawings used to construct the calibration structure. Not all components of the calibration structure required custom construction. Only those components which required custom construction are detailed here. Note that drawing scales have been reduced for printing purposes.

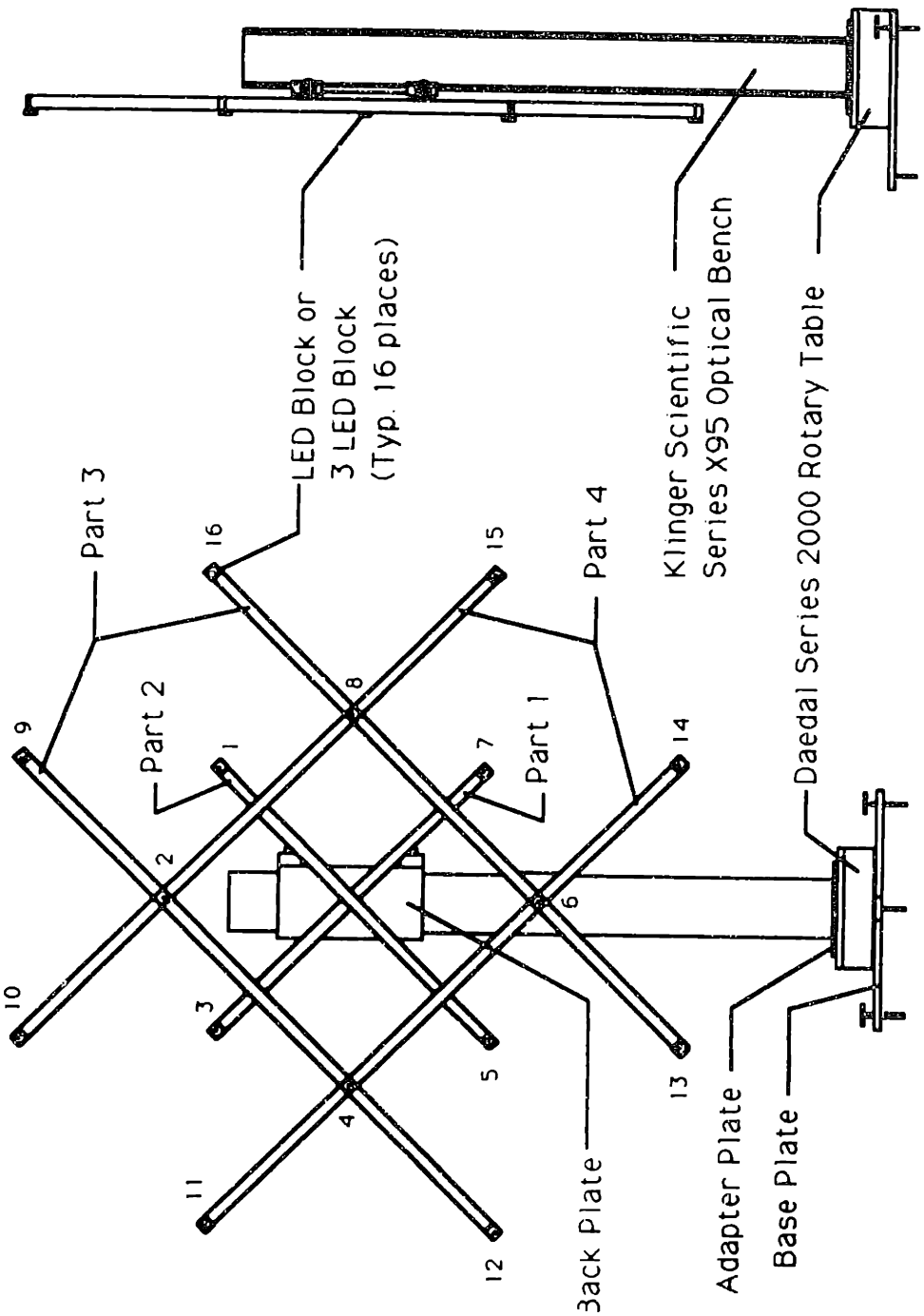


Figure D.1: Calibration structure: Key drawing

Figure D.2: Calibration structure: Part 1

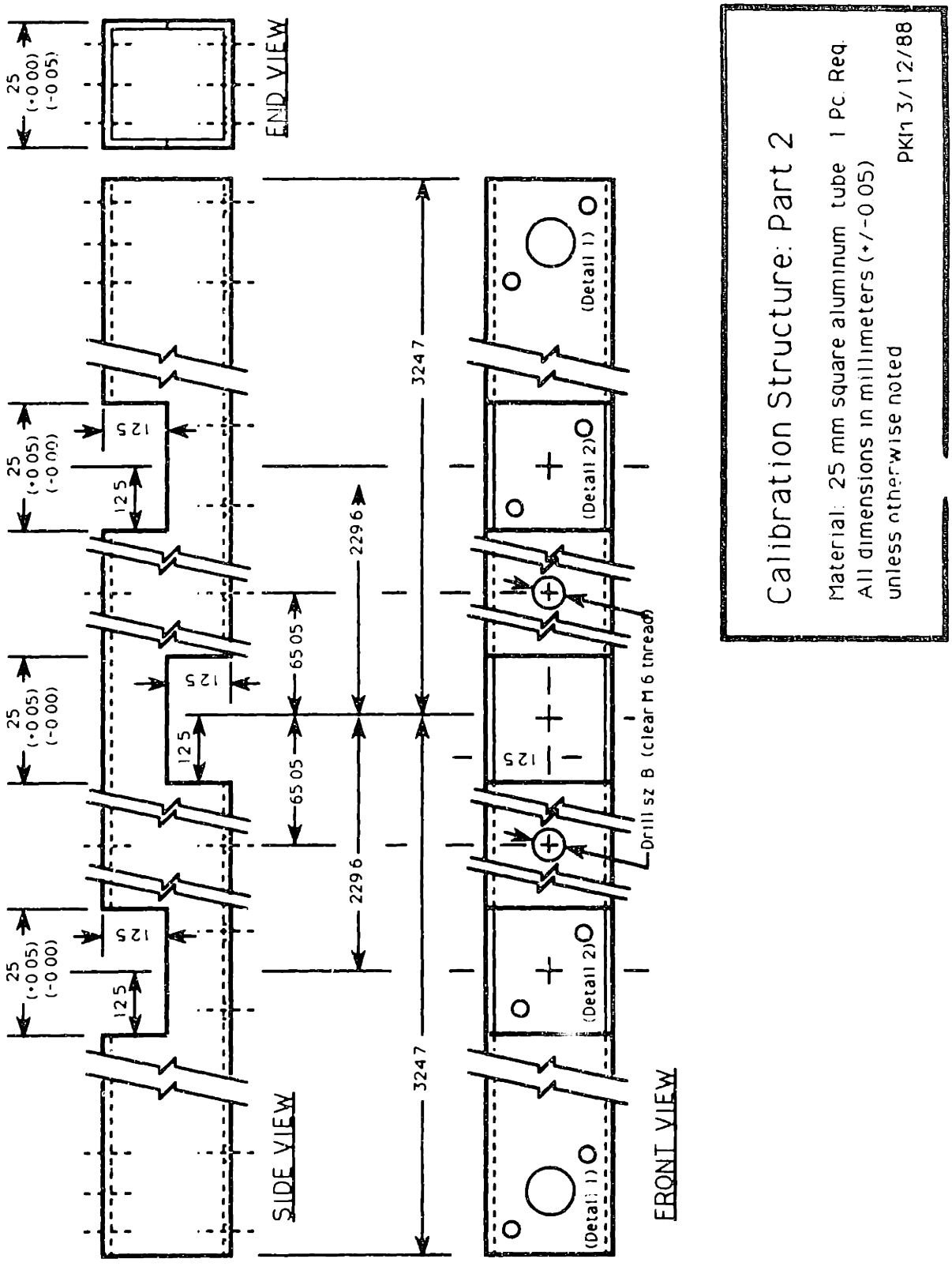


Figure D.3: Calibration structure: Part 2

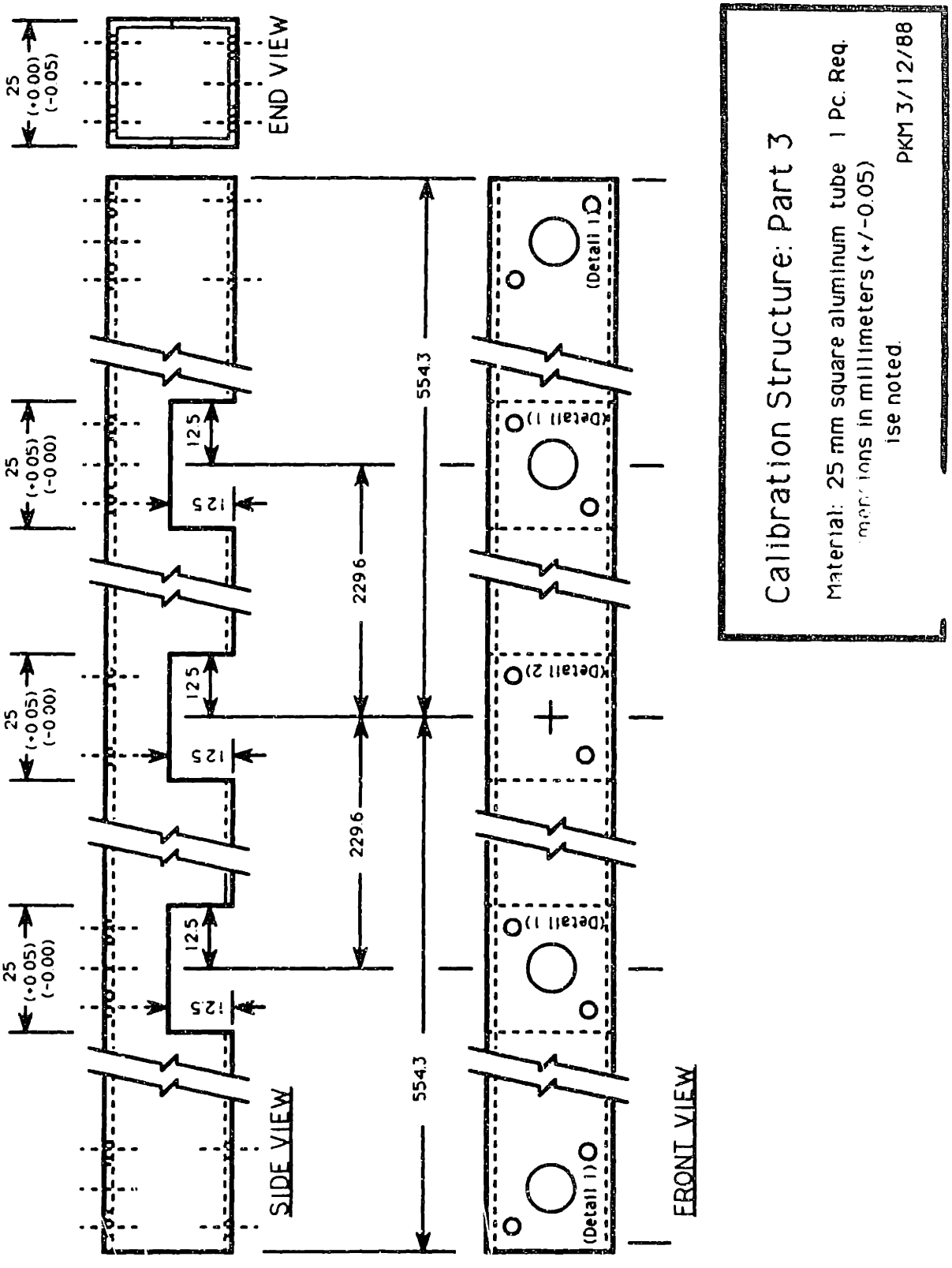


Figure D.4: Calibration structure: Part 3

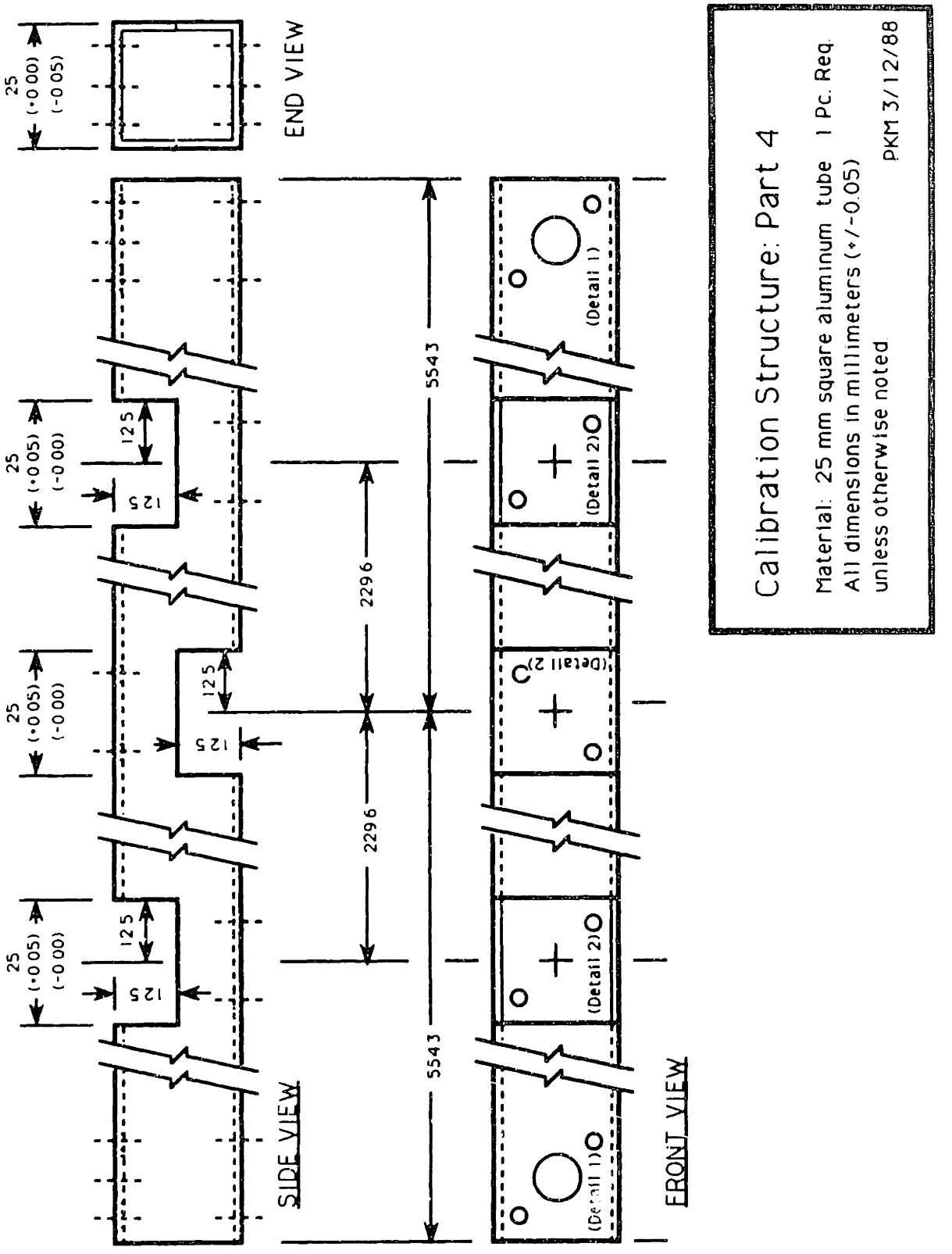


Figure D.5: Calibration structure: Part 4

Figure D.6: Calibration structure: Details

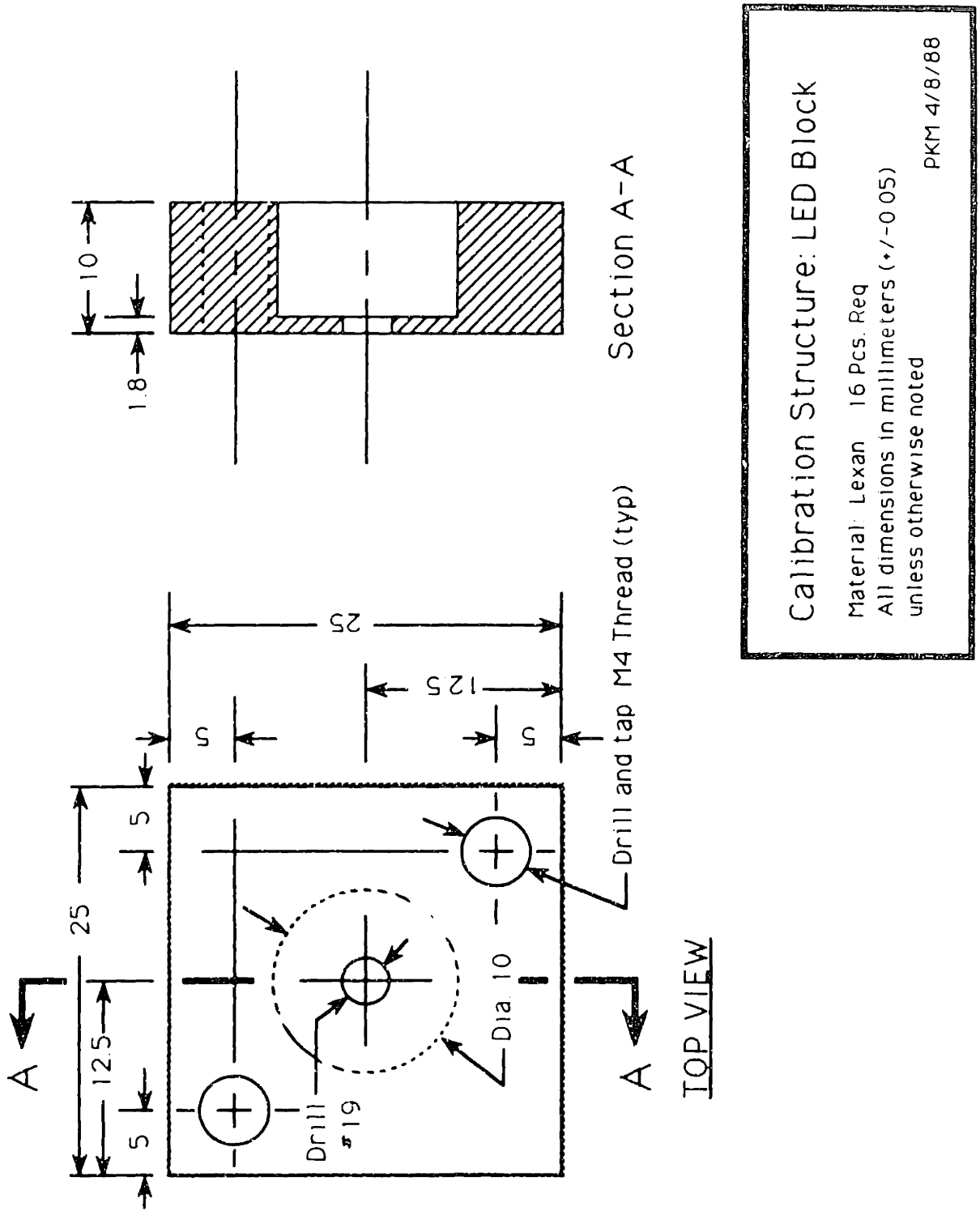


Figure D.7: Calibration structure: LED Block

Figure D.8: Calibration structure: 3 LED Block

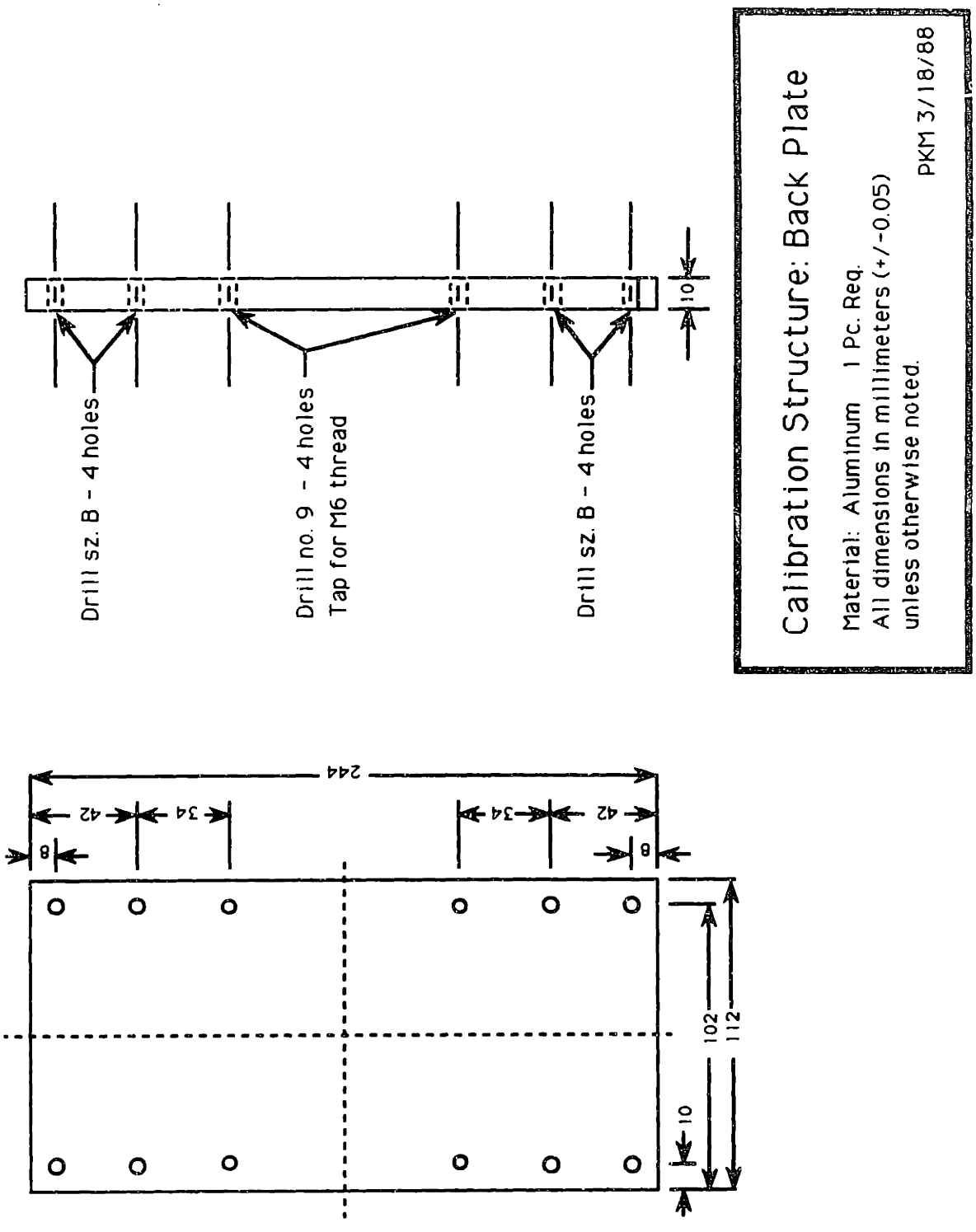


Figure D.9: Calibration structure: Back Plate

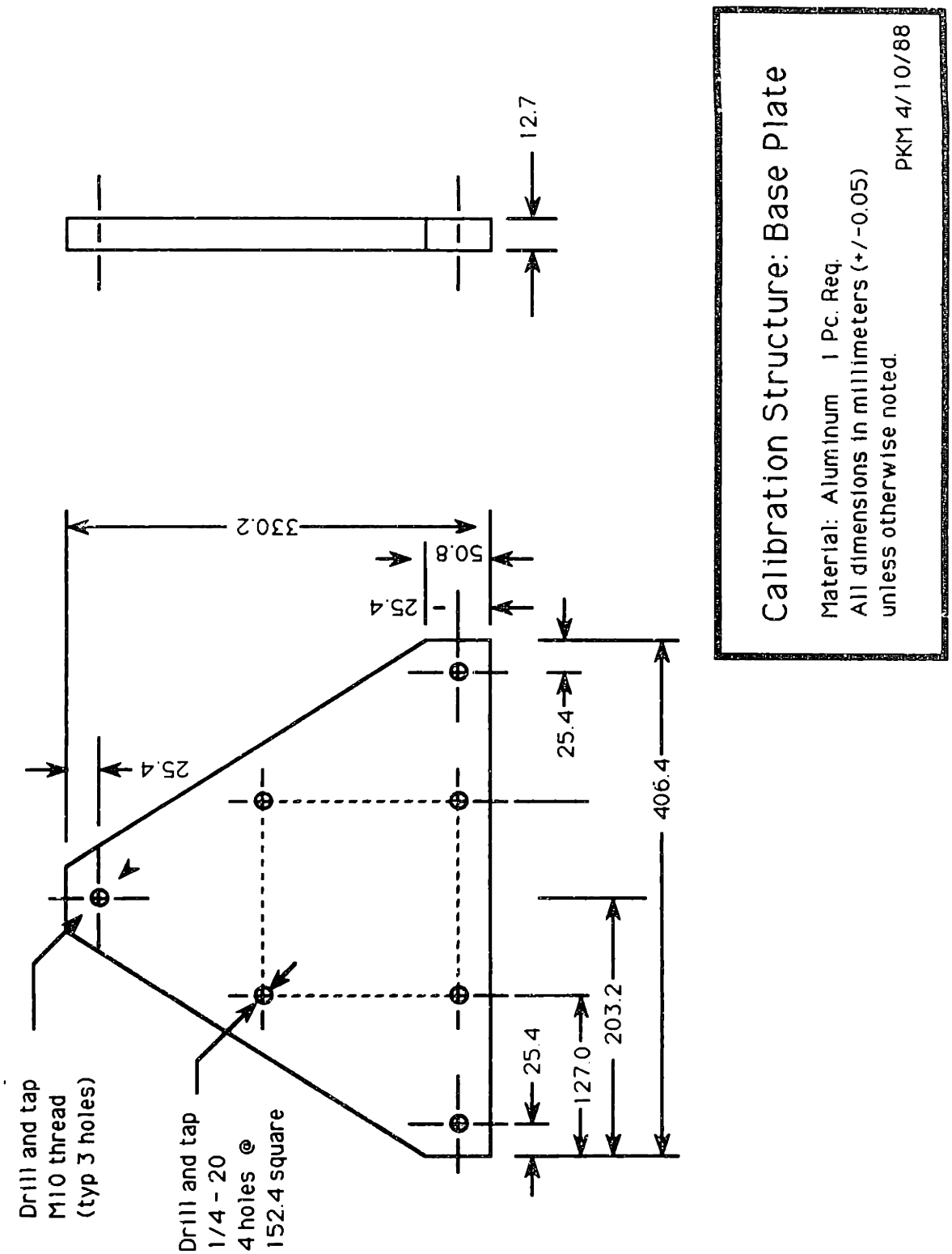


Figure D.10: Calibration structure: Base Plate

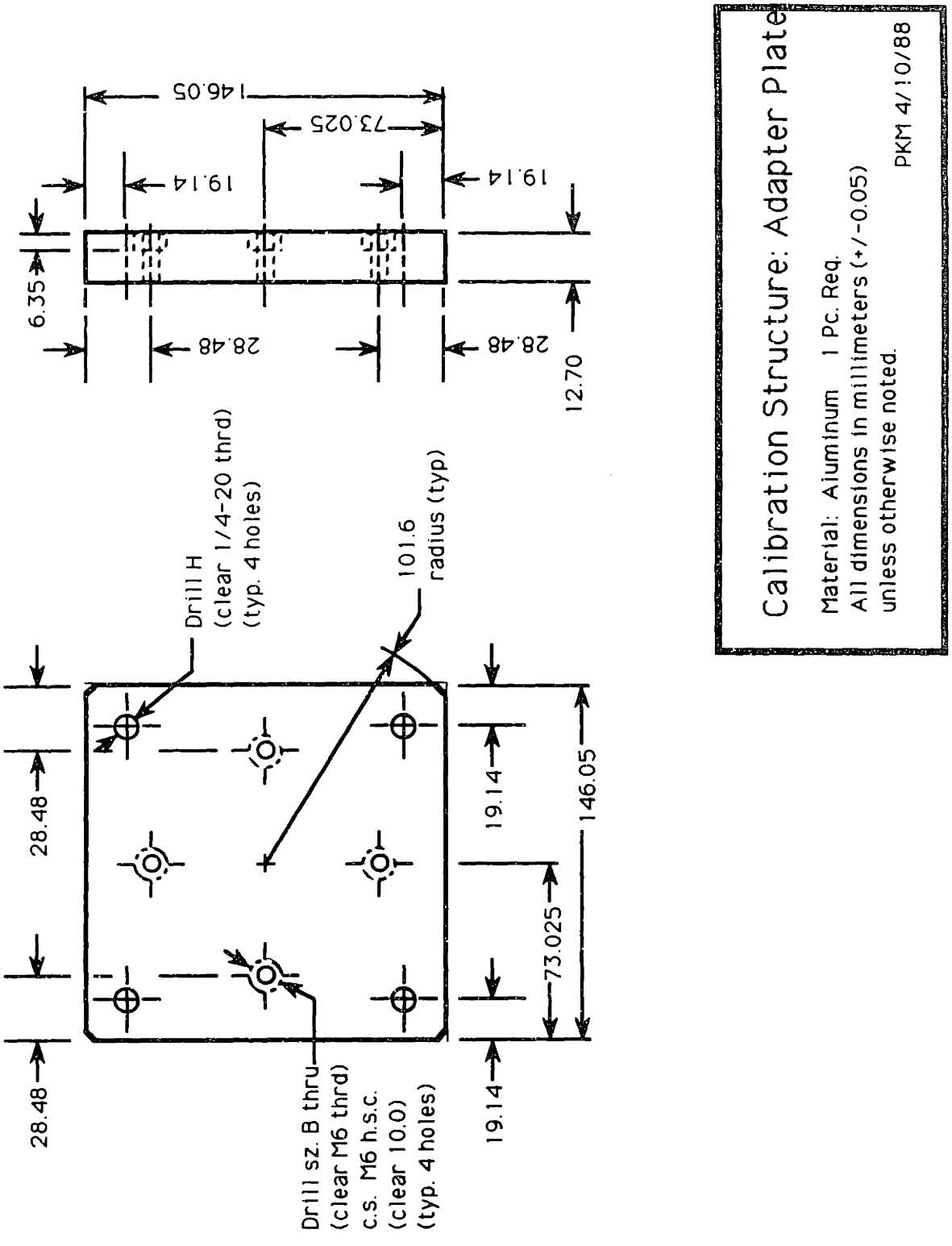


Figure D.11: Calibration structure: Adapter Plate

Appendix E

Mirror/Motor Mount Mechanical Drawings

This appendix presents the mechanical design drawings detailing the mirror/motor mount apparatus. Note that drawing scales have been reduced for printing purposes.

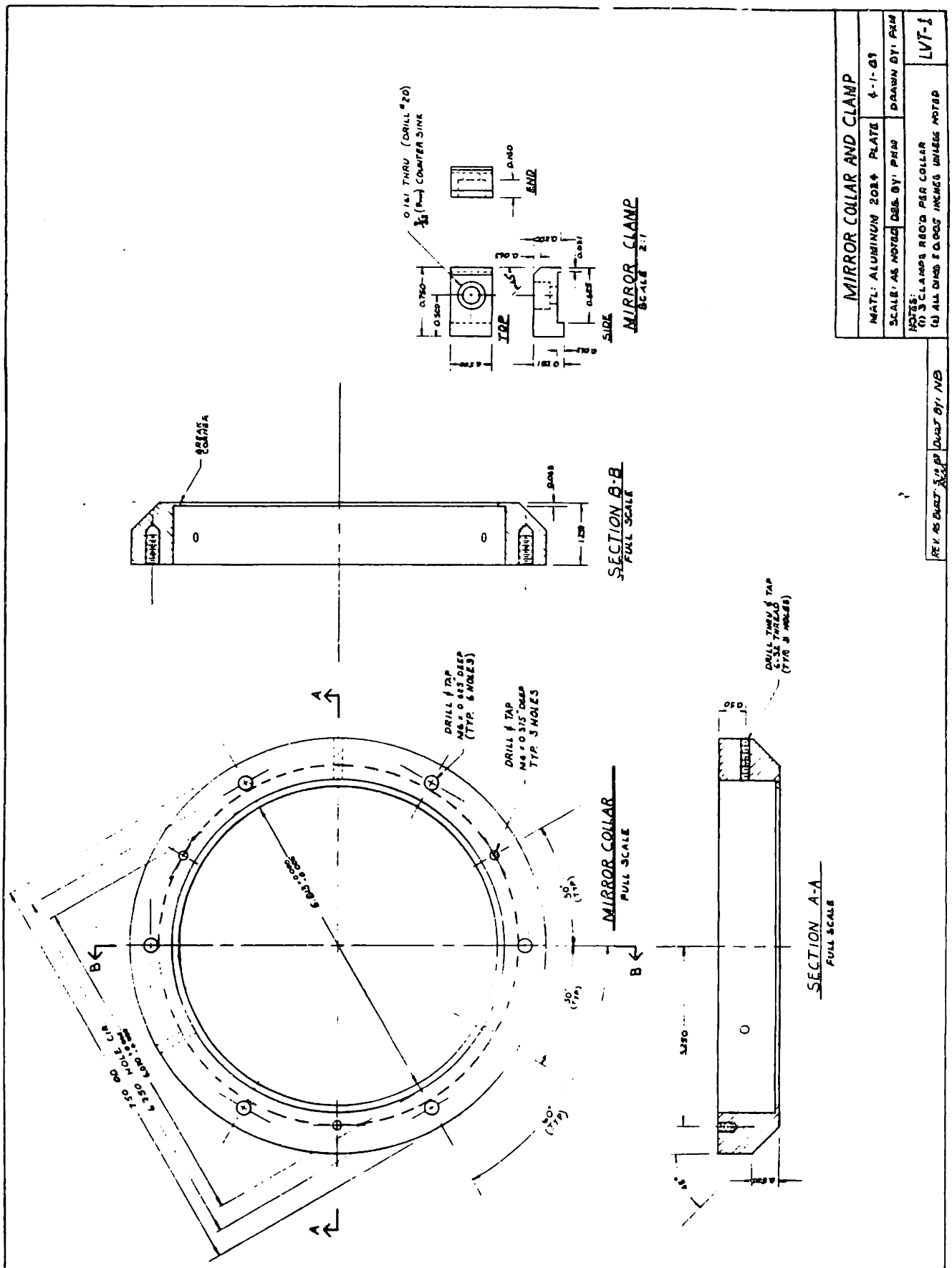


Figure E.1: Mirror collar and clamp (LVT-1)

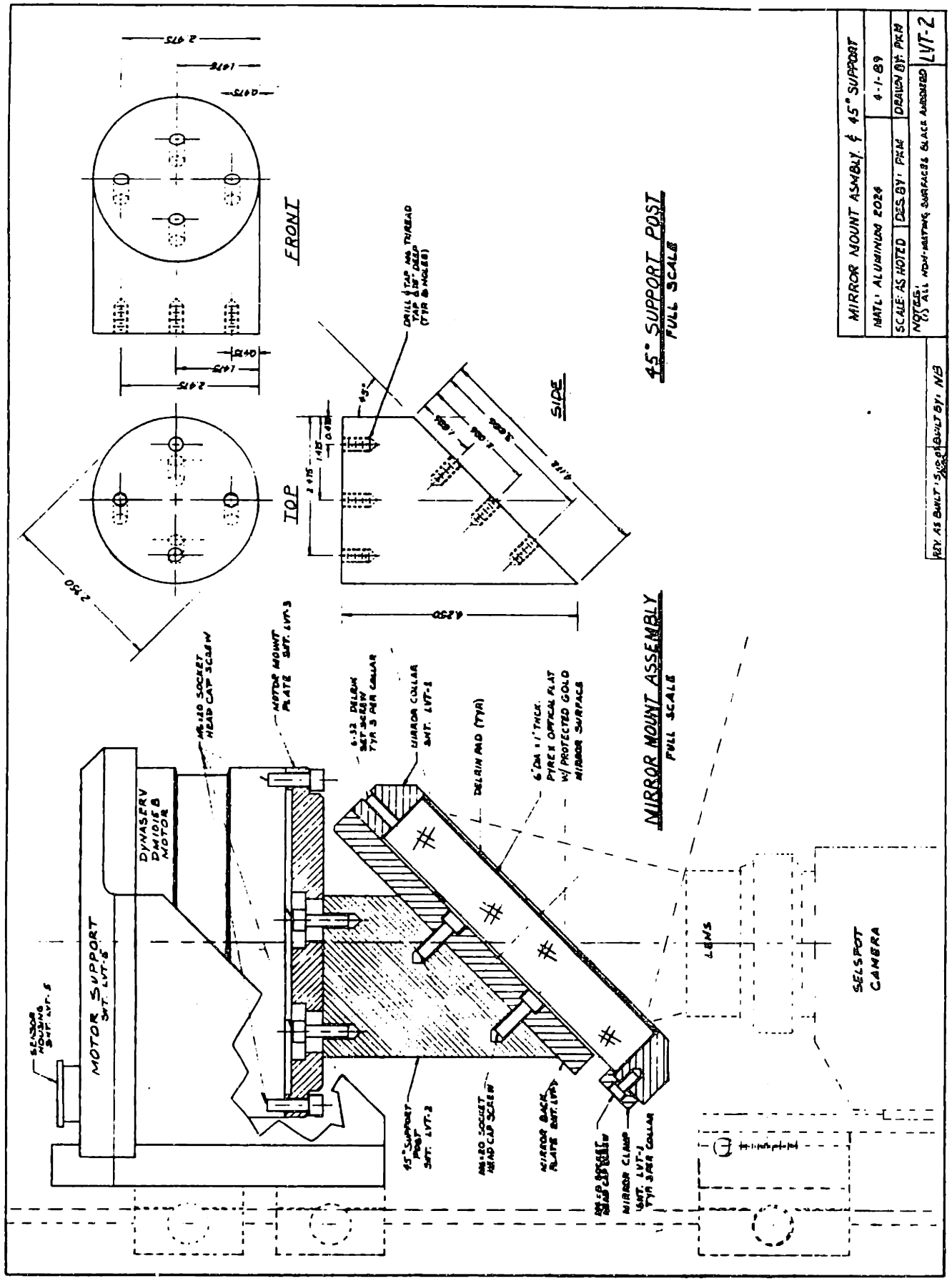


Figure E.2: Mirror mount assembly and 45 degree support (LVT-2)

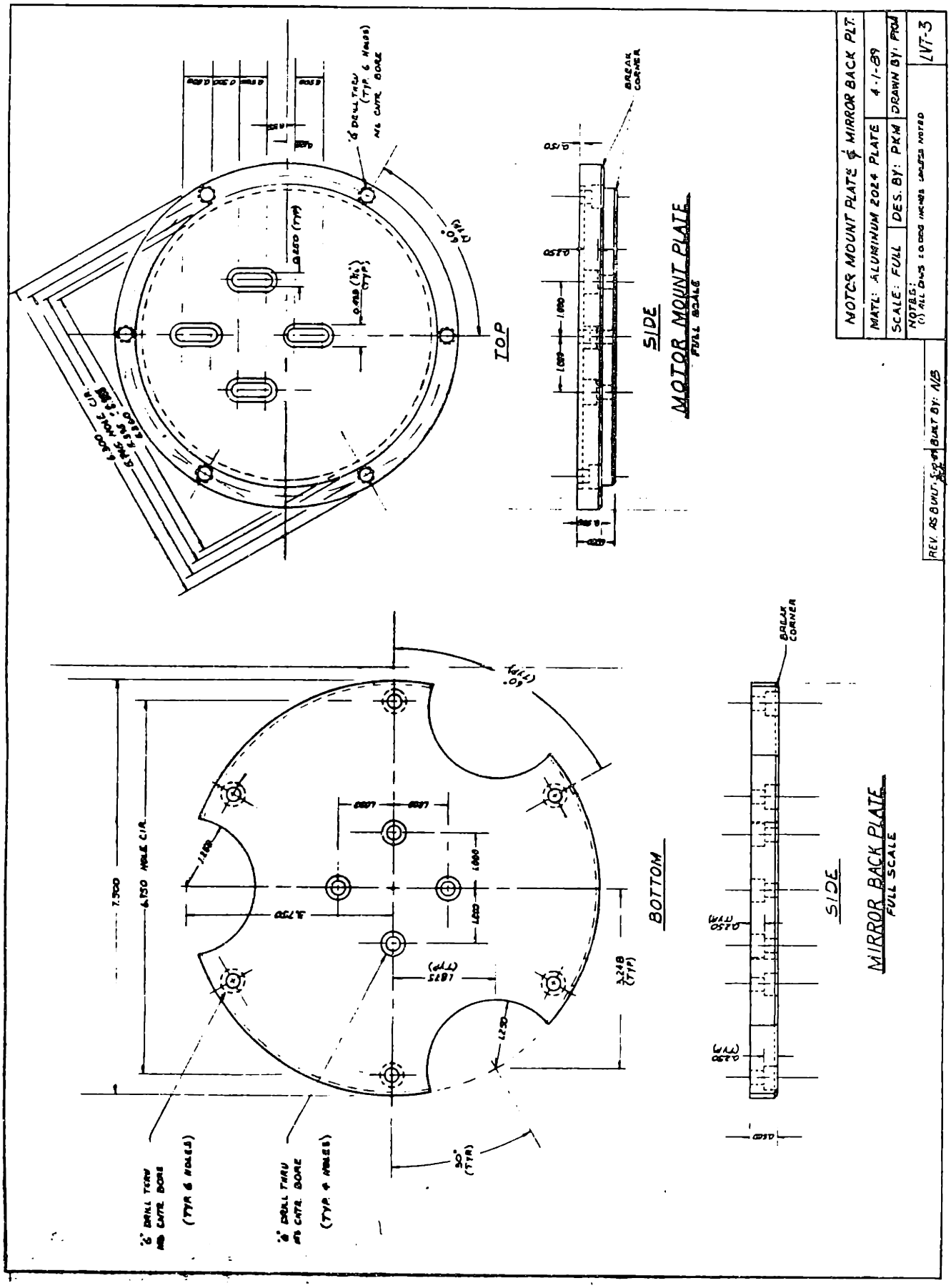


Figure E.3: Motor mount plate and mirror back plate (LVT-3)

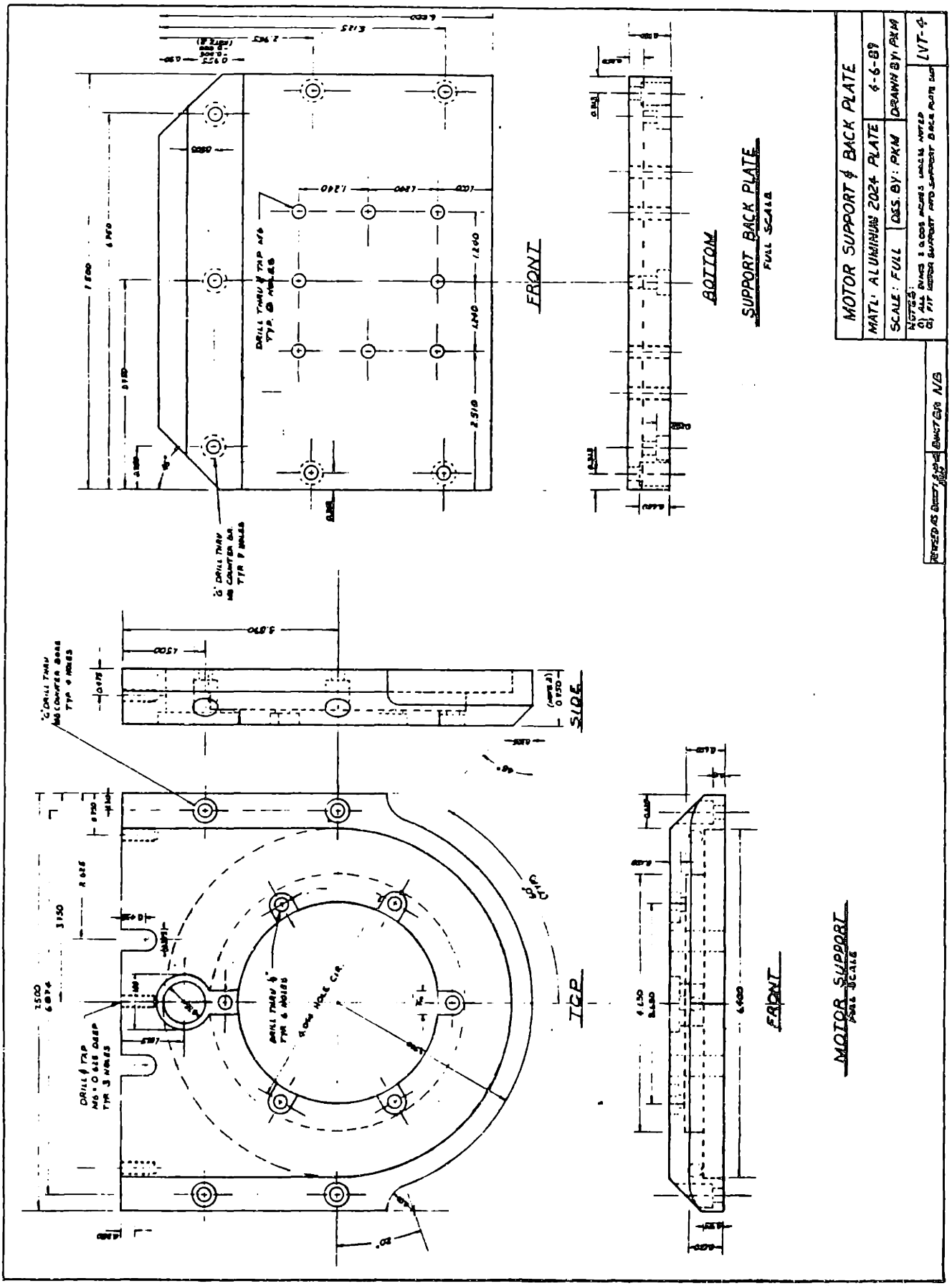


Figure E.4: Motor support and back plate (LVT-4)

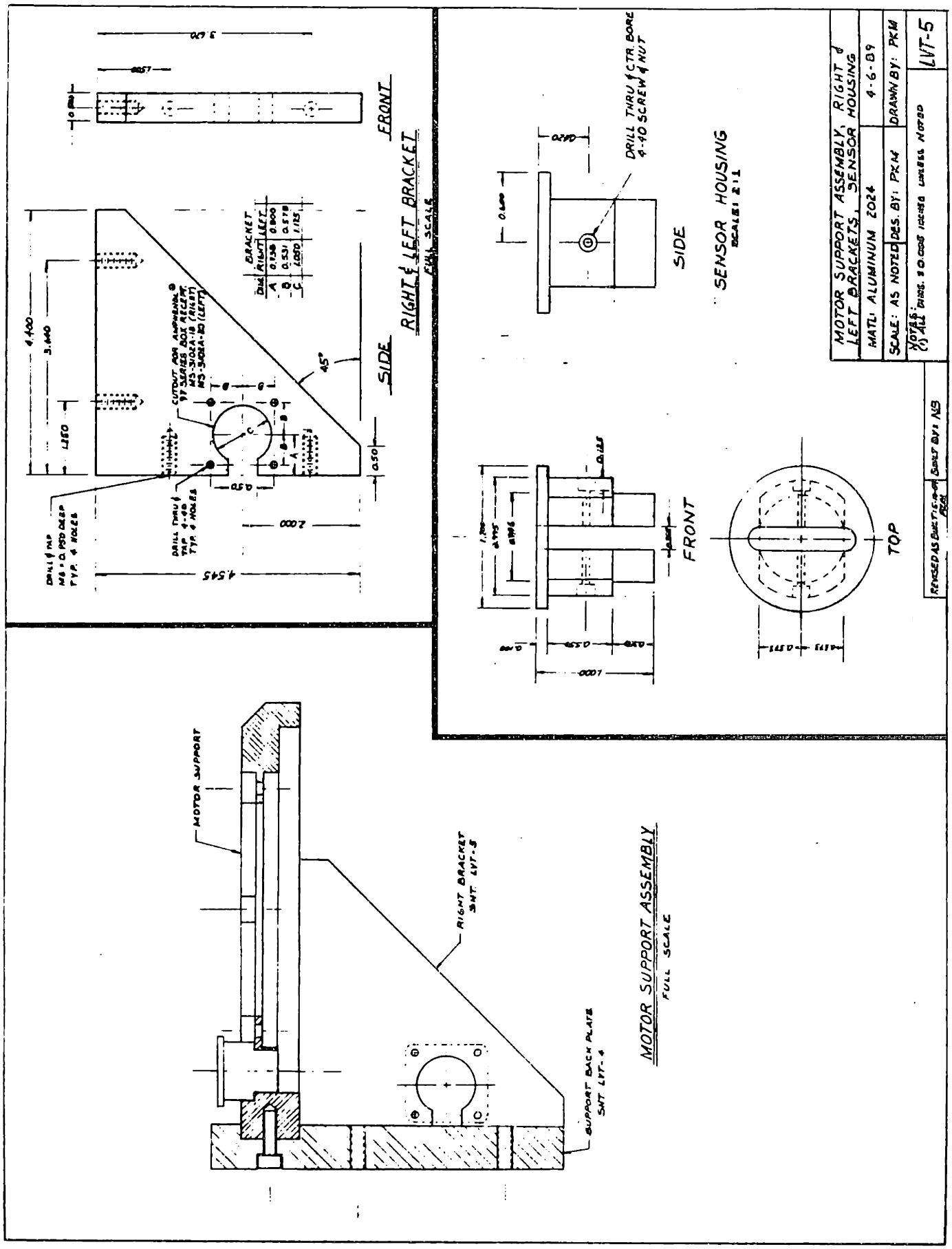


Figure E.5: Motor support assembly, right/left brackets, sensor housing (LVT-5)

Appendix F

Nomenclature

Globally defined nomenclature is listed and described below.

Symbol	*Definition*
c	camera center of projection (camera origin)
d	camera principal distance
h_i	the ith of four internal calibration interpolation functions
$\mathbf{I}$	identity matrix
l_i	the ith of 11 DLT parameters
$\mathbf{M}$	general transformation matrix from global to camera coordinates
$\mathbf{M}_{cm}$	transformation matrix from camera to mirror coordinates
$\mathbf{M}_{om}$	transformation matrix from global to mirror coordinates
$\mathbf{M}_{\Theta}$	transformation matrix of mirror through rotation angle Θ
m_{ij}	ith row and jth column component of $\mathbf{M}$
p	principal point of camera
$\mathbf{R}$	reflection matrix
$\mathbf{u}$	raw 2D image space vector having components $(u, v)^T$

$\bar{\mathbf{u}}$	corrected 2D image space vector having components $(\bar{u}, \bar{v})^T$
$\mathbf{u}'$	$\bar{\mathbf{u}}$ normalized with d
$\Delta\mathbf{u}$	2D internal image correction vector having components $(\Delta u, \Delta v)^T$
$\mathbf{u_p}$	2D image space vector to p having components $(u_p, v_p)^T$
$\mathbf{x}_o$	vector $(x_o, y_o, z_o)^T$ identifying 3D point in object space
$\mathbf{x}_c$	vector $(x_c, y_c, z_c)^T$ identifying 3D point in camera space
$\mathbf{x}_m$	vector $(x_m, y_m, z_m)^T$ identifying 3D point in mirror coordinates
$\mathbf{x}_{oc}$	camera (c) relative to global having components $(x_{oc}, y_{oc}, z_{oc})^T$
$\mathbf{x}_{co}$	global origin relative to camera having components $(x_{co}, y_{co}, z_{co})^T$
$\mathbf{x}_{cm}$	mirror relative to camera having components $(x_{cm}, y_{cm}, z_{cm})^T$
$\mathbf{x}_{om}$	mirror relative to global coordinates having components $(x_{om}, y_{om}, z_{om})^T$